

# Molecular Population Genetics of *Drosophila* Subtelomeric DNA

Jennifer A. Anderson,<sup>\*,1</sup> Yun S. Song<sup>\*,†</sup> and Charles H. Langley<sup>\*</sup>

<sup>\*</sup>Section of Evolution and Ecology, University of California, Davis, California 95616 and <sup>†</sup>Computer Science Division and Department of Statistics, University of California, Berkeley, California 94720

Manuscript received October 9, 2007  
Accepted for publication November 2, 2007

## ABSTRACT

DNA sequence surveys in yeast and humans suggest that the forces shaping telomeric polymorphism and divergence are distinctly more dynamic than those in the euchromatic, gene-rich regions of the chromosomes. However, the generality of this pattern across outbreeding, multicellular eukaryotes has not been determined. To characterize the structure and evolution of *Drosophila* telomeres, we collected and analyzed molecular population genetics data from the X chromosome subtelomere in 58 lines of North American *Drosophila melanogaster* and 29 lines of African *D. melanogaster*. We found that *Drosophila* subtelomeres exhibit high levels of both structural and substitutional polymorphism relative to linked euchromatic regions. We also observed strikingly different patterns of variation in the North American and African samples. Moreover, our analyses of the polymorphism data identify a localized hotspot of recombination in the most-distal portion of the X subtelomere. While the levels of polymorphism decline sharply and in parallel with rates of crossing over per physical length over the distal first euchromatic megabase pairs of the X chromosome, our data suggest that they rise again sharply in the subtelomeric region ( $\approx 80$  kbp). These patterns of historical recombination and geographic differentiation indicate that, similar to yeast and humans, *Drosophila* subtelomeric DNA is evolving very differently from euchromatic DNA.

POPULATION geneticists aspire to understand the evolutionary forces shaping patterns of molecular polymorphism and divergence. The recent increase in the quantity of DNA sequence data from protein-coding regions has led to considerable advances in our understanding of the forces governing the evolution of such regions. For example, genomic surveys have revealed functional variants as well as classes of genes enriched for the signature of natural selection (*e.g.*, BUSTAMANTE *et al.* 2005; NIELSEN *et al.* 2005; BEGUN *et al.* 2007). In contrast, regions of the genome containing fewer genes but an increased density of repetitive sequences have been less amenable to population and molecular evolutionary analysis. These regions, typically found adjacent to telomeres and centromeres, are collectively referred to as the heterochromatin. Because heterochromatic DNA is difficult to clone, sequence, and assemble, much less is known about the structure of these regions. Thus, there is almost no foundation for effective population or comparative genomic analyses. Indeed, only recently have researchers begun to identify the sparsely distributed protein-coding sequences embedded in heterochromatin (reviewed in YASUHARA and WAKIMOTO 2006).

Most population and molecular evolutionary surveys from telomeric regions come from yeast and primates (but see recent work on *Arabidopsis* in KUO *et al.* 2006). These studies show that telomeric DNA is evolving very differently from the gene-rich euchromatin. For example, the telomeric regions of several human chromosomes are dramatically different from the corresponding regions in chimpanzees (*e.g.*, TRASK *et al.* 1998), showing large-scale rearrangements, as well as extensive differentiation of highly repeated elements. In addition, the telomeric regions of the human and yeast genomes exhibit unusually high levels of within-species structural and nucleotide polymorphism (see PRYDE *et al.* 1997 and MEFFORD and TRASK 2002 for reviews).

Although data from other organisms are scarce, we have recently found that the most-distal regions of *Drosophila* subtelomeres are evolving quite rapidly between *Drosophila melanogaster* and its close relatives, *D. simulans* and *D. yakuba*, which is consistent with observations of primates (J. A. ANDERSON, S. E. CELNIKER and C. H. LANGLEY, unpublished results). Figure 1 depicts a typical telomeric region of *D. melanogaster*; at a large scale, the structure of this genomic region resembles the structure typical of other eukaryotes. The most-distal portion contains an array of “telomeric” repeats, which, in yeast and humans, consist of chains of a short G-rich motif that is added onto the ends of a chromosome by the enzyme telomerase (for a review, see BLACKBURN 1992). However, *Drosophila* has evolved an alternative mechanism involving the insertion of the transposable

Sequence data from this article have been deposited with the GenBank Data Libraries under accession nos. EU290758–EU291973.

<sup>1</sup>Corresponding author: Department of Neurology, University of California, 1515 Newton Ct., Davis, CA 95618. E-mail: janders@ucdavis.edu

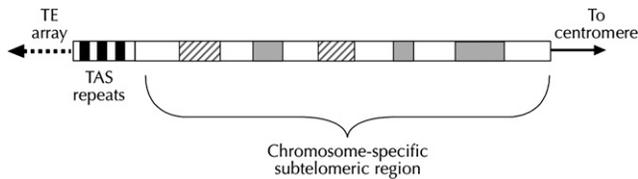


FIGURE 1.—A typical *Drosophila* telomeric region. The left arrow points to the most terminal repeats (the transposable elements *HeT-A* and *TART*), while the right arrow points in the direction of the centromere. The subtelomere is composed of two main regions: the distal repetitive portion, or the TAS, and the centromere-proximal portion, labeled as the “chromosome-specific subtelomeric region.” Within this region, coding sequence is shaded, whereas repetitive sequence is depicted as diagonal stripes.

elements *HeT-A* and *TART* into chromosome ends (RUBIN 1978; YOUNG *et al.* 1983; RENKAWITZ-POHL and BIALOJAN 1984; TRAVERSE and PARDUE 1988; BIESSMANN *et al.* 1990; LEVIS *et al.* 1993). The region that extends from the end of these telomeric arrays into the euchromatin is also structurally conserved across eukaryotes and is referred to as the subtelomere (CHAN and TYE 1983a,b; BROWN *et al.* 1990; WILKIE *et al.* 1991; KARPEN and SPRADLING 1992; LEVIS *et al.* 1993; LOUIS 1995; WALTER *et al.* 1995; FLINT *et al.* 1997; PRYDE *et al.* 1997; MEFFORD and TRASK 2002). The most-distal region of the subtelomere contains a mosaic of several different repeat elements, including arrays of minisatellites and dispersed middle repetitive sequences, some of which are shared by telomeres associated with different chromosome arms. This region is often referred to as the TAS, for *telomere-associated sequences*. We found that the TAS appears to be evolving especially rapidly between closely related *Drosophila* species (J. A. ANDERSON, S. E. CELNIKER and C. H. LANGLEY, unpublished results). Proximal to the TAS is a region of chromosome-unique sequence, which primarily consists of noncoding, unique DNA, although protein-coding regions as well as various middle repetitive elements and satellite DNA have also been found embedded in this region. Although this region appears less dynamic than the TAS, we detected elevated levels of divergence at both the structural and nucleotide levels relative to average divergence estimates from the *Drosophila* euchromatin (J. A. ANDERSON, S. E. CELNIKER and C. H. LANGLEY, unpublished results).

From a population genetics perspective, the subtelomeric portion of the telomere is of primary interest to us, as the mutation rate of the most-distal portion of the telomere (containing telomeric repeats) is so high that the structure of this region is likely to be different among cells of the same individual. Observations by several researchers have suggested that the *D. melanogaster* X chromosome subtelomere is polymorphic. For example, ROBERTS (1979) noted morphological differences at the tips of polytene X chromosomes between different strains of *D. melanogaster*: some telomeres showed

additional, polytenized bands, whereas others appeared to exhibit “terminal deletions.” In addition, AJOKA and EANES (1989) detected differences in the number of P-element insertions present at the X chromosome tip, suggesting that the presence of an insertion hotspot may vary among strains.

