The division of the major songbird radiation into Passerida and 'core Corvoidea' (Aves: Passeriformes) — the species tree vs. gene trees

MARTIN IRESTEDT & JAN I. OHLSON

Submitted: 19 July 2007
Accepted: 26 November 2007
doi:10.1111/j.1463-6409.2007.00321.x

Irestedt, M. & Ohlson, J. I. (2008). The division of the major songbird radiation into Passerida and 'core Corvoidea' (Aves: Passeriformes) — the species tree vs. gene trees. — *Zoologica Scripta*, *37*, 305–313.

The knowledge of evolutionary relationships among oscine songbirds has been largely improved in recent years by molecular phylogenetic studies. However, current knowledge is still largely based on sequence data from a limited number of loci. In this study, we re-evaluate relationships among basal lineages within the 'core Corvoidea' and Passerida radiations, by adding additional loci to previously published data. The trees obtained from the individual genes suggest incongruent topologies. Especially the positions of Callaeatidae (wattlebirds), Cnemophilidae (satinbirds) and Melanocharitidae (longbills and berrypeckers) vary among the trees, but RAG-1 is the only gene that unambiguously suggested a 'core Corvoidea' affinity for these taxa. Analyses of various combined data sets show that the phylogenetic positions for Callaeatidae, Cnemophilidae and Melanocharitidae largely depend on which genes that have been combined. As the RAG-1 gene has contributed to a majority of the phylogenetic information in previous studies, it has deeply influenced previous molecular affinities of these taxa. Based on the current data, we found a reasonable support for a Passerida affinity of Callaeatidae and Cnemophilidae, contrary to previous molecular studies. The position of Melanocharitidae is more unstable but a basal position among Passerida is congruent with a deletion observed in the glyceraldehyde-3-phosphodehydrogenase (GAPDH) loci. Molecular clock estimations conducted on the combined data sets were generally found to be similar, but for some divergences significant differences were found. These results illustrate the potential problem of phylogenies predominantly based on characters from one or a few loci, and exemplify the importance of well-supported phylogenies before reasonable time estimates of passerine divergences could be achieved.

Corresponding author: *Martin Irestedt*, *Molecular Systematics Laboratory*, *Swedish Museum of Natural History*, *P.O. Box 50007*, *SE–104 05 Stockholm*, *Sweden. E-mail: martin.irestedt@nrm.se*
Jan I. Ohlson, *Department of Vertebrate Zoology*, *Swedish Museum of Natural History*, *P.O. Box 50007*, *SE–104 05 Stockholm*, *Sweden. E-mail: jan.ohlson@nrm.se*

Introduction

Nearly half of the extant bird species in the world belong to the oscine (songbird) lineage which constitutes the majority of perching birds (Passeriformes), the linage is thus by far the most abundant avian radiation (Monroe & Sibley 1993). It includes many familiar birds; some of which are among the most well-studied birds in avian science. A robust phylogeny of oscines has therefore been highly desirable as a framework for various comparative studies and to better understand the biogeographical history of songbirds. Phylogenetic evidence from morphological studies (e.g. Beecher 1953; Ames 1971; Raikow 1978) has been limited, largely because convergent evolution to similar feeding specializations is commonly found in perching birds. During the 1980s, Charles G. Sibley and coworkers made the first serious molecular attempt, by using DNA–DNA hybridization data, to resolve the passerine tree and other avian relationships (Sibley & Ahlquist 1990). One main conclusion was that oscine birds could be divided in two reciprocally monophyletic clades: Corvida and Passerida. However, the avian DNA–DNA hybridization network by Sibley and coworkers has been heavily criticized on methodological grounds (Cracraft 1987; Houde 1987; Harshman 1994). Subsequent independent molecular studies (Barker *et al.* 2002, 2004; Ericson *et al.* 2002a,b) suggest that

Corvida *sensu* Sibley *et al*. is polyphyletic, as several deep lineages confined to the Australian region (e.g. Menuridae, Ptilonorhynchidae, Meliphagidae, Pomatostomidae and Orthonychidae) are subsequent sisters to all other oscines. The monophyly of Passerida on the other hand seems well supported as a lineage nested within Corvida *sensu* Sibley *et al*. (Barker *et al*. 2002, 2004; Ericson *et al*. 2002a).

Barker *et al*. (2002, 2004) were the first to publish a phylogeny of passerine birds based on DNA sequence data, with an almost complete family representation. Their results suggest that a majority of traditional corvidan birds form a clade that is the sister to Passerida (their 'core Corvoidea'). Furthermore, Barker *et al*. (2004) found that Callaeatidae (wattlebirds) in New Zeeland, Cnemophilidae (satinbirds) and Melanocharitidae (longbills and berrypeckers) in New Guinea, represent basal lineages in their 'core Corvoidea' radiation. Several studies have further shown that Petroicidae and Picathartidae are basal members of the Passerida radiation (Ericson *et al*. 2002b; Ericson & Johansson 2003; Barker *et al*. 2004; Jønsson *et al*. 2007). No morphological characters unambiguously support this subdivision, but an amino acid insertion in the c-*myc* gene has been proposed as a synapomorphy for Passerida (Ericson *et al*. 2000; Ericson & Johansson 2003).

Many internal relationships among 'core Corvoidea' and Passerida are poorly supported in the phylogeny by Barker *et al*. (2004). This is particularly the case for deeper relationships within the 'core Corvoidea'. However, independent studies based on other genes overall support the Barker *et al*. (2004) phylogeny, but relationships that are in topological conflict have also been found both within the 'core Corvoidea' (e.g. Fuchs *et al*. 2006) and Passerida (e.g. Ericson & Johansson 2003). Possibly the most intriguing of these is the amino acid insertion in c-*myc* (Ericson *et al*. 2000; Ericson & Johansson 2003) that is not supported as having a single unique origin in the Barker *et al*. (2004) phylogeny. However, that other studies find conflicting relationships to those suggested by Barker *et al*. (2004) is hardly surprising. The Barker *et al*. phylogeny in essence represents a RAG-1 gene tree (75% of the nucleotide data) and gene trees are not necessarily genealogically identical to the species tree (Maddison 1997). In this study, we evaluate the relationships of deep branches in 'core Corvoidea' and Passerida suggested by Barker *et al*. (2004), by adding independent DNA sequence data to their data set. We examine trees from individual genes and various combined data sets, as well as amino acid changes and indel events. We also examine how the phylogenetic results affect molecular clock estimates.

