A GENERAL DESCRIPTION OF ADDITIVE AND NONADDITIVE ELEMENTS OF SPERM COMPETITIVENESS AND THEIR RELATION TO MALE FERTILIZATION SUCCESS

Leif Engqvist[1,2,3]

[1]Evolutionary Biology, Bielefeld University, Bielefeld, Germany

[2]Centre for Ecological and Evolutionary Studies, Rijksuniversiteit Groningen, The Netherlands

[3]E-mail: leif.engqvist@uni-bielefeld.de

A complete understanding of male reproductive success, and thus sexual selection, often requires an insight into male success in sperm competition. Genuine conclusions on male sperm competitiveness can only be made in real competitive situations. However, statistical analyses of sperm competitiveness from fertilization success data have been shown to be problematic. Here, I first outline a comprehensive general description of the different additive and nonadditive elements relevant for the outcome of sperm competition staged between two males. Based on this description, I will highlight two main problems that are frequently encountered in experiments aiming at estimating sperm competitiveness. First, I focus on potential problems when using standardized competitors versus random mating trials, because trials with standardized competitors do not allow generalization if male–male interactions are important. Second, I illustrate the necessity to analyze data on the logit scale rather than on raw proportions, because only the logit scale allows a clean separation of additive and nonadditive effects (i.e., male × male and female × male interactions).

KEY WORDS: Experimental design, female-by-male interaction, interacting phenotypes, male-by-male interaction, multiple mating, reproductive success, sexual selection, sperm competition.

Sperm competition is a strong selective force that has been shown to have considerable effects on many aspects of male reproductive biology (Birkhead et al. 2009; Birkhead and Møller 1998; Parker 1970; Simmons 2001). When females mate with more than one male and sperm from rival males compete to fertilize the ova, male reproductive success will not only be affected by a male's ability to achieve matings, but also by the success of his sperm in the subsequent competition for fertilizations. Thus, in addition to conspicuous male weaponry and secondary sexual ornamentation alleged to have evolved through intra- and intersexual selection (Andersson 1994), female polyandry will instigate selection on male ability to outcompete the sperm of other males (Parker

1970; Simmons 2001). Similar to variation in male mating success (Andersson 1994; Andersson and Simmons 2006), it seems that males in many species also differ in their sperm competitive ability (Dziuk 1996; Engqvist et al. 2007; García-González and Simmons 2005; Keller and Reeve 1995; Lewis and Austad 1990; Sherman et al. 2009; Simmons and Parker 1992). To fully understand sexual selection and the relative contribution of pre- and postcopulatory processes to the variation in male reproductive success, we need to be able to quantify sperm competitiveness.

Differences in sperm competitive ability can be manifested in various traits such as for instance sperm number, velocity, or viability (Engqvist et al. 2007; Gage and Morrow 2003;

García-González and Simmons 2005; Snook 2005). However, ultimately what matters for male reproductive success is the actual outcome of sperm competition. A male's ability to outcompete others males' sperm is thus often measured in actual competitive situations, that is by assigning paternity to offspring of females that have mated with several males (e.g., Engqvist et al. 2007; Fricke et al. 2010; Sakaluk and Eggert 1996). Sperm competition success is often quantified as P_2, defined as the proportion of offspring sired by the second mating partner of a given female in double mating trials (Boorman and Parker 1976). Alternatively, one can also analyze the P_1-value, which quantifies the proportion of offspring sired by the first of the two mates. The average values of P_1 and P_2 across males reflects the population level sperm precedence pattern (Simmons and Siva-Jothy 1998) and P_2 inevitably will equal $1 - P_1$. Sperm precedence values for individual males do not follow this constraint. A male can in principle be relatively successful both when mating first and last. Or, for instance, have a disproportionately high success when mating last but lower relative success when mating first. Thus, male sperm competitive ability is not always a single specific attribute, but needs to be specified regarding mating sequence. In this respect, individual P_1-values quantify males' sperm competition "defense" abilities and P_2 measures the sperm competition "offence" abilities of males (see, e.g. Fricke et al. 2010). For example, a typical offense characteristic is a male's ability to displace the sperm from previous female mating partners, whereas the ability to resist such displacement would be a defense characteristic (Clark 2002). An important consequence is that an assay of males' offensive sperm competitiveness will be influenced by the sperm defense properties of the male competitors and vice versa. This is an inevitable consequence of the fact that the competition outcome can only be measured in the interaction of multiple individuals, and this will potentially reduce the statistical confidence in the estimates of sperm competitiveness considerably (García-González 2008b; García-González and Evans 2011). In fact, the outcome of sperm competition between two males will be affected by quite a number of different factors such as (i) the species'/populations' average sperm precedence (Simmons and Siva-Jothy 1998); (ii) variation across females in their influence of sperm precedence across males (Clark and Begun 1998); (iii) between-male differences in sperm competitiveness (Lewis and Austad 1990; Prout and Bundgaard 1977; Sherman et al. 2009; Simmons and Parker 1992); (iv) between-male variation in order effects, that is whether different males are relatively superior in defense or offence functions (Fricke et al. 2010; Michalczyk et al. 2010); (v) female × male interactions (Bjork et al. 2007; Clark 2002; Clark et al. 1999; Miller and Pitnick 2002; Wilson et al. 1997); (vi) male × male interactions (Bjork et al. 2007; Clark et al. 2000; Prout and Bundgaard 1977); and even (vii) female × male × male interactions are possible

(Bretman et al. 2004; Tregenza and Wedell 2002). To increase statistical power and minimize potential confounding effects, it has been suggested to perform sperm competition assays using standardized competitors (Droge-Young et al. 2012; Fricke et al. 2010; García-González 2008b; García-González and Evans 2011). The presented arguments in favor of using standard competitors in the estimation of sperm competition ability have been twofold: First, to minimize the sampling variation induced by picking competitors that differ in their overall effects and/or their relative qualities in defensive/offensive sperm competition (García-González 2008b; García-González and Evans 2011); second to minimize variation induced by male × male or female × male interactions (Droge-Young et al. 2012; Fricke et al. 2010; García-González and Evans 2011).

