DOI: 10.1534/genetics.108.093807

Molecular Population Genetics and Evolution of Drosophila Meiosis Genes

Jennifer A. Anderson,*,[1] William D. Gilliland†,[1,2] and Charles H. Langley*

*Department of Evolution and Ecology, University of California, Davis, California 95616 and †Stowers Institute for Medical Research, Kansas City, Missouri 64110

Manuscript received July 9, 2008
Accepted for publication October 28, 2008

ABSTRACT

While many functional elements of the meiotic process are well characterized in model organisms, the genetic basis of most of the natural phenotypic variation observed in meiotic pathways has not been determined. To begin to address this issue, we characterized patterns of polymorphism and divergence in the protein-coding regions of 33 genes across 31 lines of *Drosophila melanogaster* and 6 lines of *Drosophila simulans*. We sequenced genes known to be involved in chromosome segregation, recombination, DNA repair, and related heterochromatin binding. As expected, we found several of the genes to be highly conserved, consistent with purifying selection. However, a subset of genes showed patterns of polymorphism and divergence typical of other types of natural selection. Moreover, several intriguing differences between the two Drosophila lineages were evident: along the *D. simulans* lineage we consistently found evidence of adaptive protein evolution, whereas along the *D. melanogaster* lineage several loci exhibited patterns consistent with the maintenance of protein variation.

MEIOSIS is an essential part of sexual reproduction. While many meiotic pathways and genes are highly conserved across distantly related eukaryotes, others appear to diverge quite rapidly, even among individuals of the same species (for a review, see Gerton and Hawley 2005). For example, chiasmata are required for the proper segregation of meiotic homologs in most sexual organisms, and the main proteins mediating this process, such as the homologs of *spo11* and *Rad51*, are ubiquitous (Zickler and Kleckner 1999). However, alternative systems that successfully segregate chromosomes in the absence of crossing over have evolved in several species (*e.g.*, White 1973; Hawley *et al.* 1992), and it is possible that the evolution of lineage-specific genes has accompanied the appearance of these achiasmate pathways.

Recent studies have provided evidence that sequence variation in genes regulating meiosis can explain a portion of the variation in the fidelity of chromosome segregation found in natural populations. For example, Zwick *et al.* (1999) detected genetic variation in rates of nondisjunction among naturally occurring variants of *nod*, a chromokinesin required for achiasmate chromosome segregation in *Drosophila melanogaster* (Baker and Carpenter 1972; Carpenter 1973; Zhang and Hawley 1990; Zhang *et al.* 1990; Afshar *et al.* 1995a, b; Coop *et al.* 2008; Kong *et al.* 2008). In particular, they found two intermediate-frequency variants that were significantly associated with increased levels of female achiasmate nondisjunction in a sensitized background. To resolve this paradox (apparently deleterious alleles segregating at intermediate frequencies), they invoked the presence of true meiotic drive elements and their modifiers.

True meiotic drive can occur when meiosis is asymmetric, in the sense that all of the meiotic products are not included in the pronucleus of a gamete, as is the case in female multicellular eukaryotes (Sandler and Novitski 1957). Any variant that can increase a chromosome's probability of becoming part of the pronucleus therefore confers an immediate selective advantage. Zwick *et al.* (1999) argued that if a drive element (or its modifier) confers a deleterious pleiotropic effect, such as a high rate of chromosome nondisjunction, its selective advantage may be outweighed by its deleterious effect when it reaches intermediate or high frequencies in a population. Indeed, it is expected that many components of the meiotic pathway may experience strong selective pressure for faithful and accurate transmission of meiotic products every generation. Such factors are predicted to affect the relative fitness of a drive element by suppressing its segregation advantage (Sandler and Novitski 1957; Hiraizumi *et al.* 1960; Charlesworth and Hartl 1978; Zwick *et al.* 1999). Under these circumstances, a drive element would rarely fix, and a high level of nondisjunction, along with polymorphism at the driving and modifier loci, such as that observed for *nod*, would be maintained.

This interplay between meiotic drive and fidelity has also been invoked to explain the rapid evolution of Drosophila centromeric DNA (Henikoff *et al.* 2001).

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under accession nos. FJ218497–FJ219586.

[1]These authors contributed equally to this work.

[2]*Corresponding author:* Stowers Institute for Medical Research, 1000 E. 50th St., Kansas City, MO 64110. E-mail: WDG@stowers-institute.org

The observation of a high rate of turnover in centromere satellite composition accompanied by rapid protein evolution at the centromere-specific histone, *cid*, led investigators to propose a coevolution scenario, wherein centromere sequence evolution, spurred by meiotic drive, is tracked by evolution at the *cid* locus (Csink and Henikoff 1998; Malik and Henikoff 2001; Malik and Henikoff 2002). Indeed, genetic evidence from several species has demonstrated that centromeres can differ in their relative "strengths" during meiosis (*e.g.*, Novitski 1955; Dawe and Cande 1996; Pardo-Manuel De Villena and Sapienza 2001). If the meiotic competition among centromeres produces deleterious effects, such as increased rates of nondisjunction, that can be suppressed to some extent by the protein encoded by *cid*, then periodic selective sweeps at both the centromeric sequence and the coevolving *cid* locus are predicted (Henikoff *et al.* 2001).

In addition, Novitski (1951) provided evidence that 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 (1951) referred to this phenomenon as nonrandom disjunction, a form of true meiotic drive. His result suggested that natural telomeres may also differ in their abilities to orient and move toward the pronucleus during female meiosis.

