Common Pattern of Evolution of Gene Expression Level and Protein Sequence in *Drosophila*

Sergey V. Nuzhdin,* Marta L. Wayne,† Kristy L. Harmon,* and Lauren M. McIntyre‡

*Evolution and Ecology, University of California at Davis; †Department of Zoology, University of Florida, Gainesville; and ‡Computational Genomics, Purdue University, West Lafayette, Indiana

Sequence divergence scaled by variation within species has been used to infer the action of selection upon individual genes. Applying this approach to expression, we compared whole-genome whole-body RNA levels in 10 heterozygous *Drosophila simulans* genotypes and a pooled sample of 10 *D. melanogaster* lines using Affymetrix Genechip. For 972 genes expressed in *D. melanogaster*, the transcript level was below detection threshold in *D. simulans*, which may be explained either by sequence divergence between the primers on the chip and the mRNA transcripts or by down-regulation of these genes. Out of 6,707 genes that were expressed in both species, transcript level was significantly different between species for 534 genes (at $P < 0.001$). Genes whose expression is under stabilizing selection should exhibit reduced genetic variation within species and reduced divergence between species. Expression of genes under directional selection in *D. simulans* should be highly divergent from *D. melanogaster*, while showing low genetic variation in *D. simulans*. Finally, the genes with large variation within species but modest divergence between species are candidates for balancing selection. Rapidly diverging, low-polymorphism genes included those involved in reproduction (e.g., *Mst 3Ba, 98Cb*; *Acps 26Aa, 63F*; and *sperm-specific dynein*). Genes with high variation in transcript abundance within species included *metallothionein* and *hairless*, both hypothesized to be segregating in nature because of gene-by-environment interactions. Further, we compared expression divergence and DNA substitution rate in 195 genes. Synonymous substitution rate and expression divergences were uncorrelated, whereas there was a significant positive correlation between nonsynonymous substitution rate and expression divergence. We hypothesize that as a substantial fraction of nonsynonymous divergence has been shown to be adaptive, much of the observed expression divergence is likewise adaptive.

Introduction

The mechanism of adaptation remains a fundamental unsolved problem in evolutionary biology. Starting with allozyme analysis and continuing with the current emphasis on DNA sequencing and typing polymorphisms such as single nucleotide polymorphisms (SNPs), evolutionary biologists have simultaneously identified individual genes under a variety of selection regimes (purifying, balancing, etc.), and have made progress in understanding the relative contributions of various mechanisms to overall genetic variation within a population (Barton and Keightley 2002). Rigorous model-based testing has been developed, especially for DNA sequence variation and divergence (Hudson, Kreitman, and Aguadé 1987; McDonald and Kreitman 1991; Hey 1999). Applications of these models have been limited by small samples of genes rather than genome-wide estimates (Fay, Wyckoff, and Wu 2002; Smith and Eyre-Walker 2002). This direction of research may be extended for whole-genome analysis. With the sequence of *D. melanogaster* genome annotated (Adams et al. 2000) and *D. simulans* and *D. yakuba* soon to be at our disposal, sequence divergence between species will be estimable for all genes. When data on sequence variation within species is added, we will be able to contemplate molecular population genetics at the genomic level.

In part because of technological limitations, the analysis of genes involved in speciation and adaptation has focused on structural variation (i.e., amino acid sequence changes). Regulatory variation has been identified as a major source of evolutionary novelty (True and Haag 2001). In the postgenomic era, when new technologies are readily available to evolutionary geneticists, the field's attention is shifting to the evolution of expression at the RNA and protein levels (Greenspan 2001). However, although the approaches of studying DNA variation and divergence have solidified, studying RNA expression variation and divergence is still relatively novel. First attempts, nonetheless, have yielded exciting outcomes. For instance, in a study of the expression of 907 genes in multiple fishes from northern and southern populations of *Fundulus heteroclitus* and the sister taxon *F. grandis*, approximately 18% of genes varied between individuals within populations, with the differences magnified between populations (Oleksiak, Churchill, and Crawford 2002). The genetic component of this natural variation was impossible to extract because relatedness between fish had not been ascertained. Intriguingly, greater expression differences were observed between southern and northern populations within *F. heteroclitus* than between *F. heteroclitus* and *F. grandis*. These genes may be playing a role in adaptation to local environments, and their variation in populations is maintained by balancing selection. In a comparison of two laboratory lines of *Drosophila melanogaster* (OregonR and Samarkand), estimates of the genetic component of expression variation suggested that of 3,931 genes studied, genotype was a significant predictor of expression for 267 and genotype by sex interaction was a significant predictor for 431 genes (Jin et al. 2001). Also in *D. melanogaster*, the overlap between genes present in a qualitative trait loci (QTL) and differential expression among parental lines was used to propose candidate genes that influence ovariole number (Wayne and McIntyre 2002). These pioneer studies show the promise of expression analysis for characterizing natural variation and linking it to the evolution of phenotypes.

Interspecific divergence for expression has also been qualitatively described for *Drosophila* species. In a comparison between *D. melanogaster* and *D. simulans*, one line of each, differences between species and sexes were examined (Ranz et al. 2003). Ranking observed differences between mean expression levels, the authors found that interspecific differences often occurred in cases where males were observed to have higher absolute levels of transcript. In a study of *D. simulans*, *D. mauritiana*, and their F_1 hybrid, the F_1 hybrid male-specific genes were found to have lower expression than either of the two parents (Michalak and Noor 2003). In both studies, sex-specific effects were reported. These studies suggest avenues to explore for the identification of genes whose expression variation contributes to reproductive isolation.

Here, we compare expression between *D. melanogaster* and *D. simulans*, extending previous work by considering multiple heterozygous male genotypes. We measured expression in males of 10 genotypes of each species. In *D. simulans*, we measured each of 10 genotypes separately, and in *D. melanogaster*, we measured a pool of 10 genotypes. All genotypes (individual or pool) were isolated and hybridized three times. This design allows us to calculate and compare genetic components of line (intraspecific) and species (interspecific) variation at the transcript level. To approximate expression in wild flies as closely as possible, we used repeated heterozygous crosses between effectively isogenic lines recently extracted from nature. We studied expression in males because males contribute disproportionately to reproductive isolation in the *D. melanogaster* subgroup (i.e., Haldane's rule [Orr and Turelli 2001]) and because previous work suggests that expression differences might be more prevalent in this sex (Michalak and Noor 2003; Ranz et al. 2003). We also compare divergence between species in expression and sequence. Although the silent substitution rate is not correlated with divergence in expression, the amino acid substitution rate is positively correlated with divergence in expression. We, thus, provide a genome-wide approach to identifying candidate genes potentially responsible for adaptation and speciation in *D. simulans* and *D. melanogaster* on the basis of rapid divergence in expression and demonstrate that a large fraction of the genome may be involved in adaptation via expression. Two lines of evidence suggest that the differences in expression between species we observe are not artifacts of sequence divergence between the *D. melanogaster*–designed primers and *D. simulans* transcripts: (1) the average among hybridization correlations were the same between *D. simulans* (0.95) and *D. melanogaster* (0.94), indicating high reproducibility, and (2) the silent substitution rate (dS) is uncorrelated with expression divergence.

