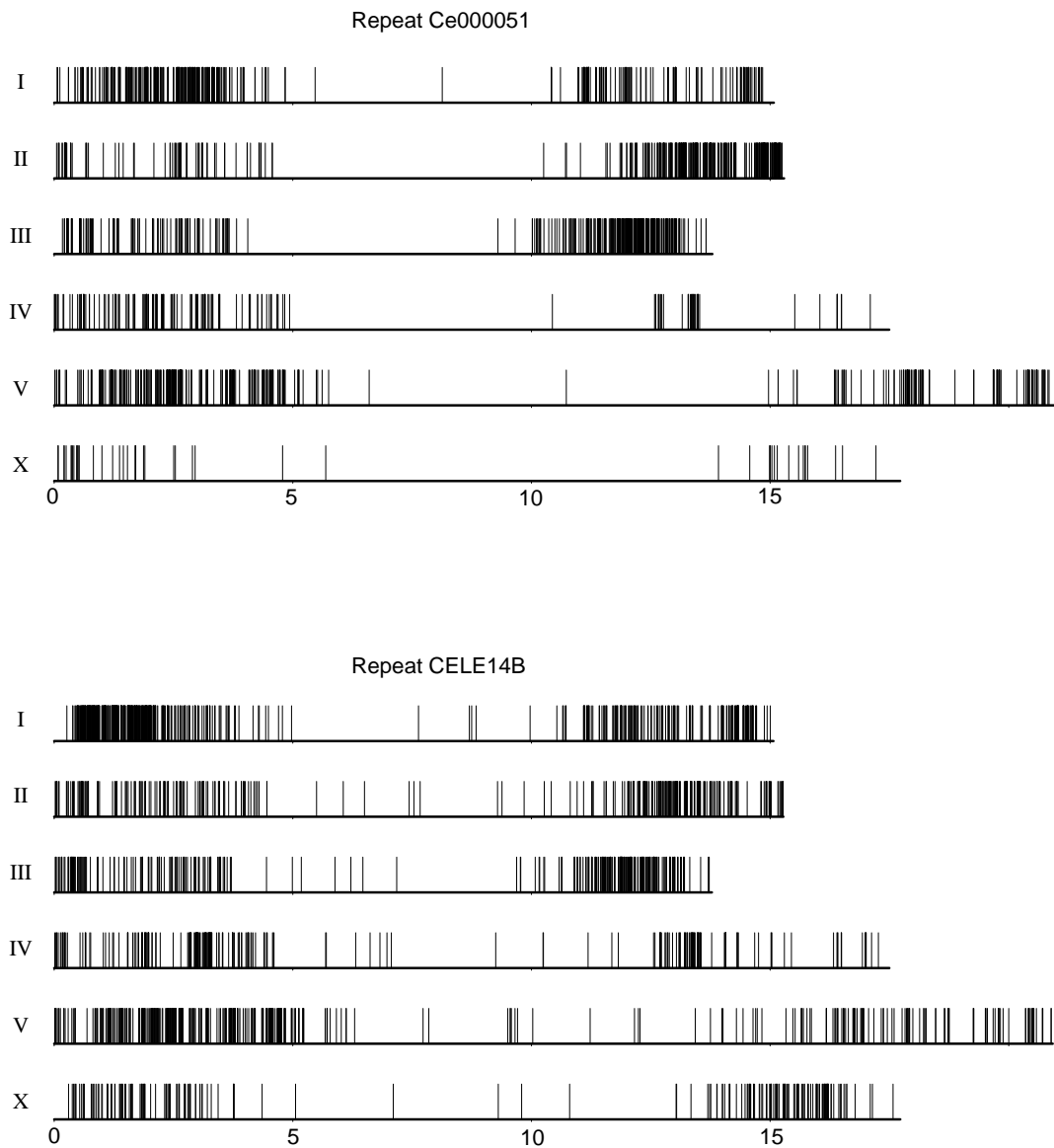
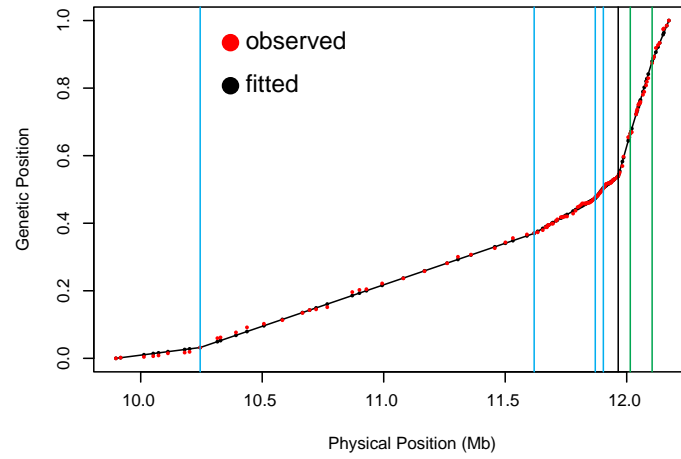