Despite these compelling examples, the magnitude and genetic basis of naturally occurring phenotypic variation in meiotic pathways has not been determined. For example, within a species the rate of crossing over can vary dramatically among different regions of the genome, among different sexes, and among individuals of the same sex (*e.g.*, Lindsley and Sandler 1977; Sniegowski *et al.* 1994; Takano-Shimizu 1999; Lynn *et al.* 2004). Moreover, the frequency and distribution of chiasmata may vary among individuals and between closely related species (Ohnishi and Voelker 1979; True *et al.* 1996; Ptak *et al.* 2005; Coop *et al.* 2008; Kong *et al.* 2008). Although genetic analysis of mutants has identified many of the genes required for meiosis, little is known about their patterns of polymorphism and divergence.

To provide a sound population genetics foundation for the investigation of meiotic processes shaping variation in eukaryotic genomes, we have characterized patterns of polymorphism and divergence in the protein-coding regions of 33 meiotic genes across 21 lines of North American *D. melanogaster*, 10 lines of African *D. melanogaster*, and 6 lines of *Drosophila simulans*. We include the African population sample as *D. melanogaster* is thought to have originated in Africa and only very recently has spread around the world and into more temperate locations, such as North America (David and Capy 1988; Lachaise *et al.* 1988). We therefore expect to capture a more complete picture of variation from our African sample, and we expect this sample to be at or close to demographic equilibrium.

TABLE 1

Genes included in the survey and their functional information

Gene	Pathway	Class
ald	CS	Ortholog
asp	CS	*ASPM*
Axs	AS	
c(2)M	DSBF/CR	LS
c(3)G	DSBF/CR	Ortholog
cav	HB, TM	LS
CG7676	C3G localization	LS
Su(var)205	HB	*HP1*
Klp3A	CS/CR	*KIF4*
Ku70	Rec/Rep/TM	*KU70*
Ku80	Rec/Rep/TM	*KU80*
mei-218	CR	LS
mei-41	CR	*ATR*
mei-P22	DSBF	LS
mei-S332	CS	LS
mei-P26	CR	
mei-W68	DSBF	*SPO11*
mei-9	Rec/Rep	*ERCC4*
mre11	TM	*MRE11*
mtrm	CS	LS
mus304	Rec/Rep	*ATRIP*
ncd	AS	
okr	Rec/Rep	*RAD54*
ord	CS	LS
polo	SA/CS	
rad50	TM	*RAD50*
smc1	SCC	*SMC1A*
spn-A	Rec/Rep	*RAD51*
spn-B	Rec/Rep	*XRCC3*
spn-D	Rec/Rep	
subito	SA/CS	
teflon	CS	LS
tefu	TM	

CS, chromosome segregation; AS, achiasmate segregation; CR, crossover regulation, DSBF, double-strand-break formation; HB, heterochromatin binding, Rec/Rep, recombination/repair; SCC, sister-chromatid cohesion; SA, spindle assembly; and TM, telomere maintenance. Meiosis lineage-specific (LS) genes are labeled. Genes possessing orthologs are labeled with their common name, if one exists, or with "ortholog" under the "Class" heading. Genes with an ambiguous class status are left unlabeled.

MATERIALS AND METHODS

Table 1 lists the genes in our survey and the pathways in which they are involved. We chose genes that have been shown to be involved in chromosome segregation, crossover regulation, double-strand break (DSB) formation, double-strand break repair, and heterochromatin binding. All of the genes encode products with important roles during meiosis, and many have well-characterized homologs in other model organisms. In addition, we chose a subset of heterochromatin-binding proteins on the basis of their roles in telomere

protection. As mentioned above, it has been shown that distinct telomeric structures can differentially influence the outcome of transmission, as measured in nonrandom disjunction assays (*i.e.*, telomeres may be meiotic drivers; Novitski 1951). Therefore, proteins that are acting to maintain telomeres may also play a role in modifying the transmission behavior/effects of telomeres, leading to unusual patterns of polymorphism and divergence. We chose *cav* as it has been shown by Schmid and Tautz (1997) to be one of the most rapidly evolving genes in Drosophila and appears to be involved in telomere protection; *tefu*, *Ku80*, *mre11*, and *rad50* are all involved in both DSB repair and telomere protection (Badugu *et al.* 2003; Ciapponi *et al.* 2004; Oikemus *et al.* 2004; Bi *et al.* 2005; Melnikova *et al.* 2005); and HP1, encoded by *Su(var)205*, is a ubiquitous heterochromatin-associated protein and one of its functions is telomere protection (Fanti *et al.* 1998).

For each locus, we obtained at least 800 bp of DNA sequence data from 21 North American lines and 10 African lines (provenance information provided as supplemental material). Most of the data are from coding regions; the noncoding data are not included in our analyses, as the number of sites is small for most genes. In addition, we obtained sequence data from 6 lines of *D. simulans* using data from the Washington University Genome Sequencing Center (http://www.genome.wustl.edu) and alignments from Begun *et al.* (2007); we then generated our own data to fill in any gaps in these sequences. Note that these lines of *D. simulans* are from different populations. We also used sequence data from the Washington University Genome Sequencing Center to include one allele of *Drosophila yakuba* in our divergence analyses using alignments provided by Begun *et al.* (2007).

DNA templates were prepared from bulk preps using CsCl gradient purification. Primers were designed for each gene with the Primer3 program (Rozen and Skaletsky 2000). DNA amplification was conducted using standard protocols, and unincorporated primers were removed using PCR purification columns (Qiagen). Sequencing of both strands was performed on an ABI 3730 capillary sequencer (http://www.appliedbiosystems.com) off of the PCR primers.

