Genet. Res., Camb. (2006), **87**, *pp.* 23–31. © 2006 Cambridge University Press
doi:10.1017/S0016672306007932 Printed in the United Kingdom

A recombinational portrait of the *Drosophila pseudoobscura* genome

DANIEL ORTIZ-BARRIENTOS[1], AUDREY S. CHANG[1,2] AND MOHAMED A. F. NOOR[1,2,*]

[1] *Department of Biological Sciences, Louisiana State University, Baton Rouge, LA 70803, USA*
[2] *DCMB Group/Biology Department, Box 91000, Duke University, Durham, NC 27708, USA*

(*Received 9 September 2005 and in revised form 2 November 2005*)

Summary

Drosophila pseudoobscura has been intensively studied by evolutionary biologists for over 70 years. The recent publication of the genome sequence not only permits studies of comparative genomics with other dipterans but also opens the door to identifying genes associated with adaptive traits or speciation or testing for the signature of natural selection across the genome. Information on regional rates of recombination, localization of inversion breakpoints distinguishing it from its sibling species *D. persimilis*, and known polymorphic markers may be useful in such studies. Here, we present a molecular linkage map of four of the five major chromosome arms of *D. pseudoobscura*. In doing so, we order and orient several sequence contig groups, localize the inversion breakpoints on chromosome 2 to intervals of 200 kilobases, and identify one error in the published sequence assembly. Our results show that regional recombination rates in *D. pseudoobscura* are much higher than in *D. melanogaster* and significantly higher even than in *D. persimilis*. Furthermore, we detect a non-significant positive correlation between recombination rate and published DNA sequence variation. Finally, the online Appendix presents 200 primer sequence pairs for molecular markers that can be used for mapping of quantitative trait loci, of which 125 are known to be polymorphic within or between species.

1. Introduction

The fruit fly *Drosophila pseudoobscura* is one of the most intensively studied model systems for understanding adaptation and speciation. This species is perhaps best known for its rich inversion polymorphism on the third chromosome. Third chromosome arrangements bear striking longitudinal clines across the species range and are associated with life-history traits such as fecundity (e.g. Anderson *et al.*, 1991). Cline frequencies are maintained despite apparently extensive gene flow among populations (Prakash *et al.*, 1969; Schaeffer & Miller, 1992; Noor *et al.*, 2000), suggesting that natural selection operates on these inversions. This species has also been studied extensively with respect to the genetic basis of traits associated with reproductive isolation, such as hybrid sterility and sexual isolation (Dobzhansky, 1936; Tan, 1946; Orr, 1987; Noor *et al.*, 2001*a*, *b*; Orr & Irving, 2001).

Much of the progress in studying adaptation and speciation in *D. pseudoobscura* came from genetic mapping studies that employed either morphological mutant markers or a moderate number of microsatellite markers. The recent availability of the *D. pseudoobscura* genome sequence (Richards *et al.*, 2005) provides an opportunity for developing a higher resolution molecular linkage map of the species. Such a linkage map can facilitate a variety of genetic and evolutionary efforts, including high-resolution mapping of quantitative trait loci (QTL) and comparisons of regional recombination rate and variation.

Several recombination maps are currently available for *D. pseudoobscura*. Anderson (1993) provided an early linkage map based on 63 morphological markers and allozymes across the genome, Orr (1995) identified linkage relationships for 23 morphological markers on the X chromosome, and Noor *et al.* (2000)

* Corresponding author. Tel: +1 (919) 6138156. Fax: +1 (919) 6138177. e-mail: noor@duke.edu

https://doi.org/10.1017/S0016672306007932 Published online by Cambridge University Press

created a molecular map using 24 microsatellite markers. Although these maps have advanced our knowledge of the adaptive history of *D. pseudoobscura* and of speciation between *D. pseudoobscura* and its sister species, new genetic maps incorporating many more molecular markers can facilitate high-resolution genome scans to compare with the physical sequence map, potentially allowing mapping of phenotype to single genes.

Here, we provide a dense recombination map and information on regional rates of recombination for four of five *D. pseudoobscura* chromosome arms. We use our linkage data to order several of the unassembled contig groups in this sequenced genome, and to localize some of the inversion breakpoints distinguishing *D. pseudoobscura* and its sister species *D. persimilis*. We also present preliminary information on recombination rates across the second chromosome of *D. persimilis* and compare them with those from *D. pseudoobscura*. Using this data, we test for an association between recombination rate and DNA sequence variation, and refine an earlier estimate of the effective population size of *D. pseudoobscura*. Finally, we present primer sequences for 200 microsatellite or other size markers, of which 125 have been shown to be variable within or between *D. pseudoobscura* and/or *D. persimilis*. In conjunction with the genome sequence of *D. pseudoobscura*, our markers and map will facilitate population genomic analysis, aid in the identification of genes involved in adaptation and speciation via QTL mapping, and ultimately help us to understand the evolutionary forces acting on this species and its close relatives.

