**Supporting Online Materials**

**Material and Methods**

***Flies and hybridizations.*** *D. melanogaster* Canton S and *D. simulans* Sim-1 (Chapel Hill, NC) flies were grown and maintained at 25 °C on glucose-cornmeal-yeast food. Males and virgin females were aged separately for 3-5 days before use. mRNA extractions were carried out with three independent sets of cultures grown and collected at different times in order to incorporate among-culture variation into the hybridizations. Total RNA from flies of both sexes and species was extracted in Trizol Reagent (GibcoBRL). Poly A+ mRNA was purified using the Oligotex Direct mRNA kit (Quiagen) and its quality checked with an Agilent 2100 Bioanalyzer. Hybridization between the oligo dT30 and the poly-A tail of the mRNA was carried out for 45 minutes, and to obtain cDNA, 2 µg of poly A+ mRNA was used as a template for reverse transcription (GibcoBRL). The product was labeled with cyanine-3 (Cy3) or cyanine-5 (Cy5) fluorochromes and purified, and the hybridizations and washes carried out as described (*S1*).

***cDNA microarrays.*** From the Drosophila Gene Collection version 1.0 of expressed sequence tags (*S2*), 5928 cDNA clones were amplified by the polymerase chain reaction and analyzed by gel electrophoresis. These products were purified and mechanically spotted onto polylysine-coated glass slides using an arrayer with a 16-pin head constructed from a design by Patrick O. Brown (http://cmgm.stanford.edu/pbrown/). Spots containing PCR products associated either with multiple bands, faint bands or with an unexpected size band were ignored in subsequent analyses. A total of 4776 clones were considered reliable. Internal quality controls included 177 known DNAs replicated 2-16 times each on the microarray.

***Genomic hybridizations.*** Genomic DNA from each sex of each species was extracted, labeled and purified (*S3*), and digested with *Dpn*II. A sample of 3.75 µg was used in competitive

hybridizations at 62 °C, with a mixture of 50 µg of salmon sperm DNA (GibcoBRL), 100 µg of yeast tRNA (Sigma) and 20 µg poly(dA-dT) (Sigma) as blocking agent. Before hybridization, the DNA was denatured at 98 °C for 2 minutes and incubated at 37 °C for 45 minutes to allow the middle repetitive fraction of the genome to reassociate. The arrays were washed as in the mRNA hybridizations.

***Imaging, data analysis, and results validation.*** Fluorescence intensity was measured with the Axon GenePix 4000A and 4000B scanners and the accompanying GenePix v. 3.0 software. The fluorescence intensity values were adjusted by subtracting local background from foreground in both channels, and the resulting values were normalized for each channel. The analysis included only those spots whose intensity value in at least one of the channels was two standard deviations above the local background. The self-self hybridization performed with cDNA (Fig. S1) showed that only a small fraction of the spots assayed on the array (0.33%) had an associated $\log_2$ (Cy5/Cy3) > 2. Pairwise correlations between replicate hybridizations in Fig. 1 ranged from 0.7 to 0.975 and from 0.846 to 0.874 when cDNA and genomic DNA were used as probes respectively. For expression data, the ratio between the fluorescence intensity of the two channels for each of the spots from each of the hybridizations was then used as input for the Bayesian analysis software (*S4*). This statistical framework assumes a Gaussian distribution of experimental error in the measurement of fluorescence intensity.

Our experimental design enabled small but reproducible differences to be identified as statistically significant. Small differences have the potential of being important for the organism (*S5, S6*), which precludes the use of arbitrary fold-change cut-offs. In our data, 56% of the statistically significant differences have associated means that differ by less than a factor of two, and for the smallest significant difference the means differ by a factor of 1.13. Since correction for multiple tests can be considered too conservative for microarray data (*S7*), we instead

analyzed random permutations of the data in order to determine the rate of false positives. We resampled with replacement the original matrix of normalized ratios for the two channels and constructed a new matrix that was then used as a new input for the Bayesian analysis software (*S4*). We observed only 5 tests that were significant by our criterion of non-overlapping 95% credible intervals out of 19104 comparisons (4776 genes × 4 comparisons per gene). Hence in the present analysis, a criterion of nonoverlapping 95% credible intervals corresponds approximately to a conventional *P* value of 0.00026.

