Evol Ecol (2012) 26:1119–1151
DOI 10.1007/s10682-012-9567-1

IDEAS & PERSPECTIVES

Meta-analysis and sexual selection: past studies and future possibilities

Michael D. Jennions · Andrew T. Kahn · Clint D. Kelly · Hanna Kokko

Received: 5 January 2012 / Accepted: 6 March 2012 / Published online: 16 March 2012

Abstract The action of sexual selection is highly variable among taxa. This creates challenges when trying to generalize (e.g. determine if a particular relationship exists based on its average strength, or if it varies in response to theoretically relevant factors). Consequently, accounting for moderating factors is likely to be crucial to explain differences in sexual selection among studies. In principle, given measures of key theoretical parameters we can predict the strength of sexual selection on different sexual signals, the benefits of mate choice, the extent of sex differences (e.g. in immune function or survival) and the likely life history trade-offs between investment into different sexual traits (e.g. sperm vs. courtship) or non-sexual traits (e.g. immune function, traits that increase longevity, parental care). How well does empirical data support theoretical expectations? First, we provide a short history of the use of meta-analysis in sexual selection studies. We present a table summarizing 94 meta-analyses that have asked questions about sexual selection or allied topics of interest to those studying sexual selection (e.g. the link between heterozygosity and fitness). Second, we list the main ways that meta-analysis has been used in sexual selection work and provide illustrative examples. Third, we provide practical advice to identify questions that are ripe for meta-analysis. We highlight 11 sexual selection topics where meta-analyses are needed (e.g. there are no meta-analyses testing game theory models of fighting contests). Finally, we discuss some general issues that will arise as the use of meta-analysis in sexual selection studies becomes more sophisticated.

Order of authorship is alphabetical.

Electronic supplementary material The online version of this article (doi:10.1007/s10682-012-9567-1) contains supplementary material, which is available to authorized users.

M. D. Jennions (✉) · A. T. Kahn · H. Kokko
Evolution, Ecology and Genetics, Research School of Biology, The Australian National University, Building 116, Canberra, ACT 0200, Australia
e-mail: michael.jennions@anu.edu.au

C. D. Kelly
Department of Ecology, Evolution and Organismal Biology, Iowa State University, Ames, IA 50011, USA

Keywords Inbreeding · Intra-specific variation · Mate choice · Mating success · Polyandry · Sex ratio · Sperm competition

Introduction

The concept of 'physics envy' describes the alleged fact that biologists would like to have theories that allow researchers to design experiments that readily distinguish between competing hypotheses. Biological data, especially in ecology and evolution, rarely provides researchers with such opportunities. Biological interactions are complex, forcing researchers to consider mutually non-exclusive hypotheses, and outcomes are often highly variable, being affected by environmental factors (Candolin and Heuschele 2008), demography (Kokko and Rankin 2006; Weir et al. 2011), historic contingencies and stochastic events (Jennions et al. 2012a). This applies to inter-specific parasite-host interactions (Poulin and Forbes 2012), or contests for resources or space (Castellanos and Verdú 2012), but it also applies to within-species interactions, especially those between the sexes over reproduction. Closely related species often have markedly different mating systems and patterns of parental care leading to variation in sexual selection. The same can be true for different populations of a single species—and even the same population in different years (review: Siepielski et al. 2009; but see Morrissey and Hadfield 2012) or stages of the season (e.g. Forsgren et al. 2004).

This creates challenges when trying to make generalizations that apply across taxa (e.g. to determine whether a particular factor is generally important), and when trying to understand why a particular feature emerges under some conditions but not others. Both tasks are important: without the ability to generalize and to infer associations that hint at causal factors of general importance, science can not proceed past individual cases of storytelling. Meta-analysis offers a potential solution. Generalizing is achieved by combining the results of different studies, with greater weight given to those offering more precise estimates, to determine the average influence of a factor of interest. We can estimate the mean 'effect size', which is a standardized measure of the strength of a relationship (Koricheva et al. 2012). Even more important is that we can include covariates, either continuous or categorical, into a meta-analysis to see if they correlate with the effect size estimates across studies. This approach, often referred to as a meta-regression (e.g. Jones et al. 2009), can identify factors that are potentially causally related to the phenomenon of interest. This is important for testing theory: advanced theory only rarely predicts a single outcome (e.g. 'females benefit by producing extra-pair young'), but rather patterns of covariation (e.g. 'production of extra-pair young should reflect variation in the costs incurred by mothers'). Meta-regression can reveal higher patterns that are undetectable in the original studies and either corroborate or refute existing theory (e.g. Griffin et al. 2005; Weir et al. 2011) or raise new research questions (see Jennions et al. 2012b).

A short history of meta-analyses of sexual selection

The first empirical meta-analyses by ecologists and evolutionary biologists appeared in 1992 (Gurevitch et al. 1992; Vanderwerf 1992). Subsequently, a key review paper aimed at evolutionary biologists highlighted the value of meta-analysis (Arnqvist and Wooster 1995). The first meta-analysis to explicitly address a sexual selection question considered

assortative mating in water striders (Arnqvist et al. 1996). The number of meta-analyses of sexual selection and allied topics has since remained static at about 4–5 per year, although there was a sudden surge to 16 meta-analyses published or available online in 2011 (Fig. 1a). In other areas of evolutionary biology and ecology one can find examples of much faster increases in the uptake of meta-analysis (e.g. host-parasite interactions; Fig. 1

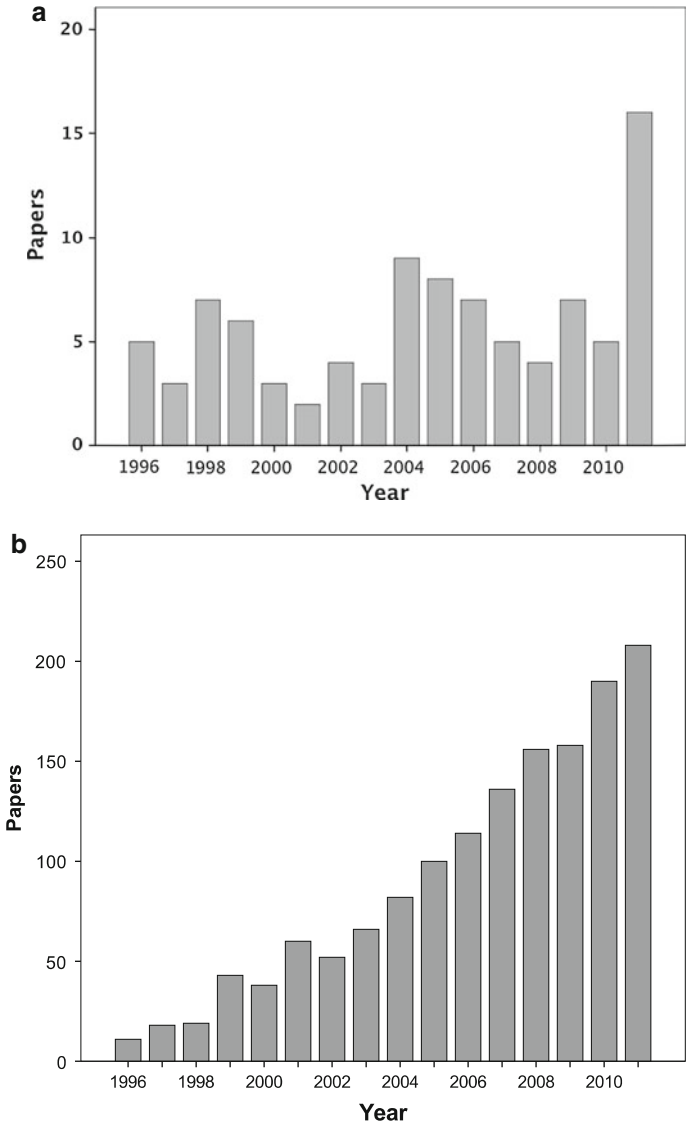

Fig. 1 **a** Papers per year that include actual meta-analyses on topics on or closely linked to the study of sexual selection ($n = 94$, see Table 1). **b** Papers per year located using the search term topic = *meta-analys** or *metaanalys** in the field categories *Ecology*, *Evolutionary Biology*, *Plant Sciences* and *Zoology* using the ISI Web of Science (Dec 31, 2011). *Note*: not all of these studies are actual meta-analyses (e.g. some might simply refer to the need for meta-analysis)

in Poulin and Forbes 2012). The total number of ecological and evolutionary meta-analysis published shows a steady annual increase (Fig. 1b; see also Jennions et al. 2012c).

Some of the earliest meta-analyses explored the role of fluctuating asymmetry (FA) in sexual selection. At the time, late 1990s, there was already disquiet as to whether FA in sexual traits offered a single characteristic that could capture much of the variation in individual quality and, hence, be a target of mate choice. Consequently, many FA studies were controversial (Leamy and Klingenberg 2005). In this atmosphere, early meta-analyses of FA and sexual selection often generated intense, even vitriolic, debates [e.g. commentaries in *J. Evol. Biol* on Møller and Thornhill (1997)]. In addition, there were claims that publication bias was a major problem in published meta-analyses (Palmer 1999; Simmons et al. 1999; but see Møller et al. 2005), which led to wider concern that the use of meta-analysis in ecology and evolution might be inappropriate due to systemic publication bias (Palmer 2000). There were also issues surrounding inclusion of unpublished work that could not be evaluated (Palmer 1999), and the lack of objective search criteria for relevant studies, with the inference that some studies were deliberately excluded from meta-analyses to bias the outcome.

In hindsight, many of the objections raised echo those from other disciplines (e.g. medicine, social sciences) when meta-analysis was first used. For example, the temporal decline in effect size found in ecological or evolutionary studies (Simmons et al. 1999; Poulin 2000; Jennions and Møller 2002) has been reported in many other fields (Koricheva et al. 2012). Similarly, a negative relationship between sample sizes and effect sizes has been reported in other fields (Jennions et al. 2012b). This pattern is expected if small sample studies are only published when they report significant results (but it can arise for many other reasons). More generally, as with any new technique, earlier studies were less rigorous and used less sophisticated approaches than those published later. For example, recent studies are far more likely to control for statistical non-independence (Nakagawa and Santos 2012) and correct for phylogeny (Chamberlain et al. in prep).

Despite these problems, several meta-analyses conducted in the 1990s were positively received, became highly cited, and probably encouraged others to attempt a meta-analysis. Notable among these were meta-analyses of the evidence for 'good genes' benefits to female choice (Møller and Alatalo 1999), the effects of polyandry on female fitness (Arnqvist and Nilsson 2000) and data on how well sexual traits predict male reproductive success in birds (Fiske et al. 1998; Møller and Ninni 1998). To date, at least 18 papers on ecological or evolutionary topics that include a meta-analysis have appeared in *Science* or *Nature*, including two on sexual selection related topics (effect of mate attractiveness on offspring sex ratio: West and Sheldon 2002; male-biased parasitism: Moore and Wilson 2002).

