Molecular Phylogeny and Divergence Times of Drosophilid Species

Claudia A. M. Russo, Naoko Takezaki, and Masatoshi Nei

Institute of Molecular Evolutionary Genetics and Department of Biology, The Pennsylvania State University

The phylogenetic relationships and divergence times of 39 drosophilid species were studied by using the coding region of the *Adh* gene. Four genera—*Scaptodrosophila, Zaprionus, Drosophila,* and *Scaptomyza* (from Hawaii)—and three *Drosophila* subgenera—*Drosophila, Engiscaptomyza,* and *Sophophora*—were included. After conducting statistical analyses of the nucleotide sequences of the *Adh, Adhr* (*Adh*-related gene), and nuclear rRNA genes and a 905-bp segment of mitochondrial DNA, we used *Scaptodrosophila* as the outgroup. The phylogenetic tree obtained showed that the first major division of drosophilid species occurs between subgenus *Sophophora* (genus *Drosophila*) and the group including subgenera *Drosophila* and *Engiscaptomyza* plus the genera *Zaprionus* and *Scaptomyza.* Subgenus *Sophophora* is then divided into *D. willistoni* and the clade of *D. obscura* and *D. melanogaster* species groups. In the other major drosophilid group, *Zaprionus* first separates from the other species, and then *D. immigrans* leaves the remaining group of species. This remaining group then splits into the *D. repleta* group and the Hawaiian drosophilid cluster (Hawaiian *Drosophila, Engiscaptomyza,* and *Scaptomyza*). *Engiscaptomyza* and *Scaptomyza* are tightly clustered. Each of the *D. repleta, D. obscura,* and *D. melanogaster* groups is monophyletic. The splitting of subgenera *Drosophila* and *Sophophora* apparently occurred about 40 Mya, whereas the *D. repleta* group and the Hawaiian drosophilid cluster separated about 32 Mya. By contrast, the splitting of *Engiscaptomyza* and *Scaptomyza* occurred only about 11 Mya, suggesting that *Scaptomyza* experienced a rapid morphological evolution. The *D. obscura* and *D. melanogaster* groups apparently diverged about 25 Mya. Many of the *D. repleta* group species studied here have two functional *Adh* genes (*Adh-1* and *Adh-2*), and these duplicated genes can be explained by two duplication events.

Introduction

The Drosophilidae is one of the most diverse and widely distributed dipteran families. This family includes nearly 3,000 species, which are divided into 61 genera (Wheeler 1986). Among these, the genus *Drosophila* is most speciose and comprises 14 subgenera and more than 1,300 species (Wheeler 1986). However, the taxonomy of this genus has been controversial. For example, drosophilids in Hawaii were once classified as several non-*Drosophila* genera (*Antopocerus, Atelodrosophila, Nudidrosophila,* etc.) (Hardy 1965). However, Kaneshiro (1974, 1976) and Carson and Kaneshiro (1976) suggested that they should be classified into two groups, one group belonging to the genus *Drosophila* and the other to *Scaptomyza.* Later, Grimaldi (1990) proposed that the Hawaiian *Drosophila* species should be raised to a rank of genus called *Idiomyia.* Furthermore, the genus *Drosophila* is subdivided into many subgenera, species groups, species subgroups, and species complexes even if we exclude so-called semispecies and subspecies (Wheeler 1986). These classifications are also controversial (see, e.g., Pelandakis et al. 1991; Powell and DeSalle 1995).

One reason for this confusing status of the *Drosophila* taxonomy is the lack of knowledge of phylogenetic relationships of the species. Grimaldi (1990) constructed a comprehensive phylogenetic tree for many species of *Drosophila* and its related genera by using a cladistic analysis of morphological characters. However, the evolutionary pattern of morphological characters is usually very complex, so that it is important to reexamine any morphological tree by using DNA sequences of which the evolutionary pattern is much simpler. In fact, Grimaldi's tree is inconsistent with the trees obtained from molecular data in several aspects. For example, the genus *Scaptomyza* is placed outside the cluster of *Drosophila* species according to his tree, but

Key words: phylogeny, alcohol dehydrogenase gene, molecular clock, divergence times, *Drosophila, Scaptodrosophila, Scaptomyza, Zaprionus.*

Address for correspondence and reprints: Masatoshi Nei, Institute of Molecular Evolutionary Genetics and Department of Biology, The Pennsylvania State University, 328 Mueller Laboratory, University Park, PA 16802. E-mail: NXM2@PSUVM.PSU.EDU

Mol. Biol. Evol. 12(3):391–404. 1995.

0737-4038/95/1203-0004$02.00

molecular data place this genus within the *Drosophila* cluster (Beverley and Wilson 1985; DeSalle 1992*a*; Thomas and Hunt 1993). Of course, this does not mean that molecular data are always better than morphological data. Molecular phylogenies are also known to be subject to various sources of errors (Nei 1991), and the phylogenetic trees of drosophilids constructed from different parts of DNA are not necessarily consistent with each other.

One problem with the previous studies of drosophilid phylogenies is that the trees obtained were not subjected to rigorous statistical tests (see, e.g., Lattorre et al. 1988; Grimaldi 1990) or that when they were subjected their reliability was rather low (see, e.g., Pelandakis et al. 1991; DeSalle 1992*a*, 1992*b*; Pelandakis and Solignac 1993; Kwiatowski et al. 1994). One exception was Thomas and Hunt's (1993) tree based on the alcohol dehydrogenase (E.C. 1.1.1.1) gene (*Adh*) sequences, which showed a high statistical reliability. Unfortunately, they examined only 11 species of which 7 were Hawaiian drosophilids, and it remains unclear how the tree is affected when more species are added.

We have therefore decided to examine the drosophilid phylogeny more thoroughly using 42 *Adh* gene sequences. There are many species in which *Adh* sequence data are available, and the extent of sequence divergence seems to be appropriate for studying the *Drosophila* phylogeny. Other sequence data such as those for mitochondrial DNA and nuclear rRNA seem to be less informative than *Adh* sequence data except for some special purposes (see below).

The main purpose of this paper is to present the results of this phylogenetic study. We will also present our estimates of the times of divergence between different species or different species clusters based on our new statistical method (N. Takezaki, A. Rzhetsky, and M. Nei, unpublished data). Since there are several duplicate copies of the *Adh* gene in drosophilids, we will also examine the pattern and times of gene duplication events.

Material and Methods

The drosophilid species used in this study were determined by the availability of *Adh* gene sequences in the literature, yet they included those belonging to major *Drosophila* subgenera and some related genera. The total number of species examined was 39, whereas the total number of *Adh* sequences was 42 because some species had duplicated genes sequenced. In this paper we follow Wheeler's (1981, 1986) classification of species except for *Scaptodrosophila lebanonensis*. This species belongs to *Drosophila* in Wheeler's classification, but we followed Grimaldi's (1990) nomenclature because it is quite different from the other *Drosophila* species according to recent studies (see, e.g., Grimaldi 1990; DeSalle 1992*b*). The names of the taxa used, the sources of gene sequences, and the GenBank accession numbers are presented in the following list. In this paper we are primarily interested in the *Adh* (*Adh-1* and *Adh-2*) genes, but we also used the *Adhr* (*Adh*-related or *Adh*-dup) gene for determining the outgroup species.

I. Genus *Scaptodrosophila*. *S. lebanonensis** (Marfany and Gonzalez-Duarte 1990; X54814) (*Adhr*, Juan et al. 1994; X63716) (1).

II. Genus *Zaprionus*. *Z. tuberculatus** (Maruyama and Hartl 1991; X63955) (2).

III. Genus *Scaptomyza*. *Sc. albovittata** (Thomas and Hunt 1993; M80925) (3).

IV. Genus *Drosophila*

A. Subgenus *Drosophila*

1. *D. immigrans* group. *D. immigrans** (*Adh* and *Adhr*, Abalat and Gonzalez-Duarte 1993; M97638) (4).
2. *D. repleta* group. (a) *D. hydei* subgroup. *D. hydei* (*Adh-1** and *2**, Menotti-Raymond et al. 1991; X58694) (4). (b) *D. mulleri* subgroup. *D. buzzatii* species complex: *D. buzzatii* (*Adh-2*, Dorit, Ayala, and Gilbert 1991†; M62743) (5). *D. eremophila* species complex: *D. mettleri* (Yum et al. 1991; M57300) (6). *D. mulleri* species complex: *D. arizonae* (*Adh-2*, Dorit, Ayala, and Gilbert 1991†; M62741) (6); *D. mayaguana* (*Adh-2*, Dorit, Ayala, and Gilbert 1991†; M62742) (6); *D. mojavensis* (*Adh-1* and *2*, Atkinson et al. 1988; X12536) (6); *D. mulleri* (*Adh-1** and *2*, Fisher and Maniatis 1985; X03048) (6); *D. navojoa* (*Adh-1*, Weaver et al. 1989; X15585) (6); *D. wheeleri* (*Adh-2**, Dorit, Ayala, and Gilbert 1991†; M62851) (6).
3. (Hawaiian) Fungus-feeders group. *D. nigra* (Thomas and Hunt 1991; M60793) (3).
4. (Hawaiian) Modified mouth-parts group. *D. mimica** (Thomas and Hunt 1991; M60792) (3).
5. Hawaiian picture-winged group. (a) *D. adiastola* subgroup. *D. adiastola** (Thomas and Hunt 1991; M60791) (3). (b) *D. grimshawi* subgroup. *D. affinidisjuncta* (Rowan and Dickinson 1988; M37262) (3). (c) *D. planitibia* subgroup. *D. differens* (Rowan and Hunt 1991; M36785) (3); *D. heteroneura** (Rowan and Hunt 1991; M36781) (3); *D. picticornis* (Rowan and Hunt 1991; M63392) (3); *D. planitibia* (Rowan and Hunt 1991; M63390) (3); D. silvestris* (Rowan and Hunt 1991; M63291) (3).

