nature Vol 438|10 November 2005|doi:10.1038/nature04114

LETTERS

A mutation accumulation assay reveals a broad capacity for rapid evolution of gene expression

Scott A. Rifkin[1,2]†, David Houle[3], Junhyong Kim[4] & Kevin P. White[1,2]

Mutation is the ultimate source of biological diversity because it generates the variation that fuels evolution[1]. Gene expression is the first step by which an organism translates genetic information into developmental change. Here we estimate the rate at which mutation produces new variation in gene expression by measuring transcript abundances across the genome during the onset of metamorphosis in 12 initially identical *Drosophila melanogaster* lines that independently accumulated mutations for 200 generations[2]. We find statistically significant mutational variation for 39% of the genome and a wide range of variability across corresponding genes. As genes are upregulated in development their variability decreases, and as they are downregulated it increases, indicating that developmental context affects the evolution of gene expression. A strong correlation between mutational variance and environmental variance shows that there is the potential for widespread canalization[3]. By comparing the evolutionary rates that we report here with differences between species[4,5], we conclude that gene expression does not evolve according to strictly neutral models. Although spontaneous mutations have the potential to generate abundant variation in gene expression, natural variation is relatively constrained.

The expression level of a gene is a polygenic, dynamic, quantitative trait, and its functional significance can be assessed by studying genetic variation both within[6,7] and between (for example, ref. 4) populations or species. Such comparisons need to be calibrated with respect to the effects of mutation. This is commonly measured by the mutational variance (V_m): that is, the per-generation increase in the variance of a trait across a population that is due only to mutation.

On the basis of studies in *D. melanogaster*[2,8], we estimate that each of our 12 mutation accumulation lines contains around 360 mutations. We measured gene expression levels during the third larval instar (before puparium formation; BPF) and at puparium formation (PF), before and after the peak of a large pulse of the hormone 20-hydroxyecdysone that triggers the start of metamorphosis (see Methods and Supplementary Fig. 1). This stage is one of substantial transcriptional activity and turnover[9,10], with broad intra- and interspecific variation in gene expression[4]. Of 11,798 genes measured, we detected significant V_m for 3,816 genes at the BPF stage, for 3,475 genes at the PF stage and for 4,658 genes overall, using a false discovery rate (FDR) of 0.05. The expression of 5,729 genes significantly differed between the two stages, although only 2,509 of these genes showed significant V_m (FDR = 0.05).

To compare the expression variability of different genes, we scaled estimates of V_m by the residual variance (V_r) to give the mutational heritability (h^2_m) (refs 1, 11). Technical variance is not a significant factor (see Supplementary Methods), so this V_r could arise from inherent physiological variability of expression and/or temporal asynchrony of the sampled individuals. The 95% interval of h^2_m for gene expression ranged from 2.7×10^{-6} to 1.2×10^{-4} with medians of 2.5×10^{-5} and 2.3×10^{-5} for BPF and PF, respectively (Fig. 1). These estimates lie at the low end of the variability spectrum, overlapping life-history traits such as viability in *D. melanogaster* and grain yield in barley[12]. Although the h^2_m values were relatively low, the transcript abundances of roughly three-fifths of the 7,878 genes with measurable mRNA levels in these two stages varied among our lines, which is ample material for rapid evolutionary change (Supplementary Data 1).

Some authors have claimed that gene expression evolves neutrally on the basis of correlations between sequence and expression divergence[13] or on the basis of functional divergence of orthologous genes between humans and mice[14]. Other studies suggest that stabilizing selection has a more important role[4,15–17]. The expected neutral divergence between species depends on the baseline mutation rate for gene expression, which we can derive from the estimates of V_m in the mutation accumulation lines (see Methods)[18]. Species divergences significantly smaller than this expectation would be consistent with either stabilizing selection within each species towards a shared (ancestral) value or other kinds of constraint that prevent neutral divergence.

