PROCEEDINGS B

rspb.royalsocietypublishing.org

Research

Cite this article: Sharp NP, Agrawal AF. 2018 An experimental test of the mutation-selection balance model for the maintenance of genetic variance in fitness components. *Proc. R. Soc. B* **285**: 20181864.
http://dx.doi.org/10.1098/rspb.2018.1864

Received: 19 August 2018
Accepted: 17 October 2018

Subject Category:
Evolution

Subject Areas:
evolution, genetics

Keywords:
mutation accumulation, standing variance, sex-specific selection, sexual antagonism, life-history traits

Author for correspondence:
Nathaniel P. Sharp
e-mail: sharp@zoology.ubc.ca

Electronic supplementary material is available online at http://dx.doi.org/10.6084/m9.figshare.c.4278797.

THE ROYAL SOCIETY PUBLISHING

An experimental test of the mutation-selection balance model for the maintenance of genetic variance in fitness components

Nathaniel P. Sharp[1,2] and Aneil F. Agrawal[2]

[1]Department of Zoology, University of British Columbia, Vancouver, Canada V6T 1Z4
[2]Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Canada

NPS, 0000-0001-6016-083X

Despite decades of research, the factors that maintain genetic variation for fitness are poorly understood. It is unclear what fraction of the variance in a typical fitness component can be explained by mutation-selection balance (MSB) and whether fitness components differ in this respect. In theory, the level of standing variance in fitness due to MSB can be predicted using the rate of fitness decline under mutation accumulation, and this prediction can be directly compared to the standing variance observed. This approach allows for controlled statistical tests of the sufficiency of the MSB model, and could be used to identify traits or populations where genetic variance is maintained by other factors. For example, some traits may be influenced by sexually antagonistic balancing selection, resulting in an excess of standing variance beyond that generated by deleterious mutations. We describe the underlying theory and use it to test the MSB model for three traits in *Drosophila melanogaster*. We find evidence for differences among traits, with MSB being sufficient to explain genetic variance in larval viability but not male mating success or female fecundity. Our results are consistent with balancing selection on sexual fitness components, and demonstrate the feasibility of rigorous statistical tests of the MSB model.

1. Introduction

The maintenance of genetic variation for fitness has been called one of the most important unresolved issues in evolutionary biology [1,2]. Despite the action of natural selection, populations harbour significant genetic variation for fitness-related traits [3]. A number of possible sources of variation exist, including deleterious mutations, beneficial alleles on their way to fixation and balancing selection, including environmental heterogeneity in selection. The relative importance of these factors is unknown, and will determine how genomes and populations evolve [4].

Deleterious mutation-selection balance (MSB) is arguably the most general explanation for genetic variance in fitness because all populations experience mutation. The question is whether MSB alone can account for the standing variation. Answering this question requires knowing how much variation would be expected from MSB alone. In a population at equilibrium, the rate at which mutation reduces mean fitness will equal the rate at which selection increases mean fitness, which in turn is equal to the genetic variance in fitness [5]. Thus, under MSB the standing genetic variance in fitness should be equal to the rate of fitness decline per generation in a mutation-accumulation experiment (mutational decline). This idea is described more formally below, and can be found in various incarnations in the literature [6–10]. If there is any other source of genetic variation apart from deleterious mutations, then the observed level of standing variation will exceed the rate of mutational decline.

Although estimates of mutational decline and standing variance exist for a number of traits (particularly in *D. melanogaster*), they are generally estimated in different experimental populations with different genetic backgrounds using different methodologies, which is far from ideal, especially given the considerable variation in some quantitative genetic parameters among populations [8,11]. In addition, the effects of new mutations are often assessed in the homozygous state; because new mutations will rarely be found in the homozygous state in a randomly mating population, it is their heterozygous effects that are most relevant. Homozygous mutational decline can only be used to infer the heterozygous equivalent after making tenuous assumptions about average dominance.

Despite these challenges, several attempts have been made to qualitatively test alternative models for the maintenance of genetic variation using data from numerous sources [7–9,11]. These investigations point to an important role of deleterious mutation in maintaining variation, but they cannot statistically evaluate the adequacy of the MSB model. While these studies have provided valuable insights, their conclusions are often limited by the quality and quantity of available data, leading these authors and others to call for experiments that directly address the MSB hypothesis [12–14]. Here, we outline some theoretical and empirical issues relevant to such experiments and use this approach to study the sources of standing genetic variation in a *D. melanogaster* laboratory population.

Populations in the laboratory, which are generally maintained in constant conditions for many generations, may be unlikely to harbour certain kinds of genetic variation because heterogeneity in selection will be minimal, and beneficial mutations on their way to fixation will be rare if the population is well adapted. (Moreover, there is reason to believe that beneficial alleles that eventually become fixed may make a relatively small contribution to standing variance [8].) However, although the environment may be relatively constant, alleles in sexual populations will be expressed in both males and females, and may be subject to intra-locus sexually antagonistic (SA) selection. This type of balancing selection may be a particularly important source of variation in well-adapted populations [15].

Although the existence of SA alleles has been established, it remains an empirical challenge to quantify their contribution to the standing genetic variance in fitness, relative to other sources of variation. If alleles under SA balancing selection are common, deleterious mutation should account for a smaller fraction of the genetic variance in sex-specific fitness components than in non-sex-specific fitness components. In other words, SA is expected to generate 'excess' genetic variance beyond that attributable to deleterious mutations. To test this idea, we collected data on standing variance and the rate of change due to mutation accumulation (MA) for larval viability, male mating success, and female fecundity. We studied heterozygous second chromosomes (approx. 37% of the genome) on a common isogenic background, and measured mutational decline and standing variance in the same way within each fitness component.

