

## RATES OF DIVERGENCE IN GENE EXPRESSION PROFILES OF PRIMATES, MICE, AND FLIES: STABILIZING SELECTION AND VARIABILITY AMONG FUNCTIONAL CATEGORIES

BERNARDO LEMOS,<sup>1,2</sup> COLIN D. MEIKLEJOHN,<sup>1,3</sup> MARIO CÁCERES,<sup>4,5,6</sup> AND DANIEL L. HARTL<sup>1,7</sup>

<sup>1</sup>Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138

<sup>2</sup>E-mail: blemos@oeb.harvard.edu

<sup>3</sup>E-mail: meiklej@fas.harvard.edu

<sup>4</sup>Department of Human Genetics, Emory University School of Medicine, 615 Michael Street, Atlanta, Georgia 30322

<sup>5</sup>Division of Neuroscience, Yerkes National Primate Research Center, Emory University, 954 Gatewood Road, Atlanta, Georgia 30329

<sup>6</sup>E-mail: mcaceres@genetics.emory.edu

<sup>7</sup>E-mail: dhartl@oeb.harvard.edu

**Abstract.**—The extent to which natural selection shapes phenotypic variation has long been a matter of debate among those studying organic evolution. We studied the patterns of gene expression polymorphism and divergence in several datasets that ranged from comparisons between two very closely related laboratory strains of mice to comparisons across a considerably longer time scale, such as between humans and chimpanzees, two species of mice, and two species of *Drosophila*. The results were analyzed and interpreted in view of neutral models of phenotypic evolution. Our analyses used a number of metrics to show that most mRNA levels are evolutionary stable, changing little across the range of taxonomic distances compared. This implies that, overall, widespread stabilizing selection on transcription levels has prevented greater evolutionary changes in mRNA levels. Nevertheless, the range of rates of divergence is large with highly significant differences in the rate and patterns of transcription divergence across functional classes defined on the basis of the gene ontology annotation (primates and mice datasets) or on the basis of the pattern of sex-biased gene expression (*Drosophila*). Moreover, rates of divergence of sex-biased genes in the contrast between *Drosophila* species show a distinct pattern from that observed in the contrast between populations of *D. melanogaster*. Hence, we discuss the time scale of the changes observed and its consequences for the relationship between variation in gene expression within and between species. Finally, we argue that differences in mRNA levels of the magnitudes observed herein could be explained by a remarkably small number of generations of directional selection.

**Key words.**—*Drosophila*, genetic drift, natural selection, neutral models, phenotypic evolution.

Received April 20, 2004. Accepted October 14, 2004.

Although many organismal differences clearly reflect adaptations to distinct environments and undoubtedly have been shaped by natural selection, there remain a large number of differences between organisms for which genetic drift cannot be unequivocally ruled out as causing the observed phenotypic diversity. In addition, many phenotypes have remained almost unchanged since the split of two species or populations; in such cases one might suggest that stabilizing selection is responsible for the observed lack of phenotypic diversity. However, statements regarding the degree of similarity among phenotypes should be made with reference to the degree of differentiation that one would expect if these phenotypes were neutrally diverging owing to the independent introduction of mutations in each species or population and their random fixation or loss by genetic drift.

In this context, neutral models of phenotypic evolution (Lande 1976, 1977; Turelli et al. 1988; Lynch 1990) were constructed as extensions of the neutral theory of molecular evolution (Kimura 1968). These phenotypic models were developed to test evolutionary hypotheses that invoke directional or stabilizing selection as forces promoting or retarding phenotypic divergence. Specifically, rates of divergence below a theoretically determined threshold indicate that stabilizing selection has prevented greater phenotypic divergence. In these cases, however, one may not distinguish between natural selection and genetic drift as the mechanism responsible for the differences in phenotypes that have ac-

cumulated within the limits set by stabilizing selection. Other approaches have been used to evaluate the relative contribution of different evolutionary forces in shaping patterns of phenotypic variation. One method compares among-population differentiation in presumably neutral markers, such as microsatellite length polymorphism, with that of the trait of interest (Bonin et al. 1996; Merila and Crnokrak 2001; Storz 2002). The distribution of neutral molecular variation within and between populations can be evaluated by means of the  $F_{ST}$  statistic (Wright 1951) and compared with the pattern observed in the phenotypic trait of interest as estimated by the  $Q_{ST}$  statistic (Spitze 1993). Another approach is to compare patterns of evolutionary fixation of phenotypic characters with patterns of fixation of genetic markers, using statistical tests framed within coalescent theory (e.g., Masta and Maddison 2002). The rationale for this approach is that directional selection reduces the amount of time needed for the fixation of selected traits, yet it does not reduce the effective population size (and thereby the fixation time) of neutral, unlinked genes.

Recent studies have reported extensive variation in gene expression levels (mRNA levels) across individuals, populations, and species (Jin et al. 2001; Sandberg et al. 2000; Enard et al. 2002; Olesiak et al. 2002; Cáceres et al. 2003; Meiklejohn et al. 2003; Ranz et al. 2003; Rifkin et al. 2003; Nuzhdin et al. 2004). Questions regarding the biological and evolutionary significance of variation in gene expression fol-

TABLE 1. Datasets analyzed in this study.

Contrast	Estimate of minimum divergence time	Minimum number of generations/year	Estimate of minimum total number of generations $\times 10^3$	Dataset
Laboratory strains of <i>Mus musculus</i> (129SvEv $\times$ C57BL/6) 4 brain tissues	60 years	5	0.6	Sandberg et al. (2000)
<i>Drosophila melanogaster</i> African $\times$ North American strains	300 years	8	4.8	Meiklejohn et al. (2003)
Human $\times$ chimpanzee (primate liver)	6 million years	0.05	600	Enard et al. (2002)
Human $\times$ chimpanzee (primate brain)	6 million years	0.05	600	Enard et al. (2002)
<i>M. musculus</i> $\times$ <i>M. spretus</i> (brain)	2 million years	2	8000	Enard et al. (2002)
<i>M. musculus</i> $\times$ <i>M. spretus</i> (liver)	2 million years	2	8000	Enard et al. (2002)
<i>D. melanogaster</i> $\times$ <i>D. simulans</i>	2 million years	8	32,000	Ranz et al. (2003)

low immediately from these studies. In this context, neutral models of phenotypic evolution can provide a yardstick against which one can interpret patterns of intrapopulation variation and evolutionary divergence in gene expression. For example, Clegg et al. (2002) used such models to show that morphological change among isolated populations of a passerine bird took place at an unexpectedly rapid pace and was, therefore, most likely to have been driven by directional selection. Using a slightly different implementation of the tests mentioned above, Rifkin et al. (2003) argued that directional selection could explain some of the gene expression differences they observed in the genus *Drosophila*. Using the same approach, Hsieh et al. (2003) estimated the fraction of differences in gene expression between humans and chimpanzees that could be attributed to the action of directional selection.

We have analyzed levels of expression variation in several published datasets that ranged from comparisons between very closely related samples, such as two recently diverged strains of mice, to comparisons across considerably longer time scales, such as between humans and chimpanzees, and between species of *Drosophila*. By comparing observed rates of divergence with expected divergence occurring exclusively by mutation and genetic drift, the relative roles of directional and stabilizing selection, genetic drift, and mutation in shaping current patterns of gene expression variation in natural populations can be evaluated, and insights into the evolutionary dynamics of gene expression variation within and among populations might be gained. A synthesis of these and other data will not only provide insights into the evolution of gene expression but, more generally, may also allow for a better understanding of the distribution of rates of divergence of other phenotypic characters.

## MATERIALS AND METHODS

### *Gene Expression Levels*

We studied variation in gene expression (mRNA) levels for several published datasets (Table 1), which included mRNA levels estimated with either cDNA or oligonucleotide arrays. These platforms use different methods to quantify mRNA abundance of thousand of genes in a biological sample. For estimates based on cDNA arrays (*Drosophila* samples; see Table 1) relative expression levels were calculated as in Ranz et al. (2003) and Meiklejohn et al. (2003) using

a Bayesian procedure (BAGEL; Townsend and Hartl 2002). BAGEL analyses result in estimates of the change in expression across samples that are normalized so that the lowest observed value is arbitrarily assigned a reference value of one. The relative expression estimates were then used in all subsequent analyses of the rate of divergence in gene expression levels in the *Drosophila* dataset.

