



# Cladistic analysis of distributions and endemism (CADE): using raw distributions of birds to unravel the biogeography of the South American aridlands

Ana Luz Porzecanski\* and Joel Cracraft

Division of Vertebrate Zoology, Department of Ornithology, American Museum of Natural History, and Department of Ecology, Evolution, and Environmental Biology, Columbia University, New York, NY, USA

## ABSTRACT

**Aim** To use published distributional and taxonomic information on Neotropical birds in a cladistic analysis of distributions and endemism (CADE) to generate a testable hypothesis of area-relationships for the arid areas of endemism, particularly those of Central South America (the ‘arid diagonal’), and to clarify the different methodologies commonly associated with parsimony analysis of endemism (PAE).

**Location** South America.

**Methods** Cladistic analysis of distributions and endemism.

**Results** We obtain a tree where the relationships of most areas are resolved, and we find support for an exclusive clade of Central South American areas, with the Caatinga as sister to both the Chaco and Cerrado.

**Main conclusions** There is a substantial amount of historical signal in avian distributions, when large numbers of taxa and multiple taxonomic levels are considered. Ecological noise and historical information are more easily distinguished in CADE analyses than they would be in PAE analyses. Based on our results we predict that among aridland birds, the Cerrado and Chaco species will be more closely related to each other than to Caatinga species.

## Keywords

Areas of endemism, CADE, Caatinga, Cerrado, Chaco, cladistic analysis, distributional analysis, parsimony analysis of endemism, South America.

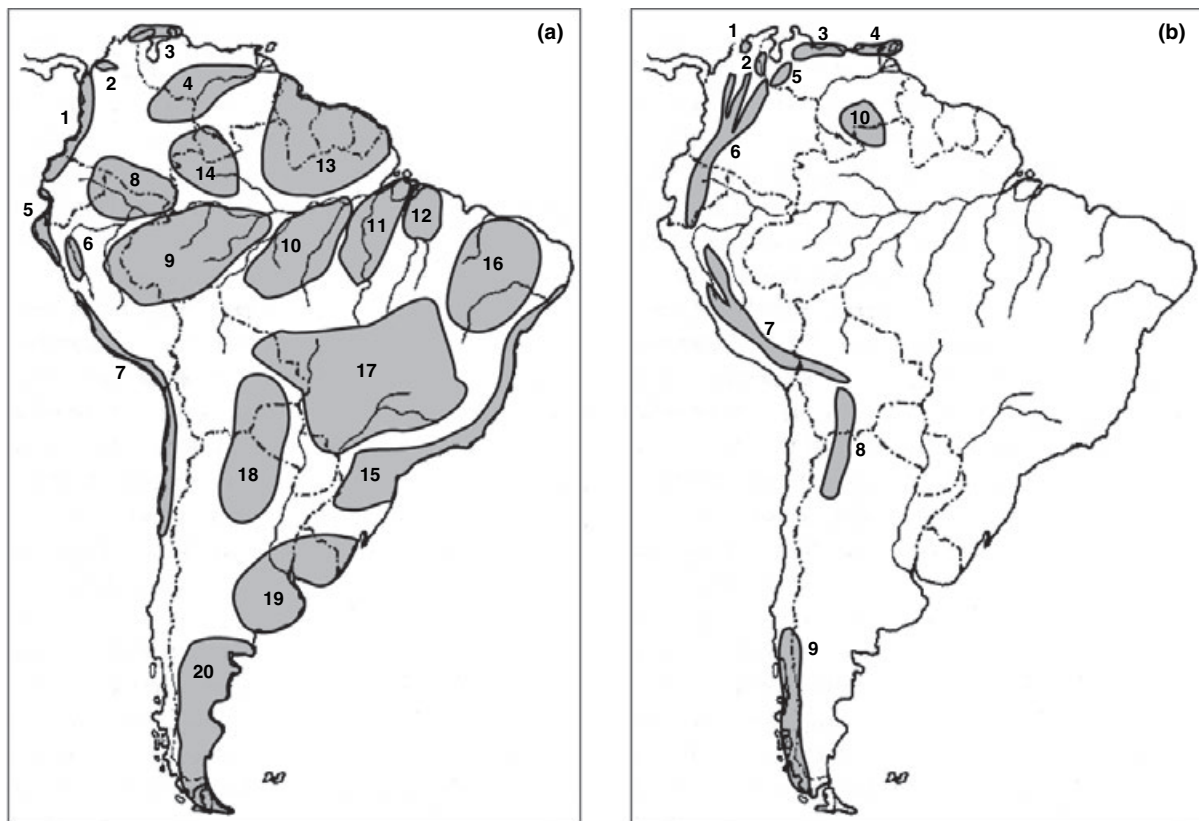
\*Correspondence and present address: Ana Luz Porzecanski, Center for Biodiversity and Conservation, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024-5192, USA.  
E-mail: alporze@amnh.org

## INTRODUCTION

The South American continent contains a variety of biomes that together harbour the most diverse biota, and the largest avifauna, on the planet (Ridgely & Tudor, 1989, 1994). Different classifications have been applied to this habitat diversity (Hueck & Seibert, 1972; Ab’Sáber, 1977; Mares, 1992; Dinerstein *et al.*, 1995), yet a key feature of the South American landscape is recognized by all: it is roughly divided into wet forest (lowland and montane rain and semideciduous forests) and arid (grasslands, savannas, shrublands, and dry forests) areas. The patterns of animal distributions associated with this mosaic of vegetation types can be exceedingly complex, but endemic sets of taxa have permitted the recognition of numerous areas of distributional

congruence, or areas of endemism. For many vertebrate groups these are largely coincident with the biomes defined on the basis of vegetation. Cracraft (1985) delineated 33 areas of endemism for the South American continent, all of them based on relatively large numbers of endemics, and Haffer (1985) performed a similar analysis focusing on the lowland forests (Fig. 1).

In spite of a long history of describing patterns of avian endemism, we still have a poor understanding of their origin and history (Cracraft, 1988; Cracraft & Prum, 1988; Bates *et al.*, 1998). Many hypotheses have been proposed to explain the high levels of diversity and endemism found in the Neotropical forests (see reviews by Haffer, 1997; Hooghiemstra & van der Hammen, 1998; Patton & Silva, 1998), and different hypotheses of area relationships have been proposed for the



**Figure 1** Areas of endemism (AE) described for birds, following Cracraft (1985) and Haffer (1985), modified from Ridgely & Tudor (1989). (a) Lowland AE: 1. Chocó, 2. Nechí, 3. Guajira, 4. Llanos, 5. Tumbesian, 6. Marañón, 7. Peruvian coastal, 8. Napo, 9. Inambari, 10. Rondônia, 11. Pará, 12. Belém, 13. Guyanan, 14. