The roles of cis- and trans-regulation in the evolution of regulatory incompatibilities and sexually dimorphic gene expression

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Published Version</td>
<td>10.1101/gr.156414.113</td>
</tr>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:22907483">http://nrs.harvard.edu/urn-3:HUL.InstRepos:22907483</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
The roles of cis- and trans-regulation in the evolution of regulatory incompatibilities and sexually dimorphic gene expression


*Genome Res.* 2014 24: 84-95 originally published online September 16, 2013
Access the most recent version at doi:10.1101/gr.156414.113

---

Supplemental Material

http://genome.cshlp.org/content/suppl/2013/11/04/gr.156414.113.DC1.html

References

This article cites 84 articles, 38 of which can be accessed free at:
http://genome.cshlp.org/content/24/1/84.full.html#ref-list-1

Open Access

Freely available online through the *Genome Research* Open Access option.

Creative Commons License

This article, published in *Genome Research*, is available under a Creative Commons License (Attribution-NonCommercial 3.0 Unported), as described at http://creativecommons.org/licenses/by-nc/3.0/.

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.
Evolutionary changes in gene expression underlie many aspects of phenotypic diversity within and among species. Understanding the genetic basis for evolved changes in gene expression is therefore an important component of a comprehensive understanding of the genetic basis of phenotypic evolution. Using interspecific introgression hybrids, we examined the genetic basis for divergence in genome-wide patterns of gene expression between Drosophila simulans and Drosophila mauritiana. We find that cis-regulatory and trans-regulatory divergences differ significantly in patterns of genetic architecture and evolution. The effects of cis-regulatory divergence are approximately additive in heterozygotes, quantitatively different between males and females, and well predicted by expression differences between the two parental species. In contrast, the effects of trans-regulatory divergence are associated with largely dominant introgressed alleles, have similar effects in the two sexes, and generate expression levels in hybrids outside the range of expression in both parental species. Although the effects of introgressed trans-regulatory alleles are similar in males and females, expression levels of the genes they regulate are sexually dimorphic between the parental D. simulans and D. mauritiana strains, suggesting that pure-species genotypes carry unlinked modifier alleles that increase sexual dimorphism in expression. Our results suggest that independent effects of cis-regulatory substitutions in males and females may favor their role in the evolution of sexually dimorphic phenotypes, and that trans-regulatory divergence is an important source of regulatory incompatibilities.

[Supplemental material is available for this article.]
ever, allometric differences in tissue size relative to body size between males and females can give the appearance of sex-biased expression in the absence of sexually dimorphic transcription when expression is assayed from whole animals (Ranz et al. 2004). Mutations may have different selective effects in males and females, due to global differences between the sexes in gene regulation and chromatin environments (Liu et al. 2005; Gelbart and Kuroda 2009), sex-specific reproductive strategies (Lawson Handley and Perrin 2007), and the need to express genes in sex-specific cells such as the gonads (Chintapalli et al. 2007). A frequently cited model of the evolution of sexually dimorphic gene expression invokes such “sexually antagonistic” alleles that are beneficial in one sex but detrimental in the other (Rice 1984). The fixation of a secondary mutation that changes expression of the antagonistic allele in the disfavored sex could simultaneously resolve the fitness conflict and generate sexually dimorphic expression. Whole-genome expression analyses have claimed to identify sexually antagonistic phenotypes in *Drosophila* (Conallon and Knowles 2005; Innocenti and Morrow 2010); however, neither study determined the genetic basis of sexually dimorphic expression, leaving key aspects of this model untested (Fry 2010).

To identify the genetic basis for regulatory divergence between closely related species, we conducted a genome-wide expression analysis in introgression hybrids between two sister species of *Drosophila*. Previous studies of gene expression in interspecific hybrids of *Drosophila* have focused largely on the F1 generation, assayed gene expression in sterile hybrid animals, and been restricted to one sex (Ranz et al. 2004; Wittkopp et al. 2004; Moehring et al. 2007; Wittkopp et al. 2008; Graze et al. 2009; Lu et al. 2010; McManus et al. 2010). In contrast, we focus here on advanced generation hybrids between *Drosophila simulans* and *Drosophila mauritiana*, which produce fertile F1 females, allowing genetic material to be moved between these species. We used microarrays to measure gene expression in genotypes derived from an inbred strain of *D. simulans* except for a small region on the third chromosome that was introgressed from an inbred strain of *D. mauritiana*. Males and females of all genotypes used here are viable and fertile, allowing us to assay expression in (relatively) phenotypically normal adults of both sexes. Our study constitutes a fine-scale genetic analysis of interspecific divergence in genome-wide gene expression and allows us to measure the effects of divergent regulatory loci in both sexes. We find that divergent gene expression caused by the introgressed genomic segments act largely via trans-regulation, that cis- and trans-regulatory divergence differ in multiple aspects of inheritance, and that cis-regulatory divergence is enriched for sexually dimorphic effects on expression.

**Figure 1.** Approach used to map factors within the introgressed region that affect gene expression. (A) Genomic regions marked with a visible transgene (P[w+]) were introgressed from *D. mauritiana* into a *D. simulans* background. Bars indicate the sex chromosomes (left) and the two major autosomes. (B) Gene expression was assayed in adult males and females from seven genotypes carrying overlapping introgressions on chromosome arm 3L. Horizontal black bars indicate the extent of the introgressed segments, and vertical lines denote the location of genotyping markers. Expression levels in the seven homozygous introgression lines are shown to the right for an exemplar gene (Est-6). (C) QTL mapping identified marker loci within the introgressions where the *D. mauritiana* allele has a detectable effect on gene expression relative to the background *D. simulans* strain.

