CMLS, Cell. Mol. Life Sci. 58 (2001) 2117–2128
1420-682X/01/142117-12 $ 1.50 + 0.20/0

CMLS Cellular and Molecular Life Sciences

Research Article

Both morphological and molecular characters support speciation of South American siskins by sexual selection

R. van den Elzen [a], J. Guillén [b], V. Ruiz-del-Valle [b], L. M. Allende [b], E. Lowy [b], J. Zamora [b] and A. Arnaiz-Villena [b,]*

[a] Alexander Koenig Research Institute and Zoological Museum, Bonn (Germany)
[b] Department of Immunology and Molecular Biology, Hospital 12 de Octubre, Universidad Complutense,
28041 Madrid (Spain), Fax +34 91 390 83 99, e-mail: aarnaiz@eucmax.sim.ucm.es; http://chopo.pntic.mec.es/biolmc

Received 11 September 2001; received after revision 12 October 2001; accepted 15 October 2001

Abstract. South American siskin radiation was studied by both mitochondrial cytochrome b (mt cyt b) DNA sequencing and homologous phenotypic characters; the latter were coded separately according to sex. Mixed phenetic and molecular (total evidence) dendrograms were constructed and the corresponding analyses suggest that speciation started in the South American siskin group with a north to south separation (*Carduelis notata/C. barbata*) along the Andean spine. A second split may have taken place around the Peruvian Andean mountains, corresponding to the present distribution pattern of *C. olivacea*. The most recent speciation events seem to have occurred in three sister species pairs: (i) *C. xanthogastra/C. atrata*, (ii) *C. magellanica/C. yarrellii*, (iii) *C. cucullata/C. crassirostris*. Accumulation of consistent characters in both morphological and molecular data at the basal nodes of the dendrograms indicate that speciation events occurred within a short period of time. Our data also suggest that speciation probably occurred by sexual selection through female mating choice in this radiation. Additionally, studies of variable amino acid residues in the mt cyt b molecule show that the three variable amino acids found are placed in the mitochondrial transmembrane region, which is also part of the hypervariable region in mammals. Each of the three amino acid changes occur in each of the three postulated evolutionary groups.

Key words. Siskins; Carduelini; *Carduelis*; cyt b; speciation mode; South America; evolution; sexual selection.

Intergeneric relationships within the *Carduelis* species complex (Aves, Carduelidae; goldfinches, siskins, greenfinches and redpolls) have been put forward by Arnaiz-Villena et al. [1] using cytochrome b (cyt b) DNA phylogeny analysis. Species grouping generally matched their present-day geographic distribution [1]. Twenty-five out of 31 extant taxa recognized by Sibley and Monroe [2] were studied. Siskins as grouped in traditional systematics [3] under the generic name of *Spinus* did not show up as monophyletic (derived from a common ancestor), but as paraphyletic in three entities: a European-North American, a North American and a North-South American disjunction, with apparent closer relationships to other non-siskin species. High bootstrap values supported a South American siskin clade and the basal position of the North American *Carduelis notata* (black-headed siskin) in that clade. Phylogenetic relationships within the South American species group could not be solved convincingly, as nodes were only weakly supported by low bootstrap values [see figs. 1 and 2 in ref. 1]. South American siskins were found to constitute the youngest clade within *Carduelis* using the fossil record of chicken and pheasant divergence time to calibrate the molecular clock; the authors suggested that the North

---

* Corresponding author.

American *notata* (or its ancestor) gave rise to a rapid South American siskin radiation when the Isthmus of Panama became established, only about 3–5 million years (my) ago.

Phylogenetic estimates of taxonomic groups are thought to be more precise if a combination of different data sets is used. All relevant character information may be combined in one single phylogenetic analysis which is named total evidence analysis [4]. Total evidence has successfully been established to ascertain the degree of relatedness for diverse organismic groups at various systematic levels [5–8], including the study of bird phylogenies [9–11]. Conditional combination [12, 13] is a restricted approach that only allows the combination of data sets that are not significantly controversial, when tested for congruence. Chippindale and Wiens [8] suggested the use of differential character weighting for unbalanced, heterogeneous data sets to accommodate different evolutionary processes; weighting itself is not unquestioned [14].

Phylogenetic hypotheses based solely on morphological data sets are often considered less informative than molecular analyses because they are believed (i) to be less objective, because characters are a priori judged before tree-building [15] and (ii) to include a large proportion of non-phylogenetic (anagenetic, i. e. adaptive or evolutionary) information content [16, 17]. Furthermore, phenetic traits are thought by some authors to be heavily prone to convergent evolution in birds, including the Carduelidae family where there is strong evidence that sexual ornamentation is correlated with both ecological factors and particular life history traits in addition to phylogeny [18]. Convergence in molecular characters has only recently been regarded as a fundamental problem in phylogenetic studies [19], but it does not apply to this microevolutionary study.

In cases of suspected incongruence between molecular and phenetic data sets, comparison of single data sets can elucidate the underlying evolutionary phenomena [20]. We want to address these conflicting arguments by: (i) proposing a phylogenetic hypothesis for a relatively recent bird clade (the South American siskins) based on total evidence from molecular and phenetic characters, (ii) testing the congruence of tree topologies based on only either single phenetic or molecular evidence, (iii) tracing patterns of character distributions for the combined data set and its separate portions and (iv) judging the plausibility and reliability of the total evidence hypothesis by analysis of geographic distribution patterns that explain speciation events in the Andes [21, 22]. Finally, the evolutionary hypothesis of sexual selection driven by mating choice will be tested based on these and other already published phylogeographic data.

Materials and methods

All phylogenetic analyses were carried out using the 4.0b4-8 test versions of PAUP* [23]. Distribution patterns rely on maps provided in the literature [24–26] supplemented by finding localities from museum specimens. The taxa used for analyses were 10 (out of 13 actually recognized) South American siskin species (Genbank accession numbers are in parentheses): *C. cucullata* (L76299), *C. atrata* (L76385), *C. olivacea* (L77871), *C. crassirostris* (L77869), *C. spinescens* (U79017), *C. magellanica* (U79016), *C. yarrellii* (U83200), *C. barbata* (L77868), *C. xanthogastra* (L76389), *C. notata* (U79019), complemented by *Fringilla coelebs* (L76609) as outgroup and reference taxon and *C. psaltria columbiana* (U78324), a sister taxon with a North American origin, but with a North to South American distribution range [27]. [See ref. 1 for the origin of DNA-analyzed species and North American *C. psaltria* accession number]. This is the relevant point to mention that one of the authors, Arnaiz-Villena in 1996, could not find *C. siemiradzkii* after intensive search along the Santa Elena peninsula (Ecuador), in Machalilla National Park and Chongon Hills. Dr. A. Agreda (Ecuador) has also searched in two different years (1999–2000) for these birds and has not been able to find them in a wide range of Ecuadorian dry woods. Either this bird is almost extinct [28] or is just a colour morph of *C. magellanica* which was collected in 1868, and its orange colour is due to feeding habits or other environmental factors [24].

Phenetic analyses

Museum skins (1884 specimens) were investigated and 22 homologous phenotypic characters were coded separately for each sex. Twenty characters were constituted from observed variation in plumage colours and ornamentations for several body regions; bill and leg colour provided two characters. Both sex-based matrices were combined in an overall phenetic matrix that included 47 characters (table 1: 22 plus 22 for each sex, 2 regarding juvenile plumage, and 1 character giving additional information on the degree of sexual dichromatism).

