Phylogenetic Signal and its Decay in Mitochondrial SSU and LSU rRNA Gene Fragments of Anisoptera

B. Misof, A. M. Rickert,* T. R. Buckley,†[1] G. Fleck,‡ and K. P. Sauer**

*Institute for Evolutionary Biology and Ecology, University of Bonn, Bonn, Germany; †School of Biological Sciences, Victoria University of Wellington, Wellington, New Zealand; and ‡Laboratoire d'Entomologie, Museum National d'Histoire Naturelle, Paris, France

The phylogeny of Anisoptera, dragonflies in the strict sense, has proven to be notoriously difficult to resolve. Based on morphological characters, several recent publications dealing with the phylogeny of dragonflies proposed contradicting inter- and intrafamily relationships. We explored phylogenetic information content of mitochondrial large-subunit (LSU) and small-subunit (SSU) ribosomal gene fragments for these systematic problems. Starting at published universal primers, we developed primer sets suitable for amplifying large parts of the LSU and SSU rRNA genes within dragonflies. These fragments turned out to harbor sufficient phylogenetic information to satisfyingly resolve intrafamily relationships, but they contain insufficient phylogenetic structure to permit reliable conclusions about several interfamily relationships. We demonstrate that decay of phylogenetic signal progresses from intrafamily to interfamily to outgroup relationships and is correlated with an increase of genetic distances. As expected, signal decay is most pronounced in fast-changing sites. Additionally, base composition among fast-changing sites significantly deviates from the expected homogeneity. Homogeneity of base composition among all included taxa was restored only after removing fast-changing sites from the data set. The molecular data tentatively support interfamily relationships proposed by the most recent publication based on morphological characters of fossil and extant dragonflies.

Introduction

In this investigation, we aimed to explore phylogenetic signal of mitochondrial 12S (small-subunit; SSU) and 16S (large-subunit; LSU) rRNA gene fragments in order to unravel phylogenetic relationships within dragonflies. Mitochondrial LSU and SSU gene fragments are routinely used in phylogenetic work. Universal primers do exist which offer the possibility of amplifying gene fragments from nearly every organism of interest (for examples, see Simon et al. 1994). In contrast, dragonflies are novel organisms in molecular systematics. Except for three recent publications on a phylogeny of the genus Ischnura (Coenagrionidae, Damselflies) (Chippindale et al. 1999), a phylogeny of the genus Calopteryx (Calopterygidae, Damselflies) (Misof, Anderson, and Hadrys 2000), and a phylogeny of the genus Libellula s.l. (Kambhampati and Charleton 1999), molecular phylogenetic work on odonates has not yet been published. This is the first attempt to establish the utility of SSU and LSU fragments in resolving the numerous open systematic questions within Anisoptera.

Depending on the author's preferences, dragonflies, Anisoptera, are subdivided into 6–15 families. Three probably monophyletic groups (based on morphological data) are species-rich, the families Gomphidae and Aeshnidae and the ''Corduliidae'' + Libellulidae complex, accounting for almost 90% of all species within dragonflies. The remaining families, Petaluridae, Austropetaliidae, Cordulegastridae, Neopetaliidae, Chlorogomphidae, Gomphomacromiidae, Macromiidae, and Synthemistidae, to name the most prominent ones, are species-poor assemblages, often restricted to small geographical areas or refugia. Imagoes of all dragonflies display a restricted ecological diversity; all dragonflies are essentially fast flying insect predators, whereas larval ecology shows a pronounced degree of diversification, from (semi)terrestrial life (e.g., Petaluridae) to life in high-current water bodies (e.g., Austropetaliidae, most Gomphidae).

Recent morphological work on a phylogeny of dragonflies (Pfau 1991; Carle 1995; Bechly 1996; Lohmann 1996; Trueman 1996; Bechly et al. 1998) left many interfamily relationships unresolved. The analyses of Pfau (1991), Carle (1995), Fleck (1996), Lohmann (1996), Trueman (1996), and Bechly et al. (1998) contradict each other in phylogenetic positions of Petaluridae, Gomphidae, Aeshnidae, Austropetaliidae, and Cordulegastridae. All of those analyses are based on morphological characters, some of them exclusively on wing characters (Trueman 1996); others are based on more exhaustive analyses of secondary sexual characters (Pfau 1991) and head characters (Fleck 1996), and others include information on fossil and extant species (Carle 1995; Lohmann 1996; Bechly et al. 1998). Even within families, most phylogenetic relationships remain obscure (e.g., within Gomphids [Carle 1986] or Libellulidae [Kambhampati and Charleton 1999]).

The phylogenetic analysis of mitochondrial SSU and LSU gene fragments therefore has the potential to assess resolution at the intrafamily and interfamily levels. Several questions have to be addressed: (1) Does the molecular analysis recover generally accepted monophyletic clades based on morphological arguments, thus supporting the interpretation of specific morphological characters? (2) What is the position of the aeshnids, pe-

---

[1] Present address: Zoology Department, North Carolina Central University.

Key words: Anisoptera, phylogeny, mitochondrial SSU and LSU rRNA, signal decay, base frequency.

Address for correspondence and reprints: B. Misof, Institute for Evolutionary Biology and Ecology, University of Bonn, An der Immenburg 1, 53121 Bonn, Germany.
E-mail: bmisof@evolution.uni-bonn.de.

Mol. Biol. Evol. 18(1):27–37. 2001

Table 1
Primers Used to Amplify Mitochondrial SSU and LSU rRNA Fragments

Name	Sequence (5′–3′)	Source
LR-J-12887	GGA GCT CCG GTT TGA ACT CAG ATC	Xiong and Kocher 1991
LR-N-13398	CGG CCG CCT GTT TAT CAA AAA CAT	Xiong and Kocher 1991
LR-J-new	GCT CCG GTT TGA ACT CAG AT	Misof, this paper
LR-N-new	AGT TCT CGC CTG TTT ATC AAA	Misof, this paper
LR-J-O2	TTA TGA CTA ATG ATT ATG CTA C	Rickert, this paper
LR-J-O3	GGT ATC TAG TTT TTT AAG AA	Rickert, this paper
LR-J-O4	TAG CTC TTC TGA AAT CGA GA	Misof, this paper
LR-N-O4rev	TCT CGA TTT CAA AAG AGC TA	Misof, this paper
SR-N-O1	GAT CTG ATG AAG GTG GAT TT	Misof, this paper
SR-N-O2	AGA TTT TGG CTC TAA AAT ATG	Rickert, this paper
SR-J-14233	AAG AGC GAC GGG CGA TGT GT	Simon et al. 1994
SR-J-14508	TAC AAA ACA GAT TCC TCT G	Simon et al. 1994
SR-N-14588	AAA CTA GGA TTA GAT ACC CTA TTA T	Simon et al. 1994
SR-N-14756	GAC AAA ATT CGT GCC AGC	Simon et al. 1994

talurids, and gomphids? All three major evolutionary lineages are characterized by a complex pattern of primitive and autapomorphic features. For example, wing structures appear to be plesiomorphic in all three groups, whereas larval ecology is specialized in petalurids to semiterrestrial life, to active freshwater predators in aeshnids and buried sit-and-wait predators in gomphids. (3) Do ‘‘Cavilabiata’’ (monophyly of Cordulegastridae, Neopetaliidae, Chlorogomphidae, Synthemistidae, Macromiidae, Corduliidae, and Libellulidae based, e.g., on the shape of the labial mask and gizzard structures) really exist, or is this a polyphyletic group? ‘‘Cavilabiata’’ are characterized by a specific larval type, a mostly semiburied sit-and-wait predator with a spoonlike labial mask which enables larvae to perform an upward stroke into the water body to rescue prey. Several groups within the assumed monophyletic ‘‘Cavilabiata’’ developed active predatory larvae, still showing the spoon-shaped labium. These larval characters are considered derived diagnostic features, which led to the inclusion of the Neopetaliidae, formerly within Aeshnoids, within ‘‘Cavilabiata’’ (Carle and Louton 1994).

