

Supporting Information

Bedford and Hartl 10.1073/pnas.0812009106

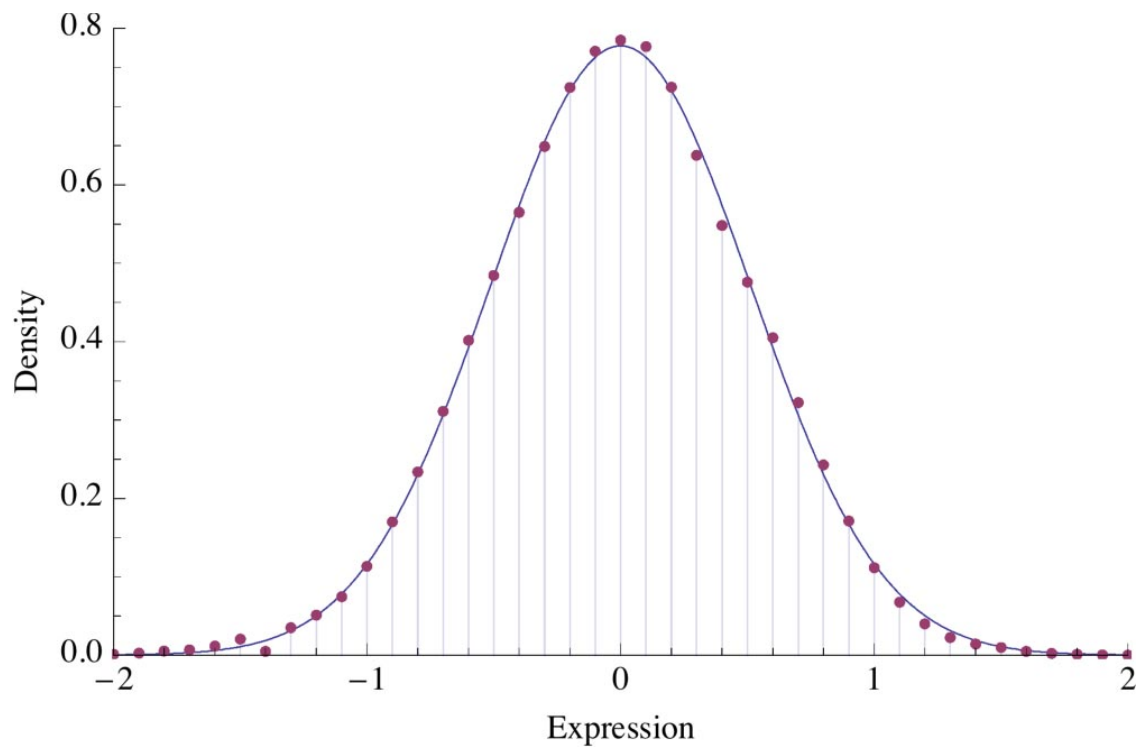


Fig. S1. Comparison of expected and simulated distribution of expression values. The equilibrium distribution of expression values expected from the OU model ($\sigma = 5.481$ and $\lambda = 57.038$) is shown as a solid line. The equilibrium distribution of expression values obtained from simulation using a strong-selection/weak-mutation model and the fitness landscape from Fig. 3 is shown as a set of discrete points. The simulation used 100,000 steps. In each step, a random mutation ($+0.1$ expression or -0.1 expression) was drawn and then checked for fixation based upon its selective coefficient. The equilibrium variance predicted by the OU model is 0.263, whereas the variance observed across the 100,000 simulated expression values is 0.265.