Materials and methods

Taxon sampling, amplification and sequencing

We examined the phylogenetic delimitation of Passerida and core Corvoidea, respectively, by analysing DNA sequence data from 36 taxa. The taxon sampling is based on previous molecular results (Barker *et al*. 2002, 2004; Ericson *et al*. 2002a; Ericson & Johansson 2003), and includes representatives from all recognized basal lineages within Passerida and 'core Corvoidea', as well as a selection of more terminal taxa within these two clades. The sampling also includes representatives from all lineages of oscines that have been suggested to be basal in relation to the 'core Corvoidea'–Passerida split. Five nuclear loci, RAG-1, RAG-2, myoglobin intron 2 (Myo), ornithine decarboxylase introns 6–7 (ODC) and glyceraldehyde-3-phosphodehydrogenase intron 11 (GAPDH), have been studied. The RAG-1 and RAG-2 sequences have been downloaded from GenBank, while sequences from Myo, ODC and GAPDH have either been sequenced for this study or downloaded from GenBank. The latter three have been chosen as they are easy to amplify from study skins (Irestedt *et al*. 2006), and have been shown to be useful to resolve avian relationships at this phylogenetic level (e.g. Jønsson *et al*. 2007). For extraction, amplification and sequencing procedures for fresh tissue/blood samples, see Irestedt *et al*. (2001, 2002), Fjeldså *et al*. (2003) and Allen & Omland (2003), while corresponding procedures for study skins are described in Irestedt *et al*. (2006) and Jønsson *et al*. (in prep). See Table 1 for the complete taxon sampling and GenBank accession numbers.

Phylogenetic inference and model selection

We used Bayesian inference (see, e.g. Huelsenbeck *et al*. 2001; Holder & Lewis 2003) to estimate the phylogenetic relationships. The models for nucleotide substitutions used in the analyses were selected for each gene individually by applying the Akaike Information Criterion (AIC, Akaike 1973) using the program MrModeltest 2.2 (Nylander 2005) in conjunction with PAUP* (Swofford 1998). Due to a rather low number of insertions in the studied genes/introns, the sequences could easily be aligned by eye. All gaps are treated as missing data in the analyses.

Posterior probabilities of trees and parameters in the substitution models were approximated with MCMC and Metropolis coupling using the program MrBayes 3.1.1 (Ronquist & Huelsenbeck 2003). Analyses were performed for (i) all the individual genes separately, (ii) the RAG-1 and RAG-2 genes combined, (iii) a concatenated data set with all genes, and (iv) a data set with all genes *except* the RAG-1 gene. In the analysis of concatenated data sets the models selected for the individual gene partition were used, but the topology was constrained to be the same. We used an unconstrained, exponential branch length prior. All chains were run for 10 million generations, with trees sampled every 100th generations. The trees sampled during the burn-in phase (i.e. before the chain had reached its apparent target distribution) were discarded, and after checking for convergence, final inference was made from the concatenated output from the two runs.

Table 1 Specimen data and GenBank accession numbers for samples used in the study.

