Article

Cis-regulatory variation in relation to sex and sexual dimorphism in *Drosophila melanogaster*

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

Affiliations:
[1]Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, ON M5S 3B2, Canada
[2]Department of Molecular Biosciences, The Wenner-Gren Institute, Stockholm University, Stockholm SE-10691, Sweden

ORCIDs:
https://orcid.org/0000-0001-8925-5066 (KG)

Keywords:
sex biased gene expression, allelic imbalance, *cis*-regulatory variation

Short title:
Sex differences in *cis*-regulatory variation

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 test for sex differences in *cis*-regulatory effects as well as examine patterns of *cis*-regulatory with respect to two other levels of variation in sexual dimorphism: (i) across genes varying 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 said variation. A sizeable proportion of *cis*-regulatory variation is observed to have sex-specific effects, with sex-dependent *cis* effects being more frequent in gonads than in heads. Finally, we detect some genes with reversed allelic imbalance between the sexes. Such variants could provide a mechanism for sex-specific dominance reversals, a phenomenon important for sexually antagonistic balancing selection.

Introduction

Sexual dimorphism is a universal feature of sexually reproducing species, with males and females differing in their appearance, physiology, life history and behaviour (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 & Ellegren, 2013). Divergent selection on the sexes is thought to be a major reason behind the evolution of sex-biased gene expression (Schärer et al., 2012; Connallon & Clark, 2011; Parsch & Ellegren, 2013; but see Houle & Cheng, 2021). Sex-specific strategies to maximise reproductive success arises from sex differences in gamete investment, known as anisogamy (Schärer et al., 2012), ultimately leading to sex-specific optima for shared traits. Thus, genes underlying such traits often experience divergent, sex-specific selection, termed 'intralocus sexual conflict' (Bonduriansky & Chenoweth, 2009). Resolution of sexual conflict at a gene could follow from the sexes independently approaching their fitness optima through the evolution of differential expression.

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 (e.g., promotors, enhancers), or through *trans*-acting factors (e.g., transcription factors), whose action typically impacts both target alleles. *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 (Lemos et al., 2008; Wray, 2007), which may influence their ability to respond to selection. In contrast, *trans* factors affect expression in multiple genes, and mutations in *trans* may be laden with pleiotropic side effects (Stern, 2000). Lastly, 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 & 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 & 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). Specifically, we investigate its relationship to sexual dimorphism in *D. melanogaster* in three ways.

First, we examine whether *cis*-regulatory variation is more common at 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 the some conditions, sexual conflict can generate balancing selection that stably maintains polymorphisms (Connallon & Clark, 2014; Kidwell et al., 1977). 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 at 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. Similarly, unbiased genes could either lack sexual conflict entirely or simply lack the appropriate variation to evolve sex-biased expression to mitigate conflict (Parsch & Ellegren, 2013). Based on population genetic analyses of human and fly data, Cheng & Kirkpatrick (2016) concluded that sexual conflict is the 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. Contrary to these ideas, Osada et al. (2017) report that the amount of *cis*-regulatory variation in *D. melanogaster* increases with the degree of female-bias in expression.

Regulatory variants that affect expression similarly in both sexes would not allow for the resolution of sexual conflict over expression via the evolution of dimorphism. Thus, our second goal is to examine the extent to which *cis*-regulatory effects differ between the sexes, measured as 'sex-dependent' allelic imbalance. Further, we are interested in whether such variants are more frequently found at sex-biased or unbiased genes. The most extreme form of sex-dependent 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 (Connallon & Chenoweth, 2019; Kidwell et al., 1977).

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, sex-dependent AI between these two tissues to discern if patterns in AI reflect the underlying levels of sexual dimorphism.

We find *cis*-regulatory variation to be pervasive in *D. melanogaster*, present to some degree in both sexes, across all categories of sex-biased gene expression, and in all tissues surveyed. We observe a positive relationship between the degree of female-bias in gene expression and the frequency of AI, implying that female-biased genes harbour greater *cis*-regulatory variation. We also detect numerous cases of sex-dependent *cis*-regulatory effects in all tissues, but most frequently in gonads. In addition, we uncover

evidence consistent with sex-reversed AI in 176 genes. Expression is much more dimorphic in gonads than heads and we find that this extends to *cis*-regulatory variation, with gonads exhibiting higher frequencies and magnitudes of sex-dependent AI.

Results

We crossed two inbred lines of *Drosophila melanogaster* (DGRP-177 and SP159N) and measured allele specific expression in heads, gonads, and whole bodies F_1s of both sexes. For whole body samples, we have F1s from both cross directions. A series of filters based on parental and F1 genomic data were applied to remove genes with possible mapping biases (Methods).

Apparent signals of parental effects

Studies suggest that parental effects on allelic expression rarely occur in *Drosophila melanogaster* (Chen et al., 2015; Coolon et al., 2012; Wittkopp et al., 2006a). Moreover, patterns consistent with parental effects may be due 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 analysed the data from each sex separately, and also performed an analysis using combined data from the sexes. 355 out of 7727 (4.6%) genes tested in females and 838 out of 8827 (9.5%; X and Y genes excluded) genes tested in males showed a significant effect of cross direction on allelic imbalance ($p < 0.05$). (The substantially higher incidence of "parental effects" in F_1 males relative to females could be due to males of the two crosses having considerable *trans*-regulatory differences attributable to complementary XY combinations.) In the combined data from both sexes, 928 of 7694 (12.1%; X and Y genes excluded) genes had a significant effect of cross direction. A total of 10568 genes were analyzed across these three tests, with 1752 (16.6%) 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 of these genes showing significant parental effects from further analysis.

Sex-dependent effects on allelic imbalance

Allelic imbalance appears to be pervasive in the *D. melanogaster* transcriptome, with more than 40% 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 or gene categories, 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 the distribution of p-values within a group, which could cause genes with similar data in different groups to have different q-values in each group.

In each tissue, numerous genes exhibit significant sex-dependent AI (Table 1). More than 25% of genes exhibited significant sex-dependent AI in gonads, the most sexually dimorphic tissue in our dataset. In contrast, only 10% of genes show sex-dependent AI in heads. We considered the possibility that our inferences of AI or sex-dependent 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. Firstly, 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 zero (Figures S2 & S3; Tables S4 & S5), indicating that mapping bias is unlikely to be a widespread problem among genes inferred to have AI. Secondly, we would expect a significant positive correlation in the sex difference in fraction of DGRP-177 reads between genomic and transcriptomic datasets. These correlations are also close to zero, remaining so even as we consider only genes with significant sex-dependent AI (Figure S4; Table S6). Thus, our inferences of AI or sex-dependent AI are unlikely to be substantially marred by mapping bias.

Table 1: Frequency of AI and sex-dependent AI.

