DOI: 10.1534/genetics.105.041947

Mechanisms of Genetic Exchange Within the Chromosomal Inversions of *Drosophila pseudoobscura*

Stephen W. Schaeffer*,[1] and Wyatt W. Anderson†

Department of Biology and Institute of Molecular Evolutionary Genetics, Pennsylvania State University, University Park, PA 16802-5301 and †Department of Genetics, University of Georgia, Athens, GA 30602

Manuscript received February 10, 2005
Accepted for publication August 11, 2005

ABSTRACT

We have used the inversion system of *Drosophila pseudoobscura* to investigate how genetic flux occurs among the gene arrangements. The patterns of nucleotide polymorphism at seven loci were used to infer gene conversion events between pairs of different gene arrangements. We estimate that the average gene conversion tract length is 205 bp and that the average conversion rate is 3.4×10^{-6}, which is 2 orders of magnitude greater than the mutation rate. We did not detect gene conversion events between all combinations of gene arrangements even though there was sufficient nucleotide variation for detection and sufficient opportunity for exchanges to occur. Genetic flux across the inverted chromosome resulted in higher levels of differentiation within 0.1 Mb of inversion breakpoints, but a slightly lower level of differentiation in central inverted regions. No gene conversion events were detected within 17 kb of an inversion breakpoint suggesting that the formation of double-strand breaks is reduced near rearrangement breakpoints in heterozygotes. At least one case where selection rather than proximity to an inversion breakpoint is responsible for reduction in polymorphism was identified.

CHROMOSOMAL rearrangements are thought to play a vital role in the evolution of populations (Dobzhansky 1944, 1949) and the formation of new species (Noor *et al.* 2001; Rieseberg 2001; Navarro and Barton 2003). The ability of chromosomal rearrangements to suppress single crossovers (Sturtevant and Beadle 1936) in heterozygous individuals can limit gene flow from one genetic background to another. Dobzhansky (1949) recognized that reduced recombination in chromosomal inversions in *Drosophila pseudoobscura* could maintain positive epistatic combinations of genes, the so-called coadapted gene complex model. Recombination suppression may limit the spread of Dobzhansky-Muller (Dobzhansky 1936; Muller 1942) incompatibility genes among different gene arrangements resulting in the formation of new species (Noor *et al.* 2001; Rieseberg 2001; Navarro and Barton 2003).

D. pseudoobscura populations have >30 different gene arrangements segregating on the third chromosome or using the standard nomenclature, Muller's element *C* (Muller 1940; Powell 1992). The gene arrangement frequencies of *D. pseudoobscura* form classical geographic clines and vary over seasons (Dobzhansky 1944, 1948). These clines have been maintained in the face of strong gene flow among populations (Riley *et al.* 1989; Schaeffer and Miller 1992; Kovacevic and Schaeffer 2000). As a force, gene flow should homogenize arrangement frequencies among populations (Slatkin 1985), but the fitness differences among gene arrangement genotypes must be quite strong to counteract the effects of the extensive gene flow. We are interested in mapping genes that are the targets of this strong selection using the pattern and organization of nucleotide polymorphism along Muller's *C*. The suppression of recombination has the potential to create a single tight linkage group, however, preventing the detection of regions subject to neutral *vs.* selective forces (Hudson *et al.* 1987).

Schaeffer *et al.* (2003) examined patterns of nucleotide sequence evolution for a set of genes distributed across Muller's element *C*. Linkage disequilibrium levels on Muller's *C* were higher than that observed on noninverted chromosomes consistent with reduced genetic flux among the different gene arrangements (Schaeffer and Miller 1993; Kovacevic and Schaeffer 2000). Despite the presence of paracentric inversions, linkage disequilibrium tended to decline with distance suggesting that some recombination in the form of gene conversion or double crossover events has broken up associations among nucleotides. In some cases, however, some long-distance associations and independence of closely linked sites suggested that linkage disequilibrium was maintained in the face of genetic exchange. These data suggest that the historical levels of gene conversion or double crossovers have been sufficient to provide some mixing among the different chromosomal backgrounds so that targets of selection can be identified.

[1]*Corresponding author:* Department of Biology, Pennsylvania State University, 208 Erwin W. Mueller Laboratories, University Park, PA 16802-5301. E-mail: sws4@psu.edu

Phylogenetic studies of the amylase gene in *D. pseudoobscura* have found evidence for gene conversion events among the gene arrangements (Popadic and Anderson 1995; Popadic *et al.* 1995), which is consistent with the linkage disequilibrium data (Schaeffer *et al.* 2003).

Gene regions that have targets of selection nearby are likely to have reduced levels of polymorphism, if directional selection is responsible for the *D. pseudoobscura* inversion frequency clines (Schaeffer *et al.* 2003). The presence of inversions can also lead to deficiencies in polymorphism even in the absence of selection. Newly emerging gene rearrangements are expected to occur at a frequency of $1/2N$ in the population. The newly inverted region will have no variation at the time of emergence and will slowly accumulate new nucleotide mutations across the chromosome. Gene conversion and crossing over will enhance this recovery process, but levels of gene conversion and crossing are not uniform across chromosomes segregating for paracentric inversions. Navarro *et al.* (1997, 2000) have shown that polymorphism near the breakpoints will be the last regions to recover from the loss of variation associated with a new inversion chromosomal mutation. Gene conversion rates are expected to be constant over the inversion, while double crossovers are expected to be restricted to more central segments of the inverted region (Navarro *et al.* 1997). Genetic flux due to gene conversion will homogenize regions across the inverted segment, but the size of the region affected will be small. Genetic flux due to crossing over has the potential to homogenize much larger regions within the inverted segment, but the central regions will tend to show the homogenizing effects of exchange more than regions near breakpoints (Navarro *et al.* 1997). Thus, gene regions near breakpoints will tend to have reduced nucleotide variation long after the initial inversion event, while central regions recover diversity more quickly.

This study uses the sequence data from seven gene regions previously published in Schaeffer *et al.* (2003) to estimate rates of genetic flux due to gene conversion and double crossovers in the inverted regions of Muller's (1940) element *C* of *D. pseudoobscura*. The chromosomes and genes examined in the reanalysis of the Schaeffer *et al.* (2003) data were collected from four populations of *D. pseudoobscura* that span an inversion frequency cline. We use the Betrán *et al.* (1997) method to identify gene conversion tracts and estimate the rate of gene conversion per base. We use the genomic sequence of Muller's *C* in *D. pseudoobscura* to infer the distance of the seven genes to breakpoints and examine the relationship between levels of differentiation and minimum breakpoint distance. We find that gene conversion rates are 2 orders of magnitude greater than the mutation rate, that gene conversion events were not detected between all pairs of gene arrangements, and that at least one locus has significantly reduced polymorphism due to adaptive selection.

MATERIALS AND METHODS

Nucleotide sequences and GenBank Accession numbers: The analyses presented here use the sequence data of seven loci generated by Schaeffer *et al.* (2003) that are found within the inverted regions of the *D. pseudoobscura* gene arrangements. Schaeffer *et al.* (2003) present where the *D. pseudoobscura* strains were collected, the methods used to generate isochromosomal strains, the gene arrangements carried by each strain, the methods used to generate the sequences, and the methods used to align the orthologous sequences. Figure 1 shows the cytogenetic phylogeny of the seven gene arrangements relevant to this study and the locations of the seven genes on the different gene arrangement backgrounds. The hypothetical gene arrangement has not been collected from natural populations, but is an important intermediate in the inversion history of *D. pseudoobscura* chromosomes. The GenBank accession numbers of the sequences used in this study are *en* (AF476326–AF476425), *exu 1* (AF476528–AF476633), *vg* (AF476818–AF476921), *Amy 1* (AF476112–AF476221); *eve* (AF476426–AF476527), *Mef 2* (AF476728–AF476817), and F6 (AF476634–AF476727). Sequences from two arrangements, Santa Cruz and Olympic, were excluded from our analyses because only one chromosome of each was sampled. The coordinates of the sequenced genes on genomic scaffolds and on the ultrascaffold of Muller's *C* are given in the supplemental information at http://www.genetics.org/supplemental (Richards *et al.* 2005). Molecular evolutionary genetic analyses were conducted using MEGA version 2.1 (Kumar *et al.* 2001) and *DnaSP* version 4.0 (Rozas *et al.* 2003).

