Molecular Phylogenetics and Evolution 58 (2011) 513–526

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Multilocus analysis of a taxonomically densely sampled dataset reveal extensive non-monophyly in the avian family Locustellidae

Per Alström [a,*], Silke Fregin [b], Janette A. Norman [c], Per G.P. Ericson [d], Les Christidis [c,e], Urban Olsson [f]

[a] Swedish Species Information Centre, Swedish University of Agricultural Sciences, Box 7007, SE-750 07 Uppsala, Sweden
[b] Vogelwarte Hiddensee, Zoological Institute and Museum, Ernst Moritz Arndt University of Greifswald, 17489 Greifswald, Germany
[c] Department of Genetics, University of Melbourne, Parkville Vic. 3052 Australia and Sciences Department, Museum Victoria, GPO Box 666, Melbourne Vic. 3001, Australia
[d] Department of Vertebrate Zoology and Molecular Systematics Laboratory, Swedish Museum of Natural History, P.O. Box 50007, SE-104 05 Stockholm, Sweden
[e] National Marine Science Centre, Southern Cross University, NSW 2450, Australia
[f] Department of Zoology, University of Gothenburg, Box 463, SE-405 30 Göteborg, Sweden

ARTICLE INFO

Article history:
Received 12 October 2010
Revised 17 December 2010
Accepted 20 December 2010
Available online 15 January 2011

Keywords:
Species trees
Gene trees
Phylogeny
Concatenation
BEST
MrBayes
Classification

ABSTRACT

The phylogeny of most of the species in the avian passerine family Locustellidae is inferred using a Bayesian species tree approach (Bayesian Estimation of Species Trees, BEST), as well as a traditional Bayesian gene tree method (MrBayes), based on a dataset comprising one mitochondrial and four nuclear loci. The trees inferred by the different methods agree fairly well in topology, although in a few cases there are marked differences. Some of these discrepancies might be due to convergence problems for BEST (despite up to 1×10^9 iterations). The phylogeny strongly disagrees with the current taxonomy at the generic level, and we propose a revised classification that recognizes four instead of seven genera. These results emphasize the well known but still often neglected problem of basing classifications on non-cladistic evaluations of morphological characters. An analysis of an extended mitochondrial dataset with multiple individuals from most species, including many subspecies, suggest that several taxa presently treated as subspecies or as monotypic species as well as a few taxa recognized as separate species are in need of further taxonomic work.

1. Introduction

The avian family Sylviidae (“Old World warblers”) has long been recognized as one of the main passerine families, although the composition has varied among authors. Traditionally, a large number of taxa were included, e.g. 60 genera and 358 species in the classification of Watson et al. (1986). Sibley and Monroe (1990), based on the DNA–DNA hybridization work by Sibley and Ahlquist (1990), split off Cisticolidae from Sylviidae, and further divided Sylviidae into the subfamilies Megalurinae, Acrocephalinae and Sylviinae. This was followed by Dickinson (2003) and Bairlein et al. (2006). Later studies, based on DNA sequence data, revised this classification. Alström et al. (2006) and Johansson et al. (2008) proposed recognition of a number of well supported major clades at family level. These authors synonymized Sylviidae with the family Timaliidae (“babblers”). Gelang et al. (2009), again based on DNA sequence data, resurrected Sylviidae, but restricted it to a clade containing mainly traditional Timaliidae species.

The subfamily Megalurinae sensu Sibley and Monroe (1990) contained the genera *Megalurus*, *Cincloramphus*, *Eremiornis*, *Amphilais*, *Megalurulus*, *Buettikoferella*, *Chaetornis*, *Graminicola* and *Schoenicola*. In contrast, the family Megaluridae sensu Alström et al. (2006) and Johansson et al. (2008) comprised the genera *Megalurus*, *Bradypterus*, *Locustella* and *Dromaeocercus*, i.e. including three of the genera placed in Acrocephalinae by Sibley and Monroe (1990). Other DNA sequence studies have shown that *Cincloramphus* and *Schoenicola* form a clade with *Bradypterus* and *Megalurus* (Beresford et al., 2005), while *Graminicola* belongs to the babbler family Timaliidae (Alström et al., 2006; Gelang et al., 2009). Beresford et al. (2005) also revealed that the aberrant *Bradypterus victorini* is not related to Megaluridae/Megalurinae.

The name Locustellinae Bonaparte, 1854, has priority over Megalurinae Blyth, 1875 (Bock, 1994: p. 152), and thus the family name Locustellidae Bonaparte, 1854 is applied in the present paper for Megaluridae sensu Alström et al. (2006) and Johansson et al. (2008). The relationships within this family are poorly known. Drovetski et al. (2004) used mitochondrial ND2 to study the relationships of all *Locustella*, two Asian and three African *Bradypterus*, and two *Megalurus*. They found the Asian *Bradypterus* and *Megalurus pryeri* nested within *Locustella*, the African *Bradypterus* in a separate clade, and *M. gramineus* on a branch on its own.

* Corresponding author.
E-mail addresses: per.alstrom@slu.se (P. Alström), silke.fregin@uni-greifswald.de (S. Fregin), janorman@unimelb.edu.au (J.A. Norman), per.ericson@nrm.se (P.G.P. Ericson), les.christidis@gmail.com (L. Christidis), urban.olsson@zool.gu.se (U. Olsson).

1055-7903/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2010.12.012

The species in Locustellidae are distributed across Africa, Eurasia and Australasia, frequenting mostly bushy, but sometimes also marshy, habitats from sea level up to above the tree limit (c. 4500 m in the Himalayas) (Bairlein et al., 2006). Most species are notoriously secretive and difficult to observe. All are non-descript, mostly various shades of brown above and at least slightly paler below; *Megalurus*, *Cincloramphus* and some *Locustella* are streaked above, some of these and some *Bradypterus* also on the underparts (Bairlein et al., 2006). *Cincloramphus cruralis* is exceptional in that the male is uniformly dark sooty brown below (Bairlein et al., 2006). Most species are fairly small, with an overall length of 13–16 cm, but some are considerably larger (22–28 cm in *Megalurus palustris*) (Bairlein et al., 2006). The songs are mostly simple but distinctive, and in general differ more than morphology among closely related species (Bairlein et al., 2006). Due to the generally cryptic plumages, there has been much confusion regarding species level taxonomy (e.g. Dickinson et al., 2000), and recent studies involving vocalizations and/or DNA have led to suggestions that some taxa currently treated as subspecies should be raised to the rank of species (e.g. Drovetski et al., 2004; Alström et al., 2008) as well as to the identification of a new cryptic species (Rasmussen et al., 2000).

In the present study, we infer the relationships of nearly all species in the family Locustellidae using one mitochondrial gene and four nuclear introns. We use traditional gene tree methods (Bayesian inference, maximum likelihood bootstrapping, parsimony bootstrapping) as well as a Bayesian species tree approach (Bayesian Estimation of Species Trees, BEST; Liu and Pearl, 2007; Liu, 2008) that accounts for lineage sorting processes that might produce discordance between gene trees. We also analyse mitochondrial DNA for a larger sample, comprising multiple individuals and several subspecies of polytypic species. A revised taxonomy is proposed based on our results.

2. Materials and methods

2.1. Study group

In total, we include 37 species from seven genera considered to belong to Locustellidae (=Megaluridae sensu Alström et al., 2006 and Johansson et al., 2008). Our sample comprises 16 species of *Bradypterus* plus cytochrome *b* (cyt*b*) sequences for three additional species (two from GenBank and one provided by Trevor Price and Udayan Borthakur; only two African and three Asian species are missing); all eight *Locustella* species; four *Megalurus* species plus cyt*b* for one more species (two species are lacking); both species of *Cincloramphus*; one of the two species of *Schoenicola*; and the monotypic genera *Dromaeocercus* and *Eremiornis*. For cyt*b*, we have in total 82 unique haplotypes, including 24 sequences from GenBank, comprising several taxa treated as subspecies of polytypic species. Sequences from four nuclear markers (ODC, myo, GAPDH, LDH) were obtained for most taxa (see Appendix A for details regarding loci coverage across the taxa).

Species level taxonomy follows Dickinson (2003) and Bairlein et al. (2006), with the exception of the recognition of *Bradypterus thoracicus kashmirensis* as a distinct species, based on a study of morphology, vocalizations and mitochondrial DNA (Alström et al., 2008).

2.2. DNA extraction and sequencing

DNA was extracted from blood, feathers, or muscle, using QIA Quick DNEasy Kit (Qiagen, Inc.) according to the manufacturer's instruction, but with 30 µl 0.1% DTT added to the initial incubation step of the extraction of feathers. We sequenced five loci: the main part of the mitochondrial cytochrome *b* gene and part of the flanking tRNA-Thr (cyt*b*); the nuclear ornithine decarboxylase exon 6 (partial), intron 6, exon 7, intron 7 and exon 8 (partial) (ODC); the entire nuclear myoglobin intron 2 (myo), the nuclear glyceraldehyde-3-phosphodehydrogenase intron 11 (GAPDH), and the complete nuclear lactate dehydrogenase intron 3 (LDH). Amplification and sequencing of cyt*b* and myo followed the protocols described in Olsson et al. (2005), of ODC Allen and Omland (2003), of GAPDH Fjeldså et al. (2003), and of LDH Fregin et al. (2009). Cyt*b* was amplified as one fragment to decrease the risk of amplifying nuclear pseudocopies (e.g. Sorensen and Quinn, 1998). DNA was also extracted from one museum specimen (*Schoenicola brevirostris*). For extraction, PCR-amplification, and sequencing procedures from this one, the procedures described in Irestedt et al. (2006) were followed, with specially designed primers obtainable from the authors upon request. All new sequences have been deposited in GenBank (Appendix A).

2.3. Phylogenetic analyses

Sequences were aligned using MegAlign 4.03 in the DNASTAR package (DNAstar Inc.); some manual adjustment was necessary for the non-coding sequences. For the nuclear loci, haplotypes were not separated, but coded as ambiguous bases.

Gene trees were estimated by Bayesian inference (BI) using MrBayes 3.1.2 (Huelsenbeck and Ronquist, 2001, 2005) according to the following: (1) All loci were analysed separately (single-locus analyses, SLAs). (2) Sequences were also concatenated, either all nuclear loci, or all loci together. In the multilocus analyses, the data were either (a) partitioned by locus, using rate multipliers to allow different rates for the different partitions (Nylander et al., 2004; Ronquist and Huelsenbeck, 2003), or (b) unpartitioned, using a homogeneous model for the entire dataset. In the analyses of all loci, species with missing data were included or excluded in various constellations. Ambiguous base pairs and indels were treated as missing data, but indels were plotted on the trees a posteriori. As outgroups, two species belonging to the family Bernieridae (*Hartertula flavoviridis* and *Thamnornis chloropetoides*) were chosen, as this family has been suggested to be sister to Locustellidae (Beresford et al., 2005; Johansson et al., 2008). Analyses were also run with 28 outgroup species, representing all families in the superfamily Sylvioidea (Alström et al., 2006; Johansson et al., 2008).

