Is Esterase-P Encoded by a Cryptic Pseudogene In *Drosophila melanogaster*?

Evgeniy S. Balakirev*,† and Francisco J. Ayala*

**Department of Ecology and Evolutionary Biology, University of California, Irvine, California 92697-2525 and †Institute of Marine Biology, Vladivostok 690041, Russia*

Manuscript received June 6, 1996
Accepted for publication August 21, 1996

ABSTRACT

We have amplified and sequenced the gene encoding Esterase-P (*Est-P*) in 10 strains of *Drosophila melanogaster*. Three premature termination codons occur in the coding region of the gene in two strains. This observation, together with other indirect evidence, leads us to propose that *Est-P* may be a pseudogene in *D. melanogaster*. *Est-P* would be a "cryptic" pseudogene, in the sense that it retains intact the coding sequence (without stop codons and other alterations usually observed in pseudogenes) in most *D. melanogaster* strains. We conjecture that the β-esterase cluster may consist in other Drosophila species of functional and nonfunctional genes. We also conjecture that the rarity of detected pseudogenes in Drosophila may be due to the difficulty of discovering them, because most of them are cryptic.

EUKARYOTIC genomes contain, in addition to functional genes, sequences that are related to particular genes but exhibit changes that inactivate their expression. Such sequences have been designated "pseudogenes" (reviewed by VANIN 1985; WEINER *et al.* 1986; WILDE 1986). Since they were first described (JACQ *et al.* 1977; FEDOROFF and BROWN 1978), pseudogenes have been identified in a variety of gene families. Some are unprocessed pseudogenes that have introns and are usually chromosomally linked to their functional counterparts, suggesting that they arise by tandem gene duplication, followed by divergence and inactivation. Processed pseudogenes lack introns and are believed to arise from reverse transcription of processed mRNA, followed by integration of the resulting DNA into a new chromosomal location in the genome. Most pseudogenes are presumably not expressed, although several are transcribed into RNA molecules (FISCHER and MANIATIS 1985; SORGE *et al.* 1990; AUBERT *et al.* 1992; CURRIE and SULLIVAN 1994; FÜRBAß and VANSELOW 1995), and processed pseudogenes that are translated into protein molecules have been discovered (MCCARREY and THOMAS 1987).

Pseudogenes along with their functional counterparts provide useful models for understanding gene evolutionary processes, since they allow the comparison of mode and rate of molecular changes in sequences with and without functional constrains (LI *et al.* 1981, 1985b; MIYATA and YASUNAGA 1981; MIYATA and HAYASHIDA 1981; GOJOBORI *et al.* 1982; LI 1982, 1983).

Only a few examples of pseudogenes are known in Drosophila. A transcriptionally active (unprocessed) alcohol dehydrogenase (*Adh*) pseudogene, closely linked to one or two functional counterparts, has been described in *D. mulleri* (FISCHER and MANIATIS 1985), *D. mojavensis* (ATKINSON *et al.* 1988), *D. mettleri* (YUM *et al.* 1991), *D. hydei* (MENOTTI-RAYMOND *et al.* 1991), *D. buzzatii* (SCHAFER 1992), *D. peninsularis*, and *D. mercatorum* (SULLIVAN *et al.* 1994). In all these species, the reading frames of the *Adh* pseudogenes contain premature termination codons, resulting from one or more frameshifts, which along with indels prevent productive translation. KYLSTEN *et al.* (1990) have found two unprocessed pseudogenes, included in the cecropin locus of *D. melanogaster*. Processed pseudogenes of alcohol dehydrogenase have been found in *D. yakuba* and *D. teissieri* (JEFFS and ASHBURNER 1991; JEFFS *et al.* 1994) and phosphoglyceromytase (*Pglum*) in *D. melanogaster* (CURRIE and SULLIVAN 1994). The *Adh* and *Pglum* processed pseudogenes lack introns and are located on different chromosomes (*Adh*) or different arms of chromosome (*Pglum*) from their functional counterparts.

COLLET *et al.* (1990) have detected a tandem duplication of the *Est-6* gene in *D. melanogaster*, denoted as *Est-P*. Both genes belong to the β-esterase gene cluster (KOROCHKIN *et al.* 1987) and are located within subsection 69 A1-A3 on the left arm of polytene chromosome *3* (PROCUNIER *et al.* 1991), with the same 5′ to 3′ orientation (*Est-P* is downstream from *Est-6*). Both genes comprise two exons, with length of 1387 and 248 bp, separated by a short intron, which is 51 and 56 bp long in *Est-6* and *Est-P*, respectively. The coding regions of the genes are separated by only 197 bp. The two genes show 64 and 60% similarity in the DNA and protein sequence, respectively (COLLET *et al.* 1990).

In Drosophila species, the β-esterase gene cluster includes two or (only in *D. pseudoobscura*) three known closely linked genes (OAKESHOTT *et al.* 1993, 1995), which have the same direction of transcription and similar exon/intron structure (YENIKOLOPOV *et al.* 1983,

Corresponding author: Francisco J. Ayala, Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697-2525. E-mail: fjayala@uci.edu

1989a; BRADY *et al.* 1990; COLLET *et al.* 1990; EAST *et al.* 1990). The distance between the coding sequences of different members of the gene cluster ranges from 197 bp (between *Est-6* and *Est-P*) in *D. melanogaster* (COLLET *et al.* 1990) and 740 bp (between *Est-A* and *Est-B*) in *D. buzzatii* (EAST *et al.* 1990), to at least 1000 bp (between *Est-S* and *Est-X*) in *D. virilis* (YENIKOLOPOV *et al.* 1989b) and ~1700 bp (between *Est-5A* and *Est-5B*) in *D. pseudoobscura* (BRADY *et al.* 1990).

There is an extensive literature devoted to different aspects of the molecular population and evolutionary genetics of *D. melanogaster Est-6* (reviewed by OAKESHOTT *et al.* 1993, 1995). EST-6 plays an important role in the reproductive biology of *D. melanogaster.* A high level of the enzyme has been detected in the ejaculatory duct of adult males (ARONSHTAM and KUZIN 1974; KOROCHKIN *et al.* 1974; KUZIN *et al.* 1975; SHEEHAN *et al.* 1979; RICHMOND *et al.* 1980; USPENSKY *et al.* 1988; KOROCHKIN *et al.* 1990). RICHMOND *et al.* (1980) have shown that EST-6 is transferred to females as a component of the seminal fluid during copulation. The adaptive significance of *Est-6* allozyme variation has been demonstrated (see reviews in RICHMOND *et al.* 1990; OAKESHOTT *et al.* 1993, 1995).

Significantly less information is available about *Est-P.* COLLET *et al.* (1990) have detected *Est-P* transcripts in *D. melanogaster* late larvae and adults of each sex, whereas the *Est-6* transcripts are found mainly in adult males. They have, accordingly, proposed distinct physiological functions for the products of the two genes.

