Cis-regulatory Variation in Relation to Sex and Sexual Dimorphism in *Drosophila melanogaster*

Prashastha Mishra [1], Tania S. Barrera [1], Karl Grieshop [1,2,3], Aneil F. Agrawal [1,*]

[1]Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, ON, Canada M5S 3B2
[2]Department of Molecular Biosciences, The Wenner-Gren Institute, Stockholm University, Stockholm SE-10691, Sweden
[3]School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK

*Corresponding author: E-mail: a.agrawal@utoronto.ca.

Accepted: October 11, 2024

Abstract

Much of sexual dimorphism is likely due to sex-biased gene expression, which results from differential regulation of a genome that is largely shared between males and females. Here, we use allele-specific expression to explore *cis*-regulatory variation in *Drosophila melanogaster* in relation to sex. We develop a Bayesian framework to infer the transcriptome-wide joint distribution of *cis*-regulatory effects across the sexes. We also examine patterns of *cis*-regulatory variation with respect to two other levels of variation in sexual dimorphism: (i) across genes that vary in their degree of sex-biased expression and (ii) among tissues that vary in their degree of dimorphism (e.g. relatively low dimorphism in heads vs. high dimorphism in gonads). We uncover evidence of widespread *cis*-regulatory variation in all tissues examined, with female-biased genes being especially enriched for this variation. A sizeable proportion of *cis*-regulatory variation is inferred to have sex-specific effects, with sex-dependent *cis* effects being much more frequent in gonads than in heads. Finally, we find some genes where 1 allele contributes to more than 50% of a gene's expression in heterozygous males but <50% of its expression in heterozygous females. Such variants could provide a mechanism for sex-specific dominance reversals, a phenomenon important for sexually antagonistic balancing selection. However, tissue differences in allelic imbalance are approximately as frequent as sex differences, perhaps suggesting that sexual conflict may not be particularly unique in shaping patterns of expression variation.

Key words: sexual dimorphism, sex-biased gene expression, allelic imbalance.

Significance

In heterozygotes, some alleles are expressed at higher levels than others, a phenomenon known as allelic imbalance (AI). Such an imbalance must be due, at least in part, to *cis*-regulatory differences between the two alleles at a locus because both alleles experience the same *trans*-regulatory environment. This study examines AI in *Drosophila melanogaster* with the aim to quantify how frequently AI occurs and the strength of these imbalances across the transcriptome. We are particularly interested in whether AI effects tend to be similar or different between the sexes, finding that they frequently differ between the sexes in gonads but much less so in heads. These results suggest that sexes can more readily evolve independently with respect to gene expression in gonads than in heads.

Introduction

Sexual dimorphism is a universal feature of sexually reproducing species, with males and females differing in their appearance, physiology, life history, and behavior (Fairbairn et al. 2007). Males and females carry similar genetic material (other than differentiated sex chromosomes in some species), and these extensive differences in

phenotype are largely the result of differential expression of a shared genome (Parsch and Ellegren 2013).

Transcriptional gene regulation is central to the evolution of sex-biased gene expression. Regulation of gene expression can occur through *cis*-acting regulatory elements linked to their target alleles or through *trans*-acting factors. *Cis*-regulatory variants have been reported to be enriched between species relative to their frequency within species in *Drosophila* (Wittkopp et al. 2004, 2008), yeast (Metzger et al. 2017), and mice (Goncalves et al. 2012), among others. This suggests that *cis*-regulatory factors are preferentially used to respond to natural selection for divergent expression. Additionally, *cis*-acting variants are more likely to have additive effects on expression than *trans*-acting variants (Wray 2007; Lemos et al. 2008), which may influence their ability to respond to selection. In contrast, *trans* factors affect expression in multiple genes, and mutations in *trans* may be constrained by pleiotropic side effects (Stern 2000). Last, at a practical level, *cis* variants are easier to study. Here, we focus entirely on regulatory variation in *cis.*

Studies indicate that *cis*-regulatory variation is highly pervasive in a wide range of taxa, including humans (Pastinen and Hudson 2004; Stranger et al. 2012), yeast (Kita et al. 2017), mice (Campbell et al. 2008), and *Arabidopsis* (Cubillos et al. 2014). As in other species, *cis*-regulatory variation is common within *Drosophila melanogaster*, with a higher abundance of variation being detected in *cis* than in *trans* (Genissel et al. 2007; Gruber and Long 2009; Osada et al. 2017). Here, we examine *cis*-regulatory variation in *D. melanogaster* by measuring differential expression between alleles, termed "allelic imbalance" (henceforth, AI). We investigate its relationship with sexual dimorphism in three ways.

First, we examine whether *cis*-regulatory variation is more common in genes with dimorphic expression (i.e. sex-biased genes). *Cis*-regulatory variation is expected to be enriched in loci undergoing sexual conflict over gene expression. Under some conditions, sexual conflict can generate balancing selection that stably maintains polymorphisms (Kidwell et al. 1977; Connallon and Clark 2014). If sex-biased genes are more likely to experience sexual conflict than unbiased genes, then one might predict *cis*-regulatory variants to be more common in sex-biased genes. However, even though sexually divergent selection in the past may be responsible for creating the dimorphic expression observed in the present, it is unclear to what extent sexual conflict may persist in genes with sex-biased expression. Sex-biased genes could represent instances of resolved or ongoing conflicts (Bonduriansky and Chenoweth 2009; Rowe et al. 2018). Similarly, unbiased genes could either lack sexual conflict entirely or simply lack the appropriate variation to evolve sex-biased expression to mitigate conflict (Parsch and Ellegren 2013). Based on population genetic analyses of human and fly data, Cheng and Kirkpatrick (2016) concluded that sexual conflict is most intense in moderately sex-biased genes; this may lead one to predict *cis*-regulatory variation being most common in genes with an intermediate sex bias. Two previous studies of *cis*-regulatory variation in *D. melanogaster* found different patterns from one another and neither matched this prediction (Osada et al. 2017; Puixeu et al. 2023).

Regulatory variants that affect expression similarly in both sexes would not allow for the resolution of sexual conflict via sex-biased gene expression. Thus, our second goal is to examine the extent to which *cis*-regulatory effects differ between the sexes, measured as "sex-dependent" AI (SD-AI). In terms of the transcriptome-wide joint distribution of *cis*-regulatory effects across the sexes, we can consider sex differences from two related perspectives: the frequency with which AI differs between sexes or the intersexual correlation in AI. The latter harkens to the intersexual genetic correlation in expression (often represented as r_{MF}), which governs the extent to which expression can evolve independently between the sexes.

The most extreme form of SD-AI occurs when there is a reversal of the dominant allele in the sexes (sex-reversed AI), whereby heterozygous males have >50% expression from one allele, while heterozygous females have >50% expression from the alternative allele. We perform additional analyses to find instances of sex-reversed AI. Cases of sex-reversed AI hold interest because they provide a potential mechanism for sex-specific dominance reversals in fitness, whereby heterozygotes of each sex could dominantly express the preferred allele for a locus under sexually antagonistic selection. Theoretical models indicate that sexually antagonistic variants exhibiting dominance reversals are preferentially maintained by balancing selection, due to partially mitigating sexual conflict (Kidwell et al. 1977; Spencer and Priest 2016; Connallon and Chenoweth 2019; Grieshop et al. 2024).

Our third goal was to examine *cis*-regulatory variation across tissues varying in their degree of sexual dimorphism. We studied expression data from whole-body, head, and gonad samples. Gonads and heads offer an interesting comparison because gonads are highly sexually dimorphic, whereas heads are much less so. We compare the frequency of AI and particularly SD-AI between these two tissues to discern whether patterns in AI reflect the underlying levels of sexual dimorphism.

Results

We crossed two inbred lines of *D. melanogaster* (DGRP-177 and SP159N) and measured allele-specific expression in heads, gonads, and whole bodies of F_1s of both sexes. The two lines were of different geographic origin to increase single nucleotide polymorphism (SNP) differences and thereby

increase our power to assign reads to haplotype origins. For whole-body samples, we had F_1s from both cross directions. A series of filters based on parental and F1 genomic data were applied to remove genes with possible mapping biases (see 'Methods' section).