Materials and Methods
Drosophila Lines

The stocks were obtained from flies caught in Wolfskill Orchard (Winters, Calif.). Ten isogenic lines of *D. melanogaster*, sib mated for 40 generations then kept as stocks, and 10 isogenic lines of *D. simulans*, in their 25th to 29th round of sib mating, were chosen for intraspecies crosses. Flies were kept in vials of standard cornmeal medium with yeast, five pairs per vial, at 25°C on a continuous light cycle for one generation. Parents were cleared after 5 days, and virgin males and females were then collected from these for round-robin crosses. Five 1-day-old to 2-day-old males of the first line were mated with five same-age females of the second line. Similarly, males of the second line were mated with females of the third line, and so on. Bottles were kept at 25°C on a continuous light cycle. Parents were cleared after 5 days. Virgin females and males were collected and then frozen after 3 days. There were two complete identical blocks.

RNA Expression and Hybridization

Three replicate chips for each of the *D. simulans* genotypes and the pool of the *D. melanogaster* genotype were used. RNA was extracted from males reared in two blocks. One replicate was made from the first block and two were made from the second block. Total RNA extractions were made with Trizol from five F_1 males for each *D. simulans* cross. Twenty males, two from each of the 10 *D. melanogaster* crosses, were pooled for extraction. We used Affymetrix Genechip arrays containing 13,966 features representing the genome of *D. melanogaster* (see Wayne and McIntyre [2002] and Michalak and Noor [2003] for other examples of use of this platform). Affymetrix Eukaryotic target preparation protocol was followed, using 5 mg of the total RNA to synthesize double-stranded cDNA. Fragmented cRNA were biotin labeled. The hybridizations and readings were made in the University of California, Davis Affymetrix core facility. Data were quantified using Affymetrix MAS5 software. Affymetrix Genechip arrays consist of a series of probes 25 bases long (25mers) from each target sequence. For the *Drosophila* array, the Affymetrix Genechip was constructed using version 1.0 of the FlyBase annotation. A probe pair consists of a "perfect match" (PM) and a "mismatch" (MM). A PM is a sequence that is 100% concordant with the sequence in FlyBase, whereas an MM is the same sequence, except that the middle nucleotide (13) of the oligonucleotide is a mismatch. The MM is expected to provide an estimate of nonspecific hybridization. Affymetrix MAS5 software estimates gene expression as the Tukey biweight sum of the (PM–MM), which is referred to as the "average difference." For the *Drosophila* Genechip, there are 14 probe pairs for each "feature." They are referred to as features rather than genes because they include predicted as well as confirmed genes. For example, of the 13,966 total features on the *Drosophila* Genechip, about 8,000 are from genes with confirmed EST/cDNA, and the remaining features are from gene prediction algorithms. If a feature does not show significantly higher signal in the PM pair than the MM pair, the feature is deemed to be "absent" based on the default parameters of the Affymetrix algorithm. The comparison between PM and MM in itself is a control for hybridization mismatches, whether because of sequence polymorphism within *D. melanogaster* or because of sequence divergence between *D. melanogaster* and

D. simulans. Further details of the Affymetrix platform and their algorithms can be found at www.affymetrix.com.

Data Analysis

Expression values for each chip were normalized to the chip median and then log transformed. Features that were considered absent, according to the Affymetrix call, for all hybridizations were excluded from further analysis ($n = 5,108$). For those transcripts absent in all replicates of *D. simulans* but present in at least one replicate of *D. melanogaster*, we report the mean expression in *D. melanogaster*. For those transcripts absent in all replicates of *D. melanogaster* but present in at least one replicate of *D. simulans*, we report the mean expression in *D. simulans*. For the remaining features, those present in both species, the transcript level was modeled in an ANOVA framework for each feature according to the model $Y_{ijk} = \mu + \varsigma_i + \lambda_{ij} + \varepsilon_{ijk}$ where Y is the normalized, log transformed transcript level, μ is the overall mean of the transcript, ς is the effect of species ($i = 1, 2$), and λ is the effect of line nested within species ($j = 1,\dots10$). We considered both species and lines random and estimated variance components using a method of moments approach where negative estimates were assumed to be zero. We then compared the F ratio for the line effect (a test of the significance of the within-line variance) to the F ratio for the species effect (a test of significance of the between-species variance). This comparison allowed us to also determine whether the differences among lines and among species were statistically significant. We also compared the variance components from the line and species. All analyses were conducted in SAS version 8.2 (SAS Institute, Cary, NC).

Cross-Species Hybridizations

Because the GeneChip is based on sequences from *D. melanogaster*, interspecific sequence divergence in primers could potentially contribute to the very interspecific differences in expression that we are attempting to measure. Accordingly, we carefully examined the behavior of cross-species hybridization. We first checked our data for within-group reproducibility. We found that the average correlation among the *D. melanogaster* hybridizations was 0.94, and the average correlation among the *D. simulans* hybridizations was 0.95. Thus, reproducibility is very similar for the two species; that is, random noise within *D. simulans* is no higher than within *D. melanogaster*.

Further, sequence divergence between *D. simulans* and *D. melanogaster* is approximately 2%. Correspondingly, we expect that for a probe of length 25, the chance of a single MM (0.02*25) is about 0.5, and so, on average, approximately half of the 25mer probes (PM oligonucleotides) should have identical sequence with between *D. melanogaster* and *D. simulans* transcripts. If the *D. simulans* transcript has a single MM with the PM oligonucleotide, it should have two MMs with the corresponding MM oligonucleotide. The PM–MM signal difference, therefore, should still reflect expression signal

higher than background noise. If this is not the case, the MAS5 algorithm will declare that feature "absent." Thus, the consequence of a poor cross-species hybridization should be a higher number of "absent" features. We compared the detection rates of the two types of hybridizations within and across species, and we found that, overall, 1,172 of the *D. melanogaster* features were absent, whereas 972 of the *D. simulans* features were absent. To minimize potential confounding effects of sequence divergence, we focus the majority of our analysis on features that are deemed "present" in both species.

