J Mol Evol (1993) 36:243–248

Journal of
Molecular Evolution
© Springer-Verlag New York Inc. 1993

Nucleotide Divergence of the *rp49* Gene Region Between *Drosophila melanogaster* and Two Species of the *Obscura* Group of Drosophila

Carmen Segarra[1,*] and Montserrat Aguadé[2]

[1] Museum of Comparative Zoology, Harvard University, Cambridge, MA 02138
[2] Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Diagonal 645, E-08071 Barcelona, Spain

Summary. A 2.1-kb *SstI* fragment including the *rp49* gene and the 3′ end of the δ-*serendipity* gene has been cloned and sequenced in *Drosophila pseudoobscura*. *rp49* maps at region 62 on the tip of chromosome II of this species. Both the coding and flanking regions have been aligned and compared with those of *D. subobscura*. There is no evidence for heterogeneity in the rate of silent substitution between the *rp49* coding region and the rate of substitutions in flanking regions, the overall silent divergence per site being 0.19. Noncoding regions also differ between both species by different insertions/deletions, some of which are related to repeated sequences. The *rp49* region of *D. pseudoobscura* shows a strong codon bias similar to those of *D. subobscura* and *D. melanogaster*. Comparison of the rates of silent (K_s) and nonsilent (K_a) substitutions of the *rp49* gene and other genes completely sequenced in *D. pseudoobscura* and *D. melanogaster* confirms previous results indicating that *rp49* is evolving slowly both at silent and nonsilent sites. According to the data for the *rp49* region, *D. pseudoobscura* and *D. subobscura* lineages would have diverged some 9 Myr ago, if one assumes a divergence time of 30 Myr for the *melanogaster* and *obscura* groups.

Key words: *rp49* gene — *Drosophila* — Sequence divergence

Present address: Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Diagonal 645, E-08071 Barcelona, Spain.
Offprint requests to: C. Segarra

Introduction

Comparison of the *rp49* coding region (which codes for a ribosomal protein) between *Drosophila subobscura* and *D. melanogaster* has shown a low rate of amino acid replacements, an indication of strong purifying selection acting against amino acid replacements (Aguadé 1988). A high level of constraint at the protein level could be expected for a structural protein like *rp49*, with many putative interactions not only with other ribosomal proteins and RNAs but also with other components of the translation apparatus. On the other hand, the *rp49* coding region has also revealed a low rate of silent or synonymous nucleotide substitutions when different coding regions were compared between species of the *melanogaster* and *obscura* groups (Sharp and Li 1989). These authors found a correlation between silent site divergence in coding regions and codon bias in synonymous codon usage, in the sense that regions which exhibited strong codon bias seemed to be evolving at a lower rate. Therefore the *rp49* coding region would seem to be highly constrained to evolve not only at nonsilent sites but also at silent sites.

Estimates of nucleotide divergence for the *rp49* region, as well as for most other regions, are based on the sequence of only one species from the *melanogaster* and the *obscura* groups. General assertions about rates of divergence at silent and nonsilent sites should be based on sequences from more than one species in each group. In the present study, the region including the *rp49* gene of *D. pseudoobscura* has been cloned and sequenced, and its sequence has been compared to those of *D. mela-*

nogaster (O'Connell and Rosbash 1984; Cadic-Jacquier and Rosbash, personal communication) and *D. subobscura* (Aguadé 1988). Comparison of the *rp49* noncoding regions between the more related species, *D. pseudoobscura* and *D. subobscura*, has allowed study of how molecular evolution has operated in these noncoding regions. Moreover, estimates of the rate of nucleotide substitutions in the *rp49* coding regions in the three pairwise species comparisons have allowed confirmation of low rates of silent and nonsilent divergence for the *rp49* region and a rough estimate of the time of divergence between *D. pseudoobscura* and *D. subobscura*.