We have amplified and sequenced *Est-P* in 10 strains of *D. melanogaster* and found three premature termination codons in two of the 10 strains, which should prevent productive translation of the gene in those strains. We propose that *Est-P* may be a pseudogene in *D. melanogaster,* which is not essential although it retains a potentially functional coding sequence in a majority of the *D. melanogaster* strains we have studied.

MATERIALS AND METHODS

Drosophila strains: The *D. melanogaster* strains were derived from wild flies collected by F. J. AYALA (October 1991) in El Rio Vineyard, Lockeford, California. The strains were made fully homozygous for the third chromosome by means of crosses with balancer stocks, as described by SEAGER and AYALA (1982). The strains were named in accordance with the superoxide dismutase electrophoretic alleles they carry, Fast (F) or Slow (S), as follows: 255S, 510S, 521S, 94F, 174F, 357F, 377F, 521F, 581F, 968F.

PCR primers, amplification, and sequencing: Total genomic DNA was extracted using the procedure described by PALUMBI *et al.* (1991). The *Est-6* and *Est-P* sequences, previously published by COLLET *et al.* (1990), were used for designing PCR and sequence primers. The amplified fragments (2.5 kb long) encompassed part of *Est-6* exon II, the intergenic region between *Est-6* and *Est-P,* the whole *Est-P* gene, and 360 bp of the 3′-untranslated region. The two primers used for the PCR amplification reactions were as follows: 5′-gataatgtaggtagtgagaaattc-3′ (forward primer), and 5′-ctatgtgactttgtggacgattgtgt-3′ (reverse primer).

All PCR reactions were carried out, as described by KWIATOWSKI *et al.* (1991), in final volumes of 100 μl containing 40 μM each of dNTP, 2.5 U AmpliTaq DNA Polymerase (Perkin-Elmer), 0.2 μM each of forward and reverse primers, buffer (Perkin-Elmer) at a final concentration of 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, and ~50 ng of template (total genomic) DNA. The mixtures were overlaid with mineral oil, placed in a DNA thermal cycler (Perkin Elmer Cetus), incubated 5 min at 95° and subjected to multiple cycles of denaturation, annealing, and extension under the following conditions: 95° for 1.0 min, 56° for 1.0 min, and 72° for 2.0 min (for the first cycle and progressively adding 3 sec at 72° for every subsequent cycle). After 30 cycles, a final 7-min extension period at 72° concluded each amplification reaction. Samples were stored at 4° for several hours or at −20° for up to 2 weeks.

One-tenth of each reaction volume was assayed on a 0.8% agarose gel. If the desired PCR product was detected, the remainder of the reaction was purified with Wizard PCR preps DNA purification system (Promega corporation). The purified PCR product was directly cloned using the TA cloning kit (Invitrogen, San Diego, CA). Cloned DNA was purified using the QIAprep plasmid preparation system (QIAGEN Inc., CA).

DNA sequencing was done by the dideoxy chain-termination technique (SANGER *et al.* 1977) with Sequenase Version 2.0 T7 DNA polymerase (Amersham Life Science Inc., USA) using ^{32}P-labeled dATP. Five internal primers (a–e) were used for sequencing (see Figure 1): (a) 5′-tacagctacgaatcgattcc-3′, (b) 5′-ggaaactatggtctaaaggat-3′, (c) 5′-tgttgtggagccgtcagatgcac-3′, (d) 5′-tagatcttcaccgaaagtatg-3′, (e) 5′-ttgtaaaaacgaggaatatgc-3′.

DNA sequence analysis: All primers were designed using the computer program DNASIS for Windows (1994, Hitachi Software Engineering Co., Ltd.). This program allowed us also to check the secondary structure of primers. Multiple alignment was carried out manually, using the program DARWIN (elaborated by Dr. ROBERT TYLER from our laboratory) and automatically, using the program CLUSTAL W (THOMPSON *et al.* 1994). Additional DNA sequences of the Drosophila β-esterase genes were obtained from the GenBank database, with the accession numbers *D. melanogaster, Est-6* and *Est-P,* M33780, M33781; *D. mauritiana, Est-6,* L10671; *D. simulans, Est-6,* L34263; *D. pseudoobscura, Est-5C* and *Est-5B,* M55907; *Est-5A,* M55908; *D. virilis, Est-S,* X70351. The *D. buzzatii Est-A* and *Est-B* sequences were obtained from EAST *et al.* (1990). The phylogenetic analysis of the esterase sequences was performed with the computer program MEGA (KUMAR *et al.* 1993). The pattern of synonymous and nonsynonymous substitutions was estimated by the method of LI (1993; see LI *et al.* 1985a), but synonymous sites were counted according to SATTA (1993).

RESULTS

Figure 1 shows the organization of the duplicated β-esterase genes in *D. melanogaster* and our sequencing strategy. The polymorphic nucleotide sites in the *Est-P* coding sequence of the 10 strains are shown in Figure 2; the polymorphic amino acid sites are shown in Figure 3. We detected three premature termination codons and one short insertion in two of the 10 *D. melanogaster* strains. One premature termination codon (TGA) occurs in the 357F strain and two tandemly located termination codons (both TGA) in the 510S strain, where they are associated with a 9-bp-long insertion. We have confirmed the presence of the stop codons in second clones, obtained from separate PCR amplifications, in-

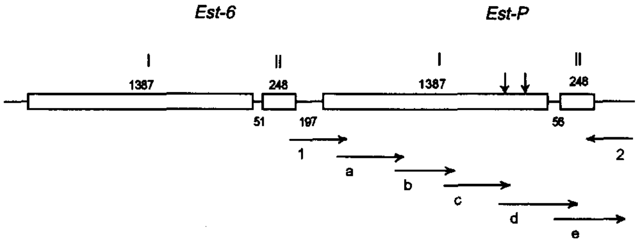

FIGURE 1.—Schematic representation of the *D. melanogaster* β-esterase gene cluster, which includes two genes: *Est-6* and *Est-P*. Coding sequences, □ marked by Roman numerals (I, II). The 5′- and 3′-untranslated regions, introns, and intergenic region are represented by thin lines. Lengths (in bp) are given above the figure for exons, and below the figure for introns and the intergenic region. The two numbered horizontal arrows (1 and 2) indicate the location and direction of the PCR amplification primers. The sequencing primers are indicated by thin arrows with lowercase letters. The small vertical arrowheads signal the position of the premature stop codons.

```
                                                     111111111111111111111111111111111
                1112233344557778899999999000023333333333333333445555555566
            13680453807856021391723345555166633677788888888889550123357712
           825689883742789489853343720127167899547901234567890012656896874
Est-P ref. TCTTAATAATTGCTCAATAAGTCAGCCAGTCTGGTGCAT---------TACATTCAACAAGAG
 255S      .........C.......C..................................C..........
 510S      ..............T........................ACATTTGATG...C..........
 521S      ......A......C.....G................................C..........
  94F      ........G.........G.................................C..........
 174F      .....C..G......................................................
 357F      CA.AT..C...TT...T...ACTTCTAGTCT..ACAGCC...........AG..TGGGG.AGA
 377F      ...............................................................
 521F      ..C...A..............................................C.....G...
 581F      ...............................................................
 968F      ..........C....G...............................................
```

