Copyright © 1998 by the Genetics Society of America

Different Forces Drive the Evolution of the *Acp26Aa* and *Acp26Ab* Accessory Gland Genes in the *Drosophila melanogaster* Species Complex

Montserrat Aguadé

Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, 08071 Barcelona, Spain

Manuscript received February 4, 1998
Accepted for publication June 16, 1998

ABSTRACT

The *Acp26Aa* and *Acp26Ab* genes that code for male accessory gland proteins are tandemly arranged in the species of the *Drosophila melanogaster* complex. An ∼1.6-kb region encompassing both genes has been sequenced in 10, 24, and 18 lines from Spain, Ivory Coast, and Malawi, respectively; the previously studied 10 lines from North Carolina have also been included in the analyses. A total of 110 nucleotide and 4 length polymorphisms were detected. Silent variation for the whole *Acp26A* region was slightly higher in African than in non-African populations, while for both genes nonsynonymous variation was similar in all populations studied. Based on F_{st} estimates no major genetic differentiation was detected between East and West Africa, while in general non-African populations were strongly differentiated from both African populations. Comparison of polymorphism and divergence at synonymous and nonsynonymous sites revealed that directional selection acting on amino acid replacement changes has driven the evolution of the Acp26Aa protein in the last 2.5 myr.

IN sexually reproducing organisms both external and internal fertilization require the interaction between proteins synthesized exclusively in the male or the female reproductive organs. In the case of external fertilization there is a relatively reduced number of gamete surface proteins that mediate species-specific gamete recognition (Palumbi 1994). The situation is more complex in organisms with internal fertilization as, in addition to fertilization, mating often triggers in the female a series of physiological and behavioral responses; also, in the case of multiple matings there may be sperm competition mediated by male-specific proteins. In both cases there is an ample field for natural selection to shape variation within and between species in the genes coding for these proteins.

The study of nucleotide variation in the genes coding for sperm lysin in abalone (Lee and Vacquier 1995) and sperm bindin in sea urchins (Metz and Palumbi 1996) has revealed the action of positive Darwinian selection in shaping their protein evolution. In the case of sperm lysin the ratio between nonsynonymous and synonymous divergence was significantly higher than one, but no polymorphism was detected in the small sample of individuals analyzed from a particular species. On the other hand, an excess of nonsynonymous substitutions between species was detected in a 40-codon domain region of sperm bindin; there was also a high level of nonsynonymous polymorphism in that same region in at least one of the species analyzed.

In Drosophila the seminal fluid contains many proteins that are synthesized in the male reproductive tissues, primarily in the accessory glands [male *Ac*cessory gland *p*roteins (Acps)], which are transferred to the female during copulation. Some of these proteins, like Acp36DE, remain in the female's genital tract and seem to participate in sperm storage (Bertram *et al.* 1996; Wolfner 1997). Other Acps (*e.g.*, Acp70A, Acp26Aa) move to the hemolymph and thereafter they interact with still unidentified female receptors and induce behavioral and physiological changes in the female. It has been also shown that variation in some of these genes can account for differences in sperm competition (Clark *et al.* 1995). In that study one of the regions that revealed a positive correlation between nucleotide variation and the ability to defend sperm displacement was the *Acp26A* region, which contains the *Acp26Aa* and *Acp26Ab* genes arranged in tandem (Monsma and Wolfner 1988). However, in a later experiment using males with a null mutation of *Acp26Aa* that produced no Acp26Aa protein (Herndorn and Wolfner 1995), the absence of Acp26Aa in these males had no effect on sperm competition. One of the well-established functions of Acp26Aa in *Drosophila melanogaster* is to stimulate egg laying for a period of 24 hr after mating (Herndorn and Wolfner 1995). Given that the sex-peptide or Acp70A also induces egg laying but for a longer period of up to 7 days (Kubli 1996), it has been proposed that Acp26Aa would increase egg deposition as opposed to Acp70A, which would stimulate the produc-

Address for correspondence: Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Diagonal 645, 08071 Barcelona, Spain. E-mail address: aguade@porthos.bio.ub.es

tion of eggs by the ovary (Kubli 1996; Wolfner 1997). On the other hand, little is known about the function of the *Acp26Ab* gene product.

Nucleotide variation in the *Acp26A* region has been surveyed in a North American population of *D. melanogaster* and also in one allele of the sibling species *D. simulans*, *D. mauritiana*, and *D. sechellia* (Aguadé *et al.* 1992). For the *Acp26Aa* gene this study revealed a high degree of amino acid replacement variation both within and between species. In fact, estimates of polymorphism and divergence at nonsynonymous sites were rather high, and of the same order as the corresponding estimates (polymorphism and divergence) at synonymous sites. Variation was not randomly distributed along the coding region, but there was an excess of both nonsynonymous and synonymous divergence in the 3′ half of the gene. The second exon of the *Acp26Aa* gene was later compared between *D. melanogaster* and *D. simulans*, and the more distantly related species of the *melanogaster* group, *D. yakuba* and *D. teissieri* (Tsaur and Wu 1997). In the different comparisons the estimates of K_a were generally higher than the corresponding estimates of K_s, pointing again to positive selection as a major force in the evolution of this gene. On the other hand, the *Acp26Ab* gene did not show the elevated nonsynonymous variation observed in the *Acp26Aa* gene.

However, variation at these two genes has been studied only in a sample from North America. To obtain a better insight into the role played by selection in the evolution of these genes, nucleotide variation at the *Acp26Aa* and *Acp26Ab* genes has been surveyed in three other populations: a European population (Montblanc, Spain) as a second representative of non-African populations, and two populations from equatorial Africa (Ivory Coast and Malawi in West and East Africa, respectively).

MATERIALS AND METHODS

Drosophila stocks: Ten isofemale lines collected in 1993 in Montblanc (Spain) were isogenized for the second chromosome by the corresponding series of crosses with a balancer stock. Twenty-four lines from Lamto (Ivory Coast) and 18 lines from Malawi, kindly provided by M. Veuille and V. Bénassi (see Bénassi *et al.* 1993; Bénassi and Veuille 1995), were used in this study. African lines were either isogenic for the second chromosome or maintained as heterozygotes over a balancer chromosome. The African isogenic lines were used directly. On the other hand, one male from each of the African lines maintained as heterozygotes was crossed with females carrying the deficiency *Df(2L)Gdh-A,dpL*, which covers cytological positions 25D7 to 26A8-9; individuals heterozygous for the deficiency and, therefore, hemizygous for the wild *Acp26A* locus were selected from the progeny. Ten lines from North Carolina, previously studied in Aguadé *et al.* (1992), were also included in the analyses.

DNA extraction, PCR amplification, and sequencing: DNA was extracted from 10 adult flies by a modification of protocol 48 in Ashburner (1989). For each of the lines an ~1.7-kb region encompassing both the *Acp26Aa* and *Acp26Ab* transcriptional units was PCR-amplified using primers 23 nucleotides long (Aguadé *et al.* 1992). The conditions for amplification were 94° for denaturing, 45° for annealing, and 65° for extension over 30 cycles. Sequencing primers were spaced on average 250 nucleotides. PCR products were made single-stranded with the lambda exonuclease procedure (Higuchi and Ochman 1989) and sequenced using Sequenase (United States Biochemical, Cleveland, OH). Alternatively, they were purified with Qiaquick columns (QIAGEN Inc., Chatsworth, CA) and sequenced using the Dye Terminator chemistry (Perkin Elmer, Norwalk, CT) according to the manufacturer's directions with slight modifications; the products were separated with an ABI 377 automated DNA sequencer (Perkin Elmer). All lines were sequenced on both strands. The sequences of *D. simulans*, *D. mauritiana*, and *D. sechellia* used in the analyses are those reported in Aguadé *et al.* (1992). The sequences newly reported in this article have been deposited in the EMBL sequence database library under accession numbers AJ231350–AJ231401.