Oligonucleotide arrays (GeneChip, Affymetrix, Santa Clara, CA) datasets were analyzed using the microarray suite (MAS) software 4.0 (Affymetrix). All arrays were first normalized to the same average intensity value of 200, based on probe sets corresponding to the 60th to 90th percentiles of the hybridization signals. Average difference (AD) values were then used as estimates of the expression levels and to calculate rates of divergence. AD equals the average difference between each of the 16 or 20 pairs of perfect match (PM) and mismatch (MM) probes for each probe set. Only genes with AD (hereafter ‘‘expression level’’) equal to or higher than 20.0 in all samples of a given comparison were included in the analysis. This procedure assures that each gene has the same statistical degrees of freedom and facilitates interpretation of the results in terms of the overall frequency distribution of evolutionary rates.

### *Rates of Phenotypic Divergence*

Several measures of the rate of phenotypic divergence of continuous traits have been used for inferring evolutionary processes (Haldane 1949; Lerman 1965; Lande 1977; Cherry et al. 1982; Gingerich 1983; Charlesworth 1984; Turelli et al. 1988; Lynch 1990; Martins 1994; Lande 2000). Analysis of rates of phenotypic divergence requires estimates of; (1) the magnitude of the change; (2) the magnitude of the variability within each population or species compared; and (3) the time of divergence between the samples compared. Haldane (1949) proposed the quantification of the rate of evolution in which one unit of change corresponds to a factor of  $e$  ( $= 2.718$ ) per million years, which he named the ‘‘darwin.’’ This metric does not take into account the variance present within the samples compared (usually species or population), which led Gingerich (1983) to introduce the haldane, a metric that is scaled by the intraspecific variability and defined as

$$H = \frac{\bar{x}_1 - \bar{x}_2}{\sigma_w t}, \quad (1)$$

where  $\bar{x}_1$  and  $\bar{x}_2$  are the mean value of species 1 and 2, respectively;  $\sigma_w$  is the standard deviation of the traits within species; and  $t$  is the time since divergence of the two species compared. The darwin and the haldane are the most commonly reported measures of evolutionary rate in phenotypic traits (Hendry and Kinnison 1999; Kinnison and Hendry 2001); both metrics, however, have the drawback that they are not clearly related to quantitative genetic parameters. In addition, several other distance metrics such as the Mahalanobis or Manhattan distances exist and could be readily employed for measuring rates of phenotypic divergence (Atchley et al. 1982, Cherry et al. 1982). As with the darwin and the haldane, it is unclear how they relate to quantitative genetic parameters.

Alternatively, metrics of evolutionary rate suggested by Lande (1976, 1977) and Turelli et al. (1988) have quantitative genetic underpinnings. The mutation-drift equilibrium (MDE) model suggested by Turelli et al. (1988) tests the neutral hypothesis of evolution by mutation and drift by the  $F$  statistic:

$$F = \frac{V_b}{V_m t}, \quad (2)$$

$t$  is the total number of generations separating the lineages,  $V_m$  is the rate of input of new genetic variance due to mutation per generation, and  $V_b$  is the between-species variance determined from an ANOVA in the following way:

$$V_b = \frac{MS_b - MS_w}{n_0}, \quad (3)$$

here  $MS_b$  and  $MS_w$  are the between-population and within-population mean squares, respectively, and  $n_0$  is the average sample size (Lynch 1988; Bjorklund 1991; Savalli 1993). Lower and upper 95% confidence limits of  $V_b$  resulting from mutation and drift can be specified as:

$$\frac{V_m t F_{0.025[n-1, \text{inf}]}}{V_b} < \text{genetic drift} < \frac{V_m t F_{0.975[n-1, \text{inf}]}}{V_b}. \quad (4)$$

Values of between-species variance above or below the 95% confidence interval may be used to argue for directional selection and stabilizing selection, respectively.

Finally, Lynch (1990) suggested another ANOVA-based metric of the rate of phenotypic divergence that is very similar to the previous one and is defined as  $\Delta = V_b/V_w t$ , where  $t$  is the total number of generations separating the two lineages and  $V_b$  and  $V_w$  are the between-species and within-species components of phenotypic variance as determined from an ANOVA (Lynch 1990). The metrics based on ANOVA approaches (Turelli et al. 1988; Lynch 1990) are equivalent and were proposed within a quantitative genetic framework to allow observed divergences to be compared with expectations based on evolution by mutation and drift. Specifically, Lynch (1990) argued that  $\Delta$  can be directly compared to estimates of mutational heritability ( $V_m/V_e$ ), which is the rate of input of new genetic variance per generation ( $V_m$ ) scaled by the environmental variance ( $V_e$ ).

Clark et al. (1995) estimated that the mutational heritability for enzyme activity is typically  $10^{-3}$ . Mukai et al. (1984) investigated the between-line variance in alcohol dehydro-

genase (ADH) activity after 300 generations of mutation accumulation and showed that the increase in variance was likely due to regulatory mutations. Lynch's (1988) reanalysis of Mukai et al. (1984) estimated  $V_m/V_e$  for ADH activity to be  $10^{-4}$ . These estimates are in agreement with the observation that  $10^{-3}$  is the typical mutational heritability for a wide range of characters, with most values usually falling between  $10^{-4}$  and  $10^{-2}$  (Lynch 1988). In view of this  $10^{-4}$  and  $10^{-2}$  are currently the best estimates of the minimum and maximum rate of divergence per generation under mutation and genetic drift, with the caveat that direct estimates of the neutral mutation rate for gene expression are lacking. Therefore, rates of divergence in expression levels below  $10^{-4}$  may suggest that stabilizing selection has prevented greater divergence, whereas rates above  $10^{-2}$  may suggest that directional selection has accelerated the divergence. We have used minimum divergence times as well as minimum generation times (Berry and Bronson 1992; Smith and Tompkins 1995; Powell 1997) when estimating the total number of generations separating the samples compared (see Table 1), which is conservative for our conclusions. Furthermore, we have calculated rates of divergence based on half the total number of generations (i.e., unidirectional divergence) separating taxa, a procedure that is also conservative for our conclusions.

#### Variation among Functional Categories

We searched for differences in intraspecific (or intrapopulation) variation and evolutionary divergence across functional categories either defined by the Gene Ontology (GO) classification (Ashburner et al. 2000) or by the pattern of expression in the sexes (Ranz et al. 2003). In the primate dataset, we tested for differences in intraspecific and interspecific variance across seven GO categories: (1) metabolic enzymes; (2) transcription factors; (3) signal transduction enzymes; (4) signal transduction proteins (not included in the class "enzyme"); (5) transporters; (6) cytoskeleton proteins; and (7) adhesion proteins. Kruskal-Wallis and median tests (Sokal and Rohlf 1995) were used to examine differences among the above categories. In the primate dataset, metrics of intraspecific (or intrapopulation) variation and evolutionary divergence were scaled by the mean square error, thus making them independent of the measurement error and therefore mean expression level of the class. Finally, in the *Drosophila* datasets and following Ranz et al. (2003) and Meiklejohn et al. (2003), we investigated the differences in the rate of divergence of genes classified by their pattern of sex-biased gene expression in 2- to 4-day-old virgin adults.

## RESULTS

### Rates of Gene Expression Divergence

We systematically compared rates of gene expression divergence measured in haldanes with those based on the ratio of variance components estimated by ANOVA. We found that these two metrics are highly correlated in all datasets ( $r^2 \geq 0.94$ ; Fig. 1). In view of this strong correlation, all subsequent discussions are based on rates of divergence estimated using variance components (ANOVA).

