Molecular Phylogenetics and Evolution 48 (2008) 667–678

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Molecular phylogenetics of cixiid planthoppers (Hemiptera: Fulgoromorpha): New insights from combined analyses of mitochondrial and nuclear genes

Paula Ceotto [a,*], Gaël J. Kergoat [b], Jean-Yves Rasplus [b], Thierry Bourgoin [a]

[a] Muséum National d'Histoire Naturelle, Laboratoire d'Entomologie and USM 601 & UMR 5202 CNRS, CP 50, F-75231 Paris, France
[b] INRA—UMR CBGP (INRA/IRD/Cirad/Montpellier SupAgro), Campus international de Baillarguet, CS 30016, F-34988 Montferrier-sur-Lez Cedex, France

ARTICLE INFO

Article history:
Received 15 November 2007
Revised 19 March 2008
Accepted 19 April 2008
Available online 26 April 2008

Keywords:
Phylogenetic inference
Cixiidae
Fulgoromorpha
Bayesian inference
Mixture models
Partitioned analyses

ABSTRACT

The planthopper family Cixiidae (Hemiptera: Fulgoromorpha) comprises approximately 160 genera and 2000 species divided in three subfamilies: Borystheninae, Bothriocerinae and Cixiinae, the later with 16 tribes. The current paper represents the first attempt to estimate phylogenetic relationships within Cixiidae based on molecular data. We use a total of 3652 bp sequence alignment of four genes: the mitochondrial coding genes Cytochrome *c* Oxidase subunit 1 (Cox1) and Cytochrome *b* (Cyt*b*), a portion of the nuclear 18S rDNA and two non-contiguous portions of the nuclear 28S rDNA. The phylogenetic relationships of 72 terminal specimens were reconstructed using both maximum parsimony and Bayesian inference methods. Through the analysis of this empirical dataset, we also provide comparisons among different *a priori* partitioning strategies and the use of mixture models in a Bayesian framework. Our comparisons suggest that mixture models overcome the benefits obtained by partitioning the data according to codon position and gene identity, as they provide better accuracy in phylogenetic reconstructions. The recovered maximum parsimony and Bayesian inference phylogenies suggest that the family Cixiidae is paraphyletic in respect with Delphacidae. The paraphyly of the subfamily Cixiinae is also recovered by both approaches. In contrast to a morphological phylogeny recently proposed for cixiids, subfamilies Borystheninae and Bothriocerinae form a monophyletic group.

1. Introduction

The worldwide-distributed planthoppers of the family Cixiidae constitute one of the largest families in the hemipteran infraorder Fulgoromorpha, comprising about 160 genera and 2000 species (Holzinger et al., 2002). Cixiids are small insects, varying from 3 to 11 mm, which are especially diverse in the tropical regions (Holzinger et al., 2002). They are intimately associated to their host plants, which serve for feeding, as a substrate for mating and oviposition, and as a protection against predators (Claridge, 1985; Sforza and Bourgoin, 1998; Wilson et al., 1994). They are mostly phloem-feeders; their nymphs usually feeding on the rootlets of various plants and living underground, whereas the adults feed and reproduce on the surface (O'Brien and Wilson, 1985). The scarce host–plant data available indicate that mono or oligophagous species are frequent, but different species belonging to the same genus may feed on distantly related plant families (Wilson et al., 1994). Their piercing and sucking feeding habits, as well as the fact that they feed on a specific plant tissue, transform them into potential vectors of phytopathogenic organisms. Several species are known as vectors of viruses, phytoplasmas and bacterium-like organisms. In Europe, several crops (beet, lavender, maize, strawberry and vineyard) suffer from different diseases caused by micro-organisms vectored by cixiid species (Danet et al., 2003; Jović et al., 2007; Sémetey et al., 2007; Sforza et al., 1998). The vectoring capacities of Cixiidae entail the need for a better understanding of their evolutionary diversification, and a phylogenetic framework might help to better understand their relationships with the phytopathogenic organisms they transmit.

Fossil records of Fulgoromorpha date from as early as the Late Permian (Shcherbakov, 2002), and from the Lower Cretaceous for Cixiidae (Grimaldi et al., 2002; Szwedo, 2004). While the monophyly of Fulgoromorpha is well supported by morphological and molecular data, that of Cixiidae has been questioned several times (Asche, 1988; Bourgoin et al., 1997; Urban and Cryan, 2007). Among the extant lineages, Cixiidae and Delphacidae are considered to be the earliest derived within the infraorder (Bourgoin et al., 1997; Urban and Cryan, 2007). Asche (1985, 1988) pointed to the absence of clear synapomorphies to define the family, mentioning that the great amount of plesiomorphy observed in cixiid characters makes it difficult to resolve the relationships between Cixiidae and Delphacidae. A maximum parsimony analysis (MP) of a molecular data set designed to estimate the relationships among Fulgoromorpha families resulted in a paraphyletic Cixiidae,

* Corresponding author.
E-mail addresses: ceotto@mnhn.fr, paulaceotto@gmail.com (P. Ceotto).

1055-7903/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2008.04.026

with Delphacidae arising from within the former family (Urban and Cryan, 2007). However, these results were not strongly supported by Bremer and non-parametric bootstrap values. Although, in the same study, Bayesian inference (BI) methods recovered a monophyletic Cixiidae, the latter finding can hardly be generalized because only four cixiid species were included in the analysis. Ceotto and Bourgoin (2008) provided the first phylogenetic hypotheses for Cixiidae based on cladistic analyses of morphological characters from 50 representative species belonging to all subfamilies and tribes. Though the monophyly of Cixiidae was recovered in their analyses, this result was not strongly supported neither by node support values nor statistical tests. Hence it appears that further studies are needed to clarify whether the Cixiidae are monophyletic or not.

Very little is known about the phylogenetic relationships within Cixiidae. Extant family subgroups have been proposed on basis of a few supposed synapomorphies, and some of them have been proposed in opposition to the group showing the putative synapomorphy (Emeljanov, 1989, 2002). Many tribes have been named without the provision of a list of characters describing the group (Emeljanov, 2002). The family is currently subdivided into three subfamilies Borystheninae, Bothriocerinae, and Cixiinae, the later including 16 tribes (Holzinger et al., 2002; Szwedo and Stroinski, 2002). To the exception of the small tribes, most of Cixiinae recognized tribes have large distributions, comprising at least three biogeographic regions (Holzinger et al., 2002). Emeljanov (2002) proposed an intuitive phylogeny for two subfamilies and most of Cixiinae tribes. Ceotto and Bourgoin (2008) phylogenetic analyses of 85 morphological characters indicated that morphological data alone is not appropriate to address the question of Cixiidae monophyly or that of its subfamilies and tribes, as a high level of homoplasy tends to mask genuine apomorphic features within Cixiidae.

To go further in our understanding of the phylogenetic relationships of the family and to better assess the monophyly of the family, the subfamilies and some of the tribes, we carried molecular phylogenetic analyses using data from two mitochondrial (Cox1 and Cyt*b*) and two nuclear genes (18S and 28S). The phylogenetic relationships among taxa were reconstructed using both BI and MP inference methods. In addition, we have chosen to use the heterogeneous nature of our molecular dataset, which includes either coding or non-coding fragments from distinct parts of the genome (mitochondrial or nuclear), as an opportunity to compare the two distinct approaches proposed to account for within data heterogeneity under a Bayesian framework. Because both approaches allow subsets of the data to evolve under distinct models and parameters, they are expected to increase the phylogenetic accuracy (Brandley et al., 2005; Nylander et al., 2004; Pagel and Meade, 2004, 2005). In the first approach (hereby referred to "*a priori* partitioning"), one needs to set partitions *a priori*, based for example on gene identities, codon position, and stem and loops locations when dealing with ribosomal genes. However, this approach cannot account for different patterns of substitution within a specific partition scheme. A second approach (hereby referred to "mixture model") was proposed by Pagel and Meade (2004) who described a mixture model for detecting pattern-heterogeneity without the need to define partitions *a priori*. In the later method, two or more qualitatively different models are fitted to each site in a gene-sequence alignment, without specifying in advance the nature of the models, their relative probabilities, or having knowledge of which sites are best fit by which model (Pagel and Meade, 2004, 2005). The main questions we intend to answer in this paper are: (1) is the Cixiidae family monophyletic? and (2) for our dataset, how do mixture models perform in comparison to *a priori* partitioning strategies?

2. Material and methods

2.1. Taxonomic sampling, extraction, amplification and sequencing

Taxonomic sampling consisted of 72 specimens representing 55 cixiid species, two species of Achilidae, four species of Delphacidae, one species of Kinnaridae, two species of Meenoplidae, and one species of Tettigometridae (first dataset). As Cixiidae are considered as one of the earliest diversified families within Fulgoromorpha, analyses with the following additional non-Fulgoromorpha outgroups were also carried: *Aetalion reticulatum* (Aetalionidae), *Froggattoides typicus* and *Pauropsalta corticinus* (Cicadidae), *Pateena polymitarior* (Schizopteridae), and *Philaenus spumarius* (Aphrophoridae) (second dataset). Within Cixiidae, 27 genera from the three recognized subfamilies were included. Ingroup taxa were selected to maximize the sampled taxonomic and biogeographic diversity of Cixiidae subfamilies and tribes. Unfortunately, the overlap between the sampling used in the present study and that used in the recent morphological phylogeny of Cixiidae (Ceotto and Bourgoin, 2008) is not large enough (only eight species in common) to allow us to carry a combined analysis. The main part of the sampling used in the morphological study came from loans of old, dry specimens from diverse Museum collections. The specimens used in the present study are listed in Table 1. Prior to DNA extraction, genitalia were removed from adults and stored as vouchers at the Muséum National d'Histoire Naturelle (MNHN) collection in Paris. DNA was extracted from alcohol-preserved specimens using Puregene DNA isolation kits (Gentra Systems, Minneapolis, USA). Each sequence was obtained from the DNA of a single specimen. In addition, sequences of two Cixiidae and one Delphacidae species available in GenBank were used in the analyses (the corresponding Accession numbers are listed in Table 1).

