UC Irvine

UC Irvine Previously Published Works

Title

Phylogeny of Drosophila and related genera: conflict between molecular and anatomical analyses.

Permalink

<https://escholarship.org/uc/item/8h10f318>

Journal

Molecular phylogenetics and evolution, 13(2)

ISSN

1055-7903

Authors

Kwiatowski, J
Ayala, FJ

Publication Date

1999-11-01

DOI

10.1006/mpev.1999.0657

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

Phylogeny of *Drosophila* and Related Genera: Conflict between Molecular and Anatomical Analyses¹

Jan Kwiatowski*† and Francisco J. Ayala†

*Institute of Botany, Warsaw University, 00-478 Warszawa, Poland; and †Department of Ecology and Evolutionary Biology,
University of California, Irvine, California 92697-2525

Received August 6, 1998; revised March 10, 1999

***Drosophila* species are extensively used in biological research; yet, important phylogenetic relationships within the genus and with related genera remain unresolved. The combined data for three genes (*Adh*, *Sod*, and *Gpdh*) statistically resolves outstanding issues. We define the genus *Drosophila* inclusively so as to include *Scaptomyza* and *Zaprionus* (considered distinct genera in the taxonomy of Wheeler, 1981) but excluding *Scaptodrosophila*. The genus *Drosophila* so defined is monophyletic. The subgenus *Sophophora* (including the *melanogaster*, *obscura*, and *willistoni* groups) is monophyletic and the sister clade to all other *Drosophila* subgenera. The Hawaiian *Drosophila* (including *Scaptomyza*) is a monophyletic group, but the subgenus *Drosophila* is not monophyletic, because the *immigrans* group is more closely related to the subgenus *Hirtodrosophila* than to other species of the subgenus *Drosophila*, such as the *virilis* and *repleta* groups.** © 1999 Academic Press

Key Words: *Chymomyza*; *Drosophila*; Hawaiian *Drosophila*; *Scaptodrosophila*; molecular systematics; *Adh*; *Gpdh*; *Sod*

INTRODUCTION

The family Drosophilidae is among the most diverse of the Diptera, encompassing more than 2500 species (Wheeler, 1986). Species of this family are often used in many areas of contemporary biological research. Understanding the phylogenetic relationships among the species is crucial to many of these studies. However, despite extensive investigations, the taxonomy of the Drosophilidae remains controversial. Wheeler's (1981, 1986) standard classification is inconsistent with the

phylogenetic relationships among the species, whether the inferences are based on morphology (Throckmorton, 1975; Grimaldi, 1990) or molecular data (DeSalle, 1992a,b; Kwiatowski *et al.*, 1994, 1997; Pélandakis and Solignac, 1993; Powell, 1997; Remsen and DeSalle, 1998; Russo *et al.*, 1995; Tatarenkov *et al.*, 1999; Thomas and Hunt, 1993). Throckmorton (1975) concluded, based on morphological, behavioral, and biogeographical data, that the standard taxonomy implies numerous paraphyletic relationships but he did not advance a new taxonomy. Grimaldi (1990) analyzed cladistically 217 morphological characters in a representative set of 120 species and advanced a revised taxonomy for the family. Notable in Grimaldi's (1990) analysis is the proposal that the subgenera *Drosophila* and *Sophophora* are sister clades and the exclusion from the genus *Drosophila* of the previous subgenus *Hirtodrosophila*, the Hawaiian species of *Drosophila* (which he classified as the genus *Idiomyia*), and the genera *Zaprionus* and *Scaptomyza*. Grimaldi's (1990) proposal contrasts with Throckmorton's (1975), who proposed an early divergence of the subgenus *Sophophora*, which would be the sister group to a complex of taxa that included the subgenus *Drosophila*, as well as *Hirtodrosophila*, *Dorsilopha*, *Zaprionus*, *Scaptomyza*, and the Hawaiian *Drosophila*.

Introduction of molecular data into taxonomic discourse should help in resolving some perplexing questions but has so far left unresolved some key relationships. Generally, molecular data (Kwiatowski *et al.*, 1994, 1997; Pélandakis and Solignac, 1993; Remsen and DeSalle, 1998; Tatarenkov *et al.*, 1999; Thomas and Hunt, 1993) contradict in very many important aspects Grimaldi's revisions of the *Drosophila* phylogeny, while they often support Throckmorton's hypothesis. The molecular studies have, however, paid relatively less attention to several taxa, including *Zaprionus*, *Scaptomyza*, *Hirtodrosophila*, and *Dorsilopha*. We have obtained the nucleotide sequence of two genes, Cu,Zn superoxide dismutase (*Sod*)² and alcohol dehydrogenase (*Adh*) in several critical species. We have combined our data with preexisting sequences of the same

¹ The sequences reported in this paper have been submitted to the GenBank/EMBL Data Libraries under Accession Nos. AF021825–AF021832 and AF022218.

