Evolution of gene expression in the *Drosophila melanogaster* subgroup

Scott A. Rifkin[1,2], Junhyong Kim[1,3] & Kevin P. White[2]

Published online 27 January 2003; doi:10.1038/ng1086

Little is known about broad patterns of variation and evolution of gene expression during any developmental process. Here we investigate variation in genome-wide gene expression among *Drosophila simulans*, *Drosophila yakuba* and four strains of *Drosophila melanogaster* during a major developmental transition—the start of metamorphosis. Differences in gene activity between these lineages follow a phylogenetic pattern, and 27% of all of the genes in these genomes differ in their developmental gene expression between at least two strains or species. We identify, on a gene-by-gene basis, the evolutionary forces that shape this variation and show that, both within the transcriptional network that controls metamorphosis and across the whole genome, the expression changes of transcription factor genes are relatively stable, whereas those of their downstream targets are more likely to have evolved. Our results demonstrate extensive evolution of developmental gene expression among closely related species.

Introduction

Although it was suggested nearly 30 years ago that differences in gene regulation may be responsible for differences between closely related species[1], the extent and nature of variation in gene expression and function during development remains a mystery. Regulating gene expression is the key step by which an organism activates the information encoded in its genome to effect developmental change, and differences in this regulation can cascade through development, resulting in different morphological or physiological character states[2–4].

The mechanistic connections among levels of phenotypic variation, for example, between spatial or temporal patterns of gene expression and morphology, determine how development constrains or channels evolution[5,6]. In some systems, developmental processes in general and gene regulation in particular are more conserved than are the phenotypes that they underlie[7–9]; in others, characters themselves are stable but development shifts underneath[10–12]. Given this complexity, it is not clear how much intra- and interspecific variation we should expect in gene expression during development, which evolutionary forces power gene expression change, or how the role of a gene in a transcriptional network might relate to how it will evolve[13].

Previous genomic studies of transcript levels during development have focused on single strains of model organisms (reviewed in ref. 14), and genome-wide surveys of standing variation in gene activity are rare[15–18]. Here we have investigated the stability of developmentally regulated gene expression during evolution in a comparative developmental context by using DNA microarrays[19] to assay the patterns of gene expression variation for 12,866 individual genes (>95% of the predicted genes in the *D. melanogaster* genome) during the onset of metamorphosis in species of the *D. melanogaster* subgroup. We infer the modes by which gene expression evolves and show that the evolution of expression varies depending on the functions of the genes.

Results

Measuring developmental change and variation in gene expression

During metamorphosis, *Drosophila* undergo substantial phenotypic changes and alterations in gene expression (T.-R. Li *et al.*, unpublished data)[20–22], but the transition itself is a conserved process. We measured genome-wide developmental changes in transcript levels at the onset of metamorphosis—that is, between late third instar larvae (18 h before puparium formation) and white prepupae (at puparium formation)—for four inbred strains of *D. melanogaster* (Canton S, Oregon R, Samarkand and Netherlands2) and one inbred strain each of *D. simulans* and *D. yakuba*. We constructed custom DNA microarrays using *D. melanogaster* coding sequences that were expected to differ on average by less than 5% from their orthologs[23] in *D. simulans* and *D. yakuba* (see Web Note A online). *D. yakuba* split from the *D. simulans* and *D. melanogaster* clade around 5.1 million years (Myr) ago, and *D. simulans* and *D. melanogaster* diverged around 2.3 Myr ago[24,25].

[1]*Department of Ecology and Evolutionary Biology, Yale University, PO Box 208106, New Haven, Connecticut 06520-8106, USA.* [2]*Department of Genetics, Yale University School of Medicine, PO Box 208005, New Haven, Connecticut 06520-8005,USA.* [3]*Present address: Department of Biology, University of Pennsylvania, Goddard 206, Philadelphia, Pennsylvania 19104-6018, USA. Correspondence should be addressed to K.P.W. (e-mail: kevin.white@yale.edu).*

Samples from all three species hybridized well to the arrays (Web Fig. A online). Because our comparisons are between developmental stages of the same lineage, minor differences in hybridization kinetics owing to sequence divergence of orthologs cancel out. We measured the developmental change for each lineage four times (or six times for Samarkand). For each gene in each lineage, we estimated the differential expression, as well as the confidence intervals around these estimates, using a general linear model and a bootstrap randomization procedure (Methods)[26].

If the estimate is significantly different from zero, the transcript shows 'developmental change' in that lineage (Fig. 1*a*). If the amount of developmental change is significantly different for a gene between at least two lineages, the gene shows significant (evolutionary) 'variation' in its expression (Fig. 1*b*). We carried out a gene-by-gene analysis of this variation as a first step toward understanding the evolution of genome-wide gene expression.

Evolutionary patterns of gene expression

Overall, transcripts from 6,742 genes (roughly half of the 12,866 genes assayed) change significantly between the time points in at least one lineage (Fig. 1*a*), with very low variation in measurements within a lineage (average s.e.m. 0.13). About half of these genes (3,457; 27% across the genome) differ significantly in the extent of developmental change between at least two lineages (Table 1).

