Fine-scale mapping of recombination rate in *Drosophila* refines its correlation to diversity and divergence

Rob J. Kulathinal*, Sarah M. Bennett†, Courtney L. Fitzpatrick†, and Mohamed A. F. Noor†‡

*Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138; and †Biology Department, Box 90338, Duke University, Durham, NC 27708

Edited by Stephen W. Schaeffer, Penn State University, University Park, PA, and accepted by the Editorial Board May 7, 2008 (received for review February 25, 2008)

Regional rates of recombination often correlate with levels of nucleotide diversity, and either selective or neutral hypotheses can explain this relationship. Regional recombination rates also correlate with nucleotide differences between human and chimpanzee, consistent with models where recombination is mutagenic; however, a lack of correlation is observed in the *Drosophila melanogaster* group, consistent with models invoking natural selection. Here, we revisit the relationship among recombination, diversity, and interspecies difference by generating empirical estimates of these parameters in *Drosophila pseudoobscura*. To measure recombination rate, we genotyped 1,294 backcross hybrids at 50 markers across the largest assembled linkage group in this species. Genome-wide diversity was estimated by sequencing a second isolate of *D. pseudoobscura* at shallow coverage. Alignment to the sequenced genome of the closely related species, *Drosophila persimilis*, provided nucleotide site orthology. Our findings demonstrate that scale is critical in determining correlates to recombination rate: fine-scale cross-over rate estimates are far stronger predictors of both diversity and interspecies difference than broad-scale estimates. The correlation of fine-scale recombination rate to diversity and interspecies difference appears to be genome-wide, evidenced by examination of an X-linked region in greater detail. Because we observe a strong correlation of cross-over rate with interspecies difference, even after correcting for segregating ancestral variation, we suggest that both mutagenic and selective forces generate these correlations, the latter in regions of low crossing over. We propose that it is not cross-overs *per se* that are mutagenic, but rather repair of DNA double-strand break precursors via crossing over and gene conversion.

variation | crossing over

One of the most influential observations in molecular evolutionary biology is the relationship between recombination rate and nucleotide diversity within species. In a pioneering study, Begun and Aquadro (1) identified a strong positive association between nucleotide polymorphism in *Drosophila melanogaster* (as measured from restriction site data of 20 loci) and recombination rate. If this relationship was caused by recombination itself being mutagenic, a similar relationship between recombination rate and nucleotide divergence to the sister species, *Drosophila simulans*, should also exist; however, this relationship was not observed. More recently, Begun *et al.* (2) examined whole-genome shotgun sequences from multiple strains of *D. simulans* and the published *D. melanogaster* genome assembly and confirmed these initial observations at an unprecedented scale. The association between recombination rate and diversity, but not divergence, in *Drosophila* is considered to be driven primarily by natural selection (3–5): fixation of positively selected variants and associated hitchhiking effects (6) and/or background selection eliminating deleterious alleles (7).

Similar studies in other taxa have yielded conflicting results. Although, in most cases, associations between recombination rate and nucleotide diversity are observed, they sometimes can be quite weak (8, 9), and in some species such as maize and humans, significant associations are observed between recombination rate and interspecies divergence (10–12). Because regions of severely reduced recombination often exhibit typical levels of interspecific divergence but reduced diversity within species, it is unequivocal that selective forces have contributed to this pattern in some species. Less clear is how much residual variation is explained by other forces, such as mutational heterogeneity associated with recombination rate.

The relative role of these forces would be better understood if a more fine-grained and reliable recombinational map was available. If there is fine-scale heterogeneity in recombination rate, a mutational relationship between recombination and nucleotide diversity or divergence would be easiest to detect by examining recombination over small, rather than large, spatial scales (assuming some temporal stability of recombination hotspots). Fine-scale heterogeneity in recombination rate has been shown in humans, yeast, and other taxa (13), including regions of *Drosophila pseudoobscura* (14). The reliability of known recombinational distances may also be problematic: it is sometimes difficult to disentangle demographic and selective effects from recombination estimates derived from linkage disequilibrium such as those found in human studies (15–17), and variations in local or genome-wide recombination rates exist within species (18–20). Available genetic distance estimates such as those from *D. melanogaster* (21) often represent a mosaic of uncontrolled crosses from various genotypes in different environments that are major determinants of cross-over rates themselves (22). Because the cross-over rate in different parts of the genome was studied in different crosses and varying conditions, perhaps only extreme differences in the cross-over rate between parts of the genome may have been accurately captured.

To address the discord among studies comparing recombination rate to diversity and divergence, we investigate the association of fine-scale cross-over rate with nucleotide diversity within species and divergence between species in the *D.*

Author contributions: M.A.F.N. designed research; R.J.K., S.M.B., and C.L.F. performed research; R.J.K. contributed new reagents/analytic tools; R.J.K. and M.A.F.N. analyzed data; and R.J.K. and M.A.F.N. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. S.W.S. is a guest editor invited by the Editorial Board.

Data deposition: The Flagstaff 1993 strain of *D. pseudoobscura* was sent to the Tucson *Drosophila* Stock Center and is maintained as stock no. 14011-0121.151. Our 454 sequence reads were submitted to National Center for Biotechnology Information Short Read Archive (SRA) under accession no. SRA000268. Standard Sanger sequences have been deposited in the EMBL/GenBank databases, with accession numbers EU479559–EU479633.