2. Materials and methods

(i) *Species, lines and rearing conditions*

Drosophila pseudoobscura and *D. persimilis* are close relatives that hybridize in nature (Dobzhansky, 1973; Powell, 1983), and some studies have suggested that introgression occurs between these species primarily in collinear regions of their genome (Machado *et al*., 2002; Machado & Hey, 2003; Hey & Nielsen, 2004). These species have six chromosome arms: two on the X chromosome (designated XL and XR), three large independent autosomes (numbered 2–4), and a 'dot' fifth chromosome which is not studied here. The species differ by fixed or nearly fixed inversions on the XL, XR and chromosome 2 arms. The third chromosome is highly polymorphic within both species, and they share one abundant arrangement (designated 'Standard').

Lines of *D. pseudoobscura* were established from individuals collected at Mather, California, in 1997 and Flagstaff, Arizona, in 1993, both of which bear the 'Arrowhead' arrangement on the third chromosome.

Lines of *D. persimilis* were established from flies collected in Mt St Helena, California, in 1993 (line MSH1993) and 1997 (line MSH3), both of which bear the 'Standard' arrangement on the third chromosome. All lines were maintained at a constant temperature (20 °C) and humidity (85%) regime in diurnal/nocturnal cycles of 12 h and reared on a mixture of agar, dextrose and yeast.

(ii) *Recombinational maps*

(a) *Pure species linkage maps*

Recombination maps were obtained by estimating recombination fractions in F_2 backcross populations between species-specific isofemale lines. Each parental line was tested for allelic differences in microsatellite loci previously published (Noor *et al*., 2000) or extracted from the *D. pseudoobscura* genome sequence (Richards *et al*., 2005), and scored for differences in amplification size of the microsatellite markers. The *D. pseudoobscura* map was made by backcrossing F_1 females of the cross between two lines to the Flagstaff 1993 line (see Ortíz-Barrientos *et al*., 2004). The *D. persimilis* map was made by backcrossing F_1 females of the cross between two lines to the MSH1993 line. We used polytene chromosome preparations of F_1 progeny to confirm that these strains also bear the same third chromosome arrangement as each other.

For each species, approximately 275 flies were genotyped for such microsatellites (see online Appendix) and a multipoint-linkage approach, as implemented in MapMaker version 3.0 (Lander *et al*., 1987), was used to generate recombination maps for the X, 2 and 4 chromosomes from *D. pseudoobscura*, and the second chromosome of *D. persimilis*. We do not investigate the third chromosome despite its complete assembly because analyses of it would be complicated by the inversion polymorphisms within these species. Recombinational distances are reported in Kosambi centimorgans (cM), and recombination rates in kilobases per centimorgan (kb/cM).

Primer sequences for markers not used in the development of this linkage map are also presented here. Some have been used in other studies by our laboratory, and the online Appendix indicates which are known to show intra- or interspecific variability. Also, where indicated just for confirmation, some markers were scored on a subset of backcross progeny rather than the full 275.

(b) *Backcross hybrid linkage maps*

Recombination maps from two interspecific backcrosses between *D. pseudoobscura* and *D. persimilis*

https://doi.org/10.1017/S0016672306007932 Published online by Cambridge University Press

were also analysed. These backcrosses have been described previously (Noor *et al*., 2001a) and correspond to an F_2 backcross to *D. persimilis* and an F_2 backcross to *D. pseudoobscura*. Recombination maps and units were estimated and reported as above.

(c) *Physical map information*

The recently published and assembled *D. pseudoobscura* genome sequence was used to anchor our recombination map to the physical map. The genome sequence is assembled into groups of contigs whose links were obtained computationally, via hybridization to BAC clones, or by presumed synteny with *D. melanogaster*. For detailed information see Richards *et al*. (2005). We use our recombinational data to predict linkage and orientation of these assembled contig groups, and to confirm or provide evidence against some of the assembly points. Positions in the specific contig groups are based on positions within release 1.0 of the *D. pseudoobscura* sequence, which can be viewed at: http://species.flybase.net/.

Using this information, we then calculate recombination rate in kilobases per centimorgan (kb/cM). Because there are sometimes small gaps between contigs within groups, our procedure will slightly *underestimate* rates of recombination (overestimating kb/cM).

(iii) *DNA sequence polymorphism data and analyses*

Published DNA sequence data from *D. pseudoobscura* were obtained from the literature (e.g. Hamblin & Aquadro, 1999; Machado *et al*., 2002). Nucleotide polymorphism data from non-coding regions were used to test whether DNA sequence polymorphism – measured as pi (π) – GC content and/or codon usage positively correlated with recombination rates. The genes or regions included were *Mlc1*, *Xdh*, *bicoid*, *rh1*, *rh3*, *trop1* (these six from Hamblin & Aquadro, 1999), DPS2001, DPS2002, DPS2003 (these three from Machado *et al*., 2002) and DPS2_1206e (second chromosome pericentromeric region; GenBank accessions DQ186852–DQ186859). Because there were no crossovers in the second chromosome pericentromeric region, we conservatively estimated its recombination rate as though one crossover product was recovered. We performed regressions and other statistical analyses presented here using StatView statistical software (SAS Institute).

3. Results

(i) D. pseudoobscura *assembly of contig groups*

We used our recombination map to order and orient groups of contigs from the incomplete assembly of the *D. pseudoobscura* X and fourth chromosomes (see Tables 1–3 and Appendix). Richards *et al*. (2005) have already connected groups from the second chromosome, and our linkage relationships generally support their assemblage. Below, we describe major findings regarding the sequence assembly provided by our recombinational maps. We refer to low numbered positions within contig groups (as designated within release 1.0 of the *D. pseudoobscura* genome sequence) as the 'base' and high numbered positions as the 'tip'.