Here, we identify 94 papers that contain formal meta-analyses (i.e. the use of effect sizes, usually weighted by the inverse of their variance or sample size) on sexual selection or allied topics (Table 1; search protocol in Electronic Appendix 1). Our approach to defining 'sexual selection or allied topics' was pragmatic. We included topics of likely interest to those studying competition for mates and mating systems (e.g. such as maternal adjustment of offspring sex ratios in response to sire attractiveness or the future level of mating competition). We also considered relationships between immune function, testosterone, parasitism and survival, because the life history trade-off between immunity and sexual traits is a major area of research (i.e. the immunocompetence handicap hypothesis; Folstad and Karter 1992; see also Poulin and Forbes 2012). We included studies on phenotypic differences between the sexes because, in the absence of sexual selection, natural selection should result in similar optimal phenotypes, and the sexes should, all else

Table 1 Summary of details of sexual selection meta-analyses ($n = 94$)

General theme and main questions	Answer	Species	Studies	Data	Phylogeny	Taxa	Reference
Females mating preferences (correlates of male mating success)							
Do Zebra finch females prefer males with redder bills?	Yes. The mean effect was significant ($r = 0.61$), but the relationship might not be causal	1	11	All	N/A	Birds	Simons and Verhulst (2011)
Do male birds with more complex songs have greater reproductive success?	Yes. The mean effect was significant ($r = 0.20$) after correcting for possible publication bias	27	44	All	Yes (7)	Birds	Soma and Garamszegi (2011)
Do epaulet colour or size affect male reproductive success in Red-winged Blackbirds?	Yes and no. Pooling across measures of male-male competition, female choice and reproductive success the mean effect was not significant for colour, but was marginally non-significant for size ($r = 0.0.8$)	1	11	All	N/A	Birds	Yasukawa et al. (2010)
Do personality (boldness, aggressiveness and willingness to explore) predict reproductive success?	Yes and no. The mean effect was significant for boldness ($r = 0.10$) and aggressiveness ($r = 0.10$) but not for exploration	3	12	Subset (f)	No	All	Smith and Blumstein (2008)
Are the relationships between male resource holding potential, resource value and reproductive success positive?	Yes, but the correlations are weaker than one might expect ($r = 0.37$–0.45)	53	63	Subset (i)	No	All	Kelly (2008)
Does bib size in house sparrows correlate with a range of fitness components?	Yes and no. Bib size increased significantly with age and was bigger in more dominant males. There was, however, no significant correlation with reproductive success or direct benefits for females	1	15	Subset (c)	N/A	Birds	Nakagawa et al. (2007)
Do male birds with more elaborate song gain greater paternity in their own nests?	No, the mean effect did not differ from zero	8	11	All	No	Birds	Garamszegi and Møller (2004)
Do female red junglefowl prefer males with larger, brighter combs and normal rather than altered plumage?	Yes and no. There was a preference based on combs but not feathers	1	14	All	N/A	Birds	Parker and Ligon (2003)

Table 1 continued

General theme and main questions	Answer	Species	Studies	Data	Phylogeny	Taxa	Reference
What is the mean strength of sexual selection on male traits in birds?	The mean effect size for examined traits was $r = 0.30$. The mean effect was much stronger for lekking than non-lekking species	80	138	All	No	Birds	Gontard-Danek and Møller (1999)
Do morphological and behavioural traits predict male mating success in lekking species?	Yes. There was a significant effect of lek attendance, display frequency, fight frequency, distance from lek center, ornamentation, size and age ($r = 0.07$–0.53)	18	32	Subset (d)	No	All	Fiske et al. (1998)
Do male birds with more elaborate sexually selected traits have greater paternity in their own nests?	Yes. The mean effect was significant ($r = 0.34$).	22	22	Subset (e)	No	Birds	Møller and Ninni (1998)
Do females prefer males with larger claws in a fiddler crab?	Yes. The mean effect was significant, but the preference was significantly stronger in summer than winter ($r = 0.73$ vs. 0.18)	1	12	All	N/A	Arthropods	Milner et al. (2010)
Correlates of male attractiveness/mating success							
Are eumelanin-based sexual traits positively correlated with fitness components in birds?	No. The mean effect was not significantly different from zero for four fitness components, but it was significant for sexually dimorphic species for clutch size	14	32	All *	Yes (3)	Birds	Meunier et al. (2011)
Is there a positive relationship between dominance and expression of plumage traits in birds?	Yes. The mean effect was significant ($r = 0.36$), as was the effect for most subsets of the data	23	37	All	No (6)	Birds	Santos et al. (2011)
Does the ventral plumage of the great tit *Parus major* differ between the sexes and change with age?	Yes. The sexes differ significantly ($r = 0.24$), and there is a change with age ($r = 0.11$)	1	10	All	N/A	Birds	Evans et al. (2010)
A range of questions about ornaments in species with mutual ornamentation	Ornamentation is positively correlated with components of fitness ($r = 0.28$). Males prefer more ornamented females ($r = 0.39$) and there is positive assortative mating ($r = 0.39$)	30	36	Subset (h)	No	All	Kraaijeveld et al. (2007)
Are carotenoid-based sexual traits better signals of condition than melanin-based ones in birds?	No. The difference was not significant, although the effect was stronger for carotenoid-based traits ($r = 0.31$ vs. 0.12)	7	17	Subset (n)	No	Birds	Griffith et al. (2006)

Table 1 continued

General theme and main questions	Answer	Species	Studies	Data	Phylogeny	Taxa	Reference
Is the blue tit's song a reliable signal of male condition?	Yes and no. Singing intensity and strophe drift were significantly correlated with measures of condition ($r = 0.13$–0.18), but six other song traits were not	1	52	All *	N/A	Birds	Parker et al. (2006)
Does bird song signal parasite prevalence and/or intensity?	Yes and no. Yes for performance-related song traits ($r = 0.26$). No for song complexity, duration or frequency	6	8	All	No	Birds	Garamszegi (2005)
Do males with more elaborate sexually selected traits have greater survivorship?	Yes. The mean effect was significant ($r = 0.13$) suggesting that these generally costly traits have condition-dependent expression	40	69	All	No	All	Jennions et al. (2001)
Is there a relationship between parasite load, condition and plumage coloration in moulting birds?	Yes. Birds with more mites had lower muscle condition ($r = 0.23$) and duller plumage ($r = 0.38$)	9	9	Subset (l)	No	Birds	Harper (1999)
Is there a negative relationship between expression of sexual traits and parasite load or immune function?	Yes. The mean effect was significant ($r = 0.13$). The effect was much stronger for immune function than parasite load	39	50	All	No	All	Møller et al. (1999)
Does male sexual ornamentation correlate with parasite load, both within and across species?	No. The mean effect did not differ from zero. Across species, however, there is a positive relationship between ornamentation and levels of parasitism ($r = 0.20$)	–	55	All *	No	All	Hamilton and Poulin (1997)
Does body size affect mating rate and lifetime mating success in odonates?	Yes. There was a significant effect of body size on mating rate, lifetime mating success and territorial success ($r = 0.09$, 0.07, 0.40 respectively)	20	33	All *	No	Insects	Sokolovska et al. (2000)
Is there a relationship between neural tissue size (song nuclei in brain) and repertoire size in birds?	Yes. Pooling across different measures of song complexity and different brain areas the mean effect was significant ($r = 0.36$), but it varied for different pairwise relationships	9	41	All	No	Birds	Garamszegi and Eens (2004)

Table 1 continued

General theme and main questions	Answer	Species	Studies	Data	Phylogeny	Taxa	Reference
Does an elevated level of yolk androgens increase the expression of sexually selected traits in adult males?	No. The mean effect was non-significant for both traits that are insensitive or sensitive to androgen ($d = 0.00$ and 0.184 respectively)	7	7	All	No	Birds	Muller and Eens (2009)
Do aspects of song under sexual selection change with age in collared flycatchers? Is this due to differential survival?	Yes and no. Repertoire size and song rate changed significantly with age ($d = 0.73$ and -0.88) but other song traits did not. The pattern was not attributable to differential survival	1	1	All	N/A	Birds	Garamszegi et al. (2007)
Does male age affect within-pair paternity (WPP) and extra-pair paternity (EPP) in birds?	Yes and no. There is a significant increase in EPP with age ($r = 0.31$), but no significant effect of age on WPP ($r = 0.07$)	N/A	51	Subset (r)	No	Birds	Cleasby and Nakagawa (2012)
Costs and benefits of mate choice, polyandry and EPCs							
Do females gain genetic benefits for their offspring by mating multiply?	Yes and no. The mean effect based on a composite measure of offspring performance was significant or marginally non-significant ($d = 0.07$–0.19)	25	41	Subset (g)	No	All	Slatyer et al. (2012)
Do female arthropods that receive more ejaculate have higher fitness?	Yes and no. There was a significant effect for fecundity ($d = 0.45$), but a negative effect for longevity ($d = 0.20$)	70	84	All	No	Arthropods	South and Lewis (2011)
Do female birds engage in extra-pair copulations for genetic benefits?	No. There is (i) no significant difference between social and extra-pair mates in sexual traits (only for age and size) and (ii) no significant difference in survivorship of extra-pair and within-pair young	–	23	Subset (d)	No	Birds	Akcay and Roughgarden (2007)
Do the direct costs of extra-pair mating explain levels of extra-pair fertilizations across songbird species?	Yes. Species with fewer extra-pair young pay a significantly higher cost for engaging in extra-pair breeding	14	14	All	Yes (1)	Birds	Albrecht et al. (2006)

Table 1 continued

General theme and main questions	Answer	Species	Studies	Data	Phylogeny	Taxa	Reference
Do female birds engage in extra-pair copulations for genetic benefits?	No. There is no significant difference in fitness components between extra-pair and within-pair young but there is a direct cost due to a decline in male care	12	14	All	No	Birds	Arnqvist and Kirkpatrick (2005)
Does female polyandry improve egg hatching success or embryo viability in arthropods?	Yes and no. The mean effect was significant for hatching success ($r = 0.19$) but not for embryo viability	12	12	All	No	Insects, Arachnids	Simmons (2005)
Do female Lepidoptera that mate with virgins males have a greater reproductive output?	Yes. The mean effect was significant ($d = 0.33$) and significantly greater for polyandrous than monandrous species ($d = 0.45$ vs. 0.25)	25	29	All	No	Insects (Lepidoptera)	Torres-Vila and Jennions (2005)
Do the benefits of remating in Lepidoptera vary with the mating system and study design?	Yes. The mean effect was significantly greater for polyandrous than monandrous species ($d = 0.77$ vs. 0.27) but there was a confounding of effect of whether the study was experimental or descriptive	25	39	All	No	Insects (Lepidoptera)	Torres-Vila et al. (2004)
Do females who mate with males with more elaborate sexually selected traits gain direct benefits?	Yes. The mean effects were significant for fertility, fecundity, male care and offspring hatching success ($r = 0.11$–0.47)	24	26	All	No	All	Møller and Jennions (2001)
Does multiple mating have a positive effect on female fitness in insects?	Yes. The benefits of greater fecundity significantly outweigh a reduction in longevity, with nuptial feeding being an important moderator	78	122	All *	No	Insects	Arnqvist and Nilsson (2000)
Is there a positive correlation between male sexual trait expression and offspring survival?	Yes. The mean effect was significant ($r = 0.12$)	22	22	All	No	All	Møller and Alatalo (1999)
Sperm allocation							
Do males adjust ejaculate size in response to the risk and intensity of sperm competition, and to female mating status and quality?	Yes and no. Predictions for sperm competition risk ($d = 0.40$) and female quality ($r = 0.32$) were supported, but that for intensity was not. Unexpectedly males transferred more sperm to virgins ($d = 0.84$)	40	46	Subset (j)	No	All	Kelly and Jennions (2011)

Table 1 continued

General theme and main questions	Answer	Species	Studies	Data	Phylogeny	Taxa	Reference
Do males release more sperm when there is a greater risk of sperm competition?	Yes. The mean effect was significant ($d = 0.32$). Males release more sperm when mating with a non-virgin and when a competitor is present	37	39	All	Yes (3)	All	DelBarco-Trillo (2011)
Is there greater variation in size for nonfertile than fertile sperm morphotypes?	Yes. Sperm size variation is significantly greater for nonfertile sperm morphs in three insect taxa (all $P < 0.004$)	150	150	All *	No	Insects	Bernasconi and Hellriegel (2005)
Differences between the sexes							
Do males and females differ in their response to alarm calls in mammals?	Yes. Males are more responsive ($d = 0.16$–0.68)	6	9	All	No	Mammals	Lea and Blumstein (2011)
Does the level of predation predict the extent of male-biased in the adult sex ratio skew in Wetas?	No. The mean sex ratio did not deviate significantly from 1:1, and there was no significant difference between high and low predation sites	6	58 pops	All	No (6)	Insect	Wehi et al. (2011)
Does the cross-sex genetic correlation differ depending upon the type of trait?	Yes. The mean correlation is usually large, but it is smaller for traits more closely associated with fitness	47	114	All	No	All	Poissant et al. (2010)
Does the effect of latitude on body size differ between the sexes? And are males more variable in size than females?	Yes. Male size changes significantly more strongly with latitude. No, males are not significantly more variable in size than females	98	–	All	Yes (2)	All	Blanckenhorn et al. (2006)
Do males suffer higher rates of parasitism than females in mammals?	Yes. The mean effect was small but significant. The male-bias in parasitism was significantly greater in more sexually dimorphic species	106	355	All	Yes (4)	Mammals	Moore and Wilson (2002)
Do males suffer higher rates of parasitism in arthropods?	No. The mean effect size was not significantly different from zero	46	61	All	No	Arthropods	Sheridan et al. (2000)
Do males have more parasites than females in birds?	No. For the full data set the difference was not significant. However, *Haemoproteus* parasites were more common in breeding females than breeding males	35	35	All	No	Birds	McCurdy et al. (1998)

