# Do male secondary sexual characters signal ejaculate quality? A meta-analysis

Brian S. Mautz[1,*,†], Anders P. Møller[2] and Michael D. Jennions[1]

[1] *Research School of Biology, Australian National University, Canberra, ACT 0200, Australia*
[2] *Laboratoire d'Ecologie, Systématique et Evolution, CNRS UMR 8079, Université Paris-Sud, Bâtiment 362, Orsay Cedex, F-91405 Paris, France*

ABSTRACT

There are two reasons why researchers are interested in the phenotypic relationship between the expression of male secondary sexual characters (SSCs) and 'ejaculate quality' (defined as sperm/ejaculate traits that are widely assumed to increase female fertility and/or sperm competitiveness). First, if the relationship is positive then females could gain a direct benefit by choosing more attractive males for fertility assurance reasons ('the phenotype-linked fertility' hypothesis). Second, there is much interest in the direction of the correlation between traits favoured by pre-copulatory sexual selection (i.e. affecting mating success) and those favoured by post-copulatory sexual selection (i.e. increasing sperm competitiveness). If the relationship is negative this could lead to the two forms of selection counteracting each other. Theory predicts that the direction of the relationship could be either positive or negative depending on the underlying genetic variance and covariance in each trait, the extent of variation among males in condition (resources available to allocate to reproductive traits), and variation among males in the cost or rate of mating. We conducted a meta-analysis to determine the average relationship between the expression of behavioural and morphological male secondary sexual characters and four assays of ejaculate quality (sperm number, viability, swimming speed and size). Regardless of how the data were partitioned the mean relationship was consistently positive, but always statistically non-significant. The only exception was that secondary sexual character expression was weakly but significantly positively correlated with sperm viability ($r = 0.07$, $P < 0.05$). There was no significant difference in the strength or direction of the relationship between behavioural and morphological SSCs, nor among relationships using the four ejaculate quality assays. The implications of our findings are discussed.

*Key words*: condition dependence, female choice, mate choice, sexual selection, sperm competition, sperm count, sperm motility, sperm length, sperm velocity, sperm viability.

CONTENTS

\* Author for correspondence (Tel: +1 613-562-5800 x 6837; Fax +1 613-562-5486; E-mail: bmautz@uottawa.ca).

† Present address: Department of Biology, University of Ottawa, 30 Marie-Curie, Ottawa, ON, K1N 6N5, Canada.

# I. INTRODUCTION

## (1) Female choice, male sexual traits and honest signals of benefits

Most exaggerated male secondary sexual characters (SSCs) increase predation risk, divert resources from other life-enhancing functions (e.g. a well-functioning immune system) or impose other fitness costs (Kotiaho, 2001). Despite these costs, such characters can be maintained through sexual selection from directional female mating preferences (Kokko, Jennions & Brooks, 2006). Males trade off a greater mating rate because they are more attractive against any resultant reduction in lifespan (Kokko, Klug & Jennions, 2012). In turn, choosy females benefit if SSCs allow them to identify males with an above-average effect on their fitness.

Males can have direct effects on female fitness *via* lifetime offspring production (i.e. provide material resources that increase her lifespan and/or fecundity per breeding event) (meta-analysis: Møller & Jennions, 2001). In addition, there are indirect effects as the male's genes influence mean offspring fitness (Andersson, 1994) through genes that elevate the fitness of sons and daughters (i.e. improve general 'viability'), or of only one sex (usually sons). The elevation of sons' fitness can be due to genes that increase attractiveness alone, or that also elevate other fitness components (the Fisher-Zahavi process: see Kokko *et al.*, 2002). There is, however, a 'grey area' between direct and indirect benefits when considering how sperm and male 'infertility' affect female fitness. For example, from an evolutionary perspective there is little difference between a sperm failing to penetrate an egg, or penetrating and the resultant zygote failing to develop. The former is, however, treated as a direct effect ('male infertility') and the later (often) as an indirect, genetic effect ('bad genes') (review: Hasson & Stone, 2009; see also Jennions & Petrie, 2000).

A male's effect on female fitness is often referred to as his 'quality' as a mate. This is a slippery term that needs a precise definition (e.g. see Hunt *et al.*, 2004 for a definition for genetic benefits). Males can affect female fitness in several ways and different aspects of male 'quality' are rarely perfectly correlated (Kokko *et al.*, 2006). For example, a male might sire offspring of above-average fitness, but

provide little parental care (e.g. Møller & Thornhill, 1998). The ideal definition of mate quality assesses the net effect of a male on his partner's fitness (Møller, 1994). In practice, many studies focus on a single aspect of quality and do not address correlations between different measures of quality. We return to this issue when considering the link between male effects on female fertility and offspring fitness.

If there is a correlation between the expression of a SSC and some aspect of male quality then a female mating preference might evolve because the trait 'honestly' signals quality (Andersson, 1994; Kokko *et al.*, 2006). The standard explanation for a reliable relationship invokes the 'handicap principle' (Zahavi, 1975). This is usually described as a mechanism where the marginal cost of investment into a sexual trait is lower for higher quality males (Grafen, 1990; empirical evidence: Møller & de Lope, 1994; Kotiaho, 2000). This allows a higher quality male to invest more into a SSC, without the resultant mating gains immediately being negated by greater costs. It has, however, been noted that this interpretation invokes an additive model of fitness (i.e. male mortality costs and mating benefits are summed as though measured in the same units of fitness) (Getty, 2006). A multiplicative model (i.e. fitness is the product of survival and mating rate) seems more appropriate but it requires a different interpretation. Here, sexual selection favours greater investment by males that more efficiently convert resources into SSCs that elevate net fitness. If higher quality males are more efficient, then females can use SSCs to assess male quality (Møller, 1994; Getty, 2006). The main problem is then how to conduct appropriate empirical tests (Murai, Backwell & Jennions, 2009): several different marginal cost relationships can generate honest signals (see box 2 and fig. 1 in Getty, 2006).

By definition, 'condition' is the major determinant of a male's ability to efficiently convert resources into sexual signals, or withstand a larger 'handicap'. SSCs that are honest signals are therefore described as condition dependent. Condition is formally defined as the pool of resources allocated to traits that enhance fitness (Rowe & Houle, 1996; Hill, 2011). This is a conceptual definition and, in practice, 'condition' is challenging to measure (Hunt *et al.*, 2004; Tomkins *et al.*, 2004). A proxy test for 'condition dependence' is whether greater food availability increases investment into the focal trait (Cotton, Fowler & Pomiankowski, 2004).

Theoretical models show that the handicap mechanism can account for honest signals of both genetic (e.g. Grafen, 1990; Iwasa, Pomiankowski & Nee, 1991) and direct benefits (e.g. Iwasa & Pomiankowski, 1999). Sexual selection theory has, however, mainly focused on genetic benefits. It has been suggested that direct benefits have been under-studied (Griffith, 2007; Wagner, 2011; but see Møller & Jennions, 2001), even though the evolution of mating preferences for direct benefits appears more likely than those for genetic benefits (Kirkpatrick & Barton, 1997), which are often very small (Møller & Alatalo, 1999; Arnqvist & Kirkpatrick, 2005).

We suggest that the relative lack of research on signals of direct benefits is partly because such benefits are often immediately assessable (e.g. territory quality, male ability to defend the female from rivals). Females do not require an intermediate signal to reveal hidden male qualities. There are, however, two notable exceptions. First, male parental care is only 'visible' in the future. This has led to theoretical ('good parent') models asking whether males should honestly signal paternal care (Hoezler, 1989; Heywood, 1989; Price, Schluter & Heckman, 1993; Schluter & Price, 1993; Iwasa & Pomiankowski, 1999; Kelly & Alonzo, 2009; Alonzo, 2012). The available empirical data are challenging to interpret. There are some consistent patterns, but there is no simple link within species between investment into male attractiveness and the level of male care, although such relationships are strong among species (Andersson, 1994). The mean relationship does not differ from zero and varies greatly among species (Møller & Jennions, 2001). This is partly explicable if sexual traits signal several aspects of male quality, whose relative importance varies among species. For example, variation in the phenotypic correlation between attractiveness and male care in birds has been attributed to species differences in the importance of genetic benefits. When genetic benefits are high (assuming that greater extra-pair paternity increases the value of more attractive sons), females prefer males with elaborate SSCs, even if they provide less care (Møller & Thornhill, 1998).

Second, ejaculate traits cannot be assessed prior to mating and female fertility (i.e. the proportion of eggs fertilized) can be affected by a male's 'ejaculate quality' (Wetton & Parker, 1991; Møller, 1994; Krokene *et al*., 1998; Wedell, Gage & Parker, 2002). It has been suggested that females prefer males with greater expression of SSCs because this signals the likelihood that she will fertilize all her eggs (Trivers, 1972; Williams, 1992; the 'phenotype-linked fertility' hypothesis: Sheldon, 1994).

## (2) Why study the link between secondary sexual characters and ejaculate traits?

Here we use a meta-analysis to test the 'phenotype-linked fertility' hypothesis. We quantify the phenotypic correlation between male secondary sexual characters (SSCs) and 'ejaculate quality' measured using four ejaculate traits putatively associated with increased female fertility ('ejaculate quality' is defined in Section I.4). Our findings complement an earlier meta-analysis of the relationship between female fertility and numerous male traits, including body size (Møller & Jennions, 2001). Conceptually, the strength of the relationship we estimate can also be envisaged as documenting the link between pre-copulatory and post-copulatory sexual selection. We document the link between SSCs that elevate mating success (i.e. pre-copulatory selection) and ejaculate traits associated with greater sperm competitiveness (i.e. under post-copulatory sexual selection). Specifically, the four ejaculate traits we use as indicators of 'ejaculate quality' are generally assumed to improve the ability to gain paternity under sperm competition (e.g. higher sperm count and greater sperm viability). Although selection on males probably favours the same ejaculatory traits in both cases, we might expect stronger selection to ensure a high share of paternity than to maximize female fertility because, at least in birds, infertility is typically in the order of 10% (Spottiswoode & Møller, 2004) while extra-pair paternity may reach more than two thirds of all offspring. Males only gain fitness when they sire offspring and, given polyandry, there is sexual conflict over paternity that can adversely affect female fertility. For example, if females usually mate multiply, then males that fertilize 70% of eggs in a competitive situation (even if at the cost of some eggs not being fertilized) will tend to do better than males that can ensure complete fertility when a female mates singly, but fertilize < 70% of eggs in the competitive situation.

There is much interest in the relationship between pre- and post-copulatory sexual selection (Birkhead & Pizzari, 2002; Evans *et al*., 2003; Hunt *et al*., 2009; Evans, 2010; Engqvist, 2011). If there is a negative correlation between traits favoured in each episode of selection this will slow the rate of evolutionary change if the correlation is genetic (Andersson & Simmons, 2006), and reduce the extent to which male reproductive success can be predicted based on mating success where the correlation is phenotypic (e.g. Danielsson, 2001).