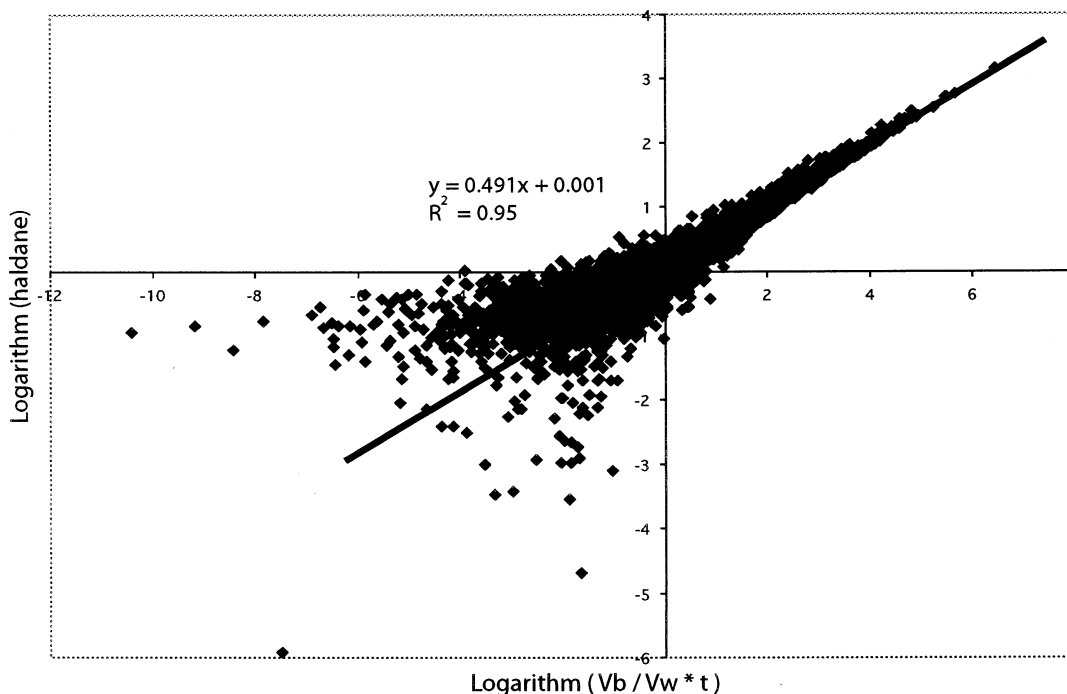


FIG. 1. Relationship between estimates of the rate of gene expression divergence.

Table 2 shows the number of genes classified into each evolutionary mode (stabilizing selection, mutation accumulation and/or random drift, and directional selection) in each dataset according to rate of divergence. One striking feature of this table is that the majority of genes (61–100%) in all datasets have diverged at a much slower rate than would be expected under unconstrained divergence in expression levels. This implies that stabilizing selection is a strong force preventing greater gene expression divergence, although we must remark that this analysis cannot distinguish whether the differences in expression levels that are observed between samples resulted from natural selection or genetic drift. Furthermore, the comparison between the two strains of mice stands out with an unusually large fraction of genes whose rate of divergence cannot be distinguished from that expected under neutrality or is larger than the neutral expectation. This

observation can be accounted for by low polymorphism in inbred strains of mice, by the small number of generations separating groups in this contrast, and/or reduced stabilizing selection in the laboratory (see Discussion).

Because minimum estimates of divergence time were used (Table 1) as well as a unidirectional mode of divergence (i.e., taking  $t$  as half of the minimum number of generations separating the taxa), our estimates of the rates of divergence are biased upward. Therefore, the conclusion about the preponderance of stabilizing selection affecting the gene expression is very robust to errors in our estimates of divergence time. Moreover, under such a conservative scenario of minimum divergence times and unidirectional divergence, we calculated the neutral mutation rate that would be compatible with assigning at least the top 30% most rapidly diverging genes to the neutral evolution (i.e., mutation accumulation through

TABLE 2. Number of features classified by rate of divergence. To be conservative and underestimate the number of features under stabilizing selection, we used lower estimates of the number generations separating the samples (see Table 1) and the unidirectional divergence (i.e., half the minimum number of generations).

Contrast	Stabilizing selection (rate < 0.0001)	Genetic drift/ mutation accumulation (0.0001 < rate < 0.01)	Directional selection (rate > 0.01)	Number of generations required for the top decile of rates to exceed 0.01
Mouse strains: hippocampus	6030 (67%)	2189 (24%)	823 (9%)	—
Mouse strains: cortex	5335 (61%)	2366 (27%)	1106 (12%)	—
Mouse strains: midbrain	6026 (66%)	2184 (24%)	885 (10%)	—
Mouse strains: cerebellum	5657 (63%)	2243 (25%)	1002 (12%)	—
<i>Drosophila melanogaster</i> : North American vs. African strains	4437 (93%)	306 (6%)	16 (1%)	230
Primates: liver	8761 (96%)	372 (4%)	0	1250
Primates: brain	9231 (99%)	95 (1%)	0	460
Mouse species: liver	6833 (100%)	0	0	300
Mouse species: brain	7252 (100%)	1 (<1%)	0	300
<i>D. melanogaster</i> vs. <i>D. simulans</i>	4758 (99%)	1 (1%)	0	650

genetic drift) class. We found that neutral mutation rates for gene expression phenotypes as low as  $10^{-8}$  (comparison between *Drosophila* species and populations),  $10^{-7}$  (comparison between species of mice, liver and brain tissues), and  $10^{-6}$  (comparison between humans and chimpanzees, liver and brain tissues) would have to be conjectured. This indicates that our conclusion of stabilizing selection is robust to errors in the estimates of the neutral mutation rate for gene expression levels and would still hold true even if the neutral mutation rate for gene expression level had been overestimated by more than 100-fold.

To infer what time scales are required to evolve the divergences in gene expression of the magnitude observed in these data, we assumed that the top 10% most rapidly evolving genes in each comparison diverged by a single pulse of directional selection producing rates of divergence above the neutral expectation. This was done by reducing the estimated divergence times until 10% of the genes had rates of divergence greater than  $10^{-2}$  per generation, which is the minimum rate of divergence associated with directional selection (see Materials and Methods). This amounts to asking how many generations of directional selection at its minimum detectable rate would be needed to produce phenotypic differences of the magnitude observed in these data. We find that, in all datasets, differences in gene expression could be accounted for directional selection maintained over a remarkably small number of generations (Table 2).

Figure 2 shows the frequency distribution of scaled divergences ( $\Delta t = V_b/V_w$ ) in the primate, mouse and fly datasets. In all datasets the distribution of rates of divergence was characterized by a large number of genes showing little or no divergence, with a decreasing number of genes in faster evolving classes. Although this is a pattern often seen in genomic studies (Luscombe et al. 2002), it is unclear whether in this instance it is mainly the by-product of the statistical distribution associated with the metric chosen or whether it reflects a more fundamental biological phenomenon. We also observed that within- and between-species components of the variance are generally correlated (Spearman rank correlation and Pearson correlation coefficients  $\geq 0.29$  in all data sets;  $P < 10^{-10}$ ); this correlation may be explained by the neutral theory expectation that variation between species may reflect variation within species or may simply result from the scaling of the variance with the mean. Nonetheless, several genes show an unusually high variance within species that does not translate into an unexpectedly high amount of variance between samples, whereas others show large amounts of between-sample variance with small variances within samples, as has been observed elsewhere (Nuzhdin et al. 2004).

The distribution of rates of divergences may be described by the relationship between the logarithm of the frequency and the logarithm of the rank, which can be fit by a linear regression. Figure 3 illustrates such an analysis for the primate dataset. This approach revealed that the distribution of rates of divergence of genes expressed in the brain has a significantly ( $P < 0.01$ ) steeper slope than that of genes expressed in the liver, thus implying an enrichment of fast-evolving genes in the liver as compared to the brain (Fig. 3A). The same approach can be used to examining the distribution of the between- and within- species components ( $V_b$ ,