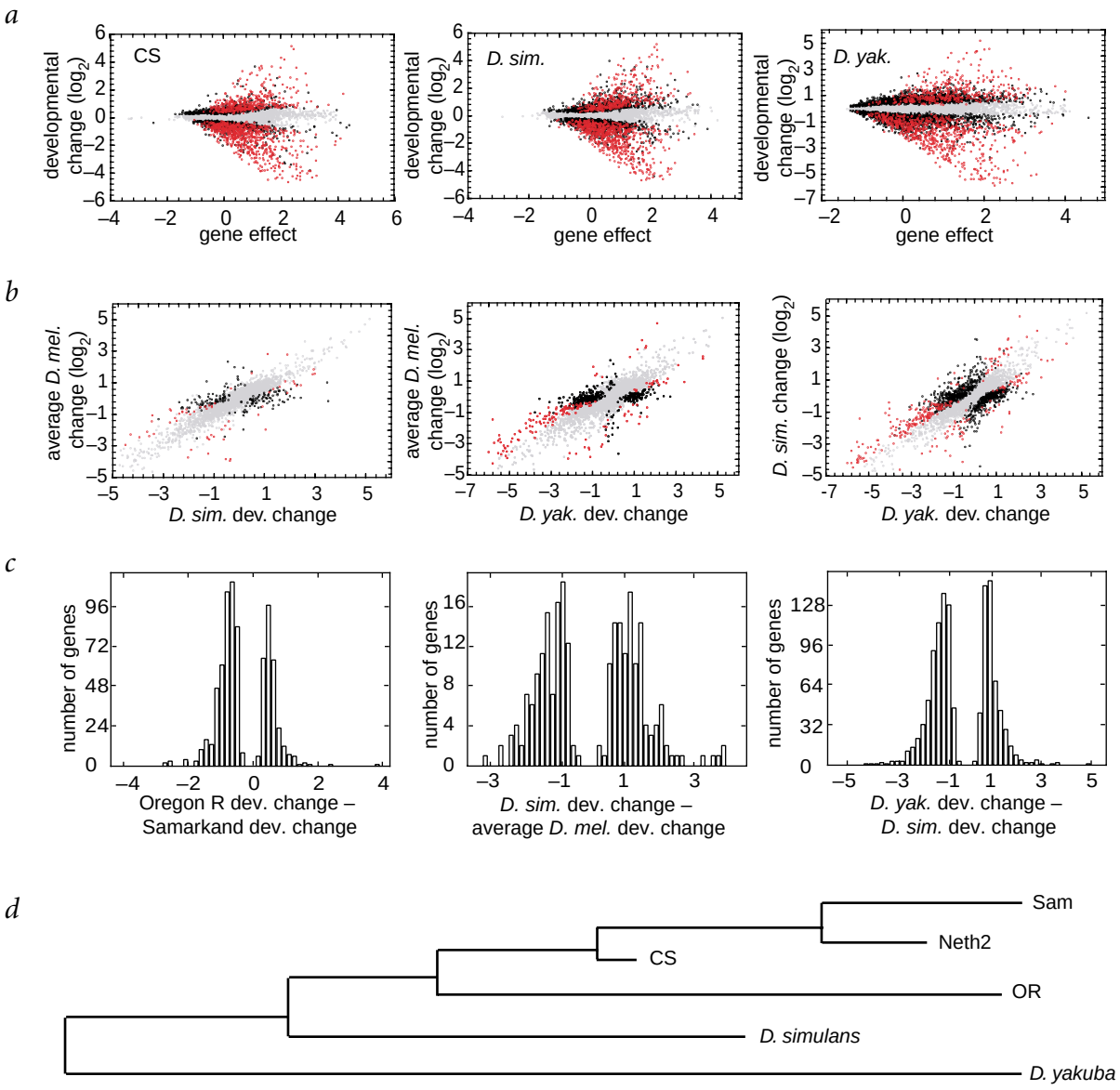


Fig. 1 Developmental and evolutionary patterns of gene expression. ***a***, Differential expression and significant developmental change. In each plot, the *x* axis indicates the gene effect (G_g) for a particular gene and the *y* axis indicates developmental change: that is, the difference between the variety-by-gene effects at puparium formation and at 18 h before puparium formation (Methods)[40]. Black dots mark transcripts that change significantly between the two stages; gray dots mark transcripts that do not change significantly; red dots mark 1,069 transcripts that significantly change in all lineages (380 increase and 689 decrease during onset of metamorphosis in all lineages). ***b***, Comparisons of the estimated developmental changes in the different species. Black dots mark transcripts that show significant differences between the two lineages; gray dots mark transcripts that do not show significant differences (Table 1); red dots mark (left to right) 58, 114 and 270 transcripts that change significantly in all lineages and also differ between the compared lineages. We averaged the values for the *D. melanogaster* strains, and only transcripts that differ between *D. simulans* (*D. yakuba*) and all of the *D. melanogaster* strains are represented. ***c***, Magnitude of the differences in changes in developmental gene expression between lineages. The average *D. melanogaster* values are the same as in *b*. ***d***, A neighbor-joining tree[49] of the lineages based on the gene count distances in Table 1. CS, Cantor S; OR, Oregon R; Sam, Samarkand; Neth2, Netherlands2.

Table 1 • Distances between lineages[a]

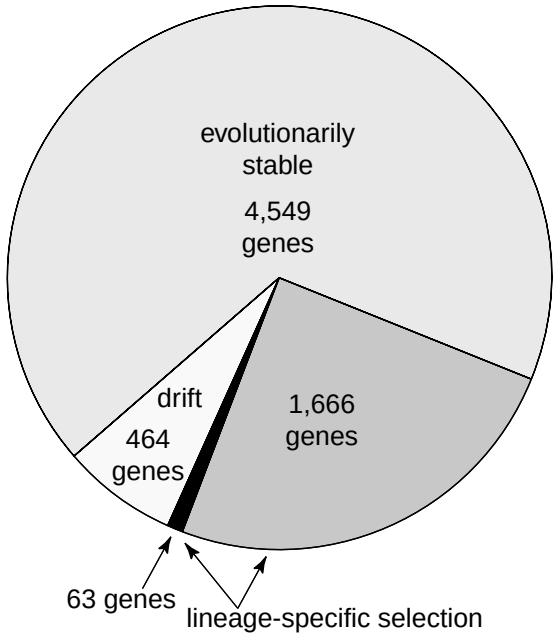
	Canton S	Oregon R	Netherlands2	*D. simulans*	*D. yakuba*
Samarkand	355	753	217	779	1,472
Canton S		549	241	629	1,023
Oregon R			807	827	1,348
Netherlands2				754	1,286
D. simulans					1,162

[a]Shown are the numbers of genes with significantly different changes in developmental expression between two lineages. For each pair-wise comparison, we considered only genes that change developmentally in either lineage: 3,457 distinct genes differ in developmental changes between at least two lineages.

