Molecular Phylogenetics and Evolution
Vol. 22, No. 1, January, pp. 101–110, 2002
doi:10.1006/mpev.2001.1028, available online at http://www.idealibrary.com on IDEAL®

Multiple Genes and the Monophyly of Ischnocera (Insecta: Phthiraptera)

Kevin P. Johnson* and Michael F. Whiting†

*Illinois Natural History Survey, Champaign, Illinois 61820; and †Department of Zoology, Brigham Young University, Provo, Utah 84602

Received March 9, 2001; revised June 29, 2001; published online December 5, 2001

Whereas most traditional classifications identify Ischnocera as a major suborder of lice in the order Phthiraptera, a recent molecular study based on one gene did not recover monophyly of Ischnocera. In this study we test the monophyly of Ischnocera using sequences of portions of three different genes: two nuclear (EF1α and 18S) and one mitochondrial (COI). Analysis of EF1α and COI sequences did not recover monophyly of Ischnocera, but these genes provided little support for ischnoceran paraphyly because homoplasy is high among the divergent taxa included in this study. Analysis of 18S sequences recovered ischnoceran monophyly with strong support. Sequences from these three gene regions showed significant conflict with the partition homogeneity test, but this heterogeneity probably arises from the dramatic differences in substitution rates. In support of this conclusion, Kishino–Hasegawa tests of the EF1α and COI genes did not reject several trees containing ischnoceran monophyly. Combined analysis of all three gene regions supported monophyly of Ischnocera, although not as strongly as analysis of 18S by itself. In sum, although rapidly evolving genes can retain some phylogenetic signal for deep phylogenetic relationships, strong support for such relationships is likely to come from more slowly evolving genes. © 2001 Elsevier Science

INTRODUCTION

Lice (Insecta: Phthiraptera) are ectoparasites of birds and mammals. The order Phthiraptera is currently divided into four suborders: Amblycera, Ischnocera, Rhyncophthirina, and Anoplura. Amblycera, Ischnocera, and Rhyncophthirina have chewing mouthparts, whereas Anoplura have sucking mouthparts. Although chewing lice have sometimes been classified as part of the insect order Mallophaga (Ferris, 1951; Kim and Ludwig, 1982), nearly all recent treatments of this group place the chewing lice and sucking lice (Anoplura) in a single order Phthiraptera (Clay, 1970; Lyal, 1985; Barker, 1994). Paraphyly of "Mallophaga" is identified on the basis of morphological characters (Lyal, 1985).

Lice in the suborder Ischnocera are permanent parasites of both birds and mammals. Ischnoceran lice feed on feathers and/or dermal debris of the host (Marshall, 1981). Phylogenies of several groups of ischnoceran lice have been constructed and compared to host phylogenies (Hafner *et al.*, 1994; Paterson *et al.*, 2000; Johnson and Clayton, 2001). These studies have generally indicated a macroevolutionary pattern consistent with cospeciation, making Ischnocera an important group for comparisons of cophylogenetic histories and relative rates of molecular evolution. However, further study will rely on the identification of groups based on classifications that reflect phylogeny, such that monophyletic groups can be recognized and selected. Higher level relationships among major groups of lice are still not understood completely. Here we address the question of ischnoceran monophyly.

Whereas the monophyly of Ischnocera is generally supported on the basis of morphology (Lyal, 1985), a recent analysis of elongation factor 1α (EF1α) sequences (Cruickshank *et al.*, 2001) failed to confirm the monophyly of Ischnocera. Paraphyly of Ischnocera would be surprising and be at odds with all previous classifications of higher taxa of lice. Cruickshank *et al.*'s (2001) result based on EF1α sequences was not simply an issue of rooting, because rooting on the suborder Amblycera (as argued by Lyal, 1985) still resulted in Anoplura being imbedded within Ischnocera.

The goal of the present study is to further test the monophyly of Ischnocera with additional gene sequences. Here we examine sequences from two nuclear (EF1α and ribosomal 18S) and one mitochondrial (cytochrome oxidase I [COI]) gene to explore whether these genes contain signal for ischnoceran monophyly. We compare gene regions in relation to their usefulness in reconstructing phylogenetic relationships by comparing overall homoplasy and levels of bootstrap support. We use this study to illustrate how differences

1055-7903/01 $35.00
© 2001 Elsevier Science
All rights reserved.

TABLE 1

Specimens Included in Study

Species	Host	Suborder	GenBank Accession Nos. COI	EF1α	18S
Campanulotes compar	*Columba livia*	Ischnocera	AF384997	AF348643	AF385036(7)
Colilipeurus colius	*Urocolius indicus*	Ischnocera	AF384998	AF385017	AF384046(7)
Cirrophthirius testudinarius	*Recurvirostra americana*	Ischnocera	AF384999	AF385018	AF385050(1)
Brueelia sp.	*Parus niger*	Ischnocera	AF385000	AF385019	AF385038(9)
Austrophilopterus subsimilis	*Ramphastos sulfuratus*	Ischnocera	AF385001	AF385020	AF385052(3)
Rhynonirmus sp.	*Scolopax* n. sp.	Ischnocera	AF385002	AF385021	AF385048(9)
Columbicola columbae	*Columba livia*	Ischnocera	AF385003	AF385022	AF385044(5)
Geomydoecus craigi	*Thomomys talpoides*	Ischnocera	AF385004	AF348667	AF385040(1)
Fulicoffula longipila	*Fulica americana*	Ischnocera	AF385005	AF385023	AF385042(3)
Anatoecus sp.	*Anas platyrhynchos*	Ischnocera	AF385006	AF385024	AF385056(7)
Oxylipeurus chiniri	*Ortalis vetula*	Ischnocera	AF348872	AF385025	AF385054(5)
Haematopinus pacochoeri	*Phacochoerus aethiopicus*	Anoplura	AF385007	AF385026	AF385058(9)
Neohaematopinus sciuri	*Sciurus carolinensis*	Anoplura	AF385008	AF385027	AF385060(1)
Haematomyzus elephantis	*Elaphus maximus*	Rhyncophthirina	AF385009	AF385028	AF385062(3)
Colimenopon urocolius	*Urocolius indicus*	Amblycera	AF385010	AF385029	AF385070(1)
Menacanthus sp.	*Penelope purpurascens*	Amblycera	AF385011	AF385030	AF385066(7)
Machaerilaemus sp.	*Hirundo abyssinica*	Amblycera	AF385012	AF385031	AF385068(9)
Dennyus hirundinis	*Apus apus*	Amblycera	AF385013	AF385032	AF385064(5)
Ricinus sp.	*Cyanocompsa parellina*	Amblycera	AF385014	AF385033	AF385072(3)
Trinoton querquedulae	*Anas platyrhynchos*	Amblycera	AF385015	AF385034	AF385074(5)
Laemobothrion atrum	*Fulica americana*	Amblycera	AF385016	AF385035	AF385076(7)

in rates of molecular substitution can be an important factor in molecular phylogenetic studies.

