Zoological Journal of the Linnean Society, 2006, **146**, ••–••. With 3 figures

Molecular systematics of the suborder Trogiomorpha (Insecta: Psocodea: 'Psocoptera')

KAZUNORI YOSHIZAWA[1]*, CHARLES LIENHARD[2] and KEVIN P. JOHNSON[3]

[1]*Systematic Entomology, Graduate School of Agriculture, Hokkaido University, Sapporo 060-8589, Japan*
[2]*Natural History Museum, c.p. 6434, CH-1211, Geneva 6, Switzerland*
[3]*Illinois Natural History Survey, 607 East Peabody Drive, Champaign, IL 61820, USA*

Received March 2005; accepted for publication July 2005

1 Phylogenetic relationships among extant families in the suborder Trogiomorpha (Insecta: Psocodea: 'Psocoptera') were inferred from partial sequences of the nuclear 18S rRNA and Histone 3 and mitochondrial 16S rRNA genes. Analyses of these data produced trees that largely supported the traditional classification; however, monophyly of the infraorder Psocathropetae (= Psyllipsocidae + Prionoglarididae) was not recovered. Instead, the family Psyllipsocidae was recovered as the sister taxon to the infraorder Atropetae (= Lepidopsocidae + Trogiidae + Psoquillidae), and the Prionoglarididae was recovered as sister to all other families in the suborder. Character states previously used to diagnose Psocathropetae are shown to be plesiomorphic. The sister group relationship between Psyllipsocidae and Atropetae was supported by two morphological apomorphies: the presence of a paraproctal anal spine and an anteriorly opened phallosome. Based on these sequence data and morphological observations, we propose a new classification scheme for the Trogiomorpha as follows: infraorder Prionoglaridetae (Prionoglarididae), infraorder Psyllipsocetae (Psyllipsocidae), infraorder Atropetae (Lepidopsocidae, Trogiidae, Psoquillidae). © 2006 The Linnean Society of London, *Zoological Journal of the Linnean Society*, 2006, **146**, 000–000.

ADDITIONAL KEYWORDS: 16S rRNA gene –18S rRNA gene – Histone 3 – morphology – phylogeny.

INTRODUCTION

The insect order Psocodea, containing over 10 000 described species, includes parasitic lice of birds and mammals (Phthiraptera) as well as free-living book lice and bark lice (Psocoptera). Psocodea is closely related to other hemipteroid orders (Hemiptera and Thysanoptera), which together with them comprise the group Paraneoptera (Yoshizawa & Saigusa, 2001). Phthiraptera and Psocoptera have previously been treated as two separate orders (Johnson, Yoshizawa & Smith, 2004), but morphological and molecular evidence reveals that the Phthiraptera is imbedded within the Psocoptera (Lyal, 1985; Yoshizawa & Johnson, 2003, 2005; Johnson *et al*., 2004), making the Psocoptera paraphyletic. Seven suborders are now generally recognized within the Psocodea: Trogiomorpha (bark lice), Psocomorpha (bark lice), Troctomorpha (book lice and bark lice), Amblycera (chewing lice), Ischnocera (chewing lice), Rhynchophthirina (chewing lice) and Anoplura (sucking lice). The phylogenetic relationships within several of the larger suborders have received recent attention. For example, phylogenies based on morphological and/or molecular data have been produced for Psocomorpha (Yoshizawa, 2002; Johnson & Mockford, 2003), Amblycera (Marshall, 2003) and Ischnocera (Smith, 2001; Smith, Page & Johnson, 2004). These studies, when combined with higher level results (Lyal, 1985; Barker *et al*., 2003; Johnson *et al*., 2004; Yoshizawa & Johnson, 2005), begin to provide a framework for the phylogenetic tree and classification of the Psocodea.

However, one suborder of Psocodea, the Trogiomorpha, containing over 340 described species, has received little phylogenetic attention, and the higher level classification of this suborder has not been well tested. The only formal phylogenetic analysis conducted to date (Perrichot *et al*., 2004) utilized relatively few characters, and the resulting tree is almost completely unresolved (Perrichot *et al*., 2004). Trogio-

*Corresponding author. E-mail: psocid@res.agr.hokudai.ac.jp

morpha is generally recognized as the sister taxon of the remaining Psocodea (Smithers, 1972; Vishnyakova, 1980; Lienhard, 1998; Johnson *et al*., 2004), and therefore understanding the phylogenetic relationships of this group is important for understanding the origin and evolution of Psocodea as a whole.

Trogiomorpha was first established as a group by Roesler (1940). Before this, Pearman (1936) recognized some family groups (= infraorders in the present sense) within 'Psocoptera', and two of these, Atropetae (including Lepidopsocidae, Trogiidae and Psoquillidae) and Psocathropetae (including Psyllipsocidae and Prionoglarididae), were assigned to Trogiomorpha by Roesler (1940, 1944) (Table 1). Pearman's classification system was roughly adopted by Roesler (1944), but Roesler united some of Pearman's families (Trogiidae and Psoquillidae into Trogiidae; Psyllipsocidae and Prionoglarididae into Psyllipsocidae). The most widely accepted current higher level classification of Trogiomorpha follows Badonnel (1951), who basically accepted Pearman's families and infraorders and assigned them to Roesler's suborder (Lienhard & Smithers, 2002). Smithers (1972) proposed a very different taxonomic classification for the Trogiomorpha, but this classification has not been accepted in any subsequent work, except by Li (2002), whose classification is largely based on Smithers (1972) (see Table 1).

Although the Trogiomorpha is smaller than the other suborders of Psocodea, and its higher level classification has not been altered for a long time, the present classification needs to be evaluated by formal phylogenetic analyses. In particular, the suborder has been diagnosed mostly based on plesiomorphic features, such as more than 20-segmented antenna, two-segmented labial palpus and three-segmented tarsi (Smithers, 1972; Mockford, 1993; Lienhard, 1998), and no definite apomorphic character has been proposed for it. Smithers (1972: 280) has even pointed out the possibility that the Psocathropetae (Psyllipsocidae + Prionoglarididae) may not be a member of the suborder but may be a sister group to all other Psocodea. Monophyly of the two trogiomorphan infraorders (Psocathropetae and Atropetae) is also poorly established, as shown by Perrichot *et al*. (2004).

In the present paper, we estimate the phylogenetic relationships among the extant families of Trogiomorpha based on the nuclear 18S rRNA and Histone 3 gene and mitochondrial 16S rRNA gene sequences. We also discuss morphological characters that are consistent with the trees resulting from the molecular data.

MATERIAL AND METHODS

Samples were selected from all extant families of the Trogiomorpha. Outgroups were selected from the suborder Psocomorpha, the infraorder Amphientometae of the suborder Troctomorpha, and the order Thysanoptera (root) (Table 2). The presence of very long insertions/deletions (indels) has been shown in the 18S and 16S subunits of lice and their relatives (the infraorder Nanopsocetae of Troctomorpha) (Yoshizawa & Johnson, 2003; Johnson *et al*., 2004), which made alignment exceedingly difficult and introduced taxa with very long branches. Therefore, no exemplars were selected from either lice or Nanopsocetae.

