ORIGINAL ARTICLE

Mating failure shapes the patterns of sperm precedence in an insect

Vicki L. Balfour[1] (iD) · Daniella Black[1] · David M. Shuker[1]

Abstract

Estimates of last male sperm precedence (P_2) are often used to infer mechanisms of sperm competition, a form of post-copulatory sexual selection. However, high levels of mating failure (i.e. copulations resulting in no offspring) in a population can lead to misinterpretations of sperm competition mechanisms. Through simulations, García-González (2004) illustrated how mating failure could cause bimodal distributions of paternity with peaks at $P_2 = 0$ and 1, under a random sperm mixing mechanism. Here, we demonstrate this effect empirically with the seed bug *Lygaeus simulans*, a species known to exhibit high levels of mating failure (40–60%), using a morphological marker to estimate paternity. Contrary to previous findings in a sister species, we did not find strong evidence for last male sperm precedence. There was a tendency towards last male precedence ($P_2 = 0.58$) but within the expected range for random sperm mixing. Instead, P_2 was highly variable, with a bimodal distribution, as predicted by García-González (2004). After taking mating failure into account, the strongest driver of paternity outcome was copulation duration. Furthermore, we found evidence that mating failure could partly be a female-associated trait. Some doubly-mated females were more likely to produce no offspring or produce offspring from two different sires than expected by chance. Therefore, some females are more prone to experience mating failure than others, a result that mirrors an earlier result in male *L. simulans*. Our results confirm that mating failure needs to be considered when interrogating mechanisms of post-copulatory sexual selection.

Significance statement

Mating failure arises when animals fail to produce offspring across their lifetime. This may be due to a failure to find a mate or a failure to produce offspring after one or more apparently successful matings. Sperm competition is when ejaculates of rival males compete to fertilize a female's eggs. Estimates of second male paternity (P_2) are often used to infer mechanisms of sperm competition (i.e. which male "wins" and how). However, García-González (2004) suggested that high levels of mating failure can skew paternity (i.e. give spuriously high/low levels of P_2) and lead to misinterpretations of these mechanisms. We carried out sperm competition experiments on *Lygaeus simulans* seed bugs using a morphological marker to estimate paternity. We show empirically that mating failure does skew patterns of paternity, causing a bimodal distribution of P_2. Therefore, by disrupting patterns of sperm competition, mating failure influences both the action of post-copulatory sexual selection and also our understanding of the mechanisms of sperm competition.

Communicated by D. J Hosken

✉ Vicki L. Balfour
 vb28@st-andrews.ac.uk

[1]   School of Biology, Harold Mitchell Building, University of St Andrews, St Andrews KY16 9TH, UK

Introduction

Post-copulatory sexual selection, which acts on traits expressed during or after copulation (Birkhead and Pizzarri 2002; Dougherty et al. 2016), is not only important for shaping some of the extreme elaborate traits that we see across the animal kingdom but also plays a role in the diversification of populations and has the potential to lead to speciation (Arnqvist et al. 2000; Birkhead and Pizzarri 2002). In insects, for example, through means of an evolutionary arms race

Springer

driven by sexual conflict (Chapman et al. 2003), post-copulatory sexual selection could act upon morphological, physiological and behavioural reproductive traits, which could, in turn, lead to reproductive isolation between populations and, hence, speciation (Arnqvist et al. 2000). Post-copulatory sexual selection is primarily thought to be driven by cryptic female choice (the ability of females to bias paternity towards males with desirable traits: Eberhard 1996; Arnqvist 2014) and sperm competition (competition between ejaculates of rival males to fertilize a female's eggs: Parker 1970; Simmons 2001).

The role of sperm competition in post-copulatory sexual selection has long been a topic of interest for evolutionary biologists (Parker 1970; Smith 1984; Birkhead and Møller 1998; Simmons 2001) and understanding the mechanisms of sperm competition are key to understanding how and why these traits are selected for and the implications this has for the evolution of the species (Simmons 2001). However, misinterpretations of the mechanisms behind sperm competition can lead to misunderstandings of the role of sperm competition in post-copulatory sexual selection (García-González 2004). Therefore, in this paper, we shall highlight and evaluate a potential source of misinterpretation of second male paternity estimates (P_2: the proportion of offspring sired by the second or last male to mate): mating failure. We will emphasize the importance of taking mating failure into account when designing and interpreting experiments that use paternity estimates to evaluate mechanisms of sperm competition.

We often use estimates of second male paternity (P_2) to make inferences about sperm usage and sperm competition mechanisms (Boorman and Parker 1976; Lewis and Austad 1990; Cook et al. 1997; Simmons 2001). For example, extremely high or low values of P_2 are indictive of mechanisms to attain high sperm precedence, such as sperm removal, sperm displacement or sperm stratification (Simmons 2001). For instance, male *Calopteryx maculata* damselflies remove rival males' sperm from the female's bursa copulatrix and spermathecal tubes using their penis, which is covered in backwards facing spines. Only then will they inseminate the female and hence gain the majority of the paternity (Waage 1979). Yellow dung flies (*Scathophaga stercoraria*) also exhibit high levels of last-male sperm precedence as a result of indirect sperm displacement as sperm are pumped in, potentially aided by female muscular contractions (Simmons et al. 1999; Hosken and Ward 2000). On the other hand, values of and around $P_2 = 0.5$ are usually interpreted as being caused by random sperm mixing (Simmons 2001). In the red flour beetle *Tribolium castaneum*, although initial paternity estimates from clutches of eggs laid 48 h after mating revealed high levels of last male precedence, P_2 values from later clutches were much lower and indictive of random sperm mixing (Lewis and Jutkiewicz 1998). Therefore, the authors suggested that there must be initial sperm stratification (the last

sperm to enter are the first to leave and fertilize the eggs; Simmons 2001) followed by random sperm mixing over time (Lewis and Jutkiewicz 1998). However, we need to be careful of the way we interpret P_2, not only due to the potential errors in the techniques used to gather the data, such as with the irradiated male technique (Simmons 2001; Simmons et al. 2006), but also due to the effect that phenomena, such as mating failure, can have (García-González 2004).

Mating failure is a phenomenon through which individuals fail to produce offspring, either due to premature death before mating can occur or due to insemination or fertilization failure during and following copulation (the latter two have been recently referred to as cryptic mating failure: Greenway et al. 2015; see also Rhainds 2010). Throughout this paper, when we talk about 'mating failure', we are referring to 'cryptic mating failure' only, i.e. the lack of offspring production following copulation. García-González (2004) termed this non-sperm representation, to describe how one or more expected ejaculates (from behavioural observations for example) were not actually present, due to a presumed failure to successfully transfer a non-negligible quantity of sperm. Here we prefer to use (cryptic) mating failure, partly as the prior term has not caught on, but our conceptualisation matches that of García-González (2004). Importantly, García-González argued that non-sperm representation can mislead us in terms of mechanisms of sperm competition. Using a series of simulation models, he showed that the occurrence of mating failure in a population can highly skew the distribution of P_2 (Fig. 1a, b). As the proportion of copulations that fail to result in offspring increase, so P_2 under a random sperm mixing mechanism changes from being normally distributed around a mean P_2 of 0.5, to becoming bimodally distributed with peaks at 0 and 1 (i.e. apparent complete first and last sperm male precedence, respectively). Interpretations of such an outcome in terms of mechanisms would likely be very different to an outcome of $P_2 = 0.5$ with a normal distribution around the mean. For example, a bimodal distribution of P_2 values may be indictive of the use of mating plugs, whereby low values of P_2 would imply that the mating plug remained intact, and high values represent cases when the mating plug was breached (Simmons and Siva-Jothy 1998; García-González 2004). Though some studies have demonstrated results that look like García-González's (2004) simulation results, none, to the best of our knowledge, have specifically tried to demonstrate the effect that mating failure has on P_2 estimates. Using *Lygaeus simulans* as a study species, here we test García-González's (2004) simulation models empirically.

