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:

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, potentially conflicting fitness optima, resulting in intralocus sexual conflict, as each sex bears a genetic 'load' of the opposite sex. To understand conflict in the genome, one strategy is to artificially select a population discordantly, against established sexual dimorphism, to reintroduce an attenuated conflict. We investigate a long-term artificial selection experiment where sexual size dimorphism has been reversed in *Drosophila melanogaster* after ~350 generations of sexually discordant size selection. We explore morphological and genomic changes to identify loci between the sexes in discordantly and concordantly selected treatments. While concordant selection maintains ancestral sexual dimorphism, discordant selection responds 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 among the sexes within each lineage identified potential sites maintained by sexual conflict. One discordant selection lineage shows a pattern of elevated genomic differentiation (F_{ST}) on chromosome 3L, consistent with the presence of maintained sexual conflict in the genome. Our results demonstrate the polygenicity of body size, as well as sexual conflict. Furthermore, our results also suggest visible 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 often have a high genetic correlation (r_{MF}) which can hinder their independent evolution (Lande 1980). The extent to which SD can evolve depends on the strength and direction of selection, 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 that favours one sex, but not 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. 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. Despite both discordant selection responding rapidly in family-based selection studies, as well as r_{MF} rapidly responding to direct selection, some previous small studies have found little response to sex-discordant selection (Tigreros and Lewis 2011). This suggests at least some barrier to divergent response in sexual dimorphism may exist in 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 a measurable phenotypic response requiring more than 100-150 generations of artificial selection. In comparison, sex-concordant selection for body size resulted in rapid phenotypic responses. These sex-concordant (hereafter, concordant) selection lineages established by Stewart & Rice have also 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 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 *diminutive* (*myc*) 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.

Populations of *D. melanogaster* harbour alleles with both sexually discordant and concordant effects on fitness (Rice and Chippindale 2001). This suggests that highly polygenic traits likely have standing genetic variation maintained by selection, or that can be selected on in a sexually antagonistic way. Variation for body size in *D. melanogaster* is highly polygenic (Carreira et al. 2009; Turner et al. 2011). Sex limited selection on males has previously been shown to incur a fitness cost to females (Prasad et al. 2007). This male limited evolution also results in a change in body size in non-selected females that is closer to the male optimum (Prasad et al. 2007). A decrease in female fitness (and body size) when males are allowed to evolve toward their own optimum without parallel female evolution suggests that there is unresolved sexual conflict in the genome of *D. melanogaster* pertaining to body size. Alleles under IASC influencing body size, can elucidate questions pertaining to how the genome responds to divergent phenotypic optimums. Selection on body size in a discordant manner has the potential to answer questions of how the shared genome overcomes the barrier of r_{MF} when selection favors opposing phenotypes.

Using lineages evolved under sexually discordant selection for size, first described in Stewart and Rice (2018), we demonstrate two important findings. First, despite selection for a trait-agnostic measure of size, trait specific patterns of SSD reversal are the norm. Second, despite the long term sexually discordant selection, we see relatively weak evidence for sexually antagonistic alleles being maintained. We do, however, identify one region potentially segregating differentially between the sexes, a possible sign of unresolved (or re-introduced) sexual conflict. Using an evolve and sequence approach we not only examined among lineage patterns of genomic differentiation, but within lineage among the sexes. We discuss our findings within the context of both the evolution of sexual dimorphism and the potential role of ongoing intra-locus sexual conflict.

Materials & Methods

Lineages

The flies used are part of a long-term experimental population for size selection (Turner et al. 2011; Stewart and Rice 2018). The selection lineages examined were started using the

outbred population, LH_M, previously adapted to the lab for over 350 generations. These flies are maintained in discrete two-week generations at moderate density (approximate 200 eggs per 10ml standard molasses food vial (Table S1). The complete methodology for maintenance of the base population has been published previously (Rice et al. 2005). For the size selection, flies are anesthetized using CO_2 and sorted using a motorized stacked sieving device in which each successive sieve is 5% smaller than the sieve above. The largest sieve used had aperture diameters of 2000μm and the smallest sieve had apertures measuring 850μm. In each generation, all flies (~1800 individuals), from a selection lineage were sorted by sieve, and 10 vials of 16 mating pairs (320 total flies per lineage) were selected based on phenotype, to populate the next generation. One treatment used only the smallest flies of both sexes (S; concordant selection), one treatment the largest of each sex (L; concordant selection), and the reversal of dimorphism (discordant selection) treatment used smallest females and largest males (E). Finally, a control treatment was populated with flies that passed through the sieves, but were not selected based on size, as a control (C). Two replicate lineages were maintained for each selection treatment. During subsequent selection for size with the established protocol, it was observed that selection can be more stringent on small flies than large because of the nature of the sieve sorting. This would imply that in the sex-discordant selection females being selected to be small were under increased selective pressure compared to males selected to be large. This would suggest that there was a greater pressure for selection in one direction and although this does not explain all results, it may influence response to sexually discordant selection.

Morphological measures and analysis

At generation 367 (August 19th, 2019) flies from each lineage were collected and stored in 70% ethanol for dissection and measurement of traits. Individuals were chosen randomly before and after the selection treatment to get an accurate estimate of size in the overall lineage. Individuals were dissected under a Leica M125 microscope and legs and thoraces were imaged at 63x using a DCF-400 camera. A minimum of 20 flies (for each sex) were dissected from each lineage by dissecting off the first (pro-thoracic) right leg, then imaging the left side of the thorax of each individual. The leg and left side of the thorax were mounted on slides (in 70% glycerol in PBS, with a small amount of phenol as a bacteriostatic agent). The femur, tibia, and first tarsal segment of each leg, and each thorax, was measured using ImageJ version 1.52q (Schneider et al. 2012). We measured thorax length as a proxy for overall size and a proxy for the general pattern of the female-biased SSD found in *Drosophila melanogaster*. We also measured each leg segment; measurement was completed from the center of the beginning of the segment to the center of the end of the segment.