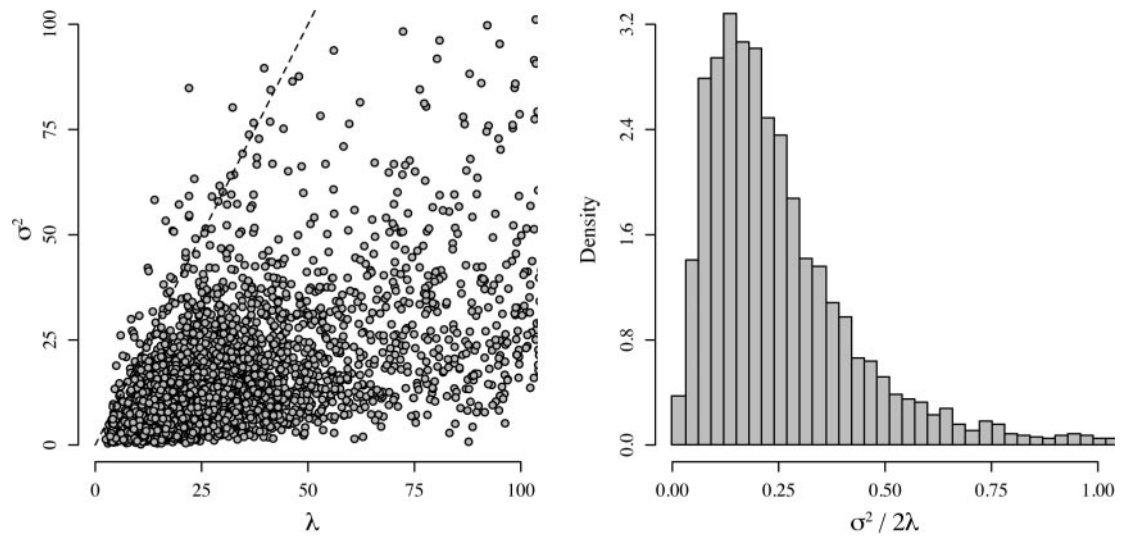


Fig. S2. Distributions of gene-specific OU parameter values. On the left, each point represents the maximum-likelihood estimate for a particular gene. The dashed line represents the neutral expectation of equilibrium variance = 1.0. The right shows the distribution of gene-specific estimates of equilibrium variance.

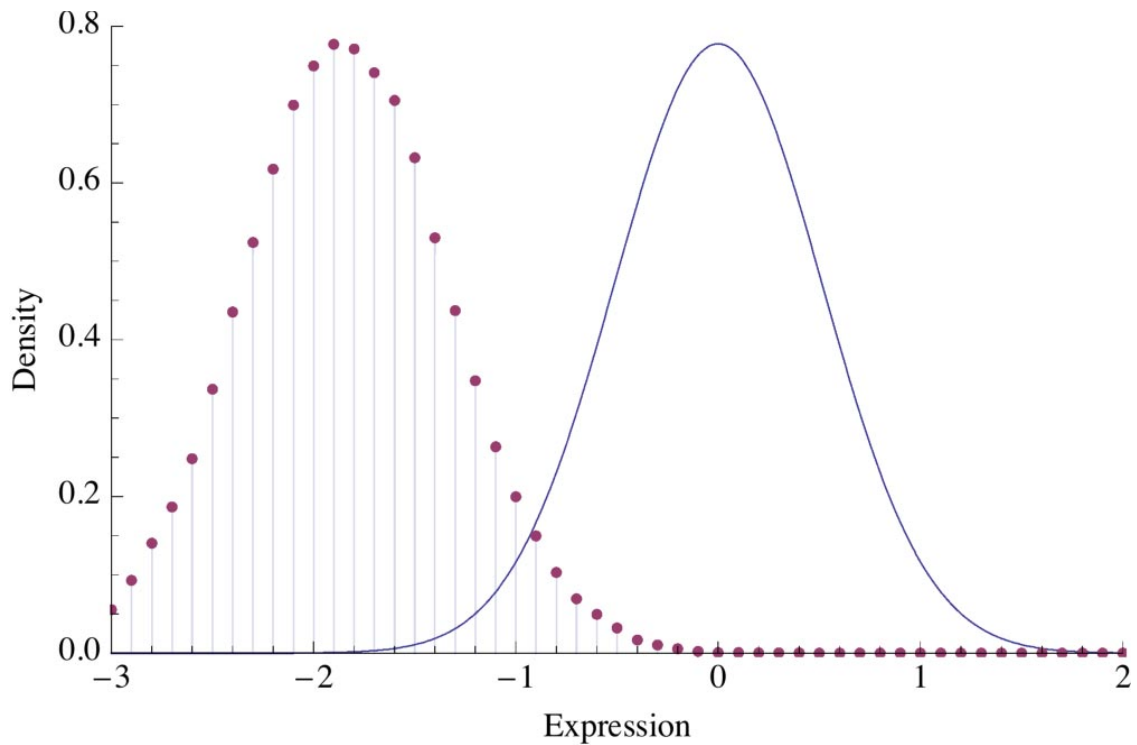


Fig. S3. As in Fig. S1, the solid line represents the OU expectation of equilibrium variance, given symmetrical mutation rates to higher expression and lower expression. However, in the simulation, the chance of a mutation reducing expression level was twice that of a mutation increasing expression level. Interestingly, asymmetrical mutation appears to shift the mean of the distribution, but not affect its variance. The equilibrium variance predicted by the OU model is 0.263, whereas the variance observed across the 100,000 simulated-expression values is 0.261.