To test incongruence within and between trees, three types of statistics were applied. The HomPart command in PAUP [23] was used to perform tests for homogeneity of partitioned data sets – male and female parsimony informative characters in the combined morphological data set as well as morphological and molecular parsimony informative characters in the combined data set of the total evidence matrix. This test corresponds to the incongruence length difference test of Farris et al. [29] that calculates the lengths of the most parsimonious trees for each matrix (partition) separately and for the combined matrix, including all characters. This first statistic assesses congruence or incongruence within trees. For comparison

Table 1. Morphological data matrix and its male, female and sexual neutral partitions.

Male characters	Female characters	Character description and coding
1 (925)	**23 (947)**	bill coloration, 0 = dark grey to black, 1 = grey to light horn coloured
2 (926)	24 (948)	head pattern, 0 = grey hood, 1 = black cap, 2 = black hood
3 (927)	25 (949)	throat patch (‘bib’), 0 = none, 1 = present
4 (928)	26 (950)	breast coloration, 0 = brown or grey, 1 = green, yellow, red, 2 = black
5 (929)	27 (951)	colour of leg feathers, 0 = white, 1 = green, yellow, red
6 (930)	28 (952)	centre of belly, 0 = white, 1 =g reen, yellow, red
7 (931)	29 (953)	pattern of flanks, 0 = no striation, 2 = striated
8 (932)	30 (954)	colour of back, 0 = not black, 1 = black
9 (933)	31 (955)	pattern of back, 0 = no striation, 1 = striated, 2 = scaly pattern
10 (934)	32 (956)	contrast of rump, 0 = none, 1 = medium, 2 = strong
11 (935)	33 (957)	contrast of head, 0 = yes, 1= no
12 (936)	**34 (958)**	primaries with white ‘window’, 0 = yes, 1 = no
13 (937)	**35 (959)**	primaries with yellow ‘window’, 0 = no, 1 = yes
14 (938)	36 (960)	secondaries with yellow ‘window’, 0 = no, 1 = yes
15 (939)	**37 (961)**	primary coverts with yellow edges, 0 = no, 1 = yes
16 (940)	**38 (962)**	base of tail feathers yellow, 0 = no, 1 = yes
17 (941)	**39 (963)**	white windows in tail feathers, 0 = yes, 1 = no
18 (942)	**43 (967)**	undertail coverts with dark spots or streaks, 0 = no, 1 = yes
19 (943)	44 (968)	undertail coverts yellow, 0 = no, 1 = yes
20 (944)	45 (969)	superciliar stripe, 0 = no, 1 = yes
21 (945)	46 (970)	ear patch, 0 = present, 1 = absent
22 (946)	47 (971)	foot coloration, 0 = pink, 1 = light to slate grey, 2 = black
Sexual neutral 40 (964)		juvenile plumage 0 = streaked, 1 = unstreaked
Sexual neutral 41 (965)		female plumage, 0 = like juvenile, 1 = different
Sexual neutral 42 (966)		sexual dichromatism, 0 = pronounced, 1 =weak

Numbers of characters in parentheses refer to numbers used in the total evidence analysis. Bold face type marks synapomorphies.

between trees, a non-parametric sign test was used (winning sites test, as provided under the Pscores command of PAUP) [23]. Topologies are tested pairwise, taking both tree lengths and character distribution patterns as measures of fit and using exact binomial tests. Additionally, we also tested for taxonomic congruence, that is for differences in the position of species along the tree topology, using a simple Spearman rank correlation test. We counted numbers of branching points in maximum parsimony (MP) trees from basal to terminal to assign ranks to taxa; trees were tested pairwise. Mean and total character differences were used (neighbour-joining analysis, NJ, 1000 bootstrap replicates) as distance measure. Exhaustive searching as well as bootstrap with random search were used in the MP analysis. All characters were set unordered.

Molecular analyses

Blood from living birds was drawn after photographing and cutting the claws locally anaesthetized with a lidocaine ointment. Blood was collected in EDTA cooled at 4 °C and frozen down until use. DNA was obtained and the mitochondrial (mt) cyt b (924 DNA bases) was amplified with primers L14841 5′-AAAAAGCTTCCATCCAACATCTCAGCATGAT-GAAA-3′ and H15767 5′-ATGAAGGGATGTTCTACTGGTTG-3′ [1]. At least two individuals per species were sequenced in order to discard any variation among individuals. Polymerase chain reaction (PCR), cloning and automatic DNA sequencing were performed as previously described [1]. At least four clones from each of two different PCRs were sequenced from each species in order to assess both PCR and DNA sequencing quality.

Mt cyt b (924 bp) was analysed from the same 12 taxa as in the morphological data set. Both distance and MP methods were applied; NJ bootstrap analysis was used based on maximum likelihood and heuristic search.

Total evidence

The total evidence method assumes that conflicts between hypotheses derived from different data sets can be put aside if they are used in combination, because the intrinsic phylogenetic signal increases by addition of several character sets. The molecular data set (924-bp cyt b) was combined with the 47 overall phenotypic character matrix; all characters were set unweighted. Unweighted data sets were compared and combined in order to explore character distributions. An NJ tree was built as was an MP tree under heuristic search on the 46 molecular plus 41 phenetic phylogenetically informative characters only, as a compromise to avoid the problem of weighting but obtaining at least a balanced number of character proportions for the two partitions. Taxonomic congruence was analysed by using a Spearman rank correlation test.

Results and discussion

Dendrogram analyses

1) Phenetic

The MP analysis did not corroborate traditional phenotype-based systematic species arrangements [2, 3, 24, 28], where hooded and capped taxa are lumped in species complexes. Only 12 of the 41 informative characters from the combined data set exhibited a consistency index (CI) of 1.00 and could thus be taken as the tree-building synapomorphies (fig. 1 A). Six each were derived from the male and the female partition. Of the12 synapomorphies, 10 characters stabilize the basal node *C. psaltria columbiana/C. barbata*, and thus the North American-South American species group disjunction. The two other synapomorphies support the clade *olivacea/notata/crassirostris* and *xanthogastra/atrata* with one character each; no morphological support is found for *magellanica.* Homoplasy obscures the signal in all branches and at all hierarchical levels of the South American species group.

The NJ bootstrap analysis (fig. 1 B) gives only a slightly different topology: monophyly of the South American clade is again supported, taxonomic congruence is found for the basal taxa *barbata* and *spinescens* and for the sister relationships of *notata/olivacea/crassirostris* as well as for the black-backed *xanthogastra/atrata*. *C. magellanica* has found a sister taxon: *cucullata*. As the capped *yarrellii* is shown at a basal hierarchical level together with the other capped taxa *barbata* and *spinescens*, the hood appears in this NJ analysis as a synapomorphic condition (note that it does not in the MP analysis). Sexual monochromatism as found in *notata, crassirostris* and *atrata* is considered as homoplasious.

