An investigation of the sex-specific genetic architecture of fitness in *Drosophila melanogaster*

Amardeep Singh, Asad Hasan and Aneil F. Agrawal[1]

Department of Ecology and Evolutionary Biology, University of Toronto, 25 Willcocks Street, Toronto, Ontario, Canada, M5S 3B2

[1]*E-mail: a.agrawal@utoronto.ca*

Abstract

In dioecious populations, the sexes employ divergent reproductive strategies to maximize fitness and, as a result, genetic variants can affect fitness differently in males and females. Moreover, recent studies have highlighted an important role of the mating environment in shaping the strength and direction of sex-specific selection. Here, we measure adult fitness for each sex of 357 lines from the *Drosophila* Synthetic Population Resource (DSPR) in two different mating environments. We analyze the data using three different approaches to gain insight into the sex-specific genetic architecture for fitness: classical quantitative genetics, genomic associations, and a mutational burden approach. The quantitative genetics analysis finds that, on average segregating genetic variation in this population has concordant fitness effects both across the sexes and across mating environments. We do not find specific genomic regions with strong associations with either sexually antagonistic (SA) or sexually concordant (SC) fitness effects, yet there is modest evidence of an excess of genomic regions with weak associations, both with SA and SC fitness effects. Our examination of mutational burden indicates stronger selection against indels and loss-of-function variants in females than males.

Introduction

The genetic architecture of fitness variation (e.g., the number and identity of loci, and their effect sizes in different contexts) remains poorly understood. Given that any locus in the genome has the potential to contribute to fitness, the genetic architecture of fitness is likely highly polygenic. However, this does not preclude the possibility that some variants have reasonably large effects. Investigating the genetic architecture of fitness may provide useful clues as to the forces responsible for the maintenance of segregating fitness variants (e.g., mutation-selection balance or various forms of balancing selection; Mackay 2001; Barton and Keightley 2002; Frankham 2009)

In dioecious populations, the genetic architecture of phenotypic traits, including fitness may differ substantially between the sexes. Sex differences in the genetic architecture of fitness depend on the sex-specific effect size and direction of segregating variants. For example, some variants may have sex-limited effects because they are sex-linked or contribute to the expression of sex-limited traits (e.g., seminal fluid proteins). Such variants can be important contributors to the variance in fitness in one sex but may have little or no influence on fitness in the other. Variants with purely sex-limited fitness effects may be relatively rare. Instead, alleles may often contribute to variance in fitness in both sexes. The evolutionary fate of such variants will depend on the strength and direction of selection they experience when expressed in each sex. In some cases, alleles at a locus may have opposing, sexually antagonistic, effects on fitness resulting in so-called 'intra-locus sexual conflict' (reviewed in Bonduriansky and Chenoweth 2009). Under specific conditions, such variants can be maintained by balancing selection, particularly when the effect sizes of variants are large (Kidwell et al. 1977; Connallon and Clark 2012). However, most genes are thought to serve similar functions in both sexes and disruptions to those functions are likely deleterious for both sexes (Rowe and Houle 1996; Whitlock and Agrawal 2009). While sexually

concordant selection should remove genetic variation at such genes (Long et al. 2012; Connallon and Clark 2013), this class of loci may serve as a large mutational target and thus contribute substantially to the standing genetic variation in fitness via mutation-selection balance.

The primary objective of this study is to better understand fitness variation within and across the sexes. We do so by measuring the "adult fitness" (i.e., fitness conditional on reaching the adult stage) of outbred *Drosophila melanogaster*, constructed from 357 recombinant inbred line (RIL) haplotypes from the *Drosophila* Synthetic Population Resource (DSPR; King et al. 2012). Ideally, one would catalog every variant and measure its fitness effect in each sex, but that ideal is not logistically feasible. Instead, we pursue three different approaches that each offer a different lens through which to view the relationship of the genetic architecture for fitness across the sexes.

First, we used a classic quantitative genetics approach to provide a summary of how genetic architecture of fitness is related between the sexes. This approach has been used in a number of laboratory (e.g., Vieira et al. 2000; Wayne et al. 2001; Pischedda and Chippindale 2006; Long et al. 2009; Collet et al. 2016) and wild (e.g., Kruuk et al. 2000; Merilä and Sheldon 2000; McCleery et al. 2004; Foerster et al. 2007) populations and is useful for asking how much additive genetic variation in sex-specific fitness exists in a population. In addition, we used this approach to estimate the intersexual genetic correlation for fitness (r_{mf}), which serves as a useful indicator of the overall concordance or antagonism of segregating genetic variation in a population. This correlation can be used to quantify the extent to which improvement in the fitness of one sex will help or hinder the improvement of the other, at least in the short term. Empirical estimates of r_{mf} have been mixed: although a survey of animal species found that, on average, estimates of r_{mf} were small but positive (Poissant et al. 2010), several studies have reported negative estimates of

104 r_{mf} of substantial magnitude (e.g., Chippindale et al. 2001; Fedorka and Mousseau 2004;
105 Brommer et al. 2007; Punzalan et al. 2014). The power and limitation of the intersexual correlation
106 is that it is a summary statistic that is predictive of short-term evolution that neither requires nor
107 reveals information about individual allelic effects. Indeed, the underlying 'genetic details' matter
108 little with respect to short-term evolution but can substantially affect evolutionary outcomes on
109 longer time-scales.

110 Next, we used genomic association mapping techniques in an attempt to gain resolution on
111 specific loci underlying sex-specific fitness. Association mapping provides a route to identifying
112 individual allelic effects and, in principle, can be used to find the number and magnitude of
113 mutations with sexually concordant, antagonistic, and sex-limited effects. Recognizing that the
114 genetic basis of fitness is likely to be highly polygenic, we are unlikely to have the statistical power
115 to identify a multitude of small effect variants. However, we were particularly interested in
116 identifying loci contributing to sexual antagonism and such variants maintained by balancing
117 selection are expected to be of larger effects (Kidwell et al. 1977; Connallon and Clark 2012),
118 possibly detectable via genome association mapping. This approach has recently proven fruitful
119 in identifying candidate sexually antagonistic loci in this species (Ruzicka et al. 2019).

120 Though association mapping is severely limited in its ability to identify small effect
121 variants, other approaches can provide insights. When sequence and fitness data are both available
122 for a collection of genotypes, then one approach is to quantify the aggregate effect of specified
123 classes that are suspected to affect fitness *a priori* (Yang et al. 2017; Brown and Kelly 2019). For
124 example, one can use this approach to ask how the total number of loss-of-function variants (LoFs)
125 in a genome affects fitness, which is equivalent to quantifying the average strength of selection
126 against segregating LoFs. We are particularly interested in comparing the selection strength in

127 males versus females. Strong competition for access to females has been postulated to increase the
128 strength of selection on males relative to females on sexually concordant variants (reviewed
129 Whitlock and Agrawal 2009), potentially increasing population mean fitness by reducing mutation
130 load (Manning 1984; Agrawal 2001; Siller 2001) and accelerating adaptation (Lorch et al. 2003).
131 Experimental studies using phenotypic marker mutations (e.g., Whitlock and Bourguet 2000;
132 Pischedda and Chippindale 2005; Stewart et al. 2005; Sharp and Agrawal 2008; Hollis et al. 2009)
133 and mutation accumulation lines (e.g., Enders and Nunney 2010; Mallet et al. 2011; Mcguigan et
134 al. 2011; Sharp and Agrawal 2013; Grieshop et al. 2016) tend to support the notion of stronger
135 selection in males than females, though there are potential concerns with these approaches (see
136 Discussion).

137 In addition to examining sex differences in the genetic architecture of fitness, the second
138 goal of this study was to investigate how the genetic architecture of fitness of each sex, as well as
139 the relationship of these architectures across the sexes, was affected by the ecological context in
140 which sexual interactions occur. Here, we estimated adult fitness of males and females in two
141 mating environments: traditional fly culturing vials and larger cages that represent lower density
142 and more structurally complex environments. Previous work has shown males and females interact
143 differently in environments similar to the two used here (Yun et al. 2017; MacPherson et al. 2018).
144 Male harm and sexual conflict are strong features of sexual interactions in standard fly vials but
145 not in the more complex cages; these differences affect the rates of nonsexual adaptation and
146 purging (Singh et al. 2017; Yun et al. 2018).

147 Our quantitative genetic analysis recovers appreciable levels of additive genetic variation
148 in adult fitness for both male and female fitness and finds that the fitness consequences of this
149 genetic variation tends to be concordant across the sexes as well as across environments. Our

150 association mapping analysis did not reveal any statistically convincing associations between
151 fitness and any specific genomic region. However, there is some evidence that both sexually
152 concordant and sexually antagonistic regions exist. Our mutation burden analysis recovers a
153 significant negative association between the burden of putatively deleterious variants and female,
154 but not male, fitness in both mating environments.

155

156 **Methods**

157 DSPR Inbred Genotypes

158 Recombinant inbred line (RIL) genotypes used for measuring sex-specific outbred fitness
159 were obtained from the *Drosophila* Synthetic Population Resource (DSPR). The creation of these
160 lines has been described in detail elsewhere (King et al. 2012a). Briefly, eight parental genotypes
161 from geographically disparate regions were crossed in a round robin design to generate an outbred
162 population that was allowed to recombine for 50 generations. Inbred genotypes were generated
163 from this outbred, synthetic population following 25 generations of fullsib mating. Following low-
164 resolution RADseq genotyping of inbred genotypes, a hidden Markov model (HMM) was
165 implemented to determine the probable parental ancestry at regularly spaced 10kb intervals (or
166 haplotype blocks) along the genome of each line. The parental ancestry of the vast majority of
167 10kb blocks in most lines can be assigned to a parental source with high certainty. Here, we
168 estimated sex-specific fitness of 357 lines. Although the majority of lines were assayed for fitness
169 in both sexes and both environments, this was not true of all lines for logistical reasons (see Table
170 S1 for number of lines assayed per sex and environment).

171 Measuring Adult Fitness of DSPR Genotypes

172 Our goal was to obtain a reasonably comprehensive measure of "adult fitness" (i.e., fitness
173 conditional on reaching the adult stage) with respect to the contexts we use for these assays. (See
174 Fig. S1 for a schematic summary of the assay procedure.) To measure sex-specific outbred adult
175 fitness of focal genotypes we first crossed females of each focal genotype to males from a common
176 outbred population that was generated by mixing all DSPR founder genotypes. Focal males and
177 females were then assayed in separate competitive fitness assays. Specifically, for each outbred
178 genotype, we combined four focal males/females with twelve competitors of the same sex from
179 an outbred laboratory population homozygous for a dominant bw^D marker. Focal and competitor
180 genotypes competed for access to sixteen unmarked individuals of the opposite sex that came from
181 an outbred laboratory population. For logistical reasons the source population for these individuals
182 differed between male and female fitness assays: in male fitness assays we used females from an
183 outbred laboratory population originally captured from a West African population (*Dahomey*
184 stock) while in the female fitness assays we used males from an outbred stock generated by mixing
185 DSPR founder genotypes.