Table 1 continued

General theme and main questions	Answer	Species	Studies	Data	Phylogeny	Taxa	Reference
Do males suffer higher rates of parasitism than females?	Yes and no. Males have significantly more parasites than female for two of three parasite types	41	50	All *	No	Mammals	Schalk and Forbes (1997)
Do Helminth parasites grow faster in male than female vertebrate hosts?	Yes and no. The mean effect was significant for nematodes ($d = 0.35$) but not cestodes	6	34	All *	No	Birds, Mammals	Poulin (1996a)
Do males suffer higher rates of Helminth parasitism than females?	Yes and no. Prevalence of parasites was significantly higher in males for some, but not all, parasite types. In contrast, there was no obvious trend for a sex-bias for the intensity of infection	–	85	All *	No	All	Poulin (1996b)
Is there a sex difference in dispersal and philopatry in damselflies?	Yes. Females seem to have significantly higher dispersal rates based on likelihood of recapture at natal and breeding sites (Odds ratio = 2.04)	18	19	All *	Yes (2)	Insects	Beirinckx et al. (2006)
Is there male-biased parasitism in a rodent?	Yes. Pooling across parasite species there was a significantly greater prevalence and abundance of parasites in males (odds ratio = 1.40; $d = 0.13$)	19	N/A	All	Yes (2)	Mammals	Matthe et al. (2010)
Is mortality higher in male than female water strider due to the expression of sexual traits?	Yes. There was a marginally significantly higher rate of male mortality during final moults ($r = 0.104$)	1	4	All	N/A	Insects	Westlake and Rowe (1999)
Are measures of immune function higher in female than in male insects?	Yes and no. Phenoloxidase activity was significantly higher in females ($d = 0.184$), but there was no significant sex difference in haemocyte count ($d = 0.014$)	11	11	Subset (o)	No	Insects	Nunn et al. (2009)
Does flea species co-occurrence (i.e. more species than expected by chance) occur more often in male than female rodents?	Yes. Co-occurrence was significantly higher in males than females (odds ratio = 3.26)	16	50	All	Yes (4)	Mammals	Krasnov et al. (2011)

Table 1 continued

General theme and main questions	Answer	Species	Studies	Data	Phylogeny	Taxa	Reference
Speciation/extinction/colonization							
Does stronger sexual selection lead to higher rates of speciation?	Yes. The mean effect was significant ($r = 0.09$), but there are several important moderating factors	–	22	All	N/A	All	Kraaijeveld et al. (2011)
Does sexual dichromatism in birds explain variation in their success at becoming established as invasive species?	Yes. Species with less sexually dichromatic plumage are significantly more likely to become established ($r = 0.18$)	–	22	All	N/A	Birds	Blackburn et al. (2009)
Sex ratios and differential allocation							
Do female birds invest more into reproduction when breeding with a more attractive mate?	Yes. The mean effect was significant ($r = 0.17$), but varied depending on the breeding system. Species with biparental care tended to produce more eggs, while those with female-only care produced larger eggs	17	45	All	Yes (5)	Birds	Horváthová et al. (2012)
Do maternal rank or the level of local resource competition affect the birth sex ratios in baboons?	No. There was no effect of population growth rate or rank on the birth sex ratio	1	36	All	N/A	Mammal (Primates)	Silk et al. (2005)
Do female produce fewer sons when the level of local mate competition is higher?	Yes. The mean effect was significant ($r = 0.58$) (*note*: this effect includes studies of local resource competition)	64	87	All	Yes (4)	All	West et al. (2005)
Do mothers in better condition produce more sons than daughters in mammals?	No. When looking at all measures the mean effect was close to zero. However, measures taken close to conception are significant predictors ($r = 0.07$–0.08)	–	68	Subset (e)	No	Mammals	Cameron (2004)
Is there general evidence that birds facultatively adjust their offspring sex ratio?	No. The mean effect did not differ from zero. However, there was a significant effect of mate quality ($r = 0.08$), laying date ($r = 0.15$) and season ($r = 0.04$)	17	–	Subset (m)	No	Birds	Ewen et al. (2004) (b)

Table 1 continued

General theme and main questions	Answer	Species	Studies	Data	Phylogeny	Taxa	Reference
Does maternal rank in primates predict whether they produce more daughters than sons?	No. The mean effect size was not significantly different from zero. There was, however, a moderating effect of sexual dimorphism. The effect was stronger in species with lower dimorphism	18	30	All	No	Mammal (Primates)	Schino (2004)
Does greater maternal dominance and/or better body condition lead to a male-biased offspring sex ratio in ungulates?	Yes. The mean effect was significant ($r = 0.09$). The effect was stronger when condition was measured before birth or based on dominance	18	37	All	No	Mammals (Ungulates)	Sheldon and West (2004)
Does maternal rank in primates predict whether they produce more daughters than sons?	No. The mean effect size did not differ from zero. There was no evidence than rank affects the birth sex ratio	15	35	All	No	Mammal (Primates)	Brown and Silk (2002)
Do females produce more sons when mated to more attractive males?	Yes. The mean effect was significant ($r = 0.19$)	8	11	All +	No	Birds, Mammals, Insects	West and Sheldon (2002)
Do more attractive males provide more parental care?	No. Across species, however, the lower the level of extra-pair paternity the more care attractive males provided	18	22	All	Yes (4)	Birds	Møller and Thornhill (1998b)
Fluctuating asymmetry							
Does sexual selection favour more symmetric males, even when correcting for publication bias?	Yes. The mean effect was significant ($r = 0.30$). There was little evidence for publication bias	–	125	All	No	All	Møller et al. (2005)
Does sexual selection favour more symmetric males?	Yes. The mean effect was significant ($r = 0.34$)	42	65	All	No	All	Møller and Thornhill (1998a) (a)
Is sexual trait size a better predictor of reproductive success than trait symmetry?	No. The mean effect was larger for symmetry than size ($r = 0.22$ vs. 0.14)	29	33	All	No	All	Thornhill and Møller (1998)
Is developmental stability significantly heritable?	Yes. The mean effect was significant ($r = 0.19$). There was, however, much subsequent debate about the validity of this finding.	17	34	All	No	All	Møller and Thornhill (1997)

Table 1 continued

General theme and main questions	Answer	Species	Studies	Data	Phylogeny	Taxa	Reference
Is there a relationship between symmetry and fitness/stress? Does this relationship differ between sexual and non-sexual traits?	Yes and no. There was a relationship between symmetry and fitness/stress but it did not differ between sexual and non-sexual traits	61	71	Subset (o)	No	All	Leung and Forbes (1996)
Assortative mating							
How strong is assortative mating in bird hybrid zones?	The mean effect was $r = 0.47$	–	58	All	No	Birds	Randler (2008)
Is there size assortative mating in water striders?	Yes. There is significant positive assortment ($r = 0.12$). Most of the variation is among studies within species (i.e. no phylogenetic signal)	7	45	All	No (6)	All	Arnqvist et al. (1996)
Immunity							
Does greater reproductive effort increase levels of parasitism and lower immune function in birds?	Yes. Greater reproductive effort significantly increases blood parasite load ($r = 0.09$) and lower immune function ($r = 0.30$)	9	20	All	No	Birds	Knowles et al. (2009)
Does immune activation suppresses plasma testosterone level in vertebrates?	Yes. There was a significant effect ($r = 0.45$) regardless of whether or not activation involved live pathogens	6	13	All	No	Birds, Mammals	Boonekamp et al. (2008)
Can we predict the extent of elevation in androgen levels in males in different situations?	Yes and no. Some predictions of the 'challenge hypothesis' were supported and other refuted.	68	68	Subset (k)	No	All	Hirschenhauser and Oliveira (2006)
Does mounting an immune response reduce survival in birds?	Yes. The mean effect was significant ($r = 0.43$)	7	12	All	No	Birds	Møller and Saino (2004)
Does elevated testosterone increase parasitism and reduce immune function?	No. The mean effect was not significantly different from zero for either relationship	18	22	All	No	Birds, Mammals, Reptiles	Roberts et al. (2004)
Is there a positive relationship between spleen size and disease level or body condition?	Yes. The mean effects were significant for both relationships ($r = 0.10$ and 0.28)	20	20	All	No	Birds	Møller et al. (1998)

Table 1 continued

General theme and main questions	Answer	Species	Studies	Data	Phylogeny	Taxa	Reference
Inbreeding, genetic diversity and fitness							
Is inbreeding depression greater in more stressful environments?	Yes. Studies in the lab are likely to underestimate inbreeding depression	27	33	All	No	All	Fox and Reed (2011)
Is there a positive relationship between heterozygosity and fitness components in wild animal populations?	Yes. There tended to be a significant mean effect ($r < 0.09$) for a range of life history, physiological and morphological traits. The effect was not stronger in inbred populations	61	105	All	No	All	Chapman et al. (2009, 2011)
Is there an effect of inbreeding on key fitness components?	Yes. There was a significant effect of inbreeding on fecundity, first year survival and survival to maturity	6	7	All	No	Birds, Mammals, Fish	O'Grady et al. (2006)
Is there a positive relationship between heterozygosity and fitness components in animals?	Yes and no. There was a significant mean effect ($r = 0.11$) for life history traits, but not for physiological and morphological traits	24	36	All	No	All	Coltman and Slate (2003)
Is there a correlation between genetic diversity and population 'fitness'?	Yes. The mean effect was significant ($r = 0.43$)	21	34	All	No	All	Reed and Frankham (2003)
Does a higher level of heterozygosity lead to higher growth rate or greater symmetry?	Yes. There were significant positive correlations (all $r < 0.20$)	16	65	All *	No	All	Britten (1996)
Do more symmetric individuals have higher levels of heterozygosity?	No. The mean effect did not differ from zero ($r = 0.04$)	20	92	Subset (p)	No	All	Vøllestad et al. (1999)
Miscellaneous							
Does the operational sex ratio predict the Intensity of male-male competition?	Yes and no. The mean effect was significant for aggressiveness, mate guarding and courtship, but not for sperm release	19	27	All *	Yes (8)	All	Weir et al. (2011)
Do partners only partially compensate when their mate provides less parental care in birds?	Yes. There was an increase in care ($d = 1.26$), but the total feeding effort by the pair was lower ($d = 1.11$) so it was only partial compensation	33	52	All	No	Birds	Harrison et al. (2009)

Table 1 continued

General theme and main questions	Answer	Species	Studies	Data	Phylogeny	Taxa	Reference
Are measures of different types of sexual behavior repeatable among individuals?	Yes. Repeatability was significant for all sexual behaviours but much higher for mating, courtship and parental care behaviours than for mating preferences (i.e. choice of same mate)	27	35	All +	No	All	Bell et al. (2009)
Does vegetation density (open vs. closed habitats) select for changes in fine-scale aspects of bird song?	Yes. There was a significant effect of habitat ($r = 0.16$–0.35) for four of five call parameters	–	26	All	No	Birds	Boncoraglio and Saino (2007)
Does breeding failure increase the likelihood of divorce in birds?	Yes. The mean effect was significant ($r = 0.11$). The effect was weaker in species with larger clutches	35	38	All	Yes (4)	Birds	Dubois and Cezilly (2002)

Species: sample sizes are the number of host species in host-parasite studies, except for Matthe et al. (2010) where it is the number of parasite species

Studies: sample sizes are based, where possible, on effects calculated at the species level. The definition of a study differs among papers. It does not necessarily conform to a different publication (e.g. if several species were examined in a single paper, each is treated as a separate study)

Data: 'All' indicates the full data set was first used to calculate a global mean effect (secondarily there were sometimes meta-analyses of subclasses of the data), 'All *' indicates the sample size is from the full data set but this was partitioned into subclasses for meta-analyses so it is an upper estimate of that in individual analyses; 'Subset' indicates the sample size is from the largest dataset used in an individual meta-analysis within the publication; 'All +' indicates the full data set used for sexual selection questions (although others were asked); the subsets were (c) status; (d) size; (e) age; (f) reproductive success; (g) fecundity; (h) ornament-fitness; (i) RHP and RV; (j) female quality; (k) parental behaviour; (l) condition-mite load; (m) mate quality; (n) carotenoids; (o) fitness; (p) PO or haemocyte count; (q) within-population studies; (r) age-EPP

Phylogeny: indicates whether or not there was an attempt to correct for phylogeny where (1) Contrasts of effect sizes (weighted by sample size), (2) variation explained at different taxonomic levels (e.g. family, order, class), (3) used Phylometa (Lajeunesse 2009), (4) Contrasts of effect sizes (but not weighted by sampling variance), (5) BPMM (Hadfield and Nakagawa 2010), (6) tested whether species identity explained variation in effect sizes, (7) used Adams (2008), (8) tested for a phylogenetic signal (Abouheif 1999)

Taxa: 'All' indicates more than three main taxa

References: (a) see also Palmer (1999), Thornhill et al. (1999); (b) for a correction see Cassey et al. (2006)

being equal, only differ in gamete size (Kokko and Jennions 2008). We also considered meta-analyses of the effects of inbreeding and the relationship between heterozygosity and fitness because of recent interest in whether mate choice/biased fertilization improves parental genetic compatibility (Griffith and Immler 2009), or whether selection favours mating with kin (Kokko and Ots 2006) or more heterozygous individuals (review: Kempenaers 2007).

