Supporting Information

Sackton et al. 10.1073/pnas.1114690108

SI Materials and Methods

Experimental Methods for Measuring Organismal Phenotypes in Y Introgression Lines. ***Lifetime male fecundity.*** To measure lifetime male fecundity, we crossed individual newly enclosed (<4-h old) males (10 replicates per line) from each of the eight Y introgression lines to 3- to 5-d-old virgin w501 *Drosophila simulans* females. Males were transferred without anesthesia to new vials with fresh virgin females at 8, 24, and 48 h, and then every 3 d until the males were 11-d old, and then every 5 d until the males were 31-d old, after which they were allowed to mate for 5 d and discarded. Total offspring were counted in each vial, and lifetime fecundity per male determined as the sum of all offspring produced across all vials.

Relative male viability. We used sex ratio as a proxy for relative male viability, which we measured by crossing males with virgin females from either 092 or w501 strains and counting the proportion of male offspring in 10 replicates of crosses with each Y introgression line male.

Offensive and defensive sperm competitive ability. Offensive sperm competition is the proportion of offspring sired by the second male to mate (P2). Defensive sperm competition is the proportion of offspring sired by the first male to mate (P1). We measured both P1 and P2 in 10 replicate vials for each of two experimental blocks, for a total of 20 datapoints per Y introgression line. In each case, we allowed virgin 092 females to mate to either experimental males (Y introgression lines) or tester w501 males for 12 h, removed the first male, allowed the females to lay eggs for 3 d, and then allowed her to mate to the alternate male. P1 and P2 are estimated as the proportion of ebony individuals among the offspring after mating with both males when the Y introgression males are the first (P1) or second (P2) male to mate.

Time to copulation and copulation duration. To measure time to copulation and copulation duration, we crossed the Y introgression lines to a wild-type *D. simulans* stock (Riverside), and performed measurements on the resulting male progeny. Single virgin males were paired with single virgin Riverside females and continuously observed for 5 h, and start and stop time of copulation were recorded in 20 pairs per Y introgression.

Statistical Analysis of Phenotypic Differences. ***Lifetime male fecundity.*** We summed across all vials for each replicate of each Y introgression line to generate an estimate of total lifetime fecundity for each male tested. We excluded males that either died before producing offspring or never produced offspring (total fecundity = 0); the number of such males does not vary with species (Fisher's Exact Test, P value = 0.78). The number of vials containing fertile males that died during the experiment does not significantly differ between species (Fisher's Exact Test, P value = 0.566); for simplicity of statistical analysis we retain data from those males. All lifetime offspring counts are square-root transformed to improve fit to normality.

We tested for differences in lifetime fecundity between lines carrying a *D. simulans* Y chromosome and lines carrying a *Drosophila sechellia* Y chromosome by comparing linear mixed models using package lme4 in R using a likelihood-ratio test to infer significance. Our alternative model includes a fixed-effect species term and a random-effect term for line within species; our null model contains only a fixed intercept and a random effect term for line within species. Standard linear models with a fixed effect of "line" within each species suggests no significant differences among line in lifetime male fecundity for either *D. simulans* or *D. sechellia*; to generate means and confidence intervals we thus pooled data across lines.

Relative male viability. To test differences in sex ratios among lines carrying a *D. simulans* Y and lines carrying a *D. sechellia* Y, we summed total male and female offspring across replicates and used a Fisher's Exact Test to determine if sex ratios between the two samples were different, separately for each experimental cross.

Sperm competitive ability. To test for differences in sperm competitive ability, we first arcsin-transformed our data to improve fit to normality, and excluded all replicates where we did not have evidence from offspring counts that both males mated to the female in question (if the proportion of ebony among the offspring is 0 or 1). We then compared linear mixed models using package lme4 in R using a likelihood-ratio test to infer significance. Our alternative model included a fixed-effect species term and a random effect term for line nested within experimental block; our null model contained only a fixed intercept and a random effect term for line nested within experimental block.

Time to copulation. The time to the start of copulation was analyzed by fitting a Cox proportional hazards mixed model with species as random effects using R package coxme. One mating pair was excluded because of male death.