What are the mechanisms creating these unusual patterns of subtelomeric polymorphism and divergence? Nonhomologous or ectopic exchange between TAS elements on different chromosomes appears to be important for generating subtelomere variation in both yeast and humans (LOUIS and HABER 1990a,b; MEFFORD and TRASK 2002; LINARDOPOULOU *et al.* 2005). In fact, while levels of homologous recombination are somewhat depressed at yeast telomeres, GOLDMAN and LICHTEN (1996) found that ectopic recombination is significantly more efficient between sequences located near telomeres relative to the rest of the chromosome. Similarly, molecular data suggest that ectopic exchange between human telomeres has occurred a number of times during the ancestry of modern populations (TRASK *et al.* 1998; BAIRD *et al.* 2000). Moreover, it has recently been reported that nonhomologous end joining (NHEJ) is a frequently used double-strand break repair pathway at the ends of human chromosomes (LINARDOPOULOU *et al.* 2005). Interestingly, while yeast exhibits a moderate reduction of homologous recombination in a small region of the telomere, *D. melanogaster* shows a more dramatic, large-scale suppression of crossing over within and proximal to telomeres (LINDSLEY and SANDLER 1977). The tip of the X chromosome (distal 5%), in particular, shows markedly reduced crossing over. Related to this is the population genetics observation that generic, euchromatic regions near the X chromosome telomere exhibit considerably reduced levels of DNA sequence polymorphism relative to regions of the genome with normal levels of crossing over, despite showing typical levels of divergence (*e.g.*, AGUADÉ *et al.* 1989; BEGUN and AQUADRO 1991; AGUADÉ and LANGLEY 1994; LANGLEY *et al.* 2000). For example, at the *Drosophila* X-linked *ewg* locus, where the rate of crossing over is essentially zero, variation is 10-fold lower than regions of normal crossing over per physical length (BRAVERMAN *et al.* 2005; see Figure 2). In contrast to this dearth of polymorphism in the euchromatic region of the X chromosome tip, the morphological data from *Drosophila* (described above) indicate that substantial telomere-associated polymorphism exists distal to *ewg* (Figure 2).

Meiotic drive is another potentially prominent mechanism shaping subtelomeric evolution. NOVITSKI (1951) showed that the distal chromosomal regions can alter a chromosome’s probability of inclusion into the pronucleus of the *Drosophila* oocyte. Using females heterozygous for a chromosome rearrangement that created homologs of unequal lengths, he showed >50% transmission of the shorter homolog. NOVITSKI referred to

this phenomenon as nonrandom disjunction, a form of meiotic drive. His result suggested that natural telomeres may differ in their abilities to orient and move toward the pronucleus during female meiosis. In theory, new mutations that confer a segregation advantage to a particular telomere should rapidly fix, which could lead to elevated levels of divergence between species accompanied by low levels of polymorphism in neighboring regions of the genome (MAYNARD SMITH and HAIGH 1974; KAPLAN *et al.* 1989). More specifically, one can envision a situation where a particular SNP or indel within the subtelomere increases its transmission (and thus its fitness) and linked regions hitchhike along as the SNP or indel rapidly increases in frequency. A similar model of directional selection due to meiotic drive was invoked to explain the rapid evolution seen at centromeric DNA in *Drosophila* (CSINK and HENIKOFF 1998; HENIKOFF *et al.* 2001; MALIK and HENIKOFF 2002). Alternatively, it is possible that the fitness of a particular meiotic driver is condition dependent, leading to the maintenance of several telomeric variants in a population (SANDLER and NOVITSKI 1957; HIRAIZUMI *et al.* 1960; ZWICK *et al.* 1999). For example, a drive element with a meiotic segregation advantage could be associated with a deleterious effect, such as a high rate of chromosome nondisjunction. Under this scenario, such a driver element would rarely reach fixation, and polymorphism could be maintained. Additionally, the relative fitnesses of telomeric variants could vary in a spatial and/or temporal fashion, also leading to the maintenance of polymorphism and possibly to rapid turnover in the structure and composition of telomeric regions.

To characterize the structure and evolution of *Drosophila* subtelomeres more thoroughly, we collected and analyzed molecular population genetics data from the *X* chromosome subtelomere in 87 lines of *D. melanogaster*. Twenty-nine of these lines were of African origin and 58 were of North American origin. We include the African population sample as *D. melanogaster* is thought to have originated in Africa and only very recently spread around the world and into more temperate locations, such as North America (DAVID and CAPY 1988; LACHAISE *et al.* 1988). Similar to observations from yeast and humans, we find that *Drosophila* subtelomeres exhibit high levels of structural and nucleotide polymorphism relative to levels observed in the linked euchromatin. In addition, our data suggest the presence of a localized hotspot of recombination in the most-distal portion of the *X* subtelomere, consistent with observations that recombination plays an important role in the evolution of subtelomeric DNA. Importantly, because we sampled a large number of naturally occurring chromosomes from an outbreeding species, our study not only corroborates previous findings from yeast and humans, but also provides a definitive demonstration of the unique population genetics properties of eukaryotic subtelomeric DNA.

**TABLE 1**  
**Amplicon information**

Amplicon	<i>D. melanogaster</i> coordinates <sup>a</sup>	Length (bp)	No. of North American alleles	No. of African alleles
1	2,311–3,086	794	53	27
2	3,635–4,330	696	55; 3 short	28; 9 short
3	7,423–7,872	450	55	28
4	7,877–8,146	270	55	27
5	12,250–12,690	441	54	29
6	18,887–19,685	799	55	29
7	24,222–25,077	856	54	28
8	29,905–30,825	921	53	28
9	35,395–36,205	811	58	28
10	47,096–47,736	641	56	29
11	57,914–58,713	800	55	29
12	63,301–64,241	941	58	29
13	71,146–71,981	836	52	29
14	76,865–77,445	581	56	29
15	101,037–101,266	230	9	8
16	102,849–103,466	618	8	8
17	110,073–110,790	718	9	8

<sup>a</sup> Coordinates for amplicons 5–17 are from release 4.3; coordinates for amplicons 1–4 correspond to those of AL031884 (*Canton-S*).

## MATERIALS AND METHODS

**Polymorphism data:** We characterized patterns of polymorphism in 17 amplicons located across the *X* chromosome subtelomere, shown in Figure 2 and listed in Table 1. As reported in Table 2, most of the data were from noncoding regions, although one amplicon included coding sequence, and two other amplicons contained UTR sequence. Primers for the most distal of the amplicons (numbered 1–4) were designed using the *X* chromosome sequence from *Canton-S*, originally part of the *Drosophila* Sequencing Project (accession no. AL031884) and sequenced by the European *Drosophila* Genome Project (BENOS *et al.* 2000). As mentioned above, *Canton-S* possesses a structurally distinct *X* subtelomere from the reference strain, *y; cn bw sp*, in that it includes a repetitive region of the subtelomere containing the canonical *X* TAS, characterized by KARPEN and SPRADLING (1992; see Figure 2). Therefore, we were able to assay unique regions of the repetitive portion of the *X* subtelomere with these amplicons (see Table 1 for a list of the amplicons, their lengths, coordinates, and number of alleles sequenced). Data from all 17 amplicons were unambiguously aligned as they were composed largely of unique DNA. Figure 2 illustrates where *Canton-S* begins to show homology with *y; cn bw sp*. For the remaining amplicons, *y; cn bw sp* sequence was used to design primers. The most centromere-proximal region assayed is ~15 kb distal of *ewg*, the most-distal, well-characterized *X*-linked euchromatic gene discussed above (BRAVERMAN *et al.* 2005; Figure 2). In total, 87 different *X* chromosome lines were amplified for most of the 17 amplicons (see Table 1 for sample sizes; see supplemental Table S1 at <http://www.genetics.org/supplemental/> for information on lines). Fifty-eight of the lines were North American (NA) and 29 were African in origin. For the NA lines, 5 were from Napa, California, 50 were from North Carolina, and 3 were laboratory stocks (*Canton-S*, *Oregon-R*, and *y; cn bw sp*). For the African lines, 11 were from Malawi and 18 from Zimbabwe. We constructed a con-