MATERIALS AND METHODS

Samples and Sequencing

We obtained representative lice from each of the four suborders within Phthiraptera (Table 1). These specimens include lice from both bird and mammal hosts. Using previous classifications (Hopkins and Clay, 1952; Eichler, 1941, 1963) and a morphologically based phylogeny (Smith, 2001), we selected a diversity of ischnoceran genera spanning the breadth of samples included in Cruickshank *et al.*'s (2001) study. The main objective of this sampling scheme was to include enough representation of louse diversity to determine whether Anoplura and/or Rhyncophthirina falls within Ischnocera. A diversity of amblyceran genera were used as a composite outgroup to root the tree as indicated by Lyal (1985). Homologous EF1α sequences were not available for a more distant outgroup, Psocoptera (Cruickshank *et al.*, 2001), so we were not able to test Lyal's (1985) suggested rooting with these data. However, analysis of 18S sequences for Psocoptera and Phthiraptera indicated a sister relationship between Amblycera and other lice (unpublished data), further supporting our choice of rooting.

Lice were stored either in 95% ethanol in a −20°C freezer or dry in a −70°C freezer. From individual lice, we extracted DNA and prepared a voucher specimen mounted on a microslide from the same specimen for identification as those described by Johnson *et al.*, (2001). We used PCR to amplify portions of the EF1α, 18S, and COI genes. Reaction conditions followed Johnson and Clayton (2000) for EF1α and COI. The primers EF1-For3 and Cho10 (Danforth and Ji, 1998) were used to amplify EF1α, and the primers L6625 and H7005 (Hafner *et al.*, 1994) were used to amplify COI. These same primers were used in sequencing reactions. We amplified and sequenced 18S as described by Whiting *et al.* (1997). We performed cycle DNA sequencing with *Taq* FS DNA polymerase using either ABI dRhodamine dye terminators or ABI Prism BigDye Terminators and followed manufacturer's protocols (Perkin–Elmer). We collected and analyzed DNA sequence data (see Table 1 for GenBank Accession numbers) using an ABI Prism 377 automated DNA sequencer (PE Applied Biosystems). We aligned and reconciled complementary chromatograms using Sequencher 3.1 (GeneCodes). We also used Sequencher to align sequences across species; final alignments were done by eye. For 18S, several indel regions were identified. To minimize the impact of alignment on phylogenetic analysis in this study, we excluded any regions of ambiguous alignment from all analyses. Whereas this had the consequence of eliminating many potentially useful regions for the 18S sequences, we wanted to eliminate alignment decisions as a potential source of difference between trees derived from different gene regions (Lutzoni, 1997).

Phylogenetic Analyses

The main goal of our study was to test the monophyly of Ischnocera using multiple gene sequences. Here we outline several analyses conducted using PAUP* (Swofford, 2000) to compare and combine phylogenetic information from the three gene regions. First, we conducted parsimony analysis on each gene region separately with all base positions unweighted and unordered. We used 20 random addition searches with TBR branch swapping in these searches (shortest trees were recovered in >90% of random addition replicates). In the case in which more than one most parsimonious tree resulted from these searches, we computed a strict consensus tree. We compared trees based on different gene regions using a symmetric difference distance (Penny and Hendy, 1985) for all possible tree comparisons. We evaluated bootstrap support (Felsenstein, 1985) for nodes in these trees using 1000 full heuristic bootstrap replicate searches. We also tested for significant signal in each data set using the permutation tail probability test (Archie, 1989; Faith and Cranston, 1991).

We compared the homogeneity of the signal from each of three gene regions by conducting a series of partition homogeneity tests (Farris *et al.*, 1994, 1995; Swofford, 2000). We conducted a test between each pair of gene regions (three tests) and among all three gene regions simultaneously (one test) using 100 replicates for each test. To evaluate trees resulting from combined analyses, we conducted parsimony searches as indicated above on all possible pairwise combinations of gene regions and again with all three regions combined in a single analysis.

As an alternative method to parsimony analysis, we used a maximum-likelihood (ML) approach to reconstruct a tree based on all three gene regions both individually and combined. Using the parsimony tree as a starting point, we used the procedure of Huelsenbeck and Crandall (1997) to select a likelihood model. We estimated parameters for a general time reversible model (GTR), with estimated base frequencies and rate heterogeneity according to a gamma distribution. We partitioned the gamma distribution into eight rate categories for these analyses. We used 10 random addition replicates with TBR branch swapping to search for the most likely trees (most likely trees were recovered in ≥90% of replicates). We also constructed bootstrap replicates for each analysis, using 100 replicates with random addition and NNI branch swapping. We compared likelihood trees using symmetric difference distances, but included comparisons of trees constructed only from single genes or all combined. With the likelihood model, we used trees from the parsimony analyses (Figs. 1 and 2) in Kishino–Hasegawa tests (Kishino and Hasegawa, 1989). With each gene region, we determined which parsimony tree was the most likely and which tree(s) could be rejected under the model.

RESULTS

Of 348 bp for EF1α, 147 (42%) were variable and 124 (36%) parsimony informative. For 18S, we included 1398 bp and 296 (21%) of these were variable, with 200 (14%) informative. For COI, 229 (60%) of 384 bp were variable, and 202 (53%) were informative. Uncorrected pairwise sequence divergences were higher for COI (18.6–39.1%) than for EF1α (11.6–28.7%). Pairwise divergences for well-aligned regions of 18S ranged from 1.1 to 12.2% for these same taxa. These divergence levels corresponded to the level of homoplasy present in each gene region with consistency indices for each gene (over the unordered parsimony tree derived from that gene) of 0.342 for COI, 0.380 for EF1α, and 0.632 for 18S. Each data set contained significant structure (all $P < 0.001$) as determined by the permutation tail probability test (Archie, 1989; Faith and Cranston, 1991).