The aim of this article is to further dissect these unexpectedly complex aspects of sperm competitiveness and how they relate to fertilization success. I will start with the formulation of a general framework that captures additive effects as well as the interaction effects. I will focus on the simplified situation of only two mating partners, because this is the classical situation that is most frequently analyzed, but the results can easily be generalized to more than two males. This formalization alone will reveal some new interesting insights. I will then use this model to assess the pros and cons of some of the presently suggested and used methods to estimate sperm competitiveness and genetic variability of this trait. We will for instance see that standardizing mating trials regarding females and competing males indeed increases the precision in estimates of sperm competitiveness, but will also potentially introduce more severe systematic errors, as the estimated sperm competitiveness is not generalizable to all mating constellations. We will also see that data needs to be analyzed on an appropriate scale, otherwise purely additive male effects can easily be mistaken for nontransitive interaction effects.

TRANSFORMATIONS AND ADDITIVE EFFECTS
As described earlier, the outcome of sperm competition between two males will be affected by a number of transitive and nontransitive effects (i.e., effects that are either general or affected by the specific combination of males and females involved). The first problem encountered when attempting to predict the resulting fertilization pattern as a combination of these different factors is that paternity, the proportion of offspring sired, is inherently nonlinear (Eggert et al. 2003; García-González 2008b; Parker et al. 1990). Fertilization success will often depend on the amount of sperm, or other aspects of sperm competitiveness, in relation to that of other competitors. The most natural way to linearize paternity values, and thus to simplify analyses, is therefore to express it as the logarithm of the odds ratio, and thus to transform paternity values to a logit scale: that is $\tilde{P}_2 = \log[P_2/(1 - P_2)]$.

Most commonly, we are interested in characterizing the overall sperm competitiveness of individual males, that is a male's competitiveness across all mating situations. In the simplest case, this will be the only factor affecting the outcome of sperm competition. Nevertheless, the outcome of sperm competition will be influenced of the competitiveness of both males involved. Let M be a vector of length n that describes the sperm competitiveness of all n male types in the population (e.g., different individuals or genotypes). The expected outcome of sperm competition between male j mating first and male k mating second can then be described by $(\tilde{P}_2)_{jk} = \mu - M_j + M_k$. Here μ describes the expected second male precedence if the males are equal competitors and corresponds to what has been termed the loading factor (Parker 1990). The simplicity of this function originates from the symmetry of the log odds ratio; note that $\tilde{P}_2 = -\tilde{P}_1$. The function can also be deduced directly from sperm competition in terms of a fair/loaded raffle: $P_2 = rs_2/(rs_2 + s_1)$ (Parker 1990). Here s_1 and s_2 corresponds to the ejaculate investment (sperm competitive ability) of the first and second male to mate and r the (dis)-advantage of the second male (loading factor). If we transform this to a logit scale, we get $\tilde{P}_2 = \log(r) - \log(s_1) + \log(s_2)$ (see also Eggert et al. 2003), which is equivalent with the equation shown above.

Until now we have ignored that males might differ in their offence and defense abilities (i.e., they might not be as good being the second male to mate as they are as the first male to mate and vice versa). One way to resolve this is to define two properties represented by the vectors M^1 and M^2 describing male competitiveness in the first and second role respectively. Then $(\tilde{P}_2)_{jk} = \mu - M_j^1 + M_k^2$. Alternatively, if we choose to focus more on overall sperm competitiveness, we can do this with M and additionally let M^O define whether a male is relatively better in the defense or the offence role (the index O refers to order of mating). In that case $(\tilde{P}_2)_{jk} = \mu - (M_j + M_j^O) + (M_k + M_k^O)$.

Finally, we can assume that individual females differ in the way second males are favored (and first males disfavored), and this is described by F. We get $(\tilde{P}_2)_{ijk} = \mu + F_i - M_j^1 + M_k^2$, alternatively $(\tilde{P}_2)_{ijk} = \mu + F_i - (M_j + M_j^O) + (M_k + M_k^O)$, for a male j mating first and male k mating second with a female i. To see how we get this, we can think about a loaded raffle in which the parameter describing the advantage of the second male to mate is specific for each female (i.e., r_i): $P_2 = r_i s_2/(r_i s_2 + s_1)$. After transforming this to logit scale we get $(\tilde{P}_2)_{ijk} = \log(r_i) - \log(s_1) + \log(s_2)$, where $\log(r_i)$ can be portioned into an overall effect μ, representing the expected second male precedence if two equal competitors mate with an "average" female and F_i, representing female specific deviations from this. Note that until now we have an entirely additive model. We now can start thinking about the possible interactions and how they need to be implemented.

FEMALE × MALE INTERACTIONS

A female × male interaction implies that certain female–male mating combinations results in a different outcome than expected from the purely additive model described earlier. Let the matrix $I^{F \times M}$ describe the way certain female–male combinations result in a biased outcome. However, females mate with two males; thus, we have to take two female–male interactions into account. Analogous to the additive model, we then get the expression $[-I_{ij}^{F \times M} + I_{ik}^{F \times M}]$ describing a simple female × male interaction scenario. Nonetheless, female × male interactions may be different depending on whether the male is mating first or second. Thus, theoretically there are two distinct female × male interactions. If we also want to account for this possibility, we again have two alternatives how to denote it. Either we can capture it by $[-I_{ij}^{F \times M^1} + I_{ik}^{F \times M^2}]$, where $I^{F \times M^1}$ and $I^{F \times M^2}$ are two separate female × male interactions describing the female × male interaction with the first and second male to mate, respectively. Alternatively the matrix $I^{F \times M}$ describes the overall female × male interaction and $I^{F \times M^O}$ the deviance from this due to mating order. Then we get $-(I_{ij}^{F \times M} + I_{ij}^{F \times M^O}) + (I_{jk}^{F \times M} + I_{jk}^{F \times M^O})$.

One important comment needs to be appended. Previously the vectors F, M, M^1, and M^2 were introduced to describe different additive female and male effects. Thus, all the weighted row and column sums of the matrices describing the possible female × male interactions will necessarily equal zero because otherwise F_i M_i etcetera would not adequately describe these overall male and female effects. It needs to be the weighted sum to account for the possibility that different male and female types occur at different frequencies in the population. As a specific example: consider the term $\sum_i p_i I_{ij}^{F \times M}$, where p_i is the frequency of type i females. This term thus describes the summed deviation of male j's sperm competitiveness due to female × male interactions. If this term deviates from zero, then M_j would no longer describe the sperm competitiveness of male type j, but instead $M_j + \sum_i p_i I_{ij}^{F \times M}$.

