Title: Sexually discordant selection is associated with trait specific morphological changes and a complex genomic response

Tyler Audet[1], Joelle Krol[1], Katie Pelletier[1], Andrew D. Stewart[2*], Ian Dworkin[1*]

1 – Department of Biology, McMaster University, Hamilton, Ontario, Canada
2 – Department of Biology, Canisius College, Buffalo, NY, USA

* Joint senior and corresponding authors: dworkin@mcmaster.ca, stewar34@canisius.edu

Author contributions:
Study Conceptualization and funding: ID
Study Design: ID
Artificial Selection and rearing: ADS
Dissections: JK
DNA extraction: TA
Image Analysis: JK, ID
Analysis: ID, TA, KP
Manuscript drafting: TA, ID
Manuscript editing: TA, ADS, ID
Manuscript revisions: TA, ID

Conflict of Interest statement: The authors declare no conflict of interests.

Upon acceptance all data and scripts will be made available via github as well as either Figshare or DRYAD. All raw sequence data will be deposited in NCBI.

Acknowledgement and Funding: This work was funded by an NSERC Discovery grant to ID, and an Ontario Graduate Scholarship to TA.

Abstract

Sexes often have differing fitness optima, potentially generating intra-locus sexual conflict, as each sex bears a genetic 'load' of alleles beneficial to the other sex. One strategy to evaluate conflict in the genome is to artificially select populations discordantly, against established sexual dimorphism, reintroducing attenuated conflict. We investigate a long-term artificial selection experiment reversing sexual size dimorphism in *Drosophila melanogaster* during ~350 generations of sexually discordant selection. We explore morphological and genomic changes to identify loci under selection between the sexes in discordantly and concordantly size selected treatments. Despite substantial changes to overall size, concordant selection maintained ancestral sexual dimorphism. However, discordant selection altered size dimorphism in a trait-specific manner. We observe multiple, possible soft selective sweeps in the genome, with size related genes showing signs of selection. Patterns of genomic differentiation between the sexes within lineages identified potential sites maintained by sexual conflict. One discordant selection lineage shows a pattern of elevated genomic differentiation on chromosome 3L, consistent with the maintenance of sexual conflict. Our results suggest measurable signs of conflict and differentially segregating alleles between the sexes due to discordant selection.

Introduction

While in many species the sexes show phenotypic differences, they must also use nearly the same genome during development to express these phenotypes. This sexual dimorphism (SD) evolves despite the fact that the sexes have a high genetic correlation (r_{MF}) for many traits, which can hinder the evolution of SD (Lande 1980). The extent to which SD can evolve depends on the strength and direction of selection, additive genetic variance, and the inter-sex genetic correlation (Lande 1980; Delph et al. 2011). Many organisms show high r_{MF}, that could limit the potential for sexually discordant evolution (Lande 1980; Poissant et al. 2010). Despite this, SD is very common in nature, in particular, sexual size dimorphism (SSD; reviewed in Fairbairn et al. 2007). In many insects including the pomace fly *Drosophila melanogaster*, females are most often the larger sex (Ashburner 1989). This dimorphism is likely due to the relative contribution of increased fecundity with increased size in females (Honěk 1993; Reeve and Fairbairn 1999). Male-biased SSD occurs in some insects, often due to the relative contribution of sexual selection, and can evolve rapidly within a clade (Emlen et al. 2005; Moczek et al. 2006; Luecke and Kopp 2019). When loci impact the phenotype in a way favouring one sex, but disadvantageous to the other, it may create intra-locus sexual conflict (IASC). This IASC combined with a high r_{MF} between the sexes, may impede the rate of evolution if genetic variation for dimorphism is not also high. Despite high r_{MF} values, family-based artificial selection has demonstrated that sex specific responses can occur relatively rapidly (Alicchio and Palenzona 1971; Bird and Schaffer 1972; Eisen and Hanrahan 1972). Not only can changes in SSD occur rapidly with family-based artificial selection, but r_{MF} has been directly selected upon and degraded in just a few generations (Delph et al. 2011). This reduction of r_{MF} allowed one sex to be selected on for size with minimal response in the unselected sex, where

sex specific selection in high r_{MF} lines resulted in a strong correlated response in the unselected sex. Although kin selection experiments demonstrates additive genetic variation for a response to sexually discordant selection, these do not reflect patterns of transmission of allelic effects in most natural populations. As such, approaches based on mass artificial selection may better reflect transmission of allelic effects in natural populations. Using strong, long-term, mass artificial selection, Stewart and Rice (2018) demonstrated that response to a sex-discordant selection pressure can occur. Stewart and Rice (2018) successfully selected in a sex-discordant manner on body size in *D*. *melanogaster* over 250 generations, with measurable phenotypic changes requiring more than 100-150 generations of artificial selection. In comparison, sex-concordant selection for body size resulted in rapid phenotypic responses. Despite discordant selection responding in family and mass selection experimental designs, one previous study using *Tribolium castaneum*, found little response to sex-discordant selection (Tigreros and Lewis 2011), despite a rapid response to sexually concordant selection on pupal mass. This suggests at least some barrier to divergent response in sexual dimorphism may exist in populations, although it may reflect the modest genetic variation in the ancestral populations and the number (7) generations of artificial selected applied.

The sex-concordant (hereafter, concordant) selection lineages established by Stewart and Rice have had their genomes sequenced at generation 100 to identify candidate SNPs associated with body size variation (Turner et al. 2011). The sex-discordant (hereafter, discordant) lineages have not previously been sequenced, and are the main focus of this current study. To date, functional genetic analyses in *Drosophila* have implicated a few pathways involved with sex determination and growth, that can contribute to SSD. Manipulation of sex-specific splice variants of *transformer* (*tra)* reduces female size, reducing (but not eliminating) SSD (Rideout et al. 2015). Increasing quantities of *tra* also showed a sex-specific increase of size in females (Rideout et al. 2015). A duplication of the *diminutive* (*myc*) gene on the X chromosome resulted in males 12-14% larger, and when paired with constitutive expression of *tra*, SSD was substantially diminished (Mathews et al. 2017). Upstream of *tra*, tissue specific depletion of *sex-lethal* (*sxl*) in neurons, led to reduction of female body size (Sawala and Gould 2017). Inhibition of insulin-signalling (IIS), had sex-specific impacts on body size, largely reducing female body size. In contrast, up-regulation of IIS increased male body size (Millington et al. 2021b). These experiments demonstrate pathways involved with the phenotypic expression of sexual dimorphism, but it is unclear whether segregating variation in these pathways contribute to natural phenotypic variation. Although alleles of large effect in any of the above genes would be exciting to find in a natural population, it is not a safe assumption that genes showing a phenotypic response in a lab setting will be the genes selected upon if a population undergoes discordant selection. Further, rapid response to selection on size as demonstrated by Stewart and Rice (2018) and Bird and Schaffer (1972) among others, suggests that a polygenic response using standing genetic variation, rather than *de novo* mutations, in growth signalling pathways is the likely mechanism for short term body size evolution. In natural populations, segregating alleles contributing to variation in sexual dimorphism interest could be being maintained in the population by selection, or simply reflect mutation-selection-drift balance.

Populations of *D. melanogaster* harbour alleles with both sexually discordant and concordant effects on fitness (Rice and Chippindale 2001). Variation for body size in *D.*

88 *melanogaster* is highly polygenic (Carreira et al. 2009; Turner et al. 2011). Sex limited selection
89 on males has been shown to incur a fitness cost to females (Prasad et al. 2007). This male
90 limited evolution also results in a change in body size in non-selected females that is closer to
91 the male optimum (Prasad et al. 2007). A decrease in female fitness (and body size) when males
92 are allowed to evolve toward their own optimum without parallel female evolution suggests
93 that there is unresolved sexual conflict in the genome of *D. melanogaster* pertaining to body
94 size. Alleles potentially under conflict in the genomes of the outbred LH_m population of *D.*
95 *melanogaster* have been identified, via examination of male mating success and female
96 fecundity (Ruzicka et al. 2019). Alleles under IASC influencing body size, can help understand
97 how the genome responds to selection with divergent phenotypic optima. Selection on body
98 size in a discordant manner has the potential to answer questions of how the shared genome
99 overcomes high r_{MF} when selection favors divergent phenotypes across the sexes.
100 Using lineages evolved under sexually discordant selection for size, first described in
101 Stewart and Rice (2018), we demonstrate two important findings. First, despite selection for a
102 trait-agnostic (*i.e.* selection on general size rather than an increase in mass, or thorax length)
103 measure of size, trait specific patterns of SSD reversal are the norm. Second, despite the long
104 term sexually discordant selection, we see relatively weak evidence for sexually antagonistic
105 alleles being maintained. We do, however, identify one region potentially segregating
106 differentially between the sexes, a possible sign of unresolved (or re-introduced) sexual conflict.
107 Using an evolve and sequence approach we not only examined among lineage patterns of
108 genomic differentiation, but within lineage among the sexes. We discuss our findings within the
109 context of both the evolution of sexual dimorphism and the potential role of ongoing intra-
110 locus sexual conflict.
111
112 ## Materials & Methods
113
114 *Lineages*
115
116 The flies used are part of long-term experimental populations for size selection (Turner
117 et al. 2011; Stewart and Rice 2018). The selection lineages examined were started using the
118 outbred population, LH_M, previously adapted to the lab for over 350 generations prior to
119 selection (and diverged from the LH_M population used in Ruzicka et al. (2019) for an unknown
120 number of generations). These flies are maintained in discrete two-week generations at
121 moderate density (approximate 200 eggs per 10ml standard molasses food vial (Table S1). The
122 complete methodology for maintenance of the base population has been published previously
123 (Rice et al. 2005). For the size selection, flies are anesthetized using CO_2 and sorted using a
124 motorized stacked sieving device in which each successive sieve is 5% smaller than the sieve
125 above. The largest sieve used had aperture diameters of 2000μm and the smallest sieve had
126 apertures measuring 850μm. In each generation, all flies (~1800 individuals), from a selection
127 lineage were sorted by sieve, and 10 vials of 16 mating pairs (320 total flies per lineage) were
128 selected based on phenotype, to generate the next generation. One treatment used only the
129 smallest flies of both sexes (S; concordant selection), one treatment the largest of each sex (L;
130 concordant selection), and the reversal of dimorphism (E; discordant selection) treatment used

131 smallest females and largest males. Finally, a control treatment was populated with flies passed
132 through sieves, but not selected based on size, as controls (C). Two replicate lineages were
133 maintained for each selection treatment. During subsequent selection for size with the
134 established protocol, it was observed that selection can be more stringent on small flies than
135 large because of the nature of the sieve sorting (hinderances due to appendages sticking out,
136 blocking passage of flies through sieve). This would imply that in the sex-discordant selection
137 females being selected to be small were under increased selective pressure compared to males
138 selected to be large. This would suggest that there was a greater pressure for selection in one
139 direction and although this does not explain all results, it may influence response to sexually
140 discordant selection.
141
142 *Morphological measures and analysis*
143
144 At generation 367 (August 19th, 2019) flies from each lineage were collected and stored
145 in 70% ethanol for dissection and measurement of traits. Individuals were chosen randomly
146 before and after the selection treatment to get an accurate estimate of size in the overall
147 lineage. Individuals were dissected under a Leica M125 microscope and legs and thoraces were
148 imaged at 63x using a DCF-400 camera. A minimum of 20 flies (for each sex) were dissected from
149 each lineage by dissecting off the first (pro-thoracic) right leg, then imaging the left side of the
150 thorax of each individual. The leg and left side of the thorax were mounted on slides (in 70%
151 glycerol in PBS, with a small amount of phenol as a bacteriostatic agent). The femur, tibia, and first
152 tarsal segment of each leg, and each thorax, was measured using ImageJ version 1.52q (Schneider
153 et al. 2012). We measured thorax length as a proxy for overall size and a proxy for the general
154 pattern of the female-biased SSD found in *Drosophila melanogaster*. We also measured each leg
155 segment; measurement was completed from the center of the beginning of the segment to the
156 center of the end of the segment.
157 Analysis of the leg and thorax measurements were completed using R v4.1.3 (R Core Team
158 2022). One individual measure of femur was removed as it's measure was ~100X smaller than
159 the mean for the trait. For all analyses of morphology, we used $\log_2$ transformed trait values to
160 facilitate inferences of proportional changes in dimorphism. We fit general linear mixed models
161 for individual traits including sex, selection and sampling and their 2nd order interactions as
162 fixed effects. We allowed sex effects to vary as a random effect of replicate lineage (i.e. random
163 slopes). Models were fit using lmer in the lme4 package v1.1.30 (Bates et al. 2015). We
164 confirmed the results with a fully multivariate mixed model for all traits. For the multivariate
165 mixed model, we allowed for random effects of sex and traits by lineage, as well as accounting
166 for within-individual variation across traits. The initial fit of this model using lmer was singular,
167 likely due to variance estimates getting "stuck" on a boundary (0). We employed two
168 approaches to deal with this. First, we employed a Bayesian extension of our model using blmer
169 in blme v1.0.5 (Chung et al. 2013). This employs weak regularizing priors (away from 0 for the
170 variances). We confirmed the stability of fixed effects using a second approach, fitting a
171 general latent-variable mixed model, to estimate reduced-rank covariance matrices for random
172 effects, as implemented in glmmTMB v1.1.7 (Brooks et al. 2017; Niku et al. 2019; Kristensen
173 and McGillycuddy 2023). Estimated marginal means, custom contrasts, and associated
174 confidence intervals were estimated using emmeans v1.8.0 (Lenth et al. 2018). Visualization

was done using ggplot2 v3.3.6 (Wickham 2018). These approaches provided very similar fixed effects estimates to each other, and to the single trait models, which were used for downstream analysis.

Despite artificial selection being "trait-agnostic" (selecting on a composite cross-sectional area, with possible hinderance from appendages), and a relatively consistent response as outlined below in the results, we examined changes in multivariate allometry among traits within each selective treatment. Commonly, the first principal component derived from a variance-covariance matrix of log transformed morphological measures captures a measure of overall size (Jolicoeur 1963; Blackith and Reyment 1971; Klingenberg 1996). It is common to assess whether the log transformed traits contribute approximately equally to size for PC1 (isometry), with expected loadings of $\frac{1}{\sqrt{p}}$, where p is the number of traits (dimensions). In addition to examining (and visualizing) these vectors from principal components analyses conducted by each selective treatment, we compared aspects of the orientation and structure of the variance-covariance matrices across treatments.

Sex ratio cross

While phenotyping flies for an unrelated experiment, a deviation of the expected 1:1 sex ratio was observed in only a single cross, between our control lineage 1 and discordant lineage 1 samples. As summarized in the results, discordant lineage 1 is also the lineage with potential evidence of maintained conflict on chromosome 3L. As such we conducted experimental crosses to examine adult sex ratios to evaluate whether there is a consistent deviation in sex ratio, potentially due to genomic conflict. At generation 464, 15 males and 15 females were taken from each treatment for single pair matings to an opposite sex individual from the originating LH_m population. The single pair reciprocal matings were allowed to lay until larvae was visible in the food (~5 days), before the F_0 pair were placed in 70% ethanol. From each F1 vial, a single sibling pair was used to generate an F2, while the rest of the F1 individuals were stored in 70% ethanol after being allowed to eclose until most pupal casings were visibly empty. Adults from the F2 generation, once most eclosed, were stored in 70% ethanol. Number of male and female offspring in F1 and F2 generations were counted.

We modeled the data for the sex ratio crosses, using a logistic regression (glm in R), with counts of males and females from each cross as the unit of sampling, with lineage and cross direction and their interactions as predictors. From the model fits, we computed estimates and their confidence intervals on the response scale using emmeans, to facilitate interpretability. We did this both with and without the reciprocal direction of genetic crosses (whether or treatment individual was the sire or dam) in the model. During the experiment to examine adult sex ratios, we noticed substantial differences in number of individual offspring. While this was not a planned analysis (and should be treated as such), we examined differences in fecundity (assessed by census of adults) fitting a general linear model (fit using lm) of number of offspring regressed onto direction of cross, selection treatments and their interactions.

Genomic sample preparation

At generation 378 (February 17th, 2020) flies from each lineage were collected in 70% ethanol and stored at -20°C (Figure 1). For each lineage and sex, flies were separated in to four pools of 25 and DNA was extracted using a column-based DNA extraction kit (DNeasy Qiagen kit, Cat # 69506). The extracted DNA from each of the four pools of 25 for distinct samples were combined such that the same concentration of DNA was added from each pool (equimolar). This resulted in pools of 100 individuals for each combination of sex and unique lineage that were sequenced. This 25 individual pooling was done due to the size of the columns not being capable of extracting from 100 individuals at once. Library preparation and sequencing was done by Génome Québec (Centre d'expertise et de services, Génome Québec). Library preps were prepared with IDT dual index adapters. Sequencing was done using Illumina NovaSeq 6000 to an average coverage of 200x with 151bp sequence fragments. Initial sequencing fell short of 200x coverage, so additional sequencing from the same libraries were conducted to 'top-up' coverage on some samples and were merged with their corresponding samples after mapping.

Bioinformatic Pipeline

Figure S1 shows a summary of the bioinformatics pipeline. A detailed summary can be found at (https://github.com/TylerAudet/SSD_workflow). Reads were trimmed using bbduk v38.90 (Bushnell 2021), aligned using BWA-mem version 0.7.8 (Li 2013), GATK v3.8 was used to mark indels, and perform local realignment. SAMtools v1.5 (Li et al. 2009) was used to convert SAM files to BAM, extract out the core genome, mark/remove read groups, and create mpileups. Three distinct mpileups were made: one with sexes and lineages separate, one with sexes merged maintaining lineages, and one where treatments were pooled together by merging both sexes and lineages together, each of these three went through the same SNP calling and file generation methods. Popoolation v1.2.2 (Kofler et al. 2011) was used to mask repetitive regions identified with Repeat Masker v4 (Smit and Hubley 2013), as well as indels for SNP calling. PoolSNP v1 (Kapun et al. 2020) was used to call variants. From this VCF file indels were masked out using custom scripts from the DrosEU pipeline (Kapun et al. 2020). Due to issues generating sync files from poolSNP VCF output files, a sync file was also generated from the mpileup using Grenedalf version 0.2.0 (Czech et al. 2023). This sync file was filtered for sites present in our SNP called VCF file to create a SNP-called sync file. For 'sexes separate' as well as 'treatments pooled' sync files, through testing, we did observe that changes in the bioinformatic pipeline such as which program was used changed the results slightly, pointing at artifacts introduced by bioinformatic programs. This suggests that great hesitancy and meticulousness must be applied when selecting programs for bioinformatic analyses, as these choices certainly change final results in, at best, a small way. Bioinformatics were done on Compute Canada servers on the Graham cluster.

