MOLECULAR PHYLOGENETICS AND EVOLUTION
Vol. 7, No. 1, February, pp. 79–93, 1997
ARTICLE NO. FY960375

Evolution of the *Drosophila obscura* Species Group Inferred from the *Gpdh* and *Sod* Genes

ELADIO BARRIO[1] AND FRANCISCO J. AYALA[2]

Department of Ecology and Evolutionary Biology, University of California, Irvine, California 92717

Received February 28, 1996; revised July 10, 1996

The *Drosophila obscura* group consists of several dozen Nearctic and Palearctic species, with a phylogeny that remains largely unresolved in spite of numerous morphological, cytogenetic, and molecular investigations. We have partially sequenced two genes, *Gpdh* (about 1000 bp) and *Sod* (about 700 bp) in 12–13 species and 1 subspecies in order to settle the issues. Difficulties in resolving the phylogeny emanate from the rapid sequence of successive radiations. Nevertheless, the following conclusions are warranted: (1) The Palearctic species include two monophyletic subgroups, *subobscura* (which may be the most ancient clade of the whole group) and *obscura,* plus two other species with unresolved phylogenetic positions, *D. bifasciata* and *D. subsilvestris;* (2) the Nearctic species form a monophyletic group consisting of two monophyletic sister clades, the *affinis* subgroup and the *pseudoobscura* subgroup. The Palearctic radiation may have resulted from adaptation to expanding temperate forests in the Old World. A second radiation occurred during the colonization of the deciduous forests of the New World by the descendants of a single lineage that soon split into the *affinis* and *pseudoobscura* subgroups.

INTRODUCTION

The *Drosophila obscura* group consists of 35 described species and has been the subject of extensive investigations concerning genetics, ecology, ethology, cytology, and evolution. Yet the phylogeny and taxonomy of the group remain controversial. Based on morphological traits, Sturtevant (1942) distinguished two subgroups: *obscura* with Palearctic and Nearctic species, and *affinis,* with mostly Nearctic species (Throckmorton, 1975). The *obscura* group also includes a set of African species, mostly discovered after 1985 and classified in the *microlabis* subgroup (Tsacas *et al.*, 1985; Cariou *et al.*, 1988; Brehm and Krimbas, 1990, 1992, 1993; Brehm *et al.*, 1991; Bachmann *et al.*, 1992; Ruttkay *et al.*, 1992), which will not be further discussed in this paper.

Extensive allozyme investigations support the split of Sturtevant's *obscura* subgroup into two subgroups, a new *obscura* subgroup encompassing the Palearctic species, and the *pseudoobscura* subgroup encompassing the Nearctic species (Lakovaara *et al.*, 1972, 1976; Marinković *et al.*, 1978; Lakovaara and Keränen, 1980; Lakovaara and Saura, 1982; see also Cabrera *et al.*, 1983; Cariou *et al.*, 1988). These studies further suggest that (1) the Nearctic *pseudoobscura* subgroup is phylogenetically closer to the (mostly) Nearctic *affinis* subgroup than to the Palearctic *obscura* subgroup; and (2) the *obscura* subgroup may not be monophyletic, but either polyphyletic or paraphyletic. One or both of these two inferences are also supported by mitochondrial DNA restriction analysis (Latorre *et al.*, 1988; González *et al.*, 1990; Barrio *et al.*, 1992) and sequencing (Ruttkay *et al.*, 1992; Beckenbach *et al.*, 1993; Barrio *et al.*, 1994), DNA-DNA hybridization (Goddard *et al.*, 1990) and cytogenetic comparison of chromosomal elements (Brehm *et al.*, 1991; Brehm and Krimbas, 1990, 1992, 1993). Barrio *et al.*, (1994) have thus proposed that the *obscura* subgroup be split into two, one retaining the name and most species, and the *subobscura* subgroup, consisting of the widely distributed *D. subobscura* and two island endemics, *D. guanche* (Canary Islands) and *D. madeirensis* (Madeira).

Barrio *et al.* (1994) have averred that the remaining phylogenetic uncertainties concerning the *obscura* group may be attributable, in part, to the special evolutionary dynamics of the *Drosophila* mitochondrial genome, but it may also be that successive phylogenetic radiations occurred very rapidly in the group and thus are difficult to discern. In the present paper we test the rapid phyletic radiation hypothesis and seek to resolve the phylogenetic uncertainties, by sequencing segments from two nuclear genes, *Gpdh* and *Sod*. *Gpdh* encodes the α-glycerophosphate dehydrogenase, which has a key role in insect flight, regenerating NADH for cellu-

[1] Current address: Department de Genètica, Universitat de València, C/Dr. Moliner, 50. 46100-València, Spain.

[2] To whom correspondence and reprint requests should be addressed. Fax: 714-824-2474. E-mail: fjayala@uci.edu.

1055-7903/97 $25.00

lar respiration through the α-glycerophosphate cycle in the thoracic flight muscle (Sacktor, 1975). *Sod* encodes the Cu,Zn superoxide dismutase, a highly specific enzyme that protects aerobic cells against the toxicity of free oxygen radicals (Fridovich, 1986).

MATERIALS AND METHODS

Drosophila Species

We use strains from 13 species and 1 subspecies belonging to the *Drosophila obscura* group, mostly obtained from the National *Drosophila* Species Stock Center at Bowling Green, OH (stock reference numbers will be given for these); the sources of the other strains are given below.

Seven species belong to the *D. obscura* subgroup, as follows. *D. subobscura* subgroup, three species: *D. subobscura* from Helsinki, Finland (strain H271 from Valencia, Spain); *D. madeirensis,* endemic in the Island of Madeira, Portugal; and *D. guanche,* endemic in the Canary Islands, Spain from Tenerife. *D. obscura* subgroup, four species: *D. ambigua* (14011-0091.1), from Port Coquitlam, British Columbia, Canada; *D. obscura* from Girona, Spain; *D. subsilvestris* of unknown origin; and *D. bifasciata* (14012-0181.0) from Lake Akan-Ko, Japan. *D. guanche, D. madeirensis, D. obscura,* and *D. subsilvestris* were kindly supplied by Dr. María Monclús, University of Barcelona. Three additional species belong to the *D. affinis* subgroup: *D. affinis* (14012-0141.0) from Crystal Lake, Hastings, NE; *D. azteca* (14012-0171.6) from Mather, CA; *D. tolteca* (14012-0201.0) from Coroico, Bolivia. Finally, three species and one subspecies are members of the *pseudoobscura* subgroup: *D. miranda* (14011-0101.1) from Mather, CA; *D. p. pseudoobscura* (14011-0121.43) also from Mather, CA; *D. p. bogotana* (14011-0121.69) from Bogotá, Colombia; and *D. persimilis* (14011-0111.2) from Mt. San Jacinto, CA.

D. melanogaster Gpdh (Bewley *et al.,* 1989) and *Sod* (Kwiatowski *et al.,* 1989) sequences are used as outgroups. Additional *Sod* and *Gpdh* sequences, used in some analyses, are from Kwiatowski *et al.* (1994 and in press, respectively).

DNA Extraction, Amplification, Cloning, and Sequencing

The *Drosophila Gpdh* gene consists of 8 exons that encode 3 isoenzymes that arise by differential expression of exons 6 to 8. The region sequenced in the present study comprises about 1000 bp of exons 3 to 6 (Fig. 1), corresponding to positions 3020 to 4241 in the *D. melanogaster* sequence (Bewley *et al.,* 1989). The *Sod* gene consists in *Drosophila* of 2 exons and one intron. The region sequenced in this study comprises about 700 bp that include most of the coding region (except the 5′ and 3′ ends) and the intron (Fig. 2), positions 438 to 1504 in the *D. melanogaster* sequence (Kwiatowski *et al.,* 1989).

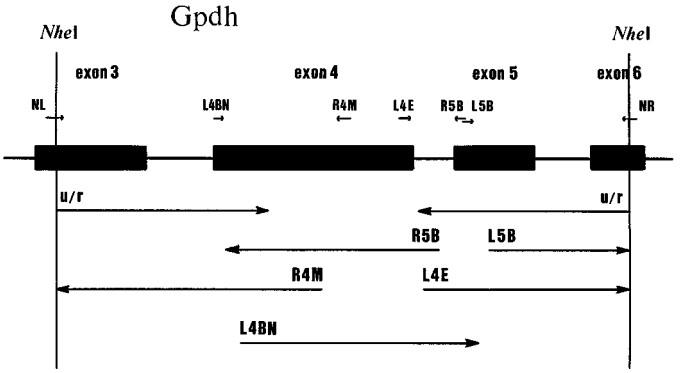

FIG. 1. Structure of the *Gpdh* gene region analyzed in the present study. The black boxes represent exons 3 to 6. The small arrows represent primers used for PCR amplification and sequencing. The amplified segment corresponds to the bracketed region between primers ML and NR. Long arrows indicate the extent and direction of sequencing.

Genomic DNA was prepared from 10 to 20 flies according to the method of Kawasaki (1990). Both gene regions were amplified by PCR using high-fidelity conditions (Kwiatowski *et al.,* 1991). Primers for PCR amplification of the *Gpdh* region (Fig. 1) were designed by comparing available sequences from other *Drosophila* species (Tominaga *et al.,* 1995): GNL, 5′-CCCGACCTGGTTGAGGCTAGCCAAGAATGC-3′; GNR, 5′-ACATATGCTCAGGGTGCTAGCGTATGCA-3′ (both contain a *Nhe*I site). Primers for the *Sod* region (Fig. 2) are described elsewhere (Kwiatowski *et al.,* 1994): N, 5′-CCTCTAGAAATGGTGGTTAAAGCTGT-NTGCGT-3′ (containing a *Xba*I site); and C, 5′-CTTGC-TGAGCTCGTGTCCACCCTTGCCCAGATCATC-3′ (containing a *Sac*I site).