Phylogenetic information in gene fragments is built up by nucleotide substitutions over time. Due to the limited number of character states, phylogenetic signal also erodes by similar processes, and its progression is dependent on rate of change. In this study, we analyzed the quantity of phylogenetic signal and its decay depending on cladistic level, genetic distance, and rate of change.

Materials and Methods

Taxon Sampling

We sampled taxa from nearly all major lineages of dragonflies. We were unable to include representatives of the family Synthemistidae, exclusively distributed in the Australian region. We further included four representatives of the Zygoptera, one representative of the Epiophlebiidae, and one representative of the Ephemeroptera (mayflies) of the family Ephemeridae as outgroup taxa. Sequences have been deposited at GenBank under accession numbers AF266042–AF266103.

DNA Extraction, PCR Amplification, and DNA Sequencing

DNA samples were prepared from individual insects by extraction of total DNA from frozen or ethanol-preserved animals. After excision of muscle tissue, voucher specimens used for DNA extractions where deposited in absolute ethanol at −20°C at the University of Bonn, Institute for Evolutionary Biology and Ecology, and at the Museum National d’Histoire Naturelle, Paris.

Total DNA was extracted with two different methods, of which the classical phenol/chloroform extraction method was preferably used on most frozen specimens (Hadrys, Balick, and Schierwater 1992). The CHELEX extraction method (Gerken et al. 1998) was used for most ethanol-preserved specimens for which the phenol/chloroform extraction did not yield template DNA suitable for PCR amplification.

PCR amplifications were performed on a Perkin Elmer 9600 (Perkin Elmer) and a Primus 96 (MWG) according to Misof, Anderson, and Hadrys (2000). Amplification of the 12S fragments was only successful using a touch-down PCR set up. The chemical composition of the PCR reaction was similar to the regular reaction conditions (Misof, Anderson, and Hadrys 2000), except that the 1–2.5 U of Taq polymerase and 50% of the sterile water were added after 2 min denaturation at 92°C. Within the first 15 cycling steps, the annealing temperature was decreased by 1°C each cycle, starting at 50°C and ending at 35°C. These initial 15 cycles were followed by 25 cycles of constant annealing temperature at 50°C and 1 min extension time at 72°C to enrich the amplification product. Primers used and developed in this investigation are listed in table 1.

Prior to sequencing, primers and unincorporated nucleotides were removed by passing PCR products through Sephadex columns (Qiagen) following the manufacturer’s protocol. Purified PCR products were sequenced directly using the ABI Big Dye Terminator Cycle sequencing kit (Perkin Elmer) and run on an Applied Biosystems 377 automated sequencer. Both strands were sequenced for every PCR product. PCR primers were used for directly sequencing the purified PCR products.

Analyses of Phylogenetic Information Content and Substitutional Saturation

Sequences were initially aligned with the CLUSTAL X program package (Higgins and Sharp 1988) and subsequently corrected by eye. We explored the quality of our alignment by varying alignment parameter specifications (gap opening cost, gap extension cost) and comparing alignments. Positions which showed variable alignments where excluded from phylogenetic analyses (Gatesy, DeSalle, and Wheeler 1993; Milinkovitch, Berube, and Palsboll 1998). Additionally, we used secondary-structure information from the 16S Van de Peer model (de Rijk, Van De Peer, and de Wachter 1997; Gutell et al., personal communication), secondary-structure information for odonates and other insects (Buckley et al. 2000), and secondary-structure information on 12S domain III (Hickson et al. 1996) to confirm identification of ambiguous alignment positions.

Prior to performing phylogenetic reconstructions, we analyzed the quality of our data set concerning phylogenetic information content and saturation of variation caused by multiple substitutions. We decided to use the approach introduced by (Lyons-Weiler, Hoelzer, and Tausch 1996), since it offers theoretical advantages over the permutation tests commonly used to check for hierarchical structure in a data matrix (e.g., Hillis 1991). This method makes use of the fact that in a hierarchically structured data matrix, the proportion between potential synapomorphies and potential informative similarity for any two taxa increases above that expected by chance alone. This offers avenues for statistical evaluation of hierarchical structure in any data set. The technique is implemented in the program RASA (for details of the method, see Lyons-Weiler, Hoelzer, and Tausch 1996). The results of RASA gave us information about the overall phylogenetic information content of the 16S and 12S data sets.

Additionally, RASA gave us information about potential long branches in the data set associated with single taxa. Taxa associated with long-branch effects were readily recognized by their significantly larger cladistic variance compared with other taxa (Lyons-Weiler and Hoelzer 1997). A significant deviation of the cladistic variance of single taxa was determined with *F*-statistics as implemented in RASA, version 2.3.

We used saturation plots to estimate areas in which phylogenetic information was expected to decay. We plotted frequencies of each substitution type against maximum-likelihood distances estimated by using the maximum-likelihood model fitted to the data (see below).

Outgroup Analysis

The selection of an optimal outgroup is crucial to derive a robust phylogenetic conclusion (from the extensive literature, see, e.g., Farris [1982] and, more recently, Milinkowitch and Lyons-Weiler [1998]). We used saturation plots to check for potential decay of phylogenetic signal caused by saturated nucleotide variation between outgroup and ingroup taxa.

In parallel, we used the RASA technique to evaluate our choice of outgroup taxa. An optimal outgroup is characterized by a relatively small number of autapomorphies and predominant plesiomorphies compared with ingroup taxa. Using this criterion, an optimal outgroup can also be selected prior to phylogenetic inference with RASA by constraining the relative synapomorphy scores (RAS) of taxon pairs with outgroup character states. A relative best outgroup maximizes the hierarchical structure in a rooted RASA analysis (Milinkowitch and Lyons-Weiler 1998). This can be verified by comparing the t_{RASA} (the test statistic in a test for homogeneity of slopes) value of the unrooted ingroup with that of the rooted ingroup analysis. With suitable outgroup assemblages, t_{RASA} increases.