Partial sequences of the nuclear 18S rRNA and Histone 3 and mitochondrial 16S rRNA genes were used for analyses. Methods for DNA extraction, PCR amplification and sequencing follow Johnson & Mockford (2003) and Johnson *et al*. (2004). Primer sets used were Ns1–Ns2a, 18Sai–18Sbi and Ns5a–Ns8P for 18S (Johnson *et al*., 2004), H3AF–H3AR for Histone 3 (Colgan *et al*., 1998), and 16Sar–16Sbr for 16S (Simon *et al*., 1994). Alignment of the protein-coding Histone 3 gene fragment was straightforward. Ribosomal DNA was aligned manually according to secondary structure models provided by Kjer (2004) for 18S and Buckley *et al*. (2000) for 16S. Indels were observed in some loop regions, and we were unable to align some of these regions confidently. Therefore, these highly variable regions were excluded from the analyses. For a few samples, we were unable successfully to amplify some genes or gene fragments because of the degraded quality of the material (generally old museum specimens of rare genera stored in 70% ethanol, Table 2). Thus, we prepared two data sets, one including and one excluding taxa with missing data. We performed the partition homogeneity test (Farris *et al*., 1994, 1995) with 1000 replicates [tree branch replication (TBR) heuristic search with ten random additions for each replicate] to compare the homogeneity of each data partition using PAUP* 4.0b10 (Swofford, 2002). We also compared parsimony [maximum-parsimony (MP)] bootstrap consensus trees estimated from each data partition to see whether any heterogeneity between data sets was reflected in strongly conflicting bootstrap topologies. Taxa with missing data were excluded from the partition homogeneity test and partitioned MP bootstrapping. The aligned data set is available online at http://insect3.agr.hokudai.ac.jp/psoco-web/data/index.html.

For both data sets, MP, maximum-likelihood (ML) and Bayesian analyses were conducted. For MP analysis, all characters were weighted equally. The substitution model for ML and Bayesian analyses was estimated using likelihood ratio tests as implemented in Modeltest 3.5 (Posada & Crandall, 1998). These tests supported use of the GTR+I+G model (unequal base frequencies: A = 0.2853, C = 0.1888, G = 0.2449, T = 0.2810; six substitution categories: AC = 0.9952, AG = 3.3787, AT = 2.9503, CG = 0.8559, CT = 4.7871,

Table 1. History of higher classification of Trogiomorpha. Fossil families are given in parentheses and marked with † (known from Cretaceous and Baltic amber, not treated in the present paper). Note: spelling of some names does not correspond to original spellings but is corrected according to ICZN spelling rules

Pearman (1936)
Group Atropetae
Family Lepidopsocidae
Family Atropidae (= Trogiidae)
Family Psoquillidae
Group Psocathropetae
[intermediate category] Psocathropida
Family Psocathropidae (= Psyllipsocidae)
[intermediate category] Scoliopsocida
Family Scoliopsyllopsididae (= Prionoglarididae)

Roesler (1944)
Suborder Trogiomorpha
Group Atropetae
Family Lepidopsocidae
Family Trogiidae
Group Psocathropetae
Family Psyllipsocidae

Badonnel (1951)
Suborder Trogiomorpha
Group Atropetae
Family Lepidopsocidae
Family Trogiidae
Family Psoquillidae
Group Psocathropetae
Family Psyllipsocidae
Family Prionoglarididae

Smithers (1972)
Suborder Trogiomorpha
Division Trogioformia
Group Perientometae
Superfamily Thylacelloidea
Family Thylacellidae
Superfamily Perientomoidea
Family Lepidopsocidae
Family Perientomidae
Group Trogietae
Superfamily Trogioidea
Family Anomocopeidae
Family Trogiidae
Superfamily Psoquilloidea
Family Psoquillidae
[Family Empheriidae]†
Division Psyllipsociformia
Group Prionoglaridetae
Superfamily Prionoglaridoidea
Family Prionoglarididae
Group Psyllipsocetae
Superfamily Speleketoroidea
Family Speleketoridae
Superfamily Psyllipsocoidea
Family Psyllipsocidae
Family Psocathropidae

Li Fasheng (2002: table 1)
Suborder Trogiomorpha
Superfamily Thylacelloidea
Family Thylacellidae
Superfamily Perientomoidea
Family Lepidopsocidae
Family Perientomidae
Superfamily Trogioidea
Family Anomocopeidae
Family Trogiidae
Superfamily Psoquilloidea
Family Psoquillidae
[Family Empheriidae]
Superfamily Prionoglaridoidea
Family Prionoglarididae
Superfamily Speleketoroidea
Family Speleketoridae
Superfamily Psyllipsocoidea
Family Psyllipsocidae
Family Psocathropidae

Lienhard & Smithers (2002)
Suborder Trogiomorpha
Infraorder Atropetae
[Family Archaeatropidae]†
[Family Empheriidae]†
Family Lepidopsocidae
Family Trogiidae
Family Psoquillidae
Infraorder Psocathropetae
Family Psyllipsocidae
Family Prionoglarididae

Yoshizawa, Lienhard, & Johnson (present paper, only extant families treated)
Suborder Trogiomorpha
Infraorder Prionoglaridetae
Family Prionoglarididae
Infraorder Psyllipsocetae
Family Psyllipsocidae
Infraorder Atropetae
Family Lepidopsocidae
Family Psoquillidae
Family Trogiidae

GT = 1; gamma distribution shape parameter = 0.5168; proportion of invariant sites = 0.4881; four rate categories). We used a heuristic algorithm with TBR branch swapping (100 replicates of random addition) to search for MP and ML trees. Bootstrap analyses were performed using 100 replicates of TBR branch swapping (with a neighbour-joining starting tree for ML bootstrapping). These searches were per-

Table 2. Taxa included in the study (family assignment of Psocodea according to Lienhard & Smithers, 2002)