‡To whom correspondence should be addressed. E-mail: noor@duke.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0801848105/DCSupplemental.

© 2008 by The National Academy of Sciences of the USA

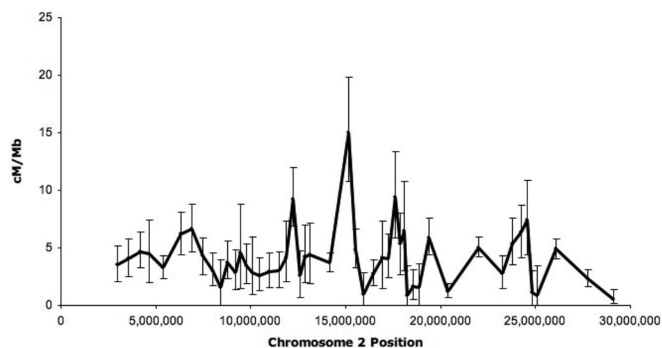

Fig. 1. Cross-over rate (in centimorgans per megabase) at various positions along the assembled *D. pseudoobscura* second-chromosome sequence, with bars indicating the 95% confidence interval of the estimated cross-over rate.

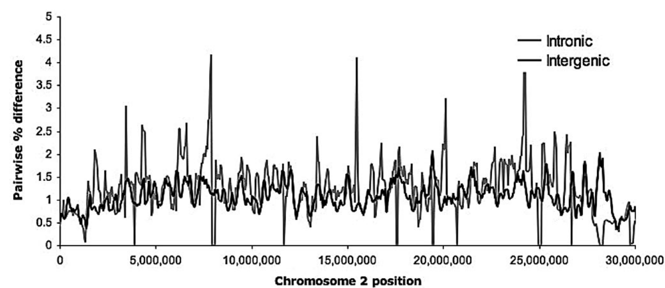

Fig. 2. Pairwise sequence difference between the genome strain of *D. pseudoobscura* and the Flagstaff 1993 line at various positions along the assembled *D. pseudoobscura* second chromosome. The gray line indicates differences at intron positions; the black line indicates differences at intergenic regions.

pseudoobscura species group. *D. pseudoobscura* has been studied extensively for patterns of nucleotide diversity and divergence (23–27), and genome sequence assemblies exist for both it and its sister species, *Drosophila persimilis* (28). We used illumina array microbeads to genotype a large F_2 backcross between two *D. pseudoobscura* strains across a chromosome, and we identified local variations in cross-over rate at a fine genomic scale in a single, controlled cross. We also generated and aligned, at low-coverage, 454 genomic reads of a second strain of *D. pseudoobscura* onto the published genome assembly (29), providing us with a genomic portrait of intraspecific nucleotide diversity. To validate patterns of nucleotide diversity and divergence with respect to recombination rate on a different part of the genome, we sequenced multiple loci by using standard technologies on chromosome XL in 10 isolates of *D. pseudoobscura* and the outgroup species, *Drosophila miranda*, over regions surveyed for cross-over rate variation (14).

Results

Cross-Over Rate Variation. We assayed the cross-over rate in 1,294 progeny from an F_2 backcross between two *D. pseudoobscura* inbred lines by using illumina microbead genotyping arrays. Because this is a new technology, we first evaluated the error rate for the genotyping. We found that cross-overs are not observed within 2 megabases (Mb) of each other in *D. pseudoobscura* (C.L.F. and M.A.F.N., unpublished data), resulting from "cross-over interference." Because of the high density of markers used in this work along the second chromosome, we can thus test for a likely type of error: we should not observe instances of one genotype flanked on both sides by an alternative genotype. This test is conservative because it ignores the possibility of biologically real gene conversion generating this observation. We found only 49 instances of such a pattern from >62,000 genotypes on this chromosome, indicating an error rate of this type of <0.1%.

Across the second chromosome intervals surveyed, we observed 0–4 cross-overs per individual, with the mode at one cross-over [see supporting information (SI) Fig. S1]. From these multilocus genotypes, we calculated a cross-over rate for each window and observed statistically significant (bootstrap two-tailed $P < 0.005$) variation; see Fig. 1) ranging from 0.88 to 15 centimorgans per Mb. Consistent with previous work (14), similar variation was also observed along the X chromosome arms, ranging from 0.91 to 22 centimorgans per Mb. In addition, we found significant associations between cross-over rate and a variety of motifs and features examined, including two motifs described to correlate with recombination rate variation, CCCCACCCC (r = 0.487, P = 0.0004) and CCTCCCT (r = 0.477, P = 0.0006), but no detectable association with transposable elements or GC content (data not shown).

