Evolutionary constraint of the shared genome

Evolutionary causes and consequences of sexual conflict and phenotypic divergence under the constraint of a shared genome

By Tyler Audet

A Thesis Submitted to the School of Graduate Studies in the Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy

A Thesis Submitted to the School of Graduate Studies in the Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy

McMaster University

Doctor of Philosophy (2025)

Hamilton, Ontario (Department of Biology)

TITLE: Evolutionary causes and consequences of sexual conflict and phenotypic divergence under the constraint of a shared genome

AUTHOR: Tyler Audet (McMaster University)

SUPERVISOR: Dr. Ian Dworkin

NUMBER OF PAGES: xi, 162

Abstract

Sexual dimorphism is common in nature. Whether it is sexual size dimorphism, or exaggerated trait dimorphism, males and females often look quite different. This sexual dimorphism arises despite a largely shared genome. To explore the evolution of sexual dimorphism, we use artificial selection, experimental evolution, and RNA-seq. We explore sex-biased gene expression in the sexually dimorphic *Drosophila prolongata* using RNA-seq analysis. We show that in *D. prolongata*, there are changes in sex-biased gene expression in developmental genes and transcription factors and changes in the magnitude and number of differentially expressed genes potentially leading to exaggerated trait development. Using artificial selection lineages with reversed sexual size dimorphism in *D. melanogaster* we identify a polygenic response to discordant selection, and a region differentially segregating between the sexes where selection pulls the sexes against their original sexual dimorphism. Using experimental evolution, we show that current hypotheses for the ecological precursors for sexually dimorphic evolution may be incomplete. Previous work has hypothesized that an ecological setting that creates potential for male monopolization of females and differential success for males who win duels would be sufficient to initiate the evolution of sexually dimorphic weapons. Here, we show that these ecological structures were insufficient to initiate sexually dimorphic evolution in *D. melanogaster* and propose that low density and opportunity for male-male signalling may be additional ecological ingredients critical for weapon evolution. Using these diverse methodologies, we expect to be able to add to our understanding of how sexual dimorphism evolves, and how a shared genome contributes to divergent phenotypes within a species.

Acknowledgements

I would like to first thank all my collaborators and undergraduates who helped tremendously with all of the works presented here, as well as works that did not make it into this thesis. Data collection for each of these experiments is far more than a one-person job, so it is without exaggeration that I can say this work would not have been possible without the help of all of these people.

I would also like to thank the graduate student community in the biology department, which has been supportive throughout my degree, especially the Dworkin lab members both past and present, who helped tremendously.

I also very much appreciate the members of my committee, Ben, Brian, and Sigal, who gave me so much amazing feedback and words of encouragement. Of course I have to thank my advisor, Ian, for always being willing to talk and supply coffee.

Finally, I want to thank my wife, Diana, who has been an unwavering support throughout this process, for which I cannot overstate my gratitude.

Contents

List of Figures

List of Tables

Declaration of authorship

Chapter 1: Most of this introductory chapter was written by myself, Julian Neves, and Dr. Ian Dworkin with plans to publish as a review on sex-biased gene expression.

Chapter 2: Study design and planning was done by Dr. Ian Dworkin, Dr. Stefan Lüpold, Dr. Jhoniel Perdigón Ferreira and Dr. Abhishek Meena, as well as myself. Sampling of tissues for RNA sequencing was done by Dr. Jhoniel Perdigón Ferreira and Dr. Abhishek Meena, as well as myself. RNA extraction and QC was done myself. Functional crosses were done by me. Phenotyping of crosses was done by John Yeom, Fatima Zaghloul, Nour Zaghloul, Arie Mizrahi, Julia Novikov, Emma Xi, and myself. Bioinformatic analyses were done by myself with the aid of scripts written by Arteen Torabi Marashi. Manuscript writing was done by myself and Dr. Ian Dworkin.

Chapter 3: Study design was done by Dr. Ian Dworkin. Bioinformatic analyses were done by myself with some scripts adapted from Dr. Katie Pelletier. Phenotyping of samples was done by Joelle Kroel. Artificial selection was done by Dr. Andrew Stewart. Manuscript drafting was done my myself, Dr. Andrew Stewart, and Dr. Ian Dworkin.

Chapter 4: Initial study design was done by Dr. Ian Dworkin and Audrey Wilson. Aggression assays were done by Dr. Reuven Dukas. Phenotyping was done by myself and Audrey Wilson, with help from Francesco Ruso. Experimental evolution was conducted by myself and Audrey Wilson. Manuscript drafting was done by myself, Audrey Wison, and Dr. Ian Dworkin.

Chapter 5: Written by myself.

Chapter 1

Introduction

1.1 Sexual selection and sexual dimorphism

The phenotypic differences between the sexes can rival or even exceed inter-specific differences. Sexual dimorphism has been an area of fascination for biologists since Darwin (1871). Striking differences in sexual dimorphism is (typically) attributable to sexual selection, which is occurs as either (or both) inter-sexual selection on one sex by the other for access to mating opportunity, or intra-sexual competition for access to mates (Andersson 1994). It has long been suggested that females are often the choosier sex (Bateman 1948). This choosiness is potentially due to the larger, more resource intensive eggs produced by females as well as the burden of offspring care which is often incurred by the female. Female choice and/or male-male competition may result in a much higher variance in male reproductive success than would occur without male competition for females. The first clear demonstration of increased variance in reproductive success in males was an experiment using *Drosophila melanogaster* by Bateman (1948), where, in the lab environment only 4% of females failed to reproduce, while 21% of males failed to reproduce. Bateman (1948) also found that fitness of males (number of offspring sired) increased with each subsequent mating at a greater rate than for females. This suggests a selective pressure on males (but not females) to mate as frequently as possible. This simplistic explanation however is mediated by more recent evidence that males can become sperm depleted and have reduced marginal gains in siring success with increased mating (Tang-Martínez 2016; Macartney et al. 2021).

Increased variance in reproductive success in males and choosiness in females (as well as remating latency for egg laying and production), potentially leads to fewer females willing to mate at any given time relative the number of males. This ratio of males:females willing to mate is the Operational Sex Ratio (OSR; Emlen and Oring 1977). This ratio is dependent broadly on parental investment which often (but not always) dictates which sex is more likely to compete for reproductive

access (Trivers 1974; Clutton-brock and Parker 1995). Deviation from a 1:1 OSR can lead to both inter- and intrasexual competition for mates. The presence of intra- or intersexual competition in males (or females in some species) can result in sexual dimorphism that favoured larger size or elaborate traits in the sex with the stronger competitive pressure (Emlen and Oring 1977). This can result in exaggerated trait evolution based on the type of selective pressure, inter-sexual mate choice resulting in ornamentation, or intra-sexual fighting resulting in weapons.

1.1.1 Evolution of weapons

Intra-sexual competition is common in many species and takes many forms (Andersson 1994; Clutton-Brock 2017). Despite the ubiquity of these contests, not all species evolve exaggerated weapons. Those that do evolve weapons have a large diversity of weapon forms and are taxonomically diverse (Emlen 2008). Contests that occur over patchy resources where the limiting sex (typically females) can be monopolized have been hypothesized to be an evolutionary catalyst for exaggerated weapon evolution (Emlen 2014). Within these resource competition sites, the resource should be laid out in a way where it results in 1:1 ‘duels’ that benefit the larger individual, rather than group scrambles. This skews the variance in reproductive success of males, exerting an evolutionary pressure to increase resource holding potential (Emlen 2014). These environmental factors are reflected in many species that have weapon exaggeration. For example, female dung beetles sequester dung either in burrows adjacent to dung piles, or they roll it away and bury it elsewhere. In species where dung is rolled, males pursue females and attempt to secure matings in large scrambles. In burrowing dung beetle species, the discrete burrow entrance creates the opportunity for large males to defend access to females within the tunnel. The opportunity to defend these tunnels means that larger males with the largest horns are better able to fend off competitors, whereas in the dung rolling species males have not evolved horns because competition occurs in scrambles (Emlen and Philips 2006). In stalk-eyed flies, males of species that guard narrow

rootlets coveted by females have exaggerated eyestalks that are used to ‘size up’ opponents in 1:1 matches, with smaller males more often retreating from the conflict. Species that do not interact on these restrictive rootlets are always more sexually monomorphic (Wilkinson and Dodson 1997). These patterns of weapon evolution appear to be well understood and intuitive, so much so that it has been argued that the study of intrasexual selection is ‘essentially complete’ (Jones and Ratterman 2009). However, research into the evolution of weapon evolution remains an active field with many open questions. Recent experimental evolution work has shown that altering the census sex ratio (as a proxy for the strength of sexual selection on males) does not result in the expected change in size dimorphism or condition dependence (Bath et al. 2023). The static allometry (log-log regression of trait size on to body size) of weapons has also been historically thought to be greater than one, meaning that as body size increases weapon size disproportionately increases in size (Kodric-Brown et al. 2006). This too may not be as straight forward as once proposed, with more recent meta-analyses showing inconsistent patterns of allometry (Voje 2016). This allometric discrepancy has been hypothesized to be related to the specific placement of the trait along a weapon/signal continuum. This continuum places weapons used exclusively as intra-sexual signals at one end and ‘pure’ weapons with no signalling component at the other, with most traits falling somewhere in between (McCullough et al. 2016; McCullough and O’Brien 2022). Many traits, such as the eyestalks of stalk-eyed flies discussed above, play an essential role in assessing the resource holding potential of competitors, and are not used directly in physical altercations. Fights in nearly all species studied begin with some form of display of size or strength (Palaoro and Peixoto 2022). Such signals and assessments of signals avoids costly interactions for males, especially when increased size disparity between combatants results in disproportionately more injuries (Emberts et al. 2021). Only a small number of species have been tested for this weapon/signal continuum relationship with static allometry, but the few examined support this continuum as being a predictor for allometry (McCullough and O’Brien 2022).

1.1.2 Evolution of ornaments

Selection on one sex (typically males) by another for a trait of interest, has resulted in some of the most iconic exaggerated traits, such as the peacock's tail. These traits are under selection of female preference, the origin of which is an outstanding question. One of the more prominent and tested hypotheses for the origin of female preference is the 'sensory bias' or 'sensory exploitation' hypothesis, which states that female preference for a trait predates male trait expression due to some other external factor such as food colour preference (Ryan 1998). The exact definitions for sensory bias are debated as well as whether the initiation of female preference has an adaptive advantage (Fuller et al. 2005). An adaptive advantage for the preference is not strictly necessary, and genetic drift, pleiotropy, and between sex correlation may be sufficient to explain the initiation of female preference (Lande 1980a). For example, females of the species *Xiphophorous maculatus* show preference for the sword tails of closely related *X. helleri* despite the fact that *X. maculatus* males do not develop swords (Basolo 1990). This suggests that preference for a trait can be present in the population and exploited by males if they evolve a trait when a pre-existing preference is present.

Once preference evolves in females and the trait that is expressed in males, many hypotheses have been put forth for the maintenance and exaggeration of these ornaments. One of the first models set out to explain the evolution of male ornaments was Fisher's model of runaway selection (Fisher 1930). Fisher envisioned a locus for preference, and a locus for trait expression which become linked through selection and evolve together. This model does not account for quantitative traits that are likely quite polygenic, nor does it account for polygenicity in preference, which was later modelled by Lande (1981). However, the fundamental idea of trait and preference becoming linked genetically is an underlying theme in many existing quantitative models (Mead and Arnold 2004). Another primary model for the maintenance of an ornamental trait is the 'good genes' model. In this model, females choose males based on their ornaments, which are an honest signal of genetic quality, and therefore their

offspring have increased viability (Petrie 1994; Møller and Alatalo 1999). This was historically held as an alternative to the Fisher process, but more recently it has been shown that these two processes may occur together (Chandler et al. 2013). These models do suffer from the inevitability of alleles for trait and preference, being under directional selection, becoming fixed in the population. If alleles for trait expression become fixed in the population, females have nothing to select on and therefore the selective advantage of displays is null, this is termed the lek paradox (Taylor and Williams 1982).

1.2 Condition dependence and genic capture

Despite the directional nature of selection acting on exaggerated traits, these secondary sexual characters often show levels of genetic variation equal to or higher than both primary sexual traits and non-sexually selected traits (Pomiankowski and Møller 1997; Kotiaho et al. 2001). The maintenance of this genetic variation despite directional selection has been an outstanding question in biology (Taylor and Williams 1982; Kirkpatrick and Ryan 1991). It has been previously demonstrated that these traits were highly condition dependent, meaning that under poor developmental conditions, sexual selected growth of the ornament or weapon decreased more dramatically than the growth of other metric traits (Andersson 1986). Rowe and Houle (1996) hypothesized that as sexually dimorphic traits evolved, they co-opted genetic mechanisms that modulate organismal condition, becoming highly polygenic and associated with developmental condition in the process. This ‘genic capture’ would both explain the association with condition, and the maintenance of high genetic variation, as trait expression became associated with condition which would be inherently polygenic. More recently, the evolution and expression of complex traits has been explored in the context of ‘omnigenic’ models (Boyle et al. 2017). This model suggests that all complex traits are controlled by effectively the entire genome, with loci having varying levels of effect on trait expression. Although these omnigenic ideas have yet to be fully explored from the context of sexual selection and sexual dimorphism, exaggerated trait expression being associated with the entire

genome would alleviate the need for a solution to a lek paradox, as trait expression would always be associated with overall genome quality and mutational input would be constantly adding potential alleles for selection. The number of genes associated with these exaggerated sexual traits does appear to be quite large as evidenced by the number of genes that, when knocked down, impact trait growth, potentially lending support to this hypothesis (Emlen et al. 2012; Warren et al. 2013; Gotoh et al. 2016; Hust et al. 2018; Rohner et al. 2021).

1.3 Sex-biased gene expression

Since morphological variation between the sexes is likely not a result of differences within the genome, the study of how gene expression varies between the sexes, sex-biased gene expression (SBGE), has emerged as a field which seeks to understand the role of gene expression in sexual dimorphism. In the past two decades, research on sexual dimorphism has increasingly focused on the promise of genomic approaches to help understand both proximate and evolutionary mechanisms resulting in sex differences in morphology (Wilkinson et al. 2013; Zinna et al. 2018), behaviour (Naurin et al. 2011; Wong et al. 2014; Khodursky et al. 2020), and disease (Zhang et al. 2011; Fass et al. 2024), among others. For example, researchers use the presence or absence of SBGE within a species as a tool for understanding the evolution of morphological or behavioural dimorphism as well as to infer sexual conflict (Connallon and Knowles 2005). With the assumption that SBGE represents the genome responding to differential selection, dimorphic evolution can be examined in the transcriptome of organisms (Pointer et al. 2013; Stuglik et al. 2014; Toubiana et al. 2021). Thus, the study of SBGE has implications for how we study the evolution of dimorphism and is an important interface between the genome and the phenotype of an organism (Ellegren and Parsch 2007; Mank 2017; Deegan and Engel 2019).

As a relatively young field coming to maturity alongside a leap forward in sequencing technology, experimental methodologies, analyses and inferential approaches are often slower to make

progress. This is an issue because it impedes the ultimate goals of large-scale biological conclusions as there is often a lack of comparability between studies. These can include major experimental concerns due to differences in tissue sampling (whole body or gonad versus non-gonad tissues) and differences with regards to the timing at which tissues are sampled (there are good arguments for sampling both juvenile and adult tissues). It can also be a result of the commonly used (but well known to be problematic) approach of generating artificial dichotomies on whether a gene shows sex-biased gene expression based on discrete p-value and/or log fold-change threshold values. Additionally, there may be subtle sampling and methodological considerations that could confound attempts at more general and comprehensive analyses. Presenting a survey of isolated genes, treated as discretely (from arbitrary thresholds using p-values for instance) differing between sexes tells us little about the extent to which SBGE meaningfully reflects or is associated with other phenotypic differences between the sexes. Taken together, this means that many studies of SBGE are not as informative or as powerful as they can be. As such, there is a need for a critical look at how we approach the study of SBGE.

Here, we review the experimental and analytical choices made in many studies of SBGE, and discuss concerns with common strategies, methodologies and assumptions. Specifically, we explore the issues with including the gonads in whole-body samples and the need for tissue-specific sampling. We then discuss some concerns with the commonly used practice of sampling tissues in adult stages, that are presented as informative of the underlying developmental processes. Next, we look at the standard way in which SBGE (and most studies of differential expression) is determined, through differential gene expression analyses of RNA-seq data between males and females, with differentially expressed genes defined as those that meet a certain p-value or log fold-change expression cut-off. In other words, instead of focusing on the question of “What is the effect/difference”, the question is posed to the much less informative question of “Is there an effect/difference”, while relying on the arbitrary (and often un-informative) null hypothesis of “no difference”. We discuss why this type of analysis, when presented on

its own (as is often done), is not very informative. Throughout this paper we provide suggestions and considerations for researchers so that they can make their studies of SBGE more meaningful and impactful, both in the experimental design stage and throughout the analysis of RNA-seq data. Importantly, we stress the importance of comparable results, so that accurate comparisons can be made between studies so that a consensus can be found in the literature.

1.3.1 Experimental considerations

1.3.1.1 Tissue specific sampling and choosing what to sample: what to sample

Choice of tissue used for a study of SBGE, particularly whether the tissue is a gonad (effectively sex-limited traits) or a whole-body sample, can introduce biases that must be considered when interpreting the results of the study. These issues have been raised before in both microarray and RNA-seq literature (Stewart et al. 2010). Unsurprisingly, gonads contribute a significant amount of SBGE in whole body sampling, potentially more-so than all other tissues included in the sample. In early experiments looking at SBGE, whole tissue sampling was the most common method of analysis (Assis et al. 2012). While non-reproductive tissues have been found to display SBGE and studies indicate that these tissues make up around 2% of sex-biased genes (Dutoit et al. 2018), it has been found that the number of sex-biased genes showing a two-fold expression change drops drastically when adults have their gonads removed (Parisi et al. 2004). While this first issue may seem obvious to many readers in the field, and fewer studies still do so this, whole-body sampling of organisms containing gonadal tissue persists (Djordjevic et al. 2021; Singh and Agrawal 2023; Mishra et al. 2024). As will be discussed further below, the extent to which this gonadal effect contributes to the whole-body estimates of SBGE might be impossible to quantify, without appropriate sampling controls. This is in large part due to the well-known issue of allometric relationships between organs (Montgomery and Mank 2016; Darolti and Mank 2022), that may differ between sexes but often goes unmeasured. Additionally, gonad specific SBGE can be sex-

dependent, in *D. melanogaster* males having a higher number of genes expressed in testis than in their soma, while females have equal gonad and somatic expression profiles (Parisi et al. 2004). Gene expression in gonad tissue also clusters by sex, because these tissues are so distinct (sex-limited), meaning gonad tissue is distinct in its expression compared to somatic tissue of either sex, as well as distinct between the sexes (Sharma et al. 2014). Gene ontology terms (GO) in gonad tissues of *D. melanogaster* also show enrichment in males for genes that control gamete generation and spermatogenesis, which could contribute to the strong bias in male expressed genes in gonad tissue (Perry et al. 2014). When whole organisms are used for RNA extraction, the inclusion of the gonads is likely to overwhelm the rest of the tissues in terms of SBGE, giving the false impression that SBGE is spread throughout all tissues when it may be largely due to the gonads. Singh and Agrawal (2023) looked at SBGE in both heads, as well as whole body in *D. melanogaster*, and found large changes in both male- and female-biased expression in whole body samples, but far more modest changes in head tissue, suggesting a bias toward 'significant' changes when whole bodies are sampled. This overwhelming bias towards gonad expression in whole body samples suggests that moving towards non-reproductive tissue-based expression studies is more biologically relevant in the majority of cases. This strategy has already been employed in larger species such as turkeys (Pointer et al. 2013), and collard fly catchers (Dutoit et al. 2018). This has also been used in some recent studies to look explicitly at sexually dimorphic somatic traits, using other non-dimorphic somatic tissues as a control on the same organism such as Rhinoceros beetles (Zinna et al. 2018) and water striders (Toubiana et al. 2021). These studies show that this somatic tissue-based sampling still provides novel insight into sex-biased gene expression, and possibly even more informative differences than whole body or gonad only sampling. This demonstrates that when exploring phenotypic differences between the sexes, non-reproductive tissue has adequate, and perhaps more informative gene expression differences. If one were specifically interested in transcriptomic differences between the sexes in the whole organism, a somewhat better

approach would be to first remove the gonads to minimize erroneous overinflation of detected sex-biased genes. Another possible way to address the issue of gonad biases, is to sequence whole body samples, but to also sequence gonad only samples to be able to look at results with and without the gonad, and account for gonad-specific expression in models, but this would not alleviate the issues of tissue allometry (Rago et al. 2020). If the sexes inherently have differences in relative tissue size, such as one sex with larger brains relative to body, then even without the gonad included the transcripts sequenced would be biased toward larger tissues relative to body.

Whether due to subtle differences in developmental timing of tissues, or different relative contribution of cell types to the tissue, there may always be some level of confounding effects (i.e. challenging to do apple to apple comparisons). As such, as a community considering the potential merits and trade-offs of single cell sequencing (scRNA-seq) to "bulk" approaches should be considered. For example, Lu and Mar (2020) investigate the expression of different cell types in the heart and brain of male and female mice and identified cell-type specific marker genes that differed between the two sexes. With this study, the authors also present a pipeline for identifying gene regulatory networks for multiple cell types from scRNA-seq data. Early work looking at scRNA-seq data by Darolti and Mank (2023) aimed to explore whether scRNA-seq data and traditional bulk tissue RNA-seq data resulted in differences in inference due to sex differences in cell type abundance. In guppies they sample sexually dimorphic pigmented skin tissue, as well as livers, heart, and gonads from males and females and sequenced at the single-cell level. Importantly, they found that the majority of cell types in reproductive tissues were sex-limited, again highlighting the fact that comparing male to female reproductive tissue is an apples-to-oranges comparison. More interestingly, when SBGE was analyzed at the cell-type level (expression for each cell type pooled together for all cells) they found that 49-77% of SBGs were not identified at the bulk-tissue level. They also found that many genes identified at the cell-type level were not identified as sex-biased, suggesting cell-type abundance is a potentially confounding factor for

SBGE at the tissue level (Darolti and Mank 2023). With its high resolution to detect changes in gene expression that would certainly otherwise be cancelled out by whole-body sampling, and even potentially tissue-specific sampling, the study of SBGE at the single-cell level is another interesting avenue for further exploration.

1.3.1.2 Sex-biased gene expression in the context of development: when to sample

Currently, there is a bias on sampling tissue from adult organisms in studies, likely due to experimental convenience. These studies are informative for asking questions about adult phenotypes, such as behaviour (Naurin et al. 2011; Wong et al. 2014; Khodursky et al. 2020). For developmental processes however, or for phenotypes formed during development, sampling during the dynamic developmental stages may be more informative. The logic for sampling adult tissues is reasonable, because at this stage obvious phenotypic dimorphism is present. However, this approach ignores the importance of development in setting the stage for adult morphological, physiological and even behavioural sexual dimorphism. As such, this sampling approach paints an incomplete picture of the changes in gene expression that result in phenotypic variability between the sexes. For example, it is known that in mammals, there is some sex biased gene expression present in embryos as early as the 8-cell stage, and the transcriptome is dynamic throughout all early ontogeny (Lowe et al. 2015).

The degree and direction of sex-bias may vary across developmental time. In a subset of mammals (mouse, rat, opossum, and rabbit) as well as chickens, the number of SBGs was highest around the transition to sexual maturity (Rodríguez-Montes et al. 2023). There were a smaller number of SBGs early in development, however the early development SBGs were associated with phenotypes that were sexually dimorphic in adulthood such as body weight, as well as gonadal sex-determination (Rodríguez-Montes et al. 2023). Few of these genes were conserved across species, suggesting conclusions that we can make about SBGE may be species specific in many cases. In the stick insect

Timema californicum the direction of sex-bias rarely changes during development and sex-biased genes early in development gradually increase in their degree of sex-bias throughout development corresponding to the degree of phenotypic sex-bias (Jelisaveta et al. 2021). A number of sex-biased genes in *T. californicum* also showed stage specific expression, which was at times related to sex-specific phenotypes such as pigmentation (Jelisaveta et al. 2021). Despite this information, much of the research done on sex-biased gene expression has occurred in adult tissues.

The sex-biased relationship between genes also seems to change throughout development. Even genes that are clearly sex-biased at one sampling time may simply be sex-biased due to their functional role rather than their sex-specific role. As an example, genes involved in gametogenesis in *Nasonia* are expressed in a male-biased direction early in development and female biased direction in late development. This effect is unlikely to be an effect having anything to do with SBGE in the context of evolution, but rather reflecting the fact that spermatogenesis occurs earlier than oogenesis in this genus. In fact, in *Nasonia*, only *feminizer* (orthologous to the *Drosophila transformer* gene) shows consistent sex-bias throughout all of development (Rago et al. 2020). All these examples are a clear demonstration that sampling during adulthood is likely too late for many questions, and sampling during development without consideration for the underlying biology may be a fool hardy task. If a dimorphic tissue is of interest, sampling during or directly prior to the development of the dimorphism is when the signal of interest is likely to be. Importantly, sampling well after the appearance of the dimorphism (such as after development) may show SBGE, but it may not be the signal of interest (i.e., not a signal from the genes responsible for producing this dimorphism as it has already been long established), but a consequence of the dimorphism in the tissue (and underlying cell types).

1.3.2 Analytical and inferential Considerations

Regardless of whether the investigation is done on whole body samples or on specific tissues, and regardless of developmental or ontogenetic timing, most studies limit their findings of what genes show SBGE based on thresholds based on a particular fold-change in expression and/or p-value cut-off. Threshold cut-off methods have been employed to detect hundreds of differentially expressed genes across many different organisms and is a useful tool across many disciplines in genetics for detecting genes of large effect associated with phenotypic variation. However, this approach has several issues. First, the threshold p-values and log fold-change values used are arbitrary and not based upon any biological significance (i.e. how much of an expression change for a transcript elicits a particular biological effect). If researchers do not account for raw transcript count number, results can be misleading. As an example, samples with 10000 transcript counts in males versus 20000 transcript counts in females or a gene with 5 counts in males and 10 in females would both have the same $\log_2$ fold-change, and depending on uncertainty estimates may have the same statistical 'significance', yet they likely have different biological implications. This approach can give an idea of how many, and which genes are differentially expressed, but not the magnitude or direction of this effect. Finally, identifying single genes that are differentially expressed between the sexes and leaving the analysis at that, ignores the fact that genes do not operate in isolation and act within expression networks, making results hard to interpret.

1.3.2.1 The issue with choosing a cut-off

Typically, gene expression studies utilize a 1.5x or 2x $\log_2$ fold-change in combination with a p-value of 0.05 (adjusted for multiple comparisons) or an FDR to define which genes are "differentially expressed" between the sexes and deemed biologically interesting. These approaches are well known to be deeply problematic as they fundamentally turn any effect size that does not meet said threshold to "0". Despite the use of arbitrary cut-offs being less frequently used in other areas of biology, in studies of genomics

they remain surprisingly common (one only need to think of traditional Manhattan plots to realize the ubiquity of such significance thresholds). With respect to the study of SBGE, significance thresholds can cause further inferential complications. For example, reproductive tissues have the most genes that are sex-specific in expression, whereas the somatic tissues have far fewer, and the magnitude of differences in somatic tissues is often much smaller meaning that threshold cut-offs may often miss interesting SBGE in the soma potentially of interest (Sharma et al. 2014). In combination with the fact that many of these studies sample whole bodies (Mishra et al. 2023; Singh and Agrawal 2023) this choice of a particular threshold is even less biologically informed because the genes that exceed these thresholds most often or by the highest margin, are most likely genes responsible for gametogenesis or gonadal development (Sharma et al. 2014).

The issue with choosing an arbitrary threshold of expression differences (*a posteriori* or *a priori*) remains equally problematic, at least without the use of biologically grounded biological priors. At face value the comparative doubling (or halving) of gene expression across treatment groups or sexes may seem sensible. However, it is rarely clear for which genes such a change would be sufficient to elicit a phenotypic response at the cellular or tissue level, ultimately impacting sexual dimorphism. While not from the SBGE literature, some previous studies have demonstrated that "key" developmental genes influencing tissues can vary more than 2X in expression across strains despite the organs remaining wild type in morphology. On the other hand, genetic perturbations that have profound impacts on organ size often can have modest expression changes (well below the 2-Fold "threshold"; Dworkin et al. 2009, 2011; Chandler et al. 2014). Dworkin et al. (2009) found that a mutation in the gene scalloped in the wing of *D. melanogaster* had profound effects on phenotype with a 50% reduction in expression, but other developmental genes differed by 8x expression between wild-type strains.

So, what should be done instead? Fundamentally it is important to be asking "what is the effect of gene X on sexual dimorphism?" rather than "Is there an effect?". This slight change in thinking can

have substantial impacts on how we interpret the results. After all, except for a few genes in the genome directly linked to sex determination, how reasonable is it to expect that there is truly "no difference in expression" between males and females in most genes. It is most likely that in most cases the expression differences are so small that they have little to no biological consequence. However, we also want to narrow down the full set of expressed genes to a manageable set for further study. As we discuss below, a thoughtful combination of considering regularization or shrinkage of estimates along with multivariate approaches can still be fruitful, allowing for subtle changes in individual gene expression to provide information for genome wide analyses.

1.3.2.2 Estimate Regularization/Shrinkage of log fold changes (LFC) and including uncertainty as weights in downstream analyses

Mean expression levels of genes vary over many levels of magnitude. This in turn can have substantial implications for the interpretation of fold differences between males and females. In particular, when examining transcripts with overall low abundance, sampling effects can result in the impression of very large expression differences (Van De Wiel et al. 2013; Love et al. 2014; Zhu et al. 2019). A difference of ~10 transcripts per million (TPM) may translate to a 2-fold change in a lowly expressed gene may, but the same difference between a highly expressed gene (1000 TPM vs. 1010 TPM) would be considered far less significant. This creates a bias in these analyses towards lowly expressed genes as being considered interesting. Compounding this bias is a tendency to not report absolute TPM values in the results section, creating statements of significance such as a log2 fold change in a gene, with no context as to whether that is 10 extra TPM or 10000. Even with the use of a negative binomial distribution or variance stabilized values for gene expression, these sampling effects can be substantial on fold changes. This is important for studies of SBGE, as the fold changes (or log2 differences) are often used for downstream analysis and inferences (often without use of SE as

appropriate weights in downstream models). When these differences are used “as is” this can result in strong confounding effects of “large” fold differences with expression abundance. As a result, the “big” differences in SBGE may reflect low abundant gene expression more often than not. As such, it is recommended to apply regularization/shrinkage to the log fold change (LFC) estimates (Van De Wiel et al. 2013; Love et al. 2014; Zhu et al. 2019). R packages such as apeglm and ashr (Stephens et al. 2016; Zhu et al. 2019) can be used from standard output of fold differences and associated measures of uncertainty, and there is a direct interface (lfcshrink) in DESeq2 (Love et al. 2014).

If LFC for sex-biased gene expression is being used for downstream analyses, we also highly recommend that authors use measures of sampling uncertainty (such as standard errors on LFC) as weights in the downstream analysis. For instance, if such a downstream analysis using LFC is being used in the context of a generalized linear model in R, then the (g)lm function can easily be supplied with the (weights =) argument and can be supplied with a vector of the (inverse) of the standard errors for LFC. As it would in other contexts (such as a meta-analysis), this would downweight the contribution of a LFC that has a high degree of uncertainty associated with it.

1.3.2.3 Preserving the context of gene function through gene network (co-expression analyses)

An issue with only looking at individual differentially expressed genes is that it isolates these genes from the genetic networks (and correlations) that they operate in. This makes results hard to interpret. Previous work suggests that expression changes in one gene have widespread consequences on the expression of others, so treating each gene as an independent variable violates these assumptions (Lemos et al. 2008). One way to preserve the network context within which genes operate is by considering the networks of gene co-expression that differ between males and females, rather than the expression of single genes. This approach, which commonly uses a weighted gene co-expression approach (WGCNA; Zhang and Horvath 2005) to cluster genes into groups of hub and node genes based

on correlations in gene expression, has been used recently in studies of sex-biased gene expression (Deegan and Engel 2019; Sutherland et al. 2019; Lopes-Ramos et al. 2020; Rago et al. 2020). In many cases, these analyses can identify significant sex-specific co-expression networks even in situations where few or no differentially expressed genes are found, suggesting that looking for differential expression of single genes alone paints an incomplete picture of the transcriptomic landscape of sexual dimorphism.

Combining analyses of differential gene expression with an analysis of sex-specific co-expression networks provides a more robust understanding of the transcriptomic changes underlying sexual dimorphism. For example, Sutherland et al. (2019) used WGCNA to identify gene-expression network (module) preservation between male and female *Salvelinus fontinalis* using expression profiles from the liver. They find that modules were generally preserved between males and females (recapitulating previous observations that modules show medium to high preservation between closely related species or sexes), and a comparison of module results to a single-gene differential expression analysis revealed that identified individual sex-biased transcripts were not overrepresented in sex-specific modules. These results suggest that a combination of classic DGE approaches and a WGCNA approach are informative. Rago et al. (2020) use WGCNA to identify sex-biased gene expression and sex-biased regulation of co-expression networks across five distinct developmental stages in the jewel wasp *Nasonia vitripennis*. They find evidence for sex-biased co-expression networks in early development even though no sex-biased genes were detected for these same stages. These genes were only detected later on in development/adulthood, when sexual dimorphism becomes more apparent and suggests that the differential expression of single-genes is more crucial for late development, although sex-biased modules are important before these differences become apparent. These two methods are complimentary, and if one has gene expression data available anyways it would be worth the extra effort to investigate gene co-expression networks as well.

1.3.2.4 Supplementing individual Gene contrasts with multivariate approaches (Vector Correlation Analyses, magnitudes)

A complementary approach to those discussed above is to use some simple methods borrowed from multivariate statistics and linear algebra (Kuruvilla et al. 2002). While much more advanced approaches can be found (Innocenti and Chenoweth 2013; McGuigan et al. 2014), to start with we recommend analysis of the magnitude of the LFC across sexes, the ratio of the magnitudes (sometimes called "alpha") and the correlation among vectors of sex-biased change. Understanding how the whole transcriptome (rather than just single genes) changes in result to SBGE is to compare vectors of effect between the sexes (Kuruvilla et al. 2002; Zinna et al. 2018). This method is advantageous because it can quantify the degree to which overall gene expression between the sexes changes at a level that is broader than even an analysis of co-expression network analysis. For example, Zinna et al. (2018) studied sex-biased gene expression between male and female Asian rhinoceros beetles, they looked at tissues with varying degrees of sexual dimorphism as well as in 'low' and 'high' condition individuals through nutritional manipulation. By examining correlations for vectors of expression changes, they were able to identify whether genes identified as sex-biased within a contrast were also expressed similarly across contrasts (across tissues, or developmental nutrition). Additionally, the vector correlations could be used to investigate how similar the gene expression differences were between two sets of contrasts, and whether the differentially expressed transcripts are more correlated than expected compared to a random subset of the full set of genes. We believe that the combination of a differential gene expression analysis, an investigation into differentially co-expression, and the use of vector correlations could give a truly in depth understanding of the transcriptomic landscape underlying sexual dimorphism.

1.4 Conclusions

The increasing accessibility of high throughput sequencing has added power to our ability to ask questions about the evolution of sex-biased gene expression and sexual dimorphism. Several long held theoretical models can be easily tested with these methods, and previously unattainable quantities of data are available for any lab interested in asking these questions. However, in the face of increasing accessibility and constantly evolving methods, it is important to critically think about the way we sample and the way we analyze these data. Tissue selection must be geared towards the questions we want to ask, and this means basing tissue choice on research hypotheses and ensuring to not include the gonads unless there is a theoretical reason for their inclusion. We also need to expand our examination of SBGE across developmental time, because the evolution of dimorphism occurs before the dimorphism is fully formed and for this reason, we must likewise sample earlier than we have thus far. In terms of thinking about how data is analyzed, it would be beneficial include analyses of gene co-expression networks to analyses of the differential expression of single genes, so that we add back in the context in which genes act. Another way in which to consider this context would be to conduct analyses of the gene vector correlations, which would allow us to understand changes with regards to gene expression at the level of whole transcriptomes. These current growing pains may be slight, and the power of RNA-sequencing is great, and for this reason it seems clear that with a small adjustment in our strategy, the field of sex-biased gene expression and dimorphism is ready for a golden age.

2. Genetic architecture of the developing forelegs of *Drosophila prolongata*; an exaggerated weapon and ornament

*This work will soon be submitted to PLoS Biology and is therefore formatted for submission.

Genetic architecture of the developing forelegs of *Drosophila prolongata*; an exaggerated weapon and ornament

Tyler Audet[1], Jhoniel Perdigón Ferreira[2], Abhishek Meena[2], Mariam Abass[1], Arteen Torabi Marashi[1], John Yeom[1], Fatima Zaghloul[1], Nour Zaghloul[1], Arie Mizrahi[1], Julia Novikov[1], Emma Xi[1], Stefan Lüpold[2]*, Ian Dworkin[1]*

Contribution:
Study design: TA, JPF, AM, SL, ID
Sample collections: TA, JPF, AM
Bioinformatics: TA, ATM
Functional crosses: TA, MA, ID
Dissections and imaging of crosses: TA, JY, FZ, NZ, AM, JN, EX
Analysis: TA, ID
Manuscript draft: TA, ID
Manuscript Editing:
Funding: SL, ID

1. Department of Biology, McMaster University, Hamilton, ON, Canada
2. Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zürich, Switzerland

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.

Acknowledgements: We would like to thank Alain Garces for Grn antibodies, Artyon Kopp for *D. carrolli* and *D. rhopaloa*, David Luecke, Yige Luo, Artyom Kopp for draft genome annotations of *D. carrolli* and *D. prolongata*. We would also like to thank the undergraduate volunteers who aided in imaging of flies. This work was funded by a National Science and Engineering Research Council discovery grant to ID and an Ontario Graduate Scholarship to TA.

Abstract

Extreme secondary sexual traits are some of the most striking phenotypes in nature. Understanding the genetics of these phenotypes has largely been focused on within species functional analyses of signalling pathways. Although useful, this does not offer insight into the evolutionary mechanisms that occur during the evolution of trait exaggeration. *Drosophila prolongata* offers a unique opportunity to explore the evolution of trait exaggeration, as it is the only species in the melanogaster species group containing male-specific foreleg exaggeration. Here, we used RNA-seq data from developing foreleg and midleg tissues at two developmental stages corresponding to early and post initiation of sexually dimorphic growth between these tissues and between closely related species. We sampled both males and females from *D. prolongata* as well as two closely related species, *D. carrolli* (diverged ~4MYA), and *D. melanogaster* (diverged ~20MYA). Using comparisons of gene expression between sexes, species, tissues, and developmental stages we found a positive relationship between the number of sex-biased genes, but not magnitude of expression difference, and the extent of phenotypic dimorphism. We identified a potential gene of large effect, *grain*, which when knocked down induces *D. prolongata*-like phenotypes in *D. melanogaster* legs. We also found very modest changes to the direction and magnitude of sex-biased gene expression in previously implicated signalling pathways for sexually dimorphic evolution, suggesting that these pathways are used to express, but not evolve, these traits of interest. This work adds to our understanding of the evolutionary mechanisms that underly trait exaggeration.

Introduction

Elaborate or exaggerated secondary sexual traits are some of the most striking in the animal kingdom. These secondary sexual traits fall along a spectrum for two roles, from weapons, used for fighting or intimidating other members of the same sex, to ornaments, used for attracting mates of the opposite sex (McCullough et al. 2016). Exaggerated weapons may evolve when ecological conditions are such that males with high resource holding potential can engage in one-on-one duels to secure disproportionate access to females (Emlen 2008). Ornaments generally evolve as a consequence of female preferences, potentially as a function of indirect benefits via good genes, runaway processes, among other mechanisms (Darwin 1871; Fisher 1930). These secondary sexual traits can be a substantial fraction of overall "size" relative to an individual's body (Kodric-Brown et al. 2006; Voje 2016). The genetic architecture underlying such a trait has been a topic of interest in evolutionary biology, with implications for the relative plausibility of some evolutionary models for sexually selected traits, to a deeper understanding of how sexual conflict and impacts of the inter-sex genetic correlation can have with respect to facilitating or hindering adaptive evolution and achieving sex-specific optima.

Early models of Fisherian runaway processes and good genes models often used a population genetic framework, assuming simple genetic architectures. However, not all of the findings of these models are supported under more biologically plausible genetic architectures, in particular a high degree of polygenicity for ornaments (Pomiankowski and Møller 1997). For instance, the nature of the Lek paradox, where it is predicted that genetic variation for the ornament should be depleted rapidly, under some forms of good genes models (Kirkpatrick and Ryan 1991). Yet, empirically, these traits often have as much, if not greater amounts of genetic variance (Pomiankowski and Møller 1997). Resolutions to this "paradox", via the evolutionary co-option of condition dependence have been proposed (Rowe and Houle 1996). However, if the trait that ultimately becomes the target of sexual selection initially has a highly polygenic basis (either with an infinitesimal or omnigenic-like distribution of allelic effects), then

there would be no paradox at all. Even strong and consistent sexual selection would not deplete genetic variance as substantial phenotypic response could occur via modest shifts in allele frequencies genome-wide, accompanied by substantial input of variation by new mutations (Rowe and Houle 1996). If the genetic architecture of exaggerated trait variation is a result of alleles of large phenotypic effect in a small number of genes, this alters predictions substantially.

The other outstanding question is on the genetic mechanisms through which sexual dimorphism is achieved. Specifically the sex-specific mode-of-action of allelic effects, and the degree to which those effects vary in magnitude across the sexes (i.e Genotype-by-Sex interactions) or are completely sex-limited in function (Testa and Dworkin 2016; Zhu et al. 2023). While some secondary sexual traits, like the sex combs among species of *Drosophila,* are sex-limited, many others represent sex specific elaborations or exaggerations. Knowing the relative contribution of sex-limited genetic effects, or amplification of magnitude or direction of genetic effects, is crucial for our understanding of the evolution of sexual dimorphism.

One potential mechanism that is implicated in the process of generating sexual dimorphism and potentially resolving sexual conflict from a largely shared genome, is the role of sex-biased gene expression during development. Since the dawn of the "functional genomics" era, sex differences in gene expression have been evaluated across many species, tissues, and time points (Ellegren and Parsch 2007; Parsch and Ellegren 2013; Ingleby et al. 2015; Mank 2017). Whole body studies of sex differences often suggested a large fraction of expressed genes showed moderate to substantial differences in expression, while more targeted, tissue specific studies, in particular during trait development, have generally been consistent with a smaller number (although often still in the hundreds) of genes showing sex differences, with the number of genes showing differences increasing at adult stages (Perry et al. 2014). While distinguishing cause from consequence in many of these studies is challenging, there are several broad inferences that have been drawn. Among sexually-dimorphic (but

not sex-limited) morphological traits, there is evidence for an increase in number and often magnitude of expression in the associated tissues during development (Zinna et al. 2018; Toubiana et al. 2021). When the sexual dimorphism involves exaggeration of size, genes related to growth control are commonly found (Moczek and Rose 2009; Wilkinson et al. 2013; Cox et al. 2017; Zinna et al. 2018; Toubiana et al. 2021). Interestingly, there is also often evidence of both male and female-biased gene expression in these tissues, even when the sexual dimorphism for size is male biased. The implication of these studies is that the sexual dimorphism is likely mediated by numerous genes that influence overall degree of sexual dimorphism.

Additionally, functional genetic studies perturbing candidate genes have become increasingly common. These studies enable assessments not only of sex differences in expression, but also of differences in sensitivity to perturbation of genes. As with transcriptional profiling, genes associated with growth signalling have been implicated in the development of sexually exaggerated traits. Due to the common pattern of condition dependence of these structures, primary candidates for exaggerate trait growth are insulin signalling and Target of Rapamycin (TOR) signaling pathways which regulate cell proliferation, growth, and metabolic processes (Warren et al. 2013). In *Drosophila melanogaster*, knockdown of insulin pathway genes reduces sexual dimorphism for body size by reducing the size of the larger sex disproportionately (Shingleton et al. 2005; Rideout et al. 2015; Sawala and Gould 2017; McDonald et al. 2021; Millington et al. 2021a,b). These previous results, as well as their potential to explain condition dependent signalling, make the insulin signalling pathway a strong candidate for pathways of interest in sexually dimorphic evolution. Genetic knockdowns of *insulin receptor* (*InR*) using RNAi in beetles leads to the reduction or loss of horns in the rhinocerous beetle *Trypoxylus dichotomus* (Emlen et al. 2012), with similar results observed for sexually dimorphic horn development in dung beetles (Casasa and Moczek 2018; Rohner et al. 2023). There is also a response in insulin signalling genes in gene expression data during development of exaggerated tissues in *T. dichotomus*

(Zinna et al. 2018). Changes in expression in higher condition individual and phenotypic response of insulin like peptide knock-down has also been shown to influence exaggerated weapon (mandible) size in the flour beetle *Gnatocerus cornutus* (Okada et al. 2019).

Somatic sexual differentiation in *Drosophila* and other holometabolous insects is largely cell-autonomous (Dobzhansky 1931; J Patterson 1938; Jahner et al. 2015), as observed in gynandromorphic individuals. In *Drosophila* this is mediated by the core sex-determination pathway initiated by the X-linked, and dosage dependent *Sex-lethal* (*sxl*) gene, resulting in a pattern of sex-specific production of downstream isoforms of genes such as *transformer* (*tra*) and *doublesex (dsx)*, and ultimately sex-biased gene expression within and among tissues (Gowen and Fung 1957; Baker and Ridge 1980; Yan and Perrimon 2015; Hérault et al. 2024), influencing sex specific patterns of growth via hormonal influences on metabolism and cell-autonomous mechanisms (Rideout et al. 2015; Mathews et al. 2017; Millington et al. 2021b; Wat et al. 2021). *dsx* mediates trait expression for the male-limited sex combs in *D. melanogaster* (Tanaka et al. 2011; Rice et al. 2019). In the dung beetles *Onthophagous taurus*, *dsx* knockdowns reduces horn size in males, and induced rudimentary horns in females. *dsx* also induces horn development in males and females of *O. sagittarius* (Kijimoto et al. 2012; Ledón-Rettig et al. 2017). In the golden stag beetle C*yclommatus metallifer*, knockdown of *dsx* results in a substantial reduction in the exaggerated male mandibles and a modest increase in the growth of female mandibles (Gotoh et al. 2014, 2016). Finally perturbation of genes involved with axis specification and limb patterning can influence sex-specific trait exaggeration, including the Hox gene *ultrabithorax* (*ubx*) in water striders legs, (Khila et al. 2009; Refki et al. 2014; Crumière and Khila 2019), and knockdowns of the gene *sex-combs reduced* (*scr*) which reduces horn size in the dung beetle *O. nigriventris* (Wasik et al. 2010). Interestingly, genes that mediate limb patterning and growth, such as *distal-less* (*dll*) have also been implicated in the development of beetle horns (Moczek and Rose 2009). The perturbation of these

genes provides important information regarding genes influencing trait expression, though not necessarily representing causal changes underlying the evolution of trait exaggeration *per se*.

While trait specific sexual size dimorphism is common among the species of the Drosophila melanogaster species subgroup (at least ~177 species) (Kopp 2006), the dimorphism is quite modest (less than 15% differences in size), and often female-biased, likely as a result of the strength and consistency of fecundity selection. *D. prolongata*, a member of the rhopaloa clade within this larger melanogaster clade is notable for the exaggerated size of the forelegs in males (Figure 1). Species of the rhopoloa clade are found in Southeast Asia (Singh 1977), and foreleg size exaggeration does not occur in other species of the rhopoloa clade (Figure 1), despite being only ~4 million years diverged from a common ancestor with *D. carrolli* which displays a more typical pattern of female-biased body size and approximately monomorphic legs (Setoguchi et al. 2014). The relative size of forelegs of *D. prolongata* are predictors for the outcome of contests between males and winners of fights disproportionately secure matings, suggesting a role as weapons (Toyoshima and Matsuo 2023). Males of *D. prolongata* also display male-biased overall body size dimorphism, unlike the majority of other Drosophilids (Rohner and Blanckenhorn 2018). *D. prolongata* forelegs are patterned in alternating black and white bands and are used by males for novel courtship behaviours including arm displays and female vibration, suggesting a role as an ornament influenced by female preference (Setoguchi et al. 2014). The closely related species *D. carrolli* does not show enlarged forelegs, female vibration, or male-biased dimorphism for body size (GOMPEL and KOPP 2018). The natural history of *D. prolongata* and *D. carrolli* are largely unknown, and therefore the evolutionary context of this trait exaggeration is an outstanding question. However, the evolution of exaggeration of a trait in just one lineage within this clade makes it an exceptional model for exploration of the evolutionary mechanisms of trait exaggeration.

Sex-biased gene expression (SBGE) changes throughout development, in both the number of sex biased genes and the direction of sex bias, and for this reason looking at SBGE over developmental time is important to get a better understanding of temporal changes in gene expression (Ingleby et al. 2015). In *D. prolongata* males, foreleg imaginal discs have on average 16% more cells than midleg imaginal discs by the end of larval development, while female leg imaginal discs are approximately equal in number between foreleg and midleg (Luecke and Kopp 2019). In contrast *D. carrolli*, both foreleg and midleg imaginal discs (matched for developmental timing) have approximately the same number of cells throughout larval development across the sexes (Luecke and Kopp 2019). This relative increase in cell number in the forelegs for male *D. prolongata* is not present in the earliest stages of the third (final) larval instar but becomes increasingly foreleg-biased as the relative amount of cell proliferation in males increases in this imaginal disc (Luecke and Kopp 2019). The genome of *D. prolongata* and *D. carrolli* have recently become available (Kim et al. 2021; Luecke et al. 2024). *D. prolongata* and *D. carrolli* have already been demonstrated to have strongly sex-biased expression for sexually selected cuticular hydrocarbons in oenocyte cells in adult *D. prolongata* (Luo et al. 2025). Luo et al. (Luo et al. 2025) identified 526 male-biased *D. prolongata* genes, 53 of which were not also biased in *D. carrolli*, a relatively smaller number of genes than would be expected under models such as genic capture that suggests traits under sexual selection become associated with a large portion of the genome. This SBGE is potentially in part due to the transposable element which they named ‘honghaier’ moving regulatory binding sites for *dsx* and *bric-a-brac 1* (*bab1*) to new loci in *D. prolongata* in front of a group of five cuticular hydrocarbon genes (Luo et al. 2025).

Here, we evaluated changes in SBGE in the developing leg imaginal discs in *D. prolongata*, *D. carrolli*, and *D. melanogaster* during two time points during the final larval instar (20 hours before wandering) and late (8 hours before wandering) in development, corresponding roughly to initiation of, and shortly after imaginal disc cell number becomes clearly sexually dimorphic in *D. prolongata*

(Luecke and Kopp 2019). Within and between species we compared imaginal disc expression profiles of both foreleg and midleg to look for foreleg specific sex-biased gene expression. We found that 38.1% of mapped genes show sex-biased gene expression in the foreleg of *D. prolongata* with 283 (0.03% of genes) showing at least a $\log_2$ fold change in *D. prolongata* forelegs, including genes previously associated with sex-biased body size and leg development in *D. melanogaster*. We followed up with functional gene knockdown in *D. melanogaster* to verify several candidates and found changes in leg size in *D. melanogaster* that phenocopy at low penetrance, the morphology of *D. prolongata* legs with the gene *grain* (*grn*). We discussed our results in the context of the evolution of sexual dimorphism and the role SBGE may play in it.

Methods

Drosophila Husbandry

The *D. prolongata* population for this experiment were collected in the Sa Pa region of Vietnam in 2018 by Dr. J. Perdigón Ferreira where ~200 field-collected individuals were used to found the population. This population had been maintained for 3 years with overlapping generations at a census number well above the founding 200 individuals to preserve genetic diversity. Flies are maintained at ~50 individual/bottle densities and offspring are redistributed among bottles each generation to minimize inbreeding. Flies are kept at 14:10hr day-night cycles at 18ºC (Ferreira and Lüpold 2022). The D. *carrolli* population for this experiment were collected in Brunei 2003 by Dr. A. Kopp and Dr. O. Barmina. They were maintained at 18ºC for at least two generations prior to this experiment to expand the population. The *Drosophila melanogaster* for this experiment was Canton-S normally maintained at 25ºC but were also kept at 18ºC for two generations before the experimental generation for the current experiment.

Larval staging, sampling and dissection

For sample collection and staging, each species was kept in plastic cages (~30cm^3) at high adult density with a yeast and orange juice paste in petri dishes as egg laying substrate. To stage larvae from their hatching time, plates were removed from cages and the yeast paste was strained through a fine mesh, enabling collecting of eggs and larvae. Eggs (but not larvae) were subsequently placed on grape juice agar plates for collection. From these egg collections, freshly hatched first instar larvae were collected hourly, for three hours. Larvae collected in this manner were considered newly hatched ±30 minutes of that collection time, and were allowed to develop on media (for recipe see 58) until sampling for dissection. Larval dissections to collect leg imaginal discs (larval tissues that develop into adult legs) occurred at two developmental time points during the third larval instar. Dissections occurred 18 hours before onset of larval wandering (6 days + 6 hours after hatch; henceforth early development) and at 8 hours prior to the onset of larval wandering (6 days + 16 hours after hatch; henceforth late development). These development times were chosen based on the findings in (Luecke and Kopp 2019), which demonstrated that sex-differences in cell proliferation in the foreleg imaginal discs (relative to the midleg disc) in *D. prolongata* only become apparent during the third larval instar, with impacts on cell number occurring from the middle to the end of the third larval instar. At the end of larval development *D. prolongata* male foreleg imaginal disc have ~16% more cells than the midleg. This contrasts with a ~2% difference between these discs in female *D. prolongata*, and no substantial differences in *D. carrolli* (Luecke and Kopp 2019). As such our time points were selected such that foreleg imaginal disc were demonstrating some male-bias in *D. prolongata* (late stage, ~2 mitotic cycles prior to the end of larval development), as well as an earlier time point, with minimal sexual dimorphism in cell number had occurred, but where the developmental signals and increase in cell proliferation was still actively occurring. These two time points were chosen to hopefully sample both the initiation of sex-biased development, as well as the SBGE in developing sexually dimorphic tissues.

Sex determination in larvae was done based on the presence of a larger gonad in male larvae (large transparent "disc" visible in the larval fat body) and was done the morning of dissections and larvae were allowed to re-enter food after they were sexed to reduce the stress response of the time required to accurately sex larvae when removed from food. Tissue dissections were done in 1X PBS solution which was kept on ice to keep it cold and reduce induction of transcriptional stress response during dissection. Larvae were dissected by removing the anterior portion of the cuticle, inverting the larval body and extracting the brain with the leg imaginal discs attached. Foreleg and midleg imaginal discs were then removed from the brain and stored separately in 200µL of RNAlater (Thermofisher scientific) at -80ºC until extraction. As leg imaginal discs are very small tissues, ~30 discs from ~15-20 individuals were used for each unique biological sample to ensure sufficient RNA for library preparation. For each species/sex/stage/tissue we aimed to collect 5 independent biological replicates for individual library preparation and sequencing.

We opted for "bulk" tissue samples for RNA sequencing due to the relevant biological differences we were evaluating as well as practicality. As seen in Figure 1, size differences between male and female D. prolongata legs are most pronounced in the femur and tibia, they are visible in most segments of the adult leg. At the stages of larval development that we dissected out the leg imaginal discs, the majority of cells in the tissue are contributing to relevant future leg structures (Kojima 2004; Schubiger et al. 2012) that show sex differences. Additionally, cost of single cell sequencing for an experiment with this many samples remains prohibitive for most labs, and challenges in quantifying expression differences from such data continue. It is also worth noting, that unlike many species in the clade, *D. prolongata* does not have the sex-limited sex comb structure on male forelegs, although they do occur in adult males for *D. melanogaster* and *D. carrolli*. The developmental precursors for the cells contributing to these structures occur at later developmental stages (beginning at the white pre-pupal stage) than we examined in this current study (Barmina and Kopp 2007; Atallah et al. 2009; Tanaka et

al. 2009, 2011) and expression differences between T1 and T2 pupal leg tissues (and across sexes) are modest at this stage (Barmina et al. 2005).