Sequence analysis: Sequences were edited for further analyses using the MacClade version 3.0.6 program (Maddison and Maddison 1992). The DnaSP version 2.80 program (Rozas and Rozas 1997) was used for most intraspecific and some interspecific analyses.

RESULTS

Nucleotide variation: Figure 1 and Table 1 give a summary of nucleotide sequence variation in the four populations surveyed. A total of 110 nucleotide and 4 length polymorphisms were detected in the 1614-nucleotide region studied (excluding alignment gaps). Five nucleotide polymorphisms segregated for three variants resulting in a total of 115 mutations segregating in the sampled populations. Only 1 length polymorphism (an 8-bp deletion from nucleotide 1575 to 1582) segregated in three populations (two African and one non-African), while the other three deletions were singletons. Nucleotide polymorphism at position 1578 was not considered in further analyses as some of the lines segregated for a deletion spanning this position.

Unlike the samples from Montblanc and North Carolina, the African samples segregated for the *In(2L)t* inversion. The region studied is located on band 26A and therefore is included in the region affected by the inversion, which extends between bands 22D3-E1 and 34A8-9. Only 1 of the 18 lines studied from Malawi presented the inversion, while 9 of the 24 lines studied from the Ivory Coast did (M. Veuille, personal communication). The 2 sets of lines of this latter population were considered together for further analyses as they were not genetically differentiated for the *Acp26A* region when using the permutation test proposed by Hudson *et al.* (1992a) and K_s^* as the test statistic.

Nucleotide variation for the whole region studied was estimated separately for each population (Table 1). The level of both total and silent polymorphism was slightly higher for the two African populations, especially when compared to the North American sample. The level of synonymous variation in the exons was higher than that of silent variation in noncoding regions. The region

between the two genes, which includes regulatory sequences of both genes, showed the lowest level of silent polymorphism.

As in the original study of this region (Aguadé *et al.* 1992), the level of nonsynonymous variation of both the *Acp26Aa* and *Acp26Ab* genes was rather high as compared to other regions of the genome; the ratio of nonsynomymous to synonymous polymorphism was higher for the *Acp26Aa* gene. Figure 2 shows the amino acid variation of the two proteins in the populations sampled. The mean number of amino acid differences between lines varied between 4 and 5 for Acp26Aa and was approximately equal to 1 for Acp26Ab. Only the sample from Lamto presented one not-in-frame 3-bp deletion in exon 2 of the *Acp26Aa* gene that resulted in one amino acid change in addition to one amino acid deletion. Polymorphisms at adjacent residues 24 and 25 involved the same 2 amino acids (Gln/Lys), and also those at nearby residues 76 and 79 (Asp/Asn). Residue 90 of the Acp26Ab protein segregated for 2 amino acid variants (Ala/Glu) at intermediate frequencies in the African populations, while it was monomorphic for Ala in the non-African samples. No amino acid polymorphism was detected in any of the five putative glycosylation sites (Monsma and Wolfner 1988) and in any of the three peptidase sites (Park and Wolfner 1995) of Acp26Aa; also the 11-amino-acid residues conserved relative to the Aplysia Egg Laying Hormone (ELH; Monsma and Wolfner 1988) were monomorphic in all the samples studied.

Genetic differentiation between populations: Table 2 summarizes the analysis of population differentiation. The F_{st} parameter, which measures the proportion of nucleotide diversity attributable to variation between populations, was estimated from the average number of differences between alleles according to Hudson *et al.* (1992b). Population differentiation was assessed by the permutation test (Hudson *et al.* 1992a) using K_s^* as the test statistic. The analysis was performed for the whole region studied considering either all variation or only silent variation, and also separately for the *Acp26Aa* and *Acp26Ab* transcription units. In all cases the comparison between the two African populations presented the lowest F_{st} estimates. These populations did not show any significant differentiation for the *Acp26Ab* transcription unit, while for the *Acp26Aa* transcription unit, and for the whole region, they were either marginally ($0.05 < P < 0.10$) or slightly ($P = 0.04$) differentiated. The non-African populations showed a significant differentiation in all cases from the Malawi population and in most cases from the Ivory Coast population.

Linkage disequilibrium: In light of the results of genetic differentiation between populations, linkage disequilibrium or pairwise gametic association was estimated separately for each population and also for the pooled African sample. Table 3 shows the total number of pairwise comparisons per population, the number of comparisons that presented a significant association (using the χ^2 test and without considering the Bonferroni correction for multiple comparisons), and the proportion of significant comparisons. In all cases this proportion was higher in the non-African than in the African samples. Also in the African samples the proportion was in all cases higher for the *Acp26Ab* transcription unit. For the two larger samples from Africa and for the pooled African sample, Table 3 also shows the number of significant comparisons when the Bonferroni correction was applied. In the pooled African sample, 8 of the 15 significant associations in the *Acp26Ab* transcription unit were in the trailer between polymorphisms 1490 and 1572; in this smaller region this represented 22% of significant associations (8 out of 36 comparisons). The trailer presented in fact three major haplotypes (represented by lines La32, La36, and La108) in the African samples. These major haplotypes were also present in the non-African samples, although only two segregated in each population. However, even in this small region (between sites 1490 and 1572) five recombination events could be inferred in the history of the African pooled sample using the four-gamete test (Hudson and Kaplan 1985), while none was in the non-African samples.

Tests of neutrality: Tests of neutrality based on polymorphism data were applied to the whole region and also to the two transcription units separately. In the samples studied neither the Tajima test (Tajima 1989) nor the Fu and Li test (Fu and Li 1993) detected any departure from neutral expectations.

Tests of neutrality based on polymorphism and divergence data were also applied. The direct relationship between levels of polymorphism and divergence in a given region, predicted by the neutral theory of molecular evolution, constitutes the common null hypothesis of these tests. Table 4 summarizes the results of the McDonald and Kreitman (MK) test (McDonald and Kreitman 1991). This test, which compares nonsynonymous and synonymous changes in a coding region, was applied separately to the *Acp26Aa* and *Acp26Ab* genes. In the *Acp26Aa* gene there was in all cases an excess of nonsynonymous differences fixed between species; this excess was significant in comparisons involving the Montblanc, Lamto, and Malawi populations, and only marginally significant in the North Carolina comparison. On the other hand, the MK test did not detect any departure from neutral expectations in the much shorter *Acp26Ab* coding region. The Hudson, Kreitman, and Aguadé (HKA) test (Hudson *et al.* 1987) was applied both for all changes and for silent changes (analysis not shown); in this test the regions compared were the two transcription units included in the fragment studied (*Acp26Aa* and *Acp26Ab*). When all changes were considered, most of the comparisons revealed a significant departure from the neutral expectations, while no departure was detected when only silent variation was considered. The difference in the results of these two