Sequence assembly for each gene was conducted using a set of perl scripts that carried out the following steps (all scripts available upon request from wdg@stowers-institute.org). First, each allele was individually assembled using Phred/Phrap/Consed (Ewing and Green 1998; Ewing *et al.* 1998; Gordon *et al.* 1998). To facilitate this, a read based on the reference sequence carrying the coding/noncoding sequence annotation as a series of Consed tags and a sequence quality of 0 for all positions was included. The assembly consensus sequence for each allele was then exported to Fasta format and aligned using ClustalW (Thompson *et al.* 1994). This alignment was then converted back into a Consed-readable file, which reconnected the annotations and the Phrap quality scores for each line. Sequence quality for all lines could then be examined simultaneously, allowing for efficient identification of regions in which additional sequence data were needed. The alignment was then trimmed at the ends, all low-quality base calls (defined as those positions where Phrap scores were <25) were converted to N's, and the alignment with annotated coding positions was exported to a Nexus file for subsequent analyses.

We used DnaSP 4.0 (Rozas *et al.* 2003) to estimate π, or the average number of pairwise differences per nucleotide, as well as the average per-site pairwise sequence difference. Pairwise divergence was measured between *D. melanogaster* and *D. simulans*, as well as between both of these species and *D. yakuba*. In addition, PAML's *codeml* was used to estimate polarized rates of nonsynonymous and synonymous divergence (d_N and d_S) for each of the three lineages (Yang 1997). We also used DnaSP to estimate F_{ST}, K_{ST}*, and S_{NN}, all measures of population differentiation between the North American and African samples (http://www.ub.es/dnasp/DnaSP32Inf.html; Hudson *et al.* 1992a,b; Hudson 2000). Note that divergence between *D. melanogaster* and *D. simulans* was measured using the African *D. melanogaster* sample. However, both samples from *D. melanogaster* were included in all other analyses.

We estimated Tajima's *D* to test for deviations from a neutral, equilibrium frequency spectrum in the two *D. melanogaster* samples (Tajima 1989). We also conducted both unpolarized and polarized McDonald–Kreitman tests (McDonald and Kreitman 1991). McDonald–Kreitman tests were not conducted for genes showing six or fewer polymorphisms. For unpolarized tests, the number of segregating sites contributed by one species population was compared to the number of fixed differences between two species for each of the genes. Data from both populations of *D. melanogaster* as well as *D. simulans* data were included in these tests. We report results of *D. melanogaster* analyses using data from each population sample separately, as well as using the combined data from the two samples. We also conducted unpolarized tests in which we pooled the polymorphism data from both species. This set of tests added no new information (relative to the former set of tests in which polymorphism data were not pooled across species) and, in fact, sometimes obscured interesting species-specific patterns of polymorphism. For these reasons, we do not include these tests in our results. For our polarized tests, we used *D. yakuba* to polarize changes along the *D. melanogaster* and *D. simulans* lineages. This kind of test can potentially be more informative than the unpolarized tests, as it assigns mutations to specific lineages.

RESULTS

McDonald–Kreitman tests: The results of the McDonald–Kreitman tests speak directly to our predictions that meiosis genes may often be associated with adaptive protein divergence or excess amino polymorphism (see above). Indeed, these tests revealed that the patterns of polymorphism and divergence for the data set as a whole and for several individual meiotic loci were consistent with our predictions.

We first pooled polarized polymorphism and divergence counts across the 33 genes to test for departures from neutrality in our entire gene set (Table 2; contingency tables for each of the genes are available in the supplemental data). Note that for the pooled analyses we combined counts of polarized mutations. We found that *D. simulans* exhibits a significantly heterogeneous distribution of synonymous and nonsynonymous polymorphism and divergence (Table 2C; $\chi^2 = 27.42$; $P < 0.0001$; d.f. $= 1$), while African *D. melanogaster* exhibits a nearly significantly heterogeneous distribution (Table 2B; $\chi^2 = 3.37$, $P = 0.067$, d.f. $= 1$). In both cases, there was an apparent excess of replacement substitutions, consistent with adaptive protein divergence (see Table 2, B and C). In contrast, the distribution observed for North American *D. melanogaster* was not significantly different from the neutral expectation (Table 2A; $\chi^2 = 0.037$; $P = 0.8475$; d.f. $= 1$).

TABLE 2

Contingency table counts pooled across the 33 genes: mutations are polarized

Change	Polymorphism	Fixation
	A. North American *D. melanogaster*	
Replacement	89	125
Silent	249	339
	B. African *D. melanogaster*	
Replacement	81	132
Silent	281	340
	C. *D. simulans*	
Replacement	129	129
Silent	449	206

All "silent" sites are synonymous.

Importantly, it is possible that the non-neutral distribution of polymorphism and divergence observed for *D. simulans* is due to an excess of silent polymorphisms, rather than an excess of replacement fixations. In fact, several data sets suggest that slightly deleterious, unpreferred polymorphisms are a significant contributor to silent polymorphism in this species (Akashi 1995; Begun *et al.* 2007). To investigate this possibility, we compared counts of replacement polymorphisms and fixations from our meiosis genes to counts of intronic polymorphisms (366) and fixations (161) from those same genes (data from Begun *et al.* 2007). We found that the pattern of adaptive protein divergence in our meiosis data set is even stronger when using intron data ($\chi^2 = 28.1$, $P < 0.0001$).