ORDER: SUBORDER Family	Species	Locality	GenBank accession no. (18S/H3/16S)
PSOCODEA: TROGIOMORPHA			
Prionoglarididae	*Prionoglaris* sp.	Greece	AY630456/DQ104773/DQ104745
Prionoglarididae	*Siamoglaris zebrina* Lienhard, 2004	Thailand	DQ104798/missing/DQ104746
Prionoglarididae	*Speleketor irwini* Mockford, 1984	USA	DQ104799/DQ104774/DQ104747
Prionoglarididae	*Sensitibilla strinatii* Lienhard, 2000	Namibia	DQ104800/DQ104775/missing
Psyllipsocidae	*Dorypteryx domestica* Smithers, 1958	Czech	AY630454/DQ104777/DQ104749
Psyllipsocidae	*Psyllipsocus oculatus* Gurney, 1943	Mexico	AY630455/DQ104776/DQ104748
Lepidopsocidae	*Echmepteryx hageni* (Packard, 1870)	USA	AY630448/DQ104782/DQ104754
Lepidopsocidae	*E. madagascariensis* (Kolbe, 1885)	Japan	AY630447/DQ104781/DQ104753
Lepidopsocidae	*Lepium* sp.	PNG	AY630450/DQ104785/DQ104758
Lepidopsocidae	*Neolepolepis occidentalis* (Mockford, 1955)	USA	AY630446/DQ104779/DQ104751
Lepidopsocidae	*Pteroxanium kelloggi* (Ribaga, 1905)	USA	AY630449/DQ104784/DQ104757
Lepidopsocidae	*Soa* sp.	PNG	DQ104802/DQ104780/DQ104752
Psoquillidae	*Rhyopsocus* sp.	USA	DQ104801/DQ104778/DQ104750
Trogiidae	*Cerobasis alpha* García Aldrete, 1986	USA	DQ104803-4(partly missing)/ DQ104787/DQ104760
Trogiidae	*Lepinotus reticulatus* Enderlein, 1905	UK	AY630452/missing/DQ104756
Trogiidae	*Lepinotus* sp.	USA	AY630451/DQ104783/DQ104755
Trogiidae	*Trogium pulsatorium* (Linnaeus, 1758)	UK	AY630453/DQ104786/DQ104759
PSOCODEA: TROCTOMORPHA			
Compsocidae	*Compsocus elegans* (Banks, 1930)	Costa Rica	AY630462/DQ104790/DQ104763
Musapsocidae	*Musapsocus* sp.	Mexico	AY630461/DQ104789/DQ104762
Troctopsocidae	*Selenopsocus* sp.	Malaysia	AY630457/DQ104788/DQ104761
PSOCODEA: PSOCOMORPHA			
Archipsocidae	*Archipsocus* sp.	Malaysia	AY630478/DQ104791/DQ104764
Epipsocidae	*Bertkauia crosbyana* Chapman, 1930	USA	AY630537/DQ104793/DQ104766
Dasydemellidae	*Matsumuraiella radiopicta* Enderlein, 1906	Japan	AY630493/DQ104797/DQ104770
Pseudocaeciliidae	*Heterocaecilius fuscus* Yoshizawa, 1996	Japan	AY630520/DQ104795/DQ104768
Philotarsidae	*Aaroniella badonneli* (Danks, 1950)	USA	AY630532/DQ104796/DQ104769
Mesopsocidae	*Mesopsocus hongkongensis* Thornton, 1959	Japan	AY630516/DQ104794/DQ104767
Hemipsocidae	*Hemipsocus* sp.	Malaysia	AY630543/DQ104792/DQ104765
THYSANOPTERA: TEREBRANTIA			
Aeolothripidae	*Franklinothrips vespiformis* Crawford, 1909	USA	AY630444/DQ104772/DQ104744
Thripidae	*Frankliniella* sp.	USA	AY630445/DQ104771/DQ104743

formed using PAUP* 4.0b10 (Swofford, 2002). Partitioned Bremer support values (Bremer, 1988; Baker & DeSalle, 1997; Baker, Yu & DeSalle, 1998) for the three gene fragments were calculated using TreeRot (Sorenson, 1999). The partitioned Bremer support values were only calculated for the tree obtained from the data set excluding taxa with missing data. Bayesian analyses were conducted using MrBayes 3.1 (Huelsenbeck & Ronquist, 2001; Ronquist & Huelsenbeck, 2003). For each data set, we ran two analyses with four chains for 2000 000 generations, and trees were sampled every 1000 generations. For Bayesian analyses, the ML score of the tree was generally stable after 20 000 generations, so we discarded the first 20 trees as burn-in. We computed a majority rule consensus tree of the remaining 1980 trees to estimate the posterior probability for nodes in the tree.

RESULTS

Data evaluation

Significant heterogeneity was not detected between 18S and 16S ($P = 0.24$) using the partition homogeneity test. However, significant heterogeneity was detected between Histone 3 and the ribosomal genes ($P = 0.001$ for Histone 3 vs. 16S, 18S, 16S + 18S). Comparisons of uncorrected pairwise distances (unaligned regions excluded) showed that 18S evolves substantially slower than 16S and Histone 3. Although the

maximum pairwise divergence for the 16S rRNA gene (33% between *Selenopsocus* sp. and *Frankliniella* sp.) was larger than the maximum for Histone 3 (26% between *Speleketor irwini* and *Frankliniella* sp.), the minimum pairwise divergence between any taxa for Histone 3 (4.5%) was larger than that for 16S (3.6% between *Pteroxanium kelloggi* and *Echmepteryx madagascariensis*). Histone 3 showed more evidence of multiple substitution when divergences were compared with 18S (Fig. 1A) than did the 16S rRNA gene (Fig. 1B). The plot of pairwise divergences for Histone 3 against 18S levelled off after about 25% Histone 3 divergence (Fig. 1A), whereas no such levelling was evident in the plot of 16S divergence against 18S divergence (Fig. 1B).

These differences were also detected in the level of homoplasy for each gene. The consistency (CI) and retention (RI) indices for Histone 3 (CI = 0.282; RI = 0.309) were low compared with those for 16S (CI = 0.453; RI = 0.475), which were lower than those for 18S (CI = 0.628; RI = 0.696) (unaligned regions excluded). Comparisons of the MP bootstrap consensus trees from each gene analysed separately revealed that 18S and 16S data produced rather well-resolved and congruent trees, but the tree resulting from Histone 3 was very poorly supported, except for a few weakly supported shallow clades (e.g. *Echmepteryx madagascariensis* + *Pteroxanium kelloggi* and monophyly of Thysanoptera; trees not shown). These results suggest that the heterogeneities between Histone 3 and ribosomal genes were probably due to different evolutionary rates and the resulting differences in underlying homoplasy, rather than different underlying phylogenetic signal (Dolphin *et al*., 2000; Barker & Lutzoni, 2002). This effect has been shown for other data sets involving comparisons of rapidly and slowly evolving genes (Johnson *et al*., 2002). Therefore, in the following analyses, we combined all data partitions into a single data matrix.

Phylogenetic analyses

All analyses based on the two data sets (including and excluding taxa with missing data) produced very similar trees (Figs 2, 3). Monophyly of the Trogiomorpha was recovered by all analyses and this result received very high support from Bayesian posterior probability (100/99% when taxa with missing data were included/excluded) and ML boostrapping (91/84%), although support from MP bootstraping was weaker (66/58%).

When taxa with missing data were included in the analyses, monophyly of the family Prionoglarididae was not recovered by ML or Bayesian analyses (indicated by the dotted line in Fig. 3). MP analysis recovered monophyly of Prionoglarididae, but placed this family as the sister of the Trogiidae + Psoquillidae clade, which is in conflict with all other results, including morphological characters (discussed below). However, when the four taxa with missing data were excluded from the analyses (Fig. 2), monophyly of Prionoglarididae was also recovered by ML and Bayesian