Appropriate substitution models were determined based on the Akaike Information Criterion (Akaike, 1974) and a hierarchical likelihood ratio test (Posada and Crandall, 1998), both calculated using MrModeltest2 (Nylander, 2004) in conjunction with PAUP* (Swofford, 2002). For all loci, posterior probabilities (PPs) were calculated under the general time-reversible (GTR) model (Lanave et al., 1984; Tavaré, 1986; Rodríguez et al., 1990), assuming rate variation across sites according to a discrete gamma distribution with four rate categories (Γ; Yang, 1994) and, for the cyt*b* data, also an estimated proportion of invariant sites (I; Gu et al., 1995). Default priors in MrBayes were used. Four Metropolis-coupled MCMC chains with incremental heating temperature 0.1 or 0.2 were run for 10–30 $\times$ 10^6 generations and sampled every 1000 generations. Chain likelihood and other parameter values and effective sample sizes (>200, generally >1000) were inspected in Tracer 1.5.0 (Rambaut and Drummond, 2009). The first 25% of the generations were discarded as "burn-in", well after stationarity of chain likelihood values had been established, and the posterior probability was estimated for the remaining generations. Every analysis was run at least twice, and the topologies and posterior probabilities compared by eye.

Species tree analysis was performed using Bayesian Estimation of Species Trees (BEST) 2.3 (Liu and Pearl, 2007; Liu, 2008). Only

species with complete data were included. Two long analyses were run, each with four Metropolis-coupled MCMC chains running 1×10^9 cycles. In addition, eight shorter analyses, each c. 7×10^7–1×10^8 replicates, were run. All analyses were sampled every 1000 generations, and the incremental heating temperature was set to 0.1. The theta prior was set to invgamma (3, 0.003) and the GeneMu prior to uniform (0.5, 1.5). The posterior distribution was summarized based on the generations with the highest, seemingly stable, likelihood values. *H. flavoviridis* and *T. chloropetoides* were again used as outgroups.

In addition, clades B, C and D identified in the BI (Figs 1–3) were analysed separately by BEST in order to try and get better convergence than in the more extensive BEST analyses. These analyses were run for 2×10^8 generations, all else being equal to the other BEST. The outgroups were the same as in the other BEST analyses.

Maximum likelihood bootstrapping (MLB) (1000 replicates) was performed on the complete dataset in Treefinder (version of October 2008; Jobb et al., 2004; Jobb, 2008) using default settings and a uniform GTR+Γ+I model. Parsimony bootstrapping (MPB) was performed in PAUP* (Swofford, 2002) on the complete dataset:

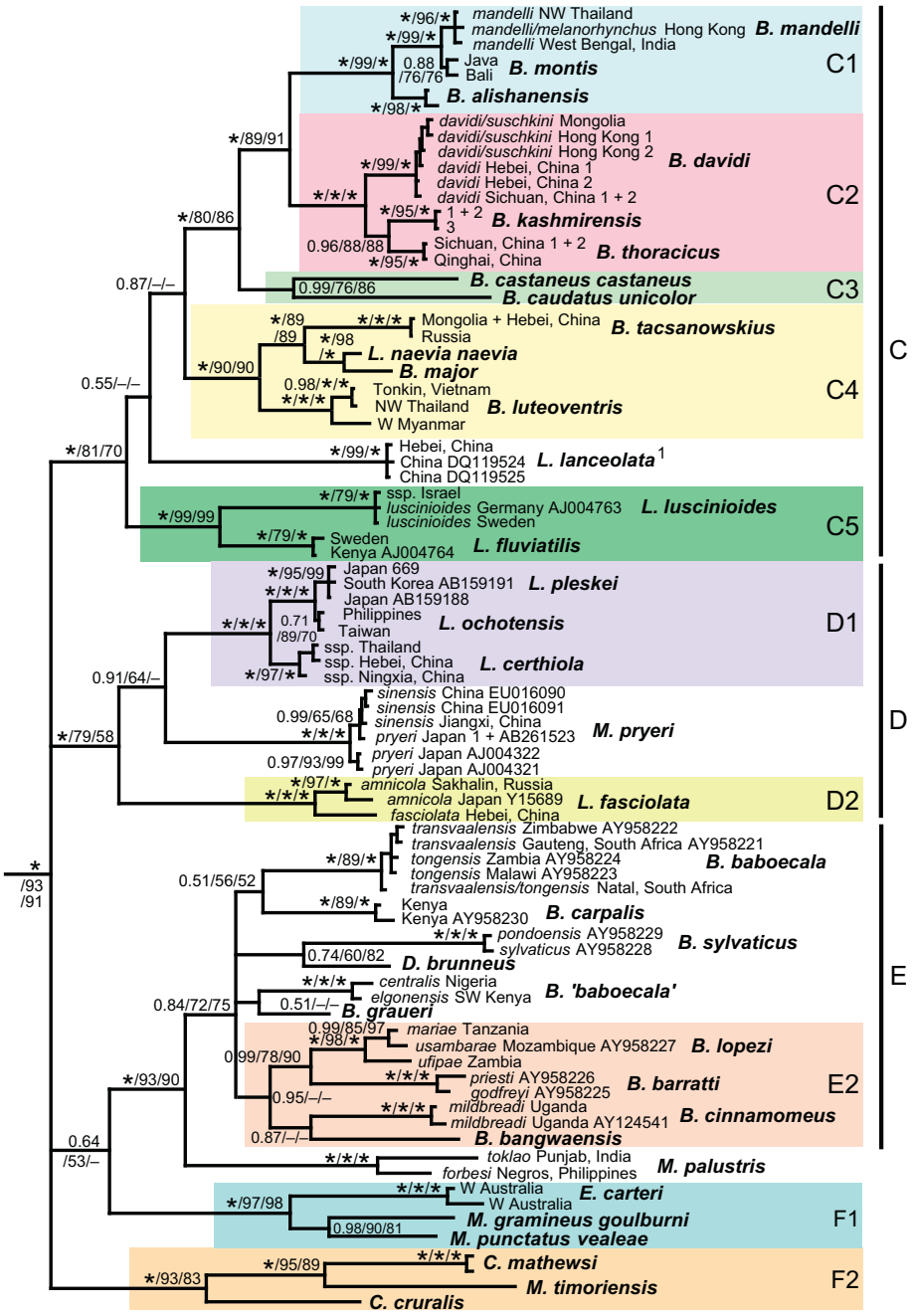

Fig. 1. Majority rule (50%) consensus tree of Locustellidae based on unique mitochondrial cytochrome *b* haplotypes, inferred by Bayesian inference under the GTR+Γ+I model. Posterior probabilities, and maximum likelihood and parsimony bootstrap values are indicated at the nodes, in this order; an asterisk represents posterior probability 1.0 or bootstrap 100%. The species for which no subspecific names are given are monotypic (except *L. lanceolata*, see below). Bars and colour shading delimit clades discussed in text. *B.* = *Bradypterus*, *C.* = *Cincloramphus*, *D.* = *Dromaeocercus*, *E.* = *Eremiornis*, *L.* = *Locustella*, and *M.* = *Megalurus*. Numbers after names are sample identifiers (e.g. *davidi* Sichuan, China 1 + 2 means *davidi* samples 1 and 2 from Sichuan, China [same haplotype], and *sinensis* China EU016090 refers to GenBank number of sequence previously used in another study; see Appendix A). [1]On geographical grounds, most likely nominate subspecies, but samples collected during migration, so subspecies *hendersonii* cannot be eliminated.

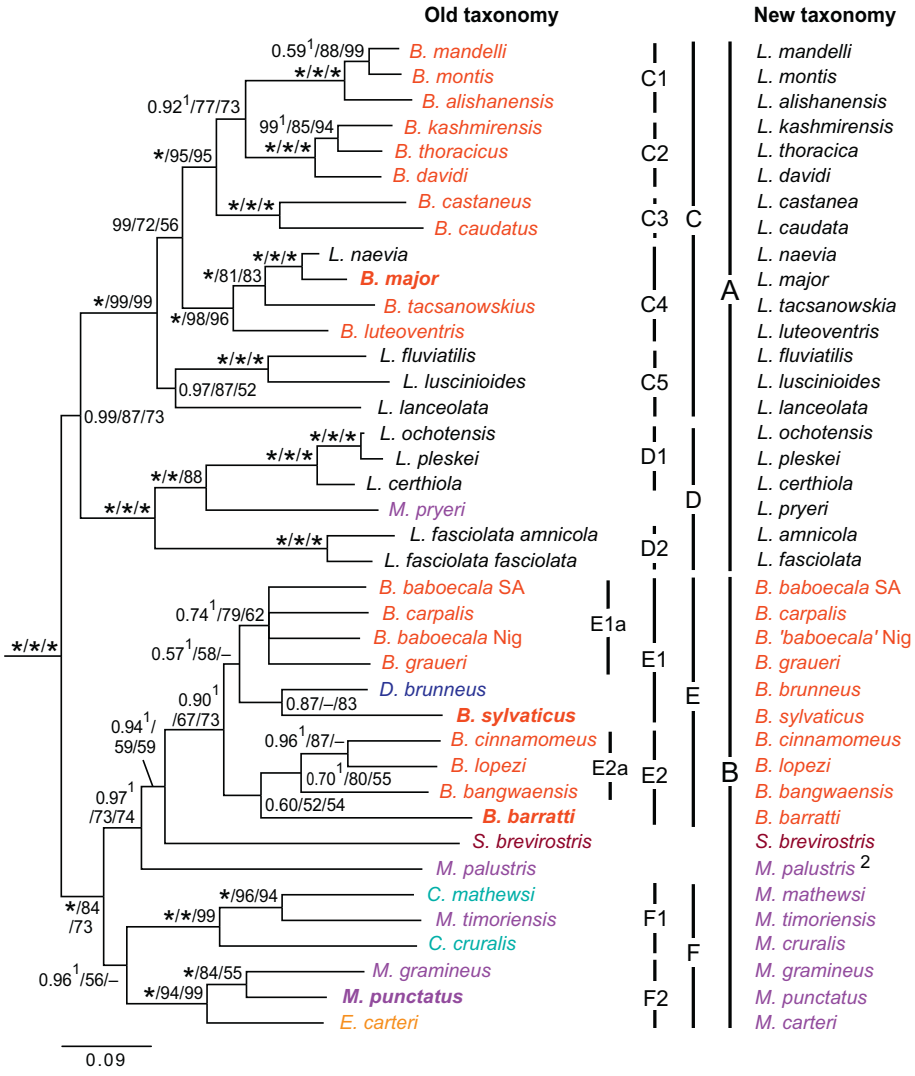

Fig. 2. Majority rule (50%) consensus tree of Locustellidae based on concatenated nuclear ODC, myoglobin, LDH and GAPDH and mitochondrial cytochrome *b* sequences, inferred by Bayesian inference, analysed in five partitions (four nuclear loci GTR+Γ, cytochrome *b* GTR+Γ+I). Colours of names indicate genus according to old taxonomy (Dickinson 2003; left) and new taxonomy proposed here (right; see Fig. 1 for explanation of abbreviations of generic names; *S.* = *Schoenicola*). Labelled bars denote clades as in Fig. 1. The four species for which only cytochrome *b* is available are in bold type. Posterior probabilities, and maximum likelihood and parsimony bootstrap values are indicated at the nodes, in this order; an asterisk represents posterior probability 1.0 or bootstrap 100%. *B. baboecala* SA and *B. baboecala* Nig refer to samples from South Africa (*transvaalensis/tongensis*) and Nigeria (*centralis*), respectively. [1]Node affected differently by different types of analyses (see Table 1). [2]See Section 4 for recognition of non-monophyletic *Megalurus*.

heuristic search strategy, 1000 replicates, starting trees obtained by stepwise addition (random addition sequence, 10 replicates), TBR branch swapping, MulTrees option not in effect (only one tree saved per replicate).

Bayes factors (Newton and Raftery, 1994; Kass and Raftery, 1995) were calculated in Tracer 1.5.0 (Rambaut and Drummond, 2009) for comparisons of alternative hypotheses in some BI analyses.