FIGURE 2.—Polymorphic nucleotide sites in the *Est-P* sequence of 10 *D. melanogaster* strains. The nucleotides are numbered (above the sequence) from the translation initiation codon. The intergenic region, intron, and 3′-untranslated region are not included. Dots indicate identical nucleotides to those in the reference sequence. Three premature stop codons are shown (even though not all sites involved are polymorphic) underlined: strain 357F at 1066–1068 and strain 510S at 1384–1386 and 1387–1389F. There is a 9-bp insertion sequence in 510S (1379–1387); these sites are represented by short dashes in the reference sequence. *Est-P* reference is from COLLET *et al.* (1990).

dicating that these alterations are not a cloning artefact. In both strains, the termination codons are located within exon I of *Est-P* (see Figures 1 and 2). The inferred polypeptide sequences are significantly shorter than the complete one (355 and 461 amino acids, instead of the 544 amino acids in the complete sequence). The missing polypeptide segments include two cysteine residues (positions 517 and 538), the catalytic histidine (position 464), and one (in the 510S strain), or three (in the 357F strain) potential N-linked glycosylation sites (positions 420, 456, 509), which are important for maintaining esterase function.

Figure 4 shows a phylogenetic tree of β-esterase coding sequences in Drosophila. In *D. pseudoobscura* and *D. buzzatii* the β-esterase genes cluster together, suggesting that they may be evolving in accordance with the concerted pattern, which is common for duplicated genes (ARNHEIM 1983; HICKEY *et al.* 1991; SHIBATA and YAMAZAKI 1995), or that the gene duplications are not very ancient. The two β-esterase genes from *D. melanogaster* show, on the contrary, much greater divergence.

The rate of synonymous substitutions is likely to be the same in the products of gene duplication, whether functional or not. The rate of nonsynonymous substitutions, however, is likely to be greater in a pseudogene than in a functional gene. Table 1 compares synonymous as well as nonsynonymous substitutions between the *Est-6* genes of *D. melanogaster, D. simulans* and *D. mauritiana,* as well as between the *D. melanogaster Est-P* and the *Est-6* gene of all three species. The relative rates of synonymous (K_s) and nonsynonymous (K_a) substitutions are examined by means of the γ parameter (SATTA 1993), which is simply the ratio K_a/K_s. This ratio is about twice as large between *Est-P* and *Est-6* genes, as between the *Est-6* genes ($t = 4.59$, $P < 0.01$), indicating that the rate of nonsynonymous substitutions is greater in *Est-P,* as expected if it were a pseudogene.

Table 2 gives the γ parameter for the 10 *Est-P* alleles as well as for the *Est-6* alleles of the same 10 strains (unpublished data of E. S. BALAKIREV and F. J. AYALA). Estimates of γ are potentially subject to large error whenever the number of substitutions between two alleles is smaller than 10 (including both synonymous and nonsynonymous sites) (SATTA 1993) as it is the case for these *Est* alleles. We have overcome this problem by counting the total numbers of substitutions between all 10 alleles of each gene, and calculating K_a, K_s and γ for the total values. Table 2 shows that γ is greater for *Est-P* than for *Est-6,* reflecting that the rate of nonsynonymous substitutions is faster in *Est-P* (K_a is about double). This is what is expected if *Est-P* is a pseudogene.

Table 3 gives the γ values for pairwise comparisons between three esterase genes of *D. pseudoobscura.* The data suggest that *Est-5A* is a pseudogene. If we combine the data of Tables 1–3, the value of γ is significantly greater for the comparisons involving the presumptive pseudogenes *Est-P* or *Est-5A* ($t = 5.5$, $P < 0.005$).

DISCUSSION

COLLET *et al.* (1990) have argued that *Est-P* is a functional gene in *D. melanogaster,* based on several lines of evidence: transcriptional activity, no premature termination codons, intact splicing sites, presence of initiation and termination codons, significantly less replacement-site than silent-site differences in the coding regions of both *Est-6* and *Est-P.* Our results suggest rather that *Est-P* is probably not a functional gene, at least not an essential one.

Two of the 10 *D. melanogaster* strains studied have premature termination codons. These strains are fully homozygous for the third chromosome, yet viable, although the termination codons will prevent productive translation. Translation of the two deficient genes would yield polypeptides that lack essential amino acids and are significantly shorter than the complete sequences (355 and 461 amino acids instead of 544 amino

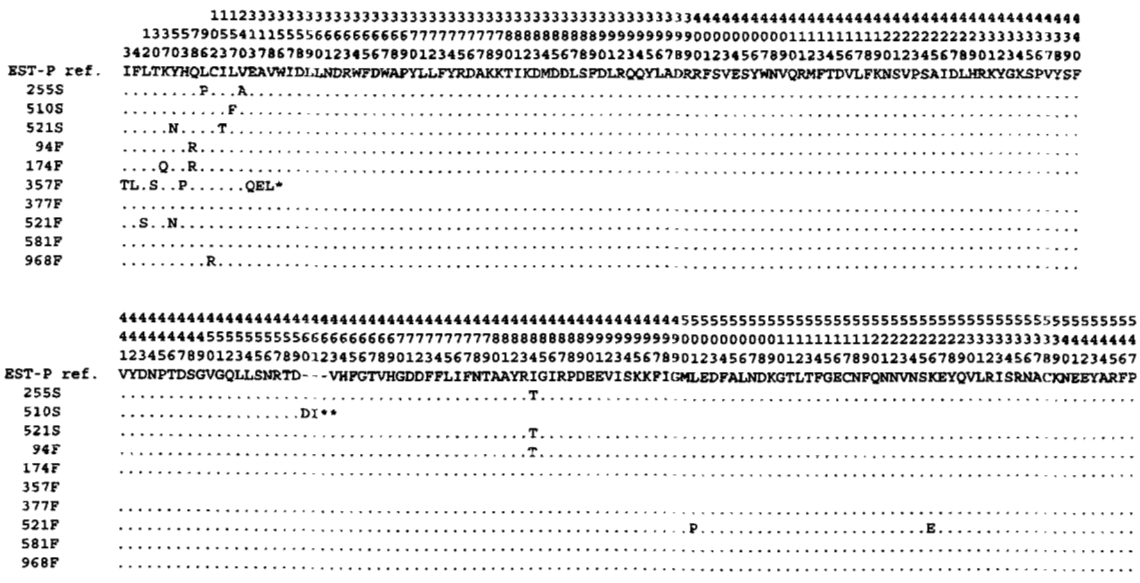

FIGURE 3.—Polymorphic sites in the inferred EST-P amino acid sequence of 10 strains of *D. melanogaster*. The three stop codons are indicated by * (356 in 357F; 462 and 463 in 510S). Other conventions as in Figure 2.

acids in the complete sequence). As noted above, the deleted polypeptide fragments include two cysteine residues, the catalytic histidine, and one (for 510S strain), or three (for 357F strain) potential N-linked glycosylation sites, all of which are important for maintaining esterase function. The catalytic histidine and the cysteine sites are strictly conserved in all Drosophila's β-esterase sequences already published (data not shown). The catalytic histidine is one of three key residues in the esterase molecule that directly effect the esterase reaction (CARTER and WELLS 1988; MYERS *et al.* 1988; SCHRAG *et al.* 1991; SUSSMAN *et al.* 1991). The cysteine residues are involved in the formation of disulfide bridges, which support the three-dimensional molecular structure, and are crucial for molecular conformational stability. It is known that the active sites of other serine hydrolases with similar catalytic mechanisms are stabilized by disulfide bridges (MYERS *et al.* 1988). *Est-6* is a glycoprotein (MANE *et al.* 1983) and contains four potential N-linked glycosylation sites, three (strain 357F) or two (strain 510S) of which are lost in *Est-P* sequences. We can assume that protein molecules lacking those sites are not functionally efficient.