**Results**

**Detecting divergent regulation in introgression genotypes**

We assayed genome-wide gene expression using microarrays in virgin adult male and female flies of seven hybrid genotypes that introduce a combined ~9.3 Mb on chromosome arm 3L from *D. mauritiana* into an otherwise *D. simulans* genetic background. Quantitative trait locus (QTL) mapping identified genes that are differentially expressed between introgression genotypes and the parental *D. simulans* strain with expression phenotypes that are statistically significantly linked to a marker within the introgressed region (Fig. 1). We refer to genes with genetic evidence for regulatory differences between species as “differently regulated.” Divergent regulation of genes located outside the introgressed region results from functionally divergent *D. mauritiana* alleles of trans-acting factors located within the introgressed region. Divergently regulated genes located within the introgressed region and close to a genetic marker linked to the regulatory effect are candidates for divergent cis-regulatory sequences.

In both sexes, we observe a significant enrichment of divergently regulated genes located within the introgressed region (Table 1) resulting from cis-regulatory divergence between these two species. This inference was confirmed by measurements of allele-specific expression via pyrosequencing in introgression heterozygotes generated by crossing introgression genotypes to the parental *D. simulans* strain (see Methods). We measured allele-specific expression of 90 genes in males; 85 of these genes also gave reproducible results in females. Among all genes assayed for allele-specific expression, 76% and 71% showed significant differences in expression between the *D. simulans* and *D. mauritiana* alleles in males and females, respectively (Supplemental Table 1), indicating cis-regulatory divergence at these loci. The allele-specific ratios obtained from pyrosequencing and the effects estimated from the microarrays are similar (Supplemental Fig. 1; p = 0.744 and 0.719 in males and females, respectively; P < 0.0001). We excluded those genes where pyrosequencing failed to confirm the microarray results from our analyses of cis-regulatory divergence.

Differentially expressed genes located within the introgressed region but unlinked from the genetic marker associated with the regulatory effect are regulated in trans. In males and females, respectively, 40% and 48% of divergently regulated genes located in the introgression are regulated in trans; however, many more introgressed genes are presumably divergently regulated by trans-factors located elsewhere in the genome. Genome-wide, we detect 665 genes in males and 490 genes in females that are divergently trans-regulated (Table 1) as a result of evolutionary divergence in a region comprising <10% of the *Drosophila* euchromatic genome.
Sexual dimorphism in divergent gene regulation

At a P-value of $2 \times 10^{-4}$, which corresponds to a false discovery rate (FDR) of 0.02 in males and 0.03 in females (see Methods), we detect 737 divergently regulated genes in males and 536 in females (Table 1), indicating that a significantly larger fraction of the genome is differentially expressed in males than in females as a result of the introgressed segments ($\chi^2 = 126$, $P < 0.001$). This excess in males is observed among genes regulated both in trans and cis, is robust to varying the FDR (Supplemental Table 2), and consists mainly of expression effects of small magnitude (Supplemental Fig. 2). Previous studies have shown that the expression of genes transcribed mainly or specifically in the testes diverges rapidly between species and is highly variable within species (Meiklejohn et al. 2003; Ranz et al. 2003; Zhang et al. 2007; Brawand et al. 2011). Although 100 genes divergently regulated in males are expressed specifically in testes in D. melanogaster (Chintapalli et al. 2007), testis-specific genes are not overrepresented among divergently regulated genes ($\chi^2 = 2.40$, $P > 0.05$) (Supplemental Table 3). Additionally, after excluding germline-specific genes, there remain significantly more divergently regulated genes in males than in females, indicating that the excess of divergent regulation seen in males cannot be solely attributed to rapid gene expression evolution in the male germline.

A large majority (80%) of genes detectably expressed in both sexes show significant divergent regulation in either males or females but not both (Supplemental Table 4). However, direct comparison of the homozygous effects of introgressed alleles on gene expression in males and females (Fig. 2A), regardless of the statistical significance of QTL linkage, suggests that requiring a significant association in both sexes overestimates the extent of genetic independence in divergent regulation between the sexes. In general, the largest expression effects are similar in males and females, although 100 genes divergently regulated in males are expressed specifically in testes in D. melanogaster (Chintapalli et al. 2007), testis-specific genes are not overrepresented among divergently regulated genes ($\chi^2 = 2.40$, $P > 0.05$) (Supplemental Table 3). Additionally, after excluding germline-specific genes, there remain significantly more divergently regulated genes in males than in females, indicating that the excess of divergent regulation seen in males cannot be solely attributed to rapid gene expression evolution in the male germline.

A large majority (~80%) of genes detectably expressed in both sexes show significant divergent regulation in either males or females but not both (Supplemental Table 4). However, direct comparison of the homozygous effects of introgressed alleles on gene expression in males and females (Fig. 2A), regardless of the statistical significance of QTL linkage, suggests that requiring a significant association in both sexes overestimates the extent of genetic independence in divergent regulation between the sexes. In general, the largest expression effects are similar in males and females,
Gene expression evolution in cis and trans

and many of these genes may show significant QTL linkage in only one sex due to limited statistical power. Among genes with a smaller magnitude of divergent expression, the effects are largely independent in males and females; we observe few antagonistic genetic effects such that genes are up-regulated in one sex and down-regulated in the other (Fig. 2A).