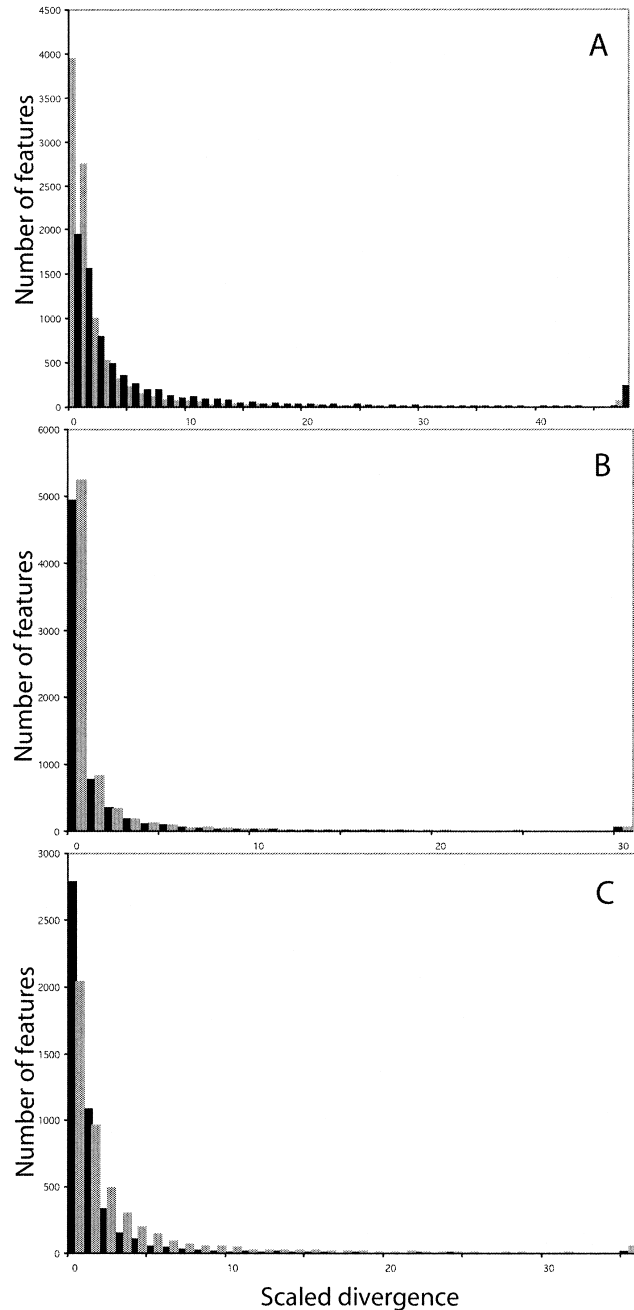


FIG. 2. Frequency distribution of scaled divergence. (A) Primates: gray, brain; black, liver; (B) mice: gray, brain; black, liver; (C) *Drosophila*: gray, *D. melanogaster* versus *D. simulans*; black, African versus North American strains.

$V_w$ ) of the variance (Fig. 3B, C). This shows that the skew of the distribution of divergence rates of genes in the liver toward higher values, relative to genes in the brain, cannot be unequivocally partitioned into an overall larger  $V_b$  in the liver or an overall larger  $V_w$  in the brain. Rather, gene specific combinations of  $V_b$  and  $V_w$  seem to be causing the apparent acceleration of genes in the liver sample, with some tendency toward both an enrichment of genes with high  $V_b$  in the liver and genes with high  $V_w$  in the brain.

Figure 4 shows the logarithm of scaled divergences in pri-

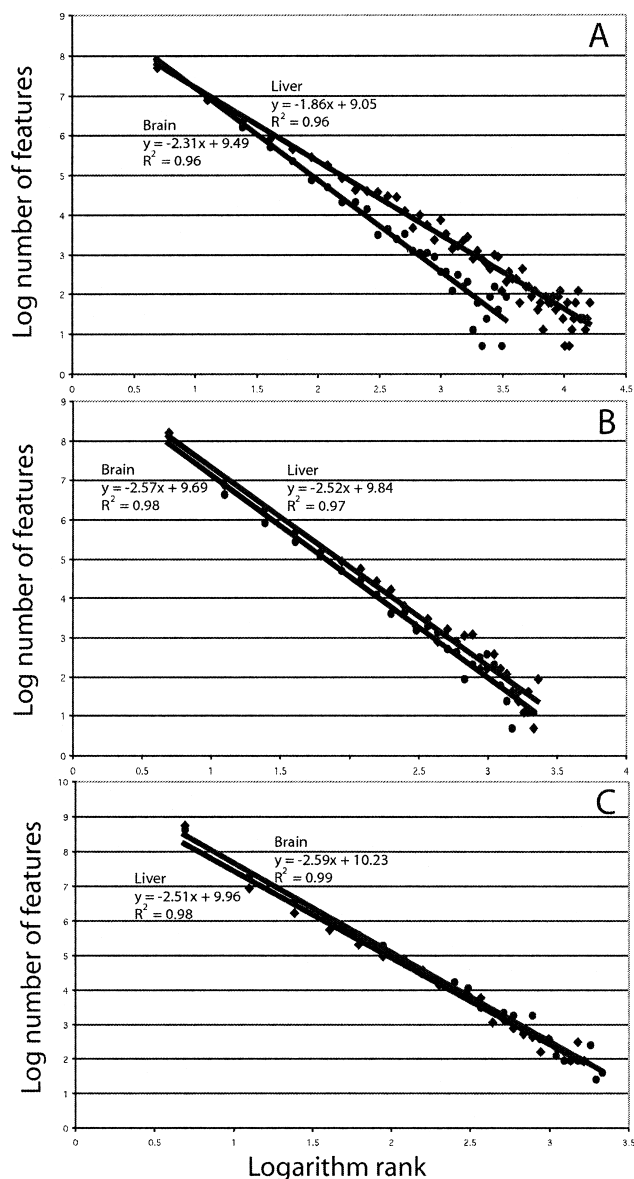


FIG. 3. Log (number of features) versus log(rank) plot for the distributions of rates of divergence in the primate dataset of Enard et al. (2002): (A) scaled divergence ( $V_b / V_w$ ); (B) between-species component of the variance ( $V_b$ ); (C) within-species component of the variance ( $V_w$ ).

mate, mice, and fly datasets. These analyses show that the distribution of the logarithm of the rate of divergence approaches normality in all datasets ( $P > 0.1$  in all samples, Kolmogorov-Smirnov test) and that these distributions are typically skewed toward an overabundance of genes with low rates of divergence (i.e., genes whose large variance within samples is not matched by an equivalent increase in the variance between samples). It is also worth pointing out that the range of rates of divergence is large, typically with greater than a 10,000-fold difference between the slowest and fastest diverging genes. Table 3 shows the summary statistics for logarithm of the rate of divergence in each dataset. Interestingly, contrasts involving longer time scales generally resulted in frequency distributions increasingly skewed to the

left (increasingly more negative coefficients of skewness). Comparisons between the closely related laboratory strains of mice, however, resulted in either a positive coefficient of skewness (three tissues) or a distribution with no skew (one tissue; hippocampus). In the primate dataset, comparison between the log(rate of divergence) of genes in the liver and brain tissues strengthens the conclusion that expression levels in the liver show indeed faster rates of divergence than in the brain ( $P < 0.001$ ; ANOVA).

#### Variation among Functional Categories

Several consistency checks of the data were made throughout the analyses. Most importantly, in one dataset (human-chimpanzee contrast of brain and liver tissue; Enard et al. 2002) we examined the correlations between metrics of variation and divergence and gene expression level (mRNA abundance). This dataset was chosen because it included two different tissues analyzed with within-species replication on oligonucleotide arrays. cDNA arrays such as those used here for *Drosophila* may not be appropriate for estimating absolute mRNA abundance and appear to have no correlation between error variance and abundance (Townsend 2003). In the human-chimpanzee comparisons, absolute transcript abundance was highly significantly correlated with the error mean square (Spearman rank correlation of  $-0.72$  and  $-0.80$  in liver and brain, respectively;  $P < 10^{-6}$  in both cases), thus indicating that the level of gene expression strongly influences measurement error when estimating mRNA abundance based on oligonucleotide arrays. Accordingly, lowly expressed transcripts are subject to significantly larger experimental errors than highly expressed ones. This effect was observed across the entire dynamic range of expression levels, which, on a log scale, typically varied from 3.0 to 8.0 in all datasets. Indeed, similar correlations between expression level and mean square error were observed even when only the top 20%, 10%, and 1% most highly expressed genes were analyzed. Genes with lower mRNA abundance in a given sample show larger error variances that could be easily misinterpreted as an inflation of their within-species (or population or strain) variance. Inflated estimates of intrasample variability result, in turn, in a metric of the rate of evolutionary divergence (or scaled divergences) that is biased downward. This may be an issue in comparisons between mice species, for which no replications were performed and, therefore, the mean square error could not be estimated.