Species	Sample ID	ODC	GAPDH	Myo	RAG-1	RAG-2	Species used in Barker *et al.* (2004)
Batis poensis	MNHN CG 1998-783	EU272120	DQ406665	AY529907	AY443263	AY443110	*B. mixta*
Bombycilla garrulus	NRM 986044	EU272128	EU272099	AY228286	AY056981	AY443111	
Callaeas cinerea	Ewen	EU272124	EU272097	EU272108	AY443317	AY443202	*Philesturnus carunculatus*
Chaetops frenatus	PFI uncat.	EF441234	EF441212	AY228289	AY443266	AY443116	
Colluricincla harmonica	MV1422	EU273356	EU273376	EU273396	AY443270	AY443124	
Coracina atriceps	WRZM1910.12.28.182	EU272118	EU272091	EU272102	AY056988	AY443127	*C. lineata*
Cormobates placens	MV E309	EF441237	EF441215	AY064731	AY443274	AY443130	*C. leucophaea*
Corvus corone	MNHN 13-16	EU272116	DQ406663	AY529914	AY056989	AY443132	
Dicrurus bracteatus	UWBM 68045	EU272113	EF052813	EF052839	AY056991	AY443140	*D. adsimilis*
Gymnorhina tibicen	MV AC78	EU272119	DQ406669	AY064741	AY443289	AY443153	
Hirundo rustica	NRM 976238	EF441240	EF441218	AY064258	AY443290	AY443155	
Hylophilus ochraceiceps	ZMUC127900	EU272109	EU272087	EU272100	AY443291	AY443156	*H. poicilotis*
Lanius collaris	MNHN 2-26	EU272112	DQ406662	AY529925	AY443293	AY443160	*L. excubitor*
Loboparadisaea sericea	NRM 566737	EU272125	EU272095	EU272106	AY443294	AY443161	
Cnemophilus loriae	NRM 569572	EU272126	EU272096	EU272107	AY443269	AY443123	
Malurus amabilis	MV C803	EF441241	EF441219	AY064729	AY057001	AY443162	*M. melanocephalus*
Melanocharis versteri	NRM 543385	EU272121	EU272092	EU272103	AY443299	AY443168	
Menura novaehollandiae	AM Lab1112	EF441242	EF441220	AY064744	AY057004	AY443171	
Monarcha melanopsis	B541, UWBM 62890	EU272114	EU272089	DQ084110	AY057006	AY443176	*M. axillaris*
Oedistoma pygmaeum	NRM 569569	EU272122	EU272093	EU272104	AY057010	AY443182	*O. iliolophum*
Oriolus xanthornus	MNHN 4–10D	EU272111	DQ406645	AY529929	AY057011	AY443184	*O. larvatus*
Orthonyx temminckii	MV B831	EF441244	EF441222	AY064728	AY443309	AY443187	
Pachycephala albiventris	ZMUC 117176	EF441245	EF441223	EF441259	AY443310	AY443188	*P. hyperythra*
Pachycephalopsis hattamensis	NRM 552153	EF441246	EF441224	EF441260	AY443311	AY443190	*P. poliosoma*
Parus major	NRM 956363	EU272127	EU272098	AY228310	AY443314	AY443197	
Pericrocotus cinnamomeus	USNM B6146	EU272117	EF052753	EF052764	AY443316	AY443200	*P. ethologus*
Picathartes gymnocephalus	LSUMZ B-19213	EF441247	EF441225	AY228314	AY057019	AY443203	
Pomatostomus temporalis	MV D257	EF441248	EF441226	AY064730	AY057023	AY443210	*P. isidorei*
Prunella modularis	NRM 976138	EF441249	EF441227	AY228318	AY057024	AY443213	*P. collaris*
Ptilonorhynchus violaceus	AM LAB1099	EF441250	EF441228	AY064742	AY057026	AY443216	
Ptilorrhoa leucosticta	NRM 84405	EF441255	EF441233	EF441261	AY443326	AY443218	*P. caerulescens*
Rhipidura rufifrons	C733, CEF239	EU272115	EU272090	DQ084100	AY443329	AY443223	*R. hyperythra*
Sturnus vulgaris	NRM 966615	EF441253	EF441231	AY228322	AY057032	AY443232	
Sylvia atricapilla	NRM 976380	EF441254	EF441232	AY228323	AY057033	AY443233	*S. nana*
Toxorhamphus poliopterus	NRM 543574	EU272123	EU272094	EU272105	AY057036	AY443238	*T. novaeguineae*
Vireo olivaceus	NRM 976766	EU272110	EU272088	EU272101	AY057041	AY443245	*V. philadelphia*

Acronyms: AM, Australian Museum, Sydney; LSUMZ, Louisiana State University, Museum of Natural Science; MNHN, Muséum National d'Histoire Naturelle, Paris; MV, Museum Victoria, Melbourne; NRM, Swedish Museum of Natural History, Stockholm; PFI, Percy Fitzpatrick Institute, Cape Town; USNM, National Museum of Natural History, Smithsonian Institution, Washington; UWBM, University of Washington, Burke Museum; WRZM, Walter Rothschild Zoological Museum, Tring; ZMUC, Zoological Museum of the University of Copenhagen.

Some Myo, GAPDH and ODC sequences, and all RAG-1 and RAG-2 sequences have been downloaded from GenBank. The RAG-1 and RAG-2 sequences that have been obtained from a different species than those used for Myo, GAPDH and ODC are indicated in the table.

All alignments of individual genes were also inspected for indel events, and the protein coding genes (RAG-1 and RAG-2) were also inspected for amino acid changes. One indel event was found to be of potential interest for the 'core Corvoidea'–Passerida split; a deletion in the GAPDH alignment. In order to examine the taxonomic distribution of this indel more carefully, all GAPDH sequences available (downloaded from GenBank and unpublished sequences by us and colleagues at the Swedish Museum of Natural History, data not shown) were inspected for this indel events. In totally 125 GAPDH sequences were examined.

Molecular rate smoothing estimates

Divergence times were estimated using the nonparametric rate smoothing method PATHd8 (Britton *et al.* 2007), which smoothes substitution rates sequentially by taking averages over paths lengths from an internode to all its descending terminals. The primary aim was not to make improved time estimates, but to investigate how different phylogenetic hypotheses affect the molecular clock estimates.

A reliable passerine calibration point based on fossil data is currently lacking, why the separation of New Zealand from Australia/Antarctica has been used as a calibration point for

the separation of *Acanthisitta* from the rest of the passerines in several studies (Barker *et al*. 2002, 2004; Ericson *et al*. 2002a)? As we lack sequences data from *Acanthisitta* for Myo, ODC and GAPDH, we arbitrary used the molecular estimate of the split of *Menura* from the rest of the oscines at 62 Mya (Barker *et al*. 2004). Divergence time estimates were made on the trees obtained from the (i) RAG-1 and RAG-2 data set, (ii) the data set with all genes, and (iii) the data set with all genes except the RAG-1 gene.

Results

Model selection and phylogenetic relationships

A priori selection of substitution models supported that the GTR + I + Γ model had the best fit for RAG-1 and RAG-2, and GTR + Γ for Myo, ODC and GAPDH. These models were used in the Bayesian analyses of the individual genes as well as in the combined analysis. After discarding the burn-in phase the inference were based on a total of 75 000–90 000 samples from the posterior for the individual genes and the combined data sets. For the phylogenetic inference, the mode of the posterior distribution of topologies was presented as a majority-rule consensus tree from each analysis (Figs 1 and 2).