The fact that two perfectly viable strains lack functional *Est-P* genes suggests that *Est-P* is not an essential gene and may be a pseudogene. There are also indirect sources of evidence that support the hypothesis that *Est-P* is a pseudogene.

1. The *Est-6* promoter has four regulatory elements and the most important one, which controls gene expression in the ejaculatory duct, is located in the region from −511 to −1132 bp, upstream from the initiation codon (LUDWIG *et al.* 1993). The intergenic region between *Est-6* and *Est-P* is only 197 bp, and thus there is not enough sequence to account for an *Est-P* promoter, if we extrapolate from *Est-6* that the *Est-P* native promoter should be longer than 197 bp. Some significant part of the promoter was lost during or after duplication. This event destroyed normal gene regulation in one of the duplicated products and determined which

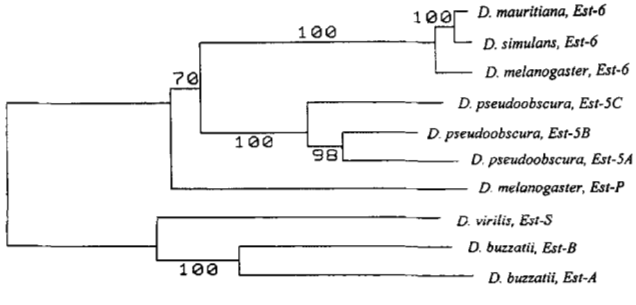

FIGURE 4.—Phylogenetic relationships of β-esterase genes in Drosophila. The sequences of *Est-A* and *Est-B* of *D. buzzatii* are from EAST *et al.* (1990). Other sequences are from GenBank (see MATERIALS AND METHODS for the accession numbers). The unrooted phylogenetic tree has been obtained by the neighbor-joining method (SAITOU and NEI 1987) using Jukes-Cantor distances for exon I. The numbers are bootstrap probability values based on 1000 replications.

TABLE 1

The rate of synonymous and nonsynonymous substitutions, and the γ parameter for intra- and interspecific comparisons between Est-P and Est-6

Comparison	K_a	K_s	γ
mel Est-P vs. mel Est-6	0.241	0.572	0.422
mel Est-P vs. sim Est-6	0.237	0.576	0.411
mel Est-P vs. mau Est-6	0.234	0.580	0.403
mel Est-6 vs. sim Est-6	0.022	0.109	0.203
mel Est-6 vs. mau Est-6	0.025	0.097	0.259
sim Est-6 vs. mau Est-6	0.009	0.041	0.212

The species abbreviations are as follows: *mel, D. melanogaster; sim, D. simulans; mau, D. mauritiana.* K_s, rate of synonymous; K_a, rate of nonsynonymous; rate of $\gamma = K_a/K_s$.

TABLE 2

Total number of synonymous and nonsynonymous substitutions among 10 alleles of *Est-P* and *Est-6* in *D. melanogaster*

Gene	N_a	N_s	K_a	K_s	γ
Est-P	34	22	0.029	0.047	0.624
Est-6	17	28	0.015	0.060	0.250

K_a, K_s and γ as in Table 1. N_s, number of synonymous; N_a, number of nonsynonymous.

TABLE 3

The rate of synonymous and nonsynonymous substitutions, and the γ parameter between three genes of *D. pseudoobscura*

Comparison	K_a	K_s	γ
Est-5A:Est-5B	0.133	0.304	0.438
Est-5A:Est-5C	0.161	0.397	0.406
Est-5B:Est-5C	0.092	0.311	0.288

Abbreviations as in Table 1.

of the two *Est-6* or *Est-P* remained an essential gene. Termination codons may then have appeared later by chance, since there were no functional constrains that would strongly select against them. Pseudogenes have been detected that have no obvious defects (such as frame-shift mutations and/or premature termination codons) and retain high sequence similarity to their functional counterparts, but lack promoter elements. The deletion of promoter parts completely inactivates the expression of these genes (SIBBALD and BLENCOWE 1990; MATTERS and GOODENOUGH 1992). A possible escape of the argument we have just advanced is to assume that a segment of the *Est-6* gene may have been highjacked into becoming the promoter of *Est-P* (see COLLET *et al.* 1990). This seems unlikely but is not, of course, impossible.

2. Members of a multigene family often share an evolutionary process called concerted or coincidental evolution (ARNHEIM 1983). Several molecular mechanisms and theoretical models have been proposed to explain patterns of concerted evolution (NAGYLAKI 1984; OHTA 1984). Both nonreciprocal (gene conversion) and reciprocal (unequal sister chromatid exchange) recombination of DNA among members of gene families contribute toward their concerted evolution, so that the members of the family retain much greater similarity than expected in terms of the time since their independent origin. The phylogenetic tree of the Drosophila β-esterase genes suggest that a gene duplication may have occurred previous to the divergence of the subgenera Sophophora (which includes *D. melanogaster* and *D. pseudoobscura*) and Drosophila (*D. buzzatii*). If so, concerted evolution might account for the greater similarity between the two (or three) paralogous genes of each species than between the ortologous genes from different species. This is the situation observed in *D. pseudoobscura* and *D. buzzatii*, whereas the divergence between the two paralogous genes is much greater in *D. melanogaster* (Figure 4). This greater divergence could, however, be accounted for if *Est-P* would have acquired a new function. We submit, however, based on the overall of evidence, that it is rather the loss of function of *Est-P* that accounts for its extensive divergence from *Est-6*.

3. COLLET *et al.* (1990) detected *Est-P* transcripts in late larvae and adults of each sex. However HEALY *et al.* (1991) failed to find the *Est-P* protein product among 22 esterases in *D. melanogaster*, assayed by polyacrylamide gel electrophoresis and isoelectric focusing. As we have mentioned above, some pseudogenes can be transcriptionally active. Thus, the presence of *Est-P* transcripts is not definitive proof of functionality, while the absence of the protein product in a functional organism is indirect evidence of lack of (at least, essential) function.