B. Subgenus *Engiscaptomyza. D. crassifemur** (Thomas and Hunt 1991; M60790) (3).
C. Subgenus *Sophophora*
   1. *D. willistoni* group. *D. willistoni** (Anderson et al. 1993; L08648) (6).
   2. *D. obscura* group. (a) *D. obscura* subgroup. *D. ambigua* (Marfany and Gonzalez-Duarte 1991a; X54813) (1); *D. guanche* (*Adh* and *Adhr*, Marfany and Gonzalez-Duarte 1993; X60113) (7); *D. madeirensis* (*Adh* and *Adhr*, Marfany and Gonzalez-Duarte 1993; X60112) (8); *D. subobscura* (Marfany and Gonzalez-Duarte 1991b; M55545) (1). (b) *D. pseudoobscura* subgroup. *D. miranda* (Schaeffer and Miller 1991; M60998) (9); *D. persimilis* strain 178* (Schaeffer and Miller 1991; M60997) (9); *D. pseudoobscura* strain AH43* (Schaeffer and Miller 1991; M60979) (*Adhr*, Schaeffer and Miller 1992; X68166) (9).
   3. *D. melanogaster* group. (a) *D. melanogaster* subgroup. *D. melanogaster* species complex: *D. mauritiana* (Cohn and Moore 1988; M19264) (10); *D. melanogaster* strain Fl-F allele* (Kreitman 1983; M17833) (4); *D. sechellia* (Coyne and Kreitman 1985; X04672) (11); *D. simulans* (Laurie, Heath, Jacobson, and Thomson 1990†; M36581) (4). *D. yakuba* species complex: *D. orena* (Bodmer and Ashburner 1984; M37837) (2); *D. teissieri* (*Adh* and *Adhr*, Ashburner 1990†; X54118) (2); *D. yakuba** (Ashburner 1990†; X54120) (2). (b) *D. montium* subgroup. *D. tsacasi* (Maruyama and Hartl 1991; X63954) (2).

In the above list the numbers in parentheses refer to the geographical distributions of the species—that is, 1, Europe; 2, Africa; 3, Hawaii; 4, cosmopolitan; 5, South America; 6, southern United States through Brazil; 7, Canary Islands; 8, Madeira Island; 9, western North America; 10, Mauritius Island; and 11, Seychelles Islands. The *Adh* sequences that are marked with an asterisk (*) were used in the branch-and-bound search for maximum-parsimony trees. A dagger (†) refers to the authors, who have not written any paper about the sequence. *Adhr* stands for the *Adh*-related gene (or *Adh*-dup), which is an ancient duplicate copy of the *Adh* gene (Schaeffer and Aquadro 1987), whereas *Adh-1* and *Adh-2* are functional duplicate genes observed in the *D. repleta* group. All nucleotide sequences of *Adh* genes were obtained from the GenBank. We used the sequences for the coding region of the *Adh* and *Adhr* genes in this study, and they were aligned by using the previous information (Sullivan et al. 1990*b*) and by inspection. The *Adh* consensus alignment was 257 codons long. The *D. melanogaster* group species had 257 codons, *Zaprionus tuberculatus* 256, and the remaining species 255.

For the phylogenetic inference, we used the neighbor-joining (NJ) (Saitou and Nei 1987), minimum-evolution (ME) (Rzhetsky and Nei 1992), and maximum-parsimony (MP) methods. Construction of the NJ trees and their bootstrap tests were conducted by using the computer programs NJBOOT2 (K. Tamura) and MEGA (Kumar et al. 1993). The estimates of the branch lengths of NJ trees were determined by the least-squares method. MEGA was also used to compute evolutionary distances and nucleotide frequencies. Estimation of ME trees and statistical tests of ME and NJ trees were conducted by using the METREE program (Rzhetsky and Nei 1992). In ME and NJ trees the statistical confidence of a particular sequence cluster was evaluated by the confidence probability (CP) that the interior branch associated is positive (CP = 1 − Type 1 error), whereas the bootstrap confidence level (BCL) gives the percentage of bootstrap trees where the same interior branch (sequence partition) as that of the original tree appears (see MEGA manual, pp. 44–45). BCL often underestimates the statistical confidence (Zharkikh and Li 1992; Hillis and Bull 1993; Sitnikova et al. 1995), so BCL tends to be lower than CP. The MP trees were constructed by the PAUP program (Swofford 1993). A heuristic search was used for the entire data set (stepwise random addition with 50 replicates), whereas the branch-and-bound search (upper bound computed via stepwise with furthest addition) was used for the sequences with asterisks in the list given above.

Results

Phylogenetic Analyses of *Adh* Gene Sequences

In the present paper we are primarily interested in the phylogenetic tree based on *Adh* gene sequences without considering *Adhr* sequences, which are available for a limited number of species. However, *Adhr* sequences will be used for finding the root of the *Adh* gene phylogeny. The alignment of the 42 *Adh* sequences showed 418 (of 771) variable and 343 parsimony-informative nucleotide sites. The *Drosophila melanogaster* group had a two–amino acid insertion at the third and fourth amino acid positions, whereas *Zaprionus tuberculatus* showed one amino acid insertion at the fourth amino acid position. The average transition/transversion ratio was 1.1 or about twice as high as the random expectation. The lowest and highest ratios were observed between *D. subobscura* and *D. madeirensis* (0.33) and between *D. mulleri*-2 and *D. arizonae*-2 (3.25), respectively. The overall frequencies of the nucleotides A, T, C, and G were 0.238, 0.224, 0.286, and 0.251, respectively. Thus,

there was not much bias in G+C content. At third codon positions, however, there was an excess of G+C content (overall about 0.70), and the extent of the excess varied with species as noted by Thomas and Hunt (1991).

Estimates of Kimura's (1980) two-parameter distances, which were used for constructing NJ and ME trees, are presented in figure 1 for all pairs of 42 drosophilid *Adh* sequences. For this set of data, the topologies of the NJ and ME trees were the same except for some minor details concerning the branching pattern within the *D. repleta* group. The NJ tree with least-squares estimates of branch lengths is presented in figure 2. In this tree *Scaptodrosophila lebanonensis* is used as the outgroup for the reason that will be mentioned in the next section. NJ trees were also constructed by using the p distance, Jukes and Cantor's (1969) distance, Tajima and Nei's (1984) distance, and Tamura's (1992) distance (see Kumar et al. 1993), but the topology remained the same. This apparently occurred because the distance values were relatively small. (Note that the largest pairwise distance in fig. 1 is 0.31, so we decided not to use amino acid sequences.) The heuristic search for MP trees generated a consensus tree, which is very similar to the NJ tree. By contrast, the branch-and-bound search with the 17 species, as indicated earlier, produced a tree with the *D. obscura* group species clustered with *D. willistoni* rather than with the *D. melanogaster* group unlike all other topologies. (The reason for this is unclear, but it could be the low G+C content of the *D. obscura* and *D. willistoni* groups.) Since this topology was also inconsistent with the tree based on the superoxidase dismutase (*Sod*) gene (Kwiatowski et al. 1994), we disregarded this tree in the following.

According to the tree in figure 2, the first major division of drosophilid species occurs between subgenus *Sophophora* (genus *Drosophila*) and the group including subgenera *Drosophila* and *Engiscaptomyza* plus genera *Zaprionus* and *Scaptomyza*. Subgenus *Sophophora* is then divided into *D. willistoni* and the clade of *D. obscura* and *D. melanogaster* groups, each of which forms a monophyletic cluster of species. This branching pattern is the same as that obtained by DNA-DNA hybridization experiments (Powell and DeSalle 1995). In the other major drosophilid group, *Zaprionus* first separates from