² Abbreviations used: *Adh*, alcohol dehydrogenase gene; *Gpdh*, glycerol-3-phosphate dehydrogenase gene; ML, maximum likelihood; MP, maximum parsimony; NJ, neighbor joining; nt, nucleotide; s.g., subgenus; *Sod*, superoxide dismutase gene.

two genes, as well as of glycerol-3-phosphate dehydrogenase (*Gpdh*), seeking to resolve the phylogeny of these and other taxa within the Drosophilidae.

MATERIALS AND METHODS

Species

Table 1 lists the 29 Drosophilidae species that we have investigated. Strains of *Chymomyza procnemis*, *D. mimica*, *D. immigrans*, *D. (Hirtodrosophila) pictiventris*, *D. (Scaptomyza) adusta*, and *D. (Zaprionus) tuberculatus* were obtained from the National Drosophila Species Stock Center at Bowling Green, Ohio; the rest were from cultures available in our laboratory. We list *Scaptomyza* and *Zaprionus*, classified as genera by Wheeler (1981), as well as *Hirtodrosophila*, *Dorsilopha*,

and *Engiscaptomyza* as *Drosophila* subgenera, following Tatarenkov *et al.* (1999), but *Scaptodrosophila* as a genus, following Grimaldi (1990) and Tatarenkov *et al.* (1999; see also Kwiatowski *et al.*, 1994, 1997; Remsen and DeSalle, 1998; Russo *et al.*, 1995).

DNA Preparation, Amplification, Cloning, and Sequencing

Genomic DNA from about 10 to 20 flies was prepared following the method of Kawasaki (1990). The amplification, cloning, and sequencing of *Sod* fragments used for most species are described elsewhere (Kwiatowski *et al.*, 1994). *Sod* fragments from *D. teissieri*, *D. paulistorum*, *D. nebulosa*, and *D. immigrans* were amplified with primers *N* (5'-CCTCTAGAAATGGTGGTTAAAGC-TGTNTGCGT-3') and *O* (5'-ACGGAAGTCTAGAAGG-

TABLE 1
The 31 Species Studied, with GenBank Accession Nos. for the Gene Sequences

Family	Genus	Subgenus	Group	Species	Accession no.		
					<i>Adh</i>	<i>Sod</i>	<i>Gpdh</i>
Drosophilidae	<i>Drosophila</i>	<i>Sophophora</i>	<i>melanogaster</i>	<i>melanogaster</i>	X78384	X13780	X14179
				<i>simulans</i>		X15685	L41248
				<i>teissieri</i>	X54118	<u>AF021828</u>	U47809
				<i>yakuba</i>	X54120		
				<i>subobscura</i>	X55545	U38233	U47877
			<i>obscura</i>	<i>pseudoobscura</i>	X62181	U47871	L41249
				<i>ambigua</i>	X54813	U47868	U47880
				<i>willistoni</i>	L08648	L13281	L37038
			<i>willistoni</i>	<i>paulistorum</i>		<u>AF021829</u>	L41648
				<i>nebulosa</i>		<u>AF021830</u>	L41250
						U37590	
				<i>saltans</i>			
				<i>virilis</i>	U26846	X13831	D10697
		<i>Drosophila</i>	<i>repleta</i>	<i>hydei</i>	X58694	U37714	L41650
				<i>mulleri</i> ^{Adh1}	X03048		
				<i>immigrans</i>	M97638	<u>AF021825</u>	
			<i>quinaria</i>	<i>guttifera</i>		<u>AF021826</u>	
				<i>mimica</i>	M60792	<u>AF022218</u>	
			Hawaiian ^b	<i>nigra</i>	M60793		
				<i>differens</i>	M36785		
				<i>heteroneura</i>	M36781		
				<i>picticornis</i>	M63392		
				<i>albovittata</i>	M80925		
		<i>Scaptomyza</i> ^a		<i>adusta</i>		<u>AF021827</u>	
		<i>Engiscaptomyza</i>		<i>crassifemur</i>	M60790		
		<i>Hirtodrosophila</i>		<i>pictiventris</i>	<u>AF021832</u>	AF021824	L41649
		<i>Dorsilopha</i>		<i>busckii</i>		U39445	R. S. Wells
		<i>Zaprionus</i> ^a		<i>tuberculatus</i>	X63955	AF021823	L37039
				<i>lebanonensis</i>	M97637	AF021822	R. S. Wells
				<i>amoena</i>		X61687	L36961
				<i>procnemis</i>	<u>AF021831</u>	AF021821	L41252
				<i>capitata</i>	Z30194	M76975	L36960
		<i>Scaptodrosophila</i> ^a					
		<i>Chymomyza</i>					
Tephritidae	<i>Ceratitis</i>						

Note. Sequences newly obtained in this study are underlined. The *busckii* and *lebanonensis* *Gpdh* sequences are courtesy of Dr. Wells and available on-line (see Materials and Methods).