Trees reconstructed from parsimony analyses of each gene independently differed (Figs. 1a–1c). Trees resulting from COI and EF1α sequences alone, while resolving relationships within Ischnocera, did not recover ischnoceran monophyly. In contrast, the strict consensus of trees constructed from parsimony analysis of 18S sequences did not resolve relationships within Ischnocera, but recovered monophyly of Ischnocera with high support (bootstrap 90%). As expected, less homoplasious genes generally provided higher bootstrap support overall for recovered nodes. Bootstrap analysis of 18S produced 10 nodes supported at over 50%, whereas analysis of EF1α and COI resulted in many fewer nodes supported at the 50% level, 3 and 1, respectively. Part of the reason that 18S trees had higher bootstrap support may stem from differences in the total length of the fragment sequenced and included in the analysis, because larger data sets tend to increase bootstrap support (Johnson and Sorenson, 1998; Johnson and Lanyon, 1999; DeFilippis and Moore, 2000). However, in this case the numbers of potentially informative characters are generally similar. Analyses of EF1α and 18S recovered monophyly of Amblycera with respect to Ischnocera, Anoplura, and Rhyncophthirina. However, parsimony analysis of COI alone did not recover monophyly of Amblycera, with one taxon (*Dennyus*) appearing within Ischnocera.

Partition homogeneity tests between gene regions were not significant for a comparison of EF1α and COI ($P = 0.42$). However, all three comparisons involving one or both of these genes in combination with 18S produced a significant partition homogeneity test (all $P < 0.01$). Parsimony analysis of two-gene combined data sets generally produced well-resolved trees (Figs. 1d–1f). In all but one case, combined analysis of two

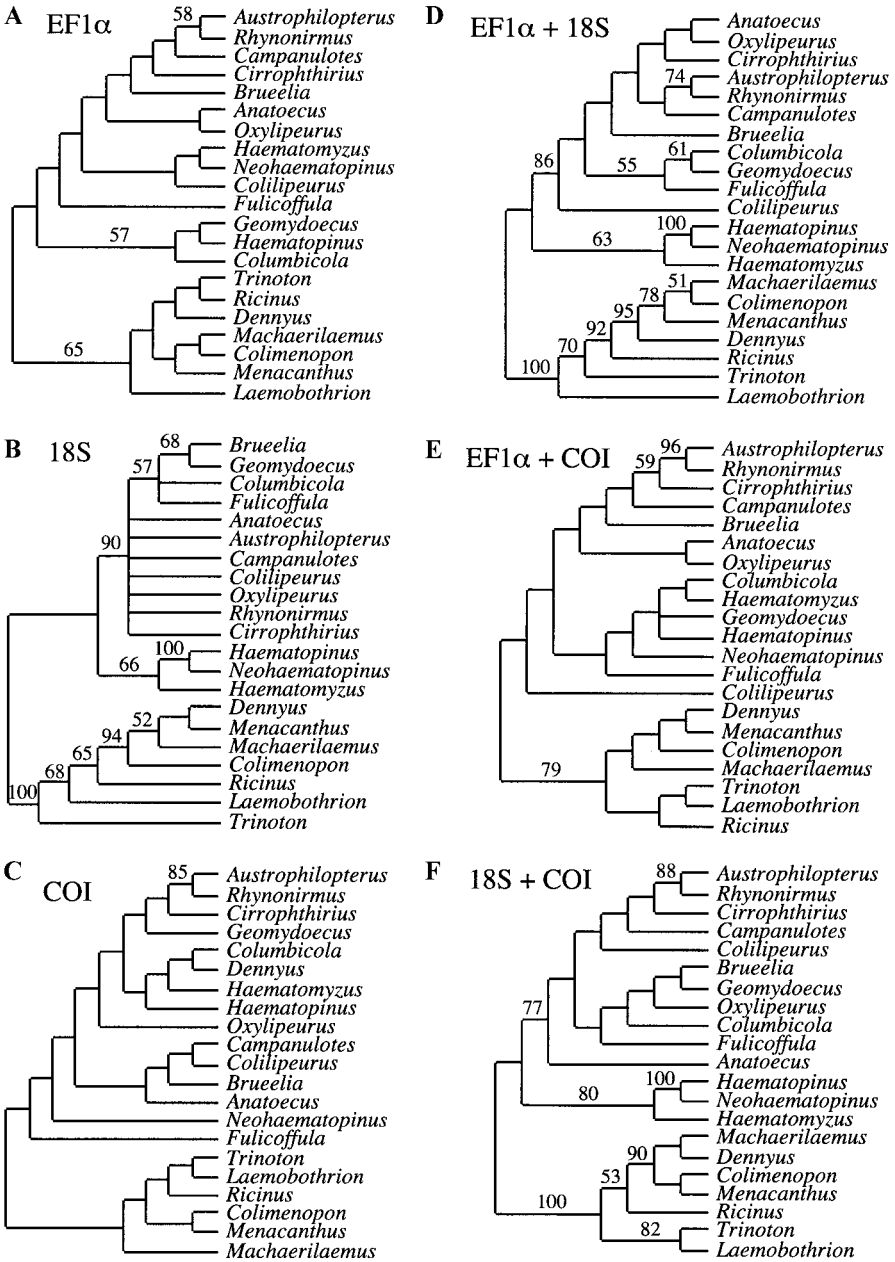

FIG. 1. Strict consensus of trees recovered from unordered parsimony analysis of (a) EF1α only [1 tree, length = 690, CI = 0.380], (b) 18S only [38 trees, length = 717, CI = 0.632], (c) COI only [1 tree, length = 1399, CI = 0.342], (d) EF1α + 18S [1 tree, length = 1443, CI = 0.495], (e) EF1α + COI [2 trees, length = 2119, CI = 0.350], and (f) 18S + COI [1 tree, length = 2165, CI = 0.430]. Numbers above branches indicate support from 1000 bootstrap replicates.

genes produced as many or more nodes supported in over 50% of bootstrap replicates than did analysis of either gene independently (Fig. 1). Combined analysis of 18S and COI sequences resulted in only 8 nodes supported in 50% of bootstrap replicates, compared to 10 nodes for 18S alone. The two combined analyses involving 18S recovered monophyly of Ischnocera (Figs. 1d and 1f), whereas combined analysis of EF1α and COI did not (Fig. 1e). All combined analyses recovered monophyly of Amblycera, and bootstrap support for this in combined trees was higher or equal to bootstrap support in trees based on independent analyses.

