ZOOLOGICAL SCIENCE **24**: 913–921 (2007)

Phylogeny of the *Drosophila immigrans* Species Group (Diptera: Drosophilidae) Based on *Adh* and *Gpdh* Sequences

Toru Katoh[1]*, Daisuke Nakaya[2], Koichiro Tamura[2] and Tadashi Aotsuka[2]

[1]*COE for Neo-Science of Natural History, Graduate School of Science, Hokkaido University, North 10 West 8, Kita-ku, Sapporo 060-0810, Japan*
[2]*Department of Biological Sciences, Tokyo Metropolitan University, 1-1 Minami-osawa, Hachioji, Tokyo 192-0397, Japan*

The *immigrans* species group in the Drosophilinae is one of the representative species groups of *Drosophila* in East Asia. Although this group constitutes a significant part of the drosophilid fauna in the Old World, only a few species have been analyzed in previous molecular phylogenetic studies. To study the phylogeny of the *immigrans* group, we analyzed the nucleotide sequences of two nuclear genes, alcohol dehydrogenase (*Adh*) and glycerol-3-phosphate dehydrogenase (*Gpdh*), for 36 drosophilid species, including 12 species of the *immigrans* group. In the resultant phylogenetic trees, 10 species of the *immigrans* group (*D. immigrans*, *D. formosana*, *D. ruberrima*, *D. albomicans*, *D. nasuta*, *D. neonasuta*, *D. pallidifrons*, *D. hypocausta*, *D. neohypocausta*, *D. siamana*) consistently formed a clade (the *immigrans* group proper), although the phylogeny within this clade did not exactly correspond to the classification of species subgroups. However, *D. annulipes* and *D. quadrilineata,* both of which belong to the *quadrilineata* subgroup of the *immigrans* group, were not included in the *immigrans* group proper. Furthermore, we obtained the unexpected result that *D. annulipes* was included in a clade comprising *Scaptomyza* and Hawaiian *Drosophila,* together with *D. maculinotata* of the *funebris* group, although the phylogenetic relationships within this clade remain uncertain and need to be substantiated with further studies. Thus, according to the present study, the *immigrans* group is polyphyletic.

Key words: molecular phylogeny, Drosophilidae, *immigrans* species group, *Adh*, *Gpdh*

INTRODUCTION

The Drosophilidae comprises a large dipteran family that includes over 3,700 described species and is widely distributed around the world (Bächli, 2006). One of these species, *D. melanogaster*, has widely been used as a model organism for various fields, including genetics, ethology, and developmental biology. To date, much knowledge has thus been accumulated on the biology of Drosophilidae. However, the phylogeny of the Drosophilidae, which should provide the basis for any evolutionary studies of drosophilids, remains controversial (Powell, 1997; Markow and O'Grady, 2006).

The taxonomy, systematics, and phylogeny of the Drosophilidae have been extensively studied using a large amount of morphological and molecular data. One of the most noteworthy studies of drosophilid phylogeny was that of Throckmorton (1975), based on internal morphology and biogeography. His phylogenetic hypothesis was subsequently widely accepted by many evolutionary biologists. However, with a cladistic analysis of external morphology, Grimaldi (1990) put forward a substantially different hypothesis. Since then, a number of molecular studies have been conducted to elucidate the phylogenetic relationships among drosophilids (DeSalle 1992; Pélandakis and Solignac, 1993; Kwiatowski *et al.*, 1994, 1997; Remsen and DeSalle, 1998; Kwiatowski and Ayala, 1999; Tatarenkov *et al.*, 1999, 2001; Katoh *et al.*, 2000; Remsen and O'Grady, 2002), and these have contributed to resolution of the evolutionary relationships among the major taxa of the subfamily Drosophilinae.

Despite these contributions, however, the phylogeny of Drosophilidae has not yet been completely resolved. One reason may be that previous molecular phylogenetic studies have encompassed a relatively narrow taxonomic range. Some groups not yet included in molecular analyses are important in addressing controversial problems in drosophilid phylogeny. In the present molecular phylogenetic study, we focused on the large *immigrans* species group of *Drosophila*. Although this group constitutes a major part of the drosophilid fauna in the Old World, only a few species from the *immigrans* group have been included in previous studies (DeSalle, 1992; Pélandakis and Solignac, 1993;

* Corresponding author. Phone: +81-11-706-3524;
Fax : +81-11-746-0862;
E-mail: tkatoh@nature.sci.hokudai.ac.jp
doi:10.2108/zsj.24.913

Russo *et al.*, 1995; Tamura *et al.*, 1995; Remsen and DeSalle, 1998; Kwiatowski and Ayala, 1999; Tatarenkov *et al.*, 1999, 2001; Yu *et al.*, 1999; Katoh *et al.*, 2000; Remsen and O'Grady, 2002; Nagaraja *et al.*, 2004; Robe *et al.*, 2005).

Throckmorton (1975) suggested that the *immigrans* group evolved within the *immigrans-Hirtodrosophila* radiation. This hypothesis has been supported by recent molecular phylogenetic studies (Kwiatowski and Ayala, 1999; Tatarenkov *et al.*, 1999, 2001; Katoh *et al.*, 2000; Remsen and O'Grady, 2002; Robe *et al.*, 2005). However, only a few species in the *immigrans* group were included in these studies, and therefore relationships between the *immigrans* and other species groups have not been resolved in detail.

The *immigrans* species group comprises 93 species and three subspecies (Toda, 2006). Except for one cosmopolitan species, *D. immigrans*, most of the species are confined to the region extending from off the African coast, through Asia, to the Pacific islands. Wilson *et al.* (1969) divided this group into the *immigrans*, *nasuta*, *quadrilineata*, *hypocausta*, and *lineosa* subgroups. Okada and Carson (1983) transferred the species of the *lineosa* subgroup into the genus *Zaprionus*. Zhang and Toda (1992) established the *curviceps* subgroup to include some new species and other species transferred from the *immigrans* subgroup. Thus, the *immigrans* group now contains five subgroups: *immigrans*, *nasuta*, *quadrilineata*, *hypocausta*, and *curviceps*.