When examining patterns of polymorphism and divergence in individual genes, we conducted tests with both polarized and unpolarized mutations; here we report the results of the unpolarized, or pairwise, McDonald–Kreitman tests, as they tend to be more powerful than polarized tests because the divergence data from both species are pooled. Moreover, the patterns revealed by the two kinds of tests were very similar (see supplemental Table S1 for the *P*-values of the unpolarized set of tests, the neutrality index (NI; Rand and Kann 1996), and the total number of tests conducted for each lineage; similarly, see supplemental Table S2 for the *P*-values and total number of tests for the polarized set of tests.) For the unpolarized set of tests, six of the genes showed significant deviations from neutrality (Table 3). Overall, *D. simulans* exhibited significant heterogeneity at four of the six genes and the *D. melanogaster* samples together showed significant tests for four of the six genes (see Table 3). In *D. simulans*, all of the significant genes deviated in the direction of excess replacement divergence (see Table 3). However, two genes showed a distinctly different signature in *D. melanogaster*: excess amino acid polymorphism (*mre11* and *rad50*; see Table 3).

To control for multiple comparisons, we calculated the false discovery rate following Benjamini and Hochberg (1995). None of the genes remained significant using this criterion. However, the total number of significant tests exceeded the number expected by chance ($P < 0.05$). For the unpolarized set of tests, *D. simulans* ($n = 31$, significant tests = 4) and North American *D. melanogaster* ($n = 22$, significant tests = 3) exceeded the number expected under neutrality, while for the polarized test set (see supplemental Table S2) *D. simulans* ($n = 24$, significant tests = 2) and African *D. melanogaster* ($n = 17$, significant tests = 1) exceeded this expectation.

Heterozygosity and divergence: Estimates of heterozygosity and divergence per gene are provided in supplemental Tables S3–S6. From a whole-genome perspective, estimates of expected heterozygosity are generally higher in *D. simulans* relative to *D. melanogaster* (*e.g.*, Aquadro *et al.* 1988; Moriyama and Powell 1996). Within *D. melanogaster*, African populations show more variation than non-African populations (*e.g.*, Andolfatto 2001; Haddrill *et al.* 2005). However, estimates of silent and replacement divergence appear to be higher along the *D. melanogaster* lineage, relative to *D. simulans* (Akashi

TABLE 3

Significant pairwise McDonald–Kreitman tests: direction of significance

Gene	North American *D. melanogaster*	African *D. melanogaster*	*D. simulans*
Klp3A[a]	Too many R fixations*	Too many R fixations*	Too many R fixations*
Ku80			Too many R fixations*
mre11	Too many R polymorphisms*		
mtrm[b]	Too many R fixations*		Too many R fixations**
ord			Too many R fixations**
rad50[c]			

R refers to replacement or nonsynonymous. *$P < 0.05$; **$P < 0.01$.

[a] Significant at the $P < 0.01$ level in the direction of too many R fixations in the combined analysis of North American and African *D. melanogaster*.

[b] Significant at the $P < 0.05$ level in the direction of too many R fixations in the combined analysis of North American and African *D. melanogaster*.

[c] Significant at the $P < 0.05$ level in the direction of too many R polymorphisms in the combined analysis of North American and African *D. melanogaster*.

TABLE 4

Comparison of genome means to meiosis gene-set means: nucleotide expected heterozygosity and polarized divergence are reported

	Meiosis means (π_S)	Genome means (π_S)	Meiosis means (π_N)	Genome means (π_N)	Meiosis means (d_S)	Genome means (d_S)	Meiosis means (d_N)	Genome means (d_N)
North American *D. melanogaster*	0.013	0.005[a]	0.001	0.001[a]	0.073	0.073[b]	0.011	0.007[b]
African *D. melanogaster*	0.020	0.022[a]	0.002	—	0.071	—	0.011	—
D. simulans	0.039	0.032[b]	0.005	0.003[b]	0.065	0.054[b]	0.013	0.007[b]
D. yakuba	—	—	—	—	0.272	0.230[b]	0.054	0.027[b]

π_S, synonymous expected heterozygosity; π_N, nonsynonymous expected heterozygosity; d_S, synonymous divergence; d_N, nonsynonymous divergence.

[a] Data are from Andolfatto (2001).

[b] Data are from Begun *et al.* (2007).

1995; Begun *et al.* 2007). These general patterns were also found in our meiosis data set, and we discuss them in greater detail below.

Table 4 reports mean estimates of the expected synonymous and nonsynonymous heterozygosity per nucleotide site (π_S and π_N) and of the polarized synonymous and nonsynonymous divergence (d_S and d_N) for our set of 33 genes. To provide a point of comparison, Table 4 also reports the genomewide means of these parameters, revealing that synonymous and nonsynonymous heterozygosity were higher in the meiosis gene set compared to the genome as a whole in both *D. melanogaster* and *D. simulans* (data are from Begun *et al.* 2007). This trend is especially clear from the North American *D. melanogaster* and *D. simulans* data, for which genomewide estimates are available (from Andolfatto 2001; Begun *et al.* 2007). Moreover, estimates of replacement divergence were almost twice as high in the meiosis gene set compared to the genomewide average for both species.

When we divide the data into lineage-specific and ortholog partitions (see Table 1), we see evidence of higher levels of protein divergence in the lineage-specific subset relative to the ortholog subset in all three population samples (Table 5). Interestingly, the estimate of amino acid divergence for the *D. simulans* lineage-specific partition is over three times higher than the estimate for the ortholog partition. In contrast, amino acid divergence estimates for the two *D. melanogaster* populations are closer to two times as high as the ortholog partition.

Site-frequency spectrum: Supplemental Table S7 reports Tajima's *D* values. This statistic measures the deviation of the site-frequency spectrum from that predicted by the neutral theory (Tajima 1989). Only two genes, *Axs* and *smc1*, were significant; both showed positive Tajima's *D* values in the North American sample. This result is indicative of a skew toward too many common alleles in North America.