	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42
1 S.lebanonensis	248	269	251	240	249	279	281	258	266	258	264	275	299	283	300	296	300	300	296	298	298	289	282	317	313	295	241	233	233	242	274	266	260	242	232	239	256	232	243	236	234
2 Z.tuberculatus		208	254	216	243	248	254	228	234	235	245	224	246	242	260	253	263	262	256	249	262	242	251	249	282	257	230	223	226	228	259	264	263	235	214	212	224	219	207	209	212
3 D.immigrans			237	215	220	235	226	209	211	215	226	212	226	215	235	231	237	237	228	231	237	224	221	264	262	290	250	246	250	251	274	274	272	272	257	266	270	268	263	257	259
4 D.wheeleri-2				42	52	73	70	59	57	59	126	94	119	121	203	198	205	204	205	208	211	206	196	249	246	293	249	245	246	261	263	273	271	257	256	259	259	253	248	248	244
5 D.mulleri-1					42	66	59	50	46	50	119	85	115	112	193	188	195	194	198	194	194	194	191	244	245	279	233	225	227	237	243	254	252	233	228	235	235	235	230	224	224
6 D.mulleri-2						56	66	43	53	63	127	86	119	115	197	196	202	202	202	202	206	202	206	253	251	284	234	228	230	248	253	263	261	256	248	246	249	247	248	240	240
7 D.mojavensis-1							70	60	59	80	141	98	127	128	219	215	222	222	224	215	222	218	215	260	264	288	265	259	262	274	278	284	280	281	270	268	270	264	266	262	264
8 D.mojavensis-2								61	36	73	128	92	121	119	203	199	206	206	211	202	207	202	206	256	262	291	259	254	256	265	268	278	272	283	267	276	278	270	268	264	264
9 D.navojoa-1									46	63	129	85	119	115	199	192	199	199	199	199	202	199	205	242	237	274	240	234	236	233	252	261	259	240	231	236	238	231	229	226	228
10 D.arizonae-2										54	116	79	109	107	191	186	193	192	200	194	200	190	202	246	239	265	245	236	238	248	256	265	259	267	248	259	261	253	252	248	248
11 D.mayaguana-2											115	70	102	104	201	196	203	203	203	203	201	194	197	258	252	286	255	251	255	263	271	281	279	261	254	267	267	257	254	250	250
12 D.mettleri												108	109	97	190	183	189	189	191	187	196	186	192	257	260	314	270	260	260	277	280	288	288	295	265	278	278	278	277	271	271
13 D.buzzatii-2													97	92	184	178	184	184	186	184	186	184	191	240	241	281	247	241	243	252	248	259	257	268	252	258	261	259	256	252	252
14 D.hydei-1														28	203	198	205	205	208	201	216	206	214	273	259	307	275	269	271	275	284	292	292	294	271	276	274	278	272	270	268
15 D.hydei-2															193	188	191	191	192	186	201	192	204	254	239	300	271	262	260	262	274	282	282	290	275	272	272	276	271	268	267
16 D.differens																5	16	16	38	54	64	57	97	228	203	297	283	281	282	305	300	310	308	310	288	298	294	298	300	292	290
17 D.planitibia																	11	11	32	49	59	51	91	221	201	299	277	276	276	299	294	303	302	301	280	292	288	292	292	284	282
18 D.heteroneura																		3	38	49	62	54	88	223	201	299	283	279	279	305	302	311	310	307	286	298	294	298	298	290	288
19 D.silvestris																			40	51	64	53	91	226	201	299	283	279	279	304	302	311	309	307	286	298	294	297	298	290	288
20 D.affinidisjuncta																				54	63	63	97	211	207	299	279	283	283	296	297	309	307	307	284	292	289	297	297	288	287
21 D.picticornis																					69	70	98	220	213	300	271	265	265	292	286	295	295	301	278	292	290	294	292	284	282
22 D.adiastola																						59	96	221	205	306	294	294	294	310	305	317	315	309	288	310	303	307	310	302	300
23 D.mimica																							79	214	188	294	269	271	273	290	288	301	299	299	274	278	276	280	290	280	280
24 D.nigra																								210	197	291	269	271	273	290	295	299	297	309	280	298	296	306	306	296	298
25 D.crassifemur																									98	289	294	292	298	274	305	310	304	288	269	271	288	286	292	282	284
26 Sc.albovittata																										290	294	290	294	286	311	314	308	304	285	299	311	303	307	301	305
27 D.willistoni																											202	204	206	232	241	243	237	270	249	259	261	256	259	249	251
28 D.miranda																												16	15	85	98	101	100	144	143	147	152	147	157	150	152
29 D.pseudoobscura																													5	86	91	94	92	147	147	152	157	152	162	155	157
30 D.persimilis																														89	94	97	95	144	147	149	154	149	159	152	154
31 D.ambigua																															98	101	99	145	133	147	165	142	152	145	149
32 D.guanche																																21	20	184	162	177	178	177	184	177	179
33 D.subobscura																																	5	182	167	177	182	177	184	177	179
34 D.madeirensis																																		181	166	176	180	176	182	175	177
35 D.tsacasi																																			102	90	98	92	95	86	90
36 D.orena																																				61	59	53	54	44	47
37 D.teissieri																																					21	50	51	44	44
38 D.yakuba																																						49	50	40	43
39 D.melanogaster																																							25	19	21
40 D.mauritiana																																								12	12
41 D.simulans																																									5
42 D.sechellia																																									

FIG. 1.—Pairwise Kimura two-parameter distances (1,000) for 42 *Adh* drosophilid sequences. All insertions/deletions were eliminated from the entire data set, and the distances were computed by using the remaining 765 nucleotides.

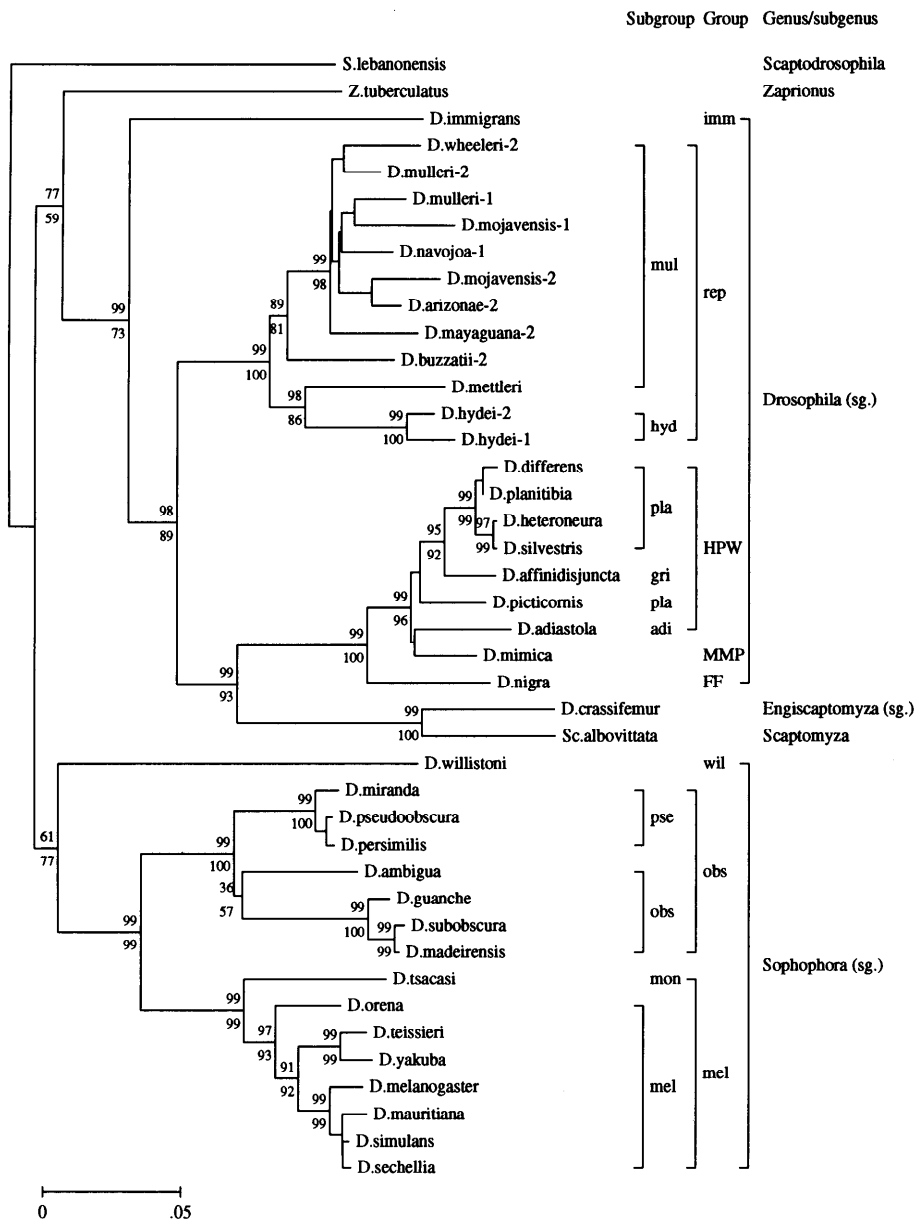

FIG. 2.—Neighbor-joining tree with Kimura two-parameter distances for 42 *Adh* nucleotide sequences. The confidence probability (CP) is shown above each interior branch tested, whereas the bootstrap confidence level (BCL) (from 1,000 replications) is shown below the branch. Even though we did not include the *Drosophila erecta Adh* sequence in this tree, it clusters with *D. orena* (CP of 97% and BCL of 95%). We used *Scaptodrosophila lebanonensis* as the outgroup for the reasons mentioned in the text. The abbreviations are as follows: mul, *D. mulleri;* hyd, *D. hydei;* pla, *D. planitibia;* gri, *D. grimshawi;* adi, *D. adiastola;* pse, *D. pseudoobscura;* obs, *D. obscura;* mon, *D. montium;* mel, *D. melanogaster;* imm, *D. immigrans;* rep, *D. repleta;* HPW, Hawaiian picture-winged; MMP, (Hawaiian) modified mouth-parts; FF, (Hawaiian) fungus feeders; wil, *D. willistoni;* 1, *Adh-1;* 2, *Adh-2;* sg., subgenus.

the other species, and then *D. immigrans* leaves the remaining group of species. This remaining group then splits into the *D. repleta* group and the Hawaiian drosophilid cluster (Hawaiian *Drosophila, Engiscaptomyza,* and *Scaptomyza*). Interestingly, *D. crassifemur* (subgenus *Engiscaptomyza*) and *Scaptomyza albovittata* form a tight cluster and clearly belong to the Hawaiian drosophilid cluster.