RNA extraction, library preparation and sequencing

RNA extractions were done using Qiagen RNeasy spin columns (CAT# 74106) with DNase treatment (CAT# 79256). To remove RNAlater, we added 1mL of cold 1X PBS to the 1.5mL micro-centrifuge tubes containing 200µL of RNAlater and centrifuged (4°C) to separate discs from solution. Supernatant was pipetted off to leave the imaginal discs and as little RNAlater as possible, and the recommended protocol for the Qiagen RNeasy kit was followed from that point. Following assessment of samples for quantity and purity, Total RNA was shipped to Centre d'expertise et de services, Génome Québec for additional QC (bio-Analyzer), library preparation and sequencing. NEB Stranded mRNA libraries with Nextera adaptor sequences were prepared for each biological sample and sequenced with an illumina NovaSeq6000 to an average of ~50 million clusters/sample (100bp Paired end sequencing). To achieve this sample depth each library was sequenced twice, split between two runs for which all (multiplexed) samples were done jointly on a single lane within a run.

Functional test of candidate genes in D. melanogaster using RNAi knockdown

To explore potential effects of candidate gene expression on leg morphology, we used the bipartite UAS-RNAi Gal4 system in *D. melanogaster* to titrate gene expression in the developing legs. Before crossing our candidate UAS-RNAi to the Pen^{NP6333}-Gal4, we crossed Pen^{NP6333}-Gal4 to a UAS-GFP.NLS strain and observed moderate GFP expression in third instar leg discs, suggesting moderate Gal4 expression during development in our target tissues. Pen^{NP6333}-Gal4 (introgressed into the SAM background marked with w^-), UAS-GFP (BDSC# 4775), *CG30457* UAS-RNAi (BDSC# 62960), *CG13285* UAS-RNAi (BDSC# 53680), *dysfusion* (*dysf)* UAS-RNAi (BDSC# 35010), grain (*grn*) UAS-RNAi (BDSC# 33746),

bric-a-brac1 (*bab1*) UAS-RNAi (BDSC# 57410), *Sox box protein 15* (*Sox15*) UAS-RNAi (BDSC# 57264), and *CG9896* UAS-RNAi (BDSC# 42587) were allowed to lay in vials for 48 hours to keep larval density low. All RNAi lines were part of the *Drosophila* Transgenic RNAi project (TRiP; (Perkins et al. 2015). Unmated females and males were collected from vials for 6 days, and then RNAi males were crossed to Pen^{NP6333}-GAL4 females (all crosses were performed reciprocally). Females in these crosses were allowed to lay eggs for 48 hours, and the flipped into backup vials and allowed to lay for another 48 hours. Vials were reared at both 25ºC, or 28ºC, as the UAS-Gal4 allows for temperature mediated titration of expression.

Offspring from each cross were collected after eclosion and cuticle sclerotization, and stored in 70% ethanol until imaging. Flies were imaged using a Leica MZ7.5 scope using a Leica IC90.E camera. For imaging, flies were dissected in mounting solution (70% glycerol in PBS, with a small amount of phenol as a bacteriostatic). Adult forelegs and midlegs were removed and imaged separately, and thorax was imaged laterally and measured from behind the tip of the scutellum to the notch created from the margin of the humeral callus and the prescutum. Each leg had length and width measured for the femur, as well as thorax length. Images were measured using ImageJ (Rueden et al. 2017).

Bioinformatic and statistical analysis

After initial QC with fastQC (version info and flags for all software available in Table S1), pairs of sequence reads were trimmed of adaptor sequences and low-quality positions using BBduk (Bushnell 2021) and then first and second runs of a given samples were concatenated. Read were then indexed, mapped within each species to genomes and assembled annotations obtained from Luecke et al. (Luecke et al. 2024) for *D. prolongata* (GCA_036346975.1), *D. carrolli* (GCA_018152295.1) and to the *D. melanogaster* assembly version 6.23 (FlyBase.org), reads were counted at the gene level using STAR and multi-mapped reads were not included in counts (Alexander Dobin and Thomas R. Gingeras 2016).

Gene level counts were imported into R (version 4.4.2) for analysis. Principal Component Analysis (PCA) was performed both within and between species for the count data (with gene lists filtered for genes that are present in all three species) using the 'rlog' variance stabilization in DeSeq2. We observed similar patterns when we used smaller subsets of varying genes (from 500 to 2000 in 250 gene increments).

Counts were modelled gene by gene within species in glmmTMB (Brooks et al. 2017) using a generalised linear mixed model with a negative binomial distribution using the following model stricture:

$$gene\ counts \sim (sex + leg + stage)^3 + diag(1|replicate{:}\,stage)$$

Unlike methods such as DESeq2, this allows us to account for uncertainty in our estimates across random effects. As model offsets we used sample specific size factors computed using the DESeq2 (Love et al. 2014) function 'estimateSizeFactors'. Estimates and contrasts were computed using emmeans (Lenth et al. 2018). Our primary focal contrasts were that of sex differences in expression, along with additional interaction contrasts (across species, time or tissue) involving sex differences. Estimated contrasts were regularized using the ashr library (Stephens et al. 2016). Counts modelling was also done using DESeq2 to confirm that model estimates (but not standard errors on estimates) were similar to those from the mixed models (Figure S1; S2).

Multivariate examination of changes in sex-biased gene expression

To explore the relative degrees of similarity in SBGE in foreleg compared to midleg within species, as well as the similarity of SBGE in forelegs between species, we conducted an analysis of vector correlation and ratio of magnitudes of expression. This analysis was done similarly (and using adapted scripts) from both Zinna et al. (Zinna et al. 2018) and Scott et al. (Scott et al. 2022). This allows us to examine where suites of genes differ (as a whole) in their direction or magnitude of SBGE, even if

individual genes are not 'significantly' different. This approach is informative when exploring expression changes for biological variables such as pathways involved with growth (for instance). Vector correlation was calculated as $r_{VC} = \frac{|a \cdot b|}{||a||x||b||}$ where vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ each represent vectors of male-female contrasts with regularization. r_{VC} values closer to 1 suggest similar direction of sex-biased expression changes, while values near 0 suggest little association. For our comparison of vector magnitudes, we calculated the ratio of magnitudes (ℓ^2 norm) of vectors of male-female SBGE contrasts using $\alpha = \frac{||a||}{||b||}$. Where higher α values suggest greater magnitudes of SBGE in a relative to b and an α of 1 would mean that the magnitude of SBGE magnitudes is similar. For this study we had two primary analyses 1) a = species 1 foreleg SBGE and b = species 2 foreleg SBGE, and 2) a = species 1 foreleg SBGE and b = species 1 midleg SBGE. All statistical analyses were done in R version 4.4.2.

Comparison of sexual dimorphism to sex-biased gene expression

We examined the relationship between sexual size dimorphism calculated as (female femur width – male femur width) + (female femur length – male femur length) for both foreleg and midleg separately, in all three species, and sex biased gene expression. We looked at the relationship between this SSD value and the magnitude of SBGE for all genes, as well as for genes with an estimated expression with a $\log_2$ fold change with 95% confidence intervals non-overlapping with the $\log_2$ fold change threshold in *D. prolongata*. Magnitude for SBGs was calculated as ℓ^2-norms of relevant vectors of effect sizes. To assess uncertainty in these estimates, we performed non-parametric bootstrap using resampling with replacement 10000 times and calculating the 95% quantiles of the distribution to generate standard percentile intervals. We also compared these SD values with the number of SBGs that had confidence intervals not overlapping zero after ashr regularization, as well as a comparison to our genes that showed a $\log_2$ fold (confidence on our estimates non-overlapping with +1 or -1) change after ashr regularization.

Molecular evolution of grain

Genomes of *D. prolongata* (GCA_036346975.1), *D. carrolli* (GCA_018152295.1), and *D. rhopaloa* (GCA_018152115.1) were downloaded from NCBI to both explore sequence changes in *grn*, as well as to look for the previously identified transposable element honghaier (Luo et al. 2025) up- or downstream of any of our candidates. The annotated *grn* sequence from *D. melanogaster* was reciprocally blasted to all three genomes to identify *grn* location, and that corresponding sequence was extracted from each genome using bedtools getfasta (Quinlan and Hall 2010) and aligned using MUSCLE (Edgar 2004b,a). This was done with each exon of *grn* individually. To look for changes in predicted enhancer sequences were first filtered for tandem repeats using tandem repeat finder (Benson 1999) using flags recommended for SCRMshaw inputs, and then predictions were made using SCRMshaw (Asma et al. 2024). SCRMshaw was trained on all available Drosophila enhancers found at: https://github.com/HalfonLab/dmel_training_sets. SCRMshaw was run using default settings on the *D. prolongata*, *D. carrolli*, and D. *rhopaloa* genomes, and predicted enhancer outputs were searched for the appearance of *grn* as the nearest downstream gene, using unix grep. To identify downstream targets of *grn* we extracted the nearest genes to identified CHIPseq targets of grain reported by the encode project (ENCSR909QHH), these genes were then used as SBGE vectors for our r_{VC} and α analyses above.

The honghaier transposable element, has been recently shown to contain both *dsx* and *bab1* binding sites as demonstrated by Luo et al. (Luo et al. 2025), and likely influences species specific changes in expression in *D. prolongata*. To search for the honghaier transposable element near genes of interest, we extracted the CDS region using the coordinates in the annotations of *D. prolongata*, *D. carrolli*, and *D. rhopaloa*, including 10000bp up- and downstream using bedtools getfasta, and blasted the honghaier sequence against each of these loci to identify potential insertion sites.

Results

Expression profiles show strong species and sex-specific clustering, with little evidence for clustering based on specific leg imaginal disc, or developmental stage.

Given the profound differences in morphology in *D. prolongata* male forelegs (Figure 1), we wanted to determine if this was reflected in broad scale changes in transcriptional profile in the foreleg imaginal disc of males of this species relative to females, other species, and in comparison to the second leg imaginal disc, whose adult morphology shows little evidence of trait exaggeration. We performed principal components analysis on all samples. There were 8008 genes that we could unambiguously assign orthologs between all three species to potentially include. Using the 1000 most variable genes as features, PC1 accounts for 48% of the variance in gene expression and clearly separates all three species. PC2 separates *D. carrolli* out from *D. melanogaster* and *D. prolongata* and accounts for 25% of the variance (Figure 2). Sex differences are associated with a combination of PC3 (5%) and PC4 (2%) (Figure 2). We further evaluated the transcriptome-wide expression profile within each species (Figure 3, Figure S3). While the sex and developmental stage and sex effects appear to group together in the PCA, we see only modest evidence of broad scale differences between the first and second leg imaginal discs, even in *D. prolongata* males.

Sex biased gene expression in D. prolongata forelegs shows species specific patterns of bias

Sex-biased genes were more numerous in *D. prolongata* than either *D. carrolli* or *D. melanogaster* (Table 1; Figure 4; S4; S5; Supplemental File 1). When these gene lists are filtered for $\log_2$ fold differentially expressed genes between the sexes in the foreleg in each sex but are not ex-biased in the midleg, *D. prolongata* shows the highest number of both male (65 compared to 34 *D. carrolli* and 17 *D. melanogaster*), and female (58 compared to 7 *D. carrolli* and 9 *D. melanogaster*) sex-biased genes.

Based on the genes identified as showing SBGE in *D. prolongata*, we wanted to evaluate to what degree the number and overall magnitude of expression differences in the developing leg imaginal discs is associated with the degree of phenotypic sexual size dimorphism (SSD) in the adult legs (Figure 4). We did not observe a strong relationship between magnitude of expression differences for the SBG and the most sexually dimorphic structures. For instance, male *D. prolongata* femurs, the trait with the greatest SSD, have similar magnitudes of SBGE as most other tissues with the highest magnitudes being in *D. carrolli* in both sexes late in development (Figure 4). The number of genes identified as showing sex-biased expression SBGs does show a positive relationship, with heightened male phenotypic SSD associated with the highest absolute number of differentially expressed genes, and this pattern holds for both the total number of SBGs (Figure 4) and the number of genes with a minimum $\log_2$-fold difference between the sexes (Figure S6).

The differentially expressed genes of primary interest for our study were genes showing sex-biased gene expression only in the foreleg imaginal discs of *D. prolongata*, and without a parallel change either in the midleg of *D. prolongata* or in either *D. carrolli* or *D. melanogaster* leg imaginal discs. Using custom interaction contrasts, we identified a set of genes with these attributes (Supplemental File 1, Figure 5). We observed sets of genes that show both male and female biased directions of expression differences in the foreleg imaginal discs of *D. prolongata*, including several genes whose orthologs influence organ specific size and shape.

RNAi knockdowns of a subset of our candidates shows changes to femur width and length as well as a qualitatively D. prolongata-like phenotype in D. melanogaster femurs

We identified all of the $\log_2$ fold SBGs in *D. prolongata* forelegs that did not show a $\log_2$ fold expression change in midlegs, and we manually curated via examination of patterns of SBGE between species, leg, and stage (Supplemental file 2) and used FlyBase to further refine these candidate genes

based on their known or presumed roles in organ size or sex-specific effects, to identify a subset of genes for further functional work. We knocked-down expression of the orthologs of these genes in the leg imaginal discs of *D. melanogaster* during larval development and examined changes in lengths and widths in adult legs. We chose to knock-down grn, *dysf*, *CG30457*, *CG13285*, *sox15*, *bab1*, and *otp* based on their SBGE patterns (Figure 5). Excitingly, we observed a small number (two) of *grn* knockdown individuals with highly enlarged femur widths, resembling the femurs of *D. prolongata* (Figure 6).

We focus on 28ºC treatment as expression knock-down in these treatments is stronger, however we do see increased femur width in 25ºC *grn* knock-downs which was primarily due to a small number of extreme measures (Figure 6). In males at 28ºC, forefemur length decreased in knockdowns of *bab1*, *dysf*, *grn*, and *sox15* (Table S2; Figure S7). In females at 28ºC forefemur length decreased in *bab1*, *grn*, *otp*, and *sox15* (Table S2; Figure S7). Forefemur width in 28ºC males decreased in *dysf*, and females at 28ºC forefemur width increased in *bab1* and *sox15* and decreased in *CG13285* (Table S2; Figure S7).

We next wanted to look at leg size changes while accounting for body (thorax) size. We did observe some reduction in thorax length relative to controls in the UAS-RNAi offspring of *bric-a-brac1* (*bab1*), and *orthopedia* (*otp*) at 28ºC in males, as well as in *bab1*, *CG13285*, *dysfusion* (*dysf*), and *sox box protein 15* (*sox15*) at 28ºC in females (Table S2; Figure S7). Adjusting for body size, forefemur length decreased in both sexes at 28ºC in *sox15* and *grn* knockdowns, and in just females in *bab1* and *otp* (Figure 7; Table S3). Forefemur width was reduced in males and increased females when *dysf* was knocked down and increased in both sexes when *sox15* was knocked down (Figure 7; Table S4).

Grain is well conserved in the closest relatives of D. prolongata

Given the intriguing observations with *grain*, a GATA transcription factor involved with growth and metabolism (Brown and Castelli-Gair Hombría 2000; Kokki et al. 2021), we examined the gene for any

possible changes suggesting a change in protein function of gene regulation. Exon alignment of *grn* between *D. prolongata*, *D. carrolli*, *D. rhopaloa,* and *D. melanogaster* show high levels of conservation. The exon 4 GATA binding site had 100% amino acid conservation, and the exon 5 GATA binding site had a few synonymous substitutions present in *D. rhopaloa*, *D. carrolli*, and *D. prolongata* compared to *D. melanogaster*. Changes outside of the binding-domain were also uncommon between *D. prolongata* and *D. carrolli* (amino acid changes: 1/169 exon 1, 0/95 exon 2, 4/276 exon 3, 0/47 exon 4, 0/47 exon 5, 0/62 exon 6, 0/62 exon 7). Cis-regulatory module prediction from SCRIMshaw showed no modules that were unique to *D. prolongata* predicted to regulate *grn*. When predicted modules were searched for the gene identifier *grn*, all predicted modules were identified in D. *carrolli* or *D. rhopaloa* as well.

The transposable element honghaier occurs up or down stream in most genes of interest in multiple species, with seven being exclusive to D. prolongata

Looking 10000bp up- and downstream, we identified 7 of our genes of interest that had a blast hit for the honghaier transposable element exclusively in *D. prolongata*. These genes were: *smydA*, *rst*, *CG11378*, *CG14075*, *CG14356*, *CG32564*, and *CG9411*. These genes do all show interesting SBGE (Supplemental file 2), but do not stand out relative to our other genes of interest.

Signalling pathways and candidate sex-biased genes appear to be expressed in similar direction and magnitude between all three species

As the changes in expression of many genes during development and growth are likely contributing, we examined changes in both direction and ratio of magnitude for vectors of sex-biased expression (i.e. contrast of expression differences between males and females) as they changed across tissues and species. Changes in the overall ratio of magnitudes of expression for SBGE between fore- and midleg imaginal discs within species is highest in *D. prolongata* but only slightly so compared to *D.*

melanogaster; early in development the change in magnitude is about 2.17X greater in the foreleg compared with the midleg, although within the range of our 95% intervals for vector correlations of randomly drawn genes (Figure 8). During the later developmental time-period the relative increase in magnitude remains (~1.62X), exceeding the 95% threshold of random vector magnitudes (Figure 8). When we separate our list of $\log_2$ fold change SBGs into male- and female-biased genes we observe negative correlations in the direction of male-biased genes early in development (r = -0.39), while the direction becomes positively correlated later (r = 0.73), which is a pattern not observed in the other species (Figure S8). We do not see this in female-biased gene vector correlations in any of three species (Figure S9).

Between species, vectors for sex-biased genes with a minimum of $\log_2$-fold change are only weakly to moderately correlated with each other (*D. prolongata* / *D. carrolli* early r_{VC} = 0.13, late r_{VC} = 0.42; *D. prolongata* / *D. melanogaster* early r_{VC} = -0.15 late r_{VC} = -0.36). The correlations (negative and positive) are somewhat greater in absolute value during the later developmental stage, although with the exception of the comparison of *D. prolongata* and *D. melanogaster* SBGE in forelegs, are not particularly extreme in comparison to random sets of genes. However, the change in the magnitude of the vector of sex-biased expression of candidate genes in the foreleg appears to be higher in *D. prolongata* than in both *D. carrolli* (11.23X greater early, and 6.36X greater late) and *D. melanogaster (7.36X greater early, and 8.92X greater late)*, while the expression magnitudes in forelegs are similar between *D. melanogaster* and *D. carrolli* (Figure 9). When we explore male- and female-biased genes separately between species, male-biased genes show a higher magnitude of expression early for both D. carrolli (11.34X greater), and *D. melanogaster* (23.31X greater) but not later in development (Figure S10). Female-biased genes also show a difference in magnitude when *D. prolongata* is compared to *D. carrolli* both early (11.01X) and late (8.92X) in development, and when *D. prolongata* is compared to *D. melanogaster* late in development (9.52X). Female-biased genes also showed a negative correlation in

the foreleg early in development compared to *D. carrolli* (r_{VC}= -0.22), and late in development compared to *D. melanogaster* (r_{VC} = -0.45) relative to expectations from randomly drawn genes (Figure S11). For *D. carrolli* female-biased genes, this represents a change in direction of the correlation, with later developmental genes becoming positively correlated with *D. prolongata* (r_{VC} = 0.53).

To verify that this magnitude difference is not explained by the fact that we are selecting genes based on their expression differences in *D. prolongata*, we extracted $\log_2$ fold differentially expressed *D. melanogaster* genes with the same thresholds used to identify our *D. prolongata* candidates and we repeated the same analysis with these *D. melanogaster* $\log_2$-fold change sex-biased 'candidates'. Using the same cutoffs as our *D. prolongata* genes of interest, we get many fewer genes with a $\log_2$-fold change in expression (17 genes in *D. melanogaster*, 121 genes in *D. prolongata*), which showed magnitudes closer to or less than one between the species (*D. prolongata* / *D. carrolli* early=1.95X greater, late=1.1X greater; *D. prolongata* / *D. melanogaster* early=0.16X, late=0.22X; Figure S12).

We also examined changes among genes within key signaling pathways involved with tissue growth within and between species. Since the insulin signalling pathway (InS) has previously been implicated with exaggerated trait growth (Emlen et al. 2006, 2012), we specifically explored this pathway both between tissues within species, and between species in the foreleg. Within species, vector correlation is similar between foreleg and midleg both early (r_{VC} = 0.63) and late (r_{VC} = 0.74) in *D. prolongata*, and magnitudes are within our 95% random vector intervals, and generally close to 1. In *D. prolongata* InS magnitude was 1.44X greater in forelegs compared with midlegs early in development and 0.74X later in development, which was similar or lower than *D. carrolli* and *D. melanogaster,* and all InS magnitude comparisons were within our distribution of random gene magnitudes (Figure S13). For other signalling pathways, vector correlations between foreleg and midleg within species were consistently high and outside our 95% interval of random gene draws, and the ratios of magnitudes

between foreleg and midleg imaginal discs were generally close to 1, meaning magnitude of SBGE between legs was approximately equivalent (Figure S13-S18).

Comparing SBGE in the InS pathway in forelegs between species, we also see generally positive vector correlations between *D. prolongata* and *D. carrolli* (early r_{VC} = 0.11; late r_{VC} = 0.65) and between *D. prolongata* and *D. melanogaster* (early r_{VC} = 0.05; late r_{VC} = 0.16), as well as magnitudes around 1 between *D. prolongata* and both *D. carrolli* (early=1.1X; late=0.85X) and *D. melanogaster* (early=1.52X; late=1.99X; Figure 10). For all other signalling pathways we explored, SBGE r_{VC} were positively correlated between species, and magnitude was around 1 (Figure S19-S24).

As *grn* is a transcription factor, we also explored its putative targets (defined as the nearest gene to reported *grn* binding sites reported by the encode project ENCSR909QHH). Changes in the magnitude of SBGE for putative grn targets in *D. carrolli* is lower than *D. melanogaster* both early (0.67X) and late (0.78X) in development. *D. prolongata* generally had higher magnitudes (1.63X greater than early *D. carrolli*; 1.06X greater than late *D. carrolli*; 2.43X greater than early *D. melanogaster*; 1.34X greater in late *D. carrolli*) but only exceeded our 95% random gene vectors in early *D. melanogaster* (Figure S25).

Discussion

The evolution of sexual dimorphism and its relationship to changes in SBGE has been a long-standing question (Parsch and Ellegren 2013; Ingleby et al. 2015; Grath and Parsch 2016). A simple good genes model would suggest a potentially small number of genes under selection in sexually selected traits, whereas various polygenic models including the genic capture model would suggest a substantial fraction of the genome mediating evolutionary changes to these traits (Rowe and Houle 1996; Pomiankowski and Møller 1997). Our results are not consistent with many genes with small expression changes contributing to a large-scale change in sex-biased transcriptional profile in the foreleg imaginal

disc of *D. prolongata* males (Figure 2, 3). Instead, we observed a modest number of genes that show SBGE in *D. prolongata* sexually exaggerated forelegs (123 DEGs out of 19330 genes in *D. prolongata*). In other studies where sexually dimorphic tissues were sampled (as opposed to whole body sampling), similarly small proportions of genes are differentially expressed (Wilkinson et al. 2013; Zinna et al. 2018; Toubiana et al. 2021). We do find a positive relationship between the number of DEGs and the degree of phenotypic dimorphism between species, but no strong relationship between magnitude of expression differences and phenotypic dimorphism in the adult (Figure 5). This relationship between number of genes and dimorphism, but such a small proportion of the genome being differentially expressed in a sex and tissue specific manner may suggest that recruitment of regulatory genes plays a role in dimorphic evolution, as co-option of regulatory genes may contribute a greater phenotypic change with fewer genes. We do find evidence that changes in a relatively small number of transcription factors (TF) may be at least partially responsible for the maintenance of the sexually exaggerated forelegs in *D. prolongata*. Of particular interest, we identify the GATA family TF *grn* as a potentially large effect gene showing a low penetrance phenotype in *D. melanogaster* legs recapitulating key aspects of the phenotype in *D. prolongata*. Little is known about developmental pathways associated with the role of *grn* in modulating adult morphology, as many mutations are homozygous lethal. However previous work has demonstrated that tissue-specific loss of function of *grn* results in shorter, wider femurs which resemble *D. prolongata* (Brown and Castelli-Gair Hombría 2000), and our functional results of *grn* knockdowns support this finding. In other work, *grn* has been demonstrated to play key roles in spiracle formation (Brown and Castelli-Gair Hombría 2000), sugar metabolism (Kokki et al. 2021), and neuron specification (Garces and Thor 2006), suggesting that as a transcription factor, *grn* plays a broad role in a number of developmental processes. Within the closest species to *D. prolongata*, *D. carrolli* and *D. rhopaloa*, we find no major changes in the protein sequence of *grn*; we find 100% conservation in the GATA binding domains between all three species and *D. melanogaster*;

and we find that its location of expression in leg imaginal discs is conserved as well, which we believe supports the idea that the level of SBGE in *grn* is responsible for its morphological impact, rather than changes to its function. We also identify strong female bias in forelegs for *dysf* and *sox15* which both show phenotypic changes to femur width or length when knocked down in *D. melanogaster* (Figure 7). *Grn*, *dysf*, and *sox15* have all been demonstrated to be regulated via notch signalling (Garces and Thor 2006; Miller et al. 2009; Córdoba and Estella 2020). The notch signalling pathway is known to influence leg morphology and joint formation in *D. melanogaster* (Rauskolb and Irvine 1999). Our results don't demonstrate large changes to the magnitude of Notch SBGE in *D. prolongata* forelegs relative to the forelegs of *D. carrolli* and *D. melanogaster* (Figure S19). We also do not see notable changes in the expression magnitude of potential targets of *grn* in the forelegs of *D. prolongata* compared to *D. carrolli* and *D. melanogaster* (Figure S25).

As well as strong morphological differences, *D. prolongata* demonstrates colouration differences between the sexes that are unique to this species within the *Drosophila* species group (Setoguchi et al. 2014). Previous work has implicated the two *bric-a-brac* paralogs (*bab1* and *bab2*) as being responsible for sex-specific pigmentation in the *D. melanogaster* thorax (Kopp et al. 2000; Williams et al. 2008). These studies have further found that *bab1/2* is under control of the female isoform of *dsx* (Kopp et al. 2000), as well as responsible for leg development and tarsal specification (Godt et al. 1993). *dsx* has also been implicated as a regulator of sexually dimorphic tibia growth in the gazelle dung beetle *Digitonthophagus gazella (Rohner et al. 2021)*. We find female biased expression in both *dsx* and *bab1* greater in *D. prolongata* than in *D. melanogaster* or *D. carrolli* and knockdowns of *bab1* in *D. melanogaster* result in sex-specific changes in both femur width and length in *D. melanogaster* (Figure 7), suggesting a potential role of the *dsx*/*bab1* gene cascade in the legs of *D. prolongata*. We caveat this result by the fact that this requires further functional work such as immunofluorescence.

We also wanted to look at broad changes in gene pathway correlation for both direction and magnitude of sets of SBGs. We found that our focal set of genes demonstrated direction of SBGE between the sexes that was less similar than vectors of equal numbers of randomly drawn genes in *D. prolongata*, but not the other two species (Figure 8; 9). We also observed a difference in magnitude for our DEGs, with a much higher magnitude of SBGE in foreleg imaginal discs of *D. prolongata* relative to *D. carrolli* and *D. melanogaster* (Figure 9). This again suggests that the extent of SBGE in the forelegs may be important for sexually dimorphic morphology in *D. prolongata*. Despite these results, the overall magnitude of SBGE does not seem to relate to the degree of morphological sexual dimorphism, although the absolute number of genes does (Figure 5).

In other systems with sexually exaggerated traits, growth pathway differences have been associated with sexual dimorphism; specifically in insulin signalling and EGFR signalling (reviewed in 90). We explored changes in both insulin and EGFR direction and magnitude between legs within species, and in forelegs between species, and found surprisingly little difference (Figure 10; S15; S20). Our sampling occurred shortly before, and shortly after the number of cells present in *D. prolongata* forelegs becomes sexually dimorphic (Luecke and Kopp 2019). However, whether this particular period of growth showed sex specific responses to insulin signaling is not currently known. Another potential reason for the lack of difference in growth signalling may be the specific traits we examined. Many of the studies implicating growth signalling look at novel traits (horns), whereas *D. prolongata* forelegs are existing traits that have been exaggerated, which may influence the extent to which important growth signalling pathways can be altered.

In other systems where male-biased dimorphism exists, it is often observed that the number of male-biased genes is greater than the number of female-biased genes (Wilkinson et al. 2013; Zinna et al. 2018). The exaggerated forelegs in *D. prolongata* however show an increase in female biased genes (Table 1). A similar pattern to this has been reported in the water strider *Microvelia longipes*, in which

the exaggerated third leg shows an increase in the number of female-biased genes relative to the mid- and foreleg which are less male-biased in length (Toubiana et al. 2021). Previous work in *D. prolongata* looking at adult pheromone profiles, found increased male-biased expression in sexually dimorphic cuticular hydrocarbon producing cells. However they do identify a number of genes that are highly female biases in this sexually dimorphic trait as well (Luo et al. 2025). The genes that we identified as male-biased in the foreleg of *D. prolongata* but not in the midleg, and also not in the other two species are mostly unnamed genes with largely unknown roles (Supplemental File 1). Although the number of female biased DEGs is higher in *D. prolongata*, we find little change to vector direction correlation or ratio if magnitudes of expression between the legs in *D. prolongata*, while we find a reversal of direction of vector correlation for SBGE in male-biased genes (Figure S8).

This work adds to our growing knowledge of how sex-specific expression of a largely shared genome can modulate sexual dimorphism. *Drosophila prolongata* is an ideal model for the evolution of sexual dimorphism due to the ease of rearing large samples, the phylogenetic novelty of the trait, and the genetic tools available in the closely related *D. melanogaster*. For these reasons, *D. prolongata* may be a particularly critical species to explore many outstanding questions in sexual selection.

References

Alexander Dobin and Thomas R. Gingeras. 2016. Optimizing RNA-Seq Mapping with STAR. Methods Mol. Biol. 1415:245–262. https://doi.org/10.1007/978-1-4939-3572-7_13

Asma, H., L. Liu, and M. S. Halfon. 2024. SCRMshaw: Supervised cis-regulatory module prediction for insect genomes. PLoS One 19:e0311752. https://doi.org/10.1371/journal.pone.0311752

Atallah, J., N. H. Liu, P. Dennis, A. Hon, D. Godt, and E. W. Larsen. 2009. Cell dynamics and developmental bias in the ontogeny of a complex sexually dimorphic trait in *Drosophila melanogaster*. Evol. Dev. 11:191–204. https://doi.org/10.1111/j.1525-142X.2009.00319.x

Baker, B. S., and K. A. Ridge. 1980. Sex and the single cell. I. On the action of major loci affecting sex determination in drosophila melanogaster. Genetics 94:383–423. https://doi.org/10.1093/genetics/94.2.383

Barmina, O., M. Gonzalo, L. M. McIntyre, and A. Kopp. 2005. Sex- and segment-specific modulation of gene expression profiles in *Drosophila*. Dev. Biol. 288:528–544. https://doi.org/10.1016/j.ydbio.2005.09.052

Barmina, O., and A. Kopp. 2007. Sex-specific expression of a HOX gene associated with rapid morphological evolution. Dev. Biol. 311:277–286. https://doi.org/10.1016/j.ydbio.2007.07.030

Benson, G. 1999. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 27:573–580. https://doi.org/10.1093/nar/27.2.573

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. 9:378–400.

Brown, S., and J. Castelli-Gair Hombría. 2000. *Drosophila grain* encodes a GATA transcription factor required for cell rearrangement during morphogenesis. Development 127:4867–4876. https://doi.org/10.1242/dev.127.22.4867

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

Casasa, S., and A. P. Moczek. 2018. Insulin signalling's role in mediating tissue-specific nutritional plasticity and robustness in the horn-polyphenic beetle Onthophagus taurus. Proc. R. Soc. B Biol. Sci. 285:20181631. https://doi.org/10.1098/rspb.2018.1631

Córdoba, S., and C. Estella. 2020. Role of Notch Signaling in Leg Development in *Drosophila melanogaster*. Pp. 103–127 *in* J. Reichrath and S. Reichrath, eds. Notch Signaling in Embryology and Cancer: Notch Signaling in Embryology. https://doi.org/10.1007/978-3-030-34436-8_7

Cox, R. M., C. L. Cox, J. W. McGlothlin, D. C. Card, A. L. Andrew, and T. A. Castoe. 2017. Hormonally Mediated Increases in Sex-Biased Gene Expression Accompany the Breakdown of Between-Sex Genetic Correlations in a Sexually Dimorphic Lizard. Am. Nat. 189:315–332. https://doi.org/10.1086/690105

Crumière, A. J. J., and A. Khila. 2019. Hox genes mediate the escalation of sexually antagonistic traits in water striders. Biol. Lett. 15:20180720. https://doi.org/10.1098/rsbl.2018.0720

Darwin, C. 1871. The Descent of Man, and Selection in Relation to Sex.

Dobzhansky, T. 1931. Interaction between female and male parts in gynandromorphs of *Drosophila simulans*. Wilhelm Roux Arch. Für Entwicklungsmechanik Org. 123:719–746. https://doi.org/10.1007/BF01380651

Edgar, R. C. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5:113. https://doi.org/10.1186/1471-2105-5-113

Ellegren, H., and J. Parsch. 2007. The evolution of sex-biased genes and sex-biased gene expression. Nat. Rev. Genet. 8:689–698. https://doi.org/10.1038/nrg2167

Emlen, D. J. 2008. The Evolution of Animal Weapons. Annu. Rev. Ecol. Evol. Syst. 39:387–413. https://doi.org/10.1146/annurev.ecolsys.39.110707.173502

Emlen, D. J., Q. Szafran, L. S. Corley, and I. Dworkin. 2006. Insulin signaling and limb-patterning: candidate pathways for the origin and evolutionary diversification of beetle 'horns.' Heredity 97:179–191. https://doi.org/10.1038/sj.hdy.6800868

Emlen, D. J., I. A. Warren, A. Johns, I. Dworkin, and L. C. Lavine. 2012. A Mechanism of Extreme Growth and Reliable Signaling in Sexually Selected Ornaments and Weapons. Science 337:860–864. https://doi.org/10.1126/science.1224286

Ferreira, J. P., and S. Lüpold. 2022. Condition- and context-dependent alternative reproductive tactic in *Drosophila prolongata*. Behav. Ecol. *33*(1), 213-221. https://doi.org/10.1093/beheco/arab127

Fisher, R. A. 1930. The genetical theory of natural selection.

Garces, A., and S. Thor. 2006. Specification of Drosophila aCC motoneuron identity by a genetic cascade involving even-skipped, grain and zfh1. Development 133:1445–1455. https://doi.org/10.1242/dev.02321

Godt, D., J.-L. Couderc, S. E. Cramton, and F. A. Laski. 1993. Pattern formation in the limbs of *Drosophila* : *bric à brac* is expressed in both a gradient and a wave-like pattern and is required for specification and proper segmentation of the tarsus. Development 119:799–812. https://doi.org/10.1242/dev.119.3.799

Gompel, N., and A. Kopp. 2018. *Drosophila* (Sophophora) *carrolli* n. sp., a new species from Brunei, closely related to *Drosophila* (Sophophora) *rhopaloa* (Diptera: Drosophilidae). Zootaxa 4434:502–510. https://doi.org/10.11646/zootaxa.4434.3.6

Gotoh, H., H. Miyakawa, A. Ishikawa, Y. Ishikawa, Y. Sugime, D. J. Emlen, L. C. Lavine, and T. Miura. 2014. Developmental Link between Sex and Nutrition; *doublesex* Regulates Sex-Specific Mandible Growth via Juvenile Hormone Signaling in Stag Beetles. PLoS Genet. 10:e1004098. https://doi.org/10.1371/journal.pgen.1004098

Gotoh, H., R. A. Zinna, I. Warren, M. DeNieu, T. Niimi, I. Dworkin, D. J. Emlen, T. Miura, and L. C. Lavine. 2016. Identification and functional analyses of sex determination genes in the sexually dimorphic stag beetle *Cyclommatus metallifer*. BMC Genomics 17:250. https://doi.org/10.1186/s12864-016-2522-8

Gowen, J. W., and S.-T. C. Fung. 1957. Determination of sex through genes in a major sex locus in Drosophila Melanogaster. Heredity 11:397–402. https://doi.org/10.1038/hdy.1957.31

Grath, S., and J. Parsch. 2016. Sex-Biased Gene Expression. Annu. Rev. Genet. 50:29–44. https://doi.org/10.1146/annurev-genet-120215-035429

Hérault, C., T. Pihl, and B. Hudry. 2024. Cellular sex throughout the organism underlies somatic sexual differentiation. Nat. Commun. 15:6925. https://doi.org/10.1038/s41467-024-51228-6

Ingleby, F. C., I. Flis, and E. H. Morrow. 2015. Sex-Biased Gene Expression and Sexual Conflict throughout Development. *Cold Spring Harbor perspectives in biology*, *7*(1), a017632. https://doi.org/10.1101/cshperspect.a01763

J Patterson. 1938. Aberrant Forms in Drosophila and Sex Differentiation. Am. Nat. 72:193–206. http://www.jstor.org/stable/2457550.

Jahner, J. P., L. K. Lucas, J. S. Wilson, and M. L. Forister. 2015. Morphological Outcomes of Gynandromorphism in Lycaeides Butterflies (Lepidoptera: Lycaenidae). J. Insect Sci. 15:38. https://doi.org/10.1093/jisesa/iev020

Khila, A., E. Abouheif, and L. Rowe. 2009. Evolution of a Novel Appendage Ground Plan in Water Striders Is Driven by Changes in the Hox Gene Ultrabithorax. PLoS Genet. 5:e1000583. https://doi.org/10.1371/journal.pgen.1000583

Kijimoto, T., A. P. Moczek, and J. Andrews. 2012. Diversification of *doublesex* function underlies morph-, sex-, and species-specific development of beetle horns. Proc. Natl. Acad. Sci. 109:20526–20531. https://doi.org/10.1073/pnas.1118589109

Kim, B. Y., J. R. Wang, D. E. Miller, O. Barmina, E. Delaney, A. Thompson, A. A. Comeault, D. Peede, E. R. D'Agostino, J. Pelaez, J. M. Aguilar, D. Haji, T. Matsunaga, E. E. Armstrong, M. Zych, Y. Ogawa, M. Stamenković-Radak, M. Jelić, M. S. Veselinović, M. Tanasković, P. Erić, J.-J. Gao, T. K. Katoh, M. J. Toda, H. Watabe, M. Watada, J. S. Davis, L. C. Moyle, G. Manoli, E. Bertolini, V. Košťál, R. S. Hawley, A. Takahashi, C. D. Jones, D. K. Price, N. Whiteman, A. Kopp, D. R. Matute, and D. A. Petrov. 2021. Highly contiguous assemblies of 101 drosophilid genomes. eLife 10:e66405. https://doi.org/10.7554/eLife.66405

Kirkpatrick, M., and M. J. Ryan. 1991. The evolution of mating preferences and the paradox of the lek. Nature 350:33–38. https://doi.org/10.1038/350033a0

Kodric-Brown, A., R. M. Sibly, and J. H. Brown. 2006. The allometry of ornaments and weapons. Proc. Natl. Acad. Sci. 103:8733–8738. https://doi.org/10.1073/pnas.0602994103

Kojima, T. 2004. The mechanism of Drosophila leg development along the proximodistal axis. Dev. Growth Differ. 46:115–129. https://doi.org/10.1111/j.1440-169X.2004.00735.x

Kokki, K., N. Lamichane, A. I. Nieminen, H. Ruhanen, J. Morikka, M. Robciuc, B. M. Rovenko, E. Havula, R. Käkelä, and V. Hietakangas. 2021. Metabolic gene regulation by Drosophila GATA transcription factor Grain. PLoS Genet. 17:e1009855. https://doi.org/10.1371/journal.pgen.1009855

Kopp, A. 2006. Basal relationships in the *Drosophila melanogaster* species group. Mol. Phylogenet. Evol. 39:787–798. https://doi.org/10.1016/j.ympev.2006.01.029

Kopp, A., I. Duncan, and S. B. Carroll. 2000. Genetic control and evolution of sexually dimorphic characters in *Drosophila*. Nature 408:553–559. https://doi.org/10.1038/35046017

Ledón-Rettig, C. C., E. E. Zattara, and A. P. Moczek. 2017. Asymmetric interactions between doublesex and tissue- and sex-specific target genes mediate sexual dimorphism in beetles. Nat. Commun. 8:14593. https://doi.org/10.1038/ncomms14593

Lenth, R., H. Singmann, J. Love, and P. Buerkner. 2018. emmeans: Estimated marginal means, aka Least-Squares Means.

Love, M. I., W. Huber, and S. Anders. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15:550. https://doi.org/10.1186/s13059-014-0550-8

Luecke, D., Y. Luo, H. Krzystek, C. Jones, and A. Kopp. 2024. Highly Contiguous Genome Assembly of Drosophila prolongata - a Model for Evolution of Sexual Dimorphism and Male-specific Innovations. *G3 Genes|Genomes|Genetics*, 14:10, jkae155. https://doi.org/10.1093/g3journal/jkae155

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. https://doi.org/10.1111/evo.13847

Luo, Y., A. Takau, J. Li, T. Fan, B. R. Hopkins, Y. Le, S. R. Ramirez, T. Matsuo, and A. Kopp. 2025. Regulatory changes in the fatty acid elongase eloF underlie the evolution of sex-specific pheromone profiles in *Drosophila prolongata*. BMC Biol. 23:117. https://doi.org/10.1186/s12915-025-02220-z

Mank, J. E. 2017. The transcriptional architecture of phenotypic dimorphism. Nat. Ecol. Evol. 1:0006. https://doi.org/10.1038/s41559-016-0006

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. https://doi.org/10.1534/genetics.116.192260

McCullough, E. L., C. W. Miller, and D. J. Emlen. 2016. Why Sexually Selected Weapons Are Not Ornaments. Trends Ecol. Evol. 31:742–751. https://doi.org/10.1016/j.tree.2016.07.004

McDonald, J. M. C., P. Nabili, L. Thorsen, S. Jeon, and A. W. Shingleton. 2021. Sex-specific plasticity and the nutritional geometry of insulin-signaling gene expression in *Drosophila melanogaster*. EvoDevo 12:6. https://doi.org/10.1186/s13227-021-00175-0

Miller, S. W., T. Avidor-Reiss, A. Polyanovsky, and J. W. Posakony. 2009. Complex interplay of three transcription factors in controlling the tormogen differentiation program of *Drosophila* mechanoreceptors. Dev. Biol. 329:386–399. https://doi.org/10.1016/j.ydbio.2009.02.009

Millington, J. W., G. P. Brownrigg, P. J. Basner-Collins, Z. Sun, and E. J. Rideout. 2021a. Genetic manipulation of insulin/insulin-like growth factor signaling pathway activity has sex-biased effects on Drosophila body size. G3 Genes|Genomes|Genetics 11:jkaa067. https://doi.org/10.1093/g3journal/jkaa067

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. 2021b. Female-biased upregulation of insulin pathway activity mediates the sex difference in Drosophila body size plasticity. eLife 10:e58341. https://doi.org/10.7554/eLife.58341

Moczek, A. P., and D. J. Rose. 2009. Differential recruitment of limb patterning genes during development and diversification of beetle horns. Proc. Natl. Acad. Sci. 106:8992–8997. https://doi.org/10.1073/pnas.0809668106

Okada, Y., M. Katsuki, N. Okamoto, H. Fujioka, and K. Okada. 2019. A specific type of insulin-like peptide regulates the conditional growth of a beetle weapon. PLoS Biol. 17:e3000541. https://doi.org/10.1371/journal.pbio.3000541

Parsch, J., and H. Ellegren. 2013. The evolutionary causes and consequences of sex-biased gene expression. Nat. Rev. Genet. 14:83–87. https://doi.org/10.1038/nrg3376

Perkins, L. A., L. Holderbaum, R. Tao, Y. Hu, R. Sopko, K. McCall, D. Yang-Zhou, I. Flockhart, R. Binari, H.-S. Shim, A. Miller, A. Housden, M. Foos, S. Randkelv, C. Kelley, P. Namgyal, C.

Villalta, L.-P. Liu, X. Jiang, Q. Huan-Huan, X. Wang, A. Fujiyama, A. Toyoda, K. Ayers, A. Blum, B. Czech, R. Neumuller, D. Yan, A. Cavallaro, K. Hibbard, D. Hall, L. Cooley, G. J. Hannon, R. Lehmann, A. Parks, S. E. Mohr, R. Ueda, S. Kondo, J.-Q. Ni, and N. Perrimon. 2015. The Transgenic RNAi Project at Harvard Medical School: Resources and Validation. Genetics 201:843–852. https://doi.org/10.1534/genetics.115.180208

Perry, J. C., P. W. Harrison, and J. E. Mank. 2014. The Ontogeny and Evolution of Sex-Biased Gene Expression in *Drosophila melanogaster*. Mol. Biol. Evol. 31:1206–1219. https://doi.org/10.1093/molbev/msu072

Pomiankowski, A., and A. P. Møller. 1997. A resolution of the lek paradox. Proc. R. Soc. Lond. B Biol. Sci. 260:21–29. https://doi.org/10.1098/rspb.1995.0054

Quinlan, A. R., and I. M. Hall. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841–842. https://doi.org/10.1093/bioinformatics/btq033

Rauskolb, C., and K. D. Irvine. 1999. Notch-Mediated Segmentation and Growth Control of the *Drosophila* Leg. Dev. Biol. 210:339–350. https://doi.org/10.1006/dbio.1999.9273

Refki, P. N., D. Armisén, A. J. J. Crumière, S. Viala, and A. Khila. 2014. Emergence of tissue sensitivity to Hox protein levels underlies the evolution of an adaptive morphological trait. Dev. Biol. 392:441–453. https://doi.org/10.1016/j.ydbio.2014.05.021

Rice, G. R., O. Barmina, D. Luecke, K. Hu, M. Arbeitman, and A. Kopp. 2019. Modular tissue-specific regulation of *doublesex* underpins sexually dimorphic development in *Drosophila*. Development 146:dev178285. https://doi.org/10.1242/dev.178285

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. https://doi.org/10.1371/journal.pgen.1005683

Rohner, P. T., and W. U. Blanckenhorn. 2018. A Comparative Study of the Role of Sex-Specific Condition Dependence in the Evolution of Sexually Dimorphic Traits. Am. Nat. 192:E202–E215. http://orcid.org/0000-0002-9840-1050

Rohner, P. T., S. Casasa, and A. P. Moczek. 2023. Assessing the evolutionary lability of insulin signalling in the regulation of nutritional plasticity across traits and species of horned dung beetles. J. Evol. Biol. 36:1641–1648. https://doi.org/10.1111/jeb.14240

Rohner, P. T., D. M. Linz, and A. P. Moczek. 2021. *Doublesex* mediates species-, sex-, environment- and trait-specific exaggeration of size and shape. Proc. R. Soc. B Biol. Sci. 288:20210241. https://doi.org/10.1098/rspb.2021.0241

Rowe, L., and D. Houle. 1996. The lek paradox and the capture of genetic variance by condition dependent traits. Proc. R. Soc. Lond. B Biol. Sci. 263:1415–1421.

Rueden, C. T., J. Schindelin, M. C. Hiner, B. E. DeZonia, A. E. Walter, E. T. Arena, and K. W. Eliceiri. 2017. ImageJ2: ImageJ for the next generation of scientific image data. BMC Bioinformatics 18:529. https://doi.org/10.1098/rspb.1996.0207

Sawala, A., and A. P. Gould. 2017. The sex of specific neurons controls female body growth in *Drosophila*. PLoS Biol. 15:e2002252. https://doi.org/10.1371/journal.pbio.2002252

Schubiger, G., M. Schubiger, and A. Sustar. 2012. The three leg imaginal discs of *Drosophila*: “Vive la différence.” Dev. Biol. 369:76–90. https://doi.org/10.1016/j.ydbio.2012.05.025

Scott, A. M., I. Dworkin, and R. Dukas. 2022. Evolution of sociability by artificial selection. Evolution 76:541–553. https://doi.org/10.1111/evo.14370

Setoguchi, S., H. Takamori, T. Aotsuka, J. Sese, Y. Ishikawa, and T. Matsuo. 2014. Sexual dimorphism and courtship behavior in *Drosophila prolongata*. J. Ethol. 32:91–102. https://doi.org/10.1007/s10164-014-0399-z

Shingleton, A. W., J. Das, L. Vinicius, and D. L. Stern. 2005. The Temporal Requirements for Insulin Signaling During Development in *Drosophila*. PLoS Biol. 3:e289. https://doi.org/10.1007/s10164-014-0399-z

Singh, B. 1977. Two new and two unrecorded species of the genus *Drosophila* Fallen (Diptera: Drosophilidae) from Shillong, Meghalaya, India. *Proc. Zool. Soc*. 30:31

Stephens, M., P. Carbonetto, D. Gerard, M. Lu, L. Sun, J. Willwerscheid, and N. Xiao. 2016. ashr: Methods for Adaptive Shrinkage, using Empirical Bayes.

Tanaka, K., O. Barmina, and A. Kopp. 2009. Distinct developmental mechanisms underlie the evolutionary diversification of *Drosophila* sex combs. Proc. Natl. Acad. Sci. 106:4764–4769. https://doi.org/10.1073/pnas.0807875106

Tanaka, K., O. Barmina, L. E. Sanders, M. N. Arbeitman, and A. Kopp. 2011. Evolution of Sex-Specific Traits through Changes in HOX-Dependent *doublesex* Expression. PLoS Biol. 9:e1001131. https://doi.org/10.1371/journal.pbio.1001131

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. https://doi.org/10.1007/s00427-016-0534-7

Toubiana, W., D. Armisén, C. Dechaud, R. Arbore, and A. Khila. 2021. Impact of male trait exaggeration on sex-biased gene expression and genome architecture in a water strider. BMC Biol. 19:89. https://doi.org/10.1186/s12915-021-01021-4

Toyoshima, N., and T. Matsuo. 2023. Fight outcome influences male mating success in *Drosophila prolongata*. J. Ethol., **41**, 119–127 https://doi.org/10.1007/s10164-023-00778-1.

Voje, K. L. 2016. Scaling of Morphological Characters across Trait Type, Sex, and Environment. Am. Nat. 187:89–98. https://doi.org/10.1086/684159

Warren, I. A., H. Gotoh, I. M. Dworkin, D. J. Emlen, and L. C. Lavine. 2013. A general mechanism for conditional expression of exaggerated sexually-selected traits. BioEssays 35:889–899. https://doi.org/10.1002/bies.201300031

Wasik, B. R., D. J. Rose, and A. P. Moczek. 2010. Beetle horns are regulated by the Hox gene, Sex combs reduced, in a species- and sex-specific manner. Evol. Dev. 12:353–362. https://doi.org/10.1111/j.1525-142X.2010.00422.x

Wat, L. W., Z. S. Chowdhury, J. W. Millington, P. Biswas, and E. J. Rideout. 2021. Sex determination gene *transformer* regulates the male-female difference in *Drosophila* fat storage via the adipokinetic hormone pathway. eLife 10:e72350. eLife Sciences Publications, Ltd. https://doi.org/10.7554/eLife.72350

Wilkinson, G. S., P. M. Johns, J. D. Metheny, and R. H. Baker. 2013. Sex-Biased Gene Expression during Head Development in a Sexually Dimorphic Stalk-Eyed Fly. PLoS ONE 8:e59826. https://doi.org/10.1371/journal.pone.0059826

Williams, T. M., J. E. Selegue, T. Werner, N. Gompel, A. Kopp, and S. B. Carroll. 2008. The Regulation and Evolution of a Genetic Switch Controlling Sexually Dimorphic Traits in *Drosophila*. Cell 134:610–623. https://doi.org/10.1016/j.cell.2008.06.052

Yan, D., and N. Perrimon. 2015. *spenito* is required for sex determination in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. 112:11606–11611. https://doi.org/10.1073/pnas.1515891112

Zhu, C., M. J. Ming, J. M. Cole, M. D. Edge, M. Kirkpatrick, and A. Harpak. 2023. Amplification is the primary mode of gene-by-sex interaction in complex human traits. Cell Genomics 3:100297. https://doi.org/10.1016/j.xgen.2023.100297

Zinna, R., D. Emlen, L. C. Lavine, A. Johns, H. Gotoh, T. Niimi, and I. Dworkin. 2018. Sexual dimorphism and heightened conditional expression in a sexually selected weapon in the Asian rhinoceros beetle. Mol. Ecol. 27:5049–5072. https://doi.org/10.1111/mec.14907

Figures

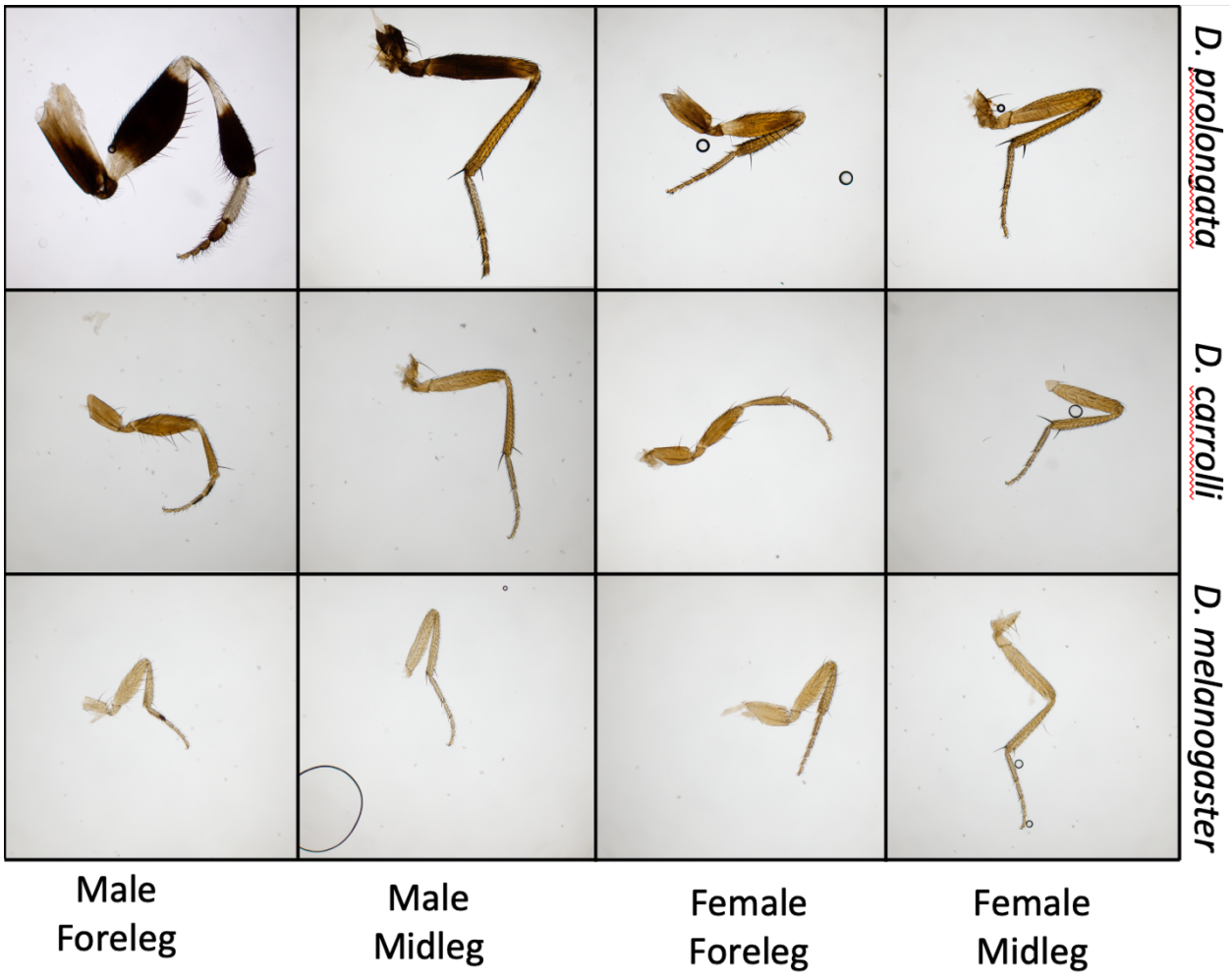

Figure 1: Images of adult fore- and midleg of each species and sex. All were sampled at juvenile stages (i.e., as imaginal discs) for RNA sequencing. All images taken at 40x magnification.