Apparent Signals of Parental Effects

Studies suggest that parental effects on allelic expression rarely occur in *D. melanogaster* (Wittkopp et al. 2006; Coolon et al. 2012; Chen et al. 2015). Moreover, patterns consistent with parental effects may be due to residual mapping bias. We performed three tests using whole-body samples from the main and reciprocal crosses to identify genes with evidence of "parental effects," seeking to remove these genes from subsequent analysis. We analyzed the data from each sex separately and also performed an analysis using combined data from the sexes. A total of 346 of 7,716 (4.5%; X genes included) genes tested in females and 829 of 8,817 (9.4%; X and Y genes excluded) genes tested in males showed a significant effect of cross direction on AI ($P < 0.05$). (The substantially higher incidence of "parental effects" in F_1 males relative to females could be due to males of the two crosses having *trans*-regulatory differences attributable to complementary XY combinations.) In the combined data from both sexes, 532 of 6,371 (8.4%; X and Y genes excluded) genes had a significant effect of cross direction. A total of 10,522 genes were analyzed across these three tests, with 1,375 (13.1%) showing evidence of parental effects in at least one of the tests. While many of these genes are expected to be false positives (5% of all genes in each test), our primary goal here is to reduce the possibility of mapping bias influencing our analyses of AI. We therefore excluded all these genes showing significant parental effects from further analysis.

Sex-Dependent Effects on AI

AI appears to be pervasive in the *D. melanogaster* transcriptome, with more than 30% of the genes tested for each tissue showing significant AI (Table 1). Given our *P*-value threshold, we expect ~5% of genes to falsely appear as showing significant AI (i.e. false positives); the observed fraction is much greater than expected under the null hypothesis. We use a *P*-value criterion for the ease of comparisons among tissues, which would be complicated when applying a false discovery rate; false discovery rate *q*-values depend not only on the data from the focal genes but also on the distribution of *P*-values within a group, which could cause genes with similar evidence for AI to have different *q*-values in each group. Nevertheless, a version of Table 1 using false discovery rate (FDR) cutoffs is shown in supplementary table S8, Supplementary Material online. As expected, using the more conservative FDR requirement reduces the overall frequency of detection of AI, but the main patterns are unchanged. In each tissue, numerous genes exhibit significant SD-AI (Table 1). About 25% of genes exhibited significant SD-AI in gonads, the most sexually dimorphic tissue in our dataset. In contrast, only 10% of genes show SD-AI in heads.

Table 1 Frequency of AI and SD-AI

Tissue	Number of genes tested	Average number of informative reads[a]	% genes with AI	% genes with SD-AI
Gonads	2,547	3,857	31.6	25.3
Heads	4,062	9,019	45.0	10.1
Whole bodies (main cross)	3,627	6,668	36.5	19.6
Whole bodies (reciprocal cross)	3,901	8,953	35.7	18.3

[a]Summing reads for each gene that could be assigned to either parental genome across the six samples (three per sex) and then averaging across genes.

We considered the possibility that our inferences of AI or SD-AI may be a consequence of mapping bias, despite our efforts to minimize such errors. If mapping bias persists in the dataset despite the filters, we would expect patterns in genomic data to be reflected in RNA-seq. First, we would expect a positive correlation in fraction of DGRP-177 reads between F1 genomic and F1 transcriptomic datasets in both the sexes. However, the correlations are very close to 0 (supplementary figs. S2 and S3 and tables S4 and S5, Supplementary Material online), indicating that mapping bias is unlikely to be a widespread problem among genes inferred to have AI. Second, we would expect a significant positive correlation in the sex difference in the fraction of DGRP-177 reads between genomic and transcriptomic datasets. These genomic-transcriptomic correlations for sex differences are also close to 0, remaining so even as we consider only genes with significant SD-AI (supplementary fig. S4 and table S6, Supplementary Material online). Thus, our inferences of AI or SD-AI are unlikely to be substantially marred by mapping bias.

The gene *doublesex* (*dsx*) is a key terminal transcription factor in the sex determination hierarchy that is involved in the establishment of most phenotypic sex differences (Camara et al. 2008) and is known to cause sex differences in expression in many genes via multiple modes of regulation (Arbeitman et al. 2016). We tested whether genes with AI, or particularly with SD-AI, were more likely to be presumed targets of *dsx* (i.e. genes with stronger DSX DNA occupancy [Clough et al. 2014]). In gonads, there was no association. However, in heads and whole bodies, presumed targets of *dsx* were more likely to have AI, though not SD-AI (supplementary table S10, Supplementary Material online).

Estimating the Distribution of AI and SD-AI

The preceding section summarizes the results from assessing the statistical significance of each gene individually. As such, Table 1 reflects the frequencies of genes for which there was statistical power to detect evidence of AI and

SD-AI. Thus, the values in Table 1 not only depend on the aspects of biology that are our real interest (frequency and effect sizes of AI and SD-AI) but also on the number of samples, sequencing depth, and the choice of statistical threshold for significance. As an alternative to the approach used to generate Table 1, we employed a Bayesian analysis to model the underlying joint distribution of AI across the sexes in each tissue type. The results are summarized in Table 2. Across all tissues, the inferred frequency of genes with AI was similar ($F_{AI} = 38\%$ to 48%) as was the magnitude of the (sex averaged) effect (~5%). However, there were major differences between tissues with respect to the frequency of sex differences in AI. In gonads, almost all genes with AI were inferred to have sex differences ($F_{SD\text{-}AI} = 94\%$), whereas in heads very few were ($F_{SD\text{-}AI} = 5\%$). In gonads, the effect sizes of SD-AI tend to be quite large relative to the sex average effect size of AI. (This is related to a result shown in a later section: AI effects tend to be larger in male than in female gonads.) The values in Table 2 represent, to our knowledge, the only available estimates of the distribution of AI/SD-AI effects but, as with any statistical model, some caution is warranted as these estimates are contingent on the assumptions underlying the framework of the modeled distribution.

AI in Relation to Sex-Biased Gene Expression

We examined how the patterns of AI vary with the degree of sex bias in expression by fitting a Bayesian model separately to gene sets binned by sex bias. For this section, we used an alternative framework for the distribution of AI effects, parameterizing the distribution in terms of the frequency with which AI occurs, F_{AI} (as in the previous section), as well as the correlation of AI effects between males and females, ρ_{MF}. There is heterogeneity in the frequency of AI with respect to sex bias (Fig. 1a; supplementary tables S11 to S13, Supplementary Material online). In heads, F_{AI} declines from female-biased to unbiased to male-biased genes. There is a similar trend in the whole-body samples, though here it is noteworthy that the inferred average effect sizes also vary with lower estimates for female-biased genes than male-biased genes (Fig. 1a; supplementary tables S11 and S12, Supplementary Material online).

The intersexual correlation of AI effects is high in heads across all sex bias categories ($\rho_{MF} > 0.9$; Fig. 1b), whereas it is much lower in gonads ($\rho_{MF} < 0.25$). In the whole-body samples, there is a striking change in ρ_{MF} across sex bias categories, going from low values for female-biased genes to high values for male-biased ones. (Applying the other model parameterization to these data, a qualitatively similar compatible pattern is manifest in that SD-AI is common among female-biased genes and much rarer for male-biased ones [supplementary table S11, Supplementary Material online]).

Sex-Dependent Reversals of AI

Instances of SD-AI may include cases where the direction of imbalance is reversed between the sexes (i.e. male expression of the DGRP-177 allele is >50%, while female expression is <50%). Cases of sex-reversed AI are of special interest but are expected to be rare as this scenario presumably involves more complex regulation. While the Bayesian models in the previous sections are designed to make inferences about the distribution using information from a large number of genes, they are not well suited to clearly identify individual genes exhibiting sex-reversed AI. To detect cases of sex-reversed AI, we analyzed genes individually, requiring significant AI effects in both sexes but of opposite direction (a compound test for which the false positive rate depends on the framing of the null hypothesis; see Supplementary Material online). We examined whole-body samples, combining data from both main and reciprocal crosses to maximize power. We tested 3,796 genes and, under the simplest null hypothesis, ~5 genes are expected to pass the test by chance (i.e. ~5 false positives); using more conservative assumptions, ~26 false positives are expected (i.e. see Supplementary Material online). With the real data, 176 of the genes passed the statistical test for sex-reversed AI, i.e. far exceeding the expected number of false positives.