Tissue	Number of genes tested	% genes with AI	% genes with sex-dependent AI
gonads	2547	41.6%	25.3%
heads	4062	41.6%	10.1%
whole bodies (main)	3627	45.3%	19.6%
whole bodies (reciprocal)	3901	41.7%	18.3%

Genes with AI are those with a significant ($p < 0.05$) intercept or 'sex' term. Genes with sex-dependent AI are the subset with a significant 'sex' term.

We examined how the patterns of AI and sex-dependent AI vary with the degree of sex bias in expression, estimated as the $\log_2$ fold change in male and female expression ($\log_2$FC). The likelihood of AI showed a significant negative relationship with sex-bias, i.e., frequency of AI decreases going from female- to male-biased genes for all tissues (Figure 1; Table 2). (The one bar deviant from this pattern in Figure 1—highly male-biased genes in heads—is based on a much smaller number of genes than other categories.) However, there exists no consistent linear relationship between sex-bias and the likelihood of sex-dependent AI in any tissue (Table S8). There was a significant negative quadratic effect of sex-bias on the likelihood of sex-dependent AI in all tissues, barring the heads (Table 2), indicating that genes with unbiased expression are more likely to have sex-dependent AI than male- and female-biased genes in both the gonads and whole bodies.

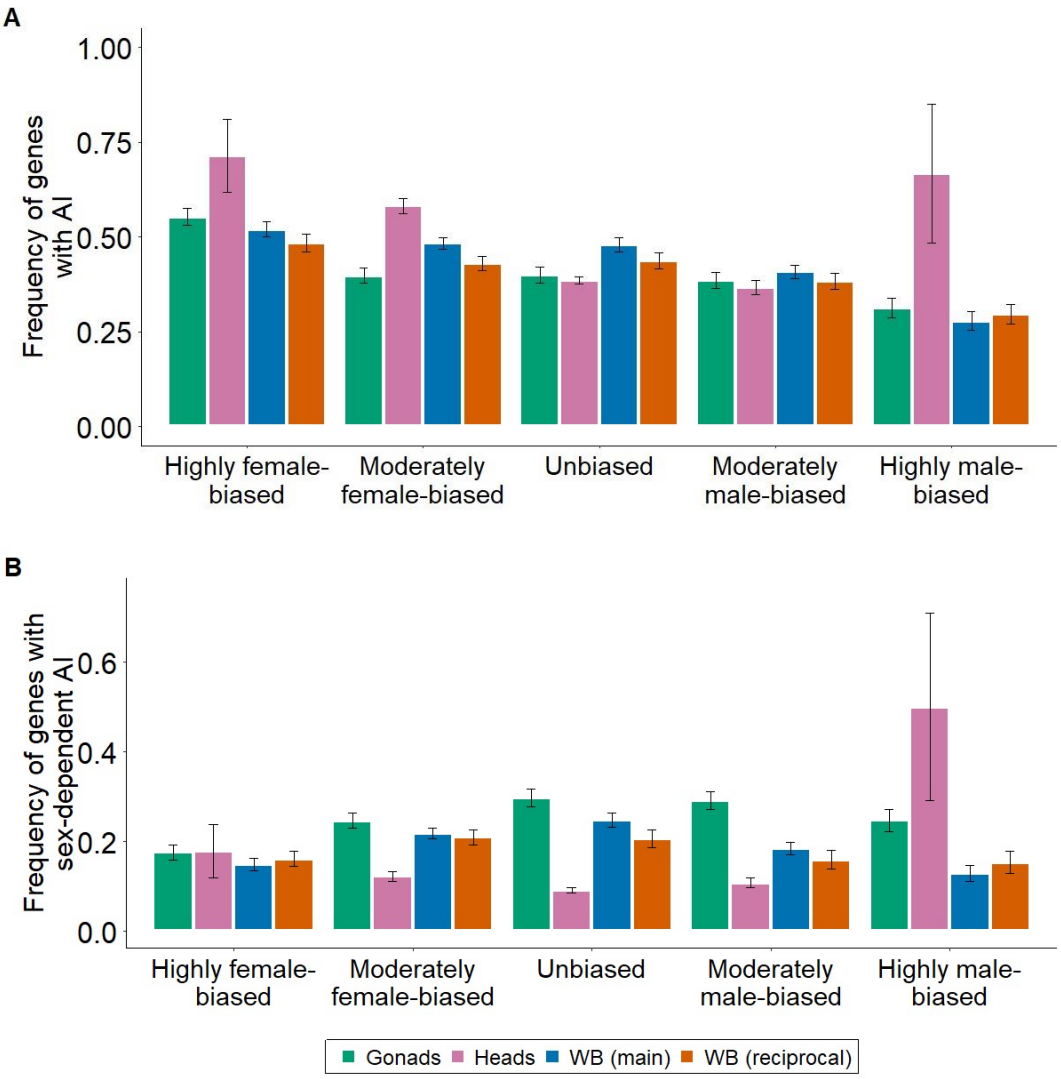

Figure 1. Frequency of genes with significant (A) allelic imbalance (AI) and (B) sex-dependent AI in each sex-bias bin. Error bars represent 95% confidence intervals. Note there are few highly male-biased genes in the heads (leading to the large error bars). WB = whole bodies.

Table 2. Results from logistic regressions modelling the effect of sex-bias (log_2FC) on the likelihood of AI or sex-dependent AI.

Tissue	Term	Effect on response variable	
		AI	Sex-dependent AI
gonads	Log_2FC	-0.160***	0.047
	$(Log_2FC)^2$	—	-0.058***
heads	Log_2FC	-0.381***	0.123
	$(Log_2FC)^2$	—	0.062
whole body (main cross)	Log_2FC	-0.129***	-0.030
	$(Log_2FC)^2$	—	-0.055***
whole body (reciprocal cross)	Log_2FC	-0.088***	-0.001
	$(Log_2FC)^2$	—	-0.017*

For analysis of sex-dependent AI, a quadratic term was included in the models. Significance is indicated as follows: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The models included a covariate (not shown) controlling for the power to detect AI effects.

The observations reported so far center around the frequencies of AI and sex-dependent AI. We examine if these patterns in frequency are reflected in the distribution of the magnitude of AI and sex-dependence in AI. We measure the magnitude of AI as $|0.5 - f|$, where f is the proportion of DGRP-177 reads as estimated by the model (Supp. Materials, Section 4 & 5). Because we are using absolute value, estimation error alone will cause an upward bias for genes where the true magnitude is zero. The measure will be more reliable (i.e., less upwardly biased) for genes with statistically significant AI. Bearing this caveat in mind, the magnitude of AI tends to be low for the most part; the vast majority of genes have magnitudes below 0.1 (Figure S8). The magnitude of AI averages around 0.04 for all genes across all tissues, remaining near 0.05 even when only considering genes with significant AI (Figure S9B).