We do not claim to provide a complete list of published sexual selection meta-analyses. Papers with 'buried' meta-analyses are easily overlooked. Indeed, less than half the papers in Table 1 have 'meta-analysis' in the title. We have, however, probably located most meta-analyses that are unambiguously related to core sexual selection questions. To keep our review manageable, however, we excluded meta-analyses exclusively on humans or plants.

The 94 meta-analyses were authored by 179 different researchers, with a mean of 2.7 authors per paper. Interestingly, given the effort required to initially familiarize oneself with how to conduct a meta-analysis, 82 % of these researchers have only coauthored a single meta-analysis. We hope that this is not a sign that the experience was too traumatic to repeat! Only a few researchers published four or more meta-analyses (L. Garamszegi, M. D. Jennions, A. P. Møller, S. Nakagawa, R. Poulin, B. C. Sheldon, R. Thornhill). Another method that shares many similarities with meta-analysis—comparative analysis using phylogenetic contrasts—appears to be far more widely used. For example, Garamszegi and Møller (2010) located 194 comparative analyses from 2003 to mid-2007 in only four journals. The restricted use of meta-analysis is surprising given that published meta-analyses on sexual selection have had great impact. For example, 16 of the meta-analyses in Table 1 have been cited over 100 times, and 14 % are among the top 100 most cited 'meta-analysis' related publications in the ISI defined fields of *ecology*, *evolution*, *plant sciences* and *zoology* (13 of 94 studies), which is double the field average of 7 % (100 of 1,474 'meta-analyses'; see Fig. 1b).

Some of the 94 publications addressed several different questions, and contain multiple meta-analyses using effect sizes drawn from partially overlapping sets of primary studies. Nevertheless, it is possible to use five broad categories to account for two-thirds of the listed meta-analyses. First, 16 studies used phenotypic correlations to identify potential life history trade-offs between male attractiveness/sexual trait expression and naturally selected traits or fitness components. Second, 15 studies tested for sex differences in key life history traits. Third, 12 studies asked which male traits are associated with greater mating success. This provides correlative, and sometimes experimental, evidence that they are targets of female choice. Four, 11 studies related female mate choice, extra-pair mating and polyandry to genetic benefits, direct benefits and naturally selected costs. Five, 10 studies tested whether offspring sex ratios are adjusted in relation to factors that might affect sons' mating success or daughters' breeding success (e.g. mate attractiveness, maternal condition).

Insights and impact of past sexual selection meta-analyses

The impact of published sexual selection meta-analyses ranges from confirming to challenging established theory. In this section we discuss how meta-analysis can help distinguish between hypotheses, what issues arise when effect sizes (which are often small) are used to test theory, and we explain how meta-regression can yield further insights if sufficient data are available.

Distinguishing between predictions or hypotheses

In science, it is often portrayed as ideal to have competing hypotheses that make different testable predictions. In nature, however, multiple factors can simultaneously influence an outcome, so a more balanced approach is to ask which factors tend to drive the observed outcomes rather than trying to completely reject the importance of others. Sexual selection theory is replete with false dichotomies. Sometimes key factors that alter the direction of a relationship are ignored or artificially fixed (e.g. the opposite ends of the Fisher-Zahavi continuum: Kokko et al. 2006b). At other times, two explanatory frameworks are falsely treated as mutually exclusive (e.g. sexual conflict vs. good genes). It is, of course, challenging to confidently state what 'drives' a system, especially when taking into account that multiple factors often interact, albeit with varying importance. Even so, identifying prevalent trends can provide key insights into the relative importance of different factors.

We illustrate this with an example. Two opposing views have been promoted as to whether females should differentially allocate resources to offspring depending on their mating partner's attractiveness. (review: Ratikanen and Kokko 2010). It has been suggested that females mated to more attractive males should increase their reproductive investment if they are likely to produce higher quality offspring with greater reproductive value ('differential allocation': Burley 1986, also widely attributed to Sheldon 2000 who did not specify the direction of adjustment). Alternatively, females mated to less attractive males might increase their reproductive investment to 'compensate' (Gowaty 2008). The unifying consideration is ultimately how parents adjust the rearing environment (i.e. resources provided). Both positive and negative relationships between male attractiveness and female investment are possible (Harris and Uller 2009). Whether one views 'differential allocation' and 'compensation' as differing merely quantitatively or also qualitatively, there is a basic empirical question: which direction occurs more often? Horváthová et al. (2012) found that the mean relationship was for significantly greater investment when mated to attractive males ($r = 0.117$ corrected for phylogeny). This finding should focus attention on the underlying causes of adjustment in effort (i.e. why does 'compensation' tend not to occur?) and, hopefully, will lead to the development of better theoretical models.

We would argue that a general insight from sexual selection meta-analyses is that one can rarely unambiguously distinguish between hypotheses. This is because the relevant hypotheses are not mutually exclusive, and the evidence being summarized is usually correlative rather than experimental. Even so, by indicating prevailing trends in nature it is often possible to estimate the relative importance of different putative causal factors.

When meta-analysis confirms expected patterns, effect sizes are often low

In many cases, meta-analysis is not used to distinguish between competing hypotheses, but simply to confirm (or refute; see next section) an already accepted assumption about an empirical relationship. The value of the exercise is thus to provide an objective measure of the average strength of the relationship. For example, Kelly (2008) asked whether, in territorial species, there is a positive correlation between male resource holding potential (RHP) and the value of the resource held, between resource value and male reproductive success, and finally, given these relationships, a positive relationship between male RHP and reproductive success. As one might expect, all three relationships were indeed significantly positive. However, the mean estimated values of r lay between 0.37 and 0.45, which reminds us that RHP cannot be expected to correlate perfectly with the resources

gained by a male and how many offspring he sires. It would be interesting to follow up and test whether RHP is a better predictor of male success in studies that included males who failed to establish a territory (resource value = 0 for these males) than in those which did not present data on such males. Some selection to maintain high RHP might remain invisible in studies that only consider variation within the subset of territorial males. Overall, however, RHP remains an imperfect predictor of male success. This fits with theoretical and experimental approaches that show that arrival order can override mild RHP differences in deciding territory ownership (Hardy and Field 1998, Kokko et al. 2006a).

In a similar vein, several studies have estimated the mean relationship between mating success and sexual dimorphic traits that are assumed to be under direct sexual selection (e.g. plumage colour, body size, song complexity). Again, the mean relationships are often modest (e.g. $r = 0.30$, Gontard-Danek and Møller 1999; $r = 0.20$, Soma and Garamszegi 2011). These examples remind us that stochasticity might play a large role in determining mating success in most systems (Jennions et al. 2012a).

This conclusion generalizes readily: mean effect sizes are often low in sexual selection studies. Even when the means are statistically significant, a univariate approach rarely explains more than 1–10 % of the variation in a trait of interest (Møller and Jennions 2002). Why are effect sizes so modest? If sexual selection theory is correct, would we not expect greater explanatory power for variables of interest? We offer five responses.

First, the problem could lie in the quality of empirical measures. Above, we highlighted an example: males who never gained a territory/mate are sometimes excluded from primary studies so that the strength of selection is underestimated. More generally, it is notoriously difficult to measure fitness. The use of proxy measures like body condition or annual survival which are fitness components, and not fitness itself, introduces measurement error and, at worst, a systematic bias that must reduce the estimated strength of actual relationships (Hunt and Hodgson 2010). An example of a bias is that lifespan or survival are often poor proxies for fitness if the fittest males have a shorter life expectancy due to trade-offs between sexual signalling and survival (Kokko 2001; Hunt et al. 2004).

Second, the effect sizes obtained will depend on whether the primary researchers used a combination of biological intuition and current theoretical expectations to pick the most relevant variables or whether they inadvertently chose less relevant proxies that will lead to a low effect size. Similarly, if a broad-brush approach is used to identify many potentially relevant factors, some will probably be irrelevant and fail to give any support to theory. For example, in the meta-analysis on RHP and territoriality, Kelly (2008) highlights such problems by discussing data on red-collared widowbirds *Euplectes ardens*. The study predicts $r \approx 1$ if collar colour is used as a determinant of territory size, but including other apparently non-important traits (e.g. tail asymmetry) deflates the mean effect size to $r = 0.52$. It is tempting to, post hoc, exclude traits that do not support the expected relationship, but this can obviously lead to Type I errors. Ultimately, more data is needed to confirm whether certain traits continue to have a low estimated effect size, suggesting that they are indeed unsuited to testing theory in the first place.

Third, most dependent variables in sexual selection studies are affected by multiple factors of interest. Consider, for example, studies where the benefit of polyandry due to elevated offspring fitness is calculated by randomly assigning varying numbers of mates to females (Slatyer et al. 2012). Even if polyandry elevates offspring fitness, there could still be much variation within each mating treatment if, for example, female size and age affect offspring fitness. The situation is even more complex if the benefits of choice are context-dependent so that some females benefit more than others (review: Schmoll 2011). In short,

a distinction must be drawn between the variation explained by a single factor (the goal of many meta-analyses) and the total variation that can be explained in any given primary study using a model building approach with multiple predictors. Peek et al. (2003) estimated that, on average, statistical models in primary ecological studies explain 47–54 % of the observed variation.

Fourth, we suggest that the role of stochastic events ('luck' in plain language) is underestimated in sexual selection studies. It is likely to generate considerable variation in mating success, especially when the sex ratio is male biased. Mating is a binary event (mate or not) so continuous variation in a male trait is imperfectly correlated with mating success (for an illustrative example see Jennions et al. 2012a; see also Kahneman 2011 for a general review of the propensity to underestimate how chance affects success). For this reason alone, effect sizes are likely to be modest in studies investigating male mating success.

Fifth, there are genuine biological differences between populations, sites and species and environmental conditions across years. The direction of a relationship can genuinely vary. In such cases, the mean effect must be smaller than that expect based on studies that show the strongest support for theory (which are often the best known studies). These situations call for an analysis of moderating variables, which we will return to shortly.

So what should we make of low effect sizes? The most important implication is that biological systems are strongly influenced by noise so that most primary studies have minimal statistical power. Failure to detect a significant influence of a factor of interest should be commonplace, regardless of whether or not it has an impact on the measured outcome. This hard truth makes it exceedingly difficult to interpret a null result in any single study: was theory refuted or was the effect simply too weak to detect despite ultimately being evolutionarily significant?

These difficulties place into perspective apparent discrepancies between studies that do and do not obtain a significant result. Historically, such differences evoke conflict: researchers often explain them away by invoking biological differences (e.g. geographic variation) or criticize the inferior methodology of rival researchers. The use of meta-analysis discourages the dichotomous interpretation of *P* values and encourages greater consideration of the distribution of the observed magnitude of a relationship. It is worthwhile remembering that studies of the same phenomena will, even in the absence of biological variation, generate a range of effect sizes purely due to sampling error (Nakagawa and Cuthill 2007). This subtle shift in outlook could profoundly direct emphasis away from over-reliance on extrapolating from individual studies and towards detecting general trends using meta-analysis.

