Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 1426–1430, February 1989
Population Biology

Amino acid polymorphisms for esterase-6 in *Drosophila melanogaster*

(nucleotide sequence variation/haplotype diversity/targets of selection)

P. H. COOKE*†‡ AND J. G. OAKESHOTT*†

*Research School of Biological Sciences, Australian National University, G.P.O. Box 475, and †Commonwealth Scientific and Industrial Research Organization, Division of Entomology, G.P.O. Box 1700, Canberra, Australian Capital Territory 2601, Australia

Communicated by Bruce Wallace, October 17, 1988

ABSTRACT **High-resolution electrophoresis has revealed 10 allozymes of esterase-6 (EC 3.1.1.1) in *Drosophila melanogaster*. The sequences of 13 isolates of the *Est6* gene covering all 10 allozymes were obtained and 52 nucleotide differences were found. Sixteen of these cause amino acid replacements, of which three result in charge differences whose size and direction are consistent with the electrophoretic mobilities of the allozymes in which they occur. The smeared electrophoretic phenotype of one allozyme can be explained by the loss of a cysteine residue involved in a disulfide bridge. Several minor mobility variants within the major F and S electrophoretic phenotypes differ by amino acid substitutions that are generally conservative for charge but not for some other properties (size, polarity, or hydrophobicity). Four amino acid differences are found among different isolates of the same allozymes and, overall, 12 amino acid haplotypes occur among the 13 isolates sequenced. Nevertheless, the most common variants within F and S are distinguished by only two amino acids (Asn/Asp at 237 and Thr/Ala at 247), and these are the most likely targets for the selection underlying complementary latitudinal clines in F and S frequencies.**

Two major difficulties have beset attempts to elucidate the adaptive significance of specific enzyme polymorphisms. The first concerns the lack of information on gametic disequilibrium between the locus of interest and variation at nearby but unknown loci; the second concerns the true amount and nature of molecular polymorphism underlying electrophoretic phenotypes (1). The first difficulty has largely been overcome for the esterase-6 (EST6; carboxylic-ester hydrolase, EC 3.1.1.1) polymorphism of *Drosophila* because parallel latitudinal clines in gene frequency occur for the same allozymes in both *D. melanogaster* and its sibling species *D. simulans* (2). These species are thought to have diverged 8–10 million years ago (3), suggesting that the *Est6* gene itself is probably the unit under selection, as intergenic gametic disequilibrium should have broken down since the divergence of the two species. This paper addresses the second question, concerning the number and nature of polymorphisms for *Est6*, by comparing the nucleotide sequences of several alleles from *D. melanogaster*.

High-resolution cellulose acetate electrophoresis has shown five major classes of mobility variants for EST6 in *D. melanogaster*, with seven minor mobility variants within the most common major classes, F (fast) and S (slow) (4). Another seven variants within F and S have been resolved by *in vitro* thermostability criteria (5), and preliminary analyses of a small number of lines by both electrophoretic and thermostability procedures suggest that the variants detected by the two techniques are not the same (4). While the adaptive significance of the many minor mobility and thermostability variants is unknown, there is biogeographical (2), biochemical (6), and behavioral (7) evidence that the major F and S classes differ under natural selection. This paper identifies the specific amino acid polymorphisms that cause the biochemical differences between the F and S classes of EST6, which are likely to be the primary targets for the selection underlying the clines.§

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

Cooke *et al.* (4) used high-resolution cellulose acetate electrophoresis to resolve 10 polymorphic EST6 allozymes in a population from Coffs Harbour, Australia. Nine allozymes giving sharp bands on electrophoresis were denoted 1 through 9 in decreasing order of anodal mobilities. Allozyme 10 had a smeared electrophoretic phenotype. These 10 allozymes cluster within the five major EST6 variants described by traditional electrophoretic procedures as follows: 1, VF; 2, F′; 3–5, F; 6–9, S; 10, U.

Eleven of the 13 *Est6* gene sequences presented here were obtained from lines characterized by Cooke *et al.* (4). One isolate each was sequenced for 9 of the allozymes. Two isolates were sequenced for allozyme 4, and these are designated 4a and 4b. The other two *Est6* sequences have been reported elsewhere. One isolate was a genomic clone from an Indiana strain homozygous for allozyme 8 (8). The other was a cDNA clone from an Oregon-R strain homozygous for allozyme 9 (9). These 2 isolates are designated 8US and 9US.

