



Phylogeography of Indonesian and Sino-Himalayan region bush warblers (*Cettia*, Aves)

Urban Olsson ^{a,*}, Per Alström ^{b,c}, Magnus Gelang ^{c,d}, Per G.P. Ericson ^b, Per Sundberg ^a

^a Department of Zoology, Göteborg University, Medicinaregatan 18, SE-413 90 Göteborg, Sweden

^b Department of Vertebrate Zoology and Molecular Systematics Laboratory, Swedish Museum of Natural History, P.O. Box 50007, SE-104 05 Stockholm, Sweden

^c Department of Systematic Zoology, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18 D, SE-752 36 Uppsala, Sweden

^d Department of Zoology, Stockholm University, SE-106 91 Stockholm, Sweden

Received 18 January 2006; revised 25 April 2006; accepted 11 May 2006

Available online 19 May 2006

Abstract

We present a hypothesis for the phylogeny and phylogeography of a group of bush warblers in the genus *Cettia*, based on parts of the mitochondrial Cytochrome *b* gene and the nuclear myoglobin intron II (in all ~1.7 kb). Ancestral areas were reconstructed by dispersal-vicariance analysis and constrained Bayesian inference. The results suggest that the insular taxa in the *Cettia vulcania* group are most closely related to *Cettia flavolivacea*, and originated from a dispersal by an ancestral population in the Himalayas towards the south, to the Sunda region. From this population, a second dispersal along a different route colonized China and northern Vietnam. Hence, the Chinese taxon *intricata* and Vietnamese *oblita*, currently allocated to *C. flavolivacea*, are more closely related to the *vulcania* group than to the other taxa in the *flavolivacea* group, and we propose that they be treated as conspecific with *C. vulcania*, restricting *C. flavolivacea* to Myanmar and the Himalayas.

© 2006 Elsevier Inc. All rights reserved.

Keywords: *Cettia*; Dispersal-vicariance analysis; Ancestral state reconstruction; Cytochrome *b*; Myoglobin intron II

1. Introduction

The bush warblers in the avian genus *Cettia* are small insectivorous birds inhabiting dense forest understorey, scrub and grassy areas, from the plains to above the tree limit. Most species are found in the mountains and foothills of southern and eastern Asia, but one extends its range to the plains of Europe and North Africa, and several taxa occur in Indonesia and on Australasian islands. Most taxa are sedentary or altitudinal migrants, but those breeding furthest to the north are medium-distance migrants. They are non-descript, brown above and paler below, with a brownish, greyish or yellowish wash and show a variably prominent pale stripe above the eye. Their tails have 10 rectrices, unlike nearly all other oscine

passerines, which have 12. Altogether approximately 45 taxa are recognized, separated into 10–14 species (Dickinson, 2003; Watson et al., 1986), but the genus has recently been shown to be non-monophyletic (Alström et al., 2006).

From the Himalayas to southern China, the three species *Cettia flavolivacea*, *Cettia fortipes* and *Cettia acanthizoides* occur sympatrically, the ranges of the latter two extending as far east as Taiwan. These are morphologically similar, but their territorial songs differ markedly (e.g. Rasmussen and Anderton, 2005). Alström et al. (in press) has proposed, based on analyses of morphology, vocalizations and DNA, that *C. acanthizoides* be split into two species, *C. acanthizoides* sensu stricto in China and *Cettia brunnescens* in the Himalayas. In the breeding season, the genus *Cettia* is virtually absent from southern Myanmar, Thailand, southern Laos, southern Vietnam, Cambodia and continental Malaysia. However, it is found again in Indonesia and East

* Corresponding author. Fax: +46 31 416729.

E-mail address: urban.olsson@zool.gu.se (U. Olsson).

Table 1
Taxonomy of *Cettia vulcania* and *Cettia flavolivacea* based on Dickinson (2003), and the present study

Dickinson (2003)	Distribution	Proposed new taxonomy
<i>Cettia vulcania</i>		
<i>C. v. vulcania</i> (Blyth, 1870)	E Sumatra, Java, Bali, Lombok (Indonesia)	<i>C. v. vulcania</i> (Blyth, 1870)
<i>C. v. flaviventris</i> (Salvadori, 1879) ^a	C–S Sumatra (Indonesia)	<i>C. v. flaviventris</i> (Salvadori, 1879)^a
<i>C. v. sepiaria</i> Kloss, 1931	N Sumatra (Indonesia)	<i>C. v. sepiaria</i> Kloss, 1931
<i>C. v. kolichisi</i> Johnstone and Darnell, 1997	Alor (Indonesia)	<i>C. v. kolichisi</i> Johnstone and Darnell, 1997
<i>C. v. everetti</i> Hartert, 1898	Timor (Indonesia)	<i>C. v. everetti</i> Hartert, 1898
<i>C. v. banksi</i> Chasen, 1935	NW Borneo (Indonesia)	<i>C. v. banksi</i> Chasen, 1935
<i>C. v. oreophila</i> Sharpe, 1888	NE Borneo (Indonesia)	<i>C. v. oreophila</i> Sharpe, 1888
<i>C. v. palawana</i> Ripley and Rabor, 1962	Palawan (Philippines)	<i>C. v. palawana</i> Ripley and Rabor, 1962
<i>Cettia flavolivacea</i>		
<i>C. f. intricata</i> (Hartert, 1909)^b	N and E Myanmar, C China	<i>C. v. intricata</i> (Hartert, 1909)^b
<i>C. f. oblita</i> (Mayr, 1941)	N Indochina	<i>C. v. oblita</i> (Mayr, 1941)
<i>Cettia flavolivacea</i>		
<i>C. f. flavolivacea</i> (Blyth, 1845)	C and E Himalayas	<i>C. f. flavolivacea</i> (Blyth, 1845)
<i>C. f. stresemanni</i> (Koelz, 1954)	Meghalaya (India)	<i>C. f. stresemanni</i> (Koelz, 1954)
<i>C. f. alexanderi</i> (Ripley, 1951) ^c	Naga to Mizoram hills (India)	<i>C. f. alexanderi</i> (Ripley, 1951) ^c
<i>C. f. weberi</i> (Mayr, 1941)	W Myanmar	<i>C. f. weberi</i> (Mayr, 1941)

^a Previously referred to as *Cettia montana sumatrana* (see Violani, 1980).

^b Includes *dulcivox* (see Dickinson, 2003).

