No evidence of sexual selection in a repetition of Bateman's classic study of *Drosophila melanogaster*

Patricia Adair Gowaty[a,b,c,1], Yong-Kyu Kim[d,e], and Wyatt W. Anderson[d,1]

[a]Department of Ecology and Evolutionary Biology and [b]Institute of the Environment, University of California, Los Angeles, CA 90095; [c]Smithsonian Tropical Research Institute, DPO AA 34002; [d]Department of Genetics, University of Georgia, Athens, GA 30602; and [e]Department of Biology, Emory University, Atlanta, GA 30322

Contributed by Wyatt W. Anderson, May 11, 2012 (sent for review January 8, 2012)

We are unique in reporting a repetition of Bateman [Bateman AJ (1948) *Heredity* (*Edinb*) 2:349–368] using his methods of parentage assignment, which linked sex differences in variance of reproductive success and variance in number of mates in small populations of *Drosophila melanogaster.* Using offspring phenotypes, we inferred who mated with whom and assigned offspring to parents. Like Bateman, we cultured adults expressing dramatic phenotypes, so that each adult was heterozygous-dominant at its unique marker locus but had only wild-type alleles at all other subjects' marker loci. Assuming no viability effects of parental markers on offspring, the frequencies of parental phenotypes in offspring follow Mendelian expectations: one-quarter will be double-mutants who inherit the dominant gene from each parent, the offspring from which Bateman counted the number of mates per breeder; half of the offspring must be single mutants inheriting the dominant gene of one parent and the wild-type allele of the other parent; and one-quarter would inherit neither of their parent's marker mutations. Here we show that inviability of double-mutant offspring biased inferences of mate number and number of offspring on which rest inferences of sex differences in fitness variances. Bateman's method overestimated subjects with zero mates, underestimated subjects with one or more mates, and produced systematically biased estimates of offspring number by sex. Bateman's methodology mismeasured fitness variances that are the key variables of sexual selection.

genetic parentage | monogamy

Bateman's study (1) of within-sex selection in *Drosophila melanogaster* is a foundational paper in sexual selection, second only to Darwin's pioneering book (2); it empirically anchored within-sex variance in number of mates (V_{NM}) as a key correlate of variance in reproductive success (V_{RS}) and as the metric of sexual selection. Bateman said his results showed that male number of mates (NM) was more variable than female NM; male reproductive success (RS) was more variable than female RS; and RS in males, but not in females, was because of NM. His conclusions were: sexual selection acted primarily on males through female choice and through male competition and profligacy in mating, so that some males mated more frequently than others, producing higher V_{RS} among males than among females because of the positive relationship between number of mates and reproductive success for males, but not for females.

Bateman's (1) paper was cited relatively infrequently before its rediscovery by Trivers (3), who used Bateman's results to buttress his arguments that the sex-differential cost of reproduction selectively favored coy, discriminating females and competitive, ardent males. After Trivers (3), citation of Bateman soared (4), as it did again after Arnold (5) discussed "Bateman's Principles" as corollaries of sex differences in behavior and fitness variances. Given its paradigmatic status, Bateman's paper has inspired further studies of V_{NM} and V_{RS} (6, 7), many of which are consistent with Bateman's main conclusions. Despite consistency in some studies and the apparent simplicity of Bateman's original design, Bateman's methods, the generality of Bateman's conclusions, and their implications are controversial (3, 8–11) (*SI Text*), suggesting that "Bateman's Principles" might be better phrased as "Bateman's Hypotheses." Thus, it is interesting that the present report is unique in being a replication of Bateman's experiment that explicitly used his methodology of inferring who mated with whom by assigning parentage to offspring inheriting dramatic parental mutant phenotypes. Here we show that Bateman's methodology violated an assumption crucial to the reliability of his inferences: the methodology obscured some observations so that some matings that occurred were not counted, thus overestimating the number of subjects with no mates to an unknown degree and underestimating the number of subjects with one or more mates, also to an unknown degree. Inaccurate counts of number of mates and number of offspring per adult thus biased estimates of NM and V_{NM}, making conclusions based upon NM and V_{NM}, such as those from plots of the relationship of NM to RS, unreliable and potentially misleading.

Bateman's experiment was conceptually simple (1, 4), and used the only method of genetic parentage assignment available in the 1940s: heritable, dramatic, and phenotypically obvious genetic mutations to identify the parents of offspring in small, replicated trial populations. Unlike modern molecular genetic studies, in which it is theoretically possible to assign paternity and maternity to all offspring, in Bateman's study only some offspring carried the phenotypic markers of their parents, limiting Bateman's inferential power relative to what is possible in modern molecular genetic studies of parentage. Bateman's experiment involved first the production of heterozygous-dominant adults carrying a "marker mutation" as one allele at their marker locus and a wild-type gene as the other allele of the marker locus. Within a population, regardless of their sex, each adult was phenotypically distinct: no adult was homozygous at its one marker locus and each adult carried only wild-type alleles at all other marker loci (Table S1).

Each offspring has both a mother and a father, which guarantees that the frequency of offspring inheriting parents' marker mutations is the Mendelian expectation when parents are heterozygotes at two different loci (Tables S2 and S3) and provides a simple way to check the assumption that inviability of combinations of parental marker alleles in offspring did not significantly affect counts of NM or RS.

Some of Bateman's trials used three individuals of each sex, others five individuals of each sex. For populations with three of each sex, there were six phenotypically distinct individuals regardless of their sexes; similarly, in populations with five of each

Author contributions: P.A.G., Y.-K.K., and W.W.A. designed research; P.A.G. analyzed data; Y.-K.K. supervised technicians who carried out the experiment, and entered observations into computerized records; and P.A.G. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

[1]To whom correspondence may be addressed. E-mail: gowaty@eeb.ucla.edu or wyatt@uga.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1207851109/-/DCSupplemental.

sex, there were 10 phenotypically distinct individuals. Bateman placed replicate sets of these potential breeders in small bottles, each constituting a separate population from which fitness variances (population parameters) could be calculated. He allowed subjects to associate and mate for 3 or 4 d, and then discarded the adults. As pupae eclosed, he collected offspring and scored the presence or absence of parental marker mutations, from which he inferred parentage. He inferred NM for each subject from offspring who inherited a phenotypically obvious dominant marker allele from each parent [i.e., the "double mutant ($M^{\varphi}M^{\sigma}$) offspring," from which one can calculate the V_{NM} for each small study population]. He calculated the RS of each adult as the sum of its $M^{\varphi}M^{\sigma}$ offspring plus those "single mutant" offspring, $M^{\varphi}w^{\sigma}$ or $w^{\varphi}M^{\sigma}$, which inherited the mother's mutation but not the father's, and vice versa. Bateman used ANOVA to test for effects of parental age, marker phenotypes, and sex differences in V_{RS} summed over sets of populations.

The crucial assumption of Bateman's method is that there is no reduction of offspring viability from inheritance of parental markers, particularly when offspring inherit a mutation from each of its parents ($M^{\varphi}M^{\sigma}$). $M^{\varphi}M^{\sigma}$ offspring are the only offspring from which NM for each adult could be inferred using Bateman's method.

Three Explanations for Bateman's Data

Today there are at least three hypotheses explaining the observed V_{NM} and V_{RS} of potential parents in Bateman's original experiment: (*i*) Inherited parental mutations with effects on viability resulted in missing offspring that biased counts of NM (4) (Tables S1–S5). If there are unbiased descriptions of who mated with whom, then it is reasonable to evaluate two other nonmutually exclusive hypotheses: (*ii*) stochastic demography (chance effects on survival and reproduction) in the absence of mate choice or within-sex behavioral or physiological competition resulted in observed V_{NM} and V_{RS} (10, 11); and (*iii*) sexual selection among males resulted in observed sex differences in NM and RS.

To reject hypothesis (*i*) of viability effects on RS, it is necessary only to demonstrate that observed frequencies of offspring phenotypic types—$M^{\varphi}M^{\sigma}$, $M^{\varphi}w^{\sigma}$, $w^{\varphi}M^{\sigma}$, and $w^{\varphi}w^{\sigma}$—occur in frequencies expected under Mendel's rules (Tables S1–S5). Under the assumption of no viability effects on offspring of inheriting parental marker mutations, (*i*) half of the offspring from each subject adult are identifiable and equal for mothers and fathers, and (*ii*) every set of parents in Bateman's experiment must produce similar frequencies of four types of offspring, as can easily be seen in Tables S6–S8: one-quarter must be $M^{\varphi}M^{\sigma}$, double-mutant offspring, inheriting a marker allele at the locus uniquely associated with each parent; one-quarter must be $M^{\varphi}w^{\sigma}$, single-mutant offspring, inheriting the marker only at their mother's marker locus but the wild-type allele at their father's marker locus; one-quarter must be $w^{\varphi}M^{\sigma}$, single-mutant offspring inheriting the wild-type allele only from their mother's marker locus but the marker allele from their father's marker locus; and one-quarter must be $w^{\varphi}w^{\sigma}$ offspring, inheriting neither of their parents' marker mutations.

Differential mating success of some individual adults over others, either because of sexual selection or stochastic demography, cannot cause deviations in expected Mendelian frequencies of $M^{\varphi}M^{\sigma}$, $M^{\varphi}w^{\sigma}$, $w^{\varphi}M^{\sigma}$, and $w^{\varphi}w^{\sigma}$ of one-quarter each (*SI Text* and Tables S6–S8), because each offspring has one mother and one father. When viability effects on offspring are ruled out, and if V_{NM} is significantly greater or less than stochastic demography predicts (10–12), one might with additional data claim a role for sexual selection in the differential within-sex mating success of males and females. If observations are consistent with viability effects on offspring, the conclusion is that sexual selection caused V_{NM} would be unjustified, because the data would be inadequate for tests of sexual selection. Similar logic organizes preliminary tests of marker suitability in modern molecular genetic parentage assignments (13–15).

Here we report the results of a comparison of offspring marker phenotypes from a two-part study. In the first part (Table S9) we tested the crucial predictions about viability effects using data (Tables S10–S13) from our repetition of Bateman's experiment. Our questions included: Were mothers and fathers equally represented among the offspring from each population and did offspring inherit parental mutations in the expected one-quarter frequencies of $M^{\varphi}M^{\sigma}$, $M^{\varphi}w^{\sigma}$, $w^{\varphi}M^{\sigma}$, or $w^{\varphi}w^{\sigma}$? In the second part, we confined breeders to monogamous pairs and compared the observed numbers of offspring in each phenotype class (data in Table S14) to those in our replication of Bateman's multimale and multifemale populations. Mate choice and behavioral or physiological competition over mates are not possible in enforced monogamous pairs. Observations of fewer than expected $M^{\varphi}M^{\sigma}$ offspring from the monogamous pairs would be evidence against the utility of Bateman's method for evaluating the hypotheses of sexual selection and stochastic demography.

Results

The repetition of Bateman's experiment shows that some parental genotypes (Fig. S1) were more common in offspring than others, consistent with hypotheses of sexual selection, demographic stochasticity, and differential offspring survival. However, bias in the methodology is obvious in that mothers were statistically significantly less often counted as parents than fathers, a biological impossibility in diploid sexual species. Of the 8,093 offspring, 3,350 (41%) expressed the mother's marker but 3,646 (45%) expressed the father's marker and the difference was statistically significant (Fig. 1*A* and Fig. S1).

The proportions of offspring in each of the phenotypic classes departed strongly from Mendelian expectations: among the 8,093 offspring in our 46 replicated populations 2,343 (29%) were $w^{\varphi}w^{\sigma}$ offspring; 2,401 (30%) were $w^{\varphi}M^{\sigma}$ offspring; 2,102 (26%) were $M^{\varphi}w^{\sigma}$ offspring; and 1,247 (15%) were $M^{\varphi}M^{\sigma}$ offspring (Table S13).