Genomic clones of the *Est6* gene were extracted from the 11 Australian lines by standard procedures using bacteriophage λgt10 as the vector (10, 11) and the *Est6* cDNA clone as a probe (9). Each *Est6* gene isolated was subcloned in both orientations into phage M13mp19. Both strands of each *Est6* gene were sequenced (12) from 24 base pairs (bp) 5′ of the initiation codon, through the 1387 bp of exon I, the 51-bp intron, the 248 bp of exon II, and 44 bp of untranslated 3′ sequence.

RESULTS

Fig. 1 shows the 52 nucleotide substitutions found among the 13 *Est6* isolates. Of the 36 substitutions that do not lead to amino acid replacements, 1 occurs in the 5′ untranslated region, 26 in exon I, 1 in the intron, 3 in exon II, and 5 in the 3′ untranslated region. The proportion of variant noncoding sites is consistent across these five sequence categories ($\chi^2_4 = 3.96$, $P > 0.05$) and averages 7.5 ± 1.2% overall. Sixteen

Abbreviation: EST6, esterase-6.

‡To whom reprint requests should be addressed at present address: Department of Biology, Jordan Hall, Indiana University, Bloomington, IN 47401.

§The sequences reported in this paper are being deposited in the EMBL/GenBank data base (accession no. J04167).

NUCL. DIFF.	A.A. DIFF.	EST6 ISOLATE 1	2	3	4a	4b	5	10	6	7	8	8US	9	9US
5′ UNTRANSLATED														
A→G -2		·	·	·	■	·	·	·	·	·	·	·	·	NA
EXON I														
C→T 71	T→I 3	·	·	·	■	·	■	·	·	·	·	·	·	·
A→G 198		·	·	■	·	·	·	·	·	·	·	·	·	·
G→T 204		·	·	·	·	·	■	·	·	·	·	·	·	·
T→G 220	S→A 53	·	·	■	·	·	·	·	·	·	·	·	·	·
A→G 318		■	·	■	■	■	·	·	·	■	·	·	·	·
C→T 324		·	·	■	·	·	·	·	·	·	·	·	·	·
C→T 333		·	·	·	■	·	·	·	·	·	·	·	·	·
C→T 378		■	·	·	■	■	·	·	·	·	·	·	·	·
A→T 387		·	·	·	·	·	■	·	·	·	·	·	·	·
T→A 390		·	·	■	■	·	·	·	·	·	·	·	·	·
A→T 511	R→W 150	■	·	·	·	·	·	·	·	·	·	·	·	·
T→C 591		■	·	■	■	■	·	·	·	·	·	·	·	·
T→C 705		■	·	■	■	■	·	·	·	·	·	·	·	·
A→G 772	N→D 237	■	·	■	■	■	■	■	■	·	·	·	·	·
A→G 802	T→A 247	■	·	■	■	■	■	■	■	■	·	·	·	·
C→A 804		·	·	■	·	·	·	·	·	·	·	·	·	·
C→T 849		■	·	■	■	■	·	■	■	■	·	·	·	·
T→C 898	F→L 279	·	·	·	·	·	·	·	■	·	·	·	·	·
T→C 906		·	·	·	■	■	■	·	·	·	·	·	·	·
A→G 909		·	·	·	■	■	■	·	·	·	·	·	·	·
T→C 910		·	·	·	■	■	■	·	·	·	·	·	·	·
T→C 966		·	·	■	·	·	·	■	■	■	·	·	·	·
T→C 990		·	·	·	·	·	■	·	·	·	·	·	·	·
T→C 1067	I→T 335	·	·	·	·	·	■	·	·	·	·	·	·	·
G→T 1089	E→D 342	·	·	·	·	·	■	·	·	·	·	·	·	·
T→C 1111	Y→H 350	·	·	·	·	·	■	·	·	·	·	·	·	·
C→T 1122		·	·	·	·	·	■	·	·	·	·	·	·	·
A→T 1145	K→M 361	·	■	·	·	·	·	·	·	·	·	·	·	·
G→C 1182	E→D 373	·	·	■	·	·	·	·	·	·	·	·	·	·
C→A 1209		·	·	■	·	·	·	·	·	·	·	·	·	·
T→G 1225	L→V 388	·	·	·	·	■	·	·	·	·	·	·	■	·
C→T 1234		·	·	·	·	·	■	·	·	·	·	·	·	·
C→T 1239		·	·	·	·	·	·	·	·	■	·	·	·	·
C→T 1350		·	·	■	·	·	·	·	·	·	·	·	·	·
G→A 1356		·	·	·	·	·	·	·	·	■	·	·	·	·
C→G 1359		·	·	·	·	·	·	·	·	■	·	·	·	·
T→C 1368		·	·	·	·	■	·	·	·	·	·	·	■	·
C→T 1374		·	·	·	·	■	·	·	·	·	·	·	■	·
INTRON I														
T→A 1405		■	·	·	·	■	·	■	■	·	·	·	■	NA
EXON II														
A→G 1446		■	·	·	■	■	·	■	■	■	·	·	■	·
T→C 1452		·	·	·	■	·	·	·	·	·	·	·	·	·
G→A 1526	R→K 471	·	·	·	·	·	·	■	■	·	·	·	·	·
T→G 1573	S→A 487	·	·	·	■	·	·	·	·	■	·	·	■	·
A→G 1607	N→S 498	·	·	·	·	·	·	·	·	·	·	·	·	■
G→A 1620		·	■	·	·	·	■	·	·	·	·	·	·	·
G→A 1655	C→Y 514	·	·	·	·	·	·	■	·	·	·	·	·	·
3′ UNTRANSLATED														
C→A 1691		·	·	·	■	·	■	·	·	·	·	·	■	NA
C→G 1703		·	·	·	·	·	·	■	■	·	·	·	·	NA
A→C 1704		·	·	·	■	·	·	·	·	·	·	·	■	NA
A→C 1706		·	·	·	·	·	■	·	·	·	·	·	·	NA
A→T 1707		·	·	·	■	·	·	·	·	·	·	·	■	NA

