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] 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
[3]School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ, United Kingdom

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

Keywords:
sexual dimorphism, sex biased gene expression, allelic imbalance

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

Author Contributions: PM, KG, TB performed dissections and RNA extractions. PM performed bioinformatic data processing. PM and AA did the data analysis. PM, KG, and AA did the writing of the manuscript. AA conceived the experiment.

Data availability: Upon acceptance, RNAseq data will be submitted to the Sequence Read Archive (NIH) and analysis scripts will be made available via GitHub.

Conflict of Interest: Authors have no conflicts of interest.

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.

Funding:
This work was supported by the Natural Sciences and Engineering Research Council of Canada (AFA).

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 use this approach to quantify transcriptome-wide sex differences in *cis*-regulatory effects as well as examine patterns of *cis*-regulatory variation 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 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 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 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; Wittkopp et al. 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). 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 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 to sexual dimorphism 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 (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 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 (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 & 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. 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' allelic imbalance. 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 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 (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, sex-dependent AI between these two tissues to discern if patterns in AI reflect the underlying levels of sexual dimorphism.

122 **Methods**

123 Sample preparation and sequencing

124 We crossed two inbred lines of *D. melanogaster* to obtain F_1 individuals for the study. DGRP-177
125 belongs to the Drosophila Genetic Reference Panel, a suite of genotypes derived from natural populations
126 of *D. melanogaster* in North Carolina (Mackay et al. 2012). SP159N was derived from a South African
127 population, as part of the Drosophila Population Genomics Project (Pool et al. 2012; Lack et al. 2016).
128 We crossed DGRP-177 males with SP159N females ("main cross"), and also performed a reciprocal
129 cross. All flies were reared and maintained at 25°C on a 12-hr light-dark cycle. After eclosion, virgin F_1
130 males and females were collected and kept in separate vials for 3 days. Vials that were observed to have
131 eggs laid in them were discarded. The F_1 individuals from the main cross (DGRP-177 males × SP159N
132 females) were used to obtain samples of whole body flies, heads, and gonads, with three replicate samples
133 of each sex*tissue-type combination. From the reciprocal cross, only whole body samples were collected
134 (three replicates of each sex). Testes and ovaries were obtained by dissection of anaesthetised whole flies,
135 performed over ice, with the flies placed in PBS buffer. Each replicate for gonads consisted of 10-12
136 individuals, while each replicate for whole flies and heads consisted of 5-7 individuals. Heads and whole
137 flies were flash-frozen in liquid nitrogen and then stored at -80°C as they awaited RNA extraction;
138 gonads were stored at -80°C upon dissection. RNA extraction was performed using ThermoFisher
139 PicoPure RNA Isolation Kit. Paired-end sequencing at a read length of 200 bp was performed using
140 Illumina NovaSeq6000. In addition, DNA was extracted from the parents and F_1 individuals using the
141 Qiagen DNeasy Blood and Tissue Kit. Paired-end whole genome sequences were obtained at a read
142 length of 100 bp using Illumina HiSeqX.
143

144 Genotype-specific references

145 Genotype-specific references are constructed from the reference genome by incorporating single
146 nucleotide polymorphisms (SNPs) for each parental genotype separately to reduce mapping bias when
147 analysing allele-specific expression (Rozowsky et al. 2011; Graze et al. 2012). We called variants (SNPs
148 and indels) for both DGRP-177 and SP159N using the GATK Best Practices workflow (McKenna et al.
149 2010; Auwera et al. 2013). The whole genome sequences for DGRP-177 and SP159N were aligned to the
150 *D. melanogaster* Release 6 reference genome (dos Santos et al. 2015) using BWA (Li and Durbin 2009)
151 with default parameters. The resulting alignment file was processed using GATK tools to sort reads and
152 mark PCR/optical duplicates. Variants were called by using GATK HaplotypeCaller and the resulting
153 VCF file was separated into two VCFs, one each for indels and SNPs. We obtained a BED file of
154 coordinates corresponding to heterozygous SNP calls and regions around indels for each parental genome.

155 Depending on the length of the indels, the following coordinates were included in the BED file: 10 bp
156 around indels of length ≤ 6 bp, 20 bp around indels of length > 6 bp but ≤ 12 bp, and 100 bp around indels
157 > 12 bp. The coordinates were then masked in the reference genome with N's. We masked sites without
158 high certainty of a homozygous SNP call in both parents, i.e., any sites where the major allele was < 95%.
159 We also masked sites where the depth of coverage fell below 10. Following recommendations by GATK,
160 the remaining variants—the homozygous SNPs—were subject to hard-filtering based upon quality
161 measures (Table S1). The filtered SNPs corresponding to DGRP-177 and SP159N were incorporated
162 separately into the masked reference genome, to create genotype-specific references for each parent.
163

164 Competitive mapping

165 For each sample, RNA-seq reads were aligned to both genotype-specific references using STAR v2.7
166 (Dobin et al. 2013) with default parameters. Following a similar method to Sánchez-Ramírez & Cutter
167 (2021), a Python-based competitive read mapping approach was used to obtain allele-specific read counts
168 from RNA-seq data for the F_1 heterozygotes. The two BAM alignments yielded by STAR were sorted and
169 indexed by read names. The alignment score (AS) and number of mismatches (nM) of each read when
170 aligned to the two references were used to assign each read's origin as follows: a) if a read had a higher
171 AS when aligned to a given parent's genotype-specific reference, it was ascribed as belonging to that
172 genotype, b) if the read has equal AS for both alignments, the read was assigned to the genotype that
173 yielded a lower nM, and c) if AS and nM do not differ between the two alignments for a read, it was
174 assigned to neither genotype, i.e., an "ambiguous" read. (We also tried the competitive alignment using
175 only nM to ascribe parental origin and obtained results similar to that from using both AS and nM.)
176 Ambiguous reads were excluded from further analysis. The percentage of such informative (i.e., non-
177 ambiguous) reads ranged from 18% to 43% for the samples examined (Table S2). Competitive read
178 mapping yielded two alignment files, corresponding to allele-specific reads of each parent. Finally, allele-
179 specific read counts were obtained from those BAM alignments using HTSeq-count v0.12 (Anders et al.
180 2015), with secondary or chimeric alignments being ignored when aggregating reads.
181

182 Bias in mapping can occur for multiple reasons and lead to errors in estimation of allele-specific
183 expression (Degner et al. 2009; León-Novelo et al. 2014). Genomic DNA from F_1s should have 50% of
184 reads mapping to each parent in the absence of mapping bias. As described in the Supplementary
185 Material, we used genomic DNA from male and female F_1 samples to conservatively filter out genes with
186 possible mapping bias (Supplementary Text S1, Figures S1-S3, Tables S3-S5).
187