The frequencies are a significant departure from the expected one-quarter frequencies (likelihood ratio $\chi^2 = 463.1$, $df = 3$, $P < 0.0001$) with the highest contribution to χ^2 coming from the $M^{\varphi}M^{\sigma}$ category. Of the 46 populations, 44 had fewer than 20% (range from 6.9%) $M^{\varphi}M^{\sigma}$ offspring (Table S13). No population had a frequency of $M^{\varphi}M^{\sigma}$s over 24.3%. The binomial probably that all 46 populations would have $M^{\varphi}M^{\sigma}$ frequencies under 25% is 1.42×10^{-14}.

Biased estimates of NM are obvious from inconsistencies between the inferences allowed from double-mutant offspring (i.e., who mated with whom) and single-mutant offspring that provided an estimate of the number of additional offspring a given individual had. Some subject adults seemed to have zero mates (their markers did not appear in $M^{\varphi}M^{\sigma}$ offspring) but did in fact mate, because their markers appeared in $M^{\varphi}w^{\sigma}$ or $w^{\varphi}M^{\sigma}$ offspring. Among the subjects in our replicate, 21 (12.7%) females and 43 (25.9%) males were binned in the category "zero mates" (based on $M^{\varphi}M^{\sigma}$ offspring). However, 4 (19%) of the zero-mating females had 17 offspring (based on $M^{\varphi}w^{\sigma}$) from whom it was impossible to infer the father; and 15 (35%) of the zero-mating males (based on $M^{\varphi}M^{\sigma}$ offspring) had 245 offspring scored from $w^{\varphi}M^{\sigma}$ offspring from whom it was impossible to infer the mother.

Reasoning that one could use the $M^{\varphi}M^{\sigma}$ offspring along with $w^{\varphi}M^{\sigma}$ and $M^{\varphi}w^{\sigma}$ offspring to estimate RS might be justified if inviability effects of different parental marker combinations were similar. The frequencies of observed combinations of specific parental alleles in $M^{\varphi}M^{\sigma}$ offspring were statistically significantly different, indicating that some parental marker combinations were more deleterious than other combinations [for parental

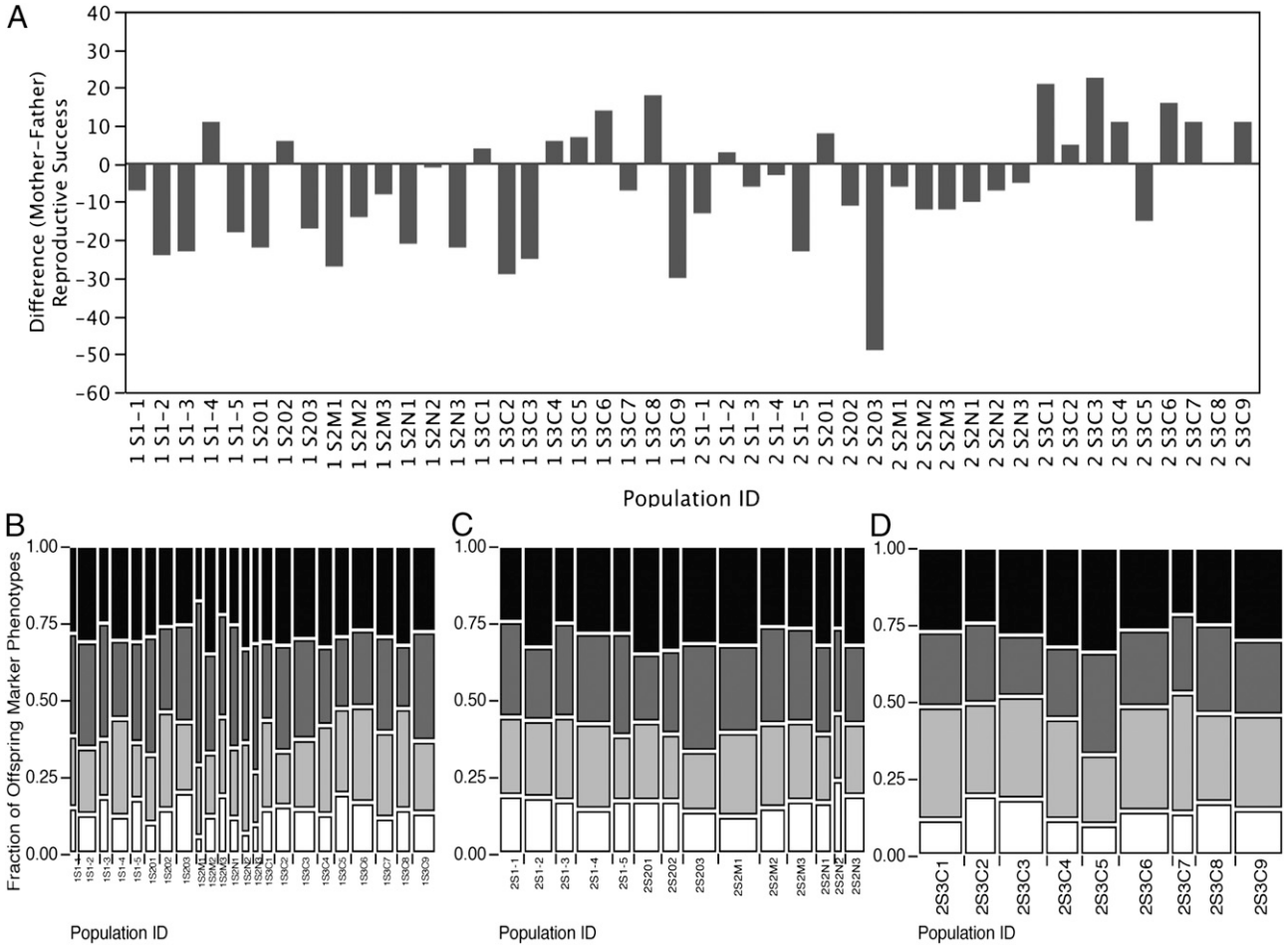

Fig. 1. (*A*) The distribution of differences in RS of parent (fathers minus mothers) is significantly different from zero. RS for mothers was Σ $M^{♀}M^{♂}$+ $w^{♀}M^{♂}$ and for fathers was Σ $M^{♀}M^{♂}$+ w $M^{♂}$. The actual estimate of difference in the repetition is 6.4 ± 15.67 (SD), *df* = 45, *t* test = 2.784, $P > |t| = 0.0078$ and signed-rank test = 226, $P > |t| = 0.0091$. It is logically impossible in sexual diploid species for more offspring to have fathers than mothers. Bateman's method of estimating RS produced a systematic bias with males having more offspring than females, a bias that could have inappropriately decreased the estimate of maternal RS, producing inaccurate estimates of sex differences in "Bateman gradients" (4). There were seemingly more fathers' children than mothers' children in Bateman's original experiment as well. (*B–D*) The frequency distributions of double-mutant ($M^{♀}M^{♂}$ indicated by white bars), single-mutant ($M^{♀}w^{♂}$ indicated by light gray bars and $w^{♀}M^{♂}$ indicated by dark gray bars), and no mutant ($w^{♀}w^{♂}$ indicated by black bars) offspring in 46 trials that replicated Bateman. Frequencies of offspring mutant combinations in each population are in Tables S10–S13. Parent marker sets: (*B*) ♀♀ = Pm, H, Mé and ♂♂ = B, Cy, Mc; likelihood ratio $\chi^2 = 260.5$, *df* = 3, $P < 0.0001$. (*C*) ♀♀ = Hw, Pm, Sb, H, Mé and ♂♂ = B, Cy, ap^{Xa}, Bl, Mc; likelihood ratio $\chi^2 = 157.1$, *df* = 3, $P < 0.0001$. (*D*) ♀♀ = B, Cy, Mc and ♂♂ = Pm, H, Sb; likelihood ratio $\chi^2 = 89.1$, *df* = 3, $P < 0.0001$. Within each marker set and for almost all populations there were significantly fewer than 25% double-mutant offspring, indicating that offspring inviability confounded estimates of NM and V_{NM} (see also Table S9), potentially biasing any estimates of Bateman gradients.

sets ♀♀ = B, Cy, Mc and ♂♂ = Pm, H, Sb, the χ^2 likelihood ratio = 35.6, *df* = 8; $P < 0.0001$; for parental sets ♀♀ = Hw, Pm, Sb, H, Mé and ♂♂ = B, cy, ap^{Xa}, Bl, Mc, the χ^2 likelihood ratio = 492.6; *df* =18, $P < 0.0001$; for parental sets ♀♀ = Pm, H, Mé and ♂♂ = B, Cy, Mc, the χ^2 likelihood ratio 597.6, *df* =7, $P < 0.0001$]. Some parental marker combinations thus were more deleterious to offspring viability than others, further compromising their reliability as unbiased estimators of among-individual within-sex and between-sex differences in RS and V_{RS}.

A contingency analysis of parental marker phenotypes in offspring ($M^{♀}M^{♂}$, $w^{♀}M^{♂}$, $M^{♀}w^{♂}$, $w^{♀}w^{♂}$) from populations (Table S13) and monogamous pairs (Table S14) consisting of ♀♀, each heterozygous for either Hw, Pm, Sb, H, Mé markers, and of ♂♂, each heterozygous for either B, Cy, ap^{Xa}, Bl, or Mc markers showed that offspring were statistically significantly different from the expected frequencies of one-quarter in each class in a two-by-four contingency test (likelihood ratio $\chi^2 = 27.8$, *df* = 3, $P < 0.0001$) and in a one-way contingency test summed over each treatment type. The overall distribution (summed across populations and monogamous pairs) was also significantly different from one-quarter in each class ($P < 0.0001$). The largest deviation from the expected frequency of one-quarter was in $M^{♀}M^{♂}$ offspring. In the populations 16.5% of offspring were $M^{♀}M^{♂}$; in the monogamous pairs 20.3% were $M^{♀}M^{♂}$. In monogamous pairs and the populations the distribution of $M^{♀}M^{♂}$, $M^{♀}w^{♂}$, $w^{♀}M^{♂}$, and $w^{♀}w^{♂}$ genotypes was significantly different from expected, and in both treatments the largest contribution to χ^2 was from the $M^{♀}M^{♂}$ class (Tables S13 and S14). The deficits in the frequency of $M^{♀}M^{♂}$ offspring from monogamous pairs were similar to those observed in the Bateman-like populations.

Discussion

Assumption of No Viability Effects on Offspring of Some Parental Marker Combinations Was Not Met by Our Data. Reduced viability of $M^{♀}M^{♂}$ offspring was near ubiquitous in the Bateman-like

populations and monogamous pairs. Reduced viability precludes any conclusions about the existence or force of V_{NM} on V_{RS}, even though, like Bateman's (1) results, the repetition seems to show that some mothers and some fathers had more mates and offspring than others (Fig. S1). However, concluding that sexual selection affected V_{NM} is unwarranted because the assumption that 25% of offspring be $M^{♀}M^{♂}$ was violated (Fig. 1 *B–D* and Table S13).

Lack of viability produced the significant differences in offspring assigned to fathers and mothers (Fig. 1*A*), and affected apparent RS for each inferred parent. Number of offspring for mothers and fathers must be equal in diploid sexual organisms; and importantly, because means and variances are positively correlated, the differences in RS for fathers and mothers would also bias estimates of sex differences in V_{RS}. In most populations there were more offspring assigned to fathers than to mothers (Fig. 1*A*), which might have produced erroneous conclusions of greater V_{NM} among males than among females. Lower frequencies of $M^{♀}M^{♂}$s also produced biases affecting inferences of NM. Such biases arise because of missing $M^{♀}M^{♂}$s (Fig. 1 *B–D* and Table S13), the only offspring class from which NM per adult could be estimated in either Bateman's or the present experiment, obscuring mating that occurred and did not occur in both sexes. The bias causes inaccuracies in the counts of individuals with zero, one, two, or three or zero, one, two, three, four, or five mates (depending on the number of potential mates in a given population), because the bias necessarily overestimated individuals with zero mates and underestimated individuals with one, two, or three mates or one, two, three, four, or five mates (depending on population size).