As noted earlier, the frequency of AI has a marked negative relationship with the sex bias in all tissues (Figure 1, Table 2). Sex-bias has a negative effect on the magnitude of AI in gonads and heads, but there is a positive effect in whole bodies (Figure S9, Table S9). This discrepancy between pattern observed for frequency and magnitude of AI in whole bodies suggests that the patterns observed with respect to frequencies are not driven solely by the power to detect AI as a function of its magnitude. In whole bodies, AI is more common among female-biased than unbiased genes but, conditional on AI being detected, the magnitude of AI is not larger among female-biased than unbiased genes.

We measure the magnitude of sex difference in AI as the absolute difference in AI between the sexes. Again, the use of the absolute value will cause an upward bias for genes where the true value is zero but magnitude estimates for sex differences in AI should be more meaningful for genes with significant sex-dependent AI. We find the magnitude of sex difference in AI to be low for most genes, but generally higher in gonads than heads (Figure 2A). Additionally, the average sex difference in magnitude of AI increases when we narrow our gene set to only those with significant AI, and further, to only genes with significant sex-dependence in AI (Figures 2A, S11). This trend is especially marked in gonads. In whole bodies and heads, the magnitude of the sex difference in AI tends to be lower for unbiased or weakly-biased genes compared to more strongly sex-biased genes (Fig S11; Table S10). In gonads, there is a slight decline in magnitude of the sex difference in AI going from female- to male-biased genes; however, this pattern is clearest when considering only those genes with significant sex-dependent AI (Fig S11C).

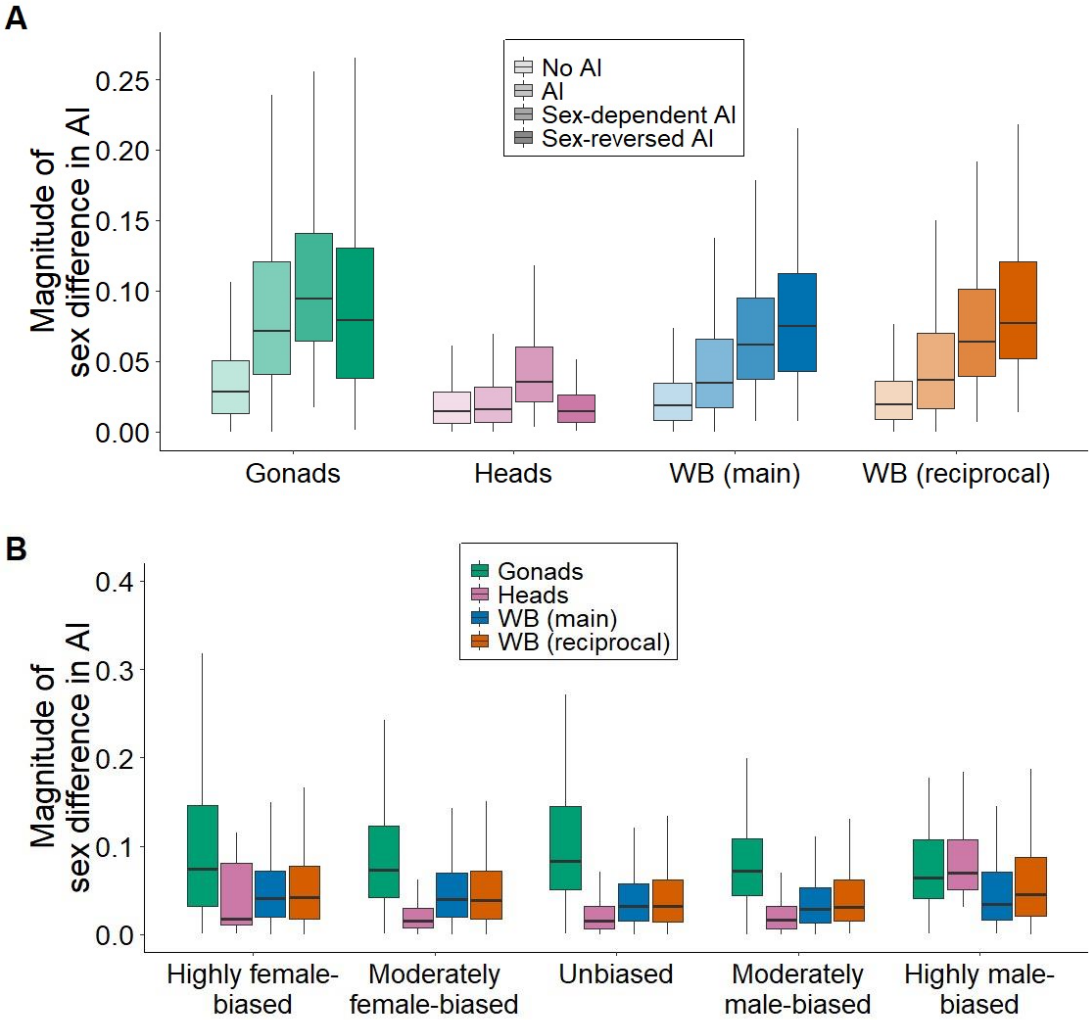

Figure 2: (A) Magnitude of the sex difference in AI with respect to statistical category (B) Magnitude of sex difference in AI stratified by sex-bias for genes with significant AI. The magnitude of the sex difference in AI is measured as $|f_{male} - f_{female}|$ where f_i is the fraction of expression from the DGRP-177 allele in sex i. WB = whole bodies.

<u>Sex-dependent reversals of allelic imbalance</u>

Instances of sex-dependent allelic imbalance 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%). To detect such cases of sex-reversed AI, we examined whole body samples, combining data from both main and reciprocal crosses. Of the 3796 genes tested in this fashion, 176 (4.6%) genes showed significantly sex-reversed AI. Using relatively conservative assumptions, we expect to find ~26 cases by chance (i.e., false positives; Supp. Materials, section 13).

We further examined these 176 genes using the head and gonad data to see if 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; in contrast, 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, a small minority of only 35 out of the 160 genes tested had sex-reversed point estimates of AI, and roughly half (only 18) of these were in the same direction as in the whole bodies, with the remaining half (17) in the opposite direction relative to whole bodies, i.e., not different form that expected by chance.

Genes with sex-reversed AI would be expected to have higher sex differences in the magnitude of AI relative to other genes with sex-dependent AI. Comparing the 176 genes with sex-reversed AI to other classes of genes, we observe this trend in the whole bodies (Figure 2A). Even so, for these 176 genes, the sex difference in the magnitude of AI averages ~7.5%, indicating a lack of drastic sex differences in allelic expression on average.

Tissue-dependent patterns of allelic imbalance

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 tested for tissue-dependent effects on AI using data from heads and gonads. We first test for AI and tissue-dependent AI separately in each sex. Approximately 50% of the genes tested in males had significant AI, while females exhibited AI in ~35% of genes tested. Males also had a higher proportion of genes with significant tissue-dependent AI than females (Table 3).