Finding an Appropriate Substitution Model

We used the approach described by Sullivan, Holsinger, and Simon (1995), Frati et al. (1997), and Buckley, Simon, and Chambers (2000) to select an appropriate substitution model for further phylogenetic analysis and distance estimation. We made use of the MODELTEST program (Posada and Crandall 1998), which implements a model test routine based on the likelihood ratio statistics and the Akaike information criterion (Akaike 1974), in combination with PAUP, version 4.0b3a*, maximum-likelihood calculations. The likelihood ratio test (Goldman 1993*a*, 1993*b*; Yang et al. 1994) indicated that every increase in model complexity significantly improved fit to data. In total, we tested 42 different models, including rate heterogeneity and number of invariable sites. Parameters of the model (substitution parameters, shape of gamma distribution, proportion of invariable sites) were estimated from the data set using a neighbor-joining tree generated with JC-distances. The general time reversible (GTR) + %I + Γ model was found to be of adequate complexity. The Akaike information criterion implied the same conclusion.

Estimation of Rate Heterogeneity

Estimation of rate heterogeneity is relatively insensitive to topology as long as long branches in the tree are correctly identified (Sullivan, Holsinger, and Simon 1996). We compared shape estimates for six alternative phylogenies of Anisoptera and found no deviation in shape parameters. Subsequently, shape estimates were used to calculate posterior probabilities that a site belonged to a given rate category (Uzzell and Corbin 1971; Yang 1994) assuming a gamma distribution with eight categories by using PAUP, version 4.0b3a*. Due to different rates of change, character partitions associated with single-site rates are expected to display differing levels of resolution within the data set but are expected to support compatible topologies, assuming that homoplasy is distributed randomly (this does not necessarily have to be the case; see Naylor and Wesley Brown 1998). Fast-changing characters might introduce predominant noise to phylogenetic reconstructions due to an increased likelihood of multiple substitutions. We tested this hypothesis with two different approaches. First, we successively subtracted character sets associated with single-site-specific rates from the data set and calculated the remaining phylogenetic signal of the reduced data sets using RASA. The increase/decrease of the test parameter t_{RASA} was plotted against the number of included character sets. If a character subset harbors predominantly random structure, the test parameter t_{RASA} is expected to reflect this through an increase after removal of that character set. Second, we compared robustness of tree topologies reconstructed with single character partitions using bootstrapping and Bremer support (Bremer 1988; Baker and DeSalle 1997) implemented in TreeRot, version 2 (Sorensen 1999). Bremer support allowed us to study how fast support for nodes of interest eroded.

Phylogenetic Reconstructions

Prior to phylogenetic reconstructions, we tested for homogeneity of base frequencies among taxa using the χ^2 test as implemented in PAUP, version 4.0b3a* (which ignores correlations due to phylogenetic structure); over all sites, over parsimony-informative sites only, without constant sites (parsimony-uninformative and constant sites will mislead the χ^2 test); and over each character subset associated with single-site rates separately. All phylogenetic reconstructions were performed with PAUP, version 4.0b3a* (Swofford 1999). We performed maximum-parsimony, minimum-evolution, and maximum-likelihood reconstructions from the pruned data set (without fast-changing characters) and compared results with those of a logDet (Lockhart et al. 1994) reconstruction on the complete data set. LogDet reconstructions were performed on the complete data set, with invariable sites removed in proportion to the base frequencies estimated from constant sites (Waddell and Steel 1997).

Prior to the maximum-parsimony analyses, we evaluated the congruence of the two data sets with an ILD test (Farris et al. 1995; Hillis, Moritz, and Mable 1996; Swofford et al. 1996), as implemented in PAUP, version 4.0b3a*. Parsimony reconstructions were performed with heuristic searches on parsimony-informative characters, and gaps treated as missing characters, with random addition of taxa for 100 replications and tree bisection-reconnection (TBR) branch swapping. Bootstrapping was conducted on parsimony-informative characters for a total of 10,000 replicates of heuristic searches, with random addition of taxa.

We used the combined data set in our maximum-likelihood analysis. Separate analyses for the 12S and 16S fragments all found that a GTR + %I + Γ model fitted the data best. We assumed that deviation from the model parameters between the two sequenced gene fragments would not dramatically influence the results of the combined analyses. Heuristic tree searches under the maximum-likelihood criterion were performed with random addition of taxa for one replication and TBR branch swapping. Since tree reconstruction with maximum likelihood is extremely time-consuming, we restricted our search to the puzzling technique to evaluate statistical support (number of quartets = 101,270; number of puzzling steps = 100,000), being aware that quartet puzzling support tends to overestimate the robustness of nodes (Strimmer and von Haeseler 1996; Cao, Adachi, and Hasegawa 1998).

Minimum-evolution trees were reconstructed using maximum-likelihood distances of the pruned data set employing the GTR + I% + Γ model. Bootstrapping was performed with 10,000 replicates.

Results

PCR Amplification

Starting from a preliminary alignment of a restricted number of 16S sequences, we constructed a new 16S primer set which reliably amplified approximately 500 bp of every dragonfly we tested (table 1). Additionally, we used alignments of 12S and 16S fragments to construct primers for the amplification of a bridge between the 12S and 16S fragments (table 1). These primers amplify this target within every dragonfly family except for the Macromiidae. In total, we amplified roughly 1,290 bp from the 16S gene fragment and 635 bp from the 12S gene fragment. Sequences are deposited in GenBank under accession numbers AF266042–AF266103.

Alignment

The net number of unambiguously aligned positions totaled 1,540, and the net number of parsimony-informative characters was 645 for the combined data set. The data matrix is available on request.

Analyses of Phylogenetic Information Content and Substitutional Saturation

A hierarchical structure is clearly present in the combined data matrix (t_{RASA} = 10.13076, df = 857). Removing *Ephemera* sp., *Platycnemis pennipes*, *Calopteryx splendens*, *Chalcolestes viridis*, and *Heteragrion* sp. as the most distant outgroup taxa enhances the footprint of a hierarchical structure in the data matrix (t_{RASA} = 20.34925, df = 699) (fig. 1). For the complete data set, no single taxon exhibited an excess of cladistic variation typical for long-branch effects.

In figure 2, a saturation plot for the 16S gene fragment is presented. Saturation plots using maximum-likelihood distances suggest strong saturation of transitional substitutions for both fragments. Detectable saturated nucleotide variation in both fragments is largely restricted to relationships between outgroup and ingroup taxa, with one exception, that of *Epiophlebia superstes* and Anisoptera.

Finding an Optimal Outgroup

Saturation plots and the analysis of hierarchical structure suggest that removal of *Ephemera* sp. and the zygopteran taxa will enhance phylogenetic signal among ingroup taxa. Assuming that *E. superstes* is indeed an

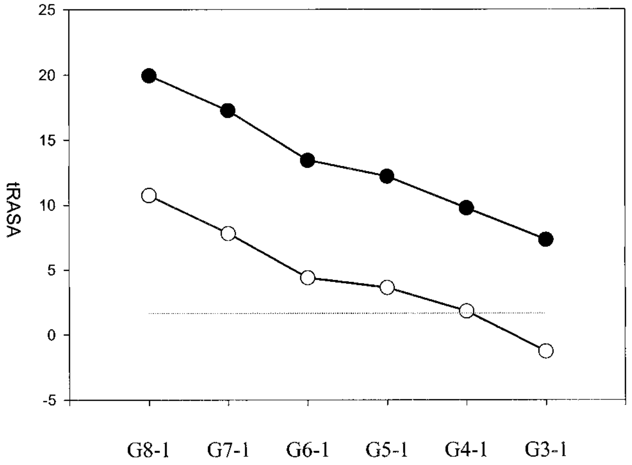

FIG. 1.—Phylogenetic signal of successively reduced data matrices. The t_{RASA} values for the complete data matrix and successively pruned matrices are plotted from left to right. Black circles represent the data set without *Ephemera* sp. and the zygopterous taxa; white circles represent the complete data set with all potential outgroup taxa. The dotted horizontal line marks the significance level for the Student's *t*-distribution at infinity ($\alpha = 1.645$, $P = 0.05$). The complete data matrix (g8–1) without *Ephemera* sp. and the zygopterous taxa has a t_{RASA} value of 19.94; removing character subsets associated with single-site rate estimates reduces t_{RASA} step by step. The character subset g3–1 still contains significant structure, with $t_{RASA} = 7.31$ (df = 677). The data set with all potential outgroup taxa (white circles) starts at a much lower t_{RASA} value (10.72) and becomes indistinguishable from a random matrix at character subset g4–1.

outgroup taxon to Anisoptera, it clearly emerges as the most suitable outgroup for rooting the anisopteran clade.