It seems there is little way to know, using Bateman's methodology, how to fairly apportion subjects to the categories of zero or more mates or to calculate reasonable estimates of V_{NM} or sex differences in V_{NM} in the populations. Nineteen percent of zero-mated females and 35% of zero-mated male subjects did mate, because their marker genes appeared in $M^{♀}w^{♂}$ and $w^{♀}M^{♂}$ offspring, an incongruity that demonstrates that using Bateman's method overestimates the number of individuals with zero mates, but simultaneously underestimates those with more than zero mates. Almost twice as many males as females are inappropriately binned in the zero-mated category (from $M^{♀}M^{♂}$ offspring), inaccurately inflating male V_{NM} and perhaps inappropriately biasing conclusions of sex differences in V_{NM}. Bateman's method mismeasures the key variables of sexual selection.

Is There an Unbiased Way to Estimate Number of Mates, RS, V_{NM}, and V_{RS} from the Data in Our Repetition? One might consider culling the data, retaining only those offspring with a father and mother in the $M^{♀}M^{♂}$ class, but this would reduce the total number of subject adults in each population, in some cases inappropriately biasing the adult sex ratio and eliminating altogether the class of individuals with zero mates. Readers would then argue that assessing the zero mating class is essential and at the heart of measuring sexual selection via female choice and among male competition. Eliminating the zero class from an analysis of the force of stochastic demography would likely render that test suspect as well.

Viability Deficits also Occurred in Monogamous Pairs in Which Sexual Selection Could Not Occur. The similarity in the frequencies of offspring phenotypes from populations and monogamous pairs provides experimental consistency, justifying the conclusion of unreliable inferential power and emphasizing the weakness of Bateman's methodology for evaluation of sexual selection. In the monogamous pairs the $M^{♀}M^{♂}$ deficit could not have resulted from male-male competitive interactions or from female choice of alternative mates, leaving only the hypothesis that the inviability caused the deficits in $M^{♀}M^{♂}$s. The deficit of $M^{♀}M^{♂}$ offspring was higher in the populations than in the monogamous pairs, an effect that could be a result of the higher number of females laying eggs: offspring competitive effects per vial were likely much higher in populations than in monogamous pairs.

Data in the Repetition Are Unable to Test Predictions of Sexual Selection. Bateman's method was flawed in our repetition of it, as it was in his study (Tables S1–S5). In the replication, it would be unjustifiable and misleading to: (*i*) estimate V_{NM} for either sex, (*ii*) test for sex differences in V_{NM}, (*iii*) test for sex differences in RS and V_{RS}, (*iv*) assess the relationship of NM to RS in either sex, or (*v*) quantify sex differences in the slope of NM on RS.

Were Bateman's Data Biased and Unable to Test Predictions of Sexual Selection? We endeavored to use exactly the same mutant lines Bateman used. All but one of Bateman's mutant lines is available today (Table 1). It is difficult to know how much the mutant lines changed in the 60 y between Bateman's experiment and the repetition. However, Bateman (1) indicated that 7 of 10 marker mutations were homozygous-lethal. That Bateman's subject adults carried mutant markers that were homozygous-lethal originally stimulated the hypothesis (4) that $M^{♀}M^{♂}$s inheriting a dramatic or sometimes disfiguring mutation at the mother's marker locus and a different mutation from the father's marker locus would suffer inviability that could bias counts of NM and RS. The first demonstration (4) of a lack of viability came from Bateman's own data (Tables S1–S5), using the only population for which he reported a complete record of offspring phenotypes. Table S4 is a replica of a table in Bateman (1); Table S5 shows that the $M^{♀}M^{♂}$s in that population were significantly fewer than one-quarter, and the RS of females is greater than males. Our repetition of Bateman's experiment also replicated similar biases to those apparent in Bateman's table (see table 4 in ref. 1). As his table contains the only offspring genotypes and their frequencies available from his paper (Table S4), and assuming that the population in Bateman's published table was a representative example of his overall data, it is probably safe to assume that Bateman's original experimental methodology produced biased results not too dissimilar from the biased results of this repetition.

Previous reexamination (4) of the data in Bateman's paper also showed that despite the pattern in the one population for which he published all of the observations, overall in his original experiment more offspring had fathers than mothers, prima facie evidence of bias in his original data, not dissimilar from the biases that emerged when we repeated his methodology (Fig. 1*A*).

Did Bateman Know About the Problem of $M^{♀}M^{♂}$s? Bateman did realize that viability effects of the inherited marker alleles could

Table 1. Mutant *D. melanogaster* stocks used in the present repetition

Chromosome	Symbol	Name	Stock number
I	Hw	Hairy-wing	102024 (Kyoto)
	B	Bar	2969 (Bloomington)
II	Bw^{V1}* (=Pm)	Plum	380 (Bloomington)
	Cy*	Curly	1430 (Bloomington)
	Bl*	Bristle	237 (Bloomington)
	ap^{Xa}*	Apterous-Xasta	Extracted from Mc
III	Sb*	Stubble	2539 (Bloomington)
	Mé*	Moire	894 (Bloomington)
	H*	Hairless	515 (Bloomington)
	Mc	Microcephalous	101603 (Kyoto)
Wild-type	Oregon-RS		4269 (Bloomington)

Bateman used CyL^4 in his experiments; we replaced CyL^4 with ap^{Xa}.
*Homozygous lethal.

create methodological biases. He said "... assuming the complete viability of all the marker genes, half the progeny of each fly could be identified" (1). Bateman also noted that the $M^{\varphi}M^{\sigma}$, double-mutant class of offspring was the only offspring class from which he could infer NM and V_{NM}, but he did not report a test for the effects of offspring viability deficits in $M^{\varphi}M^{\sigma}$. Thus, it seems Bateman did not actually check whether his observed frequencies of $M^{\varphi}M^{\sigma}$ were consistent with frequencies required by Mendelian expectations of inheritance of alleles at multiple loci when parents are heterozygous for dominant alleles, each at a different locus. Such a check of expected frequencies of parental combinations (Tables S1–S8) would have revealed the problem of his methodology to Bateman, to those who cite him, and to the legions of graduate students who have read the paper since it was published.

If the population for which Bateman did provide all observations (Tables S1–S5) was representative of his other populations, he would have systematically overestimated the number of adult subjects with zero mates and underestimated the number of adult subjects with one or more mates: his data would be inappropriate for tests of the predictions of sexual selection. If so, Bateman's conclusions about (*i*) V_{NM} for either sex, (*ii*) sex differences in V_{NM}, (*iii*) sex differences in RS, (*iv*) the relationship of number of mates to RS in either sex, and (*v*) sex differences in the slope of individuals that varied in number of mates on RS would be inaccurate to some unknown degree.

Conclusions

We conclude from our repetition of Bateman's experiment and from the evidence in his paper reviewed herein and elsewhere (4, 8–10, 16), that he had relatively weak evidence for his conclusions that (*i*) sexual selection acted primarily on males through female choice and male competition and profligacy in mating, and (*ii*) some males mated more frequently than others, producing higher V_{RS} among males than among females.

Data Do Not Allow Tests of Sexual Selection Predictions. We conclude also that there is no basis in the replication for testing the predictions of stochastic demography and sexual selection. Like Bateman, we did not observe copulations, so we do not have an independent measure of NM to substitute for biased observations from $M^{\varphi}M^{\sigma}$. Further resolution of the possibility of simultaneously acting sexual selection and chance effects of demography on NM requires another kind of experiment. The best way to sort out these possibilities would be to repeat Bateman's original design, varying the numbers of potential breeders and duration of the time available for mating, while using: (*i*) wild-type adults instead of mutants, thus eliminating the basic problem with Bateman's method; (*ii*) observations of behavior to document number of mates per individual; (*iii*) correlations between bearers' wild-type traits and their NM; and (*iv*) genetic inferences of parentage from molecular markers neutral with respect to offspring viability. Such studies would provide a basis for (*v*) testing the effects of demographic stochasticity (5–9) on fitness variances, and simultaneously testing for sexual selection and other deterministic effects on RS. Variation in NM and RS are insufficient to demonstrate selection without critical evaluation of alternative explanations.

Of course, it remains possible that in the replicated populations both viability effects on offspring of inherited parental mutations, stochastic demography, and sexual selection acting through NM could have simultaneously operated. However, neither Bateman's original experiment nor our replicated populations used methods that can answer that question.

The Future. Are there implications for other studies of NM, V_{NM}, RS, and V_{RS}? Recent studies relying on molecular genetic markers of parentage are less likely to bias offspring survival than the mutants Bateman or we used. However, even the most-unbiased molecular markers provide only partial information about NM, because mating does not guarantee offspring production: absence of offspring is not necessarily absence of mating. Thus, we urge future investigators to include behavioral observations for inferences of NM.

We are left wondering why earlier readers failed to spot the inferential problems with Bateman's original study. The main implication we take from the present study is one earlier critics (8, 9) made: The paradigmatic power of the world-view (16) captured in Bateman's conclusions and the phrase "Bateman's Principles" (5) may dazzle readers, obscuring from view methodological weaknesses and reasonable alternative hypotheses explaining V_{NM} and V_{RS}.

Methods

A comparison of our methods with Bateman's (1) is given in Table S15.

Subjects, Stocks, and Crossing Schemes. Bateman's crossing scheme for producing adult subjects was simple but labor-intensive. First, he cultured virgin subjects for sets of trials he labeled series 1, 2, and 3. He crossed heterozygous mutant females to wild-type males of the Oregon-RS *Drosophila melanogaster* strain, producing 50% heterozygotes and 50% wild-types. To culture subjects for his series 4, 5, and 6, Bateman crossed mutant females to inbred Oregon males for 200 generations, followed with backcrossing of flies for three and six generations. Because the series 4, 5, and 6 were time-intensive and reduced background genetic variation through use of inbred males and multigeneration backcrosses, we limited our repetition to the simpler, least time-intensive culturing scheme of Bateman's first three series, which also provided subjects with more wild-type background genetic variation.

To generate subjects we used stocks of female mutant *D. melanogaster* (Table 1). We backcrossed the female mutants to wild-type male Oregon-RS males to replicate Bateman's culturing scheme for his series 1, 2, and 3.

Each subject was a heterozygote that carried a single dominant gene for distinctive phenotypes (markers), each unique in their population (Table S1). The genetic mutations include those that Bateman used (see Bateman's table 2 in ref. 1) except for CyL, which is as far as we can tell, no longer available (see Table 1). Instead of CyL we used ap^{Xa}, which like CyL affects wing morphology. As many as seven of the mutants Bateman used were homozygous lethal. Bateman lists five markers as lethal in homozygous condition; two others were probably homozygous-lethal because Bateman designated them as "same or similar" to related homozygous-lethal markers. The marker lines we used for the current experiment contain seven marker genes that are also homozygous-lethal. We crossed mutant females with wild-type males of the Oregon-RS strain in mass cultures. From these crosses, we collected virgin mutant males and females to use in two temporally separate sets of experiments that we labeled "experiment one: series 1, series 2, and series 3" and "experiment two: series 1, series 2, series 3." The three series (Table S9) are similar to Bateman's series 1, 2, and 3 (1). Recall that in both Bateman's and our replication, no offspring could be homozygous-dominant at any of the parents' marker loci because each offspring always received a wild-type allele for the marker alleles from the opposite-sex parent (Table S1).

