Univerzita Palackého v Olomouci

Přírodovědecká fakulta

Dlouhodobý výzkum hnízdní biologie dutinových pěvců

Habilitační práce

Obor: Zoologie

Miloš Krist

Katedra zoologie a ornitologická laboratoř PřF UP v Olomouci

Obsah

Úvod

Hnízdní biologie patří k tradičně studovaným odvětvím ornitologie (Lack 1968). Dílem to může být proto, že období rozmnožování je kritické pro šíření vlastních genů a je tak svým způsobem z celého ptačího roku nejzajímavější. Dílem to může být i z praktických důvodů. V období rozmnožování jsou ptáci snadněji pozorovatelní, neboť jsou vázáni k jednomu místu – hnízdu, kde vychovávají mláďata. Vůbec nejsnadněji se pozorují druhy, které se dají k hnízdění nalákat nabídkou nějakého zdroje, třeba hnízdních budek. Právě dutinoví hnízdiči se proto stali modelovou skupinou, jejímž studiem se objasnila řada evolučně-ekologických otázek.

Přes tento dlouhodobý výzkumný zájem je v oboru hnízdní biologie stále co objevovat. Souvisí to i s rozvojem jiných vědních oborů a technologií, které umožňují podrobnější a přesnější studium hnízdění. Revoluci tak například přinesly metody molekulární genetiky. Ty prokázaly, že většina ptačích druhů není čistě monogamní, jak se domníval David Lack (1968), ale že si oba partneři navzájem zahýbají (Griffith et al. 2002), což vede k silnému pohlavnímu výběru (Sheldon & Ellegren 1999). Citlivé metody analytické chemie zase umožnily měřit koncentrace některých substancí, jako třeba karotenoidů nebo hormonů, které mohou i ve stopovém množství být důležité pro vývoj a konkurenceschopnost mláďat (Blount et al. 2000, Groothuis et al. 2005). Mnoho nových poznatků o hnízdění lze nyní získat díky rozvoji monitorovací techniky, jako jsou videokamery (Weidinger 2009), vysílačky (Ward et al. 2014), nebo RFID čtečky (Bonter & Bridge 2011).

Rodičovské péče a pohlavního výběru se týká i většina článků, jež jsem zahrnul do této habilitační práce. Každý článek je doplněn stručným komentářem, ve kterém uvádím zejména jinde nedostupné informace. Třeba motivaci proč článek vznikl a jak se v průběhu revizí měnil. Případně se zamýšlím nad jeho přínosem pro obor nebo zmiňuji některé zajímavé momenty, které práci na článku doprovázely. Tyto komentáře tedy nemají charakter další vědecké práce ani netvoří syntézu (review) publikovaných prací, čímž se tento spis možná liší od zaužívané formy habilitačních prací. Domnívám se však, že tato forma dobře vyhovuje § 72, odstavce 3 zákona 111/1998 o vysokých školách, který uvádí, že habilitační práce buď přináší nové vědecké poznatky nebo je souborem uveřejněných vědeckých prací, které jsou doplněny komentářem.

Většina přiložených článků popisuje výsledky terénního výzkumu na populaci lejska bělokrkého (*Ficedula albicollis*), kterou sleduji od roku 1997, kdy jsem začal pracovat na diplomové práci. Tehdy mě můj vedoucí, prof. Stanislav Bureš, seznámil s amatérským ornitologem Honzou Stříteským, jenž měl ve smrkovém lese na Velkém Kosíři u Prostějova vyvěšeno na dvou plochách asi 120 hnízdních budek. V průběhu let jsme budkové plochy rozšiřovali a přesouvali ze smrkových porostů do dubových, které lépe odpovídají nárokům lejska. V současnosti je na Kosíři asi 380 budek, které jsou rozmístěny na pěti plochách v dubových lesích. V nich hnízdí ročně kolem 100 párů lejska bělokrkého, 130 párů sýkory koňadry (*Parus major*), 40 párů sýkory modřinky (*Cyanistes caeruleus*) a 15 párů brhlíka lesního (*Sitta europaea*). Detailně monitorujeme jen populaci lejska bělokrkého, u ostatních druhů většinou se spolupracovníky a studenty stíháme jen zjišťovat základní data o jejich hnízdění. Sýkory koňadry se proto týkají jen dvě práce z tohoto souboru (články 5 a 7), který je dále doplněn jedním review (článek 10) a dvěma teoretickými pracemi (články 3 a 6), jež se zabývají rodičovskou péčí.

Obr. 1. Vlevo: kolega Honza Stříteský kontroluje obsah budky ve smrkovém lese (1998). Vpravo: v současnosti jsou všechny hnízdní budky na Velkém Kosíři vyvěšené ve svahových doubravách.

Komentáře k jednotlivým článkům

Článek 1: Krist, M. (2004) Importance of competition for food and nest-sites in aggressive behaviour of Collared Flycatcher *Ficedula albicollis*. *Bird Study*, 51, 41-47.

Tento článek je založen na datech, která jsem sbíral pro svou diplomovou práci. Téma diplomové práce jsem měl možnost vybrat si sám. To bylo na jednu stranu příjemné, protože jsem si mohl zvolit, co mě zajímalo. Na druhou stranu jsem měl v pátém semestru studia malý přehled o dění v oboru, takže pro mě bylo obtížné vymyslet si vhodné téma. Vedoucí mé práce, prof. Stanislav Bureš, mi tehdy dal přečíst nějaké práce o lejscích, na kterých se podíleli čeští kolegové. Nejvíce mě zaujal článek Krále & Bičíka (1992). Ten se zabýval agresivitou lejsků bělokrkých proti sýkoře koňadře, která je jeho konkurentem o dutiny. Navrhl jsem tedy, že provedu podobný pokus, což mi vedoucí posvětil a dále mi doporučil testovat kromě konkurence o dutiny také konkurenci o potravu. Finanční prostředky katedry byly v té době zřejmě dost omezené. Byl jsem proto rád, že mi ze sbírek gymnázia v Kroměříži, kde jsem dříve studoval, zdarma zapůjčili k provádění pokusů tři dermoplastické preparáty. Jednalo se o sýkoru koňadru, kterou jsem chtěl použít pro simulaci hnízdního a potravního konkurenta, pěnkavu obecnou (*Fringilla coelebs*), která měla simulovat konkurenta o potravu a pěvušku modrou (*Prunella modularis*), se kterou si lejsek o nic nekonkuruje.

V průběhu této práce jsem dostal od vedoucího jedno velmi užitečné poučení. Předběžná analýza dat mi ukázala, že lejsci reagovali třepotavými lety na pěvušku více než na pěnkavu. Když jsem tedy roztřeseně hlásil, že to nevychází podle předpokladů, odpověděl mi vedoucí lakonicky: „No a?" Tato poznámka mi pomohla prozřít a uvědomit si, že neexistují špatné výsledky, mohou být jen špatné metody. Toto poučení se snažím předávat i svým nynějším studentům, kteří mají často podobný problém, jako jsem měl tehdy já. Vyšší frekvenci třepotavých letů proti atrapě pěvušky jsem nakonec v manuskriptu vysvětlil tak, že toto chování není jen projevem agrese, ale slouží i jako signál partnerovi, že u hnízda není něco v pořádku. Manuskript byl nejdříve ve třech různých redakcích (Animal Behaviour, Journal of Avian Biology, Ibis) odmítnut. Jedna z připomínek recenzentů byla, že hypotéza o signalizační funkci třepotavých letů je divná a že jí nevěří. Raději jsem tedy toto chování, které nejde snadno interpretovat, v další verzi rukopisu již neanalyzoval. Nyní myslím, že je to trochu škoda, protože jsem o správnosti své „signalizační" hypotézy stále přesvědčen. Nakonec tato hypotéza byla

podložena stovkami hodin pozorování v terénu a potvrzuje se mi i dnes, když se studenty provádíme podobné pokusy v rámci předmětu „Cvičení z etologie".

Tento článek je také můj jediný, který je zcela založen na přímém pozorování zvířecího chování. V následujících pracech jsem více využíval přístrojové techniky, např. videokamer, dataloggerů, spektrofotometru, RFID čteček aj. Díky těmto přístrojům je sběr dat efektivnější. Na druhou stranu ale přímé pozorování přináší i neplánované, přesto ale zajímavé (tzv. anekdotické) informace. Ty sice většinou nejdou rigorózně analyzovat, ale i tak umožňují vytvořit si lepší obrázek o chování studovaného druhu. Pozoroval jsem tak například společný třepotavý let pěti lejsků (tedy asi dvou rodičů a tří cizích ptáků) právě proti atrapě pěvušky, převzetí budky s hnízdem uhelníčka lejskem, kdy samec odháněl uhelníčky a samice zatím nosila materiál na hnízdo, nebo odnesení utržené hlavy z atakovaného preparátu koňadry do vzdálenosti asi 15 m od hnízda. Cennou hlavu jsem po delším hledání našel, nasadil zpět na drátek a atrapu dále používal.

Další připomínka recenzentů byla, že každým pokusem provádím jen pseudoreplikaci, protože mám jen jednu či dvě atrapy od jednoho druhu konkurenta. S touto výtkou souhlasím. Mnohem lepší, než dělat stovky pokusů s pěti atrapami tří druhů, by bylo provést méně pokusů s více druhy, které by reprezentovaly každou zamýšlenou kategorii. Tedy hnízdního konkurenta, potravního konkurenta a nekonkurující druh. To jsem si ale při plánování pokusu neuvědomil a stejně bych těžko získal preparáty více druhů. Navíc v té době bylo použití dvou druhů (jeden jako treatment a jeden jako kontrola) naprosto běžné. Nyní je metodologie již pokročilejší a podobné pokusy bývají obvykle prováděny s atrapami více druhů (např. Santema et al. 2020).

Obr. 2. Samci lejsků bělokrkých (vpravo) bránili hnízdo před konkurenty intenzivněji než samice (vlevo).

Článek 2: Krist, M., Remeš, V., Uvírová, L., Nádvorník, P. & Bureš, S. (2004) Egg size and offspring performance in the collared flycatcher (*Ficedula albicollis*): a within-clutch approach. *Oecologia*, 140, 52-60.

Článek 3: Krist, M. & Remeš, V. (2004) Maternal effects and offspring performance: in search of the best method. *Oikos*, 106, 422-426.

Tyto dva články začaly sérii prací, ve kterých jsme se zabývali otázkou, do jaké míry mohou samice ovlivňovat fenotyp svých potomků tím, že změní velikost nakladených vajec. Články byly zároveň prvními výsledky mé doktorské práce. V té jsem původně chtěl studovat rozdíly ve velikosti vajec, z kterých se líhnou samci a samice. Zdrojem inspirace pro tuto myšlenku byla zřejmě kniha Tima Clutton-Brocka (1991), kterou jsem tehdy četl. Když jsem tuto možnost diskutoval se svým vedoucím,

prof. Burešem, tak ho napadlo, že by daný problém šel studovat pomocí líhně a značení nově vylíhlých mláďat.

Tento technický nápad jsme začali rychle realizovat. Bohužel však ne úplně šťastným způsobem, kdy jsme chtěli zajímavé výsledky získat s co nejmenšími možnými náklady. Našli jsme tedy inzerát soukromého chovatele papoušků, který nabízel k prodeji líheň vlastní výroby. Tu jsme zakoupili, nachystali čtyři autobaterie, od Lesů ČR vyprosili zapůjčení staré maringotky a na začátku května 2000 do líhně v ní umístěné dodali první snůšky lejsků. Za dva dny se ukázalo, že líheň má větší spotřebu energie, než můžeme zajistit pomocí čtyř autobaterií. Pokus jsme tedy pro ten rok ukončili, čímž jsem ale ztratil jednu sezónu svého doktorátu.

Na přípravu další sezóny jsme měli více času, takže jsme s autobateriemi počítali jen na noc a přes den provozovali líheň z malé elektrocentrály. Pokus tak již v roce 2001 proběhl, ale brzy se ukázaly další mouchy našeho systému. Termostat byl nedokonalý, teplota se musela neustále hlídat, vejce manuálně otáčet a hlavně každé tři hodiny ve dne v noci kropit, protože je líheň silně vysoušela. Líhnivost tak byla i přes všechnu snahu výrazně menší, než je tomu v přírodě. Další sezónu jsme koupili jiný termostat, který nejdříve fungoval lépe, ke konci sezóny se ale zasekl úplně a v líhni se nekontrolovatelně zvyšovala teplota. Naštěstí se to stalo ve dne, my si toho brzy všimli a roznesli snůšky zpět do hnízd. Tím projekt kontrolovaného líhnutí definitivně skončil. Jeden kolega z pražské přírodovědy mi tyto zážitky později okomentoval větou: „My jsme se naučili nešetřit." Kdybychom toto pravidlo v té době střídmě užívali i my, mohla být naše práce méně stresová, rychlejší a přinést lepší výsledky.

Molekulární část práce nutná pro testování vztahu mezi velikostí vajec a pohlavím mláďat byla časově náročná. DNA se izolovala ještě zdlouhavým a poněkud nebezpečným způsobem – fenolchloroformovou reakcí. Také následná PCR a elektroforéza, která rozlišovala pohlaví na základě různé délky amplifikovaných sekvencí genu CHD, probíhala dost pomalu. Jako zaujatého frekventanta kurzu „Cytologie a cytotaxonomie", který vedl prof. Jaroslav Doležel z Ústavu experimentální botaniky, mě napadlo, zda by nešlo pohlaví určovat rychleji a levněji s využitím průtokové cytometrie. Tato metoda by mohla rozlišit samce od samic na základě rozdílné velikosti pohlavních chromozomů W a Z. Profesora Doležela nápad zaujal a tak ho s námi vyzkoušel. Jistý rozdíl v obsahu DNA mezi samci a samicemi byl ve výsledcích patrný, nebyl však natolik výrazný, aby umožnil jednoznačné určení pohlaví u všech jedinců. Museli jsme se tedy spolehnout na výše uvedenou tradiční molekulární metodu.

Ještě než byly tyto molekulární analýzy provedeny, napadlo nás, že bychom mohli novým způsobem testovat, jak velikost vejce ovlivňuje fitness mláďat. Tento vztah byl většinou studován metodou tzv. cross-fosteringu, kdy se prohodily snůšky mezi hnízdy a sledovalo se, zda průměrná velikost vajec ovlivňuje fitness mláďat. Prohození vajec odstranilo možnou matoucí korelaci mezi kvalitou rodičů či teritorií a fitness mláďat. Tato metoda byla v té době považována za zlatý standard studia podobných mateřských efektů. Uvědomili jsme si, že rozdílná kvalita rodičů a teritorií by nepředstavovala problém ani tehdy, kdyby se vztah mezi velikostí vajec a fitness mláďat testoval na úrovni jednotlivých vajec/mláďat v hnízdech. V něčem by dokonce taková kontrola byla lepší, protože všechna mláďata by měla podobný genetický základ, na rozdíl od mezisnůškových cross-fosteringových pokusů.

Provedli jsme tedy vnitrosnůškový test a zjistili, že velikost vajec nemá vliv na kvalitu mláďat. To byl rozdíl proti většině cross-fosteringových studií. Tento rozdíl jsme interpretovali tak, že předchozí studie nadhodnocovaly vliv velikosti vajec na fitness mláďat. Sepsaný rukopis jsme zaslali do časopisu Journal of Animal Ecology. Jeden z recenzentů ale zřejmě nepochopil náš argument o výhodě

4

vnitrosnůškového přístupu oproti tradičním cross-fosteringům a editor následně náš rukopis odmítl. Pokusili jsme se ještě toto nedorozumění editorovi vysvětlit a případně tak jeho rozhodnutí zvrátit. To se nám ovšem nepodařilo. Kromě námitek, které měli recenzenti, se editorovi také nelíbilo, že máme jen krátkodobé měřítko fitness mláďat – jejich přežívání a růst do vyvedení.

Zkusili jsme zaslat rukopis ještě jednou v dané podobě do časopisu Ecology. Editorem zde byl Tony Williams, který se sám podobnými otázkami zabýval. Ten náš rukopis odmítl s argumentem, že mezisnůšková variabilita ve velikosti vajec je větší než ta vnitrosnůšková a tedy může mít na fitness mláďat větší vliv. Pochopili jsme tedy, že musíme naše úvahy o možných matoucích genetických korelacích mezi velikostí vajec a fitness mláďat buď zcela opustit, nebo mnohem lépe vysvětlit. Nakonec jsme se rozhodli pro rozdělení článku na dva – jeden metodický a druhý datový. V tom datovém jsme se již nezabývali trade-offem mezi velikostí vajec a jejich počtem, ale jen dopadem vnitrosnůškové variability na fitness jednotlivých mláďat. Vzhledem ke zdržení způsobenému předchozími odmítnutími jsme navíc rukopis také mohli doplnit o dlouhodobější měřítko mláděcí fitness – jejich recruitmentem do hnízdící populace. Oba tyto manuskripty pak byly bez zásadních připomínek akceptovány v Oecologii a Oikosu.

Úvahy prezentované v našem metodickém článku jsme v době jeho přípravy považovali za velmi novátorské. Kromě jedné krátké poznámky v práci Magratha (1992) nikdo z ornitologů studujících význam velikosti vajec pro fitness mláďat neuvažoval, že by tento vztah mohl být tažen korelacemi s jinými mateřskými znaky. Několik let po publikaci našeho článku jsme ale našli zmínku o tomto možném matoucím faktoru i v diskusi starší práce provedené na broucích (Fox 1997). Stejný autor ovšem tento možný problém již nezmínil ve své známější přehledové práci (Fox & Czesak 2000). Zdá se tedy, že daný problém nebyl nikdy u hmyzu uvažován jako příliš zásadní. Možná i proto, že u bezobratlých s krátkým reprodukčním cyklem se používaly i sofistikovanější metody studia mateřských efektů, než byly jednoduché cross-fosteringové experimenty rozšířené v té době v ornitologické komunitě.

Obr. 3. Inkubace vajec v domácky vyrobené líhni byla náročná kvůli nedokonalé kontrole teploty a vlhkosti. Zde je líheň pro ilustraci naplněna atrapami vajec, skutečné líhnutí jsme bohužel nikdy nevyfotili.

Článek 4: Krist, M., Nádvorník, P., Uvírová, L. & Bureš, S. (2005) Paternity covaries with laying and hatching order in the collared flycatcher *Ficedula albicollis*. *Behavioral Ecology and Sociobiology*, 59, 6-11.

Kolem roku 2000 bylo publikováno několik vlivných studií, které se zabývaly diferenciální alokací. To je proces, při kterém jeden z rodičů investuje více do potomků, pokud má atraktivního partnera (Sheldon 2000). U ptáků tak bylo například pozorováno, že samice hnízdící s atraktivními samci kladou větší vejce (Cunningham & Russel 2000), nebo do nich ukládají více testosteronu (Gil et al. 1999). Uvědomili jsme si, že podobný mechanismus, který tyto studie popisovaly na úrovni mezi snůškami, by mohl fungovat i v rámci jedné snůšky. Samice většiny druhů ptáků totiž svým partnerům zahýbají. Ve snůšce jsou pak mláďata zplozená dvěma samci – sociálním partnerem a extra-párovým samcem, který bývá často atraktivnější. Použili jsme tedy naše data získaná líhnutím lejsků pro test této nové hypotézy. Vejce, ze kterých se líhla extrapárová mláďata, však byla stejně velká jako ta oplozená sociálním otcem. Sepsali jsme manuskript, jehož myšlenku size recenzenti chválili, zároveň ale zkritizovali malou velikost vzorku, na kterém byla naše negativní zjištění postavena. Článek byl tedy v této podobě odmítnut nejdříve v časopise Ecology Letters a potom nadvakrát v Journal of Avian Biology, kde zvlášť posuzovali jeho původní a revidovanou verzi.

Myšlenkami jsme v tomto manuskriptu byli příliš upnuti na samičí investice a jejich návratnost. Proto jsme nejdříve sledovali jen energeticky náročné investice do velikosti vajec. Po zmíněných dvou odmítnutích nás však napadlo rozšířit rukopis o vztah mezi paternitou a pořadím vejce v sekvenci kladení. Následná analýza pak prokázala, že se extra-párová mláďata často objevují v prvních nakladených vejcích, kdežto v těch posledních se již téměř nevyskytují. Uvědomili jsme si, že tento mechanismus může mít pro pochopení pohlavního výběru u ptáků dost zásadní význam. V té době se totiž za nejlepší vysvětlení samičího zahýbání považoval zisk dobrých genů pro potomky. A za nejsilnější důkaz tohoto genetického zisku se považovala větší fitness extra-párových mláďat ve srovnání s mláďaty párovými (Sheldon et al. 1997, Johnsen et al. 2000, Sheldon 2000). Pokud jsou však extra-párová mláďata kladena v časných vejcích, mohou mít vyšší fitness z důvodu dřívějšího líhnutí, ne z důvodu lepších genů. Domnělý nejsilnější důkaz zdůvodňující samičí zahýbání u ptáků tak padl.

Původní manuskript jsme ambiciozním způsobem přepsali, ale stejně se nám ho nedařilo nikde udat. Postupně byl odmítnut v dalších třech redakcích (Proc R Soc Lond B, Nature, Behavioral Ecology). V této fázi jsme ho dali přečíst i Benu Sheldonovi, se kterým jsme si již dříve korespondovali. Ten nám přes své velké zaneprázdnění nezištně pomohl vypilovat argumenty, které byly do značné míry protikladné k jeho předchozím úvahám. Tuto jeho pomoc a neobvyklou otevřenost k cizím myšlenkám stále velmi obdivujeme a ceníme si jí. Manuskript pak byl přijat k publikaci v časopise Behavioral Ecology and Sociobiology.

Cítili jsme ale, že náš vzorek je příliš malý na to, aby na něm mohla být založena významná změna uvažování v oboru. Sami se dokonce díváme s jistou nedůvěrou na závěry studií s malým vzorkem a pozitivními výsledky, jejichž publikace může být ovlivněna i tzv. publikačním bias ve prospěch signifikantních výsledků. Chtěli jsme si naše závěry ověřit na větším vzorku hnízd, ale zároveň jsme již nechtěli pokračovat ve stresujícím líhnutí mláďat v inkubátoru. Vymysleli jsme proto nový způsob, jak určit, ze kterých vajec se líhnou extra-párová a párová mláďata (viz článek 11). Na konferenci Mezinárodní společnosti pro behaviorální ekologii (ISBE), která proběhla v roce 2008 na Cornellově univerzitě, jsem referoval o tomto probíhajícím experimentu. Setkal jsem se tam s Oscarem Vedderem, který představil jejich studii inspirovanou naším původním článkem. Na větším vzorku hnízd a lepšími metodami potvrdili u sýkor modřinek to, co jsme dříve zjistili u lejsků. Jejich článek byl pak publikován

v předním časopise a stal se známějším (Magrath et al. 2009). I tak ale považuji náš objev a původní příspěvek za významný pro proběhlou změnu v uvažování, proč samice svým partnerům zahýbají (viz i Kempenaers 2009, Forstmeier et al. 2014).

Obr. 4. Extrapárová mláďata se líhla první a měla tak před svými polovičními sourozenci konkurenční výhodu.

Článek 5: Remeš, V. & Krist, M. (2005) Nest design and the abundance of parasitic *Protocalliphora* blow flies in two hole-nesting passerines. *Ecoscience*, 12, 549-553.

V průběhu práce na našich hlavních tématech doktorských prací jsme se s kolegou Laďou Remešem domluvili na tom, že provedeme na lokalitě Velký Kosíř společný experiment, ve kterém prohodíme hnízda mezi lejsky bělokrkými a sýkorami koňadrami. Hnízda těchto dvou druhů se hodně liší – lejsci staví hnízda ze suché trávy a sýkory z mechu a zvířecí srsti. Předpokládali jsme, že tyto rozdíly budou ovlivňovat izolační vlastnosti hnízd i množství ektoparazitů, kteří se v hnízdech schovávají. Prohození hnízd se u budkových populací celkem nabízí, protože všechna hnízda mají stejný tvar a velikost, takže jde o snadnou manipulaci, která je ovšem nezbytná pro separaci efektu ptačího druhu od typu hnízda. Pokud víme, tak jsme byli první, kdo podobný experiment provedl. Již si přesně nevzpomínáme, je ovšem možné, že jsme se sami inspirovali u našeho kolegy Karla Weidingera, který v té době plánoval podobný experiment s hnízdy otevřeně hnízdících druhů pro odlišení vlivu velikosti a umístění hnízda na riziko predace (Weidinger 2004). Naše studie se naopak asi stala inspirací pro další podobný pokus, který o pár let později provedla španělská skupina (Moreno et al. 2009).

Terénní část této práce mi díky drobnému detailu docela utkvěla v paměti. Výzkum jsme prováděli v letech 2002–2003. První z těchto sezón jsem ještě paralelně pokračoval v líhnutí lejsků v líhni. Potřeboval jsem tedy nechat nějaká hnízda lejsků nemanipulovaná. Dostupných budek pro prohazování hnízd tak moc nebylo. Použili jsme proto i hnízda z plochy, kde jsme měli budky vyvěšené ve čtyřech různých výškách (Krist a Stříteský 2008). Lejskům se nejvíce líbilo v těch nejvyšších, které byly umístěny přes 4 m nad zemí. Kolega Honza Stříteský používal pro jejich kontrolu hliníkový žebřík, který si ale vždy vozil z domu autem. Protože finanční prostředky jsme měli velmi omezené, rozhodli jsme se s Laďou žebřík si vyrobit. Našli jsme dva rovné doubky, porazili je a žebřík z nich stloukli. Až u hotového výrobku jsme si uvědomili potíž s jeho váhou. I ve dvou jsme měli s jeho

nošením po ploše co dělat. Za jeden ze svých největších terénních výkonů proto považuji fakt, že jsem to jednou nebo dvakrát, když kolega do terénu nemohl, zvládl dokonce sám.

Po skončení terénní části výzkumu jsme nejdříve zpracovali data o množství ektoparazitů v hnízdech. Zaměřili jsme se na dva nejčastější parazity, tedy mouchy rodu *Protocalliphora* a blechy rodu *Ceratophyllus*. Typ hnízda neměl vliv ani na jednoho z těchto parazitů. Naopak jejich abundance závisela na prostředí. Blech bylo více ve smrčinách a much zase v doubravách. Mouchy také více parazitovaly koňadry bez ohledu na typ hnízda. Sepsaný manuskript byl dvakrát po sobě odmítnut, nejprve v Canadian Journal of Zoology a poté v Journal of Avian Biology. Recenzentům se nelíbilo hlavně to, že jsme budky před provedením experimentu nezbavili stávajích parazitů. To by mohl být problém v případě blech, které se v budkách vyskytují dlouhodobě, včetně zimního období. Nakonec jsme se proto rozhodli část o blechách z manuskriptu vyškrtnout, abychom dále jeho přijetí nekomplikovali. Seškrtanou práci jsme poslali do časopise Ecoscience, kde byla dříve publikována hypotéza o významu hnízdního materiálu pro výskyt ektoparazitů (Bauchau 1998), kterou jsme nyní testovali. Možná i díky této návaznosti zdejší editor náš manuskript bez větších potíží přijal.

Po přijetí této práce jsme začali zpracovávat i data z teplotních dataloggerů. Chtěli jsme testovat vliv typu hnízda na inkubační rytmus obou studovaných druhů. Tuto část práce jsme nakonec nikdy nedokončili, protože se stále objevovaly nové priority. Také jsme cítili, že dat nemáme zase tolik, aby nám mohly dát velmi přesvědčivý obraz o inkubačním rytmu, který je stejně jako ostatní behaviorální znaky mezi jedinci hodně variabilní. Bylo by vhodné dataset ještě rozšířit. To by sice nyní s dostupnými modernějšími dataloggery (i-buttony) bylo snazší, ale i tak se nám potřebný čas zatím nepodařilo vyšetřit.

Obr. 5. Hnízda sýkory koňadry (vlevo) a lejska bělokrkého (vpravo) jsou postavena z rozdílných materiálů. Larvy parazitických much však byly v obou typech hnízd stejně početné.

Článek 6: Krist, M. (2006) Should mothers in poor condition invest more in daughter than in son? *Ethology Ecology & Evolution*, **18, 241-246.**

Jak už jsem uvedl, původním záměrem mé doktorské práce bylo zjistit, zda existuje dimorfismus ve velikosti vajec, z kterých se líhnou samci a samice. V průběhu studia jsem se snažil udělat si přehled v teoriích, které se zabývaly rodičovskými investicemi a poměrem pohlaví. Zjistil jsem, že je to poměrně komplikovaná oblast, kde existuje mnoho teoretických modelů, které ale často nejsou ještě dobře empiricky otestované. Případně nemají empirici jasno v tom, co vlastně teoretické modely predikují. To byl dokonce i případ jednoho z nejznámějších modelů, který se zabývá investicemi do synů a dcer na základě mateřské kondice (Trivers & Willard 1973). Téměř 30 let po jeho publikaci se empirici v předním etologickém časopise stále dohadovali, co vlastně model předpokládá a predikuje (Cameron

& Linklater 2002, Carranza 2002). Abych si ujasnil, kdo z daných autorů má vlastně pravdu, nasimuloval jsem si s pomocí bratrů Remešových nějaká data. Dalším promýšlením této otázky jsem dospěl až k jednoduchému grafickému modelu. Nakonec jsem získal pocit, že by bylo vhodné snažit se model opublikovat a pomoci tak v tříbení myšlenek i dalším empirikům.

Sepsaný článek jsem nabídl do redakce Animal Behaviour, kde byl na základě dvou recenzí odmítnut. Prvním recenzentem byl zřejmě Juan Carranza, jehož předchozí závěry můj článek popíral a který si stále trval na svém. Druhý recenzent byl Stuart West, tedy špička oboru, který namítal, že již na dané téma existuje formální model (Frank 1987) a ten můj tedy nepřináší nic nového. Tento názor jsem ale nesdílel, protože jsem sám několikrát model Franka (1987) četl a stejně jsem si z něho nebyl schopný odnést závěr, který by rozřešil výše uvedený spor. Dokonce jsem věděl, že Frankův formální model má problém chápat a testovat i jeden špičkový empirik spolupracující právě se Stuartem Westem. Rozhodl jsem se proto vytrvat a článek jsem poslal do časopisu Ethology. Tamější editor, Klaus Reinhold, článek odmítl s tím, že by o jeho přijetí mohl znovu uvažovat, kdybych ho rozšířil o Fisherovu teorii rovné alokace. To mi však jednak nepřipadalo pro daný účel relevantní a jednak bych to ani nedokázal.

Proto jsem další pokus namířil do méně známého časopisu Ethology Ecology & Evolution, kde byl článek nakonec přijat. Bohužel však v redakci neprošel jazykovou úpravou, jak tomu bylo u našich předchozích článků a v angličtině jsou tak chyby snižující srozumitelnost práce. V té době ještě nebylo úplně zvykem redakce vyvázat se z odpovědnosti za jazykovou stránku. Dnes je tomu již jinak, opravení angličtiny se bere za zodpovědnost autorů, kteří za to také musí zaplatit. Při sazbě navíc v článku vznikla ještě další chyba – vypadla část popisku osy y u jednoho z obrázků. Částečně jsem to zavinil sám, protože jsem do redakce neposlal obrázek ve formátu pdf nebo eps, ale jen v ppt. Částečně šlo také o chybu redakce, která mi vysázený článek neposlala na korekturu. Kromě poučení z těchto chyb jsem ještě pochopil, že bude lepší neplést se dále teoretikům do práce. Vynaložené úsilí je velké, ale výsledné dílo stejně není dostatečně kvalitní a oceněné. Teoretici by však podle mého názoru zase měli své připravované manuskripty dávat na komentáře i empirikům, kteří by posoudili jejich srozumitelnost. Modely jsou nakonec určeny hlavně těm, kteří je mají v přírodě otestovat.

Článek 7: Remeš, V., Krist, M., Bertacche, V. & Stradi, R. (2007) Maternal carotenoid supplementation does not affect breeding performance in the Great Tit (*Parus major*). *Functional Ecology*, 21, 776-783.

V nultých letech vrcholil zájem o možné adaptivní mateřské efekty ve formě ukládání mikronutrientů do kladených vajec (Blount et al. 2000, Groothuis et al. 2005). Rozhodli jsme se proto s kolegou Laďou Remešem tuto myšlenku otestovat na sýkorách. Části sýkor jsme dávali do hnízd před a při kladení krmení s luteinem a na kontrolní část jen podobné krmení bez luteinu. Předpokládali jsme, že na experimentálních hnízdech budou sýkory dávat poskytnutý lutein do žloutků. Hlavně nás ale zajímalo, zda se bude lišit kvalita mláďat na experimentálních a kontrolních hnízdech. Po skončení terénní části experimentu jsme nejdříve potřebovali ověřit, zda naše suplementace skutečně ve vejcích zvýšila obsah luteinu. Nejprve jsme nabídli spolupráci na tomto výzkumu odborníkům z místní katedry analytické chemie, kteří měli zkušenosti a příslušné přístrojové vybavení potřebné pro dané analýzy. Ti však neměli o spolupráci příliš velký zájem. Oslovili jsme proto italského vědce Riccarda Stradiho, kterému se myšlenka líbila a tak nám koncentrace karotenoidů ve vejcích spolu se svým technikem Vittoriem Bertacchem zanalyzoval. Hodně jsme si oddechli, když se ukázalo, že náš experiment byl

efektivní a obsah luteinu byl opravdu ve vejcích z experimentálních hnízd výrazně vyšší než ve vejcích z hnízd kontrolních.

Experiment jsme prováděli na hnízdech všech druhů sýkor. Suplementovali jsme tak kromě nejčastějších koňader i modřinky a uhelníčky. U těchto dvou druhů jsme ale měli poměrně malý vzorek, proto jsme je nakonec do analýz nezahrnuli. Nyní mi to přijde škoda, protože i malý vzorek je lepší než nic. Tehdy jsme ale zřejmě byli ovlivněni naší předchozí zkušeností z recenzí článku 5, který byl odmítán kvůli metodické nedokonalosti jedné své části. Proto jsme v tomto následujícím článku raději druhy s malým vzorkem vůbec nepoužili.

Kromě efektů suplementace na růst a imunitu mláďat jsme původně zamýšleli testovat i její možný vliv na zbarvení peří. Proto jsme si vytiskli barevný standard, u kterého jsme vždy tři mláďata z hnízda vyfotili. Bohužel se zde však projevila naše nezkušenost se získáváním podobných dat. Jednak jsme si standard vytiskli na inkoustové tiskárně, což vedlo k jeho poškození deštěm. Ještě závažnější chyba ale byla, že jsme se snažili o standardní intenzitu osvětlení a tak jsme při fotografování používali blesk. Až v průběhu sezóny jsme si všimli, že na některých fotografiích blesk mláďata i standard přesvětlil a tak jejich žlutou barvu prakticky vymazal. Fotky tedy byly pro analýzy barvy mláďat nepoužitelné a my museli tuto proměnnou v manuskriptu oželet.

Přestože naše výsledky byly negativní – nenašli jsme skoro žádný efekt naší suplementace karotenoidy na mláďata ani dospělce – článek byl recenzenty pozitivně hodnocen a tak byl přijat v prvním časopise, kam jsme ho nabídli. Je ale fakt, že jeden z recenzentů požadoval, abychom naše nesignifikantní výsledky doplnili power analýzou, která by ukázala riziko chyby druhého typu. My jsme místo toho doplnili naše odhady efektů o konfidenční intervaly, což jsme považovali za vhodnější metodu, než jakou je power analýza (Hoenig & Heisey 2001). Recenzenta to úplně neuspokojilo, stejně jako naše zmínka, že naše studie má podobnou velikost vzorku a tudíž i sílu testu jako předchozí studie, které měly pozitivní výsledky. Podle jeho názoru byl tento fakt irelevantní. To je ale trochu sporné, protože to ukazuje na riziko vzniku publikačního bias. Všechny studie by se měly posuzovat stejně, bez ohledu na směr výsledků. Zmínku o stejné velikosti vzorků jsme ale raději obětovali a vyřadili, což již editorovi spolu s konfidenčními intervaly stačilo a power analýzu jsme tak nemuseli doplňovat.

Obr. 6. Mláďata sýkory koňadry prospívala bez ohledu na množství luteinu, které jim matky vložily do žloutků.

Článek 8: Krist, M. & Grim, T. (2007) Are blue eggs a sexually selected signal of female collared flycatchers? A cross-fostering experiment. *Behavioral Ecology and Sociobiology*, **61, 863-876.**

Ptačí vejce jsou často zbarvená modře nebo zeleně. Toto zbarvení zřejmě neslouží k maskování snůšky před predátory (Weidinger 2001) a tak bylo ještě v roce 2002 považováno za těžko vysvětlitelné (Underwood & Sealy 2002). O rok později již ale vše bylo jinak. Moreno & Osorno (2003) publikovali elegantní hypotézu, po jejímž přečtení jsem si říkal, jak je možné, že to nenapadlo mě. Modrozelené zbarvení vajec totiž vysvětlili pomocí diferenciální alokace (Sheldon 2000). Ta je v tomto případě obrácena trochu naruby. Pohlaví, které má diferenciálně alokovat na základě kvality partnera, jsou totiž samci. Ti by měli více krmit mláďata samic, která kladou vejce se sytější barvou. Barvou vajec ukazují tyto samice svou kvalitu a samci by se tak měli o jejich mláďata lépe starat (Moreno & Osorno 2003).

Hned rok po zveřejnění této hypotézy navíc daní autoři přišli s prvními empirickými výsledky, které ji potvrzovaly. Zjistili, že samci lejska černohlavého více krmí mláďata, která se vylíhla z barevnějších snůšek (Moreno et al. 2004). Jejich studie však měla slabinu. Šlo v ní totiž jen o korelativní vztah, který mohl být dán i různou kvalitou teritorií – na dobrých teritoriích mohlo být dost potravy pro produkci sytě zbarvených vajec i pro vysokou frekvenci krmení mláďat. Rozhodli jsme se proto hned v následující sezóně jejich hypotézu otestovat experimentálně. Vejce jsme již při kladení měnili mezi hnízdy, abychom randomizovali vztah mezi jejich barvou a kvalitou rodičů či teritorií. Jako měřítko samčích investic jsme použili nejen jejich frekvenci krmení, ale i intenzitu, s jakou bránili hnízda před predátorem. Také jsme testovali některé předpoklady Morenovy a Osornovy hypotézy. Tedy zda existuje vztah mezi barvou vajec a kvalitou samic či mláďat a zda se barva vajec mění v průběhu kladení.

Manuskript jsme sepsali ještě v daném roce a zaslali ho do redakce Behavioral Ecology. Přes poměrně dobré recenze ho editorka odmítla bez možnosti resubmitace. Článek jsme tedy částečně upravili podle připomínek recenzentů a nabídli ho časopisu Behavioral Ecology and Sociobiology, kde již byl přijat. V diskusi manuskriptu jsme se hodně věnovali důvodům, které mohly vést k rozdílům mezi našimi výsledky a výsledky Juana Morena a jeho skupiny. Jedním z důvodů, které jsme uvažovali, byl i rozdíl v popisu barvy vajec. Ten jsme oproti předchozí studii vylepšili, protože jsme do něj zahrnuli i UV část spektra. Recenzenti byli ale k významu UV části spektra pro ptačí chování dost skeptičtí a poukázali na to, že tento metodický rozdíl můžeme s našimi daty přímo otestovat. Udělali jsme to a museli jim dát za pravdu. Korelace mezi barvou bez přihlédnutí a s přihlédnutím k UV spektru byla vysoká a rozdíly ve výsledcích minimální. UV část spektra tedy nebyla pro ptáky tak důležitá, jak to z dobových prací vypadalo.

Morenova a Osornova hypotéza zaujala více výzkumníků a její testování pomocí cross-fosteringových experimentů se proto stalo dost populární. Provedené experimenty ale měly jedno skryté úskalí (Riehl 2011). Prohození vajec mezi hnízdy sice zruší korelaci mezi kvalitou rodičů/teritorií a barvou vajec, nikoli však mezi barvou vajec a kvalitou mláďat, která se z nich vylíhnou. Je tedy možné, že samci krmí více mláďata z barevnějších vajec, protože taková mláďata hlasitěji žadoní. To jsme si při plánování našeho experimentu neuvědomili, přestože jsme se podobnými úvahami v případě výzkumu velikosti vajec sami zabývali (viz článek 3). Problém lze vyřešit druhým cross-fosteringem vylíhlých mláďat (Stoddard et al. 2012) nebo jejich navrácením do původních hnízd (Hodges et al. 2020). Náš článek má tedy v tomto ohledu oproti novějším pracím slabinu. Na druhou stranu některé jeho silné stránky nebyly dosud zopakovány, možná i kvůli jejich časové náročnosti. Například jsme zatím jako jediní měřili i obranu hnízda nebo prohazovali vejce po jednom, hned v den jejich nakladení. Tento náročnější postup může být přitom dost důležitý. Po začátku inkubace totiž již vejce nemusí být pro samce tak dobře pozorovatelná, neboť jsou většinu času schována pod sedící samicí.

Obr. 7. Barvu lejsčích vajec jsme měřili spektrofotometrem, který zaznamenával i UV část spektra. Tu ptáci sice na rozdíl od lidí vidí, ale pro jejich chování neměla žádný zvláštní význam.

Článek 9: Krist, M. (2009) Short- and long-term effects of egg size and feeding frequency on offspring quality in the collared flycatcher (*Ficedula albicollis*). *Journal of Animal Ecology*, 78, 907-918.

V tomto článku jsem se zabýval otázkami, které jsme s kolegou Laďou Remešem nastolili v naší předchozí, teoreticky zaměřené práci (článek 3). Tedy zda velká vejce poskytují mláďatům vyšší fitness, nebo zda je tento vztah ovlivněn jinými proměnnými, třeba postnatální péčí rodičů nebo poděděnými vlohami. Provedl jsem proto dva typy cross-fosteringových experimentů, jeden na úrovni celých hnízd a druhý na úrovni jednotlivých vajec (viz i článek 11). Kromě toho jsem analyzoval i data z nemanipulovaných snůšek. Sledoval jsem prospívání mláďat jak v hnízdech, tak i po jejich vyvedení. V modelech testujících vliv velikosti vajec na prospívání mláďat jsem statisticky kontroloval pro vliv morfologie a krmení rodičů. Zjištěné efekty jsem porovnával s naší předchozí vnitrosnůškovou studií (článek 2).

Dlouhý manuskript jsem nabídl k publikaci do časopisu Journal of Animal Ecology. V tomto časopise jsme neuspěli s předchozím pokusem (článek 2) i proto, že jsme neměli dlouhodobější měřítko mláděcí fitness. Nyní jsem tedy doufal, že by tamějšího editora tato nová studie, kde již dlouhodobější měřítko bylo, mohla zaujmout. Zřejmě se tak i stalo, protože editor sice článek odmítl, ale jeho resubmitaci tentokrát povolil. Recenzenti ale měli dost námitek, článek byl pro ně těžko srozumitelný, dost technický a přehlcený různými výsledky a srovnáváním jednotlivých efektů. Musel jsem ho tedy značně zjednodušit a zkrátit. Odstranil jsem tedy část týkající se mixovaného cross-fosteringu i kvazi-metaanalytické srovnání s naší předchozí studií. V době přípravy tohoto manuskriptu jsem stejně již pracoval na skutečné metaanalýze (článek 10) a také jsem pokračoval se sběrem dat pro mixovaný cross-fosteringový experiment (článek 11). Proto mi vyškrtnutí těchto částí nijak nevadilo. Po těchto rozsáhlých úpravách byl článek přijat. Je to dokonce tzv. „cover paper", neboť na obálce časopisu je fotka samice lejska bělokrkého s potravou. Snímek pořídil a redakci na mou žádost poskytl amatérský fotograf přírody, pan Libor Šejna.

V této studii nepředstavovaly přímé genetické efekty ani rodičovské krmení problém pro odhad vlivu velikosti vejce na fitness mláďat. Přesto si ale myslím, že by bylo vhodné pro tyto potenciální matoucí

faktory v budoucích studiích kontrolovat, než jen doufat, že výsledky neovlivňují. Třeba u hmyzu bylo prokázáno, že bez podobné kontroly je význam velikosti vejce pro mláďata nadhodnocen (Fox 1997).

Obr. 8. Mláďata, která krmili oba rodiče, byla v šesti dnech větší, než ta, která krmil jen jeden z nich. Samotná frekvence krmení však na mláďata neměla pozorovatelný vliv.

Článek 10: Krist, M. (2011) Egg size and offspring quality: a meta-analysis in birds. *Biological Reviews*, 86, 692-716.

Při přípravě článků, které se zabývaly významem velikosti vajec pro fitness mláďat, mě napadalo mnoho dalších otázek. Například, zda velikost vejce více ovlivňuje přežívání mláďat nebo jejich velikost. Jestli tyto efekty vydrží až do dospělosti, nebo postupně vymizí. Nebo zda jsou efekty velikosti vajec závislé na tom, sledujeme-li mláďata v rámci jednoho hnízda nebo mezi hnízdy. Uvědomil jsem si, že tuto vznikající zvědavost bych mohl zkusit uspokojit tím, že provedu meta-analýzu publikovaných studií. Existovalo sice starší review Tonyho Williamse (1994), to ale bylo provedeno tzv. vote-counting způsobem, kdy se počítalo, kolik studií je signifikantních. To ovšem není úplně dobrá metoda, protože signifikance záleží jak na velikosti efektu, která má biologický význam, tak i na velikosti vzorku, která ovšem žádný biologický význam nemá. Navíc od publikace Williamsova review již uběhlo dost času, takže se nakumulovala řada nových studií.

Na meta-analýze jsem začal pracovat někdy v roce 2007. Tušil jsem, že poměrně dlouho může trvat i prosté získání všech článků, které by mohly obsahovat potřebná data. Tento odhad se nakonec ukázal jako správný, protože takových článků jsem potřeboval získat téměř 700. Tyto články jsem pak pročítal a data z nich vypisoval. Práce mi šla ale velmi pomalu. Článků bylo mnoho, někde data chyběla a bylo třeba kontaktovat autory o jejich doplnění. Případně jsem musel vymýšlet způsoby, jak zjistit velikost efektu u některých studií s neobvyklým designem. Snažil jsem se pro tuto mravenčí práci získat nějakého spolupracovníka, ale všichni oslovení kolegové mou nabídku spolupráce odmítli. Párkrát jsem to v průběhu vypisování dat již chtěl vzdát a raději zpracovávat vlastní data získaná z terénu, ale nakonec jsem úkol dotáhl do konce. Výsledná meta-analýza byla v té době v ekologii asi vůbec největší, co se týče počtu zahrnutých studií. Nabídl jsem ji k publikaci časopisu Biological Reviews, kde vyšla předchozí velká review o ptačích vejcích (Williams 1994, Christians 2002).

Recenzenti sice měli k článku dlouhou řadu připomínek, ty ale byly spíše technického charakteru. Jeden z recenzentů například výrazně kritizoval délku Metod. Zejména pro časopis jako je Biological Review

mu jejich popis připadal neadekvátně dlouhý a navrhoval je přesunout do Apendixu. Protože v dané době se v tomto časopise ještě moc meta-analýz nepublikovalo, nevěděl jsem, jak s touto připomínkou naložit. Preferoval jsem ale popis metod přímo v článku. Proto jsem se nejdříve zeptal editora, zda by to povolil. Ten s tím souhlasil, tak jsem v tomto ohledu ponechal článek nezměněný. Jinak jsem provedl dost úprav (rebuttal měl 58 bodů), třeba jsem v jedné analýze musel zohlednit fylogenezi mezi zahrnutými 162 druhy. Bylo to ještě před publikací komprehenzivní fylogeneze ptáků (Jetz et al. 2012), proto jsem při konstrukci fylogenetického stromu musel vycházet z více dílčích zdrojů. S fylogenezí jsem pracoval poprvé, tudíž to pro mě bylo docela zábavné i poučné. Výsledky však nebyly fylogenetickými vztahy mezi druhy nijak ovlivněny.

Meta-analýza odpověděla na otázky, které jsem si původně kladl, ale řadu dalších zase naopak nastolila. Zajímavým výsledkem třeba bylo zjištění, že je vliv vejce na mláďata silnější v laboratorních podmínkách. To by se dalo vysvětlit lepší kontrolou matoucích faktorů v laboratoři. Teoretické modely ale jinak předpovídají, že velikost vejce bude mít větší význam pro mláďata ve špatném prostředí, kdežto v dobrém prostředí, jako bylo i to laboratorní, by na velikosti vajec nemělo příliš záležet. Rozhodli jsme se proto vztah mezi kvalitou postnatálního prostředí a efektem velikosti vejce na mláďata studovat podrobněji (články 13 a 14).

Článek 11: Krist, M. & Munclinger, P. (2011) Superiority of extra-pair offspring: maternal but not genetic effects as revealed by a mixed cross-fostering design. *Molecular Ecology*, 20, 5074-5091.

Znalost vejce, ze kterého se dané mládě vylíhlo, je užitečná pro testování celé řady hypotéz (viz články 2–4). Líhnutí mláďat v líhni sice takové výsledky přineslo, práce ale byla kvůli roznášení vylíhlých mláďat zpět do hnízd a technickým problémům velmi náročná. Přemýšleli jsme tedy, jestli by tato zajímavá data nešla získat i jiným způsobem. Napadlo nás řešení kombinující terénní experiment a molekulární metody. V letech 2006-2009 jsme tedy prohazovali vejce mezi hnízdy tak, aby každé nové hnízdo obsahovalo jen jedno vejce z hnízda původního. Protože samice lejsků nekladou svá vejce do cizích hnízd, matka je v naší populaci vždy jistá. Když jsme tedy u mláďat posléze určili maternitu, věděli jsme také, z jakého vejce mládě pochází. Experiment byl pro sběr dat daleko efektivnější než předchozí kontrolované líhnutí.

Smyslem tohoto článku bylo: 1) Popsat tuto novou možnost zisku zajímavých ekologických dat pomocí molekulárních metod. 2) Ověřit, zda vztah mezi paternitou a pořadím kladení, který jsme zjistili v předchozí menší studii (článek 4), bude pozorovatelný i na větším nezávislém vzorku. 3) Otestovat, zda mají extra-párová mláďata vyšší fitness, pokud se líhnou synchronně s mláďaty párovými. Stejně jako v článku 4 jsme i na tomto vzorku zjistili, že extra-párová mláďata jsou častější v prvních vejcích a chybí v těch posledních. Prospívání mláďat nezáleželo na paternitě, což ukázalo, že extra-párová mláďata nenesou žádnou genetickou výhodu a důvody pro samičí zahýbání tedy musí být jiné, než zisk „dobrých genů". Sepsaný článek jsme zaslali do časopisu Molecular Ecology. Prvnímu recenzentovi se práce líbila a navrhl jen „minor revision". Druhý recenzent ale argumentoval, že náš experiment nemůže na otázku genetických efektů odpovědět, protože pro jejich test musí být poloviční sourozenci vychováváni stejnými rodiči v jednom hnízdě, a navrhl tak odmítnutí článku. Třetí recenzent, kterým byl Bart Kempenaers, měl námitky velmi podobné druhému recenzentovi. Nakonec ale napsal, že vzhledem k tomu, kolik nám to muselo dát práce, doufá, že se mýlí a že to s námi rád prodiskutuje. Navrhl proto článek odmítnout, ale povolit jeho resubmitaci. Takto nakonec editor i rozhodl.

Barta Kempenaerse se nám nakonec podařilo o smysluplnosti našeho experimentu přesvědčit. Opět nám v tom pomohl i Ben Sheldon, na kterého jsme se obrátili s dotazem, zda si myslí, že je náš experiment pro test dané hypotézy relevantní. Ben nám odpověděl, že se domnívá, že je to ještě lepší než srovnání polovičních sourozenců v normálních hnízdech, protože testujeme případné genetické benefity v různém environmentálním pozadí, což je pro následné úvahy o selekci na samičí chování výhodou. V dopise nám dovolil tento názor citovat i Bartu Kempenaersovi. Výsledkem celé obsáhlé korespondence a dlouhého rebuttalu tak nakonec bylo, že Bart Kempenaers navrhl přijetí revidovaného rukopisu, kdežto druhého recenzenta naše argumenty nepřesvědčily a tak navrhl jeho odmítnutí. Editor pak měl asi nelehké rozhodování, ale článek nakonec přijal.

Byli jsme rádi, že je naše práce publikována v tomto předním ekologickém časopise. Doufali jsme, že myšlenky, které jsme zde prezentovali, pomohou obor posunout dále a že metoda tzv. mixovaného cross-fosteringu se stane užitečným nástrojem pro studium mateřských a genetických efektů u dalších druhů. To se ale nestalo, článek víceméně zapadl. Možná vypadá metoda mixovaného cross-fosteringu moc složitě, i když my si v praxi ověřili, že je snazší než kontrolované líhnutí mláďat. Práci by možná také bývalo prospělo, kdybychom ji rozdělili na metodickou a datovou část, jako jsme to udělali u článků 2 a 3.
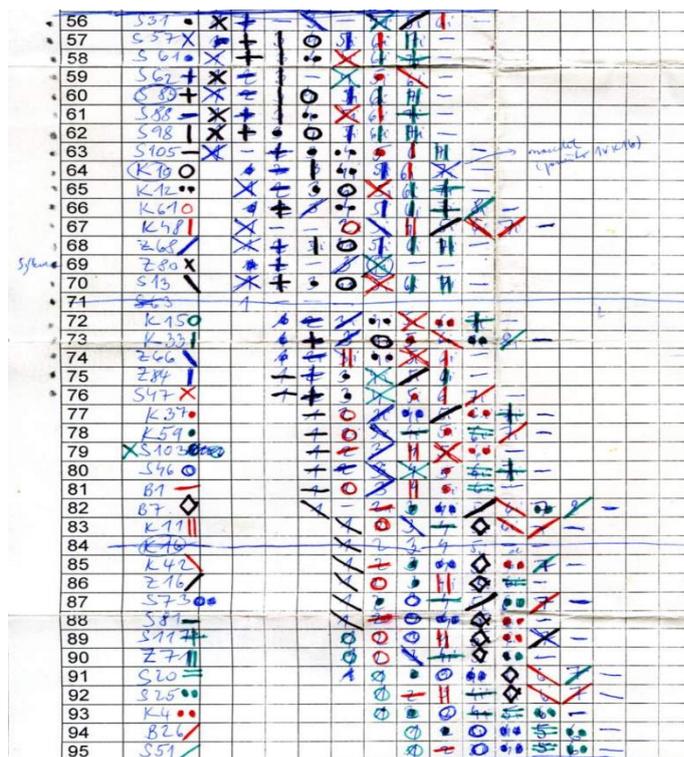
Obr. 9. Terénní protokol pro mixovaný cross-fostering ukazující hnízda (řádky), dny (sloupce) a jednotlivá snesená vejce (buňky). Symboly vyznačují hnízda, kam byla vejce po skončení kladení přenesena. Mixovaný cross-fostering lze provádět na velkém vzorku a přehled ve stovkách vajec si lze snadno udržet pomocí několika barevných symbolů. Přes efektivitu tohoto přístupu ho však jiní výzkumníci zatím nepoužili.

Článek 12: Bowers, E. K., Munclinger, P., Bureš, S., Kučerová, L., Nádvorník, P. & Krist, M. (2013) Cross-fostering eggs reveals that female collared flycatchers adjust clutch sex ratios according to parental ability to invest in offspring. *Molecular Ecology*, 22, 215-228.

Původním záměrem mé doktorské práce, kterou jsem obhájil v roce 2004, bylo studovat poměr pohlaví ve snůškách lejska bělokrkého. Tato data však ještě v roce 2011 stále nebyla publikovaná. Nejdříve jsme tento úkol odkládali kvůli práci na jiných tématech. Poté jsme se rozhodli doplnit původní dataset,

získaný kontrolovaným líhnutím mláďat v letech 2001-2002 (viz článek 2), o další data získaná mixovaným cross-fosteringem v letech 2006–2009. Tento druhý dataset byl k dispozici v lednu 2011, kdy mi kolega Pavel Munclinger zaslal genetická data. Začali jsme tedy pracovat na článku o mixovaném cross-fosteringu (viz článek 11), zároveň jsme ovšem mysleli i na publikaci týkající se sexuální alokace. Věděl jsem, že zpracovat toto druhé téma do sezóny 2012 nestihnu, protože jsme ve Vlastivědném muzeu museli pracovat na nové přírodovědné expozici.

V té době mě zaujala nová publikace, která se věnovala sexuální alokaci u střízlíka zahradního, *Troglodytes aedon* (Bowers et al. 2011). Pod vedením dvou zkušených školitelů tuto práci sepsal magisterský student Keith Bowers. Protože se mi jejich práce líbila, nabídl jsem Keithovi spolupráci na zpracování našich dat o poměru pohlaví u lejsků. Keithe nabídka potěšila a po domluvě se svými školiteli spolupráci přijal. Dost dlouho jsme potom po e-mailu ladili detaily, jak data analyzovat, jaké hypotézy testovat a jak argumentovat. Keith pak pod mým dohledem data zanalyzoval a článek sepsal. Myslím, že šlo o spolupráci výhodnou pro obě strany. Já byl každopádně s přístupem tohoto mladého talentovaného vědce neobyčejně spokojen. Byl velmi chápavý, přes své mládí již hodně sečtělý a na mé komentáře a připomínky reagoval pružně a rychle. Měl otevřenou mysl a nebyl pro něj problém pozměnit názor a do analýz nebo textu zakomponovat nějakou další perspektivu.

Manuskript jsme nejdříve nabídli redakci American Naturalist, tam ho ale editorka po vlastním pročtení odmítla, protože jí nepřipadal dostatečně inovativní. Článek měl navíc při této submitaci jiný titulek: „Females adjust clutch sex ratios, but not the sex of individual offspring: an experimental study in the collared flycatcher" a editorka nás upozornila, že nejde v pravém smyslu slova o experiment. Za experiment by považovala jen práci, ve které bychom manipulovali nějaký faktor, na který by ptáci reagovali změnou poměru pohlaví ve svých snůškách. Tento argument nás ale příliš nepřesvědčil, tak jsme článek bez úprav poslali do časopisu Journal of Animal Ecology. I zde byla tato práce odmítnuta editorem bez recenze. Rozhodli jsme se proto pro její zaslání do Molecular Ecology. Zde byl článek poslán třem recenzentům. Dva z nich ho hodnotili velmi pozitivně, třetí byl kritičtější, ale v zásadě se i jemu práce docela líbila. Dost přesně však zopakoval i připomínku ohledně experimentální manipulace nezávislé proměnné, kterou měla editorka v American Naturalist. Museli jsme tedy uznat, že na tomto názoru asi něco bude. Spolu se spoustou dalších úprav jsme proto i titulek změnili a slovo "experiment" z něj vypustili. Editor pak již ani nevyžadoval další kolo recenzí a článek přijal.

Obr. 10. Pohlaví mláďat nezáviselo na sekvenci kladení. Mláďata z prvních i posledních vajec tak byla se stejnou pravděpodobností samci i samice.

Článek 13: Krist, M. & Munclinger, P. (2015) Context dependence of maternal effects: testing assumptions of optimal egg size, differential, and sex allocation models. *Ecology*, **96, 2726-2736.**

Z naší meta-analýzy vyplynulo, že není dobře známo, zda dopad rodičovské péče na mláďata závisí na kvalitě prostředí (viz článek 10). Také jsme se zabývali otázkami, zda rodičovské investice souvisí s atraktivitou partnerů (článek 8) nebo pohlavím potomků (články 6, 12). Uvědomili jsme si, že všechny tyto problémy jsou si podobné v tom, že předpokládají závislost výnosů z rodičovských investic na daném kontextu a spadají tak do jednoho vysvětlujícího rámce (viz i Ratikainen et al. 2018). Napadlo nás také, že naše data získaná mixovaným cross-fosteringem se dají použít i pro test těchto příbuzných hypotéz. Protože tento experiment probíhal již v letech 2006–2009, mohli jsme dokonce sledovat vliv rodičovských investic i na celoživotní reprodukční úspěch potomků, což se v předešlých studiích podařilo spíše výjimečně.

Analýza dat pak prokázala závislost mateřských efektů na kvalitě postnatálního prostředí, jak implicitně předpokládají teorie optimální velikosti vejce i diferenciální alokace/kompenzace. Naopak jsme nezjistili, že by mláďata samčího pohlaví byla na intenzitě rodičovské péče závislejší než mladé samice, jak předpokládá teorie sexuální alokace (Trivers & Willard 1973). Sepsaný manuskript jsme nabídli do časopisu Ecology. Recenzenti i Tony Williams, který opět (viz komentář k článku 2) dostal náš rukopis na starost, oceňovali naše dlouhodobá data i inovativní typ experimentu. Všichni tři však měli námitky proti tomu, abychom průměrnou hmotnost mláďat v hnízdech označovali jako „environmental quality", jak jsme v první verzi rukopisu dělali. Nelíbilo se jim to hlavně proto, že hmotnost mláďat může být ovlivněna i mateřskými efekty, nejen kvalitou prostředí v užším slova smyslu. To bylo naštěstí spíše jen sémantické nedorozumění.

Samozřejmě jsme si byli vědomi možnosti mateřských efektů. Ale i tyto mateřské efekty lze z hlediska mláďat označit jako vliv prostředí, jak se to dělá v kvantitativní genetice. Naše vysvětlení jsme tedy vylepšili a pro jistotu ještě proměnnou „environmental quality" přejmenovali na zřejmě méně kontroverzní „postnatal conditions". Těmito změnami se nám podařilo recenzenty i editora o smysluplnosti našich analýz přesvědčit. Jeden nově přizvaný recenzent nás ale upozornil na to, že interpretace našich dat závisí na tom, jestli atraktivní samci pečují o své potomky lépe než ti méně preferovaní. Tuto perspektivu jsme proto do článku přidali, i když naše závěry, týkající se teorie diferenciální alokace, pak byly poněkud těžkopádné.

Obr. 11. Při provádění mixovaného cross-fosteringu byla lejsčí vejce uložena na bezpečném místě, zatímco samice inkubovaly jen atrapy. Vejce z původních hnízd si byla velikostně i barevně podobná. Jejich výměnou jsme zvýšili variabilitu v rámci nových snůšek, což nám napomohlo při testování interakce tohoto faktoru s kvalitou prostředí.

Článek 14: Krist, M., Janča, M., Edme, A. & Dzuro, R. (2015) Are prenatal maternal resources more important in competitive than in benign postnatal environments? *Auk*, **132, 577-583.**

V této práci jsme se zabývali otázkou, zda je pro ptáky výhodné klást velká vejce, pokud se mláďata vyvíjejí v kompetitivním prostředí. Pokud by tomu tak bylo, dala by se variabilita ve velikosti vajec mezi populacemi a případně i mezi druhy vysvětlit jako adaptace k danému prostředí (Fox & Czesak 2000, Riesch et al. 2014). Pokud by se naopak vliv velikosti vajec na fitness mláďat v jednotlivých typech prostředí nelišil, naznačovalo by to, že mezipopulační variabilita vzniká spíše v důsledku nějakých omezení (constraints) a nemá tedy adaptivní význam. Podobným problémem jsme se částečně zabývali i v předchozí studii (článek 13). Ta však byla obecněji zaměřena a navíc jsme v ní nemanipulovali kvalitu prostředí, ve kterém se mláďata vyvíjela, jen jsme použili cross-fosteringový experiment pro přesnější zjištění korelací mezi velikostmi vajec a kvalitou mláďat v hnízdech.

V této studii jsme ale manipulovali kvalitu prostředí pomocí nestejných cross-fosteringů, čímž jsme provedli experiment v úzkém slova smyslu (viz komentář k článku 12). Hnízda se zvětšeným počtem mláďat měla představovat kompetitivní prostředí, kdežto ta se zmenšeným počtem prostředí s malou konkurencí (viz i Bonisoli-Alquati et al. 2008). Naše manipulace byla v tomto směru úspěšná, hnízda s větším počtem mláďat měla horší reprodukční parametry. Hlavně nás ale zajímalo, zda je v datech patrná interakce mezi treatmentem (zvětšená vs. zmenšená hnízda) a velikostí vejce. Žádnou výraznou interakci jsme však nezjistili. Velká vejce byla pro kvalitu mláďat stejně prospěšná ve zvětšených i zmenšených hnízdech. Naše výsledky tak ukázaly spíše na možnost, že mezipopulační variabilita ve velikosti vajec vzniká v důsledku evolučních omezení.

Článek jsme poslali do časopisu Journal of Avian Biology, kde byl ovšem bez recenze odmítnut. Naše další volba byl časopis Auk. Zde byl manuskript dvěma recenzenty hodnocen docela pozitivně, takže ho editor po množství drobnějších úprav přijal k publikaci. Jeden z recenzentů se domníval, že jde o důležité téma, které bude stimulovat další podobný výzkum. To byl i náš záměr, protože podobných studií je u ptáků velmi málo (viz článek 10). Zatím to ale spíš vypadá, že tato naše práce zcela zapadla.

Obr. 12. Kvalita mláďat z velkých hnízd, jako bylo toto se sedmi 13-denními mláďaty, nezávisela na velikosti vajec více, než tomu bylo u malých hnízd.

V roce jsme 2012 jsme s kolegy Tomášem Albrechtem z Ústavu biologie obratlovců a Pavlem Munclingerem z Přírodovědecké fakulty Univerzity Karlovy získali grant GAČR, který se měl zabývat post-kopulačním pohlavním výběrem u pěvců. Pro pomoc při řešení tohoto grantu jsem hledal nějakého spolupracovníka. Tím se nakonec stala francouzská studentka Anaïs Edme, jež na podzim 2013 začala v Olomouci pod mým vedením své doktorské studium. Post-kopulační pohlavní výběr zahrnuje procesy jako je kompetice spermií a kryptická samičí volba, které byly v té době u ptáků velmi málo probádané. Nechtěli jsme však s psaním článků čekat, až budeme mít nasbíráno dost dat o morfologii lejsčích spermií, protože pak by Anaïs mohla mít problém dokončit svou práci v řádném termínu. Napadlo nás, že dataset získaný v předchozích letech mixovaným cross-fosteringem by šlo využít pro další analýzy extra-párové paternity, neboť ta s post-kopulačním pohlavním výběrem úzce souvisí.

Do té doby jsme totiž studovali hlavně rozdíly mezi polovičními sourozenci v hnízdech s namíchanou paternitou, tedy ultimátní výsledek extrapárových kopulací (články 4 a 11). Teď jsme se proto chtěli zaměřit na proximátní rovinu a zjistit, s jakými samci vlastně samice svým sociálním partnerům zahýbají. Nechtěli jsme ale zůstat jen u jednoduchého srovnání fenotypu sociálních a extra-párových otců, protože takových testů již bylo v jiných populacích provedeno více (Akçay & Roughgarden 2007). Málokdo se však zabýval otázkou, jestli se extra-párový partner nejen liší od sociálního partnera, ale zda je i samicí nenáhodně vybrán ze samců, kteří se v jejím okolí vyskytují. Přidali jsme proto do našich srovnání i časoprostorové hledisko. Tento aspekt naší práce oba recenzenty v Journal of Avian Biology zaujal a proto práci editorovi doporučili. Museli jsme ještě vylepšit logickou strukturu článku, ale poté již editor Simon Griffith, který se sám extra-párovou paternitou zabýval (Griffith et al. 2002), naši práci přijal.

V práci jsme trochu překvapivě zjistili, že podvádění samci mají větší ornament – křídelní skvrnu, než samci, kterým zůstaly samice věrné. Na druhou stranu starší samci, kteří mají tuto skvrnu větší, zplodili více extra-párových mláďat v jiných hnízdech. Často u svých sousedů. Pro interpretaci těchto výsledků by bylo dobré vědět, kdo vlastně extra-párové kopulace iniciuje. Létají samice aktivně za cizími samci, nebo jen akceptují jejich (ne)vítaný zájem? A proč se tomu vlastně domácí samci více nebrání a samice si nehlídají? Odpovědi na tyto mechanistické otázky zatím často chybějí, protože kopulační chování se studuje mnohem hůře než jeho výsledek – paternita, která se celkem rychle ověří z kapky krve. Díky technologickému pokroku se však i možnosti studia těchto behaviorálních mechanismů v poslední době zvyšují. Zálety do cizích teritorií se dají dobře dokumentovat například pomocí automatické telemetrie (Ward et al. 2014). Další možností automatického záznamu jsou v případě druhů hnízdících v dutinách RFID čtečky (Schlicht et al. 2015). Právě touto technologií plánujeme v budoucnu zjišťovat pohyby samců i samic lejsků mimo jejich teritoria (viz Závěr).

Obr. 13. Staří samci s velkou křídelní skvrnou vychovávají ve svých hnízdech více extra-párových mláďat, ale taky jich více zplodí v jiných hnízdech.

Článek 16: Edme, A., Zobač, P., Opatová, P., Šplíchalová, P., Munclinger, P., Albrecht, T. & Krist, M. (2017) Do ornaments, arrival date, and sperm size influence mating and paternity success in the collared flycatcher? *Behavioral Ecology and Sociobiology*, **DOI 10.1007/s00265-016-2242-8.**

V předchozích dvou experimentech, ve kterých jsme se zabývali extra-párovou paternitou (kontrolované líhnutí v letech 2001-2002 a mixovaný cross-fostering v letech 2006-2009), jsme nebyli schopni zjistit celkový reprodukční úspěch samců. Krev jsme totiž pro genetické analýzy odebírali jen mláďatům z hnízd, které byly do těchto dvou experimentů zahrnuty. Mohli jsme tak sice popsat ztráty paternity ve vlastních hnízdech samců, ale nikoli jejich úspěch v cizích (nevzorkovaných) hnízdech. Tuto informaci jsme tak chtěli doplnit, abychom získali komplexnější obrázek o kompetici spermií a pohlavním výběru v naší populaci. Zároveň jsme se již dlouhou dobu zajímali o to, zda se může lišit funkce bílých ornamentů lejsků mezi populacemi. Například čelní skvrna se ukazovala jako důležitý znak pro samičí volbu sociálního (Qvarnström et al. 2000) i extrapárového (Sheldon et al. 1997) partnera ve Švédsku, nikoli však v Maďarsku (Hegyi et al. 2002, Rosivall et al. 2009) nebo v naší populaci (Krist 2002, článek 15). Závěry ze švédské populace byly založeny jak na korelativních (Sheldon et al. 1997) tak i na experimentálních datech, kdy výzkumníci manipulovali velikost čelní skvrny (Qvarnström et al. 2000). Jinde se však podobný experimentální přístup dosud nevyzkoušel. Rozhodli jsme se tedy posledně jmenovanou studii částečně replikovat.

V roce 2013 jsme proto s mnoha terénními spolupracovníky chytali samce lejsků hned při jejich příletu na lokalitu. Odebrali jsme jim vzorek spermií, zmanipulovali velikost čelní skvrny a testovali vliv těchto proměnných na celkový reprodukční úspěch samců, tedy včetně zisků extrapárové paternity v cizích hnízdech. Před touto sezónou jsme si na zebřičkách vyzkoušeli, jak na peří drží dvě bílé barvy – Tippex, který používali kolegové ve Švédsku, a Alteco. Zjistili jsme, že Alteco je trvanlivější, proto jsme lejsky manipulovali touto barvou. Naše výsledky byly z velké části negativní. Délka spermií ani naše manipulace čelní skvrny neměla na reprodukční úspěch samců vliv. Pokud tam nějaký vztah byl, tak spíše v neočekávaném směru – samci s nabělenými čely se samicím moc nelíbili. Mohlo tomu tak být proto, že tato umělá bílá skvrna, ať již natřená Altecem nebo Tippexem, měla reflektanční spektrum dost odlišné od normálního bílého peří.

Manuskript jsme poslali do časopisu Behavioral Ecology and Sociobiology, kde tři recenzenty zaujal. Měli sice řadu drobnějších připomínek, které jsme pak v revizích zakomponovali, ale oceňovali náš experimentální přístup, detailní popis samčího reprodukčního úspěchu i diskusi možných metodických problémů spojených s manipulací ptačích ornamentů. Jeden z recenzentů také chválil naše promyšlené srovnání s předchozí studií Qvarnström et al. (2000). Za nějaký čas nás proto překvapilo, že Anně Qvarnström a jejímu týmu naše studie naopak nepřipadala užitečná, i když testovali na stejném druhu stejnou otázku – také se zabývali vlivem velikosti čelní skvrny a délky spermií na reprodukční úspěch samců (Ålund et al. 2018). Na rozdíl od nás švédský tým opět našel statisticky významné závislosti. Bylo to ovšem jen ve dvoucestných nebo dokonce trojcestných (spolu s věkem samců) interakcích těchto faktorů. My vlastně nevíme, jestli nejsou podobné vztahy i v našich datech, protože jsme do našich modelů zahrnuli délku spermií a manipulaci čelní skvrny jen jako hlavní efekty. V souladu s varováním Forstmeiera & Schielzetha (2011) se totiž obáváme, že s testováním velkého množství interakcí, které nejsou a priori predikovány, značně roste riziko chyby I. druhu a tedy nalezení signifikantních výsledků, jež vznikají pouhou náhodou. Proto zásadně testujeme jenom interakce a priori predikované, jako byla ta mezi manipulací čelní skvrny a sezónou. Touto interakcí jsme totiž replikovali analýzu z předchozí studie (Qvarnström et al. 2000).

Obr. 14. Manipulace čelní skvrny neměla výrazný vliv na reprodukční úspěch samců. Vlevo: zmenšená skvrna, uprostřed: přirozená skvrna, vpravo: zvětšená skvrna.

Článek 17: Edme, A., Zobač, P., Korsten, P., Albrecht, T., Schmoll, T. & Krist, M. (2019) Moderate heritability and low evolvability of sperm morphology in a species with high risk of sperm competition, the collared flycatcher *Ficedula albicollis*. *Journal of Evolutionary Biology*, 32, 205–217.

V sezóně 2013 jsme si ověřili, že jsme schopni účinně prochytávat populaci samců už při jejich příletu na lokalitu. Příletová data samců jsou cenná sama o sobě (článek 18). Navíc jenom část samců je možné chytit i následně při hnízdění (článek 16). Protože jsme potřebovali velký vzorek samců pro studium migrace pomocí geolokátorů (Briedis et al. 2018 a,b) i pro studium morfologie jejich spermií (tato práce), rozhodli jsme se v jarních odchytech pokračovat ještě v následujících třech sezónách (2014–2016). Samcům jsme při příletu i při následném hnízdění odebírali vzorek spermií s cílem zjistit, zda je tento znak v rámci sezóny i mezi sezónami opakovatelný a zda je heritabilní. Předpokládali jsme, že heritabilita by u lejsků, jež mají vysoké procento extrapárových mláďat, mohla být nízká kvůli silné selekci na optimální fenotyp spermie (Lifjeld et al. 2010).

Anaïs se kvantitativně genetickým analýzám učila na své stáži v Bielefeldu (Německo), kde pobývala v laboratoři Tima Schmolla. Na budování tzv. „animal modelu" dále spolupracovala i s Peterem Korstenem. Analýza dat skutečně prokázala nižší heritabilitu délky spermií u lejska než tomu bylo u zebřičky, kde asi na délce spermií tolik nezáleží, protože zebřičky jsou si věrnější a nedochází tak u nich ke kompetici spermií. Kromě tohoto genetického aspektu jsme také zjistili, že se spermie v průběhu sezóny trochu prodlužovaly. Sepsanou práci jsme nabídli časopisu Evolution, kde byla editorkou Annou Qvarnström odmítnuta na základě jedné negativní recenze. Recenzent pochyboval o kvalitě našich dat, protože jsme chytali lejsky jen před a po snášení vajec. Tedy ne při vlastním kladení, kdy jsou spermie nejvíce potřeba. Nám se zase nelíbilo, že náš rukopis měl jen jednoho recenzenta. Postěžovali jsme si na tento fakt hlavní editorce, Mario Servedio, kterou jsme žádali o posouzení naší práce ještě druhým recenzentem. Mario Servedio nám vysvětlila, že i časopis tohoto ranku mívá problém sehnat recenzenty a přenechala rozhodnutí, zda přizvat dalšího recenzenta, na subject editorce Anně Qvarnström. Ta nám pak odpověděla, že přizvání druhého recenzenta by bylo zbytečné, když ten první měl tak zásadní připomínky.

Druhý pokus jsme směřovali do Journal of Animal Ecology. Subject editor náš manuskript důkladně pročetl a okomentoval. K publikaci ho ale nedoporučil, protože mu v něm chybělo propojení dat o heritabilitě spermií s jejich efektem na fitness samců. Dalším časopisem, kam jsme rukopis poslali, byl Journal of Evolutionary Biology. Zde byl po recenzích odmítnut s možností resubmitace. Recenzenti požadovali dost úprav, včetně zohlednění možných mateřských efektů v našem kvantitativně genetickém modelu. Protože Anaïs v danou dobu již měla obhájenou svou dizertaci a nepracovala

ve vědě, neměla na takovou větší re-analýzu čas. Nakonec jsem se rozhodl tyto analýzy provést sám, i když se mi z počátku do toho moc nechtělo. Do té doby jsem totiž ještě neanalyzoval data v R-ku, což ale byla v případě této kvantitativně genetické studie nutnost. Nyní jsem rád, že jsem tento svůj odpor musel prolomit a seznámil se s analytickými i grafickými možnostmi, které R-ko nabízí. Po vyhovění recenzentům byl již náš rukopis přijat a stal se tak první prací, která stanovila heritabilitu morfologie spermií v přírodní populaci.
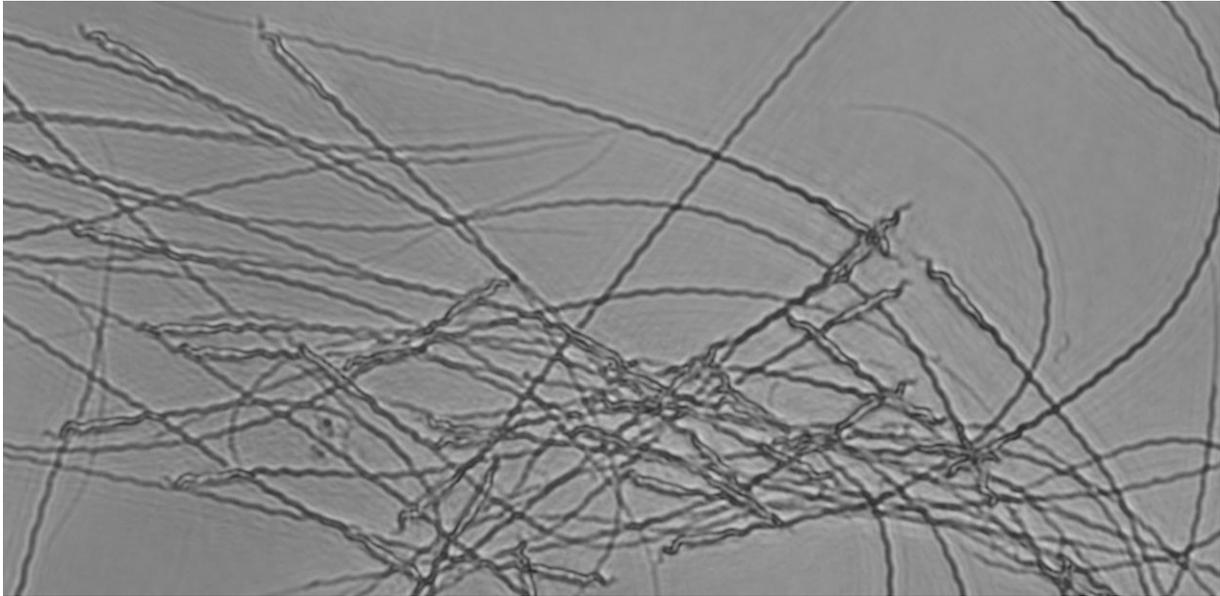
Obr. 15. Délka spermií byla opakovatelná, heritabilní a v sezóně se zvětšovala. Foto: Anaïs Edme, zvětšení 400x.

Manuskript 18: Krist, M., Munclinger, P., Briedis, M. & Adamík, P. (2020) Genetic regulation of avian migration timing: combining candidate gene and quantitative genetic approaches in a long-distance migrant. Zasláno do časopisu Oecologia.

V době globálního oteplování se migrantům vyplácí přilétat na hnízdiště dříve, neboť tak mohou obsadit dobrá teritoria bez velkého rizika, že je zahubí chladné počasí. Navíc se posunuje i fenologie jejich kořisti, třeba housenek, kterých by později v sezóně již nemuselo být dostatek (Both et al. 2006). Migranti na kratší vzdálenost se při načasování odletu ze zimoviště mohou řídit změnami fotoperiody a teploty. Migranti na delší vzdálenost tyto klíče nemusí mít k dispozici a tak se předpokládá, že jejich roční cyklus je silněji kontrolován jejich vnitřními hodinami (Åkesson et al. 2017). Aby tyto druhy mohly reagovat na dlouhodobé změny klimatu, musí mít jejich populace genetickou variabilitu na lokusech, které kódují vnitřní hodiny. Jedním z lokusů, kterým byla v této souvislosti v poslední době věnována pozornost, je CLOCK. Ten primárně řídí denní cykly, ale zároveň možná ovlivňuje i načasování migrace (Bazzi et al. 2015).

Rozhodli jsme se tyto hypotézy otestovat na našich dvou datasetech. Jednak jsme měli příletová data na hnízdiště zjištěná při jarních odchytech samců, jednak jsme část z těchto samců sledovali pomocí geolokátorů. Na obou datasetech jsme zjišťovali vztah mezi načasováním migrace a alelickou variabilitou na kandidátských genech pro roční cykly. Variabilitu v příletových datech jsme navíc rozložili na jednotlivé kauzální komponenty pomocí animal modelu. Přílety na hnízdiště byly pro jednotlivé samce opakovatelné, ale nebyly heritabilní. Opakovatelnost příletů tedy byla dána spíše

trvalými účinky prostředí než geneticky. V souladu s tímto výstupem z kvantitativně genetického modelu jsme nezjistili žádný vztah mezi načasováním migrace a alelickou variabilitou na testovaných kandidátských genech. Naše výsledky tedy byly na rozdíl od mnoha předchozích studií (přehled viz Ralston et al. 2019) negativní, přestože jsme pro testování použili podobnou nebo i větší velikost vzorku.

Manuskript jsme nejprve nabídli k publikaci časopisu Molecular Ecology. Zde ho však editorka Yanhua Qu odmítla bez zaslání na recenze. Měla proti němu více námitek, její hlavní argument ale zněl, že naše výsledky jsou negativní a tudíž nepřináší nic nového. Kvůli našemu výcviku v meta-analytickém uvažování nás tento argument skoro pohoršil. Napsali jsme proto hlavnímu editorovi, Lorenu Riesebergovi, že podobná praxe odmítání manuskriptů na základě směru výsledků vede k publikačnímu bias, což má negativní dopady na pokrok oboru. Hlavní editor nám sice odpověděl, že máme pravdu a že to na příští redakční radě proberou, nicméně tím to pro něj skončilo. Napodruhé jsme proto naši práci nabídli do Functional Ecology. Zde ji na recenze poslali. Recenzenti měli k našemu rukopisu řadu kritických připomínek, ani jeden ho zřejmě k přijetí nedoporučil, takže odmítnutí editorem nebylo příliš překvapivé. Zarazilo nás ovšem, že i tento editor ve svém odůvodnění zmínil fakt, že naše výsledky jsou nesignifikantní. To utvrdilo naše podezření, že časopisy s vysokým impakt faktorem si toto měřítko hlídají i za cenu publikačního bias.

Po zkrácení manuskriptu a doplnění analýz podle doporučení recenzentů jsme proto článek poslali do časopisu Oecologia, kde nedávno vyšla podobně zaměřená studie s nesignifikantními výsledky (Parody-Merino et al. 2019). Věříme proto, že tato redakce bude naši studii posuzovat na základě jejího designu, metod a velikosti vzorku, nikoliv na základě směru výsledků.

Obr. 16. Data z geolokátorů potvrdila naši schopnost chytit lejsky brzy po příletu na lokalitu. Načasování migrace nebylo heritabilní a nemělo vztah ke studovaným kandidátským genům.

Závěr

Výzkum na budkových populacích ptáků lze lépe plánovat, než je tomu u volně žijících druhů. U těch je totiž třeba nejprve najít hnízda a pak ještě doufat, že je predátoři nezlikvidují. Pokud tedy nechceme studovat právě hnízdní predaci (Weidinger 2009). Hnízdní budky se však musí udržovat a když je naším záměrem získat dostatečně velký vzorek originálních dat, tak bývá i terénní práce na budkové populaci docela únavná rutina. Výzkumník se pak často ke konci sezóny těší, až si v klidu sedne do pohodlné kanceláře.

Většina z přiložených článků je založena na nějakém krátkodobém pokusu či pozorování. Je to dáno tím, že šlo o výzkum v relativně nově vytvořené a sledované populaci. Z počátku zde proto ani nebyla nashromážděna žádná dlouhodobá data. Recenzenti si ale často cenili i střednědobých dat, třeba těch o přežívání mláďat. Měřit skutečnou fitness jedinců je totiž stále dost obtížné a proto se častěji používají jen její krátkodobé náhražky, které ovšem kvůli různým trade-offům o skutečné fitness zas tolik nemusí vypovídat (Hunt et al. 2004). Zpoždění publikace za provedeným pokusem, ať již kvůli nedostatku času na přípravu rukopisu nebo kvůli jeho předchozímu odmítání, nám tak několikrát umožnilo odhadnout celkovou fitness jedinců. Díky tomu jsme pak tyto výsledky prosadili do kvalitních časopisů.

Dlouhodobý výzkum je potřeba, pokud chceme určovat kvantitativně genetické parametry jako je třeba heritabilita morfologických (článek 17) nebo behaviorálních (manuskript 18) znaků. Pro správný odhad těchto parametrů musíme mít dobře popsaný rodokmen dané populace. V našem případě jsme příbuznost mezi jedinci dané generace odhadovali pomocí rodokmenu jež sahal deset generací do minulosti. V dnešní době se sice stává reálnou alternativou stanovení příbuzenské matice na základě molekulárních markerů (Speed & Balding 2015), ale tato metoda je oproti jednoduché konstrukci rodokmenu na základě znalosti populační historie drahá.

Bez dlouhodobých dat se dále neobejdeme, pokud nás zajímají změny v reprodukčních parametrech nebo fenologii populace třeba vzhledem k měnícímu se klimatu (Clutton-Brock & Sheldon 2010). Pro tyto analýzy stačívají jen základní bionomická data jako je doba hnízdění nebo velikost snůšky, která si sbírá každý výzkumník. Důležité pro inference je v tomto případě právě časoprostorové hledisko. Před časem proto vznikla na ekologickém institutu v nizozemském Groningenu iniciativa nazvaná SPI-Birds Network and Database (Culina et al. 2020), která se snaží centralizovat data z jednotlivých lokalit. V současnosti je zde zahrnuto již přes 70 populací, včetně té námi studované na Velkém Kosíři. Naše data tak mohou v budoucnosti posloužit pro různé studie, které srovnávají hnízdní biologii pěvců na velké časoprostorové škále.

Při dlouhodobé studii je vždy dobré snažit se zaznamenat co nejvíce detailů o hnízdění. Užitečnost takových podrobných dat se někdy může ukázat až po delší době. My jsme například v roce 2019 byli osloveni holandským kolegou Jelmerem Samploniem, který se zajímá o interakce mezi sýkorami a lejsky, konkrétně, kolik lejsků sýkory v budkách zabíjely v minulosti a kolik je to teď. Globální změny klimatu totiž mohou měnit fenologii různých druhů různým způsobem a tak se mohou měnit i jejich vzájemné interakce (Adamík & Král 2008). Jelmer byl proto nadšen, když jsme mu poslali údaje o 73 lejscích zabitých sýkorami na Velkém Kosíři mezi lety 1998–2019.

Z výše uvedeného vyplývá, že dlouhodobý monitoring populací přináší možnosti, jež nejsou jen součtem možností v jednotlivých sezónách, ale výrazně tento součet převyšují. Na druhou stranu ale takový výzkum má i svá specifická úskalí (Clutton-Brock & Sheldon 2010). Nutné je totiž zajistit jeho průběžné financování a také spolupráci mezi generacemi výzkumníků. Mladší kolegové do práce přináší

nový elán a nápady a později mají převzít hlavní odpovědnost za udržení výzkumu dané populace. Letos jsme proto podali grantový návrh, v případě jehož financování bychom mohli studovat hnízdní biologii dutinových pěvců mnohem efektivněji než dříve. Již od roku 2015 totiž spolupracujeme s firmou EC-Elektronik, která pro nás vyvíjí a vyrábí RFID čtečky. Ty máme nyní umístěné na všech 95 budkách jedné plochy. Tato technologie nám zatím umožnila studovat chování mláďat okolo vyvádění (Tajovská 2019). V navrhovaném projektu počítáme s přidáním RFID čteček na budky dalších dvou přilehlých ploch a intenzivnějším čipováním dospělců i jejich mláďat a to jak u lejsků, tak i u obou druhů sýkor. Tento projekt by umožnil zodpovědět mnoho dalších zajímavých otázek ohledně rodičovské péče, výběru partnera a prostředí u dutinových pěvců. Návrh projektu počítá i s financováním doktoranda, který by se v ideálním případě později mohl stát hlavním garantem pokračování tohoto dlouhodobého výzkumu.

Obr. 17. RFID čtečky nabízejí nové možnosti studia hnízdní biologie dutinových pěvců.

Poděkování

Literatura

Adamík, P. & Král, M. (2008) Climate- and resource-driven long-term changes in dormice populations negatively affect hole-nesting songbirds. *Journal of Zoology,* **275,** 209-215.

Akcay, E. & Roughgarden, J. (2007) Extra-pair paternity in birds: review of the genetic benefits. *Evolutionary Ecology Research,* **9,** 855-868.

Akesson, S., Ilieva, M., Karagicheva, J., Rakhimberdiev, E., Tomotani, B. & Helm, B. (2017) Timing avian long-distance migration: from internal clock mechanisms to global flights. *Philosophical Transactions of the Royal Society B-Biological Sciences,* **372,** 20160252.

Bauchau, V. (1998) Comparison of parasitism level in two sympatric passerines: the pied flycatcher and the great tit. *Ecoscience,* **5,** 164-171.

Bazzi, G., Ambrosini, R., Caprioli, M., Costanzo, A., Liechti, F., Gatti, E., Gianfranceschi, L., Podofillini, S., Romano, A., Romano, M., Scandolara, C., Saino, N. & Rubolini, D. (2015) Clock gene polymorphism and scheduling of migration: a geolocator study of the barn swallow *Hirundo rustica*. *Scientific Reports,* **5**, 12443.

Blount, J.D., Houston, D.C. & Møller, A.P. (2000) Why egg yolk is yellow. *Trends in Ecology & Evolution,* **15,** 47-49.

Bonisoli-Alquati, A., Martinelli, R., Rubolini, D. & Saino, N. (2008) Sex-specific effects of albumen removal and nest environment manipulation on barn swallow nestlings. *Ecology,* **89,** 2315-2324.

Bonter, D.N. & Bridge, E.S. (2011) Applications of radio frequency identification (RFID) in ornithological research: a review. *Journal of Field Ornithology,* **82,** 1-10.

Both, C., Bouwhuis, S., Lessells, C.M. & Visser, M.E. (2006) Climate change and population declines in a long-distance migratory bird. *Nature,* **441,** 81-83.

Bowers, E.K., Sakaluk, S.K. & Thompson, C.F. (2011) Adaptive sex allocation in relation to hatching synchrony and offspring quality in house wrens. *American Naturalist,* **177,** 617-629.

Briedis, M., Hahn, S., Krist, M. & Adamík, P. (2018a) Finish with a sprint: Evidence for time-selected last leg of migration in a long-distance migratory songbird. *Ecology and Evolution,* **8,** 6899-6908.

Briedis, M., Krist, M., Král, M., Voigt, C.C. & Adamík, P. (2018b) Linking events throughout the annual cycle in a migratory bird: non-breeding period buffers accumulation of carry-over effects. *Behavioral Ecology and Sociobiology,* **72**, 93.

Cameron, E.Z. & Linklater, W.L. (2002) Sex bias in studies of sex bias: the value of daughters to mothers in poor condition. *Animal Behaviour*, **63**: F5-F8.

Carranza, J. 2002. What did Trivers and Willard really predict? *Animal Behaviour*, **63**: F3-F4.

Christians, J.K. (2002) Avian egg size: variation within species and inflexibility within individuals. *Biological Reviews,* **77,** 1-26.

Clutton-Brock, T. (1991) The evolution of parental care. *Princeton University Press*, New Jersey.

Clutton-Brock, T. & Sheldon, B.C. (2010) Individuals and populations: the role of long-term, individual-based studies of animals in ecology and evolutionary biology. *Trends in Ecology & Evolution,* **25,** 562-573.

Culina, A., Adriaensen, F., Bailey, L. D., Burgess, M. D., Charmantier, A., Cole, E. F., … Visser, M. E. (2020, August 11). Connected data landscape of long-term ecological studies: the SPI-Birds data hub. https://doi.org/10.32942/osf.io/6gea7

Cunningham, E.J.A. & Russell, A.F. (2000) Egg investment is influenced by male attractiveness in the mallard. *Nature,* **404,** 74-77.

Forstmeier, W., Nakagawa, S., Griffith, S.C. & Kempenaers, B. (2014) Female extra-pair mating: adaptation or genetic constraint? *Trends in Ecology & Evolution,* **29,** 456-464.

Fox, C.W. (1997) Egg-size manipulations in the seed beetle *Stator limbatus*: consequences for progeny growth. *Canadian Journal of Zoology,* **75,** 1465-1473.

Fox, C.W. & Czesak, M.E. (2000) Evolutionary ecology of progeny size in arthropods. *Annual Review of Entomology,* **45,** 341-369.

Frank, S.A. (1987) Individual and population sex allocation patterns. *Theoretical Population Biology,* **31,** 47-74.

Gil, D., Graves, J., Hazon, N. & Wells, A. (1999) Male attractiveness and differential testosterone investment in zebra finch eggs. *Science,* **286,** 126-128.

Griffith, S.C., Owens, I.P.F. & Thuman, K.A. (2002) Extra pair paternity in birds: a review of interspecific variation and adaptive function. *Molecular Ecology,* **11,** 2195-2212.

Groothuis, T.G.G., Muller, W., von Engelhardt, N., Carere, C. & Eising, C. (2005) Maternal hormones as a tool to adjust offspring phenotype in avian species. *Neuroscience and Biobehavioral Reviews,* **29,** 329-352.

Hegyi, G., Török, J. & Toth, L. (2002) Qualitative population divergence in proximate determination of a sexually selected trait in the collared flycatcher. *Journal of Evolutionary Biology,* **15,** 710-719.

Hodges, K.E., Mortimer, N.T., Vrailas-Mortimer, A.D., Sakaluk, S.K. & Thompson, C.F. (2020) Connecting the dots: avian eggshell pigmentation, female condition and paternal provisioning effort. *Biological Journal of the Linnean Society,* **130,** 114-127.

Hoenig, J.M. & Heisey, D.M. (2001) The abuse of power: The pervasive fallacy of power calculations for data analysis. *American Statistician,* **55,** 19-24.

Hunt, J., Bussiere, L.F., Jennions, M.D. & Brooks, R. (2004) What is genetic quality? *Trends in Ecology & Evolution,* **19,** 329-333.

Jetz, W., Thomas, G.H., Joy, J.B., Hartmann, K. & Mooers, A.O. (2012) The global diversity of birds in space and time. *Nature,* **491,** 444-448.

Johnsen, A., Andersen, V., Sunding, C. & Lifjeld, J.T. (2000) Female bluethroats enhance offspring immunocompetence through extra-pair copulations. *Nature,* **406,** 296-299.

Kempenaers, B. (2009) Behavioural ecology: cuckolder eggs come first. *Current Biology,* **19,** R364-R366.

Král, M. & Bičík, V. (1992) Nest defense by the collared flycatcher (*Ficedula albicollis*) against the great tit (parus-major). *Folia Zoologica,* **41,** 263-269.

Krist, M. (2002) Geographic variation in the role of secondary sexual traits in sexual competition: the case of collared flycatcher *Ficedula albicollis*. 5313 slov, nepublikovaný manuskript.

Krist, M. & Stříteský, J. (2008) Ovlivňuje výška dutiny její atraktivitu a bezpečnost pro sekundární dutinové hnízdiče? *Zprávy Vlastivědného muzea v Olomouci*, 293-295, 74-80.

Lack, D. (1968) Ecological adaptations for breeding in birds, *Chapman & Hall*, New York.

Lifjeld, J.T., Laskemoen, T., Kleven, O., Albrecht, T. & Robertson, R.J. (2010) Sperm length variation as a predictor of extrapair paternity in passerine birds. *Plos One,* **5,** e13456.

Magrath, M.J.L., Vedder, O., van der Velde, M. & Komdeur, J. (2009) Maternal effects contribute to the superior performance of extra-pair offspring. *Current Biology,* **19,** 792-797.

Magrath, R.D. (1992) The effect of egg mass on the growth and survival of blackbirds: a field experiment. *Journal of Zoology,* **227,** 639-653.

Moreno, J., Merino, S., Lobato, E., Ruiz-De-Castaneda, R., Martinez-De la Puente, J., Del Cerro, S. & Rivero-De Aguilar, J. (2009) Nest-dwelling ectoparasites of two sympatric hole-nesting passerines in relation to nest composition: an experimental study. *Ecoscience,* **16,** 418-427.

Moreno, J. & Osorno, J.L. (2003) Avian egg colour and sexual selection: does eggshell pigmentation reflect female condition and genetic quality? *Ecology Letters,* **6,** 803-806.

Moreno, J., Osorno, J.L., Morales, J., Merino, S. & Tomas, G. (2004) Egg colouration and male parental effort in the pied flycatcher *Ficedula hypoleuca*. *Journal of Avian Biology,* **35,** 300-304.

Parody-Merino, A.M., Battley, P.F., Conklin, J.R. & Fidler, A.E. (2019) No evidence for an association between Clock gene allelic variation and migration timing in a long-distance migratory shorebird (*Limosa lapponica baueri*). *Oecologia,* **191,** 843-859.

Qvarnström, A., Pärt, T. & Sheldon, B.C. (2000) Adaptive plasticity in mate preference linked to differences in reproductive effort. *Nature,* **405,** 344-347.

Ralston, J., Lorenc, L., Montes, M., DeLuca, W.V., Kirchman, J.J., Woodworth, B.K., Mackenzie, S.A., Newman, A., Cooke, H.A., Freeman, N.E., Sutton, A.O., Tauzer, L. & Norris, D.R. (2019) Length polymorphisms at two candidate genes explain variation of migratory behaviors in blackpoll warblers (*Setophaga striata*). *Ecology and Evolution,* **9,** 8840-8855.

Ratikainen, I.I., Haaland, T.R. & Wright, J. (2018) Differential allocation of parental investment and the trade-off between size and number of offspring. *Proceedings of the Royal Society B-Biological Sciences,* **285.**

Riehl, C. (2011) Paternal investment and the "sexually selected hypothesis" for the evolution of eggshell coloration: revisiting the assumptions. *Auk,* **128,** 175-179.

Riesch, R., Plath, M., Schlupp, I., Tobler, M. & Langerhans, R.B. (2014) Colonisation of toxic environments drives predictable life-history evolution in livebearing fishes (Poeciliidae). *Ecology Letters,* **17,** 65-71.

Rosivall, B., Szöllösi, E., Hasselquist, D. & Török, J. (2009) Effects of extrapair paternity and sex on nestling growth and condition in the collared flycatcher, *Ficedula albicollis*. *Animal Behaviour,* **77,** 611-617.

Santema, P., Valcu, M. & Kempenaers, B. (2020) Exposure to predator models during the fertile period leads to higher levels of extra-pair paternity in blue tits. *Journal of Animal Ecology,* **89,** 647-657.

Schlicht, L., Valcu, M. & Kempenaers, B. (2015) Male extraterritorial behavior predicts extrapair paternity pattern in blue tits, *Cyanistes caeruleus*. *Behavioral Ecology,* **26,** 1404-1413.

Sheldon, B.C. (2000) Differential allocation: tests, mechanisms and implications. *Trends in Ecology & Evolution,* **15,** 397-402.

Sheldon, B.C. & Ellegren, H. (1999) Sexual selection resulting from extrapair paternity in collared flycatchers. *Animal Behaviour,* **57,** 285-298.

Sheldon, B.C., Merilä, J., Qvarnström, A., Gustafsson, L. & Ellegren, H. (1997) Paternal genetic contribution to offspring condition predicted by size of male secondary sexual character. *Proceedings of the Royal Society B-Biological Sciences,* **264,** 297-302.

Speed, D. & Balding, D.J. (2015) Relatedness in the post-genomic era: is it still useful? *Nature Reviews Genetics,* **16,** 33-44.

Stoddard, M.C., Fayet, A.L., Kilner, R.M. & Hinde, C.A. (2012) Egg speckling patterns do not advertise offspring quality or influence male provisioning in great tits. *Plos One,* **7**, e40211.

Tajovská, M. (2019) Proces vyvádění u lejska bělokrkého. Přírodovědecká fakulta UP, diplomová práce.

Trivers, R.L. & Willard, D.E. (1973) Natural selection of parental ability to vary sex-ratio of offspring. *Science,* **179,** 90-92.

Underwood, T.J. & Sealy, S.G. (2002) Adaptive significance of egg coloration. In: Deeming, D.C. (ed) Avian incubation: behaviour, environment, and evolution. *Oxford University Press*, New York, pp 280–298.

Ward, M.P., Alessi, M., Benson, T.J. & Chiavacci, S.J. (2014) The active nightlife of diurnal birds: extraterritorial forays and nocturnal activity patterns. *Animal Behaviour,* **88,** 175-184.

Weidinger, K. (2001) Does egg colour affect predation rate on open passerine nests? *Behavioral Ecology and Sociobiology,* **49,** 456-464.

Weidinger, K. (2004) Relative effects of nest size and site on the risk of predation in open nesting passerines. *Journal of Avian Biology,* **35,** 515-523.

Weidinger, K. (2009) Nest predators of woodland open-nesting songbirds in central Europe. *Ibis,* **151,** 352-360.

Williams, T.D. (1994) Intraspecific variation in egg size and egg composition in birds - effects on offspring fitness. *Biological Reviews,* **69,** 35-59.

Přiložené články

Článek 1: Krist, M. (2004) Importance of competition for food and nest-sites in aggressive behaviour of Collared Flycatcher *Ficedula albicollis*. *Bird Study*, 51, 41-47.

Importance of competition for food and nest-sites in aggressive behaviour of Collared Flycatcher *Ficedula albicollis*

MILOS KRIST

Laboratory of Ornithology, Faculty of Science, Palacky University, Tr. Svobody 26, 771 46 Olomouc, Czech Republic and Museum of Natural History, Republiky 5, 771 73 Olomouc, Czech Republic

Capsule Using an experimental approach, this study disentangles effects of two important sources on the elicitation of aggressive interspecific behaviour.

Aims To investigate experimentally whether competition for nest cavities and/or food can explain aggressive behaviour between competing species.

Methods The Collared Flycatcher's *Ficedula albicollis* responses to mounts of Great Tit *Parus major* (nest-site and food competitor), Chaffinch *Fringilla coelebs* (food competitor) and Dunnock *Prunella modularis* (control species) were tested. Trials were performed near flycatchers' nests during nest building, incubation and care of nestlings.

Results The intensity of dive attacks and frequency of contact attacks showed that the aggression of Collared Flycatchers decreased in the direction: Great Tit > Chaffinch > Dunnock. The difference in aggressiveness was greater between Great Tit and Chaffinch than between Chaffinch and Dunnock. Aggression directed toward Great Tits increased from the nest building to the incubation stage and then decreased markedly in the nestling stage. Males were more aggressive than females.

Conclusion These results suggest that competition for nest-sites, and to a lesser extent for food, may be of an interference nature and that the reproductive value hypothesis can only partly explain differences in the Collared Flycatcher's defensive behaviour found between sexes and stages of the breeding cycle.

Food and nest-sites are the most important resources in interspecific competition in birds (Newton 1998). Interspecific competition for food is mostly considered to be exploitation (i.e. competition without direct behavioural interactions between competing individuals) whereas that for nest-sites is considered to be interference (i.e. competition with direct behavioural interactions between competing individuals). However, departures from this rule are quite common, at least in the case of food competition. Substantial aggression directed toward heterospecifics competing for food can eventually lead to interspecific territoriality (Orians & Willson 1964) which is very clear evidence of interference. However, interspecific territoriality is costly and so preferable only in specific conditions. Among those is life in a structurally simple habitat (Reed 1982) or exploiting stratified food sources (Dearborn 1998), both of which preclude niche differ-

entiation. Newton (1998) listed 27 studies in which at least 32 species pairs maintained interspecific territories.

In the case of nest-sites, there is little doubt about the nature of competition. However, there is no support from experimental research that is directly suggestive of interference. So far, most studies have explored changes in local densities of secondary cavity-nesting birds after manipulation of nest-site availability – usually after nest boxes were provided. Newton (1998) listed 34 such studies. In all of them local densities of secondary cavity-nesting birds responded in the expected direction to manipulation and the same conclusion can be drawn – cavities are the limiting resource and thus subject to intensive intraspecific and interspecific competition. The latter was documented by complementary changes in the breeding densities of two or more competing species (Minot & Perrins 1986, Gustafsson 1988).

Although such correlations can suggest the existence of competition, they cannot reveal its nature. To inves-

Email: krist@prfnw.upol.cz

tigate whether interference or exploitation is acting it is necessary to pay attention to a proximate level of interspecific relationships, i.e. to document aggressive behaviour in at least one species involved. So far the evidence of such aggression is mainly observational (Slagsvold 1975, Gustafsson 1988, Winge & Järvi 1988, Meek & Robertson 1994, Merilä & Wiggins 1995). Experimental studies are few in this area. Järvi *et al.* (1978) tested mutual aggressive responses of Pied Flycatchers *Ficedula hypoleuca* and Great Tits *Parus major* to playback of the other species' song. Král & Bicík (1992) investigated behaviour of Collared Flycatchers *Ficedula albicollis* towards a dummy of the Great Tit placed on their nest boxes. Both studies concluded that the observed aggression was selected due to nest-site competition. However, there was no appropriate control treatment in either of these studies. *Ficedula* flycatchers compete strongly with tits not only for nest-sites but also for food (Slagsvold 1975, Gustafsson 1987, Sasvári *et al.* 1987). Thus, trials without controls for the possible confounding effect of food competition cannot adequately explain the function of observed interspecific aggression. Similarly, Martin & Martin (2001) revealed mutual aggression between two warbler species which compete for food as well as for nest-sites but their study also was not specifically designed to conclude which source caused this aggression. I know of no other study investigating the role of nest-site competition in aggressive behaviour of birds.

The main aim of the present study was to investigate whether competition for nest cavities and/or food can explain aggressive behaviour between competing species. To test these ideas, the aggressive responses of Collared Flycatchers to the Great Tit (nest-site as well as food competitor), the Chaffinch *Fringilla coelebs*, (food competitor) and the Dunnock *Prunella modularis*, (neither food nor nest-site competitor) were compared. A prediction of decreasing aggression in the direction Great Tit > Chaffinch > Dunnock was made. In addition, changes in aggression of Collared Flycatchers during the breeding cycle and differences between the sexes were examined. Aggression directed toward nest-site competitors should increase during the breeding cycle, because of the increasing reproductive value of the brood. Changes in aggression caused by competition for food are more difficult to predict because both food availability and food demands presumably increase during the breeding cycle. Thus, it is not clear in which stage of the breeding cycle food competition should be most severe.

METHODS

Study area and subjects

The study was conducted in Velky Kosir area (49°32′N, 17°04′E, 370–450 m asl), central Moravia, Czech Republic, during 1997–2000. Experiments were conducted on two nest-box plots. Both are about 12 ha in area, each provided with approximately 60 nest boxes. The vegetation of these plots consisted of structurally very simple managed spruce *Picea abies* forest. The spruces were about 90 years old and 25 m tall. Other species of tree were very scarce, represented mainly by pine *Pinus sylvestris* and birch *Betula pendula*. The shrub understorey was not developed. The herb level, consisting mainly of *Calamagrostis villosa*, covered only about 40% of the ground.

The Collared Flycatcher is a small (about 13 g), cavity-nesting, migrant passerine species with sexually dimorphic plumage. It easy adopts nest boxes for breeding (about 40 Collared Flycatcher pairs nested each year on the study plots). Collared Flycatchers forage mainly in the canopy, less in the shrubs or on the ground. They obtain food by sallying out from a perch after flying prey or picking up directly from leaves and twigs (Cramp & Perrins 1993). Diet brought to nestlings consists mainly of Lepidoptera larvae, Diptera, Hymenoptera and Araneida (Bures 1986, Cramp & Perrins 1993). The Chaffinch (an open-nesting species) and Great Tit (a cavity-nesting species) forage in a way similar to that of the Collared Flycatcher and also the food brought to their young overlaps broadly with that of the Collared Flycatcher (Cramp & Perrins 1993, 1994). Moreover, food competition has been directly proven between the Collared Flycatcher and the Great Tit (Gustafsson 1987, Sasvári *et al.* 1987) and between the Great Tit and the Chaffinch (Reed 1982). The Dunnock (an open-nesting species), on the other hand, differs from all three species mentioned above largely in its foraging techniques and the food brought to young. This species is mainly a ground forager and its diet consist mainly of harvestmen (Opilionidea), colembolan (Entomobryidae), Auchenorrhyncha and Lepidoptera larvae occurring near the ground (Geometridae) (Cramp 1988, Kristín 1989). All three bird species chosen for experiments are of similar size (Great Tit 20 g, Chaffinch 23 g and Dunnock 21 g) and were present on the study plots. Stuffed specimens of two males of the Great Tit, two males of the Chaffinch and a Dunnock (sexually monomorphic species) of unknown sex were used for experiments. All specimens were stuffed in a similar posture.

Experimental design

Experiments were conducted in three stages of the breeding cycle: nest building, incubation and care for nestlings. The dummy was placed either directly on the nest box, or on the wooden stick about 50 cm high placed at a distance of 2 m from the nest-box. Experiments were conducted on 94 nests. Each of them was tested with all three dummies, placed at a single distance (nest box or 2 m), on the same day. The order of presentation was random. Successive presentations were separated by an interval at least 45 min (75 min in the case of the former contact attack).

After the dummy had been installed, focal nest observations were made at a distance about 25 m. After the male or female had approached to a close vicinity of the nest and so presumably noticed the dummy, the observational interval began. It lasted 10 min for each focal bird. For each individual the number of dive attacks against the dummy and the latency to first dive attack were recorded. In cases when an individual did not attack the dummy at all, the latency was stated as 600 s (i.e. the length of observational interval). This method should hold type I errors below the stated level ($\alpha = 0.05$) and results may be a bit conservative in this respect.

Sometimes physical contact was made with the dummy when it was attacked. In this case the trial was stopped and the dummy removed to prevent it from destruction. Such trials were classified as 'contact attack occurred' and that was the only information involved in statistical analysis, because of the artificially shortened observational interval.

Statistical analysis

Two response variables (dive attacks and contact attacks) were used for the statistical analyses. Dive attacks was the first principal component from the principal component analysis (PCA) performed on two original variables (number of dives and the latency to first dive attack) which were highly correlated. Before PCA was performed, variables were ln- (number of attacks) or box-cox- (latency) transformed to achieve better approximation to the required normal distribution. The composite variable has eigenvalue 1.84 (92% variance explained) and the following factor loadings: number of dive attacks (0.71), and latency to first dive attack (–0.71). The second response variable was binomial (contact attack occurred yes/no). The two response variables were analysed separately. Because

the two models test the same hypothesis, a sequential Bonferroni adjustment (Rice 1989) was applied to reduce the type I error rate.

In both models the following factors with their two-way interactions were included as fixed effects: species, sex, breeding stage and distance. Since several trials were performed on individual nests, nest was treated as a random factor in both models. To select the best model, a two-step method was used. First, in the fixed part of the model, non-significant interactions were eliminated by backwards selection. Second, the random part of the model (covariance structure) was selected according to Akaike's information criterion (AIC). In the case of dive attacks, the best model contained random slopes (within nests) for sex and distance. In the case of contact attacks, inclusion of random factors caused an increase in AIC and therefore the final model contained only the fixed part. Denominator degrees of freedom were computed using Satterthwaite's method (Littell *et al.* 1996). Dive attacks were analysed using PROC MIXED, contact attacks using the GLIMMIX macro for SAS. Whenever a factor with more than two levels was significant, Tukey's HSD tests were computed. All tests were computed in SAS (Littell *et al.* 1996).

RESULTS

As judged by dive attacks, both male and female Collared Flycatchers behaved most aggressively against the Great Tit, less aggression was directed toward Chaffinch and the Dunnock was the least attacked dummy (Fig. 1 & Table 1; Tukey's HSD tests: Great Tit
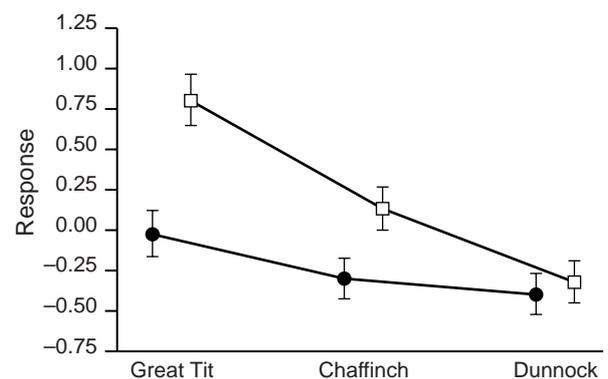
Figure 1. Least squares means ± se for the response of male (□) and female (●) Collared Flycatcher to the three species studied. The response is the first principal component comprising the number of dive attacks and latency to the first dive attack. A higher value represents more dive attacks and a shorter latency to the first dive attack.

Table 1. Type III tests of fixed effects. Response was the composite variable derived from a principal component analysis on latency to first dive attack and number of dive attacks. Backwards elimination procedure; only interactions significant after applying a Bonferroni correction were retained in the presented model.

	NDF	DDF	*F*	*P*
Species	2	621.0	22.8	< 0.001
Sex	1	94.1	13.4	< 0.001
Breeding stage	2	671.0	2.7	0.069
Distance	1	69.6	8.3	0.005
Species∗sex	2	615.0	5.9	0.003

NDF, numerator degrees of freedom; DDF, denominator degrees of freedom.

vs. Chaffinch, $P < 0.001$; Great Tit vs. Dunnock, $P < 0.001$; Chaffinch vs. Dunnock, $P = 0.005$). The occurrence of the most aggressive behaviour (contact attacks) supported this pattern (Fig. 2 & Table 2). The Great Tit was contacted significantly more frequently than either Chaffinch ($P < 0.001$) or Dunnock ($P < 0.001$). There was no significant difference between Chaffinch and Dunnock in this respect ($P = 0.19$).

Males were the more aggressive sex. They contacted dummies more often (Table 2) and performed stronger dive attacks (Table 1). In the latter case this difference was highest when the Great Tit was presented, less when the Chaffinch presented and virtually disappeared when the Dunnock presented (Fig. 1). Incubation was the stage when the probability of attacking the dummy physically was highest (Table 2; Tukey's HSD tests: incubation vs. nestling, $P < 0.001$; incubation vs. nest building, $P = 0.020$; nest building vs. nestling $P = 0.55$). Similarly to contacts, dive attacks were strongest at the incubation stage too. However, because the result of the mixed model was
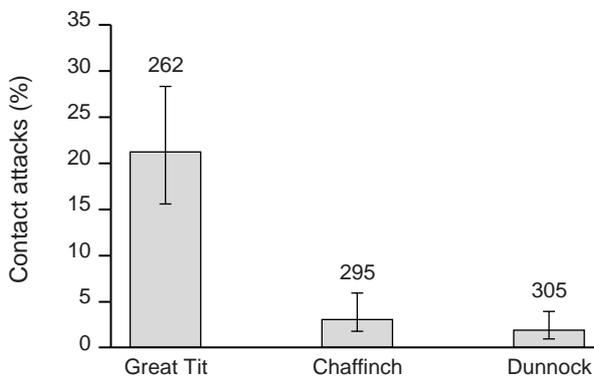
Figure 2. Least squares means ± 95% confidence limits of occurrence of contact attacks against the three species studied. Values are adjusted for sex, breeding stage and distance. The sample size for each species is indicated above the bar.

Table 2. Results of GLIMMIX model performed on contact attacks. Probability of a contact attack occurring was modelled. Backwards elimination procedure; all interactions were non-significant and therefore deleted. Dunnock, male, nestling stage, and the distance of 2 m are the reference levels for the factors species, sex, breeding stage, and distance, respectively.

a. Type III analysis of effects

	DDF	*F*	*P*
Species	855	41.0	< 0.001
Sex	855	26.1	< 0.001
Breeding stage	855	7.6	< 0.001
Distance	855	16.3	< 0.001

b. Analysis of REML estimates

Factor	Level	Parameter estimate	se
Intercept		−4.32	0.437
Species	Great Tit	2.65	0.367
	Chaffinch	0.55	0.426
Sex	Female	−1.30	0.255
Breeding stage	Nest building	0.20	0.336
	Incubation	1.09	0.280
Distance	Nest box	1.16	0.287

DDF, denominator degrees of freedom; REML, restricted maximum likelihood.

marginally non-significant (Table 1), multiple comparisons were not computed.

DISCUSSION

Food and nest-site competition

In research on interspecific competition it is commonly assumed that competition for nest-sites is of an interference nature. However, hitherto there has been no support from experimental research for this idea. This study revealed that dive attacks decreased in the direction Great Tit > Chaffinch > Dunnock as predicted under the hypothesis that nest-site competition as well as food competition contribute to aggressive behaviour of Collared Flycatcher. The difference in the intensity of dive attacks was greater between the Great Tit and the Chaffinch than between the Chaffinch and the Dunnock. The same was true for the occurrence of contact. These results imply that competition for nest-sites contributes more to aggressive behaviour of Collared Flycatchers than competition for food.

Competition for nest cavities has been considered to be low in primeval forests in the past because of high availability of cavities. Further it has been considered that at present we can observe competition for cavities

in nature only due to decrease in hole numbers caused by intensive silvicultural management (Walankiewicz 1991). However, the life-history traits of some North American (Martin 1993a) as well as European (Mönkkönen & Martin 2000) secondary cavity-breeding birds are best explained by the limited breeding opportunities hypothesis (Martin 1993b). This fact suggests that competition for tree cavities played an important role in the evolutionary past of these species. Nest defence behaviour is possibly less apt to reflect the situation hundreds of years ago, because of its potential for rapid evolution through learning and cultural transmission (cf. Maloney & McLean 1995). On the other hand, studying this behaviour is the only way to reveal underlying mechanisms involved in competition. Aggressive behaviour directed specifically toward the nest competitor in this experimental study, although it cannot be treated as evidence of long-acting selection pressure, proves the interference nature of competition for nest cavities.

Ficedula flycatchers are not usually assumed to defend large feeding territories (Cramp & Perrins 1993). However, in contrast to this traditional view, older males of the Pied Flycatcher were found to possess territories with a greater abundance of invertebrate food than subordinate yearlings (Huhta *et al.* 1998). This suggests that the abundance of food can influence intraspecific and so possibly also interspecific behaviour. The latter was proved in this study on the Collared Flycatcher when the food competitor was attacked more strongly than the non-competing species. However, the fact that the differences in aggression were greater between the Great Tit and Chaffinch than between Chaffinch and Dunnock indicates that food competition plays only a minor role in aggressive interactions between tits and flycatchers. Moreover, the effect of food on aggression may be specific for a simple habitat such as that on my study plots. As Orians & Willson (1964) stated, a simple habitat can prevent niche differentiation, consequently competition for food is more pronounced and can lead to interspecific territoriality among competing bird species. Reed (1982), for example, found Great Tit to be interspecifically territorial with Chaffinch on Scottish islands whereas such a pattern did not emerge in the more structured habitats on the mainland.

Sex differences

Males were the more aggressive sex in this study. They contacted dummies more frequently and performed stronger dive attacks compared to females. In the case of dive attacks, the difference between sexes was greatest with the Great Tit, less with the Chaffinch and almost non-existent with the Dunnock.

The role of sex in nest defence against nest competitors has been investigated just once (Král & Bicík 1992). Unlike this study, Král & Bicík (1992) did not find any differences in the nest defence between sexes. These contrasting results can be explained by a different assessment of the intensity of defence. Král & Bicík (1992) combined records of watchfulness and aggressiveness whereas here only aggression was recorded, making the interpretation more straightforward. Overall, higher aggression of males against food competitors (Great Tit and Chaffinch) is not surprising. As in most other birds, males in Collared Flycatcher are the sex predominantly engaged in territory defence (Cramp & Perrins 1993). The finding that the largest difference in aggression between sexes was seen with the Great Tit can be explained in terms of residual reproductive value, because the Great Tit, in contrast to Chaffinch and Dunnock, can destroy a flycatcher's nest. Greater intrasexual competition in males, as suggested by a small proportion of yearling males in the breeding population in contrast to a great one in females (Král & Pithart 1995, M. Krist, unpubl. data), results in their lower renesting potential and residual reproductive value. Consequently, males should defend the brood more aggressively (Montgomerie & Weatherhead 1988). In addition, if the territory defence and residual reproductive value alone are responsible for the sex differences in aggressiveness, as suspected, then aggression toward a non-competing bird (Dunnock) should be low for both males and females, as was found.

Breeding cycle

The frequency of contact attacks decreased in the direction incubation > nest building > nestling stage. Due to the very low frequency of contacts against the Chaffinch and the Dunnock (Fig. 2) this pattern is driven by flycatchers' responses toward the Great Tit. Parental investment is predicted to increase in the course of the breeding cycle due to an increase in both the reproductive value of the current brood (Andersson *et al.* 1980) and a feedback stimulus of the nest contents to a parent (McLean & Rhodes 1992). A large number of studies of nest defence against predators have confirmed this prediction (Montgomerie & Weatherhead 1988). In contrast to the large number of

studies investigating breeding-cycle changes of nest defence against predators, just two such studies were done with nest competitors, although nest competitors might pose the same threat for the brood as nest predators. In the first, Winge & Järvi (1988) found no evidence that past parental investment influenced the success of nest defence in Great Tits. However, all their observations were performed during the nest building stage, where the amount of variation in individuals' residual reproductive values was probably too small to answer this question adequately. In the second, the nest defence of Collared Flycatchers against the Great Tit was observed to be stronger later in the breeding cycle (Král & Bicík 1992). However, their results may reflect watchfulness rather than aggression, the former surely being very low at the egg-laying stage.

The pattern of contact attack variation in this study provided only partial support for a parental investment/ feedback hypothesis. Aggression against nest competitors increased from the nest building to incubation stage as predicted but then decreased markedly. This decrease can probably be explained as follows. Although both nest competitors and nest predators pose theoretically the same danger for the brood, they probably behave in a different way. Nestlings are presumably more valuable prey for nest predators than eggs, because of their greater energy content. For nest competitors, on the other hand, the presence of young can complicate the take-over of a nest hole for two reasons. First, it may be difficult for them to build their own nest on that with living young because of their activity. Second, if young are much more valuable for parents than eggs, increased aggression from nest hole owners will deter an intruder to try to take over such a hole. The latter also concerns nest predators but the difference is in the duration of experiencing such aggression. Predation can happen very quickly, whereas take-over is necessarily a lengthy act and thus costs of suffering aggression from hole owners would be probably much greater. Such conditions could favour an increase in defence against nest predators occurring near the nest with nestlings, as has been repeatedly confirmed (Montgomerie & Weatherhead 1988), since the presence of a predator near such a nest means a high risk of nest predation. However, the presence of a nest-site competitor near a nest containing nestlings is, in contrast to a nest containing eggs, probably only a random event and does not mean that it is trying to take over the hole. This hypothesis explains why the take-over of cavities occurs just before incubation (Slagsvold 1975, Meek & Robertson 1994), while

aggression directed toward the nest competitor occurring in the vicinity of nest was found to be low during the nestling stage.

REFERENCES

Andersson, M., Wiklund, C.G. & Rundgren, H. 1980. Parental defence of offspring: a model and an example. *Anim. Behav.* **28:** 536–542.

Bures, S. 1986. Composition of the diet and trophic ecology of the collared flycatcher (*Ficedula albicollis albicollis*) in three segments of groups of types of forest geobiocenoses in central Moravia (Czechoslovakia). *Folia Zool.* **35:** 143–155.

Cramp, S. (ed.) 1988. *The Birds of the Western Palearctic,* Vol. 5. Oxford University Press, Oxford.

Cramp, S. & Perrins, S.M. (eds) 1993. *The Birds of the Western Palearctic,* Vol. 7. Oxford University Press, Oxford.

Cramp, S. & Perrins, S.M. (eds) 1994. *The Birds of the Western Palearctic,* Vol. 8. Oxford University Press, Oxford.

Dearborn, D.C. 1998. Interspecific territoriality by a rufous-tailed hummingbird (*Amazilia tzacatl*): effects of intruder size and resource value. *Biotropica* **30:** 306–313.

Gustafsson, L. 1987. Interspecific competition lowers fitness in collared flycatchers *Ficedula albicollis*: an experimental demonstration. *Ecology* **68:** 291–296.

Gustafsson, L. 1988. Inter- and intraspecific competition for nest holes in a population of the collared flycatcher *Ficedula albicollis. Ibis* **130:** 11–16.

Huhta, E., Jokimäki, J. & Rahko, P. 1998. Distribution and reproductive success of the Pied Flycatcher *Ficedula hypoleuca* in relation to forest patch size and vegetation characteristics; the effect of scale. *Ibis* **140:** 214–222.

Järvi, T., Radesäter, T. & Jakobsson, S. 1978. Aggressive responses of two hole-nesting passerines, *Parus major* and *Ficedula hypoleuca*, to the play-back of sympatric species song. *Ornis Fenn.* **55:** 154–157.

Král, M. & Bicík, V. 1992. Nest defence by the collared flycatcher (*Ficedula albicollis*) against the great tit (*Parus major*). *Folia Zool.* **41:** 263–269.

Král, M. & Pithart, K. 1995. [Stárí lejska belokrkého (*Ficedula albicollis*).] *Sylvia* **31:** 69–74 (in Czech).

Kristín, A. 1989. [Ernährung der Nestlinge der syntopischen Arten Zilzalp (*Phylloscopus collybita*) und Heckenbraunelle (*Prunella modularis*).] *Folia Zool.* **38:** 349–362 (in German).

Littell, R.C., Milliken, G.A., Stroup, W.W., & Wolfinger, R.D.

1996. *SAS System for Mixed Models.* SAS Institute Inc, Cary: NC.

Maloney, R.F. & McLean, I.G. 1995. Historical and experimental learned predator recognition in free-living New Zealand robins. *Anim. Behav.* **50:** 1193–1201.

Martin, T.E. 1993a. Evolutionary determinants of clutch size in cavity-nesting birds: nest predation or limited breeding opportunities? *Am. Nat.* **142:** 937–946.

Martin, T.E. 1993b. Nest predation and nest-sites: new perspectives on old patterns. *BioScience* **43:** 523–532.

Martin, P.R. & Martin, T.E. 2001. Behavioral interactions between coexisting species: song playback experiments with wood warblers. *Ecology* **82:** 207–218.

McLean, I.G. & Rhodes, G. 1992. Enemy recognition and response in birds. In: Power, D.M. (ed.) *Current Ornithology*, Vol. 8: 173–212. Plenum Press, New York.

Meek, S.B. & Robertson, R.J. 1994. Interspecific competition for nestboxes affects mate guarding in eastern bluebirds, *Sialia sialis. Anim. Behav.* **47:** 295–302.

Merilä, J. & Wiggins, D.A. 1995. Interspecific competition for nest holes causes adult mortality in the collared flycatcher. *Condor* **97:** 445–450.

Minot, E.O. & Perrins, C.M. 1986. Interspecific interference competition: nest-sites for blue and great tits. *J. Anim. Ecol.* **55:** 331–350.

Mönkkönen, M. & Martin, T.E. 2000. Sensitivity of comparative analyses to population variation in trait values: clutch size and cavity excavation tendencies. *J. Avian Biol.* **31:** 576–579.

Montgomerie, R.D. & Weatherhead, P.J. 1988. Risks and rewards of nest defence by parent birds. *Q. Rev. Biol.* **63:** 167–187.

Newton, I. 1998. *Population Limitation in Birds.* Academic Press, London.

Orians, G.H. & Willson, M.F. 1964. Interspecific territories of birds. *Ecology* **45:** 736–745.

Reed, T.M. 1982. Interspecific territoriality in the chaffinch and great tit on islands and the mainland of Scotland: playback and removal experiments. *Anim. Behav.* **30:** 171–181

Rice, W.R. 1989. Analyzing tables of statistical tests. *Evolution* **43:** 223–225.

Sasvári, L., Török, J. & Tóth, L. 1987. Density dependent effects between three competitive bird species. *Oecologia* **72:** 127–130.

Slagsvold, T. 1975. Competition between the great tit *Parus major* and the pied flycatcher *Ficedula hypoleuca* in the breeding season. *Ornis Scand.* **6:** 179–190.

Walankiewicz, W. 1991. Do secondary cavity-nesting birds suffer more from competition for cavities or from predation in a primeval deciduous forest? *Nat. Areas J.* **11:** 203–212.

Winge, K. & Järvi, T. 1988. Nest-hole defence by the great tit against pied flycatcher intrusion: a test of the 'parental investment' and the 'fighting ability' hypotheses. *Oikos* **51:** 364–366.

Článek 2: Krist, M., Remeš, V., Uvírová, L., Nádvorník, P. & Bureš, S. (2004) Egg size and offspring performance in the collared flycatcher (*Ficedula albicollis*): a within-clutch approach. *Oecologia*, 140, 52-60.

POPULATION ECOLOGY

**Miloš Krist · Vladimír Remeš · Lenka Uvírová ·
Petr Nádvorník · Stanislav Bureš**

Egg size and offspring performance in the collared flycatcher (*Ficedula albicollis*): a within-clutch approach

Abstract Adaptive within-clutch allocation of resources
by laying females is an important focus of evolutionary
studies. However, the critical assumption of these studies,
namely that within-clutch egg-size deviations affect off-
spring performance, has been properly tested only rarely.
In this study, we investigated effects of within-clutch
deviations in egg size on nestling survival, weight,
fledgling condition, structural size and offspring recruit-
ment to the breeding population in the collared flycatcher
(*Ficedula albicollis*). Besides egg-size effects, we also
followed effects of hatching asynchrony, laying sequence,
offspring sex and paternity. There was no influence of egg
size on nestling survival, tarsus length, condition or
recruitment. Initially significant effect on nestling mass
disappeared as nestlings approached fledging. Thus, there
seems to be limited potential for a laying female to exploit
within-clutch egg-size variation adaptively in the collared
flycatcher, which agrees with the majority of earlier
studies on other bird species. Instead, we suggest that
within-clutch egg-size variation originates from the effects
of proximate constraints on laying females. If true,
adaptive explanations for within-clutch patterns in egg
size should be invoked with caution.

Keywords Cross-fostering · Intraclutch · Maternal
effects · Nestling growth · Offspring fitness

M. Krist (✉)
Museum of Natural History,
nám. Republiky 5,
771 73 Olomouc, Czech Republic
e-mail: krist@prfnw.upol.cz
Tel.: +420-585515128
Fax: +420-585222743

M. Krist · V. Remeš · S. Bureš
Laboratory of Ornithology, Palacký University,
tř. Svobody 26,
771 46 Olomouc, Czech Republic

L. Uvírová · P. Nádvorník
Department of Cell Biology and Genetics, Palacký University,
Šlechtitelů 11,
783 71 Olomouc, Czech Republic

P. Nádvorník · S. Bureš
Department of Zoology, Palacký University,
tř. Svobody 26,
771 46 Olomouc, Czech Republic

Introduction

Within-clutch allocation of resources by a laying female is
an important topic in evolutionary ecology. In addition to
studies examining the allocation of resources in relation to
laying order (e.g. O'Connor 1979; Slagsvold et al. 1984;
Wiggins 1990; Williams et al. 1993a; Cichoń 1997;
Viñuela 1997; Hillström 1999), increasing attention is
being paid to possible adaptive allocation in relation to egg
sex (Weatherhead 1985; Leblanc 1987a; Mead et al. 1987;
Teather 1989; Andersson et al. 1997; Cordero et al. 2000,
2001; Rutkowska and Cichoń 2002; Blanco et al. 2003;
Cichoń et al. 2003; Magrath et al. 2003). By allocating
resources differentially in relation to laying order, females
may enhance/impair survival of the later hatching chicks
(Slagsvold et al. 1984) or favour chicks with the highest
reproductive value (Williams et al. 1993a). By targeting
resources to eggs of a particular sex, the female may also
obtain two types of benefits. First, in sexually dimorphic
species, she may boost performance of the smaller sex to
prevent it from starvation due to competition for food with
the larger sib of the other sex (Anderson et al. 1997).
Second, when in good condition she may increase her
fitness by investing selectively resources to the sex with
larger variance in reproductive success (Trivers and
Willard 1973).

The critical assumption of the adaptive allocation of
resources within a given clutch is that the amount of the
invested resources has consequences for offspring fitness.
For example, it is usually assumed that the larger the egg,
the higher the fitness of the offspring that hatches from this
egg. This assumption has been most often tested by the
cross-fostering approach when eggs/nestlings are swapped

between nests and performance of offspring in relation to mean egg size of the clutch is analysed (we know of 16 such studies; e.g. Bize et al. 2002; Pelayo and Clark 2003). However, to test the assumption of the within-clutch adaptive allocation specifically, it is better to examine effects of within-clutch deviations in egg size on the performance of individual offspring. First, variation in egg size is typically much greater between females than within clutches of individual females (Christians 2002). Thus, cross-fostering studies work with the egg-size variability that is most probably not available to laying females when allocating resources within a clutch. Second, direct competition between sibs for resources supplied by parents, monopolisation of these resources by dominant sibs and selective parental feeding in relation to offspring size are common within broods (Budden and Wright 2001). Cross-fostering studies working on the between-female level do not take into account these within-family relations and thus may not reliably estimate egg-size effects present on the within-brood level (see also Nilsson and Svensson 1993). Third, positive covariation between direct and maternal pathways of the determination of offspring phenotype may exist, which would lead to overestimation of egg-size effects in cross-fostering design (Krist and Remeš 2004).