Analysis of the leg and thorax measurements were completed using R version 4.1.3 (R Core Team 2022). One individual measure of femur was removed as it's measure was ~100X smaller than the mean for the trait. For all analyses of morphology, we used $\log_2$ transformed trait values. We fit general linear mixed models for individual traits including, sex, selection and sampling and their 2nd order interactions as fixed effects, allowing sex effects to vary as a random effect of replicate lineage (i.e. random slopes). Models were fit using lmer in the lme4 package version 1.1.30 (Bates et al. 2015). We confirmed the results with a fully multivariate

mixed model for all traits. For the multivariate mixed model, we allowed for random effects of sex and traits by lineage, as well as accounting for within-individual variation across traits. The initial fit of this model using lmer was singular, likely due to variance estimates getting "stuck" at the boundary (0). We employed two approaches to deal with this. First, we employed a Bayesian extension of our model using blmer in blme v1.0.5 (Chung et al. 2013). This employs weak regularizing priors (away from 0 for the variances). We confirmed the stability of the fixed effects using a second approach, fitting a general latent variable mixed model, to estimate reduced-rank covariance matrices for random effects, as implemented in glmmTMB v1.1.7 (Brooks et al. 2017; Niku et al. 2019; Kristensen and McGillycuddy 2023). These approaches provided very similar fixed effects estimates to each other, and to the single trait models, which were used for downstream analysis.

Despite 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.

Statistics were conducted in R version 4.1.3 (R core team 2022). Estimated marginal means, custom contrasts, and associated confidence intervals were estimated using emmeans version 1.8.0 (Lenth et al. 2018). Visualization was done using ggplot2 version 3.3.6 (Wickham 2018).

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.

Bioinformatics and analyses

Figure 2 shows a summary of the bioinformatics pipeline, a more specific summary can be found at (https://github.com/TylerAudet/SSD_workflow). Reads were trimmed using bbduk version 38.90 (Bushnell 2021), aligned using BWA-mem version 0.7.8 (Li 2013), GATK version 3.8 was used to mark and realign around indels, and SAMtools version 1.5 (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 the sexes and replicates separate, one with sexes merged maintaining replicates, and one where treatments were pooled together by merging both sexes and replicates together, each of these three went through the same SNP calling and file generation methods. Popoolation version 1.2.2 (Kofler et al. 2011) was used to mask repetitive regions identified with Repeat Masker version 4 (Smit and Hubley 2013), as well as indels for SNP calling. PoolSNP version 1 (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 pool-seq 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, both F_{ST} and nucleotide diversity (π) were calculated using Grenedalf v0.2.0 (Czech et al. 2023). To identify variants that potentially contribute to phenotypic divergence between treatments, we estimated F_{ST} between control treatments and SSD-reversal treatments. As our primary goal was on the changes to SSD, not overall size. We also examined F_{ST} between Large and Small treatments to follow up on Turner et al. (2011) and see if their findings were replicated after an additional ~275 generations of artificial selection on body size. To identify sites where genetic differentiation is most likely to be explained by selection rather than drift, a Cochran–Mantel–Haenszel (CMH) statistic was calculated using ACER version 1.0 (Spitzer et al. 2020) for each treatment and replicate. Since an ancestral sample was not preserved, drift between control and discordant treatments were used and the generations of drift was set as 700 generations from the control treatment (~generations to convergence x2). Looking for regions of low diversity may suggest selection on a particular SNP or nearby a particular SNP. The regions with the lowest 10% of π (Control cut-off was $\pi < 0.091$, discordant cut-off was $\pi < 0.01$, large cut-off was $\pi < 0.005$, small cut-ff was $\pi < 0.01$, in all treatments this resulted in 268 windows selected) values were extracted, and the overlap with the top 5% of F_{ST} (discordant) vs. control cut-off was F_{ST} 0.51 with 621 windows meeting this criteria; large vs. small cut-off was F_{ST} 0.78 with 621 windows meeting that criteria) and SNPs with a CMH p-value < 0.01 (132290 SNPs in control vs. discordant, 115747 SNPs in large vs. small) were extracted and annotated for both our discordant treatment compared to all concordant treatments as well as the large versus small comparison. The CMH test returns a list of alleles with frequencies that are statistically different in one population compared to another, and these p-values were adjusted with a Benjamini & Hochberg correction using *p.adjust()* in R (Benjamini and Hochberg 1995). Four files were created with regions of interest: 1) "discordant replicate 1 region" All sites in and elevated region of F_{ST} in replicate 1 of our discordant treatment on chromosome 3L (explained below; supplemental file 1), 2) "discordant sites" all regions with the top 95% of F_{ST} between

discordant and control as well as the bottom 10% of π values for the pooled discordant treatments, and all windows of CMH with a p-value below 0.01 between in discordant when control is used as the ancestral genome (supplemental file 2) 3) "Large sites" all regions with the top 95% of F_{ST} between large and small as well as the bottom 10% of π values for the pooled large treatments, and all windows of CMH with a p-value below 0.01 between large and small (supplemental file 3) 4) "Small sites" all regions with the top 95% of F_{ST} between large and small (same file as "large sites") as well as the bottom 10% of π values for the pooled small treatments, and all windows of CMH with a p-value below 0.01 between large and small (same file as "large sites; supplemental file 4). These sites were extracted from the SNP called VCF file using bedtools version 2.30.0 (Quinlan and Hall 2010), and 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 then extracted the top 20 GO terms for each subset and further filtered our sites of interest by those that are present in the top 20 GO terms. Gene annotations were conducted using SNPeff version 5.1 (Cingolani et al. 2012). 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.

Results

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

Consistent with the previously published results for overall "body size" (Stewart and Rice 2018), we observed responses to artificial selection for all measured traits. Thorax was measured to estimate body size and had a significant relationship between size and selection lineage ($\chi^2 = 242.72$ $p < 0.0001$; Figure 3; Table S2; Figure S1). Concordant artificial selection for decreased size reduced thorax size relative to control lineages by ~25% in females (26% in males), while selection for increased size increased thorax size by ~9.5% for both sexes. In the discordant selection lineages, we found a ~10% decrease in female thorax length, and minimal (~0%) change in males. We generally observed a similar pattern for leg segments. Small treatment males and females decreased in length relative to controls, to a similar degree (~30% for femur, ~24% for tibia and tarsus; Figure 4). Large treatment males and females increased in size somewhat more modestly relative to controls (~2% for femur, 6% for tibia, 8% for tarsus). For sexually discordant selection, females decreased in size relative to controls (~14% for femur, ~7% tibia and tarsus). Males from the discordant lineage decreased more modestly relative to controls (~8% for femur, < 1% for tibia, ~3.5% tarsus). While some trait changes have confidence intervals that overlap zero, all traits and treatments seemed to respond in size in the expected direction of effect (Figure 3;4; Table S3; S4; S5).

