Simultaneous Estimation of Additive and Mutational Genetic Variance in an Outbred Population of *Drosophila serrata*

Katrina McGuigan,*,1 J. David Aguirre,† and Mark W. Blows*

*School of Biological Sciences, The University of Queensland, Brisbane QLD 4072, Australia, and †Institute of Natural and Mathematical Sciences, Massey University, Albany, Auckland 1011, New Zealand

ORCID IDs: 0000-0002-0525-2202 (K.M.); 0000-0001-7520-441X (J.D.A.)

ABSTRACT How new mutations contribute to genetic variation is a key question in biology. Although the evolutionary fate of an allele is largely determined by its heterozygous effect, most estimates of mutational variance and mutational effects derive from highly inbred lines, where new mutations are present in homozygous form. In an attempt to overcome this limitation, middle-class neighborhood (MCN) experiments have been used to assess the fitness effect of new mutations in heterozygous form. However, because MCN populations harbor substantial standing genetic variance, estimates of mutational variance have not typically been available from such experiments. Here we employ a modification of the animal model to analyze data from 22 generations of *Drosophila serrata* bred in an MCN design. Mutational heritability, measured for eight cuticular hydrocarbons, 10 wing-shape traits, and wing size in this outbred genetic background, ranged from 0.0006 to 0.006 (with one exception), a similar range to that reported from studies employing inbred lines. Simultaneously partitioning the additive and mutational variance in the same outbred population allowed us to quantitatively test the ability of mutation-selection balance models to explain the observed levels of additive and mutational genetic variance. The Gaussian allelic approximation and house-of-cards models, which assume real stabilizing selection on single traits, both overestimated the genetic variance maintained at equilibrium, but the house-of-cards model was a closer fit to the data. This analytical approach has the potential to be broadly applied, expanding our understanding of the dynamics of genetic variance in natural populations.

KEYWORDS mutation-selection balance; heterozygous; middle-class neighborhood; animal model

MUTATION is the ultimate source of genetic variation, and determining the properties of new mutations is a necessary step in understanding the maintenance of genetic variation, adaptation, the evolution and maintenance of sex, and extinction risks in small populations (Lande 1975, 1995; Kondrashov and Turelli 1992; Schultz and Lynch 1997; Whitlock 2000; Zhang and Hill 2002; Johnson and Barton 2005). The mutational heritability of morphological and life-history traits, as well as fitness itself, has been estimated for several species of insects and crop plants (Houle *et al.* 1996; Lynch *et al.* 1999; Halligan and Keightley 2009). These

studies have indicated that mutational heritability typically falls within the range 0.0001–0.005 (Houle *et al.* 1996; Lynch *et al.* 1999), with a typical persistence time of new mutations of 50 generations for life-history traits and 100 generations for morphological traits (Houle *et al.* 1996). However, if we are to understand the contribution of mutation to processes operating in natural populations evolving in the vicinity of their adaptive optimum, we need to evaluate the consequences of mutations in unmanipulated high-fitness populations under natural levels of recombination and outcrossing, a goal that has yet to be fully realized (Lynch *et al.* 1999).

Two general approaches have been used to characterize the input of new mutational variance each generation. These experimental approaches, which both reduce the opportunity for selection to operate, thereby allowing drift to strongly influence the fate of new mutations, differ in the extent to which they manipulate the standing genetic variance in the population. In mutation-accumulation (MA) divergence experiments, a highly

inbred ancestral population is subdivided into many replicate lines, each maintained through inbreeding, and the divergence (variation) among lines is a measure of mutational variance (Lynch and Walsh 1998). The high level of inbreeding imposed in MA experiments results in the assessment of homozygous effects of new mutations against a homozygous genetic background. In some taxa (*e.g., Drosophila melanogaster*), manipulative tools such as balancer chromosomes can be used to accumulate mutations in only part of the genome and in heterozygotes (Lynch and Walsh 1998). Although MA experiments have been employed effectively in a variety of taxa to estimate mutational variance (Houle *et al.* 1996; Lynch *et al.* 1999; Halligan and Keightley 2009), they may not be well suited to investigating the persistence of new mutations (Barton 1990) because it is selection against the mutation in heterozygote form that largely determines the fate of new mutations in natural populations. Furthermore, MA experiments, initiated from a single highly inbred line, do not provide estimates of the genetic variance segregating in the base population, which is necessary for a direct comparison between the mutational and genetic variance from a single population (Houle *et al.* 1996) to understand the selection acting against new mutations.

The second experimental approach, the middle-class neighborhood (MCN) breeding design, overcomes the drawback inherent in MA experiments of accumulating and measuring the effects of new mutations in a highly inbred genetic background. MCN experimental designs begin from an outbred, randomly mating population and reduce the opportunity for natural selection to operate by enforcing equal contributions to the next generation from all families (Shabalina *et al.* 1997; Lynch *et al.* 1999; Mack *et al.* 2001; Yampolsky *et al.* 2001). Although MCN experiments provide a more relevant estimate of the effects and fate of new mutations than do MA designs, they can suffer from other problems (Keightley *et al.* 1998; Lynch *et al.* 1999). In particular, direct estimation of the mutational variance from MCN experiments in the presence of substantial genetic variance in the base population has been highlighted as a key limitation of this design (Lynch *et al.* 1999). One approach to solving this problem has been to impose a quantitative genetic breeding design simultaneously in both the MCN population and the ancestral population from which the MCN was initially derived and using these data to demonstrate higher genetic variance (owing to new mutations) in the MCN population (Roles and Conner 2008). The logistical challenge of applying a large quantitative genetic breeding design in two populations with a single generation (common environment) is likely to limit the application of MCN designs used in this way despite their benefit in terms of allowing the estimation of mutational variance in seminatural populations undergoing random mating and recombination.

Although not yet widely applied, an analytical solution to the problem of estimating the mutational variance in a population with substantial standing genetic variance is available. Wray (1990) adapted the numerator relationship matrix

used in Henderson's mixed-model methodology (the animal model) (Henderson 1976), which uses pedigree information to estimate the additive genetic variance in measured traits, to account for the input of new mutations each generation. The animal model has facilitated recent advances in understanding of microevolution in natural populations where individual phenotypes and relatedness among individuals are known (Kruuk 2004; Wilson *et al.* 2010). Estimation of mutational variance through an animal model effectively addresses the major limitation of MCN designs identified by Lynch *et al.* (1999), allowing mutational variance to be estimated within the more natural context of an outbred population. Specifically, by fitting both the mutational and additive genetic numerator relationship matrices in these models, it is possible to explicitly partition the mutational variance from the standing genetic variance and to account for changes in the standing genetic variance over the generations of mutation accumulation, particularly the loss of variance owing to inbreeding or to the response to selection under laboratory culture. Pragmatically, by collecting phenotypic data from a modest population size each generation is logistically more tractable than attaining large sample sizes within a single generation, such as might be necessary when applying single-generation breeding designs to simultaneously estimate the genetic variance in the ancestral and derived MCN populations. Analysis of phenotypes across multiple pedigreed generations maintains the statistical power to detect effects. The combination of the MCN design with Wray's mutational numerator relationship matrix within an animal model analysis has the potential to provide relevant estimates of mutational variance across a wide array of taxa not previously considered amenable to MA studies.

The estimation of mutational effects in heterozygous form in an outbred genetic background should more closely reflect the strength of selection acting against mutations when they first appear in a natural population. In particular, the estimation of both mutational and additive genetic variance from the same data, within the same model, will provide the most direct and biologically relevant determination of the average selection against mutations ($s = V_M/V_A$). Although prominent in the theoretical literature (Kondrashov and Turelli 1992), these parameters have been determined relatively few times, and current estimates typically depend on genetic and mutational parameters measured in different populations (and in some cases at different times using different experimental protocols in different laboratories), with mutational parameter estimates derived from inbred lines (Houle *et al.* 1996; Denver *et al.* 2005; McGuigan *et al.* 2014). The simultaneous estimation of standing genetic and mutational variance from the same outbred population also allows a direct quantitative assessment of the theoretical models of mutation-selection balance that attempt to explain equilibrium levels of additive genetic variance (Turelli 1984; Lynch and Hill 1986; Bürger 2000; Johnson and Barton 2005).

Despite its promising features, the animal model approach to estimating mutational variance has been applied on few

occasions and only in highly manipulated populations (Keightley and Hill 1992; Casellas and Medrano 2008; Casellas *et al.* 2010). Casellas and colleagues adapted the approach of Wray (1990) to estimate the mutational variance in reproductive traits in a pedigreed inbred laboratory population of mice (Casellas and Medrano 2008) and an experimental flock of sheep (Casellas *et al.* 2010) within a Bayesian framework. Here we report animal model estimates of mutational parameters for 19 traits in an outbred population of *Drosophila serrata* measured over 22 pedigreed generations of mutation accumulation under a MCN breeding design. We use these estimates to make a direct quantitative test among competing models of the maintenance of genetic variance under mutation-selection balance.

Materials and Methods

Study system and experimental design

Our experiment was conducted on a population of *D. serrata* collected from The University of Queensland's St. Lucia campus in Brisbane, Australia. The population was initiated from 33 inseminated wild-caught females (polyandry is typical in the wild: see Frentiu and Chenoweth 2008) and underwent eight generations of random mating as a mass-bred population (initiated from >1000 individuals in generations 2–8). This population was maintained on a 14-day-per-generation cycle in a constant-temperature facility set at 25° with a 12:12 light-dark regime and reared on a yeast-agar-sugar medium (Rundle *et al.* 2005). After eight generations of random mating, we initiated a paternal half-sibling breeding design with 120 sires each mated to two dams (Hine *et al.* 2014); to establish the MCN population, one son and one daughter per full-sibling family were randomly paired to found the first generation of the mutation accumulation experiment, generating 240 pairs in the MCN population.