Among population Genomic Analyses

To identify variants that potentially contribute to phenotypic divergence between treatments, we examined three measurements of population differentiation between populations or genetic diversity within populations. To assess variation between populations

262 we used two related approaches, based upon both F_{ST}, and a modified Cochran–Mantel–
263 Haenszel (CMH) statistic. To assess variation within populations to identify potential regions
264 that have undergone a selective sweep, we performed windowed computation of nucleotide
265 diversity (π). We calculated F_{ST} between control treatments and discordant treatments using
266 the Nei estimation in Grenedalf v0.2.0 in 10kb windows along each chromosome (Czech et al.
267 2023), which corrects for sample size and pooled sequencing size, known sources of error in
268 parameter estimation for pooled sequencing. Although our primary goal was on the changes to
269 SSD, not overall size, we also examined F_{ST} between Large and Small treatments to follow up on
270 the initial analysis in Turner et al. (2011) and see if their findings were replicated after an
271 additional ~275 generations of artificial selection on body size. To do so we ran the sequence
272 data from generation 100 through our pipeline and compared genes of interest from F100 to
273 the genes identified at F378. Effective population size (N_e) was estimated as N_e=181 for all
274 samples using poolSeq v0.3.5 (Taus et al. 2017) using the formulas proposed in (Jónás et al.
275 2016) that are intended to work with pooled data as well as control for additional sources of
276 sampling variance in allele frequencies that could bias N_e. Using this N_e, a CMH statistic was
277 calculated using ACER version 1.0 (Spitzer et al. 2020) for each treatment and replicate as well.
278 Since an ancestral sample was not preserved, drift between control and discordant treatments
279 (or large versus small) were used, and the generations of drift was set at 700, from the control
280 treatment (~generations to convergence x2).
281 For our samples, given the large number of generations of selection, a hard selective
282 sweep would likely result in regions of low nucleotide diversity, which we measured as π using
283 Grenedalf v0.2.0 in 10kb windows for our control, discordant, and concordant samples. This
284 resulted in us having: per sample π values for 10kb windows, F_{ST} values for 10kb windows
285 between discordant and control and between large and small, and SNP-by-SNP adjusted p-
286 values from a CMH test. Using these values, regions with a small π value (< 0.0005) were
287 extracted as interesting, and were extracted along with the top 5% of F_{ST} (discordant versus
288 control cut-off was F_{ST} 0.51 with 621 windows meeting this criteria; large vs. small cut-off was
289 F_{ST} 0.78 with 621 windows meeting that criteria) and SNPs with a CMH adjusted p-value < 0.01
290 (using *p.adjust* in R; Benjamini and Hochberg 1995; 132290 SNPs in control vs. discordant,
291 115747 SNPs in large vs. small) were extracted. The overlap between all windows or SNPs of
292 interest was intersected with Bedtools v2.31.0 (Quinlan and Hall 2010) to get SNPs of interest
293 Unfortunately, most software (including Grenedalf) does not account for hierarchical structure
294 of populations (common in natural populations), nor the expected sources of drift and other
295 forms of sampling variance among replicate lineages (within treatments) in evolve and
296 resequence experiments. As such, the approach we used above (merging replicate lineages),
297 does not account for sampling variation. To account for this, for the subset of "candidate" SNPs
298 we identified we fit SNP specific logistic regression models, with SNP counts for each replicate
299 lineage representing units of sampling, and with selective treatments as the predictor.
300 Treatment contrast between either discordant versus control, or large versus small treatments
301 was obtained using emmeans. As this approach is relatively slow, we only did this for the sites
302 we identified as likely candidates (discordant = 20014, large = 10649, small = 13952). Four files
303 were created with final SNPs of interest: 1) "discordant replicate 1 region" All sites in and
304 elevated region of F_{ST} in replicate 1 of our discordant treatment on chromosome 3L (explained
305 below; supplemental file 1), 2) "discordant sites" all interesting SNPs in the discordant

treatment (supplemental file 2) 3) “Large sites” (supplemental file 3) 4) “Small sites” (same F_{ST} windows used as “large sites”; supplemental file 4). These sites were extracted from the SNP called VCF file using bedtools v2.30.0 (Quinlan and Hall 2010). Using this VCF file, a sync file was created, from which an allele frequency table for all sites of interest was made using Grenedalf, with the reference column set as the females from control replicate 1 major alleles.

Sites of interest found were annotated (Supplemental files 1,2,3,4). Gene ontology (GO) enrichment analyses were conducted using TopGO version 1.0 on each of these four subsets (Alexa et al. 2006). We also explored enriched GO terms using Gowinda v1.12 (Kofler and Schlötterer 2012), which controls for gene length and uses permutation to reduce false positives. Using 1000000 permutations in Gowinda we returned zero significant GO terms in our discordant and small SNPs of interest, and a single term in our large SNPs (GO:0030953: astral microtubule organization). Gene annotations were conducted using SNPeff version 5.1 (Cingolani et al. 2012).

Within population Genomic Analyses to identify regions of genomic conflict.

To identify genomic regions that may harbour variants contributing to IASC, we examined patterns of genomic differentiation between males and females from the same population and generation, for all of the evolved lineages (Cheng and Kirkpatrick 2016; Lucotte et al. 2016; Kasimatis et al. 2019). While simulation studies have suggested this approach generally has low power, the experimental design used for artificial, sexually discordant selection is well suited for this particular approach. We computed allele frequency tables for each sex within each lineage, along with FST between males and females within lineages in Grenedalf. As summarized in the results, we observed a region on chromosome 3L showing elevated F_{ST}, in a single discordant treatment lineage (E1). To assess whether this elevated region of F_{ST} could be accounted for by sampling variation, we simulated 1000000 sites, sampling the full range of allele frequencies used in our analyses, and simulating allele frequencies for male and female samples drawn from a common allele frequency in each simulation. We then allowed sequencing depth (based on the approximate empirical distribution) to vary for each sex. We plotted simulated allele frequencies for males and females, and over-plotted observed allele frequencies for each lineage, to determine the male-female differences in allele frequencies in this chromosomal region were extreme relative to distributions under our simulations that modeled sources of sampling variance. Additionally, we conducted a logistic regression (glm) of major and minor allele counts between the sexes. From this we obtained contrasts and confidence intervals using emmeans.

Results

Selection for size resulted in a trait dependent dimorphism when selection was sex-discordant

Consistent with the previous results for overall “body size” (Stewart and Rice 2018), all measured traits responded to artificial selection in the expected directions. Thorax length was measured as a proxy for body size and had a significant relationship between size and selection

350 lineage (χ^2 = 242.72 p <0.0001; Figure 2; Table S2; Figure S2; Figure S3) with some lineage
351 specific variation (Figure S3). Concordant artificial selection for decreased size reduced thorax
352 length relative to control lineages by ~25% in females (26% in males), while selection for
353 increased size increased thorax length by ~9.5% for both sexes. In the discordant selection
354 lineages, we found a ~10% decrease in female thorax length, and minimal (~0%) change in
355 males. We observed similar patterns for length of leg segments. Small treatment males and
356 females decreased in length relative to controls, to a similar degree (~30% for femur, ~24% for
357 tibia and tarsus; Figure 3). Large treatment males and females increased in size somewhat more
358 modestly relative to controls (~2% for femur, 6% for tibia, 8% for tarsus). For sexually
359 discordant selection, females decreased in size relative to controls (~14% for femur, ~7% tibia
360 and tarsus). Males from the discordant lineage decreased more modestly relative to controls
361 (~8% for femur, < 1% for tibia, ~3.5% tarsus). While some trait changes have confidence
362 intervals that overlap zero, all traits and treatments seemed to respond in size in the expected
363 direction of effect (Figure 2;3; Table S3; S4; S5).
364
365 *Sexual dimorphism and multivariate allometry diverged only under discordant, sex-specific*
366 *selection*
367
368 Despite substantial changes in overall size, SSD remained female biased in all
369 concordant selection treatment groups (Figure 2; 3), showing little change in dimorphism
370 relative to control (Figure 3). In contrast, the sexually discordant selection treatment, resulted
371 in a substantial reduction in the amount of ancestral female biased dimorphism, relative to the
372 controls (Figure 3). Specifically, thorax and tarsus lengths have evolved to be essentially
373 monomorphic, while femur and tibia lengths are now male-biased (~5% increase) in size. The
374 change in the amount of dimorphism was significant for all traits but tarsus length (Figure 3).
375 Given the results described above, it is not surprising we observed substantial changes
376 in patterns of multivariate allometry (Figure 4; Figure S2, S3, S4). We observe a substantial
377 change in the sex-discordant lineages, which deviates from the isometry vector (Table 1) that is
378 observed for other *Drosophila melanogaster* populations (Shingleton et al. 2009). We further
379 compared covariance matrices across selective treatments using the Krzanowski correlation.
380 The discordant selection lineages show a reduced correlation to the control lineages (r_{Krz} = 0.86)
381 compared with the patterns observed for both concordant selection treatments (r_{Krz} = 0.97 and
382 $r_{Krz} = 0.99$ for large and small respectively).
383
384 *Genomic differentiation and nucleotide diversity (π) between selection treatments*
385
386 F_{ST} between control and discordant selection treatments was highly variable, but
387 included substantial differentiation across the genome, and many regions at, or near, fixation
388 (Figure 5A). Notably, average F_{ST} is quite high across the chromosomes, consistent with a
389 substantial impact of genetic drift on allele frequencies. When F_{ST} is calculated between
390 discordant and either of the concordant (small or large) treatments, we similarly see elevated
391 F_{ST} and multiple regions at, or near fixation (Figure S5; S6). In our discordant selection versus
392 controls, the highest mean F_{ST} is found on chromosome 2L (mean F_{ST} per chromosome, 2L =
393 0.215, 2R = 0.177, 3L = 0.144, 3R = 0.171, X = 0.205). In the Large versus Small F_{ST} comparison,

again there were many regions showing high differentiation (Figure 5B). The chromosome with the highest mean F_{ST} in the large versus small comparison was the X chromosome (2L = 0.277, 2R = 0.312, 3L = 0.207, 3R = 0.183, X = 0.354). The highest 95% of F_{ST} windows (window size = 10000bp) for discordant vs. control, as well as large vs. small, were intersected with windows for nucleotide diversity under 0.0005 (windows size = 10000bp) and with SNPs with a CMH (Figure S7; S8) an fdr adjusted p-value less than 0.01. To confirm the effects for SNPs of interest, we modelled allele frequency changes with logistic regression between treatments, examining odds ratio vs. $F_{ST,}$ verifying high F_{ST} correlated to large odds ratios (Figure S9; S10; S11). All candidate SNPs were annotated (Supplemental files 1;2;3;4). After manual curation, several genes with known sex-specific size effects were identified including: *dMyc* (*myc*), *Hairless* (*H*), *Insulin-like receptor* (*InR*), *Regulator of cyclin A1* (*Rca1*), and *stunted* (*sun*). Of our candidate discordant genes, 11/295 (excluding inter-genic SNPs, and lncRNA) have both a known size phenotype as well as a sex-limited phenotype. Many other genes with known effects on aspects of body or trait size were also observed for both the discordant and concordant comparisons (supplemental files 1;2;3;4). For the concordant selective lineages, we examined genes in our list overlapping with those identified from generation 100 of selection Turner et al. (2011). The only gene that overlapped between the analysis of generation 100 and 378 was *Nop1-like* (*Nop17l*) in the small vs. large selective treatments. Modulation of expression of *Nop17l* in developing wing tissues reduces size of the wing (Bennett et al. 2006). If we examine overlapping genes excluding the logistic regression analysis, additional genes overlap between F100 and F378 concordant treatments (Supplemental file 5;6).

Both biological process and molecular function Gene ontology term enrichment was varied

The candidate genomic regions identified above were then examined for enrichment of GO terms across biological process and molecular function. For the discordant sites of interest, the top biological process GO terms pertained to perception of sweet taste, eye development, and sodium ion transport (Table S6). Within the top 20 GO terms there were also terms associated with development (positive regulation of imaginal disc growth), as well as sex-specific terms (female germ-line stem cell population maintenance). For molecular function, top results included cation:cation antiporter activity, lipid transporter activity, and sweet taste receptor activity (Table S9).

For the large selection treatment, circadian behavior, pupal development, and synaptic growth at neuromuscular junctions were the top biological processes (Table S7). Within the top 20 results, *Wnt* signalling also appeared twice (canonical *Wnt* signaling pathway, positive regulation of *Wnt* signaling pathway). *Wnt* signalling was also in the top 20 molecular function in the large treatment regions of interest (Table S10). The Small treatment regions of interest returned similar GO terms, with both circadian behavior, and pupal development being at the top of the list as well (Table S8). *Wnt* signaling also appeared in the small treatment regions of interest for biological processes as well as molecular functions (Table S11). It is worth noting, permuted GO enrichment analysis with Gowinda did not return significant terms, except for our large treatment which only returned astral microtubule organization as an enriched GO term in our large SNP file.

Intersexual genetic differentiation in one sexually discordant lineage may suggest the maintenance of intra-locus sexual conflict

Given the long-term nature of the discordant selection, these lineages may be especially useful to identify signatures of IASC. We examined F_{ST} between males and females within each lineage. As autosomes spend equal time in males and females it is expected intersexual F_{ST} should be close to zero in most circumstances, absent strong sexual conflict (Cheng and Kirkpatrick 2016; Kasimatis et al. 2019). For all concordant selection treatments (Control, Large and Small), we found mean chromosomal F_{ST} to be near zero (mean of C1 = 0.0015, C2 = 0.0019, Figure 6A;S12; L1 = 0.0012, L2 = 0.0015 Figure S13;S14; S1 = 0.0023, S2 = 0.0014, Figure S15;S16). For the discordant selection treatment replicate 2, we also find a mean chromosomal F_{ST} to be near zero (E2 = 0.0011, Figure S17). For replicate 1 of our discordant selection treatment, however, all chromosomal mean F_{ST} is near zero (E1 = 0.0019) except for an elevated section of ~3.4Mb on chromosome 3L between position 18100000 and 21600000, where mean F_{ST} for the rises to 0.005, with elevated SNPs showing a distinct peak of F_{ST} nearing 0.1, with a couple of windows reaching F_{ST} of 0.25 (Figure 6B). We did not identify any common samples with inversions on chromosome 3L (*In(3L)P, 3L133in, 3L165in, 3L096, 3L105, and 3L058)*, nor novel structural usingl DELLY (Rausch et al. 2012) contributing to this elevated region of F_{ST}. We identified SNPs in this E1 lineage on chromosome 3L that were 3 standard deviations about the mean between sex F_{ST} (Supplemental file 1).

We confirmed that this region of elevated intersex F_{ST} in discordant replicate 1 was indeed an outlier using custom simulations designed to account for various sources of sampling variation (Figure S18-S25). All male—female comparisons follow a similar pattern to those of the simulations, with the exception of discordant replicate 1, which appears extreme relative to simulated values (Figure S18). We further modelled all SNPs on chromosome 3L to look for clusters of significant SNPs within our elevated F_{ST} region (Figure S26-S33). Our discordant lineage replicate 1 shows a high number of significant SNPs within our elevated region when modelled (Figure S28), while no other lineage has a clear cluster of significant SNPs in this region. We also looked for genes that overlap with previously identified conflict genes in the establishing population of LH_m (Ruzicka et al. 2020), however we found no overlap. If we exclude our modelled SNPs and look for overlap in genes that overlap without identified high F_{ST} we find a single named gene, Formin-like (*Frl*), which does not have any known body size or sex-limited phenotypes.

Sex ratio and fecundity may suggest conflict in the discordant lineages genomes

Crossing individuals from all treatment groups in reciprocal single-pair matings to the LH_m 'ancestor' resulted in ~1:1 sex ratio in the F_1. The exception to this was Control replicate 1 males crossed to LH_m females (C1: F-M = 0.455, CI = 0.426-0.485; Figure 7), as well as within lineage crosses within both lineages of the discordant treatment (E1: F-M 0.550, CI = 0.517-0.582; E2: F-M 0.550, CI = 0.509-0.589; Figure 7). The deviation in sex ratio in control replicate 1 is also observed in the F2 (C1: F-M = 0.457, CI = 0.427- 0.487; Figure 8). The sex ratios of the F_2s from crosses within both discordant lineages both show male bias, althougwithh the confidence intervals for discordant replicate 1 cross (E1: F-M = 0.557, CI = 0.521- 0.592; Figure 8) not

482 overlapping the 1:1 expectation. We also observed a deviation from the expected sex ratio in
483 our large replicate 1 male to LH_m female cross (L1: M-F = 0.459, CI = 0.429- 0.489; Figure 8), and
484 our discordant replicate 1 male to LH_m female cross (E1: M-F = 0.453, CI = 0.426- 0.480; Figure
485 8). This effect in the discordant cross was in the opposite direction of the discordant replicate 1
486 pure cross.
487 We examined whether the direction of the cross influenced the sex ratio, by adding
488 cross direction as a predictor in the model. In the F_1 generation, overall direction of cross had
489 very modest impacts on sex ratios (χ^2 = 1.13, df = 1, p = 0.29). There was some evidence of
490 deviation from the expected sex ratio when direction was accounted for when control 1 sired
491 the cross (C1: M-F = 0.455, CI = 0.426 - 0.485). In the F_2 generation, the direction of the F_0 cross
492 (sire vs dam) had a modest impact on sex ratios (χ^2 = 5.67, df = 1 p = 0.017). Specifically, when
493 the treatment lineage was the sire in the initial cross for discordant replicate 1, and large
494 replicate 1 (E1: M-F = 0.453, CI = 0.426-0.480; L1: M-F = 0.459, CI = 0.429-0.489). Control
495 replicate 1 crosses appeared to deviate from expected sex ratios regardless of which parent
496 served as sire vs dam.
497 While it was not a planned experiment, we observed a possible difference in fecundity
498 between treatments. The preliminary results from the analysis from the F_1, when the dam was
499 from either discordant lineage, showed reduced fecundity (Figure S34). In the F_2 generation
500 there didn't appear to be an influence of direction of cross, and discordant lineages had the
501 lowest fecundity (Figure S35).
502
503 ## Discussion
504
505 Sexual dimorphism (SD) evolves frequently, despite r_{MF} generally being high within
506 species, in particular for morphological traits (Lande 1980; Poissant et al. 2010). In the presence
507 of sex-specific optimal phenotypes, this r_{MF} has the potential to generate intra-locus sexual
508 conflict through the 'load' on the opposite sex (Fairbairn et al. 2007). Species that evolve
509 changes in SD must overcome any hurdles due to a high r_{MF} as well as ensuing genomic conflict.
510 In the long-term, this conflict may reach an equilibrium, with sex-biased alleles creating as near
511 an optimum phenotype for each sex as possible or be resolved entirely. However, the
512 reintroduction of sex-discordant selection should disrupt this equilibrium and generate
513 additional genomic conflict. Examining the response to sexually discordant selection for body
514 size across the genome after selection provides an opportunity to identify genomic regions
515 undergoing conflict. In this study we utilize long-term artificially selected lineages (Stewart and
516 Rice 2018), selected for body size either in sexually concordant or discordant manner. As
517 discussed in detail below, in addition to observing trait specific changes in patterns of sexual
518 size dimorphism, we see potential evidence for the maintenance of polymorphisms consistent
519 with unresolved conflict.
520
521 *Concordant selection lineages maintained ancestral patterns of SSD despite selection on size,*
522 *while discordant selection lineages responded in a trait specific manner*
523

The lineages under concordant artificial selection for size responded in the expected directions, for all traits, and changed size proportionally maintaining SD for all traits. This proportional response is consistent with relatively high r_{MF} for morphological traits between the sexes, as previously observed (Cowley and Atchley 1988; Reeve and Fairbairn 1996). If r_{MF} had been low in the starting population, we may have observed variation in responses between the sexes, with one sex responding to selection faster than the other, and the overall SSD changing from the 'baseline', as has been shown when r_{MF} is intentionally degraded (Delph et al. 2011). Instead, within concordant treatments, SSD was largely maintained despite substantial changes in size, in particular, under selection for smaller body sizes. For the discordant artificial selection treatment, SD changed trait-specifically (Figure 2). Of the four traits measured for this study, one (thorax length) showed a substantial degree of female biased dimorphism in control lineages (difference of ~0.16, in $\log_2$ μm, or $\frac{\bar{F}}{\bar{M}} \approx 1.12$), while leg traits (femur, tibia and basi-tarsus lengths) were closer to sexually monomorphic (differences of 0.029, 0.023, 0.074, or ratios of 1.02, 1.02, 1.05 respectively). Under discordant artificial selection, all traits saw a reduction in female biased dimorphism, with the magnitude of change varying by trait (Figure 3). Dimorphism for thorax length reduced to the greatest degree (difference of ~0.012, in $\log_2$ μm, or $\frac{\bar{F}}{\bar{M}} \approx 1.01$), compared to dimorphism for leg lengths (-0.076, -0.087, 0.011, or ratios of 0.95, 0.94, 1.01). This is perhaps most clearly illustrated in the changes in patterns of multivariate allometry observed in the sexually discordant lineages (Figure 4). Traits with higher levels of sexual dimorphism tend to have lower r_{MF} than non-dimorphic traits (Poissant et al. 2010), which may partially account for this. However, this may also reflect the manner in which artificial selection was applied (via a sieve), which likely results in a complex composite "size measure" for selection on cross-sectional area, of which a trait like thorax length may contribute substantially more than any individual leg measure. Alternatively, the genetic correlations between the sexes may differ across traits. This might explain why traits like femur and tibia reversed sexual dimorphism when selected on indirectly, however, tarsus and thorax (thorax being the most directly selected on) did not fully reverse dimorphism.