We compared ratios of allelic expression in males and females for the 85 genes assayed by pyrosequencing in both sexes. While allelic expression is overall positively correlated between males and females (Fig. 3; \( P = 0.586, P < 0.0001 \)), 37 genes (44%) have ratios of allelic expression that significantly differ between the sexes (\( P < 0.02, \text{FDR} = 0.05 \)). These differences are largely of degree, and the most extreme examples of sexually dimorphic allele-specific expression are genes with strong allelic biases in one sex and little bias in the other. These extreme cases include three genes with modestly but nonsignificantly elevated expression from the \( D. maeritiana \) allele in males, but strongly biased expression from the \( D. simulans \) allele in females (Fig. 3). These three genes are located in the same 262-kb chromosomal region (Fig. 4), and this clustering of genes with sex-reversed allele-specific expression is significantly different from random expectation (permutation test, \( P = 0.0025 \)). This 262-kb region includes two additional genes that could not be assayed by pyrosequencing in females, but that show significantly greater expression from the \( D. maeritiana \) allele in males; four of these five genes are expressed exclusively in testes, and the fifth is accessory-gland specific in \( D. melanogaster \). These observations suggest that cis-regulatory control of gene expression diverges across much of this 262-kb region has diverged between these two species in a sex-specific manner.

Among genes with significant QTL associations in at least one sex, we find that both the prevalence and magnitude of sexually dimorphic expression differs between cis- and trans-regulatory divergence (Fig. 2B). First, the proportion of genes with significantly different magnitudes of divergent regulation in males versus females (as assessed by nonoverlapping 95% confidence intervals) is greater in cis (55%) than in trans (36%) (Fisher’s exact test, \( P_{\text{FET}} = 0.0009 \)). Second, the median absolute value of the difference between males and females in divergent regulation is −1.4-fold greater among genes regulated in cis than in trans (Fig. 4; Mann–Whitney test, \( P_{\text{MW}} < 0.0001 \)). Greater sexual dimorphism is associated with larger expression effects (averaged over males and females), although this effect is weaker among divergently trans-regulated genes (\( P = 0.174, P < 0.0001 \)) than it is among divergently cis-regulated genes (\( P = 0.544, P < 0.0001 \)) or among the allele-specific expression ratios obtained from pyrosequencing (\( P = 0.487, P < 0.0001 \)) (Supplemental Fig. 3).

In contrast to the homozygous effects of introgressed alleles (Fig. 2B), the dominance parameter estimates (Haley and Knott 1992) from the QTL analysis are largely independent between males and females (Fig. 2C), indicating that the effects of introgressed alleles are more similar between the sexes when homozygous than when heterozygous. However, as with the homozygous effects, dominance estimates are more disparate between males and females for introgressed cis-regulatory variants than factors acting in trans. There is a weak but statistically significant positive relationship between male and female dominance parameters among introgressed trans-factors (\( P = 0.089, P = 0.008 \)), in contrast to a weak and marginally significant negative correlation between male and female dominance parameters associated with introgressed cis-regulatory divergence (\( P = -0.22, P = 0.046 \)) (Fig. 2C). This indicates that, on average, expression effects of cis-regulatory divergence in introgression heterozygotes deviate from an additive expectation in the opposite direction in males and females.

Most gene expression traits are oligogenic or polygenic (Brem et al. 2002; Schadt et al. 2003; Brem and Kruglyak 2005; Gibson and Weir 2005); as a consequence, gene expression divergence between the parental \( \text{sim}B \) and \( \text{mau}12 \) strains is expected to often result from the combined effects of divergence at multiple loci. Differences between the effects of introgressed factors (relative to \( \text{sim}B \)) and the effects of genome-wide divergence in \( \text{mau}12 \) (relative to \( \text{sim}B \)) indicate how the effects of introgressed factors on gene expression are modified by \( D. maeritiana \) alleles at other loci. We compared the extent of sexually dimorphic expression resulting from introgressed factors with estimates of sexually dimorphic expression between \( \text{sim}B \) and \( \text{mau}12 \) at divergently regulated genes. Among genes divergently regulated by trans-factors, the effects of the introgressed factors are significantly less sexually dimorphic than the effects of whole-genome divergence (\( P_{\text{MW}} < 0.0001 \)) (Fig. 5). In contrast, the median magnitude of sexual di-

**Figure 3.** The ratio of expression from the \( D. maeritiana \) and \( D. simulans \) alleles in introgression heterozygotes is plotted for females versus males for 85 genes assayed by pyrosequencing. Three genes located in a region with sex-reversed allele-specific expression are outlined in red. Error bars indicate 95% confidence limits.

**Figure 4.** Allele-specific expression among genes assayed by pyrosequencing within the introgressed region. (A) Log2 ratios of allelic expression \( (D. maeritiana/D. simulans) \) in males and females. Trend lines were obtained by loess smoothing using a span of 0.2. The candidate region containing local sexually dimorphic cis-regulatory effects is indicated in gray. (B) Differences in allelic expression between males and females as a function of location. The trend line was obtained by loess smoothing using a span of 0.2. The candidate region with sexually dimorphic cis-regulatory effects is apparent as a net excess of \( D. maeritiana \) allele expression in males, due to strong overexpression of the \( D. simulans \) allele in females.
Divergent trans-regulation is enriched for dominant D. mauritiana alleles that cause up-regulation

In addition to the differences in sexual dimorphism, cis- and trans-regulatory divergences differ substantially in other aspects of their inheritance and evolution between species. First, the magnitude of expression effects, as measured in fold-change, is significantly larger for cis-regulatory divergence than trans-regulatory divergence (Supplemental Table 5). Second, we observe that D. mauritiana alleles of trans-regulatory factors tend to up-regulate target genes, while there is no significant bias toward up- or down-regulation as a result of cis-regulatory divergence (Table 2). This bias is most pronounced among the largest expression effects, as 42/45 genes divergently trans-regulated twofold or more are up-regulated by D. mauritiana alleles. Third, in both sexes, the effects of cis-regulatory divergence are largely additive—the effects on expression in introgression heterozygotes are approximately half the effects in introgression homozygotes—whereas trans-acting factors are more often dominant (Fig. 6). Furthermore, among genes divergently regulated in trans, we observe a sex-by-species bias in dominance relationships; expression in introgression heterozygote males, but not females, is more similar to expression in introgression homozygotes than to simB, indicating that, in males, introgressed D. mauritiana alleles of trans-regulatory factors are more frequently dominant than their homologous D. simulans alleles (Table 3).