The MCN population was maintained through 240 mating pairs in each generation for a further 21 generations under the same conditions to which it had been adapted when brought into the laboratory from the wild. A total of 10,440 animals were present in the pedigree. A single male and female offspring per family were randomly paired with a nonsibling of the opposite sex as 2-day-old virgins to initiate the next generation. Full pedigree information was recorded for both sexes in every generation.

We determined that the average inbreeding coefficient in generation 22 was 1.198% (corresponding to a 0.0006 increase per generation), slightly lower than in previous MCN studies in *Drosophila* (0.001–0.002 per generation) (Shabalina *et al.* 1997; Yampolsky *et al.* 2001). On average, 5% of pairs failed to contribute offspring to the next generation; when pairs failed to produce any offspring, three offspring from some (randomly chosen) families were used to ensure that the census population size remained constant (240 pairs). Failure to reproduce might reflect viability selection; similarly low levels of reproductive failure were reported from

a MCN experiment in radishes, where 3% of plants failed to produce seeds (Roles and Conner 2008), and in *D. melanogaster*, where 1–2% of pairs failed to reproduce each generation (Shabalina *et al.* 1997).

After being left to mate for 3 days, females were discarded, and larvae were left to develop, while males were removed and assayed for cuticular hydrocarbons (CHCs) and wing characteristics. *D. serrata* CHCs, waxy lipids present on the cuticle, are under directional sexual selection through mate choice but also have consequences for nonsexual fitness (Higgie and Blows 2007; Chenoweth *et al.* 2010; Hine *et al.* 2011; McGuigan *et al.* 2011; Delcourt *et al.* 2012). CHC profiles of individual males were determined using standard gas chromatography (GC) protocols (Sztepanacz and Rundle 2012). The abundance of nine compounds was characterized by the area under the respective peak (Figure 1A); to account for nonbiological variation in abundances, we determined the proportion of each of the nine CHCs and transformed these proportions to log contrasts [where (Z,Z)-5,9-$C_{24:2}$ was the peak used as the divisor] for analysis as appropriate for compositional data (Aitchison 1986; Blows and Allan 1998). Wing size and shape were characterized via nine wing landmarks (Figure 1B) (McGuigan and Blows 2013). Wing size was characterized as centroid size (CS), the square root of the sum of squared distances of each of the nine landmarks to their centroid, a metric commonly used to describe size (Rohlf 1999). Wing size is closely correlated with body size in *Drosophila* (Robertson and Reeve 1952; Partridge *et al.* 1987), and although there is no direct evidence of sexual selection on size, it is genetically associated with sexual fitness in *D. serrata* (McGuigan and Blows 2013). Wing shape was characterized via 10 interlandmark distances measured on aligned landmarks (after CS was taken into account). Like size, male wing shape does not appear to be a direct target of female mate choice but is pleiotropically associated with male sexual fitness and viability (McGuigan *et al.* 2011; McGuigan and Blows 2013).

Throughout this paper, we assume that our experimental population has reached equilibrium within the 30 generations after being founded from the field. Although this is unlikely to be strictly true, the approach to equilibrium is more rapid when the initial starting population has a high genetic variance and will be attained within a few dozen generations in populations of size $N_e < 1/4\mu$, where μ is the genic mutation rate (Lynch and Hill 1986). In our case, the founding population was sampled from a large field population, and in our experimental MCN design, N_e was $2N - 2 = 958$ (Crow and Kimura 1970), enforced for the 21 generations of MA. Assuming that μ is in the range 10^{-5} (Lynch and Walsh 1998) to 10^{-10} (Haag-Liautard *et al.* 2007; Keightley *et al.* 2009, 2014, 2015), the size of our experimental population sits comfortably within this limit.

Data analysis and parameter estimation

The base population (*i.e.*, the sires of the half siblings) were not assayed for CHCs or wings. In subsequent generations,
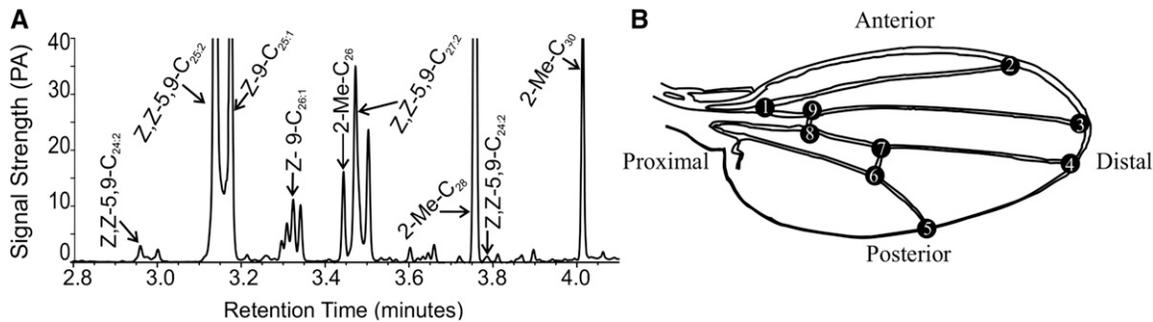
Figure 1 The traits analyzed for mutational and additive genetic variance. (A) Gas chromatography trace for male *D. serrata*. Expression of nine CHCs was quantified by determining the area under the respective peak. (B) Wing shape was characterized as interlandmark distances based on the positions of nine landmarks (Table 1).

some males died before they could be phenotyped (but after siring offspring); further, delicate wings of some males were badly damaged and could not be measured. Of the 5040 males present in the pedigree in generations 1–21, CHC data were collected for a total of 4945 individuals, and wing data were collected for 4236 individuals. The pedigree and phenotype data are available at http://dx.doi.org/10.6084/m9.figshare.1542535.

The eight CHC log contrasts, wing CS, and the 10 wing-shape traits were each analyzed individually. The CHCs were measured by GC in three separate blocks, within which samples were randomized among generations: block 1 = generations 1–6; block 2 = generations 7–12; and block 3 = generations 13–21. Because CHC means differed among GC blocks and block was confounded with generation, we first mean centered the CHC data within each GC block prior to analysis. No such effects occurred for wings, and wing data were not centered before analysis.

For each of the 19 traits, we partitioned the phenotypic variation into additive, mutational, and residual components in a univariate animal model

$$y_{ijkl} = \mu + t_i + a_j + m_k + e_l \tag{1}$$

where μ is the overall mean, t indicates generation, a is the additive genetic breeding value, m is the mutation breeding value, and e is the residual error. The variance components for additive and mutation effects were calculated as the variance in breeding values scaled by the corresponding relationship matrix $V_A = \mathbf{A}\sigma_a^2$ and $V_M = \mathbf{M}\sigma_m^2$ (Kruuk 2004; Casellas and Medrano 2008). For the error variance, we assumed that errors were uncorrelated among individuals, and therefore, $V_R = \mathbf{I}\sigma_r^2$, where $\mathbf{I}$ is an identity matrix. Although, owing to breeding failures, a small number of males each generation (5% on average in any generation; see earlier) shared a common rearing environment (and were full siblings), for reasons of computational simplicity, we do not fit rearing vial in the models. Common environment therefore might have inflated our estimate of additive genetic variance. However, we note that only a small fraction of animals shared a common environment within a generation, and this was not perpetuated among relatives across generations.

Although large pedigrees with incomplete phenotypic information (*e.g.*, for sex-limited traits in economically important agricultural taxa) are common, the pedigree structure from the MCN design differs markedly from those types of populations. Specifically, wild and agricultural populations typically consist of both full and half siblings, with overlapping generations. In the MCN design, generations are non-overlapping, and total population size remains constant because family size is experimentally controlled in each generation. This results in a strongly linear pedigree structure in which most of the genetic information comes from parent-offspring relationships (coefficient of coancestry = 0.5). We recorded maternal relationships each generation, but preliminary analyses of model (1) showed that including maternal pedigree links resulted in a dense numerator relationship matrix that required prohibitively long computational times to solve.

Another consideration when including female pedigree links is that our analysis would solve for the breeding values of the 5280 females for which we had no phenotype information (Henderson 1976), an important consideration given our prior knowledge that the genetic bases of CHCs (Gosden *et al.* 2012) and wings (McGuigan and Blows 2007) differ between the sexes. Nevertheless, analyzing model (1) with a male-only pedigree has the effect of ignoring many of the relationships resulting in coefficients of coancestry ≤ 0.125 (except grandfather-grandson, great grandfather–great grandson, etc.). We consider further in the discussion how our pedigree affected the accuracy and precision of our parameter estimates. Additionally, relationship matrices for the male-only pedigree do not account for inbreeding, but as noted earlier, inbreeding was very low in the current experiment and unlikely to have strongly influenced parameter estimates.

To estimate additive genetic variance using the animal model, we used the method of Henderson (1976) to calculate the inverse of the additive genetic numerator relationship matrix

$$\mathbf{A}^{-1} = \left(\mathbf{T}^{-1}\right)'\left(\mathbf{D}^{-1}\right)^2 \mathbf{T}^{-1} \tag{2}$$

where $\mathbf{T}^{-1}$ is a lower triangular matrix, and $\mathbf{D}^{-1}$ is a diagonal matrix. For a male-only pedigree, each row of $\mathbf{T}^{-1}$ has one

nonzero value (equal to -0.5) to the left of the diagonal linking sires and sons. If the additive relationship matrix $\mathbf{A}$ is calculated, the elements on the diagonal of $\mathbf{D}^{-1}$ are equal to the inverse of the elements on the diagonal of the Cholesky of $\mathbf{A}$.

