Zoologica Scripta

KUNGL. VETENSKAPS-AKADEMIEN
THE ROYAL SWEDISH ACADEMY OF SCIENCES

Towards a molecular phylogeny of the fungus gnat genus *Boletina* (Diptera: Mycetophilidae)

SVANTE MARTINSSON, JOSTEIN KJÆRANDSEN & PER SUNDBERG

Submitted: 17 November 2010
Accepted: 6 February 2011
doi:10.1111/j.1463-6409.2011.00474.x

Martinsson, S., Kjærandsen, J. & Sundberg, P. (2011). Towards a molecular phylogeny of the fungus gnat genus *Boletina* (Diptera: Mycetophilidae). — *Zoologica Scripta*, *40*, 272–281.

Boletina is a species rich genus of fungus gnats (Diptera: Mycetophilidae) with a mainly Holarctic distribution. The systematics within the genus has gained little attention and this is a first attempt to shed some light over the systematics of *Boletina* and to test the segregation of the genera *Saigusaia* and *Aglaomyia* from *Boletina*. The nuclear marker 28S and mitochondrial 16S, COI and CytB were amplified and sequenced for 23 taxa that were analysed separately and together with a broad sample of outgroup taxa obtained from GenBank, where also 18S sequences were added. Phylogenies were estimated using maximum likelihood, Bayesian inference and parsimony. We strengthen the hypothesized sister-group relationship between *Docosia* and *Boletina*, but the genus *Boletina* as currently delimited appears to be paraphyletic and nested in a clade together with *Aglaomyia*, *Coelosia* and *Gnoriste*. The genus *Saigusaia*, on the other hand, seems to be well separated from *Boletina*. The *Boletina erythropyga* species group is consistently found as a distinct basal clade within *Boletina* s.l. The results obtained are otherwise ambiguous both for the taxa in focus and in some analyses globally with a statistically supported total breakdown of the traditional higher classification into tribes, subfamilies and even families. Interestingly, this breakdown almost disappeared when additional 18S sequences were added.

Corresponding author: *Svante Martinsson, Department of Zoology, University of Gothenburg, Box 463, SE-405 30 Göteborg, Sweden. E-mail: svante_martinsson@yahoo.se*
Jostein Kjærandsen, Museum of Zoology, Lund University, Helgonavägen 3, SE-223 62 Lund, Sweden. E-mail: jostein.kjaerandsen@zool.lu.se
Per Sundberg, Department of Zoology, University of Gothenburg, Box 463, SE-405 30 Göteborg, Sweden. E-mail: per.sundberg@zool.gu.se

Introduction

The fungus gnat genus *Boletina* Staeger, family Mycetophilidae, consists of about 160 known species (Evenhuis *et al.* 2008) mainly confined to the Holarctic region, but representatives of the genus are found in the Oriental region as well (Bechev 2000). Larvae of *Boletina* are considered mainly saproxylobiontic (Jakovlev 1995) and larval habitats include decaying wood and fruiting bodies of fungi (Økland 1999; Jakovlev *et al.* 2008). There are also a few records of exposed mineral soil, mosses and liverworts as larval habitats (Edwards 1925; Økland 1999), but it is not clear whether they were really larval macrohabitats or just pupation places. Adults of *Boletina* are found in a wider range of habitats than most species in other genera of the family (Hutson *et al.* 1980), including alpine and arctic environments where some species can be very abundant (JK unpublished data). Other species of *Boletina* are abundant both in boreal and nemoral forests where adults can make up a considerable proportion of Malaise trap and window trap samples (Russell-Smith 1979; Økland 1994; Jakovlev 1995). As such the genus may be a suitable model group to study the effects of climatic change in the northern hemisphere and arctic, but both its systematics and biology remain insufficiently outlined. A number of new *Boletina* species have recently been described from the Nordic region (Zaitzev & Polevoi 1995; Polevoi & Hedmark 2004; Jakovlev & Penttinen 2007) and more species awaits description (Kjærandsen *et al.* 2007), but so far only a few subgroups within the genus have been revised (Zaitzev & Polevoi 2001; Zaitzev *et al.* 2005).

The genus *Boletina* was traditionally placed in the subfamily Gnoristinae, but the classification and delimitations of higher taxa in the family has been debated for decades (e.g. Tuomikoski 1966; Väisänen 1986; Søli *et al.* 2000; see Gammelmo 2004 for a overview of the traditional classifications). All recent phylogenetic studies indicate that the subfamilial and tribal classification of fungus gnats, as traditionally accepted, includes non-monophyletic groups

© 2011 The Authors • Zoologica Scripta © 2011 The Norwegian Academy of Science and Letters,

and the higher phylogeny of fungus gnats remains both largely unresolved and contentious (Søli 1997a; Baxter 1999; Hippa & Vilkamaa 2005, 2006; Amorim & Rindal 2007; Rindal *et al.* 2009a). Recent studies agree, however, to retain *Boletina* in subfamily Gnoristinae where it is placed in a clade together with *Gnoriste* Meigen (Søli 1997a; Baxter 1999; Rindal *et al.* 2009a).

A close relationship between *Docosia* Winnertz (traditionally placed in subfamily Leiinae) and *Boletina* was first suggested by Baxter (1999) and further supported by Rindal *et al.* (2009a). Baxter (1999) studied the phylogeny of the Sciaroidea based on 12S rRNA sequences, and included three species of *Boletina* that appeared to be non-monophyletic nested together with *Gnoriste*. He found that the *Boletina–Gnoriste* group formed a clade together with *Docosia*, but well separated from *Coelosia* Winnertz. Rindal *et al.* (2009a), who analysed the 16S, 18S and 28S markers, found that *Docosia* together with *Palaeodocosia* Meunier may form the sister-group of a clade consisting of *Boletina*, *Gnoriste* and *Coelosia*.

Vockeroth (1980) segregated the genera *Saigusaia* Vockeroth and *Aglaomyia* Vockeroth from *Boletina* based on morphological characters such as wing venation, setation of thoracic and abdominal sclerites and characters in the male and female terminalia. In the morphological phylogeny by Søli (1997a) *Saigusaia* appeared in a clade together with *Synapha* Meigen and *Palaeodocosia,* well separated from *Boletina*. The genus *Aglaomyia* has not been included and tested in any phylogenetic studies as yet.