When meta-analysis refutes conventional wisdom, new alternatives are often identified

Meta-analysis does not always confirm a prior prejudice. A general insight from sexual selection meta-analyses (see the occasional 'No' answer in Table 1) is that it is easy to be misled by a few high profile studies into believing that a prediction is well supported. Support is often weaker than assumed.

To start with a simple example, meta-analysis has shown that, despite some high profile primary studies that reported a significant effect of certain male traits on mating success, the mean effect did not differ from zero when looking at all the available studies. For example, Nakagawa et al. (2007) found no significant relationship between bib size and reproductive success in house sparrows *Passer domesticus* in a meta-analysis of results from 12 populations. Similarly, Garamszegi and Møller (2004) found that song complexity

did not predict a male's within-pair paternity, although Soma and Garamszegi (2011) subsequently showed that song complexity was, on average, weakly but significantly positively correlated with total reproductive success ($r = 0.20$) (but this was not readily attributable to increased success at gaining extra-pair paternity).

More important, perhaps, are cases where meta-analysis fails to detect a theoretically predicted relationship that is assumed to be widely corroborated. For example, Roberts et al. (2004) tested some key assumptions of the immunocompetence handicap hypothesis (ICCH; Folstad and Karter 1992). The ICCH suggests that a major cost constraining elaboration of sexually selected traits in vertebrates, making their expression covary with male quality, is that testosterone elevates sexual trait expression but reduces the ability to repel parasites. There was, however, no evidence from experimental manipulations of testosterone levels that testosterone is immunosuppressive (based on immune function measures like white blood cell counts or subsequent parasite loads) when data were analyzed at the species level. In another meta-analysis, Boonekamp et al. (2008) showed, albeit with a very small data set, a potential causal link in the *reverse* direction: experimental immune system activation suppresses testosterone levels. The primary studies they analysed had previously received little attention, and it is unlikely that a case for this 'reverse causality' explanation could have been built without a meta-analysis. Given the implications of these findings for the well-studied ICCH, it is surprising that no follow up meta-analyses have re-examined the findings of these two meta-analyses.

Another example can be found in game theory models of sperm competition. Models predict that males will increase ejaculate size when sperm competition risk is high (e.g. when a rival is also likely to mate with the female), but decrease ejaculate size when the intensity of sperm competition is greater (e.g. when many rivals are present because the marginal returns per sperm decline). Two meta-analyses recently confirmed the 'risk' prediction (Kelly and Jennions 2011; DelBarco-Trillo 2011), but for the 'intensity' prediction, the mean effect size did not differ from zero (Kelly and Jennions 2011). If sperm competition theory is internally correct, such that the conclusions follow from the model assumptions, then a failure to meet the predictions can highlight two problems: the assumptions might not be met in nature and/or there is a problem in how theory is being tested or applied. It might be that, while effects of 'risk' and 'intensity' both exist due to high measurement error, by chance only one has been detected. Alternatively, males might not perceive the experimental manipulation as intended by the researcher. For example, focal males might not use brief exposure to different numbers of males as a measure of the likely intensity of sperm competition. It could also be that many tests are simply on species where males do not facultatively adjust ejaculates to short-term changes in the intensity of sperm competition, thereby deflating the overall effect size.

This last point is an important consideration: is sexual selection theory as a whole incorrect if specific predictions do not apply in some species? Consider the fact that, outside of the insects, hermaphroditic species are common but never appear to have precopulatory ornaments (Lukas Schärer, pers.comm). It would be counterproductive to include many hermaphroditic species in a meta-analysis testing predictions about ornament evolution as it could 'dilute' estimates of the mean effect for more appropriate study taxa.

This is one reason why it is worthwhile looking for heterogeneity in effect sizes and attempting to identify sources of variation in effect sizes. A theory is not best supported when effect sizes are strongest, but when effect sizes can be predicted based on explanatory variables. This is why it is also worthwhile to quantify the magnitude of phylogenetic effects (Lajeunesse 2009). A strong phylogenetic signal indicates that the theory being tested is potentially more applicable to some taxa (those where the effect is stronger) than

others. If so, researchers can address a new set of questions. Could historic contingencies, combined with phylogenetic inertia, be sufficient to create a pattern where some taxa meet the theory's assumptions better than others? For example, in sex allocation studies some taxa might simply lack any mechanisms that allow them to bias the offspring sex ratio. Alternatively, there could be predictable variation in species properties across taxa that make it readily understandable why evolutionary outcomes follow theoretical predictions better in some lineages than others (e.g. local mate competition could have a stronger impact on sex ratios in taxa that lack efficient dispersal).

Meta-regression: understanding variation in effect sizes

Meta-analysis provides us with an overview of the location and distribution of effect sizes (i.e. mean and associated confidence interval), and if the number of species is small this might be the only realistically achievable goal. Many early studies simply reported the mean effect size and, at best, calculated effect sizes for a few subsets of the data (i.e. tested the influence of categorical factors). However, when sample sizes permit, meta-regression provides techniques to investigate differences among studies. This helps to overcome the problem outlined earlier that low mean effect sizes do not necessarily mean that there is nothing biologically interesting going on. Meta-regression can be used to test whether variation in effect sizes among studies is not solely attributable to sampling error (i.e. there is significant heterogeneity) and can partly be explained by biological and/or methodological moderators. The use of meta-regression with multiple predictors has become more common in recent sexual selection studies following recent software advances (see Nakagawa and Santos 2012), and we highlight its importance with three examples.

Identifying sources of variation can prevent misinterpretation of a non-significant or low mean effect as evidence that underlying theory does not apply in nature. Consider the general finding that birds increase their reproductive effort when mated to attractive males (Horváthová et al. 2012). Closer inspection showed that the magnitude of the 'attractiveness' effect varies for different types of maternal investment. There was a significant increase for some traits (clutch size, egg size and maternal feeding rate), but not others (levels of immuno-stimulants and androgens in eggs). Given the problem outlined earlier that including many irrelevant traits can dilute a real effect to yield a non-significant mean effect, it is an interesting conjecture to consider what would have happened if many more studies had existed on androgens, and the true effect for androgen is zero. Would the initial 'null' result based on the full dataset have discouraged the researchers enough to make them forgo the more advanced step of meta-regression? One hopes not, as the meta-regression not only revealed effects of 'attractiveness' on clutch size, egg size and feeding responses, but also exposed an interesting difference between species with bi-parental and female-only care. Species with female-only care showed a significantly greater propensity to increase egg size, while those with bi-parental care tended to increase clutch size. This finding makes intuitive sense if the ability of a single parent to care for a larger brood is limited. The finding also raises wider questions about the likely patterns that will arise in non-avian taxa with different patterns of parental care that can now be tested based on a priori predictions (Ratikanen and Kokko 2010). [For another good example of a study that shows how a non-significant mean effect is not necessarily a refutation of theory see Griffin et al. (2005). They show that facultative adjustment of offspring sex in cooperative breeders in response to the number of existing helpers can be explained by variation in the extent to which helpers actually elevate parental fitness].

Another example is provided by studies of extra-pair mating in birds. The cost of extra-pair mating due to reduced parental care by a cuckolded male varies widely among species. Small genetic benefits (Møller and Alatalo 1999; Slatyer et al. 2012) have led some to argue that extra-pair activity is male driven (Arnqvist and Kirkpatrick 2005, but see Eliassen and Kokko 2008; Griffith 2007; Schmoll 2011 for problems of measuring and interpreting the patterns obtained). If male behaviours (seeking and protecting paternity) were the only factors determining the proportion of extra-pair young, effect sizes should not predictably vary with the fitness consequences for females. However, Albrecht et al. (2006) showed in a phylogenetic controlled meta-analysis that rates of extra-pair fertilization decrease with increasing costs (due to lower male care) of multiple paternity for females. As with any across-study or species comparison, the evidence is correlative, but this finding clearly makes it difficult to maintain that the behaviour of only one sex (female or male) determines the distribution of paternity.

Finally, meta-regression can highlight important methodological differences between studies. We have already mentioned, in the context of RHP and territoriality (Kelly 2008), the potential for a meta-regression to estimate how important it is to include 'failed' males with no resources when estimating selection on RHP. In the context of sexual conflict over parental care, there was significant among-study heterogeneity in the compensatory increase in the focal parent's feeding rate of offspring in response to reduced care by the partner (Harrison et al. 2009). Inclusion of experimental treatment as a moderating factor showed that the increase in care was greater when the mating partner was removed as opposed to experimentally manipulated so that his/her level of care was reduced ($d = 1.69$ vs. 0.50). This study highlights the basic importance of taking into account the role of methodology when estimating the strength of the relationship.

Great potential, but do we have enough species?

Attempts to evaluate the generality of an argument, and identify causal factor relies on sufficient coverage of evolutionary outcomes from many taxa. As with much sexual selection research, there is a strong bias towards meta-analyses exclusively of birds studies (36 % in Table 1), with far fewer meta-analyses exclusively devoted to mammals (11 %) or arthropods (14 %). More than 34 % of the meta-analyses in Table 1 did, however, use data sets covering four or more major taxa (e.g. birds, mammals, reptiles, fish, frogs, amphibians, insects or arthropods), but closer inspection shows that there was often still a strong sampling bias towards one or two taxa.

Of the 94 studies, 10 were single species studies. Of the remainder, there were 75 studies that provided detailed enough information about the number of species used to tabulate it (Table 1). We noted the number of species in the largest meta-analysis per publication, if the original publication clearly estimated separate mean effects for different subsets of the data (which could vary in the number of species for which data were available). If it was not possible for us to deduce the number of species of the most extensive meta-analysis in this way, we have simply noted the total number of species, and it should be kept in mind that the number of species in each meta-analysis might then be smaller.

The mean and median numbers of species in these 75 meta-analyses were 29.2 and 20, respectively. Only 11 publications included data on more than 50 species. These numbers should alert the reader to limitations of current datasets. As we have noted earlier, the value of a meta-analysis is much enhanced when there is sufficient data to test for moderating variables. Such an approach allows researchers to address questions of the type 'why does

our phenomenon of interest appear stronger in studies/species where, say, populations are at higher density' which, if answered, advance a field more than a simple statement that 'the phenomenon appears to exist (or not)'. The capacity to test for moderating variables is limited due to low statistical power if we have only a few cases at hand. It is important to consider the level of analysis when testing for moderators (e.g. do we want to look for variation across studies or species?). In general, especially given the need to formally correct for phylogenetic relatedness (Lajeunesse 2009), we would argue that most researchers are interested in explaining variation across species. In meta-analysis, the statistical power to test whether the mean effect differs from the null value depends on study samples sizes that affect the precision of estimates of effect size and how variation among estimates is then distributed at different hierarchical levels (e.g. within and among species variation). This is information that we do not present in Table 1, partly because large differences in sample sizes among studies can make the interpretation of the mean sample size per study misleading. In contrast, simply put, the ability to detect significant moderators of variation across species will depend on the number of species examined (no matter how precisely each species mean value is estimated).

Handy hints to find a sexual selection topic to meta-analyse

Here, we assist those keen to conduct a meta-analysis but unsure where to begin. We provide general strategies to identify fruitful areas for meta-analysis and identify neglected areas in sexual selection.

Coming up with ideas for meta-analyses: some random thoughts

There are some simple strategies to identify areas where data exists, but meta-analyses are lacking. Reviews that include tables that 'vote count' studies that did or did not report a significant result usually indicate that there is sufficient empirical data to conduct a meta-analysis. More generally, any area that is the subject of extensive narrative reviews with tables of studies is likely to be ripe for meta-analysis [e.g. condition-dependence of female mate choice (Cotton et al. 2006); whether male-male competition and female choice select for different traits (Hunt et al. 2009)]. It is generally fruitful to identify statements that are repeatedly encountered but only backed up by citing specific, often high-profile, studies. This usually suggests that there is, as yet, no published meta-analysis. Even if the statements are 'obviously' true, it is unlikely that they have been objectively validated, or the relative magnitude of the relationship quantified. For example, it is often claimed that there is an 'ownership advantage' or 'prior residency effect' during territorial fights (Kokko et al. 2006a). Table A1 in Kokko et al. (2006a) suggests this is true but, if so, there is still scope to account for variation in the size of the residency effect which is important to explain the evolution of residency and migration (Kokko 2011).