Estimates of the recombination parameter: We used the Hudson (1987) method to estimate the recombination parameter $C = 4Nc$, where N is the effective population size and c is the neutral recombination rate per generation. Hudson's (1987) C is a composite estimate of genetic flux due to both crossing over and gene conversion. This method uses the site and haplotype variances for the site and haplotype heterozygosities to estimate $g(C, n)$, where $g(C, n)$ is the normalized difference in the two heterozygosities that depends on C and n, the sample size. We estimated a 95% confidence interval for C with bootstrap replication. We generated a pseudo-random sample by choosing aligned nucleotides from the original data set with replacement. Each random data set was used to estimate C with the Hudson (1987) method. A total of 10,000 random data sets were generated and the 95% confidence interval was determined from the ordered list of C values. A confidence interval was not determined for loci that had $g(C, n)$ values at the theoretical minimum or maximum values (see Figure 1 in Hudson 1987). We predict that genes within inverted regions of the third chromosome will have lower values of C when estimated across all gene arrangements.

Detection of gene conversion: Shared polymorphisms between gene arrangements could arise either through parallel mutations or via gene conversion. The hypergeometric distribution was used to estimate the probability that polymorphisms shared among gene arrangements are due to parallel mutations on both backgrounds (Rozas and Aguadé 1994). Rejection of the parallel mutation hypothesis suggests that genetic exchange is responsible for the shared polymorphisms.

Chromosomal inversions provide an ideal setting in which to detect gene conversion events because suppression of recombination leads to differentiation of the chromosomes. When rare gene conversion events occur between differentiated chromosomes, diagnostic nucleotides are swapped between gene arrangements providing a signal to detect the exchange (Rozas and Aguadé 1994; Betrán *et al.* 1997). We used the methods of Betrán *et al.* (1997) as implemented in DnaSP 4.0 (Rozas *et al.* 2003) to detect the numbers and lengths of gene

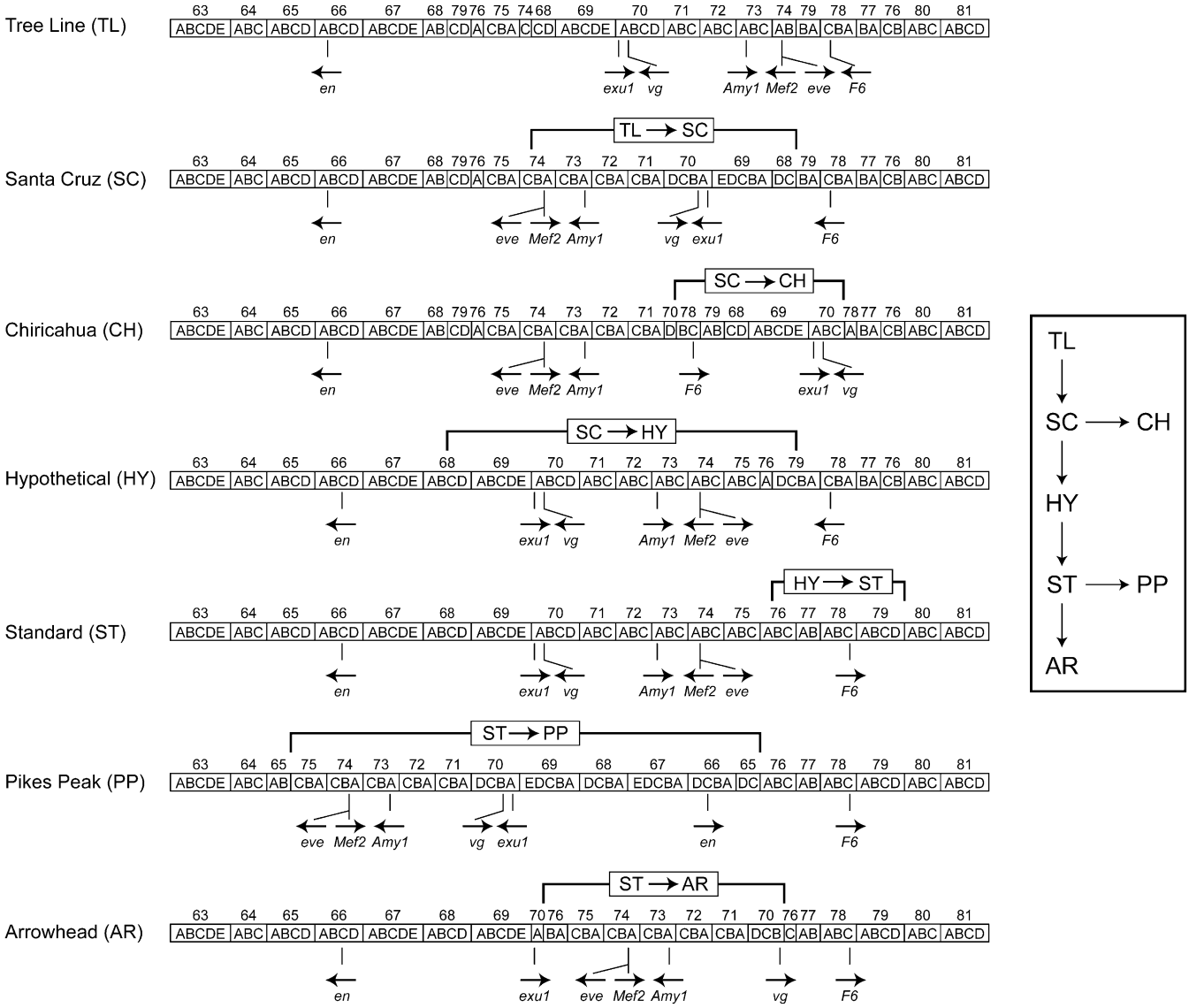

FIGURE 1.—Physical map of seven gene arrangements on Muller's *C* of *D. pseudoobscura*. The inset shows the inferred origin of six arrangements from the Tree Line common ancestor (POWELL 1992). The cytological sections and subsections are shown for each gene arrangement. The chromosomal segment that was inverted from the immediate ancestor is indicated above each cytological map and the positions and orientations of the seven sequenced gene regions are shown underneath each cytological map. The genes are *en*, engrailed; *exu1*, exuperantia 1; *eve*, even skipped; *Mef2*, myocyte enhancing factor 2; *Amy1*, amylase 1; *vg*, vestigial; and F6, the intergenic region between CG15611 and CG15605.

conversion tracts exchanged between pairs of gene arrangements. Conversion tracts were detected in pairwise comparisons of the Arrowhead, Pikes Peak, Standard, Chiricahua, and Tree Line arrangements and the source and recipient chromosomes were noted for each event. We observed several cases where different strains shared the same conversion tract (example 1 in the supplemental information at http://www.genetics.org/supplemental). In these cases, we assumed that the two copies of this conversion tract resulted from a frequency increase rather than from two independent conversion events. Only a single instance of conversion tracts such as these was included in later analyses.