**TABLE 2**  
**Summary statistics across the 17 X chromosome amplicons**

Amplicon	North American $\pi$	African $\pi$	North American Tajima's $D$	African Tajima's $D$	African <i>vs. sim</i> divergence	Coding/UTR (bp)
1	0.00231	0.00190	0.00261	-2.24067*	N/A	N/A
2	0.00512	0.00223	-0.86985	-1.41435	N/A	N/A
3	0.00331	0	2.57980*	N/A	N/A	N/A
4	0.00278	0.00026	1.23805	-1.51420	N/A	N/A
5	0.00305	0.00047	-0.32752	-1.73263*	N/A	N/A
6	0.00125	0.00017	2.25378*	-1.50906	0.06980	N/A
7	0.00215	0.00171	-0.53373	-0.56098	0.07340	N/A
8	0.00203	0.00051	-0.42544	-0.95891	0.12070	N/A
9	0.00013	0.00043	-1.68179	-1.32136	0.06120	Coding, 250
10	0.00120	0.00086	-0.70092	-0.11246	0.16190	N/A
11	0.00146	0.00121	-1.10982	-0.65443	0.08310	N/A
12	0.00049	0	-0.59201	N/A	0.05910	UTR, 333
13	0.00220	0.00017	-0.21356	-1.50906	0.12140	N/A
14	0.00097	0.00076	-0.35397	0.98512	0.09317	N/A
15	0.00255	0	-1.51297	N/A	<i>sim</i> deleted	UTR, 230
16	0.00118	0.00037	0.06935	-1.05482	0.07349	N/A
17	0.0018	0.00080	0.6021	1.34164	0.09483	N/A

N/A, not applicable. \* $P < 0.05$ .

sensus allele of the 6 *D. simulans* lines (using data produced by the Washington University Genome Sequencing Center; <http://www.genome.wustl.edu>) to calculate divergence in amplicons that were clearly homologous between the two species; due to a higher turnover rate of subtelomeric repetitive sequence, these regions were limited to the nonrepetitive, more centromere-proximal portion of the subtelomere (amplicons 5–17; see Figure 2; Table 2). In addition, we sequenced these 6 *D. simulans* lines for 7 of the same centromere-proximal amplicons (amplicons 5–11) to obtain estimates of subtelomeric heterozygosity for this species (to achieve complete coverage of the strains for each of the regions). Three of these *D. simulans* lines originated from the Old World (Africa and Madagascar), and three of them were from the New World.

For the four most-distal amplicons, we failed to achieve amplification for a small subset of the lines. Several different primer pairs and PCR conditions were employed in an attempt to determine whether these failures to amplify represented real structural differences or artifacts. On the basis of these experiments as well as the fact that the amplicons reside within the subtelomeric TAS region, which is known to be evolving quite rapidly between *Drosophila* species (J. A. ANDERSON, S. E. CELNIKER and C. H. LANGLEY, unpublished results), we feel confident that these failures do indeed represent real structural divergence (more below).

**Polymorphism data analysis:** DnaSP 4.0 (ROZAS *et al.* 2003) was used to estimate  $\pi$ , or the average number of pairwise differences per nucleotide, and Tajima's  $D$  statistic. These estimates were based on all sites for each amplicon. Consequently, for three of the amplicons, noncoding sites as well as some coding or UTR sites were included in these estimates (see Tables 1 and 2). In addition, we quantified indel heterozygosity per nucleotide across the entire region.

The programs HapBound-GC and SHRUB-GC (SONG *et al.* 2006) were employed to infer recombination events in the evolutionary history of the chromosomes in our sample. Given a specified maximum gene-conversion tract length, HapBound-GC and SHRUB-GC compute lower and upper bounds, respectively, on the minimum total number of single-crossover and gene-conversion recombinations needed to

derive an input set of SNP sequences. We tried varying the maximum conversion tract length from 0 to 500 and assessed the relative contribution of gene conversions compared to cross-overs. We also used the program LDhat version 2.0 (MCVEAN *et al.* 2004) to study variation in the crossover recombination rate along the X chromosome subtelomere. LDhat uses the composite-likelihood method of HUDSON (2001) in conjunction with the reversible jump MCMC scheme to estimate variable crossover rates. The method requires a lookup table containing two-locus likelihoods. With the population-scaled mutation rate  $\theta = 0.01$ , we generated a lookup table specific to our data set. (We also tried using  $\theta = 0.001$ . The overall pattern of variation in the estimated crossover rate did not differ very much between the two settings of  $\theta$ .) We ran the program for 10 million MCMC updates, ignored the first 1 million updates, and took samples every 3000 updates. As suggested in the manual, we tried a series of block penalties.

## RESULTS

In total, we assayed 11,403 X subtelomeric base pairs in our *D. melanogaster* samples. Table 2 reports summary statistics for each of the amplicons. Estimates were calculated separately for the NA and African data due to evidence for subdivision between the two samples; importantly, there was no evidence of subdivision within either of these samples (data not shown). In addition, the summary estimates are plotted above their corresponding chromosomal locations in Figure 2.

**North America:** Overall, levels of nucleotide heterozygosities in our NA sample are one to two orders of magnitude higher (Wilcoxon test,  $z = -2.77$ ,  $P < 0.01$ ) than those observed in the adjacent euchromatin of the X chromosome tip (see BRAVERMAN *et al.* 2005 and references therein; Table 2). In addition, although heterozygosity appears highest in the most-distal subtelomere

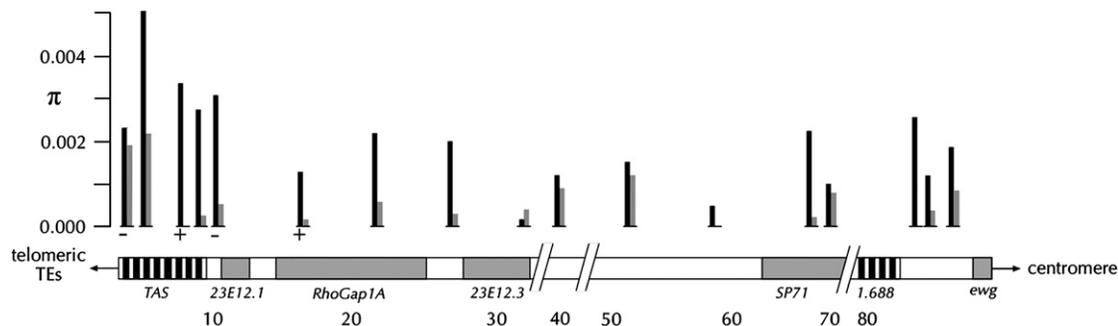


FIGURE 2.—The subtelomeric region of the *D. melanogaster* X chromosome. Expected heterozygosity ( $\pi$ ) in the 14 amplicons are reported above their corresponding locations on the chromosome. Significant positive values of Tajima's  $D$  are represented by a plus (+), while significant negative values are represented by a minus (–). Estimates for African (shading) and NA (solid) samples are reported separately. Annotated genes are shaded and labeled, and the TAS and the 1.688 satellite repeat sequence regions are shown in a striped pattern and labeled.

meric amplicons, it does not taper off appreciably in the more centromere-proximal amplicons, with the exception of those containing protein-coding sequence. In these cases, we observed an  $\sim 10$ -fold decrease in heterozygosity when collectively analyzing both silent and replacement sites (see Figure 2; Table 2).

