

EVOLUTION AND PHYLOGENY OF THE DIPTERA: A MOLECULAR PHYLOGENETIC ANALYSIS USING 28S rDNA SEQUENCES

MARKUS FRIEDRICH¹ AND DIETHARD TAUTZ²

¹California Institute of Technology, Division of Biology 156-29, Pasadena, California 91125, USA;

E-mail: mf@cco.caltech.edu

²Zoologisches Institut der Universität München, Luisenstrasse 14, 80333 München, Germany;

E-mail: tautz@zi.biologie.uni-muenchen.de

Abstract.—Portions of the large ribosomal subunit RNA gene (28S rDNA) encompassing the D1 and the D7 region were obtained from 16 dipteran species and families to reconstruct early phylogenetic events in the order Diptera. For outgroup comparison, the corresponding sequences were used from representative taxa of the Siphonaptera, Mecoptera, and Lepidoptera. A subset of 488 unambiguously alignable sites was analyzed with respect to important sequence evolution parameters. We found (1) sequence variability is significantly higher in double-stranded sites than in single-stranded sites, (2) transitions are close to saturation in most pairwise sequence comparisons, (3) significant substitution rate heterogeneity exists across sites, and (4) significant substitution rate heterogeneity exists among lineages. Tree reconstruction was carried out with the neighbor joining, maximum parsimony, and maximum likelihood methods. Four major subgroups are consistently and robustly supported: the Brachycera, the Culicomorpha, the Tipulomorpha sensu stricto, and the hitherto controversial Bibionomorpha sensu lato, which includes the families Sciaridae, Mycetophilidae, Cecidomyiidae, Bibionidae, Scatopsidae, and Anisopodidae. The phylogenetic relationships within or among these subclades and the positions of the families Psychodidae and Trichoceridae were not robustly resolved. These results support the view that the mouthparts of extant dipteran larvae evolved from a derived ground state characterized by subdivided and obliquely moving mandibles. Furthermore, sequence divergence and the paleontological record consistently indicate that a period of rapid cladogenesis gave rise to the major dipteran subgroups. [Character state polarity; Diptera; *Drosophila*; evolution; maximum likelihood ratio test; molecular phylogeny; ribosomal DNA; systematics.]

Because of their abundance and the characteristic modification of the hind wings to halteres, representatives of the holometabolous insect order Diptera are familiar to all of us. The presumably best known dipteran is the fruitfly *Drosophila melanogaster*, which has become one of the most intensively studied model organisms in modern biology (Ashburner, 1989; Lawrence, 1992). Evolution of the genus *Drosophila* has been the subject of several morphological and molecular phylogenetic studies (DeSalle and Grimaldi, 1991; Pélandakis and Solignac, 1993). In recent years, various dipteran clades other than *Drosophila* have been studied using molecular phylogenetic techniques (Vossbrinck and Friedman, 1989; Xiong and Kocher, 1991; Raich et al., 1993; Miller et al., 1996; Pawlowski et al., 1996; Tang et al., 1996; McPheron and Han, 1997; Smith and Bush, 1997). Nonetheless, the basal relationships within the Diptera have not been ad-

dressed with molecular methods. This scarcity of molecular studies addressing the higher systematics of the Diptera is contrasted by a rich body of morphological studies. The first accounts of dipteran phylogeny appeared more than 150 years ago (Lameere, 1906). A large number of studies followed, which derived phylogenetic evidence from cytology, paleontology, or morphology (White, 1949; Rohdendorf, 1964; Hackman and Väistönen, 1985). Most influential were the numerous contributions by Hennig (1948, 1968, 1981), which profited from the application of the principles of phylogenetic systematics (Hennig, 1965). Hennig's final conclusions, which were largely drawn from the analysis of adult characters, predominated dipteran systematics for a long time. The first major attempt to revise dipteran phylogeny was put forward by Wood and Borkent (1989), in which they considered both adult and larval characters. Their study provided the

background for a series of publications on more specific aspects of dipteran phylogeny (Courtney, 1990; Griffiths, 1990, 1994; Oosterbroek and Theowald, 1991; Krzeminski, 1992b; Sinclair, 1992; Sinclair et al., 1994). Most recently, Oosterbroek and Courtney (1995) undertook a new major effort towards dipteran phylogeny by revising and extending the morphological characters used by Wood and Borkent (1989), Courtney (1991), Krzeminski (1992b), and Sinclair (1992) and by applying computer analysis in tree reconstruction.

Nonetheless, many important issues of higher dipteran systematics have remained controversial. There also is an ongoing interest in comparative and evolutionary aspects of this group (Shaw and Meinertzhagen, 1986; Schmidt-Ott et al., 1994; Curtis et al., 1995; Sander, 1996; Melzer et al., 1997). Because of this situation, we investigated the potential of partial sequences from the nuclear large ribosomal RNA (rRNA) gene (28S rDNA) to resolve the early phylogenetic events in dipteran evolution.

Advances and Conflicts in Higher Systematics of the Diptera

The basic concepts of dipteran systematics date back to the early 19th century. At that time, the Diptera were already subdivided into Nematocera, flies with long and evenly segmented antennae, and Brachycera, flies with short antennae built of modified flagellomeres (Latreille, 1802; Macquart, 1834). The >85,000 extant brachycerous species, subdivided into 110 families, stand out against the remaining 35,000 species, thought to represent 27 families (Schumann, 1992) traditionally united as Nematocera. This taxon is generally assumed to be paraphyletic with respect to the Brachycera (Hennig, 1968, 1981; Wood and Borkent, 1989; Oosterbroek and Courtney, 1995). Nonetheless, arguments have been made that support a monophyletic origin of the Nematocera (Ulrich, 1991). The Brachycera is a well-supported monophyletic group, which is subdivided into Cyclorrhapha and Orthorrhapha

(Woodley, 1989). The orthorrhaphous flies are thought to represent a paraphyletic assembly with respect to the Cyclorrhapha, which form a conspicuous monophyletic subgroup supported by numerous synapomorphies (McAlpine, 1989).

The relationships between the major dipteran subgroups are a matter of ongoing controversy. Part of the problem is the uncertainty about the sister group of the Diptera. Candidate groups are the Mecoptera (scorpion flies) (Hennig, 1969; Mickoleit, 1969), the Siphonaptera (fleas) (Boudreaux, 1979; Wood and Borkent, 1989), a subclade including the Mecoptera and the Siphonaptera (Kristensen, 1975), or the mecopteran family Nannochoristidae (Tillyard, 1929; Imms, 1944). Unfortunately, molecular studies based on rDNA sequences, which support either the Lepidoptera (Pashley et al., 1993; Friedrich and Tautz, 1997) or the Strepsiptera (Chalwatzis et al., 1996; Whiting et al., 1997), have failed to convincingly clarify this issue. These studies suffer from the presence of significant substitution rate differences and base compositional shifts in strepsipteran and dipteran rRNA genes, which severely reduce the success of molecular phylogenetic tree estimation (Carmean and Crespi, 1995; Friedrich and Tautz, 1997; Huelsenbeck, 1997). A further reason for controversy over higher dipteran systematics is the general lack of phylogenetically informative character states, although a large number of character state complexes such as general morphology, chaetotaxy, cytology, ecology, and paleontology have been explored (see Hackman and Väistönen, 1982; Wood and Borkent, 1989; Oosterbroek and Courtney, 1995).

To outline the major controversies in dipteran phylogeny, we focus on the concepts of Hennig (1969, 1973, 1981), Wood and Borkent (1989), and Oosterbroek and Courtney (1995). Furthermore, the range of groups discussed is restricted to those included in the present molecular analysis (Table 1). Setting aside terminological differences, all of these studies converge on five major dipteran subgroups (Fig. 1): the Brachycera, the Tipulomorpha or

TABLE 1. Dipteran and outgroup species included in this study. New accession numbers associated with this study are given in bold.

Species	Family	Subgroup	Order	EMBL accession no.
<i>Tipula paludosa</i> Meigen	Tipulidae	Tipulomorpha	Diptera	X93387, X93405
<i>Limonia nebulosa</i> Meigen	Limoniidae	Tipulomorpha	Diptera	X93369, X93378
<i>Bradyisia coprophila</i> Lintner	Sciaridae	Bibionomorpha	Diptera	X93391, X93402
<i>Bolitophila cinerea</i> Mayer	Mycetophilidae	Bibionomorpha	Diptera	X93365, X93373
<i>Dilophus febrilis</i> L.	Bibionidae	Bibionomorpha	Diptera	X93366, X93375
<i>Clinodiplosis cilicrus</i> (Kieffer)	Cecidomyiidae	Bibionomorpha	Diptera	X93363, X93372
<i>Chironomus tentans</i> F.	Chironomidae	Culicomorpha	Diptera	X93383, X93412
<i>Simulium euryadminsterium</i> Davies	Simuliidae	Culicomorpha	Diptera	X93368, X93377
<i>Aedes albopictus</i> Mit.	Culicidae	Culicomorpha	Diptera	L22060
<i>Culex pipiens</i> L.	Culicidae	Culicomorpha	Diptera	X93384, X93403
<i>Psychoda cinerea</i> Banks	Psychodidae	Psychodomorpha	Diptera	X93386, X93404
<i>Trichocera regelationis</i> L.	Trichoceridae		Diptera	X93370, X93379
<i>Anapausis inermis</i> Ruth	Scatopsidae		Diptera	X93364, X93374
<i>Sylvicola fenestratus</i> Scop.	Anisopodidae		Diptera	X93367, X93376
<i>Drosophila melanogaster</i> Meigen	Drosophilidae	Brachycera	Diptera	M21017
<i>Tabanus sudeticus</i> Zell.	Tabanidae	Brachycera	Diptera	X93362, X93371
<i>Manduca sexta</i> L.			Lepidoptera	X93382, X93408
<i>Archaeopsylla erinacei</i> Bouchée			Siphonaptera	X93381, X93407
<i>Panorpa communis</i> L.			Mecoptera	X93380, X93406

crane-fly-like flies, the Psychodomorpha or moth-midge-like flies, the Bibionomorpha or fungus-gnat-like flies, and the Culicomorpha or mosquito-like flies. The composition of some of these clades, however, differs considerably among dipterists. The basic incongruency concerns the Psychodomorpha and the Bibionomorpha. The Bibionomorpha sensu stricto, which traditionally includes the families Bibionidae (March-flies), Sciaridae (dark fungus gnats), Cecidomyiidae (gall midges), and Mycetophilidae (fungus gnats), is generally accepted (Fig. 1). Hennig (1981) also included the families Anisopodidae (wood gnats) and Scatopsidae within the Bibionomorpha, which one may then refer to as Bibionomorpha sensu lato (Fig. 1a). Wood and Borkent, however, included the Anisopodidae and the Scatopsidae together with the Trichoceridae (winter gnats) and the Psychodidae (moth-midges) in the Psychodomorpha, referred to hereinafter as Psychodomorpha sensu W&B (Fig. 1b). Although the Psychodomorpha in the system of Hennig (1981) includes the Psychodidae and a number of other families not included in this study (Fig. 1a), Oosterbroek and Courtney extended the Psychodomorpha sensu W&B by including the families Ti-

pulidae and Limoniidae and the Brachycera (Fig. 1c). Traditionally, however, the families Tipulidae, Limoniidae, and Trichoceridae are considered basal families of the Diptera. According to Hennig (1981), for example, these three families represent the most basal subgroup, here termed Tipulomorpha sensu lato (Fig. 1a). Because the monophyly of the Tipulomorpha sensu lato was rejected by Wood and Borkent in favor of including the Trichoceridae in the Psychodomorpha sensu W&B (Fig. 1b), the general consensus reduces to consider the Tipulidae and Limoniidae sister groups (Fig. 1). The corresponding clade is termed Tipulomorpha sensu stricto. A further major subgroup, for which all authors agree on component families, is the Culicomorpha (Fig. 1). Regarding the culicomorphous families Culicidae (mosquitos), Chironomidae (water midges), and Simuliidae (black flies), most authors assume the latter two to be more closely related.

