
Non-monophyly of the avian genus *Seicercus* (Aves: Sylviidae) revealed by mitochondrial DNA

URBAN OLSSON, PER ALSTRÖM & PER SUNDBERG

Accepted: 22 October 2003

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.

The phylogeny of all species and nearly all subspecies of *Seicercus* and representatives of all subgenera in *Phylloscopus* was estimated based on two mitochondrial genes. According to the gene tree, and supported by non-molecular data, *Seicercus* belongs in three separate clades. Two of these include only taxa currently classified as *Seicercus*, while the third comprises *S. xanthoschistos* and *P. occipitalis*. These results suggest that both *Seicercus* and *Phylloscopus* are paraphyletic. The gene tree suggests two more cases of non-monophyly: (1) the ‘*S. burkii* complex’ is separated into two different clades, one of which also includes *S. affinis* and *S. poliogenys*; (2) two populations of *S. affinis intermedius* are more closely related to *S. affinis ocularis* than to a third population of *intermedius*. A recent proposal to split the ‘*S. burkii* complex’ into six species is corroborated, as is the recognition of the taxon *cognitus* as a colour morph of *S. affinis intermedius*. Our study also revealed unexpectedly large genetic divergences between three different populations of the monotypic *S. poliogenys*, indicating the presence of cryptic species. Our results underscore the importance of dense sampling at the specific and infraspecific levels in intrageneric phylogenetic studies.

Urban Olsson, Department of Zoology, Göteborg University, Box 463, SE-405 30 Göteborg, Sweden. E-mail: urban.olsson@zool.gu.se

Per Alström, Department of Systematic Zoology, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18 D, SE-752 36 Uppsala, Sweden. E-mail: per.alstrom@ebc.uu.se

Per Sundberg, Department of Zoology, Göteborg University, Box 463, SE-405 30 Göteborg, Sweden. E-mail: per.sundberg@zool.gu.se

Introduction

The two genera *Seicercus* and *Phylloscopus* (Aves, Sylviidae) are traditionally believed to be closely related (e.g. Ticehurst 1938; Watson *et al.* 1986; Sibley & Ahlquist 1990). *Seicercus* is usually considered to comprise seven species: *S. burkii*, *S. poliogenys*, *S. affinis*, *S. castaniceps*, *S. grammiceps*, *S. montis* and *S. xanthoschistos* (e.g. Watson *et al.* 1986; Sibley & Monroe 1990; Inskipp *et al.* 1996). Watson *et al.* (1986) divided these into 29 subspecies. Alström & Olsson (1999) and Martens *et al.* (1999) suggested that *S. burkii* is a complex of sibling species, and described one new species each. Alström & Olsson (2000) concluded that this complex consists of six species (Table 1). Watson *et al.* (1986) recognized 45 species in the genus *Phylloscopus*, but descriptions of new species and taxonomic rearrangements have increased this number to at least 55 species (Irwin *et al.* 2001).

The generic allocation of the species now in *Seicercus* has changed over time, and several species have been moved between *Seicercus* and *Phylloscopus*. The phylogeny of *Phylloscopus* is relatively well known (Richman & Price 1992;

Helbig *et al.* 1995, 1996; Richman 1996), and part of the *S. burkii* complex has been analysed by Martens *et al.* (1999) and Alström & Olsson (2000). However, no previous study has dealt with all of the species in *Seicercus*, and none has tested the monophyly of *Seicercus*.

Here we test the monophyly of *Seicercus*, and the presumed sister group relationship between *Seicercus* and *Phylloscopus*, and also the classification by Alström & Olsson (2000), by presenting hypotheses based on two mitochondrial genes [cytochrome *b* (*cyt b*) and ribosomal 12S]. For some taxa, non-molecular data were also used to estimate the phylogeny. The classification by Martens *et al.* (1999) was evaluated by Alström & Olsson (2000) and is not discussed any further here.

Materials and methods

Twenty-nine populations of *Seicercus* were sampled from 16 different localities (Appendix), including representatives of all groups in the *S. burkii* complex identified by Alström & Olsson (1999). This represents all species and nearly all

Table 1 Taxonomy of the *Seicercus burkii* complex following Alström & Olsson (2000), with comments based on the present study.

Species	Subspecies	Comments
<i>burkii</i> (Burton, 1836)	Monotypic	Distinctness and species status confirmed.
<i>tephrocephalus</i> (Anderson, 1871)	Monotypic	Distinctness and species status confirmed.
<i>omeiensis</i> Martens, Eck, Päckert & Sun, 1999	Monotypic	Distinctness and species status confirmed.
<i>soror</i> Alström & Olsson, 1999	Monotypic	Distinctness and species status confirmed.
<i>whistleri</i> Ticehurst, 1925	<i>whistleri</i> (Ticehurst, 1925) <i>nemoralis</i> (Koelz, 1954)	Distinctness and species status confirmed. Unlike Alström & Olsson (2000), taxon <i>nemoralis</i> restricted to west Myanmar and adjacent parts of India, <i>whistleri</i> throughout Himalayas.
<i>valentini</i> (Hartert, 1907)	<i>valentini</i> (Hartert, 1907) <i>latouchei</i> (Bangs, 1929)	Distinctness and species status confirmed. Status of <i>latouchei</i> uncertain.

subspecies currently recognized in the genus *Seicercus*. Seventeen *Phylloscopus* were included, representing the eight subgenera in the genus as defined by Watson *et al.* (1986). *Sylvia atricapilla atricapilla* and *Acrocephalus scirpaceus scirpaceus* were used as outgroups; *Acrocephalus* is believed to be one of the closest relatives of *Phylloscopus*, while *Sylvia* is somewhat further away (Sibley & Ahlquist 1990).

DNA was obtained either from blood samples or feathers collected from live birds (Appendix). Blood samples of 5–50 µL were stored in buffer (10% EDTA, 1% NaF, 1% Tris pH 7.4, 1% thymol). Two extraction protocols were followed for blood samples. Either a slightly modified standard phenol:chloroform extraction or by using the QIAgen Blood Kit, according to the manufacturer's recommendations. Feathers were extracted with either the QIAamp Mini Kit or the QIAamp DNEasy Kit, following the manufacturer's recommendations, with the exception that 30 µL of 0.1% DTT was added to the first incubation step in order to dissolve the feathers and increase the DNA yield.