^c Identical to *weberi*, *fide* Rasmussen and Anderton (2005). Taxa included in the present analysis in bold.

Malaysia, where only one taxon is present at each locality, making affinities and taxonomic status difficult to interpret.

The phylogenetic position of the polytypic *Cettia vulcania* from Indonesia and Palawan (Philippines) is unknown (Table 1), and its taxonomic history confusing. The nominate subspecies was originally described from Java as *Sylvia montana* Horsfield, 1821. This name was later amended to *Sylvia vulcania* Blyth, 1870, as *S. montana* was preoccupied (by *S. montana* Wilson, 1812 = *Dendroica virens*). Nevertheless, the names *Cettia montanus* or *Cettia montana* have been used for *C. vulcania* in more recent literature (e.g. Vaurie, 1959). Moreover, Vaurie's (1959) circumscription of *C. montana* includes *C. fortipes*. Most recent authors (e.g. Dickinson, 2003; Inskipp et al., 1996; Watson et al., 1986; White and Bruce, 1986) follow Wells (1982) in treating *C. vulcania* and *C. fortipes* as separate species, although Watson et al. (1986) considered them to be each other's nearest relatives and form a superspecies.

We here analyze the phylogeny of the montane bush warblers of southern continental Asia and Indonesia west of the Wallace Line, based on the mitochondrial (mt) Cytochrome *b* gene and the nuclear myoglobin intron II. We reconstruct the ancestral distributions by dispersal-vicariance analysis and constrained Bayesian inference, and revise the taxonomy of the group.

2. Material and methods

2.1. Study group

Blood or feathers were collected from live birds representing 13 *Cettia* taxa from 9 localities in China, India, Indonesia, Japan, Malaysia, Myanmar, Spain and Vietnam (Fig. 1, Appendix A). From two localities (West Bengal, India and

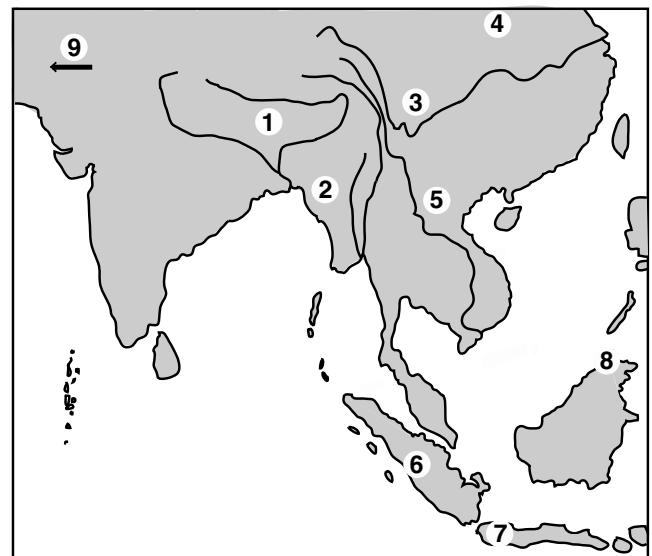


Fig. 1. Geographic origin of the samples used in this study. The localities are: (1) West Bengal, India (origin of *Cettia brunnescens*, *Cettia flavolivacea flavolivacea* and *Cettia fortipes fortipes*); (2) Chin Hills, Myanmar (origin of *Cettia flavolivacea weberi*); (3) Sichuan, China (origin of *Cettia acanthizoides acanthizoides*, *Cettia flavolivacea intricata* and *Cettia fortipes davidiana*); (4) Shaanxi, China (origin of *Cettia diphone canturians*); (5) Tonkin, Vietnam (origin of *Cettia flavolivacea oblita*); (6) Sumatra, Indonesia (origin of *Cettia vulcania flaviventris*); (7) Java, Indonesia (origin of *Cettia vulcania vulcania*); (8) Borneo, Malaysia (origin of *Cettia vulcania oreophila*). Locality (9) indicates Spain and Sweden (origin of *Cettia cetti* and *Phylloscopus trochilus trochilus*, respectively), outside the map.

Sichuan, China) we include representatives of three sympatric taxa. A pilot study revealed that the closest relative of *C. vulcania* appears to be *C. flavolivacea*, which prompted us to include a fuller range of taxa from this species. *Phylloscopus* is one of the closest relatives of *Cettia* (Alström et al., 2006), and *Phylloscopus trochilus* was used as an outgroup.

2.2. DNA extraction and sequencing

Extraction, amplification and sequencing of the Cytochrome *b* gene and myoglobin intron II were performed as described in Olsson et al. (2005).

2.3. Phylogenetic analyses

The sequences were aligned using MegAlign 4.03 in the DNASTAR package (DNA STAR, Inc.), with some manual adjustment for the myoglobin intron II sequences; the algorithm used was Clustal V (Higgins et al., 1992). Molecular phylogenies were estimated by Bayesian inference (BI) using MrBayes 3.1 (Huelsenbeck and Ronquist, 2001, 2005). The choice of model was determined based on the Akaike Information Criterion (Akaike, 1973) and a hierarchical likelihood ratio test (Posada and Crandall, 1998), both calculated in MrModeltest2 (Nylander, 2004). Phylogenies were estimated for the Cytochrome *b* and myoglobin intron II data independently, and for the concatenated sequences. For the Cytochrome *b* data set, posterior probabilities were calculated under a general time-reversible (GTR) model (Lanave et al., 1984; Rodríguez et al., 1990; Tavaré, 1986), assuming rate variation across sites according to a discrete gamma distribution with four rate categories (Γ_4 ; Yang, 1994) and an estimated proportion of invariant sites (I; Gu et al., 1995). For the myoglobin intron II data, the model selected by the Akaike Information Criterion was HKY (Hasegawa et al., 1985), and by the hierarchical likelihood ratio test K80 (Kimura, 1980). In the BI of the concatenated sequences, the data were partitioned, such that the non-coding myoglobin intron II and the protein-coding Cytochrome *b* were analyzed separately, using rate multipliers to allow different rates for the different partitions (Nylander et al., 2004; Ronquist and Huelsenbeck, 2003). Four Metropolis-coupled MCMC chains were run for 5 million generations and sampled every 100 generations. The first 100,000 generations, before the chain reached apparent stationarity (burn-in), were discarded and the posterior probability estimated for the remaining generations. Every analysis was repeated four times, starting from random trees, and the results compared to ascertain that the chains had reached the same target distributions (as suggested by Huelsenbeck et al., 2002). The samples from the stationary phases of the independent runs were pooled to obtain the final results.