Within-lineage, across-sex comparison provides possible evidence of intra-locus sexual conflict in the genome of one discordant lineage

To identify regions under possible intra-locus sexual conflict we compared male-female F_{ST} within each lineage, and conducted several follow up analyses, to identify possible SNPs showing subtle, intersexual, within-generation distortions in allele frequency. There may be small genomic regions or possibly single SNPs showing signs of different allele frequencies between the sexes, which may be evidence of unresolved conflict in the genome. Neither control lineages, nor the sexually concordant lineages showed any evidence for elevation in F_{ST} (Figure 6A;S12-S17). However, one discordantly selected lineage (E1; Figure 6B), has a region on chromosome 3L with F_{ST} reaching as high as >0.2. Other comparisons of this nature have yielded smaller male-female F_{ST} estimates, which led us to evaluate whether it was an artifact of our pipeline (Cheng & Kirkpatrick 2016; Lucotte et al. 2016; Flanagan & Jones 2016; Dutoit et al. 2017; however see Sylvestre et al. 2023). Modifying SNP calling methods, F_{ST} estimation procedures, and filtering methods all yielded the same spike in F_{ST} in the E1 lineage. This

567 elevated region of allele frequency differentiation was also extreme for only this lineage in
568 comparisons to simulations accounting for a number of sources of sampling variation and
569 overall allele frequencies differences. If it was a spurious peak of elevated F_{ST} as a result of
570 technical aspect of the pipeline, or mis-mapping of reads from a sex chromosome, other
571 replicate lineages would likely also show elevated regions. We consider the most likely
572 explanation to be due to the sustained nature of sexually discordant selection in our
573 experiment. Compared with most other studies examining M-F differences in allele frequencies
574 to identify potential loci under conflict, the selection applied to the discordant lineages in this
575 experiment is both strong, and sustained across hundreds of generations. Kasimatis et al.
576 (2019) suggests that for a locus to show this degree of asymmetry it would require an
577 exceedingly high magnitude of selection. In their model, this would require near 40% mortality
578 each generation to maintain the distortion in M-F allele frequencies at a locus. The nature of
579 artificial selection as applied in the current experiments means well over 40% (~80% of
580 individuals with the current selection design) of flies each generation (those not at phenotypic
581 extremes) are discarded, essentially 'dead' in the selective sense. Alternatively, Kasimatis et al.
582 (2019) also simulated subsamples of 50 males and 50 females with antagonistic selection and
583 found that sampling substantially increased variability in F_{ST} values due to the fact that
584 genomes under antagonistic selection tend to have increased numbers of intermediate
585 frequency alleles. However, our simulations suggest this is not a likely explanation for the
586 observed results (Figure S18-S25). Why this region in discordant treatment replicate E1 shows
587 this distortion, and the contribution of intra-locus conflict, requires further exploration.
588 Given the strange peak in this one discordant replicate, we extracted all SNPs three
589 standard deviations above mean F_{ST} on chromosome 3L, that also were significant in the
590 logistic regression (which better accounts for sampling variance). We filtered for SNPs that
591 appear in both discordant replicates and do not occur in any concordant lineage to look for
592 possible replicated mechanisms of divergence in phenotype. We found fewer candidate genes
593 than in any other comparison (Supplemental file 1), but multiple genes present in the WNT
594 pathway, including *frizzled2* (*fz2*). We also found candidate SNPs in the MAPK pathway.
595 Interestingly, none of the male-female discordant treatments SNPs showed evidence for
596 genetic differentiation between artificial selection treatments. Similar to other comparisons
597 however, many candidate genes have mutational phenotypes listed on flybase as "abnormal
598 size" so this could be suggestive again that response to selection is highly polygenic with alleles
599 of small effect "inching" towards a phenotype rather than "sprinting".
600 Finally, we explored genes found by (Ruzicka et al. 2019), identified as potentially under
601 sexual conflict. These genes were considered in conflict based on sex-specific fitness
602 (competitive mating success in males and competitive fecundity in females). An important
603 caveat in our comparison is that although our lineages are from the same initial lab population
604 (LH_m), our treatment groups have gone through nearly 400 generations of strong selection from
605 the ancestral population. It is also unknown how many generations of divergence occurred in
606 the ancestral LH_m populations prior to both experiments. In our comparison between genes
607 identified in Ruzicka et al. (2019) those identified in our study, only the gene Formin-like (*frl*)
608 overlaps. *Frl* is upstream of our region of interest and does not have any previously identified
609 body size or sex-specific phenotypes. We may not see overlap because Ruzicka et al.(2019)
610 identified conflict alleles based on specific proxies of fitness, whereas ours may reflect a more

611 indirect effect due to discordant selection on size. However, we do observe some evidence
612 that adult sex ratios are male skewed in the discordant lineages, and possibly maintained even
613 in crosses to the LH_m "ancestor", as well as potential reductions in fecundity (Figure 7, 8, S34,
614 S35).
615
616 *Discordant size selection likely has a polygenic basis, involving a number of genes that influence*
617 *body size in a sex-specific manner*
618
619 Given the number of generations of artificial selection, and the modest population sizes
620 each generation, genetic drift will have an overwhelming impact on any genome scan
621 performed with the evolved lineages. As such, any interpretation of these genome scans should
622 be tempered by this impact. With that being said, the response is broadly consistent with a
623 polygenic genetic basis, with thousands of SNPs showing evidence of genetic differentiation
624 across the genome. In 378 generations of selection and more than 125 generations after clear
625 evidence of phenotypic reversal of SSD, the opportunity for a mutation (or ancestrally
626 segregating, relatively rare variants) of large phenotypic effect occurring was possible.
627 However, we find no clear evidence for alleles of large effect that have swept to fixation in
628 parallel across both lineages of discordant selection. The chance of identifying alleles of large
629 effect is attenuated by the impact of drift making many SNPs fix randomly in each lineage.
630 Possibly constrained by genetic correlation, the response to discordant selection appears to be
631 a slow and steady crawl towards a response rather than large leaps forward. The lineage
632 specific increase in discordant replicate 1 that does not appear in replicate 2 could imply a
633 novel mutation that is under conflict in this replicate. This elevated F_{ST} is shown in a large region
634 of chromosome 3L that includes hundreds of SNPs however, so to identify a novel mutation
635 responsible requires further investigation. This slow crawl towards a response is supported by
636 the fact that Stewart and Rice (2018) saw a negligible phenotypic response to discordant
637 selection during at least the first 100 generations, with the more substantial responses
638 occurring after this. Large effect alleles with sex-specific impacts on size have been identified in
639 functional studies, however the viability cost in the opposite sex may make these large effect
640 alleles prohibitive (Rideout et al. 2015; Millington et al. 2021a). Although it is tempting to
641 speculate that these small effect alleles are a likely way for SSD to evolve in natural
642 populations, we must also caveat that it may be the case that these alleles of such small effect
643 may be weeded out by viability or fecundity selection in the wild, which are much weaker in
644 this tightly controlled artificial selection experiment. In fact, Testa and Dworkin (2016)
645 demonstrated that mutations with a sex-specific response are rare in two distinct genetic
646 backgrounds. Carreira et al. (2009) however, found that many genes associated with size
647 change reduced SD when a P-element was inserted. This highlights the complexity of the
648 genetic architecture of SSD.
649 Although Testa and Dworkin (2016) found that mutations with sex-specific effect on size
650 were rare, they did identify mutations in the EGFR pathway with sex-specific effects. In our sex-
651 discordant lineages we found a number of genes with known sex-specific size effect (d*Myc, H,*
652 *InR, RcA1, sun*). Follow up on these genes to determine the role they play in response to
653 discordant selection is required. Interestingly, we also found a number of enriched GO terms
654 for short term memory, circadian rhythm, and nervous system functions, which don't have any

clear link to body size, but natural populations varying in body size also found to have enrichments for GO terms associated with memory & learning, as well as clock genes (Fabian et al. 2012). This could be further supporting that response to selection is very polygenic and involves a number of, likely highly pleiotropic genes. Other GO terms that were enriched were related to gustatory responses such as sweet taste and response to nutrients which may suggest that a response in how larvae are responding to their environment (eating more) may play a role in size changes, but this requires more follow up.

We also compared out genes of interest to the previously published gene list from Ruzicka et al. (2019) mentioned above. In our comparison we still find modest overlap (11 named genes and 8 unnamed genes). Two genes in particular stand out however, discs large 1 (*dlg1*) which has a known increased size phenotype as well as courtship and oogenesis phenotypes, and *smrter* (*smr*) which has known decreased size phenotypes as well as a sex-limited reduced fecundity phenotype in females. Although the caveats of this comparison mentioned above still stand, it is interesting that we find genes under potential conflict coming up in these two studies both using LHm. Optimistically, this could suggest genes under conflict are maintained in the population and are being captured in both studies exploring sexual conflict, but more work must be done to confirm this. Both of these genes are located on the X chromosome which has been suggested to be enriched in conflict alleles (Rice 1984).

Concordant selection lineages show a response in growth pathways, but do not align with previous analyses

The concordant selection lineages (small and large) showed SNPs of interest in multiple growth-related pathways. Interestingly, we found no clear overlap with analyses done nearly 300 generations prior, by Turner et al. (2011). This result could be due to a few possibilities: 1) inconsistencies with bioinformatics tools and pipeline choices, 2) improvements in sequencing and software refining our more recent search, 3) potentially most biologically interesting, this could be due to multiple 'soft' selective sweeps generating allelic turn-over phenomena in the genes under active selection for size. To circumvent the first possibility, we re-analyzed the F100 data used by Turner et al. (2011) with our pipeline, adjusted for coverage and quality. Importantly (and not surprisingly given when the experiment was done), the sequence coverage per sample is very modest (coverage ~25x) in comparison to our current study (~200x). Using our pipeline, the F100 data had no genes with known size phenotypes overlap with our large genes of interest (Supplemental File 5). In the gene list overlap with our small treatment genes of interest, we retrieved Mnt (*mnt*), which has a known increased body size phenotype, we also identified saxophone (*sax*), which has an abnormal size phenotype, and potentially interestingly we recovered Tousled-like kinase (*tlk*), which has a known decreased body size phenotype (Supplemental file 6). The fact that we recover overlapping size related genes in the small treatment, but not the large treatment is likely due to the stronger selection in this treatment.

Although our candidate genes differ from those listed by Turner et al. (2011), both analyses found genes in related general pathways such as ecdysone signalling and the EGFR pathway. This could be evidence that the standing genetic variation present in the starting population (LH_M) could be fueling the clearly polygenic response and new mutation or

stochastic changes in allele frequency are continuously altering which genes are used to response to the selection over such a large time frame. The pathways both we and Turner et al. (2011) identify have been previously implicated as being under selection in naturally occurring clines in *D. melanogaster*, and therefore may provide interesting insight into natural variation in body size. We identified a substantial number of genes after conservative filtering (126 in Large and 101 in Small) that appeared to be under selection in our concordant selection lineages. A number of these genes have known size phenotypes, however since such a large number of genes in the genome impact body size, it is difficult to say if these are directly responsible for response to selection or are mere coincidence (Carreira et al. 2009; Turner et al. 2011). Additional experimental work will be necessary to distinguish between these possibilities.

Our results are both an important exploration of genomic conflict using artificial selection, as well as a follow up on one of the longest (in generations) artificial selection experiment in animals. Although with such strong and persistent selection, our results may not mirror natural settings, it does demonstrate one possibility of how the genome may respond to discordant selection. We also manage to demonstrate a likely differentially segregating region of the genome in a discordantly selected lineage, possibly a distinct autosomal region of sexual conflict in the genome. Although these works require further experimentation to narrow down and validate alleles of interest, we suggest that we have demonstrated at least one route for the genome to respond to discordant selection and sexual conflict.

743

744 ## Literature Cited

745

746 Alexa, A., J. Rahnenfuhrer, and T. Lengauer. 2006. Improved scoring of functional groups from

747 gene expression data by decorrelating GO graph structure. Bioinformatics 22:1600–

748 1607.

749 Alicchio, R., and D. L. Palenzona. 1971. Changes of sexual dimorphism values in *Drosophila*

750 *Melanogaster*. Ital. J. Zool. 38:75–84.

751 Ashburner, M. 1989. Drosophila. A laboratory handbook. Cold spring harbor laboratory press.

752 Cold Spring Harbor Laboratory Press.

753 Bates, D., M. Machler, B. Bolker, and S. Walker. 2015. Fitting Linear Mixed-Effects Models Using

754 lme4. J. Stat. Softw. 67:1–48.

755 Benjamini, Y., and Y. Hochberg. 1995. Controlling the False Discovery Rate: A Practical and

756 Powerful Approach to Multiple Testing. J. R. Stat. Soc. Ser. B Methodol. 57:289–300.

757 Bennett, D., E. Lyulcheva, and L. Alphey. 2006. Towards a Comprehensive Analysis of the

758 Protein Phosphatase 1 Interactome in Drosophila. J. Mol. Biol. 364:196–212.

759 Bird, M. A., and H. E. Schaffer. 1972. A Study of the Genetic Basis of the Sexual Dimorphism for

760 Wing Length in *Drosophila melanogaster*. Genetics 72:475–487.

761 Blackith, R. E., and R. A. Reyment. 1971. Multivariate morphometrics.

762 Brooks, M., E., K. Kristensen, K. Benthem J. ,van, A. Magnusson, C. Berg W., A. Nielsen, H. Skaug

763 J., M. Mächler, and B. Bolker M. 2017. glmmTMB Balances Speed and Flexibility Among

764 Packages for Zero-inflated Generalized Linear Mixed Modeling. R J. 9:378.

765 Bushnell, B. 2021. BBTools: BBMap short read aligner, and other bioinformatic tools.

766 Carreira, V. P., J. Mensch, and J. J. Fanara. 2009. Body size in Drosophila: genetic architecture,
767 allometries and sexual dimorphism. Heredity 102:246–256. Nature Publishing Group.
768 Cheng, C., and M. Kirkpatrick. 2016. Sex-Specific Selection and Sex-Biased Gene Expression in
769 Humans and Flies. PLOS Genet. 12:e1006170.
770 Chung, Y., S. Rabe-Hesketh, V. Dorie, A. Gelman, and J. Liu. 2013. A Nondegenerate Penalized
771 Likelihood Estimator for Variance Parameters in Multilevel Models. Psychometrika
772 78:685–709.
773 Cingolani, P., A. Platts, L. L. Wang, M. Coon, T. Nguyen, L. Wang, S. J. Land, X. Lu, and D. M.
774 Ruden. 2012. A program for annotating and predicting the effects of single nucleotide
775 polymorphisms, SnpEff. Fly (Austin) 6:80–92.
776 Cowley, D. E., and W. R. Atchley. 1988. Quantitative Genetics of Drosophila Melanogaster. II.
777 Heritabilities and Genetic Correlations between Sexes for Head and Thorax Traits.
778 Genetics 119:421–433.
779 Czech, L., J. P. Spence, and M. Expósito-Alonso. 2023. grenedalf: population genetic statistics for
780 the next generation of pool sequencing. arXiv preprint.
781 Delph, L. F., J. C. Steven, I. A. Anderson, C. R. Herlihy, and E. D. Brodie III. 2011. Elimination of a
782 Genetic Correlation Between the Sexes Via Artificial Correlational Selection. Evolution
783 65:2872–2880.
784 Eisen, E., and J. Hanrahan. 1972. Selection for Sexual Dimorphism in Body Weight of Mice. Aust.
785 J. Biol. Sci. 25:1015–24.

786 Emlen, D. J., J. Marangelo, B. Ball, and C. W. Cunningham. 2005. Diversity in the Weapons of
787 Sexual Selection: Horn Evolution in the Beetle Genus Onthophagus (coleoptera:
788 Scarabaeidae). Evolution 59:1060–1084.
789 Fabian, D. K., M. Kapun, V. Nolte, R. Kofler, P. S. Schmidt, C. Schlötterer, and T. Flatt. 2012.
790 Genome-wide patterns of latitudinal differentiation among populations of Drosophila
791 melanogaster from North America. Mol. Ecol. 21:4748–4769.
792 Fairbairn, D. J., W. U. Blanckenhorn, and T. Székely (eds). 2007. Sex, size, and gender roles:
793 evolutionary studies of sexual size dimorphism. Oxford University Press, Oxford ; New
794 York.
795 Honěk, A. 1993. Intraspecific Variation in Body Size and Fecundity in Insects: A General
796 Relationship. Oikos 66:483–492. [Nordic Society Oikos, Wiley].
797 Jolicoeur, P. 1963. 193. Note: The Multivariate Generalization of the Allometry Equation.
798 Biometrics 19:497–499. [Wiley, International Biometric Society].
799 Jónás, Á., T. Taus, C. Kosiol, C. Schlötterer, and A. Futschik. 2016. Estimating the Effective
800 Population Size from Temporal Allele Frequency Changes in Experimental Evolution.
801 Genetics 204:723–735.
802 Kapun, M., M. G. Barrón, F. Staubach, D. J. Obbard, R. A. W. Wiberg, J. Vieira, C. Goubert, O.
803 Rota-Stabelli, M. Kankare, M. Bogaerts-Márquez, A. Haudry, L. Waidele, I. Kozeretska, E.
804 G. Pasyukova, V. Loeschcke, M. Pascual, C. P. Vieira, S. Serga, C. Montchamp-Moreau, J.
805 … and J. González. 2020. Genomic Analysis of European Drosophila melanogaster
806 Populations Reveals Longitudinal Structure, Continent-Wide Selection, and Previously
807 Unknown DNA Viruses. Mol. Biol. Evol. 37:2661–2678.