We further examined these 176 genes using the head and gonad data to see whether similar patterns can be observed in narrower tissue samples, i.e. gonads or heads. In the gonads, 155 of 176 genes had sufficient data to be examined for AI, whereas 160 genes could be examined in the heads. Of the 155 genes examined in the gonads, the majority (109) had sex-reversed point estimates of AI in the same direction as observed with the whole bodies, whereas only 5 showed sex-reversed point estimates in the opposite direction relative to that found in the whole bodies. Thus, for the putative set of genes with sex-reversed AI, there exists a fair degree of concordance in the number and direction of sex-reversed AI between the gonads and whole bodies. In contrast, in heads, only 35 of the 160 genes tested had sex-reversed point estimates of AI, and roughly half (only 17) of these were in the same direction as in the whole bodies, with the remaining half (18) in the opposite direction relative to whole bodies, i.e. not different from that expected by chance.

Tissue-Dependent Patterns of AI

In our previous analyses, we noted marked differences in the frequency of AI between heads and gonads. Though the analyses used different sets of genes for each tissue, the results suggest that AI may differ between tissues. To investigate tissue differences in AI more formally, we applied the Bayesian model to estimate the joint distribution of AI effects in heads and gonads, separately for

Table 2 Parameter estimates [with 95% high posterior density interval] for the joint distribution of AI effects across sexes

	Frequency of genes with AI, F_{AI}	Average magnitude of sex-averaged AI[a], σ_{AI}	Frequency of genes with SD-AI among genes with AI, $F_{SD\text{-}AI}$	Average magnitude of sex difference in AI[b], $\sigma_{SD\text{-}AI}$	Overdispersion parameter, υ
Gonads	0.48 [0.45, 0.52]	0.053 [0.0496, 0.056]	0.94 [0.90, 1.00]	0.105 [0.099, 0.114]	0.006 [0.005, 0.006]
Heads	0.42 [0.40, 0.45]	0.050 [0.054, 0.060]	0.05 [0.01, 0.08]	0.074 [0.051, 0.106]	0.001 [0.001, 0.001]
Whole body (main cross)	0.45 [0.44, 0.47]	0.049 [0.048, 0.051]	0.75 [0.71, 0.79]	0.059 [0.058, 0.061]	0.001 [0.001, 0.002]
Whole body (reciprocal cross)	0.38 [0.35, 0.40]	0.055 [0.053, 0.058]	0.66 [0.60, 0.71]	0.065 [0.060, 0.070]	0.003 [0.002, 0.003]

[a]σ_{AI} is the standard deviation of sex-averaged effects, which is modeled as approximately normally distributed with mean 0. [b]$\sigma_{SD\text{-}AI}$ is the standard deviation of sex differences in AI effects, which is modeled as approximately normally distributed with mean 0.

each sex (Table 3). In both sexes, we infer that ~40% of genes exhibit AI, and >80% of these have effects that differ between gonads and heads. Using the alternative Bayesian model to characterize the distribution in terms of the between-tissue correlation in effects (supplementary table S14, Supplementary Material online), we find this correlation is quite low ($\rho_{head\text{-}gonad} \approx 0.25$) and that the average magnitude of effect sizes is larger in gonads than in heads (gonads ≈ 0.09; heads ≈ 0.05).

Analogous to our earlier test for sex-reversed AI, we tested for tissue-reversed AI, separately in each sex. We find 33 instances of possible tissue-reversed AI between gonads and heads in females and 94 such instances in males. We estimate that 11 of these genes are likely false positives in females and 9 in males (see Supplementary Material online). The much higher number of putative tissue-reversed AI in males may be a result of tissue differences in AI being larger in males, leading to easier detection of reversals in AI.

Discussion

This study focuses on *cis*-regulatory variation in relation to sexual dimorphism in *D. melanogaster*. We utilize a divergent cross between North American (DGRP-177) and South African (SP159N) genotypes to maximize the number of polymorphic sites in the F_1, thus increasing our power to detect allele-specific expression. To some extent, the patterns of AI observed here may reflect interpopulation differences in expression or may be idiosyncratic to the specific lines used here. As with any study of genetic variation, our results will be influenced by the interplay between the capacity of mutation to generate regulatory variation and selection to filter it. Our goal was to explore sex-specific patterns of *cis*-regulatory variation, but our study was not designed to assess the relative importance of different evolutionary factors (e.g. mutation–selection–drift balance, balancing selection within populations, and divergent selection among populations) in causing these patterns.

When analyzing individual genes, AI was detected in 30% to 45% of genes analyzed (Table 1). These values are considerably higher than the values (6% to 17%) reported by Puixeu et al. (2023). The greater divergence of haplotypes used in our study may have contributed to this difference; we crossed fly lines from different continents, whereas they crossed lines from within a single population. However, experimental and statistical differences are likely major contributors to this discrepancy. Rather than relying on the proportion of genes that reach significance in individual gene tests that depend heavily on statistical power, we also developed a Bayesian approach to directly estimate the frequency of genes with AI, finding that the frequency is quite high (38% to 48%, Table 2).

In heads and whole-body samples, there is an enrichment of *cis*-regulatory variants in female-biased genes. Osada et al. (2017) and Puixeu et al. (2023) also examined *cis*-regulatory variation in relation to sex bias. Some of the patterns reported in those two studies appear similar, while others appear different from ours or from each other. However, it is not possible to directly compare those studies with our own or with each other because of differences in methodology use to investigate these patterns. An advantage of our approach is that patterns can be investigated with respect to biologically interpretable properties such as the frequency of AI and average effect sizes.

The higher frequency of AI among female-biased genes that we detected may suggest that these genes are more frequently under balancing selection or tend to experience weaker purifying selection. Sexually antagonistic selection is one possible reason genes could experience balancing selection or weaker purifying selection. Cheng and Kirkpatrick (2016) suggested that genes with intermediate levels of sex-biased expression—both male and female biased—would be most likely to experience sexually antagonistic selection. As we do not observe an elevated frequency of AI for genes with intermediate levels of sex bias, our results would seem to indicate that either that sexually antagonistic selection does not drive the pattern we observe or that the sexually antagonistic selection is not distributed across the genome as Cheng and Kirkpatrick (2016) suggested. While it is most intuitive to think of sexually antagonistic selection causing balancing selection or weak purifying selection, it could cause strong purifying selection if selection in one sex is much stronger than the other. Thus, the suggestion of