Portions of the mitochondrial genes Cytochrome *c* Oxidase subunit 1 (Cox1) and Cytochrome *b* (Cyt*b*), a portion of the 18S rDNA (18S) and two non-contiguous portions of the 28S rDNA (28S) corresponding to domains D4–D5 and D6–D7 were amplified with primers listed in Table 2. All polymerase chain reactions (PCR) were conducted in 50 μl reaction volume, containing 2.5–1.5 mM $MgCl_2$, 0.08 mM of each deoxyribonucleotide triphosphate (dNTP), 0.7 μM of each primer and 1 U *Taq* DNA polymerase (Qiagen GmbH, Germany). The PCR cycling program was 3 min at 94 °C, 30 cycles of 1 min at 48–60 °C (48 °C for 2183-UEA8 primers, 54 °C for CB1-CP2 and D4–D5 primers, 58 °C for SSE-SSM primers and 60 °C for both pairs of 18S rDNA primers), and 1–1.5 min at 72 °C, followed by a final step of 7 min at 72 °C. Sequencing was carried out with an ABI 3730 automated sequencer (Applied Biosystems) on both strands. The same primers used for amplification were used for the sequencing reactions. The resulting sequences were assembled using Bioedit version 7.0.5.3 (Hall, 1999) and deposited in GenBank (Accession numbers EU183547–EU183743). As indicated in Table 1, PCR failures have resulted in some missing data entries.

2.2. Sequence alignment

Unlike the sequences of coding genes (Cox1 and Cyt*b*), the sequences of ribosomal genes exhibited variations in length. Their alignment was performed using ClustalX (Thompson et al., 1997), with default options, and then reviewed and corrected by eye. Highly variable regions of 18S and 28S that differed in sequence length across the sampled taxa were removed due to extreme ambiguity in possible alignments. Preliminary analyses (not shown) have indicated that removal had only minor effects on the results of phylogenetic analyses. Two non-contiguous regions with a combined length of 50 bp of the D4–D5 domain of the 28S, one 400 bp region of the D6–D7 domain of the 28S, and a region of 70 bp of the 18S were thus excluded. For all genes but COI ($P = 0.02$), no deviation from stationa-

Table 1
Taxon sampling

Subfamily Tribe Species	Collection location	GenBank Accession No.				
		COI	Cyt*b*	18S	28S (D4–D5)	28S (D6–D7)
Borystheninae						
Borysthenes sp.	Thailand	EU183610		EU183557	EU183686	EU183731
Bothriocerinae						
Bothriocera sp. 1	Martinique	EU183603	EU183642	EU183577	EU183670	EU183711
Bothriocera sp. 2	Belize	EU183604		EU183581	EU183675	EU183716
Cixiinae						
Andini Emeljanov, 2002						
Andes insolitus Muir, 1925	Singapour					EU183727
Andes insolitus Muir, 1925	Singapour				EU183684	
Andes simplex Muir, 1925	Cambodia	EU183608		EU183568	EU183685	EU183729
Andes sp.	Laos	EU183597	EU183640	EU183547	EU183663	EU183703
Cixiini Spinola, 1839						
Achaemenes intersparsus Jacobi, 1917	Madagascar	EU183598		EU183575		EU183704
Cixius nervosus Linne, 1758	France	EU183619	EU183656	EU183586	EU183692	
Cixius pallipes Fieber, 1876	France	EU183602	EU183641		EU183666	
Cixius similis Kirschbaum, 1868	France	EU183620	EU183657	EU183588	EU183693	EU183740
Cixius wagneri China, 1942	France	EU183621	EU183659		EU183694	
Tachycixius pilosus (Olivier, 1791)	France		EU183637	EU183587		EU183739
Tachycixius pilosus (Olivier, 1791)	Pays-Bas	AF304407				
Trirhacus discrepans Fieber, 1876	France	EU183612	EU183652	EU183584	EU183689	EU183734
Duiliini Emeljanov, 2002						
Duilius tenuis Stål, 1858	Namibia					EU183700
Eucarpiini Emeljanov, 2002						
Eucarpia granulinervis (Muir, 1913)	Australia		EU183636			
Mnemosynini Szwedo, 2004						
Mnemosyne sp.	Singapour			EU183556		
Oecleini Muir, 1922						
Borbonomyndus pandanicola Attié, Bourgoin & Bonfils, 2002	Reunion Island	EU183593		EU183571		EU183735
Colvanalia taffini (Bonfils, 1983)	Vanuatu	EU183613		EU183560	EU183690	
Eumyndus metcalfi Synave, 1956	Madagascar		EU183626			
Haplaxius crudus (Van Duzee, 1907)	USA				EU183676	EU183717
Haplaxius crudus (Van Duzee, 1907)	USA	EU183606		EU183553	EU183680	EU183721
Haplaxius crudus (Van Duzee, 1907)	Mexico	EU183616			EU183691	
Haplaxius deleter (Kramer, 1979)	Costa Rica	EU183605	EU183631	EU183552	EU183679	EU183720
Haplaxius skarphion (Kramer, 1979)	Mexico			EU183570	EU183682	EU183725
Nesomyndus australis Jacobi, 1917	Madagascar	EU183599				
Nymphocixia caribbea Fennah, 1971	Cuba	EU183615		EU183561		
Oecleus productus Metcalf, 1923	USA		EU183647		EU183678	EU183719
Oecleus sp.	Belize		EU183649		EU183662	
Pentastirini Emeljanov, 1971						
Hyalesthes scotti (Ferrari, 1882)	France			EU183565		
Melanoliarus humilis (Say, 1830)	Canada	EU183611	EU183651	EU183559	EU183688	EU183733
Melanoliarus placitus (Van Duzee, 1912)	USA	EU183607				EU183724
Melanoliarus vicarius (Walker, 1851)	USA	EU183601		EU183550		EU183699
Melanoliarus sp. 1	USA				EU183669	EU183709
Melanoliarus sp. 2	French Guiana				EU183672	EU183713
Melanoliarus sp. 3	Belize			EU183554		EU183722
Melanoliarus sp. 4	Virgin Islands		EU183632	EU183555	EU183683	EU183726
Oliarus hamatus Löcker, 2006	Australia	EU183617	EU183654	EU183562	EU183661	EU183738
Oliarus hamatus Löcker, 2006	Australia		EU183653			
Oliarus nosibeanus Jacobi, 1917	Madagascar	EU183591	EU183627	EU183574		
Oliarus sp. 1	Thailand		EU183635	EU183558	EU183687	EU183732
Oliarus sp. 2	New Caledonia	EU183600		EU183549	EU183665	EU183706
Oliarus sp. 3	Cambodia	EU183595	EU183650	EU183569	EU183698	
Oliarus sp. 4	Cambodia	EU183609	EU183634		EU183660	EU183730
Ozoliarus sp.	Australia	EU183618	EU183655	EU183563		EU183702
Pentastiridius sp.	France		EU183630		EU183668	EU183708
Pentastiridius sp.	France	EU183594	EU183624		EU183697	EU183743
Reptalus cuspidatus (Fieber, 1876)	Russia		EU183629		EU183667	EU183707
Reptalus panzeri (Low, 1883)	France			EU183585		EU183737
Reptalus quinquecostatus (Dufour, 1833)	France			EU183564		
Pintaliini Metcalf, 1938						
Cubana sp. 1	Martinique		EU183638	EU183576		EU183710
Cubana sp. 2	Belize	EU183590	EU183645	EU183551	EU183673	EU183714
Cubana sp. 2	Belize		EU183644			
Cubana sp. 3	Virgin Islands				EU183677	EU183718
Pintalia alta Osborn, 1935	Virgin Islands			AY744804		AY744838
Pintalia bicaudata Muir, 1934	French Guiana		EU183648	EU183582	EU183681	EU183723
Pintalia sp. 1	French Guiana			EU183578	EU183671	EU183712

(*continued on next page*)

Table 1 (*continued*)

Subfamily Tribe Species	Collection location	GenBank Accession No.				
		COI	Cyt*b*	18S	28S (D4–D5)	28S (D6–D7)
Pintalia sp. 2	Belize				EU183674	EU183715
Pintalia sp. 2	Belize		EU183646	EU183580		
Pintalia sp. 3	French Guiana	EU183592	EU183643	EU183579		
Incertae sedis						
Meenocixius virescens Attié, Bourgoin & Bonfils, 2002	Reunion Island	EU183614	EU183639	EU183572		EU183736
Achilidae						
Achilidae gen.sp.	Laos	EU183596	EU183623	EU183573		
Cixidia parnasia (Stål, 1859)	France		EU183658			
Delphacidae						
Conomelus anceps (Germar, 1821)	France			EU183548		EU183701
Delphax inermis (Ribaut, 1934)	France				EU183696	EU183742
Megamelus notula (Germar, 1830)	France	EU183622		EU183566	EU183695	EU183741
Notodelphax gillettei	USA					DQ532594
Kinnaridae						
Kinnaridae gen. sp.	Thailand		EU183633	EU183583		EU183728
Meenoplidae						
Nisia sp.	Madagascar		EU183628		EU183664	EU183705
Meenoplidae gen. sp.	Laos		EU183625			
Tettigometridae						
Tettigometra longicornis Signoret, 1866	France	EU183589		EU183567		