Material and Methods

The AH162 *D. pseudoobscura* genome library (Schaeffer et al. 1987) was screened according to Benton and Davis (1977). A 0.8-kb *DdeI* fragment that includes the *rp49* coding region of *D. subobscura* (Aguadé 1988) was used as probe. Three positive phages, λ*Dpseudo rp49-1*, λ*Dpseudo rp49-2*, and λ*Dpseudo rp49-3*, were identified and their DNA was purified (Maniatis et al. 1982). After digestion with several restriction enzymes and analysis by Southern blot (using a 1.6-kb *AvaI-EcoRI* fragment from *D. subobscura* as probe), a 2.1-kb *SstI* fragment of one of the positives, λ*Dpseudo rp49-1*, was subcloned into M13MP18 (Yanisch-Perron et al. 1985). A set of nested deletions of each of two recombinant clones which had the insert in opposite orientations (MP18-6 and MP18-12) was obtained according to Henikoff (1984). Subclones with decreasing lengths were sequenced by the "dideoxy" method (Sanger et al. 1977; Biggin et al. 1983) using modified T7 DNA polymerase (Tabor and Richardson 1987). Both strands (clones MP18-6 and MP18-12) were completely sequenced and on average each nucleotide was sequenced 7.5 times.

Sequences were aligned using the IBI/Pustell sequence analysis programs placing insertions and deletions to minimize the number of differences.

Cytological location of the *rp49* region in *D. pseudoobscura* was determined by in situ hybridization using as probe the 2.1-kb *SstI* fragment labeled with biotin (Langer-Safer et al. 1982). The pUC38 clone of *D. subobscura* (Aguadé 1988) was used as a control.

Results

The complete sequence of a 2.1-kb *SstI* fragment including the *rp49* region of *D. pseudoobscura* is presented in Fig. 1. The reported sequence includes the 5′ flanking region (1,364 bp), the *rp49* coding region with two exons (93 and 309 bp long) and one intron (72 bp), the 3′ intergenic region (236 bp), and the last 45 bp of the δ gene included in the *serendipity* complex (Vincent et al. 1985). Comparison with consensus sequences as well as with the sequence of this region in *D. subobscura* and *D. melanogaster* has allowed identification of putative signal sequences in *D. pseudoobscura* (Fig. 1). There is ambiguity in the location of the putative CAAT box (position 1,287–1,293) and putative transcription initiation site (position 1,362–1,368) due to differences with the other sequences.

The region sequenced in *D. subobscura* (Aguadé 1988) can be completely aligned with that of *D. pseudoobscura*. However, only coding regions (*rp49* and δ-*sry*) can be aligned between *D. melanogaster* and any of these species, since flanking regions have accumulated a high number of nucleotide substitutions and insertions/deletions.

The 5′ flanking region of *D. pseudoobscura*, like that of *D. subobscura*, shows different repeated short motifs between positions 581 and 621 (TGG, CTT, GTT). Likewise, the *rp49* intron of *D. pseudoobscura* presents the motif TAGTG repeated almost invariantly four times (positions, 1,504–1,508 and 1,519–1,533), while in *D. subobscura* the repeated motif is AATGG (position 1,509–1,524) (Fig. 1).

The *D. pseudoobscura* and *D. subobscura* sequences differ by 28 insertions/deletions, all located in flanking regions or in the intron, their size varying between 1 and 38 bp. Of the 21 length differences between species longer than 1 bp, seven are related to repeated sequences, i.e., length differences starting at positions 607, 675, 771, 805, 1,019, 1,347 and 2,141 (Fig. 1).

The rate of substitutions in the different noncoding regions and the rate of silent substitutions in the *rp49* coding region when comparing *D. pseudoobscura* and *D. subobscura* are summarized in Table 1. The number of silent sites in the *rp49* coding region has been estimated according to Li et al. (1985). These rates do not differ significantly in the different regions ($\chi^2 = 1.94$, 3 *df*, $P > 0.3$) (Table 1).