Acp26Aa

	-22	-8	1	39	52	53	69	71	109	117	121	137	150	151	154	176	205	207	212	217	228	234	241	270	275	306	307	316	335-7	348	383	392	406	417	420	447	468	581	586	666	693	701	702	714	743	750	753	775	813	834	846	849	867	874	902	906	911	960	961
		*				*							*						*				*		*				*			*			*	*	*												*	*			*		*	*	*	*	
				n		n					n	n		n	n	n	n		n	n			n		n		n	n	n		n	n	n					n	n			n			n							n		n					
NC1	T	A	C	T	A	A	C	T	C	G	C	G	G	C	C	T	C	G	A	C	T	G	G	G	A	C	G	G	T	G	G	C	A	T	G	T	G	T	C	C	T	G	C	C	G	T	C	C	G	C	G	G	G	C	G	C	C	G	A
NC2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	C	T	A	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NC3	A	.	T	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	A	.	.	.	C	.	.	A	.	.	.	.	.	T	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.
NC4	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	A	.	.	A	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NC5	A	.	T	.	C	.	.	.	.	.	.	A	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	C	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NC6	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	A	.	.	A	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NC7	A	.	T	.	C	.	.	.	.	.	.	A	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	C	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NC8	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	A	.	.	A	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NC9	A	.	T	.	C	.	.	.	T	.	.	.	.	.	.	A	.	.	.	A	.	.	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NC10	A	.	T	.	C	.	.	.	.	.	.	A	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	C	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Mo13a	A	.	T	.	C	.	.	.	.	.	.	A	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	C	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Mo29b	A	.	T	.	C	.	.	.	T	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	C	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Mo34a	A	.	T	.	C	.	.	.	.	.	.	A	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	C	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Mo36a	A	.	T	.	C	.	.	.	T	.	.	.	.	.	.	A	.	.	.	A	.	.	.	.	.	.	.	.	.	.	A	.	C	C	.	.	.	.	.	T	.	T	A	A	.	.	T	T	.	.	A	.	.	T	.	.	.	.	.
Mo37a	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	A	.	.	.	.	A	.	C	C	.	.	.	.	.	T	.	T	A	A	.	.	T	T	.	.	A	.	.	T	.	.	.	.	.
Mo40b	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Mo47a	.	.	.	.	C	.	.	.	.	.	.	A	.	.	.	.	.	.	.	G	.	.	.	.	.	A	.	.	.	.	A	.	C	C	.	.	.	.	.	T	.	T	A	A	.	.	T	T	.	.	A	.	.	T	.	.	.	.	.
Mo52b	A	.	T	.	C	.	.	.	T	.	.	.	.	.	.	A	.	.	.	A	.	.	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Mo79b	A	.	T	.	C	.	.	.	T	.	.	.	.	.	.	A	.	.	.	A	.	.	.	.	.	.	.	.	.	.	A	.	C	C	.	.	.	.	.	T	.	T	A	A	.	.	T	T	.	.	A	.	.	T	.	.	.	.	.
Mo80b	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
La3	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	A	.	.	.	A	C	A	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	T	.	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	.	.	.
La10	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	C	T	A	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	G
La13	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	A	.	.	A	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
La14	.	.	.	.	.	.	.	.	.	.	A	A	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	C	T	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G
La15	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	A	.	.	.	A	C	A	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	.	.	.
La25	A	.	T	.	C	.	.	.	.	.	.	A	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	C	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
La27	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	C	.	.	.	.	.	A	.	.	.	C	.	.	.	.	.	.	C	T	A	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
La28	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
La31	.	.	.	.	.	.	.	.	.	.	A	A	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	C	T	A	.	A	.	.	.	.	.	.	T	.	.	.	.	.	.	G
La32	.	.	.	.	C	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	A	.	.	A	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	C	.
La36	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	C	.	.	.	.	.	.	.	.	C	C	.	.	.	C	T	.	C	T	A	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
La37	A	.	T	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	A	.	.	.	C	.	.	.	A	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
La46	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
La54	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	C	.	.	.	.	.	.	.	.	C	C	.	.	.	C	.	.	C	T	A	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
La58	.	.	T	.	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
La60	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	C	T	A	.	A	.	.	.	.	.	.	T	.	.	.	.	.	.	G
La62	.	.	.	.	C	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	C	.	.	.	.	.	A	.	G	.	C	.	.	.	.	.	.	.	T	A	A	.	G	T	.	.	.	.	.	.	.	.	.	.	.	.
La105	A	.	T	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	C	.	.	.	.	.	T	.	T	A	A	.	.	T	.	.	.	.	.	.	T	.	.	.	.	.
La106	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
La108	A	.	T	.	C	.	.	.	.	.	.	A	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	d	.	A	.	C	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
La116	A	T	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	A	A	C	.	.	A	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.
La118	.	.	.	.	C	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	T	A	A	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.
La120	.	.	.	.	.	C	.	.	T	.	.	.	.	.	.	.	.	.	.	A	.	.	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	T	A	.	A	.	.	.	.	A	.	.	.	.	.	.	.	.	.
La125	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	T	.	C	T	A	.	A	.	.	.	.	.	.	T	.	.	.	.	.	.	G
Ma3	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	A	.	.	G	T	.	.	.	.	.	.	.	.	.	.	.	.
Ma6	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	C	T	A	.	A	.	.	.	.	.	.	T	.	.	.	.	.	.	.
Ma11	.	.	.	.	.	.	.	.	.	.	.	A	A	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	T	A	A	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.
Ma18	.	.	.	.	.	.	.	.	T	A	.	A	.	.	.	.	G	.	.	A	.	.	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Ma20	.	.	.	.	.	.	G	C	T	A	.	A	.	.	.	.	G	.	.	A	.	.	.	T	.	.	.	.	.	.	.	.	.	C	.	.	.	C	T	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Ma21	A	.	T	.	C	.	.	.	T	.	.	.	.	.	.	A	.	.	.	A	.	.	.	C	.	.	.	.	.	.	.	.	.	C	.	.	A	.	.	.	C	T	A	.	A	.	.	.	.	.	.	T	.	.	.	.	.	.	G
Ma23	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	A	.	.	.	C	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Ma24	.	.	.	.	.	.	.	.	.	.	A	A	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	C	T	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G
Ma31	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	A	.	.	.	.	.	.	.	C	.	.	.	.	.	.	C	T	A	.	.	.	T	.	.	.	.	.	.	.	.	.	T	A	.
Ma35	.	.	.	.	C	.	.	.	.	.	.	A	.	.	.	.	.	.	.	G	.	.	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	T	A	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Ma37	.	.	.	.	.	.	G	C	T	A	.	A	.	.	.	.	G	.	.	A	.	.	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Ma43	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	C	.	.	.	.	.	A	.	.	.	C	.	.	.	.	.	.	C	T	A	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Ma50	.	.	T	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	T	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Ma53	.	.	.	.	.	.	G	C	T	A	.	A	.	.	.	.	G	.	.	A	.	.	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	T	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Ma56	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	T	A	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.
Ma57	.	.	.	.	.	.	.	.	T	.	.	A	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	C	T	A	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.
Ma60	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	A	.	.	C	.	.	C	.	.	.	.	.	.	A	.	C	C	A	.	.	.	.	.	.	.	.	.	.	.	T	T	.	.	.	.	.	.	.	T	.	.	.
Ma74	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	A	G	A	C	.	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	T	A	A	.	.	T	.	.	.	.	.	A	.	.	.	.	.	.

Figure 1.—Nucleotide and length polymorphisms in the *Acp26Aa* and *Acp26Ab* transcription units in four natural populations of *D. melanogaster.* The polyadenylation signal of the *Acp26Aa* gene (Monsma and Wolfner 1988) was considered for simplicity the limit between transcription units *Acp26Aa* and *Acp26Ab.* Polymorphisms were numbered as in Aguadé *et al.* (1992). Lines from North Carolina (NC), Montblanc (Mo), Lamto (La), and Malawi (Ma) are presented sequentially. Thick-bordered boxes above the sequence information represent the two exons of *Acp26Aa* and the second exon of *Acp26Ab,* while thin-bordered boxes represent the single intron of each gene. Dots represent the same nucleotide as in the first sequence. n, Nonsynonymous; *, singletons as referred to the complete data set; d, deletion. The extent of deletions is indicated by the location of their first and last nucleotide. Nucleotide polymorphisms included in a region for which a deletion is segregating are included in square brackets.