4. In functional genes most nonsynonymous substitutions are deleterious and are eliminated from a population by negative selection. As a consequence, in functional genes the rate of nonsynonymous substitutions is generally significantly lower than the rate of synonymous substitutions. Pseudogenes are not subject to functional constraints, and thus synonymous and nonsynonymous substitutions rates are expected to be the same, and the expected ratio of the rates, $\gamma = K_a/K_s$, should gradually become one (KIMURA 1983). However, the γ values of a pseudogene and a paralogous functional gene may be fairly similar, if the duplication occurred recently, or if the pseudogene became nonfunctional recently. The possibility of gene conversion between the pseudogene and its functional counterpart also has to be taken into account, since it will contribute to conserve the similarity of the γ values.

Be that as it may, if we observe that γ is greater between a putative pseudogene and its functional counterpart than between two corresponding functional genes, this will be evidence favoring the interpretation that the putative pseudogene is indeed a pseudogene. This will be the case whether the comparison is made between species or between alleles of the same species. Both tests, as summarized in Tables 1 and 2, favor the hypothesis that *Est-P* is a pseudogene.

If we assume that the rate of nonsynonymous substitutions is constant, the K_s values in Table 1 conform to the accepted phylogeny of the *melanogaster* group species, in which *D. simulans* and *D. mauritiana* are sister species, and *D. melanogaster* is sister to the cluster of the other two (*e.g.*, POWELL and DESALLE 1995, Figure 4). The average K_s for the three *Est-P:Est-6* comparisons in Table 1 is 0.576, or 5.6 times the average K_s between *D. melanogaster* and the other two species. If we assume that the *melanogaster* and (*simulans* + *mauritiana*) lineages diverged 2.5 million years ago (POWELL and DESALLE

1995, Table IV), the *Est-P:Est-6* duplication would have occurred 14 million years ago. This inference must, however, be received with considerable caution for reasons mentioned above, particularly the possibility of gene conversion between *Est-P* and *Est-6*. If we accept the duplication date of 14 million years ago, it follows that the phylogeny based on nucleotide distances (Figures 4) is in error (introduced by different rates of nucleotide evolution in different genes), since the divergence between the *obscura* and *melanogaster* lineages occurred about 40 million years ago (see, *e.g.*, Table IV in POWELL and DESALLE 1995).

An *Est-6* null allele has been found in the carnation eye color (*car*, *X*-linked, recessive) laboratory stock of *D. melanogaster* [sc ec ev et^6 v g^2 f/FM_3 y^{31d} sc^8 dm B 1] (JOHNSON *et al.* 1966). If null alleles were common for *Est-6*, we might propose that this gene also is a pseudogene. However no *Est-6* null alleles have been observed in natural populations (VOELKER *et al.* 1980; LANGLEY *et al.* 1981; OAKESHOTT *et al.* 1989). Moreover, GILBERT and RICHMOND (1982) have shown that *Est-6* null-allele males have significantly lower reproductive fitness than typical (wild-type) *Est-6* males. SAAD *et al.* (1994) have, moreover, found strong association between the activity level of EST-6 and the reproductive fitness in *D. melanogaster* wild-derived lines. These results (see extensive discussion in OAKESHOTT *et al.* 1993, 1995) strongly suggest that *Est-6* in *D. melanogaster* is engaged in adaptively significant physiological function and therefore cannot be a pseudogene. The null allele observed in a laboratory stock (JOHNSON *et al.* 1966) may represent a spontaneous null mutation that would have very low probability to survive in nature. It may be likely that this spontaneous null mutation arose as a consequence of hybrid dysgenesis (SVED 1979; PRUDHOMMEAU and PROUST 1990; GEORGIEV and YELAGIN 1992), which increases mutation rates in specific crosses that involve strains long cultured in the laboratory. In addition to the *Est-6* null allele, JOHNSON *et al.* (1966) found another null allele for a separate esterase locus, *Est-C* in the same *D. melanogaster* stock. An abundance of null alleles certainly may occur in hybrid dysgenesis crosses. Because of the crosses used in our study to generate homozygous strains, the question arises whether our *Est-P* "null" alleles may have arisen by hybrid dysgenesis. This hypothesis seems unlikely, given that no other null alleles have been found in the numerous El Rio strains of *D. melanogaster* assayed in our laboratory for *Sod, Xdh, Gpdh* and other genes, even though all strains were made homozygous by the same extraction and maintenance procedures (HUDSON *et al.* 1994; F. J. AYALA and collaborators, unpublished results).

The original loss of function of *Est-P* in *D. melanogaster* may have happened owing to the deletion of part of the promoter region when the *Est-6* duplication occurred. After that, due to lack of functional constrains, null mutations, leading to stop codons would be tolerated and persist in some individuals. A relative abundance of premature stop codons is likely to arise only as a result of frameshift mutation. Nevertheless, stop codons are likely to be rare even in pseudogenes or nonessential genes. First, we note that the rate at which a null mutation becomes fixed in a population is highly dependent on population size and mutation rate (LI 1980, 1983). In a very large population, a null mutation may never become fixed (FISHER 1935; NEI and ROYCHOUDHURY 1973; LI 1980, 1983). Moreover, MODIANO *et al.* (1981) have shown that the pattern of codon usage in human globin genes is such that codons that can mutate by single nucleotide substitution to a termination codon are never used whenever the corresponding amino acid is specified also by other triplets. Thus, the number of opportunities to mutate to an untranslatable codon is reduced to the minimum compatible with the amino acid composition of a gene. They propose that living organisms have adopted mechanisms that considerably reduce mutations with especially severe consequences (MODIANO *et al.* 1981). The combination of these two factors (*i.e.*, fixation rate of null mutations and pattern of codon usage) may explain why the *Est-P* gene lacks premature stop codons in most individuals, despite the fact that the *Est-6/Est-P* duplication event is probably as old as the divergence of the *repleta* and *melanogaster* group (COLLET *et al.* 1990), which has been estimated to have occurred 44 million years ago by some authors (KWIATOWSKI *et al.* 1994), but 60–80 million years ago by others (*e.g.*, THROCKMORTON 1975). We may consider *Est-P* as a "cryptic" pseudogene, because only two strains out of 10 studied strains exhibit stop codons. In all other strains the coding sequence of *Est-P* is intact, without premature stop codons or other alterations typically observed in pseudogenes.

In spite of extensive molecular genetic studies, few pseudogenes have been found in Drosophila (see Introduction). This might be because many of them are cryptic, for similar reasons as we have just pointed out for *Est-P* in *D. melanogaster*. Particularly, the β-esterase gene cluster in other Drosophila species may also consist of functional and nonfunctional genes. The results displayed in Table 3, for pairwise comparisons between the β-esterase genes of *D. pseudoobscura*, suggest that *Est-5A* may be a pseudogene. Negative evidence supporting this conjecture is that translational products have not been identified for this gene (BRADY *et al.* 1990). The similarity of K_s values in Table 3 suggests that the two gene duplications giving rise to *Est-5A, Est-5B*, and *Est-5C* in *D. pseudoobscura* may have occurred in rapid succession, with the *Est-5A:Est-5B* being the most recent. If we assume that K_s is constant between *D. melanogaster* and *D. pseudoobscura*, the β-esterase duplications would have occurred ~8 million years ago (assuming again 2.5 million years for the divergence between *D. melanogaster* and *D. simulans*). The same reservations pointed out above, particularly the possibility of gene conversion, indicate that the inferred time of divergence between the *D. pseudoobscura* β-esterase

genes should be received with much caution. Indeed, BRADY *et al.* (1992; see also OAKESHOTT *et al.* 1995) have proposed that a gene duplication occurred before the divergence of *D. melanogaster* and *D. pseudoobscura,* so that *D. melanogaster Est-P* and *D. pseudoobscura Est-5A* are orthologous to each other, whereas *D. melanogaster Est-6* is orthologous to *D. pseudoobscura Est-5B* (and *Est-5C*).