2. Theoretical background

The theoretical background motivating this experiment has been described by others [6–10], and we present a summary of the major points here. We present an additive model formulation below; an analogous multiplicative formulation where traits are measured on a log scale is given in the electronic supplementary material.

In a large panmictic population, the equilibrium frequency of a deleterious mutation at a given locus will be $q^* \approx \mu/(hs)$, where μ is the mutation rate and hs is the coefficient of selection against heterozygotes [16]. This assumes that $\mu \ll hs$ so $q^* \ll 1$; thus, the frequency of heterozygotes is approximately $2q^*$ and the frequency of homozygous mutants is negligible. If the non-mutant value of trait z is k, and the heterozygous trait value is $k(1-a)$, at equilibrium the expected trait value is $\bar{z} = k(1-2a(\mu/hs))$. We assume that the effect of a mutation on trait z is some fraction c of its effect on total fitness, i.e. $a = chs$, so that $\bar{z} = k(1-2\mu c)$.

Consider the additive genetic variance in trait values relative to the trait mean, $\sigma_z^2 = V[z/\bar{z}]$, also called 'evolvability' [3]. If mutation is the only source of variation, at equilibrium σ_z^2 at one locus will be $\sigma_z^2 \cong 2\mu ac$, ignoring terms of $O(\mu^2)$. Summing over n loci in the genome, if the variance at each site is small and assuming there is no epistasis [9] and no covariance between the mutation rate and the mutational effect, the total standing genetic variance in relative z is $\sigma_z^2 \cong UE[ac]$ where U is the mutation rate per genome. When the trait is fitness itself, $c = 1$, $a = hs$, and $\sigma_z^2 \cong UE[hs]$. Note this is equal to the rate of change in mean relative fitness due to one generation of mutation in the absence of selection, ΔM_w, which can be estimated in an MA experiment [17]. This result implies that, when the effects of mutation and selection are at equilibrium, the rate of decline in fitness due to mutation must equal the rate of increase in fitness due to selection, which is given by the additive genetic variance in fitness [5], or, for a trait, the additive genetic covariance between the trait and fitness [18]. A corresponding result is that deleterious mutations contribute little in the way of dominance variance, relative to additive variance, so that most genetic variance will be additive under MSB ([10], pp. 184–185). Under the null hypothesis that deleterious mutations completely account for standing variation, we expect no 'excess' variance, i.e. $\eta_w \equiv \sigma_w^2 - \Delta M_w = 0$. Thus $\eta_w > 0$ is an indication that there is more standing genetic variation than can be explained by MSB alone.

For traits that are components of fitness, standing variance will also depend on c:

$$\sigma_z^2 \cong UE[ac] = UE[a]E[c] + UC[a,c] = \Delta M_z \bar{c} + UC[a,c],$$

where $C[a,c]$ is the covariance between a and c. While ΔM_z can be estimated directly in an MA experiment, the other terms are more difficult to determine. Observations of positive mutational correlations among fitness components (e.g. [19–22]) suggest pleiotropy is generally positive, such that the average deleterious mutation will have a smaller effect on a single fitness component than on total fitness, i.e. $\bar{c} \in [0,1]$, although this is not necessarily the case for all loci. Positive pleiotropy also suggests that $C[a,c]$ will be positive and not too large, i.e. mutations with greater effects on a given fitness component will tend to have greater effects on total fitness. Therefore, ΔM_z alone would overestimate σ_z^2 because $\bar{c} < 1$, but would underestimate σ_z^2 because $UC[a,c] > 0$. However, given realistic values the first bias will tend to be larger, such that ΔM_z will tend to overestimate σ_z^2, particularly when $\bar{c}$ is not very close to 1, which is likely to

be the case; a formal analysis of these biases is given in [8]. This bias means that the rate of mutational decline will tend to overestimate the expected variance in a given trait under MSB because the frequency of mutations will be reduced through selection on additional fitness components that are not accounted for by ΔM_z (i.e. $\bar{c} < 1$). The test of MSB for a fitness component, $\eta_z \equiv \sigma_z^2 - \Delta M_z = 0$, will therefore be conservative.

It may be possible to reduce this conservative bias by estimating $\bar{c}$. If j components of fitness are measured, a very simple estimate of $\bar{c}$ for trait z is the mutational decline in the trait relative to the mutational decline in all fitness components, i.e. $\bar{c}_z \sim \Delta M_z / \Sigma_j \Delta M_j$. This assumes that fitness components are multiplicative but that mutational effects on individual traits are sufficiently small that the total fitness effect can be approximated via the sum across all fitness components. Even when all fitness components have been measured, this approach can over- or underestimate $\bar{c}_z$ (because $E[a]/E[s] \neq E[a/s]$). If an important fitness component has not been measured, $\Sigma_j \Delta M_j < \Delta M_w$, and so $\Delta M_z / \Sigma_j \Delta M_j$ will tend to overestimate $\bar{c}_z$. Again, this would lead to a conservative test of $\eta'_z \equiv \sigma_z^2 - \Delta M_z \bar{c}_z = 0$. We obtained estimates of σ_z^2 and ΔM_z for three fitness components (traits), as described below to test the MSB hypothesis.

3. Material and methods

(a) Overview

Our goal was to estimate the rate of mutational decline, ΔM_z, and the standing additive genetic variance in relative trait values, $\sigma_z^2 = V[z]/\bar{z}^2$, for each of three major fitness components: viability, female fecundity, and male mating success. We examined each trait in a number of MA lines, in their corresponding controls, and in an outbred laboratory population (Dahomey). In each case, focal second chromosomes were tested in the heterozygous state on a common isogenic background, allowing us to compare traits and sources of variance while minimizing background effects. All reported estimates refer to haploid second chromosomes. Details of line preparation and trait measurement protocols are given in the electronic supplementary material, and crossing procedures are shown in electronic supplementary material, figures S1 and S2.

