

UC Irvine

UC Irvine Previously Published Works

Title

Erratic Evolution of Glycerol-3-Phosphate Dehydrogenase in *Drosophila*, *Chymomyza*, and *Ceratitis*

Permalink

<https://escholarship.org/uc/item/42s3f5sw>

Journal

Journal of Molecular Evolution, 44(1)

ISSN

0022-2844

Authors

Kwiatowski, Jan
Krawczyk, Michal
Jaworski, Michal
[et al.](#)

Publication Date

1997

DOI

10.1007/pl00006126

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Erratic Evolution of Glycerol-3-Phosphate Dehydrogenase in *Drosophila*, *Chymomyza*, and *Ceratitis*

Jan Kwiatowski,^{1,2} Michal Krawczyk,^{1,2} Michal Jaworski,^{1,2} Douglas Skarecky,¹ Francisco J. Ayala¹

¹ Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92717-2525, USA

² Institute of Botany, Warsaw University, 00-478 Warsaw, Poland

Received: 13 June 1996 / Accepted: 16 August 1996

Abstract. We have studied the evolution of *Gpdh* in 18 fruitfly species by sequencing 1,077 nucleotides per species on average. The region sequenced includes four exons coding for 277 amino acids and three variable-length introns. Phylogenies derived by a variety of methods confirm that the nominal genus *Zaprionus* belongs within the genus *Drosophila*, whereas *Scaptodrosophila* and *Chymomyza* are outside. The rate of GPDH evolution is erratic. The rate of amino acid replacements in a lineage appears to be 1.0×10^{-10} /site/year when *Drosophila* species are considered (diverged up to 55 million years ago), but becomes 2.3×10^{-10} when they are compared to *Chymomyza* species (divergence around 60 My ago), and 4.6×10^{-10} when species of those two genera are compared with the medfly *Ceratitis capitata* (divergence around 100 My ago). In order to account for these observations, the rate of amino acid replacement must have been 12 or more times greater in some lineages and at some times than in others. At the nucleotide level, however, *Gpdh* evolves in a fairly clocklike fashion.

Key words: GPDH — Homoplasy — Molecular clock — Protein evolution — Synonymous versus replacement evolution

The sequences reported in this paper have been submitted to the GenBank/EMBL data libraries under accession numbers L36960, L36961, L37038, L37039, L41248-52, L41647-50, U47808, and U47809

Abbreviations: aa = amino acid; ADH = alcohol dehydrogenase; GPDH = glycerol-3-phosphate dehydrogenase; nt = nucleotide; SOD = superoxide dismutase

Correspondence to: F.J. Ayala; e-mail fjayala@uci.edu

Introduction

The nicotinamide-adenine dinucleotide (NAD)-dependent cytoplasmic glycerol-3-phosphate dehydrogenase (GPDH, EC 1.1.1.8) plays a crucial role in insect flight metabolism through its key position in the glycerophosphate cycle, which provides energy for flight in the thoracic muscles of *Drosophila* (O'Brien and MacIntyre 1978). In *Drosophila melanogaster* the *Gpdh* gene is located on chromosome 2 (O'Brien and MacIntyre 1972) and consists of eight coding exons (Bewley et al. 1989; von Kalm et al. 1989). It produces three isozymes by differential splicing of the last three exons (Cook et al. 1988). In the mouse, two different isozymes are encoded by two separate loci (Kozak et al. 1982).

The polypeptide chain of GPDH can be divided into two main domains: one is NAD binding, the other is catalytic. In the rabbit enzyme, the former is determined by the first 118 amino acids (Otto et al. 1980). The enzyme is known to be evolutionarily conserved (Bewley et al. 1989), displaying very low heterozygosity within or variation among *Drosophila* species (Lakovaara et al. 1977). The catalytic domain seems to be less conserved (Bewley et al. 1989). Here we present analysis of a *Gpdh* gene region comprising most of the coding sequence (768 bp out of 831 bp) of exons 3–6 in 18 species of *Drosophila* and related genera, corresponding to the whole catalytic domain plus 45 codons of the NAD-binding domain. We have also sequenced the intervening introns, which jointly have an average length of 309 nucleotides per species.

Table 1. Taxonomy of the 18 fruitfly species according to Wheeler (1981)

Family	Genus	Subgenus	Group	Species
<i>Drosophilidae</i>	<i>Drosophila</i>	<i>Sophophora</i>	<i>melanogaster</i>	<i>melanogaster</i>
				<i>simulans</i>
				<i>teissieri</i>
				<i>obscura</i>
				<i>pseudoobscura</i>
				<i>miranda</i>
		<i>Drosophila</i>	<i>virilis</i>	<i>guanche</i>
				<i>willistoni</i>
				<i>paulistorum</i>
				<i>nebulosa</i>
				<i>virilis</i>
				<i>repleta</i>
<i>Dorsilopha</i>	<i>Hirtodrosophila</i>	<i>Scaptodrosophila</i> ^a	<i>busckii</i>	
			<i>pictiventris</i>	
			<i>lebanonensis</i>	
<i>Zaprionus</i>	<i>Chymomyza</i>		<i>tuberculatus</i>	
			<i>amoena</i>	
			<i>procnemis</i>	
			<i>capitata</i>	
<i>Tephritidae</i>	<i>Ceratitis</i>			

^a Raised to genus category in the revision by Grimaldi (1990)

Materials and Methods

Species. The 18 species studied are listed in Table 1. The strains of *Chymomyza amoena*, *Chymomyza procnemis*, *Zaprionus tubercultatus*, *D. pictiventris*, and *D. virilis* proceed from the National Drosophila Species Stock Center at Bowling Green, Ohio; *D. hydei* proceeds from the Stock Center at Indiana University, Bloomington. The other *Drosophila* species are cultured in our laboratory. For the source of the *Ceratitis capitata* DNA see Kwiatowski et al. (1992a).

DNA Preparation, Amplification, Cloning, and Sequencing. Genomic DNA is prepared from about 10–20 flies, following the method of Kawasaki (1990). The *Gpdh* gene fragment is amplified by PCR, cloned into the pCRII vector from the Invitrogen TA-cloning kit, and sequenced using standard methods (Ausubel et al. 1987), as previously described (Kwiatowski et al. 1994). The PCR fragments are obtained with primers for the beginning of exon 3 and the end of exon 6 (Fig. 1), derived from published sequences of *D. melanogaster* and *D. virilis*; namely L3: 5'-GTTCTAGATCTGGTTGAGGCTGCCAAGAA-3', and R6: 5'-ACATATGCTCTAGATGATTGCGTATGCA-3'. The *Gpdh* gene fragment from *Ceratitis* is obtained in two overlapping pieces; one by means of the degenerate primers EG: 5'-GARGGDAAYTTYTGYGARAC-3' and NH: 5'-TACATRTGYTCNGGRTGTT-3' (derived from the conserved protein fragments EGNFCE and NHPEHM, respectively); and the second obtained with a specific primer SP: 5'-CAGAGTCCTCGACCACAACCACCGAA-3', derived from the first gene fragment, and the primer FV: 5'-TTCGTCGRCCGCAYCARTTYAT-3', derived from the oligopeptide FVVPHQFI. In addition to these primers and to the standard M13, Uni, and Rev primers, we use five primers for sequencing. L4b: 5'-CCATGCGCGTGCTGATGGG-3'; L4e: 5'-GATCTTATCACGACGTGTTA-3'; L5: 5'-CGTGTCTCTGAGGCTTTTGT-3'; R4: 5'-ACAGCCGCCTTGGTGTGTGTCGCCCA-3'; and R5: 5'-GTGGCCGCAATCGTCGCGTTTC-3'. The DNA coding sequences for exons 3–6 of *D. busckii* and *D. lebanonensis* have been kindly provided by Dr. Spencer Wells.

Sequences were obtained from single PCR-amplified clones. We have estimated that the substitution error introduced by this procedure is 3×10^{-4} (Kwiatowski et al. 1991), which would amount to four erroneous nucleotide determinations in our whole data base, which should have negligible consequences for the analyses done in this pa-

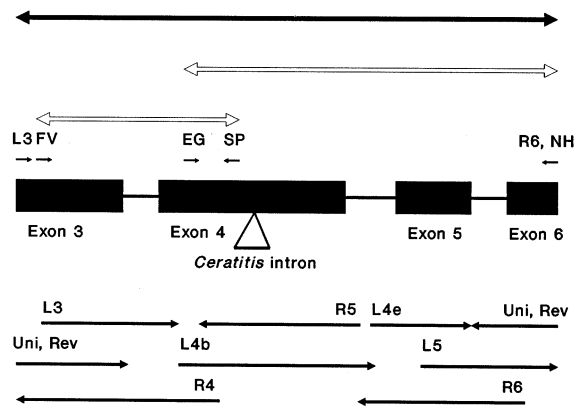


Fig. 1. Structure of the *Gpdh* gene region and strategy for amplification, cloning, and sequencing. The black boxes represent exons. The double-pointed black arrow on the top represents the 823 bp amplified, cloned, and sequenced in the Drosophilid species; the two hollow arrows just below it represent the two fragments amplified in *Ceratitis capitata*. The short thin arrows on top of the exons represent the amplification primers. The arrows at bottom indicate the extent and direction of sequencing. A triangle indicates the position of an extra intron found in *Ceratitis capitata*.

per. Moreover, all nonsilent substitutions were confirmed by sequencing second clones obtained by separate amplifications.

Sequence Analysis. Corrections for superimposed and back replacements (PAM; Dayhoff 1978) were made with the CLUSTAL V program (Higgins et al. 1992). Kimura two-parameter distances (K2) (Kimura 1980) and Jukes-Cantor distances (Jukes and Cantor 1969) were estimated with the MEGA 1.0 program (Kumar et al. 1993). Synonymous and nonsynonymous substitutions were calculated according to Li (1993) as implemented in programs written by A. Zharkikh. Neighbor-joining trees (Saitou and Nei 1987) were obtained with the MEGA program. The maximum-parsimony (Fitch 1971) and maximum-likelihood (Felsenstein 1981) trees were obtained and tested according to Templeton (1983) and Kishino and Hasegawa (1989), respectively, using the PHYLIP 3.57c package programs DNAPARS and DNAML (Felsenstein 1989).