186 All flies were collected at approximately the same age (~2-3 days post eclosion) and held
187 separately by sex in vials sprinkled with live yeast in groups of sixteen for 24 hours prior to the
188 start of fitness assays. All flies for a replicate were then combined and allowed to interact for 6
189 days in either a standard *Drosophila* culture vials (28 x 95 mm, approximately 60 mL) that
190 contained 7ml of food sprinkled with live yeast or a larger, complex mating cages. The cages were
191 polypropylene ULINE® deli containers (114.3 x 76.2 mm, approximately 500 mL) with two food
192 cups containing 7mL of food each, sprinkled with live yeast. Replicates were maintained in a
193 12L:12D light cycle at 25°C. Fresh media, sprinkled with live yeast was provided on day one and
194 three (fresh media was not provided on day three in cage assays due to logistical reasons). On day

195 six, all females were removed and, to reduce subsequent larval densities, divided approximately
196 evenly into two vials sprinkled with live yeast and allowed to oviposit for ~24 hours. Starting ten
197 days later, and repeating for a total of three days, offspring from the two oviposition vials were
198 collected, pooled and frozen daily. The vast majority of offspring emerged on the first two days of
199 collection. Following collection, we randomly sampled at least 100 offspring and calculated the
200 proportion of total offspring with wildtype (red) eye colour in this sample. We have five replicate
201 assays for most combinations of line, sex, and environment. Due to human error, some lines have
202 more than five replicates while others have fewer. Fitness assays were conducted in a blocked
203 design.

204 Though we consider these measures to be a reasonable representation of "adult fitness",
205 they are indirectly affected by juvenile survivorship because we assess the fitness of adult males
206 and adult females by measuring the number of offspring that survive to the adult stage. However,
207 we expect variation in juvenile survival to make only a small contribution to the genetic variance
208 in our "adult fitness" measures for several reasons. First, we attempted to keep offspring density
209 from being high so that survivorship should be high. Second, the contribution of the genetic
210 variance in offspring survivorship to the variance in focal adult fitness will be diluted by 50%
211 because offspring only have half the genes from the focal parent. Third, previous studies have
212 indicated that genetic variation for juvenile survivorship is small, especially compared to that of
213 adult fitness components (e.g., Leips and Mackay 2000; Sharp and Agrawal 2018). Though we
214 expect the influence of juvenile fitness effects to be small, to the extent they contribute to the
215 genetic variance, they will upwardly bias estimates of the intersexual genetic correlation (r_{mf})
216 because both surviving sons and daughters contribute to fitness measures of each sex.

217 <u>Quantitative Genetic Analysis of Fitness in the DSPR</u>

To estimate the sex-specific heritability in fitness in each mating environment, we fit separate mixed models for each mating environment to estimate the sex-specific additive genetic and residual variances. We modeled the proportion of offspring descended from focal (rather than competitor) flies as a binomial variable using *MCMCglmm* (Hadfield 2010). Sex was included as a fixed effect whereas block and RIL genotype were included as random effects. The model assumed the two sexes had separate and independent variances due to block (as the sexes were assayed in separate blocks with different competitors). On the logit scale, the model estimates RIL effects via a bivariate Gaussian distribution $N(0, \boldsymbol{G})$ where

$$\boldsymbol{G} = \begin{bmatrix} \sigma^2_{RIL:m} & \sigma_{RIL:m,f} \\ \sigma_{RIL:m,f} & \sigma^2_{RIL:f} \end{bmatrix}$$

is a genetic covariance matrix. The diagonal elements are the variances among RILs for male and female fitness, which are proportional to the additive genetic variances; the off-diagonal is the among-RIL intersexual covariance, which is proportional to the additive genetic covariance. (Our estimated matrix is proportional, rather than equal, to the traditional $\boldsymbol{G}$ matrix because our unit of observation is the fitness of groups of four focal individuals sharing a RIL haploid genome rather than a single individual; see Supplemental Material for derivation). *MCMCglmm* allows for additive overdispersion via inclusion of a random effect (on the logit scale) for each observation; we assumed the sexes had independent variances in this regard. Models were fit using extended parameter priors (e.g., *V*=*diag*(2), *nu*=1, *alpha.mu* = c(0, 0), *alpha.V* = 1000**diag*(2)), which are recommended when genetic variances are small (de Villemereuil 2018). (As a simple check, we permuted the replicate fitness values with respect to RIL identity within each block for each sex and ran the same analysis; inferred genetic variances were an order of magnitude smaller.) Because

239 the values returned from *MCMCglmm* are on the link scale, we used *QGglmm* (de Villemereuil *et*
240 *al* 2016) to obtain estimates on the observed scale.

241 Posterior distributions for parameters of interest were made from the MCMC chain of
242 model terms, following conversion to the observed scale and any required subsequent calculation
243 (e.g., as for heritability). Because we measure fitness from replicates containing four individuals,
244 the additive genetic variance for sex j is $\sigma^2_{A:j} = \sigma^2_{RIL:j}/8$ where $\sigma^2_{RIL:j}$ is the variance (on the
245 observed scale) across all replicates associated with RIL identity (see Supplementary Material).
246 Heritability in sex j is given by

247
$$h^2_j = \frac{\sigma^2_{RIL:j}}{2\left(\sigma^2_{e:j} + \frac{\sigma^2_{RIL:j}}{4}\right)}$$

248 where $\sigma^2_{e:j}$ is the non-genetic variance across all replicates for sex j. Evolvability for sex j was
249 calculated as $e_{\mu:j} = \sigma^2_{A:j}/\mu^2_j$ where μ_j is the mean adult fitness measure of an individual of sex j.
250 The intersexual genetic correlation for fitness (r_{mf}) is calculated as

251
$$r_{mf} = \frac{\sigma_{RIL:m,f}}{\sqrt{\sigma^2_{RIL:m}\,\sigma^2_{RIL:f}}}$$

252 Reported point estimates are means of the posterior and 95% credibility intervals (CI) were
253 calculated using the *HPDInterval* function in *MCMCglmm*.

254 Though our estimates of r_{mf} are intended to represent the intersexual genetic correlation
255 with respect to "adult" fitness, they are indirectly affected by juvenile survivorship because we
256 assess the fitness of adult males and adult females by measuring the number of offspring that
257 survive to the adult stage. Genetic variation for offspring survivorship will upwardly bias our

258 estimate of the intersexual correlation for adult fitness. However, we expect this effect to be small
259 for several reasons. First, we attempted to keep offspring density from being high so that
260 survivorship should be high. Second, the contribution of the genetic variance in offspring
261 survivorship to the variance in focal adult fitness will be diluted by 50% because offspring only
262 have half the genes from the focal parent. Third, previous studies have indicated that genetic
263 variation for juvenile survivorship is small, especially compared to that of adult fitness components
264 (e.g., Leips and Mackay 2000; Sharp and Agrawal 2018).

265 As a separate question, we were interested in the relationship between the genetic
266 architectures of a given sex across the two mating environments. Using the fitness data for a single
267 sex from each of the two environments, we employed a similar approach to estimate the within-
268 sex cross-environment genetic correlation for fitness (i.e., $r_{env,m}$ and $r_{env,f}$ for males and females,
269 respectively).

270 Association Mapping of Fitness and Indices of Sexual Antagonism and Concordance

271 We briefly outline methods for genome association mapping and provide details below. To
272 identify genome associations with sex-specific fitness as well as loci underpinning sexually
273 antagonistic and concordant variation, we first identified univariate traits for fitness and indices of
274 sexual antagonism and concordance. Next, we defined a statistical model for mapping genotypic
275 variation to trait variation and a significance threshold was established to identify loci putatively
276 associated with fitness or axes of antagonism/concordance.

277 To map phenotypic variation in male and female fitness we calculated the mean relative
278 fitness, separately for each sex, of each RIL. To associate loci with sexual antagonism and sexual
279 concordance, we defined an 'axis of antagonism' (i.e., SA Axis) and an 'axis of concordance' (i.e.,

280 SC Axis) following the method by Berger et al. (2014), which we illustrate in Fig. S2. The *SC*
281 value of a RIL is computed as $SC = (w_f + w_m)/\sqrt{2}$ where w_f and w_m are the estimates of mean
282 female and male relative fitness. This value represents the distance of a genotype from mean in
283 the bivariate space of male and female relative fitnesses in the direction of equal male and female
284 fitness effects (i.e., the 1 to 1 line). RILs with above (below) average fitness for both sexes are
285 expected to have positive (negative) values of SC. The *SA* value of a RIL is calculated as $SA =$
286 $(w_f - w_m)/\sqrt{2}$. This value represents the distance of a genotype from mean in the direction of
287 equal but opposite female and male fitness effects (i.e., the $w_f = -w_m$ line). RILs with above (below)
288 average values for female fitness but the opposite for male fitness are expected to have positive
289 (negative) values of SA. Our primary interest was with respect to the SA index, but we
290 independently mapped four traits—SA index and SC index, as well as male and female fitnesses—
291 separately for each environment (i.e., 8 total traits).

292 A feature of the DSPR is that the mosaic haplotype structure of each RIL is known to a
293 high degree of certainty in most cases. Specifically, at 10kb intervals along the genome of each
294 RIL, the additive probability of each of the eight parental genotypes is known from a Hidden
295 Markov Model (King et al. 2012a). To associate loci with each of the eight traits we regressed, at
296 each 10kb haplotype block, the RIL trait values against the additive probability for each of the
297 eight parental lines. Specifically, we fit the mixed model:

298 $$\bar{Z}_i = \mu + \sum_{j=1}^{l} \beta_j G_{ij} + \psi_i + e_{ij}$$

299 where $\bar{Z}_i$ is the mean trait value of the *i*th RIL and μ is the grand trait mean across all sampled
300 RILs, G_{ij} is fixed effect of the additive probability of the haplotype block of the *i*th RIL having

301 descended from parental genotype j, β_j is the additive genetic effect of the jth parental genotype
302 on the trait, ψ_i is the random effect of RIL genotype that is assumed to follow a Gaussian
303 distribution such that $\psi_i \sim \sigma_{kl} \sim N(0, \boldsymbol{K})$ where $\boldsymbol{K}$ is the relatedness matrix of the DSPR lines (see
304 below), l is the number of parental genotypes (max = 8), that contributed a cumulative additive
305 probability of at least 5% at each haplotype block i across all RIL genotypes (see below) and e_{ij}
306 specifies the residual variance of the ith RIL at the jth haplotype block and is assumed to follow a
307 Gaussian distribution with a mean of zero and unknown variance. At each 10kb haplotype block
308 i, we excluded any RIL genotypes whose most probable parental ancestor did not have a
309 cumulative additive probability of at least 5% among all RIL genotypes. *P*-values were obtained
310 from a likelihood ratio test against a null model where $\beta_j = 0$ for all j.