188

189 Parental effects in AI

190 Parental effects on allele-specific gene expression in *Drosophila* are believed to be rare (Wittkopp et al.
191 2006; Coolon et al. 2012; Puixeu et al. 2023). Still, certain forms of mapping bias could manifest as
192 apparent parental effects. To detect real or apparent "parental" effects, and then remove such genes from
193 further analyses, we applied generalised linear models to the whole body samples from the main and
194 reciprocal crosses utilising the R package *lme4* v.1 (Bates et al. 2007). The first test was applied
195 separately to each sex. Only genes that had read counts >30 in all the 6 replicates were considered. For
196 each gene, we analysed the proportion of reads assigned to the DGRP-177 genome with a quasibinomial
197 model, including cross direction (main versus reciprocal) as a fixed effect. The intercept term of this
198 model indicates allelic imbalance in the absence of an effect of cross direction. Additionally, we analysed
199 the data for both sexes in a single model, including sex and "cross direction" as fixed effects, but without
200 an interaction term. This test was applied to all genes with a total read count >20 in all 6 male and female
201 replicates and average read count >30. All genes that had a significant cross-direction effect ($p < 0.05$)
202 were excluded from subsequent analyses.

203

204 Inference of allelic imbalance and sex differences in allelic imbalance

205 For our analyses examining AI and sex-depenent AI (hereafter "SD-AI"), the data for each "tissue" type
206 (gonads, heads, whole bodies) were analysed separately. Whole bodies from the main and reciprocal
207 crosses were also analysed separately for two reasons. First, as we expect similar patterns in the whole
208 body analyses from each cross direction, any major discrepancies could serve as a cause for concern.
209 Second, the whole body results are comparable to results for gonads and heads because all the analyses
210 are based on three replicates per sex for each tissue type. Only genes with at least 30 assignable reads in
211 each of the replicates were included in the analysis for a given tissue. For each gene, we analysed the
212 proportion of reads assigned to the DGRP-177 genome with a quasibinomial model, including sex as a
213 fixed effect and using sum contrasts. In this model, the intercept term indicates the sex-averaged AI effect
214 and the "sex" term indicates the difference in AI between sexes (i.e., a significant "sex" term is evidence
215 of sex-dependent AI).

216

217 Inference of tissue effects on allelic imbalance

218 We performed analyses contrasting heads and gonads, parallel to our analyses of the two sexes.
219 Separately for each sex, we first analysed the gonad and head data with a quasibinomial model including
220 a tissue term. Significance of the intercept term is taken as evidence of (tissue-averaged) AI. A significant
221 tissue term ($p < 0.05$) indicates tissue-dependent AI. Unexpectedly, we noted substantial differences in

222 the frequency of AI and tissue-dependent AI in the male and female analyses. Subsequently, we examined
223 those genes for which there was sufficient data in both sexes using a quasibinomial model that included
224 sex, tissue, and their interaction. The interaction term indicates sex differences in tissue-dependent AI.
225
226 Inference of sex- and tissue-reversed allelic imbalance
227 The strongest form of sex-dependent AI—sex-reversed AI—occurs when one allele is responsible for
228 greater than 50% expression in males, while the other allele is more highly expressed in females. To
229 maximize our power, we jointly analysed the whole body data from the main and reciprocal crosses. All
230 genes with read counts <20 in any of the replicates were excluded from analysis. Additionally, we
231 excluded genes where the average read count across all the 12 replicates was <30. We then tested each
232 sex separately for AI, using a quasibinomial model that only fit an intercept term, which signified the AI
233 effect. A gene was said to have sex-reversed AI if it fulfilled the following conditions: (i) it had a
234 significant AI effect in females, (ii) it had a significant AI effect in males, and (iii) the AI effect in the
235 sexes were in opposite directions. The number of false positives expected from this procedure depends on
236 the assumptions one makes for a null hypothesis. We considered several possibilities to arrive at a
237 conservative estimate of the number of false positive cases of dominance reversal (Supplementary Text
238 S1).
239
240 Likewise, we used data from gonads and heads to detect cases of tissue-reversed AI in males and females
241 separately. Again, after excluding genes with low read counts as above, we tested each tissue separately
242 for AI using a quasibinomial model where the intercept term represented AI. A gene was said to have
243 tissue-reversed AI if it showed significant AI in both heads and gonadal tissue, and the direction of the AI
244 was opposite between the two tissues. The number of false positive cases of tissue-reversed AI were
245 estimated in a manner analogous to that for sex-reversed AI (Supplementary Text S1).
246
247 Modeling the distribution of AI and SD-AI
248 In the earlier section, we described analyzing each gene individually to test for the presence of AI and
249 SD-AI. As with any set of frequentist analyses, the results will be subject to both false positives and false
250 negatives. Moreover, estimates of the average magnitudes of AI and SD-AI using "significant" genes will
251 be subject to bias due to the "winner's curse". As an alternative, we used a Bayesian approach to model
252 the joint distribution of AI effects across sexes, separately for each tissue.
253
254 As the true features of this distribution are unknown, there many ways one could choose to model it. Our
255 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 non-zero 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 non-zero sex-averaged AI, $F_{SD\text{-}AI}$; and (iv) the standard deviation in the 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 non-zero 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 data set 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 this (and other) filtering steps non-randomly excludes 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.

<u>Differential gene expression analysis</u>

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

290 were assigned to one of five sex-bias bins, as follows: highly female-biased ($-\infty < \log_2FC \leq -2$),
291 moderately female-biased ($-2 < \log_2FC \leq -0.5$), unbiased ($-0.5 \leq \log_2FC \leq 0.5$), moderately male-biased
292 ($0.5 < \log_2FC \leq 2$), highly male-biased ($2 < \log_2FC < \infty$). Genes with relatively high uncertainty (i.e.,
293 standard error for $\log_2FC > 1.5$) were not assigned to any of these bins . The four tissue types—heads,
294 gonads, whole bodies (main cross) and whole bodies (reciprocal cross) —were separately analysed for
295 sex-biased expression.
296