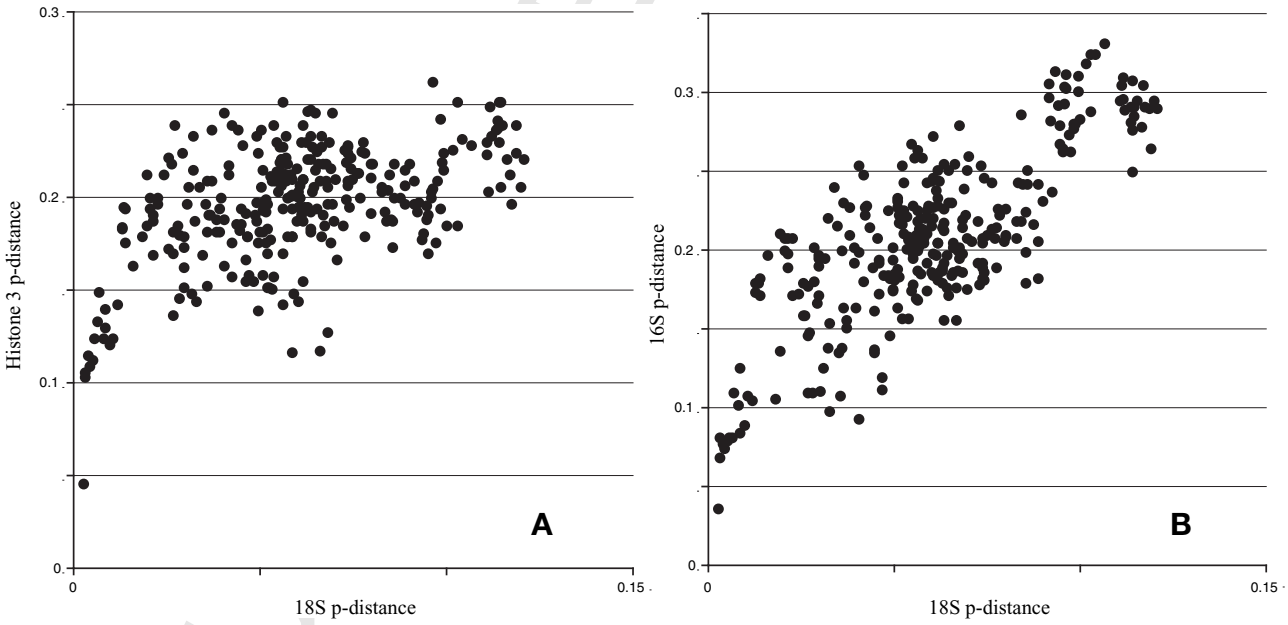

Figure 1. Plot of uncorrected pairwise distance (p-distance) in 18S vs. Histone 3 (A) and 18S vs. 16S (B). Taxa with missing data are not included.

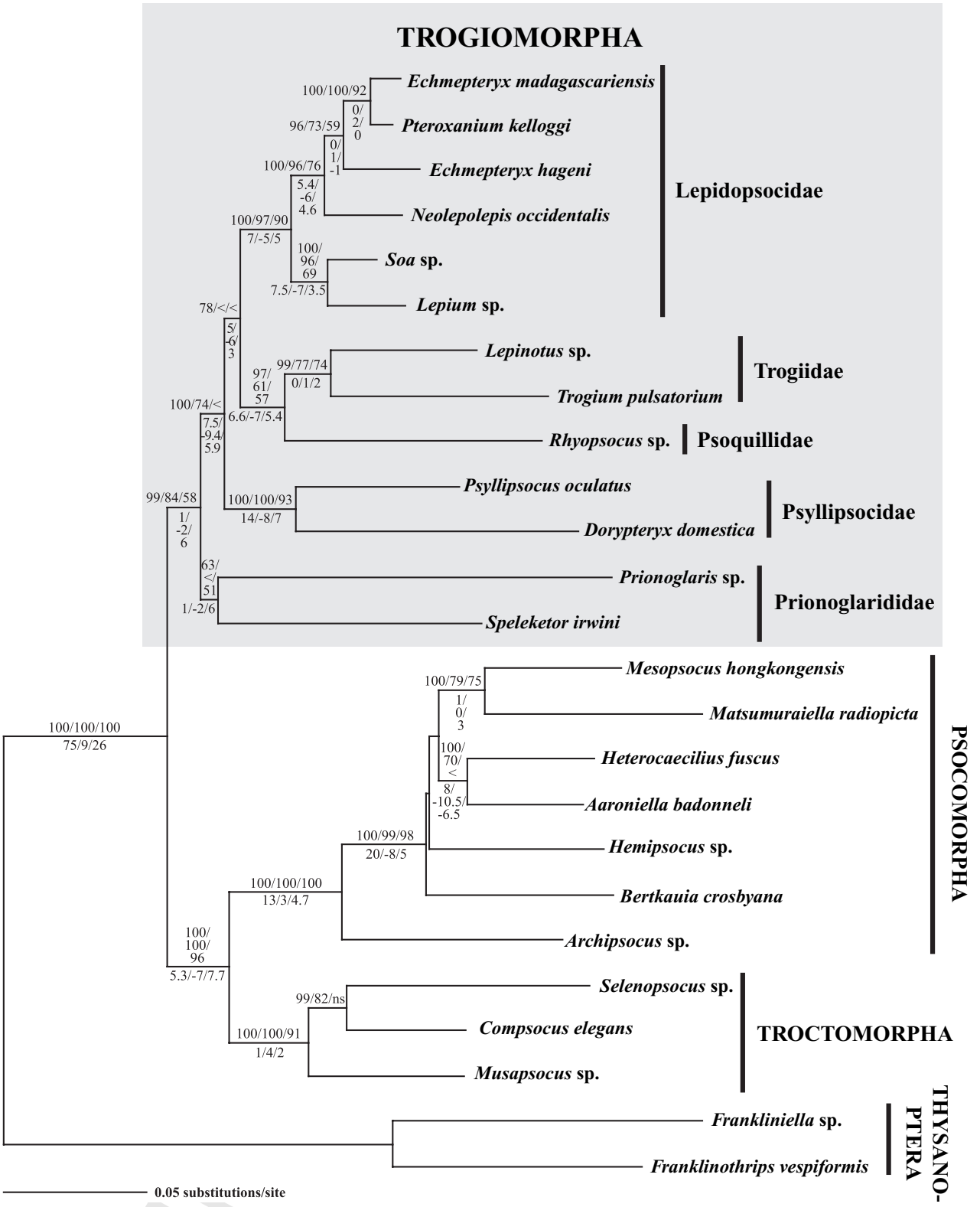

Figure 2. ML tree estimated from the data set excluding taxa with missing data. Branch lengths are proportional to ML estimated branch lengths. Numbers above the nodes are Bayesian posterior probabilities and ML and MP bootstrap support values, respectively. The numbers below the nodes indicate partitioned Bremer supports of the 18S/Histone 3/16S genes, respectively.