Studies testing effects of within-clutch deviations on offspring performance have been done less frequently than cross-fostering studies (we know of nine studies; e.g. Howe 1976; Amat et al. 2001). They often compared average egg size of surviving versus non-surviving siblings instead of looking at individual offspring. Moreover, they did not, for the most part, control for the factors that are known to affect offspring performance. First, all nest-mates may be affected in the same way by brood-level factors (between-year variation in the quality of breeding conditions, advancement of breeding season, territory quality). These factors may be included in the analyses of egg-size effects on offspring performance to reduce unexplained variation and thus increase statistical power of the main test. Second, nest-mates may differ in performance due to hatching asynchrony (Magrath 1990), laying order (Ylimaunu and Järvinen 1987), sex (Becker and Wink 2003), paternity (Sheldon et al. 1997) or concentration of androgens in eggs (Schwabl 1993). These individual-level factors may be, in contrast to brood-level factors, correlated with within-clutch differences in egg size and as such directly confound any relationship between the latter and offspring performance.

To test the assumption of adaptive within-clutch allocation of resources, we examined effects of within-clutch deviations in egg size on individual offspring performance in the collared flycatcher (*Ficedula albicollis*), a small migratory passerine. Besides egg-size effects, we also followed effects of other individual-level factors including hatching asynchrony, laying order, offspring sex and paternity, which makes our study well suited for separating an independent effect of egg size on offspring performance. We examined effects of these factors on nestling survival, weight, fledgling condition, structural size and offspring recruitment to the breeding population. In addition, we also controlled statistically for some brood-level covariates to render the analyses of egg-size effects more powerful.

Materials and methods

Field methods

The study was conducted in Velký Kosíř forest (49°32′N, 17°04′E, 370–450 m a.s.l.), central Moravia, the Czech Republic, in 2001–2003. In the study area, there were five plots with the total number of about 350 nest-boxes. Three plots were located in coniferous (*Picea abies*) and the other two in deciduous (*Quercus petrea*) forest. Approximately 60 pairs of collared flycatchers bred in the nest-boxes each year.

The study area was visited daily during the breeding season. Each egg was numbered with a waterproof felt-pen and measured to the nearest 0.01 mm with a digital calliper on the day it was laid. Egg volume was calculated using the formula: volume=$0.51 \times$length$\times$width^2 (Hoyt 1979). Two measures of width were taken in two perpendicular directions and their average was used as a measure of width. After 10–13 days of initiation of incubation, eggs were taken from nests, put into a thermo-box and then within 10 min of transfer placed into individual compartments in an incubator. Plastic dummy eggs were put into the nests for females to incubate. The method was successful since only one out of 38 artificial clutches was abandoned. Temperature in the incubator ranged between 37 and 39°C, humidity between 40 and 70%. The incubator was checked for newly hatched young at least every 3 h throughout the day and night. Hatching time was recorded for each chick. When hatching was not directly observed, hatching time was approximated as the midpoint between the check when the egg was hatched and the preceding check when the egg was still unhatched. As soon as possible, hatchlings were returned to their nest of origin. The mean time (±SD) which elapsed between hatching and the return of the hatchling to the nest was 2.95±2.33 h (range 10 min–10 h). The longer time periods occurred when the young hatched in the evening and starved until sunrise, which is also the case under natural conditions. To ensure that the delay did not affect our results, we included the time elapsed between hatching and returning the hatchling to the nest ("time to return" hereafter) as a covariate into our models (see below). Before their return, the claws of hatchlings were marked by nail-varnish to enable individual recognition. Nestlings were checked daily until they were 13 days old, i.e. close to fledging. Every day, nestlings were weighed to the nearest 0.25 g with a Pesola spring balance and re-marked if needed. Nestlings were ringed when about 7 days old, blood sampled (about 25 µl) by brachial venipuncture at 10–13 days and their tarsi were measured (to the nearest 0.01 mm) at 13 days. Blood samples were transferred to 1 ml of Queen's lysis buffer (Seutin et al. 1991). Dead nestlings were taken from nests and conserved in 70% ethanol. Putative parents were caught with nest-box traps while feeding nestlings and their blood was sampled in the same way as for nestlings.

Each year nearly all adults breeding in the study area were captured and checked for rings. Thus for the young fledglings in 2001, 2 years of potential recapture as breeding adults were available, but only 1 year for the young fledgling in 2002. It is certain that some individuals that had ultimately recruited to the breeding population were not discovered by us and thus we erroneously treated them as non-recruits. However, under the assumption that the dispersal and the probability of starting breeding in the second year of life are not biased with respect to egg size, our subsample of the recruits is representative. The first part of the assumption seems to be realistic as breeding dispersal is unbiased with respect to other offspring traits such as fledgling weight or tarsus length in this species (Pärt 1990). Concerning the second part of the assumption, it is possible that superior individuals already

start breeding in the second year of life while individuals in bad condition are "floaters" at this time but are recruited a year later, in their third year of life. This would lead to over-representation of individuals in good condition in our sample of recruits and thus overestimation of egg-size effects on recruitment (to the extent that egg size positively affects condition and probability of early breeding). However, this possibility makes our conclusions even more conservative (see below).

Sex and paternity

Nestling sex and parentage were determined using standard methods for the collared flycatcher (Sheldon and Ellegren 1996). In short, DNA was extracted from blood or tissue samples using the phenol-chloroform method. Sex was determined by polymerase chain reaction amplification of the CHD gene using primers P2 and P8 (Griffiths et al. 1998), followed by polyacrylamid electrophoresis. The method was completely accurate: sex of about 60 adults of known sex was determined rightly in all cases. Parentage was determined by comparing genotypes of putative parents and nestlings at three microsatellite loci: FhU2, FhU3 and FhU4. Their combined exclusion power is about 96% in the collared flycatcher (Sheldon and Ellegren 1996). This means that in about 4% of cases nestlings sired by an extra-pair male are erroneously concluded to be sired by the pair male. It was not possible to determine sex and parentage in three offspring due to their disappearance from the nest or decay of tissues.

Samples and statistics

Out of 224 artificially incubated eggs originating from 38 nests, 180 hatched, which represents hatchability of 80.4%. Only nests in which either all or all but one young hatched were used in this study. This ensured a natural level of sibling competition in the studied nests. Mean egg volume of the clutch did not differ between the two groups of nests (nests with high hatchability, mean egg volume $\pm$SE=1620.9$\pm$23.5 mm^3, n=29; nests with low hatchability, mean egg volume$\pm$SE=1635.3$\pm$42.1 mm^3, n=9; t=0.3, P=0.77). Further, only nests where both parents were captured, allowing the determination of parentage, were used. Consequently, 121 chicks hatched in 22 nests remained for the analyses. Hatchability in these nests was 92.4%, which equals the natural level (Cramp and Perrins 1993; M. Krist, unpublished data). Clutch sizes were six, five and seven eggs in 19, two and one nest, respectively. All clutch sizes were pooled for the analyses. Nevertheless, results were virtually the same when only six-egg clutches were used (results not shown). Analyses of tarsus length, nestling mass and fledgling condition (residuals from the regression of 13-day body mass on tarsus length; weight=0.852+0.669 tarsus, n=89, P=0.008, r^2=0.079), were based only on nestlings that subsequently fledged, because nestlings that died did not exhibit normal growth for several days before death (i.e. their mass remained constant or even decreased when the mass of their sibs increased). The only exception were young from three nests that were abandoned at the end of the nestling phase, probably due to depredation of parents. These young grew normally before their abandonment and were included in the analysis of nestling mass up to the day before strong mass recession was recorded.

Because the aim of this study was to analyse the effects of intraclutch egg-size variation, egg volume was converted to relative egg volume (hereafter termed "egg size"). This was computed as egg volume minus the mean egg volume of the clutch (i.e. centring). In this way, between-clutch variation is removed and relative egg volume then represents egg-size variation within clutches. To enable comparison between nests, hatching time was computed for every nestling as follows. The value of zero was assigned to the first-hatched young. Time (in hours) elapsed between hatching of the first young and every subsequent nest-mate was assigned to the latter. The resulting variable is hereafter termed "hatching asynchrony".

To assess the effect of egg size on offspring performance, five models were fitted. The response variables in these models were nestling survival (binomial variable; fledged versus not fledged), recruitment to the breeding population (binomial variable; recruited versus not recruited; only young that fledged successfully were used for this test), fledgling tarsus length, fledgling condition and nestling mass, respectively. The predictor variables were as follows. Firstly, egg size, hatching asynchrony, laying sequence, sex and paternity were retained in all the models as fixed effects of interest. The only exception was the model for nestling survival, which was fitted without paternity because all extra-pair young fledged. In this latter case, maximum likelihood estimates of effects may not exist and thus the validity of the model fit would be questionable. The reason for including the above variables in all final models was that their effects on offspring performance after controlling for the other factors are not known and that is why they may be of interest. Secondly, mean egg volume of the clutch, year, advancement of the breeding season (standardised between years by subtracting the median date of egg-laying in the particular year from the actual egg-laying date), and the time elapsed between hatching and returning of the hatchling to the nest were included in initial models as fixed-effects covariates. To test also for the possibility that the effect of egg size on the response variables depends on brood-level variables, interactions of egg size with year, breeding season and mean egg volume of a clutch were initially fitted in all models. Covariates and their interactions with egg size were selected according to Akaike's information criterion (AIC). The final model was that with the lowest number of parameters from the series of models which had AIC between the smallest value and the smallest value+2 (see Burnham and Anderson 1998). Thirdly, nest was included as a random effect to control for dependence of data points within nests. Denominator df were computed using Satterthwaite's method. Recruitment and survival were analysed using GLIMMIX macro of SAS (generalised linear mixed model with binomial error and logit link), the other models were fitted using PROC MIXED.

The model for nestling mass was more complex than the other models. First, nestlings were weighed each day until the brood was 13 days old (brood age zero is the day the first egg of a clutch hatched). Hence, an individual nestling was treated as a second random factor nested within nest (i.e. higher-level factor) and age of the brood as an additional fixed effect (for the rationale of the model, see Singer 1998). Second, all the interactions of brood age with fixed effects of interest were initially included in the model to investigate whether the effect of independent variables changes as young grow. Interactions were selected according to AIC as described above.

Hatching asynchrony, laying order and egg size were positively correlated (Fig. 1). Such correlations between independent variables in multiple regression (multicollinearity) could reduce the power of the analyses. To assess the influence of multicollinearity on our significance tests, we looked at variance inflation factors (VIF) for individual predictors. Predictors with VIF<10 are generally accepted as giving unbiased results (Chaterjee et al. 2000). Recently, it was suggested that even VIF as small as 2 might bias results (Graham 2003). In our analyses, VIF for paternity in the three models with the response normally distributed were between 2 and 2.5, VIF for all other predictors in all models were <2. Thus, multicollinearity should not have seriously biased our significance tests. Moreover, in contrast to significance tests, parameter estimates are always unbiased even when multicollinearity is high (Freckleton 2002). All tests are two-tailed and were computed in SAS (SAS Institute 2000).

Results

Egg size was more variable between than within clutches ($F_{21,109}$=12.48, r^2=0.71). Intraclutch egg-size variation thus represents about 29% of the total variation, which is the same as the mean figure for 26 studies reviewed by
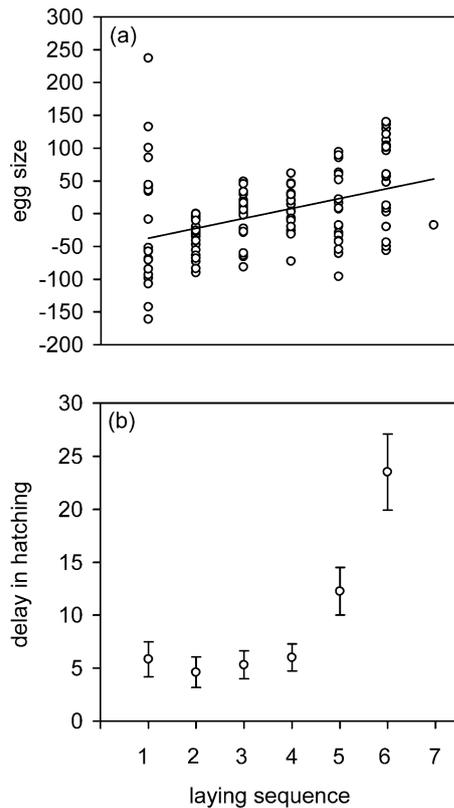
Fig. 1 a Egg size as a function of the laying sequence. The regression equation is: egg size (mm^3)=−52.19+14.96 laying sequence ($F_{1,129}$=24.39, $P<0.001$, r^2=0.16). **b** Time (mean±SE; hours) elapsed between hatching of the first-hatched egg in a clutch and the egg at a particular position in the laying sequence of the same clutch in six-egg clutches. The time is never zero because in different nests the first-hatched egg was at a different position (usually one, two, three, or four) in the laying sequence

Christians (2002). The largest egg in a clutch was on average (±SD) 177.2±68.7 mm^3 larger than the smallest one (range 79.9–332.8 mm^3). Expressed as a percent of the smallest egg, this gives 11.6±4.3% (mean±SD) and 5.0–19.4% (range). The difference between the largest and the smallest mean egg volume of a clutch was 431.0 mm^3, which is 31.2% of the smaller mean egg volume. Egg size increased with laying sequence (Fig. 1a). In six-egg clutches, hatching order reflected laying order in such a way that penultimate and ultimate eggs hatched on average about 7 and 18 h later than preceding eggs, respectively (Fig. 1b). Extra-pair paternity was detected in ten of the 22 nests (45%); extra-pair offspring comprised 30 of 128 nestlings/eggs (23.4%). Sex ratio was male-biased: 80 males/48 females (r=0.63). Hatchability was unrelated to egg size (hatched eggs, −0.10±5.89 mm^3 (mean±SE), n=121; unhatched eggs, 1.15±20.50 mm^3, n=10; t=−0.06, P=0.95).

Egg size did not affect probability of recruitment, nestling survival, fledgling tarsus length and fledgling condition (Table 1). Nestling mass was initially strongly affected by egg size. However, this effect declined steadily as nestlings grew and was not statistically significant after the brood was >9 days old (Table 1, Figs. 2a, b, 3). Weaker

correlation between the mass of 1-day-old nestlings and egg size is probably a methodological artefact caused by the fact that the mass was measured only to the nearest 0.25 g. Therefore, the same mass was assigned to many 1-day-old nestlings although in fact they differed in mass. From a brood age of 3 days onwards, this level of precision was fully adequate as nestlings were much heavier.

In contrast to egg size, hatching asynchrony had a strong effect on offspring performance. Late-hatching young were smaller and had poorer survival than early-hatching young (Table 1). On the other hand, the young from later eggs in the laying sequence were larger than the young from earlier eggs (Table 1). Time to return of offspring to the nest negatively affected offspring survival probability (Table 1). At the time of fledging, extra-pair young tended to be in better condition than young sired by social mates, and sons were in better condition than daughters (Table 1). As it was impossible to assess the effect of paternity on offspring survival while controlling for the effects of other predictors in the generalised linear model (see Materials and methods), we computed at least the probability that extra-pair and pair offspring differ in their survival by Fisher's exact test. Extra-pair young tended to survive better than young sired by social mates (n=101, P=0.067).

To assess the validity of the statistically non-significant results concerning egg size, we computed 95% confidence
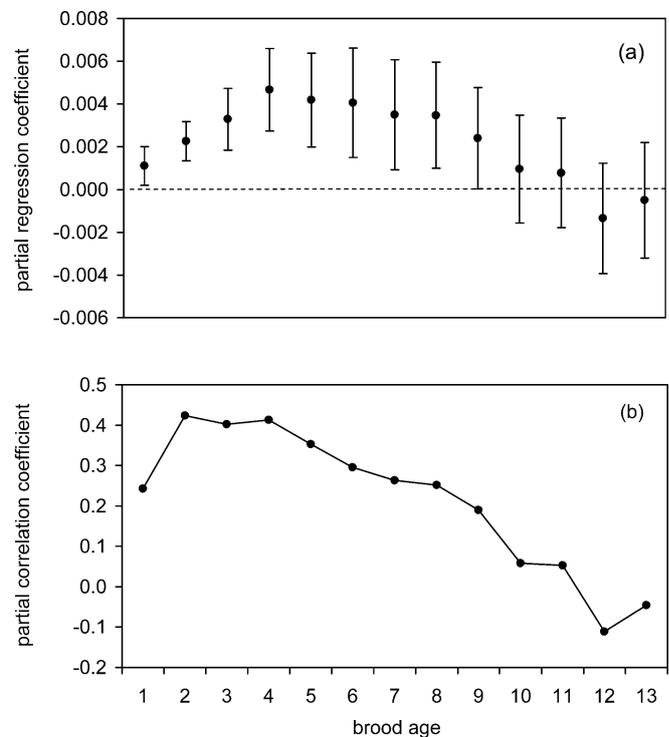
Fig. 2a, b Nestling mass as a function of the relative egg size in the course of the nestling period. Displayed are **a** partial regression (±95% confidence intervals) and **b** partial correlation coefficients between the two , partialled with respect to hatching asynchrony, laying order, offspring sex, paternity and nest. Brood age is in days. Brood age zero is the day the first egg of a clutch hatched

Table 1 Parameter estimates and type III *F*-tests of fixed effects for recruitment[a], nestling survival[a], mass, fledgling tarsus length and fledgling condition

	Estimate	SE	*df*[b]	*F*	*P*
Recruitment					
Intercept	−2.976	1.198	83		
Relative egg size	−0.00102	0.00703	83	0.02	0.886
Laying sequence	0.1002	0.335	83	0.09	0.766
Hatching asynchrony	0.0250	0.0460	83	0.30	0.589
Sex	−0.375	0.941	83	0.16	0.691
Paternity	0.260	0.944	83	0.08	0.784
Nestling survival					
Intercept	8.07	1.62			
Relative egg size	0.00868	0.00649	91.6	1.79	0.184
Laying sequence	0.203	0.305	91.1	0.44	0.507
Hatching asynchrony	−0.263	0.0522	93.7	25.42	<0.001
Sex	−0.253	0.659	88.4	0.15	0.702
Time to return	−0.479	0.140	90.5	11.77	<0.001
Nestling mass					
Intercept	1.474	0.224			
Brood age	1.173	0.0323	19.2	1,271.53	<0.001
Relative egg size	0.00456	0.00127	1254	12.91	<0.001
Laying sequence	0.0670	0.0364	164	3.39	0.067
Hatching asynchrony	−0.0664	0.00662	150	100.60	<0.001
Sex	−0.115	0.105	160	1.20	0.275
Paternity	0.0918	0.150	188	0.37	0.542
Relative egg size × brood age	−0.00034	0.000164	574	4.32	0.038
Tarsus length					
Intercept	16.53	1.49			
Relative egg size	−0.00111	0.000699	66.6	2.52	0.118
Laying sequence	0.0744	0.0341	68.1	4.76	0.033
Hatching asynchrony	−0.0203	0.00573	74.6	12.58	<0.001
Sex	0.130	0.0990	67.2	1.74	0.191
Paternity	−0.0107	0.135	81.7	0.01	0.937
Mean egg size	0.00192	0.000915	15.1	4.39	0.054
Condition					
Intercept	0.669	0.385			
Relative egg size	0.000175	0.00124	66.4	0.02	0.888
Laying sequence	−0.0644	0.0604	67.3	1.14	0.290
Hatching asynchrony	0.00135	0.0103	70.7	0.02	0.896
Sex	−0.398	0.174	67.6	5.24	0.025
Paternity	0.465	0.253	78.8	3.39	0.069
Year	−1.06	0.495	16.8	4.62	0.046

[a]In the models for recruitment and nestling survival the probabilities for recruitment/survival are modelled. The model for recruitment does not contain a random factor, all other models contain a random intercept for nest. The variables are coded as follows—sex: male = 0, female = 1; paternity: young sired by social male = 0, by extra-pair male = 1; year: 2002 = 0, 2001 = 1

[b]Numerator *df* were always 1

intervals for standardised effect sizes of egg size on fledgling mass, condition and tarsus length (Fig. 4). This method may be preferable to commonly used power analysis because confidence intervals have several advantages as compared to power analysis in evaluating non-significant results (e.g. Steidl et al. 1997; Hoenig and Heisey 2001). Effects were predicted for mean difference between the largest and the smallest egg in the clutch, i.e. 177.2 mm^3 (11.6% of the smaller egg). We multiplied this value by the parameter estimate for the effect of egg size on a particular trait and its 95% confidence limits and standardised by dividing them by the SD of the particular trait. Cohen (1988) suggested a convention that the values of standardised effects of 0.2, 0.5 and 0.8 could be treated

as small, medium and large effects, respectively, when two groups are compared. Thus, our data suggest that the difference between the largest and the smallest eggs within clutches could at most cause only small positive effects in fledgling mass and condition (Fig. 4).

Discussion

Egg-size effects

Despite controlling for a number of potentially confounding variables (hatching asynchrony, sex, paternity and laying sequence), we found no effect of within-clutch
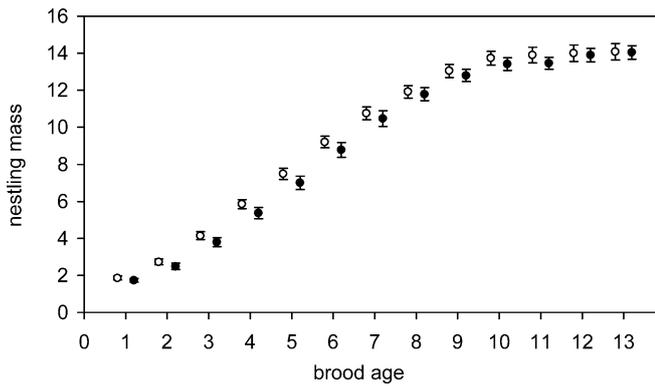
Fig. 3 Nestling mass (mean±2SE, grams) in relation to brood age (days) for nestlings from large (greater than clutch mean) (*open circle*) and small (smaller than clutch mean) (*filled circle*) eggs. Brood age zero is the day the first egg of a clutch hatched
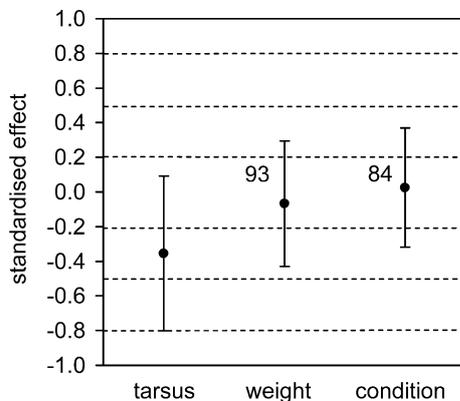
Fig. 4 Standardised effects (estimate±95% confidence interval) caused by the mean difference between the largest and smallest egg in the clutch (177.2 mm^3) for fledgling tarsus length, mass and condition. The *dashed lines* refer to small (0.2), medium (0.5) and large (0.8) standardised effects as suggested by Cohen (1988). *Numbers* by the confidence intervals are probabilities (in percent) that the size of the standardised effect is below that indicated by the *dashed line*

deviations in egg size on recruitment, nestling survival, fledgling condition and fledgling tarsus length of the collared flycatcher. Similarly, the initially strong effect of the egg size on the nestling mass diminished steadily as young were growing, resulting in fledgling mass being independent of egg size. Moreover, the fact that interactions between egg size and some brood-level factors (year, advancement of breeding season and mean egg volume of the clutch) did not improve the model fit substantially, suggests that the lack of egg-size effects holds true in a wide range of external conditions. Of course, it is possible that within-clutch deviations in egg size affected some component of offspring fitness that we did not measure (e.g. immunocompetence). However, fledgling condition, mass and structural size (indicated by tarsus length) are generally assumed to be the traits with the greatest influence on fitness later in life (Gebhardt-Henrich and Richner 1998). This may be true in the collared flycatcher because tarsus length and condition at fledging were found to be under strong directional selection in this species

(Kruuk et al. 2001; Merilä et al. 2001). Furthermore, the narrow confidence intervals for egg-size effects suggest that the lack of effect was not caused by small sample size, at least in the three measures of offspring size for which such confidence intervals can be determined (as the response variable is distributed normally).

As this is a correlative study, it is impossible to derive definite conclusions about causal relationships. However, we have suggested previously that the within-clutch approach is the most powerful one from the range of non-experimental approaches available for the analysis of egg-size effects (Krist and Remeš 2004). Our conclusions of no egg-size effects could be compromised only when other pre-laying maternal effects affecting offspring performance (e.g. concentration of carotenoids or steroids in egg yolk) are negatively correlated with within-clutch deviations in egg size thus tending to cancel each other out (see Krist and Remeš 2004). Although this problem is solvable only by experimental manipulation of egg size, we believe that such counteractive pre-laying maternal effects are unlikely. Thus although the definite answer to the question of egg-size effects on offspring performance can be derived only from a manipulative study, results of our detailed study suggest that within-clutch variation in egg size is unimportant for offspring performance in the collared flycatcher.

So far only a few studies have examined effects of within-clutch differences in egg size on offspring survival (Leblanc 1987b; Grant 1991; Williams et al. 1993b; Dawson and Clark 1996; Amat et al. 2001) and even fewer on their effects on offspring mass in birds (Howe 1976; Anderson et al. 1997; Erikstad et al. 1998; Magrath et al. 2003). Of the four studies on the effect of intraclutch egg-size variation on offspring mass, egg size was not influential in one study (Magrath et al. 2003), was fully confounded with laying order in another one (Erikstad et al. 1998) and influential in the remaining two studies (Howe 1976; Anderson et al. 1997). Out of the five studies on offspring survival, only one (Amat et al. 2001) found a positive effect of larger eggs. Thus, our finding of no effects of within-clutch deviations in egg size on components of offspring performance is in accordance with most previous studies, which suggests that adaptive explanations of intraclutch egg-size patterns should be invoked with caution.

An alternative to adaptive explanations is that intra-clutch egg-size patterns can be primarily generated by proximate constraints on egg development. For example, in studies that found the effect of embryo sex on egg size, the difference ranged from 1 to 3.5% of the smaller egg (Mead et al. 1987; Anderson et al. 1997; Cordero et al. 2000, 2001; Blanco et al. 2003; Magrath et al. 2003). It seems unlikely to us that the 1.3% difference in egg size in favour of males found in the house sparrow (*Passer domesticus*) (Cordero et al. 2000) would be sufficient to improve offspring fitness significantly when, in the collared flycatcher, a species with similar nestling development and breeding biology, a 11.6% difference was unimportant to the offspring. Instead, the sex of the

embryo could affect in some, yet unknown, way the deposition of albumen into the egg in the oviduct thus causing the slight difference in egg size. Similarly, a number of proximate constraints and processes could generate complex patterns of egg-size variation with laying sequence. These include the depletion of endogenous reserves in capital breeders (Pierotti and Bellrose 1986), gearing up physiologically for egg production (Parsons 1976), changing hormonal levels at the onset of egg laying and incubation (Mead and Morton 1985), changing the ambient temperature before egg laying (Ojanen et al. 1981; Järvinen and Ylimaunu 1986), a time constraint for laying early (Slagsvold and Lifjeld 1989; Nilsson and Svensson 1993) and a changing food supply in income breeders (Perrins 1970). A different degree of importance of individual proximate factors in different species/populations could be responsible for a rich diversity of intraclutch egg-size variation patterns found in birds.

Moreover, the relative importance of individual proximate factors may change systematically with body size, breeding habitat or latitude. In an influential study, Slagsvold et al. (1984) showed that large bird species lay small last eggs (brood-reduction strategy) and small species lay big last eggs (brood-survival strategy). They argued that this pattern is adaptive because large birds are less vulnerable to nest predation than small species. However, the alternative explanation could be that this pattern is driven by proximate constraints. Large species are often capital breeders, whereas small species are income breeders (see Meijer and Drent 1999). Since endogenous reserves may be depleted during laying, large species could lay relatively small last eggs. In contrast, small species must forage for all nutrients that they deposit into eggs and since food supply increases at the time of laying from day-to-day (Perrins 1970), they may gather more resources at the end of laying resulting in relatively large last eggs. This hypothesis seems to explain the observed pattern, i.e. a gradual change in egg size with laying sequence (Howe 1976; Wiggins 1990; Cichoń 1997; Hillström 1999; this study), better than the adaptive hypothesis, which predicts that only the last eggs hatching asynchronously should be larger/smaller (Howe 1976; Slagsvold et al. 1984).

Hatching asynchrony, laying sequence and paternity

As in many other studies (reviewed in Magrath 1990), hatching asynchrony was an important determinant of both nestling size and survival. This strong effect as compared to no effect of egg size was probably caused by much greater mass differences in nestling hierarchies, generated by hatching asynchrony, than by egg-size differences, as was found in several other species (e.g. Magrath 1992; Viñuela 1996). Nestlings hatching from ultimate, and to a lesser extent penultimate, eggs hatched usually later (Fig. 1) and thus were disadvantaged. On the other hand, laying sequence per se tended to affect positively both

nestling mass and fledgling tarsus length suggesting that egg composition could be responsible. The concentration of several egg components can change with laying sequence, including antibodies and carotenoids (Saino et al. 2002a, b), or steroid hormones (e.g. Schwabl 1993). Steroid hormones are the most probable candidates causing the larger size of nestlings from later-laid eggs. First, their concentration has recently been found to increase in laying sequence in a number of species (see Whittingham and Schwabl 2002). Second, steroids enhance the development of muscles important for begging (Lipar 2001) and nestling growth (Schwabl 1996).

There was a marginally non-significant effect of paternity on fledgling condition with extra-pair young tending to be in better condition than young sired by the social mate (Table 1), which is in agreement with an earlier study on the same species by Sheldon et al. (1997). This finding supports the hypothesis that extra-pair mates are genetically better than social mates (Sheldon 2000; but see Colegrave 2001) which would have important implications for the theory of sexual selection (Griffith et al. 2002). However, it is still not clear whether a difference in offspring performance in relation to paternity is directly caused by good genes inherited from extra-pair sires or female favouritism of extra-pair young (Gil et al. 1999; Cunningham and Russell 2000). Nevertheless, our study is the first in which some potential pathways of female favouritism (i.e. egg size, hatching asynchrony) were controlled on the within-clutch level. Recently two other studies have demonstrated that, on the between-female level, genetic benefits of mate choice (Parker 2003) and polyandry (Kozielska et al. 2004) exist even after controlling for female favouritism. Although extra-pair young tended to survive better in this study it is even more unclear whether this tendency is due to any genetic effect, because in this test no other factors were controlled for. It may be that this tendency was driven by the effect of hatching asynchrony, since extra-pair young hatch earlier than their half-sibs in this species (M. Krist, P. Nádvorník, L. Uvírová, S. Bureš, unpublished data).

Conclusions

It has been often claimed that females exploit intraclutch egg-size variation adaptively in relation to offspring sex and laying order. However, the assumption of increasing offspring performance with increasing egg size has been rarely properly tested. We found no influence of intraclutch egg-size variation on offspring performance. Although a definite conclusion can be made only after manipulative studies are performed, our results strongly suggest that the assumption does not hold in the collared flycatcher, and this finding is in agreement with most previous studies. Consequently, there seems to be limited potential for female birds in many species to exploit within-clutch variation in egg size adaptively. Instead, we suggest that intraclutch egg-size variation most often has

no adaptive significance and is caused by proximate constraints on laying females.

References

Amat JA, Fraga RM, Arroyo GM (2001) Intraclutch egg-size variation and offspring survival in the kentish plover *Charadrius alexandrinus*. Ibis 143:17–23

Anderson DJ, Reeve J, Bird DM (1997) Sexually dimorphic eggs, nestling growth and sibling competition in american kestrels *Falco sparverius*. Funct Ecol 11:331–335

Becker PH, Wink M (2003) Influences of sex, sex composition of brood and hatching order on mass growth in common terns *Sterna hirundo*. Behav Ecol Sociobiol 54:136–146

Bize P, Roulin A, Richner H (2002) Covariation between egg size and rearing condition determines offspring quality: an experiment with the alpine swift. Oecologia 132:231–234

Blanco G, Martínez-Padilla J, Serrano D, Dávila JA, Viñuela J (2003) Mass provisioning to different-sex eggs within the laying sequence: consequences for adjustment of reproductive effort in a sexually dimorphic bird. J Anim Ecol 72:831

Budden AE, Wright J (2001) Begging in nestling birds. In: Nolan V, Thompson CF (eds) Current ornithology. Kluwer, New York, pp 83–118

Burnham KP, Anderson DR (1998) Model selection and inference: a practical information-theoretic approach. Springer, Berlin Heidelberg New York

Chaterjee S, Hadi AS, Price B (2000) Regression analysis by example. Wiley, New York

Christians JK (2002) Avian egg size: variation within species and inflexibility within individuals. Biol Rev 77:1–26

Cichoń M (1997) Egg weight variation in collared flycatchers *Ficedula albicollis*. Ornis Fenn 74:141–147

Cichoń M, Dubiec A, Stoczko M (2003) Laying order and offspring sex in blue tits *Parus caeruleus*. J Avian Biol 34:355–359

Cohen J (1988) Statistical power analysis for behavioral sciences. Erlbaum, N.J.

Colegrave N (2001) Differential allocation and "good genes"—comment from Colegrave. Trends Ecol Evol 16:22–23

Cordero PJ, Griffith SC, Aparicio JM, Parkin DT (2000) Sexual dimorphism in house sparrow eggs. Behav Ecol Sociobiol 48:353–357

Cordero PJ, Viñuela J, Aparicio JM, Veiga JP (2001) Seasonal variation in sex ratio and sexual egg dimorphism favouring daughters in first clutches of the spotless starling. J Evol Biol 14:829–834

Cramp S, Perrins CM (1993) The birds of the western Palearctic. Oxford University Press, Oxford

Cunningham EJA, Russell AF (2000) Egg investment is influenced by male attractiveness in the mallard. Nature 404:74–77

Dawson RD, Clark RG (1996) Effects of variation in egg size and hatching date on survival of lesser scaup *Aythya affinis* ducklings. Ibis 138:693–699

Erikstad KE, Tveraa T, Bustnes JO (1998) Significance of intraclutch egg-size variation in common eider: the role of egg size and quality of ducklings. J Avian Biol 29:3–9

Freckleton RP (2002) On the misuse of residuals in ecology: regression of residuals vs. multiple regression. J Anim Ecol 71:542–545

Gebhardt-Henrich SG, Richner H (1998) Causes of growth variation and its consequences for fitness. In: Starck JM, Ricklefs RE (eds) Avian growth and development: evolution within the altricial-precocial spectrum. Oxford University Press, Oxford, pp 324–339

Gil D, Graves J, Hazon N, Wells A (1999) Male attractiveness and differential testosterone investment in zebra finch eggs. Science 286:126–128

Graham MH (2003) Confronting multicollinearity in ecological multiple regression. Ecology 84:2809–2815

Grant MC (1991) Relationship between egg size, chick size at hatching and chick survival in the whimbrel *Numenius phaeopus*. Ibis 133:127–133

Griffith SC, Owens IPF, Thuman KA (2002) Extra pair paternity in birds: a review of interspecific variation and adaptive function. Mol Ecol 11:2195–2212

Griffiths R, Double MC, Orr K, Dawson RJG (1998) A DNA test to sex most birds. Mol Ecol 7:1071–1075

Hillström L (1999) Variation in egg mass in the pied flycatcher, *Ficedula hypoleuca*: an experimental test of the brood survival and brood reduction hypotheses. Evol Ecol Res 1:753–768

Hoenig JM, Heisey DM (2001) The abuse of power: the pervasive fallacy of power calculations for data analysis. Am Stat 55:19–24

Howe HF (1976) Egg size, hatching asynchrony, sex and brood reduction in the common grackle. Ecology 57:1195–1207

Hoyt DF (1979) Practical methods of estimating volume and fresh weight of birds' eggs. Auk 96:73–77

Järvinen A, Ylimaunu J (1986) Intraclutch egg-size variation in birds: physiological responses of individuals to fluctuations in environmental conditions. Auk 103:235–237

Kozielska M, Krzemińska A, Radwan J (2004) Good genes and the maternal effects of polyandry on offspring reproductive success in the bulb mite. Proc R Soc Lond B 271:165–170

Krist M, Remeš V (2004) Maternal effects and offspring performance: in search of the best method. Oikos (in press)

Kruuk LEB, Merilä J, Sheldon BC (2001) Phenotypic selection on a heritable size trait revisited. Am Nat 158:557–571

Leblanc Y (1987a) Relationships between sex of gosling and position in the laying sequence, egg mass, hatchling size and fledgling size. Auk 104:73–76

Leblanc Y (1987b) Egg mass, position in the laying sequence and brood size in relation to canada goose reproductive success. Wilson Bull 99:663–672

Lipar JL (2001) Yolk steroids and the development of the hatching muscle in nestling european starlings. J Avian Biol 32:231–238

Magrath RD (1990) Hatching asynchrony in altricial birds. Biol Rev 65:587–622

Magrath RD (1992) Roles of egg mass and incubation pattern in establishment of hatching hierarchies in the blackbird (*Turdus merula*). Auk 109:474–487

Magrath MJL, Brouwer L, Komdeur J (2003) Egg size and laying order in relation to offspring sex in the extreme sexually size dimorphic brown songlark, *Cinclorhamphus cruralis*. Behav Ecol Sociobiol 54:240–248

Mead PS, Morton ML (1985) Hatching asynchrony in the mountain white-crowned sparrow (*Zonotrichia leucophrys oriantha*): a selected or incidental trait? Auk 102:781–792

Mead PS, Morton ML, Fish BE (1987) Sexual dimorphism in egg size and implications regarding facultative manipulation of sex in mountain white-crowned sparrows. Condor 89:798–803

Meijer T, Drent R (1999) Re-examination of the capital and income dichotomy in breeding birds. Ibis 141:399–414

Merilä J, Kruuk LB, Sheldon BC (2001) Cryptic evolution in a wild bird population. Nature 412:76–79

Nilsson JA, Svensson E (1993) Causes and consequences of egg mass variation between and within blue tit clutches. J Zool 230:469–481

O'Connor RJ (1979) Egg weights and brood reduction in the european swift (*Apus apus*). Condor 81:133–145

Ojanen M, Orell M, Väisänen A (1981) Egg size variation within passerine clutches: effects of ambient temperature and laying sequence. Ornis Fenn 58:93–108

Parker TH (2003) Genetic benefits of mate choice separated from differential maternal investment in red junglefowl (*Gallus gallus*). Evolution 57:2157–2165

Parsons J (1976) Factors determining the number and size of eggs laid by the herring gull. Condor 78:481–492

Pärt T (1990) Natal dispersal in the collared flycatcher: possible causes and reproductive consequences. Ornis Scand 21:83–88

Pelayo JT, Clark RG (2003) Consequences of egg size for offspring survival: a cross-fostering experiment in ruddy ducks (*Oxyura jamaicensis*). Auk 120:384–393

Perrins CM (1970) The timing of birds' breeding seasons. Ibis 112:242–255

Pierotti R, Bellrose CA (1986) Proximate and ultimate causation of egg size and the third-chick disadvantage in the western gull. Auk 103:401–407

Rutkowska J, Cichoń M (2002) Maternal investment during egg laying and offspring sex: an experimental study of zebra finches. Anim Behav 64:817–822

Saino N, Bertacche V, Ferrari RP, Martinelli R, Møller AP, Stradi R (2002a) Carotenoid concentration in barn swallow eggs is influenced by laying order, maternal infection and paternal ornamentation. Proc R Soc Lond B 269:1729–1733

Saino N, Ferrari RP, Martinelli R, Romano M, Rubolini D, Møller AP (2002b) Early maternal effects mediated by immunity depend on sexual ornamentation of the male partner. Proc R Soc Lond B 269:1005–1009

SAS Institute (2000) SAS online document, version 8. SAS Institute, Cary, N.C

Schwabl H (1993) Yolk is a source of maternal testosterone for developing birds. Proc Natl Acad Sci USA 90:11446–11450

Schwabl H (1996) Maternal testosterone in the avian egg enhances postnatal growth. Comp Biochem Physiol A 114:271–276

Seutin G, White BN, Boag PT (1991) Preservation of avian blood and tissue samples for DNA analyses. Can J Zool 69:82–90

Sheldon BC (2000) Differential allocation: tests, mechanisms and implications. Trends Ecol Evol 15:397–402

Sheldon BC, Ellegren H (1996) Offspring sex and paternity in the collared flycatcher. Proc R Soc Lond B 263:1017–1021

Sheldon BC, Merilä J, Qvarnström A, Gustafsson L, Ellegren H (1997) Paternal genetic contribution to offspring condition predicted by size of male secondary sexual character. Proc R Soc Lond B 264:297–302

Singer JD (1998) Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. J Educat Behav Stat 24:323–355

Slagsvold T, Lifjeld JT (1989) Constraints on hatching asynchrony and egg size in pied flycatchers. J Anim Ecol 58:837–849

Slagsvold T, Sandvik J, Rofstad G, Lorentsen O, Husby M (1984) On the adaptive value of intraclutch egg-size variation in birds. Auk 101:685–697

Steidl RJ, Hayes JP, Schauber E (1997) Statistical power analysis in wildlife research. J Wildl Manage 61:270–279

Teather KL (1989) Sex and egg size in great-tailed grackles. Condor 91:203–205

Trivers RL, Willard DE (1973) Natural selection of parental ability to vary the sex ratio of offspring. Science 179:90–92

Viñuela J (1996) Establishment of mass hierarchies in broods of the black kite. Condor 98:93–99

Viñuela J (1997) Adaptation vs. constraint: intraclutch egg-mass variation in birds. J Anim Ecol 66:781–792

Weatherhead PJ (1985) Sex-ratios of red-winged blackbirds by egg size and laying sequence. Auk 102:298–304

Whittingham LA, Schwabl H (2002) Maternal testosterone in tree swallow eggs varies with female aggression. Anim Behav 63:63–67

Wiggins DA (1990) Sources of variation in egg mass of tree swallows *Tachycineta bicolor*. Ornis Scand 21:157–160

Williams TD, Lank DB, Cooke F (1993a) Is intraclutch egg-size variation adaptive in the lesser snow goose? Oikos 67:250–256

Williams TD, Lank DB, Cooke F, Rockwell RF (1993b) Fitness consequences of egg-size variation in the lesser snow goose. Oecologia 96:331–338

Ylimaunu J, Järvinen A (1987) Do pied flycatchers *Ficedula hypoleuca* have brood-survival or brood-reduction strategy? Ornis Fenn 64:10–15

Článek 3: Krist, M. & Remeš, V. (2004) Maternal effects and offspring performance: in search of the best method. *Oikos*, **106, 422-426.**

Maternal effects and offspring performance: in search of the best method

Miloš Krist, Museum of Natural History, nám. Republiky 5, CZ-771 73 Olomouc, Czech Republic (krist@prfnw.upol.cz). − MK and Vladimír Remeš, Laboratory of Ornithology, Palacký University, tř. Svobody 26, CZ-771 46 Olomouc, Czech Republic.

Traditionally, maternal effects have been treated as a source of troublesome environmental variance that confounds our ability to accurately estimate the genetic basis of the traits of interest. However, the adaptive significance of maternal effects is currently at the centre of the attention of ecologists. Thus, in turn, the genetic basis of traits has become a troublesome source of the genetic resemblance that confounds our ability to accurately estimate the maternal effects of interest. This fact is, however, less widely realized among ecologists. We demonstrate this on the example of studies investigating egg-size effects on offspring performance in birds. Traditionally this relationship is being studied by cross-fostering of eggs or young and it is claimed that this design is able to separate the effects of egg size per se. However, a positive covariation between the direct effects of genes and the maternal effects exists for many studied traits, which may result in overestimation of the egg-size effects on offspring performance in cross-fostering studies. Within-clutch comparisons or direct experimental manipulation of the egg size are the approaches that do not suffer from such covariation and therefore give less biased estimates of the egg-size effects than cross-fostering studies.

Maternal effects in animal ecology

Offspring phenotype is determined by genes and the environment. Besides the direct effect of genes and the environment, maternal effects often play a significant role. Previously, maternal effects have been treated as the source of confusion in determining precisely quantitative genetic parameters (Falconer 1989). However, it is now widely appreciated that they can play an important role in driving the dynamics of evolution and population growth. Specifically, by introducing time lags into both these processes, they may lead to unpredictable evolutionary trajectories (Kirkpatrick and Lande 1989) and destabilization of population dynamics, e.g. in the form of population cycles or decaying oscillations (Ginzburg 1998, Beckerman et al. 2002). Moreover, at the individual level, maternal effects may influence offspring fitness and thus serve both offspring and parents as adaptations. This adaptive significance of maternal effects has recently become a popular focus of ecological

and evolutionary studies (Mousseau and Fox 1998). Traditionally, maternal effects have been studied in domesticated species by complex analyses (Lynch and Walsh 1998), which is not easily applicable to free-ranging animals. Moreover, these analyses just partition the variance in the focal trait and determine which part of this variance can be ascribed to maternal effects in general. Lande and Price (1989) devised a regression method based on Kirkpatrick and Lande (1989) that is able to isolate maternal effects specific for certain maternal traits. However, this method requires that all the maternal characters exerting maternal effects be included in the analyses (rather unrealistic condition) and is not readily applicable to sex-limited characters. Below we evaluate other methods for studying maternal effects employed in wild populations, including sex-limited characters, on the example of the effects of egg size on offspring performance in birds.

Studying egg-size effects on offspring performance is important for two areas of evolutionary ecology. One, life-history theory predicts a trade-off between number and quality of offspring produced from limited resources. Two, potentially adaptive allocation of limited resources among siblings within a clutch is widely studied in a broad range of species. The critical assumption in both these cases is that the amount of resources allocated to an egg has an effect on offspring performance. Consequently, egg-size effects on offspring performance are among the most frequently studied topics in the area of maternal effects and birds are the taxon in which these effects have been studied most often. We are aware of at least 60 studies dealing with this question in birds, 41 of which were reviewed by Williams (1994). However, the approaches usually employed do not control for potential confounding factors. Consequently, despite high research attention, results of many studies estimating egg-size effects on offspring performance may be biased.

Quantitative genetic framework

To demonstrate how the effect of egg size per se on offspring performance can be derived, it is useful to frame the problem in the quantitative genetics terms. From the quantitative genetics perspective, the phenotypic value of each trait can be partitioned into the components attributable to genes (genotypic value) and the environment (environmental deviation) (Falconer 1989). The genotypic value can be further partitioned into the breeding value (determined by additive effect of genes), the dominance deviation (interactions of alleles within the same locus) and the interaction deviation (interactions of alleles between loci, i.e. epistasis). In this basic framework, maternal effects are subsumed under the environmental effects, because they are defined as the non-genetic influence of the maternal phenotype on the offspring's phenotype (Kirkpatrick and Lande 1989).

For our purposes, egg size is singled out as the maternal effect of interest, whereas all other maternal effects (e.g. parental feeding, concentration of testosterone in the egg) and pure environmental effects (e.g. weather, food supply) are grouped together as the offspring environment. Consequently, the phenotypic value of each offspring's trait (z_x) can be viewed as being determined by three sources: the direct effect of genes (G_{ox}), the offspring's environment E_x, and the egg size (S_m) which is itself compounded of an environmental (E_{mw}) and a genetic (G_{mw}) component (Fig. 1). Here, subscript x denotes an offspring and w mother; o is direct pathway of determining offspring phenotype, m denotes maternal (indirect) pathway through egg size. Throughout, we assume that direct and indirect genetic
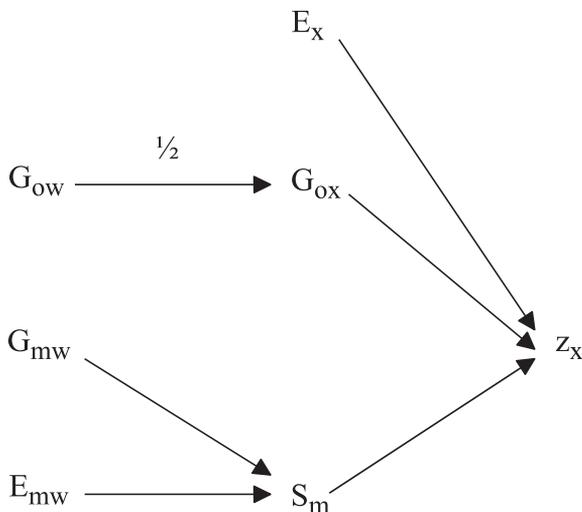

Fig. 1. Path diagram depicting the determination of the phenotype z_x of an individual x by direct genetic effects G_{ox}, environmental effects E_x (including also all parental effects except of egg size) and egg size S_m. The mother of x is denoted by w. O denotes direct pathway, m denotes maternal pathway. Modified from Lynch and Walsh (1998).

effects, G_{ox} and G_{mw}, contain additive effects but do not exhibit dominance or epistasis. Note that E_x has neither o nor m subscript, because it includes both direct and maternal effects, and S_m has neither w nor x subscript, because it is the trait of both the mother and the offspring.

If the three sources determining the traits of offspring were uncorrelated, a simple statistical technique such as the linear regression would give a reliable estimate of egg-size effects on any of offspring traits (except of daughter's egg size). However, when these sources are correlated, experimental or statistical techniques will be needed to separate egg-size effects per se.

A review of methods

In 1990 it was suggested that the size of the egg a female lays might be positively correlated with the quality of her territory or subsequent parental feeding rate to young (i.e. positive $CovE_xS_m$ exists) and that this correlation can be removed by experimental swapping of clutches/broods between nests – a technique known as cross-fostering (Amundsen and Stokland 1990, Reid and Boersma 1990). In both these studies, the authors found that the size of the eggs which foster mothers originally laid was a more influential determinant of offspring traits at fledging than the size of the egg from which the offspring actually hatched. This suggests that when cross-fostering is not performed, egg-size effects on offspring performance are highly overestimated. Therefore cross-fostering became a very popular technique to study egg-size effects on offspring performance (we know of 16 such studies performed to date).

However, although cross-fostering decouples much of correlation of egg size with offspring environment ($CovE_xS_m$, below, Table 1), it does not deal with the possible covariation between direct effect of genes and egg size ($CovG_{ow}S_m$). Yet, this covariation is likely to be large in many cases because of the choice of 'fitness' measure usually employed in studies of this kind. In practice, instead of studying directly the effects of egg size on offspring fitness, we usually study the effects on some correlate of fitness. Morphological traits of fledglings such as tarsus length or body weight are frequently used as these correlates. However, female size is frequently positively correlated with the size of eggs she lays (Christians 2002) and at the same time morphological traits are highly heritable (Merilä and Sheldon 2001). This means that the size of the trait in the offspring is highly influenced by direct effect of additive genes but this effect is ascribed (to the extent to which additive genes for mother body size are involved in the correlation between female size and egg size) to the effect of egg size in the cross-fostering design. In principle this may be a problem in every studied trait including offspring

Table 1. Summarization of relative merits and shortcomings of the individual approaches. $CovE_xS_m$ is a covariance between environmental effects and egg size, $CovG_{ow}S_m$ is a covariance between direct additive genetic effects and egg size.

Approach	Controls for	Remains uncontrolled	Main use
Observational	Nothing	$CovE_xS_m$ $CovG_{ow}S_m$	Do not use
Cross-fostering	Partially $CovE_xS_m$	$CovE_xS_m$ (pre-laying parental effects and differential feeding of young) $CovG_{ow}S_m$	Use with caution (egg size-number trade-off)
Within-clutch	Partially $CovE_xS_m$ Fully $CovG_{ow}S_m$	$CovE_xS_m$ (pre-laying parental effects and differential feeding of young)	Within-clutch adaptive allocation of resources
Manipulation	Partially $CovE_xS_m$ Fully $CovG_{ow}S_m$	$CovE_xS_m$ (differential feeding of young)	For all purposes

survival. Cross-fostering thus does not reveal effects of egg size per se, which was realized only rarely (Magrath 1992, Styrsky et al. 1999) and was not mentioned in the most recent cross-fostering studies (Hipfner et al. 2001, Bize et al. 2002, Pelayo and Clark 2003). The covariation $CovG_{ow}S_m$ can arise either through $CovG_{ow}G_{mw}$ if, for example, the same gene facilitates the conversion of food into yolk in a female and food into flesh in a nestling (Magrath 1992) or through $CovG_{ow}E_{mw}$ if, for example, larger females or males (more precisely individuals with larger breeding values for body size) attain better territories which enable females to produce larger eggs. These covariations have not been quantified in birds so far, however, in mammals it has been found that the genetic covariation between direct and maternal pathways of determining the offspring phenotype might be high (Riska et al. 1985, McAdam et al. 2002). Thus it does not seem reasonable to assume that a similar covariation does not exist in other taxa.

The direct effect of additive genes may be controlled statistically by including the midparent value of the trait as a covariate in the analysis of egg-size effects on the same trait in offspring. Although this has been done with the maternal value of the trait in some observational studies (Larsson and Forslund 1992, Potti and Merino 1994), it has never been done in any cross-fostering study investigating egg-size effects on offspring performance. In some cases addition of such a covariate may be relatively easy – for example when investigating egg-size effects on fledging tarsus length, which is fully grown at the time of fledging in many species. However, this might be very burdensome when investigating traits that can be measured only in offspring (e.g. growth rate) and impossible when investigating survival of offspring up to recruitment, because all parents were successfully recruited.

However, two other approaches that do not suffer from $CovG_{ow}S_m$ can be used to investigate egg-size effects on offspring performance. First, effects of egg size on offspring performance may be compared within clutches. So far, this approach has been used less often than the cross-fostering approach (we know of nine studies using the within-clutch approach, e.g. Dawson and Clark 1996, Amat et al. 2001), and its advantage over cross-fostering was not mentioned in any of these studies. Among-clutch variation in egg size may be removed by centring egg sizes within clutches, i.e. by subtracting mean egg size of the clutch from the actual size of every egg in the clutch. Resulting values represent within-clutch variation and as such are then used in the statistical analyses. Given that chromosomes segregate at random in meiosis, $CovG_{ox}S_m$ is zero among full-sibs. Non-zero $CovG_{ox}S_m$ could arise if the female was able to recognize which allele of an allelic pair had come to the ovum in meiotic division and targeted resources accordingly or to control the outcome of meiosis in relation to the size of ovum to be ovulated. Such high female control, however, seems highly unlikely for alleles on autosomes or homologous parts of sex-chromosomes. On the other hand, targeting of resources might perhaps work in relation to genes that are located at non-homologous parts of sex-chromosomes, such as sex-determining genes, as suggested by studies demonstrating differences in egg size between the sexes (Cordero et al. 2000).

In the within-clutch approach, territory and parents are the same for all sibs and that is why also $CovE_xS_m$ is controlled to a similar degree as in cross-fostering design (Table 1). Strictly speaking, however, $CovE_xS_m$ need not be zero both in within-clutch and cross-fostering design. Firstly, egg size may be correlated with other pre-laying maternal effects, for example concentration of hormones, antibodies or carotenoids in the egg. In this case the amount of these compounds would increase allometrically with egg size (slope of the regression of the amount of a compound on egg size would differ from one) contrary to the situation when it would increase isometrically with egg size (slope would equal one). In the latter case, the amount of these compounds may be treated as being a part of the egg size. Secondly, parents may feed more intensely small (or large) young which hatched from small (or large) eggs. This effect may be

stronger in within-clutch approach, because offspring are raised in the same nest and larger sibs may monopolize resources supplied by parents. Moreover, the problem that is specific for the within-clutch approach is that within-clutch (and also within-female) variation in egg size is usually much smaller than inter-female variation. On average, differences in egg size within clutches explain only 30% of the total egg-size variation (Christians 2002). Thus effects of great differences in egg size, which exist at the population level, cannot be directly estimated by within-clutch approach. On the other hand, there are many studies investigating adaptive allocation of resources among individual eggs within a clutch in relation to laying order (Slagsvold et al. 1984) and sex (Cordero et al. 2000). These studies rely on the assumption that within-clutch variability in egg size has some consequences for offspring performance. This assumption may be properly tested by the within-clutch approach outlined above. Possible monopolization of resources by larger siblings and small differences in egg size within clutches are not problems in this context, because they are inherent features of the relationships among young within a clutch.

The second approach to remove $CovG_{ow}S_m$ is the direct manipulation of egg volume. This method is the best way to elucidate potential effects of egg size on offspring performance (Sinervo et al. 1992). In birds, it was used, to our best knowledge, only twice on domesticated species under laboratory conditions (Hill 1993, Finkler et al. 1998). In these studies, certain part of the albumen or yolk of unincubated eggs was removed by a syringe and a needle. Such an egg size manipulation removes also the potential correlation between egg size and other pre-laying maternal effects, which is an additional advantage compared to the other approaches. Given the strengths of this approach, it could become a powerful tool in elucidating effects of egg size on offspring performance also in populations of wild-ranging birds. However, although invasive egg sampling and manipulation have been successfully applied to some wild bird species (Lipar 2001, Saino et al. 2003), rather high egg mortality encountered in the study on hens mentioned above (Finkler et al. 1998) seems to question broad applicability of this method. Moreover, we have no information on how big changes in egg volume in comparison with the natural egg-size variability are within the reach of this method, while at the same time keeping egg mortality within acceptable limits. Both these issues remain to be addressed in studies on wild species. The manipulative approach also suffers from the possibility that parents may feed their young selectively with respect to their size, which also means to the size of the egg they hatched from. This could be controlled for by statistically controlling for the amount of food brought to individual offspring by their parents or by hand-rearing of the young (Anderson et al. 1997).

Conclusions

In this comment we evaluated relative merits and shortcomings of the different approaches to the study of egg-size effects on offspring performance in birds (Table 1). It has been accepted that the cross-fostering design is better than the simple observational approach and thus it became a standard methodological tool. We argue that there are even better approaches that should give less biased estimates of egg-size effects: the within-clutch approach and the direct experimental manipulation of egg-size. These approaches are relatively readily applicable to free-ranging populations of animals and plants. Therefore, further studies using these approaches would be valuable for better understanding of the evolution and impact of maternal effects and also for the evaluation of how much the traditional approaches for studying adaptive maternal effects suffer from uncontrolled confounding factors. We focused our attention on egg-size effects in birds because this is one of the best-studied systems in the area of maternal effects and much effort has been devoted to it. However, general logic of our argument applies equally well to other traits and other taxa.

References

Amat, J. A., Fraga, R. M. and Arroyo, G. M. 2001. Intraclutch egg-size variation and offspring survival in the kentish plover *Charadrius alexandrinus*. – Ibis 143: 17–23.

Amundsen, T. and Stokland, J. N. 1990. Egg size and parental quality influence nestling growth in the shag. – Auk 107: 410–413.

Anderson, D. J., Reeve, J. and Bird, D. M. 1997. Sexually dimorphic eggs, nestling growth and sibling competition in american kestrels *Falco sparverius*. – Funct. Ecol. 11: 331–335.

Beckerman, A., Benton, T. G., Ranta, E. et al. 2002. Population dynamic consequences of delayed life-history effects. – Trends Ecol. Evol. 17: 263–269.

Bize, P., Roulin, A. and Richner, H. 2002. Covariation between egg size and rearing condition determines offspring quality: an experiment with the alpine swift. – Oecologia 132: 231–234.

Christians, J. K. 2002. Avian egg size: variation within species and inflexibility within individuals. – Biol. Rev. 77: 1–26.

Cordero, P. J., Griffith, S. C., Aparicio, J. M. et al. 2000. Sexual dimorphism in house sparrow eggs. – Behav. Ecol. Sociobiol. 48: 353–357.

Dawson, R. D. and Clark, R. G. 1996. Effects of variation in egg size and hatching date on survival of lesser scaup *Aythya affinis* ducklings. – Ibis 138: 693–699.

Falconer, D. S. 1989. Introduction to quantitative genetics. – Longman Scientific & Technical.

Finkler, M. S., Van Orman, J. B. and Sotherland, P. R. 1998. Experimental manipulation of egg quality in chickens: influence of albumen and yolk on the size and body composition of near-term embryos in a precocial bird. – J. Comp. Physiol. B 168: 17–24.

Ginzburg, L. R. 1998. Inertial growth: population dynamics based on maternal effects. – In: Mousseau, T. A. and Fox, C. V. (eds), Maternal effects as adaptations. Oxford Univ. Press, pp. 42–53.

Hill, W. L. 1993. Importance of prenatal nutrition to the development of a precocial chick. – Develop. Psychobiol. 26: 237–249.

Hipfner, J. M., Gaston, A. J. and Storey, A. E. 2001. Food supply and the consequences of egg size in the thick-billed murre. – Condor 103: 240–247.

Kirkpatrick, M. and Lande, R. 1989. The evolution of maternal characters. – Evolution 43: 485–503.

Lande, R. and Price, T. 1989. Genetic correlations and maternal effect coefficients obtained from offspring–parent regression. – Genetics 122: 915–922.

Larsson, K. and Forslund, P. 1992. Genetic and social inheritance of body and egg size in the barnacle goose (*Branta leucopsis*). – Evolution 46: 235–244.

Lipar, J. L. 2001. Yolk steroids and the development of the hatching muscle in nestling european starlings. – J. Avian Biol. 32: 231–238.

Lynch, M. and Walsh, B. 1998. Genetics and analysis of quantitative traits. – Sinauer Associates, Inc.

Magrath, R. D. 1992. The effect of egg mass on the growth and survival of blackbirds: a field experiment. – J. Zool. 227: 639–653.

McAdam, A. G., Boutin, S., Réale, D. et al. 2002. Maternal effects and the potential for evolution in a natural population of animals. – Evolution 56: 846–851.

Merilä, J. and Sheldon, B. C. 2001. Avian quantitative genetics. – In: Nolan, V. (ed.), Current ornithology, Kluwer Academic/Plenum Publishers.

Mousseau, T. A. and Fox, C. W. 1998. Maternal effects as adaptations. – Oxford Univ. Press.

Pelayo, J. T. and Clark, R. G. 2003. Consequences of egg size for offspring survival: a cross-fostering experiment in ruddy ducks (*Oxyura jamaicensis*). – Auk 120: 384–393.

Potti, J. and Merino, S. 1994. Heritability estimates and maternal effects on tarsus length in pied flycatchers, *Ficedula hypoleuca*. – Oecologia 100: 331–338.

Reid, W. J. and Boersma, P. D. 1990. Parental quality and selection on egg size in the magellanic penguin. – Evolution 44: 1780–1786.

Riska, B., Rutledge, J. J. and Atchley, W. R. 1985. Covariance between direct and maternal genetic effects in mice, with a model of persistent environmental influences. – Genet. Res. 45: 287–297.

Saino, N., Ferrari, R., Romano, M. et al. 2003. Experimental manipulation of egg carotenoids affects immunity of barn swallow nestlings. – Proc. R. Soc. Lond. B 270: 2485–2489.

Sinervo, B., Doughty, P., Huey, R. B. et al. 1992. Allometric engineering: a causal analysis of natural selection on offspring size. – Science 258: 1927–1930.

Slagsvold, T., Sandvik, J., Rofstad, G. et al. 1984. On the adaptive value of intraclutch egg-size variation in birds. – Auk 101: 685–697.

Styrsky, J. D., Eckerle, M. P. and Thompson, C. F. 1999. Fitness-related consequences of egg mass in nestling house wrens. – Proc. R. Soc. Lond. B 266: 1253–1258.

Williams, T. D. 1994. Intraspecific variation in egg size and egg composition in birds: effects on offspring fitness. – Biol. Rev. 69: 35–59.

Článek 4: Krist, M., Nádvorník, P., Uvírová, L. & Bureš, S. (2005) Paternity covaries with laying and hatching order in the collared flycatcher *Ficedula albicollis*. *Behavioral Ecology and Sociobiology*, 59, 6-11.

ORIGINAL ARTICLE

Miloš Krist · Petr Nádvorník · Lenka Uvírová ·
Stanislav Bureš

Paternity covaries with laying and hatching order in the collared flycatcher *Ficedula albicollis*

Abstract Females in most bird species engage in extra-pair copulations. Although this behaviour is widespread, benefits for females of doing so are less understood. The most widely accepted hypothesis is that they improve their previous choice of social partner and gain genetic benefits for their offspring. Some evidence for this comes from studies that find that extra-pair young (EPY) have greater fitness than their half-sibs. However, this might be also caused by maternal, non-genetic effect, a possibility that remains largely untested. Here we test whether EPY are laid in larger eggs or eggs laid early in the laying sequence in the collared flycatcher (*Ficedula albicollis*). The size of eggs bearing EPY and within-pair young (WPY) did not differ, however, EPY were laid in early eggs and consequently hatched earlier than WPY. As hatching asynchrony is a strong determinant of offspring size and survival in many species, including collared flycatcher, our results suggest that a caution is needed when paternal genetic effects are to be inferred from comparison of naturally occurring half-sibs.

M. Krist (✉)
Museum of Natural History,
nám. Republiky 5,
771 73 Olomouc, Czech Republic

M. Krist · S. Bureš
Laboratory of Ornithology, Palacký University,
tr. Svobody 26,
771 46 Olomouc, Czech Republic

P. Nádvorník · L. Uvírová
Department of Cell Biology and Genetics, Palacký University,
Šlechtitelu 11,
783 71 Olomouc, Czech Republic

P. Nádvorník · S. Bureš
Department of Zoology, Palacký University,
tr. Svobody 26,
771 46 Olomouc, Czech Republic

Introduction

Female birds frequently copulate with extra-pair males (Jennions and Petrie 2000; Griffith et al. 2002). Typically, females gain no direct benefits from extra-pair copulations (EPCs) and that is why genetic benefits are usually supposed to drive such a behaviour (Jennions and Petrie 2000; Griffith et al. 2002). In the prevailing view the very convincing evidence for this would be the finding that extra-pair young (EPY) had greater fitness than within-pair young (WPY) (Kempenaers et al. 1997; Sheldon et al. 1997; Krokene et al. 1998; Strohbach et al. 1998; Lubjuhn et al. 1999; Johnsen et al. 2000; Sheldon 2000; Whittingham and Dunn 2001; Griffith et al. 2002; Johnsen et al. 2003; Schmoll et al. 2003; Charmantier et al. 2004; Kleven and Lifjeld 2004). However, recently it has been suggested that greater fitness of EPY cannot be regarded as an evidence of their genetic superiority as they may be favoured by maternal (non-genetic) effects. Females might, for example, put EPY in larger eggs (cf. Cunningham and Russell 2000) or in eggs with greater concentration of testosterone (cf. Gil et al. 1999). Such differential allocation would be beneficial for females if extra-pair sires were more attractive than cuckolded males (Burley 1986; Sheldon 2000) which is indeed commonly the case (e.g. Møller and Ninni 1998; Sheldon and Ellegren 1999).

Nowadays many studies exist that have confirmed that differential allocation indeed exists on the level of breeding attempts (review in Sheldon 2000). When mated with attractive mate females were found to invest more in current reproductive attempt either by producing more offspring (e.g. Balzer and Williams 1998) or by investing more resources per individual offspring (Gil et al. 1999; Cunningham and Russell 2000; Kolm 2001; Saino et al. 2002a; Rutstein et al. 2004). The latter mechanism is available also for females to discriminate between EPY and WPY, i.e. to perform within-brood differential allocation. However, this possibility remains largely untested although in fact it may be more strongly selected than allocation between attempts, as the difference in quality of mates may be compared simultaneously, in contrast to between-attempt

allocation, when this must be only predicted (together with the probability of surviving to future reproduction).

We know only one study that tested for within-brood differential allocation (Westneat et al. 1995). In that study, females did not feed EPY more than their-half sibs. This is perhaps not very surprising, because birds in general seem to be unable to discriminate related and unrelated nestlings in own broods (Kempenaers and Sheldon 1996). Thus the potential for females to favour young sired by more attractive males would seem to be restricted to pre-hatching period. This idea was the target of two theoretical assessments. In the first, Birkhead et al. (2000) concluded that it may be unfeasible for females to perform such subtle differential allocation. In the second, Sheldon (2000) argued that female differential allocation to EPY could be selected only if EPY are genetically superior to WPY. Although Sheldon's argument has raised some debate (Colegrave 2001; Cunningham and Russell 2001; Gil and Graves 2001; Sheldon 2001), the idea that every difference between half-sibs must be caused by genetic differences between sires is still widely accepted (see above for references).

Here we show that in the collared flycatcher (*Ficedula albicollis*) EPY are laid in eggs of the same size as WPY but are put into early eggs in the laying sequence and consequently have advantage of earlier hatching over the WPY. This might lead to improved performance of EPY as compared to WPY.

Methods

Study area and field methods

The study was conducted in the Velký Kosíř area (49°32′N, 17°04′E, 370–450 m a.s.l.), central Moravia, the Czech Republic in 2001 and 2002. In the study area, there were five plots with a total number of about 350 nest boxes. Three plots were situated in coniferous (*Picea abies*) and the other two in deciduous (*Quercus petraea*) forest. Approximately 60 pairs of collared flycatchers breed in the nest boxes each year. The collared flycatcher is a small, migratory passerine with sexually dimorphic plumage. Males are contrastingly black and white, while females are inconspicuously brownish. Females lay and solely incubate clutches of 4–8 eggs. Extra-pair young occur in a high proportion of the nests (33–43%) in various populations of the species (Sheldon and Ellegren 1999; Veen et al. 2001). Most of females in the collared flycatcher (M. Krist, personal observation) and the closely related pied flycatcher *Ficedula hypoleuca* (Potti 1998) begin incubation before clutch completion which results in partially asynchronous (sensu Stoleson and Beissinger 1995) hatching of most broods (Potti 1998).

The study area was visited daily during the breeding season. Each egg was numbered according to its position in the laying sequence with a waterproof felt-pen and measured to the nearest 0.01mm with a digital calliper on the day it was laid. We calculated egg volume using the formula: volume=$0.51\times$length$\times$width^2 (Hoyt 1979). Two measures

of width were taken in two perpendicular directions and their average was used as a measure of width. Ten to 13 days after the initiation of incubation, the eggs were taken from the nests, put into a thermo-box and then within 10 min of transfer placed into individual compartments in the incubator. Temperature in the incubator ranged between 37–39°C, humidity between 40–70%. In the meantime, plastic dummy eggs were put into the nests for females to incubate. Only one out of 38 artificial clutches was abandoned and the overall hatching success in the incubator was 80.4%. This is lower value than the natural one (about 93%; M. Krist, unpublished data). Therefore, for the analysis of relationship between hatching order and paternity only nests, in which at most two eggs failed to hatch, were used. Hatchability in these nests was 88.1%. This slightly lower hatchability could negatively effect our results if laying order interacted with paternity in determination of probability of an egg to hatch. However, this was not the case ($F_{1,66}$=0.11; P=0.741) which indicates that our results were not affected by a bit lower hatchability of eggs.

Incubator was checked for newly hatched young at least every 3h throughout the day and night. Hatching time was recorded for each chick. When hatching was not directly observed, hatching time was approximated as the midpoint between the check when the egg was hatched and the preceding check when the egg was still unhatched. Soon after hatching (mean±SD: 2.95±2.33 hours), the young were marked with nail-varnish on their claws to enable individual recognition and returned to their nest of origin. Nestlings were controlled daily and re-marked as necessary. They were ringed at about 7 days of age and blood sampled (about 25μ) by brachial venipuncture at 10–13 days. Blood samples were transferred to 1 ml of solution of Queen's lysis buffer. Dead nestlings and unhatched embryos were taken from nests and conserved in 70% ethanol.

Putative parents were caught with nest-box traps while feeding nestlings and blood sampled in the same way as the nestlings. In putative fathers size of three ornamental traits (forehead patch width, forehead patch height and wing patch) which were previously shown to be under sexual selection (Sheldon and Ellegren 1999; Qvarnström et al. 2000) was measured (to the nearest 1 mm). The white patch on the wing was measured as the maximal length of visible white on the sixth primary. Of the 27 putative fathers for which paternity of their offspring was determined, 4 were 1-year olds and, therefore, in subadult plumage; others were older and in adult plumage.

Paternity

The paternity of nestlings was determined using standard molecular methods for the collared flycatcher (e.g. Sheldon and Ellegren 1996; Veen et al. 2001; Michl et al. 2002). In short, DNA was extracted from blood or tissue samples using the phenol–chloroform method. Parentage was determined by comparing the genotypes of putative parents and nestlings at three microsatellite loci: FhU2, FhU3, and

FhU4. Offspring that mismatched their social father at one or more loci were considered to have been fathered by an extra-pair fertilisation. Putative mothers matched their offspring in all cases which indicates that intraspecific brood parasitism and mutations in loci coding for microsatellites are rare or absent in our population. The combined exclusion power of the marker set was shown to be about 96% (Sheldon and Ellegren 1996). The relatively low exclusion power indicates that extra-pair paternity went undetected in a few cases. However, there is no reason to suspect that such cases were, in respect to egg size and laying/hatching order, anything other than a random subsample of all nestlings. Therefore, our results should not be adversely affected by this slight inadequacy. Genetic fathers were not assigned to extra-pair young because the sampling of males on study plots was not complete and at the same time the probability of genotype matching between randomly chosen male and an extra-pair young is relatively high when our markers are used (Sheldon and Ellegren 1996). Thus errors could be done if the male who matched extra-pair young was treated as its genetic father.

Data analysis

To test whether the size of ornaments predicts proportion of EPY in male own nest, model in which sizes of ornaments were included as independent variables and proportion of extra-pair young in male own nest as dependent variable was fitted using PROC GENMOD (brood size=trial, number of extra-pair young in the brood=event, binomial error structure, and logit link function). Because of the marked difference between adult and subadult plumage in this species (e.g. Sheldon and Ellegren 1999), paternal age was included in the above model as the categorical factor (adult plumage vs. subadult plumage).

Eggs hatched in incubator were from clutches ranging in size from five to seven. To enable pooling of nests with various clutch sizes into single analysis, we ranked each egg into one of three categories for analysis: (1) ultimate egg, (2) penultimate egg, and (3) precedent eggs. The same three categories were used for chicks to define their position in hatching order. There is strong biological reason to categorize eggs in this way: most flycatchers begin incubation before the last or penultimate egg is laid (Potti 1998), which results in nearly synchronous hatching of precedent eggs while penultimate and especially ultimate eggs hatch later (Krist et al. 2004). As we were interested in within-clutch relations, egg size was converted to relative egg size by subtracting mean egg volume from the actual volume of the particular egg (i.e. centring). The resulting variable thus represents intraclutch egg-size variation.

To test for a bias in paternity with respect to egg size and laying/hatching order two models were fitted in GLIMMIX macro (generalised linear mixed model with binomial error, logit link, and Kenwardroger method for denominator degrees of freedom). Only nests, in which mixed paternity was detected, were used for these tests. In the first model paternity was the dependent variable, relative egg size and laying order explanatory fixed variables, and nest explanatory random variable. Initially, interaction between laying order and relative egg size was also tested. It was non-significant ($F_{1,81}=0.14$; $P=0.712$) and therefore omitted from the final model. In the second model paternity was the dependent variable, hatching order explanatory fixed and nest explanatory random variable. Laying and hatching order may be modelled either as categorical or continuous variables. Initially we launched both models. Then we chose the models with smaller AIC for presentation. In both cases models with continuous variables are presented based on the above criterion. However, results were similar and significant also when these variables were modelled as categorical factors. All statistical tests are two-tailed and were computed in SAS, version 9.1 (SAS Institute 2004).

Results

In total, 38 broods were raised in an incubator. Subsequently, both putative parents were caught at 27 nests. Extra-pair young were detected in 14 of these nests (51.9%) and comprised 40 of 165 (24.2%) offspring with determined paternity. Paternity was not assigned to five offspring due to their disappearance from the nest or tissue decay.

None of the male ornamental traits predicted proportion of EPY in male own nest after correction for the overdispersion of the data (statistics were adjusted by dispersion parameter: deviance/df) had been applied (male age: $F_{1,22}=0.08$, $P=0.782$; forehead patch height: $F_{1,22}=1.99$, $P=0.173$; forehead patch width: $F_{1,22}=0.86$, $P=0.363$; wing patch: $F_{1,22}=0.18$, $P=0.676$). Mean egg volume did not differ between nests with single and mixed paternity [mixed paternity (mean±SE): 1612.3 ± 33.7 mm^3; single paternity: 1595.5 ± 35.0mm^3; df=25; t=0.34; $P=0.734$].

In 14 nests with mixed paternity, 40 young were sired by extra-pair males, 45 young by social males and the paternity of two young remained undetermined. In these nests, proportion of EPY decreased in the laying order ($F_{1,82}=6.06$, $P=0.016$; Fig. 1) while the relative size of eggs bearing EPY (mean±SE: -5.82 ± 11.29 mm^3) and WPY (2.56 ± 10.65 mm^3) did not differ ($F_{1,82}=0.59$, $P=0.442$; Fig. 2). Effect of nest as random factor was estimated to be zero in this model. Hatching order was strongly correlated with the laying order ($r=0.71$, $N=59$, $P<0.001$), resulting in the fact that EPY hatched earlier in the hatching sequence than their half-sibs ($F_{1,48}=6.79$, $P=0.012$; Fig. 1). Effect of random factor was small also in this model (estimate±SE: 0.022 ± 0.448, $Z=0.05$, $P=0.480$).

Discussion

Our study is the first test whether females discriminate against WPY in the pre-hatching period. We did not find evidence for differential resource allocation as measured by egg size (Fig. 2). Note that the scale of y-axis is small because relative egg volume is displayed. Thus in the nest with the largest difference between mean volumes of eggs
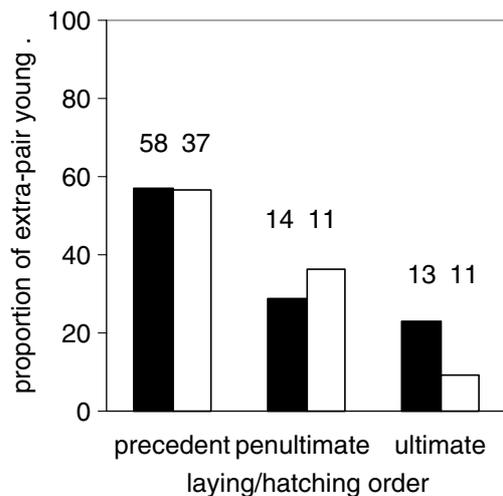
Fig. 1 Proportion of extra-pair young in particular position in laying (filled bars) and hatching (open bars) order in nests with mixed paternity. Numbers indicate sample size
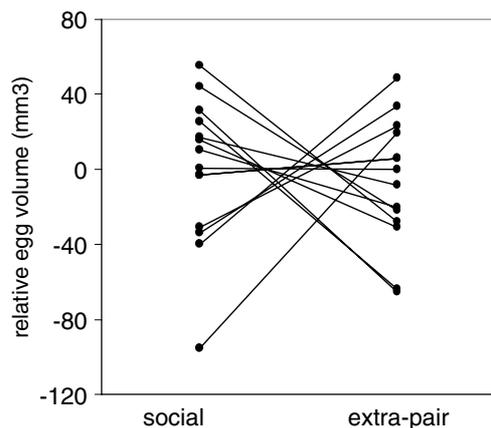
Fig. 2 Mean volumes of eggs fertilized by either the social or extra-pair males. Lines connect values in the same nest. Relative egg volume is given (see methods)

bearing WPY and EPY this difference is just 104.7 mm^2 which is likely to be too small difference to cause any effect on offspring performance (Krist et al. 2004). It should be stressed, however, that our measure of resource investment in eggs was only a crude one as it relies on size of egg as a surrogate of investment. Recently, however, it has been found that females may change also egg composition in dependence on mate attractiveness. The substances that were found to be differentially allocated at the level of breeding attempts include steroid hormones (Gil et al. 1999), carotenoids (Saino et al. 2002a), and immunoglobulins (Saino et al. 2002b). Future studies of differential within-brood allocation should therefore target also on egg composition, although it may be difficult to imagine mechanism that would enable such allocation (Birkhead et al. 2000).

In contrast to non-significant results regarding differential resource investment through egg size, we have found that in the collared flycatcher EPY are put in earlier eggs in the laying sequence than WPY and consequently also

hatch earlier in the hatching sequence than WPY. As in most other altricial birds (Magrath 1990; Stoleson and Beissinger 1995), hatching asynchrony is a strong determinant of offspring size and survival in the collared flycatcher (Krist et al. 2004) as well as in the closely related pied flycatcher (Hillström 1999). Thus, our finding suggests that EPY might have greater fitness than WPY as a result of earlier hatching.

If mothers did favour EPY through increased resource *investment,* in egg size (Cunningham and Russell 2000) or carotenoids (Saino et al. 2002a), for example, this could be in fact regarded as another evidence for genetic superiority of EPY (Sheldon 2000) as it is hard to imagine other reason for female differential investment in half-sibs. However, in the case when the advantage to EPY is driven by hatching asynchrony, the situation is markedly different. Although by proper timing of copulations and onset of incubation females might still favour EPY because of their greater reproductive value (Sheldon 2000), non-adaptive scenarios are also possible. This is because both timing of extra-pair copulations and onset of incubation are, in contrast to differential resource investment into chicks of different paternity, under diverse selective pressures. Females might, for example, pursue EPCs early in the fertile period to insure the whole clutch against social mate infertility and begin incubation before the completion of the clutch to reduce the risk of nest predation (or from about 20 another reasons, see Magrath 1990; Stoleson and Beissinger 1995). EPY then might have greater fitness as a result of early hatching regardless of their genetic quality.

Our study was not designed to discriminate whether EPY are genetically better than WPY. Consequently it is not clear whether females may gain fitness benefits by earlier hatching of EPY. The only available information for assessment of potential genetic difference between half-sibs forms our observation that the occurrence of EPY was unrelated to social mates' ornaments. This suggests that females did not perform EPCs to improve the previous choice of social partner to obtain 'good genes' for their offspring (see Møller 1992). However, this conclusion must be regarded with caution as more informative test would involve paired comparison of cuckolded fathers and their cuckolders. Unfortunately, we are unable to perform such test as genetic fathers could not be reliably assigned in this study. Moreover, EPY might be genetically better than WPY as a result of inherited complementary genes (Jennions and Petrie 2000; sensu Piálek and Albrecht 2005) in which case no relationship between paternal ornaments and offspring fitness would be expected. In fact, we have previously shown using similar sample of nestlings that EPY tended to be in better condition than WPY even after hatching asynchrony is controlled for (Krist et al. 2004), which might suggest that EPY are really genetically better than WPY. So, the question about adaptive value of earlier hatching of EPY for females remains unresolved for the present.

Nevertheless, the possibility that earlier hatching of EPY is only by-product of selection pressures operating in other contexts suggests that fitness difference between half-sibs should not be used to infer paternal genetic

effects, unless artificial fertilization is performed in vitro (Barber and Arnott 2000; Sheldon et al. 2003; Neff 2004; Evans et al. 2004) or hatching asynchrony is controlled for (Johnsen et al. 2000). In fact, as hatching asynchrony is process operating only within broods, comparison of half-sibs to infer paternal genetic effects may be less reliable approach than that using correlation between paternal trait (that is involved in female mate choice) and offspring fitness, provided that fathers are disassociated from their offspring (see Norris 1993). This is because in the latter method estimate of paternal effect is based mainly on offspring from single-paternity broods (as far as these are more frequent than those with mixed paternity). In this subsample of broods hatching asynchrony does not bias estimate of paternal genetic effects.

Clearly more studies are needed to test for non-genetic differences between naturally occurring half-sibs to see how general this phenomenon is. Remarkably, however, our results are in accord with two another published pieces of evidence. First, in an experimental study on the same species it was found that extra-pair copulations are, in contrast to within-pair ones, avoided late in the female fertile period (Michl et al. 2002). Such pattern could lead to under-representation of EPY in late eggs as was found in our study. Second, EPY were placed in early eggs also in another passerine, the house sparrow (Cordero et al. 1999), in which hatching is like in the collared flycatcher partially asynchronous (see Veiga 1992). Taken together, these results suggest that non-genetically based difference in performance of naturaly occurring half-sibs may be more widespread than was previously thought.

References

Balzer AL, Williams TD (1998) Do female zebra finches vary primary reproductive effort in relation to mate attractiveness? Behaviour 135:297–309

Barber I, Arnott SA (2000) Split-clutch IVF: a technique to examine indirect fitness consequences of mate preferences in sticklebacks. Behaviour 137:1129–1140

Birkhead T, Schwabl H, Burke T (2000) Testosterone and maternal effects: integrating mechanisms and function. Trends Ecol Evol 15:86–87

Burley N (1986) Sexual selection for aesthetic traits in species with biparental care. Am Nat 127:415–445

Charmantier A, Blondel J, Perret P, Lambrechts MM (2004) Do extra-pair paternities provide genetic benefits for female blue tits *Parus caeruleus*? J Avian Biol 35:524–532

Colegrave N (2001) Differential allocation and good genes: comment from Colegrave. Trends Ecol Evol 16:22–23

Cordero PJ, Wetton JH, Parkin DT (1999) Within-clutch patterns of egg viability and paternity in the House Sparrow. J Avian Biol 30:103–107

Cunningham EJA, Russell AF (2000) Egg investment is influenced by male attractiveness in the mallard. Nature 404:74–77

Cunningham EJA, Russell AF (2001) Differential allocation and 'good genes': comment from Cunningham and Russell. Trends Ecol Evol 16:21

Evans JP, Kelley JL, Bisazza A, Finazzo E, Pilastro A (2004) Sire attractiveness influences offspring performance in guppies. Proc R Soc Lond B 271:2035–2042

Gil D, Graves J, Hazon N, Wells A (1999) Male attractiveness and differential testosterone investment in zebra finch eggs. Science 286:126–128

Gil D, Graves J (2001) Differential allocation and good genes: comment from Gil and Graves. Trends Ecol Evol 16:21–22

Griffith SC, Owens IPF, Thuman KA (2002) Extra pair paternity in birds: a review of interspecific variation and adaptive function. Mol Ecol 11:2195–2212

Hillström L (1999) Variation in egg mass in the Pied Flycatcher, *Ficedula hypoleuca*: an experimental test of the brood survival and brood reduction hypotheses. Evol Ecol Res 1:753–768

Hoyt DF (1979) Practical methods of estimating volume and fresh weight of birds' eggs. Auk 96:73–77

Jennions MD, Petrie M (2000) Why do females mate multiply? A review of the genetic benefits. Biol Rev 75:21–64

Johnsen A, Andersen V, Sunding C, Lifjeld JT (2000) Female bluethroats enhance offspring immunocompetence through extra-pair copulations. Nature 406:296–299

Johnsen A, Delhey K, Andersson S, Kempenaers B (2003) Plumage colour in nestling blue tits: sexual dichromatism, condition dependence and genetic effects. Proc R Soc Lond B 270:1263–1270

Kempenaers B, Sheldon BC (1996) Why do male birds not discriminate between their own and extra-pair offspring? Anim Behav 51:1165–1173

Kempenaers B, Verheyen GR, Dhondt AA (1997) Extrapair paternity in the blue tit (*Parus caeruleus*): female choice, male characteristics, and offspring quality. Behav Ecol 8:481–492

Kleven O, Lifjeld JT (2004) Extrapair paternity and offspring immunocompetence in the reed bunting, *Emberiza schoeniclus*. Anim Behav 68:283–289

Kolm N (2001) Females produce larger eggs for large males in a paternal mouthbrooding fish. Proc R Soc Lond B 268:2229–2234

Krist M, Remeš V, Uvírová L, Nádvorník P, Bureš S (2004) Egg size and offspring performance in the collared flycatcher (*Ficedula albicollis*): a within clutch approach. Oecologia 140:52–60

Krokene C, Rigstad K, Dale M, Lifjeld JT (1998) The function of extrapair paternity in blue tits and great tits: good genes or fertility insurance? Behav Ecol 9:649–656

Lubjuhn T, Strohbach S, Brun J, Gerken T, Epplen JT (1999) Extra-pair paternity in great tits (*Parus major*): a long term study. Behaviour 136:1157–1172

Magrath RD (1990) Hatching asynchrony in altricial birds. Biol Rev 65:587–622

Michl G, Török J, Griffith SC, Sheldon BC (2002) Experimental analysis of sperm competition mechanisms in a wild bird population. Proc Natl Acad Sci USA 99:5466–5470

Møller AP (1992) Frequency of female copulations with multiple males and sexual selection. Am Nat 139:1089–1101

Møller AP, Ninni P (1998) Sperm competition and sexual selection: a meta-analysis of paternity studies of birds. Behav Ecol Sociobiol 43:345–358

Neff BD (2004) Increased performance of offspring sired by parasitic males in bluegill sunfish. Behav Ecol 15:327–331

Norris K (1993) Heritable variation in a plumage indicator of viability in male great tits *Parus major*. Nature 362:537–539

Piálek J, Albrecht T (2005) Choosing mates: complementary versus compatible genes. Trends Ecol Evol 20:63

Potti J (1998) Variation in the onset of incubation in the pied flycatcher (*Ficedula hypoleuca*): fitness consequences and constraints. J Zool 245:335–344

Qvarnström A, Pärt T, Sheldon BC (2000) Adaptive plasticity in mate preference linked to differences in reproductive effort. Nature 405:344–347

Rutstein AN, Gilbert L, Slater PJB, Graves JA (2004) Mate attractiveness and primary resource allocation in the zebra finch. Anim Behav 68:1087–1094

Saino N, Bertacche V, Ferrari RP, Martinelli R, Møller AP, Stradi R (2002a) Carotenoid concentration in barn swallow eggs is influenced by laying order, maternal infection and paternal ornamentation. Proc R Soc Lond B 269:1729–1733

Saino N, Ferrari RP, Martinelli R, Romano M, Rubolini D, Møller AP (2002b) Early maternal effects mediated by immunity depend on sexual ornamentation of the male partner. Proc R Soc Lond B 269:1005–1009

SAS Institute (2004) SAS/STAT 9.1 User's guide. SAS Institute Inc., Cary

Schmoll T, Dietrich V, Winkel W, Epplen JT, Lubjuhn T (2003) Long-term fitness consequences of female extra-pair matings in a socially monogamous passerine. Proc R Soc Lond B 270:259–264

Sheldon BC, Ellegren H (1996) Offspring sex and paternity in the collared flycatcher. Proc R Soc Lond B 263:1017–1021

Sheldon BC, Merilä J, Qvarnström A, Gustafsson L, Ellegren H (1997) Paternal genetic contribution to offspring condition predicted by size of male secondary sexual character. Proc R Soc Lond B 264:297–302

Sheldon BC, Ellegren H (1999) Sexual selection resulting from extrapair paternity in collared flycatchers. Anim Behav 57:285–298

Sheldon BC (2000) Differential allocation: tests, mechanisms and implications. Trends Ecol Evol 15:397–402

Sheldon BC (2001) Differential allocation and good genes: response from Sheldon. Trends Ecol Evol 16:23

Sheldon BC, Arponen H, Laurila A, Crochet PA, Merilä J (2003) Sire coloration influences offspring survival under predation risk in the moorfrog. J Evol Biol 16:1288–1295

Stoleson SH, Beissinger SR (1995) Hatching asynchrony and the onset of incubation in birds, revisited. In: Power DM (ed) Current ornithology 12. Plenum Press, New York, pp 191–270

Strohbach S, Curio E, Bathen A, Epplen JT, Lubjuhn T (1998) Extrapair paternity in the great tit (*Parus major*): a test of the "good genes" hypothesis. Behav Ecol 9:388–396

Veen T, Borge T, Griffith SC, Saetre GP, Bureš S, Gustafsson L, Sheldon BC (2001) Hybridization and adaptive mate choice in flycatchers. Nature 411:45–50

Veiga JP(1992) Hatching asynchrony in the house sparrow: a test of the egg-viability hypothesis. Am Nat 139:669–675

Westneat DF, Clark AB, Rambo KC (1995) Within-brood patterns of paternity and paternal behavior in red-winged blackbirds. Behav Ecol Sociobiol 37:349–356

Whittingham L, Dunn PO (2001) Survival of extrapair and within-pair young in tree swallows. Behav Ecol 12:496–500

Článek 5: Remeš, V. & Krist, M. (2005) Nest design and the abundance of parasitic *Protocalliphora* blow flies in two hole-nesting passerines. *Ecoscience*, 12, 549-553.

Nest design and the abundance of parasitic *Protocalliphora* blow flies in two hole-nesting passerines[1]

Vladimír REMEŠ[2], Laboratory of Ornithology, Palacky University, Tr. Svobody 26, CZ-771 46 Olomouc,
  Czech Republic, e-mail: remes@prfnw.upol.cz

Miloš KRIST, Laboratory of Ornithology, Palacky University, Tr. Svobody 26, CZ-771 46 Olomouc, Czech
  Republic, and Museum of Natural History, nam. Republiky 5, CZ-771 73 Olomouc, Czech Republic.

Abstract: Ectoparasites dwelling in bird nests regularly reduce reproductive success and condition of breeding birds. Thus, establishing the factors that determine the abundance of ectoparasites is important for better understanding of reproductive trade-offs and life history evolution in birds. A recent hypothesis states that interspecific differences in the abundance of ectoparasites may be caused by nest composition. For example, great tits (*Parus major*) have nests made of mosses and fur, whereas *Ficedula* flycatchers have nests made of grasses, bast, and bark, and tits are more infested by nest-dwelling ectoparasites than flycatchers. We swapped nests between pairs of great tits and collared flycatchers (*F. albicollis*) during egg-laying or early incubation and counted parasitic *Protocalliphora* blow flies at the end of breeding to test this hypothesis experimentally. We controlled statistically for habitat (oak *versus* spruce forest), brood size, season, year, and mean nestling weight before fledging. We found a significant effect of bird species (tit > flycatcher), habitat (oak > spruce), and year. There was no effect of nest type. Consequently, the hypothesis ascribing the different abundance of ectoparasites in great tits and collared flycatchers to different nest composition was not supported by our study.

Keywords: blow fly, ectoparasites, *Ficedula*, nest design, *Parus*, *Protocalliphora*.

Résumé : Les ectoparasites qui infestent les nids réduisent souvent le succès reproducteur et la condition physique des oiseaux nicheurs. Il est donc important d'identifier les facteurs qui influencent leur abondance pour mieux comprendre les caractéristiques de la reproduction et l'évolution du cycle vital chez les oiseaux. Selon une hypothèse récente, les différences interspécifiques dans l'abondance des ectoparasites pourraient être associées aux matériaux de construction des nids. Par exemple, les nids de la mésange charbonnière (*Parus major*) sont fabriqués de mousses et de poils alors que ceux des gobemouches (*Ficedula* spp.) sont faits de graminées, de liber et d'écorce. Or, on sait que les mésanges sont plus infestées par des ectoparasites que les gobemouches. Nous avons échangé les nids construits par des couples de mésanges charbonnières avec ceux construits par des gobemouches à collier (*F. albicollis*) pendant la ponte ou au début de l'incubation. Nous avons par la suite compté les mouches du genre *Protocalliphora* à la fin de la nidification pour tester l'hypothèse mentionnée plus haut. Nous avons pu contrôler de façon statistique l'habitat (forêt de chênes ou d'épicéas), la taille de la couvée, la saison, l'année et le poids moyen des jeunes avant l'envol. Nous avons trouvé un effet significatif pour l'espèce d'oiseau (mésange > gobemouche), l'habitat (chênes > épicéas) et l'année. Le type de nid n'a eu pour sa part aucun effet sur les ectoparasites. En conséquence, l'hypothèse d'un lien entre l'abondance d'ectoparasites et les matériaux de construction des nids n'est pas supporté par les résultats de cette étude.

Mots-clés : design des nids, ectoparasites, *Ficedula*, mouches, *Parus*, *Protocalliphora*.

Nomenclature: Sabrosky, Bennett & Whitworth, 1989; Cramp & Perrins, 1993.

Introduction

Ectoparasites dwelling in bird nests and feeding on the blood of nestlings and adults are an important group of parasites. In hole-nesting birds they include fleas (Siphonaptera), flies (Diptera), and mites (Acarina). Ectoparasites can cause lowered breeding performance and nest desertion in adults (Oppliger, Richner & Christe, 1994), negatively affect growth and condition of nestlings (Eeva, Lehikoinen & Nurmi, 1994; Merino & Potti, 1995; Puchala, 2004; review in Møller, Allander & Dufva, 1990), and reduce lifetime reproductive success of hosts (Fitze, Tschirren & Richner, 2004). Moreover, they also serve as vectors of internal

parasites and bacterial and viral diseases (Bowman *et al.,* 1997). Thus, knowledge of the factors that the determine abundance of these ectoparasites in nests is critical for better understanding of reproductive trade-offs and life history evolution in birds (Clayton & Moore, 1997).

Abundance of nest-dwelling ectoparasites varies significantly among individuals within a given bird host species, but even bigger variation in both prevalence and parasite abundance is found between sympatric host species (Bennett & Whitworth, 1992; Whitworth & Bennett, 1992). One of the hypotheses suggested to explain interspecific differences in ectoparasite load posits that the differences are caused by differences in nest design (Bauchau, 1998), such as differences in details of nest construction and nest composition (Hansell, 2000). Fresh plant material in the

nest, for example, may have negative effects on parasite abundance (Clark & Mason, 1985; Petit *et al.,* 2002; but see Dawson, 2004). Nest design may thus affect (1) demography of ectoparasite populations within nests (larval mortality, competition, growth), because demography is driven by various parameters of the living environment (*e.g.,* humidity; Heeb, Kölliker & Richner, 2000) and the living environment depends on nest composition, and/or (2) attractiveness of the particular nest type for laying/dispersing to females of parasites, which may cue on specific features of the nest.

In this study we tested the "nest design" hypothesis on great tits (*Parus major*) and collared flycatchers (*Ficedula albicollis*). The two species build nests of very different composition: while tit nests are composed of moss and feather/hair, flycatcher nests consist of dry grass, bast, and pieces of bark (Cramp & Perrins, 1993). At the same time, the abundance of larvae of parasitic flies (*Protocalliphora* spp.) is regularly higher in nests of the great tit than in *Ficedula* flycatchers (Bauchau, 1998). We experimentally switched nests between great tit and collared flycatcher pairs and followed the effects of this treatment on the abundance of parasitic flies in nests. In this way we were able to separate independent effects of nest design and species. We also statistically controlled for a number of other factors, including habitat, year, season, brood size, and mean nestling weight before fledging.

In line with the "nest design" hypothesis we hypothesized that the lower number of ectoparasites in *Ficedula* flycatchers is caused by the composition of their nests (presence of bast and bark with toxic secondary compounds; Pearce, 1996) and, thus, predicted that there would be fewer parasitic flies in the *Ficedula* nests regardless of the host species actually dwelling in the nest.

Methods

Great tit and collared flycatcher are small, hole-nesting, insectivorous passerines breeding widely in various types of woodlands. They readily accept nest-boxes for breeding. Great tits are year-round residents, whereas collared flycatchers are long-distance migrants wintering in Africa. These two species differ in brood size [great tit: median 11 (range 7–14), collared flycatcher: 6 (4–7), timing of breeding (great tits start breeding at mid-April, collared flycatchers at the end of April), body mass (great tit: *ca* 18 g; collared flycatcher: *ca* 14 g, V. Remeš & M. Krist, unpubl. data)]. Otherwise their breeding ecology is similar.

This study was conducted in 2002-2003 in the Velký Kosíř area in the eastern Czech Republic (49° 32' N, 17° 04' E, 300-450 m asl). We studied great tits and collared flycatchers on six nest-box plots, of which three were placed in spruce (*Picea abies*) and the other three in an oak (*Quercus* spp.) forest, in both cases interspersed with birch (*Betula pendula*) and pine (*Pinus silvestris*). Each plot had 50-90 nest-boxes. In early spring (before tits started nest building), nest-boxes were checked and cleaned (old nests were removed). From mid-April to mid-June, as a part of a larger study, we followed basic breeding biology of both species.

To be able to separate independent effects of nest design and species *per se* on the abundance of nest-dwell-

ing flies, we switched nests between great tit and collared flycatcher pairs. As a control treatment, we swapped nests between pairs of the same species. Thus, nests of all hosts were swapped. Two experimental (2 × flycatcher–tit) and two control (tit–tit, flycatcher–flycatcher) manipulations were made on the same day each time. We strove to make manipulations as early as possible in the breeding cycle. However, tits breed earlier than flycatchers. Consequently, manipulations on flycatcher nests were made during egg laying, whereas in tits we made the manipulations up to the fourth day after the clutch was complete (mean ± SD number of days from laying of the first egg to nest manipulation was 2.5 ± 2.97 in flycatchers and 12.5 ± 1.86 in tits). We increased our sample for flycatchers by also using some nests of tits in late incubation. In these experimental pairs, we followed nest-type effects only in the flycatcher, not in the great tit.

In the week following fledging, we collected nests and placed them into plastic bags that were sealed so that no fly could escape. We collected only nests from which at least one young had fledged. Within four weeks of collection, we opened the bags, took the nests to pieces, and counted the number of larvae, pupae, and adult flies (if they had emerged from pupae in the meantime). Time between fledging of young and collection of nests was not the same for all the nests. However, collection of nests within one week of fledging is a standard procedure in ectoparasite research (Eeva, Lehikoinen & Nurmi, 1994; Birdblowfly.com). More importantly, even when some flies had dispersed immediately after fledging, before a nest was collected, they could be counted by counting empty pupae, which are very conspicuous and cannot be overlooked.

Two species of parasitic flies were identified in the nests: *Protocalliphora azurea* and *P. falcozi*. Both species are regularly found in the nests of European cavity nesters (Hurtrez-Boussès *et al.*, 1997; Wesołowski, 2001). All *Protocalliphora* species overwinter as adults and do not lay eggs in host nests until young birds hatch (Gold & Dahlsten, 1989; Sabrosky, Bennett & Whitworth, 1989). Thus, our experimental procedure of switching the nests during egg laying or early incubation was sufficient to separate independent effects of nest type. The two species of blow flies were lumped together for further analyses for two reasons (see also Hurtrez-Boussès et al., 1997). First, all *Protocalliphora* flies (except *P. braueri*; Eastman, Johnson & Kermott, 1989) are intermittent feeders that feed on the blood of nestlings and in the meantime dwell in the nest substrate (Sabrosky, Bennett & Whitworth, 1989). Moreover, *P. azurea* and *P. falcozi* are of similar body size (9-11 mm and 8-10 mm, respectively; Grunin, 1970), so their effects on hosts can be expected to be similar. Second, it was not possible to identify all the flies to species because not all individuals emerged from pupae and to our best knowledge only adults can be identified in European *Protocalliphora* flies.

When analyzing the abundance of flies, we first fit a generalized linear model with Poisson error distribution and log link, which is usually suitable for count data. However, our data were strongly overdispersed (deviance/df = 23.26),

which is common in parasitology (Wilson & Grenfell, 1997). Thus, we used negative binomial error distribution and log link, which led to a reasonable dispersion index of data (deviance/df = 1.50). All these analyses were done in PROC GENMOD in SAS (SAS Institute, 2000). Initially, we fit a full model with the following explanatory effects: nest type (tit *versus* flycatcher nest), species (tit *versus* flycatcher), habitat (spruce *versus* oak forest), year (2002 *versus* 2003), brood size (number of hatched nestlings), season (Julian hatching date), and mean nestling weight before fledging (in grams, day 13 after hatching in flycatcher, day 15 in tit); we also included all two-way interactions between nest type and all other factors. The final model was selected by backward elimination of non-significant terms, except for the two main factors of interest (nest type and species), which were retained in the model regardless of their significance. Hatching date, brood size, and mean nestling weight were standardized by subtracting the value of a given nest from the mean of a given species (*i.e.,* the values were standardized within species). However, the results were the same with non-standardized values. Test statistics (χ^2-values) and *P*-values reported in Results for non-significant terms are from the backward elimination procedure just before the particular term (being the least significant) was removed from the model. Values for significant factors and/or factors of interest (*i.e.,* nest type and species) are from the final model.

The rationale for the inclusion of the above-mentioned variables was as follows. Nest type and species were the main factors of interest. Other factors were included as covariates to reduce unexplained variation and thus the power of the main test. Season and habitat could affect flying activity of the flies (through temperature, humidity, etc.). Brood size and mean nestling weight could affect survival and growth of larvae (by determining the amount of blood available for feeding). Alternatively, the latter factors could affect oviposition behaviour of fly females.

Results

In total, 13 experimental and 17 control tit pairs and 20 experimental and 19 control flycatcher pairs were available for the analyses. Sample sizes differ between experimental and control treatments because some nests were abandoned or depredated. There was a significant effect of species (χ^2 = 5.54, *P* = 0.019), habitat (χ^2 = 9.00, *P* = 0.003), and year (χ^2 = 5.99, *P* = 0.014) on the abundance of parasitic *Protocalliphora* flies in nests (Figure 1). Neither nest type (χ^2 = 0.28, *P* = 0.595, Figure 1) nor the interaction of nest type with species (χ^2 = 1.29, *P* = 0.256, Figure 2) had a significant influence. Similarly, there was no significant effect of brood size (χ^2 = 0.01, *P* = 0.905), mean nestling weight (χ^2 = 0.85, *P* = 0.358), season (χ^2 = 0.94, *P* = 0.332), or any interaction of nest type with other factors (all χ^2-values < 0.63, all *P*-values > 0.431).

Discussion

We experimentally tested the hypothesis that nest design is responsible for interspecific differences in ectoparasite infestation in two species of hole-nesting passerines,
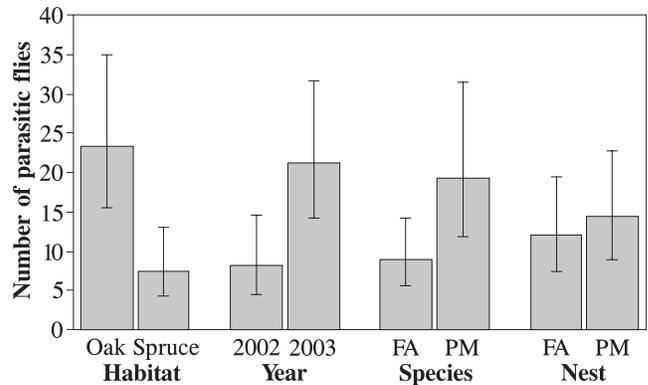

FIGURE 1. Number (least squares means ± 95% confidence limits) of parasitic *Protocalliphora* flies in the nests of great tit and collared flycatcher in relation to habitat, year, species, and nest type. FA = collared flycatcher, PM = great tit. Statistical tests are reported in Results.
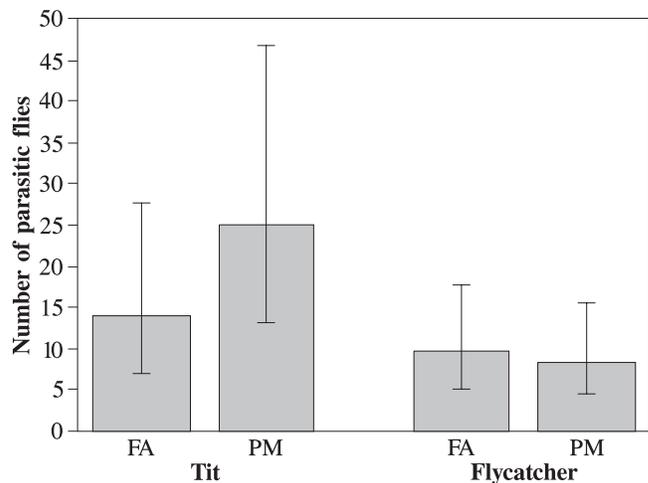

FIGURE 2. Number (least squares means ± 95% confidence limits) of parasitic *Protocalliphora* flies in the great tit (Tit) and collared flycatcher (Flycatcher) according to nest type (FA = collared flycatcher, PM = great tit). Statistical tests are reported in Results.

the great tit and the collared flycatcher (Bauchau, 1998). There was no influence of nest type on the intensity of infestation by parasitic *Protocalliphora* flies. We did, however, find a significant effect of species, habitat, and year.

Females of parasitic *Protocalliphora* flies overwinter as adults and lay their eggs in nests during the nestling phase of the host breeding cycle (Gold & Dahlsten, 1989; Sabrosky, Bennett & Whitworth, 1989). Larvae hatch within 2 d, feed on blood of nestlings while dwelling in the nest substrate, and after one to two weeks of growth pupate to complete the life cycle (Sabrosky, Bennett & Whitworth, 1989; Bennett & Whitworth, 1991). Thus, in these flies both active choice of a certain nest type by females and demography of larvae within host nests (competition, growth rate, mortality) may play a significant role in determining their abundance in relation to nest type.

Since there was no effect of nest type on fly abundance, it is likely that neither of the two possible processes played any role: fly females did not cue on nest composition when selecting their oviposition site, and demographic processes among larvae did not influence their abundance in relation

to nest type. Alternatively, these two processes may have counteracted each other in determining the abundance of larvae: flies may have selected the type of nest in which their larvae had worse performance; however, such a maladaptive habitat choice seems unlikely to evolve (but see Remeš, 2000). Nevertheless, the absence of any effect of nest type on the abundance of flies is rather puzzling. It is, for example, known that pine bark and bast contain toxic secondary compounds with a strong potential to negatively affect ectoparasites (Pearce, 1996; Bauchau, 1998). As flycatchers use this material to build their nests, this should have led to higher ectoparasite abundance in tit nests.

In contrast to nest type, there was a strong effect of species *per se* on the abundance of parasitic flies: tits were more intensely infested regardless of nest type. Tits and flycatchers differ in brood size, timing of breeding, nestling weight, and the nestling period duration (see Methods), which could in principle cause the difference between species. For example, the greater number of young in the great tit and the longer time that great tit young remain in the nest could mean that more food is available for parasitic larvae, which could lead to their higher abundance. However, those factors that we measured and included in the models had no influence on the abundance of flies within species as evidenced by their non-significance when used as standardized factors in the analysis (see Results). Thus, it seems unlikely that any of these is the causal factor behind the effect of species. This effect may have several more subtle explanations. First, adult flycatchers may be more capable of behavioural anti-parasite defences, for example in the form of nest cleaning by catching laying females and/or parasitic larvae (see also Hurtrez-Boussès *et al.,* 2000; Tripet, Glaser & Richner, 2002). Second, flycatcher nestlings may be more resistant to parasitism and fly larvae suffer greater mortality because of more effective immune defence. Third, the preference of laying females for certain bird species may significantly alter patterns of ectoparasite infestation. The preference for certain species of hosts (here great tits) may have arisen, for example, from better performance of fly larvae on their nestlings (for whatever reason, *e.g.,* different skin thickness, resistance to parasitism, length of the nestling period, etc.). Our study was not suited to revealing the proximate mechanism of the effect of species. However, given the strong effect of species *per se*, it would be interesting to find out which mechanism is responsible.

Habitat was an important determinant of the abundance of flies: they were more abundant in the oak forest than in the spruce forest. All three oak plots were situated on warmer and drier southern slopes, whereas spruce plots were situated either on the top of the hill (two of them) or on the colder and more humid northern slope (one plot). Although known effects of weather (temperature and ambient humidity) on fly abundance are in accord with this difference (Merino & Potti, 1996), many uncontrolled factors differing between the two forest types may have had an influence.

In summary, there was no effect of nest type (nest of great tit *versus* collared flycatcher) on the abundance of nest-dwelling parasitic *Protocalliphora* flies. Thus, the hypothesis ascribing different levels of ectoparasite infestation between the great tit and *Ficedula* flycatchers to nest design (Bauchau, 1998) was not supported by our experimental study.

Literature cited

Bauchau, V., 1998. Comparison of parasitism level in two sympatric passerines: The pied flycatcher and the great tit. Écoscience, 5: 164–171.

Bennett, G. F. & T. L. Whitworth, 1991. Studies on the life history of some species of *Protocalliphora* (Diptera: Calliphoridae). Canadian Journal of Zoology, 69: 2048-2058.

Bennett, G. F. & T. L. Whitworth, 1992. Host, nest, and ecological relationships of species of *Protocalliphora* (Diptera: Calliphoridae). Canadian Journal of Zoology, 70: 51-61.

Birdblowfly.com. [Online] URL: http://birdblowfly.com/, accessed 7 March 2005.

Bowman, A. S., L. B. Coons, G. R. Needham & J. R. Sauer, 1997. Tick saliva: Recent advances and implications for vector competence. Medical and Veterinary Entomology, 11: 277-285.

Clark, L. & J. R. Mason, 1985. Use of nest material as insecticidal and anti-pathogenic agents by the European starling. Oecologia, 67: 169-176.

Clayton, D. H. & J. Moore, 1997. Host-Parasite Evolution. Oxford University Press, Oxford.

Cramp, S. & C. M. Perrins, 1993. The Birds of the Western Palearctic. Volume VII. Oxford University Press, Oxford.

Dawson, R. D., 2004. Does fresh vegetation protect avian nests from ectoparasites? An experiment with tree swallows. Canadian Journal of Zoology, 82: 1005-1010.

Eastman, M. D., L. S. Johnson & L. H. Kermott, 1989. Ectoparasitism of nestling house wrens, *Troglodytes aedon*, by larvae of the blow fly *Protocalliphora braueri* (Diptera: Calliphoridae). Canadian Journal of Zoology, 67: 2358-2362.

Eeva, T., E. Lehikoinen & J. Nurmi, 1994. Effects of ectoparasites on breeding success of great tits (*Parus major*) and Pied Flycatchers (*Ficedula hypoleuca*) in an air pollution gradient. Canadian Journal of Zoology, 72: 624-634.

Fitze, P. S., B. Tschirren & H. Richner, 2004. Life history and fitness consequences of ectoparasites. Journal of Animal Ecology, 73: 216-226.

Gold, C. S. & D. L. Dahlsten, 1989. Prevalence, habitat selection, and biology of *Protocalliphora* (Diptera: Calliphoridae) found in nests of mountain and chestnut-backed chickadees in California. Hilgardia, 57: 1-19.

Grunin, K. Y., 1970. Family Calliphoridae. Pages 607-624 *in* G. Y. Bey-Bienko (ed.). A Key to the Invertebrates of the European Part of USSR V, Part 2. Nauka, Leningrad (in Russian).

Hansell, M., 2000. Bird Nests and Construction Behaviour. Cambridge University Press, Cambridge.

Heeb, P., M. Kölliker & H. Richner, 2000. Bird-ectoparasite interactions, nest humidity, and ectoparasite community structure. Ecology, 81: 958-968.

Hurtrez-Boussès, S., P. Perret, F. Renaud & J. Blondel, 1997. High blowfly parasitic loads affect breeding success in a Mediterranean population of blue tits. Oecologia, 112: 514-517.

Hurtrez-Boussès, S., F. Renaud, J. Blondel, P. Perret & M. J. Galan, 2000. Effects of ectoparasites of young on parents' behavior in a Mediterranean population of blue tits. Journal of Avian Biology, 31: 266-269.

Merino, S. & J. Potti, 1995. Mites and blowflies decrease growth and survival in nestling pied flycatchers. Oikos, 73: 95-103.

Merino, S. & J. Potti, 1996. Weather dependent effects of nest ectoparasites on their bird hosts. Ecography, 19: 107-113.

Møller, A. P., K. Allander & R. Dufva, 1990. Fitness effects of parasites on passerine birds: A review. Pages 269-280 *in* J. Blondel, A. Gosler, J. D. Lebreton & R. H. McCleery (eds.). Population Biology of Passerine Birds: An Integrated Approach. Springer-Verlag, Berlin.

Oppliger, A., H. Richner & P. Christe, 1994. Effect of an ectoparasite on lay date, nest-site choice, and hatching success in the great tit (*Parus major*). Behavioural Ecology, 5: 130-134.

Pearce, R. B., 1996. Antimicrobial defences in the wood of living trees. New Phytologist, 132: 203-233.

Petit, C., M. Hossaert-McKay, P. Perret, J. Blondel & M. M. Lambrechts, 2002. Blue tits use selected plants and olfaction to maintain an aromatic environment for nestlings. Ecology Letters, 5: 585-589.

Puchala, P., 2004. Detrimental effects of larval blow flies (*Protocalliphora azurea*) on nestlings and breeding success of tree sparrows (*Passer montanus*). Canadian Journal of Zoology, 82: 1285-1290.

Remeš, V., 2000. How can maladaptive habitat choice generate source-sink population dynamics? Oikos 91: 579-582.

Sabrosky, C. W., G. F. Bennett & T. L. Whitworth, 1989. Bird blow flies (*Protocalliphora*) in North America (Diptera: Calliphoridae) with notes on the Palearctic species. Smithsonian Institution Press, Washington, DC.

SAS Institute, 2000. SAS Online Doc, version 8. SAS Institute Inc., Cary, North Carolina.

Tripet, F., M. Glaser & H. Richner, 2002. Behavioural responses to ectoparasites: Time-budget adjustments and what matters to blue tits *Parus caeruleus* infested by fleas. Ibis, 144: 461-469.

Wesołowski, T., 2001. Host-parasite interactions in natural holes: Marsh tits (*Parus palustris*) and blow flies (*Protocalliphora falcozi*). Journal of Zoology, 255: 495-503.

Whitworth, T. L. & G. F. Bennett, 1992. Pathogenicity of larval *Protocalliphora* (Diptera: Calliphoridae) parasitizing nestling birds. Canadian Journal of Zoology, 70: 2184-2191.

Wilson, T. & B. T. Grenfell, 1997. Generalized linear modelling for parasitologists. Parasitology Today, 13: 33-38.

Článek 6: Krist, M. (2006) Should mothers in poor condition invest more in daughter than in son? *Ethology Ecology & Evolution*, 18, 241-246.

Should mothers in poor condition invest more in daughter than in son?

Miloš Krist [1,2,3]

[1] *Museum of Natural History, nám. Republiky 5, 771 73 Olomouc, Czech Republic*
[2] *Department of Zoology, Faculty of Science, Palacký University, Tř. Svobody 26, 771 46 Olomouc, Czech Republic*

Although Trivers & Willard (1973) model is among the most influential ideas in evolutionary ecology, some of its predictions are subject of controversies. On one hand the prediction that mothers in poor condition should produce more daughters while those in good condition should produce sons is generally accepted and reasonably well supported. On the other hand, it is not clear how maternal condition should affect subsequent investment into the two sexes. Using graphical model and simple simulation I show here that mothers in poor condition should invest more in daughters. I point out further that the continuing confusion around the predictions of the model is caused partly by imprecise formulation in the original Trivers & Willard (1973) paper and partly by poor discrimination of researchers between current value of the progeny versus the marginal returns from additional investment.

KEY WORDS: marginal returns, parental investment, sex allocation, sex ratio, simulation, Trivers-Willard model.

One of the most influential ideas in evolutionary biology is the Trivers-Willard model (TWM) which claims that optimal parental investment into male and female offspring depends on maternal condition (Trivers & Willard 1973). When in good condition, mothers should produce the sex with greater variance in reproductive success (usually male), while mothers in poor condition should produce female offspring (Trivers & Willard 1973). Although major part of their seminal paper is concerned exclusively with sex ratio adjustment, in their concluding paragraph authors extended the idea to differential investment into the sexes. Despite long-lasting popularity of TWM, many empirical studies looking for a TW effect conducted tests that were not covered by predictions of the original model (Cameron & Linklater 2002, Carranza 2002). This might be caused by the fact that TWM was verbal argument and thus various semantic interpretations of its assumptions and predictions are possible (Carranza 2002). This problem relates mainly to the more general prediction of TWM concerning differential investment in the sexes. In fact, although Carranza (2002) and Cameron & Linklater (2002) agree on the point that TWM was frequently misin-

[3] (Tel.: +420-585515128, Fax: +420-585222743; E-mails: krist@prfnw.upol.cz, krist@vmo.cz).

terpreted, they disagree on whether females in poor condition should invest more in sons (CARRANZA 2002) or daughters (CAMERON & LINKLATER 2002). Here, using graphical model similar to that used by CARRANZA (2002) and CAMERON & LINKLATER (2002) and simple simulation, I show that females in poor condition should invest more resources into the sex with the less variable reproductive value (usually females, but see LEIMAR 1996, HEWISON et al. 2005, KRÜGER et al. 2005). I also suggest that the confusion about differential investment in the two sexes may arise from imprecise formulation in TRIVERS & WILLARD (1973) paper.

    The conditions required for TW effect to occur are graphically displayed in Fig. 1a. (1) The reproductive value of current offspring of either sex (raw fitness return) depends positively on the amount of resources devoted to it (this statement comprises the first two assumptions of TWM; I will use the term resources rather than maternal condition which was used by TRIVERS & WILLARD (1973) since the former is more general). (2) Fitness-return curves for female and male offspring cross which represents the third assumption of TWM. To make the model more
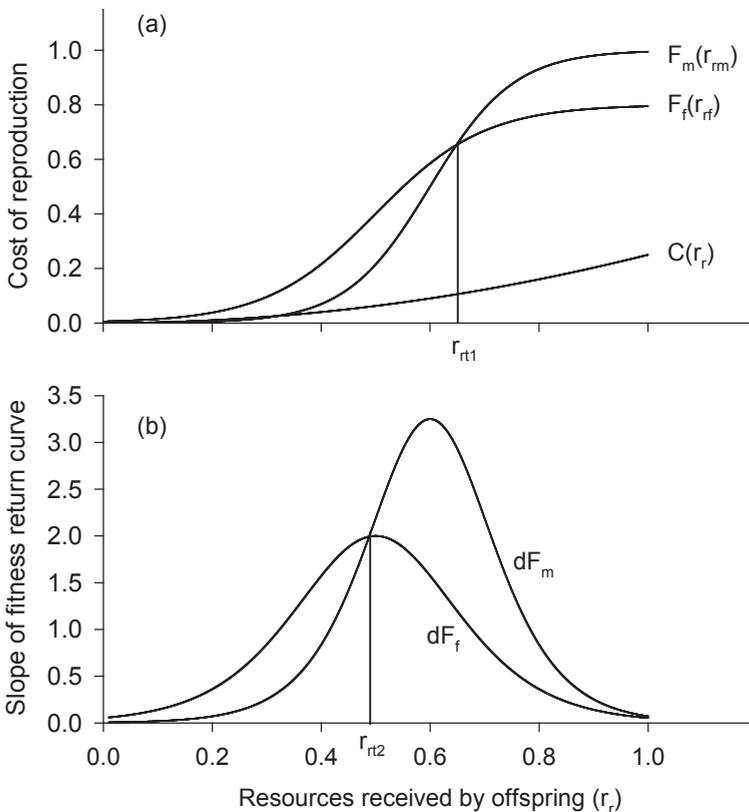
Fig. 1. — (a) The reproductive value of individual male (F_m) and female (F_f) offspring and the cost of reproduction (C) in relation to amount of resources (r_r) devoted to offspring. R_{rt1} represents threshold for sex ratio change. Functions shown are $F_m=1/(1+exp(-13(r_r-0.6)))$, $F_f=0.8/(1+exp(-10(r_r-0.5)))$, and $C=0.25r_r^2$. (b) The slopes of male (dF_m) and female (dF_f) fitness return functions from Fig. 1a. R_{rt2} represents the threshold for sex-biased investment. See text for more details.

biologically realistic, I included two additional parameters in my simulation. First, cost of reproduction is considered (Fig. 1a) because, in iteroparous organism, there is a trade-off between current and future reproduction. Second, as parents could be constrained in their investment decisions if they did not recognize offspring sex at a very early stage, I incorporated into the simulation a minimal amount of resources (m), which must be expended per offspring before its sex is known to parents. The inclusion of the cost of reproduction into the simulation means that it is important to discriminate between resources that are available for parents for current reproductive attempt (r_a) and those that are ultimately expended by parents (r_r). The fitness and cost curves drawn here (Fig. 1a) are similar to those used by CARRANZA (2002) and CAMERON & LINKLATER (2002). Their general shape is biologically plausible: raw fitness returns are likely to be diminishing at least at high levels of parental effort while costs probably accelerate with increasing parental effort (RATNIEKS 1996, LESSELLS 1998). Here it is necessary to stress that for purposes of this illustrative example the shape of fitness return curves is more important than that of cost curve. This is because costs are paid per breeding attempt and as such effect only total amount of resources that parents allocate to the current brood while do not effect within-brood allocation of resources (see LESSELLS 1998).

Now imagine that the species for which the fitness-return curves are drawn has fixed clutch/litter size of two and total amount of resources available per breeding attempt (r_a) varying between 0 and 2 for individual parents. Further assume that parents can maximize their fitness either (1) by varying primary sex ratio with no further differential investment to conceived young or (2) by differential investment in male and female offspring with no ability to change primary sex ratio. For simplicity, assume that sex ratio is fixed at unity in the latter case. Under scenario (1) the solution to parental problem of how to maximize fitness is quite simple. Fitness return curves for female and male offspring cross at 0.65 r_r (Fig. 1a). This value represents a threshold under which it is more profitable to produce daughters while sons are more profitable when resources devoted to them are over this threshold. Consequently parents that are constrained (by the amount of available resources and cost of reproduction) to expend less than 1.3 resources per breeding attempt should produce daughters while those that can afford to expend more than 1.3 resources should produce sons. Under scenario (2) the solution is less apparent. Using simple computer program I computed optimal investment in the two sexes for 420 combinations of r_a (0.1-2.0 with incremental step of 0.1) and m (0-1 with incremental step of 0.05). The simulation was based on fitness and cost curves displayed in Fig. 1a; optimal investment is maximizing parental fitness. The results of the simulation are shown graphically in Fig. 2.

First, it can be seen from Fig. 2 that for all combinations of r_a and m that lie over diagonal connecting points [0.1, 0.05] and [2.0, 1.0] investment in the two sexes is equal. This is unsurprising result; parents simply have not enough resources to perform differential allocation. Second, equal investment is always favoured when m ≥ 0.8. This is because in this illustrative example the largest achievable net fitness returns are at investment 0.81 resources in son and 0.71 in daughter. Investments greater than 1.52 per attempt, which would be required in order to invest differentially when m ≥ 0.8, are not favoured because of high cost of reproduction. Third, greater allocation of resources to female offspring is favoured under some conditions. Fourth, these conditions are restricted to small amount of available resources (r_a); female-biased investment is favoured only when the resources are less than 1.0. Remember that under the scenario (1), females should be produced up to 1.3 r_r.

The reason for this difference between the two scenarios lies in the fact that under scenario (1) the sex with larger current value, which is determined by the amount of resources received by each offspring (r_r), is produced. However, under scenario 2 current value of the progeny is unimportant for parental decisions concerning the allocation of resources. What is important is marginal return from additional investment (LAZARUS 2002). Marginal returns depend on slopes of fitness return curves. The slopes of fitness return curves (determined as their first derivations) from Fig. 1a are shown in Fig. 1b. As for the fitness return curves, the slopes of the fitness return curves also cross. The slope of female function is greater than that of male at small values of r_r but smaller at large values of r_r. However, the crossing point of slopes lies at a value of r_r equal to 0.49 (i.e. at smaller value than crossing point of the fitness functions). This difference indicates that when the amount of resources that can be expended per offspring is between the two crossing points (thresholds) it would pay to produce females but invest more in son whenever one is born. Consider, for example, the case when r_a = 1.1 and m = 0.5. After an equal investment of 0.5 in both sexes the current value of daughter is 0.40 while that of son only 0.21. However, the marginal return for additional investment of 0.1 r_r is 0.19 for daughter and 0.29 for son. So investment will be male-biased under these conditions. Importantly, this result does not depend on exact shape of fitness return curves. When fitness return curves cross, which is the third assumption of TWM, slopes of these fitness functions will also cross but the point of crossing will always be at a lower value of r_r.

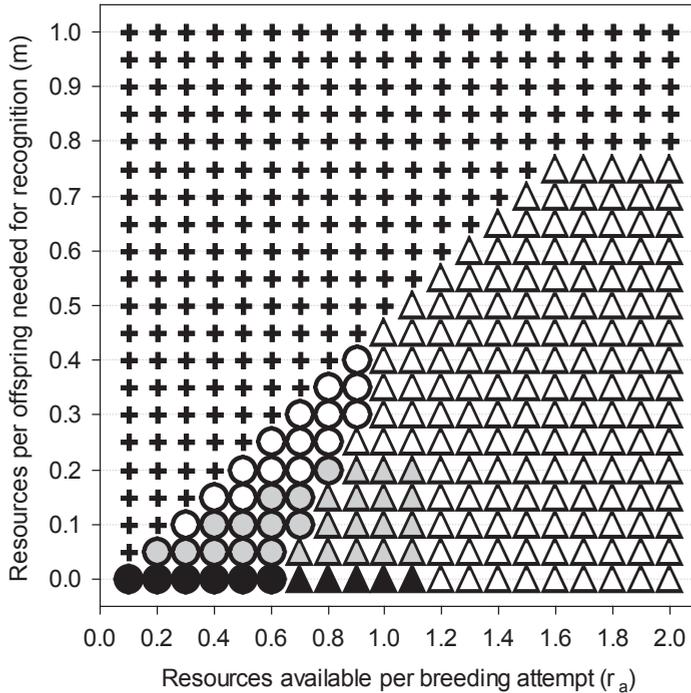Fig. 2. — Optimal investment to the two sexes as determined by r_a and m. The meaning of the symbols is as follows: (black circle) 100% investment in daughter, (black triangle) 100% investment in son, (grey circle) 99.9-75.0% investment in daughter, (grey triangle) 99.9-75.0% investment in son, (white circle) 74.9-50.1% investment in daughter, (white triangle) 74.9-50.1% investment in son, (+) equal investment in the two sexes.

The continuing confusion whether females in poor condition should invest more in daughters than sons may arise from male-biased formulation of the third assumption in original TRIVERS & WILLARD (1973) paper. The authors wrote (p. 90): 'Finally assume that such adult differences in condition affect male reproductive success (RS) more strongly than they affect female RS'. As explained above, this assumption should be formulated more generally (as TRIVERS & WILLARD (1973) did on page 91) because greater dependence of male reproductive success on one end of continuum of available resources (maternal condition) implies that the reverse is true at the other end of continuum. In other words, at the poor end of the resource continuum there will be more variance in reproductive success among female offspring, although at the level of the whole population there is more variance in reproductive success among males. Unfortunately, some subsequent reviews used the imprecise formulation from page 90 as the third assumption of TWM (e.g. HARDY 1997, HEWISON & GAILLARD 1999) which may have contributed to confusion that surrounds the model and consequently to 'sex bias in studies of sex bias' (see CAMERON & LINKLATER 2002).

It is apparent from Fig. 2 that if resources were evenly distributed in the population, as was the case with this simulation, parents should bias their favoritism to male offspring more frequently than toward females. On the contrary, under the same conditions, the sex ratio would be biased toward female sex ($r = 0.35$) as indicated by the position of crossing point of fitness return curves. Of course, even distribution of resources in the population is unlikely. However, it would be unreasonable to think about how different distribution of resources could change population allocation patterns in this illustrative example because the shape of resource distribution interacts with mating system in determining fitness return curves (CHARNOV 1982, FRANK 1987). Nevertheless, on the population level both female-biased sex ratios (WADE et al. 2003) and greater investment in male offspring (MAYNARD SMITH 1980) is expected when males have greater variance in fitness than females.

Testing the predictions of TWM is very popular, however, the amount of testing is in contrast to very limited evidence in support of its third assumption that fitness of male and female offspring depends differently on the amount of resources devoted to them (for exception see CLUTTON-BROCK et al. 1986). Such empirical studies are very difficult to perform, however, they would be very useful both for the prediction of how individual parents should divide resources between the sexes and for better understanding of sex allocation at the population level. It is important to realize, however, that TWM applies to species with small or invariant clutch/litter size. In species with large and variable clutch/litter sizes parents may maximize their fitness by changing the number of offspring produced which, of itself, may even lead to an allocation to the two sexes that is the opposite of the TWM prediction (WILLIAMS 1979, FRANK 1990).

REFERENCES

CAMERON E.Z. & LINKLATER W.L. 2002. Sex bias in studies of sex bias: the value of daughters to mothers in poor condition. *Animal Behaviour* 63: F5-F8.

CARRANZA J. 2002. What did Trivers and Willard really predict? *Animal Behaviour* 63: F3.

CHARNOV E.L. 1982. The theory of sex allocation. *Princeton: Princeton University Press*.

CLUTTON-BROCK T.H., ALBON S.D. & GUINESS F.E. 1986. Great expectations: dominance, breeding success and offspring sex. *Animal Behaviour* 34: 460-471.

FRANK S.A. 1987. Individual and population sex allocation patterns. *Theoretical Population Biology* 31: 47-74.

FRANK S.A. 1990. Sex allocation theory for birds and mammals. *Annual Review of Ecology and Systematics* 21: 13-55.

HARDY I.C.W. 1997. Possible factors influencing vertebrate sex ratios: An introductory overview. *Applied Animal Behaviour Science* 51: 217-241.

HEWISON A.J.M. & GAILLARD J.M. 1999. Successful sons or advantaged daughters? The Trivers-Willard model and sex-biased maternal investment in ungulates. *Trends in Ecology & Evolution* 14: 229-234.

HEWISON A.J.M., GAILLARD J.M., KJELLANDER P., TOIGO C., LIBERG O. & DELORME D. 2005. Big mothers invest more in daughters — reversed sex allocation in a weakly polygynous mammal. *Ecology Letters* 8: 430-437.

KRÜGER O., RADFORD A.N., ANDERSON C. & LIVERSIDGE R. 2005. Successful sons or superior daughters: sex-ratio variation in springbok. *Proceedings of the Royal Society of London (B, Biological Sciences)* 272: 375-381.

LAZARUS J. 2002. Human sex ratios: adaptations and mechanisms, problems and prospects, pp. 287-311. In: Hardy I.C.W., Edit. Sex ratios: concepts and research methods. *Cambridge: Cambridge University Press*.

LEIMAR O. 1996. Life-history analysis of the Trivers and Willard sex-ratio problem. *Behavioral Ecology* 7: 316-325.

LESSELLS C.M. 1998. A theoretical framework for sex-biased parental care. *Animal Behaviour* 56: 395-407.

MAYNARD SMITH J. 1980. A new theory of sexual investment. *Behavioral Ecology and Sociobiology* 7: 247-251.

RATNIEKS F.L.W. 1996. Evolution of unstable and stable biparental care. *Behavioral Ecology* 7: 490-493.

TRIVERS R.L. & WILLARD D.E. 1973. Natural selection of parental ability to vary the sex ratio of offspring. *Science* 179: 90-92.

WADE M.J., SHUSTER S.M. & DEMUTH J.P. 2003. Sexual selection favors female-biased sex ratios: The balance between the opposing forces of sex-ratio selection and sexual selection. *The American Naturalist* 162: 403-414.