297 **Results**

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

302 Apparent signals of parental effects
303 Studies suggest that parental effects on allelic expression rarely occur in *Drosophila melanogaster*
304 (Wittkopp et al. 2006; Coolon et al. 2012; Chen et al. 2015). Moreover, patterns consistent with parental
305 effects may be due to residual mapping bias. We performed three tests using whole body samples from
306 the main and reciprocal crosses to identify genes with evidence of "parental effects," seeking to remove
307 these genes from subsequent analysis. We analyzed the data from each sex separately, and also performed
308 an analysis using combined data from the sexes. 346 out of 7716 (4.5%; X genes included) genes tested in
309 females and 829 out of 8817 (9.4%; X and Y genes excluded) genes tested in males showed a significant
310 effect of cross direction on allelic imbalance ($p < 0.05$). (The substantially higher incidence of "parental
311 effects" in F_1 males relative to females could be due to males of the two crosses having *trans*-regulatory
312 differences attributable to complementary XY combinations.) In the combined data from both sexes, 532
313 of 6371 (8.4%; X and Y genes excluded) genes had a significant effect of cross direction. A total of
314 10522 genes were analyzed across these three tests, with 1375 (13.1%) showing evidence of parental
315 effects in at least one of the tests. While many of these genes are expected to be false positives (5% of all
316 genes in each test), our primary goal here is to reduce the possibility of mapping bias influencing our
317 analyses of AI. We therefore excluded all these genes showing significant parental effects from further
318 analysis.
319

320

321 Sex-dependent effects on allelic imbalance

322 Allelic imbalance appears to be pervasive in the *D. melanogaster* transcriptome, with more than 30% of
323 the genes tested for each tissue showing significant AI (Table 1). Given our *p*-value threshold, we expect
324 ~5% of genes to falsely appear as showing significant AI (i.e., false positives); the observed fraction is
325 much greater than expected under the null hypothesis. We use a *p*-value criterion for the ease of
326 comparisons among tissues, which would be complicated when applying a false discovery rate; false
327 discovery rate *q*-values depend not only on the data from the focal genes but also the distribution of *p*-
328 values within a group, which could cause genes with similar evidence for AI to have different *q*-values in
329 each group. Nevertheless, a version of Table 1 using FDR cut-offs is shown in Table S8. As expected,
330 using the more conservative FDR requirement reduces the overall frequency of detection of AI but the
331 main patterns are unchanged.

332 In each tissue, numerous genes exhibit significant sex-dependent allelic imbalance, SD-AI (Table 1).
333 About 25% of genes exhibited significant SD-AI in gonads, the most sexually dimorphic tissue in our
334 dataset. In contrast, only 10% of genes show SD-AI in heads. We considered the possibility that our
335 inferences of AI or SD-AI may be a consequence of mapping bias, despite our efforts to minimize such
336 errors. If mapping bias persists in the dataset despite the filters, we would expect patterns in genomic data
337 to be reflected in RNA-seq. Firstly, we would expect a positive correlation in fraction of DGRP-177 reads
338 between F1 genomic and F1 transcriptomic datasets in both the sexes. However, the correlations are very
339 close to zero (Figures S2 & S3; Tables S4 & S5), indicating that mapping bias is unlikely to be a
340 widespread problem among genes inferred to have AI. Secondly, we would expect a significant positive
341 correlation in the sex difference in the fraction of DGRP-177 reads between genomic and transcriptomic
342 datasets. These correlations are also close to zero, remaining so even as we consider only genes with
343 significant sex-dependent AI (Figure S4; Table S6). Thus, our inferences of AI or SD-AI are unlikely to
344 be substantially marred by mapping bias.
345

346

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

Tissue	Number of genes tested	% genes with AI	% genes with sex-dependent AI
gonads	2547	31.6%	25.3%
heads	4062	45.0%	10.1%
whole bodies (main)	3627	36.5%	19.6%
whole bodies (reciprocal)	3901	35.7%	18.3%

347
348
349
350 Estimating the distribution of AI and SD-AI
351 The preceding section summarizes the results from assessing the statistical significance of each gene
352 individually. As such, Table 1 reflects the frequencies of genes for which there was statistical power to
353 detect evidence of AI and SD-AI. Thus, the values in Table 1 not only depend on the aspects of biology
354 that are our real interest (frequency and effect sizes of AI and SD-AI), but also the number of samples,
355 sequencing depth, and the choice of statistical threshold for significance. As an alternative to the approach
356 used to generate Table 1, we employed a Bayesian analysis to model the underlying joint distribution of
357 AI across the sexes in each tissue type. The results are summarized in Table 2. Across all tissues, the
358 inferred frequency of genes with AI was similar (F_{AI} = 38-48%) as was the magnitude of the (sex-
359 averaged) effect (~5%). However, there were major differences between tissues with respect to the
360 frequency of sex-differences in AI. In gonads, almost all genes with AI were inferred to have sex
361 differences ($F_{SD\text{-}AI}$ = 94%), whereas in heads very few were ($F_{SD\text{-}AI}$ = 5%). In gonads, the effect sizes of
362 SD-AI tend to be quite large relative to the sex-average effect size of AI. (This is related to a result shown
363 in a later section: AI effects tend to be larger in male than female gonads.) The values in Table 2
364 represent, to our knowledge, the only available estimates of the distribution of AI/SD-AI effects but, as
365 with any statistical model, some caution is warranted as these estimates are contingent on the assumptions
366 underlaying the framework of the modelled distribution.
367
368

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

	Frequency of genes with AI, F_{AI}	Average magnitude of sex-averaged AI[1], σ_{AI}	Frequency of genes with SD-AI among genes with AI, $F_{SD\text{-}AI}$	Average magnitude of sex difference in AI[2], $\sigma_{SD\text{-}AI}$	Overdispersion parameter, υ
Gonads	0.48	0.053	0.94	0.105	0.006
	[0.45, 0.52]	[0.0496, 0.056]	[0.90, 1.00]	[0.099, 0.114]	[0.005, 0.006]
Heads	0.42	0.050	0.05	0.074	0.001
	[0.40, 0.45]	[0.054, 0.060]	[0.01, 0.08]	[0.051, 0.106]	[0.001, 0.001]
Whole Body (main)	0.45	0.049	0.75	0.059	0.001
	[0.44, 0.47]	[0.048, 0.051]	[0.71, 0.79]	[0.058, 0.061]	[0.001, 0.002]
Whole Body (reciprocal)	0.38	0.055	0.66	0.065	0.003
	[0.35, 0.40]	[0.053, 0.058]	[0.60, 0.71]	[0.060, 0.070]	[0.002, 0.003]

371 [1] σ_{AI} is the standard deviation of sex-averaged effects, which is modelled as approximately
372 normally distributed with mean 0.
373 [2] $\sigma_{SD\text{-}AI}$ is the standard deviation of sex differences in AI effects, which is modelled as
374 approximately normally distributed with mean 0.
375
376
377 <u>AI in relation to sex-biased gene expression</u>
378 We examined how the patterns of AI vary with the degree of sex bias in expression by fitting a Bayesian
379 model separately to gene sets binned by sex bias. For this section, we used an alternative framework for
380 the distribution of AI effects, parameterizing the distribution in terms of the frequency with which AI
381 occurs, F_{AI} (as in the previous section), as well as the correlation of AI effects between males and
382 females, ρ_{MF}. There is heterogeneity in the frequency of AI with respect to sex bias (Fig. 1A; Tables S10-
383 12). In heads, F_{AI} declines from female-biased to unbiased to male-biased genes. There is a similar trend
384 in the whole body samples, though here it is noteworthy that the inferred average effect sizes also vary
385 with lower estimates for female- than male-biased genes (Fig. 1A; Table S10-11).
386
387
388