The Replicated Populations. In the present repetition, there were 166 adult virgin females and 166 adult virgin males in 46 small, even-sex ratio populations of either 6 or 10 uniquely marked adults. We recorded the phenotype of all adult offspring (n = 8,093). Twice daily for 14 d, beginning on the day of first eclosion, we collected offspring from culture bottles. We sexed, genotyped, and scored individual flies while they were under CO_2 anesthetization. Interaction between maker genes sometimes affected our ability to assign parents from the expressed phenotypes of offspring. As Bateman reported, he found it difficult or impossible to unambiguously phenotype some offspring. It is a curiosity of Bateman's experiment that he used as marker mutations in the same population, but in opposite sexes, markers that affected the same characters and thereby handicapped his ability, as it did ours, to unambiguously assign mothers and fathers to some offspring who might have simultaneously expressed phenotypes that would be obscured by the other parental marker. In the present repetition, the interaction between Mc and Me or Pm caused identification for eye color in double-mutant offspring to be coded incorrectly, because some of offspring were completely eyeless so that that they would be scored as

single mutants rather than double mutants. We were able to identify offspring with the Mc:Pm genotype/phenotype, but it was impossible to identify Mc:Me genotype/phenotype because most of offspring were eyeless. It is unknown how often this happened in Bateman's original study and we have no way to estimate his error rate.

Monogamy Trials. To study variation in the frequencies of offspring phenotypes in the absence of sexual selection, we also performed a monogamy experiment in which we placed males and females (each a dominant heterozygote at a unique marker locus but homozygous wild-type at their partner's marker locus) in pairs (Table S14). We crossed 3-d-old flies (one ♀ and one ♂) in a vial (five replicates per combination). We held the flies as pairs for 3 d, after which we discarded all of the males and transferred individual females into new vials daily for 8 d. We counted all flies hatching from individual vials for 5 d and scored the phenotype of each individual offspring. Offspring phenotypes could have included offspring with a mutation from each parent (in which case we would score the offspring as $M^{♀}M^{♂}$ and specifically with an indicator of the mutant from mother and the mutant from father; for example, HB), a mutation from only one parent (e.g., Hw or wB), or wild-type from both parents (ww). We then compared offspring mutant phenotypes in the 25 sets of monogamous pairs with those occurring in the subset of populations that included five females and five males with the same marker mutations as in the monogamous pairs.

Tests of Marker Neutrality. To test if the marker genes were unbiased and neutral with respect to our questions about the V_{NM} and the V_{RS}, we characterized all offspring as having a dominant marker gene from mother ($M^{♀}$), a wild-type gene at mother's marker locus ($w^{♀}$), a dominant gene at father's marker locus ($M^{♂}$), or a wild-type gene at father's marker locus ($w^{♂}$). That is, we binned each offspring in general terms $M^{♀}M^{♀}$, $M^{♀}w^{♂}$, $w^{♀}M^{♂}$, $w^{♀}w^{♂}$. Neither stochastic nor sexually selected effects on number of mates can create deviations in the expected Mendelian frequencies of offspring characterized in terms of their inheritance of dominant or wild-type alleles from each parent (*SI Text* and Tables S6–S8). Assuming no viability effects on offspring who inherited both parents' marker mutations, offspring must occur in the following frequencies: 25% $M^{♀}M^{♀}$, 25% $M^{♀}w^{♂}$, 25% $w^{♀}M^{♂}$, and 25% $w^{♀}w^{♂}$. Data and tests for all populations in our series may be found in Table S13.

Tables S9–S12 show the frequency distributions of offspring genotypes and mutant combination phenotypes in the 48 populations we studied.

We used JMP to perform contingency analyses and produce figures. We set a priori significance at $\alpha \leq 0.05$.

What We Did Not Do. We did not provide tables of "observed" matings and reproductive success similar to Bateman's (1) or an analysis of the relationship between NM and RS or of sex differences in V_{RS} because we showed that the assumption of no viability effects of Bateman's methodology was violated, rendering the measurements of NM and RS unreliable.

ACKNOWLEDGMENTS. We thank Malin Ah-King, Margaret Anderson, Dan Blumstein, John Byers, Sergio Castrezana, Don Dewsbury, Lee C. Drickamer, Brant Faircloth, Mark Friedman, Greg Grether, Thierry Hoquet, Therese Markow, Steve Hubbell, Allen Moore, Sahotra Sarkar, Zuleyma Tang-Martinez, and Judy Stamps for comments and useful discussions; and Kyungsun "Sun" Kim for her work in carrying out the experiments. National Science Foundation Grants IBN-9631801, IBN-0911606, and IOS-1121797 each provided partial support for the study.

1. Bateman AJ (1948) Intra-sexual selection in *Drosophila*. *Heredity (Edinb)* 2:349–368.
2. Darwin C (1871) *The Descent of Man and Selection in Relation to Sex* (John Murray, London).
3. Trivers RL (1972) Parental investment and sexual selection. *Sexual Selection and the Descent of Man*, ed Campbell B (Aldine, Chicago, IL), pp 136–179.
4. Snyder BF, Gowaty PA (2007) A reappraisal of Bateman's classic study of intrasexual selection. *Evolution* 61:2457–2468.
5. Arnold SJ (1994) Bateman principles and the measurement of sexual selection in plants and animals. *Am Nat* 144:S126–S149.
6. Jones AG, Arguello JR, Arnold SJ (2002) Validation of Bateman's principles: A genetic study of sexual selection and mating patterns in the rough-skinned newt. *Proc Biol Sci* 269:2533–2539.
7. Jones AG, Rosenqvist G, Berglund A, Arnold SJ, Avise JC (2000) The Bateman gradient and the cause of sexual selection in a sex-role-reversed pipefish. *Proc Biol Sci* 267:677–680.
8. Dewsbury DA (2005) The Darwin-Bateman paradigm in historical context. *Integr Comp Biol* 45:831–837.
9. Tang-Martinez Z, Ryder TB (2005) The problem with paradigms: Bateman's worldview as a case study. *Integr Comp Biol* 45:821–830.
10. Sutherland WJ (1985) Chance can produce sex differences in mating success and explain Batean's data. *Anim Behav* 33:134–1352.
11. Hubbell SP, Johnson LK (1987) Environmental variance in lifetime mating success, mate choice, and sexual selection. *Am Nat* 130:91–112.
12. Gowaty PA, Hubbell SP (2009) Reproductive decisions under ecological constraints: It's about time. *Proc Natl Acad Sci USA* 106(Suppl 1):10017–10024.
13. Marshall TC, Slate J, Kruuk LE, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. *Mol Ecol* 7:639–655.
14. Weinberg W (1908) Über den Nachweis der Vererbung beim Menschen [On proof of inheritance in humans]. *Jahreshefte des Vereins für vaterländische Naturkunde in Württemberg* 64:368–382.
15. Hardy GH (1908) Mendelian proportions in a mixed population. *Science* 28:49–50.
16. Knight J (2002) Sexual stereotypes. *Nature* 415:254–256.

Supporting Information

Gowaty et al. 10.1073/pnas.1207851109

SI Text

Prior Criticisms of Presentation and Interpretation of Results in Bateman (1948). An important prior methodological concern is that Bateman (1) did not observe behavior (2) and therefore was able to make only a minimum estimate of number of mates (NM) from those offspring that expressed marker mutations from each parent, the $M^{\varphi}M^{\sigma}$ offspring. Dewsbury (2) pointed out that later authors misquoted Bateman when they said that he had observed behavior. Because Bateman did not observe behavior, he was unable to directly evaluate whether or how frequently females chose males or males chose females; or how frequently males fought or females fought. His conclusions that females were choosy but males indiscriminate and competitive were, in fact, speculations (2). Bateman showed two curves comparing sex differences in the effects of NM on reproductive success (RS) (see figures 1 a and b of ref. 1), one showing that multiple mates did not increase female RS above what females achieve with one mate (which is most frequently republished in textbooks) (3); and the other showing female RS increasing with NM. Graphing Bateman's data without arbitrary partitioning showed that male and female RS increased with NM (4). Bateman may have had the first evidence for an advantage for females of multiple insemination (4), but he apparently did not recognize it.

Long before more recent criticisms, Sutherland (5) and Hubbell and Johnson (6) challenged Bateman's study with models of random mating under stochastic demography. Despite assuming random mating and no sexual selection, the two models produced sex differences in variance in number of mates (V_{NM}) under assumptions of chance effects on individual survival and individual encounters with potential mates only because of stochastic effects on demography. These models resulted in V_{NM} that were not statistically different from Bateman's (4, 5) results, necessarily lessening the force of Bateman's arguments about the adaptive significance of V_{NM} (4) and emphasizing that more information than Bateman had is necessary to infer sexual selection. Therefore, two important explanations for Bateman's data have long existed: sexual selection and stochastic demography.

An even more fundamental criticism of Bateman's study lies is the possibility that despite its theoretical elegance, the methodology was flawed. Bateman's method assumes that there are no viability deficits for offspring inheriting marker alleles, either from one parent or both parents simultaneously (4). Examination of one of Bateman's tables (1), which listed the mutations he used for adult markers genes and their effects on the viability of carriers, indicates that most of the marker genes were "homozygous lethal" (4). Each of Bateman's adults was heterozygous at its "marker" locus (Table S1), so each subject carried only one mutant allele (i.e., was not homozygous and would have been expected to survive the period of the experiments). Importantly, because each adult carried markers at different loci (Table S1), none of the offspring in Bateman's experiment could be homozygous for any of the unique parental markers, and thus none could have been homozygous-lethal at any locus. However, as consideration of Tables S6–S8 demonstrate, one-quarter of the offspring, even under mating biases produced by sexual selection, should be "double-mutants, $M^{\varphi}M^{\sigma}$" that inherited a marker mutation at each of its parents' marker loci.

If the frequency of $M^{\varphi}M^{\sigma}$ offspring was less than one-quarter, a critical assumption about the reliability of Bateman's method would be violated. Did double-mutant offspring suffer reduced viability or lethal effects from carrying two parental markers? For example, would an offspring survive who inherited the dominant allele at the "microcephalus, Mc" locus from its mother and the dominant allele at "hairless, H" locus from its father, so that it had both "reduced or absent eyes" and "missing hairs from its head and other body parts"? If such double mutant, $M^{\varphi}M^{\sigma}$ offspring died before eclosion, matings between Mc females and H males would be invisible using Bateman's method. This invisibility of some matings would have confounded mating success with offspring viability, producing inaccurate estimates of NM and V_{NM}, because it was only from double-mutant offspring, $M^{\varphi}M^{\sigma}$, that Bateman could have inferred who mated with whom.