Nucleotide Diversity Within Species. By comparing the generated 454 sequences of a selected strain against the published genome sequence assembly of *D. pseudoobscura*, we can estimate patterns of nucleotide diversity along chromosome arms. Fig. 2 depicts the average pairwise difference between the two sequences for introns and intergenic regions along the second chromosome. Intergenic pairwise differences ranged from 0.18 to 2.1% (mean = 1.1%), whereas intronic pairwise differences ranged from 0.0 to 4.2% (mean = 1.2%). The greater variance across windows of intron data may reflect the smaller amounts of sequence alignment assayed per window (5,704 to 48,147 bp vs. 45 to 25,636 bp, with some windows bearing no intronic sequence alignment). Nonetheless, the two pairwise difference measures were strongly correlated, even if only nonoverlapping windows were examined (n = 149, r = 0.444, $P < 0.0001$). This correlation of fine-scale diversity measures held even if the 5-Mb windows at the centromeric and telomeric ends were excluded (n = 99, r = 0.364, P = 0.0002).

We identified an association between diversity and fine-scale cross-over rate variation when examining windows 500 kb or smaller (intergenic: n = 25, r = 0.588, P = 0.0020, see Fig. 3*A*; intronic: n = 24, r = 0.488, P = 0.0156). Importantly, none of these windows was within 4 Mb of the centromeric or telomeric ends of the assembled sequence, so this pattern is not affected by global patterning of recombination. We tested whether cross-over rate estimated from fine-scale windows (500 kb or less) vs. broad-scale windows (2 Mb) was a better predictor of nucleotide diversity, and we found that the association was much stronger in the former (intergenic: n = 25, r_{fine} = 0.588 P_{fine} = 0.0020, r_{broad} = 0.205 P_{broad} = 0.3260; intronic: n = 24, r_{fine} = 0.488 P_{fine} = 0.0156, r_{broad} = 0.045 P_{broad} = 0.8322). This much stronger association of diversity with cross-over rate estimated from fine-scale windows did not change when we performed partial correlations with or without GC content as an additional covariate. Fine-scale cross-over rate predicted sequence diversity within *D. pseudoobscura* far better than broad-scale cross-over rate.

Sequence Differences Between *D. pseudoobscura* and *D. persimilis*. Fig. S2 depicts the average pairwise difference between the published assemblies of *D. pseudoobscura* and *D. persimilis* for introns and intergenic regions along the second chromosome. Again, these two measures were highly correlated (n = 151, r = 0.574, $P < 0.0001$). Sequence differences between species (intergenic mean = 2.5%, intronic mean = 2.4%) averaged more than twice the diversity we observed within species. Nonetheless, diversity within species and interspecies difference were strongly

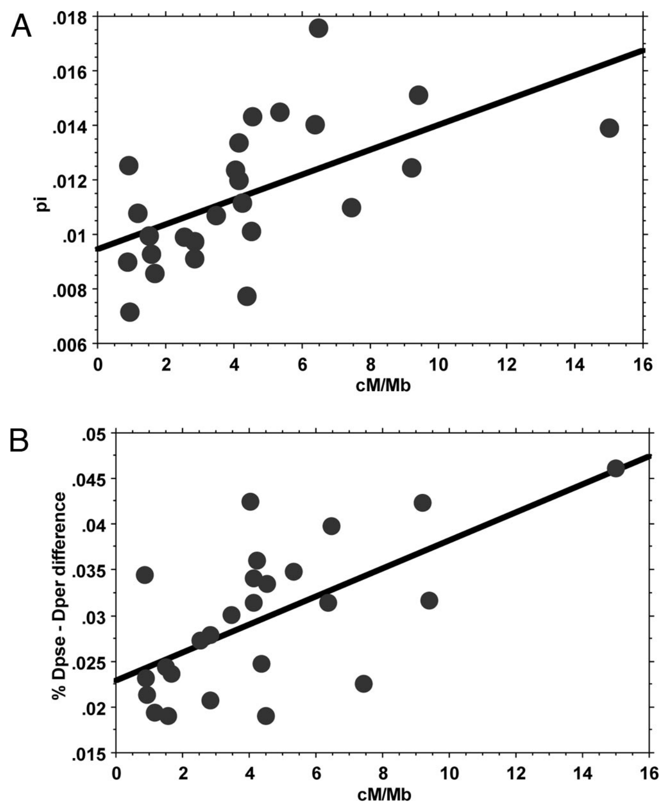

Fig. 3. Association between cross-over rate (in centimorgans per megabase) for second chromosome windows <500 kb in size. (*A*) Pairwise sequence difference between the genome and Flagstaff 1993 strains of *D. pseudoobscura*. (*B*) Pairwise sequence difference between the *D. pseudoobscura* and *D. persimilis* genome sequence assemblies.

correlated (nonoverlapping 200-kb windows; intergenic: $n = 153$, $r = 0.538$, $P < 0.0001$; intronic: $n = 149$, $r = 0.430$, $P < 0.0001$). For the remainder of the analyses, we present only intergenic sequences, although results are qualitatively similar with the available intron sequences.

Sequence differences between the two species were correlated with fine-scale cross-over rate variation (Fig. 3*B*: $n = 25$, $r = 0.635$, $P = 0.0006$), perhaps even more strongly than sequence diversity within species. However, this stronger association of interspecies difference may be driven in part by the greater depth of sequencing of *D. persimilis* relative to our low-coverage 454 sequence of a second *D. pseudoobscura* strain. Hence, we limited the *D. persimilis* dataset to just those bases for which we also had an aligned *D. pseudoobscura* 454 nucleotide filtered against possible alignment or sequencing errors (see *Materials and Methods*). The relationship to fine-scale cross-over rate was essentially unchanged ($n = 25$, $r = 0.607$, $P = 0.0013$).