The *rp49* coding region in *D. pseudoobscura* has been compared with both that of *D. subobscura* (Aguadé 1988) and that of *D. melanogaster* (O'Connell and Rosbash 1984; Cadic-Jacquier and Rosbash, personal communication). Nucleotide divergence for the *rp49* coding region has been estimated separately for silent (K_s) and nonsilent (K_a) nucleotide substitutions according to Li et al. (1985) (Table 2).

The *rp49* coding region of *D. pseudoobscura*, like those of *D. subobscura* and *D. melanogaster*, shows a strong codon bias. Both the general χ^2 ($\chi^2 = 93.48$, 39 *df*, $P < 0.001$) and the "scaled" χ^2 proposed by Shields et al. (1988) to take into account the number of codons ($\chi^2_{sc} = 0.72$) indicate a strong codon bias or unequal use of synonymous codons. *D. pseudoobscura* also shows a high G+C content at the third position (0.72), an indication (as suggested by these authors) of strong codon bias. In order to test whether the three species show a similar codon bias for the *rp49* region, base composi-

```
              *         *         *         *         *         *
   1 GAGCTCCCTGTATCCATGATGGGCCACTACATATCTGCATAGGGATTCAAATATCGCGTT

              *         *         *         *         *         *
  61 TCATTATTTGCTATATTATTGTTTATATATTTCGAGTGCTGCGGCTCTCACCTGAAAGCA

              *         *         *         *         *         *
 121 TTTGAGGTATTGGAGAAATGTTGATGGAATTTCGAGTCCGCCAGCAGCGTTTGCACGCGT

              *         *         *         *         *         *
 181 TCAAAGGGCAGGAGAATGGATTCAGCACTACCCGCCACTACGCCGGCAATCACCTTGGCA

              *         *         *         *         *         *
 241 CCGTAGTCATTCAGGCGATAGTCTTCCACCAGATATCTCCTCGTGCCATCAAACACTCCA

              *         *         *         *         *         *
 301 AACATTATGGACAACGAAATGGTTTTCTGTGCCAATGGAGGCAGCATTCCTCGGTACAGA

              *         *         *         *         *         *
 361 AAGCTAAGCCCCTCATGCCGCAGCTGACCGAAGGCAGATTTGATGGGCACGCCATGAAGC

              *         *         *         *         *         *
 421 ATTTGCCGAAAGATCATTTTGTAGATGGGATATGTGGCGGAAATGTTGACAAACGCTGCT

              *         *         *  G A    *         *C        *
 481 CCGCATCCGCAGGCGAACTCCTCCCACTGGAAGGCGCCAAAGAAGCGCTCTGTGAAGATG

              * G  CCGA *       GA*   G CC  T    TG   *  AGG    *
 541 CTCGCATGGCTAGA---TGATACACCCAGCTTTAGTGCTGTGGGTGGTGGTGCTCTTCTT

           GCTGTT   G   *   T G   *         G   T     *       C *
 601 GTTGTT------GTTTTTGTTGCGTATGCATTCTGTGCCTCGGAGTCATCATCGTCTTTC

              *     CTGTAG     A  *   A     *  T      *      T  *
 661 ATGTTGCTACTGTAG------TGGCTGGTGCGATAATGTGTGCAATTGTCACCTCGCTAC

            C *GAC   -  *  A G T AG-        *T        *-----   A*
 721 CATTGTTTACAGTGGTAACTCCGCACACCTAACACTCGGCGTGTGATTTGTTTGTTTTGA

     A  A ATT *         *    GAACC*         *    ------   G G-- *
 781 TTTTTTACTAAATTATATAATAAA-----GAACCACGGTATTACAACTAGGAATGCAACG

              --  CAGACGGA        *--------------GG   * C       *
 841 TGCTATCGAATTGTCT-----GTATATCGTTTGATAGTTTTAGAATCGATATATCGATAC

     G ACA C  *T   A   T*  C C  CC*CTA   A  TAA  T    A      C  T
 901 AAGATTAAATGTAATAAACTAGTTGTAGAGTGGTTTTAAGGGTTGTAAACAATATAATTG

      GT A GC *    GTG A*        T*    C   C*  G T    * T CC   CT
 961 AAATTTTAATTGTAACATGGCATTTTAAAAGATGTACCTCTTACAAGAATTGGTAAAA--

     AAA  G   *   G     *     A   A  TAAC  G*     G   AA        *
1021 ---TTATTGTTTACCACTATCGATAGCCAGAA--GACAACTCTGGTCGAG-CTGTCAAAA