In an animal model context, new mutations will generate variances and covariances among animals that cannot be traced to the base population because they are the result of mutations accruing in generations subsequent to that ancestral base. To estimate the mutational variance, we used the method of Wray (1990) with the modifications outlined in Casellas and Medrano (2008) to calculate the numerator relationship matrix for mutational effects

$$\mathbf{M} = \sum_{k=0}^{t} \mathbf{A}_k \qquad (3)$$

where t is the number of generations, and $\mathbf{A}_k$ is the relationship matrix for additive effects attributable to mutations arising in generation k. The elements of $\mathbf{A}_k$ describe the additive genetic relationships among individuals in generations after generation k if ancestors in generations before k are ignored. Our pedigree considered 22 nonoverlapping generations; hence $\mathbf{M}$ was calculated for generations 0–21 [for an illustration of this procedure, see Wray (1990)].

Because $\mathbf{M}^{-1}$ and $\mathbf{A}^{-1}$ have the same pattern of nonzero values, it is possible to adapt Henderson's method for calculating $\mathbf{A}^{-1}$ [Equation (2)] to calculate $\mathbf{M}^{-1}$ (Henderson 1976). Following Wray (1990),

$$\mathbf{M}^{-1} = \left(\mathbf{T}^{-1}\right)'\left(\mathbf{E}^{-1}\right)^2\mathbf{T}^{-1} \qquad (4)$$

where $\mathbf{E}^{-1}$ is a diagonal matrix with elements equal to the inverse of the elements on the diagonal of the Cholesky of $\mathbf{M}$, and $\mathbf{T}^{-1}$ is the same lower triangular matrix as used to calculate $\mathbf{A}^{-1}$ earlier.

We fit model (1) using the MCMCglmm package (Hadfield 2010) in R (v. 3.0.1). The MCMCglmm analysis was implemented for all traits individually with 10,050,000 iterations, 50,000 burn-in, and a thinning interval of 10,000, resulting in 1000 posterior estimates of the parameters for each trait. Priors for the location parameters were normally distributed around a mean of zero and a variance of 10^8. The scale parameters for additive and mutational variance were chosen based on vague prior knowledge of additive genetic and mutational heritability for morphological traits. For the residual and additive genetic variance components, priors were inverse-Wishart distributed with degrees of freedom equal to 1 and scale parameters equal to two-thirds and one-third of the phenotypic variance, respectively, corresponding to an additive genetic variance of $h^2 = 0.3$ (Mousseau and Roff 1987; Roff and Mousseau 1987). For the mutational variance, we used a prior conforming to the noncentral F-distribution with the scale hyperparameter equal to one one-hundredth of the residual variation. It is recommended that the scale hyperparameter be a value larger than what would be typically expected (Gelman 2006; Hadfield 2014); hence we specified $h_M^2 = 0.01$ based on the largest mutational heritability observed for morphologic traits in *D. melanogaster* (Houle *et al.* 1996; Lynch *et al.* 1999). In preliminary analyses, we investigated the robustness of our posterior estimates to our prior specifications; varying the values for the scale parameter and degrees of freedom, we observed little change in the parameter estimates. Even a nonsensical prior, where we switched the scale parameter for the prior variance of additive effects and mutational effects (*i.e.*, a prior where $h^2 = 0.01$ and $h_M^2 = 0.3$), gave posterior distributions that overlapped the posterior distribution from the prior based on vague prior knowledge (as described earlier).

MCMCglmm estimates of variance components are bounded at zero, and thus a hypothesis test for the presence of mutational or genetic variance based on examining whether posterior density intervals of variance components overlap zero is uninformative. One approach to determine whether estimates are greater than zero is to compare them to the variation expected to be observed through random sampling (Hadfield *et al.* 2010; Aguirre *et al.* 2014). To do this, we randomized the pedigree and reestimated V_A and V_M from the observed phenotypic data. Estimates of V_A and V_M from the randomized pedigree data give an indication of the amount of variation we can expect by chance, given our experimental design and our statistical model. Because our estimates of observed V_A and V_M were derived using an animal model based on a multigenerational pedigree with nonoverlapping generations, to remove the biological structure captured by the pedigree, we randomized individuals among families within generations to randomly break identity-by-descent relationships within the data and then recalculated $\mathbf{M}^{-1}$ and $\mathbf{A}^{-1}$. We repeated the pedigree randomization 100 times to account for any uncertainty introduced by the randomization procedure. Then, for every randomization, we replaced the observed $\mathbf{M}^{-1}$ and $\mathbf{A}^{-1}$ matrices in model (1) with the randomized (null) $\mathbf{M}^{-1}$ and $\mathbf{A}^{-1}$ matrices. The models based on the randomized pedigrees had the same priors, burn-in, and thinning intervals as those for the observed data, except that because we had 100 randomized data sets, we only stored 10 MCMC samples (after burn-in and thinning intervals) from each null data set per trait to give 1000 posterior samples for null V_A and null V_M. Overall, the null V_A and V_M estimates for each trait indicated low autocorrelation ($r < 0.1$) among MCMC samples within and among randomizations and reasonable convergence ($R < 1.23$) (Gelman and Rubin 1992).

Finally, since the $\mathbf{A}$ and $\mathbf{M}$ relationship matrices in Wray's model (Wray 1990) share many relationships, we investigated the joint behavior of the estimates of V_A and V_M arising from model (1). We estimated the correlation between the observed estimates of V_A and V_M for each trait, as well as the correlation between the null V_A and null V_M estimates (Supporting Information, Table S1). Observed estimates of V_A and V_M were uniformly negatively correlated, while the null estimates were essentially uncorrelated, although there

was some bias toward very small negative correlations. Correlations among V_A and V_M estimates therefore are generated by the biological signal in the data and not the statistical model.

Data availability

Data are available through figshare: http://dx.doi.org/10.6084/m9.figshare.1542535.

Results

Estimates of V_A and V_M in an outbred population

Median heritability ranged from 0.073 to 0.499 for the CHC traits and from 0.316 to 0.514 for wing traits (Table 1), suggesting low to moderate heritability for all traits. To determine the statistical support for the presence of additive genetic variance, we compared the posterior distribution from the observed data to the null expectation for each trait generated by the randomization of the pedigree structure (Table S2). Figure 2A illustrates the outcome of this approach for one trait, (Z,Z)-5,9-$C_{29:2}$. For this trait, the posterior distribution of V_A estimated from the observed data is distinct from the null, with a median substantially above the null median (Figure 2A). The posterior distribution of the null data has a nonzero median, as expected from the constraint, inherent in the estimation method, to return a positive variance component. However, the density of posterior values of the null close to zero is greater than that for the observed data (Figure 2A). The 95% posterior confidence intervals of the null and observed distributions do not overlap (Figure 2A), indicating significance of the V_A estimate at the conventional $\alpha = 0.05$ level. All but three traits [(Z)-9-$C_{25:1}$, (Z)-9-$C_{26:1}$, and ILD 5-6] displayed significant V_A values using this approach (Table S2).

Mutational heritability h_M^2 is calculated by scaling the mutational variance by the environmental variance (Lynch and Walsh 1998), representing the instantaneous rate of increase in heritability when there is no genetic variation. When estimating V_M from MA lines derived from a common inbred ancestor, only environmental variation and mutational variance contribute to phenotypic variation, and mutational variance will be a small component of the total variance such that the environmental variance essentially corresponds to the total phenotypic variance. Under a MCN breeding design, standing genetic variance can be substantial, here comprising up to 50% of the phenotypic variance (Table 1). We therefore departed from the standard approach and scaled the mutational variance by the total phenotypic variance to calculate mutational heritability ($h_M^2 = V_M/V_P$).

Our estimates of h_M^2 ranged from 0.0010 to 0.0068 for 18 of the 19 traits, with one trait (ILD 5-6) having an h_M^2 of over twice the upper value of this range, 0.018 (Table 1). Both the null and observed posterior distributions of V_M are truncated at zero, an example of which is shown for the trait (Z,Z)-5,9-$C_{29:2}$ in Figure 2B. There are two consequences of this posterior distribution. First, although the median V_M of the

Table 1 Median additive and mutational heritability, the selection coefficient, and the coefficient of mutational variance for each trait (defined in Figure 1)

Trait	$h_A^{2\ a}$	$h_{M\_obs}^{2\ a}$	$h_{M\_adj}^{2\ b}$	s^c	$CV_M{}^d$
(Z,Z)-5,9-$C_{25:2}$	0.1459	0.00104	0.00060	0.00413	—
(Z)-9-$C_{25:1}$	0.0729	0.00128	0.00060	0.00845	—
(Z)-9-$C_{26:1}$	0.1383	0.00264	0.00192	0.01410	—
2-Me-C_{25}	0.4430	0.00184	0.00120	0.00270	—
(Z,Z)-5,9-$C_{27:2}$	0.4985	0.00176	0.00107	0.00212	—
2-Me-C_{28}	0.2647	0.00157	0.00091	0.00343	—
(Z,Z)-5,9-$C_{29:2}$	0.2216	0.00238	0.00172	0.00788	—
2-Me-C_{30}	0.3376	0.00226	0.00150	0.00441	—
Size	0.3420	0.00541	0.00459	0.01349	0.00159
ILD 1-8	0.3499	0.00294	0.00199	0.00573	0.00289
ILD 2-4	0.4352	0.00189	0.00111	0.00254	0.00114
ILD 2-5	0.3161	0.00374	0.00291	0.00916	0.00158
ILD 2-8	0.4120	0.00687	0.00611	0.01483	0.00249
ILD 3-7	0.3365	0.00529	0.00440	0.01327	0.00185
ILD 3-9	0.3775	0.00429	0.00353	0.00950	0.00147
ILD 4-5	0.3439	0.00569	0.00482	0.01393	0.00245
ILD 5-6	0.3621	0.01771	0.01693	0.04632	0.00637
ILD 5-8	0.4391	0.00643	0.00567	0.01319	0.00239
ILD 7-9	0.5138	0.00528	0.00454	0.00872	0.00362