It is difficult to predict the phenotypic relationship between male SSCs and sperm competitiveness (mini-review: Engqvist, 2011). There is good evidence that SSCs are highly condition dependent (Cotton *et al*., 2004). There is also evidence that some ejaculate properties vary due to factors such as food availability (Perry & Rowe, 2010) and inbreeding (Fitzpatrick & Evans, 2009; references in Beausoleil *et al*., 2012, p. 975). These factors probably affect the resources a male can allocate to reproduction (i.e. condition *sensu* Rowe & Houle, 1996), suggesting that some aspects of ejaculates are condition dependent. This might imply a positive correlation because males in good condition can invest more into both SSCs and ejaculate traits (Williams, 1992; Sheldon, 1994; Malo *et al*., 2005). Similarly, free radicals that cause oxidative stress negatively affect the expression of SSCs and ejaculate traits (Blount, Møller & Houston, 2001; Dowling & Simmons, 2009; Helfenstein *et al*., 2010). Again, this seemingly implies a positive relationship when males vary in their ability to acquire/utilize antioxidants. Unfortunately, the situation is more complex.

First, there are always trade-offs when life-history traits depend upon the same resources (Roff, 2002). Investment

into elaborate SSCs could reduce the availability of resources for other reproductive traits affecting ejaculate quality such as testes (Simmons & Emlen, 2006) or ejaculate volume (Kelly, 2008; Engqvist, 2011). Given this, we might predict a negative relationship between SSCs and ejaculate quality (Parker, 1998; Evans, 2010). As with all life-history traits the phenotypic relationship ultimately depends on the relative extent of variation among males in resource acquisition and allocation strategies (van Noordwijk & de Jong, 1986). For example, if variation in acquisition is high and that in allocation modest, the phenotypic relationship will be positive, despite the evolutionary trade-offs (Reznick, Nunney & Tessier, 2000; Roff & Fairbairn, 2007).

Second, Tazzyman et al. (2009) explicitly modeled whether continuous variation among males in available resources (i.e. in 'condition') affects optimal investment into ejaculates (specifically, sperm number). They conclude that variation in condition alone does not cause the optimal ejaculate size to vary among males. However, if males vary in the costs paid to acquire a mate then males with lower costs are predicted to invest less per ejaculate: the rate of return per sperm decreases with greater investment per ejaculate (i.e. asymptotes) and the maximal value of each ejaculate is lower for males that can mate more often. The situation is analogous to spending less time foraging in a patch when new patches are more often encountered (i.e. the marginal value theorem). Males with more elaborate SSCs probably pay lower mate-acquisition costs (e.g. mate sooner, thereby reducing the time cost), suggesting that there will be a negative correlation between SSCs and ejaculate quality (see Bussière, Basit & Gwynne, 2005). Tazzyman et al. (2009) excluded the initial cost of investment into SSCs from their main model. They note, however, that if included this could generate an initial correlation between the resources available to invest into ejaculates and SSCs that might affect mating costs, and hence sperm number.

Third, we generally expect negative genetic correlations between fitness-enhancing traits. Selection leads to fixation of alleles that elevate the expression of both traits, removing standing genetic variation that creates positive genetic correlations. Ultimately, the only genetic variation is for alleles that have antagonistically pleiotropic effects on the focal traits (for a more sophisticated multivariate analysis see Blows, 2007). Several recent studies have reported negative genetic correlations between male attractiveness and investment into ejaculates (e.g. Evans, 2010; Engqvist, 2011) although others report a significantly positive correlation (Simmons & Kotiaho, 2002; Hosken et al., 2008). If negative genetic correlations are more common, when environmental variation in trait expression is low, we predict a negative phenotypic correlation (because the relationship depends on trait covariation due to genetic and shared environmental factors).

## (3) Phenotypic relationship between pre-copulatory and post-copulatory selection

There are several ways to quantify the relationship between the expression of traits under pre-copulatory and post-copulatory selection (Engqvist, 2011). Some studies directly examine the relationship between specific SSCs and female fertility. These tend to report a weak, but positive relationship (see table 1 in Møller & Jennions, 2001). Other studies examine the relationship between mating success or male attractiveness (i.e. investment into SSCs) and estimates of sperm competitiveness based on share of paternity. This has yielded correlations that range from significantly positive (e.g. Pilastro et al., 2004; Pitcher, Rodd & Rowe, 2007; Hosken et al., 2008) to negative (e.g. Danielsson, 2001; Demary & Lewis, 2007). It is worth noting, however, that cryptic female choice rather than 'sperm competitiveness' per se can affect post-copulatory success (Andersson & Simmons, 2006). It has therefore been suggested that researchers should look at specific ejaculate traits with a known or assumed causal effect on the ability to gain paternity (Evans, 2010, p. 3196).

Some studies have investigated the relationship between SSCs and specific ejaculate traits associated with 'ejaculate quality' (defined below). These studies report both significantly positive (e.g. Matthews, Evans & Magurran, 1997; Locatello et al., 2006; Calhim et al., 2009, Chargé et al., 2010) and significantly negative (Engen & Folstad, 1999; Liljedal, Folstad & Skarstein, 1999) correlations, while other studies fail to detect a significant correlation (e.g. Birkhead & Fletcher, 1995; Birkhead et al., 1997; Birkhead, Fletcher & Pellatt, 1998; Hettyey, Herczeg & Hoi, 2009; Gasparini et al. 2010a; Lifjeld et al., 2012). This wide variation in outcomes (and the seemingly greater number of studies measuring this relationship than the others described) motivated us to conduct a meta-analysis of this relationship.

## (4) What is ejaculate quality?

In our meta-analysis we do not quantify directly the relationship between a male's SSCs and his ability to gain paternity when females mate multiply (i.e. sperm or, more accurately, ejaculate competitiveness), or his effect on female fertility (only measurable if a female solely uses sperm from the focal male, i.e. she is either a virgin or a mated female without stored sperm). This is partly because appropriate data are rare. It is easier to measure ejaculate traits than to quantify share of paternity using molecular techniques. Instead we quantify the relationship between SSCs and sperm traits associated with 'ejaculate quality'. We implicitly assume that these traits predict ejaculate competitiveness and male effects on female fertility.

Although the terms 'ejaculate quality' or 'sperm quality' are used widely (e.g. Snook, 2005; title of Evans, 2010) there is no standard definition. Recent sexual selection reviews have highlighted problems with the use of the term 'quality' in other areas, leading to more precise definitions of 'individual quality' (Wilson & Nussey, 2010; Hill, 2011; Lailvaux & Kasumovic, 2011) and 'male genetic quality' (Hunt et al., 2004). Here we define ejaculate quality as high if females benefit because of increased fertility (i.e. higher egg fertilization rate) and/or because the ejaculate is more competitive which, if heritable, will increase the likelihood

that a female's sons gain paternity under sperm competition (i.e. a genetic benefit). It is possible that ejaculate traits might not simultaneously maximize sperm competitiveness and female fertility because of sexual conflict. For example, some seminal chemicals that kill rival sperm might reduce female fertility. We argue, however, that the four ejaculate traits that we use to measure 'ejaculate quality' are likely always to have positive (or neutral) effects on female fertility and ejaculate competitiveness (see Snook, 2005).

The four traits we use as indices of higher ejaculate quality are: a greater number of sperm, increased sperm viability, faster swimming sperm and longer sperm (see online Appendix 2 for how various ejaculate assays are assigned to each category). We briefly review the available evidence to support the use of each of these assays.

First, there is strong evidence that greater sperm number (e.g. higher spermatocrit) increases share of paternity. Almost all sperm competition theory makes this assumption (Parker & Pizzari, 2010), and studies of phenotypic plasticity show that males produce larger ejaculates when the sperm competition risk is greater (Kelly & Jennions, 2011). There is less evidence directly relating sperm numbers to fertility in wild animals, but a low sperm count reduces the likelihood of pregnancy in domesticated mammals and humans (references in Hasson & Stone, 2009, p. 9).

Second, there is good evidence that greater sperm viability (e.g. proportion of sperm that are motile) also increases a male's share of paternity (Fry & Wilkinson, 2004; García-González & Simmons, 2005; Smith, 2012; but see Sherman *et al.*, 2008; Sherman, Wapstra & Olsson, 2009), or is associated with greater sperm competitiveness (Hunter & Birkhead, 2002; Firman & Simmons, 2010; Gomez-Montoto *et al.*, 2011; Rowe & Pruett-Jones, 2011). This could be because only viable sperm can fertilize eggs so that, in effect, sperm viability is a 'weighting' factor needed to 'count' correctly the number of competitive sperm (but see Simmons *et al.*, 2003; Gage & Morrow, 2003). We are unaware of any evidence directly relating sperm viability to fertility in wild animals. In humans and domestic animals, however, a reduced likelihood of pregnancy or low fertilization rates is associated with a greater proportion of abnormal, dead or immotile sperm (World Health Organization, 1999; see Hasson & Stone, 2009).

Third, sperm swimming speed (e.g. straight-line velocity, curvilinear velocity, distance travelled) is often measured in studies quantifying 'ejaculate quality' (e.g. Evans, 2010). It is generally assumed that faster sperm are more likely to achieve fertilization, and that speed is therefore positively correlated with competitiveness (e.g. Ball & Parker, 1996). A comparative study of cichlid fishes showed that sperm swimming speed is faster in species with higher levels of sperm competition (Fitzpatrick *et al.*, 2009). Studies of individual species have also reported a positive relationship between sperm swimming speed and ejaculate competitiveness (e.g. Birkhead *et al.*, 1999; Gage *et al.*, 2004; Linhart *et al.*, 2005; Pizzari *et al.*, 2008; Gasparini *et al.*, 2010b; Boschetto, Gasparini & Pilastro, 2011; Beausoleil *et al.*, 2012; but see

Dziminski *et al.*, 2009; Smith, 2012). By extension, it is assumed that faster swimming sperm will elevate fertility because this increases the likelihood of sperm reaching eggs before they die. There is, for example, evidence that female fertility is determined by sperm swimming speed in several taxa (e.g. Mortimer, Pandya & Sawers, 1986; Froman *et al.*, 1999; Levintan, 2000; Malo *et al.*, 2005).