2) Molecular

Topology of species in the maximum parsimony analysis generally contradicts the morphological hypothesis: monophyly of the South American clade is affirmed, but *notata*, a terminal taxon in the phenetic hypothesis, appears at its base, only then followed by *barbata* (fig. 2 A) and no morphological sister taxon relationship finds its molecular counterpart. Of 924 codon sites, only 46 were parsimony informative (table 2), the rest were either constant (762) or autapomorphies (116). Sixteen codon sites are consistent (CI = 1) and thus tree-building synapomorphies (fig. 3). As in the phenetic analysis, most of these

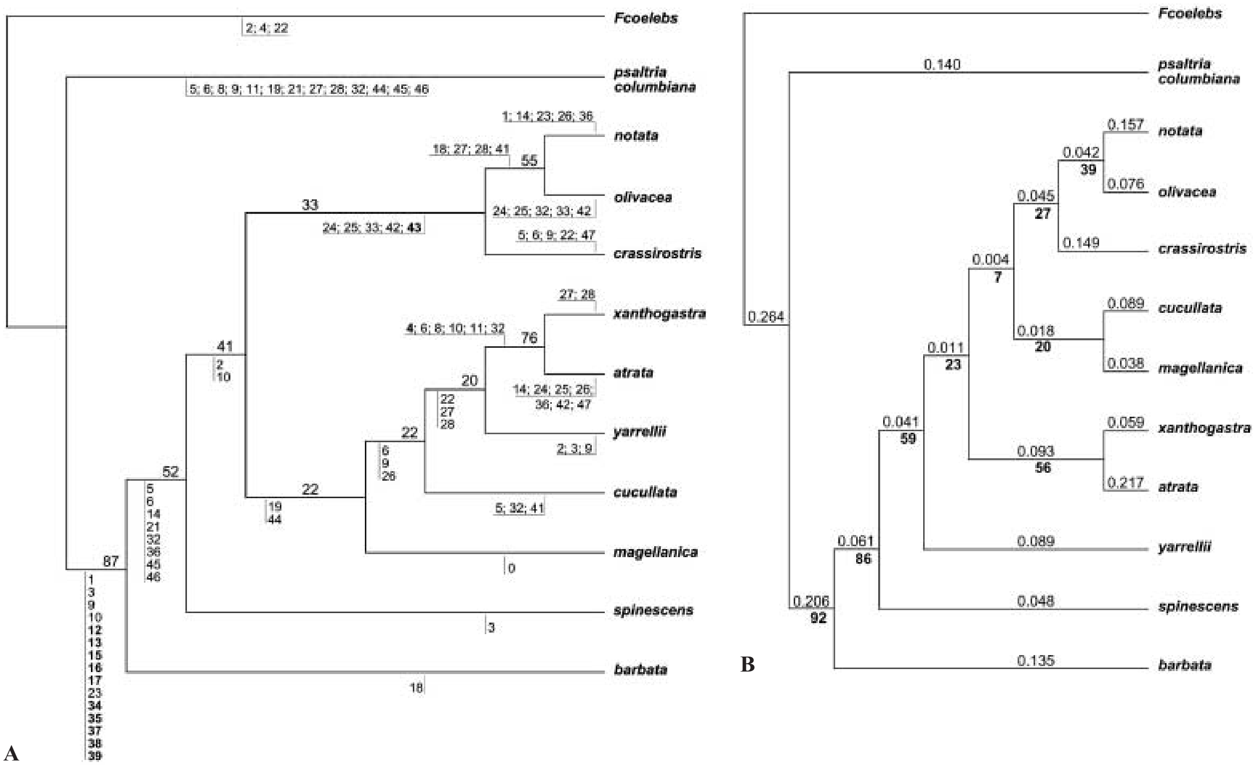

Figure 1. *A* Most parsimonious bootstrap consensus tree [97 steps, CI=0.5258, retention index (RI) = 0.4945] for 41 parsimony-informative morphological characters. Bold type indicates unambiguous (non-homoplasious) synapomorphies (ACCTRAN optimization). Bootstrap percentages from 1000 replicates are shown in larger type above branch lines. *B* NJ bootstrap consensus tree (minimum evolution score = 1.98587) based on 47 morphological characters and 1000 replicates. Branch lengths are given above the line, bootstrap percentages are below lines in bold type.

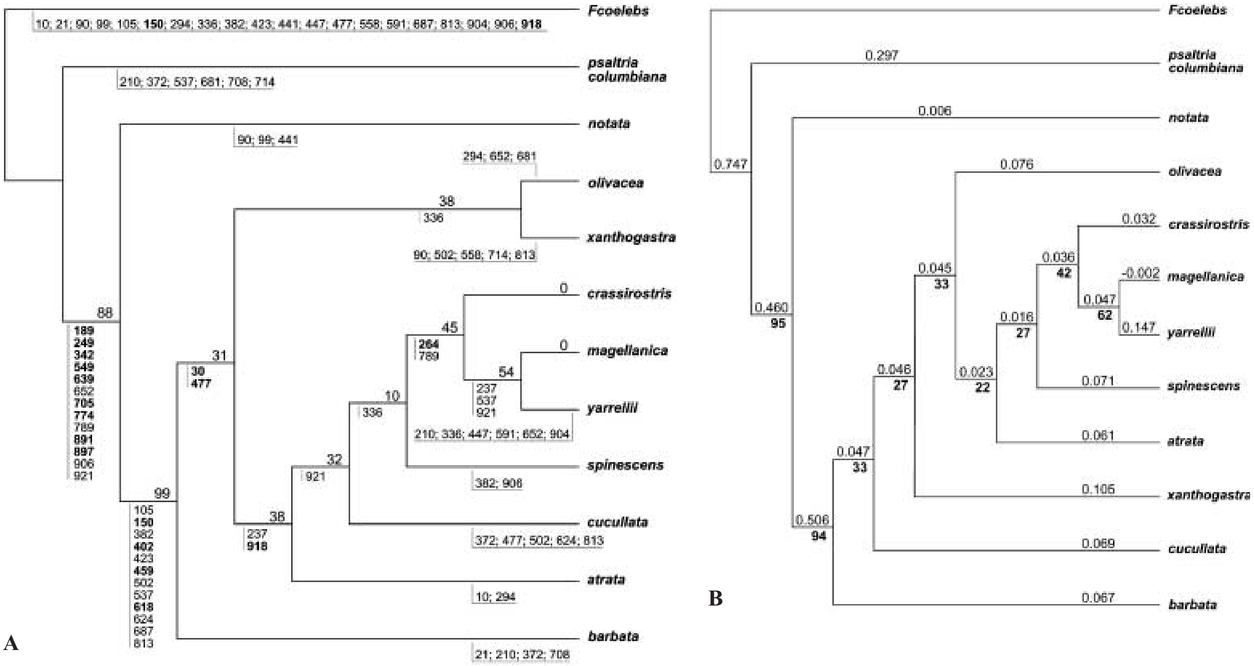

Figure 2. *A* Most parsimonious bootstrap consensus tree (93 steps, CI=0.5699, RI=0.5238) for 46 parsimony-informative molecular characters. Unambiguous synapomorphies (ACCTRAN optimization) in bold type below branch lines; bootstrap percentages from 1000 replicates above lines. *B* NJ bootstrap consensus tree (maximum-likelihood score) based on 924 cyt b base pairs. Branch lengths are indicated above, bootstrap values are below lines in bold.

Table 2. Phylogenetically informative sites found in the cyt b gene (1143 bp) are shown.