Sexual dimorphism and multivariate allometry diverged only under discordant, sex-specific selection

Despite substantial changes in overall size, SSD remained female biased in all concordant selection treatment groups (Figure 3; 4), showing little change in dimorphism relative to control (Figure 4). In contrast, the sexually discordant selection treatment, resulted in a substantial reduction in the amount of ancestral female biased dimorphism, relative to the controls (Figure 4). Specifically, thorax and tarsus lengths have evolved to be essentially monomorphic, while femur and tibia lengths are now male-biased (~5% increase) in size. The change in the amount of dimorphism was significant for all traits but tarsus length (Figure 4).

Given the results described above, it is not surprising we observed substantial changes in patterns of multivariate allometry (Figure 5; Figure S3). We observe a substantial change in the sex-discordant lineages, which deviates from the isometry vector (Table 1) that is observed for other *Drosophila melanogaster* populations (Shingleton et al. 2009). We further compared covariance matrices across selective treatments using the Krzanowski correlation. The discordant selection lineages show a reduced correlation to the control lineages (r_{Krz} = 0.86) compared with the patterns observed for both sexually-concordant selection treatments (r_{Krz} = 0.97 and $r_{Krz} = 0.99$ for large and small respectively).

Genomic differentiation and nucleotide diversity (π) between selection treatments

F_{ST} between control and discordant selection treatments was highly variable, but included substantial differentiation across the genome, and many regions at, or near, fixation (Figure 6A). Notably, average F_{ST} is quite high across the chromosomes, consistent with a substantial impact of genetic drift on allele frequencies. When F_{ST} is calculated between discordant and either of the concordant (small or large) treatments, we similarly see elevated F_{ST} and multiple regions at, or near fixation (Figure S3; S4). In our discordant selection versus controls, the highest mean F_{ST} is found on chromosome 2L (mean F_{ST} per chromosome, 2L = 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 6B). 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 the lowest 10% of π windows (windows size = 50000bp) as well as SNPs with a CMH p-value below 0.01. All these SNPs were then annotated (Supplemental files 1;2;3;4). After manual curation, several genes with sex-specific size effects were identified such as: dMyc (*myc*), Hairless (*H*), Insulin-like receptor (*InR*), Regulator of cyclin A1 (*Rca1*), and stunted (*sun*). Many other genes with size phenotypes were also observed for both the discordant and concordant comparisons (supplemental files 1;2;3;4).

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).

Intersexual genetic differentiation in one sexually discordant lineage may suggest the maintenance of intralocus 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 7A;S5; L1 = 0.0012, L2 = 0.0015 Figure S7;S8; S1 = 0.0023, S2 = 0.0014, Figure S9;S10). For the discordant selection treatment replicate 2 we also find a mean chromosomal F_{ST} to be near zero (E2 = 0.0011, Figure S6). 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 on chromosome 3L, where the mean F_{ST} for the whole chromosome arm rises to 0.005, with elevated SNPs showing a distinct peak of F_{ST} nearing 0.1 chromosome 3L, with a couple windows reaching 0.25 (Figure 7B). 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).

Discussion

Sexual dimorphism (SD) evolves frequently, despite r_{MF} generally being high within species, in particular for morphological traits (Lande 1980; Poissant et al. 2010). In the presence of sex-specific optimal phenotypes, this r_{MF} has the potential to generate intra-locus sexual conflict through the 'load' on the opposite sex (Fairbairn et al. 2007). Species that evolve changes in SD must overcome this r_{MF} hurdle and ensuing genomic conflict. In the long-term, this conflict may reach an equilibrium, with sex-biased alleles creating as near an optimum phenotype for each sex as possible. However, the reintroduction of sex-discordant selection should disrupt this equilibrium and create tension in the genome. Looking at the response to

discordant selection in the genome after selection provides an opportunity to identify active conflict in the genome and look for the signature of selection in a genome under conflict. In this study we utilize long-term artificially selected lineages (Stewart and Rice 2018), selected for body size either in sexually concordant or discordant manner. As discussed in detail below, in addition to observing trait specific changes in patterns of sexual size dimorphism, we see potential evidence for the maintenance of polymorphisms consistent with unresolved conflict.

Concordant selection lineages maintained ancestral patterns of SSD despite selection on size, while discordant selection lineages responded in a trait specific manner

The lineages under concordant artificial selection responded in the direction that we expected, for all traits, and changed size proportionally in a way that maintained SD for all traits. This proportional response is strongly suggestive of a 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 would expect to see response in a sex-specific manner, 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, we see both sexes within a treatment maintaining SSD despite strong selection on size. For the discordant artificial selection treatment, SD changed in a trait specific manner (Figure 3). Of the four traits measured for this study, one (thorax length) showed a relatively substantial degree of female biased dimorphism in the controls (difference of ~0.16, in $\log_2$ μm, or $\frac{\bar{F}}{\bar{M}} \approx 1.12$), while all the leg traits (femur, tibia and basi-tarsus lengths) were much 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 4). 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 the 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 5). 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), 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 would 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 provide 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, to identify possible SNPs showing subtle, within-generation distortions