Relating Expression Divergence to Gene Annotation and Sequence Divergence

We used the available annotation (FlyBase annotation 3.0 citation) to separate the complete list of genes into two groups: directly annotated genes (AGs) and not yet annotated genes (NAGs). A gene was considered directly annotated when it carried a name indicating that mutations of large phenotypic effect have been detected. NAGs consisted of genes that were given no name based on phenotype, but only have CG numbers. We compared the effects of including genes whose function is predicted solely by homology in the NAG classification and found that inclusion or exclusion of these genes had no effect on the analyses described below. We expect that, in general, genes with classical descriptions based on observed mutant phenotypes might be under different selective constraints than genes without previously identified mutations, given the long history of mutagenesis in *D. melanogaster*. Thus, we might expect expression divergence to be quantitatively different between these groups of genes if expression divergence is related to selective constraint.

Likewise, if expression divergence is related to selective constraint, we might expect patterns of sequence divergence to be correlated with expression divergence. Specifically, if genes are evolving rapidly in amino acid space, we might expect them to also be evolving rapidly in expression space. In contrast, silent substitutions are expected to be nearly neutral and, thus, represent a control for overall sequence divergence but should not be correlated with expression divergence if expression is generally under selection. We identified 270 sequences that had been evaluated for synonymous and nonsynonymous substitution rates in the literature. Of these, 237 were represented on the array. For 156 of these 237 features, the Affymetrix presence-absence algorithm scored expression as present in both species, 21 features as "present" only in *D. simulans*, and 18 features as "present" only in *D. melanogaster*. Forty-two genes were "absent" in both species. The 156 features where expression was found to be "present" in both species were tested for correlation between synonymous and nonsynonymous substitution rate and the species component of variance using Pearson's correlation coefficient.

Although it is most conservative to consider only genes expressed in both species, the most biologically interesting genes may be expected to come from the groups where expression is "absent" in one of the two species because this group includes genes with the most rapid

Table 1
Transcript Level in *D. melanogaster* of the Genes Not Detected to Be Expressed in *D. simulans*

Gene Name	Expression Level in *D. melanogaster*
Male-specific transcript 35Ba	3.14
Accessory gland peptide 63F	2.94
Necrotic, serpin, antifungal humoral response	2.43
Serine protease inhibitor 3	2.21
Shanty	2.16
Sperm-specific dynein intermediate chain	1.66
Accessory gland-specific peptide 26Aa	1.65
Tumor-suppressor protein 101	1.46
Glutathione S transferase E1	1.44
Odorant-binding protein 56a	1.33
Iron regulatory protein 1B	1.30
Catecholamines up	1.20
Mitochondrial ribosomal protein L16	1.18
Skittles, 1-phosphatidylinositol-4-phosphate 5-kinase	1.11
Trypsin 29F, Serine proteases	1.08
Rab-protein 4, RAB small monomeric GTPase	1.04
Odorant-binding protein 19b	1.04
Hyperkinetic, voltage-sensitive potassium channel	1.02
Protein on ecdysone puffs	0.97
Casein kinase IIβ_2 subunit	0.97
Phospholipase A$_2$ activator protein	0.96
Mitochondrial ribosomal protein S18a	0.95
Ankyrin	0.94
Rhythmically expressed gene 3, dihydropyrimidine dehydrogenase	0.91
Prospero, transcription factor	0.90

expression divergence. Thus, we examined the set of 195 (156 + 21 + 18) for correlation between synonymous and nonsynonymous substitution. For "absent" features, we were concerned that the average difference might not be a valid measure of expression level. However, the alternative would be to assume that the expression observed in one species or the other is effectively zero and to substitute a zero value for the estimated average difference. The impact of substituting a zero will be more extreme than using the estimated average difference values, which are by definition positive, thus, biasing our results in the direction of finding significant divergences. Accordingly, we proceeded with the average difference values, as this was a more conservative procedure.

We calculated two measures of expression divergence. First, we considered the variance component estimates of the between-species component of variance for the 156 genes. We compared this mean difference with a second measure of expression divergence, the absolute value of the observed difference between the mean expression levels of the two species to determine whether these metrics were comparable. The correlation between these two metrics was excellent (0.91). We then proceeded to perform ANOVA analyses and calculate variance components for the 39 genes that were "absent" in one of the two species. We compared the variance component estimates for this subset of features, with the expression divergence as calculated by mean average difference, and again found excellent correspondence between these two

measures (0.91). We then repeated our comparison of synonymous and nonsynonymous substitution rate and the species component of variance on the full set of 195 genes.

Results and Discussion
Genes Not Found Expressed Either in *D. simulans* or in *D. melanogaster*

Expression levels were individually measured for 10 genotypes of *D. simulans* and for a sample pooled from 10 genotypes of *D. melanogaster*. On Affymetrix chips, each gene is represented by multiple unique probes. Gene expression level is reconstructed from the numerous hybridization signals from all probes for each gene (De Gregorio et al. 2001), minimizing the effect of sequence divergence. Of the 13,966 features on the chip, 5,108 genes were absent in both species, 972 genes were detected only in *D. melanogaster*, 1,179 genes were detected only in *D. simulans*, and 6,707 genes were detected in both species. Genes only expressed in *D. melanogaster* were sorted by mean observed transcript level in *D. melanogaster*. The 25 annotated genes with the highest expression in *D. melanogaster* are presented in table 1. The remaining 947 genes are presented in table 1 in Supplementary Material online.

Why do we not detect transcription for the above genes in *D. simulans*? Given that the oligonulceotide sequences on the Affymetrix chip were designed for *D. melanogaster*, it is possible that genes identified as having high expression divergence are, in fact, hybridizing poorly to the chip becuse of sequence divergence between *D. melanogaster* and *D. simulans*. There are several lines of evidence that argue against this interpretation across the sample, although it is no doubt true for some genes. Most importantly, the Affymetrix software calculates expression values as a function of the ratio of a "perfect match" between the oligonucleotide probe and the cDNA sequence relative to a deliberate "mismatch," rather than as an absolute value. This ratio should thus be robust to sequence variation between species elsewhere in the oligonucleotide sequence. Further, the reproducibility between the two species is virtually identical (0.94 in *D. melanogaster* and 0.95 in *D. simulans* [see *Materials and Methods*). Thus, if we expect hybridization error to be manifested as random noise in the sample, there is no evidence of a between-species difference for noise. Finally, as argued below, the silent substitution rate dS is uncorrelated with our estimate of expression divergence among genes. Because sequencing of the *D. simulans* genome has already been initiated (C. H. Langley, personal communication), the hypothesis of sequence divergence will be directly testable in the future.