Although the observed differences between species were more than twice those found within species, the association of recombination to interspecies differences could reflect shared variation between the species either through lineage sorting or introgression. As such, our analysis would not necessarily reflect an association between recombination and "divergence" *per se*. However, we can crudely "correct" the interspecies difference for each window by subtracting from it the intraspecies diversity (30) and then determine whether the residual net difference still correlates with the cross-over rate. The lack of significant population structuring (outside of chromosome 3) in this species (27, 31, 32) suggests that the diversity we observed is typical for that within *D. pseudoobscura*. Further, if we focus our analyses on the 7-Mb region of this chromosome that differs by an inversion that is fixed between these two species [and shows many fixed differences and greater average sequence divergence between them (24, 33)], we still observe a significant association between residual divergence and fine-scale cross-over rate ($n = 10$, $r = 0.730$, $P = 0.0165$).

Additionally, four noncoding regions sampled by Machado *et al.* (24) include >500 bp of aligned sequence fall within windows 500 kb or smaller surveyed for recombination in our study. These loci (2M12, 2M16, 2M17, and 2M18) were surveyed in six to eight strains of *D. pseudoobscura* and one or more strains of *D. miranda*, a species more distantly related than *D. persimilis*. We find that our cross-over rate estimates predict both average pairwise sequence difference within species (π) within *D. pseudoobscura* ($r = 0.979$, $P = 0.0212$) and difference from *D. miranda* ($r = 0.962$, $P = 0.0376$) by using the data obtained in that study.

Ultrafine Scale Analyses of Variation in Relation to Cross-Over Rate. An earlier study (14) documented ultrafine scale variation in cross-over rate along a 2-Mb region of the XL chromosome arm in *D. pseudoobscura*, where half of the windows examined were smaller than 50 kb. To examine the generality of the results described above by using fewer intervals but more rigorous testing, we sequenced 10 strains of *D. pseudoobscura* and one strain of the more distant outgroup species *D. miranda* over seven intergenic windows studied by Cirulli *et al.* (14). All of these windows are within 1 Mb of each other, and five are within 150 kb.

Across these windows, ultrafine-scale cross-over rate was strongly correlated with average pairwise sequence difference within species (π): $r = 0.902$, $P = 0.0055$ (Fig. S3). *D. miranda* differs from *D. pseudoobscura* by a chromosomal inversion disrupting this region (34), yet despite potential effects of this rearrangement on patterns of recombination, average divergence to *D. miranda* was still correlated with ultrafine-scale cross-over rate in *D. pseudoobscura*: $r = 0.771$, $P = 0.0426$ (Fig. S4).

Discussion

We applied two genomic tools to study the association of recombination rate and nucleotide variation in *D. pseudoobscura*: low-coverage next-generation sequencing technology and high-throughput genotyping via microbead arrays. We detected fine-scale structuring across the second chromosome of this species, constituting ≈20% of the assembled genome, in cross-over rate, nucleotide diversity, and interspecies sequence difference from *D. persimilis*. Fine-scale structuring was correlated across all three of these measures, in contrast to similar studies of the *D. melanogaster* group but consistent with studies on humans. We also identified sequence motifs whose abundance was correlated with the recombination rate variation we observed. We confirmed the association of recombination rate with diversity and interspecies sequence difference in a more detailed but smaller-scale examination of chromosome XL over multiple *D. pseudoobscura* strains and the outgroup species *D. miranda*. The correlation between cross-over rate and interspecies sequence difference in particular is predicted if crossing over itself is mutagenic.

The effect of scale is very apparent in our results, and it explains some contradictory findings. Earlier studies of this same chromosome in *D. pseudoobscura* failed to observe a significant association between recombination and nucleotide diversity (35) and suggested that "rates of recombination are quite uniform" (36). In this work, when we estimated cross-over rates from broad-scale, 2-Mb windows, we observed very little variation in cross-over rate (range 3.0–5.0 centimorgans per Mb) and no significant association with nucleotide diversity. However, this apparent homogeneity masks evolutionarily relevant fine-scale

EVOLUTION

heterogeneity. Studying such parameters at larger, imprecise scales may also explain why studies of different species have come to different conclusions with respect to whether recombination is associated with patterns of nucleotide diversity or divergence.

Reconciling a Possible Association of Recombination to Divergence in *Drosophila*. Begun *et al.* (2) also recently correlated fine-scale recombination rates along the X chromosome of *D. melanogaster* to nucleotide variation in *D. simulans* (Spearman's $\rho = 0.45$, $P = 8.5 \times 10^{-8}$) but found that the relationship to interspecies nucleotide divergence was weak (Spearman's $\rho = 0.17$, $P = 0.03$). Although these associations are certainly real, we suggest that they result from large-scale recombinational variation rather than fine-scale variation and that they bear reexamination. If only the very-low-recombination telomeric region of the X chromosome (37) is excluded from their analysis (18 of 130 datapoints at the telomeric end, all points with cross-over rate <1 centimorgans per Mb), the associations of fine-scale recombination rate to variation and divergence are both nonsignificant. This negative result likely reflects some noise in the estimates of fine-scale recombination rates used: these rates were based on a conglomeration of results from various *D. melanogaster* studies done on different laboratory strains and in different laboratory conditions. This conglomeration is problematic because different intervals were not surveyed under the same genetic and environmental conditions. Consistent with our hypothesis, the authors noted some windows bearing "negative" estimates of centimorgans per kilobase, indicative of noise associated with pooling disparate data. Hence, these recombination estimates perhaps primarily capture extreme differences in recombination rate (such as the comparatively low recombination at the telomeric region) rather than more subtle ones.