^a *Scaptodrosophila* was classified by Wheeler (1981) as a subgenus of *Drosophila* but has been raised to genus by Grimaldi (1990). *Scaptomyza* and *Zaprionus* are classified as genera by Wheeler (1981); in this paper we shall refer to them, as well as to *Engiscaptomyza*, *Hirtodrosophila* and *Dorsilopha*, as subgenera within the genus *Drosophila*.

^b The Hawaiian *Drosophila* groups are: modified mouthparts (*mimica*); white tip scutellum (*nigra*); and picture wing (*differens*, *heteroneura*, *picticornis*).

GCTTTTGGGCTTTGCCACCTG-3'), resulting in 441 nt of coding sequence. The *Sod* of *D. (S.) adusta* was obtained with the previous *N* primer and a *C* primer, (5'-CTTGCTGAGCTCGTGTCCACCCTTGCCCAGATCATC-3'), resulting in 345 nt of coding sequence.

The *Adh* gene fragments were 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 576-nt-long coding fragment from *C. procnemis* was obtained with primers *L3* (5'-GAACTGAAGGCAAT(CT)AATCC(AC)AA-3'), derived from a conserved protein region ELKAINPK, corresponding to the beginning of exon 2, and *R1* (5'-TTAGATGCC(GC)GA(AG)TCCCA(AG)TG(TC)T-TGGTCCA-3'), coming from the end of exon 3 WSKHWDSGI(STOP). The fragment of *D. pictiventris* was obtained with a single *R2* primer, shorter than *R1* by six bases from the 3'-end, by adjusting the temperature to 48°C, which produced a fragment starting from the end of intron 1 and ending at the end of exon 3 (642 nt of coding DNA). For sequencing, in addition to the PCR primers, standard M13 Uni and Rev primers were used, as well as primers specific for *Adh IL* (5'-GTGAC(CT)GG(CT)TT(CT)AATGCCAT-3') and *IR* (5'-ATGGCATT(AG)AA(AG)CC(AG)GTCAC-3').

Sequence Analysis

Neighbor-joining (NJ) trees (Saitou and Nei, 1987) were obtained with the MEGA 1.0 program (Kumar *et al.*, 1993). Trees of maximum parsimony (MP) (Fitch, 1971) and maximum likelihood (ML) (Felsenstein, 1981)

were obtained and tested according to Templeton (1983) and Kishino and Hasegawa (1989), respectively, using the PHYLIP 3.572c package programs DNAPARS and DNAML (Felsenstein, 1989). We give equal weights to all sites in MP trees and use empirical base frequencies and a transition/transversion ratio of 2 in the ML calculations. No saturation was observed for the third codon position for *Adh* and *Sod* genes (Tatarenkov *et al.*, 1999) or for synonymous nucleotide replacement in the *Gpdh* gene (Ayala *et al.*, 1996). Kimura's (1980) two-parameter distances were used for constructing NJ trees. Using other distance measures for NJ trees or changing the transition/transversion ratio in ML trees did not affect the topologies in any substantial way. Our alignments of the coding sequences are available online at <http://www.bot.uw.edu.pl/~jmkwiato/aln.html> or upon request from the first author. The bootstrap values are based on 1000 replications. The SEQBOOT, DNAPARS, and CONSENSE programs of PHYLIP were used to achieve this for the MP trees. We consider *Scaptodrosophila* as an outgroup for the *Drosophila* species used in this paper. There is ample evidence, both molecular (Kwiatowski *et al.*, 1994, 1997; Remsen and DeSalle, 1998; Russo *et al.*, 1995; Tatarenkov *et al.*, 1999) and morphological (Grimaldi, 1990), that this is the case with respect to *Drosophila* but not to *Chymomyza*.

RESULTS

Figure 1 shows a neighbor-joining (NJ) tree of *Adh*, which includes sequences available in the literature

FIG. 1. NJ tree with Kimura two-parameter distances for 21 *Adh* sequences. Numbers above branches are percentage bootstrap values based on 1000 replications.