Given the fundamental differences in the composition the Bibionomorpha and the Psychodomorpha, a comparison concerning the relationships among the major dipteran subdivisions is hardly possible. Nonetheless, Hennig (1981) and Wood and Borkent (1989) favored the Tipulomorpha

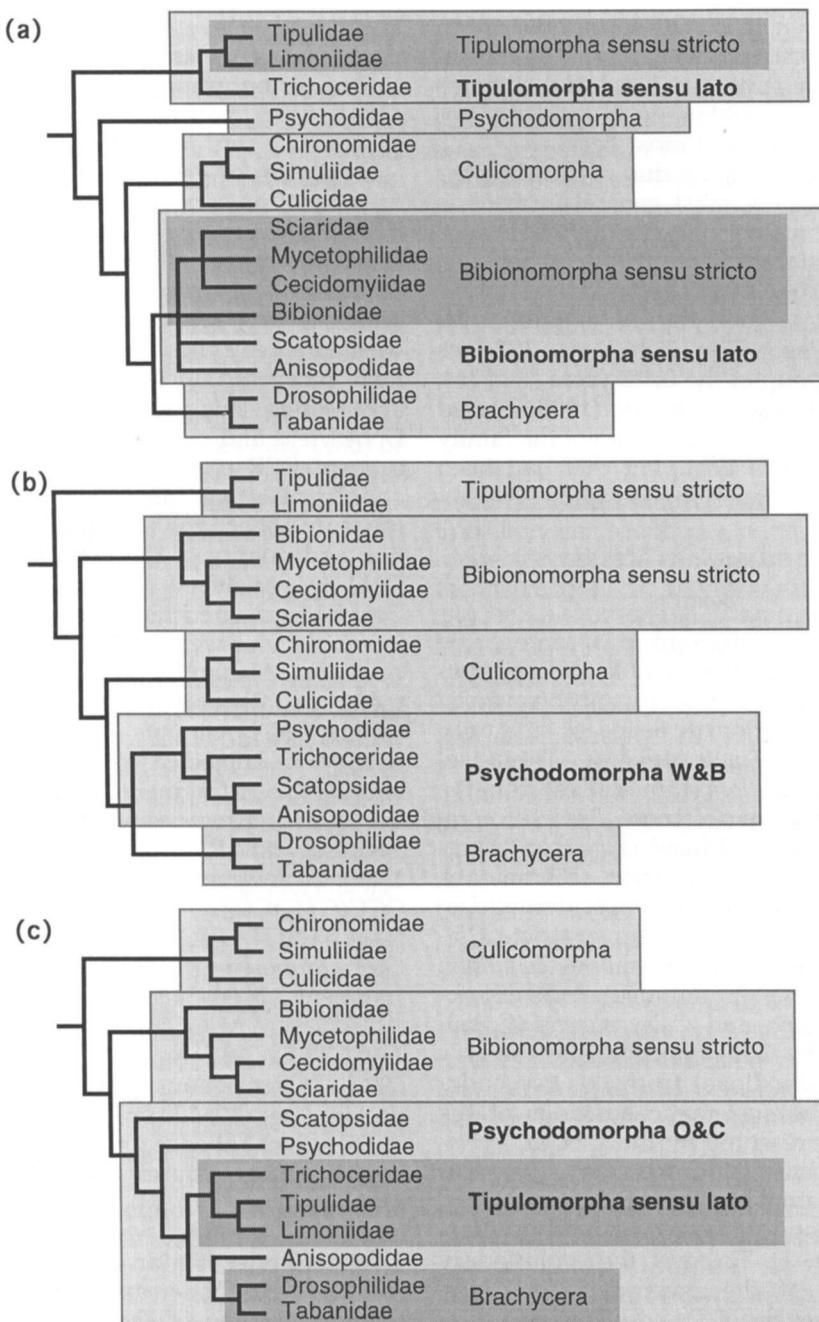


FIGURE 1. Phylogeny of the major dipteran subgroups and families covered in this study. Controversial subgroups are printed bold. Major subgroups are indicated by shaded boxes. Subgroups nested within higher clades are indicated by darker boxes. (a) Hennig, 1973, 1981. (b) Wood and Borkent, 1989. (c) Oosterbroek and Courtney, 1995.

sensu lato or the Tipulomorpha sensu stricto, respectively, as the most basal major dipteran subgroup, and Oosterbroek and Courtney (1995) found evidence for the Culicomorpha as most basal.

The phylogenetic status of the Brachycera, which is of most general interest, is similarly controversial. Hennig (1981) suggested a sister-group relationship between Bibionomorpha sensu lato and the Brachycera (Fig. 1a). Wood and Borkent (1989) favored the Psychodomorpha sensu W&B as the sister group of the Brachycera (Fig. 1b). Oosterbroek and Courtney (1995) placed the Brachycera as sister group of the family Anisopodidae, a grouping that had been proposed previously by Woodley (1989).

MATERIALS AND METHODS

Samples

Each major subgroup was represented by at least one species (Table 1). Representatives of families with uncontroversial systematic position in terms of subgroup allocation are *Tipula paludosa* (Tipulidae) and *Limonia nebulosa* (Limoniidae) from the Tipulomorpha sensu stricto, *Bradysia coprophila* (Sciaridae), *Bolitophila cinerea* (Mycetophilidae), *Dilophus febrilis* (Bibionidae), and *Clinodiplosis cilicrus* (Cecidomyiidae) from the Bibionomorpha sensu stricto, *Chironomus tentans* (Chironomidae), *Simulium euryadmiricum* (Simuliidae), *Aedes albopictus* (Culicidae), and *Culex pipiens* (Culicidae) from the Culicomorpha, and *Psychoda cinerea* (Psychodidae) from the Psychodomorpha. The systematic positions of the families represented by the species *Trichocerca regelationis* (Trichoceridae), *Anapausis inermis* (Scatopsidae), and *Sylvicola fennestralis* (Anisopodidae) are currently controversial (Fig. 1). To cover the evolutionary divergence of the Brachycera, the published sequences of *Drosophila melanogaster* (Drosophilidae, Cyclorrhapha) (Tautz et al., 1988) were retrieved from the EMBL database, and a representative (*Tabanus sudeticus*) of one of the basal families, the Tabanidae (Woodley, 1989), was chosen for sequence analysis. Representatives of the insect orders Lepidoptera (*Manduca sexta*),

Mecoptera (*Panorpa communis*), and Siphonaptera (*Archaeopsylla erinacei*) were included for outgroup comparison.

Methods

Extraction of genomic DNA from dried or alcohol-preserved specimens was carried out according to Gustincich et al. (1991) with the following modifications. DNA was not precipitated but was spin-dialyzed in Ultra free tubes (Millipore), washed three times with 300 µl H₂O, and then recovered in 25–100 µl Tris-EDTA (TE) buffer. This procedure increased the DNA yield and reduced inhibition in subsequent PCR reactions when dried or alcohol-preserved insect specimens from museum collections were used. Concentration and quality of the extracted genomic DNA was analyzed on 0.8% agarose gels. Dried insect specimens generally yielded low-molecular-weight DNA fragments of ≤600 bp. Alcohol-preserved insect specimens usually yielded high-molecular-weight DNA. Aliquots of the DNA extractions were subjected to PCR reactions to amplify DNA fragments corresponding to the following regions of the *Drosophila melanogaster* 28S rRNA sequence (Tautz et al., 1988): D1 fragment, 3338–3650 (primers: 5'-CCC(C/G)CGTAA(T/C)TTAACATAT-3', 5'-ACTCTCTATTCA(A/G)AGTTCTTT(G/C)-3', annealing temperature 60°C); D7 fragment, 5000–5464 (primer: 5'-CTGA AGTGGAGAAGGGT-3', 5'-GACTTCCT TACCTACAT-3', annealing temperature 60°C). PCR reactions were set up with approximately 10–100 ng DNA, 0.2 mM dNTPs, 2 µM of each primer, and 2 mg/ml bovine serum albumin in a buffer of 67 mM Tris/HCl (pH 8.8), 2 mM MgCl₂ (Pääbo, 1990). Cycling was performed in a Thermo Cycler (Perkin Elmer) with 5 min at 94°C for initial denaturation, followed by manual addition of Taq DNA polymerase (Cetus) on the block, then 20 cycles with 1 min denaturation at 93°C, 1 min annealing at the temperature specified for each primer pair above, and 1 min elongation at 72°C. This reaction sequence was followed by 20 analogous cycles with a 2-min 72°C elongation step. Amplified products were

examined on 1.5% agarose gels. DNA fragments were excised from the agarose gel, eluted using glassmilk (Quiaex Kit, Quiagen), and redissolved in 25 μ l TE buffer. About 100 ng of template DNA was used per direct sequencing reaction, which was carried out according to Casanova et al. (1990). DNA fragments were sequenced in both directions with the amplification primers and the following internal primers for the D7 fragment: 5'-AGGGTTTCTGTGAACAG-3', 5'-TTCCAAACC(A/C)TATCTC-3', 5'-CGATTTCAAGGTCC-3'. As judged by gel electrophoresis, the D7a expansion segments (Hancock et al., 1988) of *Tabanus sudeticus* and *Panorpa communis* are about 300 bp longer than those in all other species investigated. These expansion regions were not sequenced completely because no phylogenetically valuable information for the problem addressed here was expected.

Initial multiple alignments were produced by hand and, alternatively, using the CLUSTAL V program applying default settings (Higgins and Sharp, 1988). The resulting alignments were compared and divergent regions were rechecked, taking conservation of secondary structure models previously published for *Drosophila melanogaster* into account (Hancock et al., 1988; Rousset et al., 1991). Nucleotides 162–188 in the *Drosophila melanogaster* 28S D1 sequence provided here (Appendix 1) were accidentally omitted from the correspondent sequence and the secondary structure published by Hancock et al. (1988) and (Tautz et al., 1988) but were verified by direct sequencing (Friedrich, 1995).

Pairwise sequence divergence and pairwise transition (Ts) and transversion (Tv) sequence divergence were determined using subroutines of PAUP 3.0 (Swofford, 1993). The numbers for variable and phylogenetically informative sites were determined using MacClade (Maddison and Maddison, 1992). Stationarity of nucleotide composition was tested using the STATIO program (Rzhetsky and Nei, 1995). Maximum likelihood (ML) estimation of tree likelihoods and substitution parameters were carried out using PAML (Yang,

1996b). ML tree search was carried out using PHYLIP (Felsenstein, 1995) applying the Global Rearrangement option. Heuristic maximum parsimony (MP) tree search was carried out in PAUP applying the TBR option for branch swapping. Estimation of evolutionary distances and neighbor joining (NJ) analysis (Saitou and Nei, 1987) were carried out in PHYLIP. Branch probability (BP) values were determined by nonparametric bootstrapping (Felsenstein, 1985) based on 1,000 replicate data sets.

RESULTS

Multiple Alignment

Correct alignment of homologous positions between sequences is the first prerequisite for reliable molecular phylogenetic reconstruction. In the present study, the 28S D1 and D7 sequence regions had to be aligned among 16 distantly related dipteran species and 3 outgroup species representing lepidopterans, siphonapterans, and mecopterans. 28S rDNA expansion segment regions characteristically accumulate insertion and deletion events over time. These length-variable regions, which were dispersed among several strongly conserved blocks in the present set of sequences, were difficult to align and showed little congruence between hand- and computer-generated alignment (Appendix 1). Reexamination with respect to evolutionarily conserved secondary structure elements as indicated by compensatory substitutions in stem-forming regions helped to refine and confirm the final alignment.

The secondary structure elements supported in our set of sequences were compared with the recently published secondary structure models of *Drosophila melanogaster* and *Aedes albopictus* (Schnare et al., 1996). Other than some minor differences, the stem assumed to be formed by the sites 68–72 and 104–108 in our alignment of the 28S D7 fragment is different from these published models (Appendix 1). The same applies for the stem assumed to be formed by the sites 157–162 and 167–172, and a part of the stem

formed by sites 146–156 and 261–273 in our 28S D7 fragment alignment (Appendix 1). Nonetheless, 488 sites, which included no gaps in any of the dipteran taxa, were aligned with confidence. This subset was considered for molecular phylogenetic analysis.

Sequence Divergence

Of the 488 sites included in the analysis, 211 (43%) are variable and 159 sites (33%) are phylogenetically informative for the dipteran taxa. Within the Diptera, sequence divergence ranges from 2.9% to 24.3% of pairwise sequence divergence (Table 2). Relatively high sequence divergences are observed when representatives of the Culicomorpha (*Chironomus*, *Simulium*, *Culex*, *Aedes*) are compared with other dipteran species (17.4–24.3%). Otherwise, pairwise sequence divergence does not exceed 15.2% within the Diptera. The sequence divergence among the culicomorphous families is also high (16.7–21.1%); as expected, sequence divergence between the closely related mosquito species *Aedes albopictus* and *Culex pipiens* is the lowest observed in the data set (2.9%). The range of sequence divergence observed for the culicomorphous representatives within the Diptera even overlaps with that observed between the dipteran species and the outgroup taxa (22.7–31.0%), indicating elevated substitution rates in the culicomorphous lineages.

If the extracted sequence sites are subdivided into those derived from double-stranded (DS) and single-stranded (SS) regions, 281 sites (58%) are located in putative DS regions and the remaining 207 (42%) are in SS regions. About two thirds of the variable sites (142) are found in DS regions. Sequence variability in DS regions is significantly higher than that in SS regions (χ^2 , $P < 0.005$) if the occurrence of compensatory substitutions in DS regions is neglected. Assuming that all substitutions in DS regions are correlated with a complementary substitution, the number of independent substitution events in the DS regions would reduce to 71 versus 70 complementary pairs of sites where no

TABLE 2. Matrix of pairwise percentage of sequence divergence for dipteran and outgroup taxa.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1. <i>Drosophila</i>	—	8.3	12.9	11.7	12.1	13.6	12.5	12.7	14.4	13.5	13.5	10.8	21.4	21.3	20.6	19.1	24.8	24.1	26.8
2. <i>Tabanus</i>	—	15.2	13.8	13.5	14.3	14.6	14.0	14.5	15.4	11.9	22.1	22.3	22.8	19.4	25.8	24.8	26.6	—	—
3. <i>Bradyvius</i>	—	5.6	3.9	6.8	5.7	6.3	13.9	13.3	14.5	10.2	20.3	20.3	18.1	18.5	25.7	25.4	28.8	—	—
4. <i>Crinodiplosis</i>	—	4.2	6.9	5.6	6.0	13.1	13.1	11.8	9.2	19.6	20.4	17.4	18.7	23.7	24.4	28.1	—	—	—
5. <i>Bolitophilidae</i>	—	4.4	4.6	5.8	12.1	11.8	11.8	11.8	7.3	20.2	20.4	17.6	18.4	23.2	23.5	27.7	—	—	—
6. <i>Anapausis</i>	—	—	6.0	8.2	12.9	12.3	13.1	10.6	21.2	20.9	18.5	18.7	23.6	24.4	28.0	—	—	—	—
7. <i>Diaphlus</i>	—	—	7.6	14.4	12.7	13.9	9.8	20.7	21.1	18.1	18.9	25.3	25.4	28.2	—	—	—	—	—
8. <i>Sylvestris</i>	—	—	—	11.9	13.6	13.3	10.7	21.3	21.8	18.7	18.2	23.7	23.9	27.0	—	—	—	—	—
9. <i>Psychoda</i>	—	—	—	—	14.7	14.8	11.7	23.3	24.3	17.7	19.8	23.6	23.8	26.9	—	—	—	—	—
10. <i>Tipula</i>	—	—	—	—	—	9.8	12.0	23.1	22.7	19.4	20.5	23.3	23.2	27.0	—	—	—	—	—
11. <i>Limonia</i>	—	—	—	—	—	9.0	22.0	22.8	19.7	19.3	23.5	23.5	27.5	—	—	—	—	—	—
12. <i>Trichocera</i>	—	—	—	—	—	—	20.0	19.8	20.2	18.9	22.7	22.5	25.9	—	—	—	—	—	—
13. <i>Aedes</i>	—	—	—	—	—	—	—	—	—	2.9	19.6	20.1	29.6	29.9	30.9	—	—	—	—
14. <i>Culex</i>	—	—	—	—	—	—	—	—	—	—	21.1	30.4	31.0	31.2	31.6	—	—	—	—
15. <i>Chironomus</i>	—	—	—	—	—	—	—	—	—	—	—	27.5	29.4	30.4	30.4	30.4	—	—	—
16. <i>Simulium</i>	—	—	—	—	—	—	—	—	—	—	—	29.3	29.5	30.4	30.4	30.4	4.5	12.1	—
17. <i>Panorpa</i>	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	10.8
18. <i>Archaeopsylla</i>	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
19. <i>Manduca</i>	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—

substitution is observed. Under this most conservative assumption, DS regions are still significantly more variable than SS regions ($P < 0.05$).

