



Published in final edited form as:

J Mol Evol. 2009 July ; 69(1): 42–53. doi:10.1007/s00239-009-9250-5.

Estimation of fine-scale recombination intensity variation in the *white-echinus* interval of *D. melanogaster*

Nadia D. Singh^{*}, Charles F. Aquadro, and Andrew G. Clark

Department of Molecular Biology and Genetics, Cornell University, Ithaca NY 14853

Abstract

Accurate assessment of local recombination rate variation is crucial for understanding the recombination process and for determining the impact of natural selection on linked sites. In *Drosophila*, local recombination intensity has been estimated primarily by statistical approaches, estimating the local slope of the relationship between the physical and genetic maps. However, these estimates are limited in resolution, and as a result, the physical scale at which recombination intensity varies in *Drosophila* is largely unknown. While there is some evidence suggesting as much as a 40-fold variation in crossover rate at a local scale in *D. pseudoobscura*, little is known about the fine-scale structure of recombination rate variation in *D. melanogaster*. Here, we experimentally examine the fine-scale distribution of crossover events in a 1.2 Mb region on the *D. melanogaster* X chromosome using a classic genetic mapping approach. Our results show that crossover frequency is significantly heterogeneous within this region, varying ~ 3.5 fold. Simulations suggest that this degree of heterogeneity is sufficient to affect levels of standing nucleotide diversity, although the magnitude of this effect is small. We recover no statistical association between empirical estimates of nucleotide diversity and recombination intensity, which is likely due to the limited number of loci sampled in our population genetic dataset. However, codon bias is significantly negatively correlated with fine-scale recombination intensity estimates, as expected. Our results shed light on the relevant physical scale to consider in evolutionary analyses relating to recombination rate, and highlight the motivations to increase the resolution of the recombination map in *Drosophila*.

Keywords

Recombination rate; hotspot; crossover frequency; *D. melanogaster*

INTRODUCTION

Meiotic recombination rate, which describes the frequency with which genetic information is exchanged between homologous chromosomes during meiosis, is a key population genetic parameter. Given the well-described effects of genetic hitchhiking (Gillespie 2000; Maynard Smith and Haigh 1974) and background selection (Charlesworth et al. 1993) on levels of standing variation, it is now clear that recombination rate has a major impact on the extent to which natural selection affects linked sites. As recombination rate increases, not only does the efficacy of selection increase due to reduced interference among sites (Hill and Robertson 1966), but theory also predicts that linked variability will be reduced due to the fixation of beneficial variants or purging of deleterious variants to a degree that decreases with increased rates of recombination.

^{*}To whom correspondence should be addressed (Department of Molecular Biology and Genetics, 233 Biotechnology Building, Cornell University, Ithaca NY 14853; nds25@cornell.edu).

However, recombination rate varies widely at several scales. For instance, both the frequency and distribution of crossover events can differ markedly even among closely related species, as seen in *Drosophila* (e.g. True et al. 1996) and primates (Ptak et al. 2005; Ptak et al. 2004; Wall et al. 2003; Winckler et al. 2005). In addition, within humans, recombination rates have been shown to vary among populations (Fearnhead and Smith 2005) and individuals (Coop et al. 2008). In *Drosophila* as well, recombination rates have been shown to vary among strains (Brooks and Marks 1986). Furthermore, recombination intensity, or recombination rate scaled to physical distance (in units of centiMorgans per megabase pair, abbreviated here as cM/Mb) is heterogeneous within the genome; in *D. melanogaster*, for instance, the X chromosome and the autosomes exhibit striking differences, with a telomere-associated depression in recombination intensity on the X chromosome, but little or no such effect on the autosomes (True et al. 1996).

While these represent coarse-scale differences in recombination intensity, crossover frequency appears to vary at a fine scale as well. Recombination hotspots have been observed in a wide range of taxa including *Arabidopsis* (Drouaud et al. 2006), maize (Brown and Sundaesan 1991; Dooner et al. 1985; Fu et al. 2002), yeast (for review see Lichten and Goldman 1995; Petes 2001), humans (Myers et al. 2005), non-human primates (Ptak et al. 2005; Ptak et al. 2004; Wall et al. 2003; Winckler et al. 2005) and mice (Guillon and de Massy 2002; Kauppi et al. 2007). In humans, recombination hotspots are generally defined as regions with local increases in recombination intensity relative to background or the surrounding region, with the magnitude of the local increase between ten and several thousand fold (e.g. Li et al. 2006). Another defining feature of these hotspots is that the majority of recombination events occur within these regions; in humans, for instance, up to 80% of recombination occurs in only 10–20% of the sequence (Myers et al. 2005), and in studies mapping individual meiotic exchange events, about 60% of all such exchanges occur in previously identified hotspots (Coop et al. 2008).

It is generally thought that the recombinational landscape in *Drosophila* does not resemble that found in humans, at least in this latter regard. This is primarily due to previously reported patterns of linkage disequilibrium (LD), which decays relatively rapidly, and homogeneously, in *Drosophila* (e.g. Langley et al. 2000; Long et al. 1998; Palsson et al. 2004; Tataronkov and Ayala 2007). The absence of long haplotype blocks in *Drosophila* has thus been interpreted as being inconsistent with a recombinational landscape of highly localized hotspots in which the majority of recombination events takes place. However, it remains to be seen whether there is significant heterogeneity in recombination intensity at a fine scale in *D. melanogaster*, with fluctuations in the magnitude of recombination intensity rivaling that found in humans.

For most applications, inferences of recombination intensity in *D. melanogaster* have employed statistical methods that rely on the relationship between the physical and genetic maps. There are several examples of this type of approach, which integrate this map information over different physical scales. The adjusted coefficient of exchange (Kindahl 1994), for instance, is based on the relationship between the physical and genetic maps at a megabase scale; other sliding window approaches have been taken as well, with variable window sizes (e.g. Carvalho and Clark 1999; Hey and Kliman 2002). Polynomial regression approaches estimate recombination intensity as the derivative of a polynomial curve describing genetic location as a function of physical location. This method suffers from the fact that the estimate of recombination intensity is sensitive with the degree of the polynomial equation, which has varied among studies (Comeron and Kreitman 2000; Kliman and Hey 1993; Marais et al. 2001; Singh et al. 2004). These methods, as well as traditional genetic approaches, have revealed substantial variation in recombination rate at a broad scale, with regions of highly suppressed recombination such as centromeric regions, and recombination intensity increasing with increasing distance from the centromere (e.g. Lindsley and Sandler 1977). For instance,

the pericentromeric heterochromatin of the X chromosome has a genetic map distance of only 0.04 cM, although it may comprise up to 1/2 of total X chromosome DNA, while the remainder of the chromosome has a total genetic map length of 65 cM (Lindsley and Sandler 1977; Roberts 1965).

Although it is clear that recombination intensities in *D. melanogaster* are heterogeneous, the resolution of statistical inference of recombination intensity is limited by the genetic map, and the statistical methodology leaves considerable uncertainty about the scale at which recombination rate variation occurs. A critically unanswered question is whether the wide range in recombination intensity evident at a broad scale in *D. melanogaster* is recapitulated at a fine scale. Classical half-tetrad analysis has been used to produce fine-scale maps of only a few regions of the genome (e.g. Clark et al. 1988; Hilliker and Chovnick 1981; Hilliker et al. 1980), but this approach is not readily applied to any arbitrary chromosomal location. While there is some evidence in *D. pseudoobscura* suggesting up to 40-fold variation in the rate of crossovers at a local scale (Cirulli et al. 2007), little is known about the fine-scale structure of recombination intensity variation in *D. melanogaster*.

Here we experimentally evaluate heterogeneity in recombination intensity at a local scale in a segment of the X chromosome of *D. melanogaster*. We generated a strain carrying mutations in both *white* and *echinus*; these X-linked genes are separated by approximately 1.2 Mb and 4 cM. We focused on this region in part because prior regression-based approaches have been suggestive of moderate to high recombination intensity in this region (Hey and Kliman 2002; Kindahl 1994); were there localized regions in *D. melanogaster* that were subject to high levels of recombination, our chances of capturing such a region would likely be bolstered by focusing on a region with high recombination intensity at a broad scale. In addition, two different recombination estimators, which integrate genetic and physical map information at different scales, are somewhat incongruous in this region, which gave us confidence that there might be local heterogeneity in recombination intensity in this region in particular. Using a two-step crossing scheme, we were able to screen hundreds of thousands of flies to identify over 500 males containing a single recombination event in this region. By genotyping SNP markers across this region in our recombinant flies, we have obtained a detailed picture of the local recombinational landscape.