Fourth, several studies show that longer, larger sperm increase share of paternity (Radwan, 1996; LaMunyon & Ward, 1999; Miller & Pitnick, 2002; Oppliger *et al.*, 2003; Pattarini *et al.*, 2006). There is also comparative evidence from several taxa (Lepidoptera, mammals, frogs, fish, primates, snakes) positively linking sperm length to sexual selection (for references see García-González & Simmons, 2007; Fitzpatrick *et al.*, 2009; Tourmente *et al.*, 2009, 2011a,b). This has led to the situation where: '*Collectively, it has been hypothesized that longer sperm might increase the competitive potential of an ejaculate or promote female sperm choice because longer sperm may, among other reasons, swim faster, live longer, be more effective in sperm displacement within the female reproductive tract, or indicate higher male quality*' (Snook, 2005). (quote from García-González & Simmons, 2007, p. 817). It should, however, be noted that some studies show that smaller sperm are more competitive (e.g. Gage & Morrow, 2003; García-González & Simmons, 2007), or failed to detect an effect of sperm size on share of paternity (e.g. Simmons *et al.*, 2003, Gage *et al.*, 2004). A comparative study also reported a negative relationship between sperm length and sperm competition in fish (Stockley *et al.*, 1997; but see Fitzpatrick *et al.*, 2009). Given these findings, it is possible that the direction of the effect size we report is 'incorrect' for some species. Even so, in the absence of species-specific data it is necessary to make consistent, broad assumptions. We therefore follow the 'collective' view that larger sperm indicate higher ejaculate quality. As data accumulate it might be necessary to reverse the sign of effect sizes for species where shorter sperm are shown to be more competitive. To our knowledge there is currently little evidence linking sperm size to female fertility. If, however, larger sperm are more competitive this implies that they are better at locating eggs. We therefore assume that fertility is positively related to sperm length.

## (5) Motivation for meta-analysis

We conducted a meta-analysis of the phenotypic relationship between secondary sexual characters (SSCs) and 'ejaculate quality' because of high heterogeneity in the available empirical findings (Section I.3) and because theory does not readily make *a priori* predictions. We intended to shed some light on (*i*) the extent to which male SSCs signal ejaculate quality; (*ii*) the likely variation among males in their ability to acquire and assimilate resources affecting expression of SSCs and ejaculate quality; and (*iii*) factors that affect the strength or direction of the phenotypic relationship between SSC expression and aspects of 'ejaculate quality' (i.e. sources of heterogeneity in effect sizes). For example, are behavioural traits better than morphological traits at signaling male ejaculate quality? Or are SSCs better at

signaling some ejaculate traits than others? Finally, by conducting a systematic review we can assess the extent to which there is sufficient data to draw general conclusions, and suggest profitable directions for future empirical research.

## II. METHODS

### (1) Database

We used a two-pronged approach to locate studies. First, we performed a key word (TOPIC) search in the *ISI Web of Knowledge*. We used two sets of key words related to either (*i*) sexual selection (e.g. female choice, attractiveness), or (*ii*) secondary sexual characteristics (hereafter SSCs) (e.g. plumage, antler, horn, badge) to locate studies. These lists were then cross-referenced (using the 'AND' option) to ensure that the located studies additionally referred to ejaculate characteristics. We used three main ejaculate terms: 'sperm', 'semen', and 'ejaculate', which were combined with any of 12 trait properties: quality, size, velocity, viability, number, length, speed, motility, mobility, morphology, longevity, and (ab)normal. The exact search term combinations are listed as see online Appendix S1. The term 'fertility' was excluded as a search term because the number of studies in the initial database then became unmanageable (> 20000 papers). Second, we identified five papers that are generally regarded as relevant landmark papers. Three are early studies that tested empirically the relationship between SSCs and ejaculate characteristics (Birkhead & Fletcher, 1995; Birkhead *et al.*, 1998; Liljedal *et al.*, 1999), one is the original paper that introduced the hypothesis that SSCs signal a male's effect on female fertility (Sheldon, 1994), and the fifth is a highly cited review of 'ejaculate quality' which emphasizes that traits other than sperm number (e.g. sperm size and velocity) also affect ejaculate competitiveness (Snook, 2005). We then performed a forward search to compile a list of papers citing any of these five works.

In combination the two search approaches yielded approximately 2900 papers. We did not add papers that we subsequently encountered by chance because these are likely to be drawn from more visible, higher impact journals that might preferentially publish studies with significant results (Murtaugh, 2002). We also did not solicit unpublished datasets from colleagues, again to reduce the risk of biasing our estimates of effect sizes (Jennions *et al.*, 2012). To finalize the data checking and analysis protocols we excluded any papers identified after 30 May 2011. We are aware that other papers suitable for inclusion have appeared since (e.g. Beausoleil *et al.*, 2012), and these will need to be included in any update of the current meta-analysis. An initial cull of the approximately 2900 papers was performed to reduce the number that we had to examine directly. We inspected titles, the journal of publication and, where possible, read the online abstract (see Côté *et al.*, 2012). Publications that obviously did not fit the inclusion criteria were omitted at this stage. We were left with 258 potentially suitable papers that required closer inspection. These papers were downloaded and read to determine if they met our inclusion criteria.

### (2) Inclusion criteria

We had three inclusion criteria: (*i*) is the study within the scope of our question? (*ii*) Is the study design appropriate? (*iii*) Does the paper include extractable data?

For criterion 1, papers had to address our main study question: is there a relationship between the expression of a male SSC and an ejaculate characteristic? Authors of the focal papers defined a range of traits as SSCs, including horn size, plumage colouration, song rates, courtship rates, social status, condition and body size (see online Appendix S2). As we were searching for studies related to sexual selection, we relied on the authors of the original paper to define a trait as a SSC (and, by extension, a trait that has been, or still is, under sexual selection). We were specifically interested in sexually dimorphic traits and only included traits that are, generally speaking, not expressed by females (e.g. male courtship colours or advertisement calls) or differ qualitatively between the sexes (e.g. major differences in horn shape). We excluded traits such as social dominance, body condition and, most importantly, body size that are not usually treated as SSCs (Andersson, 1994).

We were interested in four assays of 'ejaculate quality': sperm number, sperm swimming speed, sperm size, and sperm viability (see online Appendix S2 for sub-categories). We did not include measures of non-sperm characteristics of ejaculates such as seminal chemicals, even though these might affect female fertility and/or male competitiveness under sperm competition (review: Cameron, Day & Rowe, 2007), because these relationships are far less clear than those for the four ejaculate traits we do consider. Animals from all taxa were potentially suitable for inclusion in our database, with the exception that we (*i*) excluded studies of agricultural and domesticated animals as artificial selection and artificial insemination have probably altered the correlation between SSCs and ejaculate quality, and (*ii*) excluded studies on humans (e.g. Soler *et al.*, 2003). We were interested in SSCs rather than attractiveness *per se* so we exclude studies that only report on composite measures of attractiveness (e.g. 'mating latency' in Klaus *et al.*, 2011).

For criterion 2, we included observational studies that used either categorical or continuous approaches to assess male SSCs. In categorical studies, the authors placed males into discrete classes (e.g. large or small ornaments), measured ejaculate quality for each class, and tested for statistical differences between SSC categories. In the other studies the authors presented the correlation between a SSC and an ejaculate assay for a random sample of males. The majority of studies with usable data came from the second group. Most studies we included used ejaculates that were stripped from males, rather than sampled from female genital tracts. This potentially reduces noise in the data as studies have shown that males strategically allocate resources to ejaculates depending on female attractiveness/quality (e.g. Cornwallis & Birkhead, 2007; meta-analysis: Kelly & Jennions, 2011).

Of course, there are other potential environmental sources of variation among males that could affect ejaculate traits (e.g. perception of sperm-competition risk). These could either introduce noise into the relationship between SSC and ejaculate quality, or be integral to the natural relationship. For example, if there is a general correlation between male SSC and sperm competition risk (e.g. less-attractive males tend not to mate with virgins) or mating rate (e.g. Tazzyman *et al.*, 2009), it should not be controlled for experimentally when quantifying the natural correlation between SSC and ejaculate quality.

Criterion 3 meant that some studies were omitted because of ambiguous information about sample sizes, effect size magnitude or the direction of the effect. Missing information makes it impossible to calculate an effect size and/or its variance. This problem is often associated with non-significant results (with, on average, smaller effect sizes) because *P* values, test statistics or the direction of the effect more often go unreported in such cases (e.g. Cassey *et al.*, 2004). This could lead to true effect sizes being overestimated.

We have included the full details of our literature search so that readers can evaluate whether our protocols are likely to have biased effect size estimates. We make no claim to have located all relevant studies. Instead, we simply assume that our search protocol yielded an unbiased sample of the effect sizes available in the published literature. If so, our estimated effect sizes should, within the limits of sampling error, agree with those obtained if all available studies had been located.

## (3) Meta-analysis statistical methods

Test statistics were extracted from summary tables, the text or figures (using *ImageJ ver. 1.43*). For each study, we used the software package *Metawin 2.0* (Rosenberg, Adams & Gurevitch, 1997) to convert the test statistic to Fisher's $Z$ ($Z_r$) (Borenstein *et al.*, 2009). We calculated the variance [$= 1/(N-3)$] based on the number of males used to calculate the original correlation. We give only a brief overview of our statistical methods (details in Kelly & Jennions, 2011; Slatyer *et al.*, 2011).

All analyses were conducted at the species level (i.e. one effect size per species to calculate global means). If we initially had more than one effect size estimate for a species, we calculated the mean effect and its associated variance using a hierarchical approach to combine data (Mengersen, Jennions & Schmid, 2012). To start, for each study we calculated a single mean effect size and its associated variance for a given SSC category–ejaculate trait pairing (e.g. several correlations between sperm viability and one or more morphological SSC). If the same males contributed two or more effects (e.g. sperm viability was measured at different times after ejaculation) the variance was calculated conservatively by assuming that the correlation between the estimates was $r = 1$ (equation 1 in Slatyer *et al.*, 2011). If there was still more than one independent effect size estimate per study for a given SSC – sperm trait combination (i.e. estimates from different sets of males were made), we

calculated the mean effect and its variance using a fixed-effects model. We then combined within-study estimates based on whether we wished to combine or separate the two categories of SSC and/or four sperm trait types (see below). Finally, we generated a weighted mean and its variance for each species using a standard fixed-effects model to combine study-level estimates.

We took four approaches when combining effects within and then across studies to generate species-level effects. First, we calculated separate effect size estimates for each of the four 'ejaculate quality' measures (sperm quantity, viability, size and speed) regardless of SSC category. Second, we calculated separate effect size estimates for each SSC category (behavioural or morphological) irrespective of the sperm trait. Third, for each SSC category we calculated separate effect size estimates for each 'ejaculate-quality' measure. Finally, we calculated a single effect size per species pooling across both SCC category and 'ejaculate quality' measure. For each of the four datasets, once we had species-level effect size estimates we ran separate random-effects models in *Metawin 2.0* to estimate the global mean effect.

All effect sizes were calculated so that a positive value indicates a case where more attractive males with greater expression of a SSC had larger values for the 'ejaculate quality' measure. We assume (see Section I.4) that large values (i.e. more sperm, more viable sperm, larger sperm and faster sperm) increase competitiveness under sperm competition and potentially elevate female fertility. The null hypothesis was always an effect size of zero. To test for statistical significance, we inspected the bootstrapped 95% confidence interval (calculated in *MetaWin 2.0*) for the mean weighted effect size to see whether it included zero. To test for sources of variation in effect sizes other than sampling error the total heterogeneity ($Q_T$) is reported assuming it has a $\chi^2$ distribution ($d.f. = N-1$). Mean weighted effect sizes ($Z_r$) were back-converted and expressed as *r* for presentation in tables.

We performed two analyses to test for the effect of potential moderator variables on effect sizes. First, we tested whether effect sizes differed between behavioural and morphological SSCs. Second, we tested for a difference in effect sizes between two taxa that had sufficiently large sample sizes: birds and fish. We performed comparisons by testing for significant between-group heterogeneity ($Q_b$) using randomization tests based on 999 iterations. We initially coded studies for other moderator variables but sample sizes were too small to warrant subsequent analyses.

We used three indirect methods to quantify possible publication bias. First, we calculated the correlation between standardized effect sizes and their variance ($r_{bias}$) (Begg & Mazumdar, 1994). Second, we estimated the number of 'missing' studies using the 'trim and fill' method (Duval & Tweedie, 2000), which assumes that a plot of effect size on sample size should be symmetrical if there is no bias. We then recalculated the mean effect after adding values for putative 'missing' studies. Third, we tested for temporal trends by calculating the Spearman's correlation between study-level effect sizes and year of publication. We also

performed a metaregression in *Metawin 2.0* where effect size estimates are weighted by their variance and year of publication is the predictor covariate. The significance of the influence of year was calculated using a randomization approach (Rosenberg *et al.*, 1997).

Our analyses were performed either with all the available data or excluding three studies (all on guppies, *Poecilia reticulata*) with large sample sizes ($N = 450$, 163, 106; the next three largest studies had $N = 90$, 83, 73). This ensured that the three studies with a relatively large weighting did not overly influence our conclusions. The results were quantitatively almost identical whether these studies were included or excluded, so we only present analyses based on the full dataset.

## III. RESULTS

### (1) Dataset

In total we calculated 228 effect sizes from 38 studies of 21 species (all data are in, see online Appendix S2). For behavioural SSCs there were 48 effect sizes from 13 studies that generated 17 species-level effect sizes when calculated separately for each 'ejaculate quality' measure, and seven species-level effect sizes when pooling across 'ejaculate quality' measures. For morphological SSCs there were 180 effect sizes from 31 studies that generated 21 species-level effect sizes when calculated separately for each 'ejaculate quality' measure, and 17 species-level effect sizes when pooling across 'ejaculate quality' measures (see online Appendix S3 for species-level effect sizes). In total, the dataset was based on analysis of ejaculates from 2299 males.

### (2) Pooling across sperm traits and secondary sexual characters

For the 21 species examined the mean effect size was $r = 0.06$, which is a small effect not significantly different from zero

(95% CI: $-0.04$ to 0.16). The variation among species-level effects did not differ significantly from that expected due to sampling error ($Q_T = 48.33$, $d.f. = 20$, $P = 0.17$). There was no significant difference in the mean effect size between the two taxa with the largest sample sizes, namely fish and birds ($Q_b = 0.46$, $d.f. = 1$, $P = 0.51$, $N = 8$, 8).

### (3) Analyses categorized by sperm-trait type

When data were analyzed separately for each sperm trait type, the only mean effect size that differed significantly from zero was for sperm viability ($r = 0.07$; 95% CI: $0.03–0.17$). Sperm viability was significantly higher in males with greater expression of secondary sexual characters. The mean effect sizes for the remaining three measures of ejaculate quality (sperm quantity, size or speed) did not differ significantly from zero (Table 1). There was no significant difference in the mean effect sizes across the four ejaculate quality measures ($Q_b = 0.87$ $d.f. = 3$, $P = 0.86$).

Similar trends were observed when data were analyzed separately for behavioural and morphological SSCs (Table 1).

### (4) Behavioural secondary sexual characters

The mean effect size for behavioural SSCs was not statistically different from zero ($r = 0.10$; 95% CI: $-0.09$ to 0.27), and the variation was no more than expected due to sampling error ($Q_T = 14.19$, $d.f. = 6$, $P = 0.58$). When effect sizes were analyzed separately for each ejaculate quality measure, the only mean effect size estimate that was significantly greater than zero was for sperm viability ($r = 0.07$; 95% CI: $0.00–0.22$). Males with greater expression of behavioural SSCs had significantly higher sperm viability. None of the estimates of the mean effect size for the other three ejaculate quality measures was significantly different from zero. There was no significant difference in effect sizes among the four ejaculate quality measures ($Q_b = 1.18$, $d.f. = 3$, $P = 0.76$).

Table 1. Summary of sample size ($N$) and mean effect sizes (Pearson's $r$) with 95% bias-corrected bootstrapped confidence intervals for the relationship between secondary sexual characters (SSC) and sperm quality at the species level of analysis

| SSC | Ejaculate trait | $N$ | Mean | 95% CI | $Q_T$ | $P$ |
|---|---|---|---|---|---|---|
| All | All | 21 | 0.06 | ($-0.04$ to 0.16) | 48.33 | 0.17 |
| All | Viability | 10 | **0.07** | **(0.03 to 0.17)** | — | — |
| All | Quantity | 14 | 0.09 | ($-0.06$ to 0.23) | — | — |
| All | Size | 6 | 0.03 | ($-0.22$ to 0.15) | — | — |
| All | Speed | 11 | 0.04 | ($-0.11$ to 0.14) | — | — |
| Behavioural | All | 7 | 0.10 | ($-0.09$ to 0.27) | 14.19 | 0.58 |
| Behavioural | Viability | 5 | **0.07** | **(0.00 to 0.22)** | — | — |
| Behavioural | Quantity | 7 | 0.17 | ($-0.08$ to 0.37) | — | — |
| Behavioural | Size | 3 | 0.23 | ($-0.09$ to 0.33) | — | — |
| Behavioural | Speed | 2 | 0.00 | ($-0.10$ to 0.17) | — | — |
| Morphological | All | 17 | 0.02 | ($-0.09$ to 0.13) | 42.18 | 0.11 |
| Morphological | Viability | 8 | **0.06** | **(0.03 to 0.15)** | — | — |
| Morphological | Quantity | 11 | 0.01 | ($-0.15$ to 0.15) | — | — |
| Morphological | Size | 6 | 0.01 | ($-0.24$ to 0.14) | — | — |
| Morphological | Speed | 12 | 0.03 | ($-0.09$ to 0.12) | — | — |

Estimates in bold were significantly different from zero ($P < 0.05$).

### (5) Morphological secondary sexual characters

The mean effect size for morphological SSCs was not significantly different from zero ($r = 0.02$; 95% CI: $-0.09$ to 0.13), with the variation in effect sizes being no more than expected due to sampling error ($Q_T = 42.19$, $d.f. = 32$, $P = 0.11$). When effect sizes were examined separately for each 'ejaculate quality measure', only sperm viability was significantly positively related to SSC expression ($r = 0.06$, 95% CI: 0.03–0.15). The effect size estimates did not differ significantly across the four ejaculate quality measures ($Q_b = 0.87$, $d.f. = 3$, $P = 0.87$).

Finally, there was no significant difference in effect size estimates for behavioural and morphological SSCs ($Q_b = 0.70$, $P = 0.46$).

### (6) Publication bias

For a species-level analysis based on pooling across all SSC–ejaculate quality measures there was little evidence for publication bias. There was no significant correlation between effect size estimates and their variance ($r_{bias} = -0.27$, $N = 22$, $P = 0.23$), and the 'trim and fill' analysis indicated that there were no 'missing' studies. For analyses of behavioural SSCs there was also no significant correlation between effect size estimates and their variance ($r_{bias} = 0.0$, $N = 7$, $P = 1.00$). A 'trim and fill' analysis did not indicate any 'missing studies'. For analyses of morphological SSCs there was no significant correlation between effect size estimates and their variance ($r_{bias} = -0.36$, $N = 17$, $P = 0.16$). A 'trim and fill' analysis suggested that there was one 'missing study'. The 'corrected' mean effect if this putative study was included was still non-significant at $r = 0.03$ (95% CI: $-0.09$ to 0.17).

We found no significant relationship between year of publication and effect size estimates when conducting separate analyses for each of the four ejaculate-quality measures,

regardless of whether or not we weighted the estimates by their sampling variance (Table 2). If we calculated a single effect per study there was, however, a marginally non-significant decline in effect size with year of publication using either an unweighted ($r_s = -0.30$, $N = 36$, $P = 0.07$) or weighted approach (slope $= -0.015$, $P = 0.07$). This indirect evidence suggests that there might be weak publication bias.

## IV. DISCUSSION

### (1) General findings

We found that male secondary sexual characters (SSCs) do not predict the values of four sperm traits that are assumed to be assays of ejaculate quality. With a mean effect size of $r = 0.06$, the expression of male SSC accounted for < 1% of variance in any given sperm trait. Even so, it is worth noting that small effects can have major evolutionary consequences, and the effect size reported here is similar in magnitude to 'good genes' effects (Møller & Alatalo, 1999), although smaller than that for other direct benefits associated with measures of male attractiveness (Møller & Jennions, 2001). Typically, meta-analyses in comparable areas of biology investigate relationships where researchers can account for 5–10% of the variance in a focal variable based on a single, predictor trait (Møller & Jennions, 2002).

Interestingly, when the data were divided into eight mutually exclusive subgroups (i.e. split by SSC type then by sperm trait), there were more positive relationships than expected by chance (8 of 8, binomial test, $P < 0.01$) (Table 1). This trend for a positive relationship between SSC and ejaculate quality could be related to variation in resource acquisition (i.e. condition) being higher than variation in allocation (van Noordwijk & de Jong, 1986; Reznick *et al.*, 2000; Roff & Fairbairn, 2007). It is, however, uncertain

Table 2. Summary of indirect tests for publication bias using species-level effects

| | $r_{bias}$ | $P$ | Trim and fill | | |
| | | | Missing studies | Mean (corrected) | 95% CI |
| --- | --- | --- | --- | --- | --- |
| (a) | | | | | |
| All | $-0.27$ | 0.24 | 0 | — | — |
| Behaviour | 0.00 | 1.00 | 0 | — | — |
| Morphology | $-0.36$ | 0.16 | 1 | 0.03 | $-0.09$ to 0.17 |
| | Unweighted | | Weighted | | |
| | $r_s$ | $P$ | Slope | $P$ | |
| (b) | | | | | |
| All | $-0.30$ | 0.07 | $-0.01$ | 0.07 | — |
| Viability | 0.17 | 0.53 | 2.08 | 0.72 | — |
| Quantity | $-0.02$ | 0.92 | $-2.30$ | 0.27 | — |
| Size | $-0.21$ | 0.55 | $-3.11$ | 0.27 | — |
| Speed | $-0.22$ | 0.35 | $-2.17$ | 0.13 | — |

(a) Tests for publication bias based on the correlation between mean effect size *versus* variance and the number of 'missing studies' based on the 'trim and fill' method. (b) Unweighted and weighted correlation between mean effect size and year of publication within each sperm category. Samples sizes are given in Table 1.

how often ejaculate traits are condition dependent (Pitnick, Hosken & Birkhead, 2009). There is some evidence that certain assays of ejaculate quality are condition dependent, especially those related to sperm viability (e.g. Skau & Folstad, 2003, 2005) (for evidence for phenotypic plasticity in sperm size in response to environmental changes that might affect male condition, see Crean & Marshall, 2008; Pitnick et al., 2009).