MALE × MALE INTERACTIONS

A male × male interaction implies that male sperm competitiveness will depend on the male competitor. Note that this is distinctly different from the additive model, where male *fertilization success* will depend on the capability of the male competitor but not the male's competitiveness *per se*. The important distinction between male sperm competitiveness and fertilization success has already been pointed out by others (García-González 2008a,b; García-González and Evans 2011), but needs to be emphasized here again. Male fertilization success is in principle always nontransitive—paternity in sperm competition against a poor competitor will not be similar to success in competition against superior competitors (García-González 2008b). However, an

interaction effect on sperm competitiveness would indicate that the *relative* success will change depending on male competitor. We will return to this distinction later on. Let us first describe the male × male interaction by the following matrix $I^{M \times M}$. Each element $I_{jk}^{M \times M}$ in this matrix describes the deviation from the expected outcome between a male j mating first and male k mating second based on male j's and k's overall sperm competition defense and offense, respectively. In contrast to the female × male interaction, this matrix thus fully describes all potential outcomes including order effects, and there is no need for any additional description. Note, however, that in the absence of any order effects, a male j's competiveness against male k must be the opposite of male k's competiveness against male j, thus, $I_{jk}^{M \times M} = -I_{kj}^{M \times M}$ including the special case $I_{jj}^{M \times M} = 0$. Deviance from this would indicate order effects. Yet again, the weighted row and column sums of the male × male interaction matrix are constrained to equal zero.

FEMALE × MALE × MALE INTERACTIONS

Treatments of interactions in the outcome of sperm competition are usually limited to descriptions of female × male and male × male interactions (Bjork et al. 2007; Clark 2002; Clark et al. 1999, 2000; Miller and Pitnick 2002; Wilson et al. 1997). Yet in sperm competition there are at least three parties involved, and therefore a three-way interaction needs to be considered as well (see, e.g., Bretman et al. 2004; Tregenza and Wedell 2002). A female × male × male interaction would mean that the male × male interaction described above depends on the female they are mating with, or equivalently, that any female × male interaction can be different depending on the identity of the other male competitor. Let $I^{F \times M \times M}$ describe this interaction. This interaction will have to share the same characteristics as the male × male interaction $I^{M \times M}$ described above.

We can now finally describe the outcome of sperm competition with the expression

$$(\tilde{P}_2)_{ijk} = \mu + F_i - M_j^1 + M_k^2 - I_{ij}^{F \times M^1} + I_{ik}^{F \times M^2} + I_{jk}^{M \times M} + I_{ijk}^{F \times M \times M},$$

or alternatively

$$(\tilde{P}_2)_{ijk} = \mu + F_i - \left(M_j + M_j^O\right) + \left(M_k + M_k^O\right) - \left(I_{ij}^{F \times M} + I_{ij}^{F \times M^O}\right) + \left(I_{ik}^{F \times M} + I_{ik}^{F \times M^O}\right) + I_{jk}^{M \times M} + I_{ijk}^{F \times M \times M}.$$

In the following, I will use the former notation. Furthermore, the description here is in the perspective of the second male to mate, but we could equally well have chosen to describe this with emphasis on the first males' success as $(\tilde{P}_1)_{ijk} = -(\tilde{P}_2)_{ijk}$.

EXPECTED VALUES AND VARIANCES IN STANDARDIZED AND NONSTANDARDIZED ASSAYS

Now we can turn our attention to the important question how estimates of individual sperm competitiveness and their variances will be influenced by different experimental designs. For illustrative purposes, we limit our considerations to offensive sperm competition ability; however similar analyses can be performed for sperm competition defense with identical conclusions. If we start with assays using random male competitors and random female mating partner, the expected value of male k's offensive sperm competitiveness can be computed as

$$E\langle(\tilde{P}_2)_k\rangle = \sum_{i,j} p_i q_j \left(\mu + F_i - M_j^1 + M_k^2 - I_{ij}^{F \times M^1} + I_{ik}^{F \times M^2} + I_{jk}^{M \times M} + I_{ijk}^{F \times M \times M}\right) = \mu + M_k^2,$$

where p_i and q_j gives the frequency of type i females and type j males, respectively. Thus, this gives (as anticipated) a completely unbiased estimate of male k's sperm competitiveness in the offensive role (see also Fig. 1). The observed sampling variance of male k's offensive sperm competitiveness will equal

$$\text{Var}\langle(\tilde{P}_2)_k\rangle = \sum_{i,j} p_i q_j \left((\tilde{P}_2)_{ijk} - \left(\mu + M_k^2\right)\right)^2$$

$$= \text{Var}\langle M^1\rangle + \text{Var}\langle F\rangle + \text{Var}\langle I^{F \times M^1}\rangle + \text{Var}\langle I_{*k}^{F \times M^2}\rangle + \text{Var}\langle I_{*k}^{M \times M}\rangle + \text{Var}\langle I_{**k}^{F \times M \times M}\rangle + \text{Cov}\langle\ldots\rangle$$

The $\text{Cov}\langle\ldots\rangle$-term refers to all possible covariances between male and female (interaction) effects, which we can safely ignore here and henceforward, because these should be negligible in an experimental arrangement. (However, they may be important in a natural setting where females and males mate non-randomly.) Furthermore, we have ignored any residual variance attributable to any measurement errors but these should be independent of any of the causal sources of variation.

In a mating design with a standardized rival male ($sm = $ *s*tandard *m*ale) the expected value of male k's sperm competitiveness will be given by

$$E\langle(\tilde{P}_2)_k|sm\rangle = \mu + M_{sm}^1 + M_k^2 + I_{sm\,k}^{M \times M},$$

which must not necessarily provide an unbiased estimate of M_k^2. Thus, dependent on the magnitude of the male-by-male interaction, this can represent a strongly biased estimate of true male sperm competitiveness (see Fig. 1B). Yet, in the absence of any nontransitive effects the expected value will equal