Nucleotide Composition

Most models of sequence evolution assume stationarity of nucleotide composition over time (Gillespie, 1986). The impact of base compositional shifts between taxa on the reliability of phylogenetic reconstruction has been discussed (Hasegawa and Hashimoto, 1993; Lockhart et al., 1994). For the present data set, nucleotide composition was determined for DS and SS regions separately (Appendix 2). Consistent with the results of a recent study (Friedrich and Tautz, 1997), a considerable difference between the dipteran taxa and the outgroup taxa was observed in the DS regions. A mean (SD) AT content of 49.4% (2.4%) in the Diptera contrasts with 34.9% (1.3%) in the outgroup taxa. The analogous comparison in the SS regions also shows a tendency towards elevated AT content in the Diptera with 64.4% (1.6%) versus 59.9% (1.3%) in the outgroup taxa.

We tested stationarity of base composition for DS and SS regions using the test developed by Rzhetsky and Nei (1995), which takes into account the phylogenetic relatedness of the sequences being compared. As expected, when the outgroup taxa are included, stationarity of base composition is rejected for DS sites ($P < 0.001$) but not for SS sites. However, stationarity is not rejected for both SS sites and DS sites when the three outgroup taxa are excluded from the test. Because the stationarity assumption can be upheld for the ingroup taxa, it is unlikely that nucleotide composition will have a strong influence on the reconstruction of the higher ingroup taxa.

Degree of Substitutional Saturation

Given the considerable degree of sequence divergence between many of the distantly related dipteran taxa, it appeared necessary to analyze the loss of phylogenetic information due to multiple hits. The degree of substitutional saturation was

characterized by comparing observed pairwise Ts and Tv divergences of all possible species combinations to those expected at substitutional equilibrium, Ts_{sat} and Tv_{sat} , using the approximating formulas: $Ts_{sat} = 2f(\pi_T\pi_C + \pi_A\pi_G)$ and $Tv_{sat} = 2f\pi_C\pi_G$ (Hasegawa et al., 1985). In this context, f is the probability that a given site is variable and π is the stationary nucleotide composition per base; π was set equal to the mean percentage of each nucleotide across all taxa, and f was roughly approximated by setting it equal to the partition of variable sites observed in a given set of sites. Because DS and SS sites differed significantly in sequence variability, saturation analysis was carried out for each subset separately.

Pairwise Ts divergences were plotted against Tv divergences in relation to Ts_{sat}/Tv_{sat} (Fig. 2). The comparison indicates a faster increase of Ts over Tv in DS sites but not in SS sites, which is consistent with the presence of a strong DS site-specific Ts bias previously found in a study of 28S rRNA evolution in closely related *Drosophila* species (Rousset et al., 1991). For sequence saturation, it becomes obvious that in the DS sites Ts have reached the range of substitutional equilibrium for the majority of pairwise sequence comparisons whereas Tv have remained clearly distant from Tv_{sat} (Fig. 2a). In the SS sites, however, Ts and Tv are affected to same degree by multiple hits (Fig. 2b). In approximately one third of the pairwise sequence comparisons, both classes of substitutions have reached the range of expected substitutional equilibrium.

To see which aspects of tree topology might be most affected by the considerable amount of substitutional saturation observed, we subdivided the data according to taxon combinations (Fig. 2). The strongest impact was on sequence divergences between ingroup and outgroup taxa. Furthermore, those pairwise sequence comparisons within the Diptera that involved at least one culicomorphous taxon yielded a distinct cluster of data points that are closer to saturation than are any of the data derived from the rest of the possible pairwise sequence combinations among

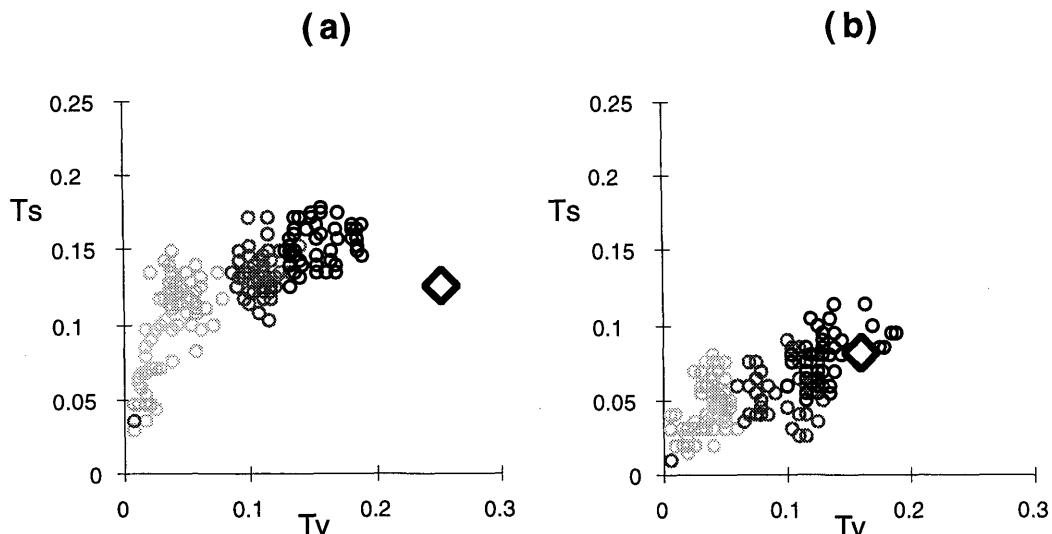


FIGURE 2. Observed average transition (T_s) and transversion (T_v) divergence per nucleotide. Light circles = pairwise comparisons between dipteran taxa except Culicomorpha; medium circles = pairwise comparisons between dipteran taxa including culicomorphous taxa; heavy circles = pairwise comparisons between dipteran and outgroup taxa. Heavy diamond = $T_{s_{\text{sat}}}/T_{v_{\text{sat}}}$. (a) DS sites. f is assumed to be 0.51, and stationary nucleotide composition is $\pi_A = 0.21$, $\pi_C = 0.26$, $\pi_T = 0.25$, $\pi_G = 0.28$. (b) SS sites. f is assumed to be 0.33, and stationary nucleotide composition is $\pi_A = 0.37$, $\pi_C = 0.14$, $\pi_T = 0.26$, $\pi_G = 0.23$.

the Diptera. The only exception in this respect is the data point obtained from the closely related culicomorphous taxa *Aedes* and *Culex*, which falls into the cluster of data derived from nonculicomorphous Diptera. Thus overall, the 28S rDNA sequences seem to have retained phylogenetic information in the majority of ingroup taxa included. Nonetheless, ambiguous results may be anticipated concerning the position of the Culicomorpha within Diptera and concerning the root of the tree based on the extremely diverged outgroup sequences. The rooting problem reduces the chances of accurately reconstructing the most basal splits among the Diptera.

Rate Heterogeneity across Sites

Substitution rates across sites are unequally distributed in most sequences and may best be fitted to a gamma distribution (Uzzell and Corbin, 1971). The shape parameter α of the gamma distribution, which is negatively correlated with the extent of rate heterogeneity, is commonly used as a measure for rate heterogeneity

across sites. In real sequences, typical values for α range from 0.1 to 1, and the importance of accommodating rate heterogeneity across sites in tree estimation has been well established in recent years (Yang, 1996a).

To test the present data for the significance of rate heterogeneity across sites, we employed the likelihood ratio test (see Huelsenbeck and Rannala, 1997). As the null hypothesis (H_0), rate homogeneity across sites was assumed; the alternative hypothesis (H_1) was rate heterogeneity across sites. In this case, two times the difference in logarithmic likelihood ($2\Delta \log L$) is approximately χ^2 distributed with one degree of freedom. Rate heterogeneity was accommodated under the F84 model of sequence evolution (Felsenstein, 1995) by applying the discrete gamma model (Yang, 1994), which approximates the gamma distribution by allowing a number of different rate categories with equal probability of occurrence along the sequence (F84-d Γ). We allowed four rate categories, which gives sufficiently good approximation of the

TABLE 3. Likelihood ratio test results for rate heterogeneity across sites and among lineages for dipteran sequences.

Tree	H_0			H_1			$\log L_1$	$2\Delta \log L$	P
	Model	Clock	$\log L_0$	Model	Clock				
T-MP ₁₉	F84	no	-3681.6	F84-dΓ	no	-3454.5	454.2	<0.001	
T-MP ₁₉	F84-dΓ	yes	-3557.4	F84-dΓ	no	-3454.5	205.9	<0.001	
T-MP ₁₆	F84-dΓ	yes	-2934.1	F84-dΓ	no	-2915.1	37.9	<0.001	
T-MP ₁₂	F84-dΓ	yes	-2102.1	F84-dΓ	no	-2088.3	27.4	<0.005	

gamma distribution (Yang, 1994). Because rate heterogeneity is overestimated if an unrealistic topology such as a star tree is assumed (Sullivan et al., 1996), we used the single most-parsimonious tree derived from the data (T-MP₁₉) (not shown) for ML estimation of rate heterogeneity across sites. Accordingly, if the F84-dΓ sequence evolution model is assumed, substitution rates are gamma distributed across the 28S rDNA sequences with an α of 0.37 (rates: 0.013, 0.16, 0.7, 3.13). As expected, the likelihood ratio test shows that accounting for rate heterogeneity across sites (F84-dΓ) improves the likelihood of the T-MP₁₉ topology significantly (Table 3). Thus, there is significant rate heterogeneity across sites in the present set of 28S rDNA sequences that must be accounted for in tree estimation.

Rate Heterogeneity among Lineages

Strong differences in substitution rates among lineages can cause tree estimation methods to become misleading (Felsenstein, 1978; Huelsenbeck and Hillis, 1993). The rDNA of dipterans evolves significantly more quickly than does that of most other insects (Carmean et al., 1992; Friedrich and Tautz, 1997). There was thus reason to assume significant taxon-specific substitution rate differences between ingroup and outgroup taxa in the present data. In addition, the sequence variability analysis indicated accelerated substitution rates in the culicomorphous lineages (Table 2). We therefore also tested homogeneity of substitution rates among lineages using likelihood ratio test statistics (Felsenstein, 1981).

The null hypothesis of rate homogeneity among lineages, i.e., a molecular clock, was compared with the alternative hy-

pothesis, which allows for rate heterogeneity among lineages. Thereby $2\Delta \log L$ is χ^2 distributed with number of species included minus two degrees of freedom. If all taxa are included (T-MP₁₉), $2\Delta \log L$ is very large and the molecular clock is rejected with high significance (Table 3). If the three outgroup taxa are excluded from the test (T-MP₁₆), $2\Delta \log L$ drops considerably but the molecular clock is still rejected with very high probability. Even if the four quickly evolving culicomorphous ingroup species are excluded from the test (T-MP₁₂), the molecular clock is rejected, although with reduced probability. These results indicate that the outgroup taxa and the culicomorphous taxa contribute most of the substitution rate heterogeneity among lineages, reinforcing the caveats derived from the saturation analysis regarding the inference of the position of the Culicomorpha and the accuracy of tree rooting.

The ML ratio test rejects rate homogeneity among the dipteran lineages when the taxa with most obviously accelerated substitution rates are excluded from the test. Previous relative rate tests of 28S rDNA sequences, which were carried out according to Wu and Li (1985), did not reject the molecular clock when the substitution rates of nonculicomorphous dipteran taxa were compared (Friedrich and Tautz, 1997). This discordance might be explained by the fact that the χ^2 distribution is an approximate null distribution of the likelihood ratio test statistics. It therefore seems advisable to determine the null distribution of the likelihood ratio test statistics by Monte Carlo simulation (Goldman, 1993) to confirm the present result.

Tree Estimation

Tree estimation was guided by the objective of accounting for sequence evolution parameters that have a possible bearing on the success of tree estimation methods, such as rate heterogeneity across sites and unequal Ts and Tv rates. We employed the MP method to construct an initial topology for the estimation of the relevant substitution parameters. Although MP is misleading in cases of extreme rate heterogeneity among lineages (Felsenstein, 1978), this method has the advantage of requiring no explicit assumptions and being efficient in recovering long true branches (Huelsenbeck and Hillis, 1993). To account for the correlated sequence evolution of complementary sites in DS regions, MP tree reconstruction was carried out weighting DS regions 0.8 over SS regions as proposed by Dixon and Hillis (1993). Under these assumptions, MP finds a single most-parsimonious tree (T-MP₁₉) with a length of 5,686 (consistency index = 0.587; retention index = 0.621). Several clades in this tree are highly supported, as indicated by bootstrap analysis (Fig. 3).