Combined parsimony analysis of all three genes simultaneously produced a completely resolved tree (Fig. 2). This tree contained 10 nodes supported at over 50% by bootstrap replicates, the same number as analysis of 18S alone. Whereas monophyly of Ischnocera was sup-

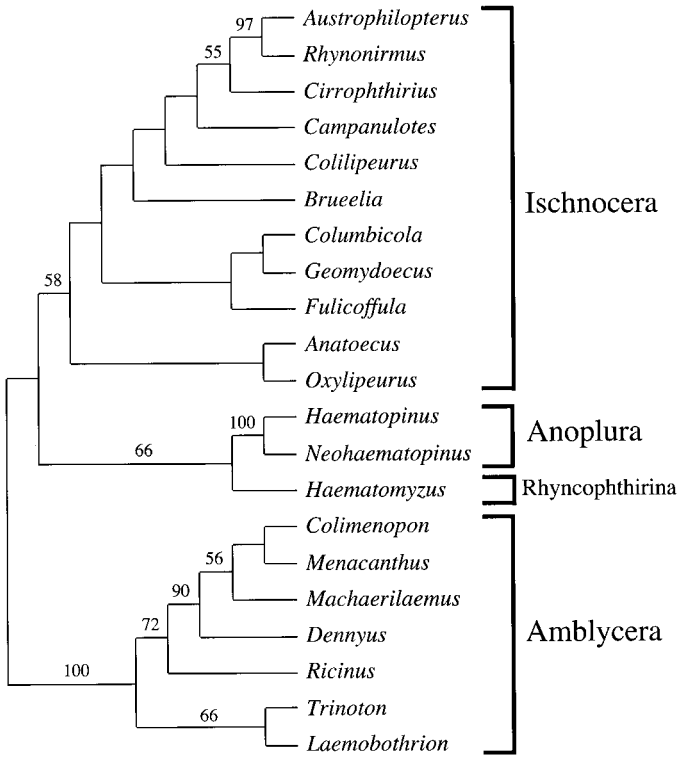

FIG. 2. Single most parsimonious tree from parsimony analysis of combined EF1α, 18S, and COI gene sequences (length = 2900, CI = 0.412). Numbers above branches indicate support from 1000 bootstrap replicates.

ported in this tree (bootstrap 58%), support for this was less than that in analysis of 18S alone.

Symmetric difference distances between trees from various parsimony analyses are given in Table 2. In general, combined analyses produced more similar trees than did independent analyses. Combined trees involving 18S were generally more similar to the 18S trees than to trees constructed from other genes. Trees from individual gene analyses showed similar differences from each other. The smallest tree to tree differences were obtained in comparisons of trees resulting from combination of two genes and the tree from combined analysis of all genes.

Maximum-likelihood analyses of separate gene regions produced a result with respect to ischnoceran monophyly similar to that of parsimony analysis (Fig. 3). One major difference between parsimony and likelihood analysis of COI was that likelihood recovered monophyly of Amblycera, whereas parsimony did not. In other respects, likelihood analyses of individual genes produced trees similar to those of the parsimony analyses. Maximum-likelihood searches of combined gene regions produced a tree (Fig. 4) that was quite similar to the combined parsimony tree. Monophyly of Amblycera, Ischnocera, and Anoplura was recovered in this tree. Differences from the parsimony tree involved rearrangements among terminal Ischnocera. Bootstrap analysis recovered nine nodes with support greater than 50%, including monophyly of Ischnocera. In comparisons of likelihood trees (Table 2), trees constructed from different gene regions were more similar to each other than similar comparisons of parsimony trees. In addition, EF1α and COI ML trees were more similar to the combined ML tree than the same comparisons in the parsimony analyses. The tree similarity between 18S ML tree and the combined ML tree was within the range of that from the parsimony analysis. Interestingly, unlike parsimony analysis, in which the COI tree was the least similar to the combined parsimony tree, the ML COI tree was the individual tree most similar to the combined ML tree.

Kishino–Hasegawa tests (Kishino and Hasegawa, 1989) involving the parsimony trees for each gene region under a likelihood model indicated that, in general, trees resulting from searches involving the gene under investigation were not significantly worse than the best tree (Table 3). The exception to this was the combined EF1α + 18S + COI tree for the EF1α data. In every case trees resulting from parsimony searches not involving the gene under consideration were significantly worse ($P < 0.05$). Interestingly, the EF1α + 18S + COI combined tree was the best tree under the model for the COI data set. Several trees, indicating monophyly of Ischnocera, are not rejected by the EF1α and COI data sets. The combined EF1α + 18S tree contains a monophyletic Ischnocera but is not rejected by a Kishino–Hasegawa test of the EF1α data. In addition, the combined 18S + COI and combined EF1α +

TABLE 2

Symmetric Difference Distance between Parsimony Trees

Tree comparison	Symmetric difference distance	
	Parsimony trees	ML trees
EF1α vs COI	34	27
EF1α vs 18S	32–34	31
COI vs 18S	35–36	30
EF1α vs COI + 18S	20	—
18S vs EF1α + COI	28–30	—
COI vs EF1α + 18S	34	—
EF1α vs EF1α + 18S	30	—
EF1α vs EF1α + COI	22–24	—
18S vs EF1α + 18S	18–24	—
18S vs 18S + COI	20–22	—
COI vs EF1α + COI	28	—
COI vs 18S + COI	28	—
EF1α vs EF1α + 18S + COI	26	25
18S vs EF1α + 18S + COI	22–24	24
COI vs EF1α + 18S + COI	28	20
EF1α + COI vs EF1α + 18S + COI	22	—
EF1α + 18S vs EF1α + 18S + COI	14	—
COI + 18S vs EF1α + 18S + COI	12	—