in allele frequency. We hypothesized that there may be small regions or possibly single SNPs with elevated F_{ST}, which may be evidence of conflict in the genome. Neither control lineages, nor the sexually concordant lineages showed any substantial evidence for elevation in F_{ST} (Figure 7A;S5;S6;S7;S8;S9;S10. However, we observed one discordant replicate lineage (E1; Figure 7B), that had a reasonably large 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). However, modifying our SNP calling methods, F_{ST} calculation software, and filtering methods all yielded the same spike in F_{ST} in the E1 lineage. Additionally, it is only this lineage which shows this pattern. If it was a spurious peak of elevated F_{ST} as a result of technical aspect of the pipeline, other replicate lineages would likely also show elevated regions. We consider the most likely explanation to be due to the sustained nature of sexually discordant selection in our experiment. Compared with most other studies examining M-F distortions in allele frequencies, to identify potential loci under conflict, the selection applied to this lineage is both strong and sustained across hundreds of generations. Kasimatis et al. (2019) suggests that for a locus to show this degree of asymmetry it would require an exceedingly high strength of selection. In their model, this would require near 40% mortality every single generation to maintain the distortion in allele frequencies at the locus. However, the nature of artificial selection does mean well over 40% (~80% of individuals with the current selection design) of the flies from each generation (those not at the extremes of phenotype) are discarded, essentially 'dead' in the selective sense. Alternatively, Kasimatis et al. also simulated subsamples of 50 males and 50 females with antagonistic selection and found that sampling substantially increased variability in F_{ST} values due to the fact that genomes under antagonistic selection tend to have increased numbers of intermediate frequency alleles. Whether this region in discordant treatment replicate E1 is truly an area of divergent alleles requires further exploration.

Given the strange peak in this one discordant replicate, we extracted all SNPs three standard deviations above the mean F_{ST} on the 3L chromosome. We filtered for SNPs that appear in both discordant replicates and do not occur in any concordant lineage to look for possible replicated mechanisms of divergence in phenotype. We found fewer candidate genes than in any other comparison (Supplemental file 1), but multiple genes present in the WNT pathway, including *frizzled2* (*fz2*). We also found candidate SNPs in the MAPK pathway. Interestingly, none of the male-female discordant treatments SNPs appear in the between treatment comparisons. Similar to other comparisons however, many candidate genes have mutational phenotypes listed on flybase as "abnormal size" so this could be suggestive again that response to selection is highly polygenic with alleles of small effect "inching" towards a phenotype rather than "sprinting".

Discordant size selection likely has a polygenic basis, involving a number of genes that influence body size in a sex-specific manner

Given the number of generations of artificial selection, and the modest population sizes each generation, genetic drift will have an overwhelming impact on any genome scan performed with the evolved lineages. As such, any interpretation of these genome scans should

be tempered by this impact. With that being said, discordant selection treatments were compared to concordant treatments, and the response is broadly consistent with a polygenic response, with thousands of SNPs appearing between treatments and F_{ST} being high across the genome. In 378 generations of selection and more than 125 generations after SSD-reversal, the opportunity for an allele of large effect to occur was possible. However, we find no clear evidence for alleles of large effect that have swept to fixation in both replicate lineages of discordant selection, but the chance of identifying alleles of large effect is attenuated by the impact of drift making many SNPs fix randomly in each lineage. Possibly constrained by genetic correlation, the response to discordant selection appears to be a slow and steady crawl towards a response rather than large leaps forward. The lineage specific increase in discordant replicate 1 that does not appear in replicate 2 could imply a novel mutation that is under conflict in this replicate. This elevated F_{ST} is shown in a large region of chromosome 3L that includes hundreds of SNPs however, so to identify a novel mutation responsible requires further investigation. This slow crawl towards a response is supported by the fact that Stewart & Rice (2018) saw a negligible phenotypic response to discordant selection during at least the first 100 generations, with the more substantial responses occurring after this. Alleles of large effect in one sex may be more likely to be in pathways tightly shared between the sexes

such as insulin signalling, and therefore many alleles of small effect may be better able to gently modulate this genomic tug-of-war, without major fitness costs. Large effect alleles with sex-specific impacts on size have been identified in functional studies, however the viability cost in the opposite sex may make these large effect alleles prohibitive (Rideout et al. 2015; Millington et al. 2021a). Although it is tempting to speculate that these small effect alleles are a likely way for SSD to evolve in natural populations, we must also caveat that it may be the case that these alleles of such small effect may be weeded out by viability or fecundity selection in the wild, which are much weaker in this tightly controlled artificial selection experiment. In fact, Testa and Dworkin (2016) demonstrated that mutations with a sex-specific response are rare in two distinct genetic backgrounds. Carreira et al. (2009) however, found that many genes associated with size change reduced SD when a P-element was inserted. This highlights the complexity of the genetic architecture of SSD.

Although Testa and Dworkin (2016) found that mutations with a sex-specific effect on size were rare, they did identify mutations in the EGFR pathway with sex-specific effects. In our sex-discordant lineages we found a number of genes with known sex-specific size effect (d*myc, H, InR, RcA1, sun*). Follow up on these genes to determine the role they play in response to discordant selection is required. Interestingly we also found a number of enriched GO terms 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.

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. Although our candidate genes differ from those listed by Turner et al., 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. 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 very 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).

Literature Cited

Alexa, A., J. Rahnenfuhrer, and T. Lengauer. 2006. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. Bioinformatics 22(13):1600–1607.

Alicchio, R., and D. L. Palenzona. 1971. Changes of sexual dimorphism values in *Drosophila Melanogaster*. Bolletino Zool. 38:75–84.

Ashburner, M. 1989. Drosophila. A laboratory handbook. Drosoph. Lab. Handb. Cold Spring Harbor Laboratory Press.

Bates, D., M. Machler, B. Bolker, and S. Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. J. Stat. Softw. 67.

Benjamini, Y., and Y. Hochberg. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. R. Stat. Soc. Ser. B Methodol. 57:289–300. [Royal Statistical Society, Wiley].

Bird, M. A., and H. E. Schaffer. 1972. A Study of the Genetic Basis of the Sexual Dimorphism for Wing Length in *Drosophila melanogaster*. Genetics 72:475–487.

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

Brooks, M., E., K. Kristensen, K. Benthem J. ,van, A. Magnusson, C. Berg W., A. Nielsen, H. Skaug J., M. Mächler, and B. Bolker M. 2017. glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. R J. 9:378.