Table 2
Names, sequences, and references of primers used

Gene	Name of primer	Sequence of primer (5′ → 3′)	Reference
COI	2183	CAACATTTATTTTGATTTTTTGG	Simon et al. (1994)
	UEA 8	AAAAATGTTGAGGGAAAAATGTTA	Lunt et al. (1996)
Cyt*b*	CB1	TATGTACTACCATGAGGACAAATATC	Jermiin and Crozier (1994)
	CP2	CTAATGCAATAACTCCTCC	Harry et al. (1998)
18S	574	GCCGCGGTAATTCCAGCT	Bourgoin et al. (1997)
	E21	CTCCACCAACTAAGAACGG	
	18S-mid	GATACCGCCCTAGTTCTAACC	
	2200	CGGCAGGTTCACCTACGG	
28S (D4–D5)	D4–D5f	CCCGTCTTGAAACACGGACCAAGG	Belshaw and Quicke (2002)
	D4–D5r	GTTACACACTCCTTAGCGGA	
28S (D6–D7)	28S EE	CCGCTAAGGAGTGTGTAA	Cryan et al. (2000)
	28S MM	GAAGTTACGGATCTARTTTG	

rity in nucleotide composition was detected by the χ^2 tests implemented in PAUP* version 4.0b10 (Swofford, 2003).

2.3. Maximum parsimony

Incongruence among all pairs of the studied genes was assessed by the incongruence length difference test (ILD; Farris et al., 1994), as implemented in PAUP*, with all uninformative characters excluded (Lee, 2001). Since the results of the ILD tests were not significant ($P > 0.05$), all sequences were concatenated in a single matrix. In addition, due to the presence of missing data entries resulting from PCR and sequencing failures, the combined matrix allowed us to broaden the scope of the separate analyses in order to obtain more synthetic and direct comparisons. These PCR and sequencing failures were a recurrent problem in this study. Some missing data may be due to damaged DNA of old samples conserved in poor conditions. However, in other cases, PCR failed for a few genes only and failure was consistent over several trials using different DNA extracts, concentrations, and PCR conditions. Wiens (2003, 2005, 2006) has shown through simulation studies that the benefits of including taxa with missing data in phylogenetic analyses usually overcome the associated disadvantages. For this reason we included the taxa with missing entries in the matrix.

MP analyses were conducted under TNT version 1.1 (Goloboff et al., 2003), which implements new algorithms of tree search developed for improving search efficiency and speeding up phylogenetic analyses. Analyses were performed with all substitutions equally weighted, gaps treated as a fifth character, maximum number of trees set to 1000, and the "new technology" algorithms: random sectorial searches with default options, 40 cycles of drift accepting suboptimal rearrangements with maximum fit difference of 2, 40 cycles of ratchet and five rounds of tree-fusing. Default settings were used for other options. Confidence in each node was assessed by 1000 replicates of non-parametric bootstrapping (Felsenstein, 1985) and Bremer support (BS) indices (Bremer, 1988, 1994). Partitioned Bremer support (PBS) values (Baker and DeSalle, 1997) were also estimated to assess the respective contribution of each gene to the support of nodes.

2.4. Bayesian inference

BI analyses were carried out using MrBayes version 3.1.2 (Ronquist and Huelsenbeck, 2003) and BayesPhylogenies version 1.0 (Pagel and Meade, 2004). In MrBayes analyses, three partitioning strategies were defined *a priori*: strategy P1, which corresponds to an analysis without partitions; strategy P2, which implements a partition for each gene (with the two non-contiguous regions

of 28S being treated as different genes); and strategy P3, which uses one partition for each nuclear gene, and three partitions for the mitochondrial genes (one partition per codon position was used). The lack of a secondary structure model for hemipteran 28S rDNA prevented us from partitioning these genes with reference to stems and loops. As emphasized by Niehuis et al. (2006), the use of extant secondary structure models in distantly related groups might lead to errors in secondary structure estimations and interpretations of stems and loops. To standardize the treatment of all ribosomal genes and to avoid possible bias, we decided not to take into account the secondary structure of 18S rDNA, for which a hemipteran-based model has been proposed (Ouvrard et al., 2000). Best-fit models of evolution for each gene were determined by using the Akaike information criterion (AIC), as implemented in Modeltest version 3.0.6 (Posada and Crandall, 1998). The General time reversible (GTR) + I + G model (Gu et al., 1995; Yang, 1994) was indicated as the best-fit model for all genes, and so this model was used in all subsequent analyses. Two independent BI runs were carried out, each one with four chains (with incremental heating) of 5,000,000 generations, with random starting trees, default priors (but temperature set to 1.0) and trees sampled every 100 generations. Log-likelihood scores were plotted to determine the number of trees to be discarded as burn-in. A conservative burn-in of 12,500 trees (corresponding to 1,250,000 generations) was adopted for all partitioning strategies and the remaining trees used to construct the majority-rule consensus trees. As suggested by Brandley et al. (2005), we have used a Bayes factor (B_F)-based statistics ($2\ln B_F$) to choose among the different partitioning strategies. Following Kergoat et al. (2007), a more conservative threshold was used to take into account the number of parameters of each competing partitioning strategy (instead of using a fixed threshold of 10).

Phylogenetic relationships were also inferred through the use of mixture models, as implemented in the program BayesPhylogenies (Pagel and Meade, 2004). Mixture models accommodate cases in which different sites in the alignment evolve in qualitatively distinct ways. Unlike traditional partitioned strategies, mixture models account for data heterogeneity without requiring prior-knowledge of within data differences in evolutionary patterns. Analyses with $nQ + \Gamma$ mixture models, where n varied between one and six independent rate matrices (Qs), were performed. As suggested by Pagel and Meade (2004), we have also used a GTR model in our dataset. The Markov chain Monte Carlo (MCMC) method was used with four chains of 5,000,000 iterations with print frequency each 1000 iterations. A summary of the parameters given in BayesPhylogenies output can be obtained using the *sump* command in MrBayes. Using this command in MrBayes also allows for standardizing the calculation of the harmonic means of analyses using MrBayes and BayesPhylogenies.

B_F were also used for comparisons among analyses with the different rate matrices, as well as among the *a priori* partitioning strategies and the mixture model approach. In all BI analyses, the robustness of clades was assessed by clade posterior probabilities (CPP).

2.5. Hypothesis testing

An *a priori* hypothesis that follows the recently proposed classification for Cixiidae (Holzinger et al., 2002) was compared to *a posteriori* hypothesis (i.e. one of the topologies obtained here; the unconstrained tree). The alternative constrained topologies—at least two alternative trees for each hypothesis—were built using Mesquite version 1.12 (Maddison and Maddison, 2006). Comparisons were made with the likelihood-based non-parametric Shimodaira–Hasegawa test (SH test; Shimodaira and Hasegawa, 1999), which enables comparisons between *a priori* (not derived from the data being tested) and *a posteriori* hypotheses (Goldman et al., 2000). The re-estimated log-likelihood (RELL) method (Kishino et al., 1990), as implemented in PAUP*, was used to resample the log-likelihoods (1000 replicates) in the SH tests.

3. Results

Main results of the BI analyses with the first and the second datasets are very similar. However, the addition of the more distantly related outgroups seem to have had a negative effect on the parsimony results (see also Smith, 1994; Lyons-Weiler et al., 1998), as the trees derived from the MP analysis with the second dataset are poorly resolved in comparison to those derived from the first dataset. For that reason, we only present and discuss the results of the analyses with the dataset including only the Fulgoromorpha outgroups here.