This establishes expression level as a relevant correlate that must be taken into account whenever comparing patterns of evolutionary divergence in gene expression level across classes whose mean (or median) expression levels differ, at least on oligonucleotide arrays. Highly significant differences in absolute transcript abundance across GO classes were detected in both primate and mice datasets. Notably, mRNAs of genes coding for metabolic enzymes and cytoskeleton proteins are significantly more highly expressed than those of genes coding for transcription factors, signal transduction and cell adhesion ( $P < 0.01$ , Kruskal-Wallis, Median test). After correcting the rates of divergence by the mean square error we found that metabolic enzymes stand out as the most rapidly diverging evolutionary class and that transcription fac-

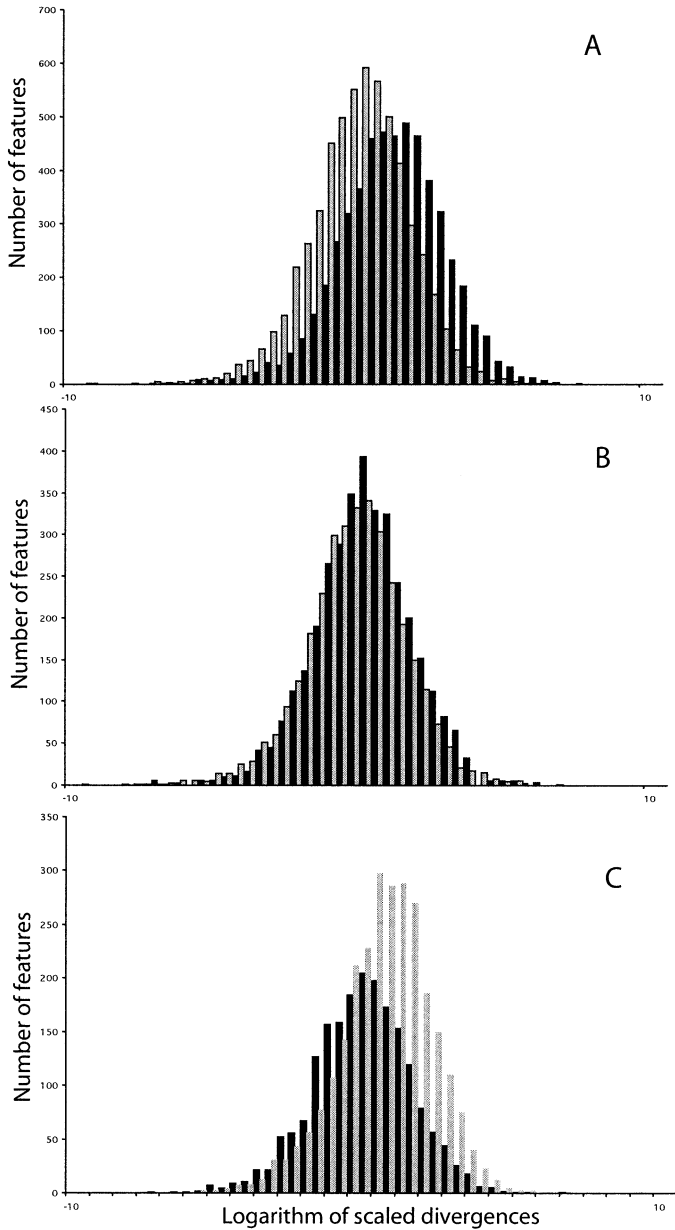


FIG. 4. Logarithm of scaled divergences ( $V_b / V_w$ ). (A) Primates: distributions for genes expressed in the brain (gray) and liver (black); (B) mouse: distributions for genes expressed in the brain (gray) and liver (black); (C) *Drosophila*: *D. melanogaster* versus *D. simulans* (gray) and African vs. North American (black) contrasts.

tors and signal transduction enzymes are more evolutionary stable both in the brain and liver tissues ( $P < 0.01$ , Kruskal-Wallis and Median test). This trend was observed across all datasets for which the mean square error could be estimated. Proteins involved in cell adhesion are an interesting discrepancy between the pattern observed in the brain and liver of primates because cell adhesion proteins show an apparent acceleration in the brain relative to the other classes ( $P < 0.05$ , Median test), whereas such relative acceleration is not observed in cell adhesion proteins expressed in the liver.

We analyzed the variability among mean rate of gene expression divergence in genes with sex-biased expression in

*Drosophila*. Our analyses agree with the conclusion of Meiklejohn et al. (2003) that male-biased genes are both more divergent between species and more variable within species than female-biased and unbiased genes. In addition, we extend their observation to the comparison between African and North American strains of *D. melanogaster*. In both contrasts (*D. melanogaster* vs. *D. simulans* and African vs. North American strains) male-biased genes showed larger within-group components of variance than female-biased and unbiased-genes ( $P < 10^{-5}$ , Kruskal-Wallis and Median tests), while female-biased genes were the least variable within groups in both contrasts. In addition, in both contrasts (*D. melanogaster* vs. *D. simulans* and African vs. North American strains) male-biased genes showed the largest variation between groups, whereas female-biased genes were the least variable.

Interestingly, however, our analysis of the distribution of rates of divergence of genes classified by their pattern of expression in the sexes shows that, in the comparison between species, the rate of divergence of male biased genes is significantly slower than that of female-biased and unbiased genes ( $P < 0.001$ , ANOVA on log rate of divergence;  $P < 0.05$ , Kruskal-Wallis and Median tests; Table 4). This is in sharp contrast with the comparison between African and North American strains in which male biased genes show a significantly faster rate of evolution than female biased genes ( $P < 0.05$ , ANOVA on log rate of divergence;  $P < 0.05$ , Kruskal-Wallis and Median tests; Table 4). These results illustrate how differences in the time scale of the comparison (i.e., between species or populations) might shift our expectations of how within-group variation is matched by between-group divergence (see Discussion).

Finally, genes classified accordingly to their pattern of expression in the sexes also show relevant differences in the relationship between the skewness of the distribution of  $\log(\Delta)$  versus time scale of the contrast (Table 4). Specifically, the distribution of  $\log(\Delta)$  of genes showing male- and female-biased expression shows increased negative (left) skew in the comparison across larger time scales. This fits with the trend observed in Table 3 regarding the relationship between the skew of the distribution of  $\log(\text{rate of divergence})$  and the time scale of the contrast. However, genes with unbiased gene expression in the sexes do not show increased negative skewness in the contrast between the two *Drosophila* species relative to that between African and North American strains.

## DISCUSSION

### *Pervasive Stabilizing Selection Prevents Greater Gene Expression Divergence*

Interpreted in view of neutral models of phenotypic evolution, our results regarding the rate of evolutionary divergence in gene expression across populations and species indicate that stabilizing selection has been a major force preventing substantially greater differentiation in expression patterns than currently observed. Most genes had rates of divergence well below the neutral interval of evolution exclusively by mutation and genetic drift, which together with the underestimation of divergence times and the assumption

TABLE 3. Summary statistics for the distribution of the logarithm of the rate of divergence ( $\Delta$ ) in each contrast. Mean, standard deviation, skewness and kurtosis are shown.

Contrast	Mean log (rate of divergence)	Standard deviation log(rate of divergence)	Skewness	Kurtosis
Mouse strains: hippocampus	-6.43	1.85	0.00	0.85
Mouse strains: cortex	-6.14	1.83	0.18	0.76
Mouse strains: midbrain	-6.30	1.86	0.20	1.22
Mouse strains: cerebellum	-6.16	1.83	0.19	0.88
<i>Drosophila melanogaster</i> : North American vs. African strains	-9.57	1.67	-0.26	0.68
Primates: liver	-12.66	1.85	-0.27	0.64
Primates: brain	-13.31	1.71	-0.36	1.09
Mouse species: liver	-16.31	1.75	-0.08	1.91
Mouse species: brain	-16.30	1.75	-0.02	1.98
<i>D. melanogaster</i> vs. <i>D. simulans</i>	-16.42	1.63	-0.4	1.24

of unidirectional divergence indicate that our conclusion of widespread stabilizing selection is robust. Recently, Rifkin et al. (2003) used a slightly modified implementation of neutral models of phenotypic evolution to estimate the fraction of genes whose divergence between species of *Drosophila* could be inferred to be due to directional selection. Following Rifkin et al. (2003), Hsieh et al. (2003) implemented similar modifications to estimate the fraction of genes that diverged due to directional selection between humans and chimpanzees. Our analyses differ from the previous ones in that these authors corrected their observed ratios of divergence to polymorphism (equivalent in our implementation to  $V_b/V_w$ ) by the expectations  $(2V_m t)/(2N_e V_m)$  (cf. Lynch and Hill 1986; Falconer and Mackay 1996, pp. 271), thus canceling out the mutational variance. The corrected ratios, however, are highly sensitive to the specific combination of population size and time since divergence, and the estimated proportion of genes will vary accordingly. Therefore, the large errors associated with the estimation of parameters of the models make them unsuited to providing precise estimates of the fraction of genes that diverged by directional selection. Although quantitative in nature, models of neutral phenotypic evolution are generally regarded as suitable for providing only qualitative information concerning directional selection on phenotypic traits (Turelli et al. 1988). Accordingly, these analyses presented here, as well as the results of Hsieh et al.