(b) Experimental lines

We measured ΔM_z in lines that accumulated mutations for 52 generations. The MA procedure used to generate these lines is described elsewhere [23]. Briefly, an initially isogenic focal second chromosome marked with *bw* was replicated into many independent MA lines. Each MA line was maintained by crossing a single heterozygous male to four outbred stock females each generation, thereby maximizing the amount of drift and rendering selection ineffective; marker alleles were used to identify the relevant chromosomes (electronic supplementary material, figure S1A,B). The same initial chromosome was also maintained in three separate control populations of 450 adults each. This moderate population size should limit MA. Following MA, crosses were performed to situate the focal second chromosomes from 51 MA lines and 57 control chromosomes on an isogenic background (electronic supplementary material, figure S1C; any MA lines with evidence of recombination during MA [24] were not included in this study). For each line, we assessed each trait in six replicates for a total of approximately 678 replicates per trait. The lines used in this experiment are inferred to contain an average of 18.6 mutations each [24]. In the homozygous state, these MA chromosomes cause significantly reduced viability [23] and adult reproductive fitness [20].

We measured σ_z^2 among second chromosomes derived from virgin females collected at random from the outbred laboratory population. Crosses were performed to situate each chromosome on the same isogenic background described above (electronic supplementary material, figure S2A), minimizing among-line selection and ensuring that each line ultimately contained only a single focal chromosome haplotype. In addition, second chromosomes remained heterozygous throughout the experiment, minimizing selection against genotypes bearing recessive lethal or deleterious alleles. For each of 133 such lines, we assessed each trait in six replicates, for a total of approximately 798 replicates per trait.

(c) Trait measurements

Viability was estimated as the probability of survival to adulthood in competition with a standard genotype. Standing variance in viability was assessed in two blocks of 67 and 66 lines respectively. Male mating success was estimated as the proportion of standard females that mated (as assessed via offspring production) with focal males in competition with marked males, under male-biased sex ratio conditions, in a period of time that did not allow females to mate multiply. These assays were performed in two blocks for the MA group, with non-overlapping sets of mutant and control lines in each block (block 1 : 29 control lines, 25 MA lines; block 2 : 28 control lines, 26 MA lines). Similarly, two blocks were performed for the assay of standing variance in this trait, of 67 and 66 lines respectively. Female fecundity was estimated as early-life egg production, where focal females competed with standard females for access to a limited amount of live yeast [25]. Note that our measures of male mating success and female fecundity are independent of larval viability. Details on trait measurements are given in the electronic supplementary material.

(d) Data analysis

To estimate means and variance components on a log scale (see the electronic supplementary material) we fit generalized linear mixed models using *MCMCglmm* [26] in *R* [27], using non-informative priors (variances$\sim$0, $nu = -2$), and a burn-in phase of 10^6 iterations. Subsequent iterations were stored such that the autocorrelation among stored values was less than 0.1 for all model parameters and the final number of stored iterations was 10^4. To test for a difference between two parameters of interest (e.g. ΔM_z and σ_z^2) we sampled without replacement from the posterior distribution of each parameter and determined the 95% credible interval (CI) for the distribution of their differences (e.g. $\eta_z \equiv \sigma_z^2 - \Delta M_z$). We consider 95% CIs that do not include zero to indicate a significant difference.

The response variable for viability was the ratio of focal offspring, n_{focal}, to standard offspring, n_{standard}. When modelled using a binomial link function, the model scale $\text{logit}(n_{\text{focal}}/(n_{\text{focal}} + n_{\text{standard}}))$ is equivalent to $\log(n_{\text{focal}}/n_{\text{standard}})$. The response variable for female fecundity was number of eggs produced, modelled using a Poisson (log) link. The response variable for male mating success was the ratio of the number of females that mated with focal males to the number of females that mated with standard males in each replicate, modelled using a binomial link. Each trait and group of lines (MA lines, MA controls, outbred lines) was modelled separately, with a random effect of genotype (line).

Block did not have a significant effect on viability in outbred lines ($p = 0.78$), and was dropped from that model. Block had a significant effect on male mating success in outbred lines ($p < 1 \times 10^{-5}$), and controls ($p < 0.05$), and a marginally non-

significant effect in MA lines ($p = 0.06$). However, there was no significant difference in genetic variance between blocks in any of these groups (5000 bootstrap replicates, all $p > 0.35$). We, therefore, modelled male mating success with a main effect of block on the intercept only.

We estimated σ_z^2 for each trait as the among-line (genetic) variance among outbred lines on the log scale. We estimated ΔM_z as $(E[y_{\text{control}}] - E[y_{\text{MA}}])/52$, where $y = \log(z)$ and 52 is the number of MA generations, as well as the mutational variance, $\Delta V_z = V[y_{MA}]/52$. The 'excess' variance in a trait beyond that expected under MSB alone was estimated as $\eta_z = \sigma_z^2 - \Delta M_z$. When comparing excess variance across traits we standardized η_z by the expected standing variance, ΔM_z (i.e. $\beta_z = \eta_z/\Delta M_z$). To account for selection on multiple traits, we estimated $E[c_z]$ for each trait as $\bar{c}_z = \Delta M_z/(\Delta M_v + \frac{1}{2}\,\Delta M_m + \frac{1}{2}\,\Delta M_f)$, where the subscripts v, m, and f represent viability, male mating success, and fecundity, respectively, and the coefficients of $\frac{1}{2}$ reflect the sex-limited nature of some fitness components. We then estimated excess variance as $\eta'_z = \sigma_z^2 - \Delta M_z \bar{c}_z$.

