TÜBİTAK

Turkish Journal of Zoology

http://journals.tubitak.gov.tr/zoology/

Research Article

Turk J Zool
(2014) 38: 285-297
© TÜBİTAK
doi:10.3906/zoo-1303-15

Is the subfamily Eriosomatinae (Hemiptera: Aphididae) monophyletic?

Xing-Yi LI[1,2], Li-Yun JIANG[1], Ge-Xia QIAO[1,*]

[1]Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing, P.R. China
[2]University of Chinese Academy of Sciences, Beijing, P.R. China

Received: 13.03.2013 • **Accepted:** 09.12.2013 • **Published Online:** 21.03.2014 • **Printed:** 18.04.2014

Abstract: Eriosomatinae, the gall-forming aphid subfamily, traditionally consists of 3 tribes, Eriosomatini, Pemphigini, and Fordini. However, the phylogenetic relationships among these tribes remain controversial, which has made it difficult to conduct further investigation regarding the evolution of galls and host alternations in this group. We analyzed the molecular phylogeny of the subfamily Eriosomatinae, combining sequences from 2 mitochondrial genes (*COI* and *COII*) and 2 nuclear genes (*EF-1α* and *LWO*). The reconstructions were implemented based on single-gene and multigene datasets through 3 different reconstructing algorithms, respectively; analyses with 5 different out-groups were also conducted. Results revealed a large paraphyletic clade, in which there were 4 out-groups clustering between Eriosomatini and the other 2 tribes. However, the monophyly of the 3 tribes was well supported by the obtained trees, respectively.

Key words: Eriosomatinae, molecular phylogeny, monophyly, paraphyletic group

1. Introduction

The aphid subfamily Eriosomatinae (Hemiptera: Aphididae), once known as Pemphiginae (Blackman and Eastop, 1994) or Pemphigidae (Zhang et al., 1999), is widely distributed in the Holarctic and Oriental regions and is composed of 310 valid species belonging to 48 genera (Remaudière and Remaudière, 1997). Eriosomatinae is typically known for inducing galls on primary host plants and shows a heteroecious holocyclic life history (i.e. seasonal host alternation and cyclical parthenogenesis) and host–plant specificity. According to the similarity of host alternation and galling habits, Eriosomatinae and Hormaphidinae have been considered sister groups (Heie, 1987) and both subfamilies have been considered primitive among Aphididae according to certain morphological characteristics, e.g., the 3-faceted eyes observed in aptera, the reduced antennal segments, and vestigial siphunculi (Zhang et al., 1999, referred to as Pemphigidae and Hormaphididae in their taxonomic system).

Based on morphological and biological evidence, Eriosomatinae was divided into 3 tribes, Eriosomatini, Pemphigini, and Fordini (Remaudière and Remaudière, 1997; Nieto Nafría et al., 2011). Fordini consists of 2 subtribes, Melaphidina and Fordina (Remaudière and Remaudière, 1997). Pemphigini also has 2 subtribes, Pemphigina and Prociphilina (Blackman and Eastop, 1994), and it has been proposed that Eriosomatini should be divided into 2 subgroups, as well (Zhang et al., 1999). The systematic status of Eriosomatinae had been proposed since the inchoate taxonomic systems of aphids were built, but it is worth noting that several recent subfamilies, such as Hormaphidinae, Phloeomyzinae, Mindarinae, and Anoeciinae, were considered closely related to Eriosomatinae or even been placed into "Pemphiginae" during the development history of the aphid taxonomic systems (Ren et al., 2006).

Monophyly of Eriosomatinae was proposed based on the unique synapomorphies of sexual females and males (Heie, 1987). However, several studies, which reconstructed phylogenies of Eriosomatinae based on either morphological or molecular evidence, did not support the monophyly of Eriosomatinae. A cladistic phylogeny of the subfamily Eriosomatinae was produced using 25 morphological and ecological characters among 28 genera distributed in China, but the monophyly of Eriosomatinae was not supported (Zhang and Chen, 1999) (Figure 1a). Molecular phylogeny of Aphidoidea based on mitochondrial 12S and 16S rDNA sequences (*12S*) involving abbreviated sampling efforts showed no clear topology of Aphididae beyond the tribal taxonomic category, and most of the tribes clustered in parallel to form a large paraphyletic group, including the 3 tribes of

* Correspondence: qiaogx@ioz.ac.cn

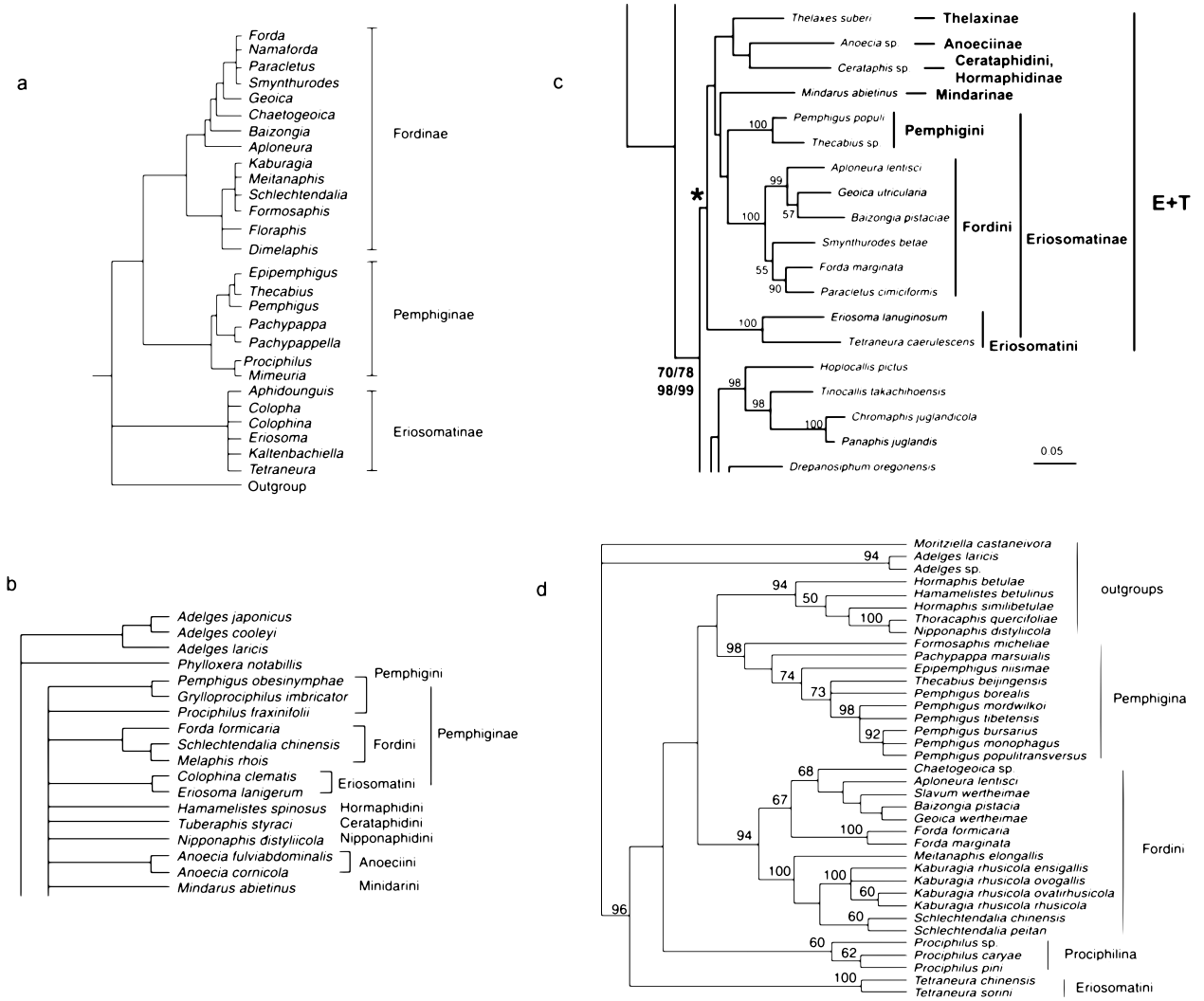

Figure 1. Previous phylogenetic hypotheses for the Eriosomatinae: a) Zhang and Chen (1999), based on morphology; b) Moran and von Dohlen (2000), based on *12S*, partial topology, 8 sampled species; c) Ortiz-Rivas et al. (2010), based on *LWO* combined with other genes, partial topology, 10 sampled species; d) Zhang and Qiao (2008), based on *EF-1α*, 25 sampled species, but only 2 species in Eriosomatini.