Our linkage maps are presented in Tables 1–3. The left arm of the X chromosome (XL) is composed of several contig groups. The tip of XL group 1a appears to be relatively close to the telomere. We linked the middle of group XL1a to the tip of group XL1e. XL group 3a then appears to be closer to the centromere than does the base of XL group 1e. Along the right arm of the X chromosome (XR), the base of XR group 6 appears to be closest to the centromere, and hence close to XL3a. Moving towards the telomere are XR groups 9 and then XR group 8, both with their tips towards the telomere. Furthest telomeric on the XR was XR group 3a. The second chromosome was fully assembled, and our linkage map largely supported this assembly (but see below). Finally, we linked two groups on the fourth chromosome: the base of group 1 links to the tip of group 5.

Table 1. *Linkage map of the* Drosophila pseudoobscura *X chromosome*

Marker	Group	Base position	cM	Total cM	kb/cM
DPSX008	XLgroup1a	7482817	15·9	15·9	51
DPSX055	XLgroup1a	6670913	15·8	31·7	
DPSX046	XLgroup1e	12316097	9·7	41·4	147
DPSX002	XLgroup1e	10886336	13·8	55·2	90
DPSX035	XLgroup1e	9645269	16·9	72·1	184
DPSX022	XLgroup1e	6536519	12·9	85	51
DPSX056	XLgroup1e	5875602	29·1	114·1	159
DPSX031	XLgroup1e	1261285	20·4	134·5	
DPSX023	XLgroup3a	2508535	23·1	157·6	
DPSX036	XRgroup6	4293648	8·2	165·8	186
DPSX043	XRgroup6	5821217	22·4	188·2	
DPSX047	XRgroup9	2756260	17·0	205·2	83
DPSX007	XRgroup9	4162734	13·8	219	35
DPSX048	XRgroup9	4641714	14·8	233·8	
DPSX024	XRgroup8	1150746	10·7	244·5	150
DPSX030	XRgroup8	2746654	20·3	264·8	
DPSX021	XRgroup8	7481819	7·1	271·9	78
DPSX021A1	XRgroup8	6926410	14·0	285·9	135
DPSX037N	XRgroup8	5029631	9·0	294·9	
DPSX052	XRgroup3a	1185942			

Indicated are the marker names, the assembly group numbers, the base positions within that group, recombinational distance (in centimorgans, cM), and estimates of the recombination rate (in kilobases per centimorgan, kb/cM). A horizontal line indicates our suggested position for a misassembly of a contig group.

https://doi.org/10.1017/S0016672306007932 Published online by Cambridge University Press

Table 2. *Linkage map of the* Drosophila pseudoobscura *second chromosome*

Marker	Base position	cM	Total cM	kb/cM
DPS2028	1515879	13·4	13·4	
DPS2014	2938447	[2·2]	[15·6]	[342]
DPS2027	3690753	10·6	26·2	
DPS2017	4751214	16·7	42·9	113
DPS2019	6637997	17·6	60·5	152
DPS2018	9313692	8·7	69·2	112
DPS2026	10287490	11·9	81·1	164
DPS2011	12243695	27·8	108·9	153
DPS2022	16496624	4·0	112·9	220
DPS2021	17375999	10·7	123·6	97
DPS2024	18409605	4·8	128·4	148
DPS2012	19119203	7·0	135·4	89
bcd	19738796	10·2	145·6	183
DPS2031	21600547	33·6	179·2	110
DPS2015	25288491	9·9	189·1	112
gld	26400179	12·4	201·5	249
DPS2016	29486344	0·0	201·5	>2082
DPS2_1206e	30694200			

Indicated are the marker names, the base positions within that group, recombinational distances (in centimorgans, cM) and estimates of the recombination rate (in kilobases per centimorgan, kb/cM). The recombinational estimates of one marker are in brackets because its position disagreed between the sequence assembly and our linkage assembly.

Table 3. *Linkage map of the* Drosophila pseudoobscura *fourth chromosome*

Marker	Group	Base position	cM	Total cM	kb/cM
DPS4033	4group1	5068643	14·0	14·0	36
DPS4032	4group1	36859	4·2	18·2	
DPS4034	4group5	2327441	4·6	22·8	44
DPS4003	4group5	2124594	2·2	25·0	130
DPS4A8	4group5	1838553	3·7	28·7	42
DPS4A4	4group5	1682437	–	–	

Indicated are the marker names, the assembly group numbers, the base positions within that group, recombinational distances (in centimorgans, cM) and estimates of the recombination rate (in kilobases per centimorgan, kb/cM).