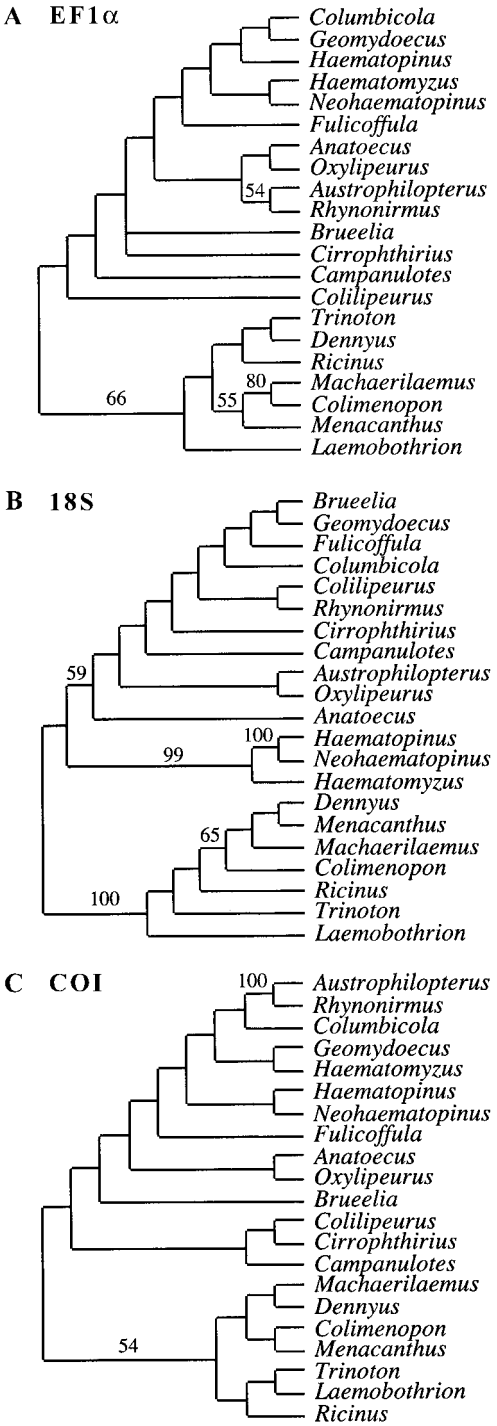

FIG. 3. Trees recovered from likelihood analysis using GTR + G model for (a) EF1α only [$-$ln L = 3173.98; A = 0.258, C = 0.273, G = 0.240, T = 0.229; A-C = 0.593, A-G = 5.097, A-T = 2.224, C-G = 0.740, C-T = 7.805, G-T = 1.0; shape parameter = 0.177], (b) 18S only [$-$ln L = 5463.60; A = 0.258, C = 0.236, G = 0.268, T = 0.237; A-C = 3.025, A-G = 3.785, A-T = 2.426, C-G = 1.113, C-T = 5.423, G-T = 1.0; shape parameter = 0.114], and (c) COI only [$-$ln L = 5245.57; A = 0.321, C = 0.120, G = 0.175, T = 0.385; A-C = 2.032, A-G = 46.03, A-T = 4.828, C-G = 43.09, C-T = 83.72, G-T = 1.0; shape parameter = 0.177]. Numbers above branches indicate bootstrap support from 100 bootstrap replicates.

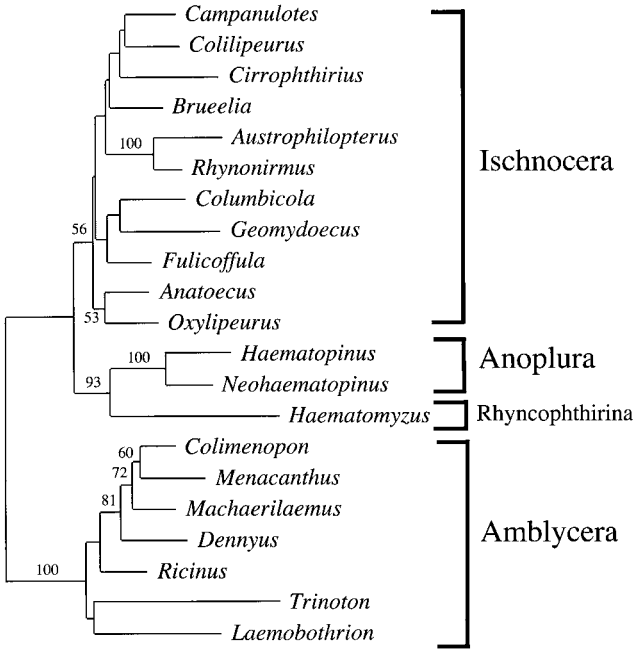

FIG. 4. Most likely tree from maximum-likelihood analysis of combined EF1α, 18S, and COI gene sequences ($-$ln Likelihood = 14,546.43). Numbers above branches indicate bootstrap support from 100 bootstrap replicates. Model includes estimated unequal base frequencies (A = 0.275, C = 0.204, G = 0.240, T = 0.230), six substitution categories (A-C = 1.565, A-G = 5.456, A-T = 3.009, C-G = 1.827, C-T = 6.653, G-T = 1.0), and rate heterogeneity according to a gamma distributions (shape parameter = 0.141) with eight rate categories.

18S + COI trees both recover monophyly of Ischnocera, but neither is rejected by the COI data. In contrast, Kishino–Hasegawa tests involving the 18S data reject all trees not containing monophyly of Ischnocera.

DISCUSSION

Phylogenetic analysis of 18S sequence data supports monophyly of the louse suborder Ischnocera, whereas analysis of either EF1α or COI data does not. Support for trees derived from these genes generally increased as homoplasy decreased. Combined analysis of all three genes supports monophyly of Ischnocera, but support for monophyly is reduced over analysis of 18S data alone. Whereas the combined evidence tree for phthirapteran taxa is completely resolved and well supported, there are important questions that arise from analyses of sequences from these three genes.

Do Gene Trees Reflect the Underlying Species Tree?