Clade support for the unweighted data set was also assessed by parsimony bootstrapping in PAUP* 4.08b (Swofford, 2001), under the following settings: heuristic search strategy, starting trees obtained via random stepwise addition, 10 replicates, followed by TBR branch swapping, MulTrees option in effect, gaps treated as fifth base, 10,000 replicates.

2.4. Phylogeographic analyses

A dispersal-vicariance analysis (Ronquist, 1997) was performed on the all compatible 50% consensus tree,

using DIVA 1.1a (Ronquist, 1996). The areas of analysis were defined based on a priori knowledge of distributional and taxonomic patterns in the group. Area A refers to the entire Palearctic region; area B was defined as the Himalayan range east to the Salween River; area C refers to the Chin Hills in western Myanmar; area D are the ranges of the widely disjunct populations in Indonesia and Palawan (the Philippines); area E is China, east of the Salween River and north Vietnam. The outgroup taxa were chosen primarily for rooting the phylogeny, not for determining the ancestral area of the entire ingroup. To achieve this, many more taxa in the genus *Cettia*, and related taxa would have to be added. Ancestral distributions were reconstructed first by an unrestricted exact search, and then with the maximum number of unit areas constrained to two. This constraint of the number of ancestral areas assumes that the dispersal ability of the ancestor was no higher than that of the descendants, and that the ancestor was not present in more than two unit areas (Sanmartín, 2003).

An ancestral state reconstruction (ASR) (Huelsenbeck and Bollback, 2001; Ronquist, 2004) was performed as a series of separate constrained Bayesian inference for each node on the concatenated sequences, based on the all compatible 50% consensus tree. The distribution data were included in a separate partition, assuming the evolutionary rate to be proportional to that of the other data. Otherwise, models and settings were the same as for the phylogenetic analyses, but only one analysis of 2 million generations was run per node. The most basal nodes were not included in this analysis, as the taxon sampling is incomplete in this part of the tree, rendering the results uninformative and potentially misleading.

3. Results

3.1. Sequence characteristics

DNA sequences for a contiguous 1038 base pair (bp) portion of the Cytochrome *b* gene (corresponding to positions 14,995–16,033 in the chicken mitochondrial sequence (Desjardins and Morais, 1990)) and 690–711 bp of the myoglobin intron II, including the complete intron II and flanking regions of exons 2 (13 bp) and 3 (10 bp) corresponding to positions 8,413,341–8,414,282 in the chicken chromosome 1 sequence (GenBank Accession No. NW 060209, International Chicken Genome Sequencing Consortium, 2004) were obtained from 14 populations of *C. vulcania*, *C. flavolivacea* and *C. fortipes*, and single samples of *C. acanthizoides*, *C. brunnescens*, *Cettia diphone*, *Cettia cetti* and the outgroup *P. trochilus*. The Cytochrome *b* sequence was amplified as one fragment to decrease the risk of amplifying nuclear pseudocopies of the gene (e.g. Arctander, 1995; Quinn, 1997; Quinn and White, 1987; Sorensen and Quinn, 1998). No unexpected start or stop codons that could indicate the presence of nuclear copies are present in the Cytochrome *b* sequences.

There were 1755 characters in the combined data matrix, of which 180 (10%) were parsimony informative. The fragments were easily aligned, although there were 10 positions with indels in the myoglobin intron II sequence, of which 9 were singletons and one 12 bases long.

3.2. Relationships

The trees resulting from BI of the independent Cytochrome *b* and myoglobin intron II data sets analyzed separately are shown in Fig. 2. Both are congruent, adding support to a hypothesis that they may represent the true organism phylogeny, not just gene-trees. The trees estimated by BI and parsimony bootstrapping of the concatenated data sets have identical topologies (Fig. 3).

Except for the taxa *intricata* (central China) and *oblita* (northern Vietnam), the phylogenetic analyses unambiguously place the different populations in clades congruent with the most widely recognized species (Table 1). There is no support for a sister relationship between *C. fortipes* and *C. vulcania*. Instead, the taxa allocated to *C. vulcania* are

part of the *C. flavolivacea* clade (clade A, Fig. 3). In this clade, there is a basal dichotomy separating the five populations from central China, north Vietnam and Indonesia (clade A1, Fig. 3) from the three populations from the Himalayas and west Myanmar (clade A2, Fig. 3). The positions of *intricata* (central China) and *oblita* (Vietnam) render *C. flavolivacea* and *C. vulcania* non-monophyletic.

3.3. Phylogeography

The unrestricted exact DIVA search resulted in a single optimal reconstruction, which required four dispersals; the ancestral area reconstruction of the basal nodes included all five areas. The second search, in which the maximum number of unit areas were constrained to two, resulted in 1280 optimal reconstructions, requiring seven dispersals. The optimal area reconstructions at each ancestral node for the second search are summarized in Fig. 4. With unit areas constrained to two, the ancestral area reconstruction in nodes 12 and 13 were restricted to area A, alone or in combination with one of the other areas. It is a well-known