GTR+Γ+I corrected pairwise divergences for the cyt*b* dataset (excluding outgroup species) were calculated in Treefinder (version of October 2008; Jobb et al., 2004; Jobb, 2008). Positions where one or more taxa had ambiguous nucleotides were deleted from the matrix, and incomplete sequences were excluded, or the ends were trimmed, so that all sequences used in the comparisons comprised 982 base pairs.

2.4. Summary of abbreviations

BI – Bayesian inference (MrBayes); BIC – Bayesian inference (MrBayes) of concatenated sequences; cyt*b* – cytochrome *b* gene; GAPDH – glyceraldehyde-3-phosphodehydrogenase intron 11; LDH – lactate dehydrogenase intron 3; MLB – maximum likelihood bootstrap; MPB – parsimony bootstrap; myo – myoglobin intron 2; ODC – ornithine decarboxylase (mainly) introns 6–7; PP – posterior probability; SLA – single-locus analysis.

3. Results

3.1. Sequence characteristics

We obtained a contiguous ≤730 base pair (bp) stretch of the ODC, ≤709 bp of myo, ≤510 bp of the LDH, ≤375 bp of the GAPDH, and ≤1038 bp of cyt*b*. No unexpected stop codons, indels, or distinct double peaks in the chromatograms that would indicate the presence of nuclear pseudogenes were found in the coding cyt*b* sequences, except for one sequence from GenBank: the *Locustella fasciolata amnicola* sequence with Genbank No. Y15689 contains a stop codon (AGA). However, this individual proved to be sister to the other individual of the same taxon, and these two form the sister clade to *L. fasciolata fasciolata*, so either the paralogue is of recent origin or the G in the stop codon is a misreading in

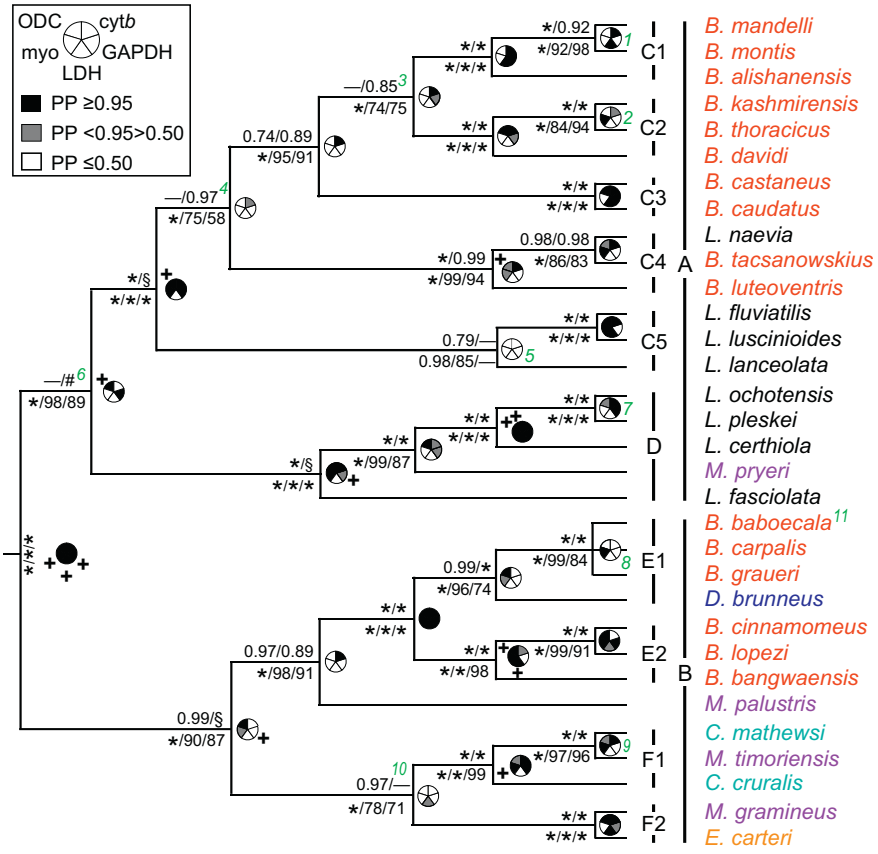

Fig. 3. Consensus phylogeny of Locustellidae based on analyses of nuclear ODC, myoglobin, LDH and GAPDH and mitochondrial cytochrome *b* sequences. Only species for which all loci are available are included. Values above branches represent species tree inferences (BEST posterior probability: entire dataset/separate analyses of clades B–D), and values below branches represent analyses based on concatenation (Bayesian posterior probability [unpartitioned: GTR+Γ+I]/maximum likelihood bootstrap/parsimony bootstrap). * indicates posterior probability 1.0 or bootstrap 100%; § indicates result from BEST not applicable, since only two distant outgroups were used in separate analyses of clades B–D; # indicates that no BEST was performed on clade A separately. Pie charts at nodes denote support in single-locus analyses (see explanation in upper left corner); + indicates further support by indel (in the locus whose pie is adjacent to the +; cf. Supplementary Fig. 3). Labelled bars denote clades as in Figs. 1 and 2. Colours of names indicate genus according to old taxonomy (Dickinson, 2003); see Figs. 1 and 2 for explanation of abbreviations of generic names. [1]In single-locus analysis of both cyt*b* and LDH this clade has PP 1.0, whereas in analysis of ODC clade (*B. mandelli*, *B. alishanensis*) has PP 0.99. [2]In single-locus analysis of cyt*b* and myo, this clade has PP 0.86 and 0.99, respectively, whereas in analysis of ODC, clade (*B. davidi*, *B. kashmirensis*) has PP 1.0. [3]In BEST analysis of entire dataset, clade C2 is sister to clade C3 with PP 0.80. [4]In BEST analysis of entire dataset, clades C1–C3, C4 and C5 form a trichotomy. [5]Not supported by any single-locus analysis. [6]In BEST of complete dataset, clade D is sister to the rest with PP 0.67. [7]In myo tree, *L. ochotensis* and *L. certhiola* are sisters with PP 1.0. [8]In LDH tree, clade (*B. baboecala*, *B. graueri*) has PP 0.97, while in myo tree (*B. baboecala*, *B. carpalis*) has PP 0.94. [9]In GAPDH tree, *C. cruralis* and *M. timoriensis* are sisters with PP 0.99. [10]See text. [11]*B. baboecala* concerns subspecies *transvaalensis/tongensis* from Natal, South Africa.

the original sequence. Including all outgroup taxa, the aligned ODC sequences comprise 762 characters, of which 203 (27%) are parsimony informative; myo 758 characters, 151 (20%) parsimony informative; LDH 541 characters, 164 (30%) parsimony informative; GAPDH 400 characters, 95 (24%) parsimony informative; and cyt*b* 1078 characters, 445 (41%) parsimony informative. Including all outgroups, the combined ODC, myo, LDH and GAPDH (hereafter nuclear) data set contains 2461 characters, of which 613 (25%) are parsimony informative, and the total data set comprises 3539 characters, of which 1058 (30%) are parsimony informative. Excluding the outgroup taxa, the aligned ODC sequences comprise 734 characters, of which 73 (9.9%) are parsimony informative; myo 709 characters, 43 (6.1%) parsimony informative; LDH 513 characters, 40 (7.8%) parsimony informative; GAPDH 379 characters, 35 (9.2%) parsimony informative; and cyt*b* 1038 characters, 348 (33.5%) parsimony informative. Excluding outgroups, the combined ODC, myo, LDH and GAPDH (hereafter nuclear) dataset contains 2335 characters, of which 191 (8.2%) are parsimony informative, and the total dataset comprises 3373 characters, of which 539 (16.0%) are parsimony informative. The cyt*b* dataset comprising multiple samples of most species includes 384 parsimony-informative characters (37.0%).

3.2. Single-locus analyses

The tree containing multiple cyt*b* haplotypes for most Locustellidae species, including several subspecies and species for which only cyt*b* is available, is overall well resolved (88% of species nodes), except for some, mostly deep, internal nodes (Fig. 1).

The trees based on single-locus analyses (hereafter SLAs) of single sequences per species vary in resolution: 85% of the nodes are bifurcating in the cyt*b* tree, 75% in the ODC tree, 59% in the myo tree, 68% in the LDH tree, and 55% in the GAPDH tree (Supplementary Fig. 1; see also Fig. 3, where SLAs are shown in pie charts). Although the resolution varies among these trees, they generally agree fairly well, and there are few strongly supported topological conflicts. Only four conflicting reconstructions receive ≥0.95 posterior probability (PP) in different trees: (1) *Bradypterus thoracicus* and *Bradypterus kashmirensis* are sisters according to myo (0.99) and cyt*b* (0.86; 0.96 in analysis with multiple individuals, Fig. 1), while *B. kashmirensis* and *Bradypterus davidi* are sisters according to ODC (1.0); (2) *Bradypterus mandelli* and *Bradypterus montis* are sisters according to LDH and cyt*b* (both 1.0), whereas *B. mandelli* and *Bradypterus alishanensis* are sisters according to ODC (0.99); (3) *Locustella ochotensis* and *Locustella pleskei* are sisters in the

ODC (0.57), LDH (0.98) and cyt*b* trees (1.0), while *L. ochotensis* and *Locustella certhiola* are sisters in the myo tree (1.0); and (4) *Cincloramphus mathewsi* and *Megalurus timoriensis* are sisters according to ODC (0.87), myo (1.0) and cyt*b* (1.0), whereas *C. cruralis* and *M. timoriensis* are sisters according to GAPDH (0.99).

3.3. Concatenated multilocus analyses

In the BI analyses with 28 outgroup species, representing all major clades in Sylvioidea, the ingroup is shown to be monophyletic, with strong support (Supplementary Fig. 3). Within Locustellidae, the BI trees based on concatenation of all loci vary among analyses in topology and support for certain clades, depending on partitioning of data, and inclusion or exclusion of species with missing data (Table 1). For example, the BIC including all species (also those with incomplete data) in five partitions results in a tree with all except two nodes bifurcating, and all of these except nine having PP $\geq$0.95 (Fig. 2), whereas the unpartitioned BIC including only species with complete data results in a tree with all except one node bifurcating with PP $\geq$0.95 (Fig. 3). Bayes factor comparisons of the partitioned and unpartitioned analyses (all else being equal) are shown in Table 1. Also the MLB and MPB trees vary depending on whether species with missing data are included or excluded (Figs. 2 and 3).

Of the 29 clades recovered in the unpartitioned BIC of species with complete data (Fig. 3), five are found in only one of the SLAs, six in two SLAs, seven in three SLAs, eight in four SLAs, and two in all five SLAs (cf. Fig. 3 and Supplementary Fig. 1). The clade comprising C1–C4, which is recovered in only one of the SLAs, has considerably higher posterior probability in the analysis of all loci than in the SLA (1.0 and 0.82, respectively). The inclusion of *Locustella lanceolata* in clade 5 receives PP >0.95, despite that it is not found in this clade in any of the SLAs, and is poorly supported in this clade by MLB and MPB.

3.4. Species tree analyses

The BEST analyses of the complete dataset (Fig. 3) had convergence problems, despite a large number (up to 1×10^9) of iterations. The analysis with the highest likelihood values reached a plateau after c. 7×10^8 generations, but then dipped again after c. 9×10^8 generations, so it is uncertain if it ever reached its target distribution. The other analysis of the complete data was still rising in likelihood values at the end of the analysis. Some of the eight shorter analyses (c. 7×10^7–1×10^8 replicates) appeared to reach stationarity, based on their likelihood plots, but all had significantly lower likelihoods than the best 1×10^9 run; the latter is strongly favoured over the shorter run with the highest likelihood (ln Bayes factor: 115; cf. Kass and Raftery, 1995). BEST analyses of subsets of species, corresponding to clades B, C and D in the BIC (Fig. 3), appeared to converge considerably faster. In general, the results of these analyses are similar to those of the BEST of all species. However, there are some pronounced differences between the two types of BEST analyses (most notably at the nodes indicated by 3, 4, 5 and 10 in Fig. 3).