Test of Homogeneity of Base Frequencies

The Γ shape parameter indicated strong rate heterogeneity within both rRNA fragments. Of 1,540 characters, 129 characters of the 16S fragment and 50 characters of the 12S fragment were estimated to belong within the fastest rate category (G8) of the discrete gamma distribution. The remaining 410 sites were estimated to fall within rate categories 1–7 (G1–G7) (for rate estimates, see table 2).

Table 2
Single-Site Rate Estimates[a]

Category	Relative Rate
1	0.0244
2	0.1151
3	0.2558
4	0.4514
5	0.7216
6	1.1123
7	1.7516
8	3.5677

[a] Without *Ephemera* sp., Zygoptera.

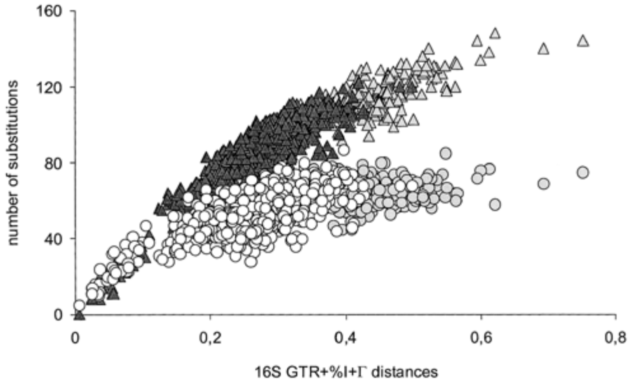

FIG. 2.—Saturation plot of 16S gene fragments. Numbers of transitions (circles) and transversions (triangles) for each pairwise comparison are plotted against maximum-likelihood distances for the 16S fragment. White circles and dark triangles depict ingroup relationships, including *Epiophlebia superstes*; gray circles and triangles depict ingroup-to-outgroup relationships excluding *E. superstes*. Maximum-likelihood distances range from 2% to roughly 77%. At least for transversions, nucleotide variation does not appear to be saturated. Transitions show saturation for ingroup-to-outgroup relationships. The 12S fragment displays a similar pattern.

When all characters were included, we found no significant deviation from homogeneity of base frequencies among taxa (table 3). Exclusion of the constant sites led to a significant deviation of homogeneity of base frequencies. Exclusion of parsimony-uninformative sites enhanced the pattern of base frequency heterogeneity. Excluding character subset G8 (fastest sites) restored homogeneity of base frequencies among taxa with and without parsimony-uninformative/constant sites. Separate analysis of character subset G8 yielded a significant deviation from the expected homogeneity of base frequencies.

Phylogenetic Signal Associated with Character Partitions

In RASA, adding character sets of single-site rates to the data set improved phylogenetic signal in each step (fig. 1). The test parameter t_{RASA} did not reach a local maximum. Fast-changing characters, particularly character set G8, did not diminish phylogenetic signal as inferred by the RASA method.

Parsimony analyses of both character partitions (G8 and G1–G7) resulted in two incongruent topologies. Support for the most parsimonious trees of character set G8 quickly decayed for the major nodes, and Bremer

Table 3
Tests for Homogeneity of Base Frequencies

Characters	No. of Characters	χ^2 (df = 120)	*P*
All	1,540	79.129	0.9985
All, excluding constant	800	157.803	0.0118
All, parsimony-informative only	589	201.42	0.0001
Set G8	138	199.07	0.0001
Set G7-1	1,402	45.63	1.0000
Set G7-1, parsimony-informative only . .	455	132.01	0.2138
Set G7-1, excluding constant	662	101.13	0.8935
Set G7	160	133.80	0.1837
Set G6	102	70.20	0.9999
Set G5	102	48.22	1.0000
Set G4	80	46.94	1.0000
Set G3	165	22.46	1.0000
Set G1	775	3.12	1.0000

support (BS) values were generally low, except for the monophyly of the Aeshnidae (BS = 9) and Cordulegastridae (BS = 20) and two subgroups within the gomphids. Gomphids and libellulids appear to be polyphyletic, petalurids appear to be paraphyletic, and Neopetaliidae and Chlorogomphidae appear to be the most basal taxa within Anisoptera. Bootstrap support did not climb beyond 40% in any node except for Aeshnidae and Cordulegastridae, and two subgroups within gomphids (Ophiogomphus + Onychogomphus, and Gomphus s.l.). Apparently, character set G8 is devoid of strong phylogenetic structure beyond some intrafamily relationships and might even contain misleading signal due to the nonstationarity of base frequencies across taxa.

For character set G7-1, the monophyly of Australopetaliidae + Aeshnidae, Gomphidae, Corduliidae s.l., and Libellulidae was robust against immediate decay; however, the monophyly of ‘‘Cavilabiata’’ quickly decayed (within three steps). Bootstrap support was generally high within families and low at the interfamily level.

In figure 3, the decrease in Bremer support is plotted against the logDet distances. Additionally, the total number of substitutions is also plotted against logDet distances. As can be expected, the median of decay indices starts out higher in fast-changing sites (fig. 3*B*) and quickly decreases above logDet distances of 12% in both fast-changing (G8) (fig. 3*B*) and slow-changing (G1–G7) (fig. 3*A*) characters. However, variance in Bremer support is much higher for slow-changing characters. Bremer support decreases before saturation of nucleotide substitution becomes apparent. Thus, both character partitions seemed to have experienced erosion of phylogenetic signal above logDet distances of 15%, although this was less pronounced in character set G7-1.

Phylogenetic Reconstructions with and Without Fast-Changing Sites

Without *Ephemera* sp. and the zygopteran taxa, 589 parsimony-informative positions remained for the complete set of characters. A minimum-evolution tree using logDet distances yielded robust reconstructions at the intrafamily level, judged by bootstrap resampling (fig. 4). Reconstructions of interfamily relationships did not experience notable support. Exceptions are the clear sister group relationship of Phyllopetalia (Austropetaliidae) and Aeshnidae and the monophyly of Macromiidae, Corduliidae, and Libellulidae (the MCL complex) with Chlorogomphidae.