WILLIAMS G.C. 1979. The question of adaptive sex ratio in outcrossed vertebrates. *The Proceedings of the Royal Society of London (B, Biological Sciences)* 205: 567-580.

Článek 7: Remeš, V., Krist, M., Bertacche, V. & Stradi, R. (2007) Maternal carotenoid supplementation does not affect breeding performance in the Great Tit (*Parus major*). *Functional Ecology*, 21, 776-783.

Maternal carotenoid supplementation does not affect breeding performance in the Great Tit (*Parus major*)

VLADIMÍR REMEŠ,*† MILOŠ KRIST,*‡ VITTORIO BERTACCHE§ and RICCARDO STRADI§

Department of Zoology, Palacký University, Tř. Svobody 26, 77146 Olomouc, Czech Republic, ‡*Museum of Natural History, Nám. Republiky 5, 77173 Olomouc, Czech Republic*, §*Istituto di Chimica Organica 'A. Marchesini' Università degli Studi di Milano, via Venezian 21, 20133 Milan, Italy*

Summary

1. Carotenoids are micronutrients with many beneficial health-related effects. They are effective antioxidants and stimulants of the immune system. Carotenoids cannot be synthesized in animals and must be obtained from food. As such, they may limit reproductive output and performance, and on the proximate level mediate reproductive trade-offs.

2. We studied carotenoid limitation in wild Great Tits (*Parus major*) by supplementing prelaying and laying females with lutein, the most abundant carotenoid in this species. We followed the effects of this supplementation on egg yolk carotenoid composition, and offspring and parental performance.

3. Females transferred the supplemented lutein into egg yolks, increasing lutein concentration to the upper limit of naturally occurring concentrations in control pairs. Concentrations of zeaxanthin, β-carotene and α-carotene did not differ between supplemented and control pairs.

4. Effects on offspring and parental performance were generally absent or weak. There were no effects on timing of laying, clutch size, hatching success, nestling survival, nestling mass (day 6 and 14), tarsus length or T-cell mediated immune response. Males on supplemented nests fed their young more than those on control nests. There was no positive effect on female feeding or mass.

5. Negligible effects of lutein supplementation on offspring and parental performance might be explained by high natural abundance of carotenoids or other antioxidants, where additional carotenoids bear no strong advantage to the birds. Additionally, conflicting results of different studies may be explained by species-specific features of their life-histories.

Introduction

Animals are expected to allocate limited resources among competing bodily functions so as to maximize fitness. During reproduction, mothers face a fundamental decision of how much resources to invest into current reproductive bout, and how much to retain for maintenance and future reproduction. Besides elaborate postnatal parental care females also invest heavily during the prenatal period into the fabrication of eggs. Besides energy needed for embryo development, eggs are packed with many valuable resources, including

†Author to whom correspondence should be addressed.
E-mail: vladimir.remes@upol.cz

antibacterial enzymes, antibodies, hormones and carotenoids (Blount, Houston & Møller 2000).

Carotenoids are a group of several hundred biologically active compounds with many important biological functions in signalling and physiology (Møller *et al.* 2000). They are widely used in the colouring of bird plumage and bare parts (Olson & Owens 2005). Carotenoids enhance the intensity of both cell-mediated and humoral immune response (Chew & Park 2004). They are also effective scavengers of reactive oxygen species (ROS) that arise during metabolic processes (Krinsky 2001). ROS are free radicals and non-radical oxygen-containing molecules that are able to damage proteins, lipids and DNA (de Zwart *et al.* 1999), a condition called oxidative stress that has been implicated

in the etiology of many diseases and ageing (Crimi *et al.* 2006). ROS are removed by the multifaceted antioxidant system that includes enzymes (e.g. catalase, superoxide dismutase), water-soluble antioxidants (ascorbic acid, glutathione) and fat-soluble antioxidants (vitamin E, carotenoids; Sies & Stahl 1995). Carotenoids can be synthesized only by plants, certain bacteria and fungi, while animals must ingest them with their food. As a 'diet-dependent' resource used in both signalling and physiology, they are a good candidate for mediating life-history trade-offs (Blount 2004).

Birds deposit carotenoids into egg yolks and the amount varies markedly within and among species (Hargitai *et al.* 2006). Some variation among yolks in the concentration of carotenoids can be explained by laying order (e.g. Blount *et al.* 2002a; Saino *et al.* 2002), year (Hargitai *et al.* 2006) or habitat (Hõrak, Surai & Møller 2002; Cassey *et al.* 2005). Studies on both domestic and wild birds demonstrated higher yolk carotenoid concentrations in mothers supplemented by carotenoid-rich diet (Blount *et al.* 2002b; Bortolotti *et al.* 2003; Biard, Surai & Møller 2005; McGraw, Adkins-Regan & Parker 2005; Ewen *et al.* 2006; Berthouly, Helfenstein & Richner 2007). However, we still do not understand whether these patterns represent active deposition of carotenoids by mothers (Blount *et al.* 2002a, Blount *et al.* 2002b; Royle, Surai & Hartley 2003) or simply reflect their supply in the diet (Partali *et al.* 1987).

Carotenoids are responsible for the typical yellow to orange colour of yolks and have important physiological functions. They reduce the susceptibility of yolk lipids to peroxidative damage (Blount *et al.* 2002b) and later protect developing embryo from oxidative stress (Surai, Noble & Speake 1996). This is important because birds grow very fast and their intense metabolism makes them especially vulnerable to oxidative damage (Blount *et al.* 2000). Upon hatching, yolk-derived carotenoids can affect the susceptibility of hatchling tissues to oxidative damage (Surai *et al.* 1996), the ability of chicks to accumulate dietary carotenoids in their body (Koutsos *et al.* 2003), or parameters of their immune function (Haq, Bailey & Chinnah 1996; Koutsos, López & Klasing 2006). All these studies were on domestic hens. Three studies on passerines suggest that similar effects may exist in this group of birds. McGraw *et al.* (2005) found that carotenoid supplementation of females enhances hatching and fledging success in captive Zebra Finches (*Taeniopygia guttata*). Under wild conditions Biard *et al.* (2005) found that young hatching from the eggs of carotenoid-supplemented females had longer tarsi at hatching and more leukocytes in their blood during growth. Berthouly *et al.* (2007) found out that maternally derived carotenoids can help nestlings cope with stress.

To advance our understanding of potential carotenoid limitation in wild birds, we performed a carotenoid-supplementation study in wild Great Tits (*Parus major* Linnaeus 1758). Reproducing parents face a trade-off of how many limited resources to allocate into current reproductive bout *vs* self-maintenance and future reproduction. In this study we focused on the potential for carotenoid limitation in the current reproductive bout. We provided Great Tit pairs with a lutein-rich supplement before and during egg laying and followed the effects of this supplementation on yolk carotenoid concentrations, and reproductive and parental performance. We tested three possible scenarios: (1) Supplemented females do not increase yolk lutein concentration and the intensity of parental care does not change. This would mean that parents are not limited during current reproduction. (2) Supplemented females increase yolk lutein with no effects on offspring performance and no effects on the intensity of parental care. In this scenario, parents are not limited in their current reproductive bout and the female bird just channels surplus micronutrients into the eggs. (3) Supplemented females do increase yolk lutein concentration with positive effects on offspring performance and/or parents care more intensely. This would demonstrate carotenoid limitation during current reproduction.

Methods

FIELD WORK

Great Tits are small, insectivorous, resident passerines that breed in nest holes during April–June in various woodland types. We studied them in 2004 on six nest-box plots (400 nest-boxes in total) in the Velký Kosʼř area in the eastern Czech Republic (49°32′N, 17°04′E, 300–450 m a.s.l.). Three plots were in a sessile oak (*Quercus petraea*) forest, the other three in a Norway spruce (*Picea abies*) forest. Before birds started breeding nest-boxes were checked and cleaned.

We visited nest-boxes daily to determine the start of nest building and egg laying. We marked eggs daily by a water-proof pen. Before and during egg laying we supplemented experimental tit pairs with 25 mg of CWS lutein (DSM Nutritional Products (Basel, Switzerland), composition: 7% of lutein, 1% DL-α-tocopherol, 1% ascorbyl palmitate, 18% fish gelatine, 46% sucrose, 2% sodium ascorbate and 25% corn starch), which means 1·75 mg of lutein daily. According to the information given in Partali *et al.* (1987; *c.* 3·3 µg of carotenoids per one lepidopteran larva) this makes daily increase in carotenoid intake equivalent to *c.* 530 lepidopteran larvae. Control pairs were supplemented with a placebo lacking lutein with otherwise identical composition. We started with 33 experimental and 27 control pairs. Both lutein and placebo were enclosed in a pill made from animal fat (*c.* 0·6 g) and put into a plastic cup (diameter 3 cm, height 2 cm). It was put inside the nest-box inhabited by the focal tit pair, *c.* 5 cm above the nest rim on one side of the nest-box. Supplemental units were freshly prepared every evening and

stored at –20 °C and in the dark until use the following day. To increase the attractiveness of the pill, we always added five meal-worms into the plastic cup. A pill was supplemented daily until egg laying was terminated. We started the supplementation on the day when tits started to bring animal fur into the moss base of the nest being built. We did not start the supplementation earlier, because females may switch between nest-boxes in the earlier phases of nest building. All the pairs which we started to supplement continued in breeding. On average 2·4 pills (SD = 2·6, $N = 60$) had already been eaten on the day the first egg of the clutch was laid (supplemented: $2·4 \pm 2·8$, $N = 33$; control: $2·5 \pm 2·3$, $N = 27$; $F < 0·1$, $P = 0·818$). Supplementation was regularly taken by birds, only 7 out of 812 pills remained uneaten the next day. We did not monitor the nest-boxes and thus we do not know the relative share of the sexes in the consumption of the supplement. When incubation commenced, apart from seven cases we collected the egg laid on that day (i.e. the last egg in the laying sequence that could have been collected without having been incubated for more than a few hours) and stored it at –20 °C before further analysis. On average, 11·5 (SD = 2·8, $N = 53$) pills had been eaten by the birds in each nest before the collected egg was laid (supplemented: $11·5 \pm 3·1$, $N = 28$; control: $11·5 \pm 2·5$, $N = 25$; $F < 0·1$, $P = 0·984$). Average position in the laying sequence of this egg was 10·0 (SD = 2·0, $N = 53$; supplemented: $10·1 \pm 2·1$, $N = 28$; control: $9·9 \pm 1·9$, $N = 25$; $F = 0·1$, $P = 0·726$).

To recognize the young hatching from late eggs that had the greatest probability of being affected by the supplementation, we frequently visited nests around the expected time of hatching. At most nests, we identified and marked nestlings hatched from late eggs in the laying sequence by clipping their dawn feathers. We weighed all the young when they were 6 days old and again when they were 14 days old. On day 14 we also measured their tarsi. On day 13, we measured the thickness of the right wing web of three young per nest (those that hatched from late eggs, if known) with a pressure-sensitive gauge (model PK-1012E, Mitutoyo, Tokyo, Japan) and then injected it with 0·09 mg of phytohaemagglutinin (L-8754, Sigma-Aldrich, St. Louis, MO, USA) in 25 µL of phosphate buffered saline. We re-measured the wing web 24 h later (± 2 h). We always measured the wing web twice and took the average. T-cell mediated immune response was quantified as the difference in the wing web thickness measured 1 day after the injection and a day before. On average 8·8 pills (SD = 3·0, $N = 46$; supplemented: $9·2 \pm 3·0$, $N = 25$; control: $8·1 \pm 2·9$, $N = 21$; $F = 1·7$, $P = 0·194$) had been eaten by parents before the eggs from which the young that were scored for immune response originated were laid (mean position in the laying sequence = 7·4, SD = 2·1, $N = 46$; supplemented: $7·3 \pm 2·0$, $N = 25$; control: $7·4 \pm 2·2$, $N = 21$; $F < 0·1$, $P = 0·985$). In some nestlings, we did not know their exact position in the laying order. In such

cases, we assigned the average of the possible positions for the chick (e.g. if we knew that the chick hatched either from egg 7 or 8, its position in the laying order was assigned to be 7·5).

During incubation we captured females, weighed them on a spring Pesola balance (to the nearest 0·25 g), and measured their tarsus with a digital caliper (to the nearest 0·01 mm). We also quantified male and female feeding rate per hour when nestlings were 7–11 days old (median = 9 days). We set up a camera *c.* 5–10 m from the nest-box and filmed feeding activity for 75 min. We then discarded the first 15 min of the recording and counted number of feeds provided by male and female during subsequent 60 min.

ANALYSIS OF CAROTENOIDS

In the yolk of the collected eggs we determined concentrations of lutein, zeaxanthin, α-carotene and β-carotene. To extract carotenoids, weighted amount of egg yolk (on an average of 200 µL) was homogenized with 2 mL of a mixture of 4% NaCl solution and ethanol (1 : 1, v/v) followed by sonication for 7 min. We then added 3 mL of hexane and further homogenized for 5 min. Yolk was then drawn into the tubes and centrifuged at 6000 r.p.m. for 5 min. After centrifugation hexane was collected and the extraction was repeated three times. Hexane extracts were combined and evaporated under N_2 at room temperature, and the residue was dissolved in 1·5 mL of acetonitrile : dichloromethane (1 : 1, v/v) and centrifuged. The supernatant was used for carotenoid determination. Carotenoids were determined by high performance liquid chromatography equipped with a binary LC pump Model 250 (Perkin Elmer, Norwalk, CT, USA), using two sequential LICHROCARTTM PUROSPHERTM RP18 columns (250 × 4nn I.D) maintained at 40 °C by a column block heater. A mobile phase of acetonitrile : methanol (85 : 15) and acetonitrile : dichloromethane : methanol (70 : 20 : 10, v/v) in linear gradient elution with PDA detection (Series 200, Perkin Elmer) at 450 nm was used. Peaks were identified and quantified using reference carotenoids kindly supplied by Carotenature (Lupsingen, Switzerland).

DATA ANALYSIS

Data were analysed by general linear models. For every response variable (i.e. offspring and parental traits), we fit a separate model with treatment as the main predictor of interest and other explanatory variables that have been previously shown to be important as covariates. Initially, we included habitat (oak *vs* spruce) and season (Julian date of the first egg, date 1 = 1 January) as covariates to all models. We analysed these response variables (additional covariates in parentheses): carotenoid concentration, hatching success (clutch size), nestling survival until day 14 and nestling mass at day 6 (brood size at hatching), nestling

tarsus length at day 14 (brood size at day 14, female tarsus length), nestling body mass at day 14 and T-cell mediated immunocompetence (brood size at day 14, tarsus length at day 14), clutch size (female tarsus length), female body mass (brood size, female tarsus length, day of the nest cycle when captured), and male and female feeding rate per hour (brood size at feeding, hour of day, age of the young). In the case of nestling mass and tarsus length we used the young originating from late eggs that had the greatest chance to be affected by supplementation (the three nestlings used for the PHA test, see above). However, analyses using mean values for all the young in the nest generated identical results (results not shown). Initially, we also included interactions between the treatment and all other factors. We gradually removed non-significant predictors beginning with interactions until only significant factors remained in the model (at $\alpha = 0.05$), with the exception of treatment. It was always retained as the main factor of interest. Variables were checked for departures from normality and appropriately transformed if necessary. We checked the reliability of our results by calculating standardized effect sizes (difference in least squares means of the dependent variable between supplemented and control groups/ SD of the total sample) with their 95% confidence limits. All tests were performed in JMP software of SAS Institute, Cary, NC, USA.

Results

EGG YOLK CAROTENOIDS

Egg yolk concentration of lutein was significantly increased in experimental nests ($F_{1,51} = 20.4$, $P < 0.001$; Fig. 1). This increase was within physiological levels experienced by birds in this population: average
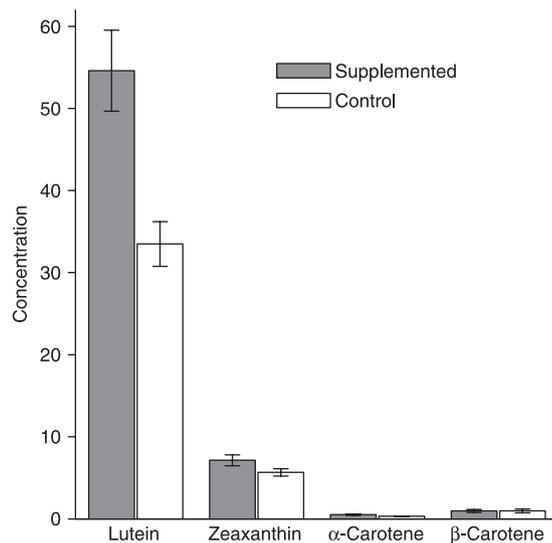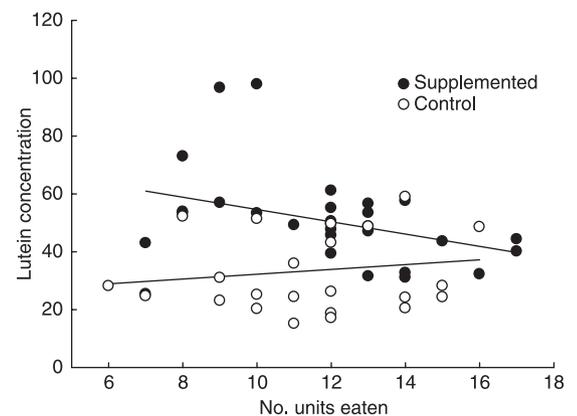
concentration in experimental eggs was $54.6\ \mu g\ g^{-1}$, whereas one-third (8 of 25) of control nests had lutein concentration very close to this value ($49\ \mu g\ g^{-1}$ or higher, maximum value in control eggs was $58.9\ \mu g\ g^{-1}$). Although zeaxanthin tended to have higher concentration in experimental pairs, no other carotenoid differed between experimental and control pairs: zeaxanthin ($F_{1,51} = 3.4$, $P = 0.073$), α-carotene ($F_{1,47} = 2.2$, $P = 0.145$) and β-carotene ($F_{1,47} < 0.1$, $P = 0.944$). Surprisingly, there was a negative relationship between the number of supplementation units eaten and lutein concentration in yolk in experimental pairs ($r = -0.48$, $P = 0.009$, $N = 28$; Fig. 2). However, the significance of this relationship was caused by one outlying nest and disappeared after its exclusion ($r = -0.33$, $P = 0.097$, $N = 27$). There was no significant relationship in control pairs ($r = 0.13$, $P = 0.535$, $N = 25$; Fig. 2). Since the overall effect of supplementation was positive (i.e. increased carotenoid concentration in experimental as compared to control pairs; Fig. 1), it is rather difficult to explain this negative correlation. It might be possible that in females that ate too many units and therefore had ingested a greater amount of lutein this interfered in some way with the incorporation into the egg yolk. However, this is only speculation and further research with precise doses of lutein would be needed to solve this puzzle.

OFFSPRING TRAITS

There was no effect of supplementation on offspring performance-related traits, including hatching success ($F_{1,49} = 0.3$, $P = 0.597$), nestling survival from hatching to day 14 of age ($F_{1,47} = 0.2$, $P = 0.684$), nestling mass at day 6 ($F_{1,47} = 0.7$, $P = 0.402$), nestling mass at day 14 ($F_{1,42} = 0.3$, $P = 0.571$), nestling tarsus length at day 14 ($F_{1,42} = 0.5$, $P = 0.478$) and T-cell mediated immunocompetence ($F_{1,42} = 0.5$, $P = 0.484$; Table 1).

On the other side, these traits were significantly related to some covariates. Hatching success was

Fig. 1. Mean concentration (± 1 SE, in µg/g) of egg yolk carotenoids in the Great Tit in lutein-supplemented and control pairs.

Fig. 2. Relationship between lutein concentration in egg yolk (µg/g) and number of supplemental units eaten fit separately for lutein-supplemented and control pairs. One extreme value (lutein concentration = 156·7 µg g^{-1}, no. of units eaten = 5) is omitted from the figure (see Results).

Table 1. Least squares means (1SE) of offspring and parental traits in lutein-supplemented and control pairs. LS means are from final models with only significant covariates retained
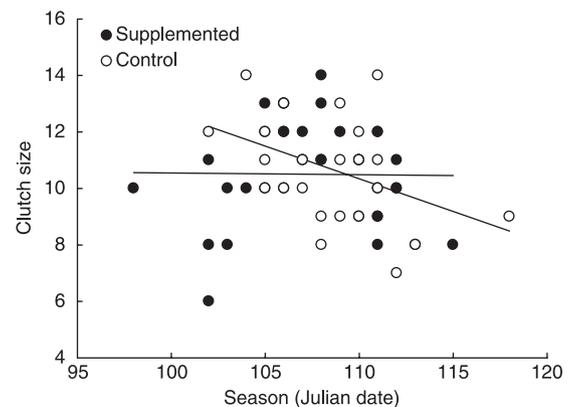
	LS means (SE)	
	Supplemented	Control
Offspring traits		
Hatching success (proportion hatched)	0·86 (0·012)	0·88 (0·014)
Nestling survival (per 14 days)	0·75 (0·055)	0·78 (0·058)
Nestling mass day 6 (g)	8·9 (0·20)	9·2 (0·22)
Nestling mass day 14 (g)	16·9 (0·22)	17·0 (0·22)
Nestling tarsus length (mm)	22·6 (0·11)	22·7 (0·12)
Nestling T-cell immunocompetence (mm)	0·52 (0·026)	0·54 (0·026)
Parental traits		
Clutch size (no. of eggs)	10·2 (0·30)	10·3 (0·31)
Laying date (Julian date)	107·2 (0·66)	108·6 (0·73)
Female feeding rate (per hour)	9·2 (1·23)	8·9 (1·31)
Male feeding rate (per hour)	16·2 (1·61)	11·1 (1·72)
Female body mass (g)	19·0 (0·17)	19·3 (0·17)

Fig. 3. Relationship between clutch size and season (Julian date of the first egg, date 1 = 1 January) fit separately for lutein-supplemented and control pairs.

frequencies were not intercorrelated ($r = -0·17$, $P = 0·242$, $N = 47$). Female body mass did not differ between treatments ($F_{1,50} = 1·8$, $P = 0·184$), whereas it was higher in the oak habitat than in the spruce habitat ($F_{1,50} = 10·5$, $P = 0·002$), scaled positively with female tarsus length ($F_{1,50} = 20·5$, $P < 0·001$), and negatively with the day of the nest cycle at capture ($F_{1,50} = 32·6$, $P < 0·001$; whole model: $F_{4,50} = 19·9$, $P < 0·001$, $R^2 = 0·61$).

CHECKING THE RELIABILITY OF THE RESULTS

In four experimental nests, the concentration of egg yolk lutein was higher than the highest value in any control nest. In these nests, unnaturally high doses of lutein could have toxic effects on nestlings. Then, mixing of beneficial (physiologically high levels) and toxic (pharmacological levels) effects of egg carotenoids in one analysis could have prevented any beneficial effects showing up. Thus, we repeated all the above analyses without those four nests. However, the results did not change. It seems that any potentially harmful effects of very high doses of carotenoids did not compromise our analyses.

For the standardized effect sizes (with confidence intervals) of carotenoid supplementation treatment on offspring and parental performance traits see Fig. 4.

Discussion

To investigate carotenoid limitation on egg formation and reproduction in wild birds, we supplemented prelaying and laying female Great Tits with lutein, the most abundant egg yolk carotenoid in this species (Partali *et al.* 1987). We showed that this supplementation had a clear effect on egg composition, because yolks of supplemented females had significantly more lutein than those of control females (Fig. 1). Strong effect of lutein supplementation on its concentration in egg yolk is not surprising. Increased concentrations of yolk carotenoids in females supplemented with carotenoids in their diet were demonstrated in both captive

positively related to clutch size ($F_{1,49} = 6·2$, $P = 0·016$; whole model: $F_{2,49} = 3·3$, $P = 0·045$, $R^2 = 0·12$), nestling survival was higher in the oak than in the spruce habitat ($F_{1,47} = 15·0$, $P < 0·001$; whole model: $F_{2,47} = 7·5$, $P = 0·001$, $R^2 = 0·24$) and body mass at day 6 was negatively related to brood size ($F_{1,47} = 4·3$, $P = 0·043$; whole model: $F_{2,47} = 2·6$, $P = 0·089$, $R^2 = 0·10$). Further, body mass at day 14 was higher in the oak than in the spruce habitat ($F_{1,42} = 9·2$, $P = 0·004$) and positively related to tarsus length ($F_{1,42} = 42·7$, $P < 0·001$; whole model: $F_{3,42} = 19·3$, $P < 0·001$, $R^2 = 0·58$), tarsus length at day 14 was positively related to both brood size ($F_{1,42} = 5·9$, $P = 0·019$) and female tarsus length ($F_{1,42} = 6·7$, $P = 0·013$; whole model: $F_{3,42} = 5·5$, $P = 0·003$, $R^2 = 0·28$) and T-cell immunocompetence was positively related to brood size ($F_{1,42} = 11·7$, $P = 0·001$; whole model: $F_{2,42} = 5·9$, $P = 0·006$, $R^2 = 0·22$).

PARENTAL TRAITS

There was a significant effect of lutein supplementation on clutch size but it depended on season (interaction: $F_{1,50} = 8·3$, $P = 0·006$). In control nests, clutch size decreased with season whereas in experimental nests it changed nonlinearly – at first it increased, whereas later (after Julian day 105) it decreased in a similar way to control clutches (Fig. 3). Clutch size was furthermore positively affected by female tarsus length ($F_{1,50} = 6·2$, $P = 0·017$) and was larger in the oak as compared to the spruce forest ($F_{1,50} = 16·1$, $P < 0·001$; whole model: $F_{5,50} = 6·1$, $P < 0·001$, $R^2 = 0·38$). Experimental and control pairs did not differ in their timing of breeding ($F_{1,58} = 2·1$, $P = 0·157$).

Female feeding rate did not differ between treatments ($F_{1,45} < 0·1$, $P = 0·869$, $R^2 < 0·01$) whereas males on experimental nests fed more frequently than males on control nests ($F_{1,45} = 4·7$, $P = 0·036$, $R^2 = 0·09$; Table 1). No other factors were significant in the analysis of feeding rates. Male and female feeding

Fig. 4. Standardized effects (with 95% CIs) of treatment on offspring and parental, performance-related traits. Vertical dashed lines denote small (0·2), medium (0·5) and large (0·8) effects, respectively, according to Cohen (1988). For the definition of the traits, see Methods.

© 2007 The Authors.
Journal compilation
© 2007 British
Ecological Society,
Functional Ecology,
21, 776–783

(summarised in Bortolotti *et al.* 2003; McGraw *et al.* 2005) and wild birds (Blount *et al.* 2002b; Biard *et al.* 2005; Ewen *et al.* 2006; Berthouly et al. 2007).

On the other hand, subsequent effects on offspring performance were negligible, whereas the effects on parental traits were slightly stronger. Clutch size and male feeding responded significantly to the supplementation, and although non-significant, confidence intervals for timing of laying included a strong negative effect (i.e. advancement of laying date; see Fig. 4), which suggests that it might have gone undetected because of low statistical power. The effect on clutch size was only apparent in the interaction with season (Fig. 3). These results by and large conform to the scenario 2 (see Introduction). In this scenario, parents are not carotenoid limited in their current reproductive bout and the female bird just channels surplus micronutrients into the eggs. However, males on supplemented nests fed their offspring more than males on control nests. This could mean that males might have been limited in the intensity of parental care, which would conform to the scenario 3. Parents may also have been limited by carotenoids in their self-maintenance and future reproduction. However, based on our data we were not able to test this possibility and it remains an interesting challenge for future work.

We showed that clutch size decreased with the advancement of the laying season in control pairs whereas it changed nonlinearly in supplemented pairs (Fig. 3). This pattern of clutch size change with the season is quite puzzling. Whereas one study found a beneficial effect of carotenoid supplementation on laying potential of females (in Lesser Black-backed Gulls, *Larus fuscus*, Blount *et al.* 2004) other experiments did not demonstrate any effects (Biard *et al.* 2005; McGraw *et al.* 2005). Moreover, since there was no significant effect of carotenoid supplementation on laying date, we have currently no explanation for the pattern found.

Another interesting result is the higher feeding rate of males on supplemented nests as compared to control nests. This may have been caused by better male condition if they also consumed the supplement, in which case they would be carotenoid limited in their current reproductive bout. Alternatively, they may have been willing to increase paternal investment in supplemented broods where supplementation may have made either females or offspring more attractive and worthy of increased investment. In this case this result would not be indicative of carotenoid limitation in males but rather of differential allocation of parental effort (see Sheldon 2000). However, the plausibility of this explanation is decreased by the finding in a recent study of the Great Tit that the young supplemented by carotenoids are not more attractive to parents and the parents do not increase their investment (Tschirren, Fitze & Richner 2005). If proved by further studies, higher feeding rate of males on supplemented nests would be an interesting observation since we currently know virtually nothing about possible relationships between male carotenoid supply, health and physiology, and paternal investment in birds (Blount 2004).

There are several nonexclusive explanations for weak to absent positive effects of our supplementation on offspring performance. Confidence intervals for the effect of supplementation on offspring traits did not embrace either middle or large positive effects (standardised effects of 0·5 and 0·8, respectively, according to Cohen 1988; see Fig. 4). There is a possibility that there were small positive effects (standardized effect size of 0·2) that we were not able to detect with our sample size. However, if there were any important (i.e. middle or large) positive effects of extra carotenoids in eggs on offspring performance, we would have been able to detect them.

Three biologically interesting explanations seem to be worth discussing. First, the detectability of potential effects may depend upon the amount of carotenoids already present in the egg. All eggs may have been supplied with carotenoids to such an extent that any increase brought about by our supplementation had no detectable health and performance related benefits for the offspring. It is known that beneficial effects of carotenoids are dose-dependent, increasing with increasing amounts supplemented but later reaching a plateau (Alonso-Alvarez *et al.* 2004).

Second, the antioxidant system of birds consists of an integrated system of substances, including enzymes, water-soluble and fat-soluble antioxidants. Vitamin E is a fat-soluble antioxidant present in bird egg yolk, including the Great Tit (Hõrak *et al.* 2002). It is transferred from egg yolk to the developing young and increases resistance to oxidative damage of tissues (Surai, Noble & Speake 1999). Our supplementation included small amounts of α-tocopherol (see Methods). It could be possible that α-tocopherol enhanced the antioxidant system of developing chicks in both experimental and control nests to such an extent that its further enhancement by lutein in experimental nests was not detectable. However, in such a case, young

birds would have to be much more sensitive to α-tocopherol than to lutein. Great Tit yolk (*c.* 0·35 g, V. Remeš, unpublished data) contains about 54 µg of α-tocopherol (*c.* 155 µg g^{-1}; Hõrak *et al.* 2002) and about 13 µg of lutein (*c.* 35–38 µg g^{-1}; this study, average for control pairs; Hõrak *et al.* 2002). We supplemented about 250 µg of α-tocopherol (4·6 × the amount in one yolk) and 1750 µg of lutein (135 × the amount in one yolk) daily. Thus, we supplemented about 29 times more intensely with lutein than with α-tocopherol. Accordingly, this explanation does not seem likely. However, factorial experiments supplementing laying mothers with different antioxidant system-enhancing micronutrients (e.g. Surai 2000) in the wild will be needed to resolve this issue.

Third, conflicting results of our study and previous ones could be explained by different study organisms. For instance, Biard *et al.* (2005) studied Blue Tits (*P. caeruleus*). In these smaller birds clutch mass comprises a relatively larger proportion of female body mass than in the closely related Great Tit. Thus, these authors suggest that laying females in this species need relatively more carotenoids and are thus more carotenoid limited than Great Tits (see also Biard, Surai & Møller 2006). This view concerns limitation during acquisition of resources. On the other hand, species differ in their resolution of the trade-off between current and future reproduction based on their position on the slow-fast life-history continuum (Ghalambor & Martin 2001). This might drive the species-specific patterns of allocation of acquired (supplemented) micronutrients between offspring and self-maintenance. Life-history differences between species together with carotenoid supply in the environment might thus be responsible for conflicting results of carotenoid-supplementation experiments. For further development of this area, it will be critical to perform similar supplementation studies on various species differing in their life-history strategies, while at the same time also following allocation of supplemented carotenoids to self-maintenance and future reproduction.

In general, carotenoid-supplementation studies that generated clear and strong positive effects on offspring performance were either performed in captivity (McGraw *et al.* 2005) or carotenoids were injected directly into the eggs in the wild (Saino *et al.* 2003). Food supplementation studies made in the wild up to now generated weak and unconvincing results (Biard *et al.* 2005; Berthouly *et al.* 2007; this study). Decisiveness of these weak results becomes even lower in the light of the number of statistical tests often performed with inherently increased probability of statistical error and finding false relationships (see also de Ayala, Martinelli & Saino 2006). This is surprising given the many beneficial health-related effects of carotenoids (see above, but see McCall & Frei 1999). It may be more difficult to detect beneficial effects of carotenoids in the wild because of less well controlled experimental conditions. Wild birds may also have

enough carotenoids to provide to their young with resulting sufficient antioxidant protection. Further experimental increase of carotenoids may then operate in the plateau region of the dose-dependent relationship between carotenoid concentration and beneficial effects (Alonso-Alvarez *et al.* 2004). Experiments with carotenoid-deplete and carotenoid-replete eggs in semi-natural conditions could help to resolve this issue (e.g. Koutsos *et al.* 2003). Alternatively, the antioxidant system of the young may be supported by other anti-oxidants to a sufficient level, again precluding any potentially beneficial effects of carotenoids to show up. These thoughts are in line with weak beneficial effects of direct supplementation of nestling food with dietary carotenoids (Biard *et al.* 2006) and with vitamin E (de Ayala *et al.* 2006) in the wild. Moreover, the effects detected by Biard *et al.* (2006) differed between species, and the authors suggested that the potential for beneficial effects of supplemental antioxidants might vary with the life-history strategy of the particular species. Similar species-specific effects were found in the relationships between carotenoids, immune function and male ornamentation in birds (summarized by Blount 2004). More studies on diverse species taking into account broader spectrum of antioxidants and employing more sophisticated study design are clearly needed to resolve these interesting issues.

References

Alonso-Alvarez, C., Bertrand, S., Devevey, G., Gaillard, M., Prost, J., Faivre, B. & Sorci, G. (2004) An experimental test of the dose-dependent effect of carotenoids and immune activation on sexual signals and antioxidant activity. *American Naturalist* **164**, 651–659.

de Ayala, R.M., Martinelli, R. & Saino, N. (2006) Vitamin E supplementation enhances growth and condition of nestling barn swallows (*Hirundo rustica*). *Behavioral Ecology and Sociobiology* **60**, 619–630.

Berthouly, A., Helfenstein, F. & Richner, H. (2007) Cellular immune response, stress resistance and competitiveness in nestling great tits in relation to maternally transmitted carotenoids. *Functional Ecology* **21**, 335–343.

Biard, C., Surai, P.F. & Møller, A.P. (2005) Effects of carotenoid availability during laying on reproduction in the blue tit. *Oecologia* **144**, 32–44.

Biard, C., Surai, P.F. & Møller, A.P. (2006) Carotenoid availability in diet and phenotype of blue and great tit nestlings. *Journal of Experimental Biology* **209**, 1004–1015.

Blount, J.D. (2004) Carotenolds and life-history evolution in animals. *Archives of Biochemistry and Biophysics* **430**, 10–15.

Blount, J.D., Houston, D.C. & Møller, A.P. (2000) Why egg yolk is yellow. *Trends in Ecology and Evolution* **15**, 47–49.

Blount, J.D., Houston, D.C., Surai, P.F. & Møller, A.P. (2004) Egg-laying capacity is limited by carotenoid pigment availability in wild gulls *Larus fuscus. Proceedings of the Royal Society of London B* **271**, S79–S81.

Blount, J.D., Surai, P.F., Houston, D.C. & Møller, A.P. (2002b) Patterns of yolk enrichment with dietary carotenoids in gulls: the roles of pigment acquisition and utilization. *Functional Ecology* **16**, 445–453.

Blount, J.D., Surai, P.F., Nager, R.G., Houston, D.C., Møller, A.P., Trewby, M.L. & Kennedy, M.W. (2002a) Carotenoids and egg quality in the lesser black-backed gull *Larus fuscus*: a supplemental feeding study of maternal effects. *Proceedings of the Royal Society of London B* **269**, 29–36.

Bortolotti, G.R., Negro, J.J., Surai, P.F. & Prieto, P. (2003) Carotenoids in eggs and plasma of red-legged partridges: effects of diet and reproductive output. *Physiological and Biochemical Zoology* **76**, 367–374.

Cassey, P., Ewen, J.G., Boulton, R.L., Blackburn, T.M., Møller, A.P., Biard, C., Olson, V. & Karadas, F. (2005) Egg carotenoids in passerine birds introduced to New Zealand: relations to ecological factors, integument coloration and phylogeny. *Functional Ecology* **19**, 719–726.

Chew, B.P. & Park, J.S. (2004) Carotenoid action on the immune response. *Journal of Nutrition* **134**, 257S–261S.

Cohen, J. (1988) *Statistical power analysis for behavioral sciences.* Lawrence Erlbaum Associates, Hillsdale, NJ.

Crimi, E., Sica, V., Williams-Ignarro, S., Zhang, H., Slutsky, A.S., Ignarro, L.J. & Napoli, C. (2006) The role of oxidative stress in adult critical care. *Free Radical Biology and Medicine* **40**, 398–406.

Ewen, J.G., Thorogood, R., Karadas, F., Pappas, A.C. & Surai, P.F. (2006) Influences of carotenoid supplementation on the integrated antioxidant system of a free living endangered passerine, the hihi (*Notiomystis cincta*). *Comparative Biochemistry and Physiology A* **143**, 149–154.

Ghalambor, C.K. & Martin, T.E. (2001) Fecundity-survival trade-offs and parental risk taking in birds. *Science* **292**, 494–497.

Haq, A.U., Bailey, C.A. & Chinnah, A. (1996) Effect of β-carotene, canthaxanthin, lutein, and vitamin E on neonatal immunity of chicks when supplemented in the broiler breeder diets. *Poultry Science* **75**, 1092–1097.

Hargitai, R., Matus, Z., Hegyi, G., Michl, G., Tóth, G. & Török, J. (2006) Antioxidants in the egg yolk of a wild passerine: differences between breeding seasons. *Comparative Biochemistry and Physiology B* **143**, 145–152.

Hõrak, P., Surai, P.F. & Møller, A.P. (2002) Fat-soluble antioxidants in the eggs of great tits *Parus major* in relation to breeding habitat and laying sequence. *Avian Science* **2**, 123–130.

Koutsos, E.A., Clifford, A.J., Calvert, C.C. & Klasing, K.C. (2003) Maternal carotenoid status modifies the incorporation of dietary carotenoids into immune tissues of growing chickens (*Gallus gallus domesticus*). *Journal of Nutrition* **133**, 1132–1138.

Koutsos, E.A., López, J.C.G. & Klasing, K.C. (2006) Carotenoids from in ovo or dietary sources blunt systemic indices of the inflammatory response in growing chicks (*Gallus gallus domesticus*). *Journal of Nutrition* **136**, 1027–1031.

Krinsky, N.I. (2001) Carotenoids as antioxidants. *Nutrition* **17**, 815–817.

McCall, M.R. & Frei, B. (1999) Can antioxidant vitamins materially reduce oxidative damage in humans? *Free Radical Biology and Medicine* **26**, 1034–1053.

McGraw, K.J., Adkins-Regan, E. & Parker, R.S. (2005) Maternally derived carotenoid pigments affect offspring survival, sex ratio, and sexual attractiveness in a colorful songbird. *Naturwissenschaften* **92**, 375–380.

Møller, A.P., Biard, C., Blount, J.D., Houston, D.C., Ninni, P., Saino, N. & Surai, P.F. (2000) Carotenoid-dependent signals: indicators of foraging efficiency, immunocompetence or detoxification ability? *Avian and Poultry Biology Reviews* **11**, 137–159.

Olson, V.A. & Owens, I.P.F. (2005) Interspecific variation in the use of carotenoid-based coloration in birds: diet, life history and phylogeny. *Journal of Evolutionary Biology* **18**, 1534–1546.

Partali, V., Liaaenjensen, S., Slagsvold, T. & Lifjeld, J.T. (1987) Carotenoids in food chain studies. 2. The food chain of *Parus* spp monitored by carotenoid analysis. *Comparative Biochemistry and Physiology B* **87**, 885–888.

Royle, N.J., Surai, P.F. & Hartley, I.R. (2003) The effect of variation in dietary intake on maternal deposition of antioxidants in zebra finch eggs. *Functional Ecology* **17**, 472–481.

Saino, N., Bertacche, V., Ferrari, R.P., Martinelli, R., Møller, A.P. & Stradi, R. (2002) Carotenoid concentration in barn swallow eggs is influenced by laying order, maternal infection and paternal ornamentation. *Proceedings of the Royal Society of London B* **269**, 1729–1733.

Saino, N., Ferrari, R., Romano, M., Martinelli, R. & Møller, A.P. (2003) Experimental manipulation of egg carotenoids affects immunity of barn swallow nestlings. *Proceedings of the Royal Society of London B* **270**, 2485–2489.

Sheldon, B.C. (2000) Differential allocation: tests, mechanisms and implications. *Trends in Ecology and Evolution* **15**, 397–402.

Sies, H. & Stahl, W. (1995) Vitamins E and C, β-carotene, and other carotenoids as antioxidants. *American Journal of Clinical Nutrition* **62**, 1315S–1321S.

Surai, P.F. (2000) Effect of selenium and vitamin E content of the maternal diet on the antioxidant system of the yolk and the developing chick. *British Poultry Science* **41**, 235–243.

Surai, P.F., Noble, R.C. & Speake, B.K. (1996) Tissue-specific differences in antioxidant distribution and susceptibility to lipid peroxidation during development of the chick embryo. *Biochimica et Biophysica Acta* **1304**, 1–10.

Surai, P.E., Noble, R.C. & Speake, B.K. (1999) Relationship between vitamin E content and susceptibility to lipid peroxidation in tissues of the newly hatched chick. *British Poultry Science* **40**, 406–410.

Tschirren, B., Fitze, P.S. & Richner, H. (2005) Carotenoid-based nestling colouration and parental favouritism in the great tit. *Oecologia* **143**, 477–482.

de Zwart, L.L., Meerman, J.H.N., Commandeur, J.N.M. & Vermeulen, N.P.E. (1999) Biomarkers of free radical damage applications in experimental animals and in humans. *Free Radical Biology and Medicine* **26**, 202–226.

Článek 8: Krist, M. & Grim, T. (2007) Are blue eggs a sexually selected signal of female collared flycatchers? A cross-fostering experiment. *Behavioral Ecology and Sociobiology*, **61, 863-876.**

ORIGINAL PAPER

Are blue eggs a sexually selected signal of female collared flycatchers? A cross-fostering experiment

Miloš Krist · Tomáš Grim

Abstract Impressive variation in egg colouration among birds has puzzled evolutionary biologists for a long time. The most frequently studied selective forces moulding egg colouration—predation and brood parasitism—have either received little empirical support or may play a role in only a minority of species. A novel hypothesis has suggested that egg colour may be significantly influenced by sexual selection. Females may deposit a blue-green pigment biliverdin into eggshells instead of using it for themselves as a powerful antioxidant. By this handicap, females may signal their quality to males, which are then hypothesized to increase their paternal effort. We experimentally tested the hypothesis in the collared flycatcher (*Ficedula albicollis*), a species laying blue-green eggs. We cross-fostered clutches between nests to disentangle effects of female/territory quality and egg colour on paternal effort and nestling quality. The results supported two assumptions of sexual signalling through egg colour hypothesis: Blue pigment seems to be a limited resource for females, and female quality is positively correlated with the intensity of the blue-green colour. However, we did not find support for the main prediction of the hypothesis, as male parental effort parameters (feeding frequencies to nestlings and intensity of nest defence) were unrelated to egg colour. We discuss possible reasons for the discrepancy between our results and previous correlative analyses that supported the hypothesis that blue egg colour may be a postmating, sexually selected signal in females.

M. Krist (✉)
Museum of Natural History,
Nám. Republiky 5,
771 73 Olomouc, Czech Republic
e-mail: milos.krist@upol.cz

M. Krist · T. Grim
Department of Zoology, Palacký University,
tř. Svobody 26,
771 46 Olomouc, Czech Republic

Introduction

Egg colouration has been attracting the attention of biologists for decades, and various hypotheses have been generated to explain the variation in this trait (Underwood and Sealy 2002). Recent research indicates that pigments may physically strengthen the eggshell (Gosler et al. 2005). However, the majority of suggested explanations propose that egg colouration primarily has a signalling function to either heterospecifics or conspecifics. On the heterospecific level, most attention has been paid to nest predators and brood parasites. Eggs may have cryptic colours to preclude predators from locating eggs/nests (Tinbergen et al. 1962). Brood parasitism may select for small intraclutch variation but large interclutch variability in host egg colouration (Øien et al. 1995; Soler and Møller 1996), and hosts may select for mimetic eggs in parasites by rejecting eggs differing from their own (Davies and Brooke 1988; Grim and Honza 2001). On the intraspecific level, egg colouration may be important for the recognition of ones own eggs in colonial birds (Birkhead 1978). By laying a pale egg last in a clutch, females may signal that they have started to incubate, which would diminish the opportunities for successful conspecific brood parasitism (Yom-Tov 1980; Ruxton et al. 2001).

The support from observations and experiments for each of these hypotheses is equivocal (Underwood and Sealy 2002). Thus, the large amount of variation in egg colour remains unexplained. For example, in a carefully designed study by Weidinger (2001), open-cup nests baited with different coloured (white, blue and cryptic) eggs had equal survival rates despite the fact that the study was conducted in a population where most nests were eventually destroyed by nest predators and where one could therefore expect strong selection for egg crypsis.

One of the most striking—and most puzzling—egg colours is blue (Underwood and Sealy 2002). The evolutionary significance of blue eggs was previously studied in the context of predation and brood parasitism. Götmark (1992) and Weidinger (2001) found no support for a hypothesized cryptic function of blue eggs (blending with specific nest micro-environment). On the other hand, there is some evidence that the blue egg morph in the common cuckoo (*Cuculus canorus*) is a counter-adaptation against host egg discrimination in host species laying bluish eggs (Moksnes et al. 1995). However, this does not explain why the hosts themselves lay blue eggs in the first place. Consequently, until recently, blue-green colouration of eggs was considered to be a mystery (Underwood and Sealy 2002).

Recently, a new hypothesis based on intraspecific signalling has been proposed to explain variation in egg colours (Moreno and Osorno 2003). According to this hypothesis, females colour their eggs using costly pigments to signal their quality and, consequently, the quality of their progeny to their mates. The costs of such extended phenotype should be outweighed by benefits from increased paternal care into the current brood. Such an increase of paternal care in high quality offspring is in turn predicted by the differential allocation hypothesis (Burley 1986; Sheldon 2000). Moreno and Osorno (2003) paid special attention to blue-green eggs when formulating their hypothesis because biliverdin, which causes the blue colour of eggs (Mikšík et al. 1996), has also a strong antioxidant activity (McDonagh 2001; Kaur et al. 2003). Thus, the deposition of biliverdin into eggs may signal female capacity to control free radicals despite the handicap (Moreno and Osorno 2003). Such a signal would be in principle similar to colouration of plumage with other strong antioxidants such as carotenoids. At present, the latter topic has received considerable attention in respect to male signalling (e.g. Hill 2002; McNett and Marchetti 2005).

Female signalling through ornamentation of plumage has been neglected both in studies of mate choice (Amundsen 2000; Hill 2002) and differential allocation (Sheldon 2000). An interest in this topic is currently growing, and some evidence has already been found to suggest that female plumage ornaments are important for male mate choice (Hill 2002; Griggio et al. 2005), paternal effort (Hill 2002; Pilastro et al. 2003) and even sperm allocation (Pizzari et al. 2003). Although female signalling through colour of plumage and eggs is similar in principle, colour of eggs may be under stronger sexual selection than that of plumage, as the first signal might be more informative than the latter. Although many types of pigments appearing in plumage may have antioxidant capacities similar to the egg pigment biliverdin (McGraw 2005), they may be indicative of either health or pigment availability at the time of molting. However, molting is usually separated from breeding by considerable time periods. Consequently, at the time of breeding, males can assess only genetic or persistent environmental quality from female plumage. In contrast to plumage ornaments, egg colour may be indicative of the females' current physiological state, condition and immunity.

To date, five correlational and one experimental study have tested the sexual selection hypothesis of evolution of the blue colouration of eggs (SSEC hypothesis). Three of the correlative and the one experimental study tested the validity of the assumption of the hypothesis, namely whether egg colour was correlated with female and/or offspring quality. In the pied flycatcher (*Ficedula hypoleuca*), a species laying blue eggs, it has been found that the saturation of egg colour was positively correlated with female immunocompetence (Moreno et al. 2005) and amount of maternal yolk antibodies (Morales et al. 2006). Blue-green chroma of eggs was also positively correlated with female condition in another species laying blue eggs, the bluebird *Sialia sialis* (Siefferman et al. 2006). Food supplementation before and during egg laying enabled female pied flycatchers to lay more saturated eggs, which suggests that egg colour depends on female current nutritional state (Moreno et al. 2006).

So far, only two correlative studies tested the main prediction of the sexually selected egg colour (SSEC) hypothesis, i.e. whether egg colour predicts paternal effort. Moreno et al. (2004) found that males of the pied flycatcher fed broods hatched from eggs with more saturated colour with greater frequency. Soler et al. (2005), in a comparative study, found positive correlations between blue colour of eggs, mating system and duration of nestling period. The two latter variables were used by Soler et al. (2005) as surrogates of intraspecific variation in paternal effort, a variable that should affect the evolution of female signalling.

Thus, all studies performed so far have given results that are consistent with the SSEC hypothesis. However, the causality of the observed relationships is unclear because of the correlative nature of most of these studies. Moreover, in three studies on the pied flycatcher (Morales et al. 2006; Moreno et al. 2004, 2005), the reflectance spectra of eggs were

measured at wavelengths above 360 or 400 nm. As most passerines are able to see even shorter wavelengths (Cuthill et al. 2000), further studies that include the UV part of the spectrum when measuring egg colour are particularly needed.

The main aim of this study was to experimentally test for a causal link between egg colour and paternal effort. We performed the study on the collared flycatcher (*Ficedula albicollis*), which lays, similarly to its sister species, the pied flycatcher, blue-green eggs. As the collared flycatcher is a hole-nester, the potential confounding effects of both predation and interspecific brood parasitism on egg colour can be rejected a priori. To test for causality between egg colour and paternal effort, we cross-fostered freshly laid eggs between pairs of nests. Consequently, egg colour was randomized (in the extent of 92%, see "Materials and Methods") with respect to parental and territory quality. The UV range was included in the measuring of egg spectra to deal with the presumed UV vision of flycatchers. Three measures of paternal effort were scored (feeding frequencies to young and old nestlings and nest defence against a nest predator) to enable a more detailed estimate to be obtained. In addition, we tested for relationships between egg colour and offspring survival, morphology and T cell-mediated immunity. We also paid attention to female parents to test whether the egg colour is correlated with any other female traits.

If the SSEC hypothesis (Moreno and Osorno 2003) holds, egg colour of cross-fostered, but not original eggs, should predict paternal effort. If the colour of original eggs predicted paternal effort, this would suggest one of the following: (1) Males adjust their effort according to some female trait that is correlated with the colour of her eggs; (2) high quality females laying eggs with more saturated colours mate assortatively with high quality males that are able to provide superior parental care; and (3) the relationship is driven by territory quality: On superior territories, it may be less costly both to lay eggs with saturated colours and feed young with greater frequency. If offspring quality was correlated with colour of cross-fostered eggs (i.e. eggs from which the young actually hatched), this would imply either genetic or environmental (i.e. egg composition) quality of eggs with saturated colour or differential male effort. In contrast, if colour of original eggs was predictive of offspring quality, this would imply that parental or territory quality is correlated with egg colour.

Materials and methods

Field methods

We conducted the study in the Velký Kosíř area (49°32′N, 17°04′E, 300–400 m a.s.l.) of the Czech Republic in 2005.

In the study area, there were approximately 300 nest-boxes in oak (*Quercus petraea*) forest. We conducted the experiment with collared flycatcher, a small migratory passerine that easily adopts nest-boxes for breeding. Females usually lay one egg per day and solely incubate clutches of 4–8 eggs. Similar to the sister species, the pied flycatcher (Moreno et al. 2005), eggs are unspotted and brightly blue-green in colour. Both collared flycatcher parents feed nestlings with invertebrate food for about 15 days until fledging.

We conducted cross-fostering experiments among 70 occupied nest-boxes. We cross-fostered eggs between pairs of nests in which laying began on the same day ($n=30$ nest pairs) or which differed in laying date by 1 day ($n=5$ nest pairs). We cross-fostered eggs on the day they were laid with the exception of five nests in which the first eggs were moved the day after they were laid. We recorded the time of both the removal of the original egg and the addition of the cross-fostered egg. We performed the exchange of first eggs of the clutches within 2 h, during which, nest-box entrances were blocked by sticks to ensure that no parent was aware that its nest was empty. Subsequent exchanges were done within a more variable timeframe (0–10 h), during which, the nest-box entrances were not blocked. We continued with exchanges on daily basis until the clutches in both nests had been completed. Original and cross-fostered clutch size was the same in 16 pairs of experimental nests; they differed by one in 14 pairs, by two in four pairs and by three in one pair. Where appropriate, we controlled for the difference in original and cross-fostered clutch size in statistical analyses. No experimental nest was abandoned in the course of laying or incubation.

The experimental design ensured that in the incubation stage, there were only cross-fostered eggs in a focal nest. However, in the laying stage, there were both original and cross-fostered eggs in focal nests. Under the assumption that eggs were visible from 6 A.M. to 8 P.M. in the laying stage (before the last egg was laid), the number of egg-hours for which the original eggs were exposed in the nests was 41.7 ± 9.9 (mean±SD), whereas the respective figure for cross-fostered eggs was 195.0 ± 62.0. Some females began to incubate before their clutch was complete. As eggs are less visible when females are incubating, we also counted the number of egg-hours for which eggs were exposed in nests before females began continuous incubation. For original eggs, this figure was 35.7 ± 10.9 and that for cross-fostered eggs was 143.8 ± 77.6. Taken together, these data suggest that the colour of cross-fostered eggs should have an overwhelming effect on paternal behaviour when compared with that of original eggs. Moreover, we can reasonably expect males to examine egg colours after clutch completion to avoid any assessment bias resulting from assessing an only partially laid clutch. In another

system where birds assess egg appearance (hosts of brood parasites), this happens as a rule after clutch completion (Davies and Brooke 1988).

We began to check nests 2 days before the presumed hatching to determine hatching date and hatching success. To estimate cell-mediated immune response of nestlings, we injected them with 0.1 mg phytohaemagglutinin in 20 µl of physiological saline solution into the right wing web when they were 12 days old. Before injection, we took two measurements of wing web thickness (to the nearest 0.01 mm) with a thickness gauge (Mitutoyo Quick-Mini) that was adjusted to push with constant pressure of 1 N. We re-measured wing web thickness 24 h (±2 h) after injection. Both measures were highly repeatable (before injection: $r=$ 0.797, $F_{251,252}=8.84$, $p<0.001$; after injection: $r=0.972$, $F_{249,250}=70.38$, $p<0.001$). Therefore, we calculated cell-mediated immune response for each chick as the difference in average thickness of wing web after injection minus the average thickness of wing web before injection. At age 13 days, we also weighed nestlings with Pesola spring balance (to the nearest 0.25 g), measured their tarsus with digital calliper (to the nearest 0.01 mm) and recorded whether ectoparasitic mites (*Dermanyssus gallinoides*) were present in the nest. We also captured adults while feeding nestlings and weighed and measured them in the same way as nestlings.

Measures of parental investment

To investigate parental investment in the current brood, we recorded two principal types of investment. First, we recorded parental feeding frequencies of young at two different ages. We videotaped nest-boxes with cameras for 70 min on day 6 and 12 of the nestling period (hatching day = day 0). Feeding frequency was determined separately for each sex as the number of visits to the nest-box per hour starting 10 min after the beginning of the tape recording. We disregarded the first 10 min of recording to minimize the effect of disturbance due to installation of cameras. Collared flycatchers in our study area readily resume feeding regimes within minutes after nest-box checks (personal observations).

Second, we recorded parental nest defence against a predator of eggs and young. At a variable age of young (5–13 days), we simulated intrusion of the great spotted woodpecker (*Dendrocopos major*), which is a common nest predator on our study plots, in the close vicinity of flycatchers' nests. We chose this species because it is not a threat to parents themselves, which could confound results (we were interested in the parental defence of their nestlings, not of parents themselves). Natural interactions between our study species and the great spotted woodpecker are common in our study area (own observations) and

confirm that woodpeckers are perceived as a threat by flycatchers. This is further confirmed by the fact that the frequency of contact attacks against the woodpecker dummy was much greater (approximately tenfold) than that against a control species used in a previous study which was done according to a similar methodology (Krist 2004). We used a stuffed specimen of female woodpecker in a posture simulating foraging on the bark. We attached the dummy to the bark of the tree at approximately 0.5 m below the nest-box of a tested pair. The dummy was oriented (looked) towards the nest-box. Before the dummy had been placed, a camera was installed that videotaped the vicinity of the nest-box. After the experiment began, the observer retreated into the shelter that was installed at least 20 m from the focal nest-box.

After a member of the pair approached close vicinity of the nest and so presumably noticed the dummy, observation of this individual began. Observations for the later-arriving individual began after its arrival. Observations lasted 5 min for each focal bird, which is the recommended length for this type of study (Sealy et al. 1998). During that time, we recorded for each individual the number of dive/contact attacks against a dummy and the latency from arrival to the first dive/contact attack. We focused our attention on dive/contact attacks because these are presumably the most risky behaviours with the greatest efficiency in deterring nest predators (see also Krist 2004; Grim 2005). Further, we scored the overall response of each individual on an ordinal scale: (1) normal feeding or silent watching on the dummy, (2) fluttering against a dummy and/or vocalization and (3) dive or contact attack against a dummy. We checked our field diary notes against the video recordings to ensure that no behaviour was overlooked when quantifying parental response for analyses.

Egg colour measurements

We measured the colour of the eggs on the day they were laid before they were transferred to foster nests. Egg colour was measured by Avantes spectrometer (AvaSpec-2048) which was configured for measurements in the UV-visible part of the spectrum. The light source for measurements was PX-2 pulsed xenon lamp. The spectrometer and the lamp were connected by a bifurcated fibre optic cable to a reflection probe, which consisted of seven optic fibres. Six of them transfer light from the PX-2 lamp to a measured surface, and the seventh one transfers reflected light to the spectrometer. The reflection probe was fixed in a probe holder at an angle of 45° and at a distance of approximately 1 cm from the measured surface. We placed eggs side by side on the measurement port (ellipse with axes 7 and 9 mm) in the probe holder. We then covered the eggs with a black cap to prevent ambient light from confusing measure-

ments. To improve the signal/noise ratio, each spectrum was obtained as the mean of 100 readings with one reading over 10-ms integration time. We measured each egg twice (on the opposite sides), and the mean from the two measurements was used in subsequent analyses. All data were generated relative to a white standard (WS-2 Avantes).

We calculated three values from the reflectance spectra. (1) Brightness or total reflectance was obtained as the summed reflectance at each 1-nm interval from 301 to 700 nm. (2) Blue-green chroma (BGC hereafter) was calculated as the reflectance between 401–600 nm divided by the total reflectance. We focused on this part of the spectrum because biliverdin, the main eggshell pigment of blue-green eggs (Mikšík et al. 1996), absorbs light weakly at this spectral range but strongly at shorter (<400 nm) and longer (>600 nm) wavelengths (Falchuk et al. 2002), which causes the blue-green appearance of eggs. (3) Hue was calculated as $\arctan\{[(Q_G - Q_{UV})/Q_T]/[(Q_R - Q_B)/Q_T]\}$ where Q_T denotes brightness (summed reflectance between 301–700 nm), Q_{UV} is the summed reflectance in the UV area of the reflectance spectrum (301–400 nm), Q_B is the summed reflectance in blue area of the spectrum (401–500 nm), Q_G is the summed reflectance in green area of the spectrum (501–600 nm) and Q_R is the summed reflectance in the red area of the spectrum (601–700 nm). The method for calculating hue was basically the same as that used by Saks et al. (2003), with the exception that we used the whole spectral range visible to birds (i.e. 301–700 nm) in the calculations (for a similar approach see McNett and Marchetti 2005; Montgomerie 2006). Note that in such method, boundaries of segments differ from those which are used in human vision studies. This is because the segment method (Endler 1990) on which the calculation of our index of hue was based requires division of the whole visible spectrum into four segments of the same length regardless of the range of the spectrum (see also Montgomerie 2006). Eggs with greater values of the hue (i.e. less negative values) have their peak of reflectance at shorter wavelength than those with more negative values.

All three colorimetric variables described above should be affected by pigment concentration, however, in unsaturated colour such as that of flycatchers eggshells, chroma should be most indicative of pigment concentration (Andersson and Prager 2006). Indeed chroma has been found to predict concentration of carotenoids in the plumage of greenfinches *Carduelis chloris* (Saks et al. 2003) and biliverdin in eggshell of pied flycatcher (Moreno et al. 2006). Consequently, as BGC should be most closely linked to biliverdin content of eggshells, we would expect that it should be the main cue for male assessment of female quality and, therefore, the main predictor of paternal behaviour.

Mean brightness of eggs in a clutch was strongly correlated with mean BGC ($r=-0.835$, $n=70$, $p<0.001$), whereas mean hue was not correlated with either mean brightness ($r=0.009$, $n=70$, $p=0.944$) or mean BG chroma ($r=-0.179$, $n=70$, $p=0.138$). The strong negative correlation between mean brightness and mean BGC suggests that these two measures are largely from the same trait. Therefore, we did not use brightness for statistical testing. We focused our attention on BGC, as this measure should be more indicative of biliverdin content in eggshells (see above) and was used for the testing of the SSEC hypothesis in three recent studies (Morales et al. 2006; Moreno et al. 2006; Siefferman et al. 2006).

Statistical analyses

Separate models were fitted for effects of BGC and hue on paternal effort and nestling performance. In each of these models, the spectral quality in question of both original and cross-fostered eggs was the main factor of interest and was, therefore, always retained in the final models. To reduce unexplained variation and increase the power of the test, we added potentially important covariates as predictors in the initial models. These covariates were backward eliminated on the basis of their significance. Thus, only significant covariates remained in the final models. Different classes of covariates were entered into models with different response variables. In models where the response variable was feeding rate, we controlled for hour, ambient temperature and brood size when feeding frequency was recorded. We fitted separate models for paternal feeding frequency at the two ages of young, as these were only weakly correlated ($r=0.222$, $n=45$, $p=0.143$). In models where the response variable was nest defence, we controlled for offspring age and brood size on the day of the experiment. In models where the response variable was hatching/fledging success or nestling immunity, we controlled for laying date of original clutch (1st May=1), total feeding frequency at age 12 days, number of parents attending the nest, presence of mites (binary variable) and the difference between original and cross-fostered clutch size. We included the difference in clutch sizes in the model to control for a potential effect of enlargement or reduction of clutch size caused by our cross-fostering design. We also included the number of parents attending the nest in addition to total feeding frequency, as we assumed that a parent rearing the brood alone may reduce the quality or quantity of food delivered per visit. In models where the response variable was fledging weight or tarsus length, we included all covariates as in the former models, as well as midweight or midtarsus of genetic parents, respectively. The weight of parents was adjusted for nestling age when a parent was captured by including residuals from the regression of parental weight on nestling

age, rather than entering the actual parental weight into the model (*female weight* $= 14.39 - 0.14 \times age$; $F_{1,53}=19.27$, $p<0.001$; *male weight* $= 13.66 - 0.09 \times age$; $F_{1,47}=7.13$, $p=0.010$). The reason for the inclusion of the parental characteristic into these models is that morphological traits are usually highly heritable (Merilä and Sheldon 2001), which means they would affect the response variable. If parental morphology was correlated with the colour of the original eggs, its inclusion among predictors would be needed to avoid spurious results (Krist and Remeš 2004). We used brood means in models investigating the effect of predictors on nestling performance (immunity, weight and tarsus length). We did not adjust nestling weight for nestling tarsus length, as these variables were only weakly correlated ($r=0.205$, $n=50$, $p=0.152$). Only the young that subsequently fledged were included in the computation of brood means.

To investigate the relationship between female characteristics and the colour of the eggs she lays, we computed correlations between spectral qualities of original eggs and female age, morphology (tarsus length and condition) and some reproductive parameters (egg size, clutch size and laying date). We knew the exact age of 37 females that we had ringed as nestlings in previous years. As this variable was not normally distributed, we assessed the relationship between female age and egg colour by non-parametric Spearman rank correlation. We decided to use female condition rather than weight for analyses, as the latter was correlated with tarsus length ($r=0.371$, $n=54$, $p<0.001$). Thus, condition was determined as residuals from the regression of female weight on tarsus length and age of young at female capture (*weight* $= -2.748 + 0.857 \times tarsus - 0.155 \times age$; $F_{2,51}=20.89$, $p<0.001$). To test for a potential association between egg colour and maternal effort (intensity of nest defence and feeding frequency), we fitted models where the colour of the original eggs was an independent factor of interest, and maternal effort was the response variable. We fitted models instead of computing correlations because it was easier to control for covariates in the former statistical design. In females, feeding frequencies recorded at the two ages of young were correlated ($r=0.465$, $n=45$, $p=0.001$). Therefore, we performed a principal components analysis (PCA) and used the first principal component (73.3% of variation explained) as a measure of maternal feeding frequency. Consequently, we included in initial models, hour (daytime), ambient temperature and brood size recorded on both days of measuring the feeding frequency. In the model for female intensity of nest defence, we included the same class of covariates as for male nest defence.

To investigate the effect of laying order on the two spectral measures, we first subtracted clutch mean from the actual value of the spectral measure (e.g. actual BGC of the sixth egg minus average BGC of eggs in that nest). By this method (centering), we received values that were more comparable between clutches and, consequently, the power of statistical tests relating to intraclutch variation in egg colour was increased. We fitted two types of models. In the first group, we used actual laying order as the predictor variable. In the second group, we used relative laying order, which may be more appropriate when different clutch sizes are pooled into a single analysis. In our sample, clutch size varied between 4–8. We therefore categorized eggs into the following categories: last, penultimate, pre-penultimate and precedent eggs. In the precedent category, one to four eggs were pooled depending on clutch size.

We fitted models with continuous response variables in JMP (SAS Institute 1995) and models with categorical or ordinal response in PROC GENMOD (SAS Institute 1999). We determined repeatabilities from variance components (PROC VARCOMP; SAS Institute 1999); the accompanying *F*-tests are the results of one-way analysis of variance. We determined prospective power of our tests in PROC POWER (SAS Institute 1999). When hatching and fledging success were the response variables, these were included in the models in event/trial syntax where the number of eggs/hatchlings in individual nests were trials, and the number of hatchlings/fledglings in individual nests were events in models for hatching and fledging success, respectively. These models had binomial error structure, logit link function and statistics corrected for overdispersion by Pearson chi-square/*df*. We assessed the paternal effort in nest defence in two models: (1) The response variable was the overall intensity of nest defence (ordinal variable; multinomial error structure and cumulative logit link function). (2) In the subgroup that attacked the dummy by dive or contact, the number of attacks and the latency to the first attack were correlated ($r=-0.431$, $n=23$, $p=0.040$). Therefore, we performed a PCA and used the first principal component (71.6% variability explained) for subsequent analyses. Greater values of PC1 indicate more attacks and a shorter latency to the first attack.

The colour of original and cross-fostered eggs was correlated within individual nests (BGC: $r=0.278$, $n=70$, $p=0.020$; hue: $r=0.175$, $n=70$, $p=0.147$). This might be caused by the fact that we cross-fostered eggs between pairs of nests with the same laying date (which is unavoidable in this sort of experimental study) and, due to time constraints, usually between nests that were located in the same of the four study plots. Correlations among predictors in multiple regression (collinearity) might reduce the power of statistical tests (Quinn and Keough 2002). To assess the effect of collinearity on our results, we looked at variance inflation factors (VIF) for our predictors in individual models with continuous response variable. Predictors with VIFs <10 are generally accepted as giving unbiased results (Quinn and

Keough 2002). Recently, it has been suggested that VIF as small as two might substantially bias results (Graham 2003). All VIFs were smaller than 1.2 in our analyses, therefore, collinearity did not seriously bias our results. It is important to realize that despite the positive correlation between the colour of original and cross-fostered eggs, more than 92% of variation in cross-fostered BGC was uniquely generated by our cross-fostering experiment (i.e. this part of variation was independent on colour of original eggs). The respective figure for cross-fostered hue is 97%.

Results

The reflectance spectra of eggs of the collared flycatcher have a bimodal shape. The major peak of reflectance lays in the blue-green part of the spectrum which agrees with the human-perceived colour of these eggs, whereas the minor one lays in the UV part of the spectrum which is invisible to humans (Fig. 1a). Both hue and BGC changed significantly throughout the laying order. Relative hue increased linearly ($F_{1,445}=100.2$, $p<0.001$, $R^2=0.184$; Fig. 1c), whereas relative BGC decreased non-linearly in the laying order (quadratic regression; $F_{2,444}=241.1$, $p<0.001$, $R^2=0.520$; Fig. 1b). These relationships hold even when the relative laying order is used for these tests (Hue: $F_{1,445}=95.4$, $p<0.001$, $R^2=0.177$; BGC-quadratic regression: $F_{2,444}=365.1$, $p<0.001$, $R^2=0.622$).

Despite the significant intra-clutch trends, both spectral measures were also significantly repeatable within clutches (BGC: $r=0.509$, $F_{69,377}=6.71$, $p<0.001$; Hue: $r=0.315$, $F_{69,377}=4.05$, $p<0.001$), which is a premise of signalling function at a clutch-level. Therefore, we used clutch means of spectral measures in subsequent analyses. The summed reflectance in the UV part of the spectrum was most variable between clutches (SD=326.8). We use SD instead of coefficient of variation as an estimate of variability in particular segments of spectrum, as the use of the latter parameter would be statistically flawed (see Dale 2006). Summed reflectance in the red (286.2), blue (248.0) and green (235.1) part of the spectrum was less variable. The above differences in standard deviations should not be greatly affected by precision differences of measuring apparatus in particular parts of spectra, as they are based on repeated measurements of eggs in a clutch (8–16 measurements depending on clutch size).

Fig. 1 Relationship between laying order and **a** shape of mean reflectance spectra (laying order is indicated by numbers alongside the reflectance curves) and **b**–**c** individual spectral qualities (centered within clutches); lines of best fit are displayed. See text for statistical tests
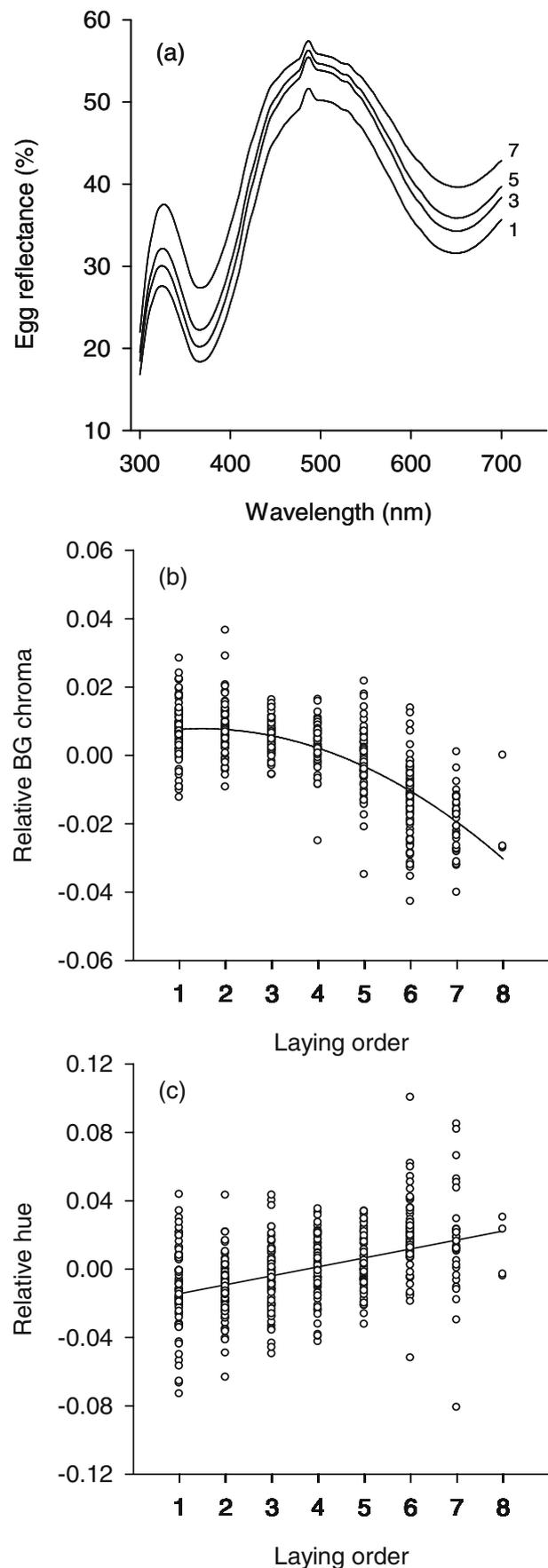
Egg colour and paternal effort

Three nests were not attended by the male at age 6 days and five at age 12 days (including the three nests that were already not attended at age 6 days). One nest was not attended by the female at age 12 days. We did not include these nests into the analyses of feeding frequencies presented below, as they might be caused either by polygyny or by predation on adults, which are factors that might have little relevance to the SSEC hypothesis. Moreover, inclusion of these cases would also be problematic from the statistical point of view because they would represent outliers in analyses of feeding frequencies. Nevertheless, nests attended by males at age 12 days did not differ from those where males were absent (BGC_{CF}: $t=$ 1.97, $p=0.055$; BGC_{OR}: $t=0.70$, $p=0.488$; Hue_{CF}: $t=-0.76$, $p=0.451$; Hue_{OR}: $t=0.68$, $p=0.500$; $df=49$ in all cases; subscripts CF and OR refer to cross-fostered and original eggs, respectively). The only result approaching statistical significance was in the opposite direction than predicted by the SSEC hypothesis, as BGC_{CF} was lower in nests attended by the male (mean±SE, $0.603±0.0019$) than in nests without males ($0.615±0.0057$).

In nests attended by males, paternal feeding frequency at nestling age 6 days was not related to any spectral quality of original or cross-fostered eggs, whereas it was negatively affected by hour of observation (Table 1; Fig. 2a). Similarly, no spectral quality predicted paternal feeding frequency at age 12 days, whereas this was positively affected by brood size (Table 1). The only result approaching significance was in the opposite direction than predicted by SSEC hypothesis. Males fed young less frequently in nests where original eggs were highly saturated (Fig. 2b). The overall intensity of nest defence by males was not affected by any spectral quality of original or cross-fostered eggs (Table 1; Fig. 2c). In the subgroup of males that attacked the dummy, BGC_{OR} and brood size negatively affected the intensity of attacks (Table 1; Fig. 2d), whereas the hue of both cross-fostered and original eggs did not predict the intensity of attacks (Table 1).

Egg colour and offspring performance

Out of 70 experimental clutches, 56 hatched. Complete hatching failure was caused by marten (*Martes martes*) predation of incubating female (12 cases) and by clutch infertility (two cases) and, therefore, has no relevance to the SSEC hypothesis. In 56 hatched clutches, 339 eggs hatched and 20 failed to hatch. No variable was predictive of hatching success in these nests (Table 2). Out of 56 hatched broods, 50 fledged and six suffered total failure. Egg colour did not differ in the two categories of nests (BGC_{CF}: $t=$ -1.20, $p=0.236$; BGC_{OR}: $t=0.41$, $p=0.681$; Hue_{CF}: $t=0.27$,

Table 1 Effects of BGC and hue of cross-fostered (subscript CF) and original (subscript OR) eggs on four measures of paternal effort. Presented models include statistically significant covariates

	Model number	DF	F*	P
Feeding frequency 6				
BGC_{CF}	1	1, 49	2.76	0.103
BGC_{OR}	1	1, 49	0.02	0.899
Time of day	1	1, 49	5.50	0.023
Hue_{CF}	2	1, 49	0.77	0.383
Hue_{OR}	2	1, 49	<0.01	0.993
Time of day	2	1, 49	5.15	0.028
Feeding frequency 12				
BGC_{CF}	3	1, 41	0.70	0.409
BGC_{OR}	3	1, 41	3.57	0.066
Brood size	3	1, 41	7.08	0.011
Hue_{CF}	4	1, 41	0.23	0.633
Hue_{OR}	4	1, 41	1.89	0.177
Brood size	4	1, 41	7.82	0.008
Intensity of nest defence				
BGC_{CF}	5	1, 47	2.00	0.157
BGC_{OR}	5	1, 47	0.92	0.336
Hue_{CF}	6	1, 47	0.02	0.898
Hue_{OR}	6	1, 47	0.04	0.850
PC1 attacks				
BGC_{CF}	7	1, 19	0.64	0.435
BGC_{OR}	7	1, 19	10.48	0.004
Brood size	7	1, 19	5.34	0.032
Hue_{CF}	8	1, 20	0.55	0.469
Hue_{OR}	8	1, 20	0.06	0.813

* In models 5 and 6 χ^2 is presented

$p=0.786$; Hue_{OR}: $t=-1.79$, $p=0.079$; $df=54$ in all cases). Within 50 fledged nests, 249 young fledged and 52 died. Fledging success was lower in nests that were attended by only one parent, infested by mites and initiated late in the season. In contrast to these significant effects of covariates, no spectral quality predicted fledging success (Table 2).

Nestling tarsus length was positively affected by mid-tarsus length of genetic parents and negatively affected by laying date in both models. From the spectral qualities, only BGC_{OR} had a significant and positive effect on nestling tarsus length (Table 2; Fig. 3b). Nestling weight (Fig. 3a) and T cell mediated immunity (Fig. 3c) were not affected by either covariates or spectral qualities (Table 2).

Original eggs and female traits

Female age was positively associated with BGC of original eggs ($r_s=0.399$, $n=37$, $p=0.014$; Fig. 4a), whereas hue of original eggs was unrelated to female age ($r_s=0.214$, $n=37$, $p=0.204$). The relationship between female morphology and the colour of eggs that she laid was marginally non-significant (condition: BGC_{OR}, $r=0.248$, $p=0.070$, Fig. 4b;

Fig. 2 Relationship between BG chroma of original (*solid diamonds, solid lines*) and cross-fostered (*open circles, dashed lines*) eggs and four measures of paternal effort. Regression lines are displayed. The displayed data and fitted regression lines are uncontrolled for effects of covariates. See Table 1 for statistical tests
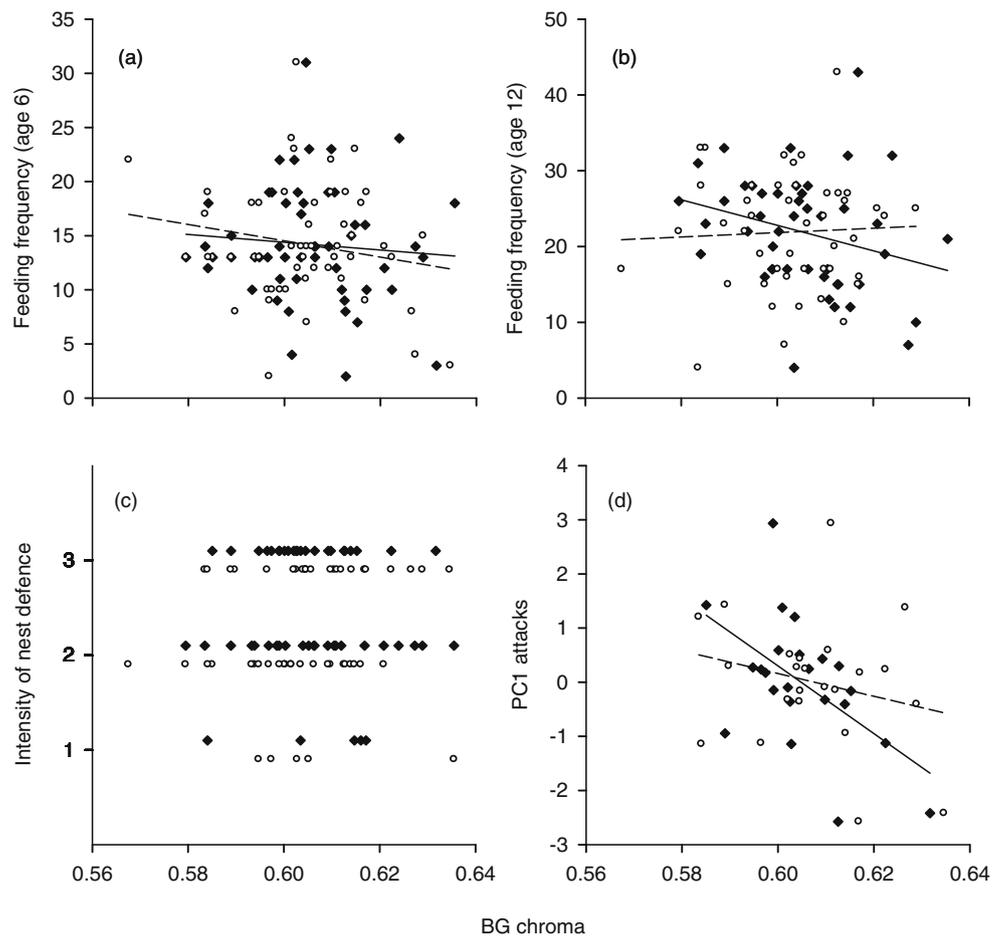
Hue$_{OR}$, $r=-0.136$, $p=0.328$; tarsus length: BGC$_{OR}$, $r=0.251$, $p=0.068$; Hue$_{OR}$, $r=0.169$, $p=0.223$; $n=54$ in all cases).

Female feeding frequencies were positively affected by brood size and hour of feeding at nestling age 6 days but unrelated to colour of original eggs (BGC$_{OR}$: $F_{1,42}=1.17$, $p=0.286$, brood size at age six: $F_{1,42}=21.99$, $p<0.001$, Fig. 4c; Hue$_{OR}$: $F_{1,41}=0.24$, $p=0.629$, brood size at age six: $F_{1,41}=17.43$, $p<0.001$, hour of feeding at age six: $F_{1,41}=4.47$, $p=0.041$). Neither spectral quality of original eggs predicted the overall intensity of nest defence by females: $\left(BGC_{OR}: \chi^2_{1,52}=0.06, p=0.809;\ Hue_{OR}: \chi^2_{1,52}=0.24, p=0.62\right)$.

Mean egg volume of the clutch was positively correlated with BGC of that clutch ($r=0.261$, $n=70$, $p=0.029$; Fig. 4d), but was unrelated to the hue of that clutch ($r=0.114$, $n=70$, $p=0.348$). In contrast to mean egg volume, neither clutch size (BGC$_{OR}$: $r=-0.035$, $n=70$, $p=0.772$; Hue$_{OR}$: $r=0.113$, $n=70$, $p=0.350$) nor laying date (BGC$_{OR}$: $r=0.156$, $n=70$, $p=0.199$; Hue$_{OR}$: $r=0.023$, $n=70$, $p=0.853$) was related to egg colour.

The relationships that we found between egg colour and female morphology and egg volume were probably not mediated through female age, as the age was not related to

any of the former variables (condition: $r_s=0.053$, $n=37$, $p=0.758$; tarsus length: $r_s=-0.002$, $n=37$, $p=0.989$; egg volume: $r_s=0.069$, $n=37$, $p=0.686$).

Discussion

In our experiments, we found that egg colour changed in the laying order; hue increased linearly, whereas BGC decreased non-linearly in the course of laying. We also found correlations between egg colour and female traits. Eggs with higher BGC were laid by older females and, independently of this, by females tended to be in better condition and to have longer tarsi. Saturation of egg colour was also positively associated with the mean egg volume that females laid. These findings are consistent with the assumptions of the SSEC hypothesis (Moreno and Osorno 2003). Firstly, the blue biliverdin pigment colouring eggs may be a limited source for laying females. Secondly, female quality was positively associated with egg colour. However, in contrast to the tentative support for the assumptions of the SSEC hypothesis, we have found no evidence in support of its prediction. None of our three

Table 2 Effects of BGC and hue of cross-fostered (subscript CF) and original (subscript OR) eggs on five measures of offspring performance. Presented models include significant covariates

	Model number	DF	F	P
Hatching success				
BGC$_{CF}$	1	1, 53	0.66	0.419
BGC$_{OR}$	1	1, 53	1.74	0.192
Hue$_{CF}$	2	1, 53	0.90	0.347
Hue$_{OR}$	2	1, 53	0.25	0.619
Fledging success				
BGC$_{CF}$	3	1, 45	0.04	0.838
BGC$_{OR}$	3	1, 45	0.64	0.426
Laying date	3	1, 45	7.51	0.006
Number of parents	3	1, 45	18.41	<0.001
Hue$_{CF}$	4	1, 36	2.51	0.122
Hue$_{OR}$	4	1, 36	0.22	0.641
Number of parents	4	1, 36	12.01	0.001
Presence of mites	4	1, 36	4.27	0.046
Nestling tarsus length				
BGC$_{CF}$	5	1, 32	0.59	0.449
BGC$_{OR}$	5	1, 32	6.67	0.015
Laying date	5	1, 32	15.47	<0.001
Parental midtarsus	5	1, 32	9.60	0.004
Hue$_{CF}$	6	1, 32	1.73	0.198
Hue$_{OR}$	6	1, 32	1.30	0.264
Laying date	6	1, 32	12.33	0.001
Parental midtarsus	6	1, 32	9.67	0.004
Nestling weight				
BGC$_{CF}$	7	1, 47	1.77	0.190
BGC$_{OR}$	7	1, 47	0.98	0.326
Hue$_{CF}$	8	1, 47	1.02	0.319
Hue$_{OR}$	8	1, 47	2.52	0.119
Nestling immunity				
BGC$_{CF}$	9	1, 47	<0.01	0.952
BGC$_{OR}$	9	1, 47	0.33	0.566
Hue$_{CF}$	10	1, 47	2.78	0.102
Hue$_{OR}$	10	1, 47	2.81	0.100

measures of paternal care was associated with the colour of cross-fostered eggs. Furthermore, we did not find any positive association between paternal care and colour of original eggs. The only significant relationship was in the unexpected direction, as males defended nests more intensely in which original eggs were less saturated.

Similar to our study, Moreno et al. (2005) previously found that eggs decrease in colouration with laying order in the closely related pied flycatcher. Both adaptive and non-adaptive explanations were previously proposed to explain the pale colour of last eggs. Yom-Tov (1980) suggested that
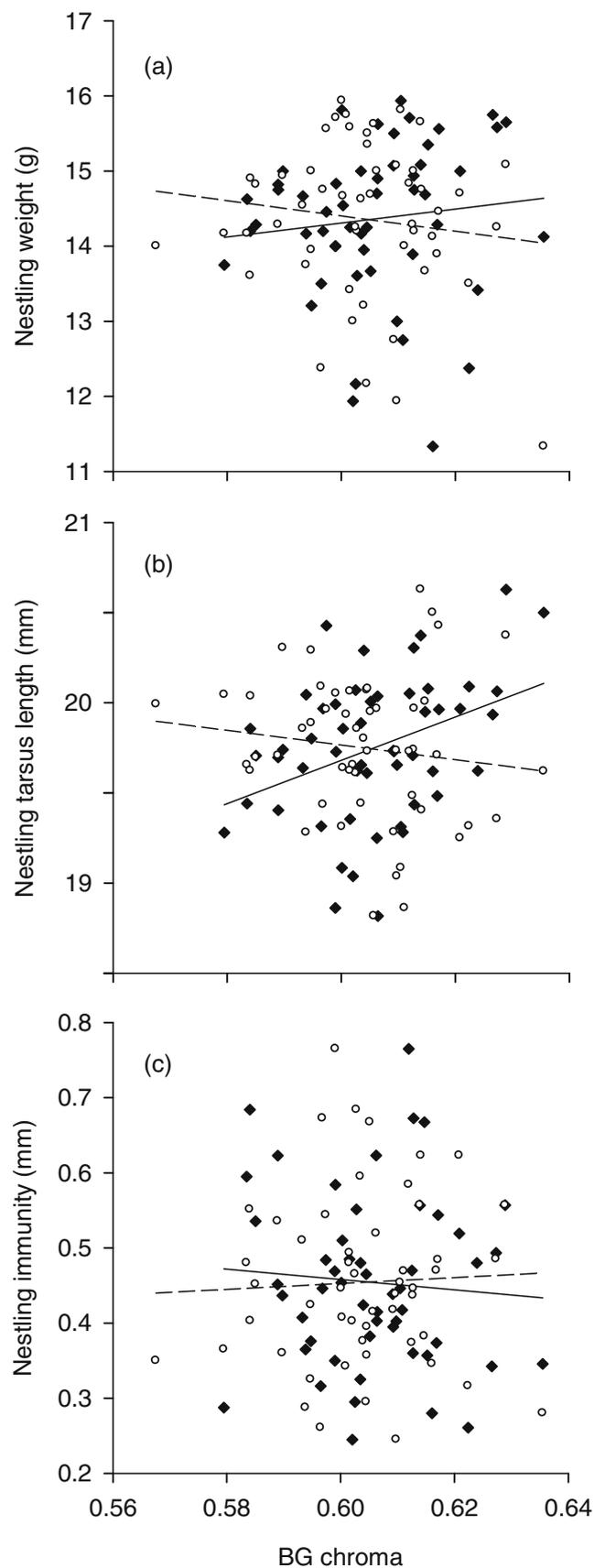
Fig. 3 Relationship between BG chroma of original (*solid diamonds and lines*) and cross-fostered (*open circles, dashed lines*) eggs and three measures of nestling performance. The displayed data and fitted regression lines are uncontrolled for effects of covariates. See Table 2 for statistical tests

Fig. 4 Relationship between BG chroma of original eggs and female age (**a**), condition (**b**), feeding frequency (**c**) and mean egg volume (**d**). The displayed data and *fitted lines* are uncontrolled for effects of covariates. The standard (reduced) major axes from model II regressions (Sokal and Rohlf 1995) are fitted instead of ordinary regression lines, as there is no clear causality between the pairs of variables. See text for statistical tests
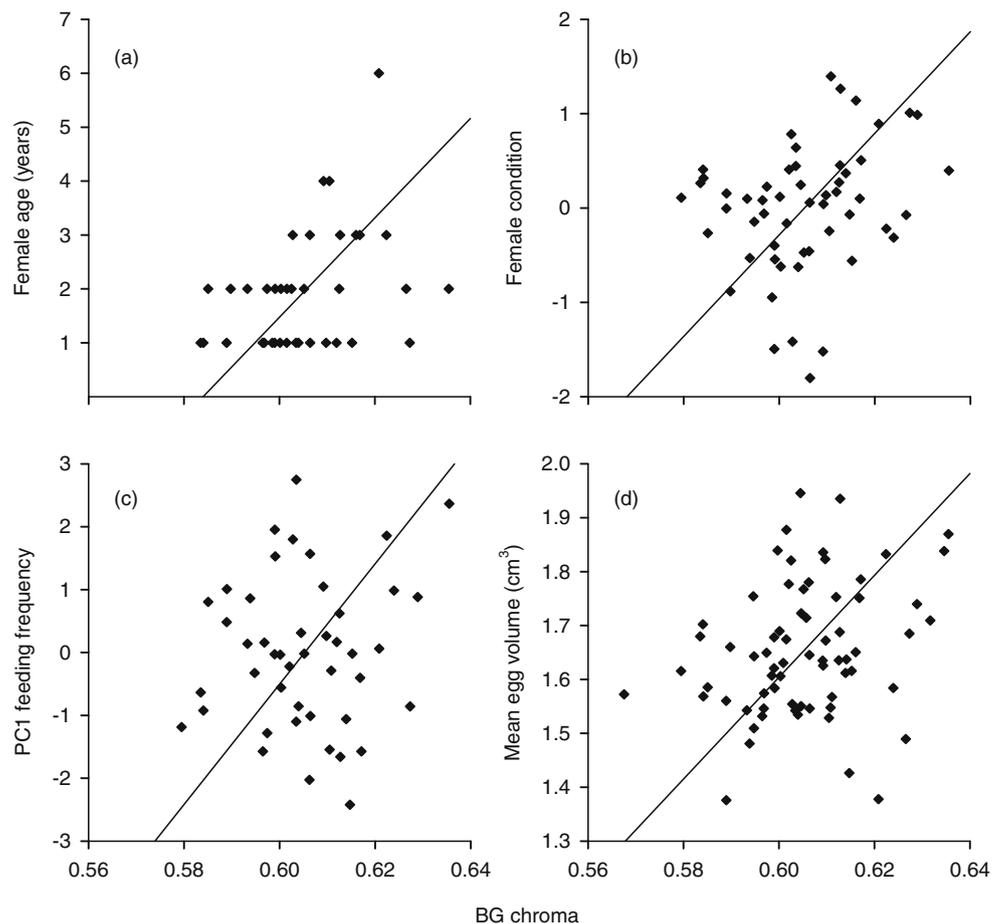
laying a pale last egg may be adaptive for females to prevent conspecific brood parasitism. By laying a pale last egg, females might indicate to conspecific females that they are already incubating the clutch. Conspecifics should then avoid laying a parasitic egg in the nest which contains a pale egg, as the probability of successful hatching would be diminished. Ruxton et al. (2001) formalized this idea and showed that such signalling may indeed theoretically work under some conditions. An alternative, non-adaptive explanation for the occurrence of pale eggs at the end of the laying sequence is that pigments deposited in eggshells are depleted in the course of laying (Nice 1937). As a gradual change in eggshell hue was detected, our data, therefore, do not provide support for the adaptive scenario of Yom-Tov (1980) which predicts only the last egg to be of a different appearance (Ruxton et al. 2001). Blue-green chroma decreased at an accelerating rate in the laying order; however, the decrease was evident as early as in the penultimate egg (results not shown). Moreover, in our population, there is no evidence for conspecific brood parasitism (see Krist et al. 2005) which suggests little opportunity for the evolution of such an anti-parasitic signalling system. However, the decrease in colour of eggs

with laying order may still be adaptive for females for other reasons. For example, if pigments strengthen eggshells (Gosler et al. 2005), females may allocate more pigments to early laid eggs because the risk of breakage may be greater for them. As our study was not designed to test for such a possibility, we can conclude that our data are only consistent with the assumption of the SSEC hypothesis that eggshell pigments are a limited source for females (see Moreno and Osorno 2003).

To be an honest signal for males, egg colour should be correlated with female and/or offspring quality (Moreno and Osorno 2003). We have found that females laying presumably more costly (saturated) eggs tended to be in better condition and to have longer tarsi. They also laid larger eggs and were older than females laying less saturated eggs. These results are generally in the direction expected under the SSEC hypothesis and as such provide some support for its assumption. For example, egg size has been suggested to be an indicator of female quality in another hole-nester with altricial young, the great tit (Hörak et al. 1997). Older individuals may be regarded as superior to young ones either due to increased breeding experiences (Cichoń 2003) or because they have already proved their

survival abilities (Mauck et al. 2004). Furthermore, we have found that young that hatched in nests where originally laid eggs were more saturated had longer tarsi, which is a trait positively associated with fitness in this species (Kruuk et al. 2001). As this finding was independent of total feeding frequency of the young, it suggests that some unmeasured type of parental or territory quality is associated with egg colour. However, as the colour of cross-fostered eggs did not affect offspring performance, we can exclude the possibility that egg colour was associated with offspring genetic quality or superior egg composition which would be a more direct support for the SSEC hypothesis.

Four previous studies looked for an association between colour of eggs and other female traits (Moreno et al. 2005, 2006; Morales et al. 2006; Siefferman et al. 2006). All of them found some evidence that egg colour depends on female quality. Similar to our study of the collared flycatcher, female condition was also positively associated with blue-green chroma of eggs in the pied flycatcher (Morales et al. 2006) and the bluebird (Siefferman et al. 2006). Siefferman et al. (2006) found that older females laid more saturated eggs, which is in agreement with our results but in contrast to those of Moreno et al. (2005) who found that females who were 5 years or older laid less pigmented eggs than younger individuals. In the pied flycatcher, immunocompetence of females (Moreno et al. 2005) and eggs (Morales et al. 2006) was positively associated with saturation of egg colour. In the same species, food-supplemented females laid more saturated eggs than control ones, which suggests a direct link between female nutritional state and egg colour (Moreno et al. 2006). Taken together, the results of these studies support the view that egg colour contains some degree of information about female and/or offspring quality.

However, in this experimental study, we did not find any support for the main prediction of the SSEC hypothesis, as no type of paternal care was related to colour of cross-fostered eggs. Perhaps the most parsimonious explanation for this negative result would be the flycatchers' inability to perceive subtle differences in egg colour in poor light conditions inside cavities. However, neither was paternal care positively related to saturation of original eggs. Taken together, these two negative results are in contrast to those of a previous correlative study on the pied flycatcher (Moreno et al. 2004) in which males fed young hatching from more saturated eggs with a higher frequency. Therefore, we need to look for such explanations of our negative results that would be consistent with positive results of Moreno et al. (2004). There is a number of such potential explanations.

First, the two studies were carried out on different species, and it is possible that differential allocation of paternal effort evolved in one but not in the other species.

However, we consider such an explanation unlikely. The collared flycatcher is a sister species of the pied flycatcher from which it divided in relatively recent past, and the reproductive isolation between the two species is currently incomplete (Saetre et al. 1997).

Second, the discordance might be caused simply by sampling effect, either by a type II error in our study or a type I error in Moreno et al. (2004) study. However, as we have used a somewhat larger sample size than Moreno et al. (2004), statistical power to detect effect of the magnitude detected by these authors ($r=0.38$ and $r=0.41$) was relatively high in our study (feeding frequency at age 6 days: power=0.82 and 0.88 for the two effect sizes, respectively; feeding frequency at age 12 days: power=0.75 and 0.82, respectively). Type I error rate in the study of Moreno et al. (2004) has been controlled at the conventional level ($\alpha<0.05$). Therefore, this explanation for the discordance results is also unlikely.

Third, in the study of Moreno et al. (2004), males cared for their own nestlings, whereas in our study, parents cared for unrelated cross-fostered offspring. This could, in theory, cause the absence of predicted effects of egg colour on paternal care in our study. However, such an explanation is highly unlikely because birds were so far never shown to discriminate against unrelated conspecific nestlings hatched in their nests (Kempenaers and Sheldon 1996).

Fourth, whereas one of our three indicators of paternal effort was exactly the same as the one of Moreno et al. (2004), feeding rates to 12-day-old nestlings recorded for an hour, the description of egg colour differed between the two studies. Moreno et al. (2004) recorded egg reflectance in the human-visible spectrum only, whereas we also measured egg reflectance in the UV spectrum (301–400 nm). This difference might be of importance, as most passerines see also in UV part of spectrum (Cuthill et al. 2000). This seems to be also the case in the pied flycatcher, as females in this species mate preferentially with males with high UV reflectance of plumage (Siitari et al. 2002). To test directly for the possibility that different spectral range caused discordance results of the two studies, we computed a second estimate of blue-green chroma ($BGC_{401-700}$) that was restricted to the spectral range used by Moreno et al. (2004). $BGC_{401-700}$ was calculated as the reflectance between 401–600 nm divided by the total reflectance between 401–700 nm. This estimate was very strongly correlated with our original estimate of BGC, which was based on the spectral range 301–700 nm ($r=0.976$, $n=70$, $p<0.001$). Such high correlation means, at least in the context of biliverdin signalling, that the information contained in the UV part of the spectrum is redundant to that of human-visible part of spectrum. Consequently, different results of the two studies cannot be explained by the use of different spectral range.

Fifth, it may be possible that males increase paternal care only when superior egg colour is accompanied by superior female quality. In a natural situation, these traits may covary, whereas our cross-fostering approach separated them. However, this would mean that in such a scenario, egg colour per se is not predictive of paternal effort.

Finally, it may be that the two flycatchers' populations diverged in the type of paternal care that is differentially allocated. In the Spanish population, feeding frequency seems to be the plastic trait with an effect on offspring performance (Moreno et al. 2004), whereas in our population, some dimension of paternal care that we did not measure (for example quality of diet) might evolve this plasticity. Remarkably, feeding frequency was important for offspring performance in Spanish population (Moreno et al. 2004), whereas we did not detect effect of total feeding frequency on any offspring trait. This hypothesis requires further testing.

In conclusion, results of the studies performed so far have given evidence supporting the view that the egg colour is informative about female and offspring quality. Such a correlation is a critical assumption of the SSEC hypothesis (Moreno and Osorno 2003). However, in this experimental study, we have found no evidence for increased paternal effort with more saturated egg colour. This suggests that the correlation between female quality and egg colour is not strong enough to select for the probably costly discriminatory abilities in males to allocate paternal care differently according to varying egg colour. Males can perhaps assess their offsprings' reproductive value more precisely and allocate their care more appropriately according to some more direct and informative cue, such as offspring size. This explanation implies that either egg colouration is not very costly or serves other functions (e.g. eggshell strengthening; Gosler et al. 2005). However, as our results are in contrast with those of previous studies on related species and the egg colouration could, in theory, be even more informative about female quality than colouration of its plumage (see "Introduction"), tests of the SSEC hypothesis might be a fruitful area for further research.

References

Amundsen T (2000) Why are female birds ornamented? Trends Ecol Evol 15:149–155

Andersson S, Prager M (2006) Quantifying colors. In: Hill GE, McGraw K (eds) Bird coloration I: mechanisms and measurements. Harvard University Press, Cambridge, pp 41–89

Birkhead TR (1978) Behavioural adaptations to high density nesting in the common guillemot *Uria aalge*. Anim Behav 26:321–331

Burley N (1986) Sexual selection for aesthetic traits in a species with biparental care. Am Nat 127:415–445

Cichoń M (2003) Does prior breeding experience improve reproductive success in collared flycatcher females? Oecologia 134:78–81

Cuthill IC, Partridge JC, Bennett ATD, Church SC, Hart NS, Hunt S (2000) Ultraviolet vision in birds. Adv Study Behav 29:159–214

Dale J (2006) Intraspecific variation in coloration. In: Hill GE, McGraw K (eds) Bird coloration II: function and evolution. Harvard University Press, Cambridge, pp 36–86

Davies NB, Brooke ML (1988) Cuckoos versus reed warblers: adaptations and counteradaptations. Anim Behav 36:262–284

Endler JA (1990) On the measurement and classification of color in studies of animal color patterns. Biol J Linn Soc 41:315–352

Falchuk KH, Contin JM, Dziedzic TS, Feng ZL, French TC, Heffron GJ, Montorzi M (2002) A role for biliverdin IX alpha in dorsal axis development of *Xenopus laevis* embryos. Proc Natl Acad Sci USA 99:251–256

Gosler AG, Higham JP, Reynolds SJ (2005) Why are birds' eggs speckled? Ecol Lett 8:1105–1113

Götmark F (1992) Blue eggs do not reduce nest predation in the song thrush *Turdus philomelos*. Behav Ecol Sociobiol 30:245–252

Graham MH (2003) Confronting multicollinearity in ecological multiple regression. Ecology 84:2809–2815

Griggio M, Valera F, Casas A, Pilastro A (2005) Males prefer ornamented females: a field experiment of male choice in the rock sparrow. Anim Behav 69:1243–1250

Grim T (2005) Host recognition of brood parasites: implications for methodology in studies of enemy recognition. Auk 122:530–543

Grim T, Honza M (2001) Differences in behaviour of closely related thrushes (*Turdus philomelos* and *T. merula*) to experimental parasitism by the common cuckoo *Cuculus canorus*. Biologia 56:549–556

Hill GE (2002) A red bird in a brown bag. Oxford University Press, New York

Hörak P, Mand R, Ots I (1997) Identifying targets of selection: a multivariate analysis of reproductive traits in the great tit. Oikos 78:592–600

Kaur H, Hughes MN, Green CJ, Naughton P, Foresti R, Motterlini R (2003) Interaction of bilirubin and biliverdin with reactive nitrogen species. FEBS Lett 543:113–119

Kempenaers B, Sheldon BC (1996) Why do male birds not discriminate between their own and extra-pair offspring? Anim Behav 51:1165–1173

Krist M (2004) Importance of competition for food and nest-sites in aggressive behaviour of collared flycatcher *Ficedula albicollis*. Bird Study 51:41–47

Krist M, Remeš V (2004) Maternal effects and offspring performance: in search of the best method. Oikos 106:422–426

Krist M, Nádvorník P, Uvírová L, Bureš S (2005) Paternity covaries with laying and hatching order in the collared flycatcher *Ficedula albicollis*. Behav Ecol Sociobiol 59:6–11

Kruuk LEB, Merilä J, Sheldon BC (2001) Phenotypic selection on a heritable size trait revisited. Am Nat 158:557–571

Mauck RA, Huntington CE, Grubb TC (2004) Age-specific reproductive success: evidence for the selection hypothesis. Evolution 58:880–885

McDonagh AF (2001) Turning green to gold. Nat Struct Biol 8:198–200

McGraw KJ (2005) The antioxidant function of many animal pigments: are there consistent health benefits of sexually selected colourants? Anim Behav 69:757–764

McNett GD, Marchetti K (2005) Ultraviolet degradation in carotenoid patches: live versus museum specimens of wood warblers (Parulidae). Auk 122:793–802

Merilä J, Sheldon BC (2001) Avian quantitative genetics. In: Nolan V (ed) Current ornithology. Kluwer, New York, pp 179–255

Mikšík I, Holáň V, Deyl Z (1996) Avian eggshell pigments and their variability. Comp Biochem Physiol B 113:607–612

Moksnes A, Røskaft E, Tysse T (1995) On the evolution of blue cuckoo eggs in Europe. J Avian Biol 26:13–19

Montgomerie R (2006) Analyzing colors. In: Hill GE, McGraw K (eds) Bird coloration I: mechanisms and measurements. Harvard University Press, Cambridge, pp 90–147

Morales J, Sanz JJ, Moreno J (2006) Egg colour reflects the amount of yolk maternal antibodies and fledging success in a songbird. Biol Lett 2:334–336. DOI 10.1098/rsbl.2006.0471

Moreno J, Osorno JL (2003) Avian egg colour and sexual selection: does eggshell pigmentation reflect female condition and genetic quality? Ecol Lett 6:803–806

Moreno J, Osorno JL, Morales J, Merino S, Tomás G (2004) Egg colouration and male parental effort in the pied flycatcher Ficedula hypoleuca. J Avian Biol 35:300–304

Moreno J, Morales J, Lobato E, Merino S, Tomás G, Martínez-de la Puente J (2005) Evidence for the signaling function of egg color in the pied flycatcher Ficedula hypoleuca. Behav Ecol 16:931–937

Moreno J, Lobato E, Morales J, Merino S, Tomás G, Martínez-de la Puente J, Sanz JJ, Mateo JM, Soler JJ (2006) Experimental evidence that egg color indicates female condition at laying in a songbird. Behav Ecol 17:651–655

Nice MM (1937) Studies in the life history of the song sparrow, volume I. A population study of the song sparrow. Trans Linn Soc N Y 4:1–247

Øien IJ, Moksnes A, Røskaft E (1995) Evolution of variation in egg color and marking pattern in European passerines: adaptations in a coevolutionary arms race with the cuckoo, Cuculus canorus. Behav Ecol 6:166–174

Pilastro A, Griggio M, Matessi G (2003) Male rock sparrows adjust their breeding strategy according to female ornamentation: parental or mating investment? Anim Behav 66:265–271

Pizzari T, Cornwallis CK, Lovlie H, Jakobsson S, Birkhead TR (2003) Sophisticated sperm allocation in male fowl. Nature 426:70–74

Quinn GP, Keough MJ (2002) Experimental design and data analysis for biologists. Cambridge University Press, Cambridge

Ruxton GD, Broom M, Colegrave N (2001) Are unusually colored eggs a signal to potential conspecific brood parasites? Am Nat 157:451–458

Saetre GP, Moum T, Bureš S, Král M, Adamjan M, Moreno J (1997) A sexually selected character displacement in flycatchers reinforces premating isolation. Nature 387:589–592

Saks L, McGraw K, Hõrak P (2003) How feather colour reflects its carotenoid content. Funct Ecol 17:555–561

SAS Institute Inc. (1995) JMP Statistics and graphics guide. SAS Institute Inc., Cary

SAS Institute Inc. (1999) SAS Online Doc, version 8. SAS Institute Inc., Cary

Sealy SG, Neudorf DL, Hobson KA, Gill SA (1998) Nest defense by potential hosts of the brown-headed cowbird: methodological approaches, benefits of defense, and coevolution. In: Rothstein SI, Robinson SK (eds) Parasitic birds and their hosts. Oxford University Press, New York, pp 194–211

Sheldon BC (2000) Differential allocation: tests, mechanisms and implications. Trends Ecol Evol 15:397–402

Siefferman L, Navara KJ, Hill GE (2006) Egg coloration is correlated with female condition in eastern bluebirds (Sialia sialis). Behav Ecol Sociobiol 59:651–656

Siitari H, Honkavaara J, Huhta E, Viitala J (2002) Ultraviolet reflection and female mate choice in the pied flycatcher, Ficedula hypoleuca. Anim Behav 63:97–102

Sokal RR, Rohlf FJ (1995) Biometry: the principles and practice of statistics in biological research. W. H. Freeman and Company, New York

Soler JJ, Møller AP (1996) A comparative analysis of the evolution of variation in appearance of eggs of European passerines in relation to brood parasitism. Behav Ecol 7:89–94

Soler JJ, Moreno J, Aviles JM, Møller AP (2005) Blue and green egg-color intensity is associated with parental effort and mating system in passerines: support for the sexual selection hypothesis. Evolution 59:636–644

Tinbergen N, Broekhuysen GJ, Feekes F, Houghton JCW, Kruuk H, Szulc E (1962) Egg shell removal by the black-headed gull, Larus ridibundus L.: a behaviour component of camouflage. Behaviour 19:74–117

Underwood TJ, Sealy SG (2002) Adaptive significance of egg coloration. In: Deeming DC (ed) Avian incubation: behaviour, environment, and evolution. Oxford University Press, New York, pp 280–298

Weidinger K (2001) Does egg colour affect predation rate on open passerine nests? Behav Ecol Sociobiol 49:456–464

Yom-Tov Y (1980) Intraspecific nest parasitism in birds. Biol Rev 55:93–108

Článek 9: Krist, M. (2009) Short- and long-term effects of egg size and feeding frequency on offspring quality in the collared flycatcher (*Ficedula albicollis*). *Journal of Animal Ecology*, 78, 907-918.

Short- and long-term effects of egg size and feeding frequency on offspring quality in the collared flycatcher (*Ficedula albicollis*)

Miloš Krist[1,2]

[1]*Museum of Natural History, nám. Republiky 5, 771 73 Olomouc, Czech Republic; and* [2]*Department of Zoology and Laboratory of Ornithology, Faculty of Science, Palacky University, Svobody 26, 771 46 Olomouc, Czech Republic*

Summary

1. Despite the central importance for life-history theory, egg-size effects on offspring fitness are still considered ambiguous. Most previous studies were only observational and consequently might suffer from uncontrolled correlations between egg size and parental/territory quality. Even after cross-fostering is performed, direct genetic effects and parental adjustment of post-natal care might confound our estimates of egg-size effects per se.

2. I performed a full cross-fostering experiment in the collared flycatcher (*Ficedula albicollis*) exchanging the whole clutches between pairs of nests. I statistically controlled for direct genetic effects and parental feeding frequencies. I followed young until recruitment to estimate the long-term effects of egg size and parental provisioning. In addition, I compared the effects obtained in the cross-fostering experiment with those obtained from a set of unmanipulated nests.

3. Egg size per se affected offspring morphology in both the short and long term, while having no effect on offspring survival and immunity. Egg-size effects were not confounded by parental post-natal care and direct genetic effects.

4. The number of care-givers was an influential predictor of nestling performance. Apart from the variation caused by this factor, feeding frequencies had no consistent effect on offspring performance.

5. Fitness benefits of large eggs may be difficult to establish due to variation of egg-size effects between years and habitats. Feeding frequency may affect offspring state but offspring state may also affect feeding frequency. Varying causality between feeding rate and offspring state may preclude the detection of a positive effect of the former on the latter.

Key-words: altricial chick, parental effort, parental investment, propagule, provisioning behaviour

Introduction

In theory, parents might maximize their fitness either through many small offspring with low survival prospects or through fewer large offspring with higher viability (Smith & Fretwell 1974). This trade-off between the number and size of offspring is one of the central tenets of life-history theory (Roff 1992). The trade-off is most easily studied in oviparous taxa, as investment per offspring can be easily estimated and is probably under high maternal control. In birds, clutch size is, unsurprisingly, a strong predictor of the number of fledglings

in both observational (e.g. Gustafsson & Sutherland 1988) and experimental studies (e.g. Gustafsson and Sutherland 1988; reviews of clutch enlargement studies in Roff 1992; Van der Werf 1992). A positive effect of egg size on offspring fitness is, however, more equivocal (Williams 1994). Nestling size is usually strongly determined by egg size shortly after hatching but this effect often rapidly diminishes as young grow older with no effect at the time of fledging (e.g. Smith, Ohlsson & Wettermark 1995; Krist *et al*. 2004; reviewed in Williams 1994). Moreover, most of the studies reviewed by Williams (1994) were observational. Causality of observed egg-size effects is therefore questionable. Egg size may be correlated with parental or territorial quality which could independently affect both egg size and offspring post-natal

*Correspondence author. E-mail: milos.krist@volny.cz

size or survival (Reid and Boersma 1990, Bolton 1991). Some authors even use egg size as a surrogate measure of overall female quality (Hörak, Mänd & Ots 1997).

At least three methods exist for controlling parental quality while testing for egg-size effects (Krist & Remeš 2004). First, the best solution in terms of control of confounding factors is experimental manipulation of egg size. Such studies were performed in insects (Fox 1997), reptiles (Sinervo *et al.* 1992; Olsson, Wapstra & Olofsson 2002), domesticated birds in laboratory conditions (Hill 1993; Finkler, Van Orman & Sotherland 1998) and only recently in nondomesticated birds either in the laboratory (Wagner & Williams 2007) or in the wild (Ferrari, Martinelli & Saino 2006; Alquati *et al.* 2007). A potential drawback of these studies is the low hatchability of manipulated eggs (Finkler *et al.* 1998; Alquati *et al.* 2007; Wagner & Williams 2007; but see Ferrari *et al.* 2006) which might be a side effect of an invasive experimental procedure rather than a direct consequence of a smaller egg.

Second, in within-brood comparisons (e.g. Howe 1976; Krist *et al.* 2004) parental and territory quality are the same for all broodmates and so cannot confound egg-size effect estimates (see Krist & Remeš). In theory, females might gain from differential allocation of resources to eggs in relation to embryo sex (Cordero *et al.* 2000), paternity (Krist *et al.* 2005), or hatching asynchrony (Howe 1976; Slagsvold *et al.* 1984). In reality, female ability to manipulate offspring phenotype adaptively by within-clutch differential allocation might be constrained. Within-clutch variation in egg size is usually rather small (Christians 2002) which might explain why intraclutch egg-size effects are limited (Krist *et al.* 2004).

Finally, it is possible to exchange clutches between pairs of nests that differ in egg size and to simultaneously study the effects of parental quality (as assessed by size of laid eggs) and egg size per se on offspring performance (Amundsen & Stokland 1990; Magrath 1992). Cross-fostering offers the possibility of studying experimentally the consequences of naturally occurring variability among clutches, which is usually greater than that among eggs within a clutch (Christians 2002). Consequently, this approach became a popular methodological tool for studies aimed at the trade-off between egg size and number not only in birds (19 studies performed to date – reviewed in Silva *et al.* 2007) but also in mammals (e.g. Oksanen *et al.* 2001) and insects (e.g. Lock *et al.* 2007). However, despite the advantages, this method has some potential drawbacks (Krist & Remeš 2004). First, cross-fostering does not decouple potential correlations between egg size and embryo genes. Consider for example nestling mass or tarsus length, traits which are often tested for their sensitivity to egg size. These traits are usually highly heritable (Merilä & Sheldon 2001) which means that larger mothers will have larger offspring as a result of inherited additive genes. Larger mothers have often been found to lay larger eggs (reviewed in Christians 2002). If these conditions occur together we might overestimate egg-size effects on offspring (Fox 1997; Krist & Remeš 2004). The solution to the problem may be relatively simple at least for some traits. Direct effects of genes

might be controlled for by including a parental midvalue of the studied trait among predictors of offspring traits (Krist & Remeš 2004). A second potential drawback is common to all methods – parents may adjust their post-natal care according to offspring state. They might, for example, feed chicks hatching from small eggs with a higher rate to compensate for their small initial size (Krist & Remeš 2004). In such a case, egg-size effects would be underestimated. Again, the confounding influence of post-natal parental care may be controlled for by inclusion of feeding rate among predictors of offspring performance (Krist & Remeš 2004).

Effects of feeding rate on offspring phenotype are additionally interesting in their own right. At present, feeding frequencies are studied mainly in the area of sexual selection. It is often tested whether parents adjust feeding frequency to the current value of the brood which might be affected by extra-pair paternity (review in Sheldon 2002) or mate quality (e.g. Moreno *et al.* 2004, Nakagawa *et al.* 2007). Such studies assume that feeding frequencies have an effect on the offspring but rarely test for these effects in details. Studies testing the validity of this assumption are rather limited (Royama 1966; Nur 1984; Ardia 2007; Schwagmeyer & Mock 2008). Furthermore, it must be noted here that if it is important to control for feeding frequency while testing for egg-size effects, the same is true when the main variable of interest is feeding frequency. I am not aware of any study that has simultaneously tested for effects of both egg size and feeding frequency on offspring quality.

Both studies of prenatal and post-natal parental effects are usually limited to short-term effects (i.e. when chicks are still in the nest). This is unfortunate since young often have the largest mortality shortly after fledging (Naef-Daenzer, Widmer & Nuber 2001). Consequently, survival of fledglings is usually the principal source of variation in female reproductive success (Clutton-Brock 1988; see also Mattsson & Cooper 2007). Not taking into account the stage when the opportunity for selection is at its highest might lead to biased conclusions (see also Russell *et al.* 2007a). I know of only two studies that have tested for the effects of feeding frequency on probability of offspring recruitment into the breeding population in birds (MacColl & Hatchwell 2003; Schwagmeyer & Mock 2008) and no such study among those that experimentally tested for egg-size effects on offspring quality.

To advance our understanding of how egg size affects offspring phenotype, I performed a cross-fostering experiment. I exchanged whole clutches between pairs of collared flycatchers' nests and followed offspring until recruitment in the following years. I also captured parents and included their morphological traits in statistical models to control for possible direct genetic effects while estimating egg-size effects on offspring morphology. Furthermore, to control for the potentially confounding influence of post-natal care, I scored parental feeding frequencies and modelled their effects on offspring quality. In addition, I compared effects obtained in the cross-fostering experiment with those obtained from a set of unmanipulated nests.

Materials and methods

GENERAL METHODS

I conducted the study in the Velký Kosíř area (49°32′N, 17°04′E, 300–400 m a.s.l.), Czech Republic during 2000–08. In the study area, there were approximately 300 nest boxes in spruce [*Picea abies* (Linnaeus) Karsten, 2000–2005] and oak [*Quercus petraea* (Mattuschka) Liebl, 2002–2008] forest. I conducted the experiments with the collared flycatcher (*Ficedula albicollis* Temminck), a small migratory passerine that easily adopts nest boxes for breeding. Females usually lay one egg per day and solely incubate clutches of three to eight eggs. Both collared flycatcher parents feed nestlings with invertebrate food for about 15 days until fledging.

During the time of laying, I monitored the nests daily and using a water-proof pen, marked newly-laid eggs with respect to laying sequence. I measured all eggs with a digital calliper to the nearest 0·01 mm. I computed egg size using Hoyt's (1979) formula: egg size = $0.51 \times breadth^2 \times length$. I visited most nests daily around the presumed time of hatching. I weighed nestlings to the nearest 0·25 g using a Pesola spring balance at age 6 and 13 days. At age 13 days, I also measured their maximal wing length to the nearest 1 mm using a ruler and tarsus to the nearest 0·01 mm using a digital calliper. To obtain an estimate of T-cell mediated immune response, at age 12 days, I injected the right wing web of each chick with 0·1 mg phytohaemagglutinin in 20 μL of physiological saline solution. Before injection, I took two measurements of wing web thickness (to the nearest 0·01 mm) with a thickness gauge (Mitutoyo Quick-Mini, Kawasaki, Japan) that was adjusted to push with a constant pressure of one Newton. I re-measured wing web thickness 24 h (± 2 h) after injection. I calculated cell-mediated immune response of each chick as the difference in average thickness of wing web after injection minus the average thickness of wing web before injection. Using traps, I captured parents while feeding nestlings and measured them as per nestling measurements, except they were not injected with PHA. All young were ringed at age 6–13 days with aluminium rings. Since recruitment rate was high (see Appendix), I was able to recapture many young as breeding adults in subsequent years.

At age 6 and 12 days, I monitored parental feeding frequencies using video-cameras that were set a few metres apart from the nest box. I monitored provisioning behaviour for 70–185 min (mean ± SD: 87 ± 23 and 88 ± 22 min at age 6 and 12 days respectively). I disregarded the first 10 min of recording to minimize the effect of disturbance on estimates of parental feeding frequencies. The collared flycatchers were minimally influenced by the video-cameras and resumed their feeding behaviour just minutes after their installation.

CROSS-FOSTERING

I exchanged whole clutches between pairs of nests. I performed this experiment during 2005–07. In 2005, I exchanged each egg on the day it was laid for purposes of another study (Krist & Grim 2007). So during this year, nest pairs were matched by laying date but may have differed with respect to clutch size which was unknown at the commencement of exchanges. Consequently, clutch size differed by three in one pair, by two in four pairs, by one in 13 pairs and was the same in 14 pairs of nests. Clutch size manipulation (cross-fostered clutch minus original clutch) was added as a covariate in all statistical models (see below). During 2006–07, I exchanged complete clutches 1–4 days after clutch completion between nest pairs matched by the clutch size but

with date of laying of the last egg differing by up to two days. As a result, in the complete sample (2005–07), nest pairs differed in the date of the last egg by four days in one case, by three in two cases, by two in six cases, by one in 26 cases and did not differ in 50 cases. The difference between laying dates in cross-fostered and original nests was added as covariate in all analyses (see below). To account for the manipulation of clutches and laying date potentially biasing results, I repeated the analyses with nests that were not manipulated in these respects (*n* = 87 nests). Results from these models were both qualitatively and quantitatively similar to those from models fitted on the whole sample and therefore only the latter are presented.

I performed exchanges within 22 min (mean ± SD = 9·32 ± 4·87) during 2006–07. Hatchability was the same as in unmanipulated clutches (Appendix) which indicates that the handling of incubated eggs did not negatively affect embryo survival.

STATISTICAL ANALYSES

All analyses are based on clutch/brood means. As the trade-off between number and size of offspring is solved at the level of individual females, clutch/brood means are appropriate levels for statistical analyses. In agreement with this view, most previous cross-fostering studies were analysed in this way (e.g. Magrath 1992; Smith *et al*. 1995; Russell *et al*. 2007b). Moreover, I would be unable to perform analyses based at the level of individual offspring, since due to high hatching synchrony these cannot be attributed to eggs.

Predictors

To test for effects of egg-size and feeding frequency on offspring performance, I fitted several models which had similar sets of predictors. (1) the size of cross-fostered egg (i.e. eggs from which nestlings actually hatched), (2) the size of original eggs (i.e. eggs that were originally laid on territory. I included size of original eggs in the models as they might be correlated with some aspect of parental or territory quality. In many previous cross-fostering studies, this factor had an influence on offspring performance (reviewed in Silva *et al*. 2007). (3) Feeding frequency per hour at age 6 days, (4) Feeding frequency per hour at age 12 days. Feeding frequencies were adjusted for a number of external factors and for brood size (Table 1) so that they represented a per-chick feeding frequency obtained under more standardized conditions. I included both feeding frequencies among predictors as they were only very weakly correlated (Table 2). (5) Number of care-givers. Most nests were attended by both parents but approximately one-ninth of them were attended by a single parent (Appendix), which was usually the female (only one brood was attended by the male only). I included this variable among predictors since I expected that a single parent may also decrease the quality of delivered diet and not just the feeding frequency. (6) Midvalue of genetic parents. I included this variable into the model to control for direct genetic effects (see Krist & Remeš 2004). When the response variable was offspring wing or tarsus length, the midvalue of genetic parents was just the mean of maternal and paternal wing and tarsus length respectively. In the models where the response variable was offspring weight (at age 6 and 13 days and when breeding), parental weight was adjusted to the time of their capture; so residuals from the regression of parental weight on offspring age at the time of capture were used [weight of females = $14.48 - 0.125 \times$ offspring age (in days), $F = 53.61$, $P < 0.001$, $R^2 = 0.171$, $n = 262$; weight of males = $13.38 - 0.033 \times$ offspring age (in days),

Table 1. Dependence of feeding frequencies on brood size and some external factors. Feeding frequencies used in all other statistical tests are residuals from these models

	Feeding frequency at 6 days			Feeding frequency at 12 days		
	DDF	*F*	*P* (R^2)	DDF	*F*	*P* (R^2)
Full model			(21·7)			(54·8)
Brood size	+139	35·90	<0·001	+133	105·49	<0·001
Time of day	−139	3·32	0·071	−133	11·25	0·001
Ambient temperature	−139	1·62	0·206	−133	5·46	0·021
Cloudiness	−139	0·36	0·548	+133	0·60	0·438

Cloudiness was estimated on following scale: 0 = cloudless, 1 = a few clouds, 2 = many clouds, 3 = overcast. The statistical direction of the effect of each individual predictor is indicated by a sign.

$F = 2·54$, $P = 0·112$, $R^2 = 0·011$, $n = 232$]. (7) Year entered the models as a categorical factor. (8) Laying date was fitted as a Julian date. (9) Clutch size manipulation (cross-fostered clutch – original clutch). (10) Laying date manipulation (laying date of the last egg in a cross-fostered clutch – laying date of the last egg in an original clutch).

Predictors 1–6 were the main factors of interest and as such were always retained in the final models. Predictors 7 and 8 were covariates that are usually very important for offspring performance. Grafen and Hails (2002, p. 218) recommend to retain such factors in final model regardless of their actual significance. I followed this guideline. Predictors 9 and 10 were included in the initial models only to ensure that my manipulation of clutch size and laying date did not bias the results. As they were otherwise of little interest in the present study, they were retained in the final models only if they were statistically significant. Since feeding frequencies cannot have a causal effect on hatchability, they were not included among predictors of the respective model. This was also the case for feeding frequency at age 12 days and nestling weight at age 6 days. Furthermore, I did not include parental wing length among the predictors of fledgling wing length since the latter is not fully grown at the time of fledging.

Note that inclusion of feeding frequencies together with size of original eggs among model predictors means that the latter variable functions here only as a surrogate of parental/territory quality other than that which is manifested in feeding frequencies. Similarly, the causal effect of the feeding frequencies on offspring performance could be underestimated if feeding frequencies were strongly correlated with the size of the original eggs. However, these correlations were weak (Table 2). To be sure that inclusion of the size of the original eggs among predictors did not bias estimates of feeding frequencies, I re-fitted all models without the former variable. Estimates of effect of feeding frequencies from these models were very similar to those from the models including size of original eggs, so only the latter are presented.

To test how biased the estimates of egg-size effects would be if feeding frequencies and genetic effects were not controlled for, I refitted all models without the midvalues of parental traits and feeding frequencies. Results of these models are only shown as a part of Fig. 1a.

Responses

I fitted models testing for the effects on offspring survival as a logistic regression (event/trial syntax) in Proc Genmod (SAS Institute 2003). (i) Hatchability: number of hatched eggs/clutch size.

(ii) Fledging success: number of fledged young/number of hatched young. (iii) Recruitment rate: number of recruited/number of fledged. Nests with predation events and with total failures were excluded from these survival analyses except for recruitment rate that was also based on nests with no recruits. I acknowledge that a small egg size might cause total failure at this stage (fledging–recruitment), in contrast to failure at the nestling stage which is usually caused by parental abandonment of the nest. These models had a binomial error structure, logit link function and statistics corrected for overdispersion by Deviance/df.

I fitted models testing for effects on offspring morphological/life-history traits as general linear models in JMP (SAS Institute 1995). Young that did not fledge were excluded from computation of brood means. For offspring traits measured up to fledging, each survived nestling contributed one value for the computation of brood means. This was not the case for traits measured on recruits. Some offspring bred in a greater number of years and consequently offered more than one value for computation of brood means. I used all measurements to utilize all available information. Consequently, brood means of traits measured on recruits are based on 1–8 recruit-measurements. Before averaging, I removed the effect of some confounding variables from the recruits' values. I removed year effect from all measurements, sex effect from the recruits' wing length and weight, as collared flycatchers are slightly dimorphic in these traits (Przbylo *et al.* 2000; M. Krist unpublished data), and the phase of breeding (i.e. offspring age when the recruit was captured) from weight of recruits. I used residuals from the (multiple) regression of the variable of interest on confounding factors listed above instead of raw measured values for computation of brood means. In the case of laying date, within-year median, not mean laying date was subtracted from actual laying date. I used only female recruits when testing egg-size effect on clutch size, as there is no evidence of a male effect on clutch size in this species (Sheldon, Kruuk & Merilä 2003).

Unmanipulated nests

I analysed data obtained on unmanipulated nests by basically the same methods as described above for the cross-fostering experiment except that some variables were not measured: feeding frequencies, nestling weight at 6, and fledging immunity.

Presentation

I present the main results in the form of effect sizes with 95% confidence intervals (Figs 1, 3). Effect size with its confidence interval is

Table 2. Correlation matrix of *predictors* of various models in the cross-fostering experiment. Pearson correlation coefficient is given. Sample size varies between 112–170. Correlation coefficients for the two categorical variables (year and number of care-givers) were computed from an *F*-test (Rosenthal 1994). A positive correlation with number of care-givers means that the value of the other character was greater if the nest was attended by both parents. In the case of year, the correlation just indicates the strength of the relationship and is presented without a sign. Variables given in italics were measured at recipient nests. Variables given in bold were measured on paired donor nests. Variables given in normal font were the same (year) or were very similar (clutch size, laying date; see Methods) for nests of the same pair

	Cross-fostered egg size	Original egg size	Feeding frequency at 6 days	Feeding frequency at 12 days	Number of care-givers	Year	Lay date	Clutch size	Mid-parent tarsus	Mid-parent weight	Mid-parent wing
Cross-fostered egg size		0·074	0·205	−0·034	−0·026	0·115	0·327	−0·081	0·167	0·292	−0·006
Original egg size			0·152	−0·087	−0·064	0·115	0·330	−0·168	0·163	0·143	0·024
Feeding frequency at 6 days				0·003	0·176	0·038	0·164	−0·105	0·061	0·324	−0·003
Feeding frequency at 12 days					−0·019	0·137	−0·028	−0·003	0·044	−0·140	−0·130
Number of care-givers						0·012	−0·139	0·059	−0·127	0·007	0·034
Year							0·279	0·099	0·136	0·249	0·129
Lay date								−0·257	0·137	0·206	−0·158
Clutch size									0·012	−0·123	0·029
Mid-parent tarsus										0·257	0·098
Mid-parent weight											0·317
Mid-parent wing											

more informative and therefore preferable method of summarizing results than conventionally used *P* values (Nakagawa & Cuthill 2007). Effect sizes presented here are correlation coefficients that were computed from the *F*-values of the respective statistical tests (see Table S1) according to the formula: $r = \sqrt{[F/(F + DDF)]}$, where DDF are error (denominator) degrees of freedom (Rosenthal 1994, p. 237). I set confidence limits for the correlation coefficient using z-transformation (Sokal & Rohlf 1995, p. 577).

Results

EGG-SIZE EFFECTS

Cross-fostering design

The size of cross-fostered eggs (i.e. eggs from which young hatched) was only weakly correlated with that of the original eggs (Table 2) which suggests that cross-fostering was efficient in decoupling potential correlations between egg size and parental/territory quality. This low correlation also means that it is reasonable to include size of both original and cross-fostered eggs as predictors in the same models. Similarly, correlations among other variables used as predictors were not very strong (the highest one being $r = 0·330$; Table 2). Consequently, fitted models are unlikely to be affected by multicollinearity among predictors.

The size of cross-fostered eggs had generally positive effects on offspring performance both in the short and long term (Fig. 1a; Table 3). These effects were weak (i.e. correlation coefficient up to 0·1; Cohen 1988) to medium ($r = 0·3$) in magnitude. They were generally higher in reduced models (i.e. models without parental midvalues and feeding frequencies), but this upward bias was negligible in magnitude (Fig. 1a). The size of the original eggs (i.e. surrogate of parental/territory quality) had a generally smaller (and statistically unrecognizable) effect on offspring performance than the size of the cross-fostered eggs (Fig. 1a; Table 3).