The detected gene conversion tracts ≥2 nucleotides were used to estimate a maximum likelihood estimate of ϕ, the probability that a conversion tract extends to an additional nucleotide (BETRÁN *et al.* 1997, Equations 8, 6, and 4). Then ϕ can be used to estimate $E(N)$, the expected true tract length, where $E(N) = 1/(1 - \phi)$. The maximum likelihood estimate of ϕ can be used to derive an estimate of the gene conversion rate from the expected number of true gene conversion events $E(k_T) = \Phi E(T) = HcE(T)$ (see BETRÁN *et al.* 1997, pp. 95–96, for details of how to estimate c), where $\Phi = Hc$ is the exchange rate due to conversion between the gene arrangements per generation, H is the average heterozygosity for gene arrangements x and y, c is the conversion rate between the two arrangements per generation, and $E(T)$ is the sum of branch lengths of the sampled genealogy in units of generations (HUDSON 1990). Heterozygosity H for particular gene arrangement pairs is expected to vary widely for each gene arrangement pair given the stable geographic clines observed in *D. pseudoobscura* (DOBZHANSKY 1944; ANDERSON *et al.* 1991). For each pair of gene arrangements, we used the maximum heterozygosity estimate for each gene arrangement pair across 48 populations in the 1980 sample of *D. pseudoobscura*

populations (Anderson *et al.* 1991). c_{bp} is then defined as the probability of a nucleotide being transferred between arrangements per generation (see Betrán *et al.* 1997, p. 95).

The inference of gene conversion events with the Betrán *et al.* (1997) method for the *D. pseudoobscura* gene arrangement polymorphism should be interpreted with caution for several reasons. One, the mathematical models developed by Betrán *et al.* (1997) were designed to detect gene conversion tracts between two gene arrangements that differ by a single inversion event. *D. pseudoobscura* has multiple gene arrangements on Muller's *C* that differ by one to four overlapping inversion events. Thus, a conversion tract could be detected with multiple pairwise comparisons of phylogenetically similar gene arrangements (example 2 in the supplemental information at http://www.genetics.org/supplemental). In example 2, a conversion tract in the Chiricahua strain was detected with Standard, Arrowhead, and Pikes Peak chromosomes and the Standard arrangement was chosen as the source of the event because Standard shared the most polymorphic nucleotides in common.

Second, the average length of the aligned sequences in the seven gene regions is 443 nucleotides and conversion tracts may extend beyond the boundary of the sequenced region leading to larger estimates of gene conversion rates. Thus, one should view the estimates of gene conversion found in this study as minimum estimates.

Coordinates of the seven loci on the different gene arrangement backgrounds: Because the levels and types of genetic exchange are sensitive to the position within an inverted region, we wished to estimate the distance between each locus and the inversion breakpoints. The complete genome of *D. pseudoobscura* was determined in an inbred strain that carried the Arrowhead gene arrangement (Richards *et al.* 2005). This sequence was used to map the breakpoints of the Standard to Arrowhead inversion (Richards *et al.* 2005). The proximal and distal Standard to Arrowhead breakpoints occur ~8.9 and 14.9 Mb from the centromere, respectively. We approximated the coordinates for the pairs of breakpoints of the five other inversion events on the basis of the breakpoint locations on the cytogenetic map of Muller's element *C* (see Dobzhansky and Sturtevant 1938, Plate I). The sections and subsections were transferred to an image of the cytogenetic map of Muller's element *C* (Kastritsis and Crumpacker 1966). The distance of each subsection was estimated from the photomicrograph and used to determine the proportion of each subsection relative to the total map length. The Standard to Arrowhead inversion (end of 70A to the end of 76B, Figure 1) was used as the calibration point to relate the physical and cytological map; *i.e.*, 8.35 cm on the photomicrograph is 6.0 Mb on the physical map. This calibration was then used to predict the beginning and end points of each cytogenetic section and subsection. The predicted subsection locations of *orange* (*cinnabar*), *engrailed*, *exuperantia 1*, *myocyte enhancing factor 2*, *even skipped*, *Amylase 1*, *patched*, *vestigial*, and F6 predicted by the calibration were in good agreement with the physical map positions (Tan 1937; Aquadro *et al.* 1991; Schaeffer *et al.* 2003). We could now estimate the physical distance of seven genes to the inversion breakpoint of any gene arrangement by reorganizing the sections and subsections on the basis of the cytological maps of the different inversions (Dobzhansky 1944).

RESULTS

Estimates of the mutation and recombination parameters: Table 1 shows estimates of the neutral mutation (Θ) and recombination (C) parameter for the seven loci in five different gene arrangement backgrounds and across all gene arrangements. We predicted that genes within the inverted regions would have lower estimates of C across all arrangements. The observed estimate of C does tend to be less in genes within the inverted regions of the third chromosome, but the confidence limits on C overlap for all genes, reflecting the large variance that C estimates will have with a given Θ and sample size (Hudson 1987).

Estimates of C tended to be greater within gene arrangements than C estimated for all gene arrangements. There were examples where the opposite was true. In general, the estimates of recombination to mutation rate (c/μ) show that several recombination events occur for each mutation event. Overall, the large variance in C estimates is likely to be due to the relatively small number of segregating sites within each locus resulting in reduced confidence in the estimate of the recombination parameter (Hudson 1987).

Tests of parallel mutation: Figure 2 shows the number of segregating sites within each gene arrangement, the number of shared polymorphisms observed in the pairwise comparisons of five gene arrangements, and the results of the parallel mutation test. All seven loci had cases in which polymorphisms were shared between at least two gene arrangements. The number of shared polymorphisms between gene arrangements tended to decrease for genes that are often within the inverted regions of the third chromosome. A total of 25 of the 70 pairwise comparisons of the five gene arrangements at the seven loci rejected the parallel mutation hypothesis for shared polymorphisms. Six of the seven loci had at least one case in which the number of shared polymorphisms rejected a model of parallel mutations. The *exu1* locus had low levels of polymorphism within all gene arrangement backgrounds, resulting in no or few shared polymorphisms. The lone case of a shared polymorphism in *exu1* failed to reject the parallel mutation hypothesis. Loci that are generally outside of inverted regions in gene arrangement heterozygotes (*en* and F6) tended to have more cases in which shared polymorphisms reject the parallel mutation hypothesis. Loci that are often within inverted regions in gene arrangement heterozygotes (*eve*, *Mef 2*, *Amy 1*, and *vg*) had fewer instances of shared polymorphism, but nine of these cases rejected the parallel mutation hypothesis. A reduction in shared polymorphisms is not surprising for genes that are often in inverted regions because recombination suppression prevents the movement of nucleotides between chromosomal backgrounds, but the shared polymorphisms that are observed are likely the result of genetic flux.

Gene conversion: We used the Betrán *et al.* (1997) method for detecting gene conversion to examine the 25 pairwise comparisons in which the shared polymorphisms rejected the parallel mutation model. In 16 of these 25 pairwise comparisons genetic exchange events were detected. The average probability of sites being

TABLE 1

Estimates of the neutral mutation and recombination parameters for seven genes located on the third chromosome of *Drosophila pseudoobscura*

