Recombining without Hotspots: A Comprehensive Evolutionary Portrait of Recombination in Two Closely Related Species of *Drosophila*

Caiti S. Smukowski Heil[1,2,*], Chris Ellison[3], Matthew Dubin[1], and Mohamed A.F. Noor[1]

[1]Biology Department, Duke University

[2]Genome Sciences Department, University of Washington

[3]Department of Integrative Biology, University of California, Berkeley

*Corresponding author: E-mail: cssh@uw.edu.

Accepted: September 3, 2015

Data deposition: This project has been deposited at Dryad under the accession doi:10.5061/dryad.sh5cs, and at GenBank under the accessions SRP007802 and PRJNA277849.

Abstract

Meiotic recombination rate varies across the genome within and between individuals, populations, and species in virtually all taxa studied. In almost every species, this variation takes the form of discrete recombination hotspots, determined in some mammals by a protein called PRDM9. Hotspots and their determinants have a profound effect on the genomic landscape, and share certain features that extend across the tree of life. *Drosophila*, in contrast, are anomalous in their absence of hotspots, PRDM9, and other species-specific differences in the determination of recombination. To better understand the evolution of meiosis and general patterns of recombination across diverse taxa, we present a truly comprehensive portrait of recombination across time, combining recently published cross-based contemporary recombination estimates from each of two sister species with newly obtained linkage-disequilibrium-based historic estimates of recombination from both of these species. Using *Drosophila pseudoobscura* and *Drosophila miranda* as a model system, we compare recombination rate between species at multiple scales, and we suggest that *Drosophila* replicate the pattern seen in human–chimpanzee in which recombination rate is conserved at broad scales. We also find evidence of a species-wide recombination modifier(s), resulting in both a present and historic genome-wide elevation of recombination rates in *D. miranda*, and identify broad scale effects on recombination from the presence of an inversion. Finally, we reveal an unprecedented view of the distribution of recombination in *D. pseudoobscura*, illustrating patterns of linked selection and where recombination is taking place. Overall, by combining these estimation approaches, we highlight key similarities and differences in recombination between *Drosophila* and other organisms.

Key words: recombination, evolution, *Drosophila*, linkage disequilibrium, hotspots.

Introduction

Homologous meiotic recombination is a crucial mechanistic and influential evolutionary process. The creation of a physical link between homologous chromosomes during meiosis I promotes proper segregation of chromosomes, preventing aneuploidy and cell death. Evolutionarily, recombination breaks down linkage between loci, allowing sites to segregate independently, and thereby facilitating adaptation and the purging of deleterious mutations. Indeed, recombination is often touted as the fundamental benefit of sexual reproduction. Nonetheless, measuring recombination elicits many challenges. Researchers have used a myriad of ways to estimate recombination rate, which can broadly be grouped into two categories: Assessing direct, current rates or capturing indirect, historical rates (Coop and Przeworski 2007; Baudat et al. 2013). The most common approach for assessing direct, current rates entails genotyping individuals from known pedigrees or controlled crosses, which provides a coarse estimate of recombination rate in a single meiosis or small number of meioses. However, this method is costly, laborious, and very difficult to achieve fine scale resolution, often involving a trade-off between number of markers and number of

individuals to survey. In contrast, with the rise of population sequencing data over the past decade, a statistical approach has been developed and now widely used that estimates the population recombination parameter, ρ, to computationally infer genome-wide, population-averaged recombination rate (Hudson 2001; McVean et al. 2002, 2004; Slatkin 2008). The approach can achieve very fine scale resolution, unmasking genomic features and leading to the discovery of determinants of recombination rate variation.

The development of recombination maps in a wide range of species through both direct and indirect methods has dramatically enhanced our understanding of where recombination takes place in the genome, what factors are responsible, and the selective pressures governing changes in recombination rate over time. For example, one of the most fundamental patterns observed is the existence of variation in recombination rate across the genome, and between populations and species in every organism examined (Nachman 2002; Dumont and Payseur 2008; Paigen and Petkov 2010; Smukowski and Noor 2011). Moreover, this variation in recombination rate is nonrandom, and dependent upon the physical scale at which it is examined. For instance, in many organisms, the broad scale recombination landscape is shaped by large scale structural features of the chromosomes, whereas the fine scale variation manifests as "hotspots": Discrete genomic regions, typically 1–2 kb in length in which recombination increases order(s) of magnitude above the background rate (Lichten and Goldman 1995; Petes 2001; Kauppi et al. 2004; Coop and Przeworski 2007).

With an increasing number of recombination maps developed for related species, it has become apparent that selective pressures governing recombination rate are also dependent on physical scale. Recombination rates observed at the megabase scale or greater exhibit conservation between closely related species (Ptak et al. 2005; Dumont and Payseur 2008; Paigen et al. 2008; Backstrom et al. 2010; Meznar et al. 2010; Stevison and Noor 2010; Smukowski and Noor 2011; McGaugh and Noor 2012), whereas recombination hotspots show nearly no overlap between closely related species (Ptak et al. 2004; Winckler et al. 2005; Auton et al. 2012). Some of the conservation at broad scales may be due to the necessity in most species for at least one crossover per chromosome to ensure proper segregation of chromosomes during meiosis (Roeder 1997; Pardo-Manuel de Villena and Sapienza 2001; Dumont and Payseur 2008; Fledel-Alon et al. 2009). This mandates a minimum level of recombination, as a crossover rate that is too low is likely to be highly deleterious. In contrast, too much recombination breaks apart advantageous allele combinations, threatens genomic integrity, and may introduce mutations at hotspots (Kauppi et al. 2004; Coop and Przeworski 2007; Rattray et al. 2015). Therefore, on a broad scale, recombination is likely constrained between a lower bound dictated by chromosome segregation, and an upper bound determined by maintaining genomic integrity. At fine scales, the recombination landscape is likely shaped by a phenomenon known as the "hotspot paradox" (Boulton et al. 1997; Jeffreys and Neumann 2002, 2005), which describes the self-destructive nature of the resolution of meiotic crossover products. A method to counteract the loss of hotspots was unknown until the recent discovery of the zinc finger histone-methyltransferase, *Prdm9*, which controls the distribution of recombination in mice and humans (Baudat et al. 2010; Parvanov et al. 2010). Positive selection on *Prdm9*'s zinc fingers has created new DNA binding residues, thereby initiating new populations of hotspots (Oliver et al. 2009; Baudat et al. 2010; Berg et al. 2010, 2011; Myers et al. 2010; Brick et al. 2012; Baker et al. 2015), and resulting in rapid turnover of hotspots and fine scale recombination rate divergence between individuals, populations, and species who have *Prdm9* (Ptak et al. 2004, 2005; Winckler et al. 2005; Jeffreys and Neumann 2009; Berg et al. 2010, 2011; Myers et al. 2010; Hinch et al. 2011; Smagulova et al. 2011; Auton et al. 2012; Brick et al. 2012; Baker et al. 2015).

Furthermore, *Prdm9* may also determine the distribution of recombination events across the genome, possibly directing recombination machinery away from promoters and other functional genomic elements (Myers et al. 2005; Smagulova et al. 2011; Brick et al. 2012). In some of the organisms lacking *Prdm9* such as dogs, plants, and yeast, (Oliver et al. 2009; Ponting 2011; Axelsson et al. 2012), hotspots are often found overlapping promoter regions/transcription start sites (TSS), presumably due to the open chromatin structure allowing recombination machinery access to the DNA (Petes 2001; Lichten 2008; Mancera et al. 2008; Pan et al. 2011; Auton et al. 2013; Choi et al. 2013; Hellsten et al. 2013). These observations set up a model in which organisms without *Prdm9* are opportunistic: They form recombination hotspots in "windows of opportunity," most often nucleosome depleted regions around promoters/TSS. On the other hand, in organisms with *Prdm9*, *Prdm9* most likely functions to introduce meiotic-specific H3K4me3 modifications and directs recombination machinery to initiate at these locations away from genomic elements (Hayashi et al. 2005; Grey et al. 2011).

Drosophila stands out from these systems in its unique recombination landscape. Like yeast and plants, *Drosophila* are missing *Prdm9* (Birtle and Ponting 2006; Oliver et al. 2009; Heil and Noor 2012), but unlike these species, *Drosophila* lack highly localized recombination hotspots, the only system known to be missing such hotspots besides *Caenorhabditis elegans* (Hey 2004; Singh et al. 2009, 2013; Comeron et al. 2012; Miller et al. 2012; Manzano-Winkler et al. 2013; Kaur and Rockman 2014). Moreover, *Drosophila* recombination is sex-specific (males do not experience crossing over), and chromosome pairing and synapsis proceed normally in its absence (Hawley et al. 1992; Hawley et al. 2002; McKim et al. 2002). *Drosophila* recombination poses an intriguing system to further explore such

questions as: How is recombination distributed in a species lacking hotspots? Are recombination rates conserved or divergent at varying scales? Are there additional differences between *Drosophila* recombination and recombination in other taxa?