      C       *         *         *         *         *         *
1081 TGTAGTGCGGCCAACTTCACATCCCTTTTCTCTTTTTGACGTGCTCTTTACAATTAGC

              * C       *         *  T      *    TC   *  C  T  A*
1141 GGTTCGTTTTTGCTATTAAATCTGAATATTTCCTAAGCATCCGAGTAACATGTTGGCTCT

       AC C A *         *CAT      * A T     *      T  *         G
1201 TTCTAAGCTAAATAAATGCAGTAACAGTGAAGCAAGACACTGGCGCAGTAATTAAATCTA

     AG      C* TT      *       C * T  →  ← *         *    GGTGG*
1261 TATTTCCATGGCATCCTCAAGAAAATGTCAACCAAACTTTGTTGACATTTTCTCA---CA

        C GA  *       C *     AATCTACAA    G*     GTG *         *
1321 GGAAGACACAATAAATAAAAATCGCG--------ATATTCAATCAAGATACTGATGAGTG

                                      C  C        A   A
         T TT * A     G *         *         *         *         *
1381 GTCTCAACTATCAAATTCCAGCTCCAAAATGACGATTCGCCCAGCGTACCGGCCCAAGAT
                                 MetThrIleArgProAlaTyrArgProLysIl

      G G               G          C              A     T   C
              *         *         *C        *     C   *         *
1441 CATCAAGAAGCGCACCAAGCACTTCATCCGTCACCAGTCGGATCGTTATGCCAAGTTGTC
     eIleLysLysArgThrLysHisPheIleArgHisGlnSerAspArgTyrAlaLysLeuSe

        A     -  GG   GG*A   G  A GC  T T C-----------*         *
1501 GGTTAGTAATATAAAATATAGTGTAGTTTAGTGAAAGATAAACCATATACTAATCATGGT
     r

                                      C  G     C           CG
           T   T C      *         *     A   *         *   A     *
1561 CTTTAT-GTT-GTAGCACAAATGGCGCAAGCCTAAGGGTATTGACAACAGAGTGCGTCGT
                    HisLysTrpArgLysProLysGlyIleAspAsnArgValArgArg

      C         A  G- TC                   C  A  G        C     C
              *    G   C*         *        T*         *   C     *
1621 CGCTTCAAGGGCCAATACTTGATGCCCAACATCGGTTACGGCTCCAACAAGCGTACCCGT
     ArgPheLysGlyGlnTyrLeuMetProAsnIleGlyTyrGlySerAsnLysArgThrArg

       C           C  A        G              C     C
              *         *         *         *         *         T
1681 CATATGCTGCCCACTGGCTTCAAGAAATTCCTGGTGCACAATGTGCGAGAGCTGGAGGTC
     HisMetLeuProThrGlyPheLysLysPheLeuValHisAsnValArgGluLeuGluVal

          C            G T      CG     G  C    G         C
          C   *        A*         *        C*         *         *
1741 CTGCTTATGCAGAACCGCGTCTACTGCGGTGAGATCGCTCACGCCGTCTCCTCAAAGAAG
     LeuLeuMetGlnAsnArgValTyrCysGlyGluIleAlaHisAlaValSerSerLysLys

             G  T     G                  G    C  C  C  C     C  T
              *       G *         *         *    G    *      T  *
1801 CGCAAGGAAATCGTCGAACGCGCCAAGCAGCTGTCGATCCGTCTAACGAATCCCAACGGA
     ArgLysGluIleValGluArgAlaLysGlnLeuSerIleArgLeuThrAsnProAsnGly

       C
              *         *         *        G*         *      T  *
1861 CGTCTGCGTTCTCAAGAGAACGAGTAACCTTAAGATTGAAATGCTCTTTGTCGAGTT-AT
     ArgLeuArgSerGlnGluAsnGlu

              A          CC       *         *     G C *  C      T
1921 TGTAACGTGGTCGGAATACAAAATTTGAAACGTTAAAAGTGAAAATCGAAAGTTGTTATA

     AT  G--------------------------------------C G   TA A A    *
1981 CCTTCAATATTAATATGTTCCACAGAATCTAGAAAATCGAATATCTCTTCGAGATCATTG

     T ---------------------      TC    TC  * AATTATACGATGAATTA *
2041 GTTTTTGATTTTAATAACATGAGTTTATTATGTTA--AGTTG----------------CA

       TG  GTGGGTAAGG   *     T   G    -    AAAGTAAAAATGA       *
2101 CAAAAAA----------GGGAATATTACTAAAGTACAAGTC------------AAAGTCTA

        TCA GAC  G        A     C     G  T  G
        G     *  A      * G     C * A     A  ¡¡¡
2161 TTTTAGAACGACGTCGTGATCAAATTGTAAGTGGTGCTCGAGCTC
```