$$E\langle(\tilde{P}_2)_k|sm\rangle = \underbrace{\mu + M_{sm}^1}_{=\mu'} + M_k^2 = \mu' + M_k^2,$$
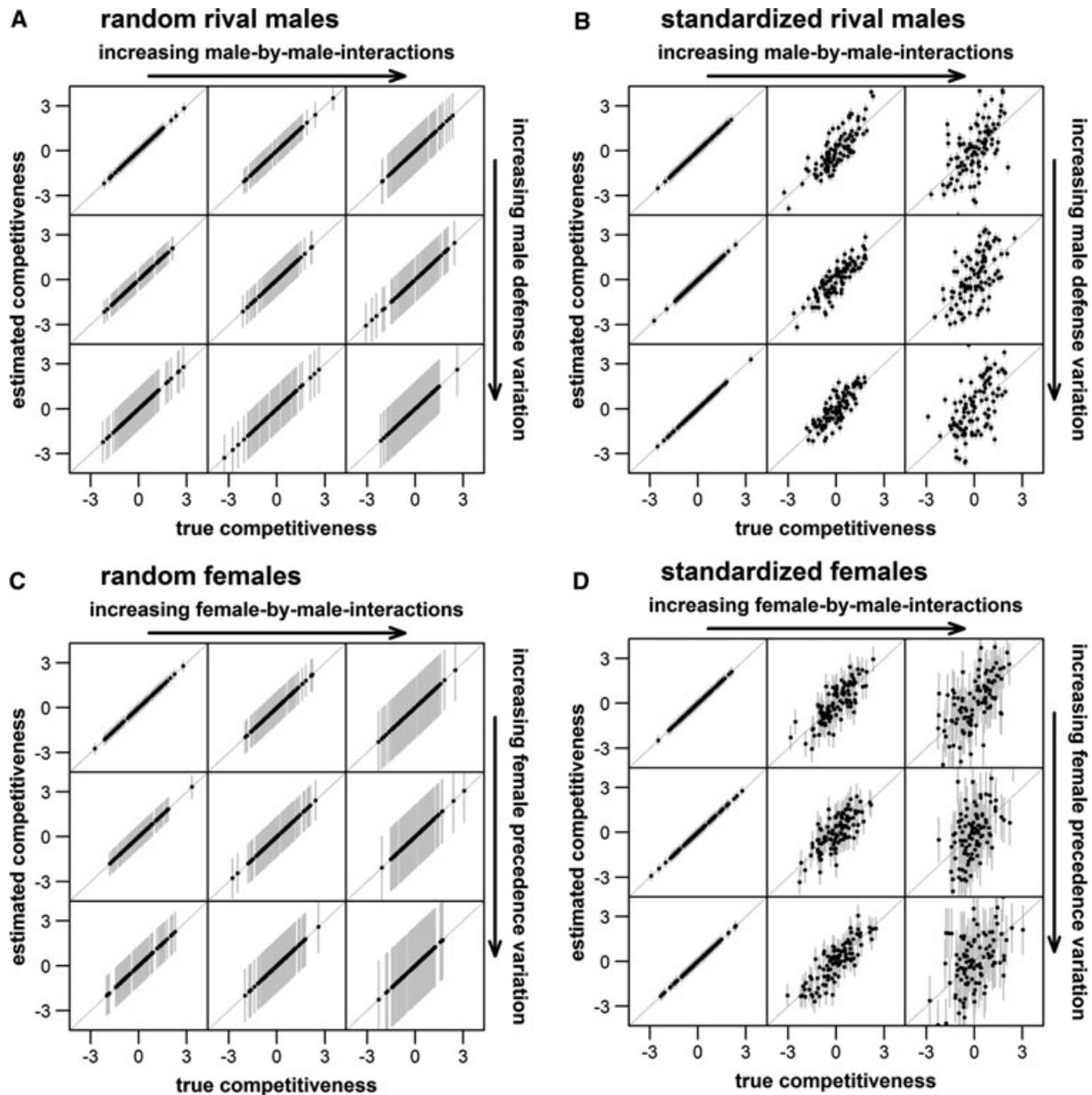
Figure 1. Illustration of expected estimates of offensive sperm competitiveness (points) and their variances (SD: grey bars) in relation to true offensive competitiveness in assays using (A) random rival males, (B) standardized rival males, (C) random females, and (D) standardized females. The plots show simulated values for 100 males, respectively. In all simulations an average overall fair raffle was assumed ($\mu = 0$), and the variation in male offensive competitiveness (Var$\langle M^2 \rangle$) equaled 1. In (A) and (B), the remaining parameter values were as follows Var$\langle F \rangle = 0.1$; Var$\langle M^1 \rangle = [0, 0.5, 2]$ (from top to bottom); Var$\langle I^{M \times M} \rangle = [0, 0.5, 2]$ (from left to right); Var$\langle F \times M \rangle = 0$; Var$\langle F \times M \times M \rangle = 0$. In (c) and (d), $\mu = 0$ (fair raffle); Var$\langle M^1 \rangle = 0.1$; Var$\langle F \rangle = [0, 0.5, 2]$ (from top to bottom); Var$\langle I^{F \times M} \rangle = [0, 0.5, 2]$ (from left to right); Var$\langle M \times M \rangle = 0$; Var$\langle F \times M \times M \rangle = 0$. Note that using random males and females will introduce variation in estimates but no bias, whereas standardized matings will reduce estimate variation but may result in strongly biased estimates when female–male or male–male interactions are present. Please also note that the effect of the female × male interactions on estimate variation is larger than the effect of the male × male interactions because in these particular examples the female × male interactions are not specific to male ordering (i.e., $I^{F \times M^1} = I^{F \times M^2}$) and thus will affect both offensive and defensive males.

and hence be unbiased, because the term μ' will be equal for all males and we can generalize trials using a standard male competitor to the population of all possible competitors. Then such a mating design will give more precise estimates of males sperm competitiveness (cf. Fig. 1A,B), because the variance will be devoid of any components attributable to the first male: $\text{Var}\langle(\tilde{P}_2)_k|sm\rangle = \text{Var}\langle F \rangle + \text{Var}\langle I_{*sm}^{F \times M^1}\rangle + \text{Var}\langle I_{*k}^{F \times M^2}\rangle + \text{Var}\langle I_{*sm\,k}^{F \times M \times M}\rangle + \text{Cov}\langle \ldots \rangle$ (In case of no male-by-male
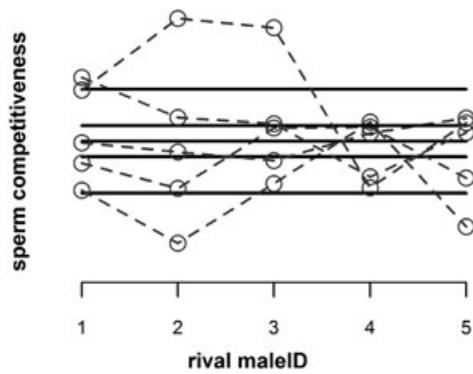
Figure 2. A schematic illustration of the impact of male–male interactions on the variance in sperm competitiveness. The dark solid lines represent the genuine sperm competitiveness of five male types across all matings, and the grey points represents the sperm competitiveness of these types in all different combinations. One can clearly see that if we limit our attention to specific rival males (as in a design with standard competitors) male–male interactions (deviations from the solid lines) will generally increase the variance in sperm competitiveness. However, this must not always be so. With male four as rival male, there is a negative covariance between male sperm competitiveness and the male–male interaction (strong competiors do relatively poor and poor do relatively good). This will reduce the variance in competitiveness. However, if there is strong positive covariance between male sperm competitiveness and the interaction effect (as with male two as rival), this will increase the variance even more. (These are simulated data where the interaction variance is 40% of the total variance.)

interactions, the interaction variation—$\mathrm{Var}\langle I_{*k}^{M \times M} \rangle$—will of course also be zero.).