The primers used to amplify the *cytb* gene were L-14995 and H-16065 according to Helbig *et al.* (1996). The primers L-1091 and H-1478 were used for the 12S gene. In one case, the primers L-14995 and H-1478 were combined. The polymerase chain reaction (PCR) cycling parameters for double-stranded amplification were: 2 min 30 s at 95 °C; 60 cycles of 30 s at 95 °C; 30 s at 45–58 °C; 2 min at 72 °C; 7 min at 72 °C. Reactions were carried out in a PTC-100 programmable thermal controller (MJ Research, Inc.). The amplified products were purified using the QIAquick PCR Purification Kit (QIAGEN). Sequencing was performed using Cy5-labelled primers on an ALF-Express automatic sequencer (Pharmacia). Sequencing reactions for ALF-Express were performed using the ThermoSequenase Sequencing Kit (Amersham) with a two-step cycle: 2 min at 96 °C; 20 cycles of 30 s at 95 °C; 40 s at respective annealing/extension temperature. The *cytb* sequencing primers were: H-15557 (Richman & Price 1992), H-15298, L-15320 and L-15722 (Helbig *et al.* 1995). For each individual, most of

both strands were sequenced twice, from different sequence reactions.

The sequences were easily aligned using MEGALIGN 4.03 in the DNASTAR package (DNASTAR Inc); some manual adjustment was necessary among the 12S sequences. Molecular phylogenies were estimated by parsimony and maximum likelihood using PAUP* (Swofford 2001) and by Bayesian inference using MRBAYES (Huelsenbeck & Ronquist 2001a,b). The sequences were analysed separately and combined. Clade support for the unweighted data set was assessed by parsimony bootstrapping, heuristic search strategy, 10 000 replicates, starting trees obtained by stepwise addition (random addition sequence, 10 replicates, 10 trees held at each replicate), tree bisection reconnection (TBR) branch swapping; gaps were coded as missing data or as a fifth base in different searches.

Posterior probability (clade credibility) values were calculated under a general time-reversible model with gamma-distributed rates over sites and an estimated proportion of invariant sites (GTR + Γ + I), and a sequencing error set to 1%. The *cytb* sequences were also analysed under a model with site-specific gamma-distributed rates over sites (GTR + Γ + SS). Four Metropolis-coupled MCMC chains with incremental heating temperature 0.2 were run for 1 000 000 generations and sampled every 100 generations. The first 70 000–90 000 generations (depending on the data set), before the chain reached apparent stationarity, were discarded, and the posterior probability estimated for the remaining generations. Every analysis was repeated five times, starting from random trees, and the results compared to ascertain that the chains had reached stationarity.

A maximum likelihood analysis was performed iteratively. A starting tree was generated by heuristic search under the parsimony criterion (using PAUP* default settings). On this tree the parameters were estimated under the GTR + Γ + I model. The values of the parameters were then fixed, and the tree was used as the starting tree in a TBR search for a maximum likelihood tree. The parameters were re-optimized on the resulting maximum likelihood tree and the procedure was

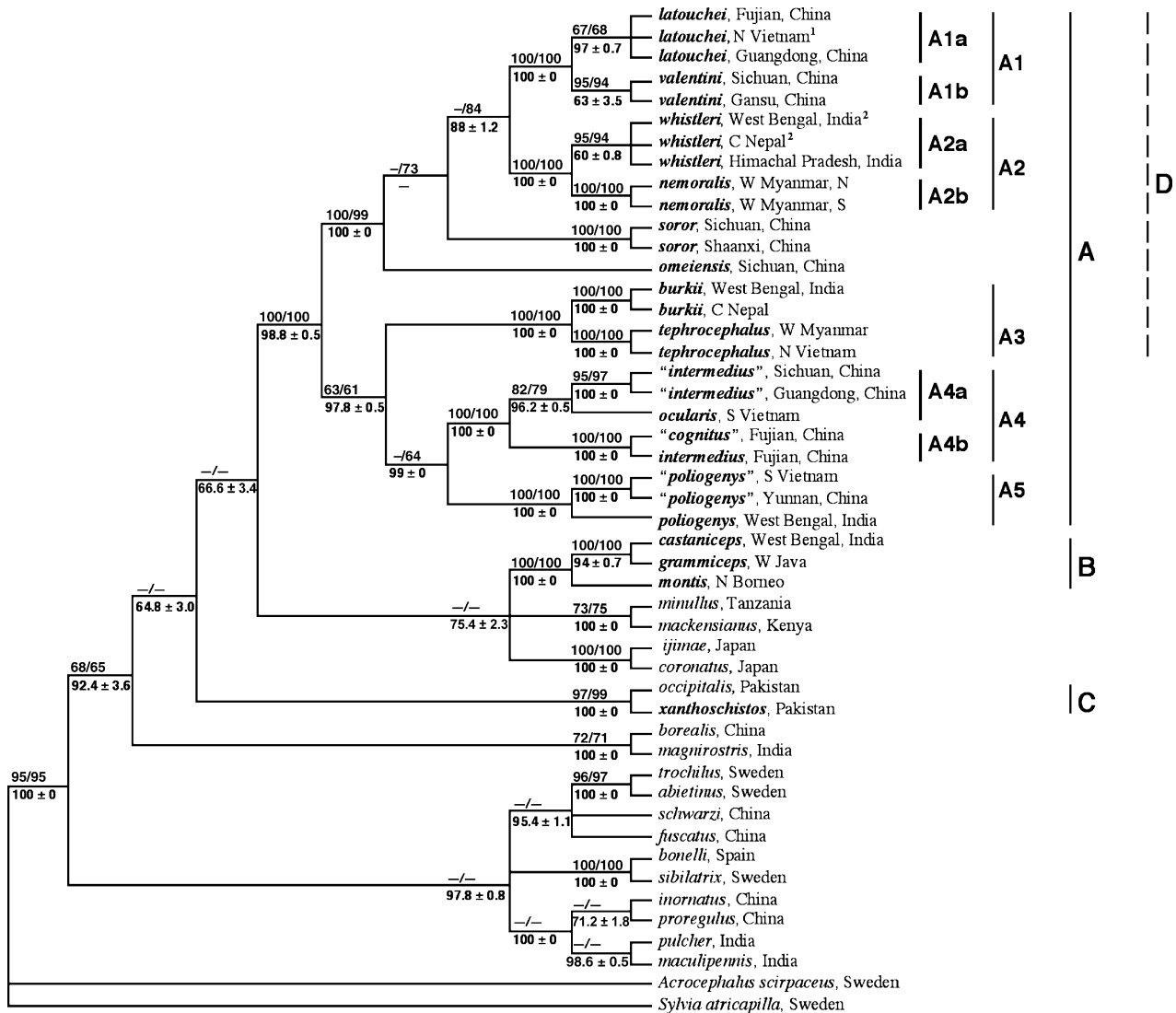


Fig. 1 Tree of *Seicercus* and representatives of the eight subgenera in *Phylloscopus* [as defined by Watson *et al.* (1986)], with *Acrocephalus* and *Sylvia* as outgroups. Posterior probability values ($\geq 60\%$; 46 000 trees) resulting from the analysis of the concatenated cytochrome *b* and 12S sequences under a GTR + Γ + I model are shown below the nodes (mean of five analyses \pm standard deviation). Bootstrap values ($\geq 50\%$; 10 000 replicates) are shown above the nodes; values to the left are based on cytochrome *b* and 12S sequences and values to the right are based on sequence data combined with non-molecular characters. The taxa traditionally placed in *Seicercus* are shown in bold and represented by the vertical bars marked A, B and C (excluding *occipitalis*). The dashed vertical bar marked D represents the '*S. burkii*' complex. ¹Treated as *valentini* by Alström & Olsson 1999. ²Treated as *nemoralis* by Alström & Olsson 1999.