Alternatively, the above genes might be silenced or down-regulated in *D. simulans*. It has been argued that duplicated genes acquire new functions (Lynch and Force 2000), whereas preexisting genes become less essential to none essential or obsolete (Wagner and Schwenk 2000; Yang, Gu, and Li 2003). Here, we ask whether expression in *D. simulans* is more frequently absent for the genes not known to be essential in *D. melanogaster*. Many decades of genetic analysis of *D. melanogaster* have yielded a list

of genes whose mutations have strong phenotypic effects—annotated genes (AGs). The rest of genes we call not yet directly annotated (NAGs). If there are nonessential genes in *Drosophila*, they will not be represented among AGs, but they will be represented among NAGs. We hypothesize that the NAGs would, on average, experience lesser selection constraints; that is, their transcript level and sequence will evolve faster. Correspondingly, we expect larger proportion of NAGs (and smaller proportion of AGs) among features whose expression is detected in *D. melanogaster* but not in *D. simulans* in comparison with the genes expressed in both species. As expected, AGs were significantly underrepresented among genes without detectable expression in *D. simulans* in comparison with genes found expressed in both species (13.8% compared with 18.7%, $\chi^2 = 11.06$). Thus, expression of genes with large-effect mutations diverges more slowly than the expression of not yet annotated genes.

Interestingly, the two genes with the highest expression in *D. melanogaster*, as well as two additional genes from table 1, have testes-specific function. Fast evolution of testes-specific genes between species is expected for multiple reasons. Hybrid male sterility is frequently the first mechanism of postmating isolation to evolve during speciation because of partial recessivity of Muller-Dobzhansky incompatibilities, faster evolution of the X chromosome, faster evolution of genes involved in spermatogenesis, or combinations of these mechanisms (Orr and Turelli 2001). Further, antagonistic coevolution between the sexes, in which males manipulate females into accelerated egg laying after mating to maximize male fitness at the expense of female fitness and females evolve resistance to male manipulation to maximize their own fitness, should drive the faster evolution of reproduction related functions (Holland and Rice 1999). Other examples of this effect include accelerated sequence divergence (Swanson and Vacquier 2002), and expression divergence (Ranz et al. 2003). Fast sequence divergence of odorant-binding proteins might be explained by their involvement in adaptation to new environment or in mating choice (Hekmat-Scafe et al. 2002; Vogt 2002). Finally, the genes involved in host defense from parasites (like *necrotic*) diverge faster as a result of Red Queen mechanisms (Begun and Whitley 2002). The sequence divergence of other genes from table 1 has also been studied: ankyrins (Maine, Lissemore, and Starmer 1995), serine protein inhibitors (Okuyama, Tachida, and Yamazaki 1997), and casein kinase (Kalmykova, Dobritsa, and Gvozdev 1997). Noteworthy, the latter study implied fast evolution of new function, casein kinase, of the *Su(Ste)* region of the Y chromosome.

At first glance, the larger number of genes with detectable expression in *D. simulans* but not in *D. melanogaster* (1,172) is surprising. Indeed, if sequence divergence decreases hybridization signal, *D. melanogaster* genes should have more frequently appeared overexpressed than unexpressed. One explanation is that for the genes with a transcript level close to the detection threshold, the power of detection (finding at least one array with the "present" call [see *Materials and Methods*]) is much higher in *D. simulans* than in *D. melanogaster* because of
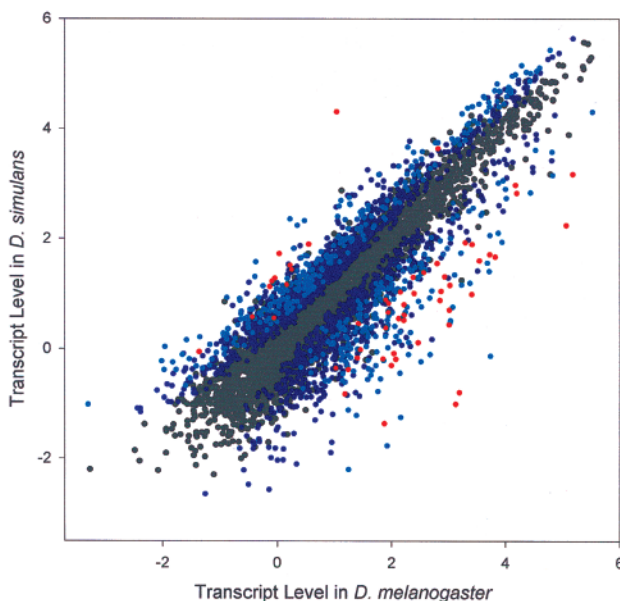
Fig. 1.—Expression levels in *D. melanogaster* compared with *D. simulans*. The genes with difference in expression levels significant at the level $P < 8.24 \times 10^{-6}$ (Bonferroni) in red, at the level $P < 0.001$ in light blue, and at the level $P < 0.05$ in blue.

the 10-fold larger number of arrays for *D. simulans*. In fact, of the 1,172 features detected in *D. simulans*, 453 are detected in exactly one of the replicates, and an additional 475 are present in less than 20% of the arrays. There were only 53 features that were consistently (more than 50% of the time) detected in *D. simulans* (see table 2 in Supplementary Material online for a complete list of genes) that were not detected in *D. melanogaster*. Which of them are truly turned off in *D. melanogaster* is difficult to conclude in this study design.

Divergence Between Transcript Levels of Genes Found Expressed in Both Species

For the 6,707 genes present in both species, the mean expression in *D. melanogaster* is strongly correlated with mean expression in *D. simulans* (figure 1, the transcript levels were strongly correlated among genes; $r^2 = 0.87$, $P < 0.0001$). However, the difference in expression levels between species was significant (using the ANOVA model described in *Materials and Methods*) for 2,294 genes out of 6,707: for 1,760 genes at the level $P < 0.05$ (colored blue in figure 1), an additional 478 genes were significant at a threshold of 0.001 (colored light blue in figure 1), and 56 features are significant at the Bonferroni level $P < 7.46 \times 10^{-6}$ (i.e., Bonferroni corrected 0.05/6,707 and colored red in figure 1). Significance tests were performed on average differences where each individual chip is standardized to the median. Thus, an overall difference in the strength of the hybridization signal is eliminated as a source for differences between *D. simulans* and *D. melanogaster*. Genes are, therefore, significant for the divergence test if they had either higher estimated standardized expression in *D. melanogaster* than in *D. simulans* or lower standardized expression in *D. melanogaster*

Table 2
Genes Significantly Different in Expression Between *D. melanogaster* and *D. simulans*

Gene Name	Expression Level in *D. melanogaster*	Expression Level in *D. simulans*	Significance of Expression Difference
Glyceraldehyde 3 phosphate dehydrogenase 1	5.18	3.17	9.23×10^{-9}
High mobility group protein D	0.17	1.16	1.50×10^{-7}
Minute (2) 21AB	3.42	1.90	6.93×10^{-7}
Cut up, microtubule motor	2.13	0.55	1.27×10^{-6}
Kettin, cell adhesion	0.22	1.51	1.83×10^{-6}
ATP synthase-β	1.88	−1.36	2.71×10^{-6}
Vacuolar H[+] ATPase G-subunit	2.82	3.63	3.07×10^{-6}
Odorant-binding protein 57c	1.19	−0.83	3.12×10^{-6}
Casein kinase I, α-subunit	0.55	1.90	3.13×10^{-6}
Male-specific RNA 98Cb	3.30	1.93	5.28×10^{-6}
Eukaryotic initiation factor 4G	1.25	−0.37	5.75×10^{-6}
Tetratricopeptide repeat protein 2	1.40	−0.17	6.72×10^{-6}