The choice of experimental design will not only have an effect on the point estimates of sperm competitiveness from individual males/male types and their variances. The estimates of variation *across* males, such as genetic variation, will also be affected. The mean second male sperm competitiveness across all males against a standardized competitor will be given by $\bar{E}\langle \tilde{P}_2 | sm. \rangle = \mu + M_{sm}^1$. Thus, the estimated variance across all males can be computed as

$$\mathrm{Var}\langle \tilde{P}_2 | sm \rangle = \sum_i p_i (E\langle (\tilde{P}_2)_i | sm \rangle - \bar{E}\langle \tilde{P}_2 | sm \rangle)^2$$
$$= \mathrm{Var}\langle M^2 \rangle + \mathrm{Var}\langle I_{sm*}^{M \times M} \rangle + 2\mathrm{Cov}\langle M^2, I_{sm*}^{M \times M} \rangle.$$

Hence in case of noteworthy male × male interactions, this is thus generally an overestimate of the true variance across males given by $\mathrm{Var}\langle M^2 \rangle$, because it will be inflated by the interaction variance (Fig. 2, see also Fig. 1B). However it is not always an overestimate: it can also be an underestimate if the covariance in the expression above is strongly negative (see Fig. 2 for details).

We have also seen that whenever we have female variation in precedence, this term (i.e., $\mathrm{Var}\langle F \rangle$) will inflate the variance of the

estimator, $E\langle(\tilde{P}_2)_k\rangle$. It can therefore be tempting to perform sperm competition assays in standardized female environments, by for instance using females from identical isogenic lines for all mating experiments. This approach, however, has the same potential drawback as the mating design with standardized rival males. Consider a mating design using standardized females ($sf = s$tandard $f$emale). The expected value of male k's sperm competitiveness will then be given by $E((\tilde{P}_2)_k | sf) = \mu + F_{sf} + M_k^2 + I_{sf\,k}^{F \times M^2}$. This estimate will only represent true sperm competitiveness as long as $I_{sf\,k}^{F \times M^2} = 0$ for all k. Any existing female × second-male interaction will bias these estimates (see Fig. 1c,d). (Corresponding assays of male defense sperm competition abilities will be biased by the female × first-male interaction, and using both standard rivals and females will introduce bias from any female × male × male interaction.) Furthermore, variance estimates will generally be overestimates because, as shown above, using standardized female environments will include the female × male interaction variation $\mathrm{Var}\langle I_{sf*}^{F \times M^2} \rangle$ (see also Fig. 1D).

DIFFICULTIES WITH ADDITIVE VERSUS INTERACTION EFFECTS

The presented model can be used to identify separate additive and interaction effects regarding both sperm male offense and defense competitiveness and female tendency to bias paternity towards males based on mating order. The crucial step in this procedure has been to transform paternity values to a logit scale. This is the natural scale when the outcome (here paternity) will be a probability distribution depending on the relative values of predictors, as is the case in a sperm competition raffle. As we have seen, both male sperm competitiveness and genuine female effects can be seen as additive effects on this scale. Nevertheless, most sperm competition studies have been analyzed as male fertilization success on the ordinary P_i-scale. Yet unfortunately, independent male and female sperm competition effects will not operate additively on this scale. Thus, the fertilization success of a male in relation to other males will be inherently nontransitive to other situations if rival males' sperm competitiveness or the sperm precedence pattern of females varies. In a statistical analysis, male and female additive effects on sperm competitiveness will instead appear as male × male and female × male interactions. This is illustrated in Figure 3. In competition with a male with poor sperm competition defense abilities, both good and intermediate offense males will achieve similarly high fertilization success, much higher than males with poor sperm competition offense. In contrast, in competition with a male with very good sperm competition defense abilities, intermediate and poor males will be similar and have much lower fertilization success than males with superior offense abilities (Fig. 3A). Although in this example the difference in male sperm competitiveness remained
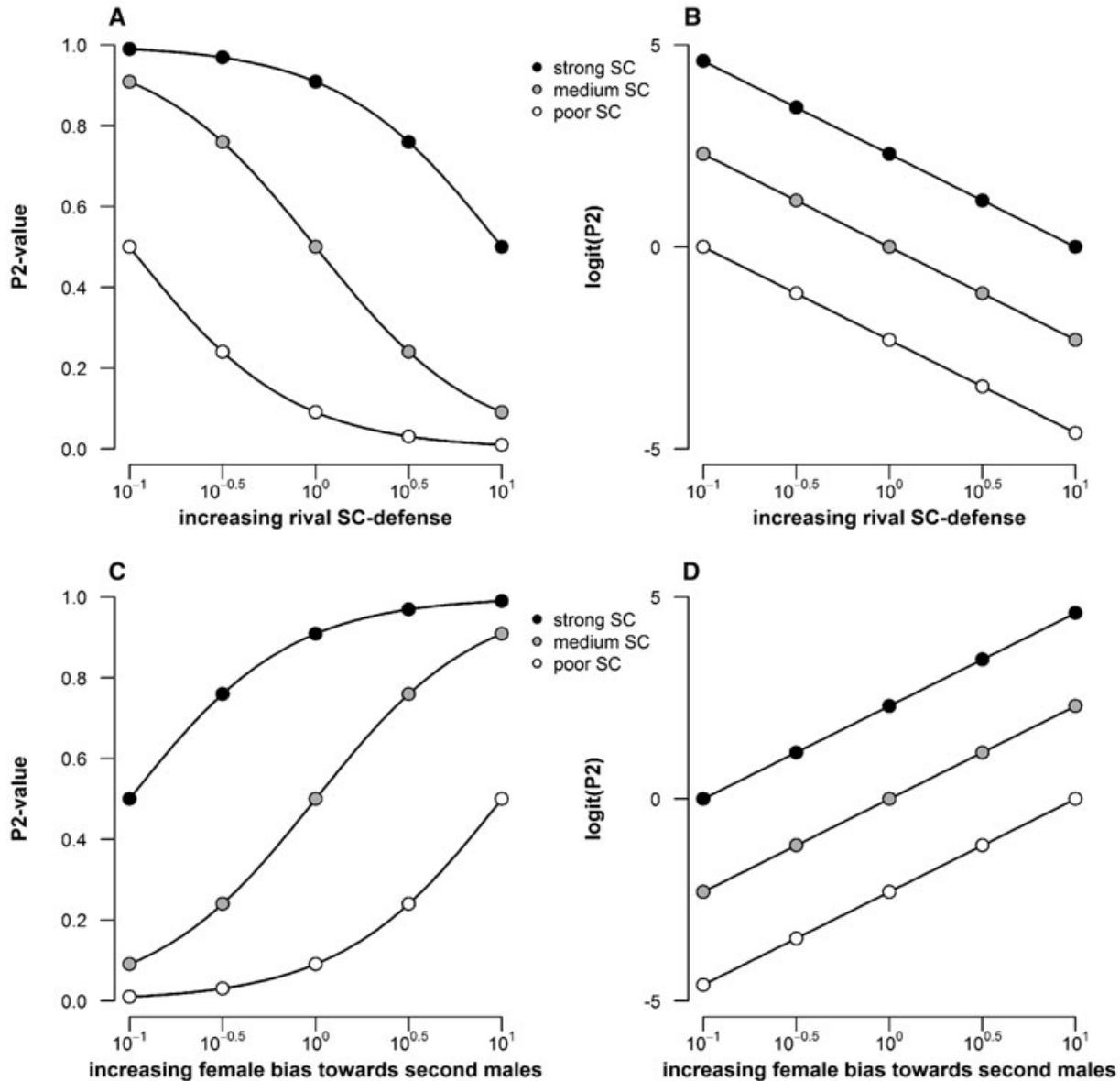
Figure 3. A comparison of P_2-values with the respective values after logit transformation $\tilde{P}_2$ (= logit[P_2]). Shown are the expected outcome with a fundamental raffle based assumption: $P_2 = \frac{rs_2}{s_1 + rs_2}$. Poor, medium, and strong sperm competitiveness (SC) refers to the offensive capability (i.e., s_2) of three distinct male types and equals 0.1, 1, and 10, respectively. In (A) and (B), sperm precedence is constant and equals $r = 1$ (fair raffle). In (C) and (D), r varies between females but male defense is constant and equals $s_1 = 1$. Note that although the differences in male sperm offense remains constant, this is only visible using the logit scale. Also note that there will be a linear relationship between logit-tranformed P_2-values and log-transformed male competitiveness, and log-tranformed values for female r-values (female-driven bias toward second males), respectively.