Population subdivision: Estimates of population subdivision between North America and Africa are reported in supplemental Table S8. Five genes in our set showed F_{ST} values ≥ 0.35. When we compared our estimates to those reported in Haddrill *et al.* (2005), the meiosis gene set as a whole exhibited fairly typical levels of differentiation between North American and African populations. Interestingly, *mei-P26*, an *X*-linked gene in a region of normal crossing over, showed an extremely high F_{ST} value of 0.79. Two silent sites within *mei-P26* have fixed variants that are unique to the African sample; it appears that *D. simulans* and North American *D. melanogaster* share the ancestral state at these sites, and all of the African *D. melanogaster* alleles possess a derived state.

DISCUSSION

In an effort to establish a population genetic foundation for the investigation of naturally occurring variation in meiotic functions, we surveyed patterns of polymorphism and divergence across 33 candidate meiotic and

TABLE 5

Comparison of lineage-specific means to ortholog means: nucleotide diversity and polarized divergence

	Lineage-specific means (π_S)	Ortholog means (π_S)	Lineage-specific means (π_N)	Ortholog means (π_N)	Lineage-specific means (d_S)	Ortholog means (d_S)	Lineage-specific means (d_N)	Ortholog means (d_N)
North American *D. melanogaster*	0.011	0.015	0.002	0.001	0.078	0.072	0.018	0.007
African *D. melanogaster*	0.018	0.021	0.005	0.002	0.083	0.067	0.022	0.007
D. simulans	0.029	0.042	0.006	0.004	0.070	0.064	0.027	0.007

π_S, synonymous diversity; π_N, nonsynonymous diversity; d_S, synonymous divergence; d_N, nonsynonymous divergence.

related heterochromatin-binding loci in *D. melanogaster* and *D. simulans*. In addition to characterizing several highly conserved genes, we discovered a subset of genes showing patterns consistent with rapid protein evolution as well as a subset containing an excess of amino acid polymorphism. Beyond this heterogeneity, we observed strikingly higher levels of mean nonsynonymous divergence relative to genomewide averages in both North American *D. melanogaster* and *D. simulans*, and we detected unusually high levels of population structure between North American and African samples at several loci.

Population sample differences: Although patterns of polymorphism and divergence in the meiosis gene set were striking in all three population samples, several differences were evident. For example, when we pooled polarized counts of polymorphisms and fixations across the entire data set, we found evidence for adaptive protein divergence in *D. simulans* but not in the two *D. melanogaster* samples. These differences could be attributed to several factors. For example, relative to both samples of *D. melanogaster*, there were more *D. simulans* variants in the pooled gene data set, consistent with previous studies that have uncovered higher levels of heterozygosity in *D. simulans* compared to *D. melanogaster* (*e.g.*, Aquadro *et al.* 1988; Moriyama and Powell 1996). As a consequence, there is additional power to reject the null hypothesis of neutrality in *D. simulans*. This is consistent with the observation that heterogeneity for the entire gene set was found to be nearly statistically significant in the African *D. melanogaster* sample, which contained more pooled variants than the North American sample. In addition, we observed an apparent excess of amino acid polymorphism at several *D. melanogaster* loci, especially in the North American sample. Thus, in the pooled analyses any signal of adaptive protein divergence along the *D. melanogaster* lineage could have been counterbalanced by the excess of amino acid polymorphism.

Related to this is the possibility that the observed differences may reflect disparate selective regimes being imposed on the different populations. In fact, North American *D. melanogaster* showed a tendency toward an excess of genes exhibiting too many amino acid polymorphisms, whereas *D. simulans* (and to some extent African *D. melanogaster*) tended toward an excess of genes exhibiting too many amino acid fixations. This pattern was also evident when we calculated the NI, a qualitative indication of the direction and degree of departure from neutrality (Rand and Kann 1996; see supplemental Table S1). NI values >1.0 indicate a tendency to show an excess of amino acid polymorphism, whereas values <1.0 suggest an excess of divergent amino acid substitutions (Rand and Kann 1996). Comparison of the entire set of NI values for all three populations revealed that the North American *D. melanogaster*–*D. simulans* contrast is nearly significantly different ($\chi^2 = 3.20$, $P = 0.074$). Indeed, North American *D. melanogaster* showed a greater proportion of NI values >1.0 (14/26), whereas *D. simulans* showed a greater proportion <1.0 (8/27; see supplemental Table S1). The North American–African *D. melanogaster* contrast was less striking, although like *D. simulans*, African *D. melanogaster* showed a greater proportion of NI values <1.0 (9/25; see supplemental Table S1).

As discussed above, depending on the fitness effects of meiotic drive variants as well as their modifiers, both observations of adaptive protein divergence and excess protein polymorphism may be consistent with the presence of such elements in populations of *D. simulans* and *D. melanogaster*. Indeed, one can imagine a scenario in which competition for preferential meiotic transmission may drive the sequential fixation of different drivers if the fitness landscape is such that only a single type is most fit at any particular point in time. As different variants enter the population, relative fitnesses may shift, allowing other variants to sweep through. Such a scenario of directional selection due to meiotic drive is consistent with the general pattern of adaptive protein divergence seen in *D. simulans*. Alternatively, North American *D. melanogaster* drive elements may be experiencing balancing selection where the fitness varies over time so that a variant rarely fixes, predicting the presence of several common drivers as well as their meiotic modifiers (Sandler and Novitski 1957). Regardless of the explanation(s), it is clear that our data set contains several potentially important patterns that merit further investigation.