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

Carreira, V. P., J. Mensch, and J. J. Fanara. 2009. Body size in Drosophila: genetic architecture, allometries and sexual dimorphism. Heredity 102:246–256. Nature Publishing Group.

Cheng, C., and M. Kirkpatrick. 2016. Sex-Specific Selection and Sex-Biased Gene Expression in Humans and Flies. PLOS Genet. 12:e1006170.

Chung, Y., S. Rabe-Hesketh, V. Dorie, A. Gelman, and J. Liu. 2013. A Nondegenerate Penalized Likelihood Estimator for Variance Parameters in Multilevel Models. Psychometrika 78:685–709.

Cingolani, P., A. Platts, L. L. Wang, M. Coon, T. Nguyen, L. Wang, S. J. Land, X. Lu, and D. M. Ruden. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. Fly (Austin) 6:80–92.

Cowley, D. E., and W. R. Atchley. 1988. Quantitative Genetics of Drosophila Melanogaster. II. Heritabilities and Genetic Correlations between Sexes for Head and Thorax Traits. Genetics 119:421–433.

Czech, L., J. P. Spence, and M. Expósito-Alonso. 2023. grenedalf: population genetic statistics for the next generation of pool sequencing. arXiv preprint. arXiv:2306.11622

Delph, L. F., J. C. Steven, I. A. Anderson, C. R. Herlihy, and E. D. Brodie III. 2011. Elimination of a Genetic Correlation Between the Sexes Via Artificial Correlational Selection. Evolution 65:2872–2880.

Eisen, E., and J. Hanrahan. 1972. Selection for Sexual Dimorphism in Body Weight of Mice. Aust. J. Biol. Sci. 25:1015–24.

Emlen, D. J., J. Marangelo, B. Ball, and C. W. Cunningham. 2005. Diversity in the Weapons of Sexual Selection: Horn Evolution in the Beetle Genus Onthophagus (coleoptera: Scarabaeidae). Evolution 59:1060–1084.

Fabian, D. K., M. Kapun, V. Nolte, R. Kofler, P. S. Schmidt, C. Schlötterer, and T. Flatt. 2012. Genome-wide patterns of latitudinal differentiation among populations of Drosophila melanogaster from North America. Mol. Ecol. 21:4748–4769.

Fairbairn, D. J., W. U. Blanckenhorn, and T. Székely (eds). 2007. Sex, size, and gender roles: evolutionary studies of sexual size dimorphism. Oxford University Press, Oxford ; New York.

Honěk, A. 1993. Intraspecific Variation in Body Size and Fecundity in Insects: A General Relationship. Oikos 66:483–492.

Jolicoeur, P. 1963. 193. Note: The Multivariate Generalization of the Allometry Equation. Biometrics 19:497–499.

Kapun, M., M. G. Barrón, F. Staubach, D. J. Obbard, R. A. W. Wiberg, J. Vieira, …, J. González. 2020. Genomic Analysis of European Drosophila melanogaster Populations Reveals Longitudinal Structure, Continent-Wide Selection, and Previously Unknown DNA Viruses. Mol. Biol. Evol. 37:2661–2678.

Kasimatis, K. R., P. L. Ralph, and P. C. Phillips. 2019. Limits to Genomic Divergence Under Sexually Antagonistic Selection. G3 Genes, Genomes, Genetics 9:3813–3824.

Klingenberg, C. P. 1996. Multivariate Allometry. Pp. 23–49. Advances in Morphometrics. Springer US, Boston, MA.

Kofler, R., R. V. Pandey, and C. Schlotterer. 2011. PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). Bioinformatics 27:3435–3436.

Kristensen, K., and M. McGillycuddy. 2023. The AR(1) covariance structure. https://cran.rstudio.com/web/packages/glmmTMB/vignettes/covstruct.html

Lande, R. 1980. Sexual Dimorphism, Sexual Selection, and Adaptation in Polygenic Characters. Evolution 34:292–305.

Lenth, R., H. Singmann, J. Love, and P. Buerkner. 2018. emmeans: Estimated marginal means, aka Least-Squares Means. https://cran.r-project.org/web/packages/emmeans/emmeans.pdf

Li, H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv:1303.3997

Li, H., B. H, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, and R. Durbin. 2009. The Sequence Alignment/Map (SAM) Format and SAMtools. *bioinformatics*, *25*(16), 2078-2079.

Luecke, D. M., and A. Kopp. 2019. Sex-specific evolution of relative leg size in Drosophila prolongata results from changes in the intersegmental coordination of tissue growth. Evolution 73:2281–2294.

Mathews, K. W., M. Cavegn, and M. Zwicky. 2017. Sexual Dimorphism of Body Size Is Controlled by Dosage of the X-Chromosomal Gene Myc and by the Sex-Determining Gene tra in Drosophila. Genetics 205:1215–1228.

Millington, J. W., G. P. Brownrigg, C. Chao, Z. Sun, P. J. Basner-Collins, L. W. Wat, B. Hudry, I. Miguel-Aliaga, and E. J. Rideout. 2021a. Female-biased upregulation of insulin pathway activity mediates the sex difference in Drosophila body size plasticity. eLife 10:e58341.

Millington, J. W., L. W. Wat, Z. Sun, P. J. Basner-Collins, G. P. Brownrigg, and E. J. Rideout. 2021b. A low sugar diet enhances *Drosophila* body size in males and females via sex-specific mechanisms. *Development*, *149*(6), dev200491.

Moczek, A. P., T. E. Cruickshank, and A. Shelby. 2006. When Ontogeny Reveals What Phylogeny Hides: Gain and Loss of Horns During Development and Evolution of Horned Beetles. Evolution 60:2329–2341.

Niku, J., F. K. C. Hui, S. Taskinen, and D. I. Warton. 2019. gllvm: Fast analysis of multivariate abundance data with generalized linear latent variable models in r. Methods Ecol. Evol. 10:2173–2182.

Poissant, J., A. J. Wilson, and D. W. Coltman. 2010. Sex-Specific Genetic Variance and the Evolution of Sexual Dimorphism: A Systematic Review of Cross-Sex Genetic Correlations. Evolution 64:97–107.