An important question when comparisons of independent phylogenetic data sets are made is whether all data sets support the same underlying phylogeny. For gene sequence data sets, this question revolves around whether gene trees reflect species trees (Pamilo and Nei, 1988; Moore, 1995; Maddison, 1997). Trees for the

TABLE 3

Kishino–Hasegawa Tests of Likelihood Models Using Parsimony Trees

Tree	−ln Likelihood	*P* value
EF1α data		
EF1α	3187.49	best tree
EF1α + COI	3199.51	0.23
†EF1α + 18S	3201.64	0.09
†EF1α + 18S + COI	3211.83	0.033*
†18S + COI	3227.99	0.0028*
†18S	3241.85	0.0004*
COI	3270.19	0.0001*
18S data		
†18S	5473.15	best tree
†EF1α + 18S	5478.43	0.61
†EF1α + 18S + COI	5480.80	0.52
†18S + COI	5486.15	0.17
COI	5971.20	0.0001*
EF1α + COI	5693.69	0.0001*
EF1α	5697.45	0.0001*
COI data		
†EF1α + 18S + COI	5271.51	best tree
EF1α + COI	5273.13	0.89
†18S + COI	5276.67	0.50
COI	5284.65	0.47
EF1α	5300.50	0.009*
†EF1α + 18S	5302.34	0.018*
†18S	5333.76	0.0001*

* Significant at the 0.05 level.
† Tree contains monophyletic Ischnocera.

same taxa derived from different genes can differ from each other simply because of uncertainties in phylogenetic reconstruction and not any real differences in the underlying phylogenies (Bull *et al.*, 1993). In general, gene trees are expected to mirror the species tree. However, there are biological processes that can lead to differences between gene trees, including hybridization (de Queiroz, 1993; Mason-Gamer and Kellogg, 1996; Johnson and Sorenson, 1999), differences in lineage sorting (Pamilo and Nei, 1988), and gene duplication (Slowinski and Page, 1999). A central problem in systematic studies involving multiple gene data sets is whether a single phylogeny underlies all data sets or whether different data sets reflect different gene histories (de Queiroz *et al.*, 1995).

A popular approach to the analysis of multiple data sets is condition combination (Bull *et al.*, 1993; de Queiroz *et al.*, 1995). Under this approach, data partitions are first assessed independently to determine whether each is consistent with the same underlying phylogeny. If this is the case, the data partitions are then combined in a single combined evidence analysis. If the data sets support differences in underlying phylogenies, the data sets are analyzed separately, but one is left with few means to arbitrate between them over the species phylogeny. In practice, the partition homogeneity test (Farris *et al.*, 1994, 1995; Swofford, 2000) is commonly employed to test for significant heterogeneity between data sets. Although advocates of conditional combination (Bull *et al.*, 1993; de Queiroz, 1993) suggest that genes should not be combined in analyses if they are significantly heterogeneous, combined analysis in the case of our study is quite instructive.

Partition homogeneity tests of the three gene regions indicated that the 18S data were significantly different from the other two genes. When significant incongruence is detected between gene regions there are several potential explanations. First, the underlying gene trees of one or more genes could be different from the species tree and different from that of other genes. Different gene trees might be caused by the vagaries of lineage sorting among different loci, such that some loci may not reflect the species tree. However, in practice we suspect that lineage sorting differences will be difficult to detect because differences between loci in the pattern of lineage sorting can occur only when the time separating speciation events is short (Lanyon, 1988; Moore, 1995). Detection of such differences at deep taxonomic scales is likely to be difficult because of multiple substitution. Lineage sorting differences among genes in old lineages would probably not result in the detection of significant incongruence.

Differences between gene trees can also arise through hybridization and the differential fixation of alternative alleles from different parental lineages. Whereas this process is certainly possible, it seems unlikely that hybridization could be detected (for the same reason as the lineage sorting problem) unless it occurred between distantly related taxa.

Gene trees may also differ because of gene duplication, such that paralogous genes are sequenced rather than homologs. Base composition and substitution properties for our COI sequences suggest that these sequences are indeed mitochondrial in origin rather than a nuclear duplicate, as has been found in some vertebrate taxa (Sorenson and Fleischer, 1996). A duplication of the EF1α gene is known from Hymenoptera (Danforth and Ji, 1998) and Diptera (Hovemann *et al.*, 1998), but this gene is believed to be in single copy in most other insects. The 18S gene is present in many copies in the nuclear genome, but this gene generally undergoes concerted evolution such that most copies are converted to a single sequence (Hillis and Dixon, 1991). We also consider it unlikely that 18S would have undergone events in molecular evolution such that it just coincidentally matched traditional taxonomy on the question of ischnoceran monophyly. In addition to these considerations, incongruence between gene trees resulting from duplication events should be evidenced by a conflict in bootstrap topologies. This was not the case for trees constructed from these three genes independently.

A second possible explanation for significant incongruence between 18S and the two protein-coding genes

is alignment error. Alignment decisions can potentially alter the signal and phylogeny inferred from the alignment (Lutzoni, 1997). Most cases of significant differences between gene regions detected by the partition homogeneity tests involve genes that require alignment (e.g., structural genes) (Sullivan, 1996; Poe, 1996; Lutzoni, 1997; Whiting *et al.,* 1997). This is the case with our data sets also. However, we took care to exclude from our analyses those regions of 18S whose alignment was not straightforward. Thus, we are confident that differences in phylogenies are not the result of alignment decisions.

A final possibility is that rather than detecting a difference in underlying phylogenetic signal, the partition homogeneity test has detected rate heterogeneity among partitions. Recent simulation studies of the partition homogeneity test (Dolphin *et al.,* 2000; Barker and Lutzoni, 2001) indicate that when differences in substitution rates in gene regions are dramatic, a significant result may be achieved even when the underlying phylogeny is the same. In these cases, differences between trees constructed from different gene regions will generally not receive strong bootstrap support. Because 18S has a substantially slower substitution rate than EF1α or COI, rate differences may have caused the significant partition homogeneity test. No nodes that conflicted between gene regions did so with support $>$60% for any two gene regions, again suggesting that rate differences, and not real differences in underlying phylogeny, were responsible for the significant partition homogeneity test results.

The partition homogeneity test operates in a parsimony framework; thus, methods that take into account rate differences might be able to illuminate whether such heterogeneity might be responsible. Likelihood analyses of each gene region separately did produce more similar trees than did parsimony analysis of each gene separately. In addition, Kishino–Hasegawa tests indicated that some topologies containing Ischnocera monophyly could not be ruled out with EF1α or COI data, whereas 18S rejected topologies without Ischnocera monophyly. In addition, a parsimony tree combining all data received a higher likelihood score under the COI data than did a parsimony tree constructed from COI alone. Taking all this evidence together, we suggest that the incongruence between genes is an artifact of dramatically different substitution rates and nodes differing with poor support, rather than reflecting real underlying differences in gene trees.

Is Ischnocera Monophyletic?