PCR products were cleaned with the Wizard PCR

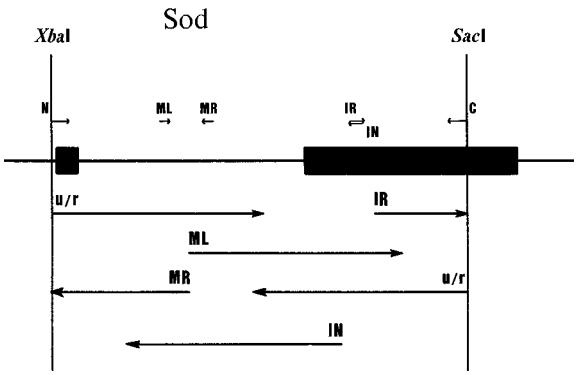

FIG. 2. Structure of the Cu,Zn Sod gene region. The black boxes represent the two exons, and the small arrows indicate the position of the primers used for PCR amplification and sequencing. The amplified segment corresponds to the bracketed region between primers N and C. The long arrows indicate the extent and direction of sequencing.

Preps DNA purification kit (Promega) and cloned into plasmids. The *Gpdh* region was cloned in pUC19 using the compatibility between the *Nhe*I and *Xba*I restriction sites present in the PCR primers and the cloning vector. Endonuclease *Xba*I was included in the ligation reaction, thus increasing approximately 10 times the cloning efficiency. The *Sod* region was cloned either into pUC19 by standard procedures or into the pCRII vector from the TA-cloning kit (Invitrogen) according to the manufacturer's protocol.

One clone of each gene was sequenced by the Sanger's dideoxynucleotide sequencing method for denatured double-stranded plasmid DNA by using the Sequenase v. 2.0 DNA sequencing kit (United States Biochemicals–Amersham). Compressions and ambiguities were solved by multiple sequencing of both strands. When single nucleotide substitutions appeared in only one species, this could be due to mistakes introduced during the PCR reaction by the *Taq* polymerase; additional clones from different PCR reactions were then sequenced to check the presumptive substitutions. Five oligonucleotides were used to sequence the *Gpdh* gene region (Fig. 1): L4BN, 5′-CCATGYGCTGTCTTGATGGG-3′; L4E, 5′-GATCTGATCACGACGTGTTAC-3′; L5E, 5′-CGCGTCTGAGGCGTTTGT-3′; R5B, 5′-CTCAGAGACGCGGCGGTTACGGCCAC-3′; and R4M, 5′-ACAGCCGCCTTGGTGTTGTCGCC-3′. Four oligonucleotides were used for sequencing the *Sod* gene region (Fig. 2): IN, 5′-GACATGCAKCCGTTRGTGTTG-3′; IR, 5′-GACAACACCAAYGGCTGCATGTC-3′; ML, 5′-TGGAATTGATGAATATTGC-3′; and MR 5′-GAGCTGCGCACTGTTATTSGAC-3′. In addition, we used the standard M13/pUC sequencing oligonucleotides.

The Genbank accession numbers for the *Gpdh* and *Sod* sequences are as follows (the two numbers are for *Gpdh* and *Sod*, respectively): *affinis* (U47874, U47879); *ambigua* (U47880, U47868); *azteca* (U47875, U47866); *bifasciata* (U47883, U47869); *guanche* (U47878, U47889); *madeirensis* (U47890, U47887); *miranda* (U47882, U47870); *obscura* (U47881, U47892); *persimilis* (U47886, U47873); *p. bogotana* (U47891, U47872); *p. pseudoobscura* (U47885, U47871); *subobscura* (U47877, U47888); *subsilvestris* (U47884, *Gpdh* only); *tolteca* (U47876, U47867).

Phylogenetic Inference

Sequences were aligned using the CLUSTAL V program (Higgins and Sharp, 1988). For phylogenetic analysis, we used the neighbor-joining (NJ) (Saitou and Nei, 1987), maximum-parsimony (MP) (Fitch, 1971), and maximum-likelihood (ML) (Felsenstein, 1981) methods. NJ trees and their statistical tests were performed using the computer program MEGA 1.0 (Kumar *et al.*, 1993). The statistical confidence of a particular cluster of sequences in the NJ trees was evaluated by the bootstrap test (1000 pseudoreplicates). Bootstrap values were corrected, whenever possible, by using the complete-and-partial bootstrap (CPB) technique (Li and Zharkikh, 1995). The MP trees and their bootstrap tests (1000 pseudoreplicates), and the ML trees, were obtained using the programs DNAPARS, SEQBOOT, and DNAML, respectively, implemented in the PHYLIP package 3.5 for Windows (Felsenstein, 1993). Alternative trees were compared by the maximum parsimony test proposed by Templeton (1983) and developed by Felsenstein (1985), and by the maximum likelihood test of Kishino and Hasegawa (1989). These tests were performed with the programs DNAPARS and DNAML, respectively, from the PHYLIP package.

RESULTS

Phylogenetic Analyses of Gpdh Sequences

The partial sequences of the *Gpdh* gene correspond to a region that ranges in size from 952 bp, in *D. affinis*, to 1010 bp, in *D. subsilvestris*, distributed as follows (Fig. 1): 168 bp from exon 3, 66–104 bp from intron III; 373 bp from exon 4, 54–74 bp from intron IV, 154 bp from exon 5, 64–81 bp from intron V, and 64 bp from exon 6.

Due to the difficulty of unambiguously aligning the *Gpdh* intron sequences from the *obscura* species with those from the outgroup, we examine separately the coding sequences and the intron sequences. *D. melanogaster* is only included in the alignment of the coding sequences.

The 15 *Gpdh* coding sequences yield 759 nucleotide positions aligned for all species. Nucleotide polymorphisms occur at 152 (20.0%) positions, of which 78 (10.3%) are phylogenetically informative. Nine variable sites correspond to first codon positions (4 are informative), 142 to third codon positions (74 informative), and only one to a second position. Only five sites (1 first, 1 second, and 3 third codon positions) show nonsynonymous substitutions, corresponding to replacements in four amino acid sites (two nonsynonymous substitutions occur at the first and second position of the same amino acid site). There are no amino acid indels. There is not much bias in G+C content: the overall nucleotide frequencies are 0.250 A, 0.255 T, 0.230 C, and 0.265 G, although there is some excess of G+C content at third codon positions (60.8%). Given the absence of substantive composition bias, we estimate nucleotide divergence according to Kimura's (1980) two-parameter model (Table 1, above diagonal). The phylogenetic tree derived by the neighbor-joining method is shown in Fig. 3A. The first branching of this NJ tree shows two monophyletic subgroups, *affinis* and *obscura*, consistent with the classification proposed by Sturtevant in 1942. The *obscura* subgroup splits into two monophyletic complexes, one corresponding to the Nearctic and the other to the Palearctic species. The next branching splits the *subobscura* triplet from the rest of the Old World *obscura* species. These branchings are supported by high bootstrap values.

TABLE 1

Pairwise Nucleotide Divergences for All Substitutions (Upper Triangle), or Only Transversions (Lower Triangle) between *Gpdh* Coding Sequences Estimated According to Kimura's (1980) Two-Parameter Model

Species	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. *melanogaster*		.1682	.1705	.1685	.1662	.1613	.1659	.1662	.1675	.1604	.1503	.1661	.1677	.1694	.1696
2. *subobscura*	.0412		.0066	.0120	.0365	.0323	.0296	.0379	.0666	.0665	.0814	.0550	.0593	.0578	.0550
3. *madeirensis*	.0369	.0040		.0133	.0379	.0337	.0323	.0393	.0694	.0679	.0829	.0564	.0607	.0592	.0564
4. *guanche*	.0383	.0026	.0040		.0324	.0282	.0269	.0310	.0636	.0622	.0800	.0507	.0550	.0536	.0507
5. *ambigua*	.0426	.0147	.0161	.0120		.0093	.0160	.0200	.0520	.0506	.0696	.0338	.0380	.0366	.0338
6. *obscura*	.0397	.0147	.0161	.0120	.0026		.0120	.0187	.0505	.0491	.0681	.0352	.0394	.0380	.0352
7. *subsilvestris*	.0455	.0147	.0161	.0120	.0080	.0080		.0146	.0548	.0534	.0738	.0366	.0380	.0366	.0366
8. *bifasciata*	.0426	.0174	.0188	.0147	.0107	.0107	.0080		.0592	.0591	.0798	.0394	.0436	.0422	.0394
9. *affinis*	.0469	.0188	.0229	.0188	.0202	.0202	.0229	.0202		.0133	.0325	.0492	.0535	.0520	.0492
10. *azteca*	.0484	.0202	.0215	.0174	.0188	.0188	.0215	.0215	.0040		.0242	.0492	.0535	.0521	.0492
11. *tolteca*	.0469	.0215	.0229	.0188	.0174	.0174	.0229	.0229	.0053	.0040		.0623	.0696	.0681	.0652
12. *miranda*	.0440	.0161	.0174	.0134	.0093	.0093	.0120	.0120	.0161	.0147	.0161		.0093	.0079	.0053
13. *p. pseudoobscura*	.0455	.0174	.0188	.0147	.0107	.0107	.0107	.0134	.0174	.0161	.0174	.0040		.0013	.0040
14. *p. bogotana*	.0455	.0174	.0188	.0147	.0107	.0107	.0107	.0134	.0174	.0161	.0174	.0040	.0000		.0026
15. *persimilis*	.0440	.0161	.0174	.0134	.0093	.0093	.0120	.0120	.0161	.0147	.0161	.0026	.0013	.0013	

Synonymous substitutions, however, can become saturated, which will introduce bias in a phylogeny when most substitutions are synonymous (presumably owing to strong purifying selection against amino acid replacements), as is the case in the *obscura* group. One escape from this bias is to rely exclusively on conservative changes (Swofford and Olsen, 1990; Hillis *et al.*, 1993), that is, in our case, transversions, which accumulate at slower rates than transitions (Brown *et al.*, 1982; Moritz *et al.*, 1987). (Nonsynonymous substitutions also accumulate at a slower rate than synonymous substitutions, but only seven, distributed over five sites, occur in the *Gpdh* data set, which makes them insufficient for the present purposes.) The nucleotide distances based on transversions, according to Kimura's (1980) two-parameter model, are presented in Table 1, below the diagonal. The NJ tree derived from these values is shown in Fig. 3B. This tree reverses the sequence of the first three branchings: first the *subobscura* triplet splits from the rest (CPB value of 90%), then the *affinis* subgroup, and finally the Nearctic *pseudoobscura* subgroup splits from the remainder Palearctic *obscura*. However, the three subgroups, *affinis*, *obscura*, and *pseudoobscura*, appear practically as a polytomy (CPB between 40 and 77%).