Duration of copulation. Duration of copulation was analyzed by comparing a linear mixed model including a species term and random effect for line within species against one with a fixed intercept and a random effect for line within species using lme4. Pairs that did not copulate or did not finish copulation before the end of the 5-h observation period were excluded.

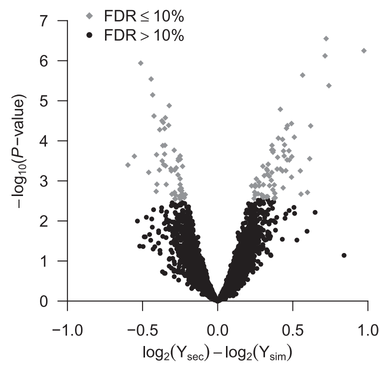

Fig. S1. Volcano plot plotting $-\log_{10}$ (*P* value) against $\log_2$ fold-change for the contrast between Y[sec] and Y[sim]. Genes with positive values of fold-change are more highly expressed (up-regulated) in the presence of the heterospecific Y (Y[sec]) and those with negative values are less highly expressed (down-regulated). Gray diamonds show values significant at a 10% false-discovery rate.

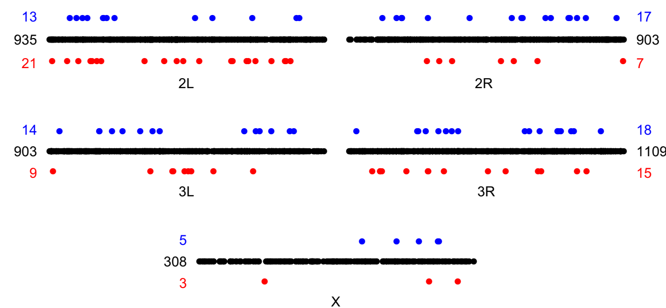

Fig. S2. Distribution of differentially regulated genes across the genome. For each chromosome, the midpoint of each probe is plotted. Significantly up-regulated (Y[sec] vs. Y[sim]) probes are colored blue, significantly down-regulated (Y[sec] vs. Y[sim]) probes are colored red, and nonregulated probes are colored black.

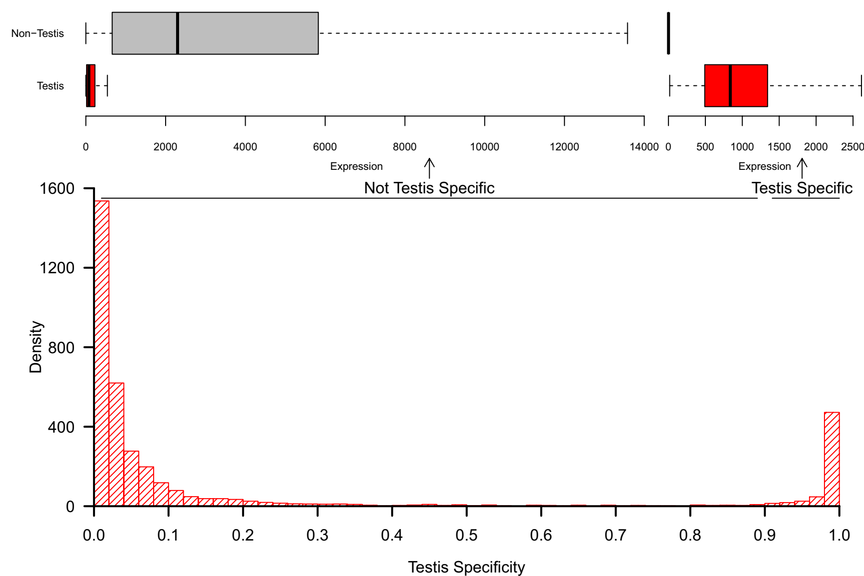

Fig. S3. Our tissue-specificity index accurately discriminates testis-specific genes from other genes. (*Lower*) Histogram of testis specificity for all genes on the array and with FlyAtlas data. We define testis-specific genes as those having a testes-specificity index >0.9. (*Upper*) Boxplots showing comparison of testis expression and total expression in all other tissues for "testis-specific" genes and "nontestis-specific" genes.