(2003) and Rifkin et al. (2003), point most forcefully to the preponderance of stabilizing selection hindering greater divergence in transcription profiles. Finally, we note that while our analysis does not distinguish between natural selection and genetic drift as the mechanism of choice for explaining the differences in phenotypes that have accumulated within the limits set by stabilizing selection, recent implementations of neutral models of phenotypic evolution may overcome these difficulties (Marroig and Cheverud 2004).

In the context of size differences between the smallest living mammal (least shrew at 3 g) and the largest (blue whale at  $10^8$  g), Gingerich (2001) argued that starting from an average sized ancestral mammal and assuming that the size diverged randomly at a rate of 0.1 standard deviations per generation, the whole size range of extant mammals could be acquired in a modest time scale of fewer than  $10^5$  generations, which is only 1% of the time available for all mammalian evolution. Similarly, gene expression levels as estimated with oligonucleotide arrays have a range of 5 expression units (minimum 3, maximum 8) on a log scale. Assuming a maximum within-species standard deviation of 0.1 units (equivalent to a coefficient of variation of less than 3% when gene expression level is 3), the difference between minimum and maximum expression levels is about 50 standard deviations. Therefore, if we conservatively assume that the expression level of a hypothetical gene evolved from its min-

TABLE 4. Summary statistics for the distribution of the logarithm of the rate of divergence ( $\Delta$ ) in *Drosophila* genes classified according to their pattern of expression in the sexes (Meiklejohn et al. 2003; Ranz et al. 2003). Mean, standard deviation, skewness, and kurtosis are shown. MBG, male-biased genes; FBG, female-biased genes; UBG, non-sex-biased genes.

		<i>D. melanogaster</i> vs. <i>D. simulans</i>	<i>D. melanogaster</i> African vs. North American strains
Mean log(rate of divergence)	MBG	-16.0	-8.6
	FBG	-15.6	-8.9
	UBG	-15.7	-8.8
Standard deviation log(rate of divergence)	MBG	1.63	1.65
	FBG	1.66	1.69
	UBG	1.56	1.64
Skewness	MBG	-0.50	-0.30
	FBG	-0.43	-0.17
	UBG	-0.33	-0.34
Kurtosis	MBG	0.88	0.88
	FBG	0.90	0.61
	UBG	2.10	0.70

imum value to its maximum possible value (3–8), at a rate equal to 0.05 standard deviations per generation, we may then ask how many generations it would take for the expression level of this gene to achieve such change. It turns out that even at such an unrealistically small rate (0.05 standard deviations per generation) it would require only about 1000 generations for a transition from minimum to maximum expression level to occur. Indeed, much faster rates of divergence are observed in nature (Endler 1986; Hendry and Kinnison 1999) and phenotypic models have shown that divergences on the order of several standard deviations can often be accomplished in a small number of generations (e.g., Kirkpatrick 1982). Moreover, Ferea et al. (1999) and Toma et al. (2002) showed that large differences in gene expression can be accomplished in just a few generations of artificial selection. Finally, we note that the long time scales available for evolutionary change makes the detection of positive selection based on rates of divergence highly unlikely (Lemos et al. 2001). This is because cellular, biochemical, and pleiotropic constraints presumably define a range of possible phenotypic variation that is too narrow relative to the potential for evolutionary variation due to the combination of the neutral mutation rate and the time available for divergence. Consequently, on the time scales considered here in comparisons among species, genes evolving neutrally within such a range are expected to have reached the maximum and/or minimum expression levels possible, perhaps multiple times.

It is worth remarking upon whether gene expression levels should be interpreted as normally distributed, polygenic, quantitative genetic traits. This issue is relevant because our analysis assumes that gene expression variation is a quantitative character whose variation is determined by a large number of factors. Our approach follows from studies in which enzyme activity (Clark 1994; Clark et al. 1995) and gene expression level (Brem et al. 2002; Schadt et al. 2003) have been successfully treated as a quantitative trait. Nevertheless, the genetic architecture (i.e., number and identity of loci, size of their effect and epistatic interactions, etc.) underlying variation in gene expression still remains an open question that has not yet been resolved empirically.

Our conclusion that divergence in expression levels is constrained primarily by stabilizing selection rather than mutation is in contrast to a recent study by Khaitovich et al. (2004), which suggested that stabilizing selection may be a weak force affecting gene expression variation. This suggestion was based on the observation that gene expression divergence in 23 expressed pseudogenes cannot be distinguished from that of more than 12,000 functional genes (Khaitovich et al. 2004). It is possible, however, that this set of pseudogenes represents a poor standard for neutral evolution of expression levels. This is because these pseudogenes have retained active expression throughout more than 6 million years since the split between humans and chimpanzees, as well as the fact that they were selected based on their high sequence conservation between these primates. This suggests that these sequences are likely to be functional, as was recently shown to be the case for the mouse pseudogene *Makorin1-p1* (Hirotsune et al. 2003; Podlaha and Zhang 2004), which results in a lethal phenotype when deleted. Moreover, the function of *Makorin1-p1* is mediated precisely through the concentra-

tion of its mRNA transcript, while it is not translated into a functional protein product (Hirotsune et al. 2003).

Our analyses cannot distinguish whether the divergence in gene expression levels we observe between species and populations are the result of directional selection or the random fixation of selectively equivalent alleles. Preliminary analysis by Khaitovich et al. (2004) suggested that divergence in expression levels might accumulate linearly with time, a pattern that was interpreted as indicating that gene expression levels evolve according to a neutral model. However, a linear accumulation of evolutionary changes over short time scales is not in disagreement with our inference of strong stabilizing selection on expression levels, as long as evolutionary changes within the bounds defined by stabilizing selection resulted primarily from genetic drift. Furthermore, our evidence for the predominance of stabilizing selection on transcription levels as well as physical limits to variation in expression levels also suggest that such a linear relationship would be unlikely to hold across longer time scales. In view of the confounding influence of sequence divergence on gene expression divergence as measured by oligonucleotide arrays (Hsieh et al. 2003), it remains unclear over which time scales gene expression divergence might accumulate linearly with time.

Indeed, the increasingly negative skew observed in comparisons between more divergent groups underscores the relevance of the time scale in our understanding of the distribution of rates of divergence. Accordingly, the distribution of  $\log(\text{rates of divergence})$  is more skewed to the left (smaller combinations of  $V_b/V_w$ ) in the contrast between the two sibling species of *Drosophila* than in the contrast between North American and African strains of *D. melanogaster*. Finally, analyses of sex-biased genes in contrasts within and between species suggest that not only the time scale of the contrast is relevant but also the class of the gene influences its pattern of variation within and between groups. Rates of divergence of male-biased genes were slower in the comparison between *D. melanogaster* versus *D. simulans* but faster in the comparison between African versus North American strains of *D. melanogaster* than that of female-biased and unbiased-genes. That is, although male-biased genes are indeed more divergent between species, their scaled divergence ( $V_b/V_w$ ) is not. This might be due to stabilizing selection over long evolutionary times, as in the comparison between species, which imposes a limit to the amount of between-species variance one can possibly observe. This results in male-biased genes having the slowest rate of divergence in the comparison between *D. melanogaster* and *D. simulans* due to an unusually large variance within species that is not proportionally reflected in the between-species variance. Taken together, these observations underscore our suggestion that the time scale of the comparison is relevant to our understanding of how within-group variation is converted to between-group divergence.