Given similarities between meta-analysis and comparative analysis it should also be obvious that almost every published comparative analysis could be re-analyzed and 'converted' into a meta-analysis simply by taking into account sampling variance/measurement error (Garamszegi and Møller 2010).

Already existing meta-analyses should be periodically updated. Many early sexual selection meta-analyses had small sample sizes, with little potential to test for moderator factors, and estimates of mean effects had broad confidence intervals. As data accumulates

it is possible to address these limitations (e.g. compare Coltman and Slate 2003 and Chapman et al. 2009). Inspection of Table 1 reveals some obvious cases of meta-analyses that could be revisited. Another reason to redo earlier meta-analyses is that they were often generous in the types of primary studies included. Specifically, experimental and observational studies were often pooled, when only the former provide strong evidence for causality. The opportunity therefore exists to conduct meta-analyses with more restrictive datasets (ideally only experiments) to reach more robust conclusions about causality. For example, Slatyer et al. (2012) tested whether polyandry confers genetic benefits using only experimental studies where the number of mates was varied and the number of matings per female stayed constant. In contrast, Møller and Alatalo (1999) looked at a range of sources of experimental and correlative evidence for 'good genes' due to mate choice.

It is not always necessary to think big. Some topics are so well studied that it is an overwhelming task to conduct a meta-analysis (e.g. the effect of body size on fight outcome). In such cases, a more restrictive meta-analysis is a pragmatic solution, such as confining the meta-analysis to certain taxa. It is also worth noting that a meta-analysis does not have to resolve major theoretical questions. As with primary studies there is value to tackling modest questions. A good start is to list the basic questions you ask in your own study system. So, for example, based on our own empirical work we might ask: Do larger fiddler crabs wave faster than smaller ones? Do male and female crickets differ in their ability to withstand an immune challenge? How are different measures of immune function in insects correlated? Do male fish avoid larger sexual competitors when deciding which females to court? If you work on a well-studied species remember that even species-specific questions can be meta-analysed (see Table 1). At the extreme, it can even be informative to conduct a meta-analysis of a single study system if the same type of data is repeatedly collected over time (or space) (e.g. Milner et al. 2010).

Neglected areas in sexual selection

Any mismatch between the topics covered in Table 1 and how many primary studies there are in these areas provides a clue as to which areas are neglected. Here we provide a 'shopping list'.

1. Intra-sexual selection: Fighting behaviour is extensively studied and there are many theoretical models (Briffa and Sneddon 2010), but we did not locate meta-analyses explicitly addressing these models. Obvious topics are: the winner/loser effect; the residency advantage; the role of body or weapon size in determining fight outcome; the relationship between body size differences and fight duration or escalation.
2. What traits show a life history trade-off with greater male attractiveness? Phenotypic correlations between male attractiveness and many key life history traits have been well researched but the general trends, and sources of variation, have not been identified by meta-analysis. For example, what is the relationship between attractiveness and: social dominance, post-copulatory reproductive success (i.e. fertilization under sperm competition) or metabolic rate? Ideally, one should address questions about evolutionary trade-offs using data on genetic correlations (e.g. Evans 2010) or effects of experimental manipulations (Reznick et al. 2000). Meta-analyses of these data provide the strongest evidence for which trade-offs are most important.

3. What are the average values of key genetic parameters measured in sexual selection studies (review: Chenoweth and McGuigan 2010)? For example, what is the mean difference between the axis of directional selection on male sexual traits and the axis of maximum additive genetic variation in attractiveness (see Blows 2007)? What is the mean heritability of specific male traits? Can we explain variation in the genetic correlation between female mating preferences and preferred male traits or net attractiveness? What is the genetic correlation between male attractiveness and lifespan, or between the strength of female mating preferences and lifespan?
4. Many studies measure phenotypic selection (review: Kingsolver et al. 2012), but there has been little exploration of the importance of the main population parameters invoked by theoreticians to account for variation in the intensity of sexual selection (but see Weir et al. 2011). For example, how well do the adult or operational sex ratio, variance in male mating success, population density, the difference in 'time out' between the sexes after mating, the division of parental care and type of 'mating system' explain variation in the strength of selection on sexual traits. More specifically, can these parameters account for variation in the level of mate choosiness or of direct physical competition for mates?
5. There are many statements in the literature about key differences between sexual and non-sexual traits in allometry, phenotypic variability and level of phenotypic plasticity (condition-dependence) that have not yet been meta-analysed (*reviews*: Pomiankowski and Møller 1995; Cotton et al. 2004; Kodric-Brown et al. 2006: Bonduriansky 2007).
6. Several experimental evolution studies have investigated the evolution of traits in both sexes in the presence/absence of sexual selection (i.e. enforced monogamy vs. polygamy) (review: Edward et al. 2010). The average effect of removing sexual selection on different types of traits has not yet been formally quantified.
7. There is much interest in the effect of inbreeding on mate choice. Some key questions are: Do females prefer heterozygous males based on either conventional mate choice or biased paternity when mating multiply (Kempenaers 2007)? Does inbreeding have a more detrimental effect on sexually than naturally selected traits (Cotton et al. 2004)? Do females mate with or bias paternity towards non-related males (i.e. avoid inbreeding)? If not, is this because inclusive fitness gains outweigh the direct costs of inbreeding depression (Kokko and Ots 2006)?
8. An extreme form of phenotypic plasticity in relation to sexual selection is to change sex. We are, however, unaware of any meta-analyses that directly investigate how the timing or direction of sex change is related to factors of theoretical importance (but see Molloy et al. 2008 for a meta-analysis of an intriguing pattern).
9. Several meta-analyses have investigated which secondary sexual traits predict male mating success (see Table 1). But what predicts success under sperm competition? There are now several ejaculate or sperm traits that seem to be of potential importance (Snook 2005), but their relative influence has not been quantified using meta-analysis.
10. There are general claims about the types of individuals that are preferred as mates that have not been subject to a meta-analysis. For example, do females prefer older males? Do males prefer virgin females or larger females? Is the Coolidge effect common?
11. Sexual selection theory makes predictions about how much each sex will invest in parental care (Kokko and Jennions 2008), but there is no theory to predict the total reproductive effort by each sex. Do the reproductive budgets of females tend to exceed those of males (see Hayward and Gillooly 2011)?

What next?

We conclude by highlighting a few general issues. While it is tempting to conduct a large and complicated meta-regression with several predictor variables, it is worth bearing in mind the available data. In general, we think it is prudent to think in terms of the number of species in the database when trying to generalise. A meta-regression should be no more complicated than the equivalent model one would run for a primary analysis with repeated measurement of individuals to explain individual level traits (i.e. treat species as analogous to individuals). Although it is possible to look at moderator effects at lower levels using a hierarchical model (Hadfield and Nakagawa 2010), the important theoretical questions in most meta-analyses are nearly always about the distribution of effect sizes across species, and the power to test species level hypotheses depends on sampling at the appropriate level.

When interpreting a meta-analysis one should be cautious in drawing inferences from non-significant results given low statistical power and the possibility that studies of irrelevant species, traits or sets of conditions 'dilute' the effect size and makes it difficult to detect whether the theory is sometimes applicable (or is generally applicable but only yields a non-zero effect in some circumstances). The opposite problem applies if researchers cherry-pick study species that appear most likely to support their favourite theory. For example, in sexual selection studies there is a tendency to work on species that seem most likely to generate the observed pattern. When studying mate choice, researchers have historically favoured species that are sexually dimorphic, lek or mate polygynously. Likewise, studies of 'sexual conflict' tend to be on species where males are readily observed harassing females. A related problem is an understandable tendency to work on common species where large sample sizes can be collected. 'Missing' species and variation in sampling effort are consequently non-randomly distributed across phylogenies (Garamszegi and Møller 2011).

Several of the meta-analyses in Table 1 look at numerous moderator variables (and even their interactions). The risk of 'statistical fishing' arises. Even if the researchers used, say, information criteria to select the best model, this does not eliminate the underlying problem that when more moderators are examined there is a greater chance that the final model will contain some moderators with a significant effect due to type I errors (review: Forstmeier and Schielzeth 2011; Freckleton 2009; for the use of model averaging see Grueber et al. 2011). Finally, as with any statistical model, the reader should ask whether the meta-regression was tested for outliers, or points that had undue leverage. Given low sample sizes, sensitivity analyses seem appropriate.

The existence of publication bias remains controversial (reviews: Jennions et al. 2012b; Nakagawa and Santos 2012). Selective reporting of results is clearly an issue that can affect meta-analysis (e.g. Cassey et al. 2004). In sexual selection studies, however, we suspect that narrow sense publication bias is unlikely to be a major problem, because of the wide range of views held by researchers (e.g. whether or not 'good genes' are important). The one notable exception to this statement is that there is likely to be an issue with genetic parameter estimates. Quantitative genetic experiments that report negative heritabilities or those close to zero are probably less likely to be published. This is not least because current research is focused on questions about multivariate genetic variation, genetic correlations and 'constraints' on evolution (review: Blows 2007). A lack of additive genetic variation means that there is no point reporting genetic correlations or engaging in more sophisticated analyses. This undermines the main framework used to write up quantitative genetic studies of sexual selection. Of course, claims about publication bias need to be directly

tested. There are many indirect tests (see Nakagawa and Santos 2012), but these are always open to alternative explanations and are often inapplicable if there is heterogeneity in effect sizes among the studies being analyzed (which is true in most cases). Direct tests that track the publication fate of completed studies are required. To date, there has been only one direct test for publication bias in sexual selection (Møller et al. 2005).

Finally, meta-analysis is a powerful way to summarize what we know, but it should not blind us to other sources of evidence. It is easy to become blinkered and assume that the only useful data is that collected in a uniform manner, and amenable to statistical tests that allow effect sizes can be calculated. There are, however, often lines of evidence based on idiosyncratic experimental approaches that illuminate the general importance of a factor of interest, but are unsuited to inclusion in a formal meta-analysis (see table 6 in Slatyer et al. 2012). Similarly, there are cases where a mechanistic understanding of a biological phenomenon might provide a more powerful explanation for a relationship than the 'black box' approach often taken in sexual selection studies of gathering correlative data. Ideally, these forms of evidence should be presented in a synthetic review as a complement to quantitative meta-analysis. Systematic, methodological advances help science to become more rigorous, but progress is never an entirely mechanical or statistical endeavour.

Acknowledgments We thank Isobel Booksmythe and Brian Mautz for feedback on a draft of this paper. We thank Lazlo Garamszegi and an anonymous reviewer for their constructive comments. We thank the many colleagues who have chatted with us about meta-analyses, most notably the members of the NCEAS working group on meta-analysis in ecology and evolution (especially Kerrie Mengersen, Jessica Gurevitch and Julia Koricheva). Finally, we extend special thanks to Shinichi Nakagawa for his patience.