Results

Gene Structure and G+C Content

The *Gpdh* gene in *Drosophila* consists of eight exons with a total 1,100 nt of coding sequence. We have amplified a segment that extends from the seventh base of exon 3 to the third base from the end of exon 6, accounting for 823 of the 831 coding nucleotides (Fig. 1). We have also sequenced the intervening introns, amounting to 309 nt per species on the average. In *Drosophila* the lengths of exons 3 to 6 are, successively, 207, 373, 154, and 97 bp (Bewley et al. 1989; Tominaga et al. 1992; von Kalm et al. 1989). The PCR fragments have the same structure and identical exon length for the 17 drosophilid species in our sample. *Ceratitis capitata* has an additional intron that splits the coding sequence corresponding to exon 4 in the drosophilid gene. We have also found an intron at the homologous site in another dipteran species, *Calliphora*, as well as in the mouse, which suggests that this additional intron is the ancestral condition that has been lost in the family Drosophilidae.

Figure 2 shows 768 bp of coding sequence (256 aa) for the Drosophilidae species and 729 bp for *C. capitata* (shorter owing to the different location of the 3'-primers L3 and FV, in Fig. 1). Table 2 gives the intron lengths, which are mostly short (between 54 and 83 bp), conforming to typical *Drosophila* intron sizes (Hawkins 1988; Mount et al. 1992), except for intron 4, which is substantially longer in the three species of the *D. melanogaster* group and in *Ceratitis*, and also, but not as long, in the three *willistoni* group species.

The G+C content of the *Gpdh* coding sequence is close to 50% in most species, but somewhat higher (55.8–56.7%) in the three *melanogaster*-group species (Table 3), which have particularly high G+C content in the third codon position (76.1–78.9%). All other species exhibit higher G+C content in the third than in the other two codon positions, but *Ceratitis* and the two *Chymomyza* species (as well as *D. nebulosa*) remain close to 50% in the third sites. Similar patterns have been observed in other *Drosophila* genes such as *Adh* (Starmer and Sullivan 1989) and *Sod* (Kwiatowski et al. 1992b), in which the species of the *D. melanogaster* group exhibit great G+C excess relative to other *Drosophila* species, and particularly relative to *Chymomyza* and *Ceratitis* (Kwiatowski et al. 1992b). The excess G+C content of *Sod* is, as in *Gpdh*, greater in the third than in the other two coding positions and particularly so in the *melanogaster*-group species (Kwiatowski et al. 1994).

Table 3 shows the G+C content for each intron and exon. *Drosophila* introns have on average 17% more A+T content than adjacent exons (Csank et al. 1990; Mount et al. 1992); in the case of *Sod* the difference reaches 35% (Kwiatowski et al. 1992b). In the case of *Gpdh*, the A+T content of intron 3 is not very different

from exon 3 (except in *Ceratitis*) and intron 5 is not very different from exon 6 (except in *Chymomyza* and *Ceratitis*), but exon 4 and exon 5 have on average much lower A+T content than their adjacent introns.

The different patterns of G+C content observed in *Gpdh* in *Ceratitis* and *Chymomyza* relative to *Drosophila* (and *Zaprionus*) and in the three *melanogaster*-group species relative to the other *Drosophila* species is consistent with earlier findings for other genes that codon use and G+C content have a strong species-specific component in fruit flies (Grantham et al. 1980; Kwiatowski et al. 1992b; Sharp et al. 1988; Starmer and Sullivan 1989).

Gpdh Evolution and Phylogeny

Figure 2 displays the 768 coding nucleotides sequenced in the 18 species (729 in *Ceratitis*). Table 4 gives the number of nucleotide differences between species pairs based on the 729 bp sequenced in all 18 species (lower triangle) as well as the number of differences between the inferred amino acid sequences (upper triangle). We have also calculated genetic distances (not shown) between species following the method of Jukes and Cantor (1969) using the complete data set displayed in Fig. 2. A tree based on the Jukes-Cantor distances obtained by the neighbor-joining (NJ) method (Saitou and Nei 1987) is displayed in Fig. 3. A tree based on Kimura's two-parameter distance (see below) has identical topology and virtually identical bootstrap values. These trees show the well-established monophyly (100% bootstrap values) of each of the *melanogaster*, *obscura*, and *willistoni* species groups. These three groups are classified within the *Sophophora* subgenus, although the *willistoni* group diverged from the others shortly after the origination of the *Sophophora* clade (Kwiatowski et al. 1994). In Fig. 3, the *D. willistoni* group is less closely related to the *melanogaster* and *obscura* clades than to other non-*Sophophora* species, but this relationship is not statistically valid (28% bootstrap value).

The two *Chymomyza* species and *D. lebanonensis* are outside the clade that includes the other *Drosophila* species and *Zaprionus* (70% bootstrap), consistent with results obtained with other genes and supporting the classification of *lebanonensis* (and other *Scaptodrosophila*) as a new genus (Grimaldi 1990; Kwiatowski et al. 1994). *D. pictiventris* (subgenus *Hirtodrosophila*) and *Zaprionus tuberculatus* are more closely related to species of the *Drosophila* subgenus than to the *Sophophora* subgenus, consistent with results obtained with the *Sod* gene (Kwiatowski et al. 1994), although in the present case, the relationships are not statistically robust.

The phylogeny represented in Fig. 3 leaves unresolved important relationships, such as whether *Chymomyza* and *Scaptodrosophila* (*D. lebanonensis*) are sister clades, and the phylogenetic relationships between *Zaprionus* and the subgenera *Drosophila*, *Dorsilopa*, and