Eriosomatinae (von Dohlen and Moran, 2000) (Figure 1b). Results based on subunit 6 of the F-ATPase complex (*ATP-6*) showed a similar topology (Martínez-Torres et al., 2001); similarly, in the phylogenetic tree based on 2 nuclear genes (*LWO* and *EF-1α*) and 2 mitochondrial genes (*COII* and *ATP-6*), the 3 tribes of Eriosomatinae (10 sampled species) clustered polyphyletically with Hormaphidinae, Mindarinae, Anoeciinae, and Thelaxinae together to form an "E+T" clade (Ortiz-Rivas et al., 2010) (Figure 1c). Additionally, a phylogeny of Eriosomatinae based on *EF-1α* was paraphyletic, as well (Zhang and Qiao, 2008) (Figure 1d).

The phylogeny of each tribe was also proposed, in which Eriosomatini and Fordini were found to be monophyletic but Pemphigini was not. According to the phylogeny based on 52 morphological characters, Eriosomatini was regarded as a monophyletic group (Sano and Akimoto, 2011). Additionally, Inbar et al. (2004) demonstrated the monophyly of Fordina (Fordini in their paper), inferred from *COI* and *COII* from 14 species. The monophyly of Fordini was also supported based on *COI* and *EF-1α* data (Zhang and Qiao, 2007a, Fordinae in their paper). In addition, Zhang and Qiao (2007b) confirmed the position of genus *Formosaphis* Takahashi 1925 within Pemphigini rather than in Fordini based on *EF-1α* sequences, but the monophyly of Pemphigini was not supported (Zhang and Chen, 1999).

Therefore, the phylogeny of the subfamily Eriosomatinae needs further investigation. In this study, we sampled 42 species in 24 genera of Eriosomatinae, including most

genera of Eriosomatinae distributed in China and all other genera, with sequences available in the GenBank; utilized 4 gene markers to build multigene datasets; and reconstructed the phylogeny within Eriosomatinae with maximum parsimony (MP), maximum likelihood (ML), and Bayesian inference (BI) algorithms.

2. Materials and methods

2.1. Sampling

Samples for DNA extraction and polymerase chain reaction (PCR) were collected by the authors and colleagues in recent years, mostly from galls on the primary hosts, and other sequences were directly downloaded from the NCBI database. Collection information for these samples, including locations, sample numbers, and collection dates, is shown in the Appendix. Except for specimens for slide-mounting that were stored in 70% ethanol, all other specimens were stored in 95% or 100% ethanol. All samples and voucher specimens were deposited in the National Zoological Museum of China, Institute of Zoology, Chinese Academy of Sciences, Beijing, China. Three to 5 individuals per sample were made into slide-mounted specimens for species identification. Specimens were identified according to their main morphological diagnostic features, and compared with previously identified specimens of corresponding species. Species names of each sample were also provided in the Appendix.

Among the 64 samples obtained, 50 corresponded to in-groups and the other 14 to out-groups. The in-groups covered 42 species (34 identified, 8 unidentified assigned as sp.) in 24 genera of Eriosomatinae. Representative species of related subfamilies of the "E+T" clade in Aphididae (Ortiz-Rivas et al., 2010) (such as Hormaphidinae, Phloeomyzinae, Anoeciinae, Mindarinae, and Thelaxinae) and sister groups to Aphididae (such as Adelgidae and Phylloxeridae) were chosen as multiple out-groups, including *Hormaphis similibetulae*, *Ktenopteryx eosocallis*, *Nipponaphis distyliicola*, and *Ceratoglyphina bambusae* in Hormaphidinae; *Mindarus keteleerifoliae* and *Mindarus abietinus* in Mindarinae; *Phloeomyzus passerinii* in Phloeomyzinae; *Anoecia* sp. in Anoeciinae; *Kurisakia querciphila* and *Thelaxes suberi* in Thelaxinae; *Pineus armandicola* and *Adelges laricis* in Adelgidae; and *Phylloxerina salicis* and *Daktulosphaira vitifoliae* in Phylloxeridae, some of which were utilized as out-groups in previous studies (Zhang et al., 1999; Ortiz-Rivas et al., 2004, 2010; Zhang and Qiao, 2008). However, the interrelations of the out-groups and the in-group still need investigation.

2.2. Sequence obtainment

One aphid individual per sample was selected for molecular experiments. The classical method of CTAB (hexadecyltrimethylammonium bromide) extraction was applied to obtain whole-genome DNA from each individual. The genes *COI*, *COII*, *EF-1α*, and *LWO* were amplified. These genes were applied widely in the molecular phylogenetic studies of aphids and are easy to use in PCR. Additionally, *LWO* was shown to be effective in the phylogenetic studies of Aphididae (Ortiz-Rivas et al., 2010). The primers are listed in Table 1 and the conditions of the amplification reactions followed the instructions provided in relevant references. Sequencing of the obtained PCR products was carried out by the commercial sequencing department of the Beijing Genomics Institute (Beijing, China) using the corresponding primers for amplification. However, there were a number of nuclear sequences whose sequencing was problematic, and so we purified their PCR products and cloned them into DH-5α, then resent 1 mL of bacterial fluid for sequencing after confirming the successful conversion of the segment.

Table 1. Primers used for the amplification of each gene.

Genes	Primers	Primer sequences (5'-3')	Reference
COI	LepF	ATTCAACCAATCATAAAGATATTGG	Foottit et al., 2008
	LepR	TAAACTTCTGGATGTCCAAAAAATCA	
COII	2993+	CATTCATATTCAGAATTACC	Normark, 1999
	A3772	GAGACCATTACTTGCTTTCAGTCATCT	
EF-1α	EF3	GAACGTGAACGTGGTATCAC	von Dohlen et al., 2002
	EF2	ATGTGAGCAGTGTGGCAATCCAA	
LWO	OPSETF1	GGYRTYACNATTTTYTTCTTRGG	Ortiz-Rivas et al., 2010
	OPSETR1	GANCCCCADATYGTNAATAAYGG	

2.3. Alignment and multisequence properties

The chromatograms obtained from sequencing were edited and assembled using SEQMAN-II, one of the applications in DNASTAR 5.0 (DNASTAR, Madison, WI, USA). For nuclear genes, introns are not appropriate for phylogenetic analysis, and so intron-splicing was applied based on the GT-AG rule along with reference cDNA segments from species of Eriosomatinae (GenBank accession numbers DQ493839 and AM996856 for *EF-1α* and *LWO*, respectively). Mitochondrial sequences were also translated into amino acid sequences to check for the presence of termination codons (usually UGA for eukaryotes) to avoid the introduction of pseudogenes. It was notable that either 1 or 2 of the 4 genes was not present in a few samples, although this did not negatively impact the analysis. Multiple alignments for each gene were conducted using the accessory CLUSTAL-W application in BIOEDIT 7.0 (Hall, 2004). The aligned data were then imported into MEGA 4.0 (Tamura et al., 2007) for analyses of the nucleotide composition, phylogenetically informative sites, and distances between species. Additionally, the saturation of the third codon position for each gene was tested in DAMBE 5.3.8 via an implemented method and transition/transversion plots (Xia and Xie, 2001; Xia et al., 2003).