As mentioned earlier, Thomas and Hunt (1993) constructed a phylogenetic tree of 11 sequences from genera *Scaptodrosophila* and *Scaptomyza* plus *Drosophila* subgenera *Drosophila, Engiscaptomyza,* and *Sophophora* as well as from Hawaiian *Drosophila.* The branching pattern of the 11 sequences agrees with that of our NJ tree, though we have used 42 sequences. They also stated that *Zaprionus* is placed on the interior

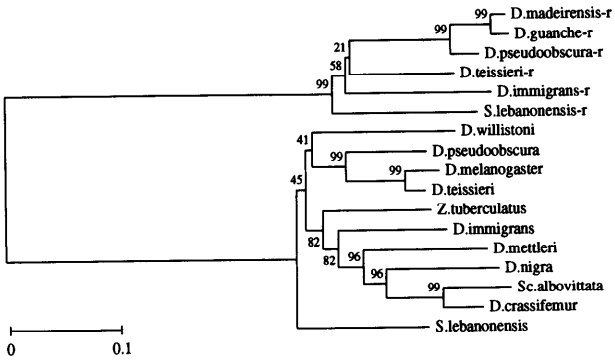

FIG. 3.—Neighbor-joining tree with Kimura two-parameter distances for 17 *Adh* and *Adhr* sequences. The total number of nucleotides used was 765 after elimination of insertions/deletions. CP is given for each interior branch.

branch between subgenera *Drosophila* and *Sophophora.* Therefore, our study supports their general conclusion. Juan et al.'s (1994) unrooted tree for 18 *Adh* sequences is also in agreement with the topology of our tree.

However, our CP values give a high confidence for many of the species clusters. Our study also gives the evolutionary relationships of several newly investigated species groups such as *D. immigrans* and *D. willistoni* together with many other species. It is interesting that all the species belonging to most species groups and subgroups form monophyletic clusters, though the number of species examined is not always large.

In our tree the BCL and CP values are generally very high and close to each other. In some of the deep branches, however, BCL is considerably lower than CP, suggesting that BCL gives an underestimate of statistical confidence when a large number of sequences is examined (Sitnikova et al. 1995). In the case of closely related species such as *D. mauritiana, D. simulans,* and *D. sechellia* or some *D. mulleri* subgroup species, the *Adh* gene does not contain enough phylogenetic information to resolve the branching pattern (see also Jeffs et al. 1994).

Determination of the Outgroup Species

As mentioned above, the phylogenetic tree in figure 2 was constructed on the assumption that *Scaptodrosophila* is the outgroup. This assumption is consistent with the phylogenetic trees obtained by using a clear-cut outgroup such as mosquito (DeSalle 1992*b*), *Leucophenga* (Pelandakis and Solignac 1993), and the medfly (Kwiatowski et al. 1994). Particularly the loss of an intron in the *Sod* gene in non-*Scaptodrosophila* and non-*Chymomyza* drosophilid species strongly suggest that *Scaptodrosophila* is an outgroup (Kwiatowski et al. 1994). Nevertheless, statistical tests of the trees constructed by these authors did not really confirm this assumption. (The confidence interval tests conducted by Kwiatowski et al. for the maximum-likelihood tree are not very reliable, because these tests are known to be too liberal; see Tateno et al. 1994.) Furthermore, the branch lengths for *Zaprionus tuberculatus, D. immigrans,* and *D. willistoni* suggested that these species could also be outgroups. We have therefore examined this problem in more detail.

To test our hypothesis, we first examined the phylogenetic tree for the *Adh* and *Adhr* genes. Since the duplication of the *Adh* and *Adhr* genes is ancient (Schaeffer and Aquadro 1987), these genes were expected to give some answer to our problem. The NJ tree obtained is presented in figure 3 together with the CP values for interior branches. This tree shows that *Scaptodrosophila* is an outgroup for both *Adh* and *Adhr* sequences. Unfortunately, however, the CP values for the interior branch that separates *Scaptodrosophila* from other drosophilids is too low in both cases.

We then reanalyzed DeSalle's (1992*a,* 1992*b*) sequence data for a segment of mitochondrial DNA (rRNA and ND1 region). This data set included not only drosophilid flies but also a mosquito species (*Aedes albopictus*). The NJ tree obtained is presented in figure 4. This tree shows that *Scaptodrosophila* is an outgroup of the other drosophilids with a 96% CP value. A similar analysis of the NJ tree for Pelandakis and Solignac's (1993) data for a nuclear rRNA gene was also conducted. The results obtained (fig. 5) again support our hypothesis with a 96% CP value. These results together with the loss of an intron of the *Sod* gene in the genus *Drosophila* strongly suggest that *Scaptodrosophila* is a true outgroup.

Comparison with Other Phylogenies of Drosophilids

As mentioned earlier, Grimaldi (1990) constructed a phylogenetic tree for many different groups of dro-

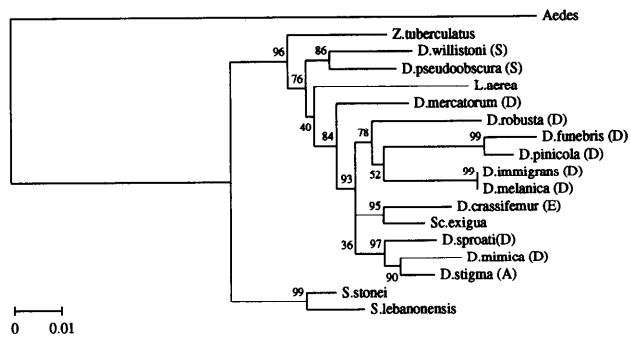

Fig. 4.—Neighbor-joining tree constructed with a 905-bp segment of mitochondrial DNA (rRNA and ND1 region) (data from DeSalle 1992*a,* 1992*b*). Pairwise Kimura two-parameter distances were used. CP is given for each interior branch. D, subgenus *Drosophila;* S, subgenus *Sophophora;* E, subgenus *Engiscaptomyza;* A, subgenus *Antopocerus* (Hawaiian); Z, genus *Zaprionus;* L, genus *Liodrosophila;* Sc, genus *Scaptomyza;* S, genus *Scaptodrosophila.*

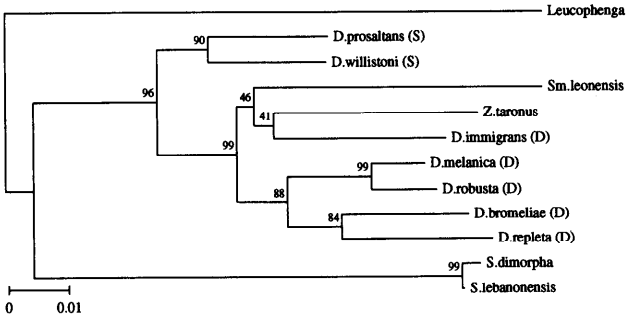

FIG. 5.—Neighbor-joining tree constructed with the 28S nuclear rRNA gene (data from Pelandakis and Solignac 1993). The total number of nucleotides used was 327. Since the sequences are closely related to each other, p distances were used. CP is given for each interior branch. D, subgenus *Drosophila;* S, subgenus *Sophophora;* Sm, genus *Samoaia;* Z, genus *Zaprionus;* S, genus *Scaptodrosophila.*

sophilids by using mainly morphological characters. One of the most conspicuous differences between his tree and ours is that the genus *Scaptomyza* is located outside the genus *Drosophila* in his tree, whereas it is inside the genus in our tree (fig. 2). In our tree, *Scaptomyza* belongs to a cluster which may be called the Hawaiian drosophilids, and this cluster is highly significant. Actually, all molecular data available support the clustering of *Scaptomyza* and subgenus *Drosophila* (Beverley and Wilson 1984; DeSalle 1992*a;* Pelandakis and Solignac 1993; Thomas and Hunt 1993). Therefore, this cluster now seems to be firmly established, which is consistent with Throckmorton's (1975) view that all Hawaiian drosophilids originated from a single ancestral species. If this is the case, it would be interesting to study how the morphology (male genitalia: Throckmorton 1962, 1966; egg ultrastructures: Kambysellis 1993) of *Scaptomyza* diverged so rapidly from that of other Hawaiian drosophilids. However, the clustering of *Scaptomyza* with Hawaiian drosophilids in figure 2 can also be explained by the hypothesis of separate introduction of the *Scaptomyza* and *Engiscaptomyza* group and the other Hawaiian drosophilid group, as suggested by Thomas and Hunt (1991).

Another conspicuous difference between Grimaldi's (1990) and our trees is that in his tree the genus *Zaprionus* is located outside the genus *Drosophila,* whereas in our tree it branches off from the interior branch between *Drosophila* subgenera *Drosophila* and *Sophophora.* Our branching pattern is again supported by the studies of Thomas and Hunt (1993) and Kwiatowski et al. (1994), but it is inconsistent with that of DeSalle (1992*a,* 1992*b;* see also fig. 4). Our tree is also somewhat different from that of Pelandakis and Solignac (1993) with respect to the phylogenetic location of *Zaprionus* (5 non-*tuberculatus* species). According to their tree, *Zaprionus* is located within subgenus *Drosophila* and is closely related to *D. immigrans* and *D. repleta* groups (see also fig. 5). However, statistical support of our branching pattern (CP = 77, BCL = 59) is not very high, so a further study of the phylogenetic location of *Zaprionus* seems to be necessary (see also Powell and DeSalle 1994).