Tajima's  $D$  measures the deviation of the polymorphic site frequency spectrum from that predicted by the neutral theory for a population in neutral mutation-drift equilibrium (TAJIMA 1989). Tajima's  $D$  was significantly positive in two of the NA subtelomeric amplicons, indicating a skew in the frequency spectrum toward too many common alleles (Table 2; Figure 2). Consistent with this result is our observation that for most of the amplicons, we found several haplotypes segregating at intermediate frequencies. We found considerable haplotype structure across amplicons as well (see below).

We detected 14 polymorphic indels in our data set. This estimate is conservative, as we included only unambiguous instances of this type of polymorphism; we excluded instances with marginal sequence quality or incomplete breakpoint information. In the NA population, the overall expected indel heterozygosity is estimated to be 0.0002. Relative to regions of normal crossing over, this estimate is roughly an order of magnitude lower (e.g., OMETTO *et al.* 2005).

We also discovered structural polymorphism in the distal amplicons; for example, two of the NA lines show copy-number differences at a TAS repeat (2 NA lines and 13 African lines), and three of the more-distal amplicons are not present in other lines (see MATERIALS AND METHODS; Figure 2; Table 2). One such line is the reference strain, *y;cn bw sp*, which does not possess amplicons 1–4 (Figure 2).

**Africa:** Overall, heterozygosity is lower in Africa than in North America (Table 2; Figure 2). Indeed, for most amplicons, the NA sample shows greater than twofold higher levels of heterozygosity, and the distributions of heterozygosity are significantly different between these two samples (Wilcoxon test,  $z = 3.41$ ;  $P < 0.001$ ; Table

2). Similar to the NA trend, levels of polymorphism are highest in more-distal amplicons and do not taper off in the centromere-proximal amplicons. Furthermore, subtelomeric heterozygosity is generally higher than that observed in the adjacent euchromatin (e.g., at *ewg*  $\pi \approx 0.00035$  in Africa; BRAVERMAN *et al.* 2005; see Table 2), although the difference is not significant (Wilcoxon test,  $z = -0.4113$ ;  $P < 0.65$ ).

For most amplicons, the African sample contains one predominant haplotype and one or two rare types, in contrast to the situation in the NA sample where several haplotypes segregate at intermediate frequencies for many of the amplicons. Not surprisingly, then, we detected a significant negative skew in the frequency spectrum in a subset of the African amplicons, and more importantly, Tajima's  $D$  across the subtelomere is significantly higher in the NA sample than in the African sample (Wilcoxon test  $z = -2.084$ ;  $P < 0.05$ ; Table 2; Figure 2).

The African and NA samples exhibit approximately equal levels of expected indel heterozygosity,  $\sim 0.002$  indels/nucleotide. Furthermore, the African sample contains significantly greater levels of subtelomeric structural polymorphism than those observed in the NA sample. For example, the TAS copy-number polymorphism mentioned above and seen in 2 of the 58 NA lines is present in 13 of 29 African lines (see above; Figure 2;  $\chi^2 = 23.2$ ;  $P < 0.001$ ). Similarly, a polymorphic length difference in another of the amplicons is present in just 3 of 58 NA lines, while 9 of the 29 African lines possess it (Figure 2;  $\chi^2 = 10.88$ ;  $P < 0.01$ ). Clearly, these patterns of structural polymorphism are distinct from those found at the nucleotide level (see above).

***D. simulans* polymorphism:** Although *D. simulans* is generally more polymorphic than *D. melanogaster* at the nucleotide level across much of the genome (AQUADRO *et al.* 1988; MORIYAMA and POWELL 1996), we detected just a single SNP in our survey of this species for seven different X subtelomere amplicons in a sample of six chromosomes. However, our data do not rule out the possibility of an increase in subtelomeric heterozygosity

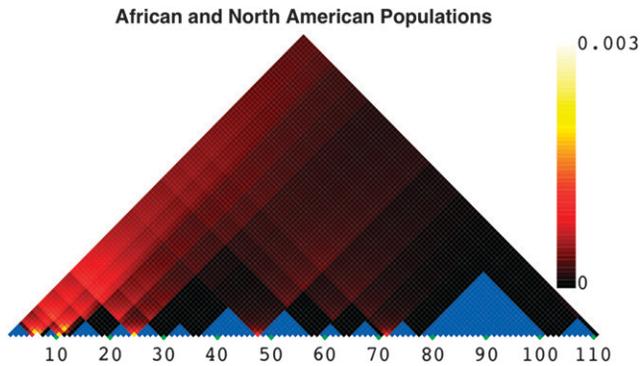


FIGURE 3.—Crossover recombination density plot for both populations of *D. melanogaster*. Each square corresponds to a 1-kb window, beginning at the most-distal end of the Xsubtelomere. Blue squares correspond to regions with no data; yellow squares correspond to the highest density of crossover events; black corresponds to the lowest density of crossover events.

more distally in *D. simulans*. Nevertheless, it is noteworthy that this subtelomeric pattern is comparable to observations from distal euchromatic regions (e.g., MARTIN-CAMPOS *et al.* 1992), in which *D. simulans* also showed less polymorphism than *D. melanogaster*.

We investigated polymorphism over a larger region using *D. simulans* data from the Washington University Genome Sequencing Center's whole genome shotgun (WGS) project (<http://www.genome.wustl.edu>; <http://www.dpgp.org/simulans>). More specifically, we looked at subtelomeric and closely linked euchromatic sequence data at the five major chromosome arms of this species (see supplemental Table S2 at <http://www.genetics.org/supplemental/> for a list of coordinates). Similar to our analysis of *D. melanogaster*, we divided the *D. simulans* WGS data into New World and Old World partitions. Unfortunately, the degree of sequence coverage in these regions is quite variable and often low. However, in general, we find that the Old World sample is more polymorphic than the New World sample across all major chromosome arms (data not shown). This pattern is similar to what has been noted for much of the genome; however, it is the opposite of the pattern that we detected at *D. melanogaster* Xsubtelomeres (Table 2).

**Divergence:** In regions where it could be assayed (amplicons 5–13), estimates of per-site sequence divergence between African *D. melanogaster* and *D. simulans* were up to threefold higher than those calculated from the neighboring euchromatin (see Table 2; 0.056 for *ewg*; BRAVERMAN *et al.* 2005). This result is consistent with our study investigating patterns of subtelomeric molecular evolution between *D. melanogaster* and two of its close relatives, *D. simulans* and *D. yakuba* (J. A. ANDERSON, S. E. CELNIKER and C. H. LANGLEY, unpublished results).