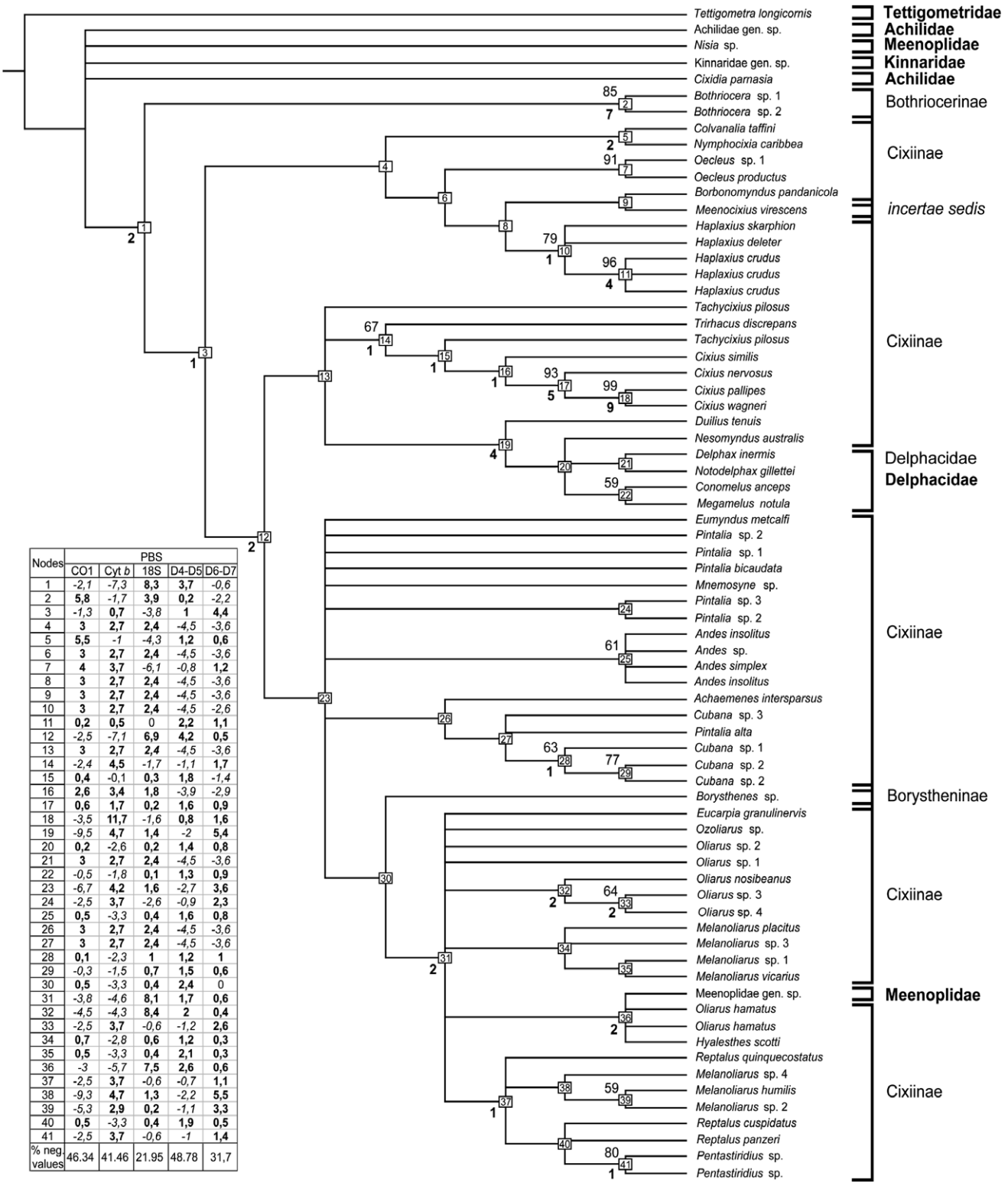
3.1. Maximum parsimony

The parsimony analysis resulted in 18 most-parsimonious topologies (5302 steps; consistency index of 0.39; retention index of 0.54). The majority rule consensus tree, with non-parametric bootstrap (>50%), BS and PBS values is shown in Fig. 1. Interestingly, the important number of negative values of the PBS that are recovered for all loci indicates a high level of conflicting data, which have not been detected by the previous ILD tests. These negative PBS values are scattered throughout the tree and among the different loci, hence no particular area of conflict could be detected in the tree.

3.2. Bayesian inference

For the partitioning strategies with partitions determined *a priori*, the most complex model (P3) appears as optimal (Table 3). B_F-based statistics also indicate that partitioning the data by genes provides a better fit than not partitioning at all ($2\ln B_F = 936.6$; $P = 0.00$). The tree derived from the analysis with the *a priori* partitioning strategy P3 is shown in Fig. 2.

In the comparisons between mixture models, we have found that likelihood scores systematically increased with the number of estimated matrices (Table 4). The B_F-based statistics are also always positive, and significant ($P < 0.05$), in the comparisons that involve analyses with n and $n + 1$ matrices. We need to look at other parameters to decide how many matrices we are going to keep. Overparameterization may be indicated by precipitous decline in the improvements of overall likelihood scores, marked increase in the average standard deviation of the rate parameters, and small weights assigned to superfluous matrices (Pagel and Meade, 2004, 2005). Increase in score is not very important when going from 4Q to 5Q and 5Q to 6Q, but the average standard deviation of rate parameters do not abruptly increase (Table 4). The behavior of the assigned weights is somewhat challenging. Whereas in 3Q assigned weights for the three matrices are proportionate, in 4Q two matrices received considerably low weights (Table 5). However, the improvement in likelihood scores when going from 3Q to 4Q is an important one, so that based on likelihood scores 4Q should be selected. These results indicate that the choice of the number of matrices to estimate may be rather subjective. However, in our dataset only small differences among trees estimated under 1–6$Q + \Gamma$ models were detected, and they involved alternative placements of weakly supported nodes. For this reason, and because we compare mixture models results to *a priori* partitioning strategies with up to six partitions, we selected the 6$Q + \Gamma$ model (Fig. 3).



Nodes	PBS				
	CO1	Cyt *b*	18S	D4-D5	D6-D7
1	*-2,1*	*-7,3*	**8,3**	**3,7**	*-0,6*
2	**5,8**	*-1,7*	**3,9**	**0,2**	*-2,2*
3	*-1,3*	**0,7**	*-3,8*	**1**	**4,4**
4	**3**	**2,7**	**2,4**	*-4,5*	*-3,6*
5	**5,5**	*-1*	*-4,3*	**1,2**	**0,6**
6	**3**	**2,7**	**2,4**	*-4,5*	*-3,6*
7	**4**	**3,7**	*-6,1*	*-0,8*	**1,2**
8	**3**	**2,7**	**2,4**	*-4,5*	*-3,6*
9	**3**	**2,7**	**2,4**	*-4,5*	*-3,6*
10	**3**	**2,7**	**2,4**	*-4,5*	*-2,6*
11	**0,2**	**0,5**	0	**2,2**	**1,1**
12	*-2,5*	*-7,1*	**6,9**	**4,2**	**0,5**
13	**3**	**2,7**	**2,4**	*-4,5*	*-3,6*
14	*-2,4*	**4,5**	*-1,7*	*-1,1*	**1,7**
15	**0,4**	-0,1	**0,3**	**1,8**	*-1,4*
16	**2,6**	**3,4**	**1,8**	*-3,9*	*-2,9*
17	**0,6**	**1,7**	**0,2**	**1,6**	**0,9**
18	*-3,5*	**11,7**	*-1,6*	**0,8**	**1,6**
19	*-9,5*	**4,7**	**1,4**	*-2*	**5,4**
20	**0,2**	*-2,6*	**0,2**	**1,4**	**0,8**
21	**3**	**2,7**	**2,4**	*-4,5*	*-3,6*
22	*-0,5*	*-1,8*	**0,1**	**1,3**	**0,9**
23	*-6,7*	**4,2**	**1,6**	*-2,7*	**3,6**
24	*-2,5*	**3,7**	*-2,6*	*-0,9*	**2,3**
25	**0,5**	*-3,3*	**0,4**	**1,6**	**0,8**
26	**3**	**2,7**	**2,4**	*-4,5*	*-3,6*
27	**3**	**2,7**	**2,4**	*-4,5*	*-3,6*
28	**0,1**	*-2,3*	**1**	**1,2**	**1**
29	*-0,3*	*-1,5*	**0,7**	**1,5**	**0,6**
30	**0,5**	*-3,3*	**0,4**	**2,4**	0
31	*-3,8*	*-4,6*	**8,1**	**1,7**	**0,6**
32	*-4,5*	*-4,3*	**8,4**	**2**	**0,4**
33	*-2,5*	**3,7**	*-0,6*	*-1,2*	**2,6**
34	**0,7**	*-2,8*	**0,6**	**1,2**	**0,3**
35	**0,5**	*-3,3*	**0,4**	**2,1**	**0,3**
36	*-3*	*-5,7*	**7,5**	**2,6**	**0,6**
37	*-2,5*	**3,7**	*-0,6*	*-0,7*	**1,1**
38	*-9,3*	**4,7**	**1,3**	*-2,2*	**5,5**
39	*-5,3*	**2,9**	**0,2**	*-1,1*	**3,3**
40	**0,5**	*-3,3*	**0,4**	**1,9**	**0,5**
41	*-2,5*	**3,7**	*-0,6*	*-1*	**1,4**
% neg. values	46.34	41.46	21.95	48.78	31,7

Fig. 1. Majority rule consensus of 18 most-parsimonious trees (length = 5302; CI = 0.39; RI = 0.54) of the Cixiidae family. Bootstrap values (>50%) are above branches and Bremer decay indexes below branches. Node numbers are indicated. The numbers in the table are the contribution of the Cox1, Cyt*b*, 18S rDNA, domain D4–D5 of 28S rDNA, and domain D6–D7 of 28S rDNA partitions to the Bremer support values (results of the Partitioned Bremer Support analysis). The table also indicates the percentage of negative values for each partition.

Comparisons between 1–6$Q + \Gamma$ mixture models with three different *a priori* partition strategies of the data indicate that mixture models perform better. Analyses with five and six matrices have significantly higher B_F than the optimal *a priori* partitioning strategy (Table 6). For that reason, we used the 6$Q + \Gamma$ topology as the unconstrained tree for the tests of alternative topologies hypotheses. As we use the SH test (a model-based test), the use of a tree obtained through a model-based method as the unconstrained topology seems more methodologically coherent to us. In addition, this tree was also preferred to the MP trees because

Table 3
Comparison of results of different *a priori* partitioning strategies using B_F

Partitioning strategy	Number of partitions	Mean log-likelihood	Harmonic mean	$2\ln B_F$
P1	1	−26,295	−26,337.4	
P2	5	−25,827.4	−25,869.1	936.6
P3	6	−25,564.7	−25,611.7	514.8

Last column corresponds, respectively, to comparisons between P1 and P2, as well as P2 and P3 strategies.

the results of the MP analyses have been likely biased by the incongruence in the dataset, detected by the PBS.