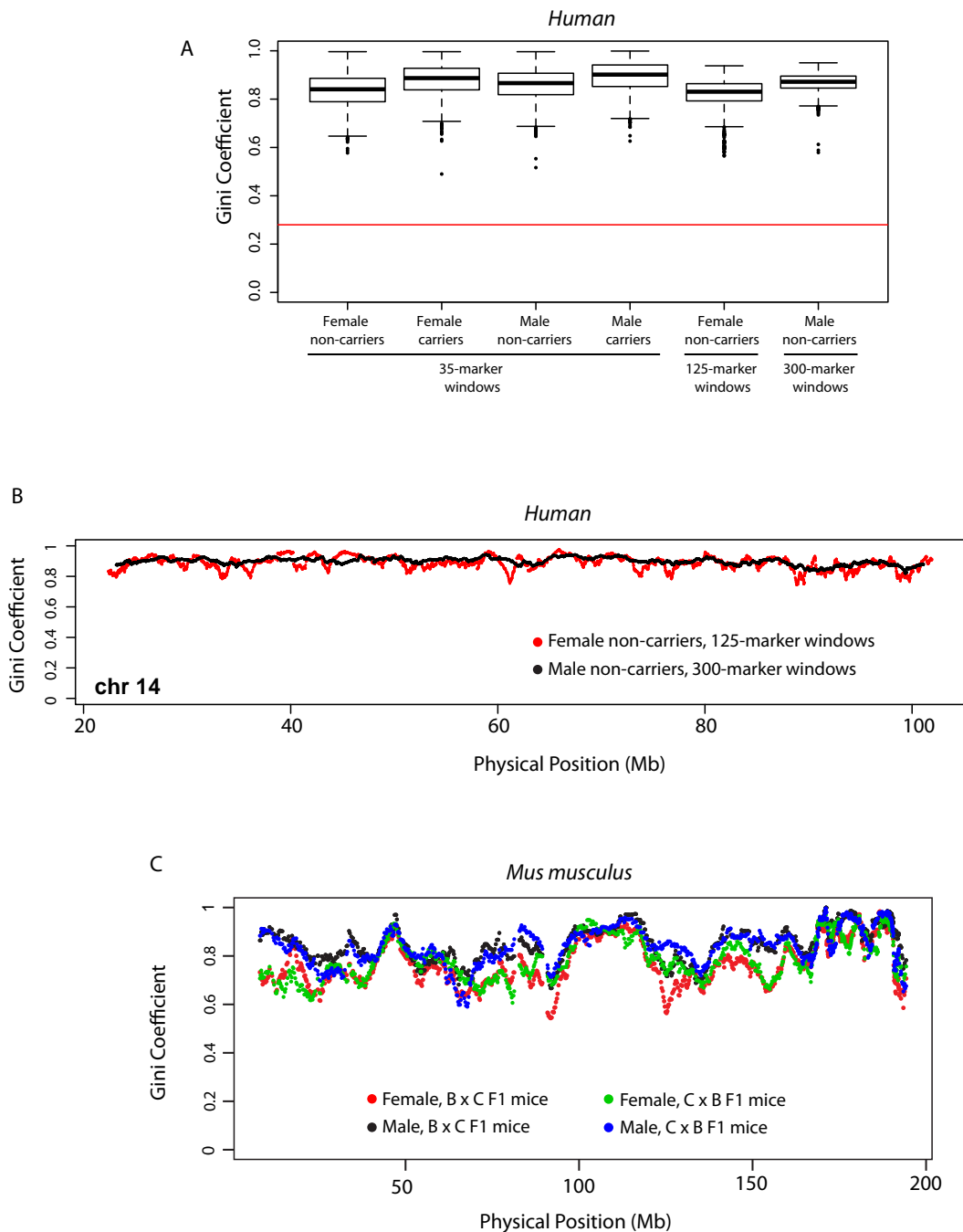
**Figure S1** Genotype analysis pipeline.