Our results show that crossover frequency is significantly heterogeneous in this region of the *D. melanogaster* X chromosome, with recombination intensities varying ~3.5 fold across this 1.2 Mb interval. Moreover, fine-scale recombination intensity appears correlated with codon bias in the expected direction, which is suggestive of a biological significance to this fine-scale variation. Simulation results indicate that within this range of recombination intensities, nucleotide diversity and empirical estimates of recombination intensity are weakly but significantly positively correlated under a variety of models of selection. Empirical estimates of recombination intensity do not, however, correlate with available estimates of standing nucleotide variation in this region. Our simulation results suggest that data from more loci are needed to confidently detect such correlations given the modest magnitude of heterogeneity in recombination intensity observed for this region. These results contribute to the emerging picture that recombinational landscapes of *Drosophila* species are significantly heterogeneous at a fine scale, and that the magnitude of the heterogeneity may vary among species. This variation has important implications for understanding the physical scale at which natural selection can affect the evolutionary trajectories of linked sites.

MATERIALS AND METHODS

Fly Strains

The double mutant strain of *D. melanogaster* used in this experiment contained two X-linked recessive mutations with visible phenotypes, corresponding to mutations in the *white* (*w*) and *echinus* (*ec*) genes. This strain was constructed by crossing a *white* mutant line (stock number 145 from Bloomington Drosophila Stock Center) with an *echinus* mutant line (stock number 32 from Bloomington Drosophila Stock Center) and screening for recombinant phenotypes in the F2 generation. We subsequently created an extracted X-chromosome (*i.e.* a line in which all of the X chromosomes are identical) *w-ec* line by crossing our initial *w-ec* line to the FM7a balancer stock. The wild-type line used in this study (UgX54A) corresponds to an X-extracted line from a Ugandan population of *D. melanogaster*. We screened for inversions by crossing these two parental lines and performing standard polytene chromosome squashes of the F1 progeny; these lines appear to be inversion-free relative to one another. We determined M versus P cytotype by crossing these two lines to lines with known cytotypes at 29° C and dissecting ovaries from female progeny; the *w-ec* line appears to have the M cytotype while the African line has the P cytotype.

Experimental Crosses

We used a two-step crossing scheme to generate the recombinant males used in this study (Figure 1). All crosses were set up in bottles on standard yeast-glucose media, and involved 20 virgin females and 20 males. A total of 160 bottles of crosses were established in the first round; these crosses were conducted at 21° C because this reduces the effects of hybrid dysgenesis induced by the M/P cytotype incompatibility of the two parental lines (Kidwell et al. 1977). There were a total of 258 bottles of crosses established in the second round, which all involved virgin females 24–36 hours old. These crosses were incubated at 25° C for the first five days, and were held at 21° C thereafter. The initial incubation temperature (25° C) was chosen because crossover frequency appears to increase above 22° C, at least for some regions (Ashburner 1989). While 25° C is a temperature at which MxP crosses results in hybrid dysgenesis to some extent, these effects are reduced in the F1 generation. That is, the incidence of hybrid dysgenesis in MPxM crosses (F1 daughters of an MxP cross mated to an M male) is markedly reduced relative to MxP crosses (Engels 1979). However, the effects of the cytotype incompatibility are not completely mitigated in the F1 generation, which suggests that M/P effects might play a limited role in some of our findings. The shift to 21° C was necessitated by incubator space limitations. Parents were cleared on day 4, and F2 progeny were scored on day 18, where day 0 is the day the crosses were set up. Males and females were counted, and males were scored for the recombinant phenotype. 536 recombinant males were collected (273 *+ec* males and 263 *w+* males).

SNP Markers

Eight single nucleotide polymorphisms were used to genotype the recombinant males generated in the course of this experiment. These correspond to fixed differences between our two parental lines. These SNPs were ascertained by Sanger sequencing of several noncoding regions in the 1.2 Mb region separating the *white* and *echinus* genes. With the exception of one marker (4647), all SNPs were located in intergenic regions. These markers were designed prior to the initiation of our experimental crosses. Our initial list of homogenously spaced candidate noncoding loci was refined based on ascertained SNPs among lines as well as pyrosequencing assay constraints, yielding a less-homogenously spaced set of markers that were ultimately used for genotyping. The absolute physical positions (Release 5.1 of the *D. melanogaster* genome) of our markers including our visible markers (*white*, 40355, 25645, 29649, 28369, 4647, 479, 52943 and 52792, *echinus*) were 2.69, 2.79, 2.83, 2.88, 2.98, 3.18, 3.30, 3.44, and 3.73 Mb, respectively. The distance among markers was therefore 95, 44, 46, 102, 69, 131,

122, 144, and 286 kb, respectively. More precise mapping information is available upon request. Note that these are the physical distances in the iso-1 reference strain.

Genotyping Recombinant Males

DNA was extracted from single recombinant males using a standard squish protocol. Each fly was flash frozen in liquid nitrogen, crushed with a pestle and subsequently immersed in a buffered solution (10mM Tris-Cl pH 8.2, 1mM EDTA, 25mM NaCl, 200µg/ml Proteinase K). This was incubated at 37° C for 30 min, and then at 95° C for 2 min to inactivate the Proteinase K.

Individuals were PCR amplified using primers designed by Pyrosequencing™ Assay Design Software (Biotage, Sweden). One primer in each amplifying primer pair was biotinylated. Amplifying conditions were as follows: 95° C/5 min, followed by 45 cycles of 95° C/15 s, 55° C/30 s, 72° C/15 s, with a final extension at 72° C for 5 min. Biotinylated PCR products were immobilized onto streptavidin-coated beads by combining 3 µl beads, 40 µl binding buffer (10 mM Tris-HCl; 2 M NaCl; 1 mM EDTA; 0.1% Tween 20) and 22 µl H₂O with 15 µl PCR product and vortexing for 10 min at room temperature. The immobilized DNA was washed for 5 s in 70% EtOH, the DNA was denatured for 5 s in 0.2 M NaOH, and was washed in 10 mM Tris-Acetate for 5 s; these three steps were performed using a vacuum prep tool (Biotage, Sweden). The DNA was immersed in 40 µl of 0.3 µM sequencing primer (designed by Pyrosequencing™ Assay Design Software (Biotage AB, Sweden)) that had been diluted to that concentration (from 10 µM) using annealing buffer (20 mM Tris-Acetate; 2 mM Mg-Acetate; pH 7.6). This was subsequently incubated at 85° C for 2 min. The PyroGold (Biotage, Sweden) substrate, enzyme and dNTPs were added to the PSQ cartridge at half strength (diluted 1:1 with distilled water), and the SNPs were genotyped using a PSQ 96MA Pyrosequencer and SNP software (Biotage, Sweden). For those individuals that posed challenges with one or more SNP genotyping reactions (this occurred in ~20% of individuals), genotypes were inferred where possible based on the assumption of a single recombination event in this region. Importantly, the distribution of crossovers in this 1.2 Mb region when only those flies with full genotype information were considered did not significantly differ from the distribution of crossovers in this region when all flies (including those with inferred genotypes) were considered, which suggests that this inference procedure is not systematically biasing our results.

Recombination Maps

We used a multi-point linkage approach to infer recombination maps in our two classes of recombinant males as implemented in MapMaker version 3.0 (Lander et al. 1987). Recombination rates are presented as cM (using Haldane's function), and recombination intensities are presented in cM/Mb.

Genomic Correlates of Recombination Intensity

To correlate local recombination intensity with genomic features, we estimated both codon bias and intronic GC content. We retrieved the sequences of all genes located in this region based on Release 5.1 of the *D. melanogaster* genome. We concatenated all exonic sequence within individual genes, and estimated codon bias for each gene using standalone implementation of codonW (downloaded from <http://codonw.sourceforge.net>). We used the codons defined as preferred in *D. melanogaster* to estimate the frequency of optimal codons (FOP) for each gene in this region. We similarly concatenated all of the intronic sequence within individual genes to estimate intronic GC content in a gene-by-gene fashion. We placed each gene into one of the nine intervals defined by our SNP markers based on its physical position, and assigned each gene and intron the recombination intensity estimate for that interval.