Fig. 1. Nucleotide sequence of the *rp49* gene region from *D. pseudoobscura*. The putative TATA box (AATAAATA) and polyadenylation signal (AATACA) are *boxed*. A palindromic motif found in the 5′ flanking region is indicated by *arrows*. The nucleotide substitutions with the sequence of *D. subobscura* are shown immediately above the sequence from position 513 to the end. On the top line of the *rp49* coding region (positions 1,409–1,501 and 1,576–1,884), the nucleotide substitutions with respect to the sequence of *D. melanogaster* are also shown. *Dashes* (-) indicate missing nucleotides. *Upside-down exclamation marks* (¡) indicate nucleotides not compared at the end of the sequence. The amino acid sequence of ribosomal protein 49 is presented below the nucleotide sequence. Amino acid replacements between *D. subobscura* and *D. melanogaster* are *underlined* and those between *D. pseudoobscura* and *D. melanogaster* are *bold printed*. The nucleotides that in the complementary strand correspond to the polyadenylation signal and stop codons of the δ-*sry* gene are *underlined*. The *bold-printed* nucleotides correspond to the reverse and complementary sequence of the 3′ end of the δ-*sry* gene coding region. On the top line of this coding region (position 2161 to the end) the nucleotide substitutions with respect to the sequence of *D. melanogaster* are also shown. The C-terminal end of the protein in *D. pseudoobscura* has the amino acid sequence GLU-LEU-GLU-HIS-HIS-LEU-GLN-PHE-ASP-HIS-ASP-VAL-VAL-LEU-LYS-COOH.

tion at the third position of exons (excluding Met, Trp, and stop codons) has been compared ($\chi^2 = 0.61$, 6 *df*, $P > 0.3$).

In situ hybridization reveals that the *rp49* single copy gene is located in the more distal region (62) of chromosome II of *D. pseudoobscura* as described in the photographic map of Kastritsis and Crumpacker (1966).

Table 1. Rate of silent nucleotide substitutions in the *rp49* coding region and rate of substitutions in noncoding regions between *D. pseudoobscura* and *D. subobscura* estimated according to Jukes and Cantor (1969): standard errors are given in parentheses (Kimura and Otha 1972)

Region	No. sites compared	No. substitutions	Corrected substitution rate
5′ flanking region	822	140	0.1932 (0.0170)
Coding region			
Exons	89	14	0.1765 (0.0488)
Intron	60	14	0.2795 (0.0793)
Intergenic region	170	27	0.1785 (0.0356)
Total	1,141	195	0.1939 (0.0144)

Table 2. Rate of silent (K_s) and nonsilent (K_a) substitutions per site in the *rp49* coding region estimated according to Li et al. (1985): standard errors are given in parentheses