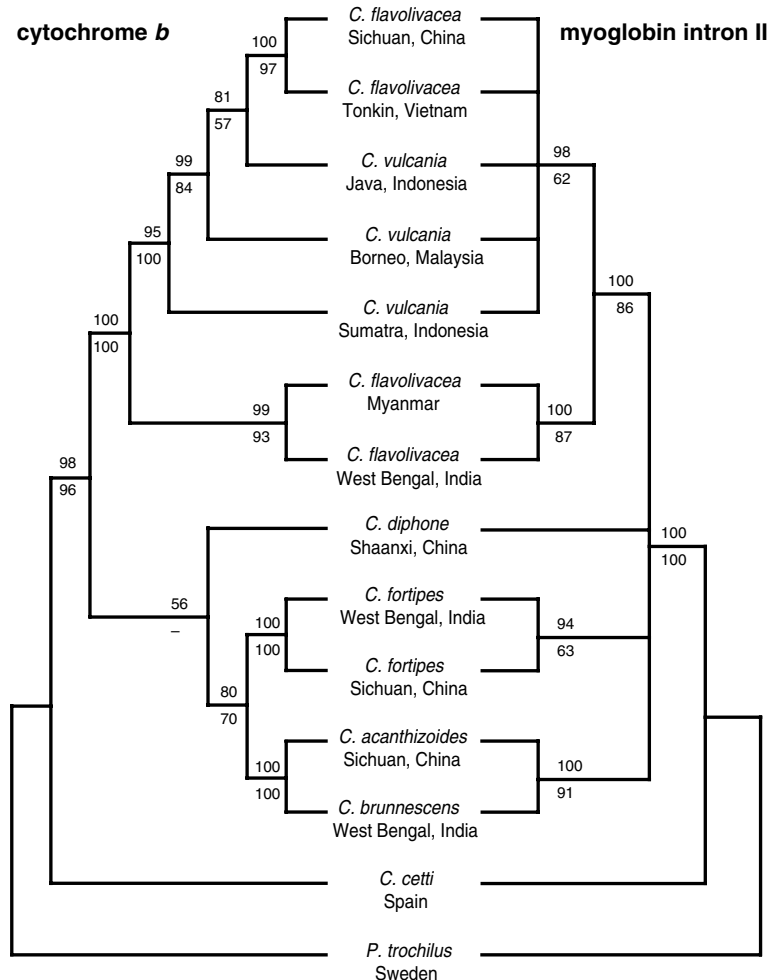


Fig. 2. Trees based on the independent markers Cytochrome *b* (1038 bp) and myoglobin intron II (717 bp), analyzed separately by Bayesian inference under the GTR + Γ + I and HKY models, respectively. Posterior probabilities ($\geq 50\%$; 141,000 trees) and parsimony bootstrap support are indicated above and below, respectively, the nodes.

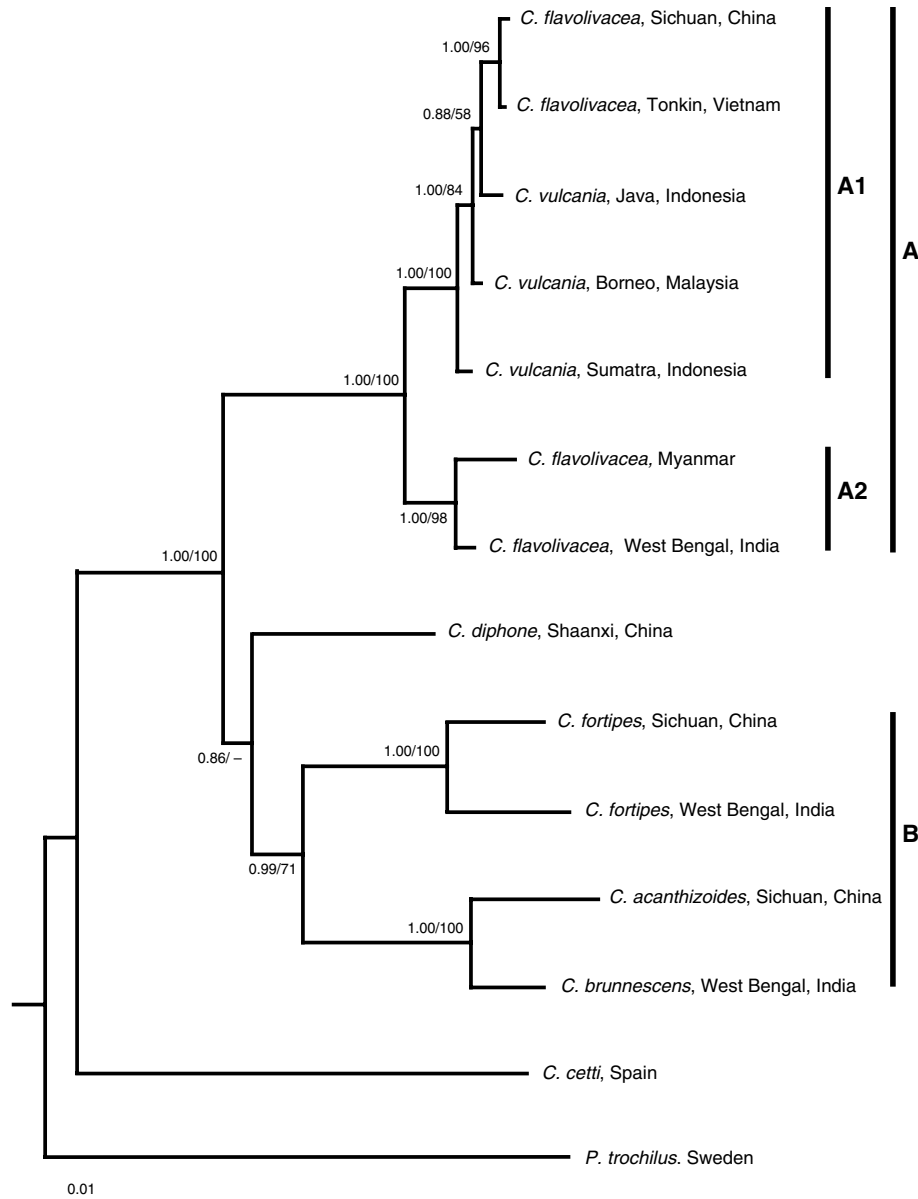


Fig. 3. Phylogram based on combined Cytochrome *b* and myoglobin intron II sequences (1755 bp) divided into two partitions estimated by Bayesian Inference under the GTR + Γ + I and HKY models, respectively. Posterior probabilities ($\geq 50\%$; 141,000 trees; left) and parsimony bootstrap support (right) are indicated next to the nodes. Vertical bars and letters denote clades that are discussed in the text.

drawback of DIVA that there is a tendency for ancestral area optimizations to become less reliable towards the root (Barber and Bellwood, 2005).

In the ASR, the probability for each ancestral state expected by chance is 0.2. We divided the results into three categories: (1) those with lower limit of 95% credibility interval above 0.2 are regarded as supported as ancestral areas; (2) those which include 0.2 within their 95% credibility interval do not deviate from a result expected by chance, and are regarded as neither supporting nor rejecting the area as ancestral; (3) for those with upper limit of 95% credibility interval below 0.2, the areas are regarded as rejected as ancestral. For Clade A (Fig. 3), the ASR (Table 2) is congruent with the DIVA results, based on the all compatible 50% consensus tree of the concatenated sequences, suggest-

ing that the Indonesian taxa originated from a dispersal from the Himalayas, and that Myanmar was colonized later, also from the Himalayas. The probabilities for the ancestral states of nodes 1 and 2 are inconclusive, with no area supported by the lower limit of 95% credibility interval being above 0.2. However, the highest probabilities are compatible with the DIVA results.