Table 3. AI differences between heads and gonads.

Sample	Number of genes tested	% of genes with AI	% of genes with tissue-dependent AI	% of genes with sex differences in tissue-dependent AI
Males (no X/Y genes)	3222	51.1%	36.1%	–
Females	2601	35.4%	21.0%	–
Combined	2418	60.8%	44.1%	25.9%

Analyses were conducted either by sex or using both sexes ("combined"). Genes with AI are those with any significant ($p < 0.05$) term. Genes with tissue-dependent AI are the subset with a significant 'tissue' term (or 'tissue-×-sex' term for combined analysis). Genes with sex differences in tissue-dependent in AI are the subset in the combined analysis with a significant 'tissue-×-sex' term.

In the analysis above, we examined the sexes separately to focus exclusively on tissue-dependency without the complicating factor of sex. It also maximizes the number of genes that could be included (as many genes lack sufficient expression to be analyzed in both sexes). However, Table 4 show an obvious discrepancy between the sexes: both AI and tissue-dependent AI occur more frequently in males than females. To examine this sex difference more formally, we analyzed the data for both sexes in a single model, using genes with sufficient data in both the sexes (n = 2418). This allowed us to discern both tissue- and sex-dependent AI, as well as tissue-by-sex interaction effects on AI. Over 60% of genes tested showed significant AI, with ~26% (n = 627) exhibiting a significant sex-by-tissue interaction on AI (Table 4). The latter class of genes have tissue-dependent AI that varies in magnitude and/or direction between the sexes. Focussing on these 627 genes, we find that although the magnitude of AI is reasonably similar across the sexes in the heads, it is notably dissimilar in the gonads, with females having smaller magnitudes of AI in gonads than males (Fig. S15). As a consequence, the magnitude of tissue differences in AI is significantly stronger in males than females (mean = -0.2, $p < 0.001$; Fig. S16). Additionally, the tissue-averaged magnitude of AI is marginally higher in males than females (mean = -0.01, $p < 0.001$; Fig S17). Taken together, these observations indicate that the discrepant patterns of AI across tissues and sexes are largely driven by sex differences in AI in the gonads.

Additionally, we test females and males separately for cases of tissue-reversed allelic imbalance. We find 33 instances of possible tissue-reversed AI between female gonads and heads, and 94 such instances in males. We estimate that 11 of these genes are likely false positives in females and 9 in males (Supp. Materials, section 12). The much higher number of putative tissue-reversed AI in males, again, 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 utilise a divergent cross between North American (DGRP-177) and South African (SP159N) genotypes to maximize the number of polymorphic sites in the F1, thus increasing our power to detect allele-specific expression. To some extent, the patterns of allelic imbalance observed here may reflect inter-population differences in expression or may be idiosyncratic to the specific lines used here. Nonetheless, some findings of ours are supported by results from studies utilizing different experimental designs. As with any study of genetic variation, our results will be influenced by the interplay between mutation and selection. The observations reported herein depend on the capacity of mutation to generate regulatory variation. Some regulatory variants may be eliminated through purifying selection, or, in contrast, be stably maintained by balancing selection. 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, divergent selection among populations) in causing these patterns. With this caveat, we interpret our results within the limits of our study design.

About 40% of genes analysed in the F_1 heads, gonads and whole bodies showed AI, pointing to considerable *cis*-regulatory divergence between the two parental lines. In all the tissues, we find a relative enrichment of *cis*-regulatory variation in female-biased genes. This accords with Osada et al.'s (2017) analysis of allele-specific expression in *Drosophila*; despite their design focusing on within-population variation among multiple DGRP lines, they too reported greater *cis*- (and *trans*-) regulatory variation with increasing female-bias. The higher frequency of AI among female-biased genes may suggest that these genes are more frequently under sexual conflict, leading to maintenance of *cis*-regulatory variants. However, this interpretation appears to contradict previous findings in humans and *Drosophila*, where the intensity of conflict was reported to be highest in genes with moderate male- or female-bias (Cheng & Kirkpatrick, 2016). Alternatively, the greater frequency of *cis*-regulatory variants in female-biased genes may result from variation in the strength of purifying selection (i.e., female-biased genes may experience weaker purifying selection on expression level). Genes with *cis*-regulatory variation have previously been shown to be subject to weaker purifying selection in *Capsella* (Steige et al., 2017), but it is unclear if this further varies with sex bias. Mutational bias could also offer an alternate explanation; female-biased genes could carry greater scope for *cis*-regulatory mutations (e.g., if such genes tend to have larger or more complex regulatory regions). However, we cannot ascertain which of these possibilities explain the observed pattern.

Notably, the *cis* effect is small in magnitude in most genes where we observe a significant allelic imbalance, with a large majority of genes exhibiting allelic expression differences below 10%. This finding is general to all the tissues examined. 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 & 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., 2012a) 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.

A *cis*-regulatory variant can give rise to sexually dimorphic expression if it interacts with the sex-specific *trans*-regulatory environment to affect expression. We uncover many instances of sex-dependent *cis*-regulatory effects in all tissues, but with a greater frequency in gonads than in heads. Because gonads are markedly more sexually dimorphic than heads, it is likely that the cellular environment constituting the *trans*-regulatory machinery is also more dimorphic in the gonads. We speculate that this allows greater opportunity for alternative *cis*-regulatory sequences to be differentially utilised in the gonads (i.e., there is a mutational bias to generate variants that cause sex-dependent AI in the gonads more readily than in the heads).

The high frequency of sex-dependent AI suggests that there is a fair amount of scope for the evolution of sexually dimorphic expression. The evolution of sex-biased gene expression has often been attributed to sexual conflict in the literature (Connallon & Knowles, 2005; Parsch & Ellegren, 2013). This assertion rests on two ideas. First, sexual conflict results when sexes have different fitness optima for expression of a gene. Second, the evolution of sex-biased expression at such a gene allows the sexes to move closer to their respective fitness optima, alleviating sexual conflict. Thus, genes with sex-biased expression may represent instances of sexual conflict, either past or present. However, unbiased genes may also be under sexual conflict (Innocenti & Morrow, 2010), which may persist because of a lack of sex-specific regulatory variation required to mitigate such conflict. Perhaps it may be harder for mutations with sex-specific regulatory effects to arise around unbiased genes. Contrary to this, however, we find sex-dependent *cis*-regulatory variation to be fairly frequent among unbiased genes, with such sex-specific *cis*-effects being marginally more common in unbiased genes than male- or female-biased genes for all tissues except the heads. Of course, we do not know anything about sex differences in selection on those genes. Consequently, even though we detect numerous cases of sex-dependent AI, we do not know if sexual conflict occurs at those genes or, if it does, whether the effects of such variants are in the direction useful to mitigating conflict. For example, if a *cis*-regulatory variant increases expression in males but does not affect expression in females, this would not help resolve sexual conflict in an unbiased gene where only females are favoured to have increased expression.