Species compared	K_s	K_a
D. pseudoobscura/ *D. subobscura*	0.1769 (0.0506)	0.0032 (0.0032)
D. subobscura/ *D. melanogaster*	0.5611 (0.1047)	0.0273 (0.0095)
D. pseudoobscura/ *D. melanogaster*	0.6415 (0.1220)	0.0239 (0.0089)

Table 3. Rate of silent (K_s) and nonsilent (K_a) substitutions per site estimated according to Li et al. (1985) in different genes completely sequenced in *D. pseudoobscura* and *D. melanogaster:* standard errors are given in parentheses

Genes[a]	No. codons compared	K_s	K_a
rp49[1]	133	0.6415 (0.1220)	0.0239 (0.0089)
Adh[2]	253	0.6211 (0.0819)	0.0520 (0.0097)
Gld[3]	611	1.0708 (0.0983)	0.0565 (0.0066)
Amy[4]	493	0.4025 (0.0423)	0.0678 (0.0080)
Xdh[5]	1334	1.1866 (0.0704)	0.0709 (0.0050)
UO[6]	344	0.8670 (0.0988)	0.0814 (0.0105)
Gart[7]	1352	1.2133 (0.0697)	0.0890 (0.0057)
bicoid[8]	478	1.0150 (0.1005)	0.1395 (0.0121)
pcp[9]	183	1.0522 (0.1758)	0.1616 (0.0216)
Est[10]	541	1.4120 (0.1422)	0.1778 (0.0131)

[a] Data sources: [1]O'Connell and Rosbash (1984), Cadic-Jacquier and Rosbash (personal communication), present work; [2]Kreitman (1983); Schaeffer and Aquadro (1987); [3]Krasney et al. (1990); [4]Brown et al. (1990); [5]Riley (1989); [6]Friedman et al (1992); [7]Henikoff and Eghtedarzadeh (1987), Henikoff et al. (1983, 1986), Henikoff and Furlong (1983); [8]Seeger and Kaufman (1990); [9]Henikoff and Eghtedarzadeh (1987), Henikoff et al. (1986); [10]Brady et al. (1990).

Discussion

Comparison of two closely related species, *D. pseudoobscura* and *D. subobscura,* offers the possibility to align noncoding regions and study mechanisms involved in its divergence. The sequences of *D. pseudoobscura* and *D. subobscura* present seven length differences associated with repeats. This result seems to indicate that slipped strand mispairing (Levinson and Gutman 1987), a mechanism that might be generating length variants within *D. subobscura* (Rozas and Aguadé 1990), might also have caused or be causing length differences between *D. subobscura* and *D. pseudoobscura.*

The present sequence of *D. pseudoobscura* enlarges the sample of genes completely sequenced in both *D. pseudoobscura* and *D. melanogaster* (Table 3). Comparison of K_s and K_a rates among these genes allows one to establish whether a given region is more or less strongly constrained to vary at nonsilent and/or silent sites. Nonsilent divergence is lowest for the *rp49* region, confirming the important role of purifying selection in the evolution of this gene; only the *hsp82* region that has only been partially sequenced in *D. pseudoobscura* (Blackman and Meselson 1986) shows an even lower nonsilent divergence (K_a = 0.0182, SE = 0.0046). Estimated silent divergences for the *Amy, Adh,* and *rp49* regions (as well as for the *hsp82* region —K_s = 0.6120, SE = 0.0726−) are low as compared to the other genes (Table 3); as pointed out by Sharp and Li (1989) and by Brown et al. (1990), these four genes also exhibit a high codon bias that might be slowing down their evolution at silent sites. The *rp49* region would not only be subject to purifying selection against amino acid replacement substitutions but to selection pressure for the use of certain codons within synonymous class.