2.4. Molecular phylogenetic analysis

Before the 4 genes were combined, analyses of single-gene datasets were conducted through different methods. In the 4-gene-combined analyses, which were assigned as the major studies, the datasets were partitioned into genes through the BI method, but concatenated to form a sequential supergene through the analyses using the MP and the ML algorithms. The MP analysis was performed with PAUP* 4.0b10 (Swofford, 2003), with all sites weighted equally, gaps treated as missing data, 1000 random-addition sequences, and tree bisection reconnection branch-swapping. After a 50% majority-rule consensus tree (con-tree) was yielded, a nonparametric bootstrap test was performed with 1000 pseudoreplicates under a heuristic search strategy and 100 random-addition sequences in each pseudoreplicate to examine the topology. Before we carried out the analyses of ML and BI methods, the most appropriate nucleotide substitution models for each gene were estimated using JMODELTEST 0.1.1 (Posada, 2008) under the Akaike information criterion and the Bayesian information criterion, respectively. The ML analysis was conducted in PHYML 3.0 (Guindon et al., 2010) under a custom model with optimized nucleotide frequencies, substitution rates, and gamma distribution. The tree topology was optimized based on the nearest neighbor interchange and subtree prune and regraft search strategies using 5 random starting trees obtained from NJ estimation. A nonparametric bootstrap test was then performed with 100 replicates to examine the tree topology. BI analysis was performed with MRBAYES 3.1.2 (Ronquist and Huelsenbeck, 2003). The models and parameters for each gene were unlinked, while the topologies were linked during the analysis. Two separate reactions with 4 chains (3 heated chains and 1 cold chain) were run with a random starting tree, and it proceeded for 10 million Markov chain Monte Carlo (MCMC) generations with sampling every 1000 generations until the average standard deviation of the split frequencies became lower than 0.01. Of the 10,001 total trees sampled in each reaction, 2500 trees were discarded as burn-in samples. The remaining trees were used to generate a 50% majority-rule con-tree, in which the percentage given on a node or branch indicates the posterior probability. All of the yielded trees were browsed and edited with FIGTREE 1.3.1 (Rambaut, 2009).

In addition to the major studies, 5 separate analyses were implemented with 1 of the 5 subfamilies as the out-group while the others were excluded, including Hormaphidinae, Phloeomyzinae, Anoeciinae, Mindarinae, and Thelaxinae. This was assigned as the test of the out-group, which was intended to examine the actual position of each out-group and the stability of the paraphyletic topology of the subfamily Eriosomatinae obtained in the major studies. Trees were rooted by Adelgidae and Phylloxeridae.

3. Results

3.1. Sequence characteristics

For all samples, the sequenced segments of *COI*, *COII*, *EF-1α*, and *LWO* were approximately 700, 800, 1100, and 870 bp, respectively. After alignment and splicing, the partial sequences of *Leu*-tRNA were discarded, and the sequences of *COI*, *COII*, *EF-1α*, and *LWO* were 680, 672, 762, and 543 bp, respectively. The total length of the combined dataset was 2657 bp. All sequences were submitted to GenBank and the accession numbers are listed in the Appendix. Among the combined 2657-bp dataset of the in-groups, 1634 bp were conserved, 1023 bp were variable, and 835 bp were parsimony-informative. When aligned with out-groups, there were 1495 bp of conserved, 1162 bp of variable, and 953 bp of parsimony-informative sites. Additionally, the average base frequencies were 34.9% T, 16.0% C, 33.4% A, and 15.7% G for in-groups and 34.7% T, 16.2% C, 33.3% A, and 15.8% G for out-groups. Information on the datasets and statistics for the sequences are listed below in Table 2.

3.2. Reconstructed phylogenies

According to the results based on single-gene datasets, there were no clear resolutions higher than the subtribe category and the in-group together with some out-groups formed a comb-like topology. Additionally, the nodal supports from the results based on *LWO* were relatively higher than those of the other 3 genes. It was worth noting that the bootstrap tests were quite time-consuming when run with single-gene datasets.

Table 2. Datasets, statistics of sequences, and number of sequences including out-groups.

Genes/ datasets	Number of sequences	Alignment length (nucleotides/ amino acids)	Variable/ informative sites in nucleotides	Variable/ informative sites in amino acids	Average p-distances (in-groups/all)	Saturation in 3rd positions
COI	55	680/226	293/251	63/37	0.114/0.118	Little
COII	57	672/223	344/274	110/76	0.125/0.133	Little
EF-1α	57	762/254	261/225	28/18	0.084/0.094	No
LWO	41	543/181	262/203	79/44	0.126/0.140	Little
All	64	2657/884	1162/953	282/175	0.091/0.102	-

However, the results of the major studies were much better. MP analysis yielded 14 most parsimonious trees (MPTs). The consistency index value was 0.275 and the retention index was 0.536. The topologies of the 14 MPTs were almost identical (Figure 2) and the out-groups of Mindarinae and Thelaxinae were inserted into the in-group between Eriosomatini and the other 2 tribes, while the in-group clustered into 3 clades corresponding to the 3 tribes (Eriosomatini, Pemphigini, and Fordini). The nonparametric bootstrap test sampled 7439 trees, and the 50% majority-rule con-tree showed a similar topology to the MPTs, except that Hormaphidinae and Anoeciinae were also inserted into the in-group and a comb-like topology was present. The bootstrap values within each tribe were relatively high, but those of the node E-root and the node O+P+F were low. The ML tree yielded the "GTR+I+Γ" custom model, and the estimated substitution rates were 6.8692 (A-C), 7.9478 (A-G), 6.1693 (A-T), 4.2327 (C-G), 67.3953 (C-T), and 1.0000 (G-T). The Γ shape parameter was equal to 0.663, and the proportion of invariant sites was 0.466. The ML topology was quite similar with the MPTs, but Hormaphidinae and Anoeciinae were also inserted into the in-group. The ML con-tree also showed a comb-like topology (Figure 3), because the bootstrap values for the node E-root and the node O+P+F were also low. The topology of the BI tree was similar to that of the MPTs and ML tree, except that some out-groups were clustering together. However, the Bayesian tree exhibited much higher values of nodal posterior probabilities. According to the final con-tree (Figure 4), Pemphigini first clusters with Fordini and then clusters with some out-groups, but the nodal support values are not high. Eriosomatini clusters near the root with low nodal supports. Thus, Eriosomatinae forms a large paraphyletic group with the out-groups. However, the three tribes of the subfamily Eriosomatinae each form a monophyletic group with high nodal supports. The monophyletic clades of subtribes, such as Fordina and Melaphidina, are also presented with high supports. It is worth noting that *Formosaphis* clustered at the root part of the Pemphigini clade.