To address these questions, we have developed both empirical and linkage disequilibrium (LD)-based estimates of recombination in each of two closely related species, *Drosophila pseudoobscura* and *Drosophila miranda,* to better understand the recombination landscape and how it changes over evolutionary time. Empirical estimates of recombination rate are available for both species from McGaugh et al. (2012), and here, we present genome-wide LD-based estimates of recombination (through LDhelmet; Chan et al. 2012). We show that LD-based recombination maps generally resemble current recombination rates in these species, and demonstrate moderate conservation between species over 3 Myr divergence. We find that *D. miranda* recombination rates have been and still are higher than those of *D. pseudoobscura*, suggestive of a global recombination modifier or modifiers, and we uncover potential historic effects of an inversion on broad-scale recombination rate. Our fine scale recombination rate estimates illuminate where recombination takes place in the *Drosophila* genome, and confirm the lack of highly localized recombination hotspots in this taxon. In total, this work highlights key similarities and differences of *Drosophila* recombination compared with other organisms, adding a more thorough understanding to the evolution of meiosis and recombination across disparate taxa.

Materials and Methods

Population Sequencing and Variant Calling

We used 11 whole-genome sequences of *D. pseudoobscura* (MV2-25 [reference], Mather32, MSH24, MSH9, TL, PP1134, PP1137, AFC12, FLG14, FLG16, FLG18) previously sequenced and analyzed by McGaugh et al. (2012) and which are available at http://pseudobase.biology.duke.edu (last accessed October 1, 2015) and GenBank (accession numbers: AFC12 SRX091462, Flagstaff14 SRX091308, Flagstaff16 SRX091303, Flagstaff18 SRX091310, Mather32 SRX091461, MatherTL SRX091324, MSH9 SRX091465, MSH24 SRX091463, PP1134 SRX091323, PP1137 SRX091311). Additionally, we used 11 whole-genome sequences of *D. miranda* (Lines: MA28, MAO101.4, MAO3.3, MAO3.4, MAO3.5, MAO3.6, ML14, ML16, ML6f, SP138, and SP235), previously sequenced and provided by Doris Bachtrog's lab (GenBank accession number PRJNA277849). There is little to no population structure present in *D. pseudoobscura*, allowing us to treat flies collected from various regions across the North American west as panmictic (Prakash et al. 1969; Schaeffer and Miller 1992; Noor et al. 2000). All sequences of *D. pseudoobcura* and *D. miranda* were aligned to the *D. pseudoobscura* reference sequence v2.9 using bwa (Li and Durbin 2009). Variants were called for the *D. miranda* genome sequences using the Genome Analysis Tool Kit (GATK) v2.7-2 to remove duplicates and locally realign reads (McKenna et al. 2010; DePristo et al. 2011), and samtools v0.1.18 to call single nucleotide variants (Li et al. 2009). We used custom Perl scripts to convert the variant output files into aligned species-specific individual chromosome FASTA files for input to LDhelmet, see below (2, 4group1, 4group2, 4group3, 4group4, XLgroup1a, XLgroup1e, XLgroup3a, XLgroup3b, XRgroup3a, XRgroup5, XRgroup6, XRgroup8). Note that chromosome 3 was not included in the analysis due to known segregating inversions in *D. pseudoobscura* (Fuller et al. 2014), and chromosome 4group5 was excluded due to poor sequence coverage.

Recombination Estimates

Empirical estimates of recombination for *D. pseudoobscura* and *D. miranda* were obtained from McGaugh et al. (2012) (table 2). Population sequencing-based estimates of recombination were determined using the LDhelmet program (Chan et al. 2012), a statistical approach designed for *Drosophila* which estimates the population recombination parameter ρ using a reversible jump Markov Chain Monte Carlo mechanism (rjMCMC), as in LDhat (McVean et al. 2002, 2004). The program calculates $\rho = 4N_e r$ (where N_e is the effective population size and r is the recombination rate per generation) to estimate the amount of recombination needed in the population to produce the observed levels of LD under a given model. LDhelmet is specifically tailored to address issues in *Drosophila* that may be problematic for programs such as LDhat, like a magnitude fold higher background recombination rate, higher single nucleotide polymorphism (SNP) density, and a large portion of the genome influenced by positive selection (Sella et al. 2009; Sattath et al. 2011). The program was run individually for each chromosome of each species using default parameters, with the exception that we individually estimated theta from each of our data sets ($\theta_{pse} = 0.01$, $\theta_{mir} = 0.0058$) using $\theta = S/a$, where $a = \sum_{i=1}^{n-1} \frac{1}{i}$, and created our own mutation transition matrices. Following the procedure of Chan et al. (2012), we ran the program for 1,000,000 iterations after a burn-in of 100,000 iterations. To assess the choice in block penalty, which controls the smoothing of the maps, we ran LDhelmet with block penalties from 10 to 100 for chromosome 2 of each species as above. The difference between different block penalty choices was negligible (supplementary tables S1 and S2, Supplementary Material online), so we report results using a block penalty of 50.

To identify whether LDhelmet can detect fine scale variation, we simulated data using the program msHOT (Hellenthal and Stephens 2007) under two recombination

landscapes (100 data sets each): A flat recombination rate of $\rho = 10$/kb and a rate of $\rho = 10$/kb with a hotspot 2 kb in width with strength ten times the background rate. We postprocessed the outputs using Seq-Gen (Rambaut and Grassly 1997) and the mutation transition matrix for *D. pseudoobscura*. To reconstruct the recombination maps of simulated data, we ran LDhelmet using a block penalty of 50 for 250,000 iterations after 50,000 iterations of burn-in (supplementary fig. S13, Supplementary Material online).

To compare empirical recombination estimates to LD-based recombination estimates, the recombination estimate from LDhelmet was corrected for distance and averaged over a given interval (intervals from McGaugh et al. 2012, approximately ~200 kb). The empirical chromosomal recombination average (cM/Mb) was divided by the total average LD-based recombination rate for a chromosome to get a single conversion factor for each chromosome. Each interval's average LD-based estimate was multiplied by the chromosomal conversion factor to get an approximation of recombination rates in the units centiMorgan per Megabase (cM/Mb) per Chan et al, and comparable to the methodology followed for chimpanzee recombination conversion (Auton et al. 2012; Chan et al. 2012). We additionally assessed the correlation between empirical and LD-based recombination rates using a Kendall ranked test of unconverted rho/bp and cM/Mb (supplementary table S3, Supplementary Material online).

To compare LD-based recombination estimates between species at different physical scales, we took the output of LDhelmet (which gives an output of ρ/bp between SNP pairs) and took the average ρ/bp over 50-, 100-, 250-, and 500-kb intervals. We plotted the regression between *D. pseudoobscura* and *D. miranda* at each physical scale and reported the Spearman rank test unless otherwise noted (Pearson's correlation coefficient was quite similar). To test the equality of variance between recombination rates averaged at different scales, we used a Levene test with the R packages "car" (John Fox, author) and "reshape2" (Wickham 2007) (supplementary table S4, Supplementary Material online).

Analyses of Genomic Correlates

Transcription Start Sites

Locations of *D. pseudoobscura* TSS were obtained from Main et al. (2013). The recombination estimates from LDhelmet were averaged over 5,000-bp intervals for the 200-kb regions flanking either side of a TSS. Genome-wide data were aggregated to create figure 3. We confirmed that the effect seen around TSS is not due to a lower SNP density caused by conservation of promoter regions, by comparing SNP density in 20-kb regions flanking TSS to 20-kb regions where no TSS are present.

Proportion Recombination

A Perl script was used to sort recombination rates and distances to build a list of the proportion of recombination occurring in the proportion of sequence (Stevison L, unpublished data). These data were plotted using a Lorenz curve from the R package "ineq" (Achim Zeileis, Christian Kleiber). The Gini coefficient was also calculated using this package as in Kaur and Rockman (2014).

Introns and Exons

Locations of exons and introns and relative positions in a gene were extracted from *D. pseudoobscura* v2.9 annotations from FlyBase (St Pierre et al. 2014). The recombination estimate from LDhelmet was corrected for distance and averaged over the given interval, then aggregated to give genome-wide totals for each exon and intron position within a gene.

Nucleotide Diversity

We calculated pairwise nucleotide diversity (π) at 4-fold degenerate sites using custom Perl scripts, excluding sites where an insertion or deletion was found in any line.

Data Archiving

All data sets used to create the figures or tables in this study are archived in Dryad, doi:10.5061/dryad.sh5c3.

Results

How Well Do LD-Based Linkage Maps Reflect Present-Day Measures of Recombination?