repeated until the topology did not change. The choice of model for the model-based analyses was determined by a likelihood ratio test (MODELTEST 3.06, Posada & Crandall 1998) for the maximum likelihood analyses and MRMODELTEST 1.0b (Nylander 2002) for the Bayesian inference analyses.

Pairwise sequence divergences were calculated in PAUP* using maximum likelihood distances corresponding to the GTR + Γ + I model (the shape parameter and the estimated proportion of invariable sites were obtained from a Bayesian

analysis, as PAUP* cannot estimate these parameters under the distance criterion).

The taxa in clade A in Fig. 1 were also investigated by non-molecular data. Nine morphological, vocal and ecological characters were selected from Alström & Olsson (1999) and, for *S. affinis* spp. and *S. poliogenys*, from our own unpublished studies (Table 2). These were analysed by parsimony bootstrapping, both together with the molecular data in a 'total evidence' analysis and separately. The bootstrap analysis was performed in the same way as for the molecular data (see above).

Table 2 Data matrix of morphological, vocal and ecological characters.

	1	2	3	4	5	6	7	8	9
<i>Seicercus affinis ocularis</i>	0	1	1	0	1	0	0	1	0
<i>Seicercus affinis intermedius</i>	0	1	1	0	1	0	0	1	0
<i>Seicercus poliogenys</i>	0	2	1	0	1	0	0	1	0
<i>Seicercus tephrocephalus</i>	1	1	0	1	0	0	0	1	0
<i>Seicercus burkii</i>	1	1	0	1	0	0	0	1	0
<i>Seicercus omeiensis</i>	1	0/1	0	0	0	0	0	1	0
<i>Seicercus soror</i>	1	0	0	0	0	1	0	0	0
<i>Seicercus w. whistleri/nemorialis</i>	0	0	0	0	1	1	1	0	1
<i>Seicercus v. valentini/latouchei</i>	0	0	0	0	1	1	1	0	1

1, Pale wing-bar: (0) usually distinct; (1) usually indistinct or absent. 2, Lateral crown-strips: (0) indistinct anteriorly, distinct posteriorly; (1) distinct throughout; (2) relatively indistinct throughout. 3, Eye-ring broken above eye: (0) no; (1) yes. 4, Eye-ring broken at rear: (0) no; (1) yes. 5, White on outer web of outermost tail feather: (0) usually none or very little; (1) usually distinct. 6, Presence of constant introductory element in song: (0) no; (1) yes. 7, Unique introductory element in song: (0) no; (1) yes. 8, Song has rattles (0) no; (1) yes. 9, Breeds rather near treeline/summits: (0) no; (1) yes.

Results

We obtained up to 1038 bp of the *cytb* gene and 395–396 bp of the 12S gene from 29 populations of *Seicercus* and 17 species of *Phylloscopus*. Slowly evolving nuclear pseudogenes of mitochondrial origin (Numts) are known from birds (e.g. Quinn & White 1987; Arctander 1995; Quinn 1997; Allende *et al.* 2001; Klitgaard Nielsen & Arctander 2001). As Numts are under relaxed evolutionary constraint, stochastic substitutions causing stop codons, frameshift mutations in protein-coding sequences, or unexpected base substitutions leading to the replacement of one amino acid for another with different properties (e.g. a change from a polar to a non-polar amino acid) in otherwise conserved regions are expected to occur more frequently than in mitochondrial DNA (e.g. Bensasson *et al.* 2001). We found no stop codons or frameshift mutations, but noted instances of unexpected amino acid changes in *omeiensis* #1 from Hebei, *latouchei* from Guangdong, *intermedius* from Sichuan (two individuals) and Guangdong and all samples of *poliogenys* (especially the one from India), *montis*, *grammiceps* and *castaniceps*. However, most of these sequences were obtained from feathers, which are expected to have a much higher ratio of mitochondrial to nuclear genome copies than blood (Sorenson & Quinn 1998). Moreover, for one of these (*S. poliogenys* from India) we amplified an approximately 3 kbp fragment. The *cytb* sequence obtained from this fragment was identical to the one amplified by the standard method. The base composition, position bias and transition/transversion ratio of the same (Table 3) are in agreement with mitochondrial genes (Vawter & Brown 1986; Arctander 1995; Klitgaard Nielsen & Arctander 2001). [However, Allende *et al.* (2001) found similar base composition in *cytb* and Numts in sparrows, *Passer*.]

Table 3 Values of parameters estimated in the Bayesian analysis (GTR + Γ + SS) of the cytochrome *b* gene for the taxa in Fig. 1.

Parameter	Mean	Variance
r_{GT}	1.00	0.00
r_{CT}	40.56	122.0
r_{CG}	0.59	0.17
r_{AT}	1.74	0.62
r_{AG}	38.36	234.3
r_{AC}	1.17	0.17
π_A	0.34	0.0002
π_C	0.44	0.0002
π_G	0.07	0.00005
π_T	0.15	0.00005
r_1	0.09	0.0003
r_2	0.002	0.000001
r_3	2.91	0.0003
ti/tv	17.5	

Transition/transversion ratio (ti/tv) calculated using the rate parameters. r_{GT} , r_{CT} , r_{CG} , etc. are the estimated rates of change between G and T, C and T, C and G, respectively. r_{GT} is arbitrarily set to 1.00 as a reference. π_A , π_C , π_G and π_T are the estimated stationary frequencies of A, C, G and T, respectively. r_1 , r_2 and r_3 are the estimated average rates of change at first, second and third codon positions, respectively.