Notwithstanding the discrepancies between the different BEST analyses, the BEST phylogeny conforms rather well in topology with the BIC trees, although the support values are lower on average in the BEST than in the BIC (Fig. 3). However, in a few cases either of the two types of BEST analyses fails to recover a clade found with strong support in the BIC (indicated by 3, 4, 5, 6 and 10 in Fig. 3). One of these clades (5) is not recovered in any SLA, two (4, 10) are found with PP <0.95 in single SLAs, and two (3, 6) are recovered in two SLAs (one or both with PP $\geq$0.95). All of these cases concern nodes in which the two types of BEST analyses disagree (except node 6, which was only analysed in the complete dataset).

3.5. Major clades in Locustellidae

The following refers to the tree based on the BIC of all loci and all species (Fig. 2) and the species tree based on fewer species (Fig. 3). In the BIC tree, Locustellidae is separated into two major, well supported clades (A and B), whereas the BEST analysis of the complete dataset does not recover clade A (no separate BEST analysis of clade A was done). Clade A comprises all of the *Locustella* (Eurasia: Palearctic), Asian (Oriental) *Bradypterus* and *Megalurus pryeri* (Asia: Palearctic), while clade B contains the African *Bradypterus*, the monotypic Malagasy genus *Dromaeocercus*, the African *Schoenicola brevirostris*, the two Australian species of *Cincloramphus*, the monotypic Australian genus *Eremiornis* and four species of *Megalurus* (south Asia to Australasia). Clade A is further divided into clades C and D, which are both strongly supported in all analyses. The former includes a mix of *Locustella* and *Bradypterus*

Table 1
Comparison of Bayesian inference (MrBayes) of all five loci concatenated. "All species" refers to analyses including also taxa for which one or more loci is missing (*L. fasciolata amnicola*, *B. sylvaticus*, *B. barratti*, *S. brevirostris* and *M. punctatus*) and *B. baboecala* Nig (=Nigeria). "Species with complete data" refers to analyses excluding these taxa. "Partitioned" refers to analyses in five partitions (four nuclear loci GTR+Γ, cytochrome *b* GTR+Γ+I); "unpartitioned" refers to analyses of unpartitioned data (GTR+Γ+I). Bayes factors for comparisons are given at bottom of table.

Clade	All species		Species with complete data	
	Partitioned	Unpartitioned	Partitioned	Unpartitioned
(*B. mandelli*, *B. montis*)	0.59	1.0	0.57	1.0
(*B. thoracicus*, *B. kashmirensis*)	0.99	1.0	0.99	1.0
(C1, C2)	0.92	1.0	0.91	1.0
E	0.90	0.97	1.0[b]	1.0[b]
E1	0.57	0.67	1.0[c]	1.0[c]
E1a	0.74	0.81	1.0[a]	1.0[a]
E2a	0.70	0.85	1.0	1.0
(*B. cinnamomeus*, *B. lopezi*)	0.96	0.98	1.0	1.0
(*S. brevirostris*, E)	0.94	≤0.50	n.a.[d]	n.a.[d]
(*M. palustris*, (*S. brevirostris*, E))	0.97	0.70	1.0[d]	1.0[d]
F	0.96	0.80	1.0[e]	1.0[e]
	ln Bayes factors = 451 in favour of partitioned		ln Bayes factors = 211 in favour of unpartitioned	

[a] *B. baboecala* Nig (=Nigeria, subspecies *centralis*) not included.
[b] *B. baboecala* Nig, *B. sylvaticus* and *B. barratti* not included.
[c] *B. baboecala* Nig and *B. sylvaticus* not included.
[d] *S. brevirostris* not included.
[e] *M. punctatus* not included.

nested among each other, while the latter contains five *Locustella* and, nested among them, *M. pryeri*.

Clade B is divided into clades E and F, with *S. brevirostris* and *M. palustris* as sequential sisters to clade E. The positions of *S. brevirostris* and *M. palustris* vary among analyses, and due to missing data, the former was excluded from the species tree analyses. Clade E is well supported and comprises the African *Bradypterus* and, nested among them, *Dromaeocercus brunneus*. The support for clade F varies among analyses (none to strong) and is only recovered in one of the SLAs.

3.6. Minor clades in Locustellidae

The following refers to the BIC tree of all loci and all species (Fig. 2) and the species tree based on fewer species (Fig. 3), unless otherwise noted. All of clades C1–C5 are well supported, except for the inclusion of *L. lanceolata* in clade C5 (see above). All of them, except C4, contain either *Locustella* or Asian *Bradypterus* species, while C4 includes both genera. As has already been noted, the topologies of clades C1 and C2 are incongruent between different SLAs. For *Bradypterus major* only cyt*b* is available (from a museum specimen from 1931; Appendix A), and accordingly, its position in clade C4 rests on this locus alone.

The relationships within clade D are strongly supported, although the sister relationship between *L. ochotensis* and *L. pleskei* is contradicted by the myo analysis. Clades E1 and E2, which comprise African *Bradypterus* and the Malagasy *Dromaeocercus*, are both well supported in the analyses of the species for which all loci are available. However, inclusion of two species for which only cyt*b* sequences are available (*Bradypterus sylvaticus* and *Bradypterus barratti*) markedly reduces the support for these clades (Fig. 2).

In a BIC analysis of cyt*b* including one sequence per species, *M. palustris* is sister to *Bradypterus graueri* in clade E with poor support (PP 0.66; not shown) and with a branch 2.5 times as long as any other branch in clade E. In contrast, inclusion of two different subspecies of *M. palustris* place these two in a sister clade to clade E (Supplementary Fig. 1a), as is also the case in the cyt*b* tree containing multiple haplotypes (Fig. 1).

Clades F1 and F2 are both strongly supported in all analyses, also when *Megalurus punctatus*, for which only cyt*b* is available, is included (Figs 2 and 3). The former clade comprises one of the two *Cincloramphus* and one *Megalurus* as sisters and the other *Cincloramphus* as sister to these, while the second clade contains two *Megalurus* as sisters and the monotypic genus *Eremiornis* as sister to these.

3.7. Indels

Several clades are supported by apparently synapomorphic indels in the alignments of the nuclear loci (Fig. 3, Supplementary Figs. 2 and 3). Within Locustellidae, two clades have two indels each, and another eight clades have one indel each. All of the clades supported by indels have PP 1.0 in the BIC of all loci.

3.8. Intra- and interspecific cytochrome b divergences

The variation in cyt*b*, as indicated by branch lengths (Fig. 1) and GTR+Γ+I corrected distances, is generally low within the species for which we have multiple samples, including the polytypic species (0.0–1.2%, mean 0.5% ± 0.4; $n = 33$ pairwise comparisons, excluding the cases below) compared to the taxa treated as different species (4.0–29.4%, mean 19.3% ± 4.4; $n = 1435$ pairwise comparisons, excluding the cases below). However, there are a few cases of unexpectedly large intraspecific divergences (Table 2). In contrast, two pairs of taxa that are usually treated as separate species differ comparatively slightly (Table 2).

Table 2
Large intraspecific or small interspecific cytochrome *b* divergences (GTR+Γ+I corrected, 982 base pairs). Mean and standard deviation in parentheses; numbers given in parentheses in left column.

Taxa	GTR+Γ+I corrected distances (%)
Intraspecific divergences	
B. baboecala tongensis/transvaalensis (South Africa; 1) vs. *B. b. centralis* (Nigeria; 1) and *B. b. elgonensis* (Kenya; 1)[a]	10.2 – 10.9 (10.6 ± 0.5)
Locustella fasciolata fasciolata (1) vs. *L. f. amnicola* (1[b])	5.0
B. lopezi mariae (Tanzania; 1) vs. *B. l. ufipae* (Zambia; 1)	3.6
M. palustris toklao (India; 1) vs. *M. p. forbesi* (Philippines; 1)[c]	6.5
B. luteoventris[d] W Myanmar (1) vs. Thailand (1) and Vietnam (1)	2.7 (±0)
E. carteri[d] from same locality (<2.5 km apart; 2)	2.0[e]
Interspecific divergences	
B. mandelli[d] (2) vs. *B. montis*[d] (2)	1.3–1.7 (1.5 ± 0.2)
L. pleskei[d] (3) vs. *L. ochotensis*[d] (2)	0.5–0.8 (0.7 ± 0.2)

[a] Non-sisters.
[b] Sequence from GenBank (Y15689) suspected of being a nuclear copy not included.
[c] Based on 898 base pairs.
[d] Monotypic.
[e] Based on 714 base pairs.

4. Discussion

4.1. Comparison of methods

In the present study there is comparatively little incongruence between different SLAs, with only four nodes in the Locustellidae clade being strongly incongruent. Accordingly, as expected, there is little difference between the trees reconstructed via species tree approaches and concatenation, and no signs of the former receiving additional signal from the likelihood function of gene trees given a species tree (cf. Edwards et al., 2007; Brumfield et al., 2008; Liu et al., 2008; Liu and Edwards, 2009; Edwards, 2009). Although several nodes have low statistical support in the BEST trees, with few exceptions they nevertheless recover the same topology as the trees inferred via concatenation. In spite of the slight differences between the species tree and gene tree approaches, we consider the former to be a step forward compared to the latter, since it accounts for the ubiquitous heterogeneity in gene trees, thereby providing more realistic support than concatenation for nodes with incongruence among loci or instances where all or most of the signal in a multilocus analysis comes from a single locus.

In the trees inferred from the BEST of the complete dataset, 92% of the nodes have PP ≥0.95 (mean PP 0.98; nodes with ≤0.5/50% support excluded), whereas BIC recover all nodes with PP >0.95 (mean 1.0). Accordingly, our results confirm the prediction that statistical confidence is generally lower in species trees than in trees estimated via concatenation (Edwards, 2009), as has also been found in other empirical studies employing BEST (e.g. Belfiore et al., 2008; Thomson et al., 2008). Although we found BEST to yield PP ≥0.95 for nodes that were only strongly supported in one SLA, at least three independent SLAs with PP ≥0.50 or two with PP ≥0.95 for a certain node were required for BEST to consistently attain PP ≥0.95 for that node (even if strongly contradicted by one other SLA). Similar results were obtained by Edwards et al. (2007) based on simulated data. However, Brumfield et al. (2008) and Liu et al. (2008) reported BEST inferring a species tree that was corroborated by independent data despite that this was not found by any of five SLAs. In spite of the increase in phylogenetic signal in species tree analyses compared to concatenation (e.g. Edwards et al., 2007; Brumfield et al., 2008; Eckert and Carstens, 2008; Liu and

Edwards, 2009; Edwards, 2009), it seems advisable to treat clades that are found in only one (or even no) SLA with caution, even when these have high BEST support. In the present study, this concerns nodes 4 and 10 in Fig. 3, although in these cases the BEST support varies between the complete and incomplete datasets (see below).

A few species tree reconstructions need to be commented on. The nodes marked by 3 and 4 in Fig. 3 are not recovered by the BEST of the complete dataset, whereas both these clades are inferred in one or two SLAs (and not contradicted by any other SLA), in all analyses of concatenated data, and in the BEST of clade C on its own. Moreover, the most basal node in clade A (indicated by 6 in Fig. 3), which is well supported by two SLAs, one indel, and all analyses of concatenated data is not recovered by BEST (only complete dataset analysed). In contrast, BEST places clade D as sister to the rest of the ingroup, with low support. With respect to clade A, the topology inferred by the majority of the analyses seems more probable based on morphology and biogeography (cf. Bairlein et al., 2006). In all these cases, it seems possible (even likely) that the BEST analyses of the complete data have not reached their target distributions (see Section 4.2, below).