Acp26Ab

	1034	1042	1043	1060	1064	1074	1120	1143	1155	1197	1218	1260	1267	1276	1278	1308	1320	1332	1331	1344	1350	1378	1383	1407	1409	1418	1421	1426	1427	1452	1458	1459	1463	1464	1469-74	1490	1491	1499	1503	1511	1512	1516	1523	1533	1548	1554	1555	1557	1561	1570	1572	[1578]	1575-82
			*	*		*				*			*	*			*							*					*	*	*	*	*	*	*			*				*	*			*	*		*	*			
																		n				n		n	n																												
NC1	C	G	G	C	A	T	T	G	T	C	C	T	T	C	A	C	G	A	A	A	C	A	G	G	C	G	T	C	A	T	T	G	T	T	C	A	T	C	G	T	A	T	T	A	G	A	C	A	T	C	T	C	C
NC2	.	.	.	.	.	.	C	.	.	.	.	C	.	.	.	T	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NC3	.	.	.	.	.	.	C	.	C	.	.	.	.	.	.	.	.	C	.	.	.	G	A	.	.	C	G	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	G	.	.	C	T	.
NC4	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	T	.	C	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	A	.	.	.	.	C	.	.	.	.	.	.	.	d	d
NC5	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NC6	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	T	.	C	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	A	.	.	.	.	C	.	.	.	.	.	.	.	d	d
NC7	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NC8	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	T	.	C	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	A	.	.	.	.	C	.	.	.	.	.	.	.	d	d
NC9	.	.	.	.	.	.	C	.	C	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	A	.	.	.	.	C	.	.	.	.	.	.	.	d	d
NC10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Mo13a	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Mo29b	.	.	.	.	.	.	C	.	C	.	.	C	.	.	.	T	T	C	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	A	C	.	.	.	.	.	.	.	G	.	.	C	.	.
Mo34a	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Mo36a	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	A	C	.	.	.	.	.	.	.	G	.	.	C	.	.
Mo37a	.	.	.	.	T	.	C	.	.	.	.	C	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Mo40b	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	C	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Mo47a	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Mo52b	.	.	.	.	.	.	C	.	C	.	.	C	.	.	.	.	.	C	.	T	T	G	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Mo79b	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	A	C	.	.	.	.	.	.	.	G	.	.	C	.	.
Mo80b	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.
La3	.	.	.	.	.	.	C	.	C	.	.	.	.	.	.	.	.	C	G	.	.	G	A	.	.	C	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	C	.	.
La10	.	.	.	.	.	.	C	.	C	.	.	.	.	.	.	.	.	C	G	.	.	G	A	.	.	C	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	G	.	.	C	.	.
La13	.	.	.	.	T	.	.	.	C	.	.	.	.	.	.	.	.	C	G	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	A	.	.	.	G	C	.	.	T	.	.	.	.	.	.
La14	.	.	.	.	.	.	C	.	C	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	T	.	.	A	C	.	.	.	.	.	.	.	G	.	.	C	.	.
La15	.	.	.	.	.	.	C	.	C	.	.	.	.	T	.	.	.	C	G	.	.	G	A	.	.	C	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	C	.	.
La25	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	T	.	.	.	.	.	.	.	.	.	.	.	.
La27	.	.	.	.	.	.	C	.	C	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
La28	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	C	.	.	T	T	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	A	C	.	.	.	.	.	.	.	G	.	.	C	.	.
La31	.	.	.	.	.	.	C	.	C	.	.	.	.	.	.	.	.	C	G	.	.	G	.	.	A	.	.	.	.	.	.	.	.	.	.	T	.	.	A	C	.	.	.	.	.	.	.	G	.	.	C	.	.
La32	.	.	.	.	.	.	C	.	C	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	T	.	.	A	C	.	.	.	.	.	.	.	G	.	.	C	.	.
La36	.	.	.	A	.	.	C	.	C	.	T	C	.	.	G	.	.	C	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	A	.	A	.	.	.	.	C	.	.	.	.	.	.	.	d	d
La37	.	.	.	.	.	.	C	.	C	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	C	.	.
La46	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.
La54	.	.	.	.	.	.	C	.	C	.	T	C	.	.	G	.	.	C	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	A	.	A	.	.	.	.	C	T	.	.	.	.	.	.	d	d
La58	.	.	.	.	.	.	C	.	C	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	C	.	.
La60	T	.	.	.	.	.	C	.	C	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	C	.	.
La62	.	.	.	.	.	.	C	A	C	.	.	C	.	.	.	.	.	C	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	A	C	.	.	.	.	.	.	.	.	.	.	C	.	.
La105	.	.	.	.	.	.	C	.	G	.	.	C	.	.	.	T	.	C	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	A	C	.	.	.	.	.	.	.	G	.	.	C	.	.
La106	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	C	.	.	.	G	.	.	A	.	.	.	.	.	.	.	.	.	.	T	.	.	A	C	.	.	.	.	.	.	.	G	.	.	C	.	.
La108	.	.	.	.	T	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
La116	.	.	.	.	.	.	C	.	C	.	.	C	.	.	.	T	.	C	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	T	.	.	.	.	T	.	.	A	.	C	.	.
La118	.	.	.	.	.	.	C	A	C	.	.	C	.	.	.	.	.	C	G	.	.	G	.	.	A	.	.	.	.	.	.	.	.	.	.	T	.	.	A	C	.	.	.	.	.	.	.	G	.	.	C	.	.
La120	.	.	.	.	.	.	C	.	C	T	.	.	.	.	.	.	.	C	.	.	T	G	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	A	C	.	.	.	.	.	.	.	G	.	.	C	.	.
La125	T	.	.	.	.	.	C	.	C	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	C	.	.
Ma3	.	.	.	.	.	.	C	A	C	.	.	C	.	.	.	.	.	C	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Ma6	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	G	.	.	A	.	.	.	.	.	.	.	.	.	.	.	A	.	A	.	.	.	.	C	.	.	.	.	.	.	.	d	d
Ma11	.	.	.	.	T	.	C	.	C	.	.	C	C	.	.	.	.	C	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	A	C	.	.	.	.	.	.	.	G	.	.	C	.	.
Ma18	.	.	.	.	.	C	C	.	C	.	.	.	.	.	.	T	.	C	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	A	C	.	.	.	.	.	.	.	G	.	.	C	.	.
Ma20	.	.	.	.	T	.	.	.	C	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Ma21	.	.	.	.	.	.	C	.	C	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	T	.	.	A	C	.	.	.	.	.	.	.	G	.	.	C	.	.
Ma23	.	.	.	.	.	.	C	.	C	.	.	C	.	.	.	T	.	C	G	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	A	C	.	.	.	.	.	.	.	G	.	.	C	.	.
Ma24	.	.	.	.	.	.	C	.	C	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	T	.	.	A	C	.	.	.	.	.	.	.	G	.	.	C	.	.
Ma31	.	.	.	.	T	.	C	.	C	.	.	.	.	.	.	T	.	C	.	T	T	G	A	.	A	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	C	T	.
Ma35	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	C	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	A	C	.	.	.	.	.	.	.	G	.	.	C	.	.
Ma37	.	.	.	.	.	.	C	.	C	.	.	.	.	.	.	.	.	C	.	.	.	G	.	.	A	.	.	.	.	.	.	.	.	.	.	.	A	.	A	.	.	.	.	C	.	.	.	.	.	.	.	d	d
Ma43	.	.	.	.	.	.	C	.	C	.	.	C	.	.	.	T	.	C	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	A	C	.	.	.	.	.	.	.	G	.	.	C	.	.
Ma50	.	.	.	.	.	.	C	.	C	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	T	.	.	A	C	.	.	.	.	.	.	.	G	.	.	C	.	.
Ma53	.	C	T	.	.	.	.	.	C	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	d	A	C	A	d	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Ma56	.	.	.	.	.	.	C	A	C	.	.	.	.	.	.	.	.	C	.	.	.	G	.	.	A	.	.	.	.	.	.	.	.	.	.	.	A	.	A	.	.	.	.	C	.	.	.	.	.	.	.	d	d
Ma57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	G	.	.	A	.	.	.	.	.	.	.	.	.	.	.	A	.	A	.	.	.	.	C	.	.	.	.	.	.	.	d	d
Ma60	.	.	.	.	.	.	C	.	C	.	.	C	.	.	.	T	.	C	.	.	T	G	.	A	.	.	.	A	C	C	.	.	.	.	.	T	.	.	A	C	.	C	.	.	.	.	.	G	.	.	.	.	.
Ma74	.	C	.	.	.	.	C	.	C	.	.	.	.	.	.	.	.	C	.	.	T	G	.	.	A	.	.	.	.	.	.	.	.	.	.	T	.	.	A	C	.	.	.	.	.	.	.	G	.	.	C	.	.