In general, our findings offer little support for the phenotype-linked fertility hypothesis (contra: Trivers, 1972; Williams, 1992; Sheldon, 1994; Griffith, 2007). The strongest (and only significant) phenotypic correlation was a positive one between SSC and sperm viability (i.e. proportion of sperm alive, motile or normal). Of the four 'ejaculate quality' assays, sperm viability, along with sperm number, is arguably the one that is most likely to affect female fertility and ejaculate competitiveness based on the currently available evidence (see Section I.4). Theory predicts that when greater effort is required to obtain a mating (as expected for males with less-attractive SSCs), this will increase investment per ejaculate (Tazzyman et al., 2009; for a slightly different explanation of this principle see Engqvist, 2011). If correct, our finding of a positive relationship of SSC with sperm viability suggests that any change in investment occurs in other ejaculate traits. The most obvious candidate is a lower sperm count ('sperm depletion') for attractive males that have high mating success (e.g. Preston et al., 2001). This negative relationship might not have been detected in our meta-analysis, however, because it only occurs when males are in their natural context and mating freely. In most studies, males were held in captivity prior to ejaculate collection with ample time to replenish sperm supplies.

## (2) Missing empirical data

We conducted a systematic review so that some of the important conclusions we draw relate to data availability. Clearly, despite several reviews stressing the need for more information about the link between pre-copulatory and post-copulatory sexual selection (e.g. Birkhead & Pizzari, 2002), there is still a lack of empirical data. Although there were sufficient studies ($N = 38$) and observations ($N = 228$) to produce a reasonable estimate of the mean effect size, it was not possible to test for sources of variation with high statistical power (see Table 1). For example, we had almost no power to test for taxonomic differences. Fish and birds had the largest sample sizes (both $N = 8$ species) so the failure to detect a significant difference in effect size between these taxa is not very informative. The same issue of minimal statistical power arises for our attempts to test for variation in effect sizes for different SSC types, or based on the sperm- or ejaculate-trait assay that was used.

Many more studies are needed to identify sources of variation in the phenotypic relationship between male SSCs and ejaculate and sperm traits. Studies with modest sample sizes are worthwhile if the goal is to obtain sufficient effect size estimates to look at a higher level for broad patterns of variation. If the small mean effect sizes we report here

are reasonable estimates of true effect sizes, however, then very large samples will be required for any single study to detect a significant relationship between the expression of a male secondary sexual trait and an ejaculate measure. For example, a sample of over 800 males is required to detect a significant relationship with 80% power when $r = 0.10$ (Cohen, 1988). By comparison, the mean number of males sampled in the 38 studies we used was below 57. We therefore suggest that the main value of future studies will be to contribute to detecting general patterns, rather than providing precise estimates of specific relationships.

Arguably the most important missing data is which ejaculate traits actually influence sperm competitiveness and fertility. Are we defining 'ejaculate quality' correctly? There are, for example, surprisingly few studies that have experimentally controlled for sperm number (arguably the best empirically supported source of variation in fertilization success) to test the effects of sperm size, speed and viability on sperm competitiveness. One solution is to 'circumvent' intermediate measurement of ejaculate traits and quantify directly the relationship between SSC and fertility or paternity. Unfortunately it is generally more expensive to assign paternity than to measure ejaculate traits. There is a clear need to move beyond single-species studies to identify the extent to which we can generalise findings to other taxa. For example, are the effects of sperm length on sperm competitiveness generally similar across taxa, or are they only important in certain groups?

The situation is even worse for fertility, where almost all the evidence for ejaculate traits that affect fertility comes from domesticated animals or humans (Hasson & Stone, 2009). More studies are needed of wild animals recording the effect of a single ejaculate on female fertility. In practice this requires controlled breeding experiments, otherwise female multiple mating (e.g. extra-pair copulations) and sperm storage make it difficult to assign fertility effects to a specific male.

## (3) What do secondary sexual characters signal?

Previous meta-analyses (the least-biased compilations of primary studies) provided tentative evidence for the honesty of SSCs because their expression is positively correlated with social dominance (Santos, Scheck & Nakagawa, 2011), survivorship (Jennions, Møller & Petrie, 2001), parasite load (Harper, 1999; Møller, Christe & Lux, 1999; but see Hamilton & Poulin, 1997; Garamszegi, 2005), immune function (Møller et al., 1999) and offspring survival (Møller & Alatalo, 1999). So, why are SSCs not honest indicators of ejaculate quality? A generic answer is that male 'quality' is a broad concept because there are many ways in which males can elevate female fitness. If these different 'qualities' are uncorrelated, it is impossible for a single sexual trait to signal each aspect of male quality reliably. We then expect to see females paying attention to those SSCs that most reliably indicate factors that have a major effect on female fitness. If, as some argue, most males are highly fertile (e.g. Spottiswoode & Møller, 2004), there are minimal fitness gains derived by identifying males that elevate female

fertility (for a review of male infertility, see Hasson & Stone, 2009). A possible exception is species with inbreeding effects (Spottiswoode & Møller, 2004). Similarly, the extent to which females benefit from identifying males with more competitive ejaculates is unclear. If competitiveness is heritable (review: Simmons & Moore, 2009) then females that can identify and mate with males with competitive ejaculates should increase the competitive fertilization success of their sons. In general, however, indirect benefits are expected to be small so that female mating preferences for genetic benefits, and hence male investment in signals thereof, should be weak unless the costs of female choice are very low (Kirkpatrick & Barton, 1997). In addition, there is an assumption that sperm competitiveness is positively genetically correlated with net fitness. This might be false if there is, say, sexually antagonistic pleiotropy (e.g. Foerster *et al.*, 2007).

In conclusion, we have shown that differences in the expression of secondary sexual characters favoured by pre-copulatory sexual selection accounts for $< 1\%$ of the variance in four oft-used measures of 'ejaculate quality' thought to be under post-copulatory sexual selection. Unless polyandrous females actively bias paternity towards more attractive males via cryptic female choice (i.e. paternity does not depend directly on male ejaculate traits), males with high fitness during the initial episode of sexual selection associated with mate acquisition are no more likely to gain a disproportionate share of fertilizations. Because the SSC – ejaculate quality relationship is not negative there is no evidence that post-copulatory success directly counters selection on traits favoured by conventional female choice. The lack of a correlation will, however, introduce variation into the relationship between male mating success and the number of offspring sired, weakening the strength of net selection on male secondary sexual traits from that estimated using mating success.

(3) The correlations did not differ significantly between behavioural and morphological secondary sexual characters, or for different types of sperm traits, or between taxa. However, our ability to detect differences between taxa, secondary sexual character types, and the four sperm-quality assays had very low statistical power due to small sample sizes.

(4) There was little or no evidence for publication bias in the literature.

(5) Evidence for the actual relationship between ejaculate traits described as assays of 'ejaculate quality' and their effects on female fertility and sperm competitiveness is limited. Far more studies are needed that test directly which sperm traits elevate fertility and sperm competitiveness. Given the reported variation in the direction of the relationship across studies it is equally important to determine whether there are predictable differences across taxa and/or selective environments (see Tolle & Wagner, 2011) that can explain heterogeneity in study outcomes.

(6) Our meta-analysis attempts to synthesize quantitatively the available data for the relationship between SSCs and ejaculate quality. We note, however, that a meta-analysis is only as strong as the available data. The weaknesses with our study mirror the empirical and theoretical issues in the literature. Many studies of 'ejaculate quality' make assumptions that have little empirical support. In addition, inspection of primary papers shows that they often make species-specific claims about the relationships between specific ejaculate traits and, in turn, their relationships with sperm competitiveness and/or fertility. It is currently difficult to determine whether these are complex *post hoc* explanations to account for the current dataset, or reflect genuine heterogeneity among study systems and species. The only way to assess this is to collect sufficient data to determine whether general patterns exist and specific trends hold across relevant taxa.

## V. CONCLUSIONS

(1) The mean effect size for the relationship between secondary sexual characters and sperm/ejaculate traits associated with ejaculate quality for fertility benefits was small, regardless of which subset of the data was analysed (range: $r = 0.01–0.23$), but always positive. However, the mean effect size was not significantly different from zero. There is thus little support for SSCs being honest indicators of ejaculate quality *contra* predictions of the phenotype-linked fertility or fertility-benefits hypotheses.

(2) Given that the four sperm-trait assays we documented are also assumed to quantify 'ejaculate quality' with respect to sperm competitiveness, our results suggest that there is no phenotypically detectable trade-off between pre-copulatory and post-copulatory sexual selection. Thus a male's pre-copulatory mating success (which is affected by SSC expression) is likely to predict poorly his post-copulatory success if relative fertilization success can be predicted based on the sperm assays (see point 5 below).

## VI. ACKNOWLEDGEMENTS

## VII. REFERENCES

Alonzo, S. H. (2012). Sexual selection favors male parental care, when females can choose. *Proceedings of the Royal Society B: Biological Sciences* **279**, 1784–1790.

Andersson, M. B. (1994). Sexual Selection. Princeton University Press, Princeton.

Andersson, M. & Simmons, L. W. (2006). Sexual selection and mate choice. *Trends in Ecology & Evolution* **21**, 296–302.

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. *The American Naturalist* **165**, S26–S37.

BALL, M. & PARKER, G. (1996). Sperm competition games: external fertilization and 'adaptive' infertility. *Journal of Theoretical Biology* **180**, 141–150.

BEAUSOLEIL, J.-M. J., DOUCET, S. M., HEATH, D. D. & PITCHER, T. E. (2012). Spawning coloration, female choice and sperm competition in the redside dace, *Clinostomus elongatus*. *Animal Behaviour* **83**, 969–977.

BEGG, C. B. & MAZUMDAR, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics* **50**, 1088–1101.

BIRKHEAD, T. R., BUCHANAN, K. L., DEVOOGD, T., PELLATT, E., SZÉKELY, T. & CATCHPOLE, C. K. (1997). Song, sperm quality and testes asymmetry in the sedge warbler. *Animal Behaviour* **53**, 965–971.

BIRKHEAD, T. R. & FLETCHER, F. (1995). Male phenotype and ejaculate quality in the Zebra Finch *Taeniopygia guttata*. *Proceedings of the Royal Society B: Biological Sciences* **262**, 329–334.

BIRKHEAD, T. R., FLETCHER, F. & PELLATT, E. J. (1998). Sexual selection in the zebra finch *Taeniopygia guttata*: condition, sex traits and immune capacity. *Behavioral Ecology and Sociobiology* **44**, 179–191.