Under mutation-drift equilibrium, in the absence of selection, the expected difference between phenotypic values in two lineages is $\sqrt{(2tV_m)}$, where t is the number of generations since their common ancestor[18]. Using our estimates for V_m, we calculated the expected differences in changes in expression levels from BPF to PF between

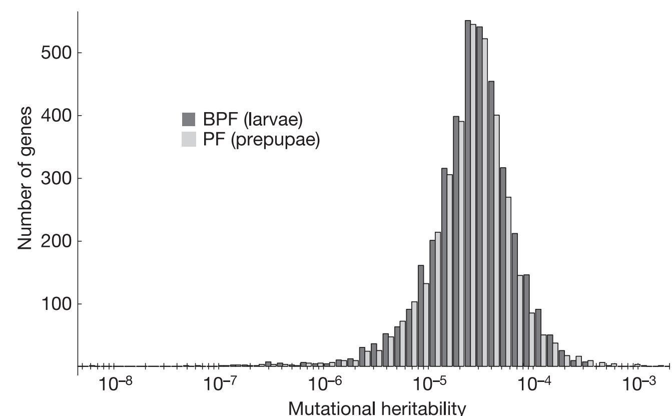

Figure 1 | Mutational heritability. Shown are histograms of the mutational heritability (h^2_m) at the two stages BPF and PF.

[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]Department of Biological Science, Florida State University, Tallahassee, Florida 32306-1100, USA. [4]Department of Biology, University of Pennsylvania, 203 Goddard Laboratories, 415 South University Avenue, Philadelphia, Pennsylvania 19104, USA. †Present address: Department of Organismic and Evolutionary Biology, Harvard University, 16 Divinity Avenue, Cambridge, Massachusetts 02138, USA.

© 2005 Nature Publishing Group

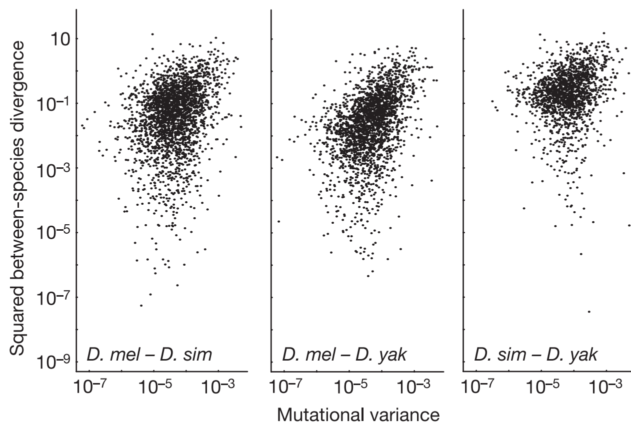

Figure 2 | Mutational variance and differences between species. Shown are log–log plots of V_m of the developmental change between BPF and PF (*x* axis) versus the squared difference of this change between species (*y* axis). Measurements of *D. melanogaster* used in the between-species comparisons are from the data in this study, as are the V_m estimates. Measurements of *D. simulans* and *D. yakuba* are from an extended version of published data[4,5] (Supplementary Methods). r_s for genes with significant V_m in either stage: 0.27 (*D. melanogaster* versus *D. simulans*), 0.45 (*D. melanogaster* versus *D. yakuba*), 0.22 (*D. yakuba* versus *D. simulans*).