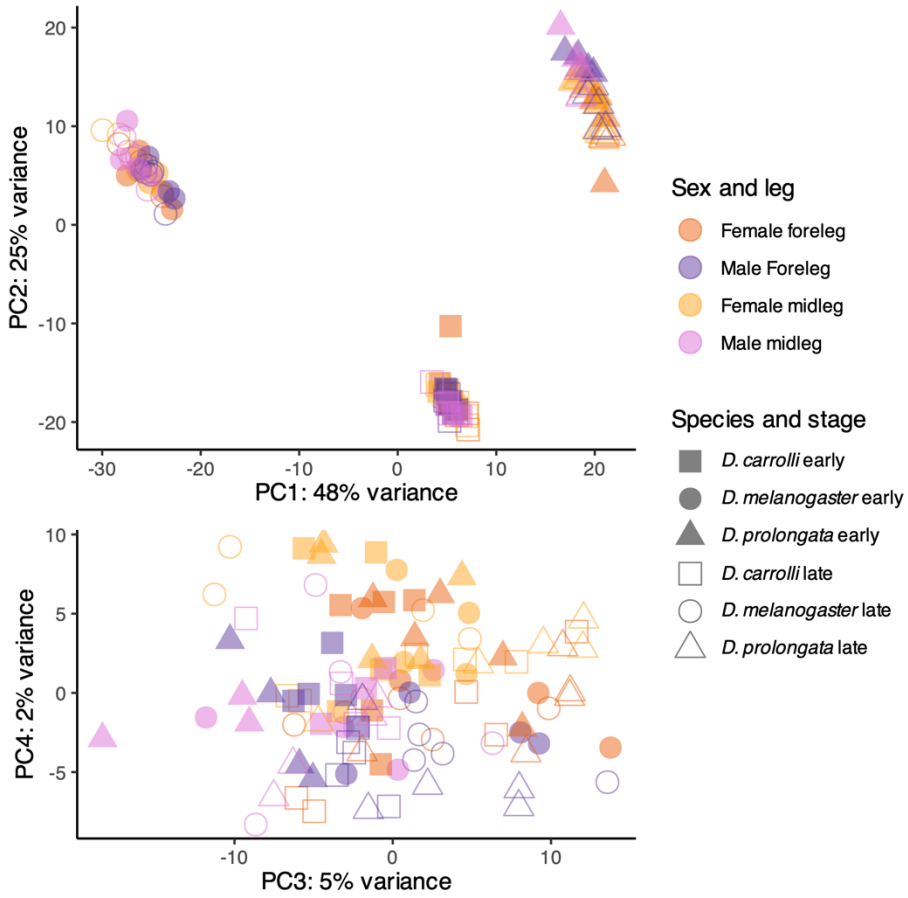

Figure 2: PCA projection of developmental stage, leg, sex, and species for the top 1000 most variable genes in *D. prolongata*, *D. carrolli*, and *D. melanogaster* with an rlog regularization applied.

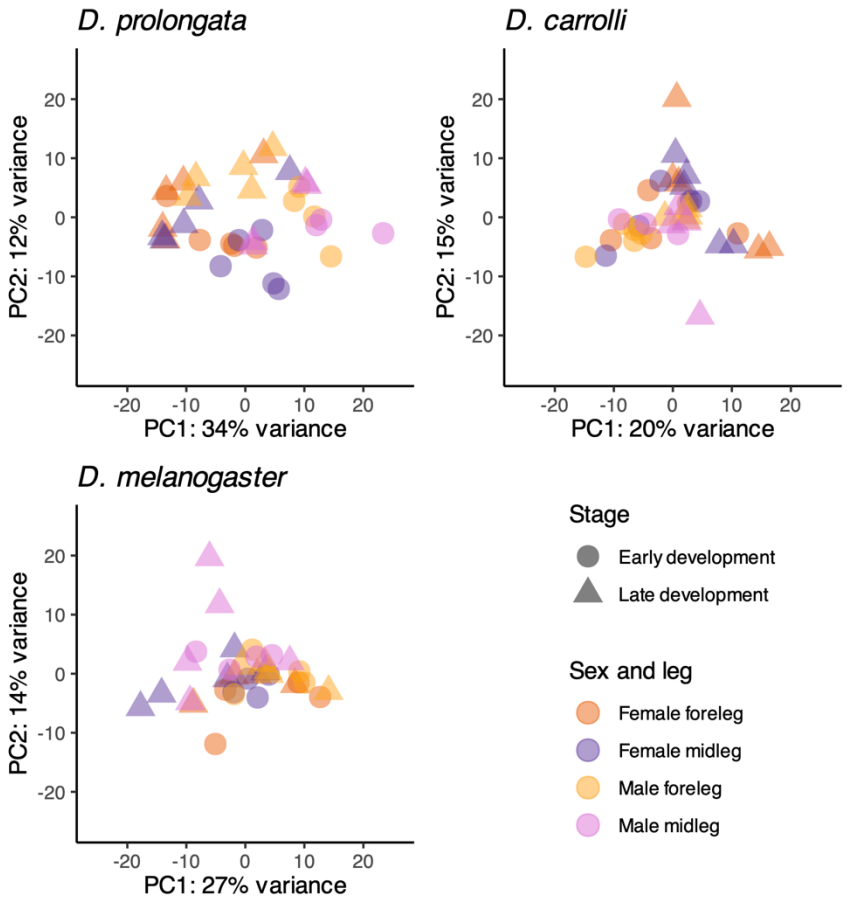

Figure 3: PCA projection of developmental stage, leg, and sex within species for the top 1000 most variable genes in *D. prolongata*, *D. carrolli*, and *D. melanogaster* with a regularized log transformation applied.

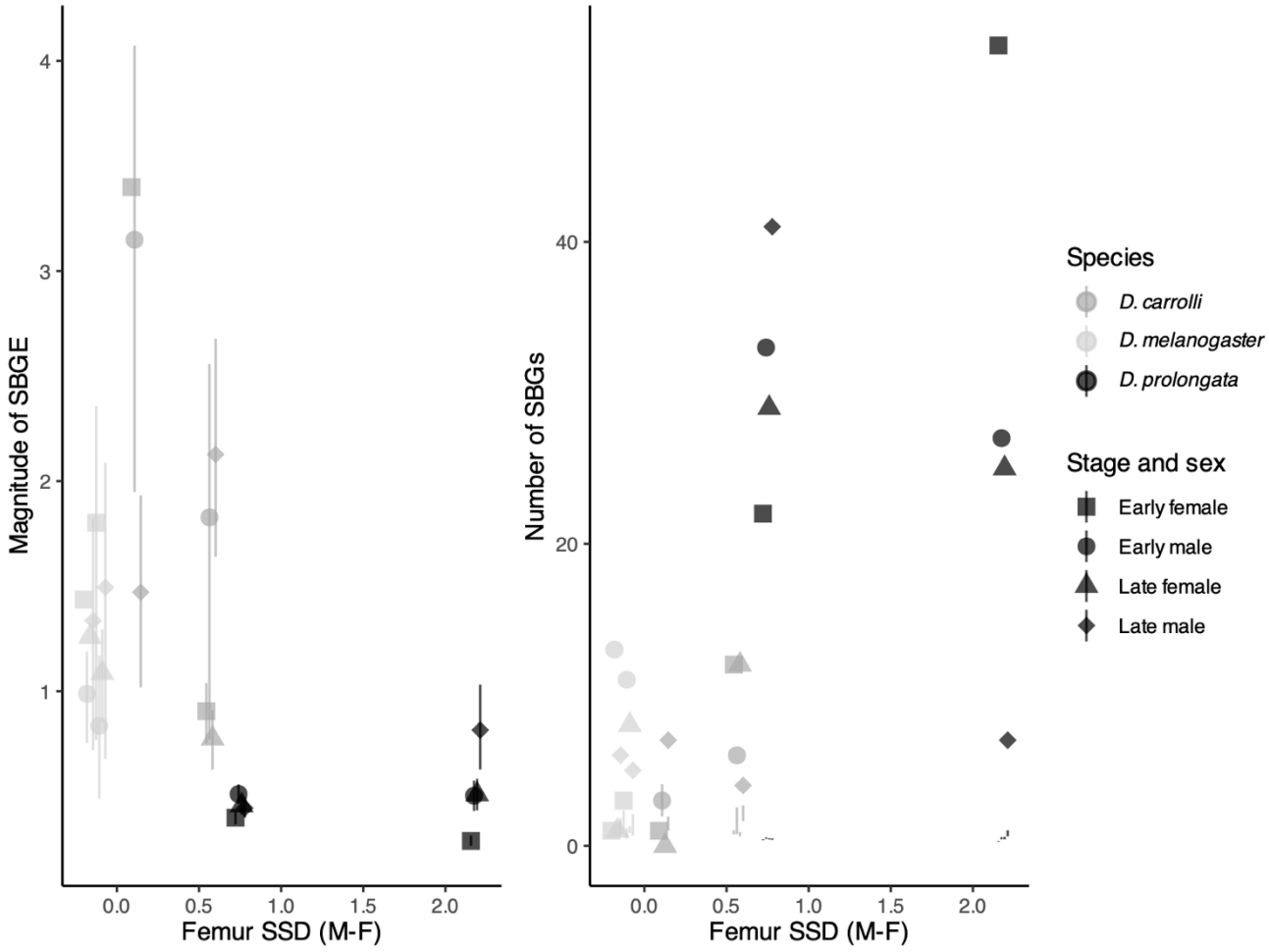

Figure 4: Relationship of the magnitude of expression (SBGE) changes (left) or number (right) for sex-biased genes (SBG) and degree of adult sexual size dimorphism, measured as female – male trait size. Magnitude and number of SBGs represent the number (or magnitude) of male (female) biased genes at each developmental stage in each species.

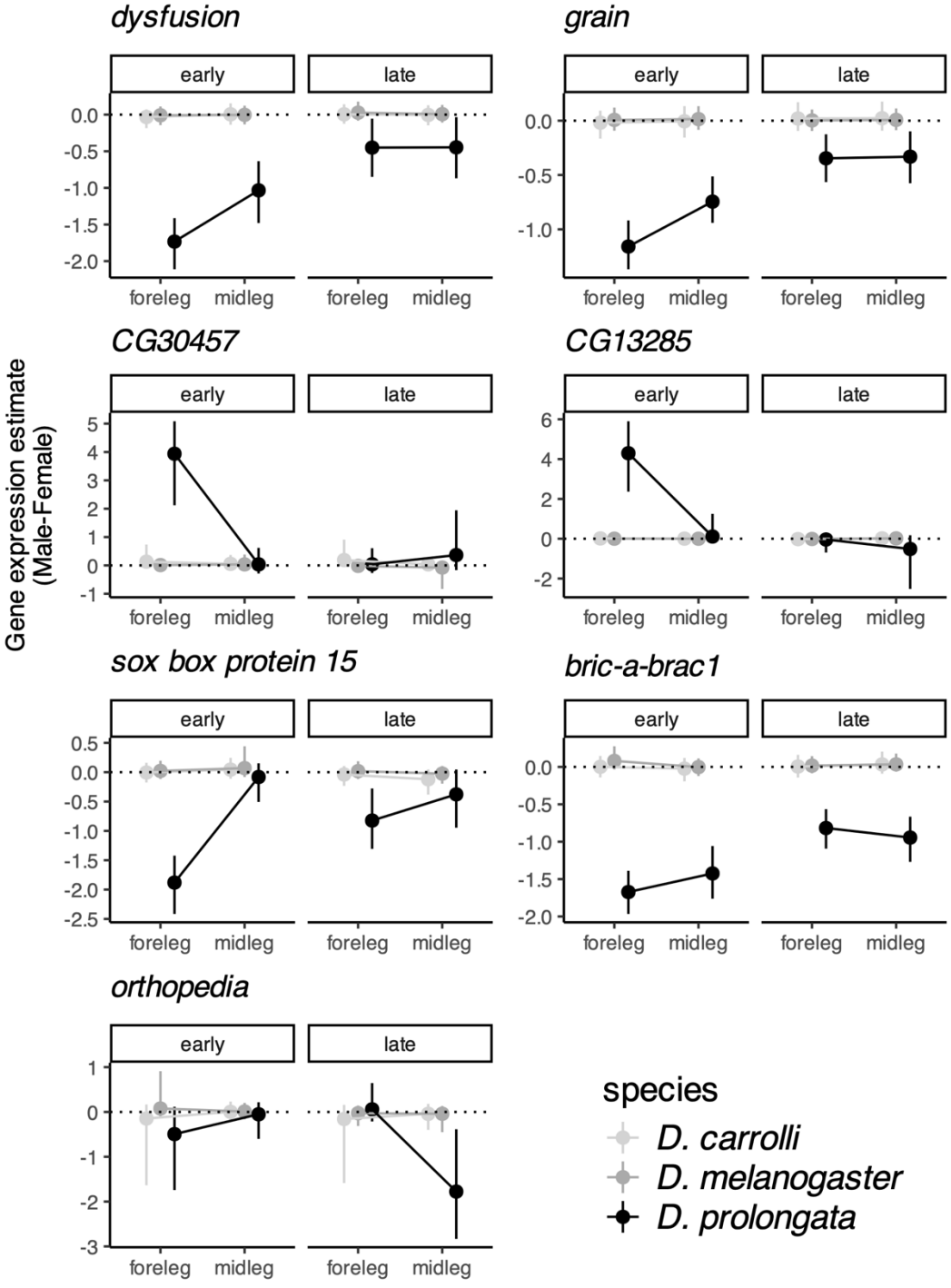

Figure 5: Interaction plots of sex-biased gene expression for candidate genes chosen for RNAi-mediated gene knockdown. Male-female contrast (regularized estimates) of gene expression (log_2) on the y-axis and both stage and leg are shown separate for each gene and species.

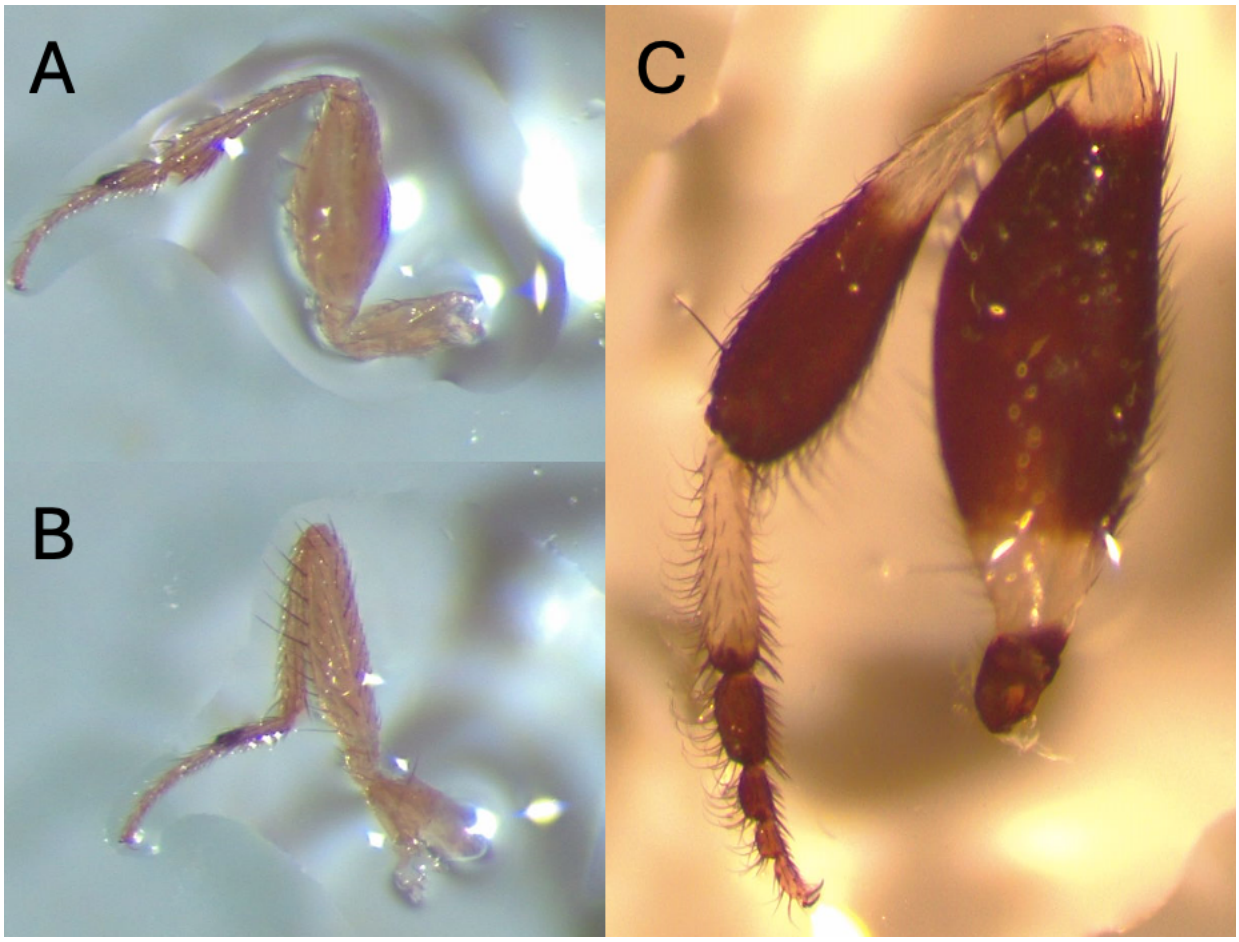

Figure 6: Male foreleg in A) *grn* knockdown *D. melanogaster* B) control *D. melanogaster*. C) wild-type *D. prolongata.* All images are at the same scale.

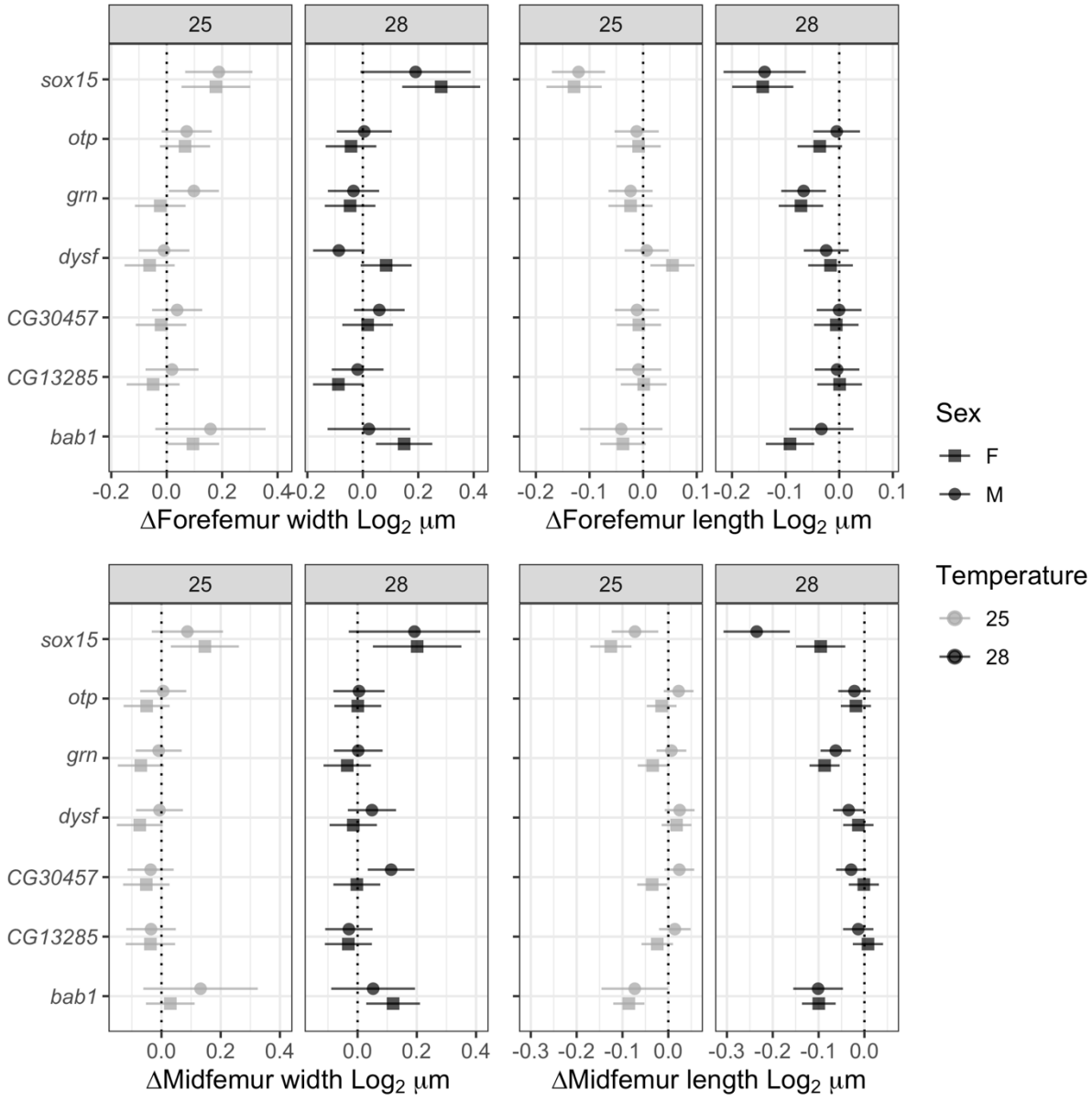

Figure 7: Log_2 leg traits size relative to thorax for UAS-RNAi strains crossed to NP3666-GAL4 contrasted with control crosses at both 25ºC and 28ºC. Contrasts calculated from model estimates using emmeans.

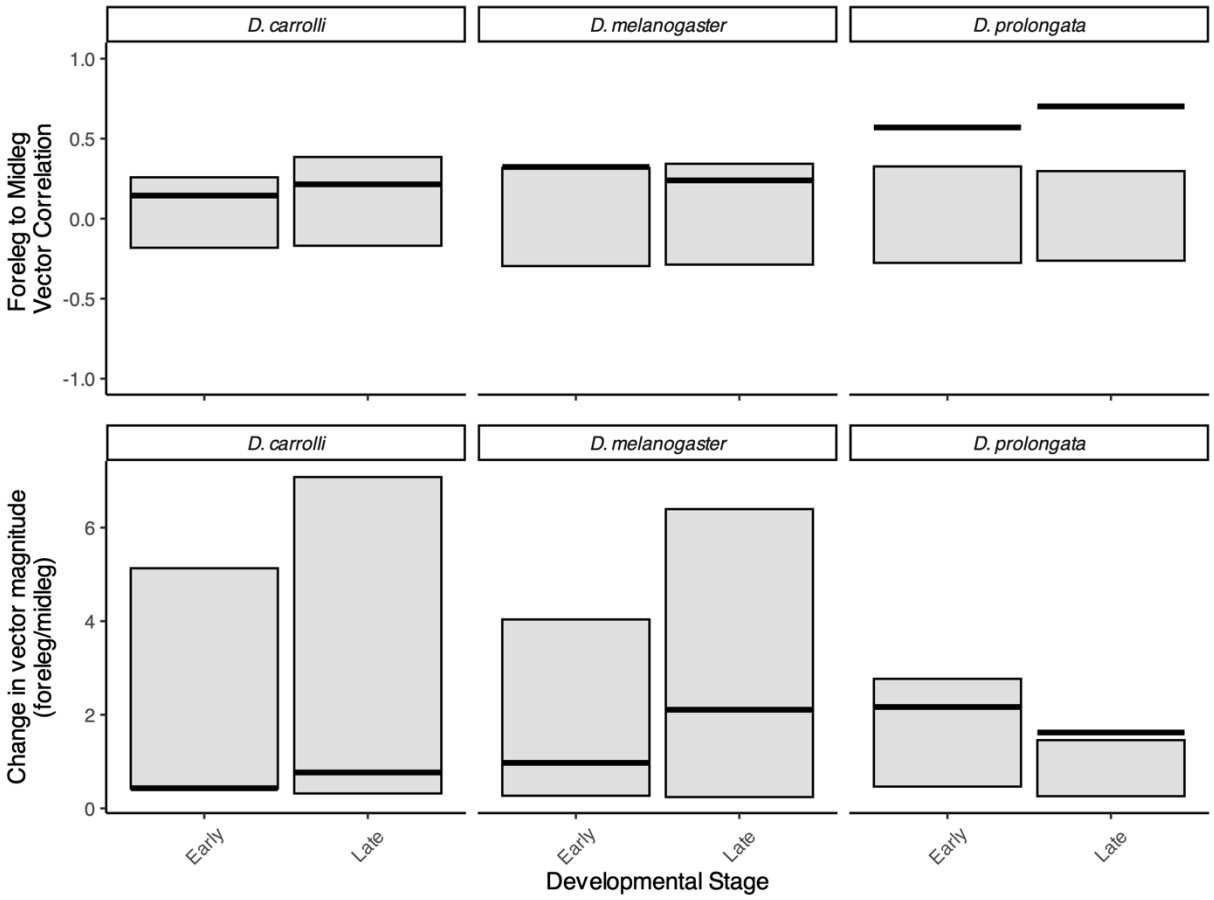

Figure 8: Changes in direction and magnitude of expression for vectors of SBGs, between foreleg and midleg imaginal discs. Based on ~80 SBGs with a minimum 2-fold difference (one log_2 unit) in *D. prolongata,* that also share orthologs in both *D. carrolli* and *D. melanogaster*. Top row shows correlations of vectors (degree of shared direction of effects) for SBGE in foreleg compared to midleg within all three species. Bottom row is the relative change in magnitude of SBGE in foreleg relative to midleg imaginal discs within each species for the same set of genes.

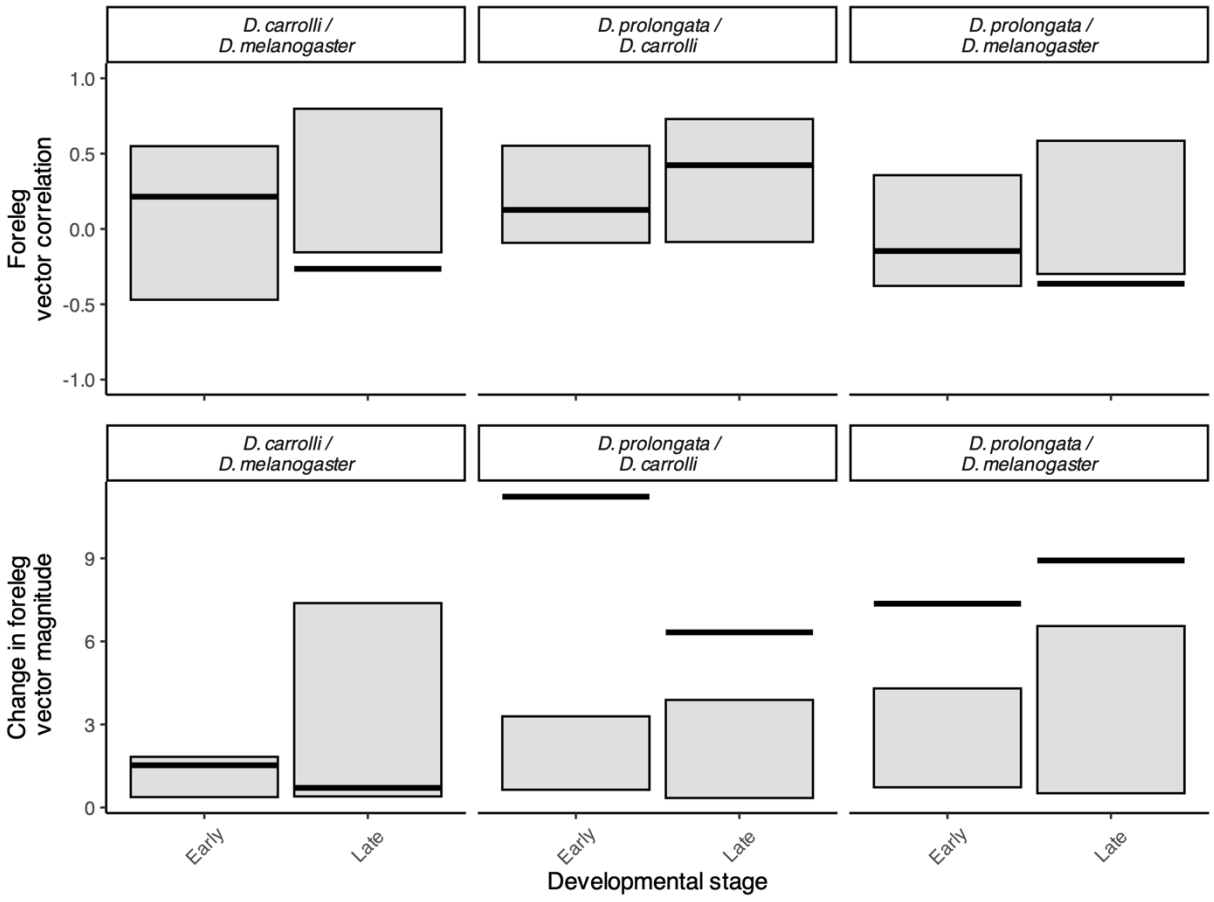

Figure 9: Degree of shared direction and changes in magnitude of vectors for sex-biased genes between species in the foreleg imaginal disc. Based on ~80 SBGs with a minimum 2-fold difference (one log_2 unit) in *D. prolongata,* with orthologs in *D. carrolli* and *D. melanogaster*. Top row shows degree of shared direction (expressed as a correlation) of SBGE in foreleg between *D. prolongata* and the other species, the first column shows *D. carrolli* compared to *D. melanogaster* as a point of reference. Bottom row is the change in magnitude of SBGE in foreleg between each species. Values greater than one represent a relative increase in magnitude of SBGE for this set of genes in *D. prolongata* relative to the other species.

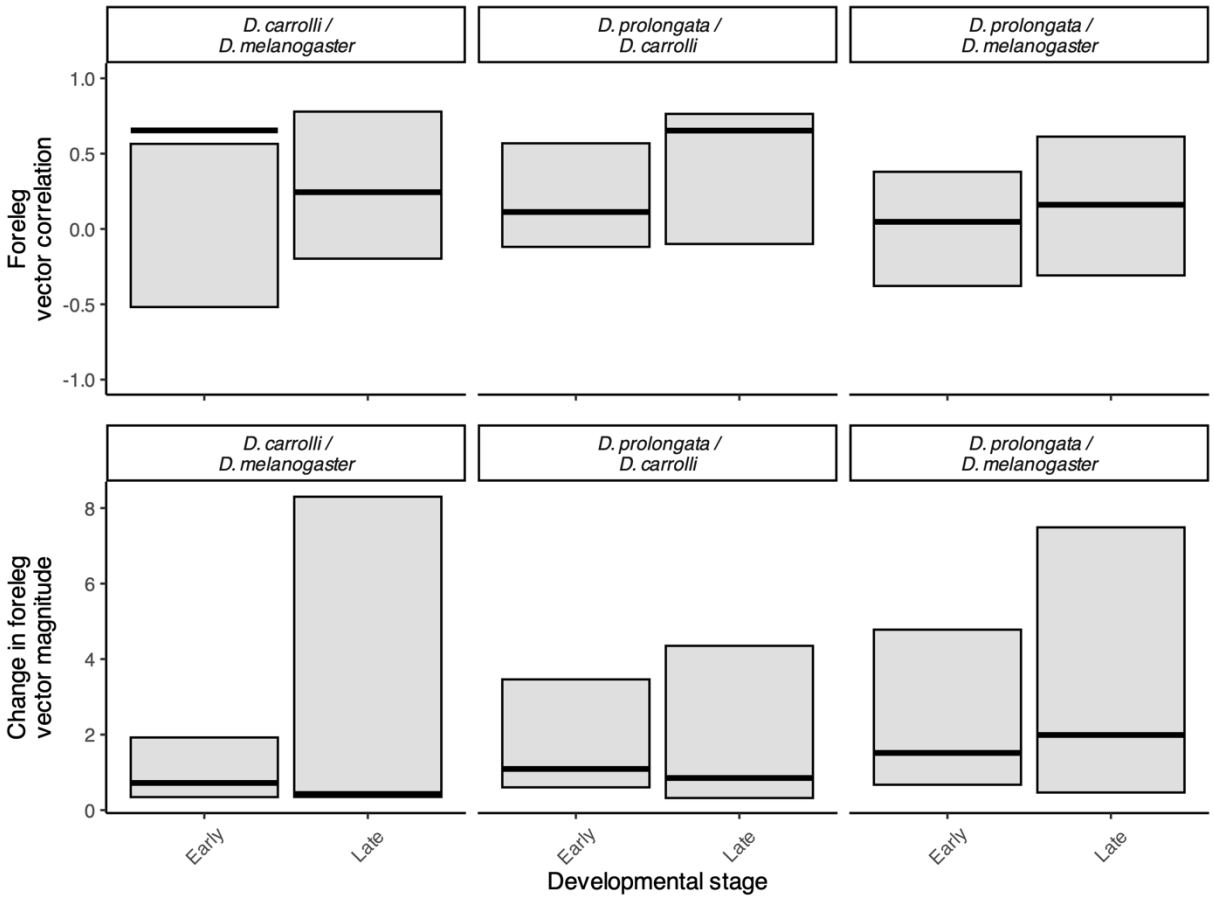

Figure 10: Degree of shared direction, and changes in magnitude for Insulin signalling genes between species in the foreleg. Top row shows direction of SBGE in foreleg compared between *D. prolongata* and the other species, and the first column showing *D. carrolli* compared to *D. melanogaster* as a point of reference. Bottom row is the magnitude of SBGE in foreleg between each species.

Table 1: Number of genes with at least a log_2-fold change in each species for each sex (Log_2 cut-off based on upper- or lower-bound of the 95% confidence interval). Foreleg genes are those that are not log_2 fold changed in midleg as well. Numbers in parentheses represent sex-biased genes with regularized 95% CIs that do not overlap zero.

Species	Female biased	Male biased	Female biased foreleg	Male biased foreleg
D. prolongata	173 (2820)	175 (2845)	58 (1920)	65 (2248)
D. carrolli	19 (1062)	35 (1745)	7 (930)	34 (1590)
D. melanogaster	39 (1559)	21 (1412)	9 (1189)	17 (1123)

3. Sexually discordant selection is associated with trait-specific morphological changes and a complex genomic response

Audet, T., Krol, J., Pelletier, K., Stewart, A. D., & Dworkin, I. (2024). Sexually discordant selection is associated with trait-specific morphological changes and a complex genomic response. *Evolution*, *78*(8), 1426-1440.

Evolution, 2024, **78(8)**, 1426–1440
https://doi.org/10.1093/evolut/qpae071
Advance access publication 9 May 2024
Original Article

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

Tyler Audet[1], Joelle Krol[1], Katie Pelletier[1,3], Andrew D. Stewart[2], Ian Dworkin[1,]

[1]Department of Biology, McMaster University, Hamilton, ON, Canada
[2]Department of Biology, Canisius University, Buffalo, NY, United States
[3]Present address: Institut de Biologie de l'Ecole Normale Superieure, CNRS, Paris, France

Corresponding authors: Department of Biology, McMaster University, Hamilton, ON, Canada. Email: dworkin@mcmaster.ca; Department of Biology, Canisius University, Buffalo, NY, United States. Email: stewar34@canisius.edu

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 (SD), 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 SD. 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 selected lineage shows a pattern of elevated genomic differentiation between males and females on chromosome 3L, consistent with the maintenance of sexual conflict. Our results suggest visible signs of conflict and differentially segregating alleles between the sexes due to discordant selection.

Keywords: evolutionary genomics, genetic variation, morphological evolution, artificial selection, sexual selection, sex

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 r_{MF} (Delph et al., 2011; Lande, 1980). Many organisms show high r_{MF}, which could limit the potential for sexually discordant evolution (Lande, 1980; Poissant et al., 2010). Despite this, SD is very common in nature, particularly, 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 & 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; Luecke & Kopp, 2019; Moczek et al., 2006). When loci impact the phenotype in a way favoring 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 is not sufficiently high. Despite high r_{MF} values, family-based artificial selection experimental designs, where phenotypic selection is based on either within or among family trait values, have demonstrated that sex-specific responses can occur relatively rapidly (Alicchio & Palenzona, 1971; Bird & Schaffer, 1972; Eisen & Hanrahan, 1972; Kaufmann et al., 2021). 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 for size with minimal response in the unselected sex, while sex-specific selection in high r_{MF} lines resulted in a strongly correlated response in the unselected sex. Although family-based selection experiments demonstrate additive genetic variation for a response to sexually discordant selection, these do not reflect the transmission of allelic effects in most natural populations. As such, approaches based on mass artificial selection may better reflect the transmission of allelic effects in natural populations because alleles of interest may be rarer, and each round of selection will be subject to drift and recombination with random haplotypes from other individuals. Using strong, long-term, mass artificial selection,

Received August 31, 2023; revisions received April 12, 2024; accepted May 7, 2024

Associate Editor: Charles Baer; Handling Editor: Tim Connallon

Stewart and Rice (2018) demonstrated that a response to a sex-discordant selection pressure can occur. Stewart and Rice (2018) successfully selected body size in *D. melanogaster* in a sex-discordant manner over 250 generations, with a measurable phenotypic response requiring more than 100 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 & Lewis, 2011) despite a rapid response to sexually concordant selection on pupal mass. This suggests barriers to divergent response in SD may exist in some populations, although it may reflect modest genetic variation in the ancestral populations and the limited number (7) of generations of artificially 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 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 influence SSD. Manipulation of sex-specific splice variants of the *transformer* (*tra)* gene reduces female size, reducing (but not eliminating) SSD (Rideout et al., 2015). Increasing *tra* expression showed sex-specific increases in size in females (Rideout et al., 2015). A duplication of the *diminutive* (*dyc*) 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 a reduction in female body size (Sawala & Gould, 2017). Inhibition of insulin signaling had sex-specific impacts on body size, largely reducing female body size. In contrast, upregulation of inhibition of insulin-signaling increased male body size (Millington et al., 2021b). These experiments demonstrate pathways involved with phenotypic expression of SD, but it is unclear whether segregating variation in these pathways contributes 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 on segregating genetic variation, rather than *de novo* mutations, likely mediates short-term responses for body size evolution. In natural populations, segregating alleles contributing to variation in SD interest could be 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 & Chippindale, 2001). Variation in body size in *D. melanogaster* is highly polygenic (Carreira et al., 2009; Turner et al., 2011). Sex-limited selection of males has been shown to incur a fitness cost to females (Prasad et al., 2007). This limited male evolution also results in a change in body size in nonselected females, which 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 potentially under conflict in the genomes of the outbred LH_M population of *D. melanogaster* have been identified via examination of male mating success and female fecundity (Ruzicka et al., 2019). Alleles under IASC influencing body size can help understand how the genome responds to selection with divergent phenotypic optima. The selection of body size in a discordant manner has the potential to answer questions about how the shared genome overcomes high r_{MF} when selection favors divergent phenotypes across the sexes.

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 (i.e., selection on general size rather than an increase in mass or thorax length) 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 reintroduced) sexual conflict. Using an evolve and sequence approach, we examined lineage patterns of genomic differentiation and within-lineage among the sexes. We discuss our findings within the context of both the evolution of SD and the potential role of ongoing intra-locus sexual conflict.

Methods

Lineages

The populations used are part of a long-term experiment on size evolution (Stewart & Rice, 2018; Turner et al., 2011). 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 2-week generations at moderate density (approximate 200 eggs per 10 ml standard molasses food vial; Supplementary Table S1). The complete methodology for the maintenance of the base population has been published previously (Rice et al., 2005). For size selection in each generation, 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 2,000 µm, and the smallest sieve had apertures measuring 850 µm. In each generation, all flies (~1,800 individuals) from a selection lineage were sieve sorted, and 10 vials of 16 mating pairs (320 total flies per lineage) selected based on phenotype were used to generate 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 the smallest females and largest males (E). Finally, a control treatment (C) was populated with flies that passed through the sieves but were not selected based on size. Two independent replicate lineages were maintained for each selection treatment. During subsequent selection for size with the established protocol, it was observed that selection could be more stringent on small flies than large ones because of the nature of the sieve sorting (hindrances due to appendages sticking out, blocking the passage of flies through the

sieve). This would imply that in the sex-discordant selection, females selected to be small were under increased selective pressure compared to males selected to be large. This suggests greater pressure for selection in one direction, and although this does not explain all the results, it may influence responses to sexually discordant selection.

Morphological measures and analysis

At generation 367 (August 19, 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 63× 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 were 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 was completed using R v4.1.3 (R Core Team, 2021). One individual measure of the femur was removed as it's measure was ~100× smaller than the mean for the trait. For all analyses of morphology, we used $\log_2$ transformed trait values to facilitate inferences of proportional changes in dimorphism. We fit general linear mixed models for individual traits, including sex, selection, and sampling, and their 2nd order interactions as fixed effects. We allowed sex effects to vary as a random effect of replicate lineage (i.e., random slopes). Models were fit using *lmer* in the *lme4* package v1.1.30 (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" on a 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 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; Kristensen & McGillycuddy, 2023; Niku et al., 2019). Estimated marginal means, custom contrasts, and associated 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 the artificial selection being "trait-agnostic" (selecting on a composite cross-sectional area, with possible hindrance 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 (Blackith & Reyment, 1971; Jolicoeur, 1963; Klingenberg, 1996). It is common to assess whether the log-transformed traits contribute approximately equally to the size of PC1 (isometry), with expected loadings of $1/\sqrt{p}$, where p is the number of traits (dimensions). In addition to examining (and visualizing) these vectors from principal component analyses conducted by each selective treatment, we compared aspects of the orientation and structure of the variance-covariance matrices across treatments.

Evaluating sex ratio of lineages and crosses among them

While phenotyping flies for an unrelated experiment, a deviation of the expected 50/50 sex ratio was observed in a cross between control lineage 1 and discordant lineage 1 samples. As summarized in the results, discordant lineage 1 is the lineage with potential evidence of maintained conflict on chromosome 3L. As such, we conducted experimental crosses to examine adult sex ratios and 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 LH_M population. Single-pair reciprocal matings were allowed to lay until larvae were visible in the food (~5 days) before F_0 pairs were placed in 70% ethanol. From each F_1 vial, a single-pair was used to generate an F_2, while the rest of the F_1 individuals were stored in 70% ethanol after being allowed to eclose until most pupal casings were visibly empty. Adults from the F_2 generation, once mostly eclosed, were stored in 70% ethanol. The number of male and female offspring in F_1 and F_2 generations were counted.

We modeled the data for the sex ratio crosses using 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 not the individual was treated as a sire or dam) in the model. During the experiment to examine adult sex ratios, we noticed substantial differences in the number of individual offspring. While this was not a planned analysis (and should be treated as such), we examined differences in fecundity (assessed by the census of adults) fitting a general linear model (fit using lm) of the number of offspring regressed onto the direction of the cross, selection treatments, and their interactions.

Genomic sample preparation

At generation 378 (February 17, 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 into four pools of 25 individuals, 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 was combined so that the same concentration of DNA was added from each pool (equimolar). This resulted in pools of 100 individuals for each sequenced

Figure 1. Example samples from control and discordant lineages. Stereomicroscope images of Control and discordant males and females (25× magnification). Flies chosen at random from populations stored and imaged in 70% ethanol.

combination of sex and unique lineage. 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). Libraries were prepared with IDT dual index adapters. Sequencing was done using Illumina NovaSeq 6000 to an average coverage of 200× with 151 bp sequence fragments. Initial sequencing fell short of 200× coverage, so additional sequencing from the same libraries was conducted to "top-up" coverage on some samples and were merged with their corresponding samples after mapping.

Bioinformatic pipeline

Supplementary Figure S1 shows a summary of the bioinformatics pipeline. A detailed summary can be found at (https://github.com/DworkinLab/Audet_etal_Evolution_2024). 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 et al., 2013–2015), 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 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, pre-print). This sync file was filtered for sites present in our SNP-called VCF file to create an SNP-called sync file. For "sexes separate" as well as "treatments pooled" sync files. Through testing, we observed 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 was 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, we used two related approaches, F_{ST} as a measure of the magnitude of change in allele frequencies and a modified Cochran–Mantel–Haenszel (CMH) statistic that incorporates sources of sampling variation and genetic drift common to evolve and re-sequence experiments. To assess variation within populations and to identify potential selective sweeps, we performed windowed computation of nucleotide diversity (π). We calculated F_{ST} between control treatments and discordant treatments using Grenedalf v0.2.0 in 10 kb windows along each chromosome (Czech et al., 2023), which corrects for sample sizes and pooled sequencing size, known sources of error in parameter estimation for pooled sequencing. We also examined F_{ST} between Large and Small treatments to follow up on the initial analysis in Turner et al. (2011) to see if their findings were replicated after an

additional ~275 generations of artificial selection on body size. To do so, we ran the sequence data from generation 100 through our pipeline and compared genes of interest from F100 to the genes identified at F378.

To calculate effective population size (N_e) we used poolSeq v0.3.5 (Taus et al., 2017). We estimated N_e between Control 1 and Discordant 1 treatments (N_e = 181) and between large 1 and small 1 (N_e = 69) and used these to calculate CMH statistics with ACER version 1.0 (Spitzer et al., 2020). We also evaluated masking all sites that were fixed in our time point 0 (control for control vs. discordant, large for large vs. small) using the checkSNP() function. This resulted in higher N_e estimated (C1 vs. E1 = 207.90, C2 vs. E2 = 224.51; L1 vs. S1 = 90.65, L2 vs. S2 = 139.76) so we used the previous smaller estimates. N_e calculations for the first replicate comparison were done using formulae proposed by Jónás et al. (2016) and are intended to work with pooled data as well as control for additional sources of sampling variance in allele frequencies that could bias N_e. Since an ancestral sample was not preserved, drift between control and discordant treatments (or large versus small) was used, and the generations of drift were set at 750 from the control treatment (~generations to convergence ×2).

For our samples, given the large number of generations of artificial selection, selective sweeps would likely result in regions of low nucleotide diversity, which we measured as π using Grenedalf v0.2.0 in 10 kb windows for our control, discordant, and concordant samples. This resulted in us having per sample π values for 10 kb windows, F_{ST} values for 10 kb windows, and SNP-by-SNP adjusted *p*-values from a CMH test. Using these values, regions with a small π value (<0.0005) were extracted as interesting, and were extracted along with the top 5% of F_{ST} (discordant vs. control cutoff was F_{ST} 0.51 with 621 windows meeting these criteria; large vs. small cutoff was F_{ST} 0.78 with 621 windows meeting that criteria) and SNPs with a CMH adjusted *p*-value <0.01 (using *p.adjust* in R (Benjamini & Hochberg, 1995); 16,218 SNPs in control vs. discordant, 115,747 SNPs in large vs. small) were extracted. The overlap between all windows or SNPs of interest was intersected with Bedtools v2.31.0 (Quinlan & Hall, 2010) to get SNPs of interest. Unfortunately, most software (including Grenedalf) does not account for the hierarchical structure of populations (common in experimental designs like artificial selection), nor the expected sources of drift and other forms of sampling variance among replicate lineages (within treatments) in evolve and re-sequence experiments. As such, the approach we used above (merging replicate lineages) does not account for sampling variation. To partially account for this, for the subset of "candidate" SNPs we identified, we fit SNP-specific logistic regression models, with SNP counts for each replicate lineage representing units of sampling and with selective treatments as the predictor. Treatment contrast between either discordant versus control or large versus small treatments was obtained using *emmeans*. As this approach is computationally slow, we only did this for the sites we identified as likely candidates (discordant = 20,014, large = 10,649, small = 13,952). From the overlap of all these analyses, we created a file containing a region of interest on chromosome 3L in discordant replicate 1, the interesting discordant sites (Supplementary File 2), interesting large sites (Supplementary File 3), and interesting small sites (Supplementary File 4). These sites were extracted from the SNP-called VCF file using bedtools v2.30.0 (Quinlan & 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 control replicate 1 major allele for the above-mentioned modeling.

Sites of interest were annotated (Supplementary Files 1, 2, 3, and 4). Gene ontology (GO) enrichment analyses were conducted Gowinda v1.12 (Kofler & Schlötterer, 2012), which controls for gene length and uses permutations to reduce false positives. 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 within the same experimental lineage and generation for each of the evolved lineages (Cheng & Kirkpatrick, 2016; Kasimatis et al., 2019; Lucotte et al., 2016). Previous simulation analyses have suggested that identifying IASC using intersex F_{ST} generally has low power (Kasimatis et al., 2019). However, the experimental design used for artificial, sexually discordant selection is well suited for this particular approach given strong and persistent selection across many generations. We computed allele frequency tables for each sex within each lineage, along with F_{ST} between males and females within lineages in Grenedalf (window size 5 kb). As summarized in the results, we observed a region on chromosome 3L showing elevated F_{ST}, in a single discordant treatment lineage (E1).

In addition to the possibility that this elevated region on chromosome 3L was due to maintained sexual conflict, there are several other possible explanations that we evaluated. To assess whether this elevated region of F_{ST} could be accounted for simply by sampling variation, we simulated 1,000,000 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 whether 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*.

To determine the extent to which demographic influences of lab adaptation and artificial selection in the context of empirically derived estimates of recombination along the chromosome arm could account for the elevated region of F_{ST} between males and females, we performed evolutionary simulations using SLiM v4.1-4.2 (Haller & Messer, 2023). We broke the simulation into two phases. The "burn-in" phase to simulate patterns of variation along chromosome 3L in the natural population from which LH_M was derived to serve as the ancestral population, followed by simulations to assess demographic impacts due to the founding and maintenance of the LH_M population and artificial selection lineages derived from it. To account for variation in recombination rates and their influence on evolutionary dynamics on chromosome 3L, we used empirically derived estimates of a recombination rate map for chromosome

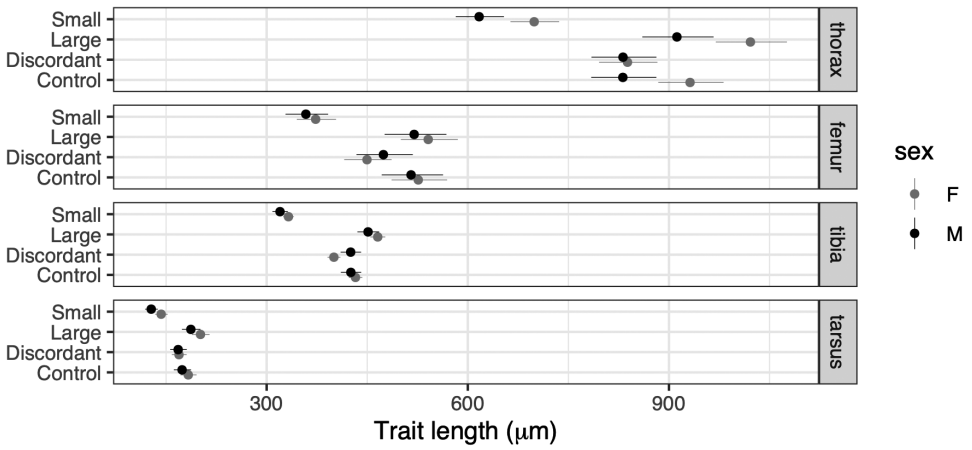

Figure 2. Trait-specific evolution of size among artificially selected lineages. Model estimates for each trait across treatments and sex. Measurements were $\log_2$ transformed for model fit back-transformed for plotting. Error bars are 95% CIs.

3L (Comeron et al., 2012). As recombination does not naturally occur in *D. melanogaster* males, we simulated sex-specific recombination rates, using the above recombination map for females and 0 for males. For mutation rates, we used the average of empirically derived estimates, 3.2e-9 (Cingolani et al., 2012; Keightley et al., 2009, 2014) as our per-site estimate. To capture the expected variation for chromosome 3L from the ancestral North American population used to find LH_M, we used an estimated $N_e = 5.54 \times 10^5$ (Arguello et al., 2019). This phase of the simulation was done for 10,000 generations. To speed up computation, we use the protocol suggested in the SLiM manual, rescaling mutation and recombination rates so that the number of individuals that needed to be simulated each generation was reduced by a factor of 20. The recombination rate was rescaled as $r' = 0.5 \times (1 - (1 - 2r)^n)$, where r is the recombination rate, and n is the rescaling factor. This initial simulation is found in the script "SeperateSexesBurnIn.slim." Thus, this first simulation represents the simulation of genomes (chromosome 3L) of individuals from the ancestral North American population from which all samples were derived. For the second part of the simulations (accounting for demographic effects of the founding and maintenance of LH_M and the artificial selection lineages), we sampled 1,792 individuals (1:1 sex ratio) from the ancestral population and maintained this population size for 360 discrete generations (capturing demographics of the LH_M population). Mutation and recombination rates were used as described above, without the need for parameter rescaling. Following this, we then simulated a population maintained at 320 individuals (1:1 sex ratio) to capture the population size of artificial selection lineages for an additional 377 generations. At the end of each simulation run ("SeperateSexesOutputSamples_Run_OnePop.slim"), we randomly sampled 200 chromosomes each for males and females, outputted allele counts for segregating sites, and after parsing the data into an appropriate format, computed windows of F_{ST} across the length of the simulated chromosome 3L using the same approach as discussed above for empirical data. We performed 100 simulations to assess the variation in variability and within-generation, within-lineage F_{ST} among males and females.

Results

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

Consistent with the previous results for overall "body size" (Stewart & Rice, 2018), all measured traits responded to artificial selection in the expected directions. Thorax length was measured as a proxy for body size and showed a clear relationship between size and selection lineage ($\chi^2 = 242.72$, $p < 0.0001$; Figure 2; Supplementary Table S2; Supplementary Figures S2 and S3) with some lineage-specific variation (Supplementary Figure S3). Concordant artificial selection for decreased size reduced thorax length relative to control lineages by ~25% in females (26% in males), while selection for increased size increased thorax length by ~9.5% for both sexes. In discordant selection lineages, we found a ~10% decrease in female thorax length and a minimal (~0%) change in males. We observed similar patterns for the length of 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 3). 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 the femur, ~7% for the tibia and tarsus). Males from the discordant lineage decreased more modestly relative to controls (~8% for the femur, <1% for the tibia, ~3.5% for the 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 (Figures 2 and 3; Supplementary Tables S3, S4, and S5).

SD 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 (Figures 2 and 3), showing little change in dimorphism relative to control (Figure 3). 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 3). 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 varied somewhat by trait but with consistent patterns of change (Figure 3).