The aim of this study was to estimate phylogenies with focus on *Boletina* based on multiple molecular markers, and to test the monophyly of the genus and also whether *Saigusaia* and *Aglaomyia* deserve their status as separate genera. In addition, we aimed at testing the placement of the *Boletina erythropyga* group that in our view seems morphological as well separated from the rest of the genus as is *Aglaomyia* and *Saigusaia*.

Material and methods

Material

The study is based on specimens from the collections at the Museum of Zoology, Lund University (MZLU) (mainly from the Nordic region), and some *Aglaomyia* specimens that were obtained on the courtesy of T. Saigusa (Fukuoka, Japan) and from Hokkaido University Museum, Japan. Fresh specimens were obtained from recent Malaise trap and sweep net collections. In some cases, where fresh material was unavailable, older museum material from MZLU both dry pined and stored in alcohol, were tried. All specimens examined were recorded with unique identification codes prefixed by "SPM–" in a BIOTA 2.04 DATABASE (Colwell 2007). A large amount of material was studied and specimens within *Boletina* were chosen to obtain a broad sample for sequencing. Samples of 16 species of *Boletina*, two species of *Aglaomyia*, and two species of *Saigusaia* were included as the ingroup. Based on previous studies (Søli 1997a; Baxter 1999; Rindal *et al.* 2009a) *Gnoriste longirostris*, *Coelosia truncata* and one undescribed species of *Docosia* were added as potential outgroups. Table 1 lists specimens sequenced in this study, together with GenBank accession numbers. Additional collection data and depository for the voucher specimens can be found in Appendix S1.

DNA extraction and amplification

Numerous DNA sequence markers have been used to infer phylogenetic relationships in insects. The genes COI, 16S, 18S and EF-1α have been proposed as a standard for molecular insect phylogenies (Catrino *et al.* 2000). The genes 28S, EF-1α and opsin may be useful for resolving within family-level relationships whereas mitochondrial loci and ITS seem to be a better choice for lower level phylogenies (Rokas *et al.* 2002). Based on these recommendations and previous molecular works on fungus gnats (Rindal *et al.* 2007, 2009a,b) we settled on the nuclear marker 28S and the mitochondrial markers 16S, COI and added CytB to be tested for the first time.

The abdomen except terminalia were used for DNA extraction, the rest of the specimens were stored in alcohol as hologenophore vouchers [following the terminology proposed by Pleijel *et al.* (2008)] deposited in MZLU. DNA was extracted using the Dneasy Blood & Tissue Kit (Qiagen, Hilden, Germany), following the included protocol 'Purification of Total DNA from Animal Tissues (Spine-Column Protocol)'. PCR amplification of segments from the nuclear 28S and the mitochondrial 16S, CytB and COI genes follows standard procedures, using the primers listed in (Table S1). The amplification programme for the 28S gene was 95 °C for 5 min, followed by 35 cycles of 95 °C for 40 s, 52 °C for 40 s and 72 °C for 1 min and a final extension step at 72 °C for 8 min. The amplification programme for the 16S gene was 95 °C for 5 min, followed by 35 cycles of 95 °C for 30 s, 45 °C for 30 s and 72 °C for 1 min and a final extension step at 72 °C for 8 min. The amplification programme for the COI gene was 95 °C for 5 min, followed by 35 cycles of 95 °C for 40 s, 45 °C for 45 s and 72 °C for 1 min and a final extension step at 72 °C for 8 min. The amplification programme for the CytB gene was 94 °C for 3 min, followed by 35 cycles of 94 °C for 45 s, 40 °C for 45 s and 72 °C for 1 min and a final extension step at 72 °C for 10 min. PCR amplifications were performed using PuRe-

Table 1 List of specimens, used for DNA extraction, with specimen identification number and GenBank accession numbers. More information about the specimens is listed in the Supporting information. The specimens are males unless otherwise noted

Species	Collecting date	Id no.	GenBank accession numbers			
			16S	CytB	28S	COI
Aglaomyia ingrica	14–30 Jun 2001	SPM-057685	–	HQ230430	–	–
Aglaomyia sp.	16 Jul 1967	SPM-057687	–	–	–	–
Aglaomyia sp. ♀	16 Jul 1967	SPM-057688	–	–	–	–
Boletina antoma	13 Jul 1964	SPM-034388	–	–	–	–
Boletina atridentata	5–9 Aug 1997	SPM-057487	HQ230386	HQ230428	HQ230410	HQ230449
Boletina cordata	6 Aug to 25 Sep 1996	SPM-010002	HQ230371	HQ230418	HQ230393	HQ230436
Boletina cordata	14 Jul to 18 Aug 2004	SPM-009053	HQ230369	HQ230416	HQ230391	HQ230434
Boletina erythropyga	14 Jul to 18 Aug 2004	SPM-057363	HQ230383	HQ230426	HQ230407	HQ230449
Boletina gripha	20 Jun to 10 Jul 2003	SPM-015009	–	HQ230422	HQ230400	–
Boletina gripha	18 Aug to 20 Sep 2004	SPM-057362	HQ230382	–	HQ230406	HQ230448
Boletina gripha	24 Sep to 10 Nov 2006	SPM-015271	HQ230378	–	HQ230401	HQ230444
Boletina gripha	8 Nov 2003	SPM-057361	HQ230381	–	HQ230405	HQ230447
Boletina gripha	2 Mar 2008	SPM-033462	HQ230380	HQ230424	HQ230403	HQ230446
Boletina griphoides	19 May 2005	SPM-009627	HQ230370	HQ230417	HQ230392	HQ230435
Boletina hedstroemi	6 Aug to 25 Sep 1996	SPM-033556	–	–	–	–
Boletina hedstroemi	18 Aug to 9 Sep 2005	SPM-014978	HQ230377	–	HQ230399	HQ230443
Boletina nigricans	26 Aug to 6 Oct 2004	SPM-008071	HQ230368	HQ230414	–	–
Boletina nigricans	5–9 Aug 1997	SPM-057488	HQ230387	HQ230429	HQ230411	HQ230452
Boletina nitida	12 Jun to 2 Jul 1999	SPM-015885	–	–	–	–
Boletina nitida	2–28 Jul 1999	SPM-009911	–	–	–	–
Boletina sahlbergi	14 Jul to 18 Aug 2004	SPM-057369	HQ230384	HQ230427	HQ230408	HQ230449
Boletina sahlbergi	21 Jun 2004	SPM-057401	HQ230385	–	HQ230409	–
Boletina sciarina	31 May to 13 Jul 2004	SPM-007450	HQ230367	HQ230413	HQ230390	HQ230433
Boletina sciarina	2–28 Jul 1999	SPM-009909	–	–	–	–
Boletina subtrivittata	2 Oct 2006	SPM-032921	HQ230379	HQ230423	HQ230402	HQ230445
Boletina trispinosa	14 Oct to 11 Nov 2003	SPM-012711	HQ230373	–	HQ230395	HQ230438
Boletina trivittata	1–31 May 2004	SPM-006015	–	HQ230412	HQ230388	HQ230431
Boletina trivittata	26 Aug to 6 Oct 2004	SPM-008072	–	HQ230415	–	–
Boletina villosa	25 Jun to 4 Nov 2000	SPM-034311	–	HQ230425	HQ230404	–
Boletina villosa	18 Aug to 9 Sep 2005	SPM-014966	HQ230374	HQ230419	HQ230396	HQ230440
Boletina spA.	18 Aug to 9 Sep 2005	SPM-014967	HQ230375	HQ230420	HQ230397	HQ230441
Coelosia truncata	6–23 Oct 2005	SPM-013025	–	–	–	HQ230439
Docosia spA.	17 Aug to 25 Sep 2004	SPM-010122	HQ230372	–	HQ230394	HQ230437
Gnoriste longirostris	12 Jul to 17 Aug 2004	SPM-014972	HQ230376	HQ230421	HQ230398	HQ230442
Saigusaia flaviventris	31 May to 13 Jul 2004	SPM-006076	–	–	HQ230389	HQ230432
Saigusaia cincta	22–30 Jul 1982	SPM-034567	–	–	–	–