In addition, we compiled nucleotide polymorphism and divergence data from previous reports, yielding a dataset of 45 loci, each of which was sampled in both African and non-African populations. These correspond to genes and noncoding regions within our 1.2 Mb interval and were taken from several published sources (Bauer DuMont and Aquadro 2005; Jensen et al. 2007; Ometto et al. 2005; Pool et al. 2006), though some unpublished data were included (V. L. Bauer DuMont and C. F. Aquadro, personal communication). For each gene, we normalized estimates of average pairwise nucleotide diversity by divergence with *D. simulans* in an effort to control for variation in mutation rate and selective constraint among loci. We assigned each locus to one of the nine intervals based on its physical map location in Release 5.1 of the *D. melanogaster* genome, and ascribed to each locus the recombination intensity estimate for that interval.

Simulations

We conducted simulations using SFS_CODE (<http://sfscode.sourceforge.net/>; Hernandez 2008). Each replicate consisted of forward-evolving a 5000 bp segment of noncoding DNA, and sampling 10 diploid individuals at the end of each simulation run. A random 500 bp window within this 5000 bp segment was chosen for each replicate, and nucleotide diversity ($n=20$ alleles) was calculated for this window. 10,000 iterations were implemented for each parameter set. In our explored parameter space, the population recombination parameter ρ (per site) varied from 0.01 to 1, the selection parameter $\gamma=|2Ns|$ varied from 0.1 to 10, and $\theta=2N_e\mu$ was constant at 0.004. The distribution of selective effects followed a three-point mass model, with a single γ for advantageous and deleterious mutations. The probability that a novel mutation is advantageous ranged from 0.05 to 0.15, and the probability that a novel mutation is deleterious was between 0.1 and 0.5; the remaining mutations were neutral.

RESULTS AND DISCUSSION

We empirically quantified the fine-scale structure of recombination intensity in a 1.2 Mb region of the *D. melanogaster* X chromosome. This region corresponds to the interval between the *white* (*w*) and *echinus* (*ec*) genes, and based on broad-scale recombination estimators, appears subject to moderate levels of recombination ($\sim 2\text{--}3$ cM/Mb). We crossed an X-extracted *D. melanogaster* line containing both visible mutations to a wild-type African line, and backcrossed the F1 female progeny to males of the *w-ec* parental line to generate 190,642 F2 progeny. The 91,539 F2 males were scored for both phenotypic markers, and males with a single recombination event in this interval should carry one visible mutation but not the other. In total, 536 recombinant male progeny were recovered, 273 with the *+ec* phenotype, and 263 with the *w+* phenotype. Importantly, this phenotypic screen will only capture those individuals with an odd number of crossover events, so we were unable to include double crossovers in our fine-scale mapping of the crossover distribution in this region. However, triple crossovers could be detected by this approach, as could gene conversion events, although none were detected in this screen.

Our mapping experiment suggests that the recombinational distance between *white* and *echinus* is 0.59 cM for this pair of chromosomes, which contrasts with the previously reported average distance of 4.0 cM (Lindsley and Grell 1967). This latter distance is based on amalgamating map data determined from multiple laboratories, likely corresponding to crossing different lines under different conditions. The laboratory lines were most likely also of North American origin. Given the extensive variability in rates of crossing over among lines (as much as twofold, see Brooks and Marks 1986), and the sensitivity of recombination rates to factors such as maternal age (Bridges 1927; Chadov et al. 2000; Lake 1984; Redfield 1964; Redfield 1966; Stern 1926) and temperature (Grushko et al. 1991; Plough 1917; Plough 1921; Smith 1936; Stern 1926), this difference between our new and the previous estimates of

recombination distance is perhaps not surprising. However, our estimated average recombination intensity in this 1.2 Mb region is only 0.49 cM/Mb, which is also markedly lower than estimates inferred from other approaches such as the Adjusted Coefficient of Exchange or a regression polynomial approach, both of which estimate recombination intensity in this region as 2–3 cM/Mb. We do not believe that this decrease in recombination intensity is due to the lack of double crossovers being taken into account in our experimental screen, as given crossover interference, it seems unlikely that these events play a significant role at this physical scale in *Drosophila* (Cirulli et al. 2007). More likely, this reduction in crossover intensity reflects variation among lines in crossover frequency. The divergence (including small insertions and deletions) between the Ugandan line and the lab strain may also result in a lower rate of genetic exchange, as might chromosomal rearrangements too small to be scored cytologically (Sturtevant and Beadle 1936).

Finally, the M/P cytotype differences between our two parental lines may have played a role in the reduction in overall recombination intensity detected in our experiment. While the extent of hybrid dysgenesis in this second round of crosses (MPxM) is expected to be reduced relative to an MxP cross, we did these crosses at a temperature where the effects of hybrid dysgenesis can be manifested (although not to the same degree as they would at higher temperature) (Kidwell et al. 1977). Cytotype differences have been known to play a role in rates of crossing over in F1 daughters of MxP crosses (Kidwell 1977), so it is possible that the M/P cytotype differences between our two parental lines could contribute to our observed reduced rate of crossing over relative to previous reports. In the future, we hope to explore this discrepancy further, and until it becomes clear how much of our observed reduction in recombination intensity is due to M/P cytotype differences versus other contributing factors, we do not believe that this result is necessarily biologically significant. Rather, we focus on the magnitude of the variation in recombination intensity within this region, independent of the previously published map distance, as we believe them to be more robust to cytotype differences between our parental lines. While overall rates of recombination may be decreased in our experimental crosses due to M/P effects, we do not have any reason to expect *a priori* that M/P crosses alone will generate fine-scale structure of recombination intensity.

Fine-scale mapping of crossover distribution

By genotyping each of our 536 recombinant males at eight SNP markers (in addition to *w* and *ec*), we were able to localize the single crossover event in each fly to one of nine intervals. We constructed recombination maps for this region in each of our two classes of recombinants (*w* + and +*ec*) using Mapmaker. Because Mapmaker is run in a likelihood framework, we can statistically test whether estimates of recombination distances were significantly different between these two classes of recombinant flies. We compared the log-likelihood of the recombination map based on combining the *w*+ and +*ec* flies to the sum of the log-likelihoods of the maps based on each class alone. Twice the difference in these log-likelihoods should be approximately chi-square distributed with nine degrees of freedom in this fully nested comparison. Because a more highly parameterized model in which each recombinant class is allowed to have its own recombination map is not a significantly better fit to the data than is a model in which these recombinant classes are combined ($P = 0.49$, χ^2 test), this suggests that the recombination maps are not significantly different between the *w*+ and +*ec* flies. Moreover, estimates of recombination intensity (recombination distance divided by physical distance) based on +*ec* flies are significantly correlated with those based on *w*+flies (Spearman's $\rho = 0.67$, $P = 0.0499$). As a consequence, we combine the two classes of recombinants and present a single recombination intensity map in this region (Figure 2).

Within this region, the distribution of recombination events (taking into account physical interval size) is significantly different from uniform ($P = 0.004$, χ^2 test). The range of

recombination intensities spans 0.25–0.86 cM/Mbp, which corresponds to ~3.5-fold variation in recombination intensity in this 1.2 Mb region. Although it appears as though there are localized regions with increased recombination intensity, we are reluctant to deem them ‘hotspots’ given that the magnitude of the local increase appears mild and because patterns of linkage disequilibrium in *Drosophila* are inconsistent with a preponderance of recombination occurring in a restricted fraction of the sequence. To distinguish these regions from hotspots such as those found in humans, which can exhibit massive local increases in recombination intensity, and in which a disproportionate amount of recombination events occur, we will refer to these regions in *Drosophila* as ‘recombination peaks.’