To test the validity of the classification presented in Table 1 and Fig. S2 at a biological level, we compared the present results with previously existing data for *D. melanogaster*. First, *a priori* well known sex-biased genes were examined showing also the expected sex-biased expression including *Acp36DE*, a male accessory gland protein; *qtc*, involved in male courtship; *ocn*, involved in spermatogenesis; and *osk* and *tos*, both female-specific genes involved in oogenesis (*S8*). Second, we contrasted a random sample of 50 male-biased genes and 50 female-biased genes with sex-biased expression in *D. melanogaster* above 2 fold in our experiment with other cDNA microarrays experiments comparing adults with approximately the same age as those used here (*S9*). All the female-biased genes and 48 of the male-biased genes showed the same sex bias in expression in both experiments. Finally, we reanalyzed the data in Ref. *S10* with the same statistical framework used here and compared the magnitude and direction of sex-biased gene expression to that of those genes in common with the present work. Over a 5% of the genes were discordantly classified in relation to their sex-bias in both experiments, being the global correlation of 0.79.

Newly deposited information in FlyBase (*S8*) indicates that a small proportion of cDNAs in the microarrays may be represented more than once. The presence of these potentially duplicated clones does not affect the results reported here.

***Functional analysis.*** Functional classification of differentially expressed genes was carried out using the informatics resources of the Gene Ontology Consortium database (www.geneontology.org) and the Kyoto Encyclopaedia of Genes and Genomes database (www.genome.ad.jp/kegg). Of the genes on the microarray, 2112 were assigned a molecular function, 913 a role in a biological process, 1054 a cellular component, and 434 a regulatory pathway. Overall, 61.5% of genes on the array have information in at least one category and often in more than one category. Significant overrepresentation of particular gene functions or categories among the genes found differentially expressed was determined using *P* values derived from the hypergeometric distribution implemented in the computer program GeneMerge (*S11*). Briefly, the probability *P* of finding a particular number *r* of genes associated with a particular function, process, localization, or pathway, within a set of *k* up-regulated genes is given by the expression (*S12*):

$$P(r \mid N, p, k) = \frac{\binom{pN}{r}\binom{(1-p)N}{k-r}}{\binom{N}{k}}$$

where, then *p* is the fraction of genes on the array that are associated with the particular class under examination, and *N* is the total number of genes on the microarray. Because over-representation for all possible classes within function, process, localization or pathway is assessed, a correction is necessary to account for significantly enriched classes that will invariably occur by chance when multiple tests are performed. Thus, in each analysis, a Bonferroni correction based on the number of classes examined, was applied (*P* value × number of classes examined). For example, 355 (*r*) of the 1441 (*k*) genes significantly up-regulated in female flies are involved in cell growth and/or maintenance, which are associated with GO term GO:0008151 (previous figures, as those hereafter, take into account the fact that several clones