We also used *MCMCglmm* to examine the standing and mutational genetic correlations among traits using priors expected to be uninformative for random effect correlations (inverse-Wishart; *G*: variances = 0.02, covariances = 0, *nu* = 4; *R*: variances ~ 0, covariances = 0, *nu* = −2; *MCMCglmm* vignette). Models were conducted with a burn-in phase of 10^6 iterations followed by 10^8 iterations with a thinning interval of 6000, leaving 16 667 values in the posterior distribution such that autocorrelation among stored values was less than 0.1 for all model parameters. Trait correlations were determined after transformation to the original measurement scale by integration. To compare mutational and standing correlations we sampled 10^5 values from each posterior distribution. We also estimated phenotypic correlations among line means between pairs of traits, which will differ from the genetic correlations due to estimation error (attenuation). For male mating success line means were first standardized by subtracting the block mean. Phenotypic correlations were compared between groups by bootstrapping with 10^4 replicates.

(e) Maximum likelihood

Inferences from the analyses above assume the ideal MA control. In reality, the control populations for the MA experiment could potentially evolve over the course of MA. Though purifying selection will prevent deleterious mutations from fixing in control populations, such mutations will appear and segregate at a low level, reducing the average fitness of the controls below their initial value, downwardly biasing our estimates of ΔM. To estimate the most probable initial control values we fit a maximum-likelihood model (see the electronic supplementary material), where both the MA and control trait means and variances depend on the underlying mutation rate and the effects of mutations on each trait. We evaluated the likelihood of a given set of parameters using generalized linear mixed model functions implemented in *lme4* [28], using Nelder-Mead optimization with *bbmle* [29], repeating the optimization 50 times with random starting values.

4. Results

Means and genetic variances for each trait on the original scale of measurement are given in electronic supplementary material, table S1, and estimates of quantitative genetic values are shown in figure 1 and electronic supplementary material, table S2. Our estimates of mutational decline and standing variance are consistent with previously reported values [9,11,17]. For each trait, we examined the posterior

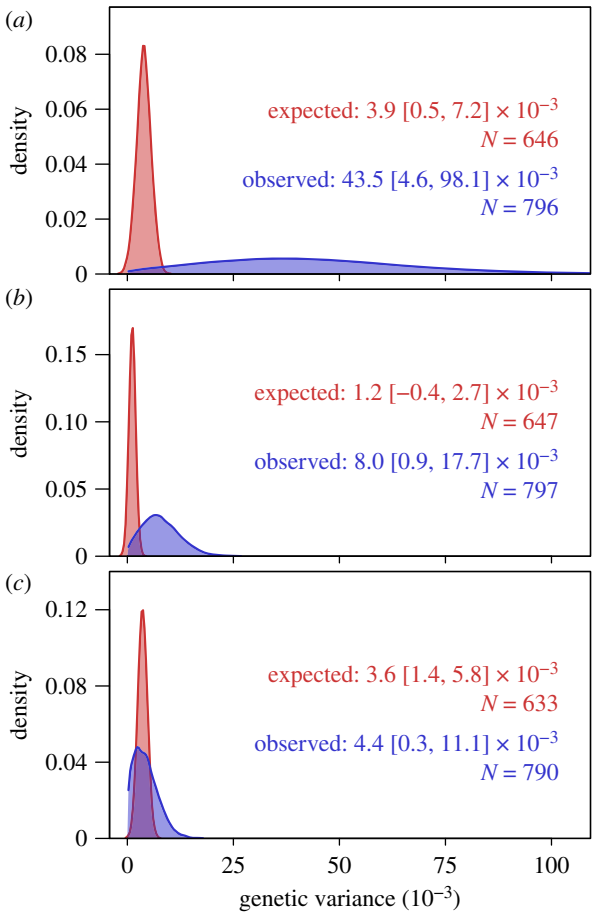

Figure 1. Standardized posterior density distributions for the genetic variance expected under mutation-selection balance (red) and the genetic variance observed (blue) for male mating success (*a*), female fecundity (*b*), and larval viability (*c*). The text in each panel describes the posterior means and 95% credible intervals based on quantiles; sample sizes (*N*) refer to total replicate trait measurements across either 133 outbred lines or 108 MA/control lines (51 MA lines plus 57 control lines).

distribution of observed standing variances and the variance expected under MSB (figure 1). We scale excess variance η_z to mean expected variance $E[\Delta M_z]$ to facilitate comparison among traits. We cannot reject MSB for larval viability, where excess variance is low ($\eta_l/E[\Delta M_l]$: mean 0.22, CI [−1.17, 2.16]). However, we find evidence for excess variance in male mating success, where the 95% CI does not overlap zero ($\eta_m/E[\Delta M_m]$: mean 10.16, CI [0.11, 24.15]). The 95% CI for female fecundity slightly overlaps zero ($\eta_f/E[\Delta M_f]$: mean 5.81, CI [−0.34, 14.26]) but 96% of the posterior probability mass is associated with excess variance ($\eta_f > 0$). The combined evidence from the adult fitness components indicates significant excess variance (i.e. $\frac{1}{2}\,\eta_m/E[\Delta M_m] + \frac{1}{2}\,\eta_f/E[\Delta M_f] > 0$; mean 7.99, CI [1.78, 16.06]), and the excess variance in the adult fitness components was significantly greater than the excess variance for viability (i.e. $\frac{1}{2}\,\eta_m/E[\Delta M_m] + \frac{1}{2}\,\eta_f/E[\Delta M_f] - \eta_v/E[\Delta M_v] > 0$; mean 7.77, CI [1.30, 16.08]). Our point estimates indicate that MSB explains only 9% and 15% of the standing genetic variance in male mating success and female fecundity, respectively, but 82% of the variance in larval viability.