Without fast-changing sites, 455 parsimony-informative positions remained. Maximum parsimony, minimum evolution, and the use of maximum-likelihood distances and maximum-likelihood reconstructions employing a GTR + %I + Γ model yielded a picture comparable to that of the logDet reconstruction (fig. 5). Resolutions within families were similar for maximum-parsimony and minimum-evolution trees; however, puzzle support for the maximum-likelihood tree was partly at odds with bootstrap support of the other reconstructions. This latter observation is not surprising, because (Cao, Adachi, and Hasegawa 1998) have already demonstrated that puzzle support values are inflated relative to bootstrap proportions and therefore cannot be interpreted in the same manner as bootstrap values.

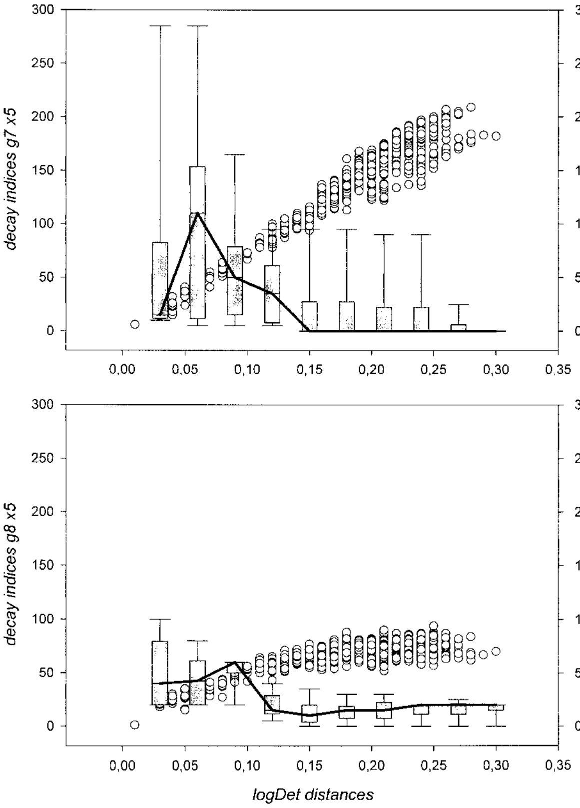

FIG. 3.—Decay of signal in relation to logDet distances and number of substitutions. *A,* The total number of substitutions for character set G7-1 (white circles) is plotted against logDet distances estimated from the complete character set. Additionally, corresponding to every taxon pair, decay values of the most recent common nodes of most parsimonious trees of this data set are plotted against logDet distances. LogDet distances are pooled into 10 classes to get a variance of decay values for each distance class. For example, decay values of corresponding distances from 0% to 3% are pooled. The variance of decay values in each pool is presented as box plots with 25% lower and 75% upper quartiles in gray and minimum and maximum values as bars. The median of decay values is shown as a continuous line. Decay indices are presented in a fivefold larger scale. Within character set G7-1, decay values reach a maximal median value at roughly 6% logDet distances and decline to 0 at 15% logDet distances. Substitutional saturation is not detectable in this plot. *B,* Same plot, but with characters of set G8 only. Again the total number of substitutions between each taxon pair and corresponding decay values of most-parsimonious trees for this data set are plotted against logDet distances of the complete character set. The median of decay indices starts out higher and declines to a level of 3–4 at 12% logDet distances. Saturation of substitutions is detectable.

Discussion

Phylogenetic Implications

We explored resolution in intrafamily and interfamily relationships. It turned out that, first, our molecular approach confirmed the monophyly of the morphologically recognized families and promises to be a valuable tool to resolve many of the open intrafamily relation-

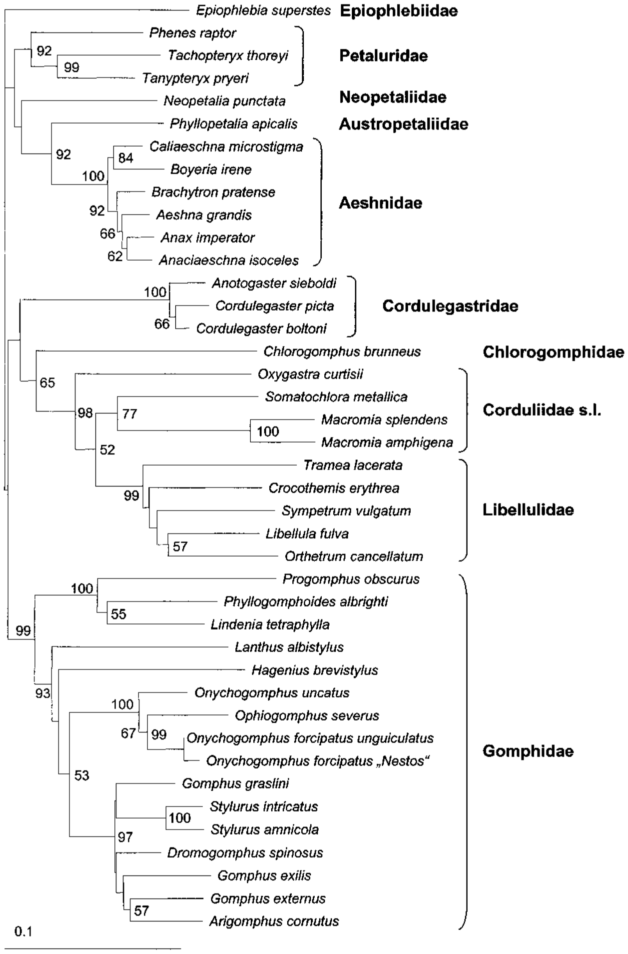

FIG. 4.—LogDet reconstruction for the complete data set. The reconstruction was rooted with *Epiophlebia superstes.* Bootstrap values above 50% are given at nodes. Branch lengths are proportional to logDet estimates.

ships. Within Petaluridae, Gomphidae, Aeshnidae, and Corduliidae s.l., our molecular data suggested a reinterpretation of plesiomorphic and apomorphic character states with consequences for the interpretation of character evolution for the entire group. For example, *Tachopteryx thoreyi* and *Tanypteryx pryeri* are sister taxa, contrary to recently published extensive revisions of petalurid dragonflies based on morphological data of fossil and extant representatives (Nel et al. 1998).