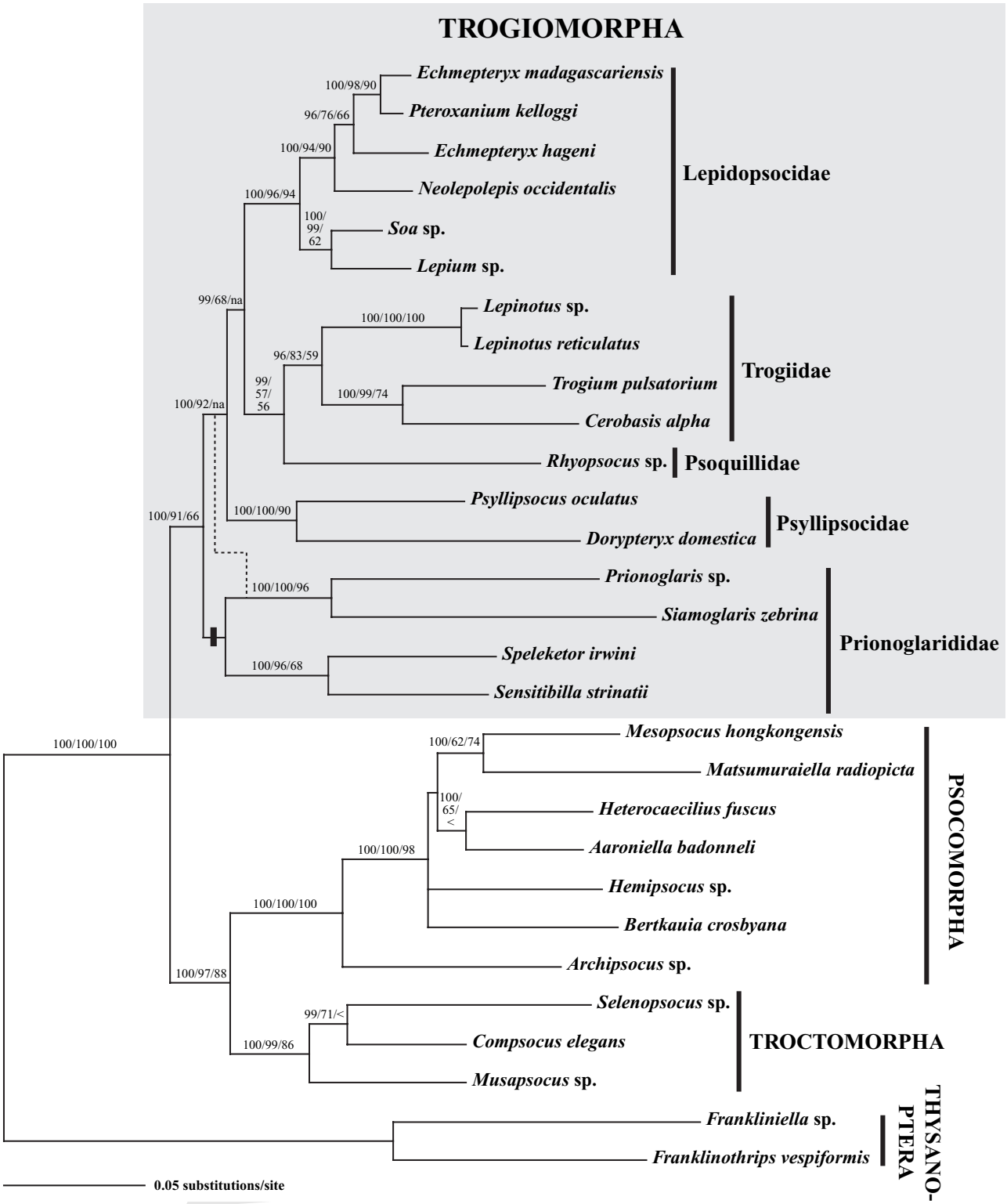

Figure 3. ML tree estimated from the data set including taxa with missing data constraining the monophyly of Prionoglarididae. Unconstrained ML analysis does not recover monophyly of Prionoglarididae (indicated by the dotted line). Branch lengths are proportional to ML estimated branch lengths. Support values for the nodes (Bayesian posterior probability/ML bootstrap/ML bootstrap, respectively) are from an unconstrained analysis.

analyses, and all analyses placed Prionoglarididae as the sister of the remainder of Trogiomorpha.

Different topologies were obtained by including and excluding taxa with missing data. Missing data can reduce the accuracy of phylogenetic estimation (Platnick, Griswold & Coddington, 1991; Novacek, 1992; Kitching *et al*., 1998). As discussed below, monophyly of Prionoglarididae was also supported by morphological characters and thus we also conducted an ML analysis including taxa with missing data, but constraining the monophyly of Prionoglarididae. The resulting tree (Fig. 3) was compatible with other trees estimated from the data set excluding taxa with missing data (Fig. 2). The constrained ML tree, as well as trees estimated from the data set excluding taxa with missing data, indicated that branch supporting monophyly of Prionoglarididae was very deep and short. Two well-supported clades were identified within Prionoglarididae, corresponding to the subfamilies Speleketorinae (*Speleketor* + *Sensitibilla*) and Prionoglaridinae (*Prionoglaris* + *Siamoglaris*) (Lienhard, 2004).

Monophyly of the infraorder Psocathropetae (= Prionoglarididae + Psyllipsocidae) was rejected by all analyses, and the family Psyllipsocidae was placed as the sister group of the infraorder Atropetae (= Lepidopsocidae + Trogiidae + Psoquillidae). This relationship was supported very strongly by Bayesian posterior probability and ML bootstrapping, but received weaker support from MP bootstrapping. Monophyly of Psyllipsocidae was also strongly supported by all analyses.

Monophyly of the infraorder Atropetae was recovered across all analyses, except for the MP analysis that included taxa with missing data. Within Atropetae, the Trogiidae and Psoquillidae were recovered as sister taxa across all analyses. Monophyly of the families within Atropetae was also supported. However, only one representative was selected from the Psoquillidae, so the monophyly of this family could not be tested.

Of the representatives of the Lepidopsocidae, two exemplars (*Neolepolepis* and *Pteroxanium*) were selected from the Echinopsocinae. However, monophyly of Echinopsocinae was not recovered by any analysis, because *Echmepteryx* (Lepidopsocinae) was always embedded within the Echinopsocinae. Bootstrap support and posterior probability of Echinopsocinae + *Echmepteryx* were very high. Monophyly of *Echmepteryx* was also not recovered, and *Echmepteryx madagascariensis* was closer to *Pteroxanium kelloggi* than to *E. hageni*: i.e. monophyly of Lepidopsocinae was not supported either. Two exemplars, *Lepium* and *Soa*, were selected from the Perientominae, and monophyly of the subfamily was well supported.

DISCUSSION

Data evaluation

The partition homogeneity test (Farris *et al*., 1994, 1995) revealed no significant heterogeneity between 18S and 16S. By contrast, significant heterogeneity was detected between Histone 3 and 18S and 16S. Although the partition homogeneity test has been widely used to test whether different gene partitions are consistent with the same phylogeny (e.g. Lecointre & Deleporte, 2005), the test is known to be sensitive to different evolutionary rates between sequences (Dolphin *et al*., 2000; Barker & Lutzoni, 2002; Johnson & Whiting, 2002; Johnson *et al*., 2002).

Comparisons of uncorrected pairwise distance of each gene and levels of homoplasy (CI and RI) reveal that the substitution rate for 18S is lower than for 16S, which is in turn lower than for Histone 3. As shown in Figure 1, Histone 3 exhibits substantial multiple substitution, 25% sequence divergence, whereas such an effect is not evident for 16S. Bootstrap analyses of each gene region separately (trees not shown) reveal that Histone 3 does not have a strong phylogenetic signal by itself, whereas 18S and 16S do. Histone 3 is one of the most conservative genes at the amino-acid level (Page & Holmes, 1998), and thus almost all substitutions are observed at the third codon position, which make this gene prone to extensive multiple substitution and homoplasy. Partitioned Bremer support values show that Histone 3 provides concordant information with 18S and 16S for shallow clades, but not for deep clades. Thus, we conclude that the significant heterogeneity between Histone 3 and the ribosomal genes is due to different evolutionary rates rather than different phylogenetic histories (Dolphin *et al*., 2000; Barker & Lutzoni, 2002; Darlu & Lecointre, 2002; Johnson & Whiting, 2002; Johnson *et al*., 2002; Yoshizawa, 2004).