Taq Ready-To-Go (GE Healthcare, Chalfont St. Giles, UK) and were carried out as 25 μL reactions. All PCR products were tested for the presence of amplified products on 1% agarose gels. The obtained PCR products were purified using ExoTAP (Exonuclease I and FastAP Thermosensitive Alkaline Phosphatase) (Werle *et al.* 1994) following the protocol provided by the producer (Fermentas, Burlington, Canada) and sequenced by Macrogen (Geumcheon-Gu, Seoul, Korea).

Alignment

Sequence segments were assembled into consensus sequences using GENEIOUS PRO v4.8 (Drummond *et al.* 2009). The sequences were aligned using the Geneious Alignment with default settings in GENEIOUS PRO v4.8. Additional 16S, 28S and 18S sequences from Rindal *et al.* (2007, 2009a,b) representing an additional species of *Boletina* and a broad sample of taxa within Mycetophilidae were obtained from GenBank (GenBank accession numbers can be found in Table S2). The extended 16S, 18S and 28S matrices were aligned using MAFFT (online version, available from/mafft.cbrc.jp/alignment/server/) using L-INS-i strategy (Katoh *et al.* 2005), some bases considered arbitrary aligned were removed from analyses. Alignments were combined using PHYUTILITY v2.2.1 (Smith & Dunn 2008). Six datasets were created, one for each gene, one concatenated dataset consisting of 16S and 28S sequences, including additional sequences from GenBank and a dataset were also 18S sequences from GenBank were added.

Phylogenetic reconstruction

The single gene datasets were analysed using maximum likelihood (ML), Bayesian inference and parsimony. The extended dataset including 16S and 28S sequences was analysed using ML and Bayesian inference whereas the dataset also including 18S sequences was only analysed using ML. MRBAYES 3.1.2 (Huelsenbeck & Ronquist 2001; Ronquist & Huelsenbeck 2003) was used for the Bayesian analyses. The extended 16S and 28S dataset were partitioned according to genes and coding regions (COI and CytB) according to codon positions. Models were selected with the Akaike Information Criterion (AIC) (Akaike 1974) as implemented in MRMODELTEST v.2.2 (Nylander 2004). The best fitting models for the first, second and third position of COI were HKY + Γ, F81 + I and GTR + I+ Γ respectively and for the first, second and third positions of CytB were HKY + Γ, HKY + I and GTR + I respectively, for 16S were GTR + I + Γ the best fitting model and for 28S were SYM + Γ the best fitting model. For the extended dataset were GTR + I + Γ the best fitting model for both 16S and 28S. The COI dataset was analysed both with third codon position included and excluded. Bayesian analyses were performed with two independent runs with 4 MC^3 in each run. The analyses run under 10 million generations with sampling every 1000th generation, the extended analysis of the combined 16S and 28S dataset ran for 20 million generations. In analyses with more than one partition the priors for branch length were adjusted to 1/10th of the default value in order to reduce the problem with unrealistic long branch-lengths that can occur in Bayesian analyses (see Marshall 2010). To diagnose convergence and examined performance of the MCMC:s AWTY (Nylander *et al.* 2008) was used in order to decide when the chains had reached the stationary phase, and were sampling from the posterior probability. The initial 'Burnin-phase' was discarded. ML analysis were performed with PHYML 3.0 (Guindon & Gascuel 2003; Guindon *et al.* 2010) as implemented at ATGC Montpellier bioinformatics platform (http://www.atgc-montpellier.fr/) for the COI, 16S, CtyB and extended combined datasets were GTR model with I and Γ parameters estimated from the data used, for 28S was HKY85 with I and Γ parameters estimated from the data used. Branch support in ML analyses was calculated with approximative Likelihood Ratio Test (aLRT) (Anisimova & Gascuel 2006) in PHYML. Parsimony Bootstrap analysis where conducted using PAUP*4.0b10 (Swofford 2002) on single gene datasets, non-parametric Bootstrap (Felsenstein 1985) were performed with 10 000 pseudoreplicates with 10 replicates within each. Gaps was treated as missing data and all characters were unordered and with equal weight. Trees obtained from analyses were drawn with FIGTREE v1.3.1 (Rambout 2009).

Results

DNA extraction and amplification

Among a total of 36 DNA extractions 16S were successfully amplified in 22 specimens, 28S were successfully amplified in 23 specimens, COI were successfully amplified in 22 specimens and CytB were successfully amplified in 19 specimens (Tables 1 and 2). This rather low success rate was due to problems with extracting DNA from older museum material. Alignments were straightforward for the new sequences obtained in this study, whereas the alignment including additional sequences obtained from GenBank were more complicated due to the occurrence of indels as can be expected when more divergent sequences is included. Some bases considered arbitrary aligned were thus removed from further analyses, the dataset including additional 16S and 28S from GenBank was 659 bp long (16S 361 bp and 28S 298 bp long) and included 83 taxa. The dataset also including 18S sequences was 1476 bp long (16S 361 bp, 28S 298 bp and 18S 817 bp long) and included 85 taxa. Datasets are available at TreeBASE (treebase.org, accession URL: http://purl.org/phylo/treebase/phylows/study/TB2:S11166).