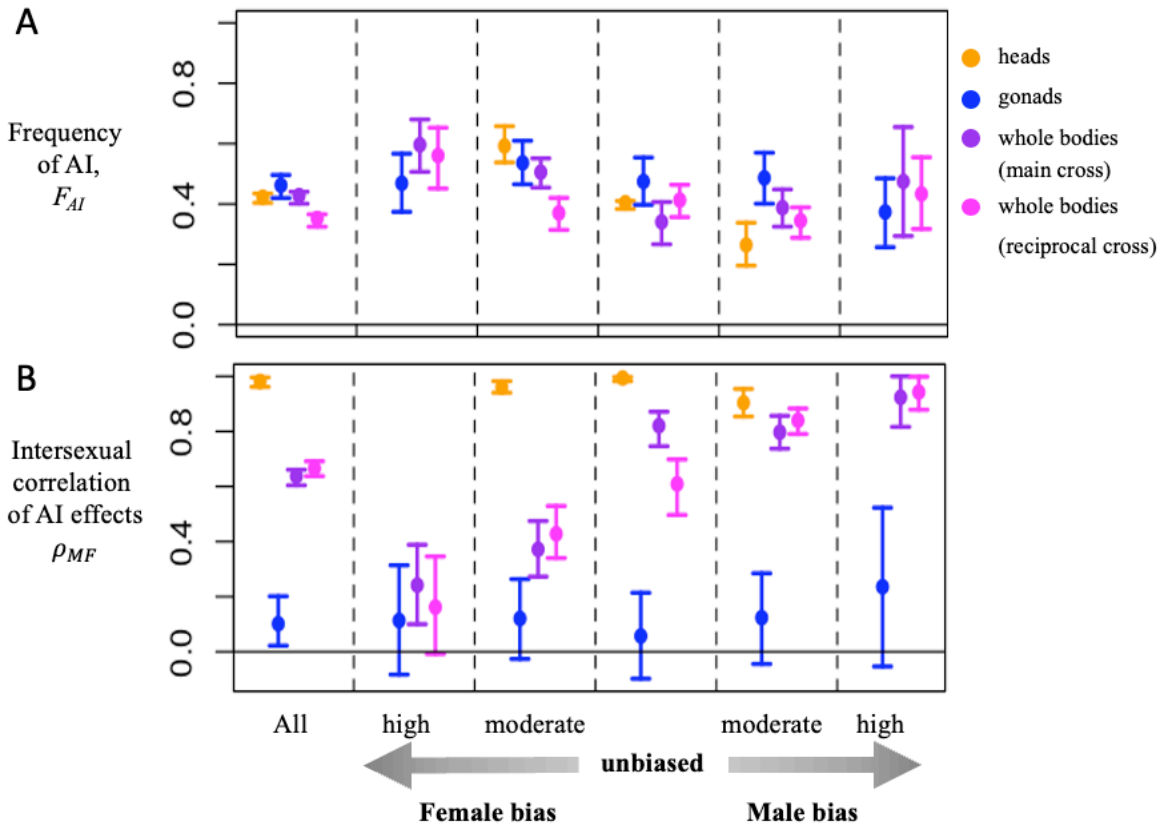

Figure 1. Model estimates of (A) the frequency of AI, F_{AI}, and (B) the intersexual correlation of AI effects, ρ_{MF}. The leftmost section in each panel shows results irrespective of sex-bias; the remaining sections show results stratified by sex bias: highly female-biased ($-\infty < \log2FC \leq -2$), moderately female-biased ($-2 < \log2FC \leq -0.5$), unbiased ($-0.5 \leq \log2FC \leq 0.5$), moderately male-biased ($0.5 < \log2FC \leq 2$), highly male-biased ($2 < \log2FC < \infty$). No analysis was performed for high sex-bias categories for heads as there were too few genes. Error bars represent 95% high density interval of the posterior. See Table S11 for other model parameters.

399

400 The intersexual correlation of AI effects is high in heads across all sex bias categories ($\rho_{MF} > 0.9$; Fig.
401 1B) whereas it is much lower in gonads ($\rho_{MF} < 0.25$). In the whole body samples, there is a striking
402 change in ρ_{MF} across sex-bias categories, going from low values for female-biased genes to high values
403 for male-biased ones. (Applying the other model parameterization to these data, a qualitatively similar
404 compatible pattern is manifest in that SD-AI is common among female-biased genes and much rarer for
405 male-biased ones; Table S10).

406

407 Sex-dependent reversals of allelic imbalance

408 Instances of sex-dependent allelic imbalance may include cases where the direction of imbalance is
409 reversed between the sexes (i.e., male expression of the DGRP-177 allele is >50%, while female
410 expression is <50%). Cases of sex-reversed AI are of special interest but are expected to be rare as this
411 scenario presumably involves more complex regulation. While the Bayesian models in the previous
412 sections are designed to make inferences about the distribution using information from a large number of
413 genes, they are not well-suited to clearly identify individual genes exhibiting sex-reversed AI. To detect
414 cases of sex-reversed AI, we analyzed genes individually, requiring significant AI effects in both sexes
415 but of opposite direction. We examined whole body samples, combining data from both main and
416 reciprocal crosses to maximize power. Of the 3796 genes tested in this fashion, 176 (4.6%) genes showed
417 significantly sex-reversed AI. Using relatively conservative assumptions, we expect to find ~26 cases by
418 chance (i.e., false positives; see Supplementary Material).