Individual genes: A total of six genes exhibited significantly non-neutral patterns of polymorphism and divergence in one or more of the population samples (Table 3). Although we cannot assert that all of these genes have experienced non-neutral forces in their histories (multiple test problem), we did observe a greater number of significant tests than expected by chance. It is also important to note that for 1 of the 33 total genes we did not observe a sufficient number of polymorphisms to conduct any McDonald–Kreitman tests (see materials and methods). Similarly, for 9 of the 33 genes, we were able to conduct tests in only one of the population samples (see contingency tables and supplemental Table S1). Clearly then, additional data will be necessary to complete this story. At this point in time, however, we can conclude that meiotic genes tend to show non-neutral patterns of evolution. Perhaps significantly, it appears that our 6 significant genes fall into two major functional categories: telomere binding/maintenance and chromosome segregation. Intriguingly, Turner *et al.* (2008) recently reported strong differentiation at telomere-binding and chromosome segregation loci between temperate and subtropical populations of *D. melanogaster*.

Telomere binding: As is true for many eukaryotes, the significant telomere-binding genes identified in this

study are known to have roles in DNA metabolism as well. For example, the protein products of the *mre11* and *rad50* genes form a complex involved in several DNA repair pathways (Ciapponi *et al.* 2004; for a recent review, see Cahill *et al.* 2006), and they also play a prominent role in Drosophila telomere maintenance by regulating the localization of the telomere-capping proteins, HOAP and HP1, to the ends of chromosomes (Bi *et al.* 2004; Ciapponi *et al.* 2004). In this study, *mre11* exhibited excess amino acid polymorphism in the North American *D. melanogaster* sample, while *rad50* showed excess amino acid polymorphism in the combined North American and African *D. melanogaster* samples. It is possible that distinct telomere structures are bound more (or less) efficiently by different variants of these telomere proteins, leading to the maintenance of amino acid variation at these loci. Two pieces of data support this scenario: first, we find that all of the replacement polymorphisms in the North American population fall within the *mre11* DNA-binding domain of the protein (the situation is less clear for *rad50*), and second, high levels of polymorphism have been observed in *D. melanogaster* subtelomeric regions (Anderson *et al.* 2008).

Ku80 is also involved in both DNA repair and telomere maintenance (for a recent review, see Slijepcevic and Al-Wahiby 2005; Melnikova *et al.* 2005). In contrast to *mre11* and *rad50*, however, the protein product of *Ku80*, along with *Ku70*, regulates Drosophila telomere length (Melnikova *et al.* 2005). Evidence of too many amino acid fixations along the *D. simulans* lineage may be consistent with *Ku80* modifying the effects of telomeric meiotic drivers in this species. In fact, as described above, Novitski (1951) showed that the lengths of distal chromosomal regions can alter a chromosome's probability of inclusion into the pronucleus of the Drosophila oocyte. Consistent with this scenario of different subtelomeric variants sweeping through the population, J. A. Anderson, S. E. Celniker and C. H. Langley (unpublished results) found that much of the subtelomeric region exhibits rapid divergence in both *D. melanogaster* and *D. simulans*. They, too, found evidence of strong differentiation between temperate and subtropical populations for these categories.

Chromosome segregation: The significant genes functioning in chromosome segregation act through distinct pathways. Both protein products of the genes *mtrm* and *ord* are involved in pairing during chromosome segregation. Mutant alleles of *mtrm* show dominant disruption of achiasmate segregation, whereas *ord* has been shown to be involved in both male and female chromosome segregation (Miyazaki and Orr-Weaver 1992; reviewed in McKim *et al.* 2002). The phenotypes of both of these mutants show increased levels of meiotic chromosome nondisjunction, and the sequences of both genes exhibit elevated amino acid divergence in *D. simulans* and *D. melanogaster*, again suggesting that they may act to modify the effects of meiotic drive elements.

Remarkably, evidence of rapid protein evolution was found in all analyses for *Klp3A* (see Table 3). The protein encoded by this locus is known to be a kinesin motor, serving roles in both meiosis and mitosis (Williams *et al.* 1997). The major meiotic defect observed in *Klp3A* mutants is the failure either of the female pronucleus to be specified or of the migration of the male and female pronuclei toward each other (Williams *et al.* 1997). In addition, Cook *et al.* (1997) identified *Klp3A* as among the set of *trans*-acting factors necessary for centromere function (and therefore chromosome segregation) in their study of centromere-defective minichromosomes. These roles suggest that *Klp3A* could modify the effects of meiotic drive elements, such as simple "centromere strength" (Novitski 1955). Alternatively, its role in pronuclear specification and/or migration suggests that it may mediate which chromosomes are included in the female pronucleus.

Functional studies: In addition to motivating further population genetics studies with more extensive sampling, our data provide the necessary information to carefully design preliminary experiments testing the effects of natural variants. For example, genes involved in chromosome segregation, such as *mtrm* and *ord*, could be subjected to nondisjunction assays, similar to the assay described for *nod* (Zwick *et al.* 1999). In addition, genes that serve a role in crossover regulation are obvious candidates for an assay designed to detect differences in the rate of recombination in particular genomic intervals, and genes that may modify the transmission of meiotic drivers (such as centromeric or telomeric variants), based on their distinct patterns of polymorphism and divergence, can be tested using a nonrandom disjunction assay. Finally, it is worth noting that because several of our genes appear pleiotropic, it is plausible that testing the phenotypes of naturally occurring variants may more readily reveal interesting phenotypic effects that analyses of complete loss-of-function alleles cannot.