The *immigrans* group is defined by only one morphological character, the presence of a row of stout spinules on the inner side of the foreleg femur (Fig. 1). Aside from this character, the group shows considerable morphological variation. Furthermore, Wakahama *et al.* (1983) reported that the karyotype of the *quadrilineata* subgroup is substantially different from that of most of the *immigrans* group, suggesting that the former may be divergent from the rest of the group. Thus, several problems remain regarding the phylogeny of the *immigrans* group.

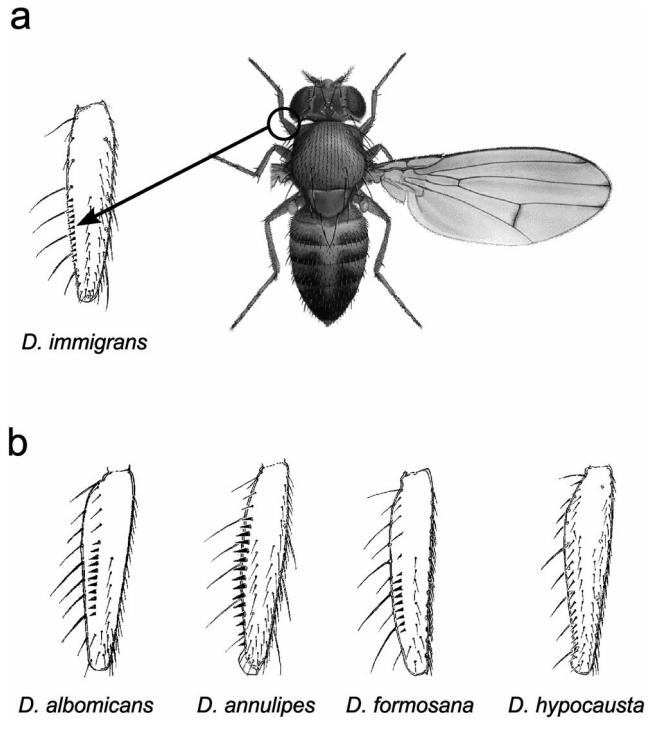

Fig. 1. A row of stout spinules on the foreleg femur is a morphological character that defines the *immigrans* species group of *Drosophila*. (a) Position of the row of stout spinules. (b) Morphology of the stout spinules observed in some species of the *immigrans* group.

To study the phylogeny of the *immigrans* group, we analyzed the nucleotide sequences of two nuclear genes, alcohol dehydrogenase (*Adh*) and glyserol-3-phosphate dehydrogenase (*Gpdh*). Both genes have been extensively utilized and have generally performed well in phylogenetic reconstructions of Drosophilinae (Tamura *et al.*, 1995; Kwiatowski *et al.*, 1997, Kwiatowski and Ayala, 1999; Tatarenkov *et al.*, 1999; Katoh *et al.*, 2000; Goto and Kimura, 2001). We included in our analysis newly determined sequences as well as homologous sequences from GenBank. In our phylogenetic reconstructions, most species of the *immigrans* group formed a clade in the *immigrans-Hirtodrosophila* lineage. Nevertheless, some incongruencies with the accepted taxonomy were also found.

MATERIALS AND METHODS

Specimens

Our study included 36 drosophilid species, among which 13 *Adh* and 16 *Gpdh* sequences from 17 species were newly determined (Table 1). The utilized *D. guttifera* strain was originally obtained from the National *Drosophila* Species Resource Center at Bowling Green State University and maintained at Tokyo Metropolitan University. The specimens of *D. maculinatata* and *S. pallida* were kindly provided by Prof. Masahito T. Kimura and Dr. Yao-Guang Hu (Hokkaido University), respectively. The other 14 species sequenced in this study were obtained from stocks maintained at Tokyo Metropolitan University.

DNA extraction, PCR amplification, cloning, and sequencing

Genomic DNA was extracted by the method of Steller (1990) or Boom *et al.* (1990), with some modifications. PCR amplification, cloning, and sequencing of *Adh* gene followed the methods described in Katoh *et al.* (2000). PCR amplification of the *Gpdh* gene was performed using the primers L3 (5'-GTT CTA GAT CTG GTT GAG GCT GCC AAG AA-3') and R6 (5'-ACA TAT GCT CTA GAT GAT TGC GTA TGC A-3') of Kwiatowski *et al.* (1997). Amplifications were carried out in 10-μl reaction volumes, each containing 1X Ex Taq buffer (Takara Bio), 200 μM each dNTP, 0.5 μM each primer, 0.25 U Ex Taq (Takara Bio), and approximately 10 ng of genomic DNA, with the following cycle conditions: 95°C for 3 min, followed by 35 cycles of 95°C for 30 s, 60°C for 90 s, and 72°C for 90 s, followed by a 7-min extension at 72°C. PCR products were directly sequenced, or cloned into pGEM-T Easy vector (Promega) using *E. coli* DH5α as the host. Sequences were determined in both directions using a BigDye Terminator Sequencing Kit (PE Applied Biosystems) and an ABI 3100-Avant Genetic Analyzer, according to the manufacturers' protocols.

Data analysis

We examined the sequences for coding regions that included 711 and 768 sites for *Adh* and *Gpdh*, respectively. Nucleotide sequences were aligned with Clustal X 1.83 (Thompson *et al.*, 1997) or with the Clustal algorithm implemented in MEGA 3.1 (Kumar *et al.*, 2004), and the resultant alignments were checked by eye. No alignment gaps occurred in the regions analyzed. Aligned sequence data were then imported into PAUP* 4.0b10 (Swofford, 2003) for phylogenetic analysis. In order to determine whether

Table 1. Taxa included in the study and Genbank accession numbers for *Adh* and *Gpdh* sequences. Sequences obtained in this study are underlined.