311 To estimate the genetic relatedness matrix of lines assayed, we calculated the average
312 pairwise relatedness for all 357 inbred genotypes over all haplotype blocks. Specifically, for each
313 pair of genotypes i and j we calculated the genetic relatedness (κ_{ij}) as

314 $$\kappa_{ij} = \frac{\sum_{n=1}^{m} \sum_{k=1}^{8} p_{ikn} p_{jkn}}{m}$$

315 where p_{ikn} is the additive probability that haplotype block (i.e., a 10kb interval) n in genotype i
316 originated from ancestral genotype k and m is the total number of haplotype blocks where
317 genotypic information is present in both inbred genotypes.

318 As reported below, we were unable to find strong statistical support for individual
319 haplotype blocks contributing to sexual concordance or antagonism. A plausible explanation is
320 that effect sizes are low relative to environmental variation, so it becomes unlikely to find the very
321 low *p*-values required when controlling for multiple testing. To explore the data, we shifted our

focus to testing for an excess of haplotype blocks with *p*-values below a given critical value, p_{crit}. In principle, the expected number of such blocks under the null hypothesis of no true genetic effects (i.e., the number of false positives, *x*) should be $E[x] = p_{crit} n_{total}$ where n_{total} is the number of blocks tested. If the tests were independent, it would be straightforward to calculate the full probability distribution of *x* under the null hypothesis and properly test whether the observed number of 'significant' blocks $n_{p \leq pcrit}$ is significantly larger than expected under the null hypothesis. There are two potential issues with this approach. First, it relies on the *p*-value from each individual test accurately reflecting the chance of a false positive under the null hypothesis (i.e., that all the assumptions of the mixed model are well met). Second, the tests are not independent due to linkage disequilibrium. Permutations offer a conceptually simple (though computationally intensive) solution to both issues. We permuted line identity (i.e., genome-wide genotype) with respect to the set of fitness values (i.e., the 8 traits) and reanalyzed the data. From each permutation, we recorded the *p*-value for each block (and for each trait) and then calculated an *x* value (i.e., number of 'significant' blocks) for a given statistical criteria ($p_{crit} \in \{0.001, 0.01, 0.05, 0.1\}$). To reduce computational time, we only tested every 7th block; the same blocks were tested in the unpermuted and permuted data sets. (There is little additional information gained from testing all blocks because of the high linkage disequilibrium among neighboring blocks.) 2077 permutations were performed.

False discovery rates are an alternative approach to the multiple testing problem inherent in genomic analyses but are not necessarily a perfect solution. The reliability of these methods depend not only on the validity of the underlying *p*-values, but can depend on the independence of tests as well as the shape of the distribution of *p*-values. The *R* package *qvalue* (Storey et al. 2021) was used to convert *p*-values into *q*-values (i.e., FDR rate) based on a previously described method

345 (Storey and Tibshirani 2003). We calculated q-values with the unpermuted data as well as each
346 permuted data set. As we did with p-values, we compare the number of blocks with a q-value
347 below a given threshold, q_{crit}, to the distribution under the null hypothesis as determined from the
348 permutations.

349

350 *Gene Ontology (GO) Enrichment Analysis*

351 We conducted a gene ontology (GO) enrichment analysis of genes found in "genomic
352 clusters" that are associated with male and female fitness and the SA and SC axes. To help control
353 for strong linkage between loci, we employed a clustering algorithm to define genomic clusters
354 that are associated with each trait. For each chromosome arm, our clustering procedure began at
355 the block furthest away from the centromere. We used a sliding window of 3.5 Mb to identify
356 windows (or clusters) that had at least one haplotype block with $p < 0.1$. When a significant
357 haplotype block was identified, we examined the immediately proceeding 3.5 Mb window. If this
358 window contained any haplotype blocks that reached statistical significance, these blocks was
359 subsumed into the previous cluster. This procedure was repeated until no haplotype blocks in the
360 proceeding 3.5 Mb window reached statistical significance. Thus, a cluster was defined as a set of
361 statistically significant haplotype blocks that are closer than 3.5 Mb to each other. After obtaining
362 genomic clusters associated with each of the eight traits (see above), we used Ensembl's *Variant*
363 *Effect Predictor* (*VEP*; McLaren et al. 2010) to find genes that fall 5Kb upstream or downstream
364 of the leading haplotype block in each cluster (i.e., the haplotype block with the lowest p-value).
365 This analysis used a liberal $p < 0.1$ cut off for statistical significance of haplotype blocks but in
366 many cases the leading block for a cluster would have a much lower p-value. Next, we used
367 *GOrilla* (Eden et al. 2009) to test for an enrichment of GO terms in genes associated with each

368 trait relative to a background set of all genes annotated in the *Drosophila* genome from our *VEP*
369 analysis. *p*-values for GO terms were converted to *q*-values and a *q*-value threshold of 0.05 was
370 used to assess statistical significance.

371

372 Selection on Putatively Deleterious Variants in the DSPR lines

373 Our pipeline to identify putatively deleterious variants (i.e., single nucleotide
374 polymorphisms [SNPs] and insertion/deletion [indel] variants) involved three steps. First, we
375 imputed the underlying genomic sequence at sites segregating for SNPs and indels in the DSPR
376 for all 357 focal genotypes assayed in this study. Following quality control steps, we attempted to
377 enrich for sites likely to be deleterious by filtering out any variants segregating in either of two
378 geographically disparate outbred populations because variants that are deleterious should be rare
379 in nature. Finally, we mapped the variants to functional categories and retained only those variants
380 falling within coding regions and whose functional consequence could reasonably be expected to
381 be deleterious. The details of these steps are provided in the Supplementary Material.

382 Following these steps, we obtained a list of three classes of rare and putatively deleterious
383 variants segregating in coding regions: loss-of-function (LoF) variants, indels, and variants
384 resulting in radical amino acid changes. We scored each inbred genotype for the number of variants
385 of each class that they carry. We termed the number of putatively deleterious variants of carried
386 by a genotype to be that genotype's "mutational burden". Note that the estimated number of
387 deleterious variants for each RIL will, on average, be slightly below the true number because a
388 small fraction of genomic blocks are excluded from each RIL. The "uncounted" variants are a
389 small source of measurement error in the estimate of burden but are not expected to create biases

390 in the effects of interest (i.e., direction and magnitude of variants or sex differences in these
391 effects). Two genotypes were clear outliers with respect to the number of SNPs causing radical
392 amino acid changes (Figure S3) and as result these two genotypes were removed from further
393 analysis.

394 To estimate the fitness effects of mutational burden, data from both sexes in both
395 environments were analyzed in a single model. We used *MCMCglmm* (with family =
396 "multinomial2") to perform a mixed model analysis that included n_{LoF}, n_{indel}, $n_{radical}$ (i.e., numbers
397 of variants of each class) as well as fixed effects of sex, environment, sex-×-environment and the
398 interaction of these with variant numbers. The random effects (and priors) were similar to the
399 models described in the quantitative genetics section. Because both environments were included
400 here, this model allowed for separate and independent 'block' and 'unit' (replicate) variances for
401 each of the four sex-×-environment combinations. The model also allowed for RIL-associated
402 variances for each sex-×-environment combination as well as all of their covariances. The expected
403 fitness, on the observed scale, of a replicate of sex i in environment j with a given mutational
404 burden can be calculated as:

405 $$W_{i,j}[n_{LoF}, n_{indel}, n_{rad}] = \frac{1}{1 + exp\left(-Z_{i,j}[n_{LoF}, n_{indel}, n_{rad}]\right)}$$

406 with

407 $$Z_{i,j}[n_{LoF}, n_{indel}, n_{rad}] = \beta_0 + \beta_{sex}X_i + \beta_{env}Y_j + \beta_{sex\times env}X_iY_j$$

408 $$+n_{LoF}\left(\beta_{LoF} + \beta_{sex\times LoF}X_i + \beta_{env\times LoF}Y_j + \beta_{sex\times env\times LoF}X_iY_j\right)$$

409 $$+n_{indel}\left(\beta_{indel} + \beta_{sex\times indel}X_i + \beta_{env\times indel}Y_j + \beta_{sex\times env\times indel}X_iY_j\right)$$

410 $$+n_{rad}(\beta_{rad} + \beta_{sex\times rad}X_i + \beta_{env\times rad}Y_j + \beta_{sex\times env\times rad}X_iY_j)$$

411 where the β's are the estimated model coefficients (on the logit link scale) and X_i is an indicator
412 variable for sex (X_i = 0 for females and 1 for males) and Y_j is an indicator variable for sex (Y_j = 0
413 for complex and 1 for simple). Selection against LoFs for sex *i* in environment *j* can be estimated
414 as

415 $$s_{i,j:LoF} = \frac{W_{i,j}[\bar{n}_{LoF} + 1, \bar{n}_{indel}, \bar{n}_{rad}]}{W_{i,j}[\bar{n}_{LoF}, \bar{n}_{indel}, \bar{n}_{rad}]} - 1$$

416 Selection for the other two variant classes can be estimated in the same fashion. Given our
417 interests, for each variant class, we focus on three measures: (i) selection averaged over both sexes
418 and environments: $s = (s_{f,comp} + s_{f,simp} + s_{m,comp} + s_{m,simp})/4$, (ii) sex differences in
419 selection: $\Delta_{sex} = s_f - s_m$ (this is averaging over environments, i.e., $s_f = (s_{f,comp} + s_{f,simp})/2$);
420 and (iii) selective differences between environments $\Delta_{env} = s_{comp} - s_{simp}$ (this is averaging over
421 sexes, i.e., $s_{comp} = (s_{f,comp} + s_{m,comp})/2$). For each of the three variant classes, we estimated
422 s, Δ_{sex} and Δ_{env} from each saved iteration in the MCMC chain to generate posterior distributions.

423 **Results**

424 Quantitative Genetic Analysis

425 We recover modest additive genetic variation (σ_A^2) for adult fitness (measured as number
426 of offspring produced) in both sexes when assayed in simple ($\sigma_{A,f}^2$ = 4.01 [95% C.I.: 3.03, 5.03];
427 $\sigma_{A,m}^2$ = 8.57 [6.11, 11.02]) and complex ($\sigma_{A,f}^2$ = 8.02 [5.79, 10.10]; $\sigma_{A,m}^2$ = 12.90 [9.08, 17.19])
428 mating environments. We confirmed by visual inspection that the posterior distribution for each
429 genetic variance component was unimodally distributed with similar mode and mean values. These

genetic variances correspond to modest estimates of heritability (simple environment: $h_f^2 = 0.15$ [0.10, 0.19], $h_m^2 = 0.102$ [0.074, 0.134]; complex environment: $h_f^2 = 0.15$ [0.09, 0.20], $h_m^2 = 0.09$ [0.04, 0.14]) and evolvability (simple environment: $e_{\mu:f} = 0.05$ [0.04, 0.06], $e_{\mu:m} = 0.06$ [0.04, 0.08]; complex environment: $e_{\mu:f} = 0.08$ [0.06, 0.10], $e_{\mu:m} = 0.08$ [0.05, 0.11]).