4. Discussion

4.1. Phylogeography

The most probable phylogeographic history of the *C. vulcania/flavolivacea* clade, supported by both DIVA and ASR, is that the Greater Sundas were colonized from the

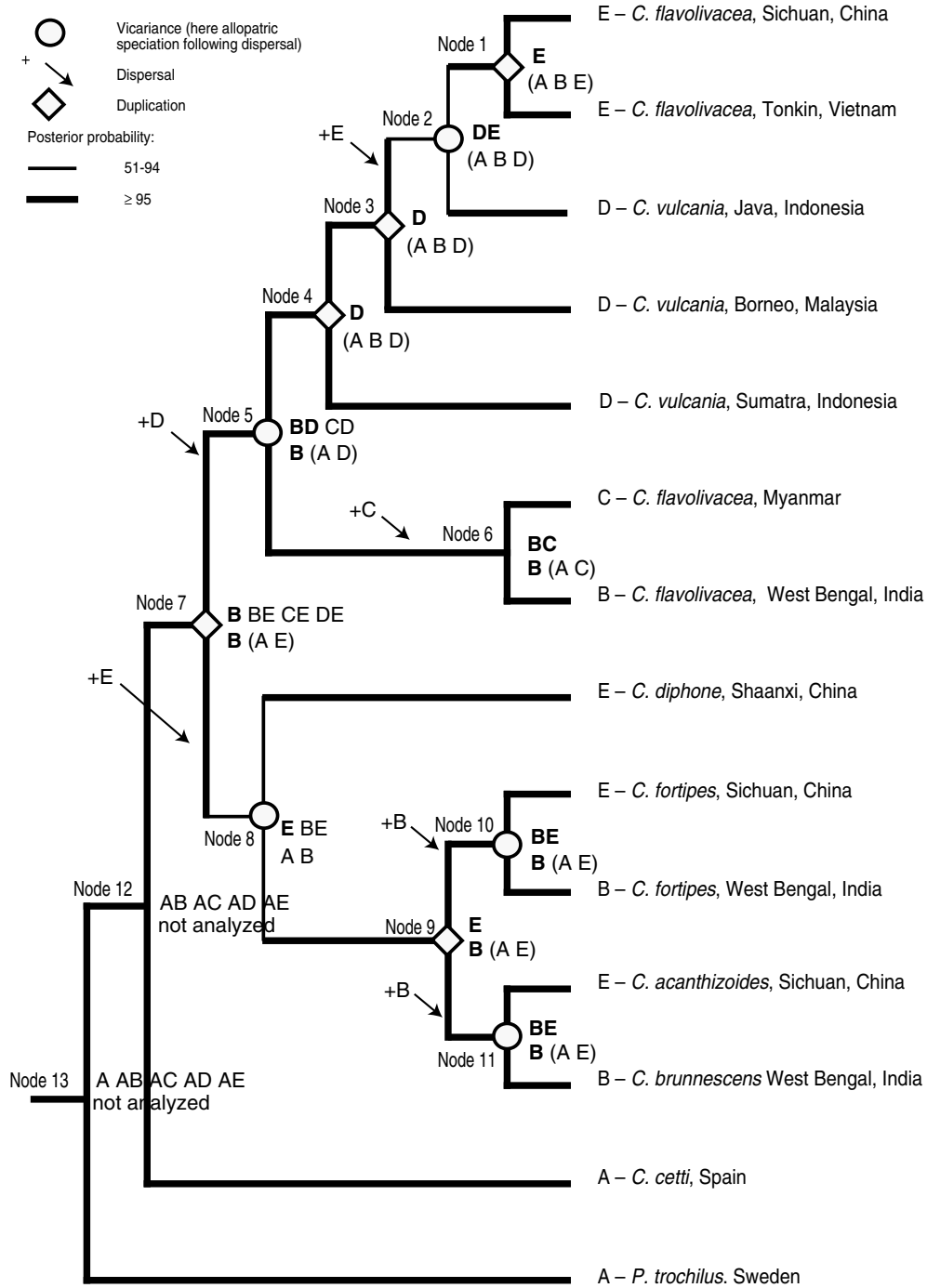


Fig. 4. Summary of the optimal reconstructions of ancestral distributions of the montane bush-warblers of southern Asia and Indonesia west of the Wallace Line based on dispersal-vicariance analysis (DIVA) and Bayesian ancestral state reconstruction (ASR). The phylogeny is the all compatible 50% consensus tree based on the Bayesian Inference of the concatenated Cytochrome *b* and myoglobin intron II sequences. The optimal distributions are given to the right of each node, those based on the exact DIVA search with maximum unit areas restricted to 2 above, and those based on the ASR below. The DIVA reconstruction required 7 dispersal events, which are indicated by the colonized area given as “+[area □]”. Vicariance events are indicated by a circle; duplication events (speciation within a defined area) are indicated by rhombs. No extinction events are inferred. The probabilities of the ancestral states calculated by the ASR with lower limit of 95% credibility interval above 0.2 are indicated in bold. Those which include 0.2 within their 95% credibility interval are given in parentheses. Rejected alternatives with upper limit of 95% credibility interval below 0.2 are not shown.

Himalayas in a first dispersal event, and that Myanmar, in spite of being geographically closer to the Himalayas, was colonized at a later stage (Fig. 5). The topology of the phylogram (Fig. 3) suggests an initial period of isolation of the

Indonesian founder population, followed by rapid radiation, including a dispersal back to mainland Asia. The DIVA indicates that a series of duplication events took place within Indonesia (area D, Fig. 5). It is highly unlikely