The use of statistical programs on population sequencing data has drastically accelerated the study of recombination rate variation, particularly at fine scales. However, although LD-based inference of recombination rate is a powerful tool, this method has its drawbacks too: LD is influenced by many other factors apart from recombination, including demography, selection, and genetic drift (Slatkin 2008). Additionally, LD represents a historical population-averaged recombination rate whereas empirical methodologies typically survey only one population for one generation of meiosis. Thus, LD-based maps can potentially deviate from direct, single-generation empirical estimates of recombination for several reasons (Jeffreys et al. 2005; Reed and Tishkoff 2006; Coop et al. 2008; Khil and Camerini-Otero 2010). Therefore, to first clarify how well our LD-based maps replicate empirical recombination rates, we created genome-wide LD-based recombination maps for *D. pseudoobscura* and *D. miranda*. The LD-based recombination maps were made with the program LDhelmet (Chan et al. 2012), similar to the widely used program LDhat (McVean et al. 2004), but explicitly designed for *Drosophila* to account for biological and technical differences in the data sets, such as a much

higher background recombination rate than human. We confirmed that estimates of recombination from LDhelmet correlated strongly with a simple direct measure of LD decay in our data set (supplementary fig. S1, Supplementary Material online). We then compared the LD-based recombination estimates with those from the empirical maps generated by McGaugh et al. (2012). The empirical maps were generated by genotyping SNP markers at 150–200 kb intervals for several of the largest chromosome groups for two populations of *D. pseudoobscura*, and one of *D. miranda*, and several selected regions of 20 and 5 kb of chromosome 2 for one population of *D. pseudoobscura* (Manzano-Winkler et al. 2013).

Recombination rates obtained from these two strategies are significantly correlated at a broad scale (~200 kb, table 1, fig. 1 and supplementary table S3 and figs. S2–S4, Supplementary Material online), although empirical-LD-based correlations are generally stronger in *D. pseudoobscura*. A possible explanation for this difference is that LD-based rates typically do not capture extreme values seen in the empirical landscape (fig. 1 and supplementary figs. S2–S4, Supplementary Material online), which may explain why the correlation is weaker in the elevated recombination map of *D. miranda*. The particular strains used in the *D. miranda* empirical study may also have an atypical recombination distribution relative to the historical population, or that there could be greater among-strain variation across *D. miranda*. *D. miranda* XLgroup1a in particular stands as an outlier (Spearman's rho = 0.353), although the empirical map in this region is broken into three pieces due to failure of markers, potentially introducing more error into these estimates. We also have empirical estimates for a small portion of *D. pseudoobscura* at finer scales of 20 and 5 kb; the correlation between empirical and LD-based maps persists at 20 kb (Spearman's rho = 0.411, Pearson's $r = 0.68$, supplementary fig. S5, Supplementary Material online), but there are too little data at the 5-kb scale to be conclusive (supplementary fig. S6, Supplementary Material online).

Few studies have compared empirical and LD-based estimates, but similar results were obtained from LDhelmet in *Drosophila melanogaster* (Raleigh, NC population; ~200 kb: 2L $r = 0.73$, 2R $r = 0.75$, 3L $r = 0.61$, 3R $r = 0.73$, X $r = 0.71$ [Chan et al. 2012]). Chan et al. (2012) used data from Singh et al. (2009) and FlyBase (McQuilton et al. 2012), at resolutions of approximately 100–200 kb, so we present the first data obtained from LDhelmet demonstrating correspondence in LD-based and empirical rates at scales under 100 kb. One of the only other comparisons is from human, where the correlation between LD-based and empirical recombination rates is nearly perfect at the megabase scale ($r = 0.98$) (Myers et al. 2005). Our results and those of Chan et al. (2012) clearly exhibit a lower correlation than observed in human; however, our comparisons are done at the approximately 200-kb scale where there is more variation in recombination rate (see below, table 3). Furthermore, empirical–empirical correlations between different populations of *D. pseudoobscura* are only moderately higher than the LD-based–empirical correlations (table 2), suggesting that the LD-based maps are capturing much of the recombination variation in these species.

Table 1
Empirical versus LD-Based Comparisons of Recombination Rate

	Interval Size Average, Median (Mb)	Number of Intervals	Average Empirical Recombination Rate (cM/Mb)	Average LD-Based Recombination Rate (cM/Mb)	Correlation (Spearman's rho, *P* Value)
2					
Drosophila pseudoobscura Flagstaff	0.215, 0.142	140	3.55	3.71	0.654, <0.0001
Drosophila pseudoobscura Pikes Peak	0.194, 0.149	158	3.47	3.75	0.601, <0.0001
Drosophila miranda	0.189, 0.147	154	4.85	4.87	0.593, <0.0001
XLgroup1a					
Drosophila pseudoobscura Flagstaff	0.215, 0.148	7	3.76	5.17	—
Drosophila pseudoobscura Pikes Peak	0.211, 0.198	24	3.32	3.47	0.747, 0.0005
Drosophila miranda	0.247, 0.202	18	4.95	5.82	0.353, 0.1806
XRgroup6					
Drosophila pseudoobscura Flagstaff	0.164, 0.143	68	3.94	3.66	0.636, <0.0001
Drosophila pseudoobscura Pikes Peak	0.166, 0.146	71	4.44	4.55	0.679, <0.0001
Drosophila miranda	0.163, 0.149	74	5.40	5.67	0.588, <0.0001
XRgroup8					
Drosophila pseudoobscura Flagstaff	0.269, 0.172	32	4.52	4.03	0.805, <0.0001
Drosophila pseudoobscura Pikes Peak	0.248, 0.209	36	4.72	4.47	0.627, <0.0001
Drosophila miranda	0.200, 0.187	34	6.12	6.06	0.411, 0.0182

Note.—Empirical recombination rates were obtained from McGaugh et al. (2012). LD-based recombination rates were averaged over empirical windows and converted to cM/Mb (see Materials and Methods). A dash indicates not enough data to assess the correlation.

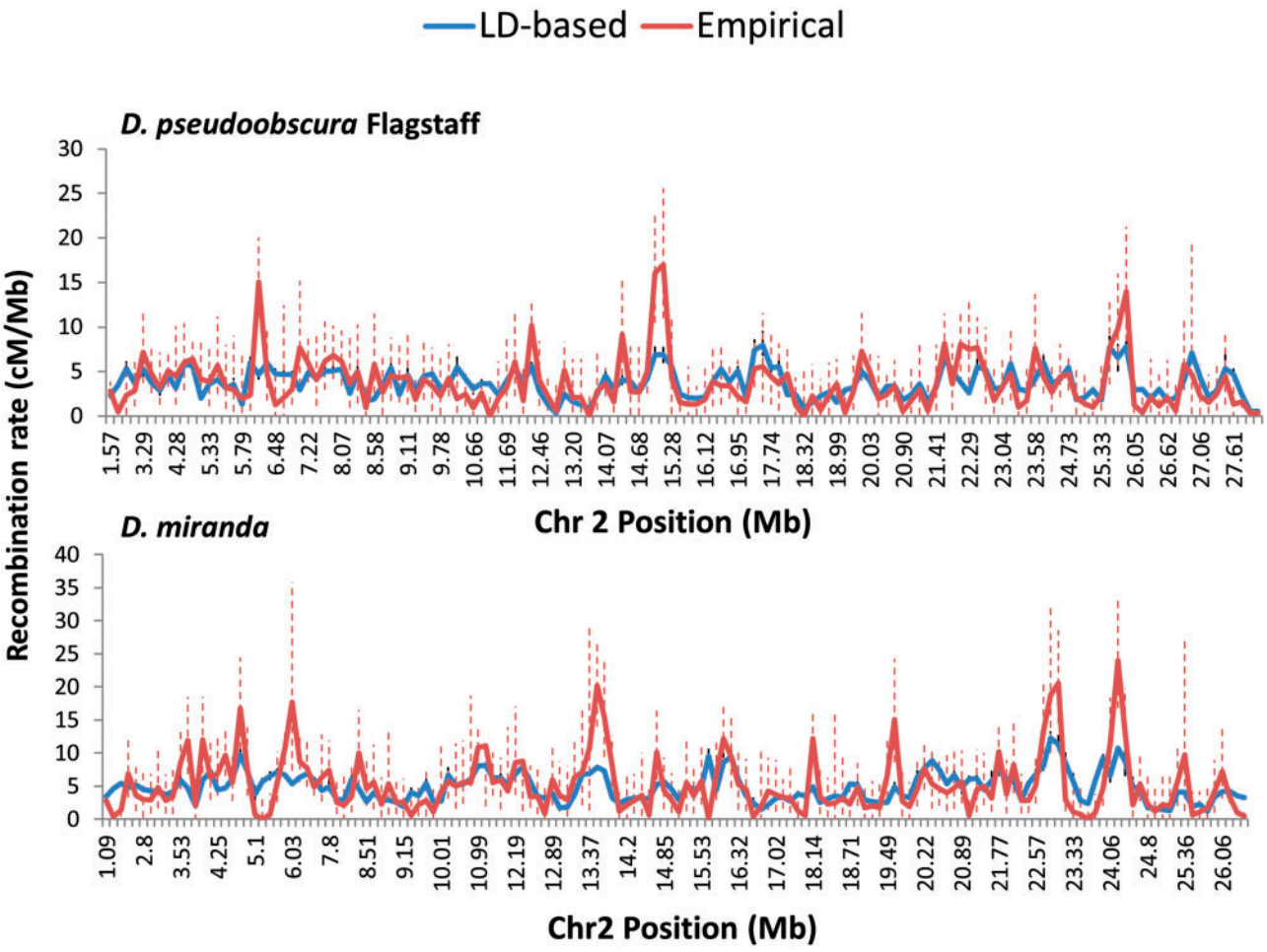

Fig. 1.—A comparison of empirical and LD-based estimations of recombination rate. Empirical recombination rates (red) are plotted with LDhelmet-derived recombination rates (blue) across chromosome 2 for *D. pseudoobscura* Flagstaff (top) and *D. miranda* (bottom). All recombination rates are reported in cM/Mb. Error bars depict 95% confidence intervals for the empirical (red, dashed), and LD-based (black, solid).