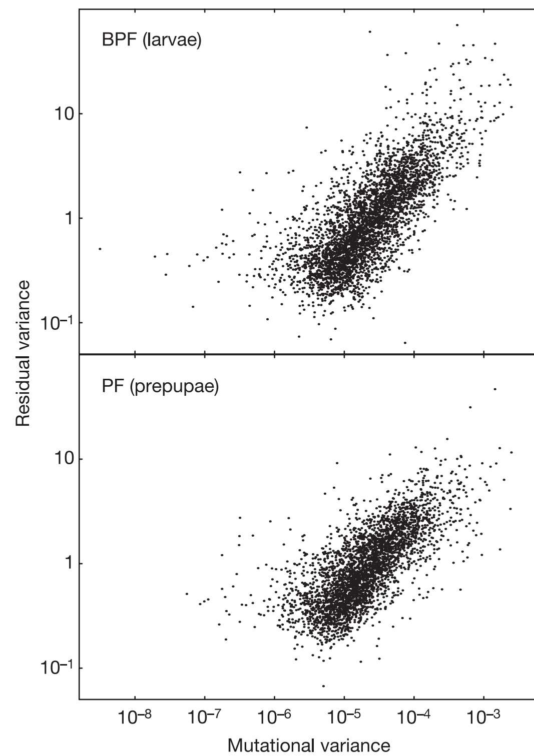

Figure 3 | Opportunity for canalization. Shown are log–log plots of V_m versus V_r. Top, BPF; $r_s = 0.75$. Bottom, PF; $r_s = 0.74$.

D. melanogaster, *Drosophila simulans* and *Drosophila yakuba* on the basis of the divergence times of 2.3 (*D. melanogaster–D. simulans*) and 5.1 (*D. melanogaster/simulans–D. yakuba*) million years with 10 generations per year[16,18]. The observed differences between these species[4,5] (Supplementary Methods) were less than expected (Methods) for almost all genes for which we could measure significant V_m, suggesting that stabilizing selection has a larger role than drift in shaping the evolution of gene expression.

Genes with higher V_m tended, however, to have larger between-species differences (Fig. 2), and patterns of variation across functional groups were also similar between the mutation accumulation and interspecific data. Genes encoding transcriptional regulators had significantly low variability in both stages (multiple test, corrected $P < 10^{-4}$), and the variability of enzymes and structural molecules was high (corrected $P = 0.12$ and 0.058, respectively, at PF). Between species, genes encoding transcriptional regulators and signal transducers were far less variable in their differential expression than those encoding enzymes and structural proteins[4,16]. Despite the overall pattern of stabilizing selection, greater mutational input could drive interspecific variation to be higher for some genes than for others. Alternatively, for genes for which changes in expression are deleterious, stabilizing selection may reduce the phenotypic effects of perturbations, canalizing the expression and thereby restricting the potential for gene expression evolution.

Stabilizing selection is most effective at driving canalization when phenotypic variation is environmentally induced[3]. At both developmental stages, the correlations between V_r and V_m for gene expression are high (Spearman's rank correlation coefficient, r_s: 0.75 at BPF, 0.74 at PF; Fig. 3). If environmental and genetic perturbations affect the mechanisms of gene expression in similar ways[19], for example by affecting the reaction kinetics between transcription factors and binding sites, then organisms probably have the same genetic basis for coping with them[20]. Although a high correlation between V_m and V_r does not necessarily lead to canalization[12,21,22], the evidence for stabilizing selection makes it likely that canalization of gene expression would evolve.

Overall, expression levels and h^2_m are slightly positively correlated ($r_s = 0.11$; Supplementary Fig. 2). When developmental information is taken into account, however, h^2_m is inversely related to levels of expression. Among the subset of genes with different h^2_m in the two stages, those with increasing expression (higher at PF than at BPF) had higher h^2_m at the earlier stage (median $h^2_m = 6.2 \times 10^{-5}$) than at the later stage (median $h^2_m = 2.0 \times 10^{-5}$). Conversely, those with decreasing expression had lower h^2_m at the earlier stage (median $h^2_m = 1.4 \times 10^{-5}$) than at the later stage (median $h^2_m = 4.9 \times 10^{-5}$) ($r_s = -0.55$; Fig. 4). As in interspecific comparisons[4], as genes are upregulated, their expression becomes more stable and less susceptible to mutations. As genes are downregulated or their transcripts degrade (indicating that large amounts of the proteins that they encode are no longer important for cellular function), variability of expression increases.