T-MP₁₉ was then used for ML estimation of sequence evolution parameters assuming the F84-d Γ model of sequence evolution, which allows for different Ts and Tv rates, unequal stationary base composition, and rate heterogeneity across sites (Yang, 1994; Felsenstein and Churchill, 1996). Under these assumptions, a Ts/Tv rate ratio parameter κ of 1.44 and an α parameter of 0.37 were estimated. Employing these parameters in the ML tree search resulted in a best ML F84-d Γ tree ($\log L = -3442.5$) that included all of the ingroup clades that were highly supported with MP (Fig. 3). To get an estimation of branch probabilities independent from MP, bootstrap analysis was carried out applying the NJ method. Evolutionary distances (K80- Γ) were estimated according to Jin and Nei (1990), combining the two-parameter sequence evolution model (Kimura, 1980) with a continuous gamma distribution model to account for rate heterogeneity across sites.

The same substitution parameter settings were used as in the ML tree search.

Both MP and NJ K80- Γ find strong support for the monophyly of the Diptera versus the outgroup taxa chosen (BP = 100) (Fig. 3). To some extent, this result may be due to the episodically accelerated rate of rDNA evolution in the stem group in the Diptera (Friedrich and Tautz, 1997), which leads to an exceptionally high number of molecular synapomorphies as is clearly reflected in the extremely long branch that joins the outgroup taxa with the Diptera in the ML F84-d Γ phylogram.

Within the Diptera, three major clades of higher systematic level are very strongly supported (BP > 95). One of these includes *Drosophila melanogaster* and *Tabanus sudeticus*, which represent the subgroup Brachycera. A second clade includes six taxa: *Bradysia coprophila*, *Dilophus febrilis*, *Bolitophilidae cinerea*, *Clinodiplosis cilicrus*, *Anapausis inermis*, and *Sylvicola fenestralis*. The first four of these represent the families Sciariidae, Bibionidae, Bolitophilidae, and Cecidomyiidae, respectively, and thus the Bibionomorpha sensu stricto. The remaining two species represent the Scatopsidae and Anisopodidae, respectively. The whole clade therefore corresponds to the controversial Bibionomorpha sensu lato. The third significantly supported major subclade includes *Chironomus tentans*, *Simulium euryadmiricum*, *Culex pipiens*, and *Aedes albopictus*, thus representing the Culicomorpha. A fourth major clade that is strongly supported (BP > 80) is the Tipulomorpha sensu stricto, represented by *Tipula paludosa* and *Limonia nebulosa*.

All branches interconnecting these four major clades and the two remaining taxa, *Psychoda cinerea* and *Trichocera regelationis*, are poorly supported by the data (BP < 60) and to a large extent not consistently resolved by the different methods. In the ML F84-d Γ tree for example, the Tipulomorpha sensu stricto splits off most basally from the rest of Diptera, whereas in the NJ K80- Γ tree and in T-MP₁₉, the Culicomorpha is most basal (not shown). Obviously, the phylogenetic information available allows inference of the composition of major dip-

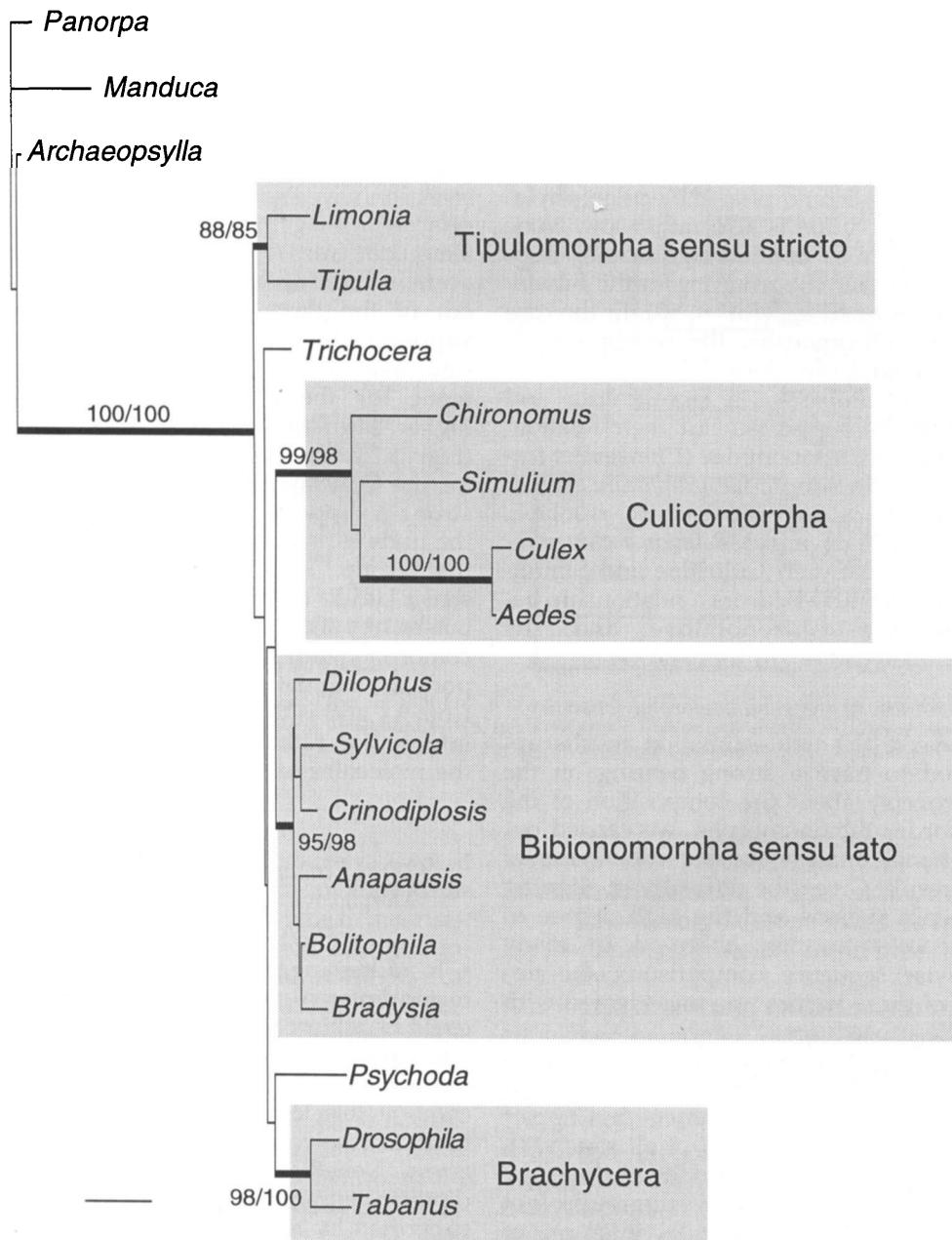


FIGURE 3. ML F84-d Γ dipteran phylogram. Bold branches are strongly supported by MP and NJ K80- Γ . Numbers at bold branches are BP for MP and NJ K80- Γ , respectively. Branches without numbers are supported by BP < 80 and may not be consistently reconstructed by NJ K80- Γ or MP. Significantly supported major subgroups are indicated by shaded boxes. Line at the bottom represents 0.1 substitutions per base.

teran infraorders but not the branching events among them. Such ambiguity in the estimation of the basal relationships was expected based on the results of the sequence saturation analysis.

Within the Bibionomorpha sensu lato and the Culicomorpha, molecular phylogenetic resolution is also rather low. None of the methods applied significantly supported further subgrouping in the Bibionomorpha sensu lato ($BP < 50$). In the case of the Culicomorpha, the monophyly of the two mosquito species, *Culex pipiens* and *Anopheles albopictus*, is significantly and consistently supported but their relationship to the Chironomidae (*Chironomus tentans*) and the Simuliidae (*Simulium euryademinicum*) is not consistently resolved. The ML F84-d Γ and MP favor a closer relationship between Culicidae and Simuliidae, but NJ K80- Γ favors a relationship between the Chironomidae and the Simuliidae ($BP = 73$) (not shown).

Robustness of the Bibionomorpha sensu lato

Because the tree estimation results appeared to have a strong bearing on the controversy about the composition of the infraorder Bibionomorpha, we carried out additional analyses of the robustness of this result. Given the difficulty in aligning sequence regions and the high degree of sequence saturation observed in many pairwise sequence comparisons, the impact of these factors was investigated with simple procedures.

To test for alignment effects, phylogenetic reconstruction was carried out based on the CLUSTAL V alignment and applying NJ K80 after removal of all sites with gaps. This procedure produced the same results concerning major subgroups (not shown). Thus, the molecular inference of the major subgroups is robust against changes in the more divergent parts of the alignment.

Weighted MP was applied to investigate the influence of homoplasy. Because T_v were far from saturation for most pairwise sequence comparisons in both DS and SS sites (Fig. 2), this class of substitutions was expected to be less affected by homoplasy

than T_s . A search for MP trees based on T_v only yielded two equally most-parsimonious trees (Fig. 4a), which include the three major subgroups that are significantly supported by unweighted MP or NJ K80- Γ (Fig. 3). For comparison, MP reconstruction was also carried out based on T_s only, which again yielded two most-parsimonious trees (Fig. 4b). These trees, however, include only the Brachycera as one out of the three otherwise significantly supported higher clades. The effect of T_s saturation was expected to be most dramatic for the Culicomorpha given the higher saturation level in this subclade (Fig. 2). The lack of support by T_s -only MP for the Culicomorpha, which is otherwise strongly supported, is thus consistent with the finding that T_s are largely saturated in this group. Finding the Bibionomorpha sensu lato to be recovered by T_v -only MP but not by T_s -only MP, as with the Culicomorpha, indicates that the strong support in the data predominantly derives from T_v and is thus based on phylogenetic information rather than on convergence at the molecular level.

DISCUSSION

Tree Evaluation with Reference to Morphological Hypotheses of Higher Dipteran Systematics

In addressing the earliest splits among the major dipteran groups, no robust resolution could be achieved with any of the three tree estimation methods employed. Thus, the molecular data do not allow reliable inferences at this level, which proved similarly hard to resolve using morphological characters. Nonetheless, there are three major dipteran infraorders that are significantly supported by all of the tree reconstruction methods applied. Given that consistency among tree reconstruction methods correlates positively with the accuracy of estimated trees (Kim, 1993) and that tree estimation bias problems are unlikely, these major subgroups may be considered as reliably inferred (Fig. 5). In each case, their composition is in accordance with groups that had been postulated previously on the

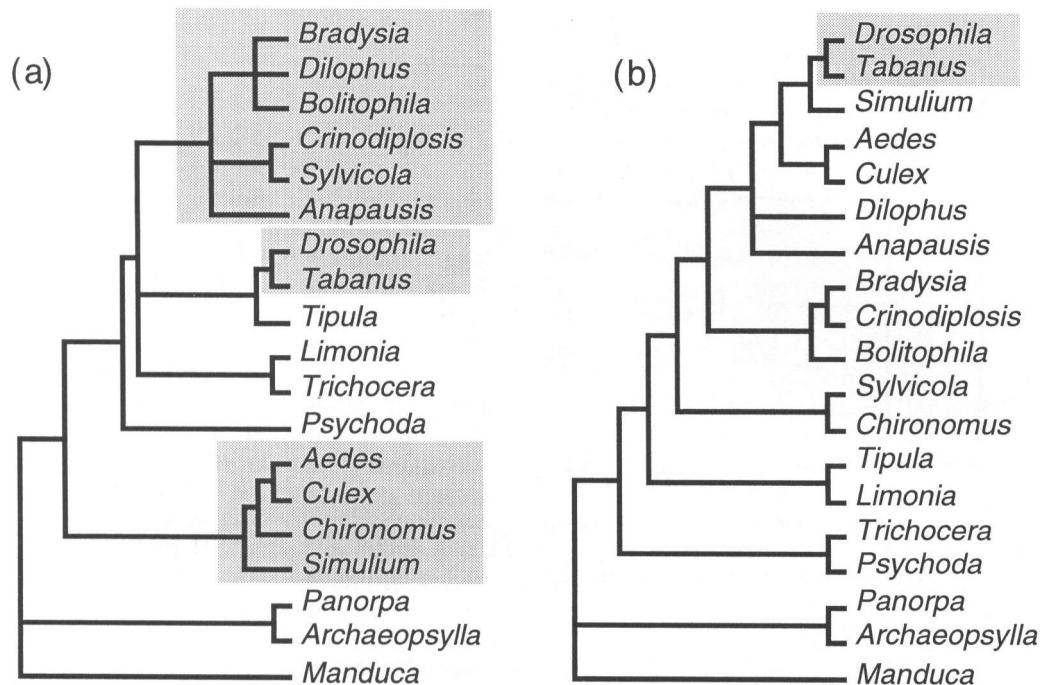


FIGURE 4. Weighted MP dipteran trees. Higher taxa that are significantly supported when all character state changes are considered are indicated by shaded boxes. (a) Strict consensus of two shortest MP trees based on Tvs only and weighting DS sites 0.8 over SS sites. (b) Strict consensus of two shortest MP trees based on Ts only and weighting DS sites 0.8 over SS sites.

basis of morphological characters. Two of these clades, the Brachycera and the Culicomorpha, are groups on which most authors agree because they are clearly supported by a large number of morphological characters (Hennig, 1981; Wood and Borkent, 1989; Oosterbroek and Courtney, 1995). For the monophyly of the Bibionomorpha sensu lato, however, no clearly derived morphological character states could be put forward in spite of overall similarity of the adult phenotype. Hennig (1981) favored the Bibionomorpha sensu lato based on reductional similarities in the wing venation. More recently, many authors have followed Wood and Borkent (1989), who rejected the monophyly of the Bibionomorpha sensu lato (Courtney, 1990; Griffiths, 1990; Sinclair, 1992; Oosterbroek and Courtney, 1995). However, our results clearly support the monophyly of the controversial Bibionomorpha sensu lato and indicate that the Anisopodidae or Scatopsidae are less closely re-

lated to the Psychodidae than to the Bibionomorpha sensu stricto, as implied in the Psychodomorpha sensu W&B or in the Psychodomorpha sensu Oosterbroek and Courtney.