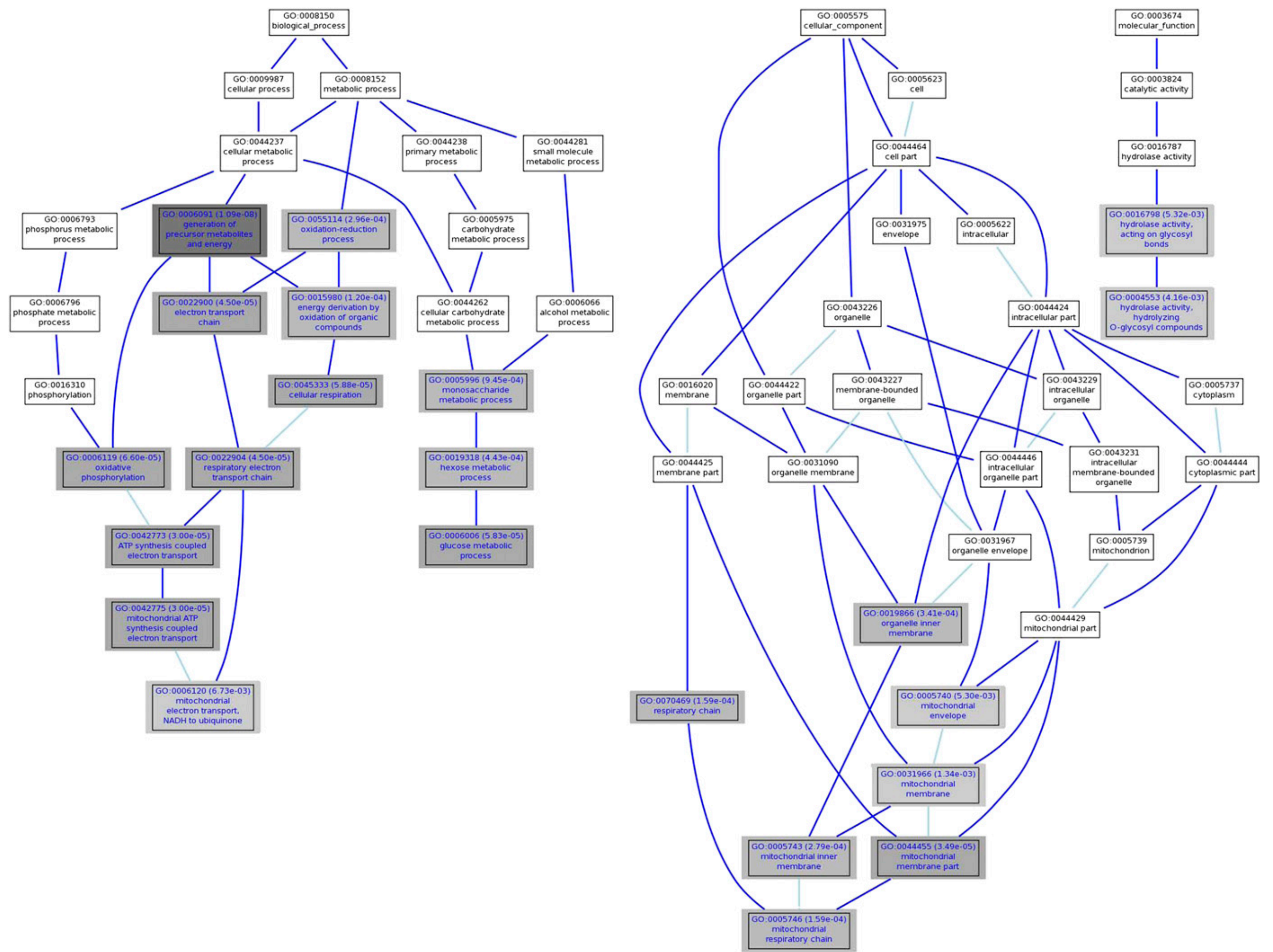

Fig. S4. Gene ontology network diagram showing all nodes significantly overrepresented (using the term-enrichment tool in AmiGO) in the up-regulated class at a Bonferroni-corrected *P* value of 0.05.