					Accession No.	
Genus	Subgenus	Species group	Subgroup	Species	*Adh*	*Gpdh*
Drosophila	*Drosophila*	*immigrans*	*hypocausta*	*hypocausta*	AB261133	AB261145
				neohypocausta	AB261134	AB261146
				siamana	AB261135	AB261147
			immigrans	*formosana*	AB261131	AB261143
				immigrans	M97638	AB261142
				ruberrima	AB261132	AB261144
			nasuta	*albomicans*	AB033642	AB261148
				nasuta	AB261137	AB261149
				pallidifrons	AB261136	AB261150
				neonasuta	AB261138	AB261151
			quadrilineata	*annulipes*	AB261129	AB261152
				quadrilinetata	AB261130	AB261153
		funebris		*funebris*	AB033643	AB261154
				maculinotata	AB261140	AB261156
		guttifera		*guttifera*	AB261139	AB261155
		virilis		*virilis*	AB033640	D10697
		repleta	*hydei*	*hydei*	X58694	L41650
		picture wing	*planitibia*	*picticornis*	M63392	AY006450
				heteroneura	M36781	AY006454
				silvestris	M63291	AY006453
				planitibia	M63390	AY006458
				differens	M36785	AY006455
	Sophophora	*melanogaster*	*melanogaster*	*melanogaster*	M17833	X14179
				simulans	M36581	L41647
				teissieri	X54118	U47809
		obscura	*subobscura*	*guanche*	X60113	U47808
			pseudoobscura	*miranda*	M60998	L41251
				pseudoobscura	M60979	L41249
		willistoni		*willistoni*	L08648	L41248
				paulistorum	AB026529	L41648
	Dorsilopha			*busckii*	AB261141	*
Hirtodrosophila				*pictiventris*	AB026530	L41649
Scaptomyza	*Parascaptomyza*			*pallida*	AB033645	AB261157
Zaprionus	*Zaprionus*			*tuberculatus*	X63955	L37039
Chymomyza				*procnemis*	AB026521	L41252
Scaptodrosophila				*lebanonensis*	X54814	*

*Sequences were obtained from Kwiatowski and Ayala (1999).

nucleotide composition bias occurred among taxa, χ^2 goodness-of-fit tests were performed on the sequence data. The *Adh* and *Gpdh* data were analyzed separately and then combined for a simultaneous analysis. Before the data sets were concatenated, we performed the incongruence length difference (ILD) test (Farris *et al.*, 1994), which is referred to as the partition homogeneity test (PHT) in PAUP*, to detect possible incongruence between the two data sets. The test was implemented under parsimony with 1,000 heuristic search replicates for each of which 100 starting trees were generated by random stepwise addition, in order to generate the null distribution. A summary of the characteristics of each data partition used in this study is presented in Table 2.

Unweighted maximum parsimony (MP) trees were obtained through 1,000 heuristic search replicates, with starting trees generated by random sequence addition, followed by the tree bisection reconnection (TBR) branch swapping. Bootstrap values (Felsenstein, 1985) for the MP tree were determined from 1,000 pseudoreplicates, for each of which an MP tree was obtained through 100 heuristic search replicates with random sequence addition and TBR branch swapping.

Maximum likelihood (ML) trees were obtained by TBR branch swapping, starting with a topology given by the neighbor-joining (NJ) method (Saitou and Nei, 1987). Parameters for ML analysis were selected on the basis of the Akaike information criterion (AIC) (Akaike, 1974) implemented in Modeltest 3.7 (Posada and Crandall, 1998). The optimal models found from the analysis are listed in Table 3. Bootstrap values for the ML trees were calculated from 200 pseudoreplicates analyzed by the nearest neighbor interchange (NNI) searches, with the starting topology given by a NJ tree.

χ^2 tests of base frequency revealed significant compositional heterogeneity for *Adh* and the combined data sets (Table 2). Jermiin *et al.* (2004) cautioned that compositional heterogeneity can mislead both MP and ML methods. To assess the effect of heterogeneity on our phylogenetic inferences, we used LogDet-paralinear (LogDet) distances to construct minimum evolution (ME) trees. As with the ML analyses, ME trees were obtained by TBR branch swapping, with the starting topology given by a NJ tree. Bootstrap values for the ME trees were obtained from 1,000 pseudoreplicates

Table 2. Characteristics of data partitions used in this study.

Partition	TS	VS	PIS	Base frequencies				Compositional heterogeneity	
				%A	%C	%G	%T	χ^2	*P*
All data	1479	684	585 (39.6%)	24.3	25.9	25.5	24.3	191.68 (df, 105)	0.0000005*
Adh	711	405	341 (48.0%)	23.6	28.1	25.2	23.1	149.37 (df, 105)	0.003*
Gpdh	768	279	244 (31.8%)	25.1	23.8	25.8	25.3	80.47 (df, 105)	0.964

TS, total sites; VS, variable sites; PIS, parsimony informative sites.
* Significant nucleotide composition bias.

Table 3. Optimal substitution models for the *Adh*, *Gpdh*, and combined data sets, selected by AIC in Modeltest 3.7 (Posada and Crandall, 1998).

Partition	Model	I	G	Base frequencies	Rate matrix
All data	TrN+I+G	0.4815	1.3514	A=0.2551	A-C=1.0000
				C=0.2873	A-G=2.5683
				G=0.2302	A-T=1.0000
				T=0.2274	C-G=1.0000
					C-T=5.7815
					G-T=1.0000
Adh	GTR+I+G	0.3554	1.2167	A=0.2389	A-C=1.4523
				C=0.3007	A-G=2.4209
				G=0.2368	A-T=1.0101
				T=0.2236	C-G=1.0691
					C-T=5.4022
					G-T=1.0000
Gpdh	TrNef+I+G	0.5952	1.4939	Equal frequencies	A-C=1.0000
					A-G=3.6883
					A-T=1.0000
					C-G=1.0000
					C-T=8.0773
					G-T=1.0000

GTR, general time reversible model (Tavaré, 1986); TrN, Tamura-Nei model (Tamura and Nei, 1993); TrNef, Tamura-Nei model with equal base frequencies (Posada and Crandall, 1998); I, proportion of invariant sites; G, gamma distribution shape parameter.

subjected to TBR branch swapping, with the starting topology of each given by a NJ tree.