BIRKHEAD, T. R., MARTINEZ, J., BURKE, T. & FROMAN, D. (1999). Sperm mobility determines the outcome of sperm competition in the domestic fowl. *Proceedings of the Royal Society B: Biological Sciences* **266**, 1759–1764.

BIRKHEAD, T. R. & PIZZARI, T. (2002). Evolution of post-copulatory sexual selection. *Nature Reviews Genetics* **3**, 262–273.

BLOUNT, J. D., MØLLER, A. P. & HOUSTON, D. C. (2001). Antioxidants, showy males and sperm quality. *Ecology Letters* **4**, 393–396.

BLOWS, M. W. (2007). A tale of two matrices: multivariate approaches in evolutionary biology. *Journal of Evolutionary Biology* **20**, 1–8.

BORENSTEIN, M., HEDGES, L. V., HIGGINS, J. P. T. & ROTHSTEIN, H. R. (2009). *Introduction to Meta-Analysis*. First Edition. Wiley, New York.

BOSCHETTO, C., GASPARINI, C. & PILASTRO, A. (2011). Sperm number and velocity affect sperm competition success in the guppy (*Poecilia reticulata*). *Behavioral Ecology and Sociobiology* **65**, 813–821.

BUSSIÈRE, L., BASIT, H. & GWYNNE, D. (2005). Preferred males are not always good providers: female choice and male investment in tree crickets. *Behavioral Ecology* **16**, 223–231.

CALHIM, S., LAMPE, H. M., SLAGSVOLD, T. & BIRKHEAD, T. R. (2009). Selection on sperm morphology under relaxed sperm competition in a wild passerine bird. *Biology Letters* **5**, 58–61.

CAMERON, E., DAY, T. & ROWE, L. (2007). Sperm competition and the evolution of ejaculate composition. *The American Naturalist* **169**, E158–E172.

CASSEY, P., EWEN, J. G., BLACKBURN, T. M. & MØLLER, A. P. (2004). A survey of publication bias within evolutionary ecology. *Proceedings of the Royal Society B: Biological Sciences* **271**, S451–S454.

CHARGÉ, R., SAINT JALME, M., LACROIX, F., CADET, A. & SORCI, G. (2010). Male health status, signaled by courtship display, reveals ejaculate quality and hatching success in a lekking species. *Journal of Animal Ecology* **79**, 843–850.

COHEN, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Second Edition. Routledge Academic, Hillsdale.

CORNWALLIS, C. K. & BIRKHEAD, T. R. (2007). Changes in sperm quality and numbers in response to experimental manipulation of male social status and female attractiveness. *The American Naturalist* **170**, 758–770.

CÔTÉ, I., CURTIS, P., ROTHSTEIN, H. & STEWART, G. (2012). Gathering data: searching the literature and selection criteria. In *Handbook of Meta-Analysis in Ecology and Evolution* (eds J. KORICHEVA, J. GUREVITCH and K. MENGERSEN). Princeton University Press, Princeton, in press.

COTTON, S., FOWLER, K. & POMIANKOWSKI, A. (2004). Do sexual ornaments demonstrate heightened condition-dependent expression as predicted by the handicap hypothesis? *Proceedings of the Royal Society B: Biological Sciences* **271**, 771–783.

CREAN, A. J. & MARSHALL, D. J. (2008). Gamete plasticity in a broadcast spawning marine invertebrate. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 13508–13513.

DANIELSSON, I. (2001). Antagonistic pre- and post-copulatory sexual selection on male body size in a water strider (*Gerris lacustris*). *Proceedings of the Royal Society B: Biological Sciences* **268**, 77–81.

DEMARY, K. C. & LEWIS, S. M. (2007). Male courtship attractiveness and paternity success in *Photinus greeni* fireflies. *Evolution* **61**, 431–439.

DOWLING, D. K. & SIMMONS, L. W. (2009). Reactive oxygen species as universal constraints in life-history evolution. *Proceedings of the Royal Society B: Biological Sciences* **276**, 1737–1745.

DUVAL, S. & TWEEDIE, R. (2000). Trim and fill: a simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* **56**, 455–463.

DZIMINSKI, M. A., ROBERTS, J. D., BEVERIDGE, M. & SIMMONS, L. W. (2009). Sperm competitiveness in frogs: slow and steady wins the race. *Proceedings of the Royal Society B: Biological Sciences* **276**, 3955–3961.

ENGEN, F. & FOLSTAD, I. (1999). Cod courtship song: a song at the expense of dance? *Canadian Journal of Zoology* **77**, 542–550.

ENGQVIST, L. (2011). Male attractiveness is negatively genetically associated with investment in copulations. *Behavioral Ecology* **22**, 345–349.

EVANS, J. P. (2010). Quantitative genetic evidence that males trade attractiveness for ejaculate quality in guppies. *Proceedings of the Royal Society B: Biological Sciences* **277**, 3195–3201.

EVANS, J. P., PITCHER, T. E. & MAGURRAN, A. E. (2002). The ontogeny of courtship, colour and sperm production in male guppies. *Journal of Fish Biology* **60**, 495–498.

EVANS, J. P., ZANE, L., FRANCESCATO, S. & PILASTRO, A. (2003). Directional postcopulatory sexual selection revealed by artificial insemination. *Nature* **421**, 360–363.

FIRMAN, R. C. & SIMMONS, L. W. (2010). Experimental evolution of sperm quality via postcopulatory sexual selection in house mice. *Evolution* **64**, 1245–1256.

FITZPATRICK, J. L. & EVANS, J. P. (2009). Reduced heterozygosity impairs sperm quality in endangered mammals. *Biology Letters* **5**, 320–323.

FITZPATRICK, J. L., MONTGOMERIE, R., DESJARDINS, J. K., STIVER, K. A., KOLM, N. & BALSHINE, S. (2009). Female promiscuity promotes the evolution of faster sperm in cichlid fishes. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 1128–1132.

FOERSTER, K., COULSON, T., SHELDON, B. C., PEMBERTON, J. M., CLUTTON-BROCK, T. H. & KRUUK, L. E. B. (2007). Sexually antagonistic genetic variation for fitness in red deer. *Nature* **447**, 1107–1110.

FROMAN, D., FELTMANN, A., RHOADS, M. & KIRBY, J. (1999). Sperm mobility: a primary determinant of fertility in the domestic fowl (*Gallus domesticus*). *Biology of Reproduction* **61**, 400–405.

FRY, C. & WILKINSON, G. (2004). Sperm survival in female stalk-eyed flies depends on seminal fluid and meiotic drive. *Evolution* **58**, 1622–1626.

GAGE, M., MACFARLANE, C., YEATES, S., WARD, R., SEARLE, J. & PARKER, G. (2004). Spermatozoal traits and sperm competition in Atlantic salmon: relative sperm velocity is the primary determinant of fertilization success. *Current Biology* **14**, 44–47.

GAGE, M. & MORROW, E. (2003). Experimental evidence for the evolution of numerous, tiny sperm via sperm competition. *Current Biology* **13**, 754–757.

GARAMSZEGI, L. (2005). Bird song and parasites. *Behavioral Ecology and Sociobiology* **59**, 167–180.

GARCÍA-GONZÁLEZ, F. & SIMMONS, L. W. (2005). Sperm viability matters in insect sperm competition. *Current Biology* **15**, 271–275.

GARCÍA-GONZÁLEZ, F. & SIMMONS, L. W. (2007). Shorter sperm confer higher competitive fertilization success. *Evolution* **61**, 816–824.

GASPARINI, C., MARINO, I. a. M., BOSCHETTO, C. & PILASTRO, A. (2010a). Effect of male age on sperm traits and sperm competition success in the guppy (*Poecilia reticulata*). *Journal of Evolutionary Biology* **23**, 124–135.

GASPARINI, C., SIMMONS, L. W., BEVERIDGE, M. & EVANS, J. P. (2010b). Sperm swimming velocity predicts competitive fertilization success in the green swordtail *Xiphophorus helleri*. *PLoS ONE* **5**, e12146.

GASPARINI, C., PERETTI, A. V. & PILASTRO, A. (2009). Female presence influences sperm velocity in the guppy. *Biology Letters* **5**, 792–794.

GETTY, T. (2006). Sexually selected signals are not similar to sports handicaps. *Trends in Ecology & Evolution* **21**, 83–88.

GOMEZ-MONTOTO, L., MAGANA, C., TOURMENTE, M., MARTIN-COELLO, J., CRESPO, C., LUQUE-LARENA, J. J., GOMENDIO, M. & ROLDAN, E. R. S. (2011). Sperm competition, sperm numbers and sperm quality in muroid rodents. *PLoS ONE* **6**, e18173.

GRAFEN, A. (1990). Biological signals as handicaps. *Journal of Theoretical Biology* **144**, 517–546.

GRIFFITH, S. C. (2007). The evolution of infidelity in socially monogamous passerines: neglected components of direct and indirect selection. *The American Naturalist* **169**, 274–281.

HAMILTON, W. J. & POULIN, R. (1997). The Hamilton and Zuk hypothesis revisited: a meta-analytical approach. *Behaviour* **134**, 299–320.

HARPER, D. G. C. (1999). Feather mites pectoral muscle condition wing length and plumage coloration of passerines. *Animal Behaviour* **58**, 553–562.

HASSON, O. & STONE, L. (2009). Male infertility, female fertility and extrapair copulations. *Biological Reviews* **84**, 225–244.

HELFENSTEIN, F., LOSDAT, S., MØLLER, A. P., BLOUNT, J. D. & RICHNER, H. (2010). Sperm of colourful males are better protected against oxidative stress. *Ecology Letters* **13**, 213–222.

HETTYEY, A., HERCZEG, G. & HOI, H. (2009). Testing the phenotype-linked fertility hypothesis in male Moor Frogs (*Rana arvalis*) exhibiting a conspicuous nuptial colouration. *Amphibia-Reptilia* **30**, 581–586.

HEYWOOD, J. (1989). Sexual selection by the handicap mechanism. *Evolution* **43**, 1387–1397.

HILL, G. (2011). Condition-dependent traits as signals of the functionality of vital cellular processes. *Ecology Letters* **14**, 625–634.

HOEZLER, G. (1989). The good parent process of sexual selection. *Animal Behaviour* **38**, 1067–1078.

HOSKEN, D. J., TAYLOR, M. L., HOYLE, K., HIGGINS, S. & WEDELL, N. (2008). Attractive males have greater success in sperm competition. *Current Biology* **18**, R553–R554.

HUNT, J., BREUKER, C. J., SADOWSKI, J. A. & MOORE, A. J. (2009). Male–male competition, female mate choice and their interaction: determining total sexual selection. *Journal of Evolutionary Biology* **22**, 13–26.