[a] Median heritability ($h_A^2 = V_A/V_P$; $h_M^2 = V_M/V_P$) estimated by applying the relevant equation to each of the 1000 posterior estimates and then determining the median of those values.
[b] Median $h_{M\_adj}^2$ = median $h_{M\_observed}^2$ − median $h_{M\_null}^2$.
[c] s = (median V_{M_obs} − median V_{M_null})/V_A, where estimates of V_M can be found in Table S2.
[d] $CV_M = \sqrt{\text{median } V_{M\_observed} - \text{median } V_{M\_null}}/\bar{X}$, where $\bar{X}$ is the trait mean. CV_M was not estimated for CHCs because the CHC data structure (log contrasts) is not suitable to mean scaling (Hansen *et al.* 2011; Houle *et al.* 2011).

observed data for (Z,Z)-5,9-$C_{29:2}$ is higher than the median V_M of the null data (Figure 2B), clearly, a standard test of significance of V_M from the null would fail to reject the hypothesis of no mutational variance for this trait (and other traits) (Table S2). The null distribution tails off more abruptly than the distribution from the observed data, which has more posterior samples with larger estimates of V_M and a greater density of values sitting further away from zero (Figure 2B). Across all 19 traits, the median of the observed data was always greater than the median of the null data (Figure 3 and Table S2), a result that is highly unlikely to have occurred by chance. The consistency of greater mutational variance in the observed data over the null therefore strongly suggests that we were able to detect mutational variance in this experiment, although our power to demonstrate statistical significance for each individual trait was very limited. We further consider this question of power in the *Discussion*.

The second consequence of the truncated posterior distribution of observed and null analyses is that the same process that results in a push away from the zero boundary in the null distribution will also inflate the estimate the observed median V_M to some extent. Here we chose to explicitly account for this known inflation bias in our estimates using the null model estimates to correct the observed data estimates (McGuigan *et al.* 2014). Although the same phenotypic data were used in both observed and null model analyses (and V_P
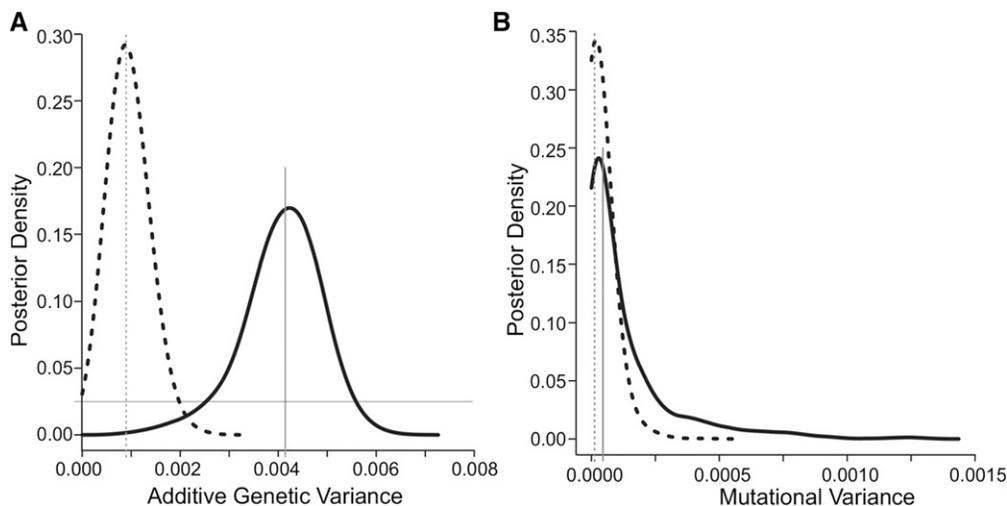
Figure 2 Distribution of posterior estimates for additive genetic and mutational variance (both multiplied by 1000) in the CHC (Z,Z)-5,9-C$_{29:2}$. The posterior distributions of the observed data (solid line) and from the randomized pedigree (dashed line) are plotted for (A) the additive genetic variance and (B) the mutational variance. The median values of each are indicated by the vertical lines through the respective distribution. The thin solid horizontal line in A indicates 2.5% of the distribution; the upper values of the null (dashed) distribution and the lower values of the observed (solid) distribution cross below this line, indicating that the 95% posterior confidence intervals of the two distributions do not overlap, providing statistical support at $P < 0.05$.

is the same for both) (Table S2), the posterior samples are not paired. Therefore, we cannot apply the bias correction to the 1000 posterior samples but rather apply it to the summary statistic of that distribution, the median. Thus, we determined the median h_M^2 from each distribution and from these medians calculated $h_{M\_adj}^2 = h_{M\_observed}^2 - h_{M\_null}^2$ (Table 1). We similarly corrected estimates of V_M scaled by $V_A(s)$ or trait mean CV_M by first correcting the median estimate of V_M (V_{M_adj} = median $V_{M\_observed}$ − median V_{M_null}) and then dividing by median $V_A(s)$ or taking the square root and then dividing by the median trait mean CV_M (Table 1). In the next section, we compare these adjusted estimates of mutational variance scaled by the total phenotypic variance, additive genetic variance, or the trait mean with published estimates.

Comparison of magnitudes of mutational variance

We compared our adjusted estimates of mutation with those reported for a range of morphological and life-history traits in other plant and animal taxa (Figure 4). Wing traits closely matched the observed median mutational heritability for morphological traits in other taxa, while CHCs were more similar in median heritability to life-history traits (Figure 4A). To our knowledge, we present the first estimate of mutational variation in CHCs or similar pheromone traits. Mutational variation in wing shape and size has been estimated previously from MA experiments in *D. melanogaster* (Santiago *et al.* 1992; Houle *et al.* 1996; Houle and Fierst 2013). The average h_M^2 for *D. melanogaster* wing dimensions [in Table 1 of Houle *et al.* (1996)] was 0.00202 (CV_M = 0.00136) compared with the slightly higher estimates in this study for *D. serrata* wing centroid size of 0.00459 (0.00159) and median estimate of the 10 wing-shape traits of 0.00447 (0.00239) (Table 1). Houle and Fierst (2013) reported mutational heritability of

wing shape from estimates in both heterozygote (0.001) and homozygote (0.0024) form, both again slightly lower than our estimates.

When mutational variance is scaled by the standing genetic variance, one can calculate the average selection against a mutant heterozygote ($s = V_M/V_A$) (Barton 1990). Following the pattern for mutational heritability, median s was typically higher for wing traits (0.0132) than for CHCs (0.0043) (Table 1 and Figure 4B). For both types of traits, the median s was comparable to that reported previously for morphological traits and slightly below that reported for life-history traits (Figure 4B). Notably, there was a single wing trait (distance between landmarks 5 and 6) (Figure 1B) with s markedly higher than for other traits (Table 1). This was the only trait with a median selection coefficient above 0.021, which represents the median level of selection against viability mutations (Houle *et al.* 1996).

Tests of mutation-selection balance theory

With estimates of standing genetic and mutational variance from the same outbred population, we can begin to make direct quantitative assessments of the theoretical models of mutation-selection balance that attempt to explain the levels of V_A maintained in the presence of stabilizing selection and mutation. For our traits, a positive relationship exists between the magnitudes of V_A and of V_M (Figure 5), and we now determine whether this relationship is consistent with drift alone or direct stabilizing selection on individual traits. The trait ILD 5-6, with the outlying estimate of h_M^2, has not been included in the following analyses because it would have had a highly disproportionate influence on the model-fitting approach employed; we note, however, that the pattern for this trait was consistent with the predicted pattern under the house-of-cards model generated from the other 18 traits.
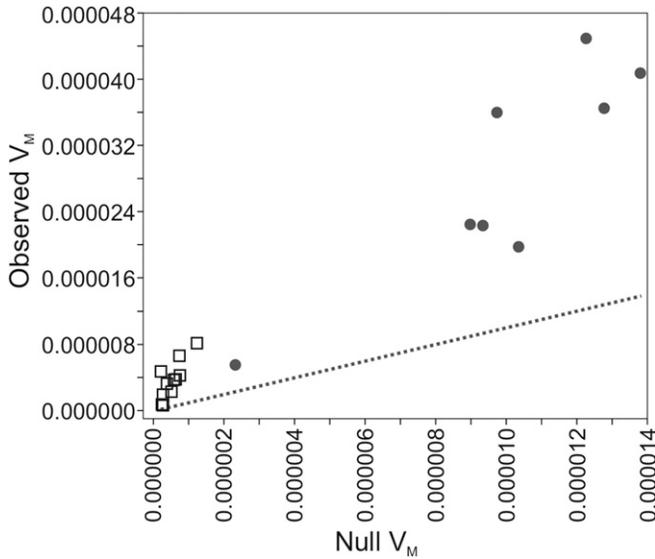
Figure 3 The median mutational variance estimated for CHCs (shaded circles) and wing traits (squares) from the observed (y-axis) and null analyses, where the phenotypic data were the same, but the relationships in the pedigree were broken by randomization. The dotted line indicates the 1:1 relationship expected if the observed analysis returned the same estimate of V_M as the null.