419

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

431

432 Tissue-dependent patterns of allelic imbalance

433 In our previous analyses, we noted marked differences in the frequency of AI between heads and gonads.
434 Though the analyses used different sets of genes for each tissue, the results suggest that AI may differ
435 between tissues. To investigate tissue differences in AI more formally, we applied to the Bayesian model
436 to estimate the joint distribution of AI effects in heads and gonads, separately for each sex (Table 3). In
437 both sexes, we infer that ~40% of genes exhibit AI and > 80% of these have effects that differ between
438 gonads and heads. Using the alternative Bayesian model to characterize the distribution in terms of the
439 between-tissue correlation in effects (Table S13), we find this correlation is quite low ($\rho_{head\text{-}gonad} \approx 0.25$)
440 and that the average magnitude of effect sizes is larger in gonads than heads (gonads ≈ 0.09; heads ≈
441 0.05).
442
443 **Table 3.** Parameter estimates [with 95% high posterior density interval] for the joint distribution of AI
444 effects across heads and gonads.
445

	Frequency of genes with AI, F_{AI}	Average magnitude of tissue-averaged AI[1], σ_{AI}	Frequency of genes with Tissue-Dependent AI among genes with AI, $F_{TD\text{-}AI}$	Average magnitude of tissue difference in AI[2], $\sigma_{TD\text{-}AI}$	Overdispersion parameter, $\boldsymbol{\upsilon}$
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]

446 [1] σ_{AI} is the standard deviation of sex-averaged effects, which is modelled as approximately normally
447 distributed with mean 0.
448 [2] $\sigma_{TD\text{-}AI}$ is the standard deviation of tissue differences in AI effects, which is modelled as approximately
449 normally distributed with mean 0.
450
451 Analogous to our earlier test for sex-reversed AI, we tested for tissue-reversed AI, separately in each sex.
452 We find 33 instances of possible tissue-reversed AI between gonads and heads in females, and 94 such
453 instances in males. We estimate that 11 of these genes are likely false positives in females and 9 in males
454 (see Supplementary Material). The much higher number of putative tissue-reversed AI in males may be a
455 result of tissue differences in AI being larger in males, leading to easier detection of reversals in AI.
456
457 **Discussion**
458 This study focuses on *cis*-regulatory variation in relation to sexual dimorphism in *D. melanogaster*. We
459 utilise a divergent cross between North American (DGRP-177) and South African (SP159N) genotypes to
460 maximize the number of polymorphic sites in the F_1, thus increasing our power to detect allele-specific
461 expression. To some extent, the patterns of allelic imbalance observed here may reflect inter-population
462 differences in expression or may be idiosyncratic to the specific lines used here. As with any study of

463 genetic variation, our results will be influenced by the interplay between the capacity of mutation to
464 generate regulatory variation and selection to filter it. Our goal was to explore sex-specific patterns of *cis*-
465 regulatory variation but our study was not designed to assess the relative importance of different
466 evolutionary factors (e.g., mutation-selection-drift balance, balancing selection within populations,
467 divergent selection among populations) in causing these patterns.
468
469 When analyzing individual genes, allelic imbalance was detected in 30-45% of genes analyzed (Table 1).
470 These values are considerably higher than the values (6-17%) reported by Puixeu et al. (2023). The
471 greater divergence of haplotypes used in our study may have contributed to this difference; we crossed fly
472 lines from different continents whereas they crossed lines from within a single population. However,
473 experimental and statistical differences are likely major contributors to this discrepancy. Rather than
474 relying on the proportion of genes that reach significance in individual gene tests which depend heavily
475 on statistical power, we also developed a Bayesian approach to directly estimate the frequency of genes
476 with AI, finding that the frequency is quite high (38-48%, Table 2).
477
478 In heads and whole body samples, there is an enrichment of *cis*-regulatory variants in female-biased
479 genes. Osada et al. (2017) and Puixeu et al. (2023) also examined *cis*-regulatory variation in relation to
480 sex bias. Some of the patterns reported in those two studies appear similar while others appear different
481 from ours or from each other. However, it is not possible to directly compare those studies to our own or
482 to each other because of differences in methodology use to investigate these patterns. An advantage of our
483 approach is that patterns can be investigated with respect to biologically interpretable properties such as
484 the frequency of AI and average effect sizes.
485
486 The higher frequency of AI among female-biased genes that we detected may suggest that these genes are
487 more frequently under balancing selection or tend to experience weaker purifying selection. Sexually
488 antagonistic selection is one possible reason genes could experience balancing selection or weaker
489 purifying selection. Cheng and Kirkpatrick (2016) suggested that genes with intermediate levels of sex-
490 biased expression—both male- and female-biased—would be most likely to experience sexually
491 antagonistic selection. As we do not observe an elevated frequency of AI for genes with intermediate
492 levels of sex-bias, our results would seem to indicate that either that sexually antagonistic selection does
493 not drive the pattern we observe or that the sexually antagonistic selection is not distributed across the
494 genome as Cheng and Kirkpatrick (2016) suggested. While it is most intuitive to think of sexually
495 antagonistic selection causing balancing selection or weak purifying selection, it could cause strong
496 purifying selection if selection in one sex is much stronger than the other. Thus, the suggestion of Cheng

497 and Kirkpatrick (2016) is not at necessarily at odds with our results if, for example, sexually antagonistic
498 selection results in weak purifying selection for intermediately female-biased genes but strong purifying
499 selection for intermediately male-biased genes.
500
501 Purifying selection (regardless of whether it involves sexual antagonism) is believed to be the most
502 common form of selection and genes with *cis*-regulatory variation have previously been shown to be
503 subject to weaker purifying selection in *Capsella* (Steige et al. 2017). Strongly female-biased genes in *D.*
504 *melanogaster* have somewhat elevated levels of non-synonymous diversity (Singh and Agrawal 2023; but
505 see details within), suggesting they may tend to be under weaker purifying selection and providing a
506 potential explanation for the increased frequency of AI among these genes. A non-selective alternative
507 explanation for our observed pattern is mutational bias; female-biased genes could have greater scope for
508 *cis*-regulatory mutations (e.g., if such genes tend to have larger or more complex regulatory regions). We
509 cannot ascertain which of these possibilities explain the observed pattern.
510
511 In general, *cis* effects are small in magnitude, with a large majority of genes exhibiting allelic expression
512 differences below 5-10%. This finding is general to all the tissues examined, though average effects were
513 notably higher in gonads than heads (Table S13). Others have also found frequent, weak *cis*-regulatory
514 effects in other species, including *Capsella* (Steige et al. 2017), mice (Crowley et al. 2015) and
515 *Saccharomyces* (Zhang and Emerson 2019). However, there have been reports of strong *cis* effects in *D.*
516 *melanogaster* (León-Novelo et al., 2018), interspecific *Drosophila* hybrids (Graze et al. 2012) and natural
517 populations of flycatchers (Wang et al. 2017), though the detection of low-effect variants may be
518 encumbered by the lack of statistical power in some of these studies. An advantage of the Bayesian model
519 approach we used is that it should allow us to infer the average effect size without the bias ("winner's
520 curse") from gene-by-gene analyses.
521
522 Using gene-by-gene analyses, we found numerous instances of sex-dependent *cis*-regulatory effects in all
523 tissues, though these were much more common in gonads than heads (Table 1). The inference from the
524 Bayesian analysis indicated that almost all (>94%) *cis*-regulatory variants affect the sexes differently in
525 gonads. Puixeu et al. (2023) also found a substantially higher frequency genes with sex differences in *cis*-
526 regulatory effects in gonads than heads. The distinct regulatory architecture of the gonads, manifest in the
527 high frequency of *cis*-regulatory variants with sex-dependent effects, should allow expression in gonads
528 to evolve relatively independently between the two sexes whenever there is selective pressure to do so.
529 By contrast, in heads we estimated that only a small fraction (~5%) of *cis*-regulatory effects have sex-
530 dependent effects, restricting the potential for independent evolution of expression between the two sexes.