Phylogenetically informative DNA base sites (from base 97)	11122 1239905813 0100950907	2223333444 4693478024 9446222231	4445555566 4570345912 7972798184	6666777778 3588001781 9217584493	889999 990012 174681
F. coelebs coelebs	CCCCCCACTA	TATCTCCACT	CCACCTCCAC	CCCCTCTTTC	TCGATA
C. psaltria columbiana	TT.TTTC.C.	..CA.TT.TC	T.C.A.TT..	..TT.TC..A	..ACA.
C. notata notata	TT...TCT..	C.CAC.T.T.	T.C..CTT..	TT.TC..CCA	CTATAG
C. olivacea	TTTTT.TT..	C..GC..G.C	TTTTTCTTGT	T.T.C..CCT	CTATAG
C. crassirostris crassirostris	TTTTT.TT.G	CGCGC..G.C	TTTTTCTTGT	TT..C..C.T.	CTATG.
C. xanthogastra xanthogastra	TTT.T.TT..	C.CGC..G.C	TTT.TC.TGT	TT..C.CCC.	CTATAG
C. cucullata	TTTTT.TT.G	C.CACT.G.C	TTC.TCTTG.	TT..C..CC.	CTATG.
C. atrata	.TTTT.TT.G	C..AC..G.C	TTTTTCTTGT	TT..C..CCT	CTATGG
C. magellanica magellanica	TTTTT.TT..	CGCGC..G.C	TTTT.CTTGT	TT..C..C.T	CTATGG
C. spinescens spinescens	TTTTT.TT.G	C.CGC.TG.C	TTTTTCTTGT	TT..C..CCT	CTACG.
C. yarrellii	TTTTT.TTC.	CGCAC..G.C	.TTT.CT.GT	T...C..C.T	CT.TGG
C. barbata	T..TT.TTC.	C.CACT.G.C	TTCTTCTTGT	TT..CT.CCT	CTATAG
Position in codon	1333333333	3333331333	3331333333	3133333333	331333
Corresponding aminoacid residues (fig. 3)	11 3346668901 6922572521	1111111111 1234456677 5004660639	1112222222 8890111234 1510158980	2222222223 4556667990 5091780053	333333 233333 914489

924 bp have been sequenced (from 97 to 1020). Bold type marks synapomorphies. The corresponding true codon or aminoacid residues in the mt cyt b molecule are also shown (see fig. 3).

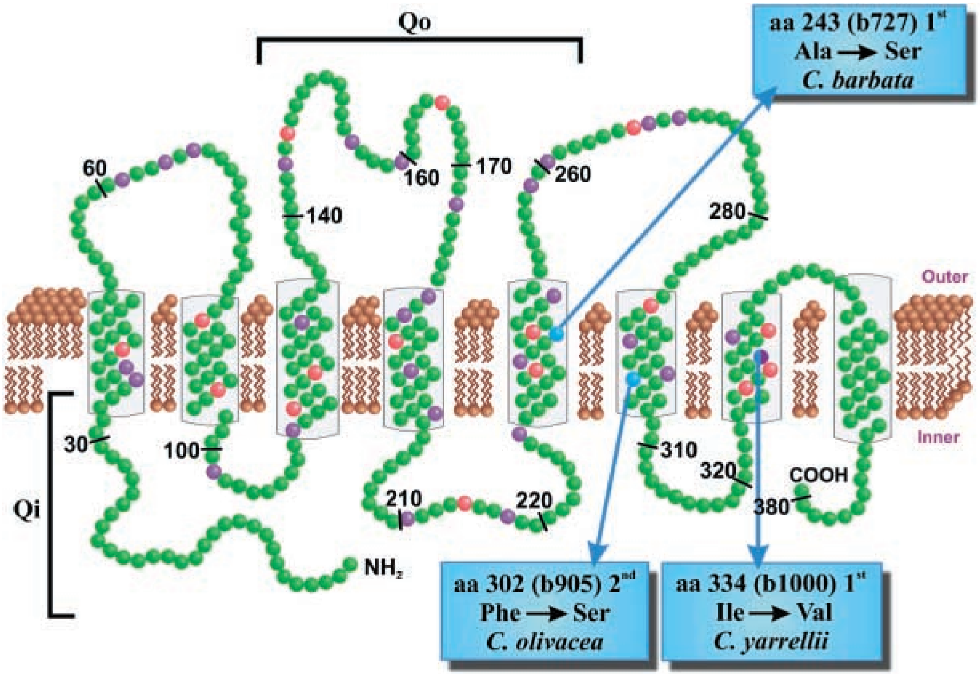

Figure 3. Cyt b molecule showing changes in South American siskins with respect to *Fringilla coelebs* and *Carduelis psaltria*, which are also included. Both purple and pink dots are phylogenetically informative residues (45). Pink amino acid residues (16) are those whose third base codons have a $CI = 1$. Note that productive changes (blue) occur in *C. barbata* (first postulated speciation event), *C. olivacea* (second postulated speciation event) and *C. yarrellii* (third postulated speciation event). Siskins are compared to *C. notata* for obtaining the three productive changes. See discussion and fig. 5 (b, DNA base; aa, amino acid).

sites are accumulated at basal nodes: 9 support the South American split, and 4 the next following basal node, 2 appear at medium levels and only 1 character at a terminal node. *C. magellanica* gets no character support, nor does *crassirostris* (as in the phenetic analysis).

The NJ hypothesis differs to a greater extent from the MP tree (fig. 2B) than the phenetic analyses differ from each other. Only one taxonomic unit (*crassirostris/magellanica/yarrellii*) is maintained in both NJ and MP molecular trees.

3) Total evidence

A heuristic search was carried out using the combined 87 parsimony-informative character matrix to give one bootstrap tree (fig. 4A). Its topology is unique, and the species arrangements are intermediate to both analyses of the single phenotypic and molecular data sets. *C. psaltria* and *F. coelebs* (outgroup taxa) are joined to *C. notata* at the basal position of the South American species lineage, and to *barbata* as the next relative, thus following the molecular model. *C. olivacea* at a basal topology is unique. These results are also supported by the NJ tree (fig. 4B): the sister relationships *crassirostris/cucullata/spinescens* are also maintained. The cluster *magellanica/yarrellii* is also seen in the molecular analysis and the unit *xanthogastra/atrata* is obtained in the phenetic data analysis. Again, synapomorphies are clustered at the basal nodes. Two molecular synapomorphies that supported the tree in the single data set analysis were lost and changed to homoplasies (fig. 4A, table 1). Furthermore, one character was lost from the morphological portion, but two others were gained, thus resulting on 15 molecular to 13 phenetic synapomorphic characters (fig. 4B, table 2).

Comparison of analyses

1) Tree length statistics

A homogeneity test (incongruence-length difference test [29]) assessed congruence between the male and female partitions of the phenetic data set (three sex-independant characters were added to both partitions) at $p < 0.27$, but not for the molecular versus phenetic portions in the total evidence data analyses, that differed significantly at $p < 0.001$.