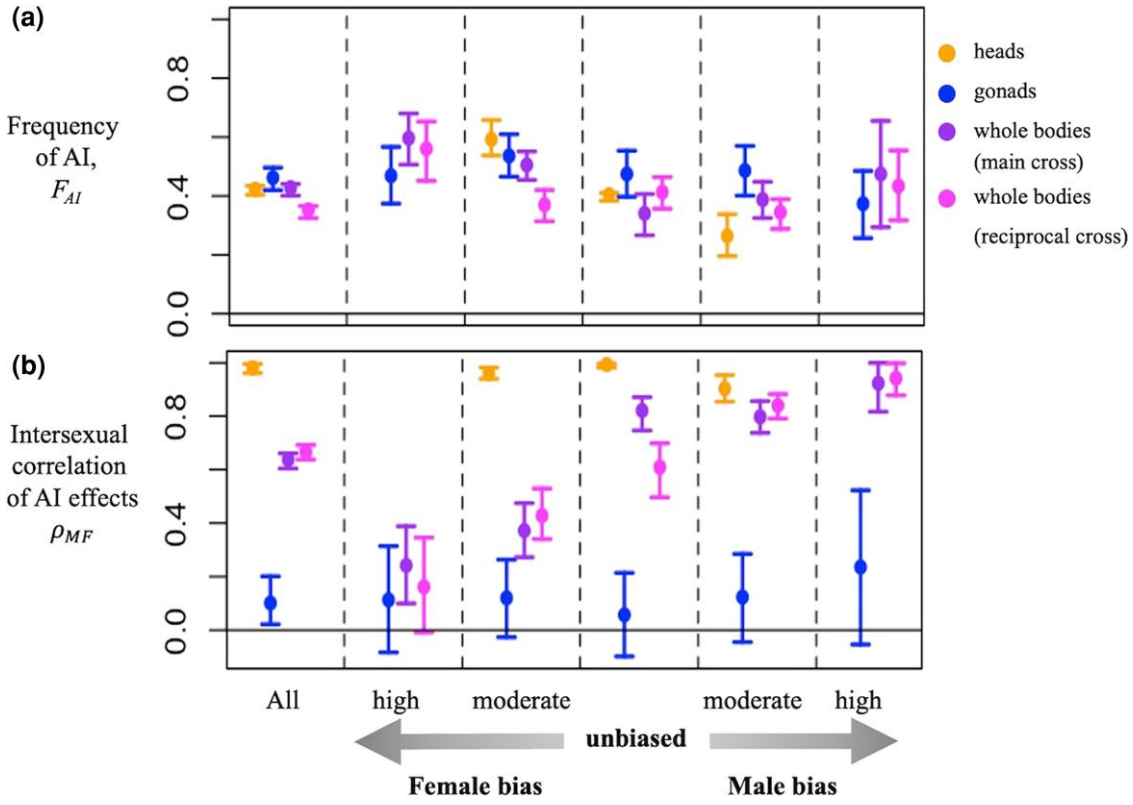

Fig. 1. Model estimates of a) the frequency of AI, F_{AI}, and b) the intersexual correlation of AI effects, ρ_{MF}. The leftmost section in each panel shows results irrespective of sex bias; the remaining sections show results stratified by sex bias: highly female biased ($-\infty < \log2FC \leq -2$), moderately female biased ($-2 < \log2FC \leq -0.5$), unbiased ($-0.5 \leq \log2FC \leq 0.5$), moderately male biased ($0.5 < \log2FC \leq 2$), and highly male biased ($2 < \log2FC < \infty$). No analysis was performed for high sex bias categories for heads as there were too few genes. The error bars represent 95% high density interval of the posterior. See supplementary table S12, Supplementary Material online for other model parameters.

Cheng and Kirkpatrick (2016) is not at necessarily at odds with our results if, for example, sexually antagonistic selection results in weak purifying selection for intermediately female-biased genes but strong purifying selection for intermediately male-biased genes.

Purifying selection (regardless of whether it involves sexual antagonism) is believed to be the most common form of selection, and genes with *cis*-regulatory variation have previously been shown to be subject to weaker purifying selection in *Capsella* (Steige et al. 2017). Strongly female-biased genes in *D. melanogaster* have somewhat elevated levels of nonsynonymous diversity (Singh and Agrawal 2023; but see details within), suggesting that they may be under weaker purifying selection and providing a potential explanation for the increased frequency of AI among these genes. A nonselective alternative explanation for our observed pattern is mutational bias; female-biased genes could have greater scope for *cis*-regulatory mutations (e.g. if such genes tend to have larger or more complex regulatory regions). We cannot ascertain which of these possibilities explain the observed pattern.

In general, *cis* effects are small in magnitude, with a large majority of genes exhibiting allelic expression differences below 5% to 10%. This finding is general to all the tissues examined, though average effects were notably higher in gonads than in heads (supplementary table S14, Supplementary Material online). Others have also found frequent, weak *cis*-regulatory effects in other species, including *Capsella* (Steige et al. 2017), mice (Crowley et al. 2015), and *Saccharomyces* (Zhang and Emerson 2019). However, there have been reports of strong *cis* effects in *D. melanogaster* (León-Novelo et al. 2018), interspecific *Drosophila* hybrids (Graze et al. 2012), and natural populations of flycatchers (Wang et al. 2017), though the detection of low-effect variants may be encumbered by the lack of statistical power in some of these studies. An advantage of the Bayesian model approach we used is that it should allow us to infer the average effect size without the bias (winner's curse) from gene-by-gene analyses. It is worth acknowledging that it is largely unknown whether small differences in expression result in biologically relevant phenotypic differences. A practical limitation of studies of multicellular eukaryotes is that very small phenotypic effects will be undetectable in any logistically feasible experiment but could still be subject to selection at sufficiently large population sizes. Indeed, population genetic inference has

Table 3 Parameter estimates [with 95% high posterior density interval] for the joint distribution of AI effects across heads and gonads

	Frequency of genes with AI, F_{AI}	Average magnitude of tissue-averaged AI[a], σ_{AI}	Frequency of genes with tissue-dependent AI among genes with AI, $F_{TD\text{-}AI}$	Average magnitude of tissue difference in AI[b], $\sigma_{TD\text{-}AI}$	Overdispersion parameter, υ
Female	0.37 [0.33, 0.41]	0.053 [0.049, 0.057]	0.94 [0.88, 1]	0.094 [0.086, 0.101]	0.0039 [0.0036, 0.0042]
Male	0.44 [0.41, 0.47]	0.050 [0.048, 0.053]	0.84 [0.78, 0.91]	0.099 [0.094, 0.105]	0.0012 [0.001, 0.0014]

[a]σ_{AI} is the standard deviation of sex-averaged effects, which is modeled as approximately normally distributed with mean 0. [b]$\sigma_{TD\text{-}AI}$ is the standard deviation of tissue differences in AI effects, which is modeled as approximately normally distributed with mean 0.

provided evidence of purifying and stabilizing selection on expression (e.g. Rifkin et al. 2005; Metzger et al. 2017; Steige et al. 2017; Glassberg et al. 2019), indicating that there are likely to be fitness consequences associated with expression variation; selection on small changes in expression may be weak (and hard to detect with experiments) but is not necessarily effectively neutral.

Using gene-by-gene analyses, we found numerous instances of sex-dependent *cis*-regulatory effects in all tissues, though these were much more common in gonads than in heads (Table 1). The inference from the Bayesian analysis indicated that almost all (>94%) *cis*-regulatory variants affect the sexes differently in gonads. Puixeu et al. (2023) also found substantially higher frequency genes with sex differences in *cis*-regulatory effects in gonads than in heads. The distinct regulatory architecture of the gonads, manifest in the high frequency of *cis*-regulatory variants with sex-dependent effects, should allow expression in gonads to evolve relatively independently between the two sexes whenever there is selective pressure to do so. By contrast, in heads we estimated that only a small fraction (~5%) of *cis*-regulatory effects has sex-dependent effects, restricting the potential for independent evolution of expression between the two sexes.

It is unsurprising that the frequency of SD-AI in whole bodies is intermediate between the estimates inferred for gonads and heads. Relatedly, the intersexual correlation of AI effects, ρ_{MF}, in whole bodies is intermediate between gonads and heads. However, there is a striking pattern of ρ_{MF} in relation to sex bias in whole bodies that is not apparent in either gonads or heads: ρ_{MF} steadily increases from strongly female-biased to strongly male-biased gene categories. This pattern seems similar to a pattern observed for the intersexual genetic correlation (r_{MF}) in whole-body gene expression (Singh and Agrawal 2023) in which the average r_{MF} increases from strongly female-biased to strongly male-biased gene categories. It should be noted ρ_{MF} and r_{MF} are not defined at the same "level"; ρ_{MF} is a correlation in AI effects between sexes estimated from a set of genes, whereas each gene has its own r_{MF} value—the intersexual correlation in expression level—estimated among a set of genotypes. Nonetheless, one would expect that if ρ_{MF} was low (high) for a class of genes, then the average r_{MF} for that class of genes would also be low (high) if *cis*-regulatory effects were a major source of the genetic variation in expression. It is unclear why this pattern in ρ_{MF} occurs in whole bodies but not heads or gonads, and it is also unknown if r_{MF} is related to sex bias in tissues other than whole bodies.

Reversals in the direction of AI between the sexes are the most extreme form of SD-AI, and we detect such reversals at 176 loci in whole-body samples. Puixeu et al. (2023) also documented some examples of sex-reversed AI (though using less stringent criteria to assess false positives). These are intriguing because heterozygotes of each sex will predominantly express alternative alleles for these genes. If the coding regions of the associated alleles were under sexually antagonistic selection, this type of *cis*-regulatory variation provides a potential mechanism by which sex-specific dominance reversals could occur (Grieshop et al. 2024). Dominance reversals have been invoked in the theoretical literature to explain the maintenance of genetic variation (Kidwell et al. 1977; Curtsinger et al. 1994; reviewed in Grieshop et al. 2024), though it has been unclear whether dominance reversals would be biologically plausible. Recently, this idea has gathered increasing empirical support with the discovery of sex-specific dominance reversal for traits likely under sexually divergent selection in salmonids (Barson et al. 2015; Pearse et al. 2019) as well as signals of dominance reversals for fitness in seed beetles (Grieshop and Arnqvist 2018). Sex reversals in AI potentially provide a mechanism, but it remains to be seen whether such *cis*-regulatory effects underlie any cases of dominance reversals for fitness.