It is unsurprising that the frequency of sex-dependent 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 sex-dependent 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 (Curtsinger et al., 1994; Kidwell et al., 1977; 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.

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

565 temperature (Chen et al. 2015; Li and Fay 2017), tissue (Pinter et al. 2015) , or metabolic state (Shih and
566 Fay 2021). Of the two dimensions we examined—sex and tissue—neither stood out as harbouring
567 dramatically more variation than the other (i.e., sex-dependent AI was very common in gonads as was
568 tissue-dependent AI in both sexes). We were motivated to look for sex-dependent AI, at least in part,
569 because of the possibility that sexually antagonistic selection might help maintain such variants.
570 However, the qualitatively similar frequency of tissue-dependent AI suggests that either sexual
571 antagonism is playing little role in maintaining regulatory variation genome-wide or that antagonistic
572 pleiotropy across tissues is similarly common. Regardless, our results add to the existing literature that
573 *cis*-regulatory effects are often context dependent in magnitude and sometimes even in direction.
574 Moreover, different dimensions of context can interact to affect allelic expression, e.g., we find a sizeable
575 number of genes with tissue-by-sex interaction effects on AI. It is plausible that this thinking extends to
576 other contexts (e.g., sex-dependent *cis*-regulatory effects may be sensitive to temperature).
577
578 Allele-specific expression analysis can be plagued by inaccurate mapping of reads to the reference. We
579 attempted minimize the influence of mapping bias in several ways. We constructed genotype-specific
580 references and ascertained their efficacy by competitively mapping parental genomic reads to the
581 references. We tested F_1 genomic data for significant deviation from the expected read count ratio of 1:1
582 and removed any such gene from further consideration. We also tested for apparent 'parental effects' on
583 allelic expression, which may ensue from any residual mapping bias, and excluded any such genes.
584 Among our retained genes, the percentage of reads assigned to the DGRP-177 allele are uncorrelated
585 between genomic and transcriptomic data, whereas a positive correlation would be expected if mapping
586 bias was a substantial problem. The correlation remains negligible even if we only consider those genes
587 with significant AI in expression. These steps are designed to allay the effects of mapping bias on our
588 inferences, with evidence suggesting they have been largely effective (Supplementary Material).
589 Nonetheless, we cannot entirely rule out the possibility that mapping bias may persist in the dataset within
590 a small number of retained genes. While we believe our major findings are robust, the results for
591 individual genes would benefit from further experimental validation.
592
593 As we infer *cis*-regulatory polymorphisms from allelic imbalance (i.e., relative expression) in
594 heterozygotes, an additional note of caution is needed with respect to total expression. If allele *A* is
595 expressed more than *B* in an *A/B* heterozygote, it is reasonable to assume that total expression of this gene
596 would be greater in *A/A* than *B/B* homozygotes. However, this may not always be the case owing to
597 negative feedback loops in transcription. Further experimental testing—including both homozygotes—

598 would be needed to evaluate how often observed AI causes the predicted effects on total expression,
599 though this is extremely difficult to do at scale while controlling for genetic background effects.
600
601 In summary, our data indicate that *cis*-regulatory variants are common and their effects are often
602 dependent on sex and tissue. The frequency of *cis*-regulatory variants varies non-randomly with respect to
603 sex-biased gene expression as well as across tissues. Further work remains to be done to more thoroughly
604 examine variants at individual loci as well as to understand the evolutionary forces shaping patterns of
605 variation genome-wide.
606
607

608 **Literature Cited**

609 Anders S, Pyl PT, Huber W. 2015. HTSeq--a Python framework to work with high-throughput
610 sequencing data. Bioinformatics. 31(2):166–169. doi:10.1093/bioinformatics/btu638.
611 Auwera GA, Carneiro MO, Hartl C, Poplin R, del Angel G, Levy-Moonshine A, Jordan T, Shakir K,
612 Roazen D, Thibault J, et al. 2013. From FastQ Data to High-Confidence Variant Calls: The
613 Genome Analysis Toolkit Best Practices Pipeline. Current Protocols in Bioinformatics. 43(1).
614 doi:10.1002/0471250953.bi1110s43. [accessed 2021 Aug 20].
615 https://onlinelibrary.wiley.com/doi/10.1002/0471250953.bi1110s43.
616 Barson NJ, Aykanat T, Hindar K, Baranski M, Bolstad GH, Fiske P, Jacq C, Jensen AJ, Johnston SE,
617 Karlsson S, et al. 2015. Sex-dependent dominance at a single locus maintains variation in age at
618 maturity in salmon. Nature. 528(7582):405–408. doi:10.1038/nature16062.
619 Bates, D., Maechler, M., Bolker, B. and S. Walker. 2015. Fitting Linear Mixed-Effects Models Using
620 *lme4*. Journal of Statistical Software, 67(1), 1-48. doi:10.18637/jss.v067.i01.
621 Bonduriansky R, Chenoweth SF. 2009. Intralocus sexual conflict. Trends in Ecology & Evolution.
622 24(5):280–288. doi:10.1016/j.tree.2008.12.005.
623 Campbell CD, Kirby A, Nemesh J, Daly MJ, Hirschhorn JN. 2008. A survey of allelic imbalance in F1
624 mice. Genome Res. 18(4):555–563. doi:10.1101/gr.068692.107.
625 Chen J, Nolte V, Schlötterer C. 2015. Temperature Stress Mediates Decanalization and Dominance of
626 Gene Expression in Drosophila melanogaster. PLoS Genet. 11(2):e1004883.
627 doi:10.1371/journal.pgen.1004883.
628 Cheng C, Kirkpatrick M. 2016. Sex-specific selection and sex-biased gene expression in humans and
629 flies. PLoS Genet. 12(9):e1006170. doi:10.1371/journal.pgen.1006170.
630 Connallon T, Chenoweth SF. 2019. Dominance reversals and the maintenance of genetic variation for
631 fitness. PLoS Biol. 17(1):e3000118. doi:10.1371/journal.pbio.3000118.
632 Connallon T, Clark AG. 2014. Balancing Selection in Species with Separate Sexes: Insights from Fisher's
633 Geometric Model. Genetics. 197(3):991–1006. doi:10.1534/genetics.114.165605.
634 Coolon JD, Stevenson KR, McManus CJ, Graveley BR, Wittkopp PJ. 2012. Genomic imprinting absent
635 in Drosophila melanogaster adult females. Cell Reports. 2(1):69–75.
636 doi:10.1016/j.celrep.2012.06.013.
637 Crowley JJ, Zhabotynsky V, Sun W, Huang S, Pakatci IK, Kim Y, Wang JR, Morgan AP, Calaway JD,
638 Aylor DL, et al. 2015. Analyses of allele-specific gene expression in highly divergent mouse
639 crosses identifies pervasive allelic imbalance. Nat Genet. 47:353–360. doi:10.1038/ng.3222.