Recombination Rates between Species

Broad Scale Recombination Rate Patterns

Present-day empirically assayed recombination rate was previously shown to be conserved at broad scales (~200 kb) between *D. pseudoobscura* and *D. miranda* (table 2) (McGaugh et al. 2012), and with our LD-based recombination maps, we recapitulate and extend these results. We detect and confirm genome-wide broad scale conservation of recombination rates between species (table 3, fig. 2 and supplementary figs. S8 and S9, Supplementary Material online), but with several notable exceptions. First, *D. miranda* LD-based recombination rates remain higher than those of *D. pseudoobscura* (supplementary figs. S8 and S9, Supplementary Material online), supporting the conclusion that the putative global recombination modifier (or modifiers) identified in empirical work persists species-wide in one or both of these species (McGaugh et al. 2012). Estimates of the effect of the modifier(s) are quite similar between empirical and LD-based data, recombination rate is 1.26× higher on chromosome 2 in *D. miranda* (compared with 1.28–1.32 higher from empirical work), and 1.71× higher on chromosome XR (empirical: 1.4–1.47). Moreover, this difference persists despite a difference in effective population sizes: *D. pseudoobscura* N_e is thought to be several times that of *D. miranda* (Bachtrog and Andolfatto 2006; Loewe et al. 2006; Jensen and Bachtrog 2011).

Another pattern apparent in the broad scale recombination landscape is that some chromosome segments exhibit a weaker correlation of recombination rate between species, potentially indicative of regional differences. For example, chromosome arm XL group1e is a clear outlier in the comparison between *D. pseudoobscura* and *D. miranda* (table 3, fig. 2 and supplementary figs. S8 and S9, Supplementary Material online). Interestingly, the species differ by two inversions on the XL chromosome arm: A large 12-Mb inversion that corresponds almost precisely with group1e and a small, approximately 4-Mb inversion that took place following the large

inversion at the distal end of group1e (Anderson et al. 1977; Schaeffer et al. 2008). This is suggestive that an inversion caused historic broad scale changes in recombination rate.

Fine Scale Recombination Rate Patterns

Although recombination rates are generally conserved at broad scales, they are divergent at fine scales in the few organisms in which they have been examined. Analysis of fine scale recombination in *Drosophila* has been relatively limited to isolated genomic regions in *D. melanogaster* (Schweitzer 1935; Singh et al. 2009, 2013; Miller et al. 2012; although see Comeron et al. 2012), and *D. pseudoobscura* (Cirulli et al. 2007; McGaugh et al. 2012; Manzano-Winkler et al. 2013), and has not been compared between species. With LD-based maps, we are able to characterize the genome-wide fine-scale recombination landscape in *D. pseudoobscura* and *D. miranda* for the first time.

To test for the presence of recombination hotspots in *D. pseudoobscura*, we first used simulations to test the power of LDhelmet to detect hotspots. Simulations with a similar recombination rate as obtained in our LD-based recombination estimates with and without a hotspot (strength ten times the background rate) show that LDhelmet can indeed detect the presence of fine scale variation in recombination rate (supplementary fig. S13, Supplementary Material online). We then searched for 500–5,000 bp regions of *D. pseudoobscura* in which the recombination rate exceeded ten times the background recombination rate in a similar manner to Chan et al. (2012). Chan et al. detected 21 total hotspots in two different populations of *D. melanogaster* and were able to independently verify 10. In *D. pseudoobscura*, we discovered 19 hotspots (supplementary fig. S10, Supplementary Material online); however, we did not verify these hotspots independently with alternative means, so their presence should be seen as provisional, and these isolated hotspots remain both far less in total number and in magnitude than those seen in other organisms (e.g., an estimated 60,000 hotspots in humans). This conclusion remains consistent with previous data, confirming that the recombination landscape of *Drosophila* possesses fine scale variation, but does not exhibit recombination hotspots like those found in humans, yeast, and other species. Recombination rates remain uniform over larger physical distances in *D. miranda*, thereby restricting the resolution at which we can examine recombination rates in this species, and the level at which we can compare rates between species. The effective population size of *D. miranda* is predicted to be much smaller than *D. pseudoobscura* (Jensen and Bachtrog 2011), and the number of segregating sites in *D. miranda* is almost half of *D. pseudoobscura* ($\theta_{pse} = 0.01$, $\theta_{mir} = 0.0058$). However, our study still achieves the finest scale genome-wide recombination map of both *D. pseudoobscura* and *D. miranda* to date, and we are still able to observe patterns from the relatively fine scale of 50 kb.

For example, in comparing recombination maps across different scales, variation and amplitude in recombination rate appear to increase at finer scales: For every chromosome assembly but one, the correlation coefficient between the species was weaker in 50-kb windows than in every larger window size (100, 250, or 500 kb; see table 3). A Levene test for equality of variance supports this observation of greater variation at small scales, with recombination rate estimates more variable at 50-kb scale than 500-kb scale (supplementary table S1, Supplementary Material online, *D. pseudoobscura* $P < 0.001$; *D. miranda* $P = 0.007$). It is difficult to discriminate between a biological phenomenon and larger error associated with sampling at smaller scales as the cause of this discrepancy, but worth noting that the same trend is seen in human–chimpanzee regressions (Auton et al. 2012). Although this hypothesis can be tested in humans and chimpanzees by examining for shared or nonshared hotspots, as described above, this is not possible in *Drosophila* since *Drosophila* lack recombination hotspots.

Nonetheless, this phenomenon was previously noted in several *Drosophila* recombination studies (McGaugh et al. 2012; Singh et al. 2013). In McGaugh et al. (McGaugh et al. 2012), the 20-kb recombination map revealed recombination rates ranging from 3.5 to 21.2 cM/Mb, whereas the approximately 200-kb map measured rates of 5.6 and 4.4 cM/Mb covering the same region(17.5–17.7 Mb of chromosome 2). If we examine this same region in the LD-based maps (supplementary fig. S7, Supplementary Material online), at position 17.5 Mb with resolution of 50 kb, the recombination rate in *D. pseudoobscura* is 0.32 ρ/bp, whereas in the 500-kb resolution map, the recombination rate is 0.16 ρ/bp, mirroring the empirical results. More generally, recombination rates peak at 0.32 ρ/bp in the 50-kb chromosome 2 map for both species, and top out at 0.20 ρ/bp (*D. miranda*) and 0.14 ρ/bp (*D. pseudoobscura*) in the 1-Mb chromosome 2 map. It follows that correlations of recombination rate between two species are significantly higher at 500-kb scale than at 50-kb scale ($P = 0.0094$), illustrating stronger conservation at broad scales (table 3 and supplementary figs. S7–S9, Supplementary Material online). Our observation could suggest that, even in the absence of *Prdm9* and hotspots, there is less constraint at finer scales, although if true, the mechanism behind this in *Drosophila* remains unclear.

The Distribution of Recombination across the Genome

Looking at the proportion of recombination that occurs in a proportion of sequence reveals the degree to which recombination events cluster across the genome. In humans, recombination is highly punctate, with about 80% of recombination happening in less than 20% of the sequence (Myers et al. 2006), or a Gini coefficient of approximately 0.8 (Kong et al. 2010; Kaur and Rockman 2014). As previously discussed, *Drosophila* lack hotspots characteristic of other organisms

Table 2

Empirical Recombination Rate Comparisons

	Interval Size Average, Median (Mb)	Number of Intervals	Correlation (Pearson's *r*, *P* Value)
2			
Drosophila pseudoobscura Flagstaff–Pikes Peak	0.262, 0.170	115	0.718, <0.0001
Drosophila pseudoobscura Flagstaff–*Drosophila miranda*	0.247, 0.152	115	0.621, <0.0001
Drosophila pseudoobscura Pikes Peak–*Drosophila miranda*	0.235, 0.150	123	0.669, <0.0001
XRgroup6			
Drosophila pseudoobscura Flagstaff–Pikes Peak	0.232, 0.202	47	0.916, <0.0001
Drosophila pseudoobscura Flagstaff–*Drosophila miranda*	0.228, 0.194	48	0.656, <0.0001
Drosophila pseudoobscura Pikes Peak–*Drosophila miranda*	0.197, 0.164	60	0.617, <0.0001
XRgroup8			
Drosophila pseudoobscura Flagstaff–Pikes Peak	0.388, 0.280	21	0.873, <0.0001
Drosophila pseudoobscura Flagstaff–*Drosophila miranda*	0.297, 0.257	20	0.776, <0.0001
Drosophila pseudoobscura Pikes Peak–*Drosophila miranda*	0.296, 0.249	23	0.612, <0.0001

NOTE.—Empirical recombination rates were obtained from McGaugh et al. (2012). Intervals were condensed between each pairing to correctly assess recombination rates as per McGaugh et al. (2012).