We cannot completely exclude the possibility that the association we detected between recombination and interspecies sequence difference the contribution of differences in shared ancestral polymorphism, rather than postspeciation divergence *per se*. However, the continued strong associations even after subtracting the nucleotide diversity within species (30, 38) and focusing specifically on regions separated by a fixed chromosomal inversion reduces the likelihood of this explanation somewhat. Relatedly, Machado *et al.* (24) surveyed multiple strains of both species at six loci within this inverted region: all six loci had between 5 and 21 fixed differences each, and only one locus bore any shared polymorphisms, resulting in an overall frequency of shared polymorphism not inconsistent with recurrent mutation. The inversion is also strongly associated with hybrid sterility in laboratory crosses (39), so its contents cannot have readily introgressed between the species. Finally, the results from our analyses using seven loci on chromosome XL and four loci on chromosome 2 on the more distant species, *D. miranda*, further support a possible relationship between cross-over rate and divergence. Hence, based on the data available, we tentatively conclude that rates of crossing over are associated with levels of nucleotide divergence between *D. pseudoobscura* and *D. persimilis*.

Mechanistic vs. Selective Explanations for the Association of Recombination and Genetic Variation. In contrast to broader-scale studies in the *D. melanogaster* species group, studies on human variation have suggested that recombination exerts a local and direct influence on diversity, and one study came to the controversial conclusion that "there is no need to invoke indirect effects of natural selection (such as hitchhiking and background selection) to explain the observed correlations" (40). The association we observe between cross-over rate and divergence is also consistent with mechanistic, rather than selective, hypotheses connecting the two processes. Some lines of evidence support the hypothesis that meiotic recombination may be mutagenic. Mutation rates are higher in meiosis than mitosis in yeast (41), and more recent studies demonstrated that misincorporations occurring during DNA synthesis accompanying mitotic recombinational repair of double-strand breaks (DSBs) elevate the local mutation rate by 100-fold relative to S phase replication (42).

Nonetheless, application across species of a "mutagenic cross-over" hypothesis fails to predict an observation well established in *Drosophila* molecular evolutionary genetics: nucleotide sequence divergence between *D. melanogaster* and *D. simulans* in regions of severely restricted crossing over is very similar to that in regions of normal crossing over (see analyses and review in ref. 43). This is also true for divergence between *D. pseudoobscura* and *D. miranda* on the nonrecombining dot chromosome relative to regions of high recombination (25). In contrast to regions of heterogeneous cross-over frequencies, characterizing these particular areas as regions of low recombination is almost certainly reliable. These observations suggest that a simple mutagenic cross-over hypothesis for explaining its association to divergence is wrong in *Drosophila*, or at least insufficient because regions of low recombination possess a substantive number of diverged sites compared with all other genomic regions. Note that this present work was confined to regions of normal to high recombination; telomeric and centromeric ends were not surveyed.

A minor yet sensible modification may allow a mechanistic hypothesis to explain the lack of association between cross-over rate and divergence in the *D. melanogaster* subgroup and perhaps across species in general. Cross-overs are initiated by DSBs in DNA. Such DSBs can be repaired through the formation of a double Holliday junction that can resolve into a cross-over. That said, alternative means for repairing DSBs do not result in cross-overs (including alternate resolutions of double Holliday junctions or the distinct process of synthesis-dependent strand annealing) and often instead form gene conversion tracts. In yeast, two recent studies have shown "substantial DSB activity in pericentromeric regions, in which cross-over formation is largely absent" (44, 45). In *D. melanogaster*, several studies have similarly shown evidence for extensive gene conversion in regions of severely restricted crossing over (46–48), consistent with DSBs occurring in these regions but being preferentially repaired by non-cross-over means. Also consistent with this hypothesis, mutations in some genes that drive meiotic cross-over formation cause significant increases in non-cross-over gene conversion (49).

We propose a hypothesis that, rather than crossing-over itself being mutagenic, DNA DSB repair more generally, including mechanisms resulting in gene conversion or crossing over, predisposes regions of the genome to higher mutation rates. Misincorporation during synthesis in multiple kinds of DSB repair could produce such mutations. The distribution of cross-overs in *Drosophila* recombinational linkage maps may correlate with the distribution of DSBs across much of the genome but fail to correlate with DSBs in regions where the cross-over rate is severely reduced, including centromeric regions, the tip of the X chromosome, and the nonrecombining dot chromosome. This hypothesis is consistent with the interpretation of early recombination nodules in female meiosis as precursors of both types of recombination event and the fact that these are as frequent in regions where crossing over is suppressed as in regions with normal frequencies of crossing over (50).