plus those of *D. (Hirtodrosophila) pictiventris* and *Chymomyza procnemis* obtained in this study. This tree is essentially congruent with the tree obtained previously for 42 Drosophilidae species using similar methods (Russo *et al.*, 1995). As shown in Fig. 1, *Scaptodrosophila* and *Chymomyza* are sister clades to a complex clade that includes all other Drosophilidae. The early divergence of *Chymomyza* and *Scaptodrosophila* is not herein robust but it has been confirmed in previous molecular analyses (Kwiatowski *et al.*, 1994, 1997; Remsen and DeSalle, 1998; Tatarenkov *et al.*, 1999). The rest of the species split into two clades, one of which includes all species of the subgenus *Sophophora* (groups *melanogaster*, *obscura*, and *willistoni*). The other clade is composed of the subgenus *Drosophila*, including the Hawaiian *Drosophila* (*Idiomyia*, sensu Grimaldi, 1990), as well as *Scaptomyza*, *Zaprionus*, and *Hirtodrosophila* (*D. pictiventris*). The bootstrap value supporting this clade is 76%. The maximum likelihood and maximum parsimony methods give identical trees, which differ from the tree in Fig. 1 only in that *Zaprionus* splits off first and then *D. immigrans* and *D. pictiventris* branch off as a pair from the rest of the clade. The monophyly of the clade containing *Zaprionus*, *Hirtodrosophila*, *Scaptomyza*, and the subgenus *Drosophila* is supported in the MP analysis by 66% occurrences. Our results differ from those of an analysis of a shorter fragment of the *Adh* gene with fewer species (DeSalle, 1992a), which places *Hirtodrosophila* as a sister clade to the complex of the *Sophophora* and *Drosophila* subgenera. This shorter *Adh* fragment is not available in GenBank, although it has, apparently, been used in a later analysis that places *Hirtodrosophila* together with *Drosophila* s.g. (sensu lato), leaving *Sophophora* outside (Remsen and DeSalle, 1998, Fig. 2A). Therefore, despite some claims (DeSalle and Grimaldi, 1992; Powell and DeSalle, 1995), the only molecular evidence supporting the position of *Hirtodrosophila* outside the *Drosophila* genus seems to come from only one set of mitochondrial sequences (DeSalle, 1992a). The early divergence of the *Sophophora* subgenus from the rest of the genus *Drosophila* is, nevertheless, consistent with other results (Kwiatowski *et al.*, 1997; Pélandakis and Solignac, 1993; Remsen and DeSalle, 1998; Russo *et al.*, 1995). This has been, moreover, recently corroborated by the occurrence of a single-codon deletion in *Ddc* (dopa-decarboxylase) in all *Sophophora* species but not in any other *Drosophila* species or Drosophilidae genera (Tatarenkov *et al.*, 1999).

A controversial relationship revealed by the *Adh* phylogeny is the close association of all Hawaiian *Drosophila* (*Idiomyia*), *Scaptomyza*, and *Engiscaptomyza* (*D. crassifemur*) with the rest of the *Drosophila* subgenus (Thomas and Hunt, 1993; Russo *et al.*, 1995). This particular association is favored by Throckmorton (1975) and by a mtDNA phylogeny (DeSalle, 1992a,b)

but has been challenged by Grimaldi (1990). DeSalle (1992a), however, favors the monophyly of the subgenus *Drosophila* as sister clade to the Hawaiian *Drosophila*, contrary to Throckmorton (1975) and to the phylogeny in Fig. 1, which shows *D. immigrans* outside a clade that includes most other *Drosophila* s.g. species as well as the Hawaiian *Drosophila*. We have obtained *Sod* sequences of the two Hawaiian species, *D. (S.) adusta* and *D. mimica*, as well as *D. immigrans*, seeking additional evidence on these issues (see below).

Figure 2 displays an NJ tree based on *Sod* sequences. This tree is very similar to the tree in Fig. 1 and it is identical to the *Adh* trees obtained with the ML and MP methods, with respect to the species being represented. The *Sod* tree confirms a close association of *Scaptomyza* with the rest of the Hawaiian *Drosophila* (represented here by *D. mimica*) as well as the association of *Hirtodrosophila* (*D. pictiventris*) with *D. immigrans*, which together with *Zaprionus* form a clade with the Hawaiian *Drosophila*, *Scaptomyza*, and the other species of the subgenus *Drosophila*. However, the bootstrap values supporting these particular relationships are not satisfactory, which is the same situation obtained with 28S rRNA (Pélandakis and Solignac, 1993) and *Gpdh* (Kwiatowski *et al.*, 1997). The *Sod* ML tree differs from the NJ tree in that *Zaprionus* splits after (*D. guttifera* (*D. immigrans*, *D. pictiventris*)). The MP tree is almost identical to the NJ tree in Fig. 2 but instead of the cluster (*D. guttifera* (*D. immigrans*, *D. pictiventris*)), the MP tree has ((*D. guttifera*, *D. immigrans*) *D. pictiventris*). However, the bootstrap values are even lower than in the NJ tree.