Table 3

LD-Based Comparison of Recombination Rates between *Drosophila pseudoobscura* and *Drosophila miranda*

Chromosome	Size (bp)	Scale (kb)			
		50	100	250	500
2	30,794,191	0.476	0.503	0.496	0.555
		$P < 0.0001$	$P < 0.0001$	$P < 0.0001$	$P < 0.0001$
4		0.613	0.656	0.723	0.793
		$P < 0.0001$	$P < 0.0001$	$P < 0.0001$	$P < 0.0001$
group1	5,278,962	0.266	0.327	0.296	—
		$P < 0.0001$	$P = 0.0024$	$P = 0.0697$	
group2	1,235,211	0.361	—	—	—
		$P = 0.0347$			
group3	11,692,073	0.691	0.753	0.784	0.829
		$P < 0.0001$	$P < 0.0001$	$P < 0.0001$	$P < 0.0001$
group4	6,587,037	0.621	0.658	0.787	—
		$P < 0.0001$	$P < 0.0001$	$P < 0.0001$	
XL		0.519	0.535	0.575	0.632
		$P < 0.0001$	$P < 0.0001$	$P < 0.0001$	$P < 0.0001$
group1a	9,151,740	0.646	0.684	0.753	0.899
		$P < 0.0001$	$P < 0.0001$	$P < 0.0001$	$P < 0.0001$
group1e	12,523,135	0.166	0.166	0.148	0.094
		$P = 0.0077$	$P = 0.0418$	$P = 0.2303$	$P = 0.4705$
group3a	2,690,836	0.387	0.449	—	—
		$P < 0.0001$	$P < 0.0001$		
group3b	386,183	—	—	—	—
XR		0.570	0.657	0.679	0.670
		$P < 0.0001$	$P < 0.0001$	$P < 0.0001$	$P < 0.0001$
group3a	1,468,912	0.372	0.513	—	—
		$P = 0.0014$	$P = 0.0059$		
group5	740,492	–0.025	—	—	—
		$P = 0.8898$			
group6	13,314,419	0.570	0.679	0.705	0.740
		$P < 0.0001$	$P < 0.0001$	$P < 0.0001$	$P < 0.0001$
group8	9,212,921	0.650	0.740	0.766	0.681
		$P < 0.0001$	$P < 0.0001$	$P < 0.0001$	$P < 0.0001$

NOTE.—Spearman's rho and *P* value for LD-based recombination rates of *D. pseudoobscura* and *D. miranda*. Recombination rates are averaged over a given interval size, first for the whole chromosome (all chromosome groups concatenated), then for each chromosome group individually. Dashes indicate not enough data to assess the correlation.

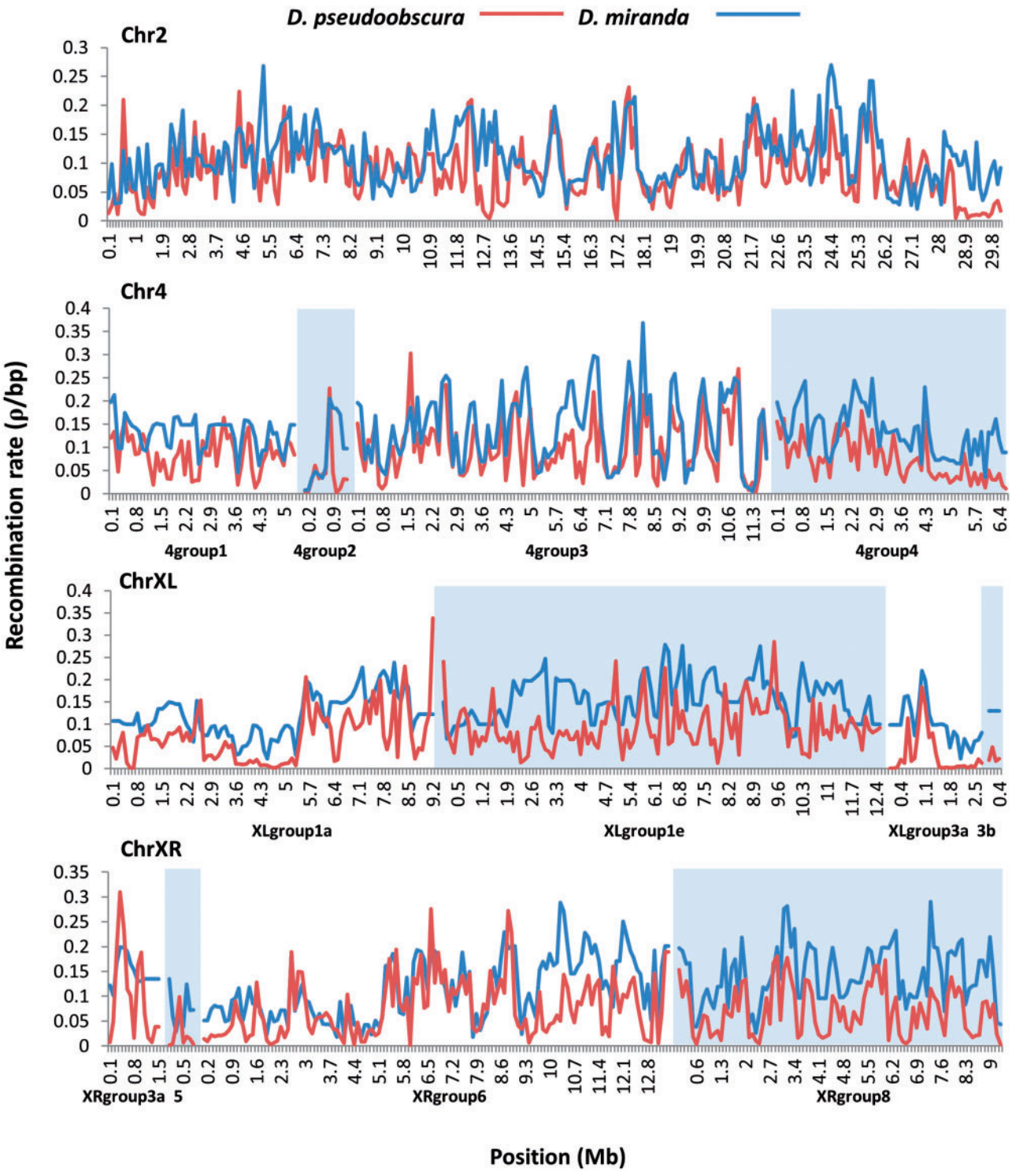

FIG. 2.—A comparison of LD-based estimations of recombination rate between two species. LD-based *D. pseudoobscura* recombination rates (red) are plotted with LD-based *D. miranda* recombination rates (blue) across the genome. Average rates were calculated for 100-kb window sizes. Chromosomes 4, XL, and XR are split into multiple groups, each labeled on the *x* axis, according to the reference assembly for *D. pseudoobscura* (denoted with alternating background color). All recombination rates are reported in ρ/bp. Error bars are not shown for clarity.

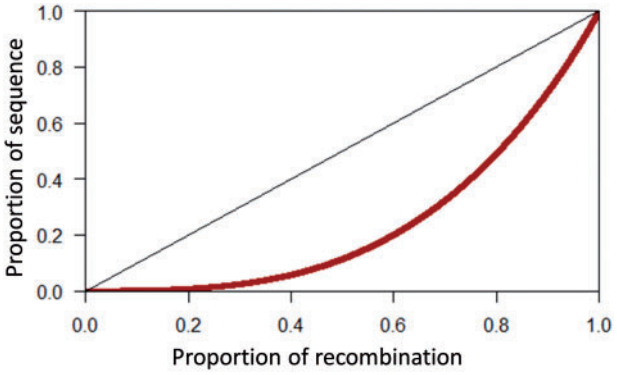

Fig. 3.—The concentration of fine-scale recombination rate in *D. pseudoobscura*.

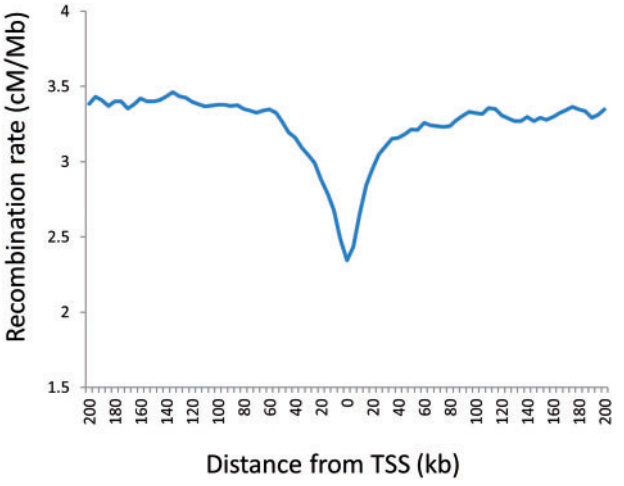

Fig. 4.—Recombination near TSS in *D. pseudoobscura*. Recombination rate was averaged for 5-kb windows for 200-kb flanking TSS across the genome.