Another incongruence between the morphological and molecular trees is the phylogenetic location of the Hawaiian drosophilid *D. picticornis.* Although morphological characters have placed this species in the *D. planitibia* subgroup, our tree puts it outside the subgroup. Actually, *D. picticornis* seems to be somewhat different from other members of the *D. planitibia* subgroup, since it lays eggs in tree saps rather than in rotten barks as the other members do (M. Kambysellis, personal communication). Our grouping is again the same as that of Rowan and Hunt (1991) with *Adh* gene sequences, and it is statistically supported at a highly significant level. If this conclusion proves to be correct for other genes as well, it seems necessary to modify the classification of Hawaiian picture-winged drosophilids.

Our phylogeny of the *D. repleta* species group is also inconsistent with the classification based on morphology to some extent. Thus, *D. mettleri,* which belongs to the *D. mulleri* subgroup, forms a tight cluster with *D. hydei,* a species belonging to the *D. hydei* subgroup.

In our tree *D. willistoni* belongs to the subgenus *Sophophora,* as in the case of morphological classification. According to Pelandakis and Solignac's (1993) tree, however, it is closer to genera *Scaptodrosophila* and *Chymomyza* than to other *Drosophila* species, and thus the subgenus *Sophophora* is polyphyletic rather than monophyletic. Therefore, the *Adh* and rRNA gene trees are contradictory with each other, though both trees have low statistical support for the *D. willistoni* branch.

On the basis of morphological characters, the *D. melanogaster* subgroup is divided into two species complexes. The first is the *D. melanogaster* species complex, including *D. melanogaster, D. simulans, D. sechellia,* and *D. mauritiana.* The second is the *D. yakuba* species complex, including *D. orena, D. erecta, D. yakuba,* and *D. teissieri* (Lemeunier et al. 1986). However, our tree places *D. orena* outside all other *D. melanogaster* subgroup members with a high CP value. Late in the preparation of this paper, the *Adh* sequence from *D. erecta* became available (Jeffs et al. 1994). Although we did not include it in figure 2, it clusters significantly (97% CP, 95% BCL) with *D. orena,* and this cluster is located outside the *D. yakuba–D. teissieri* cluster. The same clustering pattern has been reported in other molecular studies (Solignac et al. 1986; Pelandakis and Solignac 1993; Jeffs et al. 1994). These results suggest that the *D. melanogaster* subgroup may not be divided into only

two species complexes as previously thought (Lemeunier et al. 1986).

Divergence Times

To estimate the approximate times of divergence between species or species groups, we first applied our (N. Takezaki, A. Rzhetsky, and M. Nei, unpublished data) method of testing the heterogeneity of evolutionary rate among different lineages and eliminated the sequences which evolved significantly (at the 1% level) faster or slower compared with the average rate. We then constructed a linearized tree under the assumption of a constant rate of evolution. (Elimination of deviant branches is not essential for the construction of a linearized tree.) For this purpose, we used only third codon positions, because nucleotide substitutions at first and second codon positions were apparently subjected to stronger purifying selection than those at third positions. The evolutionary distances for third positions were on average about three times higher than those for all codon positions. However, the G+C content at third positions varies from species to species. We therefore used Tajima and Nei's (1984) distance measure to estimate the number of nucleotide substitutions.

Figure 6 shows the linearized tree obtained. The branch lengths in this tree were estimated by the least-squares method. (Details of these procedures will be published elsewhere.) Our test of rate heterogeneity eliminated *Z. tuberculatus* and *D. orena*. It also showed that the *D. pseudoobscura* subgroup evolved significantly slower than the average rate. However, we included the sequences for this subgroup, because they are biologically important. *Drosophila heteroneura* was excluded, because it had the same nucleotide sequence as *D. planitibia* at third codon positions. The tree in figure 6 has three multifurcating nodes, one each for the *D. mulleri* subgroup, the *D. planitibia* subgroup, and the *D. obscura* group. These multifurcating nodes were produced because of the constraint due to the assumption of rate constancy (molecular clock). However, the tree in figure 2 shows that the branching pattern of the species involved in a multifurcating node is not statistically resolved even when the data for all three codon positions are used.

To estimate the times of divergence between species, we have to know the rate of nucleotide substitution. In the case of drosophilids, this rate can be estimated by using information on the times of island formation in Hawaii. Rowan and Hunt (1991) have argued that the most useful geological dating is that for the formation of Kauai, the oldest island in the Hawaiian Archipelago (5.1 Mya). Interestingly, 98% of the Hawaiian drosophilids are endemic to single islands (Hardy 1974). Thus, *D. picticornis* is found only in Kauai, *D. differens* in Molokai (1.9 Mya), *D. affinidisjuncta* and *D. planitibia*

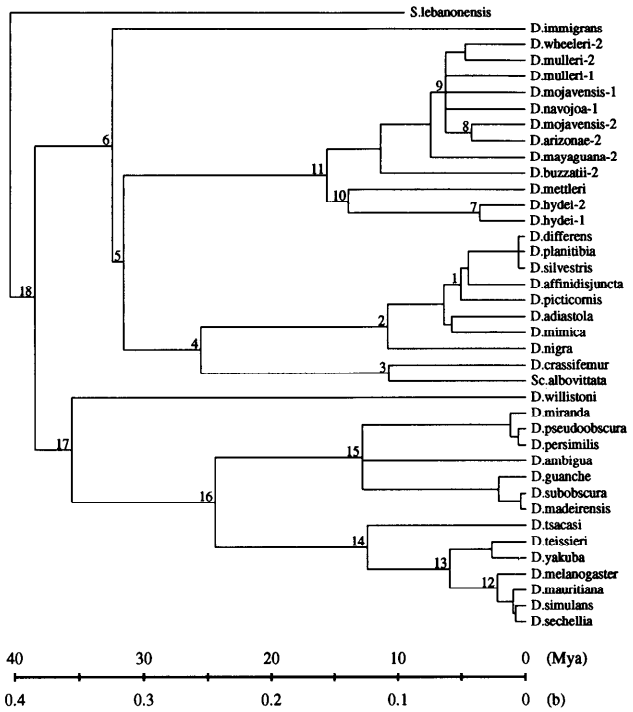

FIG. 6.—Linearized tree with divergence time estimates for 39 *Adh* drosophilid sequences. This tree was constructed by using the topology in fig. 2. *b* stands for the branch length.

in Maui (1.3 Mya), and *D. heteroneura* and *D. silvestris* in the island of Hawaii (0.5 Mya) (McDougall 1979; Carson and Yoon 1982). In this paper, therefore, we assumed that the node (1) in figure 6 corresponds to 5.1 Mya. (*Drosophila nigra* and *D. mimica* are supposed to have diverged from the *D. planitibia* subgroup before the formation of Kauai; Beverley and Wilson 1984.) This assumption gives an estimate of the rate of nucleotide substitution per site per year per lineage equal to 1.0×10^{-8}. This rate is slightly lower than the estimates of the rate of synonymous substitution (about 1.5×10^{-8}) by Moriyama (1987), Moriyama and Gojobori (1992), and Rowan and Hunt (1991), but this is reasonable because the nucleotide substitution at third codon positions is subjected to purifying selection to some extent. In this paper we did not use synonymous substitutions because it was not so easy to construct a linearized tree with this distance measure. (We need information on the variances and covariances of all pairwise distances). The time scale given in figure 6 was obtained by using the rate of 1.0×10^{-8}.

Table 1 shows the estimates of divergence times for 18 pairs of sequences or sequence clusters. This table shows that the split of subgenera *Drosophila* and *Sophophora* occurred about 40 Mya. This estimate is considerably lower than the one (60 Mya) obtained from immunological distance data (Beverley and Wilson

Table 1
Divergence Time Estimates

Taxa Compared	Time
Dnsophila picticornis vs. *D. silvestris*	5.1
D. nigra vs. *D. silvestris*	11.0 ± 1.53
Hawaiian *Scaptomyza* vs. *D. crassifemur*	10.9 ± 1.71
D. crassifemur vs. Hawaiian picture-winged	26.1 ± 2.87
Hawaiian drosophilids vs. *D. repleta* group	32.2 ± 3.04
D. immigrans vs. *D. repleta* group	33.1 ± 3.16
D. hydei Adh-1 vs. *Adh-2*	3.6 ± 0.90
D. mojavensis Adh-2 vs. *D. arizonae Adh-2*	4.2 ± 0.99
D. mulleri Adh-1 vs. *Adh-2* and *D. mojavensis Adh-1* vs. *Adh-2*	6.5 ± 0.90
D. mettleri vs. *D. hydei*	14.2 ± 2.04
D. mettleri vs. *D. mulleri* subgroup	15.9 ± 1.64
D. simulans vs. *D. melanogaster*	2.3 ± 0.65
D. melanogaster vs. *D. yakuba*	6.1 ± 1.12
D. montium vs. *D. melanogaster* subgroups	12.7 ± 1.88
D. obscura vs. *D. pseudoobscura* subgroups	13.1 ± 1.74
D. obscura vs. *D. melanogaster* groups	24.9 ± 2.88
D. willistoni vs. *D. melanogaster* groups	36.3 ± 4.26
Drosophila vs. *Sophophora* subgenera	39.2 ± 3.35

NOTE.—All time estimates are based on the assumptions that *D. picticornis* and *D. silvestris* diverged 5.1 Mya. The standard errors for the branch lengths were used to calculate the standard errors of the time estimates, and they are given in Mya.