These differences in gene expression are consistent with phylogenetic relationships based on biogeography and sequence data: roughly equal numbers of genes differ between *D. yakuba* and both *D. simulans* and the *D. melanogaster* strains, whereas *D. simulans* is more diverged from the *D. melanogaster* strains than the *D. melanogaster* strains are from each other (Table 1 and Fig. 1*d*).

The transcripts that differ between lineages are not the same in all pair-wise comparisons; the transcript sets overlap but are not nested. For example, the 1,162 transcripts that differ between *D. yakuba* and *D. simulans* and the 1,023 transcripts that differ between *D. yakuba* and Canton S have only 608 transcripts in common. On average, only 26% of the significantly different transcripts of two pairs of lineages overlap. Of the 6,742 genes that change in expression during onset of metamorphosis in at least one lineage, expression for only 1,069 genes increases or decreases in all lineages; most developmentally changing genes are not shared across all of the strains and species that we examined.

The magnitude of variation in developmental gene expression change is generally lower within the *D. melanogaster* strains than between species (Fig. 1*c*). For example, about 38% of the interspecific differences in expression are more than twofold (involving 1,142 genes), whereas only 18% of the intraspecific differences (involving 302 genes) are more than twofold. As would be expected from the phylogenetic relationships among these species, *D. simulans* is more like the *D. melanogaster* strains than is *D. yakuba*. To identify the evolutionary processes generating these extensive gene expression differences among lineages, we examined the variation in developmental changes in gene expression across this clade in more detail.



Categorization of transcripts into evolutionary modes

The developmental change in expression of a particular transcript is a polygenic quantitative character, and its variation across a clade is the product of evolutionary forces experienced during the history of the lineages. To determine whether changes in gene expression during *Drosophila* development are evolving primarily by stabilizing selection, by lineage-specific selection or by drift, we examined their patterns of intra- and interspecific variation (Methods).

Three broad patterns emerged. First, across all six lineages, some genes are evolutionarily stable, showing little variation. Such a pattern is consistent with stabilizing selection across the whole clade (or very low mutational variance). Second, other genes vary little within strains of *D. melanogaster*, but their expression in *D. simulans* and *D. yakuba* differs, suggesting lineage-specific selection in the clade. Third, the developmental changes in expression of some genes are variable within *D. melanogaster*. This pattern could be consistent with either a mutation-drift model for the evolution of gene expression or lineage-specific selection. We devised a sequence of three formal tests to classify genes into these three evolutionary modes according to their intra- and interspecific variation.

We first tested each transcript for evidence of evolutionary stability. In the ideal case of no error in measurement, we would expect the variation to be zero only under infinitely strong stabilizing selection (or zero mutation-drift variance). We devised a test based on the actual measurement error, in which we assessed whether the amounts of variation in the six lineages were greater than that expected from a distribution that would be produced by strong stabilizing selection (Methods). We found that 4,549 of the 6,742 developmentally changing transcripts are evolutionarily stable, consistent with strong stabilizing selection or low mutational variance (Fig. 2). Because this test examines the variation across the whole clade, it differs from the pair-wise comparisons above (1,330 of the 3,457 evolutionarily changing genes identified by pair-wise comparison are under stabilizing selection across the clade).

The remaining 2,193 transcripts are too variable to be under strong stabilizing selection. To determine the evolutionary forces that produce this variation, we used the *D. melanogaster* intraspecific data to divide the genes into those that are polymorphic within the species and those that are stable within *D. melanogaster*.

Fig. 2 Evolutionary modes. Transcripts from 6,742 of the 12,866 genes assayed (52%) change developmentally in at least one lineage. We failed to reject evolutionary stability for 67% of these 6,742 genes. Of the remaining 2,193 genes, 1,666 (25% of the genes whose transcripts showed developmental change) vary little within the four *Drosophila melanogaster* strains examined, implying directional selection. We tested whether the patterns of variation for the remaining 527 genes are consistent with a mutation-drift model for the evolution of differential gene expression. The patterns of variation for 63 genes (0.9% of the developmentally changing genes) are inconsistent with drift, suggesting that they have been subject to lineage-specific selection. We grouped these with the 1,666 genes identified above. Developmental changes in transcript levels of 464 genes (7% of the developmentally changing genes) have evolved in a manner consistent with drift.

Variation across the whole clade with maintenance of stability within *D. melanogaster* implies that a gene's variation comes from species-specific differences. We applied the same test of stability as above, considering only the four *D. melanogaster* strains. For 527 genes, we rejected the hypothesis that the gene expression levels in the *D. melanogaster* lineages are drawn from the same distribution. The remaining 1,666 transcripts differ significantly across the six lineages but are consistent with low variation within *D. melanogaster*. This pattern of variation indicates possible species-specific selection (Fig. 2).

If the developmental changes in transcript levels are in mutation-drift equilibrium, then we would expect the distance between the mean *D. melanogaster* expression changes and the *D. yakuba* and *D. simulans* expression changes to be a function of the mutational variance and the time since divergence[27–29]. Using the *D. melanogaster* variation to estimate the mutational variance, we tested whether the distances between species conformed to these expectations (Methods). We rejected the null hypothesis of neutral evolution for 63 genes (Fig. 2). These genes, which are polymorphic within *D. melanogaster*, are presumably subject to lineage-specific selection within the clade, and we grouped them with the 1,666 genes identified in our second test described above. The expression of the remaining 464 genes is evolving in a pattern consistent with neutral evolution modeled by a diffusion process (Fig. 2 and Methods).

Evolutionary modes of regulatory genes and their targets

We examined evolution of gene expression in the network of genes that controls metamorphosis. The release of the steroid hormone 20-hydroxyecdysone (ecdysone) triggers the onset of metamorphosis by activating a cascade of transcription factors, which in turn selectively activate target genes. More than 30 years ago, experiments showed that chromosomal puffing patterns during metamorphosis differ between members of the *D. melanogaster* subgroup in both the timing and size of puffs[30,31]; however, these differences could not be linked to individual genes before the advent of molecular cloning in *Drosophila*[32], and the relation between puff size and transcript levels remains unclear[33].