Finally, a Bayesian analysis was performed using MrBayes 3.1.2 (Huelsenbeck and Ronquist, 2001). Parameters for this analysis were selected on the basis of the AIC test implemented in MrModeltest 2.2 (Nylander, 2004). A Markov-Chain Monte-Carlo (MCMC) search was performed with four chains, each of which was run for 3,000,000 generations. Trees were sampled every 100 generations, and those of the first 500,000 generations were discarded as burn-in and ensured that a stable likelihood had been reached. A consensus of sampled trees was computed, and the posterior probability for each interior branch was obtained to assess the robustness of the inferred relationships.

RESULTS

Phylogenetic analysis

A summary of optimality values for the MP, ML, and ME analyses of *Adh*, *Gpdh*, and the combined data set is presented in Table 4. Regardless of the optimality criterion used, the individual and combined analyses yielded tree topologies that were generally in agreement with one another, although the individual analyses tended to yield less-resolved results than the combined analysis (data not shown). The ILD test yielded no significant incongruence between the *Adh* and *Gpdh* data sets (sum of tree length, 3029; *P*=0.20).

The ML tree obtained using the combined *Adh* and *Gpdh* data is shown in Fig. 2. *Chymomyza procnemis* and *Scaptodrosophila lebanonensis* were used as outgroups, as in previous phylogenetic studies of Drosophilinae (Kwiatowski *et al.* 1994, 1997; Remsen and DeSalle 1998; Kwiatowski and Ayala 1999; Tatarenkov *et al.*, 1999; Katoh *et al.*, 2000). Table 5 gives the bootstrap and posterior probability values, corresponding to the numbered clades shown in Fig. 2, for the trees obtained by MP, ME, and Bayesian analyses. Almost all clades with >50% bootstrap support in Fig. 2 were also well supported in the other analyses (Table 5). Thus, the tree in Fig. 2 was largely in agreement with those obtained by the other methods.

The phylogenetic relationships indicated in Fig. 2 are generally congruent with those detected by other studies of Drosophilinae using nuclear gene sequences (Tamura *et al.*, 1995; Kwiatowski and Ayala, 1999; Tatarenkov *et al.*, 1999, 2001; Katoh *et al.*, 2000). Two large clades (17 and 23) were detected in the ingroup. Clade 23 corresponds to the subgenus *Sophophora*, in which the *melanogaster* and *obscura* groups (clades 18 and 20, respectively) emerged as sister clades, with the *willistoni* group (clade 22) as the sister group to the *melanogaster* + *obscura* clade (21). Clade 17 contained the remaining ingroup species, including members of the subgenera *Drosophila* (species belonging to the

Table 4. Summary of optimality values for the MP, ML, and ME analyses.

Partition	No. of MP trees		-ln *L*	ME-score
All data	2	(Tree length = 3054, CI = 0.372, RI = 0.602)	15014.61488	1.94778
Adh	1	(Tree length = 1845, CI = 0.382, RI = 0.604)	8738.14844	2.46761
Gpdh	47	(Tree length = 1184, CI = 0.365, RI = 0.613)	6099.48765	1.46678