FIG. 1. Nucleotide and inferred amino acid differences among the 13 *Est6* isolates. Isolate 8 is taken as the reference and differences are indicated by ■. Nucleotides are numbered from the A of the initiation codon and amino acids from the first residue of the mature protein. One-letter symbols are used for the amino acids. Isolate 9US was a cDNA clone (9), thus NA = not available.

nucleotide substitutions result in amino acid replacements, 12 in exon I and 4 in exon II. The proportion of polymorphic amino acid replacement sites is 1.2 ± 0.3%, which is consistent across the two exons ($\chi^2_1 = 0.18$, $P > 0.05$) and about one-sixth the proportion of polymorphic noncoding sites above ($\chi^2_1 = 49.1$, $P < 0.001$). (Statistical analyses were based on the 11 Australian lines only.)

Fig. 2 shows the locations of the 16 inferred amino acid polymorphisms in the 544-residue primary sequence of the EST6 protein (8). The first 21 NH_2-terminal residues, which constitute a signal peptide (9), were invariant, as were two regions immediately surrounding the aspartic and serine residues at positions 160 and 188. These two residues may participate directly in the catalytic mechanism, and the designated regions surrounding them are highly conserved across several esterases (8, 13).

EST6 is a glycoprotein (14), and one of four potential N-linked glycosylation sites is disrupted by the Ser/Ala polymorphism at 487. This site may not normally be glycosylated because the Ser/Ala polymorphism has no obvious effect on electrophoretic mobility or staining intensity.

EST6 contains six cysteine residues that form disulfide bonds (14). Although not established empirically, the pairings for the cysteine residues in EST6 can be predicted from homologies with the known pairings of the cysteine residues in some cholinesterases (Fig. 2; refs. 8 and 15). One of the 16 EST6 amino acid polymorphisms involves Cys-514. This residue is lost in allozyme 10, and, consistent with the disruption of the disulfide bond, this alloyzme has a smeared electrophoretic phenotype indicative of reduced conformational stability.

Fig. 3 classifies the 16 amino acid polymorphisms for differences in charge, molecular volume, polarity, and hydrophobicity. Three polymorphisms, Phe/Leu at 279, Tyr/His at 350, and Asn/Ser 498, are each conservative for all four physicochemical properties. Six others are each nonconservative for one property, and the remaining 7 are each nonconservative for two.

Fig. 4 shows a hydropathy plot (18) for mature EST6 protein. The broad pattern of the plot is similar for all allozymes. Consideration of both the number and nature of the amino acid polymorphisms suggests that hydrophilic regions likely to be on or near the protein surface are more variable than hydrophobic regions likely to be internal. For example, only two polymorphisms occur in relatively hydrophobic regions: Phe/Leu at 279, which is conservative for all four physicochemical properties above, and Glu/Asp at 342, which is nonconservative for molecular volume only. On the other hand, five polymorphisms, Thr/Ile at 3, Ser/Ala at 53, Tyr/His at 350, Lys/Met at 361, and Glu/Asp at 373, lie in strongly hydrophilic regions, whereas three more, Arg/Trp at 150, Asn/Asp at 237, and Thr/Ala at 247, lie in short, relatively hydrophilic segments within an otherwise hydrophobic zone. Seven of these eight polymorphisms are nonconservative for at least one physicochemical property, and four are nonconservative for two.