Given the results described above, it is not surprising we observed substantial changes in patterns of multivariate allometry (Figure 4; Supplementary Figures S2, S3, and S4). 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

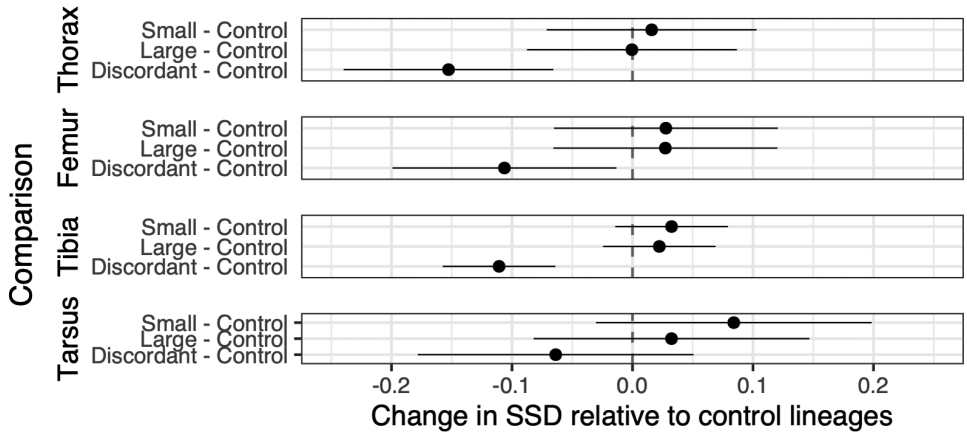

Figure 3. Substantial changes in sexual size dimorphism only occur under sexually discordant selection on size. Contrasts represent a proportional change in sexual size dimorphism for each artificially selected treatment, in comparison to controls, by trait. Modeled using $\log_2$ transformed length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

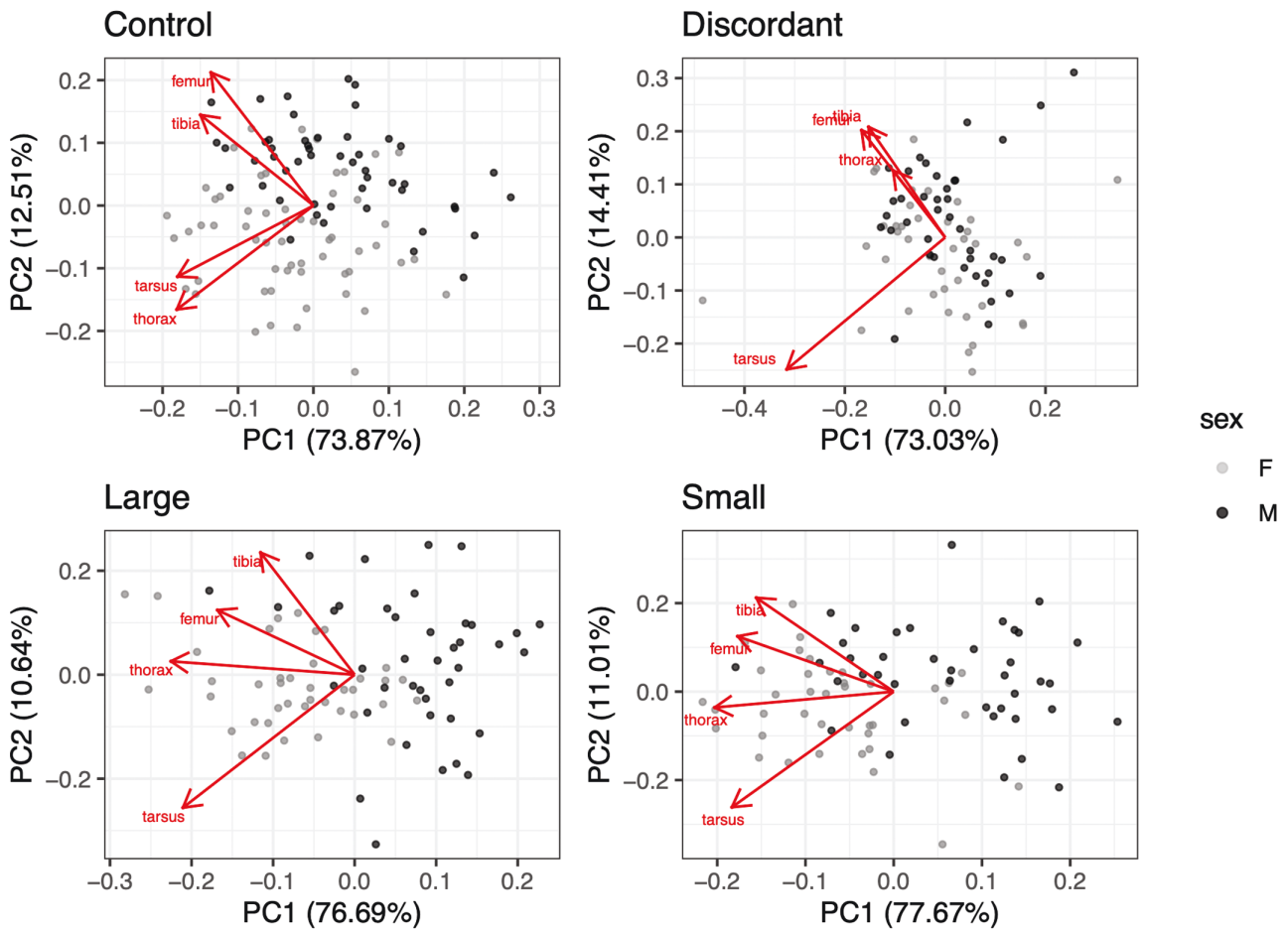

Figure 4. Sexually discordant selection alters patterns of multivariate allometry across sex. Represented as biplots, magnitudes and direction of the loadings for traits are superimposed onto PC1 and PC2. Log_2 transformed length measures were used.

the genome and many regions at or near fixation (Figure 5A). 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 (Supplementary Figures S5 and S6). 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 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). For comparison, we also provide plots of the CMH statistic (Supplementary Figures S7 and S8). To confirm the effects for SNPs of interest, we modeled allele frequency changes with logistic regression between treatments, examining odds ratio VS. F_{ST}, and verifying high F_{ST} correlated to large odds ratios (Supplementary Figures S9, S10, and S11). 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 the body or trait size were also observed for both the discordant and concordant comparisons (Supplementary Files 1, 2, 3 and 4). For the concordant selective lineages, we examined

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

Note. 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. Only the lineages artificially selected discordantly (between the sexes) show substantial changes in multivariate allometry.

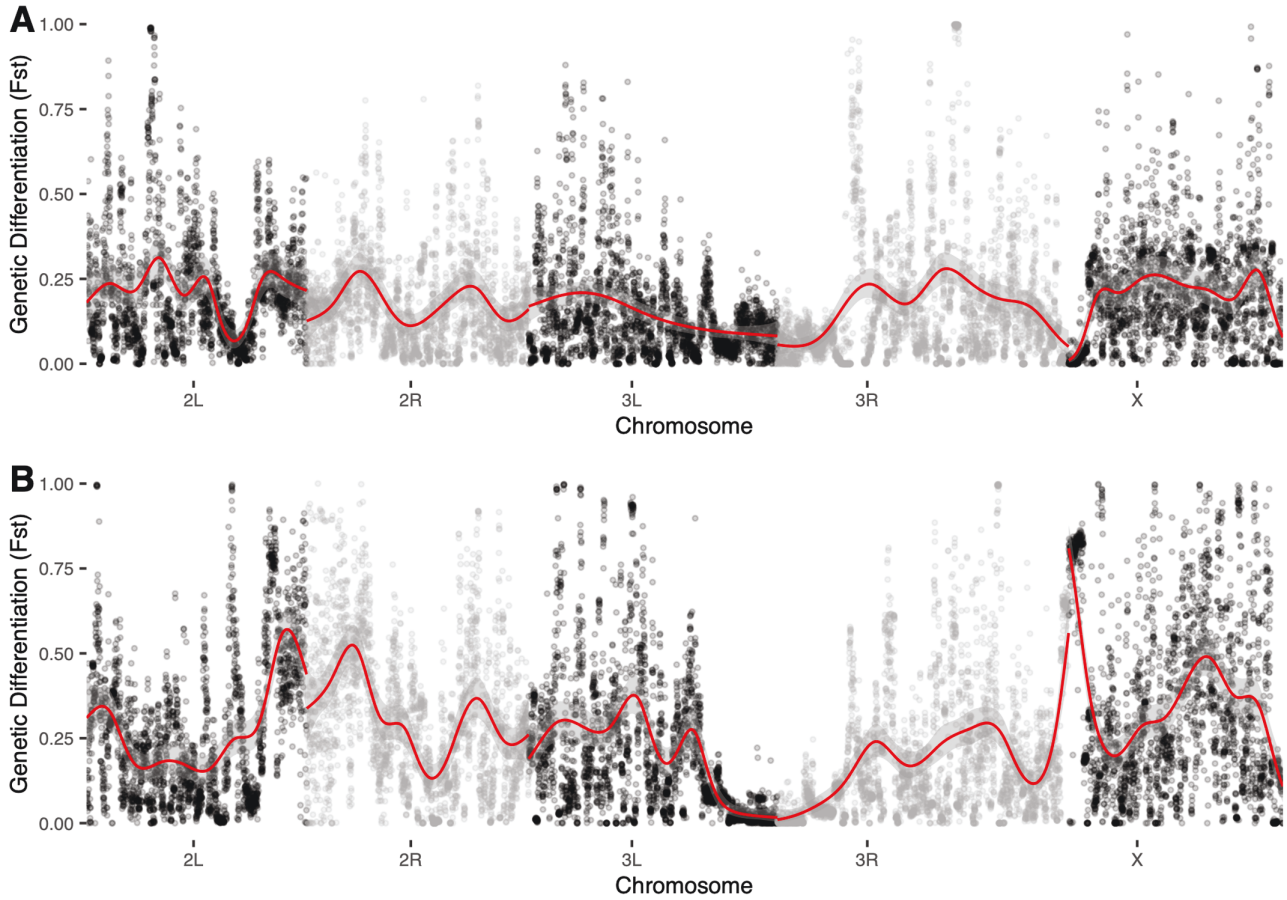

Figure 5. Genomic divergence among artificially selected treatments. Genome-wide F_{ST} (10,000 bp windows). Chromosomal trends for F_{ST} (binomial, gamm) in red. (A) Discordant selection compared to control treatments. (B) Large compared to small treatments.

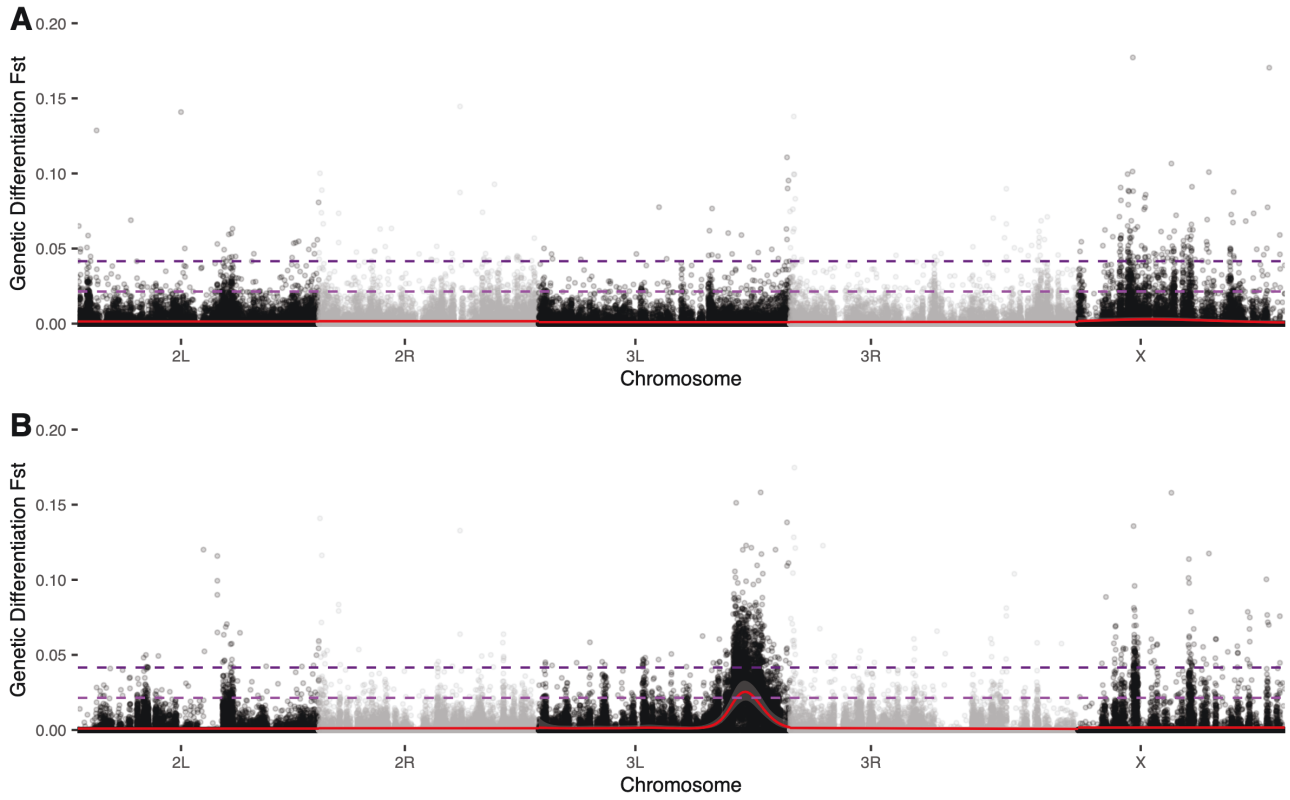

Figure 6. Within-lineage, within-generation, genome-wide F_{ST} (10,000 bp windows) between males and females. Chromosomal trends for F_{ST} (binomial, gamm) in red. Maximum simulated value (neutral evolution, SLiM v4.2) in dark purple, 95th quantile of simulated values in light purple. (A) Male vs. Female for control treatment (replicate lineage 1), showing a trend $\cong 0$, across the genome. (B) Male vs. Female F_{ST} values for sexually discordant treatment (replicate 1) showing a region on chromosome 3L with elevated F_{ST}. Within the region of interest, the maximum simulated value was 0.02, and the 95th quantile of simulated values in that region was 0.001.

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*). Modulation of the expression of Nop17l in developing wing tissues reduces the size of the wing (Bennett et al., 2006). If we examine overlapping genes excluding the logistic regression analysis, 22 additional genes or SNPs overlap between F100 and F378 concordant treatments (Supplementary Files 5 and 6).

The candidate genomic regions identified above were then examined for enrichment of GO terms. Permuted GO enrichment analysis with Gowinda did not return significant terms, except for our large treatment, which only returned the term astral microtubule organization (GO:0030953) as an enriched GO term.

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

Given the long-term nature of discordant selection in this experiment, these lineages may be useful in identifying signatures of IASC. 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 & 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, Supplementary Figures S13 and S14; S1 = 0.0023, S2 = 0.0014, Supplementary Figures S15 and S16). For the discordant selection treatment replicate 2, we also found a mean chromosomal F_{ST} to be near zero (E2 = 0.0011, Supplementary 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 ~3.4 Mb on chromosome 3L between position 18,100,000 and 21,600,000, where mean F_{ST} 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 inversions on chromosome 3L (*In(3L)P*, *3L133in*, *3L165in*, *3L096*, *3L105*, and *3L058*), nor novel structural variants using 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 above the mean between sex F_{ST} (Supplementary File 1).

We confirmed that this region of elevated intersex F_{ST} in discordant replicate 1 was indeed an outlier using several types of simulations. SLiM simulations, as well as a custom simulation designed to account for various sources of sampling variation (Supplementary Figures S18–S26), as well as chromosome-wide simulations accounting for aspects of variation in recombination rates and demographic effects of the populations. All male-female comparisons follow a similar pattern to the simulations, with the exception of discordant replicate 1, which appears extreme relative to all simulations (Supplementary Figure S19). We further modeled all SNPs on chromosome 3L to look for clusters of significant

SNPs within our elevated F_{ST} region (Supplementary Figures S27–S34). Our discordant lineage replicates 1 shows a high number of significant SNPs within our elevated region when modeled (Supplementary Figure S29), 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. Ruzicka et al. (2019) also used LH_M to explore sexual conflict loci, but their LH_M flies have diverged from the LH_M used in this experiment for an unknown number of generations. If we exclude our modeled 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.

Evolved changes in sex ratio and fecundity may suggest conflict in the discordant lineages

Crossing all treatment groups in single-pair matings reciprocally to the LH_M "ancestor" resulted in ~1:1 sex ratio in the F_1. The exception to this was Control replicate 1 male crossed to LH_M females (C1: F-M = 0.455, CI = 0.426–0.485; Figure 7), as well as within-lineage crosses in both discordant treatments (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 the control replicate 1 is also observed in F_2 (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 show male bias, and the confidence intervals for discordant replicate 1 cross do not overlap the 1:1 expectation (E1: F-M = 0.557, CI = 0.521–0.592; Figure 8). We also observed a deviation from the expected sex ratio in our large replicate 1 male to LH_M female cross (L1: M-F = 0.459, CI = 0.429–0.489; Figure 8), and our discordant replicate 1 male to LH_M female cross (E1: M-F = 0.453, CI = 0.426–0.480; Figure 8). This effect in the discordant cross was in the opposite direction of the discordant replicate 1 pure cross.

Next, we explored whether the direction of the cross had an effect on the sex ratio by adding cross-direction as a predictor in the model. In the F_1 generation, the overall direction of the cross had very modest impacts on sex ratios ($\chi^2 = 1.13$, $df = 1$, $p = 0.29$). There was some evidence of deviation from the expected sex ratio when the direction was accounted for when control 1 sired the cross (C1: M-F = 0.455, CI = 0.426–0.485). In the F_2 generation, the direction of the F_0 cross (sire vs. dam) had a modest impact on sex ratios ($\chi^2 = 5.67$, $df = 1$, $p = 0.017$). Specifically, when the treatment lineage was the sire in the initial cross for discordant replicate 1 and large replicate 1 (E1: M-F = 0.453, CI = 0.426–0.480; L1: M-F = 0.459, CI = 0.429–0.489). The control replicate 1 cross appeared to deviate from the expected sex ratio regardless of which parent served as sire or dam.

While it was not a planned experiment, while counting flies for sex ratio crosses, we observed a possible difference in fecundity between treatments. The preliminary results from the analysis from the F_1, when the dam was from either discordant lineages, showed reduced fecundity (Supplementary Figure S35). In the F_2 generation, there did not appear to be a direction of cross-effect, but discordant treatments had the lowest fecundity (Supplementary Figure S36).

Discussion

Sexual dimorphism 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 any hurdles due to a high r_{MF} 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, or be resolved entirely. However, the reintroduction of sex-discordant selection should disrupt this equilibrium and generate additional genomic conflict. Examining the response to sexually discordant selection for body size across the genome after selection provides an opportunity to identify genomic regions undergoing conflict. In this study, we utilize long-term artificially selected lineages (Stewart & Rice, 2018), selected for body size either in a sexually concordant or discordant manner. As discussed in detail below, in addition to observing trait-specific changes in patterns of SSD, we see potential evidence for the maintenance of polymorphisms consistent with unresolved conflict.

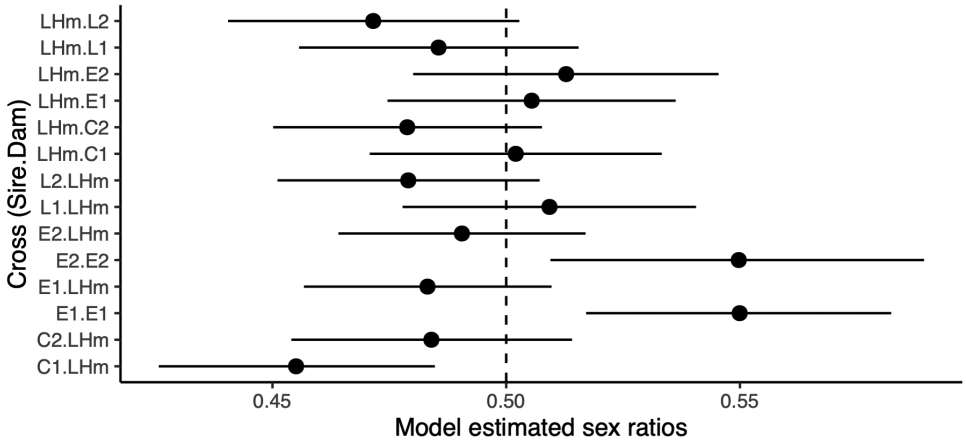

Figure 7. F_1 offspring sex ratios from all treatments crossed to the founder population as well as both discordant lineages crossed "pure." The cross label has Sire on the left and Dam on the right of the cross identifier. Dashed line marks expected 1:1 sex ratio. LHm = LHm population. L = large selection; C = control; E = discordant selection. Numbers following population labels are replicate lineages.

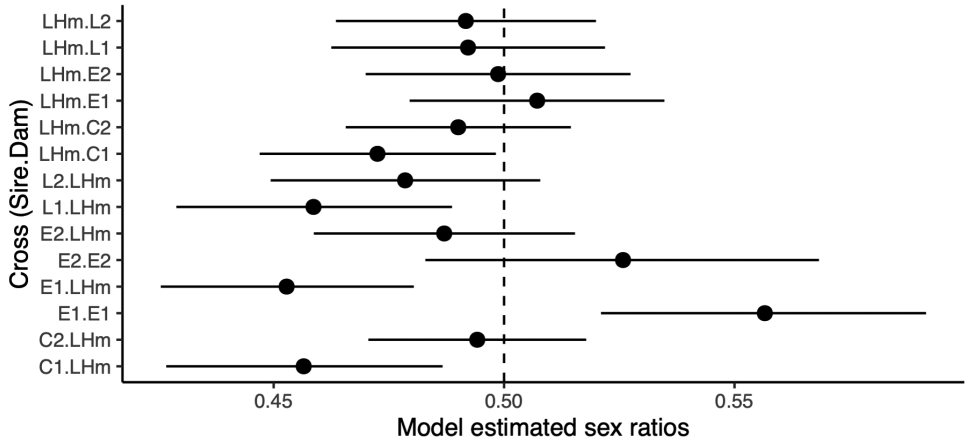

Figure 8. F_2 offspring sex ratios from all treatments crossed to the founder population as well as both discordant lineages crossed "pure." The cross label has Sire on the left and Dam on the right of the cross identifier. The dashed line marks the expected 50/50 sex ratio. Label codes as in Figure 7.

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 for size responded in the expected directions for all traits (Figure 2) and changed size proportionally, maintaining SD for all traits (Figure 3). This proportional response is consistent with relatively high r_{MF} for morphological traits, as previously observed (Cowley & Atchley, 1988; Reeve & Fairbairn, 1996). If r_{MF} had been low in the starting population, we might 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). Within concordant treatments, SSD was largely maintained despite substantial changes in size. For the discordant selection treatment, SD changed trait specifically (Figures 2 and 3). 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 $\bar{F}/\bar{M} \approx 1.12$), while leg traits (femur, tibia, and basitarsus 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). The dimorphism for thorax length reduced to the greatest degree (difference of ~0.012, in $\log_2$ µm, or $\bar{F}/\bar{M} \approx 1.01$), 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 SD tend to have lower r_{MF} than nondimorphic 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 SD when selected indirectly; however, tarsus and thorax (thorax being the most directly selected) did not fully reverse dimorphism.

Within-lineage, across-sex comparisons 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 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 of elevated F_{ST} (Figure 6A; Supplementary Figures 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; Dutoit et al., 2018; Flanagan & Jones, 2017; Lucotte et al., 2016; however see Sylvestre et al., 2023) or the demographics of the populations. Modifying SNP calling methods, F_{ST} estimation procedures, and filtering methods, all yielded the same spike in F_{ST} in the E1 lineage. This elevated region was extreme for this lineage in comparison to simulations accounting for a number of sources of sampling variation and overall allele frequencies as well as chromosome-wide simulations of the population. If it was a spurious peak of elevated F_{ST} as a result of a technical aspect of the pipeline or mis-mapping of reads from a sex chromosome, 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 differences in allele frequencies to identify potential loci under conflict, the selection applied to the discordant lineages in this experiment is both strong and sustained across hundreds of generations. Kasimatis et al. (2019) suggest that for a locus to show this degree of asymmetry, an exceedingly high magnitude of selection is required. In their model, this would require nearly 40% mortality each generation to maintain the distortion in M-F allele frequencies at a locus. The nature of artificial selection as applied in the current experiments means well over 40% (~80% of individuals with the current selection design) of flies in each generation (those not at phenotypic extremes) 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. However, our simulations suggest this is not a likely explanation for the observed results (Supplementary Figures S18–S26). Why this region in discordant treatment replicate E1 shows this distortion and the contribution of intra-locus conflict requires further exploration.

Given the strange peak in this one discordant replicate, we extracted all SNPs three standard deviations above mean F_{ST} on chromosome 3L, which also were significant in the logistic regression (which better accounts for sampling variance). 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 (Supplementary File 1), but multiple genes were 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 showed evidence for genetic differentiation between artificial selection treatments. Many candidate genes have mutational phenotypes listed on flybase as "abnormal size," suggestive that the response to selection is highly polygenic with alleles of small effect "inching" toward a phenotype rather than "sprinting."

Finally, we explored genes found by Ruzicka et al. (2019), which were identified as potentially being in sexual conflict. These genes were considered in conflict based on sex-specific fitness (competitive mating success in males and competitive fecundity in females). An important caveat in our comparison is that although our lineages are from the same initial lab population (LH_M), our treatment groups have gone through nearly 400 generations of strong selection from the ancestral population. It is also unknown how many generations of divergence occurred in the ancestral LH_M populations prior to both experiments. In our comparison between genes identified in Ruzicka et al. (2019) and those identified in our study, only the gene Formin-like (*frl*) overlaps. *Frl* is upstream of our region of interest and does not have any previously identified body size or sex-specific phenotypes. We may not see overlap because Ruzicka et al. (2019) identified conflict alleles based on specific proxies of fitness, whereas ours may reflect more indirect effects due to discordant selection on size. However, we do observe some evidence that adult sex ratios are male-skewed in the discordant lineages and possibly maintained even in crosses to the LH_M "ancestor," as well as potential reductions in fecundity (Figures 7 and 8; Supplementary Figures S35 and S36).

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 of 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, the response is broadly consistent with a polygenic basis, with thousands of SNPs showing evidence of genetic differentiation across the genome. In 378 generations of selection, and more than 125 generations after clear evidence of phenotypic reversal of SSD, the opportunity for a mutation (or ancestrally segregating relatively rare variants) of large phenotypic effect occurring was possible. We find no clear evidence for alleles of large effect that have swept to fixation in parallel across both lineages of discordant selection. The chance of identifying alleles of large effect is attenuated by the impact of drift, making many SNPs fixed 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, identifying a novel mutation that is responsible requires further investigation. This slow crawl toward a response is supported by the fact that Stewart and Rice (2018) saw a negligible phenotypic response to discordant selection during at least the first 100 generations, with more substantial responses occurring after this. 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 (Millington et al., 2021a; Rideout et al., 2015). 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 sex-specific effects 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 effects (d*Myc*, *H*, *InR*, *RcA1*, *sun*). Follow-up on these genes to determine their role in response to discordant selection is required.

We also compared our 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; *discs large 1* (*dlg1*), which has a known increased size phenotype as well as courtship and oogenesis phenotypes, and *smrtr* (*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 LH_M. Optimistically, this could suggest genes under conflict that are stable in the population and are being captured in these studies that explore sexual conflict, but more work must be done to confirm this. Both of these interesting genes are also 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 a previous analysis

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, and (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 reanalyzed the F100 data used by Turner et al. 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 ~25×) in comparison to our current study. Using our pipeline, the F100 data had no genes with known size phenotypes overlapping with our large genes of interest (Supplementary File 5). In the gene list overlapping 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 (Supplementary 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., both analyses found genes in related general pathways such as ecdysone signaling 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 mutations 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).

Our results are both an important exploration of genomic conflict using artificial selection and a follow-up on one of the longest (in generations) artificial selection experiments 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.

Supplementary material

Supplementary material is available online at *Evolution*.

Data availability

All phenotypic data and scripts are available on GitHub (https://github.com/DworkinLab/Audet_etal_Evolution_2024) along with a static copy on DRYAD (https://doi.org/10.5061/dryad.6t1g1jx6k). All raw genomic sequence data is deposited in the NCBI SRA under BioProject PRJNA1107500.

Author contributions

Study conceptualization and funding: I.D.; Study design: I.D.; Artificial selection and rearing: A.D.S.; Dissections: J.K.; DNA extraction: T.A. and K.P.; Image analysis: J.K. and I.D.; Analysis: I.D., T.A., and K.P.; Manuscript drafting: T.A. and I.D.; Manuscript editing: T.A., A.D.S., and I.D.; Manuscript revisions: T.A. and I.D.

Funding

This work was funded by an NSERC Discovery grant to I.D. and an Ontario Graduate Scholarship to T.A.

Conflict of interest: Editorial processing of the manuscript was done independently of I.D., who is an Associate Editor of *Evolution*. The other authors declare no conflict of interest.

Acknowledgments

We thank the associate editor Dr. Charles Baer, the handling editor Dr. Tim Connallon, and two anonymous reviewers for feedback that has improved this manuscript.

References

Alicchio, R., & Palenzona, D. L. (1971). Changes of sexual dimorphism values in *Drosophila Melanogaster*. *Italian Journal of Zoology*, *38*, 75–84.

Arguello, J. R., Laurent, S., & Clark, A. G. (2019). Demographic history of the human commensal *Drosophila melanogaster*. *Genome Biology and Evolution*, *11*(3), 844–854. https://doi.org/10.1093/gbe/evz022

Ashburner, M. (1989). *Drosophila. A laboratory handbook*. Cold Spring Harbor Laboratory Press.

Bates, D., Machler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bennett, D., Lyulcheva, E., Alphey, L., & Hawcroft, G. (2006). Towards a comprehensive analysis of the protein phosphatase 1 Interactome in *Drosophila*. *Journal of Molecular Biology*, *364*(2), 196–212. https://doi.org/10.1016/j.jmb.2006.08.094

Bird, M. A., & Schaffer, H. E. (1972). A study of the genetic basis of the sexual dimorphism for wing length in *Drosophila melanogaster*. *Genetics*, *72*(3), 475–487. https://doi.org/10.1093/genetics/72.3.475

Blackith, R. E., & Reyment, R. A. (1971). *Multivariate morphometrics*. Academic Press.

Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Mächler, M., & Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*, *9*, 378.

Bushnell, B. (2021). *BBTools: BBMap short read aligner, and other bioinformatic tools*. SourceForge.

Carreira, V. P., Mensch, J., & Fanara, J. J. (2009). Body size in *Drosophila*: Genetic architecture, allometries and sexual dimorphism. *Heredity*, *102*(3), 246–256. https://doi.org/10.1038/hdy.2008.117

Cheng, C., & Kirkpatrick, M. (2016). Sex-specific selection and sex-biased gene expression in humans and flies. *PLoS Genetics*, *12*(9), e1006170. https://doi.org/10.1371/journal.pgen.1006170

Chung, Y., Rabe-Hesketh, S., Dorie, V., Gelman, A., & Liu, J. (2013). A nondegenerate penalized likelihood estimator for variance parameters in multilevel models. *Psychometrika*, *78*(4), 685–709. https://doi.org/10.1007/s11336-013-9328-2

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X., & Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly*, *6*(2), 80–92. https://doi.org/10.4161/fly.19695

Comeron, J. M., Ratnappan, R., & Bailin, S. (2012). The many landscapes of recombination in *Drosophila melanogaster*. *PLoS Genetics*, *8*(10), e1002905. https://doi.org/10.1371/journal.pgen.1002905

Cowley, D. E., & Atchley, W. R. (1988). Quantitative genetics of *Drosophila melanogaster*. II. Heritabilities and genetic correlations between sexes for head and thorax traits. *Genetics*, *119*(2), 421–433. https://doi.org/10.1093/genetics/119.2.421

Czech, L., Spence, J. P., & Expósito-Alonso, M. (2023). grenedalf: Population genetic statistics for the next generation of pool sequencing. *arXiv*, 2306.11622.

Delph, L. F., Steven, J. C., Anderson, I. A., Herlihy, C. R., & Brodie, E. D. III (2011). Elimination of a genetic correlation between the sexes via artificial correlational selection. *Evolution*, *65*(10), 2872–2880.

Dutoit, L., Mugal, C. F., Bolívar, P., Wang, M., Nadachowska-Brzyska, K., Smeds, L., Yazdi, H. P., Gustafsson, L., & Ellegren, H. (2018). Sex-biased gene expression, sexual antagonism and levels of genetic diversity in the collared flycatcher (Ficedula albicollis) genome. *Molecular Ecology*, *27*(18), 3572–3581. https://doi.org/10.1111/mec.14789

Eisen, E., & Hanrahan, J. (1972). Selection for sexual dimorphism in body weight of mice. *Australian Journal of Biological Sciences*, *25*(5), 1015–1024. https://doi.org/10.1071/bi9721015

Emlen, D. J., Marangelo, J., Ball, B., & Cunningham, C. W. (2005). Diversity in the weapons of sexual selection: Horn evolution in the beetle genus Onthophagus (Coleoptera: Scarabaeidae). *Evolution*, *59*(5), 1060–1084.

Fairbairn, D. J., Blanckenhorn, W. U., & Székely, T. (Eds.). (2007). *Sex, size, and gender roles: Evolutionary studies of sexual size dimorphism*. Oxford University Press.

Flanagan, S. P., & Jones, A. G. (2017). Genome-wide selection components analysis in a fish with male pregnancy. *Evolution*, *71*(4), 1096–1105. https://doi.org/10.1111/evo.13173

Haller, B. C., & Messer, P. W. (2023). SLiM 4: Multispecies eco-evolutionary modeling. *American Naturalist*, *201*(5), E127–E139. https://doi.org/10.1086/723601. The University of Chicago Press

Honěk, A. (1993). Intraspecific variation in body size and fecundity in insects: A general relationship. *Oikos*, *66*(3), 483–492. https://doi.org/10.2307/3544943

Jolicoeur, P. (1963). 193. Note: The multivariate generalization of the Allometry equation. *Biometrics*, *19*(3), 497–499. https://doi.org/10.2307/2527939

Jónás, A., Taus, T., Kosiol, C., Schlötterer, C., & Futschik, A. (2016). Estimating the effective population size from temporal allele frequency changes in experimental evolution. *Genetics*, *204*(2), 723–735. https://doi.org/10.1534/genetics.116.191197

Kapun, M., Barrón, M. G., Staubach, F., Obbard, D. J., Wiberg, R. A. W., Vieira, J., Goubert, C., Rota-Stabelli, O., Kankare, M., Bogaerts-Márquez, M., Haudry, A., Waidele, L., Kozeretska, I., Pasyukova, E. G., Loeschcke, V., Pascual, M., Vieira, C. P., Serga, S., Montchamp-Moreau, C., … González, J. (2020). Genomic analysis of European *Drosophila melanogaster* populations reveals longitudinal structure, continent-wide selection, and previously unknown DNA viruses. *Molecular Biology and Evolution*, *37*(9), 2661–2678. https://doi.org/10.1093/molbev/msaa120

Kasimatis, K. R., Ralph, P. L., & Phillips, P. C. (2019). Limits to genomic divergence under sexually antagonistic selection. *G3 (Bethesda, Md.)*, *9*(11), 3813–3824. https://doi.org/10.1534/g3.119.400711

Kaufmann, P., Wolak, M. E., Husby, A., & Immonen, E. (2021). Rapid evolution of sexual size dimorphism facilitated by Y-linked genetic variance. *Nature Ecology and Evolution*, *5*(10), 1394–1402. https://doi.org/10.1038/s41559-021-01530-z

Keightley, P. D., Ness, R. W., Halligan, D. L., & Haddrill, P. R. (2014). Estimation of the spontaneous mutation rate per nucleotide site in a *Drosophila melanogaster* full-sib family. *Genetics*, *196*(1), 313–320. https://doi.org/10.1534/genetics.113.158758

Keightley, P. D., Trivedi, U., Thomson, M., Oliver, F., Kumar, S., & Blaxter, M. L. (2009). Analysis of the genome sequences of three *Drosophila melanogaster* spontaneous mutation accumulation lines. *Genome Research*, *19*(7), 1195–1201. https://doi.org/10.1101/gr.091231.109

Klingenberg, C. P. (1996). Multivariate allometry. In L. F. Marcus, M. Corti, A. Loy, G. J. P. Naylor, & D. E. Slice (Eds.), *Advances in morphometrics* (pp. 23–49). Springer US.

Kofler, R., Pandey, R. V., & Schlotterer, C. (2011). PoPoolation2: Identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics*, *27*(24), 3435–3436. https://doi.org/10.1093/bioinformatics/btr589

Kofler, R., & Schlötterer, C. (2012). Gowinda: Unbiased analysis of gene set enrichment for genome-wide association studies. *Bioinformatics*, *28*(15), 2084–2085. https://doi.org/10.1093/bioinformatics/bts315

Kristensen, K., & McGillycuddy, M. (2023). *The AR(1) covariance structure*. The Comprehensive R Archive Network.

Lande, R. (1980). Sexual dimorphism, sexual selection, and adaptation in polygenic characters. *Evolution*, *34*(2), 292–305. https://doi.org/10.1111/j.1558-5646.1980.tb04817.x

Lenth, R., Singmann, H., Love, J., & Buerkner, P. (2018). *emmeans: Estimated marginal means, aka least-squares means*. The Comprehensive R Archive Network.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv*, *1303.3997*.

Li, H., B, H., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). *The Sequence Alignment/Map (SAM) format and SAMtools*. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Lucotte, E. A., Laurent, R., Heyer, E., Ségurel, L., & Toupance, B. (2016). Detection of allelic frequency differences between the sexes in humans: A signature of sexually antagonistic selection. *Genome Biology and Evolution*, *8*(5), 1489–1500. https://doi.org/10.1093/gbe/evw090

Luecke, D. M., & Kopp, A. (2019). Sex-specific evolution of relative leg size in *Drosophila prolongata* results from changes in the intersegmental coordination of tissue growth. *Evolution*, *73*(11), 2281–2294. https://doi.org/10.1111/evo.13847

Mathews, K. W., Cavegn, M., & Zwicky, M. (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*(3), 1215–1228. https://doi.org/10.1534/genetics.116.192260

Millington, J. W., Brownrigg, G. P., Chao, C., Sun, Z., Basner-Collins, P. J., Wat, L. W., Hudry, B., Miguel-Aliaga, I., & Rideout, E. J. (2021a). Female-biased upregulation of insulin pathway activity mediates the sex difference in *Drosophila* body size plasticity. *eLife*, *10*, e58341. https://doi.org/10.7554/eLife.58341

Millington, J. W., Wat, L. W., Sun, Z., Basner-Collins, P. J., Brownrigg, G. P., & Rideout, E. J. (2021b). A low sugar diet enhances *Drosophila* body size in males and females via sex-specific mechanisms. *Physiology*, *149*(6), dev200491. https://doi.org/10.1242/dev.200491

Moczek, A. P., Cruickshank, T. E., & Shelby, A. (2006). When ontogeny reveals what phylogeny hides: Gain and loss of horns during development and evolution of horned beetles. *Evolution*, *60*(11), 2329–2341.

Niku, J., Hui, F. K. C., Taskinen, S., & Warton, D. I. (2019). gllvm: Fast analysis of multivariate abundance data with generalized

linear latent variable models in r. *Methods in Ecology and Evolution*, *10*(12), 2173–2182. https://doi.org/10.1111/2041-210x.13303

Poissant, J., Wilson, A. J., & Coltman, D. W. (2010). Sex-specific genetic variance and the evolution of sexual dimorphism: A systematic review of cross-sex genetic correlations. *Evolution*, *64*(1), 97–107. https://doi.org/10.1111/j.1558-5646.2009.00793.x

Prasad, N. G., Bedhomme, S., Day, T., & Chippindale, A. K. (2007). An evolutionary cost of separate genders revealed by male-limited evolution. *The American Naturalist*, *169*(1), 29–37. https://doi.org/10.1086/509941

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*(6), 841–842. https://doi.org/10.1093/bioinformatics/btq033

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rausch, T., Zichner, T., Schlattl, A., Stütz, A. M., Benes, V., & Korbel, J. O. (2012). DELLY: Structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics*, *28*(18), i333–i339. https://doi.org/10.1093/bioinformatics/bts378

Reeve, J. P., & Fairbairn, D. J. (1996). Sexual size dimorphism as a correlated response to selection on body size: An empirical test of the quantitative genetic model. *Evolution*, *50*(5), 1927–1938. https://doi.org/10.1111/j.1558-5646.1996.tb03580.x

Reeve, J. P., & Fairbairn, D. J. (1999). Change in sexual size dimorphism as a correlated response to selection on fecundity. *Heredity*, *83*(Pt 6), 697–706. https://doi.org/10.1046/j.1365-2540.1999.00616.x. Nature Publishing Group

Rice, W. R. (1984). Sex chromosomes and the evolution of sexual dimorphism. *Evolution*, *38*(4), 735–742. https://doi.org/10.1111/j.1558-5646.1984.tb00346.x

Rice, W. R., & Chippindale, A. K. (2001). Intersexual ontogenetic conflict. *Journal of Evolutionary Biology*, *14*(5), 685–693. https://doi.org/10.1046/j.1420-9101.2001.00319.x

Rice, W. R., Linder, J. E., Friberg, U., Lew, T. A., Morrow, E. H., & Stewart, A. D. (2005). Inter-locus antagonistic coevolution as an engine of speciation: Assessment with hemiclonal analysis. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(Suppl. 1), 6527–6534. https://doi.org/10.1073/pnas.0501889102

Rideout, E. J., Narsaiya, M. S., & Grewal, S. S. (2015). The sex determination gene transformer regulates male-female differences in drosophila body size. *PLoS Genetics*, *11*(12), e1005683. https://doi.org/10.1371/journal.pgen.1005683

Ruzicka, F., Dutoit, L., Czuppon, P., Jordan, C. Y., Li, X. -Y., Olito, C., Runemark, A., Svensson, E. I., Yazdi, H. P., & Connallon, T. (2020). The search for sexually antagonistic genes: Practical insights from studies of local adaptation and statistical genomics. *Evolution Letters*, *4*(5), 398–415. https://doi.org/10.1002/evl3.192

Ruzicka, F., Hill, M. S., Pennell, T. M., Flis, I., Ingleby, F. C., Mott, R., Fowler, K., Morrow, E. H., & Reuter, M. (2019). Genome-wide sexually antagonistic variants reveal long-standing constraints on sexual dimorphism in fruit flies. *PLoS Biology*, *17*(4), e3000244. https://doi.org/10.1371/journal.pbio.3000244

Sawala, A., & Gould, A. P. (2017). The sex of specific neurons controls female body growth in Drosophila. *PLoS Biology*, *15*(10), e2002252. https://doi.org/10.1371/journal.pbio.2002252

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, *9*(7), 671–675. https://doi.org/10.1038/nmeth.2089

Shingleton, A. W., Estep, C. M., Driscoll, M. V., & Dworkin, I. (2009). Many ways to be small: Different environmental regulators of size generate distinct scaling relationships in Drosophila melanogaster. *Proceedings Biological Sciences*, *276*(1667), 2625–2633. https://doi.org/10.1098/rspb.2008.1796

Smit, A. F. A., Hubley, R., & Green, P. (2013–2015). *RepeatMasker Open-4.0*. http://www.repeatmasker.org

Spitzer, K., Pelizzola, M., & Futschik, A. (2020). Modifying the Chi-square and the CMH test for population genetic inference. *The Annals of Applied Statistics*, *14*(1), 202–220.

Stewart, A. D., & Rice, W. R. (2018). Arrest of sex-specific adaptation during the evolution of sexual dimorphism in *Drosophila*. *Nature Ecology and Evolution*, *2*(9), 1507–1513. https://doi.org/10.1038/s41559-018-0613-4

Sylvestre, F., Mérot, C., Normandeau, E., & Bernatchez, L. (2023). Searching for intralocus sexual conflicts in the three-spined stickleback (*Gasterosteus aculeatus*) genome. *Evolution*, *77*(7), 1667–1681. https://doi.org/10.1093/evolut/qpad075

Taus, T., Futschik, A., & Schlötterer, C. (2017). Quantifying selection with pool-seq time series data. *Molecular Biology and Evolution*, *34*(11), 3023–3034. https://doi.org/10.1093/molbev/msx225

Testa, N. D., & Dworkin, I. (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. *Development Genes and Evolution*, *226*(3), 159–171. https://doi.org/10.1007/s00427-016-0534-7

Tigreros, N., & Lewis, S. M. (2011). Direct and correlated responses to artificial selection on sexual size dimorphism in the flour beetle, *Tribolium castaneum*. *Journal of Evolutionary Biology*, *24*(4), 835–842. https://doi.org/10.1111/j.1420-9101.2010.02221.x

Turner, T. L., Stewart, A. D., Fields, A. T., Rice, W. R., & Tarone, A. M. (2011). Population-based resequencing of experimentally evolved populations reveals the genetic basis of body size variation in *Drosophila melanogaster*. *PLoS Genetics*, *7*(3), e1001336. https://doi.org/10.1371/journal.pgen.1001336

Wickham, H. (2018). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York.

4. The role of resource defensibility in facilitating sexually selected weapon evolution: An experimental evolution test

Audet, T., Wilson, A., Dukas, R., & Dworkin, I. (2025). The role of resource defensibility in facilitating sexually selected weapon evolution: an experimental evolution test. *Evolution*, qpaf073.

Evolution, 2025, **XX(XX)**, 1–17
https://doi.org/10.1093/evolut/qpaf073
Advance access publication 9 April 2025
Original Article

The role of resource defensibility in facilitating sexually selected weapon evolution: an experimental evolution test

Tyler Audet[1,†], Audrey Wilson[1,†], Reuven Dukas[2], Ian Dworkin[1]

[1]Department of Biology, McMaster University, Hamilton, Ontario, Canada
[2]Cognitive Ecology Group, Department of Psychology, Neuroscience and Behaviour, McMaster University, Hamilton, Ontario, Canada
Corresponding author: Department of Biology, McMaster University, 1280 Main Street W, Hamilton, Ontario, Canada L8S 4L8. Email: dworkin@mcmaster.ca
[†]Co-first author contribution.

Abstract

Animal weapons have evolved multiple times, primarily for battling for access to mates. Despite intra-sexual selection being common, exaggerated weapons have evolved relatively rarely. So why do exaggerated weapons not evolve more commonly? It has been hypothesized that three conditions are necessary for evolution of exaggerated weapons: high variance in reproductive success, patchy, high-value resources, and spatial environments conducive to one-on-one competition. Here, we test this hypothesis by performing experimental evolution in *Drosophila melanogaster*, utilizing heterogeneous environments where conditions facilitating territorial defense and opportunities for competitive interactions vary. We examine changes in sexually dimorphic morphology and male aggression that are predicted to occur, based on this model. We also examine whether condition dependence for sexual dimorphism has evolved after 35 and 75 generations of experimental evolution. Aggression did increase, albeit modestly, in environments that facilitate resource defense. Morphological changes are modest although with some trait-specific changes to allometry, generally in the opposite direction of our predictions. Condition dependence trends in the opposite direction from those predicted by our hypothesis as well. We discuss our results in the context of the necessary conditions for the evolution of exaggerated weapons, and if, and when condition dependence may evolve.

Keywords: weapon evolution, sexual selection, experimental evolution, trait exaggeration, *Drosophila melanogaster*

Introduction

The intensity of sexual competition has long been attributed to the accessibility of a limiting sex (typically females) by a less limiting sex (typically males; Bateman, 1948; Darwin, 1871). In large part due to the consequences of anisogamy, including female receptivity and parental investment, the Operational Sex Ratio (OSR) of populations can often be male-biased. This male-biased OSR leads to various forms of competition for access to mates including sperm competition, mate harm, and intra-sexual aggression, which can result in a high variance in the number of matings individual males acquire relative to females in the population (Bath et al., 2021; Nandy et al., 2013; Sepil et al., 2022). The ability to differentially access mates can become more intense depending on the mating system of the organism in question, as well as ecological constraints (Emlen & Oring, 1977). The Environmental Potential for Polygamy (EPP) has been hypothesized to depend on the ability to defend multiple mating partners, or resources desired by multiple mates, by an individual (Emlen & Oring, 1977).

In mating systems where EPP is high, and males compete for access to mates, intra-sexual selection may be intense and resource-defense polygyny may create a skew in which males must fight for access to mates, favoring aggressive interactions. If the resource is sufficiently high value, males with high Resource-Holding Potential (RHP) should be more willing to escalate agonistic interactions (Hurd, 2006). A pattern has also been present in the literature suggesting species with exaggerated weapons tend to be more aggressive than closely related species, or individuals of the same species without the trait exaggeration (Boisseau et al., 2020; Kudo et al., 2017; Moczek & Emlen, 2000). To date, to the best of our knowledge, the evolution of aggression in relation to the evolution of weapon exaggeration has yet to be studied. It is known that males without obvious trait exaggeration, such as the pomace fly, *Drosophila melanogaster*, can be aggressive and show resource defense/territoriality (Chen et al., 2002; Dow & Schilcher, 1975; Guo & Dukas, 2020; Hoffmann & Cacoyianni, 1990). This includes both the use of threatening signals, such as wing displays, and physical altercations mediated by use of the front legs and heads (details discussed in the *Methods* section), which themselves are genetically correlated (Baxter et al., 2023). *Drosophila melanogaster* selected for increased territoriality, showed increased mating success and longevity under some conditions (Hoffmann & Cacoyianni, 1989). Other *Drosophila spp*, such as Hawaiian *Drosophila* show trait exaggeration (Spieth, 1981), in particular, hypercephaly, lekking behavior, and substantial aggression (Kudo et al., 2017). *Drosophila prolongata* has exaggerated male forelegs used in male-male combat, and outcomes of contests influence mating success (Toyoshima & Matsuo, 2023). Although the native mating substrates for *D. prolongata* are

Received July 12, 2024; revisions received March 17, 2025; accepted April 1, 2025

Associate Editor: Jessica Abbott; Handling Editor: Tim Connallon

currently unknown, the intensity of aggression in this species suggests that aggressive interactions are a necessary precursor for changes in male-biased sexual dimorphism. Increased competition for mates and associated increases in variance in reproductive success may lead to the evolution of increased aggression and contests, and the exaggeration of traits used in aggressive interactions.

The evolution of exaggerated weapons is far rarer than the occurrence of male-male competition for mating (Palaoro & Peixoto, 2022; Voje, 2016). The relatively rare evolution of exaggerated weapons begs the question as to why, when, and how weapon exaggeration evolves. Building on the foundations laid out in Emlen and Oring (1977), Emlen (2008, 2014) hypothesized that three explicit conditions are necessary for the precursors of exaggerated weapon evolution in males. First, there must be competition for access to females, likely in a way that creates asymmetry in access to mates, generating increased variance in male reproductive success. This may result in males who expend resources into trait (weapon) expression, as investing in increased RHP may be crucial for reproductive success. Second, there are limiting, localized (patchy) resources required by females. If resources are distributed abundantly throughout the environment, there is little benefit in defending one patch if there is a plethora nearby of equal value that females may visit instead. Discrete patches of limiting resources result in predictable locations that females must visit, and therefore specific locations to defend. Finally, the layout of these resources must be such that males compete in duels, or one-on-one fights. If resource patches are sufficiently large (spatially) resource defense may be impossible, as many males attempt to usurp the dominant male at the same time. In this scenario, competition becomes a scramble and there is likely no direct benefit in being the strongest because a scramble does not necessarily reward the largest weapon, but rather the fastest male to secure a mating. These conditions correlate with observations of extant species with exaggerated weapons (Emlen & Philips, 2006). For instance, *Onthophagus spp.* (dung beetles) use two strategies to sequester dung for their larvae; they may roll dung away from the source and bury it elsewhere, or they may dig tunnels adjacent to the dung source. If dung is rolled away, males scramble to fight over the dung ball largely out in the open, and many males may compete at once. Tunnels, however, restrict access to dung in a way where males interact in one-on-one competition for access to the female who requires the dung for egg laying. Phylogenetically, only in lineages where males interact in these restricted spaces allowing one-on-one duels have horns evolved, and they have never evolved when competition occurs as a scramble (Emlen & Philips, 2006). Diospidae (stalk-eyed flies) also have species with exaggerated male eyestalks used in aggressive signaling, as well as species with only rudimentary eyestalks. Sexually dimorphic species form nocturnal clusters on rootlets where males are able to control access to multiple females, and due to the linear nature of the rootlets, interactions between males occur one-on-one with the larger male typically winning (Wilkinson & Dodson, 1997). The monomorphic species appear to not display the same clustering behavior that allows male-male competition for access (Wilkinson & Dodson, 1997). Although the conditions laid out by Emlen (2014) correspond to weapon evolution in some taxa, little has been done to experimentally test if these ecological conditions are necessary and/or sufficient to initiate weapon evolution, or contribute to increased male-biased sexual size dimorphism (SSD).

To test Emlen's (2008; 2014) hypothesis for the ecological conditions necessary for weapon evolution, we performed experimental evolution using *D. melanogaster*. We generated three experimental environments, including two conditions conducive for males to attempt to defend food resource desired by females (food optimized to maximize female fecundity). These defensible resources should result in male-male competition for monopolization of the resource and an increase in mating success of males with high RHP. Size of resource patches was based on prior research that demonstrated that when resource access (via food patch size) was limited, male *D. melanogaster* would increasingly perform resource defense (Hoffmann & Cacoyianni, 1990). We generated three environments where individuals had spatially constrained access to resource patches, potentially facilitating increased one-on-one contests in males. One where patches were easily accessible and too large to easily defend, one where patches occurred in sizes conducive to resource defense attempts, and one facilitating the opportunity for one-on-one contest, via restricted openings leading to resource (Figure 1; Supplementary Figure S1). Previous work using this setup demonstrated that depending on the spatial constraint of resources and opportunities for sexual selection, there was variation in efficacy of selection to purge deleterious alleles (Wilson et al., 2021). Based on the hypothesis set out by Emlen (2008, 2014), and the fact that male *D. melanogaster* extensively use their forelegs (prothoracic) in agonistic interactions (Chen et al., 2002), we predicted evolution of increasingly male-biased SSD for these legs, and an evolutionary increase in the allometric slope of this leg (relative to overall body size). Specifically, we predicted rank order changes based on three experimental environments we set up. We predicted a related response in wing length, as wings are used in threat displays. In the environment without small, defensible resource patches (large open resources, which single males cannot defend, termed 'no territory', NT, hereafter), scramble competition tends to dominate (Hoffmann & Cacoyianni, 1990). When resources are readily available in relatively large patches, scramble competition is classically believed to be the most common mating system in *D. melanogaster* (Hoffmann & Cacoyianni, 1990; Partridge et al., 1987; Spieth, 1974) with recent work suggesting that interference competition may represent a substantial fraction of mating interactions (Baxter et al., 2018). We did not predict any substantial evolutionary changes in SSD or trait allometries in this environment, as a result of either of these mating strategies. In the second environment, with open, but defensible resource patches, termed "unconstrained territories" (UCT), we predicted modest evolutionary increases in aggression, evolution towards male-biased SSD and positive allometry in the legs. In the final environment, with spatially constrained access to resource patches, termed 'spatially constrained territories' (SCT), we predicted increased magnitude of evolutionary changes in aggression, SSD, and positive allometry. We also evaluated traits for evidence of the evolution of increased condition dependence in male forelegs and wings relative to other traits. An association between trait exaggeration and overall condition has been suggested to result in a heightened condition-dependent response to environmental perturbation in exaggerated secondary sexual traits, meaning that when conditions are poor the trait reduces in size to a greater degree than other non-secondary sexual traits (Rowe & Houle, 1996). This pattern has been observed in most systems where it has been studied

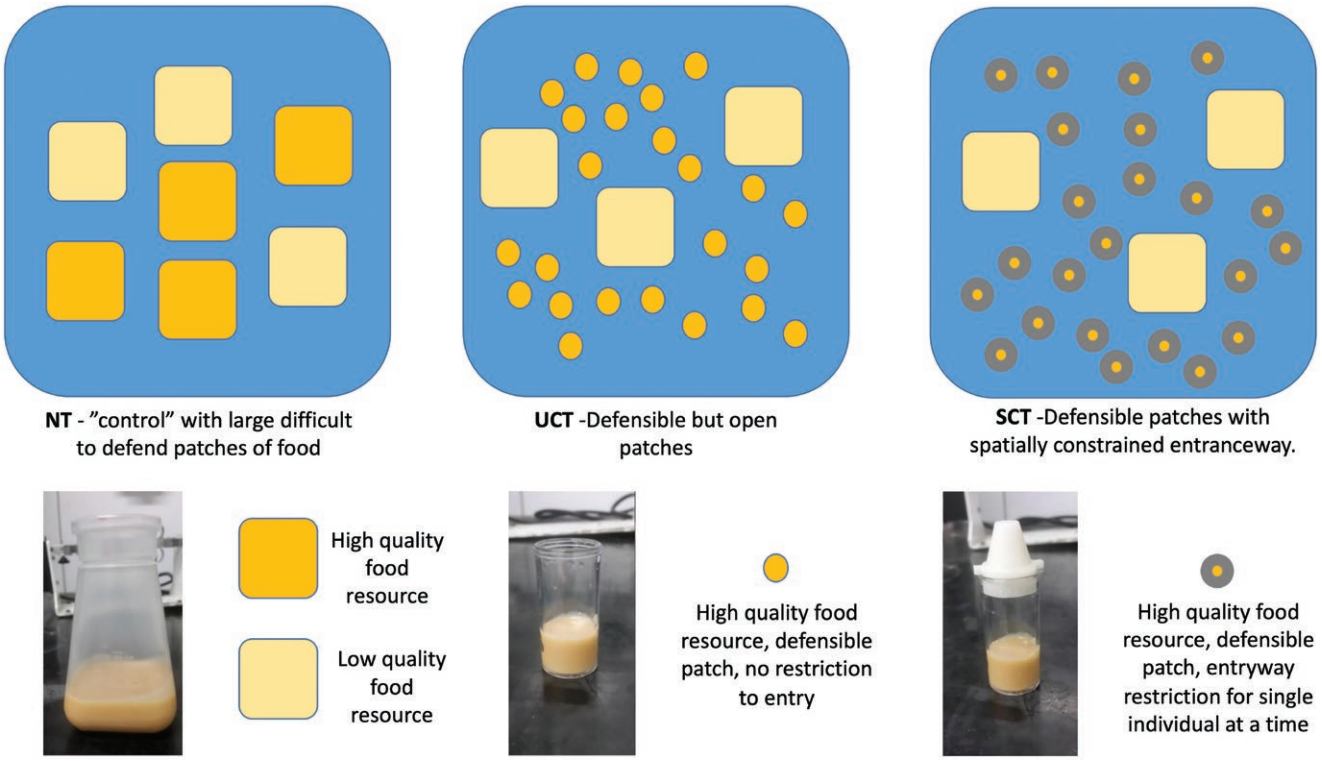

Figure 1. Top-down layout of high- and low-quality food resource structure in each territorial treatment. Upper row: Large squares represent open, easily accessible food, small circles represent food containers of a size *D. melanogaster* has been demonstrated to defend, and small circles with a grey ring represent resources with restricted access points. Darker objects represent high-quality food and lighter objects represent diluted food. Placement of food containers in each cage was random each generation. Lower row: side view shows the food containers and accessibility of resource. Smaller vials were decreased further in height after generation 10 to reduce the amount of space within the vial for flies to occupy (see *Methods* section).

(Bonduriansky, 2007; David et al., 2000; Johns et al., 2014), with few exceptions (although see: Ceballos & Valenzuela, 2011; Fairbairn, 2005; Perdigón Ferreira et al., 2023). For this reason, it could be anticipated that the evolution of condition dependence co-evolves with trait exaggeration. We predicted rank order differences in condition dependence across the three experimental environments (NT < UCT < SCT). While we observed modest changes in male aggression in the predicted direction, our results relating to changes in SSD, allometry, and condition dependence were not generally consistent with our predictions. We discuss our findings in the context of the evolution of trait exaggeration.

Methods

Environmental treatments

Three environmental treatments were created to encapsulate the ideas of the three conditions for weapon evolution proposed by Emlen (2014). All three environments had the same surface area of high and low-quality food, and approximately equal volume of food. The "high-quality" resources have been optimized for female fecundity based on nutritional geometry studies (Jensen et al., 2015; Lee et al., 2008; Maklakov et al., 2008; Reddiex et al., 2013; Tatar, 2011), and were developed as highly desirable resources for female oviposition. Fecundity selection is an important driver of overall fitness in *D. melanogaster* (Tanaka & Yamazaki, 1990), and evidence supports that it is likely one of the major contributors to the ancestral pattern of female-biased SSD (Honěk, 1993; Reeve & Fairbairn, 1999). As such, female access to the high-quality food contributes to an individual's ability to maximize fitness. "Low-quality" resources are dilutions of the high-quality resource, and for all three experimental environments, provided as four 177 ml *Drosophila* culture bottles filled with ~50 ml of a 25% dilution of the high-quality resource. At generation 48, the low-quality resource was changed to a 10% dilution, as it was observed that flies were possibly beginning to adapt to this dilution, evidenced by a few normal-sized larvae present in the diluted media. The purpose of these diluted food resources was to allow adult individuals to feed (even if the media could not generally support larval growth). In this way, individuals were not competing for resources necessary for adult survival, but for resources necessary to maximize reproduction. We note that we regularly observed individuals in each population using low-quality resources for feeding, and some egg laying, but that larvae developing on them were small, developmentally delayed, and did not eclose as adults fast enough to contribute to the following discrete generation.

While all experimental environments shared the same amount (based on surface area) and proportions of high- and low-quality food, they differed based on patchiness, accessibility, and defensibility of resources (Figure 1). The "non-territory" (NT) treatment was designed to have large, open, easily accessible food patches that are difficult to defend as territories (Hoffmann & Cacoyianni, 1990). NT effectively serves as a "control" environment, mimicking conditions of the lab-domesticated population in spatial structure, and where scramble competition in *Drosophila* dominates (Hoffmann & Cacoyianni, 1990). The "unconstrained territory" (UCT) treatment was designed to have small, defensible,

but accessible patches. That is, while they are defensible, they are easy to access by multiple rival individuals, harder to monopolize, and potentially less likely to result in frequent duels. Territorial behavior at this patch size has been previously demonstrated in laboratory conditions by Hoffmann and Cacoyianni (1990), and in natural populations by Dukas (2020). The third environment, the "spatially constrained territory" (SCT) treatment, has the same patch sizes and numbers as UCT but access to the food patch is constrained via 3D-printed "funnel caps" on each resource patch, restricting the opening (accessibility) to the resource (see Supplementary Figure S1; and Wilson et al., 2021). These funnel caps are 3D-printed plastic conical fittings that cover the opening of short vials, sitting on top with a beveled edge leading to a smaller entrance at the peak with a 4 mm opening. The addition of the funnel cap facilitates opportunities for defending and holding resources and provides additional physical spaces that could encourage agonistic encounters for control of the resource to occur as one-on-one battles. The 4 mm aperture is large enough that two males could pass each other, but small enough that one male can harass and attempt to restrict access to other males. Each environmental treatment was set up in mesh BugDorm-4F3030 cages (30 cm^3) with the specific setup as follows: the NT treatment had four *Drosophila* culture bottles (177 ml, 5.5 cm length and width for a total surface area of 30.25 cm^2 for a total of 121 cm^2 total high-quality food surface area) each containing ~50 ml of the high-quality food resource with four drops of a yeast and orange juice mixture placed on top to attract females (Dweck et al., 2013). The UCT treatment had 25 open vials (height of 32 mm, 25 mm outer diameter, ~22 mm inner diameter, 4.8 cm^2 surface area for a total of 120 cm^2 of high-quality food surface area), close to the optimal 20 mm diameter that promotes resource-defense polygyny (Hoffmann & Cacoyianni, 1990), each with a single drop of yeast-orange juice paste on the food surface. The SCT treatment had the same set-up as UCT (including surface area and volume of resources), except each vial had a 3D-printed funnel cap (22 mm diameter, 25 mm height, and 4 mm opening) to restrict access to the vial (resource patch) with a smaller entrance. Vial heights were reduced for UCT and SCT treatments from 95 to 32 mm at generation 10 of experimental evolution to reduce the amount of space between the top of the funnel caps and the surface of the resource in the SCT treatment. This was done to increase defensibility of the resources as initial monitoring of these vials showed high adult densities. Pipe cleaners were wrapped around the tops of bottles and vials containing high-quality food resources as perching sites. The outline of the territorial treatments and food containers can be seen in Figure 1 and Supplementary Figure S1.

Experimental evolution population maintenance

Populations were created by collecting virgin females and males from a large outbred, lab-domesticated population, initiated from a large collection (several thousand individuals) from Fenn Valley Winery (FVW), Michigan in 2010 (GPS co-ordinates: 42.578919, −86.144936). This population had adapted to lab conditions for ~160 generations, prior to the initiation of this experiment. Thus, confounding effects of concurrent selection for lab adaptation would be minimized (Harshman & Hoffmann, 2000). From this population, 300 males and 300 females were placed into a cage, set up with one of the three environmental treatments. This was done with four replicate cages for each treatment, resulting in 12 lineages total (four independent lineages per treatment). Populations were maintained at 12L:12D cycles at 21 °C with 60% relative humidity in a Conviron walk-in chamber (CMP6050). The populations were kept on a 13- to 15-day schedule depending on emergence times, such that each population had about an equal amount of adults contributing to the next generation (census size was not measured directly). After the initial populations were placed into their respective treatments, adults were allowed to mate and lay eggs for 3 days. After this period, the media with eggs and larvae was removed from these cages and placed into new cages (without adults) to allow for development and eclosion. Development and eclosion occurred over a 10- to 12-day period. Once the new generation of adults emerged, old food was removed, and new food was placed in these cages with the set-up described above, and the cycle was repeated. This timeframe was used, as it was too short for the emergence of the rare individuals who developed on the low-quality resource (which had very few pupae regardless). For each generation, the new food was placed into the cages in a random distribution, and the cages were placed onto racks in a random order, such that each population varied in position in the walk-in chamber each generation.

Assessment of male competitive fitness across the environmental treatments

To assess how potential spatial constraints and opportunities for territoriality interact to influence variance in male mating success, we performed an experiment to assess male competitive fitness. We predicted that males reared on high-quality resources would have increased competitive fitness (compared with the common competitor) than those deprived of food during their terminal growth period (and are therefore smaller), and these differences would increase with opportunity for territoriality. For this experiment (summer 2020), we used the "ASW" population of *Drosophila melanogaster* established from 600 field-collected females during the spring and summer of 2018 at various sites near Hamilton, Ontario, Canada, and maintained a census size above 2,000 individuals each generation. For more details about this population, please see (Scott et al., 2022). The spontaneous X-linked *crossveinless*^{1} (*cv*^{1}) mutation was introgressed into this population to serve as a visible marker to assess competitive fertilization. This visible mutation was chosen as previous work in the lab demonstrated that it had relatively modest deleterious effects in comparison to many visible markers. To manipulate male quality, we placed 50 eggs onto high-quality food and allowed high-quality flies to develop to eclosion at 24 °C before collecting virgin flies, while low-quality males were removed from the food 2 days prior to pupation as described below. After eclosion, flies were stored in individual vials prior to the experimental manipulations. One individual focal ASW male of either high- or low-quality, and one *cv*^{1} male were then aspirated into a test cage (355 ml plastic containers) with high-quality resource corresponding to the treatments described above as well as an open, low-quality resource patch (high-quality food diluted to 25%), sufficient for adult feeding and hydration, but insufficient to support proper growth. NT treatments contained a 1 oz plastic cup with ~35 mm surface area of high-quality food, UCT contained a ~22 mm surface area vial of high-quality food, and SCT contained the same food vial as the UCT treatment

but with the restricted funnel cap with a 4 mm opening as described above. Focal and cv^1 males were allowed to settle into their environment for 2 hr before a virgin homozygous *crossveinless*^{1} female was introduced. Because of the X-linked recessive nature of cv^1, females sired by a cv^1 male would be phenotypically crossveinless, and females sired by the focal ASW males would be phenotypically wild-type. Offspring from each treatment were allowed to develop and then collected for phenotyping to count the number of phenotypically crossveinless and wild-type female offspring.

Condition manipulation

At generations 35 and 75 of experimental evolution, two 177 ml *Drosophila* food culture bottles containing high-quality food were placed in each environmental treatment after the initial 3-day egg-laying period for population maintenance. These bottles were removed after 7 hr to keep egg density low and were kept at 21 °C. Upon emergence of adults, 20–25 pairs were placed in three containers per replicate with a 2% apple-juice agar plate with a drop of orange juice yeast paste on the surface. Eggs were collected and placed into vials containing high-quality food resource at a density of 50 eggs per vial. For each of the 12 populations, 16 vials of eggs were collected and were split into three condition cohorts to undergo food deprivation protocol (Stillwell et al., 2011). The purpose of the food deprivation was to manipulate organismal condition, generating size differences, by limiting the nutritional content available to the larvae during growth phases of development. The first condition cohort (0) has normal food availability throughout larval development. Condition cohorts 2 and 1 each have successively increased days of food restriction before the end of larval development, with cohort 1 spending 1 day before the end of larval development without food, and cohort 2 with 2 days without food. Condition cohort 0 consisted of four vial replicates and developed on food for 6 days, cohort 1 consisted of five vial replicates and was left to develop on food for 5 days, and cohort 2 consisted of seven vial replicates and was left to develop on food for 4 days. After these time periods, the larvae from cohorts 1 and 2 were removed by adding 5 ml of a 40% sucrose solution to each vial and shaking for 20 min. Once the larvae were loose from the food, they were collected using a fine paintbrush and placed into a new vial containing a water-moistened cotton ball. The larvae continued development at 21 °C and upon eclosion and sclerotization, 50 individuals of each sex amongst all vials from each condition cohort and population were collected and stored in 70% ethanol for morphometric measurements.

Morphological measurements

Traits chosen for morphological measurement are based on previous research demonstrating their involvement in aggressive interactions. Primary among these is the foreleg (prothoracic leg) which has been shown to be involved with numerous aspects of aggressive behaviors such as thrusting, boxing/fencing, and lunges (Chen et al., 2002; Dow & Schilcher, 1975; Dukas, 2020; Rohde et al., 2017). In addition to measurement of the forelegs, we also measured thorax length as a proxy for body size, as well as head width and wing length. Wing threats are used in aggressive displays (Chen et al., 2002), and there is evidence of a genetic correlation between displays and fighting in *Drosophila* (Baxter et al., 2023). However, there is no evidence that wings are used as weapons directly. As such, wing length would potentially differ in its response in contrast to the legs, which are directly used in physical combat. While the majority of aggressive interactions are among males (Jacobs, 1960), females do sometimes display agonistic interactions with one another, often associated with the defense of a high-quality food resource (Nilsen et al., 2004; Ueda & Kidokoro, 2002). Current evidence is not consistent with the outcome of female-female contests resulting in winner–loser hierarchies (Nilsen et al., 2004). While they share some of the same aggressive behaviors with males, their frequency of these differs substantially, and additionally will use "headbutting" (Nilsen et al., 2004; Zwarts et al., 2012), rarely seen in males. As such, head width was also included as a trait in our study.

Of the flies collected, 20 individuals of each sex of each cohort and treatment combination were dissected for imaging and subsequent measurement. Flies were dissected and images of the head, thorax, wing, and foreleg were taken with a Leica M125 stereoscope with a Leica DFC400 digital camera at magnifications of 50× or 63×, depending on the trait. Measurements of head width, thorax, wing length, wing width, femur, tibia, and first tarsal segment were conducted using ImageJ (1.53e) software (Rueden et al., 2017).

Aggression assays

To assess aggression, at generation 60 we removed 20 females from each treatment and replicated after the 4 days of territorial exposure. These females were allowed to lay eggs in two vials on 2 consecutive days, and density was controlled by culling excess eggs. We sexed newly eclosed males, placed them individually in food vials, and conducted the aggression assays when the males were 3 days old. Assay methodology is described in Baxter and Dukas (2017) and summarized here. Two males of the same treatment and lineage were placed in arenas 3 cm in diameter with a patch of standard food 1.3 cm in diameter and a 3 mm ball of yeast and grapefruit juice. After the two males were added to the arenas, they were video-recorded for 15 min using Logitech c920 cameras. We ran eight trials per lineage for a total of 32 trials for each of the three treatments. BORIS software (Friard et al., 2016) was used to score the footage with observers blind to fly treatment. Observers recorded wing threats, single male aggression (lunging or holding), and reciprocal male aggression (boxing or tussling).

Statistical analyses

All analyses were done in R version 4.1.3. Response variables and the continuous predictor of thorax length were $\log_2$ transformed. The predictor variable of $\log_2$ (thorax length) was mean-centred to aid model interpretation. Linear mixed models were fit using the glmmTMB version 1.1.4 package in R (Brooks et al., 2017). Both intercept and influence of thorax length were allowed to vary as random effects of replicate lineage nested within evolutionary treatment (thorax| Replicate). A similar model was also fit with starvation cohort as a predictor to determine if an interaction between the allometric coefficient and cohort existed to control for changes in allometry due to our starvation protocol. Plotting of observations identified several possible outliers, so models were run with and without outliers. Estimates were found to be similar, so outliers were included. Confidence intervals and a priori, custom contrasts used for inferences, were determined using emmeans version 1.8.0 (Lenth et al., 2018).

Data evaluating competitive fertilization success was modeled using a logistic generalized linear mixed model in glmmTMB with the counts of wild-type ("successes") and crossveinless ("failures") female offspring sired from a focal male as the response variable, and environmental treatment, male quality, and their interaction as predictor variables. We also included random effects of experimental block and cage. We used emmeans to extract estimates and confidence intervals for treatment contrasts to our control treatment (NT). *p*-values from these estimates were adjusted using the Dunnett X method for two tests.

Data to evaluate changes in aggression were also modeled in glmmTMB, with evolutionary treatment and observer as fixed effects, while lineage nested within treatment, day of experiment, and camera were modeled as independent random effects. Counts of lunges were modeled as Poisson with zero inflation. Threat duration is semi-continuous with zeros, and as such was modeled according to a Tweedie distribution (Tweedie power parameter estimated as ≈ 1.6). To confirm that the results (and the presence of many zeroes in the threat duration) were not unduly impacting model inferences (specifically treatment contrasts), we fit a similar model to the one above, but using a hurdle-Gamma, using the zero-inflated Gamma distribution in glmmTMB. This modeling strategy showed similar patterns of changes in aggression, of more modest magnitude. We used a log link for these models. Visualization was done using ggplot2 (Wickham, 2018).

Results

Territorial restriction leads to increased fertilization success in high-quality males relative to non-territorial controls

To test how our territorial treatments influence variance in male reproductive success, we challenged both high- and low-quality focal males against marked tester males (crossveinless) in each territorial treatment. Consistent with our prediction, we observed an increase in competitive fertilization success in high-quality males in the SCT treatment relative to the NT treatment (Figure 2; odds ratio (OR) of 4.25, 95% CIs: 1.54–11.7, *SE*: 1.937, *Z*-ratio: 3.17, $p = 0.003$). Consistent with the prediction there was a modest increase in siring success in the UCT treatment relative to the NT treatment (Figure 2; OR: 3.09, CIs: 1.01 –9.46, *SE*: 1.557, *Z*-ratio: 2.24, $p = 0.048$). As expected, the differences in the environmental treatments had very modest influence on competitive siring success with low-quality focal males (SCT/NT OR: 0.98, CIs: 0.34–2.79, *SE*: 0.46, *Z*-ratio: –0.05, $p = 0.99$; UCT/NT OR 0.24, CIs: 0.06–0.89, *SE*: 0.142, *Z*-ratio: –2.41, $p = 0.03$). We also conducted an analysis of deviance (type II Wald χ^2) for the interaction term between environmental treatment and male quality, demonstrating that the magnitude of difference in siring success across territorial treatments was greater in the high-condition males ($\chi^2 = 8.78$, $df = 2$, $p = 0.012$, Supplementary Table S1).

Sexual size dimorphism (SSD) modestly changed for some traits after 75 generations of experimental evolution

To test our prediction that territorial restriction would induce changes in sexual dimorphism due to intra-sexual competition, we measured female–male sexual dimorphism at generation 35 and 75 in all traits (Figure 3; Supplementary Figures S2 and S3). We also modeled changes in SSD between condition cohorts as well as treatment (Figure 4). We observed no substantial changes in sexual dimorphism for any leg trait at either generation 35 (Figure 3A) or generation 75 (Figure 3B). At generation 75 we saw a change in SSD as a response to condition in the wing (Figure 4B). This effect appears to be due to a decrease in size of female wings. A similar effect was observed in femur length with a change in SSD when condition was accounted for (Figure 4B), which appeared to be due to a decrease in female femur size in UCT and SCT treatments in high condition but converging on similar trait values in low condition (Figure 4B). We also observed a change in SSD in tarsus when condition is accounted for (Figure 4B). Interestingly, this appears to be due to a lower condition response in tarsus in both sexes in UCT and SCT treatment, which is the opposite of the predicted trend (Figure 4B). A model with generation (and associated interactions) as predictors of sexual dimorphism for each trait contrasted between UCT or SCT and NT showed similar results when generation is accounted for (Supplementary Figure S4).

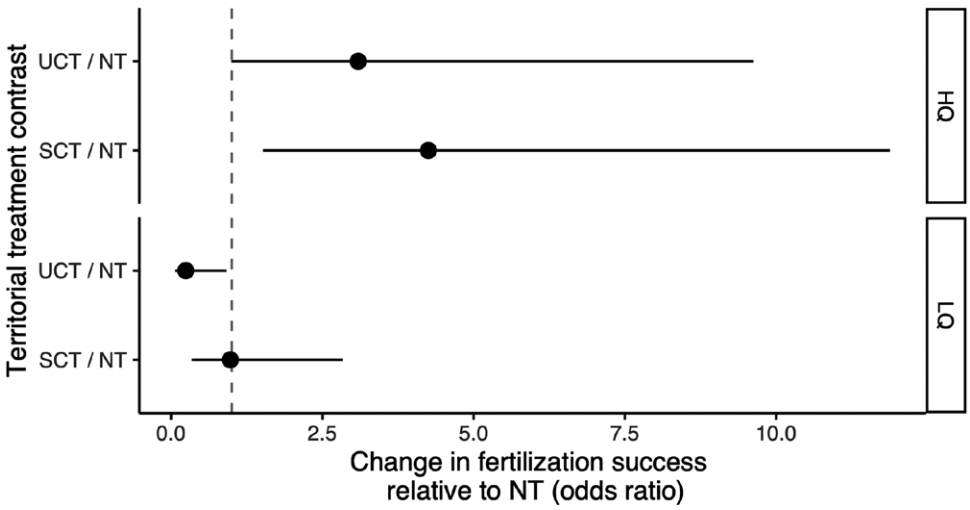

Figure 2. Odds ratio change between the territorial SCT and UCT treatments relative to the control NT treatment for both high- and low-quality male fertilization success against a crossveinless competitor. Model estimates plotted with 95% confidence intervals. HQ = high-quality males, LQ = low-quality males.

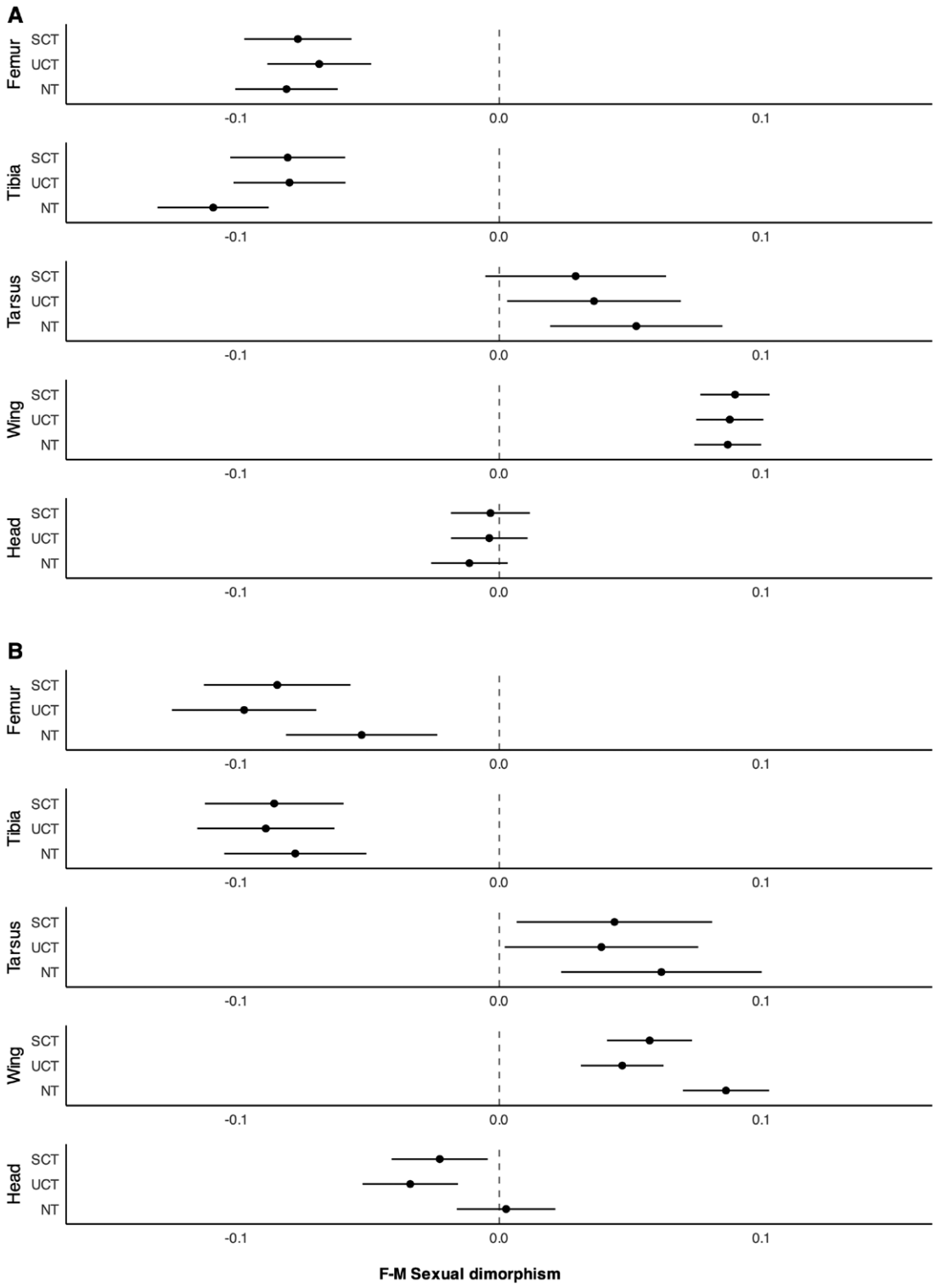

Figure 3. Trait- and treatment-specific sexual size dimorphism, for fully fed individuals, measured as difference between estimates of female–male trait sizes for (A) generation 35 and (B) generation 75. Response variables were $\log_2$ transformed. Model estimates plotted with 95% confidence intervals.

After 75 generations of experimental evolution, allometry changes were modest, and in the opposite direction of predictions

To test the prediction that the two territorially restricted treatments (UCT and SCT) would result in the evolution of increased allometric slopes (forelegs ~ thorax) relative to NT, due to intra-sexual competition among males, we modeled the allometric slope for all measured traits at both generation 35 and generation 75 (Figure 5). At generation 35, the magnitude of changes in allometric relationships was modest for any trait relative to thorax size (Figure 5A; Supplementary Figure S5). Female head width in the SCT treatment had a lower allometric slope relative to NT, but with no concordant response observed in males (Figure 5A; Supplementary Figure

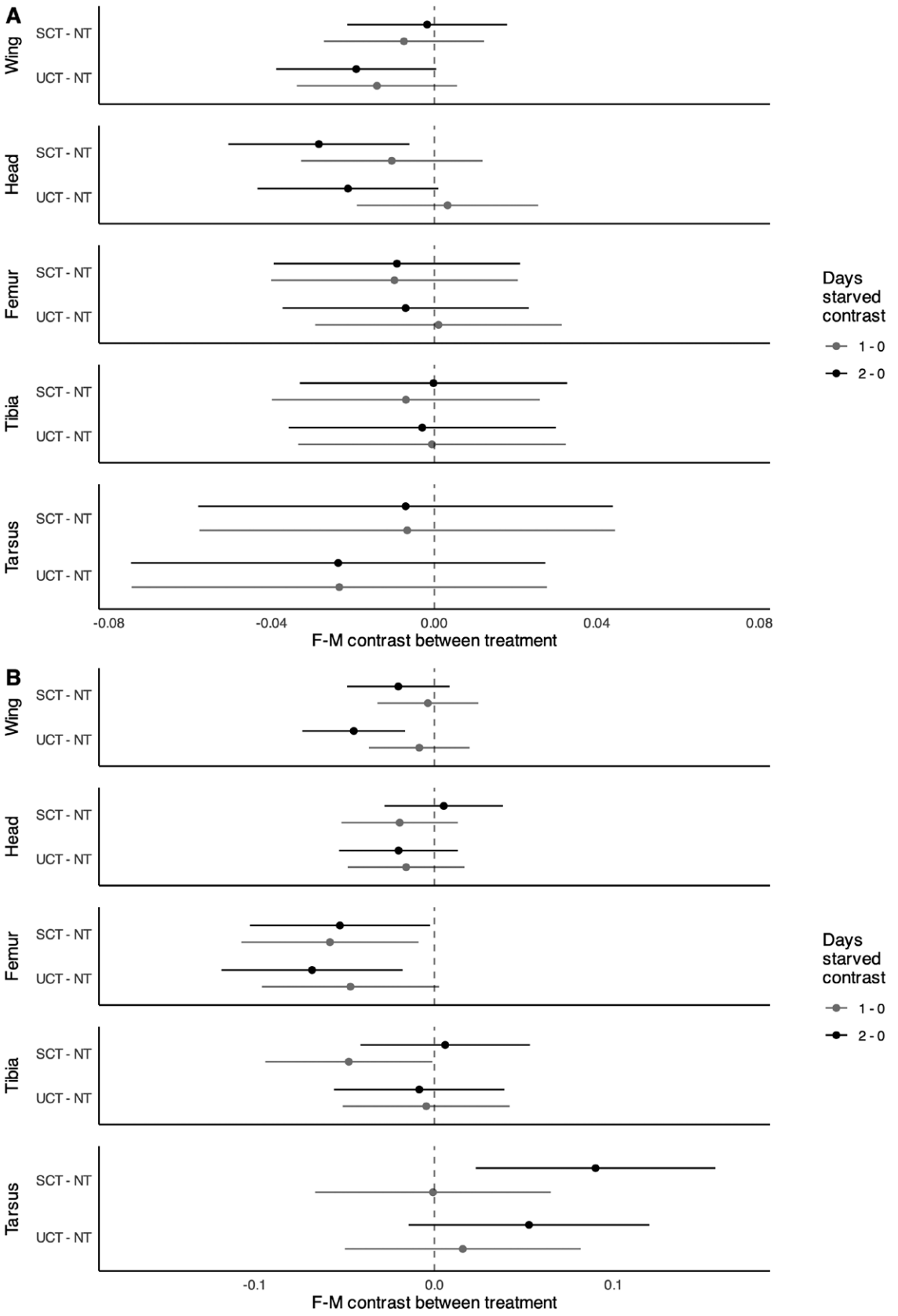

Figure 4. Trait-specific changes among evolutionary treatments in sexual size dimorphism, relative to the NT treatment, measured as female trait size–male trait size for (A) generation 35 and (B) generation 75. Dimorphism was both contrasted between our non-restricted (NT) treatment and both territorial treatments (UCT and SCT) and food restriction condition treatment was contrasted with the non-restricted, high-quality treatment within each territorial treatment. Response variables were $\log_2$ transformed. Model estimates plotted with 95% confidence intervals.

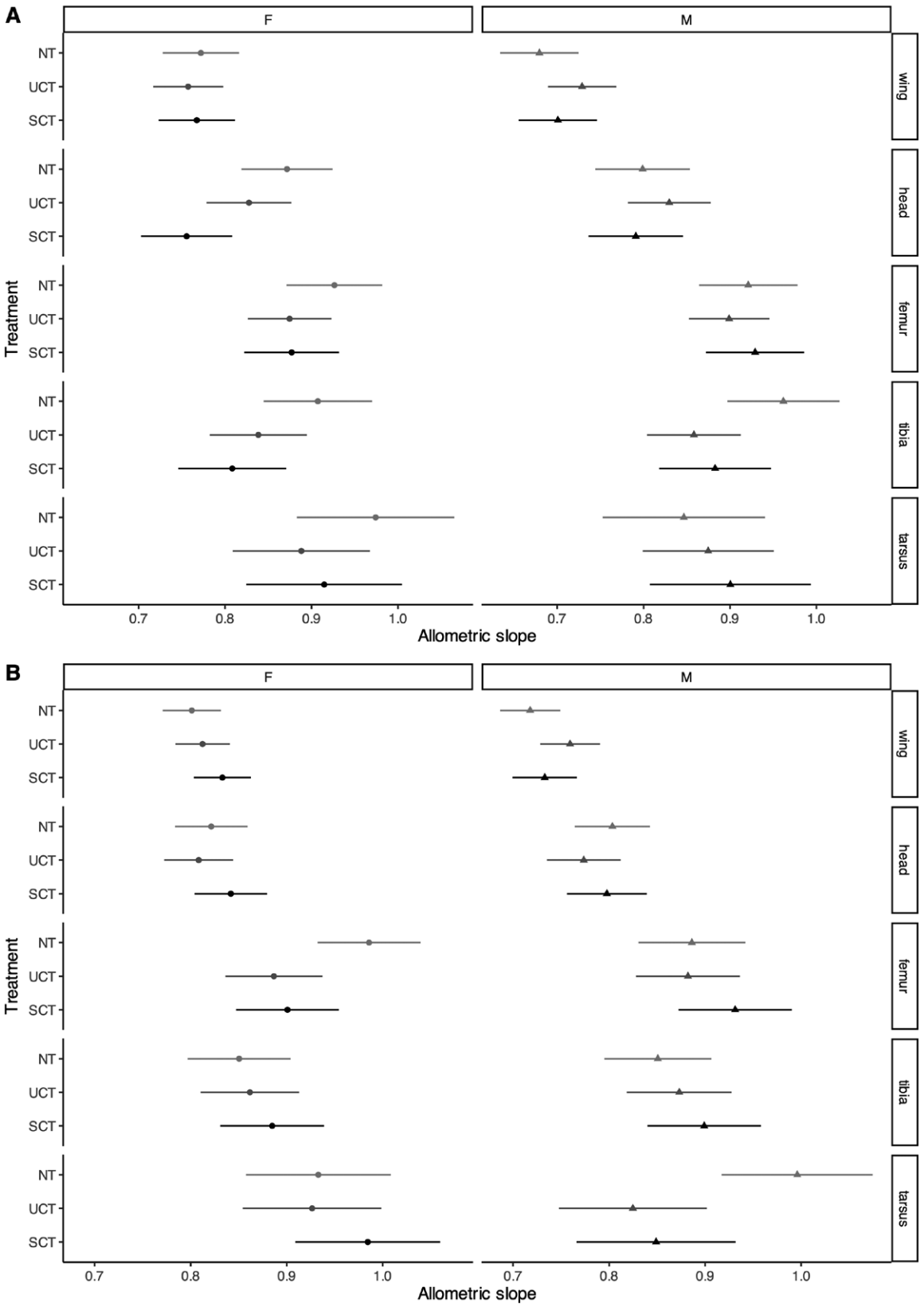

Figure 5. Allometric slopes for territorial treatments in each sex and each trait obtained from model estimated slopes of the response of trait size with $\log_2$ transformed thorax length as a predictor for (A) generation 35 and (B) generation 75. Bars represent 95% confidence intervals for model estimated slopes based on $\log_2$ transformed responses.

S5). There was also some modest evidence for an interaction between sex, treatment, and head allometry in generation 35 ($\chi^2 = 6.61$, $df = 2$, $p = 0.04$). At generation 75 there were still no major evolutionary changes in the allometric slope for most traits (Figure 5B; Supplementary Figure S6). The allometric slope of femur seems to have decreased in females for both the SCT and UCT treatments relative to NT, but not in males (Figure 5B; Supplementary Figure S6), and there was a slightly significant sex by treatment by allometry interaction ($\chi^2 = 7.34$, $df = 2$, $p = 0.03$). The allometric slope of tarsus decreased in males of both UCT and SCT treatments, but not in females. To look at potential changes in thorax size relative to controls, we modeled thorax length for both generations 35 and 75 (Supplementary Figures S7 and S8). We did not observe substantial difference between thorax size between our territorial treatments (SCT, UCT) and control (NT) within generation.

After 75 generations, condition dependence in the tarsus of males decreased, with other traits showing inconsistent and modest responses

To test our prediction that an increase in condition dependence occurs concordantly with increases in sexual selection and trait exaggeration, we measured trait size responses to food deprivation (during development) in each treatment. At generation 35, all treatments and sexes showed the expected decrease in size with food deprivation (Figure 6A; Supplementary Figure S2). However, the slope of the response to manipulating condition shows very modest changes between treatments for most traits (Figure 6A). At generation 75, the overall reduction in size with deprivation is still observed (Figure 6B; Supplementary Figure S3). The response is slightly less compared to generation 35 (Figure 6A), but these experiments were done 2 years apart, performed by two separate individuals, so we do not recommend directly comparing between generations. The slope of the condition dependence appeared to increase for multiple traits in both sexes (a possible reduction of condition dependence) relative to NT (Figure 6). This increase was observed for wing size for both sexes, there was also an interaction between sex, treatment, and condition in the wing, this third-order interaction appears to be due to the UCT condition contrast between the sexes ($\chi^2 = 8.30$, $df = 2$, $p = 0.016$; Supplementary Table S2). There also appeared to be a minor decrease in condition dependence in the head in females. In the tarsus there was an interaction between sex, treatment, and condition dependence, this third-order interaction appears to be due to the SCT condition contrasts between the sexes ($\chi^2 = 7.09$, $df = 2$, $p = 0.03$; Supplementary Table S3). And both sexes seemed to be less condition-dependent in the UCT treatment relative to the NT treatment (Figure 6). Males were also less condition-dependent in the SCT treatment relative to NT (Figure 6).

Modest increases in aggression in flies evolving in territorial treatments

In an experiment performed at generation 60, there was an increase in the duration of threatening wing displays in males in the UCT and SCT treatment (Figure 7A). There was also an increase in the number of lunges engaged by males of the UCT treatment relative to the NT treatment (Figure 7B). These responses both appear to be due to increases in the number of rare, highly aggressive males (Figure 7).

Discussion

According to the hypothesis laid out by Emlen (2008; 2014), the environmental precursors required for the evolution of exaggerated weapons include: the possibility for males to defend a limiting resource required by females, intra-sexual competition favoring larger males or larger weapons, and these defendable and patchy resources are located in a way that facilitates one-on-one contests between males in which there is a winner and a loser. These three requirements should facilitate resource defense polygyny and differential reproductive success that is conducive to larger males who bear the largest weapons siring the bulk of the offspring. The experimental manipulations used in this study were developed to simulate these conditions in laboratory settings and to partition some of the salient features regarding spatial constraints in access to the resources from overall defensibility and patchiness. The goal was to simulate conditions similar to what is observed in natural systems with resource defense polygyny, for instance, the guarding of tunnel entrances by dung beetle males (Emlen, 1997), tree trunk fissures guarded by males and used as oviposition sites in antler flies (*Protopiophila litigate*; Dodson, 1997), or the competition among males for overhanging rootlets with oviposition sites in stalk-eyed flies (Wilkinson & Dodson, 1997). We demonstrate that these environmental conditions were sufficient to increase variance in male siring success with increasing territorial restriction, as predicted (Figure 2).

We predicted evolutionary increases in the allometric slopes in the legs of *D. melanogaster* males, generally used for intra-sexual combat (Chen et al., 2002; Dow & Schilcher, 1975; Dukas, 2020; Rohde et al., 2017), in our UCT and SCT treatments, due to increased opportunity for territorial defense and increased mating success for males who defend a territory. While we did observe "statistically significant" evolutionary changes in morphology, broadly speaking, the results from our experiment were not consistent with these predictions. We did not observe consistent evolutionary increases in male-biased dimorphism in the legs, allometric coefficients did not evolve substantially, and decreased for some traits, in all treatments compared to our non-territorial (NT) control treatment (Figure 5). Recent work has argued that "pure" weapons (used solely for combat), but not for threat signaling, may not be under selection for increases in allometric slope, whereas "threat signals" would evolve such a response (McCullough & O'Brien, 2022). In addition to its locomotory role, the legs of *D. melanogaster* are used in physical combat, but not to our knowledge as a threat signal. However, this weapon-signal continuum suggests aggressive signaling traits, such as wings in *D. melanogaster*, would be under selection for increased allometric slope, which we also do not observe (Figure 5; Chen et al., 2002; McCullough & O'Brien, 2022). We highlight the importance of working from pre-established hypotheses and predictions in experimental evolution studies, as these experiments are conducive to evolutionary changes (as we observed), potentially unrelated to the purpose of the experiment.

Exaggerated weapons often display sensitivity to condition proportionally greater than non-exaggerated traits, an increase in condition dependence (Bonduriansky & Day, 2003; Rowe & Houle, 1996). This increased condition dependence in exaggerated traits has been hypothesized to occur as a signal of overall quality. Larger weapons are a possible

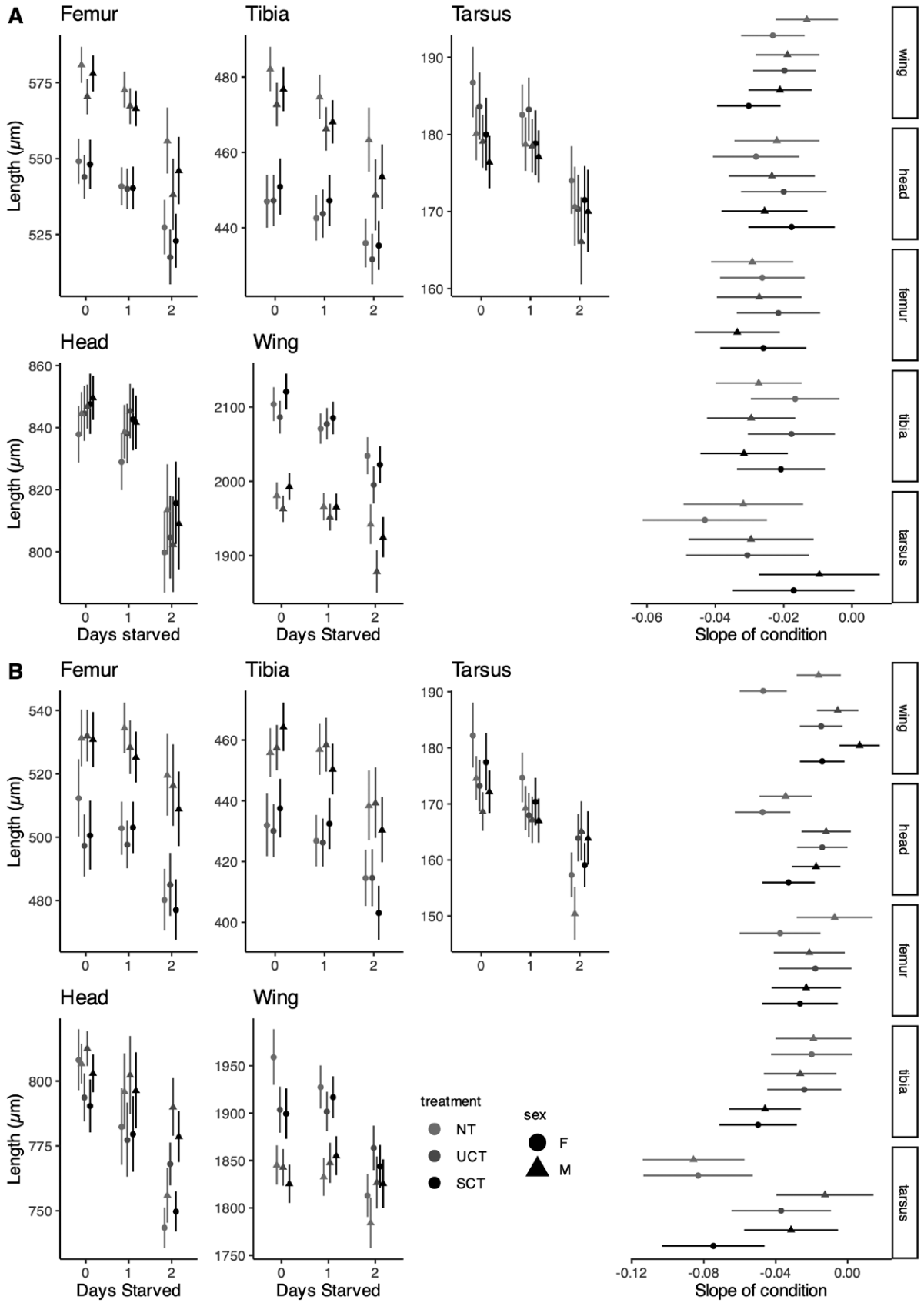

Figure 6. Condition response in trait size by treatment and sex and slope of response to condition by trait for (A) generation 35 and (B) generation 75. Trait size was $\log_2$ transformed for model fit, and back transformed for plotting. Estimates for trait size were run with condition as an ordinal predictor and used only for plotting, slope of condition was estimated with condition treatment as a continuous variable and was used for size and condition inferences.

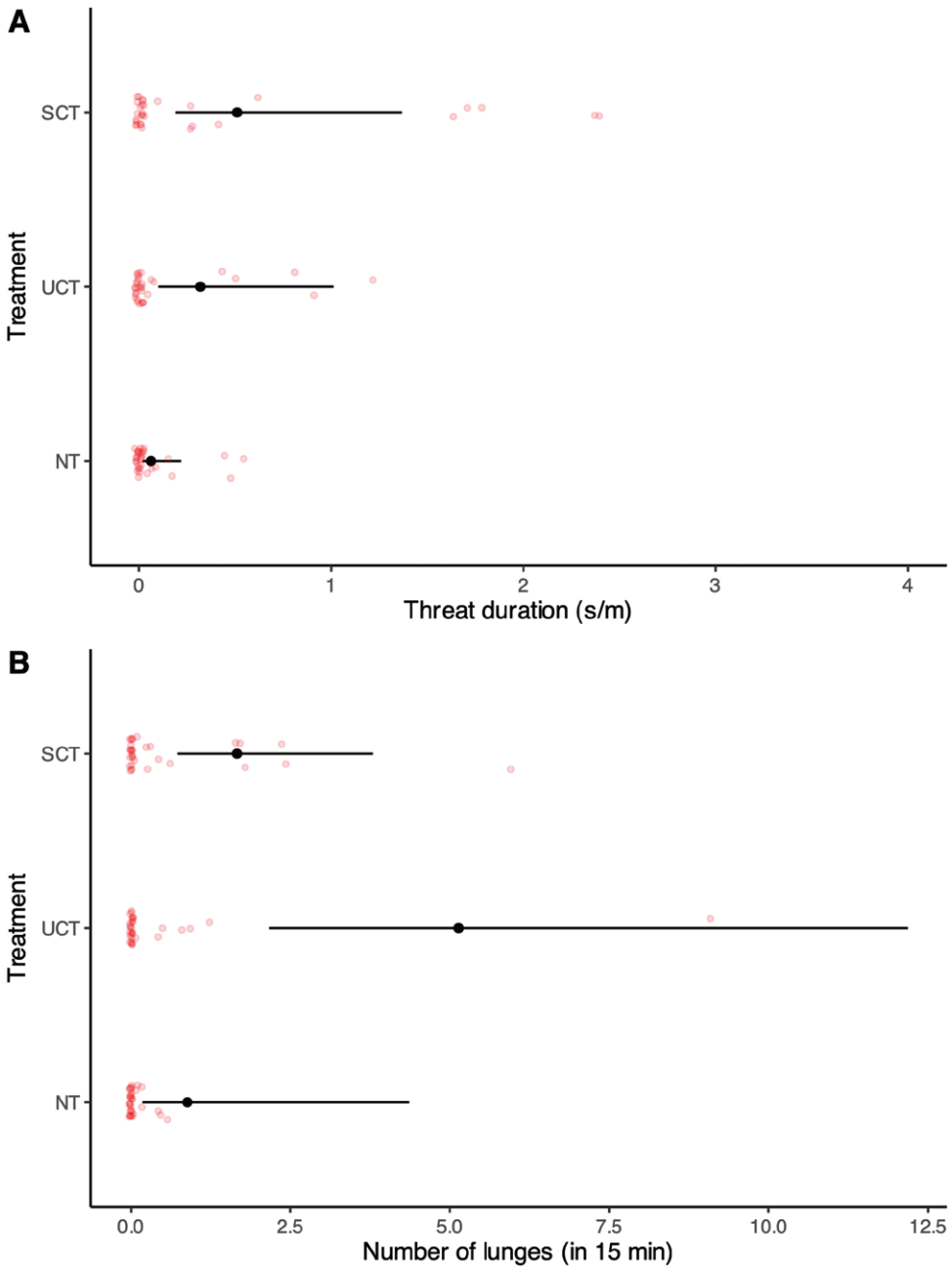

Figure 7. Estimated marginal means with 95% confidence intervals for aggression assays. (A) Estimated duration of threat displays for each treatment. In this panel, two values above 4 are not shown. (B) Model estimates for the number of lunges.

indicator that a male has high-quality alleles that can produce a large trait (Andersson, 1994; Kirkpatrick, 1996; Zahavi, 1975). We found increased condition dependence in female traits and decreases in condition dependence in males, contrary to our expectations (Figure 6; Supplementary Figure S6). This reduced condition dependence was observed in the wing, head, and tarsus with no change in the femur or tibia, which were the traits predicted to respond to territorial selection. Although this runs counter to our prediction, even if we did see trait exaggeration it is not known at what stage we would expect condition dependence to evolve. Under a good genes model where variation for the indicator trait is oligogenic, condition dependence may evolve later in the evolution of trait exaggeration once the trait has already begun to increase in size and potentially after a depletion of segregating variation for the "trait." It may also be the case that trait exaggeration occurs most often for traits that already show heightened condition dependence because they already act as reliable indicators before exaggeration, as suggested by Johnstone et al. (2009). In this case, we would not expect to observe changes in condition dependence with increased sexual selection or SSD for the forelegs. Because the forelegs have additional locomotory functions, and their composite nature has been historically shaped by natural selection, it may be more constrained in co-opting condition dependence into trait expression, in comparison to a novel trait like a beetle horn. Despite the additional locomotory functions however, a number of insects do display exaggerated legs, suggesting

that if selection for leg exaggeration is high enough it can overcome this barrier (Singh, 1977; Zeh et al., 1992).

If Emlen's (2014) hypothesis for the three key conditions for weapon evolution, and the evolution of condition dependence to maintain sexually selected traits is true, there are experimental reasons we may have failed to detect morphological change that should be considered. Primarily among them is the density of individuals in our experiment. Although Hoffmann and Cacoyianni (1990) demonstrated resource defense polygyny in *D. melanogaster* at a similar resource patch diameter used in this experiment, they also found that at higher densities, defense was abandoned and flies reverted to scramble competition. Our findings may suggest that for trait exaggeration to initiate, populations may require generally low density to avoid interactions where resource defense is not a viable option due to pressures from multiple competitors at once. If we limited the census size, however, it would necessarily reduce genetic variation in our experiment and possibly diminish the likelihood of selecting on alleles of interest. The other possible experimental impact is also due to uncontrolled density creating possible limits on food for larvae. Although we provided both high-quality food as well as a low-quality food resource for adults, larvae in the high-quality food appeared to be growing under high-density conditions in the food. This high larval density may have created additional countervailing selection pressure on size due to larval competition. Although we did not test fecundity, the observation that density was high may suggest that fecundity selection and larval viability selection predominated, rather than sexual selection on males. With very high-quality food provided and a high adult density, there may have been a strong enough selection pressure on females to lay as many eggs as early as possible to "beat the rush." This could counter impacts of sexual selection on males in the treatments. In a previous study in our laboratory using similar spatial-environment treatments to assess selective dynamics of the purging of deleterious alleles, we observed that the impact of the spatial-environment varied among deleterious mutations influencing a variety of traits (Wilson et al., 2021). Third, we caveat that this experiment was done in a simplified lab environment, using mesh cages with prepared food in vials and bottles, and does not directly reflect the natural life history of species that have evolved exaggerated traits. Although experimental evolution is a useful tool to explore theory and test hypotheses, we acknowledge these potential limitations. However, we point out that such simplified models used in experimental evolution are an important and highly successful approach for identifying causal relationships between targets and agents of selection (Kassen, 2024; Kawecki et al., 2012; Reznick & Travis, 2019).

A large portion of the literature has been dedicated to the idea that weapons or traits used for resource (or mate) defense develop positive allometries (allometric coefficient > 1; Eberhard et al., 2018; Kodric-Brown et al., 2006). This has been shown in a number of species including beetle species (Kawano, 1995), cervids (Lemaître et al., 2014), stalk-eyed flies (Baker & Wilkinson, 2001), and many others (Voje, 2016). While weapons are often associated with positive allometries, it is by no means always the case (McCullough & O'Brien, 2022; Voje, 2016). In our study, the allometry of leg traits is generally the greatest in magnitude (although with slopes still less than 1), in comparison to other traits. Femur length, while phenotypically plastic, scales approximately isometrically with overall size in response to changes in density or food availability (Pesevski, 2021; Shingleton et al., 2009). Although we saw a general trend towards smaller size and lower allometric coefficients in the legs, confidence intervals often overlapped substantially across treatments and both sexes, with generally small magnitudes across evolutionary treatments (slope changes). This may simply reflect the constraints to sexually dimorphic evolution due to genetic correlation (r_{MF}) between the sexes (Lande, 1980). Previous studies applying strong, sexually discordant, mass artificial selection on male and female body size, found it took more than 100 generations to observe sex-specific responses, while sex-concordant responses were much faster (Audet et al., 2024; Stewart & Rice, 2018). This suggests that 75 generations of what is likely modest selective pressure (in terms of changes to sexual selection per se, as a result of the spatial environments employed), may have been insufficient number of generations for a sex-specific response (Stewart & Rice, 2018). In plants, it has been shown that the breakdown of r_{MF} can result in rapid changes in *SD* (Delph et al., 2011), but this reduced constraint was imposed by strong family-based selection. In general, homologous traits between the sexes have a very high genetic correlation, providing some constraints on the potential to evolve dimorphically (Poissant et al., 2010). The legs in *D. melanogaster*, although used in intra-sexual competition even under all the predicted circumstances proposed by Emlen, are still constrained by the shared genome, and the biomechanical function of legs in locomotion. Hence, the time required to adapt morphological changes with a shared genome with experimental evolution may be significantly longer than what is experimentally feasible. This is of course opposed by the fact that there was at least a small change in SSD in both head and wing, but not the focal trait of leg. r_{MF} is trait-specific, and response to sexually discordant selection occurs in a trait-specific manner, which has been previously suggested (Audet et al., 2024; Poissant et al., 2010).

The one trait that we observed to respond in a consistently sex-specific way was the tarsus of males (Figure 5B). It appears as though after 75 generations of evolution, male tarsi have become more hypo-allometric and less condition-dependent. This may be due to the sex-limited structure on the tarsi of males, the sex comb. The presence of sex combs *in D. melanogaster* is important for mating success (Ng & Kopp, 2008), and the length of the tarsus is directly related to sex comb number (Combs, 1937). For this reason, the presence and necessity of sex combs may be creating a physical barrier to decreased size evolution and condition-dependent response in males.