Because our estimates of σ_z^2 are constrained to be non-negative but our estimates of ΔM_z are not, it is possible for this difference to create a bias towards finding positive η_z. ΔM_z represents the expected amount of variance so negative

Table 1. Summary of genetic and phenotypic (line mean) correlation estimates for each pairwise trait combination, for mutation-accumulation (MA) lines and the standing population.

group	traits[a]	genetic correlation [95%CI[b]]	line mean correlation [95%CI[c]]
MA	$m-f$	0.118 [−0.518, 0.679]	0.077 [−0.211, 0.354]
	$m-v$	0.117 [−0.514, 0.674]	0.000 [−0.278, 0.301]
	$f-v$	0.079 [−0.458, 0.577]	−0.045 [−0.339, 0.245]
standing	$m-f$	0.172 [−0.348, 0.606]	0.197 [0.045, 0.328]
	$m-v$	0.182 [−0.300, 0.598]	0.096 [−0.107, 0.329]
	$f-v$	0.278 [−0.108, 0.600]	0.206 [0.036, 0.368]

[a]Traits are viability (v), male mating success (m), and female fecundity (f).
[b]Bayesian credible interval.
[c]Confidence interval.

values are arguably non-sensical. Repeating the analysis with negative ΔM_z values replaced with zero or removed does not qualitatively affect the results.

The comparisons above ignore two likely sources of bias, which will act in opposite directions. First, mutations will generally have greater effects on total fitness than their effects on any one fitness component (i.e. $\bar{c} < 1$), which would lead us to underestimate excess variance. Second, the appearance of deleterious mutations in control populations could not be prevented in our experiment, which would lead us to underestimate ΔM and overestimate excess variance. To address these issues we first used maximum likelihood to estimate the original trait values of the MA control genotype (electronic supplementary material, table S3), and used these adjusted control values to re-estimate ΔM for each trait. Next, we used the revised ΔM values to estimate $\bar{c}$ for each trait (finding $\bar{c}_m = 0.68$, $\bar{c}_f = 0.35$, and $\bar{c}_v = 0.49$), and tested the null hypothesis $\eta'_z \equiv \sigma_z^2 - \Delta M_z \bar{c}_z = 0$. The net effect of these adjustments does not substantially alter the results described above, and we obtain similar evidence for excess variance ($\eta'_z / E[\Delta M_z \bar{c}_z] > 0$) in male mating success (mean 7.63, CI [0.14, 18.09]) and female fecundity (mean 2.64, CI [−0.06, 6.36]; 97% of posterior probability mass is associated with excess variance) but not larval viability (mean 0.65, CI [−0.60, 2.46]). There is significant excess variance combining evidence from adult traits (mean 5.14, CI [1.06, 10.67]), and greater amounts of excess variance for adult traits than for viability (mean 4.48, CI [0.04, 10.15]). Our adjusted estimates indicate that MSB explains 12%, 33%, and 86% of the standing genetic variance in male mating success, female fecundity, and larval viability, respectively.

If the presence of excess standing genetic variance in adult sexual traits reflects the presence of SA alleles we would expect to find a negative standing genetic correlation between these traits, or at least that the standing genetic correlation is lower than the mutational correlation. However, the genetic correlations (table 1) are all positive and not significantly different from zero, and there is no evidence of a difference between mutational and standing genetic correlations for any pair of traits (CIs for all differences overlap zero). Although the phenotypic correlations are generally smaller than the genetic correlations (figure 2 and table 1), we detected significant standing correlations between viability and female fecundity ($r = 0.21$, $t = 2.41$, $p < 0.05$), and between male mating success and female fecundity ($r = 0.20$, $t = 2.30$, $p < 0.05$). There is no evidence of a difference between mutational and standing phenotypic correlations for any pair of traits (bootstrapping; $p > 0.40$).

5. Discussion

We tested the hypothesis that genetic variance is maintained by MSB, and our results indicate the presence of additional sources of genetic variance in adult sexual fitness components, but not viability. The strengths of our study are that we examined several major fitness components, measured mutational decline and standing genetic variance using a common genetic background and measurement procedures, and that we avoided assumptions regarding dominance by measuring the fitness effects of mutations in the heterozygous state.

We are not aware of any other study with this combination of features. A partial exception is a study of the nematodes *Caenorhabditis elegans* and *C. briggsae* [30], which examined the ratio of the mutational variance to the standing genetic variance ($\Delta V_z / \sigma_z^2$) for lifetime fitness and body size across populations of each species. This ratio approximates the average strength of selection against heterozygous mutations under MSB [6]. However, if some standing variance is due to non-mutational sources then these values will be downwardly biased, and our results suggest that this could be the case for some traits, particularly male mating success. Huang *et al.* [31] used a similar approach to compare mutational and standing variance in quantitative traits in *D. melanogaster* (e.g. bristle number), and concluded that simple models of mutation-stabilizing-selection balance are insufficient to account for standing variance.