constant across situations, the difference in fertilization success did not. The true pattern of sperm competiveness will emerge only when analyzing the results on a logit scale (Fig 3B). Analyzing the data conventionally as paternity share (i.e., P_i values) will not change the ordering of males with regard to sperm competition abilities, but such analyses will be much more prone to finding nonexisting interaction effects, in this case a male × male interaction.

A similar effect will arise when females vary in the way second males are favored/disfavored in relation to first males. Consider a female with a very strong bias in paternity toward the first male to mate. In that case, poor and intermediately strong second males will have comparably low paternity. Conversely, in mating trials using females with a very strong second male bias, intermediate males will almost be as successful as strong sperm competitors and much better than poor ones (Fig. 3C). Again in

a situation where male and female effects act additively, analyses using fertilization success would reveal them as female × male interactions. Thus, to disentangle additive from interaction effects, it is essential to transform fertilization success, measured as P_i, to a different scale before proceeding. In this example, we see that a logit scale would be an excellent choice (Fig. 3D).

Another major benefit of working with logit-transformed P_i-values is that their variance is independent of the underlying sperm precedence pattern. In contrast, the variance in P_i-values will always be smaller when there is strong sperm precedence (see also Fig. 3). Thus, everything else being equal, estimates of genetic variance in paternity success (P_i) will always be larger when there is a fair rather than a loaded raffle. Nevertheless, one must also be aware that there are some potential complications associated with transformations to logit scale. First, the rational for using it, is that the mechanism of sperm competition is based on a raffle (i.e., the outcome can at least approximately be described by the formula $P_2 = rs_2/(rs_2 + s_1)$). In cases of strong violations from this assumption, for instance a "winner takes it all"-scenario, where a superior competitor will fertilize the majority of offspring irrespective of the rivals' competitiveness, the model described here might not be ideal. Yet this does not necessarily mean that such situations call for an analysis on conventional P_i-values, but simply that other transformations than the logit might provide a better fit. Second, if one is interested in quantitative genetic parameters, such as the heritability of sperm competitiveness, it is important to note that transformations will change the estimates and interpretations of these parameters (see, e.g., García-González et al. 2012). However, this is not a big concern here, as measuring sperm competitiveness (conventionally) as a P_i-value involves a transformation itself. Actually, the suggested logit transformation is an attempt to retransform these values to a more natural and meaningful scale. Nevertheless, when interpreting and comparing these estimates, one should be aware that following a logit transformation of paternity values, we end up with an estimate of sperm competitiveness that is on a log-scale (see also Fig. 3).

Discussion

In this contribution, the aim has been to resolve some unclear issues and highlight a few difficulties when analyzing data related to male sperm competitiveness. Three topics have been emphasized: (i) the occurrence of more female × male interactions than generally acknowledged; (ii) the potential drawback of using standardized males and females in experimental assays aiming at quantifying sperm competitiveness if interaction effects exist; and (iii) the need to transform P_i-values to the logit scale to disentangle additive from interaction effects.

It has been accepted for some time that male success in sperm competition may be nontransitive from one female to the next (Clark et al. 1999; Wilson et al. 1997). Yet although, for instance, a female type may bias paternity toward certain males when in their offensive role, it must not necessarily follow that these males are favored in their defense role as first males. Thus, female × male interactions have the potential to be specific to mating order. Most female × male interactions have been described for *Drosophila* (Bjork et al. 2007; Clark 2002; Clark et al. 1999; Miller and Pitnick 2002), a taxon in which sperm displacement is a major feature in sperm competition. Hence, naturally many studies have focused on second males' success, and a female × male interaction on second males' displacement success is consequently truly a female × second-male interaction ($I^{F \times M^2}$). Identifying and disentangling the exact nature of female × male interactions in sperm competition will thus provide a future challenge for empiricists. This is also true for female × male × male interactions. Indeed, female × male × male interactions have been demonstrated (Bretman et al. 2004; Tregenza and Wedell 2002), but they are generally not included in typical sperm competition studies. In addition, we must also be aware that this relatively complicated description of sperm competition can nonetheless only provide an accurate picture in the simple case with two male competitors. On an experimental level, it still remains to be investigated how offense, defense and interaction effects are affected by adding more competing males (see, e.g., Drnevich 2003; Lewis and Jutkiewicz 1998; Morrow et al. 2005; Zeh and Zeh 1994).