Furthermore, among the 5 separate analyses of the test of the out-group, only that with Phloeomyzinae showed the monophyly of Eriosomatinae, and the nodal supports for the clade of the in-group were 1/100/87 (BI/ML/MP). However, the other 4 analyses showed the paraphyly of Eriosomatinae, and the corresponding out-group was inserted into the in-group between Eriosomatini and the other 2 tribes (Hormaphidinae and Anoeciinae) or between Fordini and Pemphigini (Mindarinae and Thelaxinae). The basal node supports through BI/ML/MP methods are presented in Table 3. It is worth noting that the node P did not exist in some analyses when *Formosaphis* clustered in parallel with Fordini, the other Pemphigini taxa, and out-groups (Mindarinae and Thelaxinae). However, the other Pemphigini taxa formed a monophyletic group in most trees with high nodal supports.

4. Discussion

4.1. Eriosomatinae is not monophyletic

All the trees obtained in the analyses with single-gene datasets were totally comb-like, which was similar to previous results based on single-gene datasets (von Dohlen and Moran, 2000; Martínez-Torres et al., 2001). This would be due to the powerlessness of single-gene datasets, rather than the species tree originally being comb-like, whereby thus the nodal supports would be totally poor. Meanwhile, the support values in the major studies based on multigene datasets were significantly improved, especially within each subtribe. It was interesting that the 14 MPTs obtained were almost identical, which provided valuable references for the actual positions of the related out-groups on the phylogeny. Comparing the tree topologies from the major studies through BI/ML/MP reconstructions, it was concluded that the positions of some out-groups were flexible. Mindarinae and Thelaxinae were constantly inserted into the in-group, which suggested that the monophyly of Eriosomatinae might be in doubt. However,

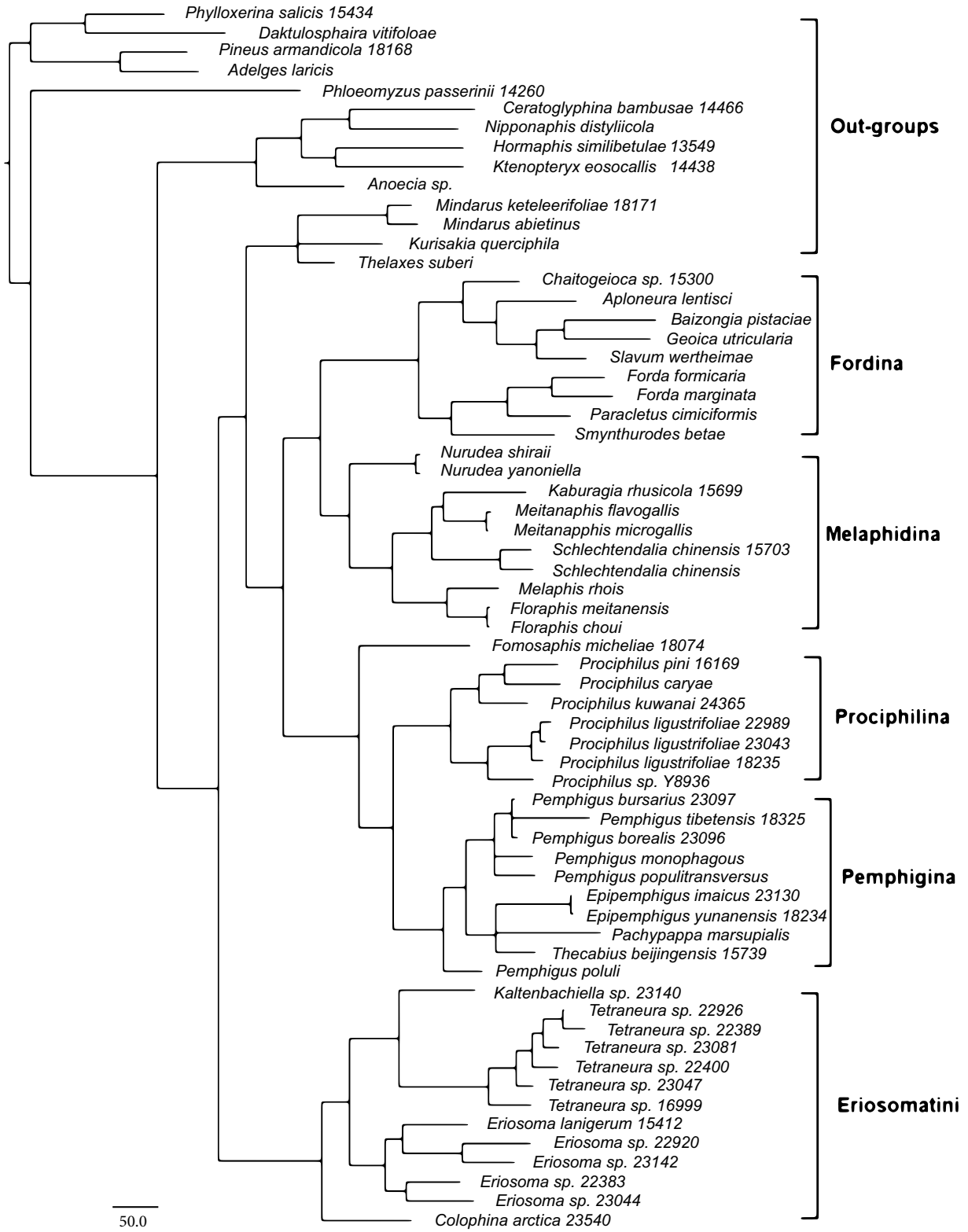

Figure 2. The strict consensus tree of the 14 MPTs from the MP analysis combining sequences from 2 mitochondrial genes (*COI* and *COII*) and 2 nuclear genes (*EF-1α* and *LWO*). Note that the topologies of the MPTs were almost identical.