Figure 1.—*Continued.*

applications of the HKA test is consistent with the excess of nonsynonymous divergence in the *Acp26Aa* gene already detected using the MK test.

DISCUSSION

Pattern of nucleotide variation: In the few studies that compare nucleotide variation between East African and non-African populations, the estimated nucleotide variation (measured as π or θ) was much higher in the African populations, while the level of linkage disequilibrium was higher in the non-African samples (Begun and Aquadro 1993, 1995). The analysis of silent (and presumably neutral) variation within the Adh^S allele in samples from both East and West Africa and in non-African samples revealed a major differentiation of non-

TABLE 1

Nucleotide polymorphism

Region	No. sites	North Carolina			Montblanc			Lamto			Malawi		
		S	π	θ	S	π	θ	S	π	θ	S	π	θ
Whole region													
All	1614	34(2)	0.008	0.007	39(1)	0.010	0.008	78(15)	0.011	0.013	79(20)	0.012	0.014
Silent[a]	784	23(2)	0.010	0.010	27(1)	0.014	0.012	54(12)	0.015	0.018	59(17)	0.018	0.022
Acp26Aa													
5′	89	2	0.012	0.008	2	0.012	0.008	3(1)	0.009	0.009	2	0.004	0.007
Synonymous	174	7(1)	0.010	0.014	9	0.023	0.018	13(1)	0.018	0.020	20(5)	0.024	0.330
Nonsynonymous	615	10	0.006	0.006	10	0.008	0.006	20(3)	0.009	0.009	16(2)	0.006	0.008
Intron	56	1	0.006	0.006	1	0.010	0.006	2	0.007	0.010	3	0.021	0.016
3′	119	0	0.000	0.000	0	0.000	0.000	3(2)	0.004	0.007	5(3)	0.005	0.012
Silent	438	10(1)	0.008	0.008	12	0.013	0.010	21(4)	0.011	0.013	30(8)	0.015	0.020
Acp26Ab													
5′	67	0	0.000	0.000	1	0.003	0.005	3(1)	0.006	0.012	4(2)	0.011	0.017
Synonymous	55	3	0.023	0.019	7(1)	0.033	0.045	9(2)	0.033	0.044	7(1)	0.031	0.037
Nonsynonymous	215	2	0.005	0.003	2	0.005	0.003	4	0.006	0.005	3(1)	0.005	0.004
Intron	61	2	0.015	0.012	2	0.017	0.012	4	0.012	0.018	3	0.014	0.014
3′	163	8(1)	0.015	0.017	5	0.014	0.010	17(5)	0.025	0.027	15(6)	0.025	0.027
Silent	346	13(1)	0.013	0.013	15(1)	0.015	0.015	33(8)	0.020	0.025	29(9)	0.022	0.024

Singletons (referred to the complete data set) are given in parentheses.

[a] Silent in noncoding regions and synonymous in coding regions, S, π, and θ refer to the number of segregating sites, nucleotide diversity (Nei 1987), and the Watterson parameter or expected nucleotide heterozygosity (Watterson 1975), respectively.

	Acp26Aa																								Acp26Ab			
	5	10*	14	19	24	25	32	42	44*	46	54*	65*	76	79	85*	86*	101	104	109	167	169	207	221	253	261	64	80	89*	90
NC1	Leu	Ileu	Leu	Ser	Gln	Gln	Leu	Leu	Asn	Pro	Ala	Asp	Asp	Asp	Tyr	Pro	Ser	Pro	Ileu	Leu	Leu	Ser	Arg	Glu	Pro	Gln	Ileu	Met	Ala
NC2	.	.	.	.	.	.	.	.	.	Thr	.	.	.	.	.	.	.	.	.	.	.	Ileu	Lys	.	.	.	Val	.	.
NC3	.	.	.	.	.	.	.	.	.	Thr	.	.	Asn	.	.	.	.	.	Leu	.	.	.	.	.	.	His	Val	.	.
NC4	.	.	.	.	.	.	.	.	.	.	.	.	.	Asn	.	.	Asn	.	.	.	.	.	.	.	.	His	Val	.	.
NC5	.	.	.	Asn	Lys	.	.	.	.	.	.	.	.	.	.	.	Asn	.	Leu	.	.	.	.	.	.	.	.	.	.
NC6	.	.	.	.	.	.	.	.	.	.	.	.	.	Asn	.	.	Asn	.	.	.	.	.	.	.	.	His	Val	.	.
NC7	.	.	.	Asn	Lys	.	.	.	.	.	.	.	.	.	.	.	Asn	.	Leu	.	.	.	.	.	.	.	.	.	.
NC8	.	.	.	.	.	.	.	.	.	.	.	.	.	Asn	.	.	Asn	.	.	.	.	.	.	.	.	His	Val	.	.
NC9	.	.	.	.	.	.	Gln	.	.	Thr	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	His	.	.	.
NC10	.	.	.	Asn	Lys	.	.	.	.	.	.	.	.	.	.	.	Asn	.	Leu	.	.	.	Lys	.	.	.	.	.	.
13a	.	.	.	Asn	.	Lys	.	.	.	.	.	.	.	.	.	.	Asn	.	Leu	.	.	.	.	.	.	.	Val	.	.
29b	.	.	.	.	.	.	Gln	.	.	.	.	.	.	.	.	.	Asn	.	Leu	.	.	.	Lys	.	.	His	Val	.	.
34a	.	.	.	Asn	.	Lys	.	.	.	.	.	.	.	.	.	.	Asn	.	Leu	.	.	.	Lys	.	.	.	Val	.	.
36a	.	.	.	.	.	.	Gln	.	.	Thr	.	.	.	.	.	.	Asn	.	Leu	.	.	Ileu	.	.	Ser	.	.	.	.
37a	.	.	.	.	.	.	.	.	.	Ala	.	.	.	.	.	.	Asn	.	Leu	.	.	Ileu	.	.	Ser	.	.	.	.
40b	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	His	Val	.	.
47a	.	.	.	Asn	.	.	.	.	.	Ala	.	.	.	.	.	.	Asn	.	Leu	.	.	Ileu	.	.	Ser	.	.	.	.
52b	.	.	.	.	.	.	Gln	.	.	Thr	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	His	Val	.	.
79b	.	.	.	.	.	.	Gln	.	.	Thr	.	.	.	.	.	.	Asn	.	Leu	.	.	Ileu	.	.	Ser	.	.	.	.
80b	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Asn	.	.	.	.	.	Lys	.	.	.	.	.	.
L3	.	.	.	Asn	.	.	Gln	.	.	Thr	.	.	.	.	.	.	.	.	.	.	Phe	.	Lys	.	Ser	His	Val	.	.
L10	.	.	.	Asn	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Ileu	Lys	.	.	His	Val	.	.
L13	.	.	.	.	.	.	.	.	.	.	.	.	.	Asn	.	.	Asn	.	.	.	.	.	.	.	.	His	Val	.	.
L14	.	.	Ileu	Asn	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Ileu	.	.	.	His	.	.	Glu
L15	.	.	.	Asn	.	.	Gln	.	.	Thr	.	.	.	.	.	.	.	.	.	.	.	.	Lys	.	Ser	His	Val	.	.
L25	.	.	.	Asn	.	Lys	.	.	.	.	.	.	.	.	.	.	Asn	.	Leu	.	.	.	.	.	.	.	.	.	.
L27	.	.	.	.	.	Lys	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Ileu	Lys	.	.	.	Val	.	.
L28	Ser	.	.	.	.	.	.	.	.	Ala	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	His	Leu	.	.
L31	.	.	Ileu	Asn	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Ileu	Lys	Asp	.	His	Val	.	Glu
L32	.	.	.	.	.	.	.	.	.	.	.	.	.	Asn	.	.	Asn	.	.	.	.	.	.	Asp	.	His	.	.	Glu
L36	.	.	.	.	.	.	.	.	.	Thr	.	.	.	.	.	.	.	.	Leu	Ser	Phe	Ileu	.	.	.	His	.	.	.
L37	.	.	.	.	.	.	Gln	.	.	Thr	.	.	.	Asn	.	.	.	.	.	.	.	.	Lys	.	.	His	.	.	.
L46	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
L54	.	.	.	.	.	.	.	.	.	Thr	.	.	.	.	.	.	.	.	Leu	Ser	.	Ileu	.	.	.	His	.	.	.
L58	.	.	.	.	.	.	Gln	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	His	.	.	.
L60	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Ileu	Lys	Asp	.	His	.	.	Glu
L62	.	.	.	.	.	Lys	.	.	.	.	.	.	.	.	.	.	.	Arg	.	.	.	Ileu	.	.	.	His	Val	.	.
L105	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Asn	.	.	.	.	Ileu	.	.	Ser	His	Val	.	.
L106	Ser	.	.	.	.	.	.	.	.	Ala	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	His	Val	.	Glu
L108	.	.	.	Asn	.	Lys	.	.	.	.	.	.	.	.	d	Ser	Asn	.	Leu	.	.	.	.	.	.	.	.	.	.
L116	.	.	.	.	.	.	.	.	.	.	Thr	.	Asn	.	.	.	.	.	.	.	.	.	.	.	.	His	Val	.	.
L118	.	.	Ileu	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Ileu	.	.	.	His	Val	.	Glu
L120	.	Leu	.	.	.	.	.	.	.	Thr	.	.	.	.	.	.	.	.	.	.	.	Ileu	Lys	.	.	His	Val	.	.
L125	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Phe	Ileu	Lys	Asp	.	His	.	.	Glu
Ma3	.	.	.	.	.	.	.	.	.	.	.	Gly	.	.	.	.	.	.	.	.	.	Ileu	.	.	.	His	Val	.	.
Ma6	.	.	.	Asn	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Ileu	Lys	Asp	.	His	Val	.	Glu
Ma11	.	.	.	Asn	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Ileu	.	.	.	His	.	.	.
Ma18	.	.	.	Asn	.	.	.	Val	.	Thr	.	.	.	.	.	.	.	.	.	.	.	.	Lys	.	.	His	Val	.	.
Ma20	.	.	.	Asn	.	.	.	Val	.	Thr	.	.	.	.	.	.	.	.	.	Ser	Phe	.	Lys	.	.	His	.	.	.
Ma21	.	.	.	.	.	.	Gln	.	.	Thr	.	.	.	.	.	.	.	.	.	.	.	Ileu	Lys	Asp	.	His	.	.	Glu
Ma23	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	His	Val	.	.
Ma24	.	.	Ileu	Asn	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Ileu	.	.	.	His	.	.	Glu
Ma31	.	.	.	.	.	.	.	.	.	Thr	.	.	.	.	.	.	.	.	.	.	.	Ileu	.	.	.	His	Val	.	Glu
Ma35	.	.	.	Asn	.	.	.	.	.	Ala	.	.	.	.	.	.	.	.	.	.	.	Ileu	Lys	.	.	His	.	.	.
Ma37	.	.	.	Asn	.	.	.	Val	.	Thr	.	.	.	.	.	.	.	.	.	.	.	.	Lys	.	.	His	Val	.	Glu
Ma43	.	.	.	.	.	Lys	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Ileu	Lys	.	.	His	Val	.	.
Ma50	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Ileu	.	.	.	His	.	.	Glu
Ma53	.	.	.	Asn	.	.	.	Val	.	Thr	.	.	.	.	.	.	.	.	.	.	.	Ileu	.	.	.	His	.	.	.
Ma56	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Ileu	.	.	.	His	Val	.	Glu
Ma57	.	.	.	Asn	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Ileu	.	Asp	.	His	Val	.	Glu
Ma60	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Asn	.	Leu	.	.	.	.	.	.	His	Val	Ileu	.
Ma74	.	.	.	Asn	.	.	.	.	Ser	Thr	.	.	.	.	.	.	.	.	.	.	.	Ileu	.	.	.	His	Val	.	Glu
	NC	C	C	C	R	R	NC	C	C	NC	NC	R	R	R		NC	C	R	C	NC	C	NC	C	C	NC	R	C	C	R