Figure 3 is an NJ tree obtained with the combined sequences. A potential benefit of combining data from several loci when testing phylogenetic hypotheses is that the phylogenetic signal weakly present in some genes becomes amplified (Baker and DeSalle, 1997). This approach improved the recovery of monophyletic groups within Drosophilidae even when incongruent data were combined (Remsen and DeSalle, 1998). The trees in Figs. 1 and 2 are very similar. Moreover, G + C content, codon usage, and Kimura's distances of *Sod* and *Adh* genes are similar for most of the species considered (Starmer and Sullivan, 1989; Kwiatowski *et al.*, 1992, 1994; Russo *et al.*, 1995). The bootstrap support for the (*D. virilis*, *D. hydei*) clade in Fig. 3 is 100%; for (*D. mimica*, *Scaptomyza*) it is 99%; for the clade of the above four, it is 99%; for (*D. immigrans*, *D. pictiventris*), it is 88%; and for the clade of all of them plus *Zaprionus*, it is 82%. The only unresolved question within this clade is the position of *Zaprionus*.

We have evaluated several competing phylogenies by applying the test of Templeton (1983) for MP; and the Kishino-Hasegawa test (Kishino and Hasegawa, 1989) for the ML trees. In particular, we are interested in the effect of placing *Sophophora* closer to *Drosophila* s.g. than *Zaprionus* and *Hirtodrosophila*. As expected, sev-

FIG. 2. NJ tree with Kimura two-parameter distances for 22 *Sod* sequences. Bootstrap values based on 1000 replications are above the branches.

eral similar hypotheses (Fig. 4, Trees 1–3) are, on the basis of the *Sod* and *Adh* sequences, equal by the Templeton test criteria, whether the *Adh* and *Sod* data are used separately or combined (Table 2). Two other hypotheses with *Drosophila* s.g. closer to *Zaprionus*

and *Hirtodrosophila* than to *Sophophora* have been proposed by Remsen and DeSalle (1998) (Fig. 4, Tree 4) and by Throckmorton (1975) (Fig. 4, Tree 5). The former (Tree 4) is best for the *Adh* data but is rejected by the *Sod* data. The latter hypothesis (Tree 5) is rejected by

FIG. 3. NJ tree with Kimura two-parameter distances for *Adh* and *Sod* sequences combined for 15 species. *Scaptomyza* refers to *adusta* for *Sod* but *albovittata* for *Adh*. *Chymomyza* is *C. amoena*. Bootstrap values based on 1000 replications are above the branches.

FIG. 4. Eight competing phylogenetic hypotheses. *Scaptodrosophila* is represented by *S. lebanonensis*; *Chymomyza* by *C. amoena*; *Scaptomyza* as in Fig. 3; *virilis/repleta* by *D. virilis* and *D. hydei*; *Zaprionus* by *D. tuberculatus*; *Sophophora* by a clade (*D. willistoni* (*D. melanogaster* (*D. subobscura*, *D. pseudoobscura*))); *Hawaiian Drosophila* by *D. mimica*; *Hirtodrosophila* by *D. pictiventris*.

the *Sod* and *Adh* data, either separately or combined. Similarly, the phylogenies having *Sophophora* closer to the *Drosophila* s.g. than *Zaprionus* and *Hirtodrosophila*, proposed by Grimaldi (1990), DeSalle (1992a), and Powell and DeSalle (1995) (Fig. 4, Trees 6–8), are all rejected by the Templeton test using the *Adh* and *Sod* data either separately or combined (Table 2). The ML tests give similar results.

We further assess the position of *Zaprionus* and *Hirtodrosophila* relative to *Sophophora* and *Drosophila* s.g. using DNA sequences of glycerol-3-phosphate dehydrogenase. Although this protein evolves in an erratic

way in the Drosophilidae family, the Kimura genetic distance of nucleotide sequences changes monotonically with time and the phylogeny based on *Gpdh* sequences (Kwiatowski *et al.*, 1997) produces a topology very similar to those of the *Sod* and *Adh* trees (Figs. 1–3), although *Hawaiian Drosophila*, *Scaptomyza*, and *D. immigrans* are not represented in the *Gpdh* data set. However, similar to the *Sod* tree (Fig. 2), the separation of the *Hirtodrosophila*, *Zaprionus*, and *Drosophila* s.g. cluster from *Sophophora* is not supported by high bootstrap values (Kwiatowski *et al.*, 1997).