A fourth clearly and consistently supported higher infraorder is the Tipulomorpha sensu stricto, which is consistent with the general consensus on this group (Hennig, 1981; Wood and Borkent, 1989; Oosterbroek and Theowald, 1991). Regarding the family Trichoceridae, however, which according to many authors is the sister group of the Tipulomorpha sensu stricto (Hennig, 1981; Griffiths, 1990; Oosterbroek and Theowald, 1991; Oosterbroek and Courtney, 1995), our results provide no further clue. Nonetheless, the rejection of the Psychodomorpha sensu W&B rules out the only morphology-based hypothesis that had questioned the Tipulomorpha sensu lato.

Thus, despite limitations with respect to phylogenetic resolution and taxon choice,

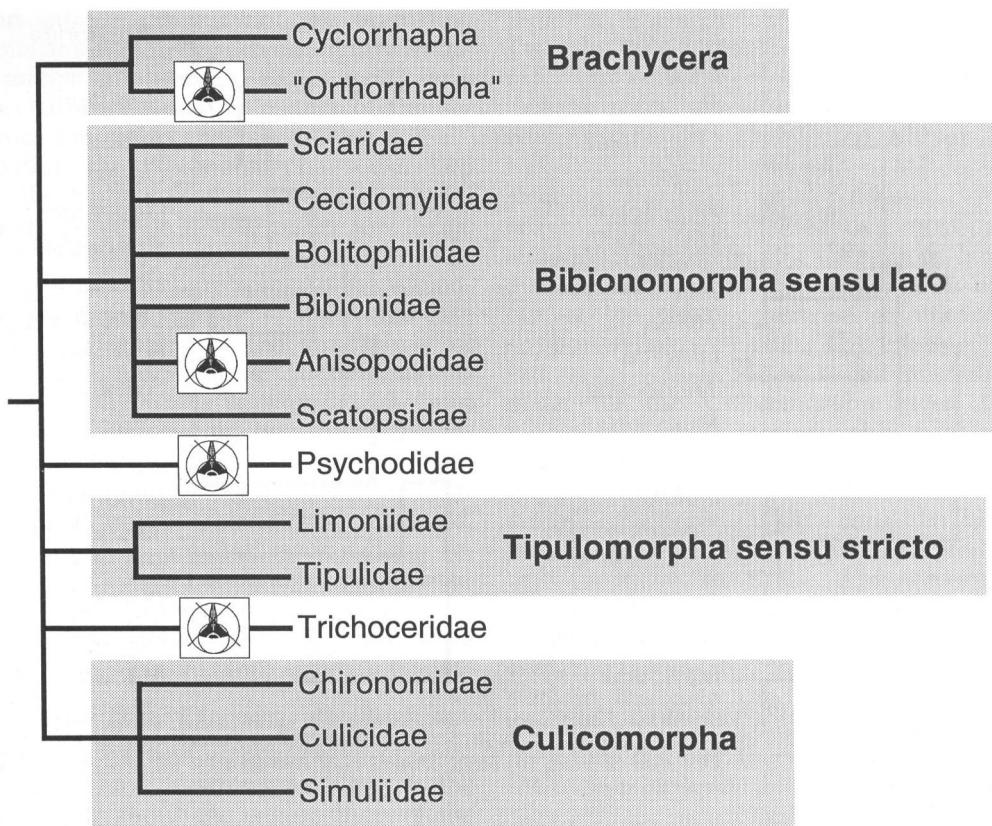


FIGURE 5. Consensus tree of major dipteran subgroups reliably reconstructed in this study. Occurrence of the obliquely moving and subdivided larval mandible types in a lineage is indicated by boxed symbols. Higher subgroups are indicated by shaded boxes.

the present results have important implications for morphological studies of dipteran systematics. A further conclusion relates to the controversial discussion of the sister group of the Brachycera (Griffiths, 1994). The family Anisopodidae was proposed as sister group of the Brachycera (Woodley, 1989; Oosterbroek and Courtney, 1995), which would render the Bibionomorpha sensu lato paraphyletic with respect to the Brachycera (Fig. 1c). The high support for the monophyly of the Bibionomorpha sensu lato, however, excludes this possibility. Although the position of the Brachycera differed among the molecular trees, in no tree was monophyly of the Nematocera supported. Thus, the Brachycera/Bibionomorpha sensu lato sister-group hypothesis remains, which for dif-

ferent reasons had previously been proposed by Hennig (1981) and by Colless and McAlpine (1970).

In conclusion, the molecular phylogenetic approach was powerful enough to provide a new framework for future morphological studies of dipteran phylogeny. The present study is complemented by a recent analysis of the infraorder Culicomorpha based on a different set of 28S rDNA sequences (Pawlowski et al., 1996). Although the basal relationships of the Culicomorpha have remained ambiguous as in our study, significant support for a sister-group relationship between the Simuliidae and the problematic family Thaumaleidae was found, which is noteworthy given that some authors proposed to include the Thaumaleidae in infraorders different from the Cul-

icomorpha (Hackman and Väisänen, 1982; Colless and McAlpine, 1991).

Evolution of the Dipteran Larval Mandible: A Paradigm Case for the Argument of Commonality

Of central importance for the contemporary discussion of higher dipteran systematics is the interpretation of the evolutionary pathways of the larval mouthparts in the Diptera. Many groups in the system of Wood and Borkent (1989) rely on arguments derived from this character state complex. Because some of the characters such as labral shape and mandibular swinging orientation may well be correlated (Teskey, 1981), the discussion may be focused on the long-standing dissent over the evolution of the larval mandibles. The underlying problem of determining character state polarity has been more or less explicitly discussed by several authors. Because neither new paleontological nor new ontogenetic data are at hand, one still depends on indirect evidence, which is commonly derived from outgroup comparison. The larvae of other holometabolous insect orders usually possess solid and horizontally moving mandibles, as is also the case for the nymphal stages of hemimetabolous insect orders (Snodgrass, 1935). This type of larval mandible is regarded as a groundplan feature of the holometabolous insects. Similar types of mandibles can be found in representatives of several dipteran families such as the Tipulidae, Limoniidae, Bibionidae, Mycetophilidae, and Sciaridae. Based on outgroup comparison, many comparative morphologists (Goethgebuer, 1925; Snodgrass, 1935; Cook, 1949; Gouin, 1959) have considered this character state ancestral for the Diptera (Fig. 6a). Some authors, however, emphasized the occurrence of strikingly similar obliquely or vertically moving and subdivided mandibles in representatives of distantly related families such as the Trichoceridae, Anisopodidae, and Psychodidae (Edwards, 1926; Anthon, 1943; Schremmer, 1951). They interpreted the phylogenetically wide distribution of this character state within the Diptera as due to evolutionary conser-

vation of an ancestral ground state (Fig. 6b). Hennig (1948, 1981) realized the contradiction between adult and larval character states in higher dipteran systematics and explicitly adopted the view of those who favored obliquely and subdivided mandibles as symplesiomorphic. Formally, the argument underlying this hypothesis of character state polarity corresponds to the principle of commonality or ingroup distribution (de Jong, 1980). In general, the commonality criterion is considered less stringent than outgroup comparison (Kitching, 1992) because it depends on some reliable knowledge of the ingroup phylogeny.

In the present case, this knowledge is provided by the molecular phylogenetic analysis. The distribution of larval mandible character states can be mapped on a cladogram that includes the dipteran subgroups that were reliably inferred (Fig. 5). The combined occurrence of the two character states, subdivision and oblique or vertical movement of the larval mandible, can be noted in representatives of at least four major dipteran lineages, the Brachycera, the Bibionomorpha sensu lato, the Psychodomorpha, and the Trichoceridae. The controversial homology of mandibular structures in the Cyclorrhapha excludes this group from the inference (Sinclair, 1992; Griffiths, 1994). Moreover, larvae of the limoniid genera *Pilaria* and *Ullomorpha* possess subdivided and obliquely moving mandibles that are considered to be structurally very different from the subdivided mandibles found in the other dipteran infraorders (Oosterbroek and Theowald, 1991; Oosterbroek and Courtney, 1995). Oblique mandible movement, however, has been proposed to be ancestral for the Tipulomorpha sensu stricto (Oosterbroek and Theowald, 1991; Oosterbroek and Courtney, 1995). If restricted to uncontroversial evidence, the wide distribution of subdivided and obliquely moving mandibles among the Diptera is confirmed and may be interpreted to indicate plesiomorphy according to the criterion of commonality.

Thus, during the evolution of the dipteran stem group the solid and horizontally

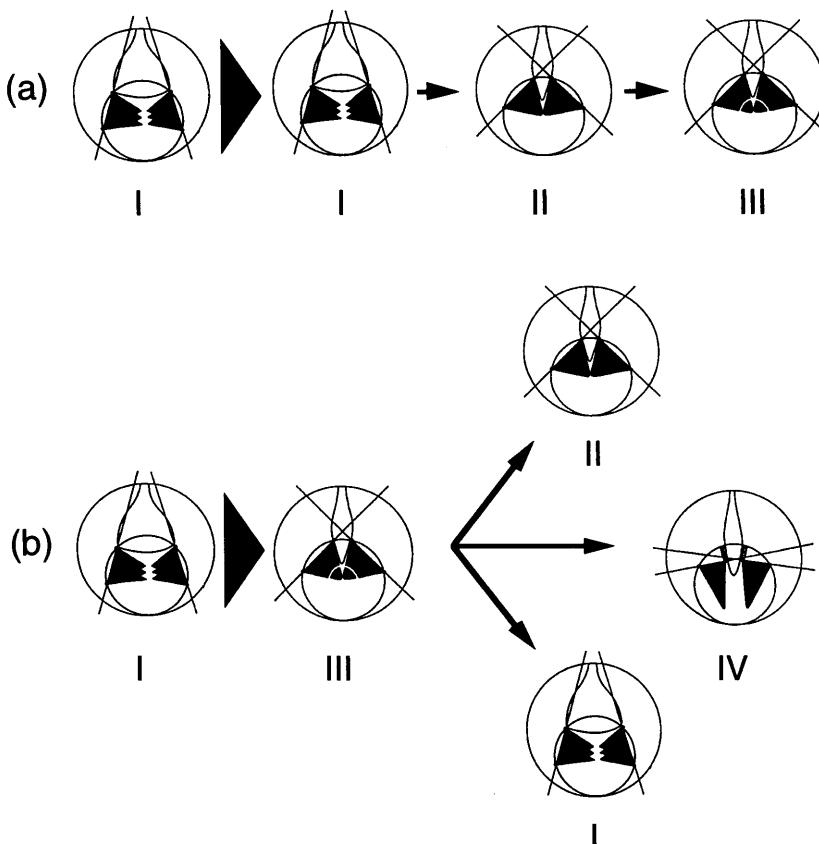


FIGURE 6. Hypotheses of the evolution of the dipteran larval mandibles. Thick arrows indicate transformation pathways in the dipteran stem group; thin arrows indicate transformation pathways in the dipteran crown group. (a) Gradual evolution of solid, horizontally moving mandibles (I) of the dipteran stem group into solid, obliquely or vertically moving mandibles (II) and subsequently into subdivided, obliquely moving mandibles (III). (b) Transformation of solid, horizontally moving mandibles into subdivided, obliquely moving mandibles during evolution of the dipteran stem group, followed by diversification of this ground plan into various types of mandibles observed in modern dipterans, such as solid, vertically moving mandibles (IV).

moving larval mandible type seen in larvae of most other holometabolous insect orders was most probably transformed into the derived type of subdivided and obliquely moving mandibles (Fig. 6b). The solid and horizontally moving biting mandibles in representatives of the families Bibionidae or Sciaridae must then be considered examples of striking convergence.

Early Dipteron Diversification according to Molecular and Fossil Record

The lack of molecular phylogenetic resolution among the major dipteran subgroups is striking when contrasted with

the strong support for the Diptera as such and for each major subgroup in this study. The considerable amount of sequence saturation in the data may obscure the basal relationships. The branching pattern in the molecular phylogram however provides evidence for yet an alternative explanation for the lack of resolution (Fig. 3). The branches that join the major groups are very short, suggesting that the diversification of the major subgroups represents a period of rapid cladogenesis in the evolution of the Diptera. A similar picture emerged from a recent analysis of longer 28S rDNA sequences from a sample of spe-

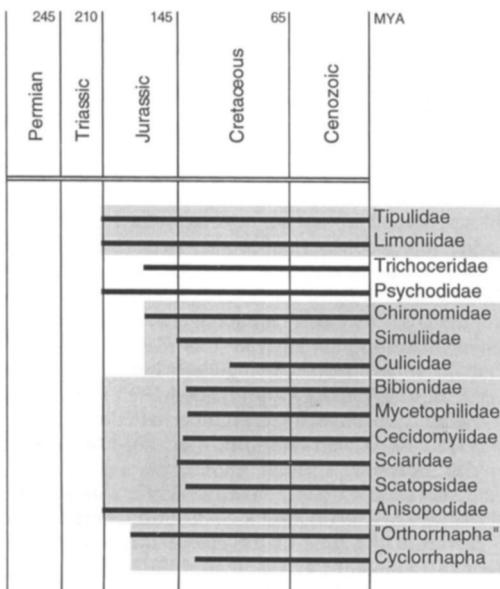


FIGURE 7. Paleontological record of dipteran families or subgroups included in this study. MYA = million years ago. Higher subgroups are indicated by shaded boxes.

cies included here (Friedrich and Tautz, 1997). This pattern of divergence can be compared with the paleontological data on earliest occurrences of dipteran families in the fossil record (Fig. 7).