Lygaeus simulans are a species of promiscuous bug, which have previously been recorded to have high levels of mating failure (40–60%; Tadler et al. 1999; Micholitsch et al.
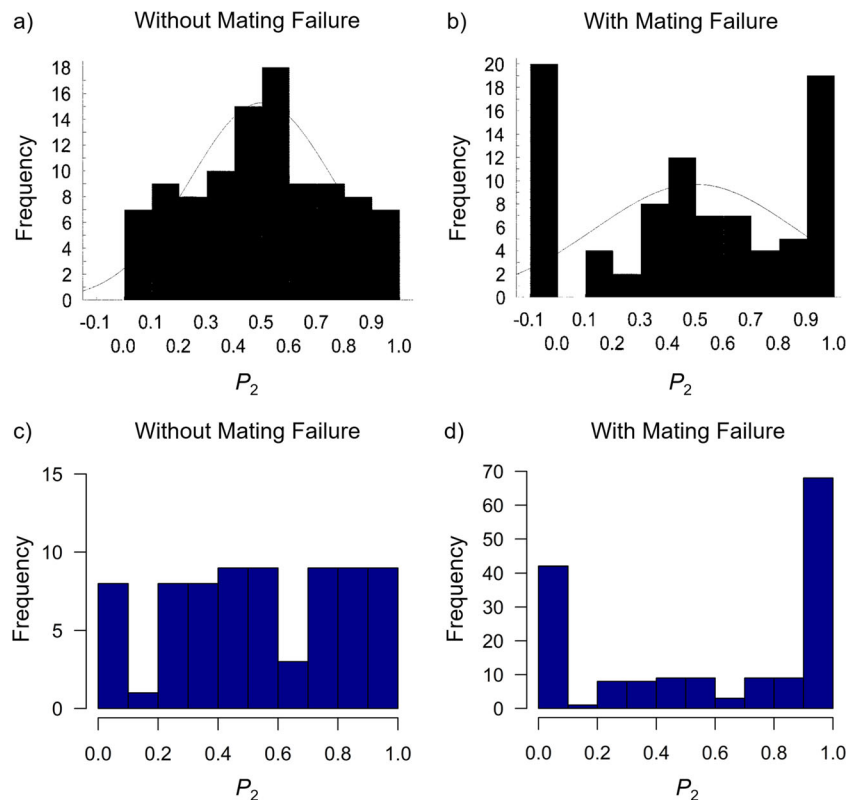
Fig. 1 Distribution of P_2 values due to mating failure during García-González's simulations (**a**) and (**b**) and during this experiment for treatments PPW and PWP combined (**c**) and (**d**). **a** Mating failure = 0%, this panel shows the simulated distribution of P_2 values assuming no mating failure ($N = 300$ simulations of a sample size of 100). The line represents a normal curve derived from the distribution, with a standard deviation (SD) beyond which the probability of obtaining a distribution with a greater SD is $P = 0.05$. **b** Mating failure = 30%, this panel shows the simulated distribution of P_2 values assuming 30% mating failure ($N =$ 300 simulations of a sample size of 100). In this case, the line again represents a normal curve derived from the distribution, with a Standard Deviation (SD) beyond which the probability of obtaining a distribution with a greater SD is $P = 0.95$ (adapted from Fig. 5, García-González 2004). **c** Only females that produced both wild-type and pale nymphs and hence did not experience any mating failure (mating failure = 0%; $N = 73$). **d** All females that produced nymphs including those that experienced mating failure with one of the males she copulated with (mating failure = 47.1%; $N = 166$)

2000; Dougherty and Shuker 2014; Greenway and Shuker 2015; Greenway et al. 2017). These bugs make an excellent study species for investigating the effects of mating failure on paternity outcome, not only because they are easy to rear in the lab, mate multiply and express these high levels of mating failure (Burdfield-Steel and Shuker 2014; Greenway and Shuker 2015; Greenway et al. 2017) but also because we can now use a genetically-based colour polymorphism as a morphological marker to estimate paternity (Balfour et al. 2018). Importantly, mating failure in this species has been found to be almost exclusively due to insemination failure: dissections showed that 92.6% of females that exhibited mating failure had received no sperm during copulation (Greenway et al. 2017).

We aimed to answer the following questions. First, do *L. simulans* exhibit last male sperm precedence, as has previously been found in the sister species *L. equestris* (Sillén-Tullberg 1981)? Second, do high levels of mating failure cause highly bimodal patterns of paternity (i.e. many pairs with either $P_2 = 1$ or $P_2 = 0$) as predicted by

García-González's (2004) simulations? Third, is mating failure a female-associated trait in *L. simulans*? Previous work has shown that mating failure is a male-associated trait in this species (i.e. some males were repeatedly more likely to fail to inseminate females than other males: Greenway and Shuker 2015; Greenway et al. 2017). By double-mating our focal females in this experiment, we can ask whether the number of females that failed from two matings to produce offspring was higher than that expected by chance, given the level of mating failure in the population. Finally, in addition to investigating the effect of mating failure on patterns of paternity, we also explored other factors that might affect paternity, including copulation duration, latency to mating, male and female body size and the number of offspring produced by focal females. By doing so, this experiment allowed us to further our understanding of post-copulatory sexual selection in this species, building on from previous work (Dougherty and Shuker 2014; Dougherty et al. 2015; Dougherty and Shuker 2016; Greenway and Shuker 2015; Greenway et al. 2017).

Methods and materials

Husbandry

Lygaeus simulans were collected in Tuscany, Italy, in 2008 and 2009, and transferred to the Shuker Lab at the University of St Andrews. In the laboratory, the bugs were kept in population cages ($30 \times 15 \times 15$ cm plastic boxes) and provided with an ad libitum supply of sunflower seeds, cotton wool for habitat and two cotton-plugged tubes of distilled water (25 ml), which were changed once a week (all water provided to the bugs mentioned below was likewise distilled water). The bugs were kept in an incubator at 29 °C on a 22:2 h light:dark cycle to prevent the onset of reproductive diapause. A minimum of two replicates of the population cages were kept at any one time. New population cages were created by transferring around 50 bugs from across each instar from two separate population cages into a new cage. This is to enhance gene flow and limit inbreeding depression.

In 2013, a population cage of pale colour morph *L. simulans* was created using pale mutants, which appeared in the wild-type population cages from 2012 onwards, in addition to F3 generation pale nymphs from an experiment carried out by L.R. Dougherty (unpublished data). Since then, a minimum of two population cages of pale *L. simulans* have been maintained in the lab (Balfour et al. 2018). The locus for colour morph in *L. simulans* is inherited in a Mendelian fashion, with a dominant wild-type allele resulting in the typical red and black aposematic colouration and a recessive pale mutant allele resulting in paler green-brown coloured bugs (Balfour et al. 2018). Here, we used this colour morph as a phenotypic marker to carry out the following experiment and assign paternity to the offspring sired.

To obtain virgin bugs for the following experiment, we made up nymph boxes ($20 \times 10 \times 8$ cm plastic boxes) by collecting late instar nymphs from population cages and transferring them, using an aspirator, to the boxes. We supplied the nymphs with a cotton-plugged water tube (25 ml), an ad libitum supply of sunflower seeds and a piece of cotton wool for habitat. Pale and wild-type individuals were housed separately.

We checked the nymph boxes every 2–3 days for newly eclosed adults and these were separated by sex into same sex tubs ($108 \times 82 \times 55$ mm plastic deli tubs) with a maximum of 10 individuals per tub and provided with an ad libitum supply of sunflower seeds, a cotton-plugged water tube (7 ml) and a piece of cotton wool for habitat. This was to ensure that all bugs used in the following experiment were virgins as *L. simulans* males become sexually mature after a few days post-eclosion and females become sexually mature around 7 days post-eclosion.

Experimental procedure

We paired focal pale females with a male for 6 h on day 1 and then with a different male for 6 h on day 2. We allowed females to lay eggs for 7 days, and then nymphs were counted a further 7 days after this. Details of each of these stages are given below.

There were four treatments: PPW, PWP, PPP and PWW (with first letter denoting the female's phenotype, second letter denoting the phenotype of the male on day 1 and the third letter denoting the phenotype of the male on day 2; P represents pale and W represents wild-type; Fig. 2). Note, due to the dominance relationships of the pale mutant, all focal females were pale. Sample sizes were $N = 248, 246, 173$ and 161, respectively. Replicates for each treatment were assigned haphazardly. We tried to equalize approximately the number of replicates of each treatment across days as males and females eclosed, whilst also taking advantage of individuals of a given phenotype as they eclosed and matured.