Table 2
Statistics from Bayesian ancestral states reconstruction

Ancestral state	Mean	Variance	95% Cred. Interval		Median
			Lower	Upper	
Node 1					
Palaearctic	0.212744	0.013023	0.000070	0.312490	0.270160
Himalayas	0.212744	0.013023	0.000070	0.312490	0.270160
Myanmar	0.085097	0.002084	<i>0.000028</i>	<i>0.124996</i>	0.108064
Indonesia	0.092341	0.003318	<i>0.001332</i>	<i>0.133260</i>	0.110459
China	0.397074	0.099487	0.125000	0.998119	0.236765
Node 2					
Palaearctic	0.291235	0.003770	0.038472	0.312500	0.312475
Himalayas	0.295250	0.003665	0.050519	0.312500	0.312487
Myanmar	0.116564	0.000596	<i>0.016437</i>	<i>0.125000</i>	0.124990
Indonesia	0.176015	0.025734	0.124997	0.844430	0.125034
China	0.120936	0.001273	<i>0.029278</i>	<i>0.132167</i>	0.124997
Node 3					
Palaearctic	0.226857	0.016048	0.000019	0.312500	0.310011
Himalayas	0.226862	0.016045	0.000020	0.312500	0.310011
Myanmar	0.090743	0.002568	<i>0.000008</i>	<i>0.125000</i>	0.124004
Indonesia	0.360396	0.122938	0.125000	0.999826	0.131500
China	0.095143	0.002682	<i>0.000106</i>	<i>0.125005</i>	0.124305
Node 4					
Palaearctic	0.297051	0.001971	0.142984	0.312500	0.311970
Himalayas	0.297052	0.001970	0.142984	0.312500	0.311970
Myanmar	0.118820	0.000315	<i>0.057194</i>	<i>0.125000</i>	0.124788
Indonesia	0.168077	0.015336	0.125000	0.596665	0.126475
China	0.119000	0.000306	<i>0.059192</i>	<i>0.125000</i>	0.124796
Node 5					
Palaearctic	0.282055	0.004947	0.059392	0.312500	0.312500
Himalayas	0.328292	0.003745	0.279581	0.529492	0.312500
Myanmar	0.116926	0.000424	<i>0.047880</i>	<i>0.125000</i>	0.125000
Indonesia	0.155528	0.007393	0.125000	0.435401	0.125000
China	0.117198	0.000742	<i>0.036751</i>	<i>0.125000</i>	0.125000
Node 6					
Palaearctic	0.192823	0.015473	0.008833	0.312500	0.260397
Himalayas	0.542226	0.058311	0.312500	0.943547	0.422129
Myanmar	0.109628	0.001712	0.027503	0.196899	0.124244
Indonesia	0.078059	0.002387	<i>0.005718</i>	<i>0.125000</i>	0.104216
China	0.077264	0.002457	<i>0.004305</i>	<i>0.125000</i>	0.104159
Node 7					
Palaearctic	0.295276	0.002618	0.100010	0.312500	0.312500
Himalayas	0.321273	0.001311	0.310232	0.432320	0.312500
Myanmar	0.118620	0.000372	<i>0.046675</i>	<i>0.125000</i>	0.125000
Indonesia	0.122459	0.000196	<i>0.077509</i>	<i>0.133406</i>	0.125000
China	0.142372	0.003709	0.125000	0.338206	0.125000
Node 8					
Palaearctic	0.311252	0.000169	0.312179	0.312500	0.312500
Himalayas	0.313045	0.000069	0.312500	0.312500	0.312500
Myanmar	0.124543	0.000024	0.124891	0.125000	0.125000
Indonesia	0.124848	0.000013	0.125000	0.125000	0.125000
China	0.126311	0.000237	0.125000	0.125162	0.125000
Node 9					
Palaearctic	0.301863	0.001725	0.144939	0.312500	0.312500
Himalayas	0.318071	0.000885	0.312051	0.394828	0.312500
Myanmar	0.121044	0.000246	<i>0.063354</i>	<i>0.125000</i>	0.125000
Indonesia	0.123554	0.000130	<i>0.095798</i>	<i>0.129135</i>	0.125000
China	0.135469	0.002234	0.125000	0.274286	0.125000
Node 10					
Palaearctic	0.267671	0.006070	0.035538	0.312500	0.312500
Himalayas	0.344214	0.005613	0.287260	0.576086	0.312500
Myanmar	0.107180	0.000959	<i>0.014863</i>	<i>0.125000</i>	0.125000

Table 2 (continued)

Ancestral state	Mean	Variance	95% Cred. Interval		Median
			Lower	Upper	
Indonesia	0.108000	0.000884	<i>0.018427</i>	<i>0.125000</i>	0.125000
China	0.172936	0.010231	0.124964	0.492347	0.125000
Node 11					
Palaearctic	0.287686	0.003958	0.067502	0.312500	0.312500
Himalayas	0.342767	0.006445	0.312500	0.616529	0.312500
Myanmar	0.115127	0.000627	<i>0.027763</i>	<i>0.125000</i>	0.125000
Indonesia	0.115528	0.000586	<i>0.030735</i>	<i>0.125000</i>	0.125000
China	0.138893	0.003815	0.111725	0.310529	0.125000

Supported ancestral areas (with lower limit of 95% credibility interval above 0.2) indicated in bold. Rejected ancestral areas (with upper limit of 95% credibility interval below 0.2) indicated in italics.

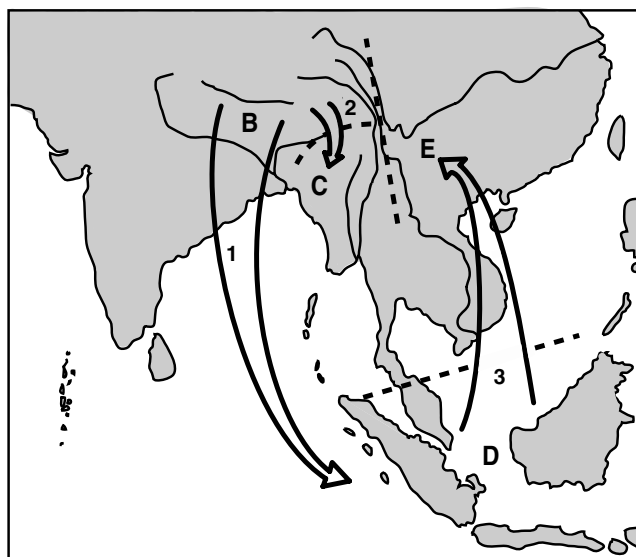


Fig. 5. Map visualizing the inferred major dispersal events in the *Cettia vulcanialflavolivacea* group. Dashed lines show borders between unit areas. Arrows indicate approximate start and end points of dispersals, but do not aspire to show the exact route. Area A is not indicated as it is mostly outside the map.

that these were true sympatric speciations, but from the available data it is not possible to determine whether they resulted from dispersals within Indonesia, or vicariance as a result of previously continuous land masses being separated by post-glaciation rises in sea level.