Combined analysis of all gene regions under either parsimony or likelihood frameworks supports monophyly of Ischnocera, although weakly (bootstrap $<$60%). Parsimony analysis of 18S data alone strongly supports monophyly of Ischnocera (bootstrap $>$90%). Given concerns over the congruence between gene regions, what confidence can be placed on ischnoceran monophyly? Combination of either EF1α or COI with 18S in parsimony analysis reduces bootstrap support for ischnoceran monophyly compared to bootstrap analysis of 18S alone. Typically, combination of data sets is expected to increase bootstrap support (Johnson and Sorenson, 1998; Johnson and Lanyon, 1999; DeFillipis and Moore, 2000). For example, support for a sister relationship between *Austrophilopterus* and *Rhynonirmus* increases when more data are combined, even though this relationship is not recovered by analysis of 18S alone. Combination of 18S with either EF1α or COI sequences involves combination of the less homoplasious data set with a more homoplasious data set. Combining data sets of drastically different levels of homoplasy can result in lowered overall bootstrap support compared to the least homplasious data set (Johnson and Clayton, 2000). Bootstrap values underestimate confidence levels when homoplasy is high (Zarkikh and Li, 1992, 1995), so this is probably at least part of the explanation for the reduced support. Bootstrap support may also be reduced when genes of different signal are combined (Mason-Gamer and Kellogg, 1996). However, given that Kishino–Hasegawa tests of EF1α or COI alone do not reject ischnoceran monophyly, we suggest that increased homoplasy is the most likely explanation for reduced support in combined analysis.

Further support for combination of data sets comes from examination of tree similarities between separate analyses and combined analyses. In almost all cases, combination of gene regions makes trees more similar than separate analyses (Table 2). For example, comparisons of trees derived from a single gene region are more similar to trees based on combination of the other two gene regions than are comparisons of trees based on single genes to each other (Table 2). Thus, combination of data improves congruence between gene regions. In addition, maximum-likelihood analysis, which takes into account rate differences, produces trees that are more similar between gene regions than does parsimony analysis. This result suggests that differences between trees in unordered parsimony analyses (and the corresponding partition homogeneity tests) may result more from underlying differences in substitution properties than from any real difference in underlying phylogenies.

CONCLUSIONS

Whereas EF1α and COI do recover some relationships with strong support among major lineages of Phthiraptera, generally, support is weak for deeper arrangements. In addition, trees based on these two genes do not agree with strongly supported relationships identified by analyses of 18S. Both EF1α and COI have relatively high substitution rates compared to

18S and the consequent multiple substitutions are likely to give rise to problems of long branch attraction. However, inclusion of these sequences in combination with 18S still provides support for relationships strongly supported by 18S. In addition, combination of gene regions resolves relationships not resolved in analyses of 18S by itself. Although highly homoplasious genes are not likely to retain strong signal for deep phylogenetic relationships, these genes do still retain an ability to resolve relationships at the terminal portions of the tree.

ACKNOWLEDGMENTS

The following individuals assisted in providing samples for this project: L. Durden, R. Moyle, J. Weckstein, R. J. Adams, E. Osnas, S. Meagher, D. Clayton, and S. Al-tamimi. Jared Helm provided assistance in sequencing 18S. R. J. Adams prepared slide mounts of voucher specimens and assisted with the identifications. R. DeSalle, M. Hardman, V. Smith, C. Taylor, and two anonymous reviewers provided helpful comments on the manuscript. This work was supported in part by NSF CAREER DEB-9703003 to D. H. Clayton and by NSF Grants DEB-9615269, DEB-9806349, and NSF CAREER award DEB-9983195 to M. F. Whiting. We thank the DNA Sequencing Facility at the University of Utah, supported in part by NCI Grant 5p30CA42014.

REFERENCES

Archie, J. W. (1989). A randomization test for phylogenetic information in systematic data. *Syst. Zool.* **38:** 239–252.

Barker, F. K., and Lutzoni, F. M. (2001). Spurious rejection of phylogenetic congruence by the ILD test: A simulation study. *Syst. Biol.*, in press.

Barker, S. C. (1994), Phylogeny and classification, origins, and evolution of host associations of lice. *Int. J. Parasitol.* **24:** 1285–1291.

Bull, J. J., Huelsenbeck, J. P., Cunningham, C. W., Swofford, D. L., and Waddell, P. J. (1993). Partitioning and combining data in phylogenetic analysis. *Syst. Biol.* **42:** 384–397.

Clay, T. (1970). The Amblycera (Phthiraptera: Insecta). *Bull. Br. Mus.* (*Nat. Hist. Entomol.)* **24:** 75–98.

Cruickshank, R. H., Johnson, K. P., Smith, V. S., Adams, R. J., Clayton, D. H., and Page, R. D. M. (2001). Phylogenetic analysis of elongation factor 1 alpha identifies major groups of lice (Insecta: Phthiraptera). *Mol. Phylogenet. Evol.* **19:** 202–215.

Danforth, B. N., and Ji, S. (1998). Elongation factor-1α occurs as two copies in bees: Implications for phylogenetic analysis of EF-1α sequences in insects. *Mol. Biol. Evol.* **15:** 225–235.

DeFilippis, V. R., and Moore, W. S. (2000). Resolution of phylogenetic relationships among recently evolved species as a function of amount of DNA sequence: An empirical study based on woodpeckers (Aves: Picidae). *Mol. Phylogenet. Evol.* **16:** 143–160.

de Queiroz, A. (1993). For consensus (sometimes). *Syst. Biol.* **42:** 368–372.

de Queiroz, A., Donoghue, M. J., and Kim, J. (1995). Separate versus combined analysis of phylogenetic evidence. *Annu. Rev. Ecol. Syst.* **26:** 657–681.

Dolphin, K., Belshaw, R., Orme, C. D. L., and Quicke D. L. J. (2000). Noise and incongruence: Interpreting results of the incongruence length difference test. *Mol. Phylogenet. Evol.* **17:** 401–406.

Eichler, W. D. (1941). Zur Klassifikation der Lauskerfe (Phthiraptera Haeckel: Rhyncophthirina, Mallophaga und Anoplura). *Arch. Naturgesch. (Leipzig) N. F.* **10:** 345–398.

Eichler, W. D. (1963). Phthiraptera. 1. Mallophaga. *Bronns Klassen Ordnungen Tierreichs.* Akademische Verlagsgesellschaft, Geest and Portig K.-G. **5,** (3), **7,** Leipzig.

Faith, D. P., and Cranston, P. S. (1991). Could a cladogram this short have arisen by chance alone?: On permutation tests for cladistic structure. *Cladistics* **7:** 1–28.