On day 1, we randomly paired focal virgin females (8–14 days old) with a virgin male (8–14 days old) in a Petri dish (55 mm diameter) for 6 h, scoring for mating (yes/no) every 15 min. Bugs were said to be mating when they adopted the back-to-back copulatory position. We allowed pairs that stopped mating after less than three checks (< 30 min) to mate again. We separated pairs that stopped mating after having been observed in-copula for three consecutive checks or more (> 30 min). This is because the minimum duration for successful sperm transfer in this species is approximately 30 min (Gschwentner and Tadler 2000); therefore, bugs that were observed in copula for two checks or less were recorded as not having mated. Only pairs that mated for three consecutive checks or more were recorded as having mated. For the analyses, we rounded everything up so pairs that mated for three consecutive checks (30–45 min) were recorded as having a copulation duration of 45 min. Along with copulation duration, mating latency—the time taken to initiate copulation—was recorded. After 6 h, we separated any pairs still mating by gently brushing their genitalia with a paintbrush. We transferred males to individual labelled Eppendorf tubes and froze these at − 18 °C for future measurements. We placed females in individual tubs ($108 \times 82 \times 55$ mm) with 15–20 sunflowers seeds and a cotton-plugged water tube (7 ml), and we returned these to the incubator overnight.

On day 2, we once more paired the focal females with a new virgin male (9–15 days old) following the same procedure as outlined for day 1. At the end of the mating trial, we again froze males in Eppendorf tubes at − 18 °C and returned females to their individual tubs and placed them back in the incubator. No females laid eggs in the 18 h between being paired with males on days 1 and 2, so females were returned to their individual tubs rather than given a fresh tub. Females that did not mate on either one day or both days were kept.
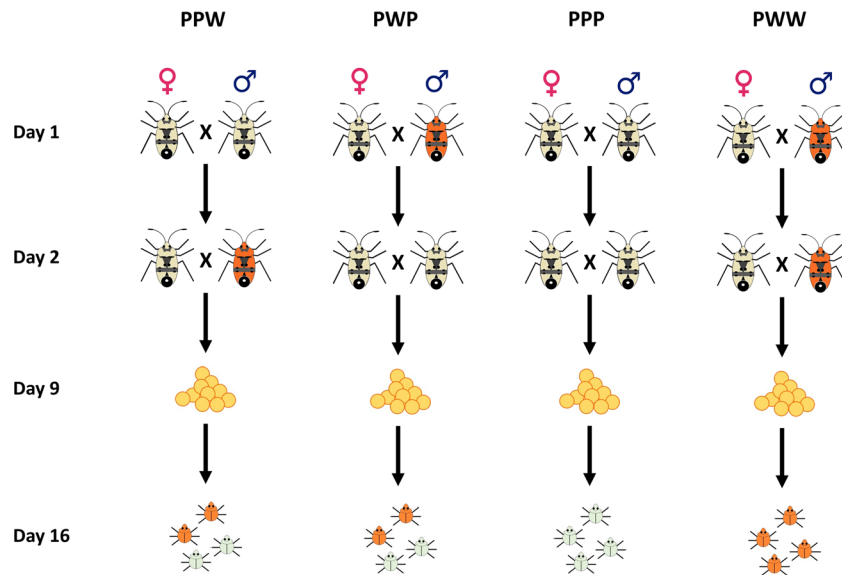
Fig. 2 Schematic of the experimental design, illustrating the four treatments (PPW, PWP, PPP and PWW; W: wild-type males/nymphs = dark/solid; P: pale males/females/nymphs = light/hatched). The colour morph of the virgin males that females were paired to on days 1 and 2 for each of the treatments is shown. Tubs were checked for the presence/absence of eggs on day 9, and then, on day 16, tubs were frozen for 24 h, and any nymphs present were scored for colour morph and counted. If there is equal paternity between sires, there should be a 1:1 ratio of wild-type:pale morph nymphs in treatments PPW and PWP. If a female only mates with pale males, then all offspring should be pale (treatment PPP). If a female only mates with wild-type males, then all offspring should be wild-type (treatment PWW)

This was to allow comparisons of mating failure rates and number of offspring produced between singly-mated and doubly-mated females with females that did not mate acting as a control group. Further treatment codes represent these females, with the letter Z in the treatment meaning that a female did not mate, i.e. PWZ means the female mated with a wild-type male on day 1 but did not mate with any male on day 2. Final sample sizes for each of these treatments are shown in Table 1.

On day 9, we removed females from their individual tubs and froze them in Eppendorf tubes at $-18\,^{\circ}\text{C}$ for later measurements. We also scored tubs for the presence/absence of eggs, discarding any tubs without eggs. We then returned tubs to the incubator for a further 7 days to allow nymphs to emerge. On day 16, all tubs were then frozen for a minimum of 24 h at $-18\,^{\circ}\text{C}$. We then scored tubs for the presence/absence of nymphs and counted any nymphs present according to colour morph.

Measurements

We (DB) measured the body length of all the bugs after thawing using a dissecting microscope fitted with an eyepiece micrometre. We measured the length from the tip of the snout to the tip of the wings, dorsal side up. We re-measured 74 bugs (22 females and 26 males from day 1 and 24 males from day 2), blind to the original measurements, to check measurement reliability. Our measurements were highly repeatable (intra-class correlation coefficient: $r = 0.954$; one-way ANOVA:

$F_{73,74} = 42.09$, $P < 0.001$; Lessells and Boag 1987). Pale males were significantly larger (mean = 10.84 ± 0.02 mm) than wild-type males (mean = 10.70 ± 0.02 mm; GLM: $F_{1,1559} = 28.5$, $P < 0.001$), being on average 0.14 mm longer. As expected, there was pronounced sexual dimorphism with females (mean = 11.60 ± 0.02 mm) being much larger than males ($F_{1,2339} = 1257$, $P < 0.001$).

Analysis

We removed 23 bugs from the data set because the female died before the end of the mating trial on day 2. We removed a further 22 data points due to a lack of males or male death prior to day 2 or due to bug escapes or missing data. Finally, we excluded 18 data points from the analysis because of the 'wrong' morph of nymphs appearing in the F1 or nymphs being present when the pairs were not observed to mate for > 30 min (the putative minimum time needed for sperm transfer). Reasons to explain these data anomalies include heterozygous males being present in small numbers in the wild-type population cages and eggs being accidentally transferred via forceps from one box to another (see also Balfour et al. 2018 for discussion, including evidence for the pale morph segregating at very low frequency in wild-type lab populations) or, indeed, sperm being transferred by males in copulations < 30 min in duration. Therefore, the final sample sizes were $N = 229$, 241, 160 and 161 for treatments PPW, PWP, PPP and PWW, respectively.

Table 1 Number of females in each treatment that succeeded or failed to produce offspring and the proportion of females that experienced mating failure (proportion failed). For treatment codes, first letter denotes female phenotype, second letter the male mated with on day 1, third letter the male mated with on day 2, P = pale, W = wild-type, Z = did not mate. Note—for treatments PPP, PWW, PWP and PPW, this is the proportion that failed to produce offspring after 2 copulations

Treatment	N	Number of females that failed or succeeded in having offspring		
		Failed	Succeeded	Proportion failed
PZZ	104	104	0	1.00
PZP	112	41	71	0.37
PPZ	45	24	21	0.53
PWZ	44	26	18	0.59
PZW	89	50	39	0.56
PPP	82	23	59	0.28
PWW	93	19	74	0.20
PPW	109	30	79	0.28
PWP	113	26	87	0.23
All treatments	791	343	448	0.43

A total of 25,277 nymphs were counted and scored for colour morph across all treatments. The paternity of the offspring from doubly-mated females was discerned by the colour morph of the nymphs as described above. P_2 was calculated as $\frac{r_2}{(r_1+r_2)}$, with r_2 representing the number of offspring sired by the second male to mate and r_1 the number of offspring sired by the first male to mate (García-González 2004). We here assumed that P_2 values of 1 and 0 were due to mating failure occurring during the first and second matings, respectively. We accept here that this may lead to some over-estimation of mating failure if these values occurred due to other processes (e.g. sperm were successfully transferred by a given male, but no eggs were fertilized by that male's sperm), but we believe that this will be generally representative for our data. As shown in the "Results", a mechanism that typically leads to the complete exclusion of one or other male's ejaculate from the fertilization set looks unlikely. Further comments on this can be found in the "Discussion".

Mating failure was scored according to the presence or absence of offspring. Singly- and doubly-mated females that produced no nymphs were recorded as having experienced mating failure. Doubly-mated females that produced only one colour morph of nymph, and hence $P_2 = 0$ or $P_2 = 1$, were recorded as having experienced mating failure with one of the males she was mated to (the one that sired no offspring). Doubly-mated females that produced nymphs of both colour morphs did not experience mating failure with either male they were mated to.

