FILE S1

Supporting Information

Identification of probes for allele specific expression

Previous studies have shown that SNPs and or in/dels between probe sequences and target sequences can cause loss of signal intensity in one allele (RONALD et al. 2005). In this experimental design, sequence differences between *D. melanogaster* and *D. simulans* may result in reduced or absent signal from *D. simulans* as the tiling array probes were designed based upon the *D. melanogaster* sequence. Exons were BLASTed on *D. melanogaster* 4.2 genome to map to their position to the aligned scaffold for *D. simulans* (http://www.dpgp.org; BEGUN et al. 2007) and then the aligned *D. melanogaster*, *D. simulans* set of sequences was excised. Micro-alignments between the *D. simulans* lines and *D. melanogaster* results were performed using clustal-W. If there was no sequence divergence, the probe was classified as no-SNP. If there were multiple SNPs or a single SNP between probe bases 9 and 17 between *D. melanogaster* and all *D. simulans* lines with available sequence, the probe was classified as containing a SNP. All other configurations (*i.e.* in/dels, missing sequence, non-fixed differences) were considered missing data relative to the SNP status (168,467 probes with SNPs, 206,492 probes with no SNPs and 367,752 probes unclassified).

To make maximum use of the array, an alternative method is needed to identify probes which have detectable hybridization polymorphisms between species. For each gene, the signals for all probes targeting the same set of transcripts were ranked within the sample. Probes that were in the top two thirds on the *D. melanogaster* gDNA array, but changed to the bottom third on both *D. simulans* gDNA arrays were identified as allele-specific probes (ASE). All remaining probes (those that do not change rank between species) with detectable levels of hybridization signal for all six gDNA arrays were used to estimate overall gene expression (OE). Performance of the rank based probe classification for ASE probes was then evaluated against the expected probe classification based on the polymorphism data and the method was found to perform well. For each gene, the signals for all ASE probes targeting the same set of transcripts were grouped together into an ASE probe set (Table 1). If an ASE probe set did not contain at least three detected probes it was removed from further analyses. A probe set representing overall expression level was developed in the same way. Details are in the methods section.

Quality control filters

To implement the logical approach of WITKOPP et al. (2004) both overall expression (OE) and allele-specific expression (ASE) must be estimated for the same set of transcripts. Further, if no overall expression is detected, then the contribution of the allele-specific expression cannot be assessed. Therefore, if OE for a probe set was not detected in at least one of the parents and in both hybrids, *cis* and *trans* effects for that set of transcripts were not tested. There were 5,682 sets of transcripts (representing 5,366 genes) that met these criteria for OE probe sets and had ASE probe sets.

As a final check on the quality of ASE probes, the following quality control filters were implemented. First, undetected ASE probe sets were eliminated. Allele-specific expression signal for a probe set was considered to be absent on a given array if less than 50% of the probes had detectable signal for the RNA arrays. If allele-specific expression for a probe set was not detected in 2 or more replicates it was considered absent for that genotype. If allele-specific expression was not detected in *D. melanogaster*, $F1_w$, or $F1_c$ than it was removed before further analyses (1,063 probe sets removed).

Second, to assess probe hybridization bias, the ASE signal was compared to the OE in the *D. melanogaster* RNA samples for each probe set (OE = ASE). If the the null hypothesis was rejected the ASE probes were removed from analysis (1,572 remaining probe sets failed this test). ASE can be equal to or less than OE, but should never be significantly greater than the overall expression level. This was tested using a one tailed t-test (only 6 probe sets were filtered at this stage). Finally, probe sets with significant differences in cross-hybridization between gDNA and cDNA samples, were eliminated if the null hypothesis ASE_{RNA} - OE_{RNA} = ASE_{DNA} - OE_{DNA} for the *D. melanogaster* samples was rejected (582 probe sets failed this test).

GO annotation enrichment tests

All possible GO terms were tested for enrichment among species, *cis* and *trans* significant tests using a chi-square test of association between GO category and significance (Rivals et al., 2007). For each comparison, species, *cis*, or *trans* the reference list was all genes that were tested for the particular effect using a Fisher's exact test. For example for the enrichment test for the *cis* effect, the reference list is all genes tested for *cis* effects. To control for the multiplicity in testing all GO categories an FDR correction was calculated. Several categories were enriched for species at a nominal p-value of 0.05 that were related to the defense system, olfaction and nervous system. This led to the creation of the broader functional categories. Defense related genes annotated in GO with key words " defense", "virus", "antioxidant", "antimicrobial", "antibacterial", "immune" "antiviral" were grouped into a single 'Defense' category. In addition, genes reported as defense related in LEMAITRE AND HOFMAN (2006) and SACKTON et al. (2007) were included in the 'Defense' category. An olfaction category was created from all GO biological process categories containing the descriptors 'olfactory', 'olfaction' and 'odorant'. A category for neurological genes was created from all GO biological process categories containing the descriptors 'nerve', 'nervous', and 'neuro'(All GO categories corresponding to the defense, olfaction and neurological categories are reported in Table S5). Association between these constructed functional groups with species differences, *cis* effects and *trans* effects were tested using a Fisher's exact test. It is difficult to determine the appropriate correction for these p-values, since these tests are not independent from the initial set of enrichment tests. The raw p-values are reported, and the interpretation is cautious with regards to these results.