Phylogenetic analyses

The genus *Boletina* is not found monophyletic with respect to *Coelosia* and *Gnoriste* in any analyses. In the clade consisting of these three genera the *B. erythropyga* group is found monophyletic and forms the sister-group to all other *Boletina*, *Coelosia* and *Gnoriste*, except in the COI analyses (Fig. 1). The species *B. trivittata* and *B. subtrivittata* are found as sister-groups to all other species in this clade, in the extended and 28S analyses (Fig. 2) *B. trivittata* and *B. subtrivittata* forms a monophyletic clade. In the extended analyses (Fig. 3) this clade is found in a trichotomy with *Coelosia* and other *Boletina* including *Gnoriste*, and these clades have high support in the Bayesian analysis while somewhat lower support in the ML analysis. The genus *Saigusaia* is well separated from *Boletina* in the 28S and in the extended analyses. In the extended analyses *Saigusaia* appears as a sister-group to *Synapha* according to Bayesian inference and the ML analysis when 18S

Table 2 Details of alignment of the nuclear 28S and mitochondrial 16S, COI and CytB genes

Gene	No. of specimens	Length of alignment	AT content	Variable positions	Parsimony informative sites
28S	23	295	54.5%	32	12
16S	21	420	78.1%	86	59
COI	22	611	65.7%	297	181
CytB	19	395	71.2%	167	128

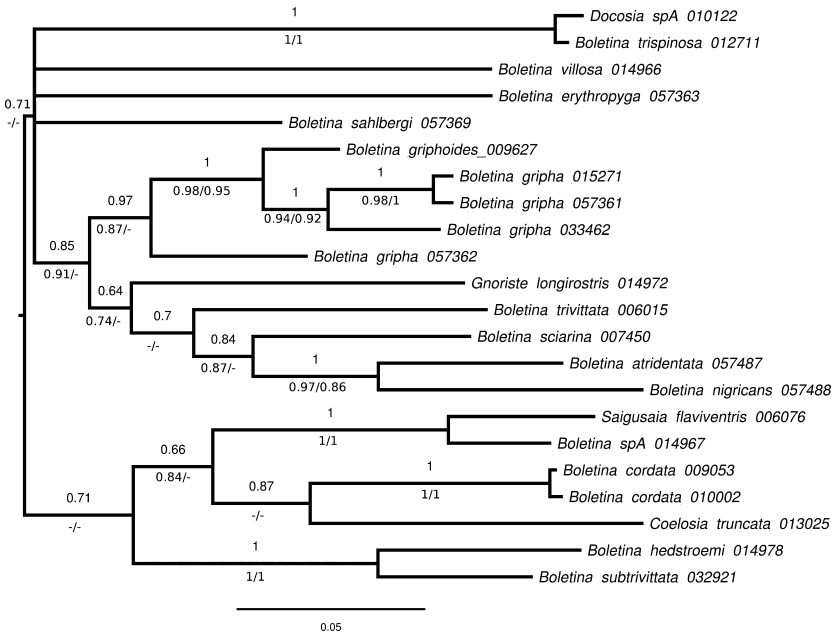

Fig. 1 Maj-Rules tree for Boletina obtained with Bayesian inference, based on COI sequences. Support values are given in the form: Above branches, Bayesian posterior probabilities, below branches aLRT from ML analysis/ Bootstrap proportions from MP analysis.

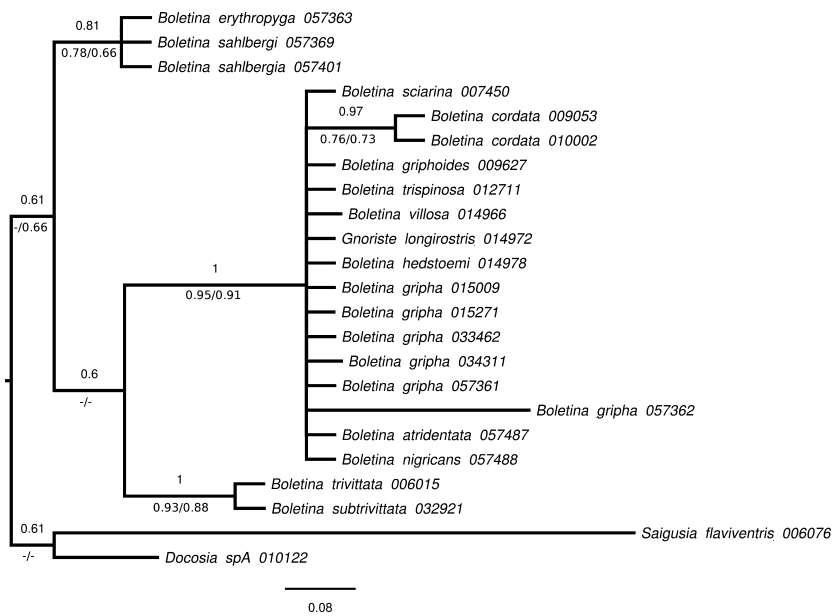

Fig. 2 Maj-Rules tree for Boletina obtained with Bayesian inference, based on 28S sequences. Support values are given in the form: Above branches, Bayesian posterior probabilities, below branches aLRT from ML analysis/ Bootstrap proportions from MP analysis.

sequences were added. Without 18S the ML analysis moved *Saigusaia* more basally to a rather questionable position in a clade mostly consisting of members of the subfamily Mycetophilinae. The genus *Aglaomyia* was only successfully sequenced for CytB and is there found within *Boletina* (Fig. 3). In all analyses *B. gripha* appears paraphyletic in respect to *B. griphoides*, this indicating that the name *B. gripha* may hide more than one species.