The phylogram (Fig. 3) indicates minor differences between the Chinese and Vietnamese populations, and both these are slightly divergent from the Indonesian populations. We interpret this as the result of rapid recent dispersal over area E (Fig. 5), probably from a source population that occupied a restricted range for some time before expanding the range.

Outside clade A, taxon sampling is less complete, and probably affects the reliability of the results of the DIVA and ASR in clade B negatively. Furthermore, the position of *C. diphone* is not well supported. Alternative positions, such as sister to clade A, or sister to both clades A and B (Fig. 3) results in contradictory DIVA results in clade B,

and increased uncertainty in the basal nodes. We thus consider the results concerning the phylogeographic history of *C. diphone* and clade B (Fig. 3) unreliable, based on the present analyses.

4.2. Taxonomic implications

The previous treatment of *C. vulcania* and *C. fortipes* as conspecific (Vaurie, 1959) or forming a superspecies (Watson et al., 1986) are not consistent with the phylogeny presented here. Instead, a close relationship between *C. vulcania* and *C. flavolivacea* is strongly supported (Figs. 1, 2).

The taxonomic treatment of the taxa in clade A (Fig. 3) at the species level is to some extent arbitrary, and dependent on the choice of species definition. It is consistent with the phylogeny to consider either one (*C. flavolivacea*, by priority; International Commission on Zoological Nomenclature, 1999), two (*C. flavolivacea* and *C. vulcania*—corresponding to the two main clades) or several species. We prefer the second alternative, and propose that the Chinese *intricata* and Vietnamese *oblita*, currently allocated to *C. flavolivacea*, are treated as conspecific with the Sunda populations, under the name *C. vulcania* (Table 1). *C. flavolivacea* is thus restricted to the Himalayan and Myanmar populations.

Appendix A

List of samples (in alphabetical order), with geographic origin and GenBank accession numbers

Taxon	Locality	Museum No.	Region	GenBank No.
<i>Cettia acanthizoides</i>	Sichuan, China	NRM-20056599	Cyt <i>b</i>	DQ673886
			Myo	DQ673899
<i>Cettia brunnescens</i>	West Bengal, India	NRM-20056602	Cyt <i>b</i>	DQ673887
			Myo	DQ673900
<i>Cettia cetti cetti</i>	Spain	NRM-20066034	Cyt <i>b</i>	DQ673876
			Myo	DQ673889
<i>Cettia diphone canturians</i>	Shaanxi, China	NRM-20046811	Cyt <i>b</i>	DQ008510
			Myo	DQ008562
<i>Cettia flavolivacea intricata</i>	Sichuan, China	NRM-20066035	Cyt <i>b</i>	DQ673880
			Myo	DQ673893
<i>Cettia flavolivacea oblita</i>	Tonkin, Vietnam	NRM-20066036	Cyt <i>b</i>	DQ673881
			Myo	DQ673894
<i>Cettia flavolivacea weberi</i>	Chin Hills, Myanmar	NRM-20066037	Cyt <i>b</i>	DQ673882
			Myo	DQ673895
<i>Cettia flavolivacea flavolivacea</i>	West Bengal, India	NRM-20066038	Cyt <i>b</i>	DQ673883
			Myo	DQ673896
<i>Cettia fortipes fortipes</i>	West Bengal, India	NRM-20066039	Cyt <i>b</i>	DQ673884
			Myo	DQ673897
<i>Cettia fortipes davidiana</i>	Sichuan, China	NRM-20066040	Cyt <i>b</i>	DQ673885
			Myo	DQ673898
<i>Cettia vulcania vulcania</i>	Java, Indonesia	NRM-20066041	Cyt <i>b</i>	DQ673878
			Myo	DQ673891

(continued on next page)

The genetic divergence between populations from north Vietnam (*oblita*) and central China (*intricata*) is slight. Normal within-population variation in the Cytochrome *b* gene in other Old World warblers is below 0.5% (e.g. Drovetski et al., 2004; Helbig et al., 1996; Martens et al., 2004; Olsson et al., 2004, 2005). Divergence in excess of 0.5% usually coincides with geographic separation, and thus possibly reduced gene flow (Olsson et al., 2005). In this case the divergence between *oblita* and *intricata* is 0.3% or less, which indicates that there is either no reduction in gene flow, or that such reduction is of very recent origin. Consequently, the taxonomic recognition of *oblita* is not supported by the genetic evidence presented here. However, we have not evaluated the morphological differences that lead to the original recognition of these taxa, and “if the samples are clearly different in one or several respects” (Mayr, 1969, p. 188) taxonomic recognition may still be justified from a morphological point of view.

Acknowledgments

We are grateful to Johan Nylander for valuable help with the ASR, and to Juan Carlos Atienza for providing the sample of *Cettia cetti*. This study was financially supported by the Swedish Research Council (Grants Nos. 621-2001-2773 to P.G.P.E. and 621-2002-3952 to P.S.).

Appendix A (continued)

Taxon	Locality	Museum No.	Region	GenBank No.
<i>Cettia vulcania flaviventris</i>	Sumatra, Indonesia	NRM-20066042	Cyt <i>b</i> Myo	DQ673877 DQ673890
<i>Cettia vulcania oreophila</i>	Borneo, Malaysia	NRM-20066043	Cyt <i>b</i> Myo	DQ673879 DQ673892
<i>Phylloscopus trochilus</i> <i>trochilus</i>	Sweden	NRM-20066044	Cyt <i>b</i> Myo	DQ673875 DQ673888

NRM = Swedish Museum of Natural History, Stockholm, Sweden. Taxonomy follows Watson et al. (1986), except for *Cettia brunnescens* (Alström et al., in press).