#### *Rates of Divergence in Primate Brain and Liver*

A considerable amount of morphological divergence in the brain has occurred since the split between the human and chimpanzee lineages (Carroll 2003). This deep differentiation

might be reflected in transcription profiles of human and chimpanzee brains. However, rates of regulatory evolution as interpreted from the perspective of neutral evolution do not seem to be exceptionally high in the brain. In addition, Caceres et al. (2003) and Hsieh et al. (2003) re-analyses of the brain and liver data of Enard et al. (2002) suggest that, on a gene-by-gene basis, the liver shows more statistically significant differences between humans and chimpanzees than does the brain, a finding that is confirmed by our conclusion that expression levels of genes transcribed in the liver have evolved at a faster rate than the expression levels of genes in the brain.

How can one reconcile the observations that gene expression levels are more stable in the brain than in the liver with the commonly held expectation of faster evolutionary divergence in brain features? Three possible explanations for the different rates of divergence observed in the brain and liver datasets should be considered. First, we cannot exclude the possibility that a gene-by-gene description of the patterns of polymorphism and divergence fails to capture the full complexity of transcriptional states. In other words, multivariate transcriptional states between human and chimpanzee brains might be fundamentally different and form the basis for differences in brain functioning between the two primates.

Second, gene expression could show less interindividual variation in the liver than in the brain, with the latter being more responsive to environmental variability (larger genotype-by-environment interaction). In experiments in which the environment cannot be controlled, as in comparisons between humans and chimpanzees, different genotype-by-environment interaction across tissues can confound interpretation of both within- and between-group variance components, as well as evolutionary rates. Our analysis of the distribution of variance components within and between species in the liver and brain tissues suggests that global differences in genotype-by-environment interactions cannot fully explain differences in rates of expression evolution between tissues. A similar distribution of between-species variances (and within-species variances) in both tissues does not allow us to pinpoint either of these as the major cause of the larger rate of divergence observed in the liver. It seems to result, instead, from gene-specific combinations of  $V_b$  and  $V_w$  (Figs. 3B, C).

Third, large differences in human and chimpanzee liver might have resulted from selection for distinct dietary needs. This possibility is relevant in view of the fact that diet has changed dramatically during hominid evolution (Aiello and Wheeler 1995; Johns 1999; Teaford and Ungar 2000) and must have imposed a strong selection pressure to change the metabolic capabilities of humans and chimpanzees (Nebert and Dieter 2000; Aiello and Wells 2002). Evolutionary differences between human and chimpanzee diets might be reflected as a faster rate of evolutionary divergence of liver expressed genes as compared to genes expressed in the brain. Additional evidence supporting the interpretation that diet differences between humans and chimpanzees might have been a major selective pressure comes from recent findings by Clark et al. (2003), who detected positive selection driving the divergence of several genes involved in energy metabolisms in the human lineage.

Finally, we note that even though little morphological change has occurred in the liver of humans and chimpanzees, gene expression variation might be more tightly connected to functional variation in the liver than in the brain. The brain, on the other hand, might be an organ in which intercellular organization, morphological attributes, or higher level features may be more accurate predictors of functional divergence. In other words, humans and chimpanzee brains have clearly diverged in terms of morphological, cognitive, and behavioral features whose low-level correlates might be found at the physiological or neuro-anatomical levels, rather than at the level of transcription profile variation in adult brains.

#### ACKNOWLEDGMENTS

We are grateful to the useful suggestions of M. Nachman and three anonymous reviewers. This work was funded by National Institutes of Health grants to DLH.

#### LITERATURE CITED

- Aiello, L. C., and J. C. K. Wells. 2002. Energetics and the evolution of the genus *Homo*. *Annu. Rev. Anthropol.* 31:323–338.
- Aiello, L. C., and P. Wheeler. 1995. The expensive-tissue hypothesis; the brain and the digestive-system in human and primate evolution. *Curr. Anthropol.* 36:199–221.
- Ashburner, M., C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Davis, K. Dolinski, S. S. Dwight, J. T. Eppig, M. A. Harris, D. P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J. C. Matese, J. E. Richardson, M. Ringwald, G. M. Rubin, G. Sherlock. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* 25:25–29.
- Atchley, W. R., E. V. Nordheim, F. C. Gunsett, and P. L. Crump. 1982. Geometric and probabilistic aspects of statistical distance functions. *Syst. Zool.* 31:445–460.
- Bamshad, M. J., S. Mummidi, E. Gonzalez, S. S. Ahuja, D. M. Dunn, W. S. Watkins, S. Wooding, A. C. Stone, L. B. Jorde, R. B. Weiss, and S. K. Ahuja. 2002. A strong signature of balancing selection in the 5' cis-regulatory region of CCR5. *Proc. Natl. Acad. Sci., USA.* 99:10539–10544.
- Berry, R. J., and F. H. Bronson. 1992. Life-history and bioeconomy of the house mouse. *Biol. Rev. Camb. Philos. Soc.* 67:519–550.
- Bjorklund, M. 1991. Evolution, phylogeny, sexual dimorphism and mating system in the grackles (*Quiscalus* spp.: Icterinae). *Evolution* 45:608–621.
- Bonnin, I., J. M. Prosperi, and I. Olivieri. 1996. Genetic markers and quantitative genetic variation in *Medicago truncatula* (Leguminosae): a comparative analysis of population structure. *Genetics* 143:1795–1805.
- Brem, R. B., G. Yvert, R. Clinton, and L. Kruglyak. 2002. Genetic dissection of transcriptional regulation in budding yeast. *Science* 296:752–755.
- Caceres, M., J. Lachuer, M. Zapala, J. C. Redmond, L. Kudo, D. H. Geschwind, D. J. Lockhart, T. M. Preuss, and C. Barlow. 2003. Elevated gene expression levels distinguish human from non-human primate brains. *Proc. Natl. Acad. Sci. USA* 100:13030–13035.
- Carroll, S. B. 2003. Genetics and the making of *Homo sapiens*. *Nature* 422:849–857.
- Charlesworth, B. 1984. Some quantitative methods for studying evolutionary patterns in single characters. *Paleobiology* 10:308–318.
- Cherry, L. M., S. M. Case, J. G. Kunkel, J. S. Wyles, and A. C. Wilson. 1982. Body shape metrics and organismal evolution. *Evolution* 36:914–933.
- Clark, A. G. 1994. A neutrality test for continuous characters based on levels of intraspecific variation and interspecific divergence.