Finally, the relationship between the effects of introgressed factors and whole-genome divergence in expression levels differs significantly between cis- and trans-regulation (Fig. 7), indicating that introgressed cis- and trans-factors are affected differently by mau12 alleles at loci outside the introgressed region. Homozygous introgressed cis-regulatory effects predict well the effects of whole-genome divergence on expression levels at those loci in both sexes ($p_{male} = 0.75$, $p_{female} = 0.85$, $P < 0.0001$). In contrast, the effects of introgressed trans-acting factors are significantly poorer predictors of whole-genome divergence in expression levels. In males, the correlation between homozgyous effects due to trans-regulatory divergence and interspecific divergence, although positive ($r = 0.40$, $P < 0.0001$), is significantly weaker ($P < 0.0001$) than for cis-regulatory divergence. In females, there is a significant negative correlation between trans-regulatory introgression effects and whole-genome divergence ($r = -0.27$, $P < 0.0001$)—on average, D. mauritiana alleles at loci outside the introgressions reverse the effects of introgressed trans-factors in females.

Previous studies have found that in crosses both within (Gibson et al. 2004) and between (Ranz et al. 2004; Landry et al. 2005; McManus et al. 2010) Drosophila species, many genes are expressed in offspring at levels outside the range of both parental genotypes. Such transgressive segregation is consistent with compensatory evolution within species and regulatory incompatibilities among species (Ranz et al. 2003; Ranz et al. 2004; Landry et al. 2005), some of which may contribute to the loss of fitness in hybrids (Michalak and Noor 2003; Michalak and Noor 2004; Good et al. 2010). We refer to those genes expressed in introgression hybrids at levels outside the range of the parental species as “misexpressed.” Substantial majorities of genes in both sexes are misexpressed in introgression hybrids—58% and 72% of divergently regulated genes in males and females, respectively (this difference in the proportion of misexpression between males and females is statistically significant; $P_{ET} < 0.0001$). We observe an enrichment of misexpressed genes as a result of trans-regulatory divergence, although this pattern is only significant in females (Table 4).

Tissue-specific genes diverge more and are more sexually dimorphic in expression than broadly expressed genes

The ubiquity of a gene’s expression is an important determinant of rates of both sequence and expression evolution. Inducible genes in single-celled organisms (Basehoar et al. 2004; Tirosh et al. 2006) and genes expressed in a restricted set of cell types in multicellular
organisms (Duret and Mouchiroud 2000; Larracuente et al. 2008) evolve significantly more rapidly than ubiquitously expressed genes. Consistent with these earlier findings, breadth of expression across tissues and organs has a significant impact on the patterns of divergence and sex-biased expression in our data. The magnitude of the effects of introgressed factors on gene expression, whole-genome expression divergence between sim and mau12, the magnitude of allele-specific expression in introgression heterozygotes, and the difference in these effects between males and females are all larger among tissue-specific genes than broadly expressed genes (Supplemental Table 6); these results generally hold for both cis- and trans-regulatory divergence. As whole animals were used for expression analysis, the expression effects among tissue-specific genes are almost certainly underestimated (Chintapalli et al. 2007), suggesting that the effects of a trans-regulatory factor that affects the expression of multiple effectors of the humoral immune response distributed throughout the genome.

The expression of these 10 immunity genes is positively correlated between the sexes in introgression genotypes ($p = 0.78, P = 0.008$ (Supplemental Fig. 4), indicating that, in a mostly D. simulans genetic background, the D. mauritiana allele of this trans-acting factor has similar effects in males and females. In contrast, relative expression of these genes is uncorrelated between mau12 males and females ($P = 0.32$); males show variable expression, and females show ≥ twofold down-regulation and mis-expression. These observations are consistent with a model whereby D. mauritiana alleles at one or more loci outside the introgressed region decouple the effects of this trans-regulator in males and females, mainly by reversing its effects in females.

### Table 3. Dominance patterns by sex, genetic control, and direction of regulation

<table>
<thead>
<tr>
<th>Dominancea</th>
<th>mau allele</th>
<th>sim allele</th>
<th>Codominant</th>
<th>% codominantb</th>
<th>P-value species biasc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>trans-regulated</td>
<td>Up-regulated</td>
<td>168</td>
<td>33</td>
<td>177</td>
</tr>
<tr>
<td></td>
<td>Down-regulated</td>
<td></td>
<td>108</td>
<td>21</td>
<td>158</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td></td>
<td>276</td>
<td>54</td>
<td>335</td>
</tr>
<tr>
<td></td>
<td>cis-regulated</td>
<td>Up-regulated</td>
<td>10</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>Down-regulated</td>
<td></td>
<td>3</td>
<td>5</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td></td>
<td>13</td>
<td>6</td>
<td>53</td>
</tr>
<tr>
<td>Females</td>
<td>trans-regulated</td>
<td>Up-regulated</td>
<td>70</td>
<td>52</td>
<td>127</td>
</tr>
<tr>
<td></td>
<td>Down-regulated</td>
<td></td>
<td>68</td>
<td>114</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td></td>
<td>138</td>
<td>166</td>
<td>186</td>
</tr>
<tr>
<td></td>
<td>cis-regulated</td>
<td>Up-regulated</td>
<td>3</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Down-regulated</td>
<td></td>
<td>3</td>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td></td>
<td>6</td>
<td>10</td>
<td>30</td>
</tr>
</tbody>
</table>

*a Dominance was assigned to the mau or sim allele if the 95% confidence intervals around the $d$ estimate did not overlap 0; confidence intervals overlapping 0 were classified as codominant.

*bThe proportion of codominant alleles is significantly greater for expression traits regulated in cis than in trans in both sexes ($\chi^2$ test, $P < 0.0005$).