The results here, although not aligned with either Emlen's hypotheses for weapon evolution or the condition dependence hypothesis, are consistent with other experimental evolution studies. These experiments show results inconsistent with previous hypotheses for responses to experimental evolution when sexual selection dynamics are altered via manipulating sex ratio (Bath et al., 2021, 2023; Edmunds et al., 2021; Sepil et al., 2022). By manipulating census sex ratio, Bath et al. (2021, 2023) and Sepil et al. (2022) attempted to modulate the intensity of sexual selection in a long-term experimental population to explore the potential consequences of heightened sexual selection. Consistent with our observations (Figure 7), these experiments found modest increases in male aggression with increased sexual competition and a more substantial increase in female

post-mating aggression (Bath et al., 2021). They also observed no change in male condition dependence but increased female condition dependence. Despite a very different experimental approach, their results, similar to ours, show that in the context of heightened sexual selection, male condition dependence does not increase, while female condition dependence does (Bath et al., 2023). In male-biased sex ratio populations (a proxy for increased sexual competition), it appeared that males decreased investment in pre-copulatory investment, but also did not increase mating duration, which would be suggestive of post-copulatory investments (Sepil et al., 2022). These results combined with ours may suggest that the condition-dependence model of sexually dimorphic evolution may rely on assumptions of a relatively simple genetic architecture and that quantitative traits may require very strong and consistent directional selection to show a response. This has been suggested previously, and even modeled. Johnstone et al. (2009) suggest that strong sexual selection for an increase in trait size does not require the evolution of condition dependence, and if it does, this may be a brief increase in condition dependence that reduces over time. Our results, as well as Bath et al. (2023), suggest that the evolution of condition dependence is not a necessary mechanism to maintain genetic variation for traits under persistent sexual selection.

Here, we used the hypothesis established in Emlen (2008, 2014) to explore the question, why do exaggerated weapons evolve? We did not find support for this hypothesis being sufficient to explain the initiation of weapon exaggeration in *D. melanogaster*. So then, why do exaggerated weapons evolve? We do find evidence for increasing aggression when the opportunities for territoriality are present (Figure 7), but not for trait exaggeration. One potential missing factor may be intra-sexual signaling. In stalk-eyed flies, large intimidating eyestalks are used to "size-up" opponents, with the male with the shorter eyestalks often backing down, rarely escalating to contact. In instances where the males are similarly matched in size and altercations escalate, the eyestalks are not used for offensive purposes, instead they fight with their legs (Panhuis & Wilkinson, 1999). It has been hypothesized that the largest traits relative to body size are often used more as a signal of size to potential opponents rather than for a function in fighting, and "pure weapons" may tend to be smaller relative to body size (McCullough & O'Brien, 2022). This may suggest that it is not simply 'Why do exaggerated weapons evolve?', but the question of what the specific pressures on that weapon may be key. In *D. melanogaster*, there may not be a pressure to exaggerate if the size of the legs does not mediate intra-sexual signaling. A second potential amendment that may be of consideration for the hypothesis of weapon evolution is local density. In Hoffmann and Cacoyianni (1989), *D. melanogaster* reduced their propensity to defend territories when density was increased. This may suggest that territorial structure may lay the foundation for weapon evolution, but the dynamics of the competition may be an important consideration for predicting exaggerated weapon evolution. For this reason, for the evolution of specifically exaggerated weaponry, the interaction of intra-sexual signaling with other ecological factors must be included in the Emlen hypothesis.

Supplementary material

Supplementary material is available online at *Evolution*.

Data availability

Data and scripts are available at https://doi.org/10.6084/m9.figshare.28687235.v2 (static copy) and https://github.com/DworkinLab/Audet_Wilson_etal_Evolution_2025.

Author contributions

Study conceptualization and funding: I.D.; Study design: I.D., A.W.; Artificial selection and rearing: A.W., T.A.; Male mating success experiment: A.W.; Aggression experiment: R.D.; Image analysis: A.W, T.A.; Analysis: I.D., A.W., and T.A.; Manuscript drafting: T.A., A.W., and I.D.; Manuscript editing: T.A., A.W., I.D., and R.D.; Manuscript revisions: I.D., T.A., R.D., and A.W.

Funding

This work was funded by National Science and Engineering Research Council (Canada) Discovery and Discovery Accelerator grants to ID and RD and Ontario Graduate Scholarships to TA and AW.

Conflict of interest: The authors declare no conflict of interest.

Acknowledgments

Special thanks to Dr. Douglas Emlen for such stimulating discussions that led us to do this study. We thank Francesco Ruso for helping dissecting. We thank associate editor Dr. Jessica Abbott, and the reviewers for thoughtful suggestions on the study and the manuscript.

References

Andersson, M. (1994). *Sexual selection*. Princeton University Press.

Audet, T., Krol, J., Pelletier, K., Stewart, A. D., & Dworkin, I. (2024). Sexually discordant selection is associated with trait-specific morphological changes and a complex genomic response. *Evolution*, *78*(8), 1426–1440. https://doi.org/10.1093/evolut/qpae071

Baker, R. H., & Wilkinson, G. S. (2001). Phylogenetic analysis of sexual dimorphism and eye-span allometry in stalk-eyed flies (Diopsidae). *Evolution*, *55*(7), 1373–1385. https://doi.org/10.1111/j.0014-3820.2001.tb00659.x

Bateman, A. J. (1948). Intra-sexual selection in *Drosophila*. *Heredity*, *2*(Pt. 3), 349–368. https://doi.org/10.1038/hdy.1948.21

Bath, E., Edmunds, D., Norman, J., Atkins, C., Harper, L., Rostant, W. G., Chapman, T., Wigby, S., & Perry, J. C. (2021). Sex ratio and the evolution of aggression in fruit flies. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, *288*(1947), 20203053. https://doi.org/10.1098/rspb.2020.3053.

Bath, E., Rostant, W., Ostridge, H. J., Smith, S., Mason, J. S., Rafaluk-Mohr, T., Mank, J. E., Chapman, T., & Perry, J. C. (2023). Sexual selection and the evolution of condition-dependence: An experimental test at two resource levels. *Evolution*, *77*(3), 776–788. https://doi.org/10.1093/evolut/qpac066

Baxter, C., & Dukas, R. (2017). Life history of aggression: Effects of age and sexual experience on male aggression towards males and females. *Animal Behaviour*, *123*, 11–20. https://doi.org/10.1016/j.anbehav.2016.10.022

Baxter, C., Mentlik, J., Shams, I., & Dukas, R. (2018). Mating success in fruit flies: courtship interference versus female choice. *Animal Behaviour*, *138*, 101–108. https://doi.org/10.1016/j.anbehav.2018.02.010

Baxter, C. M., Shams, I., Dworkin, I., & Dukas, R. (2023). Genetic correlation between aggressive signals and fighting. *Biology Letters*, *19*, 20220616. https://doi.org/10.1098/rsbl.2022.0616

Boisseau, R. P., Ero, M. M., Makai, S., Bonneau, L. J. G., & Emlen, D. J. (2020). Sexual dimorphism divergence between sister species is associated with a switch in habitat use and mating system in thorny devil stick insects. *Behavioural Processes*, *181*, 104263. https://doi.org/10.1016/j.beproc.2020.104263

Bonduriansky, R. (2007). The evolution of condition-dependent sexual dimorphism. *The American Naturalist*, *169*(1), 9–19. https://doi.org/10.1086/510214

Bonduriansky, R., & Day, T. (2003). The evolution of static allometry in sexually selected traits. *Evolution*, *57*(11), 2450–2458. https://doi.org/10.1111/j.0014-3820.2003.tb01490.x

Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Nielsen, C. B. W. A., Skaug, J. H., Mächler, M., & Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for sero-inflated generalized linear mixed modeling. *The R Journal*, *9*, 378.

Ceballos, C. P., & Valenzuela, N. (2011). The role of sex-specific plasticity in shaping sexual dimorphism in a long-lived vertebrate, the snapping turtle *Chelydra serpentina*. *Evolutionary Biology*, *38*(2), 163–181. https://doi.org/10.1007/s11692-011-9117-8

Chen, S., Lee, A. Y., Bowens, N. M., Huber, R., & Kravitz, E. A. (2002). Fighting fruit flies: A model system for the study of aggression. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(8), 5664–5668. https://doi.org/10.1073/pnas.082102599

Combs, J. D. (1937). Genetic and environmental factors affecting the development of the sex-combs of *Drosophila melanogaster*. *Genetics*, *22*(4), 427–433. https://doi.org/10.1093/genetics/22.4.427

Darwin, C. (1871). *The descent of man, and selection in relation to sex*. John Murray. https://doi.org/10.1037/12293-000

David, P., Bjorksten, T., Fowler, K., & Pomiankowski, A. (2000). Condition-dependent signalling of genetic variation in stalk-eyed flies. *Nature*, *406*(6792), 186–188. https://doi.org/10.1038/35018079

Delph, L. F., Steven, J. C., Anderson, I. A., Herlihy, C. R., & Brodie III, E. D. (2011). Elimination of a genetic correlation between the sexes via artificial correlational selection. *Evolution*, *65*(10), 2872–2880. https://doi.org/10.1111/j.1558-5646.2011.01350.x

Dodson, G. N. (1997). Resource defense mating system in antlered flies, *Phytalmia spp*. (Diptera: Tephritidae). *Annals of the Entomological Society of America*, *90*(4), 496–504. https://doi.org/10.1093/aesa/90.4.496

Dow, M. A., & Schilcher, F. V. (1975). Aggression and mating success in *Drosophila melanogaster*. *Nature*, *254*(5500), 511–512. https://doi.org/10.1038/254511a0

Dukas, R. (2020). Natural history of social and sexual behavior in fruit flies. *Scientific Reports*, *10*, 21932. https://doi.org/10.1038/s41598-020-79075-7

Dweck, H. K. M., Ebrahim, S. A. M., Kromann, S., Bown, D., Hillbur, Y., Sachse, S., Hansson, B. S., & Stensmyr, M. C. (2013). Olfactory preference for egg laying on citrus substrates in Drosophila. *Current Biology*, *23*, 2472–2480. https://doi.org/10.1016/j.cub.2013.10.047

Eberhard, W. G., Rodríguez, R. L., Huber, B. A., Speck, B., Miller, H., Buzatto, B. A., & Machado, G. (2018). Sexual selection and static allometry: The importance of function. *The Quarterly Review of Biology*, *93*(3), 207–250. https://doi.org/10.1086/699410

Edmunds, D., Wigby, S., & Perry, J. C. (2021). A resource-poor developmental diet reduces adult aggression in male *Drosophila melanogaster*. *Behavior, Ecology and Sociobiology*, *75*(7), 110. https://doi.org/10.1007/s00265-021-03050-z

Emlen, D. J. (1997). Alternative reproductive tactics and male-dimorphism in the horned beetle *Onthophagus acuminatus* (Coleoptera: Scarabaeidae). *Behavioral Ecology and Sociobiology*, *41*(5), 335–341. https://doi.org/10.1007/s002650050393

Emlen, D. J. (2008). The evolution of animal weapons. *Annual Review of Ecology, Evolution, and Systematics*, *39*(1), 387–413. https://doi.org/10.1146/annurev.ecolsys.39.110707.173502

Emlen, D. J. (2014). Reproductive contests and the evolution of extreme weaponry. *The evolution of insect mating systems*, Vol. *27* Pages 92–105. Oxford Academic.

Emlen, D. J., & Philips, T. K. (2006). Phylogenetic evidence for an association between tunneling behavior and the evolution of horns in dung beetles (Coleoptera: Scarabaeidae: Scarabaeinae). *The Coleopterists Bulletin*, *60*, 47–56. https://doi.org/10.1649/0010-065X(2006)60[47:PEFAAB]2.0.CO;2

Emlen, S. T., & Oring, L. W. (1977). Ecology, sexual selection, and the evolution of mating systems. *Science*, *197*(4300), 215–223. https://doi.org/10.1126/science.327542

Fairbairn, D. J. (2005). Allometry for sexual size dimorphism: Testing two hypotheses for Rensch's rule in the water strider Aquarius remiges. *The American Naturalist*, *166*(S4), S69–S84. https://doi.org/10.1086/444600

Friard, O., Gamba, M., & Fitzjohn, R. (2016). BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution*, *7*(11), 1325–1330. doi:10.1111/2041-210X.12584

Guo, X., & Dukas, R. (2020). The cost of aggression in an animal without weapons. *Ethology*, *126*(1), 24–31. https://doi.org/10.1111/eth.12956

Harshman, L. G., & Hoffmann, A. A. (2000). Laboratory selection experiments using *Drosophila*: What do they really tell us? *Trends in Ecology & Evolution*, *15*(1), 32–36. https://doi.org/10.1016/s0169-5347(99)01756-5

Hoffmann, A. A., & Cacoyianni, Z. (1989). Selection for territoriality in *Drosophila melanogaster*: Correlated responses in mating success and other fitness components. *Animal Behaviour*, *38*(1), 23–34. https://doi.org/10.1016/s0003-3472(89)80062-4

Hoffmann, A. A., & Cacoyianni, Z. (1990). Territoriality in Drosophila melanogaster as a conditional strategy. *Animal Behaviour*, *40*(3), 526–537. https://doi.org/10.1016/s0003-3472(05)80533-0

Honěk, A., & Honek, A. (1993). Intraspecific variation in body size and fecundity in insects: A general relationship. *Oikos*, *66*(3), 483–492. https://doi.org/10.2307/3544943

Hurd, P. L. (2006). Resource holding potential, subjective resource value, and game theoretical models of aggressiveness signaling. *Journal of Theoretical Biology*, *241*(3), 639–648. https://doi.org/10.1016/j.jtbi.2006.01.001

Jacobs, M. E. (1960). Influence of light on mating of *Drosophila melanogaster*. *Ecology*, *41*(1), 182–188. https://doi.org/10.2307/1931952

Jensen, K., McClure, C., Priest, N. K., & Hunt, J. (2015). Sex-specific effects of protein and carbohydrate intake on reproduction but not lifespan in *Drosophila melanogaster*. *Aging Cell*, *14*(4), 605–615. https://doi.org/10.1111/acel.12333

Johns, A., Gotoh, H., McCullough, E. L., Emlen, D. J., & Lavine, L. C. (2014). Heightened condition-dependent growth of sexually selected weapons in the rhinoceros beetle, *Trypoxylus dichotomus* (Coleoptera: Scarabaeidae). *Integrative and Comparative Biology*, *54*(4), 614–621. https://doi.org/10.1093/icb/icu041

Johnstone, R. A., Rands, S. A., & Evans, M. R. (2009). Sexual selection and condition-dependence. *Journal of Evolutionary Biology*, *22*(12), 2387–2394. https://doi.org/10.1111/j.1420-9101.2009.01822.x

Kassen, R. (2024). *Experimental evolution and the nature of biodiversity*. Oxford University Press.

Kawano, K. (1995). Horn and Wing allometry and male dimorphism in giant rhinoceros beetles (Coleoptera: Scarabaeidae) of tropical Asia and America. *Annals of the Entomological Society of America*, *88*(1), 92–99. https://doi.org/10.1093/aesa/88.1.92

Kawecki, T. J., Lenski, R. E., Ebert, D., Hollis, B., Olivieri, I., & Whitlock, M. C. (2012). Experimental evolution. *Trends in Ecology & Evolution*, *27*(10), 547–560. https://doi.org/10.1016/j.tree.2012.06.001

Kirkpatrick, M. (1996). Good genes and direct selection in the evolution of mating preferences. *Evolution*, *50*(6), 2125–2140. https://doi.org/10.1111/j.1558-5646.1996.tb03603.x

Kodric-Brown, A., Sibly, R. M., & Brown, J. H. (2006). The allometry of ornaments and weapons. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(23), 8733–8738. https://doi.org/10.1073/pnas.0602994103

Kudo, A., Shigenobu, S., Kadota, K., Nozawa, M., Shibata, T. F., Ishikawa, Y., & Matsuo, T. (2017). Comparative analysis of the brain

transcriptome in a hyper-aggressive fruit fly, *Drosophila prolongata*. *Insect Biochemistry and Molecular Biology*, *82*, 11–20. https://doi.org/10.1016/j.ibmb.2017.01.006

Lande, R. (1980). Sexual dimorphism, sexual selection, and adaptation in polygenic characters. *Evolution*, *34*(2), 292–305. https://doi.org/10.1111/j.1558-5646.1980.tb04817.x

Lee, K. P., Simpson, S. J., Clissold, F. J., Brooks, R., Ballard, J. W. O., Taylor, P. W., Soran, N., & Raubenheimer, D. (2008). Lifespan and reproduction in *Drosophila*: New insights from nutritional geometry. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(7), 2498–2503. https://doi.org/10.1073/pnas.0710787105

Lemaître, J. F., Vanpé, C., Plard, F., & Gaillard, J. M. (2014). The allometry between secondary sexual traits and body size is non-linear among cervids. *Biology Letters*, *10*, 20130869. https://doi.org/10.1098/rsbl.2013.0869

Lenth, R., Singmann, H., Love, J., & Buerkner, P. (2018). emmeans: Estimated marginal means, aka Least-Squares Means.https://rvlenth.github.io/emmeans/ Accessed August 4, 2024.

Maklakov, A. A., Simpson, S. J., Zajitschek, F., Hall, M. D., Dessmann, J., Clissold, F., Raubenheimer, D., Bonduriansky, R., & Brooks, R. C. (2008). Sex-specific fitness effects of nutrient intake on reproduction and lifespan. *Current Biology*, *18*(14), 1062–1066. https://doi.org/10.1016/j.cub.2008.06.059

McCullough, E. L., & O'Brien, D. M. (2022). Variation in allometry along the weapon-signal continuum. *Evolutionary Ecology*, *36*(4), 591–604. https://doi.org/10.1007/s10682-022-10158-9

Moczek, A. P., & Emlen, D. J. (2000). Male horn dimorphism in the scarab beetle, *Onthophagus taurus*: Do alternative reproductive tactics favour alternative phenotypes? *Animal Behaviour*, *59*(2), 459–466. https://doi.org/10.1006/anbe.1999.1342

Nandy, B., Gupta, V., Sen, S., Udaykumar, N., Samant, M. A., Ali, S. Z., & Prasad, N. G. (2013). Evolution of mate-harm, longevity and behaviour in male fruit flies subjected to different levels of interlocus conflict. *BMC Evolutionary Biology*, *13*, 212. https://doi.org/10.1186/1471-2148-13-212

Ng, C. S., & Kopp, A. (2008). Sex combs are important for male mating success in *Drosophila melanogaster*. *Behavior Genetics*, *38*(2), 195–201. https://doi.org/10.1007/s10519-008-9190-7

Nilsen, S. P., Chan, Y. -B., Huber, R., & Kravitz, E. A. (2004). Gender-selective patterns of aggressive behavior in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 12342–12347. https://doi.org/10.1073/pnas.0404693101

Palaoro, A. V., & Peixoto, P. E. C. (2022). The hidden links between animal weapons, fighting style, and their effect on contest success: A meta-analysis. *Biological Reviews of the Cambridge Philosophical Society*, *97*(5), 1948–1966. https://doi.org/10.1111/brv.12877

Panhuis, T. M., & Wilkinson, G. S. (1999). Exaggerated male eye span influences contest outcome in stalk-eyed flies (Diopsidae). *Behavioral Ecology and Sociobiology*, *46*(4), 221–227. https://doi.org/10.1007/s002650050613

Partridge, L., Ewing, A., & Chandler, A. (1987). Male size and mating success in *Drosophila melanogaster*: The roles of male and female behaviour. *Animal Behaviour*, *35*(2), 555–562. https://doi.org/10.1016/s0003-3472(87)80281-6

Perdigón Ferreira, J., Rohner, P. T., & Lüpold, S. (2023). Strongly sexually dimorphic forelegs are not more condition-dependent than less dimorphic traits in *Drosophila prolongata*. *Evolutionary Ecology*, *37*(3), 493–508. https://doi.org/10.1007/s10682-022-10226-0

Pesevski, M. (2021). *Influence of environmental variation on sexual dimorphism in Drosophila morphology among adaptively diverged populations and in an inter-specific comparative context* [*Doctoral dissertation McMaster University*]. http://hdl.handle.net/11375/26936

Poissant, J., Wilson, A. J., & Coltman, D. W. (2010). Sex-specific genetic variance and the evolution of sexual dimorphism: A systematic review of cross-sex genetic correlations. *Evolution*, *64*(1), 97–107. https://doi.org/10.1111/j.1558-5646.2009.00793.x

Reddiex, A. J., Gosden, T. P., Bonduriansky, R., & Chenoweth, S. F. (2013). Sex-specific fitness consequences of nutrient intake and the evolvability of diet preferences. *The American Naturalist*, *182*(1), 91–102. https://doi.org/10.1086/670649

Reeve, J. P., & Fairbairn, D. J. (1999). Change in sexual size dimorphism as a correlated response to selection on fecundity. *Heredity*, *83* (6), 697–706. https://doi.org/10.1046/j.1365-2540.1999.00616.x

Reznick, D. N., & Travis, J. (2019). Experimental studies of evolution and eco-evo dynamics in guppies (*Poecilia reticulata*). *Annual Review of Ecology, Evolution, and Systematics*, *50*(1), 335–354. https://doi.org/10.1146/annurev-ecolsys-110218-024926

Rohde, P. D., Gaertner, B., Ward, K., Sørensen, P., & Mackay, T. F. C. (2017). Genomic analysis of genotype-by-social environment interaction for *Drosophila melanogaster* aggressive behavior. *Genetics*, *206*(4), 1969–1984. https://doi.org/10.1534/genetics.117.200642

Rowe, L., & Houle, D. (1996). The lek paradox and the capture of genetic variance by condition dependent traits. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, *263*, 1415–1421. https://doi.org/10.1098/rspb.1996.0207

Rueden, C. T., Schindelin, J., Hiner, M. C., DeZonia, B. E., Walter, A. E., Arena, E. T., & Eliceiri, K. W. (2017). ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics*, *18*(1), 529. https://doi.org/10.1186/s12859-017-1934-z

Scott, A. M., Dworkin, I., & Dukas, R. (2022). Evolution of sociability by artificial selection. *Evolution*, *76*(3), 541–553. https://doi.org/10.1111/evo.14370

Sepil, I., Perry, J. C., Dore, A., Chapman, T., & Wigby, S. (2022). Experimental evolution under varying sex ratio and nutrient availability modulates male mating success in *Drosophila melanogaster*. *Biology Letters*, *18*, 20210652. https://doi.org/10.1098/rsbl.2021.0652

Shingleton, A. W., Estep, C. M., Driscoll, M. V., & Dworkin, I. (2009). Many ways to be small: Different environmental regulators of size generate distinct scaling relationships in *Drosophila melanogaster*. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, *276*(1667), 2625–2633. https://doi.org/10.1098/rspb.2008.1796

Singh, B. (1977). Two new and two unrecorded species of the genus Drosophila Fallen (Diptera: Drosophilidae) from Shillong, Meghalaya, India. *Proceedings of the Zoological Society*, *30*, 31.

Spieth, H. T. (1974). Courtship behavior in *Drosophila*. *Annual Review of Entomology*, *19*, 385–405. https://doi.org/10.1146/annurev.en.19.010174.002125

Spieth, H. T. (1981). *Drosophila heteroneura* and *Drosophila silvestris*: Head shapes, behavior and evolution. *Evolution*, *35*(5), 921–930. https://doi.org/10.1111/j.1558-5646.1981.tb04958.x

Stewart, A. D., & Rice, W. R. (2018). Arrest of sex-specific adaptation during the evolution of sexual dimorphism in *Drosophila*. *Nature Ecology and Evolution*, *2*(9), 1507–1513. https://doi.org/10.1038/s41559-018-0613-4

Stillwell, R. C., Dworkin, I., Shingleton, A. W., & Frankino, W. A. (2011). Experimental manipulation of body size to estimate morphological scaling relationships in *Drosophila*. *Journal of Visualized Experiments*, *56*, 56-3162. https://doi.org/10.3791/3162

Tanaka, T., & Yamazaki, T. (1990). Fitness and its components in *Drosophila melanogaster*. *Idengaku Zasshi*, *65*(6), 417–426. https://doi.org/10.1266/jjg.65.417

Tatar, M. (2011). The plate half-full: Status of research on the mechanisms of dietary restriction in *Drosophila melanogaster*. *Experimental Gerontology*, *46*(5), 363–368. https://doi.org/10.1016/j.exger.2010.12.002

Toyoshima, N., & Matsuo, T. (2023). Fight outcome influences male mating success in *Drosophila prolongata*. *Journal of Ethology*, *41*, 119. https://doi.org/10.1007/s10164-023-00778-1

Ueda, A., & Kidokoro, Y. (2002). Aggressive behaviours of female *Drosophila melanogaster* are influenced by their social experience and food resources. *Physiological Entomology*, *27*(1), 21–28. https://doi.org/10.1046/j.1365-3032.2002.00262.x

Voje, K. L. (2016). Scaling of morphological characters across trait type, sex, and environment: A meta-analysis of static allometries. *The American Naturalist*, *187*(1), 89–98. https://doi.org/10.1086/684159

Wickham, H. (2018). *Ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York.

Wilkinson, G. S., & Dodson, G. N. (1997). Function and evolution of antlers and eye stalks in flies. In J. C. Choe, & B. J. Crespi (Eds.), *The evolution of mating systems in insects and arachnids* (pp. 310–328). Cambridge University Press.

Wilson, A. E., Siddiqui, A., & Dworkin, I. (2021). Spatial heterogeneity in resources alters selective dynamics in *Drosophila melanogaster*. *Evolution*, *75*(7), 1792–1804. https://doi.org/10.1111/evo.14262

Zahavi, A. (1975). Mate selection—A selection for a handicap. *Journal of Theoretical Biology*, *53*(1), 205–214. https://doi.org/10.1016/0022-5193(75)90111-3

Zeh, D. W., Zeh, J. A., & Tavakilian, G. (1992). Sexual selection and sexual dimorphism in the harlequin beetle *Acrocinus longimanus*. *Biotropica*, *24*(1), 86–96. https://doi.org/10.2307/2388476

Zwarts, L., Versteven, M., & Callaerts, P. (2012). Genetics and neurobiology of aggression in *Drosophila*. *Fly*, *6*(1), 35–48. https://doi.org/10.4161/fly.19249

5. Conclusions

The evolution of exaggerated sexually dimorphic traits has been a long-standing question in evolutionary biology. The field, while still elaborating and refining models, has moved more to functional genomic approaches recently. Exploration of sex-biased gene expression specifically has been a key tool in understanding the expression of these traits, as well as in some cases the evolution. This work has unfortunately been somewhat held back by inconsistent methodology, and at times problematic sampling practices. Here we aimed to first set out best practices for analysis and sampling in sex-biased gene expression (introduction).we then utilize these set-out practices to explore the evolution of the sexually exaggerated forelegs in the up-and-coming model *D. prolongata*, and closely related species (chapter 2).we then utilize artificial selection as well as evolve and sequence methods to look for alleles under sexual conflict during artificial selection on sexual dimorphism (chapter 3). Finally, we use experimental evolution to test the prevailing hypothesis on the environmental precursors to weapon evolution (chapter 4).

Using the principles outlines in chapter 1, we explore the evolution of an exaggerated sexually dimorphic trait in the *Drosophila prolongata* as well as the non-exaggerated forelegs of the closely related species *D. carrolli* and *D. melanogaster* (chapter 2). We found the potential large effect gene *grain*, when knocked-down using RNAi, creates phenotypically *D. prolongata*-like leg phenotypes when in *D. melanogaster*. We also identified a positive relationship with number, but not magnitude, of sex-biased genes and morphological sexual size dimorphism. There was also no substantial change in how growth signalling pathways seem to be used, based on changes magnitude of expression or direction of expression between legs or between *D. prolongata* and *D. carrolli*. Previous work looking at sex-biased gene expression has implicated growth signalling (Emlen et al. 2012; Warren et al. 2013; Zinna et al. 2018), limb patterning (Hust et al. 2018), or sex-determination (Rohner et al. 2021) pathways as likely

candidates for the evolution of sexual dimorphism. These pathways are doubtlessly important for the expression exaggerated traits, but unlikely candidates for the initiation of sexually dimorphic evolution. These pathways would suggest a simplistic route to sex-specific morphological evolution, where selection on major developmental pathways escapes the constraint of the shared genome and results in spectacular morphological complexity. Theory however has predicted that these traits are likely linked to many loci, tightly linked to condition (Rowe and Houle 1996). A polygenic response appears more likely, as strong directional selection on sexually selected traits would require many alleles to select on, or alleles would quickly become fixed and there would be no sustained target for females to select on (Rowe and Houle 1996). We find some support for this model, with a positive relationship between number of sex-biased genes and morphological sexual dimorphism in leg traits. We also identified large effect genes, and many transcription factors with sex-biased expression in exaggerated forelegs. This result may suggest that these traits are highly polygenic, with a few 'key players' that demonstrate disproportionately large effects on the phenotype.

In chapter 3, We utilized an ongoing long-term artificial selection experiment in *D. melanogaster* to explore how the sexes overcome the constraint imposed by the shared genome to change sexual size dimorphism. Because the sexes largely share their genome, evolution towards differing phenotypic optima requires a tug-of-war. This is, at least partly, constrained by the genetic correlation between the sexes, r_{MF} (Lande 1980b). Sex-biased gene expression is a potential way to ease this conflict; however, regulation of gene expression may occur with allele frequency changes during discordant selection. We sequenced pools of individuals who have been selected discordantly for body size between the sexes, to reverse sexual size dimorphism (Stewart and Rice 2018). We found a large proportion of the genome had allele frequency changes when compared to control lines, suggesting a highly polygenic response to discordant selection. We also identified a region on chromosome 3L that appears to be segregating differentially between the sexes. This differential segregation has previously been modelled and

suggested to be unlikely to persist in nature (Kasimatis et al. 2019). The specific regimen of selection allows this to occur in the discordant selection lineage with >80% mortality each generation and strong persistent selection on the founders of the next generation (Kasimatis et al. 2019). The presence of these differentially segregating alleles in just one replicate may be suggestive of a novel inversion allowing maintenance of a conflict region under differential segregation (McAllester and Pool 2025). And recent work has also found potential conflict alleles segregating at differential frequencies between the sexes in sticklebacks (Sylvestre et al. 2023). More work is required however to confirm that this region is still under differential segregation, and to isolate the cause.

In chapter 4 we utilize experimental evolution to test the prevailing hypothesis for the evolution of sexually selected weapons. The potential for polygamy due to environmental layout may increase competition among males (Emlen and Oring 1977). If competition is structured in a way where females or resources required by females may be monopolized it may result in intense competition to increase resource holding potential in males. Emlen (2008) hypothesized that if fighting between males favours larger size, occurs in one-on-one duals, and this pressure to access mates results in high variance in male reproductive success, the evolution of weapons may occur. To test this hypothesis myself and Audrey Wilson conducted 75 generations of experimental evolution in *D. melanogaster* replicating these environmental precursors. We found that despite the experimental evolution set-up creating variance in reproductive success favouring larger males, and our experimental treatments resulting in increased aggression in males, only very modest changes in size or morphology occurred, and often in females rather than males. Previous work increasing sexual selection pressure via census sex ratio has also found unexpected responses in females and a lack of change in male body size despite these hypotheses (Sepil et al. 2022; Bath et al. 2023). This suggests that amendments to the hypothesis may be necessary. In our experimental evolution treatments, density was uncontrolled. Previously, willingness to defend to territories is both resource size and density dependent (Hoffmann and Cacoyianni 1990). It may be the

case that resource layout must be patchy enough that local density is particularly low before males are willing to actively engage in defense behaviour.

Altogether this work adds to our growing understanding of the genomic architecture of sexual dimorphism and the constraint that the shared genome imposes. Further work using the outlined best practices for RNAseq experiments should hopefully both increase our understanding of sex-biased gene expression and its relationship to exaggerated trait evolution, as well as provide standardized data that allows meta-analyses to draw stronger more generalizable conclusions. Further, as new genomic techniques become more accessible and affordable such as single cell RNA sequencing, CHiPseq, and ATACseq, these will better be able to give insights into specific regulatory and tissue specific changes that occur during the evolution of exaggerated traits.

References

Andersson, M. 1986. Evolution of condition-dependent sex ornaments and mating preferences: sexual selection based on viability differences. Evolution 40:804–816. https://doi.org/10.1111/j.1558-5646.1986.tb00540.x

Andersson, M. 1994. Sexual Selection. Princeton University Press.

Assis, R., Q. Zhou, and D. Bachtrog. 2012. Sex-Biased Transcriptome Evolution in Drosophila. Genome Biol. Evol. 4:1189–1200. https://doi.org/10.1093/gbe/evs093

Basolo, A. L. 1990. Female Preference Predates the Evolution of the Sword in Swordtail Fish. Science 250:808–810. https://doi.org/10.1126/science.250.4982.808

Bateman, A. J. 1948. Intra-sexual selection in Drosophila. Heredity 2:349–368. https://doi.org/10.1038/hdy.1948.21

Bath, E., W. Rostant, H. J. Ostridge, S. Smith, J. S. Mason, T. Rafaluk-Mohr, J. E. Mank, T. Chapman, and J. C. Perry. 2023. Sexual selection and the evolution of condition-dependence: an experimental test at two resource levels. Evolution 77:776–788. https://doi.org/10.1093/evolut/qpac066

Boyle, E. A., Y. I. Li, and J. K. Pritchard. 2017. An Expanded View of Complex Traits: From Polygenic to Omnigenic. Cell 169:1177–1186. https://doi.org/10.1016/j.cell.2017.05.038

Chandler, C. H., S. Chari, D. Tack, and I. Dworkin. 2014. Causes and Consequences of Genetic Background Effects Illuminated by Integrative Genomic Analysis. Genetics 196:1321–1336. https://doi.org/10.1534/genetics.113.159426

Chandler, C. H., C. Ofria, and I. Dworkin. 2013. Runaway Sexual Selection Leads to Good Genes. Evolution 67:110–119. https://doi.org/10.1111/j.1558-5646.2012.01750.x

Clutton-Brock, T. 2017. Reproductive competition and sexual selection. Philos. Trans. R. Soc. B Biol. Sci. 372:20160310. https://doi.org/10.1098/rstb.2016.0310

Clutton-brock, T. H., and G. A. Parker. 1995. Sexual coercion in animal societies. Anim. Behav. 49:1345–1365. https://doi.org/10.1006/anbe.1995.0166

Connallon, T., and L. L. Knowles. 2005. Intergenomic conflict revealed by patterns of sex-biased gene expression. Trends Genet. 21:495–499. https://doi.org/10.1016/j.tig.2005.07.006

Darolti, I., and J. E. Mank. 2023. Sex-biased gene expression at single-cell resolution: cause and consequence of sexual dimorphism. Evol. Lett. 7:148–156. https://doi.org/10.1093/evlett/qrad013

Darwin, C. 1871. The Descent of Man, and Selection in Relation to Sex.

Deegan, D. F., and N. Engel. 2019. Sexual Dimorphism in the Age of Genomics: How, When, Where. Front. Cell Dev. Biol. 7. https://doi.org/10.3389/fcell.2019.00186

Djordjevic, J., Z. Dumas, M. Robinson-Rechavi, T. Schwander, and D. J. Parker. 2022. Dynamics of sex-biased gene expression over development in the stick insect Timema californicum. Heredity. 129:113–122. https://doi.org/10.1038/s41437-022-00536-y

Dutoit, L., C. F. Mugal, P. Bolívar, M. Wang, K. Nadachowska-Brzyska, L. Smeds, H. P. Yazdi, L. Gustafsson, and H. Ellegren. 2018. Sex-biased gene expression, sexual antagonism and levels of genetic diversity in the collared flycatcher (*Ficedula albicollis*) genome. Mol. Ecol. 27:3572–3581. https://doi.org/10.1111/mec.14789

Dworkin, I., J. A. Anderson, Y. Idaghdour, E. K. Parker, E. A. Stone, and G. Gibson. 2011. The Effects of Weak Genetic Perturbations on the Transcriptome of the Wing Imaginal Disc and Its Association With Wing Shape in Drosophila melanogaster. Genetics 187:1171–1184. https://doi.org/10.1534/genetics.110.125922

Dworkin, I., E. Kennerly, D. Tack, J. Hutchinson, J. Brown, J. Mahaffey, and G. Gibson. 2009. Genomic Consequences of Background Effects on scalloped Mutant Expressivity in the Wing of *Drosophila melanogaster*. Genetics 181:1065–1076. https://doi.org/10.1534/genetics.108.096453

Ellegren, H., and J. Parsch. 2007. The evolution of sex-biased genes and sex-biased gene expression. Nat. Rev. Genet. 8:689–698. https://doi.org/10.1038/nrg2167

Emberts, Z., W. S. Hwang, and J. J. Wiens. 2021. Weapon performance drives weapon evolution. Proc. R. Soc. B Biol. Sci. 288:20202898. https://doi.org/10.1098/rspb.2020.2898

Emlen, D. J. 2014. Reproductive contests and the evolution of extreme weaponry. *The evolution of insect mating systems*, *27*, 92-105.

Emlen, D. J. 2008. The Evolution of Animal Weapons. Annu. Rev. Ecol. Evol. Syst. 39:387–413. https://doi.org/10.1146/annurev.ecolsys.39.110707.173502

Emlen, D. J., and T. K. Philips. 2006. Phylogenetic Evidence for an Association Between Tunneling Behavior and the Evolution of Horns in Dung Beetles (Coleoptera: Scarabaeidae: Scarabaeinae). Coleopt. Bull. 60:47–56. https://doi.org/10.1649/0010-065X(2006)60[47:PEFAAB]2.0.CO;2

Emlen, D. J., I. A. Warren, A. Johns, I. Dworkin, and L. C. Lavine. 2012. A Mechanism of Extreme Growth and Reliable Signaling in Sexually Selected Ornaments and Weapons. Science 337:860–864. https://doi.org/10.1126/science.1224286

Emlen, S. T., and L. W. Oring. 1977. Ecology, Sexual Selection, and the Evolution of Mating Systems. Science 197:215–223. https://doi.org/10.1126/science.327542

Fass, S. B., B. Mulvey, R. Chase, W. Yang, D. Selmanovic, S. M. Chaturvedi, E. Tycksen, L. A. Weiss, and J. D. Dougherty. 2024. Relationship between sex biases in gene expression and sex biases in autism and Alzheimer's disease. Biol. Sex Differ. 15:47. https://doi.org/10.1186/s13293-024-00622-2

Fisher, R. A. 1930. The genetical theory of natural selection.

Gotoh, H., M. Ishiguro, H. Nishikawa, S. Morita, K. Okada, T. Miyatake, T. Yaginuma, and T. Niimi. 2016. Molecular cloning and functional characterization of the sex-determination gene doublesex

in the sexually dimorphic broad-horned beetle Gnatocerus cornutus (Coleoptera, Tenebrionidae). Sci. Rep. 6:29337. https://doi.org/10.1038/srep29337

Hoffmann, A. A., and Z. Cacoyianni. 1990. Territoriality in Drosophila melanogaster as a conditional strategy. Anim. Behav. 40:526–537. https://doi.org/10.1016/S0003-3472(05)80533-0

Hust, J., M. D. Lavine, A. M. Worthington, R. Zinna, H. Gotoh, T. Niimi, and L. Lavine. 2018. The Fat-Dachsous signaling pathway regulates growth of horns in *Trypoxylus dichotomus*, but does not affect horn allometry. J. Insect Physiol. 105:85–94. https://doi.org/10.1016/j.jinsphys.2018.01.006

Innocenti, P., and S. F. Chenoweth. 2013. Interspecific Divergence of Transcription Networks along Lines of Genetic Variance in Drosophila: Dimensionality, Evolvability, and Constraint. Mol. Biol. Evol. 30:1358–1367. https://doi.org/10.1093/molbev/mst047

Jelisaveta, D., D. Zoé, R.-R. Marc, S. Tanja, and P. D. James. 2021. Dynamics of sex-biased gene expression during development in the stick insect *Timema californicum*. *Heredity*, *129*(2), 113-122. https://doi.org/10.1038/s41437-022-00536-y

Jones, A. G., and N. L. Ratterman. 2009. Mate choice and sexual selection: What have we learned since Darwin? Proc. Natl. Acad. Sci. 106:10001–10008. https://doi.org/10.1073/pnas.0901129106

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. https://doi.org/10.1534/g3.119.400711

Khodursky, S., N. Svetec, S. M. Durkin, and L. Zhao. 2020. The evolution of sex-biased gene expression in the Drosophila brain. Genome Res. 30:874–884. hhttps://doi.org/10.1101/gr.259069.119

Kirkpatrick, M., and M. J. Ryan. 1991. The evolution of mating preferences and the paradox of the lek. Nature 350:33–38. https://doi.org/10.1038/350033a0

Kodric-Brown, A., R. M. Sibly, and J. H. Brown. 2006. The allometry of ornaments and weapons. Proc. Natl. Acad. Sci. 103:8733–8738. https://doi.org/10.1073/pnas.0602994103

Kotiaho, J. S., L. W. Simmons, and J. L. Tomkins. 2001. Towards a resolution of the lek paradox. Nature 410:684–686. https://doi.org/10.1038/35070557

Kuruvilla, F. G., P. J. Park, and S. L. Schreiber. 2002. Vector algebra in the analysis of genome-wide expression data. Genome Biol. **3**, research0011.1. https://doi.org/10.1186/gb-2002-3-3-research0011

Lande, R. 1981. Models of speciation by sexual selection on polygenic traits. Proc. Natl. Acad. Sci. 78:3721–3725. https://doi.org/10.1073/pnas.78.6.3721

Lande, R. 1980a. Sexual Dimorphism, Sexual Selection, and Adaptation in Polygenic Characters. Evolution 34:292–305. https://doi.org/10.2307/2407393

Lande, R. 1980b. The genetic covariance between characters maintained by pleiotropic mutations. Genetics 94:203–215. https://doi.org/10.1093/genetics/94.1.203

Lemos, B., L. O. Araripe, and D. L. Hartl. 2008. Polymorphic Y Chromosomes Harbor Cryptic Variation with Manifold Functional Consequences. Science 319:91–93. https://doi.org/10.1126/science.1148861

Lopes-Ramos, C. M., C.-Y. Chen, M. L. Kuijjer, J. N. Paulson, A. R. Sonawane, M. Fagny, J. Platig, K. Glass, J. Quackenbush, and D. L. DeMeo. 2020. Sex Differences in Gene Expression and Regulatory Networks across 29 Human Tissues. Cell Rep. 31:107795. https://doi.org/10.1016/j.celrep.2020.107795

Love, M. I., W. Huber, and S. Anders. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15:550. https://doi.org/10.1186/s13059-014-0550-8

Lowe, R., C. Gemma, V. K. Rakyan, and M. L. Holland. 2015. Sexually dimorphic gene expression emerges with embryonic genome activation and is dynamic throughout development. BMC Genomics 16:295. https://doi.org/10.1186/s12864-015-1506-4

Lu, T., and J. C. Mar. 2020. Investigating transcriptome-wide sex dimorphism by multi-level analysis of single-cell RNA sequencing data in ten mouse cell types. Biol. Sex Differ. 11:61. https://doi.org/10.1186/s13293-020-00335-2

Macartney, E. L., V. Zeender, A. Meena, A. N. De Nardo, R. Bonduriansky, and S. Lüpold. 2021. Sperm depletion in relation to developmental nutrition and genotype in Drosophila melanogaster. Evolution 75:2830–2841. https://doi.org/10.1111/evo.14373

Mank, J. E. 2017. The transcriptional architecture of phenotypic dimorphism. Nat. Ecol. Evol. 1:0006. https://doi.org/10.1038/s41559-016-0006

McAllester, C. S., and J. E. Pool. 2025. The potential of inversions to accumulate balanced sexual antagonism is supported by simulations and Drosophila experiments. eLife 12:RP93338. https://doi.org/10.7554/eLife.93338.4

McCullough, E. L., C. W. Miller, and D. J. Emlen. 2016. Why Sexually Selected Weapons Are Not Ornaments. Trends Ecol. Evol. 31:742–751. https://doi.org/10.1016/j.tree.2016.07.004

McCullough, E. L., and D. M. O'Brien. 2022. Variation in allometry along the weapon-signal continuum. Evol. Ecol. 36:591–604. https://doi.org/10.1007/s10682-022-10158-9

McGuigan, K., J. M. Collet, E. A. McGraw, Y. H. Ye, S. L. Allen, S. F. Chenoweth, and M. W. Blows. 2014. The Nature and Extent of Mutational Pleiotropy in Gene Expression of Male Drosophila serrata. Genetics 196:911–921. https://doi.org/10.1534/genetics.114.161232

Mead, L. S., and S. J. Arnold. 2004. Quantitative genetic models of sexual selection. Trends Ecol. Evol. 19:264–271. https://doi.org/10.1016/j.tree.2004.03.003

Mishra, P., H. D. Rundle, and A. F. Agrawal. 2024. The evolution of sexual dimorphism in gene expression in response to a manipulation of mate competition. Evolution 78:746–757. https://doi.org/10.1093/evolut/qpae004

Møller, A. P., and R. V. Alatalo. 1999. Good-genes effects in sexual selection. Proc. R. Soc. Lond. B Biol. Sci. 266:85–91. https://doi.org/10.1098/rspb.1999.0607

Montgomery, S. H., and J. E. Mank. 2016. Inferring regulatory change from gene expression: the confounding effects of tissue scaling. Mol. Ecol. 25:5114–5128. https://doi.org/10.1111/mec.13824

Naurin, S., B. Hansson, D. Hasselquist, Y.-H. Kim, and S. Bensch. 2011. The sex-biased brain: sexual dimorphism in gene expression in two species of songbirds. BMC Genomics 12:37. https://doi.org/10.1186/1471-2164-12-37

Palaoro, A. V., and P. E. C. Peixoto. 2022. The hidden links between animal weapons, fighting style, and their effect on contest success: a meta-analysis. Biol. Rev. 97:1948–1966. https://doi.org/10.1111/brv.12877

Parisi, M., R. Nuttall, P. Edwards, J. Minor, D. Naiman, J. Lü, M. Doctolero, M. Vainer, C. Chan, J. Malley, S. Eastman, and B. Oliver. 2004. A survey of ovary-, testis-, and soma-biased gene expression in *Drosophila melanogaster* adults. Genome Biol. 18. https://doi.org/10.1186/gb-2004-5-6-r40

Perry, J. C., P. W. Harrison, and J. E. Mank. 2014. The Ontogeny and Evolution of Sex-Biased Gene Expression in *Drosophila melanogaster*. Mol. Biol. Evol. 31:1206–1219. https://doi.org/10.1093/molbev/msu072

Petrie, M. 1994. Improved growth and survival of offspring of peacocks with more elaborate trains. Nature 371:598–599. https://doi.org/10.1093/molbev/msu072

Pointer, M. A., P. W. Harrison, A. E. Wright, and J. E. Mank. 2013. Masculinization of Gene Expression Is Associated with Exaggeration of Male Sexual Dimorphism. PLOS Genet. 9:e1003697. https://doi.org/10.1371/journal.pgen.1003697

Pomiankowski, A., and A. P. Møller. 1997. A resolution of the lek paradox. Proc. R. Soc. Lond. B Biol. Sci. 260:21–29. https://doi.org/10.1098/rspb.1995.0054

Rago, A., J. H. Werren, and J. K. Colbourne. 2020. Sex biased expression and co-expression networks in development, using the hymenopteran *Nasonia vitripennis*. PLoS Genet. 16:e1008518. https://doi.org/10.1371/journal.pgen.1008518

Fuller C., Houle D., and Travis J. 2005. Sensory Bias as an Explanation for the Evolution of Mate Preferences. Am. Nat. 166:437–446. https://doi.org/10.1086/444443

Rodríguez-Montes, L., S. Ovchinnikova, X. Yuan, T. Studer, I. Sarropoulos, S. Anders, H. Kaessmann, and M. Cardoso-Moreira. 2023. Sex-biased gene expression across mammalian organ development and evolution. Science 382:eadf1046. https://doi.org/10.1126/science.adf1046

Rohner, P. T., D. M. Linz, and A. P. Moczek. 2021. Doublesex mediates species-, sex-, environment- and trait-specific exaggeration of size and shape. Proc. R. Soc. B Biol. Sci. 288:20210241. https://doi.org/10.1098/rspb.2021.0241

Rowe, L., and D. Houle. 1996. The lek paradox and the capture of genetic variance by condition dependent traits. Proc. R. Soc. Lond. B Biol. Sci. 263:1415–1421. https://doi.org/10.1098/rspb.1996.0207

Ryan, M. J. 1998. Sexual Selection, Receiver Biases, and the Evolution of Sex Differences. Science 281:1999–2003. https://doi.org/10.1126/science.281.5385.1999

Sepil, I., J. C. Perry, A. Dore, T. Chapman, and S. Wigby. 2022. Experimental evolution under varying sex ratio and nutrient availability modulates male mating success in *Drosophila melanogaster*. Biol. Lett. 18:20210652. https://doi.org/10.1098/rsbl.2021.0652

Sharma, E., A. Künstner, B. A. Fraser, G. Zipprich, V. A. Kottler, S. R. Henz, D. Weigel, and C. Dreyer. 2014. Transcriptome assemblies for studying sex-biased gene expression in the guppy, *Poecilia reticulata*. BMC Genomics 15:400. https://doi.org/10.1186/1471-2164-15-400

Singh, A., and A. F. Agrawal. 2023. Two forms of sexual dimorphism in gene expression in *Drosophila melanogaster* : their coincidence and evolutionary genetics. 40: 5 msad091. https://doi.org/10.1093/molbev/msad091

Stephens, M., P. Carbonetto, D. Gerard, M. Lu, L. Sun, J. Willwerscheid, and N. Xiao. 2016. ashr: Methods for Adaptive Shrinkage, using Empirical Bayes.

Stewart, A. D., A. Pischedda, and W. R. Rice. 2010. Resolving Intralocus Sexual Conflict: Genetic Mechanisms and Time Frame. J. Hered. 101:S94–S99. https://doi.org/10.1093/jhered/esq011

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. https://doi.org/10.1038/s41559-018-0613-4

Stuglik, M. T., W. Babik, Z. Prokop, and J. Radwan. 2014. Alternative reproductive tactics and sex-biased gene expression: the study of the bulb mite transcriptome. Ecol. Evol. 4:623–632. https://doi.org/10.1002/ece3.965

Sutherland, B. J. G., J. M. Prokkola, C. Audet, and L. Bernatchez. 2019. Sex-Specific Co-expression Networks and Sex-Biased Gene Expression in the Salmonid Brook Charr Salvelinus fontinalis. G3 Genes|Genomes|Genetics 9:955–968. https://doi.org/10.1534/g3.118.200910

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. https://doi.org/10.1093/evolut/qpad075

Tang-Martínez, Z. 2016. Rethinking Bateman's Principles: Challenging Persistent Myths of Sexually Reluctant Females and Promiscuous Males. J. Sex Res. 53:532–559. https://doi.org/10.1080/00224499.2016.1150938

Taylor, P. D., and G. C. Williams. 1982. The lek paradox is not resolved. Theor. Popul. Biol. 22:392–409. https://doi.org/10.1016/0040-5809(82)90052-1

Toubiana, W., D. Armisén, C. Dechaud, R. Arbore, and A. Khila. 2021. Impact of male trait exaggeration on sex-biased gene expression and genome architecture in a water strider. BMC Biol. 19:89. https://doi.org/10.1186/s12915-021-01021-4

Trivers, R. L. 1974. Parent-Offspring Conflict. Am. Zool. 14:249–264. https://doi.org/10.1093/icb/14.1.249

Van De Wiel, M. A., G. G. R. Leday, L. Pardo, H. Rue, A. W. Van Der Vaart, and W. N. Van Wieringen. 2013. Bayesian analysis of RNA sequencing data by estimating multiple shrinkage priors. Biostatistics 14:113–128. https://doi.org/10.1093/biostatistics/kxs031

Voje, K. L. 2016. Scaling of Morphological Characters across Trait Type, Sex, and Environment. Am. Nat. 187:89–98. https://doi.org/10.1086/684159

Warren, I. A., H. Gotoh, I. M. Dworkin, D. J. Emlen, and L. C. Lavine. 2013. A general mechanism for conditional expression of exaggerated sexually-selected traits. BioEssays 35:889–899. https://doi.org/10.1002/bies.201300031

Wilkinson, G. S., and G. N. Dodson. 1997. Function and evolution of antlers and eye stalks in flies. Pp. 310–328 The Evolution of Mating Systems in Insects and Arachnids.

Wilkinson, G. S., P. M. Johns, J. D. Metheny, and R. H. Baker. 2013. Sex-Biased Gene Expression during Head Development in a Sexually Dimorphic Stalk-Eyed Fly. PLoS One 8:e59826. https://doi.org/10.1371/journal.pone.0059826

Wong, R. Y., M. M. McLeod, and J. Godwin. 2014. Limited sex-biased neural gene expression patterns across strains in Zebrafish (*Danio rerio*). BMC Genomics 15:905. https://doi.org/10.1186/1471-2164-15-905

Zhang, B., and S. Horvath. 2005. A General Framework for Weighted Gene Co-Expression Network Analysis. Stat. Appl. Genet. Mol. Biol. 4. https://doi.org/10.2202/1544-6115.1128

Zhang, Y., K. Klein, A. Sugathan, N. Nassery, A. Dombkowski, U. M. Zanger, and D. J. Waxman. 2011. Transcriptional Profiling of Human Liver Identifies Sex-Biased Genes Associated with Polygenic Dyslipidemia and Coronary Artery Disease. PLoS One 6:e23506. https://doi.org/10.1371/journal.pone.0023506

Zhu, A., J. G. Ibrahim, and M. I. Love. 2019. Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. Bioinformatics 35:2084–2092. https://doi.org/10.1093/bioinformatics/bty895

Zinna, R., D. Emlen, L. C. Lavine, A. Johns, H. Gotoh, T. Niimi, and I. Dworkin. 2018. Sexual dimorphism and heightened conditional expression in a sexually selected weapon in the Asian rhinoceros beetle. Mol. Ecol. 27:5049–5072. https://doi.org/10.1111/mec.14907

Chapter 2 supplement

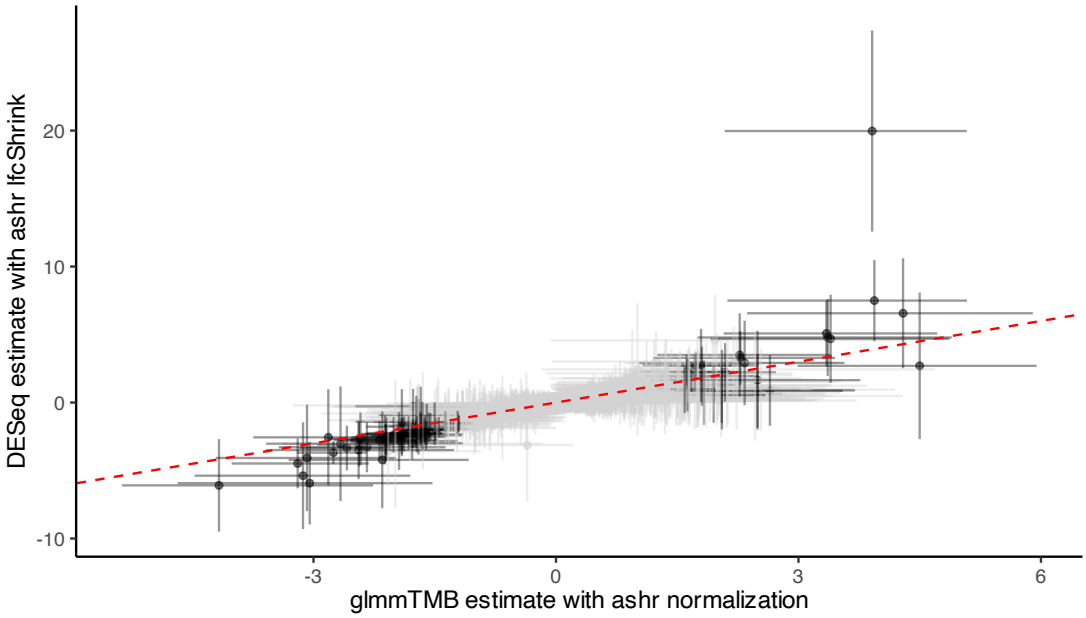

Figure S1: DeSeq2 compared glmmTMB estimate for log_2 fold change foreleg genes in *D. prolongata* early in development comparison. glmmTMB estimates regularized with ashr and adjusted 95% confidence intervals, DESeq estimate intervals are 2*SE. Red dotted line at a slope of 1 and intercept of zero.