808 Kasimatis, K. R., P. L. Ralph, and P. C. Phillips. 2019. Limits to Genomic Divergence Under
809 Sexually Antagonistic Selection. G3 GenesGenomesGenetics 9:3813–3824.
810 Klingenberg, C. P. 1996. Multivariate Allometry. Pp. 23–49 *in* L. F. Marcus, M. Corti, A. Loy, G. J.
811 P. Naylor, and D. E. Slice, eds. Advances in Morphometrics. Springer US, Boston, MA.
812 Kofler, R., R. V. Pandey, and C. Schlotterer. 2011. PoPoolation2: identifying differentiation
813 between populations using sequencing of pooled DNA samples (Pool-Seq).
814 Bioinformatics 27:3435–3436.
815 Kofler, R., and C. Schlötterer. 2012. Gowinda: unbiased analysis of gene set enrichment for
816 genome-wide association studies. Bioinformatics 28:2084–2085.
817 Kristensen, K., and M. McGillycuddy. 2023. The AR(1) covariance structure. https://cran.r-
818 project.org/web/packages/glmmTMB/vignettes/covstruct.html
819 Lande, R. 1980. Sexual Dimorphism, Sexual Selection, and Adaptation in Polygenic Characters.
820 Evolution 34:292–305. [Society for the Study of Evolution, Wiley].
821 Lenth, R., H. Singmann, J. Love, and P. Buerkner. 2018. emmeans: Estimated marginal means,
822 aka Least-Squares Means.
823 Li, H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.
824 arXiv.
825 Li, H., B. H, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, and R. Durbin.
826 2009. The Sequence Alignment/Map (SAM) Format and. bioinformatics.
827 Lucotte, E. A., R. Laurent, E. Heyer, L. Ségurel, and B. Toupance. 2016. Detection of Allelic
828 Frequency Differences between the Sexes in Humans: A Signature of Sexually
829 Antagonistic Selection. Genome Biol. Evol. 8:1489–1500.

830 Luecke, D. M., and A. Kopp. 2019. Sex-specific evolution of relative leg size in *Drosophila*
831 *prolongata* results from changes in the intersegmental coordination of tissue growth.
832 Evolution 73:2281–2294.
833 Mathews, K. W., M. Cavegn, and M. Zwicky. 2017. Sexual Dimorphism of Body Size Is Controlled
834 by Dosage of the X-Chromosomal Gene Myc and by the Sex-Determining Gene tra in
835 Drosophila. Genetics 205:1215–1228.
836 Millington, J. W., G. P. Brownrigg, C. Chao, Z. Sun, P. J. Basner-Collins, L. W. Wat, B. Hudry, I.
837 Miguel-Aliaga, and E. J. Rideout. 2021a. Female-biased upregulation of insulin pathway
838 activity mediates the sex difference in Drosophila body size plasticity. eLife 10:e58341.
839 Millington, J. W., L. W. Wat, Z. Sun, P. J. Basner-Collins, G. P. Brownrigg, and E. J. Rideout.
840 2021b. A low sugar diet enhances *Drosophila* body size in males and females via sex-
841 specific mechanisms. Physiology.
842 Moczek, A. P., T. E. Cruickshank, and A. Shelby. 2006. When Ontogeny Reveals What Phylogeny
843 Hides: Gain and Loss of Horns During Development and Evolution of Horned Beetles.
844 Evolution 60:2329–2341.
845 Niku, J., F. K. C. Hui, S. Taskinen, and D. I. Warton. 2019. gllvm: Fast analysis of multivariate
846 abundance data with generalized linear latent variable models in r. Methods Ecol. Evol.
847 10:2173–2182.
848 Poissant, J., A. J. Wilson, and D. W. Coltman. 2010. Sex-Specific Genetic Variance and the
849 Evolution of Sexual Dimorphism: A Systematic Review of Cross-Sex Genetic Correlations.
850 Evolution 64:97–107.

851 Prasad, N. G., S. Bedhomme, T. Day, and A. K. Chippindale. 2007. An Evolutionary Cost of
852 Separate Genders Revealed by Male-Limited Evolution. 9.
853 Quinlan, A. R., and I. M. Hall. 2010. BEDTools: a flexible suite of utilities for comparing genomic
854 features. Bioinformatics 26:841–842.
855 Rausch, T., T. Zichner, A. Schlattl, A. M. Stütz, V. Benes, and J. O. Korbel. 2012. DELLY: structural
856 variant discovery by integrated paired-end and split-read analysis. Bioinformatics
857 28:i333–i339.
858 Reeve, J. P., and D. J. Fairbairn. 1999. Change in sexual size dimorphism as a correlated
859 response to selection on fecundity. Heredity 83:697–706. Nature Publishing Group.
860 Reeve, J. P., and D. J. Fairbairn. 1996. Sexual size dimorphism as a correlated response to
861 selection on body size: an empirical test of the quantitative genetic model. Evolution
862 50:1927–1938.
863 Rice, W. R. 1984. Sex Chromosomes and the Evolution of Sexual Dimorphism. Evolution 38:735–
864 742. [Society for the Study of Evolution, Wiley].
865 Rice, W. R., and A. K. Chippindale. 2001. Intersexual ontogenetic conflict. J. Evol. Biol. 14:685–
866 693.
867 Rice, W. R., J. E. Linder, U. Friberg, T. A. Lew, E. H. Morrow, and A. D. Stewart. 2005. Inter-locus
868 antagonistic coevolution as an engine of speciation: Assessment with hemiclonal
869 analysis. Proc. Natl. Acad. Sci. 102:6527–6534. Proceedings of the National Academy of
870 Sciences.

871 Rideout, E. J., M. S. Narsaiya, and S. S. Grewal. 2015. The Sex Determination Gene transformer
872 Regulates Male-Female Differences in Drosophila Body Size. PLOS Genet. 11:e1005683.
873 Public Library of Science.
874 Ruzicka, F., L. Dutoit, P. Czuppon, C. Y. Jordan, X.-Y. Li, C. Olito, A. Runemark, E. I. Svensson, H.
875 P. Yazdi, and T. Connallon. 2020. The search for sexually antagonistic genes: Practical
876 insights from studies of local adaptation and statistical genomics. Evol. Lett. 4:398–415.
877 Ruzicka, F., M. S. Hill, T. M. Pennell, I. Flis, F. C. Ingleby, R. Mott, K. Fowler, E. H. Morrow, and M.
878 Reuter. 2019. Genome-wide sexually antagonistic variants reveal long-standing
879 constraints on sexual dimorphism in fruit flies. PLOS Biol. 17:e3000244.
880 Sawala, A., and A. P. Gould. 2017. The sex of specific neurons controls female body growth in
881 Drosophila. PLOS Biol. 15:e2002252. Public Library of Science.
882 Schneider, C. A., W. S. Rasband, and K. W. Eliceiri. 2012. NIH Image to ImageJ: 25 years of image
883 analysis. Nat. Methods 9:671–675. Nature Publishing Group.
884 Shingleton, A. W., C. M. Estep, M. V. Driscoll, and I. Dworkin. 2009. Many ways to be small:
885 different environmental regulators of size generate distinct scaling relationships in
886 Drosophila melanogaster. Proc. R. Soc. B Biol. Sci. 276:2625–2633. Royal Society.
887 Smit, A., and R. Hubley. 2013. RepeatModeler Open-1.0.
888 Spitzer, K., M. Pelizzola, and A. Futschik. 2019. Modifying the Chi-square and the CMH test for
889 population genetic inference: adapting to over-dispersion. arXiv.
890 Stewart, A. D., and W. R. Rice. 2018. Arrest of sex-specific adaptation during the evolution of
891 sexual dimorphism in Drosophila. Nat. Ecol. Evol. 2:1507–1513.

892 Sylvestre, F., C. Mérot, E. Normandeau, and L. Bernatchez. 2023. Searching for intralocus sexual
893 conflicts in the three-spined stickleback (*Gasterosteus aculeatus*) genome. Evolution
894 77:1667–1681.
895 Taus, T., A. Futschik, and C. Schlötterer. 2017. Quantifying Selection with Pool-Seq Time Series
896 Data. Mol. Biol. Evol. 34:3023–3034.
897 Testa, N. D., and I. Dworkin. 2016. The sex-limited effects of mutations in the EGFR and TGF-β
898 signaling pathways on shape and size sexual dimorphism and allometry in the
899 Drosophila wing. Dev. Genes Evol. 226:159–171.
900 Tigreros, N., and S. M. Lewis. 2011. Direct and correlated responses to artificial selection on
901 sexual size dimorphism in the flour beetle, *Tribolium castaneum*. J. Evol. Biol. 24:835–
902 842.
903 Turner, T. L., A. D. Stewart, A. T. Fields, W. R. Rice, and A. M. Tarone. 2011. Population-Based
904 Resequencing of Experimentally Evolved Populations Reveals the Genetic Basis of Body
905 Size Variation in Drosophila melanogaster. PLoS Genet. 7:e1001336.
906 Wickham, H. 2018. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, New
907 York.
908 R Core Team (2022). R: A language and environment for statistical computing. R Foundation
909 for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
910

Figure 1: Stereomicroscope images of Control and discordant males and females (25x magnification). Flies chosen at random from populations stored in 70% ethanol and imaged in 70% ethanol.

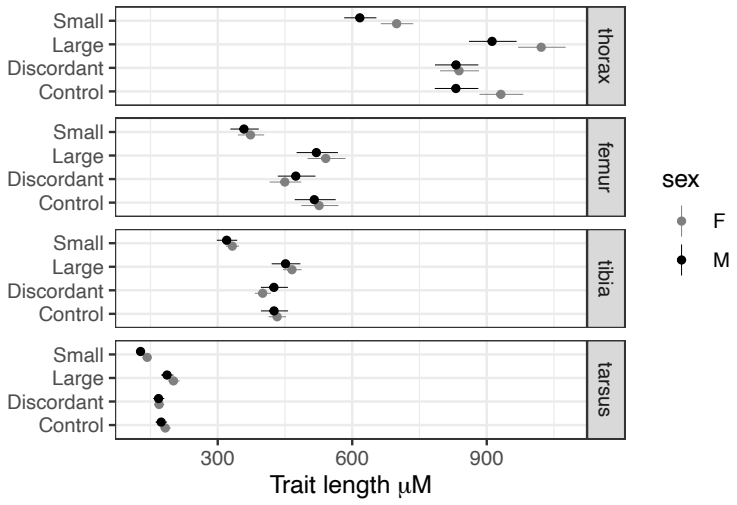

Figure 2: Model estimates for each trait, across treatments and sex. Measurements were $\log_2$ transformed for model fit, and back-transformed only for plotting. Error bars represent 95% confidence intervals.

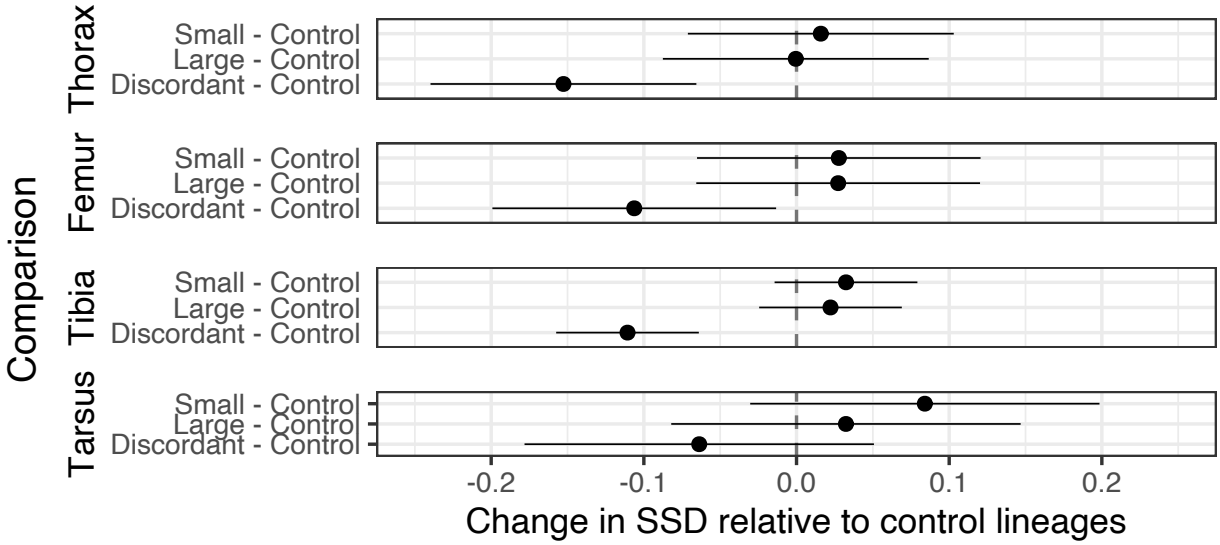

Figure 3: Substantial changes in sexual size dimorphism only occurs under discordant selection on size. Contrasts represent change in the amount of sexual size dimorphism between each artificially selected treatment, in comparison to controls, by trait. Modeled using $\log_2$ transformed length measures (μm), facilitating comparisons of proportional changes. Error bars are 95% confidence intervals for contrasts.

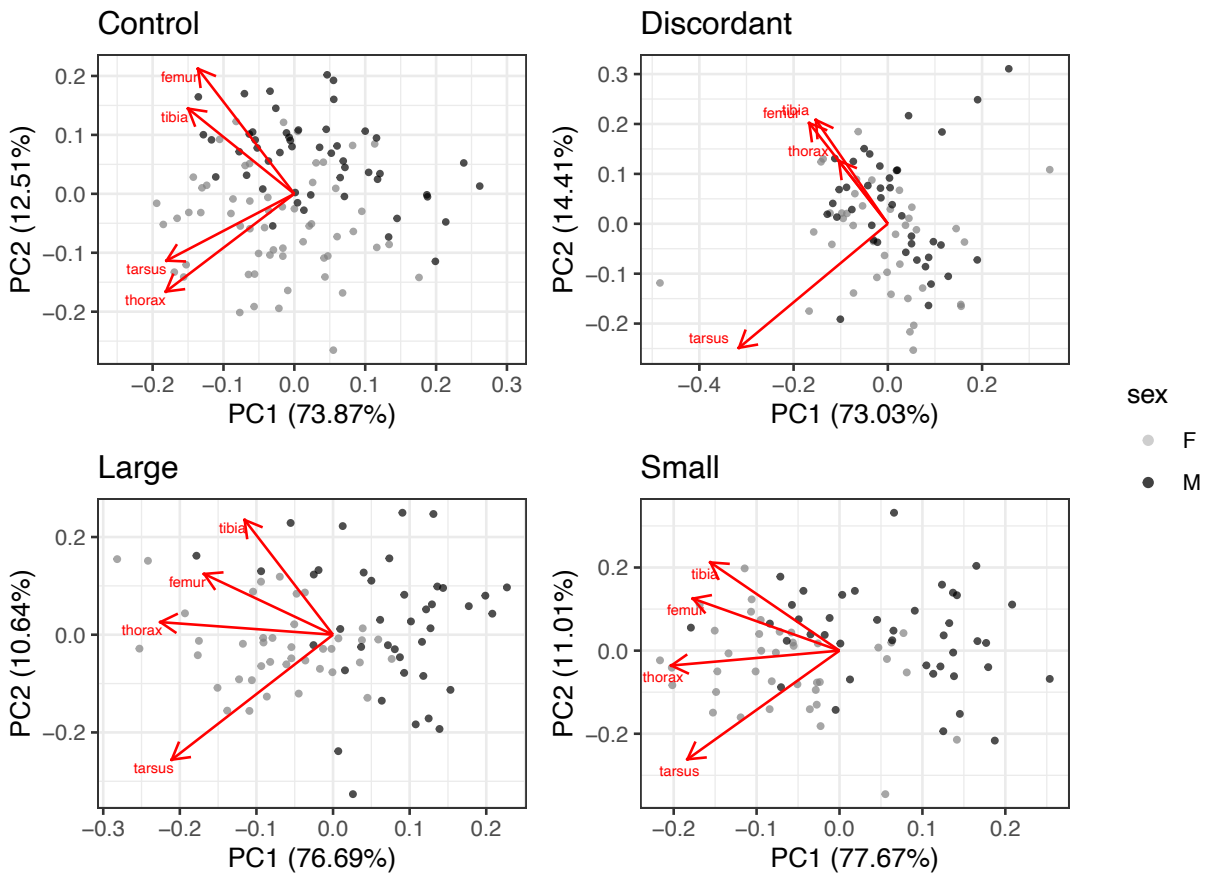

Figure 4: Sexually discordant selection alters patterns of multivariate allometry across sex. Represented as biplots, magnitudes and direction of the loadings for traits are superimposed onto PC1 and PC2. *Log*$_2$ transformed length measures were used.

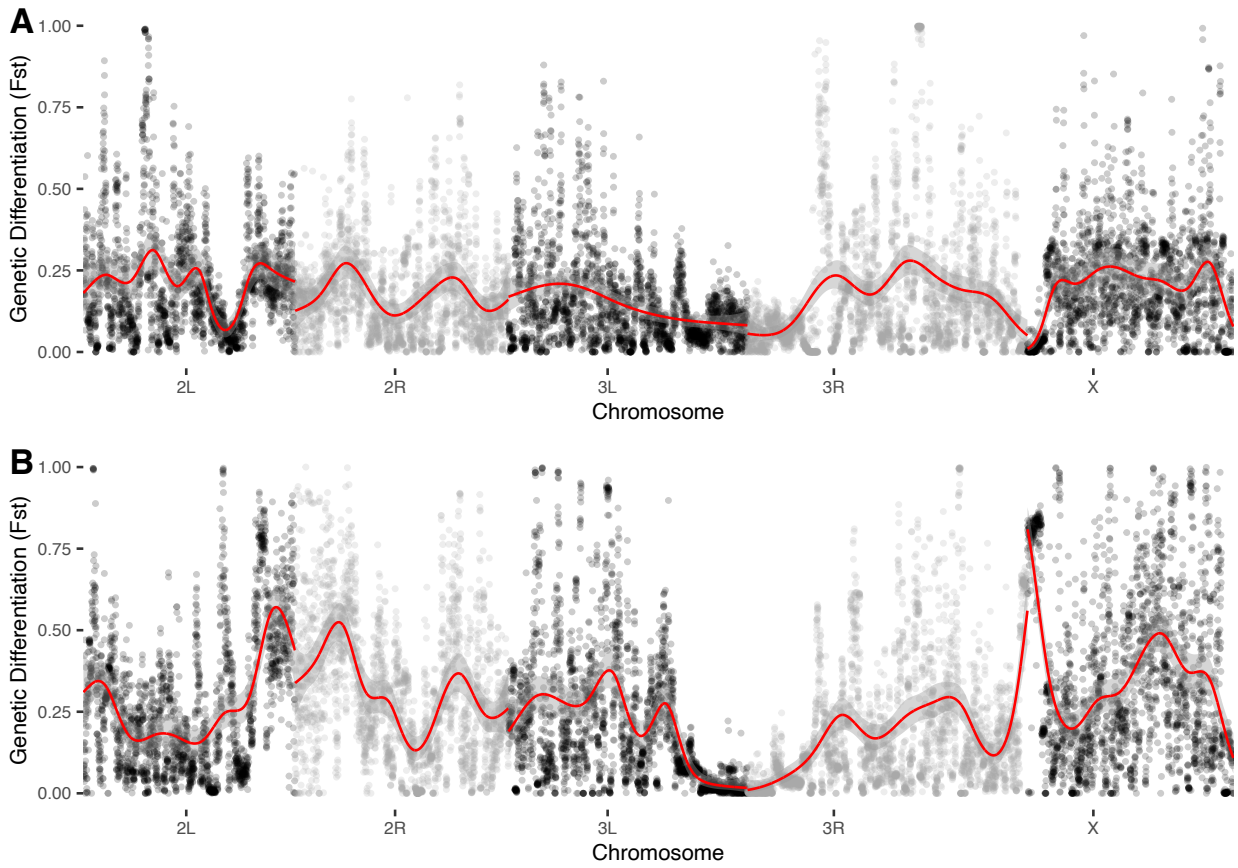

Figure 5: Genome-wide F_{ST} (10000bp windows). Chromosomal trends for F_{ST} (binomial, gamm) in red. A) Discordant selection compared to control treatments. B) Large compared to small treatments.