If the rate of silent substitutions at the coding region of *rp49* was reduced by purifying selection, one would expect this rate to be lower than the rate of substitutions in flanking regions. The lack of evidence for any difference in these rates (Table 1) might reflect the existence of constraints to vary in flanking regions due to the presence of transcription signals. In addition to the putative general sequences of eukaryotic promoters in the 5′ flanking region of *rp49* there is a palindromic sequence and a stretch of 69 nucleotides completely conserved (position 1,083–1,151) (Fig. 1) that might be contributing to this constraint.

Although estimates of nucleotide divergence for the *rp49* coding region are based on a relatively low number of codons, they have been used to obtain a rough estimate of the time of divergence between *D. pseudoobscura* and *D. subobscura,* assuming a constant rate per unit of time. Given the close relationship between these two species, silent-site substitution rates have been used for this purpose, as

they are higher and not so affected by estimation errors (Nei 1987). Although the *rp49* region shows a strong codon bias, the three species being compared do not differ significantly in their codon bias (See Results.) According to the rate of silent substitutions in the *rp49* coding region the split between the *melanogaster* and *obscura* groups would be 3.4 times older than the split between *D. pseudoobscura* and *D. subobscura*. At least two different divergence times have been proposed for the *melanogaster/obscura* split: 30 Myr (Throckmorton 1975; Moriyama 1987) and 46 Myr (Beverley and Wilson 1984). According to the data for the *rp49* region, *D. pseudoobscura* and *D. subobscura* lineages would have diverged some 9 or 14 Myr ago. Only when better estimates of divergence time between *melanogaster* and *obscura* groups and when a large collection of genomic regions sequenced in both species are available for comparison will more accurate estimates of the time of split between *D. pseudoobscura* and *D. subobscura* be obtained.

Acknowledgments. We thank R.C. Lewontin for encouragement and criticism of the present work, S.W. Schaeffer for the *D. pseudoobscura* genome library, and W-S. Li for programs. C.S. also thanks members of the Population Genetics Group in the Museum of Comparative Zoology (Harvard University), especially S.R. Kaplan for technical assistance. This work has been supported by a Fulbright fellowship awarded to C.S. by Ministerio de Educación y Ciencia, Spain.

References

Aguadé M (1988) Nucleotide sequence comparison of the rp49 gene region between *Drosophila subobscura* and *D. melanogaster*. Mol Bio Evol 5:433–441

Benton WD, Davis RW (1977) Screening λgt recombinant clones by hybridization to single plaques in situ. Science 196:180–183

Beverley SM, Wilson AC (1984) Molecular evolution in *Drosophila* and the higher Diptera II. A time scale for fly evolution. J Mol Evol 21:1–13

Biggin MD, Gibson TJ, Hong GF (1983) Buffer gradient gels and ^{35}S label as an aid to rapid DNA sequence determination. Proc Natl Acad Sci USA 80:3963–3965

Blackman RK, Meselson M (1986) Interspecific nucleotide sequence comparisons used to identify regulatory and structural features of the *Drosophila hsp82* gene. J Mol Biol 188: 499–515

Brady JP, Richmond RC, Oakeshott JG (1990) Cloning of the Esterase-5 locus from *D. pseudoobscura* and comparison with its homologue in *D. melanogaster*. J Mol Biol 7:525–546

Brown CJ, Aquadro CF, Anderson WW (1990) DNA sequence evolution of the Amylase multigene family in *Drosophila pseudoobscura*. Genetics 126:131–138

Friedman TB, Burnett JB, Lootens S, Steinmen R, Wallrath LL (1992) The urate oxidase gene of *Drosophila pseudoobscura* and *Drosophila melanogaster*. Evolutionary changes of sequence and regulation. J Mol Evol 34:62–77

Henikoff S (1984) Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene 28: 351–359

Henikoff S, Eghtedarzadeh MK (1987) Conserved arrangement of nested genes at the *Drosophila Gart* locus. Genetics 117: 711–725

Henikoff S, Furlong CE (1983) Sequence of a *Drosophila* DNA segment that functions in *Saccharomyces cerevisiae* and its regulation by a yeast promoter. Nucleic Acids Res 11:789–800