Reversals in the direction of AI between the sexes are the most extreme form of sex-dependent AI and we detect such reversals at 176 loci. 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. Dominance reversals have been invoked in the theoretical literature to explain the maintenance of genetic variation (Curtsinger et al., 1994; Kidwell et al., 1977; reviewed in Grieshop et al. 2021). Specifically in the context of sexual conflict, dominance reversals between the sexes were posited to expand the scope for maintaining variation. However, it was 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 & Arnqvist, 2018). Sex reversals in AI potentially provide a mechanism, but it remains to be seen whether such *cis*-regulatory effects underly any cases of dominance reversals for fitness.

Further insights into the evolutionary forces shaping patterns of *cis*-regulatory variation that we have reported here might come from population genetic comparisons for genes with versus without (sex-dependent) AI. We conducted a preliminary investigation of whether genes with (sex-dependent) AI are associated with possible signatures of purifying or balancing selection. We failed to find any such

association; neither π_N/π_S nor Tajima's *D* is associated with a gene's status with respect to exhibiting AI or sex-dependent AI (Tables S13-14; Section 14, Supp. Mat.).

In this study, we focussed 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; X. C. Li & Fay, 2017), tissue (Pinter et al., 2015) and metabolic state (Shih & Fay, 2021). Of the two dimensions we examined—sex and tissue—neither stood out as harbouring dramatically more variation than the other. We were motivated to look for sex-dependent 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 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 they have been largely effective (Supp. Mat., Sections 6-9). 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 our major findings are robust, the results for individual genes would benefit from further experimental validation.

As we infer *cis*-regulatory polymorphisms from allelic imbalance (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 *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 non-randomly with respect to sex-biased gene expression as well as across tissues. Further work remains to be done to more thoroughly examine variants at individual loci as well as to understand the evolutionary forces shaping 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 (Lack et al., 2016; Pool et al., 2012). We crossed DGRP-177 males with SP159N males ("main cross"), and also performed a reciprocal cross. All flies were reared and maintained at 25°C on a 12-hr light-dark cycle with a standard cornmeal fly medium. After eclosion, virgin F_1 males and females were collected and kept in separate vials for 3 days. Vials that were observed to have eggs laid in them were discarded, to avoid accidentally including flies that are not virgins. The F_1 individuals from the main cross (DGRP-177 males × SP159N females) were then 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 anaesthetised whole flies, performed over ice, with the flies placed in PBS buffer Each replicate for gonads consisted of 10-12 individuals, while each replicate for whole flies and heads consisted of 5-7 individuals. Heads and whole flies were flash-frozen in liquid nitrogen and then stored at -80°C as they awaited RNA extraction; gonads were directly stored at -80°C following dissection. RNA extraction was performed using 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 F1 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 single nucleotide polymorphisms (SNPs) identified for the genotype of interest. The use of genotype-specific references for read-mapping has been reported to reduce mapping bias when analysing allele-specific

expression (Graze et al., 2012b; Rozowsky et al., 2011). We called variants (SNPs and indels) for both DGRP-177 and SP159N using the GATK Best Practices workflow (Auwera et al., 2013; McKenna et al., 2010). 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 (H. Li & Durbin, 2009) with default parameters. The resulting alignment file was processed using GATK tools to sort reads and mark PCR/optical duplicates. Variants were called by using GATK HaplotypeCaller and the resulting 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 N's. 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 (Table S1). 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 & 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 are sorted and indexed by read names. We then compare the alignment score (AS) and number of mismatches (nM) of each read when aligned to the two references. Parental origin is ascribed to a read under the following rules: a) if a read has a higher AS when aligned to a given parent's genotype-specific reference, it is ascribed as belonging to that genotype, b) if the read has equal AS for both alignments, the read is assigned to the genotype that yielded a lower nM, and c) if AS and nM do not differ between the two alignments for a read, it is 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 to competitively align reads; results reported use both nM and AS). Ambiguous reads were excluded from the analysis, as we are only interested in reads that can be unambiguously assigned to either parental reference genomes. The percentage of such informative (i.e., non-ambiguous) reads ranged from 18% to 43% for the samples examined (Table S2). Competitive read mapping yielded two 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.

Mapping bias

Bias in mapping can lead to errors in estimation of allele-specific expression (Degner et al., 2009). Mapping bias can stem from multiple technical reasons, including sequence differences between the reads and reference genome, missing or false SNP calls and peculiarities in the alignment algorithm (L. G. León-Novelo et al., 2014). DNA controls were used to assess mapping bias in genes, and filter problematic genes from analysis. If mapping bias exists in a particular gene, paternal/maternal proportion of DNA read counts from F_1 heterozygotes will deviate from 50%. Our competitive mapping procedure was applied to genomic reads of F_1 heterozygotes (i.e., DNA controls). We aligned DNA to the genotype-specific references using BWA (H. Li & Durbin, 2009).

For each gene with a total read count of >30, we calculate the binomial probability of the observed proportion of allele-specific read counts under the null expectation of 50% of each parental allele. (Genes with a total read count ≤ 30 in either male or female genomic data set were excluded from all analyses.) A gene was considered "suspicious" (i.e., potentially subject to mapping bias) if the observed proportion of parental reads fell in either 5% tail of the binomial distribution. We separately evaluated three genomic data sets: male genomic data, female genomic data, and combined data from both sexes. Using this criterion, we expect 10% of the genes in each data set to be identified as "suspicious" by chance (i.e., false positives). In each data set, we identify more than 10% of genes tested as suspicious (males: 12%; females: 13%; combined data: 14%). Genes that were identified as suspicious in any of these three data sets (3102 out of 12045 genes = 25%) were excluded from further analysis. However, most genes had allelic coverage close to the expected 1:1 ratio (Figure S1). We also excluded genes where the estimated DGRP-177 coverage deviated from the expected frequency of DGRP-177 of 0.5 by more than 0.1, in either males or females. An additional 197 genes were excluded as a consequence of this criterion.

Despite these filters that aim to address mapping bias, it is possible that mapping bias still occurs for some of the retained genes. The list of genes identified with AI should be regarded only as candidates for genes with AI, rather than definitively being such. As an additional check, we examined the correlation between proportion of DGRP-177 in genomic and transcriptomic reads. A positive correlation is expected if mapping bias is falsely being interpreted as allelic imbalance. That correlation was near zero when including all genes as well as when considering only genes with evidence of AI (Figures S2 & S3; Tables S4 & S5). In contrast, these correlations were larger and significantly different from zero when including only those genes that exhibited mapping bias (Table S3). The sharp reduction in correlations pos t-filtering suggests that mapping bias is unlikely to be a major problem across the data set but does not preclude mapping bias affecting a relatively small set of genes.