The three polymorphisms (at residues 150, 237, and 361) that are nonconservative for charge are all in relatively hydrophilic regions likely to be on or near the protein surface, where they should affect net charge and hence electrophoretic mobility. Moreover, the distribution of these amino acid differences across allozymes (Fig. 3) suggests that they are responsible for distinguishing four of the five major mobility variants (the smeared phenotype of the fifth, U, being due to the loss of Cys-514). The greater anodal mobility of F relative to S can be explained by the polymorphism at 237 for the negatively charged aspartic residue and neutral asparagine residue, respectively. Likewise, the F′/S difference can be attributed to the polymorphism at 361 for the neutral methionine and positive lysine, and the VF/F difference to the polymorphism at 150 for the neutral tryptophan and positive arginine. Although these three polymorphisms all involve nominally similar charge changes, the F′/S mobility difference due to the Met/Lys polymorphism at 361 is about twice as great as the F/S and VF/F mobility differences. This might suggest that residue 361 is more exposed to the surrounding solvent than are residues 237 and 150, and this is indeed indicated by the hydropathy analyses (Fig. 4).

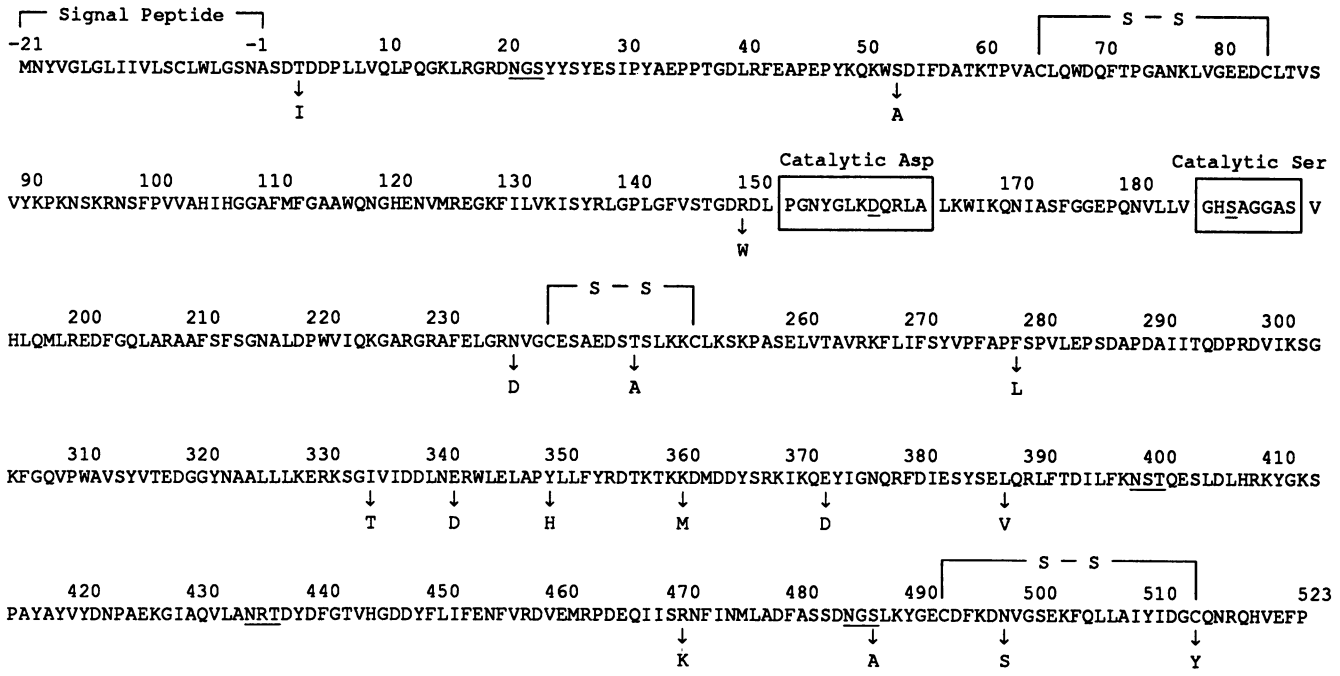

FIG. 2. Inferred amino acid sequence for isolate 8. Sixteen amino acid polymorphisms are shown with arrows. Four potential N-linked glycosylation sites are underlined; and two conserved regions surrounding key catalytic aspartic and serine residues are boxed. The three disulfide bridges (-S—S-) are shown above the sequence.