As a first step, the equilibrium genetic variance under a neutral model is given by $V_A = 2N_e V_M$ (Lynch and Hill 1986). For our experiment, where $N_e = 958$ and the median adjusted $V_M = 4.52 \times 10^{-6}$ (Table S2), this equation predicts an additive genetic variance of 8.65×10^{-3}, twenty times greater than the observed median V_A of these traits (0.397×10^{-3}) (Table S2). This neutral expectation is plotted in Figure 5 on the heritability scale and clearly demonstrates that assuming no selection acting against mutations affecting these traits results in substantial overestimation of the standing genetic variance given the observed mutational variance. The failure of this model is not surprising given the evidence for selection acting on both wing and CHC traits in this and related *Drosophila* species (Hine *et al.* 2011; McGuigan *et al.* 2011; McGuigan and Blows 2013).

Next, we considered two competing models of drift-mutation-selection balance for a single trait under direct stabilizing selection. First, assuming that the variance in heterozygous effects of mutations at each locus follows a Gaussian distribution, the Gaussian allelic approximation for mutation-selection equilibrium is given by Lynch and Lande (1993) as

$$V_A = \sqrt{\left(\frac{\ell V_s}{2N_e}\right)^2 + 2\ell V_s V_M} - \frac{\ell V_s}{2N_e} \qquad (5)$$

where ℓ is the number of loci that contribute to the trait, V_s is the strength of stabilizing selection on the trait (scaled by V_P), and h_A^2 and h_M^2 are substituted for V_A and V_M to place the prediction on the heritability scale. Although most theoretical work assumes $V_s \geq 20$, empirical estimates of this parameter are lower (median = 5) (Kingsolver *et al.* 2001, 2012; Johnson

and Barton 2005), so we fixed $V_s = 5$ to fit Equation (5) to the data. Maximum-likelihood optimization was conducted using the NLMIXED procedure in SAS (v. 9.3) (SAS Institute 2011). Figure 5 shows the fit to the data of this model ($-2\text{LL} = -23.9$, AIC = -19.9) when the maximum-likelihood estimate of $\ell = 4.2$ is used. It is important to note that this model-fitting approach is not employed to give biologically interpretable estimates of the parameter ℓ because it is completely confounded with V_s, which we fix at $V_s = 5$ to fit model (5).

Using an alternative set of assumptions known as the *house-of-cards mutational scheme*, the distribution of mutational effects at each locus is assumed to be highly leptokurtic, with most V_M contributed by rare alleles of larger effect (Johnson and Barton 2005). The stochastic house-of-cards approximation for a single trait under direct stabilizing selection is given by (Bürger 2000, equation 2.15)

$$V_A = \frac{2V_M}{V_M\left(2 + \frac{1}{U\ V_s}\right) + N_e^{-1}} \qquad (6)$$

where U is genomic per-trait mutation rate, which is thought to be on the order of $U = 0.1$ (Lynch and Walsh 1998). Again fixing $V_s = 5$ and substituting h_A^2 and h_M^2 for V_A and V_M, we fit this model to the data ($-2\text{LL} = -28.6$, AIC = -24.6), and Figure 5 shows the result when the maximum-likelihood estimate of $U = 0.061$ is used. As with ℓ, this analysis cannot provide a biologically interpretable estimate of U because it is confounded with V_s, which we fix [see model (6)]. Comparison of the Akaike information criterion (AIC) suggests that the house-of-cards model is a better fit to the observed data than the Gaussian allelic approximation. However, as we discuss later, both models tend to overestimate the predicted genetic variance when more biologically plausible parameter values are used.

Discussion

The generation of new mutational variance in natural populations has widespread evolutionary implications, and a key goal is to understand the consequences of mutation in populations evolving in the vicinity of their adaptive optimum under natural levels of recombination and outcrossing (Lynch *et al.* 1999). The mutational numerator relationship matrix developed by Wray (1990), when used within an animal model, has the potential to change our approach to this difficult task. This approach allows the estimation of mutational parameters from a wide array of taxa and traits in the same way that the widespread adoption of the animal model itself expanded the study of genetic variance beyond the laboratory or captive bred populations.

Mutational variance in an outbred population

The estimated mutational heritability in wing shape and CHCs for *D. serrata* males was very similar in magnitude to that reported from inbred-line experimental designs applied to
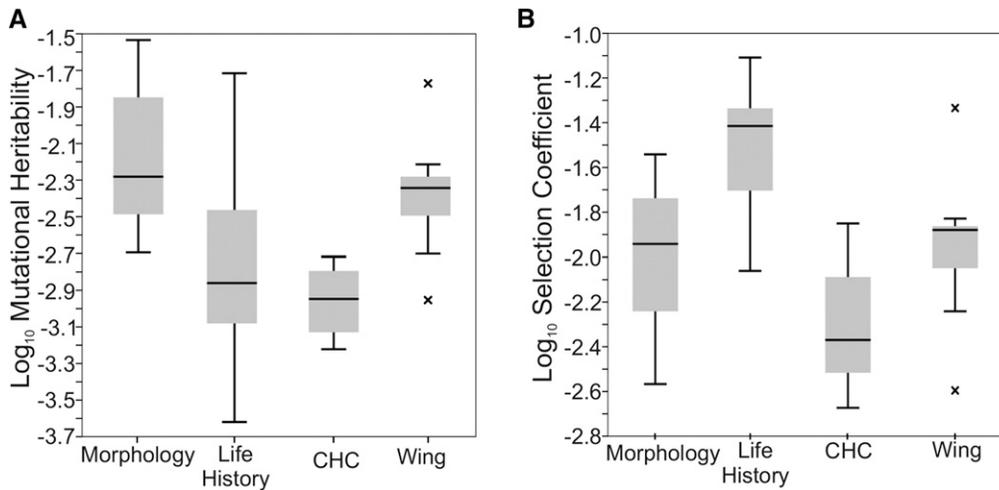
Figure 4 Mutational heritability and selection coefficients of traits in this study ("CHC" and "Wing" from Table 1) compared to those reported in Houle *et al.* (1996) ("Morphology" and "Life History"). Boxes represent the range between the first and third quartiles of the distribution, with the median shown as a black band across the box; whiskers represent up to 1.5 times the interquartile range, and traits outside this are plotted individually (crosses). Published estimates are from Table 1 of Houle *et al.* (1996) (for the subset of studies identified by Houle *et al.* as having relatively unbiased estimates of mutation, *i.e.*, studies also in their Table 3). For traits with multiple published estimates in the same species, as reported in Houle *et al.* (1996), we took the mean estimate. Houle *et al.* (1996) classified published estimates as life-history or morphological trait; the plotted published data include 9 morphological traits and 14 life-history traits from *Daphnia pulex*, *D. melanogaster*, *Tribolium castaneum*, *Mus musculus*, *Hordeum vulgare*, *Oryza sativa*, and *Zea mays*.

D. melanogaster. Lynch *et al.* (1999) reported the range of published estimates of h_M^2 to be 0.0005–0.005 for *D. melanogaster*, matching the range found here for *D. serrata* of 0.0006–0.006 with one exception. The similarity in the magnitude of h_M^2 between estimates derived from inbred and outbred experimental populations is surprising because there are two reasons that could cause h_M^2 measured in an MCN experimental design to be downwardly biased in comparison to an inbred population. First, in MCN experimental designs, the high (relative to inbred lines) within-family variance provides the opportunity for selection to act (Lynch *et al.* 1999; Moorad and Hall 2009), under our experimental design, via competition among full siblings within each family vial. Selection against new deleterious mutations then could result in an underestimation of the mutational variance. It is worth noting that selection likely operates on inbred MA divergence lines to an underappreciated extent (Keightley *et al.* 1993; Lynch *et al.* 1999; McGuigan and Blows 2013), suggesting that the comparison of estimates from the different methods is not simply one between the presence and absence of selection.

Second, in the MCN design, mutational effects are characterized in a heterozygous genetic background in contrast to studies of inbred MA lines, where effects are typically determined in homozygous form. If many mutations are recessive, heterozygous effects of mutations on the phenotype will be smaller than homozygous effects, consequently reducing the level of mutational heritability detected under the MCN design. This question of dominance has been explicitly addressed in inbred-line MA studies by crossing among MA lines to estimate mutational effects in heterozygotes. For instance, Houle and Fierst (2013) reported two to three times higher mutational heritability of wing shape when estimated from crosses within *vs.* among MA lines. Although crosses among MA lines assess dominance effects of mutations, the genetic background within which these mutations are assessed

is still relatively genetically depauperate. There are few estimates of heterozygous mutational heritability in outbred populations (see Shabalina *et al.* 1997; Roles and Conner 2008). The only other estimate that we are aware of to come from application of Wray's method (Wray 1990) in an outbred population is $h_M^2 = 0.006$ for litter size in a domesticated sheep population (Casellas *et al.* 2010). That estimate is at the upper end of reported estimates of mutational heritability in life-history traits (Houle *et al.* 1996) and at the upper range of our estimates here for morphological traits. Overall, it appears that neither source of potential downward bias (within family selection or heterozygous effects) has had a discernible impact on h_M^2 estimated in our outbred population. The magnitude of the mutational variance estimated from inbred populations therefore may be broadly representative of the variance generated by the effects of heterozygous mutations, although further studies in other taxa and traits are needed to confirm this.