Our results thus indicate that the distribution of crossover events in this small region of the *D. melanogaster* X chromosome is significantly heterogeneous. This echoes a previous finding from *D. pseudoobscura*, which suggested that fine-scale rates of crossing over are significantly different from uniform as well (Cirulli et al. 2007). Notably, this previous report in *D. pseudoobscura* was focused on an X-linked region of similar size, making the direct comparison of results appropriate. However, the height of the recombination peaks captured in the present experiment in *D. melanogaster* appears quite low, on the order of ~3.5 fold. This may in part reflect the resolution of our recombination map, given that we divided our 1.2 Mb region into 9 intervals. In a comparable study in *D. pseudoobscura* which divided a 2 Mb region into 8 intervals, recombination intensity variation was similar in magnitude, ranging from ~2–7 cM/Mbp (Cirulli et al. 2007). Closer dissection of the fine-scale structure of recombination intensity variation in *D. pseudoobscura* by additional genotyping in the interval with the highest estimated crossover rate suggested that recombination intensities can vary up to almost 40-fold (Cirulli et al. 2007). One possibility for this ascertained difference in magnitude in recombination intensity variation between *D. melanogaster* (~3.5 fold) and *D. pseudoobscura* (up to 40-fold) is that of resolution. Alternatively, this difference could reflect the generally increased rate of recombination in *D. pseudoobscura* relative to *D. melanogaster* (Hamblin and Aquadro 1999; Ortiz-Barrientos et al. 2006), which could in principle result from an increase in the number of recombination peaks or an increase in the intensity of these peaks. Further comparisons of the recombinational landscapes of *D. melanogaster* and *D. pseudoobscura* are likely to shed light on whether interspecific differences are principally driven by the frequency of recombination peaks or their intensities.

Recombination intensity and nucleotide diversity

Nucleotide polymorphism and regional rate of recombination have been shown to be positively correlated in *Drosophila melanogaster* and *D. simulans* (e.g. Begun and Aquadro 1992; Begun et al. 2007), and this could result from genetic hitchhiking (Gillespie 2000; Maynard Smith and Haigh 1974), background selection (Charlesworth et al. 1993), or some combination of both (Kim and Stephan 2000). However, the bulk of previous studies correlating nucleotide diversity and recombination rate in *Drosophila* have been limited to coarse-scale estimates of recombination rate (although see Kulathinal et al. 2008), and as a consequence, our understanding of this relationship is limited in resolution. As more fine-scale recombination rate data emerge, we can begin to dissect the physical scale at which this correlation is manifested, which can aid in disentangling the relative roles of different selective forces in generating this pattern.

To investigate whether the range of recombination intensities captured in this study is sufficiently wide to significantly impact levels of standing nucleotide diversity under different models of selection, we used a simulation approach. We used SFS_CODE (Hernandez 2008) to simulate the evolution of a 5 kb stretch of DNA, taking into account the effects of both advantageous and deleterious mutations. To make these simulations comparable to typical empirical polymorphism datasets, we randomly sampled a 500 bp window within the 5 kb

locus to estimate nucleotide diversity, and sampled 20 alleles to generate our population sample.

For the purposes of presentation, we have also included simulations with recombination intensities far exceeding those detected in this study. Our simulation results suggest that even moderately strong selection yields a significant, positive correlation between polymorphism and recombination intensities. While this is visually apparent across a range of recombination intensities spanning two orders of magnitude (Figure 5), it is also statistically significant within the narrow range of recombination intensities captured here ($\rho \approx 0.01\text{--}0.04$) (Table 1). The magnitude of the increase in diversity with increasing recombination in this range, reflected in the leftmost two values of ρ in each panel in Figure 3, appears to be small, however. In addition, the strength of the correlation appears to depend on the strength of selection; as the selection parameter increases, the dependence of polymorphism on recombination intensity increases as well (Table 1). The frequency of selective events appears to play a role as well, with more frequent selective events having a more pronounced effect on nucleotide diversity (Table 1). Overall, these results suggest that under a variety of modes of selection, recombination intensity and nucleotide diversity are positively correlated even within a modest range of recombination intensity values. However, the magnitudes of the correlation coefficients indicate that the association between recombination rate and diversity at this scale is weak, and suggest that genome-scale sampling is likely to be required to recover this pattern.

Given the availability of empirical polymorphism data in this region of the X chromosome in *D. melanogaster*, we did compare our empirically generated fine-scale recombination intensity estimates with estimates of nucleotide diversity in this region. We compiled polymorphism data from several sources, which collectively represent 45 loci across this 1.2 Mb region (which fall into 5 of our 9 intervals) sampled separately in both African and non-African populations. Because these loci include both coding and noncoding regions, we normalized our estimates of nucleotide diversity by divergence with *D. simulans*, to control for the effects of mutation rate variation. For both African populations and non-African populations, recombination intensity and normalized nucleotide diversity are not significantly correlated (Figure 4).

Our lack of significant correlation between local nucleotide diversity and recombination intensity contrasts with that recently reported in *D. pseudoobscura* (Kulathinal et al. 2008). However, local recombination intensity varied by more than an order of magnitude across the region studied in *D. pseudoobscura* (Kulathinal et al. 2008). In addition, another crucial difference between our simulation results and empirical results for *D. melanogaster* is the number of loci sampled; the significant correlations detected in our simulations were based on 10,000 iterations of a given selection/recombination model, which contrasts with the 45 loci sampled in our empirical data. To investigate whether the absence of a significant correlation in our empirical sample was due to a lack of statistical power, we sampled 11 iterations of each parameter set (for a total of 44 iterations across four recombination intensity categories corresponding to $\rho = 0.01, 0.02, 0.03, 0.04$) and examined the correlation between polymorphism and recombination intensity with these subsets. In no case was there a significant positive correlation between polymorphism and recombination intensity, suggesting that the limited number of loci sampled in our population genetic dataset may underlie our inability to capture such an association given the limited range of recombination intensities detected in this region of the *D. melanogaster* X chromosome.

Base Composition Evolution

It has previously been reported that codon bias and the GC content of noncoding sequences are significantly negatively correlated with recombination rate on the *Drosophila* X chromosome (Singh et al. 2005). This is contrary to the naïve expectation under a Hill-Robertson interference model, in which the efficacy of selection on codon bias would increase

with increasing recombination. Indeed, this expected positive correlation is seen on the autosomes specifically and in the genome in general (Kliman and Hey 1993; Marais et al. 2001; Singh et al. 2005), although the role of selection in generating this correlation remains unclear (e.g. Marais et al. 2001; Marais and Piganeau 2002; Singh et al. 2004). To examine whether the negative correlations observed at the scale of the whole X chromosome were recapitulated at a fine scale, we estimated codon bias and intronic GC content for the 66 genes within the 1.2 Mb region under study. Consistent with the negative correlation at the level of the entire X chromosome, we found that codon bias is significantly negatively correlated with fine-scale recombination rate in this 1.2 Mb region of the X chromosome (Spearman's $\rho = -0.31$, $P = 0.01$) (Figure 5a). However, intronic GC content (based on 47 intron-containing genes) was not significantly correlated with fine-scale estimates of recombination intensity in this region (Figure 5b). This difference between coding and noncoding sequences is not likely to reflect differences in power; while there are more point estimates of codon bias than there are of intronic GC content, the average length of the intronic sequence (13.7 kb) is substantially larger than the average exonic length (1.3 kb) per gene, which suggests that the intronic point estimates reflect less noise than their codon bias counterparts. More likely, the difference observed between codon bias and intronic GC content with respect to their relationship with recombination intensity reflects differences in the evolutionary forces serving to modulate base composition of coding versus noncoding sequences. While base composition at synonymous sites may be subject to selection for translational efficiency, for instance, base composition at intronic sites may be subject to different selection pressures, and/or different strengths of selection. In addition, the physical scale at which base composition is biologically relevant may differ between coding and noncoding sequences, with perhaps short-range functional significance in coding sequence and long-range importance in noncoding sequences, for instance.

CONCLUSIONS AND FUTURE DIRECTIONS

Our investigation of the fine-scale structure of recombination intensity in a 1.2 Mb region of the *Drosophila melanogaster* X chromosome revealed significant heterogeneity in the distribution of crossover events. While the magnitude of the variation in crossover frequency appears small (~3.5 fold), these data, in combination with similar results from *D. pseudoobscura* (Cirulli et al. 2007), suggest that recombination intensity in *Drosophila* does indeed vary significantly at a fine-scale. Moreover, our estimates of recombination intensity correlate with genomic features such as codon bias even at this fine-scale, which is suggestive of a biological significance to this heterogeneity. Furthermore, our simulation results show that recombination intensity and levels of nucleotide diversity should be weakly but significantly positively correlated within the range of empirical values of recombination intensity measured here, provided that sampling is sufficiently dense. While we did not find evidence for a significant association between fine-scale recombination intensity and levels of nucleotide diversity in our empirical study, and our simulation results suggest that this may be due to a lack of statistical power given the limited number of loci in this region with available population genetic data.