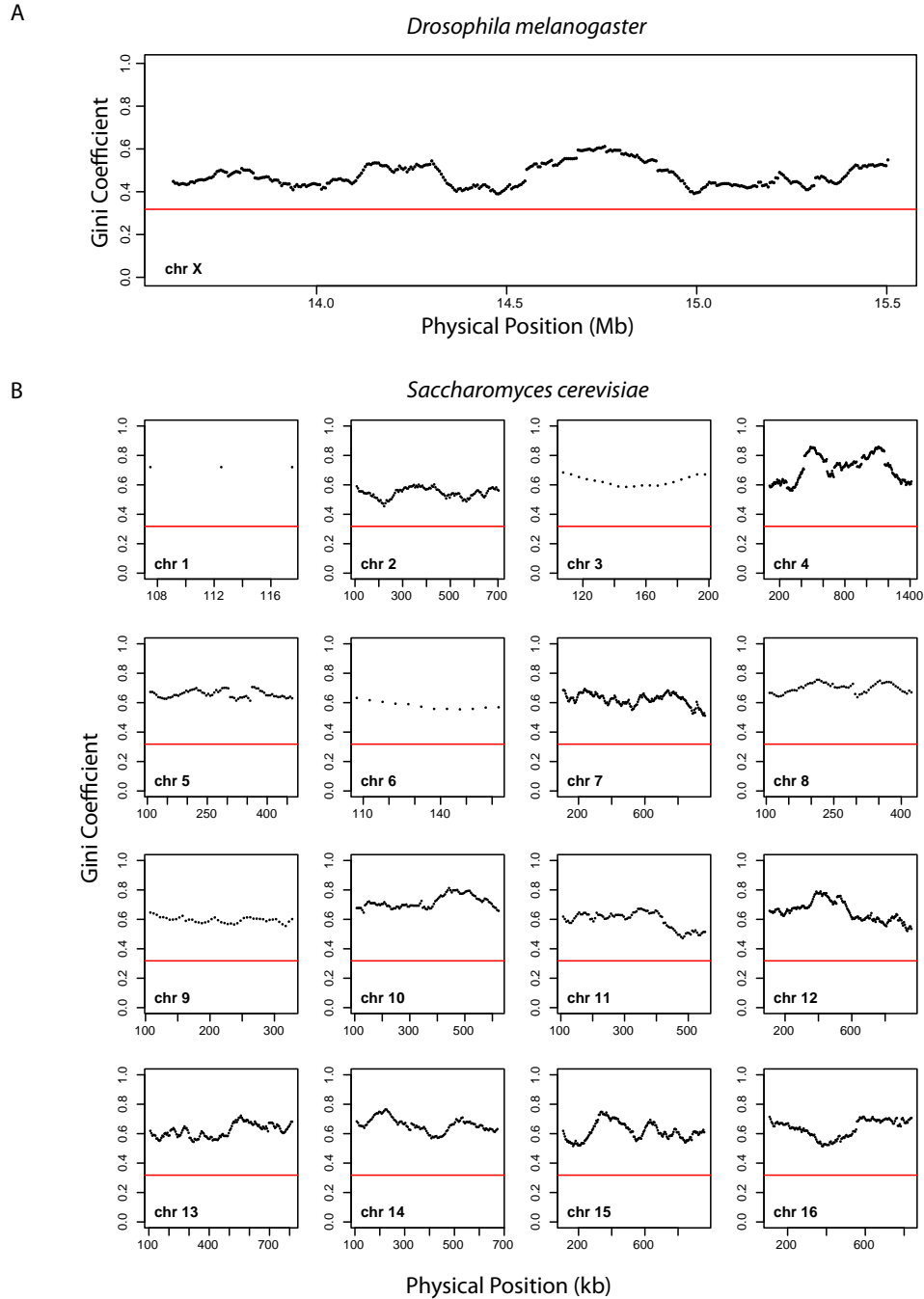
**Figure S2** Two repeat with significantly arm-biased distributions in our studied region are arm-biased across the genome. Each of the 1,901 instances of Repeat Ce000051 and the 2,298 instances of CELE14B are plotted in their physical locations.