All data was analysed using R statistical software (R Core Team 2019). One-sample Z-tests were used to determine whether P_2 values differed from 0.5 and whether the number of doubly-mated females that experienced mating failure with one male differed from expected given the overall mating failure rate. A Chi-squared test was used to test whether the observed distribution of mating failure differed from the expected distribution if mating failure was random with respect to female phenotype. Pearson's correlation coefficient was used to investigate whether there was a correlation between copulation duration and mating latency. A pairwise t test was used to test whether there was a difference between copulation duration on day 1 and 2 for individual females. Generalized linear models (GLMs) with a quasibinomial distribution and logit link function (to account for overdispersion) were used to test whether there was (i) a difference in P_2 values between treatments PPW and PWP, (ii) whether there was an effect of the difference in body length of day 1 and day 2 males on paternity, (iii) whether there was an effect of copulation duration on paternity and also (iv) whether there was an interaction between mating duration on day 1 and mating duration on day 2. GLMs with a binomial distribution and logit link function were used to test (i) the relationship between copulation duration and the likelihood of insemination success, (ii) the effect of male phenotype, (iii) day, (iv) female mating status and (v) body length on the likelihood of copulation success, (vi) the relationship between number of times mated, (vii) male phenotype and (viii) sperm competition and the likelihood of mating failure. GLMs with a normal distribution were used to test (i) the relationship between the number of times mated and the number of nymphs produced, (ii) the effect of sperm competition and (iii) male phenotype on the number of offspring sired, (iv) the effect of male

phenotype, (v) day, (vi) female mating status and (vii) body length on copulation duration and mating latency and (viii) the relationship between body length and the number of nymphs produced.

Results

Out of the 791 pairs considered in this experiment, 50.2% of females mated twice, 36.7% of females mated only once and 13.1% of females did not mate at all. When considering only pairs that mated, there were high levels of mating failure, with 47.1% of matings failing to produce offspring (see Table 1 for individual mating failure rates for each treatment). For the females that mated twice, the number of females that produced each nymph morph is illustrated in Table 2. Last male sperm precedence values are shown for treatments PPW and PWP, with an overall tendency for limited last male sperm precedence ($P_2 = 0.584$; see below).

Patterns of paternity

There was no difference between P_2 values for treatments PPW and PWP (i.e. the order in which a pale female was presented a pale male and a wild-type male) when including all females that produced nymphs (GLM: $\chi^2_1 = 5.14$, $P = 0.713$) or when considering only females that produced both wild-type and pale nymphs and hence did not experience mating failure with either male they mated with ($\chi^2_1 = 55.03$, $P = 0.098$). This indicates there were no genotype-specific order effects on paternity, i.e. pale males did not gain a larger proportion of the paternity when they were the first male to mate compared with when wild-type males were the first to mate. Therefore, for all further paternity analysis, the two treatments will be combined.

Mating failure influenced patterns of paternity. When considering all females that produced nymphs (i.e. including females that experienced mating failure with one male, as well as those that experienced no mating failure), the mean P_2 was

0.584 ± 0.038, which was significantly different from equal paternity (one-sample Z-test: $\chi^2_1 = 261.9$, $P < 0.001$). Similarly, when considering only females that had both pale and wild-type nymphs (i.e. only females that did not experience mating failure with either male she was paired with), the mean P_2 was 0.548 ± 0.058, which was also significantly different from equal paternity ($\chi^2_1 = 41.29$, $P < 0.001$), thus showing a tendency towards last male paternity but not falling outside the expected range of 0.4–0.6 for a random sperm-mixing mechanism (García-González 2004).

There was, however, considerable variation in P_2 values. When only females that experienced no mating failure were included, the distribution of P_2 was relatively uniform (Fig. 1c). However, when including females that experienced mating failure with one male, a bimodal distribution arose with peaks at 0 and 1 (Fig. 1d). These results clearly follow the patterns predicted by García-González's (2004) simulation model for the distribution of P_2 values under a random sperm mixing mechanism, for a population with high levels of mating failure (Fig. 1a, b).

For all doubly-mated females, across all treatments, females copulated for longer on day 2 (mean = 232.2 min ± 6.1 min) than on day 1 (186.8 min ± 6.1 min; Paired t test: $t_{396} = -5.74$, $P < 0.001$). Because P_2 was significantly influenced by copulation duration (see below), this may well explain the slight tendency towards last male precedence.

The copulation durations on both day 1 and day 2 were significant as main effects, negatively associated with P_2 for day 1 (GLM: Duration day 1: $\beta = -0.009 \pm 0.001$, $F_{1,164} = 77.8$, $P < 0.001$) and positively associated with P_2 for day 2 (Duration day 2: $\beta = 0.009 \pm 0.001$, $F_{1,164} = 70.1$, $P < 0.001$). This means that longer copulations on day 1 were associated with higher P_1 and lower P_2, and vice versa for longer copulations on day 2; put another way, the longer the copulation then the greater the paternity associated with that copulation. There was also a significant interaction though, as the effect of copulation duration on one day was influenced by how long the copulation was on the other day (Interaction: $F_{3,162} = 7.40$, $P = 0.007$). Unsurprisingly, paternity was positively

Table 2 Summary of the number of pairs that mated twice in each treatment and how many of these produced no offspring (complete mating failure), only one morph of nymph or both nymph morphs.

Mean P_2 shown with binomial standard errors for treatments PPW and PWP when considering all females that produced nymphs and only females that produced both types of nymphs

Treatment	Number	Type of nymphs present				$P_2 \pm$ SE	
		None	Wild-type only	Pale only	Wild-type & pale	All females that had nymphs	Only females that had both wild-type and pale nymphs
PPP	93	19	0	74	0	–	–
PWW	82	23	59	0	0	–	–
PPW	109	30	22	21	36	0.572 ± 0.056	0.603 ± 0.082
PWP	113	26	13	37	37	0.595 ± 0.053	0.492 ± 0.082
PPW and PWP	222	56	35	58	73	0.584 ± 0.038	0.548 ± 0.058

associated with the difference in copulation duration on days 1 and 2, with males that mated for longer than the other male gaining higher paternity ($\chi^2_1 = 3901$, $P < 0.001$; Fig. 3). On the other hand, the difference in body length between males that mated on day 1 and males that mated on day 2 had no effect on the P_2 ($\chi^2_1 = 0.341$, $P = 0.923$).

Mating failure

Mating failure was associated with short copulations (when considering only singly-mated females, i.e. treatments: PPZ, PZW, PZP and PZW; GLM: $\chi^2_1 = 178.8$, $P < 0.001$; Fig. 4). The shortest copulation duration that resulted in offspring was 45 min (i.e. three consecutive observation checks *in copula*). The mean copulation duration for pairs that produced offspring was 281.9 ± 6.8 min, whereas the mean copulation duration for pairs that did not produce offspring was 105.2 ± 7.3 min (these were calculated by looking only at singly-mated females, treatments: PPZ, PZW, PZP, PZW).

Doubly-mated females were less likely to experience mating failure (24.7%) than singly-mated females (48.6%; GLM: $\chi^2_1 = 42.3$, $P < 0001$; Table 3). Additionally, doubly-mated females produced significantly more nymphs (mean = 43.7 ± 1.73) than singly-mated females (mean = 26.6 ± 1.97) when including females that experienced mating failure (GLM: $F_{1,685} = 42.2$, $P < 0.001$). However, when females that produced no nymphs were removed from the data set, this difference became less pronounced (singly-mated: mean = 51.8 ± 2.42; doubly-mated: mean = 58.0 ± 1.58: $F_{1,446} = 4.89$, $P = 0.028$). Therefore, mating twice helps to reduce the chance
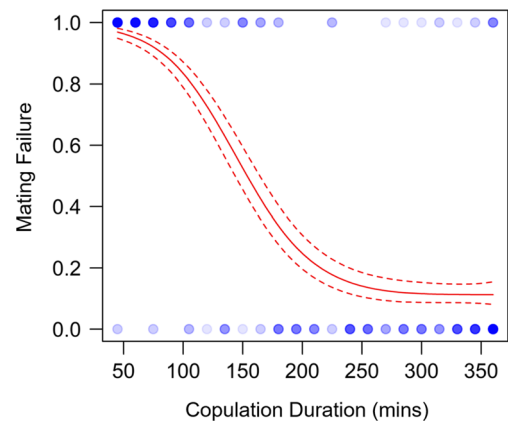
Fig. 4 Relationship between mating failure and mating duration for singly-mated females (treatments = PPZ, PWZ, PZP, PZW, $N = 734$), visualized as a cubic spline. Data are represented by circles, with colour reflecting sample size (the darker, the more replicates with that estimated copulation duration). Dashed lines indicate 1 standard error above and below the predicted line

of mating failure and also increases female fitness in terms of the number of nymphs produced (Fig. 5).