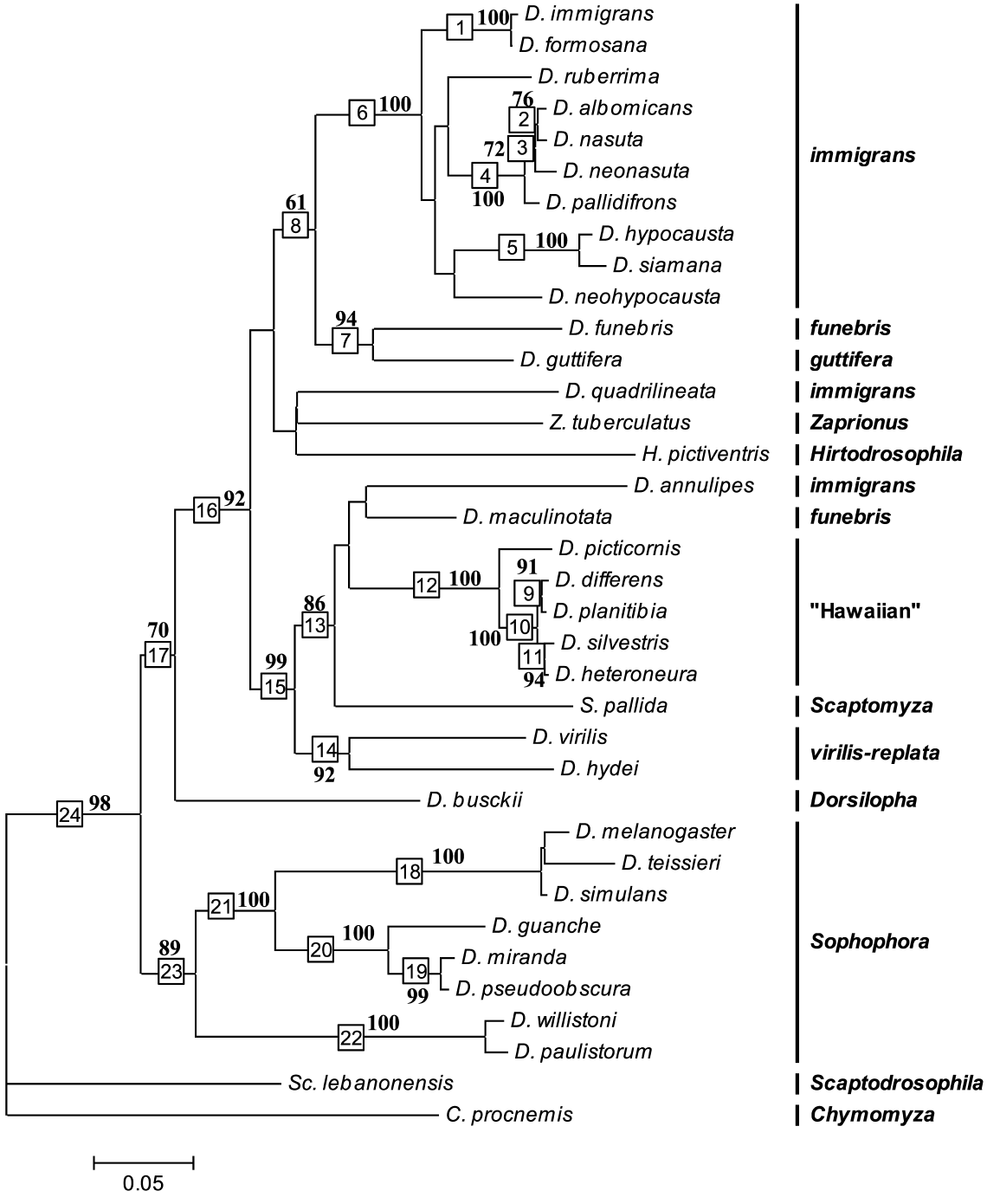

Fig. 2. Maximum-likelihood (ML) tree resulting from analyses of the combined *Adh* and *Gpdh* data sets, based on the TrN+I+G substitution model. Bootstrap values >50% are shown, determined by analysis of 200 pseudoreplicates. Numbers in squares indicate clades with >50% bootstrap support and correspond to the clade numbers given in Table 5.

Table 5. Bootstrap and posterior probability values for each of the clades shown in Fig. 2, applicable to the trees constructed using MP, ME, and Bayesian methods.

Clade	Bootstrap value		Posterior
	MP	ME	probability
1	100	100	1.0
2	65	93	0.98
3	<50	53	0.97
4	100	100	1.0
5	100	100	1.0
6	100	100	1.0
7	96	99	1.0
8	62	–	0.98
9	69	58	1.0
10	100	100	1.0
11	97	100	1.0
12	100	100	1.0
13	80	85	1.0
14	86	86	1.0
15	86	98	1.0
16	80	69	1.0
17	61	72	0.98
18	100	100	1.0
19	100	100	1.0
20	100	100	1.0
21	84	98	1.0
22	100	100	1.0
23	53	56	1.0
24	94	95	1.0

immigrans, *funebris*, *guttifera*, and *virilis-repleta* groups) and *Dorsilopha* (*D. busckii*), and of the genera *Hirtodrosophila* (*H. pictiventris*), *Zaprionus* (*Z. tuberculatus*), *Scaptomyza* (*S. pallida*), as well as Hawaiian *Drosophila* (clade 12). In clade 17, *D. busckii* was the early offshoot, comprising the sister group to clade 16, which contained the remaining species. Within clade 16, highly supported clade 15 comprised two sister clades: the *virilis-repleta* clade (14) and, another (13) that included Hawaiian *Drosophila* (clade 12), *D. annulipes*, *D. maculinotata*, and *S. pallida*. The relationships among the remaining species were less well resolved, except for the sister-group relationship between the *D. funebris* + *D. guttifera* clade (7) and clade 6 that included 10 species of the *immigrans* group. The phylogenetic positions of *H. pictiventris*, *Z. tuberculatus*, and *D. quadrilineata* were unresolved, being inconsistent among the trees deduced from different analyses and weakly supported by bootstrap values. The phylogeny of species within the *immigrans* group is described in the following section.

The *immigrans* species group

In this study, we analyzed 12 species belonging to the *immigrans* group, including three species of the *immigrans* subgroup (*D. immigrans*, *D. formosana*, and *D. ruberrima*), four species of the *nasuta* subgroup (*D. albomicans*, *D. nasuta*, *D. neonasuta*, and *D. pallidifrons*), and three species of the *hypocausta* subgroup (*D. hypocausta*, *D. neohypocausta*, and *D. siamana*). These species comprised clade 6, henceforth referred to as the *immigrans* group proper, supported by 100% bootstrap and posterior probability values (Table 5). In the MP, ML, and Bayesian analyses, clade 6 comprised the sister group to *D. funebris* + *D. guttifera* (clade 7), although this relationship (clade 8) was not supported by the ME analysis (Table 5).

Tree topologies showing the relationships within the *immigrans* group proper (clade 6) are presented in Fig. 3. Within this clade, four species of the *nasuta* subgroup (*D. albomicans*, *D. nasuta*, *D. neonasuta*, and *D. pallidifrons*) formed a clade with 100% bootstrap and posterior probability support. Outside the *nasuta* subgroup, *D. immigrans* + *D. formosana* (*immigrans* subgroup) and *D. siamana* + *D. hypocausta* (*hypocausta* subgroup) formed well supported clades. However, the phylogenetic positions of *D. ruberrima* (*immigrans* subgroup) and *D. neohypocausta* (*hypocausta* subgroup) were unstable among the phylogenetic trees deduced by different methods (Fig. 3). Thus, questions remain concerning the monophyly of both the *immigrans* and *hypocausta* subgroups.

Drosophila quadrilineata and *D. annulipes*, both of which belong to the *quadrilineata* subgroup, were neither included in the clade of the *immigrans* group proper, nor joined as monophyletic. In Fig. 2, *D. quadrilineata* was in clade 16, but its exact position within this clade was uncertain. The more striking finding was the placement of *D. annulipes* together with *D. maculinotata* (*funebris* group) in clade 13, along with *S. pallida* and Hawaiian *Drosophila*, with 80–86% bootstrap support (Fig. 2, Table 5). Thus, the *quadrilineata* subgroup emerged as polyphyletic, and distantly related to the *immigrans* group proper.

DISCUSSION

Phylogeny of the *immigrans* group within Drosophilinae

The monophyly of the *immigrans* group, and also the relationships between *quadrilineata* and the other subgroups, have been open to question. In general appearance, species of the *quadrilineata* subgroup look quite different from others of the *immigrans* group (Zhang and Toda, 1992). Furthermore, the karyotype of *D. annulipes* of the *quadrilineata* subgroup is fundamentally different from that common to other members of the *immigrans* group (Wakahama *et al.*, 1983). These observations suggest that the *quadrilineata* subgroup may be remote from the other species of the *immigrans* group. Interestingly, our phylogenetic trees placed *D. quadrilineata* and *D. annulipes* of the *quadrilineata* subgroup separately from one another and remote from the other species of the *immigrans* group.