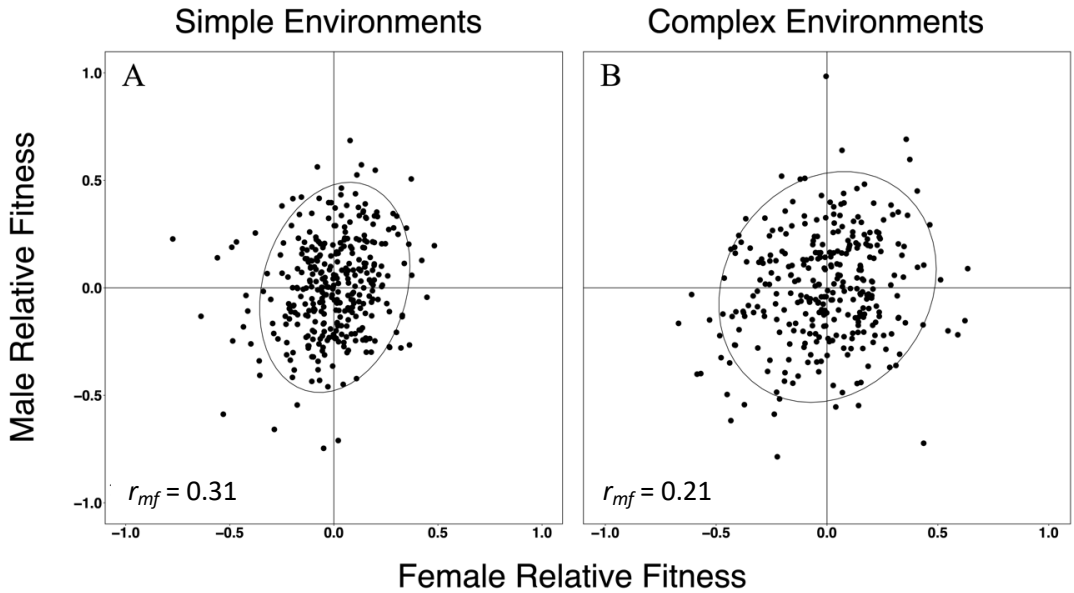

Figure 1: Male and female mean relative fitness of DSPR genotypes measured in **(A)** simple and **(B)** complex mating environments. For this figure (but not analysis), fitness measures are standardized to the within-block mean fitness for each sex and environment respectively and zero centered. Ellipses represent the 95% confidence ellipse for the distribution of sex-specific mean fitness estimates.

The intersexual genetic correlation for fitness was positive and significantly different from zero in both mating environments (Figure 1; simple: $r_{mf} = 0.31$ [0.14, 0.48]; complex: $r_{mf} = 0.21$ [0.03, 0.38]). Though positive with credibility intervals that do not overlap zero, these correlations are far from unity, suggesting a fair degree of genetic independence between the sexes. The cross-

environment correlation for fitness was also positive and substantially different from zero and unity in both sexes (Figure 2; $r_{env,f}$ = 0.28 [0.12, 0.45], $r_{env,m}$ = 0.40 [0.21, 0.57]).

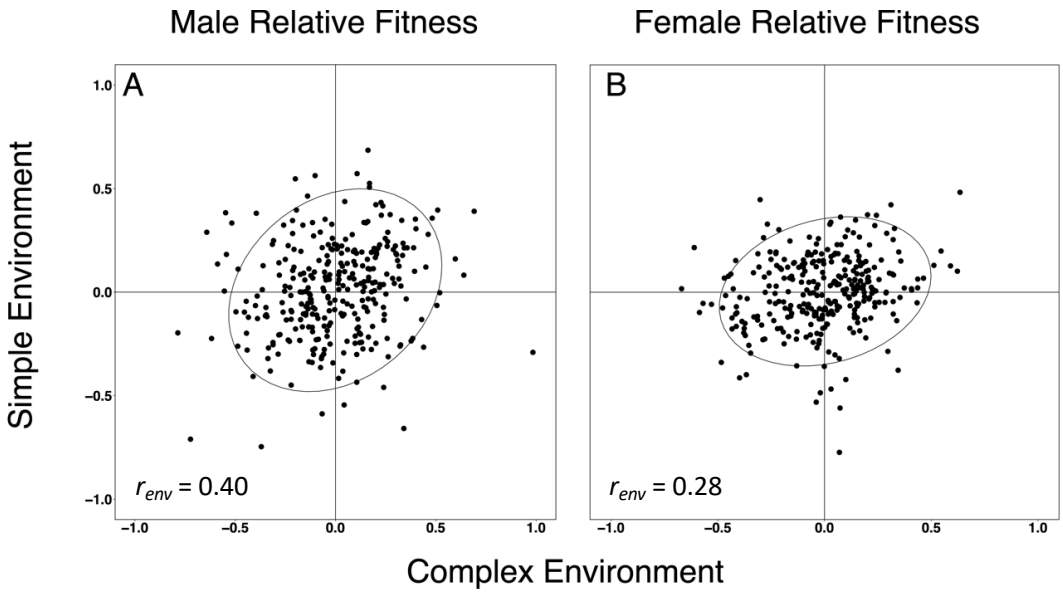

Figure 2: Mean relative fitness measured in simple and complex mating environments in **(A)** males and **(B)** Females. For this figure, fitness measures are standardized to the within-block mean fitness for each sex and environment respectively and zero centered. Ellipses represent the 95% confidence ellipse estimated from the sex-specific mean fitness estimates.

<u>Association Mapping of Sexually Antagonistic and Sexually Concordant Loci</u>

We present the results for sexual antagonism (SA) and sexual concordance (SC) here and results for male and female fitness are given in the Supplemental Materials. Manhattan plots for SA and SC in the simple and complex environments are shown in Figure 3. Few haplotype blocks have strong statistical associations with SA or SC; thus, we have little confidence in specific regions affecting SA or SC. Consequently, we shifted our focus to exploring whether the there is any evidence at all of regions associated with SA or SC.

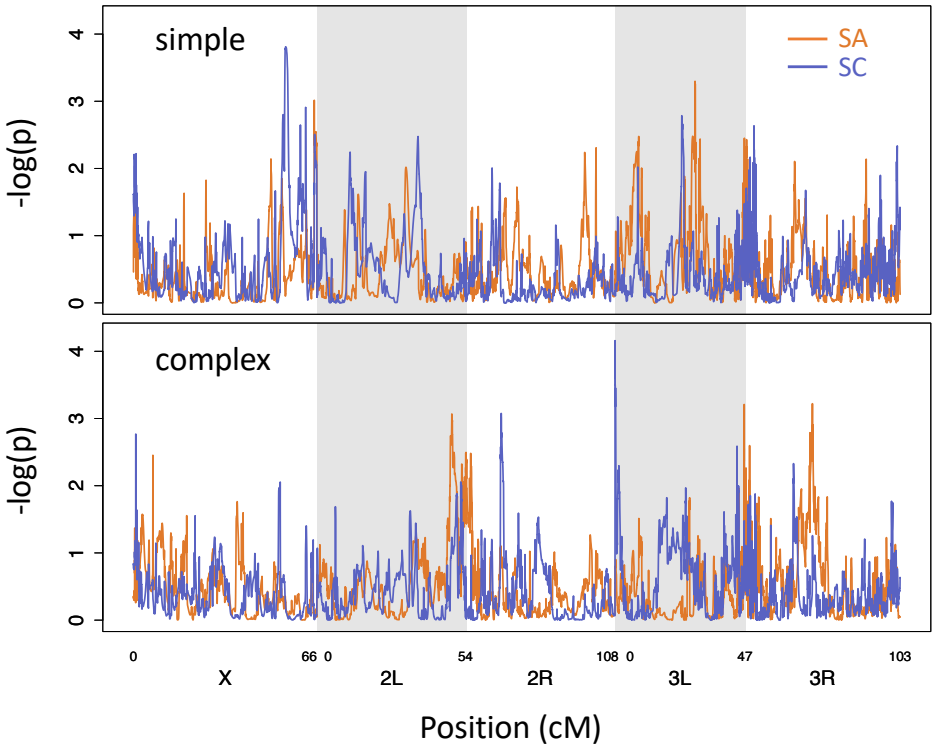

Figure 3: Manhattan plots. Strength of statistical support for associations between 10kb haplotype blocks along the genome with sexual antagonism (SA; orange) and sexual concordance (SC; purple) in DSPR lines when fitness was measured in the simple (top) or complex (bottom) mating environment.

469 As reported above, the the classical quantitative genetic analysis indicates the existence
470 genetic variation for both sexes. Because SA and SC are simple linear transformations of female
471 and male fitnesses, we can readily compute the additive genetic variances of the former from the
472 latter (simple environment: $\sigma^2_{A:SA} = 0.04$ [0.03, 0.05], $h^2_{SA} = 0.08$ [0.06, 0.11], $\sigma^2_{A:SC} = 0.07$ [0.05,
473 0.09], $h^2_{SC} = 0.15$ [0.12, 0.19]; complex environment: $\sigma^2_{A:SA} = 0.06$ [0.04, 0.08], $h^2_{SA} = 0.09$ [0.05,
474 0.13], $\sigma^2_{A:SC} = 0.09$ [0.07, 0.12], $h^2_{SC} = 0.13$ [0.08, 0.18]; see Supplementary Material for details
475 of calculations). Given the existence of this quantitative genetic variation for SA and SC, it is
476 plausible that many blocks have modest effects that do not yield extreme p-values (e.g., $p < 10^{-5}$).
477 To explore this possibility, we counted the number haplotype blocks $n_{p \leq pcrit}$ that passed a given
478 statistical requirement (e.g., $p_{crit} = 0.001$, $p_{crit} = 0.05$) and compared this number to the distribution
479 under the null hypothesis of no genetic effects as generated by permutation.

480 Consider the results using $p_{crit} = 0.05$. We first ask if there is evidence of genetic
481 associations of any sort, considering the tests for all four traits (i.e., SA and SC in both
482 environments). We find 610 tests that pass the criterion of $p \leq 0.05$ ("significant blocks"). In
483 permutations, the average number of significant blocks is about half this number (353.3) and fewer
484 than 1% of permutations result in 610 or more blocks (i.e., there are more associations than
485 expected by chance). Examining SA and SC separately, there are 338 significant associations with
486 SA considering both environments. In permuted data sets, the average number of SA blocks is
487 176.2 and fewer than 2% of permutations have 338 or more. There are 272 significant SC blocks
488 across environments though ~10% of permutations have as many or more. Examining SA and SC
489 in each of the environments separately, there are more observed blocks than the corresponding
490 average number observed for each of the four traits, but only for SA in the simple environment is

491 the observed number in the upper 5% of the tail of its corresponding permutation distribution
492 (Table 1).