Pseudogenes are named using the designation of the functional counterpart with additional prefix "ψ". We propose that the *Est-P* pseudogene in *D. melanogaster* be designated as ψ*Est-6*. To prevent confusion in the Drosophila β-esterase nomenclature, we prefer to maintain the suffix "6" in the pseudogene abbreviation, indicating its likely origin as a duplicated product of the gene designated *Est-6* in *D. melanogaster.*

We thank all members of our laboratory, particularly FRANCISCO RODRIGUEZ-TRELLES, ALBERTO GARCIA SAEZ, STEPHEN RICH, ANDREI TATARENKOV, JAN KWIATOWSKI, KEVIN BAILEY, VICTOR DEFILIPPIS, CARLOS MACHADO, HEATHER CARSTENS and DOUGLAS SKARECKY for encouragement, help, and discussion. This work is supported by National Institutes of Health Grant GM-42397 to F.J.A.

LITERATURE CITED

ARNHEIM, N., 1983 Concerted evolution of multigene families, pp. 38–61 in *Evolution of Genes and Proteins,* edited by M. NEI and R. KOEHN. Sinauer Associates, Sunderland, MA.

ARONSHTAM, A. D., and B. A. KUSIN, 1974 Formation of sexual dimorphism of *est-6* gene expression in development of *Drosophila melanogaster* and *Drosophila simulans.* Zh. Obshchei Biologii **35:** 926–933 (in Russian).

ATKINSON, P. W., L. E. MILLS, W. T. STARMER and D. T. SULLIVAN, 1988 Structure and evolution of the *Adh* genes of *Drosophila mojavensis.* Genetics **120:** 713–723.

AUBERT, D., C. BISANZ-SEYER and M. HERZOG, 1992 Mitochondrial *rps* 14 is a transcribed and edited pseudogene in *Arabidopsis thaliana.* Plant Mol. Biol. **20:** 1169–1174.

BRADY, J. P., and R. C. RICHMOND, 1992 An evolutionary model for the duplication and divergence of esterase genes in *Drosophila.* J. Mol. Evol. **34:** 506–521.

BRADY, J. P., R. C. RICHMOND and J. G. OAKESHOTT, 1990 Cloning of the esterase-5 locus from *Drosophila pseudoobscura* and comparison with its homologue in *D. melanogaster.* Mol. Biol. Evol. **7:** 525–546.

CARTER, P., and J. A. WELLS, 1988 Dissecting the catalytic triad of a serine protease. Nature **332:** 564–568.

COLLET, C., K. M. NIELSEN, R. J. RUSSELL, M. KARL, J. G. OAKESHOTT *et al.,* 1990 Molecular analysis of duplicated esterase genes in *Drosophila melanogaster.* Mol. Biol. Evol. **7:** 9–28.

CURRIE, P. D., and D. T. SULLIVAN, 1994 Structure, expression and duplication of genes which encode phosphoglyceromutase of *Drosophila melanogaster.* Genetics **138:** 353–363.

EAST, P. D., A. GRAHAM and G. WHITINGTON, 1990 Molecular isolation and preliminary characterisation of a duplicated esterase locus in *Drosophila buzzatii,* pp. 389–406 in *Ecological and Evolutionary Genetics of Drosophila,* edited by J. S. F. BARKER, W. T. STARMER and R. J. MACINTURE. Plenum Press, New York.

FEDOROFF, N. V., and D. D. BROWN, 1978 The nucleotide sequence of the repeating unit in the oocyte 5S ribosomal DNA of *Xenopus laevis.* Cold Spring Harbor Symp. Quant. Biol. **42:** 1195–1200.

FISHER, R. A., 1935 The sheltering of lethals. Amer. Nat. **69:** 446–453.

FISHER, J. A., and T. MANIATIS, 1985 Structure and transcription of the *Drosophila mulleri* alcohol dehydrogenase. Nucleic Acids Res. **13:** 6899–6917.

FÜRBAß, R., and J. VANSELOW, 1995 An aromatase pseudogene is transcribed in the bovine placenta. Gene **154:** 287–291.

GEORGIEV, P., and V. YELAGIN, 1992 Super-unstable mutations associated with P-M hybrid dysgenesis in *Drosophila melanogaster.* Genetica **87:** 17–29.

GILBERT, D. J., and R. C. RICHMOND, 1982 Esterase 6 in *Drosophila melanogaster:* reproductive function of active and null males at low temperature. Proc. Natl. Acad. Sci. USA **79:** 2962–2966.

GOJOBORI, T., W.-H. LI and D. GRAUR, 1982 Patterns of nucleotide substitution in pseudogenes and functional genes. J. Mol. Evol. **18:** 360–369.

HEALY, M. J., M. M. DUMANCIC and J. G. OAKESHOTT, 1991 Biochemical and physiological studies of soluble esterases from *Drosophila melanogaster.* Biochem. Genet. **29:** 365–388.

HICKEY, D. A., L. BALLY-CUIF, S. ABUKASHAWA, V. PAYANT and B. F. BENKEL, 1991 Concerted evolution of duplicated protein-coding genes in *Drosophila.* Proc. Natl. Acad. Sci. USA **88:** 1611–1615.

HUDSON, R. R., K. BAILEY, A. SKARECKY, J. KWIATOWSKI and F. J. AYALA, 1994 Evidence for positive selection in the superoxide dismutase (*Sod*) region of *Drosophila melanogaster.* Genetics **136:** 1329–1340.

JACQ, C., J. R. MILLER and G. G. BROWNLEE, 1977 A pseudogene structure in 5S DNA of *Xenopus laevis.* Cell **12:** 109–120.

JEFFS, P., and M. ASHBURNER, 1991 Processed pseudogenes in *Drosophila.* Proc. R. Soc. London B **244:** 151–159.

JEFFS, P. S., E. C. HOLMES and M. ASHBURNER, 1994 The molecular evolution of the alcohol dehydrogenase and alcohol dehydrogenase-related genes in the *Drosophila melanogaster* species subgroup. Mol. Biol. Evol. **11:** 287–304.

JOHNSON, F. M., B. B. WALLIS and C. DENNISTON, 1966 Recessive esterase deficiencies controlled by alleles of *Est C* and *Est 6* in *D. melanogaster.* Drosoph. Inf. Serv. **41:** 159.