The extended, concatenated 16S and 28S analyses (Fig 4; Figs S1–S2) are well resolved but with a statistically supported total breakdown of the traditional higher classification into tribes, subfamilies and even families that give ambiguous classifications both for the taxa in focus and globally. In the Bayesian analysis *Tetragoneura sylvatica* (Curtis) is found as the sister-group to Diadocidiidae and this clade is found in a trichotomy together with *Azana* Walker and all other Mycetophilidae. Whereas in the ML analysis *Azana* is found as the sister-group to a clade consisting of Diadocidiidae, Bolitophilidae and all other Mycetophilidae, a clade consisting of *Leptomorphus* Curtis,

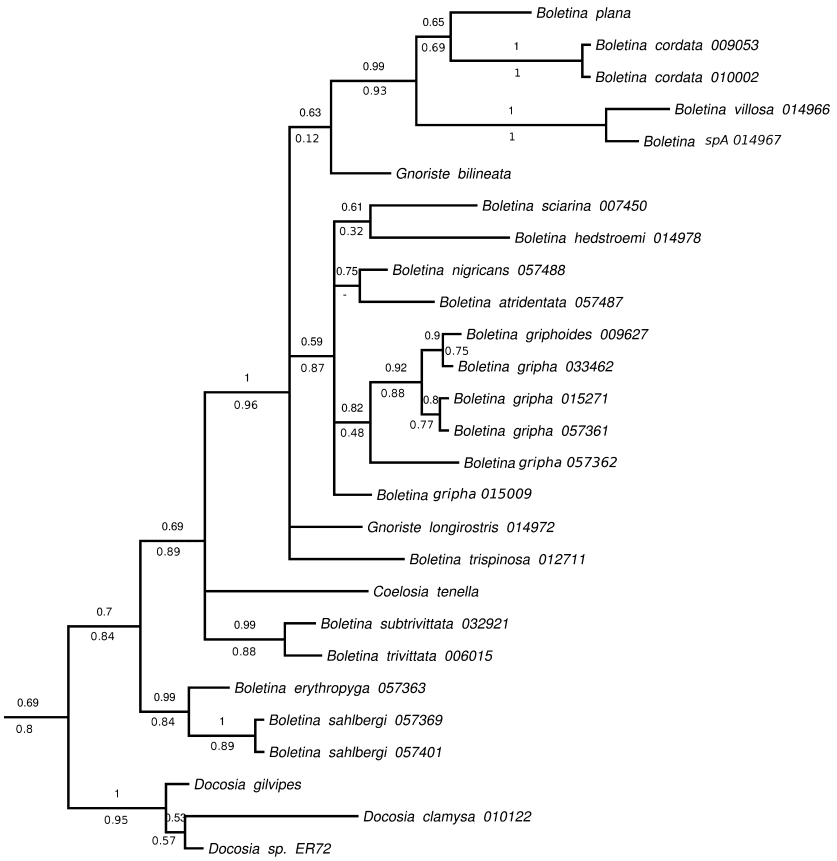

Fig. 3 Subtree of the clade consisting of Boletina, Coelosia, Docosia and Gnoriste from analysis of an extended dataset of 16S and 16 sequences (Supplementary fig. S2). Values above branches are Bayesian posterior probabilities and values below is ML aLRT supports.

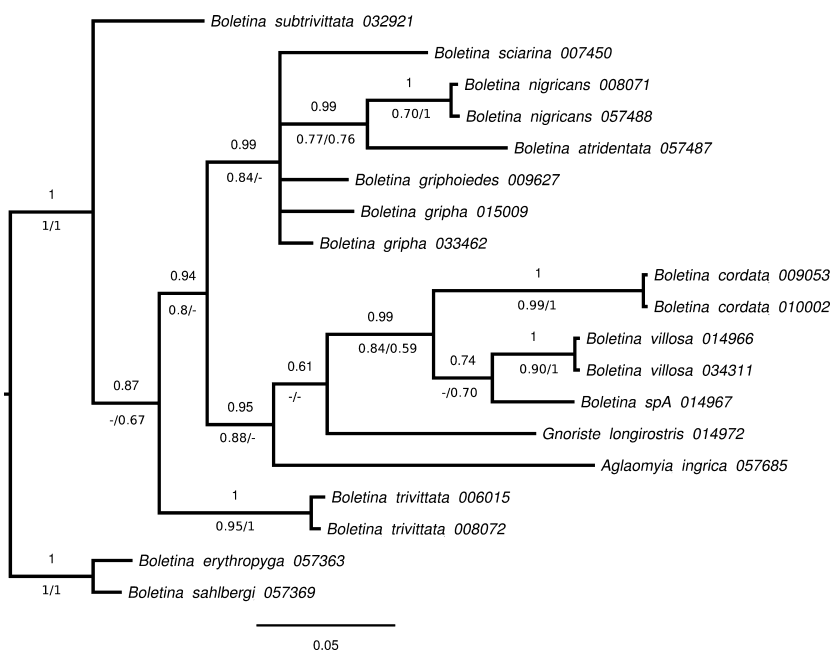

Fig. 4 Maj-Rules tree for Boletina obtained with Bayesian inference, based on CytB sequences. Support values are given in the form: Above branches, Bayesian posterior probabilities, below branches aLRT from ML analysis/ Bootstrap proportions from MP analysis.

Monoclona Mik and *Megalopelma* Enderlein is found as the sister-group to Bolitophilidae and a clade of Diadocidiidae + *Tetragoneura*. Interestingly, when the 18S sequences are added to the dataset most of these ambiguities are resolved (Fig. S3) and the result is much more in accordance with the results obtained by Rindal *et al.* (2009a). The traditional higher classification is then mainly supported and the family Mycetophilidae is retained as monophyletic with the exception of *T. sylvatica* that appears as the sister-group to the family Diadocidiidae.

Discussion

This study suggests that the genus *Boletina* as currently delimited is paraphyletic, thus supporting the findings of Baxter (1999). Søli (1997b) gave three possible synapomorphies for the clade containing *Boletina*, *Coelosia* and *Gnoriste*; viz.: palpomere three with reduced sensory pit, male abdominal segments 7 and 8 strongly reduced, and labrum well developed. These three genera differ greatly in the structure of the male terminalia, mouthparts and to some extend in wing venation and we are not able to add additional uniting morphological synapomorphies to the ones listed above.