*cP-value species bias was determined by a $\chi^2$ test comparing the proportion of mau and sim dominant alleles with a 1:1 expectation.
Discussion

We used microarray analysis of introgression hybrids between *D. mauritiana* and *D. simulans* to sample the genetic basis for genome-wide differences in gene expression between these two species. Below we describe how the evolutionary histories of these species and our approach circumscribe our study before turning to the evolutionary conclusions that we draw from our genetic analyses.

First, *D. simulans* and *D. mauritiana* diverged ~242,000 yr ago and differ at 1.6% of sites across the euchromatic portion of the genome (Garrigan et al. 2012), and by at least 1000 structural differences totaling 1 Mb of euchromatic DNA (Langley et al. 2012). Both species have large effective population sizes and high levels of nucleotide polymorphism (Kliman and Hey 1993; Nolte et al. 2013). Given the large effective population sizes of both species and the recency of their divergence, some unknown number of the genetic factors that we identify are likely polymorphisms segregating in one or both species, rather than fixed differences.

Second, *D. simulans* and *D. mauritiana* are separated by multiple systems of reproductive isolation, including premating (Coyne 1992; Coyne and Charlesworth 1997), postmating–prezygotic (Price et al. 2000), and intrinsic postzygotic isolation (Lachaise et al. 1986). It has been estimated that ~15 loci across the genome can confer complete male sterility and two or three loci cause lethality when introgressed from *D. mauritiana* into *D. simulans* (Tao et al. 2003; Cattani and Presgraves 2009). Multiple mtDNA haplotypes have recently introgressed from *D. simulans* into *D. mauritiana* (Ballard 2000; Nunes et al. 2010; Garrigan et al. 2012), and 2%–5% of 5-kb genomic segments have evolutionary histories consistent with recent gene flow between these two species (Garrigan et al. 2012). Thus, there is the opportunity for gene flow between these two species, but significant barriers to gene flow also exist in the form of both a few major and presumably many more minor hybrid incompatibilities (True et al. 1996). Due to the accumulation of intrinsic postzygotic incompatibilities, we expect that many gene expression phenotypes associated with introgressed factors are deleterious; however, we do not know the fitness effects associated with the observed expression phenotypes. Nonetheless, in view of the reduced opportunity for selection to act in hybrids relative to the parental species, we assume that introgression hybrids are more likely to have deleterious patterns of gene expression than either of the pure species genotypes.

Third, we observe that ~90% of divergently regulated genes change expression as a result of species divergence at trans-factors (Table 1). Most previous studies in *Drosophila* have inferred that cis-regulatory divergence can account for at least 50% of differences in gene expression both within (Hughes et al. 2006; Genissel et al. 2008; Lemos et al. 2008a) and between species (Wittkopp et al. 2004, 2008; McManus et al. 2010). The large proportion of trans-divergence we detect here can likely be attributed to our use of homozygous introgression hybrids, as opposed to assaying *F*₁ progeny, for two reasons. First, our introgression approach sampled regulatory divergence across 9.3 Mb and estimated the effects of this divergence on genome-wide gene expression. Due to the polygenic nature of gene expression (Brem et al. 2002; Schadt et al. 2003; Brem and Kruglyak 2005; Gibson and Weir 2005), some differences in gene expression between these species result from divergence at multiple trans-acting loci. Because some genes divergently regulated by an introgressed trans-factor are also regulated by divergence at loci not captured by these introgressions, this approach overestimates the proportion of genes genome-wide that are divergently regulated in *trans*. Second, regulatory alleles that act recessively will be masked in heterozygous *F*₁ hybrids. Previous studies (Lemos et al. 2008a; McManus et al. 2010; Zhang et al. 2011; Gruber et al. 2012) and our results suggest that recessive alleles more frequently underlie trans-regulatory divergence (Table 4); *F*₁ analyses therefore likely underestimate divergent trans-regulation. A recent mutagenesis study in *Saccharomyces cerevisiae* is also broadly consistent with the idea that genome-wide heterozygosity may mask trans-regulatory variation. Gruber et al. (2012) found that >80% of de novo mutations assayed in a common genetic background acted in *trans* to modulate reporter gene expression, and these *trans*-mutations were masked in heterozygous diploids.

### Table 4. Gene misexpression results more frequently from divergent trans-regulation than cis

<table>
<thead>
<tr>
<th>Gene expression</th>
<th>trans</th>
<th>cis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Misexpressed</td>
<td>Not</td>
</tr>
<tr>
<td><strong>Males</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All genes</td>
<td>394</td>
<td>271</td>
</tr>
<tr>
<td>≥ 1.5-fold</td>
<td>38</td>
<td>9</td>
</tr>
<tr>
<td>≥ twofold</td>
<td>15</td>
<td>5</td>
</tr>
<tr>
<td><strong>Females</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All genes</td>
<td>362</td>
<td>128</td>
</tr>
<tr>
<td>≥ 1.5-fold</td>
<td>55</td>
<td>9</td>
</tr>
<tr>
<td>≥ twofold</td>
<td>23</td>
<td>2</td>
</tr>
</tbody>
</table>