At least three-fifths of the genes expressed during these developmental stages vary significantly in these lines (FDR = 0.05), which is as many as the estimated number of mutations. Such widespread change may be due to a few mutations with broadly pleiotropic influences. However, the estimated line effects cluster into about 250 distinct patterns of expression (see Supplementary Methods). The sizes of these clusters are highly skewed, suggesting that some of the variability is due to pleiotropic mutations, but the number of distinct patterns indicates that there is the potential for substantial freedom in the directions of evolution.

Although mutational heritabilities of gene expression are an order of magnitude lower than those of many other traits[1,22], between population differences, although substantial[4], are still far lower than expected under neutral models. There are at least three possible explanations for this restriction. First, expression levels are bounded traits, both above and below. Physical limits on the transcription and degradation rates of mRNA could make it impossible to change expression enough to meet the neutral expectation. Over millions of generations, levels of gene expression evolving at the rates that we report here would rebound off these constraints, erasing any correlation between mutational variance and between-population differentiation. Second, stabilizing selection may act directly on the expression of each gene individually. The sheer number of genes varying along with the complexity of gene interactions in pathways and networks makes this unlikely to be a general explanation. Third,

© 2005 Nature Publishing Group

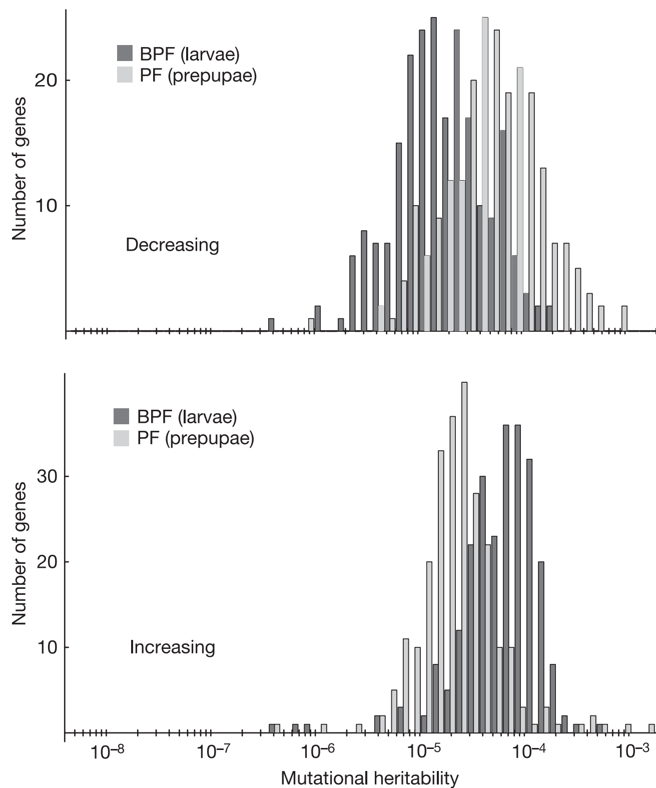

Figure 4 | Developmental context of mutational heritability. Shown are histograms of the h^2_m of genes with significant but unequal h^2_m in both stages (FDR = 0.05). Top, genes with significantly decreasing expression between the two stages; bottom, genes with significantly increasing expression between the two stages (FDR = 0.05).

network output, namely the production of a particular product at a specific place and time, may be the target of selection rather than gene expression itself. As in enzyme flux models[23], the selective effect of any particular change in gene expression may be negligible over a range of values but become substantial when the abundance of mRNA becomes rate limiting or when the variation becomes otherwise functionally relevant. Stabilizing selection, by canalizing network output against perturbations, may facilitate neutrality among members of the underlying network[24]. Gene expression would be able to tolerate a moderate number of mild mutations but would trigger strong selection if the network output were substantially affected. Such a model could also account for moderate correlations between mutational and interspecific variation even when the total level of between-species divergence is far less than expected under neutrality.