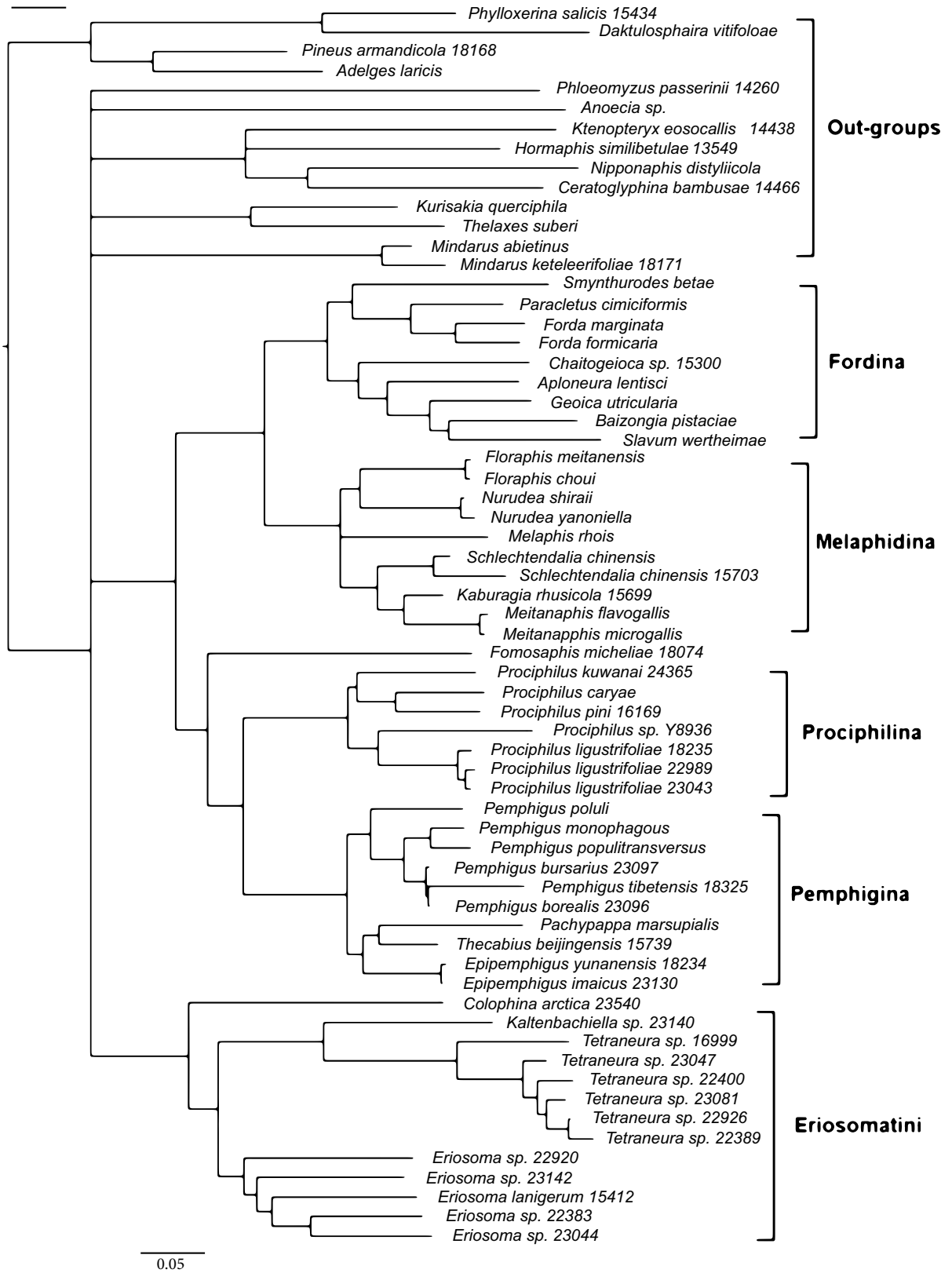

Figure 3. The 50% majority-rule consensus tree from the ML analysis combining sequences from 2 mitochondrial genes (*COI* and *COII*) and 2 nuclear genes (*EF-1α* and *LWO*). Nodal support values were omitted. Note the comb-like topology.

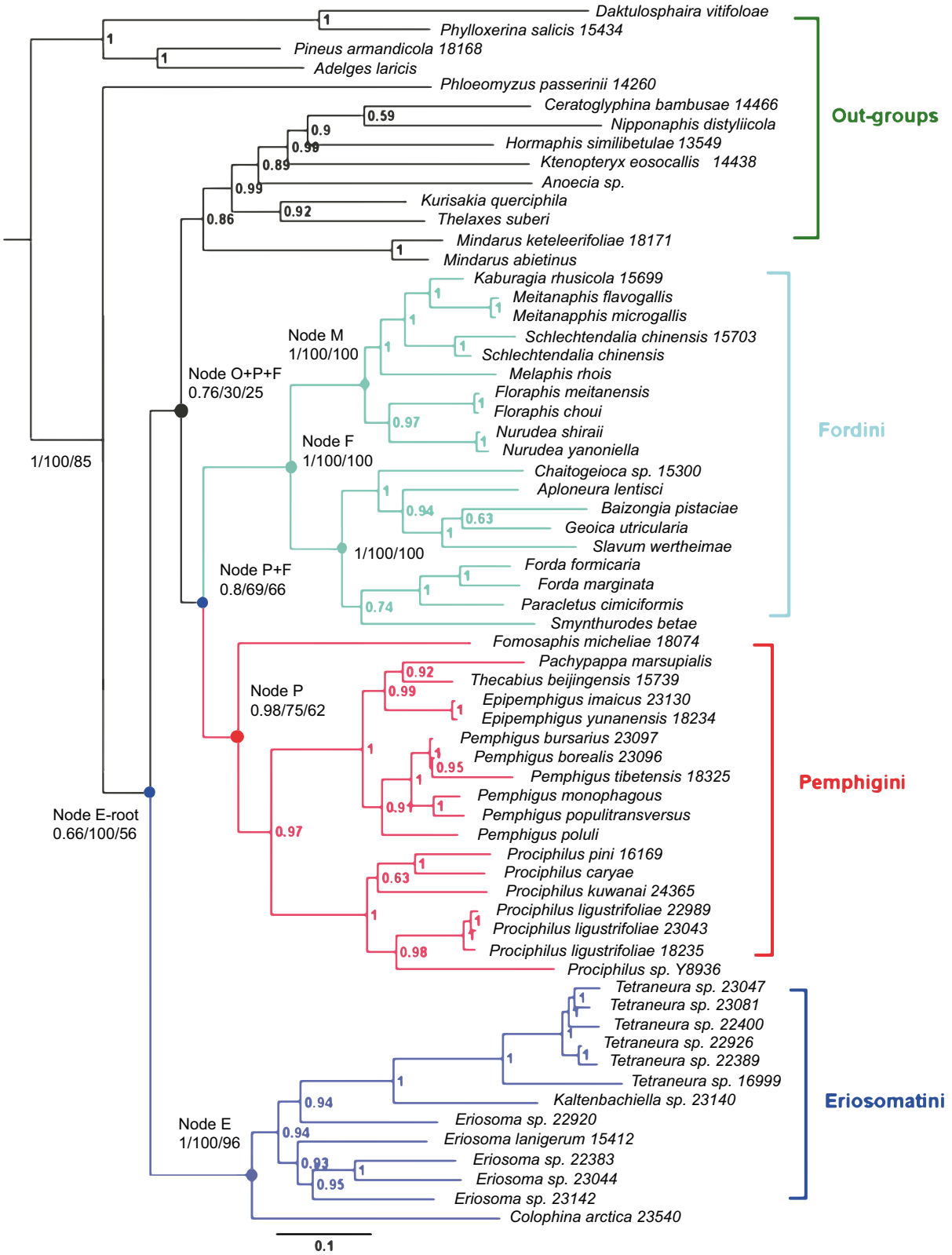

Figure 4. The Bayesian tree of Eriosomatinae combining sequences from 2 mitochondrial genes (*COI* and *COII*) and 2 nuclear genes (*EF-1α* and *LWO*). Nodes in Table 3 are marked as P = Pemphigini, E = Eriosomatini, F = Fordini, M = Melaphidina, and O = out-group. Nodal supports from different algorithms are listed in the order BI/ML/MP; the sample IDs are presented after the species names.

Table 3. Nodal supports for the basal nodes from the test of the out-group. The corresponding positions for the nodes are marked in Figure 4. "-" means the node does not exist in that analysis. Note that the node P did not exist in some analyses. P = Pemphigini, E = Eriosomatini, F = Fordini, O = out-group.