Though sex differences in AI could arise via different routes, one possibility is that such effects are mediated by *doublesex*, a transcription factor with different isoforms in the two sexes (DSX^M and DSX^F). However, we did not find that presumed targets of DSX were more likely to have SD-AI. Perhaps this is not surprising as DSX^M and DSX^F have the same N-terminus, which contains the DNA-binding domain (Camara et al. 2008). Given that DSX^M and DSX^F bind DNA similarly, the sex differences DSX mediates are thought to be due to how their differing C-termini interact with other proteins (e.g. *intersex*). Somewhat surprisingly, presumed DSX targets were more likely to have AI than nontargets. The greater frequency of expression variants in DSX targets might indicate that these genes are under weaker selective constraint compared with genes whose expression is in not sex dependent.

In this study, we focused on sex-dependent aspects of *cis*-regulatory variation because of its implications for

sexual dimorphism and sexual conflict. However, "sex" is one of many dimensions in which the context for a gene's expression can differ. *Cis*-regulatory effects can vary within other contexts, such as temperature (Chen et al. 2015; Li and Fay 2017), tissue (Pinter et al. 2015), or metabolic state (Shih and Fay 2021). Of the two dimensions we examined —sex and tissue—neither stood out as harboring dramatically more variation than the other (i.e. SD-AI was very common in gonads as was tissue-dependent AI in both sexes). We were motivated to look for SD-AI, at least in part, because of the possibility that sexually antagonistic selection might help maintain such variants. However, the qualitatively similar frequency of tissue-dependent AI suggests that either sexual antagonism is playing little role in maintaining regulatory variation genome wide or that antagonistic pleiotropy across tissues is similarly common. Regardless, our results add to the existing literature that *cis*-regulatory effects are often context dependent in magnitude and sometimes even in direction. Moreover, different dimensions of context can interact to affect allelic expression, e.g. we find a sizeable number of genes with tissue-by-sex interaction effects on AI. It is plausible that this thinking extends to other contexts (e.g. sex-dependent *cis*-regulatory effects may be sensitive to temperature).

Allele-specific expression analysis can be plagued by inaccurate mapping of reads to the reference. We attempted to minimize the influence of mapping bias in several ways. We constructed genotype-specific references and ascertained their efficacy by competitively mapping parental genomic reads to the references. We tested F_1 genomic data for significant deviation from the expected read count ratio of 1:1 and removed any such gene from further consideration. We also tested for apparent "parental effects" on allelic expression, which may ensue from any residual mapping bias, and excluded any such genes. Among our retained genes, the percentage of reads assigned to the DGRP-177 allele are uncorrelated between genomic and transcriptomic data, whereas a positive correlation would be expected if mapping bias was a substantial problem. The correlation remains negligible even if we only consider those genes with significant AI in expression. These steps are designed to allay the effects of mapping bias on our inferences, with evidence suggesting that they have been largely effective (see Supplementary Material online). Nonetheless, we cannot entirely rule out the possibility that mapping bias may persist in the dataset within a small number of retained genes. While we believe that our major findings are robust, the results for individual genes would benefit from further experimental validation.

As we infer *cis*-regulatory polymorphisms from AI (i.e. relative expression) in heterozygotes, an additional note of caution is needed with respect to total expression. If allele *A* is expressed more than *B* in an *A/B* heterozygote, it is reasonable to assume that total expression of this gene would be greater in *A/A* than in *B/B* homozygotes. However, this may not always be the case owing to negative feedback loops in transcription. Further experimental testing—including both homozygotes—would be needed to evaluate how often observed AI causes the predicted effects on total expression, though this is extremely difficult to do at scale while controlling for genetic background effects.

In summary, our data indicate that *cis*-regulatory variants are common, and their effects are often dependent on sex and tissue. The frequency of *cis*-regulatory variants varies nonrandomly with respect to sex-biased gene expression and across tissues. Further work remains to be done to more thoroughly examine variants at individual loci and to understand the evolutionary forces that shape patterns of variation genome wide.

Methods

Sample Preparation and Sequencing

We crossed two inbred lines of *D. melanogaster* to obtain F_1 individuals for the study. DGRP-177 belongs to the Drosophila Genetic Reference Panel, a suite of genotypes derived from natural populations of *D. melanogaster* in North Carolina (Mackay et al. 2012). SP159N was derived from a South African population, as part of the Drosophila Population Genomics Project (Pool et al. 2012; Lack et al. 2016). We crossed DGRP-177 males with SP159N females (main cross) and also performed a reciprocal cross. All flies were reared and maintained at 25 °C on a 12-h light-dark cycle. After eclosion, virgin F_1 males and females were collected and kept in separate vials for 3 d. Vials that were observed to have eggs laid in them were discarded. The F_1 individuals from the main cross (DGRP-177 males × SP159N females) were used to obtain samples of whole body flies, heads, and gonads, with three replicate samples of each sex × tissue-type combination. From the reciprocal cross, only whole-body samples were collected (three replicates of each sex). Testes and ovaries were obtained by dissection of anesthetized whole flies, performed over ice, with the flies placed in phosphate buffered saline, i.e., PBS buffer. Each replicate for gonads consisted of 10 to 12 individuals, while each replicate for whole flies and heads consisted of 5 to 7 individuals. Heads and whole flies were flash-frozen in liquid nitrogen and then stored at −80 °C as they awaited RNA extraction; gonads were stored at −80 °C upon dissection. RNA extraction was performed using the ThermoFisher PicoPure RNA Isolation Kit. Paired-end sequencing at a read length of 200 bp was performed using Illumina NovaSeq6000. In addition, DNA was extracted from the parents and F_1 individuals using the Qiagen DNeasy Blood and Tissue Kit. Paired-end whole-genome sequences were obtained at a read length of 100 bp using Illumina HiSeqX.

Genotype-Specific References

Genotype-specific references are constructed from the reference genome by incorporating SNPs for each parental genotype separately to reduce mapping bias when analyzing allele-specific expression (Rozowsky et al. 2011; Graze et al. 2012). We called variants (SNPs and indels) for both DGRP-177 and SP159N using the GATK Best Practices workflow (McKenna et al. 2010; Auwera et al. 2013). The whole-genome sequences for DGRP-177 and SP159N were aligned to the *D. melanogaster* Release 6 reference genome (dos Santos et al. 2015) using BWA (Li and Durbin 2009) with default parameters. The resulting alignment file was processed using GATK tools to sort reads and mark PCR/optical duplicates (PCR = polymerase chain reaction). Variants were called by using GATK HaplotypeCaller, and the resulting variant call format (VCF) file was separated into two VCFs, one each for indels and SNPs. We obtained a BED file of coordinates corresponding to heterozygous SNP calls and regions around indels for each parental genome. Depending on the length of the indels, the following coordinates were included in the BED file: 10 bp around indels of length ≤6 bp, 20 bp around indels of length >6 bp but ≤12 bp, and 100 bp around indels >12 bp. The coordinates were then masked in the reference genome with Ns. We masked sites without high certainty of a homozygous SNP call in both parents, i.e. any sites where the major allele was <95%. We also masked sites where the depth of coverage fell below 10. Following recommendations by GATK, the remaining variants—the homozygous SNPs—were subject to hard filtering based upon quality measures (supplementary table S1, Supplementary Material online). The filtered SNPs corresponding to DGRP-177 and SP159N were incorporated separately into the masked reference genome, to create genotype-specific references for each parent.

Competitive Mapping

For each sample, RNA-seq reads were aligned to both genotype-specific references using STAR v2.7 (Dobin et al. 2013) with default parameters. Following a similar method to Sánchez-Ramírez and Cutter (2021), a Python-based competitive read mapping approach was used to obtain allele-specific read counts from RNA-seq data for the F_1 heterozygotes. The two BAM alignments yielded by STAR were sorted and indexed by read names. The alignment score (AS) and number of mismatches (nM) of each read when aligned to the two references were used to assign each read's origin as follows: (i) if a read had a higher AS when aligned to a given parent's genotype-specific reference, it was ascribed as belonging to that genotype, (ii) if the read has equal AS for both alignments, the read was assigned to the genotype that yielded a lower nM, and (iii) if AS and nM do not differ between the two alignments for a read, it was assigned to neither genotype, i.e. an "ambiguous" read. (We also tried the competitive alignment using only nM to ascribe parental origin and obtained results similar to that from using both AS and nM.) Ambiguous reads were excluded from further analysis. The percentage of such informative (i.e. nonambiguous) reads ranged from 18% to 43% for the samples examined (supplementary table S2, Supplementary Material online). Competitive read mapping yielded 2 alignment files, corresponding to allele-specific reads of each parent. Finally, allele-specific read counts were obtained from those BAM alignments using HTSeq-count v0.12 (Anders et al. 2015), with secondary or chimeric alignments being ignored when aggregating reads.