We are grateful to Alisha Holloway, Kristian Stevens, and Yu-Ping Poh for using their computational expertise to help us compare our data set to the entire genome. We also thank Dan Bradford and the molecular biology core facility at the Stowers Institute for Medical Research for their sequencing expertise, and two anonymous reviewers for helpful suggestions on our analysis. In addition, Dave Begun provided many insightful suggestions regarding our analyses, as well as information on genomewide patterns of polymorphism and divergence in *D. simulans*. We are also grateful to R. Scott Hawley and the Stowers Institute for their support of this work. This research was funded by a National Science Foundation Predoctoral Graduate Fellowship to J.A.A., an American Cancer Society Postdoctoral Fellowship to W.D.G., and HG-02942-10A1 to C.H.L.

LITERATURE CITED

Afshar, K., N. R. Barton, R. S. Hawley and L. S. B. Goldstein, 1995a DNA-binding and meiotic chromosomal localization of the Drosophila *nod* kinesin-like protein. Cell **81:** 129–138.

Afshar, K., J. Scholey and R. S. Hawley, 1995b Identification of the chromosome localization domain of the Drosophila *nod* kinesin-like protein. J. Cell Biol. **131:** 833–843.

Akashi, H., 1995 Inferring weak selection from patterns of polymorphism and divergence at "silent" sites in Drosophila DNA. Genetics **139:** 1067–1076.

Anderson, J. A., Y. S. Song and C. H. Langley, 2008 Molecular population genetics of Drosophila subtelomeric DNA. Genetics **178:** 477–487.

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.

Badugu, R., M. M. Shareef and R. Kellum, 2003 Novel Drosophila heterochromatin protein 1 (HP1)/origin recognition complex-associated protein (HOAP) repeat motif in HP1/HOAP interactions and chromocenter associations. J. Biol. Chem. **278:** 34491–34498.

Baker, B. S., and A. T. Carpenter, 1972 Genetic analysis of sex chromosomal meiotic mutants in *Drosophila melanogaster*. Genetics **71:** 255–286.

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.

Benjamini, Y., and Y. Hochberg, 1995 Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. B. **57:** 289–300.

Bi, X., S. C. Wei and Y. S. Rong, 2004 Telomere protection without telomerase: the role of ATM and Mre11 in Drosophila telomere maintenance. Curr. Biol. **14:** 1348–1353.

Bi, X., D. Srikanta, L. Fanti, S. Pimppinelli, R. Badugu *et al.*, 2005 Drosophila ATM and ATR checkpoint kinases control partially redundant pathways for telomere maintenance. Proc. Natl. Acad. Sci. USA **102:** 15167–15172.

Cahill, D., B. Connor and J. P. Carney, 2006 Mechanisms of eukaryotic DNA double strand break repair. Front. Biosci. **11:** 1958–1976.

Carpenter, A. T., 1973 Meiotic mutant defective in distributive disjunction in *Drosophila melanogaster*. Genetics **73:** 393–428.

Charlesworth, B., and D. L. Hartl, 1978 Population dynamics of segregation distorter polymorphism of *Drosophila melanogaster*. Genetics **89:** 171–192.

Ciapponi, L., G. Cenci, J. Ducau, C. Flores, D. Johnson-Schlitz *et al.*, 2004 The Drosophila Mre11/Rad50 complex is required to prevent both telomeric fusion and chromosome breakage. Curr. Biol. **14:** 1360–1366.

Cook, K. R., T. D. Murphy, T. C. Nguyen and G. H. Karpen, 1997 Identification of *trans*-acting genes necessary for centromere function in *Drosophila melanogaster* using centromere-defective minichromosomes. Genetics **145:** 737–747.

Coop, G., X. Wen, C. Ober, J. K. Pritchard and M. Przeworski, 2008 High-resolution mapping of crossovers reveals extensive variation in fine-scale recombination patterns among humans. Science **319:** 1395–1398.

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.

Dawe, R. K., and W. Z. Cande, 1996 Induction of centromeric activity in maize by suppressor of meiotic drive 1. Proc. Natl. Acad. Sci. USA **93:** 8512–8517.

Ewing, B., and P. Green, 1998 Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res. **8:** 186–194.

Ewing, B., L. Hillier, M. C. Wendl and P. Green, 1998 Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res. **8:** 175–185.

Fanti, L., G. Giovinazzo, M. Berloco and S. Pimpinelli, 1998 The heterochromatin protein 1 prevents telomere fusions in Drosophila. Mol. Cell **2:** 527–538.

Gerton, J. L., and R. S. Hawley, 2005 Homologous chromosome interactions in meiosis: diversity amidst conservation. Nat. Rev. Genet. **6:** 477–487.

Gordon, D., C. Abajian and P. Green, 1998 Consed: a graphical tool for sequence finishing. Genome Res. **8:** 195–202.

Haddrill, P. R., K. R. Thornton, B. Charlesworth and P. Andolfatto, 2005 Multilocus patterns of nucleotide variability and the demographic and selection history of *Drosophila melanogaster* populations. Genome Res. **15:** 790–799.

Hawley, R. S., H. Irick, A. E. Zitron, 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.

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 segregation-distorter locus. Evolution **14:** 433–444.

Hudson, R. R., 2000 A new statistic for detecting genetic differentiation. Genetics **155:** 2011–2014.

Hudson, R. R., D. D. Boos and N. L. Kaplan, 1992a A statistical test for detecting geographic subdivision. Mol. Biol. Evol. **9:** 138–151.

Hudson, R. R., M. Slatkin and W. P. Maddison, 1992b Estimation of levels of gene flow from DNA sequence data. Genetics **132:** 583–589.

Kong, A., G. Thorleifsson, H. Stefansson, G. Masson, A. Helgason *et al.*, 2008 Sequence variants in the RNF212 gene associate with genome-wide recombination rate. Science **319:** 1398–1401.