More work is needed to assess the importance of scale both with respect to the distribution of crossover events in *Drosophila*, as well as with respect to the relationship between polymorphism and recombination intensity. Perhaps most importantly, our understanding of the magnitude in recombination intensity fluctuation at a local scale (the height of the recombination peaks) would benefit considerably from increased resolution of our recombination map. Because the intervals studied here range from 40–300 kb, we are limited in our ability to detect heterogeneity at an ultra-fine scale, such as on the order of kilobases or even tens of kilobases, as these effects might be swamped in these larger intervals. At increased resolution, we will have considerably more power to detect whether recombination peaks vary

in intensity in *D. melanogaster* (as they do in humans, for instance), and what the distribution of recombination peak width is with respect to physical distance covered. Critically, it also remains to be seen whether the heterogeneity in recombination intensity captured by our study is a general feature other *D. melanogaster* lines and populations, of the X chromosome, or of the *D. melanogaster* genome as a whole. Moreover, the extent to which recombination intensity varies at a fine-scale in other *Drosophila* species could benefit from further study, particularly in close relatives of *D. melanogaster*, and comparative analysis of the recombinational landscapes among species will shed much light on how labile recombination rates are over evolutionary time. Finally, as our recombination maps in *Drosophila* continue to gain resolution, we can determine the biologically relevant scale at which levels of nucleotide diversity are affected by recombination rates.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the entire Clark and Aquadro labs (particularly S. Hackett and H. Flores), along with B. P. Lazzaro, M. Hamblin, and J. G. Mezey, for their tireless fly sorting and scoring efforts. The authors also thank H. Flores, C. Pergusoles, A. Larracuente, and J. Werner for their assistance with screening our experimental lines for inversions, and X. Wang for his help with optimizing the pyrosequencing assays. R. Hernandez generously gave us permission to use his program SFS_CODE for our simulations, and the authors are also indebted to him for his guidance in implementing this program. This work was supported in part by an NIH NRSA (grant number 1F32GM080944-01 to NDS, CFA and AGC). The original version of this manuscript is available at www.springerlink.com, and the DOI is 10.1007/s00239-009-9250-5.

LITERATURE CITED

- Ashburner, M. *Drosophila: A Laboratory handbook*. Cold Spring Harbor: Cold Spring Harbor Laboratory Press; 1989.
- Bauer DuMont V, Aquadro CF. Multiple signatures of positive selection downstream of Notch on the tip X chromosome in *Drosophila melanogaster*. *Genetics* 2005;171:639–653. [PubMed: 16020794]
- Begun DJ, Aquadro CF. Levels of naturally occurring DNA polymorphism correlate with recombination rates in *Drosophila melanogaster*. *Nature (London)* 1992;356:519–520. [PubMed: 1560824]
- Begun DJ, Holloway AK, Stevens K, Hillier LW, Poh Y-P, Hahn MW, Nista PM, Jones CD, Kern AD, Dewey C, Pachter L, Myers E, Langley CH. Population genomics: Whole-genome analysis of polymorphism and divergence in *Drosophila simulans*. *PLoS Biology* 2007;5:2534–2559.
- Bridges CB. The relation of the age of the female to crossing over in the third chromosome of *Drosophila melanogaster*. *Journal of General Physiology* 1927;8:689–700.
- Brooks LD, Marks RW. The organization of genetic variation for recombination in *Drosophila melanogaster*. *Genetics* 1986;114:525–547. [PubMed: 3095185]
- Brown J, Sundaresan V. A recombination hotspot in the maize A1 intragenic region. *Theoretical and Applied Genetics* 1991;81:185–188.
- Carvalho AB, Clark AG. Intron size and natural selection. *Nature (London)* 1999;401:344. [PubMed: 10517631]
- Chadov BF, Chadova EV, Anan'ina GN, Kopyl SA, Volkova EI. Age changes in crossing over in *Drosophila* are similar to changes resulting from interchromosomal effects of a chromosome rearrangement on crossing over. *Russian Journal of Genetics* 2000;36:258–264.
- Charlesworth B, Morgan MT, Charlesworth D. The effect of deleterious mutations on neutral molecular variation. *Genetics* 1993;134:1289–1303. [PubMed: 8375663]
- Cirulli ET, Kliman RM, Noor MAF. Fine-scale crossover rate heterogeneity in *Drosophila pseudoobscura*. *Journal of Molecular Evolution* 2007;64:129–135. [PubMed: 17160365]
- Clark SH, Hilliker AJ, Chovnick A. Recombination can initiate and terminate at a large number of sites within the *rosy* locus of *Drosophila melanogaster*. *Genetics* 1988;118:261–266. [PubMed: 2834266]
- Comeron JM, Kreitman M. The Correlation Between Intron Length and Recombination in *Drosophila*: Dynamic Equilibrium Between Mutational and Selective Forces. *Genetics* 2000;156:1175–1190. [PubMed: 11063693]

- Coop G, Wen XQ, Ober C, Pritchard JK, Przeworski M. High-resolution mapping of crossovers reveals extensive variation in fine-scale recombination patterns among humans. *Science* 2008;319:1395–1398. [PubMed: 18239090]
- Dooner HK, Weck E, Adams S, Ralston E, Favreau M, English J. A molecular genetic analysis of insertions in the Bronze locus in maize *Molecular & General Genetics* 1985;200:240–246.
- Drouaud J, Camilleri C, Bourguignon PY, Canaguier A, Berard A, Vezon D, Giancola S, Brunel D, Colot V, Prum B, Quesneville H, Mezard C. Variation in crossing-over rates across chromosome 4 of *Arabidopsis thaliana* reveals the presence of meiotic recombination "hot spots". *Genome Research* 2006;16:106–114. [PubMed: 16344568]
- Engels WR. Hybrid dysgenesis in *Drosophila melanogaster*: rules of inheritance of female sterility. *Genetical Research* 1979;33:219–236.
- Fearnhead P, Smith NGC. A novel method with improved power to detect recombination hotspots from polymorphism data reveals multiple hotspots in human genes. *American Journal of Human Genetics* 2005;77:781–794. [PubMed: 16252238]
- Fu HH, Zheng ZW, Dooner HK. Recombination rates between adjacent genic and retrotransposon regions in maize vary by 2 orders of magnitude. *Proceedings of the National Academy of Sciences of the United States of America* 2002;99:1082–1087. [PubMed: 11792865]
- Gillespie JH. Genetic drift in an infinite population: the pseudohitchhiking model. *Genetics* 2000;155:909–919. [PubMed: 10835409]
- Grushko TA, Korochkina SE, Klimenko VV. Temperature-dependent control of crossing-over frequency in *Drosophila melanogaster*: The impact upon recombination frequency of infraoptimal and superoptimal shock temperatures in the course of early ontogeny. *Genetika* 1991;27:1714–1721. [PubMed: 1778450]
- Guillon H, de Massy B. An initiation site for meiotic crossing-over and gene conversion in the mouse. *Nature Genetics* 2002;32:296–299. [PubMed: 12244318]
- Hamblin MT, Aquadro CF. DNA sequence variation and the recombinational landscape in *Drosophila pseudoobscura*: a study of the second chromosome. *Genetics* 1999;153:859–869. [PubMed: 10511563]
- Hernandez RD. A flexible forward simulator for populations subject to selection and demography. *Bioinformatics* 2008;24:2786–2787. [PubMed: 18842601]
- Hey J, Kliman RM. Interactions Between Natural Selection, Recombination and Gene Density in the Genes of *Drosophila*. *Genetics* 2002;160:595–608. [PubMed: 11861564]
- Hill WG, Robertson A. The effect of linkage on limits to artificial selection. *Genetical Research* 1966;8:269–294. [PubMed: 5980116]
- Hilliker AJ, Chovnick A. Further observations on intragenic recombination in *Drosophila melanogaster*. *Genetical Research* 1981;38:281–296. [PubMed: 6800885]
- Hilliker AJ, Clark SH, Chovnick A, Gelbart WM. Cytogenetic analysis of the chromosomal region immediately adjacent to the *rosy* locus in *Drosophila melanogaster*. *Genetics* 1980;95:95–110. [PubMed: 6776006]
- Jensen JD, DuMont VLB, Ashmore AB, Gutierrez A, Aquadro CR. Patterns of sequence variability and divergence at the diminutive gene region of *Drosophila melanogaster*: complex patterns suggest an ancestral selective sweep. *Genetics* 2007;177:1071–1085. [PubMed: 17720938]
- Kauppi L, Jasin M, Keeney S. Meiotic crossover hotspots contained in haplotype block boundaries of the mouse genome. *Proceedings of the National Academy of Sciences of the United States of America* 2007;104:13396–13401. [PubMed: 17690247]
- Kidwell MG. Reciprocal differences in female recombination associated with hybrid dysgenesis in *Drosophila melanogaster*. *Genetical Research* 1977;30:77–88. [PubMed: 411714]
- Kidwell MG, Kidwell JF, Sved JA. Hybrid dysgenesis in *Drosophila melanogaster*: A syndrome of aberrant traits including mutation, sterility and male recombination. *Genetics* 1977;86:813–833. [PubMed: 17248751]
- Kim Y, Stephan W. Joint effects of genetic hitchhiking and background selection on neutral variation. *Genetics* 2000;155:1415–1427. [PubMed: 10880499]
- Kindahl, EC. Recombination and DNA polymorphism on the third chromosome of *Drosophila melanogaster*. Ithaca, NY: Cornell University; 1994.