The BIC analyses result in varying support depending on whether the data are partitioned or not. It is not evident which analysis is better. With respect to the analyses of all species, Bayes factor comparisons strongly favour the partitioned analysis over the unpartitioned one, whereas the opposite is true in the analyses including only species with complete data (all else being equal). It could be argued, however, that partitioned analyses are generally superior to unpartitioned analyses (e.g. Brown and Lemmon, 2007; McGuire et al., 2007), especially in cases where different loci have markedly different phylogentic signal; in the present study, cyt*b* is much more informative than the nuclear loci.

In the BIC analyses, inclusion of species for which only cyt*b* is available has varying effects in different parts of the tree. Inclusion of two such species (*B. sylvaticus*, *B. barratti*) negatively affects the support for several nodes in clade E, whereas inclusion of two other species with only cyt*b* data that are inferred to belong in clades C4 (*B. major*) and F2 (*M. punctatus*), respectively, do not appear to reduce the support for any neighbouring nodes (cf. Figs 2 and 3). These differences might be the result of the different lengths of these sequences: the sequences for *B. sylvaticus* and *B. barratti* are only 603 base pairs, whereas for *B. major* and *M. punctatus* they are 711 and 716 base pairs, respectively.

4.2. Convergence in BEST analyses

The BEST analyses of the complete dataset obviously had convergence problems, despite the large number of iterations. Even the longest run with the highest likelihood was fluctuating markedly near the end, while the other 1×10^9 run was still climbing when it terminated. The shorter BEST results from the complete dataset all had lower likelihood values, and therefore appeared not to have reached their target distributions, despite some having apparently spuriously stationary likelihood values. The differences in topology and support between the BEST of the complete dataset and the separate analyses of clades B, C and D might be due to convergence problems, especially in the more extensive dataset.

Convergence problems for BEST have been reported in other empirical studies. In a BEST of 162 genes from five species of *Zea* maize, Cranston et al. (2009) failed to reach convergence in 1.6×10^9 iterations, and Linnen and Farrell (2008) reported lack of convergence in multiple 50×10^6 generation runs for a *Neodiprion* saw fly dataset. These and the present results suggest that BEST might need to be run exceedingly long to reach the proper target distribution. Our results also emphasize that it is advisable to do multiple analyses of the same dataset to ascertain that convergence has been reached. If other analyses suggest the presence of some well corroborated monophyletic subgroups, analysing these separately, as also tested here, is likely to help BEST converge more quickly. Cranston et al. (2009) suggest that it might be possible to increase the rate of convergence by exploring MCMC parameters, using different proposal mechanisms, or perhaps by inferring starting parameters for the individual genes before beginning the joint analysis. An alternative solution might be to vary the population size (θ) prior. This has proven helpful in a study of *Sceloporus* fence lizards, in which only analyses with higher θ values ($\geq$0.015) converged (Leaché, 2009).

4.3. Phylogeny of Locustellidae

The phylogenetic hypothesis in Fig. 3 is mostly well supported by the data, although resolution of some internal nodes is uncertain. Clade A, which contains the Asian *Bradypterus*, all *Locustella*, and *M. pryeri*, is moderately or strongly supported by all analyses except BEST (only complete dataset analysed), and is further supported by one indel, albeit only by two SLAs. Clade B, which includes the African *Bradypterus*, three *Megalurus*, *Dromaeocercus*, *Cincloramphus* and *Eremiornis*, and according to the tree in Fig. 2 also *Schoenicola*, is well supported in all analyses, and receives additional support from one indel, although it is only recovered in two SLAs. These two clades make sense from a biogeographical perspective, as all of the species in clade A breed in the Palearctic or Oriental regions, whereas the species in clade B are Afrotropical/Malagasy (clade E) or Oriental to Australasian (clade F and *M. palustris*) (Bairlein et al., 2006).

Clades C, D, E, F1 and F2 are unanimously well corroborated by the data. The support for the sister relationship between F1 and F2 rests mainly or exclusively on LDH and is lacking in one of the BEST analyses. From a biogeographical and morphological perspective (Bairlein et al., 2006), this is a sensible group (but see comment on *M. palustris*, below). Clades C1–C5 are robust, except for the inclusion of *L. lanceolata* in C5 (see below). However, the relationships among these are uncertain. The relative positions of C1, C2 and C3 vary among analyses. Even the inclusion of these in the same clade is not unanimously well supported, and relies exclusively or mainly on cyt*b*. Also the position of clade C4 rests only or mainly on cyt*b*, and disagrees among different analyses.

B. major is placed in clade C4, as sister to *Locustella naevia*, with good support, in the analyses of the concatenated sequences. However, this should be considered provisional, as it is based on cyt*b* only. Similarly, the inclusion of *L. lanceolata* in clade C5 is tentative. This clade receives high BIC and MLB support, but weak or no statistical support in the species tree or MPB analyses, and is not recovered in any SLA. Also for *B. sylvaticus* and *B. barratti* only cyt*b* is available, and the precise positions within clade E are indeterminate. Regarding the latter, Bairlein et al. (2006) point out that based on similarities in morphology and habitat choice it forms a group with *Bradypterus cinnamomeus*, *Bradypterus lopezi* and *Bradypterus bangwaensis*, in agreement with our results.

The sequences for *S. brevirostris* were acquired from a museum specimen, and no cyt*b* or LDH data were obtained. Due to the missing data, this species was excluded from the species tree analyses. However, BIC places this species as sister to clade E, although with insufficient statistical support, and this position seems reasonable from a biogeographical and morphological perspective (Bairlein et al., 2006). This is further supported by analyses of the nuclear RAG-1 and RAG-2, which place *S. brevirostris* with strong support as sister to a clade with *B. barratti* and *Bradypterus baboecala* (Beresford et al., 2005).

The position of *M. palustris* as sister to *S. brevirostris* plus clade E receives mostly strong support in the different analyses of all loci combined, although this is only inferred by one SLA (cyt*b*). However, this is contradicted by analyses of RAG-1 and RAG-2, according to which *M. palustris* and *C. mathewsi* form a strongly supported clade, which is sister to a clade containing *Schoenicola* and two African *Bradypterus* (Beresford et al., 2005). The tree inferred by the present study is surprising from a morphological and vocal point of view. *M. palustris* resembles the other species of *Megalurus* (clade F) morphologically, whereas it differs in multiple aspects from *S. brevirostris* and the species in clade E (Bairlein et al., 2006). Moreover, the song of *M. palustris* is said to be similar to at least the Philippine populations of *M. timoriensis*, whereas it differs more from *S. brevirostris* and the African *Bradypterus* (Bairlein et al., 2006). In addition, *M. palustris* and the species in clade F are collectively distributed from the Indian Subcontinent via the Philippines and Indonesia to Australia and New Zealand, whereas *S. brevirostris* and the species in clade E occur in the Afrotropics (though the second species of *Schoenicola*, *S. platyurus*, is found in south India). If the position of *M. palustris* inferred here is indeed correct, this implies that the morphological evolution set off in a new direction in the lineage leading to *S. brevirostris*/clade E after these split from a most recent common ancestor with *M. palustris*, whereas the morphological divergence was much more conservative in the lineages leading to *M. palustris* and clade F.

The indels in the nuclear alignments lend further support to the inferred tree. All except one of the nodes with corroborating indel data is unanimously well supported by the different analyses. The exception concerns the most basal node in clade A, which has conflicting inferences.

Drovetski et al. (2004) used mitochondrial ND2 to study the relationships of all *Locustella*, two Asian and three African *Bradypterus* (*B. castaneus*, *B. tacsanowskius*, *B. baboecala*, *B. cinnamomeus*, *B. mariae* [=*B. lopezi mariae*]), and two *Megalurus* (*M. gramineus*, *M. pryeri*). In agreement with our results, they found that Asian *Bradypterus* and *M. pryeri* nested within *Locustella*, and African *Bradypterus* formed a separate clade, as did *M. gramineus*. The relationships within these clades conform with our cyt*b* tree.

4.4. Taxonomic implications

According to our data, the phylogeny strongly disagrees with the current taxonomy at the generic level. We propose a number of taxonomic changes (Fig. 2): (1) that the Asian species of *Bradypterus* and *M. pryeri* be placed in *Locustella*; (2) that *Bradypterus* is restricted to the species in clade E (which includes the type species of this genus, *B. baboecala*), which means that the monotypic genus *Dromaeocercus* is synonymized with *Bradypterus*; (3) that *Schoenicola* is provisionally retained, pending further studies of its affinities based on additional loci and inclusion of the other species of *Schoenicola* (*S. platyurus*, south India) and the two missing African *Bradypterus*; and (4) that *Cincloramphus* and *Eremiornis* are synonymized with *Megalurus*. The last point renders *Megalurus* non-monophyletic, since *M. palustris* is retained in *Megalurus*, and accordingly runs counter to modern taxonomic practice. However, this is a preliminary standpoint, which takes into account the phylogenetic uncertainty with respect to *M. palustris* (conflict between our data, which rest mainly or entirely on cyt*b*, on the one hand, and RAG sequence data, morphology and vocalizations, on the other hand; see above). The alternatives, to treat all of clade B as *Megalurus* (by priority) or to recognize a monotypic *Megalurus* for *palustris* (which is the type species for this genus) and referring to clade F as *Cincloramphus* (by priority), are less appealing at this stage. More data are needed to determine the position of *M. palustris* before this issue can be satisfactorily resolved.

We lack samples of the genera *Amphilais* (monotypic, Madagascar), *Megalurulus* (four species, Melanesia), *Buettikoferella* (monotypic, Timor), and *Chaetornis* (monotypic, Indian Subcontinent), which have been suggested to be closely related to *Megalurus*, and *Elaphrornis* (monotypic, Sri Lanka), which is usually placed in *Bradypterus* (e.g. Bairlein et al., 2006). Future studies will show whether these taxa are part of Locustellidae or not, though in any event they are unlikely to affect the taxonomic changes proposed here.

The taxon *pryeri* has already been suggested to belong in *Locustella* based on morphology (Morioka and Shigeta, 1993), and, as pointed out above, this has been confirmed by mitochondrial ND2 by Drovetski et al. (2004). However, the suggestion by Bairlein et al. (2006) that its two subspecies *pryeri* and *sinensis* might deserve species rank is not corroborated by the present study. The use of the generic name *Bowdleria* for *M. punctatus*, which has been advocated based on osteological characters (Olson, 1990), is not supported by our data.

Some taxa that are currently treated as conspecific appear to be sufficiently divergent (cf. e.g. Hebert et al., 2004; Lovette and Bermingham, 1999; Olsson et al., 2005) in cyt*b* to warrant species status, although this needs to be confirmed by independent data: *L. fasciolata fasciolata* vs. *L. f. amnicola* (also remarked by Drovetski et al., 2004); *B. baboecala tongensis*/*B. b. transvaalensis* vs. *B. b. centralis*/*B. b. elgonensis*; *B. lopezi mariae*/*B. l. usambarae* vs. *B. l. ufipae*; and *M. palustris toklao* vs. *M. p. forbesi*. Also the monotypic *B. luteoventris* needs further study in the light of our results. The two samples of *E. carteri* are surprisingly divergent considering that they are from the same locality, and this needs to be investigated. Conversely, the divergences between the two species pairs *L. pleskei*–*L. ochotensis* and *B. mandelli*–*B. montis* are so slight that their status as separate species need to be studied further. Slight differences between the former pair have previously been found in ND2 (Drovetski et al., 2004), and they have been treated as conspecific (e.g. Williamson, 1968). The two latter have been treated as conspecific (e.g. Watson et al., 1986), but were split based on minor differences in morphology and song (Dickinson et al., 2000).