KIMURA, M., 1983 *The Neutral Theory of Molecular Evolution.* Cambridge University Press, Cambridge.

KOROCHKIN, L., A. ARONSHTAM and N. MATVEEVA, 1974 Genetics of esterases in *Drosophila.* III. Influence of different chromosomes on esterase pattern in *Drosophila.* Biochem. Genet. **12:** 9–24.

KOROCHKIN, L. I., M. Z. LUDWIG, E. V. POLIAKOVA and M. R. PHILINOVA, 1987 Some molecular genetic aspects of cellular differentiation in *Drosophila.* Sov. Sci. Rev. F. Physiol. Gen. Biol. **1:** 411–466.

KOROCHKIN, L., M. LUDWIG, N. TAMARINA, I. USPENSKY, G. YENIKOLOPOV *et al.,* 1990 Molecular genetic mechanisms of tissue-specific esterase isozymes and protein expression in *Drosophila,* pp. 399–440 in *Isozymes: Structure, Function, and Use in Biology and Medicine,* edited by C. MARKERT and J. SCANDALIOS. Wiley-Liss, Inc. New York.

KUMAR, S., K. TAMURA and M. NEI, 1993 MEGA: molecular evolutionary genetics analysis, version 1.0. The Pennsylvania State University, University Park, PA.

KUZIN, B., A. ARONSHTAM and L. KOROCHKIN, 1975 Studies of the determination of the *est-6* gene expression in male genitals of *Drosophila melanogaster.* Ontogenez **6:** 323–330 (in Russian).

KWIATOWSKI, J., D. SKARECKY, S. HERNANDEZ, D. PHAM, F. QUIJAS *et al.,* 1991 High fidelity of the polymerase chain reaction. Mol. Biol. Evol. **8:** 884–887.

KWIATOWSKI, J., D. SKARECKY, K. BAILEY and F. J. AYALA, 1994 Phylogeny of *Drosophila* and related genera inferred from the nucleotide sequence of the CU, ZN *SOD* gene. J. Mol. Evol. **38:** 443–454.

KYLSTEN, P., C. SAMAKOVLIS and D. HULTMARK, 1990 The cecropin locus in *Drosophila;* a compact gene cluster involved in the response to infection. EMBO J. **9:** 217–224.

LANGLEY, C. H., R. A. VOELKER, A. J. LEIGH BROWN, S. OHNISHI, B. DICKSON *et al.,* 1981 Null allele frequencies at allozyme loci in natural populations of *Drosophila melanogaster.* Genetics **99:** 151–156.

LI, W.-H., 1980 Rate of gene silencing at duplicate loci: a theoretical study and interpretation of data from tetraploid fishes. Genetics **95:** 237–258.

LI, W.-H., 1982 Evolutionary change of duplicate genes, pp. 55–92 in *Isozymes: Current Topics in Biological and Medical Research,* Vol. 6, edited by M. C. RATTAZZI, J. G. SCANDALIOS and G. S. WHITT. Alan R. Liss, New York.

LI, W.-H., 1983 Evolution of duplicate genes and pseudogenes, pp. 14–37 in *Evolution of Genes and Proteins,* edited by M. NEI and R. K. KOEHN. Sinauer Associates, Sunderland, MA.

LI, W.-H., 1993 Unbiased estimation of the rates of synonymous and nonsynonymous substitution. J. Mol. Evol. **36:** 96–99.

LI, W.-H., T. GOJOBORI and M. NEI, 1981 Pseudogenes as a paradigm of neutral evolution. Nature **292:** 237–239.

LI, W.-H., C.-I. WU and C.-C. LUO, 1985a A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. Mol. Biol. Evol. **2:** 150–174.

LI, W.-H., C.-C. LUO and C.-I WU, 1985b Evolution of DNA sequences, pp. 1–94 in *Molecular Evolutionary Genetics,* edited by R. J. MACINTYRE. Plenum Press, New York.

LUDWIG, M. Z., N. A. TAMARINA and R. C. RICHMOND, 1993 Localization of sequences controlling the spatial, temporal, and sex-specific expression of the esterase 6 locus in *Drosophila melanogaster* adults. Proc. Natl. Acad. Sci. USA **90:** 6233–6237.

MANE, S. D., C. S. TEPPER and R. C. RICHMOND, 1983 Purification and characterization of esterase 6, a polymorphic carboxylesterase of *Drosophila melanogaster.* Biochem. Genet. **21:** 1019–1040.

MATTERS, G. L., and U. W. GOODENOUGH, 1992 A gene/pseudogene tandem duplication encodes a cysteine-rich protein expressed during zygote development in *Chlamydomonas reinhardtii.* Mol. Gen. Genet. **232:** 81–88.

MCCARREY, J. R., and K. THOMAS, 1987 Human testis-specific *PGK* gene lacks introns and possesses characteristics of a processed gene. Nature **326:** 501–505.

MENOTTI-RAYMOND, M., W. T. STARMER and D. T. SULLIVAN, 1991 Characterization of the structure and evolution of the *Adh* region of *Drosophila hydei.* Genetics **127:** 355–366.

MIYATA, T., and H. HAYASHIDA, 1981 Extraordinarily high evolutionary rate of pseudogenes: evidence for the presence of selective pressure against changes between synonymous codons. Proc. Natl. Acad. Sci. USA **78:** 5739–5743.

MIYATA, T., and T. YASUNAGA, 1981 Rapidly evolving mouse a globin-related pseudo gene and its evolutionary history. Proc. Natl. Acad. Sci. USA **78:** 450–453.

MODIANO, G., G. BATTISTUZZI and A. G. MOTULSKY, 1981 Nonrandom patterns of codon usage and of nucleotide substitutions in human α- and β-globin genes: an evolutionary strategy reducing the rate of mutations with drastic effects? Proc. Natl. Acad. Sci. USA **78:** 1110–1114.

MYERS, M., R. C. RICHMOND and J. G. OAKESHOTT, 1988 On the origins of esterases. Mol. Biol. Evol. **5:** 113–119.

NAGYLAKI, T., 1984 The evolution of multigene families under intrachromosomal gene conversion. Genetics **106:** 529–548.

NEI, M., and A. K. ROYCHOUDHURY, 1973 Probability of fixation of nonfunctional genes at duplicate loci. Amer. Nat. **107:** 362–372.

OAKESHOTT, J. G., P. H. COOKE, R. C. RICHMOND, A. BORTOLI, A. Y. GAME *et al.,* 1989 Molecular population genetics of structural variance of esterase-6 in *Drosophila melanogaster.* Genome **31:** 788–796.

OAKESHOTT, J. G., E. A. VAN PAPENRECHT, T. M. BOYCE, M. J. HEALY and R. J. RUSSELL, 1993 Evolutionary genetics of *Drosophila* esterases. Genetica **90:** 239–268.

OAKESHOTT, J. G., T. M. BOYCE, R. J. RUSSELL and M. J. HEALY, 1995 Molecular insights into the evolution of an enzyme; esterase 6 in *Drosophila.* TREE **10:** 103–110.