Differences in the transition/transversion ratio are important for inferring phylogenetic trees by the maximum likelihood method. The average transition/transversion ratio for the *obscura* species is 2.12, ranging from 0.5, between *D. obscura* and *D. subsilvestris*, to 5.0, between *D. affinis* and *D. tolteca*. If we consider very closely related taxa (because they better reflect the true ratio; Brown *et al.*, 1982; Hillis *et al.*, 1993), the range is about the same: a ratio of 0.7 for *subobscura* and *madeirensis*, 2.0 for *pseudoobscura* and *persimilis*, and, as noted, 5.0 for *azteca* and *tolteca*. We have, therefore, obtained ML trees using a variety of transition/transversion ratios, including 2.1, the average observed in our data set. The tree based on the average ratio (Fig. 4A) shows unlikely relationships, which also appear in other ML trees based on ratios smaller than 5. When ratios $\geq$5.0 are used, the ML trees exhibit the topology depicted in Fig. 4B. As in Fig. 3B, the *subobscura* triplet diverges first; but in contrast with that tree, the Nearctic subgroups *affinis* and *pseudoobscura* appear as sister taxa.

Maximum parsimony analysis (data not shown) yields implausible topologies, similar to the ML trees obtained with 2.1 or other low transition/transversion ratios. Thus, the MP tree shows *melanogaster* within the *affinis* subgroup, as a sister taxon of *D. tolteca*, although with low bootstrap values, suggesting excessive homoplasy.

The analyses based on only the intron (Fig. 5) sequences (which do not include the outgroup *D. melanogaster*) show the two Palearctic subgroups (*subobscura* and *obscura*) as a pair of sister clades and the two Nearctic subgroups (*affinis* and *pseudoobscura*) as another pair. These relationships are fairly similar to those in the ML tree (Fig. 4B), and in the transversion-only NJ tree (Fig. 3B), if we ignore the low-bootstrap relative positions of the *obscura* and *affinis* subgroups.

From the different analyses based on the two sets of alignments, we conclude that the 14 *obscura* taxa are grouped into four significantly monophyletic clusters: the *subobscura*, *obscura*, *affinis*, and *pseudoobscura* subgroups. We may note that *D. bifasciata* and *D. subsilvestris* typically cluster with the *ambigua–obscura* pair, but this is not well supported in most cases by bootstrap values.

We have carried out two additional statistical tests seeking to resolve phylogenetic relationships left unsettled by the previous analyses (Fig. 6). These tests are

FIG. 3. Neighbor-joining trees based on Gpdh coding sequences using all substitutions (A) or only transversion (B). Branch lengths are proportional to the scale given in substitutions per nucleotide. Percentage bootstrap values and complete-and-partial bootstrap values (in parentheses) based on 1000 pseudoreplicates are given on the nodes.

the maximum parsimony (MP) of Templeton (1983), as modified by Felsenstein (1993), and the maximum likelihood (ML) of Kishino and Hasegawa (1989), which compare phylogenies as wholes with one another. One question is the position of *D. subsilvestris* ("s" in Fig. 6), which in our analyses appears associated with *D. bifasciata* ("b" in Fig. 6), but has been previously (Lakovaara *et al.*, 1972; Gonzalez *et al.*, 1990; Barrio *et al.*, 1994) clustered with the couplet of *D. ambigua* and *D. obscura* ("ao" in Fig. 6). The other question is the branching sequence or topological configuration of the main clusters: *affinis* ("A" in Fig. 6, consisting of *affinis, azteca,* and *tolteca*), *subobscura* ("S," with *subobscura, madeirensis,* and *guanche*), and *pseudoobscura* ("P,"

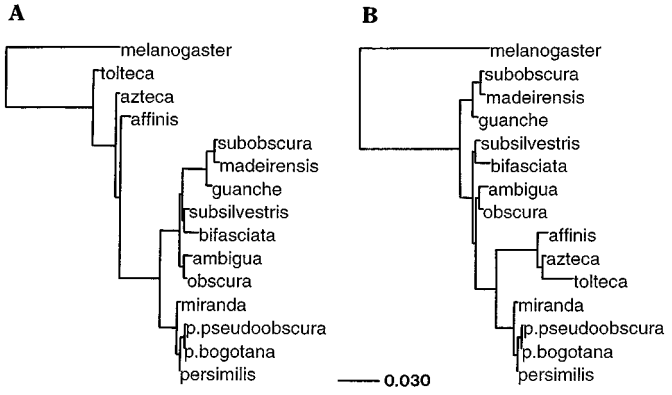

FIG. 4. Maximum-likelihood trees derived from Gpdh coding sequences. The transition/transversion rate is 2.1 in A, but 5 in B. The scale is for substitutions per site.

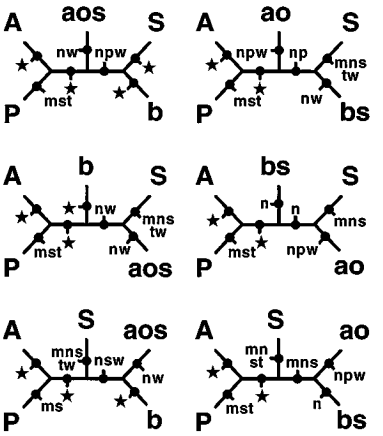

FIG. 6. Six Gpdh topologies that are statistically superior to the other possible 24 topologies, according to maximum parsimony (Templeton 1983) or maximum likelihood (Kishino and Hasegawa 1989). The lineages represented are the subgroups: *affinis* (A), *pseudoobscura* (P), and *subobscura* (S) and the species *D. ambigua* (a), *D. obscura* (o), *D. subsilvestris* (s), and *D. bifasciata* (b). These six trees are not significantly different, according to either test applied to the coding, the intron, or the total (intron + coding) sequences. The ML test based on the coding sequences including *D. melanogaster* as outgroup yields a number of superior but indistinguishable trees (of 210 possible), namely those in which the position of the root is indicated by a black dot. Also used as outgroup sequences are *D. nebulosa* (n), *D. paulistorum* (p), *D. simulans* (s), *D. teissieri* (t), and *D. willistoni* (w). Trees that share a root position indicated by a small line and letters are statistically equivalent. A star is used when all the outgroups root equivalent trees at the same position.

with *p. pseudoobscura*, *p. bogotana*, and *persimilis*), as well as the four species just mentioned.

There are 30 possible unrooted trees relating these groupings. Figure 6 displays the six configurations that are statistically superior to the other 24; but these six, however, are not significantly different from one another. This result remains the same whether we use only the coding sequences, only the introns, or both combined (which can be accomplished in the case of MP owing to the assumption of character independence). The six trees all share the association of the two Nearctic subgroups, *affinis* and *pseudoobscura*, but leave unresolved whether *D. subsilvestris* clusters with *D. bifasciata* or with the *obscura–ambigua* couplet (see Fig. 7).

The number of possible trees is 210 if we add *D. melanogaster* as an outgroup to the 30 unrooted trees, since the root can be in any of the seven branches of each tree. The ML test yields 48 trees superior to the rest, but not significantly different from one another. Of these 48 trees, 42 correspond to the six topologies

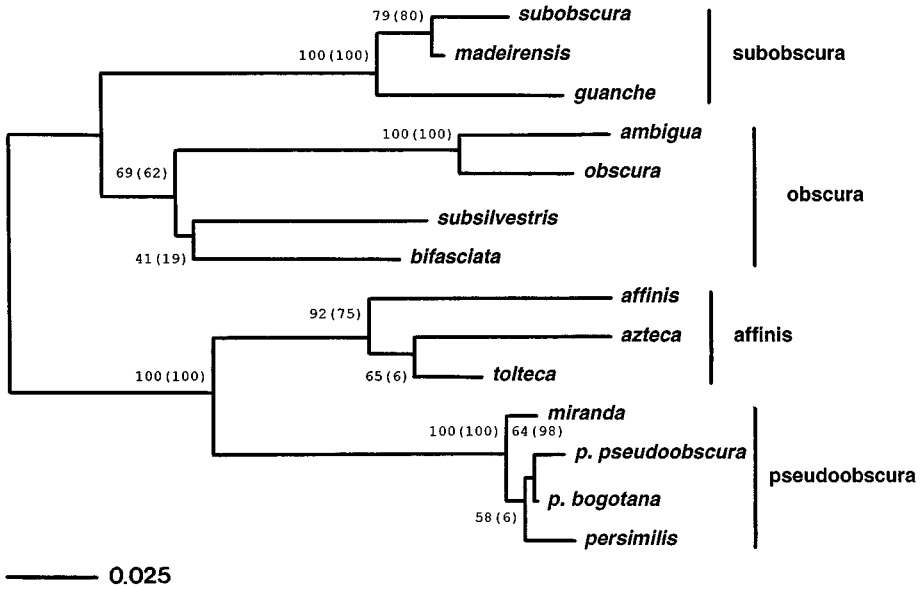

FIG. 5. Neighbor-joining tree based on Gpdh intron sequences. Branch lengths are proportional to the scale given in substitutions per nucleotide. Parent bootstrap values and the complete-and-partial bootstrap values (in parentheses) based on 1000 pseudoreplicates are given on the nodes.