The size of cross-fostered eggs had the greatest effect on offspring morphology. This holds from the age of 6 days (nestling mass) till adulthood (weight and wing length of recruits, Fig. 1a). The only measured morphological variable not affected by cross-fostered egg size was offspring mass at fledging. In contrast to morphology, the size of cross-fostered eggs had virtually no effect on offspring survival at any life

Table 3. Weighed means of effect size for main predictors. Means were weighed by the sample size for each individual effect size. Number of estimates of effect size available for computation of weighed mean (i.e. number of available predictor–response combinations) is indicated in parentheses

	Cross-fostering		Unmanipulated	
Cross-fostered egg size	0·119	(12)		
Original egg size	0·043	(12)	0·148	(10)
Feeding frequency at 6 days	−0·008	(11)		
Feeding frequency at 12 days	−0·038	(10)		
Number of care-givers	0·192	(11)		
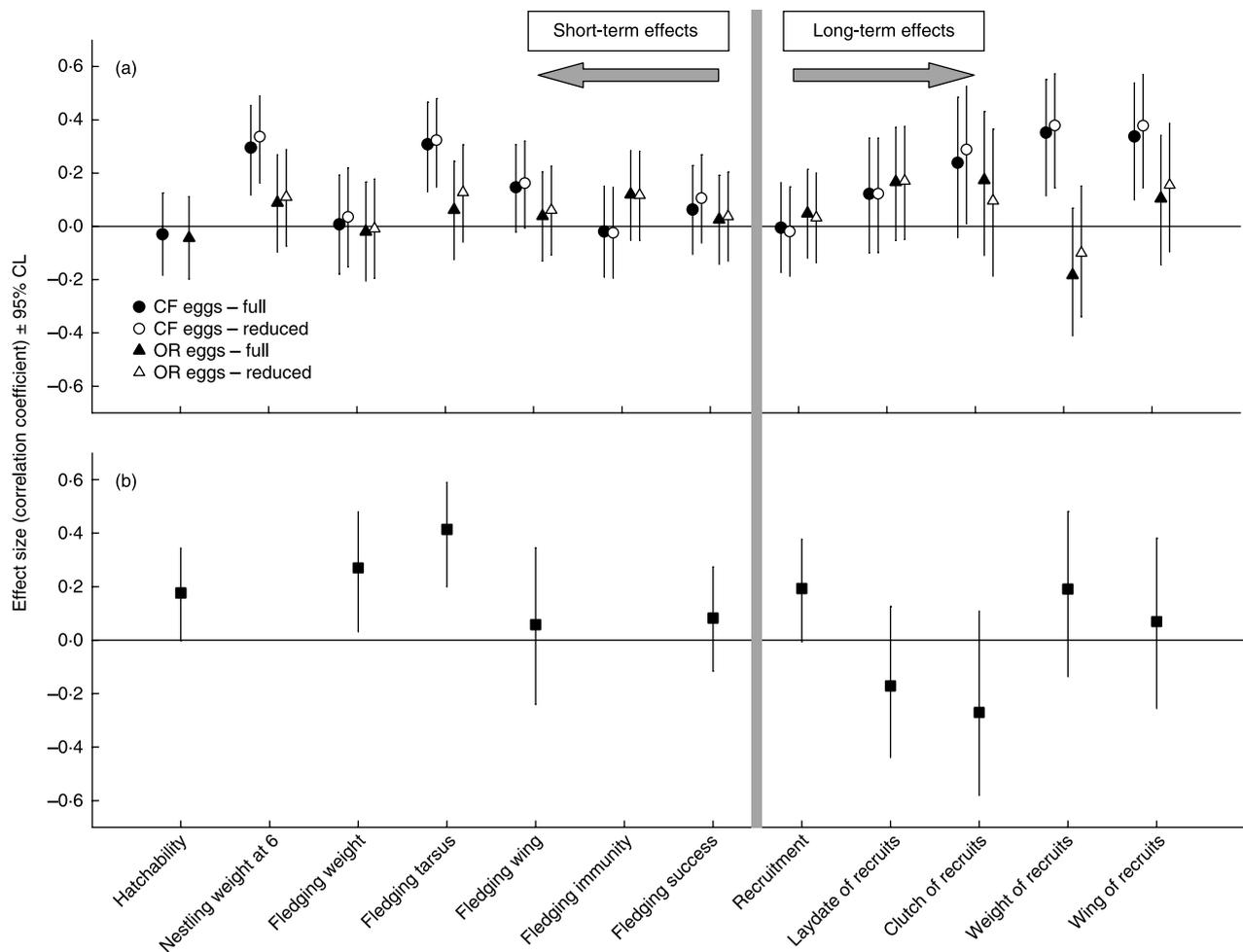
Fig. 1. Effect sizes corresponding to relationships between egg size and offspring traits. The sign of the presented effects corresponds to the sign of statistical covariation between egg size and dependent variables. The exception is the egg-size effect on laying date, where a statistically negative effect means a biologically positive effect. The sample size used for computation of effect size can be inferred from degrees of freedom of the particular test (Table S1). (a) Effects obtained in the cross-fostering experiment. The effect sizes from full and reduced models are compared. The latter models do not include the parental mid-value and feeding frequencies among predictors. Labels 'CF eggs' and 'OR eggs' denote size of the cross-fostered egg (i.e. an egg from which chicks actually hatch) and original egg (i.e. an egg originally laid on the territory) respectively. (b) Effect sizes obtained from the sets of unmanipulated nests.

stage (hatchability, fledging success and recruitment rate, Fig. 1a). Similarly, nestling PHA immunity was unaffected by the size of the eggs from which they had hatched (Fig. 1a). Life-history traits (laying date and clutch size) were positively but weakly affected by the size of cross-fostered eggs (Fig. 1a).

Unmanipulated design

The egg-size effect obtained on unmanipulated clutches should in theory be equal to the sum of effects of cross-fostered and original eggs in cross-fostering designs. This is roughly true if we compare the overall egg-size effects (cross-fostering: 0·162, unmanipulated: 0·148, Table 3), but looking at individual traits, however, would not lead to this conclusion (Fig. 1a,b). In contrast to the cross-fostering design, egg

size in unmanipulated nests had a positive effect on offspring survival (hatchability and recruitment rate) and fledging weight, but negative effect on life-history traits of recruits (Fig. 1b).

EFFECTS OF PARENTAL PROVISIONING

Approximately one-ninth of nests were attended by only one parent (usually a female) at offspring age 12 days (Appendix). Broods attended by single parents were provisioned with a lower frequency. However, the difference was relatively small which indicates that parents working alone increased their level of parental effort (Fig. 2). At age 12 days, feeding frequency per offspring was the same in the two categories of nests (Fig. 2). This was attributable to a brood-size reduction that occurred more often in nests attended by single parents (see Fig. 3b).
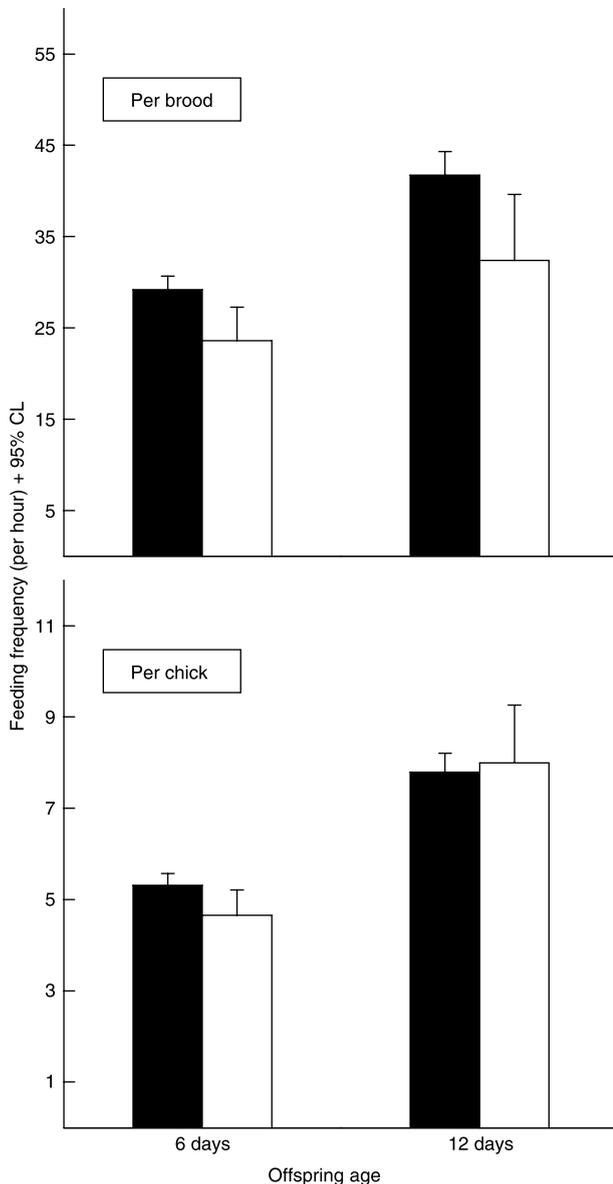
Fig. 2. Total feeding frequencies (+95% CL) to nests attended by both parents (black bars) and a single parent (white bars). Number of nests attended by a single parent was 10 and 16 at age 6 and 12 days respectively. The respective figure for nests attended by both parents was 135 and 122.

The number of care-givers had a consistent positive effect on nestling performance (Fig. 3b). However, in contrast to these clear short-term effects, the number of care-givers had no effect in the long term (Fig. 3b).

Parents fed offspring with greater frequency at age 12 days than at age 6 days (Fig. 2; Appendix). Feeding frequencies were uncorrelated between the two ages (Table 2) and had no clear directional relationship with offspring performance (Fig. 3a; Table 3). In line with expectations, feeding frequency at age 6 days had a positive effect on nestling survival (Fig. 3a). However, in contrast to my a priori expectations, morphology of fledglings was negatively correlated with the brood feeding frequency (Fig. 3a).

Discussion

EGG-SIZE EFFECTS

The size of cross-fostered eggs (i.e. eggs from which chicks had hatched) positively affected offspring morphology but had no effect on offspring survival and immunity. The size of eggs originally laid in the territory (surrogate of parental/ territory quality) had no significant effect on any of the offspring performance traits. If effects of egg size per se and parental quality were additive, their sum should be equal to egg-size effect in unmanipulated nests. However, this was usually not the case. This finding suggests that egg-size effects change temporally or differ between habitat types, since the majority of data on unmanipulated nests were obtained during different years (see Appendix) and from different habitat types (spruce and oak forest) to the data on cross-fostered nests (oak forest only).

Ultimately, the relationship between egg size and offspring fitness is of interest. This relationship is difficult to study directly which is why I studied the effects on a suite of other offspring traits that might be related to their fitness. Clearly, this relationship will differ between the studied traits. Perhaps the most direct relationships to fitness are the three traits describing offspring survival at various life stages (hatchability, fledging success, and recruitment probability). Egg-size effects on these survival traits were absent in the cross-fostering design.

Although all three above-mentioned traits describe survival, fitness may be more sensitive to change in a stage with lower mean survival. Consider this example: the probability of successful hatching, fledging, and recruitment is 0·9, 0·8, and 0·2. The probability of an egg giving a recruit is the product of the three above probabilities: $0·9 \times 0·8 \times 0·2 = 0·144$. An increase in survival of 10% added to the three stages would lead to an overall survival probability of 0·160, 0·162, and 0·216 respectively. Although this example is artificial, its values are similar to real ones (Appendix). The above conclusion is also supported by across-nest correlations between the probability of an egg giving a recruit and the survival probabilities of the three stages ($r_s = 0·057$, 0·190, and 0·849 respectively; $n = 141$, only nests with at least one recruit included, data from the cross-fostered and unmanipulated designs pooled). Thus, opportunity for selection is highest in the stage after fledging (also see Clutton-Brock 1988) which stresses the need to consider this stage when thinking about the fitness benefits of parental care.

Additionally, morphological and immunological traits of fledglings are well known to affect post-fledging survival both in *Ficedula* flycatchers (Kruuk, Merilä & Sheldon 2001; Merilä, Kruuk & Sheldon 2001; Moreno *et al.* 2005) and in birds in general (Gebhardt-Henrich & Richner 1998). On the contrary, directional selection on traits expressed later in life may be relatively strong as in the case of laying date (Sheldon *et al.* 2003; Garant *et al.* 2007), while weak at best on adult morphology (Przybylo, Sheldon & Merilä 2000) and clutch size (Sheldon *et al.* 2003; but see Garant *et al.* 2007 for a more
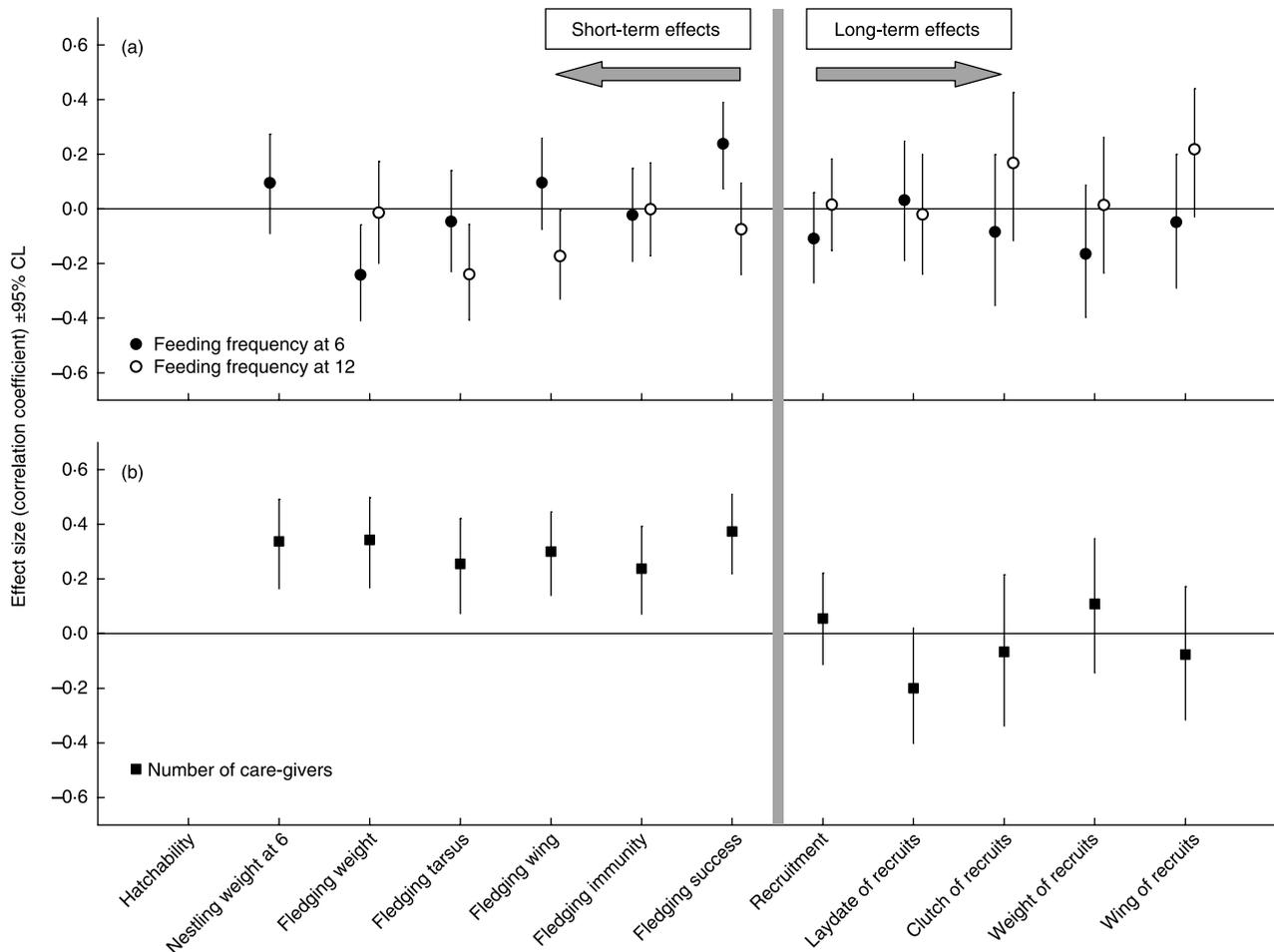
Fig. 3. Effect sizes corresponding to relationships between feeding frequencies or number of care-givers and offspring traits. The sign of the presented effects corresponds to the sign of statistical covariation between feeding frequencies and dependent variables. The exception is the effect on laying date, where a statistically negative effect means a biologically positive effect. The sample size used for computation of effect sizes can be inferred from degrees of freedom of the particular test (Table S1). (a) Effects of feeding frequencies at age 6 and 12 days on offspring performance. (b) Effects of number of care-givers on offspring performance. A positive effect means better performance of offspring from nests attended by both parents.

complicated picture). From these reasons, it may be problematic to average egg-size effects through the suite of offspring traits as I have done in the Table 3. Despite this limitation, I believe that similar illustrative information may be useful if one interprets it with caution.

The size of the eggs a female lays is usually regarded as an indicator of female or territory quality (Hörak *et al.* 1997, Silva *et al.* 2007). In most previous cross-fostering studies, the size of the original eggs had no effect on offspring traits immediately after hatching but this effect appeared later in the nestling period (reviewed in Silva *et al.* 2007). I did not find an effect of original eggs on offspring performance. It is usually assumed that the effect of the original eggs is mediated by increased post-natal care (e.g., parental feeding). If this was true in my study, the lack of the effect of original eggs would be expected since feeding rates were included in the same models as the original eggs. The latter variable then could become redundant due to the inclusion of a more direct

estimate of parental care into the statistical model. However, this was clearly not the case. The set of reduced models in which original eggs but not feeding frequencies were tested, gave very similar results (Fig. 1a).

The results of the full and reduced models were also very similar for the cross-fostered eggs. This means that although in theory, parental feeding and genetic effects may bias estimates of an egg-size effect in cross-fostering designs (Krist & Remeš 2004), this bias was not of practical importance in the present data set. To bias egg-size effect estimates substantially, parental feeding frequencies need to be both (i) strongly correlated with egg size, and (ii) highly influential for offspring performance. Neither of the two conditions arose in this study. These two conditions also did not arise simultaneously in the case of direct action of genes. Tarsus length was highly heritable (Table S1) but only mildly correlated with egg size (Table 2). Adult body mass was more strongly correlated with egg size (Table 2) but less heritable (Table S1). Similar

correlations between adult mass, tarsus length, and egg size were found in previous studies on collared (Hargitai *et al.* 2005) and pied flycatchers (Potti 1993). This suggests that in *Ficedula* flycatchers, egg-size effects may be properly estimated without the need for control of genetic effects.

However, this may not be the general rule. In mammals, for example, direct genetic and maternal effect on offspring performance were strongly positively correlated in both laboratory (Riska, Rutledge & Atchley 1985) and field settings (McAdam *et al.* 2002). In arthropods, there was usually found positive correlation between egg size and offspring performance (Fox and Czesak 2000). However, detailed studies on seed beetles revealed that offspring body size is from the most part genetically controlled with little room for maternal effects (Fox 1994). In line with this, positive correlation between egg size and offspring body size which is observable in non-experimental settings, disappear if egg mass is experimentally manipulated (Fox 1997). These results suggest that correlation between heritable body size and egg size confounds relationship between egg size and offspring body size (Fox 1997). To conclude, both correlation between parental characters (egg size, feeding frequencies, body size) and their effects on offspring quality may be species- and context specific. Consequently, more studies unravelling causal relationship between egg size and offspring quality are needed for better understanding of the trade-off between quality and number of offspring.

FEEDING FREQUENCIES

The number of care-givers was an influential predictor of nestling performance but had no effect in the long term. It is not very surprising to find that two parents provide better care than a single parent, although the opposite is also theoretically possible if parents heavily negotiate over care (McNamara *et al.* 2003). Although some support has been found for the latter scenario (Griggio & Pilastro 2007), the majority of previous studies provided results similar to our own (e.g. Markman, Yom-Tov & Wright 1996; Russell *et al.* 2007a,b), including other studies on flycatchers (Sheldon 2002; Huk & Winkel 2006). Females working alone increased the frequency of feeding visits [Fig. 2; if present, the male provided about half of all feedings (M. Krist, unpublished data)] but they did not fully compensate for the males' absence (Fig. 2). This would suggest that the poor performance of offspring in these nests is caused by an insufficient feeding frequency. However, this possibility could be excluded because feeding frequencies were controlled for in these tests. Notably, the same pattern was found in a previous study on this species (Sheldon 2002).

This, at first sight may seem somewhat paradoxical but the result can be explained in two ways. First, secondary females might decrease either the quality of delivered food or quantity per visit. Recently, it has been found that the size of delivered prey was more important for offspring fitness than just the frequency of delivery (Schwagmeyer & Mock 2008). Second, the effect of number of care-givers on offspring can in fact be driven by feeding frequencies but I might have been unable to control for this effect adequately. I measured feeding frequencies in a limited time-window while effect of number of attending parents was spread over the whole nestling period in the majority of cases. Feeding frequencies obtained in this limited time-window may only be a poor estimate of the feeding frequency over the whole nestling period. The lack of correlation between parental feeding frequencies at offspring age 6 and 12 days suggests that both of them cannot be good estimates of total feeding frequency.

On the other hand, both feeding frequencies were correlated in the expected direction with a number of other factors (Table 1, Fig. 2). Feeding frequency increased with number and age of young, which makes intuitive sense, and decreased during the day, a pattern that was also found in other previous studies (e.g. Moreno *et al.* 2004). These results suggest that feeding frequencies had some explanative power despite the limited time-window in which I measured them. It should also be mentioned here that most previous studies that dealt with feeding frequencies used a similar or a more limited time-window (but see e.g. Saino, Calza & Møller 1997; MacColl & Hatchwell 2003; Schwagmeyer & Mock 2008 for more detailed monitoring), and despite this, the majority of these studies reported a positive correlation between feeding frequency and some offspring trait.

Another explanation for the lack of clear effect of feeding frequency on offspring performance may be varying causality. Clear causality exists in the relationship between egg-size and offspring performance. Egg size can affect offspring performance while offspring performance cannot affect the size of the egg from which this offspring hatched. Unlike this case, feeding frequency can affect offspring performance but parents may also adjust feeding frequency to offspring state. Parents in nests with poorly developed young close to fledging might increase the feeding frequency to improve offspring condition at fledging, which is important for subsequent post-fledging survival (Kruuk *et al.* 2001; Merilä *et al.* 2001). Consequently, if causality goes from feeding frequency to offspring state, the sign of the correlation between the two variables would be positive, while both positive and negative correlations may arise if causality is reversed. These two effects might also cancel one another out resulting in no correlation between feeding frequency and offspring performance.

Conclusions

On the between-clutch/brood level, egg size per se (i.e. the size of cross-fostered eggs) affected offspring morphology in both the short and long term. As larger fledglings usually survive better than smaller ones (Merilä *et al.* 2001; Kruuk *et al.* 2001), this might in theory confer fitness benefits of larger eggs. However, in this study, recruitment was independent of egg size and egg size did not affect survival at earlier stages (hatchability and fledging success). Although egg size positively affected adult morphology, this trait can confer small benefits in terms of fitness (Przybylo *et al.* 2000). Con-

sequently, it is not clear whether a larger egg indeed confers *fitness* benefits for the offspring.

Egg size increased offspring hatchability and recruitment probability on unmanipulated nests. However, in this case it is not clear whether these effects were caused by egg size per se or by parental/territory quality. It is also possible that egg size has a positive effect on offspring only in some years and/or environments (see e.g. Smith & Bruun 1998). To conclude, although sample size was solid in this study, an even larger sample size and spread over more years and habitats may be needed to find conclusive evidence that, on the between-clutch level, larger eggs do indeed confer *fitness* benefits to offspring. It is also necessary to stress that egg size is only one aspect of egg quality. Egg composition, which was not target of this study, may have important consequences for offspring regardless of egg size (e.g. Blount, Houston & Møller 2000; Gil 2003; but see Remeš *et al.* 2007).

In contrast to a consistent positive effect of number of care-givers on nestling traits, there was no clear directional relationship between feeding frequency and the latter variables. This may be caused either by the limited time-window in which feeding frequencies were measured, lack of information on diet quality or parental adjustment of feeding frequencies to offspring state. Two principal improvements of studies dealing with feeding frequencies would be (i) to cover a greater portion of the nestling period by observation of parental feeding frequencies, and (ii) to measure also prey size and quality delivered to nestlings. At present, new technologies, such as automatic recorders based on PIT tag technology (e.g., Ottoson, Backman & Smith 1997) and small microcameras that are easy to position close to nestlings (Qvarnström *et al.* 2007) offer the possibility to solve these issues.

Acknowledgements

References

Alquati, A.B., Rubolini, D., Romano, M., Boncoraglio, G., Fasola, M. & Saino, N. (2007) Effects of egg albumen removal on yellow-legged gull chick phenotype. *Functional Ecology*, **21**, 310–316.

Amundsen, T. & Stokland, J.N. (1990) Egg-size and parental quality influence nestling growth in the Shag. *Auk*, **107**, 410–413.

Ardia, D.R. (2007) Site- and sex-level differences in adult feeding behaviour and its consequences to offspring quality in tree swallows (*Tachycineta bicolor*) following brood-size manipulation. *Canadian Journal of Zoology-Revue Canadienne de Zoologie*, **85**, 847–854.

Blount, J.D., Houston, D.C. & Møller, A.P. (2000) Why egg yolk is yellow. *Trends in Ecology & Evolution*, **15**, 47–49.

Bolton, M. (1991) Determinants of chick survival in the lesser black-backed gull: relative contributions of egg size and parental quality. *Journal of Animal Ecology*, **60**, 949–960.

Christians, J.K. (2002) Avian egg size: variation within species and inflexibility within individuals. *Biological Reviews*, **77**, 1–26.

Clutton-Brock, T.H. (1988) Reproductive success. *Reproductive Success: Studies of Individual Variation in Contrasting Breeding Systems* (ed. T.H. Clutton-Brock), pp. 472–485. The University of Chicago Press, Chicago, Illinois.

Cohen, J. (1988) *Statistical Power Analysis for Behavioral Sciences*. Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Cordero, P.J., Griffith, S.C., Aparicio, J.M. & Parkin, D.T. (2000) Sexual dimorphism in house sparrow eggs. *Behavioral Ecology and Sociobiology*, **48**, 353–357.

Ferrari, R.P., Martinelli, R. & Saino, N. (2006) Differential effects of egg albumen content on barn swallow nestlings in relation to hatch order. *Journal of Evolutionary Biology*, **19**, 981–993.

Finkler, M.S., Van Orman, J.B. & Sotherland, P.R. (1998) Experimental manipulation of egg quality in chickens: influence of albumen and yolk on the size and body composition of near-term embryos in a precocial bird. *Journal of Comparative Physiology B*, **168**, 17–24.

Fox, C.W. (1994) Maternal and genetic influences on egg size and larval prformance in a seed beatle (*Callosobruchus maculatus*): multigenerational transmission of a maternal effect? *Heredity*, **73**, 509–517.

Fox, C.W. (1997) Egg-size manipulations in the seed beetle *Stator limbatus*: consequences for progeny growth. *Canadian Journal of Zoology-Revue Canadienne de Zoologie*, **75**, 1465–1473.

Fox, C.W. & Czesak, M.E. (2000) Evolutionary ecology of progeny size in arthropods. *Annual Review of Entomology*, **45**, 341–369.

Garant, D., Kruuk, L.E.B., McCleery, R.H. & Sheldon, B.C. (2007) The effects of environmental heterogeneity on multivariate selection on reproductive traits in female great tits. *Evolution*, **61**, 1546–1559.

Gebhardt-Henrich, S.G. & Richner, H. (1998) Causes of growth variation and its consequences for fitness. *Avian Growth and Development: Evolution Within the Altricial-Precocial Spectrum* (eds J.M. Starck & R.E. Ricklefs), pp. 324–339. Oxford University Press, Oxford, UK.

Gil, D. (2003) Golden eggs: maternal manipulation of offspring phenotype by egg androgen in birds. *Ardeola*, **50**, 281–294.

Grafen, A. & Hails, R. (2002) *Modern Statistics for the Life Sciences*. Oxford University Press, Oxford.

Griggio, M. & Pilastro, A. (2007) Sexual conflict over parental care in a species with female and male brood desertion. *Animal Behaviour*, **74**, 779–785.

Gustafsson, L. & Sutherland, W.J. (1988) The costs of reproduction in the collared flycatcher *Ficedula albicollis*. *Nature*, **335**, 813–815.

Hargitai, R., Török, J., Tóth, L., Hegyi, G., Rosivall, B., Szigeti, B. & Szöllősi, E. (2005) Effects of environmental conditions and parental quality on inter- and intraclutch egg-size variation in the collared flycatcher (*Ficedula albicollis*). *Auk*, **122**, 509–522.

Hill, W.L. (1993) Importance of prenatal nutrition to the development of a precocial chick. *Developmental Psychobiology*, **26**, 237–249.

Hörak, P., Mänd, R. & Ots, I. (1997) Identifying targets of selection: a multivariate analysis of reproductive traits in the great tit. *Oikos*, **78**, 592–600.

Howe, H.F. (1976) Egg size, hatching asynchrony, sex and brood reduction in the common grackle. *Ecology*, **57**, 1195–1207.

Hoyt, D.F. (1979) Practical methods of estimating volume and fresh weight of bird eggs. *Auk*, **96**, 73–77.

Huk, T. & Winkel, W. (2006) Polygyny and its fitness consequences for primary and secondary female pied flycatchers. *Proceedings of the Royal Society B: Biological Sciences*, **273**, 1681–1688.

Krist, M. & Grim, T. (2007) Are blue eggs a sexually selected signal of female collared flycatchers? A cross-fostering experiment. *Behavioral Ecology and Sociobiology*, **61**, 863–876.

Krist, M. & Remeš, V. (2004) Maternal effects and offspring performance: in search of the best method. *Oikos*, **106**, 422–426.

Krist, M., Remeš, V., Uvírová, L., Nádvorník, P. & Bureš, S. (2004) Egg size and offspring performance in the collared flycatcher (*Ficedula albicollis*): a within-clutch approach. *Oecologia*, **140**, 52–60.

Krist, M., Nádvorník, P., Uvírová, L. & Bureš, S. (2005) Paternity covaries with laying and hatching order in the collared flycatcher *Ficedula albicollis*. *Behavioral Ecology and Sociobiology*, **59**, 6–11.

Kruuk, L.E.B., Merilä, J. & Sheldon, B.C. (2001) Phenotypic selection on a heritable size trait revisited. *American Naturalist*, **158**, 557–571.

Lock, J.E., Smiseth, P.T., Moore, P.J. & Moore, A.J. (2007) Coadaptation of prenatal and postnatal maternal effects. *American Naturalist*, **170**, 709–718.

MacColl, A.D.C. & Hatchwell, B.J. (2003) Heritability of parental effort in a passerine bird. *Evolution*, **57**, 2191–2195.

Magrath, R.D. (1992) The effect of egg mass on the growth and survival of blackbirds: a field experiment. *Journal of Zoology*, **227**, 639–653.

Markman, S., Yom-Tov, Y. & Wright, J. (1996) The effect of male removal on female parental care in the orange-tufted sunbird. *Animal Behaviour*, **52**, 437–444.

Mattsson, B.J. & Cooper, R.J. (2007) Which life-history components determine breeding productivity for individual songbirds? A case study of the louisiana waterthrush (*Seiurus motacilla*). *Auk*, **124**, 1186–1200.

McAdam, A.G., Boutin, S., Réale, D. & Berteaux, D. (2002) Maternal effects and the potential for evolution in a natural population of animals. *Evolution*, **56**, 846–851.

McNamara, J.M., Houston, A.I., Barta, Z. & Osorno, J.L. (2003) Should young ever be better off with one parent than with two? *Behavioral Ecology*, **14**, 301–310.

Merilä, J. & Sheldon, B.C. (2001) Avian quantitative genetics. *Current Ornithology* (ed. V. Nolan), pp. 179–255. Kluwer Academic/Plenum Publishers, New York.

Merilä, J., Kruuk, L.B. & Sheldon, B.C. (2001) Cryptic evolution in a wild bird population. *Nature*, **412**, 76–79.

Moreno, J., Osorno, J.L., Morales, J., Merino, S. & Tomas, G. (2004) Egg colouration and male parental effort in the pied flycatcher *Ficedula hypoleuca*. *Journal of Avian Biology*, **35**, 300–304.

Moreno, J., Merino, S., Sanz, J.J., Arriero, E., Morales, J. & Tomas, G. (2005) Nestling cell-mediated immune response, body mass and hatching date as predictors of local recruitment in the pied flycatcher *Ficedula hypoleuca*. *Journal of Avian Biology*, **36**, 251–260.

Naef-Daenzer, B., Widmer, F. & Nuber, M. (2001) Differential post-fledging survival of great and coal tits in relation to their condition and fledging date. *Journal of Animal Ecology*, **70**, 730–738.

Nakagawa, S. & Cuthill, I.C. (2007) Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biological Reviews*, **82**, 591–605.

Nakagawa, S., Ockendon, N., Gillespie, D.O.S., Hatchwell, B.J. & Burke, T. (2007) Does the badge of status influence parental care and investment in house sparrows? An experimental test. *Oecologia*, **153**, 749–760.

Nur, N. (1984) Feeding frequencies of nestling blue tits (*Parus caeruleus*): costs, benefits and a model of optimal feeding frequency. *Oecologia*, **65**, 125–137.

Oksanen, T.A., Jonsson, P., Koskela, E. & Mappes, T. (2001) Optimal allocation of reproductive effort: manipulation of offspring number and size in the bank vole. *Proceedings of the Royal Society B: Biological Sciences*, **268**, 661–666.

Olsson, M., Wapstra, E. & Olofsson, C. (2002) Offspring size-number strategies: experimental manipulation of offspring size in a viviparous lizard (*Lacerta vivipara*). *Functional Ecology*, **16**, 135–140.

Ottoson, U., Backman, J. & Smith, H.G. (1997) Begging affects parental effort in the pied flycatcher, *Ficedula hypoleuca*. *Behavioral Ecology and Sociobiology*, **41**, 381–384.

Potti, J. (1993) Environmental, ontogenetic, and genetic variation in egg size of pied flycatchers. *Canadian Journal of Zoology-Revue Canadienne De Zoologie*, **71**, 1534–1542.

Przybylo, R., Sheldon, B.C. & Merilä, J. (2000) Patterns of natural selection on morphology of male and female collared flycatchers (*Ficedula albicollis*). *Biological Journal of the Linnean Society*, **69**, 213–232.

Qvarnström, A., Kehlenbeck, J.V., Wiley, C., Svedin, N. & Saether, S.A. (2007) Species divergence in offspring begging intensity: difference in need or manipulation of parents? *Proceedings of the Royal Society B: Biological Sciences*, **274**, 1003–1008.

Reid, W.J. & Boersma, P.D. (1990) Parental quality and selection on egg size in the magellanic penguin. *Evolution*, **44**, 1780–1786.

Remeš, V., Krist, M., Bertacche, V. & Stradi, R. (2007) Maternal carotenoid supplementation does not affect breeding performance in the Great Tit (*Parus major*). *Functional Ecology*, **21**, 776–783.

Riska, B., Rutledge, J.J. & Atchley, W.R. (1985) Covariance between direct and maternal genetic effects in mice, with a model of persistent environmental influences. *Genetical Research*, **45**, 287–297.

Roff, D.A. (1992) *The Evolution of Life Histories: Theory and Analysis*. Chapman & Hall, New York.

Rosenthal, R. (1994) Parametric Measures of Effects Size. *The Handbook of Research Synthesis* (eds H. Cooper & L.V. Hedges), pp. 231–244. Russell Sage Foundation, New York.

Royama, T. (1966) Factors governing feeding rate, food requirement and brood size of nestling great tit *Parus major*. *Ibis*, **108**, 313–347.

Russell, A.F., Young, A.J., Spong, G., Jordan, N.R. & Clutton-Brock, T.H. (2007a) Helpers increase the reproductive potential of offspring in cooperative meerkats. *Proceedings of the Royal Society B: Biological Sciences*, **274**, 513–520.

Russell, A.F., Langmore, N.E., Cockburn, A., Astheimer, L.B. & Kilner, R.M. (2007b) Reduced egg investment can conceal helper effects in cooperatively breeding birds. *Science*, **317**, 941–944.

Saino, N., Calza, S. & Møller, A.P. (1997) Immunocompetence of nestling barn swallows in relation to brood size and parental effort. *Journal of Animal Ecology*, **66**, 827–836.

SAS Institute Inc. (1995) *JMP Statistics and Graphics Guide*. SAS Institute Inc., Cary.

SAS Institute Inc. (2003) *SAS Online Doc, 9·1*. SAS Institute Inc., Cary.

Schwagmeyer, P.L. & Mock, D.W. (2008) Parental provisioning and offspring fitness: size matters. *Animal Behaviour*, **75**, 291–298.

Sheldon, B.C. (2002) Relating paternity to paternal care. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **357**, 341–350.

Sheldon, B.C., Kruuk, L.E.B. & Merilä, J. (2003) Natural selection and inheritance of breeding time and clutch size in the collared flycatcher. *Evolution*, **57**, 406–420.

Silva, M.C., Boersma, P.D., Mackay, S. & Strange, I. (2007) Egg size and parental quality in thin-billed prions, *Pachyptila belcheri*: effects on offspring fitness. *Animal Behaviour*, **74**, 1403–1412.

Sinervo, B., Doughty, P., Huey, R.B. & Zamudio, K. (1992) Allometric engineering: a causal analysis of natural selection on offspring size. *Science*, **258**, 1927–1930.

Slagsvold, T., Sandvik, J., Rofstad, G., Lorentsen, O. & Husby, M. (1984) On the adaptive value of intraclutch egg-size variation in birds. *Auk*, **101**, 685–697.

Smith, C.C. & Fretwell, S.D. (1974) The optimal balance between size and number of offspring. *American Naturalist*, **108**, 499–506.

Smith, H.G. & Bruun, M. (1998) The effect of egg size and habitat on starling nestling growth and survival. *Oecologia*, **115**, 59–63.

Smith, H.G., Ohlsson, T. & Wettermark, K.J. (1995) Adaptive significance of egg size in the european starling: experimental tests. *Ecology*, **76**, 1–7.

Sokal, R.R. & Rohlf, F.J. (1995) *Biometry: The Principles and Practise of Statistics in Biological Research*. W.H. Freeman and Company, New York.

VanderWerf, E. (1992) Lack's clutch size hypothesis: An examination of the evidence using meta-analysis. *Ecology*, **73**, 1699–1705.

Wagner, E.C. & Williams, T.D. (2007) Experimental (antiestrogen-mediated) reduction in egg size negatively affects offspring growth and survival. *Physiological and Biochemical Zoology*, **80**, 293–305.

Williams, T.D. (1994) Intraspecific variation in egg size and egg composition in birds: Effects on offspring fitness. *Biological Reviews*, **69**, 35–59.

Supporting information

Additional supporting information may be found in the online version of this article:

Table S1. Short- and long-term effects on offspring performance. Results of statistical models.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Appendix. Means ± SD, and sample sizes (number of nests) for variables used either as predictors or responses in statistical models in the two data sets. Means and SDs were computed from clutch/brood means. Feeding frequencies are given as a per-chick per-hour value at age 6 and 12 days. Values of feeding frequencies and mid-parent weight were adjusted for statistical tests (see Methods). Here unadjusted values are given. Sample sizes for individual years are indicated in the bottom rows of the table

	Cross-fostering		Unmanipulated	
Egg size (mm^3)	1668·3 ± 131·4	170	1659·3 ± 117·2	134
Clutch size	6·31 ± 0·75	169	6·08 ± 0·78	134
Feeding frequency at 6 days	5·26 ± 1·51	145		
Feeding frequency at 12 days	7·89 ± 2·32	138		
Number of care-givers*	16/122	138		
Mid-parent tarsus (mm)	20·00 ± 0·30	136	19·89 ± 0·41	90
Mid-parent weight (g)	13·30 ± 0·64	134	13·35 ± 0·68	88
Mid-parent wing (mm)	8·16 ± 0·11	133	8·17 ± 0·12	88
Hatchability (%)†	93·9 ± 10·3	160	94·0 ± 13·3	121
Nestling weight at 6 (g)‡	9·92 ± 1·17	142		
Fledging tarsus (mm)‡	19·84 ± 0·41	151	19·84 ± 0·45	90
Fledging weight (g)‡	14·74 ± 0·97	150	14·68 ± 0·99	88
Fledging wing (mm)‡	5·10 ± 0·27	150	5·15 ± 0·26	45
Fledging immunity (mm)‡	0·51 ± 0·14	136		
Fledging success (%)†	85·6 ± 19·4	151	85·5 ± 21·0	100
Recruitment (%)¶	20·3 ± 21·6	151	13·4 ± 16·8	98
2000–2004		0		110
2005		64		7
2006		66		3
2007		40		14

*Number of nests attended by single/both parents at offspring age 12 days; †nests with complete failure were excluded; ‡nests with predation events were excluded; values are based on fledged young only; ¶nests with predation events were excluded; based only on nests with at least one young fledged.

Článek 10: Krist, M. (2011) Egg size and offspring quality: a meta-analysis in birds. *Biological Reviews*, 86, 692-716.

Egg size and offspring quality: a meta-analysis in birds

Miloš Krist[1,2,*]

[1] *Museum of Natural History, nám. Republiky 5, 771 73 Olomouc, Czech Republic*
[2] *Department of Zoology and Laboratory of Ornithology, Faculty of Science, Palacky University, tř. Svobody 26, 771 46 Olomouc, Czech Republic*

ABSTRACT

Parents affect offspring fitness by propagule size and quality, selection of oviposition site, quality of incubation, feeding of dependent young, and their defence against predators and parasites. Despite many case studies on each of these topics, this knowledge has not been rigorously integrated into individual parental care traits for any taxon. Consequently, we lack a comprehensive, quantitative assessment of how parental care modifies offspring phenotypes. This meta-analysis of 283 studies with 1805 correlations between egg size and offspring quality in birds is intended to fill this gap. The large sample size enabled testing of how the magnitude of the relationship between egg size and offspring quality depends on a number of variables. Egg size was positively related to nearly all studied offspring traits across all stages of the offspring life cycle. Not surprisingly, the relationship was strongest at hatching but persisted until the post-fledging stage. Morphological traits were the most closely related to egg size but significant relationships were also found with hatching success, chick survival, and growth rate. Non-significant effect sizes were found for egg fertility, chick immunity, behaviour, and life-history or sexual traits. Effect size did not depend on whether chicks were raised by their natural parents or were cross-fostered to other territories. Effect size did not depend on species-specific traits such as developmental mode, clutch size, and relative size of the egg, but was larger if tested in captive compared to wild populations and between rather than within broods. In sum, published studies support the view that egg size affects juvenile survival. There are very few studies that tested the relationship between egg size and the fecundity component of offspring fitness, and no studies on offspring survival as adults or on global fitness. More data are also needed for the relationships between egg size and offspring behavioural and physiological traits. It remains to be established whether the relationship between egg size and offspring performance depends on the quality of the offspring environment. Positive effect sizes found in this study are likely to be driven by a causal effect of egg size on offspring quality. However, more studies that control for potential confounding effects of parental post-hatching care, genes, and egg composition are needed to establish firmly this causal link.

Key words: altricial, birth mass, environmental quality, egg size, feeding frequency, intraclutch, juvenile survival, maternal effect, offspring fitness, yolk steroids.

CONTENTS

  * Address for correspondence E-mail: milos.krist@upol.cz; milos.krist@volny.cz

I. INTRODUCTION

Parental effects are presently a focus of interest to ecologists and evolutionary biologists (Badyaev & Uller, 2009). Of particular interest is the question of how variation in the environment provided by the parents affects offspring phenotype (e.g. Groothuis *et al.*, 2005; Green, 2008; Marshall & Keough, 2008). This relationship may be studied using two analytical approaches (Lynch & Walsh, 1998; see also McGlothlin & Brodie, 2009). In the first, offspring phenotypic variation is decomposed to its causal components by employing breeding designs, pedigree analysis, and cross-fostering (Kruuk, 2004; Kruuk & Hadfield, 2007; Wilson *et al.*, 2010). As a result we know how large the variance component of an offspring trait is due to common environmental and/or parental effects. However, we do not know which parental trait caused this variation, which is often of great interest. The second approach is a regression analysis (Lande & Price, 1989). All parental traits that causally affect offspring traits of interest should be included as predictors of the multivariate regression to find their net effects (Lande & Price, 1989). This condition can be difficult, if not impossible, to achieve. If the results are interpreted with caution however, the regression approach is useful, especially in combination with some experimental settings (Krist & Remeš, 2004).

Most studies that used the decomposition of variance method detected a significant parental component in offspring traits such as morphology (Kruuk, Merilä & Sheldon, 2001; McAdam *et al.*, 2002; DiBattista *et al.*, 2009), immunity (Soler, Moreno & Potti, 2003; Kilpimaa *et al.*, 2005; Pitala *et al.*, 2007), rate of development (Fox, 1993; Rauter & Moore, 2002; Winn, 2004), life-history (Hunt & Simmons, 2002; Fox, Czesak & Wallin, 2004; Charmantier *et al.*, 2006),

and behaviour (Forstmeier, Coltman & Birkhead, 2004). Parental effects are often contingent on environment or the timing of measurements. They may significantly differ between populations (Ardia & Rice, 2006) and years (Gebhardt-Henrich & van Noordwijk, 1994), and they typically decrease as young grow older (Charmantier *et al.*, 2006; Lindholm, Hunt & Brooks, 2006; Wilson & Réale, 2006). They may be more pronounced in poor-quality environments (McAdam & Boutin, 2003; Charmantier *et al.*, 2004). This is well supported by observations that environmental variance, to which parental effects contribute, increases while heritability decreases in poor-quality environments (Merilä & Sheldon, 2001; Charmantier & Garant, 2005). In sum, parental effects are pervasive, although their magnitude differs among offspring traits, environments and life-history stages.

Given their widespread occurrence, a logical question arises: what particular qualities of parents mediate these effects? Regression analyses often reveal the effects of parental size (Gebhardt-Henrich & van Noordwijk, 1991; Schrader & Travis, 2009), condition (Schluter & Gustafsson, 1993), age (Fox, Bush & Wallin, 2003; Berkeley, Chapman & Sogard, 2004; Bowen, 2009), diet (Bonduriansky & Head, 2007), exposure to parasites (Gallizzi & Richner, 2008), social environment (Kerrigan, 1997; Mateo, 2009), and sexual ornamentation (Griffith, Owens & Burke, 1999) on various offspring characters (see Green, 2008, for an extensive review in fishes). Although these relationships are of interest, parental characters are correlates, rather than causes of variation in offspring performance traits. Parents causally affect offspring performance by parental care (Clutton-Brock, 1991), that may take the form of selection of safe (Weidinger, 2002; Remeš, 2005) or high nutritional quality

(Agosta, 2008) oviposition sites, investment in propagule size (reviews in Williams, 1994; Bernardo, 1996), propagule quality (review in Blount, Houston & Møller, 2000; Gil, 2003; Groothuis *et al.*, 2005), incubation behaviour (Kovařík, Pavel & Chutný, 2009; Matysioková & Remeš, 2010; review in Deeming, 2002), food provisioning (Schwagmeyer & Mock, 2008; Krist, 2009), and active defence of offspring (Krist, 2004; Grim, 2005; Divino & Tonn, 2008). In sum, parents may affect offspring quality by several pathways and parental abilities to invest in these pathways are affected by conditions that they experience.

One possibility for how to give offspring a good start in life is to allocate extra resources to the propagule. However, this action is likely to be costly for a parent. The trade-off between the number and size of offspring is one of the central tenets in life-history evolution (Stearns, 1992; Roff, 2002). The first optimality model of this trade-off was given by Smith & Fretwell (1974). Their model relies on two assumptions. First, the energy available for reproduction is limited to a finite amount at any given time. Second, offspring fitness increases with parental per offspring investment. Subsequent models explored optimal parental allocation under more complex conditions. They assumed a larger dependence of offspring fitness on parental investment in harsh, competitive environments which thus select for larger propagules (Brockelman, 1975; Parker & Begon, 1986; McGinley, Temme & Geber, 1987). In addition to this between-female variation, allocation of resources within clutches has received substantial attention. First, in highly variable environments, females adopt a bet-hedging strategy and divide resources unequally within a clutch (Koops, Hutchings & Adams, 2003; Crean & Marshall, 2009; see also Geritz, 1995). Second, individual eggs may have different reproductive value due to spatial position in a clutch (Kudo, 2001), laying sequence (Williams, Lank & Cooke, 1993*a*; Riehl, 2010) or paternity (Sheldon, 2000; Krist *et al.*, 2005). By differential resource allocation, females might avoid investment in eggs with poor survival prospects (Williams *et al.*, 1993*a*; Kudo, 2001; Riehl, 2010) and facilitate (Slagsvold *et al.*, 1984; Krist *et al.*, 2005; Magrath *et al.*, 2009; Reed, Clark & Vleck, 2009; Kozlowski & Ricklefs, 2010) or counteract (Howe, 1976; Rosivall, Szöllősi & Török, 2005; Ferrari, Martinelli & Saino, 2006) within-brood competitive asymmetries caused by hatching asynchrony or paternity. All the above models assume greater fitness of large eggs. This seems to be a reasonable assumption in terrestrial habitats; in aquatic environments egg size may have a negative impact on hatching success due to limited diffusion of oxygen to developing embryos combined with a positive impact on post-hatching survival (Hendry, Day & Cooper, 2001).

Although the assumption of increasing offspring fitness with egg size seems to be reasonable, is there empirical evidence for this relationship? The relationship between egg size and offspring performance has been studied in every oviparous vertebrate class as well as in plants and many invertebrate taxa. These case studies have been reviewed as a part of wider, narrative reviews of maternal effects in plants (Roach & Wulff, 1987; Donohue & Schmitt, 1998), marine invertebrates (Marshall & Keough, 2008), arthropods (Fox & Czesak, 2000), and fish (Heath & Blouw, 1998). These reviews found positive relationships between propagule size and offspring quality. However, the amount of available data was generally too small to allow strong conclusions. Moreover, these relationships were sometimes limited to harsh environments (Donohue & Schmitt, 1998; Fox & Czesak, 2000) or early stages in the offspring life cycle (Heath & Blouw, 1998).

In birds, the relationship between egg size and offspring performance was the target of a specialized review by Williams (1994). Based on 60 studies, he found that this relationship was more evident in precocial than altricial species and in early compared to late phases in the chick-rearing period. He concluded: "There is little unequivocal evidence to date in a support of a positive relationship between egg size and offspring fitness in birds." (p. 54). His review was a narrative one and the conclusions were largely based on a comparison of a number of studies that found or did not find statistically significant egg-size effects. However, statistical significance is a poor measure of effect size since it confuses effect size and sample size. Therefore narrative and vote-counting reviews based on statistical significance of effect sizes found in primary studies are prone to errors and often lead to erroneous conclusions (Borenstein *et al.*, 2009, pp. 251–255). What is needed is a formal meta-analysis that bases the conclusions on effect size while also taking into account sample size (Arnqvist & Wooster, 1995). Despite this limitation, William's (1994) review together with the volume by Mousseau & Fox (1998), and the introduction of yolk hormones as modifiers of chick growth and behaviour by Schwabl (1993, 1996) led to an increased interest in egg-size effects in birds and a boom of publications on this topic.

The aim of the present study was to perform a meta-analysis of studies testing for the correlation between egg size and offspring quality in birds and thus provide a comprehensive, quantitative estimate of the strength of the propagule size—offspring quality relationship. This meta-analysis is based on 283 studies and 1805 estimates of effect size. The large sample size enabled testing of how the effect size depends on a number of variables. The variables included environment (captivity *versus* wild), level of variance in egg size (between-clutch, intraclutch, mixed), type of study design (e.g. cross-fostering *versus* observational), stage in offspring life cycle (egg, hatchling, nestling, post-fledging), nestling age, offspring traits (e.g. survival, morphology, immunity, growth rate), and species attributes (relative egg size, clutch size, developmental mode). After reviewing the field, I identify gaps in our knowledge, suggest avenues of further research, and discuss methodological issues related to estimation of the egg-size effect.

II. METHODS

(1) Data search and inclusion criteria

Three electronic databases were searched for studies that described the relationship between egg size and offspring

traits: *Web of Science* (since 1945), *Zoological Record* (since 1978), and *Biological Abstracts* (since 1990) with the last access on 5th October 2009. The exact search term is given in Appendix S1. In addition, reference lists of those studies that contained relevant data as well as that of Williams (1994) were searched. A third source of data was studies that were found accidentally, e.g. while reading them for other purposes. The study was included in the meta-analysis if it passed through all of the following selection criteria:

(1) The study contained a direct test of the relationship between egg size and offspring traits. A direct test means that egg size was either an independent (for example when offspring mass is regressed on egg mass) or a dependent (for example when the size of hatched and unhatched eggs is compared using a *t*-test) variable in the statistical test. Tests were not included in which categorical variables, whose levels differed in mean egg size, were used as predictors of offspring traits. These variables were for example pair experience (Ollason & Dunett, 1986), or experimental treatments such as food supplementation (Bolton, Houston & Monaghan, 1992), tamoxifen injection (Williams, 2001; Wagner & Williams, 2007), or direct manipulation of egg size (e.g. Bonisoli-Alquati *et al.*, 2008). The manipulative studies are useful for our understanding of maternal effects but at present too few such studies exist for a separate analysis and they are too different to pool with the rest of the data (see Section IV/5). Also excluded were tests that used hatchling mass as a surrogate of egg mass (e.g. Davis, 1975; O'Connor, 1975).

(2) The egg size was measured at the level of an individual egg or a clutch. Tests that correlated mean egg size measured at a higher hierarchical level such as population (Kroll & Haufler, 2007) or year (Järvinen, 1994) were excluded. These correlations were likely to be confounded by factors varying among populations or years and thus probably do not reveal the causal effect of egg size on offspring traits.

(3) The study was carried out on non-domesticated species/populations that were not kept for commercial purposes such as meat and egg production.

(4) The study involved ecologically relevant offspring traits; i.e. traits with either a known or at least assumed relationship to fitness. Tests relating egg size to neonatal body composition (see e.g. Anderson & Alisauskas, 2002) were not included since it is not clear whether it is better to have more lipids or proteins in the body.

(5) The study contained enough information to enable computation of the exact effect size and study variance (sample size) or at least an estimation of these quantities as explained in Section II/3. If this information was not evident from the published version, the authors were contacted for these details. For example, most of the data contained in a detailed study by Schifferli (1973) could not be used since means and regression coefficients given in the paper were not accompanied by standard errors or deviations.

(6) The study did not have a problematic experimental design or data analyses. For example, studies were excluded that cross-fostered clutches with large eggs for those with small eggs as egg-size effects could cancel out with parental rearing abilities in this experimental setting (e.g. Mänd, 1985; Arnold, Hatch & Nisbet, 2006). Also excluded were studies that were likely to suffer from a large multicollinearity between predictors in a multiple regression such as if egg volume, egg length, and egg breadth were tested in the same model (e.g. Adamou *et al.*, 2009). Similarly, if the study tested egg-size effects in a model also containing the interaction of egg size with another variable, the data were only included if separate estimates for different levels of the interacting factors were given or if the authors provided test statistics for the model without this interaction. It would be erroneous to use a test of the main effect as a measure of effect size when the interaction effect is included in the same model (see Engqvist, 2005).

(2) Effect size computation

Pearson's correlation coefficient (*r*) was used as a measure of effect size. If a test statistic other than a correlation coefficient was published, I converted it to *r* according to the formulae given in Rosenthal (1994). It is important to realize the potential difference between the statistical and biological direction of an effect. In this meta-analysis the two are likely to be the same for most offspring traits such as for example offspring size and survival, as both of these traits are probably positively related to offspring fitness. Therefore biological direction was considered the same as the statistical one with the exception of offspring's laying date (Krist, 2009), since earlier laying usually confers fitness benefits (e.g. Sheldon, Kruuk & Merilä, 2003).

In most published studies egg size was measured on a continuous scale and all values of egg size were used for the statistical test. However, two other approaches were quite common. First, egg size was dichotomized into categories, for example large and small eggs. Second, only part of the available egg sizes were used. Typically the test was based on only large and small eggs while middle-sized eggs were excluded. The first type of data handling is called dichotomization of the continuous independent variable while the second is called range enhancement in the independent variable. The effect size obtained in the first case is underestimated while in the second it is overestimated compared to the whole population. Therefore the effect size for these two treatments was adjusted according to the formulae given in Hunter & Schmidt (2004, p. 36–37). For adjustment to range enhancement (or range restriction) it is necessary to know the ratio of standard deviations (S.D.s) of an enhanced/restricted study to an unenhanced/unrestricted study. This was estimated using a large (100 000) sample normal distribution with S.D. = 1.
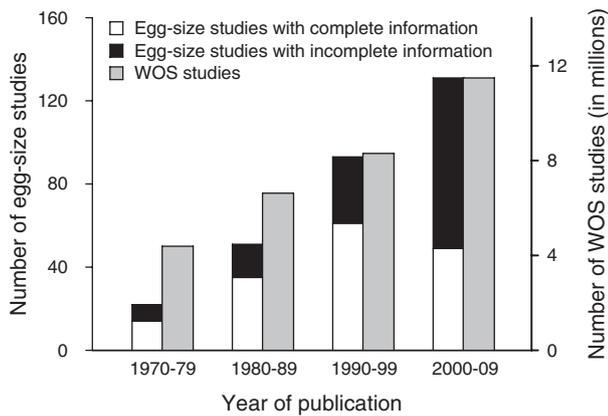
Fig. 1. Number of studies that tested the relationship between egg size and offspring quality in non-domesticated birds ($N = 297$) and the number of all studies included in the *Science Citation Index Expanded* database at *Web of Science* (WOS) in four decades. The egg-size studies with complete information published all the details needed for computation of effect size. This was not the case for studies with incomplete information.

These simulated data were restricted in the same way as they were restricted in the study in question and then the S.D. of this restricted dataset was examined. Formulae for both dichotomization and range restriction/enhancement worked well and the adjusted effect sizes were closer to real ones than were unadjusted ones. This was confirmed on a large simulated dataset (results not shown). For small samples, these adjustments also worked well on average but may have overestimated or underestimated the real effects in individual cases due to sampling variance. In four cases the adjusted r was larger than 1.00. r was set at 0.99 in these four cases. All these cases had small weight because (1) they were based on a small sample size and (2) the study variance was increased by dichotomization and range restriction. Moreover, all of them were merged with other estimates to give one estimate per study.

If the data were presented only in the form of graphs, these were scanned and the values read by one of two programs (scatterplots: *DigitizeIt*; bar plots: *Tpsdig*). If the information necessary for computation of effect size or study variance was missing, which was unfortunately quite common (Fig. 1), the authors were contacted for the missing details. Two types of information were necessary to compute effect size: (1) magnitude of the effect (this may be inferred, for example, from the F value if degrees of freedom are provided), (2) direction of effect (i.e. was the relationship between egg size and offspring trait positive or negative?). This latter information cannot be inferred from the F value, t value, chi-squared value, or P value standing alone.

(3) Estimation of effect size when published information was incomplete

The effect size was estimated if the missing details were not provided by the authors or if the authors were not located. Most often, the information was missing because the result

of the statistical test was stated as non-significant only. If the sample size was known, the upper boundary for the size of these non-significant effects could be computed. This upper boundary could be used as an estimate of effect size. More reasonable, however, was to use the value in the middle between this upper boundary and zero as an estimate of effect size. This was confirmed on a sample of 852 effect sizes from this meta-analysis which were non-significant but the magnitude and direction of the effect was known. The true mean correlation coefficient in this sample was 0.081. If the sign of negative correlations from this sample was changed to make all 852 estimates positive, the mean correlation would rise to 0.139. If we just know that these coefficients were non-significant and computed the upper boundary for them, this would equal 0.293. By this method they would be highly overestimated. The overestimation would not be so high if the latter rule is used $[(0.293 + 0)/2 = 0.146]$. Therefore this latter rule was used to estimate the magnitude of effect. The same rule was applied when the result was published as significant only—the estimated effect was in the middle between the lower boundary and 1.0.

All effects for which the direction of effect ($N = 162$ of 1805) was unknown were set as positive, resulting in an overestimation of the mean effect size. However, this overestimation is small because the magnitude of these effects is generally small. When these effects are set as positive, the mean weighted effect in the whole sample ($N = 1805$) is $r = 0.210$. If these 162 effects were set as negative, the mean effect size would only decrease to $r = 0.195$. Importantly, setting these effects as positive leads to lower overestimations of the mean effect size than if these effects with an unknown direction were excluded from the study (mean $r = 0.217$; $N = 1643$).

In sum, the magnitude of the effect or its direction was unavailable in 176 of 1805 cases (see Fig. 2). These estimated effects were included in the analyses to increase sample size and avoid selective exclusion of part of the data. However, all models were also refitted without these effects to check the sensitivity of the results to this uncertainty.

(4) Coding of moderator variables

(a) Individual-effect moderators—general

The main purpose of this review was not to find the mean effect size but to identify influential moderators of effect size. For each effect size the following 18 variables were coded.

(1) Study.
(2) Year of study publication.
(3) Species.
(4) Title—whether the title of the study included the term "egg size" or a similar term that suggested that egg size was the main focus of the study. Levels: (a) *Yes*—"egg size" appeared in the title; (b) *No*—"egg size" did not appear in title.
(5) Environment—levels: (a) *Wild*—the study was done in the wild (e.g. Parsons, 1970; Williams *et al.*,
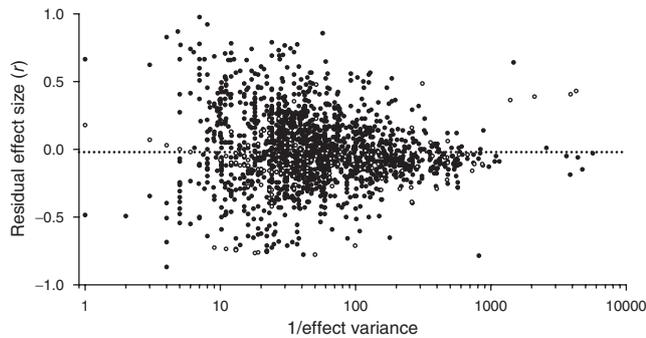
Fig. 2. Funnel plot. Residuals from the final model with the common-weighting scheme were used instead of raw effect sizes due to significant heterogenity (see Figs 3–4) in the latter. Solid circles ($N = 1629$): exact effect sizes; for these effects both exact magnitude and direction (positive or negative) were known. Open circles ($N = 176$): estimated effect sizes; either the exact magnitude or direction of these effects was unknown. The dotted line indicates weighted mean residual effect size. The line deviates slightly from zero on the y axis due to back-calculation of Zr residuals to r residuals.

1993*b*; Bogdanova, Nager & Monaghan, 2006), (*b*) *Captivity*—the study was done in captivity (e.g. Pinkowski, 1975; Ricklefs, Bruning & Archibald, 1986; Anderson, Reeve & Bird, 1997).

(6) Predictor—the predictor of offspring traits. Levels: (*a*) *Egg size*—predictor was egg size from which chicks hatched. (*b*) *Parental quality*—this was specific to the cross-fostering design, predictor was the size of eggs originally laid on the territory where cross-fostered chicks were raised (see e.g. Amundsen & Stokland, 1990; Reid & Boersma, 1990; Krist, 2009).

(7) Variance level—level at which predictor was measured. Levels: (*a*) *Between-clutch*—egg size was averaged within clutches (e.g. Schifferli, 1973; Magrath, 1992; Dawson & Clark, 2000). Also included were data on species that lay single-egg clutches (e.g. Feare, 1976; Weidinger, 1996; Silva *et al.*, 2007). (*b*) *Within-clutch*—only egg size variability at within-clutch level was used [see variable 8 for how this condition was achieved (e.g. Howe, 1976; Krist *et al.*, 2004; Maddox & Weatherhead, 2008)]. (*c*) *Total*—predictor was the size of an individual egg without taking into account which clutch it originated from (e.g. Parsons, 1970; Hořák & Albrecht, 2007; Oh & Badyaev, 2008). This was therefore a mix of within-clutch and between-clutch variation.

(8) Study design—this depended on the former variable. For variance measured at the between-clutch and total levels, three designs were distinguished: (*a*) *Cross-fostering*—eggs were cross-fostered between pairs of parents (e.g. Reid & Boersma, 1990; Styrsky, Eckerle & Thompson, 1999; Krist, 2009). (*b*) *Observational*—eggs were not cross-fostered (e.g. Lloyd, 1979; Ramos, 2001; Parker, 2002), (*c*) *Mixed*—mix of the two former designs. Statistical tests were based on pooled samples

of cross-fostered and non-cross-fostered clutches (e.g. Blomqvist, Johansson & Götmark, 1997; Selman & Houston, 1996; de Neve *et al.*, 2004) or eggs within a clutch (e.g. Ricklefs & Peters, 1981; Lessells, 1986; Williams, 1990). For variance measured at the within-clutch level, again three designs were distinguished: (*a*) *Pure*—either the mean egg size of the clutch was removed from the predictor by centring or dyads of eggs or groups of eggs were compared using a paired test within clutches (e.g. Howe, 1976; Leblanc, 1987; Krist *et al.*, 2004). All nestlings raised in a nest were siblings. (*b*) *Nest ID*—size of individual eggs was a predictor in the statistical model that also included nest identity, usually as a random factor (e.g. Ricklefs, 1984*b*; Rubolini *et al.*, 2006*b*; Whittingham, Dunn & Lifjeld, 2007). I confirmed on real data (Krist *et al.*, 2004, dataset available at Dryad Digital Repository, doi:10.5061/dryad.1758) that this approach leads to similar results as centring (results not shown). However, this result may not be generally applicable (see van de Pol & Wright, 2009) so this study design was separated from the above category. (*c*) *Not siblings*—Partial cross-fostering was done but the statistical tests were performed within broods by either of the two above approaches (pure or nest ID, smaller sample size disallowed their separation in this case; e.g. Ricklefs, 1984*a*; Rubolini *et al.*, 2006*a*; Bonisoli-Alquati *et al.*, 2008). This means that egg size variability was a mixture of the within-clutch and between-clutch variability but offspring traits might be affected by competition among nest-mates or parental food-allocation decisions within broods as in the two other within-clutch designs.

(9) Offspring stage—stage when the offspring traits were measured. Levels: (*a*) *Egg*, (*b*) *Hatchling* (measured on the day of hatching), (*c*) *Nestling*—measured while in the nest or before capable of flight. This period was defined as the time between age = 1 day and the mean fledging age for the species multiplied by 1.25. The multiplier was added to include cases when nestling development was somewhat slower than the average value for the species. (*d*) *Post-fledging*—offspring trait was measured after the mean fledging age x 1.25.

(10) Response—offspring trait that was dependent on egg size. These variables were recorded as they were named in the papers and then grouped into several broader categories. Levels: (*a*) *Hatching success* (hatched *versus* unhatched eggs, the latter may include both dead and infertile eggs; e.g. Murton, Westwood & Isaacson, 1974; Clifford & Anderson, 2002; d'Alba & Torres, 2007), (*b*) *Egg fertility* (infertile *versus* fertile eggs, the latter includes both hatched and unhatched eggs; e.g. Meathrel *et al.*, 1993; Wiebe & Bortolotti, 1995; Hernández *et al.*, 2008), (*c*) *Body mass* (e.g. Schifferli, 1973; Ricklefs, 1984*b*; Reed, Turner & Sotherland, 1999), (*d*) *Skeletal size*—e.g. tarsus, head, culmen length (e.g. Ankney, 1980; Weidinger, 1997; Isaksson, Uller &

Andersson, 2006), (*e*) *Condition*—body mass controlled for skeletal size or wing length (e.g. O'Connor, 1979; Nager, Monaghan & Houston, 2000; Silva *et al.*, 2007), (*f*) *Wing/feather length*—e.g. wing, tail, primary, rectrix length (e.g. Stempniewicz, 1980; Järvinen & Ylimaunu, 1984; Badzinski *et al.*, 2002), (*g*) *Survival*—chick survival (e.g. Parsons, 1970; Lundberg & Väisänen, 1979; Rutkowska & Cichoń, 2005); included also are cases where survival was scored as breeding success, i.e. survival from egg laying to fledging (22 effects in 14 studies; e.g. Zieliński & Bańbura, 1998; Ramos *et al.*, 2006; Louzao *et al.*, 2008), (*h*) *Activity*—e.g. locomotor performance (Goth & Evans, 2004), swimming speed (Anderson & Alisauskas, 2001), begging rate (e.g. Rubolini *et al.*, 2006*a*; Bonisoli-Alquati *et al.*, 2007), (*i*) *Immunity*—most often phytohemagglutinin (PHA)-induced immune response (e.g. de Neve *et al.*, 2004; Velando, Torres & Espinosa, 2005; Krist, 2009), but also maternal immunoglobulin G (IgG) levels and others (e.g. Pihlaja, Siitari & Alatalo, 2006), (*j*) *Growth rate of mass* (e.g. Nisbet, 1978; Amundsen, Lorentsen & Tveraa, 1996; Styrsky, Dobbs & Thompson, 2000), (*k*) *Growth rate of skeleton* (e.g. Bolton, 1991; Bitton, Dawson & O'Brien, 2006; Bogdanova & Nager, 2008), (*l*) *Growth rate of wing/feather* (e.g. Ricklefs, 1984*a*; Hipfner & Gaston, 1999; Quillfeldt & Peter, 2000), (*m*) *Life history/sexual trait*—traits measured on offspring, once they themselves became adult, e.g. clutch size, laying date, male ornaments (e.g. Cunningham & Russell, 2000; Parker, 2002; Krist, 2009).

(11) Number of variables—number of variables controlled when egg-size effects were tested. This is the sum of the covariates from the final model and those variables which were controlled by sample division into subgroups, e.g. males—females, first year—second year, first eggs—second eggs.

(*b*) *Individual-effect moderators—specific for some responses*

(12) Offspring age—age (in days) when the offspring traits were measured. Hatching day = 0. For some responses the precise age was unknown. Mean fledging age of the species was used as an estimate of age at fledging, peak mass and asymptotic mass from a fitted growth model. Age was coded for all responses with the exception of hatching success and egg fertility. Chick survival was measured between two ages (observational interval hereafter). Except for survival scored as breeding success or recruitment probability, the observational interval started with hatching in all but six cases. Due to the low variability of the start of the observational interval, age at the end of this interval was the only analysed variable.

(13) Causality of mortality—coded for hatching success and chick survival. Levels: (*a*) *Causal*—egg-size effects are likely to be causal (e.g. Lislevand *et al.*, 2005; Kontiainen *et al.*, 2008; Krist, 2009). An effort was made by the authors to control for

mortality factors that are unlikely to be affected by egg size. For example, nests which failed due to predation or abandonment were excluded from analyses. (*b*) *All losses*—factors listed in the point above were apparently not controlled for (e.g. Evans *et al.*, 2005; Budden & Beissinger, 2005; Fargallo *et al.*, 2006). (*c*) *Not-causal*—eggs/chicks died due to mortality factors that are unlikely to be affected by egg size, e.g. predation, abandonment (e.g. Hochachka, 1993; Boulton & Powlesland, 2008; Fernández & Reboreda, 2008).

(14) Type of growth measurement—levels: (*a*) *Absolute increase*—does not take into account initial size or mass differences—growth rate is measured as a slope of linear regression, or mass increment between two successive ages (e.g. Nisbet, 1978; Stokland & Amundsen, 1988; Gilbert *et al.*, 2006). (*b*) *Relative increase*—initial size or mass differences are taken into account—growth rate is measured by a growth constant from the logistic model, or chick mass is given on a logarithmic scale (e.g. Ricklefs *et al.*, 1986; Weidinger, 1997; Samelius & Alisauskas, 1999).

(*c*) *Species-specific moderators*

In addition to this individual-effect coding, some species-specific variables were recorded. These data were taken from *The Birds of the Western Palearctic* (Cramp & Perrins, 1977–1994), *The Birds of North America* (Poole, Stettenheim & Gill, 1993–2002), *Handbook of Australian, New Zealand and Antarctic Birds* (Higgins & Peter, 1990–2006) and *Handbook of the Birds of the World* (del Hoyo, Elliott & Sagartal, 1992–2006). For each species-trait combination all available data in one of these handbooks was coded and their mean was used for analyses.

(15) Development—developmental mode with levels: (*a*) *Altricial*, (*b*) *Precocial*, (*c*) *Semi-precocial* or semi-altricial, labelled hereafter as semi-precocial.

(16) Relative egg size—residuals from the regression ($\log_e$ egg volume $= -1.305 + 0.782 \times \log_e$ female body mass, $\mathcal{N} = 162$ species, $R^2 = 0.915$, $P < 0.001$) were used as an index of relative egg size. The results would be qualitatively the same if the regression was controlled for phylogeny (results not shown, see Appendix S2 for phylogeny of included species and methods of phylogenetic regression). Egg volume was computed from mean egg length and breadth, which was usually given in handbooks, according to Hoyt's (1979) formula. For four species only fresh egg mass was available. For these species egg volume was estimated based on a linear regression of egg volume on egg mass (egg volume $= 0.917 \times$ egg mass, no intercept, $\mathcal{N} = 138$ species, $R^2 = 0.998$, $P < 0.001$). For *Sterna hirundinacea* neither egg volume nor egg mass was available. Egg volume was estimated from the regression of egg volume on adult mass in five other *Sterna* species (egg volume $=$

9.900 + 0.0856 × female body mass, $\mathcal{N} = 5$, $R^2 = 0.978$, $P = 0.001$).

(17) Clutch size.

(18) Fledging age.

(5) Data analyses

(a) *Pseudoreplications, weighting schemes, and heterogeneity*

For statistical testing Pearson's *r* was transformed to Fisher's $\mathcal{Z}r$-transform using the formula given in Lipsey & Wilson (2001, p. 63). All presented values (least-square means and confidence intervals) were converted back to correlations.

The smallest units of analyses were individual effect size estimates which are called "estimates" hereafter. Since several estimates per study were usually available, the problem of pseudoreplications could arise. Two estimates were considered as clearly pseudoreplicated if they were based on the same sample and had the same combination of levels of independent variables 1 to 14 as listed above or if they differed only in variable 11. In those cases one of two selections was made. First, only one estimate was selected (blindly with respect to effect size) for analyses and the other was excluded as pseudoreplication. This was the case, for example if (1) two traits describing skeletal size such as tarsus and culmen length were tested with the same combination of levels of independent variables in the same study or if (2) the test was performed both on subsamples such as years (e.g. Williams *et al.*, 1993*b*) or laying orders (e.g. O'Connor, 1979) and on the composite sample. In this latter case, the test on the composite sample was always excluded to avoid the problem known as Simpson's paradox (see Borenstein *et al.*, 2009, p. 303–309). Second, some estimates were based on multiple contrasts. For example, survival to fledging was contrasted between groups of offspring hatched from small, medium, and large eggs. Three contrasts were computed (small–large, small–medium, medium–large) but they were not independent since each egg size category was involved in two contrasts. In this case the three estimates were merged into one composite estimate. A weighted mean was used, where weight was an inverse variance of the individual contrast. Sample size for this composite measure was the sum of the sample sizes in the three categories of eggs.

Despite the above treatments, estimates from the same study are still not independent. To take this non-independence into account, the study or the species was included as a random factor in the statistical models.

Another problem in meta-analysis is that estimates based on a large sample size should have greater weight than those based on a small sample size since the sampling error is greater in the latter case. Two types of weighting are used in meta-analysis—the fixed-effects model and the random-effects model (Borenstein *et al.*, 2009, p. 61). The fixed-effects model takes into account within-study variance only. As this model expects only one true effect size that is common to each study, it may be called the common-effect model (Borenstein *et al.*, 2009, p. 61). This notation will be used hereafter. In most instances the random-effects model is more appropriate

(Borenstein *et al.*, 2009, p. 86) since it also takes into account between-study variance, which is likely to be non-trivial in ecological studies. However, random-effects models are more difficult to compute. The main aim of the present study was to find factors, called moderators in meta-analysis, that affect the strength of the relationship between egg size and offspring traits. This type of meta-analysis is sometimes called meta-regression and the methods to solve it are not implemented in software specially developed for meta-analysis. This special software including *Comprehensive Meta-analysis* and *MetaWin*, is more oriented to the computation of mean effect size rather than on taking moderator effects into account and allows only one covariate in the computation of the mean effect size. Therefore, *SAS* software was used for analyses (SAS Institute, 2003). *SAS* enables computation of both fixed and random-effects models in meta-analysis while offering the possibility to control for many covariates (van Houwelingen, Arends & Stijnen, 2002). Unfortunately, in the case of the present analysis, sample size was too large for a random-effects model to be computed in combination with the large number of covariates as indicated by the "Out of memory" statement in the Log of *SAS*. Therefore two other analyses were conducted.

First, common-effect analysis was conducted where estimates were weighted by the inverse of their variance. This inverse variance is equal to $n - 3$ for effect size expressed as Fisher's $\mathcal{Z}r$ (Lipsey & Wilson, 2001, p.72). In the case of dichotomization and range enhancement/restriction, the variance of the estimate had to be adjusted according to the formula given in Borenstein *et al.* (2009, p. 343). Second, unweighted analysis was conducted for the following reason. The preferred method—the random-effects model—weights estimates by the sum of the within-study variance and between-study variance when the latter is the same for all estimates (Borenstein *et al.*, 2009, p. 73). Consequently, the random-effects model weights estimates more equally than common-effect meta-analysis but less equally than in unweighted analyses where weights are the same for all estimates by definition. Consequently, good congruence between common-effect and unweighted analyses would also suggest that random-effects analysis would provide similar results. Some recent meta-analyses used solely unweighted analyses (Schoech & Hahn, 2008).

Heterogeneity between effect-size estimates was assessed with the Q test and I^2 statistic. Q is the weighted sum of squares that is distributed as chi-squared with degrees of freedom equalling the number of estimates minus one (Borenstein *et al.*, 2009, p. 109–110). I^2 is the proportion of the observed variance that reflects real differences in effect sizes (Borenstein *et al.*, 2009, p. 117).

(b) *Model selection and collinearity*

First a random part of the model was selected. Either the study or the species was used as the subject within which both the intercept and the slopes of the independent variables were nested. The best covariance structure was selected according to Akaike's information criterion (AIC). It was not possible to

include both the study and the species in the same model if the slopes varied within subjects. Such a model would involve the computation of many random effects, some of which were crossed. Crossed random effects are more difficult to estimate than nested ones (West, Welch & Gałecki, 2007, p. 14). This probably explains why *SAS* was unable to fit this model. In common-effect weighting schemes the best models include the study as a subject. In unweighted analyses the best models include the species as the subject. Recently, methods have been proposed on how to include all phylogenetic information into meta-analysis (Adams, 2008; Lajeunesse, 2009). However, it may be difficult, if not impossible, to include phylogeny in such complex models as those fitted here. Therefore I did not control for phylogeny and the results should be viewed with this caveat in mind.

After a random part of the model was selected, the fixed part was selected. Starting with a full model that included the independent variables numbered 1-11 and 15-17 non-significant variables were eliminated and the final model included only the significant ones.

Large correlations between independent variables cause problems in estimation of regression coefficients and their standard errors. Models that include predictors with variance inflation factors (VIFs) less than 10 (Quinn & Keough, 2002) or 5 (Zuur, Ieno & Smith, 2007) are usually considered to give acceptable results although also more stringent criteria have been suggested (Graham, 2003). For each independent variable its VIF was estimated by the sequential method described in Zuur *et al.* (2007, p. 469). The structure of the models searching for predictors' VIFs had to be simplified (no random factors and no nested structure in the dependent and in some cases also in the independent variable). Consequently, resulting VIFs may be considered only as estimates of the true VIF in more complex models. Estimated VIFs were always less than 5 (see Tables 3, 4). Therefore, results of the presented models are unlikely to be greatly affected by multicollinearity between independent variables.

(c) *Additional models for some responses*

For some of the responses additional models were fitted. Since the two weighting schemes provided similar results for models based on all data, these additional models were fitted for specific responses with the common-effect weighting scheme only. These models included variables that were significant in the model based on all data and the variables 12-14 as the factor of interest. For simplicity, they included only the intercept in their random parts.

Whether chick age affects the correlation between egg size and offspring trait was tested on the four responses with the largest sample size (chick survival, body mass, skeletal size, and wing/feather length) in the nestling stage. All cases where survival was recorded as breeding success were excluded from the analysis of age-effect on chick survival. The prediction of age-effect differs for egg size (decreasing effect with age) and parental quality (increasing with age). Ideally, this should be tested as an interaction effect between the age and the

predictor. This was possible to do only with nestling body mass, where sufficient data for parental quality existed. For the other three responses, estimates based on parental quality were excluded from the dataset. The distribution of chick age was skewed to the right. Data points of the predictor variable that depart considerably from the rest of the distribution may strongly affect regression results. Therefore, a second set of models was fitted without points that departed more than 3 S.D. from the mean chick age (see Grafen & Hails, 2002, pp. 40–42). Chick age relative to fledging age might be a more relevant measure of chick age than the absolute age of the chick. Therefore, a third set of models was fitted in which chick age was transformed to relative age (relative age = chick age/fledging age of the species).

Whether causality of mortality affects the relationship between egg size and hatching success or nestling survival was also tested. Finally whether the type of growth measurement affects the correlation between egg size and rate of increase in chick mass was also tested.

(6) Publication bias

Publication bias is a potential problem for both narrative and meta-analytic reviews (Møller & Jennions, 2001; Borenstein *et al.*, 2009). Several methods were employed to deal with publication bias. First, bias was minimized in the included studies by (1) a comprehensive search of the literature which also included non-English studies (see Table 1). Without their inclusion the review might be especially prone to bias (Møller & Jennions, 2001; Gates, 2002); (2) contacting authors for additional details if published studies did not contain enough information to enable computation of effect size, which was most often the case if results were non-significant; (3) not excluding studies for which all necessary information to compute effect size was unobtainable. Instead the size of these effects was estimated and an analysis was conducted with and without these estimated effect sizes (i.e. sensitivity analysis; Gates, 2002).

Second, whether the included effect sizes are likely to be biased was assessed. (1) For each study, two variables that might reveal bias were coded. First, it was coded if the title of the study contained the phrase "egg size" or some similar term. In these studies egg-size effects are likely to be the main focus of the research. If publication bias was substantial, effect sizes in these studies would be larger than in studies less focused on egg-size effects. Second, year of publication

Table 1. Number of considered and used studies ordered by the language of publication.

Language	Considered	Used
English	582	278
Russian	20	0
Chinese	15	2
German	15	0
French	7	1
Other 10 languages	27	2
Total	666	283
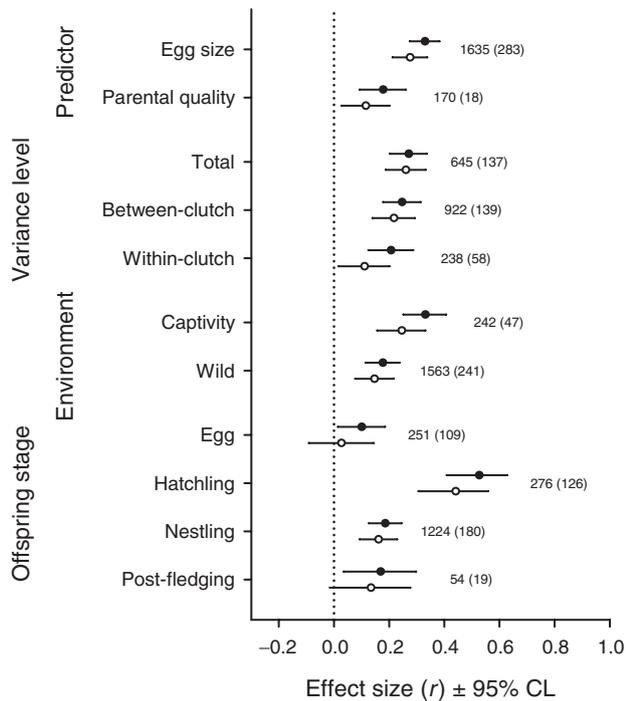
Fig. 3. Effect size for four independent variables that significantly explained variability in effect sizes. Independent variable "Variance level" was significant only in the unweighted analysis. Displayed are the least-square means (LSM) ± 95% confidence limits (CL) for each level of the independent variable. Values are controlled for independent variables that were retained in the final model. Number of estimates and studies (in parentheses) is given for each level of the independent variable. Solid circles: LSM from models with a common-effect weighting scheme. Open circles: LSM from unweighted analyses.

was also coded. It is known that relationships often fade with time which is likely to be due to publication bias (Jennions & Møller, 2002). (2) A funnel plot was constructed to assess whether there was a lack of small or negative effect sizes in small-sample size studies which would be one common form of publication bias (Møller & Jennions, 2001; Borenstein *et al.*, 2009). Residuals from the final model that included significant moderators of effect size were used for the funnel plot. If raw effect sizes were plotted, their substantial heterogeneity caused, for example, by different stages in the offspring life cycle and different responses (see Figs 3, 4), might cause asymmetry in the funnel plot even if no publication bias existed.

(7) Interpretation of effect size

For each effect size several criteria can be evaluated. (1) *Direction*—whether the relationship between egg size and offspring quality indicator is positive or negative. (2) *Absolute magnitude*—according to Cohen's (1988, pp. 77–81) convention, effect size is considered as large if $r = 0.5$, medium if $r = 0.3$, and small if $r = 0.1$. Møller & Jennions (2002) have shown that in the field of ecology and evolutionary biology the mean correlation between the major
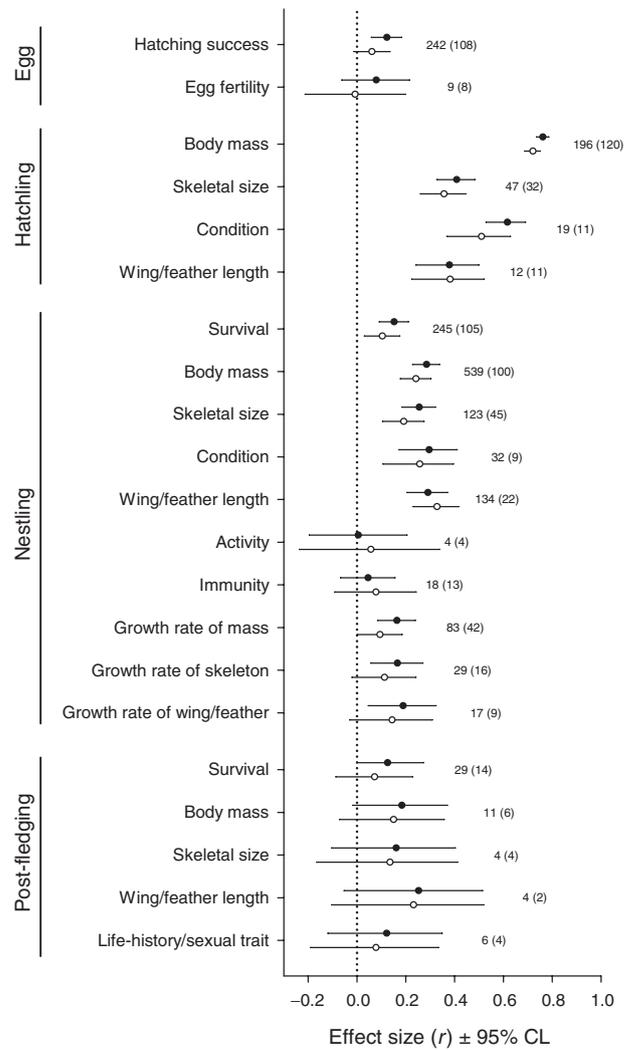
Fig. 4. Effect size for the last significant independent variable—response nested within the offspring stage. Displayed are the least-square means (LSM) ± 95% confidence limits (CL) for each level of the independent variable with the exception of hatchling immunity and activity since these LSM were based on only one estimate. Values are controlled for independent variables that were retained in the final model. Number of estimates and studies (in parentheses) is given for each level of the independent variable. Solid circles: LSM from models with a common-effect weighting scheme. Open circles: LSM from unweighted analyses.

factor of interest and the response variable is $r = 0.19$. These values can be used as yard-sticks to place results from the present study into a broader context. (3) *Precision*—effect sizes are accompanied by confidence intervals; those with narrow confidence intervals are estimated with high precision. (4) *Statistical significance*—effect sizes whose confidence intervals do not overlap zero are considered to be statistically significant. (5) *Relative magnitude*—effect sizes may differ among levels of moderator variables. Inference may be made by statistical test of the moderator variable and by visual inspection of figures with plotted effect sizes.

Table 2. Taxonomic distribution of effects included in the meta-analysis. Values are numbers of species, studies, and estimates.

Order	Species	Studies	Estimates
Charadriiformes	42	72	493
Passeriformes	41	91	515
Anseriformes	17	35	242
Procellariiformes	15	18	235
Falconiformes	9	16	65
Pelecaniformes	8	11	79
Gruiformes	8	10	60
Sphenisciformes	6	10	53
Other 10 orders	16	20	65
Total	162	283	1805

III. RESULTS

(1) Description of dataset

In sum, the three sources of data (electronic databases, reference lists and accidentally found studies) provided approximately 5000 candidate studies. If the study was carried out on bird species and published in a non-poultry-science journal, I read its abstract. Based on the reading of the abstracts, 666 studies were considered as containing potentially relevant data and their full texts were searched.

Of these studies, 297 contained data of interest (i.e. passed through selection criteria 1–4) but 14 were excluded because of problematic design, analyses etc. Appendix S1 provides a list of excluded studies and the reasons for their exclusion. Consequently, the final number of studies was 283 (also listed in Appendix S1). These studies contained 2318 estimates. After the exclusion of pseudoreplications, the final sample of estimates was reduced to 1805. These studies were carried out on 162 species distributed among 18 orders (see Appendix S2). The vast majority of data was obtained on a few waterfowl orders and passerines (Table 2).

The first study was published in 1970 (Parsons, 1970) and the number of published studies increased throughout the years (Fig. 1). This increase was steeper than the general increase in the number of scientific publications, documenting a proportional increase of interest in egg-size effects mainly during the period 1970–2000 (Fig. 1). Unfortunately, the growing interest was not accompanied by a more rigorous publication of effects. On the contrary, the number of studies that published incomplete effect sizes rose disproportionately in the last decade (Fig. 1).

Effect sizes were significantly heterogeneous ($Q = 20224.7$, d.f. $= 1804, P < 0.001$). A high proportion of observed variance reflected real differences in effect sizes ($I^2 = 91.1\%$). Even after the random effect of the study or species was accounted for, effect sizes still remained substantially heterogeneous (see Figs 3, 4).

(2) Effect of moderators in the two weighting schemes

In both the common-effect weighting scheme and the unweighted analyses the best covariance structure included slopes of independent variables nested within subjects. The subject was the study in the case of the common-effect weighting scheme and the species in the unweighted analyses. These models were better according to AIC than either the models with only a random intercept or those lacking the random part at all.

Despite different weighting methods and subjects within which slopes were allowed to vary, the two models provided quite similar results for fixed variables. In both models, the predictor, environment, offspring stage, and the response that was nested in the offspring stage were found to be significant (Tables 3, 4). The least-square means (LSM) for levels of these categorical variables were also in good congruence between the two models (Figs 3, 4), although generally unweighted analyses provided a somewhat lower LSM than the common-effect weighting scheme.

Offspring quality was more correlated with egg size from which the young hatched than with egg size that was laid in territories on which cross-fostered young were raised (a surrogate of parental quality), although the latter relationship was also significantly positive (Fig. 3).

Studies performed in captivity found a larger effect size than those carried out in the wild. In theory this is expected at the hatchling stage since hatchlings may be weighed before receiving any food in captivity (i.e. usually hatched in an incubator). On the other hand, effect size in the nestling stage is expected to be larger when food resources are scarce (McGinley et al., 1987; Smith & Bruun, 1998) which is more typical in the wild. I tested for the possibility that the effect of the environment depends on the offspring stage by inclusion of the interaction between the two variables in the final model. This interaction was not significant (common-effect model: $F_{3,396} = 2.15, P = 0.094$) and the least-square means revealed similar or larger effect sizes in captivity compared to the wild regardless of offspring stage (stage: LSM for effect size in captivity, LSM for effect size in the wild; egg: 0.014, 0.032; hatchling: 0.619, 0.450; nestling: 0.251, 0.106; post-fledging: 0.233, 0.091). Unweighted analysis produced a similar pattern (results not shown).

Unsurprisingly, hatching was the stage when effect sizes were the largest with absolute magnitude classified as "large" according to Cohen's (1988) convention (see Figs 3, 4). The lowest effect size was found for the egg stage (hatching success, egg fertility), where effects were weak, although some of them were statistically significant due to large sample size (Figs 3, 4). Effect sizes were weak to medium for both nestling and post-fledging stages, although in the latter case effect sizes were accompanied by much wider confidence intervals due to a smaller sample size (Figs 3, 4).

The largest effect of type of response was evident in the hatchling stage when egg size was much more correlated with body mass than with body condition, and especially with skeletal size and wing/feather length (Fig. 4). In the nestling stage, effect sizes were similar for all morphological traits, lower for the survival and growth rates and non-significant for activity and immunity traits (Fig. 4). In the post-fledging stage effect sizes for all traits were similar but

Table 3. Results of common-effect weighing scheme: effects of all considered predictors on effect size. For the fixed part of the model both significant and non-significant predictors are shown. The latter are presented in the order they were eliminated from the model. A random part of the model is presented in its final form. The subject is study. (Random slopes nested within study). F/Z = test statistic, NDF = numerator degrees of freedom, DDF = denominator degrees of freedom, S.E. = standard error, estVIF = estimated variance inflation factor.

	F/Z	NDF	DDF	P	Parameter	S.E.	estVIF
Random part							
Predictor	*2.19*			*0.014*	*0.00484*	*0.00221*	
Study design (Variance level)	*2.51*			*0.006*	*0.00951*	*0.00379*	
Offspring stage	*4.43*			*<0.001*	*0.0206*	*0.00464*	
Response (Offspring stage)	*4.11*			*<0.001*	*0.0101*	*0.00247*	
Residual	*23.48*			*<0.001*	*1.811*	*0.0771*	
Fixed part							
(a) Final model							
Intercept							
Predictor	25.64	1	36.6	<0.001			1.14
Environment	20.34	1	414	<0.001			1.32
Offspring stage	11.08	3	572	<0.001			2.21
Response (Offspring stage)	13.60	19	250	<0.001			1.44
Year of publication	9.11	1	278	0.003	−0.00420	0.00138	1.11
(b) Eliminated terms							
Title	0.01	1	298	0.917			1.49
Study design (Variance level)	0.38	6	134	0.892			3.02
Relative egg size	0.03	1	329	0.863	0.00736	0.0427	3.85
Developmental mode	1.15	2	300	0.319			1.93
Clutch size	1.26	1	362	0.262	0.00513	0.00457	1.33
Variance level	2.11	2	130	0.125			1.33
Number of variables	3.02	1	434	0.083	−0.0127	0.00730	1.26

Table 4. Results of unweighted analyses: effects of all considered predictors on effect size. For the fixed part of the model both significant and non-significant predictors are shown. The latter are presented in the order they were eliminated from the model. A random part of the model is presented in its final form. The subject is species. (Random slopes nested within species). F/Z = test statistic, NDF = numerator degrees of freedom, DDF = denominator degrees of freedom, S.E. = standard error, estVIF = estimated variance inflation factor.

	F/Z	NDF	DDF	P	Parameter	S.E.	estVIF
Random part							
Predictor	*1.03*			*0.150*	*0.00190*	*0.00183*	
Variance level	*3.19*			*<0.001*	*0.0163*	*0.00511*	
Offspring stage	*2.85*			*0.002*	*0.0167*	*0.00586*	
Response (Offspring stage)	*3.62*			*<0.001*	*0.0141*	*0.00390*	
Residual	*24.30*			*<0.001*	*0.0578*	*0.00239*	
Fixed part							
(a) Final model							
Intercept							
Predictor	33.53	1	22.1	<0.001			1.23
Variance level	8.09	2	115	<0.001			1.19
Environment	6.87	1	327	0.009			1.57
Offspring stage	7.29	3	506	<0.001			2.39
Response (Offspring stage)	8.79	19	278	<0.001			1.46
Number of variables	4.16	1	1210	0.042	−0.0158	0.00775	1.20
(b) Eliminated terms							
Title	0.02	1	791	0.880			1.33
Study design (Variance level)	0.49	6	872	0.818			2.65
Developmental mode	1.57	2	173	0.210			2.78
Clutch size	1.03	1	207	0.312	0.00659	0.00650	1.65
Year of publication	2.21	1	962	0.138	−0.00192	0.00130	1.41
Relative egg size	3.21	1	171	0.075	−0.0681	0.0380	1.29

significant only for survival because of the larger sample size in this variable (Fig. 4).

The two weighting schemes disagreed on the significance of variance levels, which were found to be significant in unweighted analysis but non-significant in the common-effect one. They also differed in assessing the effect of continuous variables: in the unweighted analysis the "number of variables" was retained in the final model while the "year of publication" was retained in the common-effect model. However, note that all these variables that were retained in one but were eliminated from the other model, were only eliminated at the end of the backward elimination procedure (Tables 3, 4).

Both weighting schemes agreed on the non-significance of all species-specific variables (developmental mode, relative egg size, clutch size), title of the study, and study design nested within variance levels (Fig. 5; Tables 3, 4). As the two weighting schemes provided closely similar results in the main statistical tests as described above, additional statistical tests were performed using only the common-weighting scheme.

(3) Additional moderators for some responses

Effect size did not change significantly as the young grew older if the response was chick survival or wing/feather length but decreased if the response was skeletal size or body mass (Table 5; Fig. 6). Models without extreme data points and with relative chick age fitted instead of actual age provided both qualitatively and quantitatively closely similar results (results not shown). A decrease in effect size on body mass with age was steeper if the predictor was egg size (slope in Zr units $= -0.00625$) than parental quality (slope $= -0.00156$, see Fig. 6B; test of this interaction: $F_{1,456} = 7.54, P = 0.006$).

A visual inspection of the plotted data suggested a non-linear, convex effect of chick age on the magnitude of the effect size on body mass. First, most residuals, after the age of 50 days, were positive (Fig. 6B). Second, the convex shape would be even more evident if the data on hatchlings were included: the LSM for the body mass of hatchlings is about $r = 0.7$ (Fig. 4), while the intercept for the data based on nestlings only is about $r = 0.4$ (Fig. 6B). Third, effect size was still positive in the post-fledging stage (Fig. 4). The last point also holds for chick skeletal size. The hypothesis about the non-linear relationship was only set *post hoc*, therefore the formal test was not performed. Instead, the linear lines were divided into two parts: the solid region extending over the chick ages with most of the data while the dotted line extends to the high end of the *x* axis where data were more scarce and therefore prediction was less reliable (Fig. 6).

The type of mortality did not significantly affect the effect size for either hatching success ($F_{2,113} = 0.61, P = 0.544$; level: LSM, number of estimates, number of studies: all losses: 0.070, 107, 20; causal: 0.034, 99, 57; uncausal: 0.057, 36, 12) or chick survival ($F_{2,140} = 2.50, P = 0.086$; all losses: 0.268, 96, 54; causal: 0.192, 93, 38; uncausal: 0.152, 15, 3). The same was true for the type of measurement of the growth
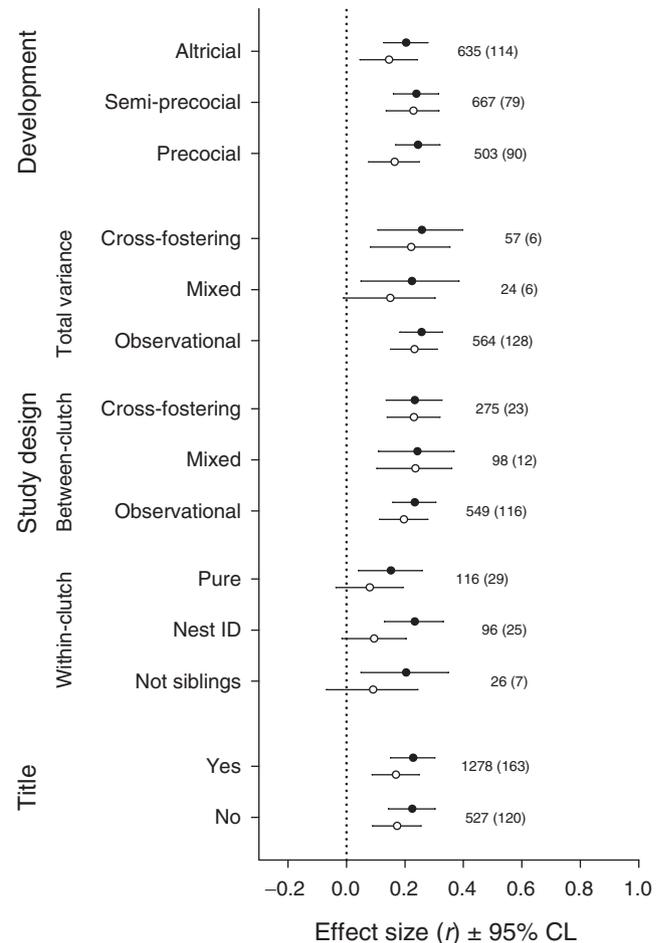
Fig. 5. Effect size for independent variables that did not significantly explain variability in effect sizes. Displayed are the least-square means (LSM) $\pm$ 95% confidence limits (CL) for each level of the independent variable. Values are controlled for those independent variables that were retained in the model at the time of exclusion of the independent variable in question. The number of estimates and studies (in parentheses) is given for each level of the independent variable. Solid circles: LSM from models with a common-effect weighting scheme. Open circles: LSM from unweighted analyses.

rate of mass ($F_{1,36.7} = 2.18, P = 0.148$; absolute increase: 0.081, 43, 21; relative increase: 0.009, 40, 21).

(4) Publication bias

(a) Avoidance of publication bias

Most of the studies considered as potentially containing data were written in English but the number of non-English-written studies was also substantial (Table 1). The consideration of non-English-written studies was intended to reduce publication bias (Gates, 2002). However in this study, bias would not arise if the search was restricted to English-written studies since the number of non-English-written studies that contained data was very small (Table 1).

Table 5. Tests of the relationship between nestling age and effect size for four nestling traits. Negative parameter (regression coefficient) means that the correlation between egg size and nestling trait decreases as the young grow older. The body mass model included interaction between predictor (egg size or parental quality) and chick age. All other tests were only based on egg size as a predictor. See text for further details. F = test statistic, NDF = numerator degrees of freedom, DDF = denominator degrees of freedom, S.E. = standard error.

Nestling trait	F	NDF	DDF	N	P	Parameter	S.E.
Survival	0.73	1	75.8	204	0.396	−0.000380	0.000445
Body mass	14.16	1	524	539	<0.001		
Skeletal size	14.41	1	93.3	111	<0.001	−0.00490	0.00129
Wing/feather length	0.14	1	116	120	0.708	0.000752	0.00201
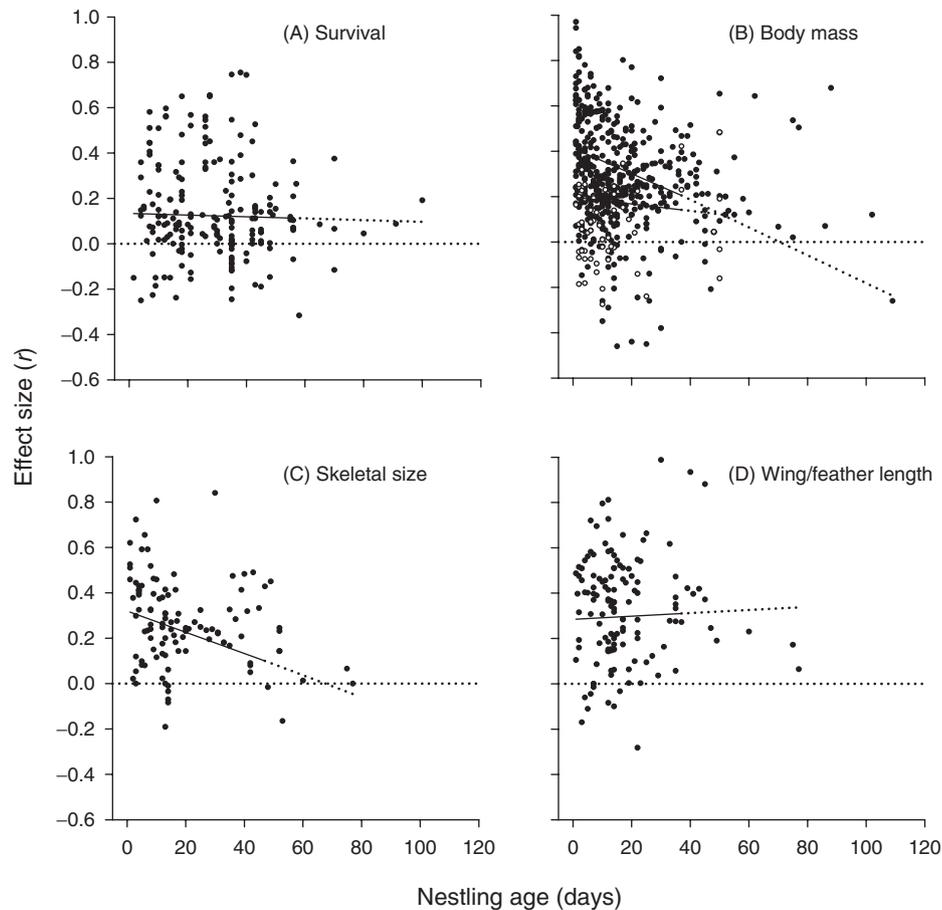
Fig. 6. Relationship between nestling age and effect size of four traits for which sufficient sample sizes were available. Solid circles: effect sizes where predictor was egg size. Open circles: effect size where predictor was parental quality (i.e. egg size laid originally on territory where cross-fostered nestlings were raised). Fitted lines are predicted from models with a common-effect weighting scheme. Solid part of lines: 90% of data is in this range of the *x* axis. Dotted part of lines: only 10% of data in this range of the *x* axis. Three outliers are not displayed for nestling survival with coordinates [275, −0.061], [275, 0.123], and [135, 0.236], all obtained on albatrosses. One outlier is not displayed for skeletal size, coordinates [12, −0.800], whose effect size is based on $N = 7$. See Table 5 for statistical tests.

Of 297 studies that contained relevant data, all information necessary to compute effect size and its variance was published in 158 studies while only incomplete information was available in 139 studies (Fig. 1). Because of poor experimental design or pseudoreplication, three studies with complete information and two with incomplete information were excluded. An attempt was made to obtain missing information from the authors of the 137 suitable studies with missing information. Of these, 62 studies provided the necessary information. The information that was provided by the authors on request is given in red italics in Appendix S3, sheet "all data". Effect sizes provided by authors on request were smaller (weighted mean $r = 0.064$, $N = 120$) than those published ($r = 0.210$, $N = 1554$, $F_{1,1672} = 48.52$,

$P < 0.001$), as was found also in other meta-analyses (Cassey *et al.*, 2004). For the remaining 75 studies, necessary information was not obtained because of the following reasons: 14 authors were not located, 20 did not have the missing information, 18 gave only initial responses, and 23 authors did not respond at all. Fortunately I was able to estimate effect size or its variance (see above for estimation methods) in 66 of these studies. Consequently, only nine more studies from all analyses had to be excluded and these studies did not seem to have extremely large or small effect sizes, so their exclusion is unlikely to bias the results.

(b) Assessment of remaining bias

The main analyses were performed with both the exact known and the estimated effect sizes. These analyses were repeated with the exact known effect sizes only to look at the sensitivity of the results to inclusion of the estimated effect sizes. The results of both models with common-effect weighting and no weighting showed that the results were robust. Only three differences were found when restricted datasets were used instead of the complete one. First, both number of variables and publication years were significant variables in unweighted models. Second, the variance level was no longer a significant variable in unweighted analysis ($P = 0.108$). Third, most effect sizes expressed as the least-square means were greater by $0.01 - 0.05$ in the restricted dataset as compared to the full dataset. As most conclusions would be the same if a restricted dataset was used, this meta-analysis is not very sensitive to publication bias.

Low publication bias is also suggested by the non-significance of the title of the study (Tables 3, 4; Fig. 5) while some upper-bias is indicated in early studies as effect size decreased with the year of publication (Tables 3, 4). However, this might be partly caused by greater control of confounding variables in more recent studies as a positive correlation between the year of the study and the number of controlled variables exists ($r = 0.278$, $P < 0.001$, $N = 1805$). The funnel plot does not indicate publication bias with respect to the sample size upon visual inspection (Fig. 2). However, Spearman rank correlation between residual effect size and inverse variance of effect size is negative and significant ($r_s = -0.096$, $P < 0.001$, $N = 1805$). Therefore the "trim and fill" method (Duval & Tweedie, 2000) was used to estimate number of estimates missing on the bottom side of the funnel plot. The L_0 estimator suggested that only three estimates were missing. Filling these three estimates with very small weight had negligible effect on the overall mean effect size (results not shown).

IV. DISCUSSION

Egg size was positively correlated with offspring quality across all stages in the offspring life cycle—from egg to post-fledging, as well as across most studied offspring traits. This analysis provides strong support for the offspring size—quality relationship since it is based on a quantitative analysis of a very large number of primary studies. The results are unlikely to be much affected by publication bias because all published, non-significant results also were included. Consequently, there was little evidence for a publication bias among the included studies. Furthermore, the trim and fill method suggests that only a few studies with small effect sizes were unpublished. This might be because positive egg-size effects are predicted by theory, and finding no effect or even a negative one is of interest and therefore reported by authors. One exception may be when the response variable is measured on a binary scale, such as hatching success or chick survival. If nearly all eggs hatch/do not hatch or all chicks survive/die, authors may not test for the relationship between egg size and mortality as it is clear that this correlation will be weak (see e.g. Bitton *et al.*, 2006; van de Pol *et al.*, 2006). So estimates of effect sizes on binary traits are likely to be somewhat upper biased.

(1) Cross-fostering and post-hatching care

Birds take care of their offspring after hatching. In theory, this could be another source of upper bias in estimates of effect sizes since parents that are able to lay large eggs may also be able to provide more food for their chicks. The correlation between egg size and chick quality might arise through the correlation of these two variables with a third, unmeasured one, such as territory or parental quality (Birkhead & Nettleship, 1982; Bolton, 1991). By cross-fostering clutches randomly between nests this latter correlation is broken and the independent effects of parental quality and egg size may be estimated simultaneously (Amundsen & Stokland, 1990; Reid & Boersma, 1990).

The present review led to an unexpected result—studies that employed a cross-fostering design did not find weaker effect sizes than observational studies despite the fact that the former also found a positive correlation between parental quality and offspring performance. This apparent paradox may be explained in two ways. First, the authors of the cross-fostering studies might select different subjects and better control the confounding variables than was done in the observational studies. This might increase the estimate of effect size in the former set. In other words, a comparison of cross-fostering and observational studies has an observational nature since the treatments were not allocated at random. Moreover, selection of nests within a treatment is also not random in the case of the cross-fostering design which requires dyads of nests. Second, parental provisioning and offspring demands may be coadapted (Wolf & Brodie, 1998; Kölliker, Brodie & Moore, 2005; Lock *et al.*, 2007). By cross-fostering, the coadapted phenotypes are disassociated which might induce changes in parental provisioning behaviour. Under some conditions, such post-hatching effects can be ascribed to the effect of the size of the cross-fostered eggs (see Krist & Remeš, 2004).

Although cross-fostering decouples the correlation between egg size and parental quality, it does not ensure zero correlation between egg size and parental post-hatching care. Such correlations may arise if parents plastically adjust

their provisioning behaviour to the offspring state that is co-determined by egg size (Krist & Remeš, 2004). Therefore, it is important to include the intensity of post-hatching care as a covariate in the analysis of pre-hatching effects (Krist & Remeš, 2004), as well as pre-hatching effects when testing for post-hatching ones (Russell *et al.*, 2007). So far, only a few studies have directly tested for the covariation between egg size and post-hatching parental care and its effect on offspring quality (Quillfeldt & Peter, 2000; Russell *et al.*, 2007; Krist, 2009). However, one of the indirect findings from this review suggests that these correlations may generally be either weak or non-existent. If parents compensated for differences in pre-hatching investment by differential provisioning, effect size should be smaller in altricial compared to precocial species (Magrath, 1992; Williams, 1994). However, the developmental mode did not predict effect size.

(2) Between-clutch *versus* intraclutch effects

There are several reasons why the level at which the variance of egg size is measured should affect the strength of effect size on offspring quality. The first is statistical. All other things being equal, less variance in the independent variable means a lower effect size (Hunter & Schmidt, 2004, pp. 37–39). In birds, egg size is variable mainly between clutches with only about 30% intraclutch variation (Christians, 2002). Consequently, for this statistical reason, effect size should decrease in the order: total > between clutch > intraclutch variance used. The second reason is ecological and more interesting. Sibling rivalry often leads to monopolization of resources by the larger siblings and starvation or even death of the smaller one (Mock & Parker, 1997; Forbes & Wiebe, 2010). Furthermore, parents may actively enhance or mitigate within-brood competitive asymmetries by differential food allocation (Krebs, 2002). The third reason is a quantitative-genetic one. Unlike between-clutch studies, intraclutch ones do not suffer from correlations of egg size with direct genetic effects (Krist & Remeš, 2004). Most often, this correlation is probably positive (see Riska, Rutledge & Atchley, 1985; McAdam *et al.*, 2002) and therefore causes an upper bias in the estimates of egg-size effects between clutches (Krist & Remeš, 2004). On the contrary, although egg size may be correlated with other pre-hatching effects in all non-manipulative designs (Krist & Remeš, 2004) this correlation may be higher in within-clutch compared to between-clutch settings (see Reed *et al.*, 2009; Kozlowski & Ricklefs, 2010) and therefore cause upper bias in egg-size effects in the former compared to the latter design.

In the present study, effect size at the within-brood level was smaller than those at total or between-broods levels, which suggests a role of smaller egg-size variation or compensating effects of parental provisioning, or increased bias in the latter two designs due to a confounding direct genetic effect. Partial resolution of these hypotheses is offered by studies that performed partial cross-fostering, increasing egg-size variation within nests, and then looked at the relationship between egg size and offspring performance within broods. If egg-size variation was highly important,

effect size should be higher in such a setting compared to a pure within-clutch design. This was not the case. However, the number of these studies was quite limited. Therefore the conclusion, that variation is of minor importance, is weak. The remaining two hypotheses are even more difficult to assess at present. The correlation between the egg size and the direct genetic effects did not upwardly bias the results of the one between-clutch study (Krist, 2009), although this effect is hypothesized to exist in frogs (Ficetola & de Bernardi, 2009; but see Dziminski & Roberts, 2006). No study looked at within-brood egg-size effects controlled for parental food provisioning. This remains a challenge for future research.

(3) Environmental quality and effect size

Effect size was generally larger in captive compared to wild populations. The strength of the selection on egg size is expected to differ between environments which ultimately may explain differences in egg sizes among populations and species (Fox & Czesak, 2000), although other factors often play an even larger role (Moles *et al.*, 2005; Martin *et al.*, 2006). Contrary to my finding of larger effect size in captivity, theoretical models usually assume a greater dependence of offspring fitness on egg size in harsh, more competitive environments (Brockelman, 1975; Parker & Begon, 1986; McGinley *et al.*, 1987). Empirical studies in non-avian taxa that manipulated the quality of the offspring environment generally supported this assumption in plants (Rey *et al.*, 2004; Quero *et al.*, 2007; review in Donohue & Schmitt, 1998), invertebrates (Fox, 2000; Agosta, 2008; Allen, Buckley & Marshall, 2008; review in Fox & Czesak, 2000), fish (Hutchings, 1991; Einum & Fleming, 1999; Bashey, 2006), and amphibians (Parichy & Kaplan, 1992; Dziminski & Roberts, 2006). However, some studies found the opposite pattern in amphibians (Semlitsch & Gibbons, 1990), reptiles (Svensson & Sinervo, 2000), and mammals (Oksanen *et al.*, 2003).

Given the theoretical importance of the concept of selection varying with environmental quality, a surprisingly limited number of studies have dealt with this problem in birds. A few observational studies found a stronger relationship between egg size and offspring quality in harsh, more competitive environments (Smith & Bruun, 1998; Styrsky *et al.*, 1999; Garant *et al.*, 2007). Only two studies were specifically designed to solve this question and manipulated the offspring post-hatching environment either by food supplementation (Styrsky *et al.*, 2000) or brood-size manipulation (Bonisoli-Alquati *et al.*, 2008). Both of these studies did not find a difference between effect size in good *versus* poor conditions. If the lack of an effect of environmental quality is a general pattern in birds, the finding of larger effect sizes in the less-competitive conditions in captivity could be explained by better control of confounding variables.

(4) Components of offspring fitness and types of studied traits

Egg size was positively correlated with nearly all studied traits across all stages in the offspring life cycle. Does this

finding mean that egg size has a positive effect on offspring fitness? In iteroparous organisms, such as birds, fitness has three main components: juvenile survival (survival from egg to sexual maturity), adult survival, and fecundity (Stearns, 1992; Roff, 2002).

Egg size likely affected the first component—juvenile survival. Chicks hatching from large eggs had enhanced components of juvenile survival such as hatching success and nestling survival. They were also significantly larger and had slightly enhanced immunity. These traits often are predictive of post-fledging survival (tarsus length: Kruuk *et al.*, 2001; body mass or condition: Merilä, Kruuk & Sheldon, 2001; Braasch, Schauroth & Becker, 2009; Tilgar *et al.*, 2010; wing length: Morrison *et al.*, 2009; immunity: Cichoń & Dubiec, 2005; Moreno *et al.*, 2005). Chicks hatching from large eggs also grow faster. This might also be positively related to juvenile survival as fast growing shortens the nestling period during which the young are vulnerable to nest predation (Remeš & Martin, 2002), although rapid growth also has costs (Metcalfe & Monaghan, 2001). In sum, these pieces of evidence suggest that egg size enhances juvenile survival but the exact magnitude of this effect is unknown since only a few studies have followed offspring up to sexual maturity.

The lack of long-term studies also means that we have nearly no knowledge of egg-size effects on the two other components of fitness that are manifested in adults. Only four out of 283 (1.4%) studies tested for the relationship between egg size and sexual or life-history traits that are related to female fecundity or male mating success (Cunningham & Russell, 2000; Parker, 2002; Krist, 2009; Zanette, Clinchy & Sung, 2009). No study tested for egg-size effects on offspring survival as adults. The lack of studies looking at long-term effects of egg size is unfortunate. As an important component of early offspring environments, egg size is likely to have consequences for offspring reproductive success, given that similar effects are often found for other components of early offspring environments (reviews in Lindström, 1999; Monaghan, 2008), such as natal brood size (Gustafsson, Qvarnström & Sheldon, 1995; Naguib, Nemitz & Gil, 2006; Alonso-Alvarez, Bertrand & Sorci, 2007), maternal nutritional condition (Gorman & Nager, 2004), and prenatal exposure to androgens (Rubolini *et al.*, 2007).

The three fitness components may be negatively correlated due to trade-offs (Schluter, Price & Rowe, 1991; Roff, 2002; Lailvaux, Hall & Brooks, 2010), or positively correlated due to differences among individuals in resource acquisition (van Noordwijk & de Jong, 1986; Reznick, Nunnev & Tessier, 2000; Vorburger, 2005). Therefore, we cannot infer fitness from knowledge of only one component of fitness (Kokko *et al.*, 2003; Hunt *et al.*, 2004; Lailvaux *et al.*, 2010). Despite the theoretical importance of egg-size effects on offspring fecundity (Marshall & Keough, 2008) or survival as an adult, these effects also have been neglected in non-avian animal taxa. They were not mentioned in reviews of fish (Green, 2008) and arthropods (Fox & Czesak, 2000) and only a few studies on offspring fecundity have been carried out in reptiles (Sinervo & Doughty, 1996), and marine invertebrates (e.g. Marshall, Bolton & Keough, 2003; Dias & Marshall, 2010). In contrast to the few studies on animals, in plants the relationship between seed size and subsequent offspring fecundity has been studied quite routinely (e.g. Stanton, 1984; Mazer, 1987; Mazer & Wolfe, 1998).

To conclude, in birds, only one component of offspring fitness—juvenile survival—has been widely studied for its dependence on egg size. Egg-size effects on offspring fecundity and adult survival remain to be tested. Similarly, the relationship between egg size and offspring global fitness, not its components, remains to be established in any animal taxa. Such a study would test something different than studies that looked at selection on egg size (Hõrak, Mänd & Ots, 1997; Garant *et al.*, 2007; Kontiainen *et al.*, 2008). These latter studies tested for the relationship between egg size and the lifetime reproductive success of individuals that laid the eggs, not those that hatched from them. This level was appropriate for their purpose since selection optimizes maternal, not offspring fitness (Marshall & Uller, 2007). By contrast, if we want to parameterize the Smith & Fretwell (1974) or other optimization models we need to know the quantitative relationship between egg size and offspring fitness (Marshall & Keough, 2008; Dias & Marshall, 2010).

Apart from the life-history stage when offspring traits were measured, morphological traits were studied most often (1121/1805 estimates, i.e. 62.1%), followed by offspring survival (28.6%) and growth rate of morphological traits (7.1%). Only a few estimates were made on chick immunity (1.1%), egg fertility (0.5%), chick behaviour/activity (0.3%), adult life-history (0.2%), and sexual traits (0.2%). This skewed distribution somewhat parallels studies of selection (Kingsolver *et al.*, 2001) and avian quantitative genetics (Merilä & Sheldon, 2001). In both of these other fields, morphological traits also were the most commonly studied with a few studies performed on behavioural and physiological traits. The difference is that in these fields, life-history traits were the second most commonly investigated traits while only one study tested their dependence on egg size (Krist, 2009). This neglect of offspring life-history traits in the field of propagule size—offspring fitness is common to other animal taxa (see above) in which generally the same kind of traits as in birds were studied. However, in reptiles the relationship between egg size and offspring locomotor performances have often been studied (e.g. Sinervo, 1990; Olsson, Wapstra & Olofsson, 2002; Warner & Andrews, 2002; Warner & Shine, 2009). This contrasts with birds where there are only two such studies (Anderson & Alisauskas, 2001; Goth & Evans, 2004). Given that increased locomotor performance may reduce predation risk and thus enhance survival (Jayne & Bennett, 1990; Warner & Andrews, 2002) and potentially enhance mating success (Byers, Hebets & Podos, 2010) these traits should also be of interest in avian research. A few studies also tested whether egg size predicts begging intensity (Anderson & Alisauskas, 2001; Gilbert *et al.*, 2006; Rubolini *et al.*, 2006a; Bonisoli-Alquati *et al.*, 2007). This offspring trait should also be of interest since begging stimulates parental provisioning (e.g. Ottosson, Backman &

Smith, 1997) that in turn can enhance or mitigate the initial effect of egg size on offspring quality (Krist & Remeš, 2004).

(5) Manipulative approaches

Egg size may be correlated with embryo genes, egg composition, and parental post-hatching care (Krist & Remeš, 2004). Experimental manipulation of egg size may uncouple most of these correlations and consequently may be the best method to infer causal effect of egg size. However, I did not include experimental studies in this meta-analysis for several reasons.

Egg size may be manipulated in two distinct ways. The indirect one utilizes females' phenotypic plasticity to lay differently sized eggs in different conditions. For example, eggs might be enlarged by exposing females to an experimentally enhanced food supply (Bolton *et al.*, 1992), a low temperature (Fischer *et al.*, 2003), or a poor host quality (Fox, 1997) in the pre-laying period. If the subsequent test of egg-size effects on offspring is done within females (see e.g. Wagner & Williams, 2007), this setting controls for genetic effects similar to within-clutch comparisons, but with an additional property that variance in egg size was increased by experimental conditions. However, this approach does not control for the other two potential confounders, egg composition and parental care. In fact, the correlation between experimentally induced changes in egg size and post-hatching care or egg composition may be even larger than in purely observational studies. It is easy to imagine, for example, that food-supplemented females are in better condition and consequently provide better care to their chicks. Due to this threat, studies that indirectly manipulated egg size were not included in the meta-analysis.

Egg size may also be manipulated directly by yolk (Sinervo, 1990; Sinervo *et al.*, 1992) or albumen (Hill, 1993; Ferrari *et al.*, 2006) removal, physical removal of part of developing follicles which effectively increase the size of those remaining (Sinervo & Licht, 1991*b*), or the manipulation of the female hormonal function involved in follicle growth (Sinervo & Licht, 1991*a*; Williams, 2001). The most direct manipulation is one that manipulates the egg size outside the female after laying. Similar manipulations were first carried out in urchins (Sinervo & McEdward, 1988; but see Marshall & Keough, 2008 for criticism of the experimental approach used in this taxon) and reptiles (Sinervo, 1990; Sinervo *et al.*, 1992). More recently, these techniques have been applied to fish (Morley *et al.*, 1999; Jardine & Litvak, 2003) and poultry (Hill, 1993; Finkler, van Orman & Sotherland, 1998) and only very recently to wild birds (Ferrari *et al.*, 2006; Bonisoli-Alquati *et al.*, 2007, 2008). These studies generally find positive relationships between egg size and offspring quality. These approaches, providing the most causal test of egg-size effects, can only suffer if the parents adjust their post-hatching care according to the state of the hatchlings (Krist & Remeš, 2004).

These manipulative studies are difficult to pool together with studies that utilized natural variation in egg size. For example, if we find the correlation r = 0.2 between egg volume and chick mass, this means that a change of 1 S.D. in egg volume causes a 0.2 S.D. change in mass. However, if we find that the removal of 1 S.D. of egg volume content causes a difference of 0.2 S.D. in body mass, how should this be interpreted? Is this effect equal to the former one? This question is difficult to answer, because egg content is not homogenous and in practise only albumen or yolk is usually removed while the other part is left intact. However, for the developing embryo it may be more relevant what proportion of albumen was removed, or how the ratio of albumen to yolk content changed (see Ferrari *et al.*, 2006), not the volume of egg removed. In other words, it is unclear how to measure the strength of the experimental treatment. Moreover, except for the whole size of the egg, this strength can only be estimated for each particular egg since the proportion of yolk and albumen cannot be determined for any individual egg if the aim is to leave the embryo alive. These difficulties do not mean that it is impossible to compare manipulative and observational studies but with only three manipulative studies available (Ferrari *et al.*, 2006; Bonisoli-Alquati *et al.*, 2007, 2008) this would not be very meaningful.

In contrast to the manipulation of already laid eggs, manipulation of developing eggs inside females is less direct since this can affect female condition, rearing abilities, and also egg composition. These effects were argued to be negligible in the case of the recently developed technique of the application of tamoxifen that functions as an antiestrogen (Wagner & Williams, 2007). However, such effects are unlikely to be fully absent. At the very least, females that laid miniaturized eggs did not pay the costs of laying large ones (see Williams, 2005; Nager, 2006) and therefore might be in better condition after laying. At worst, manipulation of hormonal metabolism might affect the deposition of hormones into eggs thus creating a strong confounding correlation between egg size and egg composition.

(6) Egg composition and effect size in other taxa

Egg composition came to the attention of avian ecologists after the publication of Schwabl's (1993, 1996) studies reporting that yolk steroids affected chick quality. Many subsequent studies found effects of the concentration of yolk androgens (reviews in Gil, 2003, 2008; Groothuis *et al.*, 2005) and carotenoids (e.g. Saino *et al.*, 2003; McGraw, Adkins-Regan & Parker, 2005; but see Remeš *et al.*, 2007) on chick performance. Given these new findings, an intriguing question arises: is egg size more or less important for chick quality than egg composition? This question has no answer yet. Results of some studies indirectly suggested that egg composition might be more important (Nager *et al.*, 2000; Reed *et al.*, 2009) while others suggested the opposite pattern (Rubolini *et al.*, 2006*a*) or found an interactive effect of egg size and composition (Romano *et al.*, 2008). This issue can be resolved by meta-analysis of composition effects and their comparison with results of the present study. Such comparison would be a necessary step to unravel by which of

these pathways females may more effectively adjust offspring phenotype.

Similarly, it would be of great interest to elucidate whether egg-size effects on offspring quality are the same, weaker, or larger in other oviparous taxa compared to birds. For example, I reviewed only a few studies on egg-size effects in fish, yet several effect sizes were larger than the largest effect size found for the same condition in birds. Einum & Fleming (2000) found in Atlantic salmon (*Salmo salar*) a correlation between egg size and juvenile body mass ($r = 0.90$ and 0.66 at juvenile age of 28 and 107 days, respectively). Similarly, juvenile survival at age 20 days was very highly correlated ($r = 0.87$ and 0.88 at high and low food levels, respectively) with egg size in brook trout, *Salvelinus fontinalis* (Hutchings, 1991). These examples suggest that egg size may be more important for offspring fitness in fish, a taxa with less-developed post-hatching parental care compared to birds.

V. CONCLUSIONS

(1) This meta-analysis is the first quantitative assessment of the relationship between propagule size and offspring quality done in any animal or plant taxon. Egg size was positively related to nearly all studied offspring traits during all stages in their life cycle. However, this research effort was severely biased to offspring morphological traits and those in the early stages in their life cycle. Only a few studies were performed on offspring behavioural, physiological, life-history, and sexual traits. Few followed the offspring until the post-fledging stage, and nearly none until sexual maturity. Consequently, evidence that juvenile survival is positively related to egg size is robust but relationships between egg size and adult survival, fecundity or global fitness of offspring are unknown at present. This remains a major challenge for further work.

(2) A major question is whether the positive relationships between egg size and offspring quality are driven by a causal effect of egg size or by some other variable that is correlated with egg size. Independent of egg size, offspring quality may be affected by parental post-hatching care, direct action of genes, and egg composition. This meta-analysis found no difference in effect size between observational and cross-fostering studies. This suggests little confusion of the effect size by parental or territory quality. Largely untested is the possibility that effect size is confounded by parental adjustment of post-hatching care, although some indirect evidence suggests that this should not be a problem. Direct genetic effects do not confound relationships at the within-clutch level which were also found to be significant, although of lower magnitude. A few studies that manipulated egg size directly in wild birds showed that egg size is related to offspring quality regardless of egg composition. In sum, these pieces of evidence suggest that relationships found in this meta-analysis are driven primarily by a causal effect of egg size. However, more studies controlling for potentially confounding variables are needed to establish firmly the causality of these relationships.

(3) The relationship between egg size and offspring traits found within broods was smaller than that found between broods. This can be caused by (*a*) less intraclutch egg-size variation, (*b*) parental within-brood compensation of a poor start by the young from small eggs through increased food-provisioning, or (*c*) a correlation of egg size with genetic effects in between-clutch settings. Further resolution of these hypotheses is impossible at present due to the scarcity of studies testing for direct genetic effects, parental provisioning, and those that decreased egg size variation at the between-clutch level or increased it at the intraclutch level.

(4) The relationship between egg size and offspring traits was larger if tested in captivity than in the wild, which can be explained by the better control of confounding variables in laboratory conditions. Larger effect size in benign laboratory conditions is opposite of what is assumed by theoretical models and usually found in observational studies in birds and experimental studies in other taxa. More studies are needed that manipulate the offspring environment. For example, studies that involve food supplementation or brood-size manipulation, and compare egg-size effects in benign and harsh conditions.

(5) At present, the effects of egg composition on chick quality are often studied. It would be of great interest to elucidate whether the offspring phenotype may be more effectively manipulated by egg size or egg composition. This can be achieved by meta-analysis in the field of egg composition and comparison of the found effect sizes with those in the present study. Similarly, meta-analyses of relationships between egg size and offspring quality in other oviparous taxa and their comparison with the present study can add to our understanding of life-history diversity among animal and plant taxa.

VI. ACKNOWLEDGEMENTS

VII. REFERENCES

ADAMOU, A. E., KOUIDRI, M., CHABI, Y., SKWARSKA, J. & BAŃBURA, J. (2009). Egg size variation and breeding characteristics of the black-winged stilt *Himantopus himantopus* in a Saharan oasis. *Acta Ornithologica* **44**, 1–7.

ADAMS, D. C. (2008). Phylogenetic meta-analysis. *Evolution* **62**, 567–572.

AGOSTA, S. J. (2008). Fitness consequences of host use in the field: temporal variation in performance and a life history trade-off in the moth *Rothschildia lebeau* (Saturniidae). *Oecologia* **157**, 69–82.

ALLEN, R. M., BUCKLEY, Y. M. & MARSHALL, D. J. (2008). Offspring size plasticity in response to intraspecific competition: an adaptive maternal effect across life-history stages. *American Naturalist* **171**, 225–237.

ALONSO-ALVAREZ, C., BERTRAND, S. & SORCI, G. (2007). Sex-specific transgenerational effects of early developmental conditions in a passerine. *Biological Journal of the Linnean Society* **91**, 469–474.

AMUNDSEN, T. & STOKLAND, J. N. (1990). Egg size and parental quality influence nestling growth in the shag. *Auk* **107**, 410–413.

AMUNDSEN, T., LORENTSEN, S. H. & TVERAA, T. (1996). Effects of egg size and parental quality on early nestling growth: an experiment with the Antarctic petrel. *Journal of Animal Ecology* **65**, 545–555.

ANDERSON, V. R. & ALISAUSKAS, R. T. (2001). Egg size, body size, locomotion, and feeding performance in captive king eider ducklings. *Condor* **103**, 195–199.

ANDERSON, V. R. & ALISAUSKAS, R. T. (2002). Composition and growth of king eider ducklings in relation to egg size. *Auk* **119**, 62–70.

ANDERSON, D. J., REEVE, J. & BIRD, D. M. (1997). Sexually dimorphic eggs, nestling growth and sibling competition in American kestrels *Falco sparverius*. *Functional Ecology* **11**, 331–335.

ANKNEY, C. D. (1980). Egg weight, survival, and growth of lesser snow goose goslings. *Journal of Wildlife Management* **44**, 174–182.

ARDIA, D. R. & RICE, E. B. (2006). Variation in heritability of immune function in the tree swallow. *Evolutionary Ecology* **20**, 491–500.

ARNOLD, J. M., HATCH, J. J. & NISBET, I. C. T. (2006). Effects of egg size, parental quality and hatch-date on growth and survival of common tern *Sterna hirundo* chicks. *Ibis* **148**, 98–105.

ARNQVIST, G. & WOOSTER, D. (1995). Meta-analysis: synthesizing research findings in ecology and evolution. *Trends In Ecology & Evolution* **10**, 236–240.

BADYAEV, A. V. & ULLER, T. (2009). Parental effects in ecology and evolution: mechanisms, processes and implications. *Philosophical Transactions of the Royal Society B-Biological Sciences* **364**, 1169–1177.

BADZINSKI, S. S., ANKNEY, C. D., LEAFLOOR, J. O. & ABRAHAM, K. F. (2002). Egg size as a predictor of nutrient composition of eggs and neonates of Canada geese (*Branta canadensis interior*) and lesser snow geese (*Chen caerulescens caerulescens*). *Canadian Journal of Zoology* **80**, 333–341.

BASHEY, F. (2006). Cross-generational environmental effects and the evolution of offspring size in the Trinidadian guppy *Poecilia reticulata*. *Evolution* **60**, 348–361.