Phylogenetic analyses and higher level systematics of Trogiomorpha

All analyses produced generally well-resolved and well-supported trees. Monophyly of Trogiomorpha was recovered throughout all analyses and received strong support by Bayesian posterior probability and ML bootstrapping. On the basis a morphological data set, Perrichot *et al*. (2004) suggested that monophyly of Trogiomorpha was supported by an autapomorphy of antennae with greater than 20 segments. However, this character state is actually plesiomorphic. The analysis of Perrichot *et al*. (2004) did not include non-psocodean outgroups and thus the root of the tree and the polarity of character states at the basal nodes cannot be determined by their analysis (Maddison, Donoghue & Maddison, 1984). Thus, the present study

provides the first support for monophyly of the suborder Trogiomorpha based on a formal phylogenetic analysis. The suborder Trogiomorpha has been characterized based on many morphological features, most of them plesiomorphic. However, three character states appear to be morphological apomorphies of the group (i.e. not present in outgroup taxa): (1) ventral and dorsal valves of gonapophyses strongly reduced; (2) external valve well developed, close to ventral midline of abdomen, forming the ovipositor; and (3) subgenital plate short, covering at most basal part of external valves.

Inclusion of taxa with missing data tends to destabilize various parts of the tree, especially the monophyly of Prionoglarididae (e.g. a paraphyletic Prionoglarididae is placed at the most basal node of the Trogiomorpha by ML and Bayesian analyses, whereas a monophyletic Prionoglarididae is placed as the sister of Trogiidae + Psoquillidae clade by MP analysis). However, two of the four samples with missing data are members of the Prionoglarididae, owing to the scarcity of freshly collected specimens of this family and the difficulty of amplifying DNA from old material (*Sensitibilla* and *Siamoglaris*; see Table 2). Inclusion of these taxa with missing data may reduce the accuracy of tree estimation (Platnick *et al*., 1991; Novacek, 1992; Kitching *et al*., 1998). Excluding these four taxa from the analyses, the monophyly of Prionoglarididae is recovered, and the family is consistently sister to the remainder of Trogiomorpha.

Based on analyses of simulation data, Wiens (2003) suggested that missing data will have no impact on the accuracy of parsimony analyses when the data contain a sufficient number of characters (e.g. more than 100 characters with up to 50% missing data). However, in the present analyses, the tree becomes highly unstable upon the inclusion of taxa with missing data, even though the present data set includes 2427 characters and the missing data comprises less than 5% of the total. For example, when taxa with missing data are included in the analysis, Prionoglarididae is imbedded within Atropetae and placed as the sister of the Trogiidae + Psoquillidae clade by parsimony. This result conflicts with all other analyses and morphological observations and thus is likely to be an artefact caused by inclusion of taxa with missing data. The simulations of Wiens (2003) created missing data randomly throughout the data matrix, whereas in practice, molecular data are more likely to be missing for entire genes (as was the case in our study). Because different genes can have different substitution properties, data sets in which entire gene sequences are missing for some taxa might more dramatically affect the phylogenetic results. Results obtained from the data set that excludes taxa with missing data are in agreement with traditional classifications, including the monophyly of Prionoglarididae and the basal placement of this family, which are also supported morphologically. The monophyly of Prionoglarididae is morphologically supported by a highly specialized male genital structure and the simplification or reduction of the lacinia in adults (Mockford, 1984; Lienhard, 2004). Therefore, we reanalysed the full data set constraining the monophyly of Prionoglarididae and present this tree as the best phylogenetic hypothesis for the Trogiomorpha (for those samples for which sequence data for at least some of the genes are available; Fig. 3).

Based on detailed morphological examination, Lienhard (2000, 2004) recognized two clades within the Prionoglarididae: Prionoglaridinae and Speleketorinae. Monophyly of both clades is well supported by analyses of the DNA sequence data. Vishnyakova (1980) tentatively placed the origin of the Trogiomorpha in the Early Jurassic. Because of the very disjunct distribution of the extant representatives of the Prionoglarididae, the four genera within the family Prionoglarididae could be interpreted as Pangaean relicts. Each of the four genera of this family is known from a different zoogeographical region (Palaearctic, Oriental, Ethiopian, Nearctic), where the few known species are very rare and usually live in caves or similar habitats (Lienhard, 2000, 2004). The present molecular trees suggest that the origin and diversification of the family is deep and possibly support this scenario. However, more data (including fossil record evidence) is required for more precise dating of the tree.

Monophyly of Psocathropetae (Prionoglarididae + Psyllipsocidae) is not supported by any analysis, and the Psyllipsocidae is consistently recovered as the sister group of Atropetae. A sister group relationship between Psyllipsocidae + Atropetae is strongly supported by ML bootstrapping and Bayesian posterior probability (74–100%), although more weakly by MP bootstrapping (less than 50%). In the most widely used classification system (Lienhard & Smithers, 2002), two families, Psyllipsocidae and Prionoglarididae, are assigned to the infraorder Psocathropetae (Table 1). This infraorder has been characterized by the following two character states (Mockford, 1993): (1) ventral and dorsal valves of gonapophyses usually present, external valve not elongated; and (2) veins CuP and A1 of forewing ending together on wing margin (nodulus). However, these character states are plesiomorphic (presence of three pairs of valves, external valve broad) or highly homoplastic (presence of nodulus) (Smithers, 1972), and no convincing autapomorphy of the infraorder is known. By contrast, we observed the following two possible synapomorphies that support a sister group relationship between Psyllipsocidae and Atropetae: (1) paraproct with anal

spine; and (2) phallosome opened anteriorly. Therefore, the infraorder Psocathropetae appears to be a paraphyletic group, and this classification should be abandoned. The monophyly of Psyllipsocidae is supported by the molecular trees and also by the following morphological autapomorphy: spermathecal sac with complex sclerifications at origin of duct, usually with an accessory vesicle (Lienhard, 1998; personal observation by C.L. on several undescribed species of *Psyllipsocus*).

The monophyly of Atropetae is recovered throughout the analyses, although the support for this clade is relatively low, except for Bayesian posterior probability obtained from the data set that includes taxa with missing data (99%). Morphologically, the monophyly of Atropetae is supported by two autapomorphies (Mockford, 1993): (1) external valves of gonapophyses elongated and partially joined together on midline by membrane, composing the ovipositor; and (2) spermathecal sac with one or two glandular accessory bodies. Within the Atropetae, a sister group relationship between Trogiidae and Psoquillidae is recovered throughout the analyses. This relationship has already been suggested by Smithers (1972), based on the presence of spermathecal accessory bodies in Psoquillidae and Trogiidae. However, the presence of homologous glandular structures in the Lepidopsocidae led Mockford (1993) to consider this character state a synapomorpy of all the Atropetae (see above). The monophyly of Trogiidae + Psoquillidae is supported by the following synapomorphies (Mockford, 1993; Lienhard, 1998): (1) pretarsal claws lacking preapical tooth; and (2) pulvillus distinctly enlarged through its whole length.

In the present analyses, only one representative of the Psoquillidae is included and thus monophyly of the family cannot be tested. Morphologically, the monophyly of Psoquillidae is supported by a character state of accessory bodies situated at the opening of spermatheca (Mockford, 1993). In this case, the character state observed in the Trogiidae [.e. spermathecal accessory bodies consisting of two denticulate plaques ('maculae') attached to spermathecal wall] is plesiomorphic. Monophyly of the other families included in the Atropetae is well supported by the molecular analyses. Morphologically, monophyly of these families is supported as follows – Lepidopsocidae: body and forewings covered with scales or dense setae; Trogiidae: wings greatly reduced and always veinless, sometimes absent.