The trees obtained from the Bayesian analyses of the individual gene partitions are all more or less topologically incongruent (Fig. 1), but certain clades are supported by all gene regions. In general, certain regions in the trees are seemingly more congruent than other, while the incongruence is worse in other areas of the trees. Of interest for this study are the relative positions of Callaeatidae, Cnemophilidae, Melanocharitidae (*Toxorhamphus*, *Oedistoma* and *Melanocharis*), Picathartidae and Petroicidae (*Pachycephalopsis*). While the RAG-1 tree suggests that all of these taxa, except Petroicidae, are basal members of core Corvoidea, the position of these taxa are not that clear in the other genes. However, several of them are generally suggested to be basal members of the Passerida radiation, while others are principally unresolved or in some occasion basal to both the Passerida and 'core Corvoidea' radiations. Of all the individual gene trees, the RAG-1 tree is also the most resolved tree and has most nodes with posterior probability values > 95%, followed by the RAG-2 gene, while the trees obtain from the intron regions are the most unresolved and have lowest number of nodes with posterior probability values > 95%. However, this is merely a consequence of the how many phylogenetically informative characters these genes have, respectively.

Variation in the molecular data set

The alignments of the protein coding genes RAG-1 and RAG-2 consists of 2872 and 1152 bp, respectively. A few amino acid indels were observed in the two RAG alignments, all RAG-1 indels were found to be autapomorphic, while two deletions in RAG-2 were shared between *Cormobates*, *Ptilonorhynchus* and *Vireo*, and *Bombycilla* and *Hirundo*, respectively. However, as these indels are of no interest for the division of Passerida and 'core Corvoidea' they are not further discussed. The alignments of the non-coding intron regions were 338 bp for GAPDH, 750 bp for Myo and 758 bp for ODC. Most indels in these more variable regions were found to be short and autapomorphic (one exception is the ODC sequence from *Rhipidura* that has a 83-bp long insertion). Some indels were also found to be incongruent with the phylogenetic tree obtained from the analysis of the combined data sets. However, these were generally found in the most variable regions and some of the single base pair insertions actually consist of different bases. Most indel events congruent with any/all combined phylogenies were found to be of limited interest for the Passerida and 'core Corvoidea' division (e.g. supporting only minor terminal clades) and will not be further discussed. However, one indel was found to be of potential interest for the 'core Corvoidea'–Passerida split; a deletion of 18 bp in the GAPDH alignment were uniquely found in all traditional Passerida species (32 taxa), *Chaetops*, *Picathartes*, *Pachycephalopsis*, Callaeatidae, *Cnemophilus*, *Loboparadisaea*, *Toxorhamphus*, *Oedistoma* and *Melanocharis* (*Lanius* and *Ptilorrhoa* have partly overlapping autapomorphic deletions in GAPDH, but they start and end at different positions). If mapping this indel on the combined trees (Fig. 2), the GAPDH deletion has its most parsimonious distribution on the tree obtained from the analyses of the combined data set of all genes expect RAG-1 (tree C, Fig. 2). No amino acid substitutions were found in the protein coding genes RAG-1 and RAG-2 that could be of importance for the 'core Corvoidea'–Passerida split.

Molecular rate smoothing estimates

Divergence estimates from the RAG-1/RAG-2 tree and the tree based on all genes except RAG-1 are generally rather similar, but for some nodes the time estimates are strikingly different. The most noticeable, is that Callaeatidae is suggested to have diverged from other oscines 30.5 Mya in the estimate from the RAG-1/RAG-2 tree, while the corresponding estimate from the tree based on all genes except RAG-1 suggest the Callaeatidae diverged from the 'Passerida' linage as early as 46 Mya.

Discussion

Species tree, gene trees and the 'core Corvoidea'–Passerida split

The trees obtained from the Bayesian analyses of the individual gene partitions are more or less topologically incongruent (Fig. 1), especially regarding the positions of Callaeatidae, Cnemophilidae, Melanocharitidae and Picathartidae. The analyses of the combined data set (Fig. 2) also shows that the