Line differences in expression:

As described in the methods, a single *D. melanogaster* genotype *y; cn bw sp,* (Mel) two *D. simulans* parental genotypes (w501, C167.4), and the resulting two hybrid genotypes ($F1_W$, $F1_C$), were assayed. Before proceeding to testing the overall effects of species, the two parental lines were compared to each other. The differences among parents were small. There were fewer than 1.5% of probe sets that showed evidence for differences between *D. simulans* parental lines. The number of differences between hybrid genotypes was even smaller, with only 0.4% showing evidence for differences between the two *D. simulans* lines (Table 3). Consequently, the effects of the *D. simulans species* were considered rather than the effects of individual lines.

Expression phenotypes possibly associated with known mutations in the sequenced lines:

Vision and body color genes are of specific interest for the lines analyzed, because two of the lines used in this study carry visible mutations, *D. melanogaster y; cn bw sp* (*y*, body color; *cn* and *bw*, eye color; *sp*, body color and wing) and *D.*

simulans w501 (*w*, eye color). Of these genes, three had detectable probe sets for which species differences could be tested (*w*, *cn*, and *bw*). Only *bw* (for transcript Fbtr0072117 only) was significant for a difference between species (the common probe set for *bw* was not significant). RANZ et al., 2004, also found greater *bw* expression in *D. simulans* relative to *D. melanogaster*, but for wild type lines. It is unclear whether the transcript level differences are at all related to the insertion responsible for the *bw* mutant phenotype. Interestingly, while the mutant genes themselves do not appear to have strong expression differences between the *D. simulans* and *D. melanogaster*, four major phototransduction genes (*Rh6*, *ninaD*, *trpl* and *rdgA*) were expressed differently between *D. simulans* and *D. melanogaster*. Additionally, two genes involved in pigment synthesis Dat and yellow-f (WITTKOPP et al. 2003) are differentially expressed and could result from the mutations in *y* and/or *sp*.

Of the four major vision genes listed in the species section only two were testable for *cis* and *trans*. These genes, *Rh6* and *trpl*, are both significant for *cis* differences. Additional genes related to vision, however, were significant for *cis* and or *trans* differences (For example, *trans: ninaC* and *cac; cis: WASp*, *norpA*, *sec6*, *trp*; *cis* and *trans: Rh3* and *Notum*; Complete lists given in Table S6). The vision functional category was constructed from all genes with the keywords 'visual', 'eye ', and 'phototransduction ' for GO biological process annotation and from genes annotated as GO cellular compartment 'rhabdomere'. Overall, while vision genes were not enriched among *cis* (or *trans*) significant genes, there were 27 vision annotated genes with significant *cis* differences and 12 *trans*. The preponderance of *cis* rather than *trans* effects for these genes suggests that these are not simple downstream effects of the eye mutations present in two of the compared genotypes (which we might expect to observe as large *trans* effects on expression). Rather than natural species differences, however, the observed regulatory differences in these genes could be due to cryptic variation only observed in the 'sensitized' mutant background. Alternatively, *cis* and *trans* differences in vision related genes may arise as a result of complex epistasis for expression, as F1 mel-sim hybrids have a partially penetrant rough eye phenotype (BARBASH *et al.* 2003) Note that a portion of these vision related genes are also sex biased or sex specifically regulated, so this pattern may be a composite effect of multiple contributing factors.

Correlation of *cis* and *trans* estimates:

Differences in expression levels between species can be due to divergence in *cis*, *trans* or a combination of these effects. As the total transcript level is not conserved in these cases, the forces which act upon the regulatory effects may be similar between the species, or they may be divergent. Species differences may result in positive correlation between *cis* and *trans* effects, if both species are subject to directional selection on overall transcript level.

Differences in *cis* and *trans* regulation may exist, even when overall transcript levels in the compared species are not different. Genes subject to stabilizing selection for transcript level can still diverge for gene regulation, accumulating compensatory changes in *cis* and *trans* effects (TRUE and HAAG 2001). Evolutionary forces acting on regulatory elements may produce correlations between *cis* and *trans* effects. Directional selection on overall expression level is likely to result in a positive correlation between *cis* and *trans* effects. Stabilizing selection may manifest as fixation of compensatory differences resulting in negative correlation between *cis* and *trans* effects. As dominance components contribute to both *cis* and *trans* effects, strong dominance might also result in the observation of positive correlation of *cis* and *trans* effects.

We compared the correlation between *cis* and *trans* estimates for various subsets of genes in order to try to understand whether signatures of selection may be visible in the correlation pattern. Since positive correlation of *cis* and *trans* effects

is predicted for genes that are different between species, and also for genes that show dominance within the species, we consider cases where either species differences or dominance are observed jointly.

There is no significant correlation between *cis* and *trans* estimates when all genes significant for both *cis* and *trans* effects are considered, (n=93 , Rho = -0.1, P = 0.35; Table S9). For the subset of these genes which also show a significant species or dominance effect are considered no correlation among *cis* and *trans* effects is observed (n = 6 , Rho = .54, P = 0.27; Table S9). However, the genes without evidence of species and dominance effects show a significant negative correlation between *cis* and *trans* estimates (n = 87 , Rho = -0.33, P = 0.002; Table S9).