Prasad, N. G., S. Bedhomme, T. Day, and A. K. Chippindale. 2007. An Evolutionary Cost of Separate Genders Revealed by Male-Limited Evolution. *The American Naturalist*, *169*(1), 29-37.

Quinlan, A. R., and I. M. Hall. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841–842.

Reeve, J. P., and D. J. Fairbairn. 1999. Change in sexual size dimorphism as a correlated response to selection on fecundity. Heredity 83:697–706.

Reeve, J. P., and D. J. Fairbairn. 1996. Sexual size dimorphism as a correlated response to selection on body size: an empirical test of the quantitative genetic model. Evolution 50:1927–1938.

Rice, W. R., and A. K. Chippindale. 2001. Intersexual ontogenetic conflict. J. Evol. Biol. 14:685–693.

Rice, W. R., J. E. Linder, U. Friberg, T. A. Lew, E. H. Morrow, and A. D. Stewart. 2005. Inter-locus antagonistic coevolution as an engine of speciation: Assessment with hemiclonal analysis. Proc. Natl. Acad. Sci. 102:6527–6534.

Rideout, E. J., M. S. Narsaiya, and S. S. Grewal. 2015. The Sex Determination Gene transformer Regulates Male-Female Differences in Drosophila Body Size. PLOS Genet. 11:e1005683.

Sawala, A., and A. P. Gould. 2017. The sex of specific neurons controls female body growth in Drosophila. PLOS Biol. 15:e2002252.

Schneider, C. A., W. S. Rasband, and K. W. Eliceiri. 2012. NIH Image to ImageJ: 25 years of image analysis. Nat. Methods 9:671–675.

Shingleton, A. W., C. M. Estep, M. V. Driscoll, and I. Dworkin. 2009. Many ways to be small: different environmental regulators of size generate distinct scaling relationships in Drosophila melanogaster. Proc. R. Soc. B Biol. Sci. 276:2625–2633.

Smit, AFA, Hubley, R. *RepeatModeler Open-1.0*. 2008-2015 <http://www.repeatmasker.org>.

Spitzer, K., M. Pelizzola, and A. Futschik. 2020. Modifying the Chi-square and the CMH test for population genetic inference: Adapting to overdispersion. The Annals of Applied Statistics 14:202–220. Institute of Mathematical Statistics.

Stewart, A. D., and W. R. Rice. 2018. Arrest of sex-specific adaptation during the evolution of sexual dimorphism in Drosophila. Nat. Ecol. Evol. 2:1507–1513.

Sylvestre, F., C. Mérot, E. Normandeau, and L. Bernatchez. 2023. Searching for intralocus sexual conflicts in the three-spined stickleback (*Gasterosteus aculeatus*) genome. Evolution 77:1667–1681.

Testa, N. D., and I. Dworkin. 2016. The sex-limited effects of mutations in the EGFR and TGF-β signaling pathways on shape and size sexual dimorphism and allometry in the Drosophila wing. Dev. Genes Evol. 226:159–171.

Tigreros, N., and S. M. Lewis. 2011. Direct and correlated responses to artificial selection on sexual size dimorphism in the flour beetle, Tribolium castaneum. J. Evol. Biol. 24:835–842.

Turner, T. L., A. D. Stewart, A. T. Fields, W. R. Rice, and A. M. Tarone. 2011. Population-Based Resequencing of Experimentally Evolved Populations Reveals the Genetic Basis of Body Size Variation in Drosophila melanogaster. PLoS Genet. 7:e1001336.

Wickham, H. 2018. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, New York.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

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: 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.

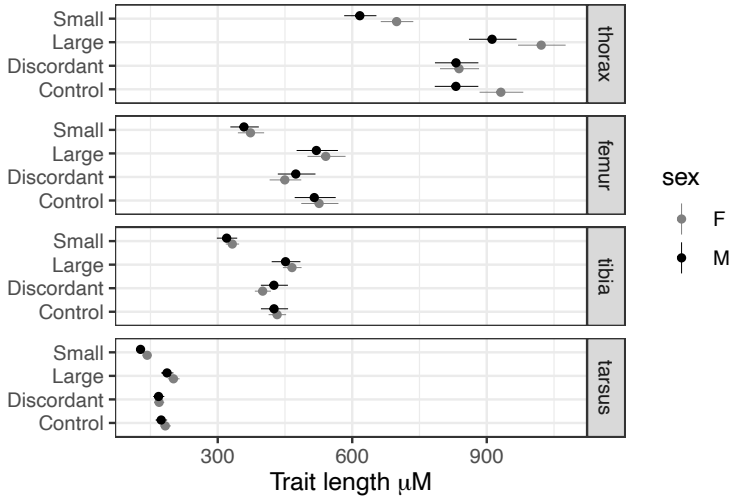

Figure 3: Model estimated means for each trait, across treatments and sex. Measurements were $\log_2$ transformed for model fit, and back-transformed only for plotting. Error bars represent 95% confidence intervals.