Our analysis shows that developmental changes in expression for several of the transcription factor genes activated early in the ecdysone network, including the ecdysone receptor (*EcR*), *Eip75B*, *Eip74EF*, *Hr78* and *crol*, are stable across the clade. By contrast, the developmental changes in transcript levels of many of their targets show evolutionary change: *Eip63E*, *Sgs3*, *Sgs4*, *Sgs5*, *Sgs8*, *ng3*, *Eig71Ec–f*, *Eig71Eh–k* and *Fbp2*, are all under lineage-specific selection; and several genes, including *ng1*, *Uro*, *Hsp23*, *Eig71Ea*, *Eig71Eb* and *Eig71Eg*, are evolving neutrally (Fig. 3).

To determine whether these results apply generally, we tested all genes for associations between the three evolutionary modes identified above and functional categories defined by the Gene Ontology project (Table 2)[34]. Across the whole clade, developmental changes in the expression of genes encoding transcription factors and signal transducers are significantly more likely to be stable than to vary. Conversely, developmental changes in the transcript levels of enzymes and structural factors are significantly likely to have experienced neutral drift. On a genome-wide scale, evolution of gene expression is most likely to occur through changes in the regulation of target genes, owing to the evolution of *cis*-regulatory regions, to the functional evolution of the coding sequences of *trans*-acting factors, or to the evolution of the expression levels of a few regulatory factors that broadly affect the transcription of target genes.

We also investigated the evolution of gene expression in individual tissues using genomic transcriptional data from these same two time points. We focused on the midgut, salivary glands and epidermis with attached connective tissue, in which sufficient numbers of developmentally changing transcripts have been identified to allow us to do adequate tests (T.-R. Li *et al.*, unpublished data). At this stage of development, these tissues express mainly non-overlapping sets of genes that reflect the

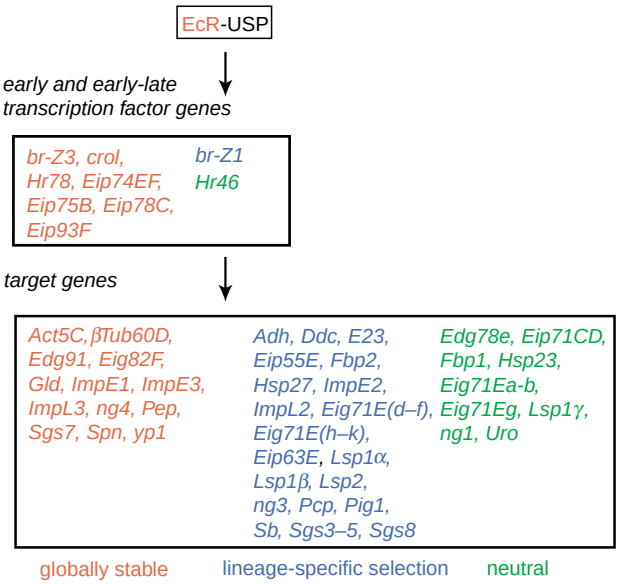


Fig. 3 Evolutionary modes of the ecdysone-regulated genes. Ecdysone binds to the EcR-USP heterodimer, which activates a transcriptional cascade[50]. A group of early transcription factor genes is activated directly by the ecdysone receptor complex, including *Eip74EF*, *Eip75B*, *Eip93F*, *br*, *crol* and *Hr78*. A small set of early-late transcription factor genes, such as *Eip78C* and *Hr46*, is also activated by ecdysone but slightly after the early genes. The coordinated actions of the ecdysone receptor complex, the early genes and the early-late genes regulate several late genes to control the progression of metamorphosis. As compared with their target genes, three times as many early and early-late transcription factor genes are evolutionarily stable.

Table 2 • Associations between evolutionary modes and gene function[a]

	Stable	Lineage-specific	Neutral	Total
Transcription factor	156/117****	16/44 ****	1/12**	173
Signal transducer	232/190****	44/72***	5/19**	281
Enzyme	1,003/1,064**	417/404 NS	157/109****	1577
Structural protein	94/112*	48/43 NS	24/11**	166

[a]Associations between Gene Ontology functional categories (http://www.godatabase.org) and evolutionary modes were assessed by 2×2 Fisher's exact tests[47]. Each cell records the number of genes observed over the number expected by random assortment. The results were significant at *$\alpha = 0.05$ (with a Dunn-Sidak multiple-test correction P-value of < 0.0043), **$\alpha = 0.01$ ($P < 0.00084$), ***$\alpha = 0.001$ ($P < 8.4 \times 10^{-5}$), ****$\alpha = 0.0001$ ($P < 8.4 \times 10^{-6}$), *****$\alpha = 1.0 \times 10^{-10}$ ($P < 8.4 \times 10^{-12}$). NS, not significant.

Table 3 • Evolutionary modes of developmentally changing transcripts in tissues[a]

	Stable	Lineage-specific	Neutral	Total
Epidermis	244/377*****	189/143****	125/38*****	558
Midgut	219/332*****	168/126***	105/34*****	492
Salivary gland	177/247*****	124/94**	65/25*****	366

[a]Associations between transcripts that significantly increased or decreased in three tissues (T.-R. Li *et al.*, unpublished data) and evolutionary mode assignments were assessed by 2×2 Fisher's exact tests[47]. Each cell records the number of genes observed over the number expected by random assortment. To be considered, each gene had to be developmentally changing in the Canton S data reported here: its dynamics had to be detectable in the whole animal. The results were significant at *$\alpha = 0.05$ (with a Dunn-Sidak multiple test correction P-value of < 0.0057), **$\alpha = 0.01$ ($P < 0.0011$), ***$\alpha = 0.001$ ($P < 0.00011$), ****$\alpha = 0.0001$ ($P < 1.1 \times 10^{-5}$), *****$\alpha = 1.0 \times 10^{-10}$ ($P < 1.1 \times 10^{-11}$).

functional state of each tissue (T.-R. Li *et al.*, unpublished data). In all three tissues, larval cells will be replaced by adult cells as metamorphosis proceeds, although this process occurs at different times in each of the tissues examined[35].