It follows that estimates of selection coefficients derived from inbred lines also may reflect the magnitude of selection against heterozygous mutations, of particular interest in models of mutation-selection balance. To the best of our knowledge, our estimates of selection coefficients (Table 1) represent the first in which both the numerator and the denominator have been estimated in the same outbred population; previous estimates of selection coefficients of new mutations were the product of combining estimates of V_M and V_A (or V_G) from different experiments on distinct populations (*e.g.*, outbred for V_A and inbred for V_M) (Houle *et al.* 1996). The ratio V_M/V_A estimated from the same data in an outbred population reflects the heterozygous effects of new mutations in an outbred genetic background, which closely resembles how mutations would affect the phenotype when they first arise in a natural population. Our estimates have a median of 0.008, comparable to the median estimate for morphologic traits reported for a range of taxa (Houle *et al.* 1996),
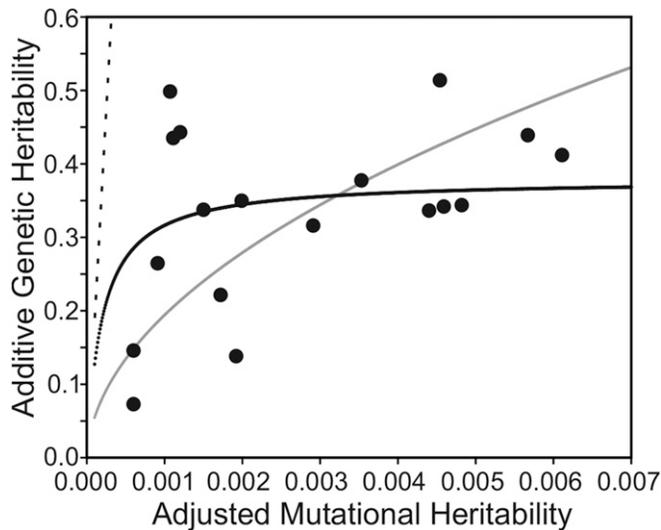
Figure 5 The observed (solid black circles) relationship between mutational (*x*-axis) and additive genetic (*y*-axis) heritability and the fit of the neutral drift (dashed black line), Gaussian approximation (gray) and house-of-cards (black) models.

suggesting an average persistence time of mutations affecting such traits of about 125 generations.

Limitations of the experimental design

Despite several lines of evidence that suggested we successfully captured the effects of mutation in an outbred population, we were unable to provide evidence for statistically significant mutational variance in any single trait. Similar problems have been encountered when applying the animal model to estimate additive genetic variation in wild populations; when estimating quantitative genetic parameters under natural conditions, data will be more variable and larger sample sizes are required to provide sufficient statistical power (Kruuk 2004). In our experiment, greater variability was introduced not by the environment (as in wild populations) but because we were focused on mutational variance while allowing standing genetic variation to persist in our experimental population, a key feature that allowed us to estimate mutational effects in the heterozygous state.

In general, and ignoring biases introduced by maternal and environmental effects (Clément *et al.* 2001; Kruuk and Hadfield 2007), in an animal model analysis of genetic variation, statistical power is determined by the size of the base population, accuracy is determined by the number of pedigree links between levels of the fixed effects (*i.e.*, generations in our case) (Kennedy and Trus 1993; Hanocq *et al.* 1996; Clément *et al.* 2001; Quinn *et al.* 2006), and precision is determined by the number of phenotypes per family (Quinn *et al.* 2006). All these properties of the experimental design are captured by the T-matrix and hence should apply to both additive genetic [Equation (2)] and mutational [Equation (4)] effects. For our particular experimental design, consisting of 120 base-population males with a pedigree maintained

for a further 21 generations and one individual measured per family, we favored accuracy, likely at the expense of power and precision. We expect that increasing the number of individuals sampled in the base population and measuring the phenotypes of multiple individuals from each family would result in greater power to detect effects and more precise estimates of mutational variance.

Mutation-selection balance in an outbred population

The processes maintaining genetic variation in natural populations remain one of the unresolved questions in evolution (Johnson and Barton 2005). Genetic variation might be maintained by balancing selection, favoring alternate alleles or phenotypes, or through the balance between the input of new mutations and their elimination via selection. Mutation is not an explicit parameter of balancing selection models, although it is the ultimate source of the genetic variance maintained under this mode of selection. Here we explored the implications of the observed magnitudes of standing and mutational variance for models of mutation-selection balance. Mutation-selection balance models of the maintenance of genetic variation have struggled to reconcile the general observations of strong stabilizing selection, high heritability, and high mutation in natural populations (Johnson and Barton 2005). Consistent with this view, both the Gaussian and house-of-cards models tended to overestimate the standing level of V_A in this population when compared to the data. Notably, we assumed strong selection ($V_s = 5$) compared to $V_s \geq 20$ that is used in most theoretical treatments (Johnson and Barton 2005) of mutation-selection balance. Weakening the strength of selection increases the equilibrium V_A predicted by these models, further worsening their fit to the observed data. Similarly, forcing the other two parameters that were free to vary, ℓ and U, to possibly more realistic values, say, $\ell > 10$ or $U = 0.1$ (Lynch and Walsh 1998), also increases the equilibrium level of V_A.

Johnson and Barton (2005) identified three different classes of mutation-selection balance models: direct stabilizing selection on individual traits, direct selection on multiple traits, and pure pleiotropy models (see also Zhang and Hill 2005). Having shown that direct selection on single traits is unlikely to adequately explain the estimated levels of additive and mutational genetic variance in this outbred population of *D. serrata*, we highlight two points concerning the other two classes of mutation-selection balance models. First, we can exclude the pure pleiotropy model as a good fit to these data based on our current estimates. Under the pure pleiotropy class of model, observed mutational heritability on the order of 0.001 implies that standardized stabilizing selection gradients on traits must be −0.001 ($V_s = 500$) or weaker to maintain *any* level of heritability (Johnson and Barton 2005, equation A25). While this strength of stabilizing selection has been observed on the genetic variance of some individual CHC traits, it varies considerably among traits and has been reported to be up to 100 times stronger on one CHC expressed in males of our study species (Delcourt *et al.* 2012). A pure pleiotropy model of mutation-selection balance is

therefore unable to explain the levels of heritability in this outbred population given the strength of stabilizing selection that has been observed on some of these traits.

Second, it is well known that the morphological traits analyzed here are not genetically independent of one another (Blows *et al.* 2004; Mezey and Houle 2005; McGuigan and Blows 2007; Houle and Fierst 2013; Hine *et al.* 2014), and such pleiotropy will tend to reduce the level of heritability that can be maintained under mutation-selection balance (Bürger 2000, p. 298). This suggests that a multivariate assessment of mutation-selection balance might be informative given that single-trait models generally overestimated the predicted heritability. An assessment of the multivariate extension of the direct selection model requires estimation of the additive and mutational variance-covariance matrices (Lande 1980), which we are currently pursuing for these traits.

How genetic variation is maintained in the presence of selection remains one of the fundamental unresolved issues in evolutionary biology and a key question given that mutation and stabilizing selection are ubiquitous evolutionary forces (Turelli 1984; Johnson and Barton 2005; Walsh and Blows 2009). Mutation-selection balance is one of the two leading explanations (along with balancing selection) for the levels of standing variance in natural populations (Barton 1990; Johnson and Barton 2005; Zhang and Hill 2005), but the data available on key mutation parameters are sparse and of low quality (Houle *et al.* 1996). Wray's insight (Wray 1990) has provided us with a way of estimating key mutational parameters in an outbred population. By combining these estimates of mutation genetic variance with estimates of standing genetic variance derived from the same population and statistical model, more detailed and internally consistent investigations of mutation-selection balance in outbred populations become feasible.

Acknowledgments

Literature Cited

Aguirre, J. D., M. W. Blows, and D. J. Marshall, 2014   The genetic covariance between life cycle stages separated by metamorphosis. Proc. Biol. Sci. 281: 20141091.

Aitchison, J., 1986   *The Statistical Analysis of Composition Data*. Chapman & Hall, London.

Barton, N. H., 1990   Pleiotropic models of quantitative variation. Genetics 124: 773–782.

Blows, M. W., and R. A. Allan, 1998   Levels of mate recognition within and between two *Drosophila* species and their hybrids. Am. Nat. 152: 826–837.

Blows, M. W., S. F. Chenoweth, and E. Hine, 2004   Orientation of the genetic variance-covariance matrix and the fitness surface for multiple male sexually selected traits. Am. Nat. 163: E329–E340.

Bürger, R., 2000   *The Mathematical Theory of Selection, Recombination, and Mutation*. Wiley, Chichester, UK.

Casellas, J., and J. R. Medrano, 2008   Within-generation mutation variance for litter size in inbred mice. Genetics 179: 2147–2155.

Casellas, J., G. Caja, and J. Piedrafita, 2010   Accounting for additive genetic mutations on litter size in Ripollesa sheep. J. Anim. Sci. 88: 1248–1255.

Chenoweth, S. F., H. D. Rundle, and M. W. Blows, 2010   The contribution of selection and genetic constraints to phenotypic divergence. Am. Nat. 175: 186–196.

Clément, V., B. Bibé, É. Verrier, J. M. Elsen, E. Manfredi *et al.*, 2001   Simulation analysis to test the influence of model adequacy and data structure on the estimation of genetic parameters for traits with direct and maternal effects. Genet. Sel. Evol. 33: 369–396.

Crow, J. F., and M. Kimura, 1970   *An Introduction to Population Genetics Theory*. Harper & Row, New York.