The *B. erythropyga* group is consistently found as the sister-group to all other included *Boletina*, *Coelosia* and *Gnoriste*, and compared with the remaining *Boletina* s.l. it could possibly deserve a separate generic status also when differences in morphology are considered (see Zaitzev & Polevoi 2001). The genus *Saigusaia* appears to be well segregated from *Boletina*, at least in the 28S and combined analyses, and is found as the sister-group to *Synapha*, a result in accordance with that of Søli (1997a) based on morphology. Unfortunately only the CytB-sequence was successfully amplified for *Aglaomyia* and we thus failed to present a rigorous test for the segregation of *Aglaomyia* from *Boletina*. The isolated CytB analysis, however, suggests a placement of *Aglaomyia* close to or within *Boletina*, and the morphology of their male terminalia could add evidence for a placement of close to the *B. trivittata* group. We would like to see better data, i.e. additional molecular markers and taxa combined with morphological studies, in order to conclude on consequences for the peculiar nested occurrence of *Gnoriste* and *Coelosia* inside *Boletina* s.l., tasks that are of importance for further work on the systematics of the entire *Boletina*-clade. It will be important also to look further into the role of *Docosia* as a potential outgroup for this clade. The sister-group relationship between *Docosia* and *Palaeodocosia* found by Rindal *et al.* (2009a) is not supported by our data.

The rather ambiguous and partly conflicting results between individual sequences and among different combinations may have several causes. The reason for the large difference in topology when 18S data where added to the combined analyses, compared to when only 16S and 28S data where used, lies probably in a combination of 18S being rather conserved between species and that it is relatively longer and thereby adds much information without adding much homoplasy to the analyses. Rokas *et al.* (2002) tested the utility of several genetic markers for reconstructing phylogenies among cynipid wasps

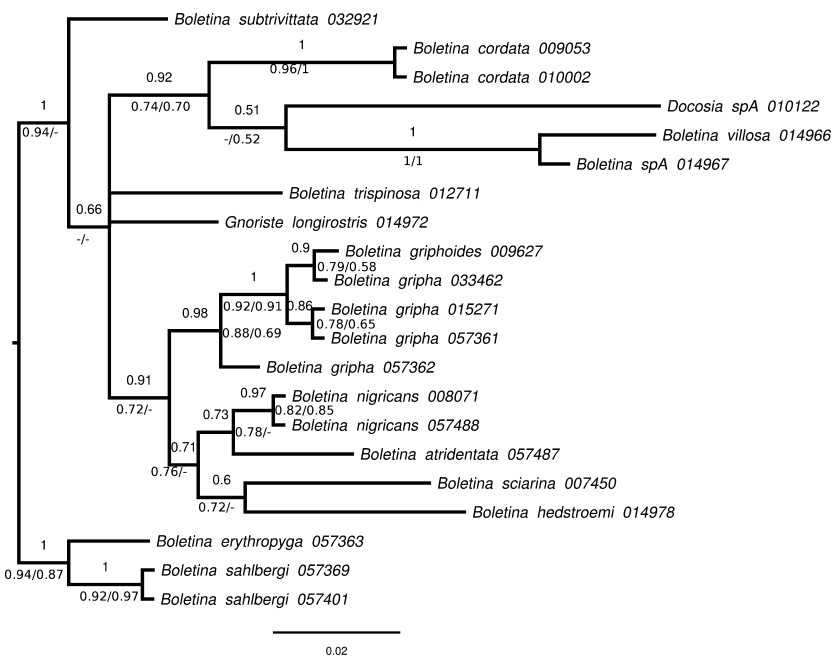

Fig. 5 Maj-Rules tree for Boletina obtained with Bayesian inference, based on 16S sequences. Support values are given in the form: Above branches, Bayesian posterior probabilities, below branches aLRT from ML analysis/Bootstrap proportions from MP analysis.

© 2011 The Authors • Zoologica Scripta © 2011 The Norwegian Academy of Science and Letters,

(Hymenoptera: Cynipidae) and showed that 18S is suitable for resolving above family-level relationships. Here, we find that 18S appears to be informative even at family level and below in Diptera.

The trees obtained from mitochondrial sequences (Figs 1 and 4–5) show more resolution nearer to terminals and lower resolution nearer to the root of the cladograms.

The differences in resolution between mitochondrial and the nuclear 28S sequences are most likely due to differences in mutation rates. In general mutation rates are higher in mitochondrial sequences (Lin & Danforth 2004) that render better resolution of more recent splits but decrease resolution nearer to the root due to increasing amount of homoplasy/saturation in more distantly related taxa. Apart from differences in resolution the trees inferred from 28S, 16S and CytB sequences are mainly congruent while trees inferred from COI differs and are harder to interpreted due to the low resolution. The particular low resolution in the COI analyses (Fig. 5) could be due to the occurrence of nuclear mitochondrial pseudogenes, 'numts' (Song *et al.* 2008), which can be indicated by additional stop codons in the sequence. In these analyses some well-supported relationships are found that are not found in the other analyses, but the possible occurrence of numts needs verification. Numts are known to occur in several animal groups (Song *et al.* 2008; Buhay 2009) and limits the use of a single sequence for DNA-barcoding as proposed by Hebert *et al.* (2003).

Weak taxonomic sampling has been shown to have significant effects on cladogram topology (e.g. Zwickl & Hillis 2002). Future studies would therefore benefit both from an extended taxon sampling, increased sampling length within each gene and from testing out a wider range of genes. In addition to 18S and EF-1α (Catrino *et al.* 2000), the gene CAD has also been studied with promising results for use in Diptera systematics (Moulton & Wiegmann 2004).

The age and storage of material affect the DNA and older material has often shown unsuited for standard DNA extraction and amplification methods due to degradation (e.g. Brammer & von Dohlen 2007; Rindal *et al.* 2007; Almeida & Danforth 2009; Reidenbach *et al.* 2009). This is likely to be the reason for the failure to extract sufficient amount of DNA from a number of specimens. The use of certain pesticides in museum collections has also shown to affect the DNA negatively (Espeland *et al.* 2010).

Acknowledgements

We are grateful to Anna Ansebo and Inger Holmqvist for assistance in the laboratory, to Hokkaido University Museum and T. Saigusa (Fukuoka, Japan) for providing specimens of the genus *Aglaomyia*. Useful suggestions and corrections to the text provided by Peter H. Kerr and an anonymous reviewer are greatly acknowledged.

This study was financially supported by a grant (to PS) from the Swedish Natural Research Council (contract 80565801). JK is financially supported by The Swedish Taxonomy Initiative.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*, 716–723.

Almeida, E. A. B. & Danforth, B. N. (2009). Phylogeny of colletid bees (Hymenoptera: Colletidae) inferred from four nuclear genes. *Molecular Phylogenetics and Evolution*, *50*, 290–309.