Figure 2.—Amino acid replacement polymorphism in Acp26Aa and Acp26Ab in four natural populations of *D. melanogaster*. Numbers indicate the amino acid residue in the *D. melanogaster* protein. Dots indicate same amino acid as in the first sequence. *, Singletons as referred to the complete data set; d, deletion; C, conservative; NC, nonconservative without charge change; R, nonconservative with charge change.

TABLE 2

Population differentiation

	Whole region		*Acp26Aa*		*Acp26Ab*	
	F_{st}	*P*	F_{st}	*P*	F_{st}	*P*
All sites						
NC-Mo	0.09	0.031	0.12	0.057	0.03	0.099
NC-La	0.14	0.001	0.08	0.023	0.20	0.000
NC-Ma	0.19	0.000	0.17	0.000	0.21	0.001
Mo-La	0.14	0.000	0.17	0.000	0.09	0.009
Mo-Ma	0.19	0.000	0.23	0.000	0.14	0.002
La-Ma	0.02	0.066	0.02	0.071	0.01	0.147
Silent sites						
NC-Mo	0.11	0.029	0.16	0.111	0.05	0.127
NC-La	0.17	0.012	0.09	0.112	0.22	0.000
NC-Ma	0.18	0.001	0.18	0.000	0.19	0.002
Mo-La	0.14	0.000	0.25	0.002	0.07	0.029
Mo-Ma	0.17	0.000	0.24	0.000	0.08	0.028
La-Ma	0.02	0.085	0.03	0.040	0.01	0.179

NC, North Carolina; Mo, Montblanc; La, Lamto; Ma, Malawi; *P*, probability.

African and West African populations relative to East African samples (Bénassi and Veuille 1995).

In this study variation both at the *Acp26Aa* and *Acp-26Ab* genes was surveyed in two non-African samples and in a subsample of the same two African populations studied by Bénassi and Veuille (1995). As in previous studies the level of variation was higher in African than in non-African populations, although the differences were not very conspicuous (Table 1). In addition, both genes presented similar levels of nonsynonymous variation in African and non-African populations. As in previous nuclear DNA studies comparing African and non-African populations, polymorphisms in the North American and European samples were a subset of those present in the two African populations. However, in contrast to previous studies (Begun and Aquadro 1993, 1995) most of the nonrare polymorphisms segregating in the African samples were also segregating in the non-African samples. Also the two African samples shared most nonunique polymorphisms.

Unlike the study of the *Adh* region (Bénassi and Veuille 1995), no major differentiation for the *Acp26A* region was detected between West and East Africa. F_{st} estimates were rather low both for the *Acp26Aa* (0.02–0.03) and *Acp26Ab* (0.01) transcription units. In the case of the *Acp26Aa* region the two populations were marginally or slightly differentiated, while in the case of the *Acp26Ab* region they did not show any differentiation. In contrast to silent variation within Adh^S, variation at the *Acp26Aa* and *Acp26Ab* regions might be governed by selection (see below). Also non-African populations showed in general a strong differentiation from both East and West Africa; this differentiation could be, as for other genes, both a result of drift during the out-of-Africa expansion and of selection acting differentially during that process.