According to Throckmorton (1975), the *immigrans* group originated within the *immigrans-Hirtodrosophila* radiation, which includes species belonging to the *quinaria* and *tripunctata* groups of *Drosophila*, *Zaprionus*, *Hirtodrosophila*, and others. In Fig. 2, *D. quadrilineata* falls into a sister clade to the *D. funebris* + *D. guttifera* clade besides the *immigrans* group proper. Furthermore, *D. annulipes* unexpectedly groups with *D. maculinotata*, and this clade is nested within a clade that includes Hawaiian *Drosophila* and *Scaptomyza*.

Although *D. annulipes* and *D. maculinotata* have been placed in the *immigrans* and *funebris* groups, respectively, there are some questions about the classification of these two species. As mentioned above, Zhang and Toda (1992)

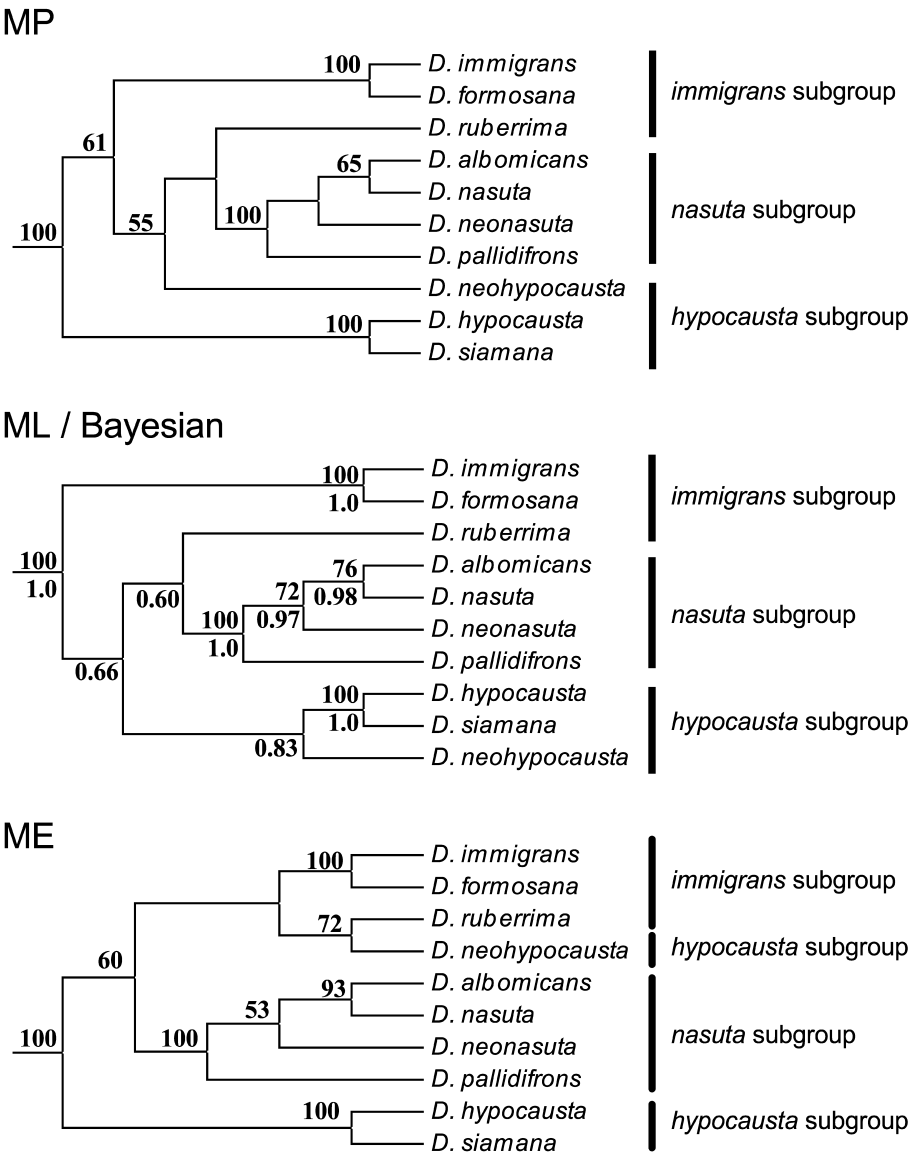

Fig. 3. Topologies of partial trees representing the relationships among the 10 species of the *immigrans* group proper, obtained from the MP, ML, ME, and Bayesian analyses. The ML and Bayesian analyses yielded the same topology. Bootstrap values >50% are shown above branches near the nodes they support. Bayesian posterior-probability values are shown below branches of the ML/Bayesian tree.

and Wakahama *et al.* (1983) have suggested that *D. annulipes* is distinct from the other species of the *immigrans* group. In addition, subsequent to Okada's (1956) original classification of *D. maculinotata* into the *funebris* group, Okada (1988) argued that this species belongs instead to an unknown species group in the *virilis* section. Thus, our study is compatible with these studies and questions the current taxonomy of these species.

This finding may be interesting when we consider the relationship between continental and the Hawaiian species of Drosophilidae. Although Hawaiian drosophilids have been well studied as an extensive adaptive radiation, hypotheses on their origin and their phylogenetic relationships to other, continental drosophilids are still in dispute. For example, on the basis of internal morphology, Throckmorton (1975) hypothesized that the Hawaiian drosophilids are a monophyletic lineage most closely related to members of the *immigrans-Hirtodrosophila* lineage. In contrast, from a cladistic analysis of external morphology, Grimaldi (1990) concluded that Hawaiian drosophilids are polyphyletic. However, recent molecular phylogenetic studies have consistently revealed Hawaiian drosophilids as a monophyletic sister group to the *virilis-repleta* lineage (Tamura *et al.*, 1995; Remsen and DeSalle, 1998; Kwiatowski and Ayala, 1999; Tatarenkov *et al.*, 1999, 2001; Katoh *et al.*, 2000; Remsen and O'Grady, 2002). Unfortunately, these molecular studies could not specify a particular species or species group as the closest sister group to the Hawaiian drosophilids. It is thus striking that our study found two continental species comprising a clade with the Hawaiian Drosophilidae.