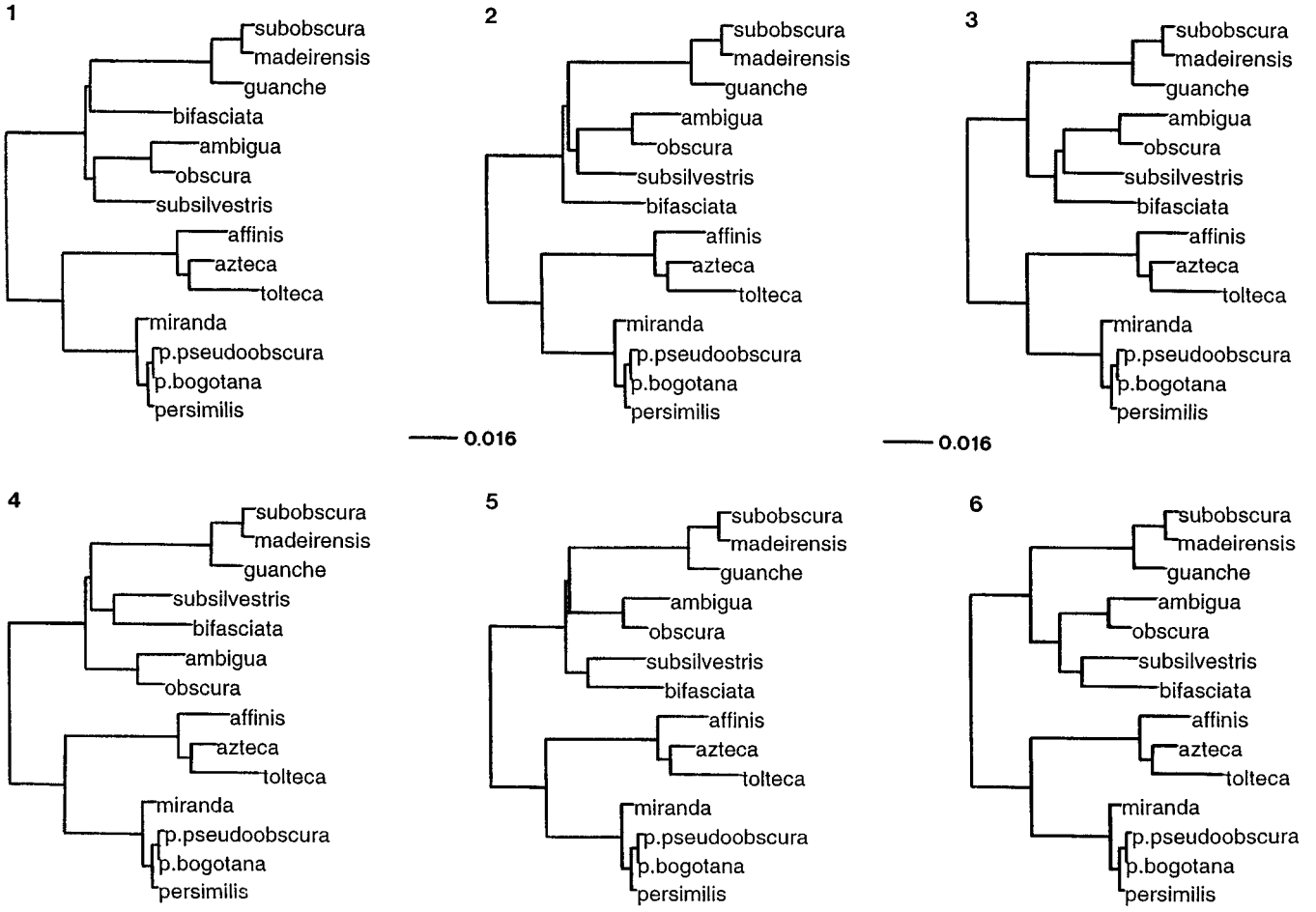

FIG. 7. Six *Gpdh* unrooted topologies that are not significantly different according to the MP and ML tests, obtained using total sequences by the maximum-likelihood method. Branch lengths are proportional to the scale given in substitutions per nucleotide.

shown in Fig. 6, where the position of the root, determined by the outgroup *D. melanogaster*, is indicated by a black dot. The MP test yields 34 superior equivalent trees, 28 of which correspond to the six topologies shown in Fig. 6. We repeated the MP test using as outgroup each of five different species of the Sophophora subgenus (the subgenus to which the *obscura* species belong), namely *D. nebulosa*, *D. paulistorum*, *D. simulans*, *D. teissieri*, and *D. willistoni* (sequences from Kwiatowski *et al.*, in press). Most of the statistically equivalent trees again correspond to the six topologies depicted in Fig. 6.

Compositional bias may influence phylogenetic topology (Saccone *et al.*, 1993). In our sequences compositional bias appears primarily in the third-codon positions, which is where most of the substitutions occur. We have used Saccone *et al.*'s (1990, Eq. 21, p. 576) χ^2 test for "stationarity" that ascertains whether nucleotide frequencies at equivalent codon positions are constant in the extant species and their ancestors. The χ^2 values are shown in Table 2. Values above 1.5 (shown in bold) are thought to violate the stationarity conditions. On the whole, the stationarity condition is satisfied for the *obscura* group species. The only exception is *D. tolteca*, which may account for the position of this species (or the *affinis* subgroup) as the first branching clade in some trees, a possible consequence of convergence with outgroup species.

Phylogenetic Analyses of Sod Sequences

Sod sequences were obtained from 12 *D. obscura* species and 1 subspecies, the same for which the *Gpdh* region was sequenced except for *D. subsilvestris*. The sequences include 42 bp (of 66) of the coding region of the first exon, the intron (320–397 bp), and 300 bp (of 396) of the second exon's coding region (Fig. 2). As in the case of *Gpdh*, intron divergence between *D. melanogaster* (725 bp) and the *obscura* species is much too high and the alignment becomes uncertain. Thus, as before, we separately consider coding sequences and intron sequences. The outgroup *D. melanogaster* is only included in the coding sequence alignment.

The alignment of the 14 *Sod* coding sequences consists of 342 nucleotide sites, of which 91 (26.6%) are

TABLE 2

Pairwise Results of the "Stationarity" χ^2-Test of Saccone *et al.* (1990) Measuring Similarity between Compositional Patterns at Third-Codon Positions in *Gpdh*

Species	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. *melanogaster*															
2. *subobscura*	**12.157**														
3. *madeirensis*	**12.942**	0.016													
4. *guanche*	**10.590**	0.140	0.186												
5. *ambigua*	**8.745**	0.468	0.538	0.175											
6. *obscura*	**9.405**	0.277	0.333	0.132	0.037										
7. *subsilvestris*	**8.720**	0.319	0.412	0.117	0.037	0.027									
8. *bifasciata*	**8.704**	0.586	0.665	0.175	0.046	0.144	0.102								
9. *affinis*	**6.258**	1.031	1.238	0.537	0.304	0.443	0.265	0.246							
10. *azteca*	**5.878**	1.114	1.352	0.656	0.429	0.549	0.343	0.390	0.020						
11. *tolteca*	**2.908**	**3.232**	**3.684**	**2.690**	**2.139**	**2.302**	**1.911**	**2.177**	0.974	0.740					
12. *miranda*	**7.423**	0.555	0.723	0.271	0.180	0.211	0.088	0.196	0.082	0.098	**1.250**				
13. *p. pseudoobscura*	**8.439**	0.533	0.654	0.148	0.148	0.231	0.122	0.065	0.154	0.246	**1.845**	0.099			
14. *p. bogotana*	**8.022**	0.572	0.711	0.188	0.157	0.239	0.116	0.086	0.102	0.174	**1.633**	0.060	0.006		
15. *persimilis*	**8.451**	0.482	0.615	0.143	0.190	0.249	0.126	0.116	0.157	0.226	**1.754**	0.076	0.009	0.009	

Note. Values above 1.5 (bold numbers) do not satisfy the stationary condition.

variable and 45 (13.7%) are informative. Sixteen variable sites are first codon positions (6 are informative), 13 are second positions (4 informative), and 62 are third positions (35 informative). Twenty-six sites (13 first and 13 second codon positions) show nonsynonymous substitutions, corresponding to replacements in 20 amino acid sites. There are no amino acid indels.

In contrast to *Gpdh, Sod* sequences have an excess of G+C (overall content 80%), which is particularly large for third codon positions (where the adenine is particularly scarce; its frequency is 28.4, 31.5, and 5.5% at first, second, and third position, respectively). Because of this composition bias, we have estimated nucleotide divergence not only according to Kimura's (1980) two-parameter model (Table 3, above the diagonal), but also according to Tajima and Nei (1984), Tamura (1992), and Tamura and Nei (1993). We have used these various distances for constructing phylogenetic trees by the NJ method, but the resulting topologies were not different from those based on Kimura's method and thus will not be displayed.

Figure 8A represents the NJ trees derived from least-squares estimates of branch lengths based on Kimura's distances. The *D. obscura* group species are clustered in four divergent lineages: the couplet *D. ambigua–D. obscura*, the *D. bifasciata* lineage, the monophyletic *subobscura* subgroup, and a monophyletic cluster formed by the two Nearctic subgroups *affinis* and *pseudoobscura*, each of which forms, in turn, a monophyletic cluster. The branching sequence of these four lineages is, however, uncertain according to the CPB test.