4.5. Dangers of morphology-based classifications

The present study underscores the well known but still often neglected problem of basing classifications on non-cladistic analyses of morphological characters. The traditional classification of these birds (e.g. Watson et al., 1896; Sibley and Monroe, 1990; Dickinson, 2003; Bairlein et al., 2006) is at variance with the phylogeny presented here regarding every single genus except *Schoenicola* (which comprises only two species, of which only one is included here). These discrepancies result from multiple cases of morphological convergence, e.g. African and Asian "*Bradypterus*", as well as several instances of strongly divergent morphological evolution, e.g. "*Dromaeocercus*", "*Eremiornis*" and "*Cincloramphus*". With regard to the latter genus, Bairlein et al. (2006) state that "striking morphological differences...suggest that the two species may not be closely related and should perhaps be placed in separate genera". The divergence between these is confirmed here, although the phylogeny implies an even more complex pattern of morphological differentiation. Molecular data have previously shown that *Graminicola bengalensis* is not closely related to Locustellidae (Alström et al., 2006; Gelang et al., 2009).

5. Conclusions

For our data, the traditional gene tree methods (Bayesian inference, maximum likelihood, parsimony) and a species tree approach (BEST) yield basically the same topology. In spite of this, we

Taxon	Locality	Sample No./Ref.	GenBank No.				
			Cyt*b*	ODC	Myo	LDH	GAPDH
Abroscopus albogularis fulvifacies	Sichuan, China	DZUG U1932	*HQ706175*	***HQ706303***	***HQ706226***	***HQ706186***	***HQ706264***
Acrocephalus arundinaceus arundinaceus	Austria		*FJ883022*	*FJ883128*	*FJ883098*	*FJ883056*	–
Acrocephalus arundinaceus zarudnyi	Xinjiang, China	NRM 20046787	–	–	–	–	***HQ706300***
Aegithalos caudatus caudatus	Sweden	NRM 976089	*AY228044*	*EU680703*	*AY228281*	***HQ706183***	*FJ357912*
Alauda arvensis arvensis	Sweden	NRM 966614	*AY228047*	*EF625336*	*AY228284*	*HQ333047*	*FJ357913*
Alophoixus pallidus annamensis/khmerensis	C Vietnam	NRM 20046822	*DQ008507*	***HQ706304***	*DQ008559*	–	–
Apalis flavida florisuga	KwaZulu-Natal, South Africa	DZUG U2204; VH B0745 (LDH)	*HQ333036*	*HQ333083*	*HQ333069*	*HQ333049*	*HQ333097*
Bernieria madagascariensis inceleber	Toliara, Madagascar	FMNH 431202	*HQ333038*	*HQ333086*	*HQ333071*	*HQ333052*	*HQ333100*
Bradypterus alishanensis	Taiwan	DZUG U1934	***HQ706133***	***HQ706310***	***HQ706232***	***HQ706192***	***HQ706272***
Bradypterus alishanensis	Taiwan	DZUG U1933	**HQ706132**	–	–	–	–
Bradypterus baboecola tongensis/transvaalensis	Natal, South Africa	NRM 20046782	*DQ008473*	*HQ333084*	*DQ008525*	*HQ333050*	*HQ333098*
Bradypterus baboecala transvaalensis	Gauteng, South Africa	Paulette Bloomer in litt.	AY958221	–	–	–	–
Bradypterus baboecala transvaalensis	Zimbabwe	Paulette Bloomer in litt.	AY958222	–	–	–	–
Bradypterus baboecala tongensis	Malawi	Paulette Bloomer in litt.	958223	–	–	–	–
Bradypterus baboecala tongensis	Zambia	Paulette Bloomer in litt.	958224	–	–	–	–
Bradypterus baboecala centralis	Nigeria	DZUG U1935	***HQ706159***	***HQ706338***	***HQ706259***	***HQ706222***	–
Bradypterus baboecala elgonensis	SW Kenya	VH A0769[a]	FJ883053	–	–	–	–
Bradypterus bangwaensis	Nigeria	DZUG U1025	***HQ706163***	***HQ706330***	***HQ706251***	***HQ706214***	***HQ706292***
Bradypterus carpalis	SW Kenya	VH A0768	***HQ706162***	***HQ706329***	***HQ706250***	***HQ706213***	***HQ706291***
Bradypterus carpalis	Kenya	Paulette Bloomer, in litt.	AY958230	–	–	–	–
Bradypterus castaneus castaneus	S Sulawesi, Indonesia	NRM 20066006	*DQ367925*	***HQ706314***	***HQ706236***	***HQ706196***	***HQ706276***
Bradypterus caudatus unicolor	Mindanao, Philippines	FMNH 392283	***HQ706140***	***HQ706315***	***HQ706237***	***HQ706197***	***HQ706277***
Bradypterus cinnamomeus cinnamomeus	Tanga, Tanzania	ZMUC 121180	–	***HQ706331***	***HQ706252***	***HQ706215***	***HQ706293***
Bradypterus cinnamomeus mildbreadi	Uganda	ZMUC 123143	***HQ706166***	–	–	–	–
Bradypterus cinnamomeus mildbreadi	Uganda	FMNH 355750	AY124541	–	–	–	–
Bradypterus davidi davidi	Hebei, China (m)	NRM 20056595	*DQ367931*	***HQ706316***	***HQ706238***	***HQ706198***	***HQ706278***
Bradypterus davidi davidi	Hebei, China (m)	NRM 20056596	DQ367932	–	–	–	–
Bradypterus davidi davidi	Sichuan, China	ZMUC 117767	DQ367933	–	–	–	–
Bradypterus davidi davidi	Sichuan, China	ZMUC 117768	DQ367934	–	–	–	–
Bradypterus davidi davidi/suschkini	Hong Kong (m)	DZUG U398	**HQ706142**	–	–	–	–
Bradypterus davidi davidi/suschkini	Hong Kong (m)	DZUG U399	**HQ706141**	–	–	–	–
Bradypterus davidi davidi/suschkini	C Mongolia (m)	NRM 20056597	DQ367935	–	–	–	–
Bradypterus graueri	Uganda	DZUG U1937	***HQ706161***	***HQ706328***	***HQ706249***	***HQ706212***	***HQ706290***
Bradypterus kashmirensis	Himachal Pradesh, India	NRM 20056593	*DQ367926*	***HQ706317***	***HQ706239***	***HQ706199***	***HQ706279***
Bradypterus kashmirensis	Himachal Pradesh, India	NRM 20056594	DQ367927	–	–	–	–
Bradypterus lopezi ufipae	Zambia	DZUG U1938	***HQ706165***	***HQ706332***	***HQ706253***	***HQ706216***	***HQ706294***
Bradypterus lopezi mariae	Tanga, Tanzania	ZMUC 05391	**HQ706164**	–	–	–	–
Bradypterus lopezi usambarae	Namuli, Mozambique	Paulette Bloomer, in litt.	AY958227	–	–	–	–
Bradypterus luteoventris	NW Thailand (m)	DZUG U1946	***HQ706144***	***HQ706319***	***HQ706241***	***HQ706201***	***HQ706281***
Bradypterus luteoventris	Tonkin, Vietnam	DZUG U1945	**HQ706143**	–	–	–	–
Bradypterus luteoventris	W Myanmar	DZUG U1944	**HQ706145**	–	–	–	–
Bradypterus major [b]	Ladakh, India	FMNH 240009	**HQ706174**	–	–	–	–
Bradypterus mandelli mandelli	West Bengal, India	DZUG U1339	***HQ706135***	***HQ706311***	***HQ706233***	***HQ706193***	***HQ706273***
Bradypterus mandelli mandelli	NW Thailand	DZUG U1941	**HQ706134**	–	–	–	–
Bradypterus mandelli mandelli/melanorhynchus	Hong Kong (m)	DZUG U1942	**HQ706136**	–	–	–	–
Bradypterus montis	Java	DZUG U1940	***HQ706137***	***HQ706312***	***HQ706234***	***HQ706194***	***HQ706274***

Bradypterus montis	Bali	DZUG U1939	**HQ706138**	–	–	–	–
Bradypterus tacsanowskius	Mongolia	UWBM 57938	*HQ333037*	*HQ333085*	–	*HQ333051*	–
Bradypterus tacsanowskius	Irkutskaya Oblast, Russia	UWBM 51699	**HQ706146**	–	*HQ333070*	–	–
Bradypterus tacsanowskius	Hebei, China (m)	NRM 20046783	DQ008474	–	–	–	*HQ333099*
Bradypterus thoracicus	Sichuan, China	ZMUC 117765	*DQ367929*	***HQ706318***	***HQ706240***	***HQ706200***	***HQ706280***
Bradypterus thoracicus	Sichuan, China	NRM 20056582	DQ367930	–	–	–	–
Bradypterus thoracicus	Qinghai, China	NRM 20056583	DQ367928	–	–	–	–
Bradypterus sylvaticus sylvaticus	George, South Africa	Paulette Bloomer, in litt.	AY958228	–	–	–	–
Bradypterus sylvaticus pondoensis	East London, South Africa	Paulette Bloomer, in litt.	AY958229	–	–	–	–
Cettia cetti cetti	France	DZUG U1936	***HQ706176***	*HQ121555*	***HQ706225***	***HQ706185***	***HQ706263***
Cinclorhamphus cruralis	South Australia	MV B.38407	***HQ706167***	***HQ706334***	***HQ706255***	***HQ706217***	***HQ706296***
Cinclorhamphus mathewsi	Victoria, Australia	MV B.24688	**HQ706169**	–	***HQ706256***	–	–
Cinclorhamphus mathewsi	South Australia	MV B.20019	***HQ706168***	***HQ706335***	–	***HQ706218***	***HQ706297***
Crossleyia xanthophrys	Madagascar	FMNH 393280	***HQ706177***	***HQ706309***	***HQ706231***	***HQ706191***	***HQ706269***
Cryptosylvicola randrianasoloi	Madagascar	FMNH 363849	***HQ706178***	***HQ706308***	***HQ706230***	***HQ706190***	***HQ706268***
Delichon urbicum	Spain	NRM 20046816	*DQ008517*	*EU680721*	*DQ008568*	*HQ333055*	*HQ333103*
Donacobius atricapilla	Paraguay	NRM 966966	*DQ008481*	*EU680723*	*DQ008533*	*HQ333054*	*FJ357915*
Dromaeocercus brunneus	Madagascar	FMNH 384749	***HQ706160***	*EU680724*	*EU680593*	***HQ706211***	***HQ706289***
Eremiornis carteri	W Australia	MV B.24551	***HQ706171***	***HQ706337***	***HQ706258***	***HQ706220***	***HQ706299***
Eremiornis carteri	W Australia	MV B.24554	**HQ706172**	–	–	–	–
Hartertula flavoviridis	Madagascar	FMNH 438721	***HQ706131***	***HQ706307***	***HQ706229***	***HQ706189***	***HQ706267***
Hippolais olivetorum	Kenya	Fregin et al. (2008)	*FJ883048*	*FJ883155*	*FJ883121*	*FJ883080*	–
Hippolais olivetorum	Bulgaria	DZUG U1947	–	–	–	–	***HQ706270***
Hirundo rustica rustica	Sweden	NRM 976238	*DQ008516*	*EF441240*	*AY064258*	–	–
Hirundo rustica rustica	Germany		–	–	–	*HQ333056*	*EF441218*
Iduna similis	Kenya	ZMUC 131329	*FJ899738*	*FJ883159*	*FJ883125*	*FJ883083*	***HQ706271***
Leptopoecile sophiae obscura	Qinghai, China	NRM 20046817	*DQ008518*	*EU680738*	*DQ008569*	***HQ706184***	***HQ706262***
Locustella certhiola ssp.	Hebei, China (m)	NRM 20046785	*DQ008476*	–	*DQ008528*	–	***HQ706286***
Locustella certhiola ssp.	Ningxia, China	DZUG U1388	**HQ706154**	***HQ706325***	–	–	–
Locustella certhiola ssp.	Thailand (m)	DZUG U1284	**HQ706155**	–	–	–	–
Locustella certhiola ssp.	Alakol, Kazakhstan	VH B0756	–	–	–	***HQ706208***	–
Locustella fasciolata amnicola	Sakhalin, Russia	UWBM 47557	***HQ706150***	***HQ706322***	***HQ706244***	***HQ706205***	–
Locustella fasciolata amnicola	Japan	Bernd Leisler, in litt.	Y15689	–	–	–	–
Locustella fasciolata fasciolata	Hebei, China (m)	DZUG U1948	***HQ706151***	***HQ706323***	***HQ706245***	***HQ706206***	***HQ706284***
Locustella fluviatilis	Kenya (m)		AJ004764	–	–	–	–
Locustella fluviatilis	Uncertain	NRM 20026297	–	*HQ121556*	–	–	–
Locustella fluviatilis	Kenya (m)	NRM 20046784	*DQ008475*	–	*DQ008527*	***HQ706203***	*HQ121546*
Locustella lanceolata ssp.[c]	Hebei, China (m)	DZUG U1949	***HQ706139***	***HQ706313***	***HQ706235***	***HQ706195***	***HQ706275***
Locustella lanceolata ssp.	China (m?)	–	DQ119524	–	–	–	–
Locustella lanceolata ssp.	China (m?)	–	DQ119525	–	–	–	–
Locustella luscinioides luscinioides	Sweden	NRM 20056589	***HQ706149***	***HQ706321***	***HQ706243***	***HQ706204***	***HQ706283***
Locustella luscinioides luscinioides	Germany	–	AJ004763	–	–	–	–
Locustella luscinioides ssp.	Israel (m)	DZUG U1950	**HQ706148**	–	–	–	–
Locustella naevia naevia	Sweden	DZUG U1951	***HQ706147***	***HQ706320***	***HQ706242***	***HQ706202***	***HQ706282***
Locustella ochotensis	Philippines (m)	DZUG U1619	***HQ706156***	***HQ706326***	–	–	–
Locustella ochotensis	Philippines (m)	DZUG U1621	–	–	***HQ706247***	–	***HQ706287***
Locustella ochotensis	Usuria, Russia	VH A0694	–	–	–	***HQ706209***	–
Locustella ochotensis	Taiwan (m)	DZUG 2101	***HQ706157***	–	–	–	–