TABLE 2
Statistical Evaluation of Eight Tree Topologies by
Means of the Templeton (1983) Test

Gene	Tree	Steps	Difference $\pm$ SE	Significantly worse than best tree?
<i>Adh</i>	1	1135	10 $\pm$ 5.5	No
	2	1127	2 $\pm$ 3.5	No
	3	1129	4 $\pm$ 5.8	No
	4	1125	—	Best
	5	1139	14 $\pm$ 7.1	Yes
	6	1181	56 $\pm$ 10.5	Yes
	7	1156	31 $\pm$ 9.4	Yes
	8	1145	20 $\pm$ 6.6	Yes
<i>Sod</i>	1	649	—	Best
	2	653	4 $\pm$ 3.2	No
	3	651	2 $\pm$ 3.5	No
	4	660	11 $\pm$ 4.8	Yes
	5	667	18 $\pm$ 6.6	Yes
	6	674	25 $\pm$ 9.0	Yes
	7	677	28 $\pm$ 8.1	Yes
	8	679	30 $\pm$ 6.6	Yes
<i>Adh</i> and <i>Sod</i>	1	1785	4 $\pm$ 5.3	No
	2	1781	—	Best
	3	1781	0 $\pm$ 5.7	No
	4	1786	5 $\pm$ 5.0	No
	5	1807	26 $\pm$ 9.5	Yes
	6	1856	75 $\pm$ 13.7	Yes
	7	1834	53 $\pm$ 12.7	Yes
	8	1825	44 $\pm$ 9.8	Yes

Note. Trees are numbered as in Fig. 4.

Figure 5 shows an NJ tree obtained with the combined sequences of *Sod* and *Gpdh*. The set of species belonging to the s.g. *Drosophila*, *Zaprionus*, *Dorsilopha* (*D. busckii*), and *Hirtodrosophila* (*D. pictiventris*) form a well-supported clade (86%), much better than when the *Sod* (Fig. 2, 49%) or *Gpdh* (Kwiatowski *et al.*, 1997) (54%) sequences are used separately. The monophyly of the s.g. *Sophophora* is not well supported, which may be attributed to the deep division between the *willistoni* group and the *melanogaster* + *subobscura* group of species (Kwiatowski *et al.*, 1994, 1997; Tatarenkov *et al.*, 1999). However, the outgroup status of *Chymomyza* and *Scaptodrosophila* relative to the rest of Drosophilidae is very well supported (95%), although the sequence of branching of the two genera remains unresolved. Although the sequences of two *Adh* genes are known for the medfly *Ceratitis*, they are very distinct from the *Drosophila* *Adh*. We, therefore, have not used herein the *Ceratitis* *Adh* genes as outgroups that might resolve the branching sequence of *Chymomyza* and *Scaptodrosophila* (Fig. 6). However, combining the *Adh*, *Sod*, and *Gpdh* sequences increases bootstrap support for a clade composed of *Zaprionus*, *Hirtodrosophila*, and *Drosophila* s.g., represented here by the *virilis/repleta* group, to a very high value of 97%. Because not all gene sequences are available for all

Drosophila species, combining the data results in a decrease of the species of the *Drosophila* s.g. in the sample. Additional work is therefore required in order to establish the position of *Zaprionus* and *Hirtodrosophila* relative to other *Drosophila* s.g. species. However, in the light of evidence presented here, the sister status of *Sophophora* relative to a clade composed of *Zaprionus*, *Hirtodrosophila*, and *Drosophila* s.g. seems to be firmly established.

DISCUSSION

The *Adh* sequences that we have analyzed manifest a deep split between two sets of taxa within the genus *Drosophila*. One clade includes the subgenus *Sophophora* (*melanogaster*, *obscura*, and *willistoni* groups). The second clade includes the other *Drosophila* subgenera but also species traditionally classified in separate genera, namely *Scaptomyza* and *Zaprionus*. Within this second clade there are three taxa that split first from the rest of the clade. These taxa are *Zaprionus*, *Hirtodrosophila* (*D. pictiventris*), and *D. immigrans*. This last species is traditionally included with the subgenus *Drosophila*, which is in the *Adh* phylogeny shown to be paraphyletic (as had been proposed by Throckmorton, 1975). The remainder of the species in the second clade form a well-defined monophyletic group consisting in turn of two well-defined clades. One clade includes the Hawaiian *Drosophila* (the group's modified mouthparts, white tip scutellum, and picture wing, as well as *Engiscaptomyza*) and *Scaptomyza*, which is thought to have originated in Hawaii, although it includes species endemic elsewhere. The other set includes two species groups traditionally included within the subgenus *Drosophila*, namely *virilis* and *repleta*. It has been proposed that the *Adh* gene is particularly useful in providing robust Drosophilidae phylogenies (Thomas and Hunt, 1993; Russo *et al.*, 1995; Powell, 1997). This is corroborated by our analysis.

There are a variety of reasons why a gene phylogeny, whether robust or not, may not well represent a species phylogeny (Brower *et al.*, 1996). Confidence, however, will tend to increase when separate unrelated genes yield similar phylogenies. The two genes that we have investigated, *Sod* (Fig. 2) and *Gpdh* (Kwiatowski *et al.*, 1997), separately or in combination with each other (Fig. 5) or with *Adh* (Figs. 3 and 6), consistently resolve certain phylogenetic relationships: (1) the outgroup status of *Chymomyza* and *Scaptodrosophila* relative to other Drosophilids (which are included within the genus *Drosophila*—sensu lato, Table 1); (2) the subgenus *Sophophora* as a sister clade to all other *Drosophila* taxa; (3) the early branching of *Zaprionus*, *Hirtodrosophila*, and *D. immigrans* relative to the remainder of the *Drosophila* clade; (4) the monophyly of all Hawaiian *Drosophila*, including the Hawaii-originated *Scapto-*

FIG. 5. NJ tree with Kimura two-parameter distances for *Gpdh* and *Sod* sequences combined from 18 species. Bootstrap values based on 1000 replications are above the branches.

myza; and (5) the monophyly of the *repleta* and *virilis* groups of the subgenus *Drosophila*.

The monophyly of the clade encompassing all *Drosophila* other than *Sophophora* (i.e., the lineages comprised by (3)–(5) above) is supported by other studies, such as Kwiatowski *et al.* (1994, *Sod*), Kwiatowski *et al.* (1997, *Gpdh*), and by other genes, such as *Ddc* (Tatarenkov *et al.*, 1999) and domains D1 and D2 of the 28S

rRNA (Pélandakis and Solignac, 1993). The clade *Hirtodrosophila*/*D. immigrans* ((3) above), which had been proposed by Throckmorton (1975), is not statistically supported by *Adh* alone but it also appears in the *Sod* phylogeny (Fig. 2) and is strongly supported (88%) by *Adh* and *Sod* combined (Fig. 3). The combination of *Sod* and *Gpdh* (Fig. 5) provides strong statistical support (86%) for the monophyly of the complex clade that

FIG. 6. NJ tree with Kimura two-parameter distances for *Adh*, *Gpdh*, and *Sod* sequences combined from 11 species.

FIG. 7. Summary tree supported by three nuclear genes, *Adh*, *Gpdh*, and *Sod*, and consistent with other nuclear-gene phylogenies but not with mitochondrial DNA analyses.

includes all the subgenera of *Drosophila* (sensu lato) other than *Sophophora*, as well as the monophyly of the whole genus (95%) in the inclusive sense that we are using in this paper.

The *Adh* or *Sod* data alone, as well as both combined, are sufficient to reject several phylogenetic hypotheses that have been proposed, namely those of Throckmorton (1975), Grimaldi (1990), DeSalle (1992), and Powell and DeSalle (1995); see Fig. 4 and Table 2. The hypothesis is not rejected by low margin by *Adh* data but is rejected by *Sod* and *Adh* + *Sod* data. The first three topologies shown in Fig. 4, however, are not statistically differentiable. Figure 7 shows a consensus tree that summarizes the data available for all three genes, *Adh*, *Sod*, and *Gpdh* (*Dorsilopha* is not represented in the *Adh* data, nor Hawaiian *Drosophila*, *Scaptomyza* and *D. immigrans* in the *Gpdh* data, Table 1). This phylogeny is consistent with the "total evidence" hypothesis of Remsen and DeSalle (1998) (Fig. 4, Tree 4), with one exception. The Remsen and DeSalle (1998) hypothesis places *Hirtodrosophila* (*D. pictiventris*) as a sister clade to (s.g. *Drosophila*, Hawaiian *Drosophila*), whereas we show *D. immigrans* as monophyletic with *Hirtodrosophila*, an association that has 88% bootstrap support when the *Adh* and *Sod* data are combined (Fig. 3). The remaining ambiguities in Fig. 7 concern the order of divergence between *Chymomyza* and *Scaptodrosophila* and between some *Drosophila* clades.

The phylogeny of Fig. 7 is also consistent with the results of Thomas and Hunt (1993), Pélandakis and Solignac (1993), Kwiatowski *et al.* (1994, 1997), Russo *et al.* (1995), and Tatarenkov *et al.* (1999) but not with mitochondrial DNA phylogenies (DeSalle, 1992a,b). It has been argued that mitochondrial DNA better reflects species phylogeny than nuclear genes (DeSalle and Giddings, 1986; Moore, 1995). The results of this

study do not warrant such claims, at least regarding *Drosophila* phylogeny, since the weight of evidence suggests otherwise.

ACKNOWLEDGMENTS

We are grateful to Marcos Antezana, Kevin Bailey, Michal Krawczyk, and Douglas Skarecky for valuable help.