We confirmed one discrepancy between our recombination-based assembly of markers and their locations according to the published DNA sequence assembly. On the XR chromosome arm, our linkage analyses suggested that DPSX037N was further from the centromere than DPSX021 (both within XR group 8), whereas the genome assembly obtained the opposite order. First, we confirmed this finding by generating a second set of backcross progeny between these *D. pseudoobscura* lines and genotyping a subset of the same microsatellites. Again, we found the same reversed orientation: there were fewer recombinants between DPSX024 and DPSX021 than between DPSX024 and DPSX037N (see Table 1 for positions). We also sequenced amplifications of DPSX021 and DPSX037N to confirm that our scoring did not result from non-specific amplification. Additionally, in a sample we tested, DPSX061 (XR group 8, position 4,662,793; see Appendix) remained closely linked to DPSX037N (XR group 8, position 5,029,631), though. These findings suggest that XR group 8 has been misassembled in the half towards the tip, beginning somewhere before the location of DPSX061. Based on the linkage data, the misassembly appears to be a simple inversion of half of XR group 8. This same erroneous XR group 8 inversion is still apparent in the version 2.0 assembly of the *D. pseudoobscura* genome.

MapMaker linkage analysis also inferred the opposite order of the published assembly for markers DPS2014 and DPS2027 on the second chromosome. In this case, we did not obtain any supporting data for or against this possible assembly error, so we have not analysed the associated links in the analyses that follow.

In addition to differences between the assembly and the recombinational maps, we found that some published *in situ* hybridization data do not match our findings or the physical assembly. For example, based on the assembly, gene *Hsp82* (group XR6, position 4,030,963) is closer to the centromere than marker DPSX009 (group XR6, base 4,925,036). This order contradicts the order reported by Machado *et al.* (2002), but it agrees with the order inferred by Kovacevic & Schaeffer (2000). Similarly, marker DPS2002, a marker localized inside a region on the second chromosome inverted between *D. pseudoobscura* and *D. persimilis* (Machado *et al.*, 2002), appears to be 1·4 MB outside the inversion on the telomeric side (position 8,068,766: see below).

(ii) *Localization of inversion breakpoints*

By comparison of recombination estimates between markers in *D. persimilis* and reference to the *D. pseudoobscura* sequence assembly, we can localize the breakpoints of inversions distinguishing these two species. We focused most of our efforts on the second chromosome, since its assembly was the most complete. In crosses between *D. persimilis* lines, we observed no recombination between markers DPS2_1109k (*D. pseudoobscura* position 9,511,535) and DPS2_3447c (*D. pseudoobscura* position 17,166,960). Similarly, we observed no recombination between markers DPS2_1109i (9,310,915) and DPS2_138c (16,834,932). However, the two pairs of markers were only loosely linked to each other in *D. persimilis* despite their extremely close proximity in *D. pseudoobscura*. The second chromosome inversion

https://doi.org/10.1017/S0016672306007932 Published online by Cambridge University Press

Table 4. *Comparative linkage maps of the* Drosophila pseudoobscura *and* D. persimilis *second chromosomes*

D. pseudoobscura				*D. persimilis*			
Marker	Base position	cM	kb/cM	kb/cM	cM	Base position	Marker
DPS2014	2938447	10·1	179	267	6·8	2938447	DPS2014
DPS2017	4751214	14·2	133	93	20·3	4751214	DPS2017
DPS2019	6637997	15·2	176	294	9·1	6637997	DPS2019
DPS2018	9313692	8·0			3·0	9310915	DPS2_1109i
DPS2026	10287490	26·9	231	304	19·7	16276347	DPS2_138b
DPS2022	16496624	3·9			3·1	10287490	DPS2026
DPS2021	17375999	9·9	112	252	4·4	17301481	DPS2_3447a
DPS2024	18409605	10·2	130	266	5·0	18409605	DPS2024
bcd	19738796	9·4	198	405	4·6	19738796	*bcd*
DPS2031	21600547	26·5	139	165	22·3	21600547	DPS2031
DPS2015	25288487					25288487	DPS2015

Indicated are the markers genotyped, positions of those markers within the *D. pseudoobscura* genome assembly, centimorgans (cM) between consecutive markers, and estimates of recombination rate (in kilobases per centimorgan, kb/cM).

breakpoints are thus close to base positions 9,400,000 and 17,000,000.

We also attempted to localize the inversions on the XL and XR chromosome arms, but this effort was hampered by the incomplete (and possibly incorrect) assembly of these chromosome arms as well as a general lack of microsatellite length variation between our *D. persimilis* lines on the X chromosome. Nonetheless, we did identify that DPSX037N (XR group 8, position 5,029,631) was 8 cM from DPSX014 (XR group 6, position 6,220,490) in *D. persimilis*. Several markers in XR group 9 were also linked to DPSX014, but on the *opposite* side from DPSX037N. This suggests that the inversion breakpoints are in XR group 6 and XR group 8 within a few megabases of these markers. The possible misassembly of XR group 8 prevents further refinement of this localization.

(iii) *Recombination rates*

(a) D. pseudoobscura

We estimated recombination rates for chromosomes X, 2 and 4 of *D. pseudoobscura*. Tables 1–3 present the distribution of recombination rates at various positions along these chromosomes. Recombination rates average 137 kb/cM across the chromosomes and do not vary dramatically within each chromosome. Mean recombination rates for intervals we examined across all chromosome arms are: XL, 121 kb/cM; XR, 105 kb/cM; second, 148 kb/cM; and fourth, 232 kb/cM. Despite general uniformity of recombination rates across the genome, we found several regions exhibiting notably different recombination rates, assuming the sequence assembly is correct. Most strikingly, the second chromosome pericentromeric region has a very low recombination rate (0/148 recombinants in a 1·2 MB span, hence >1787 kb/cM). The fourth chromosome also contains a region of comparatively low recombination rate (359 kb/cM) that inflated its overall average recombination rate, and the X and fourth chromosomes bear regions with apparently very high recombination rates (~ 50 kb/cM).