BERKELEY, S. A., CHAPMAN, C. & SOGARD, S. M. (2004). Maternal age as a determinant of larval growth and survival in a marine fish, *Sebastes melanops*. *Ecology* **85**, 1258–1264.

BERNARDO, J. (1996). The particular maternal effect of propagule size, especially egg size: patterns, models, quality of evidence and interpretations. *American Zoologist* **36**, 216–236.

BIRKHEAD, T. R. & NETTLESHIP, D. N. (1982). The adaptive significance of egg size and laying date in thick-billed murres *Uria lomvia*. *Ecology* **63**, 300–306.

BITTON, P. P., DAWSON, R. D. & O'BRIEN, E. L. (2006). Influence of intraclutch egg-mass variation and hatching asynchrony on relative offspring performance within broods of an altricial bird. *Canadian Journal of Zoology* **84**, 1721–1726.

BLOMQVIST, D., JOHANSSON, O. C. & GÖTMARK, F. (1997). Parental quality and egg size affect chick survival in a precocial bird, the lapwing *Vanellus vanellus*. *Oecologia* **110**, 18–24.

BLOUNT, J. D., HOUSTON, D. C. & MØLLER, A. P. (2000). Why egg yolk is yellow. *Trends in Ecology & Evolution* **15**, 47–49.

BOGDANOVA, M. I. & NAGER, R. G. (2008). Sex-specific costs of hatching last: an experimental study on herring gulls (*Larus argentatus*). *Behavioral Ecology and Sociobiology* **62**, 1533–1541.

BOGDANOVA, M. I., NAGER, R. G. & MONAGHAN, P. (2006). Does parental age affect offspring performance through differences in egg quality? *Functional Ecology* **20**, 132–141.

BOLTON, M. (1991). Determinants of chick survival in the lesser black-backed gull: relative contributions of egg size and parental quality. *Journal of Animal Ecology* **60**, 949–960.

BOLTON, M., HOUSTON, D. & MONAGHAN, P. (1992). Nutritional constraints on egg formation in the lesser black-backed gull: an experimental study. *Journal of Animal Ecology* **61**, 521–532.

BONDURIANSKY, R. & HEAD, M. (2007). Maternal and paternal condition effects on offspring phenotype in *Telostylinus angusticollis* (Diptera: Neriidae). *Journal of Evolutionary Biology* **20**, 2379–2388.

BONISOLI-ALQUATI, A., MARTINELLI, R., RUBOLINI, D. & SAINO, N. (2008). Sex-specific effects of albumen removal and nest environment manipulation on barn swallow nestlings. *Ecology* **89**, 2315–2324.

BONISOLI-ALQUATI, A., RUBOLINI, D., ROMANO, M., BONCORAGLIO, G., FASOLA, M. & SAINO, N. (2007). Effects of egg albumen removal on yellow-legged gull chick phenotype. *Functional Ecology* **21**, 310–316.

BORENSTEIN, M., HEDGES, L. V., HIGGINS, J. P. T. & ROTHSTEIN, H. R. (2009). *Introduction to Meta-analysis*. Chichester: John Wiley & Sons, Ltd.

BOULTON, R. L. & POWLESLAND, R. G. (2008). Variation in egg size and nest survival with female age in the South Island robin *Petroica australis*. *Ibis* **150**, 824–828.

BOWEN, W. D. (2009). Maternal effects on offspring size and development in pinnipeds. In *Maternal Effects in Mammals* (eds. D. MAESTRIPIERI & J. M. MATEO), pp. 104–132. Chicago: The University of Chicago Press.

BRAASCH, A., SCHAUROTH, C. & BECKER, P. H. (2009). Post-fledging body mass as a determinant of subadult survival in common terns *Sterna hirundo*. *Journal of Ornithology* **150**, 401–407.

BROCKELMAN, W. Y. (1975). Competition, fitness of offspring, and optimal clutch size. *American Naturalist* **109**, 677–699.

BUDDEN, A. E. & BEISSINGER, S. R. (2005). Egg mass in an asynchronously hatching parrot: does variation offset constraints imposed by laying order? *Oecologia* **144**, 318–326.

BYERS, J., HEBETS, E. & PODOS, J. (2010). Female mate choice based upon male motor performance. *Animal Behaviour* **79**, 771–778.

CASSEY, P., EWEN, J. G., BLACKBURN, T. M. & MØLLER, A. P. (2004). A survey of publication bias within evolutionary ecology. *Proceedings of the Royal Society B-Biological Sciences (Supplement)* **271**, S451–S454.

CHARMANTIER, A. & GARANT, D. (2005). Environmental quality and evolutionary potential: lessons from wild populations. *Proceedings of the Royal Society B-Biological Sciences* **272**, 1415–1425.

CHARMANTIER, A., KRUUK, L. E. B., BLONDEL, J. & LAMBRECHTS, M. M. (2004). Testing for microevolution in body size in three blue tit populations. *Journal of Evolutionary Biology* **17**, 732–743.

CHARMANTIER, A., PERRINS, C., McCLEERY, R. H. & SHELDON, B. C. (2006). Age-dependent genetic variance in a life-history trait in the mute swan. *Proceedings of the Royal Society B-Biological Sciences* **273**, 225–232.

CHRISTIANS, J. K. (2002). Avian egg size: variation within species and inflexibility within individuals. *Biological Reviews* **77**, 1–26.

CICHOŃ, M. & DUBIEC, A. (2005). Cell-mediated immunity predicts the probability of local recruitment in nestling blue tits. *Journal of Evolutionary Biology* **18**, 962–966.

CLIFFORD, L. D. & ANDERSON, D. J. (2002). Clutch size variation in the Nazca booby: a test of the egg quality hypothesis. *Behavioral Ecology* **13**, 274–279.

CLUTTON-BROCK, T. H. (1991). *The Evolution of Parental Care*. New Jersey: Princeton University Press.

COHEN, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. New Jersey: Lawrence Erlbaum Associates, Inc.

CRAMP, S. & PERRINS, C. M. (1977–1994). *The Birds of the Western Palearctic*. Oxford: Oxford University Press.

CREAN, A. J. & MARSHALL, D. J. (2009). Coping with environmental uncertainty: dynamic bet hedging as a maternal effect. *Philosophical Transactions of the Royal Society B-Biological Sciences* **364**, 1087–1096.

CUNNINGHAM, E. J. A. & RUSSELL, A. F. (2000). Egg investment is influenced by male attractiveness in the mallard. *Nature* **404**, 74–77.

D'ALBA, L. & TORRES, R. (2007). Seasonal egg-mass variation and laying sequence in a bird with facultative brood reduction. *Auk* **124**, 643–652.

DAVIS, J. W. F. (1975). Age, egg-size and breeding success in herring gull *Larus argentatus*. *Ibis* **117**, 460–473.

DAWSON, R. D. & CLARK, R. G. (2000). Effects of hatching date and egg size on growth, recruitment, and adult size of lesser scaup. *Condor* **102**, 930–935.

DEL HOYO, J., ELLIOTT, A. & SAGARTAL, J. (1992–2006). *Handbook of the Birds of the World*. Barcelona: Lynx.

DE NEVE, L., SOLER, J. J., PÉREZ-CONTRERAS, T. & SOLER, M. (2004). Genetic, environmental and maternal effects on magpie nestling-fitness traits under different nutritional conditions: a new experimental approach. *Evolutionary Ecology Research* **6**, 415–431.

DEEMING, D. C. (2002). *Avian Incubation*. Oxford: Oxford University Press.

DIAS, G. M. & MARSHALL, D. J. (2010). Does the relationship between offspring size and performance change across the life-history? *Oikos* **119**, 154–162.

DIBATTISTA, J. D., FELDHEIM, K. A., GARANT, D., GRUBER, S. H. & HENDRY, A. P. (2009). Evolutionary potential of a large marine vertebrate: quantitative genetic parameters in a wild population. *Evolution* **63**, 1051–1067.

DIVINO, J. N. & TONN, W. M. (2008). Importance of nest and paternal characteristics for hatching success in fathead minnow. *Copeia*, 920–930.

DONOHUE, K. & SCHMITT, J. (1998). Maternal environmental effects in plants. In *Maternal Effects as Adaptations* (eds. T. A. MOUSSEAU & C. W. FOX), pp. 137–158. Oxford: Oxford Unversity Press.

DUVAL, S. & TWEEDIE, R. (2000). A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association* **95**, 89–98.

DZIMINSKI, M. A. & ROBERTS, J. D. (2006). Fitness consequences of variable maternal provisioning in quacking frogs (*Crinia georgiana*). *Journal of Evolutionary Biology* **19**, 144–155.

EINUM, S. & FLEMING, I. A. (1999). Maternal effects of egg size in brown trout (*Salmo trutta*): norms of reaction to environmental quality. *Proceedings of the Royal Society of London B-Biological Sciences* **266**, 2095–2100.

EINUM, S. & FLEMING, I. A. (2000). Highly fecund mothers sacrifice offspring survival to maximize fitness. *Nature* **405**, 565–567.

ENGQVIST, L. (2005). The mistreatment of covariate interaction terms in linear model analyses of behavioural and evolutionary ecology studies. *Animal Behaviour* **70**, 967–971.

EVANS, D. M., REDPATH, S. M., EVANS, S. A., ELSTON, D. A. & DENNIS, P. (2005). Livestock grazing affects the egg size of an insectivorous passerine. *Biology Letters* **1**, 322–325.

FARGALLO, J. A., POLO, V., DE NEVE, L., MARTIN, J., DAVILA, J. A. & SOLER, M. (2006). Hatching order and size-dependent mortality in relation to brood sex ratio composition in chinstrap penguins. *Behavioral Ecology* **17**, 772–778.

FEARE, C. J. (1976). Breeding of sooty tern *Sterna fuscata* in Seychelles and effects of experimental removal of its eggs. *Journal of Zoology* **179**, 317–360.

FERNÁNDEZ, G. J. & REBOREDA, J. C. (2008). Between and within clutch variation of egg size in greater rheas. *Wilson Journal of Ornithology* **120**, 674–682.

FERRARI, R. P., MARTINELLI, R. & SAINO, N. (2006). Differential effects of egg albumen content on barn swallow nestlings in relation to hatch order. *Journal of Evolutionary Biology* **19**, 981–993.

FICETOLA, G. F. & DE BERNARDI, F. (2009). Offspring size and survival in the frog *Rana latastei*: from among-population to within-clutch variation. *Biological Journal of the Linnean Society* **97**, 845–853.

FINKLER, M. S., VAN ORMAN, J. B. & SOTHERLAND, P. R. (1998). Experimental manipulation of egg quality in chickens: influence of albumen and yolk on the size and body composition of near-term embryos in a precocial bird. *Journal of Comparative Physiology B-Biochemical, Systemic, and Environmental Physiology* **168**, 17–24.

FISCHER, K., BOT, A. N. M., BRAKEFIELD, P. M. & ZWAAN, B. J. (2003). Fitness consequences of temperature-mediated egg size plasticity in a butterfly. *Functional Ecology* **17**, 803–810.

FORBES, S. & WIEBE, M. (2010). Egg size and asymmetric sibling rivalry in red-winged blackbirds. *Oecologia* **163**, 361–372.

FORSTMEIER, W., COLTMAN, D. W. & BIRKHEAD, T. R. (2004). Maternal effects influence the sexual behavior of sons and daughters in the zebra finch. *Evolution* **58**, 2574–2583.

FOX, C. W. (1993). Maternal and genetic influences on egg size and larval prformance in a seed beetle (*Callosobruchus maculatus*): multigenerational transmission of a maternal effect? *Heredity* **73**, 509–517.

FOX, C. W. (1997). Egg-size manipulations in the seed beetle *Stator limbatus*: consequences for progeny growth. *Canadian Journal of Zoology* **75**, 1465–1473.

FOX, C. W. (2000). Natural selection on seed-beetle egg size in nature and the laboratory: variation among environments. *Ecology* **81**, 3029–3035.

FOX, C. W., BUSH, M. L. & WALLIN, W. G. (2003). Maternal age affects offspring lifespan of the seed beetle, *Callosobruchus maculatus*. *Functional Ecology* **17**, 811–820.

FOX, C. W. & CZESAK, M. E. (2000). Evolutionary ecology of progeny size in arthropods. *Annual Review of Entomology* **45**, 341–369.

FOX, C. W., CZESAK, M. E. & WALLIN, W. G. (2004). Complex genetic architecture of population differences in adult lifespan of a beetle: nonadditive inheritance, gender differences, body size and a large maternal effect. *Journal of Evolutionary Biology* **17**, 1007–1017.

GALLIZZI, K. & RICHNER, H. (2008). A parasite-induced maternal effect can reduce survival times of fleas feeding on great tit nestlings. *Oikos* **117**, 1209–1217.

GARANT, D., KRUUK, L. E. B., MCCLEERY, R. H. & SHELDON, B. C. (2007). The effects of environmental heterogeneity on multivariate selection on reproductive traits in female great tits. *Evolution* **61**, 1546–1559.

GATES, S. (2002). Review of methodology of quantitative reviews using meta-analysis in ecology. *Journal of Animal Ecology* **71**, 547–557.

GEBHARDT-HENRICH, S. G. & VAN NOORDWIJK, A. J. (1991). Nestling growth in the great tit. 1. Heritability estimates under different environmental conditions. *Journal of Evolutionary Biology* **4**, 341–362.

GEBHARDT-HENRICH, S. G. & VAN NOORDWIJK, A. J. (1994). The genetic ecology of nestling growth in the great tit: environmental influences on the expression of genetic variances during growth. *Functional Ecology* **8**, 469–476.

GERITZ, S. A. H. (1995). Evolutionarily stable seed polymorphism and small-scale spatial variation in seedling sensity. *American Naturalist* **146**, 685–707.

GIL, D. (2003). Golden eggs: maternal manipulation of offspring phenotype by egg androgen in birds. *Ardeola* **50**, 281–294.

GIL, D. (2008). Hormones in avian eggs: physiology, ecology, and behavior. *Advances in the Study of Behavior* **38**, 337–397.

GILBERT, L., WILLIAMSON, K. A., HAZON, N. & GRAVES, J. A. (2006). Maternal effects due to male attractiveness affect offspring development in the zebra finch. *Proceedings of the Royal Society B-Biological Sciences* **273**, 1765–1771.

GORMAN, H. E. & NAGER, R. G. (2004). Prenatal developmental conditions have long-term effects on offspring fecundity. *Proceedings of the Royal Society of London Series B-Biological Sciences* **271**, 1923–1928.

GOTH, A. & EVANS, C. S. (2004). Egg size predicts motor performance and postnatal weight gain of Australian brush-turkey (*Alectura lathami*) hatchlings. *Canadian Journal of Zoology* **82**, 972–979.

GRAFEN, A. & HAILS, R. (2002). *Modern Statistics for the Life Sciences*. Oxford: Oxford University Press.

GRAHAM, M. H. (2003). Confronting multicollinearity in ecological multiple regression. *Ecology* **84**, 2809–2815.

GREEN, B. S. (2008). Maternal effects in fish populations. *Advances in Marine Biology* **54**, 1–105.

GRIFFITH, S. C., OWENS, I. F. & BURKE, T. (1999). Environmental determination of a sexually selected trait. *Nature* **400**, 358–360.

GRIM, T. (2005). Host recognition of brood parasites: implications for methodology in studies of enemy recognition. *Auk* **122**, 530–543.

GROOTHUIS, T. G. G., MÜLLER, W., VON ENGELHARDT, N., CARERE, C. & EISING, C. (2005). Maternal hormones as a tool to adjust offspring phenotype in avian species. *Neuroscience and Biobehavioral Reviews* **29**, 329–352.

GUSTAFSSON, L., QVARNSTRÖM, A. & SHELDON, B. C. (1995). Trade-offs between life-history traits and a secondary sexual character in male collared flycatchers. *Nature* **375**, 311–313.

HEATH, D. D. & BLOUW, D. M. (1998). Are maternal effects in fish adaptive or merely physiological side effects? In *Maternal Effects as Adaptations* (eds. T. A. MOUSSEAU & C. W. FOX), pp. 178–201. Oxford: Oxford University Press.

HENDRY, A. P., DAY, T. & COOPER, A. B. (2001). Optimal size and number of propagules: allowance for discrete stages and effects of maternal size on reproductive output and offspring fitness. *American Naturalist* **157**, 387–407.

HERNÁNDEZ, M., GONZÁLEZ, L. M., ORIA, J., SÁNCHEZ, R. & ARROYO, B. (2008). Influence of contamination by organochlorine pesticides and polychlorinated biphenyls on the breeding of the Spanish imperial eagle (*Aquila adalberti*). *Environmental Toxicology and Chemistry* **27**, 433–441.

HIGGINS, P. J. & PETER, J. M. (1990–2006). *Handbook of Australian, New Zealand and Antarctic Birds*. Melbourne: Oxford University Press.

HILL, W. L. (1993). Importance of prenatal nutrition to the development of a precocial chick. *Developmental Psychobiology* **26**, 237–249.

HIPFNER, J. M. & GASTON, A. J. (1999). The relationship between egg size and posthatching development in the thick-billed murre. *Ecology* **80**, 1289–1297.

HOCHACHKA, W. M. (1993). Repeatable reproduction in song sparrows. *Auk* **110**, 603–613.

HOŘÁK, D. & ALBRECHT, T. (2007). Using net sacks to examine the relationship between egg size and young size in common pochards. *Journal of Field Ornithology* **78**, 334–339.

HÕRAK, P., MÄND, R. & OTS, I. (1997). Identifying targets of selection: a multivariate analysis of reproductive traits in the great tit. *Oikos* **78**, 592–600.

HOWE, H. F. (1976). Egg size, hatching asynchrony, sex and brood reduction in the common grackle. 1195–1207.

HOYT, D. F. (1979). Practical methods of estimating volume and fresh weight of bird eggs. *Auk* **96**, 73–77.

HUNT, J., BUSSIÈRE, L. F., JENNIONS, M. D. & BROOKS, R. (2004). What is genetic quality? *Trends in Ecology & Evolution* **19**, 329–333.

HUNT, J. & SIMMONS, L. W. (2002). The genetics of maternal care: direct and indirect genetic effects on phenotype in the dung beetle *Onthophagus taurus*. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 6828–6832.

HUNTER, J. E. & SCHMIDT, F. L. (2004). *Methods of Meta-analysis: Correcting Error and Bias in Research Findings*. Thousand Oaks: Sage Publications, Inc.

HUTCHINGS, J. A. (1991). Fitness consequences of variation in egg size and food abundance in brook trout *Salvelinus fontinalis*. *Evolution* **45**, 1162–1168.

ISAKSSON, C., ULLER, T. & ANDERSSON, S. (2006). Parental effects on carotenoid-based plumage coloration in nestling great tits, *Parus major*. *Behavioral Ecology and Sociobiology* **60**, 556–562.

JARDINE, D. & LITVAK, M. K. (2003). Direct yolk sac volume manipulation of zebrafish embryos and the relationship between offspring size and yolk sac volume. *Journal of Fish Biology* **63**, 388–397.

JÄRVINEN, A. (1994). Global warming and egg size of birds. *Ecography* **17**, 108–110.

JÄRVINEN, A. & YLIMAUNU, J. (1984). Significance of egg size on the growth of nestling pied flycatchers *Ficedula hypoleuca*. *Annales Zoologici Fennici* **21**, 213–216.

JAYNE, B. C. & BENNETT, A. F. (1990). Selection on locomotor performance capacity in a natural population of garter snakes. *Evolution* **44**, 1204–1229.

JENNIONS, M. D. & MØLLER, A. P. (2002). Relationships fade with time: a meta-analysis of temporal trends in publication in ecology and evolution. *Proceedings of the Royal Society of London Series B-Biological Sciences* **269**, 43–48.

KERRIGAN, B. A. (1997). Variability in larval development of the tropical reef fish *Pomacentrus amboinensis* (Pomacentridae): the parental legacy. *Marine Biology* **127**, 395–402.

KILPIMAA, J., VAN DE CASTEELE, T., JOKINEN, I., MAPPES, J. & ALATALO, R. V. (2005). Genetic and environmental variation in antibody and T-cell mediated responses in the great tit. *Evolution* **59**, 2483–2489.

KINGSOLVER, J. G., HOEKSTRA, H. E., HOEKSTRA, J. M., BERRIGAN, D., VIGNIERI, S. N., HILL, C. E., HOANG, A., GIBERT, P. & BEERLI, P. (2001). The strength of phenotypic selection in natural populations. *American Naturalist* **157**, 245–261.

KOKKO, H., BROOKS, R., JENNIONS, M. D. & MORLEY, J. (2003). The evolution of mate choice and mating biases. *Proceedings of the Royal Society of London Series B-Biological Sciences* **270**, 653–664.

KÖLLIKER, M., BRODIE III, E. D. & MOORE, A. J. (2005). The coadaptation of parental supply and offspring demand. *American Naturalist* **166**, 506–516.

KONTIAINEN, P., BROMMER, J. E., KARELL, P. & PIETIÄINEN, H. (2008). Heritability, plasticity and canalization of Ural owl egg size in a cyclic environment. *Journal of Evolutionary Biology* **21**, 88–96.

KOOPS, M. A., HUTCHINGS, J. A. & ADAMS, B. K. (2003). Environmental predictability and the cost of imperfect information: influences on offspring size variability. *Evolutionary Ecology Research* **5**, 29–42.

KOVAŘÍK, P., PAVEL, V. & CHUTNÝ, B. (2009). Incubation behaviour of the meadow pipit (*Anthus pratensis*) in an alpine ecosystem of Central Europe. *Journal of Ornithology* **150**, 549–556.

KOZLOWSKI, C. P. & RICKLEFS, R. E. (2010). Egg size and yolk steroids vary across the laying order in cockatiel clutches: a strategy for reinforcing brood hierarchies? *General and Comparative Endocrinology* **168**, 460–465.

KREBS, E. A. (2002). Sibling competition and parental control: patterns of begging in parrots. In *The Evolution of Begging: Competition, Cooperation & Communication* (eds. J. WRIGHT & M. L. LEONARD), pp. 319–336. Dordrecht: Kluwer Academic Publishers.

KRIST, M. (2004). Importance of competition for food and nest-sites in aggressive behaviour of collared flycatcher *Ficedula albicollis*. *Bird Study* **51**, 41–47.

KRIST, M. (2009). Short- and long-term effects of egg size and feeding frequency on offspring quality in the collared flycatcher (*Ficedula albicollis*). *Journal of Animal Ecology* **78**, 907–918.

KRIST, M., NÁDVORNÍK, P., UVÍROVÁ, L. & BUREŠ, S. (2005). Paternity covaries with laying and hatching order in the collared flycatcher *Ficedula albicollis*. *Behavioral Ecology and Sociobiology* **59**, 6–11.

KRIST, M. & REMEŠ, V. (2004). Maternal effects and offspring performance: in search of the best method. *Oikos* **106**, 422–426.

KRIST, M., REMEŠ, V., UVÍROVÁ, L., NÁDVORNÍK, P. & BUREŠ, S. (2004). Egg size and offspring performance in the collared flycatcher (*Ficedula albicollis*): a within-clutch approach. *Oecologia* **140**, 52–60.

KROLL, A. J. & HAUFLER, J. B. (2007). Evaluating habitat quality for the dusky flycatcher. *Journal of Wildlife Management* **71**, 14–22.

KRUUK, L. E. B. (2004). Estimating genetic parameters in natural populations using the 'animal model'. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **359**, 873–890.

KRUUK, L. E. B. & HADFIELD, J. D. (2007). How to separate genetic and environmental causes of similarity between relatives. *Journal of Evolutionary Biology* **20**, 1890–1903.

KRUUK, L. E. B., MERILÄ, J. & SHELDON, B. C. (2001). Phenotypic selection on a heritable size trait revisited. *American Naturalist* **158**, 557–571.

KUDO, S. (2001). Intraclutch egg-size variation in acanthosomatid bugs: adaptive allocation of maternal investment? *Oikos* **92**, 208–214.

LAILVAUX, S. P., HALL, M. D. & BROOKS, R. C. (2010). Performance is no proxy for genetic quality: trade-offs between locomotion, attractiveness, and life history in crickets. *Ecology* **91**, 1530–1537.

LAJEUNESSE, M. J. (2009). Meta-analysis and the comparative phylogenetic method. *American Naturalist* **174**, 369–381.

LANDE, R. & PRICE, T. (1989). Genetic correlations and maternal effect coefficients obtained from offspring-parent regression. *Genetics* **122**, 915–922.

LEBLANC, Y. (1987). Egg mass, position in the laying sequence, and brood size in relation to Canada goose reproductive success. *Wilson Bulletin.* **99**, 663–672.

LESSELLS, C. M. (1986). Brood size in Canada geese: a manipulation experiment. *Journal of Animal Ecology* **55**, 669–689.

LINDHOLM, A. K., HUNT, J. & BROOKS, R. (2006). Where do all the maternal effects go? Variation in offspring body size through ontogeny in the live-bearing fish *Poecilia parae*. *Biology Letters* **2**, 586–589.

LINDSTRÖM, J. (1999). Early development and fitness in birds and mammals. *Trends in Ecology & Evolution* **14**, 343–348.

LIPSEY, M. W. & WILSON, D. B. (2001). *Practical Meta-analysis.* Thousand Oaks: Sage Publications, Inc.

LISLEVAND, T., BYRKJEDAL, I., BORGE, T. & SÆTRE, G. P. (2005). Egg size in relation to sex of embryo, brood sex ratios and laying sequence in northern lapwings (*Vanellus vanellus*). *Journal of Zoology* **267**, 81–87.

LLOYD, C. S. (1979). Factors affecting breeding of razorbills *Alca torda* on Skokholm. *Ibis* **121**, 165–176.

LOCK, J. E., SMISETH, P. T., MOORE, P. J. & MOORE, A. J. (2007). Coadaptation of prenatal and postnatal maternal effects. *American Naturalist* **170**, 709–718.

LOUZAO, M., IGUAL, J. M., GENOVART, M., FORERO, M. G., HOBSON, K. A. & ORO, D. (2008). Spatial variation in egg size of a top predator: interplay of body size and environmental factors? *Acta Oecologica* **34**, 186–193.

LUNDBERG, C. A. & VÄISÄNEN, R. A. (1979). Selective correlation of egg size with chick mortality in the black-headed gull (*Larus ridibundus*). *Condor* **81**, 146–156.

LYNCH, M. & WALSH, B. (1998). *Genetics and Analysis of Quantitative Traits.* Sunderland: Sinauer Associates, Inc.

MADDOX, J. D. & WEATHERHEAD, P. J. (2008). Egg size variation in birds with asynchronous hatching: is bigger really better? *American Naturalist* **171**, 358–365.

MAGRATH, M. J. L., VEDDER, O., VAN DER VELDE, M. & KOMDEUR, J. (2009). Maternal effects contribute to the superior performance of extra-pair offspring. *Current Biology* **19**, 792–797.

MAGRATH, R. D. (1992). The effect of egg mass on the growth and survival of blackbirds: a field experiment. *Journal of Zoology* **227**, 639–653.

MÄND, R. (1985). On the relationship of the egg size with the growth rate and survival of the young in some Laridae species [in Russian]. *Proceedings of the Academy of Sciences of the Estonian SSR: Biology* **34**, 34–44.

MARSHALL, D. J., BOLTON, T. F. & KEOUGH, M. J. (2003). Offspring size affects the post-metamorphic performance of a colonial marine invertebrate. *Ecology* **84**, 3131–3137.

MARSHALL, D. J. & KEOUGH, M. J. (2008). The evolutionary ecology of offspring size in marine invertebrates. *Advances in Marine Biology* **53**, 1–60.

MARSHALL, D. J. & ULLER, T. (2007). When is a maternal effect adaptive? *Oikos* **116**, 1957–1963.

MARTIN, T. E., BASSAR, R. D., BASSAR, S. K., FONTAINE, J. J., LLOYD, P., MATHEWSON, H. A., NIKLISON, A. M. & CHALFOUN, A. (2006). Life-history and ecological correlates of geographic variation in egg and clutch mass among passerine species. *Evolution* **60**, 390–398.

MATEO, J. M. (2009). Maternal influences on development, social relationships, and survival behaviors. In *Maternal Effects in Mammals* (eds. D. MAESTRIPIERI & J. M. MATEO), pp. 133–158. Chicago: The University of Chicago Press.

MATYSIOKOVÁ, B. & REMEŠ, V. (2010). Incubation feeding and nest attentiveness in a socially monogamous songbird: role of feather colouration, territory quality and ambient environment. *Ethology* **116**, 596–607.

MAZER, S. J. (1987). The quantitative genetics of life-history and fitness components in *Raphanus raphanistrum* L. (Brassicaceae): ecological and evolutionary consequences of seed-weight variation. *American Naturalist* **130**, 891–914.

MAZER, S. J. & WOLFE, L. M. (1998). Density-mediated maternal effects on seed size in wild radish: genetic variation and its evolutionary implications. In *Maternal Effects as Adaptations* (eds. T. A. MOUSSEAU & C. W. FOX), pp. 323–343. Oxford: Oxford University Press.

MCADAM, A. G. & BOUTIN, S. (2003). Effects of food abundance on genetic and maternal variation in the growth rate of juvenile red squirrels. *Journal of Evolutionary Biology* **16**, 1249–1256.

MCADAM, A. G., BOUTIN, S., RÉALE, D. & BERTEAUX, D. (2002). Maternal effects and the potential for evolution in a natural population of animals. *Evolution* **56**, 846–851.

MCGINLEY, M. A., TEMME, D. H. & GEBER, M. A. (1987). Parental investment in offspring in variable environments: theoretical and empirical considerations. *American Naturalist* **130**, 370–398.

MCGLOTHLIN, J. W. & BRODIE III, E. D. (2009). How to measure indirect genetic effects: the congruence of trait-based and variance-partitioning approaches. *Evolution* **63**, 1785–1795.

MCGRAW, K. J., ADKINS-REGAN, E. & PARKER, R. S. (2005). Maternally derived carotenoid pigments affect offspring survival, sex ratio, and sexual attractiveness in a colorful songbird. *Naturwissenschaften* **92**, 375–380.

MEATHREL, C. E., SKIRA, I. J., BRADLEY, J. S. & WOOLLER, R. D. (1993). The influence of egg-size, mass and composition upon hatching success in the short-tailed shearwater *Puffinus tenuirostris* (Aves, Procellariiformes). *Journal of Zoology* **230**, 679–686.

MERILÄ, J., KRUUK, L. B. & SHELDON, B. C. (2001). Cryptic evolution in a wild bird population. *Nature* **412**, 76–79.

MERILÄ, J. & SHELDON, B. C. (2001). Avian quantitative genetics. In *Current Ornithology* (ed. V. NOLAN), pp. 179–255. New York: Kluwer Academic/Plenum Publishers.

Metcalfe, N. B. & Monaghan, P. (2001). Compensation for a bad start: grow now, pay later? *Trends in Ecology & Evolution* **16**, 254–260.

Mock, D. W. & Parker, G. A. (1997). *The Evolution of Sibling Rivalry*. Oxford: Oxford University Press.

Moles, A. T., Ackerly, D. D., Webb, C. O., Tweddle, J. C., Dickie, J. B., Pitman, A. J. & Westoby, M. (2005). Factors that shape seed mass evolution. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 10540–10544.

Møller, A. P. & Jennions, M. D. (2001). Testing and adjusting for publication bias. *Trends in Ecology & Evolution* **16**, 580–586.

Møller, A. P. & Jennions, M. D. (2002). How much variance can be explained by ecologists and evolutionary biologists? *Oecologia* **132**, 492–500.

Monaghan, P. (2008). Early growth conditions, phenotypic development and environmental change. *Philosophical Transactions of the Royal Society B-Biological Sciences* **363**, 1635–1645.

Moreno, J., Merino, S., Sanz, J. J., Arriero, E., Morales, J. & Tomás, G. (2005). Nestling cell-mediated immune response, body mass and hatching date as predictors of local recruitment in the pied flycatcher *Ficedula hypoleuca*. *Journal of Avian Biology* **36**, 251–260.

Morley, S. A., Batty, R. S., Geffen, A. J. & Tytler, P. (1999). Egg size manipulation: a technique for investigating maternal effects on the hatching characteristics of herring. *Journal of Fish Biology* **55**, 233–238.

Morrison, K. W., Hipfner, J. M., Gjerdrum, C. & Green, D. J. (2009). Wing length and mass at fledging predict local juvenile survival and age at first return in tufted puffins. *Condor* **111**, 433–441.

Mousseau, T. A. & Fox, C. W. (1998). *Maternal Effects as Adaptations*, Oxford: Oxford University Press.

Murton, R. K., Westwood, N. J. & Isaacson, A. J. (1974). Factors affecting egg-weight, body-weight and molt of woodpigeon *Columba palumbus*. *Ibis* **116**, 52–74.

Nager, R. G. (2006). The challenges of making eggs. *Ardea* **94**, 323–346.

Nager, R. G., Monaghan, P. & Houston, D. C. (2000). Within-clutch trade-offs between the number and quality of eggs: experimental manipulations in gulls. *Ecology* **81**, 1339–1350.

Naguib, M., Nemitz, A. & Gil, D. (2006). Maternal developmental stress reduces reproductive success of female offspring in zebra finches. *Proceedings of the Royal Society B-Biological Sciences* **273**, 1901–1905.

Nisbet, I. C. T. (1978). Dependence of fledging success on egg-size, parental performance and egg-composition among common and roseate terns, *Sterna hirundo* and *S. dougallii*. *Ibis* **120**, 207–215.

O'Connor, R. J. (1975). Initial size and subsequent growth in passerine nestling. *Bird-Banding* **46**, 329–340.

O'Connor, R. J. (1979). Egg weights and brood reduction in the European swift (*Apus apus*). *Condor* **81**, 133–145.

Oh, K. P. & Badyaev, A. V. (2008). Evolution of adaptation and mate choice: parental relatedness affects expression of phenotypic variance in a natural population. *Evolutionary Biology* **35**, 111–124.

Oksanen, T. A., Jokinen, I., Koskela, E., Mappes, T. & Vilpas, H. (2003). Manipulation of offspring number and size: benefits of large body size at birth depend upon the rearing environment. *Journal of Animal Ecology* **72**, 321–330.

Ollason, J. C. & Dunnet, G. M. (1986). Relative effects of parental performance and egg quality on breeding success of fulmars *Fulmarus glacialis*. *Ibis* **128**, 290–296.

Olsson, M., Wapstra, E. & Olofsson, C. (2002). Offspring size-number strategies: experimental manipulation of offspring size in a viviparous lizard (*Lacerta vivipara*). *Functional Ecology* **16**, 135–140.

Ottosson, U., Backman, J. & Smith, H. G. (1997). Begging affects parental effort in the pied flycatcher, *Ficedula hypoleuca*. *Behavioral Ecology and Sociobiology* **41**, 381–384.

Parichy, D. M. & Kaplan, R. H. (1992). Maternal effects on offspring growth and development depend on environmental quality in the frog *Bombina orientalis*. *Oecologia* **91**, 579–586.

Parker, G. A. & Begon, M. (1986). Optimal egg size and clutch size: effects of environment and maternal phenotype. *American Naturalist* **128**, 573–592.

Parker, T. H. (2002). Maternal condition, reproductive investment, and offspring sex ratio in captive red junglefowl (*Gallus gallus*). *Auk* **119**, 840–845.

Parsons, J. (1970). Relationship between egg size and post-hatching chick mortality in herring gull (*Larus argentatus*). *Nature* **228**, 1221–1222.

Pihlaja, M., Siitari, H. & Alatalo, R. V. (2006). Maternal antibodies in a wild altricial bird: effects on offspring immunity, growth and survival. *Journal of Animal Ecology* **75**, 1154–1164.

Pinkowski, B. C. (1975). Growth and development of eastern bluebirds. *Bird-Banding* **46**, 273–289.

Pitala, N., Gustafsson, L., Sendecka, J. & Brommer, J. E. (2007). Nestling immune response to phytohaemagglutinin is not heritable in collared flycatchers. *Biology Letters* **3**, 418–421.

Poole, A., Stettenheim, P. & Gill, F. (1993–2002). *The Birds of North America*, Philadelphia.

Quero, J. L., Villar, R., Marañón, T., Zamora, R. & Poorter, L. (2007). Seed-mass effects in four Mediterranean *Quercus* species (Fagaceae) growing in contrasting light environments. *American Journal of Botany* **94**, 1795–1803.

Quillfeldt, P. & Peter, H. U. (2000). Provisioning and growth in chicks of Wilson's storm-petrels (*Oceanites oceanicus*) on King George Island, South Shetland Islands. *Polar Biology* **23**, 817–824.

Quinn, G. P. & Keough, M. J. (2002). *Experimental Design and Data Analysis for Biologists*, Cambridge: Cambridge University Press.

Ramos, J. A. (2001). Seasonal variation in reproductive measures of tropical roseate terns *Sterna dougallii*: previously undescribed breeding patterns in a seabird. *Ibis* **143**, 83–91.

Ramos, J. A., Maul, A. M., Bowler, J., Wood, L., Threadgold, R., Johnson, S., Birch, D. & Walker, S. (2006). Annual variation in laying date and breeding success of brown noddies on Aride Island, Seychelles. *Emu* **106**, 81–86.

Rauter, C. M. & Moore, A. J. (2002). Evolutionary importance of parental care performance, food resources, and direct and indirect genetic effects in a burying beetle. *Journal of Evolutionary Biology* **15**, 407–417.

Reed, W. L., Clark, M. E. & Vleck, C. M. (2009). Maternal effects increase within-family variation in offspring survival. *American Naturalist*.

Reed, W. L., Turner, A. M. & Sotherland, P. R. (1999). Consequences of egg-size variation in the red-winged blackbird. *Auk* **116**, 549–552.

Reid, W. V. & Boersma, P. D. (1990). Parental quality and selection on egg size in the Magellanic penguin. *Evolution* **44**, 1780–1786.

Remeš, V. (2005). Nest concealment and parental behaviour interact in affecting nest survival in the blackcap (*Sylvia atricapilla*): an experimental evaluation of the parental compensation hypothesis. *Behavioral Ecology and Sociobiology* **58**, 326–332.

Remeš, V., Krist, M., Bertacche, V. & Stradi, R. (2007). Maternal carotenoid supplementation does not affect breeding performance in the great tit (*Parus major*). *Functional Ecology* **21**, 776–783.

Remeš, V. & Martin, T. E. (2002). Environmental influences on the evolution of growth and developmental rates in passerines. *Evolution* **56**, 2505–2518.

Rey, P. J., Alcántara, J. M., Valera, F., Sánchez-Lafuente, A. M., Garrido, J. L., Ramírez, J. M. & Manzaneda, A. J. (2004). Seedling establishment in *Olea europaea*: seed size and microhabitat affect growth and survival. *Ecoscience* **11**, 310–320.

Reznick, D., Nunnev, L. & Tessier, A. (2000). Big houses, big cars, superfleas and the costs of reproduction. *Trends in Ecology & Evolution* **15**, 421–425.

Ricklefs, R. E. (1984a). Components of variance in measurements of nestling European starlings (*Sturnus vulgaris*) in southeastern Pennsylvania. *Auk* **101**, 319–333.

Ricklefs, R. E. (1984b). Egg dimensions and neonatal mass of shorebirds. *Condor* **86**, 7–11.

Ricklefs, R. E., Bruning, D. F. & Archibald, G. W. (1986). Growth rates of cranes reared in captivity. *Auk* **103**, 125–134.

Ricklefs, R. E. & Peters, S. (1981). Parental components of variance in growth rate and body size of nestling European starlings (*Sturnus vulgaris*) in eastern Pennsylvania. *Auk* **98**, 39–48.

Riehl, C. (2010). Egg ejection risk and hatching asynchrony predict egg mass in a communally breeding cuckoo, the Greater Ani (*Crotophaga major*). *Behavioral Ecology* **21**, 676–683.

Riska, B., Rutledge, J. J. & Atchley, W. R. (1985). Covariance between direct and maternal genetic effects in mice, with a model of persistent environmental influences. *Genetical Research* **45**, 287–297.

Roach, D. A. & Wulff, R. D. (1987). Maternal effects in plants. *Annual Review of Ecology and Systematics* **18**, 209–235.

Roff, D. A. (2002). *Life History Evolution*, Sunderland: Sinauer Associates, Inc.

Romano, M., Caprioli, M., Ambrosini, R., Rubolini, D., Fasola, M. & Saino, N. (2008). Maternal allocation strategies and differential effects of yolk carotenoids on the phenotype and viability of yellow-legged gull (*Larus michahellis*) chicks in relation to sex and laying order. *Journal of Evolutionary Biology* **21**, 1626–1640.

Rosenthal, R. (1994). Parametric measures of effect size. In *The Handbook of Research Synthesis* (eds. H. Cooper & L. V. Hedges), pp. 231–244. New York: Russel Sage Foundation.

Rosivall, B., Szöllősi, E. & Török, J. (2005). Maternal compensation for hatching asynchrony in the collared flycatcher *Ficedula albicollis*. *Journal of Avian Biology* **36**, 531–537.

Rubolini, D., Martinelli, R., von Engelhardt, N., Romano, M., Groothuis, T. G. G., Fasola, M. & Saino, N. (2007). Consequences of prenatal androgen exposure for the reproductive performance of female pheasants (*Phasianus colchicus*). *Proceedings of the Royal Society B-Biological Sciences* **274**, 137–142.

Rubolini, D., Romano, M., Alquati, A. B. & Saino, N. (2006a). Early maternal, genetic and environmental components of antioxidant protection, morphology and immunity of yellow-legged gull (*Larus michahellis*) chicks. *Journal of Evolutionary Biology* **19**, 1571–1584.

Rubolini, D., Romano, M., Martinelli, R. & Saino, N. (2006b). Effects of elevated yolk testosterone levels on survival, growth and immunity of male and female yellow-legged gull chicks. *Behavioral Ecology And Sociobiology* **59**, 344–352.

Russell, A. F., Langmore, N. E., Cockburn, A., Astheimer, L. B. & Kilner, R. M. (2007). Reduced egg investment can conceal helper effects in cooperatively breeding birds. *Science* **317**, 941–944.

RUTKOWSKA, J. & CICHOŃ, M. (2005). Egg size, offspring sex and hatching asynchrony in zebra finches *Taeniopygia guttata*. *Journal of Avian Biology* **36**, 12–17.

SAINO, N., FERRARI, R., ROMANO, M., MARTINELLI, R. & MØLLER, A. P. (2003). Experimental manipulation of egg carotenoids affects immunity of barn swallow nestlings. *Proceedings of the Royal Society of London Series B-Biological Sciences* **270**, 2485–2489.

SAMELIUS, G. & ALISAUSKAS, R. (1999). Diet and growth of glaucous gulls at a large Arctic goose colony. *Canadian Journal of Zoology* **77**, 1327–1331.

SAS INSTITUTE INC. (2003). *SAS Online Doc, Version 9.1*. Cary: SAS Institute Inc.

SCHIFFERLI, L. (1973). Effect of egg weight on subsequent growth of nestling great tits *Parus major*. *Ibis* **115**, 549–558.

SCHLUTER, D. & GUSTAFSSON, L. (1993). Maternal inheritance of condition and clutch size in the collared flycatcher. *Evolution* **47**, 658–667.

SCHLUTER, D., PRICE, T. D. & ROWE, L. (1991). Conflicting selection pressures and life-history trade-offs. *Proceedings of the Royal Society of London Series B-Biological Sciences* **246**, 11–17.

SCHOECH, S. J. & HAHN, T. P. (2008). Latitude affects degree of advancement in laying by birds in response to food supplementation: a meta-analysis. *Oecologia* **157**, 369–376.

SCHRADER, M. & TRAVIS, J. (2009). Do embryos influence maternal investment? Evaluating maternal-fetal coadaptation and the potential for parent-offspring conflict in a placental fish. *Evolution* **63**, 2805–2815.

SCHWABL, H. (1993). Yolk is a source of maternal testosterone for developing birds. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 11446–11450.

SCHWABL, H. (1996). Maternal testosterone in the avian egg enhances postnatal growth. *Comparative Biochemistry and Physiology A-Physiology* **114**, 271–276.

SCHWAGMEYER, P. L. & MOCK, D. W. (2008). Parental provisioning and offspring fitness: size matters. *Animal Behaviour* **75**, 291–298.

SELMAN, R. G. & HOUSTON, D. C. (1996). The effect of prebreeding diet on reproductive output in zebra finches. *Proceedings of the Royal Society of London Series B-Biological Sciences* **263**, 1585–1588.

SEMLITSCH, R. D. & GIBBONS, J. W. (1990). Effects of egg size on success of larval salamanders in complex aquatic environments. *Ecology* **71**, 1789–1795.

SHELDON, B. C. (2000). Differential allocation: tests, mechanisms and implications. *Trends in Ecology & Evolution* **15**, 397-402.

SHELDON, B. C., KRUUK, L. E. B. & MERILÄ, J. (2003). Natural selection and inheritance of breeding time and clutch size in the collared flycatcher. *Evolution* **57**, 406–420.

SILVA, M. C., BOERSMA, P. D., MACKAY, S. & STRANGE, I. (2007). Egg size and parental quality in thin-billed prions, *Pachyptila belcheri*: effects on offspring fitness. *Animal Behaviour* **74**, 1403–1412.

SINERVO, B. (1990). The evolution of maternal investment in lizards: an experimental and comparative analysis of egg size and its effects on offspring performance. *Evolution* **44**, 279–294.

SINERVO, B. & DOUGHTY, P. (1996). Interactive effects of offspring size and timing of reproduction on offspring reproduction: experimental, maternal, and quantitative genetic aspects. *Evolution* **50**, 1314–1327.

SINERVO, B., DOUGHTY, P., HUEY, R. B. & ZAMUDIO, K. (1992). Allometric engineering: a causal analysis of natural selection on offspring size. *Science* **258**, 1927–1930.

SINERVO, B. & LICHT, P. (1991a). Hormonal and physiological control of clutch size, egg size, and egg shape in side-blotched lizards (*Uta stansburiana*): constraints on the evolution of lizard life histories. *Journal of Experimental Zoology* **257**, 252–264.

SINERVO, B. & LICHT, P. (1991b). Proximate constraints on the evolution of egg size, number, and total clutch mass in lizards. *Science* **252**, 1300–1302.

SINERVO, B. & MCEDWARD, L. R. (1988). Developmental consequences of an evolutionary change in egg size: an experimental test. *Evolution* **42**, 885–899.

SLAGSVOLD, T., SANDVIK, J., ROFSTAD, G., LORENTSEN, O. & HUSBY, M. (1984). On the adaptive value of intraclutch egg-size variation in birds. *Auk* **101**, 685–697.

SMITH, C. C. & FRETWELL, S. D. (1974). The optimal balance between size and number of offspring. *American Naturalist* **108**, 499–506.

SMITH, H. G. & BRUUN, M. (1998). The effect of egg size and habitat on starling nestling growth and survival. *Oecologia* **115**, 59–63.

SOLER, J. J., MORENO, J. & POTTI, J. (2003). Environmental, genetic and maternal components of immunocompetence of nestling pied flycatchers from a cross-fostering study. *Evolutionary Ecology Research* **5**, 259–272.

STANTON, M. L. (1984). Seed variation in wild radish: effect of seed size on components of seedling and adult fitness. *Ecology* **65**, 1105–1112.

STEARNS, S. C. (1992). *The Evolution of Life Histories*, Oxford: Oxford University Press.

STEMPNIEWICZ, L. (1980). Factors influencing the growth of the little auk, *Plautus alle* (L.), nestlings on Spitsbergen. *Ekologia Polska* **28**, 557–581.

STOKLAND, J. N. & AMUNDSEN, T. (1988). Initial size hierarchy in broods of the shag: relative significance of egg size and hatching asynchrony. *Auk* **105**, 308–315.

STYRSKY, J. D., DOBBS, R. C. & THOMPSON, C. F. (2000). Food-supplementation does not override the effect of egg mass on fitness-related traits of nestling house wrens. *Journal of Animal Ecology* **69**, 690–702.

STYRSKY, J. D., ECKERLE, M. P. & THOMPSON, C. F. (1999). Fitness-related consequences of egg mass in nestling house wrens. *Proceedings of the Royal Society of London Series B-Biological Sciences* **266**, 1253–1258.

SVENSSON, E. & SINERVO, B. (2000). Experimental excursions on adaptive landscapes: density-dependent selection on egg size. *Evolution* **54**, 1396–1403.

TILGAR, V., MÄND, R., KILGAS, P. & MÄGI, M. (2010). Long-term consequences of early ontogeny in free-living great tits *Parus major*. *Journal of Ornithology* **151**, 61–68.

VAN DE POL, M., BAKKER, T., SAALTINK, D. J. & VERHULST, S. (2006). Rearing conditions determine offspring survival independent of egg quality: a cross-foster experiment with oystercatchers *Haematopus ostralegus*. *Ibis* **148**, 203–210.

VAN DE POL, M. V. & WRIGHT, J. (2009). A simple method for distinguishing within-versus between-subject effects using mixed models. *Animal Behaviour* **77**, 753–758.

VAN HOUWELINGEN, H. C., ARENDS, L. R. & STIJNEN, T. (2002). Advanced methods in meta-analysis: multivariate approach and meta-regression. *Statistics in Medicine* **21**, 589–624.

VAN NOORDWIJK, A. J. & DE JONG, G. (1986). Acquisition and allocation of resources: their influence on variation in life-history tactics. *American Naturalist* **128**, 137–142.

VELANDO, A., TORRES, R. & ESPINOSA, I. (2005). Male coloration and chick condition in blue-footed booby: a cross-fostering experiment. *Behavioral Ecology and Sociobiology* **58**, 175–180.

VORBURGER, C. (2005). Positive genetic correlations among major life-history traits related to ecological success in the aphid *Myzus persicae*. *Evolution* **59**, 1006–1015.

WAGNER, E. C. & WILLIAMS, T. D. (2007). Experimental (antiestrogen-mediated) reduction in egg size negatively affects offspring growth and survival. *Physiological and Biochemical Zoology* **80**, 293–305.

WARNER, D. A. & ANDREWS, R. M. (2002). Laboratory and field experiments identify sources of variation in phenotypes and survival of hatchling lizards. *Biological Journal of the Linnean Society* **76**, 105–124.

WARNER, D. A. & SHINE, R. (2009). Maternal and environmental effects on offspring phenotypes in an oviparous lizard: do field data corroborate laboratory data? *Oecologia* **161**, 209–220.

WEIDINGER, K. (1996). Egg variability and hatching success in the Cape petrel *Daption capense* at Nelson Island, South Shetland Islands, Antarctica. *Journal of Zoology* **239**, 755–768.

WEIDINGER, K. (1997). Variations in growth of Cape petrel *Daption capense* chicks. *Journal of Zoology* **242**, 193–207.

WEIDINGER, K. (2002). Interactive effects of concealment, parental behaviour and predators on the survival of open passerine nests. *Journal of Animal Ecology* **71**, 424–437.

WEST, B. T., WELCH, K. B. & GAŁECKI, A. T. (2007). *Linear Mixed Models: A Practical Guide Using Statistical Software*. Boca Raton: Chapman & Hall/CRC.

WHITTINGHAM, L. A., DUNN, P. O. & LIFJELD, J. T. (2007). Egg mass influences nestling quality in tree swallows, but there is no differential allocation in relation to laying order or sex. *Condor* **109**, 585–594.

WIEBE, K. L. & BORTOLOTTI, G. R. (1995). Egg size and clutch size in the reproductive investment of American kestrels. *Journal of Zoology* **237**, 285–301.

WILLIAMS, T. D. (1990). Growth and survival in Macaroni penguin, *Eudyptes chrysolophus*: A-chicks and B-chicks—do females maximize investment in the large B-egg? *Oikos* **59**, 349–354.

WILLIAMS, T. D. (1994). Intraspecific variation in egg size and egg composition in birds: effects on offspring fitness. *Biological Reviews* **69**, 35–59.

WILLIAMS, T. D. (2001). Experimental manipulation of female reproduction reveals an intraspecific egg size-clutch size trade-off. *Proceedings of the Royal Society of London Series B-Biological Sciences* **268**, 423–428.

WILLIAMS, T. D. (2005). Mechanisms underlying the costs of egg production. *Bioscience* **55**, 39–48.

WILLIAMS, T. D., LANK, D. B. & COOKE, F. (1993a). Is intraclutch egg-size variation adaptive in the lesser snow goose? *Oikos* **67**, 250–256.

WILLIAMS, T. D., LANK, D. B., COOKE, F. & ROCKWELL, R. F. (1993b). Fitness consequences of egg-size variation in the lesser snow goose. *Oecologia* **96**, 331–338.

WILSON, A. J. & RÉALE, D. (2006). Ontogeny of additive and maternal genetic effects: lessons from domestic mammals. *American Naturalist* **167**, E23–E38.

WILSON, A. J., RÉALE, D., CLEMENTS, M. N., MORRISSEY, M. M., POSTMA, E., WALLING, C. A., KRUUK, L. E. B. & NUSSEY, D. H. (2010). An ecologist's guide to the animal model. *Journal of Animal Ecology* **79**, 13–26.

WINN, A. A. (2004). Natural selection, evolvability and bias due to environmental covariance in the field in an annual plant. *Journal of Evolutionary Biology* **17**, 1073–1083.

WOLF, J. B. & BRODIE III, E. D. (1998). The coadaptation of parental and offspring characters. *Evolution* **52**, 299–308.

ZANETTE, L., CLINCHY, M. & SUNG, H. C. (2009). Food-supplementing parents reduces their sons' song repertoire size. *Proceedings of the Royal Society B-Biological Sciences* **276**, 2855–2860.

ZIELIŃSKI, P. & BAŃBURA, J. (1998). Egg size variation in the barn swallow *Hirundo rustica*. *Acta Ornithologica* **33**, 191–196.

ZUUR, A. F., IENO, E. N. & SMITH, G. M. (2007). *Analysing Ecological Data*, New York: Springer.

VIII. SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article.

Appendix S1. Search terms and lists of studies that were included ($N = 283$) and not included ($N = 383$) in the meta-analysis. The selection criterion which disallowed inclusion of a given study into the meta-analysis is given for each excluded study.

Appendix S2. Phylogenetic relationships among species included in the meta-analysis and methods of phylogenetic regression.

Appendix S3. The dataset. This *Excel* file has two sheets labelled "analyzed data" and "all data". The sheet "analyzed data" has 23 columns with a list of variables used in statistical analyses and 1805 rows that correspond to 1805 analyzed estimates. The sheet "all data" has additional columns and 2318 rows. The additional columns contain for example: original statistics given in the published papers; formulae used to adjust effect size for dichotomization and range restriction/enhancement, and to compute Pearson's *r* from other statistics; and variables that divide study into subgroups and those that were statistically controlled for when testing for egg-size effect. The sheet "analyzed data" can be created from the sheet "all data" by selecting columns whose headings are given in red bold and rows that do not contain the phrase "pseudoreplication" ($N = 512$) or "rare design" ($N = 1$) in column named "reason for exclusion". References for the studies included in this appendix are given in Appendix S1.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Článek 11: Krist, M. & Munclinger, P. (2011) Superiority of extra-pair offspring: maternal but not genetic effects as revealed by a mixed cross-fostering design. *Molecular Ecology*, **20, 5074-5091.**

Superiority of extra-pair offspring: maternal but not genetic effects as revealed by a mixed cross-fostering design

MILOŠ KRIST*† and PAVEL MUNCLINGER‡

*Museum of Natural History, nám. Republiky 5, 771 73 Olomouc, Czech Republic, †Department of Zoology and Laboratory of Ornithology, Faculty of Science, Palacký University, tř. Svobody 26, 771 46 Olomouc, Czech Republic, ‡Department of Zoology, Faculty of Science, Charles University in Prague, Viničná 7, 128 44 Praha, Czech Republic

Abstract

Extra-pair copulations (EPC) are the rule rather than an exception in socially monogamous birds, but despite widespread occurrences, the benefits of female infidelity remain elusive. Most attention has been paid to the possibility that females gain genetic benefits from EPC, and fitness comparisons between maternal half-siblings are considered to be a defining test of this hypothesis. Recently, it was shown that these comparisons may be confounded by within-brood maternal effects where one such effect may be the distribution of half-siblings in the laying order. However, this possibility is difficult to study as it would be necessary to detect the egg from which each chick hatched. In this study, we used a new approach for egg-chick assignment and cross-fostered eggs on an individual basis among a set of nests of the collared flycatcher *Ficedula albicollis*. After hatching, chicks were ascribed to mothers and therefore to individual eggs by molecular genetic methods. Extra-pair young predominated early in the laying order. Under natural conditions, this should give them a competitive advantage over their half-siblings, mediated by hatching asynchrony. However, we experimentally synchronized hatching, and after this treatment, extra-pair young did not outperform within-pair young in any studied trait including survival up to recruitment and several indicators of reproductive success and attractiveness. We obtained only modest sample sizes for the last two traits and did not test for extra-pair success of male offspring. Thus, we cannot exclude the possibility of advantages of extra-pair young during the adult phase of life. However, our data tentatively suggest that the more likely reason for females' EPCs is the insurance against the infertility of a social mate.

Keywords: extra-pair paternity, fertility insurance, good genes, hatching asynchrony, maternal effects, offspring fitness

Received 3 May 2011; revision received 19 September 2011; accepted 28 September 2011

Introduction

Females in most socially monogamous bird species also mate with extra-pair males (Griffith *et al.* 2002), but this is potentially costly for females. Promiscuity can expose females to sexually transmitted diseases (Sheldon 1993; Kempenaers & Schlicht 2010), and furthermore, their mates may reduce their parental care

Correspondence: Miloš Krist, Fax: +420 585 222 743;
E-mails: milos.krist@upol.cz, milos.krist@volny.cz

of broods containing extra-pair young (EPY, Arnqvist & Kirkpatrick 2005). Although males are not able to recognize their young (Kempenaers & Sheldon 1996), they are able to assess paternity loss owing to indirect cues such as female behaviour during the fertile period (Sheldon 2002). The wide occurrence of extra-pair mating suggests that these costs are outweighed by some benefits. By copulating with extra-pair males, females may benefit from increased access to food resources (Gray 1997; Tryjanowski & Hromada 2005), and broods containing extra-pair young may also be

fed by extra-pair mates (Townsend *et al.* 2010). Females may also engage in extra-pair copulations (EPC) to insure against the infertility of a social mate (Sheldon 1994; Hasson & Stone 2009). Some of these direct benefits seem to be limited to special cases and life histories which is why most attention has been paid to potentially more universal indirect benefits. Females are hypothesized to improve the genetic quality of their offspring by extra-pair mating with a genetically more viable or compatible mate (reviews in Jennions & Petrie 2000; Neff & Pitcher 2005; Akçay & Roughgarden 2007; Kempenaers 2007; Mays *et al.* 2008). The strongest prediction of both of these hypotheses is the superior performance of EPY as compared to young sired by within-pair mates (WPY) (Sheldon *et al.* 1997; Griffith *et al.* 2002; Arnqvist & Kirkpatrick 2005; Hasson & Stone 2011).

The first studies comparing the quality of half-siblings appeared in the late 1990s (Kempenaers *et al.* 1997; Sheldon *et al.* 1997; Krokene *et al.* 1998) and since then many such tests have been published (reviews in Arnqvist & Kirkpatrick 2005; Akçay & Roughgarden 2007; Sardell *et al.* 2011). This research has shown that, on average, EPY do indeed outperform WPY, which suggests their genetic superiority but this indirect effect is very weak (Akçay & Roughgarden 2007). Such a small genetic benefit might be unable to outweigh the direct costs of EPC, which suggests that EPC are a male-driven strategy (Arnqvist & Kirkpatrick 2005). However, before ruling out the indirect benefits as a cause of EPC behaviour, it is vital to know whether the magnitude of the genetic effect has been estimated correctly. Three factors might have caused problems with the estimation of indirect benefits to date. First, genetic effects might be context dependent and thus manifested only under some environmental conditions (Sheldon 2000a; Schmoll *et al.* 2005; Garvin *et al.* 2006; O'Brien & Dawson 2007), hence the need for multiyear studies that test for genetic effects under a range of environmental conditions (Sheldon 2000a). Second, most previous studies were aimed at only one component of offspring fitness such as nestling quality or survival up to recruitment (Sardell *et al.* 2011), while the reproductive success of half-siblings was assessed in only one species (Schmoll *et al.* 2003, 2009). This large gap in our knowledge might contribute to an underestimation of the genetic benefits of EPC (Eliassen & Kokko 2008). Third, genetic effects might be confounded with maternal effects where, for example, EPY might be non-randomly distributed in the laying sequence (Krist *et al.* 2005). As many species initiate incubation before clutch completion (Clark & Wilson 1981; Stoleson & Beissinger 1995; Wang & Beissinger 2009), chicks from late eggs are delayed in development and thus

disadvantaged in sibling competition (Magrath 1990; Stoleson & Beissinger 1995; Kim *et al.* 2010). If EPY were laid in early eggs, they would have a maternally derived competitive advantage compared to WPY, and this pattern was recently found in three passerine species (Johnson *et al.* 2009; Magrath *et al.* 2009; Ferree *et al.* 2010). So there is a need to control the laying/hatching order when comparing fitness-relevant traits of naturally occurring half-siblings (Kempenaers & Schlicht 2010).

An investigation into the distribution of EPY in the laying order requires an assignment of chicks to eggs. This is a major challenge for field studies as many chicks may hatch in a short time interval, but several methods can be employed to overcome this difficulty. Nests can be visited several times a day (Whittingham *et al.* 2007; Ferree *et al.* 2010) or eggs can be transferred to an incubator, while supplying nests with dummy eggs to prevent nest desertion, and then the hatched chicks returned to their nests (Krist *et al.* 2004; Magrath *et al.* 2009). Both of these methods are time-consuming. Chicks can also be assigned by putting the eggs into net sacks (Hořák & Albrecht 2007) or by injecting dye below the eggshell shortly before hatching (Sockman & Schwabl 2000; Rubolini *et al.* 2005), but these two methods are limited to larger species. Lastly, incubated eggs may be frozen, which can reveal the distribution of offspring sex (Cordero *et al.* 2000) or paternity (Cordero *et al.* 1999) in relation to laying order but is naturally not suitable if chick performance is of interest. Furthermore, in most species, this invasive method would raise serious ethical concerns.

The aim of this study was to test for the potential genetic benefits of EPC in the collared flycatcher (*Ficedula albicollis*) by comparing the performance of chicks sired by extra-pair mates and social mates. We conducted a mixed cross-fostering experiment to further reveal the distribution of EPY in the laying sequence. Eggs were cross-fostered among nests on an individual basis, so that each nest contained the same number of eggs as was in its original clutch, but each egg originated from a different donor nest. Consequently, this allowed us to assign chicks to mothers and therefore to individual eggs by molecular genetic methods. The potentially confounding impact of hatching asynchrony on chick quality was minimized as all eggs in mixed clutches started being incubated at the same time. The study spanned over four years that differed in environmental conditions which should add credibility to our test if genetic benefits are context dependent. Chicks were followed until recruitment, and some indicators of their attractiveness and reproductive success were tested for dependence on their paternity status.

Methods

Study population

This study was conducted in the Velký Kosíř area, Moravia, Czech Republic (49°32′N, 17°04′E, 300–400 m a.s.l.), in 2006–2010. Experiments were performed on a collared flycatcher population breeding in approximately 350 nest boxes in an oak (*Quercus petrea*) forest. The collared flycatcher is a small (13 g) hole-breeding passerine with sexually dimorphic plumage. Females are brownish with a small white wing patch, while males have more contrasting plumage with a white forehead patch and a larger wing patch. Females usually lay 4–8 eggs, and most of them start full incubation before clutch completion, usually on the penultimate or pre-penultimate eggs (Potti 1998). The early onset of incubation results in partial hatching asynchrony of the clutch (Krist *et al.* 2004; Rosivall *et al.* 2005).

Mixed cross-fostering experiment

The mixed cross-fostering experiment was performed in 2006–2009. Nest boxes with completed nests were visited daily, and each natural egg was exchanged for a dummy one on the day it was laid. It was recorded whether the eggs were warm on each visit. The first day that the eggs were warm was labelled as the start of the continuous incubation if the eggs were not found cold on later days. Natural eggs were measured by a digital caliper, and their volume was calculated according to Hoyt's (1979) formula to the nearest 1 mm^3. They were stored in a dry safe place located at one of the study plots. The day after the last egg in the focal nest was laid, the dummy eggs were replaced by the same number of natural eggs. Each of these natural eggs was originally laid in a different nest, excluding the focal one. As a result of this manipulation, each female started to incubate the whole experimental clutch of cross-fostered eggs at the same time. Experimental clutches were not created randomly from all available eggs; typically, eggs laid earlier were also used for earlier clutches to reduce the risk of their decreased viability while left unincubated (Cook *et al.* 2003; D'Alba *et al.* 2010; Wang *et al.* 2011). Eggs were stored and therefore left unincubated for 0–7 days, which is within the range of natural variation in this species. Storage time did not affect hatchability of eggs (mean storage time for hatched eggs: 3.82 days, $n = 937$, unhatched eggs: 3.59 days, $n = 27$, $F_{1, 962} = 0.67$, $P = 0.412$).

The first nest in which the laying was completed was baited with the first eggs from other nests. If possible, eggs of the same laying order were used to create an experimental clutch. The use of earlier laid eggs was not an absolute rule as we aimed to increase the number of combinations of donor nests in experimental clutches (see Appendix S1, Supporting Information for an example of field protocol). This mixed cross-fostering design relies on the synchronized breeding of several pairs where the minimum requirement is to have eggs available from the number of nests equal to the clutch size of the female that just commenced laying. In our study population, it was possible to include nearly all nests in this type of experimental design.

Other field procedures

To detect the hatching day, experimental nests were checked from day 12 onward after experimental clutches were created. At 6 days of age, chicks were weighed (to the nearest 0.25 g), ringed and blood-sampled from the tarsal vein. By this method, only a small amount of blood was withdrawn (usually 1–5 μL) compared to the more common method of brachial venipuncture, which may have significant fitness costs (Brown & Brown 2009). The small amount of blood sampled was nevertheless fully sufficient for our aim, and we obtained genotypes from all blood samples. Unhatched eggs and dead nestlings were taken from nests and stored in ethanol. At 12 days of age, the width of the right wing web was measured twice by a thickness gauge (Mitutoyo Quick-Mini) and then injected with 0.1 mg of phytohaemagglutinin (PHA) in 20 μL of physiological saline solution. This procedure for the scoring of PHA immunity was applied only in 2006 and 2007. At 13 days, the chicks were reweighed, and the length of their tarsi (to the nearest 0.01 mm) and wings (to the nearest 1 mm) and the thickness of their wing webs (to the nearest 0.01 mm) were measured. The difference between the average thicknesses of the wing web before and after the injection was used as an estimate of the individual proinflammatory potential (see Vinkler *et al.* 2010). The nestling condition was estimated as residuals from the linear regression of the 13-day body mass on the tarsus length (body mass = −10.9 + 1.29 × tarsus length, $n = 711$, $R^2 = 0.34$).

During 2006–2010, collared flycatchers initiated (laid at least one egg) 704 breeding attempts on our study plots. We captured 520 male parents and 580 female parents while feeding nestlings. Ring identity of captured adults was used to identify which of the young from the experimental nests had been recruited to the study population. It is likely that some young that recruited to the breeding population were not recorded by us due to dispersal outside study plots and our inability to catch all breeding adults, such as due to early failure of the breeding attempt. However, we recorded 120 recruits from experimental nests giving

0.75 recruits per breeding attempt and local recruitment of 11.7% of laid eggs or 16.9% of fledged young. For comparison, in an extensively studied island population of this species, the average number of recruits per breeding attempt was 0.64 (Gustafsson 1989). High values of local recruitment in our population suggest that the status of most offspring was identified correctly and that the results on recruits are reliable.

Several traits were measured on recruits: (i) laying date (Julian day, 1st January = 1), (ii) clutch size, (iii) mean egg volume, (iv) body mass, (v) tarsus length, (vi) wing length, (vii) wing patch, calculated as the sum of visible white on primaries 3–8 as measured from the tips of the coverts to the distal part of wing (in mm), and (viii) forehead patch in males. Two pictures of the forehead patch were taken by digital camera. A ruler was aligned alongside the forehead patch on each picture, the white patch was manually encircled and the size of encircled area was computed by J-image software to the nearest 0.1 mm^2. All these procedures, including the photography, were performed twice, and the mean of the two measurements was taken as the size of the forehead patch. In experimental nests, both parents were blood-sampled by tibial venipuncture, and similar to the tarsal venipuncture in nestlings, only a small amount of blood (usually 1–5 µL) was withdrawn. All males in non-experimental nests were blood-sampled in 2007–2009.

Genotyping

DNA was extracted from blood and tissue samples using DNeasy® Blood & Tissue kit (Qiagen), and offspring and adults were genotyped at eight polymorphic microsatellite autosomal loci (Table 1). The microsatellite loci were amplified in a single multiplex PCR using fluorescently labelled primers and a Type-it® Microsatellite PCR kit (Qiagen). The reaction conditions used were 5 min. at 95 °C followed by 30 cycles of 30 s at 95 °C, 90 s at 65 °C and 30 s at 72 °C, with a final extension of 30 min at 60 °C. PCR products were mixed with GeneScan™–500 LIZ® Size Standard (Applied Biosystems) and analysed using ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems). Genotypes were then scored with the GeneMarker® version 1.9 software (Softgenetics), and loci characteristics based on allele frequencies were calculated using the Cervus version 3.0.3 (Field Genetics Ltd). The combined non-exclusion probability of this marker set was 5.67×10^{-4} for the first parent. However, locus Fhy310 had a considerable excess of homozygotes suggesting allelic dropouts or null alleles and was not used for inferences based on pairwise comparisons of offspring and candidate parents. The combined non-exclusion probability of the reduced set of seven loci was 1.48×10^{-3}.

Blood samples and genotypes were taken more than once from 147 birds owing to their breeding in multiple years or from repeated sampling of polygamous males in the same year. Some individuals were sampled twice (103), three times (37), four times (6) and one-five times, but only one genotype per individual was included for the above computation of allele frequencies and exclusion power. In total, 1235 individuals were genotyped, 1169 of them on all seven loci, 60 on six loci and six on five loci. Individuals that could not be genotyped at five loci were excluded from the analyses. The repeated genotyping of the same individuals was used to assess the frequency of genotyping errors. Two individual genotypes obtained in one sampling episode were barely readable and therefore were replaced with the genotypes obtained in the other sampling episode. Of the remaining 145 individuals, 127 (88%) had identical genotypes in all sampling episodes, 16 (11%) individuals differed at one locus and two (1.4%) differed at two loci. These errors can be ascribed to allelic dropouts or null alleles (11 cases) or a shift of 2–4 base pairs (nine cases).

Table 1 The characteristics of the marker set of eight microsatellite loci

Locus	k	N	Hobs	Hexp	F(Null)
Fhu2	20	1235	0.847	0.846	−0.0008
Cuμ04	20	1234	0.870	0.866	−0.0031
Fhy310	18	1227	0.726	0.883	+0.0968
Fhy405	28	1194	0.818	0.904	+0.0501
Fhy407	27	1234	0.919	0.910	−0.0052
Fhy428	24	1224	0.853	0.863	+0.0054
Fhy431	20	1235	0.873	0.866	−0.0040
Fhy452	32	1217	0.844	0.842	−0.0014

k, number of alleles; N, number of typed individuals; Hobs, observed heterozygosity; Hexp, expected heterozygosity; F(Null), estimated frequency of null alleles. Loci are described in Ellegren 1992 [Fhu2 (=PTC3)], Gibbs *et al.* (1999) (Cuμ04) and [Leder *et al.* (2008), Fhy310, Fhy405, Fhy407, Fhy428, Fhy431, Fhy452].

Parentage assignment

Owing to the nature of the mixed cross-fostering experiment, neither fathers nor mothers were known and maternity was assigned based on pairwise comparisons of offspring and candidate mother genotypes. The number of candidate mothers was equal to the clutch size in the focal nest, which ranged from four to eight. Conspecific brood parasitism does not occur in this population as confirmed by genetic data (Krist *et al.* 2005) and our observations that two eggs were not laid in any experimental nest during one day. Consequently, the set of candidate mothers was known with certainty, and for most nests, assignment of maternity was unequivocal by this pairwise comparison method. Typically, each offspring in the focal nest matched only one candidate mother on all loci, while all other potential combinations would involve a number of mismatches in two or more offspring–mother pairs (Appendix S2, Supporting Information).

For 17 of the offspring from eight nests, maternity was impossible to resolve with certainty as two or more offspring–mother pairs were equally likely (Appendix S2, Supporting Information). Another problem arose due to the lack of genotypes from some candidate mothers. If only one of the candidate mothers was ungenotyped for the focal nest, maternity could usually be resolved as all but one offspring matched the genotyped candidate mothers. Thus, the remaining offspring that did not match any genotyped candidate mother was assigned to an ungenotyped one. If two ($n = 17$) or three ($n = 1$) mothers were ungenotyped, the offspring that did not match any genotyped mother could not be simply assigned to the ungenotyped one.

For resolution of these cases, another approach was adopted where maternity could still be inferred based solely on an analysis of the genotypes of the candidate siblings. As the number of siblings ranged from four to eight, this analysis is more powerful than the pairwise mother–offspring comparison as it is based on a higher number of genotypes (see Walling *et al.* 2010). We ran the likelihood-based sibling analyses in the software Colony, version 2.0 (Wang 2004; Wang & Santure 2009), using also locus Fhy310. Information about the excess of homozygotes at this locus was included in likelihood calculations. For each offspring, Colony indicated the most probable parents labelled with arbitrary numbers as we did not include parental identities in these analyses. The results of these analyses confirmed the previous assignments made by offspring–mother pairwise comparisons in Cervus 3.0.3 (Kalinowski *et al.* 2007) and also resolved all controversies. For the 17 ambiguous offspring, the assignment to families was carried out based on the following results. (i) For nine offspring, Colony indicated that they were full-sibs with all other offspring of one candidate mother. (ii) Five offspring were indicated to be full-sibs of some of the offspring of one candidate mother and half-sibs with her remaining offspring, as expected for mixed-paternity broods. (iii) Three remaining offspring were indicated as half-sibs with only some offspring from one of the candidate mothers. The other offspring of this candidate mother were not indicated as sibs of the focal offspring. However, these seemingly unrelated offspring were half-sibs of those offspring related to the focal one. Of the 37 offspring with an ungenotyped mother, 20 passed into full-sib families, 12 into families with half-sibs and five into families which included some offspring unrelated to the focal one (according to likelihood algorithm).

Overall, for 846 offspring with genotyped mothers, 740 had no mismatch with their mother (87%), 99 had one mismatch (12%) and seven had two mismatches (1%). These figures are similar to those that compare genotypes obtained repeatedly for the same individuals indicating that most mismatches in mother–offspring pairs were probably caused by genotyping errors.

After maternity was determined for each offspring, paternity exclusion could begin where the distribution of the number of mismatches in offspring–social father pairs is given in Table 2. Males were considered genetic fathers if they had 0 or 1 mismatch with the offspring. For those with two mismatches, Cervus indicated a 95% probability for the male that fathered the offspring in three cases. These offspring were considered to be fathered by a social mate, and all other offspring were considered to be fathered by an extra-pair mate.

Statistical analyses

Only clutches with mixed paternity were used for the test of whether extra-pair young appear randomly in the laying order, and two analyses of this hypothesis were conducted. In the first analysis, the actual laying order and the relative size of the egg (egg volume minus mean egg volume of the clutch) were the fixed

Table 2 The distribution of the number of mismatches between 800 offspring and their social fathers

Mismatches	0	1	2	3	4	5	6	7
Cases	505	78	21	44	56	64	25	7

factors of interest, while the year and laying date in the nest of origin were included as fixed covariates. The identity of the original nest was the random factor, and the paternity was the response variable. As clutch sizes in mixed-paternity clutches ranged from four to eight (mean ± SD, 6.48 ± 0.678), the relative laying order might be a more relevant variable than the absolute laying order. Therefore, in the second model, actual laying order was replaced by the relative laying order (actual laying order divided by clutch size; Magrath *et al.* 2009; Ferree *et al.* 2010).

For the comparison of EPY and WPY performance, we used a suite of factors. These included the survival probabilities during three life stages, five morphological and one immunological trait of nestlings that predict post-fledging survival (Kruuk *et al.* 2001; Merilä *et al.* 2001; Moreno *et al.* 2005). We also compared the morphology of recruits to test whether the potential differences between within-pair and extra-pair nestlings persist until adulthood, although these adult traits are generally under weak selection (Przybylo *et al.* 2000). Furthermore, we tested the difference in three life history traits (egg size, clutch size and laying date) that affect reproductive success (e.g. Gustafsson & Sutherland 1988; Krist 2009). Specifically, laying date is under strong directional selection in this species (Sheldon *et al.* 2003), and egg size is sometimes used as a surrogate of female quality (Hõrak *et al.* 1997). Finally, we compared the size of two ornamental traits which plays a role in sexual selection (e.g. Sheldon & Ellegren 1999; Garamszegi *et al.* 2006).

Only the young originating from mixed-paternity clutches were used for the comparison of EPY and WPY performance. All statistical models where the response variable was nestling traits or offspring survival had a similar structure. Both the nest of origin and the nest of rearing were included as random factors, paternity as the fixed factor of interest, and the year and laying date in the nest of rearing as fixed covariates. Nests with complete failures were excluded when the response variable was egg hatchability and nestling survival. Models with a binary response variable had a logit link function, while those with a continuous response variable had an identity link function where the latter models were based on young that fledged. Models where the response variable was a trait measured on recruits had additional fixed covariates. All these models initially included the year when traits on recruits were measured, the age (in years) and the sex of the recruits (excluding single-sex models). The model on body mass of the recruits also included the age of the recruit's offspring at the time when the recruits were captured to control for mass recession during provisioning of nestlings (see Krist 2009). These

models were simplified by backward elimination of non-significant covariates ($P > 0.1$). The natal year, the recruit age and the breeding year were intercorrelated; thus, the reduced models are not suitable for inferences about these three intercorrelated factors. They are, however, well suited for the inference about paternity effects, which was the main target of this study. Some recruits were captured in multiple years, but only data from one randomly selected year were included to avoid pseudoreplication. All above models were fitted in SAS 9.1, Proc Glimmix (SAS Institute 2003).

The relationship between the number of eggs laid after the start of continuous incubation and the proportion of EPY in mixed-paternity clutches was tested in Proc Genmod (SAS Institute 2003). The response variable was the proportion of extra-pair young in a nest. For each nest, the number of EPY was the event and the number of all offspring with determined paternity was the trial. The year and laying date in the original nest were added as covariates.

Presentation of results

The paternity effects on offspring performance are presented in the form of effect sizes. These effect sizes are the difference in survival probabilities between EPY and WPY and the standardized difference in mean value of traits measured on a continuous scale. The means and probabilities were the least square means adjusted for the covariates retained in the final models. The difference between means was standardized by the division of the standard deviation in the pooled groups. Confidence intervals for the difference in proportions were based on formulas given in Borenstein *et al.* (2009, p. 38) and those for the mean difference on formula 16 in Nakagawa & Cuthill (2007).

Results

Hatchability and hatching asynchrony

On an individual basis during 2006–2009, 1026 eggs coming from 160 nests (2006: 23 nests; 2007: 63; 2008: 44; 2009: 30) were cross-fostered between those nests. Four of these nests containing 23 eggs were abandoned during incubation, five eggs were accidentally broken and the hatching of another seven eggs was disallowed. Of the remaining 991 eggs, 937 hatched which gave us a hatchability rate of 94.6%. This figure is very close to hatchability in natural nests (94.0%) and nests in which whole clutches were cross-fostered (93.9%; Krist 2009) indicating that egg handling did not affect egg survival.

Most nests were visited daily around the time of presumed hatching. Between two subsequent visits, 78 of

146 (53.4%) nests completely hatched. This measure of hatching asynchrony is only a crude one and does not mean that nests which were only partially hatched at the time of our visit hatched asynchronously. Most of them probably hatched during a few hours, but we had accidentally checked them in the middle of this period. Based on this crude measure, hatching asynchrony was reduced in experimental nests compared to other nests in which eggs were left in nests during egg laying (57/280 = 20.4%). About 75% of females started continuous incubation of the dummy eggs before clutch completion (Fig. 1).

Distribution of extra-pair young

Genotypes at five or more loci were obtained for 941 offspring. Genotypes were lacking for 85 eggs for the following reasons: 39 chicks disappeared from the nest, 23 eggs did not show any sign of embryo development (labelled as infertile hereafter) and therefore were not genotyped (see Arnold *et al.* 2003), nine eggs disappeared from the nest, five eggs were accidentally broken by us, five samples were lost, and in four samples, DNA had degraded due to tissue decay.

The female parent was captured and genotyped in 143 of the experimental nests, while males were identified in 135. Furthermore, seven females and 209 males were captured and genotyped in non-experimental nests during the course of this experiment. In total, 1434 genotypes were obtained for 1235 individuals as some individuals were genotyped multiple times (see Methods).

The genotype of the social father was obtained for 800 of 941 genotyped offspring. Extra-pair fathers

Fig. 1 The distribution of the number of eggs laid in experimental nests after the start of continuous incubation. Null means that continuous incubation started after the laying of the last egg.

sired 214 of these offspring (26.8%). All young were sired by a social father in 58 nests, while in eight nests, all the young were sired by extra-pair males. Mixed paternity was identified in 69 nests (51.1%). In these nests, 253 young were sired by a social father and 168 by an extra-pair father (39.9%), and the paternity of 36 young was undetermined. Two social fathers were hybrids with the pied flycatcher, and in both of these cases, all the young were sired by extra-pair males.

Laying order, infertility and incubation

In mixed-paternity clutches, the proportion of extra-pair young decreased in laying order (Table 3). This pattern was similar for the actual and relative laying order (Fig. 2). EPY were found in eggs of the same size as WPY (Table 3).

In 16 nests, 23 infertile eggs were laid. The probability that a nest contained an infertile egg tended to be lower in nests where at least one young was sired by an extra-pair father (6/77 vs. 10/58, $\chi^2 = 2.83$, $P = 0.093$).

The proportion of EPY was not larger in nests where females started incubation later ($F_{1,63} = 0.40$, $P = 0.527$) and did not depend on the year ($F_{3,63} = 1.18$, $P = 0.318$) or the laying date ($F_{1,63} = 0.28$, $P = 0.600$).

Offspring performance

Five of the morphological and one of the immunological traits measured on nestlings were closely similar for EPY and WPY (Table 4, Fig. 3). These traits were measured on a large number of nestlings originating from mixed-paternity nests. Consequently, narrow confidence intervals excluded the possibility of large or even medium superiority of EPY at this stage. Using more long-term data, no significant effect of paternity was determined for any morphological, life history and secondary sexual ornamental traits measured on recruits (Table 5, Fig. 3). Sample sizes were smaller and therefore these effect sizes are less definite. However, point estimates were also generally small at this offspring stage (Fig. 3). Paternity was not a significant predictor of offspring survival during any of the three investigated life history stages (Table 4, Fig. 3). In sum, most effect sizes were close to zero. For a few traits, point estimates suggested medium or even large effects. However, these were invariable in a negative direction (i.e. WPY > EPY: forehead patch of males, recruitment probability). In contrast to the weak effect of paternity status, nestling morphology, immunity and survival varied between years and also depended on the time of the season (Table 4).
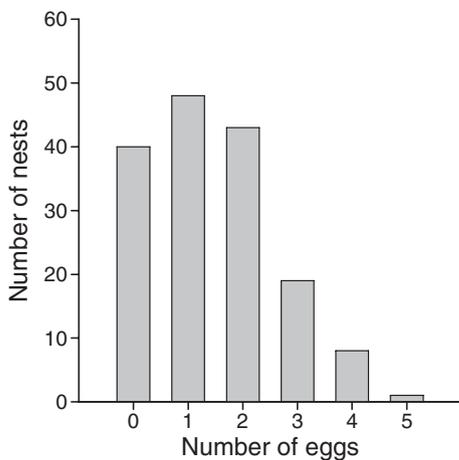
Table 3 The summary of models relating paternity to environmental and maternal variables in the nest of origin

		Absolute laying order					Relative laying order				
	NDF	DDF	Estimate	SE	F	P	DDF	Estimate	SE	F	P
Nest of origin			*0.277*	*0.178*				*0.278*	*0.178*		
Intercept			−3.86	5.34				−3.49	5.36		
Laying order	1	413	−0.227	0.0601	14.29	<0.001	413	−1.53	0.405	14.35	<0.001
Relative egg size	1	413	−0.000360	0.00166	0.05	0.828	413	−0.000350	0.00167	0.04	0.835
Laying date	1	65.5	0.0358	0.0433	0.68	0.412	65.6	0.0332	0.0434	0.58	0.448
Year	3	64.7			0.89	0.451	64.3			0.95	0.420

The random part of the models is in italics.
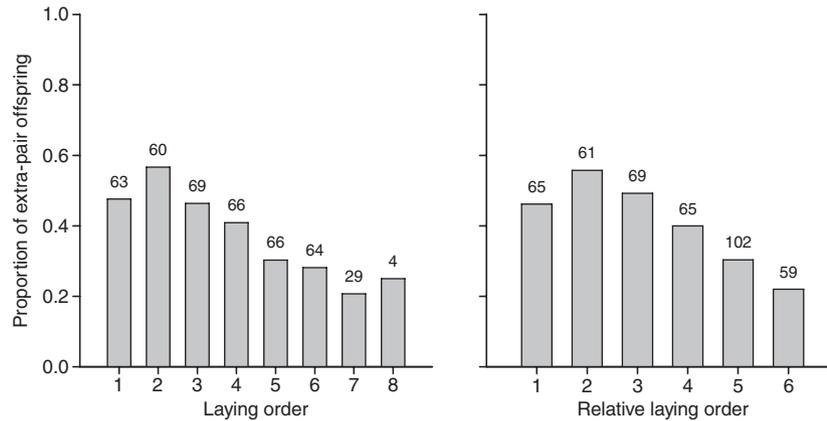
Fig. 2 The proportion of extra-pair young in mixed-paternity clutches in relation to laying order. Laying order is given as an absolute and relative value. Relative laying order is the actual laying order divided by the clutch size. The resulting continuous variable is separated into six categories to show the pattern in a clutch of modal size. The number of offspring in these categories was made approximately equal under the condition that the same value of relative laying order was not divided into two adjacent categories. The number of eggs with determined paternity is given above the bars.

Discussion

The proportion of EPY decreased with the laying order and 75% of the time females began to incubate before clutch completion, which usually led to hatching asynchrony of the brood. Therefore, under natural conditions, the occurrence of EPY in early eggs could give them a competitive advantage compared to later hatched WPY. Consequently, this maternal effect would confound an estimation of the genetic effect as might have been the case of our previous non-manipulative study (Krist *et al.* 2004). In that study, we were unable to control for hatching asynchrony when testing for the difference in survival probability between EPY and WPY. We found higher survival of EPY (Krist *et al.* 2004) that hatched from early eggs in the laying sequence (Krist *et al.* 2005), but in contrast to this result, we found no significant paternity effect on nestling morphology after hatching asynchrony was statistically controlled for (Krist *et al.* 2004).

In the present study, the confusing effect of hatching asynchrony was experimentally controlled for as females were forced to start incubation of all eggs at the same time. Extra-pair young did not outperform WPY in any studied trait in this large set of nests, suggesting that females do not obtain genetic benefits from extra-pair copulations. However, several issues need to be considered before ruling out this genetic-benefits explanation.

No genetic benefits from extra-pair paternity?

The usual approach to study genetic benefits of extra-pair paternity (EPP) is to compare performance of half-siblings that are raised in the same nest. However, in our study, half-siblings were raised in different nests, and we included nest of origin as a random factor in all analyses. Consequently, our analyses are similar in statistical principle to the usually employed paired tests (see West *et al.* 2007). However, as half-siblings were

Table 4 The dependence of offspring traits on paternity and two environmental factors

(A) Nestling morphology and immunity

	Body mass at 6 days			Body mass at 13 days			Tarsus length			Condition			Wing length			PHA immunity		
	DDF	F/estimate	P/SE	DDF	F/estimate	P/SE	DDF	F/estimate	P/SE	DDF	F/estimate	P/SE	DDF	F/estimate	P/SE	DDF	F/estimate	P/SE
Nest of origin		*0.219*	*0.0888*		*0.122*	*0.0654*		*0.0678*	*0.0224*		*0.0721*	*0.0448*		*0.0221*	*0.00625*		*0.00344*	*0.00242*
Nest of rearing		*1.14*	*0.225*		*0.910*	*0.176*		*0.149*	*0.0340*		*0.498*	*0.108*		*0.108*	*0.0167*		*0.00179*	*0.00217*
Residual		*0.543*	*0.0844*		*0.783*	*0.0912*		*0.150*	*0.0181*		*0.615*	*0.0708*		*0.0237*	*0.00307*		*0.0168*	*0.00319*
Paternity	152	0.03	0.856	250	<0.01	0.977	235	<0.01	0.977	265	<0.01	0.991	173	0.01	0.927	98	0.01	0.909
Laying date	102	2.12	0.148	121	2.84	0.095	129	8.55	0.004	114	0.02	0.884	138	16.82	<0.001	44	0.34	0.561
Year	95	23.87	<0.001	114	9.02	<0.001	108	6.75	<0.001	104	3.88	0.011	133	3.95	0.010	42	3.45	0.070

(B) Offspring survival

	Hatchability			Fledging success			Recruitment probability		
	DDF	F/estimate	P/SE	DDF	F/estimate	P/SE	DDF	F/estimate	P/SE
Nest of origin		–	–		*0*	–		*0*	–
Nest of rearing		–	–		*1.08*	*0.608*		*0.434*	*0.370*
Paternity	403	0.30	0.583	343	0.02	0.892	315	3.68	0.056
Laying date	403	0.34	0.562	175	1.91	0.169	124	3.62	0.059
Year	403	0.95	0.389	128	3.41	0.020	209	4.81	0.003

The analyses were restricted to young originating from mixed-paternity broods. Random variables and their associated values are given in italics; fixed variables and their associated values are given in common type. Values associated with random effects are variance estimates and standard errors. Values associated with fixed factors are the results of type III F-tests: DDF, denominator degrees of freedom; F, F-value; P, P-value. Statistical direction of paternity effect is given in Fig. 3. Laying date means the date of the first egg in the nest of rearing. Year was fitted as a categorical variable.

raised in different nests, they also experienced different competitive environments. Thus, we introduced noise in our data that was caused by the variable quality of rearing environments that possibly decreased the power of our test. However, this noise is unlikely to mask genetic superiority of EPY for the following reasons. First, we gathered a large sample size that, on the contrary, increased the statistical power of our tests. Our design introduced most of the noise into the morphological traits of nestlings as indicated by the high proportion of variance (40–70%) explained by the random factor 'nest of rearing' (see Table 4). These values were considerably smaller for nestling immunity (8%) and traits of recruits (0–31%) with the highest value seen for nestling wing length (70%). This test was performed on 193 WPY and 128 EPY from 69 mixed-paternity broods. If our test of paternity effects utilized only the remaining 30% of variation in wing length, its power might be comparable to the conventional test that utilizes all the variation but with a sample size 30% of our values. This sample size would be 58 WPY and 38 EPY from 21 mixed-paternity broods which are still reasonable values and comparable to many previous studies. Second, the non-significance of our tests was not caused by wide confidence intervals that would overlap zero despite the point estimates being positive. In fact, the point estimates of effect sizes were actually close to zero or negative (Fig. 3). The confidence intervals were narrow at the nestling stage while wider for traits measured on recruits. Therefore, we can safely conclude that no genetic benefits are manifested during the nestling stage. Although there was also no indication of positive genetic effects later in life, as the respective point estimates were generally negative, this possibility cannot be completely ruled out as the power of these tests was relatively small.

Apart from the limitation of increased noise in data, our design has two advantages over former approaches. First, it enabled us to control for hatching asynchrony and egg size that might otherwise cause bias in tests of genetic benefits (Magrath *et al.* 2009; Ferree *et al.* 2010). Second, eggs originating from 69 mixed-paternity clutches were distributed to 156 nests. If genetic benefits were context-dependent, the higher number of raising contexts would enhance estimation of the average effect of paternal genes which matters most for selection on female behaviour. To conclude, if mixed cross-fostering is performed with a high sample size, it represents a vital alternative to conventional approaches for estimation of paternal genetic benefits.

As noted earlier, paternal genetic effects may be context-dependent. For example, they may only be manifested under poor environmental conditions (Neff &
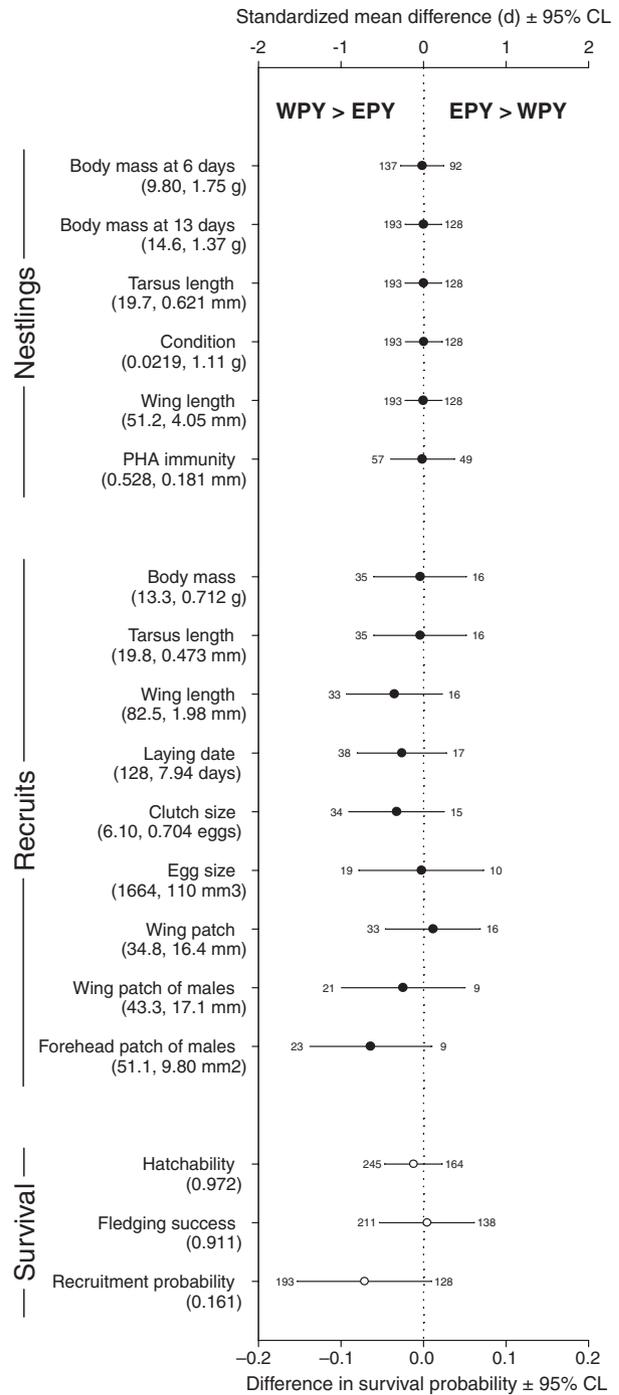

Fig. 3 The differences in performance of extra-pair and within-pair young originating from mixed-paternity nests. The number of WPY is given on the left and the number of EPY on the right of the bar. Numbers below trait labels are the mean and SD of the respective trait. Both statistics were computed on the pooled sample of all WPY and EPY from all nests. The presented SD was used for the standardization of the difference between means which is represented by filled symbols. Open symbols refer to the difference in survival probabilities.

Table 5 The dependence of recruit traits on paternity and a range of covariates

(A) Morphological and life history traits

	Body mass			Tarsus length			Wing length			Laying date			Clutch size			Egg size		
	DDF	F/estimate	P/SE	DDF	F/estimate	P/SE	DDF	F/estimate	P/SE	DDF	F/estimate	P/SE	DDF	F/estimate	P/SE	DDF	F/estimate	P/SE
Nest of origin		*0.244*	*0.189*		*0.125*	*0.0606*		*0.00247*	*0.00419*		*14.3*	*16.2*		*0.230*	*0.165*		*3503*	*4336*
Nest of rearing		*0.0230*	*0.0818*		*0.0742*	*0.0507*		*0.00203*	*0.00956*		*13.6*	*17.1*		*0.190*	*0.272*		*626*	*2752*
Residual		*0.216*	*0.144*		*0.0408*	*0.0309*		*0.0154*	*0.0102*		*43.5*	*21.8*		*0.242*	*0.302*		*5113*	*4703*
Paternity	23	**0.01**	0.905	35	**0.03**	**0.866**	38	**2.00**	0.165	52	**0.64**	0.428	39	**0.74**	**0.396**	17	**<0.01**	**0.948**
Laying date	28	0.45	0.506	37	**3.72**	**0.061**	20	2.28	0.147	30	0.01	0.929	34	1.30	0.263	8	<0.01	0.957
Year of hatching	37	0.50	0.609	38	**3.14**	**0.036**	32	**6.07**	**0.002**	39	0.10	0.909	33	0.08	0.927	21	2.48	0.109
Year of breeding	20	1.83	0.174	7	0.01	0.994	29	**10.70**	**<0.001**	44	2.17	0.106	31	1.03	0.392	12	1.57	0.248
Age	24	**4.69**	**0.010**	25	1.89	0.158	34	0.45	0.641	43	1.04	0.383	33	1.48	0.239	17	0.36	0.702
Sex	26	**7.72**	**0.010**	7	1.74	0.229	37	**2.96**	**0.094**	39	<0.01	0.996	45	1.27	0.266	20	2.15	0.158
Offspring age	38	1.60	0.213															

(B) Ornamental traits

	Wing patch			Wing patch of males			Forehead patch of males		
	DDF	F/estimate	P/SE	DDF	F/estimate	P/SE	DDF	F/estimate	P/SE
Nest of origin		*21.8*	*34.7*		*0*	*-*		*59.7*	*23.7*
Nest of rearing		*19.5*	*41.4*		*0*	*-*		*14.0*	*10.5*
Residual		*64.1*	*46.7*		*62.8*	*17.8*		*0.445*	*-*
Paternity	30	**0.25**	0.619	25	**1.50**	0.232	8	**4.56**	**0.063**
Laying date	17	0.53	0.478	24	0.15	0.705	22	0.80	0.382
Year of hatching	29	**2.83**	**0.055**	19	2.32	0.107	20	0.74	0.491
Year of breeding	24	0.28	0.761	7	1.03	0.407	8	2.54	0.128
Age	32	**10.73**	**<0.001**	25	**34.12**	**<0.001**	9	**6.27**	**0.013**
Sex	34	**28.47**	**<0.001**						

The analyses were restricted to young originating from mixed-paternity broods. Random variables and their associated values are given in italics. Fixed variables and their associated values are given in common type if removed from the final model or in bold type if they remained in the final model. Values presented for removed covariates relate to the step just before their removal. Values associated with random effects are variance estimates and standard errors. Values associated with fixed factors are the results of type III *F*-tests: DDF, denominator degrees of freedom; *F*, *F*-value; *P*, *P*-value. Statistical direction of paternity effect is given in Fig. 3. Laying date means the date of the first egg in the nest of rearing. Years and age of recruits were fitted as categorical variables.

Pitcher 2005; Schmoll *et al.* 2005; Garvin *et al.* 2006). Consequently, studies limited to only good environmental conditions may not reveal any benefit of EPY over their half-siblings. This should not have affected our study as data were collected over 4 years that differed greatly in environmental conditions, as evidenced by the strong effect of year on offspring performance. Furthermore, offspring were raised under variable conditions within years as evidenced by the significant effect of laying date on their quality/survival. There were also four factors associated with our experimental design that might have changed the quality of the environment from the offspring's perspective. First, offspring were raised by foster parents, and there may be a coadaptation between offspring demands and parental supply (Wolf & Brodie 1998). If food provisioning is under parental control, as is likely in the collared flycatcher (Ottosson *et al.* 1997), cross-fostering could lead to a negative effect on offspring (Hinde *et al.* 2010). Second, broods were composed of non-related nest-mates which might induce higher sibling competition (Boncoraglio & Saino 2008). Third, eggs were returned to the nests a day after clutch completion which likely increased the energy spent for incubation and possibly reduced the females' capacity for chick provisioning. Finally, synchronous hatching reduced sibling hierarchy which might lead to increased sibling competition (Stoleson & Beissinger 1995; Kontiainen *et al.* 2010; but see Szöllősi *et al.* 2007). In sum, a mixed cross-fostering approach might have created a relatively competitive offspring environment. However, these subtle mechanisms, if indeed operating, should be of relatively minor significance compared to large annual and seasonal effects. Overall, maternal half-siblings were exposed to varying environmental conditions in this study, and context dependence therefore cannot account for the lack of difference between EPY and WPY.

Indirect genetic benefits should ideally be tested by a comparison of true fitness between maternal half-siblings. In a detailed long-term study of song sparrows (Sardell *et al.* 2011), female EPY were found to have lower recruitment and lifespan than WPY. Despite this, the greater fitness of EPY cannot be ruled out (Sardell *et al.* 2011) as survival may be traded for reproductive success (Hunt *et al.* 2004; Head *et al.* 2005). Therefore, reproductive success should be taken into account when looking for the genetic benefits of EPC (Eliassen & Kokko 2008). In this study, we focused on both survival and several indicators of reproductive success of WPY and EPY. Survival, life history traits (clutch size, egg size, breeding time) and secondary sexual traits (forehead and wing patches) were the same or smaller for EPY, which suggests that EPY do not outperform WPY in terms of either survival or reproductive suc-cess. Dependence of reproductive success on offspring paternity status was previously tested in a single species, the coal tit *Parus ater* (Schmoll *et al.* 2003, 2009), which found that EPY produced less offspring in their lifetime than WPY (Schmoll *et al.* 2009). One caveat of both the studies on coal tits and our study on flycatchers is a lack of data on extra-pair success of male offspring which is a component of reproductive success that has been shown to be a very important part of male fitness (Webster *et al.* 1995; Albrecht *et al.* 2007; Vedder *et al.* 2011). Without taking it into account, no definite conclusion about fitness benefits of EPC can be drawn (Schmoll *et al.* 2009). In this study, one indirect finding suggests that male EPY may not have increased access to extra-pair mates. We found that EPY offspring had the same size, if not smaller, secondary sexual ornament that is important for the female choice of both a social (Qvarnström *et al.* 2000; Hegyi *et al.* 2010) and an extra-pair mate (Sheldon & Ellegren 1999; Michl *et al.* 2002). As we did not find any indication for the genetic benefit of EPC, other explanations for involvement of females in this behaviour should be explored.

Other explanations for female infidelity

Females can sometimes obtain material benefits from extra-pair mates. For example, in great grey shrikes, *Lanius excubitor*, extra-pair males gave valuable courtship gifts to females before copulation (Tryjanowski & Hromada 2005). These gifts were mainly vertebrate prey items that represented up to 65% of the female daily energy expenditure. Female red-winged blackbirds, *Agelaius phoeniceus*, were allowed to forage on territories of neighbouring males with whom they had previously copulated (Gray 1997). In cooperatively breeding American crows, *Corvus brachyrhynchos*, extra-pair sires contribute to offspring provisioning which increased the offspring's survival rate (Townsend *et al.* 2010). Similar benefits are likely to be small or absent in the collared flycatcher. Males in this species do not defend food territories (Cramp & Perrins 1993) and although courtship feeding exists (Cramp & Perrins 1993), presented invertebrate prey probably does not cover a large proportion of the female daily energy expenditure. We also did not observe offspring provisioning by two males. Therefore, large material benefits from extra-pair copulations seem to be limited to special cases such as cooperative breeding (Townsend *et al.* 2010).

Females might also copulate with extra-pair mates to ensure against potential male infertility (Sheldon 1994; Hasson & Stone 2009). Recently, azoospermia, which is a complete lack of sperm in ejaculates, was found in 2–4% of males in two passerine species (Lifjeld *et al.* 2007). Other forms of functional male infertility such as

oligospermia and asthenozoospermia were not investigated in that study. Similarly to azoospermia, these latter forms of male infertility may also contribute to female infidelity (Hasson & Stone 2009). Indirect evidence for male infertility and female insurance against it comes from studies that found an excess of broods with all offspring sired by extra-pair males (Krokene et al. 1998). This is also the case of our study where the proportion of these broods was relatively high (5.9%). Our observation that the number of EPY decreased in the laying order is consistent with the fertility insurance hypothesis. A single extra-pair copulation early in the female's fertile period could ensure the whole clutch against the potential infertility of a social mate. If a social mate was fertile, extra-pair young should disappear in later eggs as a result of subsequent inseminations from the social mate. We also found a tendency for less infertile eggs in broods with EPY which could tentatively suggest that EPC did indeed decrease the proportion of infertile eggs. However, there is a problem with this hypothesis as the opposite pattern (i.e. more infertile eggs in broods with EPY) also may be interpreted as support for fertility insurance (Wetton & Parkin 1991). One factor causing the conflicting predictions may be whether selection is currently operating or whether the population is already in equilibrium (Sheldon 1994). The second reason may be the ability of females to assess male infertility based on their appearance. If females were able to recognize the fertility of a social partner, only those with infertile partners could copulate with extra-pair males. In this case, a positive association between cuckoldry and infertility could arise. If females did not adjust their copulating behaviour to their partner's fertility, a negative association could arise. Recently, a positive correlation between sperm and plumage quality was found in the pied flycatcher, *Ficedula hypoleuca* (Calhim et al. 2009). This would suggest that females might also assess sperm quality in the collared flycatcher and therefore a positive correlation between infertility and EPY would be expected. Surely, more research is needed to elucidate whether there are fertility benefits of EPC in the collared flycatcher where special attention should be paid to sperm traits and their associations with the male phenotype. The data in hand suggest that fertility insurance might be more important for the evolution of EPC behaviour in the collared flycatcher than indirect genetic benefits.

Alternatively, females might not accrue any benefits from EPC and instead this behaviour might be driven by strong selection on male behaviour (Westneat & Stewart 2003; Arnqvist & Kirkpatrick 2005; Forstmeier et al. 2011). Males might either force or seduce (Holland & Rice 1998) females to behave in a maladaptive way.

Behavioural observations on who initiates extra-pair copulations would be very useful in assessing this hypothesis (Kempenaers & Schlicht 2010). Unfortunately, it would be tremendously difficult to conduct such observations of the collared flycatcher in the field as both within-pair and extra-pair copulations are probably rare events (Michl et al. 2002). Moreover, female extra-pair behaviour might be genetically correlated with that of males, and females may not benefit from EPP even if they actively seek EPC (Forstmeier et al. 2011). This intriguing possibility remains to be tested in the wild.

Extra-pair young and maternal effects

In contrast to no evidence for paternal genetic effects, we have documented non-random distribution of EPY in laying order. Under natural conditions, hatching is partially asynchronous in the collared flycatcher (Krist et al. 2004; Rosivall et al. 2005) owing to incubation beginning before clutch completion, as was documented in this study. Consequently, under natural conditions, EPY would appear earlier in the hatching order than WPY. As the hatching order is a strong determinant of offspring performance (Krist et al. 2004; Rosivall et al. 2005), this would likely lead to non-genetic superiority of EPY. Recently, studies on three other passerine species found a predominance of EPY early in the laying or hatching order that resulted in their better growth (Johnson et al. 2009; Magrath et al. 2009; Ferree et al. 2010). After this maternal effect was statistically controlled for, differences between half-sibs disappeared (Magrath et al. 2009; Ferree et al. 2010). We controlled hatching asynchrony experimentally, and similar to those previous studies, we did not find any genetic effects on offspring quality. The quality of the young is considered to be the only ultimate test for genetic benefits of EPC (Hasson & Stone 2011; see also Puurtinen et al. 2009). Consequently, without the control for potential maternal effects, no strong conclusions about genetic benefits of EPC can be drawn (Kempenaers & Schlicht 2010). In the few studies that controlled for hatching asynchrony, genetic effects would be overestimated without this control. This was caused by the occurrence of EPY in early eggs. However, the distribution of EPY in laying order may differ both between and within species. For example, EPY occurred in early eggs from two independent samples obtained in our population in different years (Krist et al. 2005 and this study), while no such effect was found in Polish (Wilk et al. 2008) and Hungarian (Rosivall et al. 2009) populations of the species. If EPY occurred in late eggs (Riley et al. 1995), this could lead to an underestimation of paternal genetic effects owing to the counteractive effect

of hatching asynchrony. Thus, the emerging view that the genetic benefits from EPC are small at best (Arnqvist & Kirkpatrick 2005; Akçay & Roughgarden 2007; Edler & Friedl 2008; Schmoll et al. 2009; Sardell et al. 2011) has no strong support as few previous studies have controlled for the maternal effects within broods.

The non-random distribution of EPY in mixed-paternity clutches probably arises as a result of the non-random distribution of copulations with social and extra-pair partners for which there are several potential explanations. First, it might be adaptive for females to put EPY in early or late eggs to facilitate or decrease the genetic-driven superiority of these offspring (i.e. differential allocation or compensation: Sheldon 2000b; Gowaty 2008; Horváthová et al. 2011). This explanation is unlikely for our system as EPY were not genetically superior. The second adaptive explanation may be that females copulate with an extra-pair partner before the onset of laying to ensure the whole clutch against the infertility of a social mate. Some pieces of evidence tentatively support this hypothesis. Other reasons for the non-random distribution of EPY in the laying order may be constraints imposed on female copulation behaviour. For example, if males guard their mates less in early fertile periods, females might seek extra-pair copulations mainly during this time. However, the opposite pattern of mate guarding is usually found (Westneat 1993; Pinxten & Eens 1997). Recently, another physiological mechanism that could potentially explain the excess of EPY in early eggs was proposed (Vedder et al. 2010). In that study, the addition of eggs into nests before the onset of laying induced a higher intensity of incubation and decreased the proportion of EPY in broods. This suggests that the female motivation to seek EPC decreases as the intensity of incubation increases, potentially leading to more EPY in early eggs compared to late ones (Vedder et al. 2010). Our correlative data did not support this prediction as the number of EPY was independent of female incubation during the egg laying stage. Thus, our data are consistent only with a scenario of the female actively seeking EPC for fertility insurance. However, more data on the behaviour of all three participants (female, social male and extra-pair male) are needed to elucidate the mechanisms leading to the excess of EPY in early eggs.

So far we have discussed only the non-random distribution of EPY in the laying order as a maternal effect that may confound the estimation of the genetic benefits from EPC. However, there are other maternal effects that might operate within broods. EPY might come from larger eggs, eggs with a higher concentration of hormones or carotenoids, or may be more provisioned by the parents. The last effect seems to be improbable as parents are apparently unable to recognize their own offspring (Westneat et al. 1995; Kempenaers & Sheldon 1996). Egg size (Slagsvold et al. 1984), yolk hormones (Gil 2008) and carotenoids (e.g. Saino et al. 2002; Cassey et al. 2005; Newbrey et al. 2008) often systematically vary with laying order. If the occurrence of EPY also systematically varies in the laying order, as shown in this study, it is likely that EPY may differ in egg size or composition from WPY. If such an association between prenatal nutrition and paternity was driven solely by laying order, it would be sufficient to control for this variable when comparing the performance of half-siblings. The finer targeting of nutrients based on paternity of ova seems to be unlikely (see Birkhead et al. 2000). Indeed, we found no association between egg size and paternity after the laying order had been controlled for. Although egg size (Krist 2011), yolk hormones (Gil 2008) and carotenoids (Saino et al. 2003; Biard et al. 2005; but see Remeš et al. 2007) affect offspring quality, their effect is likely to be small compared to the effect of the hatching order (Krist et al. 2004; Maddox & Weatherhead 2008; Mainwaring et al. 2010). To conclude, the hatching and/or laying order are the two maternal effects that most urgently need to be controlled for when assessing genetic benefits of EPCs via the comparison of half-siblings.

Acknowledgements

References

Akçay E, Roughgarden J (2007) Extra-pair paternity in birds: review of the genetic benefits. *Evolutionary Ecology Research*, **9**, 855–868.

Albrecht T, Schnitzer J, Kreisinger J et al. (2007) Extrapair paternity and the opportunity for sexual selection in long-distant migratory passerines. *Behavioral Ecology*, **18**, 477–486.

Arnold KE, Orr KJ, Griffiths R (2003) Primary sex ratios in birds: problems with molecular sex identification of undeveloped eggs. *Molecular Ecology*, **12**, 3451–3458.

Arnqvist R, Kirkpatrick M (2005) The evolution of infidelity in socially monogamous passerines: the strength of direct and indirect selection on extrapair copulation behavior in females. *American Naturalist*, **165**, S26–S37.

Biard C, Surai PF, Møller AP (2005) Effects of carotenoid availability during laying on reproduction in the blue tit. *Oecologia*, **144**, 32–44.

Birkhead TR, Schwabl H, Burke T (2000) Testosterone and maternal effects: integrating mechanisms and function. *Trends in Ecology & Evolution*, **15**, 86–87.

Boncoraglio G, Saino N (2008) Barn swallow chicks beg more loudly when broodmates are unrelated. *Journal of Evolutionary Biology*, **21**, 256–262.

Borenstein M, Hedges LV, Higgins JPT, Rothstein HR (2009) *Introduction to Meta-Analysis*. John Wiley & Sons Ltd, Chichester.

Brown MB, Brown CR (2009) Blood sampling reduces annual survival in cliff swallows (*Petrochelidon pyrrhonota*). *Auk*, **126**, 853–861.

Calhim S, Lampe HM, Slagsvold T, Birkhead TR (2009) Selection on sperm morphology under relaxed sperm competition in a wild passerine bird. *Biology Letters*, **5**, 58–61.

Cassey P, Ewen JG, Boulton RL *et al.* (2005) Egg carotenoids in passerine birds introduced to New Zealand: relations to ecological factors, integument coloration and phylogeny. *Functional Ecology*, **19**, 719–726.

Clark AB, Wilson DS (1981) Avian breeding adaptations: hatching asynchrony, brood reduction, and nest failure. *Quarterly Review of Biology*, **56**, 253–277.

Cook MI, Beissinger SR, Toranzos GA, Rodriguez RA, Arendt WJ (2003) Trans-shell infection by pathogenic micro-organisms reduces the shelf life of non-incubated bird's eggs: a constraint on the onset of incubation? *Proceedings of the Royal Society of London Series B-Biological Sciences*, **270**, 2233–2240.

Cordero PJ, Wetton JH, Parkin DT (1999) Within-clutch patterns of egg viability and paternity in the House Sparrow. *Journal of Avian Biology*, **30**, 103–107.

Cordero PJ, Griffith SC, Aparicio JM, Parkin DT (2000) Sexual dimorphism in house sparrow eggs. *Behavioral Ecology and Sociobiology*, **48**, 353–357.

Cramp S, Perrins CM (1993) *The Birds of the Western Palearctic*. Oxford University Press, Oxford.

D'Alba L, Shawkey MD, Korsten P *et al.* (2010) Differential deposition of antimicrobial proteins in blue tit (*Cyanistes caeruleus*) clutches by laying order and male attractiveness. *Behavioral Ecology and Sociobiology*, **64**, 1037–1045.

Edler R, Friedl TWP (2008) Within-pair young are more immunocompetent than extrapair young in mixed-paternity broods of the red bishop. *Animal Behaviour*, **75**, 391–401.

Eliassen S, Kokko H (2008) Current analyses do not resolve whether extra-pair paternity is male or female driven. *Behavioral Ecology and Sociobiology*, **62**, 1795–1804.

Ellegren H (1992) Polymerase-chain-reaction (PCR) analysis of microsatellites: a new approach to studies of genetic relationships in birds. *Auk*, **109**, 886–895.

Ferree ED, Dickinson J, Rendell W, Stern C, Porter S (2010) Hatching order explains an extrapair chick advantage in western bluebirds. *Behavioral Ecology*, **21**, 802–807.

Forstmeier W, Martin K, Bolund E, Schielzeth H, Kempenaers B (2011) Female extrapair mating behavior can evolve via indirect selection on males. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 10608–10613.

Garamszegi LZ, Rosivall B, Hegyi G *et al.* (2006) Determinants of male territorial behavior in a Hungarian collared flycatcher population: plumage traits of residents and challengers. *Behavioral Ecology and Sociobiology*, **60**, 663–671.

Garvin JC, Abroe B, Pedersen MC, Dunn PO, Whittingham LA (2006) Immune response of nestling warblers varies with extra-pair paternity and temperature. *Molecular Ecology*, **15**, 3833–3840.

Gibbs HL, Tabak LM, Hobson K (1999) Characterization of microsatellite DNA loci for a neotropical migrant songbird, the Swainson's thrush (*Catharus ustulatus*). *Molecular Ecology*, **8**, 1551–1561.

Gil D (2008) Hormones in avian eggs: physiology, ecology and behavior. *Advances in the Study of Behavior*, **38**, 337–398.

Gowaty PA (2008) Reproductive compensation. *Journal of Evolutionary Biology*, **21**, 1189–1200.

Gray EM (1997) Female red-winged blackbirds accrue material benefits from copulating with extra-pair males. *Animal Behaviour*, **53**, 625–639.

Griffith SC, Owens IPF, Thuman KA (2002) Extra pair paternity in birds: a review of interspecific variation and adaptive function. *Molecular Ecology*, **11**, 2195–2212.

Gustafsson L (1989) Collared Flycatcher. In: *Lifetime Reproduction in Birds* (ed. Newton I), pp. 75–88. Academic Press, London.

Gustafsson L, Sutherland WJ (1988) The costs of reproduction in the collared flycatcher *Ficedula albicollis*. *Nature*, **335**, 813–815.

Hasson O, Stone L (2009) Male infertility, female fertility and extrapair copulations. *Biological Reviews*, **84**, 225–244.

Hasson O, Stone L (2011) Why do females have so few extra-pair offspring? *Behavioral Ecology and Sociobiology*, **65**, 513–523.

Head ML, Hunt J, Jennions MD, Brooks R (2005) The indirect benefits of mating with attractive males outweigh the direct costs. *PLoS Biology*, **3**, 289–294.

Hegyi G, Szöllősi E, Jenni-Eiermann S *et al.* (2010) Nutritional correlates and mate acquisition role of multiple sexual traits in male collared flycatchers. *Naturwissenschaften*, **97**, 567–576.

Hinde CA, Johnstone RA, Kilner RM (2010) Parent-offspring conflict and coadaptation. *Science*, **327**, 1373–1376.

Holland B, Rice WR (1998) Chase-away sexual selection: antagonistic seduction versus resistance. *Evolution*, **52**, 1–7.

Hořák D, Albrecht T (2007) Using net sacks to examine the relationship between egg size and young size in Common Pochards. *Journal of Field Ornithology*, **78**, 334–339.

Hõrak P, Mänd R, Ots I (1997) Identifying targets of selection: a multivariate analysis of reproductive traits in the great tit. *Oikos*, **78**, 592–600.

Horváthová T, Nakagawa S, Uller T (2011) Strategic female reproductive investment in response to male attractiveness in birds. *Proceedings of the Royal Society of London Series B-Biological Sciences*, DOI: 10.1098/rspb.2011.0663.

Hoyt DF (1979) Practical methods of estimating volume and fresh weight of bird eggs. *Auk*, **96**, 73–77.

Hunt J, Bussiere LF, Jennions MD, Brooks R (2004) What is genetic quality? *Trends in Ecology & Evolution*, **19**, 329–333.

Jennions MD, Petrie M (2000) Why do females mate multiply? A review of the genetic benefits *Biological Reviews*, **75**, 21–64.

Johnson LS, Brubaker JL, Johnson BGP, Masters BS (2009) Evidence for a maternal effect benefiting extra-pair offspring in a songbird, the house wren *Troglodytes aedon*. *Journal of Avian Biology*, **40**, 248–253.

Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping

error increases success in paternity assignment. *Molecular Ecology*, **16**, 1099–1106.

Kempenaers B (2007) Mate choice and genetic quality: a review of the heterozygosity theory. *Advances in the Study of Behavior*, **37**, 189–278.

Kempenaers B, Schlicht E (2010) Extra-pair behaviour. In: *Animal Behaviour: Evolution and Mechanisms* (ed. Kappeler P), pp. 359–411. Springer, Heidelberg.

Kempenaers B, Sheldon BC (1996) Why do male birds not discriminate between their own and extra-pair offspring? *Animal Behaviour*, **51**, 1165–1173.

Kempenaers B, Verheyen GR, Dhondt AA (1997) Extrapair paternity in the blue tit (*Parus caeruleus*): female choice, male characteristics, and offspring quality. *Behavioral Ecology*, **8**, 481–492.

Kim M, Furness RW, Nager RG (2010) Hatching asynchrony is constrained by parental nest attendance during laying. *Behavioral Ecology and Sociobiology*, **64**, 1087–1097.

Kontiainen P, Pietiainen H, Karell P, Pihlaja T, Brommer JE (2010) Hatching asynchrony is an individual property of female Ural owls which improves nestling survival. *Behavioral Ecology*, **21**, 722–729.

Krist M (2009) Short- and long-term effects of egg size and feeding frequency on offspring quality in the collared flycatcher (*Ficedula albicollis*). *Journal of Animal Ecology*, **78**, 907–918.

Krist M (2011) Egg size and offspring quality: a meta-analysis in birds. *Biological Reviews*, **86**, 692–716.

Krist M, Remeš V, Uvírová L, Nádvorník P, Bureš S (2004) Egg size and offspring performance in the collared flycatcher (*Ficedula albicollis*): a within-clutch approach. *Oecologia*, **140**, 52–60.

Krist M, Nádvorník P, Uvírová L, Bureš S (2005) Paternity covaries with laying and hatching order in the collared flycatcher *Ficedula albicollis*. *Behavioral Ecology and Sociobiology*, **59**, 6–11.

Krokene C, Rigstad K, Dale M, Lifjeld JT (1998) The function of extrapair paternity in blue tits and great tits: good genes or fertility insurance? *Behavioral Ecology*, **9**, 649–656.

Kruuk LEB, Merilä J, Sheldon BC (2001) Phenotypic selection on a heritable size trait revisited. *American Naturalist*, **158**, 557–571.

Leder EH, Karaiskou N, Primmer CR (2008) Seventy new microsatellites for the pied flycatcher, *Ficedula hypoleuca* and amplification in other passerine birds. *Molecular Ecology Resources*, **8**, 874–880.

Lifjeld JT, Laskemoen T, Fossøy F, Johnsen A, Kleven O (2007) Functional infertility among territorial males in two passerine species, the willow warbler *Phylloscopus trochilus* and the bluethroat *Luscinia svecica*. *Journal of Avian Biology*, **38**, 267–272.

Maddox JD, Weatherhead PJ (2008) Egg size variation in birds with asynchronous hatching: is bigger really better? *American Naturalist*, **171**, 358–365.

Magrath RD (1990) Hatching asynchrony in altricial birds. *Biological Reviews*, **65**, 587–622.

Magrath MJL, Vedder O, van der Velde M, Komdeur J (2009) Maternal effects contribute to the superior performance of extra-pair offspring. *Current Biology*, **19**, 792–797.

Mainwaring MC, Dickens M, Hartley IR (2010) Environmental and not maternal effects determine variation in offspring phenotypes in a passerine bird. *Journal of Evolutionary Biology*, **23**, 1302–1311.

Mays HL, Albrecht T, Liu M, Hill GE (2008) Female choice for genetic complementarity in birds: a review. *Genetica*, **134**, 147–158.

Merilä J, Kruuk LB, Sheldon BC (2001) Cryptic evolution in a wild bird population. *Nature*, **412**, 76–79.

Michl G, Török J, Griffith SC, Sheldon BC (2002) Experimental analysis of sperm competition mechanisms in a wild bird population. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 5466–5470.

Moreno J, Merino S, Sanz JJ *et al.* (2005) Nestling cell-mediated immune response, body mass and hatching date as predictors of local recruitment in the pied flycatcher *Ficedula hypoleuca*. *Journal of Avian Biology*, **36**, 251–260.

Nakagawa S, Cuthill IC (2007) Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biological Reviews*, **82**, 591–605.

Neff BD, Pitcher TE (2005) Genetic quality and sexual selection: an integrated framework for good genes and compatible genes. *Molecular Ecology*, **14**, 19–38.

Newbrey JL, Reed WL, Foster SP, Zander GL (2008) Laying-sequence variation in yolk carotenoid concentrations in eggs of Yellow-headed Blackbirds (*Xanthocephalus xanthocephalus*). *Auk*, **125**, 124–130.

O'Brien EL, Dawson RD (2007) Context-dependent genetic benefits of extra-pair mate choice in a socially monogamous passerine. *Behavioral Ecology and Sociobiology*, **61**, 775–782.

Ottosson U, Backman J, Smith HG (1997) Begging affects parental effort in the pied flycatcher, *Ficedula hypoleuca*. *Behavioral Ecology and Sociobiology*, **41**, 381–384.

Pinxten R, Eens M (1997) Copulation and mate-guarding patterns in polygynous European starlings. *Animal Behaviour*, **54**, 45–58.

Potti J (1998) Variation in the onset of incubation in the pied flycatcher (*Ficedula hypoleuca*): fitness consequences and constraints. *Journal of Zoology*, **245**, 335–344.

Przybylo R, Sheldon BC, Merilä J (2000) Patterns of natural selection on morphology of male and female collared flycatchers (*Ficedula albicollis*). *Biological Journal of the Linnean Society*, **69**, 213–232.

Puurtinen M, Ketola T, Kotiaho JS (2009) The Good-genes and compatible-genes benefits of mate choice. *American Naturalist*, **174**, 741–752.

Qvarnström A, Pärt T, Sheldon BC (2000) Adaptive plasticity in mate preference linked to differences in reproductive effort. *Nature*, **405**, 344–347.

Remeš V, Krist M, Bertacche V, Stradi R (2007) Maternal carotenoid supplementation does not affect breeding performance in the Great Tit (*Parus major*). *Functional Ecology*, **21**, 776–783.

Riley HT, Bryant DM, Carter RE, Parkin DT (1995) Extra-pair fertilizations and paternity defense in house martins, *Delichon urbica*. *Animal Behaviour*, **49**, 495–509.

Rosivall B, Szöllősi E, Török J (2005) Maternal compensation for hatching asynchrony in the collared flycatcher *Ficedula albicollis*. *Journal of Avian Biology*, **36**, 531–537.

Rosivall B, Szöllősi E, Hasselquist D, Török J (2009) Effects of extrapair paternity and sex on nestling growth and condition in the collared flycatcher, *Ficedula albicollis*. *Animal Behaviour*, **77**, 611–617.

Rubolini D, Romano M, Boncoraglio G *et al.* (2005) Effects of elevated egg corticosterone levels on behavior, growth, and immunity of yellow-legged gull (*Larus michahellis*) chicks. *Hormones and Behavior*, **47**, 592–605.

Saino N, Bertacche V, Ferrari RP *et al.* (2002) Carotenoid concentration in barn swallow eggs is influenced by laying order, maternal infection and paternal ornamentation. *Proceedings of the Royal Society of London Series B-Biological Sciences*, **269**, 1729–1733.

Saino N, Ferrari R, Romano M, Martinelli R, Møller AP (2003) Experimental manipulation of egg carotenoids affects immunity of barn swallow nestlings. *Proceedings of the Royal Society of London Series B-Biological Sciences*, **270**, 2485–2489.

Sardell RJ, Arcese P, Keller LF, Reid JM (2011) Sex-specific differential survival of extra-pair and within-pair offspring in song sparrows, *Melospiza melodia*. *Proceedings of the Royal Society of London Series B-Biological Sciences*, **278**, 3251–3259.

SAS Institute (2003) *SAS Online Doc, version 9.1*. SAS Institute Inc, Cary, North Carolina.

Schmoll T, Dietrich V, Winkel W, Epplen JT, Lubjuhn T (2003) Long-term fitness consequences of female extra-pair matings in a socially monogamous passerine. *Proceedings of the Royal Society of London Series B-Biological Sciences*, **270**, 259–264.

Schmoll T, Dietrich V, Winkel W *et al.* (2005) Paternal genetic effects on offspring fitness are context dependent within the extrapair mating system of a socially monogamous passerine. *Evolution*, **59**, 645–657.

Schmoll T, Schurr FM, Winkel W, Epplen JT, Lubjuhn T (2009) Lifespan, lifetime reproductive performance and paternity loss of within-pair and extra-pair offspring in the coal tit *Periparus ater*. *Proceedings of the Royal Society of London Series B-Biological Sciences*, **276**, 337–345.

Sheldon BC (1993) Sexually transmitted disease in birds: occurrence and evolutionary significance. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, **339**, 491–497.

Sheldon BC (1994) Male phenotype, fertility, and the pursuit of extra-pair copulations by female birds. *Proceedings of the Royal Society of London Series B-Biological Sciences*, **257**, 25–30.

Sheldon BC (2000a) Environmental dependence of genetic indicator mechanisms. In: *Animal Signals: Signalling and Signal Design in Animal Communication* (eds Espmark Y, Amundsen T and Rosenqvist G), pp. 195–207. Tapir Academic Press, Trondheim.

Sheldon BC (2000b) Differential allocation: tests, mechanisms and implications. *Trends in Ecology & Evolution*, **15**, 397–402.

Sheldon BC (2002) Relating paternity to paternal care. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, **357**, 341–350.

Sheldon BC, Ellegren H (1999) Sexual selection resulting from extrapair paternity in collared flycatchers. *Animal Behaviour*, **57**, 285–298.

Sheldon BC, Merilä J, Qvarnström A, Gustafsson L, Ellegren H (1997) Paternal genetic contribution to offspring condition predicted by size of male secondary sexual character. *Proceedings of the Royal Society of London Series B-Biological Sciences*, **264**, 297–302.

Sheldon BC, Kruuk LEB, Merilä J (2003) Natural selection and inheritance of breeding time and clutch size in the collared flycatcher. *Evolution*, **57**, 406–420.

Slagsvold T, Sandvik J, Rofstad G, Lorentsen O, Husby M (1984) On the adaptive value of intraclutch egg-size variation in birds. *Auk*, **101**, 685–697.

Sockman KW, Schwabl H (2000) Yolk androgens reduce offspring survival. *Proceedings of the Royal Society of London Series B-Biological Sciences*, **267**, 1451–1456.

Stoleson SH, Beissinger SR (1995) Hatching asynchrony and the onset of incubation in birds, revisited. In: *Current Ornithology* (ed. Power DM), pp. 191–270. Plenum Press, New York.

Szöllősi E, Rosivall B, Török J (2007) Is hatching asynchrony beneficial for the brood? *Behavioral Ecology*, **18**, 420–426.

Townsend AK, Clark AB, McGowan KJ (2010) Direct benefits and genetic costs of extrapair paternity for female American crows (*Corvus brachyrhynchos*). *American Naturalist*, **175**, E1–E9.

Tryjanowski P, Hromada M (2005) Do males of the great grey shrike, *Lanius excubitor*, trade food for extrapair copulations? *Animal Behaviour*, **69**, 529–533.

Vedder O, Magrath MJL, Harts AMF *et al.* (2010) Reduced extrapair paternity in response to experimental stimulation of earlier incubation onset in blue tits. *Behavioral Ecology*, **21**, 9–15.

Vedder O, Komdeur J, van der Velde M, Schut E, Magrath MJL (2011) Polygyny and extra-pair paternity enhance the opportunity for sexual selection in blue tits. *Behavioral Ecology and Sociobiology*, **65**, 741–752.

Vinkler M, Bainová H, Albrecht T (2010) Functional analysis of the skin-swelling response to phytohaemagglutinin. *Functional Ecology*, **24**, 1081–1086.

Walling CA, Pemberton JM, Hadfield JD, Kruuk LEB (2010) Comparing parentage inference software: reanalysis of a red deer pedigree. *Molecular Ecology*, **19**, 1914–1928.

Wang JL (2004) Sibship reconstruction from genetic data with typing errors. *Genetics*, **166**, 1963–1979.

Wang JM, Beissinger SR (2009) Variation in the onset of incubation and its influence on avian hatching success and asynchrony. *Animal Behaviour*, **78**, 601–613.

Wang JL, Santure AW (2009) Parentage and sibship inference from multilocus genotype data under polygamy. *Genetics*, **181**, 1579–1594.

Wang JM, Firestone MK, Beissinger SR (2011) Microbial and environmental effects on avian egg viability: do tropical mechanisms act in a temperate environment? *Ecology*, **92**, 1137–1145.

Webster MS, PruettJones S, Westneat DF, Arnold SJ (1995) Measuring the effects of pairing success, extra-pair copulations and mate quality on the opportunity for sexual selection. *Evolution*, **49**, 1147–1157.

West BT, Welch KB, Gałecki AT (2007) *Linear Mixed Models: A Practical Guide Using Statistical Software*. Chapman & Hall/CRC, Boca Raton, Florida.

Westneat DF (1993) Temporal patterns of within-pair copulations, male mate-guarding, and extra-pair events in eastern red-winged blackbirds (*Agelaius phoeniceus*). *Behaviour*, **124**, 267–290.

Westneat DF, Stewart IRK (2003) Extra-pair paternity in birds: causes, correlates, and conflict. *Annual Review of Ecology and Systematics*, **34**, 365–396.

Westneat DF, Clark AB, Rambo KC (1995) Within-brood patterns of paternity and paternal behavior in red-winged blackbirds. *Behavioral Ecology and Sociobiology*, **37**, 349–356.

Wetton JH, Parkin DT (1991) An association between fertility and cuckoldry in the house sparrow, *Passer domesticus*. *Proceedings of the Royal Society of London Series B-Biological Sciences*, **245**, 227–233.

Whittingham LA, Dunn PO, Lifjeld JT (2007) Egg mass influences nestling quality in tree swallows, but there is no differential allocation in relation to laying order or sex. *Condor*, **109**, 585–594.

Wilk T, Cichoń M, Wolff K (2008) Lack of evidence for improved immune response of extra-pair nestlings in collared flycatcher *Ficedula albicollis*. *Journal of Avian Biology*, **39**, 546–552.

Wolf JB, Brodie ED (1998) The coadaptation of parental and offspring characters. *Evolution*, **52**, 299–308.

M. K. is behavioural ecologist with interest in parental care and sexual selection in birds. P.M. is interested in population genetics of mammals and birds, and genetics of speciation.

Data accessibility

The data supporting all analyses and results of this paper have been deposited in Dryad repository under DOI: 10.5061/dryad.984b0528.