We found that developmental changes in gene expression are significantly likely to be drifting or under lineage-specific selection in these tissues (Table 3). In the midgut, 48 of 74 developmentally changing endopeptidase genes are changing evolutionarily, whereas 7 of the 13 genes encoding cytochrome P450s are drifting. Most genes encoding structural proteins of the cuticle (17 of 23) show evolutionary changes in expression. The extreme bias for certain classes of gene to show evolution of gene expression in these tissues is associated with large changes in expression during the onset of metamorphosis. We found that the overall magnitude of developmental change is significantly correlated with evolutionary mode (Web Fig. B online). Genes with high absolute developmental changes are more likely to be subject to lineage-specific selection and drift.

The direction of expression change and a given gene's mode of expression evolution are highly associated (Table 4). Transcripts increasing during this period in these tissues are more likely to be evolutionarily stable, whereas those declining during this period are less stable. The association also holds across the whole dataset (Table 4). We propose that this pattern reflects stabilization of the induction of transcripts responsible for initiating metamorphosis, and weaker constraints on the temporal regulation of genes that are no longer needed and that do not directly affect the expression of others. In this light, the developmental changes of transcription factor transcripts should be stable—as they overwhelmingly are—and those that vary across the clade merit further attention.

Finally, we considered the evolution of the 1,069 genes that show similar developmental changes in expression across all lineages. Because these genes are developmentally regulated in a conserved developmental process in the *D. melanogaster* subgroup, the dominant mode of evolution might be expected to be stabilizing selection. Instead, we found that these genes are more likely to be drifting or under lineage-specific selection than to be stable. These 1,069 genes encode an excess of endopeptidases (85 observed/35 expected), enzymes (348/250) and structural proteins (56/26), and a paucity of transcription factors (5/27; all observations are significant, with $P < 0.001$). Overall, we found that the transcript levels of these genes are more likely to decline, than to increase, from late larva to prepupa in all of the lineages, and a quarter of those declining are drifting. By contrast, the evolutionary modes of the upregulated transcripts from this set of 1,069 developmentally changing genes are more representative of the whole dataset: we found that most upregulated genes are evolutionarily stable, fewer show lineage-specific selection and even fewer seem to be drifting. Changing developmental expression across all lineages is itself not an indicator of evolutionary stability.

Discussion

Phenotypic evolution is both constrained and driven by variation in gene function during development. Because gene expression itself varies, it is an object of evolution in its own right[36]. Determining the proximate causes of this variation—the interactions between *trans*-regulatory factors and *cis*-regulatory sequences—will be crucial for understanding how differences in gene function drive the evolution of multigenic traits[37,38]. In the *D. melanogaster* subgroup, developmental changes in gene expression vary extensively both within and between species. This variation provides abundant targets for selection and ample fuel for evolution.

Much of the genome-wide variation in expression results from changing developmental constraints, which shift the balance between different evolutionary forces. There are stronger constraints on the activation of gene expression than on the downregulation or degradation of transcripts during development. There are stronger constraints on the expression of genes that encode regulatory molecules than on the expression of genes that encode structural factors or enzymes. There are stronger constraints on genes with small changes in expression during development than on those with larger changes. Because of these constraints, interpreting the biological significance of gene expression variation and evolution will be difficult or impossible outside its developmental context.

As more studies investigate variation in genomic regulatory networks, we will be able to trace the impact of gene expression variation on network output and stability to its developmental consequences. Ultimately, genome-wide functional analyses will help to delineate the relationships among molecular variation, the organismal phenotype and specific evolutionary forces.

Table 4 • Evolution of upregulated and downregulated genes[a]

	Tissues ($P << 1.0 \times 10^{-10}$)				Whole animal ($P << 1.0 \times 10^{-10}$)			
	Stable	Lineage-specific	Neutral	Total	Stable	Lineage-specific	Neutral	Total
Increasing	267	159	55	481	2,612	717	105	3,434
Decreasing	206	220	192	618	1,899	885	321	3,085
Total	473	379	247	1,099	4,511	1,602	426	6,519

[a]Associations between the direction of developmental expression change and evolutionary stability were assessed by a *G*-test. The genes increasing in tissues are increasing in at least one Canton S tissue and decreasing in none. Similarly, the genes increasing in whole animals were increasing in at least one lineage and decreasing in none. Genes with decreasing transcript levels are more likely to be evolutionarily changing than those with increasing levels.

Methods

Materials. We grew *D. melanogaster* strains Canton S, Oregon R, Samarkand and the isofemale line Netherlands2, *D. simulans* strain 14021-0251.167, and *D. yakuba* strain 14021-0261.0 in uncrowded conditions on standard corn meal medium with 0.05% bromophenol blue. All of these lineages have been under laboratory conditions for at least 3 years (Netherlands2) and in some cases for decades (Oregon R, Canton S). At the dark-blue gut stage[39,40], on average 18 h before pupariation and at pupariation (white prepupa), we collected roughly 60 flies for each sample in SDS lysis buffer and subjected them to phenol-chloroform total RNA extraction and mRNA extraction (Promega)[40]. The mRNA was reverse-transcribed in the presence of fluorescently labeled dUTP, and the labeled cDNA was hybridized at 65 °C to the whole-genome *D. melanogaster* microarrays described below.