TABLE 3

Pairwise Nucleotide Divergences between *Sod* Coding Sequences for All Substitutions (Upper Triangle) or Only Transversions (Lower Triangle) Estimated According to Kimura's (1980) Two-Parameter Model

Species	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. *melanogaster*		.1948	.2145	.1908	.1677	.1751	.1905	.2292	.2295	.2373	.2101	.2142	.2057	.2097
2. *subobscura*	.0722		.0208	.0208	.0610	.0642	.0644	.0968	.1003	.1104	.0609	.0641	.0673	.0641
3. *madeirensis*	.0790	.0118		.0360	.0706	.0738	.0741	.1000	.1035	.1136	.0705	.0737	.0770	.0737
4. *guanche*	.0722	.0059	.0179		.0609	.0641	.0643	.0900	.0934	.1034	.0737	.0770	.0802	.0770
5. *ambigua*	.0689	.0270	.0332	.0332		.0148	.0457	.0772	.0806	.0904	.0611	.0676	.0644	.0580
6. *obscura*	.0722	.0301	.0364	.0364	.0089		.0487	.0738	.0772	.0870	.0643	.0708	.0676	.0611
7. *bifasciata*	.0756	.0209	.0270	.0270	.0059	.0089		.0807	.0842	.0941	.0645	.0645	.0614	.0614
8. *affinis*	.1000	.0491	.0556	.0556	.0332	.0364	.0270		.0088	.0208	.0577	.0641	.0610	.0546
9. *azteca*	.0965	.0459	.0524	.0524	.0301	.0332	.0240	.0029		.0148	.0610	.0674	.0643	.0578
10. *tolteca*	.1036	.0524	.0589	.0589	.0364	.0395	.0301	.0089	.0118		.0705	.0771	.0739	.0674
11. *miranda*	.0825	.0332	.0395	.0395	.0240	.0270	.0179	.0332	.0301	.0364		.0059	.0088	.0029
12. *p. pseudoobscura*	.0825	.0332	.0395	.0395	.0240	.0270	.0179	.0332	.0301	.0364	.0000		.0088	.0088
13. *p. bogotana*	.0859	.0364	.0427	.0427	.0209	.0240	.0148	.0301	.0270	.0332	.0029	.0029		.0059
14. *persimilis*	.0859	.0364	.0427	.0427	.0209	.0240	.0148	.0301	.0270	.0332	.0029	.0029	.0000	

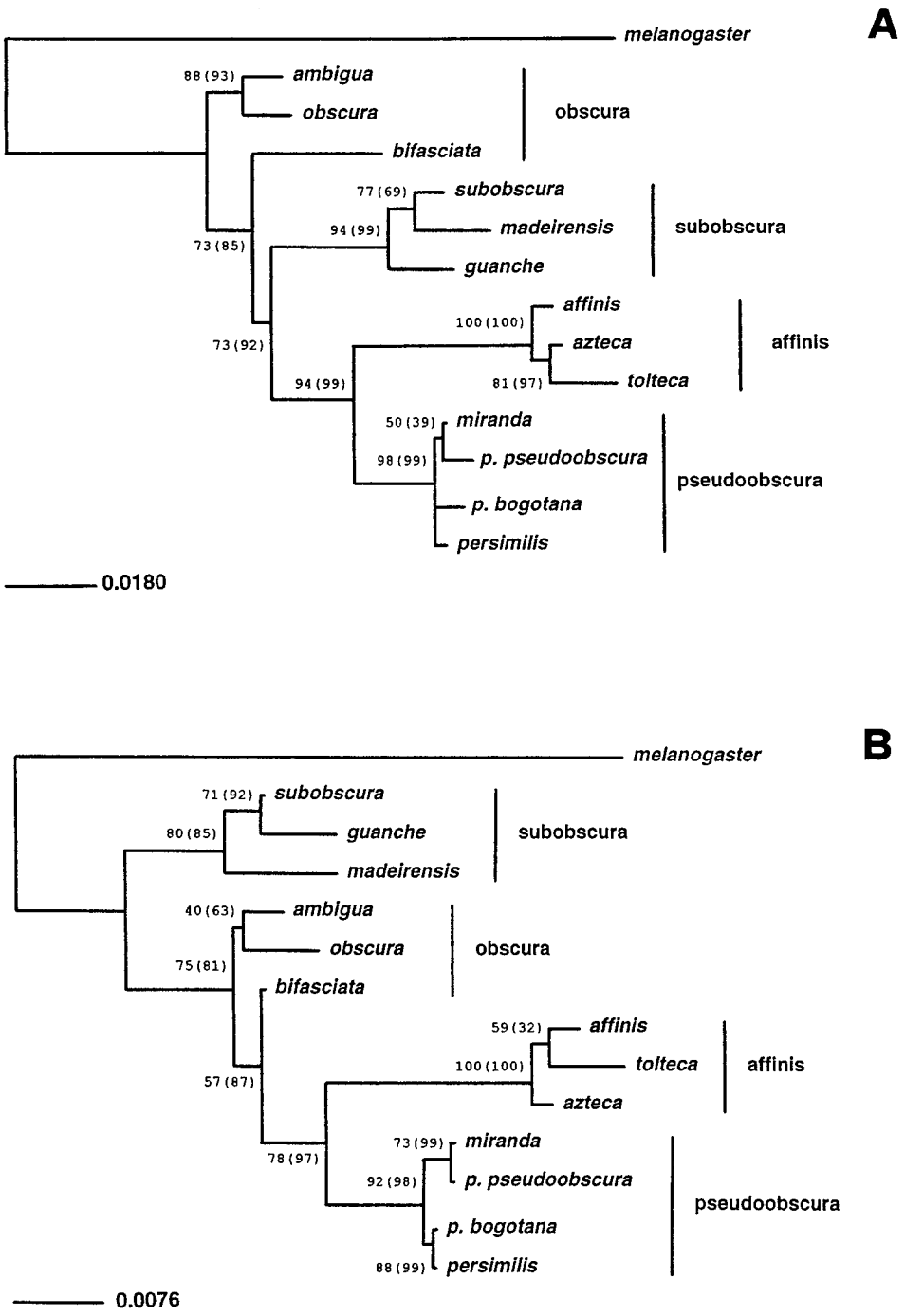

FIG. 8. Neighbor-joining trees based on *Sod* coding sequences, using all substitutions (A) or only transversions (B). Branch lengths are proportional to the scale given in substitutions per nucleotide. Percentage bootstrap values and complete-and-partial bootstrap values (in parentheses) on the nodes are based on 1000 pseudoreplicates.

To overcome the potential problem of saturation we have relied on only transversions (Table 3, below the diagonal). The NJ tree derived from the transversion divergences is shown in Fig. 8B. This tree again shows the two Nearctic subgroups as monophyletic sister clades. It differs from the previous (Fig. 8A) tree in that it shows the *subobscura* subgroup as the first clade to diverge, the *obscura* subgroup as a paraphyletic taxon with two lineages (*ambigua–obscura* and *bifasciata*), and the monophyletic Nearctic subgroups as a monophyletic clade, but without high CPB values.

In the case of protein-encoding genes, it is often useful to deal separately with synonymous and nonsynonymous substitutions, because the latter are more

likely to be selectively constrained and thus the phylogenetic signal is less likely to be erased by superimposed or back substitutions. There are 26 polymorphic nonsynonymous sites in the *Sod* data set, and a total of 45 nonsynonymous substitutions in those sites, which might warrant their separate use for phylogenetic purposes (only 5 nonsynonymous polymorphic sites appear in the *Gpdh* data set, with a total of 7 substitutions, which are obviously insufficient). Figure 9 shows the NJ trees, obtained by the unweighted method of Nei and Gojobori (1986), based on the nonsynonymous (9A) and synonymous (9B) distances reported in Table 4. Both NJ trees yield implausible topologies. The synonymous topology shown in Fig. 9B is very similar to the one in Fig. 8A, based on whole distances. The nonsynonymous tree (Fig. 9A) has low bootstrap values owing to the reduced number of substitutions, but shows *guanche* and *subobscura* (*D. subobscura* subgroups) as