- Kliman RM, Hey J. Reduced natural selection associated with low recombination in *Drosophila melanogaster*. *Molecular Biology and Evolution* 1993;10:1239–1258. [PubMed: 8277853]
- Kulathinal RJ, Bennett SM, Fitzpatrick CL, Noor MAF. Fine-scale mapping of recombination rate in *Drosophila* refines its correlation to diversity and divergence. *Proceedings of the National Academy of Sciences of the United States of America* 2008;105:10051–10056. [PubMed: 18621713]
- Lake S. Variation in the recombination frequency and the relationship of maternal age in brood analysis of the distal and centromeric regions of the X-chromosome in temperature shocked reciprocal hybrids of inbred lines of *Drosophila melanogaster*. *Hereditas* 1984;100:121–129.
- Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L. MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1987;1:174–181. [PubMed: 3692487]
- Langley CH, Lazzaro BP, Phillips W, Heikkinen E, Braverman JM. Linkage disequilibria and the site frequency spectra in the *su(s)* and *su(wa)* regions of the *Drosophila melanogaster* X chromosome. *Genetics* 2000;156:1837–1852. [PubMed: 11102378]
- Li J, Zhang MQ, Zhang X. A new method for detecting human recombination hotspots and its applications to the HapMap ENCODE data. *American Journal of Human Genetics* 2006;79:628–639. [PubMed: 16960799]
- Lichten M, Goldman ASH. Meiotic Recombination Hotspots. *Annual Review of Genetics* 1995;29:423–444.
- Lindsley DL, Grell EH. Genetic Variations of *Drosophila melanogaster*. Carnegie Inst. of Washington. 1967Pub. No. 627
- Lindsley DL, Sandler L. The genetic analysis of meiosis in female *Drosophila melanogaster*. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences* 1977;277:295–213.
- Long AD, Lyman RF, Langley CH, Mackay TFC. Two sites in the Delta gene region contribute to naturally occurring variation in bristle number in *Drosophila melanogaster*. *Genetics* 1998;149:999–1017. [PubMed: 9611209]
- Marais G, Mouchiroud D, Duret L. Does recombination improve selection on codon usage? Lessons from nematode and fly complete genomes. *Proceedings of the National Academy of Sciences, USA* 2001;98:5688–5692.
- Marais G, Piganeau G. Hill-Robertson interference is a minor determinant of variations in codon bias across *Drosophila melanogaster* and *Caenorhabditis elegans* genomes. *Molecular Biology and Evolution* 2002;19:1399–1406. [PubMed: 12200468]
- Maynard Smith J, Haigh J. The hitch-hiking effect of a favourable gene. *Genetical Research* 1974;23:23–35. [PubMed: 4407212]
- Myers S, Bottolo L, Freeman C, McVean G, Donnelly P. A fine-scale map of recombination rates and hotspots across the human genome. *Science* 2005;310:321–324. [PubMed: 16224025]
- Ometto L, Glinka S, De Lorenzo D, Stephan W. Inferring the effects of demography and selection on *Drosophila melanogaster* populations from a chromosome-wide scan of DNA variation. *Molec. Biol. Evol* 2005;22:2119–2130. [PubMed: 15987874]
- Ortiz-Barrientos D, Audrey SC, Noor MAF. A recombinational portrait of the *Drosophila pseudoobscura* genome. *Genetical Research* 2006;87:23–31. [PubMed: 16545148]
- Palsson A, Rouse A, Riley-Berger R, Dworkin I, Gibson G. Nucleotide variation in the Egfr locus of *Drosophila melanogaster*. *Genetics* 2004;167:1199–1212. [PubMed: 15280235]
- Petes TD. Meiotic recombination hot spots and cold spots. *Nature Reviews Genetics* 2001;2:360–369.
- Plough HH. The effect of temperature on crossingover in *Drosophila*. *Journal of Experimental Zoology* 1917;24:147–209.
- Plough HH. Further studies on the effect of temperature on crossing over. *Journal of Experimental Zoology* 1921;32:187–202.
- Pool JE, Bauer DuMont V, Mueller JL, Aquadro CF. A Scan of Molecular Variation Leads to the Narrow Localization of a Selective Sweep Affecting Both Afrotropical and Cosmopolitan Populations of *Drosophila melanogaster*. *Genetics* 2006;172:1093–1105. [PubMed: 16322515]

- Ptak SE, Hinds DA, Koehler K, Nickel B, Patil N, Ballinger DG, Przeworski M, Frazer KA, Paabo S. Fine-scale recombination patterns differ between chimpanzees and humans. *Nature Genetics* 2005;37:429–434. [PubMed: 15723063]
- Ptak SE, Roeder AD, Stephens M, Gilad Y, Paabo S, Przeworski M. Absence of the TAP2 human recombination hotspot in chimpanzees. *Plos Biology* 2004;2:849–855.
- Redfield H. Regional association of crossing over in nonhomologous chromosomes in *Drosophila melanogaster* and its variation with age. *Genetics* 1964;49:319–342. [PubMed: 14124947]
- Redfield H. Delayed mating and the relationship of recombination to maternal age in *Drosophila melanogaster*. *Genetics* 1966;53:593–607. [PubMed: 5919332]
- Roberts PA. Difference in behaviour in eu- and heterchromatin: crossing over. *Nature* 1965;205:725–726. [PubMed: 14287435]
- Singh ND, Arndt PF, Petrov DA. Genomic Heterogeneity of Background Substitutional Patterns in *Drosophila melanogaster*. *Genetics* 2004;169:709–722. [PubMed: 15520267]
- Singh ND, Davis JC, Petrov DA. Codon bias and Noncoding GC content correlate negatively with recombination rate on the *Drosophila* X chromosome. *Journal of Molecular Evolution* 2005;61:315–324. [PubMed: 16044248]
- Smith HF. Influence of temperature on crossing-over in *Drosophila*. *Nature* 1936;138:329–330.
- Stern C. An effect of temperature and age on crossing over in the first chromosome of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America* 1926;12:530–532. [PubMed: 16587124]
- Sturtevant AH, Beadle GW. The relations of inversions in the X chromosome of *Drosophila melanogaster* to crossing over and disjunction. *Genetics* 1936;21:554–604. [PubMed: 17246812]
- Tatarenkov A, Ayala FJ. Nucleotide variation at the *dopa decarboxylase* (*Ddc*) gene in natural populations of *Drosophila melanogaster*. *Journal of Genetics* 2007;86:125–137. [PubMed: 17968140]
- True JR, Mercer JM, Laurie CC. Differences in crossover frequency and distribution among three sibling species of *Drosophila*. *Genetics* 1996;142:507–523. [PubMed: 8852849]
- Wall JD, Frisse LA, Hudson RR, Di Rienzo A. Comparative linkage-disequilibrium analysis of the beta-globin hotspot in primates. *American Journal of Human Genetics* 2003;73:1330–1340. [PubMed: 14628290]
- Winckler W, Myers SR, Richter DJ, Onofrio RC, McDonald GJ, Bontrop RE, McVean GAT, Gabriel SB, Reich D, Donnelly P, Altshuler D. Comparison of fine-scale recombination rates in humans and chimpanzees. *Science* 2005;308:107–111. [PubMed: 15705809]