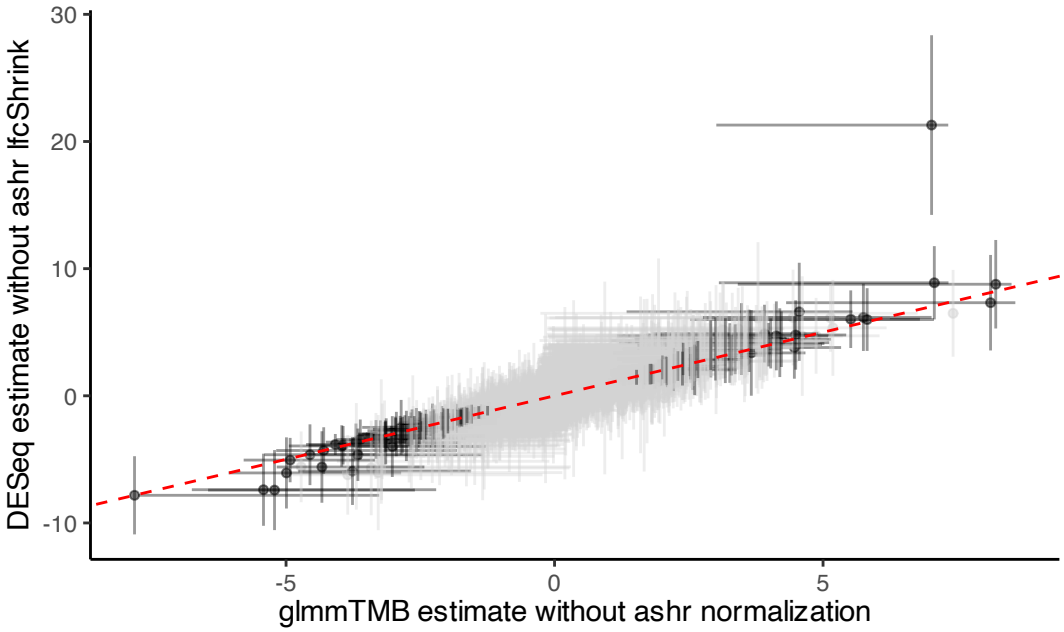

Figure S2: DeSeq2 compared glmmTMB estimate for $\log_2$ fold change foreleg genes in *D. prolongata* early in development comparison, without regularization. DESeq estimate intervals are 2*SE. Red dotted line at a slope of 1 and intercept of zero.

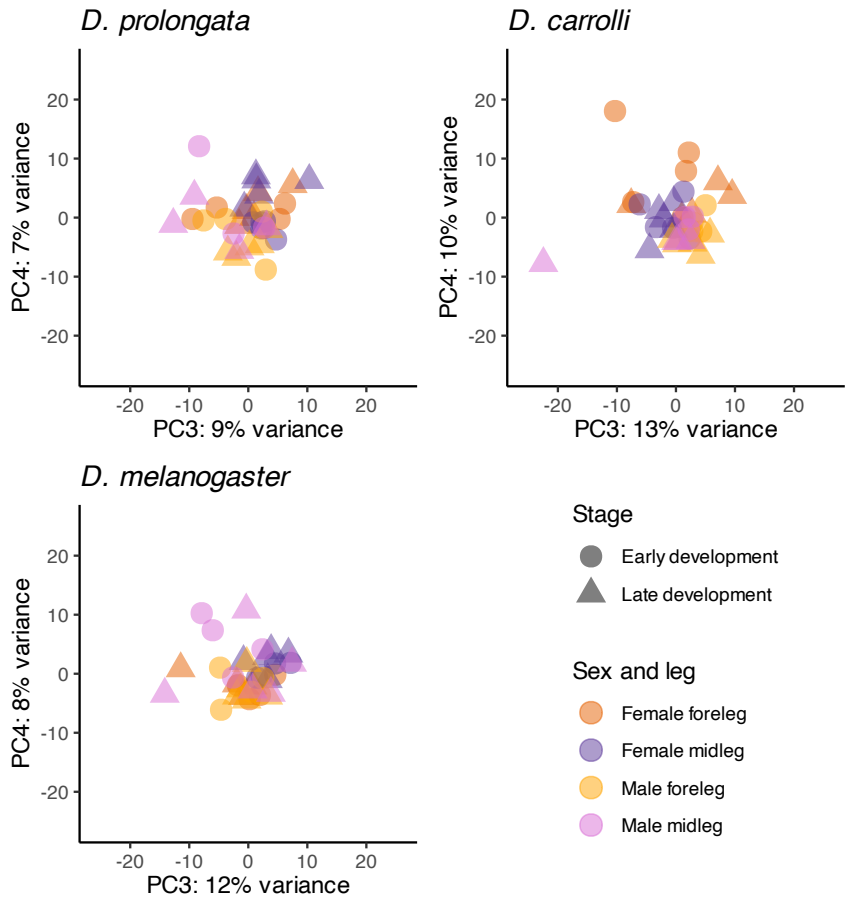

Figure S3: PC3 and PC4 within species with the top 1000 most variables genes for sex, stage, and leg.

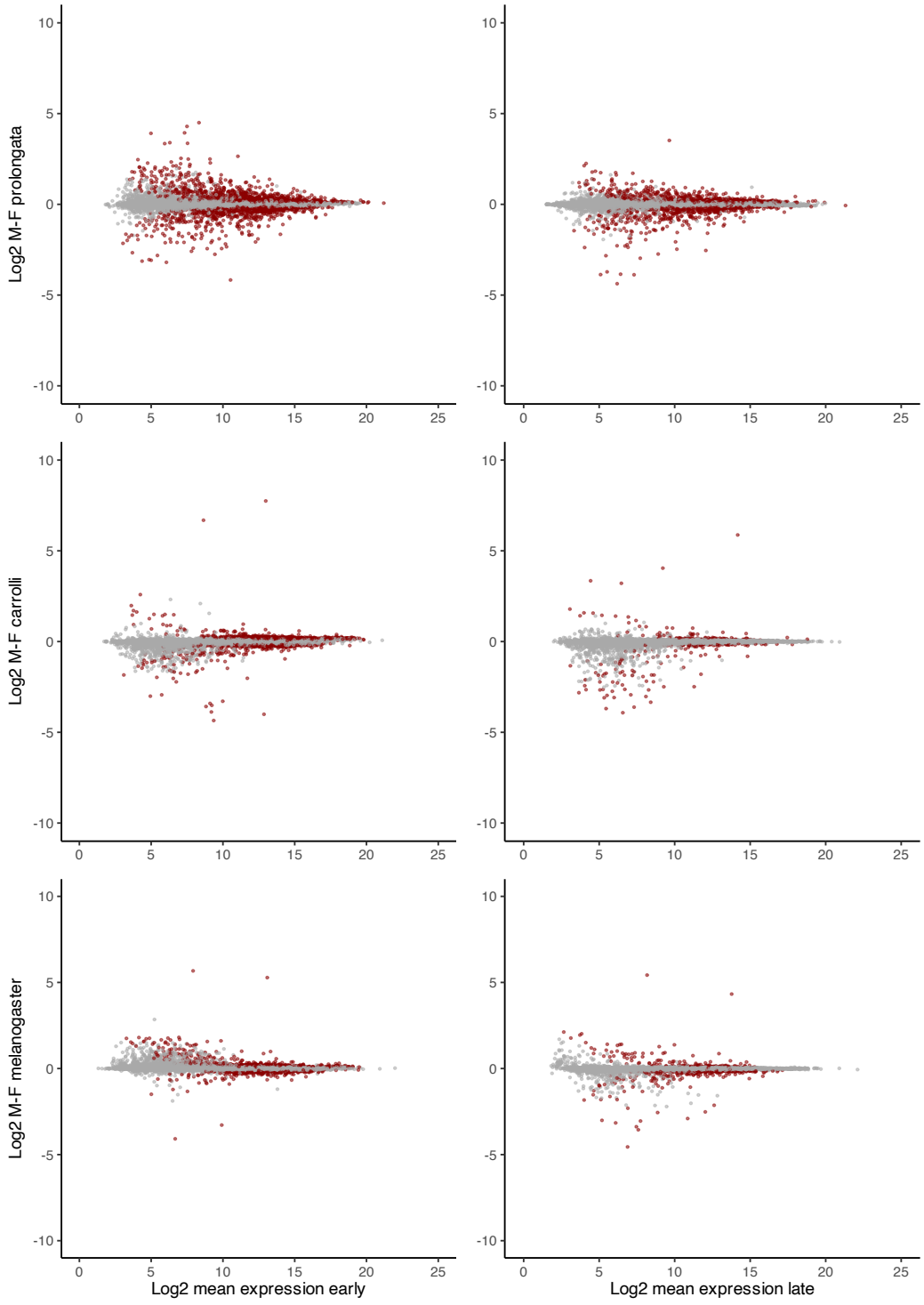

Figure S4: MA plot of sex-biased gene expression in the foreleg imaginal disc for each species. Estimated fold changes (computed with emmeans) was regularized with ashr. Red dots represent genes with ashr corrected 95% CIs that do not overlap 0.

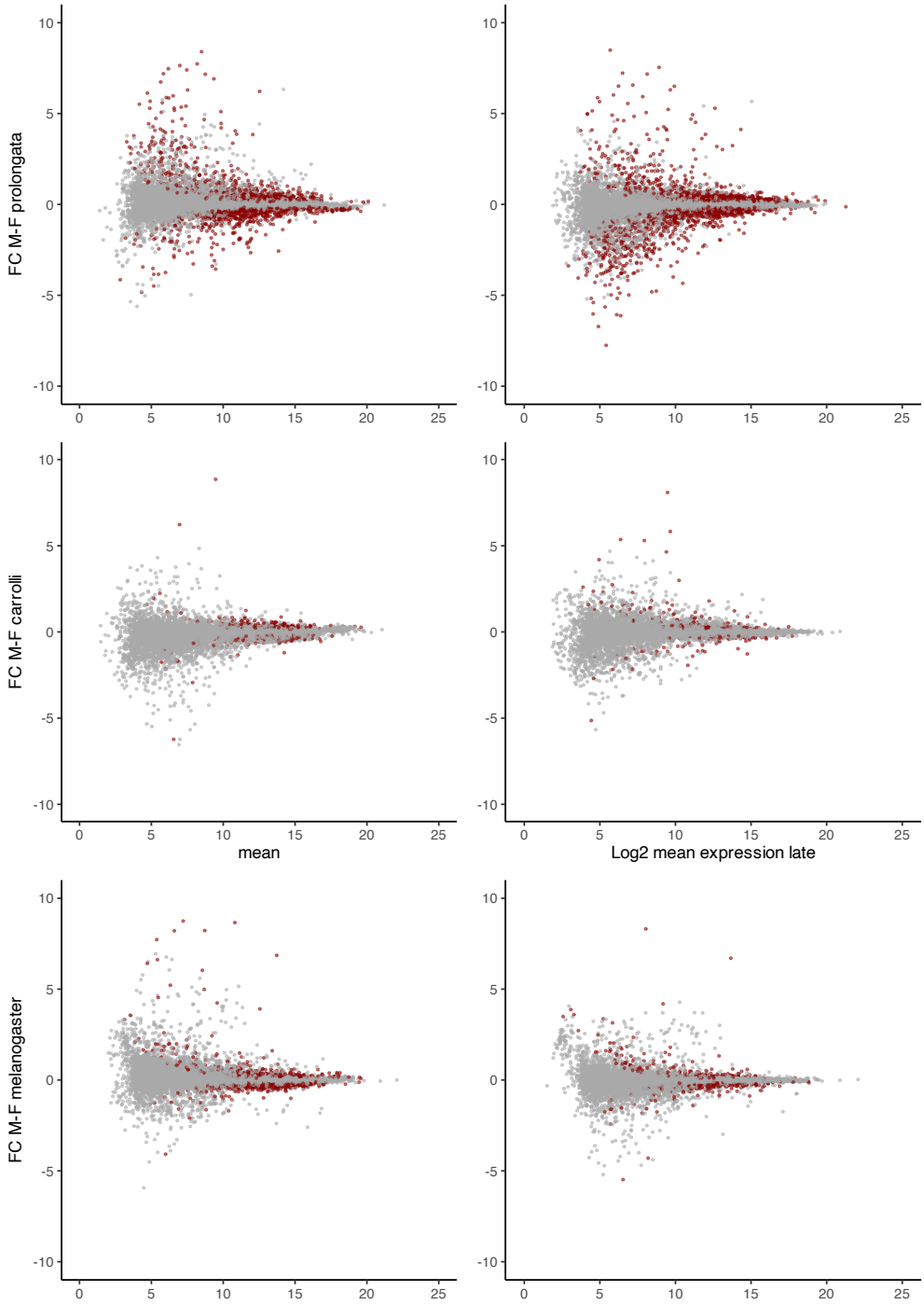

Figure S5: Figure S4: MA plot of sex-biased gene expression in the foreleg imaginal disc for each species. Estimated fold changes (computed with emmeans) was not regularized for comparison to Figure S4. Red dots represent genes whose 95% CIs that do not overlap 0.

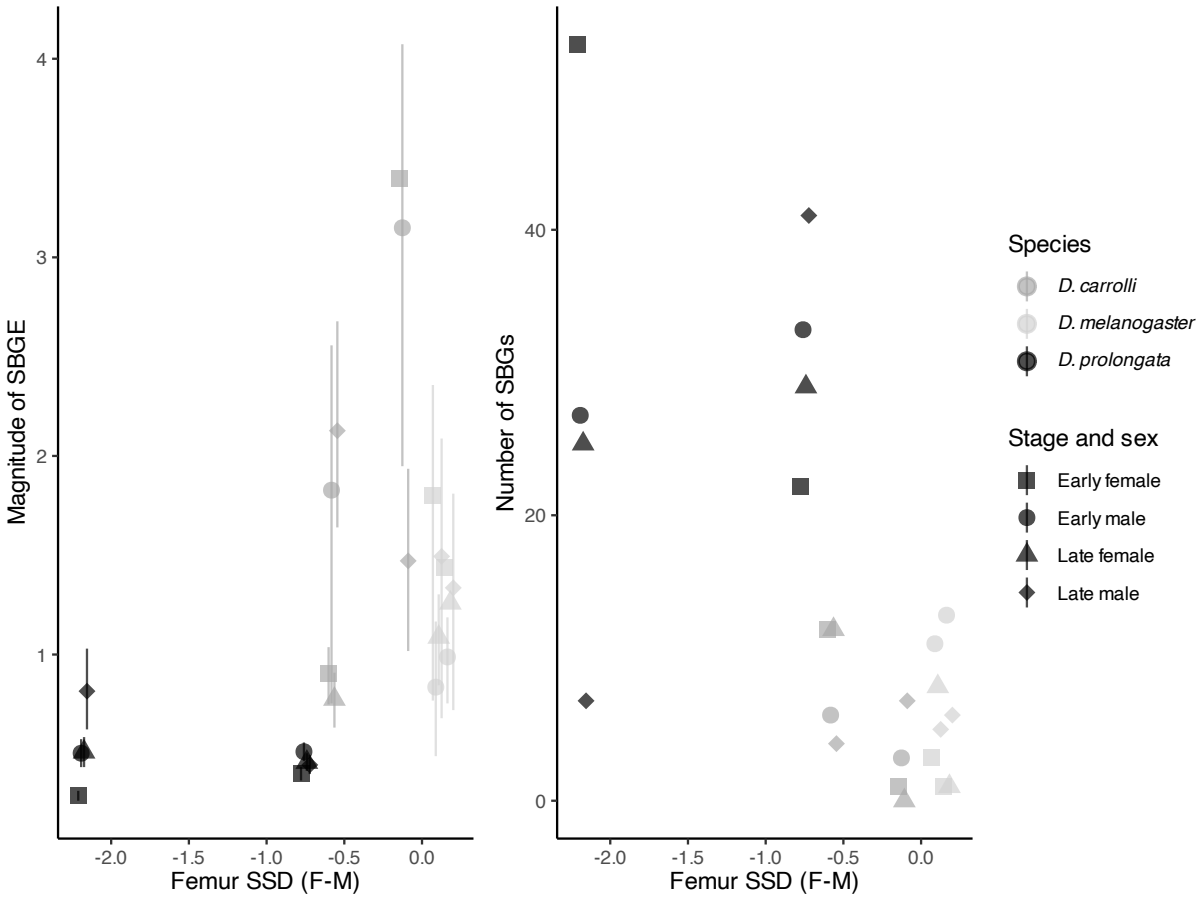

Figure S6: SSD and magnitude of SBGE or number of Log_2 SBGs

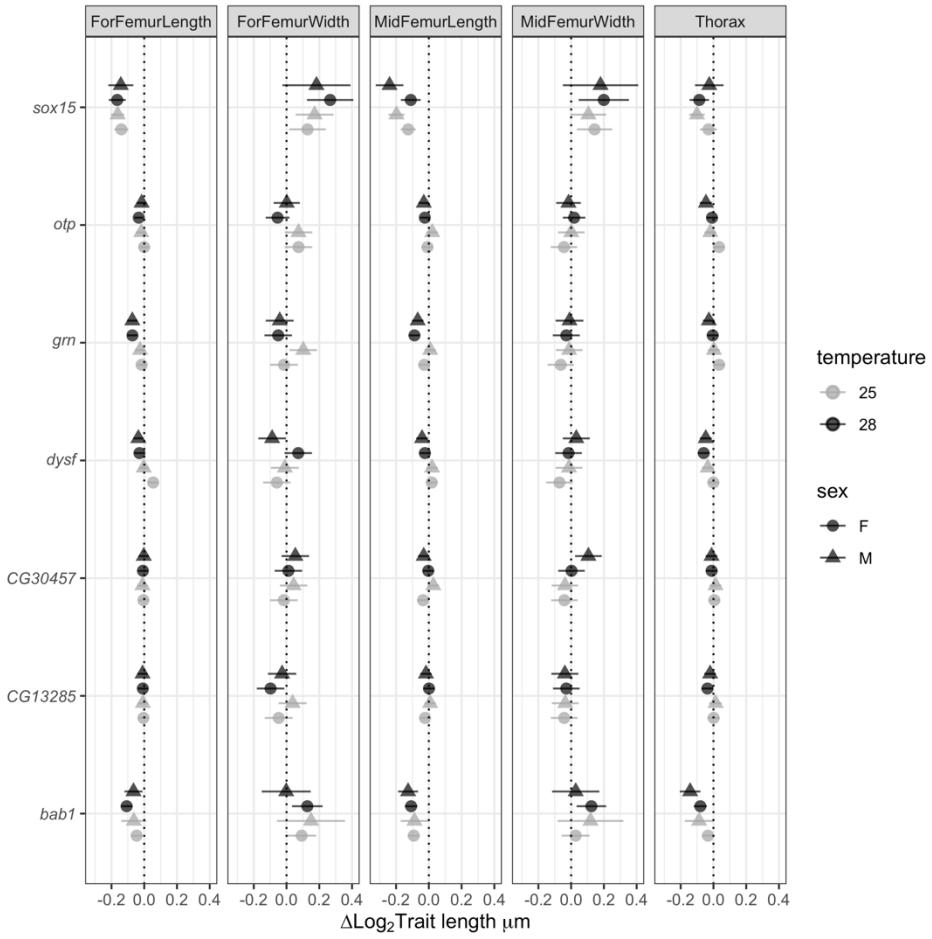

Figure S7: *pen*-Gal4 (NP6333-Gal4) knockdowns of candidate genes and their change relative to control crosses.

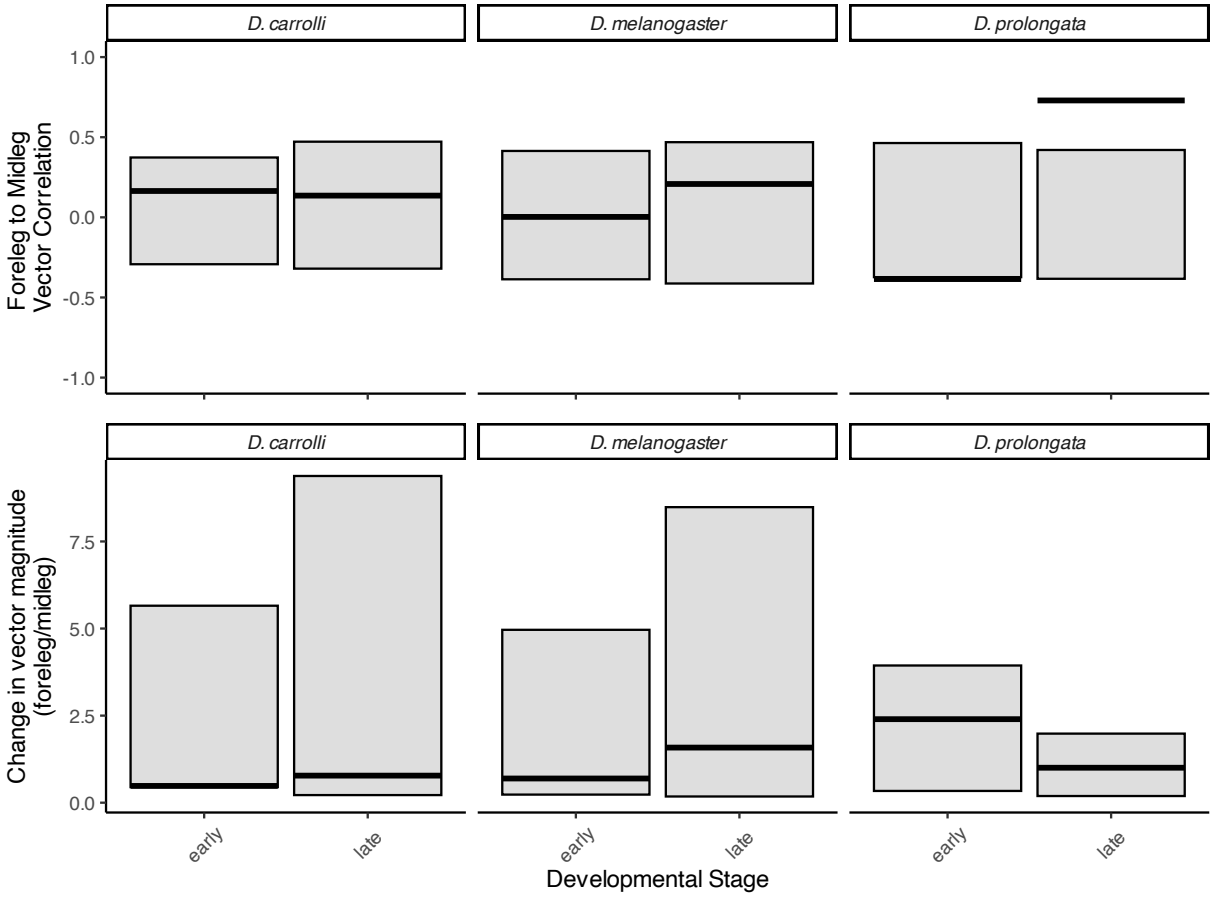

Figure S8: Vector direction and change in magnitude of log_2 fold biased genes within species between foreleg and midleg. Based on male-biased Log^2 fold change genes that overlap in all species. Top row shows direction of SBGE in foreleg compared to midleg within all three species. Bottom row is the magnitude of SBGE in foreleg relative to midleg within each species.

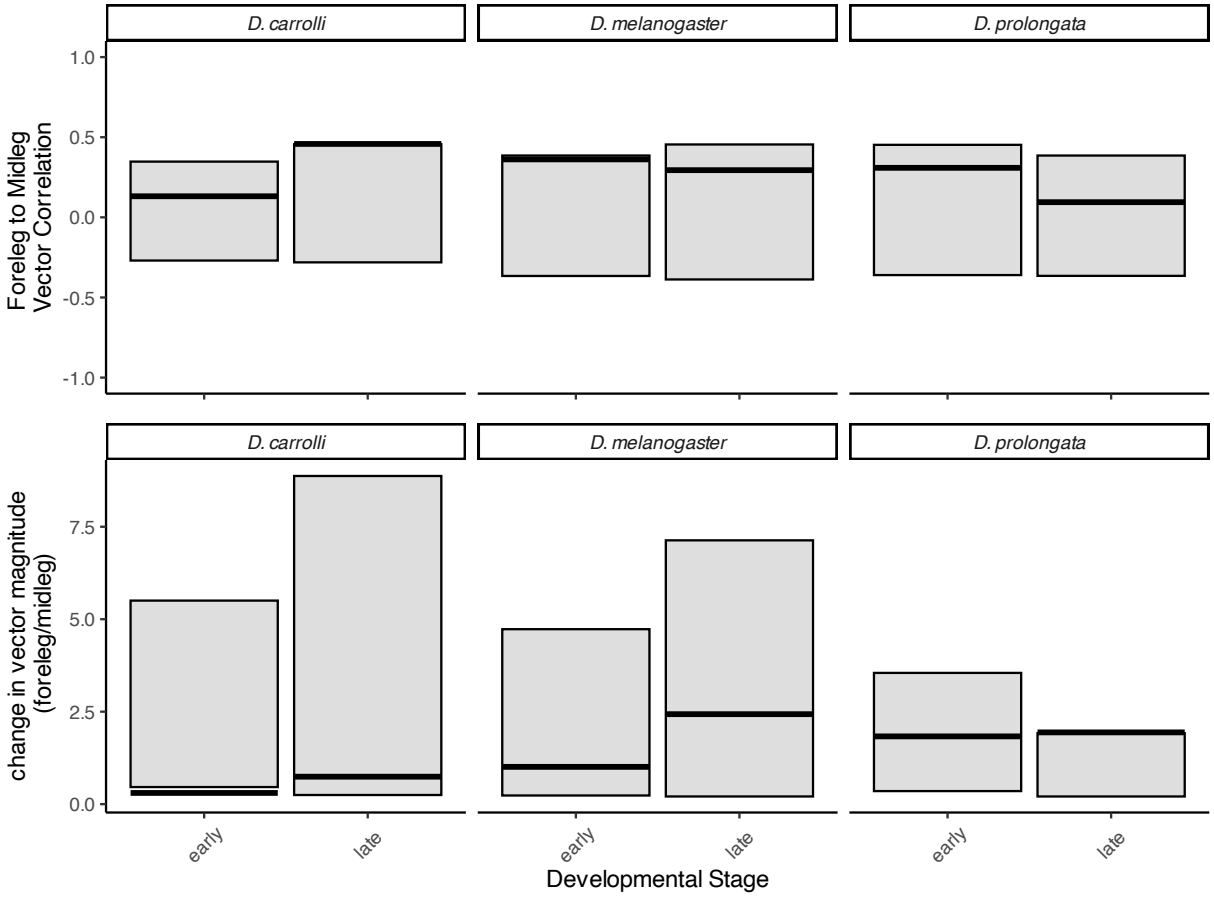

Figure S9: Vector direction and magnitude of log_2 fold biased genes within species between foreleg and midleg. Based on female-biased Log^2 fold change genes that overlap in all species. Top row shows direction of SBGE in foreleg compared to midleg within all three species. Bottom row is the magnitude of SBGE in foreleg relative to midleg within each species.

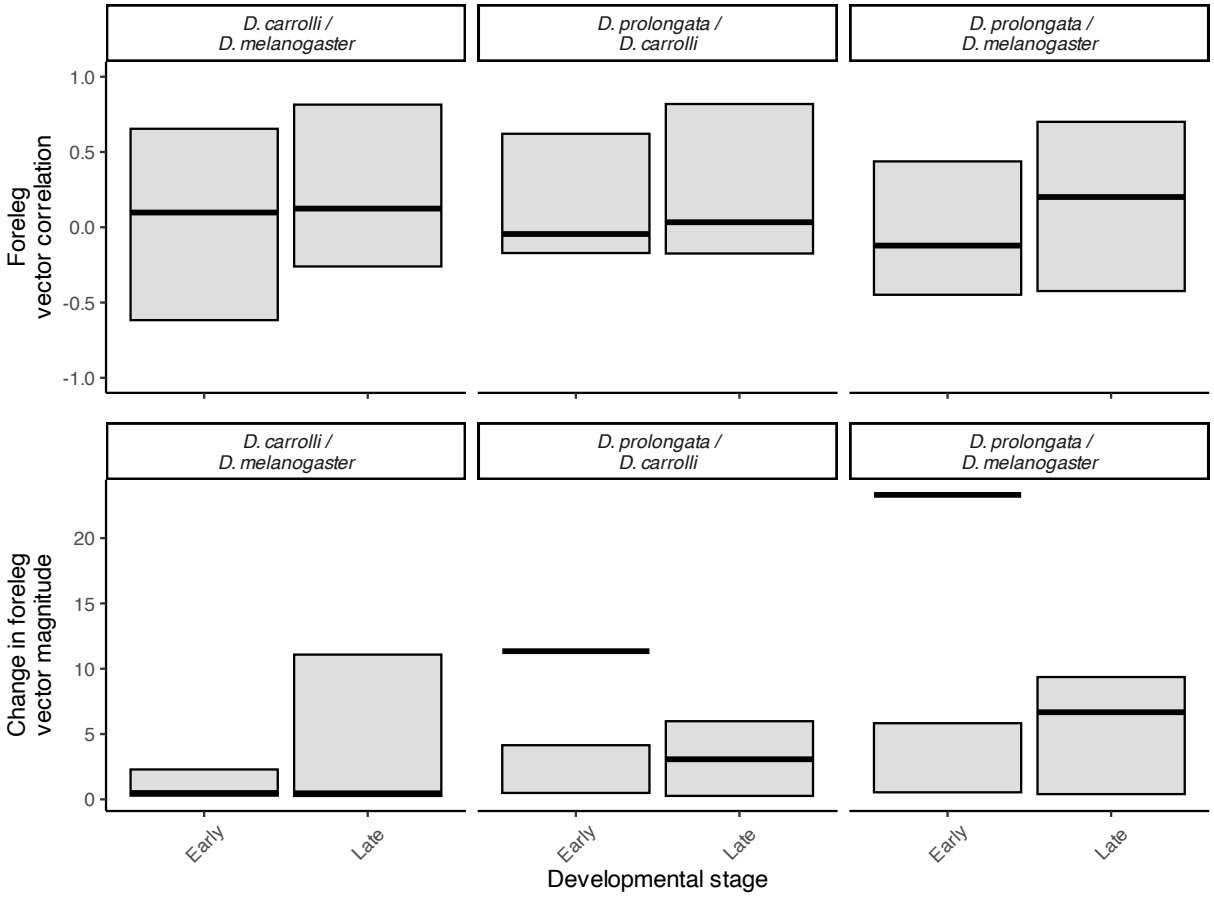

Figure S10: Direction and magnitude of male-biased Log_2 fold biased genes between species in the foreleg. Based on ~80 genes that overlap in all species. Top row shows direction of SBGE in foreleg compared between D. prolongata and the other species, the first column shows *D. carrolli* compared to *D. melanogaster* as a point of reference. Bottom row is the magnitude of SBGE in foreleg between each species.

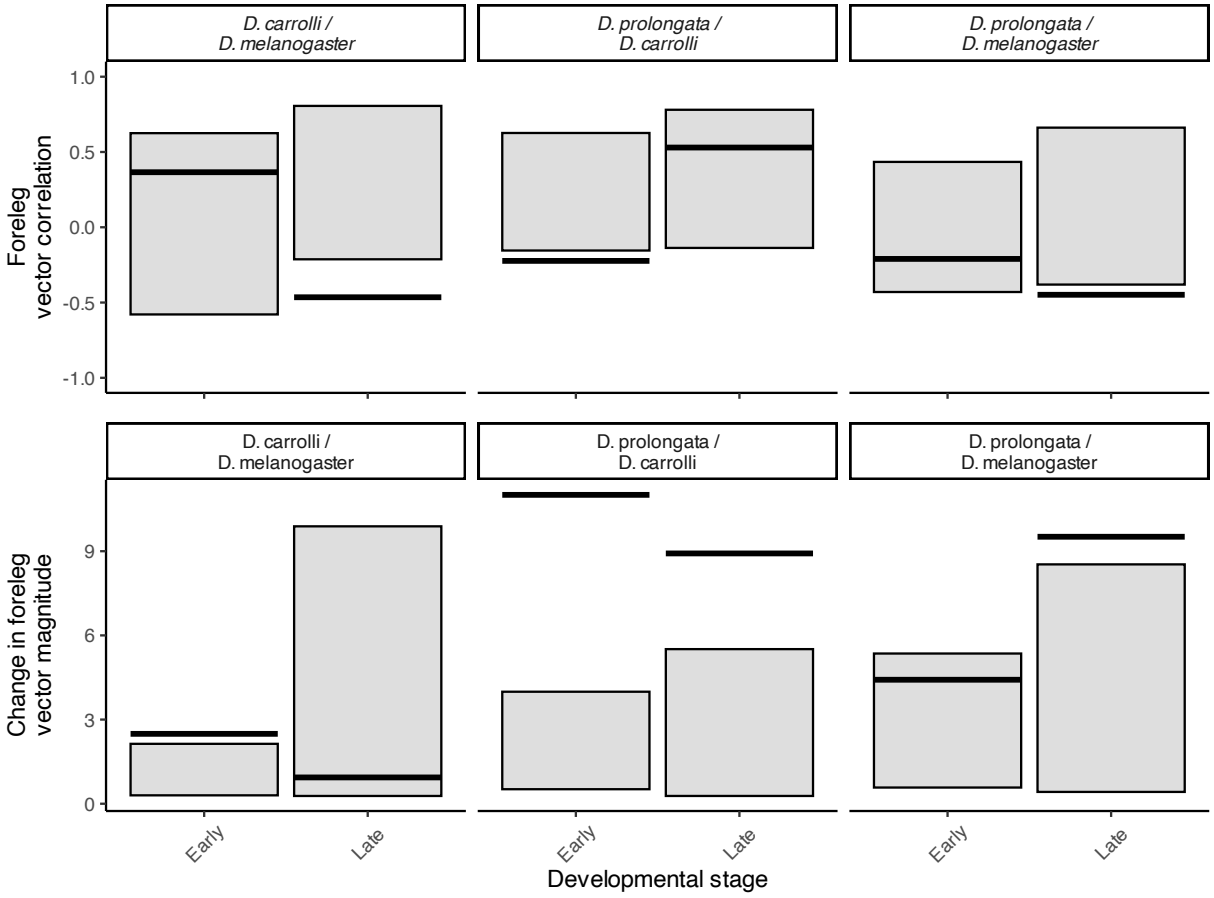

Figure S11: Direction and magnitude of female-biased Log_2 fold biased genes between species in the foreleg. Based on ~80 genes that overlap in all species. Top row shows direction of SBGE in foreleg compared between D. prolongata and the other species, the first column shows *D. carrolli* compared to *D. melanogaster* as a point of reference. Bottom row is the magnitude of SBGE in foreleg between each species.

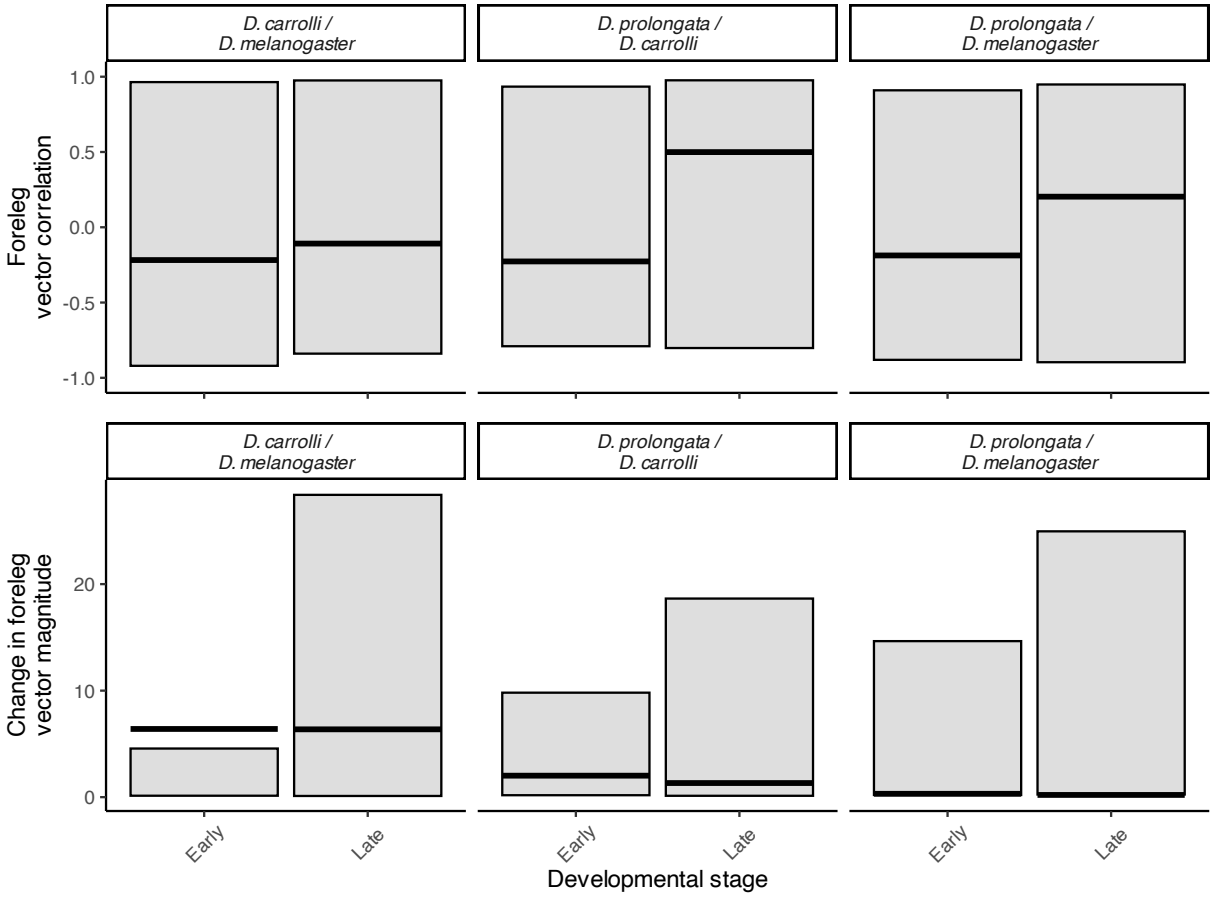

Figure S12: Direction and magnitude of *D. melanogaster* Log_2 fold biased genes between species in the foreleg. Based on ~80 genes that overlap in all species. Top row shows direction of SBGE in foreleg compared between D. prolongata and the other species, the first column shows *D. carrolli* compared to *D. melanogaster* as a point of reference. Bottom row is the magnitude of SBGE in foreleg between each species.

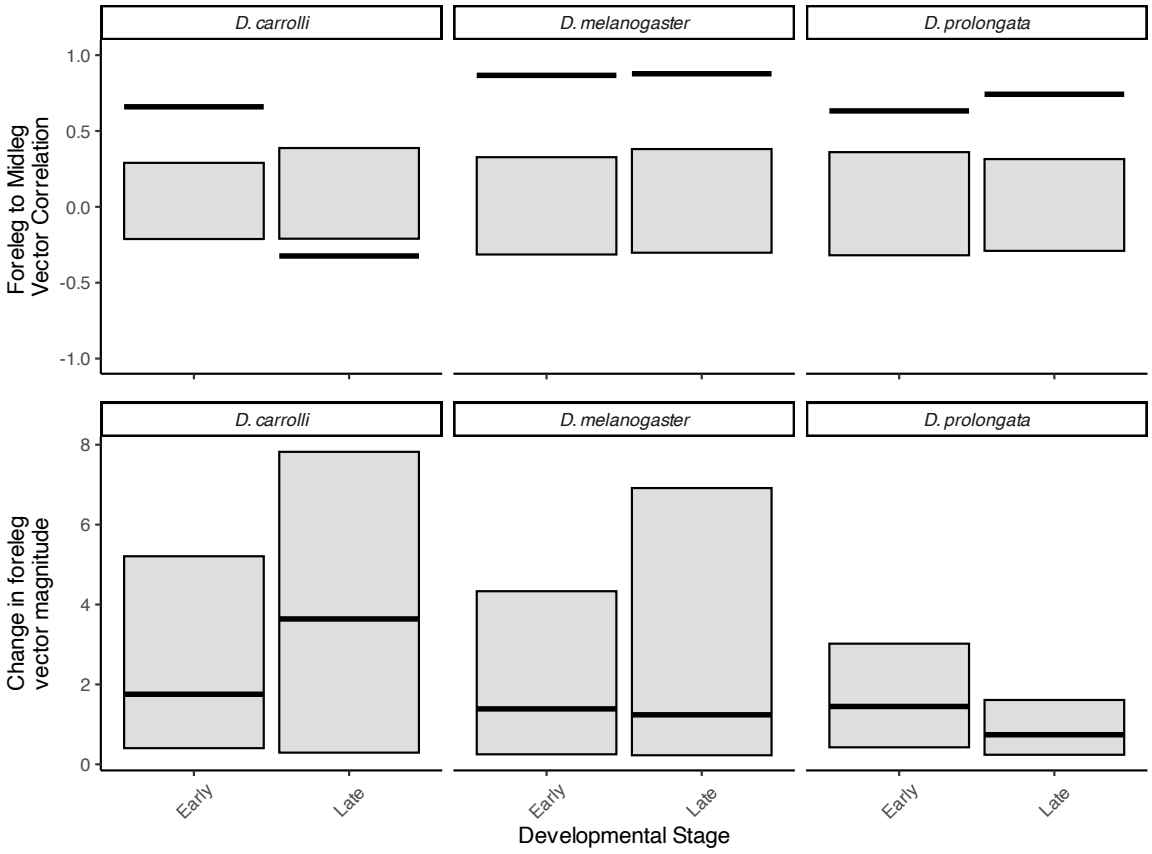

Figure S13: Vector direction and magnitude of Insulin signalling genes within species between foreleg and midleg. Based on genes that overlap in all species. Top row shows direction of SBGE in foreleg compared to midleg within all three species. Bottom row is the magnitude of SBGE in foreleg relative to midleg within each species.

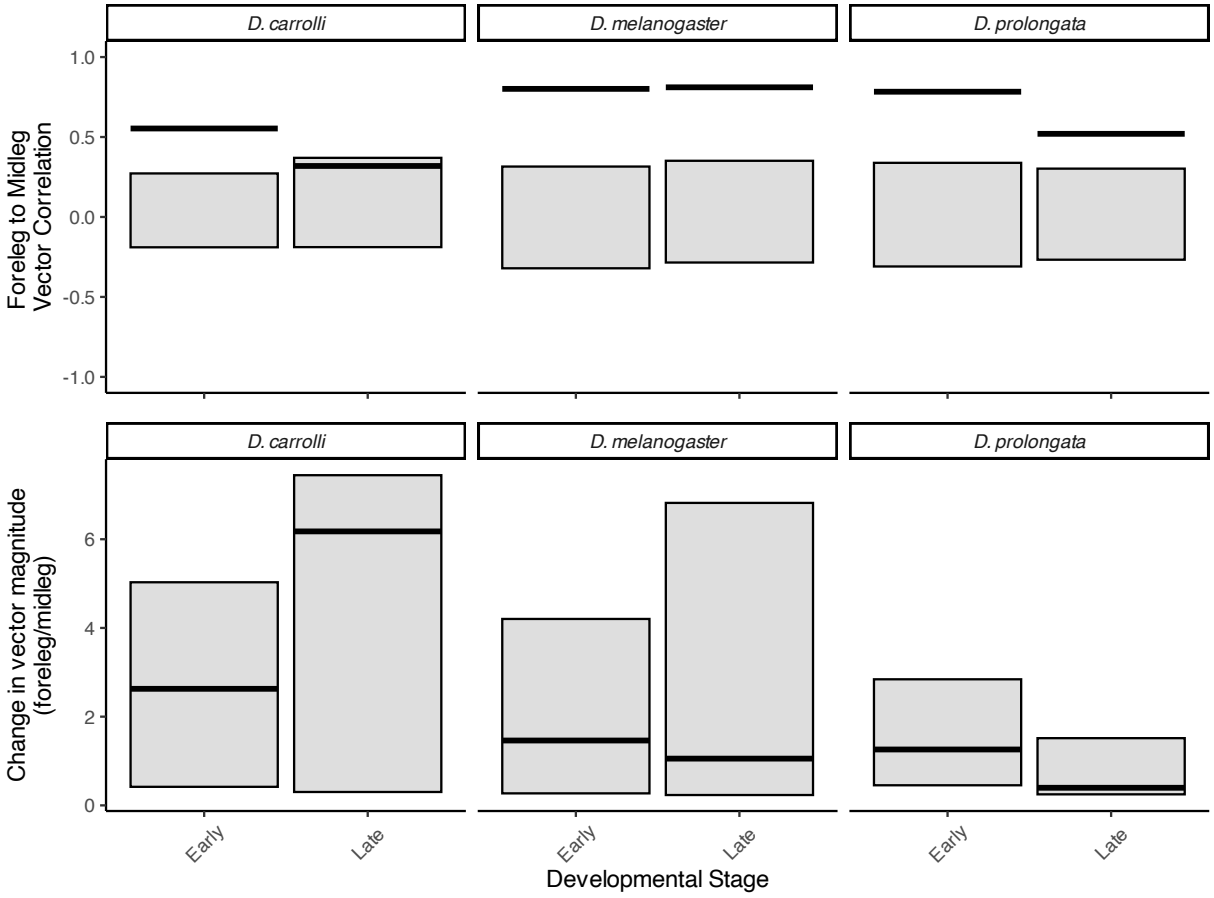

Figure S13: Vector direction and magnitude of Notch signalling genes within species between foreleg and midleg. Based on genes that overlap in all species. Top row shows direction of SBGE in foreleg compared to midleg within all three species. Bottom row is the magnitude of SBGE in foreleg relative to midleg within each species.

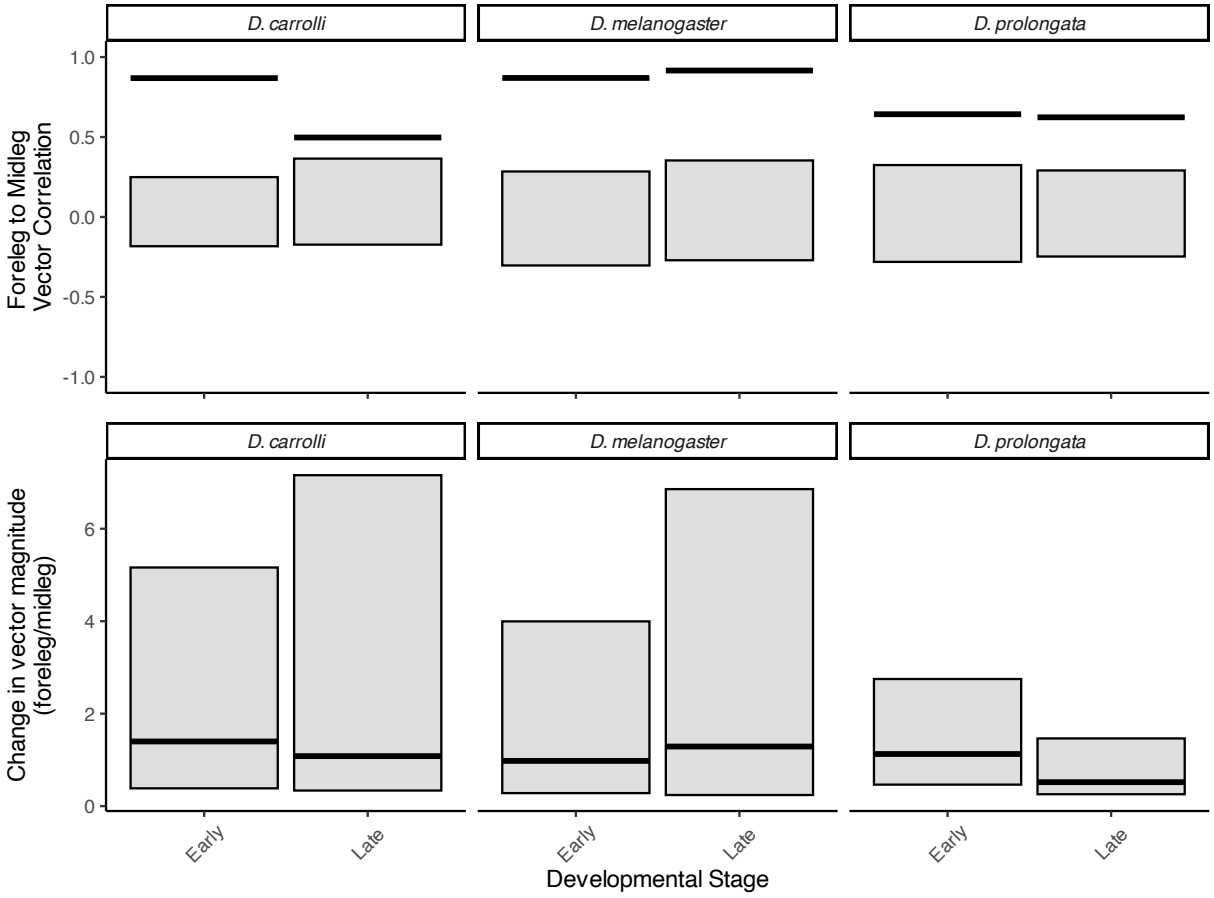

Figure S14: Vector direction and magnitude of Hippo signalling genes within species between foreleg and midleg. Based on genes that overlap in all species. Top row shows direction of SBGE in foreleg compared to midleg within all three species. Bottom row is the magnitude of SBGE in foreleg relative to midleg within each species.

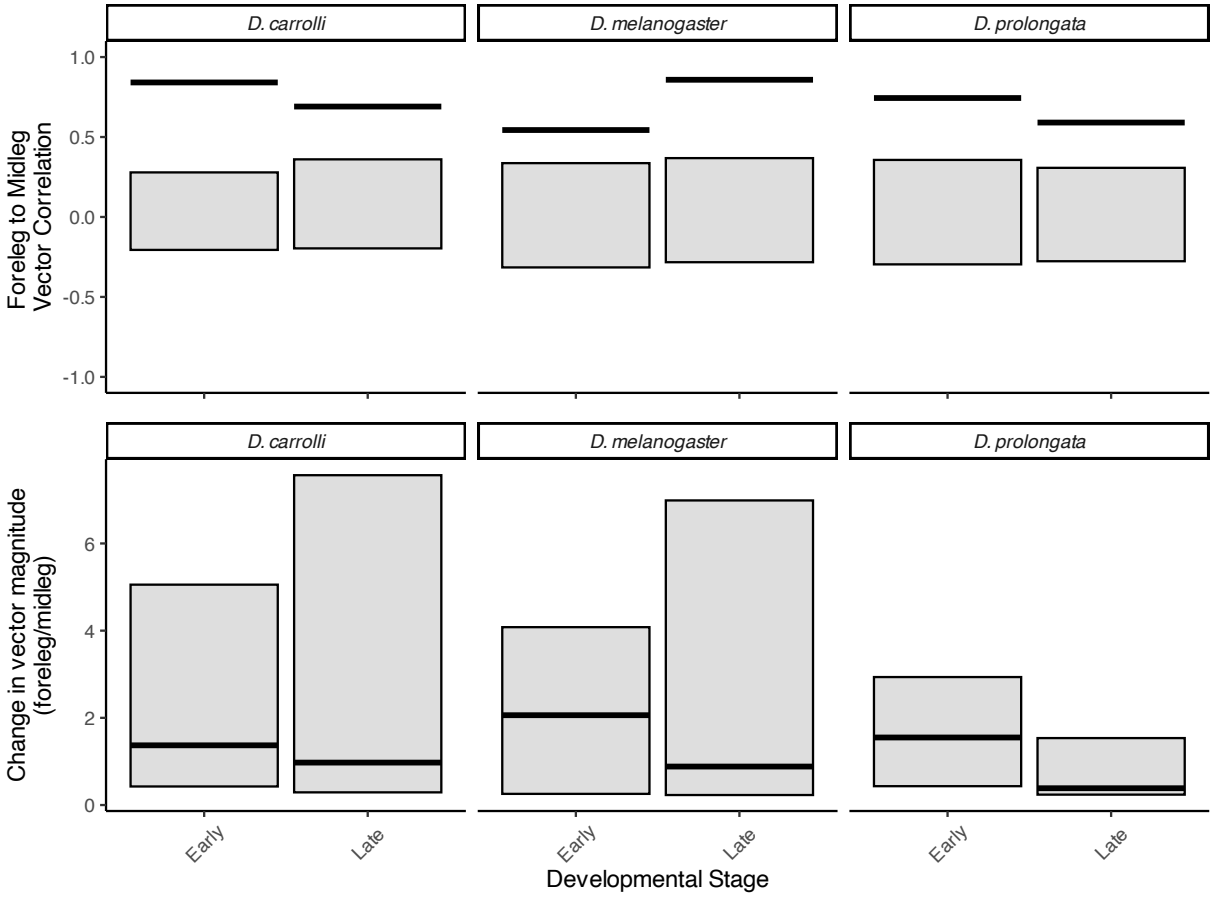

Figure S15: Vector direction and magnitude of EGFR signalling genes within species between foreleg and midleg. Based on genes that overlap in all species. Top row shows direction of SBGE in foreleg compared to midleg within all three species. Bottom row is the magnitude of SBGE in foreleg relative to midleg within each species.

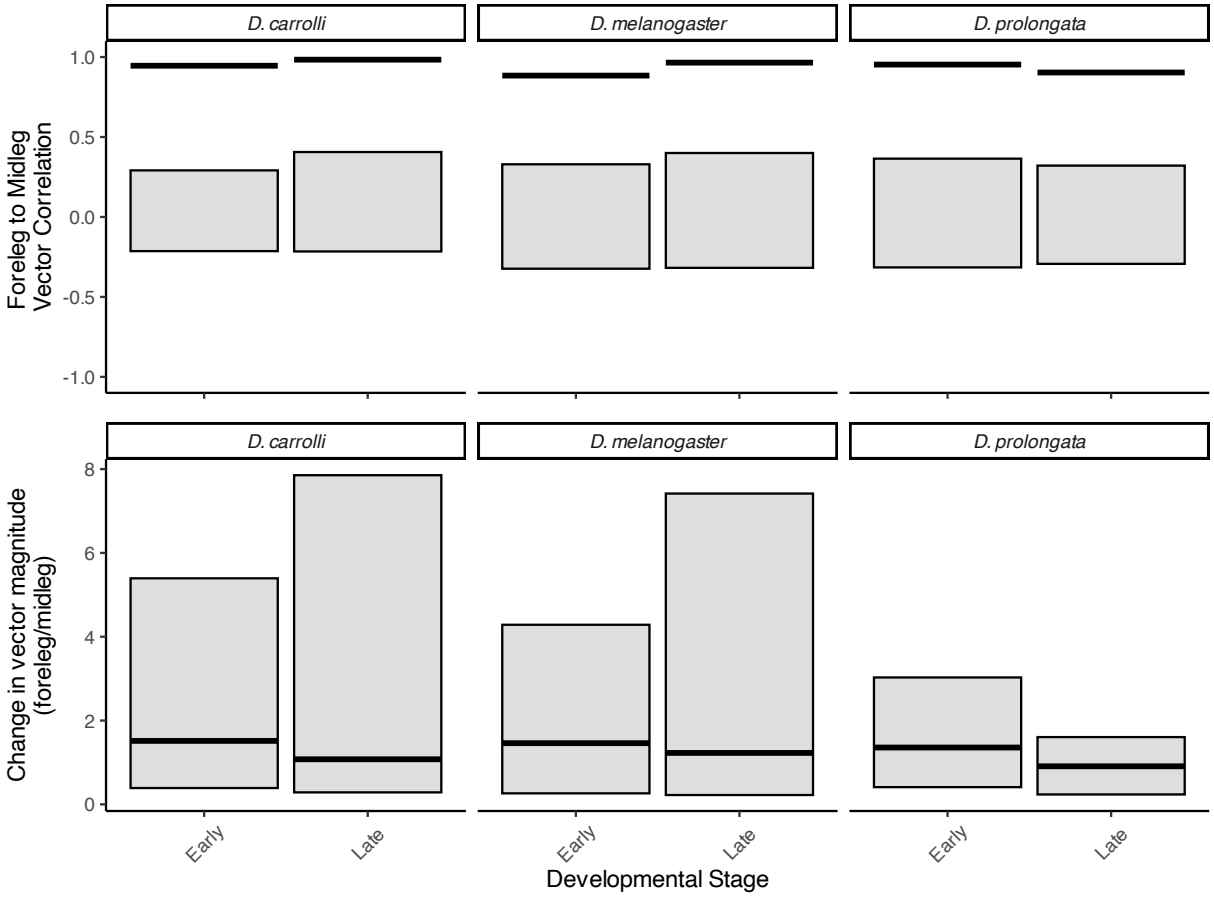

Figure S16: Vector direction and magnitude of BMP signalling genes within species between foreleg and midleg. Based on genes that overlap in all species. Top row shows direction of SBGE in foreleg compared to midleg within all three species. Bottom row is the magnitude of SBGE in foreleg relative to midleg within each species.