There are five positions with indels in the 12S gene, of which two are parsimony informative; all gaps are a single base long. The number of parsimony informative characters is 333 (32.1%) in the *cytb* gene and 46 (11.6%), including indels, in the aligned 12S sequences. The *cytb* and 12S trees (not shown) differ mainly in that the latter is considerably less resolved than the former, especially close to the terminals, as a result of the smaller number of informative sites in the more slowly evolving ribosomal gene than in *cytb*. There is slight conflict between the two trees. The aligned concatenated sequences have 1437 sites, of which 379 (26%) are parsimony informative (including indels). The trees resulting from Bayesian inference (Fig. 1) and maximum likelihood (not shown) have practically identical topologies when the frequency of occurrence is set to = 50%. The taxa presently allocated to *Seicercus* appear in three different clades (A, B, C in Fig. 1). Clades A and B include only *Seicercus*, whereas clade C comprises *S. xanthobochistos* and *P. occipitalis*. These three clades receive high bootstrap and posterior probability support, although their relationships to each other and to other clades are not confidently established.

The relationships within clade A are fairly well resolved by the molecular data. However, one of the internal nodes receives low posterior probability and bootstrap values, and three other internal nodes have low bootstrap support. The non-molecular characters help to increase the bootstrap support for three of these relationships (Fig. 1). The data suggest that *S. burkii* and *S. tephrocephalus* (clade A3) are more closely related to *S. affinis* (clade A4) and *S. poliogenys* (clade A5) than

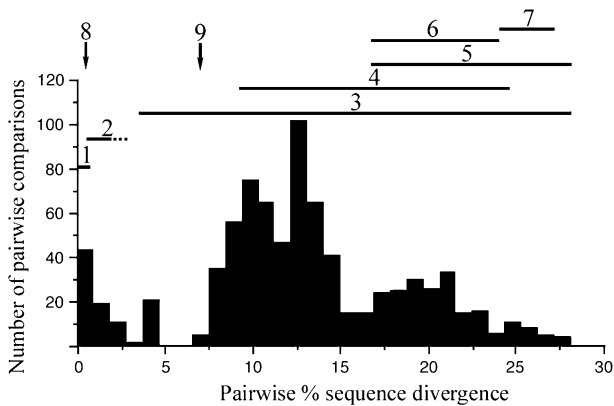


Fig. 2 Distribution of pairwise distances in the cytochrome *b* gene among the *Seicercus sensu lato* in this study [taxonomy of group D in Fig. 1 according to Alström & Olsson (2000) (cf. Table 1), of other *Seicercus* from Watson *et al.* (1986)]. Divergences estimated by maximum likelihood under the GTR + Γ + I model (shape parameter 1.3, proportion of invariable sites 0.63). (1) Within the same least-inclusive taxon, excluding *poliogenys* and comparisons between the non-monophyletic populations of *intermedius* from Sichuan/Guangdong, China vs. Fujian, China. (2) Between taxa treated as subspecies of the same species, as well as *S. poliogenys* from Yunnan, China vs. south Vietnam, and *intermedius* from Sichuan/Guangdong vs. Fujian (dashed line). (3) Between taxa classified as separate species, except *S. castaniceps castaniceps* vs. *S. grammiceps grammiceps*. (4) Between sympatric species. (5) Clade A vs. clade B. (6) Clade A vs. *xanthoschistos*. (7) Clade B vs. *xanthoschistos*. (8) *Seicercus castaniceps castaniceps* vs. *S. grammiceps grammiceps*. (9) *Seicercus poliogenys* from west Bengal, India vs. Yunnan, China/south Vietnam.

to the other taxa traditionally placed in the *S. burkii* complex (clades A1, A2, *omeiensis* and *S. soror*; marked D in Fig. 1), rendering the *S. burkii* complex non-monophyletic.

Pairwise sequence divergences between taxa are shown in Figs 2 and 3.

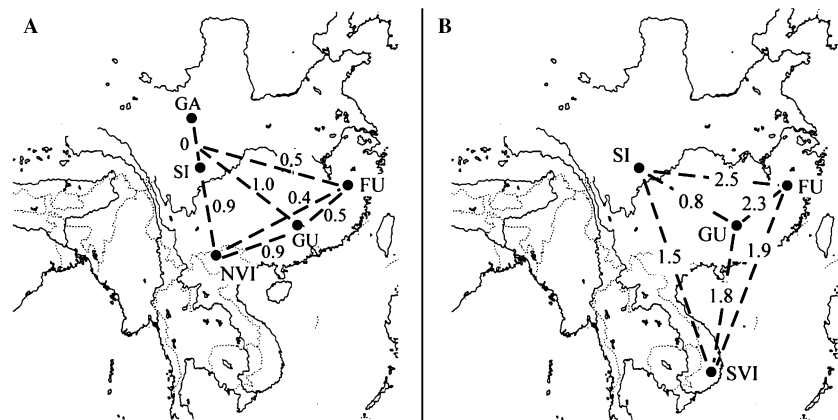


Fig. 3 A, B. Map showing the origin of our samples of —A. *Seicercus valentini valentini* (GA, SI) and *S. valentini latouchei* (FU, GU, NVI) and —B. *S. affinis intermedius* [including ‘*cognitus*’ (SI, FU, GU)] and *S. affinis ocellatus* (NVI), and the pairwise cytochrome *b* divergences between them. GA, Gansu; SI, Sichuan; FU, Fujian; GU, Guangdong; NVI, north Vietnam; SVI, south Vietnam.

Discussion

Monophyly of *Seicercus*

Mayr (in Watson *et al.* 1986) commented that ‘justification of the genus *Seicercus* is rather questionable’, and concluded that ‘the included species seem to be nothing but tropical *Phylloscopus*’. The monophyly of *Seicercus* as presently defined (e.g. Watson *et al.* 1986) is not supported by the mitochondrial gene tree (Fig. 1). This is particularly evident by the strongly supported position of *xanthoschistos* as sister to *occipitalis* and removed from clades A and B.

Clade A has high bootstrap and posterior probability values. It is also strongly supported by the diagnostic head pattern (blackish stripes on sides of crown; uniformly coloured sides of head; conspicuous pale eye-ring; extensively grey or green crown and ear-coverts; and lack of rufous coloration). This clade is composed exclusively of taxa currently placed in *Seicercus*, including the type of the genus (*burkii*). We analysed all known least-inclusive taxa that are likely to be part of this clade, except topotypical *affinis* of which we had no samples.

Clade B is well supported in the gene tree (Fig. 1). The monophyly of the three taxa in this clade, *grammiceps*, *montis* and *castaniceps*, is further corroborated by their characteristic head patterns (in addition to the first three features mentioned above for the taxa in clade A, also rufous crown, in *grammiceps* and *montis* also rufous ear-coverts) and very thin, high-pitched, simple songs, which are similar to each other but markedly different from other *Seicercus* and *Phylloscopus* (personal observations). Another 14 taxa currently treated as subspecies of *grammiceps*, *montis* and *castaniceps* (Watson *et al.* 1986) probably belong in this clade. *Seicercus grammiceps* is the type of the genus *Pycnosphrys* Strickland, 1849, and this name is thus available for this clade. However, as this name has not been used in recent times and the relationship between clade B and other clades is unclear, we refrain from explicitly assigning a name to clade B.