With the exception of allozyme 6, all the minor mobility variants within each of the major F and S classes have the same haplotype for the three charge-nonconservative amino acid polymorphisms. The minor mobility differences are presumably due to other amino acid polymorphisms that have some small effect on mobility because they alter EST6 conformation. However, it is difficult to determine which specific polymorphisms affect conformation because of the large number that are segregating.

The exceptional allozyme 6 has the same haplotype as the F variants for the three charge-nonconservative polymorphisms but is classified electrophoretically as S (albeit the "fastest" of the S variants). This suggests that one of the charge-conservative replacements characteristic of allozyme 6 causes a conformational change with a relatively large effect on electrophoretic mobility. One explanation for this might be that the altered conformation affects the degree of exposure of another, invariant, charged residue to the protein surface.

Turning to the comparison of different isolates of the same allozymes, we see that isolates 8 and 8US have the same haplotype for all 16 amino acid polymorphisms (and in fact for all 52 nucleotide polymorphisms), despite their derivation from populations on different continents. However, isolates 4a and 4b differ for three amino acid polymorphisms. Thr/Ile at 3, Leu/Val at 388, and Ser/Ala at 487, although they derive

AMINO ACID DIFFERENCE	CHARGE	MOLEC. VOLUME	POL-ARITY	HYDRO-PHOB.	HYDRO-PATHY	EST6 ISOLATE: VF 1	F' 2	F 3	F 4a	F 4b	F 5	U 10	S 6	S 7	S 8	S 8US	S 9	S 9US
Asn→Asp 237	NC	C	C	C	SUR	■	·	■	■	■	■	■	■	·	·	·	·	·
Thr→Ala 247	C	C	NC	C	SUR	■	·	■	■	■	■	■	■	■	·	·	·	·
Thr→Ile 3	C	NC	NC	C	SUR	·	·	·	■	·	■	·	·	·	·	·	·	·
Leu→Val 388	C	NC	C	C	?	·	·	·	·	■	·	·	·	·	·	·	■	·
Ser→Ala 487	C	C	NC	NC	?	·	·	·	■	·	·	·	·	■	·	·	■	·
Ser→Ala 53	C	C	NC	NC	SUR	·	·	■	·	·	·	·	·	·	·	·	·	·
Arg→Trp 150	NC	C	C	NC	SUR	■	·	·	·	·	·	·	·	·	·	·	·	·
Phe→Leu 279	C	C	C	C	INT	·	·	·	·	·	·	·	■	·	·	·	·	·
Ile→Thr 335	C	NC	NC	C	?	·	·	·	·	·	■	·	·	·	·	·	·	·
Glu→Asp 342	C	NC	C	C	INT	·	·	·	·	·	■	·	·	·	·	·	·	·
Tyr→His 350	C	C	C	C	SUR	·	·	·	·	·	■	·	·	·	·	·	·	·
Lys→Met 361	NC	C	NC	C	SUR	·	■	·	·	·	·	·	·	·	·	·	·	·
Glu→Asp 373	C	NC	C	C	SUR	·	·	■	·	·	·	·	·	·	·	·	·	·
Arg→Lys 471	C	C	C	NC	?	·	·	·	·	·	·	■	■	·	·	·	·	·
Asn→Ser 498	C	C	C	C	?	·	·	·	·	·	·	·	·	·	·	·	·	■
Cys→Tyr 514	C	NC	NC	C	?	·	·	·	·	·	·	■	·	·	·	·	·	·
MOBILITY						118	112	108	107	107	106	-	102	101	100	100	99	99
FREQUENCY (%)						4	1	1	—22—		3	1	1	6	56	NA	6	NA

FIG. 3. Physicochemical (16, 17) and electrophoretic and frequency data (4) for the 16 amino acid polymorphisms. C, conservative; NC, nonconservative. Hydropathy predictions (see Fig. 4 and text): SUR, protein surface; INT, interior; ?, inconsistent hydropathy.