Bias in mapping can occur for multiple reasons and lead to errors in estimation of allele-specific expression (Degner et al. 2009; León-Novelo et al. 2014). Genomic DNA from F_1s should have 50% of reads mapping to each parent in the absence of mapping bias. As described in the Supplementary Material online, we used genomic DNA from male and female F_1 samples to conservatively filter out genes with possible mapping bias (supplementary text S1, figs. S1 to S3, and tables S3 to S5, Supplementary Material online).

Parental Effects in AI

Parental effects on allele-specific gene expression in *Drosophila* are believed to be rare (Wittkopp et al. 2006; Coolon et al. 2012; Puixeu et al. 2023). Still, certain forms of mapping bias could manifest as apparent parental effects. To detect real or apparent "parental" effects and then remove such genes from further analyses, we applied generalized linear models to the whole-body samples from the main and reciprocal crosses utilizing the R package *lme4* v.1 (Bates et al. 2015). The first test was applied separately to each sex. Only genes that had read counts >30 in all the 6 replicates were considered. For each gene, we analyzed the proportion of reads assigned to the DGRP-177 genome with a quasibinomial model, including cross direction (main vs. reciprocal) as a fixed effect. The intercept term of this model indicates AI in the absence of an effect of cross direction. Additionally, we analyzed the data for both sexes in a single model, including sex and "cross direction" as fixed effects, but without an interaction term. This test was applied to all genes with a total read count >20 in all 6 male and female replicates and average read count >30. All genes that had a significant cross-direction effect ($P < 0.05$) were excluded from subsequent analyses.

Inference of AI and Sex Differences in AI

For our analyses examining AI and SD-AI, the data for each "tissue" type (gonads, heads, and whole bodies) were analyzed separately. Whole bodies from the main and reciprocal crosses were also analyzed separately for 2 reasons. First, as we expect similar patterns in the whole-body

analyses from each cross direction, any major discrepancies could serve as a cause for concern. Second, the whole-body results are comparable with results for gonads and heads because all the analyses are based on three replicates per sex for each tissue type. Only genes with at least 30 assignable reads in each of the replicates were included in the analysis for a given tissue. For each gene, we analyzed the proportion of reads assigned to the DGRP-177 genome with a quasibinomial model, including sex as a fixed effect and using sum contrasts. In this model, the intercept term indicates the sex-averaged AI effect, and the "sex" term indicates the difference in AI between sexes (i.e. a significant "sex" term is evidence of SD-AI). Note that only autosomal genes were used in these and all analyses described.

Inference of Tissue Effects on AI

We performed analyses contrasting heads and gonads, parallel to our analyses of the two sexes. Separately for each sex, we first analyzed the gonad and head data with a quasibinomial model including a tissue term. Significance of the intercept term is taken as evidence of (tissue averaged) AI. A significant tissue term ($P < 0.05$) indicates tissue-dependent AI. Unexpectedly, we noted substantial differences in the frequency of AI and tissue-dependent AI in the male and female analyses. Subsequently, we examined those genes for which there were sufficient data in both sexes using a quasibinomial model that included sex, tissue, and their interaction. The interaction term indicates sex differences in tissue-dependent AI.

Inference of Sex- and Tissue-Reversed AI

The strongest form of SD-AI—sex-reversed AI—occurs when one allele is responsible for >50% expression in males, while the other allele is more highly expressed in females. To maximize our power, we jointly analyzed the whole-body data from the main and reciprocal crosses. All genes with read counts <20 in any of the replicates were excluded from analysis. Additionally, we excluded genes where the average read count across all the 12 replicates was <30. We then tested each sex separately for AI, using a quasibinomial model that only fit an intercept term, which signified the AI effect. A gene was said to have sex-reversed AI if it fulfilled the following conditions: (i) it had a significant AI effect in females, (ii) it had a significant AI effect in males, and (iii) the AI effect in the sexes was in opposite directions. The number of false positives expected from this procedure depends on the assumptions one makes for a null hypothesis. We considered several possibilities to arrive at a conservative estimate of the number of false positive cases of dominance reversal (supplementary text S1, Supplementary Material online).

Likewise, we used data from gonads and heads to detect cases of tissue-reversed AI in males and females separately. Again, after excluding genes with low read counts as above, we tested each tissue separately for AI using a quasibinomial model where the intercept term represented AI. A gene was said to have tissue-reversed AI if it showed significant AI in both heads and gonadal tissue, and the direction of the AI was opposite between the two tissues. The number of false positive cases of tissue-reversed AI was estimated in a manner analogous to that for sex-reversed AI (supplementary text S1, Supplementary Material online).

Modeling the Distribution of AI and SD-AI

In the earlier section, we described analyzing each gene individually to test for the presence of AI and SD-AI. As with any set of frequentist analyses, the results will be subject to both false positives and false negatives. Moreover, estimates of the average magnitudes of AI and SD-AI using "significant" genes will be subject to bias due to the "winner's curse." As an alternative, we used a Bayesian approach to model the joint distribution of AI effects across sexes, separately for each tissue.

As the true features of this distribution are unknown, there are many ways one could choose to model it. Our goal was to approximate the distribution using relative simple models framed with respect to our key biological interests. To this end, we developed two models, each characterized by four parameters. The first model framework characterizes the distribution in terms of (i) the frequency of genes with nonzero sex-averaged AI, F_{AI}; (ii) the standard deviation in sex-averaged AI, σ_{AI}; (iii) the frequency of genes with SD-AI among those with nonzero sex-averaged AI, $F_{SD\text{-}AI}$; and (iv) the standard deviation in sex difference in AI, $\sigma_{SD\text{-}AI}$. We model AI as a deviation in expression of allele 1 from 50% and sex differences in AI as the difference of allele 1 expression in males from that in females. Because we assume there is no net directionality to AI or SD-AI effects (i.e. there is no bias across genes, on average, of which parental strain allele is more highly expressed), the average values of AI and SD-AI are zero; thus, σ_{AI} and $\sigma_{SD\text{-}AI}$ represent the average magnitudes of AI and SD-AI, respectively. The second model framework is similar but instead characterizes the distribution in terms of (i) the frequency of genes with nonzero sex-averaged AI, F_{AI} (as in the first model framework); (ii) the standard deviation in AI in females, $\sigma_{AI,F}$; (iii) the standard deviation in AI in males, $\sigma_{AI,M}$; and (iv) the intersexual correlation in AI effects, ρ_{MF}. The first model framework is motivated primarily by an interest in estimating the fraction of AI effects that differ between the sexes ($F_{SD\text{-}AI}$), whereas the second is motivated primarily by a related but different property, the intersexual correlation in AI effects, ρ_{MF}.

For both model frameworks, the likelihood function relating the observed read counts to the distribution incorporates a fifth (nuisance) parameter to allow for overdispersion. Because genes for which we have a high amount of data

(e.g. highly expressed genes) could have a disproportionate influence on the inferred distribution of AI effects, we performed the analysis on a filtered dataset for each tissue type. We removed genes in the bottom quartile of total read count for either sex and then down-sampled the read counts to equal levels for all genes (separately by sex). It is worth remembering that these (and other) filtering steps nonrandomly exclude genes from analysis (e.g. genes with low expression in either or both sexes), and thus, the inferred distribution may not be representative of such genes. We used the R package *BayesianTools* (Hartig et al. 2023) to estimate a posterior distribution. Further details about the models are provided in the supplementary text S3, Supplementary Material online.