640 Cubillos FA, Stegle O, Grondin C, Canut M, Tisné S, Gy I, Loudet O. 2014. Extensive *cis* -Regulatory
641 Variation Robust to Environmental Perturbation in *Arabidopsis*. Plant Cell. 26(11):4298–4310.
642 doi:10.1105/tpc.114.130310.
643 Curtsinger JW, Service PM, Prout T. 1994. Antagonistic pleiotropy, reversal of dominance, and genetic
644 polymorphism. American Naturalist. 144:210–228. doi:10.1086/285671.
645 Degner JF, Marioni JC, Pai AA, Pickrell JK, Nkadori E, Gilad Y, Pritchard JK. 2009. Effect of read-
646 mapping biases on detecting allele-specific expression from RNA-sequencing data. Bioinformatics.
647 25(24):3207–3212. doi:10.1093/bioinformatics/btp579.
648 Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR.
649 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 29(1):15–21.
650 doi:10.1093/bioinformatics/bts635.
651 Fairbairn DJ, Blanckenhorn WU, Székely T. 2007. Sex, size and gender roles: evolutionary studies of
652 sexual size dimorphism. Oxford University Press.
653 Genissel A, McIntyre LM, Wayne ML, Nuzhdin SV. 2007. Cis and Trans Regulatory Effects Contribute
654 to Natural Variation in Transcriptome of Drosophila melanogaster. Molecular Biology and
655 Evolution. 25(1):101–110. doi:10.1093/molbev/msm247.
656 Goncalves A, Leigh-Brown S, Thybert D, Stefflova K, Turro E, Flicek P, Brazma A, Odom DT, Marioni
657 JC. 2012. Extensive compensatory *cis-trans* regulation in the evolution of mouse gene expression.
658 Genome Res. 22(12):2376–2384. doi:10.1101/gr.142281.112.
659 Graze RM, Novelo LL, Amin V, Fear JM, Casella G, Nuzhdin SV, McIntyre LM. 2012. Allelic
660 Imbalance in Drosophila Hybrid Heads: Exons, Isoforms, and Evolution. Molecular Biology and
661 Evolution. 29(6):1521–1532. doi:10.1093/molbev/msr318.
662 Grieshop K, Arnqvist G. 2018. Sex-specific dominance reversal of genetic variation for fitness. PLoS
663 Biol. 16:e2006810. doi:10.1371/journal.pbio.2006810.
664 Grieshop K, Ho EKH, Kasimatis KR. 2024. Dominance reversals and the maintenance of genetic
665 variation. Proc B. accepted.
666 Gruber JD, Long AD. 2009. *Cis* -regulatory variation is typically polyallelic in Drosophila. Genetics.
667 181(2):661–670. doi:10.1534/genetics.108.098459.
668 Hartig F., Minunno F., S. Paul. 2023. BayesianTools: General-Purpose MCMC and SMC Samplers and
669 Tools for Bayesian Statistics. R package version 0.1.8, <https://CRAN.R-
670 project.org/package=BayesianTools>.
671 Kidwell JF, Clegg MT, Stewart FM, Prout T. 1977. Regions of stable equilibria for models of differential
672 selection in the two sexes under random mating. Genetics. 85(1):171–183.
673 doi:10.1093/genetics/85.1.171.

674 Kita R, Venkataram S, Zhou Y, Fraser HB. 2017. High-resolution mapping of *cis* -regulatory variation in
675 budding yeast. Proc Natl Acad Sci USA. 114(50):E10736–E10744. doi:10.1073/pnas.1717421114.
676 Lack JB, Lange JD, Tang AD, Corbett-Detig RB, Pool JE. 2016. A Thousand Fly Genomes: An
677 Expanded *Drosophila* Genome Nexus. Mol Biol Evol. 33(12):3308–3313.
678 doi:10.1093/molbev/msw195.
679 Lemos B, Araripe LO, Fontanillas P, Hartl DL. 2008. Dominance and the evolutionary accumulation of
680 cis- and trans-effects on gene expression. Proceedings of the National Academy of Sciences.
681 105(38):14471–14476. doi:10.1073/pnas.0805160105.
682 León-Novelo L, Gerken AR, Graze RM, McIntyre LM, Marroni F. 2018. Direct testing for allele-specific
683 expression differences between conditions. G3 Genes|Genomes|Genetics. 8(2):447–460.
684 doi:10.1534/g3.117.300139.
685 León-Novelo LG, McIntyre LM, Fear JM, Graze RM. 2014. A flexible Bayesian method for detecting
686 allelic imbalance in RNA-seq data. BMC Genomics. 15:920. doi:10.1186/1471-2164-15-920.
687 Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform.
688 Bioinformatics. 25(14):1754–1760. doi:10.1093/bioinformatics/btp324.
689 Li XC, Fay JC. 2017. Cis-Regulatory Divergence in Gene Expression between Two Thermally Divergent
690 Yeast Species. Genome Biology and Evolution. 9(5):1120–1129. doi:10.1093/gbe/evx072.
691 Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq
692 data with DESeq2. Genome Biol. 15(12):550. doi:10.1186/s13059-014-0550-8.
693 Mackay TFC, Richards S, Stone EA, Barbadilla A, Ayroles JF, Zhu D, Casillas S, Han Y, Magwire MM,
694 Cridland JM, et al. 2012. The Drosophila melanogaster Genetic Reference Panel. Nature.
695 482(7384):173–178. doi:10.1038/nature10811.
696 McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D,
697 Gabriel S, Daly M, et al. 2010. The Genome Analysis Toolkit: A MapReduce framework for
698 analyzing next-generation DNA sequencing data. Genome Research. 20(9):1297–1303.
699 doi:10.1101/gr.107524.110.
700 Metzger BPH, Wittkopp PJ, Coolon JosephD. 2017. Evolutionary Dynamics of Regulatory Changes
701 Underlying Gene Expression Divergence among Saccharomyces Species. Genome Biology and
702 Evolution. 9(4):843–854. doi:10.1093/gbe/evx035.
703 Osada N, Miyagi R, Takahashi A. 2017. *Cis* - and *Trans* -regulatory effects on gene expression in a
704 natural population of *Drosophila melanogaster*. Genetics. 206:2139–2148.
705 doi:10.1534/genetics.117.201459.
706 Parsch J, Ellegren H. 2013. The evolutionary causes and consequences of sex-biased gene expression. Nat
707 Rev Genet. 14(2):83–87. doi:10.1038/nrg3376.