Comparing expected mating failure rates with observed mating failure rates for doubly-mated females, marginally more females experienced no failure or two failures than expected by chance (Chi-squared test: $\chi^2_2 = 5.859$, $P = 0.053$; Fig. 6). Therefore, mating failure could be a female-associated trait, with some females more likely to experience mating failure than others. For males, mating failure was irrespective of whether or not a female mated once or twice, so a male was as likely to fail to sire any offspring if the female mated once or twice ($\chi^2_1 = 0.42$, $P = 0.516$).

Sexual selection

In addition to exploring mating failure and P_2, we were also able to correlate paternity success with a number of phenotypes. Across the experiment, pale males sired marginally more offspring than wild-type males when there was no sperm competition (i.e. when females were singly-mated, treatments: PPZ, PWZ, PZP and PZW; GLM: $F_{1,288} = 4.32$, $P = 0.039$) but did not sire any more offspring than wild-type males when sperm competition was present (i.e. when females were doubly-mated, treatments: PPW, PWP; $F_{1,442} = 0.23$, $P = 0.634$; Table 4). These results may be driven by pale males being less likely to experience mating failure (41.2%) than wild-type males (53.5%; GLM: $\chi^2_1 = 11.26$, $P < 0.001$). This in turn might be linked to copulation duration, as females tended to mate longer with pale males (see below). These data provide evidence therefore of sexual selection favouring pale males (i.e. assortative mating by the pale focal females). Comparing within each morph of male, unsurprisingly, pale males sired more offspring when there was no sperm competition compared with pale males that experienced sperm
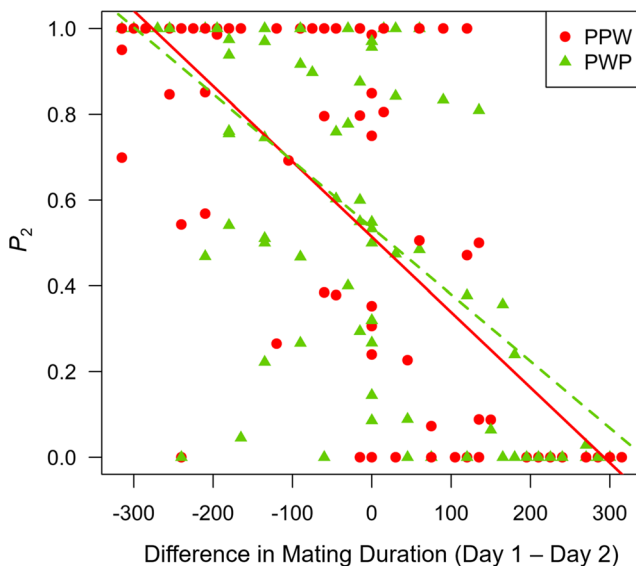

Fig. 3 Relationship between P_2 and the difference between the mating duration on days 1 and 2 for the treatments PPW and PWP ($N = 222$). For illustration, linear regressions are shown for the relationship between P_2 and mating duration difference in treatments PPW (solid line) and PWP (dashed line)

Table 3 Number of females succeeded or failed to produce offspring and the proportion of females that experienced mating failure (proportion failed). Mated 0 = all females that did not mate (for < 30 min: treatment = PZZ). Mated 1 = all treatments where females only mated on one day (for > 30 min: treatments = PPZ, PZP, PWZ, PZW). Mated 2 = all treatments where females mated on both days (for > 30 min: treatments = PPP, PWW, PWP, PPW)

Number of times a female mated	Number	Number of females that failed or succeeded in having offspring		
		Failed	Succeeded	Proportion failed
0	104	104	0	1.000
1	290	141	149	0.486
2	397	98	299	0.247

competition ($F_{1,377} = 8.09$, $P = 0.005$). However, the same was not true for wild-type males, who sired similar numbers of offspring whether they experienced sperm competition from pale males or not ($F_{1,353} = 0.32$, $P = 0.574$; Table 4).

In total, 68.5% of the males used in the experiment mated (Table 5). Pale males were more likely to mate than wild-type males (pale = 71.5%; wild-type = 65.5%; GLM: $F_{1,1578} = 6.25$, $P = 0.013$). Significantly more males mated on day 2 than on day 1 (day 1 = 61.4%; day 2 = 75.6%; $F_{1,1578} = 36.68$, $P < 0.001$) but there was no interaction between day and male phenotype on the likelihood of mating (Interaction: $F_{1,1578} = 1.15$, $P = 0.284$).

Furthermore, females copulated with pale males for longer than wild-type males. Pairs mated on average 36.8 min longer on day 2 than on day 1 ($F_{1,730} = 15.58$, $P < 0.001$) and pale males mated for 11.6 min longer than wild-type males ($F_{1,730} = 5.96$, $P = 0.015$) but there was no interaction between day and male phenotype on the mating duration (Interaction: $F_{1,730} = 0.78$, $P = 0.376$). Wild-type males initiated copulation

as quickly as pale males ($F_{1,732} = 0.60$, $P = 0.438$), however the mean latency to mate was shorter on day 2 (83.9 min ± 6.7 min) than on day 1 (118.2 min ± 11.1 min; $F_{1,732} = 25.94$, $P < 0.001$). Latency to mate was negatively correlated with copulation duration (Pearson's Correlation coefficient: $r_{732} = -0.472$, $P < 0.001$). This was likely due to the time restriction (6 h) of the experiment. Since all pairs were artificially split up at the end of the 6 h observation period, pairs that took longer to initiate copulation had a shorter maximum period of time left to copulate than pairs that initiated copulation sooner.

On day 2, males were more likely to mate with already-mated females from day 1 than virgin females (GLM: $\chi^2_1 = 24.8$, $P < 0.001$). However, for males that did mate, the mating status of females on day 2 did not significantly influence copulation duration ($F_{1,421} = 1.30$, $P = 0.254$), nor latency to mate ($F_{1,421} = 0.10$, $P = 0.757$).

Larger bugs were more likely to copulate across both days. This was true of both females (GLM: day 1: $\beta = 0.60 ± 0.13$, $\chi^2_1 = 20.9$, $P < 0.001$; day 2: $0.79 ± 0.15$, $\chi^2_1 = 28.3$, $P < 0.001$) and males ($\beta = 0.39 ± 0.10$, $\chi^2_1 = 14.4$, $P < 0.001$). Females that copulated were on average 0.19 mm larger on day 1 and 0.26 mm larger on day 2 than
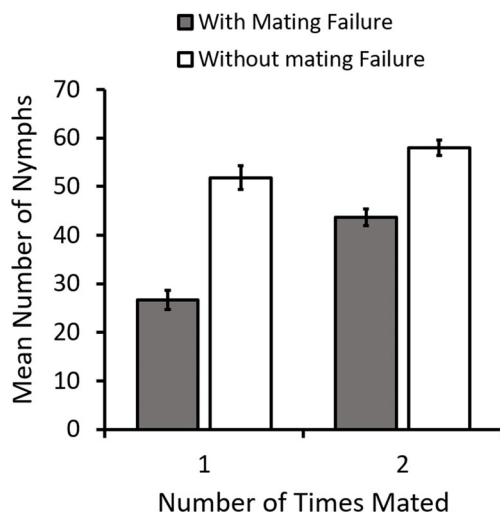