Clade C (Fig. 1), which contains the two taxa *occipitalis* and *xanthobchistos*, is well supported by the molecular data. Both *occipitalis* and *xanthobchistos* were originally placed in the genus *Phylloscopus*, but while the former is now in *Phylloscopus*, the latter is in *Seicercus* (e.g. Watson *et al.* 1986; Sibley & Monroe 1990; Inskipp *et al.* 1996). *Phylloscopus occipitalis* has never been included in the genus *Seicercus*. The plumage of *xanthobchistos* is more similar to *Phylloscopus* than to other *Seicercus*; e.g. unlike the taxa in clades A and B, both *xanthobchistos* and *Phylloscopus* show a prominent pale stripe above the eye and a dark stripe through the eye. Moreover, the song and call of *xanthobchistos* are closely similar to those of *occipitalis*, while they differ more from all other taxa currently classified as *Seicercus* (personal observations).

To conclude, mitochondrial DNA suggests that both *Seicercus* and *Phylloscopus* are non-monophyletic. The non-monophyly of these genera is further corroborated by non-molecular data which support the position of *xanthobchistos* in a clade with *occipitalis* (and other 'classic' *Phylloscopus* taxa belonging in that clade; U. Olsson, P. Alström & P. Sundberg, unpublished results). The following nomenclatural arrangements are then possible: (1) *Seicercus sensu lato* is synonymized with *Phylloscopus*, or (2) *Phylloscopus* is split into a number of genera. However, the poor support for the relationships among the different clades in the mitochondrial tree prevents a revision at present. Moreover, independent data sets, e.g. nuclear sequences, are required to corroborate or reject the present phylogenetic hypothesis. As a first step, *xanthobchistos* should be placed in *Phylloscopus*.

Monophyly of the *S. burkii* complex and relationships of *S. affinis* and *S. poliogenys*

Seicercus burkii was previously treated as a single polytypic species (e.g. Watson *et al.* 1986; Sibley & Monroe 1990; Inskipp *et al.* 1996). Although Alström & Olsson (1999, 2000) and Martens *et al.* (1999) showed that *S. burkii* is a complex of species (Table 1), they did not question the monophyly of this complex (marked by dashed vertical bar D in Fig. 1). However, maximum likelihood and Bayesian inference suggest, with a high posterior probability value, that *S. burkii* and *S. tephrocephalus* (clade A3) are more closely related to *S. affinis* and *S. poliogenys* than to the other members of the '*S. burkii* complex'. In contrast, this clade is only poorly supported in the bootstrap analysis, and is not corroborated by any of the characters in our non-molecular data set. Independent data are needed to evaluate whether or not the '*S. burkii* complex' is monophyletic.

Seicercus affinis ssp. (clade A4) and *S. poliogenys* (clade A5) are sister taxa in the maximum likelihood and Bayesian inference analyses, with high posterior probability. This clade receives < 50% bootstrap support in the mitochondrial tree, but > 60% in the 'total evidence' analysis. Morphological

evidence that this topology reveals the true taxon phylogeny is that the eye-ring is broken above the eye in both *S. affinis* ssp. and *S. poliogenys*. No other *Seicercus/Phylloscopus* shows this character state.

Relationships within the '*S. burkii* complex'

Clade A1 (Fig. 1) is a mixture of samples a priori allocated to *valentini* and *latouchei*. The samples from the northern part of the range (clade A1b), c. 350 km apart, have identical *cytb* sequences (Fig. 3A). They differ by just one substitution from a published 616 bp long *cytb* sequence of topotypical *valentini* from Shaanxi province, c. 400 km further northeast (Martens *et al.* 1999). We consider these three samples to belong to the same terminal lineage, to which we apply the name *valentini*. Clade A1a consists of three southern populations. The one from Fujian, China, represents topotypical *latouchei*. Alström & Olsson (1999) ascribed the population from northern Vietnam to *valentini*, while they did not study the one in Guangdong, China. These three populations are separated geographically from each other by c. 550–1500 km and from the populations in clade A1b by c. 800–1500 km (Fig. 3A). Clade A1a is well supported by the molecular data. *Cytb* divergences (Fig. 3A) indicate reduced gene flow between clades A1a and A1b, and perhaps also within A1a (between Guangdong and Vietnam). Among *Seicercus sensu stricto* taxa, of which we have multiple individuals (Appendix), birds from the same locality differ by up to 0.5%, and the within-taxon variation reaches 0.8% (Fig. 2). This is in agreement with a study of four *Phylloscopus* taxa where the divergence within a least-inclusive taxon did not exceed 0.4% (Helbig *et al.* 1996; Kimura two-parameter distances, which at this degree of divergence are expected to be at the most only slightly lower than figures obtained by the GTR + Γ + I model). Alström & Olsson (1999) reported slight morphological differences between *valentini* and *latouchei*. They also noted that the only topotypical *latouchei* they had tape recordings of had a different call note compared with *valentini*; however, they remarked that birds from Vietnam call like *valentini*. We tentatively apply the name *latouchei* to all three populations in clade A1a.

Alström & Olsson (1999, 2000) recognized *whistleri* from the western Himalayas and *nemoralis* from the central and eastern Himalayas and western Myanmar, and treated them as conspecific (Table 1). However, the present analysis does not support any subdivision of the Himalayan populations (clade A2a, Fig. 1). Moreover, the sequence divergence of 0.3% between the topotypical *whistleri* from Himachal Pradesh, northwestern India, and the birds from further east in the Himalayas is within the range of within-population variation (Fig. 2). In contrast, the Himalayan and western Myanmar populations differ by 1.4–2.0%, and the divergence between these is also well supported by bootstrap and posterior probability values. The type locality of *nemoralis* is

in Mizoram province, northeast India. We have no samples from that area, but judging from the geographical proximity of the western Myanmar and Mizoram populations, these are probably more closely related to each other than either is to Himalayan populations. According to Alström & Olsson (1999), birds from the western Himalayas (*whistleri sensu stricto*) differ slightly in plumage from birds from the central and eastern Himalayas. Alström & Olsson (1999) also reported differences in calls between birds from the western and eastern Himalayas, but the western type of call was also noted from the central Himalayas, i.e. apparently not correlated with the plumage variation. We have only one tape recording of the call of *nemoralis* from western Myanmar, but that differs markedly from the two call types noted by Alström & Olsson (1999) in the Himalayas. It is possible that Himalayan birds represent two lineages that have split so recently from each other that differences in mitochondrial DNA have not yet accumulated. Alternatively, introgression may have homogenized previously distinct populations. In accordance with our phylogenetic hypothesis, *contra* Alström & Olsson (1999, 2000), we restrict *nemoralis* to populations breeding in western Myanmar, Mizoram and adjacent areas southeast of the Brahmaputra River, while we provisionally unite the populations in the Himalayas under the name *whistleri*.