(b) *Comparative portrait*

We also compared the recombinational landscape of the fully assembled second chromosome between *D. pseudoobscura*, *D. persimilis* and hybrid F_2 backcrosses between the two species. We use only the 11 markers or regions surveyed in both species to prevent the higher density of markers in *D. pseudoobscura* from artefactually inflating our estimate of its recombination rate (see Table 4). Assuming the same genome size, the average recombination rate of the second chromosome is significantly higher in *D. pseudoobscura* than in *D. persimilis* (Wilcoxon signed rank test, $P=0{\cdot}025$).

Interspecies hybrid recombination rates were similar to each other irrespective of the direction of the backcross (see Fig. 1; Wilcoxon signed rank test, $P=0{\cdot}686$). Hybrid recombinational portraits for six markers along the second chromosome also reflected the presence of a region with dramatically low recombination rate. This region is known to encompass a fixed inverted segment between *D. pseudoobscura* and *D. persimilis* (Tan, 1935). When this region between markers DPS2001 and *bcd* is excluded from analysis, recombinational rates in hybrids are significantly higher than in either parental species (Wilcoxon signed rank test $P=0{\cdot}0165$ for *D. pseudoobscura* and $P=0{\cdot}0171$ for *D. persimilis*).

(iv) *Relation to DNA sequence variation*

We examined the correlation between recombination rates and nucleotide diversity (measured as pi (π)

https://doi.org/10.1017/S0016672306007932 Published online by Cambridge University Press

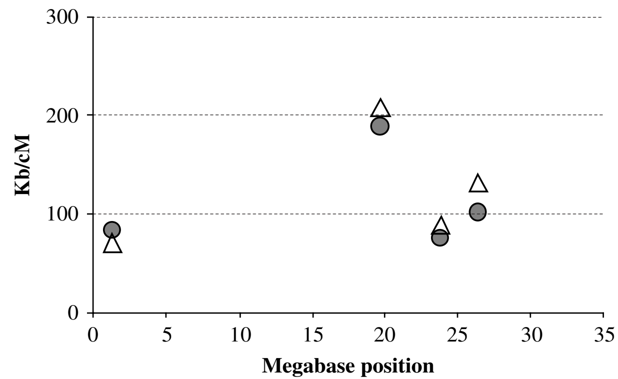

Fig. 1. Regional recombination rates across the second chromosome of *Drosophila pseudoobscura*/*D. persimilis* F_1 hybrid females. Circles indicate offspring from a backcross to *D. pseudoobscura* and triangles indicate offspring from a backcross to *D. persimilis*.

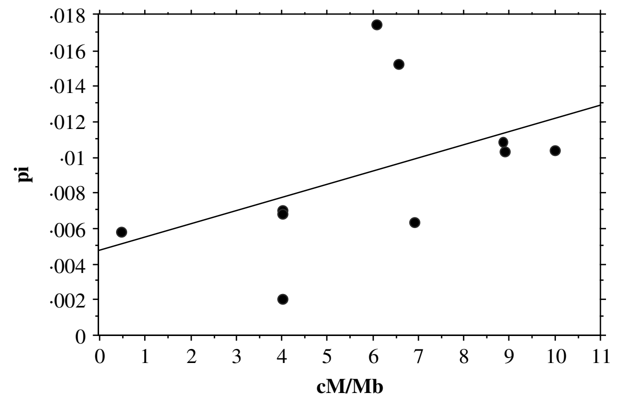

Fig. 2. Relationship between DNA sequence variation (measured as pi) to recombination rate (measured in centimorgans per megabase, cM/Mb) in *Drosophila pseudoobscura*.

from non-coding regions) by pooling DNA sequence and microsatellite variation data from several studies (e.g., Hamblin & Aquadro, 1999; Machado *et al.*, 2002). Although the slope was positive, there was no significant increase in nucleotide diversity with increasing recombination rates (see Fig. 2; $r^2=0{\cdot}22$, $P=0{\cdot}1744$). Notably, the sequenced area of the second chromosome pericentromeric region bore almost as much nucleotide variation as other regions despite its 10- to 20-fold lower recombination rate. Our measure of pi for this region was slightly elevated because of several base differences between the Mesa Verde 17 strain of *D. pseudoobscura* and the others, but these base differences were confirmed with several PCR/sequencing reactions from multiple independent DNA preparations. Excluding this sequence lowers pi for this region to 0·0035, but this revised estimate still does not produce a statistically significant regression of pi on recombination rate. Furthermore, neither codon usage nor GC content varied linearly with recombination rates (data not shown, curve estimation linear model, $P>0{\cdot}1$).