References

Abouheif E (1999) A method for testing the assumption of phylogenetic independence in comparative data. Evol Ecol Res 1:895–909

Adams DC (2008) Phylogenetic meta-analysis. Evolution 62:567–572

Akcay E, Roughgarden J (2007) Extra-pair paternity in birds: review of the genetic benefits. Evol Ecol Res 9:855–868

Albrecht T, Kreisinger J, Piálek J (2006) The strength of direct selection against female promiscuity is associated with rates of extrapair fertilizations in socially monogamous songbirds. Am Nat 167:739–744

Arnqvist G, 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:S26–S37

Arnqvist G, Nilsson T (2000) The evolution of polyandry: multiple mating and female fitness in insects. Anim Behav 60:145–164

Arnqvist G, Wooster D (1995) Metaanalysis—synthesizing research findings in ecology and evolution. Trends Ecol Evol 10:236–240

Arnqvist G, Rowe L, Krupa JJ, Sih A (1996) Assortative mating by size: a meta-analysis of mating patterns in water striders. Evol Ecol 10:265–284

Beirinckx K, Van Gossum H, Lajuenesse M, Forbes MR (2006) Sex biases in dispersal and philopatry: insights from a meta-analysis based on capture-mark-recapture studies of damselflies. Oikos 113:539–547

Bell AM, Hankison SJ, Laskowski KL (2009) The repeatability of behavior: a meta-analysis. Anim Behav 77:771–783

Bernasconi G, Hellriegel B (2005) Fertilization competence and sperm size variation in sperm-heteromorphic insects. Evol Ecol 19:45–54

Blackburn TM, Cassey P, Lockwood JL (2009) The role of species traits in the establishment success of exotic birds. Glob Change Biol 15:2852–2860

Blanckenhorn WU, Stillwell RC, Young KA, Fox CW, Ashton KG (2006) When Rensch meets Bergmann: does sexual size dimorphism change systematically with latitude? Evolution 60:2004–2011
Blows M (2007) Complexity for complexity's sake? J Evol Biol 20:39–44
Boncoraglio G, Saino N (2007) Habitat structure and the evolution of bird song: a meta-analysis of the evidence for the acoustic adaptation hypothesis. Funct Ecol 21:134–142
Bonduriansky R (2007) Sexual selection and allometry: a critical reappraisal of the evidence and ideas. Evolution 61:838–849
Boonekamp JJ, Ros AHF, Verhulst S (2008) Immune activation suppresses plasma testosterone level: a meta-analysis. Biol Lett 4:741–744
Briffa M, Sneddon LU (2010) Contest behaviour. In: Westneat DF, Fox CW (eds) Evolutionary behavioural ecology. Oxford University Press, Oxford, pp 248–265
Britten HB (1996) Meta-analyses of the association between multilocus heterozygosity and fitness. Evolution 50:2158–2164
Brown GR, Silk JB (2002) Reconsidering the null hypothesis: is maternal rank associated with birth sex ratios in primate groups? Proc Natl Acad Sci USA 99:11252–11255
Burley N (1986) Sex-ratio manipulation in color-banded populations of zebra finches. Evolution 40:1191–1206
Cameron EZ (2004) Facultative adjustment of mammalian sex ratios in support of the Trivers–Willard hypothesis: evidence for a mechanism. Proc R Soc B 271:1723–1728
Candolin U, Heuschele J (2008) Is sexual selection beneficial during adaptation to environmental change? Trend Ecol Evol 23:446–452
Cassey P, Ewen JG, Blackburn TM, Møller AP (2004) A survey of publication bias within evolutionary ecology. Proc R Soc B 271:S451–S454
Cassey P, Ewen JG, Møller AP (2006) Facultative primary sex ratio variation: a lack of evidence in birds (vol 271 pg 1277 2004). Proc R Soc B 273:3129–3130
Castellanos MC, Verdú M (2012) Meta-analysis of meta-analyses in plant evolutionary ecology. Evol Ecol (in press)
Chapman JR, Nakagawa S, Coltman DW, Slate J, Sheldon BC (2009) A quantitative review of heterozygosity-fitness correlations in animal populations. Mol Ecol 18:2746–2765
Chapman JR, Nakagawa S, Coltman DW, Slate J, Sheldon BC (2011) A quantitative review of heterozygosity-fitness correlations in animal populations (correction). Mol Ecol 20:2655
Chenoweth SF, McGuigan K (2010) The genetic basis of sexually selected variation. Ann Rev Ecol Evol Syst 41:81–101
Cleasby IR, Nakagawa S (2012) The influence of male age on within-pair and extra-pair paternity in passerines. Ibis (in press)
Coltman DW, Slate J (2003) Microsatellite measures of inbreeding: a meta-analysis. Evolution 57:971–983
Cotton S, Fowler K, Pomiankowski A (2004) Do sexual ornaments demonstrate heightened condition-dependent expression as predicted by the handicap hypothesis? Proc R Soc B 271:771–783
Cotton S, Small J, Pomiankowski A (2006) Sexual selection and condition-dependent mate preferences. Curr Biol 16:R755–R765
DelBarco-Trillo J (2011) Adjustment of sperm allocation under high risk of sperm competition across taxa: a meta-analysis. J Evol Biol 24:1706–1714
Dubois F, Cezilly F (2002) Breeding success and mate retention in birds: a meta-analysis. Behav Ecol Sociobiol 52:357–364
Edward DA, Fricke C, Chapman T (2010) Adaptations to sexual selection and sexual conflict: insights from experimental evolution and artificial selection. Phil Trans R Soc B 365:2541–2548
Eliassen S, Kokko H (2008) Current analyses do not resolve whether extra-pair paternity is male or female driven. Behav Ecol Sociobiol 62:1795–1804
Evans JP (2010) Quantitative genetic evidence that males trade attractiveness for ejaculate quality in guppies. Proc R Soc B 277:3195–3201
Evans SR, Hinks AE, Wilkin TA, Sheldon BC (2010) Age sex and beauty: methodological dependence of age- and sex-dichromatism in the great tit *Parus major*. Biol J Linn Soc 101:777–796
Ewen JG, Cassey P, Møller AP (2004) Facultative primary a lack of evidence sex ratio variation: a lack of evidence in birds? Proc R Soc B 271:1277–1282
Fiske P, Rintamaki PT, Karvonen E (1998) Mating success in lekking males: a meta-analysis. Behav Ecol 9:328–338
Folstad I, Karter AJ (1992) Parasites, bright males, and the immunocompetence handicap. Am Nat 139:603–622
Forsgren E, Amundsen T, Borg ÅA, Bjelvenmark J (2004) Unusually dynamic sex roles in a fish. Nature 429:551–554

Forstmeier W, Schielzeth H (2011) Cryptic multiple hypotheses testing in linear models: overestimated effect sizes and the winner's curse. Behav Ecol Sociobiol 65:47–55
Fox CW, Reed DH (2011) Inbreeding depression increases with environmental stress: an experimental study and meta-analysis. Evolution 65:246–258
Freckleton RP (2009) The seven deadly sins of comparative analysis. J Evol Biol 22:1367–1375
Garamszegi LZ (2005) Bird song and parasites. Behav Ecol Sociobiol 59:167–180
Garamszegi LZ, Eens M (2004) Brain space for a learned task: strong intraspecific evidence for neural correlates of singing behavior in songbirds. Brain Res Rev 4:187–193
Garamszegi LZ, Møller AP (2004) Extrapair paternity and the evolution of bird song. Behav Ecol 15:508–519
Garamszegi LZ, Møller AP (2010) Effects of sample size and intraspecific variation in phylogenetic comparative studies: a meta-analytic review. Biol Rev 85:797–805
Garamszegi LZ, Møller AP (2011) Nonrandom variation in within-species sample size and missing data in phylogenetic comparative studies. Syst Biol 60:876–880
Garamszegi LZ, Torok J, Hegy G, Szöllõs E, Rosivall B, Eens M (2007) Age-dependent expression of song in the collared flycatcher, *Ficedula albicollis*. Ethology 113:246–256
Gontard-Danek MC, Møller AP (1999) The strength of sexual selection: a meta-analysis of bird studies. Behav Ecol 10:476–486
Gowaty PA (2008) Reproductive compensation. J Evol Biol 21:1189–1200
Griffin AS, Sheldon BC, West SA (2005) Cooperative breeders adjust offspring sex ratios to produce helpful helpers. Am Nat 166:628–632
Griffith SC (2007) The evolution of infidelity in socially monogamous passerines: Neglected components of direct and indirect selection. Am Nat 169:274–281
Griffith SC, Immler S (2009) Female infidelity and genetic compatibility in birds: the role of the genetically loaded raffle in understanding the function of extrapair paternity. J Avian Biol 40:97–101
Griffith SC, Parker TH, Olson VA (2006) Melanin-versus carotenoid-based sexual signals: is the difference really so black and red? Anim Behav 71:749–763
Grueber CE, Nakagawa S, Laws RJ, Jamieson IG (2011) Multimodel inference in ecology and evolution: challenges and solutions. J Evol Biol 24:699–711
Gurevitch J, Morrow LL, Wallace A, Walsh JS (1992) A metaanalysis of competition in field experiments. Am Nat 140:539–572
Hadfield JD, Nakagawa S (2010) General quantitative genetic methods for comparative biology: phylogenies, taxonomies and multi-trait models for continuous and categorical characters. J Evol Biol 23:494–508
Hamilton WJ, Poulin R (1997) The Hamilton and Zuk hypothesis revisited: a meta-analytical approach. Behaviour 134:299–320
Hardy ICW, Field SA (1998) Logistic analysis of animal contests. Anim Behav 56:787–792
Harper DGC (1999) Feather mites pectoral muscle condition wing length and plumage coloration of passerines. Anim Behav 58:553–562
Harris WE, Uller T (2009) Reproductive investment when mate quality varies: differential allocation versus reproductive compensation. Phil Trans R Soc B 364:1039–1048
Harrison F, Barta Z, Cuthill I, Székely T (2009) How is sexual conflict over parental care resolved? A meta-analysis. J Evol Biol 22:1800–1812
Hayward A, Gillooly JF (2011) The cost of sex: quantifying energetic investment in gamete production by males and females. PlosONE 6:e16557
Hirschenhauser K, Oliveira RF (2006) Social modulation of androgens in male vertebrates: meta-analyses of the challenge hypothesis. Anim Behav 71:265–277
Horváthová T, Nakagawa S, Uller T (2012) Strategic female reproductive investment in response to male attractiveness in birds. Proc R Soc B 279:163–170
Hunt J, Hodgson D (2010) What is fitness, and how do we measure it? In: Westneat DF, Fox CW (eds) Evolutionary behavioural ecology. Oxford University Press, Oxford, pp 46–70
Hunt J, Brooks R, Jennions MD, Smith MJ, Bentsen CL, Bussière LF (2004) High-quality male field crickets invest heavily in sexual display but die young. Nature 432:1024–1027
Hunt J, Breuker CJ, Sadowski JA, Moore AJ (2009) Male-male competition, female mate choice and their interaction: determining total sexual selection. J Evol Biol 22:13–26
Jennions MD, Møller AP (2002) Relationships fade with time: a meta-analysis of temporal trends in publication in ecology and evolution. Proc R Soc B 269:43–48
Jennions MD, Møller AP, Petrie M (2001) Sexually selected traits and adult survival: a meta-analysis. Q Rev Biol 76:3–36