Others have made important contributions to this question by comparing mutational parameters and standing genetic variance from numerous sources, typically experiments using *D. melanogaster*. Houle *et al.* [7] examined the ratio of standing variance to mutational variance, $\sigma_z^2 / \Delta V_z$, which approximates the 'persistence time' of new mutations in a population under MSB [32]. They reasoned that life-history traits, which are under strong directional selection, should show reduced $\sigma_z^2 / \Delta V_z$ relative to metric traits, which tend to be under weaker selection. Consistent with MSB, they found that the persistence time for life-history traits was significantly lower than that of metric traits. Houle *et al.* [7] discuss several challenges they faced,

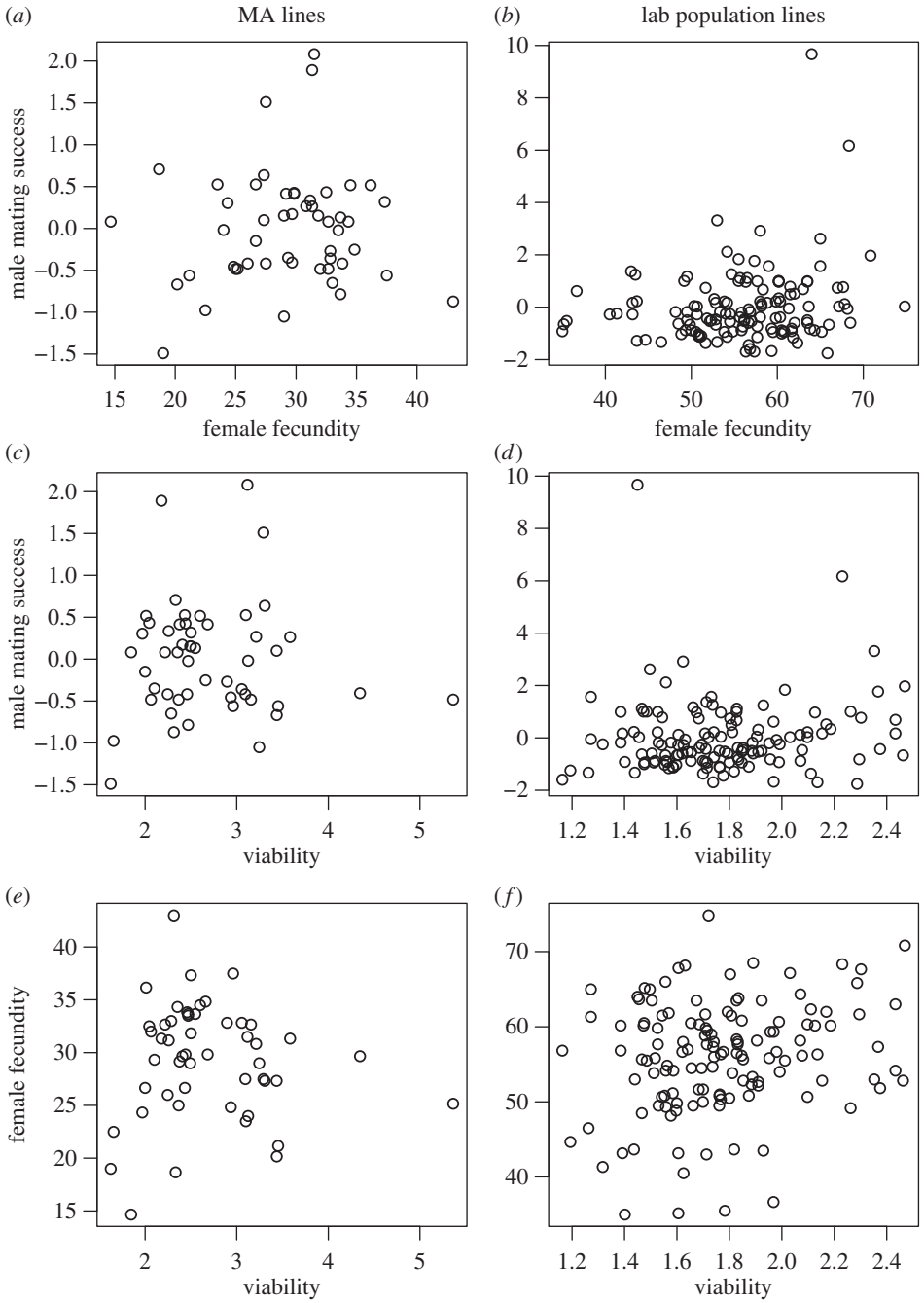

Figure 2. Phenotypic means for MA lines (*a*,*c*,*e*) and standing laboratory population lines (*b*,*d*,*f*), plotted for each pair of traits. Note that values for male mating success are standardized by the group mean within each block.

including the unavailability of estimates of σ_z^2 and ΔV_z from the same population, possible bias resulting from using estimates based on homozygous rather than heterozygous mutations, and the use of standing variance measures from populations potentially not at equilibrium. Along similar lines, Houle [11] found that σ_z^2 and ΔV_z were highly correlated across traits, as expected if σ_z^2 is largely determined by mutational input. However, the necessary standardization of traits may have introduced some autocorrelation to this relationship [11].

Charlesworth & Hughes [9] considered the relationship $\sigma_z^2 = U\overline{hs}\bar{c}_z^2$ expected under MSB (see their equation 19.6), equivalent to the relationship we tested. They used the best available estimates of U, $\overline{hs}$, and $\bar{c}$ from disparate sources to calculate expected σ^2 under MSB. By comparison with observed values of σ^2, they suggested that mutation likely contributes one-third to two-thirds of the genetic variation in a typical life-history trait. Our results (35% of variance explained on average, or 44% after adjustment for bias) are consistent with their conclusion. Using a similar approach, Charlesworth [8] concluded that most *Drosophila* populations show higher standing variance for viability than expected under MSB, but of the traits we studied viability shows the least evidence for excess variance. A strength of the meta-analyses described above [7–9,11] is that averaging over populations and traits may provide a more robust representation of the parameters of interest at the species level. However, this approach might also obscure real variation among traits or populations in the extent of non-mutational genetic variance.