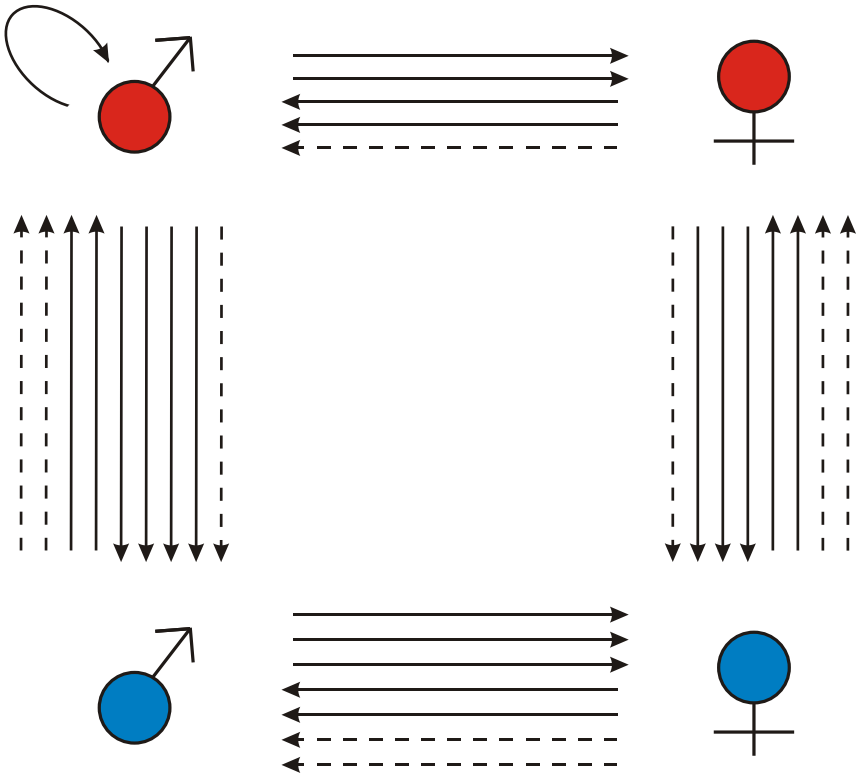
may have the same FlyBase qualifier so that they are considered only once). The number of genes on the array associated with GO:0008151 is 764 out of a total of 4527 (*N*). Accordingly, *p* equals 0.16876. Summing over the tail of the hypergeometric for all less likely cases yields a *P* value of $1.02\times10^{-20}$ for the probability that this many or more cell growth and/or maintenance genes would be present among female up-regulated genes by chance alone. Because 387 different terms were tested, $1.02\times10^{-20} \times 387$ yields a conservative, corrected *P* value of $3.93\times10^{-18}$. In the text, unless indicated, the *P* values reported are those corrected. Detailed outputs for all the analyses performed can be found as Tables S3-S10. In GeneMerge, corrected and uncorrected *P* values correspond to corrected and raw e_scores respectively.

Validation of our functional analysis was done in two ways. First, we inspected the full set of female-biased and male-biased genes in both species (Tables S9-10) for significant associations among genes involved in sex-specific activities, such as egg production in females. In good agreement with expectations and previous results (*S3*), we found that female-biased genes are enriched for cell growth and/or maintenance functions (355 genes found overexpressed / 764 genes present in our array and known to be included in that particular functional class, $P<3.9\times10^{-18}$). These include genes involved in protein biosynthesis, DNA replication and mRNA splicing (52/73, $P < 1.5\times10^{-9}$; 23/26, $P < 1.4\times10^{-5}$; 28/43, $P < 0.003$, respectively). Second, we looked for functional clustering in random samples of 200 genes, the average number of differentially expressed genes between species/sex subsets. No sets contained functional or pathway clustering with a significant, corrected *P* value. Random samples were obtained by resampling with replacement.

**Figures**

**Fig. S1.** Competitive hybridizations performed between different sexes within species and for the same sex between species. Red and blue sex symbols represent *D. melanogaster* and *D. simulans* respectively. Twenty-one hybridizations were carried out using mRNA (solid arrows) and 9 using genomic DNA (dashed arrows) as controls. The arrowhead denotes the sample labeled with Cy5; the opposite sample was labeled with Cy3. One self-self hybridization was carried out for males of *D. melanogaster* as a control. Three hybridizations were performed between females and males of the same species in order to calibrate the technique when using genomic DNA (*S3*).

**Fig. S2.** Classification of the 4776 cDNAs assayed on the relationships between males and females within species (top row) and for the same sex between species (leftmost column); controls are not included. *D. melanogaster* (red), *D. simulans* (blue). The logical operators (=, <, >) denote the relationship among the 95% credible intervals of the mean expression levels obtained by Bayesian methodology (*S4*) for the four sex-by-species combinations present in our experimental design. The number that appears in a particular cell shows the total number of genes on our cDNA microarray that fulfill the four different conditions under which each gene can be classified. The grey boxes denote 14 cells necessarily empty because their requirements are mutually incompatible.