4. Discussion

This study provides a molecular-marker-based recombinational map of *Drosophila pseudoobscura*. We used this map to assemble and order various sequenced contig groups (Richards *et al.*, 2005), and we identified at least one likely error in the published assembly. We localized the breakpoints of the second chromosome inversion distinguishing *D. pseudoobscura* and *D. persimilis* to 200–300 kb intervals. We demonstrated differences in recombination rate between these two species on this chromosome as well as between the two parent species and their hybrids. However, we failed to find an association between recombination rate and nucleotide sequence variation on one chromosome within *D. pseudoobscura*. Finally, we present primer sequences and amplicon sizes for 200 markers in this species group in online Appendix.

Several aspects of our recombinational map differ from earlier linkage maps of *D. pseudoobscura*. For example, Anderson (1993) reported the genetic size of the second chromosome to be 101 cM, and Hamblin & Aquadro (1999) similarly reported its size as 128 cM. In contrast, Noor *et al.* (2000) reported a genetic size of 204 cM for the same chromosome. Our data, using many more markers, show that the second chromosome has a genetic size of approximately 200 cM (using either ordering of markers DPS2014 and DPS2027), which agrees with Noor *et al.* (2000). The cause for the discrepancy is not obvious, particularly as the map of Hamblin & Aquadro (1999) covers most of the length of the chromosome. However, the higher density of markers in our study may have enhanced our estimate of recombination distances because double crossovers between markers would have been less frequent.

(i) *Recombination rates*

We found that recombination rates are fairly uniform across much of the genome of *D. pseudoobscura*. For example, excluding the pericentromeric region, second chromosome recombination rates range from 88 to 249 kb/cM with a weighted mean of 141 kb/cM. This uniformity was previously observed in another study that provided recombination rates for the second chromosome of *D. pseudoobscura* (Hamblin & Aquadro, 1999). In contrast, the pericentromeric region displayed a much lower recombination rate, with no recombinants in a 1·2 Mb span. We were unable to test for a similar suppression of recombination near the X chromosome centromere, but it is notable that the area surveyed within XR group 6 closest to the centromere also bore the lowest

https://doi.org/10.1017/S0016672306007932 Published online by Cambridge University Press

recombination rate (186 kb/cM). Similarly, the linkage (23·1 cM) of DPSX036 (XR position 4,293,648) and DPSX023 (XL) suggests a rather low recombination rate (greater than or equal to 186 kb/cM).

Recombination rates are higher in *D. pseudoobscura* compared with many other taxa, including its close relative *D. persimilis*. Average recombination rates in *D. pseudoobscura* can be almost 4 times larger than those observed in *D. melanogaster*: For example, the average recombination rate on the XL chromosome in *D. pseudoobscura* is 8·3 cM/Mb whereas the average recombination rate across the homologous X chromosome in *D. melanogaster* is 3·3 cM/Mb. Based on a study comparing crossover frequency in *D. melanogaster* to that in its closest relatives *D. simulans* and *D. mauritiana* (True *et al*., 1996), *D. pseudoobscura* probably has a crossover rate higher than these species as well. We recently produced a linkage map of *D. mojavensis* (Staten *et al*., 2004), which is distantly related to *D. pseudoobscura* and *D. melanogaster*, and based on the tentative genome sequence assembly for this species at this point, the average recombination rate of its homologous X chromosome is 6·7 cM/Mb (data not shown), which is still slightly lower than the average we estimated for *D. pseudoobscura*. We note again that our estimates of recombination rate in *D. pseudoobscura* are probably slight *underestimates*, suggesting the difference between this species and many others is extreme. Variation in the density and intensity of recombinational hotspots between species may explain the differences between species in recombination rate (Nachman, 2002). The high recombination rate typical of *D. pseudoobscura* should dramatically reduce the impact of interference selection on this species in general.

We also observed that recombination rates were much higher in hybrid backcross populations than in either parental species in collinear regions of the second chromosome. One explanation for the difference in recombination rates between hybrids and pure species is the interchromosomal effect produced by chromosomal rearrangements (Schultz & Redfield, 1951): recombination rates tend to be higher in collinear regions of individuals heterozygous for chromosomal inversions. Hybrid females of *D. pseudoobscura* and *D. persimilis* are heterozygous for three or four large paracentric inversions, thus providing a reason for the increased recombination rate in their genomes.

The high level of recombination (sometimes 50 kb/cM) observed in *D. pseudoobscura*, and even higher levels in hybrids between this species and *D. persimilis*, should prove useful for mapping studies. Many genetic studies are limited to big chromosomal segments and subsequent elaborated molecular genetic analyses. This problem can be alleviated with enormous sample sizes, though they are often unfeasible. Our findings imply that mapping traits in *D. pseudoobscura* will not require as many individuals to be scored relative to the numbers required in other *Drosophila* species or other model organisms. This feature, along with the variety of interesting traits found in this system, makes *D. pseudoobscura* a prime model organism for studies of adaptation and speciation.

(ii) *The effect of recombination on DNA sequence variation*

Regional variation in recombination can influence the impact of natural selection on the genome, as illustrated by the repeated correlation between nucleotide sequence variability and recombination rate (see review in Nachman, 2002). Natural selection will reduce variability in regions of low recombination because linked polymorphism is eliminated whenever a beneficial mutation sweeps in the population (Maynard Smith & Haigh, 1974) or a deleterious mutation is removed (Charlesworth *et al*., 1993). In either case, only the favoured variants (usually one or very few) residing in the selected chromosomes remain in the population while the rest are eliminated. The size of the genomic region affected will correlate inversely with the rate of recombination. As a consequence, regions of high recombination should harbour more polymorphism than regions of low recombination.