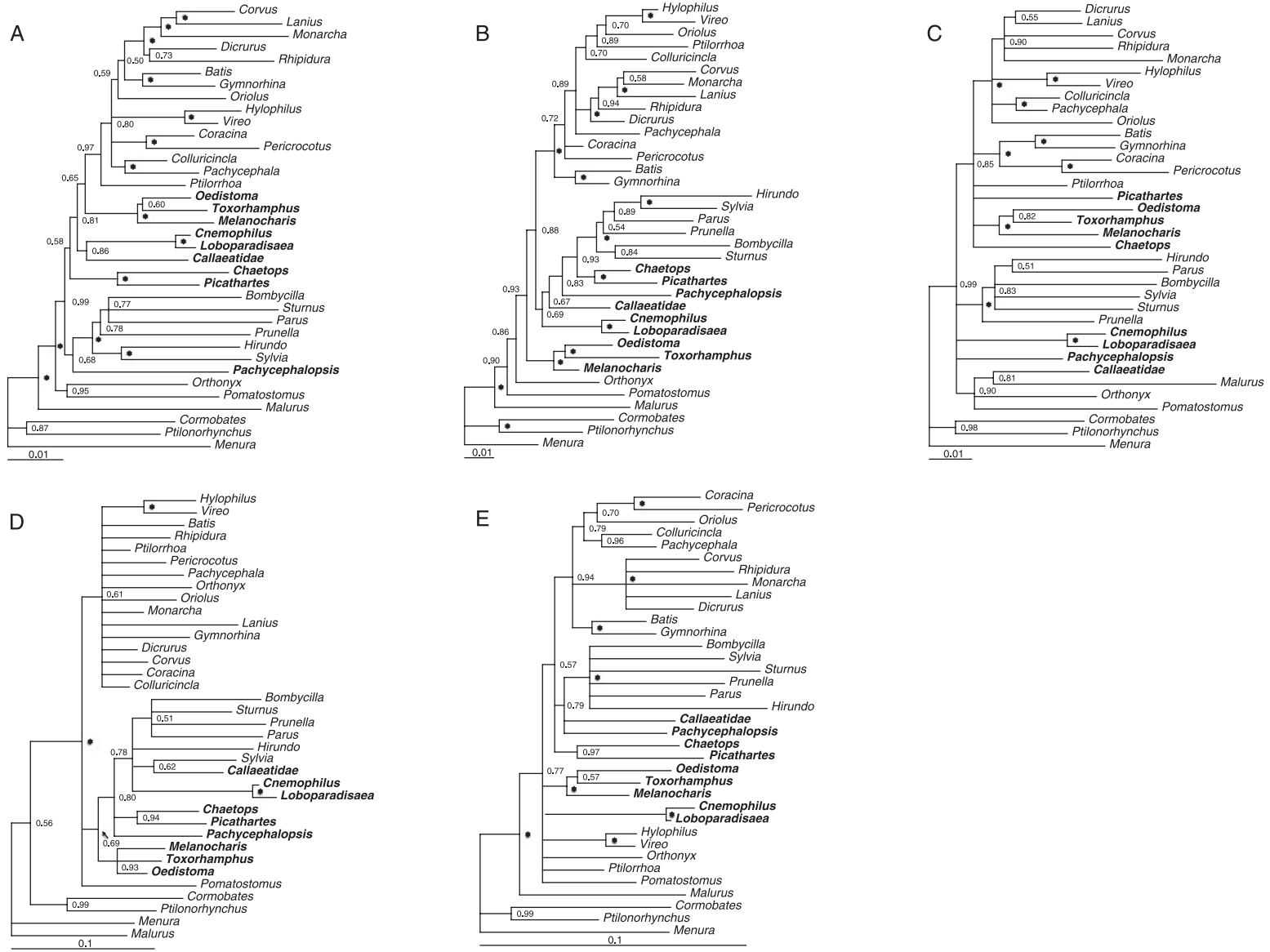

Fig. 1 A–E. The majority rule consensus trees obtained from the Bayesian analyses of the individual genes. —A. RAG-1. —B. RAG-2. —C. Myo. —D. GAPDH. —E. ODC. Posterior probability values are indicated at the node, posterior probability values of 1.00 are indicated with an asterisk. Lineages suggested by Barker *et al.* (2004) to be basal in 'core Corvoidea' (Callaeatidae, Cnemophilidae and Melanocharitidae) and in Passerida (Petroicidae and Picathartidae) are in bold type.

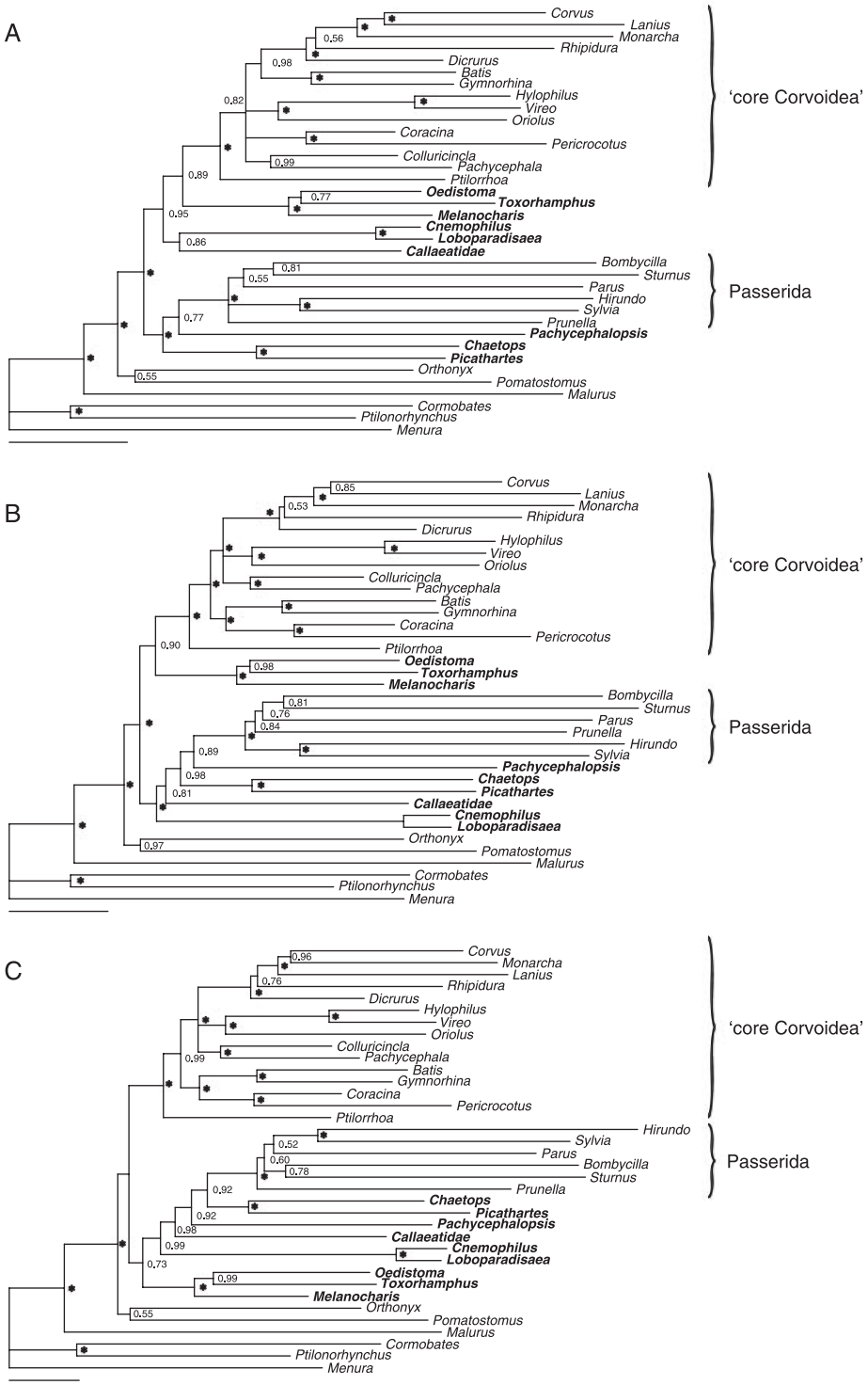

Fig. 2 A–C. The majority rule consensus trees obtained from the Bayesian analyses of combined data sets. —A. The tree obtained from the analyses of RAG-1 and RAG-2. —B. The tree obtained from the analyses of all genes (RAG-1, RAG-2, Myo, GAPDH and ODC). —C. The tree obtained from all genes except RAG-1. Posterior probability values are indicated at the node, posterior probability values of 1.00 are indicated with an asterisk. Lineages suggested by Barker *et al*. (2004) to be basal in 'core Corvoidea' (Callaeatidae, Cnemophilidae and Melanocharitidae) and in Passerida (Petroicidae and Picathartidae) are in bold type.

phylogenetic positions for these taxa largely depend on which genes that has been combined. It is therefore worth noticing that the phylogeny by Barker *et al.* (2004) is based on only two genes (RAG-1 and RAG-2) and that RAG-1 represents nearly 75% of the nucleotide data included in the study. In essence, the Barker *et al.* (2004) phylogeny is a RAG-1 gene tree, and gene trees are not necessarily topologically identical to the species tree (Maddison 1997). This is often neglected in avian molecular phylogenetics, although incongruence is a common phenomenon in molecular phylogenetic studies and has been reported at many avian levels (e.g. Degnan 1993; Alström & Ödeen 2002; Irestedt *et al.* 2004; Moyle 2004; Fjeldså *et al.* 2005; Fuchs *et al.* 2006). Observed incongruence between gene trees could be an effect of both biological processes (Maddison 1997), and various analytical factors such as the choice of optimality criterion (Huelsenbeck 1994) or taxon sampling (Graybeal 1998; Hedtke *et al.* 2006). It has also been found that current tests of incongruence are not always reliable (Sullivan 1996; Cunningham 1997). Therefore, in practice, incongruence between gene trees may be difficult to handle. Nevertheless, this is a problem that has to be considered in avian molecular systematics. Theoretically, a phylogeny based on DNA sequences from multiple independent loci should have a better chance of identifying the correct species tree than a single gene tree, by increasing the signal : noise ratio. Edwards *et al.* (2007) have demonstrated that there is a high probability of recovering the correct species tree (for eight taxa) with only three genes, if a large proportion of the genes have phylogenies that matches the species tree. On the other hand it was also shown that more than 100 genes might be needed to recover the correct species tree with reasonably high probability, if a low proportion of the gene trees are congruent with the species tree. Consequently, the information needed to resolve the correct species tree may differ significant among and within clades, due to factors such as the particular phylogenetic history (e.g. the occurrence of long and short branches within a tree and the relationship between them) and properties of the studied genes (e.g. number of variable sites, saturation, etc.).

As a consequence of this, many more loci than those available at present may be needed before a well-supported phylogeny of all oscine birds can be obtained. In the present era of genomics, it is technically possible to sequence hundreds of independent loci for all oscine families. However, in practice the funding in avian molecular systematics is limited and it is reasonable to believe that it will be long before a data set (with a sufficient number of independent loci) is available, that is satisfactorily powerful to resolve all nodes in the oscine species tree correctly. As branch lengths often vary considerably within a species tree, incongruence is likely to be more common in certain regions of a species tree (Rokas *et al.* 2003; Kubatko & Degnan 2007). It is also commonly observed that certain taxa often change positions in phylogenies when different markers are used while other taxa are more firmly placed regardless of the genes used. As most avian phylogenetic studies published at present use more than one gene, it is possible to compare individual gene trees for congruence and incongruence. Even with a data set with rather few genes, it should be possible to identify stable and unstable regions in a given species tree with some confidence. In Barker *et al.* (2004) phylogeny two genes were used, the RAG-1 and RAG-2. Although these genes are closely linked the phylogenies obtained from them are not topologically congruent. Especially, when considering the 'core Corvoidea'–Passerida split these two genes favour two different scenarios (trees A and B, Fig. 1), and if all gene trees in this study are considered (trees A–E, Fig. 1) it is obvious that the phylogenetic positions of Callaeatidae, Cnemophilidae and Melanocharitidae are difficult to assess. RAG-1 is the only gene that assigns these taxa to the 'core Corvoidea' radiation. The other gene trees are either virtually unresolved (Myo), place some of the taxa within Passerida (ODC) or place all of them within Passerida (RAG-2 and GAPDH).

When comparing the trees obtained from the combined data sets of RAG-1 and RAG-2, with the trees of all genes, and with the trees of all genes excluding RAG-1 (Fig. 2A–C), some interesting patterns are elucidated. First, the tree obtained from the RAG-1 and RAG-2 combined data set, is very similar to the individual tree from the RAG-1 gene. This is hardly surprising, as RAG-1 has more than two times, as many, parsimony informative sites as does RAG-2. In the tree obtained from the analyses of all genes, Callaeatidae and Cnemophilidae move from a basal position in the 'core Corvoidea' to become basal members of the Passerida and Melanocharitidae are placed as the basalmost clade in the 'core Corvoidea' (although with a posterior probability of only 0.90). Furthermore, the combined analysis of all genes except RAG-1, supports a tree where also the Melanocharitidae clade becomes a basal member of the Passerida, but again with low support (posterior probability 0.73). These results clearly illustrate that relationship that appear to be well supported (in this case by several thousand basepairs!) by one or a limited number of loci, could in fact be very unstable and that a few additional genes might alter the topology considerably. An obvious conclusion from these results is that many biological relationships based on molecular data from a limited number of loci need to be further substantiated by independent markers, before we can consider these phylogenetic hypotheses well supported.