In summary, *D. melanogaster* has a broad mutational capacity for changes in gene expression, in both magnitude and genomic extent, that could potentially provide ample raw material for evolutionary diversification. However, although they vary among closely related species, gene expression patterns are relatively stable[4]. In *Caenorhabditis elegans*, genetic variances of gene expression are likewise much less than the neutral expectation[17]. The convergence of this observation in two groups of organisms that diverged in the Precambrian[25] and have different reproductive and life-history strategies indicates that stabilizing selection and structural processes, including canalization, physical and developmental constraints, and correlated responses, govern gene expression evolution.

METHODS

Flies. The 12 mutation accumulation lines were randomly chosen from surviving *D. melanogaster* Ives strain (IVe-39) sub-lines described in ref. 2, after about 200 generations of mutation accumulation (Supplementary Methods). Immediately before sampling, they were rapidly expanded over not more than four generations in uncrowded conditions on standard cornmeal medium containing 0.05% bromophenol blue and sprinkled dry yeast. About 18 h before pupariation, we collected flies with dark full guts (BPF) at the stage when they start to crawl up the side of the bottle to pupariate. For the puparium formation stage, we cleared the bottles of pupae and picked newly pupariated flies in 30-min windows (Supplementary Fig. 1).

Microarray hybridizations. We designed the hybridizations to efficiently estimate stage-specific across-line and within-line variance (Supplementary Methods and Supplementary Fig. 3). Each of the 12 lines was measured eight times at each of two stages. We ground about 30 flies in TRIzol (Invitrogen), extracted total RNA by adding chloroform, extracted mRNA with oligo-dT cellulose (Ambion) and poly-prep columns (BioRad), and prepared labelled sample using the Powerscript fluorescent labelling kit (BD Biosciences) and monofunctional Cy3 and Cy5 dyes (Amersham). We printed and processed whole-genome *D. melanogaster* microarray slides as described[4], and hybridized and washed the samples according to a slightly modified version of protocol M005 from The Institute for Genomic Research (http://www.tigr.org). We scanned the slides with a GenePix 4000 series scanner (Axon) and analysed the images using Spot 2.0 (CSIRO) with modifications for manually flagging bad spots. We carried out a series of global and array-specific normalizations to remove dye-, intensity-, location-, and scale-dependent effects (Supplementary Methods).

Quantitative genetic analyses. After the global and array-specific normalizations, we fit the data for each gene individually to several linear mixed models using Python with calls to PROC MIXED[26–28] in SAS software v.8.2 (SAS Institute; Supplementary Data 2). Roughly 1% of the measurements in the data set were flagged and not used. Because PROC MIXED uses restricted maximum likelihood, missing data were not the problem that they would be in a moments-based estimation.

The gene-specific full model, allowing for stage-specific mean effects, sequence effects when several spots represent the same gene, spot effects, stage-specific across line variances and stage-specific residuals is

$$y_{ijkq} = \mu + \text{Stage}_i + \text{Sequence}_q + \text{Array(Sequence)}_{j(q)} + \text{Line(Stage)}_{k(i)} + \varepsilon_{ijk} \quad (1)$$

$$\text{Array(Sequence)} \sim N(0, \sigma^2_{A(q)}); \text{Line(Stage)} \sim N(0, \sigma^2_{L(i)}); \varepsilon_{(i)} \sim N(0, \sigma^2_{\varepsilon(i)})$$