Related out-groups	Monophyly of in-group	Nodal supports for the basal nodes (BI/ML/MP)					
		E-root	O+P+F	P+F	E	P	F
Anoeciinae	No	1/100/89	0.57/44/64	0.91/78/93	1/100/100	-	1/100/100
Hormaphidinae	No	1/100/100	0.71/53/61	0.84/66/93	1/88/90	-	1/100/99
Mindarinae	No	1/100/100	0.92/92/85	-	1/100/100	-	1/100/99
Phloeomyzinae	Yes	-	-	0.88/86/95	1/100/99	0.87/65/74	1/100/100
Thelaxinae	No	1/100/96	0.98/87/70	-	1/100/96	-	1/100/100

the basal node supports were still not high, which suggested that the topology was paraphyletic. Compared with the previous phylogenetic studies concerning similar paraphyletic topology (von Dohlen and Moran, 2000; Martínez-Torres et al., 2001; Zhang and Qiao, 2008; Ortiz-Rivas et al., 2010), there were relatively high nodal support values for the clade E-root (0.66/100/56) in our results, revealing a constant clustering identical to the "E+T" clade (Ortiz-Rivas et al., 2010). We thus conducted the test of the out-group to handle the relationships between the taxa in the "E+T" clade. In all trees obtained from the test of the out-group, only Phloeomyzinae was clustered away from the in-group with high nodal supports. The other 4 out-groups clustered into the in-group with high nodal supports, respectively. We could therefore conclude that there were certain related out-groups, i.e. Hormaphidinae, Anoeciinae, Mindarinae, and Thelaxinae, inserted between the 3 tribes of Eriosomatinae in the phylogeny, allowing us to confirm that Eriosomatinae is not monophyletic but is rather paraphyletic.

There were several possible explanations for this paraphyletic cladogram. Incompetence of datasets is one of the main possible causes, as referred to in previous studies (von Dohlen and Moran, 2000; Ortiz-Rivas et al., 2010). Molecular phylogenetics is generally based on several premises, including that the gene trees can reveal the species tree and that the molecular phylogeny is resolvable (Nei and Kumar, 2000). The aim of phylogenetic studies is to reconstruct a reasonable phylogeny, i.e. with perfect nodal support values, that could be treated as the approximate restoration of the species tree (Rannala and Yang, 2008). However, to date, no results showed even medium nodal support values for the key nodes concerning the relationships among aphid tribes. Thus, more genes are needed in improving the confidence level (high nodal support values) of the reconstructed phylogeny, as suggested in phylogenetic studies using genomic data (Dunn et al., 2006; Rannala and Yang, 2008). Furthermore, it was argued that the molecular phylogeny was originally comb-like, which might be possible evidence for the fast radiation hypothesis among the aphid tribes (von Dohlen and Moran, 2000). Obviously, little phylogenetic information could be preserved during a fast radiation process, and so there were not enough phylogenetic informative sites (i.e. nucleotide substitutions) to resolve the serried short branches. In this scenario, the molecular phylogeny of these fast-radiated taxa was finally arranged in parallel. Additionally, the fast radiation hypothesis had been partially proved in Hormaphidinae through the estimation of divergent time (Huang et al., 2012). Regardless, the classification of these aphid taxa might need adjustment in reference to the results of phylogenetic studies.

4.2. Monophyly of the 3 tribes

The position of Hormaphidinae and Anoeciinae in the major studies was slightly different to that in the test of the out-group. It was implied that Fordini and Pemphigini might be more closely related to Mindarinae and Thelaxinae, and meanwhile Eriosomatini might be more closely related to Hormaphidinae and Anoeciinae. However, these out-groups clustered together in the con-tree of the major studies, and thus the monophyly of the 3 tribes could be revealed. Taking no account of the monophyly of Eriosomatinae, it was notable that our results had several points in common with an earlier cladistic study (Zhang and Chen, 1999). For example, Eriosomatini branched at a basal node, Pemphigini was closely related to Fordini, and Fordini consisted of Fordina and Melaphidina. However, some issues remaining from the earlier study were still unresolved, such as the monophyly of Eriosomatini, Pemphigini, and Fordini.

According to our results, it was implied that Eriosomatini and Fordini were monophyletic, which was consistent with several previous studies (Inbar et al., 2004;

Zhang and Qiao, 2007b, 2008; Sano and Akimoto, 2011). The nodal support values of node E and node F were very high in different reconstructing methods, and these high support values were not affected by the changes of the out-groups. Additionally, Fordini consisted of 2 monophyletic subtribes (Fordina and Melaphidina), which was also consistent with the previous studies (Zhang and Qiao, 2007a, 2008). However, the monophyly of Pemphigini was a little problematic, because node P did not exist in a few results of the test of the out-group. This was mainly attributed to the contribution of *Formosaphis*. *Formosaphis* is a monotypic genus, and the type species, *Formosaphis micheliae* Takahashi 1925, is distributed in India, Japan, and China. It presents a hind wing typical of Pemphigini, but antennal segments III–V of the alatae exhibit many reticulate small secondary rhinaria and irregular sclerotizations identical to the Fordini species (Zhang et al., 1999). Results of the test of the out-group showed that *Formosaphis* clustered in parallel with the other Pemphigini taxa and Fordini. However, the results of the major studies through BI/ML/MP algorithms supported that *Formosaphis* clustered into Pemphigini with relatively high nodal supports, and the Pemphigini taxa formed a monophyletic clade. Therefore, our results reconfirmed that *Formosaphis* should be placed in Pemphigini rather than Fordini, which is consistent with a previous study (Zhang and Qiao, 2007b).

Meanwhile, the monophyly of Eriosomatini, Pemphigini, and Fordini is also supported by morphological characters and biological data (Heie, 1980; Blackman and Eastop, 1994; Zhang et al., 1999). Not only do the apterous viviparous females of the 3 tribes show distinctly different characters (such as the presence and absence of siphunculi, the number of wax gland plates, shape of wax cells, and so on), but so do the alatae viviparous females (such as the shape of secondary rhinaria on antennae, veins of hind wing, tor he number of gonapophyses and wax gland plates) (Table 4). Furthermore, the primary host plants of the 3 tribes show extreme specificity: Eriosomatini on Ulmaceae (*Ulums* and *Zelkova*), Pemphigini (except Prociphilina) on Salicaceae (*Populus*), and Fordini on Anacardiaceae (*Pistacia* and *Rhus*) (Blackman and Eastop, 1994). The secondary host plants of the tribe/subtribes are also distinct, in that Eriosomatini, Pemphigina, Prociphilina, Fordina, and Melaphidina mostly feed on herbal monocots, herbal dicots, Pinaceae, Graminaceae, and mosses, respectively (Zhang et al., 1999). Therefore, the monophyly of the 3 tribes is inferred from molecular data, morphological characters, and biological data.

We have concluded that Eriosomatinae is not monophyletic but is rather paraphyletic, because 4 out-group subfamilies were found to be placed between Eriosomatini and the other 2 tribes in the phylogeny. However, the monophyly of Eriosomatini, Pemphigini, and Fordini is supported by the obtained tree topologies, the morphological features, and the biological data. In addition, the tangled phylogenetic relationship between Fordini and Pemphigini was mainly attributed to the contribution of *Formosaphis*. It was reaffirmed based on more molecular data that the genus should be placed in Pemphigini rather than in Fordini.

Though the support values of the basal nodes might be improved in further studies, the phylogeny of the subfamily was still resolved well in this study. Morphological data could also provide abundant phylogenetic signals, such that it will be valuable to combine molecular and morphological data to reconstruct a total-evidenced phylogeny of the subfamily. In addition, the interrelationships of the "E+T" clade (including Eriosomatinae and 4 related subfamilies) should be the focus of other important work in the future, with more samples and larger datasets.

Table 4. Morphological characters supporting the monophyly of each tribe. Characters follow Heie (1980), Blackman and Eastop (1994), and Zhang et al. (1999). PR = primary rhinaria; SR = secondary rhinaria; WG = wax gland.