1984) but is close to the estimates obtained from DNA sequence data (Thomas and Hunt 1993; Kwiatowski et al. 1994). The split between *D. willistoni* and the *D. obscura–D. melanogaster* cluster also seems to have occurred a long time ago (36 Mya). Similarly, *D. immigrans* apparently diverged from the *D. repleta* group and Hawaiian drosophilids about 33 Mya.

Our estimate of the splitting of Hawaiian drosophilids from continental drosophilids is about 32 Mya. Grimaldi (1987) described about seven drosophilid fossil species from the early Miocene (about 23 Mya). Three of these species belong to the genus *Drosophila* and one to the genus *Scaptomyza* (from the Dominican Republic). If *Scaptomyza* originated in the Hawaiian Islands and then spread through the rest of the world (Throckmorton 1975), the minimum estimate of the time of the split between Hawaiian and continental drosophilids would be 23 Mya. Therefore, our estimate is not inconsistent with these fossil records (see also DeSalle 1992*a*). However, if *Scaptomyza* was introduced into Hawaii independently of the other Hawaiian drosophilids, then it is difficult to assess the time of the introduction of these *Scaptomyza* species.

Among the Hawaiian drosophilids, *D. crassifemur* and *S. albovittata* diverged from the rest of the Hawaiian species about 26 Mya, and then the two species diverged from each other about 11 Mya. These estimates are in good agreement with those of Thomas and Hunt (1993) and suggest that the morphology of *Scaptomyza* evolved very rapidly. There are *Scaptomyza* species which are endemic to other parts of the world, and these species may have originated in Hawaii, as mentioned above. If this is the case, the migration of the species out of Hawaii seems to have occurred during the last 11 million yr. This contradicts with the time estimate of the *Scaptomyza* fossil found in the Dominican Republic, about 23 million yr old. This apparent contradiction can be resolved, if *Engiscaptomyza* actually belongs to the taxon *Scaptomyza* or if the fossil specimen is actually an ancestral form of *Scaptomyza* and *Engiscaptomyza.* It is also possible that our molecular dating is incorrect. Some answers to these questions may be obtained if sequence data become available for other non-Hawaiian *Scaptomyza* species. How the *Scaptomyza* species migrated out of Hawaii (if they did) remains a mystery at present. (About half of this cosmopolitan genus of some 330 species occurs in Hawaii; DeSalle and Grimaldi 1993.) As mentioned earlier, Hawaiian species of *Scaptomyza* can be descendants of the second migration of drosophilids into Hawaii. Our estimates of the times of divergence for other Hawaiian drosophilid species are similar to those of Thomas and Hunt (1993), though they used the rate of synonymous substitution rather than that of third codon substitution.

Our estimate (13 Mya) of the splitting time between the *D. obscura* and *D. pseudoobscura* subgroups is much older than the estimate (6 Mya) obtained by restriction-site analysis (Lattorre et al. 1988). It should be noted that the standard error relative to the estimate of a divergence time increases as the estimate decreases (see table 1) and thus the estimates of recent divergence times are less reliable than those of older divergence times. Yet, our estimate (2.3 Mya) of the divergence time between *D. melanogaster* and *D. simulans* is very close to that (2–5 Mya) of Bodmer and Ashburner (1984) and Stephens and Nei (1985).

Gene Duplications

Many *Drosophila repleta* group species are known to have duplicate functional *Adh* genes (*Adh-1* and *Adh-2*) as well as a pseudogene (ψ*Adh*), and these genes are arranged in the order of ψ*Adh, Adh-2,* and *Adh-1* from the 5′ side of the DNA (see fig. 7). Figure 6 suggests that the duplication of the *Adh-1* and *Adh-2* genes in the *D. mulleri* subgroup, excluding *D. mettleri,* occurred 6–11 Mya. By contrast, the gene duplication in *D. hydei* seems to have occurred only about 4 Mya. Therefore, these results suggest that at least two independent gene duplications occurred in the *D. repleta* group. This view is supported by the fact that *D. mettleri* does not have two functional genes (Yum et al. 1991). Furthermore, the *Adh-2* gene in *D. hydei* is expressed from the embryonic

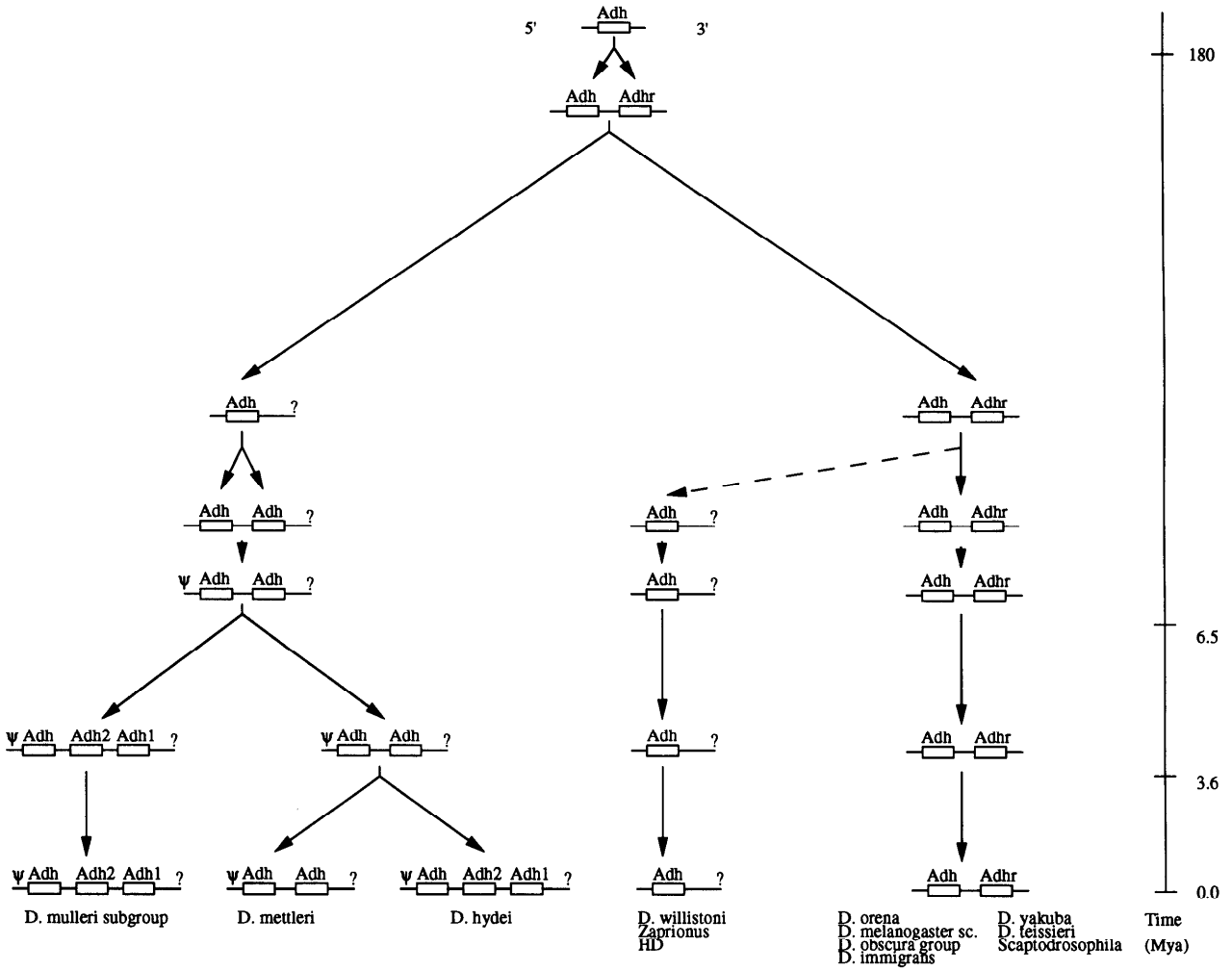

FIG. 7.—One possible scenario of the evolution of *Adh* and *Adh*-like genes by gene duplication. The question mark (?) indicates that the *Adhr* gene has not been identified in these species, either because it has not been found yet or because it has been lost. HD, Hawaiian drosophilids; sc, species complex.

stage, while the *Adh-2* gene in *D. mulleri* is first expressed in the late third instar larval stage (Sullivan et al. 1990*a*).

However, there is another possible explanation. That is, gene duplication occurred only once before the radiation of the *D. repleta* group species, but the duplicate genes have undergone occasional gene conversion as in the case of the globin genes in mammals (Zimmer et al. 1980; Hardison 1984). If gene conversion occurs between two duplicate genes occasionally, the sequence similarity between the two genes from the same species is expected to be higher than that between those from different species (Sullivan et al. 1990*b*). Under this hypothesis, the presence of one functional *Adh* gene in *D. mettleri* can be explained by assuming that one of the two copies has been lost. Menotti-Raymond et al. (1991) studied the possible occurrence of gene conversion but concluded that there was no compelling evidence for it. Therefore, the gene conversion hypothesis may not apply to the *Adh* genes in the *D. repleta* group.