Supporting information

Appendix S1 An example of the field protocol of the mixed cross-fostering design.

Appendix S2 Maternity assignment file. For each genotyped offspring, all candidate mothers are listed together with the number of mismatches between maternal and offspring genotypes. Assignments that were possible only by sibling analyses in Colony software are marked in the file.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Článek 12: Bowers, E. K., Munclinger, P., Bureš, S., Kučerová, L., Nádvorník, P. & Krist, M. (2013) Cross-fostering eggs reveals that female collared flycatchers adjust clutch sex ratios according to parental ability to invest in offspring. *Molecular Ecology*, **22, 215-228.**

Cross-fostering eggs reveals that female collared flycatchers adjust clutch sex ratios according to parental ability to invest in offspring

E. KEITH BOWERS,* PAVEL MUNCLINGER,† STANISLAV BUREŠ,‡ LENKA KUČEROVÁ,§
PETR NÁDVORNÍK§ and MILOŠ KRIST‡¶

Behavior, Ecology, Evolution, and Systematics Section, School of Biological Sciences, Illinois State University, Normal, IL, USA, †Department of Zoology, Faculty of Science, Charles University in Prague, Viničná 7, 128 44 Prague, Czech Republic, ‡Department of Zoology and Laboratory of Ornithology, Faculty of Science, Palacký University, tř. Svobody 26, 771 46 Olomouc, Czech Republic, §Department of Cell Biology and Genetics, Palacký University, Šlechtitelů 11, 783 71 Olomouc, Czech Republic, ¶Museum of Natural History, nám. Republiky 5, 771 73 Olomouc, Czech Republic

Abstract

Across animal taxa, reproductive success is generally more variable and more strongly dependent upon body condition for males than for females; in such cases, parents able to produce offspring in above-average condition are predicted to produce sons, whereas parents unable to produce offspring in good condition should produce daughters. We tested this hypothesis in the collared flycatcher (*Ficedula albicollis*) by cross-fostering eggs among nests and using the condition of foster young that parents raised to fledging as a functional measure of their ability to produce fit offspring. As predicted, females raising heavier-than-average foster fledglings with their social mate initially produced male-biased primary sex ratios, whereas those raising lighter-than-average foster fledglings produced female-biased primary sex ratios. Females also produced male-biased clutches when mated to males with large secondary sexual characters (wing patches), and tended to produce male-biased clutches earlier within breeding seasons relative to females breeding later. However, females did not adjust the sex of individuals within their clutches; sex was distributed randomly with respect to egg size, laying order and paternity. Future research investigating the proximate mechanisms linking ecological contexts and the quality of offspring parents are able to produce with primary sex-ratio variation could provide fundamental insight into the evolution of context-dependent sex-ratio adjustment.

Keywords: *Ficedula albicollis*, maternal effect, overdispersion, sex allocation, Trivers-Willard

Received 11 June 2012; revision received 20 September 2012; accepted 2 October 2012

Introduction

Despite broad interest in vertebrate sex allocation, there is still no consensus regarding whether females actively adjust the sex of their offspring following theoretical predictions (Cockburn *et al.* 2002; Cassey *et al.* 2006; Uller 2006). A number of adaptive hypotheses have been proposed to explain why reproductive females should adjust their offspring sex ratios according to

Correspondence: Miloš Krist, Fax: 420 585 222 743;
E-mail: milos.krist@upol.cz, milos.krist@volny.cz

their current state or environmental conditions, many of which are taxon-specific, reflecting the myriad ways in which selection can shape variation in the fitness of sons and daughters. Among the most commonly studied, and broadly applied, hypotheses is that of Trivers & Willard (1973), which used Bateman's Principle (Bateman 1948) to explain why females should adjust their relative production of sons and daughters under varying circumstances. Specifically, if high-quality males enjoy higher reproductive success than low-quality males and female reproductive success is less variable and falls between that of high- and low-quality

males (Krist 2006; West 2009), then mothers able to produce high-quality offspring should preferentially produce sons to produce males with high reproductive potential, but should produce daughters when unable to produce high-quality offspring because poor-quality daughters will have higher fitness than poor-quality sons (Trivers & Willard 1973). Since its inception, theoretical and empirical advances have expanded the original Trivers-Willard model to create a broad framework describing why females should produce one sex over the other across a variety of contexts, giving rise to the field of conditional sex allocation (Cockburn *et al.* 2002; West 2009).

Variation in offspring sex ratios can provide useful insight into sex-ratio studies (Williams 1979; Harmsen & Cooke 1983; Gowaty 1991). For example, if females have no control over the sex of their offspring, sex-ratio distributions should approach unity with binomial variance. However, sex-ratio variation greater than expected suggests that mothers bias the sex of their offspring towards sons or daughters more often than predicted by chance (e.g. as per the Trivers-Willard model), which could increase variation in sex ratios without necessarily affecting the average population sex ratio (Krackow *et al.* 2002; Wilson & Hardy 2002). Aside from varying the overall clutch sex ratio, females of multiparous species may also adjust the sex of individual offspring in relation to the order in which they are produced, for example to provide either sex with a competitive advantage over its siblings (Carranza 2004).

Although meta-analyses have confirmed that vertebrate mothers can adjust the sex of their offspring (West & Sheldon 2002; Cassey *et al.* 2006), opinion remains divided as to whether variation in sex ratios follows theoretical expectation (Cassey *et al.* 2006; Uller 2006). A source of this contention is attributable to a scarcity of experimentation in field studies and the reality that most studies on vertebrates are correlative and, thus, cannot control for important confounding effects. For example, in sexually size-dimorphic species, offspring sex covaries with body size and response to parental resource allocation, obscuring the assumed effect that maternal quality and reproductive effort have on offspring condition and reproductive value (Hewison & Gaillard 1999; Gaillard *et al.* 2000; Magrath *et al.* 2007).

Indeed, it is logistically difficult to study sex allocation experimentally in natural settings and to obtain sufficient a priori background on whether certain manipulations should actually affect sex allocation for different species or even populations (e.g. Rosivall *et al.* 2004; West 2009), although manipulating male attractiveness (signals of individual quality that affect female choice) is an important exception (Sheldon *et al.* 1999; Dubois *et al.* 2006; West 2009). With respect to the broad-sense Trivers-Willard model and whether parental investment ability should influence offspring sex ratios, cross-fostering experiments (e.g. Koskela *et al.* 2009; Rutkowska *et al.* 2011) can provide a reliable means of inferring cause-and-effect evidence for sex-ratio variation. For example, by cross-fostering sons and daughters prior to their receiving extensive resources from parents, the energetic demands and subsequent reproductive costs to females rearing experimentally manipulated sex ratios can shed insight into how females allocate limited energy towards the production of male and female offspring at any point in time (Rutkowska *et al.* 2011). Moreover, the condition of foster offspring that parents raise will be unrelated to the primary sex ratio the mother produced, thereby overcoming the covariance between the primary sex ratio and sex-dependent responses to parental resource allocation (Magrath *et al.* 2007). The average condition of foster young, then, provides an unbiased, functional measure of parental ability to raise fit offspring, which can be used to predict the primary sex ratio that the mother initially produced (Robert *et al.* 2010). Therefore, sex-allocation theory predicts that parents demonstrating an ability to raise foster offspring in good condition should have originally produced male-biased primary sex ratios prior to cross-fostering, whereas parents that raise foster offspring in poor condition should have produced female-biased sex ratios, all else being equal.

In this study, we investigated primary sex-ratio variation in a population of collared flycatchers (*Ficedula albicollis*). The collared flycatcher is a good model for studying sex allocation because males have higher variance in reproductive success than females as a consequence of polygyny (Qvarnström *et al.* 2003; Hegyi *et al.* 2007) and frequent extra-pair copulations (Sheldon & Ellegren 1999; Rosivall *et al.* 2009; Krist & Munclinger 2011); recent studies of our population indicate that at least 50% of females engage in extra-pair copulations and that about 25% of offspring are sired by extra-pair males (Krist *et al.* 2005; Krist & Munclinger 2011). The body mass of nestlings is known to influence their subsequent survival (e.g. Lindström 1999; Merilä *et al.* 2001) and often persists into adulthood (reviewed in Monaghan 2008), and can even carry sex-specific consequences (Gustafsson *et al.* 1995; Tilgar *et al.* 2010). Moreover, male offspring are more sensitive than female offspring to food shortage as nestlings (Rosivall *et al.* 2010), further suggesting that the fitness potential of sons and daughters varies unequally with the quality of the rearing environment (Krist 2006; Wilkin & Sheldon 2009), which is an important assumption of the Trivers-Willard model in its broader sense (West 2009). We tested several specific hypotheses falling into this broad framework:

1 We first used a cross-fostering approach to test the broad-sense Trivers-Willard model (Cockburn *et al.* 2002; West 2009) by cross-fostering eggs among nests prior to being incubated, allowing us to use the average body mass of foster offspring that parents produced as a functional measure of their ability to invest in offspring; in this case, theory predicts that parents producing male-biased primary sex ratios will raise heavier foster offspring than parents producing female-biased primary sex ratios.

2 We then tested whether females biased offspring sex ratios in relation to secondary sexual traits of their social mate; males have conspicuous white patches on the forehead and wing that are condition-dependent signals of quality (large patches indicate high quality), and are subject to inter- and intra-sexual selection (Pärt & Qvarnström 1997; Sheldon & Ellegren 1999; Török *et al.* 2003; Garamszegi *et al.* 2006). We used secondary sexual traits as opposed to morphological ones because they have been well-established as conspicuous, condition-dependent signals that are assessed by females when making mate-choice decisions. For example, females prefer males with large white patches over those with smaller patches (Sheldon & Ellegren 1999; Qvarnström *et al.* 2000; Hegyi *et al.* 2010); thus, we predicted that females mating with males possessing large secondary sexual characters would produce male-biased clutches relative to those paired to males with smaller patches because, all else being equal, sons that inherit their father's attractiveness and, thus, enjoy high mating success as adults can improve their mother's fitness to a greater extent than daughters (Fawcett *et al.* 2007).

3 Conditions for raising high-quality offspring are generally more favourable earlier within breeding seasons than they are later on (reviewed in Verhulst & Nilsson 2008), which can influence the fitness prospects for sons and daughters differently (e.g. Dijkstra *et al.* 1990; Cordero *et al.* 2001); thus, we also analysed clutch sex ratios in relation to the timing of breeding, predicting that mothers reproducing earlier within breeding seasons produce male-biased clutches relative to mothers breeding later on (see also Pen *et al.* 1999; Neto *et al.* 2011).

4 Maternal age may also be predicted to influence the sex of her offspring (Blank & Nolan 1983; Saltz 2001; Dowling & Mulder 2006; but see Hewison *et al.* 2002; Saltz & Kotler 2003); for example, if females increase reproductive effort with age or become better at rearing high-quality offspring than younger females, sex-allocation theory predicts that older females should produce male-biased clutches relative to younger, less-experienced breeders.

5 Variation in egg size or hatching order within clutches can create a competitive hierarchy among siblings, which can have long-term effects on offspring morphology, survival and reproductive success (Krist *et al.* 2004; Rosivall *et al.* 2005; Krist 2009); in such cases, females may adjust the sex of individual offspring to maximize their reproductive potential (Carranza 2004; West 2009). Thus, we tested whether females bias the sex of earlier-laid offspring within their clutches towards sons and later-laid offspring towards daughters (e.g. Badyaev *et al.* 2002; Ležalová *et al.* 2005; Bowers *et al.* 2011), and whether mothers bias the sex of large eggs towards sons and small eggs towards daughters (e.g. Mead *et al.* 1987; Cordero *et al.* 2000).

6 Lastly, because 'good genes' explanations for extra-pair mating often suggest that females do so to increase the quality of their offspring (Sheldon *et al.* 1997; Akçay & Roughgarden 2007), we also tested whether mothers bias the sex of extra-pair offspring towards sons to a greater extent than their within-pair offspring (e.g. Sheldon & Ellegren 1996; Schwarzová *et al.* 2008; Johnson *et al.* 2009).

Materials and methods

Study site and species

This study was conducted in the Velký Kosíř area, Moravia, Czech Republic (49° 32′N, 17° 04′E; 300–400 m a.s.l.), which contains 350 nest-boxes distributed in an oak (*Quercus petrea*) forest available for breeding flycatchers. Collared flycatchers are small (13 g), cavity-nesting songbirds, and females produce one brood of young per year, although males can produce multiple broods through polygyny. The species is only slightly sexually size-dimorphic in body size (Przybylo *et al.* 2000), but has sexually dimorphic plumage. Females are dull brown in colour, with a small white patch on the wing; males have a white breast and a black back and crown, with a white nape, forehead, and wing patch, which is larger in males than in females. After arriving on the breeding grounds from wintering quarters and selecting a mate, females lay one egg per day to produce a clutch of four to eight eggs (modal clutch size is six eggs). Most females begin full incubation prior to laying the last egg, generating hatching asynchrony, with last-laid eggs of the clutch hatching much later than earlier-laid eggs (Krist *et al.* 2004; Rosivall *et al.* 2005).

General field and laboratory procedures

From 2001 to 2002 and from 2006 to 2009, we visited nests daily when egg laying was expected to begin,

numbered eggs as laid using nontoxic markers, and measured their length and breadth using digital calipers, which we used to calculate egg size (mm^3) using Hoyt's (1979) formula. In 2001 and 2002, we transferred eggs to an incubator prior to hatching so we could accurately determine the position of young within the hatching sequence (80.4% of eggs transferred to the incubator subsequently hatched). We replaced natural eggs with dummy eggs, which the females readily accepted, to prevent maternal abandonment, and we used a thermobox to transfer eggs to the incubator within 10 min of their removal from the nest (see Krist et al. 2004 for details). We checked the incubator regularly for hatching and, thus, were able to return hatchlings to their nests within 2.95 ± 2.33 h (mean $\pm$ SD) of hatching. Prior to returning hatchlings to their nests, we marked their claws with nail polish for individual identification throughout the nestling period. From 2001 to 2002, at 10–13 days of age, we drew blood (~25 μL) from the nestlings' brachial veins and stored the blood samples in Queen's lysis buffer (Seutin et al. 1991) until later DNA extraction and molecular sexing based on PCR amplification of sex-linked CHD1 introns that differ in length between the Z and W sex chromosomes (Griffiths et al. 1998). From 2006 to 2009, blood samples were obtained from the tarsal veins of nestlings at 6 days of age and were stored in ethanol. Since in some cases the CHD1 intron amplification failed or produced weak bands on gels, an alternative method targeting the sex-linked ATP5A1 gene (Bantock et al. 2008) was used in 415 samples. We did not find any incongruence between the two sexing methods in 223 samples that we sexed using both the CHD1 and ATP5A1 methods. Occasionally eggs would go unhatched or a nestling would die prior to bleeding; however, because we visited nests daily during the nestling stage, we collected dead nestlings and stored them in ethanol and froze unhatched eggs for later DNA extraction from the preserved tissues.

After hatching, we captured the putative mothers and fathers while they provisioned nestlings, drew a small blood sample, and measured the size of the males' wing (mm) and forehead (mm^2) patches (details in Krist & Munclinger 2011). In 2001 and 2002, we drew blood samples (~25 μL) from the adults' brachial veins and, from 2006 to 2009, drew blood samples (1–5 μL) via tibial venipuncture. We determined paternity of offspring in 2001 and 2002 using three microsatellite loci, where offspring that mismatched their social fathers at any of the three markers were deemed to be extra-pair young (details in Krist et al. 2005). The combined exclusion power of the marker set is ~96% (Sheldon & Ellegren 1996). Thus, a few nestlings from 2001 to 2002 may have been erroneously assigned within-pair paternity,

which could add only noise, not a bias, in our results for offspring sex in relation to paternity. From 2006 to 2009, we genotyped offspring and adults at eight microsatellite loci to determine the parentage. We scored genotypes using GeneMarker software version 1.9 (Softgenetics) and used Cervus version 3.0.3 (Kalinowski et al. 2007) and Colony version 2.0 (Wang 2004) to assign parentage to the cross-fostered offspring (see Krist & Munclinger 2011 for details). Because of high recruitment (15–20% of fledged young) and annual adult return rates (~40%), we could determine the exact age (years) of 100 females (ranging from 1 to 7 years) and the minimum age for an additional 43 females. We categorized male ages into two classes (yearlings vs. older adults) as this was known without error for all the males we captured based on their plumage and is also a better way of controlling for age (rather than a male's exact age) when analysing secondary sexual characters.

Cross-fostering experiment, 2006–2009

From 2006 to 2009, we cross-fostered eggs among clutches prior to incubation, which allowed us to associate hatchlings with source eggs without the use of an artificial incubator and also provided a functional measure of parental ability to invest in offspring. On the morning each egg was laid; we numbered it, measured it (as above), and replaced it with a dummy egg. When females ceased egg production (no eggs laid on subsequent days), we replaced the dummy eggs with natural eggs so that each foster egg within a nest was produced by a different female. We made the size of a female's foster clutch the same size as her original clutch, and we randomly assigned foster eggs to clutches, but we did assign them to clutches initiated at a similar point in time, thus mimicking the natural span of time eggs would go unincubated under natural conditions (Krist & Munclinger 2011). Because each foster egg within a nest was produced by a different female, we used blood samples to assign maternity to each foster nestling (conspecific brood parasitism does not occur in this population; Krist et al. 2005) and then knew the nest from which each nestling originated and the nestling's position within its original egg-laying sequence.

After hatching, we drew a blood sample from each nestling at 6 days of age and, at 13 days of age, we weighed each nestling (0.25 g) and measured the length of their tarsi (0.01 mm). Because we cross-fostered eggs in 2006–2009 prior to being incubated, we uncoupled any correlation between the primary sex ratio and offspring condition after egg laying, which would otherwise be problematic when nestlings are sexually size-dimorphic (Magrath et al. 2007). For example, if

sons tend to be larger than daughters and can, thus, outcompete their sisters for parental resources (Magrath *et al.* 2007), variation in the mass or body condition of offspring produced would be largely attributable to sexual size-dimorphism and the relative number of sons and daughters within a nest, and not necessarily to parental effort. Indeed, by cross-fostering eggs at random, the sex ratio of foster broods was uncorrelated with the primary sex ratio that females produced (Spearman rank correlation $r_{151} = -0.05$, $P = 0.513$), and the among-nest variation in foster-brood sex ratios did not deviate from the null binomial expectation (SAS randomization test, two-tailed $P = 0.363$; see 'Data and analyses' below for details). Thus, the mean body mass of foster offspring that parents raise to fledging provides an unbiased measure of their ability to invest in offspring (see Robert *et al.* 2010 for an approach using offspring survival), which we used to predict the primary sex ratios that females produced prior to cross-fostering their eggs. We used nestling body mass rather than other morphological traits, for example, tarsus length, because of much higher heritability of the latter trait (Krist 2009), which would obscure our ability to use measures of nestling condition based on skeletal size as proxies of parental effort. Because of our cross-fostering approach, the mass of foster nestlings provides the best estimate of parental abilities, as it is mostly determined by parental effort; on the other hand, recruitment or reproductive success of foster young would be less reliable measures of parental effort because they are influenced by genes and conditions encountered later in life, after those individuals have left the nest and are, thus, outside parental control.

Data and analyses

We determined the sex of 221 offspring from 38 broods in 2001 and 2002; from 2006 to 2009, we determined the sex of 942 offspring from 160 broods. We were unable to sex three of 224 offspring from 2001 and 2002, and 61 of 1003 viable offspring from 2006–2009 because eggs and nestlings occasionally disappeared from the nest prior to obtaining their DNA, and dead nestlings' DNA occasionally degraded prior to collection. Given the high percentage of all offspring that were sexed successfully (1163 of 1227; 94.8%), our data can be considered to reflect primary sex ratios; indeed, offspring mortality prior to sexing would have to be strongly sex-biased to create departures from the primary sex ratio (Fiala 1980; Krackow & Neuhäuser 2008; West 2009). As advised by Krackow & Neuhäuser (2008), we did not exclude broods with incomplete sex-ratio data, particularly as this means analysing nonrandom subsamples of data. We were unable to assign paternity to 207 of the

1163 offspring, and so could analyse offspring sex in relation to paternity for 956 of them. Given that effect sizes in studies of vertebrate sex ratios are generally small, sample sizes often are not sufficiently high to prevent false negatives; however, the size of our data set meets the requirements ($N \approx 200$ clutches) suggested by West (2009) to minimize the probability of type II error.

Regardless of mean sex ratios, among-clutch variation in offspring sex ratios can provide insight into sex-allocation strategies (e.g. Williams 1979; Harmsen & Cooke 1983; Gowaty 1991). There are a number of valid ways to investigate variation in sex ratios (see, e.g. Krackow *et al.* 2002; Wilson & Hardy 2002; here, we compare/contrast two randomization approaches. We first analysed whether the residual deviance among clutch sex ratios ($N = 198$ clutches from 2001–2002 and 2006–2009) was greater than expected (i.e. overdispersed) under the null hypothesis using a commonly used randomization test in SAS (C. M. Lessells, personal communication), which compared the observed deviance with a null distribution of deviance values generated by 10 000 simulated data sets with similar attributes (i.e. clutch sizes and the total number of male and female offspring) as the real data set (see also Westerdahl *et al.* 1997; Ewen *et al.* 2003 and associated references). For this test, which assumes no extra-binomial variation as the null hypothesis, the proportion of simulated data sets with a deviance greater than the observed value represents the test's P-value (C. M. Lessells, personal communication). We then used a randomization test in R 2.12.1 (r-project.org) that generated a distribution of 10 000 expected clutch sex ratios and compared the variance distribution from those simulations with the observed variance (Postma *et al.* 2011). This test requires the number of offspring within a clutch, brood or litter (hereafter clutch) that were male and the number of offspring within the clutch that were sexed to create a distribution of sex ratios, each with a random number of males that is weighted by the mean proportion of males among clutches (Postma *et al.* 2011). As noted by Postma *et al.*, this approach corrects for variation in clutch size, which is particularly important when clutch sizes are small and sex-ratio distributions do not resemble a binomial distribution (see also Ewen *et al.* 2003).

We then analysed variation in the sex ratios of cross-fostered clutches ($N = 160$ clutches from 2006–2009; all clutches analysed from this time period were cross-fostered) with the number of male offspring in a clutch as the dependent variable and the number of sexed offspring within the clutch as the binomial denominator (i.e. event/trial syntax) using generalized linear mixed models (GLMMs; PROC GLIMMIX in SAS) with maternal identity as a random effect to account for nonindependence of clutches produced by the same female

(28 females produced two clutches and four females produced three clutches). In this analysis, we tested whether the mean body mass of fledged foster young predicted the primary sex ratio a female produced prior to cross-fostering; we also tested for effects of the social father's age (yearling vs. older), wing-patch size and forehead patch area along with the date of clutch initiation, maternal age and year as independent variables. Intra-specific variation in clutch size may also contribute to sex-ratio variation, although it may complicate predictions of the Trivers-Willard model (Frank 1990). For example, parents with above-average investment abilities could bias the sex ratio towards sons, as predicted in the current study, or alternatively, they could increase the number of offspring they produce without adjusting the sex ratio. Thus, we also included clutch size as an independent variable in this analysis. Brood sex ratios also have the potential to influence offspring growth (Bonisoli-Alquati *et al.* 2011); thus, we also included the sex ratio of foster broods to control for sex-specific sibling rivalry and its potentially confounding influence on our measure of parental abilities. Our model included several independent variables that could be correlated with one another; for example, maternal age was slightly correlated with breeding date, such that younger females tended to breed later than older females (Pearson's $r_{141} = -0.164$, $P = 0.050$). Thus, we investigated whether our model exhibited multicollinearity, or strong correlations among independent variables, but found no suggestion of this. Multicollinearity can be problematic because it inflates variances and standard errors associated with parameter estimates; thus, we calculated variance inflation factors (VIF) for our independent variables to assess whether any of them had experienced such an increase. A VIF greater than or equal to 10 is generally thought to suggest multicollinearity, but each VIF in our model was less than three (range = 1.1–2.59). Although our tests for overdispersion were indicative of extra-binomial variation in clutch sex ratios (see Results), the model we fitted to describe variation in clutch sex ratios (Table 1) provided a good, valid fit to the data because the residual mean deviance was very close to 1 (1.03; see Wilson & Hardy 2002); thus, we did not need scale parameters to adjust for overdispersion in that analysis.

Lastly, we analysed variation in the sex of individual offspring from 2001–2002 and 2006–2009 in relation to egg size, laying order ($N = 1162$ offspring; we did not obtain egg size for one of the 1163 sexed offspring) and paternity ($N = 956$ offspring) using a GLMM with a binomial distribution and logit link, including clutch identity as a random effect to account for nonindependence of offspring within clutches.

Table 1 Summary of effects on clutch sex ratios. See Fig. 2 for effect sizes

|	Estimate	SE	*F*	d.f.	*P*
Mass of fledged chicks	0.194	0.089	4.78	1, 108	0.031
Social father's age	−0.470	0.352	1.78	1, 108	0.185
Social father's wing patch	0.018	0.009	4.15	1, 108	0.044
Social father's forehead patch	−0.00005	0.0009	0.00	1, 108	0.957
Laying date	−0.053	0.027	3.75	1, 108	0.056
Maternal age	−0.009	0.065	0.02	1, 76.6	0.889
Clutch size	0.003	0.128	0.00	1, 108	0.979
Foster-brood sex ratio	−0.402	0.379	1.12	1, 108	0.292
Year			0.66	3, 108	0.581
Intercept	3.44	3.80			
Maternal identity	*0.035*	*0.079*			

Note: Random effects are in italics.

Presentation of data

To summarize the effects on offspring sex, we report effect sizes ± 95% confidence intervals for each variable in our analyses (Nakagawa & Cuthill 2007). We used *F*-statistics from the GLMMs described above to calculate effect sizes as correlation coefficients using the formula: $r = \sqrt{[F/(F + DDF)]}$, where DDF represents the error (denominator) degrees of freedom (Rosenthal 1994), and we calculated confidence limits using z-transformation (Sokal & Rohlf 1995). Effect sizes of 0.1, 0.3 and 0.5 are often viewed as small, medium and large, respectively (Cohen 1988). We plotted the direction of effects such that effects in support of the Trivers-Willard model are positive values and evidence against the Trivers-Willard model are less than or equal to zero.

Results

Among-clutch variation in sex ratios

The average clutch sex ratio did not differ from 1:1 (mean ± SE proportion of male offspring = 0.518 ± 0.016; binomial test Z = 1.26, $P = 0.207$), but we detected extra-binomial variation in sex ratios (Fig. 1a, b), as the relative frequency of 1:1 sex ratios was lower than expected, and the frequencies of female- and male-biased clutches were each higher than expected (Fig. 1b). Specifically, our randomization test in SAS produced 83 of 10 000 simulated data sets with a deviance greater than the observed value (Fig. 1a). Such a small proportion of simulations exceeding the observed value suggests that sex-ratio variation is significantly greater than that expected under the null
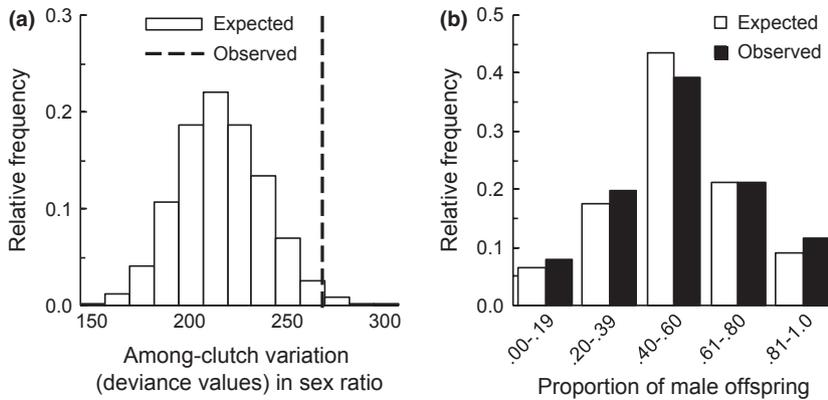
Fig. 1 Among-clutch variation in sex ratios. (a) Relative frequency distribution of expected (i.e. null) deviances generated from 10 000 simulated data sets in SAS compared with the deviance in clutch sex ratios observed in the real data (dashed line). (b) Relative frequency distribution of observed sex ratios and those expected from 10 000 randomly simulated data sets in R (see text for details).

hypothesis ($P = 0.0083$). However, the randomization test in R revealed only a marginal excess of female- and male-biased broods relative to the null expectation, as 5.26% of the 10 000 simulated data sets produced a variance greater than that observed in the real data ($P = 0.0526$).

Clutch sex ratios

The overall sex ratio a mother produced varied according to parental ability to produce high-quality offspring. Specifically, mothers that reared heavier-than-average foster fledglings with their social mates produced a greater proportion of male offspring prior to cross-fostering than females rearing lighter-than-average foster fledglings with their social mate (Table 1, Figs. 2, 3). Furthermore, females paired to males with large white wing patches were more likely to produce sons than females paired to males with smaller wing patches, after correcting for male age. Females also tended to produce male-biased clutches early within breeding seasons and female-biased clutches later, but sex ratios were uncorrelated with clutch size (fitting clutch size as a categorical variable yielded the same result). Although maternal age was slightly correlated with breeding date, the effect of breeding date was determined after correcting for differences in maternal age, which had no influence on clutch sex ratios (Table 1, Fig. 2). We also simplified the full model reported in Table 1 through stepwise elimination of nonsignificant terms ($P > 0.1$), and found that the significant results reported in Table 1 remain significant in the reduced model (see Appendix S1).

Sex of individual offspring

The sex of individual offspring did not vary according to absolute egg size or laying order within clutches (Table 2; Figs. 2, 4), and analysing offspring sex in relation to egg size and laying order relative to the rest of the clutch produced similar results (data not shown).
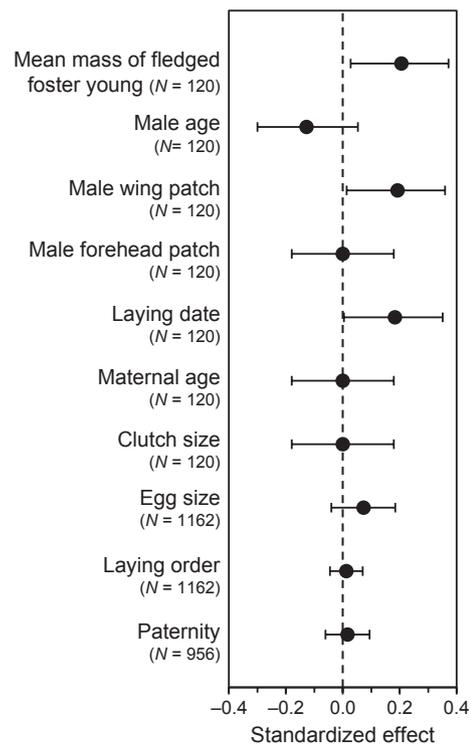
Fig. 2 Effect sizes associated with effects on clutch sex ratios and the sex of individual offspring. Standardized effects (correlation coefficients ± 95% C.L.) correspond to the relationship between parental investment abilities, maternal age and traits of the social father on clutch sex ratios, and effects of within-clutch variation in egg size, laying order and paternity on the sex of individual offspring. We plotted the direction of effects such that effects in support of the Trivers-Willard model are depicted as positive values and evidence against the Trivers-Willard model as zero and negative values (see 'Presentation of data').

We also found no association between offspring sex and paternity, as there was no difference in the sex ratio of offspring sired by within-pair males and those sired by extra-pair males (proportion of sons ± SE sired by within-pair males: 0.511 ± 0.019, extra-pair: 0.526 ± 0.031; Table 2).
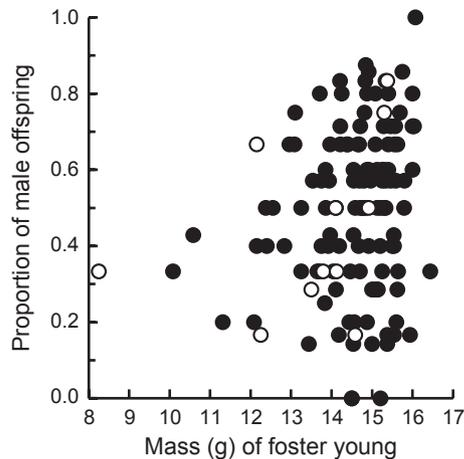
Fig. 3 Proportion of male offspring females produced in rela-
tion to the average mass of foster young raised to fledging.
Filled symbols (*N* = 120) are nests analysed by the full model
depicted in Table 1, and open symbols were excluded from the
full model because we did not capture the social father (*N* = 9)
or because we did not measure the social father's forehead
patch size (*N* = 2).

Table 2 Summary of effects on the sex of individual offspring
within clutches. See Fig. 2 for effect sizes

	Estimate	SE	*F*	d.f.	*P*
Egg size	0.0007	0.0005	1.85	1, 253.1	0.175
Laying order	−0.008	0.036	0.05	1, 901.7	0.816
Paternity	0.05	0.15	0.11	1, 628.5	0.746
Intercept	−1.035	0.806			
Nest of origin	*0.119*	*0.092*			

Note: Random effects are in italics.

Discussion

As predicted by the Trivers-Willard model, parents pro-
ducing male-biased primary sex ratios produced
heavier-than-average foster offspring, whereas parents
of female-biased primary sex ratios produced lighter-
than-average foster offspring. Thus, our results suggest
that the condition of offspring that parents are able to
raise to fledging directly predicts the primary sex ratio
a female produces. Similar to our findings, Robert *et al.*
(2010) recently cross-fostered offspring in a marsupial
mammal prior to extensive maternal investment and
found that females originally birthing a son were more
successful at weaning foster offspring than those origi-
nally birthing a daughter. Aside from parental ability,
sex-specific dispersal or philopatry of offspring can also
influence the sex ratio, particularly in species where off-
spring inherit at least part of their parents' territories or
when one sex is more likely than the other to remain at
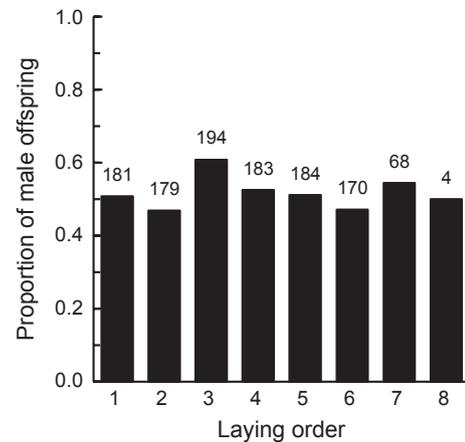
Fig. 4 Proportion of male offspring in relation to egg-laying
order. Numbers above bars represent sample sizes.

or near the natal nest (Julliard 2000; West 2009). For
example, females breeding on high-quality territories
could bias the sex ratio towards sons if those sons are
more likely than daughters to inherit the high-quality
territory. However, this hypothesis could not explain
our results because dispersal distances for both male
and female flycatcher offspring are much greater than
the size of breeding territories, meaning that there
would be no overlap between the breeding territories of
parents and their offspring (Král & Bičík 1990; Pärt
1990). Moreover, Tilgar *et al.* (2010) recently showed
that male great tit (*Parus major*) offspring actually dis-
perse farther when leaving the nest in good condition
than when in poor condition, which is the opposite of
what would be predicted if females on high-quality ter-
ritories over-produce sons so those sons will inherit the
territory.

A key assumption of the Trivers-Willard model is
that either sex will benefit or suffer to a greater extent
than the other across a range of rearing conditions
(Hewison & Gaillard 1999; Krist 2006; West 2009).
Although this is generally thought to apply to species
with greater degrees of sexual size-dimorphism than
collared flycatchers (Sheldon *et al.* 1998; Sheldon &
West 2004), Rosivall *et al.* (2010) recently showed that
the growth of sons benefits to a greater extent than that
of daughters under good rearing conditions, but suffers
to a greater extent under poorer conditions (see also
Bowers *et al.* 2011, 2012b), suggesting a sex-specific
response to environmental conditions that is not neces-
sarily attributable to overall differences in body size
(but see Sheldon *et al.* 1998). This sex-dependent devel-
opmental sensitivity satisfies an important assumption
of the Trivers-Willard model that is rarely demonstrated
(Hewison & Gaillard 1999) and, because parents able to
produce offspring in good condition bias the sex of

their offspring towards sons, suggests that sons of parents with above-average investment abilities benefit substantially from the increased reproductive effort by their parents as nestlings. Moreover, as adults, male reproductive success is more variable than that of females because of occasional polygyny (Qvarnström *et al.* 2003; Hegyi *et al.* 2007) and frequent extra-pair copulations (Sheldon & Ellegren 1999; Rosivall *et al.* 2009; Krist & Munclinger 2011), which are often determined by male quality (Sheldon *et al.* 1997; Akçay & Roughgarden 2007), thus increasing the potential for sex-specific fitness returns to parents from their adult offspring.

Although most studies have used measures of maternal body condition as an indicator of a female's ability to raise fit offspring, Trivers & Willard (1973) noted that the concept of 'condition' can be broadened to represent various components of parental abilities. Indeed, more dynamic measures of parental investment ability, for example, behavioural dominance, than static measures of morphological or physiological condition can provide a clearer picture of parental ability to raise high-quality offspring and the sex ratio they produce (Sheldon & West 2004). Here, we demonstrate that cross-fostering offspring prior to sex-differential allocation of resources to neonates and monitoring the subsequent performance of foster offspring provides a functional measure of parental abilities that is unbiased by sex-differences in size or response to maternal resource allocation (e.g. Clutton-Brock *et al.* 1985; Forsyth *et al.* 2004; Rutkowska *et al.* 2011).

We emphasize, however, that our study design does not rule out the potential for genetic and pre-hatching maternal effects to influence offspring quality. For example, variation in maternal effort may be reflected in the quality of their eggs, which is known to influence offspring growth after hatching (Gil 2008; Biard *et al.* 2009). However, the significant relationship between foster offspring mass and primary sex ratios we detected suggests that post-hatching parental effort could trump such pre-hatching maternal effects, at least in our study species. Of course, the fact that foster clutches/broods in the current study were comprised of random samples of eggs, each produced by a different female, controls for maternal effects on offspring development after hatching. Although variation in pre-hatching maternal and genetic effects could create noise in our ability to use foster offspring mass as a measure of parental effort, it is important to note that such effects could not bias our results. Therefore, the relationship between primary sex ratios and the condition of foster offspring that we report should be considered conservative. Our measure of parental abilities to raise high-quality offspring also avoids the problem of offspring sex vs. number (e.g.

Frank 1990), because parents that ultimately raise high-quality foster young are predicted to have produced sons regardless of clutch or brood size. Our egg-size data also allow us to exclude the hypothesis that mothers differentially invest in sons and daughters across the egg-laying and nestling stages. For example, it could be hypothesized that below-average investment in egg size and above-average investment in provisioning for sons, and the reverse for daughters, could create the observed correlation between primary sex ratios and condition of foster chicks. Such a strategy could, for example, allow high-quality mothers that over-produce sons to minimize the reproductive costs of egg production (e.g. Bowers *et al.* 2012a) while maximizing the provisioning effort of their social mate. However, the fact that males and females hatched from similar-sized eggs suggests that this is unlikely.

We also found that mothers biased clutch sex ratios towards sons when mated to males with large white wing patches and towards daughters when mated to males with smaller wing patches; however, we did not detect a relationship between the size of the male's forehead patch and sex ratios, which was also reported for a Swedish population of collared flycatchers (Ellegren *et al.* 1996). Secondary sexual characters can have differing effects on offspring sex ratios if they signal different kinds of information to mates (Taff *et al.* 2011), which may be the case for the two ornaments in our population. For example, in the isolated Swedish population, the size of the forehead patch is condition-dependent (Gustafsson *et al.* 1995) and functions in both intersexual (Qvarnström *et al.* 2000) and intrasexual (Pärt & Qvarnström 1997) contexts. However, in the core of the species' distribution throughout central Europe, this ornament is not condition-dependent (Hegyi *et al.* 2002) and has no function as a badge of status (Garamszegi *et al.* 2006), although it still plays a role in mate attraction (Hegyi *et al.* 2010). Unlike the forehead patch, the wing patch is a condition-dependent ornament in central Europe (Török *et al.* 2003) and also functions in both intersexual (Hegyi *et al.* 2010) and intrasexual (Garamszegi *et al.* 2006) contexts. Thus, our finding of sex-ratio bias with respect to the size of the wing patch, but not the forehead patch, is predicted by the relative roles of the two ornaments in sexual selection in central Europe. A previous study of a Hungarian population, however, did not detect sex-ratio adjustment with respect to either of these ornaments, albeit with a smaller sample size (Rosivall *et al.* 2004).

In addition to varying sex ratios according to parental investment ability and male attractiveness, mothers also tended to produce sons at a higher frequency earlier in the breeding season than mothers reproducing later on, and this was the case after correcting for maternal age,

suggesting that females adjusted their primary sex ratios according to local environmental conditions. Seasonal variation in sex allocation may provide a fitness advantage to mothers in a number of ways (Dijkstra *et al.* 1990; Pen *et al.* 1999; Freed *et al.* 2009). For example, Cordero *et al.* (2001) noted that recruitment and reproductive opportunities for daughters in their first year of life were markedly higher when produced early within breeding seasons relative to daughters produced later on, but reproductive success for sons was unrelated to the date at which they were produced; thus, mothers biased their clutches towards daughters earlier, and towards sons later, within breeding seasons. In many species, offspring fledging earlier than others within breeding seasons enjoy a fitness advantage over offspring produced later in the season, which have a narrower window of time between hatching and the onset of moult and autumn migration for migratory species or the onset of winter and reduced food availability for nonmigratory ones (Nilsson 1989, 1990; Norris 1993; Verhulst & Nilsson 2008). Indeed, seasonal variation in resource availability and offspring survival is often a very predictable pattern, which West & Sheldon (2002) noted as an essential requirement for selection to shape variation in sex ratios (see also Hasselquist & Kempenaers 2002).

Although our test for overdispersion using simulated deviance residuals suggested a high degree of overdispersion (Fig. 1a), our test using simulated variances revealed only marginal overdispersion. As noted by Postma *et al.* (2011), adjusting for variation in clutch size is critical in analyses of sex-ratio variance when clutch sizes are small and sex-ratio distributions vary widely with changes in clutch size (see also Krackow *et al.* 2002; Ewen *et al.* 2003). For example, the number of sexed offspring within broods studied by Postma *et al.* (2011) ranged from one to four, creating seven possible sex ratios, and the average number of offspring at independence was two; completely female- and male-biased broods occurred even more frequently than broods with a 1:1 sex ratio, forming a W-shaped sex-ratio distribution (Postma *et al.* 2011). Because the expected variance in sex ratios is determined by variation in clutch size, failure to account statistically for clutch size in that situation could lead researchers to over-estimate among-clutch variance in sex ratios, but our results suggest that this approach is much more conservative than the randomization test using deviance residuals. In our data set, the number of sexed offspring within a clutch ranges from one to eight, creating 23 possible sex ratios. In such cases, tests using deviance residuals are often recommended (Krackow *et al.* 2002), but increasing clutch size can lead to an increase in the risk of type I error (Ewen *et al.* 2003). A problem that

arises, then, is how to determine whether clutch-size variation in a given data set warrants using either test, and how to assess objectively the magnitude of difference between a null binomial distribution and the observed sex-ratio distribution. Developing methodology similar to the calculation of effect sizes may allow researchers to assess objectively whether or not sex-ratio variation is greater than expected under the null hypothesis.

Egg mass is commonly associated with maternal effort in birds because it positively predicts nutrient content (Nager *et al.* 2000) and offspring size and growth rates post-hatching (Krist 2009, 2011). Thus, sons may be predicted to hatch from larger eggs than daughters (Mead *et al.* 1987; Nager *et al.* 1999; Cordero *et al.* 2000). Theory also predicts that sons should not appear among later-laid eggs of the clutch in species with hatching asynchrony (Carranza 2004; see also Albrecht 2000), as delayed hatching of later-laid eggs severely disadvantages the offspring from those eggs in sibling competition with their earlier-hatching, older nest-mates (e.g. Magrath 1990; Krist *et al.* 2004). However, we did not find these predicted effects of egg size and laying order on the sex of individual offspring within clutches, despite the large sample size for testing these hypotheses. We also found no evidence that females adjusted the sex of individual offspring in relation to paternity, as also reported for Swedish (Sheldon & Ellegren 1996) and Hungarian (Rosivall *et al.* 2009) populations. If extra-pair offspring are of higher quality than within-pair offspring, the Trivers-Willard model predicts that extra-pair young should more likely be male than their within-pair half-siblings. However, Krist & Munclinger (2011) recently demonstrated that offspring sired by extra-pair males exhibit no inherent superiority over within-pair offspring in our study population. In fact, within-pair offspring recruit to the breeding population at a slightly higher rate than their extra-pair half-siblings (Krist & Munclinger 2011; see also Sardell *et al.* 2011). Thus, lack of an effect of paternity on offspring sex is not unexpected in this population.

One potential explanation for the lack of sex bias with respect to egg size and laying order may involve constraints on reproductive females. In contrast to clutch-level sex ratios, within-clutch sex adjustment in relation to egg size, laying order or paternity may require even finer mechanisms of sex determination that may not have evolved in this species, for example, if selection has simply not favoured the ability to adjust the sex of individual offspring in relation to differences in egg size or laying order. However, sex-differences in egg size (reviewed in Martyka *et al.* 2010), laying order (e.g. Ležalová *et al.* 2005; Bowers *et al.* 2011) and even paternity (Schwarzová *et al.* 2008; Johnson *et al.* 2009)

have been found in many species, suggesting that birds are not generally constrained from within-clutch sex allocation. Moreover, recent work on the house finch, *Carpodacus mexicanus*, has shown that maternal steroid hormones can act as mediators linking the order, size and sex of oocytes that are produced (Young & Badyaev 2004; Badyaev *et al.* 2008), and suggests that the ability to adjust the sex of individual offspring according to variation in egg size or laying order may not be directly subject to natural selection, but is influenced during the ontogeny of reproductive systems (Badyaev 2011). Greater integration of proximate mechanisms and ecological processes covering more populations and species will be needed to draw general conclusions about the fitness benefits, mechanisms and possible constraints in the evolution of within-clutch sex allocation.

Acknowledgements

References

Akçay E, Roughgarden J (2007) Extra-pair paternity in birds: review of the genetic benefits. *Evolutionary Ecology Research*, **9**, 855–868.

Albrecht DJ (2000) Sex ratio manipulation within broods of house wrens, *Troglodytes aedon*. *Animal Behaviour*, **59**, 1227–1234.

Badyaev AV (2011) How do precise adaptive features arise in development? Examples with evolution of context-specific sex ratios and perfect beaks. *Auk*, **128**, 467–474.

Badyaev AV, Hill GE, Beck ML *et al.* (2002) Sex-biased hatching order and adaptive population divergence in a passerine bird. *Science*, **295**, 316–318.

Badyaev AV, Young RL, Hill GE, Duckworth RA (2008) Evolution of sex-biased maternal effects in birds. IV. Intra-ovarian growth dynamics can link sex determination and sex-specific acquisition of resources. *Journal of Evolutionary Biology*, **21**, 449–460.

Bantock TM, Prys-Jones RP, Lee PLM (2008) New and improved molecular sexing methods for museum bird specimens. *Molecular Ecology Resources*, **8**, 519–528.

Bateman AJ (1948) Intra-sexual selection in *Drosophila*. *Heredity*, **2**, 349–368.

Biard C, Gil D, Karadaş F *et al.* (2009) Maternal effects mediated by antioxidants and the evolution of carotenoid-based signals in birds. *American Naturalist*, **174**, 696–708.

Blank JL, Nolan V Jr (1983) Offspring sex ratio in red-winged blackbirds is dependent on maternal age. *Proceedings of the National Academy of Sciences of the United States of America*, **80**, 6141–6145.

Bonisoli-Alquati A, Boncoraglio G, Caprioli M, Saino N (2011) Birth order, individual sex and sex of competitors determine the outcome of conflict among siblings over parental care. *Proceedings of the Royal Society B: Biological Sciences*, **278**, 1273–1279.

Bowers EK, Sakaluk SK, Thompson CF (2011) Adaptive sex allocation in relation to hatching synchrony and offspring quality in house wrens. *American Naturalist*, **177**, 617–629.

Bowers EK, Sakaluk SK, Thompson CF (2012a) Experimentally increased egg production constrains future reproduction of female house wrens. *Animal Behaviour*, **83**, 495–500.

Bowers EK, Smith RA, Hodges CJ, Zimmerman LM, Thompson CF, Sakaluk SK (2012b) Sex-biased terminal investment in offspring induced by maternal immune challenge in the house wren (*Troglodytes aedon*). *Proceedings of the Royal Society B: Biological Sciences*, **279**, 2891–2898.

Carranza J (2004) Sex allocation within broods: the intrabrood sharing-out hypothesis. *Behavioral Ecology*, **15**, 223–232.

Cassey P, Ewen JG, Møller AP (2006) Revised evidence for facultative sex ratio adjustment in birds: a correction. *Proceedings of the Royal Society B: Biological Sciences*, **273**, 3129–3130.

Clutton-Brock TH, Albon SD, Guinness FE (1985) Parental investment and sex differences in juvenile mortality in birds and mammals. *Nature*, **131**, 131–133.

Cockburn A, Legge S, Double MC (2002) Sex-ratios in birds and mammals: can the hypotheses be disentangled? In: *Sex Ratios: Concepts and Research Methods* (ed. Hardy ICW), pp. 266–286. Cambridge University Press, Cambridge.

Cohen J (1988) Statistical Power Analysis for Behavioral Sciences. Lawrence Erlbaum Associates, Inc., Hillsdale, New Jersey.

Cordero PJ, Griffith SC, Aparicio JM, Parkin DT (2000) Sexual dimorphism in house sparrow eggs. *Behavioral Ecology and Sociobiology*, **48**, 353–357.

Cordero PJ, Viñuela J, Aparicio JM, Veiga JP (2001) Seasonal variation in sex ratio and sexual egg dimorphism favouring daughters in first clutches of the spotless starling. *Journal of Evolutionary Biology*, **14**, 829–834.

Dijkstra C, Daan S, Buker JB (1990) Adaptive seasonal variation in the sex ratio of kestrel broods. *Functional Ecology*, **4**, 143–147.

Dowling DK, Mulder RA (2006) Combined influence of maternal and paternal quality on sex allocation in red-capped robins. *Journal of Evolutionary Biology*, **19**, 440–449.

Dubois NS, Kennedy ED, Getty T (2006) Surplus nest boxes and the potential for polygyny affect clutch size and offspring sex ratio in house wrens. *Proceedings of the Royal Society B: Biological Sciences*, **273**, 1751–1757.

Ellegren H, Gustafsson L, Sheldon BC (1996) Sex ratio adjustment in relation to paternal attractiveness in a wild bird population. *Proceedings of the National Academy of Sciences of the United States of America*, **93**, 11723–11728.

Ewen JG, Cassey P, King RAR (2003) Assessment of the randomization test for binomial sex-ratio distributions in birds. *Auk*, **120**, 62–68.

Fawcett TW, Kuijper B, Pen I, Weissing FJ (2007) Should attractive males have more sons? *Behavioral Ecology*, **18**, 71–80.

Fiala KL (1980) On estimating the primary sex ratio from incomplete data. *American Naturalist*, **115**, 442–444.

Forsyth DM, Tustin KG, Gaillard J-M, Loison A (2004) Fetal sex ratio variation in the highly polygynous Himalayan tahr: evidence for differential male mortality. *Behavioral Ecology*, **15**, 572–578.

Frank SA (1990) Sex allocation theory for birds and mammals. *Annual Review of Ecology and Systematics*, **21**, 13–55.

Freed LA, Cann RL, Diller K (2009) Sexual dimorphism and the evolution of seasonal variation in sex allocation in the Hawaii akepa. *Evolutionary Ecology Research*, **11**, 731–757.

Gaillard J-M, Festa-Bianchet M, Delorme D, Jorgenson J (2000) Body mass and individual fitness in female ungulates: bigger is not always better. *Proceedings of the Royal Society of London, Series B*, **267**, 471–477.

Garamszegi LZ, Rosivall B, Hegyi G, Szöllősi E, Török J, Eens M (2006) Determinants of male territorial behavior in a Hungarian collared flycatcher population: plumage traits of residents and challengers. *Behavioral Ecology and Sociobiology*, **60**, 663–671.

Gil D (2008) Hormones in avian eggs: physiology, ecology and behavior. *Advances in the Study of Behavior*, **38**, 337–398.

Gowaty PA (1991) Facultative manipulation of sex ratios in birds: rare or rarely observed? *Current Ornithology*, **8**, 141–171.

Griffiths R, Double MC, Orr K, Dawson RJG (1998) A DNA test to sex most birds. *Molecular Ecology*, **7**, 1071–1075.

Gustafsson L, Qvarnström A, Sheldon BC (1995) Trade-offs between life-history traits and a secondary sexual character in male collared flycatchers. *Nature*, **375**, 311–313.

Harmsen R, Cooke F (1983) Binomial sex-ratio distribution in the lesser snow goose: a theoretical enigma. *American Naturalist*, **121**, 1–8.

Hasselquist D, Kempenaers B (2002) Parental care and adaptive brood sex ratio manipulation in birds. *Philosophical Transactions of the Royal Society of London, Series B*, **357**, 363–372.

Hegyi G, Török J, Tóth L (2002) Qualitative population divergence in proximate determination of a sexually selected trait in the collared flycatcher. *Journal of Evolutionary Biology*, **15**, 710–719.

Hegyi G, Rosivall B, Szöllősi E, Hargitai R, Eens M, Török J (2007) A role for female ornamentation in the facultatively polygynous mating system of collared flycatchers. *Behavioral Ecology*, **18**, 1116–1122.

Hegyi G, Szöllősi E, Jenni-Eiermann S, Török J, Eens M, Garamszegi LM (2010) Nutritional correlates and mate acquisition role of multiple sexual traits in male collared flycatchers. *Naturwissenschaften*, **97**, 567–576.

Hewison AJM, Gaillard J-M (1999) Successful sons or advantaged daughters? The Trivers-Willard model and sex-biased maternal investment in ungulates. *Trends in Ecology and Evolution*, **14**, 229–234.

Hewison AJM, Gaillard J-M, Blanchard P, Festa-Bianchet M (2002) Maternal age is not a predominant determinant of progeny sex ratio variation in ungulates. *Oikos*, **98**, 334–339.

Hoyt DF (1979) Practical methods of estimating volume and fresh weight of bird eggs. *Auk*, **96**, 73–77.

Johnson LS, Thompson CF, Sakaluk SK *et al.* (2009) Extra-pair young in house wren broods are more likely to be male than female. *Proceedings of the Royal Society B: Biological Sciences*, **276**, 2285–2289.

Julliard R (2000) Sex-specific dispersal in spatially varying environments leads to habitat-dependent evolutionarily stable offspring sex ratios. *Behavioral Ecology*, **11**, 421–428.

Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, **16**, 1099–1106.

Koskela E, Mappes T, Niskanen T, Rutkowska J (2009) Maternal investment in relation to sex ratio and offspring number in a small mammal – a case for Trivers and Willard theory? *Journal of Animal Ecology*, **78**, 1007–1014.

Krackow S, Neuhäuser M (2008) Insights from complete-incomplete brood sex-ratio disparity. *Behavioral Ecology and Sociobiology*, **62**, 469–477.

Krackow S, Meelis E, Hardy ICW (2002) Analysis of sex ratio variances and sequences of sex allocation. In: *Sex Ratios: Concepts and Research Methods* (ed. Hardy ICW), pp. 112–131. Cambridge University Press, Cambridge.

Král M, Bičík V (1990) Structure and dispersal of collared flycatcher population (*Ficedula albicollis* Temm.) in Lower Jeseník mountains. *Acta Universitatis Palackianae Olomucensis, Biologica*, **30**, 161–173.

Krist M (2006) Should mothers in poor condition invest more in daughter than in son? *Ethology Ecology & Evolution*, **18**, 241–246.

Krist M (2009) Short- and long-term effects of egg size and feeding frequency on offspring quality in the collared flycatcher (*Ficedula albicollis*). *Journal of Animal Ecology*, **78**, 907–918.

Krist M (2011) Egg size and offspring quality: a meta-analysis in birds. *Biological Reviews*, **86**, 692–716.

Krist M, Munclinger P (2011) Superiority of extra-pair offspring: maternal but not genetic effects as revealed by a mixed cross-fostering design. *Molecular Ecology*, **20**, 5074–5091.

Krist M, Remeš V, Uvírová L, Nádvorník P, Bureš S (2004) Egg size and offspring performance in the collared flycatcher (*Ficedula albicollis*): a within-clutch approach. *Oecologia*, **140**, 52–60.

Krist M, Nádvorník P, Uvírová L, Bureš S (2005) Paternity covaries with laying and hatching order in the collared flycatcher *Ficedula albicollis*. *Behavioral Ecology and Sociobiology*, **59**, 6–11.

Ležalová R, Tkadlec E, Oborník M, Šimek J, Honza M (2005) Should males come first? The relationship between offspring hatching order and sex in the black-headed gull *Larus ridibundus*. *Journal of Avian Biology*, **36**, 478–483.

Lindström J (1999) Early development and fitness in birds and mammals. *Trends in Ecology and Evolution*, **14**, 343–348.

Magrath RD (1990) Hatching asynchrony in altricial birds. *Biological Reviews*, **65**, 587–622.

Magrath MJL, Van Lieshout E, Pen I, Visser GH, Komdeur J (2007) Estimating expenditure on male and female offspring in a sexually size-dimorphic bird: a comparison of different methods. *Journal of Animal Ecology*, **76**, 1169–1180.

Martyka R, Rutkowska J, Dybek-Karpiuk A, Cichoń M, Walasz K (2010) Sexual dimorphism of egg size in the European blackbird *Turdus merula*. *Journal of Ornithology*, **151**, 827–831.

Mead PS, Morton ML, Fish BE (1987) Sexual dimorphism in egg size and implications regarding facultative manipulation of sex in mountain white-crowned sparrows. *Condor*, **89**, 798–803.

Merilä J, Kruuk LEB, Sheldon BC (2001) Cryptic evolution in a wild bird population. *Nature*, **412**, 76–79.

Monaghan P (2008) Early growth conditions, phenotypic development and environmental change. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **363**, 1635–1645.

Nager RG, Monaghan P, Griffiths R, Houston DC, Dawson R (1999) Experimental demonstration that offspring sex ratio varies with maternal condition. *Proceedings of the National Academy of Sciences of the United States of America*, **96**, 570–573.

Nager RG, Monaghan P, Houston DC (2000) Within-clutch trade-offs between the number and quality of eggs: experimental manipulations in gulls. *Ecology*, **81**, 1339–1350.

Nakagawa S, Cuthill IC (2007) Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biological Reviews*, **82**, 591–605.

Neto JM, Hansson B, Hasselquist D (2011) Sex allocation in Savi's warblers *Locustella luscinioides*: multiple factors affect seasonal trends in brood sex ratios. *Behavioral Ecology and Sociobiology*, **65**, 297–304.

Nilsson J-Å (1989) Establishment of juvenile marsh tits in winter flocks: an experimental study. *Animal Behaviour*, **38**, 586–595.

Nilsson J-Å (1990) Establishment success of experimentally delayed juvenile marsh tits. *Ethology*, **85**, 73–79.

Norris K (1993) Seasonal variation in the reproductive success of blue tits: an experimental study. *Journal of Animal Ecology*, **62**, 287–294.

Pärt T (1990) Natal dispersal in the collared flycatcher: possible causes and reproductive consequences. *Ornis Scandinavica*, **21**, 83–88.

Pärt T, Qvarnström A (1997) Badge size in collared flycatchers predicts outcome of male competition over territories. *Animal Behaviour*, **54**, 893–899.

Pen I, Weissing FJ, Daan S (1999) Seasonal sex ratio trend in the European kestrel: an evolutionarily stable strategy analysis. *American Naturalist*, **153**, 384–397.

Postma E, Heinrich F, Koller U *et al.* (2011) Disentangling the effect of genes, the environment and chance on sex ratio variation in a wild bird population. *Proceedings of the Royal Society B: Biological Sciences*, **278**, 2996–3002.

Przybylo R, Sheldon BC, Merilä J (2000) Patterns of natural selection on morphology of male and female collared flycatchers (*Ficedula albicollis*). *Biological Journal of the Linnean Society*, **69**, 213–232.

Qvarnström A, Pärt T, Sheldon BC (2000) Adaptive plasticity in mate preference linked to differences in reproductive effort. *Nature*, **405**, 344–347.

Qvarnström A, Sheldon BC, Pärt T, Gustafsson L (2003) Male ornamentation, timing of breeding, and cost of polygyny in the collared flycatcher. *Behavioral Ecology*, **14**, 68–73.

Robert KA, Schwanz LE, Mills HR (2010) Offspring sex varies with maternal investment ability: empirical demonstration based on cross-fostering. *Biology Letters*, **6**, 242–245.

Rosenthal R (1994) Parametric measures of effects size. In: *The Handbook of Research Synthesis* (eds Cooper H, Hedges LV) pp. 231–244. Russell Sage Foundation, New York, New York.

Rosivall B, Török J, Hasselquist D, Bensch S (2004) Brood sex ratio adjustment in collared flycatchers (*Ficedula albicollis*):

results differ between populations. *Behavioral Ecology and Sociobiology*, **56**, 346–351.

Rosivall B, Szöllősi E, Török J (2005) Maternal compensation for hatching asynchrony in the collared flycatcher *Ficedula albicollis*. *Journal of Avian Biology*, **36**, 531–537.

Rosivall B, Szöllősi E, Hasselquist D, Török J (2009) Effects of extrapair paternity and sex on nestling growth and condition in the collared flycatcher, *Ficedula albicollis*. *Animal Behaviour*, **77**, 611–617.

Rosivall B, Szöllősi E, Hasselquist D, Török J (2010) Males are sensitive – sex-dependent effect of rearing conditions on nestling growth. *Behavioral Ecology and Sociobiology*, **64**, 1555–1562.

Rutkowska J, Koskela E, Mappes T, Speakman JR (2011) A trade-off between current and future sex allocation revealed by maternal energy budget in a small mammal. *Proceedings of the Royal Society B: Biological Sciences*, **278**, 2962–2969.

Saltz D (2001) Progeny sex ratio variation in ungulates: maternal age meets environmental perturbation of demography. *Oikos*, **94**, 377–384.

Saltz D, Kotler BP (2003) Maternal age is a predominant determinant of progeny sex ratio variation in ungulates: a reply to Hewison et al. *Oikos*, **101**, 646–648.

Sardell RJ, Arcese P, Keller LF, Reid JM (2011) Sex-specific differential survival of extra-pair and within-pair offspring in song sparrows, *Melospiza melodia*. *Proceedings of the Royal Society B: Biological Sciences*, **278**, 3251–3259.

Schwarzová L, Šimek J, Coppack T, Tryjanowski P (2008) Male-biased sex of extra pair young in the socially monogamous red-backed shrike *Lanius collurio*. *Acta Ornithologica*, **43**, 235–239.

Seutin G, White BN, Boag PT (1991) Preservation of avian blood and tissue samples for DNA analyses. *Canadian Journal of Zoology*, **69**, 82–90.

Sheldon BC, Ellegren H (1996) Offspring sex and paternity in the collared flycatcher. *Proceedings of the Royal Society of London, Series B*, **263**, 1017–1021.

Sheldon BC, Ellegren H (1999) Sexual selection resulting from extrapair paternity in collared flycatchers. *Animal Behaviour*, **57**, 285–298.

Sheldon BC, West SA (2004) Maternal dominance, maternal condition, and offspring sex ratio in ungulate mammals. *American Naturalist*, **163**, 40–54.

Sheldon BC, Merilä J, Qvarnström A, Gustafsson L, Ellegren H (1997) Paternal genetic contribution to offspring condition predicted by size of male secondary sexual character. *Proceedings of the Royal Society of London, Series B*, **264**, 297–302.

Sheldon BC, Merilä J, Lindgren G, Ellegren H (1998) Gender and environmental sensitivity in nestling collared flycatchers. *Ecology*, **79**, 1939–1948.

Sheldon BC, Andersson S, Griffith SC, Örnborg J, Sendecka J (1999) Ultraviolet colour variation influences blue tit sex ratios. *Nature*, **402**, 874–877.

Sokal RR, Rohlf FJ (1995) Biometry, 3rd edn. W. H Freeman and Company, New York, New York, USA.

Taff CC, Freeman-Gallant CR, Dunn PO, Whittingham LA (2011) Relationship between brood sex ratio and male ornaments depends on male age in a warbler. *Animal Behaviour*, **81**, 619–625.

Tilgar V, Mänd R, Kilgas P, Mägi M (2010) Long-term consequences of early ontogeny in free-living great tits *Parus major*. *Journal of Ornithology*, **151**, 61–68.

Török J, Hegyi G, Garamszegi LZ (2003) Depigmented wing patch size is a condition-dependent indicator of viability in male collared flycatchers. *Behavioral Ecology*, **14**, 382–388.

Trivers RL, Willard DE (1973) Natural selection of parental ability to vary the sex ratio of offspring. *Science*, **179**, 90–92.

Uller T (2006) Sex-specific sibling interactions and offspring fitness in vertebrates: patterns and implications for maternal sex ratios. *Biological Reviews*, **81**, 207–217.

Verhulst S, Nilsson J-Å (2008) The timing of birds' breeding seasons: a review of experiments that manipulated timing of breeding. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **363**, 399–410.

Wang J (2004) Estimating pairwise relatedness from dominant genetic markers. *Molecular Ecology*, **13**, 3169–3178.

West SA (2009) Sex Allocation. Princeton University Press, Princeton, New Jersey.

West SA, Sheldon BC (2002) Constraints in the evolution of sex ratio adjustment. *Science*, **295**, 1685–1688.

Westerdahl H, Bensch S, Hansson B, Hasselquist D, von Schantz T (1997) Sex ratio variation among broods of great reed warblers *Acrocephalus arundinaceus*. *Molecular Ecology*, **6**, 543–548.

Wilkin TA, Sheldon BC (2009) Sex differences in the persistence of natal environmental effects on life histories. *Current Biology*, **19**, 1998–2002.

Williams GC (1979) The question of adaptive sex ratio in outcrossed vertebrates. *Proceedings of the Royal Society of London, Series B*, **205**, 567–580.

Wilson K, Hardy ICW (2002) Statistical analysis of sex-ratios: an introduction. In: *Sex Ratios: Concepts and Research Methods* (ed. Hardy ICW), pp. 48–92. Cambridge University Press, Cambridge.

Young RL, Badyaev AV (2004) Evolution of sex-biased maternal effects in birds: I. Sex-specific resource allocation among simultaneously growing oocytes. *Journal of Evolutionary Biology*, **17**, 1355–1366.

Data accessibility

All data analysed as part of this paper including parental and offspring genotypes have been deposited in Dryad data repository (doi:10.5061/dryad.pt384)

Supporting information

Additional supporting information may be found in the online version of this article.

Appendix S1. Results of backward elimination procedure in the model for clutch sex ratios.

Článek 13: Krist, M. & Munclinger, P. (2015) Context dependence of maternal effects: testing assumptions of optimal egg size, differential, and sex allocation models. *Ecology*, 96, 2726-2736.

Context dependence of maternal effects: testing assumptions of optimal egg size, differential, and sex allocation models

Miloš Krist[1,2,4] and Pavel Munclinger[3]

[1]*Department of Zoology and Laboratory of Ornithology, Faculty of Science, Palacký University, tř. Svobody 26, 771 46 Olomouc, Czech Republic*
[2]*Museum of Natural History, nám. Republiky 5, 771 73 Olomouc, Czech Republic*
[3]*Department of Zoology, Faculty of Science, Charles University in Prague, Viničná 7, 128 00 Praha, Czech Republic*

Abstract. If offspring develop in adverse conditions, the maternal component of their phenotypic variation might increase due to the stronger dependence of offspring traits on parental investment. This should result in increased parental investment to individual offspring, as assumed by the model of optimal egg size. The opposite pattern, i.e., stronger dependence of offspring fitness on parental investment and consequently larger parental investment under good conditions is assumed by both the theory of differential allocation if attractive males provide material benefits, and reproductive compensation if they invest less into paternal care. Another influential idea is the Trivers-Willard model, which assumes sex-specific dependence of offspring fitness on parental investment. Here we tested these ideas by examining the effects of egg size on offspring fitness across many postnatal contexts in the Collared Flycatcher *Ficedula albicollis*. We employed a cross-fostering design that generated variation in egg size within nests and used brood means of fledgling mass as a functional measure of the quality of rearing conditions. Effects of egg size on three offspring traits, including lifetime reproductive success of recruits, were more pronounced in low-quality broods. These results support the assumption of the model of optimal egg size. Based on female preference for males providing material benefits, this pattern could support differential allocation, if attractive males invest less in paternal care, or reproductive compensation, if they invest more. By comparison, we did not find any evidence for sex specificity of fitness returns that might explain sex monomorphism of egg size in this species. The challenge for future studies will be the integration of components of parental investment and offspring fitness into their global measures and testing how the former affects the latter across gradients of postnatal conditions.

Key words: Collared Flycatcher; context dependence; environmental quality; Ficedula albicollis; maternal effects; parental care; postnatal conditions; sex allocation; sexual selection; Velký Kosíř, Czech Republic.

Introduction

Environmental quality is an increasingly popular topic in evolutionary ecology due to its potential effect on the magnitude and direction of selection (Robinson et al. 2012) and components of phenotypic variation (Hoffmann and Merilä 1999), and thus also on evolutionary change. Phenotypic variation may be greater in harsh environments due to a genotype × environment interaction, with genetic differences manifested only in poor-quality environments. This effect has been found for genetic variation due to inbreeding (Fox and Reed 2011) and extra-pair paternity (Schmoll 2011). Similarly, the environmental component of phenotypic variance might also be higher in adverse conditions. For example, offspring fitness may be more

dependent on initial resources received from parents if the overall quality of rearing conditions is low. This would lead to an increase in the maternal component of variance that contributes to environmental variance from the offspring's perspective (Mousseau and Fox 1998).

This idea has been elaborated in life history theory. The optimal solution for the trade-off between initial offspring size and number is likely to be contingent on conditions that offspring experience in their later developmental stages up to their independence. Offspring fitness should be more dependent on their initial size if postnatal conditions are poor (Fig. 1B); parents are then predicted to invest in egg or neonatal size at the expense of reduced fecundity (Brockelman 1975, Parker and Begon 1986). This prediction has recently been confirmed in fish (Bashey 2008, Rollinson and Hutchings 2013a, Riesch et al. 2014). Some studies also tested the assumption of these models that offspring fitness only strongly depends on egg or neonatal size in adverse conditions. Although most of these tests supported this

assumption (review in Krist 2011), the opposite effect has also been found across many taxa (for amphibians, see Semlitsch and Gibbons [1990]; for reptiles, see Svensson and Sinervo [2000]; for birds, see Krist [2011]; and for mammals, see Oksanen et al. [2003]). It is not clear whether this variability could be explained by random sampling variance or species-specific adaptations, or if it is a sign of a more general underlying process. Moreover, most tests have been performed in only a few environments and have used a surrogate of offspring fitness. Studies comparing egg-size effects on offspring global fitness (Krist 2011, Williams 2012) across many replicates of postnatal conditions (Rollinson and Hutchings 2013a) could offer a more reliable and general picture of how parents make trade-off decisions between offspring size and number under a variety of conditions.

More recently, and independently from the considerations of optimal offspring size, theories of differential allocation and reproductive compensation have been developed (Sheldon 2000, Gowaty et al. 2007, Gowaty 2008). These theories share one common feature with that of optimal offspring size: they also assume (largely implicitly, but see Kindsvater and Alonzo 2014) that benefits from maternal investment (e.g., egg size) into offspring depend on external conditions, in this case, the attractiveness of the male partner. Unlike rearing conditions in the offspring size model, here the quality of the partner may also contribute to offspring genetic quality, but the theory is equally valid when mates provide only material benefits (Sheldon 2000, Kindsvater and Alonzo 2014). These material benefits may be, for example, better food supply in the defended territory or more intensive paternal care. These types of benefits would contribute to the overall quality of postnatal conditions, making these cases directly comparable to the offspring size model.

However, both differential allocation and reproductive compensation may assume the opposite pattern of context dependence than the offspring size model, i.e., higher fitness returns from initial parental investment under good postnatal conditions (Fig. 1C). Whether this is the case for differential allocation or reproductive compensation depends on whether attractive or unattractive males provide material benefits. If these benefits are provided by attractive males, Fig. 1C describes the assumption of differential allocation, and Fig. 1B describes that of reproductive compensation. The opposite is true (Fig. 1C for reproductive compensation and Fig. 1B for differential allocation) if material benefits are provided by non-preferred males. In the context of differential allocation/compensation, no tests of the assumptions of context dependence of fitness returns from initial parental investment have been conducted. By contrast, predictions of these models have often been tested and more empirical support has been found for differential allocation (Horváthová et al. 2012), which is in line with the theoretical model
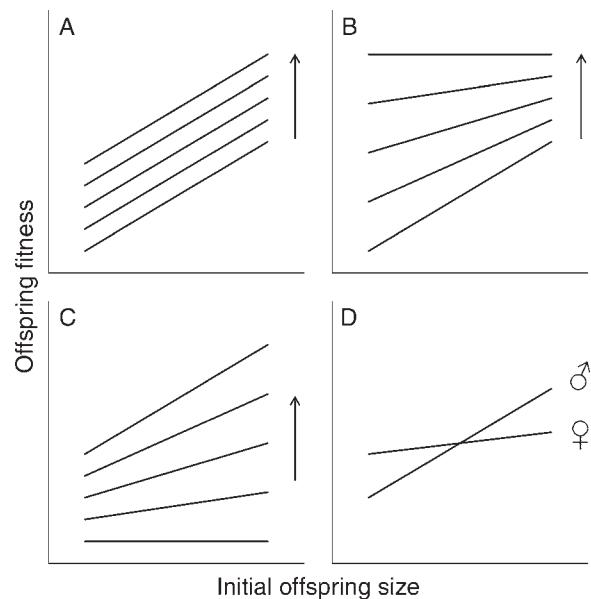

FIG. 1. Several possible scenarios of the dependence of offspring fitness on their initial size in relation to postnatal conditions: (A–C) quality of the environment or parental care or (D) offspring sex. Each line shows this relationship in a specific postnatal condition (e.g., particular quality of parental care). Arrows indicate increasing quality of postnatal conditions. (A) No interaction of initial offspring size and postnatal conditions on offspring fitness. (B) Initial offspring size is more important for fitness in poor postnatal conditions as assumed by the model of optimal egg size. This pattern is also assumed by the theory of reproductive compensation if attractive males provide more parental care or differential allocation if they provide less care. (C) Initial offspring size is more important for fitness under good postnatal conditions. This pattern is assumed by the theory of differential allocation if attractive males provide more parental care, or reproductive compensation if they provide less care. (D) Initial offspring size has sex-specific effects on fitness as assumed by the Trivers-Willard model. Linear relationships are depicted for the sake of clarity although, in reality, offspring fitness returns are likely to be nonlinear (Krist 2006, Rollinson and Hutchings 2013b).

proposed by Harris and Uller (2009), but see Ratikainen and Kokko (2010) and Kindsvater and Alonzo (2014). Clearly, more empirical as well as theoretical research is needed to resolve these controversies.

Conceptually similar to parental allocation strategies in relation to environmental conditions is the popular Trivers and Willard (1973) model of sex allocation. Here fitness returns from initial parental investment depend differentially on offspring sex (Fig. 1D). This difference in fitness returns selects for bias in primary sex ratio as well as for subsequent parental favoritism of the more sensitive sex (Krist 2006). Similarly to differential allocation and reproductive compensation, a number of studies have tested predictions of the Trivers and Willard model (West 2009), but only a few have been concerned with its assumption of sex-specific fitness functions. We know of only one study that demonstrat-

TABLE 1. Results of final models relating offspring traits to their random-effects predictors for the Collared Flycatcher *Ficedula albicollis*, in the Velký Kosíř area, Czech Republic.

Offspring trait	Nest of origin		Nest of rearing		Residual	
	Estimate	SE	Estimate	SE	Estimate	SE
Body mass at 6 days	0.0959	0.0359	1.22	0.198	0.554	0.0440
Body mass at 13 days	0.113	0.0301	0		0.590	0.0356
Tarsus length	0.0614	0.0123	0.0438	0.0100	0.144	0.00972
Wing length	0.0101	0.00232	0.0521	0.00762	0.0304	0.00207
PHA immunity	0.00519	0.00210	0.00224	0.00153	0.0218	0.00254
Hatchability	0.826	0.386	0.131	0.345		
Fledging success	0.507	0.360	0.736	0.375		
Fledging to recruitment	0		0.150	0.160		
Egg to recruitment	0		0.168	0.162		
Fitness	0.103	0.0382	0.0503	0.0312		

ed sex-specific effects of natal environment on offspring fitness (Clutton-Brock et al. 1984).

Here we employed a novel cross-fostering design to test assumptions of differential allocation, Trivers-Willard, and optimal egg size models. Specifically, we tested whether prenatal maternal investment (egg size) affects performance of male and female offspring differently across multiple postnatal contexts. In addition to some traditional short-term estimates of offspring quality, we also measured their lifetime reproductive success.

METHODS

Field and laboratory methods

We conducted this study in the Velký Kosíř area (49°32′ N, 17°04′ E, 300–400 m above sea level), Czech Republic, during 2006–2014. Study sites were on the slopes of a hill and were covered by oak (*Quercus petraea*) forest. About 400 nest boxes were provided in which about 80–100 pairs of our study species, the Collared Flycatcher (*Ficedula albicollis*), bred every year. The other common species inhabiting nest boxes were Great (*Parus major*) and Blue (*Cyanistes caeruleus*) Tits. Velký Kosíř is surrounded by agricultural landscape, which might explain the high rate of local recruitment (see Krist 2009) of flycatchers, as the opportunities to disperse in adjacent areas are limited.

The Collared Flycatcher is nearly sexually monomorphic in size but dimorphic in color. Females are brownish and males are contrastingly black and white. A small proportion of males mate polygynously (Qvarnström et al. 2003). Most males are socially monogamous, but extra-pair paternities are common (Krist and Munclinger 2011), which probably increases variance in male reproductive success over that of females. Moreover, male nestlings may be more sensitive to environmental conditions (Rosivall et al. 2010; but see Sheldon et al. 1998), which also suggests stronger dependence of their fitness on parental investment. Females solely incubate clutches of about six eggs, but both parents feed nestlings (see Plate 1) for 14 days, after which period they fledge.

Because our intention was to test for context dependence of egg-size effects, we needed to evaluate the quality of the postnatal conditions that chicks experience. As this species spends about two weeks in the nest after hatching, we decided to classify the quality of these microenvironments. One possible way to achieve this could be to use a proximate variable describing the quality of rearing conditions, such as food supply in the territory or chick feeding rates. However, by employing these proximate measures, we could miss relevant variables that actually determine the postnatal conditions of nestlings. Therefore we decided to use a functional measure of the quality of rearing conditions. We selected the mean body mass of fledglings as our estimate of the quality of postnatal conditions from the several possible performance traits. First, this variable is largely determined by the amount of food that parents deliver to their young, with only a small genetic component, which is in contrast to other fitness-related traits, such as tarsus length, that are largely heritable (Krist 2009; Table 1). Second, fledging mass is predictive of chick survival after fledging (Lindén et al. 1992). For these reasons, we considered that broods with heavy fledglings experienced good postnatal conditions, although we were not able to determine whether these were caused by, for example, a high-quality food supply or parental ability to deliver food at a high rate. Fortunately, for the purposes of this study (i.e., context dependence of early maternal effects), we did not need to distinguish between these mechanisms, because the theoretical models that we were testing also do not discriminate between purely environmental and maternal causes of the varying quality of postnatal conditions (see Hoffmann and Merilä 1999, Fox 2000, Agosta 2008).

Moreover, maternal effects such as provisioning behavior can be considered a special type of environmental variation from the offspring's perspective (Lynch and Walsh 1998, Mousseau and Fox 1998). Consequently, mean fledging mass may be considered an estimate of the quality of the offspring's environment in a broader sense. We need to stress, however, that although we considered mean fledging mass as an indicator of postnatal conditions, some prenatal effects such as egg composition (e.g., concentration of andro-

PLATE 1. Six-day-old Collared Flycathers begging for food. Photo credit: M. Krist, 7 June 2007.

gens and carotenoids) could also contribute to variation in fledging mass. However, the contribution of egg composition is likely to be small compared to the postnatal effects of territory quality, ambient conditions, and parental provisioning. Unlike egg composition, we made mean fledging mass independent of egg size through a combination of experimental and statistical techniques.

To test whether egg size has a stronger effect on individual offspring traits if they live in a poor brood microenvironment, we performed a mixed cross-fostering experiment in 2006–2009. We collected eggs on the day they were laid and exchanged them for dummy eggs. After a clutch was completed, we supplied each nest with the number of eggs equivalent to its original clutch. Each of these fostered eggs originated from a different donor nest. Use of this design enabled us to assign chicks to eggs from which they hatched, which is a necessary condition to test egg-size effects within broods. We were able to assign every chick to an egg by molecular methods because each chick in a nest had unique genetic mother. In short, maternity was resolved based on the comparison of genotypes of offspring and candidate mothers at eight microsatellite loci (for more details, see Krist and Munclinger 2011). Second, by mixing eggs from different nests, we intended to increase intraclutch egg-size variation, as egg size is normally more variable between than within clutches (Christians 2002). This should have made our within-brood tests

more powerful (Svensson and Sinervo 2000, Krist 2011). Third, each nest contained a mixture of eggs from different nests after cross-fostering. Thus, genetic composition was randomized between nests and consequently should have an even smaller effect on between-nest variation in fledging mass than is the case in natural clutches. Similarly, mean fledging mass should not be affected much by mean egg size due to the minimal variation of egg size between clutches after cross-fostering (see *Results*). Finally, as we placed a whole clutch of unincubated eggs in a recipient nest at the same time, hatching asynchrony was minimized (Krist and Munclinger 2011). This adds further rigor to our test of intraclutch egg-size effects, which under natural conditions are confused by age differences between nest mates (Krist et al. 2004, Rosivall et al. 2005). In sum, our mixed cross-fostering is a relevant tool to test these theories similarly, as was recently employed another type of cross-fostering for a study of maternal effects in turtles (Mitchell et al. 2013).

At the nestling age of 6 days, we weighed, banded, and blood sampled nestlings by tarsal venipuncture for genetic analyses. For this reason, we also collected all unhatched eggs and dead chicks. In addition to genotyping offspring at microsatellite loci, we also determined their sex by amplification of sex-linked CHD introns that differ in length between the Z and W sex chromosomes (Griffiths et al. 1998, Bowers et al. 2013). At the nestling age of 12 days, we measured the

thickness of the right patagium to 0.01 mm and injected it with 0.1 mg of PHA (phytohaemagglutinin) in 20 μL of saline solution for scoring innate immunity. We applied this procedure only in 2006 and 2007. Immunity was determined as the difference between the thickness of the patagium before and 24 hours after the injection. At 13 days, when the young are close to fledging, we again weighed them, remeasured the right patagium at the injection site, and took measurements of tarsus (to 0.01 mm) and wing length (to 1 mm).

During the course of the study (2006–2014), we strived to catch all adult flycatchers with nest box traps while they fed nestlings. Consequently, we were able to estimate recruitment rate and fitness of offspring fledged from this experiment (2006–2009). Despite the fact that recruitment rate is higher in Velký Kosíř (Krist 2009; see also *Results*) compared with other intensively studied populations of the species (Gustafsson 1989), it is likely that some individuals that we were unable to recapture dispersed rather than died. For example, they could have bred in natural cavities. Thus our results of recruitment and lifetime reproductive success should be interpreted with this limitation in mind.

We measured the fitness of offspring (F_1 generation) as the lifetime number of eggs (F_2 generation) laid by female recruits; for mate recruits, we counted lifetime number of eggs laid by their social mates. Thus, this fitness measure combines both offspring survival and fecundity. We were only able to adjust our estimate of fitness for extra-pair paternity for a minority of breeding attempts by male recruits (14 of 119). In these cases, we subtracted eggs sired by other males while adding eggs sired by focal recruits in other nests (about half of the population was sampled) when estimating male fitness. Paternity was resolved using the same set of markers as for maternity. We did not attempt to use the number of F_2 offspring in later stages (fledglings, recruits) for the estimation of F_1 offspring fitness, for two reasons. First, quantitative geneticists recommend using the number of sired zygotes as an estimate of parental fitness, not the number of fledglings or recruits, because in the latter case the fitness of parents is confused with that of their offspring (Wolf and Wade 2001, Hunt and Hodgson 2010). Second, we performed some experiments in subsequent years that could affect lifetime reproductive success of F_1 only if measured at later stages of F_2 development. For example, we switched the whole clutches after clutch completion so parents did not take care of their own chicks. We also performed experiments that could affect the survival or fecundity of focal birds (e.g., manipulation of clutch or brood size). If a focal bird had been involved in this latter type of manipulation in any year (except 2014), it was excluded from the fitness analysis.

Statistical analyses

We tested whether egg-size effects on 10 fitness-related traits of offspring depend on postnatal conditions and offspring sex. Some of these offspring traits were intercorrelated, but the correlations were not extremely large. The largest correlation was between chick mass at 6 days and fledging wing length ($r = 0.731$), which still means that about half of the variation in both traits was unique. We decided to use these traditionally (see Krist 2011) measured traits instead of some orthogonal axes of variation such as principal components due to better interpretability of the former. Each statistical model included these predictors: (1) nest of origin and nest of rearing as random factors, (2) year as a categorical factor and laying date as a covariate to explain seasonal, annual, and cohort effects, (3) egg size, offspring sex, and postnatal condition as main factors. Postnatal condition was determined for each foster nest as the mean of the 13-d body mass of all chicks. We used the residuals of mean fledging mass on mean egg size to make our measure of postnatal condition fully independent of egg size (mean mass = $15.48 - 0.000665$(mean egg volume); $F_{1, 132} = 0.17$, $P = 0,678$, $R^2 < 0.01$). Only nests from which at least two young fledged were included. This is because it was only in these nests that we were able to reasonably define fledging condition. Nests with total failures after day 13 or those with only one young fledged were excluded because the mass of 13-d-old nestlings might be outlying due to starvation. In addition, we included as a predictor (4) the interaction of egg size with offspring sex and environmental quality. If these interactions were nonsignificant ($\alpha > 0.1$), we sequentially deleted them according to their P values to enhance interpretation of the main factors (Engqvist 2005).

The model for hatchability did not include offspring sex because we were unable to sex many unhatched eggs. This model and the other tests for offspring survival (fledging success, survival from fledging to recruitment) had a binary distribution of error and logit link. Models in which the response variable was offspring morphology or PHA immunity had a normal distribution of errors and identity link. Fitness had a typical zero-inflated distribution. Most young did not recruit to the breeding population and thus had zero fitness. Therefore, we followed the suggestion of Martin et al. (2005) and divided fitness into two components. First, survival of offspring from egg to recruitment was fitted as the other binary variables. This analysis included all sexed individuals ($n = 794$). We considered recruits separately ($n = 91$). Because recruits' lifetime reproductive success (which combines both offspring survival and fecundity) had a Poisson distribution, we fitted it with log link. All tests were fitted in Proc Glimmix of SAS 9.2 (2008).

To summarize mean effects of egg size on 10 offspring traits, we computed effect sizes from models without interactions. Effect sizes were determined according to the formula $r = \sqrt{[F/(F + \text{DDF})]}$, where F is an F value and DDF are error degrees of freedom (Rosenthal 1994:237). We set confidence limits for the correlation coefficient using z transformation (Sokal and Rohlf
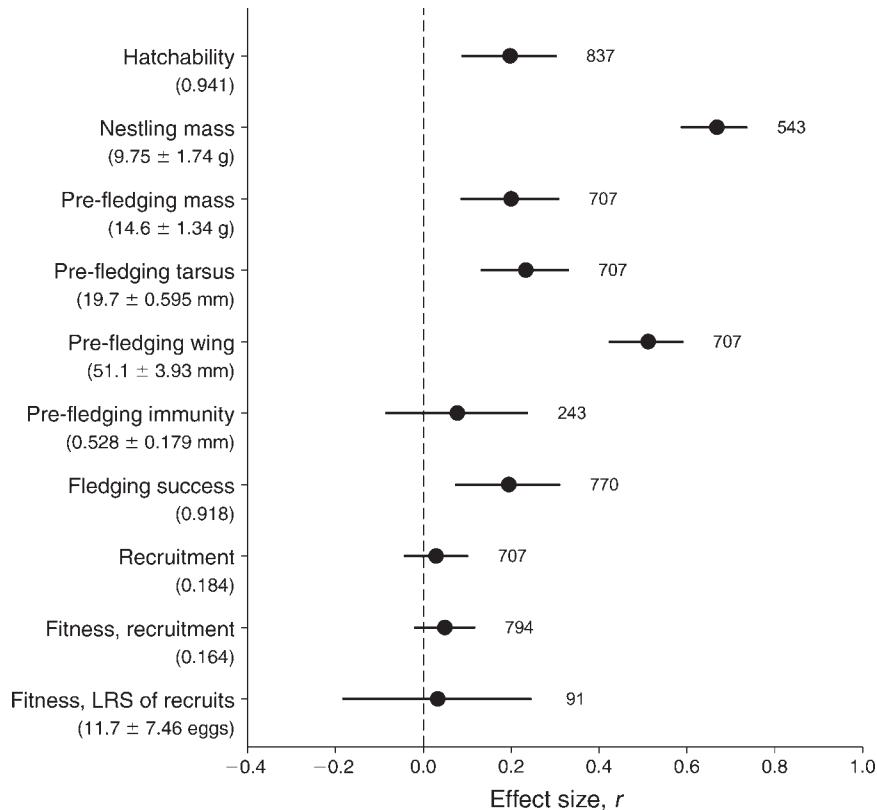
FIG. 2. Mean egg-size effects (with 95% CL) on offspring traits in the Collared Flycatcher *Ficedula albicollis*. Effect sizes were derived from models without interactive terms. Mean and SD values for offspring traits are given in parentheses. Numbers by the effect sizes indicate sample sizes. LRS is lifetime reproductive success, which combines both offspring survival and fecundity. Immunity was measured as the difference between the thickness of the patagium before and 24 h after injection with 0.1 mg of PHA (phytohaemagglutinin) in 20 μL of saline solution.

1995:577), with DDF instead of sample sizes used in that formula. This modification resulted in wider confidence intervals that better reflected true significance levels from these mixed models.

RESULTS

We cross-fostered 1026 eggs between 160 nests. Intra-clutch variation in egg size was greatly increased by our experiment. In original nests, it only explained 25.1% of the total variation in egg size, whereas this figure was 86.8% in foster nests. The coefficient of variation (mean ± SD) in egg size rose from 3.91 ± 1.91 in original clutches to 7.46 ± 2.48 in foster clutches. This experimentally caused increase in intra-clutch egg-size variation corresponds to a decrease in between-clutch variation (from 74.9% to 13.2%).

Of these 160 cross-fostered nests, 129 nests fledged at least two young and therefore enabled us to reasonably define chicks' postnatal conditions. All subsequent analyses were restricted to these nests. In this subsample of nests, 837 eggs were laid, 788 chicks hatched, 794 eggs/chicks were sexed, and 707 chicks fledged. Mean fledging mass, which we used as a measure of the quality of postnatal conditions, was highly variable between

these nests (for absolute values, range 9.9–16.4 g, 14.5 ± 1.2 g, mean ± SD; for residuals on mean egg mass, range: −4.5–2.1 g, 0.1 ± 1.2 g, mean ± SD). 130 (63 females and 67 males) of 707 fledglings recruited to the study population in subsequent years. We manipulated brood size or deployed a geolocator in at least one year for 39 of these recruits. Thus we were able to determine lifetime reproductive success for the remaining 91 recruits. Of these, only two males still bred in 2014. The rest of the recruits disappeared and probably have already died, so our estimate of their fitness should indeed reflect their lifetime reproductive success.

When averaged across all postnatal contexts, egg size had positive effects on hatchability, nestling survival, and morphology, whereas it had no effect on probability of recruitment or fitness of recruits (Fig. 2, Table 2). Some of these effects were context dependent. Egg size had positive effects on mass and tarsus length of nestlings and fitness of recruits if these were raised in low-quality conditions, but it had no such effects (Fig. 3A, B) or even slightly negative ones (Fig. 3C) when chicks were raised in high-quality conditions (Table 2). For the fitness of recruits, this pattern would hold even if the whole sample of recruits (i.e., *n* = 130) was used
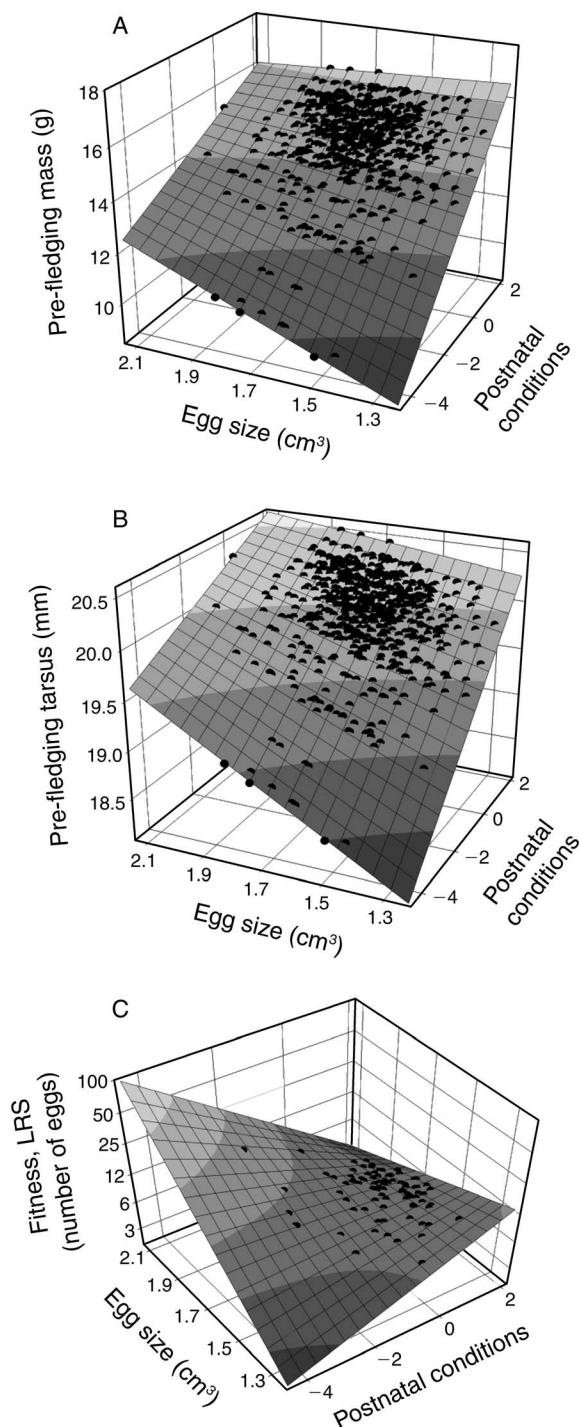
FIG. 3. Values of offspring traits predicted by the interaction between egg size and postnatal conditions (residuals from the regression of mean fledging mass in a nest on mean egg volume: 0 is average fledgling mass for the respective egg volume, 2 is body mass 2 g higher than the mean for the respective egg volume, −2 is fledgling mass 2 g lower than the mean for the respective egg volume, etc.). Points are observations projected on the predicted plot to demonstrate the distribution of data. Different shading refers to the height of the mesh on the z-axis.

(results not shown). Egg-size effects did not depend on offspring sex for any studied trait (Table 2).

DISCUSSION

In this study we generated intra-nest variation in egg size by cross-fostering and investigated whether this increased variation in prenatal conditions would interact with postnatal conditions to modify the phenotypes of individual offspring in the Collared Flycatcher. We found strong effects of egg size on offspring morphology and survival in the short term. Some of these were independent of the postnatal conditions (hatchability, fledging success, fledging wing length), whereas others were stronger in low-quality postnatal conditions (nestling mass, tarsus length). In contrast to these short-term effects, we did not find simple effects of egg size on offspring recruitment and lifetime reproductive success of recruits (Fig. 2). However, in the latter case, a significant interaction between egg size and postnatal conditions masked egg-size effects. Egg size had a positive effect on lifetime reproductive success of recruits if postnatal conditions were poor, but no effect or even a negative one if these postnatal conditions were of high quality (Fig. 3C). Negative effects of egg size in good postnatal conditions are difficult to explain. If this pattern were not simply due to sampling error, one potential explanation could be the negative side effects of yolk hormones that are typically correlated with egg size. These hormones usually induce begging and accelerate offspring growth (Gil 2003, Smiseth et al. 2011). This may be advantageous for offspring fitness in competitive postnatal conditions, but may become slightly disadvantageous in less competitive situations because rapid growth in early life may be penalized by decreased performance later in life (Metcalfe and Monaghan 2001). In contrast to context dependence of egg size in relation to postnatal conditions, we did not detect any sign of sex specificity of egg-size effects.

The stronger dependence of offspring quality on its initial size in harsh environments has often been observed in plants (Rey et al. 2004, Quero et al. 2007), invertebrates (Fox 2000, Agosta 2008), fish (Einum and Fleming 1999, Bashey 2008), and amphibians (Parichy and Kaplan 1992, Dziminski and Roberts 2006). All of these taxa provide little postnatal care, so it would be of interest to test whether a similar rule holds for taxa with postnatal care that have the potential to mitigate initial size differences. The dependence of egg-size effects on postnatal conditions such as quality of the rearing environment has been little studied in birds, which is surprising, given the popularity of the taxon for egg-size studies. From the 283 studies reviewed by Krist (2011), only three observational (Smith and Bruun 1998, Styrsky et al. 1999, Garant et al. 2007) and two experimental ones (Styrsky et al. 2000, Bonisoli-Alquati et al. 2008) tested this context dependence of egg-size effects. In these studies, the number of environments was usually limited to two and fitness measure, to short-term

TABLE 2. Results of final models relating offspring traits to their fixed-effects predictors for the Collared Flycatcher.

Offspring trait	Year F (df)	Year P	Laying date F (df)	Laying date P	Postnatal conditions F (df)	Postnatal conditions P	Offspring sex F (df)	Offspring sex P	Egg size F (df)	Egg size P	Egg size × postnatal conditions F (df)	Egg size × postnatal conditions P
Body mass at 6 days	18.12 (101)	<0.001	4.81 (100)	0.031	18.68 (94)	<0.001	8.40 (454)	0.004	169.01 (209)	<0.001		
Body mass at 13 days	0.94 (181)	0.421	2.93 (331)	0.088	39.35 (694)	<0.001	4.06 (676)	0.044	12.81 (285)	<0.001	13.65 (693)	<0.001
Tarsus length	1.94 (177)	0.124	23.74 (191)	<0.001	8.90 (680)	0.003	0.23 (620)	0.633	20.74 (344)	<0.001	3.14 (679)	0.077
Wing length	0.47 (146)	0.706	31.56 (147)	<0.001	83.61 (123)	<0.001	10.80 (575)	0.001	103.51 (291)	<0.001		
PHA immunity	8.93 (54)	0.004	4.65 (64)	0.035	4.92 (47)	0.031	2.36 (227)	0.126	0.88 (144)	0.349		
Hatchability	1.24 (166)	0.296	2.38 (176)	0.125	8.97 (193)	0.003			12.39 (304)	<0.001		
Fledging success	1.61 (116)	0.190	1.20 (187)	0.275	10.33 (66)	0.002	1.63 (762)	0.203	9.80 (248)	0.002		
Fledging to recruitment	3.58 (207)	0.015	1.49 (151)	0.224	1.88 (179)	0.172	0.02 (699)	0.890	0.59 (699)	0.441		
Egg to recruitment	4.00 (205)	0.009	1.87 (153)	0.174	2.38 (167)	0.125	1.90 (786)	0.169	0.02 (786)	0.877		
Fitness	3.16 (71)	0.030	0.04 (58)	0.843	4.54 (59)	0.037	3.22 (82)	0.077	0.43 (82)	0.514	5.63 (56)	0.021

Notes: Values associated with fixed factors are the results of type III F tests; df is the denominator degrees of freedom. Statistical directions of egg-size effects are given in Figs. 2 and 3. Laying date means the date of the first egg in the nest of rearing. PHA is injection of phytohaemagglutinin. Year was fitted as a categorical variable.

nestling quality. Coverage of broader gradients of environmental quality (Rollinson and Hutchings 2013a) as well as more ultimate fitness measures (Hunt and Hodgson 2010, Krist 2011, Williams 2012) was warranted. This is because components of offspring fitness may respond differently (Krist 2011) and sometimes even oppositely (Hendry et al. 2001) to variation in egg size. Similarly, depending on the utilized range of postnatal conditions, one could find support for both greater and smaller egg-size effects in harsh rearing conditions (see Allen et al. 2008, Koch and Meunier 2014). Results of this study that used a gradient of postnatal conditions and followed offspring through their lifetime support the view that emerged in earlier avian studies that egg size is most important for offspring fitness if the quality of the postnatal conditions is low.

Harsh environments should select for larger eggs due to higher fitness returns from large offspring (Brockelman 1975, Parker and Begon 1986, McGinley et al. 1987). This was indeed observed in a few empirical studies done on fish (Bashey 2008, Rollinson and Hutchings 2013a, Riesch et al. 2014). This pattern might be more general and responsible for large egg-size variation between populations and even species of various oviparous taxa (Fox and Czesak 2000). However, one point that seems to be neglected by both theorists and empiricists is the cost of egg size in relation to environmental conditions. If costs of laying large eggs were larger in harsh environments, as one might expect, then higher returns in terms of offspring fitness still might not lead to the evolution of large eggs, as natural

selection will favor the egg size that maximizes maternal, not offspring fitness (Marshall and Uller 2007).

Our results of stronger dependence of offspring fitness on initial maternal investment in poor postnatal conditions are relevant to the theories of differential allocation and reproductive compensation. They are in line with a recent theoretical model predicting that females should invest more in initial offspring size if they are paired with males providing poor paternal care, but should invest less if paternal care is good (Kindsvater and Alonzo 2014). This result can be called reproductive compensation if females prefer males that provide good care, but differential allocation if they prefer males providing less care (Kindsvater and Alonzo 2014). Similar to their results, we found support for a scenario depicted in Fig. 1B, which could also be called reproductive compensation if attractive males improve offspring postnatal conditions, or differential allocation if these good conditions are provided by unattractive males. Both attractive (e.g., Szöllösi et al. 2009, Moreno et al. 2013) and unattractive (e.g., Qvarnström 1997, Sanz 2001) males have been found to provide material benefits to offspring or females. Therefore it is difficult to assess whether patterns found by Kindsvater and Alonzo (2014) and the present study represent support for differential allocation or reproductive compensation. Future studies should integrate information about female preferences for male traits, including paternal care, and their allocation decisions and test possible interactive effects between them on offspring fitness.

The assumption of the Trivers-Willard model is sex specificity of fitness returns from initial parental investment (Trivers-Willard 1973, Krist 2006, Fig. 1D). This

assumption has rarely been tested, especially with more ultimate measures of offspring fitness (Komdeur 2012; but see Clutton-Brock et al. 1984, Bowers et al. 2015). Here we did not find any evidence for sex-specific fitness returns from egg size. However, a limitation of our measure of male fitness was our inability to take into account most of the extra-pair paternity and polygyny. Thus our results are not fully conclusive with respect to male reproductive success. Despite this, we can refute the possibility of sex-specific fitness returns due to differential sensitivity of male nestlings (Rosivall et al. 2010).

Most previous studies of sex-specific effects of egg size in birds were concerned with short-term effects on nestling survival or morphology. Egg size had a larger effect on male than female offspring in some of these studies (Anderson et al. 1997, Arnold and Griffiths 2003, Bonisoli-Alquati et al. 2008), whereas others reported variable directions of sex-specific effects (Love and Williams 2011) or no difference between sexes at all (Whitehead et al. 1990, Magrath et al. 2003, Fargallo et al. 2006). Whenever there are sex-specific fitness returns from larger eggs, this should lead to the evolution of larger eggs for the more sensitive sex. However, a recent meta-analysis found no evidence of sexual dimorphism in bird eggs (Rutkowska et al. 2014). Instead, variation between primary studies was explained solely by sampling error. This raises the question of whether sampling variance can also explain the variation in sex-specific fitness returns from egg investment. This could be tested after the accumulation of more data. In the particular case of the Collared Flycatcher, we previously did not find evidence of egg sexual dimorphism, although clutch sex ratios varied as predicted by the Trivers and Willard model (Bowers et al. 2013). We concluded that this difference might be caused by a constraint on female allocation within broods (Bowers et al. 2013), but results of this study suggest that this pattern can be explained simply by a lack of benefits from sex-specific allocation.

Our study examined the specific case of egg-size effects, but the assumptions of both the Trivers-Willard and differential allocation models relate to parental investment in general. Future studies should test sex specificity of fitness returns arising from other types of parental investment, such as egg composition, nestling provisioning, or post-fledging care, which is a poorly understood but badly needed part of parental investment (Grüebler and Naef-Daenzer 2010, Williams 2012). For example, parents may feed offspring for an additional two weeks after fledging in the Collared Flycatcher (Cramp and Perrins 1993). Because parents might compensate for a small investment in one stage (e.g., eggs) by a larger investment in another stage (e.g., nestling provisioning), this could confound results of studies limited to only one stage (Krist and Remeš 2004), just as looking at one component of fitness may not always provide the same results as when global fitness is considered (Hunt and Hodgson 2010). Al-

though compensation between prenatal and postnatal maternal effects has not been found in a few empirical studies (e.g., Krist 2009, 2011), it is difficult to draw broad conclusions from these findings, because studies specifically testing these compensatory effects are rare. Therefore we suggest that future studies should strive to integrate all components of parental care and fitness into their global measures and test how total parental effort modifies true offspring fitness across environmental gradients. However, this would likely be a very challenging task, especially for species with a prolonged period of parental care.

Literature Cited

Agosta, S. J. 2008. Selection on offspring size varies within and among families in relation to host nutritional quality. Evolutionary Ecology 22:71–83.

Allen, R. M., Y. M. Buckley, and D. J. Marshall. 2008. Offspring size plasticity in response to intraspecific competition: an adaptive maternal effect across life-history stages. American Naturalist 171:225–237.

Anderson, D. J., J. Reeve, and D. M. Bird. 1997. Sexually dimorphic eggs, nestling growth and sibling competition in American Kestrels Falco sparverius. Functional Ecology 11: 331–335.

Arnold, K. E., and R. Griffiths. 2003. Sex-specific hatching order, growth rates and fledging success in jackdaws Corvus monedula. Journal of Avian Biology 34:275–281.

Bashey, F. 2008. Competition as a selective mechanism for larger offspring size in guppies. Oikos 117:104–113.

Bonisoli-Alquati, A., R. Martinelli, D. Rubolini, and N. Saino. 2008. Sex-specific effects of albumen removal and nest environment manipulation on Barn Swallow nestlings. Ecology 89:2315–2324.

Bowers, E. K., P. Munclinger, S. Bureš, L. Kučerová, P. Nádvorník, and M. Krist. 2013. Cross-fostering eggs reveals that female collared flycatchers adjust clutch sex ratios according to parental ability to invest in offspring. Molecular Ecology 22:215–228.

Bowers, E. K., C. F. Thompson, and S. K. Sakaluk. 2015. Persistent sex-by-environment effects on offspring fitness and sex-ratio adjustment in a wild bird population. Journal of Animal Ecology 84:473–486.

Brockelman, W. Y. 1975. Competition, fitness of offspring, and optimal clutch size. American Naturalist 109:677–699.

Christians, J. K. 2002. Avian egg size: variation within species and inflexibility within individuals. Biological Reviews 77:1–26.

Clutton-Brock, T. H., S. D. Albon, and F. E. Guinnes. 1984. Maternal dominance, breeding success and birth sex-ratios in red deer. Nature 308:358–360.

Cramp, S., and C. M. Perrins. 1993. The birds of the western Palearctic. Oxford University Press, Oxford, UK.

Dziminski, M. A., and J. D. Roberts. 2006. Fitness consequences of variable maternal provisioning in quacking frogs

(*Crinia georgiana*). Journal of Evolutionary Biology 19:144–155.

Einum, S., and I. A. Fleming. 1999. Maternal effects of egg size in brown trout (*Salmo trutta*): norms of reaction to environmental quality. Proceedings of the Royal Society B 266:2095–2100.

Engqvist, L. 2005. The mistreatment of covariate interaction terms in linear model analyses of behavioural and evolutionary ecology studies. Animal Behaviour 70:967–971.

Fargallo, J. A., V. Polo, L. de Neve, J. Martin, J. A. Davila, and M. Soler. 2006. Hatching order and size-dependent mortality in relation to brood sex ratio composition in chinstrap penguins. Behavioral Ecology 17:772–778.

Fox, C. W. 2000. Natural selection on seed-beetle egg size in nature and the laboratory: variation among environments. Ecology 81:3029–3035.

Fox, C. W., and M. E. Czesak. 2000. Evolutionary ecology of progeny size in arthropods. Annual Review of Entomology 45:341–369.

Fox, C. W., and D. H. Reed. 2011. Inbreeding depression increases with environmental stress: an experimental study and meta-analysis. Evolution 65:246–258.

Garant, D., L. E. B. Kruuk, R. H. McCleery, and B. C. Sheldon. 2007. The effects of environmental heterogeneity on multivariate selection on reproductive traits in female great tits. Evolution 61:1546–1559.

Gil, D. 2003. Golden eggs: Maternal manipulation of offspring phenotype by egg androgen in birds. Ardeola 50:281–294.

Gowaty, P. A. 2008. Reproductive compensation. Journal of Evolutionary Biology 21:1189–1200.

Gowaty, P. A., W. W. Anderson, C. K. Bluhm, L. C. Drickamer, Y. K. Kim, and A. J. Moore. 2007. The hypothesis of reproductive compensation and its assumptions about mate preferences and offspring viability. Proceedings of the National Academy of Sciences USA 104:15023–15027.

Griffiths, R., M. C. Double, K. Orr, and R. J. G. Dawson. 1998. A DNA test to sex most birds. Molecular Ecology 7:1071–1075.

Grüebler, M. U., and B. Naef-Daenzer. 2010. Survival benefits of post-fledging care: experimental approach to a critical part of avian reproductive strategies. Journal of Animal Ecology 79:334–341.

Gustafsson, L. 1989. Collared Flycatcher. Pages 75–88 *in* I. Newton, editor. Lifetime reproduction in birds. Academic Press, London, UK.

Harris, W. E., and T. Uller. 2009. Reproductive investment when mate quality varies: differential allocation versus reproductive compensation. Philosophical Transactions of the Royal Society B 364:1039–1048.

Hendry, A. P., T. Day, and A. B. Cooper. 2001. Optimal size and number of propagules: allowance for discrete stages and effects of maternal size on reproductive output and offspring fitness. American Naturalist 157:387–407.

Hoffmann, A. A., and J. Merilä. 1999. Heritable variation and evolution under favourable and unfavourable conditions. Trends in Ecology and Evolution 14:96–101.

Horváthová, T., S. Nakagawa, and T. Uller. 2012. Strategic female reproductive investment in response to male attractiveness in birds. Proceedings of the Royal Society B 279:163–170.

Hunt, J., and D. Hodgson. 2010. What is fitness, and how do we measure it? Pages 46–70 *in* D. F. Westneat and C. W. Fox, editors. Evolutionary behavioral ecology. Oxford University Press, Oxford, UK.

Kindsvater, H. K., and S. H. Alonzo. 2014. Females allocate differentially to offspring size and number in response to male effects on female and offspring fitness. Proceedings of the Royal Society B 281(1779): 20131981.

Koch, L. K., and J. Meunier. 2014. Mother and offspring fitness in an insect with maternal care: phenotypic trade-offs

between egg number, egg mass and egg care. BMC Evolutionary Biology 14:125.

Komdeur, J. 2012. Sex allocation. Pages 171–188 *in* N. J. Royle, P. T. Smiseth, and M. Kölliker, editors. The evolution of parental care. Oxford University Press, Oxford, UK.

Krist, M. 2006. Should mothers in poor condition invest more in daughter than in son? Ethology Ecology and Evolution 18:241–246.

Krist, M. 2009. Short- and long-term effects of egg size and feeding frequency on offspring quality in the collared flycatcher (*Ficedula albicollis*). Journal of Animal Ecology 78:907–918.

Krist, M. 2011. Egg size and offspring quality: a meta-analysis in birds. Biological Reviews 86:692–716.

Krist, M., and P. Munclinger. 2011. Superiority of extra-pair offspring: maternal but not genetic effects as revealed by a mixed cross-fostering design. Molecular Ecology 20:5074–5091.

Krist, M., and V. Remeš. 2004. Maternal effects and offspring performance: in search of the best method. Oikos 106:422–426.

Krist, M., V. Remeš, L. Uvírová, P. Nádvorník, and S. Bureš. 2004. Egg size and offspring performance in the collared flycatcher (*Ficedula albicollis*): a within-clutch approach. Oecologia 140:52–60.

Lindén, M., L. Gustafsson, and T. Pärt. 1992. Selection on fledging mass in the Collared Flycatcher and the Great Tit. Ecology 73:336–343.

Love, O. P., and T. D. Williams. 2011. Manipulating developmental stress reveals sex-specific effects of egg size on offspring phenotype. Journal of Evolutionary Biology 24:1497–1504.

Lynch, M., and B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer Associates, Sunderland, Massachusetts, USA.

Magrath, M. J. L., L. Brouwer, and J. Komdeur. 2003. Egg size and laying order in relation to offspring sex in the extreme sexually size dimorphic brown songlark, *Cinclorhamphus cruralis*. Behavioral Ecology and Sociobiology 54:240–248.

Marshall, D. J., and T. Uller. 2007. When is a maternal effect adaptive? Oikos 116:1957–1963.

Martin, T. G., B. A. Wintle, J. R. Rhodes, P. M. Kuhnert, S. A. Field, S. J. Low-Choy, A. J. Tyre, and H. P. Possingham. 2005. Zero tolerance ecology: improving ecological inference by modelling the source of zero observations. Ecology Letters 8:1235–1246.

McGinley, M. A., D. H. Temme, and M. A. Geber. 1987. Parental investment in offspring in variable environments: theoretical and empirical considerations. American Naturalist 130:370–398.

Metcalfe, N. B., and P. Monaghan. 2001. Compensation for a bad start: grow now, pay later? Trends in Ecology and Evolution 16:254–260.

Mitchell, T. S., D. A. Warner, and F. J. Janzen. 2013. Phenotypic and fitness consequences of maternal nest-site choice across multiple early life stages. Ecology 94:336–345.

Moreno, J., A. Velando, S. Gonzalez-Braojos, R. Ruiz-de-Castaneda, and A. Cantarero. 2013. Females paired with more attractive males show reduced oxidative damage: possible direct benefits of mate choice in pied flycatchers. Ethology 119:727–737.

Mousseau, T. A., and C. W. Fox. 1998. The adaptive significance of maternal effects. Trends in Ecology and Evolution 13:403–407.

Oksanen, T. A., I. Jokinen, E. Koskela, T. Mappes, and H. Vilpas. 2003. Manipulation of offspring number and size: benefits of large body size at birth depend upon the rearing environment. Journal of Animal Ecology 72:321–330.

Parichy, D. M., and R. H. Kaplan. 1992. Maternal effects on offspring growth and development depend on environmen-

tal quality in the frog *Bombina orientalis*. Oecologia 91:579–586.

Parker, G. A., and M. Begon. 1986. Optimal egg size and clutch size: effects of environment and maternal phenotype. American Naturalist 128:573–592.

Quero, J. L., R. Villar, T. Maranon, R. Zamora, and L. Poorter. 2007. Seed-mass effects in four Mediterranean *Quercus* species (Fagaceae) growing in contrasting light environments. American Journal of Botany 94:1795–1803.

Qvarnström, A. 1997. Experimentally increased badge size increases male competition and reduces male parental care in the collared flycatcher. Proceedings of the Royal Society B 264:1225–1231.

Qvarnström, A., B. C. Sheldon, T. Pärt, and L. Gustafsson. 2003. Male ornamentation, timing of breeding, and cost of polygyny in the collared flycatcher. Behavioral Ecology 14:68–73.

Ratikainen, I. I., and H. Kokko. 2010. Differential allocation and compensation: who deserves the silver spoon? Behavioral Ecology 21:195–200.

Rey, P. J., J. M. Alcantara, F. Valera, A. M. Sanchez-Lafuente, J. L. Garrido, J. M. Ramirez, and A. J. Manzaneda. 2004. Seedling establishment in *Olea europaea*: seed size and microhabitat affect growth and survival. Ecoscience 11:310–320.

Riesch, R., M. Plath, I. Schlupp, M. Tobler, and R. B. Langerhans. 2014. Colonisation of toxic environments drives predictable life-history evolution in livebearing fishes (Poeciliidae). Ecology Letters 17:65–71.

Robinson, M. R., G. S. van Doorn, L. Gustafsson, and A. Qvarnström. 2012. Environment-dependent selection on mate choice in a natural population of birds. Ecology Letters 15:611–618.

Rollinson, N., and J. A. Hutchings. 2013a. Environmental quality predicts optimal egg size in the wild. American Naturalist 182:76–90.

Rollinson, N., and J. A. Hutchings. 2013b. The relationship between offspring size and fitness: integrating theory and empiricism. Ecology 94:315–324.

Rosenthal, R. 1994. Parametric measures of effect size. Pages 231–244 *in* H. Cooper and L. V. Hedges, editors. The handbook of research synthesis. Russell Sage Foundation, New York, New York, USA.

Rosivall, B., E. Szöllösi, D. Hasselquist, and J. Török. 2010. Males are sensitive: sex-dependent effect of rearing conditions on nestling growth. Behavioral Ecology and Sociobiology 64:1555–1562.

Rosivall, B., E. Szöllösi, and J. Török. 2005. Maternal compensation for hatching asynchrony in the collared flycatcher *Ficedula albicollis*. Journal of Avian Biology 36:531–537.

Rutkowska, J., A. Dubiec, and S. Nakagawa. 2014. All eggs are made equal: meta-analysis of egg sexual size dimorphism in birds. Journal of Evolutionary Biology 27:153–160.

Sanz, J. J. 2001. Experimentally reduced male attractiveness increases parental care in the pied flycatcher *Ficedula hypoleuca*. Behavioral Ecology 12:171–176.

Schmoll, T. 2011. A review and perspective on context-dependent genetic effects of extra-pair mating in birds. Journal of Ornithology 152:265–277.

Semlitsch, R. D., and J. W. Gibbons. 1990. Effects of egg size on success of larval salamanders in complex aquatic environments. Ecology 71:1789–1795.

Sheldon, B. C. 2000. Differential allocation: tests, mechanisms and implications. Trends in Ecology and Evolution 15:397–402.

Sheldon, B. C., J. Merilä, G. Lindgren, and H. Ellegren. 1998. Gender and environmental sensitivity in nestling Collared Flycatchers. Ecology 79:1939–1948.

Smiseth, P. T., M. P. Scott, and C. Andrews. 2011. Hormonal regulation of offspring begging and mediation of parent–offspring conflict. Animal Behaviour 81:507–517.

Smith, H. G., and M. Bruun. 1998. The effect of egg size and habitat on starling nestling growth and survival. Oecologia 115:59–63.

Sokal, R. R., and F. J. Rohlf. 1995. Biometry: the principles and practice of statistics in biological research. W. H. Freeman, New York, New York, USA.

Styrsky, J. D., R. C. Dobbs, and C. F. Thompson. 2000. Food-supplementation does not override the effect of egg mass on fitness-related traits of nestling house wrens. Journal of Animal Ecology 69:690–702.

Styrsky, J. D., M. P. Eckerle, and C. F. Thompson. 1999. Fitness-related consequences of egg mass in nestling house wrens. Proceedings of the Royal Society B 266:1253–1258.

Svensson, E., and B. Sinervo. 2000. Experimental excursions on adaptive landscapes: density-dependent selection on egg size. Evolution 54:1396–1403.

Szöllösi, E., B. Rosivall, D. Hasselquist, and J. Török. 2009. The effect of parental quality and malaria infection on nestling performance in the Collared Flycatcher (*Ficedula albicollis*). Journal of Ornithology 150:519–527.

Trivers, R. L., and D. E. Willard. 1973. Natural selection of parental ability to vary the sex ratio of offspring. Science 179:90–92.

West, S. A. 2009. Sex allocation. Princeton University Press, Princeton, New Jersey, USA.

Whitehead, P. J., W. J. Freeland, and K. Tschirner. 1990. Early growth of magpie geese, *Anseranas semipalpata*: sex differences and influence of egg size. Australian Journal of Zoology 38:249–262.

Williams, T. D. 2012. Physiological adaptations for breeding in birds. Princeton University Press, Princeton, New Jersey, USA.

Wolf, J. B., and M. J. Wade. 2001. On the assignment of fitness to parents and offspring: whose fitness is it and when does it matter? Journal of Evolutionary Biology 14:347–356.

SUPPLEMENTAL MATERIAL

Data Availability

Článek 14: Krist, M., Janča, M., Edme, A. & Dzuro, R. (2015) Are prenatal maternal resources more important in competitive than in benign postnatal environments? *Auk*, **132, 577-583.**

RESEARCH ARTICLE

Are prenatal maternal resources more important in competitive than in benign postnatal environments?

Miloš Krist,[1,2]* Martin Janča,[1] Anaïs Edme,[1] and Rudolf Dzuro[1]

[1] Department of Zoology and Laboratory of Ornithology, Faculty of Science, Palacký University, Olomouc, Czech Republic
[2] Museum of Natural History, Olomouc, Czech Republic
* Corresponding author: milos.krist@upol.cz

ABSTRACT

According to theoretical models, the optimal solution of the life-history trade-off between the number and size of offspring depends on the quality of the environment. Offspring size should be more important for their fitness in more competitive environments. This idea was rarely experimentally tested in taxa with prolonged periods of parental postnatal care, such as in birds. Here we manipulated the offspring rearing environment by enlarging or reducing brood size. Enlarged broods suffered greater mortality rates and raised smaller fledglings. Egg size had a significant positive effect on fledging mass and length of tarsus and a nonsignificant effect on wing length. These effects were similar in enlarged-sized as well as reduced-sized broods. We only found a tendency for the predicted interaction between treatment and egg size in the case of nestling mass where egg size had a positive effect in enlarged broods but none in reduced broods. In contrast, in one year we found an opposite interaction where egg size positively affected offspring survival only in reduced broods. More studies that manipulate the offspring rearing environment and follow offspring over the long term are needed to draw general conclusions about context-dependence of early maternal effects.

Keywords: context-dependence, egg size, environmental quality, parental investment

¿Son los recursos maternos prenatales más importantes en ambientes postnatales competitivos o benignos?

RESUMEN

De acuerdo a los modelos teóricos, la solución óptima en la historia de vida entre el número y el tamaño de los hijos depende de la calidad del ambiente. El tamaño de los hijos debería ser más importante en términos de éxito reproductivo en los ambientes más competitivos. Esta idea ha sido evaluada pocas veces experimentalmente en taxa con períodos prolongados de cuidado parental postnatal, como en las aves. Aquí manipulamos el ambiente de cría agregando o reduciendo el tamaño de la nidada. Las nidadas con crías agregadas presentaron tasas de mortalidad mayores y criaron volantones más pequeños. El tamaño del huevo tuvo efectos positivos significativos en el peso y en el largo del tarso de los pichones y tuvo un efecto no significativo en el largo del ala. Estos efectos fueron similares en las nidadas con crías agregadas reducidas de tamaño. Solo encontramos una tendencia en la interacción que predijimos entre el tratamiento y el tamaño del huevo en el caso del peso del pichón, donde el tamaño del huevo tuvo un efecto positivo en las nidadas con crías agregadas pero no en las nidadas reducidas de tamaño. Por el contrario, en un año encontramos una interacción opuesta donde el tamaño del huevo afectó positivamente la supervivencia de los pichones solo en las nidadas reducidas de tamaño. Se necesitan más estudios que manipulen el ambiente de cría de los pichones y que sigan a los pichones al largo plazo para obtener conclusiones generales sobre la dependencia del contexto de los efectos maternos tempranos.

Palabras clave: calidad ambiental, dependencia del contexto, inversión parental, tamaño del huevo.

INTRODUCTION

Parents often face the dilemma whether to invest in a limited manner to many offspring or invest heavily in a smaller number. This trade-off between offspring quality and quantity was first modeled by Smith and Fretwell (1974), who assumed diminishing fitness returns from investment in each offspring. Consequently, every offspring should only receive an intermediate amount of parental resources, and the remaining resources may be better used for nourishment of further offspring or for parental maintenance. This early model predicted one optimal offspring size both within and between families.

Subsequent models added more complexity and realism to this theory, assuming that fitness returns from parental investment may depend on environmental context. In theory, the initial maternal investment should be more important for offspring performance if they live in harsh and competitive conditions than in mild ones where all offspring have a good chance to survive (Brockelman 1975, Parker and

Begon 1986, McGinley et al. 1987, Fischer et al. 2011, Kuijper and Johnstone 2013). Consequently, harsher habitats should select for large offspring whereas more benign habitats select for small offspring. This spatial and temporal variance in selection pressures is likely the underlying cause of observed egg-size variation between and within species (Fox and Czesak 2000). Recently, this hypothesis was supported in several fish species (Einum and Fleming 1999, Bashey 2008, Rollinson and Hutchings 2013, Riesch et al. 2014).

Studying these ideas in other taxa could test their generality across different life-histories. For example, birds usually have a more-developed parental care in the form of incubation and offspring provisioning, which impacts offspring phenotypes (DuRant et al. 2013, Bowers et al. 2014) and thus could mitigate earlier maternal effects. Therefore, whether context-dependence may induce the same inter-population variability of egg size in birds as was found in fish is unclear. Many avian studies have focused on the relationship between egg size and offspring quality, yet only a few tested if egg-size effects depend on rearing conditions (Krist 2011). We know of only 3 studies that manipulated chick postnatal conditions to test for context-dependence of egg-size effects. In the first, food was supplemented to mimic a mild environment, but egg-size effects were not smaller than in natural nests (Styrsky et al. 2000). Similarly, egg-size effects were not larger in highly competitive enlarged broods (Bonisoli-Alquati et al. 2008); however, they were more pronounced in a stressful condition created by handicapping the parents (Love and Williams 2011). In contrast, between-study comparisons revealed a larger egg-size effect in captivity than in the wild, despite wild conditions likely representing a more stressful environment (Krist 2011). From these few and contradictory results, no general conclusion about context-dependence of maternal effects can be derived, and further studies are therefore warranted.

In this study we experimentally tested the assumption of adaptive scenario for between-population variation in egg size, specifically whether egg size has stronger effect on offspring performance in poor compared to good rearing environments. We created these good and poor rearing environments by experimental manipulation of brood size in the Collared Flycatcher (*Ficedula albicollis*). Enlarged broods represented poor rearing environments as a result of increased competition between nest mates (see also Forbes 2011). Despite poor offspring performance in enlarged broods, we did not find consistent evidence for stronger egg-size effects on nestling fitness-related traits in these competitive conditions.

METHODS

Study Area and Field Experiments
This study was conducted on Velký Kosíř (49°32′N, 17°04′E, 300–400 m a.s.l.), Czech Republic, in 2011, 2012, and 2014. In the breeding season, nest-boxes were monitored for

newly laid eggs of Collared Flycatchers. All eggs were measured for width and length by digital calliper to the nearest 0.01 mm, and egg volume was calculated according to Hoyt's (1979) formula: volume $= 0.51 \times$ length $\times$ breadth^2. In altricial birds, sibling competition is a major force that affects the quality of both objective and effective rearing environment (Forbes 2011); therefore, we manipulated brood size to decrease sibling competition in reduced broods and enhance it in enlarged ones. The details of how brood size was manipulated differed between years.

In 2011, complete clutches were exchanged between pairs of nests differing by one in clutch size and matched ($n = 46$) or differing by one ($n = 22$) in laying date of the last egg. Cross-fostering was usually done 2–3 days after the laying of the last egg (mean $= 3.13$, SD $= 1.87$, range 1–12). In 2012, 2-day-old chicks were exchanged between nests matched by hatching date and clutch size ($n = 56$) or differing in clutch size by 1 ($n = 12$) or 2 ($n = 2$) egg. In enlarged broods, 2 randomly chosen chicks or eggs were left and the remainder were moved to reduced broods. All young and eggs from reduced broods were fostered in enlarged broods. In cases where clutch sizes were not matched, we left a number of young in enlarged nests, which increased their brood size by 2 young or eggs. In 2014 one of the core young from a reduced brood was taken and transferred to an enlarged brood 2 days after hatching. We transferred only core young (i.e. those that did not suffer from late hatching; see Forbes 2011) to enhance the chance of their survival and thus prolonged sibling competition in enlarged broods. Nests in experimental dyads were matched by hatching date ($n = 88$) or differed by 1 day (n $= 16$). In dyads matched by hatching date, treatment was allocated randomly within dyads; in dyads differing by 1 day, the chick was always transferred from the older to the younger brood to enhance the chance of its survival among younger competitors. The 2 chicks from the original nest (in 2012) or the 1 from the foster nest (in 2014) were marked by trimming part of the down feathers on their heads.

On day 6, nestlings were ringed and weighed. We also noted which nestlings had trimmed feathers. On day 13, nestlings were reweighed and their tarsi (to 0.01 mm) and wings (to 1 mm) were measured. Thereafter, nests were checked at weekly intervals to record the fate of individual nestlings (fledged or died).

Statistical Analyses
The main predictors in all initial statistical models were mean volume of eggs in the nest, treatment (enlarged or reduced), and their interaction. We also included a 3-way interaction between treatment, egg size, and year to check whether potential context-dependence of egg-size effects was apparent only in some years, for example due to difference in our experimental designs. If these interactions
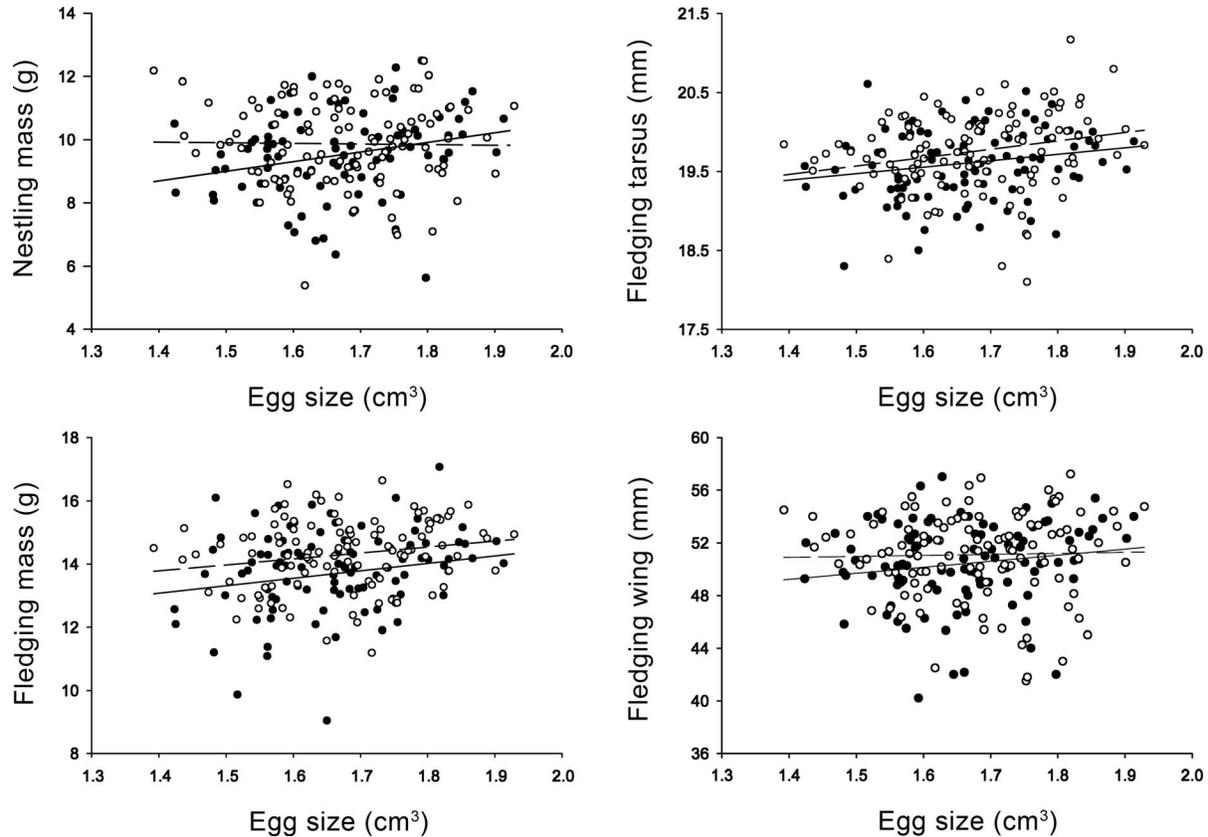
FIGURE 1. Scatterplots showing relationships between egg size and nestlings morphological traits in enlarged (solid circles and lines) and reduced (open circles and dashed lines) broods. Fitted lines are predicted from models that included interactions between treatment and egg size.

were not significant ($P > 0.05$), they were omitted from the final models, starting with the 3-way interaction. Further fixed predictors were year, fitted as a categorical factor, to control for annual effects; laying date to control for seasonal effects; and initial clutch size (see Hauber 2002 for a discussion of potential clutch-size effects). We also included identity of recipient and donor nest-boxes as random factors because some were used in several years.

Response variables were average nestling traits: body mass, length of wing and tarsus, and fledging success. These variables were averaged for focal young only; we omitted from this average the young that were left in original nests (in 2012) or that were transferred to foster nests (in 2014). Thus, mean egg size predicted traits of nestlings that were either cross-fostered (in 2011–2012) or raised in original nests (in 2014). All averages were based solely on fledged young and fledging success solely on broods from which at least one young fledged. We excluded dead chicks from computation of averages to avoid bias due to inclusion of starved nestlings that often have outlying values of morphological traits. We did not include total failures when testing for fledging success because such failures are unlikely to be caused by size of

eggs (Krist 2011); more likely reasons are predation of adults or nest abandonment. All tests were done in SAS, v9.2, Proc Glimmix. We used identity link when response variables were normally distributed (morphological traits) and logit link and event/trial syntax for binomially distributed fledging success.