3.3. Recovered topologies

In both MP and BI analyses (P3 strategy and $6Q + \Gamma$ model), Cixiidae is recovered as paraphyletic due to the placement of members of the family Delphacidae within the former family (Figs. 1–3). The three topologies differ in that the MP topology also recovers one Meenoplidae species within Cixiidae (Fig. 1), and in that the Oecleini *Eumyndus metcalfi* is not placed within Cixiidae in the P3 topology (Fig. 2). In the three topologies, the subfamily Cixiinae is also paraphyletic in respect to Delphacidae, and also because the subfamily Borystheninae is placed within it in the MP and both BI topologies. In the BI topologies the placement of the subfamily Bothriocerinae within Oecleini also renders Cixiinae paraphyletic. The subfamily Bothriocerinae and the tribe Andiini are the only groups recovered as monophyletic by both methods. In the BI analyses, Bothriocerinae arises from within the tribe Oecleini, the earliest derived Cixiidae tribe in the BI topology (Figs. 2 and 3). In the MP topology, Bothriocerinae is the most anciently diverged group within Cixiidae, followed by Oecleini (Fig. 1). In the three topologies, the *insertae saedis* genus *Meenocixius* arises from within Oecleini. This tribe is paraphyletic in both BI and in the MP analyses, as the species *Nesomyndus australis* does not cluster with the remaining Oecleini species in all trees. Pentastirini is paraphyletic in the BI and MP topologies, with the Eucarpiini *Eucarpia granulinervis* arising from within Pentastirini. The tribe Cixiini is also paraphyletic in both analyses, as *Achaemenes intersparsus* does not cluster with the remaining representative species of the tribe. The tribe Pintaliini is recovered monophyletic only in the BI analyses (Figs. 2 and 3).

3.4. Hypothesis testing

SH tests failed to reject alternative hypotheses of a monophyletic Cixiidae, as well as monophyletic Cixiini ($0.27 > P > 0.18$) and Pentastirini ($0.01 > P > 0.05$). This test also failed to reject the alternative hypothesis of *N. australis* belonging to Oecleini ($0.39 > P > 0.19$). With regard to the relationships between Bothriocerinae and Oecleini, the SH tests failed to reject the hypothesis of Bothriocerinae as sister group to Oecleini ($P = 0.14$). However, the alternative hypotheses of Oecleini and Bothriocerinae not forming a monophyletic unit were strongly rejected ($P = 0.00$). The test also rejected the hypothesis of Borystheninae and Bothriocerinae forming a monophyletic group, both when this group is tested as part of Oecleini and when it is elsewhere on the tree ($P = 0.00$).

4. Discussion

4.1. MP vs BI results

Major groups recovered by both MP and BI analyses are in general agreement (Figs. 1–3). Overall, clades recovered in MP are less supported than in BI, in agreement with results of numerous studies that have highlighted discrepancies between non-parametric bootstrap and CPP values (e.g., Alfaro et al., 2003; Erixon et al., 2003). However, even when we take these biases into account and consider as well supported those clades with non-parametric bootstrap $\geqslant 70\%$ (Hillis and Bull, 1993) and CPP $\geqslant 95\%$ (Brandley et al., 2005), MP recovered clades are slightly less supported. In addition, MP more frequently failed to recover clades whose monophyly is widely accepted, for example Delphacidae (Asche, 1988, 1990; Urban and Cryan, 2007). For heterogeneous datasets, simulations presented by Simmons et al. (2006) indicate that when parameters that accommodate heterogeneity are included in the model used to carry the analyses, BI performs generally better than MP. One of the reasons that might explain why BI performed better than MP is the possible incongruence among the different loci used, as suggested by the important level of PBS negative values for all loci (Fig. 1). The better performance of partitioned BI analyses over the non-partitioned ones might also indicate some level of incongruence in the data matrix. Although this conflict has not been detected by the ILD test, the later has been frequently criticized, either because it may give false positive results, but also because it is too conservative and may not detect incongruence among data partitions in a certain number of conditions (Darlu and Lecointre, 2002; Hipp et al., 2004; Ramírez, 2006).

Missing data may also have affected topologies resolution. Wiens (2003, 2005, 2006) showed by simulations that as long as the dataset has enough characters to enable accurate placement of incomplete taxa, missing data might not produce misleading effects on phylogeny reconstructions. However, it is difficult to assess the extent to which Wiens (2003, 2005, 2006) conclusions may be extrapolated to our dataset. While the present dataset is already large, it may be that the amount of scored characters for some of the taxa included in the analyses was not large enough to assure their accurate place in tree topologies. Another effect of missing data is that they exacerbate the problem of long-branch attraction (Wiens, 2006). This would explain the fact that MP performed worse than BI, as model-based methods tend to be inherently more robust to the problem of long-branch attraction (Huelsenbeck, 1995; Swofford et al., 2001; Holder and Lewis, 2003).

4.2. A priori partitioning vs mixture models

The analysis using the mitochondrial data partitioned by codon position and the nuclear data partitioned by genes performed better in comparison to the other *a priori* partitioning strategies. Studies based on real and simulated datasets have strengthened the idea that partitioning data according to expected differences in patterns of evolution increase accuracy of phylogenetic reconstructions and clade posterior probabilities (Brandley et al., 2005; Brown and Lemmon, 2007; Castoe and Parkinson, 2006; Kergoat et al., 2007). Our results also emphasize the importance of integrating all available biological information in analyses of data from multiple genes. We could not test the effects of partitioning stem and loops positions of ribosomal genes. However, it has been shown that partitioning data with coding genes by codon position results in comparatively greater increases in likelihood scores than partitioning based on ribosomal secondary structure (Brandley et al., 2005; Kergoat et al., 2007). For this reason, we believe that if we have tested the effects of partitioning the ribosomal genes according to stems and loops positions, general results of comparisons between *a priori* partitioning strategies and mixture models would not have been affected. In addition, Pagel and Meade (2004) applied mixture models to a 12S rDNA dataset and detected heterogeneity within stem and loop regions. They argued that partitioning these regions would miss within stem and loops variability.

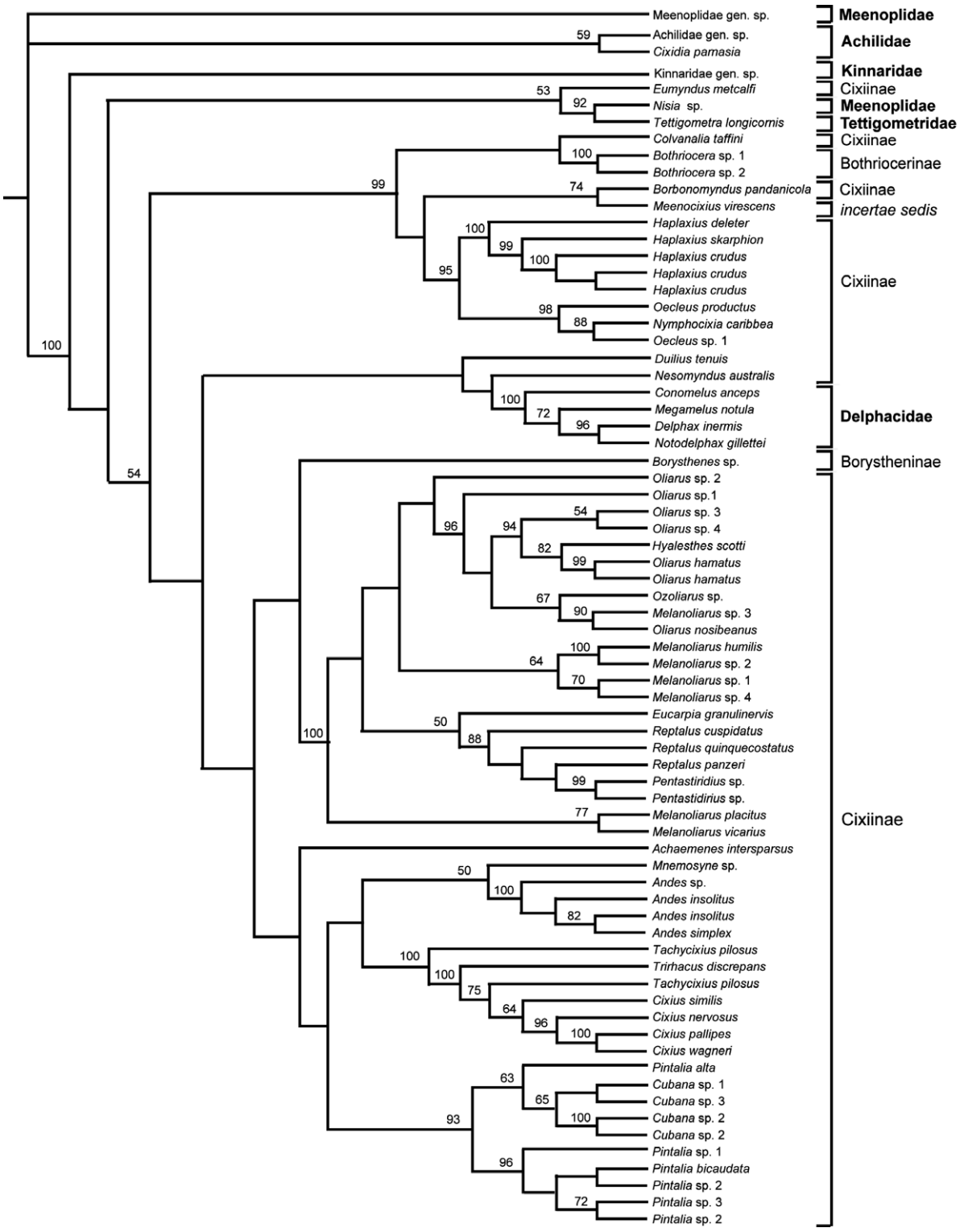


Fig. 2. Bayesian inference topology of Cixiidae derived *a priori* partitioning strategy P3 (Three partitions corresponding to 1st, 2nd and 3rd codon positions of mitochondrial genes, three partitions corresponding to the ribosomal genes: 18S rDNA, regions D4–D5 and D6–D7 of the 28S rDNA. Numbers above branches are posterior probabilities values (> 50%).