OHTA, T., 1984 Some models of gene conversion for treating the evolution of multigene families. Genetics **106:** 517–528.

PALUMBI, S., A. MARTIN, S. ROMANO, W. O. MCMILLAN, L. STICE *et al.,* 1991 The simple fool's guide to PCR, version 2.0. Department of Zoology and Kewalo Marine Laboratory, University of Hawaii, Honolulu.

POWELL, J. R., and R. DESALLE, 1995 *Drosophila* molecular phylogenies and their uses, pp. 87–138 in *Evolutionary Biology,* edited by M. K. HECHT, R. J. MACINTURE and M. T. CLEGG. Plenum Press, New York.

PROCUNIER, W. S., J. J. SMITH and R. C. RICHMOND, 1991 Physical mapping of the *Esterase-6* locus of *Drosophila melanogaster.* Genetica **84:** 203–208.

PRUDHOMMEAU, C., and J. PROUST, 1990 I-R hybrid dysgenesis in *Drosophila melanogaster*—nature and site specificity of induced recessive lethals. Mutat. Res. **230:** 135–157.

RICHMOND, R. C., D. G. GILBERT, K. B. SHEEHAN, M. H. GROMKO and F. M. BUTTERWORTH, 1980 Esterase 6 and reproduction in *Drosophila melanogaster.* Science **207:** 1483–1485.

RICHMOND, R. C., K. M. NIELSEN, J. P. BRADY and E. M. SNELLA, 1990 Physiology, biochemistry and molecular biology of the *Est-6* locus in *Drosophila melanogaster,* pp. 273–292 in *Ecological and Evolutionary Genetics of Drosophila,* edited by J. S. F. BARKER, W. T. STARMER and R. J. MACINTURE. Plenum Press, New York.

SAAD, M., A. Y. GAME, M. HEALY and J. G. OAKESHOTT, 1994 Associations of esterase 6 allozyme and activity variation with reproductive fitness in *Drosophila melanogaster.* Genetica **94:** 43–56.

SAITOU, N., and M. NEI, 1987 The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4:** 406–425.

SANGER, F., S. NICKLEN and A. R. COULSON, 1977 DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **75:** 5463–5467.

SATTA, Y., 1993 How the ratio of nonsynonymous to synonymous pseudogene substitutions can be less than one. Immunogenetics **38:** 450–454.

SCHAFER, D., 1992 Developmental and molecular studies of *Adh* in *Drosophila buzzatii.* M.Sc. Thesis, University of New England, Armidale, New South Wales.

SCHRAG, J. D., Y. LI, S. WU and M. CYGLER, 1991 Ser-His-Glu triad forms the catalytic site of the lipase from *Geotrichum candidum.* Nature **351:** 761–764.

SEAGER, R. D., and F. J. AYALA, 1982 Chromosome interactions in *Drosophila melanogaster.* I. Viability studies. Genetics **102:** 467–483.

SHEEHAN, K. B., R. C. RICHMOND and B. J. COCHRANE, 1979 Studies of esterase 6 in *Drosophila melanogaster.* III. The developmental pattern and tissue distribution. Insect Biochem. **9:** 443–450.

SHIBATA, H., and T. YAMAZAKI, 1995 Molecular evolution of the duplicated *Amy* locus in the *Drosophila melanogaster* species subgroup: concerted evolution only in the coding region and an excess of nonsynonymous substitutions in speciation. Genetics **141:** 223–236.

SIBBALD, P. R., and B. J. BLENCOWE, 1990 A completely conserved rat U6 snRNA pseudogene coding sequence is sandwiched between a cytochrome c retropseudogene and a LINE-like sequence. Nucleic Acids Res. **18:** 1063.

SORGE, J., E. GROSS, C. WEST and E. BEUTLER, 1990 High level transcription of the *glucocerebrosidase* pseudogene in normal subjects and patients with Gaucher disease. J. Clin. Invest. **86:** 1137–1141.

SULLIVAN, D. T., W. T. STARMER, S. W. CURTISS, M. MENOTTI-RAYMOND and J. YUM, 1994 Unusual molecular evolution of an *Adh* pseudogene in *Drosophila.* Mol. Biol. Evol. **11:** 443–458.

SUSSMAN, J. L., M. HAREL, F. FROLOW, C. OEFNER, A. GOLDMAN *et al.,* 1991 Atomic structure of acetylcholinesterase from *Torpedo californica:* a prototypic acetylcholine-binding protein? Science **253:** 872–879.

SVED, J. A., 1979 The "hybrid dysgenesis" syndrome in *Drosophila melanogaster.* Bioscience **29:** 659–664.

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.

THROCKMORTON, L. H., 1975 The phylogeny, ecology and geography of *Drosophila,* pp. 421–469 in *Handbook of Genetics,* Vol. 3, edited by R. C. KING. Plenum Press, New York.

USPENSKY, I., M. LUDWIG and L. KOROCHKIN, 1988 Comparative analysis of carboxyl esterase in different organs of *Drosophila* male reproductive system in the *melanogaster* species group. Zh. Obshchei Biologii **49:** 601–610 (in Russian).

VANIN, E., 1985 Processed pseudogenes: characteristics and evolution. Ann. Rev. Genet. **19:** 253–272.

VOELKER, R. A., C. H. LANGLEY, A. J. LEIGH BROWN, S. OHNISHI, B. DICKSON *et al.,* 1980 Enzyme null alleles in natural populations of *Drosophila melanogaster:* frequencies in a North Carolina population. Proc. Natl. Acad. Sci. USA **77:** 1091–1095.

WEINER, A. M., P. L. DEININGER and A. EFSTRATIADIS, 1986 Nonviral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. Ann. Rev. Biochem. **55:** 631–661.

WILDE, C. D., 1986 Pseudogenes. CRC Crit. Rev. Biochem. **19:** 323–352.

YENIKOLOPOV, G., B. KUZIN, M. EVGENIEV, M. LUDWIG, L. KOROCHKIN *et al.,* 1983 The cloning and expression of the gene encoding organ-specific *Est-S* from the genome of *D. virilis.* EMBO J. **2:** 1–7.

YENIKOLOPOV, G., R. KHECHUMIAN, B. KUZIN, L. KOROCHKIN and G. GEORGIEV, 1989a Molecular organization of the esterase multigene family of *Drosophila virilis.* Tissue-specific expression of the *Est-S* gene during the development of *D. virilis.* Genetika **25:** 406–416 (in Russian).

YENIKOLOPOV, G. N., O. A. MALEVANTSCHUK, N. I. PEUNOVA, P. V. SERGEEV and G. P. GEORGIEV, 1989b *Est* locus of *Drosophila virilis* contains two related genes. Dokl. Acad. Nauk SSSR **306:** 1247–1249 (in Russian).

YUM, J., W. T. STARMER and D. T. SULLIVAN, 1991 The structure of the *Adh* locus of *Drosophila mettleri:* an intermediate in the evolution of the *Adh* locus in the *repleta* group of *Drosophila.* Mol. Biol. Evol. **8:** 857–867.

Communicating editor: W.-H. LI