Table 3
Genes Significantly Varying for Expression Among *D. simulans* Genotypes

Gene Name	Expression Level in *D. simulans*	Line Component of Variance	Significance of Variance
robl62A, dynein ATPase	1.28	0.24	1.15×10^{-8}
PebIII, Ejaculatory bulb protein III	4.66	0.13	1.68×10^{-8}
β-Tubulin at 85D	1.19	0.17	8.18×10^{-8}
Cyp6w1, cytochrome P450	2.37	0.23	2.64×10^{-7}
Cyp4e2, cytochrome P450-4e2	2.48	0.21	7.55×10^{-7}
Transaldolase, pentose-phosphate shunt	2.86	0.10	7.93×10^{-7}
Jheh2, Juvenile hormone epoxide hydrolase 2	3.25	0.12	8.42×10^{-7}
eIF-4a, Eukaryotic initiation factor 4a	3.45	0.16	1.05×10^{-6}
Acp26Ab, Accessory gland-specific peptide	0.61	0.43	1.26×10^{-6}
Obp99c, Odorant-binding protein 99c	4.16	0.17	1.72×10^{-6}
α-Esterase 7	0.27	0.15	1.87×10^{-6}
Uro, Urate oxidase	1.09	0.43	6.88×10^{-6}

than in *D. simulans*. Of the genes significant by Bonferroni, 43 were higher in *D. melanogaster* and 13 were higher in *D. simulans* at $P < 0.001$, 176, and 282, respectively and at $P < 0.05$, 680, and 1,080, respectively. Thus, significant genes were not generally the result of disproportionately lower standardized expression in *D. simulans*, as would be expected if expression divergence were purely an artifact of sequence divergence.

The AGs significant using the Bonferroni criteria are listed in table 2. The complete list of genes and the results of the ANOVA analyses can be found in table 3 of Supplementary Material online. Interestingly, the sequence divergence of several of the genes significant at the Bonferroni level, or their close relatives, has been studied before via sequencing: ATP synthase (Wang et al. 2002), ATPase (Rand and Kann 1996; Garvey and Malcolm 2000), odorant-binding proteins (Hekmat-Scafe et al. 2002; Vogt 2002), and casein kinase (Kalmykova, Dobritsa, and Gvozdev 1997). Differences in transcript abundance might be caused by (1) expression level changing because of mutations in the promoter regions of a gene, (2) expression level changing because of evolution of transcription factors, (3) expression level remaining unchanged but the tissue where expression takes place expanded or contracted, or (4) expression level remaining unchanged but the timing of expression has diverged between species. Currently available data do not allow us to distinguish among these hypotheses.

Genetic Variation in Expression Levels

Assaying multiple genotypes in *D. simulans* in replicates allowed us to estimate genetic variance for the transcript level within species. Out of 6,707 features examined, 1,136 genes were found to be significantly

different among lines at the level $P < 0.05$, an additional 218 at the level $P < 0.001$ and 39 genes at the level $P < 7.46 \times 10^{-6}$ (see table 3 for annotated genes in this group). Genetic variation of several these genes, or their homologs, had previously being investigated, including tubulins (Akashi, Kilman, and Eyre-Walker 1998; Nielsen and Raff 2002), cytochromes P450 (Hallstrom, Magnusson, and Ramel 1982), esterases (Bublii, Imasheva, and Lazebnyi 1994), and accessory gland proteins (cf. Clark et al. [1995] and Cirrera and Aguade [1997]).

Transcript abundance of particular genes in *D. simulans* might be under stabilizing, directional, or balancing selection. Genes whose expression is under stabilizing selection should exhibit reduced genetic variation within species and reduced expression divergence between species. Expression of genes under directional selection in *D. simulans* should diverge from *D. melanogaster* while showing low genetic variation in *D. simulans*. Genes with extensive expression variation within species but modest expression divergence between species might be under balancing selection. Stabilizing, directional, or balancing selection might be a direct result of selection for cellular protein activity or an indirect result of selection for organ size (i.e., increase in cell number), despite a constant per nucleus expression level. Finally, the genes under selection relaxation might both diverge and vary. Relaxed selection might be a consequence of lesser gene activity per cell being sufficient (True and Haig 2001). While explicit molecular evolution models for the statements above await further development, we can contrast expression variation and divergence to identify the genes possibly belonging to the four different categories. We plot the line and species variance components and their F statistics in figure 2. The striking L-shape of the distribution (but see Sokal [1976]) suggests the majority of genes neither diverge nor vary (figure 2, bottom left near or at the origin, black colored). Their level of expression is likely to be controlled largely
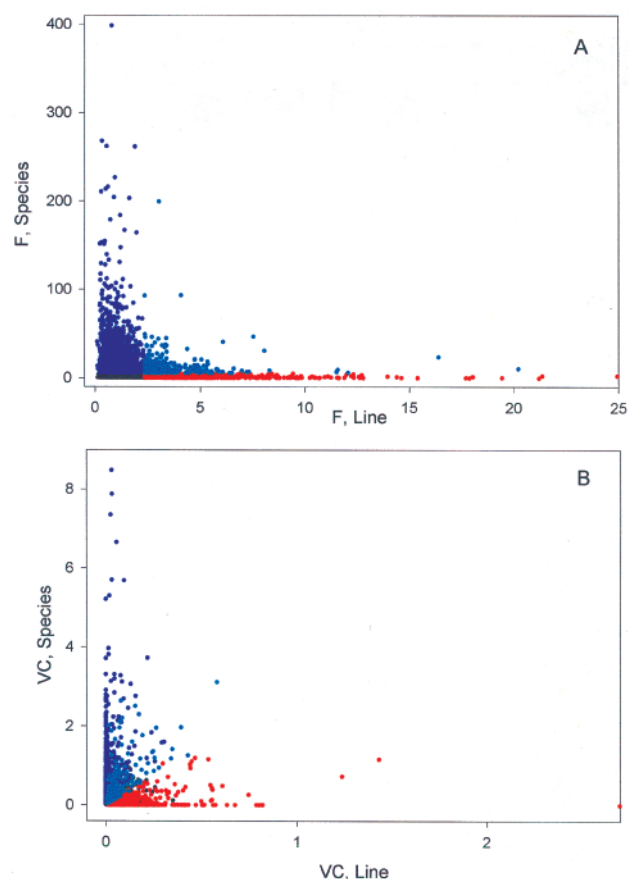
FIG. 2.—Variation in the level of expression among 10 *D. simulans* genotypes versus divergence between *D. melanogaster* and *D. simulans*. The upper plot (*A*) depicts *F*-statistic for the significance of genotype and species variance components, and the lower plot (*B*) represents the absolute value of those components. The genes exhibiting significant genetic variation among *D. simulans* genotypes are in red, those exhibiting significant genetic variation between *D. melanogaster* and *D. simulans* are in blue, and those exhibiting both conditions are in light blue. Significance was defined at a threshold of 0.05.