When comparing *a priori* partitioning strategies to the use of mixture models, the later improved the likelihood scores, as shown by B_F values (Table 6). This result indicates that pattern heterogeneity is important in our dataset and that *a priori* partitioning strategy P3 does not account for all this pattern heterogeneity. Pagel and Meade (2004) advocated the use of mixture models mainly when there is not a clear case for partitioning the dataset, but as suggested by the same authors, it may allow for unforeseen patterns of evolution to emerge. Even in cases in which *a priori* partitioning adheres to sound predictions that partitions follow different evolutionary models, heterogeneity within these different partitions may be lost when using *a priori* partitioning strategies. This may explain the better scores obtained by using a mixture models strategy with five matrices over the *a priori* partitioning

Table 4
Comparison of results under 1–6 phylogenetic mixture models using B_F

	Number of parameters	Mean log-likelihood	Mean standard deviation of rate parameters	$2\ln B_F$
1Q	6	−26,430.4	0.6	
2Q	13	−25,945.7	3.7	979.1
3Q	20	−25,710.7	0.6	458.7
4Q	27	−25,605.3	0.7	209.1
5Q	34	−25,550.3	0.7	89
6Q	41	−25,531.6	0.8	54.7

Table 5
Weights attributed by the mixture models analyses to each matrix for the analyses from two to six matrices (when only one matrix is used its weight is one)

	Q1	Q2	Q3	Q4	Q5	Q6
2Q	0.49	0.51				
3Q	0.25	0.32	0.43			
4Q	0.1	0.09	0.65	0.16		
5Q	0.64	0.08	0.08	0.05	0.15	
6Q	0.09	0.06	0.09	0.61	0.1	0.06

P3, which accounts for the important heterogeneity among codon positions but does not accommodate heterogeneity within each partition.

Even if we presume that partitioning ribosomal genes according to secondary structure would not have changed general results, part of the better performance of mixture models is probably due to heterogeneity within these genes. As they were not partitioned *a priori* according to stem and loop positions, pattern heterogeneity within ribosomal genes have only been accommodated by mixture models. To further explore the relative performance of mixture models and *a priori* partitioning as applied to ribosomal data, the approach used here could be employed to compare these two methods in terms of Bayes factors and likelihood scores.

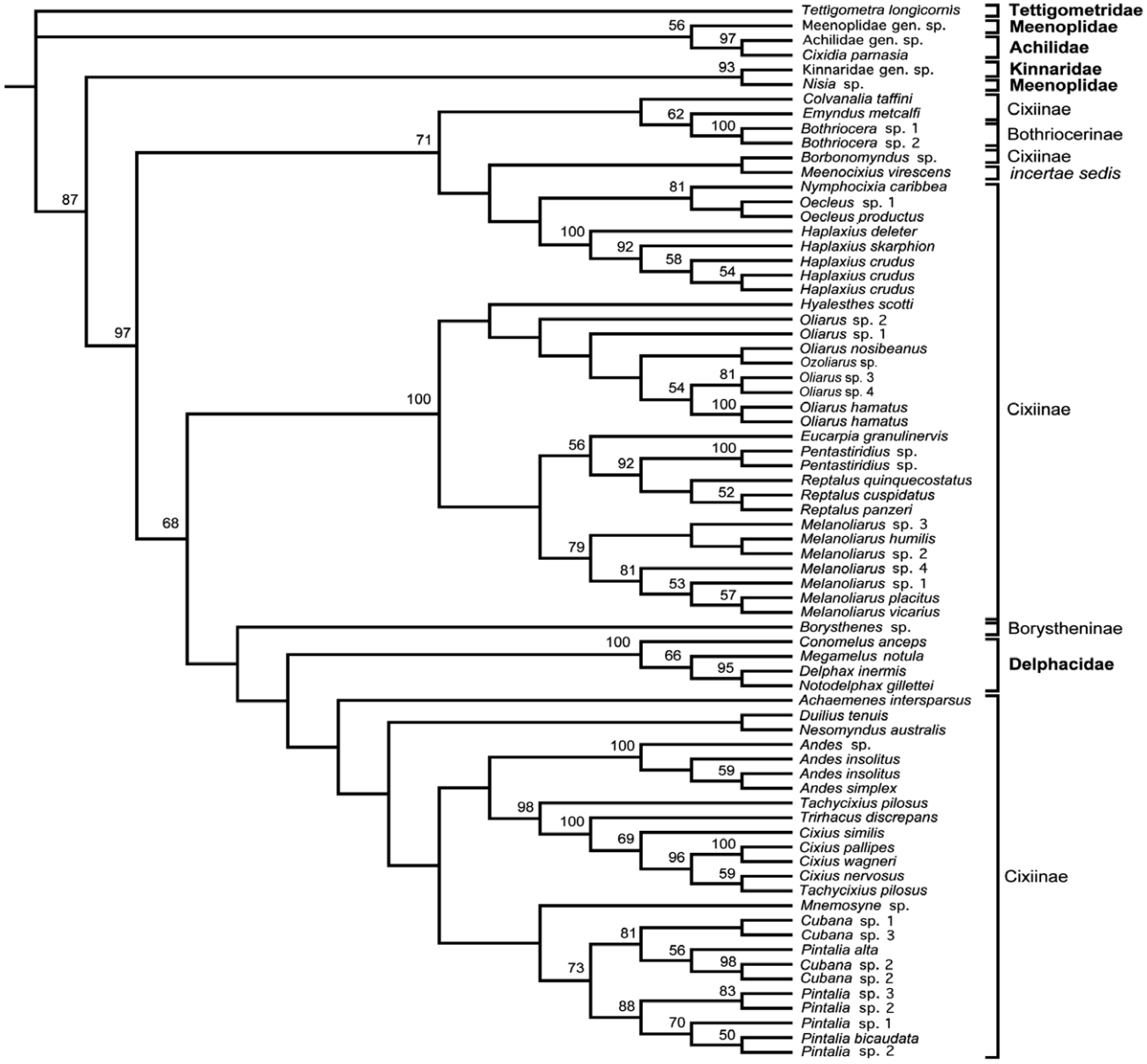


Fig. 3. Bayesian inference topology of Cixiidae derived with 6Q mixture models. Numbers above branches are posterior probabilities values (> 50%).

Table 6
Comparison of results of different *a priori* partitioning strategies and 1–6 phylogenetic mixture models using B_F

	1Q	2Q	3Q	4Q	5Q	6Q
P1	−256.26	722.84	1181.52	1390.66	1479.7	1534.4
P2	−1192.86	−213.76	244.92	454.06	543.1	597.8
P3	−1707.66	−728.56	−269.88	−60.74	28.3	83

Positive values favor mixture models and negative ones favor the *a priori* partitioning strategies.

It is interesting to notice the attributed weights of the 1–6$Q + \Gamma$ models. As mentioned before, the distribution of the weights changes from proportionate in 2–3 Q to disproportionate in 4–6 Q, with one of the matrices fitting the majority of the sites. As the increase in likelihood scores when going from 3Q to 4Q is substantial, it seems to us that this disproportion actually reflects a feature of our dataset, of which 70% corresponds to the ribosomal genes (30% of 18S rDNA and 40% of 28S rDNA). Considering the great amount of conserved or slowly evolving sites present in these genes, the most represented matrix may correspond to these conserved sites. It is difficult to assess the meaning of the matrices of lesser weight.

A good point of mixture models relates to the number of parameters. As shown by the better fit of the $5Q + \Gamma$ model over to the model of partitioning strategy P3 (six partitions; Table 6), mixture models may reduce the number of parameters to be estimated. This reduction in the number of parameters to be estimated is a good perspective for phylogenomic studies, in which the great number of genes under different evolutionary pressures and patterns may be hardly partitioned *a priori*.

4.3. Phylogeny of Cixiidae

The understanding of the relationships within Cixiidae has been mentioned as an extremely difficult problem for a long time (Asche, 1988; Ceotto and Bourgoin, 2008; Emeljanov, 2002). However, these comments have been based on studies of morphological data alone. Our study is the first attempt to recover phylogenetic relationships of Cixiidae based on molecular data. In spite of the representative number of characters present in the data (nearly 4 kb), which are expected to resolve basal nodes as well as recently diverged lineages, the recovered groups are weakly supported. In spite of this, some interesting patterns emerge.