Both *D. annulipes* and *D. maculinotata* are restricted to Asia: *D. annulipes* is widely distributed in the Oriental region, and *D. maculinotata* occurs in the subalpine zone of central Japan. These facts suggest the hypothesis that the

ancestor of the Hawaiian species came from Asia. We note, however, that this is quite speculative. Our phylogeny (Fig. 2) also allows the alternative hypothesis that both *D. annulipes* and *D. maculinotata*, or their common ancestor, originated in the Hawaiian Islands and dispersed to Asia, even though this second hypothesis requires the unlikely occurrence of two long-distance colonization events, one from a continental area to originally establish the Hawaiian drosophilids and another from Hawaii to Asia. Further studies using additional loci and taxa, especially more species of the *quadrilineata* subgroup, *Scaptomyza*, and Hawaiian *Drosophila*, will be required to resolve this issue.

Phylogeny within the *immigrans* group proper

Except for the *quadrilineata* subgroup, 10 species of the *immigrans* group constituted a clade in the *immigrans-Hirtodrosophila* lineage, which was consistent with previous morphological and molecular studies (Throckmorton 1975; Grimaldi 1990; Tamura *et al.*, 1995; Remsen and DeSalle, 1998; Kwiatwski and Ayala, 1999; Tatarenkov *et al.*, 1999, 2001; Katoh *et al.*, 2000; Remsen and O'Grady, 2002; Robe *et al.*, 2005). This clade included species belonging to three species subgroups (*hypocausta*, *immigrans*, and *nasuta*) of the *immigrans* group. In our study, however, the phylogeny of the 10 species in this clade did not exactly correspond to the classification of species subgroups. Two species of the *immigrans* subgroup (*D. immigrans* and *D. formosana*), four species of the *nasuta* subgroup (*D. albomicans*, *D. nasuta*, *D. neonasuta*, and *D. pallidifrons*), and two species of the *hypocausta* subgroup (*D. hypocausta* and *D. siamana*) formed consistent clades (Fig. 3). However, in all trees in Fig. 3, *D. ruberrima* was separate from the clade containing other species of the *immigrans* subgroup. Similarly, in the MP and ME analyses, *D. neohypocausta* was separate from the two other species of the *hypocausta* subgroup. The *hypocausta* subgroup was monophyletic only in the ML and Bayesian analyses, and bootstrap support was <50% for the ML tree. Our results do not support the monophyly of either the *immigrans* or the *hypocausta* subgroups.

The tree topology of the *immigrans* group varied among the different methods of tree reconstruction (Fig. 3). Although compositional heterogeneity can mislead both MP- and ML-based methods (Jermiin *et al.*, 2004), it seems unlikely that heterogeneity significantly affected the tree topology in this case. Our data indicated significant heterogeneity when all the taxa analyzed were included in the χ^2 test (Table 2). However, when the χ^2 test was applied only to the 10 species of the *immigrans* group proper (clade 6), no significant heterogeneity was detected (χ^2=1.249062, df=27, P=1.00).

According to Wilson *et al.* (1969), the *hypocausta* subgroup usually shows strong sexual dimorphism in body color (males are much darker than females), and the row of spinules on the inner side of the foreleg femur is poorly developed (Fig. 1). However, we have observed *D. neohypocausta* not to have the body-color dimorphism as strongly developed as in *D. hypocausta*, and to have well-developed spinules. In addition, the karyotype of *D. neohypocausta* is somewhat different from that of *D. hypocausta* (Wakahama *et al.*, 1983). These characters are consistent with those of our results (MP and ME trees in Fig. 3) that indicate *D. neohypocausta* as being remote from the other species of the *hypocausta* subgroup. On the other hand, even though the morphology and karyotype of *D. ruberrima* are not so different from those of the other species of the *immigrans* subgroup, our study did not indicate *D. ruberrima* as closely related to *D. immigrans* and *D. formosana*. This might be due to poor taxon sampling (the *immigrans* subgroup includes a total of 33 species) and/or reflect the poly- or paraphyletic nature of this subgroup. In any case, the phylogenetic positions of *D. neohypocausta* and *D. ruberrima* were not resolved by our study. The taxonomic positions of these species remain unclear, and further phylogenetic analyses of the *immigrans* group will be required to resolve them.

ACKNOWLEDGMENTS

We thank Masahito T. Kimura and Yao-Guang Hu for providing us with drosophilid samples collected from the field. We are also grateful to Matthew H. Dick, Masanori J. Toda, and two anonymous reviewers for carefully reading the manuscript and giving us many valuable comments and suggestions. This work was supported by Grants-in-Aid for Scientific Research (A) from the Ministry of Education, Japan (Nos. 12375002; 15255006) and by a MEXT Grant-in Aid for the 21st Century COE Program (J01).

REFERENCES

Akaike H (1974) A new look at the statistical model identification. IEEE Trans Automat Control 19: 716–723

Boom R, Sol CJ, Salimans MM, Jansen CL, Wertheim-van Dillen PM, Van Der Noordaa J (1990) Rapid and simple method for purification of nucleic acids. J Clin Microbiol 28: 495–503

Bächli G (2006) TaxoDros: The Database on Taxonomy of Drosophilidae, version 2006-May, last accessed 1 November 2006. http/:taxodros.unizh.ch/

DeSalle R (1992) The phylogenetic relationships of flies in the family Drosophilidae deduced from mtDNA sequences. Mol Phylogenet Evol 1: 31–40

Farris JC, Kälersjö M, Kluge AG, Bult C (1994) Testing the significance of incongruence. Cladistics 10: 315–319

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783–781

Goto SG, Kimura MT (2001) Phylogenetic utility of mitochondrial *COI* and nuclear *Gpdh* genes in *Drosophila*. Mol Phylogenet Evol 18: 404–422