(*continued on next page*)

a (*continued*)

Taxon	Locality	Sample No./Ref.	GenBank No.				
			Cyt*b*	ODC	Myo	LDH	GAPDH
Locustella pleskei	Izu isl., Japan	DZUG U1953	–	***HQ706327***	***HQ706248***	***HQ706210***	***HQ706288***
Locustella pleskei	Izu isl., Japan	DZUG U1952	***HQ706158***	–	–	–	–
Locustella pleskei	Izu isl., Japan	Takema Saitoh, in litt.	AB159188	–	–	–	–
Locustella pleskei	Deogu-do isl., South Korea	Takema Saitoh, in litt.	AB159191	–	–	–	–
Megalurus gramineus goulburni	South Australia, Australia	ANWC D224	*HQ333042*	*HQ333091*	*HQ333074*	*HQ333060*	*HQ333108*
Megalurus palustris toklao	Punjab, India	NRM 20046786	*DQ008477*	*EU680741*	*DQ008529*	***HQ706221***	*FJ357917*
Megalurus palustris forbesi	Negros, Philippines	ZMUC 02031	FJ883052	FJ883161	–	FJ883089	–
Megalurus pryeri pryeri	Japan	DZUG U1954	–	***HQ706324***	***HQ706246***	***HQ706207***	***HQ706285***
Megalurus pryeri pryeri	Japan	DZUG U1955	***HQ706152***	–	–	–	–
Megalurus pryeri pryeri	Japan	Bernd Leisler, in litt.	AJ004321	–	–	–	–
Megalurus pryeri pryeri	Japan	Bernd Leisler, in litt.	AJ004322	–	–	–	–
Megalurus pryeri pryeri	Japan	Bernd Leisler, in litt.	AJ004323	–	–	–	–
Megalurus pryeri sinensis	China (status unknown)	–	EU016090	–	–	–	–
Megalurus pryeri sinensis	China (status unknown)	–	EU016091	–	–	–	–
Megalurus pryeri sinensis	Jiangxi, China (m)	DZUG U1956	**HQ706153**	–	–	–	–
Megalurus punctatus vealeae	New Zealand	AWMM B.10962	**HQ706173**	–	–	–	–
Megalurus timoriensis	Luzon, Philippines	ZMUC 119529	***HQ706170***	***HQ706336***	***HQ706257***	***HQ706219***	***HQ706298***
Melocichla mentalis mentalis	Nigeria	NRM 20046804	*DQ008500*	*HQ333090*	*DQ008551*	–	*HQ333107*
Melocichla mentalis mentalis	Ivory Coast	VH A1550	–	–	–	*HQ333059*	–
Mirafra javanica williamsoni	Thailand	NRM 20046819	*DQ008520*	*HQ333089*	*DQ008571*	*HQ333058*	*HQ333106*
Orthotomus sutorius inexpectatus	NW Thailand	NRM 20046795	*DQ008491*	*HQ333092*	*DQ008542*	–	*HQ333109*
Orthotomus sutorius guzuratus	NW India	VH A1581	–	–	–	*HQ333061*	–
Oxylabes madagascariensis	Madagascar	FMNH 438719	***HQ706179***	***HQ706306***	***HQ706228***	***HQ706188***	***HQ706266***
Phylloscopus sindianus lorenzii	NE Turkey	DZUG U1957	***HQ706180***	***HQ706340***	–	–	–
Phylloscopus sindianus lorenzii	NE Turkey	DZUG U1958	–	–	***HQ706261***	–	–
Phylloscopus sindianus lorenzii	Caucasus	VH B0799	–	–	–	***HQ706224***	–
Phylloscopus sindianus sindianus	Pakistan	DZUG U1959	–	–	–	–	***HQ706302***
Pycnonotus barbatus inornatus	Mauretania	DZUG U2047	*HQ333043*	*HQ333093*	*HQ333075*	*HQ333062*	*HQ333110*
Schoenicola brevirostris alexinae	Kenya	NRM 569624	–	***HQ706333***	***HQ706254***	–	***HQ706295***
Seicercus tephrocephalus	W Myanmar	DZUG U1960	***HQ706182***	***HQ706339***	***HQ706260***	–	***HQ706301***
Seicercus tephrocephalus	W Myanmar	DZUG U1961	–	–	–	***HQ706223***	–
Stachyris nigriceps yunnanensis/rileyi	Ha Tinh province, C Vietnam	NRM 20026627	*HQ333045*	*HQ333095*	–	*HQ333065*	*HQ333112*
Stachyris nigriceps yunnanensis	Tonkin, N Vietnam[d]	NRM 947308	–	–	*AY228321*	–	–
Sylvia atricapilla atricapilla	Sweden	NRM 976380	–	*EF441254*	*AY887727*	–	*EF441232*
Sylvia atricapilla atricapilla	Germany	–	*Z73494*	–	–	–	–
Sylvia atricapilla atricapilla	Germany	VH A0364	–	–	–	*HQ333067*	–
Thamnornis chloropetoides	Madagascar	FMNH 436448	*HQ333046*	*HQ333096*	*HQ333077*	*HQ333068*	*FJ357923*
Xanthomixis apperti	Madagascar	FMNH 427370	***HQ706181***	***HQ706305***	***HQ706227***	***HQ706187***	***HQ706265***

[a] Same sample as FJ883162.
[b] Sequence obtained from Trevor Price and Udayan Borthakur.
[c] On geographical grounds, most likely nominate subspecies, but samples collected during migration, so subspecies *hendersonii* cannot be eliminated.
[d] According to Peter Nilsson (in litt.) (not given in published paper).

consider the latter to be a step forward, since it accounts for the ubiquitous heterogeneity in gene trees, thereby providing more realistic support than concatenation for nodes with incongruence among loci. However, BEST is computationally intense, and convergence proved difficult to attain, even in extremely long runs (up to 1×10^9 generations). We suggest that if other analyses indicate the presence of some well corroborated monophyletic subgroups, analysing these separately is likely to help BEST converge more quickly.

The phylogeny strongly disagrees with the current taxonomy at the generic level. We propose a revised classification that recognizes four instead of seven genera. One of these (*Megalurus*) is actually non-monophyletic according to our data, but we stress that this classification is tentative and takes account of the phylogenetic uncertainty (i.e. conflict between our results, which in this case rest on only one locus, and previously published data based on another locus, as well as morphology and vocalizations). Analysis of multiple cyt*b* haplotypes, including several different subspecies of polytypic species, suggests several cases where taxonomic revision is warranted.

Acknowledgements

We are most grateful to the following institutions or persons for samples: Göran Frisk, Ulf Johansson and Peter Nilsson and the Swedish Museum of Natural History, Stockholm, Sweden; Jon Fjeldså and Jan Bolding Kristensen and the Zoological Museum of the University of Copenhagen, Copenhagen, Denmark; David Willard and the Field Museum of Natural History, Chicago, USA; Museum Victoria, Melbourne, Australia; Brian Gill and the Auckland War Memorial Museum, Auckland, New Zealand; Australian National Wildlife Collections (CSIRO); Sharon Birks and the University of Washington Burke Museum, Seattle, USA; Martin Haase and Vogelwarte Hiddensee, Zoological Institute and Museum, Ernst Moritz Arndt University of Greifswald, Greifswald, Germany; Staffan Andersson; Staffan Bensch; Geoff Carey and Paul Leader; Ross McGregor; Magnus Jäderblad; Ulf Ottosson; Bo Petersson; Patrik Rhönnstad; Phil Round; Yoshimitsu Shigeta; Johan Ställberg; Trevor Price; Mats Waern; and Anders Waldenström. P.A. is also most grateful to Cheng-te Yao and the Taiwan Endemic Species Research Institute for invaluable assistance in the field. Dario Zuccon kindly sequenced *S. brevirostris*; Trevor Price and Udayan Borthakur generously provided a cytochrome *b* sequence of *B. major*; and Paulette Bloomer, Bernd Leisler and Takema Saitoh kindly provided information on geographic origin of some cytochrome *b* sequences in GenBank. Scott Edwards is gratefully acknowledged for his most valuable discussions and suggestions on BEST analyses, and for running three of the shorter BEST analyses for us; Normand David for comments on the endings of the scientific names proposed here; Edward Dickinson and Lars Larsson for various information; Peter Kennerley for comments on the Locustellidae introduction; and two anonymous referees for constructive criticism. P.A. also gratefully acknowledges the Riksmusei Vänners Linnaeus award, which has allowed him to devote time to this study. The Swedish Research Council provided financial support (Grant No. 621-2006-3194 to U.O. and 621-2007-5280 to P.E.).