Although such a correlation has been observed in several taxa ranging from plants to *Drosophila* to humans (see review in Nachman, 2002), we failed to detect a significant one in *D. pseudoobscura*. This failure is probably due to a lack of statistical power, since only 10 loci were surveyed. Additionally, our data from regions of very low recombination were limited, and it could be that the one pericentromeric region we studied was anomalous in possessing moderate levels of nucleotide sequence variation. However, the variation we observed in this region is comparable to observations in African *D. melanogaster* in regions of similarly low recombination (Andolfatto & Przeworski, 2001).

(iii) *Effective population size*

Our data also allow us to estimate the effective population size (N_e) of *D. pseudoobscura*. Previous estimates of N_e have ranged from several thousands to $4{\cdot}5\times10^6$ individuals (e.g. Dobzhansky & Wright, 1941; Powell *et al*., 1976; Schaeffer, 1995). These estimates were obtained from lethals within populations, estimates of dispersion rates or mutational and recombination parameters. The value of $4N_ec$ has been estimated to be 487·3 for *Adh* in *D. pseudoobscura*. Based on our results, the average

https://doi.org/10.1017/S0016672306007932 Published online by Cambridge University Press

recombination rate in the vicinity of *Adh* is (1/(359,413 bp/0·01 recombination events) $= 2{\cdot}782 \times 10^{-8}$ recombination events per base pair and $c = (0{\cdot}5)$ ($2{\cdot}782 \times 10^{-8}$ events per base pair) ($3{\cdot}2 \times 10^{3}$ base pairs) $= 4{\cdot}45 \times 10^{-5}$ recombination events per base pair in females. Thus, from our results, the effective population size of *D. pseudoobscura* is $4N_ec/4c = 487{\cdot}3/(4)(4{\cdot}45 \times 10^{-5}) = 2\,740\,000$. This is close to Schaeffer's (1995) estimate of 4·5 million and Leman *et al.*'s (2005) estimate of 3·7 million. It is also consistent with the role of genetic drift not being nearly as strong as envisioned by Dobzhansky & Wright (1941).

(iv) *Prospects*

In the near future, 12 species of *Drosophila* will have their genome sequences published, and molecular evolutionary studies on the genomic scale will be possible. Ideally, we will have access to both physical and genetic maps for all these species. This would provide opportunities to check with great power the effects of reduced recombination on the efficacy of selection. The production of dense recombination maps will also help in solving two of the great-unsolved problems in evolutionary biology: the genetics of adaptation and speciation. QTL mapping is one of the main approaches to studying the genetics of species differences and traits contributing to reproductive isolation. Ultimately, identifying the genes responsible for these differences and traits will allow us to explore issues such as the size of an adaptive change, the number of adaptive changes separating two species and, finally, the order of events leading to adaptation and the magnitude of their effects. Similarly, the genetics of speciation will surely be rejuvenated with the identification of genes for hybrid sterility and mating discrimination.

We thank E. Gragg and K. Zimmerman for technical assistance and B. Counterman for comments on the manuscript. This work was funded by National Science Foundation grants 0211007, 0314552, 0509780 and EPS-0346411, the State of Louisiana Board of Regents Support Fund and Louisiana Board of Regents Governor's Biotechnology Initiative grant 005.

References

Anderson, W. W. (1993). Linkage map of *Drosophila pseudoobscura*. In *Genetic Maps: Locus Maps of Complex Genomes* (ed. S. J. O'Brien), pp. 3.252–3.253. Plainview, NY: Cold Spring Harbor Laboratory Press.

Anderson, W. W., Arnold, J., Baldwin, D. G., Beckenbach, A. T., Brown, C. J., *et al.* (1991). Four decades of inversion polymorphism in *Drosophila pseudoobscura*. *Proceedings of the National Academy of Sciences of the USA* **88**, 10367–10371.

Andolfatto, P. & Przeworski, M. (2001). Regions of lower crossing over harbor more rare variants in African populations of *Drosophila melanogaster*. *Genetics* **158**, 657–665.

Charlesworth, B., Morgan, M. T. & Charlesworth, D. (1993). The effect of deleterious mutations on neutral molecular variation. *Genetics* **134**, 1289–1303.

Dobzhansky, T. (1936). Studies of hybrid sterility. II. Localization of sterility factors in *Drosophila pseudoobscura* hybrids. *Genetics* **21**, 113–135.

Dobzhansky, T. (1973). Is there gene exchange between *Drosophila pseudoobscura* and *Drosophila persimilis* in their natural habitats? *American Naturalist* **107**, 312–314.

Dobzhansky, T. & Wright, S. (1941). Genetics of natural populations V. Relations between mutation rate and accumulation of lethals in populations of *Drosophila pseudoobscura*. *Genetics* **26**, 23–51.

Hamblin, M. T. & Aquadro, C. F. (1999). DNA sequence variation and the recombinational landscape in *Drosophila pseudoobscura*: A study of the second chromosome. *Genetics* **153**, 859–869.