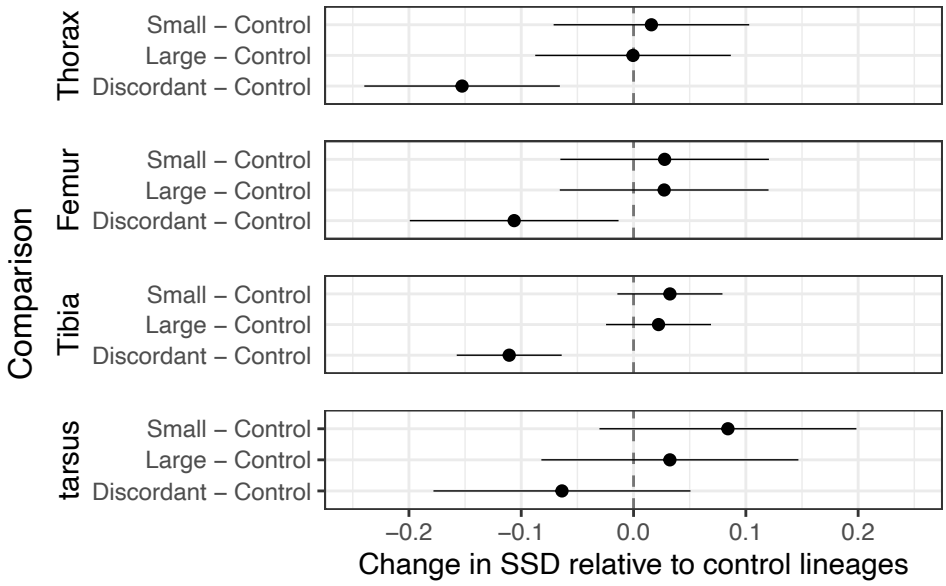

Figure 4: Substantial changes in sexual size dimorphism only occurs under discordant selection on size. Contrasts represent changes in the amount of sexual size dimorphism between each artificially selected treatment, in comparison to controls, by trait. Estimates using $\log_2$ transformed size measures to facilitate comparisons for proportional changes. Error bars are 95% confidence intervals for interaction contrasts.

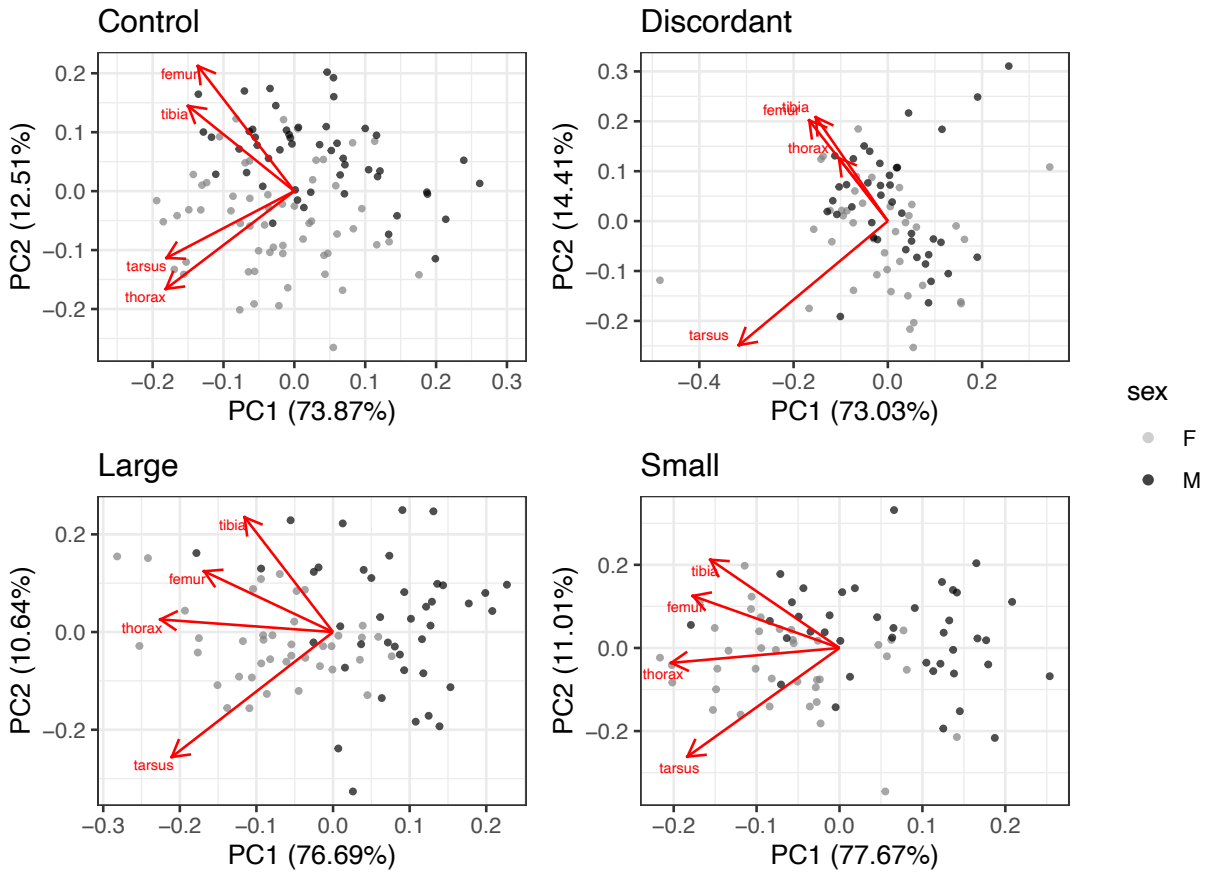

Figure 5: Sexually discordant selection alters patterns of multivariate allometry across sex. Represented as a biplot, the magnitudes and direction of the 4 traits are superimposed onto PC1 and PC2.