Table 4
Genes Strongest Varying Within *D. simulans* but Minimally Diverging Between Species

Gene Name[a]	Mean Sum of Squares Between Species	Mean Sum of Squares Within Species
Hemomucin, RNA binding	1.36×10^{-7}	0.0833
Splicing factor 1	3.40×10^{-7}	0.0533
Small ribonucleoprotein Sm D3	2.39×10^{-7}	0.0165
Ras oncogene at 85D	2.92×10^{-6}	0.0817
Intronic Protein 259	7.30×10^{-6}	0.148
Metallothionein A	8.24×10^{-6}	0.122
Iap2, Inhibitor of apoptosis 2	4.04×10^{-6}	0.0362
sesB, stress-sensitive B	1.85×10^{-5}	0.151
Adenosine 2, phosphoribosylformylglycinamidine synthase	1.66×10^{-5}	0.0931
Taf12, TBP-associated factor 12	3.18×10^{-5}	0.108
Burgundy, GMP synthase (glutamine hydrolyzing)	0.000234	0.474
Hairless, transcription corepressor	7.16×10^{-5}	0.125
Scarlet, eye pigment precursor transporter	0.000182	0.191
Scab, cell adhesion receptor	0.00115	1.07
osa, DNA binding	0.000872	0.781
Purity of essence, calmodulin binding	0.000135	0.115
Aph-4, Alkaline phosphatase 4	0.000207	0.135
Aprt, Adenine phosphoribosyltransferase	0.000281	0.165
HLH106, Helix-loop-helix protein 106	8.11×10^{-5}	0.0473
RpS26, Ribosomal protein S26	0.000112	0.0628

[a] Because statistical significance cutoff for inverse of *F* is difficult to evaluate, we report AGs from 100 with the smallest ratio of the mean sum of squares for species by the mean sum of squares for line.

by purifying selection. The genes that do not diverge while showing evidence of intraspecific variation (bottom right, red colored) point to potential adaptive intraspecific variation or balancing selection. Another group of genes—low in intraspecific variation but divergent between species (top left, blue colored)—are candidate adaptation or speciation genes. Finally, genes that have variation both within and between species (light-blue colored) might be enriched for those under relaxed selection.

As a formal procedure for identifying genes under directional and balancing selection, we used the *F*-statistic for species; that is; grouped genes by the largest or smallest ratio of the mean sum of squares for species ($MS_{species}$) to the mean sum of squares for line (MS_{line}). For the large values of this ratio, interpretation is straightforward. These are the genes for which expression divergence is much larger than expected, given their expression polymorphism, and thus, are diverging adaptively between species or are speciation genes per se (table 2). The genes with the lowest ratio of $MS_{species}$ to MS_{line} are candidates for adaptive variation within species or for balancing selection

(table 4). *Metallothionein* and *Hairless* have previously been implicated in natural variation for metal susceptibility (Symonds and Gibson 1992; Posthuma and Vanstraalen 1993) and for bristle number (Long et al. 1996; Lyman and Mackay 1998), with gene-by-environment interaction proposed to contribute to segregation of natural alleles.

Comparison of Sequence and Expression Divergence

How are expression and sequence divergence related? Out of the sequenced *D. simulans* genes summarized by Betancourt and Presgraves (2002), 237 genes are represented on the Affymetrix array, and 195 are expressed in at least one species, of which 156 are present in both species. For the latter group, we correlated transcript level divergence measured in three ways: absolute difference of transcript levels (figure 3*A*), species variance component (figure 3*B*), and *F*-statistic for expression level difference between species (figure 3*C*) with divergence in synonymous (*dS*) and nonsynonymous (*dN*) sites. Synonymous substitutions do not change the amino acid sequence and are expected to be nearly neutral (Akashi 1995); thus, these accumulate rapidly between species. We have found no significant correlation between expression divergence and *dS* (difference in transcript level 0.03, *P* = 0.67; variance component −0.06, *P* = 0.94; *F*-statistic 0.06, *P* = 0.45). In contrast, we observed a positive correlation between expression divergence and nonsynonymous substitutions, which alter proteins and thus are expected to be under stronger selection than synonymous substitutions (difference in transcript level 0.17, *P* = 0.0029;
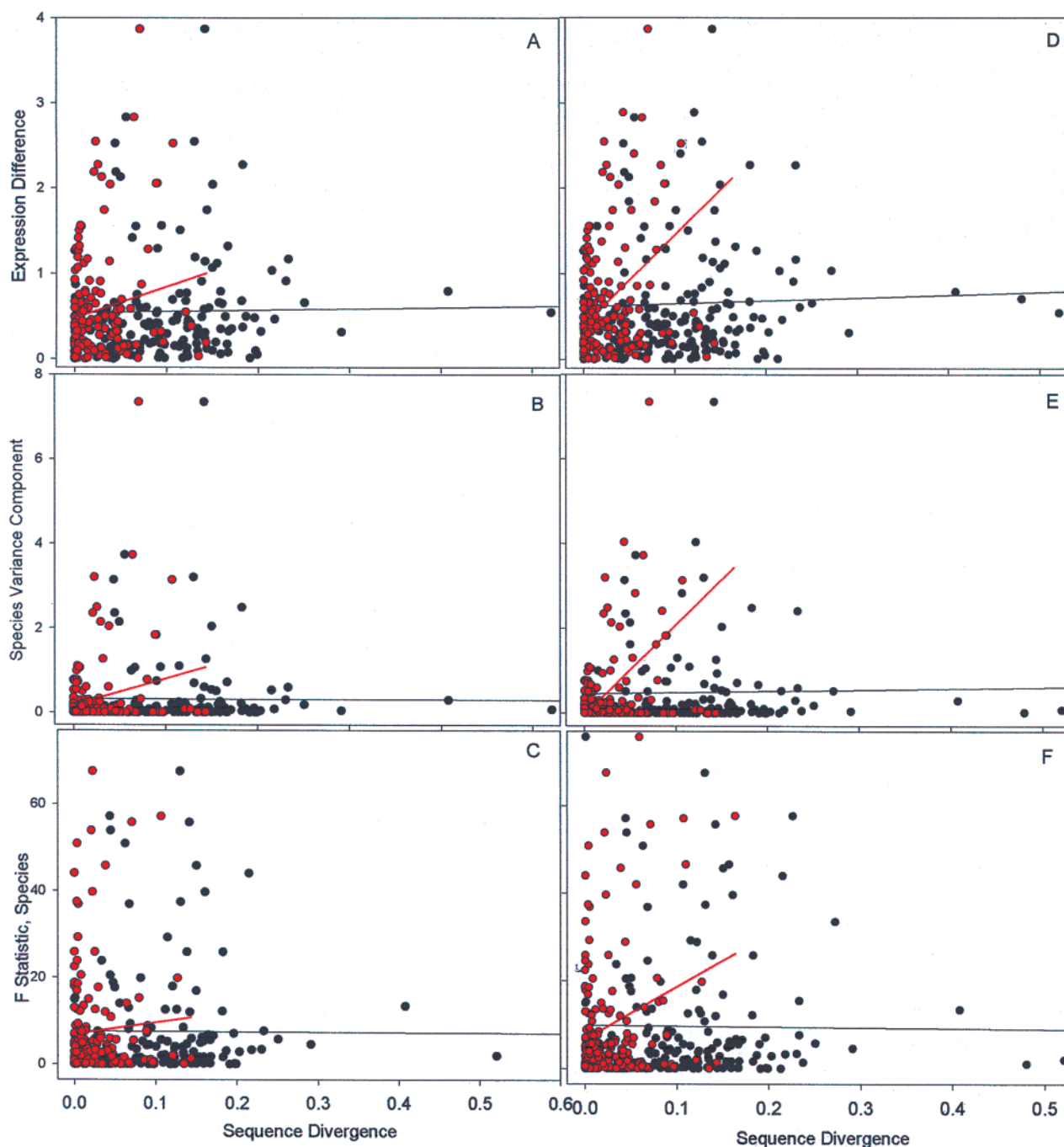
F_{IG}. 3.—Correlation of transcript level divergence: absolute difference of transcript levels (*A* and *D*), species variance component (*B* and *E*), and *F*-statistic for expression level difference between species (*C* and *F*), with divergence in synonymous (black) and nonsynonymous (red) sites. Plots on the left are based on 156 genes, and plots on the on the right are based on 198 genes, as explained in the *Results and Discussion* section.