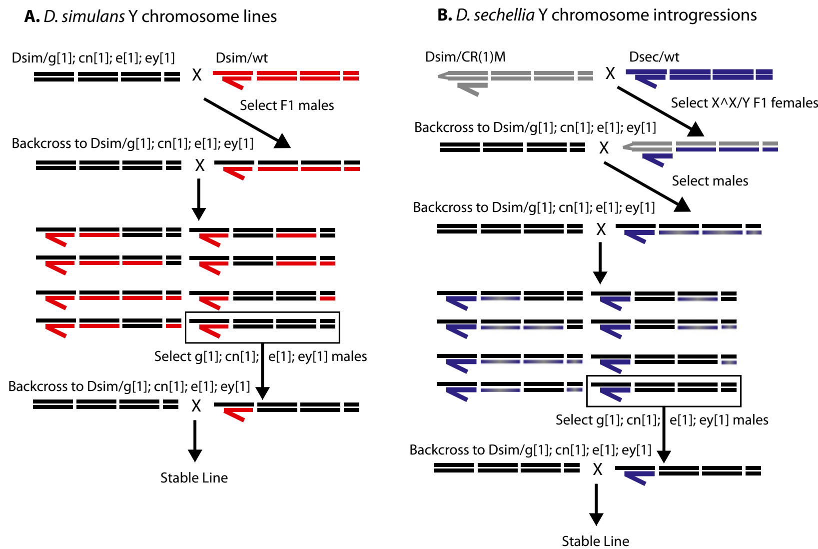

Fig. S5. Crossing scheme used to create Y chromosome extraction lines in *D. simulans* and *D. sechellia* Y chromosome introgressions. (*A*) Y[sim]/Dsim lines were created by crossing males from the a Cameroon population (chromosomes indicated in red) to University of California at San Diego Stock Center (UCSD) line 14021–0251.092 with genotype Dsim/g[1]; cn[1]; e[1]; ey[1] (chromosomes indicated in black). F1 males were backcrossed to virgin parental females and males carrying all four visible mutations used to establish stable lines. (*B*) Y[sec]/Dsim lines were created by first crossing males from *D. sechellia* stocks (from the UCSD Stock Center, indicated in blue) to *D. simulans* C(1)RM attached-X females (gray). The X^X/Y female offspring of this cross are then crossed to Dsim/g[1]; cn[1]; e[1]; ey[1] males, and the male offspring (carrying the *D. sechellia* Y chromosome and the g[1] X chromosome) backcrossed to Dsim/g[1]; cn[1]; e[1]; ey[1] females. Male offspring of this cross carrying all four markers are used to establish stable lines.

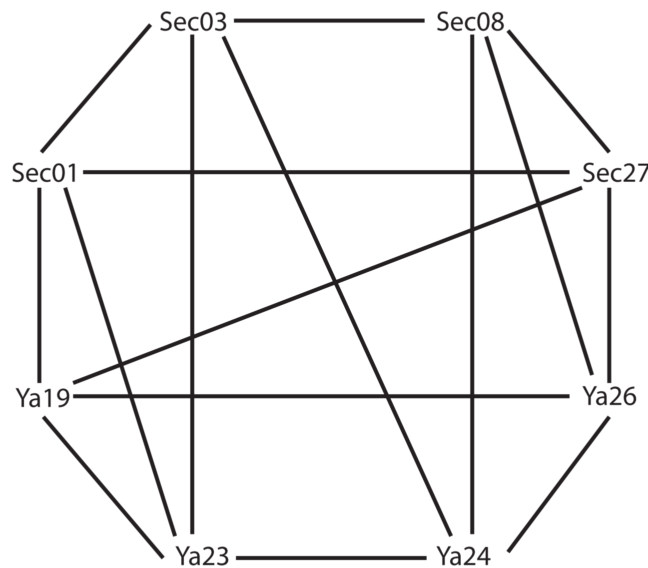

Fig. S6. Array design. Each line represents a pair of technical dye-swap replicates from an individual biological replicate. Samples are labeled as described in *Materials and Methods*.