where y_{ijkq} is the $\log_2$ measurement for a sequence q for a particular gene at stage i from line k on array j, μ denotes the grand mean, and $N(0,\sigma^2)$ means the normal distribution with variance σ^2, respectively subscripted for each component. Specifically, $\sigma^2_{A(q)}$ is the variance for a particular spot across arrays, and $\sigma^2_{L(i)}$ represents the variance across lines at stage i. The overall mean, stage and sequence effects are fixed; the array, line and error effects are random. Preliminary analyses with models including a covariance term could not estimate significant correlation between stages, possibly because of insufficient power. Because each of our samples contained about 30 flies, we multiplied the estimate of $\sigma^2_{\varepsilon(i)}$ from our model by 30, and used this residual variance (V_r) as a surrogate for the environmental variance. If there is appreciable intra-sample covariance within a line, this correction will overestimate the environmental variance. We estimated V_m as 1/400 of the across-line variance $\sigma^2_{L(i)}$ (ref. 18). On the basis of a power analysis, we would detect an across-line variance of 0.02 with a probability of 0.9 given the median within-line variance of 0.023.

We constructed a hierarchy of six alternative models that relaxed the assumptions of stage-specificity of σ^2_L, stage-specificity of σ^2_ε, and the existence of line-specific effects (Line_k). Using a series of likelihood ratio tests, we determined which model best fit the data and used parameter estimates from that model for further analyses (Supplementary Fig. 4 and Supplementary Methods). We used a false discovery rate[29] of 0.05 to account for multiple testing (Supplementary Methods). To reduce bias in our estimates, we subjected the variance estimates to a jack-knife procedure (Supplementary Methods)[30].

Comparisons to interspecific differences used an updated version of published data[4,5] (Supplementary Methods). We tested for divergence from neutrality using a two-tailed $F_{(1,11)}$-test of the ratio of squared divergence between species to the neutral expectation (FDR = 0.05). A Kruskal–Wallis test showed that functional classes of genes did not significantly differ in the ratio of V_m to squared between-species differences.

Received 26 April; accepted 2 August 2005.

1. Lynch, M. & Walsh, B. *Genetics and Analysis of Quantitative Traits* (Sinauer Associates, Sunderland, Massachusetts, 1998).
2. Houle, D. & Nuzhdin, S. V. Mutation accumulation and the effect of *copia* insertions in *Drosophila melanogaster*. *Genet. Res.* **83**, 7–18 (2004).