Amorim, D. S. & Rindal, E. (2007). Phylogeny of the Mycetophiliformia, with proposal of the subfamilies Heterotrichinae, Ohakuneinae, and Chiletrichinae for the Rangomaramidae (Diptera, Bibionomorpha). *Zootaxa*, *1535*, 1–92.

Anisimova, M. & Gascuel, O. (2006). Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. *Systematic Biology*, *55*, 539–552.

Baxter, J. E. (1999). *Phylogeny of the Sciariodea (Diptera) as Estimated From 16S and 12S Ribosomal RNA Sequences*. PhD Thesis, Berkeley: University of California.

Bechev, D. (2000). World distribution of the genera of fungus gnats (Diptera: Sciaroidea, excluding Sciaridae). *Studia Dipterologica*, 7, 543–552.

Brammer, C. A. & von Dohlen, C. D. (2007). Evolutionary history of Stratiomyidae (Insecta: Diptera): the molecular phylogeny of a diverse family of flies. *Molecular Phylogenetics and Evolution*, *43*, 660–673.

Buhay, J. E. (2009). "COI-like" sequences are becoming problematic in molecular systematic and DNA barcoding studies. *Journal of Crustacean biology*, *29*, 96–110.

Catrino, M. S., Cho, S. & Sperling, F. A. H. (2000). The current state of insect molecular systematics: a thriving tower of babel. *Annual Review of Entomology*, *45*, 1–54.

Colwell, R. K. (2007). *Biota 2 – The Biodiversity Database Manager*. Version 2.04 [computer software], Sunderland, MA, Sinauer Associates. Available via http://viceroy.eeb.uconn.edu/biota (accessed 14 December 2009).

Drummond, A. J., Ashton, B., Cheung, M., Heled, J., Kearse, M., Moir, R., Stones–Havas, S., Thierer, T. & Wilson, A. (2009). *Geneious v4.7*, Available via http://www.geneious.com (accessed 25 February 2010).

Edwards, F. W. (1925). British Fungus-Gnats (Diptera, Mycetophilidae). With a revised generic classification of the family. *Transactions of the Entomological Society of London*, *73*, 505–670.

Espeland, M., Irestedt, M., Johanson, K. A., Åkerlund, M., Bergh, J.-E. & Källersjö, M. (2010). Dichlorvos exposure impedes extraction and amplification of DNA from insects in museum collections. *Frontiers in Zoology*, 7, 8.

Evenhuis, N. L., Pape, T., Pont, A. C. & Thompson, F. C. (Eds). (2008). *Biosystematic Database of World Diptera*, *Version 10.5*. http://www.diptera.org/biosys.htm. (accessed 12 February 2010).

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

Gammelmo, Ø. (2004). Classification of Mycetophilidae (Diptera, Sciaroidea). *Norwegian Journal of Entomology*, *51*, 145–149.

Guindon, S. & Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, *52*, 696–704.

Guindon, S., Dufayard, J. F., 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. *Systematic Biology*, *59*, 301–321.

Hebert, P. D. N., Cywinska, A., Ball, S. L. & deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society London Series B*, *270*, 313–322.

Hippa, H. & Vilkamaa, P. (2005). The genus Sciarotricha gen. n. (Sciaridae) and the phylogeny of recent and fossil Sciaroidea (Diptera). *Insect Systematics & Evolution*, *36*, 121–144.

Hippa, H. & Vilkamaa, P. (2006). Phylogeny of the Sciaroidea (Diptera): the implication of additional taxa and character data. *Zootaxa*, *1132*, 63–68.

Huelsenbeck, J. P. & Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, *17*, 754–755.

Hutson, K. M., Ackland, D. M. & Kidd, L. N. (1980). Mycetophilidae (Bolitophilinae, Ditomyiinae, Diadocidiinae, Keroplatinae, Sciophilinae and Manotinae) Diptera, Nematocera. In: Fitton, M. G. (Ed.) *Handbooks for the Identification of British Insects*, Vol. IX, Part 3. 11 (pp. 1–111) London: Royal Entomological Society of London.

Jakovlev, J. (1995). Species diversity and abundance of fungivorous Diptera in forests and city parks of Russian Karelia. *An International Journal of Dipterological Research*, *6*, 335–362.

Jakovlev, J. & Penttinen, J. (2007). *Boletina dispectoides* sp.n. and six other species of fungus gnats (Diptera: Mycetophilidae) new to Finland. *Entomologica Fennica*, *18*, 211–217.

Jakovlev, J., Kjærandsen, J. & Viklund, B. (2008). Fungus gnats (Diptera: Bolitophilidae, Diadocidiidae, Ditomyiidae, Keroplatidae & Mycetophilidae) from Tyresta National Park and Nature Reserve in Sweden. *Sahlbergia*, *14*, 29–52.

Katoh, K., Kuma, K., Toh, H. & Miyata, T. (2005). MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Research*, *33*, 11–518.

Kjærandsen, J., Hedmark, K., Kurina, O., Polevoi, K., Økland, B. & Götmark, F. (2007). Annotated checklist of fungus gnats from Sweden (Diptera: Bolitophilidae, Diadocidiidae, Ditomyiidae, Keroplatidae and Mycetophilidae). *Insect Systematics and Evolution Supplements*, *65*, 1–128.

Lin, Chung-Ping & Danforth, Bryan N. (2004). How do insect nuclear and mitochondrial gene substitution patterns differ? Insights from Bayesian analyses of combined datasets. *Molecular Phylogenetics and Evolution*, *30*, 686–702.

Marshall, D. C. (2010). Cryptic failure of partitioned Bayesian phylogenetic analyses: lost in the land of long trees. *Systematic Biology*, *59*, 108–117.

Moulton, J. K. & Wiegmann, B. M. (2004). Evolution and phylogenetic utility of CAD (rudimentary) among Mesozoic-aged Eremoneuran Diptera (Insecta). *Molecular Phylogenetics and Evolution*, *31*, 363–378.

Nylander, J. A. A. (2004). *MrModeltest v2*. Program distributed by the author. Uppsala: Evolutionary Biology Centre, Uppsala University.

Nylander, J. A. A., Wilgenbusch, J. C., Warren, D. L. & Swofford, D. L. (2008). AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics*, *24*, 581–583.

Økland, B. (1994). Mycetophilidae (Diptera), an insect group vulnerable to forestry? A comparison of clearcut, managed and semi-natural spruce forests in southern Norway. *Biodiversity and Conservation*, *3*, 68–85.