variance component 0.22, $P = 0.0056$; F-statistic 0.06, $P = 0.44$). Note that for the genes studied here, the ratio of synonymous to nonsynonymous substitutions is approximately 5:1 (Betancourt and Presgraves 2002), whereas the number of synonymous to nonsynonymous sites is roughly 1:2 (Hedrick 2000). Thus, overall sequence divergence between species is governed by synonymous divergence. Because synonymous divergence appears uncorrelated with expression divergence, we conclude that overall

sequence divergence is unlikely to have overwhelming effect on estimation of expression level (see *Materials and Methods* for more explanations). However, we cannot rule out the possibility that sequence divergence contributes to our observations.

The most divergent genes are those that show expression in one species but not in another. The initial comparison of genes expressed in both species will be an underestimate of the correlation, as we are ignoring the

fastest divergent genes; that is, those with transcript levels found to be "absent" in one of the two species. We examined the behavior of the estimators carefully (see *Materials and Methods*) and then proceeded to examine the relationship between dN and dS with the full set of genes ($n = 195$). As expected, the magnitude of the effect we observe is enhanced, although the same general trend is apparent. Transcript level divergence remains uncorrelated with dS (difference in transcript level 0.05, $P = 0.43$; variance component 0.03, $P = 0.68$; F-statistic -0.01, $P = 0.84$), whereas the correlations with dN become much stronger (difference in transcript level 0.42, $P < 0.0001$; variance component 0.48, $P < 0.0001$; F-statistic 0.22, $P = 0.0018$).

We hypothesize that both dN and expression level divergence are governed by similar selective regimes. Nonsynonymous substitutions are overrepresented in genes for which selection is relaxed or in genes under directional selection (Fay, Wyckoff, and Wu 2002; Smith and Eyre-Walker 2002). If selection constraints are relaxed, mutations affecting both protein sequence and expression accumulate (including in upstream regulatory genes). We then expect to see strong sequence divergence in nonsynonymous sites and expression, as well as abundant intraspecific variation in gene sequences and expression. In contrast, if increased gene product activity is selected, mutations changing amino acids and affecting expression will be preferentially fixed. We then expect strong dN and expression divergence but little intraspecific nonsynonymous and expression variation (Hudson, Kreitman, and Aguadé 1987; Fay, Wyckoff, and Wu 2002; Smith and Eyre-Walker 2002). Further work is necessary to address the relative contributions of expression and protein sequence divergence to selection.

We have identified a group of rapidly evolving genes between closely related species. Accounting for intraspecific transcript variation allows us to select the genes whose fast evolution is explained by natural selection rather than by selection relaxation (table 2). Because a substantial fraction of dN is adaptive (Fay, Wyckoff, and Wu 2002; Smith and Eyre-Walker) and expression divergence is correlated with dN, it seems likely that a substantial fraction of expression divergence is likewise adaptive.

Literature Cited

Adams, M. D., S. E. Celniker, R. A. Holt, C. A. Evans, J. D. Gocayne, P. G. Amanatides, S. E. Scherer, and P. W. Li. 2000. The genome sequence of *Drosophila melanogaster*. Science **287**:2185–2195.

Akashi, H. 1995. Inferring weak selection from patterns of polymorphism and divergence at "silent" sites in *Drosophila* DNA. Genetics **139**:1067–1076.

Akashi, H., R. M. Kilman, and A. Eyre-Walker. 1998. Mutation pressure, natural selection, and the evolution of base composition in *Drosophila*. Genetica **103**:49–60.

Barton, N. H., and P. D. Keightley. 2002. Understanding quantitative genetic variation. Nat. Rev. Gen. **3**:11–21.

Begun, D. J., and P. Whitley. 2000. Adaptive evolution of relish, a *Drosophila* NF-kappa B/Ikappa B protein. Genetics **154**:1231–1238.

Betancourt, A. J., and D. C. Presgraves. 2002. Linkage limits the power of natural selection in *Drosophila*. Proc. Natl. Acad. Sci. USA **99**:13616–13620.

Bublii, O. A., A. G. Imasheva, and O. E. Lazebnyi. 1994. Geographic variation of *Adh, alpha-Gpdh,* and *Est-6* in Eurasian natural populations of *Drosophila melanogaster*. Genetika **30**:467–477.

Cirrera, S., and M. Aguade. 1997. Evolutionary history of the sex peptide *Acp 70A* gene region in *Drosophila melanogaster*. Genetics **147**:189–197.

Clark, A. G., M. Aguade, T. Prout, L. G. Harshman, and C. H. Langley. 1995. Variation in sperm displacement and its association with accessory gland protein loci in *Drosophila melanogaster*. Genetics **139**:189–201.

De Gregorio, E., P. T. Spellman, G. M. Rubin, and B. Lemaitre. 2001. Genome-wide analysis of the *Drosophila* immune response by using oligonucleotide microarrays. Proc. Natl. Acad. Sci. USA **98**:12590–12595.