Arr		*en*	*exu1*	*eve*	*Mef2*	*Amy1*	*vg*	F6
AR	n	39	45	40	32	43	42	37
	Θ	5.6 ± 3.1	1.0 ± 0.8	1.6 ± 1.7	1.6 ± 1.0	3.2 ± 1.7	3.8 ± 2.0	4.0 ± 2.6
	C	28.5	>10,000	0.4	>10,000	229.2	72.0	21.1
	C_{95}	4.8–36.5	ND	0.04–61.1	ND	3.2–129.8	2.2–78.4	0.9–28.4
	c/μ	5.0	ND	0.2	ND	72.2	18.9	5.3
PP	n	20	21	21	21	24	20	18
	Θ	5.1 ± 2.9	0.7 ± 0.9	3.4 ± 2.1	3.9 ± 2.3	0.7 ± 0.6	5.8 ± 5.4	4.2 ± 3.3
	C	27.7	1.6	31.9	40.2	>10,000	<0.5	4.8
	C_{95}	2.5–49.3	1.0–4.0	1.8–55.7	1.2–55.5	ND	ND	0.6–18.0
	c/μ	5.4	2.4	9.3	10.3	ND	ND	1.2
ST	n	16	15	16	13	17	16	16
	Θ	5.8 ± 3.7	1.7 ± 1.0	0.1 ± 0.3	1.2 ± 1.8	2.9 ± 2.5	8.0 ± 4.2	3.0 ± 2.3
	C	13.2	474.6	>10,000	<0.5	4.4	36.1	11.8
	C_{95}	1.7–32.4	2.4–562.5	ND	ND	0.6–15.9	5.1–54.5	0.7–36.6
	c/μ	2.3	276.9	ND	ND	1.5	4.5	3.9
CH	n	16	16	16	15	17	16	14
	Θ	7.2 ± 4.3	1.6 ± 0.9	3.8 ± 2.8	3.8 ± 2.4	2.8 ± 2.4	5.8 ± 8.3	4.2 ± 2.5
	C	17.4	>10,000	10.1	20.2	4.0	<0.5	45.6
	C_{95}	4.3–28.4	ND	0.4–21.7	29.6–495.9	0.6–17.3	ND	1.4–101.1
	c/μ	2.4	ND	2.6	5.4	1.4	ND	10.8
TL	n	6	6	6	6	6	6	6
	Θ	4.9 ± 2.8	0.9 ± 0.8	0.0 ± 0.0	5.1 ± 5.9	4.5 ± 3.2	11.7 ± 6.4	2.7 ± 2.1
	C	58.6	32.4	ND	<0.5	8.7	37.3	6.6
	C_{95}	6.9–259.3	1.6–32.4	ND	ND	0.7–256.8	6.7–74.8	1.7–33.7
	c/μ	12.0	37.4	ND	ND	2.0	3.2	2.4
All	n	99	105	101	89	109	102	94
	Sites	404 (320)	358 (166)	374 (146)	418 (309)	458 (365)	651 (292)	436 (348)
	S	49	14	25	38	48	80	53
	Θ	6.5 ± 3.4	1.3 ± 1.0	3.5 ± 2.4	5.6 ± 3.7	5.7 ± 3.0	17.0 ± 9.6	5.8 ± 3.7
	C	30.3	152.8	12.7	11.1	34.7	12.7	13.1
	C_{95}	5.3–33.1	1.3–94.6	0.6–13.1	1.3–17.1	2.8–41.3	5.4–18.4	1.6–18.8
	c/μ	4.7	115.6	3.6	2.0	6.1	0.7	2.3

Arr, gene arrangement; n, sample size; Θ, heterozygosity per gene based on the number of segregating sites ± standard error (Watterson 1975); C, estimates of the neutral recombination parameter (Hudson 1987); C_{95}, 95% confidence interval on C determined with bootstrap replication; ND, not determined.

informative of conversion events (ψ) varied from 0.002 to 0.059 and was used to detect 25 nonredundant gene conversion events in five loci (Table 2). The detected exchange events could result from double crossovers rather than from gene conversion events when (1) exchanges occur between chromosomes that differ by a single large inversion or (2) exchanges occur between chromosomes in proximal or distal regions. Of the 25 detected exchanges, all exchanges were between gene arrangements separated by two to four inversions (Table 2, Figures 1 and 3). Six events could be due to double crossovers with five events detected in the noninverted proximal region and one detected in the noninverted distal region. These events were excluded from gene conversion estimates. The remaining 19 exchange events are likely to be due to gene conversion because they occurred within one of the several loops generated by the overlapping inversions of the heterozygote. The likelihood of double crossovers occurring within the same inversion loop decreases as the number of inversion loops increases because there is a higher likelihood of producing unbalanced gametes with increasing numbers of small inversion loops (Navarro *et al.* 1997).

The most ancestral gene arrangements (Popadic and Anderson 1994), Tree Line and Chiricahua, served as the source arrangement for 11 of 19 conversion tracts compared to the more derived arrangements, Standard, Arrowhead, and Pikes Peak (Figure 3). The Chiricahua and Pikes Peak arrangements were the recipients of 11 of 19 conversion tracts (Figure 3). Seven of the 19 gene conversion tracts were observed in multiple strains (Table 2). In two of these seven cases, the conversion tracts were observed in strains that were collected from two different populations. The observed tract lengths varied from 2 to 354 nucleotides, both extreme values being observed in the *vg* locus.

The estimates of the probability of a converting tract extending to an additional site ϕ, expected true tract

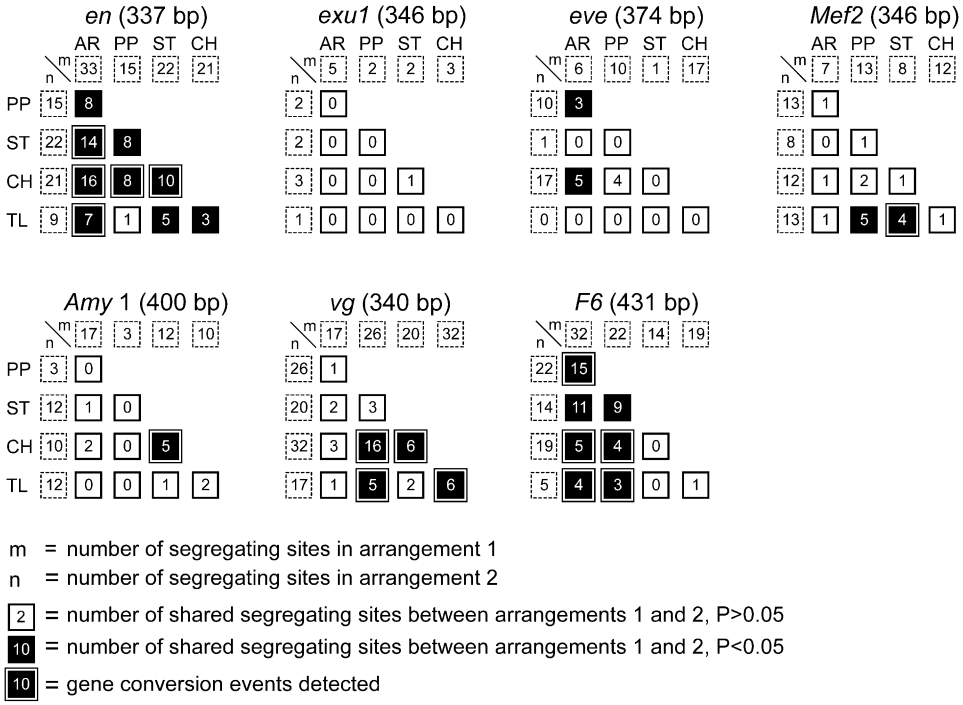

FIGURE 2.—Analysis of shared polymorphisms between gene arrangements at seven loci in *D. pseudoobscura*. Each matrix shows the numbers of polymorphic sites within a gene arrangement for each locus (dashed box). The solid or open boxes at the intersection of two gene arrangements give the numbers of shared polymorphisms for the two chromosomes. A hypergeometric distribution was used to test the hypothesis that the shared polymorphisms are due to independent mutations. The shared polymorphism box was solid if the independent mutation hypothesis was rejected at $P < 0.05$ with a correction for multiple tests (RICE 1989). The shared polymorphism box was left open when the hypergeometric test failed to reject the independent mutation hypothesis.

length $E(N)$, sum of genealogical branch lengths $E(T)$, and gene conversion flux rate $\Phi = Hc$, maximum observed heterozygosity H, and conversion rate per base pair c_{bp} for all pairwise comparisons of *D. pseudoobscura* gene arrangements are found in Table 3. The mean estimate of expected true conversion tract length is 205.8 over all loci, but the estimates vary from a low of 36.7 to a high of 330.0. The majority of comparisons suggest that 1.6 to 6.6 tracts are expected between gene arrangements.