Neither Stochastic Nor Sexually Selected Effects Can Explain Deficits in $M^{\varphi}M^{\sigma}$ Offspring. Consider the necessary frequencies if one male, carrying a unique marker mutation, failed to mate with any females: his mutation would be entirely absent from any offspring, and the number of cells in the matrix (Table S6) would be reduced from 36 to 24. Thus, the expected frequency of offspring in each cell would then be 100/24 = 4.1667, equal to 25% $M^{\varphi}M^{\sigma}$, 25% $M^{\varphi}w^{\sigma}$, 25% $w^{\varphi}M^{\sigma}$, and 25% $w^{\varphi}w^{\sigma}$. Table S7 shows another scenario of differential mating (via sexual selection or demographic stochasticity) in which if one male, again with a unique marker unlike any other males', failed to mate and also one female, also carrying a unique marker unlike any other females' markers, failed to mate, the number of cells in the matrix would be reduced to 16 and the expected frequency of each cell would be 100/16 = 6.25%, equal to 25% $M^{\varphi}M^{\sigma}$, 25% $M^{\varphi}w^{\sigma}$, 25% $w^{\varphi}M^{\sigma}$, and 25% $w^{\varphi}w^{\sigma}$. Table S8 displays expectations when a male and a female mated with others but failed to mate with each other: still the frequency of $M^{\varphi}M^{\sigma}$, $M^{\varphi}w^{\sigma}$, $w^{\varphi}M^{\sigma}$, and $w^{\varphi}w^{\sigma}$ would be 25%. Even if sexual selection (or demographic stochasticity) occurred, the frequency of $M^{\varphi}M^{\sigma}$, $M^{\varphi}w^{\sigma}$, $w^{\varphi}M^{\sigma}$, and $w^{\varphi}w^{\sigma}$ offspring must remain 25%, unless there is differential offspring survival associated with inheriting double mutations or nonindependent assortment of the marker mutations. If the frequencies of $M^{\varphi}M^{\sigma}$, $M^{\varphi}w^{\sigma}$, $w^{\varphi}M^{\sigma}$, and $w^{\varphi}w^{\sigma}$ were each 25%, one would conclude that there were no viability effects on offspring of inheriting two marker mutations. If one wants to use the $M^{\varphi}M^{\sigma}$ offspring to make unbiased estimates of who mated with whom, evaluation of $M^{\varphi}M^{\sigma}$, $M^{\varphi}w^{\sigma}$, $w^{\varphi}M^{\sigma}$, and $w^{\varphi}w^{\sigma}$ frequencies is an essential preliminary exercise toward demonstrating the presence or absence of potential observation biases, as it is only from $M^{\varphi}M^{\sigma}$ offspring that NM and V_{NM} can be inferred.

The data from Bateman's table 4 (see table 4 in ref. 1) (Table S4) reported a single experimental population of three adult males and three adult females that produced 459 offspring in total, 86 (18.7%) of whom were $M^{\varphi}M^{\sigma}$. For offspring for whom fathers could be identified, 77 had fathers heterozygous for the dominant mutation "hairless, H," 105 had fathers heterozygous for the dominant mutation "plum, Pm," and 29 had fathers heterozygous for the dominant mutation "Stubble, Sb," suggesting that Pm fathers had an advantage over the other two types of fathers. For offspring for whom mothers could be identified, 79 had mothers heterozygous for the mutation "Curly wing, Cy," 90 had mothers heterozygous for the mutation "Curly lobed, CyL," and 55 had mothers heterozygous for the mutation "Microcephalous, Mc." One-hundred and ten offspring were $w^{\varphi}w^{\sigma}$ for which it was impossible to identify either parent. One-hundred and twenty-five offspring were $w^{\varphi}M^{\sigma}$; 138 offspring were Mw. Mothers were reliably assigned to 224 offspring (48.8%), fathers to 211 (46.6%) offspring. $M^{\varphi}M^{\sigma}$s were significantly less frequent than required to avoid observation bias in NM and V_{NM}, at least from the only population for which Bateman published all observations. If all of

Bateman's populations had similar deficits in $M^{\varphi}M^{\sigma}$ offspring, his experiment would have produced unreliable results. Snyder and Gowaty (4) thus asked whether the other population trials in Bateman's experiment also had deficits in $M^{\varphi}M^{\sigma}$s. This question might have occurred to other readers aware of the fact that offspring expressing double-mutant phenotypes are often less viable than heterozygous offspring with only one mutation. Thus, we also wondered whether or not a repetition of Bateman's experiment would show experiment-wide inviability effects on double mutant offspring.

1. Bateman AJ (1948) Intra-sexual selection in *Drosophila*. *Heredity (Edinb)* 2: 349–368.
2. Dewsbury DA (2005) The Darwin-Bateman paradigm in historical context. *Integr Comp Biol* 45:831–837.
3. Tang-Martinez Z, Ryder TB (2005) The problem with paradigms: Bateman's worldview as a case study. *Integr Comp Biol* 45:821–830.
4. Snyder BF, Gowaty PA (2007) A reappraisal of Bateman's classic study of intrasexual selection. *Evolution* 61:2457–2468.
5. Sutherland WJ (1985) Chance can produce sex differences in mating success and explain Batean's data. *Anim Behav* 33:134–1352.
6. Hubbell SP, Johnson LK (1987) Environmental variance in lifetime mating success, mate choice, and sexual selection. *Am Nat* 130:91–112.

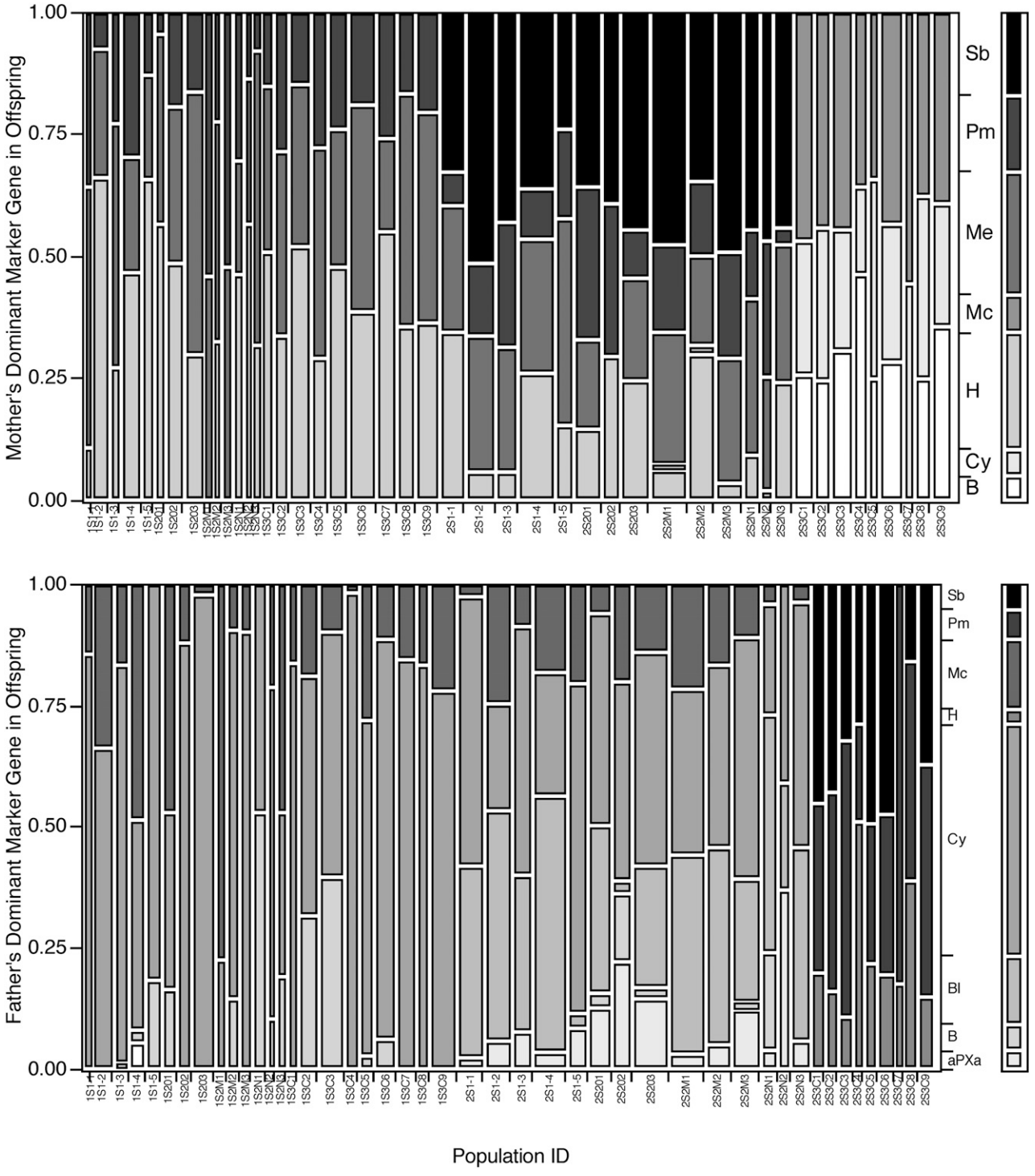

Fig. S1. The distribution of parental marker genes in offspring for each replicated population by sex of parent. (*Upper*) Mothers' marker genes in offspring; (*Lower*) fathers' marker genes in offspring. The width of the bars varies by population, because the number of adult offspring varied by population. For some populations the number of adults was three females and three males; in other populations the number of adults was five females and five males (see Table S9 for details of markers in females and in males as well as population size for each population and Tables S10–S12 for contingency analyses by population). Note that over all populations there were 10 markers for males and 10 for females (see Table 1 for a description of parental marker genes), some of which never appeared in adult offspring, an obervation consistent with possible failure of adults carrying some markers to mate at all or more likely to differential viability of offspring inheriting deleterious alleles from their parents as indicated in Fig. 1 *A*–*C*, or to both effects.

Table S1. Parental genotypes across all marker loci, showing that each subject adult was genetically and phenotypically distinct

	Adult genotypes (two alleles) at each marker locus					
Adults	Sb	Pm	H	CyL^4	Cy	Mc
$♂_1$	Sb+	++	++	++	++	++
$♂_2$	++	Pm+	++	++	++	++
$♂_3$	++	++	H+	++	++	++
$♀_1$	++	++	++	CyL^4+	++	++
$♀_2$	++	++	++	++	Cy+	++
$♀_3$	++	++	++	++	++	Mc+

Observed offspring phenotypes from a single population trial in Bateman's original paper (see table 4 in ref. 1), illustrate his method of parental assignments. ++ indicates that the alleles were wild-type. Note that we used "w" for wild-type in tables showing generalizations associated with inheriting a parental marker or the wild-type allele.

Table S2. Expected offspring genotypes/phenotypes under the assumption that all adults mated with each other and there were no inviability effects of markers on offspring

Possible paternal alleles inherited by offspring	Possible maternal alleles inherited by offspring					
	CyL^4	Cy	Mc	+ from CyL^4	+ from Cy	+ from Mc
Sb	CyL^4Sb	CySb	McSb	+Sb	+Sb	+Sb
Pm	CyL^4Pm	CyPm	McPm	+Pm	+Pm	+Pm
H	CyL^4H	CyH	McH	+H	+H	+H
+ from Sb	CyL^4+	Cy+	Mc+	++	++	++
+ from Pm	CyL^4+	Cy+	Mc+	++	++	++
+ from H	CyL^4+	Cy+	Mc+	++	++	++

++ indicates wild-type genes. In later tables that generalize genetic details to marker presence or absence, we use "w" to indicate wild-type.

Table S3. Expected frequency of offspring in four general phenotypic classes

	Maternal alleles ♀					
Paternal♂ alleles	$CyL^{4♀}$	Cy♀	Mc♀	w♀ from CyL^4	w♀ from Cy	w♀ from Mc
Sb♂	M♀M♂	M♀M♂	M♀M♂	w♀ M♂	w♀ M♂	w♀ M♂
Pm♂	M♀M♂	M♀M♂	M♀M♂	w♀ M♂	w♀ M♂	w♀ M♂
H♂	M♀M♂	M♀M♂	M♀M♂	w♀ M♂	w♀ M♂	w♀ M♂
w♂ from Sb	M♀w♂	M♀w♂	M♀ w♂	w♀ w♂	w♀ w♂	w♀ w♂
w♂ from Pm	M♀ w♂	M♀ w♂	M♀ w♂	w♀ w♂	w♀ w♂	w♀ w♂
w♂ from H	M♀ w♂	M♀ w♂	M♀ w♂	w♀ w♂	w♀ w♂	w♀ w♂

Those inheriting two markers, one from each parent M♀M♂, a marker (M) at one parent's marker locus and the wild-type (w) allele at the marker locus of the other parent (either M♀w♂ or w♀M♂), or wild-type alleles at each parent's marker locus w♀w♂. Assuming no viability effects on offspring of inherited mutations, the expected frequency of offspring in each cell would be 100/36 = 2.7778%, giving 25% double mutants (M♀M♂); 25% single mutants with the mutant allele from mother and a wild-type allele from father (M♀w♂); 25% single mutants with a mutant allele from father and a wild-type allele from mother (w♀M♂); and 25% with a wild-type allele from father and a wild-type allele from mother (w♀w♂). Note that neither stochastic demography nor sexual selection can alter the expected frequencies of M♀M♂, M♀w♂, w♀M♂, and w♀w♂.