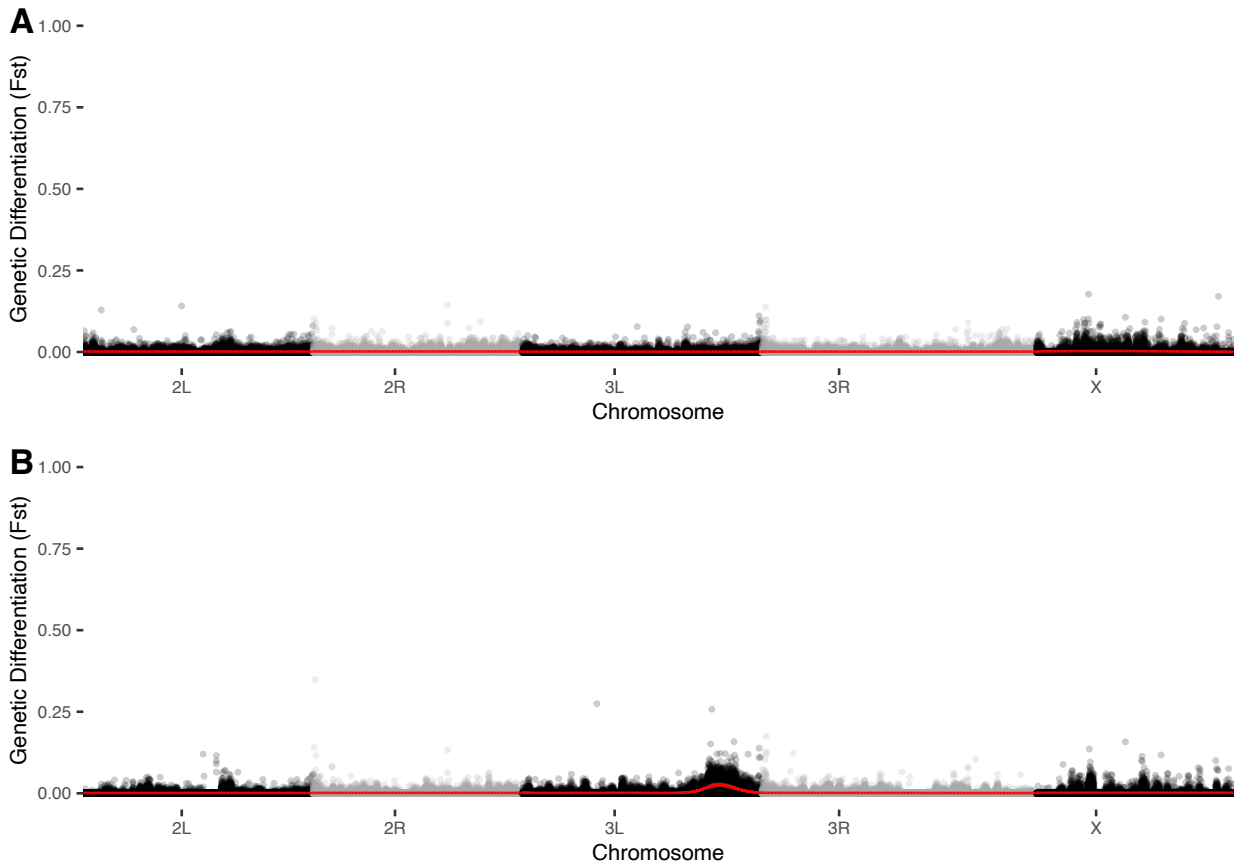

Figure 6: Genome-wide F_{ST} (10000bp windows). Chromosomal trends for F_{ST} (binomial, gamm) in red. A) Male vs. Female for control treatment (replicate 1), showing a trend $\cong 0$, across the genome. B) Male vs. Female for discordant treatment (replicate 1) with a small region on chromosome 3L with elevated F_{ST} values.

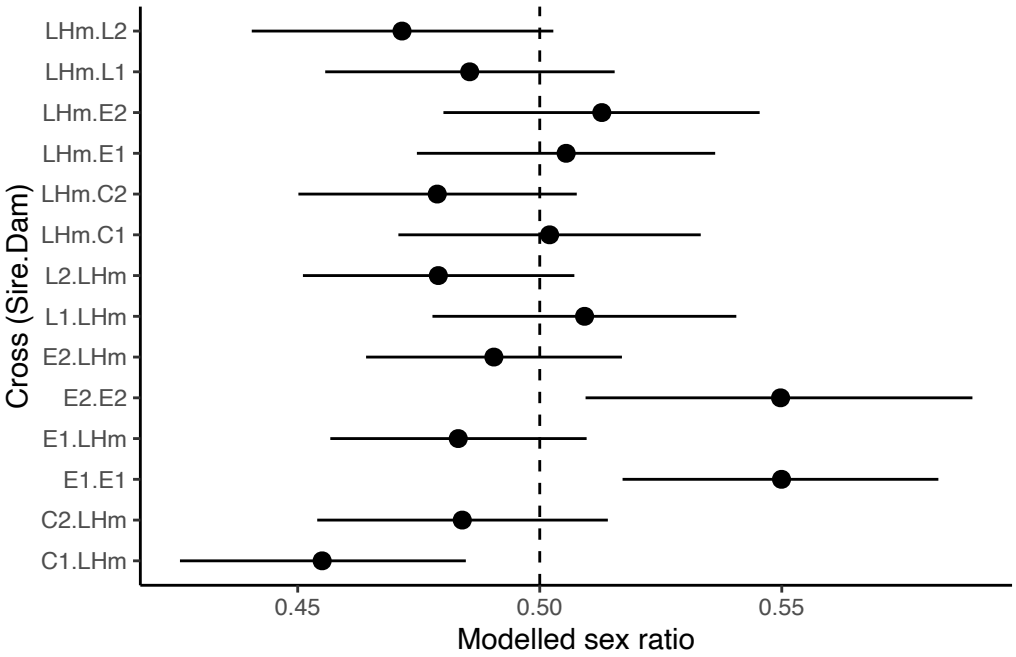

Figure 7: F1 offspring sex ratio from all treatments crossed to the founder population as well as both discordant lineages crossed 'pure'. Cross label has Sire on the left and Dam on the right of the cross identifier. Dashed line marks expected 1:1 sex ratio.

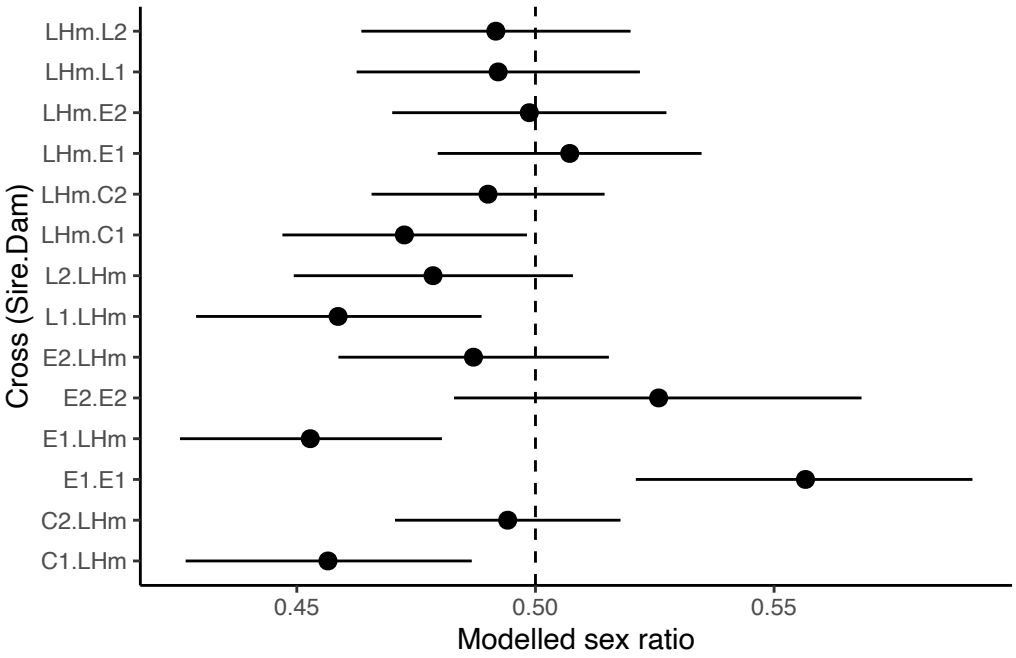

Figure 8: F2 offspring sex ratio from all treatments crossed to the founder population as well as both discordant lineages crossed 'pure'. Cross label has Sire on the left and Dam on the right of the cross identifier. Dashed line marks the expected 50/50 sex ratio.

Tables

Table 1: loadings for PC1, by treatment

Trait	Control (0.74)	Small (0.78)	Large (0.77)	Sex—discordant (0.73)
log_2(femur)	0.417	0.490	0.455	0.415
log_2(tibia)	0.460	0.430	0.312	0.380
log_2(tarsus)	0.553	0.508	0.569	0.786
log_2(thorax)	0.556	0.562	0.609	0.258

Loadings for PC1 (eigenvector 1) from variance-covariance matrices for each treatment. The values next to treatment names correspond to the proportion of variation accounted for by PC1.

Table S1: Food recipe for artificial selection lineages. All ingredients except cornmeal, yeast, and preservatives are brought to a boil for 10 minutes, yeast/cornmeal and a small amount of reserved water are added and boiled for another 10 minutes, food is cooled to 60ºC then preservatives are added before pouring ~10mL in to vials.

Water	1.2L
Agar	11.1g
Cornmeal	75.0g
Yeast	31.0g
Molasses	75mL
methyl-p-hydrobenzoic acid	1.8g
95% EtOH	17.5mL
Propionic Acid	7mL

Figure S1: Flow chart showing steps and tools involved in the bioinformatic pipeline. Version numbers, parameter settings and more detail available in text. White: bam files, grey: intermediate files, green: output results files.

Table S2: model ANOVA for thorax

	χ^2	**Df**	**Pr(>χ^2)**
sex	140.67	1	1.90E-32
selection	242.72	3	2.45E-52
sampling	18.9	1	1.37E-05
sex:selection	38.24	3	2.52E-08
sex:sampling	5.29	1	2.14E-02
selection:sampling	20.28	3	1.48E-04

Table S3: model ANOVA for femur

	χ^2	**Df**	**Pr(>χ^2)**
sex	2.26	1	1.33E-01
selection	109.62	3	1.32E-23
sampling	8.32	1	3.93E-03
sex:selection	21.6	3	7.89E-05
sex:sampling	6.13	1	1.33E-02
selection:sampling	10.59	3	1.41E-02

Table S4: model ANOVA for tibia

	χ^2	**Df**	**Pr(>χ^2)**
sex	1.03	1	3.11E-01
selection	295.82	3	7.98E-64
sampling	28.33	1	1.02E-07
sex:selection	31.91	3	5.45E-07
sex:sampling	9.11	1	2.54E-03
selection:sampling	13.71	3	3.33E-03

Table S5: model ANOVA for tarsus

	χ^2	**Df**	**Pr(>χ^2)**
sex	36.71	1	1.37E-09
selection	133.67	3	8.74E-29
sampling	2.86	1	9.10E-02
sex:selection	13.44	3	3.77E-03
sex:sampling	4.26	1	3.90E-02
selection:sampling	17.5	3	5.58E-04

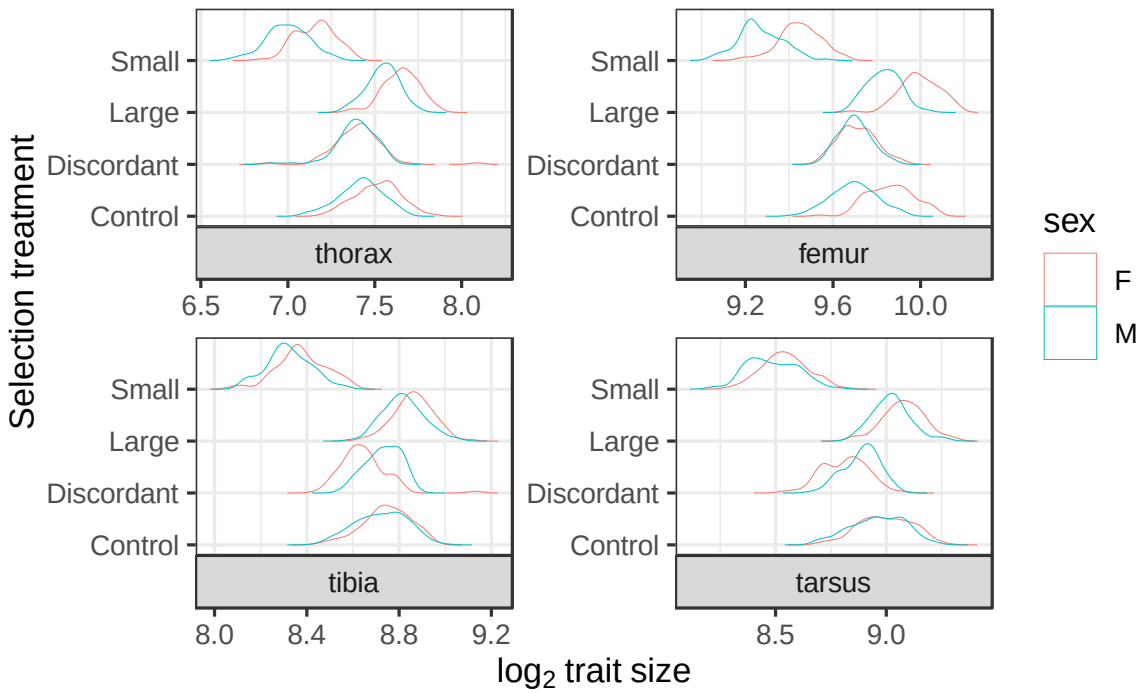

Figure S2: Density plots of raw measures (*log_2* transformed) for traits and sexes

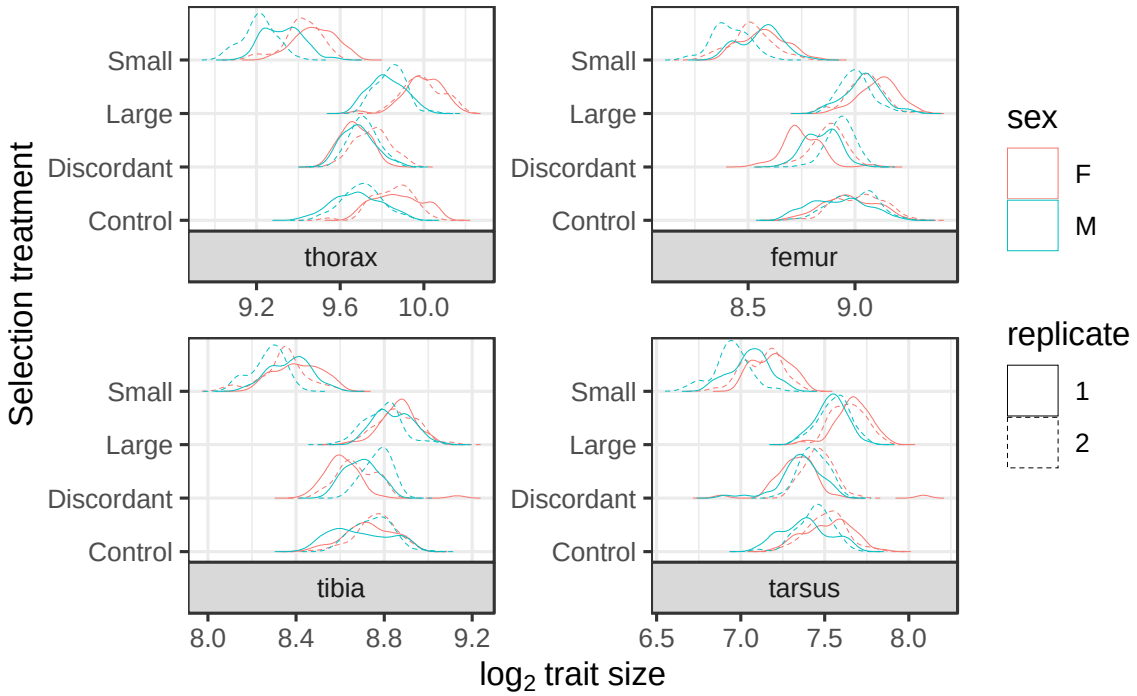

Figure S3: Density plots of raw measures (*log_2* transformed) for traits and sexes with replicates separated to show replicate heterogeneity

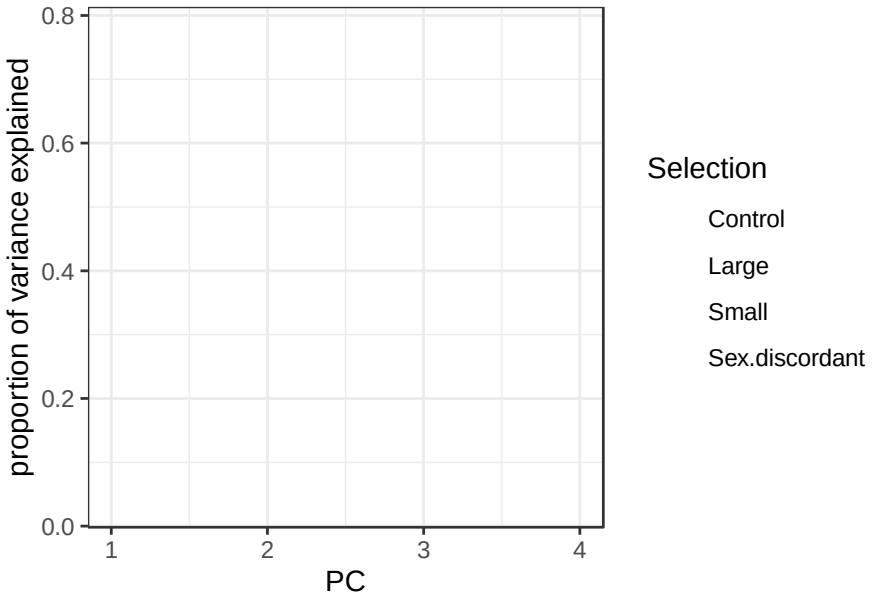

Figure S4: scree plot of proportion of variation explained by each associated eigenvector.