We printed poly-L-lysine-coated arrays on a Gene Machines arrayer using PCR fragments of ~300 bp of predicted genes from the *D. melanogaster* genome. The fragments were selected to minimize intrafamily hybridization (see Web Note A online). We processed the arrays after printing using 1-2-dichloroethane, *N*-methylimidazol and succinic anhydride[41]. We limited our analyses to the 12,866 genes common to each array. For some genes, several different sequences were spotted on the arrays (see Web Note A online). If preliminary *t*-tests indicated that the spots were not behaving similarly, we treated the spots as separate genes (such as potential splice variants) and have marked the 19 such cases in the data by 'S'. Otherwise, the measurements were averaged in the following analysis.

Microarray analysis. After scanning the arrays with a confocal laser scanner (Axon), we analyzed the images using Spot[42] without any background correction (so as not to introduce additional variance), yielding log base-2 ($\log_2$) measurements for mean intensities for each channel in each spot on the array. All subsequent analyses were coded in python with links to the *R* statistical package (see URLs below). For each array, we fitted a loess curve (span = 0.25) to the difference in intensities between dyes as a function of the mean of the intensities of the two dyes. Other groups have shown that such curves can differ between blocks on the array[43]. We did not find such differences, but, because the intensity difference between the dyes overall varied with mean spot intensity, we shifted the data by the deviation of the loess curve from the mean line to account for this array-specific nonlinear effect of dyes. The final data set consisted of this adjusted data for the four (or six) replicates per population.

To estimate the differences in gene expression between the two conditions while removing noise, we fitted the data to a general linear model of the form:

$$y_{ijkg}=\mu+A_i+D_j+G_g+AD_{ij}+AG_{ig}+VG_{kg}+\varepsilon_{ijkg}$$

where y_{ijkg} is the $\log_2$ measurement for a particular gene (g) under a particular condition (k) labeled with a particular dye (j) on a particular array (i), μ is the overall mean, A, D and G measure the effects of the single factors array, dye and gene, AD measures the array-by-dye interaction, AG measures the array-by-gene effect, VG measures the variety-by-gene effect (the effect of interest), and ε is the residual between the data and the model[44]. For comparisons in which sequence divergence might be appreciable, it is crucial to be able to separate hybridization variation caused by sequence divergence from actual differences in mRNA abundances at particular stages. In our model, sequence divergence would contribute to the G effect and should not affect the VG estimate. To minimize the effects of outliers, we estimated the effects μ, A, D and AD using the middle 80% of the measurements. We estimated missing values by filling them in with the mean for that spot on the other arrays, fitting the model, and then iterating this process using the model estimates to update the missing measurements until the whole model converged.

To assign confidence intervals to the developmental changes in expression between stages, we bootstrapped the data set by randomly swapping error terms between measurements and re-estimating the general linear model according to a described scheme[26]. A gene's expression was taken to change during development if zero did not fall within the 95% confidence interval. The confidence intervals averaged around 0.73 units (average s.e.m. 0.13) in width on a $\log_2$ scale. A gene was taken to change significantly during evolution if the confidence interval for neither lineage included the estimate of the mean developmental change for the other (see Note online).

Evolutionary mode analysis. The conventional null hypothesis for tests of evolutionary mode is a mutation-drift model that requires either extensive sampling of the patterns of divergence[45,46] or a direct estimate of mutational variance[28]. If we were to use the observed data to estimate mutational variance, derived for example from the *D. melanogaster* variation, then we would fail to reject neutrality for several cases in which stabilizing selection has limited the variation across the whole clade, including within *D. melanogaster*, either by eliminating deleterious phenotypes or by favoring some optimal phenotype. If the measurements were nearly identical, the noise variance would become the baseline for rates of evolution. To limit such false negatives, we first used an independent criterion to test whether the variation across the lineages was small. This formed the basis of our first two tests. For the remaining genes, we attempted to reject a mutation-drift model on the basis of estimating the mutational variance from the observed data. Because these lineages are laboratory strains, it is possible that the patterns of variation, especially low amounts of variation, do not wholly reflect the natural intra- and interspecific variation in this group.

Stabilizing and lineage-specific selection tests. If we could measure developmental changes in gene expression in these lineages with infinite precision, identical values would arise from either infinitely strong stabilizing selection or no mutational variance. Any deviation would indicate that stabilizing selection could not eliminate all of the variation introduced by mutation. With imperfect data, however, measurement error sets the limit for whether we can discriminate means and sets a threshold (less than infinite in this case) for the strength of stabilizing selection that we can detect. With a combination of strong stabilizing selection and low mutational variance, the estimates of developmental change across the various lineages would look as though they were six random samples from some shared error distribution around a common mean. The null hypothesis for these tests was therefore the presence of selection, and thus the number of developmental expression changes under selection might be overestimated. Estimates of mutational variance will help to overcome this caveat, enabling us to use more standard tests for selection using neutrality as the null hypothesis. Given the measurement error and the differences between this test and the pair-wise test described above, the transcript levels for a gene in two of these samples could differ from each other even if the gene were stable when considering the whole clade (1,330 of 3,457 evolutionarily different genes were under stabilizing selection). It is also possible to reject a common distribution even if the pair-wise test did not identify significant differences in a the change in expression of a gene (66 of the 1,729 genes under lineage-specific selection were not identified as different evolutionarily in the pair-wise analysis).

We estimated the common standard deviation, $\hat{\sigma}$, by averaging the confidence intervals (weighted by the sample size) for each gene across the lineages and assuming that the width of this average confidence interval corresponded to $4\hat{\sigma}$. We then estimated the average change in developmental gene expression, $\bar{x}_{1-6}$, accounting for the tree structure, and constructed the normal distribution $N(\bar{x}_{1-6}, \hat{\sigma})$. For each lineage estimate, we calculated the probability of picking it or a more extreme value from this distribution, and combined these probabilities to assess the overall significance of the pattern of gene expression for that gene[47]. If the combined probability were less than 0.05, we rejected the null hypothesis of a common distribution for the estimates.