Trees were also built from the two sex-based phenetic modules, as from the total phenetic and molecular data sets, and mapped on each other, exhibiting significant differences in pairwise winning-sites tests (table 3). Molecular and phenetic trees differed from each other with one exception. Molecular trees matched female trees if tested on the female data set. In addition the female tree differed significantly from the male tree when mapped on the male data, and phenetic trees based on male plus female characters did not differ significantly, in either tree

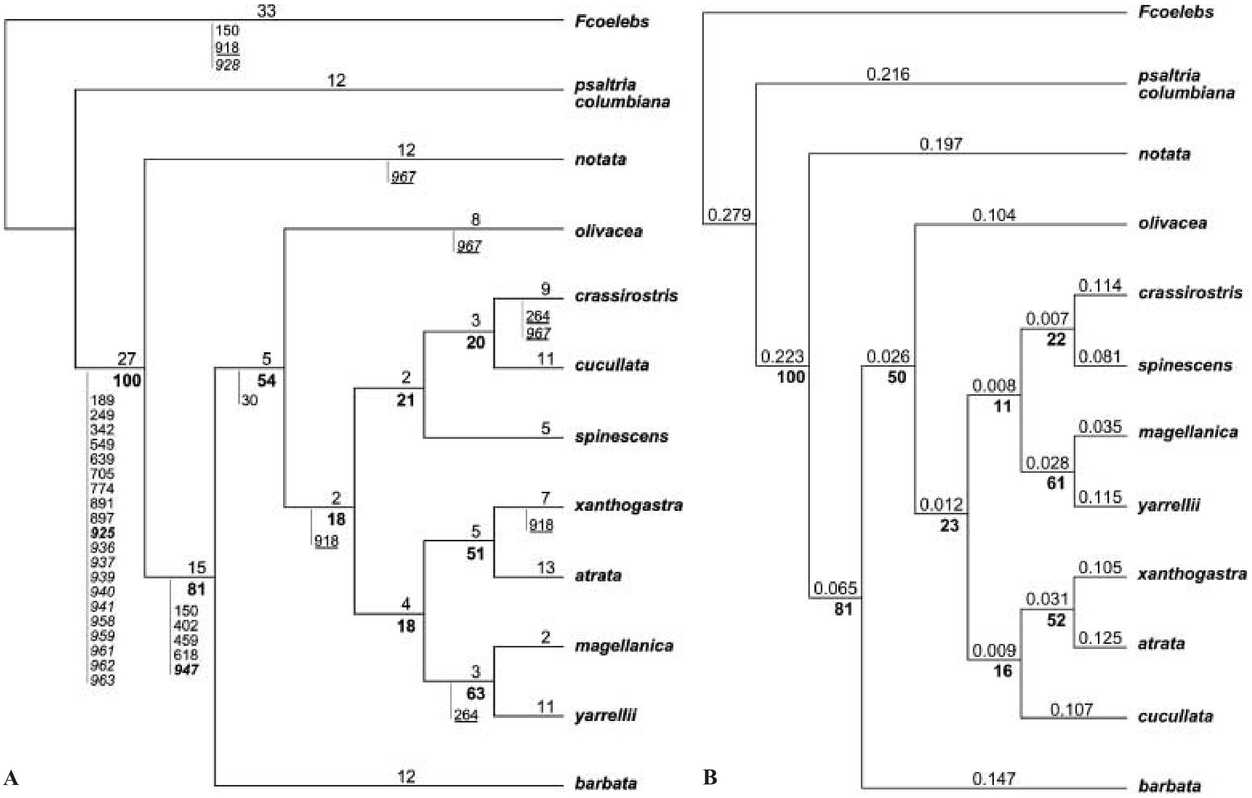

Figure 4. *A* Total evidence MP bootstrap consensus tree (1000 replicates) based on 87 parsimony-informative characters (201 steps, CI=0.5174, RI=0.4393). Only synapomorphies from the single data sets are shown. Morphological characters are in italics. Changes in character status (new synapomorphies) are in bold type. Changes to homoplasies are underlined. Branch lengths are given above the lines, bootstrap percentages are shown in bold below branches. *B* NJ bootstrap consensus tree (1000 replicates) based on 971 characters.

length or character distribution along branches, from trees built from either data set. Also the total evidence MP tree was compared to female, male, overall phenetics and molecular matrices. Only female comparison was significantly different (not shown in table 3).

If tree length is accepted as a measure of evolutionary change, incongruence between trees built from different data sets may be interpreted as differences in evolutionary steps manifested in these characters. The anagenetic plus phylogenetic information content of the phenetic and molecular data set thus differs significantly, with the exception of the female phenotype that contains the largest amount of phylogenetic information. Male and thus overall phenetics include a significantly higher proportion of anagenetic information. This would imply that morphological change is faster in male than in female ornamentation in South American siskins; it may indicate the presence of sexual selection driven by female mate choice. Also shown is that the molecular partition in the combined data set is more informative, because the molecular tree could be explained by characters from the female phenotype, but male and overall phenetics could not be reconstructed using the molecular modules, due to obvious differences in evolutionary rates.

Taxonomic congruence

The topologies of taxon arrangements in both the phenetic and molecular trees are significantly congruent with

Table 3. Results of winning-sites test for three morphological, one molecular and total evidence matrices (left to right) mapped on male, female, overall phenetic, molecular and total evidence trees (top to bottom).

Matrix	Male	Female	Overall phenetics	Molecular
Male	–	$p < 0.0352$*	$P < 1.000$	$p < 0.0034$**
Female	$p < 0.1797$	–	$P < 1.000$	$p < 0.1094$
Overall phenetics	$p < 0.5078$	$p < 0.0963$	–	$p < 0.0127$*
Molecular	$p < 0.0001$**	$p < 0.0001$**	$p < 0.0001$**	–

the total evidence hypothesis (Spearman rank correlations 0.8322, $p > 0.0058^{**}$ for the molecular; 0.6941, $p > 0.0213^{*}$ for the phenetic tree topology). Concordant taxon positions retained in all analyses are found in basal and uppermost nodes. At the base, *C. psaltria* is confirmed in all trees as a second outgroup. *Notata* (in its ancestral position) is congruent with the molecular NJ topology. *Notata* plus *barbata* placing is congruent with the molecular MP tree. The arrangement of *magellanica/yarrellii* and *xanthogastra/atrata* as terminal sister taxa is consistent in distance as well as in parsimony methods. The first species pair is retained in all molecular and the second pair in the phenetic trees. Topological incongruence between molecular and phenetic analyses on the one hand and between sexes on the other (trees not shown) is not only revealed at the deep nodes – e.g., the basal position and ancestral status is taken by *barbata* in the 'female phenetic' and overall phenetic tree instead of *notata* in the molecular or *xanthogastra/atrata* in the 'male phenetic' tree – but also in completely different arrangements of species along the most shallow branches.

Character analysis

According to cladistics [30, 31], only synapomorphies (shared derived characters) contain information about the history of speciation. These synapomorphies are indicators of monophyly in taxa, as inherited from a common ancestor . Characters or character states shared with the outgroup are plesiomorphic. Nodes are concordantly established in all analyses by the pattern already described: as synapomorphies unambiguous consistent characters are always concentrated at the deepest node. Terminal nodes have, if any, only support from single consistent characters (see figs. 1 A, 2 A, 4 A).