Table S4. Subject genotypes and observed offspring genotypes as reported in Bateman's (1) paper

	Possible maternal alleles							
Possible paternal alleles	CyL^4	Cy	Mc	+ from CyL^4	+ from Cy	+ from Mc	No. of mates per ♂	Assigned RS for each ♂♂
Sb	13	0	0		16		1	29
Pm	10	12	15		68		3	105
H	7	29	0		41		2	77
+ from Sb								
+ from Pm								
+ from H	60	38	40		110			
No. of mates per ♀	3	2	1					RS ♂♂ = 211
Assigned RS for each ♀♀	90	79	55					RS ♀♀ = 224

Table S5. Observed frequency of $M^{♀}M^{♂}$, $M^{♀}w^{♂}$, $w^{♀}M^{♂}$, and $w^{♀}w^{♂}$ calculated from Bateman's original paper (1) show a deficit in $M^{♀}M^{♂}$ indicating inviability of offspring

	Offspring in marker classes by presence or absence of maternal marker					
Possible paternal alleles	CyL^4	Cy	Mc	w from CyL^4	w from Cy	w from Mc
Sb	$M^{♀}M^{♂}$ = 86 (18.7%)			$w^{♀}M^{♂}$ = 125 (27.2%)		
Pm						
H						
w from Sb	$M^{♀}w^{♂}$ = 138 (30%)			$w^{♀}w^{♂}$ = 110 (24%)		
w from Pm						
w from H						

Table S6. Expected offspring genotypes if one male, Pm, failed to mate

	Possible maternal alleles					
Possible paternal alleles	CyL^4	Cy	Mc	+ from CyL^4	+ from Cy	+ from Mc
Sb	CyL^4 Sb	Cy Sb	Mc Sb	+ Sb	+ Sb	+ Sb
Pm						
H	CyL^4 H	Cy H	Mc H	+ H	+ H	+ H
+ from Sb	CyL^4 +	Cy +	Mc +	+ +	+ +	+ +
+ from Pm						
+ from H	CyL^4 +	Cy +	Mc +	+ +	+ +	+ +

Expected frequencies of offspring genotypes from which the frequencies of offspring classes, $M^{♀}M^{♂}$, $M^{♀}w^{♂}$, $w^{♀}M^{♂}$, and $w^{♀}w^{♂}$ can be inferred (see also Tables S1–S5). ++ indicates a wild-type allele, generalized to "ww" when translating the absence of a parental marker allele into the generalized classes in which M indicates the presence of a parental marker mutation or "w" the absence of a parental marker. Table S6 should be compared with results from Tables S2, S7, and S8. In all cases, the frequencies of offspring inheriting marker mutations from each parent, or from only one parent, or from neither parent are the same.

Table S7. Expected offspring genotypes if one female, Cy, and one male, Pm, failed to mate

	Possible maternal alleles					
Possible paternal alleles	CyL^4	Cy	Mc	w from CyL^4	w from Cy	w from Mc
Sb	CyL^4 Sb		Mc Sb	+ Sb		+ Sb
Pm						
H	CyL^4 H		Mc H	+ H		+ H
w from Sb	CyL^4 +		Mc +	+ +		+ +
w from Pm						
w from H	CyL^4 +		Mc +	+ +		+ +

Expected frequencies of offspring genotypes from which the frequencies of offspring classes, $M^{♀}M^{♂}$, $M^{♀}w^{♂}$, $w^{♀}M^{♂}$, and $w^{♀}w^{♂}$ can be inferred (see also Tables S1–S5). ++ indicates a wild-type allele, generalized to "ww" when translating the absence of a parental marker allele into the generalized classes in which M indicates the presence of a parental marker mutation or "w" the absence of a parental marker. Table S7 should be compared with results from Table S2, S6, and S8. In all cases, the frequencies of offspring inheriting marker mutations from each parent, or from only one parent, or from neither parent are the same.

Table S8. Expected offspring genotypes if one male, H, failed to mate with one female, Mc

	Possible maternal alleles					
Possible paternal alleles	CyL^4	Cy	Mc	w from CyL^4	w from Cy	w from Mc
Sb	CyL^4 Sb	Cy Sb	Mc Sb	+ Sb	+Sb	+Sb
Pm	CyL^4 Pm	Cy Pm	Mc Pm	+ Pm	+Pm	+Pm
H	CyL^4 H	Cy H		+ H	+H	
w from Sb	CyL^4 +	Cy +	Mc +	+ +	+ +	+ +
w from Pm	CyL^4 +	Cy +	Mc +	+ +	+ +	+ +
w from H	CyL^4 +	Cy +		+ +	+ +	

Expected frequencies of offspring genotypes from which the frequencies of offspring classes, $M^{♀}M^{♂}$, $M^{♀}w^{♂}$, $w^{♀}M^{♂}$, and $w^{♀}w^{♂}$ can be inferred (see also Tables S1–S5). ++ indicates a wild-type allele, generalized to "ww" when translating the absence of a parental marker allele into the generalized classes in which M indicates the presence of a parental marker mutation or "w" the absence of a parental marker. Table S8 should be compared with results from Table S2, S6, and S7. In all cases, the frequencies of offspring inheriting marker mutations from each parent, or from only one parent, or from neither parent are the same.

Table S9. Characteristics of populations in the present replication

Experiment, series, replicates, population IDs	Mutants female × male	No. of adults by sex, each with a different marker mutation	Duration of mating opportunity (d)	Ages of adults	No. of offspring produced and phenotyped
	♀♀ Pm, H, Me	3	3	Mixed age	
1-S1-1	♂♂ B, Cy, Mc				72
1-S1-2					195
1-S1-3					118
1-S1-4					184
1-S1-5					129
	♀♀ Pm, H, Me	3	3	♀♀ ♂♂	
	♂♂ B, Cy, Mc			1 1	136
1-S2N1				1 1	155
1-S2N2				1 1	186
1-S2N3				3 3	90
1-S2M1				3 3	121
1-S2M2				3 3	99
1-S2M3				6 6	123
1-S2O1				6 6	102
1-S2O2				6 6	92
1-S2O3					
	♀♀ Pm, H, Me	3	4	♀♀ ♂♂	
1-S3C1	♂♂ B, Cy, Mc			1 1	121
1-S3C2				3 1	168
1-S3C3				6 1	236
1-S3C4				1 3	154
1-S3C5				3 3	158
1-S3C6				6 3	231
1-S3C7				1 6	197
1-S3C8				3 6	154
1-S3C9				6 6	239
	♀♀ Hw, Pm, Sb, H, Mé	5	3	Mixed age	
2-S1-1	♂♂ B, Cy, ap^{Xa}, Bl, Mc				227
2-S1-2					278
2-S1-3					192
2-S1-4					332
2-S1-5					185
	♀♀ Hw, Pm, Sb, H, Mé	5	3	♀♀ ♂♂	
2-S2N1	♂♂ B, Cy, ap^{Xa}, Bl, Mc			1 1	254
2-S2N2				1 1	190
2-S2N3				1 1	326
2-S2M1				3 3	372
2-S2M2				3 3	246
2-S2M3				3 3	255
2-S2O1				6 6	165
2-S2O2				6 6	103
2-S2O3				6 6	194
	♀♀ B, Cy, Mc	3	4	♀♀ ♂♂	
2-S3C1	♂♂ Pm, H, Sb			1 1	166
2-S3C2				3 1	123
2-S3C3				6 1	169
2-S3C4				1 3	126
2-S3C5				3 3	134
2-S3C6				6 3	190
2-S3C7				1 6	85
2-S3C8				3 6	138
2-S3C9				6 6	183

Table S10. Population ID by offspring genotype mother's marker locus allele then father's locus allele

Count row %	++	+H	+Pm	+Sb	B+	BH	BPm	BSb	Cy+	CyH	CyPm	CySb	Mc+	McH	McPm	McSb	Row total
2S3C1	45	8	17	15	18	0	0	3	13	2	3	4	30	2	1	5	166
	27.11	4.82	10.24	9.04	10.84	0.00	0.00	1.81	7.83	1.20	1.81	2.41	18.07	1.20	0.60	3.01	
2S3C2	30	5	15	12	11	1	2	1	11	0	6	2	15	3	0	9	123
	24.39	4.07	12.20	9.76	8.94	0.81	1.63	0.81	8.94	0.00	4.88	1.63	12.20	2.44	0.00	7.32	
2S3C3	47	4	21	9	12	0	7	8	11	3	6	2	34	0	3	2	169
	27.81	2.37	12.43	5.33	7.10	0.00	4.14	4.73	6.51	1.78	3.55	1.18	20.12	0.00	1.78	1.18	
2S3C4	40	17	5	8	18	3	3	2	7	0	1	2	16	3	0	1	126
	31.75	13.49	3.97	6.35	14.29	2.38	2.38	1.59	5.56	0.00	0.79	1.59	12.70	2.38	0.00	0.79	
2S3C5	45	10	11	24	6	1	4	0	13	1	2	2	11	1	0	3	134
	33.58	7.46	8.21	17.91	4.48	0.75	2.99	0.00	9.70	0.75	1.49	1.49	8.21	0.75	0.00	2.24	
2S3C6	50	9	14	25	17	2	7	0	18	2	3	3	29	2	1	8	190
	26.32	4.74	7.37	13.16	8.95	1.05	3.68	0.00	9.47	1.05	1.58	1.58	15.26	1.05	0.53	4.21	
2S3C7	18	5	17	0	0	0	0	0	10	0	10	0	23	1	1	0	85
	21.18	5.88	20.00	0.00	0.00	0.00	0.00	0.00	11.76	0.00	11.76	0.00	27.06	1.18	1.18	0.00	
2S3C8	34	13	23	4	10	1	4	1	12	7	0	5	18	4	2	0	138
	24.64	9.42	16.67	2.90	7.25	0.72	2.90	0.72	8.70	5.07	0.00	3.62	13.04	2.90	1.45	0.00	
2S3C9	54	6	26	13	17	0	3	10	12	5	4	0	27	0	2	4	183
	29.51	3.28	14.21	7.10	9.29	0.00	1.64	5.46	6.56	2.73	2.19	0.00	14.75	0.00	1.09	2.19	
	363	77	149	110	109	8	30	25	107	20	35	20	203	16	10	32	1,314

Number and frequency of offspring in populations with mutant sets of subjects: Females = B, Cy, Mc; Males = Pm, H, Sb.