Differential Gene Expression Analysis

We obtained gene-level read counts by using *htseq-count* on BAM alignments of each RNA-seq sample to the *D. melanogaster* reference genome. We then estimated differential gene expression between the sexes using the R package *DESeq2* (Love et al. 2014). In estimating differential expression, we removed any genes averaging fewer than 50 reads across all the replicates for a given tissue. Differential expression at each gene is estimated as the $\log_2$ fold change in male-to-female gene expression ($\log_2$FC). Excluding a small number of genes where the $\log_2$FC had a standard error >1.5, all other genes were assigned to one of five sex-bias bins, as follows: highly female biased ($-\infty < \log_2\text{FC} \leq -2$), moderately female biased ($-2 < \log_2\text{FC} \leq -0.5$), unbiased ($-0.5 \leq \log_2\text{FC} \leq 0.5$), moderately male biased ($0.5 < \log_2\text{FC} \leq 2$), and highly male biased ($2 < \log_2\text{FC} < \infty$). Genes with relatively high uncertainty (i.e. standard error for $\log_2$FC >1.5) were not assigned to any of these bins. The four tissue types—heads, gonads, whole bodies (main cross), and whole bodies (reciprocal cross)—were separately analyzed for sex-biased expression.

Supplementary Material

Supplementary material is available at *Genome Biology and Evolution* online.

Acknowledgments

We thank Ina Anreiter for her help with dissections, Baharul Choudhury for his help with RNA extractions, Santiago Sánchez-Ramírez for bioinformatics advice, and Stephen Wright for helpful discussion. We thank John Pool and Yuheng Huang (U Madison) for providing the SP159N line.

Funding

This work was supported by the Natural Sciences and Engineering Research Council of Canada (Discovery Grant to A.F.A.).

Conflict of Interest

All authors declare no conflict of interests.

Data Availability

RNA-seq data are available at the NIH Sequence Read Archive (accession: PRJNA1173240) and analysis scripts at https://github.com/mishrap1971/Mishra_et_al_Drosophila_ASE.

Literature Cited

Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015:31(2):166–169. https://doi.org/10.1093/bioinformatics/btu638.

Arbeitman MN, New FN, Fear JM, Howard TS, Dalton JE, Graze RM. Sex differences in *Drosophila* somatic gene expression: variation and regulation by *doublesex*. G3 (Bethesda, Md.). 2016:6(7):1799–1808. https://doi.org/10.1534/g3.116.027961.

Auwera GA, Carneiro MO, Hartl C, Poplin R, del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J, et al. From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. Curr Protoc Bioinformatics. 2013:43(1110):11.10.1–11.10.33. https://doi.org/10.1002/0471250953.bi1110s43.

Barson NJ, Aykanat T, Hindar K, Baranski M, Bolstad GH, Fiske P, Jacq C, Jensen AJ, Johnston SE, Karlsson S, et al. Sex-dependent dominance at a single locus maintains variation in age at maturity in salmon. Nature. 2015:528(7582):405–408. https://doi.org/10.1038/nature16062.

Bates D, Mäechler M, Bolker B, Walker S. Fitting linear mixed-effects models using *lme4*. J Stat Softw. 2015:67(1):1–48. https://doi.org/10.18637/jss.v067.i01.

Bonduriansky R, Chenoweth SF. Intralocus sexual conflict. Trends Ecol Evol. 2009:24(5):280–288. https://doi.org/10.1016/j.tree.2008.12.005.

Camara N, Whitworth C, Van Doren M. The creation of sexual dimorphism in the *Drosophila soma*. Curr Top Dev Biol. 2008:83:65–107. https://doi.org/10.1016/S0070-2153(08)00403-1.

Campbell CD, Kirby A, Nemesh J, Daly MJ, Hirschhorn JN. A survey of allelic imbalance in F1 mice. Genome Res. 2008:18(4):555–563. https://doi.org/10.1101/gr.068692.107.

Chen J, Nolte V, Schlötterer C. Temperature stress mediates decanalization and dominance of gene expression in *Drosophila melanogaster*. PLoS Genet. 2015:11(2):e1004883. https://doi.org/10.1371/journal.pgen.1004883.

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

Clough E, Jimenez E, Kim YA, Whitworth C, Neville MC, Hempel LU, Pavlou HJ, Chen ZX, Sturgill D, Dale RK, et al. Sex-and tissue-specific functions of Drosophila *doublesex* transcription factor target genes. Dev Cell. 2014:31(6):761–773. https://doi.org/10.1016/j.devcel.2014.11.021.

Connallon T, Chenoweth SF. Dominance reversals and the maintenance of genetic variation for fitness. PLoS Biol. 2019:17(1):e3000118. https://doi.org/10.1371/journal.pbio.3000118.

Connallon T, Clark AG. Balancing selection in species with separate sexes: insights from Fisher's geometric model. Genetics. 2014:197(3):991–1006. https://doi.org/10.1534/genetics.114.165605.

Coolon JD, Stevenson KR, McManus CJ, Graveley BR, Wittkopp PJ. Genomic imprinting absent in *Drosophila melanogaster* adult