Mutation-drift test. Under mutation-drift equilibrium, the expected squared difference between the phenotypes of two lineages is $\sigma_m^2 t$, where t is the amount of time that separates them and σ_m^2 is the mutational variance. For *D. melanogaster* and *D. yakuba*, t is roughly 10.2 Myr; and for *D. melanogaster* and *D. simulans*, t is roughly 4.6 Myr[25,48]. The genetic variance within *D. melanogaster* is a function of the variance introduced since the lineages diverged and the amount of variance in the ancestral population. For an emigration from Africa about 15,000 years ago[25], ten generations per year and an effective population size of 3 million individuals[48], the amount of intraspecific genetic variance for these *D. melanogaster* strains can be approximated by their ancestral variance, which is around $2N_e\sigma_m^2$ or $6 \times 10^6\sigma_m^2$ (with units scaled to generations)[28]. The ratio of the squared difference between the phenotype of *D. yakuba* (*D. simulans*) and the mean phenotype of *D. melanogaster* to

the variance within *D. melanogaster* scaled by the ratio of their expectations, $10.2 \times 10^7\ \sigma_m^2/6 \times 10^6\ \sigma_m^2$ ($4.6 \times 10^7\ \sigma_m^2 / 6 \times 10^6\ \sigma_m^2$, should be distributed as an $F_{[1,3]}$ distribution:

$$F_{(yak-mel)} \sim \frac{\mathrm{dist}_{(yak-mel)}^2}{\mathrm{var}_{mel}} \cdot \frac{2N_e\sigma_m^2}{\sigma_m^2 t_{(yak-mel)}} \Rightarrow \frac{\mathrm{dist}_{(yak-mel)}^2}{17 \cdot \mathrm{var}_{mel}}$$

$$F_{(sim-mel)} \sim \frac{\mathrm{dist}_{(sim-mel)}^2}{\mathrm{var}_{mel}} \cdot \frac{2N_e\sigma_m^2}{\sigma_m^2 t_{(sim-mel)}} \Rightarrow \frac{3 \cdot \mathrm{dist}_{(sim-mel)}^2}{23 \cdot \mathrm{var}_{mel}}$$

We rejected neutrality for test statistics that fell in the outer 2.5% tails for either the *D. yakuba* or the *D. simulans* comparison.

URLs. Python, http://www.python.org; *R* statistical package, http://www.r-project.org; GEO database, http://www.ncbi.nlm.nih.gov/geo/. For additional information about this study and access to the array data, see http://flygenome.yale.edu/Comparative/.

Accession numbers. GEO microarray data, GSE129, GSE130, GSE131, GSE132, GSE133, GSE134.

Note: Supplementary information is available on the Nature Genetics website.

Acknowledgments

We thank A. Davis for the Netherlands2 line, and M. Feldman, B. Null, P. Lizardi, J. Leamon, P. Magwene, G. Wagner, S. Rice, J.D.Lambert, T.-R. Li, N. Carriero and members of the Kim and White laboratories for advice, support and technical help. This research was supported by the National Institutes of Health.

Competing interests statement

The authors declare that they have no competing financial interests.

Received 22 July 2002; accepted 2 January 2003.