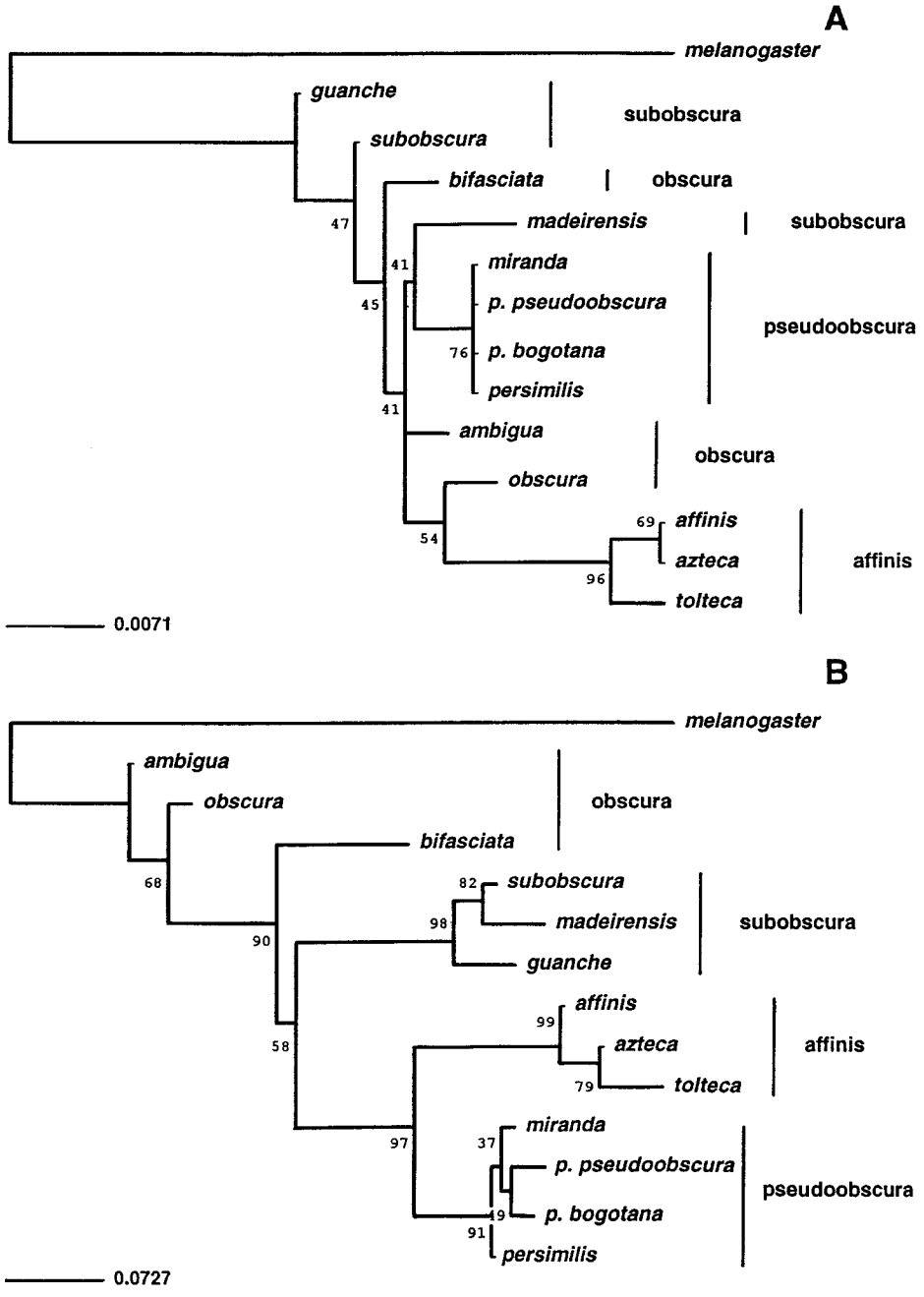

FIG. 9. Neighbor-joining *Sod* trees based on nonsynonymous substitutions (A) or synonymous (B) substitutions. The scales for branch lengths are substitutions per nonsynonymous (A) or synonymous (B) position. Percentage complete bootstrap values on the nodes are based on 1000 pseudoreplicates.

TABLE 4

Jukes-Cantor Corrected Proportion of Synonymous (Upper-Right Matrix) and Nonsynonymous (Lower-Left Matrix) Divergences between *Sod* Coding Sequences Estimated According to Nei and Gojobori's (1986) Unweighted Method

Species	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. *melanogaster*		.9448	.9448	1.029	.8071	.8621	.8071	.6184	.5654	.7786	.8621	.9023	.8240	.8621
2. *affinis*	.0965		.0382	.0646	.4015	.4009	.3403	.2849	.2856	.2948	.1819	.2142	.1979	.1662
3. *azteca*	.0965	.0000		.0381	.4230	.4224	.3602	.3033	.3041	.3135	.1979	.2309	.2142	.1819
4. *tolteca*	.0966	.0077	.0077		.4888	.4881	.4207	.3597	.3606	.3708	.2468	.2820	.2642	.2298
5. *subobscura*	.0729	.0233	.0233	.0233		.0515	.0787	.2672	.2678	.2859	.2658	.2838	.3022	.2838
6. *madeirensis*	.0879	.0273	.0273	.0273	.0116		.1067	.3033	.3041	.3231	.3018	.3206	.3399	.3206
7. *guanche*	.0687	.0272	.0272	.0273	.0038	.0155		.2495	.2502	.2678	.3210	.3403	.3602	.3403
8. *obscura*	.0771	.0193	.0193	.0194	.0115	.0154	.0154		.0386	.1840	.2668	.3033	.2849	.2492
9. *ambigua*	.0770	.0233	.0233	.0233	.0077	.0115	.0115	.0077		.1845	.2675	.3041	.2856	.2499
10. *bifasciata*	.0728	.0252	.0252	.0253	.0077	.0115	.0115	.0115	.0077		.2765	.2765	.2586	.2586
11. *miranda*	.0836	.0233	.0233	.0233	.0077	.0116	.0116	.0115	.0077	.0096		.0253	.0382	.0125
12. *p. pseudoobscura*	.0836	.0233	.0233	.0233	.0077	.0116	.0116	.0115	.0077	.0096	.0000		.0382	.0382
13. *p. bogotana*	.0836	.0233	.0233	.0233	.0077	.0116	.0116	.0115	.0077	.0096	.0000	.0000		.0253
14. *persimilis*	.0836	.0233	.0233	.0233	.0077	.0116	.0116	.0115	.0077	.0096	.0000	.0000	.0000	

the first and second species to diverge from the other *obscura* species.

The search for MP trees generated 6 equally parsimonious trees (results not shown), which have similar topologies. In all six trees, *D. obscura* and *D. ambigua* are the first to diverge, but they form a monophyletic clade in only two trees. A second radiation gives rise to three lineages corresponding to *D. bifasciata*, the *D. subobscura* subgroup, and the monophyletic cluster of the two Nearctic subgroups.

We have performed the ML analysis on 1st-plus-2nd and 3rd codon positions separately. The 1st-plus-2nd analysis yields a polytomous tree (Fig. 10A), showing *guanche* and *subobscura* as the first diverging species, just as in Fig. 9B. The ML tree based on 3rd codon positions (Fig. 10B) shows the unlikely relationships also manifest in the synonymous NJ tree (Fig. 10B).

Phylogenetic analyses of the intron sequences improve accuracy, but the lack of a reliable outgroup sequence makes moot the position of the root in the trees. Phylogenies obtained by different methods all result in similar topologies (Fig. 11). The two monophyletic Nearctic subgroups, *affinis* and *pseudoobscura*, form in turn a consistently monophyletic cluster in the MP (100% bootstrap replications) and NJ (100%) trees. The species pair *ambigua–obscura* and the *D. subobscura* subgroup are monophyletic clades (98–100% bootstrap). *D. bifasciata* clusters with the *D. subobscura* subgroup with low reliability on the basis of the intron sequences (CPB, 47%).

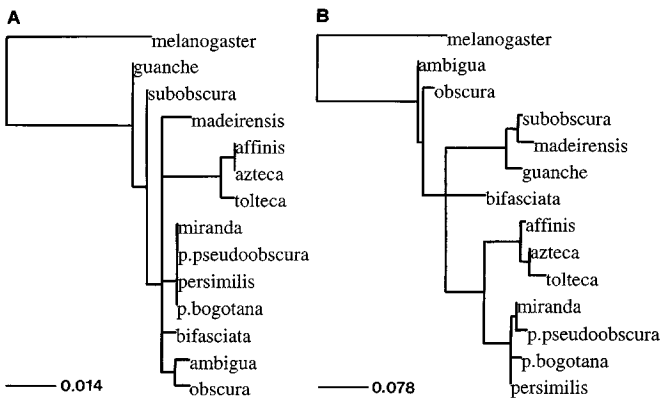

FIG. 10. Maximum likelihood *Sod* trees derived from 1st-plus-2nd (A) or 3rd (B) codon positions. The transition/transversion ratio is 1.68 in A and 1.26 in B. The scales are for substitutions per codon position.

Alternative trees representing all the branching orders among subgroups and lineages can be compared by the tests of Templeton (1983) and of Kishino and Hasegawa (1989). We tested different topologies of the set of 6 clades manifested by *Gpdh*, that is, the *affinis*, *pseudoobscura*, and *subobscura* subgroups, the couplet *ambigua–obscura*, *D. bifasciata*, and the outgroup *D. melanogaster*. There are 15 possible unrooted trees and 105 rooted trees. Of the 15 unrooted trees, 3 (Fig. 12) are superior to the others but not significantly different among them, according to either the MP or ML test for the *Sod* intron alignment, and also for the *Sod* coding region excluding the outgroup. The three trees can jointly be represented as the polytomous unrooted tree depicted in Fig. 13.

When *D. melanogaster* is included as an outgroup, there are 39 trees (of 105) statistically equivalent according to Kishino and Hasegawa's ML test, and only 10 according to Templeton's MP test. Fifteen ML trees and all 10 MP trees correspond to rooted trees derived from the 3 topologies depicted in Fig. 12, by locating the root on any branch (ML test) or on the branches connecting the Palearctic lineages (MP test). The other ML trees show either the *ambigua–obscura* pair or the

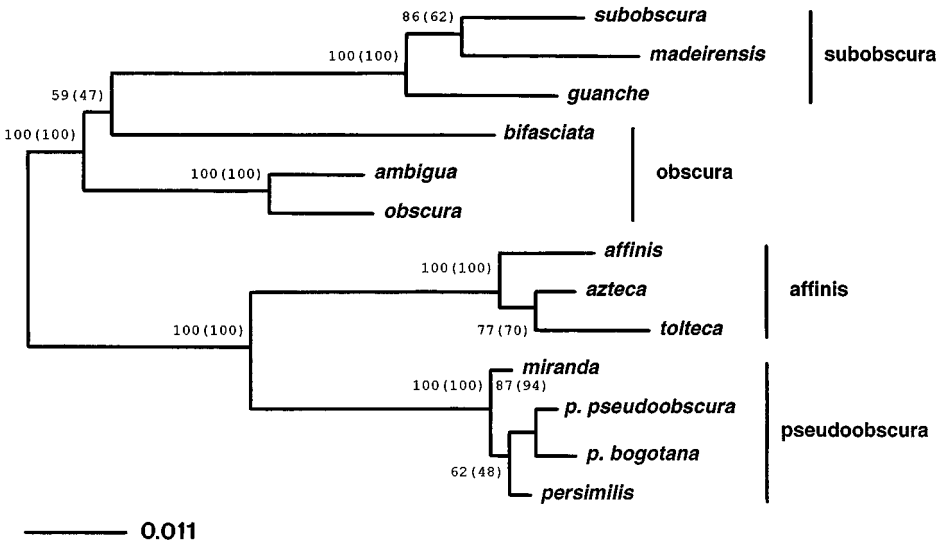

FIG. 11. Neighbor-joining trees based on *Sod* intron sequences. Branch lengths are proportional to the scale given in substitutions per nucleotide. Percentage bootstrap values and the complete-and-partial bootstrap values (in parentheses) on the nodes are based on 1000 pseudoreplicates.

subobscura subgroup as the first clade to diverge (results not shown). Similar results are obtained when other outgroup sequences are used, namely, *D. simulans*, *D. saltans*, or *D. willistoni* (sequence data from Kwiatowski *et al.*, 1994). Most rooted trees derived from the 3 unrooted trees connecting the same set of 4 clades (Nearctic species, *subobscura* subgroup, *ambigua–obscura* pair, and *D. bifasciata*) are statistically equivalent, according to Templeton's MP test. The unusual result showing *D. ambigua* as closer to the outgroups in some trees may be an artifact due to convergence. However, this cannot be due to differences in the *Sod* compositional constraints among the species of the *D. obscura* group, because *Sod* third codon positions satisfy the stationarity condition according to Saccone *et al.*'s test (results not shown).