1) Phenetic characters

Phenetic synapomorphies (table 1) form a wing and tail pattern, stabilizing the South American species group together. These characters are demonstrated during courtship display by males. Identical patterns are found in some Asiatic greenfinches and also in female siskins, but are not as exaggerated and contrasted by surrounding black plumage colours as in male South American siskins and never demonstrated in displays. Subordinate and female siskin individuals even hide these traits, as shown by van den Elzen and Sternheim [32]. This paper put forward a phylogeny, established on traits of courtship behaviour in 11 *Spinus* and 3 other *Carduelis* species using maximum parsimony. A consensus tree defined three species clusters: (i) the *Carduelis (Spinus) magellanica* group that included all hooded taxa under study (*magellanica, notata, atrata*, *cucullata*); (ii) *C. barbata* and *spinus*, that are morphologically characterized by a cap and a bib came up as entities with similar courtship behaviours and (iii) the courtship display of *C. psaltria* was closer to the European goldfinch *C. carduelis* than to other siskin taxa. Six basal taxa could not be aligned due to missing information or because they were characterized only by symplesiomorphies, and included all the capped siskin taxa under study (*C. lawrencei, tristis, yarrellii, pinus*). In summary, the behavioural analysis corroborates the NJ phenetic tree and the phenetic synapomorphies. Thus, morphological traits that have a function in courtship display and/or species recognition were found to be phylogenetically informative; this further supports the hypothesis that speciation has been driven by sexual selection in the South American siskins.

Black caps, hoods and bibs that have been used to establish arrangements of species groups in traditional systematics proved to be homoplasious. These traits may function in a social context, as most siskin species gather in large numbers outside the breeding season, also forming mixed flocks that consist of different species. Senar and colleagues [33, 34] showed in experiments that enlargements or removals of black bibs in the European siskin *C. spinus*, (a species resembling the ancestral *C. barbata* in this trait) have an effect on social status. Large bibs in *spinus* signalled overt aggression and thus dominant status. In first encounters, siskins were able to discriminate the bib and preferred to join subordinate individuals with smaller bibs. The ancestral phenotype of the South American species group was hooded and sexually monomorphic; this is inferred from the basal position and contemporary appearance of *notata.* The hood was a new feature for the clade, but developed independently in different lineages, due to convergent evolution. Less exposed body regions that apparently have no function in social or sexual context lack special patterns. Sexual monomorphism (*atrata*, *crassirostris, notata*) as well as extreme melanism (*atrata*, *xanthogastra*) occur in species living at high altitudes (from 1500 to above 3000 m). Further studies should investigate if these characters are correlated to ecological or life history traits of coexisting siskin species.

2) Molecular characters

Mitochondrial DNA sequences have proven to be helpful for defining evolutionary relationships among relatively distant and closely related birds and other species [1, 35, 36]. Cyt b is one of the best-known proteins making up complex III of the mitochondrial oxidative phosphorylation system that is encoded by the mitochondrial genome [37]. Cyt b contains two redox centers, Q_0 and Q_i, involved in electron transfer [37, 38], where mutational and evolutionary studies have facilitated the development of a model structure of cyt b as well as the definition of the sites of electron transfer and inhibitor action.

If both *F. coelebs* and *C. psaltria* are taken into account, 46 phylogenetically informative sites are found, most of them being silent positions located in the third codon position (table 2). Forty-five residues are encoded by these codons [purple and pink residues, fig. 3; (sites 904 and 906 are located in the same codon (334), see table 2]. Twenty-six of the 45 residues are located in the transmembrane domains of the molecule, 5 are located in the intramitochondrial part and 14 in the extramitochondrial part of the cyt b molecule (fig. 3). There are 16 unambiguous molecular synapomorphies (table 2, bold type; fig. 3, pink residues). All of these correspond to third codon sites and to silent positions. The residues encoded by these codons are located (12 of them) within the transmembrane segments of the protein. The regions that had been defined as Q_0 and Q_i are highly conserved within mammals and this appears to be a major contributor to the reduced evolutionary rate found in the molecule outer surface [35]. If *F. coelebs* and *C. psaltria* are not included in the calculations (i.e. if only the ten species of South American siskins are used) 18 phylogenetically informative sites and only 3 productive changes in the molecule are found (fig. 3, blue residues). These three productive changes in amino acid residues within the South American siskin group are placed in the transmembrane segments of the protein, in positions 243, 302 and 334 (in *C. barbata, C. olivacea* and *C. yarrellii*, respectively). In mammals, these particular residues have also been defined as hypervariable [35]. The present data suggest that a similar pattern also exists in birds.

Distribution patterns of South American siskins

Phylogeny may also be inferred from distribution patterns: range size and relation to ranges of other taxa. Vicariance biogeography [39] assumes that speciation proceeds in geographical isolation, the arising new species are vicariants, allopatric in distribution and similar in their ecological demands. Consequently, distribution ranges of the phylogenetically youngest sister taxa in siskins are expected to be disjunct or allopatric. An alternative hypothesis, the gradient model [40], is based on a parapatric speciation mode, where new species emerge from a continuous distribution along either climatic, geographic or specific orographic gradients. Applied to distribution patterns of a clade, the youngest sister taxa are expected to exist in a parapatric or sympatric distribution.

Speciation events of forest birds in the Andes have been thoroughly summarized by García-Moreno and Fjeldså [21]. The authors concluded that (i) the diversification of the Andean avifauna proceeded throughout the period of uplift and gradual formation of new habitats in the northern Andes since the Miocene epoch, (ii) the typical mode of speciation was both allopatric or parapatric (the latter also suggested for rodents [22, 41]) and (iii) species with small distribution ranges are in the recent clades, and sympatry and range extensions are in the older clades. This model cannot be applied without limitations to the siskin clade, because its species are neither restricted to the Andes nor are they all forest dwellers; they inhabit a variety of suitable habitats including landscapes from montane moist to dry woodland and secondary bushland at low elevations with the exception of the Amazon basin.

The majority of siskins are concentrated in mountainous habitats along the Andean Cordilleras. With the exception of a minor area of sympatry in the coastal areas of Ecuador, two major and three minor distribution hotspots (fig. 5) are localized in the Andes. Two smaller ones are focussed in the northern and north-eastern Andes, where four species occur in parapatry to sympatry, if *C. psaltria* with its South American ssp. *columbiana* is included. The major hotspot focusses in the Peruvian Andes, a centre of species richness and high avian endemism also documented for other bird taxa [42, 43]. Five of the 13 recognized extant South American siskin species are found here in mostly allopatric or parapatric breeding ranges, separated from each other either along elevational gradients of the eastern or western Andean Cordilleras, or chained on top of mountain peaks or continuously along the climatically contrasted western or eastern slopes [44]. Five species occur out of the main Andean Cordilleras, one of these reaching the Falkland Islands in the southernmost parts of the continent (*barbata*) and another the eastern lowlands in Brazil (*yarrellii*).

Timing of the siskin radiation for the Miocene as traced by molecular markers [1] meets the proposed estimate for Andean forest bird speciation [21]. The occurrence of allo- or parapatric speciation events can be traced by comparison of species pairs in the shallow branches only, as actual distribution patterns of phylogenetically older species may have changed in range or shifted over geological time. Distribution patterns observed for the youngest siskins are heterogeneous and do not allow the assignation of speciation model: present-day distribution patterns arranged according to total evidence topology suggest that speciation events started with a north to south separation in siskins (*notata/barbata*), most probably along the eastern Andean Cordilleras, where the second split (*olivacea*) is documented according to the present-day siskin distribution. Three sister species pairs (AB, CD, EFG; fig. 5) are shown to be the most recent speciation events: the first sister group comprises taxa with allopatric ranges in the northern and eastern Andes (*xanthogastra, atrata*) and the High Andes (*atrata*), the second (*magellanica/yarrellii*) exhibits a cis-trans Andean separation within the range of one species (*magellanica*) and a lowland split at the eastern end (*yarrellii*), and the third sister group exhibits a northern midland and eastern High Andean disjunction (*cucullata/crassirostris*), which is mediated by the northern Andean *spinescens*.