Using a different method based on the effect of artificial selection on the mean value and inbreeding depression in a trait [33], there is evidence that intermediate-frequency alleles contribute to variation in flower size and male fitness in

Mimulus guttatus [34,35] and early fecundity in *D. melanogaster* [14]. Although it provides a qualitative test only, this approach is relatively assumption free, and it would be valuable to apply it to compare more traits and populations. Additional evidence comes from the observation that mutation does not seem to completely explain the effects of inbreeding on life-history traits [9]. Overall, these comparisons indicate that the MSB model may often be insufficient to account for standing genetic variation. The data are inconsistent with heterozygote advantage across all traits, so any excess variance is likely due to balancing selection in the form of antagonistic pleiotropy (leading to net heterozygote advantage), heterogeneity in selection, or frequency-dependent selection [9].

We found evidence of excess variance in our study but no support for sexual antagonism as the cause, though our confidence in the latter result is low due to the uncertainty in each correlation estimate. Perhaps the excess variance in male mating success and female fecundity is due to antagonistic pleiotropy with fitness components not measured in this experiment (e.g. adult survival, sperm competitive ability). Another possible explanation arises from the fact that the conditions of the fitness assays do not represent selection on these fitness components exactly as it occurs in the laboratory population. For example, if the effects of mutations on a given fitness component had a larger effect in the assay conditions than in the laboratory population, this would inflate our estimates of mutational decline and standing variance, but would have a greater effect on standing variance, giving the false impression of excess variation. Our results could, therefore, be explained by our adult fitness assays being more selective than reality and our viability assays being less selective. Our fitness assays seem a reasonable approximation of fitness as it occurs in the laboratory population but these assays are not identical to the conditions under which the population was maintained, and do not consider the effects of multiple mating and postcopulatory sexual selection. Selection on both larval and adult fitness components can be sensitive to assay conditions [25,36]. In this respect, *Drosophila* populations with maintenance regimes that can be more precisely replicated in fitness assays would be preferable (e.g. [19,37]).

Although our data represent one of the more controlled tests to date of MSB in fitness components, several empirical limitations of our study should be considered. First, new mutations that arose in the control populations during the MA procedure cause us to underestimate mutational decline and overestimate excess variance; we attempted to account for this in our maximum-likelihood model but this is not an ideal solution. Second, we chose to measure mutational decline and standing genetic variance in three traits at once, which placed a practical limit on the number of independent genotypes we could manipulate and examine. Nevertheless, we gained some statistical power by combining evidence across traits. Third, we measured standing variation on only an autosomal part of the genome. This was partly done for practical reasons: we wanted to compare standing variation with mutational variation, which arose on the second chromosome in our MA lines, and we wanted to avoid complications due to hemizygosity in males. However, we might have detected stronger departures from MSB if we had included variation on the sex chromosomes, which may be enriched for alleles with SA effects [37,38].

Our findings indicate that, within our particular laboratory population, MSB is inadequate to explain levels of standing genetic variance in some fitness components, but not others. Comparing fitness components in this respect may lead to greater insight into the sources and genetic basis of non-mutational variation.

Data accessibility. Data from this research is available in the Dryad Digital Repository: https://doi.org/10.5061/dryad.2p7qb12 [39].

Authors' contributions. N.S. and A.A. designed experiments, analysed the data, and wrote the manuscript; N.S. performed experiments. All authors gave final approval for publication.

Competing interests. We have no competing interests.

Funding. This research was funded by the Natural Sciences and Engineering Research Council of Canada.

Acknowledgements. We thank S. Otto, M. Whitlock, and members of the Otto laboratory for helpful discussion.

References

1. Barton NH, Keightley PD. 2002 Understanding quantitative genetic variation. *Nat. Rev. Genet.* **3**, 11–21. (doi:10.1038/nrg700)
2. Lynch M, Walsh B. 1998 *Genetics and analysis of quantitative traits*. Sunderland, MA: Sinauer Associates Incorporated.
3. Houle D. 1992 Comparing evolvability and variability of quantitative traits. *Genetics* **130**, 195–204.
4. Mitchell-Olds T, Willis JH, Goldstein DB. 2007 Which evolutionary processes influence natural genetic variation for phenotypic traits? *Nat. Rev. Genet.* **8**, 845–856. (doi:10.1038/nrg2207)
5. Fisher RA. 1930 *The genetical theory of national selection*. Oxford, UK: Oxford University Press.
6. Barton NH. 1990 Pleiotropic models of quantitative variation. *Genetics* **124**, 773–782.
7. Houle D, Morikawa B, Lynch M. 1996 Comparing mutational variabilities. *Genetics* **143**, 1467–1483.
8. Charlesworth B. 2015 Causes of natural variation in fitness: evidence from studies of *Drosophila* populations. *Proc. Natl Acad. Sci. USA* **112**, 1662–1669. (doi:10.1073/pnas.1423275112)
9. Charlesworth B, Hughes KA. 1999 The maintenance of genetic variation in life-history traits. In *Evolutionary genetics: from molecules to morphology* (eds RS Singh, CB Krimbas). Cambridge, UK: Cambridge University Press.
10. Charlesworth B, Charlesworth D. 2012 *Elements of evolutionary genetics*. Greenwood Village, CO: Roberts and Co.
11. Houle D. 1998 How should we explain variation in the genetic variance of traits? *Genetica* **102/103**, 241–253. (doi:10.1023/A:1017034925212)
12. Keightley PD. 1994 The distribution of mutation effects on viability in *Drosophila melanogaster*. *Genetics* **138**, 1315–1322.
13. Lynch M, Blanchard J, Houle D, Kibota T, Schultz S, Vassilieva L, Willis J. 1999 Perspective: spontaneous deleterious mutation. *Evolution* **53**, 645. (doi:10.2307/2640707)
14. Charlesworth B, Miyo T, Borthwick H. 2007 Selection responses of means and inbreeding depression for female fecundity in *Drosophila melanogaster* suggest contributions from intermediate-frequency alleles to quantitative trait variation. *Genet. Res.* **89**, 85–91. (doi:10.1017/S001667230700866X)
15. Long TAF, Agrawal AF, Rowe L. 2012 The effect of sexual selection on offspring fitness depends on the nature of genetic variation. *Curr. Biol.* **22**, 204–208. (doi:10.1016/j.cub.2011.12.020)