Nonsynonymous variation at the *Acp26Aa* gene is governed by directional selection: When the whole gene was considered, K_a and K_s estimates between *D. melanogaster* and the closely related species *D. simulans*, *D. mau-*

TABLE 3

Linkage disequilibrium

Region	No. comparisons	No. significant comparisons[a]	% significant comparisons
Whole region			
North Carolina	190	53	0.28
Montblanc	465	73	0.16
Lamto	1378	164(17)	0.12
Malawi	630	72(5)	0.11
Africa	1711	215(19)	0.13
Acp26Aa			
North Carolina	55	16	0.29
Montblanc	190	38	0.20
Lamto	378	39(2)	0.10
Malawi	171	26(2)	0.15
Africa	528	65(9)	0.12
Acp26Ab			
North Carolina	36	9	0.25
Montblanc	55	15	0.27
Lamto	300	45(15)	0.15
Malawi	136	28(3)	0.21
Africa	325	63(15)	0.19

[a] Number of significant comparisons using the χ^2 test without and with (in parentheses) the Bonferroni correction.

TABLE 4

McDonald and Kreitman tests

Acp26Aa						*Acp26Ab*					
	F_s	F_{ns}	P_s	P_{ns}	Probability		F_s	F_{ns}	P_s	P_{ns}	Probability
NC *vs.* sim	24	78	7	9	0.109	NC *vs.* sim	2	3	3	2	0.55
NC *vs.* mau	22	80	7	9	0.075	NC *vs.* mau	9	3	3	2	0.57
NC *vs.* sech	23	79	7	9	0.091	NC *vs.* sech	8	5	3	2	0.96
Mo *vs.* sim	24	77	9	9	0.031*	Mo *vs.* sim	2	3	7	2	0.19
Mo *vs.* mau	22	78	9	9	0.021*	Mo *vs.* mau	9	3	7	2	0.89
Mo *vs.* sech	23	77	9	9	0.026*	Mo *vs.* sech	8	5	7	2	0.43
La *vs.* sim	21	76	13	19	0.042*	La *vs.* sim	2	3	9	4	0.28
La *vs.* mau	19	78	13	19	0.022*	La *vs.* mau	9	3	9	4	0.76
La *vs.* sech	20	77	13	19	0.031*	La *vs.* sech	8	5	9	4	0.69
Ma *vs.* sim	20	77	19	15	0.002**	Ma *vs.* sim	2	3	7	3	0.29
Ma *vs.* mau	18	78	19	15	0.0001***	Ma *vs.* mau	9	7	7	3	0.80
Ma *vs.* sech	19	77	19	15	0.0001***	Ma *vs.* sech	8	5	7	3	0.68
Af *vs.* sim	20	75	22	22	0.0007***	Af *vs.* sim	2	3	11	5	0.28
Af *vs.* mau	18	77	22	22	0.0002***	Af *vs.* mau	9	3	11	5	0.73
Af *vs.* sech	19	76	22	22	0.0004***	Af *vs.* sech	8	5	11	5	0.69
mel *vs.* sim	20	75	25	23	0.0002***	mel *vs.* sim	2	3	12	5	0.24
mel *vs.* mau	18	77	25	23	0.0001***	mel *vs.* mau	9	3	12	5	0.80
mel *vs.* sech	19	76	25	23	0.0001***	mel *vs.* sech	8	5	12	5	0.61

F_s, number of synonymous fixed differences between species; F_{ns}, number of nonsynonymous fixed differences between species; P_s, number of synonymous segregating sites; P_{ns}, number of nonsynonymous segregating sites. NC, North Carolina; Mo, Montblanc; La, Lamto; Ma, Malawi; Af, Africa; mel, all *D. melanogaster* lines; sim, *D. simulans*; mau, *D. mauritiana*; sech, *D. sechellia*. Probability established with a *G*-test with Williams' correction for continuity. *$0.01 < P < 0.05$; **$0.001 < P < 0.01$; ***$P < 0.001$.

ritiana, and *D. sechellia* were rather similar, and no significant excess of replacement changes could be detected (Aguadé *et al.* 1992). However, when the region corresponding to the second exon of the gene was compared between both *D. melanogaster* and *D. simulans*, and the more distantly related species *D. yakuba* and *D. teissieri* (Tsaur and Wu 1997), the average K_a/K_s ratio for the four comparisons was ~1.5; only the K_a/K_s ratio for the branch separating both species pairs was significantly >1, which constituted unambigous evidence that at least in the past the evolution of Acp26Aa had been driven by positive selection.

In this study another three populations of *D. melanogaster* were sampled. Application of the MK test to the different populations surveyed (and to the complete *D. melanogaster* sample) revealed a significant excess of nonsynonymous fixed differences between *D. melanogaster* and each of the three sibling species (Table 4). This excess of nonsynonymous divergence was also revealed by the results of the HKA tests performed comparing all variation (most tests significant) and only silent variation (all tests nonsignificant) between the *Acp26Aa* and *Acp26Ab* regions (analysis not shown). We can now assert that directional selection not only drove the evolution of Acp26Aa in the distant past (between the *D. melanogaster*-*D. simulans* split, 2.5 mya, and the *D. yakuba*-*D. teissieri* split, 6 mya) but also in the last 2.5 my.

Acp26Aa is processed in the female genital tract just after mating; the protein is cleaved three times after residues 48, 68, and 115 or 117, generating four peptides (Park and Wolfner 1995). Only the processed protein is active, and secretions from the male accessory gland main cells are needed for the processing. As one of the known functions of Acp26Aa is the stimulation of egg laying in the mated female, the last peptide would be the candidate for that function as it contains the region of similarity to the ELH in the four species of the *melanogaster* complex. To test whether positive selection has driven the evolution of the different parts of the protein, the *Acp26Aa* coding region was first divided in two fragments that included the first three peptides and the last peptide, respectively. Application of the MK test revealed an excess of fixed nonsynonymous differences in both fragments, although significance was higher in the second fragment (Table 5). To further explore the distribution of fixed and polymorphic changes, both fragments were subdivided, despite the fact that subdivision could lower the statistical power of the MK test; the first fragment was divided in two, corresponding to the regions coding for the first two peptides and for the third peptide, and the second was in two equally sized fragments (Table 5). Although the contribution of the different peptides to the function(s) of the Acp26Aa protein is not known, this analysis indicates that directional selection is governing at least the evolution of the third and fourth peptides in the *melanogaster* species

TABLE 5

McDonald and Kreitman tests along the *Acp26Aa* gene

AA 19–117	F_s	F_{ns}	P_s	P_{ns}	Probability	AA 118–264	F_s	F_{ns}	P_s	P_{ns}	Probability
mel *vs.* sim	4	32	10	14	0.007**	mel *vs.* sim	14	41	13	6	0.0011***
mel *vs.* mau	4	33	10	14	0.006**	mel *vs.* mau	13	42	13	6	0.0006***
mel *vs.* sech	4	31	10	14	0.009**	mel *vs.* sech	14	43	13	6	0.0009***
AA 19–68	F_s	F_{ns}	P_s	P_{ns}	Probability	AA 118–189	F_s	F_{ns}	P_s	P_{ns}	Probability
mel *vs.* sim	2	10	6	10	0.234	mel *vs.* sim	10	21	2	2	0.520
mel *vs.* mau	1	13	6	10	0.048*	mel *vs.* mau	8	22	2	2	0.391
mel *vs.* sech	2	10	6	10	0.234	mel *vs.* sech	9	22	2	2	0.443
AA 69–117	F_s	F_{ns}	P_s	P_{ns}	Probability	AA 190–264	F_s	F_{ns}	P_s	P_{ns}	Probability
mel *vs.* sim	2	22	4	4	0.021*	mel *vs.* sim	4	20	11	4	0.0004***
mel *vs.* mau	3	20	4	4	0.051	mel *vs.* mau	5	20	11	4	0.0004***
mel *vs.* sech	2	21	4	4	0.023**	mel *vs.* sech	5	21	11	4	0.0007***

AA*x–y*, region coding for amino acid residues *x* through *y*. F_s, number of synonymous fixed differences between species; F_{ns}, number of nonsynonymous fixed differences between species; P_s, number of synonymous segregating sites; P_{ns}, number of nonsynonymous segregating sites. mel, all *D. melanogaster* lines; sim, *D. simulans*; mau, *D. mauritiana*; sech, *D. sechellia*. Probability established with a *G*-test with Williams' correction for continuity. *$0.01 < P < 0.05$; **$0.001 < P < 0.01$; ***$P < 0.001$.

complex; it also indicates that selection might be acting differentially along the fourth peptide.