Table S11. Population ID by offspring genotype mother's marker locus allele then father's locus allele

Count row %	++	+ ap^{Xa}	+B	+Bl	+ Cy	+ Mc	H+	H ap^{Xa}	HB	H Bl	H Cy	H Mc	Hw +	Mé +	Mé ap^{Xa}	Mé B	Mé Bl	Mé Cy	Pm+	Pm ap^{Xa}	Pm B	Pm Bl	Pm Cy	Pm Mc	Sb+	Sb ap^{Xa}	Sb B	Sb Bl	Sb Cy	Sb Mc	Row total
2S1-1	55	2	0	31	35	3	19	0	0	6	10	0	0	16	0	0	7	3	5	0	0	1	1	0	18	1	0	0	14	0	227
	24.23	0.88	0.00	13.66	15.42	1.32	8.37	0.00	0.00	2.64	4.41	0.00	0.00	7.05	0.00	0.00	3.08	1.32	2.20	0.00	0.00	0.44	0.44	0.00	7.93	0.44	0.00	0.00	6.17	0.00	
2S1-2	90	4	0	32	9	22	0	0	0	1	0	6	0	21	1	0	0	12	11	2	0	0	5	0	38	0	0	23	0	1	278
	32.37	1.44	0.00	11.51	3.24	7.91	0.00	0.00	0.00	0.36	0.00	2.16	0.00	7.55	0.36	0.00	0.00	4.32	3.96	0.72	0.00	0.00	1.80	0.00	13.67	0.00	0.00	8.27	0.00	0.36	
2S1-3	48	6	0	14	32	6	4	0	0	0	1	0	0	16	0	0	4	2	12	0	0	10	0	0	20	1	0	2	12	2	192
	25.00	3.13	0.00	7.29	16.67	3.13	2.08	0.00	0.00	0.00	0.52	0.00	0.00	8.33	0.00	0.00	2.08	1.04	6.25	0.00	0.00	5.21	0.00	0.00	10.42	0.52	0.00	1.04	6.25	1.04	
2S1-4	93	2	0	58	23	14	22	0	0	0	4	11	0	25	0	0	10	4	7	2	0	3	2	1	40	1	0	6	4	0	332
	28.01	0.60	0.00	17.47	6.93	4.22	6.63	0.00	0.00	0.00	1.20	3.31	0.00	7.53	0.00	0.00	3.01	1.20	2.11	0.60	0.00	0.90	0.60	0.30	12.05	0.30	0.00	1.81	1.20	0.00	
2S1-5	52	1	0	2	40	19	8	1	0	0	2	0	0	12	4	0	0	14	10	1	0	0	2	0	9	1	0	1	6	0	185
	28.11	0.54	0.00	1.08	21.62	10.27	4.32	0.54	0.00	0.00	1.08	0.00	0.00	6.49	2.16	0.00	0.00	7.57	5.41	0.54	0.00	0.00	1.08	0.00	4.86	0.54	0.00	0.54	3.24	0.00	
2S201	88	6	1	16	31	3	11	1	2	0	0	2	0	11	1	0	0	8	18	1	0	14	1	0	25	4	0	5	4	1	254
	34.65	2.36	0.39	6.30	12.20	1.18	4.33	0.39	0.79	0.00	0.00	0.79	0.00	4.33	0.39	0.00	0.00	3.15	7.09	0.39	0.00	5.51	0.39	0.00	9.84	1.57	0.00	1.97	1.57	0.39	
2S202	64	7	7	1	22	15	12	3	4	0	3	0	0	0	0	0	0	0	15	0	0	0	8	0	14	9	1	1	2	2	190
	33.68	3.68	3.68	0.53	11.58	7.89	6.32	1.58	2.11	0.00	1.58	0.00	0.00	0.00	0.00	0.00	0.00	0.00	7.89	0.00	0.00	0.00	4.21	0.00	7.37	4.74	0.53	0.53	1.05	1.05	
2S203	103	15	2	29	45	22	17	6	0	4	0	0	0	12	2	0	2	7	9	0	2	0	0	0	26	0	0	5	18	0	326
	31.60	4.60	0.61	8.90	13.80	6.75	5.21	1.84	0.00	1.23	0.00	0.00	0.00	3.68	0.61	0.00	0.61	2.15	2.76	0.00	0.61	0.00	0.00	0.00	7.98	0.00	0.00	1.53	5.52	0.00	
2S2M1	118	3	0	43	35	26	5	1	0	3	0	0	2	27	0	0	2	11	17	1	0	7	1	0	50	0	0	8	5	7	372
	31.72	0.81	0.00	11.56	9.41	6.99	1.34	0.27	0.00	0.81	0.00	0.00	0.54	7.26	0.00	0.00	0.54	2.96	4.57	0.27	0.00	1.88	0.27	0.00	13.44	0.00	0.00	2.15	1.34	1.88	
2S2M2	63	4	0	31	29	15	16	0	0	11	0	4	2	12	0	0	1	6	10	2	0	0	4	0	27	0	0	4	5	0	246
	25.61	1.63	0.00	12.60	11.79	6.10	6.50	0.00	0.00	4.47	0.00	1.63	0.81	4.88	0.00	0.00	0.41	2.44	4.07	0.81	0.00	0.00	1.63	0.00	10.98	0.00	0.00	1.63	2.03	0.00	
2S2M3	67	11	1	20	41	5	3	0	1	0	0	0	0	18	3	0	5	2	14	1	0	6	3	0	31	0	0	0	15	8	255
	26.27	4.31	0.39	7.84	16.08	1.96	1.18	0.00	0.39	0.00	0.00	0.00	0.00	7.06	1.18	0.00	1.96	0.78	5.49	0.39	0.00	2.35	1.18	0.00	12.16	0.00	0.00	0.00	5.88	3.14	
2S2N1	53	2	14	17	13	1	1	1	0	1	1	2	0	13	0	1	7	0	7	0	0	2	0	0	16	0	0	10	3	0	165
	32.12	1.21	8.48	10.30	7.88	0.61	0.61	0.61	0.00	0.61	0.61	1.21	0.00	7.88	0.00	0.61	4.24	0.00	4.24	0.00	0.00	1.21	0.00	0.00	9.70	0.00	0.00	6.06	1.82	0.00	
2S2N2	27	13	0	6	10	0	0	1	0	0	0	0	0	5	6	0	0	0	10	0	0	0	3	0	7	0	0	6	9	0	103
	26.21	12.62	0.00	5.83	9.71	0.00	0.00	0.97	0.00	0.00	0.00	0.00	0.00	4.85	5.83	0.00	0.00	0.00	9.71	0.00	0.00	0.00	2.91	0.00	6.80	0.00	0.00	5.83	8.74	0.00	
2S2N3	62	3	0	15	30	2	11	1	0	7	0	1	0	13	0	0	8	2	2	0	0	0	1	0	19	1	0	5	11	0	194
	31.96	1.55	0.00	7.73	15.46	1.03	5.67	0.52	0.00	3.61	0.00	0.52	0.00	6.70	0.00	0.00	4.12	1.03	1.03	0.00	0.00	0.00	0.52	0.00	9.79	0.52	0.00	2.58	5.67	0.00	
	983	79	25	315	395	153	129	15	7	33	21	26	4	201	17	1	46	71	147	10	2	43	31	1	340	18	1	76	108	21	3319

Number and frequency of offspring in populations with mutant sets of subjects: Females = Hw, Pm, Sb, H, Mé; Males = B, Cy, ap^{Xa}, Bl, Mc.

Table S12. Population ID by offspring genotype mother's marker locus allele then father's locus allele

Count row %	++	+B	+Cy	+Mc	H+	HB	HCy	HMc	Mé +	Mé B	Mé Cy	Mé Mc	Pm+	PmB	PmCy	Row total
1S1-1	20	0	19	5	1	0	2	0	9	0	6	0	7	0	3	72
	27.78	0.00	26.39	6.94	1.39	0.00	2.78	0.00	12.50	0.00	8.33	0.00	9.72	0.00	4.17	
1S1-2	60	0	36	31	23	0	22	0	15	0	3	0	5	0	0	195
	30.77	0.00	18.46	15.90	11.79	0.00	11.28	0.00	7.69	0.00	1.54	0.00	2.56	0.00	0.00	
1S1-3	29	1	33	11	5	0	7	0	13	0	9	0	4	0	6	118
	24.58	0.85	27.97	9.32	4.24	0.00	5.93	0.00	11.02	0.00	7.63	0.00	3.39	0.00	5.08	
1S1-4	56	1	18	28	26	1	3	8	15	0	4	0	17	0	7	184
	30.43	0.54	9.78	15.22	14.13	0.54	1.63	4.35	8.15	0.00	2.17	0.00	9.24	0.00	3.80	
1S1-5	40	8	34	0	18	2	11	0	5	0	5	0	1	2	3	129
	31.01	6.20	26.36	0.00	13.95	1.55	8.53	0.00	3.88	0.00	3.88	0.00	0.78	1.55	2.33	
1S201	40	11	14	27	13	0	8	4	16	0	1	0	1	0	1	136
	29.41	8.09	10.29	19.85	9.56	0.00	5.88	2.94	11.76	0.00	0.74	0.00	0.74	0.00	0.74	
1S202	40	0	37	6	28	0	5	2	14	0	9	0	7	0	7	155
	25.81	0.00	23.87	3.87	18.06	0.00	3.23	1.29	9.03	0.00	5.81	0.00	4.52	0.00	4.52	
1S203	47	0	57	2	11	0	13	0	22	0	21	0	9	0	4	186
	25.27	0.00	30.65	1.08	5.91	0.00	6.99	0.00	11.83	0.00	11.29	0.00	4.84	0.00	2.15	
1S2M1	16	0	7	41	0	0	0	0	12	0	0	0	9	0	5	90
	17.78	0.00	7.78	45.56	0.00	0.00	0.00	0.00	13.33	0.00	0.00	0.00	10.00	0.00	5.56	
1S2M2	42	6	28	5	6	2	5	0	14	0	4	0	5	0	4	121
	34.71	4.96	23.14	4.13	4.96	1.65	4.13	0.00	11.57	0.00	3.31	0.00	4.13	0.00	3.31	
1S2M3	22	0	28	5	0	0	0	0	13	0	8	0	12	0	11	99
	22.22	0.00	28.28	5.05	0.00	0.00	0.00	0.00	13.13	0.00	8.08	0.00	12.12	0.00	11.11	
1S2N1	31	25	24	0	9	9	2	0	9	0	1	0	10	0	3	123
	25.20	20.33	19.51	0.00	7.32	7.32	1.63	0.00	7.32	0.00	0.81	0.00	8.13	0.00	2.44	
1S2N2	34	2	21	8	16	2	3	0	10	0	1	0	4	0	1	102
	33.33	1.96	20.59	7.84	15.69	1.96	2.94	0.00	9.80	0.00	0.98	0.00	3.92	0.00	0.98	
1S2N3	29	8	8	22	7	1	0	0	7	0	8	0	2	0	0	92
	31.52	8.70	8.70	23.91	7.61	1.09	0.00	0.00	7.61	0.00	8.70	0.00	2.17	0.00	0.00	
1S3C1	37	0	23	8	17	0	10	0	13	0	5	0	5	0	3	121
	30.58	0.00	19.01	6.61	14.05	0.00	8.26	0.00	10.74	0.00	4.13	0.00	4.13	0.00	2.48	
1S3C2	54	19	26	13	7	8	1	3	9	0	12	0	13	0	3	168
	32.14	11.31	15.48	7.74	4.17	4.76	0.60	1.79	5.36	0.00	7.14	0.00	7.74	0.00	1.79	
1S3C3	70	39	28	11	21	0	25	0	22	6	1	0	10	0	3	236
	29.66	16.53	11.86	4.66	8.90	0.00	10.59	0.00	9.32	2.54	0.42	0.00	4.24	0.00	1.27	
1S3C4	50	0	38	1	13	0	6	0	21	0	7	0	11	0	7	154
	32.47	0.00	24.68	0.65	8.44	0.00	3.90	0.00	13.64	0.00	4.55	0.00	7.14	0.00	4.55	
1S3C5	46	0	26	11	20	2	6	8	13	0	8	0	11	0	7	158
	29.11	0.00	16.46	6.96	12.66	1.27	3.80	5.06	8.23	0.00	5.06	0.00	6.96	0.00	4.43	
1S3C6	62	1	48	9	27	5	10	1	30	0	16	1	15	0	6	231
	26.84	0.43	20.78	3.90	11.69	2.16	4.33	0.43	12.99	0.00	6.93	0.43	6.49	0.00	2.60	
1S3C7	57	0	49	13	29	0	14	0	11	0	4	0	15	0	5	197
	28.93	0.00	24.87	6.60	14.72	0.00	7.11	0.00	5.58	0.00	2.03	0.00	7.61	0.00	2.54	

Number and frequency of offspring in populations with mutant sets of subjects: Females = Pm, H, Mé; Males = B, Cy, Mc.