The origin of the Diptera most probably dates back into the Upper Permian, 250 million years ago (MYA), although the fossil evidence for flies from that period is virtually nonexistent and difficult to interpret (Tillyard, 1929; Hennig, 1981; Willmann, 1989; Wootton and Ennos, 1989; Krzeminski, 1992a). Nonetheless, the finding of representatives of various dipteran subgroups in the Late Triassic (Krzeminski, 1992b; Fraser et al., 1996) quite convincingly documents the diversification of some major dipteran sublineages by the middle Mesozoic. Upper Triassic species of the crown group Diptera are known from locations in Australia (Evans, 1971) and North America (Olsen et al., 1978; Krzeminski, 1992b; Fraser et al., 1996). The North American site and a Triassic site in Russia contain a fly related to the modern nematocerous family Anisopodidae (Krzeminski, 1992a; Fraser et al., 1996).

The North American site also contains a number of additional fly taxa. Among those, Krzeminski (1992a) identified a second anisopodid-like family (Procramptonomyiidae) and representatives of the modern subgroups Limoniidae and Brachycera. Fraser et al. (1996) recently extended the list of North American Triassic dipterans by adding descriptions of psychodid fossils. Thus, representatives of four major modern dipteran subdivisions appear almost simultaneously in the paleontological record 220 MYA. Although the possibility of sampling artifact must be considered, the consistency between the paleontological and molecular data concerning a rapid diversification of the major lineages is notable. The amount of phylogenetic information that can accumulate for a clade is correlated with the time span of the existence of the respective stem lineage. The small amount of phylogenetic information documenting the earliest splits in the Diptera at both the molecular and the morphological level is thus also consistent with a rapid diversification of the major lineages.

ACKNOWLEDGMENTS

We thank Friedrich Reiss and Wolfgang Schacht for supplying specimens, for help with species determination, and for their interest, Hans Meyer for determining *Clinodiplosis ciliatus*, Monika Retzlaff for a DNA sample of *Psychoda cinerea*, Ralf Sommer for DNA samples of *Bradyzia coprophila*, *Simulium euryadiminulum*, and *Chironomus tentans*, Svante Pääbo for some of the primers, Alexander Riedel for help with PCR and sequencing, Hans Ulrich, Pjotr Oosterbroek, and Herman de Jong for help with systematic literature, Ewa Krzeminska for detecting a mistaken species name, and Kimberlee Wollter for improving the language of the final manuscript. We gratefully acknowledge critical reading of an early version of this manuscript by Bernhard Hausdorf and Roland Melzer, suggestions on the final version by Pjotr Oosterbroek and Greg Courtney, and detailed comments by Chris Simon, David Grimaldi, Timothy Collins, and an anonymous reviewer. This work was supported by a Ph.D. studentship awarded to M.F. by the University of Munich.

REFERENCES

- ANTHON, H. 1943. Der Kopfbau der Larven einiger nematocerous Dipterenfamilien: Rhypidae, Trichoceridae, Psychodidae und Ptychopteridae. Spolia Zool. Mus. Haun. 3:1–61.

- ASHBURNER, M. 1989. *Drosophila*. A laboratory handbook. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- BOUDREAUX, H. B. 1979. Arthropod phylogeny, with special reference to insects. Wiley & Sons, New York.
- CARMEAN, D., AND J. B. CRESPI. 1995. Do long branches attract flies? *Nature* 373:666.
- CARMEAN, D., L. S. KIMSEY, AND M. L. BERBEE. 1992. 18S rDNA sequences and the holometabolous insects. *Mol. Phylogen. Evol.* 2:270–278.
- CASANOVA, J. L., C. PANNETIER, C. JAULIN, AND P. KOURILSKY. 1990. Optimal conditions for directly sequencing double-stranded PCR products with Sequenase. *Nucleic Acids Res.* 18:4028.
- CHALWATZIS, N., J. HAUF, Y. VAN DE PEER, R. KINZELBACH, AND R. K. ZIMMERMANN. 1996. 18S ribosomal-RNA genes of insects—Primary structure of the genes and molecular phylogeny of the Holometabola. *Ann. Entomol. Soc. Am.* 89:788–803.
- COLLESS, D. H., AND D. K. MCALPINE. 1970. Diptera. Pages 656–740 in *The insects of Australia*, 1st edition, Volume 2 (D. F. Waterhouse, ed.). CSIRO, Melbourne Univ. Press, Melbourne.
- COLLESS, D. H., AND D. K. MCALPINE. 1991. Diptera. Pages 717–786 in *The insects of Australia: A textbook for students and research workers*, 2nd edition, Volume 2 (I. D. Naumann, P. B. Carne, J. F. Lawrence, E. S. Nielsen, J. P. Spradberry, R. W. Taylor, M. J. Whitten, and M. J. Littlejohn, eds.). CSIRO, Melbourne Univ. Press, Melbourne.
- COOK, E. F. 1949. The evolution of the head in the larvae of the Diptera. *Microentomology* 14:1–57.
- COURTNEY, G. W. 1990. Cuticular morphology of larval mountain midges (Diptera: Deuterophlebiidae): Implications for the phylogenetic relationships of the Diptera. *Can. J. Zool.* 68:556–578.
- COURTNEY, G. W. 1991. Phylogenetic analysis of the Blephariceromorpha, with special reference to mountain midges (Diptera: Deuterophlebiidae). *Syst. Entomol.* 16:137–172.
- CURTIS, D., J. APFELD, AND R. LEHMANN. 1995. *Nanos* is an evolutionarily conserved organizer of anterior-posterior polarity. *Development* 121:1899–910.
- DE JONG, R. 1980. Some tools for evolutionary and phylogenetic studies. *Z. Zool. Syst. Evolutionsforsch.* 18:1–23.
- DESALLE, R., AND D. A. GRIMALDI. 1991. Morphological and molecular systematics of the Drosophilidae. *Annu. Rev. Ecol. Syst.* 22:447–475.
- DIXON, M. T., AND D. M. HILLIS. 1993. Ribosomal RNA secondary structure: Compensatory mutations and implications for phylogenetic analysis. *Mol. Biol. Evol.* 10:256–267.
- EDWARDS, F. W. 1926. The phylogeny of nematocerous Diptera: A critical review of some recent suggestions. *Verh. Int. Kongr. Entomol.* 3:111–130.
- EVANS, J. W. 1971. Some upper Triassic Hemiptera from Mount Crosby, Queensland. *Mem. Queensl. Mus.* 6:145–152.
- FELSENSTEIN, J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. *Syst. Zool.* 19:401–410.
- FELSENSTEIN, J. 1981. Evolutionary trees from DNA sequences: A maximum likelihood approach. *J. Mol. Evol.* 17:368–376.
- FELSENSTEIN, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39:783–791.
- FELSENSTEIN, J. 1995. PHYLIP: Phylogenetic inference programs, version 3.572. Univ. Washington, Seattle.
- FELSENSTEIN, J., AND G. A. CHURCHILL. 1996. A Hidden Markov Model approach to variation among sites in rate of evolution. *Mol. Biol. Evol.* 13:93–104.
- FRASER, N. C., D. A. GRIMALDI, P. E. OLSEN, AND B. AXSMITH. 1996. A Triassic Lagerstatte from eastern North America. *Nature* 380:615–619.
- FRIEDRICH, M. 1995. Phylogenetischer Informationsgehalt und evolutionäre Eigenschaften von kerncodierten ribosomalen Genen innerhalb der Euarthropoda. Ph.D. Dissertation, Ludwig-Maximilians Univ, Munich.
- FRIEDRICH, M., AND D. TAUTZ. 1997. An episodic change of rDNA nucleotide substitution rate has occurred at the time of the emergence of the insect order Diptera. *Mol. Biol. Evol.* 14:644–653.
- GILLESPIE, J. H. 1986. Rates of molecular evolution. *Annu. Rev. Ecol. Syst.* 17:637–665.
- GOETHGEBUER, M. 1925. Contribution à l'étude des "premandibules" chez les larves des Diptères nématocères. *Encycl. Entomol. B* 2:142–152.
- GOLDMAN, N. 1993. Simple diagnostic statistical tests of models for DNA substitution. *J. Mol. Evol.* 37: 650–661.
- GOUIN, F. J. 1959. Morphology of the larval head of the Chironomidae. *Smithson. Misc. Coll.* 157:175–201.
- GRIFFITHS, G. C. D. 1990. Review of *Manual of Nearctic Diptera, Volume 3*, edited by J. F. McAlpine and D. M. Wood. *Quaest. Entomol.* 26:117–130.
- GRIFFITHS, G. C. D. 1994. Relationships among the major subgroups of the Brachycera (Diptera): A critical review. *Can. Entomol.* 126:861–880.
- GUSTINCICH, S., G. MANFIOLETTI, G. DEL SAR, AND C. SCHNEIDER. 1991. A fast method for high-quality genomic DNA extraction from whole human blood. *BioTechniques* 11:298–301.
- HACKMAN, W., AND R. VÄISÄNEN. 1982. Different classification systems in the Diptera. *Ann. Zool. Fenn.* 19:209–219.
- HACKMAN, W., AND R. VÄISÄNEN. 1985. The evolution and phylogenetic significance of the costal chaetotaxy in the Diptera. *Ann. Zool. Fenn.* 22:169–203.
- HANCOCK, J. M., D. TAUTZ, AND G. DOVER. 1988. Evolution of secondary structures and compensatory mutations of the ribosomal RNAs of *Drosophila melanogaster*. *Mol. Biol. Evol.* 5:393–414.
- HASEGAWA, M., AND T. HASHIMOTO. 1993. Ribosomal RNA trees misleading? *Nature* 361:23.
- HASEGAWA, M., H. KISHINO, AND T. A. YANO. 1985. Dating of the human ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* 22:160–174.
- HENNIG, W. 1948. Die Larvenformen der Dipteren. Eine Übersicht der bisher bekannten Jugendstadien

- der zweiflügeligen Insekten, Volume 1. Akademie Verlag, Berlin.
- HENNIG, W. 1965. Phylogenetic systematics. Illinois Univ. Press, Urbana.
- HENNIG, W. 1968. Kritische Bemerkungen über den Bau der Flügelwurzel bei den Dipteren und die Frage nach der Monophylie der Nematocera. Stuttg. Beitr. Naturk. 193:1–23.
- HENNIG, W. 1969. Die Stammesgeschichte der Insekten. Senckenberg-Buch, Volume 49. Kramer, Frankfurt am Main, Germany.
- HENNIG, W. 1973. Diptera (Zweiflügler). Pages 1–337 in Handbuch der Zoologie. Eine Naturgeschichte des Tierreichs, Volume 4 (J. G. Helmcke, D. Starck, and H. Wermuth, eds.). De Gruyter, Berlin.
- HENNIG, W. 1981. Insect phylogeny. Wiley & Sons, New York.
- HIGGINS, D., AND P. M. SHARP. 1988. CLUSTAL: A package for performing multiple sequence alignments on a microcomputer. *Gene* 73:237–244.
- HUELSENBECK, J. P. 1997. Is the Felsenstein zone a fly trap? *Syst. Biol.* 46:69–74.
- HUELSENBECK, J. P., AND D. M. HILLIS. 1993. Success of phylogenetic methods in the four-taxon case. *Syst. Biol.* 42:247–264.
- HUELSENBECK, J. P., AND B. RANNALA. 1997. Phylogenetic methods come of age: Testing hypotheses in an evolutionary context. *Science* 276:227–232.
- IMMS, A. D. 1944. On the constitution of the maxillae and labium in Mecoptera and Diptera. *J. Microscop. Sci. N.S.* 85:73–96.
- JIN, L., AND M. NEI. 1990. Limitations of the evolutionary parsimony method of phylogenetic analysis. *Mol. Biol. Evol.* 7:607–633.
- KIM, J. 1993. Improving the accuracy of phylogenetic estimation by combining different methods. *Syst. Biol.* 42:331–340.
- KIMURA, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16:111–120.
- KITCHING, I. J. 1992. The determination of character polarity. Pages 22–43 in Cladistics. A practical course in systematics (P. L. Forey, C. J. Humphries, I. J. Kitching, R. W. Scotland, D. J. Siebert, and D. M. Williams, eds.). Clarendon Press, Oxford, England.
- KRISTENSEN, N. P. 1975. The phylogeny of hexapod "orders". A critical review of recent accounts. *Z. Zool. Syst. Evolutionsforsch.* 13:1–44.
- KRZEMIŃSKI, W. 1992a. The oldest Polyneura (Diptera) and their importance to the phylogeny of the group. *Acta Zool. Cracov.* 35:45–52.
- KRZEMIŃSKI, W. 1992b. Triassic and Lower Jurassic stage of Diptera evolution. *Mitt. Schweiz. Entomol. Ges.* 65:39–59.
- LAMEERE, A. 1906. Notes sur la classification des Diptères. *Mem. Soc. Entomol. Belg.* 12:105–140.
- LATREILLE, P. A. 1802. Histoire naturelle, générale et particulière, des Crustaces et des Insectes. Pages 1–468 in Histoire naturelle par Buffon, Volume 3 (C. S. Sonnini, ed.). Buffon, Paris.
- LAWRENCE, P. A. 1992. The making of a fly. The genetics of animal design. Blackwell, London.
- LOCKHART, P. J., M. A. STEEL, M. D. HENDY, AND D. PENNY. 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. *Mol. Biol. Evol.* 11:605–612.
- MACQUART, J. 1834. Histoire naturelle des insectes. Pages 1–578 in Collection des suites à Buffon, Volume 1 (N. E. Roret, ed.). Buffon, Paris.
- MADDISON, W. P., AND D. R. MADDISON. 1992. MacClade, version 3.0. Sinauer, Sunderland, Massachusetts.
- MCALPINE, J. F. 1989. Phylogeny and classification of the Muscomorpha. Pages 1397–1518 in Manual of Nearctic Diptera, Volume 3 (J. F. McAlpine and D. M. Wood, eds.). Agriculture Canada, Ottawa.
- MCPHERON, B. A., AND H. Y. HAN. 1997. Phylogenetic analysis of North American *Rhagoletis* (Diptera: Tephritidae) and related genera using mitochondrial DNA sequence data. *Mol. Phylogenet. Evol.* 7:1–16.
- MELZER, R. R., T. ZIMMERMANN, AND U. SMOLA. 1997. Modification of dispersal patterns of branched photoreceptor axons and the evolution of neural superposition. *Cell. Mol. Life Sci.* 53:242–247.
- MICKOLEIT, G. 1969. Vergleichend-anatomische Untersuchungen an der pterothorakalen Pleurotergal-muskulatur der Neuroptera und Mecoptera (Insecta, Holometabola). *Z. Morphol. Tiere* 64:151–178.
- MILLER, B. R., M. B. CRABTREE, AND H. M. SAVAGE. 1996. Phylogeny of fourteen *Culex* mosquito species, including the *Culex pipiens* complex, inferred from the internal transcribed spacers of ribosomal DNA. *Insect Mol. Biol.* 5:93–107.
- OLSEN, P. E., C. L. REMINGTON, B. CORNET, AND K. S. THOMSON. 1978. Cyclic change in Late Triassic lacustrine communities. *Science* 201:729–733.
- OOSTERBROEK, P. F. L. S., AND G. COURTNEY. 1995. Phylogeny of nematocerous families of Diptera (Insecta). *Zool. J. Linn. Soc.* 115:267–311.
- OOSTERBROEK, P. F. L. S., AND B. THEOWALD. 1991. Phylogeny of the Tipuloidea based on characters of the larvae and pupae (Diptera, Nematocera) with an index to the literature except Tipulidae. *Tijdschr. Entomol.* 134:211–267.
- PÄÄBO, S. 1990. Amplifying ancient DNA. Pages 159–166 in PCR protocols. A guide to methods and applications (A. M. Innis, D. H. Gelfand, J. J. Sninsky, and T. J. White, eds.). Academic Press, San Diego, California.
- PASHLEY, D. P., B. A. MCPHERON, AND E. A. ZIMMER. 1993. Systematics of holometabolous insect orders based on 18S ribosomal RNA. *Mol. Phylogenet. Evol.* 2:132–142.
- PAWLowski, J., R. SZADZIEWSKI, D. KMIECIAK, J. FAHRNI, AND G. BITTAR. 1996. Phylogeny of the infraorder Culicomorpha (Diptera: Nematocera) based on 28 RNA sequences. *Syst. Entomol.* 21:167–178.
- PÉLANDAKIS, M., AND M. SOLIGNAC. 1993. Molecular phylogeny of *Drosophila* based on ribosomal RNA sequences. *J. Mol. Evol.* 37:525–543.
- RAICH, T. J., J. L. ARCHER, M. A. ROBERTSON, W. J. TABACHICK, AND B. J. BEATY. 1993. Polymerase chain reaction approaches to Culicoides (Diptera: Cer-