Tribes	Apterous viviparous female				Alatae viviparous female			
	Siphunculi	WG	Trochanter	PR	SR	Hind wing vein	WG	Gonapophyses
Eriosomatini	Ring-like, uplifted, a few setae	A central cell region	-	-	Ring-like	2 oblique veins, separated at base	A central cell region	2, often without setae
Pemphigini	Indistinct, pore-like	No central cell, 2–4 rows, a seta	Separated from femur	Mostly ciliated	Narrow, transverse, subring	2, but not separated at base	Honeycomb-like	3
Fordini	Absent or pore-like	No central cell, 6 rows	Fused with f emur	Seldom ciliated	Round, oval, irregular	2, separated at bases or closed	Often absent	2, often with setae

Acknowledgments
We thank the predecessors and colleagues of our lab who accumulated abundant samples for our study. We are indebted to FD Yang for making all the slide preparations and to Dr Z Wang for the help in the experiments. We also thank L Liu for sharing sequences of the out-groups. This work was supported by the National Science Fund for Distinguished Young Scientists (No. 31025024), the National Science Fund for Fostering Talents in Basic Research (No. J1210002), and the Ministry of Science and Technology of the People's Republic of China (MOST Grant No. 2011FY120200).

References

Blackman RL, Eastop VF (1994). Aphids on the World's Trees: An Identification and Information Guide. Wallington, UK: CAB International.

Dunn CW, Hejnol A, Matus DQ, Pang K, Browne WE, Smith SA, Seaver E, Rouse GW, Obst M, Edgecombe GD et al. (2006). Broad phylogenomic sampling improves resolution of the animal tree of life. Nature 452: 745–749.

Foottit RG, Maw HEL, von Dohlen CD, Hebert PDN (2008). Species identification of aphids (Insecta: Hemiptera: Aphididae) through DNA barcodes. Mol Ecol Resour 8: 1189–1201.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59: 307–321.

Hall TA (2004). BioEdit Sequence Alignment Editor 7.0.1. Carlsbad, CA, USA: Isis Pharmaceuticals.

Heie OE (1980). The Aphidoidea of Fennoscandia and Denmark. I. Fauna Entomologica Scandinavica 9: 1–236.

Heie OE (1987). Paleontology and phylogeny. In: Minks AK, Harrewijn P, editors. Aphids: Their Biology, Natural Enemies and Control. Amsterdam, the Netherlands: Elsevier, pp. 367–391.

Huang XL, Xiang-Yu JG, Ren SS, Zhang RL, Zhang YP, Qiao GX (2012). Molecular phylogeny and divergence times of Hormaphidinae (Hemiptera: Aphididae) indicate Late Cretaceous tribal diversification. Zool J Linn Soc-Lond 165: 73–87.

Inbar M, Wink M, Wool D (2004). The evolution of host plant manipulation by insects: molecular and ecological evidence from gall-forming aphids on *Pistacia*. Mol Phylogenet Evol 32: 504–511.

Martínez-Torres D, Buades C, Latorre A, Moya A (2001). Molecular systematics of aphids and their primary endosymbionts. Mol Phylogenet Evol 20: 437–449.

Nei M, Kumar S (2000). Molecular Evolution and Phylogenetics. Oxford, UK: Oxford University Press.

Nieto-Nafría JM, Mier-Durante MP, Pérez Hidalgo N (2011). Register of family-group taxa of Aphidoidea. In: Nieto-Nafría JM, Favret C, editors. Registers of Family-Group and Genus-Group Taxa of Aphidoidea (Hemiptera Sternorrhyncha). León, Spain: Universidad de León, Area de Publicaciones, pp. 21–80.

Normark BB (1999). Evolution in a putatively ancient asexual aphid lineage: recombination and rapid karyotype change. Evolution 53: 1458–1469.

Ortiz-Rivas B, Martínez-Torres D (2010). Combination of molecular data support the existence of three main lineages in the phylogeny of aphids (Hemiptera: Aphididae) and the basal position of the subfamily Lachninae. Mol Phylogenet Evol 55: 305–317.

Ortiz-Rivas B, Moya A, Martínez-Torres D (2004). Molecular systematics of aphids (Homoptera: Aphididae): new insights from the long-wavelength opsin gene. Mol Phylogenet Evol 30: 24–37.

Posada D (2008). jModelTest: Phylogenetic model averaging. Mol Biol Evol 25: 1253–1256.

Rambaut A (2009). FigTree 1.3.1. Available online at http://tree.bio.ed.ac.uk/software/figtree/.

Rannala B, Yang ZH (2008). Phylogenetic inference using whole genomes. Annu Rev Genom Hum G 9: 217–231.

Remaudière G, Remaudière M (1997). Catalogue des Aphididae du Monde. Homoptera Aphidoidea. Paris, France: INRA.

Ren SS, Jiang LY, Qiao GX (2006). Survey of aphids phylogeny. Acta Zootaxonomica Sin 31: 304–310.

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

Sano M, Akimoto S (2011). Morphological phylogeny of gall-forming aphids of the tribe Eriosomatini (Aphididae: Eriosomatinae). Syst Entomol 36: 607–627.

Swofford DL (2003). PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sunderland, MA, USA: Sinauer Associates.

Tamura K, Dudley J, Nei M, Kumar S (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24: 1596–1599.

von Dohlen CD, Kurosu U, Aoki, S (2002). Phylogenetics and evolution of the eastern Asian-eastern North American disjunct aphid tribe, Hormaphidini (Hemiptera: Aphididae). Mol Phylogenet Evol 23: 257–267.

von Dohlen CD, Moran NA (2000). Molecular data support a rapid radiation of aphids in the Cretaceous and multiple origins of host alternation. Biol J Linn Soc 71: 689–717.

Xia X, Xie Z (2001). DAMBE: Data Analysis in Molecular Biology and Evolution. J Hered 92: 371–373.

Xia X, Xie Z, Salemi M, Chen L, Wang Y (2003). An index of substitution saturation and its application. Mol Phylogenet Evol 26: 1–7.

Zhang GX, Chen XL (1999). Study on the phylogeny of Pemphigidae (Homoptera: Aphidinea). Acta Entomol Sin 42: 176–183.

Zhang GX, Qiao GX, Zhong TS, Zhang WY (1999). Fauna Sinica Insecta, Vol. 14. Homoptera: Mindaridae and Pemphigidae. Beijing, China: Science Press.

Zhang HC, Qiao GX (2007a). Molecular phylogeny of Fordini (Hemiptera: Aphididae: Pemphiginae) inferred from nuclear gene EF-1α and mitochondrial gene COI. B Entomol Res 97: 379–386.

Zhang HC, Qiao GX (2007b). Systematic status of the genus *Formosaphis* Takahashi and the evolution of galls based on the molecular phylogeny of Pemphigini (Hemiptera: Aphididae: Eriosomatinae). Syst Entomol 32: 690–699.

Zhang HC, Qiao GX (2008). Molecular phylogeny of Pemphiginae (Hemiptera: Aphididae) inferred from nuclear gene EF-1α sequences. B Entomol Res 98: 499–507.

Appendix. The list of samples according to the 3 tribes of Eriosomatinae and the out-groups. The GenBank accession numbers of each gene are presented. The sampling information of the downloaded sequences is missing.