493 These results are sensitive to the statistical criterion (p_{crit}) used to identify significant blocks
494 (Table 1). In most cases, the observed number of blocks is greater than expected (i.e., the means
495 of the permutation distributions) but whether the excess is substantially greater than expected
496 depends on which trait is being considered and what statistical criterion is used. Whereas the
497 results based on $p_{crit} = 0.05$ indicate an excess of associations for SA but not SC, with $p_{crit} = 0.001$
498 or $p_{crit} = 0.1$, there appears to be an excess of associations for SC but not SA. With $p_{crit} = 0.01$,
499 neither type of association is in the upper 5% of the tail of the permutation distribution.

500 The use of permutations as a basis for a null distribution is important in evaluating the data
501 from the DSPR. As an example, consider the associations tests for SA in the simple environment
502 using $p_{crit} = 0.05$. 1682 blocks are tested. Under the null hypothesis of no true associations, we
503 expect 1682*0.05 = 93.1 'significant' blocks on average; the mean of the permutations is close to
504 this value at 88.1. If the tests were independent, then the variance among permutations should be
505 1682*0.05*0.95 = 88.5; however, the observed variance—1952—is much larger. This higher
506 variance is likely due to the non-independence of tests because of linkage disequilibrium, which
507 can result in the false positive rate being much higher (or much lower) than expected under
508 independence in any given instance (i.e., individual permutation or real data). The use of q-values
509 does not solve this issue (Table S9). For example, applying the significance criterion of $q \leq 0.3$,
510 only 73% of permutations result in 0 'significant' blocks and, of the remaining 27% of
511 permutations, the average number of 'significant' blocks is 129 (range: 1-1675).

512

Table 1. The number haplotype blocks passing a specified statistical criteria. For this analysis, only every 7th haplotype block was tested (1682 blocks tested). *n* is the observed number in the real (i.e., unpermuted data). $\bar{n}_{perm}$ is the mean number of associations in permutated data sets. P_{perm} is the proportion of permutations with as many as or more than the observed number of associations (*n*); cases where the observed value is the upper 5% of the distribution (i.e., $P_{perm} < 0.05$) are shown in bold. Results using a *q*-value criterion are shown in the Supplementary Material, as are results for male and female fitness (Tables S2-4).

Criterion for 'significant' association	Metric	Any trait	SA	SC	SA in simple environment	SC in simple environment	SA in complex environment	SC in complex environment
	n	24	3	21	3	7	0	14
p_{crit} = 0.001	$\bar{n}_{perm}$	7.5	3.9	3.7	2.1	1.8	1.7	1.9
	P_{perm}	0.056	0.396	**0.025**	0.232	0.079	1	**0.025**
	n	117	54	63	34	34	20	29
p_{crit} = 0.01	$\bar{n}_{perm}$	73.2	36.2	36.9	18.3	18.7	18.0	18.3
	P_{perm}	0.117	0.184	0.131	0.148	0.152	0.333	0.201
	n	710	338	272	195	162	143	110
p_{crit} = 0.05	$\bar{n}_{perm}$	353	176	178	88.1	89.1	88.0	88.5
	P_{perm}	**0.0087**	**0.018**	0.095	**0.031**	0.084	0.127	0.274
	n	1064	479	585	261	298	218	287
p_{crit} = 0.1	$\bar{n}_{perm}$	699	348	351	173	174	174	176
	P_{perm}	**0.0077**	0.091	**0.016**	0.100	0.055	0.230	0.062

519
520 Following the procedure outlined in the Methods to generate lists of candidate regions, we
521 conducted GO enrichment analyses (Tables S5-S12). Under the statistical criterion used, no GO
522 terms were enriched among genes associated with the SA and SC in the simple environment.
523 However, we find evidence of enrichment of GO terms among genes associated with SA and with
524 SC in the complex environment. For SA, the enriched categories are biological processes that
525 affect basic cell function as well as higher levels of regulation (e.g., post-translational regulation,
526 epigenetic regulation; Table S7). For SC, a single molecular process was enriched: dipeptidyl-
527 peptidase activity (Table S8). Though the terms enriched are not readily interpretable, the existence
528 of enriched terms is another line of evidence that the association mapping results are more than
529 mere noise.

530

531 Mutational Burden

532 The assayed DSPR genotypes had on average 5.6 (range: 1-12) LOFs, 125.7 (75-172)
533 indels in coding regions, and 274.8 (39-363) radical amino acid missense variants. We estimated
534 the fitness effects of each of these three variant classes with the goal of addressing three questions:
535 What is the average fitness effect per variant (s), averaging over both sexes and both
536 environments? What is the difference in fitness effects between sexes ($\Delta_{sex} = s_f - s_m$)? What is the
537 difference in fitness effects between environments ($\Delta_{env} = s_{comp} - s_{simp}$)?

538 For the two variant types with the strongest *a priori* expectation to be deleterious—LoFs
539 and indels—the point estimates of average selection (s) are, counter to expectation, slightly
540 positive but with 95% credible intervals broadly overlapping zero (Figure 4). However, for both
541 LoFs and indels, there is evidence that deleterious selection is stronger in females than males

542 (LoFs: Δ_{sex} = -0.021 [-0.038, -0.003]; indels: Δ_{sex} = -0.0019 [-0.0038, -0.0001]). In females, adult
543 fitness is estimated to be reduced by ~ 0.7% per LoF and ~0.06% per indel; in males, the estimates
544 are of the opposite sign and with substantially higher uncertainty. For these two variant types, there
545 is no support for environmental differences (i.e., credibility intervals for Δ_{env} overlap zero). For
546 radical amino acid missense variants, average selection is estimated to be positive rather than
547 negative (s = 0.0008 [0.0001, 0.0014]). For this variant type, there is no support for sex differences
548 (credibility intervals for Δ_{sex} overlap zero) but selection appears to differ between environments
549 (Δ_{env} = -0.0015 [-0.0028, -0.0001]), with selection favouring such variants more strongly in the
550 simple than complex environment.

551

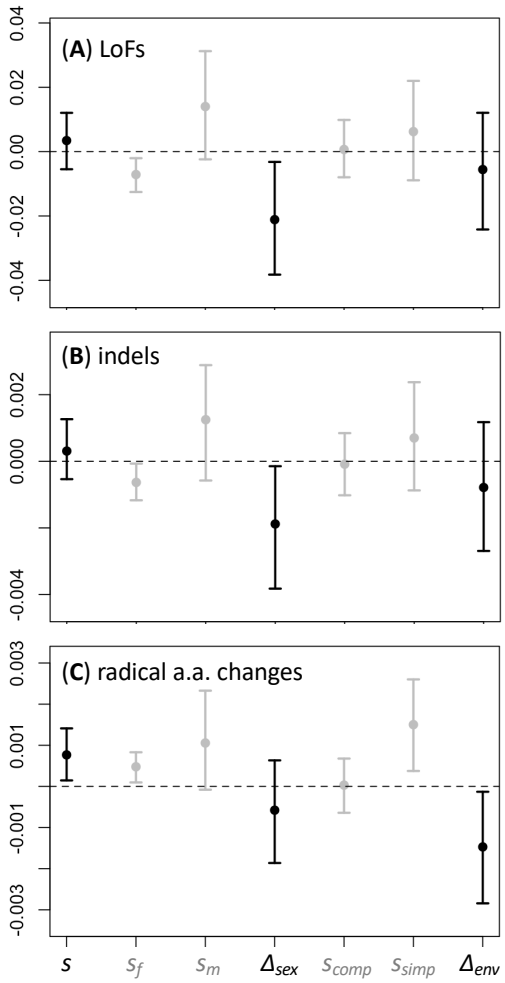

Figure 4: Fitness effects of putatively deleterious variants: (A) LoFs, (B) indels, (C) radical amino acid missense variants. s is the estimated selective effect per variant, averaging across sexes and environments (negative values indicate deleterious fitness effects); s_f (s_m) is the selection per variant in females (males), averaging across environments; $\Delta_{sex} = s_f - s_m$ is the sex difference in selection; s_{comp} (s_{simp}) is the selection per variant in the complex (simple) environment, averaging across sexes; $\Delta_{env} = s_{comp} - s_{simp}$ is the sex difference in selection.

563 **Discussion**

564 The sexes represent different selective environments: surveys of phenotypic selection have
565 found that shared traits experience persistent sex differences in both the strength (Singh and
566 Punzalan 2018) and direction (Cox and Calsbeek 2009; De Lisle et al. 2018) of selection. Thus,
567 there is considerable scope for segregating genetic to affect fitness of the two sexes differently. On
568 the other hand, the two sexes share many similar requirements so selection on many variants will
569 be similar. In this study, we used three approaches to explore different aspects of the sex-specific
570 genetic architecture of fitness in a 'synthetic' laboratory population of *Drosophila melanogaster*.

571 The history of this synthetic population should be considered in interpreting our results.
572 The RILs of the DSPR were established after first mixing eight inbred lines from different
573 geographic regions and allowing for 50 generations of recombination (King et al. 2012a,b). During
574 that period, some level adaptation to laboratory conditions is likely to have occurred. The selection
575 regime that occurred during period was not identical to either of the environments used in this
576 assay, though presumably it was more similar to those environments than to selection in many
577 environments in nature where this species is found. Thus, it is difficult to say how well adapted
578 the population was to the assay conditions. This issue is not unique to our study; most studies lack
579 a quantitatively meaningful measure of how far a population is from maximal adaptation, though
580 it is common to claim a lab population of invertebrates is "well-adapted" if it has been maintained
581 in the lab for a few tens of generations.

582 This is an important issue because the genetic variation will be shaped by the history of
583 selection. Selection is expected to drive variants that are deleterious in a given environment to low
584 frequency. Thus, a population with a history of selection in a given environment may carry many
585 fewer such variants at intermediate frequency than a population without that history of selection.

586 This issue is not only relevant to the amount of genetic variance for fitness but also the intersexual
587 genetic correlation. Populations experiencing a new environment may harbour numerous variants
588 that are deleterious for both sexes. Over time, selection is expected to remove this sexually
589 concordant variation; in contrast, variants with sexually antagonistic fitness effects are expected
590 to persist longer and make up an increasingly larger proportion of the genetic variation, resulting
591 in a less positive, or more negative, r_{mf} as the population adapts (Long et al. 2012; Connallon and
592 Clark 2013). We do not know where our population is in this process.

593 An additional issue to consider is that, following the 50 generations of recombination (and
594 lab adaptation), the RILs of the DSPR were created by 25 generations of fullsib inbreeding. During
595 this process, there could have been some purging of variants with recessive deleterious effects.
596 Variants with deleterious effects in both sexes would tend to be purged, reducing the amount of
597 sexually concordant variation. Among sexually antagonistic variants, those with large recessive
598 deleterious effects in females may be preferentially purged because the establishment of RILs
599 depends more strongly on female than male fitness (Grieshop et al. 2017). However, without a
600 series of additional assumptions, it is difficult to predict the net effect of the inbreeding process on
601 fitness variation from sexually concordant versus antagonistic variants and, thereby, the effect on
602 r_{mf}.