**Figure S3** Constant-rate subdomains. The simplest constant-rate model of the observed crossovers requires at least eight subdomains. The vertical bars mark the seven internal boundaries between domains, estimated by constrained piecewise-linear b-spline regression on the Marey Map. Blue bars bound the domains in the center, green bars in the arm.



**Figure S4 Gini coefficients in humans and mice.** Mammalian crossover rates are highly heterogeneous. See Supporting Text 1. **A.** Each box plots the Gini coefficient of 1,000 regions of human genome whose crossover frequencies were determined by KONG *et al.* (2010). “Carriers” are individuals with one or more *PRDM9* 14/15 alleles. The 35-marker intervals are constrained to have physical lengths between 180 and 250 kb. The 125-marker windows are restricted to those with genetic lengths greater than 2.5 cM, and the 300-marker windows are restricted to those with genetic lengths of greater than 3.5 cM. **B.** A sliding window of Gini coefficients along a representative human chromosome. For females, every 125-marker window is plotted, and for males, every 300-marker window. **C.** Gini coefficients for 35-marker windows along mouse chromosome I, using crossover frequencies from PAIGEN *et al.* (2008).



**Figure S5 Gini coefficients in flies and yeast.** Crossover rate heterogeneity in *Drosophila melanogaster* and *Saccharomyces cerevisiae* is higher than in *C. elegans*. See Supporting Text 1. **A.** Gini coefficients for 215-kb windows along the garnet-scaled interval of the X chromosome in *Drosophila melanogaster*, based on SINGH *et al.*'s (2013) estimates for 5-kb intervals. The red line shows the estimated Gini coefficient for our *C. elegans* arm data, after interpolating genetic positions at 5-kb spacing to match the SINGH *et al.* data. **B.** Gini coefficients for 215-kb windows along the *S. cerevisiae* genome. Each point is plotted at the center of the corresponding 215-kb window. The data are interpolated at 5-kb intervals from the observations of MANCERA *et al.* (2008). The red line is as in **A.**

Files S1-S4

Available for download at <http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.113.158857/-/DC1>

**File S1** Breakpoint.data.ws200.txt

**File S2** TestSexDifferences.R

**File S3** MareyMapLRT.R

**File S4** GammaDistributedRateSimulation.R

**File S5**  
**Supporting Text**

**Gini coefficients in model organisms**

We estimated Gini coefficients for multiple species using published datasets. Ideal comparative data are counts of crossovers per marker interval measured at high resolution from individual meioses. Because our interest is the distribution of crossovers, we excluded recombination rate datasets inferred from population genomic data (which may be confounded by selection or gene conversion or biased by distributional assumptions) and from double-strand break data (e.g., RAD51-ChIP-seq or Spo11-oligo-seq), which often depart substantially from crossover distributions because of non-crossover break repair.

We identified appropriate, publically available datasets for human (KONG *et al.* 2010), mouse (PAIGEN *et al.* 2008), *Drosophila melanogaster* (SINGH *et al.* 2013), and *Saccharomyces cerevisiae* (MANCERA *et al.* 2008).

For each dataset, we attempted to extract subsets of the data that match our *C. elegans* data in marker number, physical length, and number of observed crossovers. Generally, we found that these three desiderata could not be satisfied simultaneously, and we therefore evaluated an array of subsampling schemes.

**Human Recombination Rates**

The highest resolution dataset for humans is KONG *et al.* (2010), which includes data from more than 15,000 meioses and 290,000 markers. While our *C. elegans* data are from a single genotype (SP419xCB4856 F1), the KONG *et al.* data are from large numbers of genetically distinct individuals, rendering their recombination rate estimates population averages. KONG *et al.* showed that recombination rates differ between sexes and between *PRDM9* genotype classes. We therefore calculated Gini coefficients for each of four sex-by-genotype classes.

To generate datasets comparable to our worm data, we used two approaches. First, we sampled intervals spanned by 35 markers, the same number covered by the arm region we studied on *C. elegans* chromosome II, with the constraint that the regions must span 180-250 kb in physical length, comparable to our 214 kb arm region. We sampled 1,000 such regions at random for each of the four partitions of the data, and we found that mean and median Gini coefficients are ~0.85 in each case. No sampled region has an estimated Gini coefficient below 0.49 (Figure S3A).