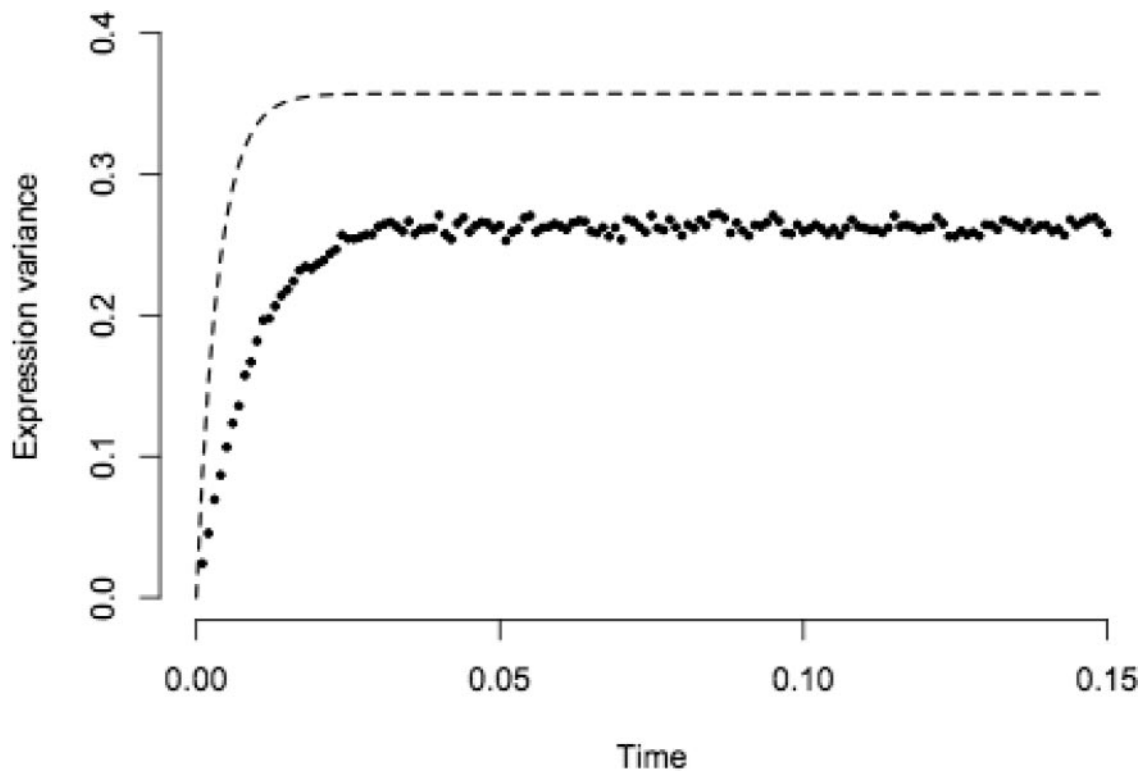


Fig. S4. Normalization reduces gene-expression divergence. The dashed line represents gene-expression variance expected under an OU process, with $\sigma = 10$ and $\lambda = 140$. Each point represents 10,000 realizations of this OU process for a specific time t . In each realization, two independent evolutions (arriving at expression values x_A and x_B) are taken from a common ancestor [$x_0 \approx N(0, 1)$]. Then, the set of x_A values and the set of x_B values are independently normalized to have mean 0 and variance 1. Expression variance is measured for each x_A, x_B pair, and the mean taken across all pairs. In this scenario, nonlinear regression estimates $\sigma = 5.506$ and $\lambda = 57.489$.