- topogonidae) identification. *J. Med. Entomol.* 30: 228–232.
- ROHDENDORF, B. B. 1964. The historical development of the Diptera. Univ. Alberta Press, Edmonton.
- ROUSSET, F., M. PÉLANDAKIS, AND M. SOLIGNAC. 1991. Evolution of compensatory substitutions through G-U intermediate state in *Drosophila* rRNA. *Proc. Natl. Acad. Sci. USA* 88:10032–10036.
- RZHETSKY, A., AND M. NEI. 1995. Tests of applicability of several substitution models for DNA sequence data. *Mol. Biol. Evol.* 12:131–151.
- SAITOU, N., AND M. NEI. 1987. The neighbor-joining method: A method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4:406–425.
- SANDER, K. 1996. Variants of embryonic patterning mechanisms in insects—Hymenoptera and Diptera. *Semin. Cell Dev. Biol.* 7:573–582.
- SCHMIDT-OTT, U., G. M. TECHNAU, AND K. SANDER. 1994. Expression of *engrailed* in embryos of a beetle and five dipteran species with special reference to the terminal regions. *Arch. Dev. Biol.* 203:298–303.
- SCHNARE, M. N., S. H. DAMBERGER, M. W. GRAY, AND R. R. GUTELL. 1996. Comprehensive comparison of structural characteristics in eukaryotic cytoplasmatic large subunit (23S-like) ribosomal RNA. *J. Mol. Biol.* 256:701–719.
- SCHREMMER, F. 1951. Die Mundteile der Brachycer-enlarven und der Kopfbau der Larven von *Stratiomyces chamaeleon* L. *Österr. Zool. Z.* 3:326–397.
- SCHUMANN, H. 1992. Systematische Gliederung der Ordnung Diptera mit besonderer Berücksichtigung der in Deutschland vorkommenden Familien. *Dtsch. Entomol. Z. N.F.* 39:103–116.
- SHAW, S. R., AND I. A. MEINERTZHAGEN. 1986. Evolutionary progression at synaptic connections made by identified homologous neurones. *Proc. Natl. Acad. Sci. USA* 83:7961–7965.
- SINCLAIR, B. J. 1992. A phylogenetic interpretation of the Brachycera (Diptera) based on the larval mandible and associated mouthpart structures. *Syst. Entomol.* 17:233–252.
- SINCLAIR, B. J., J. M. CUMMING, AND D. M. WOOD. 1994. Homology and phylogenetic implications of male genitalia in Diptera—Lower Brachycera. *Entomol. Scan.* 24:407–432.
- SMITH, J. J., AND G. L. BUSH. 1997. Phylogeny of the genus *Rhagoletis* (Diptera: Tephritidae) inferred from DNA sequences of mitochondrial cytochrome oxidase II. *Mol. Phylogen. Evol.* 7:33–43.
- SNODGRASS, R. E. 1935. Principles of insect morphology. McGraw-Hill, New York.
- SULLIVAN, J., K. E. HOLSINGER, AND C. SIMON. 1996. The effect of topology on estimates of among-site rate variation. *J. Mol. Evol.* 42:308–312.
- SWOFFORD, D. L. 1993. PAUP: Phylogenetic analysis using parsimony, version 3.0r. Illinois Natural History Survey, Champaign.
- TANG, J., K. PRUESS, E. W. CUPP, AND T. R. UNNASCH. 1996. Molecular phylogeny and typing of blackflies (Diptera: Simuliidae) that serve as vectors of human or bovine onchocerciasis. *Med. Vet. Entomol.* 10: 228–234.
- TAUTZ, D., J. M. HANCOCK, D. A. WEBB, C. TAUTZ, AND G. A. DOVER. 1988. Complete sequences of the rRNA genes of *Drosophila melanogaster*. *Mol. Biol. Evol.* 5:366–376.
- TESKEY, H. J. 1981. Morphology and terminology—Larvae. Pages 65–88 in *Manual of Nearctic Diptera*, Volume 1 (J. F. McAlpine, B. V. Peterson, G. E. Shewell, H. J. Teskey, J. R. Vockeroth, and D. M. Wood, eds.). Agriculture Canada, Ottawa.
- TILLYARD, R. J. 1929. Permian Diptera from Warner's Bay, N.S.W. *Nature* 123:778–779.
- ULRICH, H. 1991. The present state of comparative morphology of thoracic musculature in Diptera. *Int. Congr. Diptera* 2:315–326.
- UZZELL, T., AND K. W. CORBIN. 1971. Fitting discrete probability distributions to evolutionary events. *Science* 172:1089–1096.
- VOSSBRINCK, C. R., AND S. FRIEDMAN. 1989. A 28S ribosomal RNA phylogeny of certain cyclorrhaphous Diptera based upon a hypervariable region. *Syst. Entomol.* 14:417–431.
- WHITE, M. J. D. 1949. Cytological evidence on the phylogeny and classification of the Diptera. *Evolution* 3:252–260.
- WHITING, M. F., J. C. CARPENTER, Q. D. WHEELER, AND W. C. WHEELER. 1997. The Strepsiptera problem: Phylogeny of the holometabolous insect orders inferred from 18S and 28S ribosomal DNA sequences and morphology. *Syst. Biol.* 46:1–68.
- WILLMANN, R. 1989. Rediscovered: *Permotipula patricia*, the oldest known fly. *Naturwissenschaften* 76: 375–377.
- WOOD, D. M., AND A. BORKENT. 1989. Phylogeny and classification of the Nematocera. Pages 1333–1370 in *Manual of Nearctic Diptera*, Volume 3 (J. F. McAlpine and D. M. Wood, eds.). Agriculture Canada, Ottawa.
- WOODLEY, N. E. 1989. Phylogeny and classification of the "orthorrhaphous" Brachycera. Pages 371–1395 in *Manual of Nearctic Diptera*, Volume 3 (J. F. McAlpine and D. M. Wood, eds.). Agricultural Canada, Ottawa.
- WOOTTON, R. J., AND A. R. ENNOS. 1989. The implications of function on the origin and homologies of the dipterous wing. *Syst. Entomol.* 14:507–520.
- WU, C.-I., AND W.-H. LI. 1985. Evidence for higher rates of nucleotide substitution in rodents than in man. *Proc. Natl. Acad. Sci. USA* 82:1741–1745.
- XIONG, B., AND T. D. KOCHER. 1991. Comparison of mitochondrial DNA sequences of seven morphospecies of black flies (Diptera: Simuliidae). *Genome* 34:306–311.
- YANG, Z. 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Approximate methods. *J. Mol. Evol.* 39: 306–314.
- YANG, Z. 1996a. Among-site rate variation and its impact on phylogenetic analysis. *Trends Evol. Ecol.* 11: 367–372.
- YANG, Z. 1996b. PAML: Phylogenetic analysis by maximum likelihood, version 1.2. Department of Integrative Biology, Univ. California, Berkeley.

Received 30 January 1995; accepted 19 July 1997

Associate Editor: Chris Simon

APPENDIX 1
ALIGNMENT OF 28S D1 AND 28S D7 RDNA SEQUENCES

Taxon names are abbreviated to the first three letters of the genus name. Sites extracted for phylogenetic analysis are marked with a star above the alignment. DS regions are indicated by lines above the alignment. Dots indicate positions identical to that in the heading sequence. Dash = gap; N = ambiguous site. Regions of major disagreement between the final alignments and CLUSTAL V alignments are situated at positions 23–25, 39–69, 75–120, 143–150, and 176–200 of the final 28S D1 alignment and at positions 39–104, 124–130, 177–179, 188–190, 226–230, 253, 263, 301–303, and 312–315 of the final 28S D7 alignment. The alignments can be retrieved from the EBI server under accession numbers DS25432 (D1) and DS25433 (D7).

D1	1	85
Dro	AGTAGCGCGGAGCGAAAGAAAAA-CAGTTCACTAAGTCAGTCACTTGTC	--TATAT--GGCAAATGTGAGATGCAGTGTAT-GG
TabA.....T.T.....TA--T.T.....TAA.....A.A.....A.....
Bra	CAGG.T-A..CC.....CTCTCCG.....GTCC.....A.TGG.AGAG.....
Cri	CAGG.T-T..CC.....TCCAA.T--T.....ATTGG.A.....
Bol	CAGG..A..CC.....C.TCTCA.....GTCC.....TGG.AG.C.....
Ana	T.GG.T-A..C.....C.TTC.G..T--C.....A.C--A.T.G.TA.G.....
Dil	C.GG.T-A.....CTC.AA----TA.....TT.G--AG.....T.....
Syl	C.GG..A..CC.....TCTC.AA..T--GT.G..G..TT.G.AGA.....
Psy	C.GG..A..CC.....T.C.CA..G.CT--CATA.AC--C.....TG.G.....
Tip	N..T.GG..A..C.....T.....CT.....ACTG--AT.....CTC..T.G.ATG..A..T..A..AA.....
Lim	C.GG.T-A..CC.....T.A.....CTC..TC--TT.TA..C.GA.ATTG.TG..A..T..T.....CAAT.....
Tri	C.GG.T-T..CC.....C.....CTC.A.C.T--TA..C.G..TT.G.AG.....G.....A.....
Aed	T.....C.GG..G..C.....GC.....GGG.GA..CA--CTGG--T.T.TC..CCC.G.T.C.....CT.....
Cul	T.....N.....GG..G..C.....GC.....GGGTGGC.C--ACTACCCC--.GTC.ACCC.G.T.C.....CT.....
Chi	T.....A.....C.GG.T--..CC.....T.....GT.....GAT.A.A.G.....TA..T..T.T.....ATC.....T.C.....AA.....
SimGG..A..C.....GT.....AAT.GCAACT--GT.....TTGT.....ATT.....C.....CA.....
PanC.GG..G..CC.....G.A..C.A.C.G--AT..T--C..AT.G..G..A..T.....A.....
ArcN..C..GG..G..CC.....A..C.GCG.C.G--G.G..--C..ACCG..G..A..T.....TAG.....
ManC..G..TAT..CC.....G.A..C.GC..TG--A--C..GC.G.....TG.....TCG.....
D1	86	170
Dro	AGCGTCAATTCTAGTATGAGAAATTACG--ATTAACTCCTCTAAATGAGGC-CA-TTTACCCATAGAGGGTGCAGGC	-----
TabATA..T..TA.....AT.....TA--G..C.....G.....-.....-.....AT.....
BraAA..T..TA.....TA.GCA.CG.TAAT.T--G..C.....A..G.....G.....-.....T.AA.....
CriAA..T..TA.....G..G.TGT..TTCA.TAC--G..C.....A..G.....G.....-.....A.A..T..C.....
Bol	GAA..T..TA.....TA.....CGGTAGC--G..C.....A..G.....G.....-.....T..A.....
AnaAAA..T..TA.....ATPAT.CAGC.TAT--G..C.....A..G.....G.....-.....A.....
DilAA..T..TACT.TA.....A..AG..TGAT--G..C.....A..G.....G.....-.....T..A..GC.....
SylA..T.CTA..CA.....TT.CGGC..TAC--G..C.....A..G.....G.....C.....
PsyT..C..TAA..CA..TAC..AT..GA..G..T..TAGC..C.....A..G.....G.....-.....T.....
TipTA..T..TA..CTAA..TTT--TAT--G..CC.....A..G.....G.....-.....C.....
LimT..AA..TC..TA.....CAT..TGGT--GC.....T..A..G.....G.....-.....C.....
TriTAAA..T..TA.....TA.....TTCC..A..C--G..C.....A..G.....G.....-.....T..A.....
AedA..TCG..TA.....GCCG..CCGG..CGCT--G..C.....T..AA..AG.....T..-TTT.ACT.....G.....AT.....
CulAAT..TTG..CA.....GCC.....CCCGGGCG..GT--G..CC.....T..AA..AG.....T..-TTT.ACT.....G.....AT.....
ChiG..C..TA.....T..C..CATAT..AC..T..AT--A.....AG..G..C.....T..A..A.....C.....
SimTT..G..C..TA.....CA..T..T..T..T..AGTAT--G..C.....G.....-.....A..A..C.....AT.....
PanGA..TC..T..ATCCCG..GATC..TG..AGGCCG--G..C.....A..G.....G.....-.....A.....
ArcGA..C..AT..ATCTCGG..C..TCGGGGCGA--G..C.....A..G.....G.....-.....A.....
ManGT..CG-C..TCTCGTACCG..TACTC..T--G..CC.....T..G.....G..C..G.....G..T..GA.....
D1	171	255
Dro	CCGTATAACGTTAACGAT--TACTA-GATGATGT--TTCCTAAAGAGTCGTGCTTGTAGTAGTCAGCACTAACGTTGGGTG	-----
TabG..ACA..T..T..AA--A--A..A..A..TT.....T.T.....
BraGC..GC..ATTATCAG--TA..A..A.GCT.....T.....
CriG..GC..ATTGTT..TG--C..TC--A..GCT.....
BolGC..AC..GTTACCGA--TA..A..A..T.....G.....T.....
AnaGT..AC..ATTGTT..TA--A..G..TC--A.....A..T.....
DilGT..AC..ATTATCAA--TA..A..A..GCTC.....
SylGTGAC..GTTGTCGGTTGC..TG--A..A..CT--C..TT.....T.....
PsyGG..TACC..CTGT--A..T..T..TG..CTA--C..TT.....
TipGTGAC..AT..ATC..TG--TTAG--A..A..A..C..T..G.....T.....
LimGC..ACTATTGTA--T..A..G..A..T--G..ATC.....
TriGGT..AC..ATTGTTAT--TA..A..A..T--G..TT.....G.....
AedG..G..GGCGACCGA--GG..A..C..TG.....
CulGG..GGCG--GG..G..AA..T--TG.....
ChiTGTG..A..C..TG--GTA..TG..GCTATCAAC..T..G..A.....T.....
SimTGCT--A..A..TG..A..A..C..T..G.....A.....
PanA..AC..GTTACTGAT--CTTGG..G..C..C..TTT..G.....G.....T.....
ArcGCGACCGTTGCGCGT--AT..G..G..C..C..TC.....A.....G.....T.....
ManGCGACCG..GG..CGGT--GG..G..G..GAC..C..TC.....G.....G.....C.....