D. melanogaster	AACGCTGACA	TCCTGATCTT	CGTGGTGGCC	CACCAGTTCA	TTCCCAACTT	CTGCAACAG	CTACTGGGCA	AGATTAAGCC	AAATGCCATC	GCTATCTCCC	100
D. simulans	.T.C.	.T.	.T.	.T.	.T.	.T.	.T.	.C.	.C.	.C.	
D. teissieri	.C.	.C.	.C.	.C.	.C.	.C.	.C.	.C.	G.C.	.C.	
D. pseudoobscura	.T.T.A.C.T.	.A.A.T.	.A.A.T.	.T.	.T.	T.G.AT.G.	T.A.C.	.T.	.C.A.	TT	
D. miranda	.T.T.A.C.T.	.A.A.T.	.A.A.T.	.T.	.T.	T.G.AT.G.	T.A.C.	.T.	.C.A.	TT	
D. guanche	.T.T.A.C.T.	.A.A.T.	.A.A.T.	.T.	.T.	G.AT.G.	T.A.C.	.T.	.C.A.	TT	
D. willistoni	.T.T.TT.A.T.	.T.A.	.T.A.T.	.A.	.T.	.AT.G.	.T.	.C.	.T.	C.A.	
D. paulistorum	.T.T.TT.A.T.	.T.A.	.T.A.T.	.A.	.T.	.AT.G.	.T.	.C.	.T.	C.A.	
D. nebulosa	.T.C.T.TT.	.T.A.	.T.A.T.	.A.	.T.	.T.G.	.T.A.	.T.	.T.	C.A.T	
D. virilis	.T.C.T.A.A.	.T.	.A.T.A.T.	.A.A.T.	.G.A.T.G.	.C.	.T.	.T.	.C.A.		
D. hydei	.T.C.T.A.C.T.	.A.A.	.T.A.T.	.ATA.T.	.G.	.T.G.	.A.C.	.T.	.C.A.		
D. busckii	.C.T.AT.	.T.T.	.A.T.	.A.A.T.	.AT.GT.	.A.C.	.C.	.T.	.C.A.A.		
D. pictiventris	.T.C.T.A.C.T.	.G.T.	.T.A.A.T.	.AT.G.	.T.	.C.	.C.	.A.	.C.A.	AT	
D. lebanonensis	.C.T.A.C.T.	.T.	.A.T.	.T.A.G.T.	.AT.GT.	.A.A.	.C.	.T.	.C.A.A.		
Zaprionus	.T.C.T.T.C.T.	.T.T.	.T.C.G.	.T.	.AT.GT.	.T.	.C.	.T.	.C.A.A.	T.	
C. amoena	.C.TG.AT.	.T.	.C.G.T.	.T.A.A.T.	.AT.GT.	.T.A.	.C.	.T.	.C.A.A.	AT	
C. procnemis	.T.C.TG.AT.	.T.	.T.G.T.	.T.A.A.T.	.AT.G.	.T.A.	.C.	.T.	.C.A.A.	AT	
Ceratitidis	.C.G.	.C.G.	.C.G.	.C.G.	.ACT.TT.	.A.A.C.	.AG.G.	.C.T.G.			
D. melanogaster	TGATCAAGGG	CTTCGACAAG	GCCGAGGGCG	GCGGCATCGA	TCTGATCTCG	CATATAATCA	CGCGTCACCT	GAAGATCCCA	TGCGCCGTGT	TGATGGGGCC	200
D. simulans	.C.	.C.	.C.	.C.	.C.	.C.	.C.	.C.	.C.	.C.	
D. teissieri	.C.	.C.	.C.	.C.	.C.	.C.	.C.	.C.	.C.	.C.	
D. pseudoobscura	.T.	.A.A.T.	.T.	.C.T.	.C.	.T.	.A.	.T.C.	.C.	.C.	
D. miranda	.T.	.A.A.T.	.T.	.C.T.	.C.	.T.	.A.	.T.C.	.C.	.C.	
D. guanche	.T.	.A.A.T.	.T.	.C.T.	.C.	.T.	.A.	.T.T.C.	.C.	.C.	
D. willistoni	.T.	.A.	.T.	.T.T.	.C.T.	.C.	.T.	.A.	.AC	.C.	
D. paulistorum	.T.	.A.	.T.	.T.T.	.C.T.	.C.	.T.	.A.	.AC	.C.	
D. nebulosa	.T.	.T.A.	.T.	.A.T.	.C.T.	.C.	.T.	.A.	.T.A.	.C.	
D. virilis	.T.	.T.	.T.	.A.	.C.T.T.	.C.C.TT.	.A.	.T.C.	.C.	.C.	
D. hydei	.T.	.T.	.T.	.T.T.C.	.C.	.C.T.	.C.C.T.	.A.	.T.C.	.C.	
D. busckii	.T.	.T.A.T.	.T.	.A.T.	.C.T.T.	.C.C.T.	.A.	.T.C.	.C.	.C.	
D. pictiventris	.T.T.	.A.	.T.	.T.T.	.C.T.	.C.C.TT.	.A.A.C.	.TC	.C.A.	.C.	
D. lebanonensis	.A.	.T.	.T.	.A.T.T.	.C.T.T.	.C.TT.	.A.A.	.T.C.	.A.	.C.	
Zaprionus	.T.	.T.	.T.	.A.T.	.C.T.	.C.C.T.	.A.	.T.C.	.C.A.	.C.	
C. amoena	.A.T.T.	.A.	.T.T.	.A.T.C.	.C.T.T.	.T.A.	.A.	.TT.	.C.	.T.	
C. procnemis	.A.T.T.	.A.	.T.T.	.A.T.C.	.C.T.T.	.T.A.	.A.	.TT.	.C.	.T.	
Ceratitidis	.C.T.	.T.	.A.T.	.T.T.	.T.	.C.C.T.	.C.C.TT.	.T.C.	.T.	.A.	
D. melanogaster	AAACCTGGCC	AACGAGTGG	CTGAGGGCAA	CTTCTGCGAA	ACGACAATCG	GCTGCACGGA	CAAGAAGTAT	GGCAAGGTGC	TGCGCGATCT	CTTCAGGGC	300
D. simulans	.C.	.C.	.C.	.C.	.C.	.C.	.C.	.C.	.C.	.C.	
D. teissieri	.C.	.A.	.C.	.C.	.C.	.C.	.C.	.C.	.C.	.A.	
D. pseudoobscura	.C.T.	.T.A.	.C.	.T.G.	.C.T.	.T.	.A.	.T.	.C.	.C.	
D. miranda	.C.T.	.T.A.C.	.C.	.T.G.	.C.T.	.T.	.A.	.T.	.C.	.C.	
D. guanche	.C.TT.	.T.A.	.C.	.T.G.	.C.T.	.T.	.A.	.T.	.C.	.C.	
D. willistoni	.C.T.	.T.	.C.A.T.	.T.	.T.C.	.T.T.T.	.T.A.	.A.	.T.	.C.	
D. paulistorum	.C.T.	.T.	.C.A.T.	.T.	.T.C.	.T.T.T.	.T.A.C.	.A.	.T.	.C.	
D. nebulosa	.C.T.	.T.	.C.A.A.T.	.T.	.G.	.T.T.T.	.T.A.	.A.	.T.	.A.C.	
D. virilis	.T.T.	.T.	.C.A.A.T.	.T.	.G.	.T.T.T.	.T.A.	.A.	.T.	.A.C.	
D. hydei	.T.T.	.T.	.C.	.T.G.	.C.	.A.C.	.A.C.A.	.T.	.T.	.C.	
D. busckii	.C.T.	.T.	.C.	.T.T.	.C.C.T.	.A.	.A.C.A.	.T.	.T.	.C.	
D. pictiventris	.C.T.	.A.	.C.	.T.G.	.A.	.T.T.	.A.C.	.A.	.T.	.C.	
D. lebanonensis	.C.T.	.T.A.	.C.A.A.T.	.T.G.	.A.	.T.AC.	.T.A.	.A.TT.	.T.	.A.	
Zaprionus	.T.T.	.T.	.C.A.A.	.T.G.	.A.	.T.A.	.A.C.	.A.TT.	.T.	.A.	
C. amoena	.C.T.A.T.	.A.	.C.A.A.	.T.T.G.	.A.T.	.T.GT.	.A.C.	.AT	.T.	.A.A.	
C. procnemis	.C.TT.	.T.A.	.C.	.T.T.	.A.T.	.T.AT.	.A.C.	.CA.	.AT	.T.A.A.	
Ceratitidis	.C.TT.	.A.	.ATT.	.T.T.	.T.	.CGC.	.TA.	.AC.	.CCAT.	.T.A.A.	
D. melanogaster	AACCACTTCC	GCGTGGTGGT	CGTGGACGAT	GCTGATGCCG	TGGAGGTTTG	CGGTGCCCTC	AAGAACATCG	TGGCCTCGGG	CGTGGCTTT	GTTGATGGCC	400
D. simulans	.C.	.C.	.C.	.C.	.C.	.C.	.C.	.C.	.C.	.C.	
D. teissieri	.C.	.C.	.C.	.C.	.C.	.C.	.C.	.C.	.C.	.C.	
D. pseudoobscura	.T.T.	.C.T.	.T.	.C.C.	.G.	.T.C.T.G.	.T.T.	.G.	.T.G.	.G.	
D. miranda	.T.T.	.C.T.	.T.	.C.C.	.G.	.T.C.T.G.	.T.T.	.G.	.T.G.	.G.	
D. guanche	.T.T.	.C.	.A.	.C.	.C.T.	.T.	.T.	.T.G.	.C.	.C.	
D. willistoni	.T.T.	.T.T.	.C.G.	.C.	.T.	.T.	.T.T.	.T.	.T.C.	.T.	
D. paulistorum	.T.T.	.T.T.	.C.G.	.C.	.T.	.T.	.T.T.	.T.	.T.C.	.T.	
D. nebulosa	.T.T.	.T.T.	.T.C.G.	.C.	.T.	.T.	.T.T.	.T.	.T.C.	.T.	
D. virilis	.T.T.T.	.T.C.	.T.G.	.C.	.C.	.G.	.T.	.T.T.	.T.G.	.C.T.	
D. hydei	.T.T.	.C.	.C.G.	.C.	.T.	.C.	.TT.G.	.G.T.	.C.	.G.	
D. busckii	.T.T.	.T.C.T.	.T.	.CT.C.	.T.	.T.	.A.T.	.T.	.T.	.G.	
D. pictiventris	.T.C.T.	.T.T.T.	.C.	.T.A.	.G.	.A.A.G.	.A.T.T.	.A.	.T.	.G.T.	
D. lebanonensis	.T.T.	.T.	.T.C.GG.	.C.	.G.	.C.	.T.	.T.C.	.T.G.	.G.	
Zaprionus	.T.T.	.C.T.	.T.	.T.C.	.T.A.A.	.C.	.T.	.T.C.	.T.G.	.G.	
C. amoena	.T.T.T.	.T.T.G.	.T.C.	.T.	.T.	.C.T.G.	.A.	.T.A.A.T.	.T.G.	.C.T.	
C. procnemis	.T.T.T.	.T.T.G.	.T.A.	.T.	.T.	.T.T.G.	.A.	.T.A.A.T.	.T.G.	.C.T.	
Ceratitidis	.CCAT.	.T.T.G.C.G.C.T.	.T.A.	.T.G.	.T.TAC.	.T.C.	.T.C.	.T.C.	.T.	.T.	

Fig. 2. Nucleotide sequence of the *Gpdh* coding region sequenced in 18 species. The sequences show the last two nucleotides of the 5' amplification primer, start at nt 34 of the third exon, and end immediately before the 3' amplification primer. Dots indicate identical nucleotides to those of *D. melanogaster*. The sequences of *D. busckii* and *D. lebanonensis* were generously provided by Dr. Spencer Wells.

Hirtodrosophila. These relationships remain unresolved when genetic distances are estimated using a variety of additional methods, such as Kimura's two-parameter distance (Kimura 1980; see Table 5, lower triangle), Tamura and Nei's distance (1993), distances based only on transversions, only on transitions, or both, and when the phylogenies are reconstructed with methods other than NJ, such as maximum parsimony (Fitch 1971) or maximum likelihood (Felsenstein 1981).

The transition/transversion ratios for all interspecies

comparisons are 1.7–2.0 between the three species of the *melanogaster* group, which diverged 5 My ago, and about the same for most other comparisons between *Drosophila* species, corresponding to divergence times between 40 and 60 My. It is somewhat lower, about 1.5, between *Ceratitidis* and the rest of the species, corresponding to 100 My divergence (Fitch and Ayala 1994).

We have used two statistical tests seeking to resolve phylogenetic relationships left unsettled by the bootstrap results: the maximum-parsimony test of Templeton

500

D. melanogaster	TCAAGCTGGG	CGACAACACC	AAGGCGGCTG	TCATCCGGCT	GGGCCTCATG	GAGATGATTC	GCTTCGTGGA	CGTCTTCTAT	CCCGGAAGCA	AGCTGTCCAC
D. simulans			A							C
D. teissleri	.G		A							C
D. pseudoobscura	.G			T	AT	T	G	A		T
D. miranda	.G			T	AT	T	G	A		T
D. guanche	.G	A		A	T	AT	T	G	A	
D. willistoni				T	T		T	G	A	A
D. paulistorum				T	T		T	G	A	A
D. nebulosa	.AT			C	T	T	T	G	A	A
D. virilis				A	G		T	TT		A
D. hydei	.T			A	G		T	T	A	T
D. busckii				A	G		T	T	A	T
D. pictiventris	.G	T								
D. lebanonensis	.T	T								
Zaprionus	.G									
C. amoena	.T	AT								
C. procnemis	.A	A								
Ceratitis	.GGGCT									

600

D. melanogaster	CTTCTTCGAG	AGCTGCGGCG	TGGCCGATCT	GATCAGCAGC	TGTTACGGTG	GCCGTAACCG	CCGCGTCTCT	GAGGCGTTTG	TGACTTCTGG	AAAACATT
D. simulans										
D. teissleri			T							
D. pseudoobscura			T	G	C	T				
D. miranda			T	G	C	T				
D. guanche			T	G	C	T				
D. willistoni			T	C	G	T				
D. paulistorum			T	C	G	T				
D. nebulosa			T	T	T					
D. virilis	A		T	T	T					
D. hydei	T		T	T	T					
D. busckii	G									
D. pictiventris			T	A	T	T				
D. lebanonensis	A									
Zaprionus			T	T	T					
C. amoena	A		T	T	T					
C. procnemis	A									
Ceratitis	A									

700

D. melanogaster	GAGGAGCTGG	AGAAGGAAAT	GCTCAACGGC	CAGAAGCTGC	AGGGCCCACC	CACTGCCGAG	GAGGTCAACT	ACATGTGAA	GAATAAGGGT	CTGGAGGCA
D. simulans										
D. teissleri	.A									
D. pseudoobscura	.A	A	C	A						
D. miranda	.A	A	C	A						
D. guanche	.A	A	C	A						
D. willistoni	.A	AT								
D. paulistorum	.A	A								
D. nebulosa	.A	T								
D. virilis	.A	T								
D. hydei										
D. busckii										
D. pictiventris	.A	T								
D. lebanonensis	.AT									
Zaprionus										
C. amoena	.A	C	A							
C. procnemis	.A	T								
Ceratitis	.AT									

768

D. melanogaster	AATTCCCCCT	GTTTCACGGCT	ATTCAACAAA	TATGCACAAA	TCAGCTCAAG	CCTAATGATT	TAATGATG			
D. simulans										
D. teissleri										
D. pseudoobscura										
D. miranda										
D. guanche										
D. willistoni										
D. paulistorum										
D. nebulosa										
D. virilis										
D. hydei										
D. busckii	.G	T	AT							
D. pictiventris										
D. lebanonensis	.G									
Zaprionus										
C. amoena	.G	GT								
C. procnemis	.G	GT								
Ceratitis										

Fig. 2. Continued.