HUNT, J., BUSSIÈRE, L., JENNIONS, M. & BROOKS, R. (2004). What is genetic quality? *Trends in Ecology & Evolution* **19**, 329–333.

HUNTER, F. & BIRKHEAD, T. (2002). Sperm viability and sperm competition in insects. *Current Biology* **12**, 121–123.

IWASA, Y. & POMIANKOWSKI, A. (1999). Good parent and good genes models of handicap evolution. *Journal of Theoretical Biology* **200**, 97–109.

IWASA, Y., POMIANKOWSKI, A. & NEE, S. (1991). The evolution of costly mate preferences. 2. The handicap principle. *Evolution* **45**, 1431–1442.

JANHUNEN, M., RUDOLFSEN, G., KEKÄLÄINEN, J., FIGENSCHOU, L., PEUHKURI, N. & KORTET, R. (2009). Spawning coloration and sperm quality in a large lake population of Arctic charr (Salmonidae: *Salvelinus alpinus* L.). *Biological Journal of the Linnean Society* **98**, 794–802.

JENNIONS, M. D., LORTIE, C., ROSENBERG, M. S. & ROTHSTEIN, H. (2012). Publication and related biases. In *Handbook of Meta-Analysis in Ecology and Evolution*, (eds J. KORICHEVA J. GUREVITCH and K. MENGERSEN). Princeton University Press, Princeton, in press.

JENNIONS, M. D., MØLLER, A. P. & PETRIE, M. (2001). Sexually selected traits and adult survival: a meta-analysis. *The Quarterly Review of Biology* **76**, 3–36.

JENNIONS, M. D. & PETRIE, M. (2000). Why do females mate multiply? A review of the genetic benefits. *Biological Reviews* **75**, 21–64.

KELLY, C. D. (2008). Sperm investment in relation to weapon size in a male trimorphic insect? *Behavioral Ecology* **19**, 1018–1024.

KELLY, C. D. & JENNIONS, M. D. (2011). Sexual selection and sperm quantity: meta-analyses of strategic ejaculation. *Biological Reviews* **86**, 863–884.

KELLY, N. B. & ALONZO, S. H. (2009). Will male advertisement be a reliable indicator of paternal care, if offspring survival depends on male care? *Proceedings of the Royal Society B: Biological Sciences* **276**, 3175–3183.

KIRKPATRICK, M. & BARTON, N. (1997). The strength of indirect selection on female mating preferences. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 1282–1286.

KLAUS, S. P., FITZSIMMONS, L. P., PITCHER, T. E. & BERTRAM, S. M. (2011). Song and Sperm in crickets: a trade-off between pre- and post-copulatory traits or phenotype-linked fertility? *Ethology* **117**, 154–162.

KOKKO, H., BROOKS, R., McNAMARA, J. & HOUSTON, A. (2002). The sexual selection continuum. *Proceedings of the Royal Society B: Biological Sciences* **269**, 1331–1340.

KOKKO, H., JENNIONS, M. D. & BROOKS, R. (2006). Unifying and testing models of sexual selection. *Annual Review of Ecology, Evolution, and Systematics* **37**, 43–66.

KOKKO, H., KLUG, H. & JENNIONS, M. D. (2012). Unifying cornerstones of sexual selection: operational sex ratio, Bateman gradient, and the scope for competitive investment. *Ecology Letters* **15**, 1340–1351.

KORTET, R., VAINIKKA, A., RANTALA, M. J. & TASKINEN, J. (2004). Sperm quality, secondary sexual characters and parasitism in roach (*Rutilus rutilus* L.). *Biological Journal of the Linnean Society* **81**, 111–117.

KOTIAHO, J. S. (2000). Testing the assumptions of conditional handicap theory: costs and condition dependence of a sexually selected trait. *Behavioral Ecology and Sociobiology* **48**, 188–194.

KOTIAHO, J. S. (2001). Costs of sexual traits: a mismatch between theoretical considerations and empirical evidence. *Biological Reviews* **76**, 365–376.

KROKENE, C., RIGSTAD, K., DALE, M. & LIFJELD, J. (1998). The function of extrapair paternity in blue tits and great tits: good genes or fertility insurance? *Behavioral Ecology* **9**, 649–656.

LAILVAUX, S. P. & KASUMOVIC, M. M. (2011). Defining individual quality over lifetimes and selective contexts. *Proceedings of the Royal Society B: Biological Sciences* **278**, 321–328.

LAMUNYON, C. & WARD, S. (1999). Evolution of sperm size in nematodes: sperm competition favours larger sperm. *Proceedings of the Royal Society B: Biological Sciences* **266**, 263–267.

LIFJELD, J. T., LASKEMOEN, T., KLEVEN, O., PEDERSEN, A. T. M., LAMPE, H. M., RUDOFLSEN, G., SCHMOLL, T. & SLAGSVOLD, T. (2012). No evidence for pre-copulatory sexual selection on sperm length in a passerine bird. *PLoS ONE* **7**, e32611.

LILJEDAL, S., FOLSTAD, I. & SKARSTEIN, F. (1999). Secondary sex traits, parasites, immunity and ejaculate quality in the Arctic charr. *Proceedings of the Royal Society B: Biological Sciences* **266**, 1893–1898.

LILJEDAL, S., RUDOLFSEN, G. & FOLSTAD, I. (2008). Factors predicting male fertilization success in an external fertilizer. *Behavioral Ecology and Sociobiology* **62**, 1805–1811.

LINHART, O., RODINA, M., GELA, D., KOCOUR, M. & VANDEPUTTE, M. (2005). Spermatozoal competition in common carp (*Cyprinus carpio*): what is the primary determinant of competition success? *Reproduction* **130**, 705–711.

LOCATELLO, L., RASOTTO, M. B., EVANS, J. P. & PILASTRO, A. (2006). Colourful male guppies produce faster and more viable sperm. *Journal of Evolutionary Biology* **19**, 1595–1602.

MALO, A. F., ROLDAN, E. R. S., GARDE, J., SOLER, A. J. & GOMENDIO, M. (2005). Antlers honestly advertise sperm production and quality. *Proceedings of the Royal Society B: Biological Sciences* **272**, 149–157.

MASVAER, M., LILJEDAL, S. & FOLSTAD, I. (2004). Are secondary sex traits, parasites and immunity related to variation in primary sex traits in the Arctic charr? *Proceedings of the Royal Society B: Biological Sciences* **271**, S40–S42.

MATTHEWS, I. M., EVANS, J. P. & MAGURRAN, A. E. (1997). Male display rate reveals ejaculate characteristics in the Trinidadian guppy, *Poecilia reticulata*. *Proceedings of the Royal Society B: Biological Sciences* **264**, 695–700.

MENGERSEN, K., JENNIONS, M. D. & SCHMID, J. (2012). Statistical models for the meta-analysis of non-independent data. In *Handbook of Meta-Analysis in Ecology and Evolution* (eds J. KORICHEVA, J. GUREVITCH and K. MENGERSEN). Princeton University Press, Princeton, in press.

MILLER, G. & PITNICK, S. (2002). Sperm-female coevolution in *Drosophila*. *Science* **298**, 1230–1233.

MØLLER, A. P. (1994). *Sexual selection and the barn swallow*. Oxford University Press, Oxford.

MØLLER, A. P. & ALATALO, R. V. (1999). Good-genes effects in sexual selection. *Proceedings of the Royal Society B: Biological Sciences* **266**, 85–91.

MØLLER, A. P., CHRISTE, P. & LUX, E. (1999). Parasitism, host immune function and sexual selection. *The Quarterly Review of Biology* **74**, 3–20.

MØLLER, A. P. & DE LOPE, F. (1994). Differential costs of a secondary sexual character: an experimental test of the handicap principle. *Evolution* **48**, 1676–1683.

MØLLER, A. P. & JENNIONS, M. D. (2001). How important are direct fitness benefits of sexual selection? *Naturwissenschaften* **88**, 401–415.

MØLLER, A. P. & JENNIONS, M. D. (2002). How much variance can be explained by ecologists and evolutionary biologists? *Oecologia* **132**, 492–500.

MØLLER, A. & THORNHILL, R. (1998). Male parental care, differential parental investment by females and sexual selection. *Animal Behaviour* **55**, 1507–1515.

MORTIMER, D., PANDYA, I. & SAWERS, R. (1986). Relationship between human-sperm motility characteristics and sperm penetration into human cervical-mucus *in vitro*. *Journal of Reproduction and Fertility* **78**, 93–102.

MURAI, M., BACKWELL, P. R. Y. & JENNIONS, M. D. (2009). The cost of reliable signaling: experimental evidence for predictable variation among males in a cost-benefit trade-off between sexually selected traits. *Evolution* **63**, 2363–2371.

MURTAUGH, P. A. (2002). Journal quality, effect size, and publication bias in meta-analysis. *Ecology* **83**, 1162–1166.

OPPLIGER, A., NACIRI-GRAVEN, Y., RIBI, G. & HOSKEN, D. (2003). Sperm length influences fertilization success during sperm competition in the snail *Viviparus ater*. *Molecular Ecology* **12**, 485–492.

PARKER, G. A. (1998). Sperm competition and the evolution of ejaculates: towards a theory base. In Sperm Competition and Sexual Selection (eds T. R. BIRKHEAD and A. P. MØLLER), pp. 3–54. Academic University Press, New York.

PARKER, G. A. & PIZZARI, T. (2010). Sperm competition and ejaculate economics. *Biological Reviews* **85**, 897–934.

PARKER, T. H., THOMPSON, D., LIGON, J. D., SCHNEIDER, B. & BYRN, F. (2006). Does red junglefowl comb size predict sperm swimming speed and motility? *Ethology Ecology and Evolution* **18**, 53–60.

PATTARINI, J. A., STARMER, W. T., BJORK, A. & PITNICK, S. (2006). Mechanisms underlying the sperm quality advantage in *Drosophila melanogaster*. *Evolution* **60**, 2064–2080.

PERRY, J. C. & ROWE, L. (2010). Condition-dependent ejaculate size and composition in a ladybird beetle. *Proceedings of the Royal Society B: Biological Sciences* **277**, 3639–3647.

PETERS, A., DENK, A. G., DELHEY, K. & KEMPENAERS, B. (2004). Carotenoid-based bill colour as an indicator of immunocompetence and sperm performance in male mallards. *Journal of Evolutionary Biology* **17**, 1111–1120.

PILASTRO, A. & BISAZZA, A. (1999). Insemination efficiency of two alternative male mating tactics in the guppy *Poecilia reticulata*. *Proceedings of the Royal Society B: Biological Sciences* **266**, 1887–1891.

PILASTRO, A., EVANS, J. P., SARTORELLI, A. & BISAZZA, A. (2002). Male phenotype predicts insemination success in guppies. *Proceedings of the Royal Society B: Biological Sciences* **269**, 1325–1330.

PILASTRO, A., SIMONATO, M., BISAZZA, A. & EVANS, J. P. (2004). Cryptic female preference for colorful males in guppies. *Evolution* **58**, 665–669.