Fig. 5 Mean number of nymphs produced by females that mated once (singly-mated) or twice (doubly-mated), comparing all females (including both females that did and did not experience mating failure: i.e. both females that did and did not produce nymphs; grey: $N = 687$) with only those females that produced nymphs and hence did not experience mating failure (white, $N = 448$). Error bars show the standard error
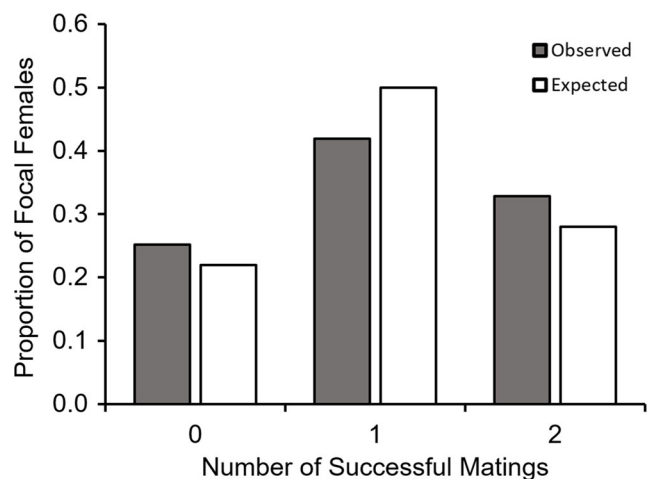

Fig. 6 The distribution of observed (grey) and expected (white) successful matings in treatments PPW and PWP ($N = 222$). A mating was deemed successful if nymphs with the same phenotype as the father were produced

Table 4 Mean number of nymphs sired by pale and wild-type males under different sperm competition conditions. No sperm competition: when females only mated with one male (singly-mated). Sperm competition: when females mated with two males (doubly-mated). Mating failure rate indicates the proportion of males that failed to sire any offspring

	Number of nymphs sired by pale males				Number of nymphs sired by wild-type males			
	N	Mean	Range	Mating failure rate	N	Mean	Range	Mating failure rate
No sperm competition	157	30.4 ± 2.7	0–113	0.414	133	22.2 ± 2.9	0–116	0.571
Sperm competition	222	21.6 ± 1.8	0–107	0.410	222	20.3 ± 2.0	0–135	0.514

females that did not, whereas males that copulated were only, on average, 0.10 mm longer than males that did not (Fig. 7).

Out of the pairs that mated, larger females were more likely to copulate for longer across both days (day 1: $F_{1,478} = 9.35$, $P = 0.002$; day 2: $F_{1,590} = 24.1$, $P < 0.001$) whereas there was no relationship between copulation duration and male body length ($F_{1,1065} < 0.001$, $P = 0.981$). Likewise, larger females engaged in copulation more quickly than smaller females on day 2 ($F_{1,590} = 9.22$, $P = 0.003$), but not on day 1 ($F_{1,478} = 0.42$, $P = 0.517$). For males, there was no such relationship between body size and mating latency ($F_{1,1065} = 0.09$, $P = 0.761$).

When considering treatments with singly-mated females (treatments: PPZ, PZP, PWZ, PZW), larger females produced significantly more nymphs than smaller females ($F_{1,288} = 39.04$, $P < 0.001$), but larger males did not sire more nymphs than smaller males (pale males [treatments PZP, PPZ]: $F_{1,154} = 0.26$, $P = 0.611$; wild-type males [treatments PWZ, PZW]: $F_{1,128} = 0.11$, $P = 0.745$).

Discussion

First, our results confirm that mating failure—in this case primarily due to failure to transfer sperm during copulation (Greenway et al. 2017)—can influence patterns of P_2, as argued by García-González (2004). Our results provide an empirical demonstration of the simulation models of García-González (2004), which predicted that, under a random sperm mixing mechanism, a population that experiences high levels of mating failure will show a strong bimodal skew in paternity with peaks at $P_2 = 0$ and $P_2 = 1$ (compare Fig. 1a, b and 1c, d). Without an appreciation of mating failure, the patterns of sperm precedence in *L. simulans* would fit with a mechanism of sperm competition such as sperm displacement. However, taking mating failure into account, the mechanism for sperm competition is more consistent with a random sperm-mixing model. Even though P_2 was significantly different from equal paternity (mean = 0.58), it did not fall outside the expected range of 0.4–0.6 for a random sperm mixing mechanism (García-González 2004). Second, there was considerable variance in the distribution of P_2 associated with the effects of copulation duration. This suggests that sperm loading (i.e. males transfer different qualities or quantities of sperm to females) might be at play (Simmons 2001). Third, our results provide some evidence that mating failure in this species could also be a female-associated trait, i.e. there is a tendency that some females are more likely to not be inseminated than others, but that this is perhaps not as strong a driver of mating failure outcome as male-associated traits have previously been found to be (Greenway and Shuker 2015; Greenway et al. 2017).

Our estimate of mating failure (leading to "non-sperm representation": García-González 2004) is indirect, as we used the production of pale and/or wild-type nymphs to assess whether sperm from one or both males was passed to a focal female. Most cryptic mating failure in *L. simulans* is associated with the transfer of negligible amounts of sperm or no sperm at all (Greenway et al. 2017). Nonetheless, we might have over-estimated mating failure (for instance if some males passed small ejaculates

Table 5 Number and proportion of males that mated on days 1 and 2 with regard to male phenotype

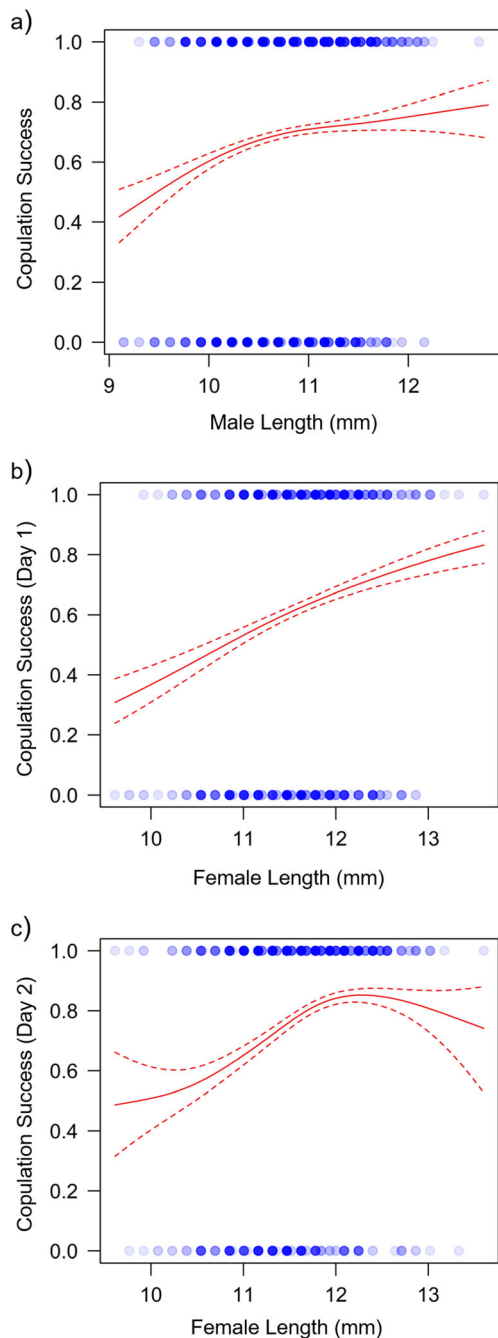| Phenotype	Sample Size	Number of males that mated (Y) or not (N)						Proportion that mated		
		Day 1		Day 2		Total		Day 1	Day 2	Total
		Y	N	Y	N	Y	N			
Pale	790	247	142	318	83	565	225	0.635	0.793	0.715
Wild-type	792	239	163	280	110	519	273	0.595	0.718	0.655
All males	1582	486	305	598	193	1084	498	0.614	0.756	0.685

Fig. 7 Relationship between copulation success (whether pairs copulated (1) or not (0)) and body length (all treatments) visualized as cubic splines. (**a**) Males ($N = 1561$), (**b**) females on day 1 ($N = 780$) and (**c**) females on day 2 ($N = 780$). Data are represented by circles, with the colour reflecting the number of individuals of the given size and mated state (darker = more replicates). Dashed lines indicate 1 standard error above and below the predicted line

and their sperm did not make it into the fertilization set of the eggs we then sampled as nymphs). However, the patterns of sperm precedence we see, once removing our estimated instances of mating failure, do not otherwise suggest clustering towards high P_1 or high P_2, so we are confident we have not missed significant amounts of 100%

first- or last-male paternity when both males successfully transfer sperm.