**Recombination:** As described in MATERIALS AND METHODS, we used HapBound-GC and SHRUB-GC to estimate the minimum total number of single-crossover and gene-conversion recombinations needed to derive

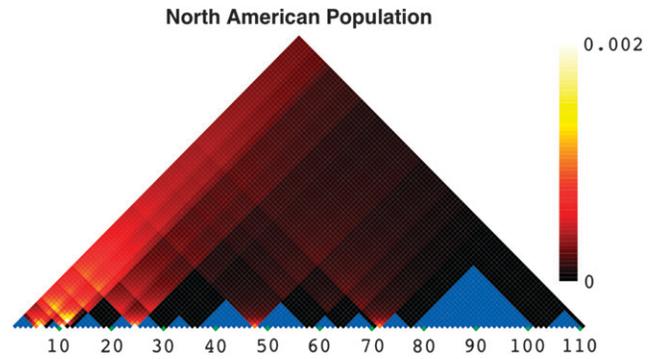


FIGURE 4.—Crossover recombination density plot for the NA data. Each square corresponds to a 1-kb window, beginning at the most-distal end of the Xsubtelomere. Blue squares correspond to regions with no data; yellow squares correspond to the highest density of crossover events; black corresponds to the lowest density of crossover events.

the sequences in our data set. These analyses revealed two main patterns. First, similar to previous reports from the euchromatic regions of the distal X chromosome, we detected evidence of considerable exchange, despite the fact that crossing over is generally thought to be severely reduced (BEGUN and AQUADRO 1995; LANGLEY *et al.* 2000). Second, exchange appears to be heterogeneous across the subtelomere, with the eight most-distal amplicons (the distal-most 31-kb region covering the TAS region) showing evidence of  $\sim 10$ -fold greater exchange than other regions of comparable size. Significant elevation in the level of recombination was observed both in the pooled data and in the data from each sample considered separately. Figures 3 and 4 show the density of detected crossovers when the maximum conversion tract length was set to 0. Only six crossovers were detected for the African data, so its density plot is not shown; about five times as many events were detected for the NA data. Importantly, although the SNP density is higher in this region of elevated recombination activity, Figure 5 suggests that SNP density alone cannot explain the presence of this putative hotspot. In fact, other regions with similar SNP densities exhibit different levels of exchange activity. More proximally, there is variation in the number of inferred recombination events across amplicons; however, no region exhibits as striking levels of inferred exchange as the distal most.

When the maximum gene-conversion tract length was set to 1, about one-third of the detected recombinations were gene-conversion events. Gene-conversion events with tracts containing a single SNP are generally indistinguishable from recurrent mutations, but it seems meaningful to note that about two-thirds of those gene-conversion events occurred in the distal-most 31-kb region discussed above. When the maximum tract length was increased to 500,  $\sim 60\%$  of all detected recombinations were gene-conversion events, about two-thirds of which had conversion tracts containing at least two SNPs. Again, about two-thirds of the total gene-conversion

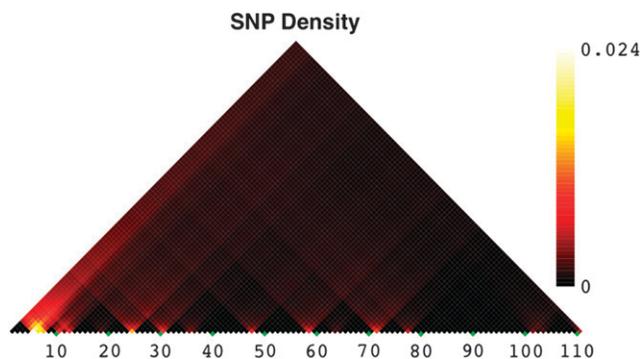


FIGURE 5.—SNP density plot using information from both populations of *D. melanogaster*. Each square corresponds to a 1-kb window, starting at the most-distal end of the X subtelomere. Yellow corresponds to the highest SNP density; black corresponds to the lowest SNP density.

events occurred in the distal-most 31-kb region. Although it is not possible to infer gene-conversion rates from these results, we can conclude that gene conversion is likely to have played an important role in the evolution of the X subtelomere.

We also estimated the crossover recombination rate per kilobase using the program LDhat. Figure 6 shows the inferred crossover rates (estimated as  $4N_e r$ /kb) plotted along the X chromosome subtelomere. Similar to the results described above, a hotspot of crossovers is apparent in the most-distal portion of the subtelomere, with the NA data showing higher levels of recombination activity than the African data. The NA data contain a distinguished crossover hotspot in the TAS region (the region to the left of  $\sim 10$  kb in Figure 6), while such a strong hotspot is absent in the African data. This is consistent with the fact that HapBound-GC and SHRUB-GC detected neither crossovers nor gene conversions in the TAS region for the African data.

As mentioned above, *Drosophila* shows a marked suppression of crossing over near the tip of the X chromosome, spanning megabases of sequence (LINDSLEY and SANDLER 1977). In contrast, our subtelomeric data come from a physical region of  $\sim 100$  kb. Therefore, our finding of an extremely localized hotspot of recombination in the distal subtelomere should not be interpreted as evidence that the large-scale suppression of crossing over is relieved in the subtelomere. Indeed, the scales of the two patterns are so different that it is difficult, and potentially misleading, to compare them.

## DISCUSSION

Our investigation of subtelomeric polymorphism in *D. melanogaster* was motivated by several pieces of data demonstrating the functionally and evolutionarily unique properties of these eukaryotic genomic regions. Indeed, the elevated polymorphism, divergence, and rates of recombination observed in yeast and primates strongly suggest that the mutational and population genetic

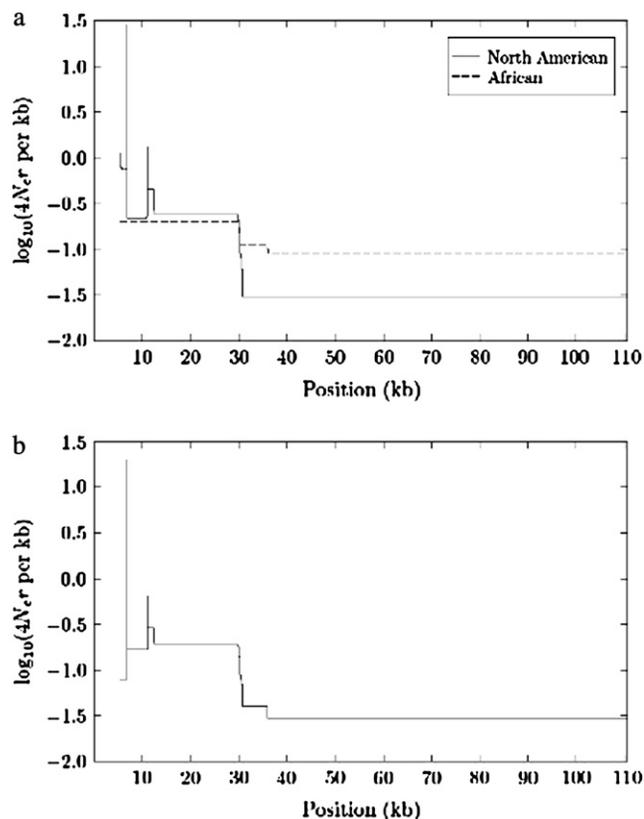


FIGURE 6.—Crossover recombination rates estimated using LDhat version 2.0. Log-transformed estimated rates are plotted along the X chromosome subtelomere. The distal-most region clearly shows an elevated level of recombination activity. In particular, the North American population possesses a distinguished crossover hotspot in the TAS region (to the left of position 10 kb). (a) Separate estimates for each of the NA and African data sets. (b) Estimates for the combined NA and African data.

processes governing subtelomeric polymorphism and divergence are different from those thought to be important in the evolution of euchromatin. Our data from nearly 90 *D. melanogaster* X chromosomes definitively demonstrate the generality of these patterns.

Relative to the neighboring euchromatin, the *D. melanogaster* X subtelomere shows high levels of nucleotide and structural polymorphism as well as divergence. This is consistent with the idea that the subtelomeric region has a higher mutation rate than the euchromatin. Indeed, studies involving yeast and humans indicate that variation in subtelomeric DNA is likely to be generated through distinct mechanisms (for a review, see PRYDE *et al.* 1997; MEFFORD and TRASK 2002; also see LINARDOPOULOU *et al.* 2005). For example, both ectopic recombination and NHEJ have been shown to be important repair pathways in subtelomeric regions (LOUIS and HABER 1990a,b; MEFFORD and TRASK 2002; LINARDOPOULOU *et al.* 2005). Similarly, it was recently found that sister-chromatid exchange during replication may be a frequent contributor to the generation of structural variation in human subtelomeres (RUDD *et al.* 2007).