Although the family-level and higher classification scheme of the Trogiomorpha is well supported by the molecular data, the present results strongly contradict some presently accepted subfamilial- or generic-level classification within the family Lepidopsocidae. Many subfamilies and genera in the Lepidopsocidae are, however, characterized by plesiomorphic (e.g. absence of scale: Thylacellinae) or highly homoplastic (e.g. brachyptery: Echinopsocinae) characters. Wing venation is also used frequently to define subfamilies or genera (e.g. hindwing with closed cell: Thylacellinae, Perientominae), but brachyptery is frequent in the family and thus such characters are less valuable for defining monophyletic groups. Therefore, revision of the subfamilial- or generic-level classification of the Lepidopsocidae is required based on more dense sampling, detailed re-examination of morphological characters and analyses of more rapidly evolving gene sequences useful for resolving more recent nodes. Smithers (1972), recently followed by Li (2002), proposed a new classification for the Lepidopsocidae (*sensu* Mockford, 1993), in which some lepidopsocid subfamilies (*sensu* Mockford, 1993) are treated as independent families or even superfamilies (see Table 1). However, such a system cannot be justified from the present results.

In conclusion, the presently accepted taxonomic classification of the suborder Trogiomorpha (Lienhard & Smithers, 2002) is well supported by DNA sequence data, except that monophyly of Psocathropetae is rejected throughout the analyses. This result is also supported by more detailed morphological observations. Based on these results, the following new higher classification is proposed for the Trogiomorpha.

PROPOSED CLASSIFICATION OF TROGIOMORPHA AND DIAGNOSES OF SUPRAFAMILIAL TAXA

The higher classification system proposed here is the direct translation of the molecular trees obtained in this study according to the annotated Linnean system method (Wiley, 1981). Diagnoses are based on Mockford (1993), Lienhard (1998) and on the present study; they are only valid for adults. Autapomorphies of each taxon are given in italics.

SUBORDER TROGIOMORPHA

Antenna generally with more than 20 segments (except an undescribed psyllipsocid from Thailand with a 19-segmented antenna; personal observation by C.L.). Hypopharyngeal filaments separate, never fused on midline (sometimes reduced in Prionoglarididae). Labial palpus two-segmented, with minute basal segment and rounded or somewhat elongated distal segment. Distal inner labral sensilla consisting of a row of five identical placoids or trichoids (Badonnel, 1977). Tarsi three-segmented. Pterostigma in forewing not thickened, completely transparent or slightly opaque. Female: *ventral and dorsal valves of gonapophyses strongly reduced or absent, external*

valves well developed and setose; subgenital plate short, covering at most basal part of external valves, which come close to ventral midline of abdomen, forming the ovipositor.

Infraorder Prionoglaridetae

Second article of maxillary palpus with (Speleketorinae) or without (Prionoglaridinae) conical spur sensillum. *Lacinia simplified or reduced in adults, sometimes much shortened or virtually absent, if of normal length apex parallel-sided and lacking distinct teeth* (NB: lacinia normally developed in nymphs, with several apical teeth). Forewing: basal segment of Sc well developed and forming a large arc, ending on R1 somewhat basally to pterostigma, delimiting a subcostal cell not reaching the wing margin; pterostigma joined to stem of radial fork by a long crossvein; nodulus present, i.e. CuP and A1 meeting on wing margin. Hindwing: vein A bifurcate. Wing pilosity reduced or virtually absent. Paraproct lacking anal spine. Female: external valve of gonapophyses broad; spermathecal sac lacking glandular accessory bodies and sclerifications at origin of duct. Male: *phallosome consisting of a cuticular sac with a pair of posterolateral processes*; basal struts of phallosome, when discernible, fused anteriorly.

Included family: Prionoglarididae.

Infraorder Psyllipsocetae

Second article of maxillary palpus without conical spur sensillum (except an undescribed psyllipsocid from Thailand, where such a sensillum is present; personal observation by C.L.). Lacinia normally developed, at least with two apical teeth. Forewing: basal segment of Sc short, ending free in membrane or joining wing margin; nodulus present, i.e. CuP and A1 meeting on wing margin (except the fossil genus *Khatangia* Vishnyakova, 1975, in which CuP and A1 reach the wing margin separately). Hindwing: vein A simple or bifurcate. Wing pilosity well developed. Paraproct with anal spine. Female: external valve of gonapophyses broad; spermathecal sac lacking glandular accessory bodies, *with complicated sclerifications at origin of duct and often with accessory vesicle*. Male: basal struts of phallosome never fused anteriorly.

Included family: Psyllipsocidae.

Infraorder Atropetae

Second article of maxillary palpus with conical spur sensillum. Wings often strongly reduced. Forewing: basal segment of Sc well developed or reduced; nodulus absent, i.e. CuP and A1 reaching wing margin separately. Hindwing: vein A simple. Wing pilosity well developed (hairs usually modified to scales in Lepidopsocidae). Paraproct with anal spine. Female: *external valves of gonapophyses elongated and partially joined together on midline by membrane, composing the ovipositor; spermathecal sac with one or two glandular accessory bodies*, attached on wall of sac (Lepidopsocidae and Trogiidae) or near origin of duct (Psoquillidae). Male: basal struts of phallosome never fused anteriorly.

Included families: Lepidopsocidae, Psoquillidae, Trogiidae.

ACKNOWLEDGEMENTS

We thank A. Garcia-Aldrete, Z. Kucerova, E. Mockford, P. Schwendinger, P. Strinati, N. Takahashi and K. Thaler for providing specimens for this study. This project was financially supported by JSPS grant 15770052 (K.Y.), NSF grant DEB-0107891 and University of Illinois Research Board (K.P.J.). Some specimens used in the paper were collected during a Tropic Asian insects inventory project conducted by O. Yata supported by JSPS grant 142550161.

REFERENCES

Badonnel A. 1951. Psocoptères. In: Grassé PP, ed. *Traité de Zoologie (10)*. Paris: 1301–1340.

Badonnel A. 1977. Sur les sensilles du bord distal du labre des Psocoptères et leur intérêt taxonomique. *Bulletin de la Société Entomologique de France* **82:** 105–113.

Baker RH, DeSalle R. 1997. Multiple sources of character information and the phylogeny of Hawaiian Drosophilids. *Systematic Biology* **46:** 654–673.

Baker RH, Yu XB, DeSalle R. 1998. Assessing the relative contribution of molecular and morphological characters in simultaneous analysis trees. *Molecular Phylogenetics and Evolution* **9:** 427–436.

Barker FK, Lutzoni FM. 2002. The utility of the incongruence length difference test. *Systematic Biology* **51:** 625–637.

Barker SC, Whiting M, Johnson KP, Murrell A. 2003. Phylogeny of the lice (Insecta, Phthiraptera) inferred from small subunit rRNA. *Zoologica Scripta* **32:** 407–414.

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

Buckley TR, Simon C, Flook PK, Misof B. 2000. Secondary structure and conserved motifs of the frequently sequenced domains IV and V of the insect mitochondrial large subunit rRNA gene. *Insect Molecular Biology* **9:** 565–580.

Colgan DJ, McLauchlan A, Wilson GDF, Livingston SP, Edgecombe GD, Macaranas J, Cassis G, Gray MR. 1998. Histone H3 and U2 sn RNA DNA sequences and arthropod molecular evolution. *Australian Journal of Zoology* **46:** 419–437.

Darlu P, Lecointre G. 2002. When does the incongruence length difference test fail. *Molecular Biology and Evolution* **19:** 432–437.