Taking mating failure into account then, our data suggest random sperm mixing with a potential role for sperm loading, since copulation duration was positively associated with increased paternity, for either the first or second male to mate. Copulation duration may be positively associated with quantity of sperm transferred (currently being tested) or it may be associated with some other aspect of sperm storage and usage by females (Eberhard 1996; Simmons 2001). We also note that the prolonged copulations in this species may also function as post-copulatory mate guarding (Alcock 1994; Sillén-Tullberg 1981), although that function was not relevant to paternity success in our experimental set-up. The slight tendency towards last male precedence could potentially be explained by the tendency for males to mate for longer on day 2 than on day 1. Longer copulations may allow more time to transfer a greater number of sperm, which males might do with already-mated females on day 2 as a response to the presence of another male's sperm, i.e. the actuality of sperm competition. Another reason for this tendency towards last male precedence could be due to sperm stratification (see discussion below). Although our results confirm that mating failure can generate a spurious bimodality in sperm precedence data, controlling for mating failure in our data generated a more uniform distribution, as opposed to a normal distribution about the mean, as predicted by García-González (2004) simulations. We will consider possible causes of variation in P_2 below, when we compare our work with previous estimates of P_2 in this genus.

Importantly, our data differ somewhat from those of Sillén-Tullberg (1981) who estimated much higher levels of last-male sperm precedence ($P_2 = 0.9$) in the sister species *L. equestris*. First, it could simply be that *L. equestris* uses a different sperm competition mechanism to *L. simulans*. This seems rather unlikely though, seeing as the two sister species appear to have very similar behaviours, life histories and genital morphologies (see Greenway 2017 for details). Second, some of the difference in estimated P_2 could be due to mating failure not being taken into account in the original study. That said, our data would not suggest such a large discrepancy in terms of mean P_2, rather a big difference in how variation is expressed about that mean. As such, the difference in sample size between the two studies could be of significance: Sillén-Tullberg's experiment had a much smaller sample size ($N = 10$ to 13 per treatment) than ours ($N = 109$ to 113 per treatment), therefore it is possible that such a high level of P_2 in the original study is a Type I error.

Third, copulation duration was found to be the greatest driver of paternity outcome in our experiment. Perhaps the key difference between our study and that of Sillén-Tullberg (1981) is that female *L. equestris* had 24 h unobserved pairing with 2 different males in her experiment. This meant that

copulations were not monitored and also that bugs had much longer to mate. If—as our data suggest—the second male can assess the mating status of the female, then longer and in particular repeated copulations by the second male in the Sillén-Tullberg study may well have produced the much higher value of P_2 observed. As such, the difference between the two studies may well have less to do with different mechanisms of sperm competition (driven by sperm loading and sperm mixing), as by the different opportunities the males had in terms of mating, having most impact in terms of the opportunity of the second male to respond to sperm competition.

A possible reason for the slight tendency towards last male precedence in our experiment could be due to stratification effects in the spermatheca. This is the idea that the last sperm to enter the spermatheca are the first to leave and hence fertilize eggs (Simmons 2001). If this is the case, then as more mixing occurs over time, the P_2 should become closer to 0.5 (Simmons 2001), such as was demonstrated in red flour beetles as discussed above (Lewis and Jutkiewicz 1998). Additionally, Haddrill et al. (2008) showed that in the two-spot ladybird, Adalia bipunctata, when females were mated to multiple males, temporal usage of sperm over time varied a lot and sperm from 'early' males to mate could still be used to fertilize eggs in later clutches. Therefore, if we had sampled nymphs that had hatched from eggs laid 2 days after the mating trials, compared with eggs laid 7 days after the mating trials, we might have expected to see a reduction in P_2 during this time. However, if we compare our results with Sillén-Tullberg's (1981) again, she collected eggs for 4 weeks after the mating trials. Therefore, time clearly did not reduce the P_2 in that instance.

Our data illustrate a bimodal pattern of P_2 due to the presence of mating failure, as was predicted in García-González's (2004) simulations. As noted above, we cannot rule out the possibility that not all the P_2 values of 0 and 1 may have been due to mating failure. However, given the extremely high rates of mating failure in singly-mated females (48.6%), and the proportion of doubly-mated females that produced no offspring at all (24.7%), the proportion of doubly-mated females that had P_2 values of 0 and 1 is, in fact, less females than expected (One-sample Z-test: $\chi^2_1 = 25.6$, $P < 0.001$; Fig. 6), arguing against an artificial inflation of mating failure.

How often might bimodal and trimodal patterns of paternity be due to mating failure in other species? Bimodal skews are prevalent in many species of Lepidoptera, but these are generally believed not to be caused by mating failure (LaMunyon and Eisner 1993; LaMunyon 1994; Cook et al. 1997; Mongue et al. 2015). On the other hand, one study in which the extreme bimodal paternity skews could be completely explained by mating failure was carried out by Evans and Magurran (2001) on the Trinidadian guppy Poecilia reticulata. Thirty percent of the double-matings in this experiment failed to result in any offspring, indicating

high levels of mating failure in the population. This is not dissimilar to the levels of mating failure in our experiment and so you would predict that there would be high numbers of double-matings in which one of the males failed to inseminate the female, hence the other male to mate received the full share of the paternity. The authors do not discuss this as a possible cause of their results, but this paper precedes García-González's (2004) paper, and, indeed, the concept of mating failure was not so well appreciated at the time.

In a couple of cases, trimodal skews of P_2 values have been reported, for example in the Australian field cricket Teleogryllus oceanicus (Simmons et al. 2006) and stalk-eyed flies Teleopsis dalmanni (Corley et al. 2006). Simmons et al. (2006) addressed the issue that this could be due to mating failure and so excluding any pairs with P_2 values of 0 and 1 during analysis to prevent the results being skewed by the occurrence of mating failure. Kock et al. (2006) took likewise cautions when analysing their paternity data on the scorpionfly, Panorpa germanica. For every male that they used in their double-mating experiment, they mated each male to a virgin female afterwards to see if she produced offspring to determine whether the male was infertile or not. Mating trials involving any males that did not produce offspring from these copulations were excluded from the analysis. This addresses the warnings that García-González (2004) gives about the risk of mating failure skewing P_2 distributions, leading to false conclusions about sperm competition mechanisms. However, Kock et al. (2006) concluded that, for the three females that had $P_2 = 1$ and the one female with $P_2 = 0$, these results could not have been due to infertile matings. The possibility that these could not have been due to mating failure should not be ruled out, however, as it could have been due to an intromission failure (García-González 2004) or even if the males found these females unattractive and chose not to inseminate them during copulation, a form of cryptic male choice, which is yet to be conclusively proven (Aumont and Shuker 2018).

A paper that cites García-González (2004) but does not clearly address his warnings about mating failure was Corley et al.'s (2006) study on stalk-eyed flies. They concluded that their trimodal patterns of P_2 showed that all modes of sperm usage were at play: sperm precedence, sperm mixing and the two of these in conjunction with one another. They mention male infertility but argue that it does not explain all of the results since their patterns of extreme P_2 values did not match up with their calculated mating failure rate. We would argue here, however, that they have not taken mating failure into account when drawing their conclusions about sperm usage (and have small sample sizes too), so have possibly drawn misleading conclusions about sperm usage in stalk-eyed flies. As such, here we want to re-emphasize the importance of taking mating failure into account when predicting mechanisms of sperm usage and sperm competition from P_2

values, as we have clearly shown empirically that P_2 values do become highly skewed as a result of mating failure, confirming the theoretical predictions laid out by García-González (2004).

Finally, we found evidence of both female and male mate choice. Females appeared to show pre-copulatory mate choice for pale males, being more likely to mate with pale than wild-type males, suggesting assortative mating for colour morph (all focal females were pale). Alternatively, it might be that pale males are more willing to mate with pale females than wild-type males are. Females also mated with pale males for longer, and this might explain why pale males sired more offspring than wild-type males under no sperm competition conditions (i.e. when females were singly-mated). As it is assumed that males control copulation duration in this species (Sillén-Tullberg 1981) this might indicate a form of post-copulatory choice in males (Arnqvist 2014). Males may also prefer females of the same phenotype to themselves, and so copulate with them for longer and potentially transfer more sperm to them, a form of cryptic male choice whereby males differentially allocate resources, such as ejaculates and nuptial gifts, to females during or after copulation depending on female phenotype (Bonduriansky 2001; Arnqvist 2014; Aumont and Shuker 2018). Males also exhibited pre- and post-copulatory choice towards larger females: they were more likely to engage in copulation, initiate copulation quicker and copulate for longer with larger than smaller females. This confirms previous findings (Dougherty and Shuker 2014). Larger females are more fecund (Balfour et al. 2018), so males will likely gain more fitness benefits from mating with large than small females. Males also appeared to be more willing to engage in copulation with once-mated than virgin females, as has previously been found in this and the sister species *L. equestris* (Sillén-Tullberg 1981; Micholitsch et al. 2000). Although this could be driven by males preferring once-mated females (perhaps because these females are more gravid as mating induces ovary development and egg maturation: Sillén-Tullberg 1984), this could also be driven by females that were virgin after day 1 being generally less willing to mate, carried over to day 2. To finish the story, larger males were also more likely to engage in copulation than smaller males. This could be due to female mate choice, or it could be that larger males were better able to manipulate females and coerce them into mating. Although highly significant, it must be noted that the difference in body size between males that mated and males that did not was very small compared with the difference in body size between females that did and did not mate. Further to this, in a previous study, larger males were more likely to succeed in siring offspring than smaller males (Greenway et al. 2017).