(1983) and the maximum likelihood of Kishino and Hasegawa (1989), which compare phylogenies as wholes with one another. The alternative phylogenies examined are displayed in Fig. 4 and the results obtained with Templeton's method are given in Table 6; *Ceratitis* is not included since it is unambiguously an outgroup relative to the drosophilids. (The results obtained with the method of Kishino and Hasegawa are similar, except that fewer statistically different phylogenies are detected.) Tree 1 (Fig. 4) requires the fewest steps but, on the basis of the *Gpdh* data alone, is not statistically better than phylogenies 2–5. If we combine the data for *Gpdh* and

Sod (Kwiatowski et al. 1994), tree 1 is again the best, but statistically no better than phylogenies 2–4. When the data for *Adh* are added, however, tree 1 is statistically better than any other phylogeny (Table 6; the *Adh* sequence is not available for *D. busckii*). The topology of tree 1 clusters the subgenera *Dorsilopha* (*D. busckii*) and *Hirtodrosophila* (*D. pictiventris*) with each other and then successively with the *Drosophila* subgenus and *Zaprionus*. This phylogeny differs from phylogeny 2, the one proposed by Grimaldi (1990; see also DeSalle and Grimaldi 1991), based on cladistic analysis of morphological information, in that Grimaldi

Table 2. Lengths of *Gpdh* introns in 16 fruitfly species^a

Species	Intron site			
	3	3a	4	5
<i>D. melanogaster</i>	65		329	68
<i>D. simulans</i>	65		328	68
<i>D. teissieri</i>	65		334	64
<i>D. pseudoobscura</i>	81		67	74
<i>D. miranda</i>	80		65	74
<i>D. guanache</i>	80		67	64
<i>D. willistoni</i>	75		167	65
<i>D. paulistorum</i>	79		171	59
<i>D. nebulosa</i>	65		212	72
<i>D. virilis</i>	70		67	67
<i>D. hydei</i>	65		55	70
<i>D. pictiventris</i>	72		63	75
<i>Zaprionus tuberculatus</i>	66		59	67
<i>Chymomyza amoena</i>	58		65	63
<i>Chymomyza procnemis</i>	59		54	62
<i>Ceratitidis capitata</i>	83	75	353	75

^a The intron sequences of *D. busckii* and *D. lebanonensis* are not available

places *D. pictiventris* outside a cluster consisting of the genera *Drosophila* (except *D. lebanonensis*, which is appropriately classified in a different genus, *Scaptodrosophila*) and *Zaprionus*; and on this basis, he has raised *Hirtodrosophila* to the genus category. Moreover, Grimaldi considers the *Sophophora* subgenus more closely related to the subgenera *Drosophila* and *Dorsilopha* than to *Zaprionus* (see tree 2), whereas tree 1 shows that *Sophophora* is outside the cluster (((*Dorsilopha* + *Hirtodrosophila*) *Drosophila*) *Zaprionus*). Based

on mitochondrial DNA sequence data, DeSalle (1992) has proposed phylogeny 7, which places *Hirtodrosophila* outside the cluster of the three genera *Drosophila*, *Zaprionus*, and *Chymomyza*, a hypothesis that is statistically rejected by the *Gpdh* data alone, or combined with *Sod*, or with *Sod* and *Adh*. Our results are also inconsistent with Throckmorton's (1975) proposal that *Chymomyza* is associated with the *Sophophora* radiation, but support his claim that the *Sophophora* radiation preceded the divergence of *Zaprionus* and the subgenera *Drosophila*, *Dorsilopha*, and *Hirtodrosophila*.

Rate of Evolution

The number of amino acid and nucleotide differences between species is given in Table 4. Values corrected for superimposed and back substitutions are given in Table 5. It is apparent that the rate of amino acid replacements is not uniform over time. Figure 5 gives the polymorphic amino acid sites and Fig. 6 (top left) displays the number of (corrected) amino acid replacements against time. Average values and times of divergence are given in Table 7. The number of amino acid replacements is zero between species of the same group, in all three species groups: *melanogaster*, *obscura*, and *willistoni*. The divergence time between species of the *melanogaster* group is 5–8 My, but it is 20–25 My between *guanache* and the other two *obscura* species or between *nebulosa* and the two other *willistoni* group species. The number of amino acid replacements for most other interspecific comparisons between species of the *Drosophila* genus is between zero and four, corresponding to divergence

Table 3. G + C content (%) in the *Gpdh* gene of dipteran species^a

Species	Gene region							Coding sequence	
	Exon 3	Intron 3	Exon 4	Intron 4	Exon 5	Intron 5	Exon 6	All sites	Third positions
<i>D. melanogaster</i>	53.4	50.8	60.0	34.9	55.8	35.3	37.3	55.8	76.1
<i>D. simulans</i>	53.4	49.2	60.9	26.8	56.5	38.2	38.8	56.4	77.7
<i>D. teissieri</i>	56.3	49.2	60.9	33.9	55.2	32.8	38.8	56.7	78.9
<i>D. pseudoobscura</i>	42.5	44.4	54.1	35.8	47.4	36.5	37.3	48.7	58.6
<i>D. miranda</i>	42.5	45.0	54.7	38.5	47.4	39.2	37.3	48.9	59.3
<i>D. guanache</i>	43.1	37.5	52.2	37.3	47.4	29.7	35.8	47.8	56.3
<i>D. willistoni</i>	42.0	32.0	51.3	28.7	45.5	27.7	35.8	46.6	52.4
<i>D. paulistorum</i>	43.1	36.7	51.7	29.8	46.1	33.9	34.3	47.2	53.5
<i>D. nebulosa</i>	42.6	33.8	49.3	26.4	44.8	29.2	29.8	45.2	48.8
<i>D. virilis</i>	45.4	34.3	53.6	43.3	49.4	26.9	35.8	49.3	59.4
<i>D. hydei</i>	45.9	44.6	55.0	30.9	48.0	32.9	31.4	49.5	59.3
<i>D. busckii</i>	43.1	—	53.9	—	50.7	—	32.8	48.9	59.4
<i>D. pictiventris</i>	44.3	31.9	53.3	30.2	46.8	28.0	35.9	48.4	56.3
<i>D. lebanonensis</i>	41.9	—	51.2	—	50.6	—	44.7	48.4	57.9
<i>Z. tuberculatus</i>	46.6	37.9	53.1	32.2	51.9	37.3	37.3	50.0	60.9
<i>C. amoena</i>	41.4	39.7	48.0	27.7	46.1	17.5	43.3	45.7	50.8
<i>C. procnemis</i>	41.3	44.1	46.1	20.4	44.2	24.2	40.3	44.1	46.9
<i>Ceratitidis</i>	45.2	25.3	49.1	20.7	48.1	20.0	37.3	47.0	53.1

^a The sequences used for the calculations are complete for all introns and exons 4 and 5; for exons 3 and 6, we use the last 174 (135 in *Ceratitidis*) and the first 67 nucleotides, respectively. The intron data for *D. busckii* and *D. lebanonensis* are not available

Table 4. Number of amino acid (above the diagonal) and nucleotide differences (below the diagonal) between 18 species in 243 codons of the *Gpdh* gene

Species	1	2	3	4	5	6	7	8	9
1 <i>D. melanogaster</i>	—	0	0	2	2	2	2	2	2
2 <i>D. simulans</i>	9	—	0	2	2	2	2	2	2
3 <i>D. teissieri</i>	27	19	—	2	2	2	2	2	2
4 <i>D. pseudoobscura</i>	106	106	114	—	0	0	2	2	2
5 <i>D. miranda</i>	105	105	113	3	—	0	2	2	2
6 <i>D. guanche</i>	104	106	117	34	37	—	2	2	2
7 <i>D. willistoni</i>	108	107	115	99	100	102	—	0	0
8 <i>D. paulistorum</i>	107	106	114	101	102	103	11	—	0
9 <i>D. nebulosa</i>	117	118	121	108	109	103	42	45	—
10 <i>D. virilis</i>	114	110	120	99	100	96	95	90	104
11 <i>D. hydei</i>	112	108	116	96	95	100	103	102	109
12 <i>D. busckii</i>	106	101	111	109	110	106	90	88	96
13 <i>D. pictiventris</i>	124	121	123	107	108	114	94	95	108
14 <i>D. lebanonensis</i>	128	126	132	117	120	117	106	105	113
15 <i>Z. tuberculatus</i>	108	106	112	102	103	102	98	99	107
16 <i>C. amoena</i>	155	155	157	133	134	130	129	128	128
17 <i>C. procnemis</i>	149	149	153	131	132	132	124	123	125
18 <i>Ceratitis</i>	165	164	175	172	171	170	156	160	165