We also discovered evidence of considerable amounts of recombination in the most-distal regions of the X chromosome in the history of the sampled X chromosomes, consistent with previous studies at the tip of the X (BEGUN and AQUADRO 1995; LANGLEY *et al.* 2000). Gene conversion has been invoked to explain these patterns and in fact our analysis indicates that much of the recombination in the subtelomeric region of the X chromosome can be attributed to gene conversion. In addition, our data suggest the possibility that a recombination hotspot exists in the distal subtelomere. Such heterogeneity on a relatively small physical scale has not previously been detected in *D. melanogaster* and thus may be unique to the telomeric regions. Given the existence of markers in this genomic region, our hotspot hypothesis should be testable through genetic experiments.

The most striking aspect of the data presented here is that the African sample is consistently and significantly less heterozygous than the NA sample. This pattern stands in stark contrast to data from the euchromatin, in which African X chromosomes have been found to exhibit higher nucleotide heterozygosity than non-African X chromosomes (*e.g.*, BEGUN and AQUADRO 1993; LANGLEY *et al.* 2000; ANDOLFATTO 2001). In addition, African subtelomeres are characterized by a skew in the frequency spectrum toward rare alleles and are generally segregating one major haplotype and several minor types over the scale of amplicons. Alternatively, NA amplicons are characterized by multiple, intermediate-frequency haplotypes. Although nonequilibrium demography associated with colonization of non-African habitats (DAVID and CAPY 1988; LACHAISE *et al.* 1988) could explain the more intermediate frequencies of variants in the NA sample, this hypothesis predicts that the common NA alleles should also be common in Africa. This is generally not observed in our data, as most of the common NA SNPs are not segregating at appreciable frequencies in the sampled African populations.

Another explanation for our finding of reduced African *vs.* NA subtelomeric heterozygosity is that some form of subtelomere-specific selection acts differently in these two geographic regions. Two possible models of subtelomeric selection are (1) subtelomeric variants compete for inclusion into the oocyte pronucleus during female meiosis (*sensu* NOVITSKI 1951, as described above) and (2) subtelomeres function to modify the pleiotropic, meiotic effects of chromosomal inversion heterozygosity. We acknowledge that the following discussion is highly speculative; however, this striking pattern deserves comment.

Competition for preferential meiotic transmission may drive the sequential fixation of different subtelomeric types if the fitness landscape is such that only a single type is most fit at any particular point in time. As different subtelomeric variants enter the population, relative fitnesses may shift, allowing other variants to sweep through. Such a scenario of directional selection

due to meiotic drive (or nonrandom disjunction, as defined above) is consistent with our observations of reduced polymorphism in Africa (relative to NA populations) and a negative skew in the frequency spectrum (AGUADÉ *et al.* 1989; BRAVERMAN *et al.* 1995). Perhaps such processes are not occurring in NA populations, although the explanation for why this may be is not obvious. Alternatively, NA subtelomeres may be experiencing balancing selection where the fitness of subtelomeres varies over time so that a variant rarely fixes, consistent with the presence of several common subtelomeres in a population (SANDLER and NOVITSKI 1957; see above).

Under a model of subtelomeric selection due to meiotic drive, we should be able to detect genetic variation in the abilities of different natural subtelomeric variants to compete against a tester chromosome in nonrandom disjunction experiments similar to those described in NOVITSKI (1951; and see above). More specifically, we expect the performance of the most common African variant (against a tester chromosome) to be the highest, and we expect its performance to be stronger in an African *vs.* NA genetic background. Indeed, the fitnesses of subtelomeric variants should be higher in the presence of co-adapted modifiers. Similarly, the relative competitive abilities of the common NA subtelomeric variants are expected to differ between assays involving a NA *vs.* African background. However, functional differences among the common NA subtelomeric variants may not be easily revealed in a simple nonrandom disjunction assay if competitiveness in meiotic drive is strongly dependent on genetic background or specific environmental parameters.

Alternatively, an obvious genetic difference between African and NA populations possibly related to subtelomeric variation is that African populations are highly polymorphic for large numbers of inversions, many of which are endemic (for a review, see LEMEUNIER and AULARD 1992). In contrast, only four common inversions, one on each of the autosomal chromosome arms, have been found in NA (and in most temperate) populations, and the frequencies of these “temperate” inversions are known to vary clinally with latitude (*e.g.*, METTLER *et al.* 1977). Interestingly, virtually all of the *D. melanogaster* inversions are a single mutational step away from the ancestral standard karyotype, which strongly suggests that despite the high diversity associated with African inversions, essentially all of them have recently invaded the species (reviewed in LEMEUNIER and AULARD 1992).

When paired with a standard chromosome during female meiosis, inversions alter the frequency and distribution of crossover events (STURTEVANT and BEADLE 1936; NOVITSKI and BRAVER 1954; THEURKAUF and HAWLEY 1992). Because the location of a particular chiasma can impact the fidelity of chromosome segregation, inversion heterozygosity has a broad and strong

influence on the outcome of meiotic transmission (KOEHLER *et al.* 1996). If subtelomeres are modifiers of the meiotic effects of inversion heterozygosity, then we might expect to see evidence of co-evolution between inversions and subtelomeres. Perhaps the recent invasion of large numbers of autosomal inversions in African *D. melanogaster* populations has led to the sequential fixation of subtelomeric variants along with associated hitchhiking effects, consistent with our observation of lower levels of African subtelomeric diversity (relative to NA subtelomeric diversity) accompanied by a negative skew in the frequency distribution (BRAVERMAN *et al.* 1995). More specifically, one can imagine a scenario in which the most-fit subtelomeric variant can accommodate the meiotic effects of the largest range of different inversion constellations encountered in African populations. As inversion frequencies shift over time, the relative fitnesses of different genotypes should also change, leading to recurrent hitchhiking events as different subtelomeric variants sweep to fixation. A prediction of this model is that we should find distinct, dominant (*i.e.*, most common) subtelomeric types in different African regions harboring distinct inversion profiles. In addition, although the explanation for why most inversions have not been able to colonize more temperate populations remains mysterious, the fact that inversion diversity in these populations is dramatically reduced relative to African diversity could indicate, according to this model, that the selective environment for subtelomeric variants is significantly different in NA (and other temperate) populations. Indeed, perhaps the particular subtelomeres that are common in NA populations represent specific adaptations for dealing with the effects of the relatively few common inversion karyotypes on meiosis. If this is the case, we would expect to see evidence of subtelomeric clines in temperate populations of *D. melanogaster*, which parallel the documented inversion clines. In fact, preliminary subtelomeric data from northern and southern Australian populations show high levels of differentiation relative to genomic averages (T. TURNER, M. LEVINE and D. BEGUN, unpublished data), consistent with clinal variation.

Although these models of subtelomeric selection are speculative, they make clear experimental predictions regarding the fitness effects (including competitiveness in meiotic drive) of the African and NA subtelomeres in the context of (1) NA *vs.* African genetic backgrounds and (2) temperate *vs.* African karyotypes. The inversion model also predicts a general correlation between inversion variation and subtelomeric variation (*i.e.*, different combinations of inversions favor different subtelomeres) that can be tested in several *Drosophila* lineages. In addition, assays testing the meiotic drive model can be extended to include other populations of *D. melanogaster*, which are likely to further inform our understanding of how selection due to meiotic drive might influence subtelomeric patterns of polymorphism.

We thank Charis Marston for her invaluable assistance with collecting the data. We also thank Kristian Stevens and Alisha Holloway for help with computational aspects of the project. Dave Begun and Sergey Nuzhdin provided useful comments on the manuscript. J.A.A. was supported by a National Science Foundation Predoctoral Graduate Fellowship. Portions of this research were supported by National Institutes of Health grants R01-HG002942 to C.H.L. and 1K99-GM080099 to Y.S.S.