However, although this modified mechanistic explanation may contribute to the association of cross-over rate and nucleotide diversity within species, particularly in regions of moderate or high recombination, we emphasize that it cannot explain the pattern observed in regions of restricted recombination. If this explanation were sufficient, then regions of severely reduced cross-over rate but substantial gene conversion would show normal levels of nucleotide sequence diversity, which they clearly do not. Hence, selective explanations, such as sweeps of rare positively selected alleles or background selection against dele-

terious mutations, must also contribute to the association of cross-over rate and diversity in regions of low recombination.

We hypothesize a two-phase model for the association of recombination rate to diversity. In regions of very low recombination, the association may be driven primarily by natural selection. Then, the residual association in regions of moderate or high recombination may be primarily mechanistic, perhaps via mutagenic properties of DSB repair in general. If supported with direct empirical data, this hypothesis reconciles the results from studies of the *D. melanogaster* species group with those from primates and mechanistic studies of the process of recombination.

This research has generated a detailed linkage map of one chromosome of *D. pseudoobscura* and a low-coverage genome sequence of a new strain, both of which will be useful for other investigators to test a variety of hypotheses in this model system. Using these data, we identified a strong correlation within species between fine-scale cross-over rate and nucleotide diversity, contrasting the conclusions of previous studies of the same chromosome in this species that used broad-scale cross-over rate. We also identify a relationship between cross-over rate and sequence difference between species (at two phylogenetic distances), and based in part on this result, we posit a hypothesis to explain that relationship. New empirical data on the genetic architecture of DNA breakage and repair, the distribution of double-strand breaks, the association of various nucleotide motifs, and the mechanism of DNA repair will have a direct bearing on understanding selective vs. neutral processes on a genome-wide scale.

Materials and Methods

454 Genome Sequencing and Read Processing. The Flagstaff 1993 line of *D. pseudoobscura* was selected for sequencing. We isolated total genomic DNA from 25 adult females by using the PureGene DNA isolation protocol. The purified DNA was then submitted to 454 Life Sciences for sequencing on a 454 FLX sequencer. 262,225 reads of an average length of 252.0 bp per read were generated. BLAT version 32 was used to align reads to the published *D. pseudoobscura* r2.0 genome assembly. Only reads that aligned uniquely to the *D. pseudoobscura* assembly with a BlastZ score of at least 3000 were used. The confidence of aligned base pairs from each 454 read was first assessed from Phred-equivalent quality scores generated by the instrument. A minimum score of 10 was required. Aligned sites were checked for consistency among multiple reads: quality scores were added if base pairs were the same, whereas sites were expunged if base pairs did not match. Three flanking base pairs on either side of each nucleotide were also checked for the presence of gaps. This reduces the impact of homopolymer tracts that may be artifacts from the pyrosequencing technology. Finally, each 7-bp window was filtered to have two or fewer diverged sites by using *D. persimilis* genome-to-genome alignments (see next paragraph). This conservative filtering method (a single unique genomic hit, a minimal quality score, multiple read consistency, no flanking gaps, and a minimum divergence within a sliding window) allowed us to make base calls of reliable quality.

Of the 8,383,880 bp from 454 reads that uniquely aligned to *D. pseudoobscura* chromosome 2 with certainty, 5,735,931 bp could be mapped to a precise position, resulting in a coverage of 18.6%. After quality filters were applied, a total of 5,568,486 bp (18.0% coverage) were aligned with confidence.

***D. pseudoobscura* Genomic Annotations.** Sequence annotation such as intergenic regions, codon positions 1–3, and intronic regions were derived from r2.0 *D. pseudoobscura* annotations from FlyBase (FB2007_04). Only sites that contain a single ontological type were used.

To search for sequence property correlates of recombination rate, sequences were run through RepeatMasker version 3.1.6 by using the ''-qq'' and ''-species *drosophila*'' parameters. Two known recombination hotspot motifs, CCTCCCT and CCCCACCCC (51), were searched in all sequences by using 1-bp sliding windows until a match is found. The next sliding window commences immediately after the identified motif. Statistical significance was evaluated through resampling following Cirulli *et al.* (14).

Diverged sites between *D. pseudoobscura* (r2.0) and *D. persimilis* (CAF1) assemblies were identified by using the genome-to-genome alignments from the Mercator/Mavid pipeline (Lio Pachtor's laboratory, *Drosophila* AAA wiki site).

Generation of Local Estimates of Recombination Rate. A large F_2 backcross was generated by crossing females from the MV2-25 inbred strain to males from the Flagstaff 1993 strain and crossing the F_1 females to MV2-25 males. All crosses were carried out at 20°C on standard sugar-yeast-agar medium. Adult flies were isolated and frozen at −80°C, and their DNA was purified by using the PureGene DNA isolation protocol. DNAs were plated into 96-well format and submitted to the Duke IGSP Genotyping Facility for SNP genotyping by using the illumina BeadXpress Reader system.

Comparing the 454 sequences from Flagstaff 1993 to the published MV2-25 line genome sequence, we also identified a panel of 100-bp sequences with a single SNP differentiating the two lines. These sequences were used to generate the genotyping microbeads by illumina. A handful of markers were excluded because they displayed very strong deviations from expected 50–50 allelic ratios among the backcross progeny indicating either segregating variation within the lines or some type of failure in design or accuracy. Of the final set used for genotyping, 49 markers were separated by distances of 130 kb to 1.7 Mb (median 466 kb), and one pair of markers was separated by 9 kb.