Jennions MD, Kokko H, Klug H (2012a) The opportunity to be misled in studies of sexual selection. J Evol Biol 25:591–598
Jennions MD, Lortie C, Koricheva J (2012b) Using meta-analysis to test ecological and evolutionary theory. In: Koricheva J, Gurevitch J, Mengersen K (eds) Handbook of meta-analysis in ecology and evolution. Princeton University Press, Princeton (in press)
Jennions MD, Lortie C, Rosenberg M, Rothstein H (2012c) Publication and related biases. In: Koricheva J, Gurevitch J, Mengersen K (eds) Handbook of meta-analysis in ecology and evolution. Princeton University Press, Princeton (in press)
Jones KS, Nakagawa S, Sheldon BC (2009) Environmental sensitivity in relation to size and sex in birds: meta-regression analysis. Am Nat 174:122–133
Kahneman D (2011) Thinking, fast and slow. Farrar, Straus and Giroux, New York
Kelly CD (2008) The interrelationships between resource-holding potential, resource-value and reproductive success in territorial males: how much variation can we explain? Behav Ecol Sociobiol 62:855–871
Kelly CD, Jennions MD (2011) Sexual selection and sperm quantity: meta-analyses of strategic ejaculation. Biol Rev 86:863–884
Kempenaers B (2007) Mate choice and genetic quality: a review of the heterozygosity theory. Adv Study Behav 37:189–278
Kingsolver J, Diamond S, Siepelski A, Carlson S (2012) Synthetic analyses of phenotypic selection in natural populations: lessons, limitations and future directions. Evol Ecol (in press)
Knowles SCL, Nakagawa S, Sheldon BC (2009) Elevated reproductive effort increases blood parasitaemia and decreases immune function in birds: a meta-regression approach. Funct Ecol 23:405–415
Kodric-Brown A, Sibly RM, Brown JH (2006) The allometry of ornaments and weapons. Proc Natl Acad Sci USA 103:8733–8738
Kokko H (2001) Fisherian and 'good genes' benefits of mate choice: how (not) to distinguish between them. Ecol Lett 4:322–326
Kokko H (2011) Directions in modelling partial migration: how adaptation can cause a population decline and why the rules of territory acquisition matter. Oikos 120:1826–1837
Kokko H, Jennions MD (2008) Parental investment, sexual selection and sex ratios. J Evol Biol 21:919–948
Kokko H, Ots I (2006) When not to avoid inbreeding. Evolution 60:467–475
Kokko H, Rankin DJ (2006) Lonely hearts or sex in the city? Density-dependent effects in mating systems. Phil Trans R Soc B 361:319–334
Kokko H, Jennions MD, Brooks RC (2006a) Unifying and testing models of sexual selection. Ann Rev Ecol Evol Syst 37:43–66
Kokko H, López-Sepulcre A, Morrell LJ (2006b) From hawks and doves to self-consistent games of territorial behavior. Am Nat 167:901–912
Koricheva J, Gurevitch J, Mengersen K (2012) Handbook of meta-analysis in ecology and evolution. Princeton University Press, Princeton
Kraaijeveld KF, Kraaijeveld-Smit JL, Komdeur J (2007) The evolution of mutual ornamentation. Anim Behav 74:657–677
Kraaijeveld KF, Kraaijeveld-Smit JL, Maan ME (2011) Sexual selection and speciation: the comparative evidence revisited. Biol Rev 86:367–377
Krasnov BR, Stanko M, Matthee SM, Laudisoit A, Leirs H, Khokhlova IS, Korallo-Vinarskaya NP, Vinarski MV, Morand S (2011) Male hosts drive infracommunity structure of ectoparasites. Oecologia 166:1099–1110
Lajeunesse MJ (2009) Meta-analysis and the comparative phylogenetic method. Am Nat 174:369–381
Lea AJ, Blumstein DT (2011) Ontogenetic and sex differences influence alarm call responses in mammals: a meta-analysis. Ethology 117:839–851
Leamy LJ, Klingenberg CP (2005) The genetics and evolution of fluctuating asymmetry. Ann Rev Ecol Evol Syst 36:1–21
Leung B, Forbes MR (1996) Fluctuating asymmetry in relation to stress and fitness: effects of trait type as revealed by meta-analysis. Ecoscience 3:400–413
Matthe S, McGeoch MA, Krasnov BR (2010) Parasite-specific variation and the extent of male-biased parasitism; an example with a South African rodent and ectoparasitic arthropods. Parasitology 137:651–660
McCurdy DG, Shutler D, Mullie A, Forbes MR (1998) Sex-biased parasitism of avian hosts: relations to blood parasite taxon and mating system. Oikos 82:303–312
Meunier JS, Pinto F, Burri T, Roulin A (2011) Eumelanin-based coloration and fitness parameters in birds: a meta-analysis. Behav Ecol Sociobiol 65:559–567
Milner RNC, Detto T, Jennions MD, Backwell PRY (2010) Experimental evidence for a seasonal shift in the strength of a female mating preference. Behav Ecol 21:311–316
Møller AP, Alatalo RV (1999) Good-genes effects in sexual selection. Proc R Soc B 266:85–91

Møller AP, Jennions MD (2001) How important are direct fitness benefits of sexual selection? Naturwissenschaften 88:401–415
Møller AP, Jennions MD (2002) How much variance can be explained by ecologists and evolutionary biologists? Oecologia 132:492–500
Møller AP, Ninni P (1998) Sperm competition and sexual selection: a meta-analysis of paternity studies of birds. Behav Ecol Sociobiol 43:345–358
Møller AP, Saino N (2004) Immune response and survival. Oikos 104:299–304
Møller AP, Thornhill R (1997) A meta-analysis of the heritability of developmental stability. J Evol Biol 10:1–16
Møller AP, Thornhill R (1998a) Bilateral symmetry and sexual selection: a meta-analysis. Am Nat 151:174–192
Møller AP, Thornhill R (1998b) Male parental care, differential parental investment by females and sexual selection. Anim Behav 55:1507–1515
Møller AP, Christe P, Erritzoe J, Maravez J (1998) Condition, disease and immune defence. Oikos 83:301–306
Møller AP, Christe P, Lux E (1999) Parasitism, host immune function and sexual selection. Q Rev Biol 74:3–20
Møller AP, Thornhill R, Gangestad SW (2005) Direct and indirect tests for publication bias: asymmetry and sexual selection. Anim Behav 70:497–506
Molloy PP, Reynolds JD, Gage MJG, Mosqueria I, Cote IM (2008) Links between sex change and fish densities in marine protected areas. Biol Conserv 141:187–197
Moore SL, Wilson K (2002) Parasites as a viability cost of sexual selection in natural populations of mammals. Science 297:2015–2018
Morrissey M, Hadfield J (2012) Directional selection in temporally replicated studies is remarkably consistent. Evolution (in press)
Muller W, Eens M (2009) Elevated yolk androgen levels and the expression of multiple sexually selected male characters. Hormones Behav 55:175–181
Nakagawa S, Cuthill IC (2007) Effect size, confidence interval and statistical significance: a practical guide for biologists. Biol Rev 779(82):591–605
Nakagawa S, Santos ESA (2012) Methodological issues and advances in biological meta-analysis. Evol Ecol (in press)
Nakagawa S, Ockendon N, Gillespie DOS, Hatchwell BJ, Burke T (2007) Assessing the function of house sparrows' bib size using a flexible meta-analysis method. Behav Ecol 18:831–840
Nunn CL, Lindenfors P, Pursall ER, Rolff J (2009) On sexual dimorphism in immune function. Phil Trans R Soc B 364:61–69
O'Grady JJ, Brook BW, Reed DH, Ballou JD, Tonkyn DW, Frankham R (2006) Realistic levels of inbreeding depression strongly affect extinction risk in wild populations. Biol Conserv 133:42–51
Palmer AR (1999) Detecting publication bias in meta-analyses: a case study of fluctuating asymmetry and sexual selection. Am Nat 154:220–233
Palmer AR (2000) Quasireplication and the contract of error: lessons from sex ratios, heritabilities and fluctuating asymmetry. Ann Rev Ecol Syst 31:441–480
Parker TH, Ligon JD (2003) Female mating preferences in red junglefowl: a meta-analysis. Ethol Ecol Evol 15:63–72
Parker TH, Barr IR, Griffith SC (2006) The blue tit's song is an inconsistent signal of male condition. Behav Ecol 17:1029–1040
Peek MS, Leffler AJ, Flint SD, Ryel RJ (2003) How much variance is explained by ecologists? Additional perspectives. Oecologia 137:161–170
Poissant J, Wilson AJ, Coltman DW (2010) Sex-specific genetic variance and the evolution of sexual dimorphism: a systematic review of cross-sex genetic correlations. Evolution 64:97–107
Pomiankowski A, Møller AP (1995) A resolution of the lek paradox. Proc R Soc B 1357:21–29
Poulin R (1996a) Helminth growth in vertebrate hosts: does host sex matter? Int J Parasitol 26:1311–1315
Poulin R (1996b) Sexual inequalities in helminth infections: a cost of being a male? Am Nat 147:287–295
Poulin R (2000) Manipulation of host behaviour by parasites: a weakening paradigm? Proc R Soc B 267:787–792
Poulin R, Forbes MR (2012) Meta-analysis and research on host-parasite interactions: past and future. Evol Ecol (in press)
Randler C (2008) Mating patterns in avian hybrid zones—a meta-analysis and review. Ardea 96:73–80
Ratikanen I, Kokko H (2010) Differential allocation and compensation: who deserves the silver spoon. Behav Ecol 21:192–200
Reed DH, Frankham R (2003) Correlation between fitness and genetic diversity. Conserv Biol 117:230–237

Reznick D, Nunney L, Tessier A (2000) Big houses, big cars, superfleas and the costs of reproduction. Trend Ecol Evol 15:421–425
Roberts ML, Buchanan KL, Evans MR (2004) Testing the immunocompetence handicap hypothesis: a review of the evidence. Anim Behav 68:227–239
Santos ES, Scheck AD, Nakagawa S (2011) Dominance and plumage traits: meta-analysis and metaregression analysis. Anim Behav 82:3–19
Schalk G, Forbes MR (1997) Male biases in parasitism of mammals: effects of study type host age and parasite taxon. Oikos 78:67–74
Schino G (2004) Birth sex ratio and social rank: consistency and variability within and between primate groups. Behav Ecol 15:850–856
Schmoll T (2011) A review and perspective on context-dependent genetic effects of extra-pair mating in birds. J Ornithol 152:S265–S277
Sheldon BC (2000) Differential allocation: tests, mechanisms and implications. Trends Ecol Evol 15:398–402
Sheldon BC, West SA (2004) Maternal dominance maternal condition and offspring sex ratio in ungulate mammals. Am Nat 163:40–54
Sheridan LAD, Poulin R, Ward DF, Zuk M (2000) Sex differences in parasitic infections among arthropod hosts: is there a male bias? Oikos 88:327–334
Siepielski AM, DiBattista JD, Carlson SD (2009) It's about time: the temporal dynamics of phenotypic selection in the wild. Ecol Lett 12:1261–1276
Silk J, Willoughby BE, Brown GR (2005) Maternal rank and local resource competition do not predict birth sex ratios in wild baboons. Proc R Soc B 272:859–864
Simmons LW (2005) The evolution of polyandry: sperm competition, sperm selection and offspring viability. Ann Rev Ecol Evol Syst 36:125–146
Simmons LW, Tomkins JL, Kotiaho JS, Hunt J (1999) Fluctuating paradigm. Proc R Soc B 266:593–595
Simons MJP, Verhulst S (2011) Zebra finch females prefer males with redder bills independent of song rate-a meta-analysis. Behav Ecol 22:755–762
Slatyer RA, Mautz BS, Backwell PRY, Jennions MD (2012) Estimating genetic benefits of polyandry from experimental studies: a meta-analysis. Biol Rev 87:1–33
Smith BR, Blumstein DT (2008) Fitness consequences of personality: a meta-analysis. Behav Ecol 19:448–455
Snook RR (2005) Sperm in competition: not playing by the numbers. Trends Ecol Evol 20:46–53
Sokolovska N, Rowe L, Johansson F (2000) Fitness and body size in mature odonates. Ecol Entomol 25:239–248
Soma M, Garamszegi LZ (2011) Rethinking birdsong evolution: meta-analysis of the relationship between song complexity and reproductive success. Behav Ecol 22:363–371
South A, Lewis SM (2011) The influence of male ejaculate quantity on female fitness: a meta-analysis. Biol Rev 86:299–309
Thornhill R, Møller AP (1998) The relative importance of size and asymmetry in sexual selection. Behav Ecol 9:546–551
Thornhill R, Møller AP, Gangestad SW (1999) The biological significance of fluctuating asymmetry and sexual selection: A reply to Palmer. Amer Nat 154:234–241
Torres-Vila LM, Jennions MD (2005) Male mating history and female fecundity in the Lepidoptera: do male virgins make better partners? Behav Ecol Sociobiol 57:318–326
Torres-Vila LM, Rodriguez-Molina MC, Jennions MD (2004) Polyandry and fecundity in the Lepidoptera: can methodological and conceptual approaches bias outcomes? Behav Ecol Sociobiol 55:315–324
Vanderwerf E (1992) Lack clutch size hypothesis—an examination of the evidence using meta-analysis. Ecology 73:1699–1705
Vøllestad LA, Hindar K, Møller AP (1999) A meta-analysis of fluctuating asymmetry in relation to heterozygosity. Heredity 83:206–218
Wehi PM, Nakagawa S, Trewick SA, Morgan-Richards M (2011) Does predation result in adult sex ratio skew in a sexually dimorphic insect genus? J Evol Biol 24:2321–2328
Weir LK, Grant JWA, Hutchings JA (2011) The influence of operational sex ratio on the intensity of competition for mates. Am Nat 177:167–176
West SA, Sheldon BC (2002) Constraints in the evolution of sex ratio adjustment. Science 295:1685–1688
West SA, Shuker DM, Sheldon BC (2005) Sex-ratio adjustment when relatives interact: a test of constraints on adaptation. Evolution 59:1211–1228
Westlake KP, Rowe L (1999) Developmental costs of male sexual traits in the water strider *Rheumatobates rileyi*. Can J Zool 77:917–922
Yasukawa K, Enstrom DA, Parker PG, Jones TC (2010) Male Red-winged Blackbirds with experimentally dulled epaulets experience no disadvantage in sexual selection. J Field Ornithol 81:31–41