REFERENCES

- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Smith, J. A., Seidman, J. G., and Struhl, K. (1987). "Current Protocols in Molecular Biology," Wiley, New York.
- Ayala, F. J., Barrio, E., and Kwiatowski, J. (1996). Molecular clock or erratic evolution? A tale of two genes. *Proc. Natl. Acad. Sci. USA* **93**: 11729–11734.
- Baker, R. H., and DeSalle, R. (1997). Multiple sources of character information and the phylogeny of Hawaiian drosophilids. *Syst. Biol.* **46**: 645–673.
- Brower, A. V. Z., DeSalle, R., and Vogler, A. (1996). Gene trees, species trees, and systematics: A cladistic perspective. *Annu. Rev. Ecol. Syst.* **27**: 423–450.
- DeSalle, R. (1992a). The origin and possible time of divergence of the Hawaiian Drosophilidae: Evidence from DNA sequences. *Mol. Biol. Evol.* **9**: 905–916.
- DeSalle, R. (1992b). The phylogenetic relationships of flies in the family Drosophilidae deduced from mtDNA sequences. *Mol. Phylogenet. Evol.* **1**: 31–40.
- DeSalle, R., and Giddings, L. V. (1986). Discordance of nuclear and mitochondrial DNA phylogenies in Hawaiian *Drosophila*. *Proc. Natl. Acad. Sci. USA* **83**: 6902–6906.
- DeSalle, R., and Grimaldi, D. (1992). Characters and the systematics of Drosophilidae. *J. Hered.* **83**: 182–188.
- 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, W. M. (1971). Toward defining the course of evolution: Minimum change for a specified tree topology. *Syst. Zool.* **20**: 406–416.
- Grimaldi, D. A. (1990). A phylogenetic, revised classification of the genera in the Drosophilidae (Diptera). *Bull. Am. Mus. Nat. Hist.* **197**: 1–139.
- 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.
- 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 of 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, Version 1.0," Pennsylvania State University, University Park, PA.
- Kwiatowski, J., Skarecky, D., Burgos, M., and Ayala, F. J. (1992). 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., 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., Krawczyk, M., Jaworski, M., Skarecky, D., and Ayala, F. J. (1997). Erratic evolution of glycerol-3-phosphate dehydrogenase in *Drosophila*, *Chymomyza*, and *Ceratitis*. *J. Mol. Evol.* **44**: 9–22.
- Moore, W. S. (1995). Inferring phylogenies from mtDNA variations: Mitochondrial-gene trees versus nuclear-gene trees. *Evolution* **49**: 718–726.
- Pélandakis, M., and Solignac, M. (1993). Molecular phylogeny of *Drosophila* based on ribosomal RNA sequences. *J. Mol. Evol.* **37**: 525–543.
- Powell, J. R. (1997). "Progress and Prospects in Evolutionary Biology: The *Drosophila* Model," Oxford Univ. Press, New York.
- Powell, J. R., and DeSalle, R. (1995). *Drosophila* molecular phylogenies and their uses. In "Evolutionary Biology" (M. K. Hecht, R. J. MacIntire, and M. T. Clegg, Eds.), Vol. 28, pp. 87–138. Plenum, New York.
- Remsen, J., and DeSalle, R. (1998). Character congruence of multiple data partitions and the origin of the Hawaiian Drosophilidae. *Mol. Phylogenet. Evol.* **9**: 225–235.
- Russo, C. A. M., Takezaki, N., and Nei, M. (1995). Molecular phylogeny and divergence times of Drosophilid species. *Mol. Biol. Evol.* **12**: 391–404.
- Saitou, N., and Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**: 406–425.
- Starmer, W. T., and Sullivan, D. T. (1989). A shift in the third-codon-position nucleotide frequency in alcohol dehydrogenase genes in the genus *Drosophila*. *Mol. Biol. Evol.* **6**: 546–552.
- Tatarenkov, A., Kwiatowski, J., Skarecky, D., Barrio, E., and Ayala, F. J. (1999). On the evolution of dopa decarboxylase (*Ddc*) and *Drosophila* systematics. *J. Mol. Evol.* **48**: 445–462.
- Templeton, A. R. (1983). Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and apes. *Evolution* **37**: 221–244.
- Thomas, R. H., and Hunt, J. A. (1993). Phylogenetic relationships in *Drosophila*: A conflict between molecular and morphological data. *Mol. Biol. Evol.* **10**: 362–374.
- Throckmorton, L. H. (1975). The phylogeny, ecology, and geography of *Drosophila*. In "Handbook of Genetics" (R. C. King, Ed.), pp. 421–459. Plenum, New York.
- Wheeler, M. R. (1981). The Drosophilidae: A taxonomic overview. In "The Genetics and Biology of *Drosophila*" (M. Ashburner, H. L. Carson, and J. N. Thompson, Jr., Eds.), Vol. 3a, pp. 1–97. Academic Press, New York.
- Wheeler, M. R. (1986). Additions to the catalog of the world's Drosophilidae. In "The Genetics and Biology of *Drosophila*" (M. Ashburner, H. L. Carson, and J. N. Thompson, Jr., Eds.) Vol. 3e, pp. 395–409. Academic Press, New York.