Genotypes were provided directly by the genotyping facility, and the recombination rate was assayed as the number of recombinant segments over the total number of segments evaluated (same as the number of individuals = 1,294). No correction was made for double cross-overs given our marker density because we found that cross-overs are not observed within 2 Mb of each other in *D. pseudoobscura* (C.L.F. and M.A.F.N., unpublished data), resulting from cross-over interference.

Sequencing and Analysis of DNA Polymorphisms. We identified seven 500- to 700-bp XL-chromosome noncoding segments within the windows surveyed for fine-scale cross-over rate heterogeneity (14). Primers were designed, and patterns of polymorphism and divergence were surveyed by using standard PCR and bidirectional BigDye cycle sequencing in 10 inbred isofemale lines of *D. pseudoobscura* and one inbred line of *D. miranda*. Sequences were aligned manually, and polymorphism and divergence statistics were identified by using DNAsp (52).

ACKNOWLEDGMENTS. We thank R. Kliman, T. Mitchell-Olds, L. Stevison, and two anonymous reviewers for comments on the manuscript. We also thank the Computational Biology Group at the Harvard FAS Center for Systems Biology; W. M. Gelbart, D. L. Hartl, R. C. Lewontin, J. Collett, T. Petes, J. Sekelsky, and members of the Hartl laboratory for constructive feedback. K. Shianna, A. Somerville, and K. R. Kommaraju provided technical assistance. This work was supported by National Science Foundation Grants 0509780 and 0715484 and National Institutes of Health Grant GM076051.

1. Begun DJ, Aquadro CF (1992) Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster. Nature* 356:519–520.
2. Begun DJ, *et al.* (2007) Population genomics: Whole-genome analysis of polymorphism and divergence in *Drosophila simulans. PLoS Biol* 5:e310.
3. Andolfatto P (2001) Adaptive hitchhiking effects on genome variability. *Curr Opin Genet Dev* 11:635–641.
4. Hudson RR (1994) How can the low levels of DNA sequence variation in regions of the *Drosophila* genome with low recombination rates be explained? *Proc Natl Acad Sci USA* 91:6815–6818.
5. Nachman MW (2002) Variation in recombination rate across the genome: Evidence and implications. *Curr Opin Genet Dev* 12:657–663.
6. MaynardSmith J, Haigh J (1974) The hitch-hiking effect of a favourable gene. *Genet Res* 23:23–35.
7. Charlesworth B, Morgan MT, Charlesworth D (1993) The effect of deleterious mutations on neutral molecular variation. *Genetics* 134:1289–1303.
8. Roselius K, Stephan W, Stadler T (2005) The relationship of nucleotide polymorphism, recombination rate, and selection in wild tomato species. *Genetics* 171:753–763.
9. Stephan W, Langley CH (1998) DNA polymorphism in *Lycopersicon* and crossing-over per physical length. *Genetics* 150:1585–1603.
10. Hellmann I, Ebersberger I, Ptak SE, Paabo S, Przeworski M (2003) A neutral explanation for the correlation of diversity with recombination rates in humans. *Am J Hum Genet* 72:1527–1535.
11. Hellmann I, *et al.* (2005) Why do human diversity levels vary at a megabase scale? *Genome Res* 15:1222–1231.
12. Tenaillon MI, U'Ren J, Tenaillon O, Gaut BS (2004) Selection versus demography: A multilocus investigation of the domestication process in maize. *Mol Biol Evol* 21:1214–1225.
13. Lichten M, Goldman AS (1995) Meiotic recombination hotspots. *Annu Rev Genet* 29:423–444.
14. Cirulli ET, Kliman RM, Noor MAF (2007) Fine-scale cross-over rate heterogeneity in *Drosophila pseudoobscura. J Mol Evol* 64:129–135.
15. Reed FA, Tishkoff SA (2006) Positive selection can create false hot spots of recombination. *Genetics* 172:2011–2014.
16. Wang N, Akey JM, Zhang K, Chakraborty R, Jin L (2002) Distribution of recombination cross-overs and the origin of haplotype blocks: The interplay of population history, recombination, and mutation. *Am J Hum Genet* 71:1227–1234.