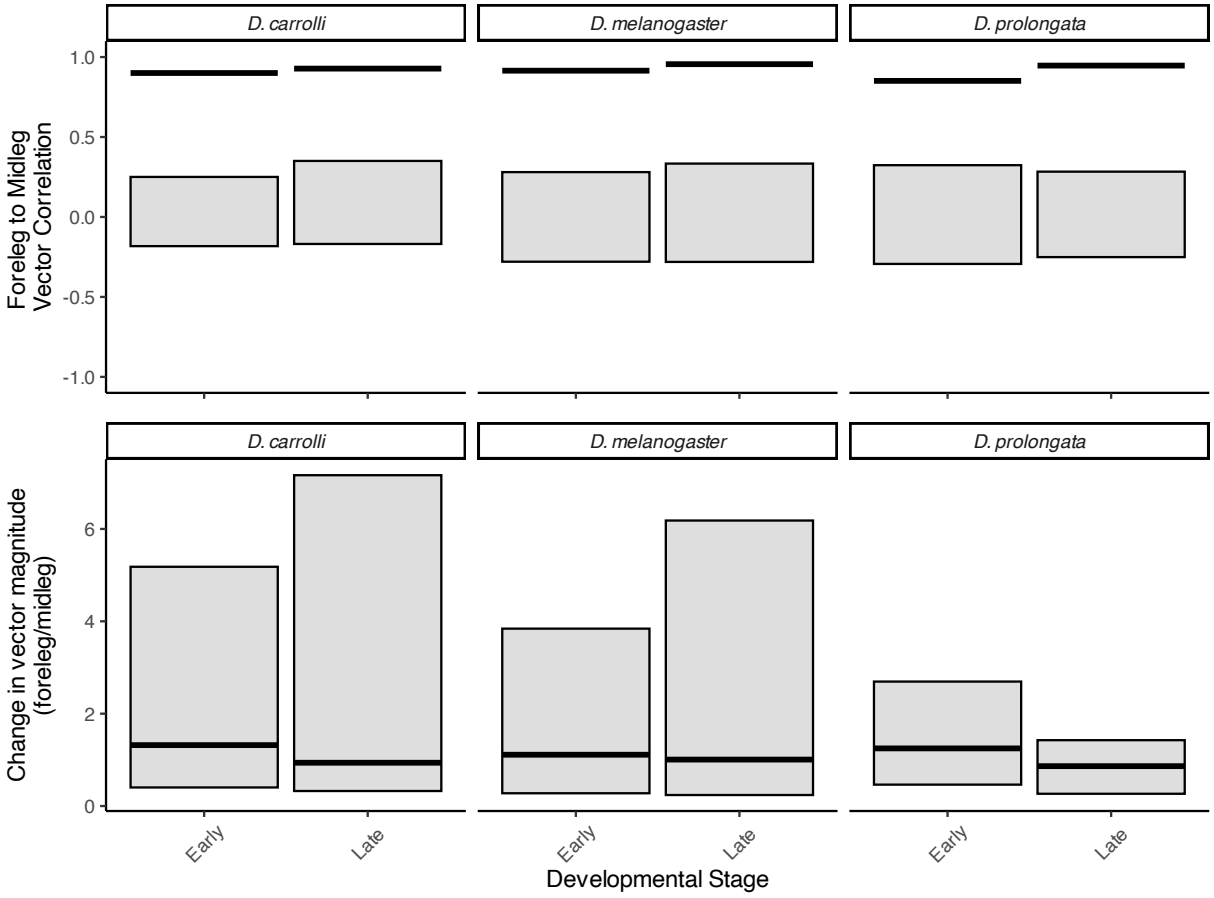

Figure S17: Vector direction and magnitude of Hedgehog signalling genes within species between foreleg and midleg. Based on genes that overlap in all species. Top row shows direction of SBGE in foreleg compared to midleg within all three species. Bottom row is the magnitude of SBGE in foreleg relative to midleg within each species.

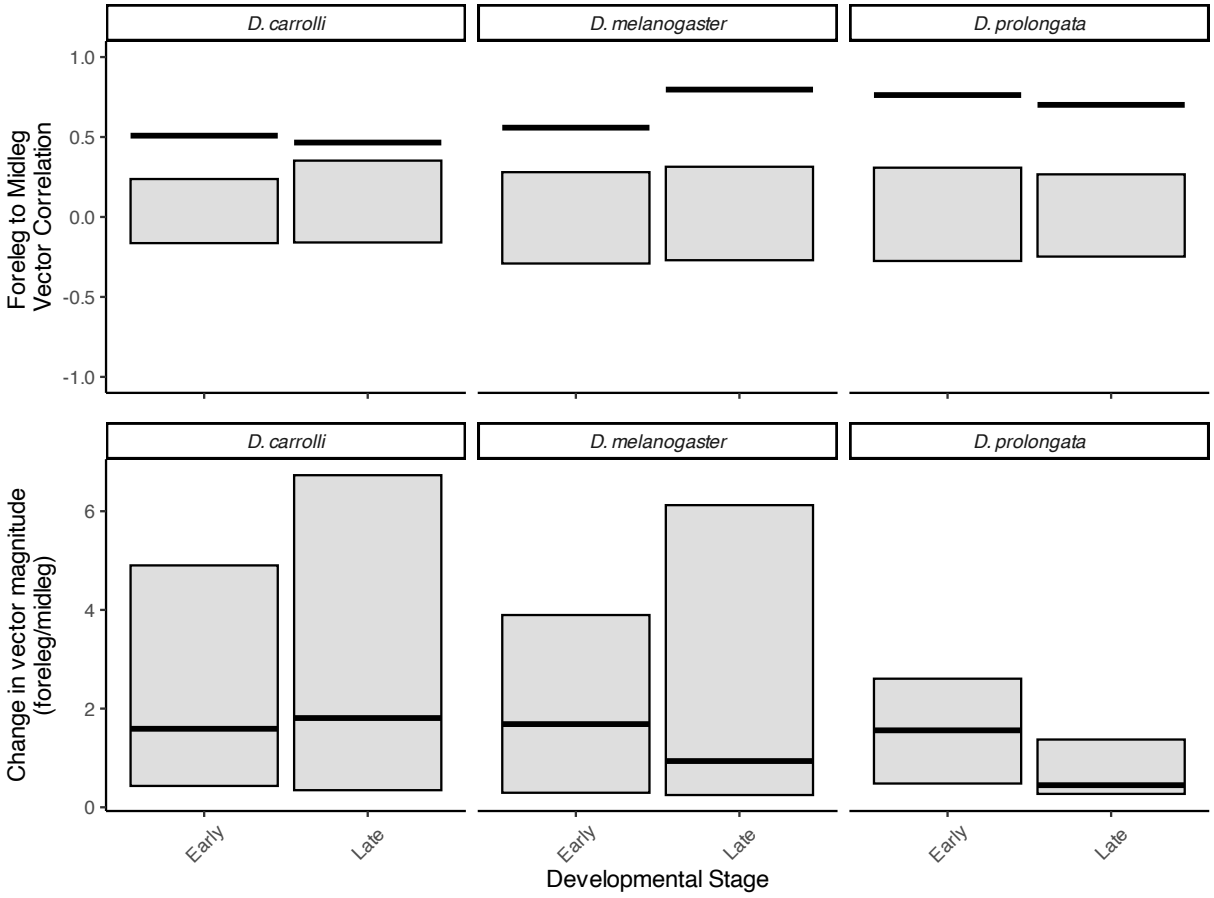

Figure S18: Vector direction and magnitude of WNT signalling genes within species between foreleg and midleg. Based on genes that overlap in all species. Top row shows direction of SBGE in foreleg compared to midleg within all three species. Bottom row is the magnitude of SBGE in foreleg relative to midleg within each species.

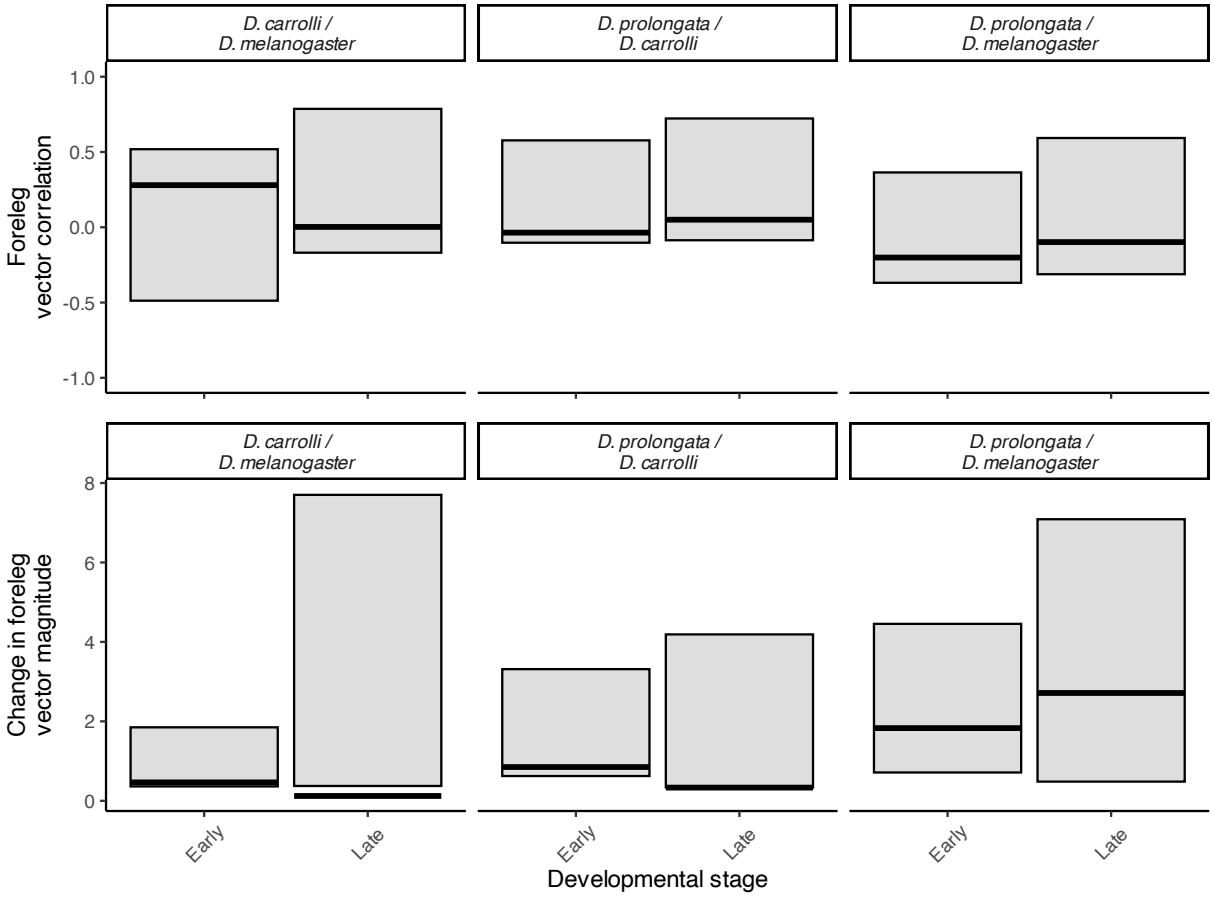

Figure S19: Direction and magnitude Notch signalling genes between species in the foreleg. Top row shows direction of SBGE in foreleg compared between *D. prolongata* and the other species, with the last column being the comparison of, the first column shows *D. carrolli* compared to *D. melanogaster* as a point of reference. Bottom row is the magnitude of SBGE in foreleg between each species.

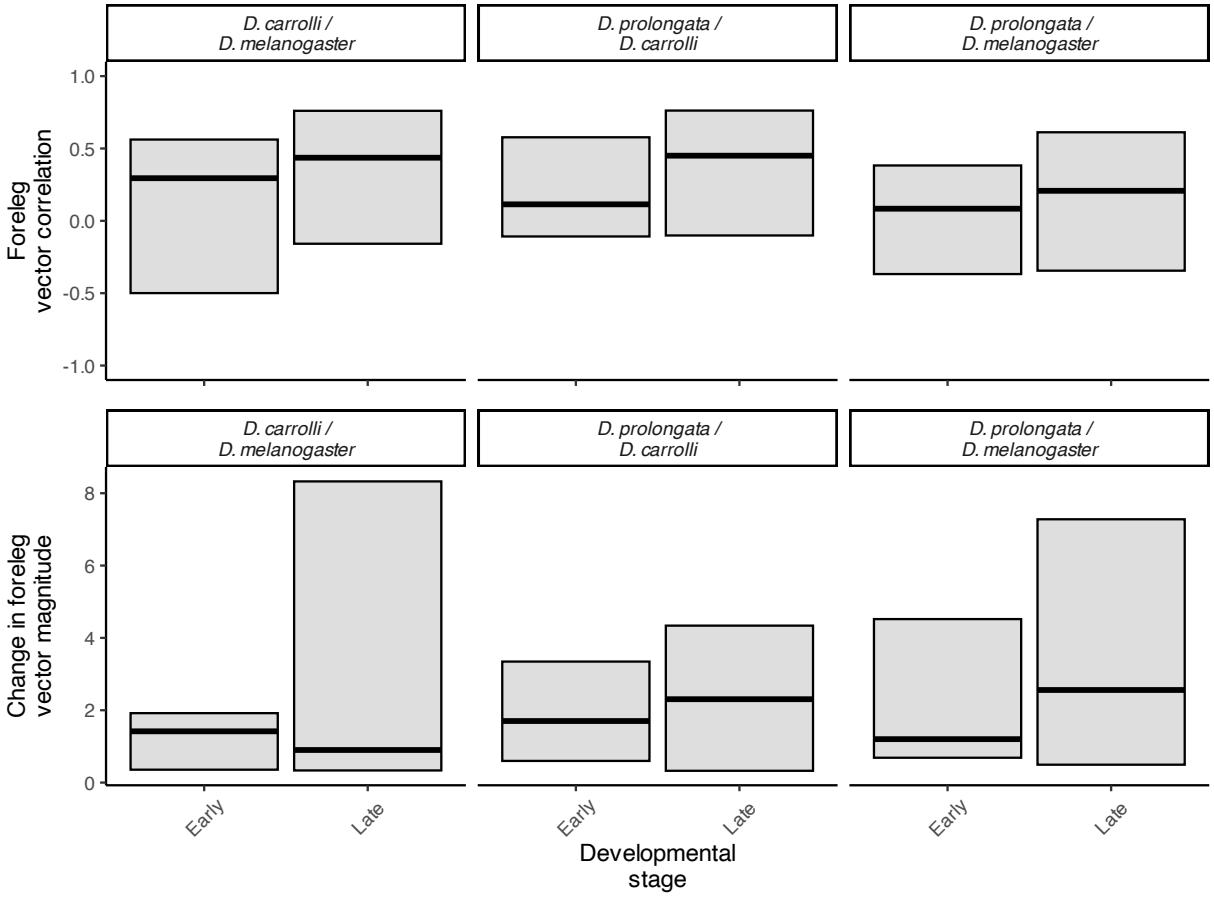

Figure S20: Direction and magnitude of EGFR signalling genes between species in the foreleg. Top row shows direction of SBGE in foreleg compared between *D. prolongata* and the other species, with the last column being the comparison of, the first column shows *D. carrolli* compared to *D. melanogaster* as a point of reference. Bottom row is the magnitude of SBGE in foreleg between each species.

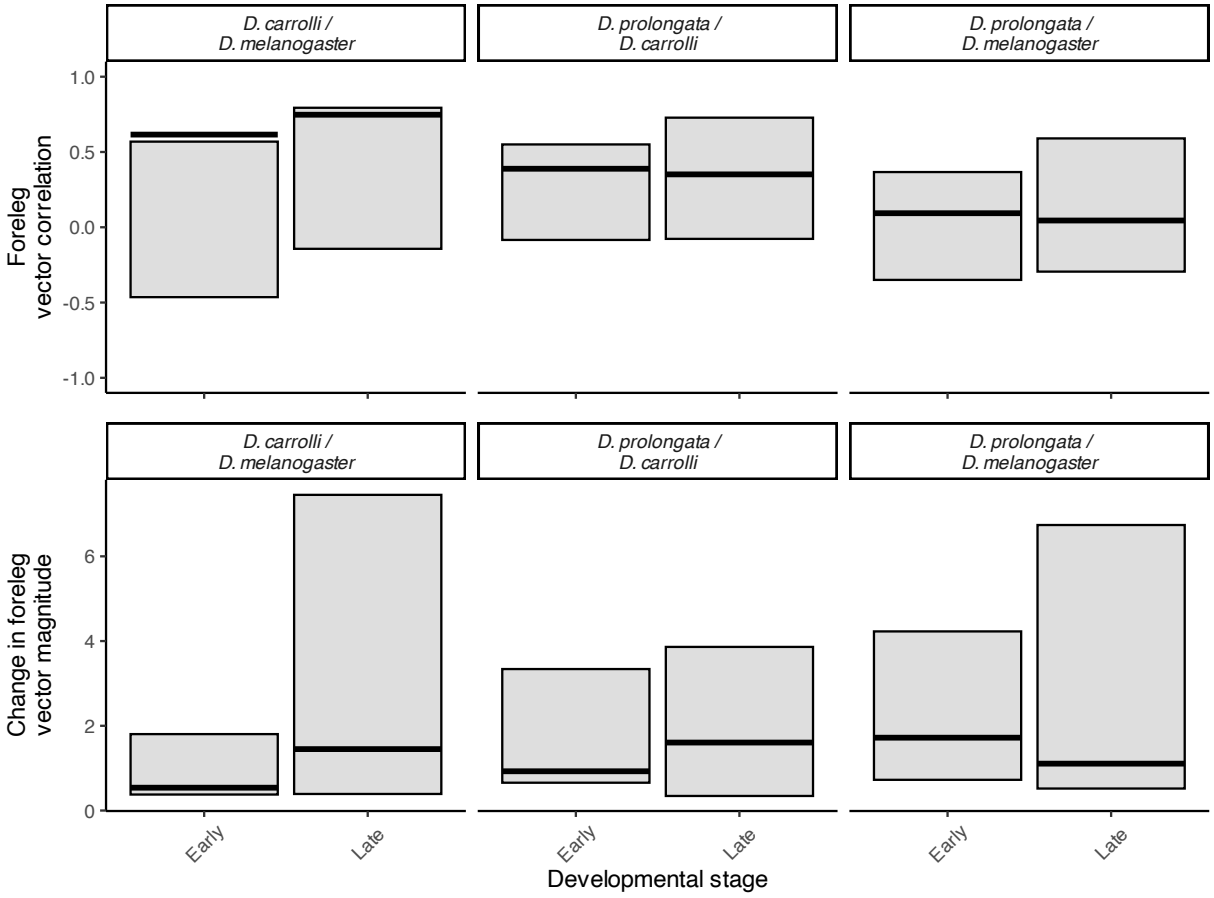

Figure S21: Direction and magnitude of Hippo signalling genes between species in the foreleg. Top row shows direction of SBGE in foreleg compared between *D. prolongata* and the other species, with the last column being the comparison of, the first column shows *D. carrolli* compared to *D. melanogaster* as a point of reference. Bottom row is the magnitude of SBGE in foreleg between each species.

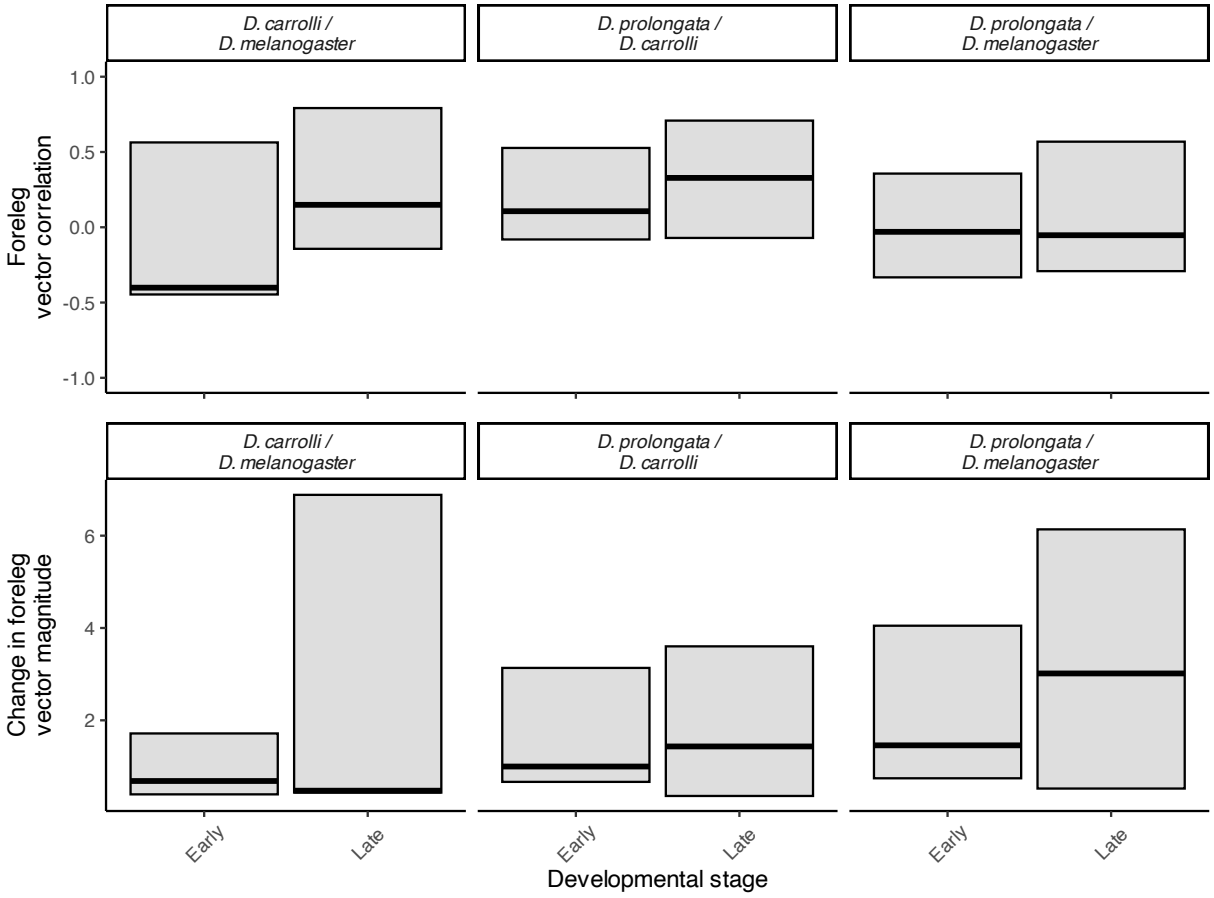

Figure S22: Direction and magnitude of WNT signalling genes between species in the foreleg. Top row shows direction of SBGE in foreleg compared between *D. prolongata* and the other species, with the last column being the comparison of, the first column shows *D. carrolli* compared to *D. melanogaster* as a point of reference. Bottom row is the magnitude of SBGE in foreleg between each species.

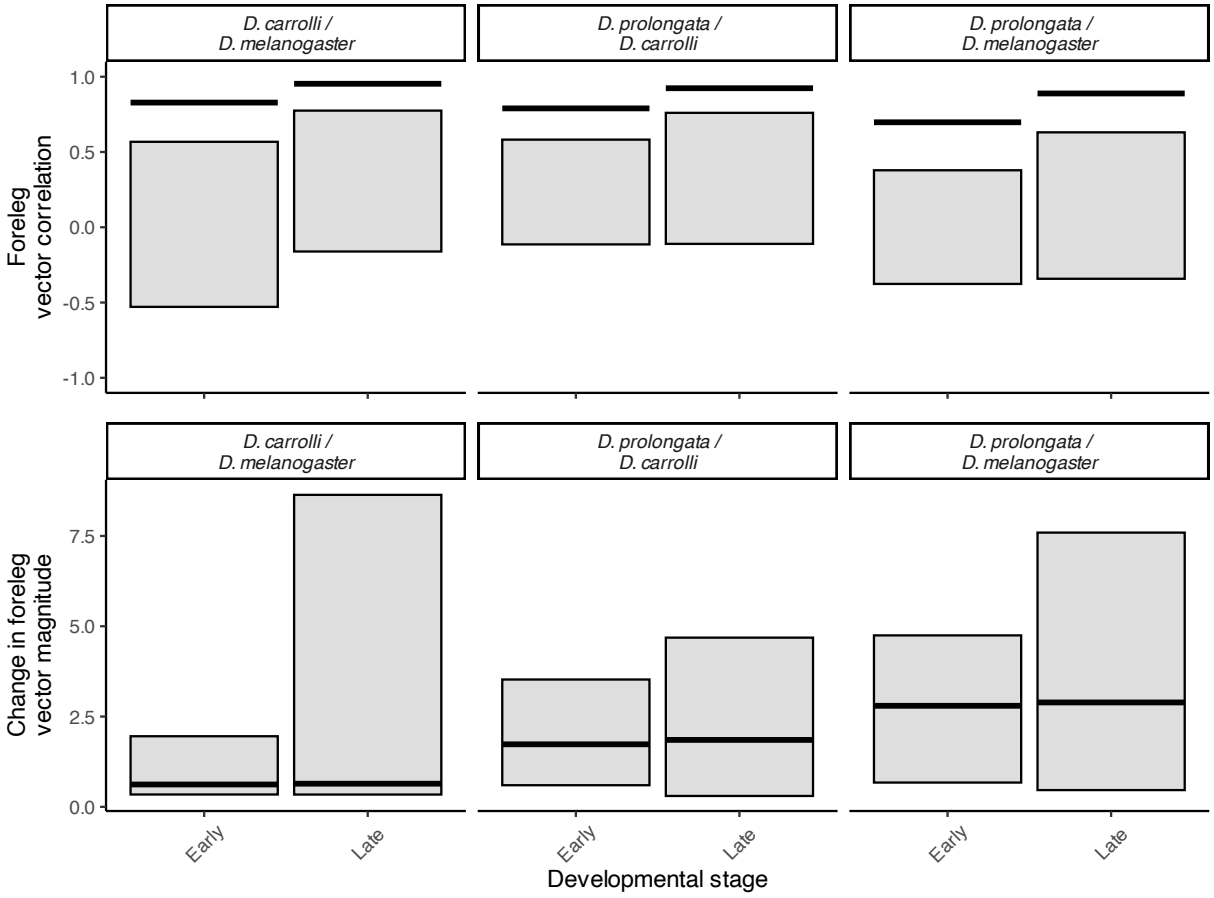

Figure S23: Direction and magnitude of BMP signalling genes between species in the foreleg. Top row shows direction of SBGE in foreleg compared between *D. prolongata* and the other species, with the last column being the comparison of, the first column shows *D. carrolli* compared to *D. melanogaster* as a point of reference. Bottom row is the magnitude of SBGE in foreleg between each species.

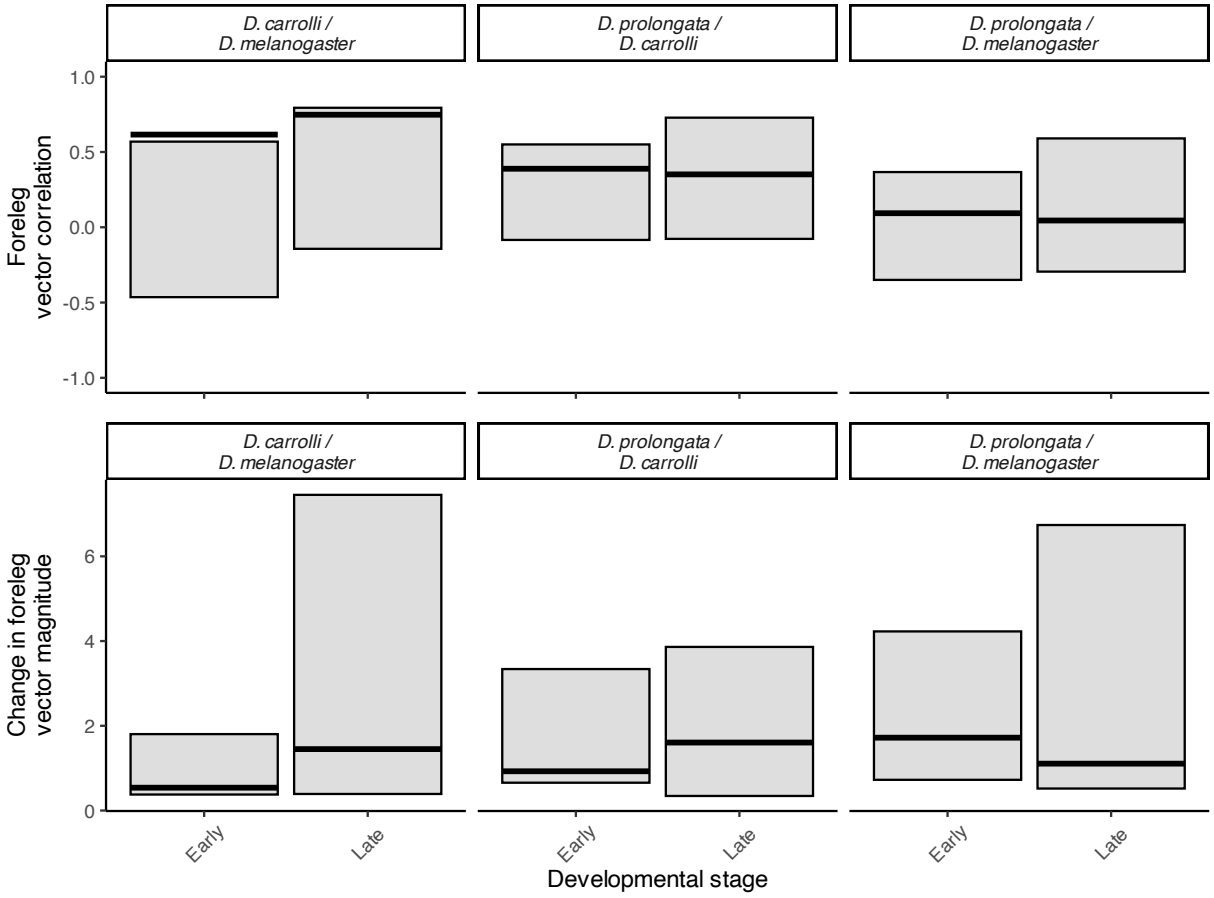

Figure S24: Direction and magnitude of Hedgehog signalling genes between species in the foreleg. Top row shows direction of SBGE in foreleg compared between *D. prolongata* and the other species, with the last column being the comparison of, the first column shows *D. carrolli* compared to *D. melanogaster* as a point of reference. Bottom row is the magnitude of SBGE in foreleg between each species.

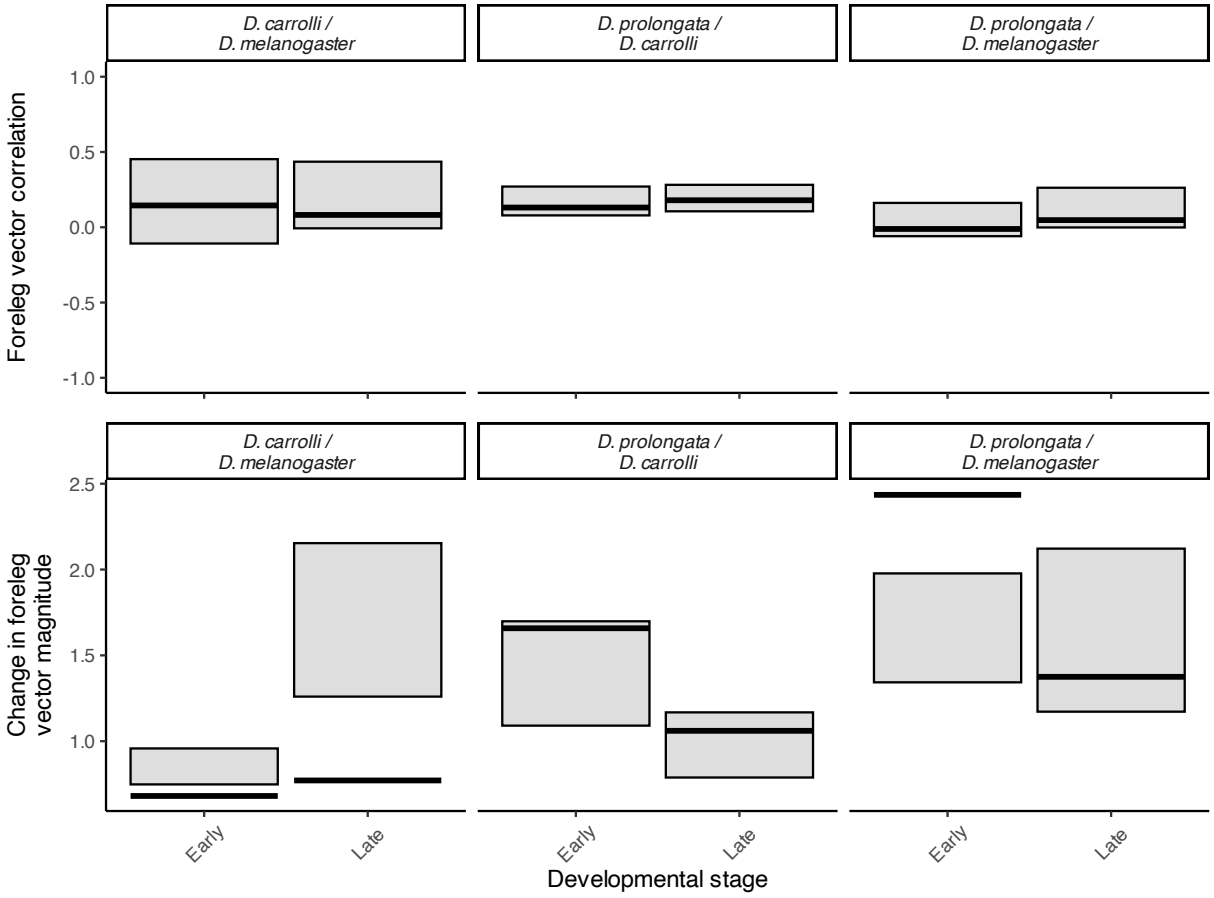

Figure S25: Direction and magnitude of putative *grn* targets between species in the foreleg. Top row shows direction of SBGE in foreleg compared between *D. prolongata* and the other species, the first column shows *D. carrolli* compared to *D. melanogaster* as a point of reference. Bottom row is the magnitude of SBGE in foreleg between each species.

Table S1: Summary of software used, versions, and any non-default flags or parameters

software	version	flags
BBduk	39.06	threads=32 ftr=100 ktrim=r k=21 mink=10 hdist=1 tpe tbo qtrim=rl trimq=15 minlength=36 rcomp=t
agat	0.9.2	convert_sp_gff2gtf.pl
STAR	2.7.11b	--quantMode TranscriptomeSAM GeneCounts --outSAMtype BAM SortedByCoordinate --sjdbFileChrStartEnd
glmmTMB	1.1.10	the_counts ~ (sex+leg+stage)^3 + diag(1 | stage:replicate) offset = normFactors family = nbinom2()
emmeans	1.10.5	
DESeq2	1.46.0	lfcShrink(res, type = “ashr”)
blast+	2.14.1	-outfmt "6 stitle sstart send sseqid" -sorthits 4 -max_target_seqs 1
bedtools getfasta	2.31	
MUSCLE	3.8.1551	
Tandem Repeat Finder	4.09.1	2 7 7 80 10 50 500 -m -h
SCRMshaw	1.1	--imm
tidyverse	2.0.0	
ashr	2.2-63	
data.table	1.16.2	

Table S2: multivariate model for all traits contrasted between UAS-RNAi genes and Pen^{NP3666}-GAL4

trait	temp	sex	code_trt.vs.ctrl1	estimate	SE	df	t.ratio	p.value
ForFemurLength	25	F	bab1 - control	-0.05	0.01	5803	-3.79	3.60e-04
ForFemurLength	25	F	CG13285 - control	-5.05e-03	0.01	5803	-0.43	0.78
ForFemurLength	25	F	CG30457 - control	-4.88e-03	0.01	5803	-0.42	0.78
ForFemurLength	25	F	dysf - control	0.05	0.01	5803	4.73	8.07e-06
ForFemurLength	25	F	grn - control	-0.02	0.01	5803	-1.36	0.30
ForFemurLength	25	F	otp - control	-4.88e-04	0.01	5803	-0.04	0.97
ForFemurLength	25	F	Sox15 - control	-0.14	0.02	5803	-8.84	8.75e-18
ForFemurLength	28	F	bab1 - control	-0.11	0.01	5803	-8.21	9.82e-16
ForFemurLength	28	F	CG13285 - control	-8.57e-03	0.01	5803	-0.74	0.49
ForFemurLength	28	F	CG30457 - control	-7.97e-03	0.01	5803	-0.69	0.49
ForFemurLength	28	F	dysf - control	-0.03	0.01	5803	-2.53	0.02
ForFemurLength	28	F	grn - control	-0.07	0.01	5803	-6.21	1.32e-09
ForFemurLength	28	F	otp - control	-0.03	0.01	5803	-3.25	2.02e-03
ForFemurLength	28	F	Sox15 - control	-0.16	0.02	5803	-8.53	1.31e-16
ForFemurLength	25	M	bab1 - control	-0.06	0.03	5803	-2.25	0.07
ForFemurLength	25	M	CG13285 - control	-8.19e-03	0.01	5803	-0.69	0.57
ForFemurLength	25	M	CG30457 - control	-0.01	0.01	5803	-0.94	0.49
ForFemurLength	25	M	dysf - control	-3.02e-03	0.01	5803	-0.26	0.79
ForFemurLength	25	M	grn - control	-0.03	0.01	5803	-2.17	0.07
ForFemurLength	25	M	otp - control	-0.02	0.01	5803	-1.57	0.21
ForFemurLength	25	M	Sox15 - control	-0.16	0.01	5803	-11.17	7.45e-28
ForFemurLength	28	M	bab1 - control	-0.07	0.02	5803	-3.22	3.01e-03
ForFemurLength	28	M	CG13285 - control	-0.01	0.01	5803	-0.89	0.44
ForFemurLength	28	M	CG30457 - control	-2.71e-03	0.01	5803	-0.23	0.82
ForFemurLength	28	M	dysf - control	-0.04	0.01	5803	-3.11	3.33e-03
ForFemurLength	28	M	grn - control	-0.07	0.01	5803	-6.20	4.31e-09
ForFemurLength	28	M	otp - control	-0.02	0.01	5803	-1.42	0.22
ForFemurLength	28	M	Sox15 - control	-0.14	0.03	5803	-5.05	1.56e-06
ForFemurWidth	25	F	bab1 - control	0.09	0.03	5803	2.80	0.02
ForFemurWidth	25	F	CG13285 - control	-0.05	0.03	5803	-1.52	0.18
ForFemurWidth	25	F	CG30457 - control	-0.02	0.03	5803	-0.53	0.62
ForFemurWidth	25	F	dysf - control	-0.06	0.03	5803	-1.90	0.10
ForFemurWidth	25	F	grn - control	-0.02	0.03	5803	-0.50	0.62
ForFemurWidth	25	F	otp - control	0.07	0.03	5803	2.36	0.04
ForFemurWidth	25	F	Sox15 - control	0.13	0.04	5803	3.07	0.01
ForFemurWidth	28	F	bab1 - control	0.13	0.03	5803	3.65	9.10e-04
ForFemurWidth	28	F	CG13285 - control	-0.10	0.03	5803	-3.13	4.08e-03
ForFemurWidth	28	F	CG30457 - control	0.01	0.03	5803	0.37	0.71
ForFemurWidth	28	F	dysf - control	0.07	0.03	5803	2.29	0.04
ForFemurWidth	28	F	grn - control	-0.05	0.03	5803	-1.63	0.12
ForFemurWidth	28	F	otp - control	-0.05	0.03	5803	-2.04	0.06
ForFemurWidth	28	F	Sox15 - control	0.27	0.05	5803	5.07	2.94e-06
ForFemurWidth	25	M	bab1 - control	0.15	0.08	5803	1.94	0.09
ForFemurWidth	25	M	CG13285 - control	0.04	0.03	5803	1.17	0.28
ForFemurWidth	25	M	CG30457 - control	0.04	0.03	5803	1.39	0.23
ForFemurWidth	25	M	dysf - control	-0.01	0.03	5803	-0.33	0.74
ForFemurWidth	25	M	grn - control	0.10	0.03	5803	3.28	3.63e-03

ForFemurWidth	25	M	otp - control	0.07	0.03	5803	2.35	0.04
ForFemurWidth	25	M	Sox15 - control	0.17	0.04	5803	4.00	4.55e-04
ForFemurWidth	28	M	bab1 - control	-1.96e-03	0.06	5803	-0.04	0.97
ForFemurWidth	28	M	CG13285 - control	-0.03	0.03	5803	-0.85	0.56
ForFemurWidth	28	M	CG30457 - control	0.05	0.03	5803	1.72	0.20
ForFemurWidth	28	M	dysf - control	-0.09	0.03	5803	-2.81	0.03
ForFemurWidth	28	M	grn - control	-0.04	0.03	5803	-1.31	0.33
ForFemurWidth	28	M	otp - control	1.46e-03	0.03	5803	0.05	0.97
ForFemurWidth	28	M	Sox15 - control	0.18	0.08	5803	2.36	0.06
MidFemurLength	25	F	bab1 - control	-0.09	0.01	5803	-6.84	3.13e-11
MidFemurLength	25	F	CG13285 - control	-0.02	0.01	5803	-1.79	0.10
MidFemurLength	25	F	CG30457 - control	-0.04	0.01	5803	-2.76	0.01
MidFemurLength	25	F	dysf - control	0.02	0.01	5803	1.39	0.19
MidFemurLength	25	F	grn - control	-0.03	0.01	5803	-2.20	0.05
MidFemurLength	25	F	otp - control	-8.63e-03	0.01	5803	-0.68	0.50
MidFemurLength	25	F	Sox15 - control	-0.13	0.02	5803	-7.43	9.01e-13
MidFemurLength	28	F	bab1 - control	-0.11	0.01	5803	-7.63	1.96e-13
MidFemurLength	28	F	CG13285 - control	1.03e-03	0.01	5803	0.08	0.94
MidFemurLength	28	F	CG30457 - control	-3.42e-03	0.01	5803	-0.27	0.92
MidFemurLength	28	F	dysf - control	-0.02	0.01	5803	-1.87	0.09
MidFemurLength	28	F	grn - control	-0.09	0.01	5803	-6.87	2.55e-11
MidFemurLength	28	F	otp - control	-0.02	0.01	5803	-2.23	0.05
MidFemurLength	28	F	Sox15 - control	-0.11	0.02	5803	-4.90	2.34e-06
MidFemurLength	25	M	bab1 - control	-0.09	0.03	5803	-2.80	0.02
MidFemurLength	25	M	CG13285 - control	8.07e-03	0.01	5803	0.62	0.53
MidFemurLength	25	M	CG30457 - control	0.03	0.01	5803	2.14	0.08
MidFemurLength	25	M	dysf - control	0.02	0.01	5803	1.54	0.17
MidFemurLength	25	M	grn - control	8.29e-03	0.01	5803	0.65	0.53
MidFemurLength	25	M	otp - control	0.02	0.01	5803	1.59	0.17
MidFemurLength	25	M	Sox15 - control	-0.20	0.02	5803	-11.00	5.15e-27
MidFemurLength	28	M	bab1 - control	-0.13	0.02	5803	-5.59	8.11e-08
MidFemurLength	28	M	CG13285 - control	-0.02	0.01	5803	-1.36	0.18
MidFemurLength	28	M	CG30457 - control	-0.03	0.01	5803	-2.44	0.02
MidFemurLength	28	M	dysf - control	-0.04	0.01	5803	-3.15	2.91e-03
MidFemurLength	28	M	grn - control	-0.07	0.01	5803	-5.22	4.34e-07
MidFemurLength	28	M	otp - control	-0.03	0.01	5803	-2.55	0.01
MidFemurLength	28	M	Sox15 - control	-0.24	0.03	5803	-7.64	1.74e-13
MidFemurWidth	25	F	bab1 - control	0.03	0.03	5803	0.86	0.39
MidFemurWidth	25	F	CG13285 - control	-0.04	0.03	5803	-1.40	0.19
MidFemurWidth	25	F	CG30457 - control	-0.04	0.03	5803	-1.39	0.19
MidFemurWidth	25	F	dysf - control	-0.07	0.03	5803	-2.39	0.06
MidFemurWidth	25	F	grn - control	-0.06	0.03	5803	-2.12	0.08
MidFemurWidth	25	F	otp - control	-0.04	0.03	5803	-1.47	0.19
MidFemurWidth	25	F	Sox15 - control	0.14	0.04	5803	3.57	2.54e-03
MidFemurWidth	28	F	bab1 - control	0.12	0.03	5803	3.68	1.65e-03
MidFemurWidth	28	F	CG13285 - control	-0.03	0.03	5803	-0.98	0.58
MidFemurWidth	28	F	CG30457 - control	2.24e-03	0.03	5803	0.07	0.94
MidFemurWidth	28	F	dysf - control	-0.02	0.03	5803	-0.53	0.70
MidFemurWidth	28	F	grn - control	-0.03	0.03	5803	-0.97	0.58
MidFemurWidth	28	F	otp - control	0.02	0.03	5803	0.71	0.67

MidFemurWidth	28	F	Sox15 - control	0.20	0.06	5803	3.50	1.65e-03
MidFemurWidth	25	M	bab1 - control	0.12	0.07	5803	1.59	0.39
MidFemurWidth	25	M	CG13285 - control	-0.04	0.03	5803	-1.19	0.41
MidFemurWidth	25	M	CG30457 - control	-0.04	0.03	5803	-1.25	0.41
MidFemurWidth	25	M	dysf - control	-0.01	0.03	5803	-0.42	0.81
MidFemurWidth	25	M	grn - control	-0.01	0.03	5803	-0.39	0.81
MidFemurWidth	25	M	otp - control	8.61e-04	0.03	5803	0.03	0.98
MidFemurWidth	25	M	Sox15 - control	0.11	0.04	5803	2.62	0.06
MidFemurWidth	28	M	bab1 - control	0.03	0.05	5803	0.52	0.70
MidFemurWidth	28	M	CG13285 - control	-0.04	0.03	5803	-1.26	0.49
MidFemurWidth	28	M	CG30457 - control	0.11	0.03	5803	3.48	3.48e-03
MidFemurWidth	28	M	dysf - control	0.03	0.03	5803	1.03	0.53
MidFemurWidth	28	M	grn - control	-9.90e-03	0.03	5803	-0.31	0.75
MidFemurWidth	28	M	otp - control	-0.02	0.03	5803	-0.58	0.70
MidFemurWidth	28	M	Sox15 - control	0.18	0.09	5803	2.10	0.13
Thorax	25	F	bab1 - control	-0.03	0.01	5803	-2.44	0.03
Thorax	25	F	CG13285 - control	9.17e-04	0.01	5803	0.07	0.97
Thorax	25	F	CG30457 - control	5.55e-03	0.01	5803	0.42	0.94
Thorax	25	F	dysf - control	-5.59e-04	0.01	5803	-0.04	0.97
Thorax	25	F	grn - control	0.03	0.01	5803	2.63	0.03
Thorax	25	F	otp - control	0.03	0.01	5803	2.59	0.03
Thorax	25	F	Sox15 - control	-0.03	0.02	5803	-1.58	0.20
Thorax	28	F	bab1 - control	-0.08	0.01	5803	-5.44	3.94e-07
Thorax	28	F	CG13285 - control	-0.04	0.01	5803	-2.76	0.01
Thorax	28	F	CG30457 - control	-0.01	0.01	5803	-0.84	0.56
Thorax	28	F	dysf - control	-0.06	0.01	5803	-4.47	2.74e-05
Thorax	28	F	grn - control	-4.58e-03	0.01	5803	-0.34	0.73
Thorax	28	F	otp - control	-8.72e-03	0.01	5803	-0.69	0.57
Thorax	28	F	Sox15 - control	-0.09	0.02	5803	-3.91	2.19e-04
Thorax	25	M	bab1 - control	-0.09	0.03	5803	-2.71	0.02
Thorax	25	M	CG13285 - control	0.01	0.01	5803	0.91	0.42
Thorax	25	M	CG30457 - control	0.01	0.01	5803	0.97	0.42
Thorax	25	M	dysf - control	-0.03	0.01	5803	-2.62	0.02
Thorax	25	M	grn - control	2.70e-03	0.01	5803	0.20	0.84
Thorax	25	M	otp - control	-0.02	0.01	5803	-1.42	0.27
Thorax	25	M	Sox15 - control	-0.10	0.02	5803	-5.76	6.27e-08
Thorax	28	M	bab1 - control	-0.14	0.02	5803	-6.09	8.20e-09
Thorax	28	M	CG13285 - control	-0.02	0.01	5803	-1.54	0.17
Thorax	28	M	CG30457 - control	-0.01	0.01	5803	-0.89	0.43
Thorax	28	M	dysf - control	-0.05	0.01	5803	-3.49	1.69e-03
Thorax	28	M	grn - control	-0.03	0.01	5803	-2.09	0.06
Thorax	28	M	otp - control	-0.05	0.01	5803	-3.28	2.43e-03
Thorax	28	M	Sox15 - control	-0.03	0.03	5803	-0.79	0.43

Table S3: FFL

contrast	temp	sex	estimate	SE	df	Lower.CL	upper.CL
bab1 - control	25	F	-0.0375389876	0.01613063	1083	-0.08000523	0.004927259
CG13285 - control	25	F	0.0011269694	0.01624377	1083	-0.04163713	0.043891070
CG30457 - control	25	F	-0.0078946882	0.01574506	1083	-0.04934585	0.033556471
dysf - control	25	F	0.0547771303	0.01564277	1083	0.01359527	0.095958990
grn - control	25	F	-0.0234483297	0.01566639	1083	-0.06469239	0.017795735
otp - control	25	F	-0.0084377792	0.01566576	1083	-0.04968016	0.032804605
Sox15 - control	25	F	-0.1288173762	0.01960156	1083	-0.18042134	-0.077213412
bab1 - control	28	F	-0.0917693423	0.01713551	1083	-0.13688106	-0.046657628
CG13285 - control	28	F	0.0007136154	0.01582474	1083	-0.04094733	0.042374557
CG30457 - control	28	F	-0.0055084783	0.01577260	1083	-0.04703214	0.036015184
dysf - control	28	F	-0.0162827001	0.01584253	1083	-0.05799046	0.025425062
grn - control	28	F	-0.0714064010	0.01577033	1083	-0.11292409	-0.029888715
otp - control	28	F	-0.0363353894	0.01577209	1083	-0.07785770	0.005186923
Sox15 - control	28	F	-0.1426575817	0.02167124	1083	-0.19971029	-0.085604878
bab1 - control	25	M	-0.0408074119	0.02918856	1083	-0.11765055	0.036035728
CG13285 - control	25	M	-0.0087280910	0.01625184	1083	-0.05151343	0.034057252
CG30457 - control	25	M	-0.0114694707	0.01565165	1083	-0.05267472	0.029735778
dysf - control	25	M	0.0066387459	0.01567410	1083	-0.03462560	0.047903094
grn - control	25	M	-0.0234330708	0.01564839	1083	-0.06462973	0.017763588
otp - control	25	M	-0.0120650896	0.01565619	1083	-0.05328228	0.029152102
Sox15 - control	25	M	-0.1204625608	0.01886189	1083	-0.17011924	-0.070805884
bab1 - control	28	M	-0.0332852146	0.02272880	1083	-0.09312210	0.026551670
CG13285 - control	28	M	-0.0039771831	0.01579660	1083	-0.04556403	0.037609667
CG30457 - control	28	M	-0.0003061673	0.01586645	1083	-0.04207691	0.041464578
dysf - control	28	M	-0.0243400768	0.01590747	1083	-0.06621881	0.017538660
grn - control	28	M	-0.0664262469	0.01587354	1083	-0.10821567	-0.024636828
otp - control	28	M	-0.0047311627	0.01651417	1083	-0.04820711	0.038744784
Sox15 - control	28	M	-0.1391134595	0.02916397	1083	-0.21589186	-0.062335061

Table S4: FFW

contrast	temp	sex	estimate	SE	df	Lower.CL	upper.CL
bab1 - control	25	F	0.094611369	0.03588458	1074	0.0001386695	0.189084068
CG13285 - control	25	F	-0.049012888	0.03624164	1074	-0.1444255952	0.046399820
CG30457 - control	25	F	-0.020289078	0.03463077	1074	-0.1114608905	0.070882734
dysf - control	25	F	-0.061504692	0.03428173	1074	-0.1517575890	0.028748205
grn - control	25	F	-0.023262941	0.03470732	1074	-0.1146362762	0.068110395
otp - control	25	F	0.066107797	0.03435827	1074	-0.0243466072	0.156562200
Sox15 - control	25	F	0.177377395	0.04708123	1074	0.0534274748	0.301327316
bab1 - control	28	F	0.148925771	0.03861650	1074	0.0472608125	0.250590729
CG13285 - control	28	F	-0.088202287	0.03480112	1074	-0.1798225607	0.003417988
CG30457 - control	28	F	0.017442770	0.03472036	1074	-0.0739648918	0.108850432
dysf - control	28	F	0.084116145	0.03495065	1074	-0.0078978100	0.176130100
grn - control	28	F	-0.045919208	0.03471467	1074	-0.1373118884	0.045473471
otp - control	28	F	-0.042676080	0.03471697	1074	-0.1340748238	0.048722664
Sox15 - control	28	F	0.282723118	0.05341550	1074	0.1420970846	0.423349151
bab1 - control	25	M	0.157870801	0.07554723	1074	-0.0410210413	0.356762644
CG13285 - control	25	M	0.019276816	0.03627292	1074	-0.0762182428	0.114771874
CG30457 - control	25	M	0.037675970	0.03431590	1074	-0.0526668937	0.128018834
dysf - control	25	M	-0.009630205	0.03473338	1074	-0.1010721412	0.081811732
grn - control	25	M	0.098036301	0.03430513	1074	0.0077218148	0.188350787
otp - control	25	M	0.071957663	0.03433067	1074	-0.0184240648	0.162339390
Sox15 - control	25	M	0.188016790	0.04601418	1074	0.0668760753	0.309157505
bab1 - control	28	M	0.021891240	0.05664375	1074	-0.1272337466	0.171016227
CG13285 - control	28	M	-0.018731631	0.03545567	1074	-0.1120751388	0.074611877
CG30457 - control	28	M	0.059363936	0.03472619	1074	-0.0320590852	0.150786957
dysf - control	28	M	-0.086570653	0.03522299	1074	-0.1793015733	0.006160266
grn - control	28	M	-0.033613239	0.03510253	1074	-0.1260270511	0.058800573
otp - control	28	M	0.005257614	0.03774336	1074	-0.0941086575	0.104623886
Sox15 - control	28	M	0.190262820	0.07550569	1074	-0.0085196511	0.389045290