| | | ♂>♀ | | | ♂<♀ | | | ♂=♀ | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | ♂>♀ | ♂<♀ | ♂=♀ | ♂>♀ | ♂<♀ | ♂=♀ | ♂>♀ | ♂<♀ | ♂=♀ |
| | ♀>♀ | 56 | 2 | 14 | 1 | 87 | 23 | 12 | 14 | 64 |
| ♂>♂ | ♀<♀ | 2 | 3 | 0 | | 1 | | | 7 | |
| | ♀=♀ | 213 | 10 | 64 | | 56 | 0 | 1 | 46 | 57 |
| | ♀>♀ | 1 | | | 3 | 7 | 8 | 1 | | |
| ♂<♂ | ♀<♀ | 54 | 0 | 2 | 0 | 53 | 2 | 21 | 6 | 66 |
| | ♀=♀ | 87 | | 0 | 1 | 92 | 48 | 60 | 2 | 108 |
| | ♀>♀ | 13 | | 0 | 0 | 183 | 85 | 17 | 0 | 28 |
| ♂=♂ | ♀<♀ | 15 | 0 | 12 | | 64 | 0 | 0 | 23 | 25 |
| | ♀=♀ | 470 | | 82 | | 964 | 161 | 121 | 99 | 1059 |

♂ ♀ *D. melanogaster* ♂ ♀ *D. simulans*

No difference in level or sex-bias of expression (2493 genes)

Difference in level but not sex-bias of expression (380 genes)

Increase or decrease in sex-biased expression (952 genes)

Gain, loss or reversal of sex-biased expression (951 genes)

**Table S2.** Genes with the same sex pattern of expression in *D. melanogaster* and *D. simulans* according to their magnitude of sex bias in gene expression.

| Factor of difference between sexes [1] | Class | | |
|---|---|---|---|
| | Female-biased | Male-biased | No sex-bias |
| ≤ 2 | 570 (101/564) | 459 (57/451) | - |
| > 2, ≤ 3 | 260 (53/259) | 79 (5/77) | - |
| > 3, ≤ 4 | 56 (13/56) | 9 (3/9) | - |
| > 4 | 96 (17/96) | 169 (12/167) | - |
| Total | 982 (184/975) | 716 (77/704) | 1401 (211/1382) |

[1] Calculated as the ratio between the mean values estimated by the Bayesian methodology used (*S4*). For each class, the fraction of genes on the chromosome X in relation to all those with a known chromosomal location (*S8*) is given in parenthesis. Autosomal genes on the chromosome 4 were excluded. For the numbers above, both species displayed the same factor of difference between sexes. An additional 525 female-biased genes and 195 male-biased genes have, although the same direction of bias, a different factor of sex-bias in gene expression in *D. melanogaster* and *D. simulans*. If the average between species of those genes is calculated and added to those shown above, the resulting figures are, for female-biased genes, and from lower to higher factor of difference: 621 (112/613), 569 (112/564), 153 (37/153), 164 (32/164), and 1507 (293/1494) for the total; and for male-biased genes: 471 (59/463), 178 (8/174), 51 (9/51), 211 (21/208), and 911 (97/896) for the total. The significance and direction of the departure in relation to the expectation do not change when the *G* tests in the text are redone with these latter figures except for genes with a factor difference less or equal 2. In that case a significant excess of male-biased genes in gene expression is detected ($G_{adj}$ = 13.4; d.f. = 1; $P < 2.6\times10^{-4}$). All *G* tests in the text related to this table were adjusted by the continuity correction (*S12*).

**References**


S1. M. B. Eisen, P. O. Brown, *Methods Enzymol.* **303**, 179 (1999).

S2. G. M. Rubin *et al.*, *Science* **287**, 2222 (2000).

S3. J. R. Pollack *et al.*, *Nat. Genet.* **23**, 41 (1999).

S4. J. P. Townsend, D. L. Hartl, *Genome Biol.* **3**, research0071 (2002).

S5. G. Gibson, *Mol. Ecol*. **11**, 17 (2002).

S6. H. Yan, et al., *Nat. Gene*t. **30**, 25 (2002).

S7. W. Jin et al., *Nat. Genet.* **29**, 389 (2001).

S8. The FlyBase Consortium. *Nucleic Acids Res.* **31**, 172 (2003).

S9. M. N. Arbeitman *et al.*, *Science* **297**, 2270 (2002).

S10. M. Parisi *et al*., *Science* **299**, 697 (2003).

S11. C. I. Castillo-Davis, D. L. Hartl, *Bioinformatics* **19**, 891 (2003).

S12. R. R. Sokal, F. J. Rohlf, *Biometry: The Principles and Practice of Statistics in Biological Research* (Freeman, New York, ed. 3, 1995).