Alström & Olsson (1999) hypothesized, based on plumage, structure, song and habitat choice/altitude, that *valentini/latouchei* (clade A1) and *whistleri/nemoralis* (clade A2) form a monophyletic group. Such a clade has a reasonably high posterior probability, and a rather high bootstrap value when non-molecular data are included in a 'total evidence' analysis, but a bootstrap value of <50% when only molecular data are analysed (Fig. 1). The positions of *soror* and *omeiensis* are only confidently resolved in the 'total evidence' analysis, where *soror* is placed as sister to clades A1/A2 with fairly high support (Fig. 1). In the maximum likelihood tree (not shown), that position is instead occupied by *omeiensis*, but the reconstruction is unreliable as the internode between *soror* and *omeiensis*/A1/A2 is extremely short (0.00028).

Alström & Olsson (1999) recognized three populations of *tephrocephalus* based on differences in morphology (all three), vocalizations (at least two) and allopatric distributions. Martens *et al.* (1999) found two of these breeding in sympatry, and described one of them as *omeiensis*. The present study supports the distinctness of *omeiensis* (Fig. 1). The similarity in plumage and song between *omeiensis* and *tephrocephalus*, that led Alström & Olsson (1999) to treat these as conspecific, is apparently due to retained plesiomorphic character states in the former. Mitochondrial data do not corroborate any subdivision of *tephrocephalus sensu stricto*. The divergence in *cytb* is only 0.2–0.3% between our samples from northern Vietnam and western Myanmar, which is comparable to within-population divergence (Fig. 2). That *tephrocephalus*

is most closely related to *burkii* is well supported by the molecular data (Fig. 1). This is further corroborated by a non-molecular character: in these two taxa the eye-ring is usually thinly broken at the rear, unlike in all other *Seicercus sensu stricto* (Table 2).

The genetic distances between the two clades *valentini/latouchei* (clade A1) and *whistleri/nemoralis* (clade A2) (7.7–10.1%) are pronounced and support their classification as two separate species, as the divergences approach or match those between sympatric species (Fig. 2). The genetic divergence between *burkii* and *tephrocephalus* (clade A3) is considerably lower (3.7–4.1%) and approaches those among taxa treated as conspecific in the present study (Fig. 2). However, there are sympatric *Phylloscopus* species with similar genetic distances (U. Olsson, P. Alström & P. Sundberg, unpublished results), so treatment of *burkii* and *tephrocephalus* as different species is acceptable in our opinion.

Relationships within the S. affinis complex

With respect to *S. affinis*, the mitochondrial gene tree (clade A4, Fig. 1) is incompatible with the current taxonomy. The geographically isolated population in southern Vietnam, *ocularis*, is usually synonymized with *affinis* (e.g. Watson *et al.* 1986), but we were unable to evaluate this in the absence of samples representing topotypical *affinis* from the eastern Himalayas.

The taxa *intermedius* and *cognitus* were described as separate species [the latter as a subspecies of *Cryptolopha* (= *Seicercus*) *burkii*] from the same locality in Fujian province, southeastern China. They are clearly different morphologically, *cognitus* being green on the median crown-stripe and the side of the crown above the eye, while *intermedius* shows a grey median crown-stripe and rear crown-sides. Stresemann (1940) proposed that *cognitus* is a colour morph of *intermedius*. The present analysis supports this conclusion. One of our two '*cognitus*' and our single *intermedius* from the same locality in Fujian, China, are identical in the genes studied, while the other '*cognitus*' from the same place differs by 0.4%. Further evidence is that '*cognitus*' is only known to occur within the range of *intermedius*, and according to our field observations (unpublished), intermediate individuals exist, the vocalizations of *cognitus* and *intermedius* are indistinguishable, and males of both forms respond to playback of the other's song.

Our three samples of *intermedius* from Guangdong (one) and Sichuan (two) provinces in southern and central China, respectively, are separated geographically from each other by c. 1100 km and the former from the Fujian population by c. 550 km (Fig. 3B). The three former birds have pairwise *cytb* divergences ranging from 0.5% (same locality) to 0.8%, and are likely to represent the same terminal lineage (Fig. 3B). However, they differ from topotypical *intermedius* from Fujian by 2.3–2.5% (Fig. 3B), and the gene tree (Fig. 1) suggests that they are phylogenetically closer to *ocularis* from

southern Vietnam than to topotypical *intermedius*. This is unexpected, as the plumage of *intermedius* is markedly different from that of *ocularis*, while the three populations of *intermedius* appear to be indistinguishable by plumage. Independent data, such as nuclear genes, are required to corroborate or reject the present hypothesis.

Variation among populations of *S. poliogenys*

Seicercus poliogenys (clade A5, Fig. 1) is considered monotypic (Watson *et al.* 1986). However, the present analysis demonstrated a divergence of 6.9–7.3% in the *cytb* gene between the topotypical Himalayan specimen and the two birds from southern Vietnam and Yunnan, China, respectively (Fig. 2). The two latter differ by 1.7%. Provided that none of these sequences is accidentally amplified nuclear pseudogenes, they represent separate lineages. The differences between the two southeastern populations and the Himalayan one are well within the range of taxa treated as separate species (e.g. almost twice as large as between *burkii* and *tephrocephalus*). More research is needed on the variation in genetics, morphology, vocalizations and distributions in *S. poliogenys*.

Taxon sampling in intrageneric phylogenetic studies

Our results suggest cases of non-monophyly and the existence of cryptic species in the genus *Seicercus*. This underscores the importance of dense sampling at the specific and infraspecific levels in intrageneric phylogenetic studies. Most previous studies have only dealt with taxa treated as species, and few have attempted to sample all terminal lineages in a genus.

Acknowledgements

This study was financially supported by grants from Uddenberg Nordingska Stiftelsen, Helge Axson Johnsons Stiftelse, Hierta Retzius Stiftelse, Yngve Sjöstedts donation/Kungliga vetenskapsakademien, Elis Wides fond, Kungliga & Hvitfeldtska Stipendieinrättningen, Olle och Signhild Engkvists stiftelser, Erna and Victor Hasselblad Foundation, and the Swedish Natural Science Research Council (to PS). We are indebted to Geoff Carey, Cheung Ho Fai, Jon Fjeldså, Magnus Gelang, Darren Irwin, Anders Jihmanner, Paul Leader, Bettina Olausson, Bo Pettersson and Yoshimitsu Shigeta for providing samples, and to Fredrik Ronquist and an anonymous referee for comments on the manuscript.