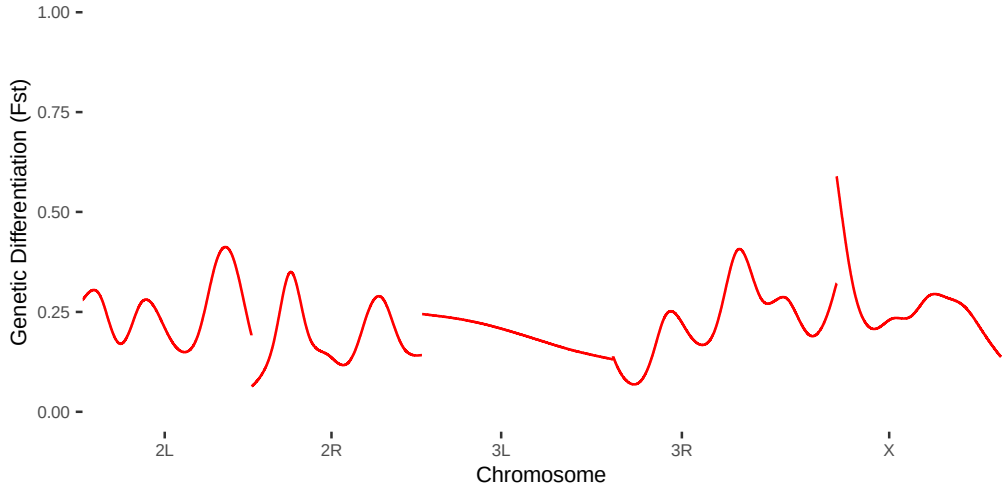

Figure S5: discordant vs. large 10000bp windows

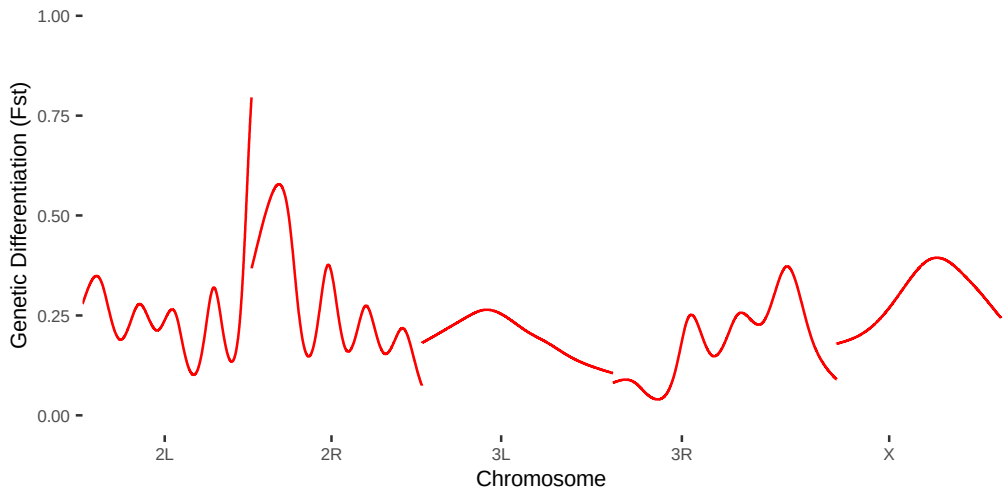

Figure S6: discordant vs. small 10000bp windows

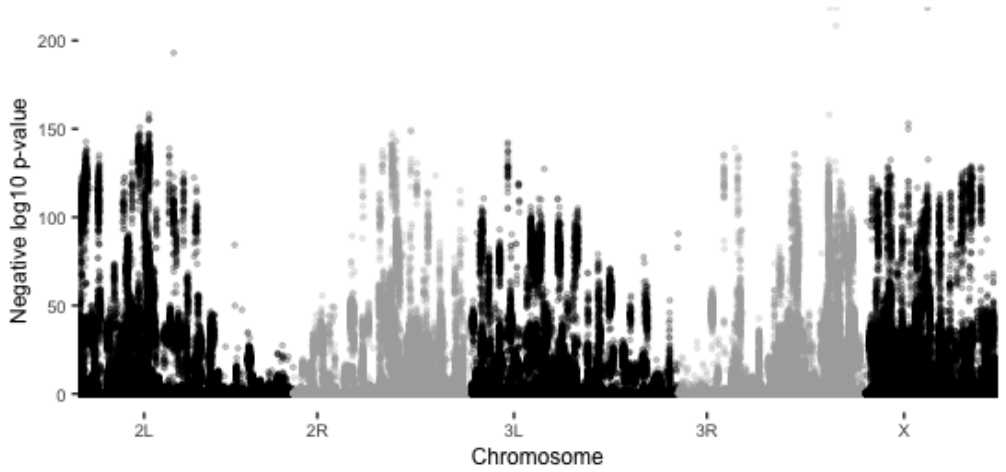

Figure S7: CMH p-values for Large vs. Small treatments

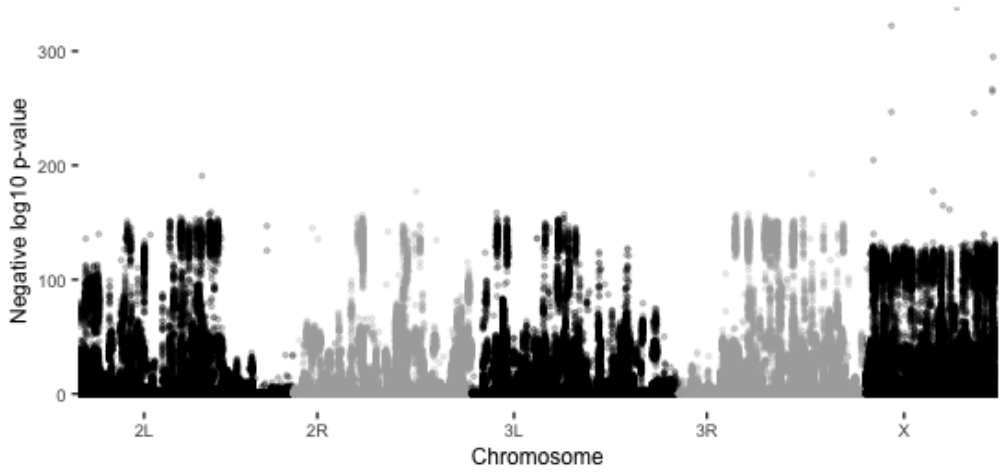

Figure S8: CMH p-values for discordant vs. control treatments

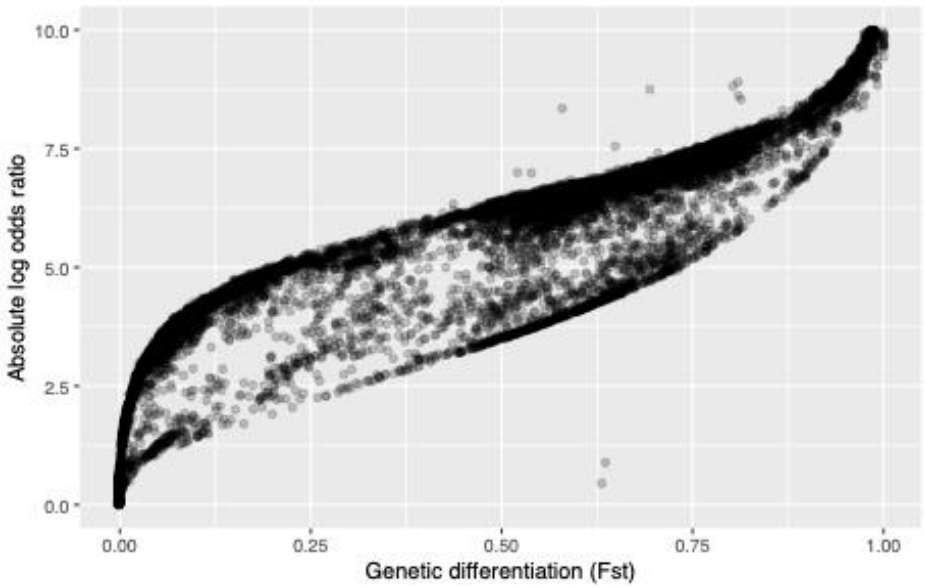

Figure S9: Correlation between F_{ST} and the log odds from our model estimates for the Discordant vs. control comparisons

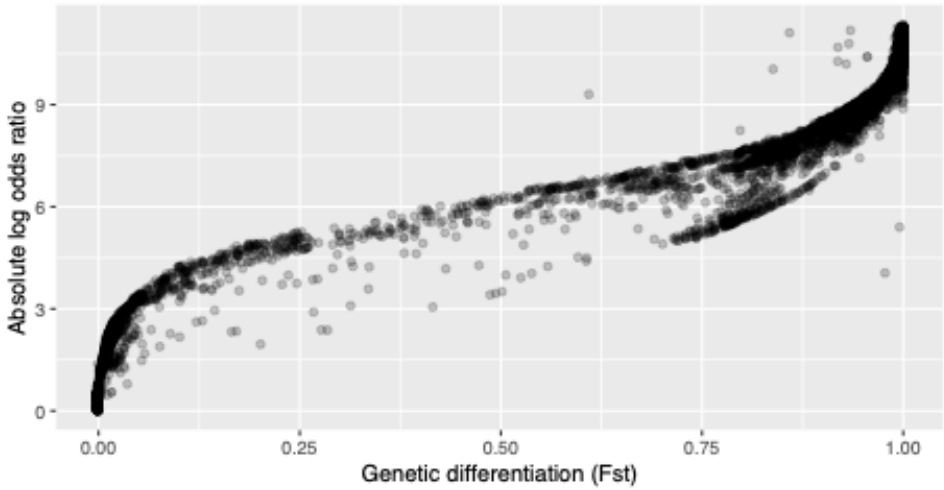

Figure S10: Correlation between Large vs. Small F_{ST} and the log odds from our model estimates for the SNPs found to be associated with the Large treatment

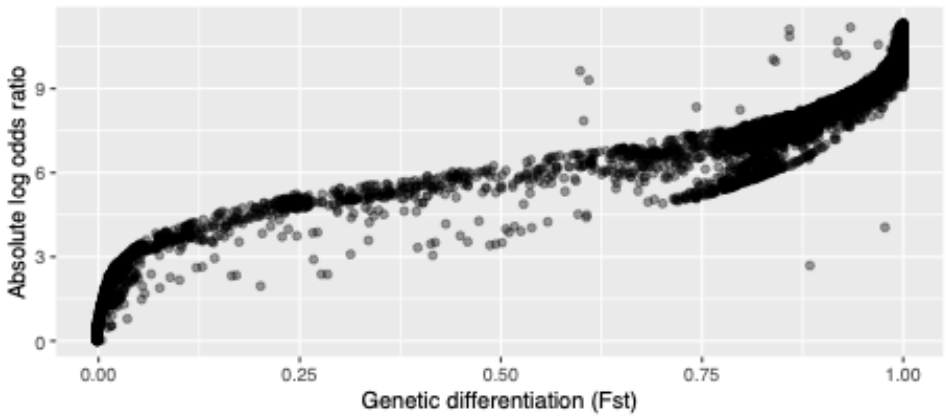

Figure S11: Correlation between Large vs. Small F_{ST} and the log odds from our model estimates for the SNPs found to be associated with the Small treatment

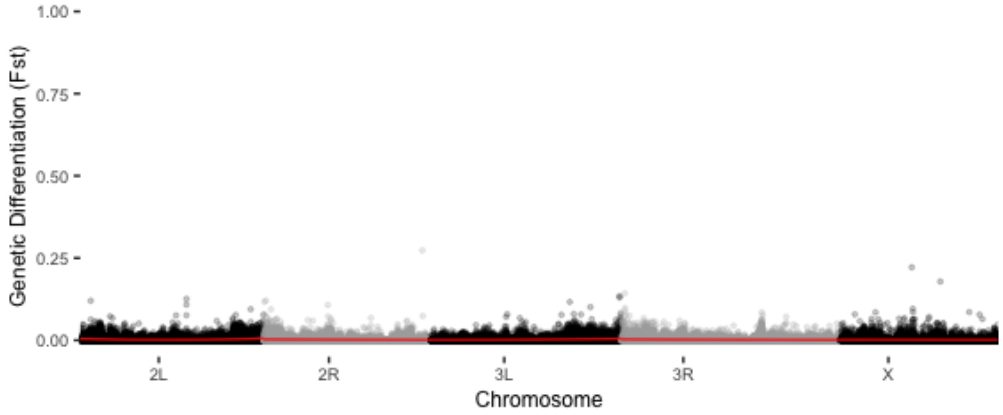

Figure S12: control replicate 2 males vs. females 1000bp

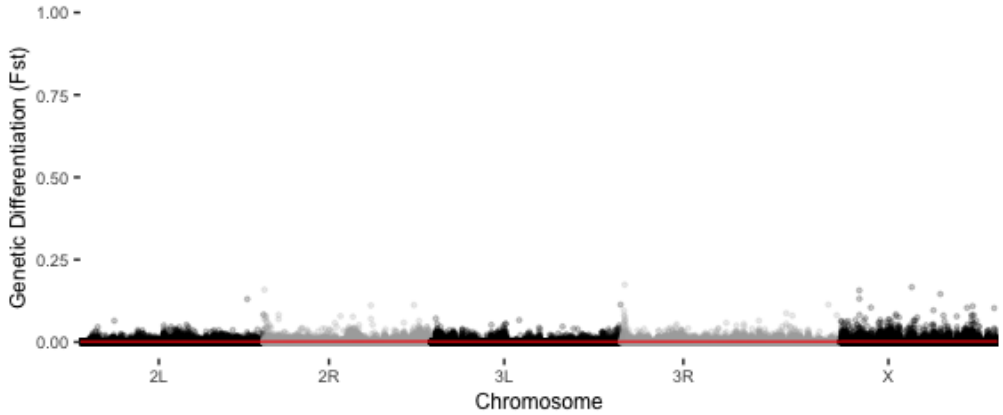

Figure S13: large replicate 1 males vs. females 1000bp

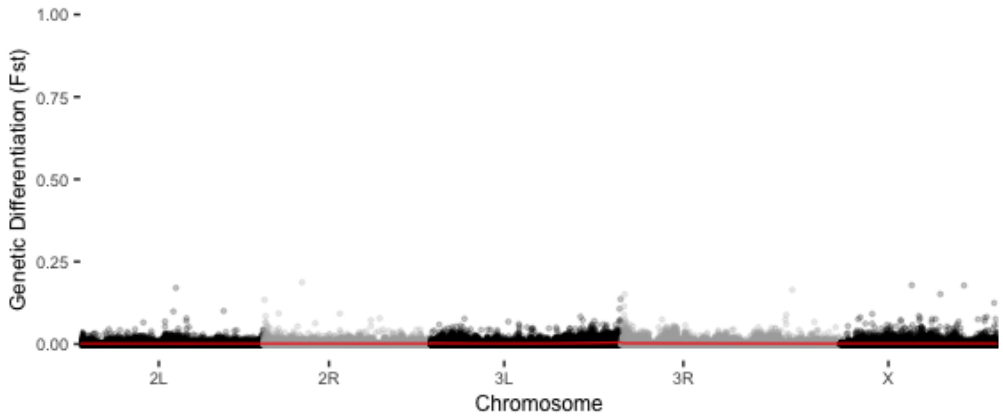

Figure S14: large replicate 2 males vs. females 1000bp

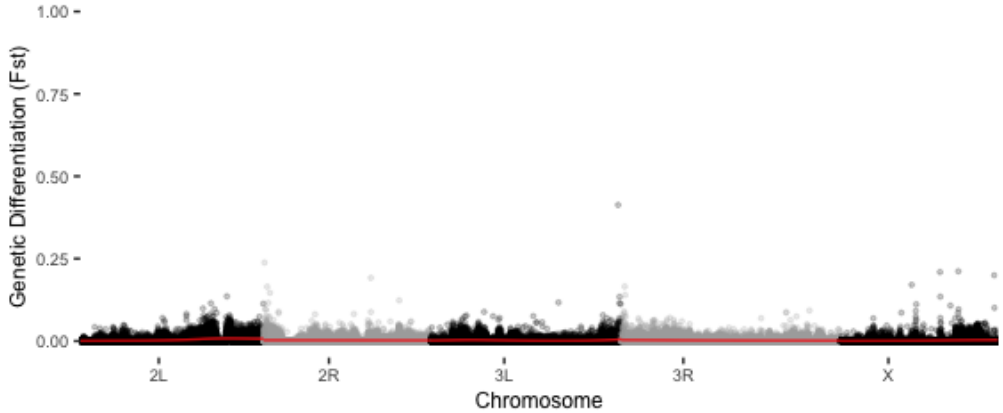

Figure S15: small replicate 1 males vs. females 1000bp

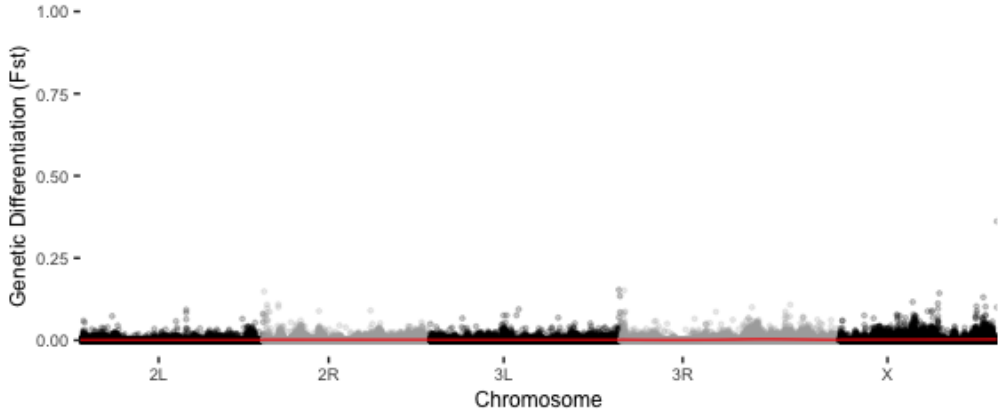

Figure S16: small replicate 2 males vs. females 1000bp

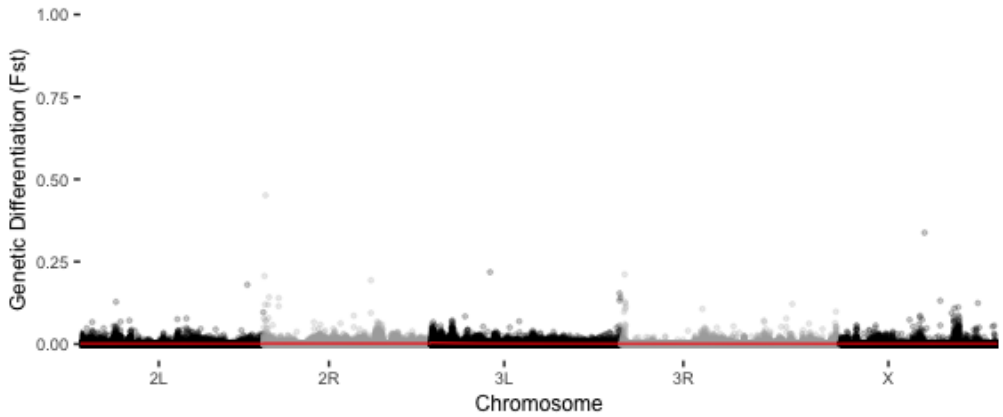

Figure S17: discordant replicate 2 males vs. females 1000bp

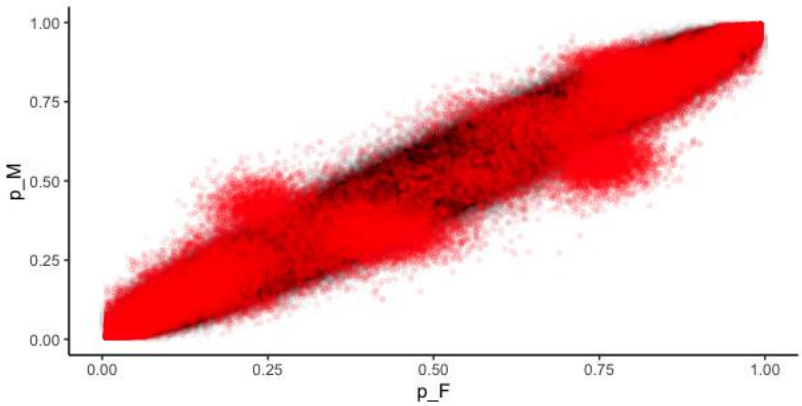

Figure S18: E1 male versus female minor allele frequency (computed as ALT/(REF+ALT)) in red plotted on top of 1000000 simulated variants with similar coverage and allele frequency distributions of our data.