Table 4. Continued

Species	10	11	12	13	14	15	16	17	18
1 <i>D. melanogaster</i>	3	2	2	0	4	3	9	8	21
2 <i>D. simulans</i>	3	2	2	0	4	3	9	8	21
3 <i>D. teissieri</i>	3	2	2	0	4	3	9	8	21
4 <i>D. pseudoobscura</i>	3	2	4	2	4	3	9	8	22
5 <i>D. miranda</i>	3	2	4	2	4	3	9	8	22
6 <i>D. guanche</i>	3	2	4	2	4	3	9	8	22
7 <i>D. willistoni</i>	3	2	4	2	5	5	9	8	21
8 <i>D. paulistorum</i>	3	2	4	2	5	5	9	8	21
9 <i>D. nebulosa</i>	3	2	4	2	5	5	9	8	21
10 <i>D. virilis</i>	—	1	5	3	4	4	8	7	21
11 <i>D. hydei</i>	67	—	4	2	3	3	7	6	20
12 <i>D. busckii</i>	79	87	—	2	5	3	8	7	20
13 <i>D. pictiventris</i>	93	93	96	—	4	3	9	8	21
14 <i>D. lebanonensis</i>	115	111	111	118	—	5	9	7	21
15 <i>Z. tuberculatus</i>	76	79	91	94	117	—	6	5	20
16 <i>C. amoena</i>	122	137	125	121	121	129	—	2	20
17 <i>C. procnemis</i>	129	134	120	125	122	129	70	—	19
18 <i>Ceratitis</i>	163	165	157	171	163	161	181	162	—

times about 42–55 My. These differences suggest a slow rate of protein evolution. But, as is apparent in Tables 5 and 7 and Fig. 6 (top left), the number of replacements between *Drosophila* and *Chymomyza* is about three times as large as between the *Drosophila* subgenera, although the time elapsed is only slightly greater, 60–65 My. The discordance of evolutionary rates is also notable when comparisons are made between *Ceratitis* and *Drosophila* species, which diverged about 100 My ago and differ by seven to ten times more replacements than the *Drosophila* species. We have drawn in Fig. 6 (top left) three lines that correspond to the rate of amino acid replacement between *Drosophila* species (2.1×10^{-10} replacements/site/year), between *Chymomyza* and *Drosophila* (5.2×10^{-10}), and between *Ceratitis* and all other

species (9.1×10^{-10}). (The best-fit regression equation for the comparisons between *Drosophila* species is $0.004 + 2.75 \times 10^{-10}$ replacements/site/year, with a fit of $r^2 = 0.43$. These rates are all replacements between species; the replacement rates per lineage will be one-half.)

The rate of nucleotide substitutions is, nevertheless, approximately constant over time as can be seen by inspecting Table 5 (lower triangle), where Kimura's two-parameter distances are given, and Fig. 6 (lower left), where Kimura's distances are plotted against time. (The regression equation is $0.015 + 2.71 \times 10^{-9}$ substitutions/site/year; $r^2 = 0.85$.) We have used the method of Li (1993) to estimate K_a and K_s , the number of nonsynonymous (amino acid) and synonymous substitutions, respectively. The results plotted against time are displayed

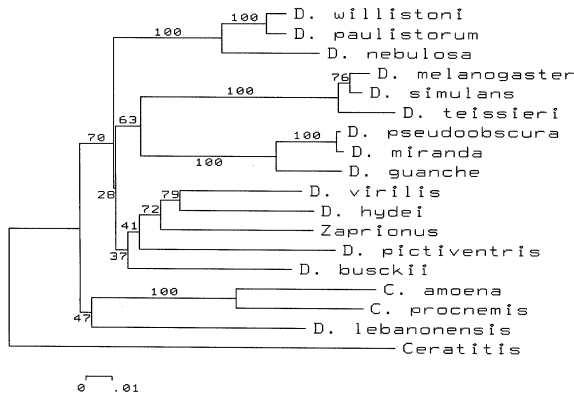


Fig. 3. Phylogeny of 18 species obtained by the neighbor-joining method (Saitou and Nei 1987) using Jukes-Cantor (1969) distances for the sequences given in Fig. 2. Numbers above the branches are percent bootstrap values based on 1,000 replications.

in Fig. 6 (right-hand panels). The rate of amino acid substitutions K_a is clearly not constant over time; indeed, it seems to be increasing exponentially over time although this perception emanates primarily from the large K_a values between *Ceratitits* and the other species. The average rate of K_a substitutions is, however, more nearly constant over time, although with large variance. (The regression equation is $-0.12 + 19.6 \times 10^{-9}$ synonymous substitutions/site/year; $r^2 = 0.67$.)

Discussion

Throckmorton (1975), based primarily on biogeographical and morphological considerations, proposed that the subgenera *Histodrosophila* (here represented by *D. pic-*

Table 5. Corrected amino acid distances (PAM) (above the diagonal) and Kimura two-parameter nucleotide distances (below diagonal) for 243 codons of the *Gpdh* gene; all values $\times 1,000$

Species	1	2	3	4	5	6	7	8	9
1 <i>D. melanogaster</i>	—	0	0	8.3	8.3	8.3	8.3	8.3	8.3
2 <i>D. simulans</i>	12	—	0	8.3	8.3	8.3	8.3	8.3	8.3
3 <i>D. teissieri</i>	38	27	—	8.3	8.3	8.3	8.3	8.3	8.3
4 <i>D. pseudoobscura</i>	165	166	180	—	0	0	8.3	8.3	8.3
5 <i>D. miranda</i>	163	164	177	4	—	0	8.3	8.3	8.3
6 <i>D. guanche</i>	162	166	185	48	53	—	8.3	8.3	8.3
7 <i>D. willistoni</i>	168	166	180	152	153	157	—	0	0
8 <i>D. paulistorum</i>	166	164	178	155	157	159	15	—	0
9 <i>D. nebulosa</i>	185	187	192	168	170	159	60	65	—
10 <i>D. virilis</i>	179	171	189	152	153	146	146	137	162
11 <i>D. hydei</i>	175	168	182	147	145	154	160	158	171
12 <i>D. busckii</i>	164	155	173	170	171	164	137	134	148
13 <i>D. pictiventris</i>	197	191	195	166	168	179	143	145	168
14 <i>D. lebanonensis</i>	203	200	211	182	188	182	164	163	178
15 <i>Z. tuberculatus</i>	167	164	175	158	159	158	150	152	167
16 <i>C. amoena</i>	256	257	261	212	214	207	205	203	203
17 <i>C. procnemis</i>	243	244	252	208	210	210	195	194	197
18 <i>Ceratitits</i>	274	272	296	287	284	283	255	263	274

Table 5. Continued

Species	10	11	12	13	14	15	16	17	18
1 <i>D. melanoagaster</i>	12.5	8.3	8.3	0	16.7	12.5	38.0	33.7	92.0
2 <i>D. simulans</i>	12.5	8.3	8.3	0	16.7	12.5	38.0	33.7	92.0
3 <i>D. teissieri</i>	12.5	8.3	8.3	0	16.7	12.5	38.0	33.7	92.0
4 <i>D. pseudoobscura</i>	12.5	8.3	16.7	8.3	16.7	12.5	38.0	33.7	96.7
5 <i>D. miranda</i>	12.5	8.3	16.7	8.3	16.7	12.5	38.0	33.7	96.7
6 <i>D. guanche</i>	12.5	8.3	16.7	8.3	16.7	12.5	38.0	33.7	96.7
7 <i>D. willistoni</i>	12.5	8.3	16.7	8.3	20.9	20.9	38.0	33.7	92.0
8 <i>D. paulistorum</i>	12.5	8.3	16.7	8.3	20.9	20.9	38.0	33.7	92.0
9 <i>D. nebulosa</i>	12.5	8.3	16.7	8.3	20.9	20.9	38.0	33.7	92.0
10 <i>D. virilis</i>	—	4.1	20.9	12.5	16.7	16.7	33.7	29.4	92.0
11 <i>D. hydei</i>	99	—	16.7	8.3	12.5	12.5	29.4	25.1	87.4
12 <i>D. busckii</i>	119	132	—	8.3	20.9	12.5	33.7	29.4	87.4
13 <i>D. pictiventris</i>	141	141	146	—	16.7	12.5	38.0	33.7	92.0
14 <i>D. lebanonensis</i>	181	174	174	185	—	20.9	38.0	29.4	92.0
15 <i>Z. tuberculatus</i>	113	118	138	143	184	—	25.1	20.9	87.4
16 <i>C. amoena</i>	193	222	199	189	192	206	—	8.3	87.4
17 <i>C. procnemis</i>	207	217	190	197	193	206	105	—	82.7
18 <i>Ceratitits</i>	271	274	258	285	269	265	310	268	—

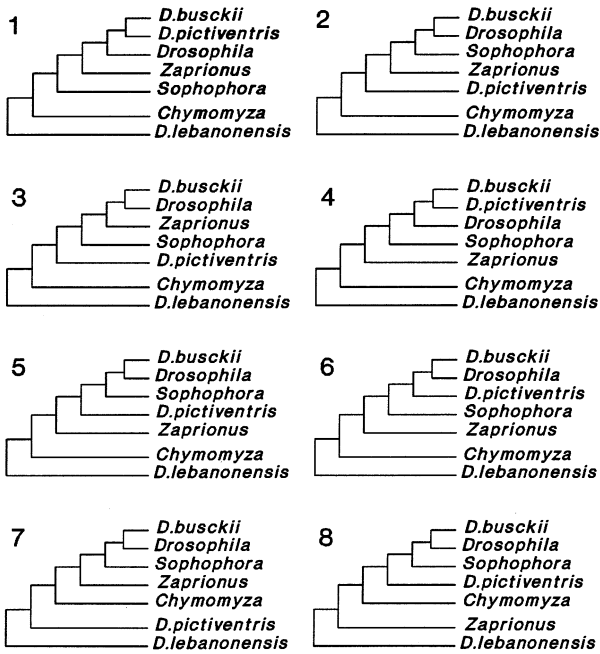


Fig. 4. Eight trees showing alternative phylogenetic hypotheses. *Drosophila* refers to the subgenus and includes *D. hydei* and *D. virilis*; the subgenus *Sophophora* includes three species of each of the *melanogaster* and *willistoni* groups; *Chymomyza* includes *C. amoena* and *C. procnemis*.

tiventris) and *Dorsilopha* (here, *D. busckii*) are phylogenetically closely related to each other and to the genus *Zaprionus*. These taxa would have been part of a radiation that included the *Drosophila* subgenus. The statistical analysis presented in Fig. 4 and Table 6 supports clustering these taxa but shows *Zaprionus* as the sister taxon to the cluster of the subgenera *Hirtodrosophila*, *Dorsilopha*, and *Drosophila* (see 1 in Fig. 4), rather than being closer to the first two subgenera than to the *Drosophila* subgenus.