Table S13. Population ID by offspring marker phenotypes

Count row %	M♀M♂	M♀w♂	w♀M♂	w♀w♂	Row total
1S1-1	11	17	24	20	72
	15.28	23.61	33.33	27.78	
1S1-2	25	43	67	60	195
	12.82	22.05	34.36	30.77	
1S1-3	22	22	45	29	118
	18.64	18.64	38.14	24.58	
1S1-4	23	58	47	56	184
	12.50	31.52	25.54	30.43	
1S1-5	23	24	42	40	129
	17.83	18.60	32.56	31.01	
1S201	14	30	52	40	136
	10.29	22.06	38.24	29.41	
1S202	23	49	43	40	155
	14.84	31.61	27.74	25.81	
1S203	38	42	59	47	186
	20.43	22.58	31.72	25.27	
1S2M1	5	21	48	16	90
	5.56	23.33	53.33	17.78	
1S2M2	15	25	39	42	121
	12.40	20.66	32.23	34.71	
1S2M3	19	25	33	22	99
	19.19	25.25	33.33	22.22	
1S2N1	15	28	49	31	123
	12.20	22.76	39.84	25.20	
1S2N2	7	30	31	34	102
	6.86	29.41	30.39	33.33	
1S2N3	9	16	38	29	92
	9.78	17.39	41.30	31.52	
1S3C1	18	35	31	37	121
	14.88	28.93	25.62	30.58	
1S3C2	27	29	58	54	168
	16.07	17.26	34.52	32.14	
1S3C3	35	53	78	70	236
	14.83	22.46	33.05	29.66	
1S3C4	20	45	39	50	154
	12.99	29.22	25.32	32.47	
1S3C5	31	44	37	46	158
	19.62	27.85	23.42	29.11	
1S3C6	39	72	58	62	231
	16.88	31.17	25.11	26.84	
1S3C7	23	55	62	57	197
	11.68	27.92	31.47	28.93	
1S3C8	23	50	32	49	154
	14.94	32.47	20.78	31.82	
1S3C9	33	55	85	66	239
	13.81	23.01	35.56	27.62	
2S1-1	43	58	71	55	227
	18.94	25.55	31.28	24.23	
2S1-2	51	70	67	90	278
	18.35	25.18	24.10	32.37	
2S1-3	34	52	58	48	192
	17.71	27.08	30.21	25.00	
2S1-4	48	94	97	93	332
	14.46	28.31	29.22	28.01	
2S1-5	32	39	62	52	185
	17.30	21.08	33.51	28.11	
2S201	44	65	57	88	254
	17.32	25.59	22.44	34.65	
2S202	33	41	52	64	190
	17.37	21.58	27.37	33.68	

Table S13. Cont.

Count row %	$M^{♀}M^{♂}$	$M^{♀}w^{♂}$	$w^{♀}M^{♂}$	$w^{♀}w^{♂}$	Row total
2S203	46	64	113	103	326
	14.11	19.63	34.66	31.60	
2S2M1	46	101	107	118	372
	12.37	27.15	28.76	31.72	
2S2M2	37	67	79	63	246
	15.04	27.24	32.11	25.61	
2S2M3	44	66	78	67	255
	17.25	25.88	30.59	26.27	
2S2N1	28	37	47	53	165
	16.97	22.42	28.48	32.12	
2S2N2	25	22	29	27	103
	24.27	21.36	28.16	26.21	
2S2N3	37	45	50	62	194
	19.07	23.20	25.77	31.96	
2S3C1	20	61	40	45	166
	12.05	36.75	24.10	27.11	
2S3C2	24	37	32	30	123
	19.51	30.08	26.02	24.39	
2S3C3	31	57	34	47	169
	18.34	33.73	20.12	27.81	
2S3C4	15	41	30	40	126
	11.90	32.54	23.81	31.75	
2S3C5	14	30	45	45	134
	10.45	22.39	33.58	33.58	
2S3C6	28	64	48	50	190
	14.74	33.68	25.26	26.32	
2S3C7	12	33	22	18	85
	14.12	38.82	25.88	21.18	
2S3C8	24	40	40	34	138
	17.39	28.99	28.99	24.64	
2S3C9	28	56	45	54	183
	15.30	30.60	24.59	29.51	
Column total	1242	2108	2400	2343	8,093
	15.35	26.05	29.66	28.95	

From the current repetition of Bateman's (1) experiment, the total number and percent of offspring inheriting mutations from both parents ($M^{♀}M^{♂}$), only mother ($M^{♀}w^{♂}$), only father ($w^{♀}M^{♂}$), or neither parent ($w^{♀}w^{♂}$). Each cell contains the count, plus the row percentage. The last column shows the number of offspring eclosed and phenotyped from each population. The last row shows the overall counts of offspring marker phenotype and total sample size. There are significantly more $w^{♀}w^{♂}$ and $w^{♀}M^{♂}$ offspring (both 29%) and significantly fewer $M^{♀}M^{♂}$ offspring (15%) than the one-quarter expected by chance (likelihood ratio $\chi^2 = 463.1$, $df = 3$; $P < 0.0001$). The greatest contribution to χ^2 was from the cell with $M^{♀}M^{♂}$ offspring, indicating that offspring who inherited a marker allele from each parent had reduced viability.

Table S14. Total number and percent of offspring from the monogamy trials

Count row %	$M^{♀}M^{♂}$	$M^{♀}w^{♂}$	$w^{♀}M^{♂}$	$w^{♀}w^{♂}$	Row total
H × ap^{Xa}	55	33	68	82	238
	23.11	13.87	28.57	34.45	
H × B	148	104	173	227	652
	22.70	15.95	26.53	34.82	
H × Bl	169	197	223	217	806
	20.97	24.44	27.67	26.92	
H × Cy	79	88	102	127	396
	19.95	22.22	25.76	32.07	
H × Mc	98	101	115	153	467
	20.99	21.63	24.63	32.76	
Hw × ap^{Xa}	42	66	92	87	287
	14.63	23.00	32.06	30.31	
Hw × B	113	13	135	107	368
	30.71	3.53	36.68	29.08	
Hw × Bl	50	47	75	115	287
	17.42	16.38	26.13	40.07	
Hw × Cy	91	104	171	187	553
	16.46	18.81	30.92	33.82	
Hw × Mc	96	86	165	147	494
	19.43	17.41	33.40	29.76	
Mé × ap^{Xa}	79	104	117	115	415
	19.04	25.06	28.19	27.71	
Mé × B	122	108	187	174	591
	20.64	18.27	31.64	29.44	
Mé × Bl	166	122	146	148	582
	28.52	20.96	25.09	25.43	
Mé × Cy	111	146	149	124	530
	20.94	27.55	28.11	23.40	
Mé × Mc	0	128	266	145	539
	0.00	23.75	49.35	26.90	
Pm × ap^{Xa}	60	84	96	87	327
	18.35	25.69	29.36	26.61	
Pm × B	67	51	96	92	306
	21.90	16.67	31.37	30.07	
Pm × Bl	59	73	70	81	283
	20.85	25.80	24.73	28.62	
Pm × Cy	56	66	71	81	274
	20.44	24.09	25.91	29.56	
Pm × Mc	30	64	85	78	257
	11.67	24.90	33.07	30.35	
Sb × B	130	135	170	143	578
	22.49	23.36	29.41	24.74	
Sb × ap^{Xa}	116	109	115	124	464
	25.00	23.49	24.78	26.72	
Sb × Bl	181	190	158	185	714
	25.35	26.61	22.13	25.91	
Sb × Cy	171	243	209	199	822
	20.80	29.56	25.43	24.21	
Sb × Mc	151	214	199	211	775
	19.48	27.61	25.68	27.23	6.46
Column Total	2440	2676	3453	3436	12,005
	20.32	22.29	28.76	28.62	

From the monogamy trials, the total number and percent of offspring inheriting each of their parents' marker alleles ($M^{♀}M^{♂}$), a marker gene from their mother and wild-type from father ($M^{♀}w^{♂}$), a wild-type allele from their mother and a marker gene from their father ($w^{♀}M^{♂}$), or a wild-type allele at each parental marker locus. We ran five replicates of each set (25 × 5) of monogamous parents.

Table S15. Methods of Bateman (1) compared with the present replication

Method	Bateman	Present replication
Culture of specific genotypic lines of *D. melanogaster* to produce subjects: dominant heterozygous individuals each bearing unique markers at different loci.	See Bateman (pp. 354 and 355 in ref. 1).	We obtained mutants from several *Drosophila* laboratories in February 2007 and allowed them to breed in our laboratory for at least four generations before extracting heterozygotes. See *Methods.*
Duration of experiment.	Not indicated in Bateman's paper, but probably 2–3 y.	Began 4/1/07 and ended 9/30/07.
Holding potential subjects in same-sex jars to guarantee that no copulations occurred before establishment of replicate populations.	See Bateman (p. 356 in ref. 1).	We held same-sex subjects bearing the same marker genes in small milk jars until establishing experimental populations of virgins.
Establishment of small replicate populations.	"Several flies of each sex were mated together in one bottle, each fly carrying a different dominant marker gene" (p. 355 in ref. 1). 64 populations from six series.	We placed virgins in small milk bottles in combinations of either 3 ♀♀ and 3 ♂♂ or 5 ♀♀ and 5 ♂♂, each constituting a replicate population. We replicated only series 1, 2, 3 of Bateman's original design. n = 46 populations.
Populations (replicated trials) were in "Series" in which (*i*) the sex of subjects bearing particular sets of marker mutations, (*ii*) number of subjects of each sex, (*iii*) number of days in which subjects could mate, and (*iv*) the age of subjects entering a population were held constant.	See Bateman (table 3, in ref. 1).	See Table S9.
Removal of subject adults from bottles, also called the "mating period" in Bateman (1).	Not mentioned in Bateman, but implied by his noted "duration of matings", i.e., the number of days in which adult subjects could mate.	After the 3- or 4-d duration of "opportunity to mate," we removed and disposed of all adult subjects.
Identification of parents when offspring were double mutants, so that they expressed two marker alleles.	"...not always possible" Bateman (see p. 355 in ref. 1)	When parental marker mutations affected the same character, say the eye, it was indeed sometimes impossible to tell if the offspring carried both parental markers. See *Methods*.
Number of days after first eclosion in which offspring were counted and phenotyped.	Not indicated in Bateman.	We collected all offspring eclosing in experimental bottles for 14 d beginning with the first day pupae eclosed
Recording of observations of offspring phenotypes.	Not described in Bateman.	We examined each offspring individually on the day it eclosed and recorded its phenotype on data sheets as well as a unique ID for each individual (in each population).
Data summaries from trial populations.	Not described in Bateman, but most likely by sorting all offspring per bottle by phenotype and tabulating results by hand	We used routines in JMP to summarize offspring phenotypes into categories of double mutants $M^{♀}M^{♂}$, single mutants $M^{♀}w^{♂}$, $w^{♀}w^{♂}$, and no mutants $w^{♀}w^{♂}$. These summaries are in Table S13.
Statistical analyses.	ANOVA to test effect of markers, marker combinations, age, population duration, and sex on number of mates and number of offspring. Bateman combined trials within series. Bateman only partially tested the assumption of offspring viability.	We began by testing for inviability effects in offspring. We analyzed each population separately because of our interest in within-population, within- and between-sex (i.e., individual) variation in NM, V_{NM}, RS, and V_{NM}. We used χ^2 to test for deviations for expected Mendelian frequencies of parental phenotypes in offspring, and t tests of difference scores to test if Bateman's methodology produced systematic biases in the number of offspring scored as having mothers versus those having fathers.
Monogamous pairs.	Bateman did not perform monogamy trials.	See *Methods.*