Both MP and BI analyses recovered Cixiidae as paraphyletic in respect to Delphacidae. The paraphyletic condition of Cixiidae has been suggested several times (Asche, 1988; Muir, 1923; Urban and Cryan, 2007). In a molecular phylogeny of the infraorder Fulgoromorpha, Delphacidae appeared as arising from Cixiidae in a parsimony analysis, whereas the same data analyzed with Bayesian methods recovered these two families as sister-groups (Urban and Cryan, 2007). Furthermore, Cixiidae was recovered as monophyletic in a recent morphological study, with low support though (Ceotto and Bourgoin, 2008). In the present study, the failure of the topologies test to reject the alternative hypothesis of Cixiidae as monophyletic unfortunately prevents us from being more incisive regarding Cixiidae paraphyly. However, the fact that both MP and BI analyses recovered Delphacidae as arising from Cixiidae can be regarded as a good indication of the paraphyly of Cixiidae. This conclusion also seems reasonable because these families share some unique characters within Fulgoromorpha: long ovipositor (Asche, 1988; Muir, 1923), and the presence of a flagelliform aedeagus in all Cixiids and in basal Delphacidae (Asche, 1990). The monophyly of the later has been supported by a phylogenetic analysis of the family, the most striking of its synapomorphies being the presence of a spur in the tibial apex (Asche, 1985).

As a result of the placement of the Eucarpiini *E. granulinervis*, the present analyses failed to recognize the monophyly of the Pentastirini tribe, recovered as monophyletic in a morphology-based phylogeny (Ceotto and Bourgoin, 2008). However, topologies tests did not reject the alternative hypothesis of monophyly for Pentastirini. It seems that the clustering of *E. granulinervis* with Pentastirini species is a consequence of the fact that the only gene that we obtained for this Eucarpiini species was Cyt*b* and that it is the only representative species of Eucarpiini in the analyses. In fact, different topologies recovered using 1–6Q mixture models and P1–P3 *a priori* partitioning strategies involve changes in the placement of four terminals, and for three of them only one sequence was obtained. Pentastirini are morphologically defined by three synapomorphies: forewings with CuA bifurcating distad of ScP + R + MA, the presence of multi-pointed setae on segment nine of females, internal margin of gonapophysis nine separated (Ceotto and Bourgoin, 2008).

Contrary to the phylogeny based on morphological data (Ceotto and Bourgoin, 2008), the subfamilies Bothriocerinae and Borystheninae did not cluster together. In addition, topologies tests significantly rejected the alternative hypothesis of a clade formed by Borystheninae + Bothriocerinae. These subfamilies share some morphological characters: the reniform antennal pedicel (also with Stenophlepsini), the presence of a subantennal carina (also with Stenophlepsini), the claval apex located in the basal half of forewings, the overlapping forewings, and the rounded apex of mesonotum. Of these characters, the only that appeared as a sound synapomorphy for the group in the morphological phylogeny of Ceotto and Bourgoin (2008) is the rounded apex of mesonotum. Cixiidae morphological characters are highly homoplastic (Asche, 1988; Ceotto and Bourgoin, 2008), so that particular characters shared by a few taxa may get disproportionate importance and support some spurious clades. Convergence might thus explain the formation of a Borystheninae + Bothriocerinae clade based on morphological features. Further studies are needed to better understand the relative position of these subfamilies and to evaluate the validity of their shared morphological characters. Our molecular results not supporting the clade Borystheninae + Bothriocerinae are in accord with the actual distribution of these subfamilies: whereas the former is found in the Ethiopian and Oriental regions, the later is distributed over the New World. However, the presence of a fossil Bothriocerinae in the Eocene Baltic amber (Szwedo, 2002) prevents us from discarding the possibility of a Borystheninae + Bothriocerinae clade.

5. Conclusions

In this study, we used a combined approach to estimate the first phylogeny of Cixiidae based on a large molecular dataset. Our results point to a paraphyletic status of the family, supporting the long suspected idea that Delphacidae might be considered as a subfamily of Cixiidae in the future. Further studies, with a more comprehensive sampling, are needed to verify this suggestion. The present paper also reinforced the previous studies that state that data partitioning significantly improves likelihood scores. More importantly, our data support the view that the performance of mixture models may overcome that of *a priori* partitioning strategies in BI methods. It seems to us that the benefits of using mixture models have been neglected, as few studies using these methods can be found in the literature. As a matter of fact, the recently developed and promising field of phylogenomic studies would probably benefit from this approach, as it is not limited by the need to define partitions *a priori*.

Acknowledgments

We are very grateful to B. Campbell, B. Nusillard, J. Constant, J. Cryan, N. Penny, J. Danet, and L. Vladimir for donation of specimens. Sequences were performed through a common project (BI4I) granted by the French sequencing center Genoscope. P. Ceotto thanks G. Genson for instruction in molecular techniques, J.-F. Martin, S. Cros-Arteil, and T. Malausa for help with molecular amplification difficulties, and expresses her gratitude to D. Bourguet for training opportunity at CBGP (Centre de Biologie et Gestion de Populations). We thank A. Estoup and T. Guillemaud for allowing us to use cluster computers in CBGP and INRA-Sophia Antipolis, respectively, and Filipe Santos for management of CBGP cluster analyses. We are very grateful to Carlos Peña for sending the scripts to calculate PBS in TNT. We thank T. Malausa for his careful reading and valuable comments on a previous version of the manuscript. The Ph.D. scholarship provided by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) to Paula Ceotto is greatly acknowledged.

References

Alfaro, M.E., Zoler, S., Lutzoni, F., 2003. Bayes or bootstrap? A simulation study comparing the performance of Bayesian Markov chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. Mol. Biol. Evol. 20, 255–266.

Asche, M., 1985. Summary: zur phylogenie der Delphacidae. Marburger Ent. Pub. 2, 366–374.

Asche, M., 1988. Preliminary thoughts on the phylogeny of Fulgoromorpha (Homoptera, Auchenorrhyncha). In: Vidano, C., Arzone, A. (Eds.), Proceedings of the Sixth Auchenorrhyncha Meeting, Turin, Italy, 1987. Consiglio nazionale delle Ricerche, Italy, pp. 47–53.

Asche, M., 1990. Vizcayinae, a new subfamily of Delphacidae with a revision of *Vizcaya* Muir (Homoptera: Fulgoroidea)—a significant phylogenetic link. Bishop Mus. Occas. Paper 30, 154–187.

Baker, R.H., DeSalle, R., 1997. Multiple sources of character information and the phylogeny of the Hawaiian drosophilids. Syst. Biol. 46, 645–673.

Belshaw, R., Quicke, D.L.J., 2002. Robustness of ancestral state estimates: evolution of life history strategy in Ichneumonoid parasitoids. Syst. Biol. 51, 450–477.

Bourgoin, T., Steffen-Campbell, J.D., Campbell, B.C., 1997. Molecular phylogeny of Fulgoromorpha (Insecta, Hemiptera, Archaeorrhyncha). The enigmatic Tettigometridae: evolutionary affiliations and historical biogeography. Cladistics 13, 207–224.

Brandley, M.C., Schmitz, A., Reeder, T.W., 2005. Partitioned Bayesian analyses, partition choice, and the phylogenetic relationships of Scincid lizards. Syst. Biol. 54, 373–390.

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

Bremer, K., 1994. Branch support and tree stability. Cladistics 10, 295–304.

Brown, J.M., Lemmon, A.R., 2007. The importance of data partitioning and the utility of Bayes factors in Bayesian phylogenetics. Syst. Biol. 56, 643–655.

Castoe, T.A., Parkinson, C.L., 2006. Bayesian mixed models and the phylogeny of pitvipers (Viperidae: Serpentes). Mol. Phylogenet. Evol. 39, 91–110.

Claridge, M.F., 1985. Acoustic behavior of leafhoppers and planthoppers: species problems and speciation. In: Nault, L.R., Rodriguez, J.G. (Eds.), The Leafhoppers and Planthoppers. John Wiley and Sons, New York, pp. 61–102.

Ceotto, P.C., Bourgoin, T., 2008. Insights into the phylogenetic relationships within Cixiidae (Hemiptera: Fulgoromorpha): cladistic analysis of a morphological data set. Syst. Entomol. 33.

Cryan, J.R., Wiegmann, B.M., Deitz, L.L., Dietrich, C.H., 2000. Phylogeny of the treehoppers (Insecta: Hemiptera: Membracidae): evidence from two nuclear genes. Mol. Phylogenet. Evol. 17, 317–334.

Danet, J.L., Foissac, X., Zreik, L., Salar, P., Verdin, E., Nourrisseau, J.G., Garnier, M., 2003. '*Candidatus phlomobacter* fragariae' is the prevalent agent of marginal chlorosis of strawberry in French production fields and is transmitted by the planthopper *Cixius wagneri* (China). Phytopathology 93, 644–649.

Darlu, P., Lecointre, G., 2002. When does the incongruence length difference test fail? Mol. Biol. Evol. 19, 432–437.

Emeljanov, A.F., 1989. On the problem of the division of the family Cixiidae (Homoptera, Cicadina). Entomol. Rev. 68, 54–67.

Emeljanov, A.F., 2002. Contribution to classification and phylogeny of the family Cixiidae. Denisia 4, 103–112.

Erixon, P., Svennblad, B., Britton, T., Oxelman, B., 2003. Reliability of Bayesian posterior probabilities and bootstrap frequencies in phylogenetics. Syst. Biol. 52, 665–673.

Farris, J.S., Källersjö, M., Kluge, A.G., Bult, C., 1994. Testing significance of incongruence. Cladistics 10, 315–319.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.