In theory, egg-size effect can be inflated if eggs are not cross-fostered due to the co-variation of egg size and parental quality (Magrath 1992, Krist and Remeš 2004). In reality, however, this inflation is usually small at best, as demonstrated by a recent meta-analysis (Krist 2011) and a previous study in this population (Krist 2009). Moreover, in this study we are interested in the difference of egg-size effects between treatments, not in absolute egg-size effect within a treatment. Thus our data, despite being based on a mix of nestlings raised in original and foster nests, are well suited for testing the interaction between egg size and quality of the post-natal environment.

RESULTS

Main results are presented as direction and size of egg effects and treatment effects (Figures 1, 2) and statistical
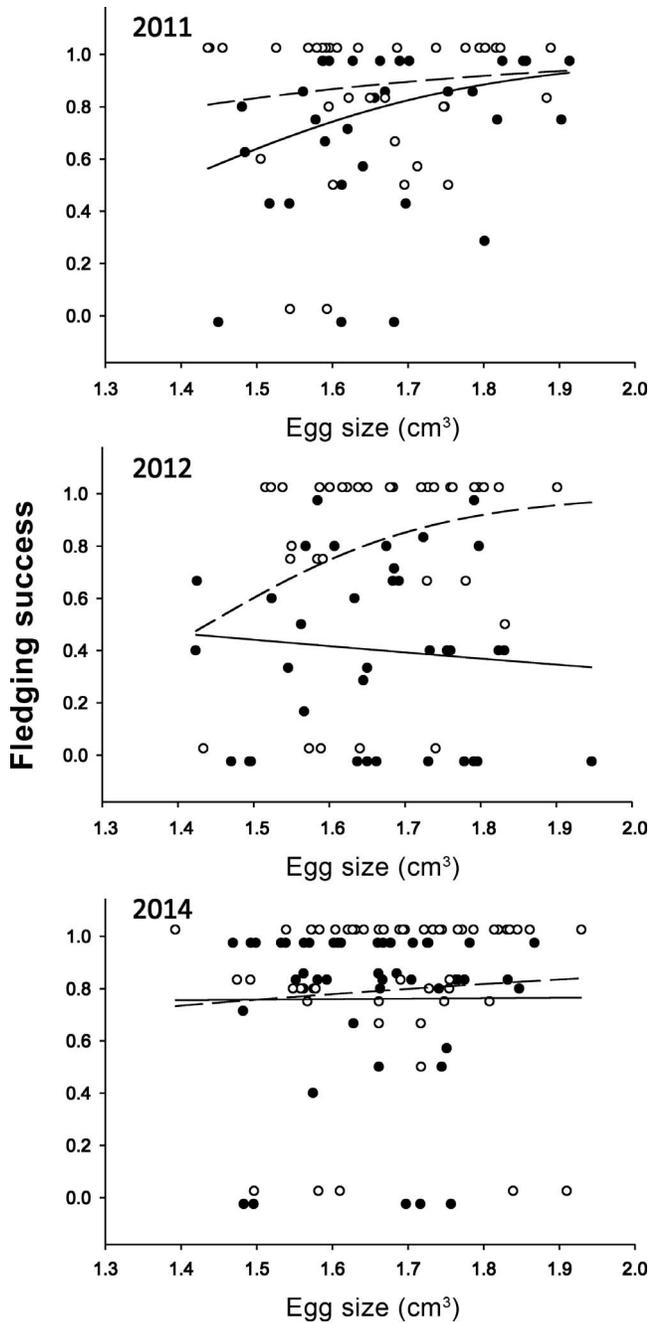
FIGURE 2. Relationship between egg size and fledging success in 3 years of the study. Broods with complete failure are also displayed, but these nests were not used for statistical tests. For nests with complete failure or success, points are slightly shifted above (reduced broods) or below (enlarged broods) from 0 and 1 to make the pattern more visible. Fitted lines are predicted from models that included interactions between treatment and egg size.

significance of these relationships (Table 1). Scatterplots show raw data; however, fitted lines were predicted from models that included the interaction between treatment and egg size. Thus, these predicted values were controlled for effects of other variables in the model. Difference in slopes of these lines would suggest different effect of egg size on offspring performance in enlarged compared to reduced broods (i.e. significant interaction between treatment and egg size). Different height of these lines would suggest treatment effect.

Because the lines are generally lower in enlarged broods, nestlings in these broods were smaller (Table 1; Figure 1) and had a lower rate of survival (Table 1; Figure 2) than those in reduced broods, indicating that our experiment created nest environments with low and high intensity of sibling competition. For example, nestlings in enlarged broods fledged at about 0.5 g lower mass than those in reduced broods, an effect independent of egg mass (Figure 1). Despite lower survival rate, final number of all fledglings was still higher in enlarged broods (total failures excluded; mean $\pm$ SE: 5.22 $\pm$ 0.14 vs. 4.03 $\pm$ 0.14, $F_{1, 209} = 37.00$, $P < 0.001$), which holds also for number of focal fledglings (4.56 $\pm$ 0.14 vs. 4.03 $\pm$ 0.13, $F_{1, 205} = 7.39$, $P = 0.007$) .

Egg size generally had positive effects on all offspring traits (Figures 1, 2), although this effect was not statistically significant for wing length (Table 1). Egg-size effects were not confounded by difference in initial clutch sizes because these 2 variables were not correlated ($r = -0.060$, $n = 242$, $P = 0.354$). We found no consistent evidence that egg-size effects are more important in competitive than in benign postnatal environments. There was only a tendency for stronger egg-size effect on nestling mass in enlarged broods (Figure 1, test of the interaction: $F_{1, 189} = 3.76$, $P = 0.054$). On the contrary, fledging success was more dependent on egg size in reduced broods in 2012, while no difference between slopes was found in 2011 and 2014 (Table 1; Figure 2). Similarly, in the case of other morphological traits, the slope of egg-size effect in reduced broods was similar to that in enlarged broods (Figure 1).

DISCUSSION

Egg size had positive effects on the offspring's fitness-related traits, indicating that parents did not fully compensate for small eggs with more intensive prenatal (incubation behaviour; DuRant et al. 2013) or postnatal (offspring provisioning; Bowers et al. 2014) care. Avian egg size thus might be as important for offspring fitness as in taxa with little parental care, such as in fish. However, unlike in fish, we did not find consistent evidence for context-dependence of egg-size effects. Although our brood size manipulation was effective in creating harsh and mild rearing environments, the slopes relating

TABLE 1. Effects of fixed factors on offspring traits. Only final models after elimination of nonsignificant interactions are shown. DDF = denominator degrees of freedom, F = F-value, and P = P-value. Statistical directions of treatment and egg-size effects are given in Figures 1 and 2. Year was fitted as a categorical variable.

	Body mass at 6 days			Body mass at 13 days			Tarsus length			Wing length			Fledging success		
	DDF	F	P	DDF	F	P	DDF	F	P	DDF	F	P	DDF	F	P
Treatment	188	3.55	0.061	198	12.63	<0.001	196	4.12	0.044	194	2.13	0.146	194	0.18	0.672
Egg size	189	2.41	0.122	198	8.61	0.004	197	10.74	0.001	192	1.68	0.196	194	1.13	0.230
Egg size * Treatment													194	0.01	0.940
Egg size * Treatment * Year													194	2.51	0.043
Year	168	2.55	0.081	198	9.34	<0.001	182	9.42	<0.001	176	8.97	<0.001	194	0.37	0.691
Laying date	190	17.58	<0.001	198	9.43	0.002	197	10.85	0.001	191	29.8	<0.001	194	8.83	0.003
Clutch size	189	2.53	0.113	198	5.53	0.020	198	0.16	0.693	192	6.07	0.015	194	5.42	0.021

offspring performance to egg size were generally similar across environments. We found only a tendency for steeper slope in enlarged broods for nestling mass but an opposite pattern (steeper slope in reduced broods) for fledging success in 1 of the 3 study years.

Egg size is often variable between populations (Fox and Czesak 2000, Martin et al. 2006, Marshall and Keough 2008). In the prevailing view, this variation should be adaptive because harsh environments select for large eggs, while in benign habitats egg size is of less importance for offspring survival. Consequently, population differences in egg size could arise either as evolved fixed strategies (Fox and Czesak 2000) or as plastic responses of individuals able to anticipate the quality of the offspring environment and adjust their eggs accordingly (e.g., Garant et al. 2007). However, experiments that manipulate the offspring-rearing environment to test its interaction with egg size on offspring and maternal fitness are rare in taxa in general (Fox and Czesak 2000) and in birds in particular (Krist 2011, Williams 2012). Thus, an alternative explanation could be that inter-population variation is largely non-adaptive and has arisen due to proximate constraints such as temperature, precipitation, and food supply at the time of egg formation. The results of our study somewhat support this nonadaptive scenario because we did not find consistent evidence for stronger egg-size effects in poor rearing environments. However, several potential issues should be considered.

First, we tested egg-size effects on traits that are correlates of offspring fitness, not on true offspring fitness itself. This major weakness is unfortunately a rule rather than an exception in this field (Krist 2011) and in other fields (Hunt et al. 2004). Fledging traits are strongly correlated with first year survival (Lindén et al. 1992, Kruuk et al. 2001); however, recently it was shown that a large part of this correlation might be caused by annual effects, not causal effects of fledgling traits (Bouwhuis et al. 2015). So the challenge for future work is to follow

offspring over the long-term, ideally through their lifetime, which is inevitably a difficult task.

Second, we experimentally created environments of 2 qualities and looked at the difference in egg-size effects between them. More elegant evidence for context-dependence would be the gradual change of egg-size effects across an environmental gradient (Rollinson and Hutchings 2013); however, such studies are usually observational because experimentally inducing environmental gradients would be challenging. Nevertheless, this experimental approach has the strength of decoupling the potential correlations between the factor of interest, here intensity of sibling competition, and other unmeasured variables.

Third, here we studied egg-size effects between broods. Although egg-size variability is usually much greater between females rather than within them (Christians 2002), which makes tests conducted on the former level more powerful (Krist 2011), nestlings from different broods do not directly compete, which might decrease the chance to reveal context-dependence of egg-size effects between nests. It is difficult to predict which of the 2 mechanisms will prevail, and thus the level of context-dependence should be easier to find. In a previous study we generated extra variation of egg size within clutches and found evidence for context-dependence of egg-size effects within broods (Krist and Munclinger personal communication). Another possibility to test environmental specificity of egg-size effects would be to compare offspring of different hatching orders. In species with substantial hatching asynchrony, marginal offspring that hatched late live in an effectively more competitive environment than core offspring (Forbes 2011). Thus, egg size should be more important for fitness of marginal than core offspring.

We concentrated our discussion on only one form of maternal effects: egg size. It is clear that there are many other types of prenatal effects, such as quality of

incubation (DuRant et. al. 2013) and amount of androgens (Gil 2008), carotenoids (Saino et al. 2003), immune factors (Saino et al. 2002), and vitamins (Matrková and Remeš 2014). All of these prenatal as well as the postnatal effects, such as nestling provisioning, may also have context-specific effects on offspring fitness; however, this possibility is even less explored than egg-size effects. Clearly more data are needed to draw more general conclusions about context-specific effect of parental investment on offspring fitness.

ACKNOWLEDGMENTS

LITERATURE CITED

Bashey, F. (2008). Competition as a selective mechanism for larger offspring size in guppies. Oikos 117:104–113.

Bonisoli-Alquati, A., R. Martinelli, D. Rubolini, and N. Saino (2008). Sex-specific effects of albumen removal and nest environment manipulation on Barn Swallow nestlings. Ecology 89: 2315–2324.

Bouwhuis, S., O. Vedder, C. J. Garroway, and B. C. Sheldon (2015). Ecological causes of multilevel covariance between size and first-year survival in a wild bird population. Journal of Animal Ecology 84:208–218.

Bowers, E. K., D. Nietz, C. F. Thompson, and S. K. Sakaluk (2014). Parental provisioning in House Wrens: Effects of varying brood size and consequences for offspring. Behavioral Ecology 25:1485–1493.

Brockelman, W. Y. (1975). Competition, fitness of offspring, and optimal clutch size. American Naturalist 109:677–699.

Christians, J. K. (2002). Avian egg size: Variation within species and inflexibility within individuals. Biological Reviews 77:1–26.

DuRant, S. E., W. A. Hopkins, G. R. Hepp, and J. R. Walters (2013). Ecological, evolutionary, and conservation implications of incubation temperature-dependent phenotypes in birds. Biological Reviews 88:499–509.

Einum, S., and I. A. Fleming (1999). Maternal effects of egg size in brown trout (Salmo trutta): Norms of reaction to environmental quality. Proceedings of the Royal Society of London, Series B 266:2095–2100.

Fischer, B., B. Taborsky, and H. Kokko (2011). How to balance the offspring quality-quantity tradeoff when environmental cues are unreliable. Oikos 120:258–270.

Forbes, S. (2011). Social rank governs the effective environment of siblings. Biology Letters 7:346–348.

Fox, C. W., and M. E. Czesak (2000). Evolutionary ecology of progeny size in arthropods. Annual Review of Entomology 45:341–369.

Garant, D., L. E. B. Kruuk, R. H. McCleery, and B. C. Sheldon (2007). The effects of environmental heterogeneity on multivariate selection on reproductive traits in female Great Tits. Evolution 61:1546–1559.

Gil, D. (2008). Hormones in avian eggs: Physiology, ecology and behavior. Advances in the Study of Behavior 38:337–398.

Hauber, M. E. (2002). Is reduced clutch size a cost of parental care in Eastern Phoebes (Sayornis phoebe)? Behavioral Ecology and Sociobiology 51:503–509.

Hoyt, D. F. (1979). Practical methods of estimating volume and fresh weight of bird eggs. The Auk 96:73–77.

Hunt, J., L. F. Bussiere, M. D. Jennions, and R. Brooks (2004). What is genetic quality? Trends in Ecology & Evolution 19:329–333.

Krist, M. (2009). Short- and long-term effects of egg size and feeding frequency on offspring quality in the Collared Flycatcher (Ficedula albicollis). Journal of Animal Ecology 78: 907–918.

Krist, M. (2011). Egg size and offspring quality: A meta-analysis in birds. Biological Reviews 86:692–716.

Krist, M., and V. Remeš (2004). Maternal effects and offspring performance: In search of the best method. Oikos 106:422–426.

Kruuk, L. E. B., J. Merilä, and B. C. Sheldon (2001). Phenotypic selection on a heritable size trait revisited. American Naturalist 158:557–571.

Kuijper, B., and R. A. Johnstone (2013). How should parents adjust the size of their young in response to local environmental cues? Journal of Evolutionary Biology 26: 1488–1498.

Lindén, M., L. Gustafsson, and T. Pärt (1992). Selection on fledging mass in the Collared Flycatcher and the Great Tit. Ecology 73:336–343.

Love, O. P., and T. D. Williams (2011). Manipulating developmental stress reveals sex-specific effects of egg size on offspring phenotype. Journal of Evolutionary Biology 24: 1497–1504.

Magrath, R. D. (1992). The effect of egg mass on the growth and survival of blackbirds: A field experiment. Journal of Zoology (London) 227:639–653.

Marshall, D. J., and M. J. Keough (2008). The evolutionary ecology of offspring size in marine invertebrates. Advances in Marine Biology 53:1–60.

Martin, T. E., R. D. Bassar, S. K. Bassar, J. J. Fontaine, P. Lloyd, H. A. Mathewson, A. M. Niklison, and A. Chalfoun (2006). Life-history and ecological correlates of geographic variation in egg and clutch mass among passerine species. Evolution 60: 390–398.

Matrková, J., and V. Remeš (2014). Vitamin E improves growth of Collared Flycatcher Ficedula albicollis young: A supplementation experiment. Journal of Avian Biology 45:475–483.

McGinley, M. A., D. H. Temme, and M. A. Geber (1987). Parental investment in offspring in variable environments – theoretical and empirical considerations. American Naturalist 130:370–398.

Parker, G. A., and M. Begon (1986). Optimal egg size and clutch size: Effects of environment and maternal phenotype. American Naturalist 128:573–592.

Riesch, R., M. Plath, I. Schlupp, M. Tobler, and R. B. Langerhans (2014). Colonisation of toxic environments drives predictable life-history evolution in livebearing fishes (Poeciliidae). Ecology Letters 17:65–71.

Rollinson, N., and J. A. Hutchings (2013). Environmental quality predicts optimal egg size in the wild. American Naturalist 182:76–90.

Saino, N., P. Dall'ara, R. Martinelli, and A. P. Møller (2002). Early maternal effects and antibacterial immune factors in the eggs, nestlings and adults of the Barn Swallow. Journal of Evolutionary Biology 15:735–743.

Saino, N., R. Ferrari, M. Romano, R. Martinelli, and A. P. Møller (2003). Experimental manipulation of egg carotenoids affects immunity of Barn Swallow nestlings. Proceedings of the Royal Society of London, Series B 270:2485–2489.

Smith, C. C., and S. D. Fretwell (1974). The optimal balance between size and number of offspring. American Naturalist 108:499–506.

Styrsky, J. D., R. C. Dobbs, and C. F. Thompson (2000). Food-supplementation does not override the effect of egg mass on fitness-related traits of nestling House Wrens. Journal of Animal Ecology 69:690–702.

Williams, T. D. (2012). Physiological Adaptations for Breeding in Birds. Princeton University Press, Princeton, NJ.

Článek 15: Edme, A., Munclinger, P. & Krist, M. (2016) Female collared flycatchers choose neighbouring and older extra-pair partners from the pool of males around their nests. *Journal of Avian Biology*, 47, 552-562.

Articles

Female collared flycatchers choose neighbouring and older extra-pair partners from the pool of males around their nests

Anais Edme, Pavel Munclinger and Milos Krist

A. Edme (anais.edme01@upol.cz) and M. Krist, Dept of Zoology and Laboratory of Ornithology, Faculty of Science, Palacky Univ., 17 listopadu 50, CZ-771 46 Olomouc, Czech Republic. MK also at: Museum of Natural History, nam. Republiky 5, CZ-771 73 Olomouc, Czech Republic. – Pavel Munclinger, Dept of Zoology, Faculty of Science, Charles Univ. in Prague, Vinicna 7, CZ-128 44 Prague, Czech Republic.

Extra-pair copulation is common among passerine birds. Females might engage in this behavior to obtain direct or indirect benefits. They may choose extra-pair males with larger ornaments, especially if they are costly to produce. Here we studied extra-pair paternity in the collared flycatcher. Genetic analysis allowed us to identify the presence or absence of extra-pair young in the focal nests, and to identify extra-pair fathers. We also identified potential males available as extra-pair sires around the nests of females who had extra-pair young. First, we tested the relationship between paternity in own nest and ornament size (wing patch and/or forehead patch), morphological traits and age of social males and females. Second, we compared the same suite of traits among social mates, extra-pair males and all potential extra-pair mates. Finally, we investigated the effect of the size of ornaments on the distance between the social nest and that of nest the extra-pair father. Contrary to our prediction, males with larger ornaments and longer wings lost more paternity in their nests. We also found that early breeders lost less paternity in their nests. Extra-pair males were older and had longer wings than social and potential extra-pair males. Females mainly obtained extra-pair mates near their nests but the distance did not vary according to ornamentation. These results could potentially be explained by differences in mate guarding strategy as older males may be more experienced in guarding their mate and attract other females more easily. More data about mate guarding and prospecting are needed to increase our understanding of mechanisms underlying the extra-pair paternity in birds.

Many birds are socially monogamous, but extra-pair paternity is widespread (Griffith et al. 2002, Westneat and Stewart 2003). This is advantageous for extra-pair males, because they obtain more offspring without providing parental care to these extra young (Møller and Ninni 1998, Sheldon and Ellegren 1999). The advantage of extra-pair copulations for females is still debated and unclear. Females do not seem to obtain large direct benefits from mating with extra-pair males, as they do not provide food to the female or parental care to the extra-pair offspring (Arnqvist and Kirkpatrick 2005). On the contrary, it may be costly for female to engage in extra-pair copulation as males uncertain of their paternity may decrease parental care toward own nest (Dixon et al. 1994, Arnqvist and Kirkpatrick 2005, but see Schnitzer et al. 2013). However, one possibility of direct benefit for females is to obtain sperm from their extra-pair mates in case of total or partial infertility of their social partners (Sheldon 1994, Krist and Munclinger 2011). Females may also obtain indirect benefits which can include better or more compatible genes for their offspring (Sheldon et al. 1997, Jennions and Petrie 2000, Neff and Pitcher 2005, Akçay and Roughgarden 2007). The good genes hypothesis (reviewed by Akçay and Roughgarden 2007) states that females may obtain genes for viability or attractiveness from their extra-pair mates. The compatible genes hypothesis

suggests that females may choose males whose genotypes are more compatible with their own and thus increase the fitness of the offspring (Tregenza and Wedell 2000, Neff and Pitcher 2005). Alternatively, female extra-pair behavior may not confer any benefits to females if it is caused by indirect selection on male behavior and cross-sex genetic correlation in this trait (Forstmeier et al. 2011, 2014).

If females strive to obtain good genes from their extra-pair mltes, these males are predicted to be superior to the social partner. Indeed, many studies show that extra-pair males possess larger ornaments than social mates and that social males with larger ornaments lose less paternity in their own nests (Kempenaers et al. 1992, Perreault et al. 1996, Richardson and Burke 1999, Johnsen et al. 2001, Bouwman et al. 2007, Albrecht et al. 2009, reviewed by Akçay and Roughgarden 2007). However, this observation is no general as opposite patterns are frequent (Johnson et al. 2002, Marshall et al. 2007). Similarly, females may also prefer extra-pair males which are older and thus signal good quality and survival ability (Akçay and Roughgarden 2007, Moreno et al. 2010, reviewed by Cleasby and Nakagawa 2012).

In addition to male traits, some environmental factors may influence the distribution of extra-pair paternity. For example, spatial and temporal availability of potential extra-pair partners should be considered (Schlicht et al. 2015).

Extra-pair males are often neighbors (Smith 1988, Neudorf et al. 1997, Mays and Ritchison 2004, Pedersen et al. 2006), but it is not always the case; extra-pair males can also come from distant territories (Woolfenden et al. 2005, Canal et al. 2012a). The factors that influence the distance between social and extra-pair nests have yet to be elucidated and need further investigation. For example, attractiveness may determine the distance a male is able to travel to obtain extra-pair copulation if he need to trade off within-pair and extra-pair paternity. Attractive males that have a lower risk of paternity loss in their own nest may spend more time outside own territory and thus visit more distant territories when searching for a possibility of extra-pair copulation. Consequently, distance between social and extra-pair nest could increase as a function of ornaments of extra-pair males. Similar pattern may occur if females, instead of males, actively seek extra-pair copulations (Double and Cockburn 2000, Tarof and Ratcliffe 2000, Pedersen et al. 2006). Such a foraying female might accept copulation with an extra-pair partner only if the male has larger ornaments than her social mate (Pedersen et al. 2006). Consequently, females that are mated to highly ornamented social partners will likely need to travel farther away to find even better extra-pair male. So in this case distance would increase with the size of ornaments of the social mate. As far as we know, these hypotheses linking male ornaments with the distance to extra-pair nest have not been tested yet.

In this study we tested which traits influenced the probability that a male was cuckolded, and which morphological traits and secondary sexual traits females use to choose extra-pair mates within an available pool of males. In many studies, social and extra-pair males have been compared, but the presence of other potential extra-pair males has rarely been taken into account. We also tested if the extra-pair males were neighbours and whether the traits of females, social or extra-pair males predicted the distance between the nests of cuckolded males and their cuckolders. We used collared flycatchers *Ficedula albicollis* as the study species due to frequent extra-pair paternity (Sheldon et al. 1997, Sheldon and Ellegren 1999), as was also found in our population (Krist et al. 2005, Krist and Munclinger 2011). Males of this species possess two secondary sexual ornaments that could play a role in sexual selection: a forehead patch (Sheldon and Ellegren 1999) and a wing patch (de Heij et al. 2011). Females can use the size of these ornaments to assess the quality of males. We predicted that 1) more ornamented males lose less paternity in their nests, 2) extra-pair males are more ornamented than social males and other males that did not sire any offspring in focal nests, 3) females choose males in their local neighborhood, and 4) the distances between the nests of social mates and extra-pair mates increase with the size of ornaments of the former group.

Material and methods

Study site and species

The study was conducted between 2006 and 2009 in Velky Kosir, Moravia, Czech Republic (49°32′N, 17°04′E, ca 350 m a.s.l). Approximately 350 nest boxes suitable for flycatchers were installed in an oak forest. Collared flycatchers *Ficedula albicollis* are small migratory passerine birds (13 g). Old males are black and white with large wing and forehead patches. Sub-adult males (i.e. those that are in the second year of their lives) have brownish primaries and their wing patches are smaller than those of adult males. Females are dull brown and white and their age cannot be reliably inferred from the plumage Males arrive earlier at the breeding site than females, in mid-April, to establish their territory. After pair-bond formation and nest building, females usually lay 4–8 eggs and start to incubate before the last eggs are laid, resulting in asynchronous hatching.

Field methods

All genotyped nests included in the present study were subject of cross-fostering for purposes of another study (Krist and Munclinger 2011). This sample constituted one quarter to one third of all initiated nests in the population per year. Consequently, we were unable to determine complete extra-pair paternity gains for males. However, as we strived to sample all males in the population, we were able to assign high proportion of extra-pair young to their genetic fathers (see Results) and compare traits of social and extra-pair fathers with those of other available males in the population.

The nests were visited daily when the onset of egg-laying was expected. The first day of egg-laying was noted (1st January = day 1). The laying order was written directly on the eggs, and we measured egg length and width with a digital caliper (± 0.01 mm). Each egg laid was replaced with a dummy egg. Mixed clutches with eggs from different nests were compiled. When the last eggs had been laid and incubation had started, the dummy eggs were replaced with a new clutch with the same size as the original (see Krist and Munclinger 2011 for more details). We refrained from checking the nests every day to avoid disturbance and desertion by females during the incubation period. The daily visits continued 10 d later, when hatching was expected. On day 6 after hatching, the chicks were marked with a numbered ring and weighed (nearest 0.25 g), and we took blood from the tarsal vein which was then stored in ethanol. On day 13, the chicks were measured (tarsus ± 0.01 mm and wing nearest 1 mm) and weighed again. Dead embryos and nestlings were collected and stored in ethanol.

During the feeding period, adults were caught and weighed, and we measured tarsal and wing length. The wing patch was measured with a ruler in the field and its size was calculated as the sum of the visible white patches on the primaries 3 to 8 from the tip of the coverts to the distal part of the wing (in mm). The forehead patch was photographed twice with a ruler aligned along it. The surface area was measured twice after being delimited to the nearest 0.1 mm² in imageJ software ver. 1.49, and the final surface area was defined as the mean of these measurements. A small amount of blood was taken from the tarsus. Male age was determined by wing plumage coloration and it was recorded as adult or sub-adult. As we wanted to assign extra-pair fathers during these field seasons, we strived to catch all males, including those on nests whose chicks were not genotyped, when they were feeding chicks. They were measured, and a blood

sample was taken to allow for genetic analyses and to identify them as genetic sires.

Genotyping

Blood samples and tissue from chicks and dead embryos were used in DNA extraction with DNeasy® Blood and Tissue Kit (Qiagen). They were then genotyped at eight polymorphic microsatellite autosomal loci. A single multiplex PCR using a fluorescently-labelled primers and a Type-it® Microsatellite PCR kit (Qiagen) were used to amplify the microsatellites. The samples were treated for 5 min at 95°C, and then in 30 cycles of 30 s at the same temperature, 90 s at 65°C, 30 s at 72°C, and finally 30 min at 60°C. GeneScan™-500 LIZ® Size Standard (Applied Biosystems) was added to the PCR products, and analyzed with ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems). GeneMarker® ver. 1.9 was used to score the genotypes, and Cervus 3.0.3 (Kalinowski et al. 2007) was used to obtain loci characteristics based on allele frequencies. The eight loci were: Fhu2 (or PTC3) (Ellegren 1992) Cuμ04 (Gibbs et al. 1999) Fhy310, Fhy405, Fhy407, Fhy428, Fhy431, and Fhy452 (Leder et al. 2008). For the first parent, the combined non-exclusion probability for that group of loci was found to be 5.67×10^{-4}. Individuals that could not be genotyped at five or more loci were excluded and the genotypes for 1235 individuals were obtained.

Parentage assignment

Due to the nature of the cross fostering design (Krist and Munclinger 2011), maternal identity was not known and had to be assigned with the use of microsatellite markers. Since the exclusion power was already high for the first parent, and each chick had 8 candidate mothers at most, this maternal assignment proved to be relatively easy and reliable (Krist and Munclinger 2011).

Seven loci were used (Fhy 310 was excluded because of the risk of null alleles) for a comparison of the chicks' genotypes and the genotype of the male paired with the genetic mother. Likelihood approach in Cervus 3.0.3 (Kalinowski et al. 2007) was used to calculate confidence of assignment based on the difference in likelihood scores between the most likely and the second most likely parent. When males matched with all or all but one locus, they were considered to be the genetic fathers. In a few cases that had two mismatches, the chick was indicated as sired by the social father with 95% confidence and we considered it as such.

In other cases, the nestlings were classified as extra-pair young. For these, we compared the offspring's genotype with all male genotypes in the population to determine extra-pair sires. We did this paternity assignment only for years 2007–2009 since in 2006 we blood-sampled only a small proportion of potential sires. All paternities were determined with a 95% maximum likelihood in Cervus Software 3.0.3 (Kalinowski et al. 2007), based on our conservative estimate that we sampled half of the male population. Extra-pair chicks in one nest were usually sired by one extra-pair male, but in eight cases clutch extra-pair paternity was shared by two or more extra-pair males.

To compare traits of social and extra-pair males, we formed groups called trios which were unique combinations of ring numbers of the social male, extra-pair male, and female (Kalinowski et al. 2007). In cases where one female had extra-pair young with two or more different extra-pair males, we considered her to be a part of two separate trios. In total, 59 trio combinations were used for the comparison of traits.

Spatial analyses and potential males

The 350 nest boxes were distributed among four different plots in the study area and we were interested in whether females chose extra-pair males randomly from all available males or if they preferred males from their own neighborhood. To precisely determine the distance between the social and extra-pair nest and social and all other available nests, a linear distance matrix was designed using GIS software Qgis 2.4.0 (<http://qgis.org>). All coordinates were presented in degrees, minutes, seconds, and a WGS 84 projection was used. Analyses were conducted separately for each year to avoid superimposition of occupied nests. We kept records of every nest occupied by collared flycatchers from each breeding season even if information about a male was missing. This information may have been missing because nests were deserted or predated, or we were unable to capture males but they were still available as potential extra-pair males. We compared the distances between the female's nest and the nest of the extra-pair father with the mean distance to all nests occupied by collared flycatchers either in the whole study area or only within the same plot. In the latter case, we excluded nine cases where the extra-pair father was in fact found on another plot. We performed the second analysis because most extra-pair sires were found within the same plot where the female bred (see Results).

We wanted to identify all potential extra-pair males available for the females who obtained extra-pair young, and compare those males with the social and actual extra-pair fathers. As the actual extra-pair father was often a close neighbor (see Results), we considered as potential extra-pair fathers only those individuals that bred as close as or closer to the focal nest than the actual extra-pair father. We draw a buffer zone around each social nest with a radius equal to the distance between the nest and the extra-pair father. We added 20 m (the mean of half the distance between two adjacent nest-boxes) to the real distance to include nests with approximately the same distances from the social to the extra-pair nests. Using the plug-in 'spatial queries ver. 0.1' in QGIS, we identified every occupied nest that lay within the focal buffer and obtained the list of potential males for a focal nest.

Moreover, we repeated this spatial analysis when we considered the temporal availability of potential extra-pair sires. Not all males breeding in the vicinity around the focal nest were indeed available as extra-pair sires since some of them arrived after the fertile period of the focal female ended. Therefore, in the second analysis we included among potential sires only males that bred at the same time or earlier as the focal female.

Statistical analyses

To analyze the relationship between social parents traits and the occurrence of extra-pair offspring in the nest

(prediction 1), we used a generalized linear mixed model (GLMM) with a binomial distribution (presence/absence of extra-pair offspring in the social father's nest) and a logit function as link. We ran this analysis with the glmer function from package lme4 ver. 1.1-6 (Bates et al. 2015) in R studio ver. 0.98.501 (R Core Team, <www.R-project.org/>). The presence or absence of extra-pair young in the nest was the model's response variable. The predictors included two morphological characteristics for social males and females (tarsal length and wing length), wing patch size for both sexes, and the forehead patch size and age of males (adults or sub-adults). We were unable to include female age among variables, despite its potential relevance for extra-pair paternity (Moreno et al. 2015), since females have no delayed plumage maturation. The year of study and the ring numbers of both males and females were used as random factors. The p-values were obtained by comparing the deviance between the full model and the model without the considered factor. The difference between the two deviances was compared with a chi-square distribution (Crawley 2007).

To compare the traits of males according to their status (social, extra-pair, potential; prediction 2) separate linear mixed models were run. Father status was used as a predictor in all models, and response variables were tarsus, wing length, wing patches, forehead patch size, or the female laying date for each respective model. For potential males, these traits were averaged within trios and consequently the number of potential males equalled the number of trios. The lmer function with the package lme4 (R package ver. 1.1-8, Bates et al. 2015) and lmerTest (R package ver. 2.0-29, Kuznetsova et al. 2015) were used. Age is a binary type of data so we did not compute averages for potential males and we used the glmer function with a logit link function instead of the lmer). The random factor was the trio identity for all models.

To test whether females chose males randomly with respect to the distance to their nests (prediction 3), we compared the distance between females' nest and the nest of extra-pair sires with the mean distance to all available nests (including the distance between the social and extra-pair nest). We used a paired t-test assuming unequal variance between groups. This analysis was run twice: firstly for all trios and secondly for only trios where the extra-pair father was from the same plot. We also used three separate linear mixed models to test whether morphological characteristics of social males, extra-pair males and females predict the distance between the social and extra-pair nest (prediction 4).

Data available from the Dryad Digital Repository: <http://dx.doi.org/10.5061/dryad.043tn> (Edme et al. 2016).

Results

Distribution of extra-pair paternity

In total, 160 nests were used for the cross-fostering experiments between 2006 and 2009. Females from 143 nests were genotyped while social males were genotyped from 135 of them. In total 941 offspring were genotyped and for 800 of them the social father was known. Extra-pair fathers sired 214 of these 800 offspring (26.8%). Three social males were hybrids with the pied flycatcher and their nests were excluded from all analyses. Consequently, sample size when testing prediction 1 was 132 nests. We also did not use data from year 2006 for paternity assignment. Our sample size for paternity assignment was then 181 extra-pair young from 2007–2009. We were able to assign genetic fathers to 114 of them (63.0%). These 114 extra-pair young formed 59 trios (unique combination of social male, female and extra-pair male). Those 59 trios represent our basic sample size for testing predictions 2, 3 and 4.

Paternity loss and adults' ornamentation

We checked for collinearity among the predictors of the model. Variance inflation factor (VIF) revealed that age and wing patches were collinear (VIF > 2) (Graham 2003) (Table 1). Therefore, we re-ran the model once without

Table 1. A test whether traits of social male and female influence the probability of having extra-pair young in a nest (prediction 1, n = 132 nests). GLMM model with presence/absence of extra-pair young in the nest (response variable) and males and females traits (predictors). Random effect was females and males ring and year. Positive estimates indicate a high probability of having extra-pair young in the nest (SE = standard error, DF = degree of freedom, VIF = variance inflation factor).

Random effects	Variance				
Male ring	0.000				
Female ring	0.618				
Year	0.008				
Residual	0.902				
Fixed effects	Estimate ± SE	F-value	DF	p-values	VIF
Intercept	−65.17 ± 23.57				
Julian day	0.216 ± 0.070	9.42	1,126	<0.001	1.26
Male age	2.38 ± 1.08	4.93	1,126	0.020	3.78
Male forehead patch	0.002 ± 0.002	1.35	1,126	0.208	1.05
Male wing patch	0.090 ± 0.028	4.92	1,126	<0.001	3.55
Male wing length	3.59 ± 1.68	0.74	1,126	<0.001	1.50
Male tarsus length	−0.066 ± 0.514	0.06	1,126	0.899	1.18
Female wing patch	−0.011 ± 0.029	0.04	1,126	0.670	1.08
Female wing length	1.61 ± 1.47	0.06	1,126	0.245	1.21
Female tarsus length	−0.432 ± 0.453	0.52	1,126	0.303	1.09

wing patches, and for a second time with patch size but without age. Wing patch was a significant predictor of loss of paternity ($F_{1,125} = 5.57$, $p < 0.001$) when tested in isolation, but age was not ($F_{1,106} = 0.12$, $p = 0.417$). We found that wing patches were larger in cuckolded males than in non-cuckolded ones (all males: 54.37 ± 13.76 mm², n = 73, and 46.58 ± 14.68 mm², n = 59, adults only: 58.77 ± 8.25 mm², n = 62, and 53.1 ± 8.59 mm², n = 45). Males who lost paternity had longer wings than non-cuckolded males (8.32 ± 0.18 cm and 8.25 ± 0.16 cm respectively) (Table 1) and the size of the forehead patch area did not significantly differ (mean cuckold vs non-cuckold 53.35 ± 9.32 mm² and 50.18 ± 9.26 mm²) (Table 1). Tarsal length was not different between the two types of males. The females who cuckolded males showed no differences in any of the traits from those who did not (Table 1). There was also a higher probability to have extra-pair young in the focal nest as the breeding season progressed (Table 1).

Comparison between social mates, extra-pair mates and potential mates

We identified 46 extra-pair males and 45 social males and females in these 59 trios. For several reasons, the number of males does not need to equal the number of trios. For example, some extra-pair mates had extra-pair young in more than one nest (range 1–4 nests). Other individuals bred and were cuckolded in more than one year. The number of potential mates (the actual extra-pair mates were excluded from the list of potential mates) in the buffer zone around the focal nest ranged from 0 to 127 males (all plots: mean $\pm$ SD = 21 ± 27 males; within plots: mean $\pm$ SD = 8.6 ± 7.4 males) or from 0 to 71 if only nests with the same or earlier laying date were considered to provide potential sires for focal extra-pair young (all plots: 10.83 ± 14.89, within plots: 5.31 ± 7.18).

Forehead patch, tarsus length and date of laying did not differ in the three categories of males (social, extra-pair, and potential mates) (Fig. 1, Table 2). However the wing
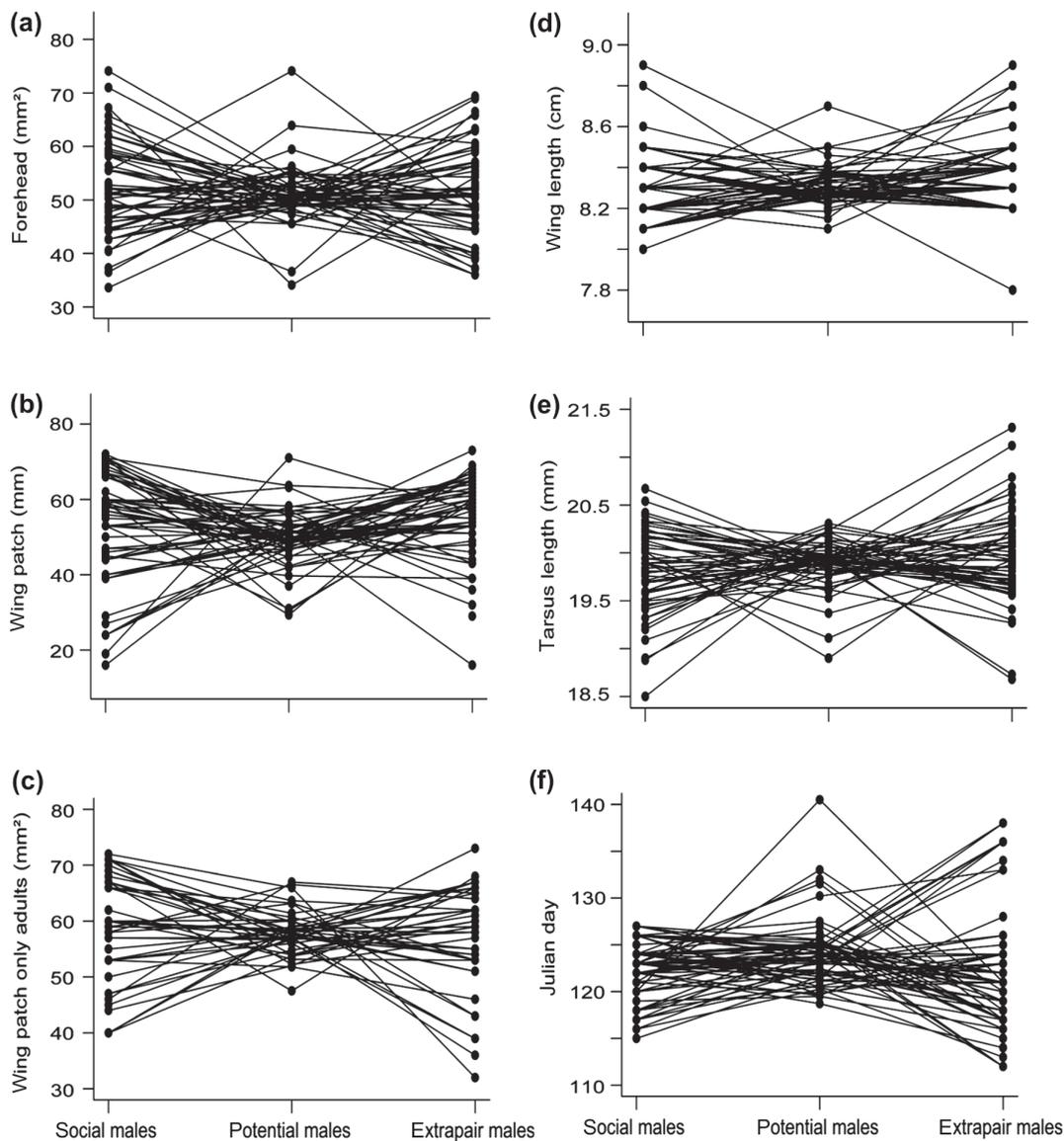

Figure 1. Comparisons of morphological and life-history traits between social, potential and extra-pair fathers (n = 59). Whereas values for social and extra-pair males are individual data points, each point for potential fathers represents the mean of all potential fathers for that social nest. Lines connect values related to the same focal nest.

Table 2. Comparison between traits of social, extra-pair and potential males (prediction 2, n = 59 trios). All males breeding in the zone around social nest were considered as potential sires. Each line shows results of separate model with male status as the predictor and morphological or life-history trait as a response variable. DF = degree of freedom. Trio identity was fitted as a random effect.

	Random effects		Fixed effects			Means ± SD		
	Trio (variance)	Residual	F test	DF	p-values	Social	Extra-pair	Potential
Forehead (mm²)	0.000	6279	0.491	2,168	0.612	52.54 ± 9.08	51.51 ± 8.64	50.40 ± 9.34
Wing patch (mm²)	0.000	129.3	4.439	2,170	0.013	52.91 ± 14.42	55.96 ± 10.91	48.90 ± 18.32
Wing patch (adults only) (mm²)	0.00	67.5	0.523	2,120	0.593	58.70 ± 9.62	56.85 ± 9.65	57.87 ± 4.05
Wing length (cm)	0.004	0.025	4.180	2,104.5	0.017	8.31 ± 0.19	8.39 ± 0.02	8.29 ± 0.17
Tarsus (mm)	0.000	0.170	1.648	2,170	0.195	19.82 ± 0.3	19.96 ± 0.49	19.94.00 ± 0.52
Julian day	1.422	21.7	3.366	2,115.5	0.037	122.25 ± 3.07	122.03 ± 6.81	123.90 ± 6.68
Age	0.000	1.09	6.282	2	< 0.001	Adults: 48 sub-adults: 11	Adult: 57 sub-adults: 2	Adults: 683 sub-adults: 255

characteristics differed among males. The wings of extra-pair mates were longer than those of social and potential mates (Fig. 1, Table 2). Wing patch size was also larger in extra-pair males than in potential mates (Fig. 1, Table 2), but there was no difference between social and extra-pair mates. When we controlled for age (adults have larger patches than sub-adults), wing patch size was no longer different (Fig. 1, Table 2). Extra-pair fathers were older than social mates (estimate ± SE = 1.876 ± 0.792, z-value = –2.369, p = 0.017) and potential extra-pair mates (estimate ± SE = 2.366 ± 0.722, z-value = –3.277, p < 0.001). There was no age difference between social mates and potential mates. Social males and extra-pair males had females who bred at the same time, but potential males had females who laid their eggs later (Fig. 1, Table 2). The results are closely similar when only males breeding at the same time or earlier are considered as potential sires of extra-pair chicks (see Methods). The only natural exception is then the breeding date that is by

definition earlier for this restricted set of potential males (Supplementary material Appendix 1).

Distance of extra-pair sires

The selection of extra-pair mates was not random, and females mostly chose an extra-pair mate close to their nest. In 50 of 59 cases, males bred within the same plot as females (chi-square = 28.491, DF = 1, p-value < 0.001). A comparison was done between the actual distance between social and extra-pair nests and that between social nests and the mean distance to all other available nests of collared flycatchers (prediction 3). We carried out this analysis twice. The first time we took into account only the trios with the extra-pair mate within the same plot as the social nest. The second time, we ran the analysis with all 59 trios which also included extra-pair sires breeding on other plots. The actual distances were shorter than the mean distances to all available nests
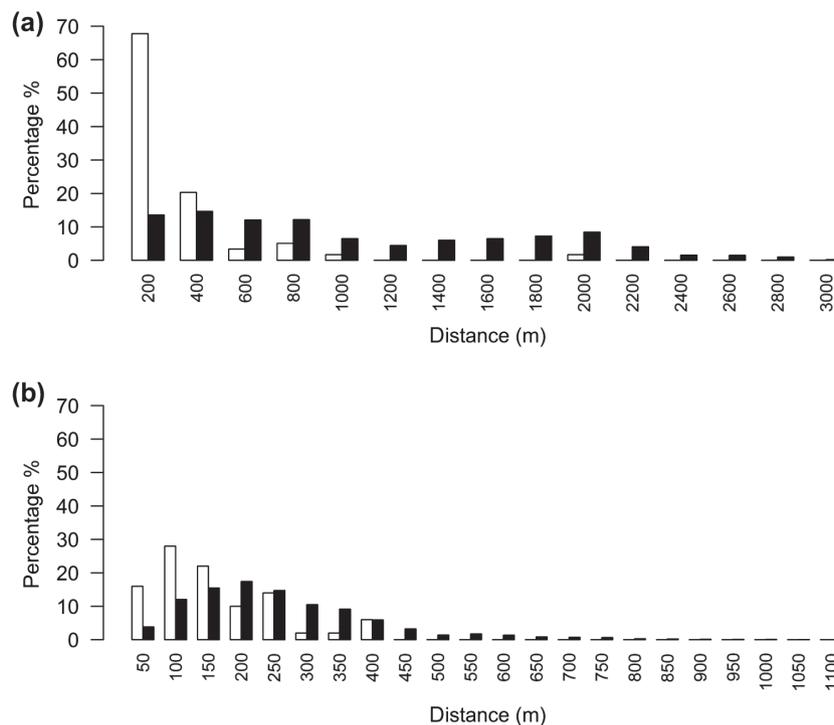

Figure 2. Distribution of distances between social nests and extra-pair nests (open bars) or social nests and all available nests (filled bars). Panel (a) considers nests in all plots, panel (b) considers only extra-pair nests and all available nests that were located on the same plot as the social nest.

in both cases (within plot comparison: 136.5 ± 90.2 vs 239.7 ± 64.7 m, t = −7.24, DF = 49, p < 0.001; all data: 222.6 ± 291.7 vs 962.8 ± 246.0 m, t = −15.9, DF = 58, p < 0.001, Fig. 2). In the test of our prediction 4 we found that neither female traits nor the extra-pair male traits influenced the actual distances between social nest and that of extra-pair father (Table 3, 4). However, social males with longer wings had their extra-pair rival closer than was the case of social males with shorter wings (Table 5).

Discussion

The aim of this study was to determine which traits influenced the distribution of extra-pair paternity in the collared flycatcher. We only genotyped nestlings in a fraction of nests but sampled the whole population of breeding males. Consequently we were able to assign genetic fathers to a high proportion of genotyped chicks which enabled us to compare traits of social and extra-pair fathers on solid sample size that was gathered over three to four years. We found that more ornamented males had more extra-pair young in their nest than less ornamented males. This contradicts with our first prediction that more ornamented males lose less paternity. Our second prediction about female preference for more ornamented extra-pair males was only partly supported since there was no difference in ornament size between the social, potential and extra-pair mates besides that caused by age. However, extra-pair males were older and had longer wings than both social and potential mates. In support of our third prediction we found that females choose extra-pair mates among close neighbors. In contrast, we found little evidence for our last prediction as distance between social and extra-pair nest is not affected by traits of either males or females.

More ornamented males are more cuckolded

Considering the male traits, only wing length and wing patches appeared to play a role in the probability of being cuckolded. However, our results are in contradiction with our prediction 1. In fact, males with longer wings and larger patches had more extra-pair offspring in their nests. In many previous studies it has been shown that more ornamented males lose less paternity (Møller and Ninni 1998, but see Sundberg and Dixon 1996, Bouwman et al. 2007)

Table 3. Distance between social and extra-pair nest as a function of female traits (n = 59). Linear mixed model with distance as response variable and female characteristics as predictors. (SE = standard error, DF = degree of freedom). A negative estimate indicates a decrease in distance.

Random effects	Variance			
Female ring	< 0.001			
Residual	89814			
Fixed effects	Estimate ± SE	DF	t-value	p-values
Intercept	2914 ± 3231	54	0.902	0.371
Female wing patch	−1.279 ± 5.158	54	−0.248	0.805
Female wing length	−220.9 ± 286.7	54	−0.771	0.444
Female tarsus length	−15.4 ± 80.9	54	−0.191	0.785
Julian day	−4.38 ± 13.42	54	−0.334	0.740

Table 4. Distance between social and extra-pair nest as a function of traits of the extra-pair male (n = 59). Linear mixed model with distance as response variable and extra-pair male characteristics as predictors (SE = standard error, DF = degree of freedom). A negative estimate indicates a decrease in distance.

Random effects	Variance			
Extra pair male ring	5980			
Residual	88654			
Fixed effects	Estimate ± SE	DF	t-value	p-values
Intercept	1072 ± 2071	42.5	0.548	0.607
Forehead patch	0.539 ± 0.568	43.6	0.948	0.348
Age	130.0 ± 281.7	46.6	0.462	0.646
Wing patch	6.15 ± 5.22	41.5	1.179	0.245
Wing length	−159.5 ± 311.7	45.7	−0.512	0.611
Tarsus length	−14.28 ± 102.20	44.3	−0.140	0.890
Julian day	1.24 ± 7.16	40.8	0.173	0.863

and increase their total reproductive success (Sheldon and Ellegren 1999). We need to stress, however, that the presence of extra-pair young in the nests of more ornamented males does not mean that their total reproductive success is lower than that of non-cuckolded males. Males who lost paternity in their own nests may have gained extra-pair paternity in other nests to compensate for the potential loss of paternity in their own. Our finding that males with longer wings lost more paternity in own nest but were favorite extra-pair fathers highlights that the same trait that causes loss of within-pair paternity may lead to the gain of extra-pair paternity. Unfortunately, we were unable to determine the total extra-pair paternity per male in this study, since we took DNA samples only from a subsample of nests that were involved in the cross-fostering experiment.

Although our results are contrary to what we predicted, they could be explained by different mate guarding strategies among males. Males with larger patches might guard their females less and instead spend time outside their territories in trials to attract other females and sire extra-pair young. On the other hand, males of a lower quality may guard their mates more intensively to avoid losing paternity (Kokko and Morrell 2005). Regardless of this, the loss of paternity depends on both male and female strategies (Kokko and Morrell 2005), and it seems difficult for males to efficiently guard their females and obtain extra-pair copulations for

Table 5. Distance between social and extra-pair nest as a function of traits of the social male (n = 59). Linear mixed model with distance as response variable and social males characteristics as predictors (SE = standard error, DF = degree of freedom). A negative estimate indicates a decrease in distance.

Random effects	Variance			
Social male ring	0.000			
Residual	8727			
Fixed effects	Estimate ± SE	DF	t-value	p-values
Intercept	5386 ± 3097	50.6	1.739	0.088
Age	−68.1 ± 165.6	50.6	−0.411	0.682
Forehead patch	−0.364 ± 0.461	50.6	−0.789	0.433
Julian day	−13.0 ± 16.7	50.6	−0.779	0.439
Wing patch	0.229 ± 4.56	50.6	0.050	0.96
Wing length	−515.5 ± 231.0	50.6	−2.231	0.030
Tarsus length	45.9 ± 99.5	50.6	0.462	0.645

themselves at the same time. However, mate guarding may be inefficient when females are highly motivated to escape their partners guard (Kokko and Morrell 2005). This might be the case in blue tits *Cyanistes caeruleus*, where mate guarding did not protect paternity of social males, (Kempenaers et al. 1992).

Moreover, a further trade-off may exist between mate guarding and other male activities such as territory defence and the solution of this trade-off may depend on male age and experience. It has been found that young collared flycatchers males defend their territories more aggressively than old males (Garamszegi et al. 2006) and thus may be more exposed to loss of paternity because they provide more opportunities for females to seek out extra-pair copulation. Studies on the role of male age in the loss or gain of paternity are contradictory (Kempenaers et al. 1992, Moreno et al. 2010, Bowers et al. 2015, reviewed by Cleasby and Nakagawa 2012) but this often seems to be an important factor. In our case, age does not appear to be an important factor by itself for the loss of paternity, but only becomes significant when linked to the size of the wing patch.

Similar to Canal et al. (2012b), we also found that males that paired with females laying earlier in the breeding season, had fewer extra-pair young. Males arrive before females to establish territories (Harnos et al. 2015), and better territories (quality of the nest for example) are occupied by more experienced males (Askenmo 1984, see also Pärt and Qvarnström 1997). It is possible that because they arrived earlier, the males who paired first are better able to secure their females from extra-pair copulation. In some species, females choose a mate based on his traits and the quality of the territory (Alatalo et al. 1986). Females may be less motivated to cuckold their mates if these are good males who arrived first at the breeding site (Potti 1998). To arrive and breed earlier could also be advantageous because when their females finish laying, those males could look for extra-pair copulations without the danger of losing paternity. On the other hand, late breeding can represent a double cost for males. First, late breeders are less able to successfully raise healthy offspring. Secondly, it may encourage females to look for higher quality extra-pair males to provide a genetic advantage to the offspring (Qvarnström et al. 2000).

Extra-pair males are older and have longer wings than other potential partners

Extra-pair partners were older males with longer wings than the mean for all available males which indicates that either females assess the quality of the males around their nests or males seek extra-pair copulations based on their phenotype. As we lack information about movements during pairing, we do not know if male or female collared flycatchers visit many different territories or how often they may engage in this behavior. We can consider several possible scenarios for extra-pair mate choice. Firstly, females could eavesdrop on male–male singing contests (Mennill et al. 2002). If males' songs provide information about male quality or territory, females may listen to them and choose to visit males who seem likely to be good extra-pair males. Secondly, females could be more prospective and do extra-territorial forays to directly obtain information on potential extra-pair mates

(Double and Cockburn 2000). In line with these two scenarios it has been found that captive females of the pied flycatchers often actively solicit extra-pair copulations (Drevon and Slagsvold 2005). Thirdly, females may stay in their territories and accept copulation from other males who intrude into her partner's territory (Hung et al. 2009, see also Moreno et al. 2015). Lastly, both females and males might leave their territories to find suitable extra-pair partners (Mays and Ritchison 2004). Unfortunately, little is known about the behavior of *Ficedula* flycatchers when seeking extra-pair copulations in the wild. Such behavioral data is badly needed but difficult to obtain. Employment of new technologies like automated radio telemetry can shed light on such secret behaviors as extraterritorial forays (Ward et al. 2013, 2014).

Females may obtain direct or indirect benefits when seeking extra-pair copulation. To obtain genetic benefits, they may copulate with extra-pair males superior to their social partner with respect to good or compatible genes (reviewed by Hsu et al. 2015). In our study, females did not choose males with larger secondary sexual ornaments that may be indicative of good genes (Neff and Pitcher 2005). There was no difference in the wing patches or forehead patches between social and extra-pair males. Moreover, we did not find evidence of superiority of extra-pair chicks in a previous study which also suggested that females did not obtain good or compatible genes for their offspring from extra-pair copulation (Krist and Munclinger 2011). However, that study was aimed mainly on short-term fitness benefits and we were unable to exclude superiority of extra-pair young in adult phase of life (Krist and Munclinger 2011). Here we found that extra-pair partners are older than social mates. By choosing older partners, females might obtain genes for offspring longevity. This possibility remains to be tested.

Besides good or compatible genes females might obtain direct benefits from extra-pair copulations. For example, female red-winged blackbirds *Agelaius phoeniceus* were found to forage in territories of their extra-pair partners (Gray 1997). But this type of direct benefit may be rare, and probably does not play a major role in collared flycatchers since this species does not defend feeding territories (Cramp and Perrins 1993). More general direct benefits from extra-pair copulation may be the insurance of the clutch against social mate infertility (Sheldon 1994), since sperm depletion or infertility may devalue the reproductive investment of females. In theory, females may be able to detect infertility or the poor sperm quality of their partners and choose extra-pair mates who possess traits indicative of good quality sperm (Sheldon 1994). However this phenotype-linked fertility hypothesis has weak empirical support since ejaculate or sperm traits are usually uncorrelated to male phenotypes (Mautz et al. 2013). In line with this we also found no difference in ornament size of extra-pair and social males despite our previous study tentatively suggested that females might engage in extra-pair copulations to insure clutch fertility (Krist and Munclinger 2011). Useful steps would be to take into consideration sperm traits to test if they are linked to male phenotype and have effect on paternity in the collared flycatcher. In the sister species, the pied flycatcher, support for the linkage between sperm traits and phenotype is mixed (Calhim et al. 2009, Lifjeld et al. 2012) while data on

effect of sperm traits on paternity are needed (Cramer et al. 2013).

Copulation with neighboring males

Some previous studies have found that extra-pair males are close neighbors (Kempenaers et al. 1992, 1995, Perreault et al. 1996, Valera et al. 2003). In agreement with our prediction, almost all of the extra-pair males bred on the same plot as the females, with a mean distance of about 136 m. There are several explanations for why extra-pair sires should be close neighbors. First, males can allow their extra-pair partners to forage in their territories (Gray 1997). Another potential benefit may be a decreased risk of predation since the presence of other males means better vigilance for predators (Gray 1997, Eliassen and Jørgensen 2014). Third, if females only need sperm to be sure of fertilizing their eggs, they may copulate with a neighbor without considering the male traits. Here also mate guarding by social males may play a role since females probably have more opportunities to meet with neighbors and such meetings may take less time then travelling to distant territories and thus have greater chance of being hidden to the guarding mate. In line with this 'opportunity' or 'random pairing' hypothesis we found little evidence for the effect of female or male traits on the distance between social and extra-pair nest. The only pattern we found was a decrease of the distance with increase of the wing length of the social male. Currently we have no explanation of this finding. Surely more studies are needed to elucidate if male and female morphological and ornamental traits generally affect spatial distribution of extra-pair paternity.

Conclusions

In summary, males with longer wings and larger patches lost more paternity in their nests while older males with longer wings were more successful in gaining paternity in other nests. This finding suggests either that male effort in seeking of extra-pair mating opportunities depend on their phenotype or that females assess the quality of potential males around their nests when choosing extra-pair mates. On the other hand, females usually choose close neighbors as extra-pair partners and the distance to their nests is not dependent on their traits. This suggests, on the contrary, that females are not very choosy when selecting extra-pair mate. This is expected if they engage in extra-pair copulation to insure against potential infertility of their mate. Further research should strive to determine the factors that influence the distance between social and extra-pair nests and the behavior of all involved participants. On the one hand, in many species it is not well known if females or extra-pair males usually travel to obtain extra-pair copulations. On the other hand, the intensity of mate guarding by social mates is also likely to vary, but similarly to extraterritorial forays has rarely been quantified (but see Double and Cockburn 2000, Stutchbury et al. 2005, Woolfenden et al. 2005). Focus on sperm traits should also provide useful information since we lack general knowledge on the linkage between sperm morphology, male attractiveness and success in sperm competition.

References

Akçay, E. and Roughgarden, J. 2007. Extra-pair paternity in birds: review of the genetic benefit. – Evol. Ecol. Res. 9: 855–868.

Alatalo, R. V., Lundberg, A. and Glynn, C. 1986. Female pied flycatchers choose territory quality and not male characteristics. – Nature 323: 152–153.

Albrecht, T., Vinkler, M., Schnitzer, J., Polakova, R., Munclinger, P. and Bryja, J. 2009. Extra-pair fertilizations contribute to selection on secondary male ornamentation in a socially monogamous passerine. – J. Evol. Biol. 22: 2020–2030.

Arnqvist, G. and Kirkpatrick, M. 2005. The evolution of infidelity in socially monogamous passerines: the strength of direct and indirect selection on extrapair copulation behavior in females. – Am. Nat. 165: 26–37.

Askenmo, C. E. H. 1984. Polygyny and nest site selection in the pied flycatcher. – Anim. Behav. 32: 972–980.

Bates, D., Maechler, M., Bolker, B. and Walker, S. 2015. Fitting linear mixed-effects models using lme4. – J. Stat. Softw. 67: 1–48.

Bouwman, K. M., Van Dijk, R. E., Wijmega, J. J. and Komdeur, J. 2007. Older male reed bunting are more successful at gaining extrapair fertilizations. – Anim. Behav. 73: 15–27.

Bowers, E. K., Forsman, A. M., Masters, B. S., Johnson, B. S. M., Johnson, L. S., Sakaluk, S. K. and Thompson, C. F. 2015. Increased extra-pair paternity in broods of aging males and enhanced recruitment of extra-pair young in a migratory bird. – Evolution 69: 2533–2541.

Calhim, S., Lampe, H. M., Slagsvold, T. and Birkhead, T. R. 2009. Selection on sperm morphology under relaxed sperm competition in a wild passerine bird. – Biol. Lett. 5: 58–61.

Canal, D., Jovani, R. and Potti, J. 2012a. Male decisions or female accessibility? Spatiotemporal patterns of extra pair paternity in a songbird. – Behav. Ecol. 23: 1146–1153.

Canal, D., Jovani, R. and Potti, J. 2012b. Multiple mating opportunities boost protandry in a pied flycatcher population. – Behav. Ecol. Sociobiol. 66: 67–76.

Cleasby, I. R. and Nakagawa, S. 2012. The influence of male age on within-pair and extra-pair paternity in passerines. – Ibis 154: 318–324.

Cramer, E. R. A., Laskemoen, T., Kleven, O., LaBarbera, K., Lovette, I. J. and Lifjeld, J. T. 2013. No evidence that sperm morphology predicts paternity success in wild house wrens. – Behav. Ecol. Sociobiol. 67: 1845–1853.

Cramp, S. and Perrins, C. M. 1993. Handbook of the birds of Europe the Middle East and north Africa. The birds of the Western Palearctic. – Oxford Univ. Press.

Crawley, M. J. 2007. The R book. – John Wiley and Sons.

de Heij, M. E., Gustafsson, L. and Brommer, J. E. 2011. Experimental manipulation shows that the white wing patch in collared flycatchers is a male sexual ornament. – Ecol. Evol. 1: 546–555.

Dixon, A., Ross, D. and Malley, S. L. C. O. 1994. Paternal investment inversely related to degree of extra-pair paternity in the reed bunting. – Nature 371: 698–700.

Double, M. and Cockburn, A. 2000. Pre-dawn infidelity: females control extra-pair mating in superb fairy-wrens. – Proc. R. Soc. B 267: 465–470.

Drevon, T. and Slagsvold, T. 2005. When and from whom do female pied flycatchers (*Ficedula hypoleuca*) solicit copulations. – Behaviour 142: 1059–1076.

Edme, A., Munclinger, P. and Krist, M. 2016. Female collared flycatchers choose neighbouring and older extra-pair partners from the pool of males around their nests. – Dryad Digital Repository, <http://dx.doi.org/10.5061/dryad.043tn>.

Eliassen, S. and Jørgensen, C. 2014. Extra-pair mating and evolution of cooperative neighbourhoods. – PLoS One 9: e99878.

Ellegren, H. 1992. Polymerase-chain-reaction (PCR) analysis of microsatellites – a new approach to studies of genetic relationships in birds. – Auk 109: 886–895.

Forstmeier, W., Martin, K., Bolund, E., Schielzeth, H. and Kempenaers, B. 2011. Female extrapair mating behavior can evolve via indirect selection on males. – Proc. Natl Acad. Sci. USA 108: 10608–10613.

Forstmeier, W., Nakagawa, S., Griffith, S. C. and Kempenaers, B. 2014. Female extra-pair mating: adaptation or genetic constraint? – Trends Ecol. Evol. 29: 456–464.

Garamszegi, L. Z., Rosivall, B., Hegyi, G., Szöllösi, E., Török, J. and Een, M. 2006. Determinants of male territorial behavior in a hungarian collared flycatcher population: plumage traits of residents and challengers. – Behav. Ecol. Sociobiol. 60: 663–671.

Gibbs, H. L., Tabak, L. M. and Hobson, K. 1999. Characterization of microsatellite DNA loci for a neotropical migrant songbird, the Swainson's thrush (*Catharus ustulatus*). – Mol. Ecol. 8: 1551–1561.

Graham, M. H. 2003. Confronting multicollinearity in ecological multiple regression. – Ecology 84: 2809–2815.

Gray, E. M. 1997. Do female red-winged blackbirds benefit genetically from seeking extra-pair copulations? – Anim. Behav. 53: 605–623.

Griffith, S. C., Owens, I. P. F. and Thuman, K. A. 2002. Extra pair paternity in birds: a review of interspecific variation and adaptative function. – Mol. Ecol. 11: 2195–2212.

Harnos, A., Nora, A., Kovacs, S., Lang, Z. and Csorgo, T. 2015. Increasing protandry in the spring migration of the pied flycatcher (*Ficedula hypoleuca*) in central Europe. – J. Ornithol. 156: 543–546.

Hsu, Y. H., Schroeder, J., Winney, I., Burke, T. and Nakagawa, S. 2015. Are extra-pair males different from cuckolded males? A case study and a meta-analytic examination. – Mol. Ecol. 24: 1558–1571.

Hung, S., Tarof, S. A. and Stutchbury, B. J. M. 2009. Extra-pair mating tactics and vocal behavior of female acadian flycatchers. – Condor 111: 653–661.

Jennions, M. D. and Petrie, M. 2000. Why do females mate multiply? A review of the genetic benefits. – Biol. Rev. 75: 21–64.

Johnsen, A., Lifjeld, J., Andersson, S., Ornborg, J. and Amundsen, T. 2001. Male characteristics and fertilisation success in bluethroats. – Behaviour 138: 1371–1390.

Johnson, L. S., Hicks, B. G. and Masters, B. S. 2002. Increased cuckoldry as a cost of breeding late for male house wrens (*Troglodytes aedon*). – Behav. Ecol. 13: 670–675.

Kalinowski, S. T., Taper, M. L. and Marshall, T. C. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. – Mol. Ecol. 16: 1099–1106.

Kempenaers, B., Verheyen, G. R., Van den Broeck, M., Burke, T., Van Broeckhoven, C. and Dhondt, A. A. 1992. Extrapair paternity results from female preference for high quality males in the blue tit. – Nature 357: 494–496.

Kempenaers, B., Verheyen, G. R. and Dhondt, A. A. 1995. Extrapair paternity in the blue tit (*Parus caeruleus*): female choice, male characteristics, and offspring quality. – Behav. Ecol. 8: 481–492.

Kokko, H. and Morrell, L. J. 2005. Mate guarding, male attractiveness, and paternity under social monogamy. – Behav. Ecol. 16: 724–731.

Krist, M. and Munclinger, P. 2011. Superiority of extra-pair offspring: maternal but not genetic effects as revealed by a mixed cross-fostering design. – Mol. Ecol. 20: 5074–5091.

Krist, M., Nádvorník, P., Uvírová, L. and Bureš, S. 2005. Paternity covaries with laying and hatching order in the collared flycatcher *Ficedula albicollis*. – Behav. Ecol. Sociobiol. 59: 6–11.

Kuznetsova, A., Brockhoff, P. B. and Christensen, R. H. B. 2015. lmerTest: tests in linear mixed effects models. – R package ver. 2.0-29, <http://CRAN.R-project.org/package=lmerTest>.

Leder, E. H., Karaiskou, N. and Primmer, C. R. 2008. Seventy new microsatellites for the pied flycatcher, *Ficedula hypoleuca* and amplification in other passerine birds. – Mol. Ecol. Resour. 8: 874–880.

Lifjeld, J. T., Laskemoen, T., Kleven, O., Pedersen, A. T. M., Lampe, H. M., Rudolfsen, G. and Slagsvold, T. 2012. No evidence for pre-copulatory sexual selection on sperm length in a passerine bird. – PLoS One 7: e32611.

Marshall, R. C., Buchanan, K. L. and Catchpole, C. K. 2007. Song and female choice for extrapair copulations in the sedge warbler, *Acrocephalus schoenobaenus*. – Anim. Behav. 73: 629–635.

Mautz, B. S., Møller, A. P. and Jennions, M. D. 2013. Do male secondary sexual characters signal ejaculate quality? A meta-analysis. – Biol. Rev. 88: 669–682.

Mays, H. L. and Ritchison, G. 2004. The effect of vegetation density on male mate guarding and extra-territorial forays in the yellow-breasted chat (*Icteria virens*). – Naturwissenschaften 91: 195–981.

Mennill, D. J., Ratcliffe, L. M. and Boag, P. T. 2002. Female eavesdropping on male song contests in songbirds. – Science 296: 873.

Møller, A. P. and Ninni, P. 1998. Sperm competition and sexual selection: a meta-analysis of paternity studies of birds. – Behav. Ecol. Sociobiol. 43: 345–358.

Moreno, J., Martínez, J. G., Morales, J., Lobato, E., Merino, S., Tomás, G., Vásquez, R. A., Möstl, E. and Osorno, J. L. 2010. Paternity loss in relation to male age, territorial behaviour and stress in the pied flycatcher. – Ethology 116: 76–84.

Moreno, J., Martinez, J. G., Gonzalez-Braojos, S., Cantarero, A., Ruiz-de-Castaneda, R., Precioso, M. and Lopez-Arrabe, J. 2015. Extra-pair paternity declines with female age and wing length in the pied flycatcher. – Ethology 121: 501–512.

Neff, B. D. and Pitcher, T. E. 2005. Genetic quality and sexual selection: an integrated framework for good genes and compatible genes. – Mol. Ecol. 14: 19–38.

Neudorf, D. L., Stutchbury, B. J. M. and Piper, W. H. 1997. Covert extraterritorial behavior of female hooded warblers. – Behav. Ecol. 8: 595–600.

Pärt, T. and Qvarnström, A. 1997. Badge size in collared flycatchers predicts outcome of male competition over territories. – Anim. Behav. 54: 893–899.

Pedersen, M. C., Dunn, P. O. and Whittingham, L. A. 2006. Extra-territorial forays are related to a male ornamental trait in the common yellowthroat. – Anim. Behav. 72: 479–486.

Perreault, S., Lemon, R. E. and Kuhnlein, U. 1996. Patterns and correlates of extrapair paternity in American redstarts (*Setophaga ruticilla*). – Behav. Ecol. 8: 612–621.

Potti, J. 1998. Arrival time from spring migration in male pied flycatchers: individual consistency and familial resemblance. – Condor 100: 702–708.

Qvarnström, A., Pärt, T. and Sheldon, B. C. 2000. Adaptive plasticity in mate preference linked to differences in reproductive effort. – Nature 405: 344–347.

Richardson, D. S. and Burke, T. 1999. Extra-pair paternity in relation to male age in Bullock's orioles. – Mol. Ecol. 8: 2115–2126.

Schlicht, L., Valcu, M. and Kempenaers, B. 2015. Spatial patterns of extra-pair paternity: beyond paternity gains and losses. – J. Anim. Ecol. 84: 518–531.

Schnitzer, J., Exnerová, A., Poláková, R., Vinkler, M., Tomášek, O., Munclinger, P. and Albrecht, T. 2013. Male ornamentation and within-pair paternity are not associated with male provisioning rates in scarlet rosefinches *Carpodacus erythrinus*. – Acta Ethol. 17: 89–97.

Sheldon, B. C. 1994. Male phenotype, fertility, and the pursuit of extra-pair copulations by female birds. – Proc. R. Soc. B 257: 25–30.

Sheldon, B. C. and Ellegren, H. 1999. Sexual selection resulting from extrapair paternity in collared flycatchers. – Anim. Behav. 57: 285–298.

Sheldon, B. C., Merilä, J., Qvarnström, A., Gustafsson, L. and Ellegren, H. 1997. Paternal genetic contribution to offspring condition predicted by size of male secondary sexual character. – Proc. R. Soc. B 264: 297–302.

Smith, S. M. 1988. Extra-pair copulations in black-capped chickadees: the role of the female. – Behaviour 107: 15–23.

Stutchbury, B., Pitcher, T. E., Norris, D. R., Tuttle, E. M. and Gonser, R. A. 2005. Does male extra-territory foray effort affect fertilization success in hooded warblers *Wilsonia citrina*? – J. Avian Biol. 36: 471–477.

Sundberg, J. and Dixon, A. 1996. Old, colourful male yellowhammers, *Emberiza citrinella*, benefit from extra-pair copulations. – Anim. Behav. 52: 113–122.

Tarof, S. A. and Ratcliffe, L. M. 2000. Pair formation and copulation behaviour in least flycatcher clusters. – Condor 102: 832–837.

Tregenza, T. and Wedell, N. 2000. Genetic compatibility, mate choice and patterns of parentage: invited review. – Mol. Ecol. 9: 1013–1027.

Valera, F., Hoi, H. and Kristin, A. 2003. Male shrikes punish unfaithful females. – Behav. Ecol. 14: 403–408.

Ward, M. P., Sperry, J. H. and Weatherhead, P. J. 2013. Evaluation of automated radio telemetry for quantifying movements and home ranges of snakes. – J. Herpetol. 47: 337–345.

Ward, M. P., Alessi, M., Benson, T. J. and Chiavacci, S. J. 2014. The active nightlife of diurnal birds: extraterritorial forays and nocturnal activity patterns. – Anim. Behav. 88: 175–184.

Westneat, D. F. and Stewart, I. R. K. 2003. Extrapair paternity in birds: causes, correlates, and conflict. – Annu. Rev. Ecol. Evol. Syst. 34: 365–396.

Woolfenden, B. E., Stutchbury, B. J. M. and Morton, E. S. 2005. Male acadian flycatchers, *Empidonax virescens*, obtain extrapair fertilizations with distant females. – Anim. Behav. 69: 921–929.

Článek 16: Edme, A., Zobač, P., Opatová, P., Šplíchalová, P., Munclinger, P., Albrecht, T. & Krist, M. (2017) Do ornaments, arrival date, and sperm size influence mating and paternity success in the collared flycatcher? *Behavioral Ecology and Sociobiology*, DOI 10.1007/s00265-016-2242-8.

CrossMark

ORIGINAL ARTICLE

Do ornaments, arrival date, and sperm size influence mating and paternity success in the collared flycatcher?

Anaïs Edme[1] · Petr Zobač[1] · Pavlína Opatová[1,2,3] · Petra Šplíchalová[4] ·
Pavel Munclinger[4] · Tomáš Albrecht[2,4] · Miloš Krist[1,5]

Abstract

Males advertise their intrinsic parental and/or genetic qualities by the size of secondary sexual ornaments. Moreover, they compete with one another for the best territory and males who arrive first at the breeding ground usually have an advantage in this competition. Females may consider multiple male qualities simultaneously and prefer the one most important for their fitness in the current context. They can further improve their fitness by selecting the best care-giver as their social mate and engaging in an extra-pair copulation with a genetically superior male. In such cases, sperm competition arises in the female reproductive tract and its outcome may be affected by the sperm morphology of both the social and extra-pair male. Here, we tested these ideas in the collared flycatcher (*Ficedula albicollis*), a species with context-dependent choice of social partners and frequent extra-pair paternity. We recorded male arrival to breeding sites, manipulated their forehead patches, and measured sperm size. In contrast to a previous study in a Swedish population, males with enlarged patches were non-significantly less successful late in the season while no such difference was found early in the season. Besides this tendential seasonal interaction, arrival date did not affect mating and paternity success or male fitness, and the same was true for sperm size. These results suggest different benefits of male ornamentation and female mate choice between populations and call for more replicated research within and between species.

✉ Miloš Krist
   milos.krist@upol.cz

1   Department of Zoology and Laboratory of Ornithology, Faculty of
    Science, Palacký University, 17 listopadu 50, CZ-771
    46 Olomouc, Czech Republic

2   Institute of Vertebrate Biology, Czech Academy of Sciences, v.v.i.,
    Květná 8, CZ-67502 Brno, Czech Republic

3   Department of Botany and Zoology, Faculty of Science, Masaryk
    University, Kotlářská 267/2, CZ-61137 Brno, Czech Republic

4   Department of Zoology, Faculty of Science, Charles University in
    Prague, Viničná 7, CZ-128 44 Prague, Czech Republic

5   Museum of Natural History, nám. Republiky 5, CZ-771
    73 Olomouc, Czech Republic

Significance statement

The fitness of a male of a migratory species might be affected by several pathways. First, early arrival should confer benefits in the form of best territory choice. Second, in a dichromatic and sexually promiscuous species, secondary sexual ornaments are considered by females both in the choice of social and extra-pair mates. Third, by modifying sperm traits, males may outmatch their rivals in sperm competition. These ideas have usually been tested in isolation. In this experimental study, we tested the joint effect of all of these factors on the genetic fitness of males. We found little evidence for the dependence of male reproductive success on either sperm morphology or plumage ornamentation which is in contrast to other populations of the species. Our study calls for replicated research both in well-established fields like mate choice and emerging ones like sperm competition.

 Springer

Introduction

Many studies have shown that females have preferences for males with more elaborate secondary sexual traits such as more diverse songs (Gentner and Hulse 2000; Drăgănoiu et al. 2002) and larger sexual ornaments (Sheldon et al. 1997). Females may have preferences for those males with superior traits because males will provide either direct benefits such as parental care or indirect genetic benefits to the offspring (Andersson 1994).

However, female mating preferences may vary according to the context (Qvarnström 2001). Indeed, females have displayed differences in preference according to social context (such as presence or absence of competitors) (Callander et al. 2012), environmental conditions (Hale 2008), and timing of breeding (Qvarnström et al. 2000). This last factor can be particularly important for migratory birds that are constrained by their migration schedule. It is common that males arrive earlier than females (Møller 2004; Tottrup and Thorup 2008). Early arrival allows males to settle on the best territories (Aebischer et al. 1996) and also to obtain females more easily as females may also use territories as cues for mate selection (Alatalo et al. 1986). Among males arriving at the same time, those with brighter and/or larger plumage ornaments usually win competitions for territories (Pärt and Qvarnström 1997; Beck 2013). Therefore, females can choose these highly ornamented males to have access to necessary resources like nest sites or food. However, choosing dominant mates at the start of the season may also be costly as such males often try to attract a secondary or extra-pair female instead of caring for the primary nest (Qvarnström 1997, 1999). Consequently, it may be advantageous to only choose males with larger ornaments later in the season as their chances to find another mate are low at that time, and thus, they are expected to invest more in the feeding of nestlings (Qvarnström et al. 2000).

In addition to the choice of social mate, females might use several other mechanisms to increase their fitness. First, they may be unfaithful to their social mate. Extra-pair young obtained with another male of superior quality may be of higher quality compared to within-pair young (Akçay and Roughgarden 2007, but see Krist and Munclinger 2011). Many studies show that females cuckold their mates with older males (Cleasby and Nakagawa 2012) and more ornamented males (Kempenaers et al. 1992; Richardson and Burke 1999; Akçay and Roughgarden 2007), though the role of ornaments remains controversial (review in Akçay and Roughgarden 2007). In contrast to female preferences for social mates, which may be context-dependent due to the trade-off between the direct and indirect benefits of mate choice (Qvarnström 2001), preferences for extra-pair males with large ornaments may be consistent during the course of the season as these males can provide only genetic benefits, and thus, there is no trade-off with their paternal care even at the start of the season.

Extra-pair copulations are a prerequisite for another process that has come to the center of attention of ecologists in recent years. Whenever females copulate with more than one male, different ejaculates compete to fertilize the eggs, which is known as sperm competition. Many factors may influence the success of sperm in fertilizing eggs: the timing of copulation (Birkhead et al. 1989), the frequency of copulation (Møller and Birkhead 1993; Mougeot 2004), and sperm traits (Snook 2005). Among these sperm traits, viability (Smith 2012), speed of swimming, (Birkhead et al. 1999), number (Laskemoen et al. 2010), and size of the sperm (Lifjeld et al. 2010; Bennison et al. 2015) may modulate the success of egg fertilization.

Although it has previously been shown that male arrival date (Aebischer et al. 1996), secondary sexual ornaments (Sheldon and Ellegren 1999), and sperm size (Bennison et al. 2015) can have fitness effects, these factors were usually tested in isolation which complicates the evaluation of their relative importance. One remarkable exception is the study of Qvarnström et al. (2000) that tested how benefits of female choice of male ornaments depend on the time of male arrival to the breeding ground. However, this study did not take sperm competition pathways of sexual selection into account. Here, we tested the effects of male ornamentation, arrival time, and sperm morphology on their ability to sire offspring and gain fitness.

We studied these questions in the collared flycatcher (*Ficedula albicollis*), a migratory bird in which males arrive on the breeding grounds before females. Males of this species display two white patches, one on the forehead and the other on the wing, that have been found to be sexually selected in Swedish population (e.g., Sheldon and Ellegren 1999; de Heij et al. 2011). However, there may be differences in the strength of sexual selection between populations. For example, large forehead patch has been found to be preferred in extra-pair mates in the Swedish (Sheldon et al. 1997; Sheldon and Ellegren 1999) but not Hungarian (Rosivall et al. 2009) or Czech (Edme et al. 2016) populations. This calls for replicative research both within and between populations to test if the differences between studies really represent differences in the strength of selection between populations, which would have important consequences for the evolution of the species (see Scordato and Safran 2014) or if they are merely caused by sampling variance.

Methods

Study site and species

This study was carried out in an oak forest with approximately 350 nest boxes that are distributed among five study plots in Velky Kosir (49°32′N, 17°04′E) in Moravia, Czech Republic,

from 2013 to 2015. Collared flycatchers (*F. albicollis*) are migratory passerine birds, and males arrive first (around mid-April) at the breeding site to obtain territories. Males are black with one white patch on the forehead and another on their wing feathers. Females selected their social mates based on both white ornaments in a Swedish population (Sheldon and Ellegren 1999; de Heij et al. 2011) and usually lay between four and eight eggs after pair bond formation. Chicks can hatch asynchronously as females start to incubate before the completion of the clutch. Both parents feed the chicks.

Adult measurements and forehead patch manipulation

In 2013, the first male arrived on April 15, and we started to trap males the following day. In total, we trapped males on 12 different days between April 16 and May 15. Each trapping day, we captured males in all empty or abandoned nest boxes with string nest box traps. We did not activate traps in nest boxes where nest material appeared unless these were apparently abandoned for several days (i.e., no progress in nest building). For individual males, we considered the first day of capture as their date of arrival. Our trapping scheme was highly efficient as the first day of capture was highly correlated ($r = 0.96$) with true arrival date as inferred from 16 males bearing geolocators in 2015 (M. Briedis et al., unpublished data). Immediately after each new male was captured, it was brought to the central site located among study plots. This transfer lasted up to half an hour.

At the central site, body mass, wing length, and tarsal length were measured. The wing patches were determined by summing the visible length of white patches on primaries 3 to 8 from the tip of the coverts to the distal part of the wing (in mm). All of these measurements were done by one person (MK). A blood sample was taken from the tarsal vein and stored in alcohol. A cloacal massage allowed us to obtain a sperm sample (see Quay 1986), which was stored in 4% formaldehyde. The age of males was determined by wing plumage as subadult males have brownish primaries. The forehead patch area was photographed two times before and another two times after the manipulation. The original patch size was computed as the mean of the two measurements before the manipulation delimited to the nearest 0.1 mm^2 in ImageJ software.

We regularly rotated among three treatments: (1) we increased the height of white forehead patches by painting black feathers with a white marker (Alteco Paint Marker no.15). Using this technique, the size of the white patch was enlarged by ca. 50% (Table 1, Supplementary Online Material). We decided to use Alteco markers instead of Tippex used in former studies (e.g., Qvarnström et al. 2000) since they proved to be more durable during our pre-experimental manipulation done on caged zebra finches. Tippex usually started to erode within a week of manipulation, while Alteco still looked good

after 7 days. Both Alteco and Tippex have similarly shaped reflectance curves that differ from natural white feathers. At low and high wavelengths, natural white reflects more than Tippex and especially Alteco (see Fig. 1). (2) Control birds were only measured and then released without any manipulation of the forehead patch. (3) We decreased the height of the white forehead patch to about half (Table 1, Supplementary Online Material) by painting it with a Copic 110 special black marker that has previously been used in flycatchers (de Heij et al. 2011). This manipulation resulted in naturally low reflectance (Fig. 1) but started to fade within a few days of manipulation. Our rotation scheme led to a random distribution of treatments among plots as indicated by a non-significant relationship between plot and frequency of treatment ($\chi^2 = 9.23$, $p = 0.324$, df = 8, $n = 73$).

Because our manipulations were relatively short-term, they could mainly affect processes at the start of the breeding season like female choice of social partners, which usually takes place during the days after arrival to breeding sites (8 days for control males on average; see Table 1, Fig. 2). However, they might be less effective for female choice of extra-pair partners which might continue for a long time after males are socially mated, although most extra-pair copulations likely take place early in the female fertile period (Krist et al. 2005; Krist and Munclinger 2011) which peaks 2 days before laying of the first egg (Lifjeld et al. 1997). In the nests attended by our control males, laying started 6 days after social mating, i.e., 14 days after male arrival.

After manipulation, males were released on the same plot as they were caught. We caught the males a second time during the feeding period, and the same measurements were taken as well as blood and sperm samples. Females were also caught during the feeding period and were measured in the same way as males except for the forehead patch.

Monitoring of reproductive success

Nests were checked daily when the first egg was expected after nest building. Each egg was marked to obtain the laying order. The width and length were measured with digital calipers (±0.01 mm). The volume of the egg was calculated as volume = $0.51 \times$ length $\times$ width^2 (Hoyt 1979). When females ended the laying sequence and began incubation, we stopped the daily checks and started once again when the hatchlings were expected (around 10 days after the last egg was laid). A blood sample was obtained from chicks 6 days after hatching, and their fate was monitored until fledging. Unhatched eggs were collected 4 days after the last chicks hatched, and embryos were stored in ethanol, as were all of the other chicks found dead before day 6. Blood and tissue samples were used for paternity analyses.

In 2014 and 2015, we captured all of the males at arrival and both sexes during the breeding season, so we were able to

Table 1 Summary of means ± SD for different traits according to patch treatment

|	Patch treatment		
	Decreased	Control	Enlarged
Number of males manipulated	51	50	52
Number of males breeding	27	25	21
Polygynous males	1	3	1
Number of males used in most analyses	24	23	20
Arrival date (Julian day, January 1 = 1)	112.3 ± 5.7	112.7 ± 7.0	113.8 ± 5.4
Mating speed (days)	10.1 ± 7.29	8.0 ± 4.01	11.4 ± 8.0
Forehead patch before treatment (mm^2)	58.7 ± 7.4	56.1 ± 11.7	57.1 ± 8.6
Forehead patch after treatment (mm^2)	29.4 ± 5.8	56.1 ± 11.7	85.9 ± 12.8
Wing patch (mm)	50.8 ± 15.5	58.9 ± 12.6	49.2 ± 11.1
Clutch size	6.16 ± 0.81	6.21 ± 0.85	5.75 ± 1.16
Egg volume (mm^3)	1658 ± 131	1603 ± 135	1604 ± 121
Number of WPY	4.65 ± 1.52	4.81 ± 2.10	4.29 ± 2.22
Number of EPY	1.43 ± 1.61	1.59 ± 2.21	1.47 ± 1.69
Total paternity	6.08 ± 2.48	6.40 ± 3.48	5.76 ± 2.30
Number of fledglings	5.08 ± 2.24	5.47 ± 2.76	4.05 ± 2.72
Number of recruits	1.00 ± 1.17	1.65 ± 1.15	1.05 ± 1.39

Sample size was 67 males except for number of WPY, number of EPY, and total paternity, which were based on $n = 62$. For polygynous males, only their primary nests were considered for the calculation of means. The exceptions were the variables "number of extra-pair young" and "total paternity" that also included young sired by the polygynous males in their secondary nests

count the number of recruits as all the chicks were ringed during the field season in 2013. We did not record whether those recruits bred during those 2 years, but only their survival since fledging. So our recruitment data concerned the number of chicks who survived and were able to come back to our field area. Despite natal fidelity being relatively high in our

study area (Krist 2009), some individuals surely dispersed and thus our estimate of recruitment represents only the lower limit of the real value.

Genotyping and parentage assignment

DNA extraction was performed with DNeasy® Blood & Tissue Kit (Qiagen) for blood samples and tissue from dead
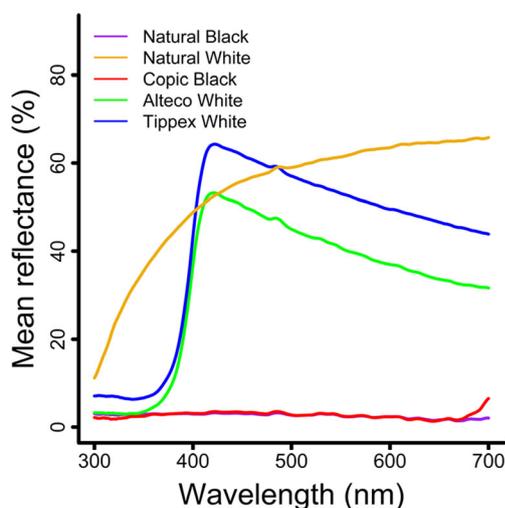
Fig. 1 Reflectance of primaries of adult males before and after coloration with black or white markers. Five measurements were taken from the feathers of two males, and the lines are averages of these five measurements. The reflectance of primaries of adult males likely closely reflects that of their foreheads but the former was easier to measure on dead birds that were available before breeding season. These dead birds were killed by great tits that destroyed their foreheads
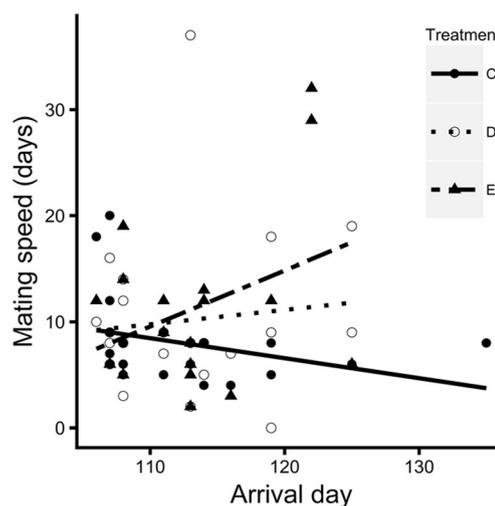
Fig. 2 Relationship between arrival date and mating speed for the three treatments of forehead patch size. Control treatment, *solid circles* and *solid line*; decreased treatment, *open circles* and *dotted line*; enlarged treatment, *triangles* and *dashed line*

embryos and chicks. All of the samples were genotyped at eight polymorphic microsatellite autosomal loci: Fhu2 (or PTC3) (Ellegren 1992), Cuμ04 (Gibbs et al. 1999), Fhy310, Fhy405, Fhy407, Fhy428, Fhy431, and Fhy452 (Leder et al. 2008). A single multiplex PCR using fluorescently labeled primers and a Type-it® Microsatellite PCR Kit (Qiagen) were used to amplify the microsatellites. The samples were treated with the following reaction conditions: 5 min at 95 °C, then in 30 cycles of 30 s at 95 °C, 90 s at 65 °C, 30 s at 72 °C, and finally 30 min at 60 °C. PCR products were mixed with GeneScan™-500 LIZ® Size Standard (Applied Biosystems) and analyzed with ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems). GeneMarker® version 1.9 was used to score the genotypes, and locus characteristics based on allele frequencies were obtained with Cervus 3.0.3 (Kalinowski et al. 2007).

We obtained the genotypes of 262 adults (104 females and 158 males). For the first parent, the combined non-exclusion probability for that group of loci was found to be 7.03×10^{-4}. We only considered the individuals that were genotyped at five loci or more for parental analysis. When female genotype was known, we compared it with its chick genotypes to check for egg dumping. One chick did not correspond to its social mother and was excluded. Secondly, when the social male was known, we compared the genotype of the male with the chicks he fed. If trio confidence (female-social male-chicks) based on Delta (difference in overall likelihood ratio scores between the most likely candidate parent and the second most likely candidate parent) and simulations of parentage was superior to 95%, we considered the chicks to be within-pair young. In cases where the mother was unknown, we took into account the duo confidence (male-chicks) with the same criterion. All chicks that were not assigned as within-pair young were classified as extra-pair young. Finally, we tried to determine the males who sired the extra-pair young. We selected all the males from the breeding season and compared their genotypes with the extra-pair chicks using the same criterion of 95% trio or duo confidence.

Sperm analyses

Two hundred and forty-two sperm samples were stored in 4% formaldehyde (152 from males at arrival and 90 during the feeding period) either at room temperature or at 8 °C in a refrigerator. We created slides for microscopy by spreading 7 μl of a sperm sample and letting it dry. The slide was then carefully rinsed with distilled water in order to remove dirt and salt crusts and air-dried again. For each sample, 20 pictures of morphologically normal-looking sperm were taken at ×400 magnification under light-field conditions using an Olympus CX41 microscope equipped with an Infinity 2 camera. If 20 sperms were not found on the first slide, a second slide was prepared. If after those two slides, no sperm at all was found,

we did not prepare a third slide. For samples where the number of sperm was between 1 and 19 sperms after two slides, an ultimate slide was analyzed to complete the number of sperm pictures. We obtained 130 samples with the required number of sperm at arrival and 39 at feeding. Heads, mid-pieces and tails were measured (μm) in ImageJ software 1.49v (see Laskemoen et al. 2010). All of these measurements were done blindly by one person (PZ). Total sperm length was calculated by adding the three parts. Mean sperm length was calculated for each male.

Statistical analyses

All statistical analyses were conducted in RStudio, version 0.99.878 (R Core Team 2014), and we used the "lm" or "glm" function from the package "Stats" (R Core Team 2014). Since males were trapped at arrival ($n = 153$) and recaptured during the feeding period ($n = 73$), it was possible to identify those who were successful at pairing and establishing a nest in that particular season. To test this, we fitted a generalized linear model with a binomial link function (glm function from Stats package in R). The response variable was mating success (obtaining a nest: yes/no), and the predictors were the arrival date, original forehead patch size, wing patch size, the relative age (adult/subadult), and treatment (enlarged, decreased, and control). We also tested the interaction between arrival date and treatment, as the effect of ornament manipulation was found to be dependent on the time of the season in a previous study (Qvarnström et al. 2000). When this interaction was non-significant, it was removed from the final model. Continuous predictors in our models (i.e., male arrival and size of original forehead patch, wing patch, and sperm length) were not strongly intercorrelated (all r between -0.4 and $+0.4$, $n = 63$), indicating that multicollinearity was not a serious problem.

Another factor that we were interested in was mating speed. We calculated this as the time between male arrival date and the start of nest building by its social female. Six out of 73 males presented a negative value for the time lapse between those two dates, indicating that we trapped them well after their arrival. These males were trapped during the searching of secondary nest sites after they had started to breed in their primary nest box. We excluded them from all analyses. Five out of 73 breeding males were polygynous, and their secondary nests were not considered in analyses of mating speed, clutch size, and egg size. So in total, 67 manipulated and breeding males were used for most of the analyses. The mating speed ranged from 0 to 37 days (see also Table 1). A linear model was run, where the response variable was mating speed and the predictors were the same as in the model for mating success.

We also tested whether females changed their early reproductive effort in respect to male secondary sexual traits, as is

predicted by the theory of reproductive allocation (Sheldon 2000; Horváthová et al. 2012). In the first model, we looked at the number of eggs the female laid. A linear model was run on the clutch size, as it had a better fit than the alternate Poisson model, and residuals from the linear model were normally distributed. The response variable was clutch size, and the predictors were the same as in the model for mating success. Second, we looked at the volume of the eggs; a linear model was run with mean egg volume as the response variable and the same predictors as in the model for mating success.

We added sperm length among predictors of the models testing for paternity success. We used the male sperm length measured at arrival. For five males, we obtained sperm only for the feeding period. As we had the mean size at arrival and feeding for 28 males, we calculated the difference between the mean sperm size at arrival (mean ± SD; 96.7 ± 3.20 μm) and the mean size at feeding (98.2 ± 2.30 μm) and subtracted this difference from the size at feeding for those five males without arrival data. In this way, we extrapolated the size of the sperm at arrival for those males. The results would be very similar if these males were excluded from the analyses. For another five males, we did not obtain enough sperm either at arrival or during breeding, and therefore, we excluded them from this analysis that was consequently based on 62 males.

The total paternity success of a male can be separated into two parts: the within-pair paternity in the social nest and the extra-pair paternity obtained in other nests. We first looked at the within-pair paternity with a generalized linear model with a quasi-binomial distribution and event/trial syntax for the response variable. In consequence, the response variable was the number of within-pair young (event) according to the clutch size (trial). In addition to predictors used in the model for mating success, we added mean sperm length and its quadratic term to all three models testing for paternity success. We added the quadratic term into models to test for the possibility of stabilizing selection on sperm size (Lifjeld et al. 2010). The extra-pair paternity was analyzed by a generalized linear model with a quasi-Poisson distribution. The response was the number of extra-pair young that males sired in all other nests in the nest box population (n = 119 nests with genotyped offspring). The predictors were the same as in the preceding model. As in all other models except of that for mating success, we tested only success of males breeding in our nest boxes. For five identified polygynous males, we included the number of young they sired in their secondary nests to their extra-pair success. This was done to be equivalent to cases where polygynous males were not identified at all as they did not feed their secondary nests. By this method, extra-pair success was overestimated while within-pair success was underestimated for polygynous males. Nevertheless, the results would be closely similar if five identified polygynous males were excluded from this model (results not shown). Moreover, this slight inadequacy did not affect the model of male total paternity because in this model the two paternity components were summed together. The model for total paternity was the same as for extra-pair paternity except for the response variable that was the total paternity. Our estimates of male extra-pair and therefore also total paternity success only reflect the lower limits of the real values since focal males might also sire offspring in natural cavities, i.e., outside our genotyped nest box population.

The number of fledglings and recruits is reflective of male fitness, so we ran two other models with the number of male genetic offspring that fledged as a response in the first model and the number of genetic offspring that were recruited (in 2014–2015) in the second model. For both models, the predictors were the same as in the model for mating success.

Results

During the arrival period, 160 males were trapped and 153 were involved in the patch manipulation experiment (52 increased, 51 decreased, and 50 for control). Seventy-three of them were recaptured when they were feeding chicks. Five of them were polygynous. We excluded secondary nests of polygynous males from analyses of mating speed, clutch size, and egg size. We also excluded six males that were caught a long time after their arrival (see "Methods" section). Consequently, our sample size for most analyses was 67 breeding males. In all models testing for paternity success, our sample size was reduced to 62 males due to missing sperm samples from 5 males. In these 62 nests, 286 within-pair young were sired by social and 67 by extra-pair mates. These 62 social males also sired 93 offspring outside their primary nests.

The males involved in our treatment arrived on average on 112.9 ± 6.0 (mean ± SD) Julian day (April 23) and required about 9.7 ± 6.6 days to pair (see Table 1 for more details). Females laid on average 6.06 ± 0.95 (mean ± SD) eggs, and the mean volume of the eggs was 1623 ± 130 mm³ (Table 1), with an average of 4.91 ± 2.60 chicks fledging from each nest (Table 1). We recaptured 83 of the nestlings in 2014 and 2015. The mean ± SD number of recruits per nest was 1.24 ± 1.26 (Table 1).

None of our main variables (arrival date, original size of male ornaments, and their experimental treatments) significantly affected male mating success (Table 2), although males with enlarged patches (21/52 = 40.4%) had a non-significantly lower mating success compared to the control group (25/50 = 50%) and males with reduced patches (27/51 = 52.9%). Similarly, males in the enlarged treatment had non-significantly lower mating speed than males in the other two treatments (Tables 1 and 2), and this seemed to be true mainly late in the season (Fig. 2), although the interaction between treatment and arrival date was marginally non-significant

Table 2 Models for mating success ($N = 153$) and the speed of mating ($N = 66$)

	Mating success			Mating speed		
	Estimate ± SE	F	P	Estimate ± SE	F	P
Intercept	−1.83 ± 3.52			−5.37 ± 19.5	0.08	0.784
Arrival date	0.001 ± 0.026	<0.01	0.954	0.079 ± 0.147	0.29	0.598
Forehead patch size	−0.0002 ± 0.0173	<0.01	0.990	−0.026 ± 0.098	0.07	0.791
Wing patch size	0.029 ± 0.202	2.04	0.155	0.094 ± 0.111	0.72	0.400
Age	0.906 ± 0.858	1.07	0.303	4.85 ± 4.64	1.10	0.299
Treatment		0.82	0.445		1.61	0.207
Treatment decreased	0.169 ± 0.410			2.38 ± 2.08		
Treatment enlarged	−0.344 ± 0.404			4.07 ± 2.28		

($p = 0.10$). We did not find any evidence for female pre-hatching differential allocation since neither egg size nor clutch size differed between treatments (Tables 1 and 3). Male success in sperm competition was not affected by their arrival date, size of original forehead patch, experimental treatment, or sperm size (Table 4, Fig. 3). Finally, we also did not find a significant effect of any predictor on male fitness as determined by the number of fledglings and recruits, although males in the enlarged treatment had somewhat poorer performance compared to those in reduced and especially control treatments (Tables 1 and 5).

Discussion

We found several lines of evidence suggesting that males in the enlarged treatment of forehead patch size might have inferior breeding performance compared to males in control and reduced treatments. They had lower mating success, it took them longer to pair, especially late in the season, and their fitness as measured by clutch size and number of fledglings and recruits was also lower than in the other two treatment groups. However, although these effects were visible in the difference between means (Table 1), they were also highly variable, which caused them to be statistically non-significant,

despite the fact that we involved the whole nest box population in our experiment and thus had a sample size comparable to many previous studies.

The manipulation of male attractiveness is a common type of experiment when studying mate choice, female investment, and paternity (Mazuc et al. 2003; Grana et al. 2012; Horváthová et al. 2012). Manipulations of ornaments in the collared flycatchers were previously done in the isolated Swedish population on the island of Gotland (Qvarnström 1999; Qvarnström et al. 2000; de Heij et al. 2011). Here, we partly replicated the forehead patch size manipulation from the Qvarnström et al. (2000) study in a Czech population of collared flycatchers. Qvarnström et al. (2000) found that females only preferred males with enlarged patches late in the season while having no strong preferences early in the season. We did not find a statistically significant interaction between pairing latency and experimental treatment of the forehead patch. If anything, there was an opposite tendency. Females in our population did not show any preferences early in the season but tended to prefer control males and males with decreased patch sizes over enlarged ones later in the season. There are several potential explanations for these different results.

First, it may be due to the type of white markers used in the experiment. We used a white paint marker (Alteco) while

Table 3 Models for clutch size ($N = 67$) and egg volume ($N = 66$)

|	Clutch size			Egg volume		
	Estimate ± SE	F	P	Estimate ± SE	F	P
Intercept	16.6 ± 2.31	51.8	<0.001	917 ± 356	6.63	0.012
Arrival date	−0.091 ± 0.017	27.4	<0.001	6.67 ± 2.68	6.20	0.015
Forehead patch size	−0.001 ± 0.011	0.03	0.871	−3.40 ± 1.82	3.50	0.066
Wing patch size	−0.0007 ± 0.013	<0.01	0.957	2.03 ± 2.07	0.96	0.331
Age	0.428 ± 0.552	0.60	0.440	71.6 ± 85.3	0.70	0.404
Treatment		1.11	0.335		2.35	0.104
Treatment decreased	−0.133 ± 0.247			76.5 ± 38.6		
Treatment enlarged	−0.391 ± 0.267			15.3 ± 41.8		

Table 4 Models for within-pair, extra-pair, and total paternity (N = 62)

	Within-pair paternity			Extra-pair paternity			Total paternity		
	Estimate ± SE	F	P	Estimate ± SE	F	P	Estimate ± SE	F	P
Intercept	−1.67 ± 1.48			0.870 ± 3.55			1.62 ± 1.38		
Arrival date	0.008 ± 0.011	0.60	0.441	−0.0003 ± 0.026	1.05	0.309	0.001 ± 0.010	0.03	0.870
Forehead patch size	0.001 ± 0.007	0.05	0.832	−0.016 ± 0.019	0.14	0.707	−0.001 ± 0.007	0.05	0.823
Wing patch size	0.006 ± 0.008	0.68	0.413	−0.005 ± 0.021	0.48	0.491	0.010 ± 0.008	1.41	0.240
Age	0.199 ± 0.334	0.35	0.554	−17.8 ± 1924	3.76	0.057	0.076 ± 0.358	0.05	0.978
Treatment		0.02	0.981		0.04	0.965		0.01	0.989
Treatment decreased	−0.004 ± 0.147			0.407 ± 0.417			−0.020 ± 0.153		
Treatment enlarged	0.024 ± 0.164			0.417 ± 0.449			−0.020 ± 0.172		
Sperm size		0.30	0.741		0.18	0.834		0.54	0.587
Sperm size: linear	−0.079 ± 0.512			−0.415 ± 1.35			0.045 ± 0.533		
Sperm size: quadratic	−0.393 ± 0.523			0.789 ± 1.25			−0.557 ± 0.547		

Qvarnström et al. (2000) used Tippex. However, this difference is unlikely to explain the opposite direction of our results as the shape of the reflectance curves of the two markers is very similar. In contrast, the shape of the reflectance curve of natural white is different from both artificial colorations (see Fig. 1). Consequently, it is possible that females can distinguish between natural and artificial white and consider only the natural one as attractive while the artificial one may be unattractive. If true, the different results could partially stem from a difference in the treatment of control groups. In our study, we did not color the control group at all, contrary to Qvarnström et al. (2000) who painted Tippex over the natural white in the same extent as was used to paint the enlarged patch group over their natural black. Consequently, females in our study might have perceived enlarged patch males as less attractive because they had the same extent of natural (and attractive) white as control males but, in addition, they had patches of artificial white that made them unattractive. In contrast, in the study of Qvarnström et al. (2000), both the enlarged patch and control groups had the same extent of artificial white, but the experimental group retained a larger extent of natural white making them more attractive.

Second, Qvarnström et al. (2000) kept the males caged for 1 day to break their dominance over their original territories. We released males immediately after patch size manipulation and thus allowed them to return to their territory without a need to fight for them once more. If nest-site competition was intense only late in the season due to the lack of unoccupied territories, then the pairing latency of large-patched males could be shorter only at the end of the season, as was found by Qvarnström et al. (2000), due to their ability to win the competition over territory (see Pärt and Qvarnström 1997). In contrast, pairing latency in our population should not be as strongly affected by male-male competition and thus directly represent female mate choice.

Finally, and most interestingly, differences in the role of ornaments in sexual selection may exist between populations (Scordato and Safran 2014). For example, it has been shown that forehead patch size is condition dependent (Gustafsson et al. 1995) and males with large forehead patches are preferred as social (Qvarnström et al. 2000) and extra-pair (Sheldon et al. 1997) partners in an isolated Swedish population. In contrast, wing patch size (Török et al. 2003) but not forehead patch size (Hegyi et al. 2002) is a condition-dependent signal important in male-male competition (Garamszegi et al. 2006) in a Hungarian population. Similarly to the Hungarian population, wing but not forehead patches played a role in extra-pair paternity in our Czech population (Edme et al. 2016). These similarities suggest a greater role of wing patches in
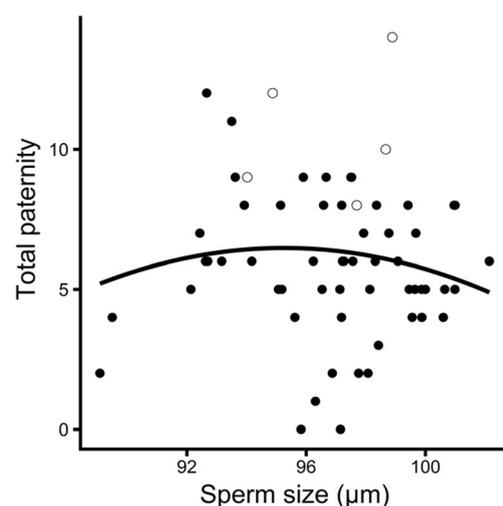
Fig. 3 Relationship between sperm size and total number of sired offspring (total paternity). *Solid circles* depict males that had only one social nest (n = 57). *Open circles* depict polygynous males (n = 5). *Fitted line* shows predicted quadratic regression

Table 5 Models for number of fledglings and recruits ($N = 67$)

	Fledgling			Recruit		
	Estimate ± SE	F	P	Estimate ± SE	F	P
Intercept	2.05 ± 1.54			1.73 ± 3.04		
Arrival date	−0.0004 ± 0.011	<0.01	0.966	−0.006 ± 0.023	0.09	0.766
Forehead patch size	−0.006 ± 0.007	0.60	0.439	−0.008 ± 0.014	0.35	0.558
Wing patch size	−0.0009 ± 0.008	0.01	0.917	−0.0007 ± 0.0176	<0.01	0.964
Age	−0.356 ± 0.393	0.83	0.365	−0.418 ± 0.815	0.27	0.605
Treatment		1.42	0.250		1.05	0.356
Treatment decreased	−0.011 ± 0.160			−0.426 ± 0.302		
Treatment enlarged	−0.269 ± 0.183			−0.407 ± 0.352		

Central Europe, the core of the distribution of the collared flycatcher. Nevertheless, females apparently paid attention to male foreheads in our population too, as they were less willing to mate with males with enlarged patches, and this was true especially late in the season. This change of mate preference with the season suggests an underlying change in costs and benefits of mating with large-patched males (Qvarnström 2001).

One explanation for plastic mate preferences may be the greater dependence of chicks on male paternal care late in the season. Consequently, females may be reluctant to pair with males that will not provide enough parental care during this difficult period of the breeding season. Highly ornamented males may invest resources into mating effort and provide less paternal care (Qvarnström 1997; Mazuc et al. 2003; Mitchell et al. 2007). Moreover, the size of the forehead patch may be used by females as an indicator of paternal care as this patch has been shown to decrease in the year following experimental increase of brood size (Gustafsson et al. 1995). Females living in populations with very limited resources may prefer males with smaller secondary sexual ornaments throughout the year (Griffith et al. 1999).

On the other hand, avoiding dominant males may also mean a loss on the side of indirect benefits if these males are genetically superior over subordinates. Therefore, females socially mated to high-quality fathers may increase the genetic component of offspring fitness by extra-pair copulation with superior males (Jennions and Petrie 2000). Extra-pair paternity is common in the collared flycatcher and is often related to secondary sexual plumage traits (Sheldon and Ellegren 1999; de Heij et al. 2011; Edme et al. 2016) as is also common in other species (Jennions and Petrie 2000; Akçay and Roughgarden 2007). However, extra-pair paternity is not determined solely by behavioral interactions among females and social and extra-pair males but also by the ability of sperm to fertilize ova, a process known as sperm competition. This area of research has been studied only recently and has yielded mixed results. Some studies have found a

relationship between sperm traits and success in extra-pair paternity (Laskemoen et al. 2010; Bennison et al. 2015) while others have not supported this idea (Cramer et al. 2013).

Here, we found neither a linear nor a non-linear effect of sperm size on within-pair or extra-pair paternity. Thus, there was no evidence of either directional nor stabilizing selection on sperm size. Stabilizing selection for optimal sperm size is hypothesized to be linked to the intensity of sperm competition between species, with the strongest selection for optimal sperm phenotype in the most promiscuous species (Lifjeld et al. 2010). Sperm competition in our population is quite intense as roughly 20–25% of young are sired by extra-pair males (Krist et al. 2005; Krist and Munclinger 2011; this study). Therefore, at first sight, our results do not seem to support the hypothesis of Lifjeld et al. (2010). However, it is tremendously difficult to predict within-species effects from comparative studies. It could be that our population has already reached evolutionary equilibrium, when the sperm size of all males might be so close to the species' optimum that any subtle differences in sperm morphology play no role in their fertilizing abilities. Moreover, other sperm traits that we did not measure might be more relevant for success in sperm competition, for instance sperm viability (Smith 2012), speed of swimming (Birkhead et al. 1999), and number of sperm cells in the ejaculate (Laskemoen et al. 2010).

We partially replicated the study of Qvarnström et al. (2000) that manipulated forehead patch size in the collared flycatcher. Contrary to the Swedish population, we did not find any evidence for female preference of males with enlarged patches late in the season. Males with artificially enlarged patches seemed to be unattractive in the Czech population, and this was especially true late in the season. We also did not find any evidence that sperm size affects within-pair or extra-pair paternity and consequently male fitness. These findings call for replicated research both in well-established fields like female mate choice with respect to male ornaments and emerging ones like sperm variation and its effect on paternity and fitness.

References

Aebischer A, Perrin N, Krieg M, Studer J, Meyer DR (1996) The role of territory choice, mate choice and arrival date on breeding success in the Savi's Warbler *Locustella luscinioides*. J Avian Biol 27:143–152

Akçay E, Roughgarden J (2007) Extra-pair paternity in birds : review of the genetic benefits. Evol Ecol Res 9:855–868

Alatalo RV, Lundberg A, Glynn C (1986) Female pied flycatchers choose territory quality and not male characteristics. Nature 323:152–153

Andersson M (1994) Sexual selection. Princeton University Press, Princeton

Beck ML (2013) Nest-box acquisition is related to plumage coloration in male and female prothonotary warblers (*Protonotaria citrea*). Auk 130:364–371

Bennison C, Hemmings N, Slate J, Birkhead T (2015) Long sperm fertilize more eggs in a bird. Proc R Soc B 282:20141897

Birkhead TR, Hunter FM, Pellatt E (1989) Sperm competition in the zebra finch, *Taeniopygia guttata*. Anim Behav 38:935–950

Birkhead TR, Martinez JG, Burke T, Froman DP (1999) Sperm mobility determines the outcome of sperm competition in the domestic fowl. Proc R Soc Lond B 266:1759–1764

Callander S, Backwell PRY, Jennions MD (2012) Context-dependent male mate choice: the effects of competitor presence and competitor size. Behav Ecol 23:355–360

Cleasby IR, Nakagawa S (2012) The influence of male age on within-pair and extra-pair paternity in passerines. Ibis 154:318–324

Cramer ERA, Laskemoen T, Kleven O, Lebarbera K, Lovette IJ, Lifjeld JT (2013) No evidence that sperm morphology predicts paternity success in wild house wrens. Behav Ecol Sociobiol 67:1845–1853

de Heij ME, Gustafsson L, Brommer JE (2011) Experimental manipulation shows that the white wing patch in collared flycatchers is a male sexual ornament. Ecol Evol 1:546–555

Drăgănoiu TI, Nagle L, Kreutzer M (2002) Directional female preference for an exaggerated male trait in canary (*Serinus canaria*) song. Proc R Soc Lond B 269:2525–2531

Edme A, Munclinger P, Krist M (2016) Female collared flycatchers choose neighbouring and older extra-pair partners from the pool of males around their nests. J Avian Biol 47:552–562

Ellegren H (1992) Polymerase-chain-reaction (PCR) analysis of microsatellites—a new approach to studies of genetic relationships in birds. Auk 109:886–895

Garamszegi LZ, Rosivall B, Hegyi G, Szöllösi E, Török J, Eens M (2006) Determinants of male territorial behavior in a Hungarian collared flycatcher population: plumage traits of residents and challengers. Behav Ecol Sociobiol 60:663–671

Gentner TQ, Hulse SH (2000) Female European starling preference and choice for variation in conspecific male song. Anim Behav 59:443–458

Gibbs HL, Tabak LM, Hobson K (1999) Characterization of microsatellite DNA loci for a neotropical migrant songbird, the Swainson's thrush (*Catharus ustulatus*). Mol Ecol 8:1551–1561

Grana SC, Sakaluk SK, Bowden RM, Doellman MA, Vogel LA, Thompson CF (2012) Reproductive allocation in female house wrens is not influenced by experimentally altered male attractiveness. Behav Ecol Sociobiol 66:1247–1258

Griffith SC, Owens IPF, Burke T (1999) Female choice and annual reproductive success favour less-ornamented male house sparrows. Proc R Soc Lond B 266:765–770

Gustafsson L, Qvarnström A, Sheldon BC (1995) Trade-offs between life history traits and a secondary sexual character in male collared flycatchers. Nature 375:311–313

Hale RE (2008) Evidence that context-dependent mate choice for parental care mirrors benefits to offspring. Anim Behav 75:1283–1290

Hegyi G, Török J, Toth L (2002) Qualitative population divergence in proximate determination of a sexually selected trait in the collared flycatcher. J Evol Biol 15:710–719

Horváthová T, Nakagawa S, Uller T (2012) Strategic female reproductive investment in response to male attractiveness in birds. Proc R Soc Lond B 279:163–170

Hoyt DF (1979) Practical methods of estimating volume and fresh weight of birds' eggs. Auk 96:73–77

Jennions MD, Petrie M (2000) Why do females mate multiply? A review of the genetic benefits. Biol Rev 75:21–64

Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Mol Ecol 16:1099–1106

Kempenaers B, Verheyen G, Van den Broeck M, Burke T, Van Broeckhoven C, Dhondt A (1992) Extra-pair paternity results from female preference for high quality males in the blue tit. Nature 357: 494–496

Krist M (2009) Short- and long-term effects of egg size and feeding frequency on offspring quality in the collared flycatcher (*Ficedula albicollis*). J Anim Ecol 78:907–918

Krist M, Munclinger P (2011) Superiority of extra-pair offspring: maternal but not genetic effects as revealed by a mixed cross-fostering design. Mol Ecol 20:5074–5091

Krist M, Nádvorník P, Uvírová L, Bureš S (2005) Paternity covaries with laying and hatching order in the collared flycatcher *Ficedula albicollis*. Behav Ecol Sociobiol 59:6–11

Laskemoen T, Kleven O, Fossøy F, Robertson RJ, Rudolfsen G, Lifjeld JT (2010) Sperm quantity and quality effects on fertilization success in a highly promiscuous passerine, the tree swallow *Tachycineta bicolor*. Behav Ecol Sociobiol 64:1473–1483

Leder EH, Karaiskou N, Primmer CR (2008) Seventy new microsatellites for the pied flycatcher, *Ficedula hypoleuca* and amplification in other passerine birds. Mol Ecol Resour 8:874–880

Lifjeld JT, Slagsvold T, Ellegren H (1997) Experimental mate switching in pied flycatchers: male copulatory access and fertilization success. Anim Behav 53:1225–1232

Lifjeld JT, Laskemoen T, Kleven O, Albrecht T, Robertson RJ (2010) Sperm length variation as a predictor of extrapair paternity in passerine birds. PLoS One 5:e13456

Mazuc J, Chastel O, Sorci G (2003) No evidence for differential maternal allocation to offspring in the house sparrow (*Passer domesticus*). Behav Ecol 14:340–346

Mitchell DP, Dunn PO, Whittingham LA, Freeman-Gallant CR (2007) Attractive males provide less parental care in two populations of the common yellowthroat. Anim Behav 73:165–170

Møller AP (2004) Protandry, sexual selection and climate change. Glob Change Biol 10:2028–2035

Møller AP, Birkhead TR (1993) Cuckoldry and sociality: a comparative study of birds. Am Nat 142:118–140

Mougeot F (2004) Breeding density, cuckoldry risk and copulation behaviour during the fertile period in raptors: a comparative analysis. Anim Behav 67:1067–1076

Pärt T, Qvarnström A (1997) Badge size in collared flycatchers predicts outcome of male competition over territories. Anim Behav 54:893–899

Quay WB (1986) Cloacal protuberance and cloacal sperm in passerine birds: comparative study of quantitative relations. Condor 88:160–168

Qvarnström A (1997) Experimentally increased badge size increases male competition and reduces male parental care in the collared flycatcher. Proc R Soc Lond B 264:1225–1231

Qvarnström A (1999) Different reproductive tactics in male collared flycatchers signalled by size of secondary sexual character. Proc R Soc Lond B 266:2089–2093

Qvarnström A (2001) Context-dependent genetic benefits from mate choice. Trends Ecol Evol 16:5–7

Qvarnström A, Pärt T, Sheldon BC (2000) Adaptive plasticity in mate preference linked to differences in reproductive effort. Nature 405: 344–347

R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org/

Richardson DS, Burke T (1999) Extra-pair paternity in relation to male age in Bullock's orioles. Mol Ecol 8:2115–2126

Rosivall B, Szöllösi E, Hasselquist D, Török J (2009) Effects of extrapair paternity and sex on nestling growth and condition in the collared flycatcher, *Ficedula albicollis*. Anim Behav 77:611–617

Scordato ESC, Safran RJ (2014) Geographic variation in sexual selection and implications for speciation in the Barn Swallow. Avian Res 5:8

Sheldon BC (2000) Differential allocation: tests, mechanisms and implications. Trends Ecol Evol 15:397–402

Sheldon BC, Ellegren H (1999) Sexual selection resulting from extrapair paternity in collared flycatchers. Anim Behav 57:285–298

Sheldon BC, Merilä J, Qvarnström A, Gustafsson L, Ellegren H (1997) Paternal genetic contribution to offspring condition predicted by size of male secondary sexual character. Proc R Soc Lond B 264:297–302

Smith CC (2012) Opposing effects of sperm viability and velocity on the outcome of sperm competition. Behav Ecol 23:820–826

Snook RR (2005) Sperm in competition: not playing by the numbers. Trends Ecol Evol 20:46–53

Török J, Hegyi G, Garamszegi LZ (2003) Depigmented wing patch size is a condition-dependent indicator of viability in male collared flycatchers. Behav Ecol 14:382–388

Tottrup AP, Thorup K (2008) Sex-differentiated migration patterns, protandry and phenology in North European songbird populations. J Ornithol 149:161–167

Článek 17: Edme, A., Zobač, P., Korsten, P., Albrecht, T., Schmoll, T. & Krist, M. (2019) Moderate heritability and low evolvability of sperm morphology in a species with high risk of sperm competition, the collared flycatcher *Ficedula albicollis*. *Journal of Evolutionary Biology*, 32, 205–217.

RESEARCH PAPER

WILEY | JOURNAL OF Evolutionary Biology ⌀eseb

Moderate heritability and low evolvability of sperm morphology in a species with high risk of sperm competition, the collared flycatcher *Ficedula albicollis*

Anaïs Edme[1] | Petr Zobač[1] | Peter Korsten[2] | Tomáš Albrecht[3,4] | Tim Schmoll[5] | Miloš Krist[1,6] (iD)

[1]Faculty of Science, Department of Zoology and Laboratory of Ornithology, Palacky University, Olomouc, Czech Republic

[2]Department of Animal Behaviour, Bielefeld University, Bielefeld, Germany

[3]Institute of Vertebrate Biology, Czech Academy of Sciences, Brno, Czech Republic

[4]Faculty of Science, Department of Zoology, Charles University in Prague, Prague, Czech Republic

[5]Evolutionary Biology, Bielefeld University, Bielefeld, Germany

[6]Museum of Natural History, Olomouc, Czech Republic

Correspondence
Miloš Krist, Department of Zoology and Laboratory of Ornithology, Faculty of Science, Palacky University, Olomouc, Czech Republic.
Email: milos.krist@upol.cz

Present address
Tim Schmoll, Institute of Evolutionary Ecology and Conservation Genomics, Ulm University, Ulm, Germany

Funding information
IGA UPOL, Grant/Award Number: IGA_PrF_2018_016; German Research Foundation, Grant/Award Number: P506/12/2472; Ministry of Education, Youth and Sports, Grant/Award Number: MSMT-56147/2012-310

Abstract

Spermatozoa represent the morphologically most diverse type of animal cells and show remarkable variation in size across and also within species. To understand the evolution of this diversity, it is important to reveal to what degree this variation is genetic or environmental in origin and whether this depends on species' life histories. Here we applied quantitative genetic methods to a pedigreed multigenerational data set of the collared flycatcher *Ficedula albicollis*, a passerine bird with high levels of extra-pair paternity, to partition genetic and environmental sources of phenotypic variation in sperm dimensions for the first time in a natural population. Narrow-sense heritability (h^2) of total sperm length amounted to 0.44 ± 0.14 *SE*, whereas the corresponding figure for evolvability (estimated as coefficient of additive genetic variation, CV_a) was 0.02 ± 0.003 SE. We also found an increase in total sperm length within individual males between the arrival and nestling period. This seasonal variation may reflect constraints in the production of fully elongated spermatozoa shortly after arrival at the breeding grounds. There was no evidence of an effect of male age on sperm dimensions. In many previous studies on laboratory populations of several insect, mammal and avian species, heritabilities of sperm morphology were higher, whereas evolvabilities were similar. Explanations for the differences in heritability may include variation in the environment (laboratory vs. wild), intensity of sexual selection via sperm competition (high vs. low) and genetic architecture that involves unusual linkage disequilibrium coupled with overdominance in one of the studied species.

KEYWORDS
evolvability, heritability, phenotypic plasticity, quantitative genetics, sexual selection, sperm competition, sperm size, wild population

1 | INTRODUCTION

Across the animal kingdom, the function of spermatozoa is to fertilize ova, and hence, its form must have been selected to maximize fertilization success. Despite this common utility, spermatozoa display very large variation in size, shape and motility (Birkhead, Hosken, & Pitnick, 2009). Sperm trait variation exists between species (Alberti, 1995; Jamieson, 1987, 1991; Lifjeld, Laskemoen, Kleven, Albrecht, & Robertson, 2010; Simpson, Humphries, Evans, Simmons, & Fitzpatrick, 2014) but also between populations of the same species (e.g. Hettyey & Roberts, 2006; Laskemoen et al., 2013; Lifjeld et al., 2012; Minoretti & Baur, 2006; Schmoll & Kleven, 2011; Schmoll, Kleven, & Rusche, 2018). Furthermore, substantial variation has been reported between individuals within populations (e.g. Cramer, Laskemoen, Kleven, & Lifjeld, 2013a; Helfenstein, Szép, Nagy, Kempenaers, & Wagner, 2008; Morrow & Gage, 2001b; Ward, 1998) as well as within individual males (Crean & Marshall, 2008; Immler, Pryke, Birkhead, & Griffith, 2010; Schmoll et al., 2018).

Part of this variation may be explained by varying intensity of post-copulatory sexual selection acting on sperm phenotypes. Both sustained directional and stabilizing selection are expected to erode genetic variation (Barton & Keightley, 2002) with the potential consequence of decreased heritability of traits under selection (Mousseau & Roff, 1987; Teplitsky, Mills, Yarrall, & Merilä, 2009; but see Houle, 1992). This process, however, may not be so straightforward, since genetic variation in traits under selection may be maintained through a number of mechanisms including, for example, trade-offs with other traits (Cattelan, di Nisio, & Pilastro, 2018; Immler et al., 2011) and mutation-selection balance (Zhang & Hill, 2005). Nevertheless, despite these mechanisms that may maintain genetic variation despite selection, empirical studies have often found that in species with intense post-copulatory sexual selection, spermatozoa were less variable in size, both between (Calhim, Immler, & Birkhead, 2007; Fitzpatrick & Baer, 2011; Kleven, Laskemoen, Fossøy, Robertson, & Lifjeld, 2008; Lifjeld et al., 2010) and within males (Fitzpatrick & Baer, 2011; Lifjeld et al., 2010; Šandera, Albrecht, & Stopka, 2013; Varea-Sánchez, Gómez Montoto, Tourmente, & Roldan, 2014), in comparison to species with a lower intensity of sexual selection. This observation is usually interpreted as evidence for strong stabilizing selection having depleted genetic variance in sperm length to favour the same optimal phenotype and consequently genotype in species with a high risk and/or intensity of sperm competition (Fitzpatrick & Baer, 2011; Lifjeld et al., 2010; Šandera et al., 2013; Varea-Sánchez et al., 2014).

The genetic sources of phenotypic variation in sperm traits have been studied in several species of insects (Baer et al., 2006; Dobler & Hosken, 2010; Morrow & Gage, 2001a; Simmons & Kotiaho, 2002; Ward, 2000) and mammals (Napier, 1961; Woolley, 1971; Woolley & Beatty, 1967), but as yet only in a single bird species: captive zebra finches (Birkhead, Pellatt, Brekke, Yeates, & Castillo-Juarez, 2005; Kim et al., 2017; Knief et al., 2017; Mossman, Slate,

Humphries, & Birkhead, 2009). Heritability estimates of sperm morphology from these studies were usually high, with often more than 50% of the phenotypic variation in sperm morphological traits attributable to additive genetic effects (Table 1). Such high values may be remarkable given that many of the studied species are important models for studying sexual selection and sperm competition [e.g. dung fly, *Scatophaga stercoraria*, Parker (1970), dung beetle, *Ontophagus taurus*, Simmons, Tomkins, and Hunt (1999), field cricket, *Gryllus bimaculatus* (Tregenza & Wedell, 1998) and house mouse, *Mus musculus* (Dean, Ardlie, & Nachman, 2006)]. In these species, females typically mate with more than one male, which should lead to stabilizing selection on sperm traits, and depletion of additive genetic variance in sperm traits (see above).

Apart from the operation of specific mechanisms maintaining genetic variation in sperm traits under selection, one general explanation for the observed high heritabilities is that all of these studies were done in the laboratory where environmental variance was highly reduced compared to natural populations. As heritability standardizes the additive genetic variance by the total phenotypic variance of a trait, it does not provide a direct clue to the absolute amount of additive genetic variation that is present in the population. To overcome this problem, an alternative quantity, the coefficient of additive genetic variation, has been proposed for comparisons of trait evolvability (the capacity for adaptive evolution) between populations (Hansen, Pélabon, & Houle, 2011; Houle, 1992; Postma, 2014). Ideally, both heritability and evolvability should be studied in the wild under a natural level of environmental variation and strength of selection. However, such studies are lacking for sperm morphological traits (see Table 1).

Besides the low environmental variance in the laboratory, some of the previous heritability estimates may be high due to a very specific architecture of loci influencing sperm morphology. This may be the case in the only studied bird species, the zebra finch *Taeniopygia guttata*, where the heritability of total sperm length was estimated to be 0.63 ± 0.11 *SE* (Mossman et al., 2009). Recent genetic studies in the zebra finch have shown that genes underlying sperm morphology are located on the Z chromosome which has a low recombination rate due to a large chromosomal inversion in some haplotypes (Kim et al., 2017; Knief et al., 2017). Consequently, these linked genes act as one "supergene" with a major effect on sperm phenotype. Moreover, this genetic variation is maintained by balancing selection as heterozygotes (one normal and one inverted Z chromosome) have the fastest sperm and greatest fertilization success (Kim et al., 2017; Knief et al., 2017). As such specific genetic architecture might be taxonomically limited to only the zebra finch and perhaps its closest relatives, studies in other avian taxa which presumably do not have Z chromosome inversions affecting sperm phenotype will be crucial for improving our understanding of the inheritance and evolution of avian sperm traits.

In addition to genetic effects, sperm traits have been found to vary with certain environmental and ontogenetic factors. For example, sperm and ejaculate traits can vary with male age. In some species, older males have longer, faster and more spermatozoa

TABLE 1 Overview of studies estimating heritability of sperm morphology. Heritability estimates are based on mean length of a particular sperm section per per sperm sample. For some studies, estimates obtained by two alternate methods are provided. Populations were studied under different conditions: laboratory (long-term maintenance of the population under laboratory conditions), captive (parental generation was collected in the wild but then it was, together with its descendants, maintained in the laboratory). CV_a is the coefficient of additive genetic variation. Number of spermatozoa indicates how many spermatozoa were used for the calculation of mean sperm size per sperm sample

Study	Species	Sperm trait	Population	Heritability (±SE)	CV_a	Number of spermatozoa
Baer et al. (2006)	*Bombus terrestris*	Total length	Captive	0.43 ± 0.15 (0.20 ± 0.09)[a]	0.023	5
Birkhead et al. (2005)	*Taeniopygia guttata*	Head length	Laboratory	0.48 ± 0.03	0.038	5
Birkhead et al. (2005)	*Taeniopygia guttata*	Midpiece length	Laboratory	0.46 ± 0.09	0.097	5
Birkhead et al. (2005)	*Taeniopygia guttata*	Flagellum length	Laboratory	0.60 ± 0.12	0.082	5
Dobler & Hosken (2010)	*Scathophaga stercoraria*	Total length	Captive	0.53 ± 0.09 (0.45 ± 0.02)[b]	0.014	15
Kim et al. (2017)	*Taeniopygia guttata*	Head length	Laboratory	0.41 (0.13–0.63)[c]	0.037	5
Kim et al. (2017)	*Taeniopygia guttata*	Midpiece length	Laboratory	0.39 (0.22–0.55)[c]	0.078	5
Kim et al. (2017)	*Taeniopygia guttata*	Tail length	Laboratory	0.59 (0.46–0.71)[c]	0.170	5
Kim et al., (2017)	*Taeniopygia guttata*	Total length	Laboratory	0.59 (0.47–0.70)[c]	0.051	5
Morrow & Gage (2001a)	*Gryllus bimaculatus*	Total length	Laboratory	0.52 ± 0.06 (0.29 ± 0.20)[d]	–	10
Mossman et al. (2009)	*Taeniopygia guttata*	Head length	Laboratory	0.51 ± 0.09[e]	–	5
Mossman et al. (2009)	*Taeniopygia guttata*	Midpiece length	Laboratory	0.68 ± 0.09[e]	–	5
Mossman et al. (2009)	*Taeniopygia guttata*	Tail length	Laboratory	0.94 ± 0.01[e]	–	5
Mossman et al. (2009)	*Taeniopygia guttata*	Flagellum length	Laboratory	0.95 ± 0.09[e]	–	5
Mossman et al. (2009)	*Taeniopygia guttata*	Total length	Laboratory	0.63 ± 0.11 (0.97 ± 0.09)[e]	–	5
Minoretti, Stoll, & Baur (2013)	*Arianta arbustorum*	Total length	Captive	0.52 ± 0.55	0.017	25–30
Napier (1961)	*Oryctolagus cuniculus*	Head length	Laboratory	0.72 ± 0.18	0.026	10
Napier, (1961)	*Oryctolagus cuniculus*	Head breadth	Laboratory	0.71 ± 0.13	–	10
Napier, (1961)	*Oryctolagus cuniculus*	Head area	Laboratory	0.74 ± 0.15	–	10
Napier, (1961)	*Oryctolagus cuniculus*	Head shape	Laboratory	0.56 ± 0.13	–	10
Simmons & Kotiaho (2002)	*Ontophagus taurus*	Total length	Captive	0.57 ± 0.31 (1.14 ± 0.61)[d]	0.028	10
Ward (2000)	*Scathophaga stercoraria*	Total length	Laboratory	0.69 ± 0.23	0.017	20
Woolley (1971)	*Mus domesticus*	Midpiece length	Laboratory	0.76 ± 0.02	0.009	10
Woolley & Beatty (1967)	*Mus domesticus*	Midpiece length	Laboratory	0.97 ± 0.36	0.011	10

[a]Based on individual spermatozoa (i.e. including within-ejaculate variability). [b]Realized heritability. [c]Credible interval. [d]Does not control for the potential linkage of sperm-controlling genes on sex chromosomes. [e]Bivariate model.