Henikoff S, Sloan JS, Kelly JD (1983) A *Drosophila* metabolic gene transcript is alternatively processed. Cell 34:405–414

Henikoff S, Keene A, Fechtel K, Fristom JM (1986) Gene within a gene: nested *Drosophila* genes encode unrelated proteins on opposite DNA strands. Cell 44:33–42

Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: Munro HW (ed) Mammalian protein metabolism. Academic Press, New York, pp 21–120

Kastritsis CD, Crumpacker DW (1966) Gene arrangements in the third chromosome of *Drosophila pseudoobscura*. I. Configurations with tester chromosomes. Heredity 57:151–158

Kimura M, Ohta T (1972) On the stochastic model for estimation of mutational distance between homologous proteins. J Mol Evol 2:87–90

Krasney PA, Carr C, Cavener DR (1990) Evolution of the glucose dehydrogenase gene in *Drosophila*. Mol Evol Biol 7: 155–177

Kreitman M (1983) Nucleotide polymorphism at the alcohol dehydrogenase locus in *Drosophila melanogaster*. Nature 304: 412–417

Langer-Safer P.R, Levine M, Ward DC (1982) Immunological method for mapping genes on *Drosophila* polytene chromosomes. Proc Natl Acad Sci USA 79:4381–4385

Levinson G, Gutman GA (1987) Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Mol Biol Evol 4:203–221

Li W-H, Wu C-I, Luo C-C (1985) A new method for estimating synonymous and non synonymous rates of nucleotide substitution considering the relative likehood of nucleotide and codon usage. Mol Biol Evol 2:150–174

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Moriyama EN (1987) Higher rates of nucleotide substitution in *Drosophila* than in mammals. Jpn J Genet 62:139–147

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

O'Connell P, Rosbash M (1984) Sequence, structure and codon preference of the *Drosophila* ribosomal protein 49 gene. Nucleic Acids Res 12:5495–5513

Riley MA (1989) Nucleotide sequence of the *Xdh* region in *Drosophila pseudoobscura* and an analysis of the evolution of synonymous codons. Mol Biol Evol 6:33–52

Rozas J, M Aguadé (1990) Evidence of genetic exchange in the *rp49* region among polymorphic chromosome inversions in *Drosophila subobscura*. Genetics 126:417–426

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74: 5463–5467

Schaeffer SW, Aquadro CF, Anderson WW (1987) Restriction-map variation in the alcohol dehydrogenase region of *Drosophila pseudoobscura*. Mol Biol Evol 4:254–265

Schaeffer SW, Aquadro CF (1987) Nucleotide sequence of the alcohol dehydrogenase region of *Drosophila pseudoobscura*: evolutionary change and evidence for an ancient duplication. Genetics 117:61–73

Seeger MA, Kaufman TC (1990) Molecular analysis of the *bicoid* gene from *Drosophila pseudoobscura*: identification of con-

served domains within coding and noncoding regions of the *bicoid* mRNA. EMBO J 9:2977–2987

Sharp PM, Li W-H (1989) On the rate of DNA sequence evolution in *Drosophila*. J Mol Evol 28:398–402

Shields DC, Sharp PM, Higgins DG, Wright F (1988) "Silent" sites in *Drosophila* genes are not neutral: evidence of selection among synonymous codons. Mol Biol Evol 5:704–716

Tabor S, Richardson CC (1987) DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. Proc Natl Acad Sci USA 84:4767–4771

Throckmorton LH (1975) The phylogeny, ecology and geography of *Drosophila*. In: King RC (ed) Handbook of genetics, vol 3. Plenum Press, New York, pp 421–469

Vincent A, Colot HV, Rosbash M (1985) Sequence and structure of the *serendipity* locus of *Drosophila melanogaster*. A densely transcribed region including a blastoderm-specific gene. J Mol Biol 186:149–166

Yanish-Perron C, Vieira J, Messing J (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13MP18 and pUC19 vectors. Gene 33:103–119

Received February 11, 1992/Revised September 1, 1992