Note that the *C. elegans* arm region has a Gini coefficient of 0.278, but we focused on the arm precisely because the data imply a relatively constant rate there. If we assess all 35-marker intervals for the *C. elegans* data, the Gini coefficient has a maximum of 0.49 where it spans the center-arm boundary (and where the rate heterogeneity is due to the two domains, not to punctate heterogeneity), and it declines to less than 0.3 on either side.

Next, we attempted to match crossover counts between the human and *C. elegans* datasets. Our arm dataset includes observations of 218 crossovers. Because KONG *et al.* report estimated genetic distances between marker pairs, rather than crossover counts, we used their sample sizes and total map lengths to estimate the genetic map length required to encompass 218 crossover observations, as detailed below.

Female noncarriers:  
8,238 meioses \* 4,246 cM map coverage = 349,758 breakpoints.  
218 = 2.65 cM.

Male noncarriers:  
5905 meioses \* 2,290 cM map coverage = 135,203 breakpoints.  
218 breakpoints = 3.69 cM.

Female Carriers:

612 meioses \* 4,257 cM map coverage = 26,057 breakpoints.  
218 breakpoints = 35 cM.

Male carriers:

502 meioses \* 2,259 cM map coverage = 11,340 breakpoints.  
218 breakpoints = 43 cM.

For both male and female carriers of the *PRDM9* 14/15 alleles, the number of observed crossovers is too low to permit useful matched-sampling comparison with our data, and we therefore focus on the non-carrier datasets. In order to sample regions with the desired number of observations, we had to increase the number of markers spanned by each sampled region. For female noncarriers, we sampled 1,000 regions spanned by 125 markers with the constraint that the regions must be 2.5 cM or greater, guaranteeing that the number of crossover observations is comparable to our 218. For male noncarriers, we sampled 1,000 regions spanned by 300 markers with the constraint that regions must be 3.5 cM or greater. In each case, mean and median Gini coefficients are above 0.8, and no region gives an estimate below 0.55 (Figure S3A).

As a check on our random sampling, we also performed a sliding window analysis, calculating the Gini coefficient for every part of the genome. With 125-marker windows scanned across the female noncarrier data, the entire genome exhibits Gini coefficients around 0.85 with modest fluctuations (Figure S3B). The single exception is a region from 6-8 Mb on Chr 8, where the Gini coefficient is around 0.35. This pattern is an artifact due to a single marker interval of more than a megabase, which dominates coefficient estimates for all windows that contain it. For male noncarrier data, scanned with 300-marker data, the global pattern is very similar, and the minimum Gini coefficient across the genome – the artifactual region on chromosome 8 – is 0.55. Overall, the male data exhibits less variability in Gini coefficient because of the larger window size.

### Mouse Recombination Data

PAIGEN et al. (2008) generated 5,472 crossovers on mouse chromosome I and localized them by genotyping 1,058 markers. They counted crossovers in F1 meioses of each sex from crosses of C57BL/6J x CAST/EiJ. The F1s were generated by reciprocal crosses to test for parent-of-origin effects. We confirmed the presence of subtle local parent-of-origin effects using our likelihood-ratio test with 1,000 permutations (females,  $p = 0.009$ ; males,  $p = 0.019$ ). Using the Kolmogorov-Smirnov test, we found no significant global effects (females,  $p = 0.058$ ; males,  $p = 0.340$ ), in keeping with the conclusion of PAIGEN et al. (2008). Differences between males and females were highly significant by both tests for both cross directions ( $p < 0.001$  in every case).

Whether we consider the four sex-by-cross datasets individually, combine them by sex, combine them by cross, or combine all the data, the Gini coefficients for 35-marker intervals are fairly similar and fluctuate around 0.8 (Figure S3C). No interval has a Gini coefficient below 0.5.

For most of chromosome I, PAIGEN et al. (2008) genotyped at low density compared to our *C. elegans* arm region. For the interval from 168.8-193.5 Mb, however, they generated high-resolution data, with median marker spacing of 48 kb. With 1,420 crossovers and 332 markers distributed across this region, segments of 51 markers are roughly comparable in observed crossover number to our *C. elegans* arm data. We calculated Gini coefficients for all 51-marker intervals across this region, and we found that the coefficient was never less than 0.68. However, these regions are all more than 2 Mb. For smaller regions, the number of observed breakpoints is lower and the variance in Gini coefficient becomes large. Even for 8-marker regions, however, the median Gini coefficients is 0.77, and coefficients below 0.3 are not observed.