603

604 Quantitative Genetics Analysis

605 Our quantitative genetics analysis recovers significant levels of additive genetic variance
606 for fitness in both sexes in the two mating environments where we assayed adult fitness. The
607 heritabilities and evolvabilities of fitness were modest in both sexes in the two mating

608 environments. These results are similar in magnitude as reported in other laboratory populations
609 of *Drosophila* (e.g., Fowler et al. 1997; Gardner et al. 2005; Pischedda and Chippindale 2006;
610 Long et al. 2009; Ruzicka et al. 2019) and other insect species (e.g., Archer et al. 2013; Berger et
611 al. 2014) but our estimates of the heritability are much higher than those reported from studies of
612 wild populations (e.g., Gustafsson 1986; Kruuk et al. 2000; Merilä and Sheldon 2000; McCleery
613 et al. 2004; Teplitsky et al. 2009; Mcfarlane et al. 2014). Low heritability estimates in wild
614 populations are likely the result of high levels environmental variance in fitness (Price and Schluter
615 1991; Houle 1992) rather than a lack of genetic variation. As noted above, the population used
616 here is not expected to be exquisitely adapted to these assay conditions. Thus, there may be more
617 fitness-affecting variants segregating at intermediate frequency relative to an adapted population
618 where directional selection is expected to have reduced σ_A^2.

619 The intersexual genetic correlation for fitness r_{mf} gives us a genome-wide summary of the
620 extent to which the shared genome constrains the sexes from reaching their sex-specific trait
621 optima (Lande 1980) at least in the short term. Our estimates of r_{mf} were positive in both mating
622 environments (Figure 1), implying that, on average, genetic variants have sexually concordant
623 fitness effects and, as a result, evolution would not be strongly constrained from improving the
624 fitness of both sexes simultaneously in either of these environments. Though estimates for r_{mf} in
625 both environments are significantly greater than 0, they are also far from 1. This intermediate value
626 for r_{mf} suggests that while sexually concordant variants may be predominant, other types of
627 variants (e.g., sex-limited and/or sexually antagonistic variants) are likely also segregating. This
628 view is consistent with our genomic association mapping analysis: the Manhattan plots (Figure 3)
629 suggest that sexually antagonistic genomic regions are present and are only modestly less frequent
630 than sexually concordant ones (though this inference should be regarded with caution because of

631 limited power). It is plausible that our estimates of positive r_{mf} reflect an overall concordance of
632 selection as a consequence of this population being relatively maladapted to the laboratory
633 conditions in which we measured fitness, which was not identical to how the lab population was
634 originally created by King et al. (2012a,b).

635 Our significantly positive estimates of r_{mf} stands in contrast to a classic study using the
636 same species by Chippindale et al. (2001) who reported a significantly negative r_{mf}. That study
637 measured fitness in a lab context to which their fly population was "well adapted." However, not
638 all studies neatly fit within the narrative in which the degree of adaptedness explains variation in
639 r_{mf}. Sharp and Agrawal (2018) estimated a positive though non-significant values for r_{mf} using
640 another *D. melanogaster* population that was "well-adapted" to the lab in which assay conditions
641 were similar to but did not exactly replicate maintenance conditions. A particularly interesting case
642 is that of Collet et al. (2016) who measured r_{mf} in two descendant populations of the fly population
643 studied by Chippindale et al. (2001). Despite being maintained and assayed similarly, the two
644 descendant populations had significantly different values of r_{mf}, with r_{mf} significantly negative
645 in one and non-significantly positive in the other. That study demonstrates that r_{mf} can evolve
646 substantially over a few hundred generations in ways that are not captured in any obvious way by
647 "adaptedness".

648 In addition to evolutionary changes in r_{mf} (i.e., changes in the frequencies of sequence
649 variants), r_{mf} can be directly affected by the environment. Studies have found that r_{mf} may be
650 highly sensitive to the ecological context in which genetic variation is expressed (e.g., Berger et
651 al. 2014; Punzalan et al. 2014; Duffy et al. 2019). However, our estimates of r_{mf} in the two mating
652 environments are reasonably similar (and confidence intervals are broadly overlapping). Given

653 that sexual interactions are quite different in these two types of environments (Yun et al. 2017), it
654 was not obvious from the outset that the two r_{mf} estimates would be so similar.

655 The similarity of r_{mf} estimates would be unsurprising if the genetic architecture of fitness
656 for each sex was the same across the two environments (i.e., if r_{env} was close to 1 for both sexes).
657 Though r_{env} was significantly positive is was also substantially less than 1 in both sexes, indicating
658 that the genetic architecture for fitness differs somewhat across environments for both sexes. This
659 is not surprising given that these two environments likely select for different phenotypic traits in
660 both sexes. We have strong reasons to believe this for males. In a different experiment, populations
661 that evolved for a few tens of generations in environments similar to those examined here had
662 diverged significantly for male fitness in a genotype-x-environment dependent manner (Yun et al.
663 2019), presumably because these alternative mating environments modulate the opportunity for
664 male harassment to result in successful mating (Yun et al. 2017). In contrast to males, females
665 showed much weaker signs of divergence following adaptation to these alternative environments.
666 As such, we might have expected a higher r_{env} for females than males but instead the point estimate
667 for males is higher (0.40 vs. 0.28) .

668 It is notable that the intersexual genetic correlations within an environment (r_{mf}) and the
669 cross-environment genetic correlations for each sex (r_{env}) are all significantly positive and of
670 similar magnitude. As mentioned above, a simple explanation for all correlations being positive is
671 that this population contains an excess of variants that are unconditionally deleterious in the lab
672 environment, perhaps because of insufficient time to adapt to general lab conditions. Although
673 particular suites of phenotypes may maximize fitness for each sex and mating environments, it is
674 plausible that many of the same morphological, physiological and behavioural traits necessary to
675 acquire fitness are similar between the sexes in both environments. Nonetheless, it is surprising

676 that r_{mf} and r_{env} estimates are of similar magnitude; roughly speaking, this suggests that the
677 differences between sexes are as large as the differences between environments with respect to
678 segregating variation. For example, knowing that a genotype confers high fitness to males in one
679 environment is approximately as good a predictor of conferring high fitness to females in that same
680 environment as it is of conferring high fitness to males in the other environment.

681

682 Association Mapping Analysis

683 Our association mapping analysis revealed few, if any, genomic regions with very strong
684 statistical associations with the traits examined. Consequently, we took a more exploratory
685 approach to evaluating whether there was any evidence of genetic associations beyond what would
686 be expected by chance, using permutations to create appropriate null distributions. The results of
687 these analyses were somewhat inconclusive. On the one hand, in almost all cases the observed
688 number of associations was larger than expected. However, the permutation-based null
689 distributions have high variance and, in only some cases, was the observed number of associations
690 unusually high relative to the corresponding null distribution. With some criteria, there was a
691 substantial excess of associations with SA but not SC; with other criteria, there was an excess of
692 associations with SC but not SA.

693 In general, the power to detect a significant genomic association is inversely related to the
694 minor allele frequency (MAF) at loci underlying trait variation. An advantage of this synthetic
695 population is that the variants segregating in this panel should have high MAFs (i.e., a neutral
696 expectation of 12.5%) compared to natural populations. However, the observed MAFs are often
697 lower than 12.5%, presumably due to 50 generations of drift and selection. Of course, statistical

power is also directly related to effect size. Because of selection, the standing variation will be depauperate of variants with large directional effects (e.g., unconditionally deleterious) segregating at high frequency. It is likely there are numerous variants that make small contributions to fitness variation, but we would not expect to detect them via this association mapping approach. Nonetheless, it seemed plausible we could detect large-effect sexually antagonistic variants if such alleles were segregating. Theory indicates that large-effect sexually antagonistic variants are more readily maintained by balancing selection than small effect ones (Kidwell et al. 1977; Connallon and Clark 2012). Our failure to observe any strong associations implies that the loci underlying fitness are unlikely to be of large effect.

Our tentative interpretation of the association study is that effect sizes are small to moderate for both SA and SC as we tend to see more modest associations than expected yet fail to find very strong associations. In addition, these data suggest that SA variation is not particularly rare relative to (detectable) SC variation as we see only modestly stronger evidence for associations with SC than SA (Table 2).

Using a permissive threshold to generate larger lists of candidate regions, we found some evidence of enriched GO terms for SA and SC in the complex but not the simple environment. In their genomic analysis of sexual antagonism, Ruzicka et al. (2019) found little evidence of enriched GO categories and those they did where not obviously linked to sexually dimorphic traits. One interpretation of these results is that sexual antagonism is diffuse affecting many types of processes and/or is especially prone to persisting in genes affecting a variety of basic functions that are under selective constraint. However, another interpretation is that studies such as ours lack power to meaningfully identify genomic regions experiencing sexually antagonistic selection. On the other hand, even though the terms identified are not readily biologically interpretable, the

721 existence of significantly enriched GO terms for both SA and SC in the complex environment
722 provide another form of evidence that the association analysis is not just noise.

723

724 Mutational Burden

725 Examining entire classes of variants, rather than individual variants, provides greater
726 power to detect smaller effects. We estimated that each additional LoF reduces female fitness by,
727 on average, ~0.7%, and each additional indel by ~0.06%. In contrast, we find no significant
728 associations between any of the three classes of putatively deleterious variants and relative fitness
729 in males in either mating environment. The results imply that selection acting on mutational burden
730 is stronger in females than in males.

731 These results are surprising considering many studies have found evidence for stronger
732 selection acting on variants expressed in males. Several studies have estimated the sex-specific
733 selection acting on marker mutations in *Drosophila* (Whitlock and Bourguet 2000; Pischedda and
734 Chippindale 2005; Stewart et al. 2005; Sharp and Agrawal 2008; Hollis et al. 2009). These studies
735 have generally found evidence in support of the hypothesis that the strength of selection tends to
736 be stronger in males. However, it is unclear whether results from visible marker mutations apply
737 more broadly. Several studies examining the fitness consequences of novel mutations via mutation
738 accumulation (MA) experiments in *Drosophila* and other species (e.g., Enders and Nunney 2010;
739 Mallet et al. 2011; Mcguigan et al. 2011; Sharp and Agrawal 2013; Grieshop et al. 2016) have also
740 generally supported the hypothesis of stronger selection through males. While an advantage of
741 using MA lines is that they capture the fitness effects of more 'typical' mutations, one disadvantage
742 is that the causative mutations are typically completely unknown. Fitness results from MA studies

743 may be dominated by a small subset of the mutations that have large effects. This is problematic
744 if such mutations are unrepresentative in other ways. For example, perhaps mutations with male-
745 limited effects are disproportionately likely to be of large effect.