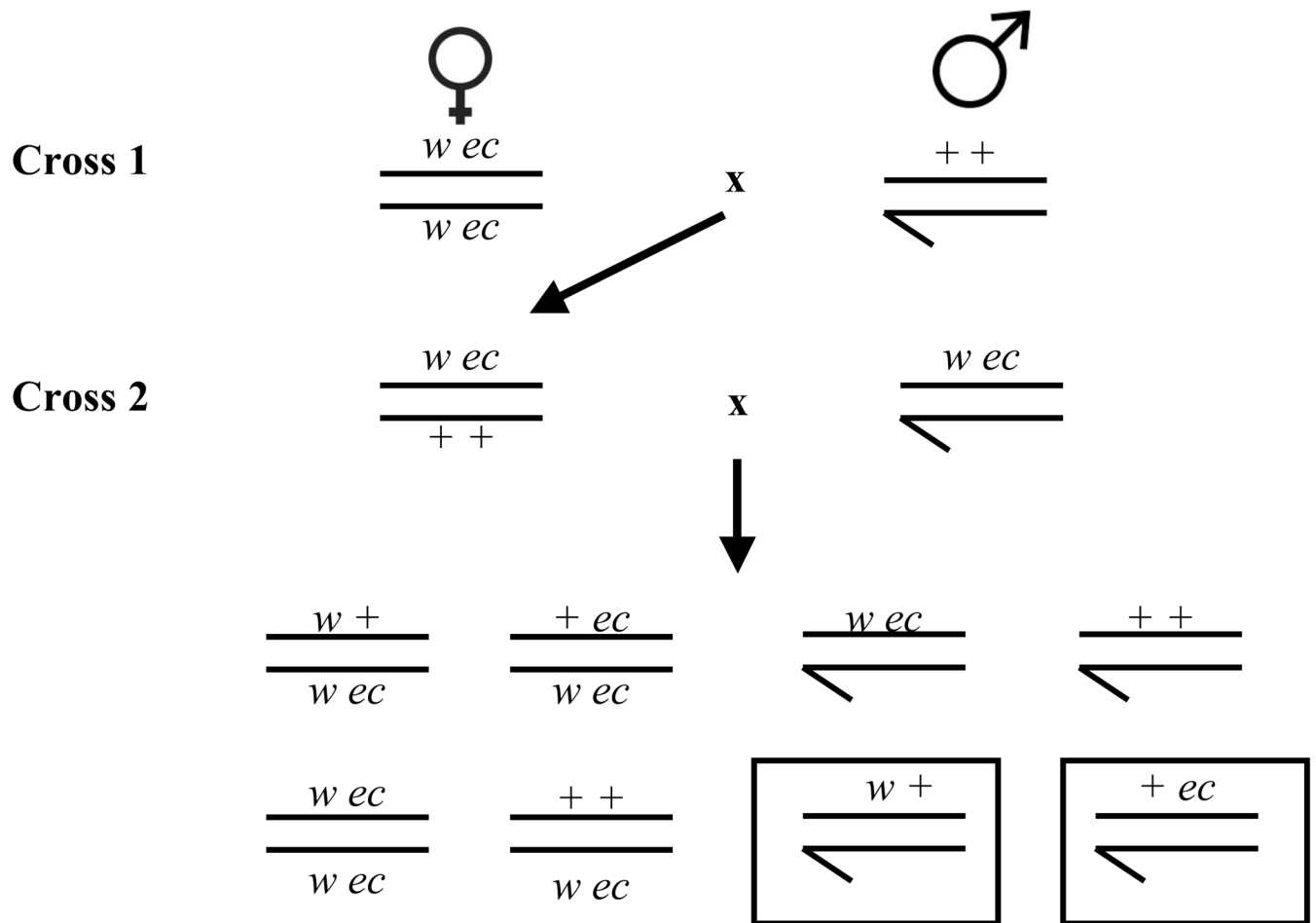


Figure 1. Schematic representation of our two-step crossing scheme. Boxed F2 male progeny correspond to the two recombinant genotypes identified by our screen; these males contain a single crossover in the region between *white* and *echinus*.

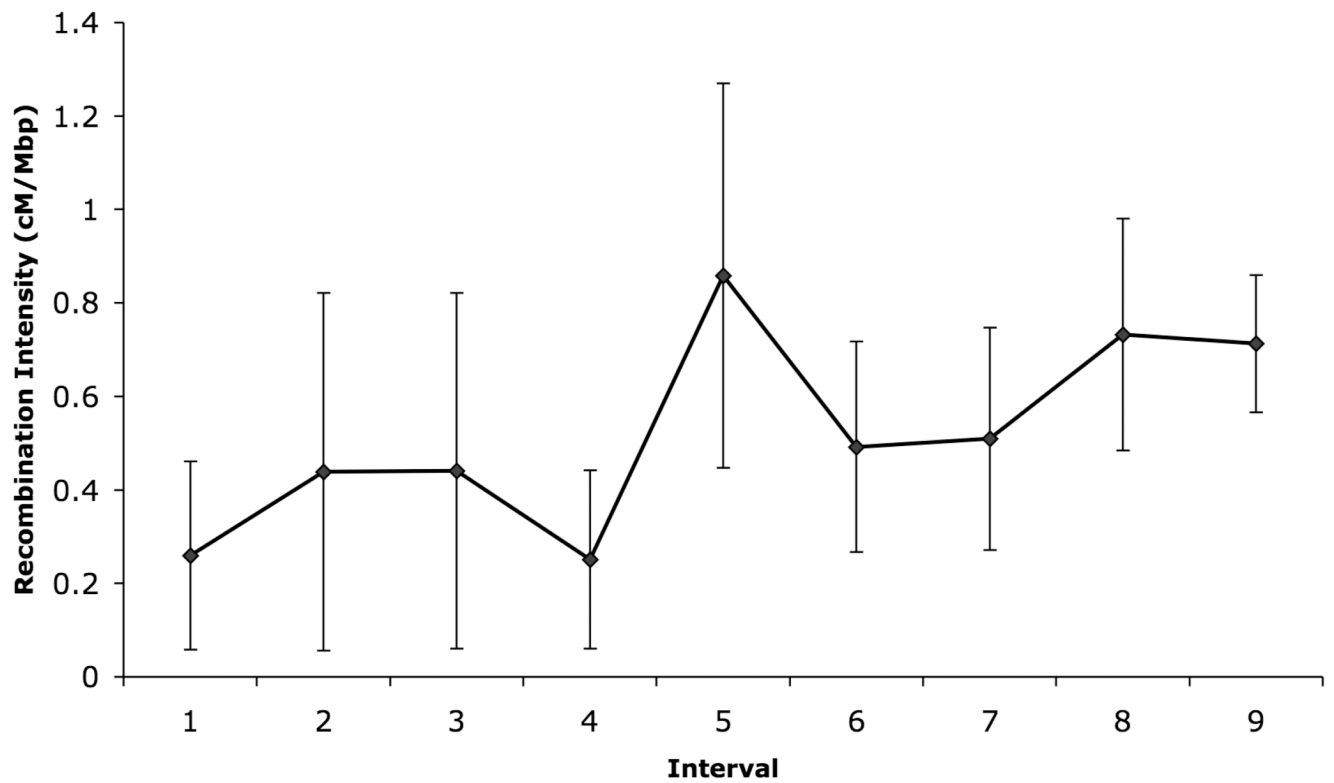


Figure 2. Recombination intensity (in cM/Mbp) as estimated by crossover frequency per physical distance. Error bars correspond to the 95% confidence interval. The intervals are defined by eight SNP markers in combination with the two visible phenotypic markers (*white* and *echinus*).