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 greater than 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 ($0.5 < \log_2\text{FC} \leq 2$), highly male-biased ($2 < \log_2\text{FC} < \infty$). Genes with relatively high uncertainty (i.e., standard error for $\log_2\text{FC} > 1.5$) were not assigned to any of these bins . Note that four tissue types—heads, gonads, whole bodies (main cross) and whole bodies (reciprocal cross) —were separately analysed for sex-biased expression. The assignment of genes to bins was primarily for graphical presentation of results. As described below, the major patterns with respect to sex-bias were examined using statistical models that treated $\log_2$FC as a continuous variable.

Parental effects in AI

Parental effects on allele-specific gene expression in *Drosophila* are believed to be rare (Coolon et al., 2012; Wittkopp et al., 2006b). 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 generalised linear models to the whole body samples from the main and reciprocal crosses utilising the R package *lme4* v.1 (Bates et al. 2007). 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 analysed the proportion of reads assigned to the DGRP-177 genome with a quasibinomial model, including cross direction (main versus reciprocal) as a fixed effect. The intercept term of this model indicates allelic imbalance in the absence of an effect of cross direction. Additionally, we analysed 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 allelic imbalance and sex effects in allelic imbalance

For our main analyses examining AI and sex differences in AI (hereafter "sex-dependent AI"), the data for each "tissue" type (gonads, heads, whole bodies) were analysed separately. Whole bodies from the main and reciprocal crosses were also analysed separately for two 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, by analysing the cross directions separately, the whole body results are more comparable to results for gonads and heads in that all the analyses are then 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 analysed the proportion of reads assigned to the DGRP-177

genome with a quasibinomial model, including sex as a fixed effect. In this model, the "sex" effect indicates sex-dependent AI and the intercept term indicates AI even in the absence of sex differences in AI. A gene was considered to have evidence of AI if either term was significant ($p < 0.05$).

We were interested in the extent to which the occurrence of AI varies among genes with different degrees of sex bias, e.g., AI could be more common among sex-biased than unbiased genes, or female-biased genes could harbour greater AI than male-biased genes, etc. We applied a logistic regression to discern the effects of sex-bias (measured as $\log_2$FC) on the likelihood of detecting AI or sex-dependent AI. To control for potential effects of gene expression levels in the power of detecting AI or sex-dependent AI, we included a term for "total expression" in the regression analysis; total expression was represented by the natural logarithm of average read counts over the six male and female replicates. Though we label the term "total expression" for simplicity, the primary purpose for its inclusion is controlling for the power to detect AI, which depends on read count. (True total expression would depend on factors other than our read count values, including the fraction of reads that are unambiguously assigned to either parent as well as gene length.) Visual inspection of the data suggested the presence of a non-linear relationship between sex-bias and the incidence of sex-dependent AI, hence in the analysis of sex-dependent AI the logistic regressions included a quadratic term for $\log_2$FC.

<u>Inference of tissue effects on allelic imbalance</u>

We performed analyses contrasting heads and gonads, parallel to our analyses of the two sexes. Separately for each sex, we first analysed the gonad and head data with a quasibinomial model including a tissue term. A significant tissue term ($p < 0.05$) indicates tissue-dependent AI; significance of either the tissue or intercept term is taken as evidence of 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 was 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.

<u>Inference of sex- and tissue-reversed allelic imbalance</u>

The strongest form of sex-dependent AI—sex-reversed AI — occurs when one allele is responsible for greater than 50% expression in males, while the other allele is more highly expressed in females. To maximize our power, we jointly analysed 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 were 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 (Supp. Materials, Section 13).

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 were estimated in a manner analogous to that for sex-reversed AI (Supp. Materials, Section 13).

Acknowledgements

We thank Ina Anreiter for help with dissections, Baharul Choudhury for 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. This work was supported by the Natural Sciences and Engineering Research Council of Canada (AFA).

References

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

Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., Banks, E., Garimella, K. V., Altshuler, D., Gabriel, S., & DePristo, M. A. (2013). From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. *Current Protocols in Bioinformatics*, *43*(1). https://doi.org/10.1002/0471250953.bi1110s43

Barson, N. J., Aykanat, T., Hindar, K., Baranski, M., Bolstad, G. H., Fiske, P., Jacq, C., Jensen, A. J., Johnston, S. E., Karlsson, S., Kent, M., Moen, T., Niemelä, E., Nome, T., Næsje, T. F., Orell, P., Romakkaniemi, A., Sægrov, H., Urdal, K., … Primmer, C. R. (2015). Sex-dependent dominance at a single locus maintains variation in age at maturity in salmon. *Nature*, *528*(7582), 405–408. https://doi.org/10.1038/nature16062

Bonduriansky, R., & Chenoweth, S. F. (2009). Intralocus sexual conflict. *Trends in Ecology & Evolution*, *24*(5), 280–288. https://doi.org/10.1016/j.tree.2008.12.005

Campbell, C. D., Kirby, A., Nemesh, J., Daly, M. J., & Hirschhorn, J. N. (2008). A survey of allelic imbalance in F1 mice. *Genome Research*, *18*(4), 555–563. https://doi.org/10.1101/gr.068692.107

Chen, J., Nolte, V., & Schlötterer, C. (2015). Temperature Stress Mediates Decanalization and Dominance of Gene Expression in Drosophila melanogaster. *PLOS Genetics*, *11*(2), e1004883. https://doi.org/10.1371/journal.pgen.1004883

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

Connallon, T., & Chenoweth, S. F. (2019). Dominance reversals and the maintenance of genetic variation for fitness. *PLOS Biology*, *17*(1), e3000118. https://doi.org/10.1371/journal.pbio.3000118

Connallon, T., & Clark, A. G. (2011). Association between Sex-Biased Gene Expression and Mutations with Sex-Specific Phenotypic Consequences in Drosophila. *Genome Biology and Evolution*, *3*, 151–155. https://doi.org/10.1093/gbe/evr004

Connallon, T., & Clark, A. G. (2014). Balancing Selection in Species with Separate Sexes: Insights from Fisher's Geometric Model. *Genetics*, *197*(3), 991–1006. https://doi.org/10.1534/genetics.114.165605

Connallon, T., & Knowles, L. L. (2005). Intergenomic conflict revealed by patterns of sex-biased gene expression. *Trends in Genetics*, *21*(9), 495–499. https://doi.org/10.1016/j.tig.2005.07.006

Coolon, J. D., Stevenson, K. R., McManus, C. J., Graveley, B. R., & Wittkopp, P. J. (2012). Genomic Imprinting Absent in Drosophila melanogaster Adult Females. *Cell Reports*, *2*(1), 69–75. https://doi.org/10.1016/j.celrep.2012.06.013