In conclusion, patterns of sperm precedence can be shaped and influenced by the occurrence of mating failure and so researchers should be careful to take this into account when using patterns of sperm precedence to make inferences about mechanisms of sperm competition. We also wish to highlight the implications that mating failure has on sexual selection through the disruption of sperm competition, and we suggest that future research should focus on furthering our understanding of the consequences that mating failure has on sexual selection.

References

Alcock J (1994) Postinsemination associations between males and females in insects: the mate-guarding hypothesis. Annu Rev Entomol 39:1–21

Arnqvist G (2014) Cryptic female choice. In: Shuker DM, Simmons LW (eds) The evolution of insect mating systems. Oxford University Press, Oxford, pp 261–274

Arnqvist G, Edvardsson M, Friberg U, Nilsson T (2000) Sexual conflict promotes speciation in insects. PNAS 97:10460–10464

Aumont C, Shuker DM (2018) Cryptic male choice. Curr Biol 28: R1171–R1189

Balfour VL, Aumont C, Dougherty LR, Shuker DM (2018) The fitness effects of a pale mutant in the aposematic seed bug *Lygaeus simulans* indicate pleiotropy between warning coloration and life history. Ecol Evol 8:12855–12866

Birkhead TR, Møller AP (1998) Sperm competition and sexual selection. Academic, San Diego

Birkhead TR, Pizzarri T (2002) Postcopulatory sexual selection. Nat Rev Genet 3:262–273

Bonduriansky R (2001) The evolution of male mate choice in insects: a synthesis of ideas and evidence. Biol Rev 76:305–339

Boorman E, Parker GA (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

Burdfield-Steel E, Shuker DM (2014) The evolutionary ecology of the Lygaeidae. Ecol Evol 4:2278–2301

Chapman T, Arnqvist G, Bangham J, Rowe L (2003) Sexual conflict. Trends Ecol Evol 18:41–47

Cook PA, Harvey IF, Parker GA (1997) Predicting variation in sperm precedence. Philos Trans R Soc Lond B Biol Sci 352:771–780

R Core Team (2019) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna https://www.R-project.org/

Corley LS, Cotton S, McConnell E, Chapman T, Fowler K, Pomiankowski A (2006) Highly variable sperm precedence in the stalk-eyed fly, *Teleopsis dalmanni*. BMC Evol Biol 6:53

Dougherty LR, Shuker DM (2014) Precopulatory sexual selection in the seed bug *Lygaeus equestris*: a comparison of choice and no-choice paradigms. Anim Behav 89:207–214

Dougherty LR, Shuker DM (2016) Variation in pre- and post-copulatory sexual selection on male genetal size in two species of lygaeid bug. Behav Ecol Sociobiol 70:625–637

Dougherty LR, Simmons LW, Shuker DM (2016) Postcopulatory sexual selection when a female mates once. Anim Behav 116:13–16

Dougherty LR, Rahman IA, Burdfield-Steel ER, Greenway EV, Shuker DM (2015) Experimental reduction of intromittent organ length reduces male reproductive success in a bug. Proc Biol Sci 282: 201507724

Eberhard W (1996) Female control: sexual selection by cryptic female choice. Princeton University Press

Evans JP, Magurran AE (2001) Patterns of sperm precedence and predictors of paternity in the Trinidadian guppy. Proc Biol Sci 268:719–724

García-González F (2004) Infertile matings and sperm competition: the effect of "nonsperm representation" on intraspecific variation. Am Nat 164:457–472

Greenway EV (2017) Causes and consequences of mating failure in the Lygaeidae. PhD thesis, University of St Andrews, Scotland

Greenway EV, Shuker DM (2015) The repeatability of mating failure in a polyandrous bug. J Evol Biol 28:1578–1582

Greenway EV, Dougherty LR, Shuker DM (2015) Mating failure. Curr Biol 25:R534–R536

Greenway EV, Balfour VL, Shuker DM (2017) Can females choose to avoid mating failure in the seed bug *Lygaeus simulans*? Anim Behav 129:61–69

Gschwentner R, Tadler A (2000) Functional anatomy of the spermatheca and its duct in the seed bug *Lygaeus simulans* (Heteroptera: Lygaeidae). Eur J Entomol 97:305–312

Haddrill PR, Shuker DM, Amos W, Majerus MEN, Mayes S (2008) Female multiple mating in wild and laboratory populations of the two-spot ladybird, *Adalia bipunctata*. Mol Biol 17:3189–3197

Hosken DJ, Ward PI (2000) Copula in yellow dung flies (*Scathophaga stercoraria*): investigating sperm competition models by histological observation. J Insect Physiol 46:1355–1363

Kock D, Epplen JT, Sauer KP (2006) Patterns of sperm use in the scorpionfly *Panorpa germanica* L. (Mecoptera: Panorpidae). Behav Ecol Sociobiol 60:528–535

LaMunyon CW (1994) Paternity in naturally-occurring *Utetheisa ornatrix* (Lepidoptera: Arctiidae) as estimated using enzyme polymorphism. Behav Ecol Sociobiol 34:404–408

LaMunyon CW, Eisner T (1993) Postcopulatory sexual selection in an arctiid moth (*Utetheisa ornatrix*). PNAS 90:4689–4692

Lessells CM, Boag PT (1987) Unrepeatable repeatabilities: a common mistake. Auk 104:116–121

Lewis SM, Austad SN (1990) Sources of intraspecific variation in sperm precedence in red flour beetles. Am Nat 135:351–359

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

Micholitsch T, Krügel P, Pass G (2000) Insemination and fertilization in the seed bug *Lygaeus simulans* (Heteroptera: Lygaeidae). Eur J Entomol 97:13–18

Mongue AJ, Ahmed MZ, Tsai MV, de Roode JC (2015) Testing for cryptic female choice in monarch butterflies. Behav Ecol 26:386–395

Parker GA (1970) Sperm competition and its evolutionary consequences in the insects. Biol Rev 45:525–567

Rhainds M (2010) Female mating failure in insects. Entomol Exp Appl 136:311–226

Sillén-Tullberg B (1981) A male 'postcopulatory' strategy in a promiscuous species *Lygaeus equestris* (Heteroptera: Lygaeidae). Behav Ecol Sociobiol 9:283–289

Sillén-Tullberg B (1984) Copulation as a determinant of non-diapause development in female *Lygaeus equestris*. Entomol Exp Appl 36: 261–264

Simmons LW (2001) Sperm competition and its evolutionary consequences in the insects. Preston University Press

Simmons LW, Siva-Jothy MT (1998) Sperm competition in insects: mechanisms and the potential for selection. In: Birkhead TR, Møller AP (eds) Sperm competition and sexual selection. Academic Press, London, pp 341–434

Simmons LW, Parker A, Stockley P (1999) Sperm displacement in the yellow dung fly, *Scatophaga stercoraria*: an investigation of male and female processes. Am Nat 153:302–314

Simmons LW, Beveridge M, Wedell N, Tregenza T (2006) Postcopulatory inbreeding avoidance by female crickets only revealed by molecular markers. Mol Biol 15:3817–3824

Smith RL (1984) Sperm competition and the evolution of animal mating systems. Academic Press, Orlando

Tadler A, Nemeschkal HL, Pass G (1999) Selection of male traits during and after copulation in the seedbug *Lygaeus simulans* (Heteroptera, Lygaeidae). Biol J Linn Soc 68:471–483

Waage JK (1979) Dual function of the damselfly penis: sperm removal and transfer. Science 203:916–918