(Gasparini, Marino, Boschetto, & Pilastro, 2010; Green, 2003; Laskemoen et al., 2010). Furthermore, sperm morphology has been found to vary with the advancement of the breeding season in natural bird populations (e.g. Calhim, Lampe, Slagsvold, & Birkhead, 2009; Cramer, Laskemoen, Kleven, & Lifjeld, 2013a; Lüpold, Birkhead, & Westneat, 2012; Schmoll et al., 2018). However, it remained unclear whether the observed variation is generally caused by phenotypic plasticity within individual males or by selective sampling of males with different sperm phenotypes at different points in time, because often only one sperm sample was taken per individual male (but see Lüpold et al., 2012; Schmoll et al., 2018). Repeated sperm sampling from the same individual at different time points during the breeding season and at different ages is essential for disentangling these two sources of variation in sperm traits (i.e. within-individual vs. between-individual variation).

In this study, we focus on a wild population of the collared flycatcher *Ficedula albicollis*, a species with a high risk of sperm competition, to partition environmental and genetic sources of phenotypic variation in sperm morphological traits using quantitative genetic methods based on a large multigenerational pedigreed data set. We demonstrate moderate heritabilities, low evolvabilities and seasonal variation of sperm dimensions, the latter resulting from individual phenotypic plasticity.

2 | MATERIALS AND METHODS

2.1 | Study population and field methods

This study was conducted in the years 1998–2015 (breeding data for construction of the pedigree) and 2012–2016 (sperm sampling). The study area is situated in Velký Kosíř (49°32'N, 17°04'E) in Moravia, Czech Republic. Five study plots are located on a hill (rising 300–400 m above sea level) covered by oak forest (*Quercus petraea*). In total, the study area harbours around 350 nest boxes (diameter of entrance 32 mm, inner dimensions 12 cm × 11 cm × 24 cm).

The collared flycatcher is a 13 g migratory passerine wintering in Southern Africa and arriving around mid-April at the breeding sites in the Czech Republic (Briedis et al., 2016). In our population, sperm competition is prevalent as approximately 25% of all chicks are sired by an extra-pair male (Edme et al., 2017; Krist & Munclinger, 2011; Krist, Nádvorník, Uvírová, & Bureš, 2005). Females usually lay 5–7 eggs in the first clutch of the season. Replacement clutches occur in cases of clutch failure, but there are no true second clutches.

We caught males twice during the breeding season. First, we trapped them at the time of their spring arrival with small traps when they entered the nest boxes while searching for potential breeding sites (median catch date: 24th April) which was about 10 days before clutch initiation (median date: 4th May). These traps were activated in all nest boxes simultaneously at variable intervals (one-day to one-week) from mid-April to mid-May depending on the number of newly arriving males on the previous catching day. If many new males had arrived, we trapped the subsequent day too, if only a few of them had arrived, the interval to next trapping was prolonged. We consider

the date of first trapping of an individual in any given year as the date of its spring arrival on the study site, which was validated by data from geolocators carried by 16 males (for details, see Edme et al., 2017). Second, we trapped males when they were feeding nestlings (median catch date: 27th May). We determined male age by plumage characteristics as either yearling (brownish primaries) or older (black primaries) (Cramp & Perrins, 1993). When males were not ringed, we gave them a ring with a unique number (Praha Ringing Centre, Czech Republic). All adults and nestlings were blood-sampled (1–5 µl) from the tarsal vein (males during the arrival and nestling feeding periods, females during the nestling feeding period only).

Sperm samples were collected by cloacal massage (Girndt, Cockburn, Sánchéz-Trójar, Løvlie, & Schroeder, 2017; Quay, 1986) of males captured during both the arrival and nestling feeding periods from 2013 to 2015. In 2012, males were only sampled during the feeding period and in 2016 only during the arrival period. Cloacal massage usually provided a small amount of liquid including faeces and/or sperm. Thus, in contrast to more invasive methods, such as testicular dissection, our method has the advantage of sampling possibly mature spermatozoa that are ready to be released into the cloaca. The disadvantage was that we did not obtain enough spermatozoa in some of the samples. All samples were stored in 10% formalin at room temperature or at 8°C in a refrigerator until processing. Such storage does not affect the size measurements of avian spermatozoa (Schmoll, Sancipri an, & Kleven, 2016). The methods employed in this study were ethically reviewed and approved by the ethical committee of Palacky University and by the Ministry of Education, Youth and Sports (licence number: MSMT-56147/2012-310) and comply with the current laws of the Czech Republic.

2.2 | Sperm morphometry

Microscopic slides were prepared after each breeding season by spreading out 7 µl of the sperm solution, letting it dry, rinsing it with distilled water and letting it dry again. To obtain a sufficient number of spermatozoa for our measurements, we prepared up to three slides per sperm sample in cases where less than 20 spermatozoa were present in the first (or second) slide. A total of 860 faecal/sperm samples were collected in the field, but 20 or more spermatozoa were found in only 600 of them (see Supporting information Table S1 for sample sizes). These 600 samples belonged to 367 different males. We took photographs of 20 morphologically normal-looking spermatozoa per sample.

Spermatozoa sampled in 2012 and 2013 were photographed using an Olympus BX51 microscope and an Olympus DP71 microscope digital camera, and those sampled from 2014 to 2016 using an Olympus CX41 microscope equipped with an Infinity 2 digital camera. In both cases, a 400x magnification under light-field conditions was used. We included microscope type as a fixed factor in all analyses to control for potential minor differences in set-up/magnification between the two microscopes. For each spermatozoon, its head, midpiece and tail were measured using ImageJ software 1.49v (Schneider, Rasband, & Eliceiri, 2012). Total sperm length was

calculated by summing up the three sperm sections, and flagellum length is the sum of midpiece and tail length. All measurements were made by the same person (PZ) to avoid measurement error resulting from inter-observer differences.

Repeatability and thereby intra-observer measurement error of sperm length was assessed by taking measurements of the same 20 spermatozoa twice for 30 randomly selected sperm samples (but only one sample per male). The 30 sperm samples (i.e. 600 spermatozoa) were measured twice in the same order. Linear mixed-effects models were fitted to obtain repeatability estimates (response variable: length of specific sperm section; fixed effect: intercept only; random effect: sperm identity). The repeatabilities of the measurements of the length of the different sperm sections ranged from 0.796 to 0.960 (Table 2), as calculated with the R package "rptR" (Stoffel, Nakagawa, & Schielzeth, 2017).

As we were mainly interested in between-sample and between-male variation in sperm morphology and strived to minimize the impact of measurement error, we disregarded within-sample variation and calculated mean sperm dimensions from 20 individual spermatozoa in each sample (but see Table 2 for within-sample repeatabilities and Supporting information Table S2 for variance components based on individual spermatozoa). The analyses presented, except for the estimation of intra-observer measurement error, were thus based on the mean sperm length per sample, which was the approach adopted by all previous studies (Table 1).

2.3 | Pedigree

The breeding ecology of the collared flycatcher on the study site has been recorded since 1998, and the population is highly philopatric (Krist, 2009). Therefore, extensive social pedigree information is available. We constructed the pedigree for the years 1998–2015 according to observations in the field where adults were considered the social parents of a brood when providing parental care in the form of nestling feeding. We pruned the pedigree with respect to available information about sperm size using the package "pedantics" in R (Morrissey, 2014). The resulting pruned pedigree contained 781 informative individuals, 14003 nonzero pairwise relatedness estimates

and had a maximum depth of 10 generations (see Supporting information Table S3 for additional pedigree statistics).

We conducted paternity analyses for part of the population in 2001–2002 (Krist et al., 2005), 2006–2009 (Krist & Munclinger, 2011) and for the whole population in 2013 (Edme et al., 2017). We did not find any case of intra-specific brood parasitism. Therefore, social mothers were considered genetic mothers in all cases. We determined paternity status (sired within-pair or extra-pair) for 1396 of 7700 offspring in the complete pedigree; 978 of them were sired by social males and 328 by extra-pair males, which constitutes a 23.5% rate of extra-pair paternity. The identity of extra-pair fathers was determined for 207 extra-pair offspring. In addition, we also assigned genetic sires for 90 offspring with unknown social fathers. We corrected the social pedigree with all of the available genetic information. As we were able to do this for only about 18% of the offspring (1396/7700), our estimates of heritability may be lower than the true values. However, for rates of extra-pair paternity up to 40%, a simulation study has shown that heritability was underestimated only by 20% as a maximum (Charmantier & Réale, 2005) or by 15% as a maximum under nonrandom extra-pair mating but lower rates of extra-pair paternity (12.5%: Firth, Hadfield, Santure, Slate, & Sheldon, 2015).

2.4 | Statistical analyses

All analyses were performed in R 3.5.1. (R Core Team, 2018). In order to investigate genetic *versus* environmental sources of variation in sperm dimensions, we fitted linear mixed-effects models which decomposed the variance according to its different sources through the inclusion of random effect terms, while accounting for various fixed effects.

As fixed effects, we included microscope type (categorical), male age (categorical: yearling/older) and sampling period (categorical: arrival/feeding). We applied within-subject centring to male age and sampling period in order to allow for distinguishing within-individual from between-individual effects (see van de Pol & Wright, 2009). The between-individual predictor is then represented by the mean per individual, whereas the within-individual

TABLE 2 Repeatabilities (±SE) of sperm dimensions calculated at various grouping levels. Repeatabilities for the same spermatozoa measured twice by the same person and repeatabilities for sperm dimensions within samples are based on measurements of individual spermatozoa. Repeatabilities within males and years are based on mean sperm size per sample (means from 20 individual spermatozoa of the sperm sample). Repeatabilities based on sperm samples, males and years are adjusted for the fixed effects of male age, period of sampling and microscope type

| | Individual spermatozoa | | Mean sperm size per sample | |
Sperm section	Measurements (N = 600 spermatozoa measured twice)	Samples (N = 12,000 spermatozoa from 600 samples)	Male identity (N = 600 samples from 367 males)	Year (N = 600 samples from 5 years)
Total	0.959 ± 0.003	0.421 ± 0.015	0.607 ± 0.042	<0.001
Flagellum	0.960 ± 0.003	0.434 ± 0.017	0.613 ± 0.043	0.028 ± 0.032
Head	0.796 ± 0.017	0.304 ± 0.014	0.356 ± 0.058	0.214 ± 0.133
Midpiece	0.960 ± 0.003	0.338 ± 0.017	0.412 ± 0.050	0.009 ± 0.019
Tail	0.912 ± 0.007	0.243 ± 0.014	0.485 ± 0.047	0.019 ± 0.020

predictor is the difference between the mean per individual and each of the individuals' measurements (van de Pol & Wright, 2009). As both the variables were categorical, we transformed them to assume the values 0 and 1 before applying within-subject centring of the data.

We built the random part of the models in two steps. Our first aim was to calculate repeatabilities of sperm dimensions defined as the proportion of the total variance accounted for by differences between various grouping levels. Therefore, we fitted sets of models that included the fixed effects listed above and one of the following random effect terms representing grouping factors: sperm sample, male identity and year for calculation of adjusted repeatabilities within samples (this analysis was based on individual spermatozoa measurements); males and year (these analyses were based on means per sperm sample). All models were fitted in the R package rptR (Stoffel et al., 2017).

Second, we built an animal model by inclusion of the pedigree-based pairwise relatedness matrix to estimate the additive genetic variance (V_a). The advantage of the animal model compared to other methods of estimating quantitative genetic parameters (like parent-offspring regression) is the use of multigenerational pedigree information even if this information is incomplete, unbalanced or complex (Kruuk, 2004; Wilson et al., 2010). The animal model further contained random effects of year and the identities of males, their mothers and nests of origin. Consequently, we were able to partition the phenotypic variance into the variance components due to additive genetic (V_a), permanent environment (V_{pe}), maternal ($V_{maternal}$) and nest-of-origin (V_{nest}) effects. The last two components might be slightly underestimated because 64 of 367 males had been exchanged between nests as eggs or young nestlings and consequently identity of original nest and genetic mother does not include nongenetic effects which potentially arose after cross-fostering. However, inclusion of identities of foster mother and foster nest instead of that of genetic mother and original nest gave very similar results (not shown). We estimated the narrow-sense heritability (h^2) for each of the focal traits, defined as the proportion of phenotypic variance (V_p) explained by additive genetic effects: $h^2 = V_a/V_p$.

In addition, we also calculated coefficients of additive genetic variation according to the formula: $CV_a = \sqrt{V_a}/\bar{x}$, where V_a is the additive genetic variance and $\bar{x}$ is the trait mean (Houle, 1992). We employed the same method for calculation of CV_a for previously published studies (Table 1). If V_a was not reported by a study, we estimated it as the product of phenotypic variance and heritability (Garcia-Gonzalez, Simmons, Tomkins, Kotiaho, & Evans, 2012; Postma, 2014). We calculated approximate standard errors of CV_a according to the formula: $SE\ (CV_a) \approx SE\ (V_a)/(2\mu \times \sqrt{V_a})$ (eq. 6 in Garcia-Gonzalez et al. (2012).

Moreover, using bivariate animal models we also estimated the covariances between different sperm sections to assess to what extent these can evolve independently of one another. We calculated the genetic correlation between traits x and y as: $r_a = COV_{a(xy)}/\sqrt{(V_{a(x)} \times V_{a(y)})}$, where COV_a is the additive genetic covariance between the two traits and V_a are their additive genetic variances

TABLE 3 Variance components from univariate and covariances from bivariate animal models for sperm total length and the sperm sections: head, midpiece, tail and flagellum (midpiece and tail combined). Subscripts a and pe refer to additive genetic and permanent environmental variance respectively. Heritabilities, genetic correlations (r_a) and coefficients of additive genetic variation (CV_a) were calculated using equations given in the Methods. Standard errors (SE) are provided with all estimates

	V_a	V_{pe}	$V_{maternal}$	V_{nest}	V_{year}	$V_{residual}$	Heritability	CV_a
Total length	4.089 ± 1.304	0.841 ± 1.214	0.200 ± 0.764	0.631 ± 0.996	0.019 ± 0.070	3.436 ± 0.318	0.444 ± 0.143	0.021 ± 0.003
Flagellum	3.831 ± 1.271	1.128 ± 1.202	0.121 ± 0.752	0.681 ± 1.031	0.263 ± 0.275	3.280 ± 0.303	0.412 ± 0.137	0.022 ± 0.004
Head	0.066 ± 0.033	0.031 ± 0.031	<0.001	<0.001	0.072 ± 0.061	0.184 ± 0.016	0.187 ± 0.096	0.023 ± 0.006
Midpiece	2.467 ± 1.326	2.388 ± 1.344	<0.001	<0.001	0.102 ± 0.177	6.800 ± 0.611	0.210 ± 0.136	0.023 ± 0.006
Tail	0.799 ± 0.854	2.632 ± 0.910	<0.001	<0.001	0.125 ± 0.146	3.572 ± 0.322	0.112 ± 0.121	0.050 ± 0.027

	COV_a	COV_{pe}	$COV_{maternal}$	COV_{nest}	COV_{year}	$COV_{residual}$	r_a	
Head–Midpiece	0.115 ± 0.147	−0.197 ± 0.147	–	–	−0.061 ± 0.093	0.006 ± 0.077	0.278 ± 0.366	
Head–Tail	−0.173 ± 0.117	0.172 ± 0.121	–	–	−0.074 ± 0.079	−0.057 ± 0.056	−0.828 ± 0.710	
Midpiece–Tail	0.188 ± 0.810	−1.24 ± 0.875	–	–	−0.170 ± 0.212	−4.33 ± 0.443	0.122 ± 0.577	

FIGURE 1 Relative contribution of particular variance components to the total phenotypic variance of sperm sections as identified by the animal model. Note that the relative contribution of V_a corresponds to heritability, whereas the relative contribution of $V_a + V_{pe}$ corresponds to repeatability
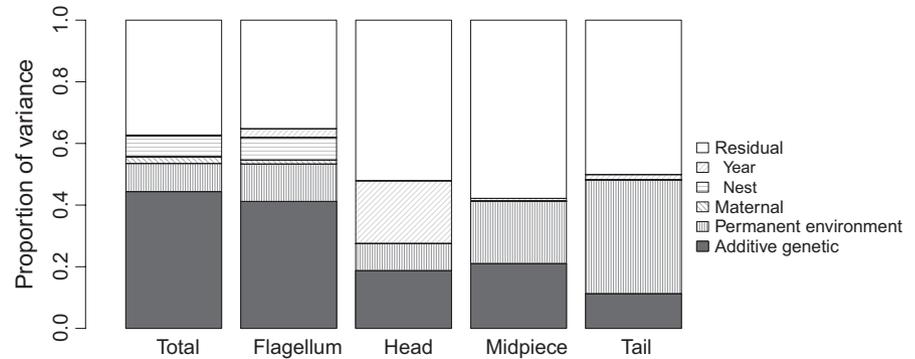
(Falconer, 1989, p. 317). For the bivariate animal models, inclusion of maternal or nest of origin identity precluded model convergence. Therefore, these models included only the random effects of year and male identity. All quantitative genetic models were fitted in the package ASReml-R, version 3.0 (Butler, 2009).

3 | RESULTS

3.1 | Repeatabilities and quantitative genetic parameters of sperm dimensions

Based on the phenotypic model, we found high within-male (between-sample) repeatability for flagellum and total length and lower repeatability for other sperm sections (Table 2). Head and tail length were also weakly repeatable within years (Table 2). As with repeatability, the heritability of total sperm length and flagellum was higher ($h^2 = 0.44 \pm 0.14$ SE and 0.41 ± 0.14 SE, respectively) compared to smaller heritabilities of separate sperm sections ($h^2 = 0.11–0.21$) with their confidence intervals (i.e. estimate ± 2 SE) overlapping zero (Table 3, Figure 1). This contrast was caused by the relatively smaller additive genetic variances for the separate sperm sections as compared with larger permanent environment variances (midpiece and tail) or annual variance (head) as well as increased unexplained, residual variance (Table 3, Figure 1). Maternal and nest-of-origin effects explained little variance in models of total sperm and flagellum length and negligible variance in models for the separate sperm sections (Table 3).

Evolvability, measured as the coefficient of additive variation (CV_a), was small for all sperm traits. The largest value (0.050 ± 0.027 SE) was found for sperm tail length, that is the trait with the lowest heritability (Table 3). The additive genetic covariance and genetic correlation between the lengths of midpiece and tail were small and positive, which is in contrast to the large, negative residual covariance of these two components (Table 3). However, all estimates of genetic covariances and correlations had large standard errors (Table 3), which suggests that our dataset had little statistical power in this respect.

3.2 | Seasonal and age effects on sperm dimensions

Although sperm dimensions did not differ between yearlings and older males, they varied significantly with sampling period (Tables 4–6).

TABLE 4 Mean length±standard deviation (SD) (µm) of sperm sections for the arrival period (median date 24th April, $n = 8500$ spermatozoa) and nestling feeding period (median date 27th May, $n = 3500$ spermatozoa)

Section	Arrival	Feeding
Total	97.99 ± 4.68	99.47 ± 4.00
Flagellum	86.99 ± 4.56	88.27 ± 3.90
Head	11.00 ± 1.04	11.19 ± 0.97
Midpiece	68.88 ± 5.98	71.03 ± 4.72
Tail	18.10 ± 5.34	17.24 ± 4.06

Specifically, head size and midpiece increased in length during the breeding season, whereas the tail demonstrated an opposite change (Tables 4 and 6). In absolute terms, the seasonal change was largest for the longest section, the midpiece, which led to longer flagella and thus longer spermatozoa in the feeding period (Tables 4–5, Figure 2). Despite high statistical significance of these seasonal effects, their magnitude was only weak to medium. For example, the average increase in total sperm length over a month was about 0.35 SD (Table 4). The effects within and between individuals were always in the same direction and usually similar in magnitude (Tables 5–6). Note that although the mean within-male effect was positive (i.e. increase in size from arrival to feeding stage), there was also considerable variation in these within-male changes (Figure 2) with about one-quarter of males exhibiting the opposite trend (Figure 2).

The length of the midpiece and tail was strongly negatively correlated at the phenotypic level ($r = -0.558$, $n = 600$, $p < 0.001$), whereas the two remaining relationships were much weaker (head–midpiece: $r = 0.105$, $p = 0.010$; head–tail: $r = -0.091$, $p = 0.026$). The strong negative phenotypic correlation between the midpiece and tail is likely to be mostly environmental in origin, given the small additive genetic covariance and small genetic correlation, but large negative residual covariance between these two components (see section above; Table 3).

4 | DISCUSSION

4.1 | Genetic sources of variation

In order to investigate environmental and genetic sources of variation in avian sperm size in a wild bird population, we estimated

TABLE 5 Fixed effects part of animal models for sperm total length and flagellum length. The effects of male age and sampling period were centred within males. Negative estimates for age indicate longer sperm for young males and positive estimates for period indicate longer sperm in the feeding period

	Total sperm length				Flagellum			
	Estimate ± SE	df	F	p	Estimate ± SE	df	F	p
Intercept	92.609 ± 1.004	18.1	8506	<0.001	82.669 ± 1.327	11.9	3880	<0.001
Age (between males)	−0.263 ± 0.352	392.6	0.6	0.456	−0.330 ± 0.352	400.5	0.9	0.349
Age (within males)	−0.027 ± 0.453	145.5	<0.1	0.952	−0.092 ± 0.461	253.1	<0.1	0.842
Microscope	2.336 ± 0.284	2.0	67.8	0.015	1.782 ± 0.556	3.3	10.3	0.043
Period (between males)	1.765 ± 0.447	278.3	15.6	<0.001	1.629 ± 0.457	387.9	12.7	<0.001
Period (within males)	1.534 ± 0.224	224.5	47.0	<0.001	1.206 ± 0.223	241.8	29.3	<0.001

repeatabilities, heritabilities and evolvabilities of sperm dimensions in the promiscuous collared flycatcher. We found within-male repeatabilities of mean sperm dimensions ranging from 0.36 to 0.61. In line with the magnitude of these within-male repeatabilities, we found sperm dimensions to be moderately heritable in our population. The heritability of total sperm length amounted to 0.45 ± 0.14 SE, whereas heritabilities for separate sperm sections ranged from 0.11 to 0.21 and their confidence intervals overlapped zero. Evolvabilities of all sperm dimensions measured as their CV_a were only around 0.02 with the exception of sperm tail length which had an estimate of 0.05.

It is important to realize that heritability estimates are conditioned on the structure of the statistical model used to calculate them because heritability standardizes additive genetic variance by the total phenotypic variance (Wilson, 2008). Evolvabilities are less sensitive to model structure since they standardize additive genetic variance by the trait mean, not its variance (Houle, 1992). For example, if we also considered within-sample variation in sperm length and used individual spermatozoa instead of the sample means for the calculation of heritabilities, these would be much lower due to larger unexplained, residual variation. The heritability of total sperm length would decrease to 0.21 ± 0.07 SE in such a model (see Supporting information Table S2), whereas evolvability would remain at the level of 0.02 ± 0.003. The increase in heritability due to within-sample averaging corresponds very closely to findings by Baer et al. (2006) in bumblebees, *Bombus terrestris* (Table 1). Similarly, heritability of total sperm length would also decrease to 0.35 ± 0.13 SE if we used mean sperm length per sample but did not fit any fixed effects that decrease the amount of residual variation (details not shown). Taken together, although we used a model that controlled for some known sources of variation and thus provided relatively high heritability estimates of sperm dimensions in the collared flycatcher, these were still substantially lower than those found in most previous studies (Table 1).

One explanation for the smaller heritability estimates in this study might be that in contrast to all previous studies, we conducted the study in the wild where many (unmeasured) environmental effects increase the residual variance and hence decrease the heritability in

the natural environment where sperm traits have been (and currently are) evolving. When we standardized the additive genetic variances by the trait means, resulting evolvabilities (CV_a) were similar to those reported by most previous studies where they usually ranged from 0.01 to 0.03 (Tables 1, 3). Such small values of CV_a but large heritabilities are typical for morphological traits (Postma, 2014). The exception in our study was sperm tail length for which the evolvability was 0.05 ± 0.027 and thus within the range found in studies on the zebra finch, where CV_a ranged from 0.04 to 0.17 (Table 1).

There are two potential explanations for why evolvabilities were higher in the zebra finch (Birkhead et al., 2005; Kim et al., 2017; Mossman et al., 2009) compared to all other species (Table 1), including most estimates in the collared flycatcher (Table 3). First, sperm competition is less prevalent in wild populations of the zebra finch [rate of extra-pair paternity (EPP): 2% of offspring (Birkhead, Burke, Zann, Hunter, & Krupa, 1990; Griffith, Holleley, Mariette, Pryke, & Svedin, 2010)] compared to the collared flycatcher that had 23% of EPP in our population. Similarly, many of the previous non-zebra finch studies were done in species experiencing high intensities of sperm competition [e.g. dung fly, Parker (1970), dung beetle, Simmons et al. (1999), field cricket (Tregenza & Wedell, 1998) and house mouse (Dean et al., 2006)]. In these species, stabilizing selection on sperm length might have been strong in their evolutionary past, which might have depleted additive genetic variance for this trait (see Calhim et al., 2007; Kleven et al., 2008; Lifjeld et al., 2010). In contrast, selection on sperm traits might have been comparatively weak in the zebra finch, resulting in higher additive genetic variances for its sperm traits. Here we have to note that direct or stabilizing selection on sperm traits was not evident in our population of the collared flycatcher (Edme et al., 2017). However, this does not mean that such selection did not operate in the evolutionary past of the collared flycatcher. Selection may also have remained undetected in the single studied year (2013) since selection often fluctuates over time (e.g. van de Pol, Brouwer, Ens, Oosterbeek, & Tinbergen, 2010) or it may appear low at present because sperm traits have already been optimized in the population.

Second, the zebra finch might in fact have a very specific architecture of genes influencing sperm morphology. Two recent studies

TABLE 6 Fixed effects part of animal models for length of sperm sections. The effects of male age and sampling period were within-male centred. Positive estimates mean longer sperm for old males and in the feeding period

	Head				Midpiece				Tail			
	Estimate ± SE	df	F	p	Estimate ± SE	df	F	p	Estimate±SE	df	F	p
Intercept	9.751 ± 0.460	4.4	449	<0.001	64.768 ± 1.323	15.1	2395	<0.001	17.932 ± 1.124	16.6	254.6	<0.001
Age (between males)	0.039 ± 0.062	456.0	0.4	0.535	−0.416 ± 0.404	440.6	1.1	0.303	0.189 ± 0.317	441.5	0.4	0.552
Age (within males)	0.091 ± 0.108	290.0	07	0.400	0.159 ± 0.644	218.1	<0.1	0.805	−0.335 ± 0.473	246.8	0.5	0.479
Microscope	0.646 ± 0.252	3.1	6.6	0.080	1.578 ± 0.450	2.6	12.3	0.048	0.120 ± 0.424	3.6	0.1	0.793
Period (between males)	0.179 ± 0.083	475.9	4.6	0.032	2.167 ± 0.521	297.4	17.3	<0.001	−0.533 ± 0.412	367.0	1.7	0.197
Period (within males)	0.345 ± 0.053	272.0	43.0	<0.001	2.403 ± 0.316	244.8	57.8	<0.001	−1.206 ± 0.231	248.5	27.3	<0.001

found that these genes are located on the Z chromosome and, due to strong linkage disequilibrium, effectively work as one "super-gene" (Kim et al., 2017; Knief et al., 2017). Consequently, the phenotype largely depends on whether an individual is a homozygote or a heterozygote for normal and inverted haplotypes (Kim et al., 2017; Knief et al., 2017) which may underlie the high evolvability and heritability of sperm morphology, as well as the negative genetic correlation between some sperm sections (Birkhead et al., 2005). As heterozygotes have an advantage in the form of the longest and fastest sperm (Kim et al., 2017; Knief et al., 2017), both haplotypes are maintained in the population.

Lower evolvabilities and heritabilities of sperm morphology in the collared flycatcher suggest more conventional, polygenic inheritance of these quantitative traits as has also been found for other morphological traits in this species (Silva et al., 2017). This view is also supported by a genomic study that found a high recombination rate within the flycatcher Z chromosome, which is in contrast to the low recombination rate in the central part of the Z chromosome in the zebra finch (Kawakami et al., 2014). Furthermore, in line with this difference in the strength of linkage disequilibrium, a strong negative genetic correlation was found between the length of the midpiece and flagellum (and therefore tail) in the zebra finch (Birkhead et al., 2005), whereas the negative phenotypic correlation between the same components that were also found in the collared flycatcher appears to be entirely of environmental origin.

Even though the genome of the collared flycatcher has been sequenced (Ellegren et al., 2012; Kawakami et al., 2014), no genes have been identified yet that influence sperm morphology in this species. Kim et al. (2017) identified several genes explaining variance in sperm morphology in zebra finches. Evaluation of the role of these candidate genes for sperm morphology in the collared flycatcher and other species would be a useful next step for improving our understanding of the genetic basis of sperm evolution under sperm competition in birds.

4.2 | Environmental sources of variation

As in some other studies of birds (Laskemoen et al., 2010; Møller et al., 2009; Rowe, Swaddle, Pruett-Jones, & Webster, 2010), we found no effect of male age on sperm size. This is in contrast to other taxa such as insects and fish where sperm size increased with male age (Gasparini et al., 2010; Green, 2003). Reasons for the different sperm size dynamics over male ontogeny in different taxa are not clear at present.

On the other hand, our results show that sperm size changed with the advancement of the breeding season. Spermatozoa were longer during the nestling period compared with the time when males arrived at the breeding site. Interestingly, the observed changes were not uniform as the head, midpiece (and as a consequence flagellum) increased in size, whereas the tail became shorter over the breeding season. These results are in agreement with previous studies (Cramer, Laskemoen, Kleven, & Lifjeld, 2013; Lüpold et al., 2012), where the length of spermatozoa or their sections also
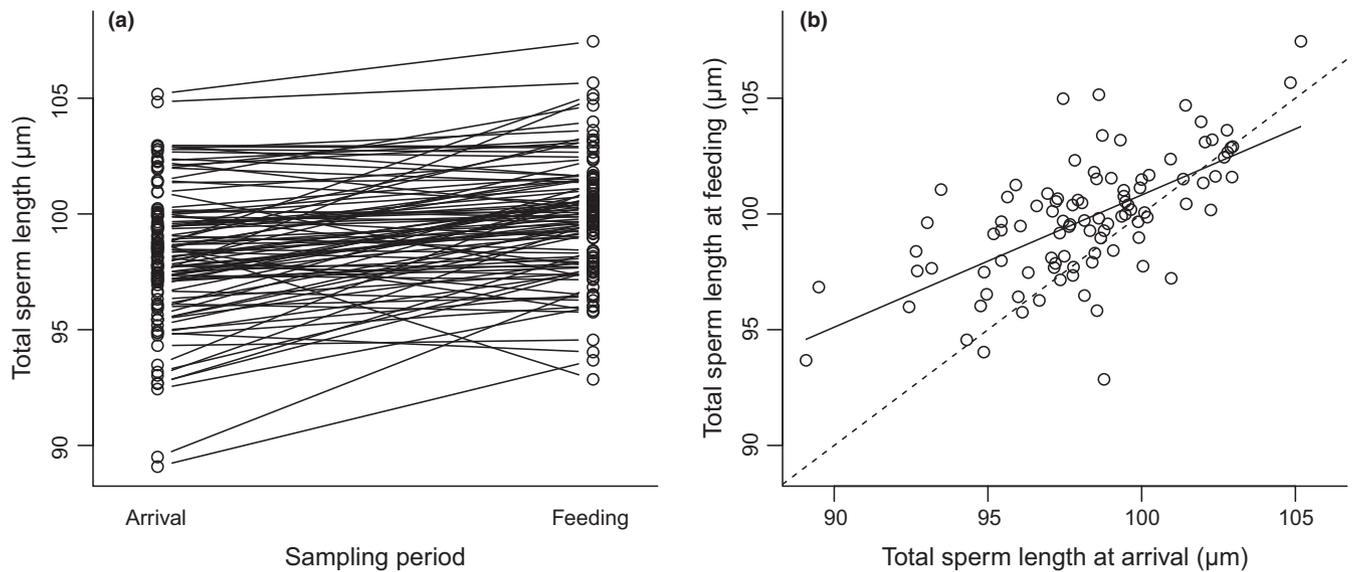
FIGURE 2 Variation of spermatozoa length (points are means of 20 spermatozoa per male) according to the period of sampling (arrival at the breeding site and feeding of the chicks) for 97 males sampled at both arrival and feeding stage in the same year. (a) Lines connect values of the same male sampled in the two periods. (b) Solid line shows the fit of Ordinary Least Squares regression, whereas the dashed one is the isometric reference line ($y = x$)

increased over the course of the season. As we sampled the same males repeatedly, we were also able to determine whether seasonal variation in sperm size was due to within- or between-individual effects. As the between-male effect was similar in magnitude to the within-male one, both of these effects may simply be explained as the result of phenotypic change with no need to invoke selective disappearance of individuals with short sperm from the breeding population. This within-male phenotypic change may either reflect adaptive adjustment or constraints imposed on sperm production.

If males made a strategic allocation of resources into sperm production, they would probably invest in sperm with higher competitive ability early in the breeding season. Later in the season, the number of receptive females drops and males are expected to invest more in parental care and moult than in sperm production. According to this hypothesis, sperm should be longer early in the season (Calhim et al., 2009), as longer sperm may swim faster and be more effective in sperm competition (Bennison, Hemmings, Slate, & Birkhead, 2015). However, our results are in contrast to this as we found sperm to be shorter at the start of the season. Moreover, we also did not find any evidence for greater success of longer sperm in sperm competition in the study population (Edme et al., 2017).

Consequently, seasonal changes in sperm length may rather be the result of a time constraint imposed on this migratory species instead of representing male strategic allocation. Although males already start to produce sperm during migration (Bauchinger, Hof, & Biebach, 2007), their spermatozoa may not have reached the maximum possible length yet just after arrival. However, all spermatozoa that we measured should have been mature in a sense of their ability to fertilize ova, since we sampled only those ready to be released by males into their cloacae. Another possibility is that males copulate at a high rate at the

beginning of the breeding season depleting their sperm stores, and therefore, on average, less mature and thus shorter spermatozoa are obtained in sperm samples early in the year. Indeed, spermatozoa tend to be shorter when ejaculates are sampled from the same male successively (Harris, Moore, & Moore, 2007; G. Crapa, M. Rusche, O. Kleven & T. Schmoll, unpublished data). This may reflect a trade-off between awaiting the availability of more mature longer sperm (longer midpiece and thus flagellum), which allows sperm to swim faster and may increase competitive fertilization success (LaMunyon & Ward, 1998; Bennison et al., 2015; but see Cramer, Laskemoen, Kleven, LaBarbera, et al., 201b; Edme et al., 2017), and copulating at a higher rate using less mature shorter sperm that are already available. The latter strategy may be adaptive if shorter sperm can be produced at a higher rate during the period when sperm competition is at its peak (Immler et al., 2011).

5 | CONCLUSIONS AND FUTURE DIRECTIONS

Further quantitative genetic studies of sperm morphology in other species in the wild would be useful for testing to what extent the heritability and evolvability of sperm morphology are related to the degree of sperm competition as well as its underlying genetic architecture. It would also be worthwhile investigating whether seasonal changes in sperm morphology are evident in migratory species only, which may be under more severe time constraints when breeding, or are found in resident species too, which would suggest that other constraints (and thereby agents of selection) are important in shaping seasonal plasticity of sperm morphology.

CONFLICT OF INTEREST

Authors declare no conflict of interest.

ORCID

Miloš Krist iD https://orcid.org/0000-0002-6183-686X

REFERENCES

Alberti, G. (1995). Comparative spermatology of Chelicerata: Review and perspective. *Mémoires du Muséum national d'histoire Naturelle*, *166*, 203–230.

Baer, B., de Jong, G., Schmid-Hempel, R., Schmid-Hempel, P., Høeg, J. T., & Boomsma, J. J. (2006). Heritability of sperm length in the bumblebee *Bombus terrestris*. *Genetica*, *127*, 11–23. https://doi.org/10.1007/s10709-005-2157-1

Barton, N. H., & Keightley, P. D. (2002). Understanding quantitative genetic variation. *Nature Reviews Genetics*, *3*, 11–21. https://doi.org/10.1038/nrg700

Bauchinger, U., Hof, T. V. T., & Biebach, H. (2007). Testicular development during long-distance spring migration. *Hormones and Behavior*, *51*, 295–305. https://doi.org/10.1016/j.yhbeh.2006.10.010

Bennison, C., Hemmings, N., Slate, J., & Birkhead, T. (2015). Long sperm fertilize more eggs in a bird. *Proceedings of the Royal Society of London B: Biological Sciences*, *282*, 20141897.

Birkhead, T. R., Burke, T., Zann, R., Hunter, F. M., & Krupa, A. P. (1990). Extra-pair paternity and brood parasitism in wild zebra finches Taeniopygia guttata, revealed by DNA fingerprinting. *Behavioral Ecology and Sociobiology*, *27*, 315–324. https://doi.org/10.1007/BF00164002

Birkhead, T. R., Hosken, D. J., & Pitnick, S. (2009). *Sperm Biology*. London: Academic Press.

Birkhead, T. R., Pellatt, E. J., Brekke, P., Yeates, R., & Castillo-Juarez, H. (2005). Genetic effects on sperm design in the zebra finch. *Nature*, *434*, 383–387. https://doi.org/10.1038/nature03374

Briedis, M., Hahn, S., Gustafsson, L., Henshaw, I., Träff, J., Král, M., et al. (2016). Breeding latitude leads to different temporal but not spatial organization of the annual cycle in a long-distance migrant. *Journal of Avian Biology*, *47*, 743–748. https://doi.org/10.1111/jav.01002

Butler, D, . 2009. Asreml: asreml() fits the linear mixed model. R package version 3.0. Retrieved from www.vsni.co.uk

Calhim, S., Immler, S., & Birkhead, T. R. (2007). Postcopulatory sexual selection is associated with reduced variation in sperm morphology. *PLoS ONE*, *2*, e413. https://doi.org/10.1371/journal.pone.0000413

Calhim, S., Lampe, H. M., Slagsvold, T., & Birkhead, T. R. (2009). Selection on sperm morphology under relaxed sperm competition in a wild passerine bird. *Biology Letters*, *5*, 58–61. https://doi.org/10.1098/rsbl.2008.0544

Cattelan, S., di Nisio, A., & Pilastro, A. (2018). Stabilizing selection on sperm number revealed by artificial selection and experimental evolution. *Evolution*, *72*, 698–706. https://doi.org/10.1111/evo.13425

Charmantier, A., & Réale, D. (2005). How do misassigned paternities affect the estimation of heritability in the wild? *Molecular Ecology*, *14*, 2839–2850. https://doi.org/10.1111/j.1365-294X.2005.02619.x

Cramer, E. R. A., Laskemoen, T., Kleven, O., LaBarbera, K., Lovette, I. J., & Lifjeld, J. T. (201a). No evidence that sperm morphology predicts paternity success in wild house wrens. *Behavioral Ecology and Sociobiology*, *67*, 1845–1853. https://doi.org/10.1007/s00265-013-1594-6

Cramer, E. R. A., Laskemoen, T., Kleven, O., & Lifjeld, J. T. (201b). Sperm length variation in House Wrens *Troglodytes aedon*. *Journal of Ornithology*, *154*, 129–138. https://doi.org/10.1007/s10336-012-0878-3

Cramp, S., & Perrins, C. M. (1993). *Handbook of the Birds of Europe, the Middle East and North Africa: Flycatchers to shrikes*, Vol. *7*. Oxford: Oxford University Press.

Crean, A. J., & Marshall, D. J. (2008). Gamete plasticity in a broadcast spawning marine invertebrate. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 13508–13513. https://doi.org/10.1073/pnas.0806590105

Dean, M. D., Ardlie, K. G., & Nachman, M. W. (2006). The frequency of multiple paternity suggests that sperm competition is common in house mice (*Mus domesticus*). *Molecular Ecology*, *15*, 4141–4151. https://doi.org/10.1111/j.1365-294X.2006.03068.x

Dobler, R., & Hosken, D. J. (2010). Response to selection and realized heritability of sperm length in the yellow dung fly (*Scathophaga stercoraria*). *Heredity*, *104*, 61–66. https://doi.org/10.1038/hdy.2009.93

Edme, A., Zobač, P., Opatová, P., Šplíchalová, P., Munclinger, P., Albrecht, T., et al. (2017). Do ornaments, arrival date, and sperm size influence mating and paternity success in the collared flycatcher? *Behavioral Ecology and Sociobiology*, *71*, 3. https://doi.org/10.1007/s00265-016-2242-8

Ellegren, H., Smeds, L., Burri, R., Olason, P. I., Backström, N., Kawakami, T., et al. (2012). The genomic landscape of species divergence in *Ficedula* flycatchers. *Nature*, *491*, 756–760. https://doi.org/10.1038/nature11584

Falconer, D. S. (1989). *Introduction to Quantitative Genetics*, 3d ed. New York: Longman.

Firth, J. A., Hadfield, J. D., Santure, A. W., Slate, J., & Sheldon, B. C. (2015). The influence of nonrandom extra-pair paternity on heritability estimates derived from wild pedigrees. *Evolution*, *69*, 1336–1344. https://doi.org/10.1111/evo.12649

Fitzpatrick, J. L., & Baer, B. (2011). Polyandry reduces sperm length variation in social insects. *Evolution*, *65*, 3006–3012. https://doi.org/10.1111/j.1558-5646.2011.01343.x

Garcia-Gonzalez, F., Simmons, L. W., Tomkins, J. L., Kotiaho, J. S., & Evans, J. P. (2012). Comparing evolvabilities: Common errors surrounding the calculation and use of coefficients of additive genetic variation. *Evolution*, *66*, 2341–2349. https://doi.org/10.1111/j.1558-5646.2011.01565.x

Gasparini, C., Marino, I. A. M., Boschetto, C., & Pilastro, A. (2010). Effect of male age on sperm traits and sperm competition success in the

guppy (*Poecilia reticulata*). *Journal of Evolutionary Biology*, 23, 124–135. https://doi.org/10.1111/j.1420-9101.2009.01889.x

Girndt, A., Cockburn, G., Sánchéz-Trójar, A., Løvlie, H., & Schroeder, J. (2017). Method matters: Experimental evidence for shorter avian sperm in faecal compared to abdominal massage samples. *PLoS ONE*, 12(8), e0182853. https://doi.org/10.1371/journal.pone.0182853

Green, K. (2003). Age-related variation in mean sperm length, in the rove beetle *Aleochara bilineata*. *Journal of Insect Physiology*, 49, 993–998. https://doi.org/10.1016/S0022-1910(03)00164-1

Griffith, S. C., Holleley, C. E., Mariette, M. M., Pryke, S. R., & Svedin, N. (2010). Low level of extrapair parentage in wild zebra finches. *Animal Behavior*, 79(2), 261–264. https://doi.org/10.1016/j.anbehav.2009.11.031

Hansen, T. F., Pélabon, C., & Houle, D. (2011). Heritability is not evolvability. *Evolutionary Biology*, 38, 258–277. https://doi.org/10.1007/s11692-011-9127-6

Harris, W. E., Moore, A. J., & Moore, P. J. (2007). Variation in sperm size within and between ejaculates in a cockroach. *Functional Ecology*, 21, 598–602. https://doi.org/10.1111/j.1365-2435.2007.01248.x

Helfenstein, F., Szép, T., Nagy, Z., Kempenaers, B., & Wagner, R. H. (2008). Between-male variation in sperm size, velocity and longevity in sand martins *Riparia riparia*. *Journal of Avian Biology*, 39, 647–652. https://doi.org/10.1111/j.1600-048X.2008.04450.x

Hettyey, A., & Roberts, J. D. (2006). Sperm traits of the quacking frog, *Crinia georgiana*: Intra- and interpopulation variation in a species with a high risk of sperm competition. *Behavioral Ecology and Sociobiology*, 59, 389–396. https://doi.org/10.1007/s00265-005-0062-3

Houle, D. (1992). Comparing evolvability and variability of quantitative traits. *Genetics*, 130, 195–204.

Immler, S., Pitnick, S., Parker, G., Durrant, K., Lüpold, S., Calhim, S., et al. (2011). Resolving variation in the reproductive tradeoff between sperm size and number. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 5325–5330. https://doi.org/10.1073/pnas.1009059108

Immler, S., Pryke, S. R., Birkhead, T. R., & Griffith, S. C. (2010). Pronounced within-individual plasticity in sperm morphometry across social environments. *Evolution*, 64, 1634–1643. https://doi.org/10.1111/j.1558-5646.2009.00924.x

Jamieson, B. G. M. (1987). *The ultrastructure and phylogeny of insect spermatozoa*. Cambridge: Cambridge University Press.

Jamieson, B. G. M., (1991). *Fish evolution and systematics: Evidence from spermatozoa. With a survey of lophophorate, echinoderm and protochordate sperm and an account of gamete cryopreservation*. Cambridge: Cambridge University Press.

Kawakami, T., Smeds, L., Backström, N., Husby, A., Qvarnström, A., Mugal, C. F., et al. (2014). A high-density linkage map enables a second-generation collared flycatcher genome assembly and reveals the patterns of avian recombination rate variation and chromosomal evolution. *Molecular Ecology*, 23, 4035–4058. https://doi.org/10.1111/mec.12810

Kim, K.-W., Bennison, C., Hemmings, N., Brookes, L., Hurley, L. L., Griffith, S. C., et al. (2017). A sex-linked supergene controls sperm morphology and swimming speed in a songbird. *Nature Ecology & Evolution*, 1, 1168–1176. https://doi.org/10.1038/s41559-017-0235-2

Kleven, O., Laskemoen, T., Fossøy, F., Robertson, R. J., & Lifjeld, J. T. (2008). Intraspecific variation in sperm length is negatively related to sperm competition in passerine birds. *Evolution*, 62, 494–499. https://doi.org/10.1111/j.1558-5646.2007.00287.x

Knief, U., Forstmeier, W., Pei, Y., Ihle, M., Wang, D., Martin, K., et al. (2017). A sex-chromosome inversion causes strong overdominance for sperm traits that affect siring success. *Nature Ecology & Evolution*, 1, 1177–1184. https://doi.org/10.1038/s41559-017-0236-1

Krist, M. (2009). Short- and long-term effects of egg size and feeding frequency on offspring quality in the collared flycatcher (*Ficedula albicollis*). *Journal of Animal Ecology*, 78, 907–918. https://doi.org/10.1111/j.1365-2656.2009.01536.x

Krist, M., & Munclinger, P. (2011). Superiority of extra-pair offspring: Maternal but not genetic effects as revealed by a mixed cross-fostering design. *Molecular Ecology*, 20, 5074–5091. https://doi.org/10.1111/j.1365-294X.2011.05337.x

Krist, M., Nádvorník, P., Uvírová, L., & Bureš, S. (2005). Paternity covaries with laying and hatching order in the collared flycatcher *Ficedula albicollis*. *Behavioral Ecology and Sociobiology*, 59, 6–11. https://doi.org/10.1007/s00265-005-0002-2

Kruuk, L. E. (2004). Estimating genetic parameters in natural populations using the "animal model". *Philosophical Transactions of the Royal Society B: Biological Science*, 35, 873–890. https://doi.org/10.1098/rstb.2003.1437

LaMunyon, C. W., & Ward, S. (1998). Larger sperm outcompete smaller sperm in the nematode *Caenorhabditis elegans*. *Proceedings of the Royal Society of London B: Biological Sciences*, 265, 1997–2002. https://doi.org/10.1098/rspb.1998.0531

Laskemoen, T., Albrecht, T., Bonisoli-Alquati, A., Cepak, J., de Lope, F., Hermosell, I. G., et al. (2013). Variation in sperm morphometry and sperm competition among barn swallow (*Hirundo rustica*) populations. *Behavioral Ecology and Sociobiology*, 67, 301–309. https://doi.org/10.1007/s00265-012-1450-0

Laskemoen, T., Kleven, O., Fossøy, F., Robertson, R. J., Rudolfsen, G., & Lifjeld, J. T. (2010). Sperm quantity and quality effects on fertilization success in a highly promiscuous passerine, the tree swallow *Tachycineta bicolor*. *Behavioral Ecology and Sociobiology*, 64, 1473–1483. https://doi.org/10.1007/s00265-010-0962-8

Lifjeld, J. T., Laskemoen, T., Kleven, O., Albrecht, T., & Robertson, R. J. (2010). Sperm length variation as a predictor of extrapair paternity in passerine birds. *PLoS ONE*, 5, e13456. https://doi.org/10.1371/journal.pone.0013456

Lifjeld, J. T., Laskemoen, T., Kleven, O., Pedersen, A. T. M., Lampe, H. M., Rudolfsen, G., et al. (2012). No evidence for pre-copulatory sexual selection on sperm length in a passerine bird. *PLoS ONE*, 7, e32611. https://doi.org/10.1371/journal.pone.0032611

Lüpold, S., Birkhead, T. R., & Westneat, D. F. (2012). Seasonal variation in ejaculate traits of male red-winged blackbirds (*Agelaius phoeniceus*). *Behavioral Ecology and Sociobiology*, 66, 1607–1617. https://doi.org/10.1007/s00265-012-1415-3

Minoretti, N., & Baur, B. (2006). Among- and within-population variation in sperm quality in the simultaneously hermaphroditic land snail *Arianta arbustorum*. *Behavioral Ecology and Sociobiology*, 60, 270–280. https://doi.org/10.1007/s00265-006-0165-5

Minoretti, N., Stoll, P., & Baur, B. (2013). Heritability of sperm length and adult shell size in the land snail Arianta arbustorum (Linnaeus, 1758). *Journal of Molluscan Studies*, 79, 218–224. https://doi.org/10.1093/mollus/eyt012

Møller, A. P., Mousseau, T. A., Rudolfsen, G., Balbontín, J., Marzal, A., Hermosell, I., & De Lope, F. (2009). Senescent sperm performance in old male birds. *Journal of Evolutionary Biology*, 22, 334–344. https://doi.org/10.1111/j.1420-9101.2008.01650.x

Morrissey, M., (2014). Pedantics: functions to facilitate power and sensitivity analyses for genetic studies of natural populations. R package version 1.5. Retrieved from https://CRAN.R project.org/package=pedantics

Morrow, E. H., & Gage, M. J. G. (2001a). Artificial selection and heritability of sperm length in *Gryllus bimaculatus*. *Heredity*, 87, 356–362. https://doi.org/10.1046/j.1365-2540.2001.00921.x

Morrow, E. H., & Gage, M. J. G. (2001b). Consistent significant variation between individual males in spermatozoal morphometry. *Journal of Zoology*, 354, 147–153. https://doi.org/10.1017/S0952836901000656

Mossman, J., Slate, J., Humphries, S., & Birkhead, T. (2009). Sperm morphology and velocity are genetically

codetermined in the zebra finch. *Evolution*, *63*, 2730–2737. https://doi.org/10.1111/j.1558-5646.2009.00753.x

Mousseau, T. A., & Roff, D. A. (1987). Natural selection and the heritability of fitness components. *Heredity*, *59*, 181–197. https://doi.org/10.1038/hdy.1987.113

Napier, R. A. (1961). Fertility in the male rabbit. I. Sensitivity of spermatozoa to handling techniques. *Journal of Reproduction and Fertility*, *2*, 246–259. https://doi.org/10.1530/jrf.0.0020246

Parker, G. A. (1970). Sperm competition and its evolutionary effect on copula duration in the fly *Scatophaga stercoraria*. *Journal of Insect Physiology*, *16*, 1301–1328. https://doi.org/10.1016/0022-1910(70)90131-9

van de Pol, M., Brouwer, L., Ens, B. J., Oosterbeek, K., & Tinbergen, J. M. (2010). Fluctuating selection and the maintenance of individual and sex-specific diet specialization in free-living oystercatchers. *Evolution*, *64*, 836–851.

van de Pol, M., & Wright, J. (2009). A simple method for distinguishing within- versus between-subject effects using mixed models. *Animal Behavior*, *77*, 753–758.

Postma, E. (2014). Four decades of estimating heritabilities in wild vertebrate populations: Improved methods, more data, better estimates? In A. Charmantier, D. Garant, & L. E. B. Kruuk (Eds.), *Quantitative Genetics in the Wild*. Oxford: Oxford University Press.

Quay, W. B. (1986). Cloacal protuberance and cloacal sperm in passerine birds: Comparative study of quantitative relations. *Condor*, *88*, 160–168. https://doi.org/10.2307/1368911

R Core Team. 2018. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/.

Rowe, M., Swaddle, J. P., Pruett-Jones, S., & Webster, M. S. (2010). Plumage coloration, ejaculate quality and reproductive phenotype in the red-backed fairy-wren. *Animal Behavior*, *79*, 1239–1246. https://doi.org/10.1016/j.anbehav.2010.02.020

Šandera, M., Albrecht, T., & Stopka, P. (2013). Variation in apical hook length reflects the intensity of sperm competition in murine rodents. *PLoS ONE*, *8*, e68427. https://doi.org/10.1371/journal.pone.0068427

Schmoll, T., & Kleven, O. (2011). Sperm dimensions differ between two Coal Tit *Periparus ater* populations. *Journal of Ornithology*, *152*, 515–520. https://doi.org/10.1007/s10336-010-0603-z

Schmoll, T., , Kleven, O., , & Rusche, M., (2018). Individual phenotypic plasticity explains seasonal variation in sperm morphology in a passerine bird. *Evolutionary Ecology Research*, forthcoming.

Schmoll, T., Sanciprian, R., & Kleven, O. (2016). No evidence for effects of formalin storage duration or solvent medium exposure on avian sperm morphology. *Journal of Ornithology*, *157*, 647–652. https://doi.org/10.1007/s10336-015-1321-3

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, *9*, 671–675. https://doi.org/10.1038/nmeth.2089

Silva, C. N. S., McFarlane, S. E., Hagen, I. J., Rönnegård, L., Billing, A. M., Kvalnes, T., et al. (2017). Insights into the genetic architecture of morphological traits in two passerine bird species. *Heredity*, *119*, 197–205. https://doi.org/10.1038/hdy.2017.29

Simmons, L. W., & Kotiaho, J. S. (2002). Evolution of ejaculates: Patterns of phenotypic and genotypic variation and condition dependence in sperm competition traits. *Evolution*, *56*, 1622–1631. https://doi.org/10.1111/j.0014-3820.2002.tb01474.x

Simmons, L. W., Tomkins, J. L., & Hunt, J. (1999). Sperm competition games played by dimorphic male beetles. *Proceedings of the Royal Society B-Biological Sciences*, *266*, 145–150. https://doi.org/10.1098/rspb.1999.0614

Simpson, J. L., Humphries, S., Evans, J. P., Simmons, L. W., & Fitzpatrick, J. L. (2014). Relationships between sperm length and speed differ among three internally and three externally fertilizing species. *Evolution*, *68*, 92–104. https://doi.org/10.1111/evo.12199

Stoffel, M. A., Nakagawa, S., & Schielzeth, H. (2017). rptR: Repeatability estimation and variance decomposition by generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *8*, 1639–1644. https://doi.org/10.1111/2041-210X.12797

Teplitsky, C., Mills, J. A., Yarrall, J. W., & Merilä, J. (2009). Heritability of fitness components in a wild bird population. *Evolution*, *63*, 716–726. https://doi.org/10.1111/j.1558-5646.2008.00581.x

Tregenza, T., & Wedell, N. (1998). Benefits of multiple mates in the cricket *Gryllus bimaculatus*. *Evolution*, *52*, 1726–1730. https://doi.org/10.1111/j.1558-5646.1998.tb02252.x

Varea-Sánchez, M., Gómez Montoto, L., Tourmente, M., & Roldan, E. R. S. (2014). Postcopulatory sexual selection results in spermatozoa with more uniform head and flagellum sizes in rodents. *PLoS ONE*, *9*, e108148. https://doi.org/10.1371/journal.pone.0108148

Ward, P. I. (1998). Intraspecific variation in sperm size characters. *Heredity*, *80*, 655–659. https://doi.org/10.1046/j.1365-2540.1998.00401.x

Ward, P. I. (2000). Sperm length is heritable and sex-linked in the yellow dung fly (*Scatophaga stercoraria*). *Journal of Zoology*, *251*, 349–353. https://doi.org/10.1111/j.1469-7998.2000.tb01085.x

Wilson, A. J. (2008). Why h2 does not always equal VA/VP? *Journal of Evolutionary Biology*, *21*, 647–650. https://doi.org/10.1111/j.1420-9101.2008.01500.x

Wilson, A. J., Réale, D., Clements, M. N., Morrissey, M. M., Postma, E., Walling, C. A., et al. (2010). An ecologist's guide to the animal model. *Journal of Animal Ecology*, *79*, 13–26. https://doi.org/10.1111/j.1365-2656.2009.01639.x

Woolley, D. M. (1971). Selection for the length of the spermatozoan midpiece in the mouse. *Genetical Research*, *16*, 261–275.

Woolley, D. M., & Beatty, R. A. (1967). Inheritance of midpiece length in mouse spermatozoa. *Nature*, *215*, 94–95. https://doi.org/10.1038/215094a0

Zhang, X., & Hill, W. G. (2005). Genetic variability under mutation selection balance. *Trends in Ecology & Evolution*, *20*, 468–470. https://doi.org/10.1016/j.tree.2005.06.010

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Manuskript 18: Krist, M., Munclinger, P., Briedis, M. & Adamík, P. (2020) Genetic regulation of avian migration timing: combining candidate gene and quantitative genetic approaches in a long-distance migrant. Zasláno do časopisu Oecologia.

Genetic regulation of avian migration timing: combining candidate gene and quantitative genetic approaches in a long-distance migrant

Miloš Krist[1] · Pavel Munclinger[2] · Martins Briedis[1] · Peter Adamík[1]

1 Department of Zoology and Laboratory of Ornithology, Faculty of Science, Palacký University, Olomouc, Czech Republic

2 Department of Zoology, Faculty of Science, Charles University, Prague, Czech Republic

Correspondence

Miloš Krist, Department of Zoology and Laboratory of Ornithology, Faculty of Science, Palacký University, Olomouc, Czech Republic

E-mail: milos.krist@upol.cz

Present address

Martins Briedis, Department of Bird Migration, Swiss Ornithological Institute, Sempach, Switzerland & Lab of Ornithology, Institute of Biology, University of Latvia, Salaspils, Latvia

Author contributions

MK and PA designed the study, PM carried out the genotyping, MB extracted the data from geolocators, MK, MB and PA did the fieldwork, MK analysed the data and wrote the first draft, all authors contributed to revisions.

Abstract

Recent climatic change advanced the phenology of organisms across trophic levels, including insectivorous birds. However, the relative contribution of evolutionary and plastic responses to the observed phenotypic change has rarely been addressed. The distinction between these two scenarios is crucial since only genetic adaptation ensures long-term sustainability of populations under environmental change. Long-distance migrants have to rely on an inherited circannual pacemaker to time the start of their spring migration. Genetic regulation of their annual cycles can be studied either via candidate genes or through decomposition of phenotypic variance in phenological traits by quantitative genetics. Here we combined both approaches to study the timing of migration in a long-distance migrant, collared flycatcher (*Ficedula albicollis*). We found that none of the four studied candidate genes (*CLOCK, NPAS2, ADCYAP1 and CREB1*) had any consistent effect on the timing of six annual cycle stages of geolocator-tracked individuals. This negative result was confirmed for spring arrivals to the breeding site by direct observations of arriving males in four consecutive years. Although male spring arrival date was significantly repeatable (R = 0.24 ± 0.08), most of this value was attributable to permanent and common environmental effects while the additive genetic variance and heritability was very small (h^2 =

0.03 ± 0.17 SE). Such a low value constrains species evolutionary adaptation to climate change and thus our study appends to those warning that populations may be threatened by the ongoing climatic change.

Keywords

CLOCK, heritability, migration, repeatability, phenology

Introduction

Climatic change has shifted phenology of plants (Franks et al. 2013, Kolářová et al. 2014) and consequently of the whole trophic cascade (Thackeray et al. 2016, Cohen et al. 2018) including herbivorous insects (Kearney et al. 2010, van Asch et al. 2013), insectivorous birds (Gienapp et al. 2007, Usui et al. 2017) and raptors (Sullivan et al. 2016, Rosenfield et al. 2017). Thanks to the availability of long time series (e.g. Kolářová et al. 2017), birds became one the most popular taxons to study phenological shifts (Knudsen et al. 2011, Radchuk et al. 2019). Rising spring temperatures advanced most avian annual cycle events but the magnitude of this recent change differed between traits and species (Rubolini et al. 2007). Breeding advanced more than spring migration (Both and Visser 2001, Weidinger and Král 2007, Pearce-Higgins and Green 2014, Tomotani et al. 2018) and the latter advanced more in short-distance migrants compared to long-distance migrants (Végvári et al. 2010, Usui et al. 2017, Lehikoinen et al. 2019).

These observations suggest that phenological shifts are often driven by plastic responses of birds to the local temperature and vegetation phenology – environmental cues that are more available to sedentary species and short-distance migrants (Helm et al. 2013, Clark et al. 2014). In contrast, temperatures at the non-breeding sites of long-distance migrants are not very indicative of phenology at their breeding sites (but see Saino and Ambrosini 2008). Therefore, these species largely rely on their internal circannual pacemaker that tells them when to start the migration (Gwinner 1996, Helm et al. 2013, Åkesson et al. 2017). Long-distance migrants can also modify the speed of migration according to temperatures or rainfalls along their route (Ahola et al. 2004, Hüppop and Winkel 2006, Both 2010, see also Haest et al. 2018) but their arrival to breeding site is largely determined by the onset of migration (Stanley et al. 2012, Ouwehand and Both 2017, Briedis et al. 2019, Schmaljohann 2019). Thus, unlike rapid plastic responses by short-distance migrants, advanced phenology of long-distance migrants may be mainly driven by evolutionary changes of their internal clocks (Pulido 2007a, Schmaljohann 2019). Both plastic changes and evolutionary responses help birds to cope with the climate change. Plastic responses are faster and thus may serve as the first aid (Charmantier et al. 2008, Vedder et al. 2013) but because plasticity has its own costs and limits (DeWitt 1998, Auld et al. 2010), ultimately evolutionary response is needed for long-term sustainability of the population (Gienapp et al. 2013, Gienapp and Brommer 2014). Most of the observed shifts in migratory phenology are likely caused by a combination of both mechanisms (van Buskirk et al. 2012), but their relative contribution has rarely been addressed (Knudsen et al. 2011, Merilä and Hendry 2014).

A prerequisite for an evolutionary response to climate change is a sufficient amount of genetic variation upon which selection may act. Laboratory experiments have demonstrated rapid evolutionary response of the timing of migration to artificial selection (Pulido et al. 2001). However, such findings need to be verified in the wild where the environmental variation may mask genotypes from the action of selection. In other words, we need to know the amount of

additive genetic variation scaled by natural amount of phenotypic variation, i.e. heritability of the trait measured in the wild. Heritability of the breeding time was often studied in the wild and was found to vary both between and within populations (reviewed by Liedvogel et al. 2012). In contrast, only a handful of studies considered the heritability of the timing of migration. They usually found substantial heritability, suggesting high evolutionary potential of this phenological trait in wild populations (Potti 1998, Møller 2001, Teplitsky et al. 2011, Arnaud et al. 2013, Tarka et al. 2015). This was recently confirmed in a German population of the pied flycatcher (*Ficedula hypoleuca*) which considerably advanced its internal annual rhythm in as few as 21 years (Helm et al. 2019).

Although heritability is a useful measure for estimating the possible rate of adaptation to environmental change, it is silent about proximate mechanisms underlying this genetic adaptation. To reveal how the genetic variation is translated to differences between phenotypes, which are visible to selection, one needs to identify the involved genes and the function of their protein products. One possibility is to adopt candidate gene approach which looks for orthologs of genes with known function in genetic models like *Drosophila* or mouse in the studied non-model organism (Fitzpatrick et al. 2005). However, these genetic models are usually short-lived which prevents to study circannual rhythms on them. Fortunately, an emerging view suggests that circannual rhythms are closely linked to circadian rhythms, the latter having very well described molecular basis (Dunlap 1999, Bell-Pedersen 2005). Steinmeyer et al. (2009) found several circadian genes with allelic variation caused short repetitive sequence (microsatellite: *CLOCK, NPAS2, ADCYAP1, CREB1*) in the blue tit (*Cyanistes caeruleus*). Apart from their role in regulation of daily cycles, these genes are also the candidates for regulation of annual cycles (Visser et al. 2010, Liedvogel et al. 2011, Helm et al. 2013, Merlin and Liedvogel 2019).

Correlations between individual circadian genotype and timing of breeding (Liedvogel et al. 2009, Caprioli et al. 2012, Bourret and Garant 2015), moult (Saino et al. 2013), and migration (Mueller et al. 2011, Peterson et al. 2013, Bazzi et al. 2015, Ralston et al. 2019) have been found in several avian populations. In contrast, other studies did not find such within-population associations between individual genotype and timing of either migration (Contina et al. 2018, Parody-Merino et al. 2019) or breeding (Liedvogel and Sheldon 2010, Dor et al. 2012, Chakarov et al. 2013). While the former tests were usually not very powerful due to the difficulty to follow birds on migration, the large sample sizes of the latter studies mean that their negative findings are more robust. Thus, the evidence for the effects of circadian genes on avian phenology is mixed and therefore this topic warrants further study (Ralston et al. 2019).

In this study we tested if there is any association between the genotype at four candidate genes (*CLOCK, NPAS2, ADCYAP1, CREB1*) and the phenology of migration in the collared flycatcher (*Ficedula albicollis*), a long-distance nocturnal migrant. First, we tracked flycatchers throughout their whole annual cycle using geolocators. Second, we directly recorded spring arrival of males to our breeding sites in four consecutive years which provided larger sample size for this phenological stage. Third, this large sample size combined with the known pedigree of our population enabled us to calculate heritability and thus the potential for evolutionary changes in male spring arrival date.

Methods

Study populations and fieldwork

We conducted this study in two close nest-box populations of the collared flycatcher (Dlouhá Loučka: 49°50′N, 17°30′E, 340–500 m asl and Velký Kosíř: 49°32′N, 17°03′E, 300–400 m asl).

Collared flycatcher is a long-distance nocturnal migrant with non-breeding residency areas in central and southern Africa, about 7000 km apart from the central European breeding sites (Briedis et al. 2016). It easily adopts nest-boxes for breeding; each of our populations hosts about 100 pairs of collared flycatchers. Other species breeding in our nest-boxes are great tit (*Parus major*), blue tit (*Cyanistes caeruleus*) and nuthatch (*Sitta europea*). Nest-boxes are attached on trees about 160 cm above ground and have inner dimensions of 22.5–25.5 x 11 x 12 cm (height x width x depth) and diameter of the nest entrance of 32 mm.

We caught collared flycatcher males at the time of their spring arrival to Velký Kosíř in 2013–2016. In this period, 381 nest-boxes were located at five plots in an oak (*Quercus petraea*) forest. We simultaneously trapped males in all these nest-boxes. Each year 11–15 catching sessions (days) were carried out. At each trapping session, the team of ca 10 field assistants activated string traps in all nest-boxes that did not contain any nest or that contained a nest without any progress in nest-building from the last nest-box check. We monitored the study area several times in a week following the first record of the collared flycatcher in the database for birdwatchers in the Czech Republic (https://birds.cz/avif). The day after we had recorded the first singing male in our study area, we started to catch flycatchers (12th–16th April in the four years). Catching lasted about a month with the last session between 15th–19th May. We started to catch flycatchers at 7–8 a.m. and finished about at 14–15 p.m. We planned the subsequent session according to the number of newly arrived males. When many new males arrived, the next session was carried out the next day or within a few days while the interval was prolonged when only a few new males arrived (see Dryad Digital Repository for the terms of all catching sessions). We started to catch flycatchers in most nest-boxes as they were empty at the start of the season. The number of nest-boxes available for catching steadily declined as the season progressed and tits and flycatchers started to breed. Each newly arrived male was brought to the field station placed in the centre of the study area for measurements, ringing and blood sampling that was done by tibial venipuncture. Males were also aged based on their plumage characteristics (second year of life or older) and then either brought back to the plot where they were caught (2013) or released directly from the field station (2014–2016).

Arrival date of each male was calculated as the mean date between the first capture and the previous catching session. We validated this method on 16 males for whom we had both catching data and the true arrival data from geolocators in 2016 (see below for methods of geolocation). Our estimate of arrival date correlated strongly with the true arrival date (r = 0.95, n = 16, Figure 1a). Further option how to check our ability to catch males soon after their arrival is to compare our estimates of arrival date to laying dates in nests attended by them. In total, 281 out of 520 males caught on arrival were found breeding in our nest-boxes. For 18 of these the estimated arrival date was later than laying date of the first egg (i.e. negative mating speed, mean = -3.28, range: -0.5 to -15.5, Figure 1b). This happens sometimes due to mate replacement, i.e. adoption of the widowed female and the nest by a newly arrived male (see Lifjeld et al. 1997, Sheldon et al. 1999), but we consider most of these cases to be caused by our inability to catch males before they attracted their primary female. After the primary female started to build a nest, we stopped catching in that nest-box and the males could only be caught if they tried to attract a secondary female. Therefore, we excluded these presumed inaccuracies from the dataset. If we assumed similarly low frequency of ca 6% (18/281*100) of these inaccuracies in the rest of the data for which we cannot make this validation since the arriving males did not breed at our plots (i.e. 520 - 281 = 239 arrivals), we might have about 15 such cases in our sample of 502 analyzed arrivals. Together with the good congruence between our estimate of arrival date and the true arrival date from geolocators, this low frequency of inaccurate values suggests that our estimates captured most of the phenotypic variation in the timing of male arrival.

Geolocator data

We used geolocators to study timing of 6 migratory stages (departure from the breeding site, Sahara crossing in autumn, arrival to the non-breeding site, departure from the non-breeding site, Sahara crossing in spring and arrival to the breeding site) of collared flycatchers. We deployed 69 geolocators (model GDL2.0 with 7mm light stalk, Swiss Ornithological Institute) on 33 adult males and 36 adult females at Dlouhá Loučka when they cared of their broods during the late nestling stage in 2013. At Velký Kosíř, we equipped adults with geolocators upon their spring arrival (139 males) or during nestling stage (18 males and 8 females) in 2014. We retrieved 29 geolocators in Dlouhá Loučka and 30 at Velký Kosíř in following years.

Geolocators recorded light intensity on an arbitrary scale ranging from 0 to 63 units, corresponding to 0 and approximately 3500 lux, respectively. We used a threshold approach (Lisovski et al. 2020) to estimate individual migration timing from the recorded light data. First, we identified sunrise and sunset times within the light-level data using 'GeoLocator' software (Swiss Ornithological Institute) and setting the light intensity threshold to 1 unit on the arbitrary scale. Second, we determined the start and end of migratory periods within each dataset using the '*changeLight*' function (parameters: q = 0.85 and minimum duration of stationary/stopover periods = 3 days) of the R-package 'GeoLight' v 1.03 (Lisovski and Hahn 2012). Timing of Sahara crossing was determined by manual inspection of raw daily light recordings to identify days when prolonged periods of uninterrupted maximal light intensities (63 arbitrary units) were recorded. Such patterns of uninterrupted maximal light recordings are conspicuous in the raw light-level data and are characteristic to non-stop diurnal flight when birds cross large ecological barriers, like the Sahara Desert (Adamík et al. 2016, including data from the collared flycatcher). We used the first day of Sahara crossing (the total duration of the crossing is 2-3 days in most cases) in all calculations.

Candidate genes

DNA was extracted from blood using DNeasy (r) Blood & Tissue kit (Qiagen). We genotyped flycatchers at four candidate genes (*CLOCK*, *NPAS2*, *ADCYAP1* and *CREB1*) for circadian/circannual cycles (Steinmeyer et al. 2009). Primers for amplification of *NPAS2* and *ADCYAP1* were the same as in Steinmeyer et al. (2009). *CREB1* was amplified using forward primer from Steinmeyer et al. (2009) and modified reverse primer AGAATAACGCAGCCCAGAGC from Bourret and Garant (2015). *CLOCK* primers were adopted from Caprioli et al. (2012). Forward primers for amplification of *CLOCK, NPAS2, ADCYAP1* and *CREB1* were labelled with the dyes 6FAM, VIC, PET and NED, respectively. All loci were amplified in a single multiplex PCR using Type-it® Microsatellite PCR kit (Qiagen) following manufacturer's protocols. Annealing temperature was 53 °C. PCR products were mixed with GeneScanTM–500 LIZ® Size Standard (Applied Biosystems) and their size was resolved using fragment analysis in 3130xl Genetic Analyzer (Applied Biosystems). Genotypes were scored with the GeneMarker® 1.9 (Softgenetics).

Pedigree

At the Velký Kosíř site most nestlings and adults were ringed each year since 1998. This provides an extensive social pedigree for this population. In some years, paternity analysis was conducted on parts of the population (2001–2002, Krist et al. 2005), (2006–2009, Krist and Munclinger 2011) or on the whole population (2013, Edme et al. 2017). There was no case of maternal error, so all maternal links in the pedigree should be accurate. In total, extra-pair paternity was detected at a rate of 23.5%. We corrected the social pedigree by the genetic information which was available

for 18% of the offspring born between 1998–2015 (Edme et al. 2019). Consequently, after this correction, probably less than 20% of paternal links remained erroneous, giving a total of less than 10% errors in the full pedigree (maternal and paternal links together). This inadequacy should not have excessive impact on our ability to estimate quantitative genetic parameters since pedigree-based animal models are robust to even higher rates of paternity errors (Charmantier and Réale 2005; see also Firth et al. 2015).

Data analyses

Allele frequencies were calculated in Cervus 3.0 (Kalinowski et al. 2007). All other statistical analyses were done in R 3.6.2 (R Core Team 2019). The relationships between genotypes at three loci (*CLOCK*, *ADCYAP1* and *CREB1*) and timing of the six core phases of the annual cycle inferred from geolocators were assessed and visualized in 18 (3x6) separate linear models with day of year of the stage ($1=1^{th}$ January) as the dependent and mean allele length as the independent variable. Three other independent variables were included in each model to reduce residual variation in day of year (locality: Kosíř vs. Loučka, sex: male vs. female and age at deployment: first year vs. second year or older). Male age was determined from plumage characteristics while that of females from ring data because all females included in this study were ringed previously. We did not test for the effect of the *NPAS2* gene because of its low allelic diversity (Table 1) and small sample size of tracked individuals. We used only one (mean allele length) of several possible genetic variables (mean, maximal and minimal allele length) for each relationship to keep the number of statistical tests reasonably small. Mean allele length has been most often used for similar investigations because variation in complex traits such as behaviour is usually more affected by additive compared to dominant or epistatic effects (Hill et al. 2008, Wolak and Keller 2014) and mean allele length encompasses information about both the shorter and longer allele.

Direct observations of male spring arrivals to the breeding site were conducted in four consecutive years. To test for the potential effects of candidate genes, we first averaged values of repeatedly sampled individuals. Before averaging, year and age effects were removed from the data by the means of using residuals from an ANOVA of arrival date on these two factors (Supplementary Table 1). Similarly as for geolocator data, average residual date was then predicted by the mean length of allele for three loci (*CLOCK, ADCYAP1* and *CREB1*) or by exact genotype (categorical factor) in the case of less variable *NPAS2* locus. Contribution of each datapoint (individual mean arrival date) was weighted by the number of observations (years) used for its calculation.

Phenotypic variance in male spring arrival date from direct observations was decomposed to its causal components by fitting an animal model in the R-package Asreml-R, version 3 (Butler 2009). Male spring arrival date was the dependent variable, age (second year or older) – fixed independent variable. The random part of the model included four random effects: year, male identity not linked to the pedigree, male identity linked to the pedigree, and identity of the rearing nest. Male identity linked to the pedigree estimates the additive genetic component, male identity not linked to the pedigree estimates permanent environment component, and identity of the nest estimates common environment effects (Wilson et al. 2010). Before entering the animal model, pedigree was pruned to contain only informative individuals (i.e. those that were either phenotyped or linked to at least two phenotyped individuals) in the R-package Nadiv (Wolak 2012). Pruned pedigree retained 1048 phenotyped individuals and their relatives. Repeatabilities and their standard errors were calculated in the Rptr package (Stoffel et al. 2017).

Results

In total, 407 individuals that were either geolocated or had a record of their spring arrival to the Velký Kosíř breeding site were genotyped at the four candidate genes. *CLOCK* alleles varied by three nucleotides. This fits with the previous studies showing that this gene is variable due to CAG repetition on its 5' end resulting in polyglutamine chain (polyQ) at the carboxyl end of the coded Clock protein. We also confirmed this by sequencing of *CLOCK* homozygotes with the length of alleles 120bp, 123bp, 126bp and 129bp. These alleles coded 10–13 glutamins and therefore can be labelled as Q_{10}–Q_{13}. *ADCYAP1* and *CREB1* were more variable compared to *CLOCK,* while *NPAS2* was the most conservative locus (Table 1). There was an excess of homozygotes at *CREB1* locus which together with the highest rate of amplification failure among loci probably suggests presence of null alleles. However, we cannot rule out an alternative explanation based on underdominance.

We caught 59 individuals equipped with geolocators in years following geolocator deployment but our final sample size for the timing of their annual cycle events varied from 26 (arrival to breeding site) to 41 (Sahara crossing in autumn and arrival to non-breeding site) due to failure rate of the devices. As has commonly been observed (Briedis et al. 2019), males were ahead of females in most stages of the annual cycle (Supplementary Table 2), although this difference was not statistically significant due to the bias of our sample toward males. Older individuals were somewhat delayed when compared to younger ones (Supplementary Table 2) but note that also young birds have already been experienced and we recorded their second migration to and back from Africa via geolocators. Site had the strongest effect on the timing of migration. Birds from Loučka population migrated before those from Kosíř (Supplementary Table 2). However, this might be also due to the difference between years since we tracked migration during 2013/2014 season in birds from Loučka but during 2014/2015 in birds from Kosíř. In contrast to these consistent effects of covariates, the effects of candidate genes on the timing of annual events were inconsistent and weak. We found only one significant relationship (p<0.05) but as this came from a sample of 18 related and somewhat dependent tests and the effect was not strong enough to be significant after accounting for multiple testing (Supplementary Table 2, Figure 2), we consider this observation likely to be a sampling error.

Catching of arriving males at the Kosíř breeding site revealed that this stage of migration was delayed by about a week in yearlings compared to older males (Supplementary Table 1). Note that spring arrival to the breeding site was the final stage of the first migration of yearling males, which is in contrast to the data from geolocators that covered the second migration of these young birds (see above). Mean timing of spring arrival varied only slightly (up to five days, Supplementary Table 1) in the four studied years, although this result was statistically significant. In contrast, male spring arrival was unrelated to mean allele length or genotype on any of the candidate loci [*CLOCK*: -0.29 ± 0.24 (estimate ± SE), $F_{1,368} = 1.40$, p = 0.237; *NPAS2*: $F_{3,366} = 0.83$, p = 0.476; *ADCYAP1*: 0.30 ± 0.28, $F_{1,367} = 1.13$, p = 0.288, *CREB1*: -0.03 ± 0.10, $F_{1,352} = 0.09$, p = 0.761; Figure 3].

In total, we recorded 502 arrivals of 372 males to Velký Kosíř breeding site by the method of their catching in nest-boxes. 266 males were caught only in one year but we had repeated records for 106 individuals (86 were caught in two years, 16 in three years and four in all four years of the study). These males were significantly repeatable in the timing of their spring arrival (r = 0.24 ± 0.08 SE, Figure 4). We further decomposed the variance in male spring arrival to its causal components with the animal model. We found the within-individual repeatability to be explained mainly by permanent and common environmental effects, while the additive genetic component was very low. Consequently, our estimate of heritability of spring arrival was also small and insignificant (Table 2).

Discussion

We did not detect any consistent relationships between genetic variation at four circadian genes and timing of any phase of the annual cycle in geolocated collared flycatchers. These negative findings, however, should to be treated with caution as they are based on tens of individuals and thus the tests have small statistical power. However, we found the same negative result in the dataset with hundreds of individuals for which we directly observed their spring arrivals to the breeding site. Taken together, our results suggest that none of the circadian genes play an important role in circannual rhythmicity in the focal migratory species. Furthermore, our quantitative genetic analysis revealed moderate repeatability (R = 0.24 ± 0.08 SE) and very small and insignificant heritability (h^2 = 0.03 ± 0.17 SE) of the timing of spring arrival to the breeding site. That suggests limited genetic variation and, therefore, low evolutionary potential of this important phenological trait.

Previous studies sometimes confirmed associations between individual circadian genotype and phenology, for instance in timing of breeding (Liedvogel et al. 2009, Caprioli et al. 2012), dispersal (Chakarov et al. 2013), migration (Mueller et al. 2011, Bazzi et al. 2015, Saino et al. 2017, Ralston et al. 2019) or moult (Saino et al. 2013, Bazzi et al. 2017). However, many other studies did not find any such relationships (Dor et al. 2011, Liedvogel et al. 2012, Peterson et al. 2013, Contina et al. 2018, Romano et al. 2018, Parody-Merino et al. 2019, this study), or found them only in an interaction with environmental variables, such as breeding density (Bourret and Garant 2015). Many of these studies were based on small sample sizes because of the difficulty to obtain phenological data throughout the entire avian annual cycle (e.g. by tracking individuals using geolocators). Moreover, even those studies that reported significant associations between circadian genotypes and phenology usually did this only for some of the tested loci and only for a subset of the genetic traits that were considered in these studies as the group (mean allele length, length of longer allele, length of shorter allele, heterozygosity, exact genotype and methylation level). Taken together, the evidence for the causal link between the timing of daily and annual cycles is relatively weak (reviewed by Ralston et al. 2019, Parody-Merino et al. 2019). This is also in line with a recent study in the great tit where genomic selection caused phenological shifts without any correlated change in the endogenous daily cycles (Verhagen et al. 2019). This indicates that these two cycles are not genetically correlated (Verhagen et al. 2019) and circadian genes cannot explain much of the phenological variation. However, migratory traits often have substantial genetic variation and heritability (Pulido 2007a). Consequently, there is a need to look for other loci that would explain variation in the key annual cycle events (Contina et al. 2018).

This is achievable currently thanks to the development of the next-generation sequencing methods (Stapley et al. 2010). Whole genome sequencing and high-density single-nucleotide polymorphism (SNP) chips can reveal sites of divergent selection and thus allelic variation between migratory phenotypes (Liedvogel et al. 2009, Merlin and Liedvogel 2019). Variation in migratory phenotypes may also be caused by the differences in gene expression. This can be tested via transcriptomic or epigenomic analyses (Stapley et al. 2010, Lafuente and Beldade 2019, Merlin and Liedvogel 2019). All these genome-wide association studies (GWAS) can reveal loci correlated with migratory phenotypes but they also have their own limitations.

GWAS often identified large genomic regions with hundreds of loci under divergent selection (Zhan et al. 2014, Lundberg et al. 2017). Many of these loci may have no causal effect on the studied phenotype. Instead, they may be linked only to another locus with an effect. Moreover, none of the loci may achieve genome-wide significance unless very large sample size is used which regularly led to false negatives in GWAS (Ioannidis et al. 2011). This conservative nature of GWAS is in contrast to the candidate gene approach that is very prone to falsely positive

results (Ioannidis et al. 2011, Kim et al. 2018). These opposite sources of errors – lack of the power in genome-wide studies vs. bias in candidate gene studies – may explain why GWAS often do not confirm expectations derived from candidate gene studies. For example, recent GWAS in the great tit did not find any association of exploration behavior with locus *DRD4* that has previously been considered to causally affect this behavior (Kim et al. 2018). Similarly, SNPs most associated with circadian period (tau) were located outside of the 12 previously considered candidate loci (Laine et al. 2019).

Further progress may be accomplished by the development of denser SNP chips or employment of whole genome-sequencing. Such high-resolution methods will enable to locate genomic regions with causal effect on the studied phenotype more precisely and thus limit the number of candidate genes to only those most promising (Delmore et al. 2015, Kim et al. 2018). This was recently done in a study of *Vermivora* warblers whose migratory phenotypes correlated with SNPs at a new candidate gene VPS 13A (Toews et al. 2019). The function of this gene is unknown in birds but authors hypothesize that it may be involved in processing of metabolic products that arise when the warbler is migrating (Toews et al. 2019). Similarly, migratory phenotypes of willow warbler were most divergent at loci related to metabolism of fatty acids (Lundberg et al. 2017). It is important to bear in the mind that different migratory phenotypes came from different populations across migratory divide. Thus they represented suite of intercorrelated traits such as different breeding and non-breeding area in addition to different migratory patterns in space and time. It is therefore not clear whether these candidate metabolic genes may also affect timing of avian migration. Nevertheless, a support for this idea is a recent finding that laying date of the great tit was most strongly associated to another gene controlling metabolism, thyroglobulin, within a breeding population (Gienapp et al. 2017). Therefore, it may be a time now to broaden our view of potential candidates in avian phenology from circadian genes only to those affecting physiology and metabolism.

Technological advances in tracking techniques like GPS and light-level geolocators provided the opportunity to repeatedly follow even small migrants through their whole annual cycle without severely compromising their fitness (Brlík et al. 2020). Repeated tracking usually revealed high individual consistency in the timing of migration. This suggests that the start of migration is controlled by an innate program (Alerstam et al. 2006, Lourenço et al. 2011, Vardanis et al. 2011, Stanley et al. 2012, Conklin et al. 2013, López-López et al. 2014, van Wijk et al. 2016). In contrast, birds were rather flexible in selection of the route (Alerstam et al. 2006, Vardanis et al. 2011, Stanley et al. 2012, López-López et al. 2014), although also an opposite pattern of consistent route and flexible timing has been found (Hasselquist et al. 2017). Repeatable timing of spring migration may be caused by individual consistency in the time of departure from the non-breeding site (Both et al. 2016). In such a case we would expect decrease of repeatability along the route with the smallest values at the time of arrival to the breeding ground (Both et al. 2016) as the progress of migration is dependent on environmental conditions en route and these often fluctuate between years (Both 2010, Briedis et al. 2017). Smaller repeatability of the timing of migration near the final destination have been found in many studies (Alerstam et al. 2006, Lourenço et al. 2011, Vardanis et al. 2011, Sergio et al. 2014, Fraser et al. 2019, but see van Wijk et al. 2016). However, the absolute difference was often small and so the evidence for the vanishing repeatability is not particularly strong. Caution is also needed when comparing repeatability between studies (Both et al. 2016). For example, smaller reported repeatability at the arrival to the breeding site compared to those at the departure from the non-breeding site might be caused by the fact that the latter values are typically obtained on tracked birds while the former is most often derived by direct observations that are prone to larger measurement errors. On the other hand, studies using direct observations often provided much larger sample size than those of tracked birds (review in Both et al. 2016).

Our estimate of repeatability of male spring arrival (r = 0.24 ± 0.08 SE) broadly fits within the values obtained by direct observations at breeding sites (Pulido 2007b, Both et al. 2016) and is actually very close to that of male pied flycatchers (*Ficedula hypoleuca*, r = 0.27 ± 0.03, Both et al. 2016). As repeatability is usually the upper limit to heritability (Falconer 1993, but see Dohm 2002) both these studies suggest rather limited evolutionary potential of arrival date in *Ficedula* flycatchers. We further decomposed the variance in arrival timing explained by male identity to its causal components and found very low additive genetic variance and heritability (h^2 = 0.03 ± 0.17 SE) and larger but still insignificant common and permanent environment components. Thus, the repeatable timing of the same individual is not much controlled by its genotype. Instead it is possible that young birds learn how to migrate at their first migration and then repeat this successful tactic in following years, a mechanism suggested also for common terns (*Sterna hirundo*) and black kites (*Milvus migrans*; Arnaud et al. 2013, Sergio et al. 2014). Alternatively, timing of migration may also be somewhat pre-determined by early condition (e.g. Pulido 2007b) as suggested by non-zero common environment component. However, these conclusions have to be treated with caution since our heritability estimate has wide standard errors despite having more than a thousand of informative individuals in our pedigree. The difficulty to determine heritability in the wild with a high precision is mirrored in a scarcity of quantitative genetic studies of avian migration. We know of only five such studies (Potti 1998, Møller 2001, Teplitsky et al. 2011, Arnaud et al. 2013, Tarka et al. 2015). All of them tested the heritability of avian spring arrival to breeding grounds, as we did, and found the heritability ranging from 0.10 to 0.54. Thus, our heritability estimate is so far the lowest one. However, much more studies are needed to see whether the different heritability estimates have something to do with the biology of the studied species or if they are merely a product of sampling variation in space and time.

The low evolutionary potential in phenological traits, as the one we found for spring arrival of collared flycatchers, would severely limit the rate of adaptive evolution and thus may have negative consequences for populations currently experiencing rapid climate change. On the other hand, low consistency of individuals in their spring arrival suggests high plasticity in their decisions when to arrive at the breeding locality and this variation would also be adaptive if corresponded to the local phenology of vegetation. Unfortunately, we were unable to test for this scenario because our study spanned only four years. In general, phenotypic plasticity indeed helps populations to persist under climate change (Charmantier et al. 2008, Gienapp et al. 2013). However, the length of this persistence depends on the costs of phenotypic plasticity (Chevin et al. 2010) and species life history (Vedder et al. 2013). Adaptive evolution is necessary for long-term persistence of populations under directional selection caused by global warming (Gienapp et al. 2013, Radchuk et al. 2019). The evolved responses enable animals to appropriately time their life cycle events without the need to maintain costly physiological machinery that is necessary for a high phenotypic plasticity (DeWitt et al. 1998, Auld et al. 2010). In addition, the cues upon which the birds plastically react may be available to resident species or short-distance migrants, but not to long-distance migrants (Both and Visser 2001). Consequently, long-distance migrants may be unable to plastically adjust their spring arrival to match their breeding to the peak of food supply (Both et al. 2006) which leads to population decline (Møller et al. 2008, Both et al. 2010, Koleček et al. 2020) and increased risk of extinction (Radchuk et al. 2019). The application of methods of quantitative genetics to a broader spectrum of species differing in life-histories, e.g. migration distance or generation time, would shed useful light on the mechanisms by which birds may adapt their phenology to the ongoing climatic change.

Acknowledgements

Data accessibility

All R-code and data files used for analyses presented in this manuscript are available on the Dryad Digital Repository

References

Adamik P et al. (2016) Barrier crossing in small avian migrants: individual tracking reveals prolonged nocturnal flights into the day as a common migratory strategy. Scientific Reports 6. doi: 10.1038/srep21560

Ahola M, Laaksonen T, Sippola K, Eeva T, Rainio K, Lehikoinen E (2004) Variation in climate warming along the migration route uncouples arrival and breeding dates. Global Change Biology 10:1610-1617. doi: 10.1111/j.1365-2486.2004.00823.x

Akesson S, Ilieva M, Karagicheva J, Rakhimberdiev E, Tomotani B, Helm B (2017) Timing avian long-distance migration: from internal clock mechanisms to global flights. Philosophical Transactions of the Royal Society B-Biological Sciences 372:16. doi: 10.1098/rstb.2016.0252

Alerstam T, Hake M, Kjellen N (2006) Temporal and spatial patterns of repeated migratory journeys by ospreys. Animal Behaviour 71:555-566. doi: 10.1016/j.anbehav.2005.05.016

Arnaud CM, Becker PH, Dobson FS, Charmantier A (2013) Canalization of phenology in common terns: genetic and phenotypic variations in spring arrival date. Behavioral Ecology 24:683-690. doi: 10.1093/beheco/ars214

Auld JR, Agrawal AA, Relyea RA (2010) Re-evaluating the costs and limits of adaptive phenotypic plasticity. Proceedings of the Royal Society B-Biological Sciences 277:503-511. doi: 10.1098/rspb.2009.1355

Bazzi G et al. (2015) Clock gene polymorphism and scheduling of migration: a geolocator study of the barn swallow Hirundo rustica. Scientific Reports 5. doi: 10.1038/srep12443

Bazzi G et al. (2017) Candidate genes have sex-specific effects on timing of spring migration and moult speed in a long-distance migratory bird. Current Zoology 63:479-486. doi: 10.1093/cz/zow103

Bell-Pedersen D et al. (2005) Circadian rhythms from multiple oscillators: Lessons from diverse organisms. Nature Reviews Genetics 6:544-556. doi: 10.1038/nrg1633

Both C (2010) Flexibility of Timing of Avian Migration to Climate Change Masked by Environmental Constraints En Route. Current Biology 20:243-248. doi: 10.1016/j.cub.2009.11.074

Both C, Bijlsma RG, Ouwehand J (2016) Repeatability in spring arrival dates in Pied Flycatchers varies among

years and sexes. Ardea 104:3-21. doi: 10.5253/arde.v104i1.a1

Both C, Bouwhuis S, Lessells CM, Visser ME (2006) Climate change and population declines in a long-distance migratory bird. Nature 441:81-83. doi: 10.1038/nature04539

Both C, Van Turnhout CAM, Bijlsma RG, Siepel H, Van Strien AJ, Foppen RPB (2010) Avian population consequences of climate change are most severe for long-distance migrants in seasonal habitats. Proceedings of the Royal Society B-Biological Sciences 277:1259-1266. doi: 10.1098/rspb.2009.1525

Both C, Visser ME (2001) Adjustment to climate change is constrained by arrival date in a long-distance migrant bird. Nature 411:296-298. doi: 10.1038/35077063

Bourret A, Garant D (2015) Candidate gene-environment interactions and their relationships with timing of breeding in a wild bird population. Ecology and Evolution 5:3628-3641. doi: 10.1002/ece3.1630

Briedis M et al. (2019) A full annual perspective on sex-biased migration timing in long-distance migratory birds. Proceedings of the Royal Society B-Biological Sciences 286. doi: 10.1098/rspb.2018.2821

Briedis M, Hahn S, Adamik P (2017) Cold spell en route delays spring arrival and decreases apparent survival in a long-distance migratory songbird. Bmc Ecology 17. doi: 10.1186/s12898-017-0121-4

Briedis M et al. (2016) Breeding latitude leads to different temporal but not spatial organization of the annual cycle in a long-distance migrant. Journal of Avian Biology 47:743-748. doi: 10.1111/jav.01002

Briedis M, Hahn S, Krist M, Adamik P (2018a) Finish with a sprint: Evidence for time-selected last leg of migration in a long-distance migratory songbird. Ecology and Evolution 8:6899-6908. doi: 10.1002/ece3.4206

Briedis M, Krist M, Kral M, Voigt CC, Adamik P (2018b) Linking events throughout the annual cycle in a migratory bird-non-breeding period buffers accumulation of carry-over effects. Behavioral Ecology and Sociobiology 72. doi: 10.1007/s00265-018-2509-3

Brlik V et al. (2020) Weak effects of geolocators on small birds: A meta-analysis controlled for phylogeny and publication bias. Journal of Animal Ecology 89:207-220. doi: 10.1111/1365-2656.12962

Butler D (2009) asreml() fits the linear mixed model. R package version 3.0

Caprioli M et al. (2012) Clock Gene Variation Is Associated with Breeding Phenology and Maybe under Directional Selection in the Migratory Barn Swallow. Plos One 7. doi: 10.1371/journal.pone.0035140

Chakarov N, Jonker RM, Boerner M, Hoffman JI, Kruger O (2013) Variation at phenological candidate genes correlates with timing of dispersal and plumage morph in a sedentary bird of prey. Molecular Ecology 22:5430-5440. doi: 10.1111/mec.12493

Charmantier A, Gienapp P (2014) Climate change and timing of avian breeding and migration: evolutionary versus plastic changes. Evolutionary Applications 7:15-28. doi: 10.1111/eva.12126

Charmantier A, McCleery RH, Cole LR, Perrins C, Kruuk LEB, Sheldon BC (2008) Adaptive phenotypic plasticity in response to climate change in a wild bird population. Science 320:800-803. doi: 10.1126/science.1157174

Charmantier A, Reale D (2005) How do misassigned paternities affect the estimation of heritability in the wild? Molecular Ecology 14:2839-2850. doi: 10.1111/j.1365-294X.2005.02619.x

Chevin LM, Lande R, Mace GM (2010) Adaptation, Plasticity, and Extinction in a Changing Environment: Towards a Predictive Theory. Plos Biology 8:8. doi: 10.1371/journal.pbio.1000357

Clark RG, Poysa H, Runko P, Paasivaara A (2014) Spring phenology and timing of breeding in short-distance migrant birds: phenotypic responses and offspring recruitment patterns in common goldeneyes. Journal of Avian Biology 45:457-465. doi: 10.1111/jav.00290

Cohen JM, Lajeunesse MJ, Rohr JR (2018) A global synthesis of animal phenological responses to climate change. Nature Climate Change 8:224-+. doi: 10.1038/s41558-018-0067-3

Conklin JR, Battley PF, Potter MA (2013) Absolute Consistency: Individual versus Population Variation in Annual-Cycle Schedules of a Long-Distance Migrant Bird. Plos One 8:9. doi:

10.1371/journal.pone.0054535

Contina A, Bridge ES, Ross JD, Shipley JR, Kelly JF (2018) Examination of Clock and Adcyap1 gene variation in a neotropical migratory passerine. Plos One 13. doi: 10.1371/journal.pone.0190859

Delmore KE et al. (2015) Genomic analysis of a migratory divide reveals candidate genes for migration and implicates selective sweeps in generating islands of differentiation. Molecular Ecology 24:1873-1888. doi: 10.1111/mec.13150

DeWitt TJ, Sih A, Wilson DS (1998) Costs and limits of phenotypic plasticity. Trends in Ecology & Evolution 13:77-81. doi: 10.1016/s0169-5347(97)01274-3

Dohm MR (2002) Repeatability estimates do not always set an upper limit to heritability. Functional Ecology 16:273-280

Dor R et al. (2012) Clock gene variation in Tachycineta swallows. Ecology and Evolution 2:95-105. doi: 10.1002/ece3.73

Dunlap JC (1999) Molecular bases for circadian clocks. Cell 96:271-290. doi: 10.1016/s0092-8674(00)80566-8

Edme A, Zobac P, Korsten P, Albrecht T, Schmoll T, Krist M (2019) Moderate heritability and low evolvability of sperm morphology in a species with high risk of sperm competition, the collared flycatcher Ficedula albicollis. Journal of Evolutionary Biology 32:205-217. doi: 10.1111/jeb.13404

Edme A et al. (2017) Do ornaments, arrival date, and sperm size influence mating and paternity success in the collared flycatcher? Behavioral Ecology and Sociobiology 71. doi: 10.1007/s00265-016-2242-8

Falconer DS (1993) Introduction to Quantitative Genetics. Longman Scientific & Technical, Essex, England

Firth JA, Hadfield JD, Santure AW, Slate J, Sheldon BC (2015) The influence of nonrandom extra-pair paternity on heritability estimates derived from wild pedigrees. Evolution 69:1336-1344. doi: 10.1111/evo.12649

Fitzpatrick MJ, Ben-Shahar Y, Smid HM, Vet LEM, Robinson GE, Sokolowski MB (2005) Candidate genes for behavioural ecology. Trends in Ecology & Evolution 20:96-104. doi: 10.1016/j.tree.2004.11.017

Franks SJ, Weber JJ, Aitken SN (2014) Evolutionary and plastic responses to climate change in terrestrial plant populations. Evolutionary Applications 7:123-139. doi: 10.1111/eva.12112

Fraser KC, Shave A, de Greef E, Siegrist J, Garroway CJ (2019) Individual Variability in Migration Timing Can Explain Long-Term, Population-Level Advances in a Songbird. Frontiers in Ecology and Evolution 7:7. doi: 10.3389/fevo.2019.00324

Gienapp P, Brommer JE (2014) Evolutionary dynamics in response to climate change. In: Charmantier A, Garant D, Kruuk LEB (eds) Quantitative Genetics in the Wild. Oxford University Press, Oxford, pp 254-273

Gienapp P, Laine VN, Mateman AC, van Oers K, Visser ME (2017) Environment-Dependent Genotype-Phenotype Associations in Avian Breeding Time. Frontiers in Genetics 8. doi: 10.3389/fgene.2017.00102

Gienapp P, Leimu R, Merila J (2007) Responses to climate change in avian migration time - microevolution versus phenotypic plasticity. Climate Research 35:25-35. doi: 10.3354/cr00712

Gienapp P, Lof M, Reed TE, McNamara J, Verhulst S, Visser ME (2013) Predicting demographically sustainable rates of adaptation: can great tit breeding time keep pace with climate change? Philosophical Transactions of the Royal Society B-Biological Sciences 368:10. doi: 10.1098/rstb.2012.0289

Gwinner E (1996) Circadian and circannual programmes in avian migration. Journal of Experimental Biology 199:39-48

Haest B, Huppop O, Bairlein F (2018) The influence of weather on avian spring migration phenology: What, where and when? Global Change Biology 24:5769-5788. doi: 10.1111/gcb.14450

Hasselquist D, Montras-Janer T, Tarka M, Hansson B (2017) Individual consistency of long-distance migration in a songbird: significant repeatability of autumn route, stopovers and wintering sites but not

in timing of migration. Journal of Avian Biology 48:91-102. doi: 10.1111/jav.01292

Helm B et al. (2013) Annual rhythms that underlie phenology: biological time-keeping meets environmental change. Proceedings of the Royal Society B-Biological Sciences 280. doi: 10.1098/rspb.2013.0016

Helm B, Van Doren BM, Hoffmann D, Hoffmann U (2019) Evolutionary Response to Climate Change in Migratory Pied Flycatchers. Current Biology 29:3714-+. doi: 10.1016/j.cub.2019.08.072

Hill WG, Goddard ME, Visscher PM (2008) Data and theory point to mainly additive genetic variance for complex traits. Plos Genetics 4:10. doi: 10.1371/journal.pgen.1000008

Huppop O, Winkel W (2006) Climate change and timing of spring migration in the long-distance migrant Ficedula hypoleuca in central Europe: the role of spatially different temperature changes along migration routes. Journal of Ornithology 147:344-353. doi: 10.1007/s10336-005-0049-x

Ioannidis JPA, Tarone R, McLaughlin JK (2011) The False-positive to False-negative Ratio in Epidemiologic Studies. Epidemiology 22:450-456. doi: 10.1097/EDE.0b013e31821b506e

Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Molecular Ecology 16:1099-1106. doi: 10.1111/j.1365-294X.2007.03089.x

Kearney MR, Briscoe NJ, Karoly DJ, Porter WP, Norgate M, Sunnucks P (2010) Early emergence in a butterfly causally linked to anthropogenic warming. Biology Letters 6:674-677. doi: 10.1098/rsbl.2010.0053

Kim JM et al. (2018) A high-density SNP chip for genotyping great tit (Parus major) populations and its application to studying the genetic architecture of exploration behaviour. Molecular Ecology Resources 18:877-891. doi: 10.1111/1755-0998.12778

Knudsen E et al. (2011) Challenging claims in the study of migratory birds and climate change. Biological Reviews 86:928-946. doi: 10.1111/j.1469-185X.2011.00179.x

Kolarova E, Matiu M, Menzel A, Nekovar J, Lumpe P, Adamik P (2017) Changes in spring arrival dates and temperature sensitivity of migratory birds over two centuries. International Journal of Biometeorology 61:1279-1289. doi: 10.1007/s00484-017-1305-5

Kolarova E, Nekovar J, Adamik P (2014) Long-term temporal changes in central European tree phenology (1946-2010) confirm the recent extension of growing seasons. International Journal of Biometeorology 58:1739-1748. doi: 10.1007/s00484-013-0779-z

Kolecek J, Adamik P, Reif J Shifts in migration phenology under climate change: temperature vs. abundance effects in birds. Climatic Change. doi: 10.1007/s10584-020-02668-8

Krist M, Munclinger P (2011) Superiority of extra-pair offspring: maternal but not genetic effects as revealed by a mixed cross-fostering design. Molecular Ecology 20:5074-5091. doi: 10.1111/j.1365-294X.2011.05337.x

Krist M, Nadvornik P, Uvirova L, Bures S (2005) Paternity covaries with laying and hatching order in the collared flycatcher Ficedula albicollis. Behavioral Ecology and Sociobiology 59:6-11. doi: 10.1007/s00265-005-0002-2

Lafuente E, Beldade P (2019) Genomics of Developmental Plasticity in Animals. Frontiers in Genetics 10. doi: 10.3389/fgene.2019.00720

Laine VN et al. (2019) The Genomics of Circadian Timing in a Wild Bird, the Great Tit (Parus major). Frontiers in Ecology and Evolution 7. doi: 10.3389/fevo.2019.00152

Lehikoinen A et al. (2019) Phenology of the avian spring migratory passage in Europe and North America: Asymmetric advancement in time and increase in duration. Ecological Indicators 101:985-991. doi: 10.1016/j.ecolind.2019.01.083

Liedvogel M, Akesson S, Bensch S (2011) The genetics of migration on the move. Trends in Ecology & Evolution 26:561-569. doi: 10.1016/j.tree.2011.07.009

Liedvogel M, Cornwallis CK, Sheldon BC (2012) Integrating candidate gene and quantitative genetic

approaches to understand variation in timing of breeding in wild tit populations. Journal of Evolutionary Biology 25:813-823. doi: 10.1111/j.1420-9101.2012.02480.x

Liedvogel M, Sheldon BC (2010) Low variability and absence of phenotypic correlates of Clock gene variation in a great tit Parus major population. Journal of Avian Biology 41:543-550. doi: 10.1111/j.1600-048X.2010.05055.x

Liedvogel M, Szulkin M, Knowles SCL, Wood MJ, Sheldon BC (2009) Phenotypic correlates of Clock gene variation in a wild blue tit population: evidence for a role in seasonal timing of reproduction. Molecular Ecology 18:2444-2456. doi: 10.1111/j.1365-294X.2009.04204.x

Lifjeld JT, Slagsvold T, Ellegren H (1997) Experimental mate switching in pied flycatchers: Male copulatory access and fertilization success. Animal Behaviour 53:1225-1232. doi: 10.1006/anbe.1996.0430

Lisovski S et al. (2020) Light-level geolocator analyses: A user's guide. Journal of Animal Ecology 89:221-236. doi: 10.1111/1365-2656.13036

Lisovski S, Hahn S (2012) GeoLight-processing and analysing light-based geolocator data in R. Methods in Ecology and Evolution 3:1055-1059. doi: 10.1111/j.2041-210X.2012.00248.x

Lopez-Lopez P, Garcia-Ripolles C, Urios V (2014) Individual repeatability in timing and spatial flexibility of migration routes of trans-Saharan migratory raptors. Current Zoology 60:642-652. doi: 10.1093/czoolo/60.5.642

Lourenco PM, Kentie R, Schroeder J, Groen NM, Hooijmeijer J, Piersma T (2011) Repeatable timing of northward departure, arrival and breeding in Black-tailed Godwits Limosa l. limosa, but no domino effects. Journal of Ornithology 152:1023-1032. doi: 10.1007/s10336-011-0692-3

Lundberg M et al. (2017) Genetic differences between willow warbler migratory phenotypes are few and cluster in large haplotype blocks. Evolution Letters 1:155-168. doi: 10.1002/evl3.15

Merila J, Hendry AP (2014) Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. Evolutionary Applications 7:1-14. doi: 10.1111/eva.12137

Merlin C, Liedvogel M (2019) The genetics and epigenetics of animal migration and orientation: birds, butterflies and beyond. Journal of Experimental Biology 222. doi: 10.1242/jeb.191890

Moller AP (2001) Heritability of arrival date in a migratory bird. Proceedings of the Royal Society B-Biological Sciences 268:203-206

Moller AP, Rubolini D, Lehikoinen E (2008) Populations of migratory bird species that did not show a phenological response to climate change are declining. Proceedings of the National Academy of Sciences of the United States of America 105:16195-16200. doi: 10.1073/pnas.0803825105

Mueller JC, Pulido F, Kempenaers B (2011) Identification of a gene associated with avian migratory behaviour. Proceedings of the Royal Society B-Biological Sciences 278:2848-2856. doi: 10.1098/rspb.2010.2567

Nagy AD, Csernus VJ (2007) The role of PACAP in the control of circadian expression of clock genes in the chicken pineal gland. Peptides 28:1767-1774. doi: 10.1016/j.peptides.2007.07.013

Ouwehand J, Both C (2017) African departure rather than migration speed determines variation in spring arrival in pied flycatchers. Journal of Animal Ecology 86:88-97. doi: 10.1111/1365-2656.12599

Parody-Merino AM, Battley PF, Conklin JR, Fidler AE (2019) No evidence for an association between Clock gene allelic variation and migration timing in a long-distance migratory shorebird (Limosa lapponica baueri). Oecologia 191:843-859. doi: 10.1007/s00442-019-04524-8

Pearce-Higgins J, Green R (2014) Birds and climate change: impacts and conservation responses. Cambridge University Press, Cambrige

Peterson MP, Abolins-Abols M, Atwell JW, Rice RJ, Milá B, Ketterson ED (2013) Variation in candidate genes CLOCK and ADCYAP1 does not consistently predict differences in migratory behavior in the songbird genus Junco. F1000Research 2:17

Potti J (1998) Arrival time from spring migration in male Pied Flycatchers: Individual consistency and familial resemblance. Condor 100:702-708. doi: 10.2307/1369752

Pulido F (2007a) The genetics and evolution of avian migration. Bioscience 57:165-174. doi: 10.1641/b570211

Pulido F (2007b) Phenotypic changes in spring arrival: evolution, phenotypic plasticity, effects of weather and condition. Climate Research 35:5-23. doi: 10.3354/cr00711

Pulido F, Berthold P, Mohr G, Querner U (2001) Heritability of the timing of autumn migration in a natural bird population. Proceedings of the Royal Society B-Biological Sciences 268:953-959. doi: 10.1098/rspb.2001.1602

Radchuk V et al. (2019) Adaptive responses of animals to climate change are most likely insufficient. Nature Communications 10. doi: 10.1038/s41467-019-10924-4

Ralston J et al. (2019) Length polymorphisms at two candidate genes explain variation of migratory behaviors in blackpoll warblers (Setophaga striata). Ecology and Evolution 9:8840-8855. doi: 10.1002/ece3.5436

Romano A et al. (2018) Circadian genes polymorphism and breeding phenology in a resident bird, the yellow-legged gull. Journal of Zoology 304:117-123. doi: 10.1111/jzo.12501

Rosenfield RN, Hardin MG, Bielefeldt J, Keyel ER (2017) Are life history events of a northern breeding population of Cooper's Hawks influenced by changing climate? Ecology and Evolution 7:399-408. doi: 10.1002/ece3.2619

Rubolini D, Moller AP, Rainio K, Lehikoinen E (2007) Intraspecific consistency and geographic variability in temporal trends of spring migration phenology among European bird species. Climate Research 35:135-146. doi: 10.3354/cr00720

Saino N, Ambrosini R (2008) Climatic connectivity between Africa and Europe may serve as a basis for phenotypic adjustment of migration schedules of trans-Saharan migratory birds. Global Change Biology 14:250-263. doi: 10.1111/j.1365-2486.2007.01488.x

Saino N et al. (2017) Migration phenology and breeding success are predicted by methylation of a photoperiodic gene in the barn swallow. Scientific Reports 7. doi: 10.1038/srep45412

Saino N et al. (2013) Timing of molt of barn swallows is delayed in a rare Clock genotype. Peerj 1. doi: 10.7717/peerj.17

Schmaljohann H (2019) The start of migration correlates with arrival timing, and the total speed of migration increases with migration distance in migratory songbirds: a cross-continental analysis. Movement Ecology 7. doi: 10.1186/s40462-019-0169-1

Sergio F et al. (2014) Individual improvements and selective mortality shape lifelong migratory performance. Nature 515:410-+. doi: 10.1038/nature13696

Sheldon BC, Davidson P, Lindgren G (1999) Mate replacement in experimentally widowed collared flycatchers (Ficedula albicollis): determinants and outcomes. Behavioral Ecology and Sociobiology 46:141-148. doi: 10.1007/s002650050603

Stanley CQ, MacPherson M, Fraser KC, McKinnon EA, Stutchbury BJM (2012) Repeat Tracking of Individual Songbirds Reveals Consistent Migration Timing but Flexibility in Route. Plos One 7:6. doi: 10.1371/journal.pone.0040688

Stapley J et al. (2010) Adaptation genomics: the next generation. Trends in Ecology & Evolution 25:705-712. doi: 10.1016/j.tree.2010.09.002

Steinmeyer C, Mueller JC, Kempenaers B (2009) Search for informative polymorphisms in candidate genes: clock genes and circadian behaviour in blue tits. Genetica 136:109-117. doi: 10.1007/s10709-008-9318-y

Stoffel MA, Nakagawa S, Schielzeth H (2017) rptR: repeatability estimation and variance decomposition by generalized linear mixed-effects models. Methods in Ecology and Evolution 8:1639-1644. doi: 10.1111/2041-210x.12797

Sullivan AR, Flaspohler DJ, Froese RE, Ford D (2016) Climate variability and the timing of spring raptor migration in eastern North America. Journal of Avian Biology 47:208-218. doi: 10.1111/jav.00692

Tarka M, Hansson B, Hasselquist D (2015) Selection and evolutionary potential of spring arrival phenology in males and females of a migratory songbird. Journal of Evolutionary Biology 28:1024-1038. doi: 10.1111/jeb.12638

Team RC (2019) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

Teplitsky C, Mouawad NG, Balbontin J, de Lope F, Moller AP (2011) Quantitative genetics of migration syndromes: a study of two barn swallow populations. Journal of Evolutionary Biology 24:2025-2039. doi: 10.1111/j.1420-9101.2011.02342.x

Thackeray SJ et al. (2016) Phenological sensitivity to climate across taxa and trophic levels. Nature 535:241-U294. doi: 10.1038/nature18608

Toews DPL, Taylor SA, Streby HM, Kramer GR, Lovette IJ (2019) Selection on VPS13A linked to migration in a songbird. Proceedings of the National Academy of Sciences of the United States of America 116:18272-18274. doi: 10.1073/pnas.1909186116

Tomotani BM et al. (2018) Climate change leads to differential shifts in the timing of annual cycle stages in a migratory bird. Global Change Biology 24:823-835. doi: 10.1111/gcb.14006

Usui T, Butchart SHM, Phillimore AB (2017) Temporal shifts and temperature sensitivity of avian spring migratory phenology: a phylogenetic meta-analysis. Journal of Animal Ecology 86:250-261. doi: 10.1111/1365-2656.12612

van Asch M, Salis L, Holleman LJM, van Lith B, Visser ME (2013) Evolutionary response of the egg hatching date of a herbivorous insect under climate change. Nature Climate Change 3:244-248. doi: 10.1038/nclimate1717

Van Buskirk J, Mulvihill RS, Leberman RC (2012) Phenotypic plasticity alone cannot explain climate-induced change in avian migration timing. Ecology and Evolution 2:2430-2437. doi: 10.1002/ece3.367

van Wijk RE, Bauer S, Schaub M (2016) Repeatability of individual migration routes, wintering sites, and timing in a long-distance migrant bird. Ecology and Evolution 6:8679-8685. doi: 10.1002/ece3.2578

Vardanis Y, Klaassen RHG, Strandberg R, Alerstam T (2011) Individuality in bird migration: routes and timing. Biology Letters 7:502-505. doi: 10.1098/rsbl.2010.1180

Vedder O, Bouwhuis S, Sheldon BC (2013) Quantitative Assessment of the Importance of Phenotypic Plasticity in Adaptation to Climate Change in Wild Bird Populations. Plos Biology 11:10. doi: 10.1371/journal.pbio.1001605

Verhagen I et al. (2019) Genetic and phenotypic responses to genomic selection for timing of breeding in a wild songbird. Functional Ecology 33:1708-1721. doi: 10.1111/1365-2435.13360

Visser ME, Caro SP, van Oers K, Schaper SV, Helm B (2010) Phenology, seasonal timing and circannual rhythms: towards a unified framework. Philosophical Transactions of the Royal Society B-Biological Sciences 365:3113-3127. doi: 10.1098/rstb.2010.0111

Weidinger K, Kral M (2007) Climatic effects on arrival and laying dates in a long-distance migrant, the Collared Flycatcher Ficedula albicollis. Ibis 149:836-847. doi: 10.1111/j.1474-919X.2007.00719.x

Wilson AJ et al. (2010) An ecologist's guide to the animal model. Journal of Animal Ecology 79:13-26. doi: 10.1111/j.1365-2656.2009.01639.x

Wolak K, Keller F (2014) Dominance genetic variance and inbreeding in natural populations. In: Charmantier A, Garant D, Kruuk LEB (eds) Quantitative Genetics in the Wild. Oxford University Press, Oxford

Wolak ME (2012) nadiv: an R package to create relatedness matrices for estimating non-additive genetic variances in animal models. Methods in Ecology and Evolution 3:792-796. doi: 10.1111/j.2041-210X.2012.00213.x

Zhan S et al. (2014) The genetics of monarch butterfly migration and warning colouration. Nature 514:317-+. doi: 10.1038/nature13812

Table 1 Allele frequency, observed and expected heterozygosity, chi-square test of Hardy-Weinberg equilibrium and estimated proportion of null alleles for four candidate genes (n = 407 individuals). K = number of alleles, nt = individuals not successfully typed at the particular locus. F(null) = estimated frequency of null alleles.

Locus	k	HObs	HExp	alelle size (bp)/frequency	X^2	p	f(null)
CLOCK	4	0.327	0.328	120/0.080, 123/0.076, 126/0.812, 129/0.030, nt/0.003	1.59	0.663	0.009
NPAS2	4	0.025	0.029	165/0.001, 168/0.001, 175/0.985, 181/0.010, nt/0.003	-	-	0.078
ADCYAP1	6	0.592	0.585	172/0.004, 176/0.006, 178/0.052, 180/0.348, 182/0.538, 184/0.048, nt/0.005	10.3	0.112	-0.009
CREB1	9	0.514	0.797	268/0.054, 270/0.063, 274/0.280, 276/0.323, 278/0.041, 280/0.059, 282/0.066, 284/0.018, 286/0.052, nt/0.044	381.1	<0.001	0.214

Table 2 Variance components of male spring arrival date calculated from an animal model. Variance components are given as both absolute values and relative to the total phenotypic variance. Relative additive genetic component is the heritability of the trait.

	Absolute VC ± SE	Relative VC ± SE
Additive genetic	1.58 ± 7.03	0.031 ± 0.170
Permanent environment	9.89 ± 8.74	0.194 ± 0.712
Common environment	2.32 ± 6.67	0.045 ± 0.201
Annual	4.68 ± 4.26	0.092 ± 0.307
Residual	32.59 ± 5.03	0.638 ± 1.247

Figure legend

FIgure 1 Validation of the catching date as the arrival date. (a) Relationship between spring arrival date (expressed as day of the year) inferred from geolocators and our estimate of arrival date given as the midpoint between the date of first catching of the male and the date of previous catching session. (b) Distribution of mating speeds. The speed is expressed as a difference between laying date of the first egg in the nest this male attends and male estimated arrival date. Negative values mean that the first egg was laid before we caught the attending male.

Figure 2 Scatterplots showing the relationships between mean allele length at three candidate loci and the timing of six annual cycle events. BS = breeding site, non-BS = non-breeding site, Sahara = timing of Sahara crossing.

Figure 3 Relationships between mean allele length (*CLOCK, ADCYAP1, CREB1*) or genotype (*NPAS2*) and male spring arrival dates to breeding grounds. Boxplot shows median and interquartile range. Symbol size corresponds to the number of years used for calculation of individual means (range 1-4).

Figure 4 Phenotypic plasticity of male spring arrival dates (day of the year). Lines connecting two adjacent years delimit the same individual (n=106 repeatedly recorded males).

Figure 1

(a) Scatter plot with x-axis "True arrival date (geolocators)" and y-axis "Estimated arrival date (catching)".
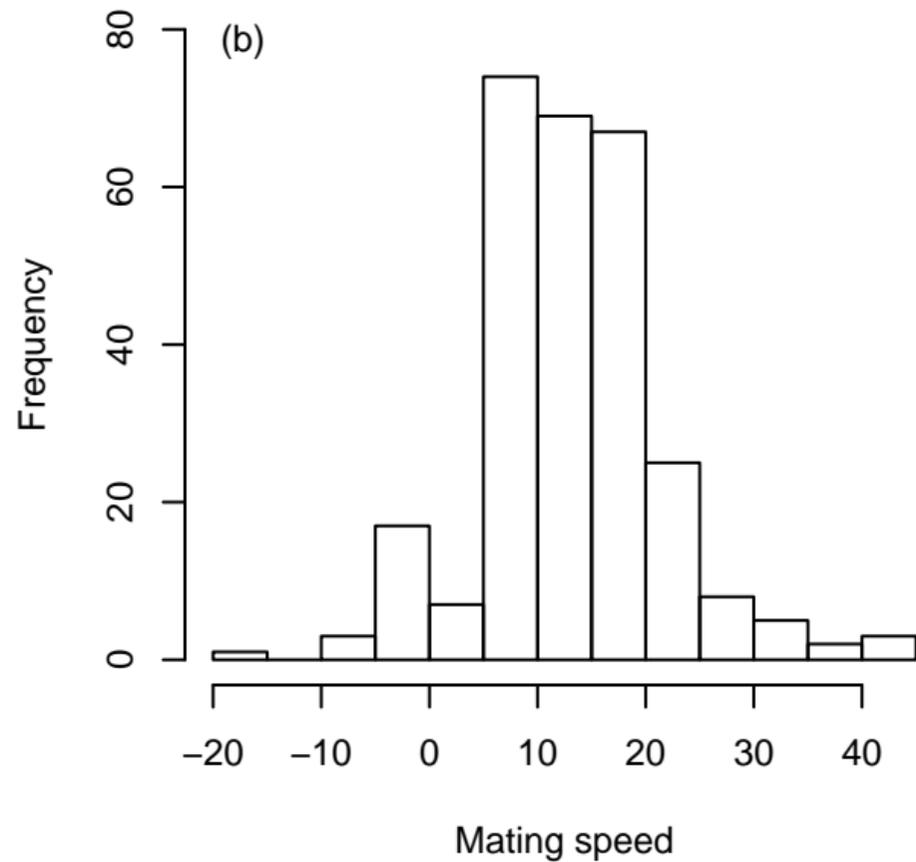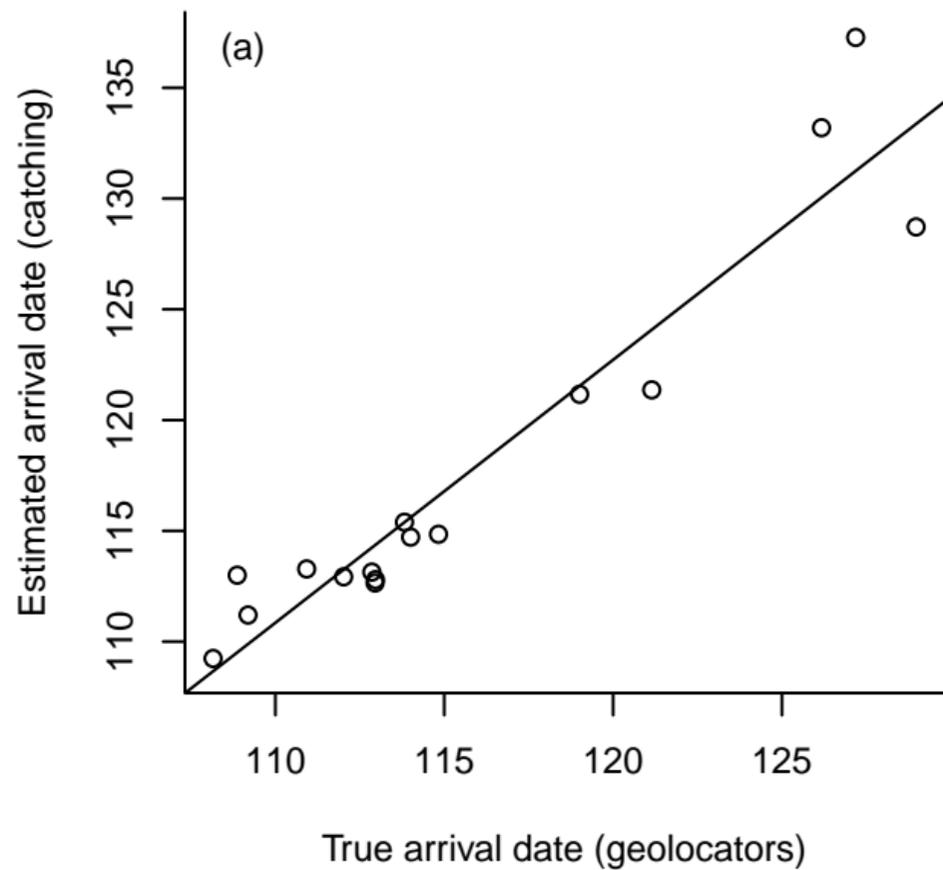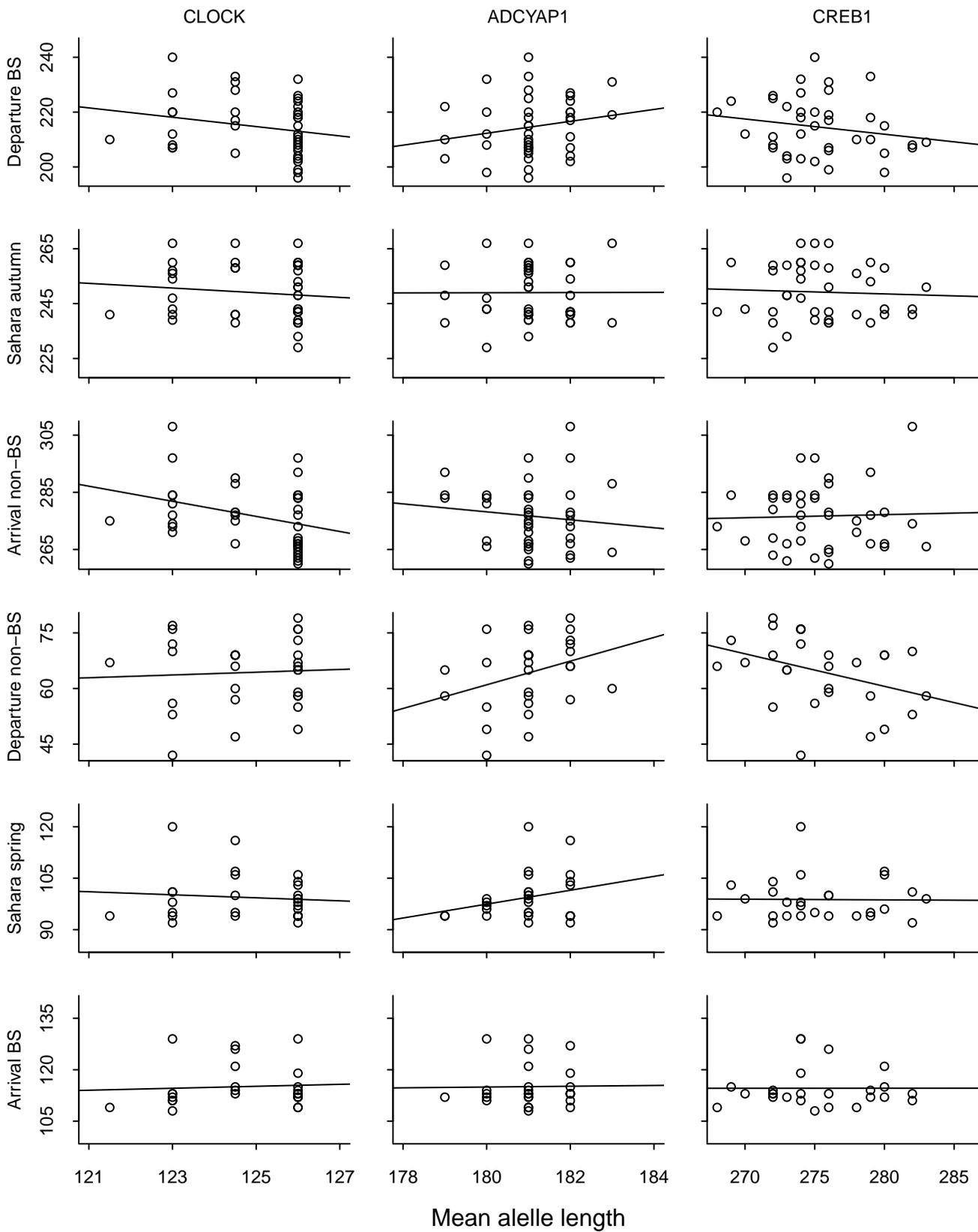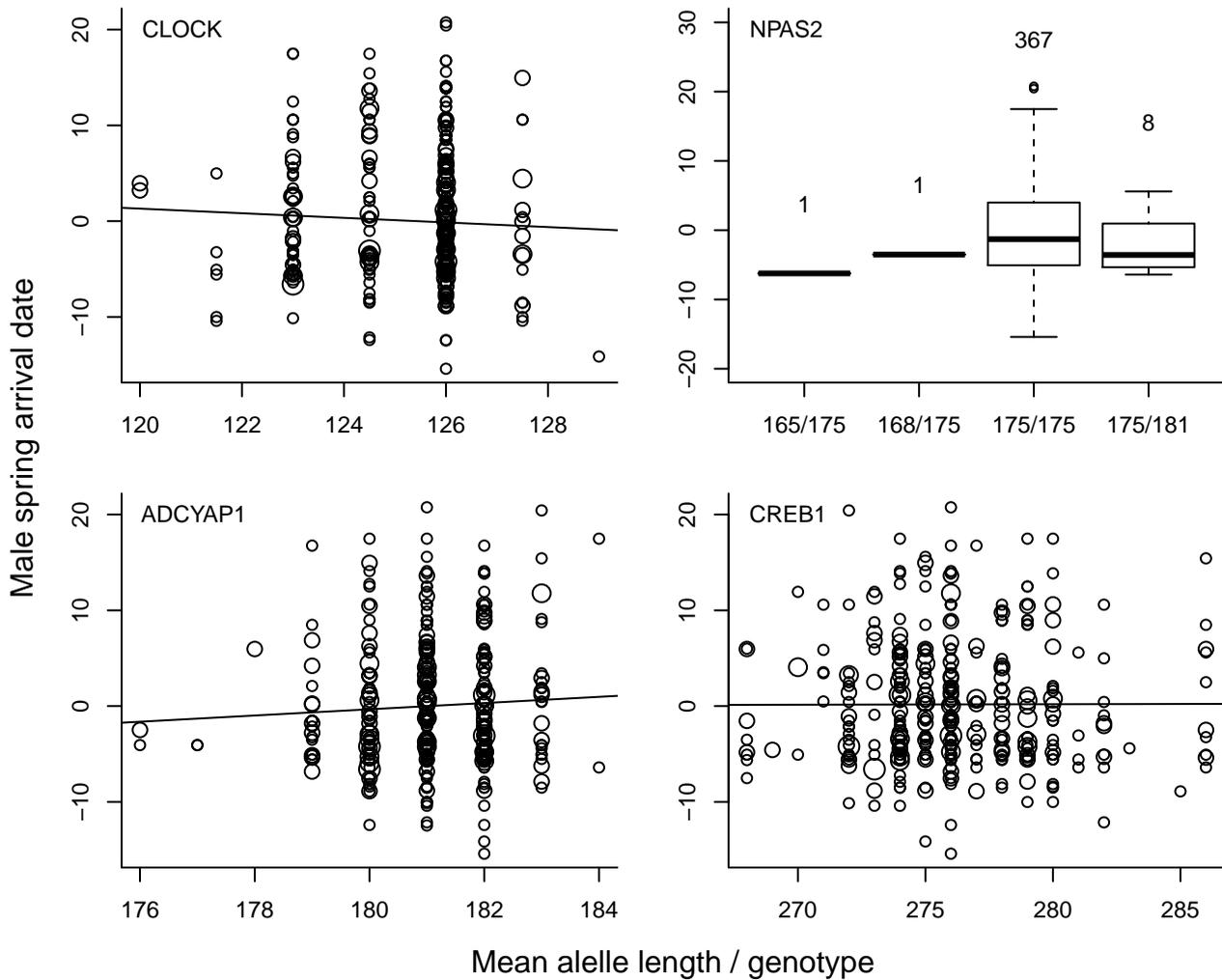
(b) Histogram with x-axis "Mating speed" and y-axis "Frequency".

Figure 2

Figure 3

Figure 4