References

- Allende, L. M., Rubio, I., Ruiz del Valle, V., Guillen, J., Martinez-Laso, J., Lowe, E., Varela, P., Zamora, J. & Arnaiz-Villena, A. (2001). The Old World sparrows (genus *Passer*) phylogeography and their relative abundance of nuclear mtDNA pseudogenes. *Journal of Molecular Evolution*, *53*, 144–154.
- Alström, P. & Olsson, U. (1999). The golden-spectacled warbler: a complex of sibling species, including a previously undescribed species. *Ibis*, *141*, 545–568.
- Alström, P. & Olsson, U. (2000). Golden-spectacled warbler systematics. *Ibis*, *142*, 495–500.
- Arctander, P. (1995). Comparison of a mitochondrial gene and a corresponding nuclear pseudogene. *Proceedings of the Royal Society of London Series B. Biological Sciences*, *262*, 13–19.
- Bensasson, D., Zhang, D.-X., Hartl, D. L. & Hewitt, G. M. (2001). Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends in Ecology and Evolution*, *16*, 314–321.
- Helbig, A. J., Martens, J., Seibold, I., Henning, F., Schottler, B. & Wink, M. (1996). Phylogeny and species limits in the Palaearctic chiffchaff *Phylloscopus collybita* complex: mitochondrial genetic differentiation and bioacoustic evidence. *Ibis*, *138*, 650–666.
- Helbig, A. J., Seibold, I., Martens, J. & Wink, M. (1995). Genetic differentiation and phylogenetic relationships of Bonelli's warbler *Phylloscopus bonelli* and green warbler *Phylloscopus nitidus*. *Journal of Avian Biology*, *26*, 139–153.
- Huelsenbeck, J. P. & Ronquist, F. (2001a). MrBayes: a program for the Bayesian inference of phylogeny, version 2.01 (<http://morphbank.ebc.uu.se/MrBayes/>).
- Huelsenbeck, J. P. & Ronquist, F. (2001b). MrBayes: Bayesian inference of phylogeny. *Bioinformatics*, *17*, 754–755.
- Inskipp, T., Lindsey, N. & Duckworth, W. (1996). *An Annotated Checklist of the Birds of the Oriental Region*. Sandy: Oriental Bird Club.
- Irwin, D. E., Alström, P., Olsson, U. & Benowitz-Fredericks, Z. M. (2001). Cryptic species in the genus *Phylloscopus* (Old World leaf warblers). *Ibis*, *143*, 233–247.
- Klitgaard Nielsen, K. & Arctander, P. (2001). Recombination among multiple mitochondrial pseudogenes from a passerine genus. *Molecular Phylogenetics and Evolution*, *18*, 362–369.
- Martens, J., Eck, S., Päckert, M. & Sun, Y.-H. (1999). The golden-spectacled warbler *Seicercus burkii* — a species swarm (Aves: Passeriformes: Sylviidae), part 1. *Zoologische Abhandlungen Museum Dresden*, *50*, 281–327.
- Nylander, J. A. A. (2002). Mrmodeltest, version 1.1b (<http://www.ebc.uu.se/systzoo/staff/nylander.html>).
- 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 D. P. Mindell (Ed.) *Avian Molecular Evolution and Systematics* (pp. 3–28). London: Academic Press.
- Quinn, T. W. & White, B. N. (1987). Analysis of DNA sequence variation. In F. Cooke & P. A. Buckley (Eds) *Avian Genetics: a Population and Ecological Approach*. London: Academic Press.
- Richman, A. D. (1996). Ecological diversification and community structure in the Old World leaf warblers (genus *Phylloscopus*): a phylogenetic perspective. *Evolution*, *50*, 2461–2470.
- Richman, A. D. & Price, T. (1992). Evolution of ecological differences in the Old World leaf warblers. *Nature*, *355*, 817–821.
- Sibley, C. G. & Ahlquist, J. E. (1990). *Phylogeny and Classification of Birds*. New Haven: Yale University Press.
- Sibley, C. G. & Monroe, B. L. Jr (1990). *Distribution and Taxonomy of Birds of the World*. New Haven: Yale University Press.
- Sorenson, M. D. & Quinn, T. W. (1998). Numts: a challenge for avian systematics and population biology. *Auk*, *115*, 214–221.
- Stresemann, E. (1940). *Ornithologische Monatsberichte*, *48*, 49–50.
- Swofford, D. L. (2001). *PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods)*, version 4.10b. Sunderland, MA: Sinauer.

Ticehurst, C. B. (1938). *A Systematic Review of the Genus Phylloscopus*. London: Trustees of the British Museum.

Vawter, L. & Brown, W. M. (1986). Nuclear and mitochondrial DNA comparisons reveal extreme rate variation in the molecular clock. *Science*, 234, 194–196.

Watson, G. E., Traylor, M. A. Jr & Mayr, E. (1986). Family Sylviidae, Old World warblers. In E. Mayr & G. W. Cottrell (Eds) *Checklist of Birds of the World*, Vol. 11 (pp. 221–262). Cambridge, MA: Museum of Comparative Zoology (Harvard University).

Appendix

Specimens used in this study.