It was shown here that sperm competition assay procedures using standardized males (and females) can lead to very biased estimators of male sperm competitiveness and inflate estimates of genetic variability, in contrast to recent arguments (Fricke et al. 2010; García-González 2008b; García-González and Evans 2011). These effects can exclusively be attributed to the presence of either male × male or female × male interactions. In such cases, sperm competitiveness measured under a specific condition can unfortunately not be generalized to sperm competitiveness under other conditions. Moreover, estimates of genetic variation will be inflated, because the interaction variance will be pooled within the variation across genotypes. However, in the absence of such interactions, the best approach is indeed to use a standardized approach, which effectively reduces the residual variation (see also García-González and Evans 2011). This will increase the precision of estimates of individual sperm competitiveness. Furthermore, it will provide more accurate estimates of heritability, because it will be prevented that a part of the variation in sperm competitiveness (from the rival male) will be pooled within the environmental variance (García-González and Evans 2011). The problem is that researchers seldom have prior knowledge of the magnitude of the interaction effects, but they need to be estimated in the same experimental process. In such cases, introduced uncertainty of sperm competition measures due to variation in

rival males' sperm competitiveness and individual females' sperm precedence patterns might simply be something we have to live with. However, a resolution can be achieved by inserting male type (e.g., individual, genotype, etc.) as a random factor in a mixed model analysis, and thereby effectively control for this variance. Indeed, such an analysis would nevertheless be necessary if the aim is to estimate the variance parameters described here and simultaneously extract point estimates of the random effects (e.g., BLUPs) characterizing male offense and defense abilities. A further inclusion of interaction random effects would be an elegant possibility to estimate the interaction variances, and thereby for instance attain separate estimates of additive and interaction genetic variation. With sufficient data, one would be able to estimate all variance components, including main effect and interactions. A fully crossed design will frequently be impracticable, but at least partially crossed designs should be possible to achieve in some study systems. If one is interested in individual phenotypic variation, this will presumably still present some problems. It would require at least some repetitive matings of the same individuals. One would therefore need to additionally control for confounding factors such as age or mating history, unless special conditions apply (e.g., broadcast spawning). If the interest is focused on genetic variance parameters, this problem can be circumvented because individuals within the sib-groups will be the repeated measures. Thus, for each sib-group, there would be a possibility to stage sperm competition assays using several different "standard" males (and possibly females), for instance by using different isogenic lines (see, e.g., Clark et al. 1999; Dowling et al. 2010). In case of interaction effects, this would enable generalization of breeding values for sperm competitiveness to situations beyond just one standard competitor, and simultaneously allow for the separate estimates of additive and interaction genetic effects.

In any case, an adequate separation of additive and interaction effects essentially requires a transformation of sperm competition success, usually measured as the proportion offspring sired. The necessity to transform P_i-values to draw conclusions on sperm competition patterns is not a novel insight as such (Eggert et al. 2003; Neff and Wahl 2004; Parker et al. 1990). In fact, the logit transformation advocated here is merely an extension of the approach suggested by Eggert et al. (2003). Transformations can either be performed prior to the analysis. Alternatively, in situations where paternity can be assigned to a small offspring number only, a normal approximation of the inherently binomially distributed data may be inappropriate. In such cases generalized models will be indispensable; fortunately a logit transformation is often an inherent feature of data transformation with a binomial error structure (see, e.g., Venables and Ripley 2002). Thus, analyzing the data using a generalized linear model with binomial errors using a logit-link function (see Michalczyk et al. 2010 for an

example related to sperm competitivenes) spontaneously resolves one of the major obstacles with fertilization success data mentioned here and elsewhere (see García-González 2008b). Another benefit of working with generalized models is that they provide algorithms for dealing with extreme values (i.e., zero or complete paternity), which otherwise cannot be transformed meaningfully to a logit scale. A comprehensive analytical methodology to sperm competition data would thus require a generalized linear mixed model (possibly with crossed random factors), which might be a challenging but absolutely feasible approach if used thoughtfully (Bolker et al. 2009; Gelman 2005; Gelman and Hill 2007).

LITERATURE CITED

Andersson, M. 1994. Sexual selection. Princeton Univ. Press, Princeton, NJ.

Andersson, M., and L. W. Simmons. 2006. Sexual selection and mate choice. Trends Ecol. Evol. 21:296–302.

Birkhead, T. B., D. J. Hosken, and S. Pitnick, eds. 2009. Sperm biology: an evolutionary perspective. Academic Press, San Diego, CA.

Birkhead, T. R., and A. P. Møller, eds. 1998. Sperm competition and sexual selection. Academic Press, San Diego, CA.

Bjork, A., W. T. Starmer, D. M. Higginson, C. J. Rhodes, and S. Pitnick. 2007. Complex interactions with females and rival males limit the evolution of sperm offence and defence. Proc. R. Soc. B 274:1779–1788.

Bolker, B. M., M. E. Brooks, C. J. Clark, S. W. Geange, J. R. Poulsen, M. H. H. Stevens, and J. S. S. White. 2009. Generalized linear mixed models: a practical guide for ecology and evolution. Trends Ecol. Evol. 24:127–135.

Boorman, E., and G. A. Parker. 1976. Sperm (ejaculate) competition in *Drosophila melanogaster*, and the reproductive value of females to males in relation to female age and mating status. Ecol. Entomol. 1: 145–155.

Bretman, A., N. Wedell, and T. Tregenza. 2004. Molecular evidence of post-copulatory inbreeding avoidance in the field cricket *Gryllus bimaculatus*. Proc. R. Soc. B 271:159–164.

Clark, A. G. 2002. Sperm competition and the maintenance of polymorphism. Heredity 88:148–153.

Clark, A. G., and D. J. Begun. 1998. Female genotypes affect sperm displacement in *Drosophila*. Genetics 149:1487–1493.

Clark, A. G., D. J. Begun, and T. Prout. 1999. Female x male interactions in *Drosophila* sperm competition. Science 283:217–220.