#### LITERATURE CITED

- AGUADÉ, M., and C. H. LANGLEY, 1994 Polymorphism and divergence in regions of low recombination in *Drosophila*, pp. 67–76 in *Non-neutral Evolution*, edited by B. GOLDING. Chapman & Hall, New York.
- AGUADÉ, M., Y. MIYASHITA and C. H. LANGLEY, 1989 Reduced variation in the *yellow-achaete-scute* region in natural populations of *Drosophila melanogaster*. *Genetics* **122**: 607–615.
- AGUADÉ, M., W. MEYERS, A. D. LONG and C. H. LANGLEY, 1994 Single-strand nucleotide variation in *Drosophila melanogaster* X chromosome. *Proc. Natl. Acad. Sci. USA* **91**: 4658–4662.
- AJOKA, J. W., and W. F. EANES, 1989 The accumulation of P-elements on the tip of the X chromosome in populations of *Drosophila melanogaster*. *Genet. Res.* **53**: 1–6.
- ANDOLFATTO, P., 2001 Contrasting patterns of X-linked and autosomal nucleotide variation in *Drosophila melanogaster* and *Drosophila simulans*. *Mol. Biol. Evol.* **18**: 279–290.
- AQUADRO, C. F., K. M. LADO and W. A. NOON, 1988 The rosy region of *Drosophila melanogaster* and *Drosophila simulans*. I. Contrasting levels of naturally occurring restriction map variation and divergence. *Genetics* **119**: 875–888.
- BAIRD, D. M., J. COLEMAN, Z. H. ROSSER and N. J. ROYLE, 2000 High levels of sequence polymorphism and linkage disequilibrium at the telomere of 12q: implications for telomere biology and human evolution. *Am. J. Hum. Genet.* **66**: 235–250.
- BEGUN, D. J., and C. F. AQUADRO, 1991 Molecular population genetics of the distal portion of the X chromosome in *Drosophila*: evidence for genetic hitchhiking of the yellow-achaete region. *Genetics* **129**: 1147–1158.
- BEGUN, D. J., and C. F. AQUADRO, 1993 African and North American populations of *Drosophila melanogaster* are very different at the DNA level. *Nature* **365**: 548–550.
- BEGUN, D. J., and C. F. AQUADRO, 1995 Evolution at the tip and base of the X chromosome in an African population of *Drosophila melanogaster*. *Mol. Biol. Evol.* **12**: 382–390.
- BEGUN, D. J., A. K. HOLLOWAY, K. STEVENS, L. W. HILLIER, Y. P. POH *et al.*, 2007 Population genomics: whole-genome analysis of polymorphism and divergence in *Drosophila simulans*. *PLoS Biol.* **5**: e310.
- BENOS, P. V., M. K. GATT, L. MURPHY, D. HARRIS, B. BARRELL *et al.*, 2000 From sequence to chromosome: the tip of the X chromosome of *D. melanogaster*. *Science* **24**: 2220–2222.
- BIESSMANN, H., S. B. CARTER and J. M. MASON, 1990 Chromosome ends in *Drosophila* without telomeric DNA sequences. *Proc. Natl. Acad. Sci. USA* **87**: 1758–1761.
- BLACKBURN, E. H., 1992 Telomerases. *Annu. Rev. Biochem.* **61**: 113–129.
- BRAVERMAN, J. M., R. R. HUDSON, N. L. KAPLAN, C. H. LANGLEY and W. STEPHAN, 1995 The hitchhiking effect on the site frequency spectrum of DNA polymorphisms. *Genetics* **140**: 783–796.
- BRAVERMAN, J. M., B. P. LAZZARO, M. AGUADÉ and C. H. LANGLEY, 2005 DNA sequence polymorphism and divergence at the *erect wing* and *suppressor of sable* loci of *Drosophila melanogaster* and *D. simulans*. *Genetics* **170**: 1153–1165.
- BROWN, W. R. A., P. J. MACKINNON, A. VILLASANTE, N. SPURR, V. BUCKLE *et al.*, 1990 Structure and polymorphism of human telomere-associated DNA. *Cell* **63**: 119–132.
- BUSTAMANTE, C. D., A. FLEDEL-ALON, S. WILLIAMSON, R. NIELSEN, M. T. HUBISZ *et al.*, 2005 Nature selection on protein-coding genes in the human genome. *Nature* **437**: 1153–1157.