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

Huelsenbeck JP, Ronquist F (2001) MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics 17: 754–755

Jermiin L, Ho SY, Ababneh F, Robinson J, Larkum AW (2004) The biasing effect of compositional heterogeneity on phylogenetic estimates may be underestimated. Syst Biol 53: 638–643

Katoh T, Tamura K, Aotsuka T (2000) Phylogenetic position of the subgenus *Lordiphosa* of the genus *Drosophila* (Diptera: Drosophilidae) inferred from alcohol dehydrogenase (*Adh*) gene sequences. J Mol Evol 51: 122–130

Kumar S, Tamura K, Nei M (2004) MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Briefings Bioinform 5: 150–163

Kwiatowski J, Ayala FJ (1999) Phylogeny of *Drosophila* and related genera: conflict between molecular and anatomical analyses. J Mol Evol 13: 319–328

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

Kwiatowski J, Krawczyk M, Jaworski M, Skarecky D, Ayala FJ

(1997) Erratic evolution of glyserol-3-phosphate dehydrogenase in *Drosophila*, *Chymomyza*, and *Ceratitis*. J Mol Evol 44: 9–22

Markow TA, O'Grady PA (2006) Drosophila: A Guide to Species Identification and Use. Elsevier, London

Nagaraja, Nagaraju J, Ranganath HA (2004) Molecular phylogeny of the *nasuta* subgroup of *Drosophila* based on 12S rRNA, 16S rRNA, and COI mitochondrial genes, RAPD and ISSR polymorphisms. Genes Genet Syst 79: 293–299

Nylander JA (2004) MrModeltest 2.2. Evolutionary Biology Centre, Uppsala University. http://www.ebc.uu.se/systzoo/staff/nylander.html

Okada T (1956) Systematic Study of Drosophilidae and Allied Species of Japan. Gihôdô, Tokyo

Okada T (1988) Taxonomic outline of the family Drosophilidae of Japan. In "Selected Papers by Toyohi Okada" Ed by K Suzuki, College of Liberal Arts, Toyama University, Toyama, pp 1–87

Okada T, Carson HL (1983) The genera *Pholticella* Duda and *Zaprionus* Coquillett (Diptera, Drosophilidae) of the Oriental region and New Guinea. Kontyû, Tokyo 51: 539–533

Pélandakis M, Solignac M (1993) Molecular phylogeny of *Drosophila* based on ribosomal RNA sequences. J Mol Evol 37: 525–543

Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. Bioinformatics 14: 817–818

Powell JR (1997) Progress and Prospects in Evolutionary Biology: The *Drosophila* Model. Oxford University Press, New York

Remsen J, DeSalle R (1998) Character congruence of multiple data partitions and the origin of the Hawaiian Drosophilidae. Mol Phylogenet Evol 9: 225–235

Remsen J, O'Grady P (2002) Phylogeny of Drosophilinae (Diptera: Drosophilidae), with comments on combined analysis and character support. Mol Phylogent Evol 24: 249–264

Robe LJ, Valente VL, Budnik M, Loreto ÉL (2005) Molecular phylogeny of the subgenus *Drosophila* (Diptera, Drosophilidae) with an emphasis on Neotropical species and groups: a nuclear versus mitochondrial gene approach. Mol Phylogenet Evol 36: 623–640

Russo CA, Takezaki N, Nei M (1995) Molecular phylogeny and divergence times of drosophilid species. Mol Biol Evol 12: 391–404

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

Swofford DL (2003) PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, MA

Steller HD (1990) Rapid small-scale isolation of *Drosophila* DNA and RNA. In "Rubin Lab Method Book" 2nd ed Ed by L Rubin, Rubin Lab, pp 185–186

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol 10: 512–526

Tamura K, Toba G, Park J, Aotsuka T (1995) Origin of Hawaiian drosophilids inferred from alcohol dehydrogenase gene sequences. In "Current Topics on Molecular Evolution" Ed by M Nei, N Takahata, Institute of Molecular Evolutionary Genetics, Pennsylvania State University, pp 9–18

Tatarenkov A, Kwiatowski J, Skarecky D, Barrio E, Ayala FJ (1999) On the evolution of *Dopa decarboxylase* (*Ddc*) and *Drosophila* systematics. J Mol Evol 48: 445–462

Tatarenkov A, Zurovcová M, Ayala FJ (2001) *Ddc* and *amd* sequences resolve phylogenetic relationships of *Drosophila*. Mol Phylogenet Evol 20: 321–325

Tavaré S (1986) Some probabilistic and statistical problems in the analysis of DNA sequences. In "Some Mathematical Questions in Biology: DNA Sequence Analysis" Ed by RM Miura, American Mathematical Society, Providence, RI, pp 57–86

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25: 4876–4882

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

Toda MJ (2006) DrosWLD-Species: Taxonomic Information Database for World Species of Drosophilidae, last accessed 1 November 2006. http://bioinfo.lowtem.hokudai.ac.jp/db/modules/stdb

Wakahama K, Shinohara T, Hatsumi M, Uchida S, Kitagawa O (1983) Metaphase chromosome configuration of the *immigrans* species group of *Drosophila*. Jpn J Genet 57: 315–326

Wilson FD, Wheeler MR, Harget M, Kambysellis M (1969) Cytogenetic relations in the *Drosophila nasuta* subgroup of the *immigrans* group of species. Univ Texas Publ 6918: 207–270

Yu H, Wang W, Fang S, Zhang YP, Lin FJ, Geng ZC (1999) Phylogeny and evolution of the *Drosophila nasuta* subgroup based on mitochondrial ND4 and ND4L gene sequences. Mol Phylogenet Evol 13: 556–565

Zhang W, Toda MJ (1992) A new species-subgroup of the *Drosophila immigrans* species group (Diptera, Drosophilidae) with description of two new species from China and revision of taxonomic terminology. Jpn J Ent 60: 839–850

(Received March 16, 2007 / Accepted May 12, 2007)