746 It remains unclear why our current study found evidence of stronger selection in females
747 than males whereas past studies using phenotypic marker mutations or MA lines found the
748 opposite. One possibility is that the genetic variation in the DSPR has been filtered by selection in
749 the wild (i.e., which variants were present in the eight founder lines for the DSPR) in contrast to
750 MA studies that capture the fitness effects of all novel mutations (excluding only extremely large
751 effect variants e.g., lethal or sterilizing variants). Variants causing strong effects in the wild (in
752 either sex) are expected to be underrepresented relative to mutational input. One possibility for the
753 discrepancy between this study and others is that fitness effects on males are reasonably consistent
754 between field and lab conditions whereas effects on females are more variable between
755 environments. Thus, variants with large fitness effects in males in the lab would be
756 underrepresented whereas those causing large female effects in the lab would not be so
757 underrepresented due to past selection in the wild. Perhaps selection is more consistent on males
758 between field and lab because their fitness is so heavily determined by their interactions with other
759 flies via mate competition whereas female fitness may be more directly affected by the
760 environment itself (e.g., temperature, humidity, food availability).

761 A second possibility for the discrepancy between this and past study is from differences in
762 how fitness was assayed. First, viability is a potentially important source of fitness variation in our
763 study but minimized in many other studies. We measured fitness of adults via offspring produced
764 after a six day 'mating interaction' period whereas other fly studies typically use much shorter
765 mating periods. Our protracted period of mating interaction allows for more fitness variation via

766 adult mortality, especially so for females; males that die during this period can still contribute
767 offspring via matings prior to death but females dying prior to the day we allowed for egg laying
768 would prevent any contribution to offspring production in our study (i.e., fitness effect of mortality
769 during the mating period is stronger on females than males).

770 While the type of variants we studied are often presumed to be deleterious, very few studies
771 have attempted to validate this presumption by examining associations with fitness. Two recent
772 studies have employed similar methods to investigate the aggregate fitness consequences of
773 putatively deleterious variants. Yang et al (2017) employed a comparatively conservative criteria
774 to filter for putatively deleterious variants in maize (*Zea mays*) by filtering SNPs based on both
775 their genomic evolutionary rate profiling (GERP) scores—a measure of evolutionary constraint on
776 a SNP (Cooper et al. 2005)—and the predicted functional consequence of variants. Brown and
777 Kelly (2019) took a more liberal approach to identify deleterious variants in yellow monkeyflower
778 (*Mimulus guttatus*) by filtering out any variant with a minor allele cut-off greater than 5%. Both
779 studies we able to identify significant associations between mutation burden and measures of
780 fitness. Our own study did not provide evidence of negative selection when we averaged over
781 both sexes. For LoFs and indels, we detected negative selection in females but not males. For
782 radical amino acid missense variants, sex-average selection was positive. The latter class of variant
783 is the one for which we have the weakest *a priori* expectation for deleterious effects. Our results
784 suggest that some of these variants may be beneficial at least in the lab environment.

785 Conclusion

786 Any study of the genetic variation in fitness should be considered with respect to the history of the
787 population in relation to the environment in which fitness was assayed. Our study likely represents
788 a modestly, but not exquisitely, adapted population. Relative to an adapted population at

789 equilibrium in a constant environment, we would expect the genetic lines used here to harbour
790 more sexually concordant variation. Despite this expected "bias" towards concordance, both the
791 quantitative genetic and association mapping results suggest sex differences in the genetic
792 architecture of fitness. The observed intersexual genetic correlations are greater than zero, yet they
793 are also substantially less than 1. Within the limits of our statistical power, there is evidence for
794 genomic regions that are sexually concordant but also regions that are sexually antagonistic. While
795 the history of the population is expected to affect the relative amounts of concordant versus
796 antagonistic variation, there is no *a priori* reason it should differentially affect the sexes with
797 respect to the fitness effects of putatively deleterious variants. Nonetheless, we find evidence of
798 stronger selection in females than males, contrary to studies of selection against phenotypic
799 markers and new mutations. While caution should be used in interpreting these results, they add
800 to the existing empirical evidence of sex differences in selective effects.

801

802

803 **Acknowledgements**: We are grateful to many undergraduate assistants who helped with data
804 collection, especially Nina Adler, Malak Bayoumi, Elenore Breslow, and Matthew Lindsay. We
805 thank Karl Grieshop, Julia Kreiner, George Sandler, and Santiago Sánchez-Ramírez for
806 statistics/bioinformatics help and advice. DSPR lines were generously provided by Stuart
807 MacDonald. Funding for this research was provided in the form of an Ontario Graduate
808 Scholarship (OGS) to AS and an NSERC Discovery Grant to AFA.

809

810

811

812

813

814 **Literature Cited**

Agrawal, A. F. 2001. Sexual selection and the maintenance of sexual reproduction. Nature 411:692–695.

Archer, C. R., S. K. Sakaluk, C. Selman, N. J. Royle, and J. Hunt. 2013. Oxidative stress and the evolution of sex differences in life span and ageing in the decorated cricket, *Gryllodes sigillatus*. Evolution. 67:620–634.

Arnqvist, G., and L. Rowe. 2005. Sexual Conflict. Princeton University Press, Princeton NJ.

Barton, N. H., and P. D. Keightley. 2002. Understanding quantitative genetic variation. Nat. Rev. Genet. 3:11–21.

Bateman, A. J. 1948. Intra-sexual Selection in *Drosophila*. Heredity. 2:349–368.

Bates, D., M. Mächler, B. M. Bolker, and S. Walker. 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67:1–48.

Berger, D., K. Grieshop, M. I. Lind, J. Goenaga, A. A. Maklakov, and G. Arnqvist. 2014. Intralocus sexual conflict and environmental stress. Evolution. 68:2184–2196.

Bonduriansky, R., and S. Chenoweth. 2009. Intralocus sexual conflict. Trends Ecol. Evol. 24:280–288.

Brommer, J. E., M. Kirkpatrick, A. Qvamström, and L. Gustafsson. 2007. The intersexual genetic correlation for lifetime fitness in the wild and its implications for sexual selection. PLoS One 2: e744.

Brown, K. E., and J. K. Kelly. 2019. Severe inbreeding depression is predicted by the "rare allele load" in *Mimulus guttatus*. Evolution. 74:587–596.

Chippindale, A. K., J. R. Gibson, and W. R. Rice. 2001. Negative genetic correlation for adult fitness between sexes reveals ontogenetic conflict in *Drosophila*. Proc. Natl. Acad. Sci. 98:1671–1675.

Collet, J. M., S. Fuentes, J. Hesketh, M. S. Hill, P. Innocenti, E. H. Morrow, K. Fowler, and M. Reuter. 2016. Rapid evolution of the intersexual genetic correlation for fitness in *Drosophila melanogaster*. Evolution. 70:781–795.

Connallon, T., and A. G. Clark. 2012. A general population genetic framework for antagonistic selection that accounts for demography and recurrent mutation. Genetics 190:1477–1489.

Connallon, T., and A. G. Clark. 2014. Evolutionary inevitability of sexual antagonism. Proc. Biol. Sci. 281:20132123.

Connallon, T., and A. G. Clark. 2013. Evolutionary inevitability of sexual antagonism. Proc. R.

Soc. B Biol. Sci. 281.

Cooper, G. M., E. A. Stone, G. Asimenos, E. D. Green, S. Batzoglou, and A. Sidow. 2005. Distribution and intensity of constraint in mammalian genomic sequence. Genome Res. 15:901–913.

Cox, R. M., and R. Calsbeek. 2009. Sexually Antagonistic Selection, Sexual Dimorphism, and the Resolution of Intralocus Sexual Conflict. Am. Nat. 173:176–187.

Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, and R. Durbin. 2011. The variant call format and VCFtools. Bioinformatics 27:2156–2158.

De Lisle, S., D. Goedert, A. Reedy, and E. Svensson. 2018. Climatic factors and species range position predict sexually antagonistic selection across taxa. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 373:20170415.

de Villemereuil P., H. Schielzeth, S. Nakagawa, and M. Morrissey. 2016. General methods for evolutionary quantitative genetic inference from generalised mixed models. Genetics 204: 1281-1294.

de Villemeruil, P. 2018. Quantitative genetic methods depending on the nature of the phenotypic trait. Annals of the New Yor Academy of Sciences. 1422: 29-47.

Duffy, E., C. R. Archer, M. D. Sharma, M. Prus, R. A. Joag, J. Radwan, N. Wedell, and D. J. Hosken. 2019. Wolbachia infection can bias estimates of intralocus sexual conflict. Ecol. Evol. 9:328–338.

Enders, L. S., and L. Nunney. 2010. Sex- specific effects of inbreeding in wild- caught *Drosophila melanogaster* under benign and stressful conditions. J. Evol. Biol. 23:2309–2323.

Fedorka, K. M., and T. A. Mousseau. 2004. Female mating bias results in conflicting sex-specific offspring fitness. Nature 429:65–67.

Foerster, K., T. Coulson, B. C. Sheldon, J. M. Pemberton, T. H. Clutton-Brock, and L. E. B. Kruuk. 2007. Sexually antagonistic genetic variation for fitness in red deer. Nature 447:1107–1110.

Fowler, K., C. Semple, N. H. Barton, and L. Partridge. 1997. Genetic variation for total fitness in *Drosophila melanogaster*. Proc. R. Soc. B 264:191–199.

Fox, J., and S. Weisberg. 2019. An R Companion to Applied Regression. Third. Sage, Thousand Oaks, CA.

Frankham, R. 2009. Genetic Architecture of Reproductive Fitness and its Consequences. Pp. 15–

877 40 *in* J. van der Werf, H.-U. Graser, and R. Frankham, eds. Adaptation and fitness in animal
878 populations: evolutionary and breeding perspectives on genetic resource management.
879 Springer US, New York.
880 Fry, J. D. 2010. The genomic location of sexually antagonistic variation: Some cautionary
881 comments. Evolution. 64:1510–1516.
882 Gardner, M. P., K. Fowler, N. H. Barton, and L. Partridge. 2005. Genetic variation for total fitness
883 in *Drosophila melanogaster*: Complex yet replicable patterns. Genetics 169:1553–1571.
884 Grieshop, K., J. Stångberg, I. Martinossi-Allibert, G. Arnqvist, and D. Berger. 2016. Strong sexual
885 selection in males against a mutation load that reduces offspring production in seed beetles.
886 J. Evol. Biol. 29:1201–1210.
887 Grieshop, K., D. Berger, and G. Arnqvist. 2017. Male-benefit sexually antagonistic genotypes
888 show elevated vulnerability to inbreeding. BMC Evol Biol 17: 134.
889 Gustafsson, L. 1986. Lifetime Reproductive Success and Heritability: Empirical Support for
890 Fisher's Fundamental Theorm. Am. Nat. 128:761–764.
891 Hadfield, J. D. 2010. MCMC Methods for Multi-Response Generalized Linear Mixed Models:
892 The MCMCglmm R Package. J. Stat. Softw. 33:1–22.
893 Hollis, B., J. L. Fierst, and D. Houle. 2009. Sexual selection accelerates the elimination of a
894 deleterious mutant in *Drosophila melanogaster*. Evolution. 63:324–333.
895 Houle, D. 1992. Comparing evolvability and variability. Genetics 130:195–204.
896 Kent, W. J., C. W. Sugnet, T. S. Furey, K. M. Roskin, T. H. Pringle, A. M. Zahler, and a. D.
897 Haussler. 2002. The Human Genome Browser at UCSC. Genome Res. 12:996–1006.
898 Kidwell, J. F., M. T. Clegg, F. M. Stewart, and T. Prout. 1977. Regions of stable equilibria for
899 models of differential selection in the two sexes under random mating. Genetics 85:171–183.
900 King, E. G., S. J. Macdonald, and A. D. Long. 2012a. Properties and power of the *Drosophila*
901 synthetic population resource for the routine dissection of complex traits. Genetics 191:935–
902 949.
903 King, E. G., C. M. Merkes, C. L. McNeil, S. R. Hoofer, S. Sen, K. W. Broman, A. D. Long, and
904 S. J. Macdonald. 2012b. Genetic dissection of a model complex trait using the *Drosophila*
905 Synthetic Population Resource. Genome Res. 22:1558–1566.
906 Kruuk, L. E. B., T. H. Clutton-Brock, J. Slate, J. M. Pemberton, S. Brotherstone, and F. E.
907 Guinness. 2000. Heritability of fitness in a wild mammal population. Proc. Natl. Acad. Sci.