17. Tenaillon MI, Austerlitz F, Tenaillon O (2008) Apparent mutational hotspots and long distance linkage disequilibrium resulting from a bottleneck. *J Evol Biol* 21:541–550.
18. Brooks LD, Marks RW (1986) The organization of genetic variation for recombination in *Drosophila melanogaster*. *Genetics* 114:525–547.
19. Coop G, Wen X, Ober C, Pritchard JK, Przeworski M (2008) High-resolution mapping of cross-overs reveals extensive variation in fine-scale recombination patterns among humans. *Science* 319:1395–1398.
20. Kong A, *et al.* (2008) Sequence variants in the RNF212 gene associate with genome-wide recombination rate. *Science* 319:1398–1401.
21. Lindsley DL, Zimm GG (1992) *The Genome of Drosophila melanogaster* (Academic, London).
22. Ashburner, M (1989) *Drosophila*: *A Laboratory Handbook* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).
23. Aquadro CF, Weaver AL, Schaeffer SW, Anderson WW (1991) Molecular evolution of inversions in *Drosophila pseudoobscura*: The amylase gene region. *Proc Natl Acad Sci USA* 88:305–309.
24. Machado CA, Haselkorn TS, Noor MAF (2007) Evaluation of the genomic extent of effects of fixed inversion differences on intraspecific variation and interspecific gene flow in *Drosophila pseudoobscura* and *D. persimilis. Genetics* 175:1289–1306.
25. Machado CA, Hey J (2003) The causes of phylogenetic conflict in a classic *Drosophila* species group. *Proc R Soc London Ser B* 270:1193–1202.
26. Schaeffer SW, Anderson WW (2005) Mechanisms of genetic exchange within the chromosomal inversions of *Drosophila pseudoobscura*. *Genetics* 171:1729–1739.
27. Schaeffer SW, Miller EL (1992) Estimates of gene flow in *Drosophila pseudoboscura* determined from nucleotide sequence analysis of the alcohol dehydrogenase region. *Genetics* 132:471–480.
28. Clark AG, *et al.* (2007) Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* 450:203–218.
29. Richards S, *et al.* (2005) Comparative genome sequencing of *Drosophila pseudoobscura*: Chromosomal, gene, and *cis*-element evolution. *Genome Res* 15:1–18.
30. Nei M (1987) *Molecular Evolutionary Genetics* (Columbia Univ Press, New York).
31. Noor MAF, Schug MD, Aquadro CF (2000) Microsatellite variation in populations of *Drosophila pseudoobscura* and *Drosophila persimilis. Genet Res* 75:25–35.
32. 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.
33. Noor MA, Garfield DA, Schaeffer SW, Machado CA (2007) Divergence Between the *Drosophila pseudoobscura* and *D. persimilis* genome sequences in relation to chromosomal inversions. *Genetics* 177:1417–1428.
34. Anderson WW, Ayala FJ, Michod RE (1977) Chromosomal and allozymic diagnosis of three species of *Drosophila. J Hered* 68:71–74.
35. Ortiz-Barrientos D, Chang AS, Noor MAF (2006) A recombinational portrait of the *Drosophila pseudoobscura* genome. *Genet Res* 87:23–31.
36. Hamblin MT, Aquadro CF (1999) DNA sequence variation and the recombinational landscape in *Drosophila pseudoobscura*: A study of the second chromosome. *Genetics* 153:859–869.
37. Aguade M, Miyashita N, Langley CH (1989) Reduced variation in the Yellow–Achaete–Scute region in natural populations of *Drosophila melanogaster*. *Genetics* 122:607–615.
38. Reed FA, Akey JM, Aquadro CF (2005) Fitting background-selection predictions to levels of nucleotide variation and divergence along the human autosomes. *Genome Res* 15:1211–1221.
39. Noor MAF, *et al.* (2001) The genetics of reproductive isolation and the potential for gene exchange between *Drosophila pseudoobscura* and *D. persimilis* via backcross hybrid males. *Evolution (Lawrence*, *Kans)* 55:512–521.
40. Spencer CCA, *et al.* (2006) The influence of recombination on human genetic diversity. *PLoS Genet* 2:e148.
41. Magni GE, Von Borstel RC (1962) Different rates of spontaneous mutation during mitosis and meiosis in yeast. *Genetics* 47:1097–1108.
42. Strathern JN, Shafer BK, McGill CB (1995) DNA synthesis errors associated with double-strand-break repair. *Genetics* 140:965–972.
43. Noor MAF, Kliman RM (2003) Variability on the dot chromosome in the *Drosophila simulans* clade. *Genetica* 118:51–58.
44. Blitzblau HG, Bell GW, Rodriguez J, Bell SP, Hochwagen A (2007) Mapping of meiotic single-stranded DNA reveals double-strand-break hotspots near centromeres and telomeres. *Curr Biol* 17:2003–2012.
45. Buhler C, Borde V, Lichten M (2007) Mapping meiotic single-strand DNA reveals a new landscape of DNA double-strand breaks in *Saccharomyces cerevisiae*. *PLoS Biol* 5:e324.
46. Andolfatto P, Wall JD (2003) Linkage disequilibrium patterns across a recombination gradient in African *Drosophila* melanogaster. *Genetics* 165:1289–1305.
47. Jensen MA, Charlesworth B, Kreitman M (2002) Patterns of genetic variation at a chromosome 4 locus of *Drosophila melanogaster* and *D. simulans. Genetics* 160:493–507.
48. Langley CH, Lazzaro BP, Phillips W, Heikkinen E, Braverman JM (2000) Linkage disequilibria and the site frequency spectra in the su(s) and su(w(a)) regions of the *Drosophila* melanogaster X chromosome. *Genetics* 156:1837–1852.
49. Blanton HL, *et al.* (2005) REC, *Drosophila* MCM8, drives formation of meiotic crossovers. *PLoS Genet* 1:e40.
50. Zickler D, Kleckner N (1999) Meiotic chromosomes: integrating structure and function. *Annu Rev Genet* 33:603–754.
51. 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.
52. Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19:2496–2497.