The range of gene conversion flux (Φ) values varies from 8.4×10^{-8} to 2.3×10^{-7}. The mean estimate of the gene conversion flux rate is 1.2×10^{-7} (Table 3). The gene conversion flux rate and estimates of population heterozygosity of the pairs of gene arrangements, H, were used to solve for the conversion rate per base per generation c_{bp}. Because the gene arrangement heterozygosity varies substantially among localities in *D. pseudoobscura* (ANDERSON *et al.* 1991), we used the maximum observed estimate of H to solve for c (supplemental information at http://www.genetics.org/supplemental). This assumes that gene conversion events are most likely to occur between gene arrangements in populations where the chromosomes have the highest probability of being heterozygous. The mean value of c_{bp} across all comparisons is 3.4×10^{-6} with a range of 4.8×10^{-8}–1.4×10^{-6} (Table 3). The range of c_{bp} values estimated from the *D. pseudoobscura* is consistent with the *rp49* locus conversion rate estimated between two gene arrangements of *D. subobscura* (BETRÁN *et al.* 1997).

Impact of genetic flux on genetic diversity: We now ask what impact the genetic flux has had on genetic diversity across the inverted regions of the third chromosome of *D. pseudoobscura*. We used the *D. pseudoobscura* genome sequence to estimate the distance of the seven genes to the nearest breakpoint for all pairs of the five gene arrangements. The 25 scaffolds of Muller's element *C* form one ordered ultrascaffold (RICHARDS *et al.* 2005). The estimated sizes of the subsections of Muller's element *C* are in the supplemental information at http://www.genetics.org/supplemental. The calibration of the sequence to the cytological map was used to predict the location of each coding region to the physical map. The breakpoint for the Standard to Arrowhead rearrangement has been mapped to the genome sequence (RICHARDS *et al.* 2005), while the positions of the breakpoint pairs for the Standard to Pikes Peak, Hypothetical to Standard, and Santa Cruz to Chiricahua events were estimated from the boundaries defined by cytological subsections (DOBZHANSKY 1944).

For the *en*, *eve*, *Mef2*, *Amy1*, *vg*, and F6 gene regions, levels of genetic differentiation were estimated among all pairs of gene arrangements with γ_{st} (NEI 1982, Equation 5). Only genes within inverted regions were used for this analysis. Figure 4 shows the relationship between minimum breakpoint distance and γ_{st}. The levels of genetic differentiation tend to be highest when genes are within 0.1 Mb of breakpoints and tend to decrease to a modest level of differentiation as the distance to the nearest breakpoint increases. The correlation coefficient, however, is not significantly different from zero ($R = 0.108$, $P = 0.409$).

SCHAEFFER *et al.* (2003) found several loci that had significant deficiencies in polymorphism within the Arrowhead, Pikes Peak, and Standard arrangements using the HUDSON *et al.* (1987) test. One locus, *vg* within

TABLE 2

Gene conversion tracts detected with the BETRÁN *et al.* (1997) approach

Gene	Source	Recipient	Sites	Strain(s)	Popn.	ψ	l	Loc	Inversion Len (Mb)
en	CH	AR	338	1	JR	0.002	206	P	
en	TL	AR	338	1	MSH	0.005	56	P	
en	TL	AR	338	1	KB	0.005	190	P	
en	CH	PP	355	1	DM	0.007	204	I	4.9 (2.6)
en	PP	CH	355	1	DM	0.007	137	I	4.9 (2.6)
en	ST	CH	337	1	JR	0.004	12	P	
en	ST	CH	337	2	KB, DM	0.004	5	P	
Mef 2	ST	TL	354	1	MSH	0.013	163	I	6.0 (1.1)
Amy 1	CH	ST	422	2	JR	0.011	83	I	4.9 (2.1)
Amy 1	ST	CH	422	2	JR, KB	0.011	73	I	4.9 (2.1)
vg	CH	PP	432	2	JR	0.059	153	I	1.9 (0.9)
vg	PP	CH	432	1	KB	0.059	354	I	1.9 (0.9)
vg	TL	PP	433	1	DM	0.036	153	I	6.1 (1.0)
vg	TL	PP	433	1	DM	0.036	263	I	6.1 (1.0)
vg	PP	TL	433	1	MSH	0.036	2	I	6.1 (1.0)
vg	CH	ST	443	2	JR	0.049	112	I	1.8 (0.8)
vg	ST	CH	443	1	KB	0.049	426	I	1.8 (0.8)
vg	CH	TL	439	1	MSH	0.026	54	I	1.9 (1.1)
F6	PP	AR	431	3	JR, KB, DM	0.002	15	D	
F6	AR	CH	431	2	JR	0.019	14	I	1.5 (0.7)
F6	TL	AR	431	1	KB	0.015	116	I	2.4 (0.7)
F6	TL	AR	431	2	KB, DM	0.015	73	I	2.4 (0.7)
F6	TL	AR	431	1	KB	0.015	25	I	2.4 (0.7)
F6	CH	PP	431	1	JR	0.019	243	I	1.4 (0.6)
F6	PP	CH	431	2	JR	0.019	14	I	1.4 (0.6)

Gene, gene where gene conversion tract was detected; Strain(s), recipient strains with the observed conversion tract; Sites, number of aligned nucleotide sites excluding gaps; Strains, number of recipient strains with the conversion tract; Popn., populations where the recipient strains were collected; ψ, probability of a nucleotide site being informative of a conversion tract; l, observed conversion tract length in nucleotides; Loc, exchange event is in a noninverted proximal segment (P), noninverted distal segment (D), or within an inverted region (I); Inversion Len, length of the inverted region in Mb where the conversion was detected and the distance to the nearest breakpoint given in parentheses.

Arrowhead, had a significant deficiency that is best explained by its proximity to an inversion breakpoint (Figure 1) where the homogenizing effect of genetic flux is expected to be low. The cases where *vg* was within 0.1 Mb of the nearest breakpoint had some of the highest γ_{st} values: 0.530, 0.608, 0.622, and 0.714. Other loci such as *Mef2* within Arrowhead or *Amy1* within Pikes have significant deficiencies in polymorphism despite the opportunity for higher levels of genetic flux in the central regions of the inversion. The reduction in polymorphism in *Mef2* within the Arrowhead arrangement is particularly striking given its close proximity to *eve*, which is located <75 kb away from *Mef2*. *eve* does not have a significant reduction in polymorphism compared to *Mef2*.

DISCUSSION

We have used sequences from seven loci on Muller's element *C* to estimate levels of genetic flux due to gene conversion among five different gene arrangements in *D. pseudoobscura*. The mean gene conversion tract length detected was 205.8 bp. This should be viewed as a minimal estimate of the conversion tract length because the gene regions sequenced had an average length of 443 bp. Observed conversion tracts could extend beyond the boundaries of the sequenced region. This tract length estimate is higher than that for two gene arrangements at the *rp49* locus in *D. subobscura* (122 bp) (BETRÁN *et al.* 1997) and less than that found from coconversion data at the *rosy* locus of *D. melanogaster* (352 bp) (HILLIKER *et al.* 1994). Our conversion tract data were used to estimate a mean conversion rate per nucleotide per generation of $c_{bp} = 3.4 \times 10^{-6}$, which is an order of magnitude greater than that found for the *rp49* locus of *D. subobscura* (2.7×10^{-7}); however, the range of values observed includes the estimate for *rp49*.