Fay, J. C., G. J. Wyckoff, and C. I. Wu. 2002. Testing the neutral theory of molecular evolution with genomic data from *Drosophila*. Nature **415**:1024–1026.

Garvey, C. F., and C. A. Malcolm. 2000. *Anopheles stephensi Dox-A2* shares common ancestry with genes from distant groups of eukaryotes encoding a *26S* proteosome subunit and in a conserved gene cluster. J. Mol. Evol. **50**:497–509.

Greenspan, R. J. 2001. The flexible genome. Nat. Rev. Genet. **2**:382–387.

Hallstrom, I., J. Magnusson, and C. Ramel. 1982. Relation between the somatic toxicity of dimethylnitrosamine and a genetically determined variation in level and induction of cytochrome P450 in *Drosophila melanogaster*. Mut. Res. **92**:161–168.

Hedrick, P. H. 2000. Genetics of populations. Jones and Bartlett Publishers, Inc., Sudbury, Mass.

Hekmat-Scafe, D. S., C. R. Scafe, A. J. McKinney, and M. A. Tanouye. 2002. Genome-wide analysis of the odorant-binding gene family in *Drosophila melanogaster*. Genome Res. **12**:1357–1369.

Hey, J. 1999. The neutralist, the fly and the selectionist. Trends Ecol Evol. **14**:35–38.

Holland, B., and W. R. Rice. 1999. Experimental removal of sexual selection reverses intersexual antagonisitc coevolution and removes reproductive load. Proc. Natl. Acad. Sci. USA **96**:5083–5088.

Hudson, R. R., M. Kreitman, and M. Aguadé. 1987. A test of neutral molecular evolution based on nucleotide data. Genetics **116**:153–159.

Jin, W., R. M. Riley, R. D. Wolfinger, K. P. White, G. Passador-Gurgel, and G. Gibson. 2001. The contributions of sex, genotype and age to transcriptional variance in *Drosophila melanogaster*. Nat. Genet. **29**:389–395.

Kalmykova, A. I., A. A. Dobritsa, and V. A. Gvozdev. 1997. The *Su(Ste)* repeat in the *Y* chromosome and *beta-CK2tes* gene encode predicted isoforms of the regulatory beta-subunity of protein kinase *CK2* in *Drosophila melanogaster*. FEBS Lett. **416**:164–166.

Long, A. D., S. L. Mullaney, T. F. Mackay, and C. H. Langley. 1996. Genetic interactions between naturally occurring alleles at quantitative trait loci and mutant alleles at candidate loci affecting bristle number in *Drosophila melanogaster*. Genetics **144**:1497–1510.

Lyman, R. F., and T. F. Mackay. 1998. Candidate quantitative trait loci and naturally occurring phenotypic variation for bristle number in *Drosophila melanogaster*: the *Delta-Hairless* gene region. Genetics **149**:983–998.

Maine, E. M., J. L. Lissemore, and W. T. Starmer. 1995. A phylogenetic analysis of vertebrate and invertebrate *Notch-related* genes. Mol. Phylogenet. Evol. **4**:139–149.

McDonald, J. H., and M. Kreitman. 1991. Adaptive protein evolution at the *Adh* locus in *Drosophila*. Nature **351**:652–654.

Michalak, P., and M. A. F. Noor. 2003. Genome-wide patterns of expression in *Drosophila* pure species and hybrid males. Mol. Biol. Evol. **20**:1070–1076.

Nielsen, M. G., and E. C. Raff. 2002. The best of all worlds or the best possible world? Developmental constraints in the evolution of beta tubulin nad the sperm tail axoneme. Evol. Devel. **4**:303–315.

Okuyama, E., H. Tachida, and T. Yamazaki. 1997. Molecular analysis of the intergenic region of the duplicated *Amy* gene is of *Drosophila melanogaster* and *Drosophila teisseri*. J. Mol. Evol. **45**:32–42.

Oleksiak, M. F., G. A. Churchill, and D. L. Crawford. 2002. Variation in gene expression within and among natural populations. Nat. Genet. **32**:261–266.

Orr, H.A., and M. Turrelli. 2001. The evolution of postzygotic isolation: accumulating Dobzhansky-Muller incompatibilities. Evolution **55**:1085–1094.

Posthuma, L., and N. M. Vanstraalen. 1993. Heavy-metal adaptation in terrestrial invertebrates: a review of occurrence, genetics, physiology and ecological consequences. Comp. Biochem. Physiol. C Pharmacol. Toxicol. Endocrinol. **106**:11–38.

Rand, D. M., and L. M. Kann. 1996. Excess amino acid polymorphism in mitochondrial DNA: contrasts among genes from *Drosophila*, mice, and humans. Mol. Biol. Evol. **13**:735–748.

Ranz, J. M., C. I. Castillo-Davis, C. D. Meiklejohn, and D. L. Hartl. 2003. Sex-dependent gene expression and evolution of the *Drosophila* transcriptome. Science **300**:1742–1745.

Smith, N. G. C., and A. Eyre-Walker. 2002. Adaptive protein evolution in *Drosophila*. Nature **415**:1022–1024.

Sokal, R. R. 1976. The Kluge-Kerfoot phenomenon reexamined. Am. Nat. **110**:1077–1091.

Swanson, W. J., and V. D. Vacquier. 2002. Reproductive protein evolution. Annu. Rev. Ecol. System. **33**:161–179.

Symonds, J. E., and J. B. Gibson. 1992. Restriction site variation, gene duplication, and the activity of *sn*-glycerol-e-phosphate dehydrogenase in *Drosophila melanogaster*. Biochem. Gen. **30**:169–188.

True, J. R., and E. S. Haag. 2001. Developmental system drift and flexibility in evolutionary trajectories. Evol. Dev. **3**:109–119.

Vogt, R. G. 2002. Odorant binding protein homologues of the malaria mosquito *Anophales gambiae*: possible orthologues of the *OS-E* and *OS-F OBP*s of *Drosophila melanogaster*. J. Chem. Ecol. **28**:2371–2376.

Wagner, G. P., and K. Schwenk. 2000. Evolutionarily stable configurations: functional integration and the evolution of phenotypic stability. Evolutionary Biology **31**:155–217.

Wang, W., F. G. Brunet, E. Nevo, and M. Long. 2002. Origin of *sphinx*, a young chimeric RNA gene in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **99**:4448–4453.

Wayne, M. L., and L. M. McIntyre. 2002. Combining mapping and arraying: an approach to candidate gene identification. Proc. Natl. Acad. Sci. USA **99**:14903–14906.

Yang, J., Z. L. Gu, and W. H. Li. 2003. Rate of protein evolution versus fitness effects of deletions. Mol. Biol. Evol. **20**:772–774.