16. Haldane J. 1937 The effect of variation on fitness. *Am. Nat.* **71**, 337–349. (doi:10.1086/280722)
17. Halligan DL, Keightley PD. 2009 Spontaneous mutation accumulation studies in evolutionary genetics. *Annu. Rev. Ecol. Evol. Syst.* **40**, 151–172. (doi:10.1146/annurev.ecolsys.39.110707.173437)
18. Price GR. 1970 Selection and covariance. *Nature* **227**, 520–521. (doi:10.1038/227520a0)
19. Mallet MA, Bouchard JM, Kimber CM, Chippindale AK. 2011 Experimental mutation-accumulation on the X chromosome of *Drosophila melanogaster* reveals stronger selection on males than females. *BMC Evol. Biol.* **11**, 156. (doi:10.1186/1471-2148-11-156)
20. Sharp NP, Agrawal AF. 2012 Male-biased fitness effects of spontaneous mutations in *Drosophila melanogaster*. *Evolution* **67**, 1189–1195. (doi:10.1111/j.1558-5646.2012.01834.x)
21. Keightley PD, Ohnishi O. 1998 EMS-induced polygenic mutation rates for nine quantitative characters in *Drosophila melanogaster*. *Genetics* **148**, 753–766.
22. Houle D, Hughes KA, Hoffmaster DK, Ihara J, Assimacopoulos S, Canada D, Charlesworth B. 1994 The effects of spontaneous mutation on quantitative traits. I. Variances and covariances of life history traits. *Genetics* **138**, 773–785.
23. Sharp NP, Agrawal AF. 2012 Evidence for elevated mutation rates in low-quality genotypes. *Proc. Natl Acad. Sci. USA* **109**, 6142–6146. (doi:10.1073/pnas.1118918109)
24. Sharp NP, Agrawal AF. 2016 Low genetic quality alters key dimensions of the mutational spectrum. *PLoS Biol.* **14**, e1002419-18. (doi:10.1371/journal.pbio.1002419)
25. Zikovitz AE, Agrawal AF. 2013 The condition dependency of fitness in males and females: the fitness consequences of juvenile diet assessed in environments differing in key adult resources. *Evolution* **67**, 2849–2860. (doi:10.1111/evo.12170)
26. Hadfield JD. 2010 MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *J. Stat. Soft.* **33**, 1–22. (doi:10.18637/jss.v033.i02)
27. R Core Team. 2014 R: a language and environment for statistical computing. (doi:10.1111/j.1365-2656.2012.01955.x/full)
28. Bates D, Mächler M, Bolker B, Walker S. 2015 Fitting linear mixed-effects models using lme4. *J. Stat. Soft.* **67**, 1–48. (doi:10.18637/jss.v067.i01)
29. Bolker B. 2017 bbmle: Tools for general maximum-likelihood estimation. https://CRAN.R-project.org/package=bbmle
30. Salomon MP *et al.* 2009 Comparing mutational and standing genetic variability for fitness and size in *Caenorhabditis briggsae* and *C. elegans*. *Genetics* **183**, 685–692. (doi:10.1534/genetics.109.107383)
31. Huang W, Lyman RF, Lyman RA, Carbone MA, Harbison ST, Magwire MM, Mackay TF. 2016 Spontaneous mutations and the origin and maintenance of quantitative genetic variation. *eLife* **5**, e14625. (doi:10.7554/eLife.14625)
32. Crow JF. 1979 Minor viability mutants in Drosophila. *Genetics* **92**, s165–s172.
33. Kelly JK. 1999 An experimental method for evaluating the contribution of deleterious mutations to quantitative trait variation. *Genet. Res.* **73**, 263–273. (doi:10.1017/S0016672399003766)
34. Kelly JK, Willis JH. 2001 Deleterious mutations and genetic variation for flower size in *Mimulus guttatus*. *Evolution* **55**, 937–942. (doi:10.1111/j.0014-3820.2001.tb00611.x)
35. Kelly JK. 2003 Deleterious mutations and the genetic variance of male fitness components in *Mimulus guttatus*. *Genetics* **164**, 1071–1085.
36. Kondrashov AS, Houle D. 1994 Genotype–environment interactions and the estimation of the genomic mutation rate in *Drosophila melanogaster*. *Proc. R. Soc. Lond. B* **258**, 221–227. (doi:10.1098/rspb.1994.0166)
37. Gibson JR, Chippindale AK, Rice WR. 2002 The X chromosome is a hot spot for sexually antagonistic fitness variation. *Proc. R. Soc. Lond. B* **269**, 499–505. (doi:10.1098/rspb.2001.1863)
38. Rice WR. 1984 Sex-chromosomes and the evolution of sexual dimorphism. *Evolution* **38**, 735–742. (doi:10.1111/j.1558-5646.1984.tb00346.x)
39. Sharp N, Agrawal A. 2018 Data from: An experimental test of the mutation-selection balance model for the maintenance of genetic variance in fitness components. Dryad Digital Repository. (doi:10.5061/dryad.2p7qb12)