References

- Akaike, H., 1973. Information theory as an extension of the maximum likelihood principle. In: Petrov, B.N., Csaki, F. (Eds.), Second International Symposium on Information Theory. Akademiai Kiado, Budapest.
- 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., Olsson, U., Rasmussen, P.C., Yao, C.-t., Ericson, P.G.P., Sundberg, P., in press. Morphological, vocal and genetic divergence in the *Cettia acanthizoides* complex (Aves: Cettiidae). *Zool. J. Linn. Soc.*
- Arctander, P., 1995. Comparison of a mitochondrial gene and a corresponding nuclear pseudogene. *Proc. R. Soc. Lond. B* 262, 13–19.
- Barber, P.H., Bellwood, D.R., 2005. Biodiversity hotspots: evolutionary origins of biodiversity in wrasses (Halichoeres: Labridae) in the Indo-Pacific and new world tropics. *Mol. Phylogenet. Evol.* 35, 235–253.
- Desjardins, P., Morais, R., 1990. Sequence and gene organization of the chicken mitochondrial genome a novel gene order in higher vertebrates. *J. Mol. Biol.* 212, 599–634.
- Dickinson, E.C. (Ed.), 2003. The Howard & Moore Complete Checklist of the Birds of the World. Christopher Helm, London.
- 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.
- 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.
- Hasegawa, M., Kishino, H., Yano, T., 1985. Dating the human-ape split by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* 22, 160–174.
- Helbig, A.J., Martens, J., Seibold, I., Henning, F., Schottler, B., Wink, M., 1996. Phylogeny and species limits in the Palearctic chiffchaff *Phylloscopus collybita* complex: mitochondrial genetic differentiation and bioacoustic evidence. *Ibis* 138, 650–666.
- Higgins, D.G., Bleasby, A.J., Fuchs, R., 1992. Clustal V: Improved software for multiple sequence alignment. *CABIOS* 8, 189–191.
- Huelsenbeck, J.P., Bollback, J.P., 2001. Empirical and hierarchical Bayesian estimation of ancestral states. *Syst. Biol.* 50, 351–366.
- Huelsenbeck, J.P., Larget, B., Miller, R.E., Ronquist, F., 2002. Potential applications and pitfalls of Bayesian inference of phylogeny. *Syst. Biol.* 51, 673–688.
- Huelsenbeck, J.P., Ronquist, F., 2001. MrBayes: Bayesian inference of phylogeny. *Bioinformatics* 17, 754–755.
- Huelsenbeck, J.P., Ronquist, F., 2005. MrBayes: a program for the Bayesian inference of phylogeny, Version 3.1. Available from: <<http://morphbank.abc.uu.se/mrbayes/>>.
- Inskipp, T., Lindsey, N., Duckworth, W., 1996. An Annotated Checklist of the Birds of the Oriental Region. Oriental Bird Club, Sandy.
- International Chicken Genome Sequencing Consortium, 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* 432, 695–716.
- International Commission on Zoological Nomenclature, 1999. International Code of Zoological Nomenclature, fourth ed., London.
- Kimura, M., 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16, 111–120.
- Lanave, C., Preparata, C., Saccone, C., Serio, G., 1984. A new method for calculating evolutionary substitution rates. *J. Mol. Evol.* 20, 86–93.
- Martens, J., Tietze, D.T., Eck, S., Veith, M., 2004. Radiation and species limits in the Asian Pallas's warbler complex (*Phylloscopus proregulus* s.l.). *J. Orn.* 145, 206–222.
- Mayr, E., 1969. Principles of Systematic Zoology. McGraw-Hill, New York.
- Nylander, J.A.A., 2004. MrModeltest2. Available from: <<http://www.abc.uu.se/systzoo/staff/nylander.html/>>.
- Nylander, J.A.A., Ronquist, F., Huelsenbeck, J.P., Nieves-Aldrey, J.L., 2004. Bayesian phylogenetic analysis of combined data. *Syst. Biol.* 53, 1–21.
- Olsson, U., Alström, P., Sundberg, P., 2004. Non-monophyly of the avian genus *Seicercus* (Aves, Sylviidae) revealed by mitochondrial DNA. *Zoologica Scripta* 33, 501–510.
- 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.
- Quinn, T.W., 1997. Molecular evolution of the mitochondrial genome. In: Mindell, D.P. (Ed.), Avian Molecular Evolution and Systematics. Academic Press, London, San Diego, pp. 3–28.
- Quinn, T.W., White, B.N., 1987. Analysis of DNA sequence variation. In: Cooke, F., Buckley, P.A. (Eds.), Avian Genetics: a Population and Ecological Approach. Academic Press, London, pp. 163–198.
- Rasmussen, P.C., Anderton, J.C., 2005. Birds of South Asia: the Ripley guide. Lynx Edicions, Barcelona.
- Rodríguez, J., Oliver, L., Marín, A., Medina, R., 1990. The general stochastic model of nucleotide substitution. *J. Theor. Biol.* 142, 485–501.
- Ronquist, F., 1996. DIVA, Version 1.1. Computer program and manual available by anonymous FTP from Uppsala University (<ftp.uu.se> or <ftp.systbot.uu.se>).
- Ronquist, F., 1997. Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. *Syst. Biol.* 46, 195–203.
- Ronquist, F., 2004. Bayesian inference of character evolution. *TREE* 19, 475–481.
- Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572–1574.
- Sanmartín, I., 2003. Dispersal vs. vicariance in the Mediterranean: historical biogeography of the Palearctic Pachydeminae (Coleoptera, Scarabaeoidea). *J. Biogeog.* 30, 1883–1897.
- Sorensen, M.D., Quinn, T.W., 1998. Numts: a challenge for avian systematics and population biology. *Auk* 115, 214–221.
- Swofford, D.L., 2001. PAUP*: Phylogenetic Analysis Using Parsimony (* and other methods), Version 4.08b. Sinauer Associates, Sunderland, MA.
- Tavaré, S., 1986. Some probabilistic and statistical problems on the analysis of DNA sequences. *Lec. Math. Life Sci.* 17, 57–86.

- Vaurie, C., 1959. The Birds of the Palearctic Fauna, Order Passeriformes. H.F. and G. Witherby, London.
- Violani, C., 1980. What is *Brachypteryx flaviventris* Salvadori? Bull. Brit. Orn. Cl. 100, 187–189.
- Watson, G.E., Traylor, M.A., Mayr, E., 1986. Family Sylviidae, Old World warblers. In: Mayr, E., Cottrell, G.W. (Eds.), Checklist to the birds of the World, Vol. XI. Museum of Comparative Zoology, Cambridge, MA.
- Wells, D.R., 1982. Biological species limits in the *Cettia fortipes* complex. Bull. Brit. Orn. Cl. 102, 57–62.
- White, C.M.N., Bruce, M.D., 1986. The birds of Wallacea (Sulawesi, the Moluccas & Lesser Sunda Islands, Indonesia). British Ornithologists' Union, London.
- Yang, Z., 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. J. Mol. Evol. 39, 306–314.