Crowley, J. J., Zhabotynsky, V., Sun, W., Huang, S., Pakatci, I. K., Kim, Y., Wang, J. R., Morgan, A. P., Calaway, J. D., Aylor, D. L., Yun, Z., Bell, T. A., Buus, R. J., Calaway, M. E., Didion, J. P., Gooch, T. J., Hansen, S. D., Robinson, N. N., Shaw, G. D., … Pardo-Manuel de Villena, F. (2015). Analyses of allele-specific gene expression in highly divergent mouse crosses identifies pervasive allelic imbalance. *Nature Genetics*, *47*(4), 353–360. https://doi.org/10.1038/ng.3222

Cubillos, F. A., Stegle, O., Grondin, C., Canut, M., Tisné, S., Gy, I., & Loudet, O. (2014). Extensive *cis* -Regulatory Variation Robust to Environmental Perturbation in *Arabidopsis*. *The Plant Cell*, *26*(11), 4298–4310. https://doi.org/10.1105/tpc.114.130310

Curtsinger, J. W., Service, P. M., & Prout, T. (1994). Antagonistic Pleiotropy, Reversal of Dominance, and Genetic Polymorphism. *The American Naturalist*, *144*(2), 210–228. https://doi.org/10.1086/285671

Degner, J. F., Marioni, J. C., Pai, A. A., Pickrell, J. K., Nkadori, E., Gilad, Y., & Pritchard, J. K. (2009). Effect of read-mapping biases on detecting allele-specific expression from RNA-sequencing data. *Bioinformatics*, *25*(24), 3207–3212. https://doi.org/10.1093/bioinformatics/btp579

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*(1), 15–21. https://doi.org/10.1093/bioinformatics/bts635

dos Santos, G., Schroeder, A. J., Goodman, J. L., Strelets, V. B., Crosby, M. A., Thurmond, J., Emmert, D. B., Gelbart, W. M., & the FlyBase Consortium. (2015). FlyBase: Introduction of the Drosophila melanogaster Release 6 reference genome assembly and large-scale migration of genome annotations. *Nucleic Acids Research*, *43*(D1), D690–D697. https://doi.org/10.1093/nar/gku1099

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

Genissel, A., McIntyre, L. M., Wayne, M. L., & Nuzhdin, S. V. (2007). Cis and Trans Regulatory Effects Contribute to Natural Variation in Transcriptome of Drosophila melanogaster. *Molecular Biology and Evolution*, *25*(1), 101–110. https://doi.org/10.1093/molbev/msm247

Goncalves, A., Leigh-Brown, S., Thybert, D., Stefflova, K., Turro, E., Flicek, P., Brazma, A., Odom, D. T., & Marioni, J. C. (2012). Extensive compensatory *cis-trans* regulation in the evolution of mouse gene expression. *Genome Research*, *22*(12), 2376–2384. https://doi.org/10.1101/gr.142281.112

Graze, R. M., Novelo, L. L., Amin, V., Fear, J. M., Casella, G., Nuzhdin, S. V., & McIntyre, L. M. (2012a). Allelic Imbalance in Drosophila Hybrid Heads: Exons, Isoforms, and Evolution. *Molecular Biology and Evolution*, *29*(6), 1521–1532. https://doi.org/10.1093/molbev/msr318

Graze, R. M., Novelo, L. L., Amin, V., Fear, J. M., Casella, G., Nuzhdin, S. V., & McIntyre, L. M. (2012b). Allelic Imbalance in Drosophila Hybrid Heads: Exons, Isoforms, and Evolution. *Molecular Biology and Evolution*, *29*(6), 1521–1532. https://doi.org/10.1093/molbev/msr318

Grieshop, K., & Arnqvist, G. (2018). Sex-specific dominance reversal of genetic variation for fitness. *PLOS Biology*, *16*(12), e2006810. https://doi.org/10.1371/journal.pbio.2006810

Grieshop, K., Ho, E. K., & Kasimatis, K. R. (2021). Dominance reversals, antagonistic pleiotropy, and the maintenance of genetic variation. *arXiv preprint arXiv:2109.01571*.

Gruber, J. D., & Long, A. D. (2009). *Cis* -regulatory Variation Is Typically Polyallelic in Drosophila. *Genetics*, *181*(2), 661–670. https://doi.org/10.1534/genetics.108.098459

Houle, D., & Cheng, C. (2021). Predicting the Evolution of Sexual Dimorphism in Gene Expression. *Molecular Biology and Evolution*, *38*(5), 1847–1859. https://doi.org/10.1093/molbev/msaa329

Innocenti, P., & Morrow, E. H. (2010). The Sexually Antagonistic Genes of Drosophila melanogaster. *PLoS Biology*, *8*(3), e1000335. https://doi.org/10.1371/journal.pbio.1000335

Kidwell, J. F., Clegg, M. T., Stewart, F. M., & Prout, T. (1977). Regions of Stable Equilibria for Models of Differential Selection in the Two Sexes under Random Mating. *Genetics*, *85*(1), 171–183. https://doi.org/10.1093/genetics/85.1.171

Kita, R., Venkataram, S., Zhou, Y., & Fraser, H. B. (2017). High-resolution mapping of *cis* -regulatory variation in budding yeast. *Proceedings of the National Academy of Sciences*, *114*(50), E10736–E10744. https://doi.org/10.1073/pnas.1717421114

Lack, J. B., Lange, J. D., Tang, A. D., Corbett-Detig, R. B., & Pool, J. E. (2016). A Thousand Fly Genomes: An Expanded *Drosophila* Genome Nexus. *Molecular Biology and Evolution*, *33*(12), 3308–3313. https://doi.org/10.1093/molbev/msw195

Lemos, B., Araripe, L. O., Fontanillas, P., & Hartl, D. L. (2008). Dominance and the evolutionary accumulation of cis- and trans-effects on gene expression. *Proceedings of the National Academy of Sciences*, *105*(38), 14471–14476. https://doi.org/10.1073/pnas.0805160105

León-Novelo, L. G., McIntyre, L. M., Fear, J. M., & Graze, R. M. (2014). A flexible Bayesian method for detecting allelic imbalance in RNA-seq data. *BMC Genomics*, *15*(1), 920. https://doi.org/10.1186/1471-2164-15-920

León-Novelo, L., Gerken, A. R., Graze, R. M., McIntyre, L. M., & Marroni, F. (2018). Direct Testing for Allele-Specific Expression Differences Between Conditions. *G3 Genes|Genomes|Genetics*, *8*(2), 447–460. https://doi.org/10.1534/g3.117.300139

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, X. C., & Fay, J. C. (2017). Cis-Regulatory Divergence in Gene Expression between Two Thermally Divergent Yeast Species. *Genome Biology and Evolution*, *9*(5), 1120–1129. https://doi.org/10.1093/gbe/evx072