The estimates of gene conversion rates from the *D. pseudoobscura* data should be viewed with caution. One of the key choices that must be made to estimate the conversion rate per base per generation is what gene arrangement heterozygosity to use. This is particularly problematic in *D. pseudoobscura* because populations vary widely in the frequencies of the different gene arrangements. Choose too a low value for the heterozygosity value and the conversion rate will be an

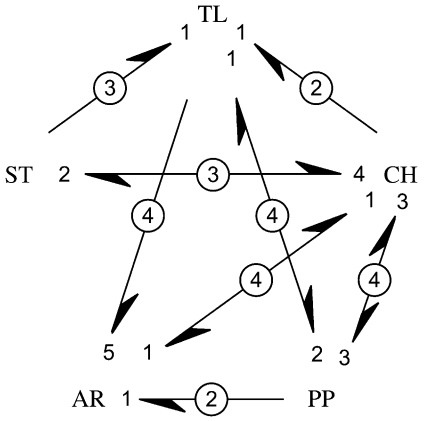

FIGURE 3.—Polarity of 25 gene conversion events among 10 pairwise possible exchanges between five gene arrangements in *D. pseudoobscura*. The five gene arrangements are given by their two-letter names: TL, Tree Line; ST, Standard; CH, Chiricahua; AR, Arrowhead; and PP, Pikes Peak. The arrows connecting the five gene arrangement pairs indicate cases in which a gene conversion event was observed with the arrowhead pointing from the source arrangement to the recipient arrangement. The arrowheads point to the number of gene conversion events that have occurred in the recipient arrangement from each of the source arrangements. The circled numbers indicate the number of inversion events that separate the two arrangements with a detected exchange event.

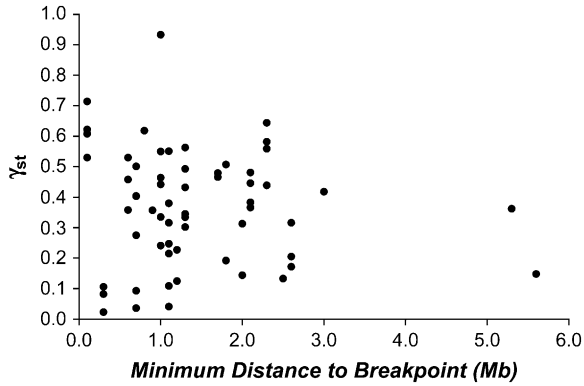

FIGURE 4.—Relationship between minimum breakpoint distance and genetic differentiation of genes between pairs of gene arrangements.

overestimate. We chose to use the maximum value for the gene arrangement heterozygosity because this assumes the maximum probability for genetic flux between arrangements.

Second, the original conversion tract data were culled to remove duplicate conversion tracts. This assumes that all conversion tracts were of unique origin and that if multiple identical tracts were observed, then the tract must have increased in frequency in the population. This seems the most parsimonious assumption, but if tracts are generated more than once, then culling the tract data would lead to an underestimate of the conversion rate. Despite these caveats, our estimates of gene conversion rates are consistent with previous estimates based on coconversion data and nucleotide sequences (HILLIKER *et al.* 1994; BETRÁN *et al.* 1997).

If we assume that each conversion tract is unique, then our observation of the occurrence of identical tracts in more than one population is consistent with the extensive gene flow that has been previously documented in this species (RILEY *et al.* 1989; SCHAEFFER and MILLER 1992; KOVACEVIC and SCHAEFFER 2000; SCHAEFFER *et al.* 2003).

SCHAEFFER *et al.* (2003) found extensive differentiation among the major gene arrangements. In addition, they found extensive linkage disequilibrium among nucleotide sites within the inverted regions; however, there was a negative relationship between linkage disequilibrium and distance between sites. MUNTE *et al.* (2005) also found extensive differentiation and linkage disequilibrium across two gene arrangements (O_{ST} and

TABLE 3

Estimates of genetic flux Φ due to gene conversion between different gene arrangements in *D. pseudoobscura*

Locus	Arr1	Arr2	ϕ	$E(N)$ (95%CI)	$E(T)$	Φ (Hc)	H	c_{bp}
en	PP	CH	0.9970	330.0 (75.1–1,041.6)	4.7×10^{7}	8.4×10^{-8}	0.054	1.4×10^{-6}
Mef 2	ST	TL	0.9963	271.7 (40.6–1,063.8)	1.6×10^{7}	1.0×10^{-7}	0.274	2.9×10^{-7}
Amy 1	ST	CH	0.9940	167.2 (36.0–531.9)	4.0×10^{7}	1.2×10^{-7}	0.280	1.6×10^{-7}
vg	PP	CH	0.9965	282.5 (77.9–900.9)	2.4×10^{7}	9.2×10^{-8}	0.054	1.1×10^{-6}
vg	PP	TL	0.9942	173.3 (57.8–473.9)	1.8×10^{7}	2.3×10^{-7}	0.110	8.3×10^{-7}
vg	ST	CH	0.9967	304.0 (83.6–970.8)	2.1×10^{7}	1.1×10^{-7}	0.280	2.6×10^{-7}
vg	CH	TL	0.9900	99.7 (14.2–390.6)	1.5×10^{7}	1.3×10^{-7}	0.257	1.2×10^{-7}
F6	AR	CH	0.9728	36.7 (4.1–39.3)	5.5×10^{7}	1.1×10^{-7}	0.197	4.8×10^{-8}
F6	AR	TL	0.9929	141.0 (40.6–390.6)	4.6×10^{7}	1.4×10^{-7}	0.238	2.0×10^{-7}
F6	PP	CH	0.9947	190.1 (48.0–609.8)	3.4×10^{7}	9.5×10^{-8}	0.054	7.8×10^{-7}
All events				205.8 (128.5–323.6)	Mean	1.2×10^{-7}		3.4×10^{-6}

Arr1, gene arrangement 1; Arr2, gene arrangement 2; ϕ, probability that a converting tract elongates to an additional nucleotide; $E(N)$, expected true gene conversion tract length with 95% confidence interval, $E(T)$, the sum of branch lengths in the genealogy; Φ, the flux of gene conversion events per generation between arrangement within locus; H, maximum heterozygosity of gene arrangements 1 and 2 across all populations (ANDERSON *et al.* 1991); c_{bp}, conversion rate per base pair per generation.

O_{3+4}) in *D. subobscura*, but did not see a reduction in linkage disequilibrium with distance. They concluded that either double crossovers do not occur or recombinants are selected against. The results of Schaeffer *et al.* (2003) and this study also suggest that double crossovers are not the likely explanation for the type of genetic exchange that reduces nonrandom associations with distance. First, most observed exchanges occurred between gene arrangements that differ by two or more inversions. In these heterozygotes, it is unlikely that two crossovers would occur in the same inversion loop because of the size and number of loops formed (Navarro *et al.* 1997). In addition, segregation of the inverted chromosomes is likely to occur via achiasmate pathway (Hawley *et al.* 1992; Hawley and Theurkauf 1993). Second, if double crossovers occurred at a high enough frequency, then we might expect to see a reduction of genetic differentiation with increasing distance from inversion breakpoints. We do have evidence that genetic differentiation is highest near inversion breakpoints from the *vg* locus, which is within 0.1 Mb of the nearest breakpoint, but differentiation levels reach a constant level in the central region of inverted regions.

Gene conversion seems to be the most likely explanation for how linkage disequilibrium decreases with distance within the gene arrangements of *D. pseudoobscura*. Although gene conversion occurs at a higher rate per base than the mutation rate, the average number of bases transferred between arrangements is small relative to the size of the inverted regions. Thus, more generations would be needed to reduce linkage disequilibrium levels with distance, suggesting that the inversion system of *D. pseudoobscura* is quite old (Aquadro *et al.* 1991).

The conversion tract data show that the sources for observed gene conversion tracts tend to be the more ancestral arrangements. Of 19 observed conversion tracts for genes in inverted regions, 11 used either the Tree Line or Chiricahua arrangements as the source of DNA (Figure 3). No conversion events were detected between either Standard and Arrowhead (ST/AR) or Standard and Pikes Peak (ST/PP) chromosomes. For gene conversion events to be detected with the Betrán *et al.* (1997) approach, an inversion mutation must generate a new arrangement, the new arrangement must accumulate nucleotide differences over time, the new arrangement must be paired with a second arrangement in a heterozygote, and genetic exchange must be initiated within the inverted region. The failure to detect a conversion event may result from (1) a lack of power where gene arrangements are not sufficiently differentiated from each other; (2) a population structure effect where some pairs of arrangements never reach appreciable frequencies in populations to form heterozygotes, where conversion events may occur; (3) a position effect where the proximity of the locus to a inversion breakpoint may inhibit gene conversion; or (4) sample sizes insufficient to detect at least one conversion event for each pair of chromosomes.