- Pp. 101–111 in B. Golding, ed. *Non-neutral evolution*. Chapman and Hall, New York.
- Clark, A. G., L. Wang, and T. Hulleberg. 1995. Spontaneous mutation rate of modifiers of metabolism in *Drosophila*. *Genetics* 139:767–779.
- Clark, A. G., S. Glanowski, R. Nielsen, P. D. Thomas, A. Kejarawal, M. A. Todd, D. M. Tanenbaum, D. Civello, F. Lu, B. Murphy, S. Ferreira, G. Wang, X. Zheng, T. J. White, J. J. Sninsky, M. D. Adams, and M. Cargill. 2003. Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios. *Science* 302: 1960–1963.
- Clegg, S. M., S. M. Degnan, C. Moritz, A. Estoup, J. Kikkawa, and I. P. F. Owens. 2002. Microevolution in island forms: the roles of drift and directional selection in morphological divergence of a passerine bird. *Evolution* 56:2090–2099.
- Enard, W., P. Khaitovich, J. Klose, S. Zollner, F. Heissig, P. Giallisco, K. Nieselt-Struwe, E. Muchmore, A. Varki, R. Ravid, G. M. Doxiadis, R. E. Bontrop, and S. Paabo. 2002. Intra- and interspecific variation in primate gene expression patterns. *Science* 296:340–343.
- Falconer, D. S., and T. F. C. Mackay. 1996. *Introduction to quantitative genetics*. Prentice Hall, Englewood Cliffs, NJ.
- Ferea, T. L., D. Botstein, P. O. Brown, and R. F. Rosenzweig. 1999. Systematic changes in gene expression patterns following adaptive evolution in yeast. *Proc. Natl. Acad. Sci. USA* 96: 9721–9726.
- Gingerich, P. D. 1983. Rates of evolution: effects of time and temporal scaling. *Science* 222:159–161.
- . 2001. Rates of evolution on the time scale of the evolutionary process. *Genetica* 112/113:127–144.
- Haldane, J. B. S. 1949. Suggestions as to quantitative measurement of rates of evolution. *Evolution* 3:51–56.
- Hendry, A. P., and M. T. Kinnison. 1999. Perspective: The pace of modern life: measuring rates of contemporary microevolution. *Evolution* 53:1637–1653.
- Hirotsune, S., N. Yoshida, A. Chen, L. Garrett, F. Sugiyama, S. Takahashi, K. Yagami, A. Wynshaw-Boris, and A. Yoshiki. 2003. An expressed pseudogene regulates the messenger-RNA stability of its homologous coding gene. *Nature* 423:91–96.
- Hsieh, W. P., T. M. Chu, R. D. Wolfinger, and G. Gibson. 2003. Mixed-model reanalysis of primate data suggests tissue and species biases in oligonucleotide-based gene expression profiles. *Genetics* 165:747–757.
- 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.
- Johns, T. 1999. The chemical ecology of human ingestive behaviors. *Annu. Rev. Anthropol.* 28:27–50.
- Khaitovich, P., G. Weiss, M. Lachmann, I. Hellmann, W. Enard, B. Muetzel, U. Wirkner, W. Ansorge, and S. Paabo. 2004. A neutral model of transcriptome evolution. *PLoS Biol.* 2:E132.
- Kimura, M. 1968. Evolutionary rate at the molecular level. *Nature* 217:624–626.
- . 1983. *The neutral theory of molecular evolution*. Cambridge Univ. Press, Cambridge, U.K.
- King, J. L., and T. H. Jukes. 1969. Non-darwinian evolution. *Science* 164:788–798.
- Kinnison, M. T., and A. P. Hendry. 2001. The pace of modern life II: from rates of contemporary microevolution to pattern and process. *Genetica* 112:145–164.
- Kirkpatrick, M. 1982. Quantum evolution and punctuated equilibria in continuous characters. *Am. Nat.* 119:833–848.
- Lande, R. 1976. Natural selection and random genetic drift in phenotypic evolution. *Evolution* 30:314–334.
- . 1977. Statistical tests for natural selection on quantitative characters. *Evolution* 31:442–444.
- . 2000. Quantitative genetics and phenotypic evolution. Pp. 335–350 in R. S. Singh and C. B. Krimbas, eds. *Evolutionary genetics*. Cambridge Univ. Press, Cambridge, U.K.
- Lemos, B., G. Marroig, and R. Cerqueira. 2001. Evolutionary rates and stabilizing selection in large-bodied opossum skulls (Didelphimorphia: Didelphidae). *J. Zool.* 255:181–189.
- Lerman, A. 1965. On rates of evolution of unit characters and characters complexes. *Evolution* 19:16–25.
- Luscombe, N. M., J. Qian, Z. Zhang, T. Johnson, and M. Gerstein. 2002. The dominance of the population by a selected few: power-law behaviour applies to a wide variety of genomic properties. *Genome Biol.* 3:R0040.1–R0040.7.
- Lynch, M. 1988. The rate of polygenic mutation. *Genet. Res.* 51: 137–148.
- . 1990. The rate of phenotypic evolution in mammals from the standpoint of the neutral expectation. *Am. Nat.* 136:727–741.
- Lynch, M., and W. G. Hill. 1986. Phenotypic evolution by neutral mutation. *Evolution* 40:915–935.
- Marroig, G., and J. M. Cheverud. 2004. Did natural selection or genetic drift produce the cranial diversification of neotropical monkeys? *Am. Nat.* 163:417–428.
- Martins, E. P. 1994. Estimating the rate of phenotypic evolution from comparative data. *Am. Nat.* 144:193–209.
- Masta, S. E., and W. P. Maddison. 2002. Sexual selection driving diversification in jumping spiders. *Proc. Natl. Acad. Sci. USA* 99:4442–4447.
- Meiklejohn, C. D., J. Parsch, J. M. Ranz, and D. L. Hartl. 2003. Rapid evolution of male-biased gene expression in *Drosophila*. *Proc. Natl. Acad. Sci. USA* 100:9894–9899.
- Merila, J., and P. Crnokrak. 2001. Comparison of genetic differentiation at marker loci and quantitative traits. *J. Evol. Biol.* 14: 892–903.
- Mukai, T., K. Harada, and H. Yoshimaru. 1984. Spontaneous mutations modifying the activity of alcohol dehydrogenase (ADH) in *Drosophila melanogaster*. *Genetics* 106:73–84.
- Nebert, D. W., and M. Z. Dieter. 2000. The evolution of drug metabolism. *Pharmacology* 61:124–135.
- Nuzhdin, S. V., M. L. Wayne, K. L. Harmon, and L. M. McIntyre. 2003. Common pattern of evolution of gene expression level and protein sequence in *Drosophila*. *Mol. Biol. Evol.* 21: 1308–1317.
- 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.
- Podlaha, O., and J. Zhang. 2004. Non-neutral evolution of the transcribed pseudogene makorin1-p1 in mice. *Mol. Biol. Evol.* 21: 2202–2209.
- Powell, J. R. 1997. *Progress and prospects in evolutionary biology: the Drosophila model*. Oxford Univ. Press, New York.
- 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.
- Rifkin, S. A., J. Kim, and K. P. White. 2003. Evolution of gene expression in the *Drosophila melanogaster* subgroup. *Nat. Genet.* 33:138–144.
- Rockman, M. V., and G. A. Wray. 2002. Abundant raw material for cis-regulatory evolution in humans. *Mol. Biol. Evol.* 19: 1991–2004.
- Sandberg, R., R. Yasuda, D. G. Pankratz, T. A. Carter, J. A. Del Rio, L. Wodicka, M. Mayford, D. J. Lockhart, and C. Barlow. 2000. Regional and strain-specific gene expression mapping in the adult mouse brain. *Proc. Natl. Acad. Sci. USA* 97: 11038–11043.
- Savalli, U. M. 1993. An application of the neutral model to the evolution of tail length in the genus *Euplectes* (Aves, Ploceidae). *Evolution* 47:696–699.
- Smith, B. H., and R. L. Tompkins. 1995. Toward a life history of the hominidae. *Annu. Rev. Anthropol.* 24:257–279.
- Sokal, R. R., and F. J. Rohlf. 1997. *Biometry*. W. H. Freeman and Company, New York.
- Spicer, G. 1993. Morphological evolution of the *Drosophila virilis* species group as assessed by rate tests for natural selection on quantitative characters. *Evolution* 47:1240–1254.
- Spitze, K. 1993. Population structure in *Daphnia obtusa*: quantitative genetics and allozymic variation. *Genetics* 135:367–374.
- Storz, J. F. 2002. Contrasting patterns of divergence in quantitative and neutral DNA markers: analysis of clinal variation. *Mol. Ecol.* 11:2537–2551.
- Teaford, M. F., and P. S. Ungar. 2000. Diet and the evolution of

- the earliest human ancestors. *Proc. Natl. Acad. Sci. USA* 97: 13506–13511.
- Toma, D. P., K. P. White, J. Hirsch, and R. J. Greenspan. 2002. Identification of genes involved in *Drosophila melanogaster* geotaxis, a complex behavioral trait. *Nat. Genet.* 31:349–353.
- Townsend, J. P. 2003. Multifactorial experimental design and the transitivity of ratios with spotted DNA microarrays. *BMC Genomics* 4:41.
- Townsend, J. P., and D. L. Hartl. 2002. Bayesian analysis of gene expression levels: statistical quantification of relative mRNA level across multiple strains or treatments. *Genome Biol.* 3:RESEARCH0071.
- Turelli, M., J. H. Gillespie, and R. Lande. 1988. Rate tests for selection on quantitative characters during macroevolution and microevolution. *Evolution* 42:1085–1089.
- Wright, S. 1951. The genetical structure of populations. *Ann. Eugen.* 15:323–354.

Corresponding Editor: M. Nachman