(Singh et al. 2009, 2013; Manzano-Winkler et al. 2013), although fine scale variation has been identified (see above), and Chan et al. (2012) even detected a handful of more typical hotspots. This is reflected in *D. pseudoobscura*'s intermediate Gini coefficient of approximately 0.50, or about 80% of recombination occurring in 50% of the sequence (fig. 3). This is similar to the Gini coefficient estimated for *D. melanogaster* (0.47), falling between *C. elegans* (0.28) and *Saccharomyces cerevisiae* (0.64) (Kaur and Rockman 2014). This supports the hypothesis that recombination in *Drosophila* is more evenly distributed across the genome, with no region experiencing a near-complete absence of recombination, as compared with humans which exhibit large stretches with essentially no recombination events (Myers et al. 2006).

Another way to examine the distribution of recombination is to examine recombination rate near genomic features, such as TSS. As previously discussed, organisms without *Prdm9* exhibit an increase in recombination around TSS, most likely due to reduced nucleosome occupancy in these regions. However, in *D. pseudoobscura*, recombination is reduced around TSS and adjacent regions up to 35 kb away (fig. 4). This is supported by findings in *D. melanogaster* as well (Chan et al. 2012), although the reduction in rates appears to persist for greater distances in *D. pseudoobscura*. Similarly, we see a slight reduction in recombination rate at the 5′-end of genes, which marginally increases with distance from the start of the coding region (supplementary fig. S11, Supplementary Material online). Recombination rate within a gene (0.060 ρ/bp) is notably decreased from the average recombination rate of 0.098 ρ/bp. This is somewhat inconsistent with reports from *D. melanogaster*, although the increase in *D. melanogaster* species seems to be driven almost entirely by gene conversion events (Comeron et al. 2012).

Patterns of Linked Selection

One of the most ubiquitous patterns in molecular evolution is the association between recombination rate and nucleotide diversity, first observed in *D. melanogaster* and since reproduced in many species, including several additional species of *Drosophila* (Begun and Aquadro 1992; Smukowski and Noor 2011; Cutter and Payseur 2013; Charlesworth and Campos 2014). The association with diversity, but not nucleotide divergence, has been interpreted as evidence that selective sweeps and/or background selection are eliminating diversity in regions of low recombination (Smith and Haigh 1974; Birky and Walsh 1988; Charlesworth et al. 1993; McVicker et al. 2009; Sattath et al. 2011; Elyashiv et al. 2014). Previous research has observed these associations at scales ranging from 50 kb to 1.5 Mb, but failed to detect an association below 50 kb (Begun and Aquadro 1992; Begun et al. 2007; Kulathinal et al. 2008; Sackton et al. 2009; Stevison and Noor 2010; Singh et al. 2013). This could potentially be explained by the size of the footprint of a selective event, such as a selective sweep, which may be undetectable at fine scales. To characterize the correlation between recombination and diversity at varying scales, we assessed diversity at 4-fold degenerate positions binned into windows ranging from 5 kb to 1 Mb (supplementary fig. S12, Supplementary Material online). Similar to Singh et al. (2013), we detect a weak, but significant correlation between recombination and diversity at very fine scales, but this correlation increases in strength with window size so that at 1 Mb, recombination explains 36% of the variance in diversity. This pattern can be explained in part by the smoothing of stochasticity in diversity and recombination rate measures in large windows, and is potentially confounded by the method of recombination estimation depending on SNP density, although this method has been used for this purpose many times (Spencer et al. 2006; McVicker et al. 2009; Chan et al. 2012). We were able to retest the correlation for a sample of windows using thinned SNP data (excluding 50% of SNPs), and did not find it to alter our results. We therefore conclude

that patterns of linked selection are indeed apparent in *D. pseudoobscura*, and the power of this selection is strongest at large scales (although significant at all scales, $P < 0.0001$).

Discussion

Leveraging existing empirical recombination estimates and combining these with LD-based estimates of recombination in *D. pseudoobscura* and *D. miranda*, we present the most comprehensive comparative portrait of recombination in a pair of sibling species to date. As recombination rates can be influenced by external factors, and the stochasticity in LD-based estimations can be unclear, assessing the correlation between laboratory empirical and LD-based recombination estimates is an important step in the interpretation of LD-based recombination estimators. Here, we report robust correlations between 1) independent estimates of LD (supplementary fig. S1, Supplementary Material online), 2) LD-based and empirical estimates, and 3) independent LD-based estimates between two species. These correlations are comparable to correlations identified in independent empirical within-species comparisons (table 2), suggesting that LD-based estimates of recombination are representing about the same level of accuracy as other methods. Furthermore, correlation between empirical and LD-based recombination estimates at approximately 20-kb intervals combined with simulations of very fine scale recombination variation indicates that LD-based methods are able to accurately detect very fine scale variation (supplementary fig. S13, Supplementary Material online).

With the development of our genome-wide fine scale recombination maps for two closely related species, we were able to identify both shared and exclusive patterns of recombination across different taxa. For example, without hotspots, it was unclear whether *Drosophila* would recapitulate the pattern observed in several other species, in which broad scale recombination rates are more conserved than fine scale recombination rates. Although we could not compare rates at the 1–2 kb level, by examining correlations at scales from 500 to 50 kb, we tentatively suggest that recombination rates are less conserved at fine scales across the genome than broader scales. This trend is complicated by binning data into various physical sizes, as there is more noise associated with the correlations at fine scales than at broader scales, where the noise is averaged out. Although we cannot rule out that the weaker conservation at fine scales is driven in part by this estimation error, if this finding is confirmed, it raises interesting implications about selection pressures on recombination that persist across distantly related taxa with widely different recombination landscapes, supporting a relaxed constraint on rates at finer scales.

To illuminate potential strain variation or local changes in recombination rate over time, we can compare historical and present-day recombination rates between species (tables 1–3). Present recombination rates between species are moderately well correlated (table 3), suggesting that recombination rate has not experienced drastic changes since these species diverged, although this analysis is limited to chromosome 2 and parts of XRgroup6 and XRgroup8. Comparing historical rates (LD-based) between species at a similar scale (table 2), we see a decrease in correlation strength compared with present day, some regions quite dramatically. Generally, this likely represents variability across strains, as the LD-based recombination rates portray a population average. In specific regions such as 4group1 and XLgroup1e, which seem to exhibit very different historical recombination patterns between species, this may indicate species-specific local changes to recombination rate. As discussed above, one potential driver of this change is chromosomal rearrangements, such as inversions. Fittingly, an inversion between *D. miranda* and *D. pseudoobscura* overlaps almost exactly with the region of greatest divergence on chromosome arm XLgroup1e. This phenomenon has been previously documented in human–chimpanzee broad scale correlations, where inverted regions showed a lower correlation in recombination rate than noninverted regions (Auton et al. 2012), supporting the conclusion that inversions influence broad scale patterns of recombination across multiple species.

Another potential modifier of recombination, originally implicated in empirical work (McGaugh et al. 2012), was confirmed here by observing genome-wide elevated recombination in historical recombination rates of *D. miranda* despite a much lower effective population size. These new data provide compelling evidence that the predicted recombination modifier(s) is a species-wide trait. Of course, modifiers of recombination are well known to *Drosophila* (Baker et al. 1976; Cattani et al. 2012), and the theory of conditions that favor increased (or decreased) recombination is rich (Barton 1995, 2010; Feldman et al. 1996; Otto and Michalakis 1998; Lenormand and Otto 2000; Otto and Barton 2001; Otto and Lenormand 2002; Martin et al. 2006). The selection pressures that *D. miranda* has faced are unknown, and unfortunately, determining the locus or loci responsible for an increase in recombination in this species is not possible through traditional mapping methods, as *D. pseudoobscura* and *D. miranda* do not interbreed.

Of course one of the most recent, well-studied determinants of recombination, *Prdm9*, is absent in *Drosophila*, just one of many other differences that set *Drosophila* recombination apart from other organisms. As has already been discussed, *Drosophila* lack hotspots of recombination, and exhibit a more consistent distribution of recombination than yeast, mouse, and human. Additionally, recombination is decreased around the start of genes, in contrast to other organisms without *Prdm9*. These results add further evidence to long-known differences in *Drosophila* meiosis, which can largely be attributed to the loss of recombination in males. Whether there is a link between the evolution of sex-specific

recombination and the patterns observed more recently is unclear, but the decreased cost of sequencing and computational approaches like LDhelmet may help elucidate this question by characterizing the recombination landscape of other sex-specific recombination systems, among other questions such as why sex-specific recombination has evolved so many times, and why recombination hotspots were lost in the lineage leading to *Drosophila* when they persist across fungi, plants, and metazoans.

In conclusion, the creation of recombination maps will continue to illuminate new perspectives on the evolution of recombination, shedding light on novel determinants of recombination, patterns of selection, and genome evolution.

Supplementary Material

Supplementary figures S1–S13 and tables S1–S4 are available at *Genome Biology and Evolution* online (http://www.gbe.oxfordjournals.org/).