Hey, J. & Nielsen, R. (2004). Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* **167**, 747–760.

Kovacevic, M. & Schaeffer, S. W. (2000). Molecular population genetics of X-linked genes in *Drosophila pseudoobscura*. *Genetics* **156**, 155–172.

Lander, E. S., Green, P., Abrahamson, J., Barlow, A., Daly, M. J., *et al.* (1987). MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* **1**, 174–181.

Leman, S. C., Chen, Y., Stajich, J. E., Noor, M. A. F. & Uyenoyama, M. K. (2005). Likelihoods from summary statistics: Recent divergence between species. *Genetics* **171**, 1419–1436.

Machado, C. A. & Hey, J. (2003). The causes of phylogenetic conflict in a classic *Drosophila* species group. *Proceedings of the Royal Society of London, Series B* **270**, 1193–1202.

Machado, C. A., Kliman, R. M., Markert, J. A. & Hey, J. (2002). Inferring the history of speciation from multilocus sequence data: the case of *Drosophila pseudoobscura* and its close relatives. *Molecular Biology and Evolution* **19**, 472–488.

Maynard Smith, J. & Haigh, J. (1974). The hitch-hiking effect of a favourable gene. *Genetical Research* **23**, 23–35.

Nachman, M. W. (2002). Variation in recombination rate across the genome: evidence and implications. *Current Opinion in Genetics and Development* **12**, 657–663.

Noor, M. A. F., Schug, M. D. & Aquadro, C. F. (2000). Microsatellite variation in populations of *Drosophila pseudoobscura* and *Drosophila persimilis*. *Genetical Research* **75**, 25–35.

Noor, M. A. F., Grams, K. L., Bertucci, L. A., Almendarez, Y., Reiland, J., *et al.* (2001*a*). The genetics of reproductive isolation and the potential for gene exchange between *Drosophila pseudoobscura* and *D. persimilis* via backcross hybrid males. *Evolution* **55**, 512–521.

Noor, M. A. F., Grams, K. L., Bertucci, L. A. & Reiland, J. (2001*b*). Chromosomal inversions and the reproductive isolation of species. *Proceedings of the National Academy of Sciences of the USA* **98**, 12084–12088.

Orr, H. A. (1987). Genetics of male and female sterility in hybrids of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* **116**, 555–563.

Orr, H. A. (1995). A new linkage map of the *D. pseudoobscura* X chromosome. *Drosophila Information Service* **76**, 127–128.

https://doi.org/10.1017/S0016672306007932 Published online by Cambridge University Press

Orr, H. A. & Irving, S. (2001). Complex epistasis and the genetic basis of hybrid sterility in the *Drosophila pseudoobscura* Bogota–USA hybridization. *Genetics* **158**, 1089–1100.

Ortíz-Barrientos, D., Counterman, B. A. & Noor, M. A. F. (2004). The genetics of speciation by reinforcement. *PLoS Biology* **2**, e416.

Powell, J. R. (1983). Interspecific cytoplasmic gene flow in the absence of nuclear gene flow: Evidence from *Drosophila*. *Proceedings of the National Academy of Sciences of the USA* **80**, 492–495.

Powell, J. R., Dobzhansky, T., Hook, J. E. & Wistrand, H. E. (1976). Genetics of natural populations. XLIII. Further studies on rates of dispersal of *Drosophila pseudoobscura* and its relatives. *Genetics* **82**, 493–506.

Prakash, S., Lewontin, R. C. & Hubby, J. L. (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.

Richards, S., Liu, Y., Bettencourt, B. R., Hradecky, P., Letovsky, S., *et al*. (2005). Comparative genome sequencing of *Drosophila pseudoobscura*: Chromosomal, gene, and cis-element evolution. *Genome Research* **15**, 1–18.

Schaeffer, S. (1995). Population genetics in *Drosophila pseudoobscura*: a synthesis based on nucleotide sequence data for the *Adh* gene. In *Genetics of Natural Populations: The Continuing Importance of Theodosius Dobzhansky* (ed. L. Levine), pp. 329–352. New York: Columbia University Press.

Schaeffer, S. W. & Miller, E. L. (1992). Estimates of gene flow in *Drosophila pseudoobscura* determined from nucleotide sequence analysis of the alcohol dehydrogenase region. *Genetics* **132**, 471–480.

Schultz, J. & Redfield, H. (1951). Interchromosomal effects on crossing over in *Drosophila*. *Cold Spring Harbor Symposia in Quantitative Biology* **16**, 175–197.

Staten, R., DixonSchully, S. & Noor, M. A. F. (2004). A microsatellite linkage map of *Drosophila mojavensis*. *BMC Genetics* **5**, 12.

Tan, C. C. (1935). Salivary gland chromosomes in the two races of *Drosophila pseudoobscura*. *Genetics* **20**, 392–402.

Tan, C. C. (1946). Genetics of sexual isolation between *Drosophila pseudoobscura* and *Drosophila persimilis*. *Genetics* **31**, 558–573.

True, J. R., Mercer, J. M. & Laurie, C. C. (1996). Differences in crossover frequency and distribution among three sibling species of *Drosophila*. *Genetics* **142**, 507–523.

https://doi.org/10.1017/S0016672306007932 Published online by Cambridge University Press