DISCUSSION

The phylogenetic relationships of the *D. obscura* group species have been intensively studied using molecular and other approaches. Restriction analysis

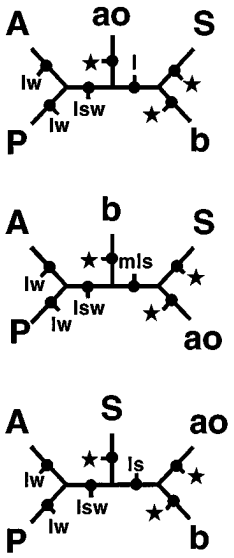

FIG. 12. Three *Sod* topologies that are statistically best, according to maximum parsimony (Templeton 1983) or maximum likelihood (Kishino and Hasegawa 1989). Species, subgroups, and other symbols are as described in the legend of Fig. 6, except that *saltans* (1) is now used as outgroup, but *nebulosa*, *paulistorum*, and *teissieri* are not.

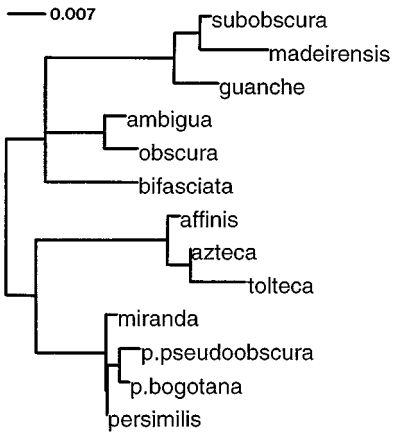

FIG. 13. Polytomous tree that incorporates the three equivalent unrooted topologies of Fig. 11. This unrooted tree is obtained using total *Sod* sequences by the maximum-likelihood method. Branch lengths are proportional to the scale given in substitutions per nucleotide.

and sequencing of mitochondrial DNA have been useful for deciphering phylogenetic relationships within subgroups but not between subgroups. Barrio *et al.* (1994) have suggested that the difficulties in solving the phylogeny of the subgroups could be due either to their rapid phyletic radiation, or to the distinctive evolutionary dynamics of the mitochondrial genome in *Drosophila*.

We have now analyzed two nuclear gene regions, *Sod* and *Gpdh*, both of which yield consistent phylogenetic interpretations. The *obscura* group appears as a monophyletic clade consisting of several lineages: (1) a Nearctic cluster with two monophyletic subgroups, *pseudoobscura* and *affinis;* (2) the *D. subobscura*-related species (*subobscura* subgroup according to Barrio *et al.*, 1994); (3) the lineage of the sister taxa *D. obscura* and *D. ambigua*, and (4) the lineage (or lineages) of *D. bifasciata* and *D. subsilvestris*. Our attempts to determine the branching order of these four (or five, if we separate *bifasciata* and *subsilvestris*) lineages fail, owing to short internode lengths which result in levels of sequence divergence among the taxa that are not differentiable (Figs. 7 and 13; Tables 1 and 3).

The statistical instability of the outgroups in the trees based on the *Sod* and *Gpdh* coding regions indicates that the phylogenies are affected by stochastic noise. The essential role of the enzymes encoded by both genes subjects them to intense purifying selection. This is clearly observed in the *Gpdh* gene, where almost all variable sites correspond to synonymous codon positions (96.7%), but also, in the *Sod* gene (71.4% synonymous substitutions). Saturation and parallel and back mutations may have randomized the sequences with respect to the common phylogenetic history that the *obscura* species share with the outgroups (which are the closest outgroups available, and members of the same *Sophophora* subgenus; see Throckmorton, 1978; Kwiatowski *et al.*, 1994). Randomization could also have clouded the relationships between subgroups and lineages, especially because they are products of rapid phyletic radiation. Clouding would occur even though the examination of the skewness of the tree-length distributions (Hillis and Huelsenbeck, 1992) for both gene regions (data not shown), and the resolution of the phylogenetic relationships between species within subgroups, indicate that there is a strong phylogenetic signal.

Rapid phyletic radiation and/or a saturation effect may account for the apparent impossibility to resolve the early phylogeny of the *D. obscura* group. The instability of the outgroups and the convergent compositional patterns observed in some species argue in favor of the randomization of the sequences over time. But the persistent inability to resolve the phylogeny with either nuclear or mitochondrial genes supports the hypothesis of rapid phyletic radiation (Hoelzer and Melnick, 1995). Very likely we are confronting both phenomena, randomization of the phylogenetic signal and short internode time lapses.

Constrained positions and conservative changes are thought to be particularly reliable in the phylogenetic analysis of DNA sequences, because their phylogenetic signal is less likely erased by parallel and back mutations and saturation (Swofford and Olsen, 1990, pp. 497–500; Hillis *et al.*, 1993; Miyamoto *et al.*, 1994). When we rely on conservative changes (transversions in the NJ and MP analyses of both genes; or, for *Sod*, NJ nonsynonymous substitutions or ML 1st-plus-2nd positions), or assume that transitions are much more frequent than transversions in the ML analyses of *Gpdh*, the phylogenetic reconstructions show the *D. subobscura* subgroup as the first taxon to diverge. These trees are not statistically robust but both genes yield the same result. We thus tentatively propose that the *D. subobscura* subgroup may have been the first to diverge from the rest. This hypothesis is consistent with the proposal that the metaphase configuration of five acrocentric chromosomes is the ancestral one in *Drosophila* (Buzzati-Traverso and Scossiroli, 1955). *D. subobscura, D. guanche,* and *D. madeirensis* have five acrocentric chromosomes (Lakovaara and Saura, 1982), whereas the prevailing condition in most *obscura* species is one metacentric X chromosome plus three acrocentric autosomes.

We hypothesize that the *obscura* group species derive from two main radiations. One resulted from adaptation to new habitats that appeared during the expansion of temperate forests in the Old World. This radiation gave rise to the *D. subobscura* subgroup, several *D. obscura* subgroup lineages, and the ancestor of the Nearctic species. Products of this first radiation may have also been the ancestors of the *D. microlabis* subgroup (Cariou *et al.*, 1988) and perhaps of several rare species of unclear taxonomic position, like *D. alpina* and *D. helvetica* (Lakovaara and Saura, 1982), and *D. inexspectata* (Tsacas, 1988). A second radiation took place during the colonization of deciduous forests of the New World by the descendants of a single colonizing lineage that gave origin to the two Nearctic subgroups, *affinis* and *pseudoobscura* (Goddard *et al.*, 1990; Beckenbach *et al.*, 1993).

In conclusion, our analysis of the *Gpdh* and *Sod* gene sequences supports two previous proposals: (1) the division of the paraphyletic former *D. obscura* subgroup in, at least, two subgroups (*obscura* and *subobscura*) (Barrio *et al.*, 1994, based on mitochondrial DNA sequences); and (2) the monophyletic origin of the Nearctic species, including both the *affinis* and *pseudoobscura* subgroups (based on DNA–DNA hybridization, Goddard *et al.*, 1990; and mitochondrial COII sequences, Beckenbach *et al.*, 1993).

ACKNOWLEDGMENTS

This work has been supported by a postdoctoral fellowship to E.B. from the "Programa General de Formación del Profesorado, Ministerio de Educación y Ciencia," Spain, and by the National Institutes of Health Grant GM42397 to F.J.A. We are grateful to Douglas W. Skarecky for technical advice and to Brigita Fody, Nelsson Becerra, and Julissa Aguilar for valuable help. Computer analyses were carried out using the "Servei de Bioinformàtica" (University of Valencia, Spain), supported by Grant PB93-0034. We are indebted to Professor W.-H. Li for the NJBOOTK2 program.

REFERENCES

Bachmann, L., Muller, E., Cariou, M. L., and Sperlich, D. (1992). Cloning and characterization of *KM190*, a specific satellite DNA family of *Drosophila kitumensis* and *D. microlabis*. *Gene* **120:** 267–269.

Barrio, E., Latorre, A., and Moya, A. (1994). Phylogeny of the *Drosophila obscura* species group deduced from mitochondrial DNA sequences. *J. Mol. Evol.* **39:** 478–488.

Barrio, E., Latorre, A., Moya, A., and Ayala, F. J. (1992). Phylogenetic reconstruction of the *Drosophila obscura* group, on the basis of mitochondrial DNA. *Mol. Biol. Evol.* **9:** 621–635.

Beckenbach, A. T., Wei, Y. W., and Liu, H. (1993). Relationships in the *Drosophila obscura* species group, inferred from mitochondrial cytochrome oxidase-II sequences. *Mol. Biol. Evol.* **10:** 619–634.