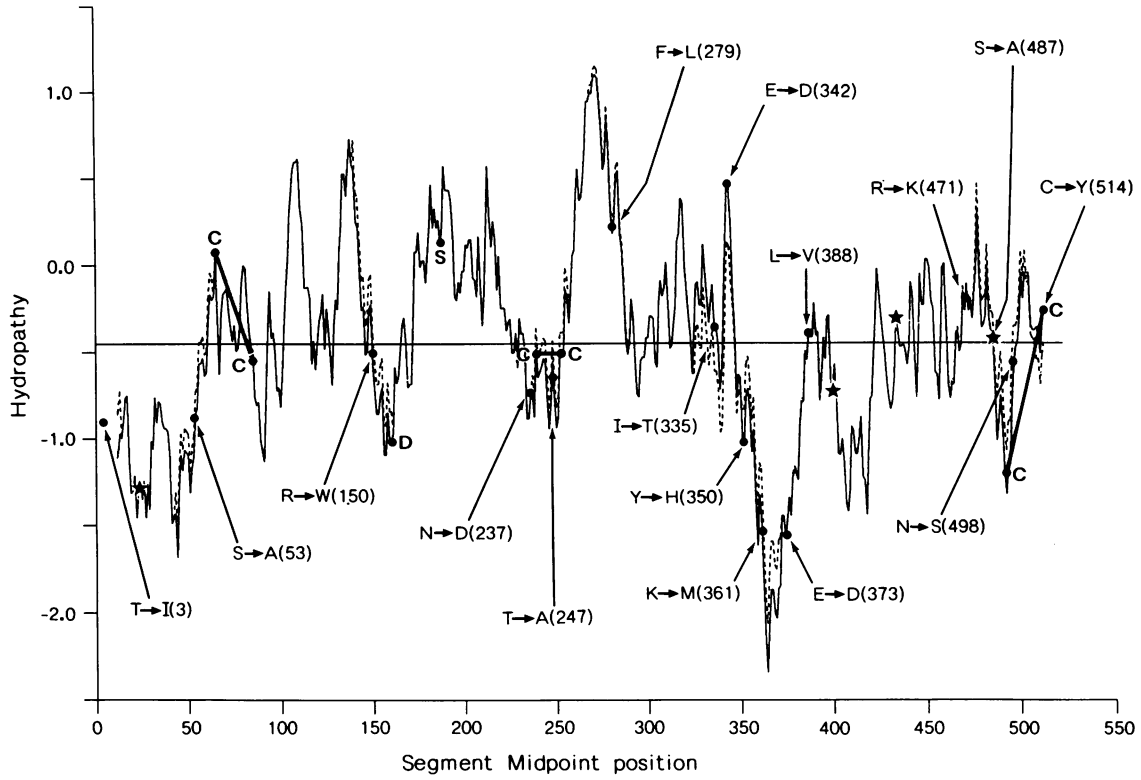

FIG. 4. Hydropathy plot for mature EST6 protein (negative values hydrophilic; segment size = 21). ★, Potential glycosylation sequence; D and S, catalytic aspartic and serine residues, respectively (see Fig. 2). Isolate 8 is shown by a full line, and variations due to the 16 amino acid polymorphisms are shown by broken lines. Disulfide bonds between cysteine (C) residues are shown by heavy lines.

from the same population. Isolates 9 and 9US also differ for three amino acid polymorphisms, the Asn/Ser at 498, as well as the two at 388 and 487, which also segregate between 4a and 4b. The Asn/Ser polymorphism at 498 must have some effect on mobility, since it alone distinguishes 9US from alloyzme 8. Likewise, either or both of the polymorphisms at 388 and 487 must have a similar effect on mobility, since they distinguish isolate 9 from allozyme 8. Our conclusion from the indistinguishable mobilities of isolates 9 and 9US is therefore not that the polymorphisms do not affect mobility, but rather that a particular mobility state can be achieved by different amino acid haplotypes. We further suggest that the same interpretation applies to the two isolates of allozyme 4, since they share two of the same polymorphisms.

DISCUSSION

There is a 6-fold higher proportion of nucleotide polymorphisms in amino acid nonreplacement sites than in replacement sites in the *Est6* gene. This suggests that purifying selection has acted to remove substantial amounts of amino acid polymorphism from EST6 and to conserve its amino acid sequence.

Sixteen amino acid polymorphisms were found among the 10 EST6 allozymes detected by high-resolution electrophoresis, and the average difference between any two allozymes was 4.1 amino acids. These results contrast directly with the findings for alcohol dehydrogenase in *D. melanogaster*, where analyses of a similarly sized data set revealed only 2 amino acid differences among the three polymorphic allozymes and no electrophoretically cryptic amino acid haplotypes at all (19, 20).