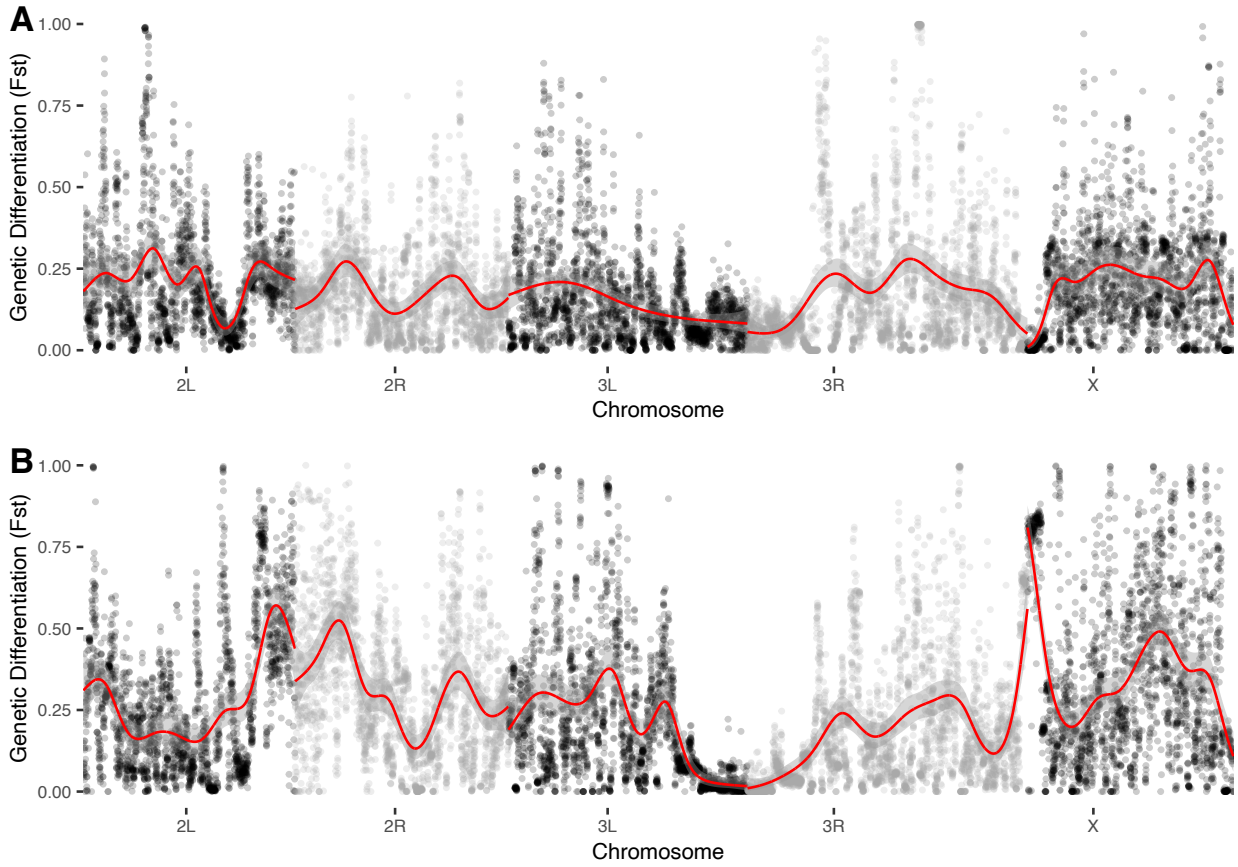

Figure 6: Genome-wide F_{ST} (10000bp windows). Chromosomal trends for F_{ST} (binomial, gamm) in red. A) Discordant selection compared to control treatments. B) Large compared to small treatments.

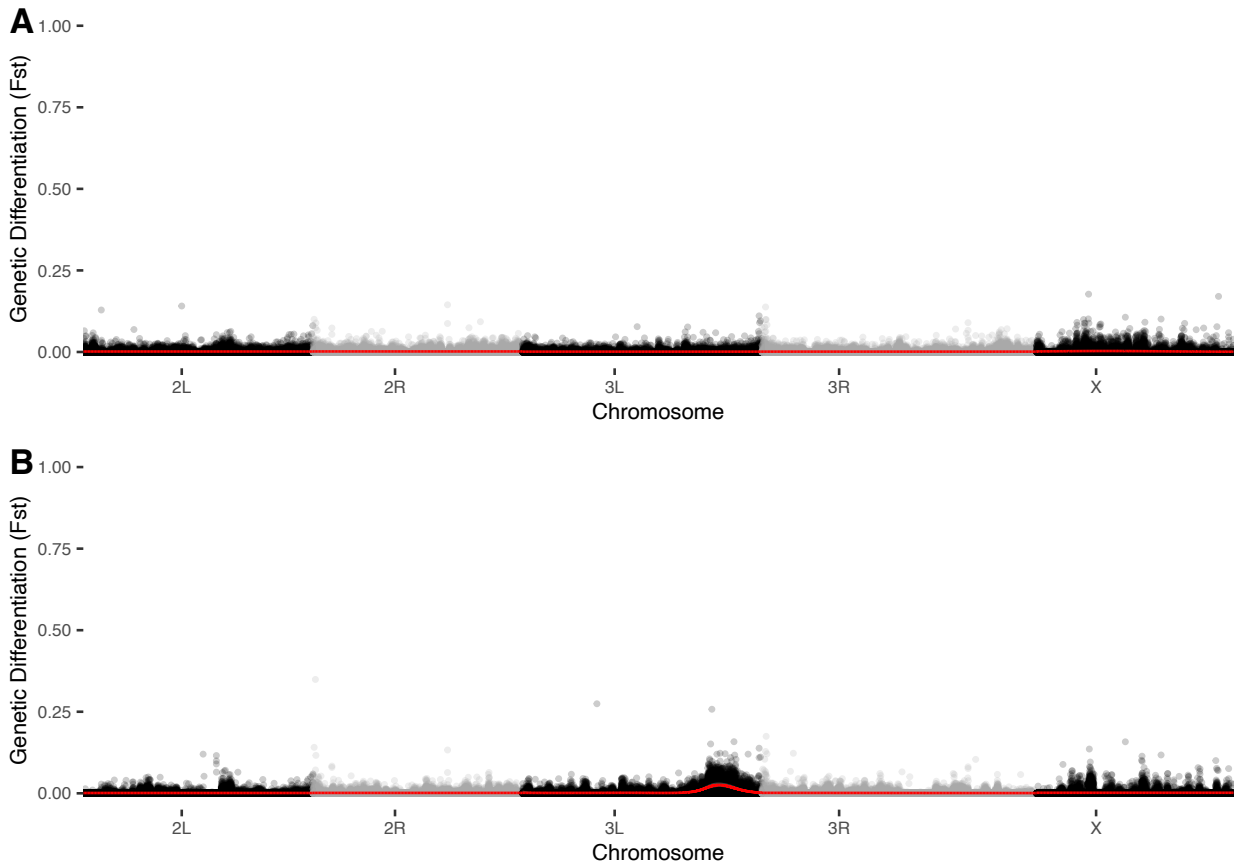

Figure 7: Genome-wide F_{ST} (10000bp windows). Chromosomal trends for F_{ST} (binomial, gamm) in red. A) Male vs. Female for control treatment (replicate 1), showing a trend ≅ 0,across the genome. B) Male vs. Female for discordant treatment (replicate 1) with a small region on chromosome 3L with elevated F_{ST} values.

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.