Dolphin K, Belshae R, Orme CDL, Quicke DLJ. 2000. Noise and incongruence: interpreting results of the incongruence length difference test. *Molecular Phylogenetics and Evolution* **17:** 401–406.

Farris JS, Källersijö M, Kluge AG, Bult C. 1994. Testing significance of congruence. *Cladistics* **10:** 315–320.

Farris JS, Källersijö M, Kluge AG, Bult C. 1995. Constructing a significance test for incongruence. *Systematic Biology* **44:** 570–572.

Huelsenbeck JP, Ronquist R. 2001. MrBayes: Bayesian inference on a phylogeny. *Bioinformatics* **17:** 754–755.

Johnson KP, Mockford EL. 2003. Molecular systematics of Psocomorpha (Psocoptera). *Systematic Entomology* **28:** 409–416.

Johnson KP, Weckstein JD, Witt CC, Faucett RC, Moyle RG. 2002. The perils of using hosts relationships in parasite taxonomy: phylogeny of the *Degeeriella* complex. *Molecular Phylogenetics and Evolution* **23:** 150–157.

Johnson KP, Whiting MF. 2002. Multiple genes and the monophyly of Ischnocera (Insecta: Phthiraptera). *Molecular Phylogenetics and Evolution* **22:** 101–110.

Johnson KP, Yoshizawa K, Smith VS. 2004. Multiple origins of parasitism in lice. *Proceedings of the Royal Society of London (B)* **271:** 1771–1776.

Kitching IJ, Forey PL, Humphries CJ, Williams DM. 1998. *Cladistics. The theory and practice of parsimony analysis*, 2nd edn. London: Oxford University Press.

Kjer KM. 2004. Aligned 18S and insect phylogeny. *Systematic Biology* **53:** 506–514.

Lecointre G, Deleporte P. 2005. Total evidence requires exclusion of phylogenetically misleading data. *Zoologica Scripta* **34:** 101–117.

Li Fasheng. 2002. *Psocoptera of China*. Beijing: Science Press.

Lienhard C. 1998. Psocoptères euro-méditerranéens. *Faune de France* **83:** xx + 1–517.

Lienhard C. 2000. A new genus of Prionoglarididae from a Namibian cave (Insecta: Psocoptera). *Revue Suisse de Zoologie* **107:** 871–882.

Lienhard C. 2004. *Siamoglaris zebrina* gen. n., sp. n., the first representative of Prionoglarididae from the Oriental Region (Insecta: Psocoptera). *Revue Suisse de Zoologie* **111:** 865–875.

Lienhard C, Smithers CN. 2002. *Psocoptera (Insecta): World catalogue and bibliography*. Instrumenta Biodiversitatis 5. Genève: Muséum d'Histoire Naturelle.

Lyal CHC. 1985. Phylogeny and classification of the Psocodea, with particular reference to the lice (Psocodea: Phthiraptera). *Systematic Entomology* **10:** 145–165.

Maddison WP, Donoghue MJ, Maddison DR. 1984. Outgroup analysis and parsimony. *Systematic Zoology* **33:** 83–103.

Marshall IK. 2003. A morphological phylogeny of four families of amblyceran lice (Phthiraptera: Amblycera: Menoponidae, Boopiidae, Laemobothriidae, Ricinidae). *Zoological Journal of the Linnean Society* **138:** 39–82.

Mockford EL. 1984. Two new species of *Speleketor* from southern California with comments on the taxonomic position of the genus (Psocoptera: Prionoglaridae). *Southwestern Naturalist* **29:** 169–179.

Mockford EL. 1993. *North American Psocoptera (Insecta).* Flora and Fauna Handbook, no. 10. Gainesville, FL: Sandhill Crane Press.

Novacek MJ. 1992. Fossils as critical data for phylogeny. In: Wheeler QD, Novacek MJ, eds. *Extinction and phylogeny*. New York: Columbia University Press, 46–88.

Page RDM, Holmes EC. 1998. *Molecular evolution. A phylogenetic approach*. Oxford: Blackwell Science.

Pearman JV. 1936. The taxonomy of the Psocoptera: preliminary sketch. *Proceedings of the Royal Entomological Society, London (B)* **5:** 58–62.

Perrichot V, Azar D, Neraudeau D, Nel A. 2004. New Psocoptera in the Early Cretaceous amber of SW France and Lebanon (Insecta: Psocoptera: Trogiomorpha). *Geological Magazine* **140:** 669–683.

Platnick NI, Griswold CE, Coddington JA. 1991. On missing entries in cladistic analysis. *Cladistics* **7:** 337–343.

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

Roesler R. 1940. Neue und wenig bekannte Copeognathengattungen. I. *Zoologischer Anzeiger* **129:** 225–243.

Roesler R. 1944. Die Gattungen der Copeognathen. *Stettiner Entomologische Zeitung* **105:** 117–166.

Ronquist F, Huelsenbeck JP. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19:** 1572–1574.

Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a comparison of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America* **87:** 651–701.

Smith VS. 2001. Avian louse phylogeny (Phthiraptera: Ischnocera): a cladistic study based on morphology. *Zoological Journal of the Linnean Society* **132:** 81–144.

Smith VS, Page RD, Johnson KP. 2004. Data incongruence and the problem of avian louse phylogeny. *Zoologica Scripta* **33:** 239–259.

Smithers CN. 1972. The classification and phylogeny of the Psocoptera. *Australian Museum Memoirs* **14:** 1–349.

Sorenson MD. 1999. *Treerot*. Boston, MA: Boston University Press.

Swofford D. 2002. *PAUP*: phylogenetic analysis using parsimony (*and other methods)*. Sunderland, MA: Sinauer Associates.

Vishnyakova VN. 1975. Psocoptera in Late Cretaceous insect bearing resins of the Taimyr. *Entomologicheskoe Obozrenie* **54:** 92–106. (English translation. *Entomological Review, Washington* **54:** 63–75).

Vishnyakova VN. 1980. The historical development of the Insecta. Superorder Psocidea. Order Psocida. *Trudy Paleontologicheskogo Instituta* **175:** 52–57.

Wiens JJ. 2003. Missing data, incomplete taxa and phylogenetic accuracy. *Systematic Biology* **52:** 528–538.

Wiley EO. 1981. *Phylogenetics. The theory and practice of phylogenetic systematics*. New York: John Wiley & Sons.

Yoshizawa K. 2002. Phylogeny and higher classification of suborder Psocomorpha (Insecta: Psocodea: 'Psocoptera'). *Zoological Journal of the Linnean Society* **136:** 371–400.

Yoshizawa K. 2004. Molecular phylogeny of major lineages of *Trichadenotecnum* and a review of diagnostic morphological characters (Psocoptera: Psocidae). *Systematic Entomology* **29:** 383–394.

Yoshizawa K, Johnson KP. 2003. Phylogenetic position of Phthiraptera (Insecta: Paraneoptera) and elevated rate of evolution in mitochondrial 12S and 16S rDNA. *Molecular Phylogenetics and Evolution* **29:** 102–114.

Yoshizawa K, Johnson KP. 2005. Morphology of male genitalia in lice and relatives and phylogenetic implications. *Systematic Entomology* **in press**.

Yoshizawa K, Saigusa T. 2001. Phylogenetic analysis of paraneopteran orders (Insecta: Neoptera) based on forewing base structure, with comments on monophyly of Auchenorrhyncha (Hemiptera). *Systematic Entomology* **26:** 1–13.

4