Second, we achieved robust reconstructions of interfamily relationships only for parts of the tree. A crucial problem turned out to be the recognition of the ''Cavilabiata,'' as is the case in morphological studies (Pfau 1991; Carle 1995; Lohmann 1996; Trueman 1996; Bechly et al. 1998). Despite an ecological diversity of larval and imaginal life history features, an apomorphic spoon-shaped development of the larval labial mask suggests a monophyly of the ''Cavilabiata.'' Morphological analyses emphasized this character as an apomorphic structure (e.g., Carle 1995; Bechly et al. 1998; in contrast to Pfau 1991), which led to the inclusion of *Neopetalia punctata* (Neopetaliidae) in the ''Cavilabiata'' (Carle and Louton 1994). The recognition of this monophylum is contradicted by an analysis of secondary sexual characters (Pfau 1991) and, additionally, leads to a paleontological paradox. The assumed most-ancestral representatives, the Cordulegastridae, the Neopetaliidae, and also the Chlorogomphidae, are unknown from fossil records (Bechly et al. 1998) in contrast to Lower Cretaceous representatives of the MCL complex (Nel and Paicheler 1994; Fleck, Nel, and Martinez-Delclòs 1999). Based on two wing characters, Bechly et al. (1998) proposed recognition of the extinct Hemeroscopidae as close relatives of the Chlorogomphidae, relaxing this paradoxical situation. Imaginal characters of Chlorogomphidae display many plesiomorphic features compared with species of the MCL complex, closely resembling representatives of the Cordulegastridae and *N. punctata.* The molecular data did not conclusively confirm the monophyly of the ''Cavilabiata,'' but it is compatible with this hypothesis. *Chlorogomphus brunneus* is at the stem of the monophyletic MCL complex, supporting a view based on predominantly larval characters (Bechly et al. 1998). Genetic distances of Chlorogomphidae to species of the MCL complex illustrate a pronounced molecular distinction of Chlorogomphidae from Neopetaliidae and Cordulegastridae not immediately apparent from morphological data. Thus, our molecular data have added structure to the stem of the group. Future inclusion of presumably basal cordulegastrid species, e.g., *Zoraena* sp. and *Chlorogomphus* ssp. and species of the families Gomphomacromiidae and Synthemistidae, will possibly further improve resolution.

Genetic distances of *E. superstes* from the anisopterous taxa imply a comparatively recent split between *E. superstes* and sampled anisopterous taxa, contrary to the view expressed in Nel et al. (1993) based on wing characters. Morphological analyses of ''Anisozygoptera'' larvae including larvae of *E. superstes* (unpublished data) support the results of the molecular analysis. Among other characters, larvae of *Epiophlebia* ssp. and Anisoptera share an absence of external larval gills. Comparison of Anisoptera and Epiophlebia larvae should provide clues as to which ecological factors induced loss of external larval gills in this monophyletic group.

Genetic distances of zygopterous taxa and anisopterous taxa are obviously inflated by saturated nucleotide variation, implying an old split between Zygoptera and Anisoptera, making them unfit outgroup taxa in our analysis; similarly, *Ephemera* sp. is unsuitable to serve as an outgroup taxon for Anisoptera. Assuming a relatively constant rate of change at the molecular level, our data indicate an ancient split between extant Zygoptera and Anisoptera + Epiophlebia. A direct relationship of extinct Permian zygopterous forms and extant Zygoptera might be worth considering.

Phylogenetic Signal Associated with Single-Site Rates

The characters assumed to change the most rapidly do not harbor phylogenetic signal sufficient for resolving most interfamily and intrafamily relationships, in

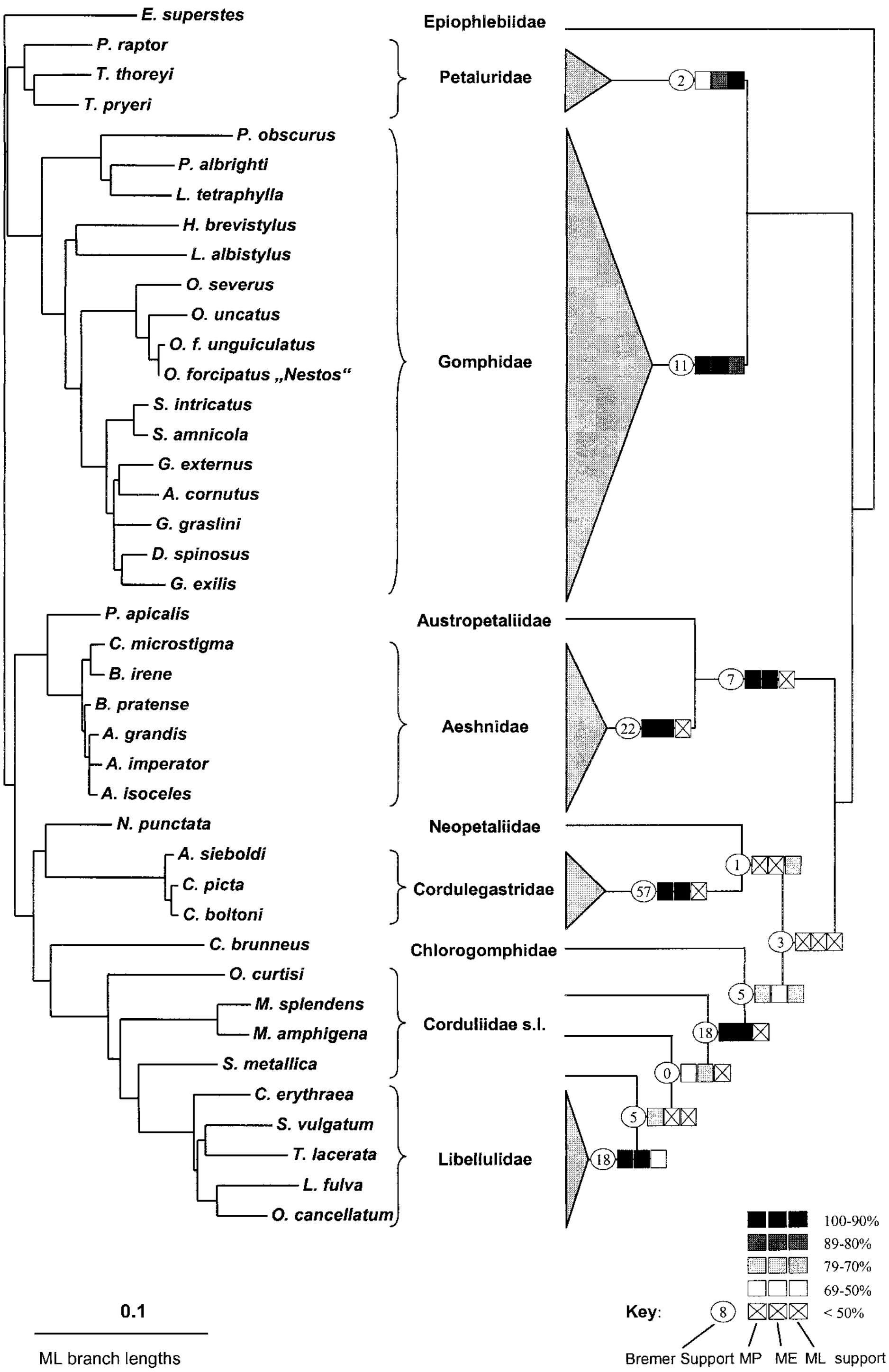

FIG. 5.—Phylogenetic reconstruction with the reduced data Set. Omitting characters with single-site rate estimates of gamma rate 8 (G8) results in largely congruent trees from all three reconstruction methods. "Cavilabiata" emerge as a monophyletic group (Neopetaliidae, Cordulegastridae, Chlorogomphidae, Corduliidae s.l., Libellulidae) and as a sister group to the Aeshnoidea (Austropetaliidae + Aeshnidae). However, bootstrap support (maximum parsimony, minimum evolution) and puzzle support (maximum likelihood) for basal anisopterous relationships are still weak, as illustrated in the schematic tree. Support values are displayed in a color code on this tree, outlining the major results of our analyses. Support values for relationships within families are omitted in this figure to improve the clarity of the drawing.