D1 256 284

Dro	GTA AACTCCATC TAAA ACTAA ATAAC		
Tab	.	.	.
Bra	.	GG	G
Cri	.	GG	
Bol	.	GG	
Ana	G	G	T
Dil	.	.	
Syl	.	GG	C
Psy	.	GG	T
Tip	.	G	G
Lim	.	GG	
Tri	T	GG	NN
Aed	T	G	C
Cul	T	G	C
Chi	T	GG	T
Sim	T	G	T
Pan	.	GG	
Arc	.	GG	G
Man	.	GG	T

D7 1 85

Dro	GAACGAAAGGGAA	TACGGTCCAA	TTCCGTAA	CTTTGTTATTAAATA	-----	TGGGCCTCGT
Tab	.	.	-	A	-----	A-----
Bra	A.G.	.	-	A	TA.T	A.A-----
Cri	.GG.	.	-	A	TA.T	A.G-----
Bol	.G.	.	-	A	TA.T	A.A-----
Ana	.G.	.	T	A	GA	A-----
Dil	A.G.	.	-	A	TA.T	A.A-----
Syl	.G.	.	-	A	TAAT	A.TA-----
Psy	.TG.	.	A	-----	A.G	-----
Tip	.G.	.	-	G.A.C	T.A.GAT	-----
Lim	.C.	.	-	A.C	AAA.GAT	ATA.TAACATTTAA
Tri	.G.	.	-	A	-----	TTCTT.TAATTACT
Aed	.G.	C.	A	G.G	G.A	CA-----
Cul	.G.	C.	A	G.G	G.A	CA-----
Chi	.G.	.	N	-----	A	-----
Sim	.G.	C.	A	G	-----	A
Pan	.GG.	C.	T	G	C.GCA.C.N.A	CAT.ATGTG.CCGCC
Arc	NNNNN	C.	T	G	C.GCA.C.G.A	ACATT.ATGTGGTTAT
Man	.G.	C.	T	G	C.GCA.C.G.A	AACA.G-----

D7 86 170

Dro	-GCTCATCCTGGCACAGGAACG--	-ACCATAAAAGAAGCCGTGAGAGATATCGGAAGAGTTTCTT
Tab	-.....G.....CA.....A.....T.....A.....G.....	
Bra	-A.....	-.....G.....T.....
Cri	-A.....	-.....A.....G.....T.....
Bol	-A.....T.....A.....	-.....G.....T.....
Ana	-.....A.....T.....A.....A.....G.....T.....A.....	CT.....
Dil	-.....A.....T.....A.....	-.....G.....T.....
Syl	-.....A.....	-.....A.....G.....T.....G.....CT.....
Psy	TGCAAGTGTTCGGTTTA.....	A.G.T--GG.....G.....AGCG.....C.....A.....
Tip	GATGANT--.....AT.....T.....	T--G.....G.....C.....G.....T.....C.....A.....
Lim	GATGAAT--.....AT.....T.....	A--.....G.....C.....T.....G.....T.....A.....
Tri	NTAATTAAAGGT--.....A.....T.....	A--.....G.....
Aed	AACCGTAGCGCCTT--.....TG.....A.....A.....T.....	TCCTTTT.....T.....CG.....AA.....A.....
Cul	ACACGGTAGCGCCTT--.....TG.....A.....A.....T.....	TCCTTT.....T.....CG.....AA.....A.....
Chi	CGAA--.....TAA.....TC.....T.....GA.....	GCT--CT.....G.....C.....T.....AAT.....G.....G.....A.....
Sim	CCATT--.....ATGT.....TC.....GA.....	GCCC--AT.....TA.....G.....TG.....A.....AA.....
Pan	-.....T.....G.....GG.....T.....CCA.....AA--GA.....C.....GG.....C.....	G.....CCA.....A.....
Arc	-.....T.....G.....GG.....T.....CCA.....AT--GA.....C.....GG.....C.....	G.....CG.....G.....
Man	CGAGT--.....T.....GA.....GG.....T.....CCA.....GT--GG.....C.....G.....C.....	G.....CCG.....

D7	171	255

Dro	TCTGTTT-TATAGCCGTA-CTACCATGGAAAGTCCTTCGCAGAGAGATATGGTAGATG-GGCTAGAAGAGCATGACATATACT-GT	
TabTT...-T.....A.....A...-AA.....AG.....-C	
Bra	.C.A-C.A.C.AA.....-C.....TAT.....G.....-CT.G.....T.....	
CriG-A.C.A.T.....-TA.....A.....TA.....-AT.G.....A.....T.....	
BolG-A.C.A.T.....-TC.....A.....GA.....-AT.G.....T.....	
AnaG-A.C.ATT.....-TC.....TA.....GA.....-AAT.G.....T.....	
DilG-A.C.A.T.....-TC.....AT.....GA.....-AT.G.....T.....	
SylGTA.C.A.T.....-TA.....AAA.....TA.....-A.T.G.....T.....T.....	
PsyT.....-T.....A.....A.....-A.....G.C.....T.....-C	
Tip-C.....A.A.....T.....A.AA.....T.....T.....	
LimAA.C.....-TA.....T.....TA.....A.....T.....	
Tri-T.....C.....-TC.....C.....A.....GA.....T.....T.....-C	
AedC.....T.AC.....CTAG.....AT.....T.GACA.....G.T.....GT.....TA.A.T.C	
CulC.....AC.....CTGG.....AT.....C.....GACAA.....G.T.....GT.....TA.A.T.C	
ChiC.....AT.....CG.....AT.....T.....ACA.....AT.G.T.....TA.A.T.	
Sim-C.....TT.....CGC.....AT.....GT.....ATA.....G.T.....G.....TC.G.T.C	
PanA.-G.GCATTTCG.-G.T.....C.....TC.....A.....G.....TTGG--AATGC.....C-G.....GT.G.G-C	
ArcA.-G.GC.....TTCG.-G.T.....C.....TC.....C.A.....G.....G.....TTGG--AA.GC.....C-G.....GT.G.G-C	
ManCC.....-G.GC.....TTCG.-G.T.....C.....A.....A.....G.....TCGG--AA.GC.....C-G.....T.G.G-C	
D7	256	340

Dro	TGTGTCG-ATATTTCTCCTCGAACCTTGAAAATTATGGTGGGG--ACACGCCAAC--TTCTCAACAGGCCGTACCAAATATCCG	
TabG.....A.....-TT.....-T	
BraA-G.....T.....C.....-C.....-	
CriA.....T.....C.....A.....-C.....N.....T.....A.....G.....	
BolA-G.....T.....C.....A.....-C.....-A.....A.....G.....	
AnaA-G.....A.....T.....C.....A.....-C.....-A.....A.....G.....	
DilA-G.....A.....T.....C.....-C.....-A.....N.....A.....G.....G	
SylA-G.....C.....T.....C.....-C.....-A.....A.....G.....	
PsyCG.....CG.....CT.....C.....CCA.....-C.....T.....A.....T.....A	
Tip-G.....C.....C.....T.....C.....AA.....-TT.....T.....-A.....G	
Lim-C.....C.....A.....C.....A.....-CT.....T.....-A.....G	
Tri-C.....A.....T.....C.....-C.....T.....-N.....A.....G.....	
Aed-C.....T.....T.....CG.....A.....ACT.....-G.....-C.....CTT.....G-A	
Cul-C.....T.....T.....CG.....A.....ACT.....-G.....-C.....CTT.....G-A	
Chi-G.....AC.....A.....T.....C.....AAA.....-AT.....A.....T.....C.....T.....CGA	
SimT.....T.....CAA.....TA.....AT.....-T.....-A.....G-A	
PanG.....T.GG.....C.....C.....CC.....G.....AGA.....-C.....TGG.....GGTG.....G.GC.G.TT.....C	
ArcG.....C-G.....C.....C.....CC.....G.....AGA.....-C.....-GG.....GGTG.....G.GC.G.TT.....C	
ManG.....C-GG.....AC.....TG.....C.....G.....TGA.....ATGT.....TGG.....AATG.....G.GC.G.TT.....C.....N	
D7	341	353

Dro	CAGCTGGTCTCCA	
TabN.....	
BraC.....	
CriC.....	
BolC.....	
AnaC.....NN	
DilC.....NN	
SylC.....	
PsyA.....	
Tip	
Lim	
Tri	
AedA.....	
CulA.....	
ChiA.....	
SimA.....	
PanA.....N	
ArcA.....	
ManA.....	

APPENDIX 2. Percentage base composition of dipteran and outgroup 28S D1 and D7 rDNA sequences in single-stranded (SS) and double-stranded (DS) sites.

Taxon	SS				DS			
	A	T	G	C	A	T	G	C
<i>Drosophila</i>	39.3	25.5	21.9	13.3	22.6	27.4	26.3	23.7
<i>Tabanus</i>	39.8	28.1	21.4	10.7	24.4	29.6	24.1	21.9
<i>Bradyzia</i>	39.8	24.0	23.0	13.3	21.5	26.7	26.7	25.2
<i>Crinodiplosis</i>	39.8	24.0	22.4	13.8	23.0	27.4	25.6	24.1
<i>Bolitophilida</i>	39.3	24.5	23.0	13.3	21.5	27.4	26.7	24.4
<i>Anapausis</i>	39.8	24.5	22.4	13.3	23.0	28.5	25.9	22.6
<i>Dilophus</i>	38.8	24.5	23.0	13.8	22.6	27.4	26.3	23.7
<i>Sylvicola</i>	39.3	25.0	23.0	12.8	20.7	27.8	27.0	24.4
<i>Psychoda</i>	38.3	26.5	23.0	12.2	20.0	25.2	27.8	27.0
<i>Tipula</i>	37.8	26.0	22.4	13.8	22.2	27.8	25.9	24.1
<i>Limonia</i>	36.7	26.0	21.9	15.3	23.0	28.9	25.6	22.6
<i>Trichocera</i>	37.8	26.5	23.5	12.2	20.7	26.3	27.8	25.2
<i>Aedes</i>	39.3	24.5	23.5	12.8	19.3	27.0	30.0	23.7
<i>Culex</i>	39.3	25.0	23.0	12.8	19.6	26.7	29.6	24.1
<i>Chironomus</i>	39.8	26.5	21.4	12.2	22.2	28.5	27.0	22.2
<i>Simulium</i>	39.3	25.5	24.0	11.2	22.2	28.5	26.7	22.6
<i>Panorpa</i>	34.7	26.5	25.0	13.8	15.6	20.7	34.4	29.3
<i>Archaeopsylla</i>	34.7	25.5	25.0	14.8	15.6	19.3	34.8	30.4
<i>Manduca</i>	32.7	25.5	27.0	14.8	14.4	19.3	35.9	30.4