Farris, J. S., Kallersjo, M., Kluge, A. G., and Bult, C. (1994). Testing significance of congruence. *Cladistics* **10:** 315–320.

Farris, J. S., Kallersjo, M., Kluge, A. G., and Bult, C. (1995). Constructing a significance test for incongruence. *Syst. Biol.* **44:** 570–572.

Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **39:** 783–791.

Ferris, G. F. (1951). The Sucking Lice. *Mem. Pacif. Cst. Entomol. Soc.*

Hafner, M. S., Sudman, P. D., Villablance, F. X., Spradling, T. A., Demastes, J. W., and Nadler, S. A. (1994). Disparate rates of molecular evolution in cospeciating hosts and parasites. *Science* **365:** 1087–1090.

Hillis, D. M., and Dixon, M. T. (1991). Ribosomal DNA: Molecular evolution and phylogenetic inference. *Q. Rev. Biol.* **66:** 411–453.

Hopkins, G. H. E., and Clay, T. (1952). "A Checklist of the Genera and Species of Mallophaga," Br. Mus. Nat. Hist., London.

Hovemann, B., Richter, S., Walldorf, U., and Cziepluch, C. (1998). Two genes encode related cytoplasmic elongation factors 1 alpha (EF-1 alpha) in *Drosophila melanogaster* with continuous and stage specific expression. *Nucleic Acids Res.* **16:** 3175–3194.

Huelsenbeck, J. P., and Crandall, K. A. (1997). Phylogeny estimation and hypothesis testing using maximum likelihood. *Annu. Rev. Ecol. Syst.* **28:** 437–466.

Johnson, K. P., Adams, R. J., and Clayton, D. H. (2001). Molecular systematics of the Goniodidae (Insecta: Phthiraptera). *J. Parasitol.* **87:**

Johnson, K. P., and Clayton, D. H. (2000). Nuclear and mitochondrial genes contain similar phylogenetic signal for pigeons and doves (Aves: Columbiformes). *Mol. Phylogenet. Evol.* **14:** 141–151.

Johnson, K. P., and Clayton, D. H. (2001). Coevolutionary history of ecological replicates: Comparing phylogenies of wing and body lice to Columbiform hosts. *In* "Tangled Trees: Phylogeny, Cospeciation, and Coevolution (R. D. M. Page, Ed.). Univ. of Chicago Press, Chicago, in press.

Johnson, K. P., and Lanyon, S. M. (1999). Molecular systematics of the grackles and allies, and the effect of additional sequence (cyt *b* and ND2). *Auk* **116:** 759–768.

Johnson, K. P., and Sorenson, M. D. (1998). Comparing molecular evolution in two mitochondrial protein coding genes (cytochrome *b* and ND2) in the dabbling ducks (Tribe: Anatini). *Mol. Phylogenet. Evol.* **10:** 82–94.

Johnson, K. P., and Sorenson, M. D. (1999). Phylogeny and biogeography of dabbling ducks (Genus: *Anas*): A comparison of molecular and morphological evidence. *Auk* **116:** 792–805.

Kim, K. C., and Ludwig, H. W. (1982). Parallel evolution, cladistics, and classification of parasitic Psocodea. *Ann. Entomol. Soc. Am.* **75:** 537–548.

Kishino, H., and Hasegawa, M. (1989). Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. *J. Mol. Evol.* **29:** 170–179.

Lanyon, S. M. (1988). The stochastic mode of molecular evolution: What consequences for systematic investigations? *Auk* **105:** 565–573.

Lutzoni, F. M. (1997). Phylogeny of lichen- and non-lichen-forming

Omphalinoid mushrooms and the utility of testing for combinability among multiple data sets. *Syst. Biol.* **46:** 373–406.

Lyal, C. H. C. (1985). Phylogeny and classification of Psocodea, with particular reference to lice (Psocodea: Phthiraptera). *Syst. Entomol.* **10:** 145–165.

Maddison, W. P. (1997). Gene trees in species trees. *Syst. Biol.* **46:** 523–536.

Marshall, A. G. (1981). "The Ecology of Ectoparasitic Insects," Academic Press, London.

Mason-Gamer, R. J., and Kellogg, E. A. (1996). Testing for phylogenetic conflict among molecular data sets in the tribe Triticeae (Gramineae). *Syst. Biol.* **45:** 524–545.

Moore, W. S. (1995). Inferring phylogenies from mtDNA variation: Mitochondrial-gene trees versus nuclear-gene trees. *Evolution* **49:** 718–726.

Pamilo, P., and Nei, M. (1988). Relationships between gene trees and species trees. *Mol. Biol. Evol.* **5:** 568–583.

Paterson, A. M., Wallis, G. P., Wallis, L. J., and Gray, R. D. (2000). Seabird and louse coevolution: Complex histories revealed by 12S rRNA sequences and reconciliation analysis. *Syst. Biol.* **49:** 383–399.

Penny, D., and Hendy, M. D. (1985). The use of tree comparison metrics. *Syst. Zool.* **34:** 75–82.

Poe, S. (1996). Data set incongruence and the phylogeny of crocodilians. *Syst. Biol.* **45:** 393–414.

Slowinski, J. B., and Page, R. D. M. (1999). How should species phylogenies be inferred from sequence data? *Syst. Biol.* **48:** 814–825.

Smith, V. S. (2001). Avian louse phylogeny (Phthiraptera: Ischnocera): A cladistic study based on morphology. *Zool. J. Linn. Soc.* **132:** 81–144.

Sorenson, M. D., and Fleishcer, R. C. (1996). Multiple independent transpositions of mitochondrial DNA control region sequences to the nucleus. *Proc. Natl. Acad. Sci. USA* **93:** 15239–15243.

Sullivan, J. (1996). Combining data with different distributions of among-site rate variation. *Syst. Biol.* **45:** 375–380.

Swofford, D. L. (2000). PAUP*: Phylogenetic analysis using parsimony, version 4.0, beta. Sinauer, Sunderland, MA.

Whiting, M. F., Carpenter, J. C., Wheeler, Q. D., and Wheeler, W. C. (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.

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

Zarkikh, A., and Li, W-H. (1995). Estimation of confidence in phylogeny: The complete-and-partial bootstrap technique. *Mol. Phylogenet. Evol.* **4:** 44–63.