908 U. S. A. 97:698–703.
909 Lack, J. B., C. M. Cardeno, M. W. Crepeau, W. Taylor, R. B. Corbett-Detig, K. A. Stevens, C. H.
910 Langley, and J. E. Pool. 2015. The *Drosophila* genome nexus: A population genomic resource
911 of 623 *Drosophila melanogaster* genomes, including 197 from a single ancestral range
912 population. Genetics 199:1229–1241.
913 Lack, J. B., J. D. Lange, A. D. Tang, R. B. Corbett-Detig, and J. E. Pool. 2016. A thousand fly
914 genomes: An expanded *Drosophila* genome nexus. Mol. Biol. Evol. 33:3308–3313.
915 Lande, R. 1980. Sexual Dimorphism, Sexual Selection, and Adaptation in Polygenic Characters.
916 Evolution. 34:292–305.
917 Long, T. A. F., A. F. Agrawal, and L. Rowe. 2012. The effect of sexual selection on offspring
918 fitness depends on the nature of genetic variation. Curr. Biol. 22:204–208.
919 Long, T. A. F., P. M. Miller, A. D. Stewart, and W. R. Rice. 2009. Estimating the heritability of
920 female lifetime fecundity in a locally adapted *Drosophila melanogaster* population. J. Evol.
921 Biol. 22:637–643.
922 Lorch, P. D., S. Proulx, L. Rowe, and T. Day. 2003. Condition-dependent sexual selection can
923 accelerate adaptation. Evol. Ecol. Res. 5:867–881.
924 Mackay, T. F. C. 2001. The Genetic Architecture of Quantitative Traits. Annu. Rev. Genet.
925 35:303–339.
926 MacPherson, A., L. Yun, T. S. Barrera, A. F. Agrawal, and H. D. Rundle. 2018. The effects of
927 male harm vary with female quality and environmental complexity in *Drosophila*
928 *melanogaster*. Biol. Lett. 14:6–9.
929 Mallet, M. A., J. M. Bouchard, C. M. Kimber, and A. K. Chippindale. 2011. Experimental
930 mutation-accumulation on the X chromosome of *Drosophila melanogaster* reveals stronger
931 selection on males than females. BMC Evol. Biol. 11:156.
932 Manning, J. T. 1984. Males and the advantage of sex. J. Theor. Biol. 108:215–220.
933 McCleery, R. H., R. A. Pettifor, P. Armbruster, K. Meyer, B. C. Sheldon, and C. M. Perrins. 2004.
934 Components of variance underlying fitness in a natural population of the great tit *Parus*
935 *major*. Am. Nat. 164:E62–E72.
936 Mcfarlane, S. E., J. C. Gorrell, D. W. Coltman, M. M. Humphries, S. Boutin, and A. G. Mcadam.
937 2014. Very low levels of direct additive genetic variance in fitness and fitness components in
938 a red squirrel population. Ecol. Evol. 4:1729–1738.

939 Mcguigan, K., D. Petfield, and M. W. Blows. 2011. Reducing mutation load through sexual
940 selection on males. Evolution. 65:2816–2829.
941 McLaren, W., B. Pritchard, D. Rios, Y. Chen, P. Flicek, and F. Cunningham. 2010. Deriving the
942 consequences of genomic variants with the Ensembl API and SNP Effect Predictor.
943 Bioinformatics 26:2069–2070.
944 Merilä, J., and B. C. Sheldon. 2000. Lifetime reproductive success and heritability in nature. Am.
945 Nat. 155:307–310.
946 Page, A. J., B. Taylor, A. J. Delaney, J. Soares, T. Seemann, J. A. Keane, and S. R. Harris. 2016.
947 SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. Microb.
948 genomics 2:e000056.
949 Parker, G. A. 1979. Sexual selection and sexual conflict. Pp. 123–166 *in* M. S. Blum and B. N. A.,
950 eds. Sexual Selection and Reproductive Competition in Insects. Academic Press, New York.
951 Pischedda, A., and A. Chippindale. 2005. Sex, mutation and fitness: Asymmetric costs and routes
952 to recovery through compensatory evolution. J. Evol. Biol. 18:1115–1122.
953 Pischedda, A., and A. K. Chippindale. 2006. Intralocus sexual conflict diminishes the benefits of
954 sexual selection. PLoS Biol. 4:2099–2103.
955 Poissant, J., A. J. Wilson, and D. W. Coltman. 2010. Sex-specific genetic variance and the
956 evolution of sexual dimorphism: A systematic review of cross-sex genetic correlations.
957 Evolution. 64:97–107.
958 Price, T., and D. Schluter. 1991. On the Low Heritability of Life-History Traits. Evolution.
959 45:853–861.
960 Punzalan, D., M. Delcourt, and H. D. Rundle. 2014. Comparing the intersex genetic correlation
961 for fitness across novel environments in the fruit fly, *Drosophila serrata*. Heredity. 112:143–
962 148. Nature Publishing Group.
963 R Development Core Team. 2019. R Development Core Team. R Foundation for Statistical
964 Computing, Vienna.
965 Ruzicka, F., M. S. Hill, T. M. Pennell, I. Flis, F. C. Ingleby, R. Mott, K. Fowler, E. H. Morrow,
966 and M. Reuter. 2019. Genome-wide sexually antagonistic variants reveal long-standing
967 constraints on sexual dimorphism in fruit flies. PLoS Biol. 17: e3000244.
968 Sharp, N. P., and A. F. Agrawal. 2018. An experimental test of the mutationselection balance
969 model for the maintenance of genetic variance in fitness components. Proc. R. Soc. B Biol.

Sci. 285.

Sharp, N. P., and A. F. Agrawal. 2013. Male-biased fitness effects of spontaneous mutations in *Drosophila melanogaster*. Evolution 67:1189–1195.

Sharp, N. P., and A. F. Agrawal. 2008. Mating density and the strength of sexual selection against deleterious alleles in *Drosophila melanogaster*. Evolution. 62:857–867.

Siller, S. 2001. Sexual selection and the maintenance of sex. Nature 411:689–692.

Singh, A., A. F. Agrawal, and H. D. Rundle. 2017. Environmental complexity and the purging of deleterious alleles. Evolution. 71:2714–2720.

Singh, A., and D. Punzalan. 2018. The strength of sex-specific selection in the wild. Evolution. 72:2818–2824.

Smith, N. G. C. 2003. Are Radical and Conservative Substitution Rates Useful Statistics in Molecular Evolution? J. Mol. Evol. 57:467–478.

Sohail, M., O. A. Vakhrusheva, J. H. Sul, S. L. Pulit, L. C. Francioli, L. H. Van Den Berg, J. H. Veldink, P. I. W. De Bakker, G. A. Bazykin, A. S. Kondrashov, and S. R. Sunyaev. 2017. Negative selection in humans and fruit flies involves synergistic epistasis. Science. 356:539–542.

Storey, J.D., A.J. Bass, D. Robinson (2021). *qvalue*: Q-value estimation for false discovery rate control. R package version 2.26.0. https://github.com/jdstorey/qvalue.

Storey, J.D., and Tibshirani R. 2003. Statistical significance for genomewide studies. Proc. Nat. Acad. Sci. U.S.A. 100:9440-9445.

Stewart, A. D., E. H. Morrow, and W. R. Rice. 2005. Assessing putative interlocus sexual conflict in *Drosophila melanogaster* using experimental evolution. Proc. R. Soc. B Biol. Sci. 272:2029–2035.

Teplitsky, C., J. A. Mills, J. W. Yarrall, and J. Merilä. 2009. Heritability of fitness components in a wild bird population. Evolution. 63:716–726.

Wade, M. J. 1979. Sexual Selection and Variance in Reproductive Success. Am. Nat. 114:742–747.

Wade, M. J., and S. J. Arnold. 1980. The intensity of sexual selection in relation to male sexual behaviour, female choice, and sperm precedence. Anim. Behav. 28:446–461.

Whitlock, M. C., and A. F. Agrawal. 2009. Purging the genome with sexual selection: Reducing mutation load through selection on males. Evolution. 63:569–582.

1001 Whitlock, M. C., and D. Bourguet. 2000. Factors affecting the genetic load in *Drosophila*:
1002 Synergistic epistasis and correlations among fitness components. Evolution. 54:1654–1660.
1003 Yang, J., S. Mezmouk, A. Baumgarten, E. Buckler, K. Guill, M. McMullen, R. Mumm, and J.
1004 Ross-Ibarra. 2017. Incomplete Dominance of Deleterious Alleles Contributes Substantially
1005 to Trait Variation and Heterosis in Maize. PLoS Genet. 13:e086132.
1006 Yun, L., M. Bayoumi, S. Yang, P. J. Chen, H. D. Rundle, and A. F. Agrawal. 2019. Testing for
1007 local adaptation in adult male and female fitness among populations evolved under different
1008 mate competition regimes. Evolution. 73:1604–1616.
1009 Yun, L., P. J. Chen, A. Singh, A. F. Agrawal, and H. D. Rundle. 2017. The physical environment
1010 mediates male harm and its effect on selection in females. Proc. R. Soc. B Biol. Sci.
1011 284:20170424.
1012 Yun, L., P. J. Chen, K. E. Kwok, C. S. Angell, R. D. Rundle, and A. F. Agrawal. 2018. Competition
1013 for mates and the improvement of nonsexual fitness. Proc. Natl. Acad. Sci. U.S.A. 115: 6762-
1014 6767.
1015 Zhang, J. 2000. Rates of conservative and radical nonsynonymous nucleotide substitutions in
1016 mammalian nuclear genes. J. Mol. Evol. 50:56–68.
1017
1018
1019
1020
1021