According to Throckmorton (1975), the radiation of the four taxa just mentioned would have occurred after separation of their stem lineage from the subgenus *Sophophora* (represented in our paper by nine species; three from each of the *melanogaster*, *obscura*, and *willistoni* groups), a claim also supported by our analysis (Fig. 4 and Table 6). In a cladistic revision of the Drosophilidae, Grimaldi (1990) concludes that *Sophophora* is a sister clade to the doublet made of the *Dorsilopha* and *Drosophila* subgenera, whereas *Zaprionus* is sister to the previous three clades, and *Hirtodrosophila* is the sister clade of the previous four (see tree 2 in Fig. 4). On the basis of these proposed phylogenetic relationships, he accepts the status of *Zaprionus* as a genus distinct from *Drosophila* and raises *Hirtodrosophila* from subgenus to genus (see also DeSalle 1992; and DeSalle and Grimaldi 1991). Our analysis does not support these claims. If the

Table 6. Statistical comparison of eight tree topologies, using the method of Templeton (1983)^a

Tree	<i>Gpdh</i>				<i>Gpdh</i> & <i>Sod</i>			
	Steps	Difference	SE	Significantly worse?	Steps	Difference	SE	Significantly worse?
1	651	—	—	—	1,299	—	—	—
2	653	2	5.7	No	1,310	11	8.7	No
3	655	4	5.5	No	1,310	11	8.4	No
4	656	5	2.6	No	1,308	9	5.0	No
5	660	9	5.0	No	1,316	17	8.3	Yes
6	661	10	4.0	Yes	1,314	15	6.9	Yes
7	669	18	8.0	Yes	1,330	31	11.1	Yes
8	670	19	7.1	Yes	1,338	39	11.1	Yes

Table 6. Continued

Tree	<i>Gpdh</i> , <i>Sod</i> , & <i>Adh</i>			
	Steps	Difference	SE	Significantly worse?
1	2,131	—	—	—
2	2,157	26	10.4	Yes
3	2,156	25	9.9	Yes
4	2,152	21	6.4	Yes
5	2,167	36	10.0	Yes
6	2,158	27	7.9	Yes
7	2,191	60	14.1	Yes
8	2,196	65	14.0	Yes

^a Comparisons are with respect to tree 1 (Fig. 4)

	111111111	111222222	2233333
	801457778	899113778	8813333
	8446902575	6350972490	364156
	IQIAETKGVN	HDACKIRSTS	TEGNPN
<i>D. melanogaster</i> FastK
<i>D. melanogaster</i> SlowK
<i>D. simulans</i>K
<i>D. teissieri</i>K
<i>D. pseudoobscura</i>K..K
<i>D. miranda</i>K..K
<i>D. guanche</i>K..K
<i>D. willistoni</i>E.....K..K
<i>D. paulistorum</i>E.....K..K
<i>D. nebulosa</i>E.....K..K
<i>D. virilis</i>E.....D...K
<i>D. hydei</i>E.....K..K
<i>D. busckii</i>K.....S.....K..K
<i>D. pictiventris</i>K..K
<i>D. lebanonensis</i>N..I..G.....K..K
<i>Zaprionus tuberculatus</i>S.....S.....K
<i>Chymomyza amoena</i>	V..S.S.A..	..ES...K.S.	...S..K
<i>Chymomyza procnemis</i>	V..S.N.A..	..ES...K.S.	...S..K
<i>Ceratitis capitata</i>	-TVSLRITPP	YESTGV.AKT	S...VQ

Fig. 5. Polymorphic amino acid sites inferred from the 18 nucleotide sequences given in Fig. 2. Site numbers (shown on top) are as in *D. melanogaster* (Bewley et al. 1989) and *D. virilis* (Tominaga et al. 1992). Dots indicate amino acids identical to those on top.

subgenus *Sophophora* is retained within the *Drosophila* genus, cladistic taxonomy requires that *Hirtodrosophila* be retained as a subgenus and that *Zaprionus* be downgraded from genus to subgenus.

Throckmorton (1975) proposes that the traditional genus *Chymomyza* represents an early branch of the *So-*

phophora radiation, a claim contradicted by our analysis, which statistically supports the position of *Chymomyza* outside the genus *Drosophila*, as proposed also by Grimaldi (1990). The phylogenetic position of *Chymomyza* outside the genus *Drosophila* (as in tree 1, Fig. 4) has been conclusively established by the presence of an extra intron in the *Sod* gene, which is also present in *Scaptodrosophila* and *Ceratitis*, but not in other *Drosophila* (or *Zaprionus*) (Kwiatowski et al. 1994). The discovery of this ancestral intron, deleted early in the evolution of the *Drosophila* genus, corroborates the phylogenetic position of *Scaptodrosophila* shown in tree 1 (Fig. 4) and thus that this taxon should be raised to the genus category, as done by Grimaldi (1990) and consistent also with Throckmorton's (1975) conclusions.

The pattern of evolution of GPDH is puzzling. There are no amino acid differences among the three members of any of the three *Drosophila* groups (Table 7), with times of divergence about 8–10 My between *D. teissieri* and the other two species of the *melanogaster* group, about 20 My between *D. guanche* and the other two species of the *obscura* group, and about 25 My between *D. nebulosa* and the two other species of the *willistoni* group. Barrio and Ayala (1997) have studied 253 GPDH amino acids (coded by exons 3–6) in 14 species of the

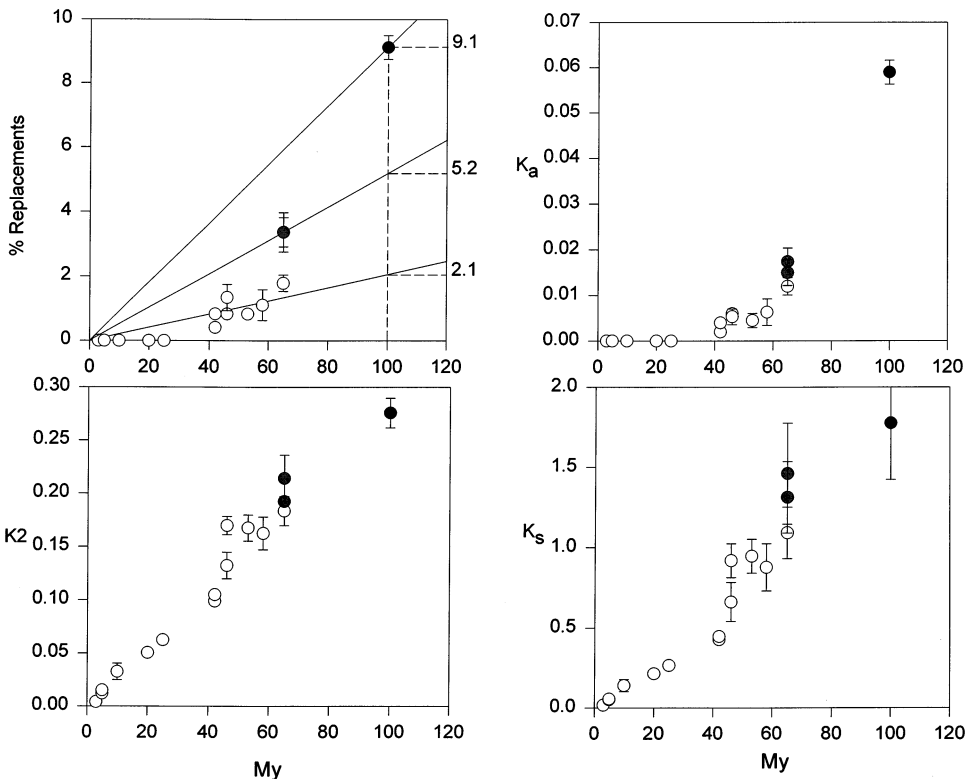


Fig. 6. Rate of *Gpdh* evolution. Time (*abscissa*) is in million years. Top left: Amino acid replacements (PAM values, Dayhoff 1978). Bottom left: Genetic distances (K2, Kimura's two-parameter method). Top right: Nonsynonymous nucleotide substitutions (K_a , Li 1993). Bottom right: Synonymous substitutions (K_s , Li 1993). The percent rates of amino acid replacement between lineages per 100 My are shown at

right of the top left diagram. The rate of 2.1 for comparisons between *Drosophila* species has been obtained by best-fit regression. The rates between *Chymomyza* and *Drosophila* (5.2) and between *Ceratitis* and the *Drosophilids* (9.1) are obtained by drawing straight lines from the origin to the relevant points.

Table 7. *Gpdh* divergence between increasingly distant species^a

Comparison	My	% amino acid replacements		% nucleotide substitutions	
		X	Per 100 My	X	Per 100 My
1. Within group	5–25	0	0	0.4–6.2	6.7–16.3
2. Between groups	45 ± 10	0.83 ± 0.00	0.9	16.8 ± 0.2	18.7
3. Between subgenera	55 ± 10	1.15 ± 0.06	1.0	15.8 ± 0.2	14.3
4. Between genera	60 ± 10	2.87 ± 0.13	2.3	20.3 ± 0.4	17.0
5. Between families	100 ± 20	9.12 ± 0.09	4.6	27.6 ± 0.3	13.8

^a The amino acid replacements are PAM-corrected; nucleotide substitutions estimated as Kimura's distances. Species included in the comparisons are: 1, between the three members of each of the *melanogaster*, *obscura*, or *willistoni* species groups; 2, between species from the three different groups; 3, between species from different subgenera (*Sophophora*, *Drosophila*, *Dorsilopha*, *Hirtodrosophila*), including *Zaprionus*, but excluding *Scaptodrosophila*; 4, the genera compared are *Chymomyza* with either *Drosophila* or *Scaptodrosophila*; 5, between *Ceratitidis* and the drosophilids. The plus/minus values are crude estimates of error for My, but SEs for replacements and substitutions. The time estimates used in this table (and in the text) are crude consensus values derived from the recent literature (see particularly Powell and DeSalle 1995, and references therein; Kwiatowski et al. 1994; Takezaki et al. 1995; Russo et al. 1995). Beverley and Wilson (1984) give somewhat larger estimates for the divergence between families (123 My vs our 120 ± 20 My) and between the subgenera *Sophophora* and *Drosophila* (62 My vs our 55 ± 10 My)

affinis and *obscura* subgroups and observed very few polymorphisms: (1) One species differs from all others at each of two sites and (2) two species differ from all others at a third site (but are identical to each other, although one belongs to the *affinis* and the other to the *obscura* subgroup). This corresponds to an average of about one-half amino acid difference between any two species, or 0.002 per site, for species that diverged 20–25 My ago. These results are consistent with previous observations that GPDH is a slowly evolving enzyme (Bewley et al. 1989; Wells 1995, 1996a,b). Comparisons among all *Drosophila* (including *Zaprionus*) species also manifest a slow rate of amino acid replacement. The number of amino acid differences between species from different groups of the *Sophophora* subgenus are two (corresponding to about 90 My of separate evolution, 45 My on each lineage) and between species from different subgenera are between zero and five (corresponding to about 110 My of separate evolution). This is a lineage rate of about 1×10^{-10} replacements/site/year, comparable to the rates of the very slowly evolving histones; e.g., 1.7×10^{-10} for H2A or H2B (Wilson et al. 1977).