The lack of power hypothesis suggests that younger inversions such as Arrowhead, Standard, and Pikes Peak do not have sufficient nucleotide differences in linkage disequilibrium to detect gene conversion events. We can test this hypothesis by examining the net number of nucleotide differences per site between the pairs of gene arrangements at the seven loci (d_A; Nei 1987, Equation 10.21) (supplemental information at http://www.genetics.org/supplemental). If this hypothesis is true, then more divergent pairs of gene arrangements will more likely show evidence for gene conversion events than less divergent gene arrangement pairs. In general, we find support for this hypothesis at the *exu1*, *eve*, *Mef2*, and F6 loci where there are virtually no net nucleotide differences between either ST/AR or ST/PP and no conversion events were detected. The lack of power hypothesis does not explain why gene conversion events were not observed in *en*, *Amy1*, or *vg* where the two derived pairs (ST/AR and ST/PP) have higher levels of nucleotide differentiation than other chromosome pairs where gene conversion events were detected, *e.g.*, Chiricahua and Tree Line in *vg*. Thus, the lack of power hypothesis does not completely explain why conversions were not detected in the ST/AR and ST/PP comparisons.

The population structure hypothesis suggests that gene conversion events are not detected among some arrangements because the heterozygosity for these chromosomes is sufficiently low to limit the opportunity for gene conversion events to occur. We can test the population structure hypothesis by examining the maximum heterozygosity observed for all pairs of chromosomes examined in this study using historical collections of *D. pseudoobscura* (1938–1983) (Anderson *et al.* 1991) (supplemental information at http://www.genetics.org/supplemental). As stated above, populations of *D. pseudoobscura* vary dramatically in their gene arrangement frequencies, and the likelihood of two gene arrangements being found in heterozygous condition in any population can be quite different. The Pikes Peak and Chiricahua arrangements had six observed gene conversion tracts although the maximum heterozygosity observed for these two arrangements was 0.034 and an average of 0.006. Standard and Arrowhead chromosomes are expected to have ample opportunity for genetic exchange because the maximum frequency of heterozygotes observed was 0.417 with a mean of 0.126 (Anderson *et al.* 1991). The Standard and Pikes Peak arrangements may have less opportunity for genetic exchange between the inversions because the maximum heterozygosity that has been observed is 0.137 with a mean of 0.017. In both cases, ST/AR and ST/PP, heterozygosity levels appear to be sufficient to have allowed gene conversion events to take place, yet no events were detected.

A position effect hypothesis may explain why gene conversion events were not detected between Standard and Arrowhead chromosomes in the vestigial locus. The proximity of *vg* locus with respect to the inversion breakpoint may act to reduce gene conversion events. The *vg* locus is located 17 kb from the distal breakpoint of the Standard to Arrowhead inversion (Richards *et al.* 2005). Models of conversion and crossing over have suggested that conversion occurs at a constant rate across an inverted region (Navarro *et al.* 1997). The lack of conversion tracts in the *vg* locus between Standard and Arrowhead may suggest that the initiation of conversion tracts via double-strand breaks is not constant across inverted regions and may decline near breakpoints as is the case for crossover events. This interpretation should be viewed with caution because the sample sizes of Standard and Arrowhead chromosomes in *vg* were relatively small (AR, $n = 42$; ST, $n = 16$). Larger samples will be necessary to test this hypothesis. Sample size may also explain why conversion tracts were not detected between Standard and Pikes Peak chromosomes.

The Chiricahua arrangement also is the recipient of most observed gene conversion tracts (6 of 19). This is interesting given that the Chiricahua arrangement is not one of the most frequent arrangements in *D. pseudoobscura* populations. Because the Chiricahua arrangement tends to be fairly differentiated from the other chromosomes, gene conversion tracts are readily detected because the transfer of DNA from any of the arrangements will be apparent. Even though the Chiricahua is observed at lower frequencies in *D. pseudoobscura* populations, it is a widely distributed chromosome that increases the opportunity for gene conversion (Dobzhansky 1944; Anderson *et al.* 1991).

What effect has the genetic exchange had on genetic variation of the genes at or near the inversion breakpoints? In general, levels of polymorphism were low near inversion breakpoints (Schaeffer *et al.* 2003) consistent with the loss of variation associated with the initial inversion mutation event (Navarro *et al.* 2000). Schaeffer *et al.* (2003) documented that several loci had a significant reduction in polymorphism that could be due to proximity to an inversion breakpoint or due to directional selection (Schaeffer *et al.* 2003). The data presented here suggest that some of the observed reductions are due to proximity to the nearest inversion breakpoint (*vg* in Arrowhead, Figure 1). The significant reduction in polymorphism in the *Mef2* gene within the Arrowhead arrangement may indeed be the target of adaptive selection. *Mef2* is located 2 Mb from the nearest breakpoint, yet it has lower levels of polymorphism with respect to divergence than *vg*. *Mef2* is located within 75 kb of the *eve* locus, a gene that has a level of polymorphism that is near that expected given its distance from the Arrowhead inversion breakpoint. It is quite striking that the pattern of polymorphism in *Mef2* has become dissociated from *eve*, given that both loci are within the inverted region of the Arrowhead inversion. These data suggest that that there is a gene near *Mef2* that may be involved with the adaptive selection that alters the frequency of the Arrowhead arrangement within the inversion frequency cline in *D. pseudoobscura* (Schaeffer *et al.* 2003). Further investigation of the genes near *Mef2* will be needed to identify the specific target of the adaptive selection.

Navarro and Barton (2003) have proposed a chromosomal model of speciation that suggests where inversions can act as a barrier to gene flow and allow the accumulation of Dobzhansky-Muller incompatibility genes (Dobzhansky 1936; Muller 1942). Inversions are likely to be a heterogeneous barrier to gene flow because genes near breakpoints are less likely to be exchanged than those in more central regions of the inversion (Hasson and Eanes 1996; Navarro *et al.* 1997, 2000). Our data are consistent with this idea although we do find evidence for the dissociation of selective effects within central regions of the inversion. This suggests that incompatibility genes that are closer to breakpoints are less likely to spread to other gene arrangement backgrounds than ones in the central regions of the chromosome.

In summary, we have used the inversion system of *D. pseudoobscura* to investigate how genetic flux occurs among the gene arrangements. We have used the Betrán *et al.* (1997) method to infer gene conversion events and found that this approach indicates that gene conversion occurs at levels that are 2 orders of magnitude greater than the mutation rate. We did find some biases in that not all possible gene arrangements were observed to have had gene conversion events even though there was sufficient variation and sufficient opportunity for exchanges to occur. Heterogeneous levels of genetic flux at inversion breakpoints resulted in reductions in polymorphism. At least one case in which selection rather than proximity to an inversion breakpoint is responsible for reduction in polymorphism has been identified.

We thank the following people for technical assistance in collecting sequence data and mapping experiments: M. Paula Goetting-Minesky, Miro Kovacevic, John R. Peoples, Jennifer L. Graybill, Jonathan M. Miller, Kyungsun Kim, Donna Sosnoski, and Julie G. Nelson. We also thank two anonymous reviewers who raised important issues that were overlooked in our original analyses and interpretations. This work was supported by two grants from the National Science Foundation, 9726285 to S.W.S. and 9729144 to W.W.A. Any opinions, findings, conclusions, or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the National Science Foundation.