Directional selection is, however, expected to cause a reduction of polymorphism in the region closely linked to the favored variant. Although all *D. melanogaster* populations harbored an important amount of nonsynonymous and synonymous polymorphisms at the *Acp26Aa* region, some regions showed a very low level of polymorphism without the corresponding reduction in the levels of fixed differences. In fact, in the sample from North Carolina the region extending from polymorphic site 417 to the end of the *Acp26Aa* transcription unit (Figure 1) presented 5 singletons out of 6 polymorphisms; also in the sample from Malawi the region spanning polymorphic sites 270 to 693 presented 10 singletons out of 11 polymorphisms. Nevertheless, it could be argued that the sample from North Carolina is not at equilibrium, and that the observed distribution of polymorphism is still reflecting its rather recent origin; this argument would generally not be made for the East African sample. Also the extent of the fragment with reduced polymorphism in Malawi (~400 bp) would be consistent with the region expected to be affected by a selective sweep in a region of high recombination like the *Acp26Aa* region. On the other hand, as the coalescence time of any given sample of sequences is $4N$ generations, selective sweeps that occurred before that time would have no effect on the observed polymorphism. Under the assumption of mutation-drift equilibrium this time can be estimated from the ratio between θ ($4N\mu$) and divergence ($2t\mu$). Using restriction map polymorphism in *D. melanogaster*, and divergence between *D. melanogaster* and *D. simulans*, this time was estimated to represent 25% of the time since the split of the *melanogaster* and *simulans* lineages (Eanes *et al.* 1996). Therefore, only selective sweeps occurring in the last ~600,000 years would have any effect on extant polymorphism in *D. melanogaster.*

The *Acp26Aa* and *Acp26Ab* genes have different evolutionary histories as expected in a high recombination region like the *Acp26A* region. The function of the Acp26Ab protein after its transfer to the female is not known; there is no evidence of selection either driving its evolution or maintaining the different protein variants. Unlike the *Acp26Aa* gene, the level of nonsynonymous variation in the *Acp26Ab* coding region is, both within and between species, roughly an order of magnitude lower than that of synonymous variation (Aguadé *et al.* 1992), indicating a certain degree of functional constraint at the protein level.

I thank Michel Veuille and Véronique Bénassi for the African lines and for sharing unpublished information on the chromosomal polymorphism of these lines. I also thank the Umeå Stock Center for the *Df(2L)Gdh-A,dpL/SM1* line, Serveis Científico-Tècnics from Universitat de Barcelona for automated sequencing facilities, and Julio Rozas, Carmen Segarra, and Griselda Ribó for comments on the manuscript. Special thanks are given to David Salguero for his excellent technical assistance. This work was supported by grants PB94-0923 from Dirección General de Investigación Científica y Técnica, Spain, and 1995SGR-577 from Comissió Interdepartamental de Recerca i Tecnologia, Generalitat de Catalunya.

LITERATURE CITED

Aguadé, M., N. Miyashita and C. H. Langley, 1992 Polymorphism and divergence in the *Mst26A* male accessory gland gene region in Drosophila. Genetics **132:** 755–770.

Ashburner, M., 1989 *Drosophila: A Laboratory Handbook.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Begun, D. J., and C. F. Aquadro, 1993 African and North American

populations of *Drosophila melanogaster* are very different at the DNA level. Nature **356:** 519–520.

Begun, D. J., and C. F. Aquadro, 1995 Molecular variation at the *vermillion* locus in geographically diverse populations of *Drosophila melanogaster* and *D. simulans.* Genetics **140:** 1019–1032.

Bénassi, V., and M. Veuille, 1995 Comparative population structuring of molecular and allozyme variation of *Drosophila melanogaster Adh* between Europe, West Africa and East Africa. Genet. Res. **65:** 95–103.

Bénassi, V., S. Aulard, S. Mazeau and M. Veuille, 1993 Molecular variation of *Adh* and *P6* genes in African populations of *Drosophila melanogaster* and its relation to chromosomal inversions. Genetics **134:** 789–799.

Bertram, M. J., D. M. Neubaum and M. F. Wolfner, 1996 Localization of the Drosophila accessory gland protein Acp36DE in the mated female suggests a role in sperm storage. Inst. Biochem. Mol. Biol. **26:** 971–980.

Clark, A. G., M. Aguadé, T. Prout, L. Harshman and C. H. Langley, 1995 Variation in sperm displacement and its association with accessory gland protein loci in *Drosophila melanogaster.* Genetics **139:** 189–201.

Eanes, W. F., M. Kirchner, J. Yoon, C. H. Biermann, I.-N. Wang *et al.*, 1996 Historical selection, amino acid polymorphism and lineage-specific divergence at the G6pd locus in *Drosophila melanogaster* and *D. simulans.* Genetics **144:** 1027–1041.

Fu, Y. X., and W-S. Li, 1993 Statistical tests of neutrality of mutations. Genetics **133:** 693–709.

Herndorn, L. A., and M. F. Wolfner, 1995 A Drosophila seminal fluid protein, Acp26Aa, stimulates egg-laying in females for one day after mating. Proc. Natl. Acad. Sci. USA **92:** 10114–10118.

Higuchi, R. G., and H. Ochman, 1989 Production of single-stranded DNA templates by exonuclease digestion following the polymerase chain reaction. Nucleic Acids Res. **17:** 5865.

Hudson, R. R., and N. L. Kaplan, 1985 Statistical properties of the number of recombination events in the history of a sample of DNA sequences. Genetics **111:** 147–164.

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

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.

Kubli, E., 1996 The Drosophila sex-peptide: a peptide pheromone involved in reproduction. Adv. Dev. Biochem. **4:** 99–128.

Lee, Y., and V. D. Vacquier, 1995 Positive selection is a general phenomenon in the evolution of abalone sperm lysin. Mol. Biol. Evol. **12:** 231–238.

Maddison, W. P., and D. R. Maddison, 1992 *MacClade: Analysis of Phylogeny and Character Evolution, Version 3.0.* Sinauer Associates, Sunderland, MA.

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

Metz, E. C., and S. R. Palumbi, 1996 Positive selection and sequence rearrangements generate extensive polymorphism in the gamete recognition protein bindin. Mol. Biol. Evol. **13:** 397–406.

Monsma, S., and M. F. Wolfner, 1988 Structure and expression of a Drosophila male accessory gland gene whose products resemble a peptide pheromone precursor. Genes Dev. **2:** 1063–1073.

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

Palumbi, S. R., 1994 Genetic divergence, reproductive isolation, and marine speciation. Ann. Rev. Ecol. Syst. **25:** 547–572.

Park, M., and M. F. Wolfner, 1995 Male and female cooperate in the prohormone-like processing of a *Drosophila melanogaster* seminal fluid protein. Dev. Biol. **171:** 694–702.

Rozas, J., and R. Rozas, 1997 DnaSP version 2.0: a novel software package for extensive molecular population genetics analysis. Comput. Appl. Biosci. **13:** 307–311.

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

Tsaur, S.-H., and C.-I. Wu, 1997 Positive selection and the molecular evolution of a gene of male reproduction, *Acp26Aa* of Drosophila. Mol. Biol. Evol. **14:** 544–549.

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

Wolfner, M. F., 1997 Tokens of love: functions and regulation of Drosophila male accessory gland products. Inst. Biochem. Mol. Biol. **27:** 179–192.

Communicating editor: A. G. Clark