Goldman, N., Anderson, J.P., Rodrigo, A.G., 2000. Likelihood-based tests of topologies in phylogenetics. Syst. Biol. 49, 652–670.

Goloboff, P., Farris, J., Nixon, K., 2003. TNT: Tree Analysis Using New Technology. Program and documentation, available from the authors and at: <http://www.zmuc.dk/public/phylogeny>.

Grimaldi, D.A., Angel, M.S., Nascimbene, P.C., 2002. Fossiliferous Cretaceous Amber from Myanmar (Burma): its rediscovery, biotic diversity and palaeontological significance. Am. Mus. Nov. 3361, 1–72.

Gu, X., Fu, Y.-X., Li, W.-H., 1995. Maximum likelihood estimation of the heterogeneity of substitution rate among nucleotide sites. Mol. Biol. Evol. 12, 546–557.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids Symp. Ser. 41, 95–98.

Harry, M., Solignac, M., Lachaise, D., 1998. Molecular evidence for parallel evolution of adaptive syndromes in fig-breeding *Lissocephala* (Drosophilidae) on the pantropical genus *Ficus* (Moraceae). J. Biogeogr. 23, 543–552.

Hillis, D.M., Bull, J.J., 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst. Biol. 42, 182–192.

Hipp, A.L., Hall, J.C., Systma, K.J., 2004. Congruence versus phylogenetic accuracy: revisiting the Incongruence Length Difference test. Syst. Biol. 53, 81–89.

Holder, M., Lewis, P.O., 2003. Phylogeny estimation: traditional and Bayesian approaches. Nat. Rev. Genet. 4, 275–284.

Holzinger, W.E., Emeljanov, A.F., Kammerlander, I., 2002. The family Cixiidae Spinola 1839 (Hemiptera: Fulgoromorpha)—a review. Denisia 4, 113–138.

Huelsenbeck, J., 1995. The performance of phylogenetic methods in simulation. Syst. Biol. 44, 17–48.

Jermiin, L.S., Crozier, R.H., 1994. The cytochrome *b* region in the mitochondrial DNA of the ant *Tetraponera rufoniger*: sequence divergence in Hymenoptera may be associated with nucleotide content. J. Mol. Evol. 38, 282–294.

Jović, J., Cvrković, T., Mitrović, M., Krnjajić, S., Redinbaugh, M.G., Pratt, R.C., Gingery, R.E., Hogenhout, S.A., Toševski, I., 2007. Roles of stolbur phytoplasma and *Reptalus panzeri* (Cixiinae, Auchenorrhyncha) in the epidemiology of Maize redness in Serbia. Eur. J. Plant Pathol. 118, 85–89.

Kergoat, G.J., Silvain, J.-F., Delobel, A., Tuda, M., Anton, K.-W., 2007. Defining the limits of taxonomic conservatism in host–plant use for phytophagous insects: molecular systematics and evolution of host–plant associations in the seed-beetle genus *Bruchus* Linnaeus (Coleoptera: Chrysomelidae, Bruchinae). Mol. Phylogenet. Evol. 43, 251–269.

Kishino, H., Miyata, T., Hasegawa, M., 1990. Maximum likelihood inference of protein phylogeny and the origin of chloroplasts. J. Mol. Evol. 30, 151–160.

Lee, M.S.Y., 2001. Uninformative characters and apparent conflict between molecules and morphology. Mol. Biol. Evol. 18, 676–680.

Lunt, D.H., Zhand, D.X., Szymura, J.M., Hewitt, G.M., 1996. The insect cytochrome oxidase I gene evolutionary patterns and conserved primers for phylogenetic studies. Insect Mol. Biol. 5, 153–165.

Lyons-Weiler, J., Hoelzer, G.A., Tausch, R.J., 1998. Optimal outgroup analysis. Biol. J. Linn. Soc. 64, 493–511.

Maddison, W.P., Maddison, D.R., 2006. Mesquite: A Modular System for Evolutionary A, version 1.12. Available at: <http://mesquiteproject.org>.

Muir, F., 1923. On the classification of the Fulgoroidea (Homoptera). Proc. Hawaiian Entomol. Soc. 5, 205–247.

Niehuis, O., Naumann, C.M., Misof, B., 2006. Identification of evolutionary conserved structural elements in the mt SSU rRNA of Zygaenoidea (Lepdoptera): a comparative sequence analysis. Organ. Divers Evol. 6, 17–32.

Nylander, J.A.A., Ronquist, F., Huelsenbeck, J.P., Nieves-Aldrey, J.L., 2004. Bayesian phylogenetic analysis of combined data. Syst. Biol. 53, 47–67.

O'Brien, L.B., Wilson, S.W., 1985. Planthopper systematics and external morphology. In: Nault, L.R., Rodriguez, J.G. (Eds.), The Leafhoppers and Planthoppers. John Wiley and Sons, New York, pp. 61–102.

Ouvrard, D., Campbell, B.C., Bourgoin, T., Chan, K.L., 2000. 18S rRNA secondary structure and phylogenetic position of Peloridiidae (Insecta: Hemiptera). Mol. Phylogenet. Evol. 16, 403–417.

Pagel, M., Meade, A., 2004. A phylogenetic mixture model for detecting pattern-heterogeneity in gene sequence or character-state data. Syst. Biol. 53, 571–581.

Pagel, M., Meade, A., 2005. Mixture models in phylogenetic inference. In: Gascuel, O. (Ed.), Mathematics of Evolution and Phylogeny. Clarenton Press, Oxford, pp. 121–142.

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

Ramírez, M.J., 2006. Further problems with the Incongruence Length Difference test: "hypercongruence" effect and multiple comparisons. Cladistics 22, 289–295.

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Shcherbakov, D., 2002. The 270 million year history of *Auchenorrhyncha* (Homoptera). Denisia 04, 29–36.

Sémétey, O., Gatineau, F., Bressan, A., Boudon-Padieu, E., 2007. Characterization of a gamma-3 Proteobacteria responsible for the syndrome "Basses Richesses" of sugar beet transmitted by *Pentastiridius* sp (Hemiptera, Cixiidae). Phytopathology 97, 72–78.

Sforza, R., Bourgoin, T., 1998. Female genitalia and copulation of the planthopper *Hyalesthes obsoletus* Signoret (Hemiptera, Fulgoromorpha, Cixiidae). Ann. Soc. Entomol. Fr. (N.S.) 34, 63–70.

Sforza, R., Clair, D., Daire, X., Larrue, J., Boudon-Padieu, E., 1998. The role of *Hyalesthes obsoletus* (Hemiptera: Cixiidae) in the occurrence of Bois Noir of grapevines in France. J. Phytopathol. 146, 549–556.

Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16, 1114–1116.

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., Flook, P., 1994. Evolution, weighting and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. 87, 651–702.

Simmons, M.P., Zhang, L.-B., Webb, C.T., Reeves, A., Miller, J.A., 2006. The relative performance of Bayesian and parsimony approaches when sampling characters evolving under homogeneous sets of parameters. Cladistics 22, 171–185.

Smith, A.B., 1994. Rooting molecular trees: problems and strategies. Biol. J. Linn. Soc. 51, 279–292.

Swofford, D.L., 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods), Version 4 beta 10. Sinauer Associates, Sunderland, MA.

Swofford, D.L., Waddell, P.J., Huelsenbeck, J.P., Foster, P.G., Lewis, P.O., Rogers, J.S., 2001. Bias in phylogenetic estimation and its relevance to the choice between parsimony and likelihood methods. Syst. Biol. 50, 525–539.

Szwedo, J., 2002. The first fossil Bothriocerinae from Eocene Baltic amber with notes on recent taxa (Hemiptera, Fulgoromorpha, Cixiidae). Deut. Entomol. Zeit. 49, 197–207.

Szwedo, J., 2004. *Autrimpus sambiorum* gen. and sp. nov. from Eocene baltic amber and notes on Mnemosynini stat. nov. (Hemiptera: Fulgoroidea: Cixiidae). Ann. Zool. 4, 567–578.

Szwedo, J., Stroinski, A., 2002. First fossil Pentastirini from Eocene baltic ambar (Hemiptera: Fulgoromorpha: Cixiidae). Ann. Zool. 52, 173–179.

Thompson, J.D., Gibson, T.J., Plewuniak, F., Jeanmougin, F., Higgins, D.G., 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882.

Urban, J.M., Cryan, J.R., 2007. Evolution of the planthoppers (Insecta: Hemiptera: Fulgoroidea). Mol. Phylogenet. Evol. 42, 556–572.

Wiens, J.J., 2003. Missing data, incomplete taxa, and phylogenetic accuracy. Syst. Biol. 52, 528–538.

Wiens, J.J., 2005. Can incomplete taxa rescue phylogenetic analyses from long-branch attraction? Syst. Biol. 54, 731–742.

Wiens, J.J., 2006. Missing data and the design of phylogenetic analyses. J. Biomed. Inform. 39, 34–42.

Wilson, S.W., Mitter, C., Denno, R.F., Wilson, M.R., 1994. Evolutionary patterns of host plant use by delphacid Planthoppers and their relatives. In: Denno, R.F., Perfect, T.J. (Eds.), Planthoppers: Their Ecology and Management. Chapman and Hall, New York, pp. 7–45.

Yang, Z., 1994. Maximum likelihood models for combined analyses of multiple sequence data. J. Mol. Evol. 42, 587–596.