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

Toru Katoh¹*, Daisuke Nakaya², Koichiro Tamura² and Tadashi Aotsuka²

¹COE for Neo-Science of Natural History, Graduate School of Science, Hokkaido University, North 10 West 8, Kita-ku, Sapporo 060-0810, Japan  
²Department 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. palidifrons*, *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;...
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, quadridilineata, 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, quadridilineata, 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 quadridilineata 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.

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). 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. guttulifera 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. maculipalata 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 CTT 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 30 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...
nucleotide composition bias occurred among taxa, $\chi^2$ 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*. 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.

$\chi^2$ 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 1. Taxa included in the study and Genbank accession numbers for Adh and Gpdh sequences. Sequences obtained in this study are underlined.

<table>
<thead>
<tr>
<th>Genus</th>
<th>Subgenus</th>
<th>Species group</th>
<th>Subgroup</th>
<th>Species</th>
<th>Adh Accession No.</th>
<th>Gpdh Accession No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drosophila</td>
<td>Drosophila</td>
<td>immigrans</td>
<td>hypocausta</td>
<td>hypocausta</td>
<td>AB261133</td>
<td>AB261145</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>neohypocausta</td>
<td>AB261134</td>
<td>AB261146</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>siarana</td>
<td>AB261135</td>
<td>AB261147</td>
</tr>
<tr>
<td></td>
<td></td>
<td>immigrans</td>
<td>formosana</td>
<td>AB261131</td>
<td>AB261143</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>immigrans</td>
<td>M97638</td>
<td>AB261142</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ruberrima</td>
<td>AB261132</td>
<td>AB261144</td>
</tr>
<tr>
<td></td>
<td></td>
<td>nasuta</td>
<td>albomicans</td>
<td>AB033642</td>
<td>AB261148</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>nasuta</td>
<td>AB261137</td>
<td>AB261149</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pallidifrons</td>
<td>AB261136</td>
<td>AB261150</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>neonasuta</td>
<td>AB261138</td>
<td>AB261151</td>
</tr>
<tr>
<td></td>
<td></td>
<td>quadrilineata</td>
<td>annulipes</td>
<td>AB261129</td>
<td>AB261152</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>quadrilineata</td>
<td>AB261130</td>
<td>AB261153</td>
</tr>
<tr>
<td></td>
<td></td>
<td>funebris</td>
<td>funebris</td>
<td>AB033643</td>
<td>AB261154</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>maculinotata</td>
<td>AB261140</td>
<td>AB261156</td>
</tr>
<tr>
<td></td>
<td></td>
<td>guttifera</td>
<td>guttifera</td>
<td>AB261139</td>
<td>AB261155</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>virilis</td>
<td>virilis</td>
<td>AB033640</td>
<td>D10697</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>repleta</td>
<td>hydei</td>
<td>X58694</td>
<td>L41650</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>picture wing</td>
<td>planitibia</td>
<td>M63392</td>
<td>AY006450</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>heteroneura</td>
<td>M36781</td>
<td>AY006454</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>silvestris</td>
<td>M36291</td>
<td>AY006453</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>planitibia</td>
<td>M63930</td>
<td>AY006458</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>differens</td>
<td>M36785</td>
<td>AY006455</td>
</tr>
<tr>
<td>Sophophora</td>
<td>melanogaster</td>
<td>melanogaster</td>
<td>melanogaster</td>
<td>M17833</td>
<td>X14179</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>simulans</td>
<td>M36581</td>
<td>L41647</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>teissieri</td>
<td>X54118</td>
<td>U47809</td>
</tr>
<tr>
<td></td>
<td></td>
<td>obscura</td>
<td>subobscura</td>
<td>X60113</td>
<td>U47808</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>guanche</td>
<td>M60998</td>
<td>L41251</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>miranda</td>
<td>M60979</td>
<td>L41249</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pseudoobscura</td>
<td>M60979</td>
<td>L41249</td>
</tr>
<tr>
<td></td>
<td></td>
<td>willistoni</td>
<td>willistoni</td>
<td>L08648</td>
<td>L41248</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>paulistorum</td>
<td>AB026529</td>
<td>L41648</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>busckii</td>
<td>AB261141</td>
<td>*</td>
</tr>
<tr>
<td>Dorsilopha</td>
<td></td>
<td>pictiventris</td>
<td>pictiventris</td>
<td>AB026530</td>
<td>L41649</td>
<td></td>
</tr>
<tr>
<td>Hirtodrosophia</td>
<td></td>
<td>pallida</td>
<td>pallida</td>
<td>AB033645</td>
<td>AB261157</td>
<td></td>
</tr>
<tr>
<td>Scaptomyza</td>
<td>Parascaptomyza</td>
<td>tuberculatus</td>
<td>tuberculatus</td>
<td>X63955</td>
<td>L37039</td>
<td></td>
</tr>
<tr>
<td>Zapionus</td>
<td>Zapironus</td>
<td>procnemis</td>
<td>procnemis</td>
<td>AB026521</td>
<td>L41252</td>
<td></td>
</tr>
<tr>
<td>Chymomyza</td>
<td></td>
<td>lebanonensis</td>
<td>lebanonensis</td>
<td>X54814</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

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

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

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.
immigrans, funebris, guttata, and virilis-repleta groups) and Dorsilophia (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. guttata 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. palidifrons), and three species of the hypocauсta subgroup (D. hypocauсta, D. neohypocauсta, 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. guttata (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. palidifrons) formed a clade with 100% bootstrap and posterior probability support. Outside the nasuta subgroup, D. immigrans + D. formosana (immigrans subgroup) and D. siamana + D. hypocauсta (hypocauсta subgroup) formed well supported clades. However, the phylogenetic positions of D. ruberrima (immigrans subgroup) and D. neohypocauсta (hypocauсta subgroup) were unstable among the phylogenetic trees deduced by different methods (Fig. 3). Thus, questions remain concerning the monophyly of both the immigrans and hypocauсta subgroups.

*Drosophila quadrilineata* and *D. annulipes*, both of which belong to the quadrilineata group, 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. guttata* 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)
Phylogeny of the *Immigrans* Species Group

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

---

**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.
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. maclinotata, 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 groups 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; Rimens and DeSalle, 1998; Kwiatowski and Ayala, 1999; Tatarenkov et al., 1999, 2001; Katoh et al., 2000; Rimens and O’Grady, 2002; Robe et al., 2005). This clade included species belonging to three species subgroups (hypocaustus, 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 hypocaustus subgroup (D. hypocaustus 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. neohypocaustus was separate from the two other species of the hypocaustus subgroup. The hypocaustus 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 hypocaustus 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 χ² test (Table 2). However, when the χ² test was applied only to the 10 species of the immigrans group proper (clade 6), no significant heterogeneity was detected (χ² = 1.249062, df = 27, P = 1.00).

According to Wilson et al. (1969), the hypocaustus 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. neohypocaustus not to have the body-color dimorphism as strongly developed as in D. hypocaustus, and to have well-developed spinules. In addition, the karyotype of D. neohypocaustus is somewhat different from that of D. hypocaustus (Wakahama et al., 1983). These characters are consistent with those of our results (MP and ME trees in Fig. 3) that indicate D. neohypocaustus as being remote from the other species of the hypocaustus 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. neohypocaustus 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


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


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


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