Taxon	Locality	Status	Regions	Genbank accession number
<i>Seicercus valentini latouchei</i>	Fujian, southeast China (c. 27.5° N, 117.5° E)	b	Cytb 12S	AY635072 AY635119
<i>Seicercus valentini latouchei</i>	North Vietnam (c. 22° N, 103.5° E)	b	Cytb 12S	AY635088 AY635135
<i>Seicercus valentini latouchei</i>	Guangdong, south China (c. 25° N, 113° E)	b	Cytb 12S	AY635086 AY635133
<i>Seicercus valentini valentini</i>	Sichuan, central China (c. 29.5° N, 103° E)	b	Cytb 12S	AY635087 AY635134
<i>Seicercus valentini valentini</i>	Gansu, central China (c. 33° N, 104° E)	b	Cytb 12S	AY635089 AY635136
<i>Seicercus whistleri whistleri</i>	West Bengal, northeast India (c. 27° N, 89° E)	b	Cytb 12S	AY635077 AY635124
<i>Seicercus whistleri whistleri</i>	Central Nepal (c. 28° N, 85.5° E)	b	Cytb 12S	AY635075 AY635122
<i>Seicercus whistleri whistleri</i>	Himachal Pradesh, northwest India (c. 32° N, 77° E)	b	Cytb 12S	AY635090 AY635137
<i>Seicercus whistleri nemoralis</i>	West Myanmar north (c. 23° N, 94° E)	b	Cytb 12S	AY635076 AY635123
<i>Seicercus whistleri nemoralis</i>	West Myanmar south (c. 21° N, 93° E)	b	Cytb 12S	AY635074 AY635121
<i>Seicercus soror</i>	Sichuan, central China (c. 29.5° N, 103° E)	b	Cytb 12S	AY635082 AY635129
<i>Seicercus soror</i>	Shaanxi, north central China (c. 33.5° N, 107.5° E)	b	Cytb 12S	AY635083 AY635130
<i>Seicercus omeiensis</i>	Sichuan, central China (c. 29.5° N, 103° E)	b	Cytb 12S	AY635078 AY635125
<i>Seicercus burkii</i>	West Bengal, India (c. 27° N, 89° E)	b	Cytb 12S	AY635065 AY635112
<i>Seicercus burkii</i>	Central Nepal (c. 28° N, 85.5° E)	b	Cytb 12S	AY635064 AY635111
<i>Seicercus tephrocephalus</i>	West Myanmar (c. 21° N, 93° E)	b	Cytb 12S	AY635085 AY635132
<i>Seicercus tephrocephalus</i>	North Vietnam (c. 22° N, 103.5° E)	b	Cytb 12S	AY635084 AY635131
<i>Seicercus 'affinis intermedius'</i>	Sichuan, China (c. 29.5° N, 103° E)	b	Cytb 12S	AY635070 AY635117
<i>Seicercus 'affinis intermedius'</i>	Guangdong, south China (c. 25° N, 113° E)	b	Cytb 12S	AY635069 AY635116
<i>Seicercus affinis ocularis</i>	South Vietnam (c. 11.5° N, 108° E)	b	Cytb 12S	AY635063 AY635110
<i>Seicercus affinis 'cognitus'</i>	Fujian, southeast China (c. 27.5° N, 117.5° E)	b	Cytb 12S	AY635066 AY635114
<i>Seicercus affinis intermedius</i>	Fujian, southeast China (c. 27.5° N, 117.5° E)	b	Cytb 12S	AY635071 AY635118
<i>Seicercus 'poliogenys'</i>	South Vietnam (c. 11.5° N, 108° E)	b	Cytb 12S	AY635080 AY635126
<i>Seicercus 'poliogenys'</i>	Yunnan, south central China (c. 25° N, 98° E)	b	Cytb 12S	AY635081 AY635128

Appendix Continued

Taxon	Locality	Status	Regions	Genbank accession number
<i>Seicercus polioagenys</i>	West Bengal, northeast India (c. 27° N, 89° E)	b	Cytb 12S	AY635079 AY635127
<i>Seicercus castaniceps castaniceps</i>	West Bengal, India (c. 27° N, 89° E)	b	Cytb 12S	AY635067 AY635113
<i>Seicercus grammiceps grammiceps</i>	West Java (c. 6.5° S, 106.5° E)	b	Cytb 12S	AY635068 AY635115
<i>Seicercus montis montis</i>	North Borneo (c. 6° N, 116.5° E)	b	Cytb 12S	AY635073 AY635120
<i>Phylloscopus ruficapilla minullus</i>	Tanzania (c. 5° S, 38° E)	b	Cytb 12S	AY635060 AY635105
<i>Phylloscopus umbrovirens mackensianus</i>	Kenya (c. 0.5° S, 36.5° E)	b	Cytb 12S	AY635062 AY635109
<i>Phylloscopus ijimae</i>	Japan Japan (c. 34° N, 139.5° E)	b	Cytb 12S	Y10741 AY635098
<i>Phylloscopus coronatus</i>	Japan (c. 36° N, 140.5° E)	b	Cytb 12S	AY635053 AY635096
<i>Phylloscopus occipitalis</i>	Northwest Frontier Prov., north Pakistan (c. 34.5° N, 73.5° E)	b	Cytb	AY635057 AY635102
<i>Seicercus xanthoschistos xanthoschistos</i>	Northwest Frontier Prov., north Pakistan (c. 34.5° N, 73.5° E)	b	Cytb 12S	AY635091 AY635138
<i>Phylloscopus borealis borealis</i>	Hebei, northeast (c. 39.5° N, 119° E)	m	Cytb 12S	AY635052 AY635094
<i>Phylloscopus magnirostris</i>	South India (c. 12° N, 77° E)	m	Cytb 12S	AY635056 AY635101
<i>Phylloscopus trochilus trochilus</i>	Germany South Sweden (c. 57.5° N, 12° E)	b	Cytb 12S	Z73492 AY635108
<i>Phylloscopus collybita abietinus</i>	Sweden North Sweden (c. 63.5° N, 20.5° E)	b	Cytb 12S	Z73479 AY635095
<i>Phylloscopus schwarzi</i>	Hebei, northeast China (c. 39.5° N, 119° E)	m	Cytb 12S	AY635061 AY635106
<i>Phylloscopus fuscatus fuscatus</i>	China Hebei, northeast China (c. 39.5° N, 119° E)	m	Cytb 12S	Y10729 AY635097
<i>Phylloscopus bonelli</i>	Spain Spain (c. 37° N, 3° W)	b	Cytb 12S	Z73490 AY635093
<i>Phylloscopus sibilatrix</i>	Germany South Sweden (c. 57.5° N, 12° E)	b	Cytb 12S	Z73491 AY635107
<i>Phylloscopus inornatus</i>	Hebei, northeast China (c. 39.5° N, 119° E)	m	Cytb 12S	AY635054 AY635099
<i>Phylloscopus proregulus</i>	Hebei, northeast China (c. 39.5° N, 119° E)	m	Cytb 12S	AY635058 AY635103
<i>Phylloscopus pulcher pulcher</i>	Yunnan, China (c. 25° N, 98° E)	b	Cytb 12S	AY635059 AY635104
<i>Phylloscopus maculipennis maculipennis</i>	West Bengal, India (c. 27° N, 89° E)	b	Cytb 12S	AY635055 AY635100
<i>Acrocephalus scirpaceus scirpaceus</i>	Germany Spain (c. 49.5° N, 0° E)	b	Cytb 12S	Z73483 AY635092
<i>Sylvia atricapilla atricapilla</i>	Germany Spain (c. 49.5° N, 0° E)	b	Cytb 12S	Z73494 AY635139

Cytb, cytochrome *b*.

In the status field, individuals sampled on their breeding grounds in the breeding season (and in all but a few cases showing territorial behaviour) are denoted b, while individuals sampled on migration or in their winter quarters are denoted m.