Rows in boldface indicate that the proportion of misexpressed genes is significantly different (Fisher’s exact test, *P* ≤ 0.01) between trans- and cis-regulated genes.
Finally, we studied a single autosomal region for effects on genome-wide gene expression. It is therefore important to consider which results are likely to generally hold for regulatory divergence between *D. simulans* and *D. mauritiana*, and which may be idiosyncratic to the evolutionary history of this particular region of the genome. We assume that the majority of cis-regulatory divergence acts gene specifically, excepting the candidate region highlighted in Figure 5. Patterns associated with divergence in cis therefore result from many independent regulatory substitutions and therefore are likely to generally apply to autosomal, euchromatic cis-regulatory divergence between these species. In both sexes, each of the 10 genetic markers (Fig. 1) showed linkage to multiple trans-regulated expression phenotypes, indicating that, at a minimum, 10 divergent factors are required to account for the trans-regulatory divergence we observe. Patterns that we observe among trans-regulated expression phenotypes linked to all markers are the result of divergence at multiple regulatory factors and are therefore also likely to reveal general features of regulatory divergence between these two species. In contrast, patterns that are driven by expression phenotypes linked to only one marker are more likely to be idiosyncratic consequences of the specific mutations captured by these introgressions. Across both sexes, 67% of genes up-regulated more than 1.5-fold in trans show linkage to the same marker, and many of these genes have annotated immune function, suggesting that the large fraction of up-regulation in these genotypes (Table 2) and the enrichment of strongly up-regulated genes with immune function (Supplemental Table 7) may be due to a single introgressed trans-regulatory factor of large effect. All other patterns we describe are more likely to be general features of regulatory divergence in *Drosophila*.

Although a bias toward up-regulation in trans has been observed among de novo regulatory mutations in *S. cerevisiae* (Gruber et al. 2012), studies in *Drosophila* have found a significant excess of genes down-regulated in interspecific *F*1 hybrids relative to both parent species (Michalak and Noor 2003; Ranz et al. 2004; Moehring et al. 2007; McManus et al. 2010). However, these studies profiled sterile individuals with sterility phenotypes that range from post-meiotic disruption of gametogenesis (Kulathinal and Singh 1998) to largely agametic (Sturtevant 1920), and the large number of underepressed genes results at least in part from assaying whole animals lacking germline cells, compared with fertile individuals from the parental species (Ranz et al. 2004; McManus et al. 2010). Both sexes of the introgression hybrids used here are fertile, and thus gross differences in tissue composition between genotypes are not expected. As mentioned above, the up-regulation in trans we detect may be attributable to a single large-effect factor; it therefore remains unclear if biases toward up- or down-regulation are, in fact, common features of gene misregulation in *Drosophila* hybrids.

The evolution of sexually dimorphic gene expression

Assaying gene expression in both males and females allowed us to identify sex-biased and sex-specific effects of regulatory divergence. We detect significantly more divergently regulated genes in males than in females despite the fact that the false discovery rate is slightly higher in females (see Methods). An excess of cis-eQTL in males was also detected in a North Carolina population of *D. melanogaster* (Massouras et al. 2012). One possible explanation for these observations is that global gene expression in *Drosophila* males is simply more sensitive than expression in females to genetic variation and divergence. For example, genetic perturbations affecting heterochromatin have more severe effects on genome-wide expression in males than females (Liu et al. 2005; Deng et al. 2009), which may reflect interactions with male-specific factors that affect both euchromatic and heterochromatin gene regulation, such as the Y chromosome (Lemos et al. 2008b), or the somatic dosage compensation complex (Deng et al. 2009). However, we caution that greater statistical power associated with the microarray experiments in males, e.g., due to greater technical variation or sensitivity to random environmental effects in females, could potentially contribute to the excess of divergently regulated genes detected in males.

It has been proposed that the evolution of sexually dimorphic gene expression is driven in part by selection to resolve sexually antagonistic fitness variation (Rice 1984). This model assumes that mutations are generally less sexually dimorphic than required for optimal fitness in both sexes, in contrast to a model in which most mutations have disparate effects in males and females, and subsequent evolution restores fitness to the disfavored sex via reducing sexual dimorphism in expression. Mutation accumulation experiments in animals have measured the effects of de novo mutations on gene expression in one sex only (Denver et al. 2005; Rifikin et al. 2005), so the degree of sexual dimorphism among new mutations affecting gene expression levels is currently unknown. We observe that, among genes regulated by divergent trans-factors, the effects of introgressed alleles on gene expression are more similar between the males and females than the effects of genome-wide divergence between *simB* and *mau12*, indicating that *D. mauritiana* alleles elsewhere in the genome increase sexual dimorphism in expression relative to the introgressed trans-factors. Assuming that reduced dimorphism in some expression phenotypes due to introgressed regulatory factors confers decreased fitness relative to the pure species genotypes, this pattern supports the scenario whereby compensatory evolution increases the sexual dimorphism of individual mutations and suggests the existence of the modifier alleles predicted by Rice's model (Rice 1984).

The median homozygous effect of introgressed cis-regulatory sequences is 40% more sexually dimorphic than the median effect of introgressed trans-regulatory factors (Fig. 5), and the effects of cis-regulatory divergence in males and females are similar in the *simB* and *mau12* backgrounds (Fig. 5; Supplemental Fig. 5). This indicates that, while *mau12* alleles at loci outside the introgressions modulate the effects of introgressed trans-factors to increase sex differences in expression, such modulation is, on average, less frequent for cis-regulatory divergence. Case studies have repeatedly identified that divergence at cis-regulatory elements is responsible for the evolution of sex-specific patterns of gene expression (Gompel et al. 2005; Jeong et al. 2006; Williams et al. 2008), and a population genomic study found that 75% of cis-eQTL are sex specific within a single *D. melanogaster* population (Massouras et al. 2012). Together with these studies, our results suggest that disparity in expression between the sexes may be a general feature of cis-regulatory control and that this disparity includes overall expression levels, in addition to the timing and spatial patterns of gene expression through development. If different expression optima in males and females are common, this may be an important pleiotropic constraint that is alleviated by cis-regulatory evolution (Stern 2000).