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.

Lindsley, D. L., and L. Sandler, 1977 Genetic analysis of meiosis in female *Drosophila melanogaster*. Philos. Trans. R. Soc. Lond. B Biol. Sci. **277:** 295–312.

Lynn, A., T. Ashley and T. Hassold, 2004 Variation in human meiotic recombination. Annu. Rev. Genomics Hum. Genet. **5:** 317–349.

Malik, H. S., and S. Henikoff, 2001 Adaptive evolution of *Cid*, a centromere-specific histone in Drosophila. Genetics **157:** 1293–1298.

Malik, H. S., and S. Henikoff, 2002 Conflict begets complexity: the evolution in centromeres. Curr. Opin. Genet. Dev. **12:** 711–718.

McDonald, J. H., and M. Kreitman, 1991 Adaptive protein evolution at the *ADH* locus in Drosophila. Nature **351:** 652–654.

McKim, K. S., J. K. Jang and E. A. Manheim, 2002 Meiotic recombination and chromosome segregation in Drosophila females. Annu. Rev. Genet. **36:** 205–232.

Melnikova, L., H. Biessmann and P. Georgiev, 2005 The Ku protein complex is involved in length regulation of Drosophila telomeres. Genetics **170:** 221–235.

Miyazaki, W. Y., and T. L. Orr-Weaver, 1992 Sister-chromatid misbehavior in Drosophila *ord* mutants. Genetics **132:** 1047–1061.

Moriyama, E. N., and J. R. Powell, 1996 Intraspecific nuclear DNA variation in Drosophila. Mol. Biol. Evol. **13:** 261–277.

Novitski, E., 1951 Non-random disjunction in Drosophila. Genetics **36:** 267–280.

Novitski, E., 1955 Genetic measures of centromere activity in *Drosophila melanogaster*. J. Cell. Comp. Physiol. **45:** 151–169.

Ohnishi, S., and R. A. Voelker, 1979 Comparative studies of allozyme loci in *Drosophila simulans* and *D. melanogaster*. 2. Gene arrangement on the 3rd chromosome. Jpn. J. Genet. **54:** 203–209.

Oikemus, S. R., N. McGinnis, J. Queiroz-Machado, H. Tukachinsky, S. Takada *et al.*, 2004 Drosophila *atm/telomere fusion* is required for telomeric localization of HP1 and telomere position effect. Genes Dev. **18:** 1850–1861.

Pardo-Manuel de Villena, F., and C. Sapienza, 2001 Female meiosis drives karyotype evolution in mammals. Genetics **159:** 1179–1189.

Ptak, S. E., D. A. Hinds, K. Koehler, B. Nickel and N. Patil, 2005 Fine-scale recombination patterns differ between chimpanzees and humans. Nat. Genet. **37:** 429–434.

Rand, D. M., and L. M. Kann, 1996 Excess amino acid polymorphism in mitochondrial DNA: contrasts among genes from Drosophila, mice, and humans. Mol. Biol. Evol. **13:** 735–748.

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.

Rozen, S., and H. Skaletsky, 2000 Primer3 on the WWW for general users and for biologist programmers, pp. 365–386 in *Bioinformatics Methods and Protocols: Methods in Molecular Biology*, edited by S. Krawetz and S. Misener. Humana Press, Totowa, NJ.

Sandler, L., and E. Novitski, 1957 Meiotic drive as an evolutionary force. Am. Nat. **91:** 105–110.

Schmid, K. J., and D. Tautz, 1997 A screen for fast evolving genes from Drosophila. Proc. Natl. Acad. Sci. USA **94:** 9746–9750.

Slijepcevic, P., and S. Al-Wahiby, 2005 Telomere biology: integrating protection with DNA damage chromosomal end response. Chromosoma **114:** 275–285.

Sniegowski, P. D., A. Pringle and K. A. Hughes, 1994 Effects of autosomal inversions on meiotic exchange in distal and proximal regions of the *X* chromosome in a natural population of *Drosophila melanogaster*. Genet. Res. **63:** 57–62.

Tajima, F., 1989 Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics **123:** 585–595.

Takano-Shimizu, T., 1999 Local recombination and mutation effects on molecular evolution in Drosophila. Genetics **153:** 1285–1296.

Thompson, J. D., D. G. Higgins and T. J. Gibson, 1994 Clustal-W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:** 4673–4680.

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

Turner, T. L., M. T. Levine, M. L. Eckert and D. J. Begun, 2008 Genomic analysis of adaptive differentiation in *Drosophila melanogaster*. Genetics **179:** 455–473.

White, M. J. D., 1973 *Animal Cytology and Evolution*, Ed. 3. Cambridge University Press, Cambridge, UK.

Williams, B. C., A. F. Dernburg, J. Puro, S. Nokkala and M. L. Goldberg, 1997 The Drosophila kinesin-like protein *klp3a* is required for proper behavior of male and female pronuclei at fertilization. Development **124:** 2365–2376.

Yang, Z. H., 1997 PAML: a program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. **13:** 555–556.

Zhang, P., and R. S. Hawley, 1990 The genetic analysis of distributive segregation in *Drosophila melanogaster*: further genetic analysis of the *nod* locus. Genetics **125:** 115–127.

Zhang, P., B. A. Knowles, L. S. B. Goldstein and R. S. Hawley, 1990 A kinesin-like protein required for distributive chromosome segregation in Drosophila. Cell **62:** 1053–1062.

Zickler, D., and N. Kleckner, 1999 Meiotic chromosomes: integrating structure and function. Annu. Rev. Genet. **33:** 603–754.

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: M. Aguadé