Delcourt, M., M. W. Blows, J. D. Aguirre, and H. D. Rundle, 2012   Evolutionary optimum for male sexual traits characterized using the multivariate Robertson-Price Identity. Proc. Natl. Acad. Sci. USA 109: 10414–10419.

Denver, D. R., K. Morris, J. T. Streelman, S. K. Kim, M. Lynch *et al.*, 2005   The transcriptional consequences of mutation and natural selection in *Caenorhabditis elegans*. Nat. Genet. 37: 544–548.

Frentiu, F. D., and S. F. Chenoweth, 2008   Polyandry and paternity skew in natural and experimental populations of *Drosophila serrata*. Mol. Ecol. 17: 1589–1596.

Gelman, A., 2006   Prior distributions for variance parameters in hierarchical models. Bayesian Anal. 1: 515–533.

Gelman, A., and D. B. Rubin, 1992   Inference from iterative simulation using multiple sequences. Stat. Sci. 7: 457–472.

Gosden, T. P., K. L. Shastri, P. Innocenti, and S. F. Chenoweth, 2012   The B-matrix harbors significant and sex-specific constraints on the evolution of multicharacter sexual dimorphism. Evolution 66: 2106–2116.

Haag-Liautard, C., M. Dorris, X. Maside, S. Macaskill, D. L. Halligan *et al.*, 2007   Direct estimation of per nucleotide and genomic deleterious mutation rates in *Drosophila*. Nature 445: 82–85.

Hadfield, J. D., 2010   MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. J. Stat. Softw. 33: 1–22.

Hadfield, J. D., 2014   MCMCglmm course notes; available at: cran.r-project.org/web/packages/MCMCglmm/vignettes/CourseNotes.pdf.

Hadfield, J. D., A. J. Wilson, D. Garant, B. C. Sheldon, and L. E. B. Kruuk, 2010   The misuse of BLUP in ecology and evolution. Am. Nat. 175: 116–125.

Halligan, D. L., and P. D. Keightley, 2009   Spontaneous mutation accumulation studies in evolutionary genetics. Annu. Rev. Ecol. Evol. Syst. 40: 151–172.

Hanocq, E., D. Boichard, and J. L. Foulley, 1996   A simulation study of the effect of connectedness on genetic trend. Genet. Sel. Evol. 28: 67–82.

Hansen, T. F., C. Pelabon, and D. Houle, 2011   Heritability is not evolvability. Evol. Biol. 38: 258–277.

Henderson, C. R., 1976   A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. Biometrics 32: 69–83.

Higgie, M., and M. W. Blows, 2007   Are traits that experience reinforcement also under sexual selection? Am. Nat. 170: 409–420.

Hine, E., K. McGuigan, and M. W. Blows, 2011 Natural selection stops the evolution of male attractiveness. Proc. Natl. Acad. Sci. USA 108: 3659–3664.

Hine, E., K. McGuigan, and M. W. Blows, 2014 Evolutionary constraints in high-dimensional trait sets. Am. Nat. 184: 119–131.

Houle, D., and J. Fierst, 2013 Properties of spontaneous mutational variance and covariance for wing size and shape in *Drosophila melanogaster*. Evolution 67: 1116–1130.

Houle, D., B. Morikawa, and M. Lynch, 1996 Comparing mutational variabilities. Genetics 143: 1467–1483.

Houle, D., C. Pelabon, G. P. Wagner, and T. F. Hansen, 2011 Measurement and meaning in biology. Q. Rev. Biol. 86: 3–34.

Johnson, T., and N. Barton, 2005 Theoretical models of selection and mutation on quantitative traits. Philos. Trans. R. Soc. Lond. B Biol. Sci. 360: 1411–1425.

Keightley, P. D., and W. G. Hill, 1992 Quantitative genetic variation in body size of mice from new mutations. Genetics 131: 693–700.

Keightley, P. D., T. F. C. Mackay, and A. Caballero, 1993 Accounting for bias in estimates of the rate of polygenic mutation. Proc. Biol. Sci. 253: 291–296.

Keightley, P. D., A. Caballero, and A. Garcia-Dorado, 1998 Population genetics: surviving under mutation pressure. Curr. Biol. 8: R235–R237.

Keightley, P. D., U. Trivedi, M. Thomson, F. Oliver, S. Kumar *et al.*, 2009 Analysis of the genome sequences of three *Drosophila melanogaster* spontaneous mutation accumulation lines. Genome Res. 19: 1195–1201.

Keightley, P. D., R. W. Ness, D. L. Halligan, and P. R. Haddrill, 2014 Estimation of the spontaneous mutation rate per nucleotide site in a *Drosophila melanogaster* full-sib family. Genetics 196: 313–320.

Keightley, P. D., A. Pinharanda, R. W. Ness, F. Simpson, K. K. Dasmahapatra *et al.*, 2015 Estimation of the spontaneous mutation rate in Heliconius melpomene. Mol. Biol. Evol. 32: 239–243.

Kennedy, B. W., and D. Trus, 1993 Considerations on genetic connectedness between management units under an animal model. J. Anim. Sci. 71: 2341–2352.

Kingsolver, J. G., H. E. Hoekstra, J. M. Hoekstra, D. Berrigan, S. N. Vignieri *et al.*, 2001 The strength of phenotypic selection in natural populations. Am. Nat. 157: 245–261.

Kingsolver, J. G., S. E. Diamond, A. M. Siepielski, and S. M. Carlson, 2012 Synthetic analyses of phenotypic selection in natural populations: lessons, limitations and future directions. Evol. Ecol. 26: 1101–1118.

Kondrashov, A. S., and M. Turelli, 1992 Deleterious mutations, apparent stabilizing selection and the maintenance of quantitative variation. Genetics 132: 603–618.

Kruuk, L. E. B., 2004 Estimating genetic parameters in natural populations using the "animal model." Philos. Trans. R. Soc. Lond. B Biol. Sci. 359: 873–890.

Kruuk, L. E. B., and J. D. Hadfield, 2007 How to separate genetic and environmental causes of similarity between relatives. J. Evol. Biol. 20: 1890–1903.

Lande, R., 1975 Maintenance of genetic variability by mutation in polygenic character with linked loci. Genet. Res. 26: 221–235.

Lande, R., 1980 The genetic covariances between characters maintained by pleiotropic mutations. Genetics 94: 203–215.

Lande, R., 1995 Mutation and conservation. Conserv. Biol. 9: 782–791.

Lynch, M., and W. G. Hill, 1986 Phentoypic evolution by neutral mutation. Evolution 40: 915–935.

Lynch, M., and R. Lande, 1993 Evolution and extinction in response to environmental change, pp. 234–250 in Biotic Interactions and Global Climate Change, edited by P. M. Kareiva,

J. G. Kingsolver, and R. B. Huey. Sinauer Associates, Sunderland, MA.

Lynch, M., and B. Walsh, 1998 *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Sunderland, MA.

Lynch, M., J. Blanchard, D. Houle, T. Kibota, S. Schultz *et al.*, 1999 Spontaneous deleterious mutation. Evolution 53: 645–663.

Mack, P. D., V. K. Lester, and D. E. L. Promislow, 2001 Age-specific effects of novel mutations in *Drosophila melanogaster*. II. Fecundity and male mating ability. Genetica 110: 31–41.

McGuigan, K., and M. W. Blows, 2007 The phenotypic and genetic covariance structure of Drosophilid wings. Evolution 61: 902–911.

McGuigan, K., and M. W. Blows, 2013 Joint allelic effects on fitness and metric traits. Evolution 67: 1131–1142.

McGuigan, K., L. Rowe, and M. W. Blows, 2011 Pleiotropy, apparent stabilizing selection and uncovering fitness optima. Trends Ecol. Evol. 26: 22–29.

McGuigan, K., J. M. Collet, S. L. Allen, S. F. Chenoweth, and M. W. Blows, 2014 Pleiotropic mutations are subject to strong stabilizing selection. Genetics 197: 1051–1062.

Mezey, J. G., and D. Houle, 2005 The dimensionality of genetic variation for wing shape in *Drosophila melanogaster*. Evolution 59: 1027–1038.

Moorad, J. A., and D. W. Hall, 2009 Mutation accumulation, soft selection and the middle-class neighborhood. Genetics 182: 1387–1389.

Mousseau, T. A., and D. A. Roff, 1987 Natural selection and the heritability of fitness components. Heredity 59: 181–197.

Partridge, L., A. Hoffmann, and J. S. Jones, 1987 Male size and mating success in *Drosophila melanogaster* and *D. pseudoobscura* under field conditions. Anim. Behav. 35: 468–476.

Quinn, J. L., A. Charmantier, D. Garant, and B. C. Sheldon, 2006 Data depth, data completeness, and their influence on quantitative genetic estimation in two contrasting bird populations. J. Evol. Biol. 19: 994–1002.

Robertson, F. W., and E. Reeve, 1952 Studies in quantitative inheritance. I. The effects of selection of wing and thorax length in *Drosophila melanogaster*. J. Genet. 50: 414–448.

Roff, D. A., and T. A. Mousseau, 1987 Quantitative genetics and fitness: lessons from *Drosophila*. Heredity 58: 103–118.

Rohlf, F. J., 1999 Shape statistics: procrustes superimpositions and tangent spaces. J. Classif. 16: 197–223.

Roles, A. J., and J. K. Conner, 2008 Fitness effects of mutation accumulation in a natural outbred population of wild radish (*Raphanus raphanistrum*): comparison of field and greenhouse environments. Evolution 62: 1066–1075.