contrast to characters assumed to change with rate 1–7. Lack of phylogenetic resolution in character set G8 must be attributed to the erosion of phylogenetic information over time due to multiple hits. In figure 3*B*, a decrease in Bremer support coincides with saturation of nucleotide substitution. At roughly 12% logDet distances, decay values of character set G8 level off at 3–4. We assume this pattern to be due to the high evolutionary rates of these characters. Simulations will show whether this interpretation is correct. Character set G7-1 exhibits a pattern similar to that of character set G8, but signal decay is delayed (fig. 3*A*). Variance of decay indices remains high up to 24% logDet distances, indicating that there is still some phylogenetic signal. It is clear that comparing genetic distances, decay indices of common nodes, and the inferred number of substitutions provides only a rough record of signal dilution. Nevertheless, these plots demonstrate that the robustness of tree reconstruction declines with or before substitutional saturation becomes detectable in both data sets. This is in contrast to recently published reports (e.g., Yang 1998) in which good performance of reconstruction methods was recorded despite extensive substitutional saturation. The disparity between Yang's results (1998) and ours might indicate that even the best evolutionary substitution model is still an insufficient description of the actual biological processes shaping the pattern of nucleotide variation within these gene fragments.

Base composition among fast-changing sites (subset G8) deviated significantly from homogeneity, further emphasizing their problematic nature (table 3) by possibly introducing misleading information. Saturation, base composition, and decay analyses imply that decay of phylogenetic signal associated with among-site rate variation is the important player in our analysis. Our findings agree with empirical results of Sullivan, Holsinger, and Simon 1995) and simulation results of Yang (1995, 1996), which demonstrated that effects of transition/transversion rate bias can be minor in comparison with those of among-site rate variation and unequal branch lengths.

Conclusions

We are able to propose solutions to some of the systematic questions formulated in the introduction. In particular, intrafamily relationships have been resolved to a large extent. However, there are clear limits of resolution within the order Anisoptera for both sequenced fragments. When 16S and 12S are combined, phylogenetic signal decays progressively from the intrafamily level to the interfamily level to the outgroup comparisons. It remains to be tested whether more exhaustive taxon sampling can compensate for progressive signal decay (compare Graybeal 1998). As crude as our analysis is, we do find the expected association between speed of signal decay, genetic distance, and estimated evolutionary rate. Phylogenetic constraints on single-site rates clearly play an important role in signal decay, which is generally not sufficiently appreciated in the current literature dealing with ribosomal fragments (an exception is Hickson et al. 1996). Generation of (un)observed nucleotide variation in non–protein-coding sequences is certainly of different complexity than that in protein-coding sequences; for example, codon biases do not play a causal role in generating the observed variance. Future investigations will elucidate the complex differences in generation of variation between protein-coding and non–protein-coding sequences and the associated loss/gain of phylogenetic information.

Acknowledgments

We are extremely indebted to the people who collected dragonflies for us: H. White (North American Gomphidae, Tramea), S. Dunkle (Phyllogomphoides, Progomphus, Tachopteryx, Heteragrion), I. Ueda (Tanypteryx, *Macromia amphigena*), A. Camousseight (Phyllopetalia, Phenes, Neopetalia), W. Heymer (Boyeria, Oxygastra, *Macromia splendens*), F. Suhling (*Onychogomphus uncatus*, *Onychogomphus forcipatus unguiculatus*, *Gomphus graslini*), and T. Aoki (Anotogaster, Chlorogomphus, Epiophlebia). B.M. was supported by a Schroedinger Postdoctoral Fellowship granted by the Austrian Science Foundation. We also acknowledge G. K. Chambers, C. Simon, and Victoria University of Wellington for funding and support.

LITERATURE CITED

Akaike, H. 1974. A new look at the statistical model identification. IEEE Trans. Autom. Contr. **19**:716–723.

Bechly, G. H. P. 1996. Morphologische Untersuchungen an Fluegelgeaeder der rezenten Libellen und deren Stammgruppenvertreter (Insecta; Pterygota; Odonata) unter besonderer Beruecksichtigung der phylogenetischen Systematik und des Grundplanes der Odonata. Petalura **2**(Special Vol.):1–402.

Bechly, G., A. Nel, X. Martinez-Delclos, and G. Fleck. 1998. Four New dragonflies from the upper jurrasic of Germany and the lower cretaceous of Mongolia (Anisoptera: Hemeroscopidae, Sonidae, and Proterogomphidae Fam. nov.). Odonatologica **27**:149–187.

Baker, R. H., and R. DeSalle. 1997. Multiple sources of character information and the phylogeny of Hawaiian Drosophilids. Syst. Biol. **46**:654–673.

Bremer, K. 1988. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. Evolution **42**:795–803.

Buckley, T. R., C. M. Simon, and G. K. Chambers. 2000. Exploring among-site rate variation models in a maximum likelihood framework using empirical data: the effects of model assumptions on estimates of topology, branch Lengths, and bootstrap support. Syst. Biol. (in press).

Buckley, T. R., C. Simon, P. K. Flook, and B. Misof. 2000. Secondary structure and conserved motifs of the frequently sequenced domains IV and V of the insect mitochondrial large subunit rRNA gene. Insect Mol. Biol. (in press).

Cao, Y., J. Adachi, and M. Hasegawa. 1998. Comment on the quartet puzzling method for finding maximum-likelihood tree topologies. Mol. Biol. Evol. **15**:87–89.

Carle, F. L. 1986. The classification, phylogeny and biogeography of the Gomphidae (Anisoptera). I. Classification. Odonatologica **15**:275–326.

———. 1995. Evolution, taxonomy, and biogeography of ancient Gondwanian Libelluloides, with comments on anisopteroid evolution and phylogenetic systematics (Anisoptera: Libelluloidea). Odonatologica **24**:383–424.

Carle, F. L., and T. A. Louton. 1994. The larva of *Neopetalia punctata* and establishment of Austropetaliidae fam. nov. (Odonata). Proc. Entomol. Soc. Wash. **96**:147–155.

Chippindale, P. T., V. K. Davé, D. H. Whitmore, and J. V. Robinson. 1999. Phylogenetic relationships of North American damselflies of the genus *Ischnura* (Odonata: Zygoptera: Coenagrionidae) based on sequences of three mitochondrial genes. Mol. Phylogenet. Evol. **11**:110–121.

de Rijk, P., Y. Van De Peer, and R. de Wachter. 1997. Database on the structure of the large ribosomal RNA. Nucleic Acids Res. **25**:117–123.

Farris, J. S. 1982. Outgroups and parsimony. Syst. Zool. **31**: 328–334.

Farris, J. S., M. Källersjö, A. G. Kluge, and C. Bult. 1995. Testing significance of incongruence. Cladistics **10**:315–319.

Fleck, G. 1996. La phylogénie des Odonata Anisoptera. Analyse de nouveaux caractéres des structures céphaliques (Insecta). D.E.A. Sys. anim. and végét. Lab. Ent. Master's thesis, Muséum National d'Histoire Naturelle, Paris.

Fleck, G., A. Nel, and X. Martinez-Delclòs. 1999. The oldest record of libellulid dragonflies from the Upper Cretaceous of Kazakhstan (Insecta: Odonata, Anisoptera). Cretaceous Res. **20**:655–658.