The affinity of Callaeatidae, Cnemophilidae and Melanocharitidae, divergence date estimates and conclusions

The phylogenetic affinity of Callaeatidae, Cnemophilidae, Melanocharitidae and other basal passerida and 'core Corvoidea'

clades deserves further attention. Additional independent loci would most likely cast more light over this part of the oscine phylogeny. However, based on current data we believe that we have a reasonable support for a Passerida affinity of Callaeatidae and Cnemophilidae, while the affinity of Melanocharitidae is more uncertain. We consider the phylogeny based on all genes except RAG-1 as the most plausible hypothesis for this part of the oscine phylogeny as: (i) it is fully congruent with the unique deletion observed in GAPDH in core Passerida, Petroicidae, Picathartidae, Callaeatidae, Cnemophilidae and Melanocharitidae; (ii) the topology is consistent with a monophyletic origin of the amino acid insertion in c-*myc* found in core Passerida and Picathartidae; (iii) the position of Callaeatidae, Cnemophilidae and Melanocharitidae is most deviant in the RAG-1 gene tree; and (iv) it is the biogeographically most parsimonious hypothesis as it requires only one major dispersal event of the Passerida branch from the Australo-Papuan region to the Old World.

To understand the evolution and biogeographical history of oscines or other avian clades, molecular clock estimations on DNA sequence data is an important and commonly used tool. Many factors such as uncertainties in calibrations (e.g. Graur & Martin 2004) none clocklike evolution, or inconsistent use of the various calibrations methods available (Peterson 2006), could make these estimates less reliable. An additional, obvious but often overlooked problem with molecular time estimates is how well the phylogenies, on which the estimates are based, represent the species phylogeny. Herein, we have shown drastically topological changes for a group of taxa when different genes have been used to interpret their affinity, and this will obviously also affect the time estimates of the divergences of these taxa. When we compared some divergence estimates from the RAG-1/RAG-2 tree and the tree based on all genes except RAG-1 most estimates were found to be rather similar, but some nodes were found where the time estimates are strikingly different. The most noticeable is that Callaeatidae is suggested to have diverged from other oscines 30.5 Mya in the estimate from the RAG-1/RAG-2 tree, while the corresponding estimate from the tree based on all genes except RAG-1 suggest the Callaeatidae diverged from the 'Passerida' linage as early as 46 Mya. Such large discrepancy is an obvious source for incorrect interpretations of biogeography or evolutionary responses to habitat changes.

The results in this study illustrate the potential problem of using a small number of genes in avian systematics, and especially when a majority of the phylogenetic informative characters have been obtained from one (or a few) loci. We advocate that individual genes always should be analysed separately (apart for the combined analysis) and inspected for topological congruence/incongruence, as this makes it possible to discriminate, with some confidence, between parts of a tree that are well and poorly supported. When possible, we also advocate selecting multiple independent loci that contribute roughly equal to the phylogenetic information, rather than using very long sequences from single loci. The results also illustrate the importance of well-supported phylogenies before reasonable time estimates of passerine divergences could be achieved.

Acknowledgements

Footpad samples have been obtained from the Department of Vertebrate Zoology, Swedish Museum of Natural History (Göran Frisk and Per G. P. Ericson). Knud A. Jønsson, Ulf S. Johansson, Dario Zuccon, Per G. P. Ericson, and three anonymous reviewers are thanked for comments on the manuscript. Pia Eldenäs, Annika Einarsson, Mattias Myrenås and Keyvan Mirbakhsh are thanked for practical support at the laboratory. The laboratory work was founded by a grant to MI from Magnus Bergvalls Stiftelse.

References

Akaike, H. (1973). Information theory as an extension of the maximum likelihood principle. In B. N. Petrov & F. Csaki (Eds) *Second International Symposium on Information Theory* (pp. 267–281). Budapest: Akademiai Kiado.

Allen, E. S. & Omland, K. E. (2003). Novel intron phylogeny (ODC) supports plumage convergence in orioles (*Icterus*). *The Auk*, *120*, 961–969.

Alström, P. & Ödeen, A. (2002). Incongruence between mitochondrial DNA, nuclear DNA and non-molecular data in the avian genus *Motacilla*: implications for estimates of species phylogenies. In P. Alström (Ed.) *Species Limits and Systematics in Some Passerine Birds*. PhD Thesis. Uppsala University.

Ames, P. L. (1971). The morphology of the syrinx in passerine birds. *Bulletin of the Peabody Museum of Natural History*, *37*, 1–194.

Barker, K. F., Barrowclough, G. F. & Groth, J. G. (2002). A phylogenetic analysis for passerine birds: taxonomic and biogeographic implications of an analysis of nuclear DNA sequence data. *Proceedings of the Royal Society of London B*, *269*, 295–305.

Barker, F. K., Cibois, A., Schikler, P. A., Feinstein, J. & Cracraft, J. (2004). Phylogeny and diversification of the largest avian radiation. *Proceedings of the National Academy of Sciences of the USA*, *101*, 11040–11045.

Beecher, W. J. (1953). A phylogeny of oscines. *The Auk*, *70*, 270–333.

Britton, T., Anderson, C. L., Jacquet, D., Lundqvist, S. & Bremer, K. (2007). Estimating divergence times in large phylogenetic trees. *Systematic Biology*, *56*, 741–752.

Cracraft, J. (1987). DNA hybridization and avian phylogenetics. *Evolutionary Biology*, *21*, 47–96.

Cunningham, C. W. (1997). Can three incongruence tests predict when data should be combined? *Molecular Biology and Evolution*, *14*, 733–740.

Degnan, S. M. (1993). The perils of single gene trees — mitochondrial versus single-copy nuclear DNA variation in white-eyes (Aves: Zosteropidae). *Molecular Ecology*, *2*, 219–225.

Edwards, S. V., Liu, L. & Pearl, D. K. (2007). High-resolution species trees without concatenation. *Proceedings of the National Academy of Sciences of the USA*, *104*, 5936–5941.