In theory, recombination among the 16 amino acid polymorphisms for EST6 could generate up to 2^{16} haplotypes. The actual number will be far fewer because gametic disequilibrium among several polymorphisms is high. Nevertheless, we note that 12 of the 13 isolates sequenced differ in amino acid haplotypes, suggesting that many more would be identified in a larger sample. We conclude that a substantial proportion of EST6 haplotypes will be cryptic, even to high-resolution electrophoretic procedures.

The two most polymorphic amino acid differences in EST6 are Asn/Asp at 237 and Thr/Ala at 247, which are in strong gametic disequilibrium with each other and with the major F/S electrophoretic mobility difference (Fig. 3). The charge difference associated with the Asn/Asp difference at 237 presumably causes the mobility differences between F and S but, given their close proximity and strong disequilibrium with each other, either or both polymorphisms could cause other biochemical differences observed between the two allozymes. Purified EST6-S is more thermostable, has greater catalytic efficiency, and is more susceptible to inhibitors than is purified EST6-F (6). Thermostability differences imply effects on conformational stability, whereas differences for catalytic and inhibition characteristics suggest effects on the active site or substrate-binding cleft of EST6. All three effects are consistent with the proximity of the two polymorphisms to a disulfide bridge between Cys-240 and Cys-252, since the active sites of other serine hydrolases with similar catalytic mechanisms are known to be stabilized by a disulfide bridge (13, 15, 21). The physicochemical basis for these effects could involve the charge difference due to the Asn/Asp substitution or the large polarity difference due to the Thr/Ala substitution (Fig. 3).

On the basis of the evidence above, we propose that one or both of the polymorphisms at 237 or 247 are the primary targets for the selection underlying the F/S latitudinal clines. We cannot dismiss the alternative hypothesis that the two polymorphisms only serve as markers for other linked variants, which are the actual targets for selection. However, none of the other 50 nucleotide polymorphisms described here are in such strong gametic disequilibrium with the F/S

difference. Moreover, the likelihood that the target of selection lies outside the sequenced region is reduced by the finding of similar F and S allozymes showing similar latitudinal clines in *D. simulans* (2). Unless the selected and marker polymorphisms are very closely linked or subject to strong epistatic selection, the disequilibrium between them would have largely decayed since speciation, and parallel clines in F and S frequencies would not now be found in the two species.

Some clues on the evolutionary history of the F/S polymorphism can be obtained from the present data. The overall level of noncoding nucleotide polymorphism is significantly greater among the F isolates than among the S isolates (32 polymorphisms among isolates 3, 4a, 4b, and 5, compared to 14 among isolates 6, 7, 8, and 9; $\chi^2_1 = 7.04$, $P < 0.01$), suggesting that F is ancestral to S. (Phenetic trees based on the nucleotide differences support this contention but are not shown here because they are confounded to an unknown extent by recombination and gene conversion.) Furthermore, within S, the two isolates of allozyme 8 have identical sequences. Yet allozyme 8 was by far the most common minor mobility variant in the population Cooke *et al.* (4) subjected to high-resolution electrophoresis. Taken together, these data suggest that allozyme 8 has both arisen and proliferated relatively recently. Clearly, these temporal changes may interact with the latitudinal clines; however, it will not be clear how until other latitudinally disparate populations of both *D. melanogaster* and *D. simulans* are analyzed for high-resolution electrophoretic variants of EST6 and some representative EST6 allozymes from *D. simulans* are sequenced.

There are three other relatively polymorphic amino acid differences among the EST6 isolates: Thr/Ile at 3, Leu/Val at 388, and Ser/Ala at 487, each of which occurs in more than one relatively common allozyme. Some of the haplotypes generated by these polymorphisms may represent the electrophoretically cryptic thermostability variants known to occur within the major F and S allozymes (4, 5). Direct evidence that the polymorphisms at positions 3, 388, and 487 generate thermostability variation awaits the sequencing of known thermostability variants, since none of the 13 isolates analyzed here have been scored for their thermostability phenotype. However, some effects on conformational stability might be expected for the three polymorphisms, which are relatively nonconservative for physicochemical properties and only one of which (the difference at residue 3) is in a clearly hydrophilic region likely to be on the protein surface.