Frati, F., C. Simon, J. Sullivan, and D. L. Swofford. 1997. Evolution of the mitochondrial cytochrome oxidase II gene in Collembola. J. Mol. Evol. **44**:145–158.

Gatesy, J., R. DeSalle, and W. Wheeler. 1993. Alignment-ambiguous nucleotide sites and the exclusion of systematic data. Mol. Phylogenet. Evol. **2**:152–157.

Gerken, T., J. Kurtz, K. P. Sauer, and T. Lubjuhn. 1998. DNA preparation and efficient microsatellite analysis from insect hemolymph. Electrophoresis **19**:3069–3070.

Graybeal, A. 1998. Is it better to add taxa or characters to a difficult phylogenetic problem. Syst. Biol. **47**:9–17.

Goldman, N. 1993*a.* Simple diagnostic statistical tests of models for DNA substitution. J. Mol. Evol. **37**:650–661.

———. 1993*b*. Statistical tests of models of DNA substitution. J. Mol. Evol. **36**:182–198.

Hadrys, H., M. Balick, and B. Schierwater. 1992. Applications of random amplified polymorphic DNA (RAPD) in molecular ecology. Mol. Ecol. **1**:55–65.

Hickson, R. E., C. Simon, A. Cooper, G. S. Spicer, J. Sullivan, and D. Penny. 1996. Conserved sequence motifs, alignment, and secondary structure for the third domain of animal 12S rRNA. Mol. Biol. Evol. **13**:150–169.

Higgins, D. G., and P. M. Sharp. 1988. CLUSTAL: a package for performing multiple sequence alignments on a microcomputer. Gene **73**:237–244.

Hillis, D. M. 1991. Discriminating between phylogenetic signal and random noise in DNA sequences. Pp. 278–295 *in* M. M. Miyamoto and J. Cracraft, eds. Phylogenetic analysis of DNA sequences. Oxford University Press, New York and Oxford.

Hillis, D., C. Moritz, and B. K. Mable. 1996. Molecular systematics. Sinauer, Sunderland, Mass.

Kambhampati, S., and R. E. Charleton. 1999. Phylogenetic relationship among *Libellula*, *Ladona* and *Plathemis* (Odonata: Libellulidae) based on DNA sequences of mitochondrial 16S rRNA gene. Syst. Entomol. **24**:37–50.

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.

Lohmann, H. 1996. Das phylogenetische System der Anisoptera (Odonata). Dtsch. Entomol. Z. **106**:209–252, 253–266.

Lyons-Weiler, J., and G. A. Hoelzer. 1997. Escaping from the Felsenstein Zone by detecting long branches in phylogenetic data. Mol. Phylogenet. Evol. **8**:375–384.

Lyons-Weiler, J., G. A. Hoelzer, and R. J. Tausch. 1996. Relative apparent synapomorphy analysis (RASA) I: the statistical measurement of phylogenetic signal. Mol. Biol. Evol. **13**:749–757.

Milinkovitch, M. C., M. Berube, and P. J. Palsboll. 1998. Cetaceans are highly specialized artiodactyls. Pp. 113–131 *in* J. G. M. Thewissen, ed. The emergence of whales: evolutionary patterns in the origin of Cetacea. Plenum, New York.

Milinkovitch, M. C., and J. Lyons-Weiler. 1998. Finding optimal ingroup topologies and convexities when the choice of outgroups is not obvious. Mol. Phylogenet. Evol. **9**:348–357.

Misof, B., C. Anderson, and H. Hadrys. 2000. A phylogeny of the damselfly genus *Calopteryx* (Odonata) using mitochondrial 16S rDNA markers. Mol. Phylogenet. Evol. **15**: 5–14.

Naylor, G. J. P., and M. Wesley Brown. 1998. Amphioxus mitochondrial DNA, chordate phylogeny, and the limits of inference based on comparisons of sequences. Syst. Biol. **47**:61–76.

Nel, A., G. Bechly, E. Jarzembowski, and X. Martínez-Delclòs. 1998. A revision of the fossil petalurid dragonflies (Insecta: Odonata: Anisoptera: Petalurida). Paleontol. Lombarda **10**:1–68.

Nel, A., X. Martinez-Delclos, J.-C. Paicheler, and M. Henrotay. 1993. Les ''Anisozygoptera'' fossiles: phylogénie et classification (Odonata). Martinia **3**:1–311.

Nel, A., and J.-C. Paicheler. 1994. Les Libelluloidea autres que Libellulidae fossiles. Un inventaire critique (Odonata, Corduliidae, Macromiidae Synthemistidae, Chlorogomphidae et Mesophlebiidae). Nouv. Rev. Entomol. **11**:321–334.

Pfau, H. K. 1991. Contributions of functional morphology to the phylogenetic systematics of Odonata. Adv. Odonatol. **5**: 109–141.

Posada, D., and K. A. Crandall. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics **124**:817–818.

Simon, C., F. Frati, A. Beckenbach, B. Crespi, H. Liu, and P. Flook. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. **87**:651–701.

Sorensen, M. D. 1999. TreeRot. Version 2. Boston University, Boston.

Strimmer, K., and A. von Haeseler. (1996). Quartet puzzling: a quartet maximum likelihood method for reconstructing tree topologies. Mol. Biol. Evol. **13**:964–969.

Sullivan, J., K. E. Holsinger, and C. Simon. 1995. Among-site rate variation and phylogenetic analysis of the 12S rRNA in Sigmodontine rodents. Mol. Biol. Evol. **12**:988–1001.

———. 1996. The effect of topology on estimates of among-site rate variation. J. Mol. Evol. **42**:308–312.

Swofford, D. L. 1999. Beta update PAUP 4.0b3a* of ''PAUP 3.1.1; phylogenetic analysis using parsimony.'' Illinois Natural History Survey, Champaign.

Swofford, D. L., G. J. Olsen, P. J. Wadell, and D. M. Hillis. 1996. Phylogenetic inference. Pp. 407–514 *in* D. M. Hillis, C. Moritz, and B. K. Mable, eds. Molecular systematics. 2nd edition. Sinauer, Sunderland, Mass.

Truemann, J. W. H. 1996. A preliminary cladistic analysis of odonate wing venation. Odonatologica **25**:59–72.

Uzzell, T., and K. W. Corbin. 1971. Fitting discrete probability distributions to evolutionary events. Science **172**: 1089–1096.

Waddell, P. J., and M. A. Steel. 1997. General time-reversible distances with unequal rates across sites: mixing Γ and inverse Gaussian distributions with invariant sites. Mol. Phylogenet. Evol. **8**:398–414.

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 rate over sites: approximate methods. J. Mol. Evol. **39**:306–314.

———. 1995. Evaluation of several methods for estimating phylogenetic trees when substitution rates differ over nucleotide sites. J. Mol. Evol. **40**:689–697.

———. 1996. The among-site rate variation and its impact on phylogenetic analyses. TREE **11**:367–372.

———. 1998. On the best evolutionary rate for phylogenetic analysis. Syst. Biol. **47**:125–133.

AXEL MEYER, reviewing editor

Accepted August 23, 2000