The largest distribution range size, a prerequisite of phylogenetically old mountain forest species [21], is found in

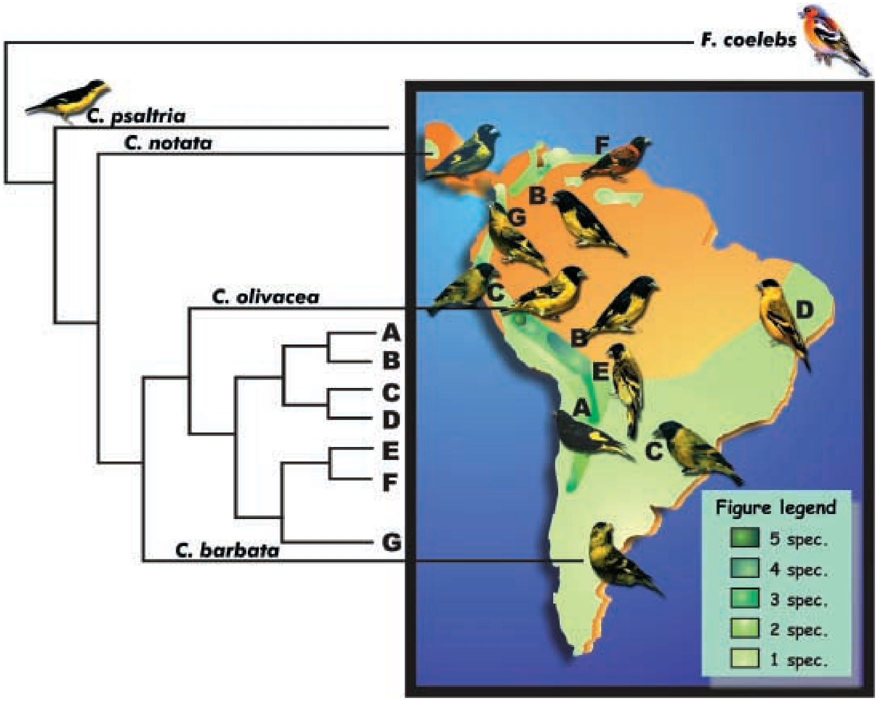

Figure 5. Total evidence MP bootstrap consensus tree mapped on South American siskin hotspots and distributions. A, *Carduelis atrata*; B, *C. xanthogastra*; C, *C. magellanica*; D, *C. yarrellii*; E, *C. crassirostris*; F, *C. cucullata*; G, *C. spinescens*. Darker-green colours show evolutionary 'hotspots' as defined by Fjeldså et al. [43] for other South American birds. 'Hotspots' hold maximum numbers of siskins species; the dark- to lighter-green gradient indicates a decrease in species numbers [43].

a terminal, phylogenetically recent species *magellanica*, which also exhibits great ecological plasticity, being found in a variety of climatic zones and habitats. Small distribution ranges, representative for recent forest species, tend in siskins to be found in species at the deeper nodes, but exceptions occur. A broad range overlap is generally only observed between 5 of 45 possible species combinations among the siskins under study and in younger as well as in phylogenetically middle-aged taxa, but not between young sister species. The distant, phylogenetically oldest siskins are not true vicariants, as their distribution ranges are disjunct, but their ecological demands differ.

In the Peruvian siskin diversity centre, species from each of the three terminal sister groups as well as their parental species *olivacea* are represented. The coincidence that taxa with relationships to the 'hotspot' area are combined by nodes with the weakest support in the total evidence tree may suggest an explosive radiation of this lineage in allo- to parapatric situations within this area and favours the gradient speciation model and differentiation of phenotypes by sexual selection in ecological rather than geographical isolation. On the other hand, the area is not only a hotspot of species richness and, in siskins, of phenetic differentiation but also of ecological instability [43]. The observed patterns of parapatric range overlap might also be based on ecological diversity that, in combination with phenotypic separation, allows coexistence and parapatry in secondary contact zones.

A speciation model driven by sexual selection through female mating choice is favoured in siskins because the morphological change is faster in male than in female ornamentation. It presently offers the most convincing explanation for our finding that morphological traits (shared in both sexes and functioning in courtship behaviour and thus species recognition) are also reflected in molecular patterns. Female phenotype correlation with the molecular tree may be explained by female selection and pressure on male traits but also by maternal mt cyt b inheritance. The convergent appearance of morphological traits (homoplasies) like black caps or hooded females may be just due to the fact that convergence would only apply to different taxa and not to sexes and/or that there is a genetic correlation between the expression in males and females. These traits are related to specific ecological and/or high-altitude conditions in restricted-range species, and fixation of alleles might be remarkably accelerated by positive phenotype matching. This is possible because siskins are not long-distance migrants and are only exposed to local movements. In a vagrant, Carduelinae species, the European crossbill, *Loxia curvirostra,* vocalizations are thought to trigger speciation [45]. The authors suggest that dialect formation and assortative mating among birds of the same vocal type facilitate the generation of new species, provided that sufficient opportunities for ecological differentiation are present.

Acknowledgements. We thank the following persons and institutions for their support in loan of materials (institutions in alphabetical order); Frank B. Gill, Academy of Natural Sciences, Philadelphia; Christine Blake, American Museum of Natural History, New York; Mark Adams, Michael Walters, Bird Group, British Museum (Natural History), Tring; Kenneth C. Parkes, Carnegie Museum, Pittsburgh; David Willard, Field Museum, Chicago; Raymond A. Paynter Jr, Harvard University, Cambridge, Mass.; James Dean, National Museum of Natural History, Smithsonian Institution, Washington; Janet Hinshaw, University of Michigan, Ann Arbor; Charles Sibley, Peabody Museum, Yale University. For support in analyses, preparation of figures and comments on the paper we also thank (in alphabetical order): Andreas Fey, Doris Hölling, Bernhard Misof, Stefanie Rick, Andreas Schmitz, Bradley Sinclair and Uwe Vaartjes (Bonn, Germany). We are also indebted to the following Spanish ornithologists: Bernardino Yebes, Gloria Gardó, Francisco Mira Chinchilla, Arturo Fernández Cagiao and Alvaro Guillén. Alberto García helped with the computer-designed figures. This work was supported in part by grants from the Spanish Ministry of Education (PM95-57, PM96-21 and PM99-23) and Madrid Regional Government (06/70/97 and 8.3/14/98).

1 Arnaiz-Villena A., Alvarez-Tejado M., Ruíz-del-Valle V., Garcia-de-la-Torre C., Varela P., Recio M. J et al. (1998) Phylogeny and rapid northern and southern hemisphere speciation of goldfinches during the Miocene and Pliocene epochs. Cell. Mol. Life Sci. **54:** 1031–1041

2 Sibley C. G. and Monroe B. L. (1990) Distribution and Taxonomy of Birds of the World, Yale University Press, New Haven, Conn.