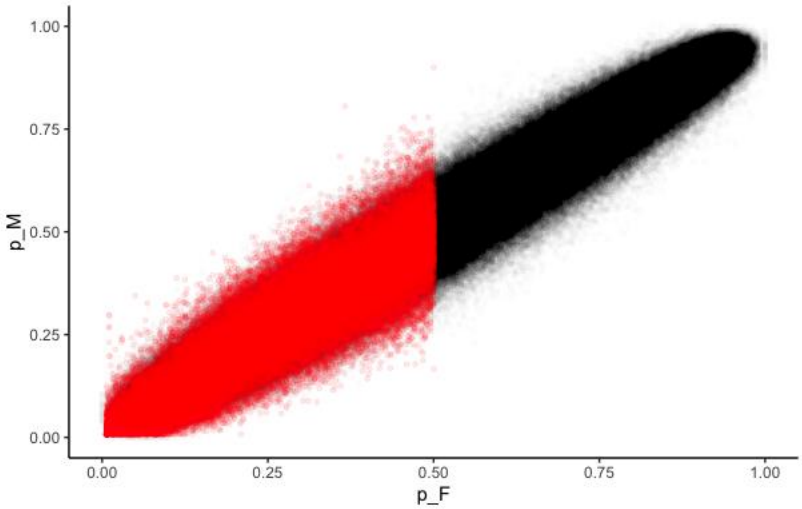

Figure S19: C1 male versus female minor allele frequency (computed as ALT/(REF+ALT)) in red plotted on top of 1000000 samples binomially sampled from a simulated pool with a similar coverage and distribution of our data. ***Please note***, the reference allele frequency for all

treatments are set to the major allele of C1 females (this sample) so no alternate alleles are expected to reach >0.5 frequency.

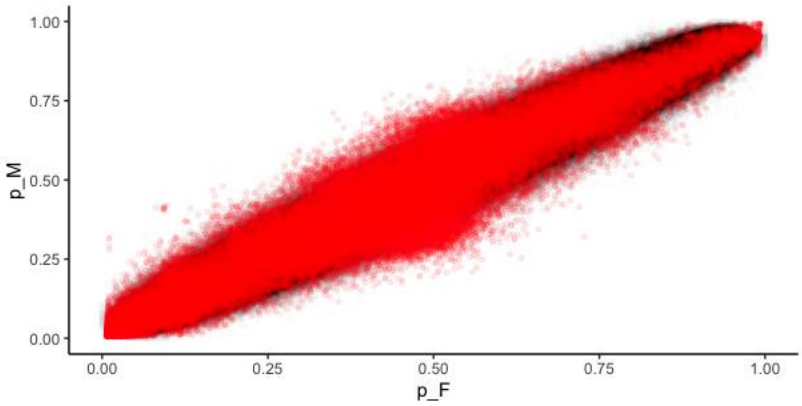

Figure S20: C2 male versus female minor allele frequency (computed as ALT/(REF+ALT)) in red plotted on top of 1000000 samples binomially sampled from a simulated pool with a similar coverage and distribution of our data.

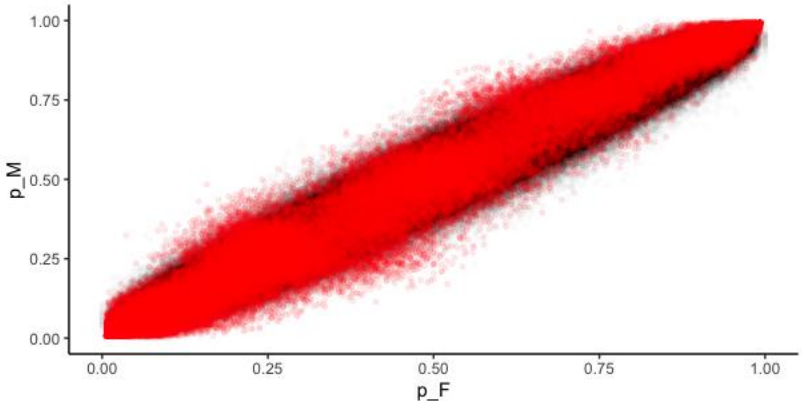

Figure S21: E2 male versus female minor allele frequency (computed as ALT/(REF+ALT)) in red plotted on top of 1000000 samples binomially sampled from a simulated pool with a similar coverage and distribution of our data.

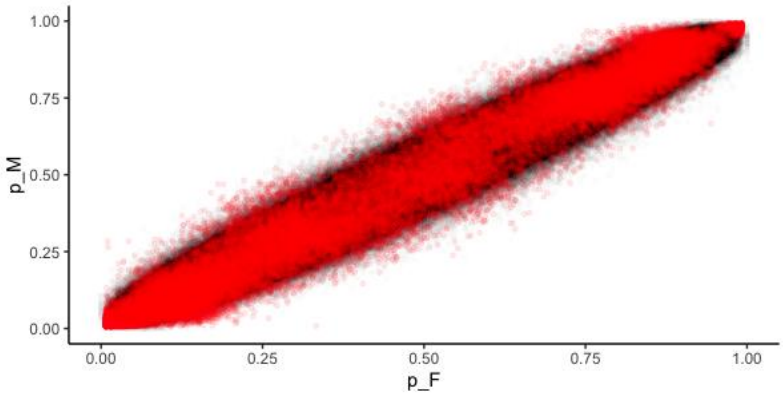

Figure S22: L1 male versus female minor allele frequency (computed as ALT/(REF+ALT)) in red plotted on top of 1000000 samples binomially sampled from a simulated pool with a similar coverage and distribution of our data.

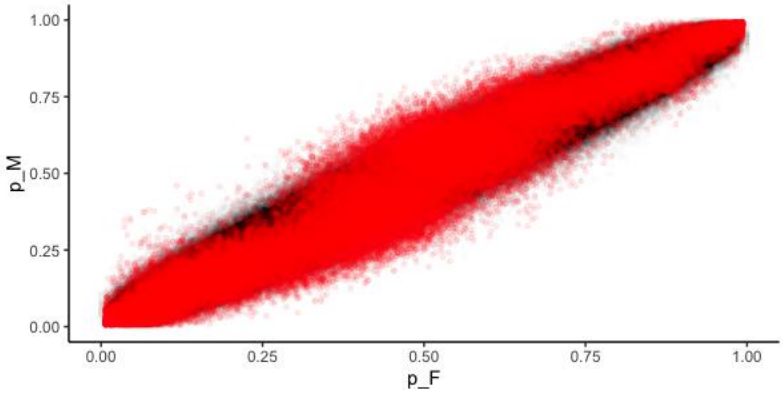

Figure S23: L2 male versus female minor allele frequency (computed as ALT/(REF+ALT)) in red plotted on top of 1000000 samples binomially sampled from a simulated pool with a similar coverage and distribution of our data.

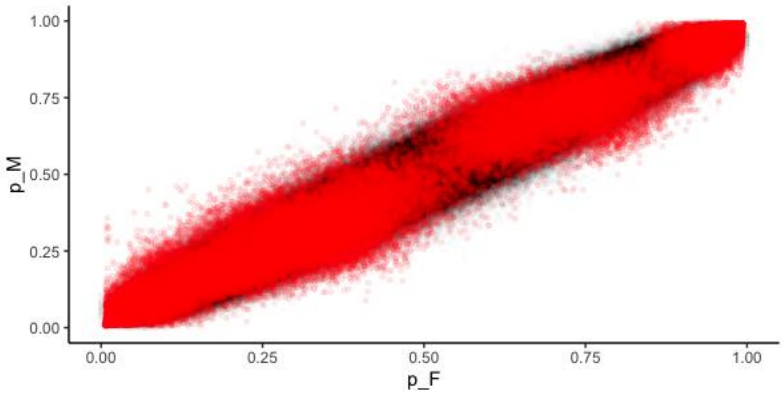

Figure S24: S1 male versus female minor allele frequency (computed as ALT/(REF+ALT)) in red plotted on top of 1000000 samples binomially sampled from a simulated pool with a similar coverage and distribution of our data.

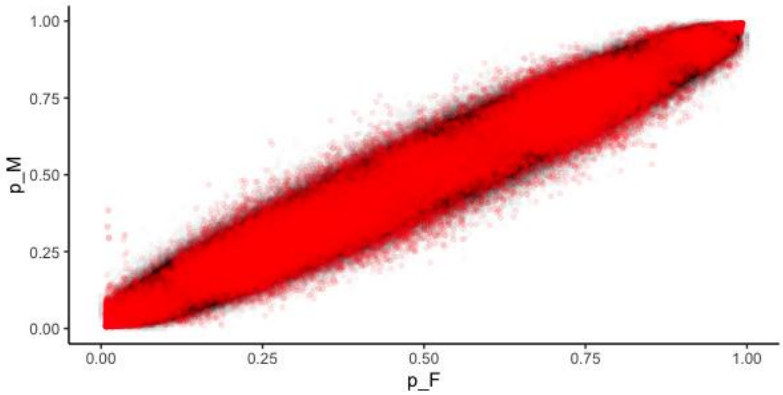

Figure S25: S2 male versus female minor allele frequency (computed as ALT/(REF+ALT)) in red plotted on top of 1000000 samples binomially sampled from a simulated pool with a similar coverage and distribution of our data.

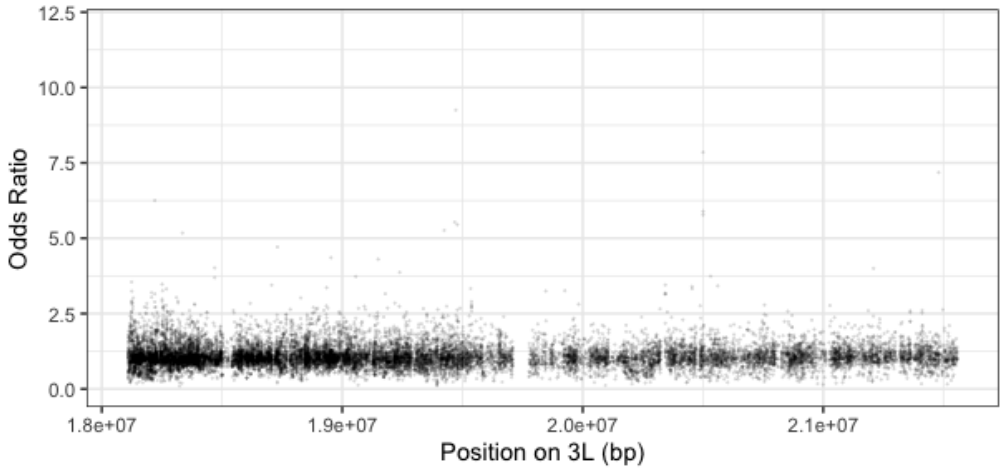

Figure S26: C1 modelled odds ratios by position on chromosome 3L within the elevated region on E1. Positions with a p-value < 0.0001 highlighted red with upper and lower 95% confidence intervals shown.

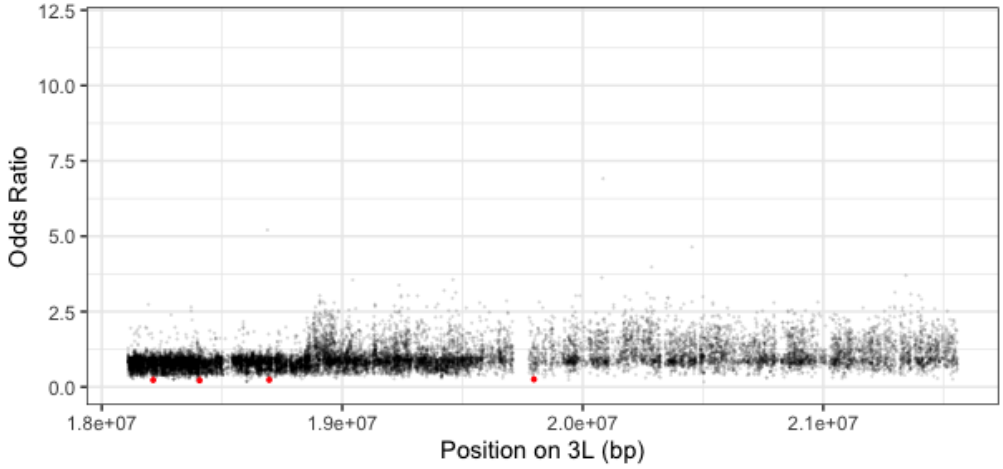

Figure S27: C2 modelled odds ratios by position on chromosome 3L within the elevated region on E1. Positions with a p-value < 0.0001 highlighted red with upper and lower 95% confidence intervals shown.

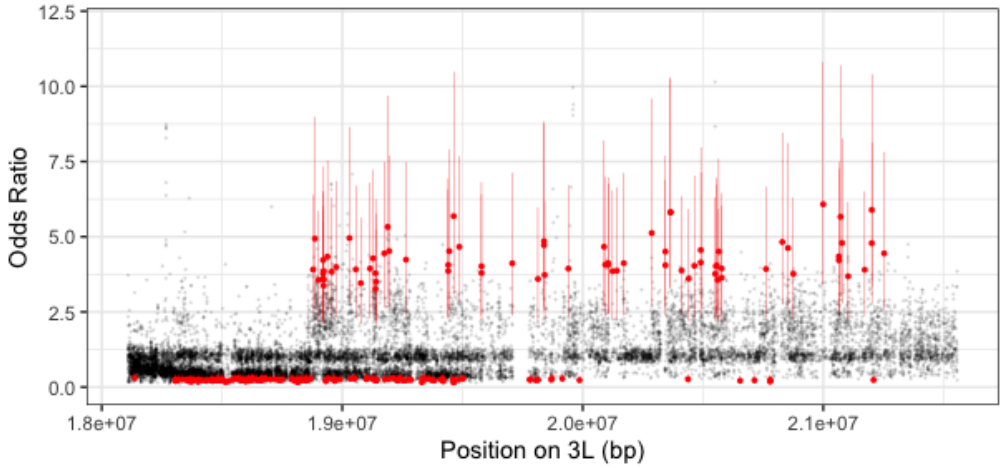

Figure S28: E1 modelled odds ratios by position on chromosome 3L within the elevated region on E1. Positions with a p-value < 0.0001 highlighted red with upper and lower 95% confidence intervals shown.

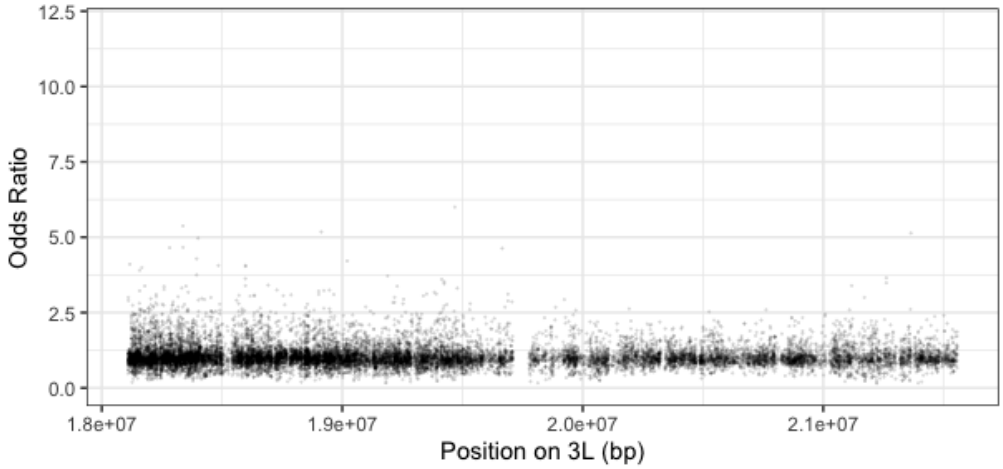

Figure S29: E2 modelled odds ratios by position on chromosome 3L within the elevated region on E1. Positions with a p-value < 0.0001 highlighted red with upper and lower 95% confidence intervals shown.

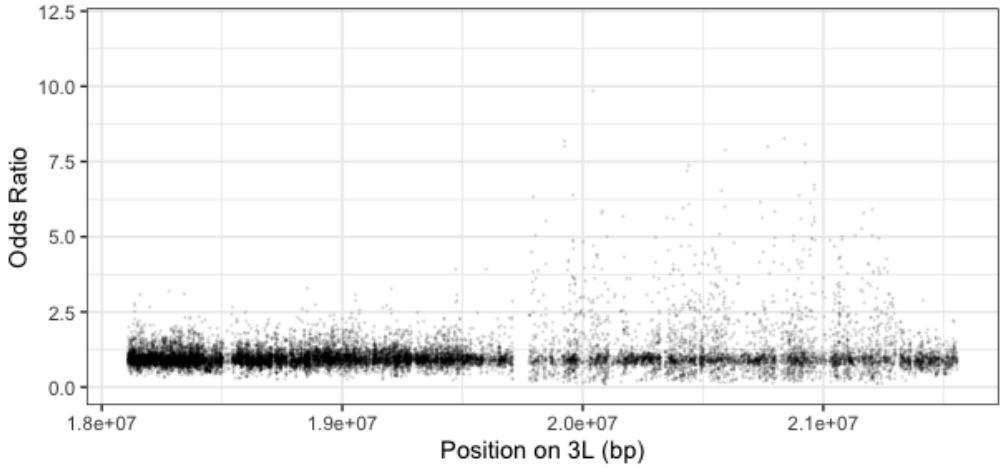

Figure S30: L1 modelled odds ratios by position on chromosome 3L within the elevated region on E1. Positions with a p-value < 0.0001 highlighted red with upper and lower 95% confidence intervals shown.