### ***Drosophila* Recomb Data**

There are few *Drosophila* datasets with dense marker sampling in the manner of our study. CIRULLI *et al.* (2007) and MCGAUGH *et al.* (2012), employing very similar designs but with sparser marker density, identified substantial recombination rate heterogeneity in *D. pseudoobscura* and *D. miranda*. COMERON *et al.* (2012), using advanced intercross lines of *D. melanogaster*, demonstrated substantial rate heterogeneity and strong genotype-dependence, with different wild isolates employing different hotspots. Here, we calculate Gini coefficients from the data of SINGH *et al.* (2013). This study collected recombinant flies using visible markers in a controlled cross of two strains, in the manner of our study. Genotyping, however, was performed on pools of recombinant flies, using pooled-sample next-generation sequencing following a hybridization-based enrichment step. Recombination frequencies between 451 SNP markers were inferred from allele frequencies in sequencing pools. The genotyping design prevents localization of individual crossovers and requires averaging across adjacent SNPs to account for noisy allele-frequency estimates. We estimated Gini coefficients for every 215-kb segment along the 2.09 Mb study region, using SINGH *et al.*'s estimates of rates for 5kb intervals (Figure S4A). The 215kb segments of *Drosophila* data yield Gini coefficients between 0.390 and 0.612 (median = 0.465). For comparison, we used linear interpolation to approximate breakpoint counts at 5kb intervals for our *C. elegans* data. The *C. elegans* arm, with 5kb interpolation, has a Gini coefficient of 0.318.

### **Yeast Recombination Data**

MANCERA *et al.* (2008) generated crossover counts with high resolution from 56 meioses in a *S. cerevisiae* cross. Because the marker number is greater than the crossover number, rates in individual intervals are not well estimated. We therefore interpolated counts at 5kb intervals and used these to calculate Gini coefficients for every 215-kb region. The median coefficient is 0.640, and the minimum is 0.455 (Figure S4B). As noted above, a 5-kb interpolation of the *C. elegans* arm data yields a Gini coefficient of 0.318.

### **Representative Data in Figure 4**

In Figure 4B of the main text, we plotted the following representative examples:

For humans, we show data for the 1,000 35-marker intervals selected from the female non-carrier dataset of KONG *et al.* with the constraint that their physical length is 180-250 kb.

For mice, we show data for every 35-marker interval spanning the whole PAIGEN *et al.* 2008 dataset from the pooled cross data.

For *D. melanogaster*, we plot all 215-kb intervals from the SINGH *et al.* 2013 5kb dataset.

For *S. cerevisiae*, we plotted Gini coefficients estimated from every 215-kb interval from the 5kb interpolation over the entire genome from the MANCERA *et al.* 2008 dataset.

LITERATURE CITED

- CIRULLI, E. T., R. M. KLIMAN and M. A. NOOR, 2007 Fine-scale crossover rate heterogeneity in *Drosophila pseudoobscura*. *J Mol Evol* **64**: 129-135.
- COMERON, J. M., R. RATNAPPAN and S. BAILIN, 2012 The many landscapes of recombination in *Drosophila melanogaster*. *PLoS Genet* **8**: e1002905.
- KONG, A., G. THORLEIFSSON, D. F. GUDBJARTSSON, G. MASSON, A. SIGURDSSON *et al.*, 2010 Fine-scale recombination rate differences between sexes, populations and individuals. *Nature* **467**: 1099-1103.
- MANCERA, E., R. BOURGON, A. BROZZI, W. HUBER and L. M. STEINMETZ, 2008 High-resolution mapping of meiotic crossovers and non-crossovers in yeast. *Nature* **454**: 479-485.
- MCGAUGH, S. E., C. S. HEIL, B. MANZANO-WINKLER, L. LOEWE, S. GOLDSTEIN *et al.*, 2012 Recombination modulates how selection affects linked sites in *Drosophila*. *PLoS Biol* **10**: e1001422.
- PAIGEN, K., J. P. SZATKIEWICZ, K. SAWYER, N. LEAHY, E. D. PARVANOV *et al.*, 2008 The recombinational anatomy of a mouse chromosome. *PLoS Genet* **4**: e1000119.
- SINGH, N. D., E. A. STONE, C. F. AQUADRO and A. G. CLARK, 2013 Fine-Scale Heterogeneity in Crossover Rate in the garnet-scalloped Region of the *Drosophila melanogaster* X Chromosome. *Genetics* **194**: 375-387.