PITCHER, T. E., DOUCET, S. M., BEAUSOLEIL, J. M. J. & HANLEY, D. (2009). Secondary sexual characters and sperm traits in coho salmon *Oncorhynchus kisutch*. *Journal of Fish Biology* **74**, 1450–1461.

PITCHER, T. E. & EVANS, J. P. (2001). Male phenotype and sperm number in the guppy (*Poecilia reticulata*). *Canadian Journal of Zoology* **79**, 1891–1896.

PITCHER, T. E., NEFF, B. D., RODD, F. H. & ROWE, L. (2003). Multiple mating and sequential mate choice in guppies: females trade up. *Proceedings of the Royal Society B: Biological Sciences* **270**, 1623–1629.

PITCHER, T. E., RODD, F. H. & ROWE, L. (2007). Sexual colouration and sperm traits in guppies. *Journal of Fish Biology* **70**, 165–177.

PITNICK, S., HOSKEN, D. J. & BIRKHEAD, T. R. (2009). Sperm morphological diversity. In Sperm Biology: An Evolutionary Perspective (eds T. R. BIRKHEAD, D. J. HOSKEN and S. PITNICK), pp. 69–149. Academic Press, Burlington.

Pizzari, T., Worley, K., Burke, T. & Froman, D. P. (2008). Sperm competition dynamics: ejaculate fertilising efficiency changes differentially with time. *BMC Evolutionary Biology* **8**, 332.

Preston, B., Stevenson, I., Pemberton, J. & Wilson, K. (2001). Dominant rams lose out by sperm depletion: a waning success in siring counters a ram's high score in competition for ewes. *Nature* **409**, 681–682.

Price, T., Schluter, D. & Heckman, N. (1993). Sexual selection when the female directly benefits. *Biological Journal of the Linnean Society* **48**, 187–211.

Radwan, J. (1996). Intraspecific variation in sperm competition success in the bulb mite: a role for sperm size. *Proceedings of the Royal Society B: Biological Sciences* **263**, 855–859.

Reznick, D., Nunney, L. & Tessier, A. (2000). Big houses, big cars, superfleas and the costs of reproduction. *Trends in Ecology & Evolution* **15**, 421–425.

Roff, D. A. (2002). Life History Evolution. Sinauer Associates Inc., Sunderland.

Roff, D. A. & Fairbairn, D. J. (2007). The evolution of trade-offs: where are we? *Journal of Evolutionary Biology* **20**, 433–447.

Rohlenová, K. & Šimková, A. (2010). Are the immunocompetence and the presence of metazoan parasites in cyprinid fish affected by reproductive efforts of cyprinid fish? *Journal of Biomedicine and Biotechnology* **2010**, 418312.

Rosenberg, M. S., Adams, D. C. & Gurevitch, J. (1997). MetaWin Statistical Software for Meta-Analysis with Resampling Tests. Sinauer Associates Inc., Sunderland.

Rowe, L. & Houle, D. (1996). The lek paradox and the capture of genetic variance by condition dependent traits. *Proceedings of the Royal Society B: Biological Sciences* **263**, 1415–1421.

Rowe, M. & Pruett-Jones, S. (2011). Sperm competition selects for sperm quantity and quality in the Australian Maluridae. *PLoS ONE* **6**(1), e15720.

Rowe, M., Swaddle, J. P., Pruett-Jones, S. & Webster, M. S. (2010). Plumage coloration, ejaculate quality and reproductive phenotype in the red-backed fairy-wren. *Animal Behaviour* **79**, 1239–1246.

Rudolfsen, G., Müller, R., Urbach, D. & Wedekind, C. (2008). Predicting the mating system from phenotypic correlations between life-history and sperm quality traits in the Alpine whitefish *Coregonus zugensis*. *Behavioral Ecology and Sociobiology* **62**, 561–567.

Santos, E. S., Scheck, A. D. & Nakagawa, S. (2011). Dominance and plumage traits: meta-analysis and metaregression analysis. *Animal Behavior* **82**, 3–19.

Schluter, D. & Price, T. (1993). Honesty, perception and population divergence in sexually selected traits. *Proceedings of the Royal Society B: Biological Sciences* **253**, 117–122.

Sheldon, B. C. (1994). Male phenotype, fertility, and the pursuit of extra-pair copulations by female birds. *Proceedings of the Royal Society B: Biological Sciences* **257**, 25–30.

Sherman, C. D. H., Wapstra, E. & Olsson, M. (2009). Consistent male-male paternity differences across female genotypes. *Biology Letters* **5**, 232–234.

Sherman, C. D. H., Wapstra, E., Uller, T. & Olsson, M. (2008). Males with high genetic similarity to females sire more offspring in sperm competition in Peron's tree frog *Litoria peronii*. *Proceedings of the Royal Society B: Biological Sciences* **275**, 971–978.

Simmons, L. W. & Emlen, D. J. (2006). Evolutionary trade-off between weapons and testes. *Proceedings of the National Academy of Sciences* **103**, 16346–16351.

Simmons, L. & Kotiaho, J. (2002). Evolution of ejaculates: patterns of phenotypic and genotypic variation and condition dependence in sperm competition traits. *Evolution* **56**, 1622–1631.

Simmons, L. W. & Moore, A. J. (2009). Evolutionary quantitative genetics of sperm. In *Sperm Biology: An Evolutionary Perspective* (eds T. R. Birkhead, D. J. Hoskens and S. Pitnick), pp. 405–434. Academic Press, Burlington.

Simmons, L., Wernham, J., García-González, F. & Kamien, D. (2003). Variation in paternity in the field cricket *Teleogryllus oceanicus*: no detectable influence of sperm numbers or sperm length. *Behavioral Ecology* **14**, 539–545.

Skau, P. A. & Folstad, I. (2003). Do bacterial infections cause reduced ejaculate quality? A meta-analysis of antibiotic treatment of male infertility. *Behavioral Ecology* **14**, 40–47.

Skau, P. A. & Folstad, I. (2005). Does immunity regulate ejaculate quality and fertility in humans? *Behavioral Ecology* **16**, 410–416.

Skinner, A. M. J. & Watt, P. J. (2007). Phenotypic correlates of spermatozoon quality in the guppy, *Poecilia reticulata*. *Behavioral Ecology* **18**, 47–52.

Slatyer, R. A., Mautz, B. S., Backwell, P. R. Y. & Jennions, M. D. (2011). Estimating genetic benefits of polyandry from experimental studies: a meta-analysis. *Biological Reviews* **87**, 1–33.

Smith, C. C. (2012). Opposing effects of sperm viability and velocity on the outcome of sperm competition. *Behavioral Ecology* **23**, 820–826.

Snook, R. S. (2005). Sperm in competition: not playing by the numbers. *Trends in Ecology & Evolution* **20**, 46–53.

Soler, C., Núñez, M., Gutiérreza, R., Núñez, J., Medina, P., Sancho, M., Álvarez, J. & Núñez, A. (2003). Facial attractiveness in men provides clues to semen quality. *Evolution and Human Behavior* **24**, 199–207.

Spottiswoode, C. & Møller, A. P. (2004). Genetic similarity and hatching success in birds. *Proceedings of the Royal Society B: Biological Sciences* **271**, 267–272.

Stockley, P., Gage, M., Parker, G. & Møller, A. (1997). Sperm competition in fishes: the evolution of testis size and ejaculate characteristics. *The American Naturalist* **149**, 933–954.

Tazzyman, S. J., Pizzari, T., Seymour, R. M. & Pomiankowski, A. (2009). The evolution of continuous variation in ejaculate expenditure strategy. *The American Naturalist* **174**, E71–E82.

Tolle, A. E. & Wagner, W. E. Jr. (2011). Costly signals in a field cricket can indicate high- or low-quality direct benefits depending upon the environment. *Evolution* **65**, 283–294.

Tomkins, J., Radwan, J., Kotiaho, J. & Tregenza, T. (2004). Genic capture and resolving the lek paradox. *Trends in Ecology & Evolution* **19**, 323–328.

Tourmente, M., Gomendio, M. & Roldan, E. R. S. (2011a). Mass-specific metabolic rate and sperm competition determine sperm size in marsupial mammals. *PLoS ONE* **6**(6): e21244.

Tourmente, M., Gomendio, M. & Roldan, E. R. S. (2011b). Sperm competition and the evolution of sperm design in mammals. *BMC Evolutionary Biology* **11**, 12.

Tourmente, M., Gomendio, M., Roldan, E. R. S., Giojalas, L. C. & Chiaraviglio, M. (2009). Sperm competition and reproductive mode influence sperm dimensions and structure among snakes. *Evolution* **63**, 2513–2524.

Trivers, R. L. (1972). Parental investment and sexual selection. In *Sexual Selection and the Descent of Man* (ed. B. Campbell), pp. 136–179. Heinemann, London.

Van Noordwijk, A. J. & De Jong, G. (1986). Acquisition and allocation of resources: their influence on variation in life history tactics. *The American Naturalist* **128**, 137–142.

Wagner, W. E. Jr. (2011). Direct benefits and the evolution of female mating preferences: conceptual problems, potential solutions, and a field cricket. *Advances in the Study Behavior* **43**, 273–319.

Wagner, W. E. Jr. & Harper, C. J. (2003). Female life span and fertility are increased by the ejaculates of preferred males. *Evolution* **57**, 2054–2066.

Wedell, N., Gage, M. J. G. & Parker, G. A. (2002). Sperm competition, male prudence and sperm-limited females. *Trends in Ecology & Evolution* **17**, 313–320.

Wetton, J. & Parker, D. (1991). An association between fertility and cuckoldry in the house sparrow, *Passer domesticus*. *Proceedings of the Royal Society B: Biological Sciences* **245**, 227–233.

Williams, G. C. (1992). *Natural Selection: Domains, Levels, and Challenges*. Oxford University Press, New York.

Wilson, A. J. & Nussey, D. H. (2010). What is individual quality? An evolutionary perspective. *Trends in Ecology & Evolution* **25**, 207–214.

World Health Organization (1999). WHO Laboratory manual for the examination of human semen and sperm-cervical mucus interactions. Fourth Edition. Cambridge University Press, Cambridge.

Zahavi, A. (1975). Mate selection - selection for a handicap. *Journal of Theoretical Biology* **53**, 205–214.

Zajitschek, S. R. K. & Brooks, R. C. (2010). Inbreeding depression in male traits and preference for outbred males in *Poecilia reticulata*. *Behavioral Ecology* **21**, 884–891.

# VIII. SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article.

**Appendix S1.** ISI Web of Science key word search terms used to identify relevant publications.

**Appendix S2.** Individual effects sizes.

**Appendix S3.** Species-level effect sizes (*i*) across all sperm and secondary sexual character (SSC) traits, (*ii*) for behavioural traits across all sperm categories, and (*iii*) for morphological traits across sperm categories.