The apparent rate of GPDH evolution increases, however, when *Chymomyza* is compared with *Drosophila*. The average number of amino acid differences is about eight (Tables 5 and 7), corresponding to 120 My of evolution, or a rate of 2.6×10^{-10} replacements/site/year. When *Ceratitidis* is compared with *Drosophila* or *Chymomyza*, the average number of replacements is about 23 (Table 4), corresponding to 200 My of evolution, or a rate of 4.6×10^{-10} replacements/site/year, still slower than the rate of 6.7×10^{-10} of cytochrome *c* (Wilson et al. 1977), a slowly evolving protein.

What is going on? One could argue that the rate of evolution in the *Chymomyza* lineage, as well as between nodes 3 and 4 in Fig. 7, is faster than within the genus *Drosophila*. But it would have to be much faster in order to account for the 2.6 times increase in the overall number of amino acid differences. If we assume 120 My of

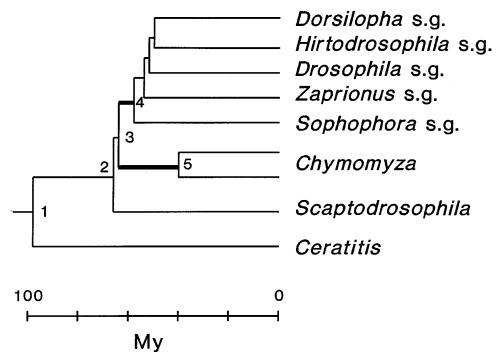


Fig. 7. Phylogeny of genera and subgenera. Thicker branches indicate postulated faster rates of *Gpdh* evolution. Numbers identify relevant nodes.

separate evolution between any *Chymomyza* and *Drosophila* species, 55 of the 120 My belong to the evolution of the *Drosophila* lineages; we are left with 65 My of evolution (55 My of the *Chymomyza* lineage, plus 10 My of the *Drosophila* lineage before the radiation of the genus, i.e., between nodes 3 and 4; see Fig. 7). But we know that the *Chymomyza* lineages have also been evolving slowly. The two *Chymomyza* species diverged about 42 My ago (Kwiatowski et al. 1994) but differ only by two amino acids (compared to about three between the somewhat older *Drosophila* subgenera) (Table 4). Thus, the time-length during which evolution would have happened faster is at a maximum $120 - 55 - 42 = 23$ My, or one-sixth of the time of separate evolution between any *Chymomyza* species and any *Drosophila* species. Thus, in order to account for the 2.6 times increase in amino acid differences between the two genera, we have to conclude that during those 23 My (represented by thicker lines in Fig. 7), the rate of GPDH was nine times faster than for the other 97 My of separate evolution between *Chymomyza* and *Drosophila* species.

This conclusion can be somewhat exacerbated by noticing that the average distance between *Scaptodro-*

sophila and *Drosophila* is about four amino acids, only slightly greater than the average between the *Drosophila* subgenera. Thus between nodes 2 and 4 (Fig. 7) the prevailing rate of evolution must have been the same as in the *Drosophila* genus. Therefore, the accelerated rate of *Chymomyza* evolution would have been restricted to the branch segment between nodes 3 and 5, when it would have been 12 times greater than at any other time in the evolution of *Drosophila* and *Chymomyza*.

We noted earlier that the apparent rate of amino acid replacements between *Ceratitidis* and *Drosophila* is about four times greater than within the *Drosophila* genus. Using the same logic as in the previous paragraph, the acceleration could have only occurred (Fig. 7) between nodes 1 and 2 and from node 1 to *Ceratitidis*, corresponding to 140 My out of the total 200 My separating the two genera. The rate of evolution corresponding to those branches would need to be about six times faster than in the rest of the tree (excluding *Chymomyza*).

It may be noted that the disparate rates of GPDH evolution observed are not strongly dependent on the particular times of divergence assumed. We have noted in the Results (see also Fig. 4 and Table 6) that the topology displayed in Fig. 7 is statistically superior to other alternatives. This topology shows that *Chymomyza* is more closely related than *Scaptodrosophila* to the *Drosophila* species. Yet the mean number of amino acid replacements (calculated from Table 5) is much greater between *Chymomyza* and *Drosophila* (33.7 ± 0.8) than between *Scaptodrosophila* and *Drosophila* (17.9 ± 0.7). It is, in any case, now well established that the *Scaptodrosophila* and *Chymomyza* lineages diverged from the *Drosophila* lineages within a relatively short time interval. Changing their branching sequence would not by itself account for the large discrepancy in the GPDH differences. The same point can be made with respect to *Ceratitidis*. The time of divergence between *Ceratitidis* and *Drosophila* assumed in Table 7 is 100 ± 20 My, not quite double the time of divergence between different *Drosophila* subgenera (55 ± 10); yet the number of amino acid replacements is 92.0 ± 0.8 , nearly six times as large. This discrepancy cannot be accounted for even if the divergence between the two families were as old as 150 My or the divergence between the subgenera as recent as 40 My, values which are outside those determined by various methods (see Powell and DeSalle 1995 and other references cited in Table 7).

The point made in the previous paragraph can also be made by reference to Fig. 6. The comparisons made and the times of divergence assumed are the same in all panels. Yet the nucleotide genetic distances (K_2) as well as the number of synonymous substitutions (K_s) are approximately linear with respect to time (bottom panels), whereas the number of amino acid replacements (PAM) or of nonsynonymous substitutions (K_a) is not (top panels).

Wells (1996b) has elucidated one idiosyncrasy of GPDH evolution: reversal and parallelism. He obtained the complete GPDH sequence of 13 *Drosophila* species and observed amino acid polymorphism at 17 of 363 sites, at nine of which two or more species were different from the rest. At four of these nine sites, parallel or reversed replacements have occurred. Wells considers these four sites hypervariable because they account for half of all interspecific amino acid replacements, even though only two different amino acids occur at these four sites in the 13 species. The degree of homoplasy, as measured by the global consistency index (Archie 1989), is significantly greater in *Drosophila* GPDH evolution than the average in ten additional proteins, one (ADH) from *Drosophila* and nine from vertebrates (Wells 1996b).

If we exclude *Ceratitidis* and *Chymomyza*, we have observed eight polymorphic amino acid sites in 15 species of *Drosophila* (including *Zaprionus*) (Fig. 5). At three sites, only one species differs from all others; these sites are uninformative for the present purposes. Four of the remaining five sites exhibit parallel or reversed replacements, namely, sites 193, 195, 314, and 336. By reference to the phylogeny in Fig. 3, the minimum number of independent amino acid replacements would be:

1. Three at site 193, from ancestral E \rightarrow D at the root of the genus and two parallel reversals D \rightarrow E at the root of the *willistoni* group and of the *Drosophila* subgenus (*virilis* and *hydei*); the next most parsimonious alternative would be to assume that D is ancestral and that four D \rightarrow E parallel replacements have occurred in *Ceratitidis*, *Chymomyza*, the *willistoni* group, and the *Drosophila* subgenus. The first alternative, in addition to requiring only three rather than four replacements, is also favored when one notes that the muscoid fly *Calliphora* has E at this site (our unpublished data), similarly as *Chymomyza* and *Ceratitidis*.
2. Three at site 195, all parallel from ancestral S \rightarrow A in the *Sophophora* subgenus (which we assume to be monophyletic, comprising the *willistoni*, *obscura*, and *melanogaster* groups), in the *Drosophila* subgenus, and in *pictiventris*. (*Calliphora*, like *Chymomyza* and *Ceratitidis*, has S at this site.) An equally parsimonious alternative is to assume ancestral S \rightarrow A in the root of the *Drosophila* genus (including *lebanonensis* but not *Chymomyza*) and two parallel reversals A \rightarrow S in *D. busckii* and *Zaprionus*; however, if we assume that the phylogenetic positions of *Zaprionus* and *pictiventris* are reversed in Fig. 3, only two S \rightarrow A parallel replacements are required.
3. Two parallel replacements at site 314, ancestral G \rightarrow K in the *obscura* and *willistoni* groups; alternatively, ancestral G \rightarrow K in the root of *Sophophora* and one reversal K \rightarrow G in the *melanogaster* group.
4. At 336 three parallel and one reverse replacement: ancestral K \rightarrow N in the *melanogaster* and *willistoni*

groups and in *busckii* + *pictiventris* (assuming these two are monophyletic) and N → K between the Fast and Slow alleles of *D. melanogaster* (the Slow allele is known to be more recent, Takano et al. 1993; Wells, 1996b); alternatively, ancestral K → N in the root of the *Drosophila* genus, and three parallel reversals N → K in the *obscura* group, the *Drosophila* subgenus, and *D. melanogaster* Slow.

Of the four homoplasious sites detected in our phylogeny, only site 336 is manifest in Wells's (1996) phylogeny, which only includes *Drosophila* species. The other homoplasious sites (28, 45, and 362) detected by Wells are not included within the *Gpdh* fragment sequenced by us.

Wells (1996b) has pointed out that the high level of homoplasia makes GPDH unsuitable for reconstructing the phylogeny of *Drosophila* species. An instance of misleading similarity occurs between the *melanogaster* group species and *D. pictiventris*, which show no amino acid differences (Table 3), although they are phylogenetically remote. GPDH is unsuitable for reconstructing the phylogeny of *Drosophila* because of the slow rate of amino acid replacements within the genus. Moreover, as already noted, the detectable rate of amino acid replacements is very erratic within the set of fruitfly species surveyed in our study, which makes GPDH unsuitable for estimating times of divergence or making any other inferences dependent on the assumption of a molecular clock.