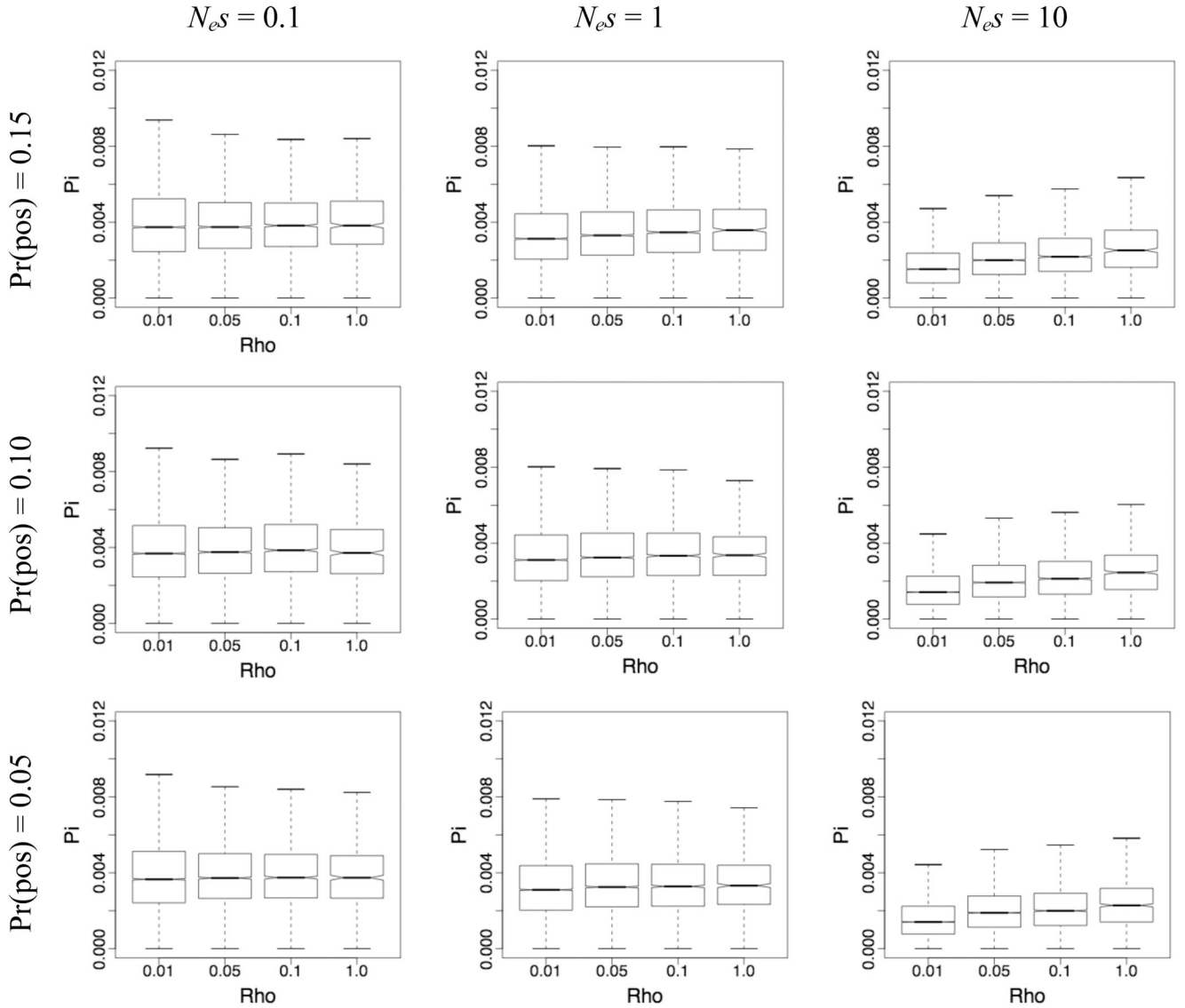


Figure 3.

Representative simulation results. Boxplots correspond to distributions of nucleotide diversity as a function of recombination rate under different parameter combinations, with the selection parameter ranging from 0.1 to 10 and the probability of a novel mutation being advantageous ranging from 0.05 to 0.15. For these plots, the probability of a novel mutation being deleterious is constant at 0.5. The range of the population recombination parameter ρ is consistent across plots, ranging from 0.01 to 1, and the leftmost two values in each plot (0.01 and 0.05) correspond roughly to the range of recombination intensity captured by our empirical study. The notch in each box corresponds to the median value, and the lower and upper edges of the box correspond to the 25th and 75th percentile, respectively. Whiskers extend to the most extreme datapoint which is no more than 1.5 times the interquartile range from the box.

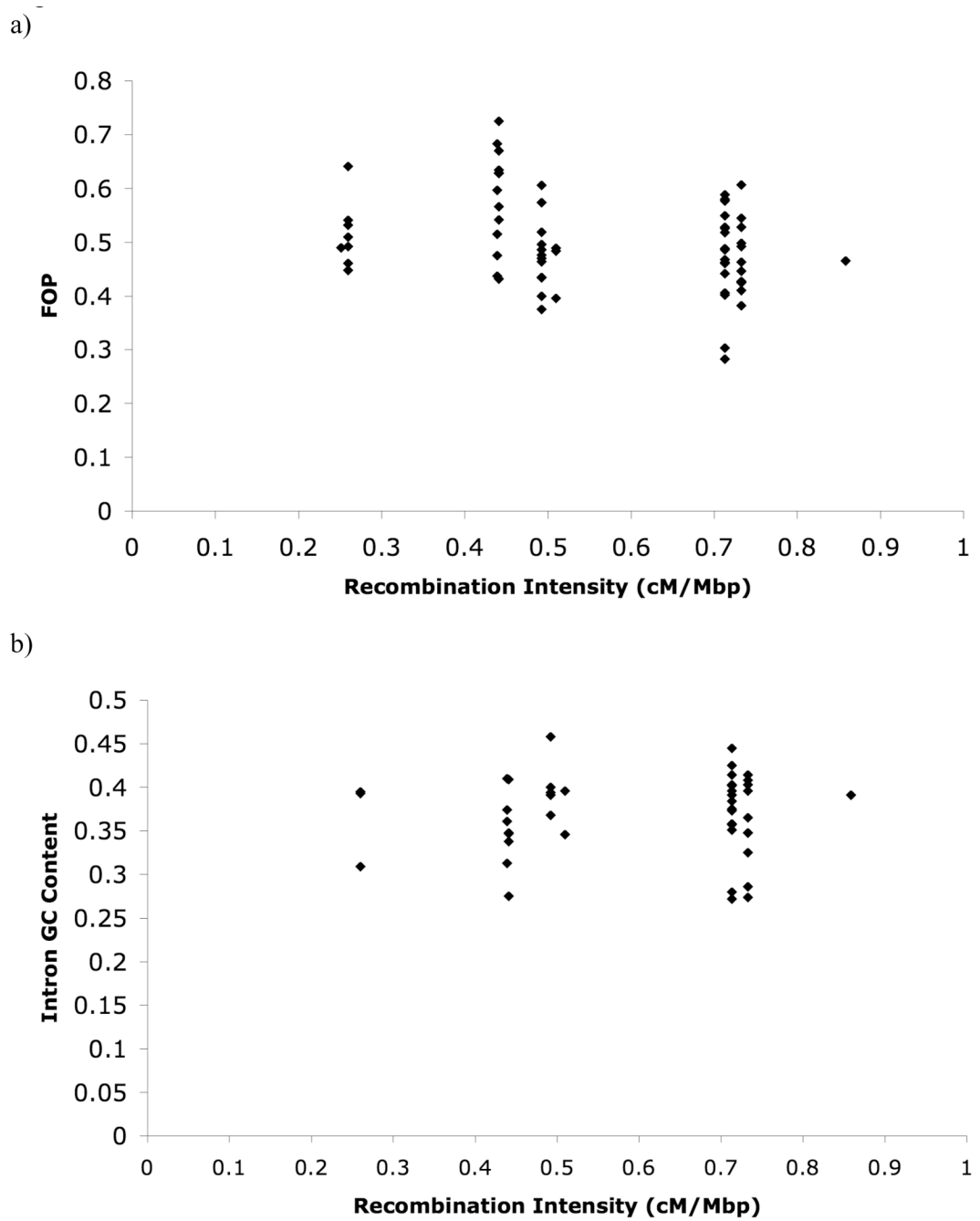


Figure 5. Scatterplot of recombination intensity and a) codon bias, as measured by the frequency of optimal codons (FOP) and b) intronic GC content.

Table 1

Representative simulation results: Selection models and correlation coefficients, based on 10,000 iterations of each of four recombination intensity values ($Rho = 0.01, 0.02, 0.03, 0.04$)

Pr(advantageous)	Pr(deleterious)	$N_e s$	Spearman's Correlation Coefficient	P-value
0.15	0.50	10	0.1258	<0.0001
0.15	0.50	1	0.032	<0.0001
0.15	0.50	0.1	0.0046	0.488
0.10	0.50	10	0.1235	<0.0001
0.10	0.50	1	0.032	<0.0001
0.10	0.50	0.1	0.0144	0.0317
0.05	0.50	10	0.12	<0.0001
0.05	0.50	1	0.0244	<0.0001
0.05	0.50	0.1	0.0181	0.0003