1. King, M.C. & Wilson, A.C. Evolution at two levels in humans and chimpanzees. *Science* **188**, 107–116 (1975).
2. Davidson, E.H. Genomic Regulatory Systems: Development and Evolution (Academic, San Diego, 2001).
3. Johnson, N.A. & Porter, A.H. Toward a new synthesis: population genetics and evolutionary developmental biology. *Genetica* **112**, 45–58 (2001).
4. Wilkins, A.S. *The Evolution of Developmental Pathways* (Sinauer, New York, 2001).
5. Dickinson, W.J. On the architecture of regulatory systems—evolutionary insights and implications. *BioEssays* **8**, 204–208 (1988).
6. Wagner, G.P. Homologues, natural kinds and the evolution of modularity. *Am. Zool.* **36**, 36–43 (1996).
7. Dickinson, W.J., Yang, Y.F., Schuske, K. & Akam, M. Conservation of molecular prepatterns during the evolution of cuticle morphology in *Drosophila* larvae. *Evolution* **47**, 1396–1406 (1993).
8. McGinnis, W., Garber, R.L., Wirz, J., Kuroiwa, A. & Gehring, W.J. A homologous protein-coding sequence in *Drosophila* homeotic genes and its conservation in other metazoans. *Cell* **37**, 403–408 (1984).
9. Alonso, C.R., Maxton-Kuechenmeister, J. & Akam, M. Evolution of Ftz protein function in insects. *Curr. Biol.* **11**, 1473–1478 (2001).
10. Wimmer, E.A., Carleton, A., Harjes, P., Turner, T. & Desplan, C. *bicoid*-independent formation of thoracic segments in *Drosophila*. *Science* **287**, 2476–2479 (2000).
11. Wagner, G.P. & Gauthier, J.A. 1,2,3 = 2,3,4: a solution to the problem of the homology of the digits in the avian hand. *Proc. Natl. Acad. Sci. USA* **96**, 5111–5116 (1999).
12. Wagner, G.P. & Misof, B.Y. How can a character be developmentally constrained despite variation in developmental pathways? *J. Evol. Biol.* **6**, 449–455 (1993).
13. Palopoli, M.F., Davis, A.W. & Wu, C.I. Discord between the phylogenies inferred from molecular versus functional data: uneven rates of functional evolution or low levels of gene flow? *Genetics* **144**, 1321–1328 (1996).
14. Reinke, V. & White, K. Developmental genomic approaches in model organisms. *Annu. Rev. Genom. Hum. G.* **3**, 153–178 (2002).
15. Cavalieri, D., Townsend, J.P. & Hartl, D.L. Manifold anomalies in gene expression in a vineyard isolate of *Saccharomyces cerevisiae* revealed by DNA microarray analysis. *Proc. Natl. Acad. Sci. USA* **97**, 12369–12374 (2000).
16. Hughes, T.R. *et al*. Functional discovery via a compendium of expression profiles. *Cell* **102**, 109–126 (2000).
17. Damerval, C., Maurice, A., Josse, J.M. & Devienne, D. Quantitative trait loci underlying gene-product variation—a novel perspective for analyzing regulation of genome expression. *Genetics* **137**, 289–301 (1994).
18. Jin, W. *et al*. The contributions of sex, genotype and age to transcriptional variance in *Drosophila melanogaster*. *Nat. Genet.* **29**, 389–395 (2001).
19. Schena, M., Shalon, D., Davis, R.W. & Brown, P.O. Quantitative monitoring of gene-expression patterns with a complementary-DNA microarray. *Science* **270**, 467–470 (1995).
20. Riddiford, L.M. in *The Development of* Drosophila melanogaster Vol. 2 (eds. Bate, M. & Arias, A.M.) 899–940 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1993).
21. White, K.P., Rifkin, S.A., Hurban, P. & Hogness, D.S. Microarray analysis of *Drosophila* development during metamorphosis. *Science* **286**, 2179–2184 (1999).
22. Arbeitman, M.N. *et al*. Gene expression during the life cycle of *Drosophila melanogaster*. *Science* **297**, 2270–2275 (2002).
23. Moriyama, E.N. & Powell, J.R. Intraspecific nuclear DNA variation in *Drosophila*. *Mol. Biol. Evol.* **13**, 261–277 (1996).
24. Lachaise, D. *et al*. Historical biogeography of the *Drosophila melanogaster* species subgroup. *Evol. Biol.* **22**, 159–225 (1988).
25. Powell, J.R. *Progress and Prospects in Evolutionary Biology. The* Drosophila *Model* (Oxford Univ. Press, New York, 1997).
26. Kerr, M.K. *et al*. Statistical analysis of a gene expression microarray experiment with replication. *Stat. Sin.* **12**, 203–217 (2002).
27. Lande, R. Natural selection and random genetic drift in phenotypic evolution. *Evolution* **30**, 314–334 (1976).
28. Lynch, M. & Hill, W.G. Phenotypic evolution by neutral mutation. *Evolution* **40**, 915–935 (1986).
29. Turelli, M., Gillespie, J.H. & Lande, R. Rate tests for selection on quantitative characters during macroevolution and microevolution. *Evolution* **42**, 1085–1089 (1988).
30. Ashburner, M. Patterns of puffing activity in salivary gland chromosomes of *Drosophila*. 3. A comparison of autosomal puffing patterns of sibling species *D. melanogaster* and *D. simulans*. *Chromosoma* **27**, 156–177 (1969).
31. Ashburner, M. & Lemeunie, F. Patterns of puffing activity in salivary-gland chromosomes of *Drosophila*.7. Homology of puffing patterns on chromosome arm 3L in *D. melanogaster* and *D. yakuba*, with notes on puffing in *D. teissieri*. *Chromosoma* **38**, 283–295 (1972).
32. Wensink, P.C., Finnegan, D.J., Donelson, J.E. & Hogness, D.S. System for mapping DNA sequences in chromosomes of *Drosophila melanogaster*. *Cell* **3**, 315–325 (1974).
33. Stowers, R.S., Russell, S. & Garza, D. The 82F late puff contains the *L82* gene, an essential member of a novel gene family. *Dev. Biol.* **213**, 116–130 (1999).
34. The Gene Ontology Consortium. Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).
35. Robertson, C.W. The metamorphosis of *Drosophila melanogaster,* including an accurately timed account of the principal morphological changes. *J. Morphol.* **59**, 351–399 (1939).
36. Lewontin, R.C. The organism as the subject and object of evolution. *Scientia* **188**, 65–82 (1983).
37. Toma, D.P., White, K.P., Hirsch, J. & Greenspan, R.J. Identification of genes involved in *Drosophila melanogaster* geotaxis, a complex behavioral trait. *Nat. Genet.* **31**, 349–353 (2002).
38. White, K.P. Functional genomics and the study of development, variation, and evolution. *Nat. Rev. Genet.* **2**, 528–537 (2001).
39. Maroni, G. & Stamey, S.C. Use of blue food to select synchronous late 3rd instar larvae. *Drosoph. Inf. Serv.* **59**, 142–143 (1983).
40. Andres, A.J. & Thummel, C.S. Methods for quantitative analysis of transcription in larvae and prepupae. *Methods Cell Biol.* **44**, 565–573 (1994).
41. Diehl, F., Grahlmann, S., Beier, M. & Hoheisel, J.D. Manufacturing DNA microarrays of high spot homogeneity and reduced background signal. *Nucleic Acids Res.* **29**, e38 (2001).
42. Yang, Y.H., Buckley, M.J., Dudoit, S. & Speed, T.P. Comparison of methods for image analysis on cDNA microarray data. *J. Comput. Graph. Stat.* **11**, 108–136 (2002).
43. Yang, Y.H., Dudoit, S., Luu, P. & Speed, T.P. Normalization for cDNA microarray data. *Proc. SPIE* **4266**, 141–152 (2001).
44. Kerr, M.K., Martin, M. & Churchill, G.A. Analysis of variance for gene expression microarray data. *J. Comput. Biol.* **7**, 819–837 (2000).
45. Martins, E.P. Estimating the rate of phenotypic evolution from comparative data. *Am. Nat.* **144**, 193–209 (1994).
46. Hansen, T.F. & Martins, E.P. Translating between microevolutionary process and macroevolutionary patterns: the correlation structure of interspecific data. *Evolution* **50**, 1404–1417 (1996).
47. Sokal, R.R. & Rohlf, F.J. *Biometry* (W.H. Freeman and Company, New York, 1995).
48. Li, Y.J., Satta, Y. & Takahata, N. Paleo-demography of the *Drosophila melanogaster* subgroup: application of the maximum likelihood method. *Genes Genet. Syst.* **74**, 117–127 (1999).
49. Swofford, D.L. PAUP*: phylogenetic analysis using parsimony (and other methods), edn 4.0 (Sinauer, Sunderland, MA, 1996).
50. Thummel, C.S. Molecular mechanisms of developmental timing in *C. elegans* and *Drosophila*. *Dev. Cell* **1**, 453–465 (2001).