Clark, A. G., E. T. Dermitzakis, and A. Civetta. 2000. Nontransitivity of sperm precedence in *Drosophila*. Evolution 54:1030–1035.

Dowling, D. K., M. Nystrand, and L. W. Simmons. 2010. Maternal effects, but no good or compatible genes for sperm competitiveness in Australian crickets. Evolution 64:1257–1266.

Drnevich, J. M. 2003. Number of mating males and mating interval affect last-male sperm precedence in *Tenebrio molitor* L. Anim. Behav. 66: 349–357.

Droge-Young, E. M., M. K. Manier, S. Lüpold, J. M. Belote, and S. Pitnick. 2012. Covariance among premating, post-copulatory and viability fitness components in *Drosophila melanogaster* and their influence on paternity measurement. J. Evol. Biol. 25:1555–1563.

Dziuk, P. J. 1996. Factors that influence the proportion of offspring sired by a male following heterospermic insemination. Anim. Reprod. Sci. 43:65–88.

Eggert, A. K., K. Reinhardt, and S. K. Sakaluk. 2003. Linear models for assessing mechanisms of sperm competition: the trouble with transformations. Evolution 57:173–176.

Engqvist, L., G. Dekomien, T. Lippmann, J. T. Epplen, and K. P. Sauer. 2007. Sperm transfer and paternity in the scorpionfly *Panorpa cognata*: large variance in traits favoured by postcopulatory episodes of sexual selection. Evol. Ecol. 21:801–816.

Fricke, C., O. Y. Martin, A. Bretman, L. F. Bussiere, and T. Chapman. 2010. Sperm competitive ability and indices of lifetime reproductive success. Evolution 64:2746–2757.

Gage, M. J. G., and E. H. Morrow. 2003. Experimental evidence for the evolution of numerous, tiny sperm via sperm competition. Curr. Biol. 13:754–757.

García-González, F. 2008a. Male genetic quality and the inequality between paternity success and fertilization success: consequences for studies of sperm competition and the evolution of polyandry. Evolution 62:1653–1665.

———. 2008b. The relative nature of fertilization success: implications for the study of post-copulatory sexual selection. BMC Evol. Biol. 8:140.

García-González, F., and J. P. Evans. 2011. Fertilization success and the estimation of genetic variance in sperm competitiveness. Evolution 65:746–756.

García-González, F., and L. W. Simmons. 2005. Sperm viability matters in insect sperm competition. Curr. Biol. 15:271–275.

García-González, F., L. W. Simmons, J. L. Tomkins, J. S. Kotiaho, and J. P. Evans. 2012. Comparing evolvabilities: common errors surrounding the calculation and use of coefficients of additive genetic variation. Evolution 66:2341–2349.

Gelman, A. 2005. Analysis of variance—why it is more important than ever. Ann. Stat. 33:1–53.

Gelman, A., and J. Hill. 2007. Data analysis using regression and multilevel/hierarchical models. Cambridge Univ. Press, Cambridge, U. K.

Keller, L., and H. K. Reeve. 1995. Why do females mate with multiple males? The sexually selected sperm hypothesis. Adv. Study Behav. 24:291–315.

Lewis, S. M., and S. N. Austad. 1990. Sources of intraspecific variation in sperm precedence in red flour beetles. Am. Nat. 135:351–359.

Lewis, S. M., and E. Jutkiewicz. 1998. Sperm precedence and sperm storage in multiply mated red flour beetles. Behav. Ecol. Sociobiol. 43:365–369.

Michalczyk, Ł., O. Y. Martin, A. L. Millard, B. C. Emerson, and M. J. G. Gage. 2010. Inbreeding depresses sperm competitiveness, but not fertilization or mating success in male Tribolium castaneum. Proc. R. Soc. B 277:3483–3491.

Miller, G. T., and S. Pitnick. 2002. Sperm-female coevolution in *Drosophila*. Science 298:1230–1233.

Morrow, E. H., A. D. Stewart, and W. R. Rice. 2005. Patterns of sperm precedence are not affected by female mating history in *Drosophila melanogaster*. Evolution 59:2608–2615.

Neff, B. D., and L. M. Wahl. 2004. Mechanisms of sperm competition: testing the fair raffle. Evolution 58:1846–1851.

Parker, G. A. 1970. Sperm competition and its evolutionary consequences in the insects. Biol. Rev. 45:525–567.

———. 1990. Sperm competition games: raffles and roles. Proc. R. Soc. B 242:120–126.

Parker, G. A., L. W. Simmons, and H. Kirk. 1990. Analysing sperm competition data: simple models for predicting mechanisms. Behav. Ecol. Sociobiol. 27:55–65.

Prout, T., and J. Bundgaard. 1977. Population genetics of sperm displacement. Genetics 85:95–124.

Sakaluk, S. K., and A.-K. Eggert. 1996. Female control of sperm transfer and intraspecific variation in sperm precedence: antecedents to the evolution of a courtship food gift. Evolution 50:694–703.

Sherman, C. D. H., E. Wapstra, and M. Olsson. 2009. Consistent male-male paternity differences across female genotypes. Biol. Lett. 5:232–234.

Simmons, L. W. 2001. Sperm competition and its evolutionary consequences in the insects. Princeton Univ. Press, Princeton, NJ.

Simmons, L. W., and G. A. Parker. 1992. Individual variation in sperm competition success of yellow dung flies, *Scatophaga stercoraria*. Evolution 46:366–375.

Simmons, L. W., and M. T. Siva-Jothy. 1998. Sperm competition in insects: mechanisms and the potential for selection. Pp. 341–434 *in* T. R. Birkhead, and A. P. Møller, eds. Sperm competition and sexual selection. Academic Press, San Diego, CA.

Snook, R. R. 2005. Sperm in competition: not playing by the numbers. Trends Ecol. Evol. 20:46–53.

Tregenza, T., and N. Wedell. 2002. Polyandrous females avoid costs of inbreeding. Nature 415:71–73.

Venables, W. N., and B. D. Ripley. 2002. Modern applied statistics with S. Springer, New York.

Wilson, N., S. C. Tubman, P. E. Eady, and G. W. Robertson. 1997. Female genotype affects male success in sperm competition. Proc. R. Soc. B 264:1491–1495.

Zeh, J. A., and D. W. Zeh. 1994. Last-male sperm precedence breaks down when females mate with three males. Proc. R. Soc. B 257:287–292.

Associate Editor: E. Morrow