Appendix A

List of samples (in alphabetical order), with GenBank accession numbers. Taxonomy follows Dickinson (2003), except for splitting of *Bradypterus davidi* and *B. kashmirensis* from *B. thoracicus* (Alström et al., 2008), inclusion of the recently described *Bradypterus alishanensis* (Rasmussen et al., 2000), and *Iduna similis* being moved from genus *Chloropeta* (Fregin et al., 2009). ANWC = Australian National Wildlife Collection (CSIRO), Canberra, Australia; AWMM = Auckland War Memorial Museum, Aukland, New Zealand; DZUG = Department of Zoology, University of Gothenburg, Göteborg, Sweden; FMNH = Field Museum of Natural History, Chicago, USA; MNHN = Muséum National d'Histoire Naturelle, Paris, France; MV = Museum Victoria, Melbourne, Australia; NRM = Swedish Museum of Natural History, Stockholm, Sweden; UWBM = University of Washington Burke Museum, Seattle, USA; VH = Vogelwarte Hiddensee, Zoological Institute and Museum, Ernst Moritz Arndt University of Greifswald, Greifswald, Germany; ZMUC = Zoological Museum of the University of Copenhagen, Copenhagen, Denmark. m = Sample collected on migration or in winter quarters. Sequences that are new to this study are in bold, and sequences included in multilocus analyses are in italics.

Appendix B. Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2010.12.012.

References

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans Auto Control. 19, 716–723.
Allen, E.S., Omland, K.E., 2003. Novel intron phylogeny supports plumage convergence in orioles (*Icterus*). Auk 120, 961–969.
Alström, P., Ericson, P.G.P., Olsson, U., Sundberg, P., 2006. Phylogeny and classification of the avian superfamily Sylvioidea. Mol. Phylogenet. Evol. 38, 381–397.
Alström, P., Rasmussen, P.C., Olsson, U., Sundberg, P., 2008. Species delimitation based on multiple criteria: the Spotted Bush Warbler *Bradypterus thoracicus* complex (Aves, Megaluridae). Zool. J. Linn. Soc. 154, 291–307.
Bairlein, F., Alström, P., Aymí, R., Clement, P., Dyrcz, A., Gargallo, G., Hawkins, F., Madge, S., Pearson, D., Svensson, L., 2006. Family Sylviidae (warblers). In: del Hoyo, J., Elliott, A., Christie, D.A. (Eds.), Handbook of the Birds of the World, vol. 12. Lynx Edicions, Barcelona., pp. 492–709.
Belfiore, N.M., Liu, L., Moritz, C., 2008. Multilocus phylogenetics of a rapid radiation in the genus Thomomys (Rodentia: Geomyidae). Syst. Biol. 57, 294–310.
Beresford, P., Barker, F.K., Ryan, P.G., Crowe, T.M., 2005. African endemics span the tree of songbirds (Passeri): molecular systematics of several evolutionary 'enigmas'. Proc. Roy. Soc. Lond., B. 272, 849–858.
Bock, W.J., 1994. History and nomenclature of avian family-group names. Bull. Am. Mus. Nat. Hist. 222, 1–283.
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.
Brumfield, R.T., Liu, L., Lum, D.E., Edwards, S.V., 2008. Comparison of species tree methods for reconstructing the phylogeny of bearded manakins (Aves: Pipridae, *Manacus*) from multilocus sequence data. Syst. Biol. 57, 719–731.
Cranston, C., Hurwitz, B., Ware, D., Stein, L., Wing, R.A., 2009. Species trees from highly incongruent gene trees in rice. Syst. Biol. 58, 489–500.
Dickinson, E.C., 2003. The Howard and Moore Complete Checklist of the Birds of the World. Christopher Helm, London.
Dickinson, E.C., Rasmussen, P.C., Round, P.D., Rozendaal, F.G., 2000. Systematic notes on Asian birds. 1. A review of the russet bush-warbler *Bradypterus seebohmi* (Ogilvie-Grant, 1895). Zool. Verh. Leiden 331, 12–64.
Drovetski, S.V., Zink, R.M., Fadeev, I.V., Nesterov, E.V., Koblik, E.A., Red'kin, Y.A., Rohwer, S., 2004. Mitochondrial phylogeny of *Locustella* and related genera. J. Avian Biol. 35, 105–110.
Eckert, A.J., Carstens, B.C., 2008. Does gene flow destroy phylogenetic signal? The performance of three methods for estimating species phylogenies in the presence of gene flow. Mol. Phylogenet. Evol. 49, 832–842.
Edwards, S.V., 2009. Is a new and general theory of molecular systematics emerging? Evolution 63, 1–19.
Edwards, S.V., Liu, L., Pearl, D.K., 2007. High-resolution species trees without concatenation. Proc. Natl. Acad. Sci. USA 104, 5936–5941.
Fjeldså, J., Zuccon, D., Irestedt, M., Johansson, U.S., Ericson, P.G.P., 2003. *Sapayoa aenigma*: a new world representative of 'Old World suboscines'. Proc. Roy. Soc. Lond. B 270 (Suppl.), 238–241.
Fregin, S., Haase, M., Olsson, U., Alström, P., 2009. Multi-locus phylogeny of the family *Acrocephalidae* (Aves: Passeriformes) – the traditional taxonomy overthrown. Mol. Phylogenet. Evol. 52, 866–878.
Gelang, M., Cibois, A., Pasquet, E., Olsson, U., Alström, P., Ericson, P.G.P., 2009. Phylogeny of babblers (Aves, Passeriformes): major lineages, family limits and classification. Zool. Scr. 35, 225–236.
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.
Hebert, P.D.N., Stoeckle, M.Y., Zemlak, T.S., Francis, C.M., 2004. Identification of birds through DNA barcodes. PLoS Biol. 2, e312. doi:10.1371/journal.pbio.0020312.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17, 754–755.
Huelsenbeck, J.P., Ronquist, F., 2005. MRBAYES version 3.1.2. <http://mrbayes.scs.fsu.edu/download.php>.
Irestedt, M., Ohlson, J.I., Zuccon, D., Källersjö, M., Ericson, P.G.P., 2006. Nuclear DNA from old collections of avian study skins reveals the evolutionary history of the Old World suboscines (Aves, Passeriformes). Zool. Scr. 35, 567–580.
Johansson, U., Fjeldså, J., Bowie, R.C.K., 2008. Phylogenetic relationships within Passerida (Aves: Passeriformes): a review and a new molecular phylogeny based on three nuclear intron markers. Mol. Phylogenet. Evol. 48, 858–876.
Jobb, G., 2008. TREEFINDER (version of October 2008). <http://www.treefinder.de>.
Jobb, G., von Haeseler, A., Strimmer, K., 2004. TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics. BMC Evol. Biol. 4, 18. doi:10.1186/1471-2148-4-18.
Kass, R.E., Raftery, A.E., 1995. Bayes factors. J. Am. Stat. Assoc. 90, 773–795.
Lanave, C., Preparata, C., Saccone, C., Serio, G., 1984. A new method for calculating evolutionary substitution rates. J. Mol. Evol. 20, 86–93.
Leaché, A.D., 2009. Species tree discordance traces to phylogeographic clade boundaries in North American fence lizards (*Sceloporus*). Syst. Biol. 58, 547–559.
Linnen, C.R., Farrell, B.D., 2008. Comparison of methods for species-tree inference in the sawfly genus *Neodiprion* (Hymenoptera: Diprionidae). Syst. Biol. 57, 876–890.
Liu, L., 2008. BEST: Bayesian estimation of species trees under the coalescent model. Bioinformatics 24, 2542–2543.
Liu, L., Edwards, S.V., 2009. Phylogenetic analysis in the anomaly zone. Syst. Biol. 58, 452–460.
Liu, L., Pearl, D.K., 2007. Species trees from gene trees: reconstructing Bayesian posterior distributions of a species phylogeny using estimated gene tree distributions. Syst. Biol. 56, 504–514.
Liu, L., Pearl, D.K., Brumfield, R., Edwards, S.V., 2008. Estimating species trees using multiple-allele DNA sequence data. Evolution 62, 2080–2091.
Lovette, I.J., Bermingham, E., 1999. Explosive speciation in the New World *Dendroica* warblers. Proc. Roy. Soc. Lond. B 266, 1629–1636.
McGuire, J.A., Witt, C.C., Altshuler, D.L., Remsen Jr., J.V., 2007. Phylogenetic systematics and biogeography of hummingbirds: Bayesian and maximum likelihood analyses of partitioned data and selection of an appropriate partitioning strategy. Syst. Biol. 56, 837–856.
Morioka, H., Shigeta, Y., 1993. Generic allocation of the japanese Marsh Warbler *Megalurus pryeri* (Aves: Sylviidae). Bull. Nat. Sci. Mus. Tokyo A 19, 37–43 (in Japanese).
Newton, M.A., Raftery, A.E., 1994. Approximate Bayesian inference with the weighted likelihood bootstrap. J. Roy. Stat. Soc., B 56, 3–48.
Nylander, J.A.A., 2004. MRMODELTEST2 version 2.2. <http://www.abc.se/~nylander>.
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.
Olson, S.L., 1990. Osteology and systematics of the fernbirds (*Bowdleria*; Sylviidae). Notornis 37, 161–171.
Olsson, U., Alström, P., Ericson, P.G.P., Sundberg, P., 2005. Non-monophyletic taxa and cryptic species – evidence from a molecular phylogeny of leaf-warblers (Phylloscopus, Aves). Mol. Phylogenet. Evol. 36, 261–276.
Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.
Rambaut, A., Drummond, A.J., 2009. Tracer version 1.5.0. <http://beast.bio.ed.ac.uk>.
Rasmussen, P.C., Round, P.D., Dickinson, E.C., Lambert, F.G., 2000. A new bush-warbler (Sylviidae, Bradypterus) from Taiwan. Auk 117, 279–289.
Rodríguez, J., Oliver, L., Marín, A., Medina, R., 1990. The general stochastic model of nucleotide substitution. J. Theo. Biol. 142, 485–501.
Ronquist, F., Huelsenbeck, J.P., 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.
Sibley, C.G., Ahlquist, J.E., 1990. Phylogeny and Classification of Birds. A Study in Molecular Evolution. Yale University Press, New Haven and London.
Sibley, C.G., Monroe Jr., B.L., 1990. Distribution and Taxonomy of Birds of the World. Yale University Press, New Haven.
Sorensen, M.D., Quinn, T.W., 1998. Numts: a challenge for avian systematics and population biology. Auk 115, 214–221.
Swofford, D.L., 2002. PAUP*: phylogenetic analysis using parsimony (*and other methods) version 4.0b10. Sinauer Associates, Sunderland, MA.
Tavaré, S., 1986. Some probabilistic and statistical problems on the analysis of DNA sequences. Lect. Mater. Life Sci. 17, 57–86.
Thomson, R.C., Shedlock, A.M., Edwards, S.V., Shaffer, H.B., 2008. Developing markers for multilocus phylogenetics in non-model organisms: a test case with turtles. Mol. Phylogenet. Evol. 49, 514–525.
Watson, G.E., Traylor Jr., M.A., Mayr, E., 1986. In: Mayr, E., Cottrell, G.W. (Eds.), Checklist of Birds of the World, vol. 11. Museum of Comparative Zoology, Cambridge, Massachusetts.
Williamson, K., 1968. Identification for Ringers. The Genera *Cettia*, *Locustella*, *Acrocephalus* and *Hippolais*, third ed. British Trust for Ornithology, Tring.
Yang, Z., 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. J. Mol. Evol. 39, 306–314.