Species name	Sample number	Collecting locality	Collecting date	GenBank accession numbers			
				COI	*COII*	*EF-1α*	*LWO*
Tetraneura sp.	22389	MONGOLIA: Ulan Bator	22 June 2009	JQ916865	JX536380	JX559469	JX559406
Tetraneura sp.	22400	Beijing City (Changping District)	4 July 2009	JQ916866	JX536381	JX559470	JX559407
Tetraneura sp.	22926	Gansu Province (Yuzhong County)	24 May 2009	JQ916872	JX536369	JX559475	JX559412
Tetraneura sp.	23047	Gansu Province (Maiji Mountain)	2 June 2009	JQ916881	JX536375	JX559481	JX559418
Tetraneura sp.	23081	Gansu Province (Maiji Mountain)	3 June 2009	JQ916885	JX536379	JX559484	JX559421
Tetraneura sp.	16999	Zhejiang Province (Putuo Mountain)	29 May 2005	JX536321	JX536384	JX559487	-
Kaltenbachiella sp.	23140	Sichuan Province (Meigu County)	10 June 2009	JX536308	JX536363	JX559488	JX559425
Eriosoma lanigerum	15412	Tibet (Linzhi Prefecture)	21 August 2003	JQ916894	JX536354	JX559489	JX627604
Eriosoma sp.	22383	MONGOLIA: Ulan Bator	22 June 2009	JX536311	JX536357	-	JX559426
Eriosoma sp.	22920	Gansu Province (Yuzhong County)	24 May 2009	JX536312	-	JX559490	JX559427
Eriosoma sp.	23044	Gansu Province (Tianshui City)	2 June 2009	JX536316	JX536360	-	JX559432
Eriosoma sp.	23142	Sichuan Province (Meigu County)	10 June 2009	JX536318	JX536359	JX559495	JX559435
Colophina arctica	23540	Beijing City (Miyun County)	1 August 2005	JQ916891	JX536364	-	JX559436
Fomosaphis micheliae	18074	Fujian Province (Wuyi Mountain)	22 October 2005	JQ916862	JX536332	DQ779152	JX559437
Prociphilus pini	16169	Beijing City (Miyun County)	13 May 2005	JQ916861	JX536327	DQ779159	JX559438
Prociphilus ligustrifoliae	18235	Yunnan Province (Lijiang City)	27 April 2006	JQ916897	JX627587	JX627602	JX627606
Prociphilus ligustrifoliae	22989	Gansu Province (Wudu County)	28 May 2009	JQ916876	JX536328	JX559496	JX559439
Prociphilus ligustrifoliae	23043	Gansu Province (Tianshui City)	2 June 2009	JQ916880	JX536329	JX559497	JX559440
Prociphilus sp.	Y8936	Hebei Province (Renqiu City)	3 July 2010	JX536290	JX536331	-	JX559441
Prociphilus kuwanai	24365	Henan Province (Miyang Conunty)	9 May 2010	JX536291	JX536330	JX559498	-
Prociphilus caryae	-	-	-	EU701858	DQ005163	DQ005161	-
Thecabius beijingensis	15739	Heilongjiang Province (Mohe County)	31 July 2004	JX536307	JX536386	JX559501	JX559444
Epipemphigus yunanensis	18234	Yunnan Province (Lijiang City)	27 April 2006	JX627585	-	JX627601	JX627605
Epipemphigus imaicus	23130	Sichuan Province (Meigu County)	9 June 2009	JX536303	JX536348	JX559506	JX559450
Pachypappa marsupialis	-	-	-	-	DQ005162	DQ005135	-
Pemphigus poluli	-	-	-	AY522907	AM748713	FM163603	-
Pemphigus monophagous	-	-	-	EU701836	AY182300	DQ779155	-

Appendix. (Continued).

Species name	Sample number	Collecting locality	Collecting date	GenBank accession numbers			
				COI	*COII*	*EF-1α*	*LWO*
Pemphigus populi-transversus	-	-	-	EU701844	AM748720	DQ779157	-
Pemphigus tibetensis	18325	Tibet (Linzhi Prefecture)	22 August 2005	-	JX536340	-	JX559451
Pemphigus borealis	23096	Sichuan Province (Louji Mountain)	5 June 2009	JX536297	JX536343	JX559514	JX559460
Pemphigus bursarius	23097	Sichuan Province (Louji Mountain)	6 June 2009	-	JX536338	JX559515	JX559461
Kaburagia rhusicola	15699	Shaanxi Province (Xixiang County)	26 June 2004	JQ916893	JX536323	DQ499612	JX559465
Schlechtendalia chinensis	15703	Sichuan Province (Emei Mountain)	2 September 2004	JQ916860	JX536326	DQ499619	JX559468
Chaitogeioca sp.	15300	Shaanxi Province (Qishan County)	14 July 2004	JX536320	JX536385	JX559517	-
Aploneura lentisci	-	-	-	AY227083	AY227092	DQ499605	AJ489289
Forda formicaria	-	-	-	AY227076	AF454629	DQ499608	AM996874
Forda marginata	-	-	-	EU701668	AY227098	DQ499609	FM177108
Melaphis rhois	-	-	-	EU701748	FJ215686	FJ215685	-
Schlechtendalia chinensis	-	-	-	JF7001701	AF454628	EU363670	-
Baizongia pistaciae	-	-	-	AY227079	AY227093	DQ499606	AJ489290
Paracletus cimiciformis	-	-	-	AY227089	AY227102	FM163597	FM177109
Geoica utricularia	-	-	-	-	AY227096	FM163600	FM177110
Smynthurodes betae	-	-	-	AY227078	AF454630	FM163598	FM177111
Slavum wertheimae	-	-	-	AY227077	AY227103	DQ499616	-
Floraphis choui	-	-	-	-	EU363665	EU363668	-
Floraphis meitanensis	-	-	-	-	EU363666	EU363669	-
Nurudea shiraii	-	-	-	-	AF454627	EU363679	-
Nurudea yanoniella	-	-	-	-	EU363667	EU363680	-
Meitanaphis flavogallis	-	-	-	-	EU363663	EU363673	-
Meitanapphis microgallis	-	-	-	-	EU363664	EU363674	-
Hormaphis similibetulae	13549	Tibet (Linzhi Prefecture)	5 July 2002	JQ920920	JX627589	JX627595	JX627608
Ktenopteryx eosocallis	14438	Fujian Province (Wuyi Mountain)	7 July 2003	JQ920921	JX627590	JX627596	JX627609
Ceratoglyphina bambusae	14466	Fujian Province (Wuyi Mountain)	10 July 2003	JX627586	JX627588	JX627594	JX627607
Phloeomyzus passerinii	14260	Tibet (Lasa City)	24 August 2003	JQ920929	JX627591	JX627597	JX627611
Mindarus keteleerifoliae	18171	Yunnan Province (Kunming City)	22 April 2006	JQ920925	-	JX627600	JX627610
Mindarus abietinus	-	-	-	FJ668265	-	FM174703	FM177107
Anoecia sp.	-	-	-	-	FM174706	AJ539463	-
Nipponaphis distyliicola	-	-	-	GU978809	AF454626	AF454614	-
Kurisakia querciphila	-	-	-	GU978801	JQ418320	-	-
Thelaxes suberi	-	-	-	-	-	FM174702	AJ489287
Phylloxerina salicis	15434	Yunnan Province (Kunming)	21 April 2006	JQ920928	JX627592	JX627598	-
Pineus armandicola	18168	Yunnan Province (Kunming)	22 April 2006	JQ920909	JX627593	JX627599	-
Adelges laricis	-	-	-	FJ502430	DQ256142	DQ493827	-
Daktulosphaira vitifoloae	-	-	-	AF307423	AF307423	EF073221	AJ489295