females. Cell Rep. 2012:2(1):69–75. https://doi.org/10.1016/j.celrep.2012.06.013.
Crowley JJ, Zhabotynsky V, Sun W, Huang S, Pakatci IK, Kim Y, Wang JR, Morgan AP, Calaway JD, Aylor DL, et al. Analyses of allele-specific gene expression in highly divergent mouse crosses identifies pervasive allelic imbalance. Nat Genet. 2015:47(4):353–360. https://doi.org/10.1038/ng.3222.
Cubillos FA, Stegle O, Grondin C, Canut M, Tisné S, Gy I, Loudet O. Extensive *cis*-regulatory variation robust to environmental perturbation in *Arabidopsis*. Plant Cell. 2014:26(11):4298–4310. https://doi.org/10.1105/tpc.114.130310.
Curtsinger JW, Service PM, Prout T. Antagonistic pleiotropy, reversal of dominance, and genetic polymorphism. Am Nat. 1994:144(2):210–228. https://doi.org/10.1086/285671.
Degner JF, Marioni JC, Pai AA, Pickrell JK, Nkadori E, Gilad Y, Pritchard JK. Effect of read-mapping biases on detecting allele-specific expression from RNA-sequencing data. Bioinformatics. 2009:25(24):3207–3212. https://doi.org/10.1093/bioinformatics/btp579.
Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013:29(1):15–21. https://doi.org/10.1093/bioinformatics/bts635.
dos Santos G, Schroeder AJ, Goodman JL, Strelets VB, Crosby MA, Thurmond J, Emmert DB, Gelbart WM; the FlyBase Consortium. FlyBase: introduction of the *Drosophila melanogaster* release 6 reference genome assembly and large-scale migration of genome annotations. Nucleic Acids Res. 2015:43(D1):D690–D697. https://doi.org/10.1093/nar/gku1099.
Fairbairn DJ, Blanckenhorn WU, Székely T. Sex, size and gender roles: evolutionary studies of sexual size dimorphism. Oxford (UK): Oxford University Press; 2007.
Genissel A, McIntyre LM, Wayne ML, Nuzhdin SV. Cis and trans regulatory effects contribute to natural variation in transcriptome of *Drosophila melanogaster*. Mol Biol Evol. 2007:25(1):101–110. https://doi.org/10.1093/molbev/msm247.
Glassberg EC, Gao Z, Harpak A, Lan X, Pritchard JK. Evidence for weak selective constraint on human gene expression. Genetics. 2019:211(2):757–772. https://doi.org/10.1534/genetics.118.301833.
Goncalves A, Leigh-Brown S, Thybert D, Stefflova K, Turro E, Flicek P, Brazma A, Odom DT, Marioni JC. Extensive compensatory *cis-trans* regulation in the evolution of mouse gene expression. Genome Res. 2012:22(12):2376–2384. https://doi.org/10.1101/gr.142281.112.
Graze RM, Novelo LL, Amin V, Fear JM, Casella G, Nuzhdin SV, McIntyre LM. Allelic imbalance in *Drosophila* hybrid heads: exons, isoforms, and evolution. Mol Biol Evol. 2012:29(6):1521–1532. https://doi.org/10.1093/molbev/msr318.
Grieshop K, Arnqvist G. Sex-specific dominance reversal of genetic variation for fitness. PLoS Biol. 2018:16(12):e2006810. https://doi.org/10.1371/journal.pbio.2006810.
Grieshop K, Ho EKH, Kasimatis KR. Dominance reversals: the resolution of genetic conflict and maintenance of genetic variation. Proc Biol Sci. 2024:291(2018):20232816. https://doi.org/10.1098/rspb.2023.2816.
Gruber JD, Long AD. *Cis* -regulatory variation is typically polyallelic in *Drosophila*. Genetics. 2009:181(2):661–670. https://doi.org/10.1534/genetics.108.098459.
Hartig F, Minunno F, Paul S. 2023. BayesianTools: general-purpose MCMC and SMC samplers and tools for Bayesian statistics. R package version 0.1.8. https://CRAN.R-project.org/package=BayesianTools.
Kidwell JF, Clegg MT, Stewart FM, Prout T. Regions of stable equilibria for models of differential selection in the two sexes under random mating. Genetics. 1977:85(1):171–183. https://doi.org/10.1093/genetics/85.1.171.
Kita R, Venkataram S, Zhou Y, Fraser HB. High-resolution mapping of *cis*-regulatory variation in budding yeast. Proc Natl Acad Sci U S A. 2017:114(50):E10736–E10744. https://doi.org/10.1073/pnas.1717421114.
Lack JB, Lange JD, Tang AD, Corbett-Detig RB, Pool JE. A thousand fly genomes: an expanded *Drosophila* genome nexus. Mol Biol Evol. 2016:33(12):3308–3313. https://doi.org/10.1093/molbev/msw195.
Lemos B, Araripe LO, Fontanillas P, Hartl DL. Dominance and the evolutionary accumulation of cis- and trans-effects on gene expression. Proc Natl Acad Sci U S A. 2008:105(38):14471–14476. https://doi.org/10.1073/pnas.0805160105.
León-Novelo L, Gerken AR, Graze RM, McIntyre LM, Marroni F. Direct testing for allele-specific expression differences between conditions. G3 (Bethesda). 2018:8(2):447–460. https://doi.org/10.1534/g3.117.300139.
León-Novelo LG, McIntyre LM, Fear JM, Graze RM. A flexible Bayesian method for detecting allelic imbalance in RNA-seq data. BMC Genomics. 2014:15(1):920. https://doi.org/10.1186/1471-2164-15-920.
Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009:25(14):1754–1760. https://doi.org/10.1093/bioinformatics/btp324.
Li XC, Fay JC. Cis-regulatory divergence in gene expression between two thermally divergent yeast species. Genome Biol Evol. 2017:9(5):1120–1129. https://doi.org/10.1093/gbe/evx072.
Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014:15(12):550. https://doi.org/10.1186/s13059-014-0550-8.
Mackay TFC, Richards S, Stone EA, Barbadilla A, Ayroles JF, Zhu D, Casillas S, Han Y, Magwire MM, Cridland JM, et al. The *Drosophila melanogaster* genetic reference panel. Nature. 2012:482(7384):173–178. https://doi.org/10.1038/nature10811.
McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010:20(9):1297–1303. https://doi.org/10.1101/gr.107524.110.
Metzger BPH, Wittkopp PJ, Coolon JD. Evolutionary dynamics of regulatory changes underlying gene expression divergence among Saccharomyces species. Genome Biol Evol. 2017:9(4):843–854. https://doi.org/10.1093/gbe/evx035.
Osada N, Miyagi R, Takahashi A. *Cis*- and *trans*-regulatory effects on gene expression in a natural population of *Drosophila melanogaster*. Genetics. 2017:206(4):2139–2148. https://doi.org/10.1534/genetics.117.201459.
Parsch J, Ellegren H. The evolutionary causes and consequences of sex-biased gene expression. Nat Rev Genet. 2013:14(2):83–87. https://doi.org/10.1038/nrg3376.
Pastinen T, Hudson TJ. Cis-acting regulatory variation in the human genome. Science. 2004:306(5696):647–650. https://doi.org/10.1126/science.1101659.
Pearse DE, Barson NJ, Nome T, Gao G, Campbell MA, Abadía-Cardoso A, Anderson EC, Rundio DE, Williams TH, Naish KA, et al. Sex-dependent dominance maintains migration supergene in rainbow trout. Nat Ecol Evol. 2019:3(12):1731–1742. https://doi.org/10.1038/s41559-019-1044-6.
Pinter SF, Colognori D, Beliveau BJ, Sadreyev RI, Payer B, Yildirim E, Wu C, Lee JT. Allelic imbalance is a prevalent and tissue-specific feature of the mouse transcriptome. Genetics. 2015:200(2):537–549. https://doi.org/10.1534/genetics.115.176263.
Pool JE, Corbett-Detig RB, Sugino RP, Stevens KA, Cardeno CM, Crepeau MW, Duchen P, Emerson JJ, Saelao P, Begun DJ, et al. Population genomics of sub-Saharan *Drosophila melanogaster*:

African diversity and non-African admixture. PLoS Genet. 2012:8(12):e1003080. https://doi.org/10.1371/journal.pgen.1003080.

Puixeu G, Macon A, Vicoso B. Sex-specific estimation of *cis* and *trans* regulation of gene expression in heads and gonads of *Drosophila melanogaster*. G3 (Bethesda). 2023:13(8):jkad121. https://doi.org/10.1093/g3journal/jkad121.

Rifkin SA, Houle D, Kim J, White KP. A mutation accumulation assay reveals a broad capacity for rapid evolution of gene expression. Nature. 2005:438(7065):220–223. https://doi.org/10.1038/nature04114.

Rozowsky J, Abyzov A, Wang J, Alves P, Raha D, Harmanci A, Leng J, Bjornson R, Kong Y, Kitabayashi N, et al. AlleleSeq: analysis of allele-specific expression and binding in a network framework. Mol Syst Biol. 2011:7(1):522. https://doi.org/10.1038/msb.2011.54.

Rowe L, Chenoweth SF, Agrawal AF. The genomics of sexual conflict. Amer Nat. 2018:192(2):274–286. https://doi.org/10.1086/698198.

Sánchez-Ramírez S, Cutter AD. CompMap: an allele-specific expression read-counter based on competitive mapping. bioRxiv 431019. https://doi.org/10.1101/2021.02.12.431019, preprint: not peer reviewed.

Shih C-H, Fay J. Cis-regulatory variants affect gene expression dynamics in yeast. eLife. 2021:10:e68469. https://doi.org/10.7554/eLife.68469.

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

Spencer HG, Priest NK. The evolution of sex-specific dominance in response to sexually antagonistic selection. Amer Nat. 2016:187(5):658–666. https://doi.org/10.1086/685827.

Steige KA, Laenen B, Reimegård J, Scofield DG, Slotte T. Genomic analysis reveals major determinants of *cis*- regulatory variation in *Capsella grandiflora*. Proc Natl Acad Sci U S A. 2017:114(5):1087–1092. https://doi.org/10.1073/pnas.1612561114.

Stern DL. Perspective: evolutionary developmental biology and the problem of variation. Evolution. 2000:54(4):1079–1091. https://doi.org/10.1111/j.0014-3820.2000.tb00544.x.

Stranger BE, Montgomery SB, Dimas AS, Parts L, Stegle O, Ingle CE, Sekowska M, Smith GD, Evans D, Gutierrez-Arcelus M, et al. Patterns of cis regulatory variation in diverse human populations. PLoS Genet. 2012:8(4):e1002639. https://doi.org/10.1371/journal.pgen.1002639.

Wang M, Uebbing S, Ellegren H. Bayesian inference of allele-specific gene expression indicates abundant cis-regulatory variation in natural flycatcher populations. Genome Biol Evol. 2017:9(5):1266–1279. https://doi.org/10.1093/gbe/evx080.

Wittkopp PJ, Haerum BK, Clark AG. Evolutionary changes in cis and trans gene regulation. Nature. 2004:430(6995):85–88. https://doi.org/10.1038/nature02698.

Wittkopp PJ, Haerum BK, Clark AG. Parent-of-origin effects on mRNA expression in *Drosophila melanogaster* not caused by genomic imprinting. Genetics. 2006:173(3):1817–1821. https://doi.org/10.1534/genetics.105.054684.

Wittkopp PJ, Haerum BK, Clark AG. Regulatory changes underlying expression differences within and between *Drosophila* species. Nat Genet. 2008:40(3):346–350. https://doi.org/10.1038/ng.77.

Wray GA. The evolutionary significance of cis-regulatory mutations. Nat Rev Genet. 2007:8(3):206–216. https://doi.org/10.1038/nrg2063.

Zhang X, Emerson JJ. Inferring the genetic architecture of expression variation from replicated high throughput allele-specific expression experiments. bioRxiv 699074. https://doi.org/10.1101/699074, preprint: not peer reviewed.

Associate editor: Toni Gossmann