LITERATURE CITED

Anderson, W. W., J. Arnold, D. G. Baldwin, A. T. Beckenbach, C. J. Brown *et al.*, 1991 Four decades of inversion polymorphism in *Drosophila pseudoobscura*. Proc. Natl. Acad. Sci. USA **88:** 10367–10371.

Aquadro, C. F., A. L. Weaver, S. W. Schaeffer and W. W. Anderson, 1991 Molecular evolution of inversions in *Drosophila pseudoobscura*: the amylase gene region. Proc. Natl. Acad. Sci. USA **88:** 305–309.

Betrán, E., J. Rozas, A. Navarro and A. Barbadilla, 1997 The estimation of the number and the length distribution of gene conversion tracts from population DNA sequence data. Genetics **146:** 89–99.

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

Dobzhansky, T., 1944 Chromosomal races in *Drosophila pseudoobscura* and *Drosophila persimilis*. Carnegie Inst. Washington Publ. **554:** 47–144.

Dobzhansky, T., 1948 Genetics of natural populations. XVI. Altitudinal and seasonal changes in certain populations of *Drosophila pseudoobscura* and *Drosophila persimilis*. Genetics **33:** 158–176.

Dobzhansky, T., 1949 Observations and experiments on natural selection in *Drosophila*, pp. 210–224 in *Proceedings of the International Congress of Genetics*, edited by G. Bonnier and R. Larsson. Berlingska Boktryckeriet, Lund, Sweden.

Dobzhansky, T., and A. H. Sturtevant, 1938 Inversions in the chromosomes of *Drosophila pseudoobscura*. Genetics **23:** 28–64.

Hasson, E., and W. F. Eanes, 1996 Contrasting histories of three gene regions associated with *In(3L)Payne* of *Drosophila melanogaster*. Genetics **144:** 1565–1575.

Hawley, R. S., and W. E. Theurkauf, 1993 Requiem for distributive segregation: achiasmate segregation in Drosophila females. Trends Genet. **9:** 310–317.

Hawley, R. S., H. Irick, A. E. Zitron, D. A. Haddox, A. Lohe *et al.*, 1992 There are two mechanisms of achiasmate segregation in Drosophila females, one of which requires heterochromatic homology. Dev. Genet. **13:** 440–467.

Hilliker, A. J., G. Harauz, A. G. Reaume, M. Gray, S. H. Clark *et al.*, 1994 Meiotic gene conversion tract length distribution within the *rosy* locus of *Drosophila melanogaster*. Genetics **137:** 1019–1026.

Hudson, R. R., 1987 Estimating the recombination parameter of a finite population model without selection. Genet. Res. **50:** 245–250.

Hudson, R. R., 1990 Gene genealogies and the coalescent process. Oxf. Surv. Evol. Biol. **7:** 1–44.

Hudson, R. R., M. Kreitman and M. Aguadé, 1987 A test of neutral molecular evolution based on nucleotide data. Genetics **116:** 153–159.

Kastritsis, C. D., and D. W. Crumpacker, 1966 Gene arrangements in the third chromosome of *Drosophila pseudoobscura*. I. Configurations with tester chromosomes. J. Hered. **57:** 150–158.

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

Kumar, S., K. Tamura, I. B. Jakobsen and M. Nei, 2001 MEGA2: Molecular Evolutionary Genetics Analysis software. Bioinformatics **17:** 1244–1245.

Muller, H. J., 1940 Bearings of the "Drosophila" work on systematics, pp. 185–268 in *The New Systematics*, edited by J. Huxley. Clarendon Press, Oxford.

Muller, H. J., 1942 Isolating mechanisms, evolution and temperature, pp. 71–125 in *Biological Symposia*, edited by T. Dobzhansky. The Jaques Catell Press, Lancaster, PA.

Munte, A., J. Rozas, M. Aguade and C. Segarra, 2005 Chromosomal inversion polymorphism leads to extensive genetic structure: a multilocus survey in *Drosophila subobscura*. Genetics **169:** 1573–1581.

Navarro, A., and N. H. Barton, 2003 Accumulating postzygotic isolation genes in parapatry: a new twist on chromosomal speciation. Evolution **57:** 447–459.

Navarro, A., A. Barbadilla and A. Ruiz, 2000 Effect of inversion polymorphism on the neutral nucleotide variability of linked chromosomal regions in Drosophila. Genetics **155:** 685–698.

Navarro, A., E. Betrán, A. Barbadilla and A. Ruiz, 1997 Recombination and gene flux caused by gene conversion and crossing over in inversion heterokaryotypes. Genetics **146:** 695–709.

Nei, M., 1982 Evolution of human races at the gene level, pp. 167–181 in *Human Genetics. Part A: The Unfolding Genome*, edited by B. Bonne-Tamir, T. Cohen and R. M. Goodman. Alan R. Liss, New York.

Nei, M., 1987 *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Noor, M. A., K. L. Grams, L. A. Bertucci and J. Reiland, 2001 Chromosomal inversions and the reproductive isolation of species. Proc. Natl. Acad. Sci. USA **98:** 12084–12088.

Popadic, A., and W. W. Anderson, 1994 The history of a genetic system. Proc. Natl. Acad. Sci. USA **91:** 6819–6823.

Popadic, A., and W. W. Anderson, 1995 Evidence for gene conversion in the amylase multigene family of *Drosophila pseudoobscura*. Mol. Biol. Evol. **12:** 564–572.

Popadic, A., D. Popadic and W. W. Anderson, 1995 Interchromosomal exchange of genetic information between gene arrangements on the third chromosome of *Drosophila pseudoobscura*. Mol. Biol. Evol. **12:** 938–943.

Powell, J. R., 1992 Inversion polymorphisms in *Drosophila pseudoobscura* and *Drosophila persimilis*, pp. 73–126 in *Drosophila Inversion Polymorphism*, edited by C. B. Krimbas and J. R. Powell. CRC Press, Ann Arbor, MI.

Rice, W. R., 1989 Analyzing tables of statistical tests. Evolution **43:** 223–225.

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

Rieseberg, L. H., 2001 Chromosomal rearrangements and speciation. Trends Ecol. Evol. **16:** 351–358.

Riley, M. A., M. E. Hallas and R. C. Lewontin, 1989 Distinguishing the forces controlling genetic variation at the *Xdh* locus in *Drosophila pseudoobscura*. Genetics **123:** 359–369.

Rozas, J., and M. Aguadé, 1994 Gene conversion is involved in the transfer of genetic information between naturally occurring inversions of *Drosophila*. Proc. Natl. Acad. Sci. USA **91:** 11517–11521.

Rozas, J., J. C. Sanchez-DelBarrio, X. Messeguer and R. Rozas, 2003 DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics **19:** 2496–2497.

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

Schaeffer, S. W., and E. L. Miller, 1993 Estimates of linkage disequilibrium and the recombination parameter determined from segregating nucleotide sites in the alcohol dehydrogenase region of *Drosophila pseudoobscura*. Genetics **135:** 541–552.

Schaeffer, S. W., P. Goetting-Minesky, M. Kovacevic, J. Peoples, J. L. Graybill *et al.*, 2003 Evolutionary genomics of inversions in *Drosophila pseudoobscura*: Evidence for epistasis. Proc. Natl. Acad. Sci. USA **100:** 8319–8324.

Slatkin, M., 1985 Gene flow in natural populations. Annu. Rev. Ecol. Syst. **16:** 393–430.

Sturtevant, A. H., and G. W. Beadle, 1936 The relations of inversions in the X chromosome of *Drosophila melanogaster* to crossing over and disjunction. Genetics **21:** 544–604.

Tan, C. C., 1937 The cytological maps of the autosomes in *Drosophila pseudoobscura*. Z. Zellforsch. Mikrosk. Anat. Bd. **26:** 439–462.

Watterson, G. A., 1975 On the number of segregating sites in genetical models without recombination. Theor. Popul. Biol. **7:** 256–276.

Communicating editor: M. Aguadé