Rundle, H. D., S. F. Chenoweth, P. Doughty, and M. W. Blows, 2005 Divergent selection and the evolution of signal traits and mating preferences. PLoS Biol. 3: 1988–1995.

Santiago, E., J. Albornoz, A. Dominguez, M. A. Toro, and C. Lopez-Fanjul, 1992 The distribution of spontaneous mutations on quantitative traits and fitness in *Drosophila melanogaster*. Genetics 132: 771–781.

SAS Institute, 2011 SAS 9.3. Cary, NC.

Schultz, S. T., and M. Lynch, 1997 Mutation and extinction: the role of variable mutational effects, synergistic epistasis, beneficial mutations, and degree of outcrossing. Evolution 51: 1363–1371.

Shabalina, S. A., L. Y. Yampolsky, and A. S. Kondrashov, 1997 Rapid decline of fitness in panmictic populations of *Drosophila melanogaster* maintained under relaxed natural selection. Proc. Natl. Acad. Sci. USA 94: 13034–13039.

Sztepanacz, J. L., and H. D. Rundle, 2012 Reducing genetic variance among high fitness individuals: Inferring stabilizing selection on male sexual displays in *Drosophila serrata*. Evolution 66: 3101–3110.

Turelli, M., 1984 Heritable genetic variation via mutation selection balance: Lerch's zeta meets the abdominal bristle. Theor. Popul. Biol. 25: 138–193.

Walsh, B., and M. W. Blows, 2009 Abundant genetic variation + strong selection = multivariate genetic constraints: a geometric view of adaptation. Annu. Rev. Ecol. Evol. Syst. 40: 41–59.

Whitlock, M. C., 2000 Fixation of new alleles and the extinction of small populations: drift load, beneficial alleles, and sexual selection. Evolution 54: 1855–1861.

Wilson, A. J., D. Reale, M. N. Clements, M. M. Morrissey, E. Postma *et al.*, 2010 An ecologist's guide to the animal model. J. Anim. Ecol. 79: 13–26.

Wray, N. R., 1990 Accounting for mutation effects in the additive genetic variance covariance matrix and its inverse. Biometrics 46: 177–186.

Yampolsky, L. Y., L. E. Pearse, and D. E. L. Promislow, 2001 Age-specific effects of novel mutations in *Drosophila melanogaster*. I. Mortality. Genetica 110: 11–29.

Zhang, X. S., and W. G. Hill, 2002 Joint effects of pleiotropic selection and stabilizing selection on the maintenance of quantitative genetic variation at mutation-selection balance. Genetics 162: 459–471.

Zhang, X. S., and W. G. Hill, 2005 Genetic variability under mutation selection balance. Trends Ecol. Evol. 20: 468–470.

Communicating editor: C. D. Jones

GENETICS

Simultaneous Estimation of Additive and Mutational Genetic Variance in an Outbred Population of *Drosophila serrata*

Katrina McGuigan, J. David Aguirre, and Mark W. Blows

Table S1. Correlation between estimates of V_A and V_M from the observed data, and the null model in which the pedigree was randomized. Traits are defined in Figure 1.

Trait	Observed Data	Null Model
Z,Z–5,9–$C_{25:2}$	-0.288	-0.001
Z–9–$C_{25:1}$	-0.238	0.026
Z–9–$C_{26:1}$	-0.590	-0.035
2–Me–C_{25}	-0.612	-0.028
Z,Z–5,9–$C_{27:2}$	-0.487	-0.045
2–Me–C_{28}	-0.483	0.011
Z,Z–5,9–$C_{29:2}$	-0.661	-0.059
2–Me–C_{30}	-0.652	-0.043
Centroid Size	-0.868	-0.040
ILD 1–8	-0.742	0.011
ILD 2–4	-0.500	-0.062
ILD 2–5	-0.741	-0.029
ILD 2–8	-0.908	-0.029
ILD 3–7	-0.864	-0.046
ILD 3–9	-0.855	-0.037
ILD 4–5	-0.871	0.008
ILD 5–6	-0.943	-0.100
ILD 5–8	-0.905	-0.090
ILD 7–9	-0.912	-0.021

Table S2. Median estimates (and 95% confidence intervals) (x1000) of variance components for additive genetic (V_A), mutational (V_M) and total phenotypic (V_P) variance from the null model and the observed data. Note, the phenotypic data was identical between the two datasets, but pedigree information was randomised in the null data. As V_P is therefore the same in the two datasets, only the median of the 1000 iterations of the observed data is shown.

Trait	V_A Null	V_A Observed	V_M Null	V_M Observed	V_P
Z,Z–5,9–$C_{25:2}$	0.2555	0.7754*	0.00232	0.00553	5.3260
	(0.1266 - 0.5050)	(0.5052 – 1.0846)	(0.00000 - 0.02410)	(0.00001 - 0.06242)	(5.0993 – 5.5363)
Z–9–$C_{25:1}$	0.0007278	1.1119	0.01035	0.01974	15.3449
	(0.3591 – 1.4126)	(0.5841 – 1.9374)	(0.00001 - 0.11958)	(0.00004 - 0.22260)	(14.6847 – 15.9783)
Z–9–$C_{26:1}$	0.6397	1.8605	0.00974	0.03597	13.4905
	(0.3322 – 1.2223)	(0. 9100 - 0.0026449)	(0.00002 - 0.10169)	(0.00007 - 0.42070)	(12.7015 – 14.0730)
2–Me–C_{25}	0.9402	8.7760*	0.01277	0.03648	19.8280
	(0.4639 - 0.0018208)	(6.6495 - 10.2711)	(0.00002 - 0.13630)	(0.00007 - 0.67487)	(18.6527 – 20.8207)
Z,Z–5,9–$C_{27:2}$	0.6328	6.3428*	0.00898	0.02244	12.7681
	(0.3187 – 1.2930)	(5.3029 - 0.0073773)	(0.00002 - 0.08130)	(0.00005 - 0.31425)	(12.0653 – 13.3346)
2–Me–C_{28}	0.6967	3.7755*	0.00934	0.02231	14.1896
	(0.3361 – 1.3762)	(2.6604 - 4.6274)	(0.00001 - 0.09250)	(0.00006 - 0.31311)	(13.4744 – 14.7522)

Z,Z–5,9–C$_{29:2}$	0.8973	4.1439*	0.01226	0.04492	18.7076
	(0.4322 – 1.6235)	(2.1588 – 5.4126)	(0.00003 - 0.13952)	(0.00008 - 0.68827)	(17.3659 - 19.5281)
2–Me–C$_{30}$	0.9045	6.1084*	0.01380	0.04073	18.0120
	(0.4500 – 1.7195)	(4.2395 – 7.3787)	(0.00002 - 0.17019)	(0.00010 - 0.61536)	(16.8318 - 18.7980)
Centroid Size	0.0776	0.5113*	0.00124	0.00814	1.4891
	(0.0385 - 0.1536)	(0.1836 - 0.6582)	(0.00000 - 0.01450)	(0.00001 - 0.12860)	(1.2636 – 1.5714)
ILD 1–8	0.0127	0.0849*	0.00023	0.00072	0.2420
	(0.0065 - 0.0259)	(0.0506 - 0.1045)	(0.00000 - 0.00232)	(0.00000 - 0.01245)	(0.2198 - 0.2535)
ILD 2–4	0.0178	0.1516*	0.00027	0.00066	0.3489
	(0.0083 - 0.0341)	(0.1185 - 0.1798)	(0.00000 - 0.00313)	(0.00000 - 0.00932)	(0.3281 - 0.3649)
ILD 2–5	0.0321	0.1949*	0.00051	0.00230	0.6137
	(0.0153 - 0.0631)	(0.0965 - 0.2475)	(0.00000 - 0.00567)	(0.00001 - 0.03507)	(0.5532 - 0.6458)
ILD 2–8	0.0502	0.3970*	0.00074	0.00663	0.9600
	(0.0257 - 0.1036)	(0.1171 - 0.4934)	(0.00000 - 0.01020)	(0.0002 - 0.10964)	(0.7714 – 1.0141)
ILD 3–7	0.0364	0.2368*	0.00063	0.00377	0.7018
	(0.0179 - 0.0707)	(0.0863 - 0.3049)	(0.00000 - 0.00728)	(0.00000 - 0.05793)	(0.6005 - 0.7422)
ILD 3–9	0.0492	0.3681*	0.00075	0.00425	0.9764
	(0.0238 - 0.0992)	(0.1619 - 0.4548)	(0.00000 - 0.00895)	(0.00000 - 0.07924)	(0.8375 – 1.0293)

ILD 4–5	0.0333	0.2252*	0.00058	0.00371	0.6529
	(0.0162 - 0.0664)	(0.0769 - 0.2855)	(0.00000 - 0.00659)	(0.00001 - 0.05762)	(0.5560 - 0.6877)
ILD 5–6	0.0137	0.0975	0.00022	0.00473	0.2670
	(0.0068 - 0.0275)	(0.0183 - 0.1301)	(0.00000 - 0.00277)	(0.00001 - 0.03404)	(0.2118 - 0.2872)
ILD 5–8	0.0243	0.2179*	0.00038	0.00325	0.4940
	(0.0118 - 0.0492)	(0.0697 - 0.2631)	(0.00000 - 0.00451)	(0.00001 - 0.05439)	(0.3947 - 0.5214)
ILD 7–9	0.0182	0.1897*	0.00028	0.00193	0.3677
	(0.0089 - 0.0365)	(0.0740 - 0.2261)	(0.00000 - 0.00368)	(0.00000 - 0.04332)	(0.2881 - 0.3890)

*significant at $P<0.05$