© 2005 Nature Publishing Group

3. Wagner, G. P., Booth, G. & Bagheri-Chaichian, H. A population genetic theory of canalization. *Evolution* **51**, 329–347 (1997).
4. Rifkin, S. A., Kim, J. & White, K. P. Evolution of gene expression in the *Drosophila melanogaster* subgroup. *Nature Genet.* **33**, 138–144 (2003).
5. Gu, Z. L., Rifkin, S. A., White, K. P. & Li, W. H. Duplicate genes increase gene expression diversity within and between species. *Nature Genet.* **36**, 577–579 (2004).
6. Oleksiak, M. F., Roach, J. L. & Crawford, D. L. Natural variation in cardiac metabolism and gene expression in *Fundulus heteroclitus*. *Nature Genet.* **37**, 67–72 (2004).
7. Morley, M. *et al.* Genetic analysis of genome-wide variation in human gene expression. *Nature* **430**, 743–747 (2004).
8. Yang, H.-P., Tanikawa, A. Y. & Kondrashov, A. S. Molecular nature of 11 spontaneous *de novo* mutations in *Drosophila melanogaster*. *Genetics* **157**, 1285–1292 (2001).
9. White, K. P., Rifkin, S. A., Hurban, P. & Hogness, D. S. Microarray analysis of *Drosophila* development during metamorphosis. *Science* **286**, 2179–2184 (1999).
10. Li, T.-R. & White, K. P. Tissue-specific gene expression and ecdysone-regulated genomic networks in *Drosophila*. *Dev. Cell* **5**, 59–72 (2003).
11. Lande, R. The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genet. Res.* **26**, 221–235 (1976).
12. Houle, D., Morikawa, B. & Lynch, M. Comparing mutational variabilities. *Genetics* **143**, 1467–1483 (1996).
13. Khaitovich, P. *et al.* A neutral model of transcriptome evolution. *PLoS Biol.* **2**, e132 (2004).
14. Yanai, I., Graur, D. & Ophir, R. Incongruent expression profiles between human and mouse orthologous genes suggest widespread neutral evolution of transcription control. *OMICS* **8**, 15–24 (2004).
15. Hsieh, W.-P., Chub, T.-M., Wolfinger, R. D. & Gibson, G. Mixed-model reanalysis of primate data suggests tissue and species biases in oligonucleotide-based gene expression profiles. *Genetics* **165**, 747–757 (2003).
16. Lemos, B., Meiklejohn, C. D., Caceres, M. & Hartl, D. L. Rates of divergence in gene expression profiles of primates, mice, and flies: stabilizing selection and variability among functional categories. *Evolution* **59**, 126–137 (2005).
17. Denver, D. R. *et al.* The transcriptional consequences of mutation and natural selection in *Caenorhabditis elegans*. *Nature Genet.* **37**, 544–548 (2005).
18. Lynch, M. & Hill, W. G. Phenotypic evolution by neutral mutation. *Evolution* **40**, 915–935 (1986).
19. Waddington, C. H. Genetic assimilation of an acquired character. *Evolution* **7**, 118–126 (1953).
20. Meiklejohn, C. D. & Hartl, D. L. A single mode of canalization. *Trends Ecol. Evol.* **17**, 468–473 (2002).
21. Houle, D. How should we explain variation in the genetic variance of traits? *Genetica* **103**, 241–253 (1998).
22. Lynch, M. *et al.* Perspective: spontaneous deleterious mutation. *Evolution* **53**, 645–663 (1999).
23. Kacser, H. & Burns, J. The control of flux. *Symp. Soc. Exp. Biol.* **27**, 65–104 (1973).
24. Hartl, D., Dykhuizen, D. & Dean, A. Limits of adaptation: the evolution of selective neutrality. *Genetics* **111**, 655–674 (1985).
25. Benton, M. J. & Ayala, F. J. Dating the tree of life. *Science* **300**, 1698–1700 (2003).
26. Wolfinger, R. *et al.* Assessing gene significance from cDNA microarray expression data via mixed models. *J. Comput. Biol.* **8**, 625–637 (2001).
27. Littell, R. C., Milliken, G. A., Stroup, W. W. & Wolfinger, R. D. *SAS System for Mixed Models* (SAS Institute, Cary, North Carolina, 1996).
28. Fry, J. D. in *Genetic Analysis of Complex Traits Using SAS* (ed. Saxton, A.) 11–34 (SAS Institute, Cary, North Carolina, 2004).
29. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57**, 289–300 (1995).
30. Sokal, R. R. & Rohlf, F. J. *Biometry* (W. H. Freeman and Company, New York, 1995).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank G. Wagner, S. Rice, J. Powell, M. Lynch, J. Fry, G. Gibson, P. Gibbs, T.-R. Li, J. D. Lambert and members of the Kim and White laboratories for suggestions, help and comments. This work was supported by a National Library of Medicine fellowship (to S.A.R.), an NIH grant (to J.K.), and grants from the W.M. Keck Foundation, the Arnold and Mabel Beckman Foundation, and the NIH and National Human Genome Research Institute (to K.P.W.).

Author Contributions S.A.R. and J.K. planned and designed the project in consultation with D.H. and K.P.W. Expression data was collected by S.A.R. using spotted microarrays developed by K.P.W. D.H. generated and maintained the mutation accumulation lines. S.A.R. and J.K. developed the computational analyses and carried out the quantitative genetics modelling. S.A.R. wrote the paper with contributions from all authors.

Author Information Microarray data have been deposited in the Gene Expression Omnibus under accession numbers GSE2126 (mutation accumulation lines), GSE2641 (technical error) and GSE2642 (comparative data update and extension). Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to J.K. (junhyong@sas.upenn.edu) or K.P.W. (kevin.white@yale.edu).

© 2005 Nature Publishing Group