Acknowledgments

The authors thank L. Stevison, S. McGaugh, and four anonymous reviewers for helpful comments on the preparation of this manuscript. They are grateful to D. Bachtrog for sharing the *Drosophila miranda* whole-genome sequences with them. They also thank Andrew H. Chan for his generous guidance and assistance with LDhelmet throughout the project. This work was funded by NSF grant 1210384 and an NSF GRF to C.S.S.H.

Literature Cited

Anderson WW, Ayala FJ, Michod RE. 1977. Chromosomal and allozymic diagnosis of three species of *Drosophila*. *Drosophila pseudoobscura*, *D. persimilis*, and D. miranda. J Hered. 68:71–74.

Auton A, et al. 2012. A fine-scale chimpanzee genetic map from population sequencing. Science 336:193–198.

Auton A, et al. 2013. Genetic recombination is targeted towards gene promoter regions in dogs. PLoS Genet. 9:e1003984.

Axelsson E, et al. 2012. Death of PRDM9 coincides with stabilization of the recombination landscape in the dog genome. Genome Res. 22:51–63.

Bachtrog D, Andolfatto P. 2006. Selection, recombination and demographic history in *Drosophila miranda*. Genetics 174:2045–2059.

Backstrom N, et al. 2010. The recombination landscape of the zebra finch *Taeniopygia guttata* genome. Genome Res. 20:485–495.

Baker BS, Carpenter AT, Esposito MS, Esposito RE, Sandler L. 1976. The genetic control of meiosis. Annu Rev Genet. 10:53–134.

Baker CL, et al. 2015. PRDM9 drives evolutionary erosion of hotspots in *Mus musculus* through haplotype-specific initiation of meiotic recombination. PLoS Genet. 11:e1004916.

Barton NH. 1995. A general-model for the evolution of recombination. Genet Res. 65:123–144.

Barton NH. 2010. Mutation and the evolution of recombination. Philos Trans R Soc Lond B Biol Sci. 365:1281–1294.

Baudat F, et al. 2010. PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. Science 327:836–840.

Baudat F, Imai Y, de Massy B. 2013. Meiotic recombination in mammals: localization and regulation. Nat Rev Genet. 14:794–806.

Begun DJ, Aquadro CF. 1992. Levels of naturally-occurring DNA polymorphism correlate with recombination rates in *Drosophila melanogaster*. Nature 356:519–520.

Begun DJ, et al. 2007. Population genomics: whole-genome analysis of polymorphism and divergence in *Drosophila simulans*. PLoS Biol. 5:2534–2559.

Berg IL, et al. 2010. PRDM9 variation strongly influences recombination hot-spot activity and meiotic instability in humans. Nat Genet. 42:859.

Berg IL, et al. 2011. Variants of the protein PRDM9 differentially regulate a set of human meiotic recombination hotspots highly active in African populations. Proc Natl Acad Sci U S A. 108:12378–12383.

Birky CW Jr, Walsh JB. 1988. Effects of linkage on rates of molecular evolution. Proc Natl Acad Sci U S A. 85:6414–6418.

Birtle Z, Ponting CP. 2006. Meisetz and the birth of the KRAB motif. Bioinformatics 22:2841–2845.

Boulton A, Myers RS, Redfield RJ. 1997. The hotspot conversion paradox and the evolution of meiotic recombination. Proc Natl Acad Sci U S A. 94:8058–8063.

Brick K, Smagulova F, Khil P, Camerini-Otero RD, Petukhova GV. 2012. Genetic recombination is directed away from functional genomic elements in mice. Nature 485:642–645.

Cattani MV, Kingan SB, Presgraves DC. 2012. Cis- and trans-acting genetic factors contribute to heterogeneity in the rate of crossing over between the *Drosophila simulans* clade species. J Evol Biol. 25:2014–2022.

Chan AH, Jenkins PA, Song YS. 2012. Genome-wide fine-scale recombination rate variation in *Drosophila melanogaster*. PLoS Genet. 8:e1003090.

Charlesworth B, Campos JL. 2014. The relations between recombination rate and patterns of molecular variation and evolution in *Drosophila*. Annu Rev Genet. 48:383–403.

Charlesworth B, Morgan MT, Charlesworth D. 1993. The effect of deleterious mutations on neutral molecular variation. Genetics 134:1289–1303.

Choi K, et al. 2013. Arabidopsis meiotic crossover hot spots overlap with H2A.Z nucleosomes at gene promoters. Nat Genet. 45:1327–1336.

Cirulli ET, Kliman RM, Noor MAF. 2007. Fine-scale crossover rate heterogeneity in *Drosophila pseudoobscura*. J Mol Evol. 64:129–135.

Comeron JM, Ratnappan R, Bailin S. 2012. The many landscapes of recombination in *Drosophila melanogaster*. PLoS Genet. 8(10):e1002905.

Coop G, Przeworski M. 2007. An evolutionary view of human recombination. Nat Rev Genet. 8:23–34.

Coop G, Wen XQ, Ober C, Pritchard JK, Przeworski M. 2008. High-resolution mapping of crossovers reveals extensive variation in fine-scale recombination patterns among humans. Science 319:1395–1398.

Cutter AD, Payseur BA. 2013. Genomic signatures of selection at linked sites: unifying the disparity among species. Nat Rev Genet. 14:262–274.

DePristo MA, et al. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 43:491–498.

Dumont BL, Payseur BA. 2008. Evolution of the genomic rate of recombination in mammals. Evolution 62:276–294.

Elyashiv E, et al. 2014. A genomic map of the effects of linked selection in *Drosophila*. arXiv:1408.5461.

Feldman MW, Otto SP, Christiansen FB. 1996. Population genetic perspectives on the evolution of recombination. Annu Rev Genet. 30:261–295.

Fledel-Alon A, et al. 2009. Broad-scale recombination patterns underlying proper disjunction in humans. PLoS Genet. 5(9):e1000658.