Økland, B. (1999). New rearing records of forest–dwelling Diptera. *An International Journal of Dipterological Research*, *10*, 133–146.

Pleijel, F., Jondelius, U., Norlinder, E., Nygren, A., Oxelman, B., Schander, C., Sundberg, P. & Thollesson, M. (2008). Phylogenies without roots? A plea for the use of vouchers in molecular phylogenetic studies. *Molecular Phylogenetics and Evolution*, *48*, 369–371.

Polevoi, A. V. & Hedmark, K. (2004). New species of the genus *Boletina* Winnertz (Diptera: Mycetophilidae) from Fennoscandia. *Entomologica Fennica*, *15*, 23–33.

Rambout, A. (2009). *FigTree v.1.3.1*. http://tree.bio.ed.ac.uk/software/figtree/ (accessed 2 June 2010).

Reidenbach, K. R., Cook, S., Bertone, M. A., Harbach, R. E., Wiegmann, B. M. & Besansky, N. J. (2009). Phylogenetic analysis and temporal diversification of mosquitoes (Diptera: Culicidae) based on nuclear genes and morphology. *BMC Evolutionary Biology*, *9*, 298.

Rindal, E., Søli, G. E. E., Kjærandsen, J. & Bachmann, L. (2007). Molecular phylogeny of the fungus gnat tribe Exechiini (Mycetophilidae, Diptera). *Zoologica Scripta*, *36*, 327–335.

Rindal, E., Søli, G. E. E. & Bachmann, L. (2009a). Molecular phylogeny of the fungus gnat family Mycetophilidae (Diptera, Mycetophiliformia). *Systematic Entomology*, *34*, 524–532.

Rindal, E., Søli, G. E. E. & Bachmann, L. (2009b). On the systematics of the fungus gnat subfamily Mycetophilinae (Diptera): a combined morphological and molecular approach. *Journal of Zoological Systematics and Evolutionary Research*, *47*, 227–233.

Rokas, A., Nylander, J. A. A., Ronquist, F. & Stone, G. N. (2002). A maximum-likelihood analysis of eight phylogenetic markers in gallwasps (Hymenoptera: Cynipidae): implications for insect phylogenetic studies. *Molecular Phylogenetics and Evolution*, *22*, 206–219.

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

Russell-Smith, A. (1979). A study of fungus flies (Diptera: Mycetophilidae) in beech woodland. *Ecological Entomology*, *4*, 355–364.

Smith, S. A. & Dunn, C. W. (2008). Phyutility: a phyloinformatics tool for trees, alignments, and molecular data. *Bioinformatics*, *24*, 715–716.

Søli, G. E. E. (1997a). The adult morphology of Mycetophilidae (s.str.), with a tentative phylogeny of the family (Diptera, Sciaroidea). *Entomologica Scandinavica Supplement*, *50*, 5–55.

Søli, G. E. E. (1997b). The systematics and phylogeny of *Coelosia* Winnertz, 1863 (Diptera, Mycetophilidae). *Entomologica Scandinavica Supplement*, *50*, 57–139.

Søli, G. E. E., Vockeroth, R. J. & Matile, L. (2000). K.4. Families of Sciaroidea. In L. Papp & B. Darvas (Eds) *Contribution to a Manual of Palaearctic Diptera. Appendix* (pp. 49–92). Budapest: Science Herald.

Song, H., Buhay, J. E., Whiting, M. F. & Crandall, K. A. (2008). Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. *PNAS*, *105*, 13486–13491.

Swofford, D. L. (2002). *PAUP*. Phylogenetic Analysis Using Parsimony (* and Other Methods)*. Version 4.0b10. Sunderland, MA: Sinauer Associates.

Tuomikoski, R. (1966). Generic taxonomy of the Exechiini (Diptera, Mycetophilidae). *Annales Entomologici Fennici*, *32*, 159–194.

Väisänen, R. (1986). The delimitation of the Gnoristinae: criteria for classification of recent European genera (Dipetra, Mycetophilidae). *Annales Zoologici Fennici*, *23*, 197–206.

Vockeroth, J. R. (1980). New genera and species of Mycetophilidae (Diptera) from the Holarctic region, with notes on other species. *The Canadian Entomologist*, *112*, 529–544.

Werle, E., Schneider, C., Renner, M., Völker, M. & Fiehn, W. (1994). Convenient single-step, one tube purification of PCR products for direct sequencing. *Nucleic Acids Research*, *22*, 4354–4355.

Zaitzev, A. I. & Polevoi, A. V. (1995). New species of fungus gnats (Diptera: Mycetophilidae) from the Kivach Nature Reserve. Russian Karelia. *Entomologica Fennica*, *6*, 185–195.

Zaitzev, A. I. & Polevoi, A. V. (2001). Holarctic species of the *Boletina erythropyga*–group (Diptera, Mycetophilidae). *Studia dipterologica*, *8*, 639–644.

Zaitzev, A. I., Jakovlev, J. & Polevoi, A. V. (2005). Palaearctic species of the *Boletina nitida*–group (Diptera: Mycetophilidae) including the description of a new species. *Studia dipterologica*, *12*, 243–253.

Zwickl, D. J. & Hillis, D. M. (2002). Increased taxon sampling greatly reduces phylogenetic error. *Systematic Biology*, *51*, 588–598.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Additional information about the specimens used for DNA extraction in this study. The data is extracted from a Biota 2.04 database. For each species and country the localities are sorted hierarchically within provinces, districts, localities and sites respectively. Unless otherwise stated the voucher specimens are deposited at the Museum of Zoology, Lund University (MZLU).

Table S1. Primers used in this study for PCR amplification and sequencing of the mitochondrial 16S, COI and CytB, and the nuclear and 28S genes.

Table S2. Additional sequences obtained from GenBank used in phylogenetic analyses, sorted in alphabetical order within families.

Fig. S1 Phylogenetic tree for Mycetophilidae obtained from Maximum likelihood analysis, based on 16S and 28S sequences. Values at nodes are aLRT support values

Fig. S2 Maj-Rule tree for Mycetophilidae obtained with Bayesian inference, based on 16S and 28S sequences. Values above branches are Bayesian posterior probabilities.

Fig. S3 Phylogenetic tree for Mycetophilidae obtained from Maximum likelihood analysis, based on 16S, 18S and 28S sequences. Values at nodes are aLRT support values.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.