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8

Mackay, T. F. C., Richards, S., Stone, E. A., Barbadilla, A., Ayroles, J. F., Zhu, D., Casillas, S., Han, Y., Magwire, M. M., Cridland, J. M., Richardson, M. F., Anholt, R. R. H., Barrón, M., Bess, C., Blankenburg, K. P., Carbone, M. A., Castellano, D., Chaboub, L., Duncan, L., … Gibbs, R. A. (2012). The Drosophila melanogaster Genetic Reference Panel. *Nature*, *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., & DePristo, M. A. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*(9), 1297–1303. https://doi.org/10.1101/gr.107524.110

Metzger, B. P. H., Wittkopp, P. J., & Coolon, Joseph. D. (2017). Evolutionary Dynamics of Regulatory Changes Underlying Gene Expression Divergence among Saccharomyces Species. *Genome Biology and Evolution*, *9*(4), 843–854. https://doi.org/10.1093/gbe/evx035

Osada, N., Miyagi, R., & Takahashi, A. (2017). *Cis*—And *Trans* -regulatory Effects on Gene Expression in a Natural Population of *Drosophila melanogaster*. *Genetics*, *206*(4), 2139–2148. https://doi.org/10.1534/genetics.117.201459

Parsch, J., & Ellegren, H. (2013). The evolutionary causes and consequences of sex-biased gene expression. *Nature Reviews Genetics*, *14*(2), 83–87. https://doi.org/10.1038/nrg3376

Pastinen, T., & Hudson, T. J. (2004). Cis-Acting Regulatory Variation in the Human Genome. *Science*, *306*(5696), 647–650. https://doi.org/10.1126/science.1101659

Pearse, D. E., Barson, N. J., Nome, T., Gao, G., Campbell, M. A., Abadía-Cardoso, A., Anderson, E. C., Rundio, D. E., Williams, T. H., Naish, K. A., Moen, T., Liu, S., Kent, M., Moser, M., Minkley, D. R., Rondeau, E. B., Brieuc, M. S. O., Sandve, S. R., Miller, M. R., … Lien, S. (2019). Sex-dependent dominance maintains migration supergene in rainbow trout. *Nature Ecology & Evolution*, *3*(12), 1731–1742. https://doi.org/10.1038/s41559-019-1044-6

Pinter, S. F., Colognori, D., Beliveau, B. J., Sadreyev, R. I., Payer, B., Yildirim, E., Wu, C., & Lee, J. T. (2015). Allelic Imbalance Is a Prevalent and Tissue-Specific Feature of the Mouse Transcriptome. *Genetics*, *200*(2), 537–549. https://doi.org/10.1534/genetics.115.176263

Pool, J. E., Corbett-Detig, R. B., Sugino, R. P., Stevens, K. A., Cardeno, C. M., Crepeau, M. W., Duchen, P., Emerson, J. J., Saelao, P., Begun, D. J., & Langley, C. H. (2012). Population Genomics of Sub-Saharan Drosophila melanogaster: African Diversity and Non-African Admixture. *PLoS Genetics*, *8*(12), e1003080. https://doi.org/10.1371/journal.pgen.1003080

Rozowsky, J., Abyzov, A., Wang, J., Alves, P., Raha, D., Harmanci, A., Leng, J., Bjornson, R., Kong, Y., Kitabayashi, N., Bhardwaj, N., Rubin, M., Snyder, M., & Gerstein, M. (2011). AlleleSeq: Analysis of allele-specific expression and binding in a network framework. *Molecular Systems Biology*, *7*(1), 522. https://doi.org/10.1038/msb.2011.54

Sánchez-Ramírez, S., & Cutter, A. D. (2021). *CompMap: An allele-specific expression read-counter based on competitive mapping* [Preprint]. Bioinformatics. https://doi.org/10.1101/2021.02.12.431019

Schärer, L., Rowe, L., & Arnqvist, G. (2012). Anisogamy, chance and the evolution of sex roles. *Trends in Ecology & Evolution*, *27*(5), 260–264. https://doi.org/10.1016/j.tree.2011.12.006

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

Steige, K. A., Laenen, B., Reimegård, J., Scofield, D. G., & Slotte, T. (2017). Genomic analysis reveals major determinants of *cis*- regulatory variation in *Capsella grandiflora*. *Proceedings of the National Academy of Sciences*, *114*(5), 1087–1092. https://doi.org/10.1073/pnas.1612561114

Stern, D. L. (2000). Perspective: Evolutionary Developmental Biology and the Problem of Variation. *Evolution*, *54*(4), 1079–1091. https://doi.org/10.1111/j.0014-3820.2000.tb00544.x

Stranger, B. E., Montgomery, S. B., Dimas, A. S., Parts, L., Stegle, O., Ingle, C. E., Sekowska, M., Smith, G. D., Evans, D., Gutierrez-Arcelus, M., Price, A., Raj, T., Nisbett, J., Nica, A. C., Beazley, C., Durbin, R., Deloukas, P., & Dermitzakis, E. T. (2012). Patterns of Cis Regulatory Variation in Diverse Human Populations. *PLoS Genetics*, *8*(4), e1002639. https://doi.org/10.1371/journal.pgen.1002639

Wang, M., Uebbing, S., & Ellegren, H. (2017). Bayesian Inference of Allele-Specific Gene Expression Indicates Abundant Cis-Regulatory Variation in Natural Flycatcher Populations. *Genome Biology and Evolution*, *9*(5), 1266–1279. https://doi.org/10.1093/gbe/evx080

Wittkopp, P. J., Haerum, B. K., & Clark, A. G. (2004). Evolutionary changes in cis and trans gene regulation. *Nature*, *430*(6995), 85–88. https://doi.org/10.1038/nature02698

Wittkopp, P. J., Haerum, B. K., & Clark, A. G. (2006a). Parent-of-Origin Effects on mRNA Expression in *Drosophila melanogaster* Not Caused by Genomic Imprinting. *Genetics*, *173*(3), 1817–1821. https://doi.org/10.1534/genetics.105.054684

Wittkopp, P. J., Haerum, B. K., & Clark, A. G. (2006b). Parent-of-Origin Effects on mRNA Expression in *Drosophila melanogaster* Not Caused by Genomic Imprinting. *Genetics*, *173*(3), 1817–1821. https://doi.org/10.1534/genetics.105.054684

Wittkopp, P. J., Haerum, B. K., & Clark, A. G. (2008). Regulatory changes underlying expression differences within and between Drosophila species. *Nature Genetics*, *40*(3), 346–350. https://doi.org/10.1038/ng.77

Wray, G. A. (2007). The evolutionary significance of cis-regulatory mutations. *Nature Reviews Genetics*, *8*(3), 206–216. https://doi.org/10.1038/nrg2063

Zhang, X., & Emerson, J. J. (2019). *Inferring the genetic architecture of expression variation from replicated high throughput allele-specific expression experiments* [Preprint]. Genetics. https://doi.org/10.1101/699074