The evolution of *Gpdh*, however, appears to be much more nearly uniform through time at the nucleotide level than at the amino acid level. When the nucleotide distances between species are calculated using Kimura's (1980) two-parameter method (Table 5, below diagonal, and Table 7), they approximate a uniform rate of evolution, as shown in Fig. 6 (lower left panel). This apparent regularity results, no doubt, from the confounding of two different rates corresponding to two sets of nucleotides. The rate of nonsynonymous nucleotide substitutions (K_a), calculated by the method of Li (1993), is clearly not a linear function of time, as shown in Fig. 6 (top right panel), whereas the rate of synonymous substitutions (K_s) is more nearly linear (Fig. 6, bottom right), although much variation is apparent.

Acknowledgments. We thank Dr. S. Wells for providing the sequences of *D. busckii* and *D. lebanonensis*; Drs. J. Felsenstein, D. Higgins, M. Nei, and A. Zharkikh for making available their computer programs; and Drs. W.M. Fitch, R.R. Hudson, and G.S. Spicer for helpful comments. This research was supported by NIH grant GM42397.

References

Archie JW (1989) Homoplasia excess ratios: new indices for measuring levels of homoplasia in phylogenetic systematics and a critique of the consistency index. *Syst Zool* 38:253–269

- Ausubel FM, Brent R, Kingston RE, Moore DD, Smith JA, Seidman JG, Struhl K (1987) *Current protocols in molecular biology*. John Wiley, New York
- Barrio E, Ayala FJ (1997) Evolution of the *Drosophila obscura* species group inferred from the *Gpdh* and *Sod* genes. *Mol Phyl Evol* 7:1–15
- 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
- Bewley GC, Cook JL, Kusakabe S, Mukai T, Rigby DL, Chambers GK (1989) Sequence, structure and evolution of the gene coding for *sn*-glycerol-3-phosphate dehydrogenase in *Drosophila melanogaster*. *Nucleic Acids Res* 17:8553–8567
- Cook JL, Bewley GC, Shaffer JB (1988) *Drosophila sn*-glycerol-3-phosphate dehydrogenase isozymes are generated by alternate pathways of RNA processing resulting in different carboxyl-terminal amino acid sequences. *J Biol Chem* 263:10858–10864
- Csank C, Taylor FM, Martindale DW (1990) Nuclear pre-mRNA introns: analysis and comparison of intron sequences from *Tetrahymena thermophila* and other eukaryotes. *Nucleic Acids Res* 18:5133–5141
- Dayhoff MD (1978) *Atlas of protein sequences and structure*. National Biomedical Research Foundation, Washington, DC
- DeSalle R (1992) The phylogenetic relationships of flies in the family Drosophilidae deduced from mtDNA sequences. *Mol Phylogenet Evol* 1:31–40
- DeSalle R, Grimaldi DA (1991) Morphological and molecular systematics of the Drosophilidae. *Annu Rev Ecol Syst* 22:447–475
- Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. *J Mol Evol* 17:368–374
- Felsenstein J (1989) PHYLIP—phylogeny inference package (version 3.2). *Cladistics* 5:164–166
- Fitch WM (1971) Toward defining the course of evolution: minimum change for a specified tree topology. *Syst Zool* 20:406–416
- Fitch WM, Ayala FJ (1994) The superoxide dismutase molecular clock revisited. *Proc Natl Acad Sci USA* 91:6802–6807
- Grantham R, Gautier C, Gouy M, Mercier R, Pavé A (1980) Codon catalog usage and the genome hypothesis. *Nucleic Acids Res* 8:r49–r62
- Grimaldi D (1990) A phylogenetic revised classification of genera in the Drosophilidae. *Bull Am Museum Nat Hist* 197:1–139
- Hawkins JD (1988) A survey on intron and exon lengths. *Nucleic Acids Res* 16:9893–9908
- Higgins DG, Bleasby AJ, Fuchs R (1992) CLUSTAL V: improved software for multiple sequence alignment. *Comput Appl Biosci* 8:189–191
- Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: Munro HN (ed) *Mammalian protein metabolism*. Academic Press, New York, pp 21–132
- Kawasaki ES (1990) Sample preparation from blood, cells, and other fluids. In: Innis MA, Gelfand DH, Sninski JJ, White TJ (eds) *PCR protocols: a guide to methods and applications*. Academic Press, San Diego, pp 146–152
- Kimura M (1980) A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. *J Mol Evol* 16:111–120
- Kishino H, Hasegawa M (1989) Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. *J Mol Evol* 29:170–179
- Kozak LP, Burkart DL, Hjorth JP (1982) Unlinked structural genes for the developmentally regulated isozymes of *sn*-glycerol-3-phosphate dehydrogenase in mice. *Dev Genet* 3:1–6
- Kumar S, Tamura K, Nei M (1993) MEGA: molecular evolutionary genetics analysis, version 1.0. The Pennsylvania State University, University Park, PA 16802
- Kwiatowski J, Skarecky D, Hernandez S, Pham D, Quijas F, Ayala FJ (1991) High fidelity of the polymerase chain reaction. *Mol Biol Evol* 8:884–887

- Kwiatowski J, Skarecky D, Ayala FJ (1992a) Structure and sequence of the Cu,Zn *Sod* gene in the Mediterranean fruit-fly *Ceratitis capitata*: intron insertion and deletion in the evolution of the SOD gene. *Mol Phyl Evol* 1:72–82
- Kwiatowski J, Skarecky D, Burgos M, Ayala FJ (1992b) Structure and sequence of the Cu,Zn superoxide dismutase gene of *Chymomyza amoena*: phylogeny of the genus and codon-use evolution. *Insect Mol Biol* 1:3–13
- Kwiatowski J, Skarecky D, Bailey K, Ayala FJ (1994) Phylogeny of *Drosophila* and related genera inferred from the nucleotide sequence of the Cu,Zn *Sod* gene. *J Mol Evol* 38:443–454
- Lakovaara S, Saura A, Lankinen P (1977) Evolution at the α -Gpdh locus in Drosophilidae. *Evolution* 31:319–330
- Li W-H (1993) Unbiased estimation of the rates of synonymous and nonsynonymous substitution. *J Mol Evol* 36:96–99
- Mount SM, Burks C, Hertz G, Stormo GD, White O, Fields C (1992) Splicing signals in *Drosophila*: intron size, information content, and consensus sequences. *Nucleic Acids Res* 20:4255–4262
- O'Brien SJ, MacIntyre RJ (1972) The α -glycerophosphate cycle in *Drosophila melanogaster*. II. Genetic aspects. *Genetics* 71:127–138
- O'Brien SJ, MacIntyre RJ (1978) Genetics and biochemistry of enzymes and specific proteins of *Drosophila*. In: Ashburner M, Wright TRF (eds) *The genetics and biology of Drosophila*. Academic Press, New York, pp 395–551
- Otto J, Argos P, Rossmann MG (1980) Prediction of secondary structural elements in glycerol-3-phosphate dehydrogenase by comparison with other dehydrogenases. *Eur J Biochem* 109:325–330
- Powell JR, DeSalle R (1995) *Drosophila* molecular phylogenies and their uses. In: Hecht MK, MacIntyre RJ, Clegg MT (eds) *Evolutionary biology*, vol 28. Plenum Press, New York, pp 87–138
- Russo CAM, Takezaki N, Nei M (1995) Molecular phylogeny and divergence times of Drosophilid species. *Mol Biol Evol* 12:391–404
- Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425
- Sharp PM, Cowe E, Higgins DG, Shields DC, Wolfe KH, Wright F (1988) Codon usage patterns in *Escherichia coli*, *Bacillus subtilis*, *Saccharomyces cerevisiae*, *Schistosaccharomyces pombe*, *Drosophila melanogaster* and *Homo sapiens*; a review of the considerable within-species diversity. *Nucleic Acids Res* 16:8207–8211
- Starmer WT, Sullivan DT (1989) A shift in the third-codon-position nucleotide frequency in alcohol dehydrogenase genes in the genus *Drosophila*. *Mol Biol Evol* 6:546–552
- Takano ST, Kusakabe S, Mukai T (1993) DNA polymorphism and the origin of protein polymorphism at the *Gpdh* locus of *Drosophila melanogaster*. In: Takahata N, Clark AG (eds) *Mechanisms of molecular evolution*. Japan Scientific Societies Press, Tokyo, pp 179–190
- Takezaki N, Rzhetsky A, Nei M (1995) Phylogenetic test of the molecular clock and linearized trees. *Mol Biol Evol* 12:823–833
- Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* 10:512–526
- Templeton AR (1983) Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and apes. *Evolution* 37:221–244
- Throckmorton LH (1975) The phylogeny, ecology and geography of *Drosophila*. In: King RC (ed) *Handbook of genetics*. Plenum Press, New York, pp 421–469
- Tominaga H, Shiba T, Narise S (1992) Structure of *Drosophila virilis* glycerol-3-phosphate dehydrogenase gene and a comparison with the *Drosophila melanogaster* gene. *Biochim Biophys Acta* 1131:233–238
- von Kalm L, Weaver J, DeMarco J, MacIntyre RJ, Sullivan DT (1989) Structural characterization of the α -glycerol-3-phosphate dehydrogenase-encoding gene of *Drosophila melanogaster*. *Proc Natl Acad Sci USA* 86:5020–5024
- Wells RS (1995) Sequence and evolution of the *Drosophila pseudoobscura* glycerol-3-phosphate dehydrogenase locus. *J Mol Evol* 41:886–893
- Wells RS (1996a) Nucleotide variation at the *Gpdh* locus in the genus *Drosophila*. *Genetics* 143:375–384
- Wells RS (1996b) Excessive homoplasy in an evolutionarily constrained protein. *Proc R Soc Lond [Biol]* 263:393–400
- Wheeler MR (1981) The Drosophilidae: a taxonomic overview. In: Ashburner M, Carson HL, Thompson JNJ (eds) *The genetics and biology of Drosophila*. Academic Press, New York, pp 1–97
- Wilson AC, Carlson SS, White TJ (1977) Biochemical evolution. *Annu Rev Biochem* 46:573–639