The enrichment of sexually dimorphic gene expression associated with cis-regulation could conceivably result from either molecular mechanisms of gene regulation or evolutionary processes. In *Drosophila*, global differences between males and females in the balance of euchromatin and heterochromatin (Deng et al.
2009) and the distribution of covalent histone modifications (Lucchesi et al. 2005) may give males and females distinct chromatin states across the genome, which could enrich sexually dimorphic effects among cis-regulatory mutations. An alternate, but not mutually exclusive, possibility is that cis-regulatory mutations that affect sexually dimorphic expression are more likely to be tightly linked to sexually antagonistic protein-coding mutations located in the same gene, potentially facilitating the spread of both mutations through a population (Rice 1984).

Dominance and the evolution of gene expression
Assaying gene expression in homozygous and heterozygous introgression genotypes allowed us to estimate the dominance of divergent regulatory factors. We observe that, in contrast to the additive effects of introgressed factors, dominance effects are independent between males and females, and are even weakly negatively correlated among cis-regulatory divergence (Fig. 2C). Furthermore, D. mauritiana trans-regulatory alleles are predominantly dominant in males only (Table 3). These observations have important implications for our understanding of sexually antagonistic fitness variation, as the dominance of sexually antagonistic mutations is a critical parameter determining their fate in populations. Sex-specific dominance resulting from X-chromosome hemizygosity in males allows X-linked sexually antagonistic alleles to segregate at higher frequencies than equivalent autosomal alleles that have similar dominance effects between males and females (Rice 1984). However, we observe extensive sex-specific dominance in expression effects associated with introgressed autosomal regulatory factors. If sex-specific dominance is common, then the fate of sexually antagonistic alleles within populations may be largely determined by the particular combination of sex-specific fitness effects and dominance relationships associated with these mutations, and sexually antagonistic fitness variation may not be restricted to the X chromosome (Fry 2010).

The dominance of gene expression changes differs significantly between divergent cis- and trans-regulation. Divergent trans-regulation is enriched for dominant or recessive alleles, while the effects of cis-regulatory divergence are intermediate in heterozygous genotypes (Fig. 6; Table 3). Additivity has been previously documented for cis-regulatory variants (Wray 2007; Lemos et al. 2008a; McManus et al. 2010; Gruber et al. 2012) and is expected in the absence of transvection due to independent transcription of homologous alleles in a diploid heterozygote (Wray 2007). As additive effects expose rare alleles to selection, this may cause fewer advantageous cis-regulatory alleles to be lost to drift, potentially enriching their fixation relative to their mutational origination as compared with trans-acting mutations (Wray 2007). Additionally, additivity may enrich cis-regulatory variants for mutations with heterozygote advantage and influence levels of standing cis-regulatory genetic variation (Sellis et al. 2011).

The contributions of cis- and trans-regulatory divergence to the evolution of gene expression
The average fold-change resulting from cis-regulatory divergence identified here is significantly larger than those resulting from trans-divergence (Supplemental Table 5), consistent with what appears to be a general pattern of larger expression effects associated with cis-regulatory variation (Brem et al. 2002; Schadt et al. 2003; Hughes et al. 2006; Lemos et al. 2008a; McManus et al. 2010; Gruber et al. 2012). The effects of divergent trans-regulatory factors are weakly correlated (in males) or anti-correlated (in females) with overall expression differences between simB and mau12, while the effects of cis-regulatory divergence are positively correlated with interspecific expression differences (Fig. 7), resulting in an excess of gene misexpression due to divergent trans-regulation (Table 4). In our experiments, divergent regulation in both cis and trans results from interspecific interactions between alleles that modulate gene expression (D. simulans trans-factors with D. mauritiana cis-regulatory sequences and D. mauritiana trans-factors with D. simulans cis-regulatory sequences, respectively). However, only divergent trans-regulation is affected by interactions between trans-factors from both species. Interspecific interactions between trans-factors, then, could be responsible for the observed mis-expression bias. If true, this suggests that misexpression rarely results from direct interactions between trans-factors and cis-regulatory sequences, but rather as a downstream consequence of interspecific interactions between alleles at two trans-regulatory factors, or between alleles at loci whose effects on gene expression are indirect (Yvert et al. 2003).

Finally, our results have implications for the evolution of gene expression phenotypes and the accumulation of incompatibilities that contribute to or follow reproductive isolation. For gene expression levels that are under stabilizing selection in both species—which is likely the majority (Rifkin et al. 2003; Denver et al. 2005; Lemos et al. 2005; Ronald and Akery 2007)—expression phenotypes that are deleterious in the introgression hybrid relative to both parental species will be enriched among misexpressed genes, where expression in the hybrid is outside the range of the parents. Such expression incompatibilities will therefore be more common among trans-regulatory divergences, and, as described above, these may result mostly from trans–trans interactions. This suggests that evolutionary turnover of cis-regulatory DNA and the proteins that interact with these sequences may contribute less to the evolution of incompatibilities, at least for the expression of euchromatic single-copy genes. Additionally, if a significant fraction of regulatory mutations select for secondary mutations that modify regulatory effects in a sex-specific manner, then sex-specific misexpression may contribute to a loss of fitness in hybrids.

Methods
Drosophila genetics and husbandry
The construction of the introgression genotypes used here was originally described in Tao et al. (2003). Stocks of a D. mauritiana w’ strain (mau12) carrying a codominant visible marker (P[w’]) on the third chromosome were crossed to a w’ strain of D. simulans, and hybrid females carrying D. mauritiana material marked by P[w’] were backcrossed to D. simulans w’ males for four generations. Introgression lines were then crossed to a w’ genotype of D. simulans (referred to hereafter as simB) with a recessively marked second chromosome, and third chromosomes carrying introgressed segments were passaged through males to ensure that the X and second chromosomes were derived entirely from simB. To eliminate mutations that may have accumulated in these lines since their construction (e.g., Denver et al. 2005; Rifkin et al. 2005; Landry et al. 2007), the introgressed D. mauritiana segments were re-extracted into the simB background using the crossing scheme shown in Supplemental Figure 6. In the final generation of the re-extraction, single males and females were paired to create sublines of each original introgression genotype. To confirm that the introgressions remained intact following two generations of re-
Microarray and pyrosequencing methods

Microarray construction, RNA extraction, cDNA synthesis, and hybridization were performed following standard protocols. Two-color competitive hybridizations followed a modified loop design (Supplemental Fig. 7) that minimizes the distance between any two genotypes (Townsend 2003). Allele-specific expression was assayed by pyrosequencing at 90 candidate genes identified from the microarrays as potentially harboring cis-regulatory variants (Supplemental Fig. 8; Supplemental Table 1). Detailed microarray and pyrosequencing methods are available as Supplemental Material.