Bewley, G. C., Cook, J. L., Kusakabe, S., Mukai, T., Rigby, D. L., and Chambers, G. K. (1989). Sequence, structure and evolution of the gene coding for sn-glycerol-3-phosphate dehydrogenase in *Drosophila melanogaster*. *Nucleic Acids Res.* **17:** 8553–8567.

Brehm, A., and Krimbas, C. B. (1990). The phylogeny of 9 species of the *Drosophila obscura* group inferred by the banding homologies of chromosomal regions. 2. Element E. *Hereditas* **113:** 157–168.

Brehm, A., and Krimbas, C. B. (1992). The phylogeny of 9 species of the *Drosophila obscura* group inferred by the banding homologies of chromosomal regions. 3. Element D. *Genome* **35:** 1075–1085.

Brehm, A., and Krimbas, C. B. (1993). The phylogeny of 9 species of the *Drosophila obscura* group inferred by the banding homologies of the chromosomal regions. 4. Element C. *Heredity* **70:** 214–220.

Brehm, A., Krimbas, C. B., Sourdis, J., and Cariou, M. L. (1991). The phylogeny of 9 species of the *Drosophila obscura* group inferred by the banding homologies of chromosomal regions. 1. Element B. *Genome* **34:** 464–471.

Brown, W. M., Prager, E. M., Wang, A., and Wilson, A. C. (1982). Mitochondrial DNA sequences of primates: tempo and mode of evolution. *J. Mol. Evol.* **18:** 225–239.

Buzzati-Traverso, A., and Scossiroli, R. E. (1955). The *obscura* group of the genus *Drosophila*. *Adv. Genet.* **7:** 47–92.

Cabrera, V. M., González, A. M., Larruga, J. M., and Gullón, A. (1983). Genetic distance and evolutionary relationships in the *Drosophila obscura* group. *Evolution* **37:** 675–689.

Cariou, M. L., Lachaise, D., Tsacas, L., Sourdis, J., Krimbas, C. B., and Ashburner, M. (1988). New African species in the *Drosophila obscura* species group: genetic variation, differentiation and evolution. *Heredity* **61:** 73–84.

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.* **17:** 368–376.

Felsenstein, J. (1985). Confidence limits on phylogenies with a molecular clock. *Syst. Zool.* **34:** 152–161.

Felsenstein, J. (1993). PHYLIP: phylogeny inference package, v. 3.5c. University of Washington, Seattle.

Fitch, W. M. (1971). Toward defining the course of evolution: minimum change for a specified tree topology. *Syst. Zool.* **20:** 406–416.

Fridovich, I. (1986). Superoxide dismutases. *Adv. Enzymol.* **58:** 61–97.

Goddard, K., Caccone, A., and Powell, J. R. (1990). Evolutionary implications of DNA divergence in the *Drosophila obscura* group. *Evolution* **44:** 1656–1670.

González, A. M., Hernández, M., Volz, A., Pestano, J., Larruga, J. M., Sperlich, D., and Cabrera, V. M. (1990). Mitochondrial DNA evolution in the *obscura* species subgroup of *Drosophila*. *J. Mol. Evol.* **31:** 122–131.

Higgins, D. G., and Sharp, P. M. (1988). CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene* **73:** 237–244.

Hillis, D. M., Allard, R. W., and Miyamoto, M. M. (1993). Analysis of DNA sequence data: Phylogenetic inference. *Methods Enzymol.* **224:** 456–487.

Hillis, D. M., and Huelsenbeck, J. P. (1992). Signal, noise, and reliability in molecular phylogenetic analysis. *J. Heredity* **83:** 189–195.

Hoelzer, G. A., and Melnick, D. J. (1995). Patterns of speciation and limits to phylogenetic resolution. *Trends Ecol. Evol.* **9:** 104–107.

Hudson, R. R., Bailey, K., Skarecky, D., Kwiatowski, J., and Ayala, F. J. (1994). Evidence for positive selection in the superoxide dismutase (*Sod*) region of *Drosophila melanogaster*. *Genetics* **136:** 1329–1340.

Kawasaki, E. S. (1990). Sample preparation from blood, cells, and other fluids. *In* "PCR Protocols: A Guide to Methods and Applications" (M. A. Innis, D. H. Gelfand, J. J. Sninsky, and T. J. White, Eds.), pp. 146–152, Academic Press, San Diego, CA.

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., and Hasegawa, M. (1989). Evaluation of the maximum likelihood estimate on the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. *J. Mol. Evol.* **29:** 170–179.

Kumar, S., Tamura, K., and Nei, M. (1993). MEGA: molecular evolutionary genetics analysis, v. 1.0. The Pennsylvania State University, University Park, Pennsylvania.

Kwiatowski, J., Patel, M., and Ayala, F. J. (1989). *Drosophila melanogaster* Cu,Zn superoxide dismutase gene sequence. *Nucleic Acids Res.* **17:** 1264.

Kwiatowski, J., Skarecky, D., Bailey, K., and Ayala, F. J. (1994). Phylogeny of *Drosophila* and related genera inferred from the nucleotide sequence of the Cu,Zn *Sod* gene. *J. Mol. Evol.* **38:** 443–454.

Kwiatowski, J., Skarecky, D., Hernández, S., Pham, D., Quijas, F., and Ayala, F. J. (1991). High fidelity of the polymerase chain reaction. *Mol. Biol. Evol.* **8:** 884–887.

Kwiatowski, J., Krawczyk, M., Jaworski, M., Skarecky, D., and Ayala, F. J. (1996). Erratic evolution of glycerol-3-phosphate dehydrogenase in *Drosophila*, *Chymomyza*, and *Ceratitis*. *J. Mol. Evol.* **43:** in press.

Lakovaara, S., and Keränen, L. (1980). Phylogeny of the *Drosophila obscura* group. *Genetika* **12:** 157–172.

Lakovaara, S., and Saura, A. (1982). Evolution and speciation in the *Drosophila obscura* group. *In* "The Genetics and Biology of *Drosophila*" (M. Ashburner, H. L. Carson, and J. N. Thompson, Eds.), pp. 1–59, Academic Press, New York.

Lakovaara, S., Saura, A., and Falk, C. T. (1972). Genetic distance and evolutionary relationships in the *Drosophila obscura* group. *Evolution* **26:** 177–184.

Latorre, A., Barrio, E., Moya, A., and Ayala, F. J. (1988). Mitochondrial DNA evolution in the *Drosophila obscura* group. *Mol. Biol. Evol.* **5:** 717–728.

Li, W.-H., and Zharkikh, A. (1995). Statistical tests of DNA phylogenies. *Syst. Biol.* **44:** 49–63.

Marinković, D., Ayala, F. J., and Andjelković, M. (1978). Genetic polymorphism and phylogeny of *Drosophila subobscura*. *Evolution* **32:** 164–173.

Miyamoto, M. M., Allard, R. W., Adkins, R. M., Janecek, L. L., and Honeycutt, R. L. (1994). A congruence test of reliability using linked mitochondrial DNA sequences. *Syst. Biol.* **43:** 236–249.

Moritz, C., Dowling, T. E., and Brown, W. M. (1987). Evolution of animal mitochondrial DNA: Relevance for population biology and systematics. *Annu. Rev. Ecol. Syst.* **18:** 269–292.

Nei, M., and Gojobori, T. (1986). Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* **3:** 418–426.

Ruttkay, H., Solignac, M., and Sperlich, D. (1992). Nuclear and mitochondrial ribosomal RNA variability in the *obscura* group of *Drosophila*. *Genetica* **85:** 131–138.

Saccone, C., Lanave, C., and Pesole, G. (1993). Time and biosequences. *J. Mol. Evol.* **37:** 154–159.

Saccone, C., Lanave, C., Pesole, G., and Preparata, G. (1990). Influence of base composition on quantitative estimates of gene evolution. *Methods Enzymol.* **183:** 570–583.

Sacktor, B. (1975). Biochemistry of insect flight. *In* "Insect Biochemistry and Function" (D. J. Candy and B. A. Kilby, Eds.), pp. 3–88, Wiley, New York.

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

Sturtevant, A. H. (1942). The classification of the genus *Drosophila* with the description of nine new species. *Univ. Texas Publ.* **4213:** 5–51.

Swofford, D. L., and Olsen, G. J. (1990). Phylogeny reconstruction. *In* "Molecular Systematics" (D. M. Hillis and C. Moritz, Eds.), pp. 411–501, Sinauer, Sunderland, MA.

Tajima, F., and Nei, M. (1984). Estimation of evolutionary distance between nucleotide sequences. *Mol. Biol. Evol.* **1:** 269–285.

Tamura, K. (1992). Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+C-content biases. *Mol. Biol. Evol.* **9:** 678–687.

Tamura, K., and 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, A. R. (1983). Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and apes. *Evolution* **37:** 221–224.

Throckmorton, L. H. (1975). The phylogeny, ecology, and geography of *Drosophila*. *In* "Handbook of Genetics" (R. C. King, Ed.), pp. 421–469, Plenum Press, New York.

Tominaga, H., Shiba, T., and Narise, S. (1995). Structure of *Drosophila virilis* glycerol-3-phosphate dehydrogenase gene and comparison with the *Drosophila melanogaster* gene. *Biochim. Biophys. Acta* **1131:** 233–238.

Tsacas, L. (1988). Une cinquième espèce africaine du groupe *Drosophila obscura* (*Diptera: Drosophilidae*). *Ann. Soc. Ent. Fr.* (*NS*) **24:** 331–335.

Tsacas, L., Cariou, M. L., and Lachaise, D. (1985). Le groupe *Drosophila obscura* en Afrique de l'Est, description de trois nouvelles espèces (*Diptera: Drosophilidae*). *Ann. Soc. Ent. Fr.* (*NS*) **21:** 413–424.