Fuller ZL, et al. 2014. Evidence for stabilizing selection on codon usage in chromosomal rearrangements of *Drosophila pseudoobscura*. G3 4:2433–2449.
Grey C, et al. 2011. Mouse PRDM9 DNA-binding specificity determines sites of histone H3 lysine 4 trimethylation for initiation of meiotic recombination. PLoS Biol. 9(10):e1001176.
Hawley R, Harris DT, Cui W, Kramer JJ, Page SL. 2002. Meiotic chromosome segregation in *Drosophila*. Mol Biol Cell. 13:150a–150a.
Hawley RS, et al. 1992. There are 2 mechanisms of achiasmate segregation in *Drosophila* females, one of which requires heterochromatic homology. Dev Genet. 13:440–467.
Hayashi K, Yoshida K, Matsui Y. 2005. A histone H3 methyltransferase controls epigenetic events required for meiotic prophase. Nature 438:374–378.
Heil CS, Noor MA. 2012. Zinc finger binding motifs do not explain recombination rate variation within or between species of *Drosophila*. PLoS One 7:e45055.
Hellenthal G, Stephens M. 2007. msHOT: modifying Hudson's ms simulator to incorporate crossover and gene conversion hotspots. Bioinformatics 23:520–521.
Hellsten U, et al. 2013. Fine-scale variation in meiotic recombination in *Mimulus* inferred from population shotgun sequencing. Proc Natl Acad Sci U S A. 110:19478–19482.
Hey J. 2004. What's so hot about recombination hotspots? PLoS Biol 2:730–733.
Hinch AG, et al. 2011. The landscape of recombination in African Americans. Nature 476:170–175.
Hudson RR. 2001. Two-locus sampling distributions and their application. Genetics 159:1805–1817.
Jeffreys AJ, Neumann R. 2002. Reciprocal crossover asymmetry and meiotic drive in a human recombination hot spot. Nat Genet. 31:267–271.
Jeffreys AJ, Neumann R. 2005. Factors influencing recombination frequency and distribution in a human meiotic crossover hotspot. Hum Mol Genet. 14:2277–2287.
Jeffreys AJ, Neumann R. 2009. The rise and fall of a human recombination hot spot. Nat Genet. 41:625–629.
Jeffreys AJ, Neumann R, Panayi M, Myers S, Donnelly P. 2005. Human recombination hot spots hidden in regions of strong marker association. Nat Genet. 37:601–606.
Jensen JD, Bachtrog D. 2011. Characterizing the influence of effective population size on the rate of adaptation: Gillespie's Darwin domain. Genome Biol Evol. 3:687–701.
Kauppi L, Jeffreys AJ, Keeney S. 2004. Where the crossovers are: recombination distributions in mammals. Nat Rev Genet. 5:413–424.
Kaur T, Rockman MV. 2014. Crossover heterogeneity in the absence of hotspots in *Caenorhabditis elegans*. Genetics 196:137–148.
Khil PP, Camerini-Otero RD. 2010. Genetic crossovers are predicted accurately by the computed human recombination map. PLoS Genet. 6:e1000831.
Kong A, et al. 2010. Fine-scale recombination rate differences between sexes, populations and individuals. Nature 467:1099–1103.
Kulathinal RJ, Bennett SM, Fitzpatrick CL, Noor MA. 2008. Fine-scale mapping of recombination rate in *Drosophila* refines its correlation to diversity and divergence. Proc Natl Acad Sci U S A. 105:10051–10056.
Lenormand T, Otto SP. 2000. The evolution of recombination in a heterogeneous environment. Genetics 156:423–438.
Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760.
Li H, et al. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079.
Lichten M. 2008. Meiotic chromatin: the substrate for recombination initiation. In: Egel DHLR, editor. Recombination and meiosis: models, means, and evolution. Berlin: Springer-Verlag. p. 165–193.
Lichten M, Goldman ASH. 1995. Meiotic recombination hotspots. Annu Rev Genet. 29:423–444.
Loewe L, Charlesworth B, Bartolome C, Noel V. 2006. Estimating selection on nonsynonymous mutations. Genetics 172:1079–1092.
Main BJ, Smith AD, Jang H, Nuzhdin SV. 2013. Transcription start site evolution in *Drosophila*. Mol Biol Evol. 30:1966–1974.
Mancera E, Bourgon R, Brozzi A, Huber W, Steinmetz LM. 2008. High-resolution mapping of meiotic crossovers and non-crossovers in yeast. Nature 454:479–485.
Manzano-Winkler B, McGaugh SE, Noor MA. 2013. How hot are drosophila hotspots? Examining recombination rate variation and associations with nucleotide diversity, divergence, and maternal age in *Drosophila pseudoobscura*. PLoS One 8:e71582.
Martin G, Otto SP, Lenormand T. 2006. Selection for recombination in structured populations. Genetics 172:593–609.
McGaugh SE, et al. 2012. Recombination modulates how selection affects linked sites in *Drosophila*. PLoS Biol. 10:e1001422.
McGaugh SE, Noor MA. 2012. Genomic impacts of chromosomal inversions in parapatric *Drosophila* species. Philos Trans R Soc Lond B Biol Sci. 367:422–429.
McKenna A, et al. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20:1297–1303.
McKim KS, Jang JK, Manheim EA. 2002. Meiotic recombination and chromosome segregation in *Drosophila* females. Annu Rev Genet. 36:205–232.
McQuilton P, St Pierre SE, Thurmond J, FlyBase Consortium. 2012. FlyBase 101-the basics of navigating FlyBase. Nucleic Acids Res. 40:D706–D714.
McVean G, Awadalla P, Fearnhead P. 2002. A coalescent-based method for detecting and estimating recombination from gene sequences. Genetics 160:1231–1241.
McVean GA, et al. 2004. The fine-scale structure of recombination rate variation in the human genome. Science 304:581–584.
McVicker G, Gordon D, Davis C, Green P. 2009. Widespread genomic signatures of natural selection in hominid evolution. PLoS Genet. 5:e1000471.
Meznar ER, Gadau J, Koeniger N, Rueppell O. 2010. Comparative linkage mapping suggests a high recombination rate in all honeybees. J Hered. 101:S118–S126.
Miller DE, et al. 2012. A whole-chromosome analysis of meiotic recombination in *Drosophila melanogaster*. G3 2:249–260.
Myers S, Bottolo L, Freeman C, McVean G, Donnelly P. 2005. A fine-scale map of recombination rates and hotspots across the human genome. Science 310:321–324.
Myers S, et al. 2006. The distribution and causes of meiotic recombination in the human genome. Biochem Soc Trans. 34:526–530.
Myers S, et al. 2010. Drive against hotspot motifs in primates implicates the PRDM9 gene in meiotic recombination. Science 327:876–879.
Nachman MW. 2002. Variation in recombination rate across the genome: evidence and implications. Curr Opin Genet Dev. 12:657–663.
Noor MA, Schug MD, Aquadro CF. 2000. Microsatellite variation in populations of *Drosophila pseudoobscura* and *Drosophila persimilis*. Genet Res. 75:25–35.
Oliver PL, et al. 2009. Accelerated evolution of the Prdm9 speciation gene across diverse metazoan taxa. PLoS Genet. 5(12):e1000753.
Otto SP, Barton NH. 2001. Selection for recombination in small populations. Evolution 55:1921–1931.
Otto SP, Lenormand T. 2002. Resolving the paradox of sex and recombination. Nat Rev Genet. 3:252–261.
Otto SP, Michalakis Y. 1998. The evolution of recombination in changing environments. Trends Ecol Evol. 13:145–151.

Paigen K, et al. 2008. The recombinational anatomy of a mouse chromosome. PLoS Genet. 4(7):e1000119.
Paigen K, Petkov P. 2010. Mammalian recombination hot spots: properties, control and evolution. Nat Rev Genet. 11:221–233.
Pan J, et al. 2011. A hierarchical combination of factors shapes the genome-wide topography of yeast meiotic recombination initiation. Cell 144:719–731.
Pardo-Manuel de Villena F, Sapienza C. 2001. Recombination is proportional to the number of chromosome arms in mammals. Mamm Genome. 12:318–322.
Parvanov ED, Petkov PM, Paigen K. 2010. Prdm9 controls activation of mammalian recombination hotspots. Science 327:835.
Petes TD. 2001. Meiotic recombination hot spots and cold spots. Nat Rev Genet. 2:360–369.
Ponting CP. 2011. What are the genomic drivers of the rapid evolution of PRDM9? Trends Genet. 27:165–171.
Prakash S, Lewontin RC, Hubby JL. 1969. A molecular approach to the study of genic heterozygosity in natural populations. IV. Patterns of genic variation in central, marginal and isolated populations of *Drosophila pseudoobscura*. Genetics 61:841–858.
Ptak SE, et al. 2004. Absence of the TAP2 human recombination hotspot in chimpanzees. PLoS Biol. 2:e155.
Ptak SE, et al. 2005. Fine-scale recombination patterns differ between chimpanzees and humans. Nat Genet. 37:429–434.
Rambaut A, Grassly NC. 1997. Seq-Gen: an application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. Comput Appl Biosci. 13:235–238.
Rattray A, Santoyo G, Shafer B, Strathern JN. 2015. Elevated mutation rate during meiosis in *Saccharomyces cerevisiae*. PLoS Genet. 11:e1004910.
Reed FA, Tishkoff SA. 2006. Positive selection can create false hotspots of recombination. Genetics 172:2011–2014.
Roeder GS. 1997. Meiotic chromosomes: it takes two to tango. Genes Dev. 11:2600–2621.
Sackton TB, et al. 2009. Population genomic inferences from sparse high-throughput sequencing of two populations of *Drosophila melanogaster*. Genome Biol Evol. 1:449–465.
Sattath S, Elyashiv E, Kolodny O, Rinott Y, Sella G. 2011. Pervasive adaptive protein evolution apparent in diversity patterns around amino acid substitutions in *Drosophila simulans*. PLoS Genet. 7:e1001302.
Schaeffer SW, et al. 2008. Polytene chromosomal maps of 11 *Drosophila* species: the order of genomic scaffolds inferred from genetic and physical maps. Genetics 179:1601–1655.
Schaeffer SW, Miller EL. 1992. Estimates of gene flow in *Drosophila pseudoobscura* determined from nucleotide sequence analysis of the alcohol dehydrogenase region. Genetics 132:471–480.
Schweitzer MD. 1935. An analytical study of crossing over in *Drosophila melanogaster*. Genetics 20:497–527.
Sella G, Petrov DA, Przeworski M, Andolfatto P. 2009. Pervasive natural selection in the *Drosophila* genome? PLoS Genet. 5(6):e1000495.
Singh ND, Aquadro CF, Clark AG. 2009. Estimation of fine-scale recombination intensity variation in the white-echinus interval of *D. melanogaster*. J Mol Evol. 69:42–53.
Singh ND, Stone EA, Aquadro CF, Clark AG. 2013. Fine-scale heterogeneity in crossover rate in the garnet-scalloped region of the *Drosophila melanogaster* X chromosome. Genetics 194:375–387.
Slatkin M. 2008. Linkage disequilibrium—understanding the evolutionary past and mapping the medical future. Nat Rev Genet. 9:477–485.
Smagulova F, et al. 2011. Genome-wide analysis reveals novel molecular features of mouse recombination hotspots. Nature 472:375–378.
Smith JM, Haigh J. 1974. Hitch-hiking effect of a favorable gene. Genet Res. 23:23–35.
Smukowski CS, Noor MAF. 2011. Recombination rate variation in closely related species. Heredity 107:496–508.
Spencer CCA, et al. 2006. The influence of recombination on human genetic diversity. PLoS Genet. 2:1375–1385.
St Pierre SE, Ponting L, Stefancsik R, McQuilton P. 2014. FlyBase 102—advanced approaches to interrogating FlyBase. Nucleic Acids Res. 42:D780–D788.
Stevison LS, Noor MAF. 2010. Genetic and evolutionary correlates of fine-scale recombination rate variation in *Drosophila persimilis*. J Mol Evol. 71:332–345.
Wickham H. 2007. Reshaping data with the reshape package. J Stat Softw 21:1–20.
Winckler W, et al. 2005. Comparison of fine-scale recombination rates in humans and chimpanzees. Science 308:107–111.

Associate editor: Laurence Hurst