Statistical analyses

Microarray log2 expression ratios were analyzed with an empirical Bayes linear model (Smyth 2004), and relative expression levels were estimated for each homozygous and heterozygous introgression genotypes (introgression/simB), and the parental simB and mau12 strains, for a total of 16 genotypes.

All flies used for RNA extraction were reared on standard cornmeal molasses media in glass vials at 25°C on a 12:12 light:dark cycle. Culture density was controlled by combining 15 males with 15 females per vial for homozygous introgression genotypes, and 10 males with 10 females for the simB, mau12, and heterozygous introgression genotypes. Flies for RNA extraction were reared in three independent cohorts. Adults were collected as virgins, and the sexes were aged separately for 1–2 d before flash-freezing in liquid nitrogen and storage at −80°C. All flies were frozen at the same time of day to control for circadian effects on gene expression. Heterozygous introgression genotypes were generated from both reciprocal crosses, i.e., virgin simB females were crossed to introgression line males, virgin introgression line females were crossed to simB males, and equal amounts of RNA extracted from the offspring of reciprocal crosses were pooled for microarray hybridization. This approach averages over any potential parent-of-origin effects on expression, and thus such effects can neither be detected nor confound our results; in any case, there is little evidence for imprinting or parent-of-origin effects on gene expression in adult Drosophila (Coolon et al. 2012).

In this model, expression in homozygous introgression genotypes is estimated by 2a and expression in heterozygous introgression genotypes by a + d. QTL analyses were done using R/qtl v1.23 (Broman et al. 2003).

The P-value cutoff for assigning a statistically significant association between expression phenotypes and genetic markers was obtained by setting the false discovery rate (FDR) (Benjamini and Hochberg 1995) to 0.02 in males, which corresponds to a nominal χ² P-value of 0.00021. This P-value cutoff was then used for both the male and female experiments, resulting in an FDR of 0.03 in females. This FDR was calculated using all linkage tests at all markers, but we considered only the marker with the best fit for each gene expression phenotype. This approach is very conservative with respect to false positives, as shown in Q–Q plots of the QTL linkage P-values (Supplemental Fig. 9). A common P-value cutoff and slightly different FDR rates across the sexes were deemed preferable to controlling the FDR to a common value in both males and females, as an FDR of 0.02 in females corresponds to a nominal P-value of 0.000065, which reduces the number of observed significant associations in females by half. Our decision to set a common P-value cutoff is therefore conservative with regard to the excess in the number of significant associations detected in males.

Normalized log2-transformed relative expression measures from pyrosequencing were analyzed in two steps. First, a mixed linear model including main effects of sex, strain, reciprocal cross, and all two-way interactions were fit for all genes separately, excepting four genes assayed using only one genotype, for which the model included sex, reciprocal cross, and their interaction. The results of these models were examined, and effects of strain and cross and their interactions were dropped if they were not significant. Based on this approach, one of three statistical models was used for each gene (see Supplemental Table 8):

1. log2 normalized ratio of allelic expression − sex + strain + ε (four genes)
2. log2 normalized ratio of allelic expression − sex + cross + sex * cross + ε (five genes)
3. log2 normalized ratio of allelic expression − sex + ε (81 genes)

All models used type III sums of squares and were run using the GLM procedure in SAS statistical software (version 10).

FlyAtlas analyses

Gene expression estimates from 20 tissues dissected from larval and adult D. melanogaster were compiled from FlyAtlas (Chintapalli et al. 2007). Affymetrix microarray probe sets with multiple matches to the genome were excluded from the analysis, and among genes with multiple probe sets, the probe set with the highest mean signal intensity across samples was selected. Signal intensities at probe sets with absent calls were set to 1, and all intensity values were log2-transformed. Mean expression values were then calculated from four replicate arrays. Tissue specificity was calculated as the τ statistic (Yanai et al. 2005).

All statistical analyses except for the pyrosequencing models were done in R (v2.14.1) (R Development Core Team 2011). Microarray analyses used the limma package (v3.10.3); standardized major axis regression was used for bivariate regression analysis with the smatr package (v3.2.4).

Data access

Microarray data have been submitted to the NCBI Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/) under accession number GSE44329. Pyrosequencing data are included in Supplemental Table 8.
Acknowledgments

We thank David Rand for his generous support, Yun Tao for providing the introgression genotypes, and Daven Presgraves and Kristi Montooth for comments that improved the manuscript. This work was supported by funds from the NSF to C.D.M. (DEB-0839348) and P.J.W. (MCB-1021398) and funds from the NIH to J.D.C. (1F32-GM089009-01A1) and D.L.H. (GM065169).

Author contributions: C.D.M., J.D.C., and P.J.W. conceived and designed the experiments; C.D.M. and J.D.C. performed the experiments; C.D.M. and J.D.C. analyzed the data; C.D.M., D.L.H., and P.J.W. contributed reagents/materials/analysis tools; and C.D.M. wrote the paper.

References

Gene expression evolution in cis and trans


Received February 15, 2013; accepted in revised form September 11, 2013.