3 Wolters H. E. (1975–1982) Die Vogelarten der Erde, Parey, Hamburg, Germany

4 Kluge A. G. (1989) A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). Syst. Zool. **38:** 7–25

5 Eernisse D. J. and Kluge A. (1993) Taxonomic congruence versus total evidence, and amniote phylogeny inferred from fossils, molecules and morphology. Mol. Biol. Evol. **10:** 1170–1195

6 Mattern M. Y. and McLennan D. A. (2000) Phylogeny and speciation of felids. Cladistics **16:** 232–253

7 McLennan D. A. and Mattern M. Y. (2001) The phylogeny of the gasteroidae: combining behavioral and morphological data sets. Cladistics **17:** 11–27

8 Chippindale P. T. and Wiens J. J. (1994) Weighting, partitioning, and combining characters in phylogenetic analysis. Syst. Biol. **43:** 278–287

9 Zink R. M. and Blackwell R. C. (1996) Patterns of allozyme, mitochondrial DNA, and morphometric variation in four sparrow genera. Auk **113:** 59–67

10 Griffiths C. (1999) Phylogeny of the Falconidae inferred from molecular and morphological data. Auk **116:** 116–130

11 Johnson K. P. and Sorenson M. D. (1999) Phylogeny and biogeography of dabbling ducks (genus *Anas*): a comparison of molecular and morphological data. Auk **116:** 792–805

12 Bull J., Huelsenbeck J., Cunningham C., Swofford D. and Waddell P. (1993) Partitioning and combining data in phylogenetic analysis. Syst. Biol. **42:** 384–397

13 Gatesy J., O'Grady P. and Baker R. H. (1999) Corroboration among data sets in simultaneous analysis: hidden support for phylogenetic relationships among higher level artiodactyl taxa. Cladistics **15 :** 271–313

14 Turner H. and Zandee R. (1995) The behaviour of Goloboff's tree fitness measure F. Cladistics **11:** 57–72

15 Sibley C. G. and Ahlquist J. E. (1987) Avian phylogeny reconstructed from comparisons of the genetic material, DNA. In: Molecules and Morphology: Conflict or Compromise? Patterson C. (ed.), Cambridge University Press, Cambridge, UK

16 Zink R. M. and Avise J. C. (1990) Patterns of mitochondrial DNA and allozyme variation in the avian genus *Ammodramus.* Syst. Zool. **39:** 148–161

17 Hedges S. B. and Maxson L. R. (1996) Re: molecules and morphology in amniote phylogeny. Mol. Phylogenet. Evol. **6:** 312–314

18 Badyaev A. V. and Hill G. E. (1999) Variation in avian sexual dichromatism in relation to phylogeny and ecology. In: Proceedings of the 22nd International Ornithological Congress, Durban, pp. 1687–1705, Adams N. J. and Slotow R. H. (eds), BirdLife South Africa, Johannesburg

19 Simmons M. P. (2000) A fundamental problem with amino-acid-sequence characters for phylogenetic analyses. Cladistics **16:** 274–282

20 McCracken K. G. and Sheldon F. H. (1998) Molecular and osteological heron phylogenies: sources of incongruence. Auk **115:** 127–141

21 García-Moreno J. and Fjeldså J. (2000) Chronology and mode of speciation in the Andean avifauna. In: Isolated Vertebrate Communities in the Tropics, Zoological Monographs 46, Bonn

22 Patton J. L. and Smith M. F. (1992) mtDNA phylogeny and Andean mice: a test of diversification across ecological gradients. Evolution **46:** 174–183

23 Swofford D. L. (1998) Phylogenetic Analysis Using Parsimony (PAUP*). Illinois Natural History Survey, Urbana

24 Clement P., Harris A. and Davis J. (eds) (1993) Finches and Sparrows: An Identification Guide, Helm and Black, London

25 Fjeldså J. and Krabbe N (eds) (1990) Birds of the High Andes, Apollobooks, Svendborg

26 Ridgely R. S. and Tudor G. (eds) (1989) The Birds of South America, volume I, The Oscine Passerines, Oxford University Press, Oxford

27 Todd W. E. C. (1926) A study of the Neotropical finches of the Genus *Spinus*. Ann. Carnegie Mus. **17:** 11–82

28 BirdLife International (2000) Threatened Birds of the World, Lynx Edicions and BirdLife International, Barcelona, Spain and Cambridge, UK

29 Farris J. S., Källerssjö M., Kluge A. G. and Bult C. (1995) Testing significance of incongruence. Cladistics **10:** 315–319

30 Hennig W. (ed.) (1950) Grundzüge einer Theorie der phylogenetischen Systematik, Deutscher Zentralverlag, Berlin, Germany

31 Hennig W. (ed.) (1966) Phylogenetic Systematics. University of Illinois Press, Urbana

32 Elzen R. van den and Sternheim B. (1992) Courtship display and systematic position of the hooded siskin, *Spinus magellanicus* (Vieill., 1805). Ornitol. Neotrop. **3:** 35–44

33 Senar J. C., Camerino M., Copete J. L. and Metcalf N. B. (1993) Variation in the black bib of the Eurasian siskin (*Carduelis spinus*) and its role as a reliable badge of dominance. Auk **110:** 924–927

34 Senar J. C. and Camerino M. (1998) Status signalling and the ability to recognize dominants: an experiment with siskins (*Carduelis spinus*). Proc. R. Soc. Lond. B **265:** 1515–520

35 Irwin D. M., Kocher T. D. and Wilson A. C. (1991) Evolution of the cytochrome b gene of mammals. J. Mol. Evol. **32:** 128–144

36 Arnaiz-Villena A., Alvarez-Tejado M., Ruíz-del-Valle V., Garcia-de-la-Torre C., Varela P., Recio M. J. et al. (1999) Rapid radiation of canaries (genus *Serinus*). Mol. Biol. Evol. **16:** 2–11

37 Hatefi Y. (1985) The mitochondrial electron transport and oxidative phosphorylation system. Annu. Rev. Biochem. **54:** 1015–1069

38 Howell N. and Gilbert K. (1988) Mutational analysis of the mouse mitochondrial cytochrome b gene. J. Mol. Biol. **203:** 607–618

39 Nelson G. and Platnick G. (eds.) (1981) Systematic and Biogeography, Cladistics and Vicariance, Columbia University Press, New York

40 Endler J. E. (ed.) (1980) Geographic Variation, Speciation and Clines, Monogr. Popul. Biol 10, 2nd ed, Princeton University Press, Princeton, NJ

41 Patton J. L., Myers P. and Smith M. F. (1990) Vicariant versus gradient models of diversification: the small mammal fauna of eastern Andean slopes of Peru. In: Vertebrates in the Tropics, Peters G. and Hutterer R. (eds), Alexander Koenig Zoological Reseach Institute and Zoological Museum, Bonn, Germany

42 Bibby C., Collar N. J., Crosby M. J., Heath M. F., Imboden C., Johnsons T. H. et al. (eds) (1992) Putting Biodiversity on the Map: Priority Areas for Global Preservation, ICBP, Cambridge, UK

43 Fjeldså J., Lambin E. and Mertens B. (1999) Correlation between endemism and local ecoclimatic stability by comparing Andean bird distributions and remotely sensed land surface data. Ecographica **22:** 63–78

44 Parker T. A., Allen Parker S. and Plenge M. A. (eds) (1982) An Annotated Checklist of Peruvian Birds, Buteo, Vermillion, South Dakota

45 Galis F. and Alphen J. M. van (2000) How fast do crossbills speciate? On assortative mating and vocalizations. Trends Ecol. Evol. **15:** 357