708 Pastinen T, Hudson TJ. 2004. Cis-Acting Regulatory Variation in the Human Genome. Science.
709 306(5696):647–650. doi:10.1126/science.1101659.
710 Pearse DE, Barson NJ, Nome T, Gao G, Campbell MA, Abadía-Cardoso A, Anderson EC, Rundio DE,
711 Williams TH, Naish KA, et al. 2019. Sex-dependent dominance maintains migration supergene in
712 rainbow trout. Nat Ecol Evol. 3(12):1731–1742. doi:10.1038/s41559-019-1044-6.
713 Pinter SF, Colognori D, Beliveau BJ, Sadreyev RI, Payer B, Yildirim E, Wu C, Lee JT. 2015. Allelic
714 imbalance is a prevalent and tissue-specific feature of the mouse transcriptome. Genetics.
715 200(2):537–549. doi:10.1534/genetics.115.176263.
716 Pool JE, Corbett-Detig RB, Sugino RP, Stevens KA, Cardeno CM, Crepeau MW, Duchen P, Emerson JJ,
717 Saelao P, Begun DJ, et al. 2012. Population genomics of sub-Saharan Drosophila melanogaster:
718 African diversity and non-African admixture. PLoS Genet. 8(12):e1003080.
719 doi:10.1371/journal.pgen.1003080.
720 Puixeu, G, Macon, A, and B Vicoso. 2023. Sex-specific estimation of *cis* and *trans* regulation of gene
721 expression in heads and gonads of *Drosophila melanogaster*. *G3: Genes, Genomes, Genetics*
722 jkad121.
723 Rozowsky J, Abyzov A, Wang J, Alves P, Raha D, Harmanci A, Leng J, Bjornson R, Kong Y,
724 Kitabayashi N, et al. 2011. AlleleSeq: analysis of allele-specific expression and binding in a
725 network framework. Mol Syst Biol. 7(1):522. doi:10.1038/msb.2011.54.
726 Rowe L, Chenoweth SF, Agrawal AF. 2018. The genomics of sexual conflict. Amer Nat. 192:274-86.
727 Sánchez-Ramírez S, Cutter AD. 2021. CompMap: an allele-specific expression read-counter based on
728 competitive mapping. Bioinformatics. [accessed 2021 Aug 18].
729 http://biorxiv.org/lookup/doi/10.1101/2021.02.12.431019.
730 dos Santos G, Schroeder AJ, Goodman JL, Strelets VB, Crosby MA, Thurmond J, Emmert DB, Gelbart
731 WM, the FlyBase Consortium. 2015. FlyBase: introduction of the Drosophila melanogaster Release
732 6 reference genome assembly and large-scale migration of genome annotations. Nucleic Acids
733 Research. 43(D1):D690–D697. doi:10.1093/nar/gku1099.
734 Schärer L, Rowe L, Arnqvist G. 2012. Anisogamy, chance and the evolution of sex roles. Trends in
735 Ecology & Evolution. 27(5):260–264. doi:10.1016/j.tree.2011.12.006.
736 Shih C-H, Fay J. 2021. Cis-regulatory variants affect gene expression dynamics in yeast. eLife.
737 10:e68469. doi:10.7554/eLife.68469.
738 Singh A, Agrawal AF. 2023. Two forms of sexual dimorphism in gene expression in *Drosophila*
739 *melanogaster*: their coincidence and evolutionary genetics. Molecular Biology and Evolution.
740 40(5):msad091.

741 Spencer HG, Priest NK. 2016. The evolution of sex-specific dominance in response to sexually
742 antagonistic selection. Amer Nat 187:658-66.
743 Steige KA, Laenen B, Reimegård J, Scofield DG, Slotte T. 2017. Genomic analysis reveals major
744 determinants of *cis*- regulatory variation in *Capsella grandiflora*. Proc Natl Acad Sci USA.
745 114(5):1087–1092. doi:10.1073/pnas.1612561114.
746 Stern DL. 2000. Perspective: Evolutionary developmental biology and the problem of variation.
747 Evolution. 54(4):1079–1091. doi:10.1111/j.0014-3820.2000.tb00544.x.
748 Stranger BE, Montgomery SB, Dimas AS, Parts L, Stegle O, Ingle CE, Sekowska M, Smith GD, Evans
749 D, Gutierrez-Arcelus M, et al. 2012. Patterns of Cis Regulatory Variation in Diverse Human
750 Populations. Barsh GS, editor. PLoS Genet. 8(4):e1002639. doi:10.1371/journal.pgen.1002639.
751 Wang M, Uebbing S, Ellegren H. 2017. Bayesian Inference of Allele-Specific Gene Expression Indicates
752 Abundant Cis-Regulatory Variation in Natural Flycatcher Populations. Genome Biology and
753 Evolution. 9(5):1266–1279. doi:10.1093/gbe/evx080.
754 Wittkopp PJ, Haerum BK, Clark AG. 2004. Evolutionary changes in cis and trans gene regulation.
755 Nature. 430(6995):85–88. doi:10.1038/nature02698.
756 Wittkopp PJ, Haerum BK, Clark AG. 2006. Parent-of-Origin Effects on mRNA Expression in *Drosophila*
757 *melanogaster* Not Caused by Genomic Imprinting. Genetics. 173(3):1817–1821.
758 doi:10.1534/genetics.105.054684.
759 Wittkopp PJ, Haerum BK, Clark AG. 2008. Regulatory changes underlying expression differences within
760 and between Drosophila species. Nat Genet. 40(3):346–350. doi:10.1038/ng.77.
761 Wray GA. 2007. The evolutionary significance of cis-regulatory mutations. Nat Rev Genet. 8(3):206–
762 216. doi:10.1038/nrg2063.
763 Zhang X, Emerson JJ. 2019. Inferring the genetic architecture of expression variation from replicated
764 high throughput allele-specific expression experiments. Genetics. [accessed 2021 Dec 17].
765 http://biorxiv.org/lookup/doi/10.1101/699074.
766
767