Ericson, P. G. P. & Johansson, U.S. (2003). Phylogeny of Passerida (Aves: Passeriformes) based on nuclear and mitochondrial sequence data. *Molecular Phylogenetics and Evolution*, *29*, 126–138.

Ericson, P. G. P., Johansson, U.S. & Parsons, T. J. (2000). Major divisions of oscines revealed by insertions in the nuclear gene c-*myc*: a novel gene in avian phylogenetics. *The Auk*, *117*, 1077–1086.

Ericson, P. G. P., Christidis, L., Cooper, A., Irestedt, M., Jackson, J., Johansson, U.S. & Norman, J. A. (2002a). A Gondwanan origin of passerine birds supported by DNA sequences of the endemic New Zealand wrens. *Proceedings of the Royal Society of London B*, *269*, 235–241.

Ericson, P. G. P., Christidis, L., Irestedt, M. & Norman, J. A. (2002b). Systematic affinities of the lyrebirds (Passeriformes: *Menura*), with a novel classification of the major groups of passerine birds. *Molecular Phylogenetics and Evolution*, *25*, 53–62.

Fjeldså, J., Zuccon, D., Irestedt, M., Johansson, U.S. & Ericson, P. G. P. (2003). *Sapayoa aenigma*: a New World representative of 'Old World suboscines'. *Proceedings of the Royal Society of London B*, *270*, 238–241.

Fjeldså, J., Irestedt, M. & Ericson, P. G. P. (2005). Molecular data reveal some major adaptational shifts in the early evolution of the most diverse avian family, the Furnariidae. *Journal of Ornithology*, *146*, 1–13.

Fuchs, J., Fjeldså, J. & Pasquet, E. (2006). An ancient African radiation of corvoid birds (Aves: Passeriformes) detected by mitochondrial and nuclear sequence data. *Zoologica Scripta*, *35*, 375–385.

Graur, D. & Martin, W. (2004). Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision. *Trends in Genetics*, *20*, 80–86.

Graybeal, A. (1998). Is it better to add taxa or characters to a difficult phylogenetic problem? *Systematic Biology*, *47*, 9–17.

Harshman, J. (1994). Reweaving the tapestry: what can we learn from Sibley & Ahlquist (1990)? *The Auk*, *111*, 377–388.

Hedtke, S. M., Townsend, T. M. & Hillis, D. M. (2006). Resolution of phylogenetic conflict in large data sets by increased taxon sampling. *Systematic Biology*, *55*, 522–529.

Holder, M. & Lewis, P. O. (2003). Phylogeny estimation: traditional and Bayesian approaches. *Nature Genetics*, *4*, 275–284.

Houde, P. (1987). Critical evaluation of DNA hybridization studies in avian systematics. *The Auk*, *104*, 17–32.

Huelsenbeck, J. P. (1994). Performance of phylogenetic methods in simulation. *Systematic Biology*, *44*, 17–48.

Huelsenbeck, J. P., Ronquist, F., Nielsen, R. & Bollback, J. P. (2001). Reverend Bayes meets Darwin: Bayesian inference of phylogeny and its impact on evolutionary biology. *Science*, *288*, 2349–2350.

Irestedt, M., Johansson, U.S., Parsons, T. J. & Ericson, P. G. P. (2001). Phylogeny of major lineages of suboscines (Passeriformes) analysed by nuclear DNA sequence data. *Journal of Avian Biology*, *32*, 15–25.

Irestedt, M., Fjeldså, J., Johansson, U.S. & Ericson, P. G. P. (2002). Systematic relationships and biogeography of the tracheophone suboscines (Aves: Passeriformes). *Molecular Phylogenetics and Evolution*, *23*, 499–512.

Irestedt, M., Fjeldså, J., Nylander, J. A. A. & Ericson, P. G. P. (2004). Phylogenetic relationships of typical antbirds (Thamnophilidae) and test of incongruence based on Bayes factors. *BMC Evolutionary Biology*, *4*, 23.

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). *Zoologica Scripta*, *35*, 567–580.

Jønsson, K. A., Fjeldså, J., Ericson, P. G. P. & Irestedt, M. (2007). Systematic placement of the enigmatic *Eupetes macrocerus* and implications for the biogeography of a main songbird radiation, the Passerida. *Biology Letters*, *3*, 323–326.

Kubatko, L. S. & Degnan, J. H. (2007). Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Systematic Biology*, *56*, 17–24.

Maddison, W. (1997). Gene trees in species trees. *Systematic Biology*, *46*, 523–536.

Monroe, B. L. & Sibley, C. G. (1993). *A World Checklist of Birds*. New Haven and London: Yale University Press.

Moyle, R. G. (2004). Phylogenetics of barbets (Aves: Piciformes) based on nuclear and mitochondrial DNA sequence data. *Molecular Phylogenetics and Evolution*, *30*, 187–200.

Nylander, J. A. A. (2005). *MRMODELTEST v.2.2*. [Program distributed by the author]. Uppsala University, Uppsala: Department of Systematic Zoology.

Peterson, A. T. (2006). Application of molecular clocks in ornithology revisited. *Journal of Avian Biology*, *37*, 541–544.

Raikow, R. J. (1978). Appendicular myology and relationships of the New World nine-primaried oscines (Aves: Passeriformes). *Bulletin of the Carnegie Museum of Natural History*, 7, 1–43.

Rokas, A., Williams, B. L., King, N. & Carroll, S. B. (2003). Genome-scale approaches to resolving incongruence in molecular phylogenies. *Nature*, *23*, 798–803.

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 the Birds of the World*. New Haven, CT: Yale University Press.

Sullivan, J. (1996). Combining data with different distributions of among-site rate variation. *Systematic Biology*, *45*, 375–380.

Swofford, D. L. (1998). *PAUP*. Phylogenetic Analysis Using Parsimony (* and Other Methods)*, *ver.4*. [Computer program and manual] Sunderland, Massachusetts: Sinauer Associates, Inc. Publisher.