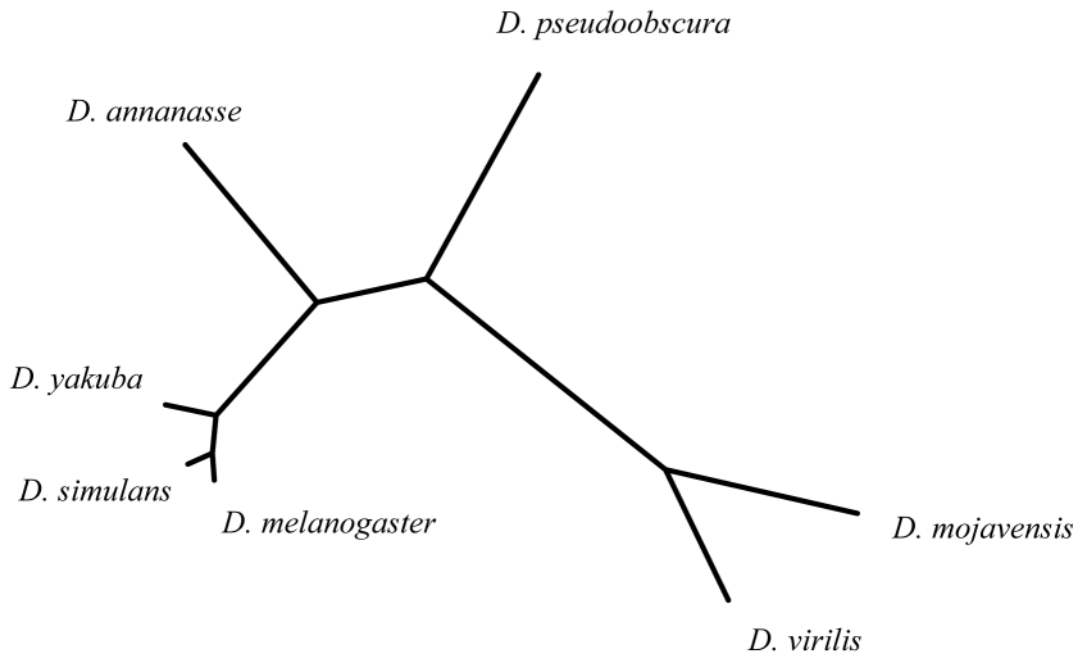


Fig. S5. Species tree for 7 *Drosophila* species. Branch lengths are proportional to amino acid substitutions per site determined by maximum-likelihood averaged across 5,380 genes. The tree shown here is: ((((((dmel: 0.008306, dsim: 0.008237): 0.011678, dyak: 0.015866): 0.046511, dana: 0.063118): 0.034376, dpse: 0.071528): 0.093932, dmoj: 0.060454, dvir: 0.044506).

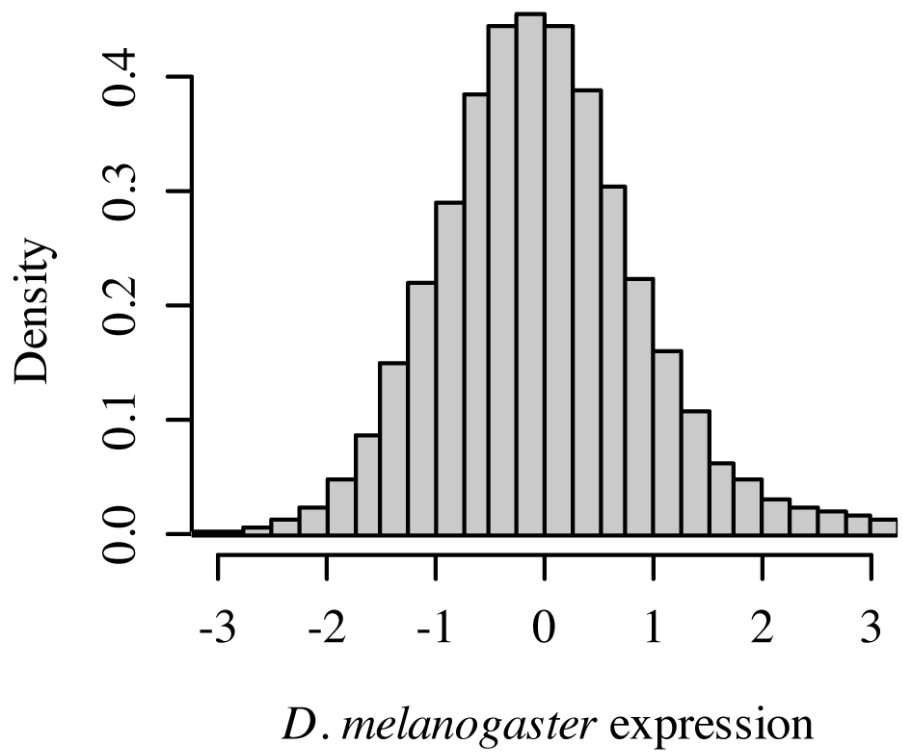


Fig. S6. Distribution of expression level across 6,085 *D. melanogaster* genes. Gene expression was determined by \log_2 probe intensity, and was normalized to have mean 0 and variance 1. Gene expression distributions for other *Drosophila* species are similar.

Other Supporting Information Files

[Dataset S1 \(PDF\)](#)

[SI Appendix](#)