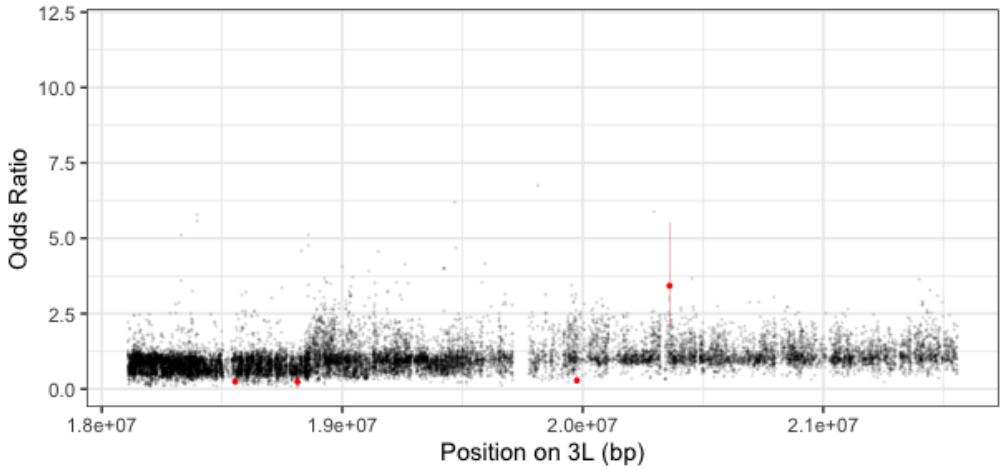

Figure S31: L2 modelled odds ratios by position on chromosome 3L within the elevated region on E1. Positions with a p-value < 0.0001 highlighted red with upper and lower 95% confidence intervals shown.

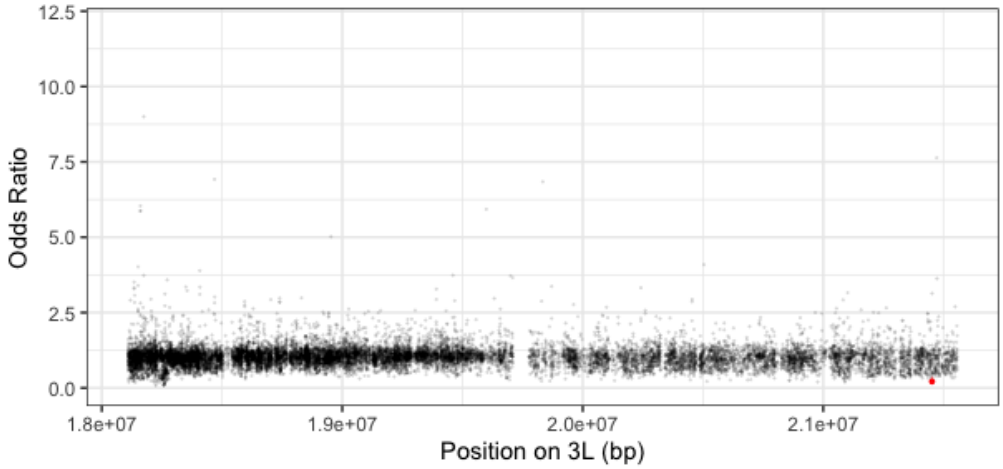

Figure S32: S1 modelled odds ratios by position on chromosome 3L within the elevated region on E1. Positions with a p-value < 0.0001 highlighted red with upper and lower 95% confidence intervals shown.

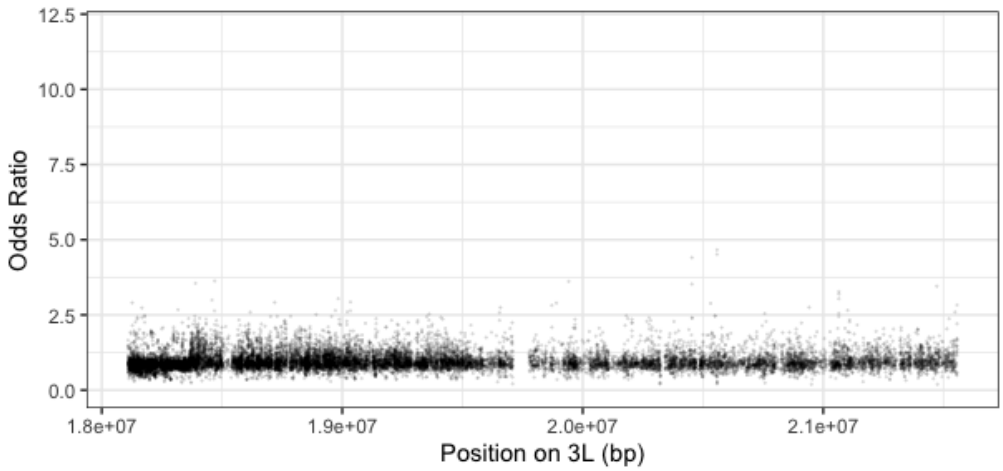

Figure S33: S2 modelled odds ratios by position on chromosome 3L within the elevated region on E1. Positions with a p-value < 0.0001 highlighted red with upper and lower 95% confidence intervals shown.

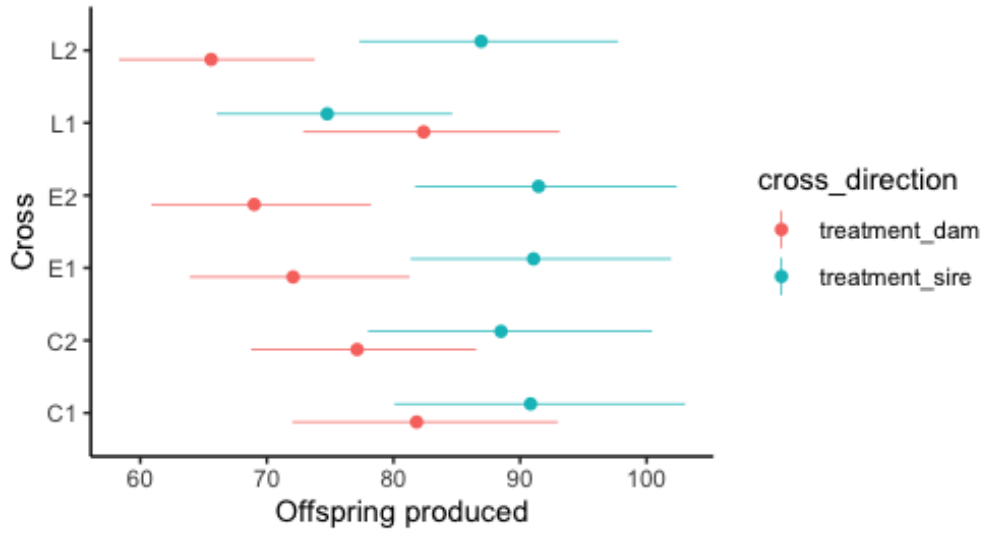

Figure S34: Fecundity measure for F1 of our sex ratio cross based on F0 treatment

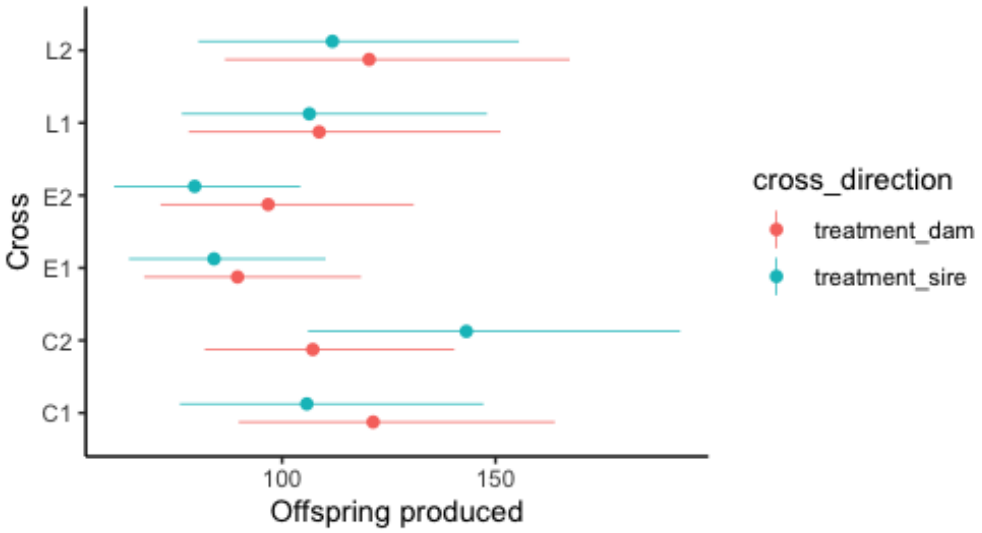

Figure S35: Fecundity measure for F2 of our sex ratio cross based on F0 treatment

Table S6: Discordant -Top 20 biological process GO terms

GO.ID	Term	Annotated	Significant	Expected	classicFisher
GO:0050916	sensory perception of sweet taste	14	5	0.48	7.20E-05
GO:0048749	compound eye development	377	24	12.95	0.00017
GO:0006814	sodium ion transport	66	11	2.27	0.0002
GO:0007424	open tracheal system development	241	19	8.28	0.00029
GO:0050912	detection of chemical stimulus involved in sensory perception of taste	23	5	0.79	0.00094
GO:0007409	axonogenesis	337	24	11.58	0.00154
GO:0045572	positive regulation of imaginal disc growth	16	4	0.55	0.0018
GO:0007615	anesthesia-resistant memory	17	4	0.58	0.00229
GO:0051124	synaptic growth at neuromuscular junction	139	10	4.77	0.00277
GO:0007614	short-term memory	29	5	1	0.0028
GO:0036099	female germ-line stem cell population ma...	29	5	1	0.0028
GO:0099024	plasma membrane invagination	26	4	0.89	0.00289
GO:0007455	eye-antennal disc morphogenesis	42	6	1.44	0.00291
GO:0007584	response to nutrient	19	4	0.65	0.00353
GO:0070839	metal ion export	10	3	0.34	0.00403
GO:0045187	regulation of circadian sleep/wake cycle, sleep	36	6	1.24	0.00424
GO:0019367	fatty acid elongation, saturated fatty acid	20	4	0.69	0.00429
GO:0034625	fatty acid elongation, monounsaturated fatty acid	20	4	0.69	0.00429

GO:0034626	fatty acid elongation, polyunsaturated fatty acid	20	4	0.69	0.00429
GO:0035158	regulation of tube diameter, open tracheal system	11	3	0.38	0.0054

Table S7: Large - Top 20 biological process GO terms

GO.ID	Term	Annotated	Significant	Expected	classicFisher
GO:0048512	circadian behavior	110	8	3.18	0.00082
GO:0035209	pupal development	17	4	0.49	0.00121
GO:0051124	synaptic growth at neuromuscular junction	139	12	4.02	0.00125
GO:0016319	mushroom body development	74	8	2.14	0.00129
GO:0007616	long-term memory	75	8	2.17	0.00141
GO:0010628	positive regulation of gene expression	191	14	5.52	0.00227
GO:0006821	chloride transport	29	4	0.84	0.00433
GO:0070828	heterochromatin organization	59	5	1.71	0.00478
GO:0060070	canonical Wnt signaling pathway	93	7	2.69	0.00545
GO:0007297	ovarian follicle cell migration	138	8	3.99	0.00546
GO:0017156	calcium-ion regulated exocytosis	13	3	0.38	0.00553
GO:0007474	imaginal disc-derived wing vein specification	57	6	1.65	0.00587
GO:0007619	courtship behavior	120	7	3.47	0.00683
GO:0045944	positive regulation of transcription by RNA polymerase II	331	18	9.57	0.00773
GO:0030177	positive regulation of Wnt signaling pathway	43	5	1.24	0.00779
GO:0035249	synaptic transmission, glutamatergic	26	4	0.75	0.00838
GO:0008586	imaginal disc-derived wing vein morphogenesis	44	5	1.27	0.00846
GO:0016311	dephosphorylation	196	9	5.67	0.0085
GO:0051260	protein homooligomerization	29	4	0.84	0.00922
GO:0031122	cytoplasmic microtubule organization	29	4	0.84	0.00922

Table S8: Small Top 20 biological process GO terms

GO.ID	Term	Annotated	Significant	Expected	classicFisher
GO:0048512	circadian behavior	110	8	3.37	0.00092
GO:0007615	anesthesia-resistant memory	17	4	0.52	0.0015
GO:0035209	pupal development	17	4	0.52	0.0015
GO:0016319	mushroom body development	74	8	2.27	0.00185
GO:0016226	iron-sulfur cluster assembly	20	4	0.61	0.00284
GO:0048675	axon extension	54	7	1.65	0.004
GO:0006821	chloride transport	29	4	0.89	0.00508

GO:1902808	positive regulation of cell cycle G1/S phase transition	12	3	0.37	0.0051
GO:0070828	heterochromatin organization	59	6	1.81	0.00532
GO:0060070	canonical Wnt signaling pathway	93	7	2.85	0.0064
GO:0045455	ecdysteroid metabolic process	40	5	1.22	0.00642
GO:0017156	calcium-ion regulated exocytosis	13	3	0.4	0.00648
GO:0008045	motor neuron axon guidance	75	7	2.3	0.00793
GO:0007616	long-term memory	75	7	2.3	0.00793
GO:0007619	courtship behavior	120	7	3.67	0.00801
GO:0030177	positive regulation of Wnt signaling pathway	43	5	1.32	0.00871
GO:0030154	cell differentiation	1703	81	52.14	0.00937
GO:0016311	dephosphorylation	196	9	6	0.01075
GO:0051124	synaptic growth at neuromuscular junction	139	11	4.26	0.01082
GO:0008016	regulation of heart contraction	29	4	0.89	0.01121

Table S9: discordant - Top 20 molecular function GO terms

GO.ID	**Term**	**Annotated**	**Significant**	**Expected**	**classicFisher**
GO:0015491	cation:cation antiporter activity	14	4	0.48	0.001
GO:0005319	lipid transporter activity	46	6	1.58	0.0032
GO:0033041	sweet taste receptor activity	30	5	1.03	0.0033
GO:0008252	nucleotidase activity	10	3	0.34	0.004
GO:0022848	acetylcholine-gated monoatomic cation-selective channel activity	10	3	0.34	0.004
GO:0009922	fatty acid elongase activity	20	4	0.69	0.0043
GO:0015141	succinate transmembrane transporter activity	11	3	0.38	0.0054
GO:0008227	G protein-coupled amine receptor activity	22	4	0.75	0.0061
GO:0004596	peptide alpha-N-acetyltransferase activiy	13	3	0.45	0.0089
GO:0046872	metal ion binding	1047	44	35.93	0.0107
GO:0070566	adenylyltransferase activity	14	3	0.48	0.011
GO:0016791	phosphatase activity	180	15	6.18	0.0119
GO:0005201	extracellular matrix structural constituent	30	4	1.03	0.0185
GO:0008201	heparin binding	17	3	0.58	0.0191
GO:0008270	zinc ion binding	518	27	17.78	0.0201
GO:0000976	transcription cis-regulatory region binding	296	21	10.16	0.0249
GO:0003777	microtubule motor activity	65	7	2.23	0.0292
GO:0042826	histone deacetylase binding	20	3	0.69	0.0296

GO:0000977	RNA polymerase II transcription regulatory region sequence-specific DNA binding	221	15	7.58	0.0364
GO:0000166	nucleotide binding	959	41	32.91	0.0365

Table S10: Large - Top 20 molecular function GO terms

GO.ID	Term	Annotated	Significant	Expected	classicFisher
GO:0004035	alkaline phosphatase activity	14	5	0.4	2.90E-05
GO:0000977	RNA polymerase II transcription regulatory region sequence-specific DNA binding	221	17	6.29	0.00034
GO:0022839	ion gated channel activity	13	3	0.37	0.00529
GO:0043565	sequence-specific DNA binding	492	30	14.01	0.0092
GO:0001056	RNA polymerase III activity	16	3	0.46	0.00973
GO:0005249	voltage-gated potassium channel activity	17	3	0.48	0.01157
GO:0017147	Wnt-protein binding	18	3	0.51	0.0136
GO:0000981	DNA-binding transcription factor activity, RNA polymerase II-specific	259	18	7.38	0.01481
GO:0001227	DNA-binding transcription repressor activity, RNA polymerase II-specific	52	5	1.48	0.01582
GO:0003714	transcription corepressor activity	54	5	1.54	0.01839
GO:0016798	hydrolase activity, acting on glycosyl bonds	98	4	2.79	0.02028
GO:0005516	calmodulin binding	37	4	1.05	0.02037
GO:0015171	amino acid transmembrane transporter activity	57	5	1.62	0.02274
GO:0016462	pyrophosphatase activity	493	21	14.04	0.02558
GO:0016831	carboxy-lyase activity	23	3	0.66	0.02659
GO:0004683	calmodulin-dependent protein kinase activity	10	2	0.28	0.03129
GO:0043015	gamma-tubulin binding	10	2	0.28	0.03129
GO:0004112	cyclic-nucleotide phosphodiesterase activity	10	2	0.28	0.03129
GO:0005254	chloride channel activity	26	3	0.74	0.03668
GO:0004860	protein kinase inhibitor activity	11	2	0.31	0.03753

Table S11: Small - Top 20 molecular function GO terms

GO.ID	Term	Annotated	Significant	Expected	classicFisher
GO:0004035	alkaline phosphatase activity	14	5	0.43	4.10E-05
GO:0043565	sequence-specific DNA binding	492	31	15	0.0022
GO:0000977	RNA polymerase II transcription regulatory region sequence-specific DNA binding	221	15	6.74	0.0023
GO:0004683	calmodulin-dependent protein kinase activity	10	3	0.3	0.0029

GO:0015171	amino acid transmembrane transporter activity	57	6	1.74	0.0039
GO:0005516	calmodulin binding	37	5	1.13	0.005
GO:0022839	ion gated channel activity	13	3	0.4	0.0064
GO:0000981	DNA-binding transcription factor activity, RNA polymerase II-specific	259	19	7.9	0.0082
GO:0001056	RNA polymerase III activity	16	3	0.49	0.0117
GO:0005249	voltage-gated potassium channel activity	17	3	0.52	0.0139
GO:0017147	Wnt-protein binding	18	3	0.55	0.0163
GO:0005525	GTP binding	155	10	4.73	0.0203
GO:0001227	DNA-binding transcription repressor activity, RNA polymerase II-specific	52	5	1.59	0.0206
GO:0016798	hydrolase activity, acting on glycosyl bonds	98	5	2.99	0.0229
GO:0003714	transcription corepressor activity	54	5	1.65	0.0239
GO:0016614	oxidoreductase activity, acting on CH-OH group of donors	121	4	3.69	0.0252
GO:0016462	pyrophosphatase activity	493	24	15.03	0.0301
GO:0016831	carboxy-lyase activity	23	3	0.7	0.0317
GO:0003924	GTPase activity	145	9	4.42	0.0333
GO:0043015	gamma-tubulin binding	10	2	0.3	0.0355