As mentioned earlier, duplication of the *Adh* and *Adhr* genes occurred before drosophilids evolved, and it is interesting to know the time of this duplication event. For this purpose, however, we cannot use all codon positions, because nucleotide substitutions between the two genes are saturated at third codon positions. We therefore computed Kimura's two-parameter distances for the first and second codon positions for all pairs of species in figure 3. Using these distance values, we constructed a linearized tree similar to that of figure 6 (data not shown). We then estimated the divergence time between the *Adh* and *Adhr* genes assuming that *Drosophila* subgenera *Drosophila* and *Sophophora* diverged 39 Mya. The estimate obtained was 180 Mya.

This suggests that the *Adh-Adhr* gene duplication occurred during the Jurassic period (before 135 Mya),

from which no Cyclorrapha (higher dipteran suborder) fossils have been found (Beverley and Wilson 1984, 1985). Since the cyclorraphan radiation is believed to have occurred about 100 Mya (Wiegmann et al. 1993), the gene duplication apparently preceded the evolution of the higher dipteran suborder Cyclorrapha.

Discussion

We have produced a phylogenetic tree for drosophilids that is more reliable statistically than previous trees. This tree does not necessarily agree with Grimaldi's (1990) tree based on morphological characters. The most conspicuous difference occurred in the phylogenetic location of *Scaptomyza,* as mentioned earlier. However, since our study is based on Hawaiian *Scaptomyza,* it remains to be seen whether the same problem arises with continental *Scaptomyza.*

It should be noted that the present study is based on data of a single gene (*Adh*) and that the tree may have been distorted by some peculiarities of the gene which we have not detected. It is therefore desirable to examine the phylogeny by using some other genes as well. At present, there are several sets of DNA sequence data that can be used for constructing a drosophilid phylogeny. Unfortunately, different investigations have used different groups of species when they studied different genes. Therefore, it is difficult to combine these data to produce a more reliable tree.

Our tree is not only statistically well supported but also largely consistent with the geographical distributions of the species, if we exclude cosmopolitan species such as *D. hydei, D. immigrans, D. melanogaster,* and *D. simulans* (Parsons and Stanley 1981). The most conspicuous of this consistency is Hawaiian drosophilids that form a monophyletic clade, as we have already discussed. The drosophilids that belong to the *D. repleta* group are endemic to North and South Americas (see Material and Methods). According to Throckmorton (1982), this group of species was derived from a lineage in Asia (at latest) about 30 Mya. Since Hawaiian drosophilids are also believed to have originated in eastern Asia, our tree is in accord to Throckmorton's (1975) conjecture of the evolution of drosophilids. It would be interesting to study how the drosophilids in eastern Asia (e.g., *D. histrio, D. confusa*) are related to the Hawaiian drosophilids and the *D. repleta* group.

Since *D. melanogaster* and *D. simulans* are believed to be recent migrants from Africa (Lemeunier et al. 1986), all species of the *D. melanogaster* subgroup apparently originated in Africa. This group of species also form a tight cluster in our tree. It has been suggested that the *D. melanogaster* subgroup originated about 17–20 Mya (Jeffs et al. 1994), when the faunal interchange between Africa and Eurasia was first possible. However, our study suggests that this event took place somewhat later (i.e., about 6–13 Mya).

By contrast, the *D. obscura* group species are geographically subdivided. The *D. pseudoobscura* subgroup species, which are endemic to western North America, do not cluster with the *D. repleta* group, another North American species group. Instead, they are closer to the *D. obscura* subgroup (from Europe). This supports Lakovaara and Saura's (1982) hypothesis that the *D. pseudoobscura* subgroup was introduced from Europe through Asia relatively recently. Our linearized tree in figure 6 suggests that this introduction occurred about 13 Mya (late Miocene).

Drosophilids have been an important group of species for studying the mechanisms of evolution (see, e.g., Carson and Kaneshiro 1976). Yet, the phylogenetic relationships of these species are not well established. To take advantage of this extremely diversified group of species for the study of evolution, it is very important to clarify the phylogenetic relationships. It will not only contribute to the development of a more reasonable drosophilid taxonomy but also to the understanding of the mechanism of the evolution of important morphological and physiological characters.

Acknowledgments

We would like to thank David Grimaldi, John Hunt, Michael Kambysellis, Jeff Powell, Steve Schaeffer, Darren Schafer, David Sullivan, Koichiro Tamura, and Richard Thomas for their valuable comments and helpful suggestions during the preparation of the paper. The paper is a part of the Ph.D. thesis of Claudia A. M. Russo at the Departamento de Genética, Instituto de Biologia da Universidade Federal do Rio de Janeiro, Brazil. Claudia Russo is sponsored by CAPES from the Education Ministry, Brazil. This study was also supported by research grants from National Institutes of Health (GM20293) and National Science Foundation (DEB 9119802) to M.N.

LITERATURE CITED

ABALAT, R., and R. GONZALEZ-DUARTE. 1993. *Adh* and *Adh*-dup sequences of *Drosophila lebanonensis* and *D. immigrans:* interspecies comparisons. Gene **126**:171–178.

ANDERSON, C. L., E. A. CAREW, and J. R. POWELL. 1993. Evolution of the *Adh* locus in the *Drosophila willistoni* group: the loss of an intron, and shift of codon usage. Mol. Biol. Evol. **10**:605–618.

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

BEVERLEY, S. M., and A. C. WILSON. 1984. Molecular evolution in *Drosophila* and the higher Diptera. II. A time scale for fly evolution. J. Mol. Evol. **21**:1–13.

———. 1985. Ancient origin for Hawaiian Drosophilidae inferred from protein comparisons. Proc. Natl. Acad. Sci. USA **82**:4753–4757.

BODMER, M., and M. ASHBURNER. 1984. Conservation and change in the DNA sequences coding for alcohol dehydrogenase in sibling species of *Drosophila.* Nature **209**:425–429.

CARSON, H. L., and K. Y. KANESHIRO. 1976. *Drosophila* of Hawaii: systematics and ecological genetics. Ann. Rev. Ecol. Syst. **7**:311–345.

CARSON, H. L., and J. S. YOON. 1982. Genetics and evolution of Hawaiian *Drosophila.* Pp. 297–344 *in* M. ASHBURNER, H. L. CARSON, and J. N. THOMPSON, eds. The genetics and biology of *Drosophila.* Vol. 3b. Academic Press, New York.

COHN, V. H., and G. P. MOORE. 1988. Organization and evolution of the *Adh* gene in *Drosophila.* Mol. Biol. Evol. **5**: 154–166.

COYNE, J., and M. KREITMAN. 1985. Evolutionary genetics of two sibling species, *Drosophila simulans* and *D. sechellia.* Evolution **40**:673–691.

DESALLE, R. 1992*a*. The origin and possible time divergence of the Hawaiian Drosophilidae: evidence from DNA sequences. Mol. Biol. Evol. **9**:905–916.

———. 1992*b*. The phylogenetic relationships of flies in the Family Drosophilidae deduced from mtDNA sequences. Mol. Phylogenet. Evol. **1**:31–40.

DESALLE, R., and D. A. GRIMALDI. 1993. Phylogenetic pattern and developmental process in *Drosophila.* Syst. Biol. **42**: 458–475.

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

GRIMALDI, D. A. 1987. Amber fossil Drosophilidae (Diptera), with particular reference to the Hispaniolan taxa. Am. Mus. Novitates **2880**:1–23.

———. 1990. A phylogenetic, revised classification of genera in the Drosophilidae (Diptera). Bull. Am. Mus. Nat. Hist. **197**:1–139.

HARDISON, R. C. 1984. Comparison of the b-like globin gene families of rabbits and humans indicates that the gene cluster 5′-e-g-d-b-3′ predates the mammalian radiation. Mol. Biol. Evol. **1**:390–410.

HARDY, D. E. 1965. Diptera: Cyclorrhapha II, Series Schizophora, Section Acalypterae I, Family Drosophilidae. Pp. 1–814 *in* E. C. ZIMMERMAN, ed. Insects of Hawaii. Vol. 12. University of Hawaii Press, Honolulu.

———. 1974. Introduction and background information. Pp. 71–80 *in* M. J. D. WHITE, ed. Genetic mechanisms of speciation in insects. Reidel, Boston.

HILLIS, D. M., and J. J. BULL. 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst. Biol. **42**:182–192.

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

JUAN, E., M. PAPACEIT, and A. QUINTANA. 1994. Nucleotide sequence of the genomic region encompassing *Adh* and *Adh*-dup genes of *D. lebanonensis* (*Scaptodrosophila*): gene expression and evolutionary relationships. J. Mol. Evol. **38**: 455–467.

JUKES, T. H., and C. R. CANTOR. 1969. Evolution of protein molecules. Pp. 21–132 *in* H. N. MUNRO, ed. Mammalian protein metabolism. Vol. 3. Academic Press, New York.

KAMBYSELLIS, M. P. 1993. Ultrastructural diversity in the egg chorion of Hawaiian *Drosophila* and *Scaptomyza:* ecological and phylogenetic considerations. Int. J. Insect Morphol. Embryol. **22**:417–446.

KANESHIRO, K. Y. 1974. Phylogenetic relationships of Hawaiian Drosophilidae based on morphology. Pp. 102–110 *in* M. J. D. WHITE, ed. Genetic mechanisms of speciation in insects. Reidel, Boston.

———. 1976. A revision of generic concepts in the biosystematics of Hawaiian Drosophilidae. Proc. Hawaii. Entomol. Soc. **22**:255–278.

KIMURA, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. **16**:111–120.