- CHAN, C. S., and B. K. TYE, 1983a A family of *Saccharomyces cerevisiae* repetitive autonomously replicating sequences that have very similar genomic environments. *J. Mol. Biol.* **168**: 505–523.
- CHAN, C. S., and B. K. TYE, 1983b Organization of DNA sequences and replication origins at yeast telomeres. *Cell* **33**: 563–573.
- CSINK, A. K., and S. HENIKOFF, 1998 Something from nothing: the evolution and utility of satellite repeats. *Trends Genet.* **14**: 200–204.
- DAVID, J. R., and P. CAPY, 1988 Genetic-variation of *Drosophila melanogaster* natural populations. *Trends Genet.* **4**: 106–111.
- FLINT, J., G. P. BATES, C. CLARK, D. WILLINGHAM, B. A. ROE *et al.*, 1997 Sequence comparison of human and yeast telomeres identifies structurally distinct subtelomeric domains. *Hum. Mol. Genet.* **6**: 1305–1314.
- GOLDMAN, A. S. H., and M. LICHTEN, 1996 The efficiency of meiotic recombination between dispersed sequences in *Saccharomyces cerevisiae* depends upon their chromosomal location. *Genetics* **144**: 43–55.
- HENIKOFF, S., K. AHMAD and H. S. MALIK, 2001 The centromere paradox: stable inheritance with rapidly evolving DNA. *Science* **293**: 1098–1102.
- HIRAIZUMI, Y., L. SANDLER and J. E. CROW, 1960 Meiotic drive in natural populations of *Drosophila melanogaster*. 3. Populational implications of the distorter locus. *Evolution* **14**: 433–444.
- HUDSON, R. R., 2001 Two-locus sampling distributions and their application. *Genetics* **159**: 1805–1817.
- KAPLAN, N. L., R. R. HUDSON and C. H. LANGLEY, 1989 The “hitch-hiking effect” revisited. *Genetics* **123**: 887–899.
- KARPEN, G. H., and A. C. SPRADLING, 1992 Analysis of subtelomeric heterochromatin in the *Drosophila* minichromosome *Dp1187* by single *P* element insertional mutagenesis. *Genetics* **132**: 737–753.
- KOEHLER, K. E., C. L. BOULTON, H. E. COLLINS, R. L. FRENCH, K. C. HERMAN *et al.*, 1996 Spontaneous X chromosome MI and MII nondisjunction events in *Drosophila melanogaster* oocytes have different recombinational histories. *Nat. Genet.* **14**: 406–414.
- KUO, H. F., K. M. OLSEN and E. J. RICHARDS, 2006 Natural variation in a subtelomeric region of Arabidopsis: implications for the genomic dynamics of a chromosome end. *Genetics* **173**: 401–417.
- LACHAISE, D., M. L. CARIOU, J. R. DAVID, F. LEMEUNIER, L. TSACAS *et al.*, 1988 Historical biogeography of the *Drosophila melanogaster* species subgroup. *Evol. Biol.* **22**: 159–225.
- LANGLEY, C. H., B. P. LAZZARO, W. PHILLIPS, E. HEIKKINEN and J. M. BRAVERMAN, 2000 Linkage disequilibrium and the site frequency spectra of the *su(s)* and *s(wa)* regions of the *Drosophila melanogaster* X chromosome. *Genetics* **156**: 1837–1852.
- LEMEUNIER, F., and S. AULARD, 1992 Inversion polymorphism in *Drosophila melanogaster*, pp. 339–405 in *Drosophila Inversion Polymorphism*, edited by C. B. KRIMBAS and J. R. POWELL. CRC Press, Boca Raton, FL.
- LEVIS, R. W., R. GANESAN, K. HOUTCHENS, L. A. TOLAR and F. SHEEN, 1993 Transposons in place of telomeric repeats. *Cell* **75**: 1083–1093.
- LINARDOPOULOU, E. V., E. M. WILLIAMS, Y. X. FAN, C. FRIEDMAN, J. M. YOUNG *et al.*, 2005 Human subtelomeres are hot spots of interchromosomal recombination and segmental duplication. *Nature* **437**: 94–100.
- LINDSLEY, D. L., and L. SANDLER, 1977 Genetic analysis of meiosis in female *Drosophila melanogaster*. *Philos. Trans. R. Soc. Lond. Ser. B* **277**: 295–312.
- LOUIS, E. J., 1995 The chromosome ends of *Saccharomyces cerevisiae*. *Yeast* **11**: 1553–1573.
- LOUIS, E. J., and J. E. HABER, 1990a The subtelomeric *Y'* repeat family in *Saccharomyces cerevisiae*: an experimental system for repeated sequence evolution. *Genetics* **124**: 533–545.
- LOUIS, E. J., and J. E. HABER, 1990b Mitotic recombination among subtelomeric *Y'* repeats in *Saccharomyces cerevisiae*. *Genetics* **124**: 547–559.
- MALIK, H. S., and S. HENIKOFF, 2002 Conflict begets complexity: the evolution in centromeres. *Curr. Opin. Genet. Dev.* **12**: 711–718.
- MARTIN-CAMPOS, J. M., J. M. COMERON, N. MIYASHITA and M. AGUADÉ, 1992 Intraspecific and interspecific variation at the *y-ac-sc* region of *Drosophila simulans* and *Drosophila melanogaster*. *Genetics* **130**: 805–816.
- MAYNARD SMITH and HAIGH, 1974 The hitch-hiking effect of a favourable gene. *Genet. Res.* **23**: 23–35.
- MCVEAN, G., S. MYERS, S. HUNT, P. DELOUKAS, D. BENTLEY *et al.*, 2004 The fine-scale structure of recombination rate variation in the human genome. *Science* **304**: 581–584.
- MEFFORD, H. C., and B. J. TRASK, 2002 The complex structure and dynamic evolution of human subtelomeres. *Nat. Rev. Genet.* **3**: 91–102.
- METTLER, L. E., R. A. VOELKER and T. MUKAI, 1977 Inversion clines in populations of *Drosophila melanogaster*. *Genetics* **87**: 169–176.
- MORIYAMA, E. N., and J. R. POWELL, 1996 Intraspecific nuclear DNA variation in *Drosophila*. *Mol. Biol. Evol.* **13**: 261–277.
- NIELSEN, R., C. BUSTAMANTE, A. G. CLARK, S. GLANOWSKI, T. B. SACKTON *et al.*, 2005 A scan for positively selected genes in the genomes of humans and chimpanzees. *PLoS Biol.* **3**: 976–985.
- NOVITSKI, E., 1951 Non-random disjunction in *Drosophila*. *Genetics* **36**: 267–280.
- NOVITSKI, E., and G. BRAVER, 1954 An analysis of crossing-over within a heterozygous inversion in *Drosophila melanogaster*. *Genetics* **39**: 197–209.
- OMETTO, L., W. STEPHAN and D. DE LORENZO, 2005 Insertion/deletion and nucleotide polymorphism data reveal constraints in *Drosophila melanogaster* introns and intergenic regions. *Genetics* **169**: 1521–1527.
- PRYDE, F. E., H. C. GORHAM and E. J. LOUIS, 1997 Chromosome ends: all the same undercaps. *Curr. Biol.* **7**: 822–828.
- RENKAWITZ-POHL, R., and S. BIALOJAN, 1984 A DNA sequence of *Drosophila melanogaster* with a differential telomeric distribution. *Chromosoma* **89**: 206–211.
- ROBERTS, P. A., 1979 Rapid change of chromomeric and pairing patterns of polytene chromosome tips in *D. melanogaster*: Migration of polytene-non-polytene transition zone? *Genetics* **92**: 861–868.
- 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.
- RUBIN, G. M., 1978 Isolation of telomeric DNA sequence from *Drosophila melanogaster*. *Cold Spring Harbor Symp. Quant. Biol.* **42**: 1041–1046.
- RUDD, M. K., C. FRIEDMAN, S. S. PARGHI, E. V. LINARDOPOULOU, L. HSU *et al.*, 2007 Elevated rates of sister chromatid exchange at chromosome ends. *PLoS Genet.* **3**: 319–323.
- SANDLER, L., and E. NOVITSKI, 1957 Meiotic drive as an evolutionary force. *Am. Nat.* **91**: 105–110.
- SONG, Y. S., Z. DING, D. GUSFIELD, C. H. LANGLEY and Y. WU, 2006 Algorithms to distinguish the role of gene-conversion from single-cross-over recombination in the derivations of SNP sequences in populations: Proceedings of Research in Computational Molecular Biology (RECOMB). *Lect. Notes Comput. Sci.* **3909**: 231–245.
- 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**: 554–604.
- TAJIMA, F., 1989 Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585–595.
- THEURKAUF, W. E., and R. S. HAWLEY, 1992 Meiotic spindle assembly in *Drosophila* females: behavior of nonexchange chromosomes and the effects of mutations in the nod kinesin-like protein. *J. Cell Biol.* **116**: 1167–1180.
- TRASK, B. J., C. FRIEDMAN, A. MARTIN-GALLARDO, L. ROWEN, C. AKINBAMI *et al.*, 1998 Members of the olfactory receptor gene family are contained in large blocks of DNA duplicated polymorphically near the ends of human chromosomes. *Hum. Mol. Genet.* **7**: 13–26.
- TRAVERSE, K. L., and M.-L. PARDUE, 1988 A spontaneously opened ring chromosome of *Drosophila melanogaster* has acquired HeT DNA sequences at both new telomeres. *Proc. Natl. Acad. Sci. USA* **85**: 8116–8120.
- WALTER, M. F., C. JANG, B. KASRAVI, B. M. MECHLER, J. M. MASON *et al.*, 1995 DNA organization and polymorphism of a wild-type *Drosophila* telomere region. *Chromosoma* **104**: 229–241.
- WILKIE, A. O. M., D. R. HIGGS, K. A. RACK, V. J. BUCKLE, N. K. SPURR *et al.*, 1991 Stable length polymorphism of up to 260 at the tip of the short arm of human chromosome 16. *Cell* **64**: 595–606.
- YASUHARA, J. C., and B. T. WAKIMOTO, 2006 Oxymeron no more: the expanding world of heterochromatic genes. *Trends Genet.* **22**: 330–338.

YOUNG, B. S., A. PESSON, K. L. TRAVERSE, C. FRENCH and M. L. PARDUE, 1983 Telomere regions in *Drosophila* share complex DNA sequences with pericentric heterochromatin. *Cell* **34**: 85–94.

ZWICK, M. E., J. L. SALSTROM and C. H. LANGLEY, 1999 Genetic variation in rates of nondisjunction: association of two naturally

occurring polymorphisms in the chromokinesin nod with increased rates of nondisjunction in *Drosophila melanogaster*. *Genetics* **152**: 1605–1614.

Communicating editor: A. D. LONG