Of the remaining 11 amino acid replacements, 10 occur in only one isolate each, and the other, Arg/Lys at 471, is confined to two relatively uncommon allozymes, 6 and 10. Some of these less common polymorphisms (Arg/Trp at 150, Lys/Met at 361, and Cys/Thr at 514) are associated with the less common major mobility variants VF, F′, and U, and others (e.g., Asn/Ser at 498) may be associated with minor mobility variants. Still others may be cryptic to electrophoresis. On physicochemical grounds it seems unlikely that 1 of these 11 less common polymorphisms, Cys/Tyr at 514, which disrupts a disulfide bridge, is neutral to natural selection.

However, it is difficult to assess the functional or adaptive significance of the other 10 amino acid replacements, although we note that as a group they are neither more nor less conservative for physicochemical properties than are the 5 more common polymorphisms (Fig. 3). That these 10 variants were detected at all in the present small sample suggests that, while they may be uncommon, they certainly are not rare. On one hand, the case for their selective neutrality is supported by theory showing that the conditions for the maintenance of many alleles by heterozygote advantage involving constant selection coefficients are unrealistically restrictive (22). On the other hand, these conditions are greatly relaxed if the selection coefficients are not constant, and a system like EST6, which affects reproductive behavior, may well be subject to variable selection pressures depending on population and ecological parameters.

We gratefully acknowledge the advice, encouragement, and assistance of Anne Game, Marion Healy, Jill Karotam, Lynne McIntyre, Mark Myers, Rollin Richmond, Robyn Russell, Marlene Saad, and Lis van Papenrecht.

1. Lewontin, R. C. (1985) *Annu. Rev. Genet.* **19,** 81–102.
2. Anderson, P. R. & Oakeshott, J. G. (1984) *Nature (London)* **308,** 729–731.
3. Easteal, S. & Oakeshott, J. G. (1985) *Mol. Biol. Evol.* **2,** 87–91.
4. Cooke, P. H., Richmond, R. C. & Oakeshott, J. G. (1987) *Heredity* **59,** 259–264.
5. Cochrane, B. J. & Richmond, R. C. (1979) *Genetics* **93,** 461–478.
6. White, M. M., Mane, S. D. & Richmond, R. C. (1988) *Mol. Biol. Evol.* **5,** 41–62.
7. Scott, D. (1986) *Evolution* **40,** 1084–1091.
8. Collet, C., Nielsen, K. M., Russell, R. J., Karl, M., Oakeshott, J. G. & Richmond, R. C. (1988) *Mol. Biol. Evol.*, in press.
9. Oakeshott, J. G., Collet, C., Phillis, R. W., Nielsen, K. M., Russell, R. J., Chambers, G. K., Ross, V. & Richmond, R. C. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 3359–3363.
10. Coen, E. S., Thoday, J. M. & Dover, G. (1982) *Nature (London)* **295,** 564–568.
11. Huynh, T. V., Young, R. A. & Davis, R. W. (1984) in *DNA Cloning Techniques: A Practical Approach*, ed. Glover, D. (IRL, Oxford), Vol. 1, pp. 49–78.
12. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
13. Myers, M., Richmond, R. C. & Oakeshott, J. G. (1988) *Mol. Biol. Evol.* **5,** 113–119.
14. Mane, S. D., Tepper, C. S. & Richmond, R. C. (1983) *Biochem. Genet.* **21,** 1019–1040.
15. MacPhee-Quigley, K. T., Vedvick, T., Taylor, P. & Taylor, S. S. (1986) *J. Biol. Chem.* **260,** 13565–13570.
16. Larsen, P. O. (1980) in *The Biochemistry of Plants*, ed. Miflin, B. J. (Academic, New York), Part 2, p. 235.
17. Taylor, W. R. (1986) *J. Theor. Biol.* **119,** 205–218.
18. Kyte, J. & Doolittle, R. R. (1982) *J. Mol. Biol.* **157,** 105–132.
19. Kreitman, M. (1983) *Nature (London)* **304,** 412–417.
20. Collet, C. (1988) *J. Mol. Evol.* **27,** 142–146.
21. Young, C. L., Barker, W. C., Tomaselli, C. M. & Dayhoff, M. O. (1978) in *Atlas of Protein Sequence and Structure*, ed. Dayhoff, M. O. (Natl. Biomed. Res. Found., Washington, DC), Vol. 5, Suppl. 3, pp. 73–93.
22. Lewontin, R. C., Ginzburg, L. R. & Tuljapurkar, S. D. (1978) *Genetics* **88,** 149–170.