KREITMAN, M. 1983. Nucleotide polymorphism at the alcohol dehydrogenase locus of *Drosophila melanogaster.* Nature **304**:412–417.

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

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

LAKOVAARA, S., and A. SAURA. 1982. Evolution and speciation in the *Drosophila obscura* group. Pp. 1–59 *in* M. ASHBURNER, H. L. CARSON, and J. N. THOMPSON, eds. The genetics and biology of *Drosophila.* Vol. 3b. Academic Press, New York.

LATTORRE, A., E. BARRIO, A. MOYA, and F. J. AYALA. 1988. Mitochondrial DNA evolution in the *Drosophila obscura* group. Mol. Biol. Evol. **5**:717–728.

LEMEUNIER, F., J. R. DAVID, L. TSACAS, and M. ASHBURNER. 1986. The *melanogaster* species group. Pp. 147–256 *in* M. ASHBURNER, H. L. CARSON, and J. N. THOMPSON, eds. The genetics and biology of *Drosophila.* Vol. 3e. Academic Press, New York.

MARFANY, G., and R. GONZALEZ-DUARTE. 1990. Nucleotide sequence of the *Adh* gene of *Drosophila lebanonensis.* Nucleic Acids Res. **18**:6706.

———. 1991*a*. The *Adh* genomic region of *Drosophila ambigua:* evolutionary trends in different species. J. Mol. Evol. **32**:454–462.

———. 1991*b*. The *Drosophila subobscura Adh* genomic region contains valuable evolutionary markers. Mol. Biol. Evol. **9**:261–277.

———. 1993. Characterization and evolution of the *Adh* genomic region in *Drosophila guanche* and *Drosophila madeirensis.* Mol. Phylogenet. Evol. **2**:13–22.

MARUYAMA, K., and D. L. HARTL. 1991. Evidence of interespecific transfer of the transposable element mariner be-

tween *Drosophila* and *Zaprionus*. J. Mol. Evol. **33**:514–524.

MCDOUGALL, I. 1979. Age of shield-building volcanism of Kauai and linear migration of volcanism in the Hawaiian Island chain. Earth Planet. Sci. Lett. **46**:31–42.

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

MORIYAMA, E. N. 1987. Higher rates of substitution in *Drosophila* than in mammals. Jap. J. Genet. **62**:139–147.

MORIYAMA, E. N., and T. GOJOBORI. 1992. Rates of synonymous substitutions and base composition of nuclear genes in *Drosophila*. Genetics **130**:855–864.

NEI, M. 1991. Relative efficiencies of different tree-making methods for molecular data. Pp. 90–128 *in* M. M. MIYAMOTO and J. CRACRAFT, eds. Phylogenetic analysis of DNA sequences. Oxford University Press, New York.

PARSONS, P. A., and S. M. STANLEY. 1981. Domesticated and widespread species. Pp. 349–393 *in* M. ASHBURNER, H. L. CARSON, and J. N. THOMPSON, eds. The genetics and biology of *Drosophila*. Vol. 3a. Academic Press, New York.

PELANDAKIS, M., D. G. HIGGINS, and M. SOLIGNAC. 1991. Molecular phylogeny of the subgenus *Sophophora* of *Drosophila* derived from large subunit of ribosomal RNA sequences. Genetica **84**:87–94.

PELANDAKIS, M., and M. SOLIGNAC. 1993. Molecular phylogeny of *Drosophila* based on ribosomal RNA sequences. J. Mol. Evol. **37**:525–543.

POWELL, J., and R. DESALLE. 1995. *Drosophila* molecular phylogenies and their uses. Evol. Biol. (in press).

ROWAN, R. G., and W. J. DICKINSON. 1988. Nucleotide sequence of the genomic region coding for alcohol dehydrogenase in *Drosophila affinidisjuncta*. J. Mol. Evol. **28**:43–54.

ROWAN, R. G., and J. A. HUNT. 1991. Rates of DNA change and phylogeny from the DNA sequences of the alcohol dehydrogenase gene for five closely related species of Hawaiian *Drosophila*. Mol. Biol. Evol. **8**:49–70.

RZHETSKY, A., and M. NEI. 1992. A simple method for estimating and testing minimum-evolution trees. Mol. Biol. Evol. **9**:945–967.

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

SCHAEFFER, S. W., and C. F. AQUADRO. 1987. Nucleotide sequence of the *Adh* gene region of *Drosophila pseudoobscura:* evolutionary change and evidence for an ancient gene duplication. Genetics **117**:61–73.

SCHAEFFER, S. W., and E. L. MILLER. 1991. Nucleotide sequence analysis of *Adh* genes estimates the time of geographic isolation of the Bogota population of *Drosophila pseudoobscura*. Proc. Natl. Acad. Sci. USA **88**:6097–6101.

———. 1992. Molecular population genetics of an electrophoretically monomorphic protein in the alcohol dehydrogenase region of *Drosophila pseudoobscura*. Genetics **132**: 163–178.

SITNIKOVA, T., A. RZHETSKY, and M. NEI. 1995. Interior-branch and bootstrap tests of phylogenetic trees. Mol. Biol. Evol. (in press).

SOLIGNAC, M., M. MONNEROT, and J.-C. MOUNOLOU. 1986. Mitochondrial DNA evolution in the *melanogaster* species subgroup of *Drosophila*. J. Mol. Evol. **23**:31–40.

STEPHENS, J. C., and M. NEI. 1985. Phylogenetic analysis of polymorphic DNA sequences at the *Adh* locus in *Drosophila melanogaster* and its sibling species. J. Mol. Evol. **22**:289–300.

SULLIVAN, D. T., P. W. ATKINSON, C. A. BAYER, and M. A. MENOTTI-RAYMOND. 1990*a*. The evolution of *Adh* expression in the Repleta group of *Drosophila*. Pp. 407–418 *in* J. S. F. BARKER, W. F. STARMER, and R. J. MACINTYRE, eds. Ecological and evolutionary genetics of *Drosophila*. Plenum, New York.

SULLIVAN, D. T., P. W. ATKINSON, and W. T. STARMER. 1990*b*. Molecular evolution of the alcohol dehydrogenase genes in the genus *Drosophila*. Evol. Biol. **24**:107–147.

SWOFFORD, D. L. 1993. PAUP: phylogenetic analysis using parsimony, version 3.1.1. Illinois Natural History Survey, Champaign.

TAJIMA, F., and M. NEI. 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 bias. Mol. Biol. Evol. **9**:678–687.

TATENO, Y., N. TAKEZAKI, and M. NEI. 1994. Relative efficiencies of the maximum-likelihood, neighbor-joining, and maximum-parsimony methods when substitution rate varies with site. Mol. Biol. Evol. **11**:261–277.

THOMAS, R. H., and J. A. HUNT. 1991. The molecular evolution of the alcohol dehydrogenase locus and the phylogeny of Hawaiian *Drosophila*. Mol. Biol. Evol. **8**:687–702.

———. 1993. Phylogenetic relationships in *Drosophila:* a conflict between molecular and morphological data. Mol. Biol. Evol. **10**:362–374.

THROCKMORTON, L. 1962. The problem of phylogeny in the genus *Drosophila*. Univ. Texas Publ. **6205**:207–343.

———. 1966. The relationships of the endemic Hawaiian Drosophilidae. Univ. Texas Publ. **6615**:335–396.

———. 1975. The phylogeny, ecology, and geography of *Drosophila*. Pp. 421–469 *in* R. C. KING, ed. Handbook of genetics. Vol. 3. Plenum, New York.

———. 1982. Pathways of evolution in the genus *Drosophila* and the founding of the *Repleta* group. Pp. 33–47 *in* J. S. J. BARKER and W. T. STARMER, eds. Ecological genetics and evolution, the cactus-yeast-*Drosophila* model system. Academic Press, Sydney.

WEAVER, J. R., J. M. ANDREWS, and D. T. SULLIVAN. 1989. Nucleotide sequence of the *Adh-1* gene of *Drosophila navojoa*. Nucleic Acids Res. **17**:7524.

WHEELER, M. R. 1981. The Drosophilidae: a taxonomic overview. Pp. 1–97 *in* M. ASHBURNER, H. L. CARSON, and

J. N. THOMPSON, eds. The genetics and biology of *Drosophila.* Vol. 3a. Academic Press, New York.

———. 1986. Additions to the catalog of the world's Drosophilidae. Pp. 395–409 *in* M. ASHBURNER, H. L. CARSON, and J. N. THOMPSON, eds. The genetics and biology of *Drosophila.* Vol. 3e. Academic Press, New York.

WIEGMANN, B. M., C. MITTER, and F. C. THOMPSON. 1993. Evolutionary origin of the Cyclorrapha (Diptera): tests of alternative morphological hypotheses. Cladistics **9**:41–81.

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

ZHARKIKH, A., and W.-H. LI. 1992. Statistical properties of bootstrap estimation of phylogenetic variability from nucleotide sequences. I. Four taxa with a molecular clock. Mol. Biol. Evol. **9**:1119–1147.

ZIMMER, E. A., S. L. MARTIN, S. M. BEVERLEY, Y. W. KAN, and A. C. WILSON. 1980. Rapid duplication and loss of genes coding for the a chains of hemoglobin. Proc. Natl. Acad. Sci. USA 77:2158–2162.

CHARLES AQUADRO, reviewing editor

Received July 28, 1994

Accepted November 15, 1994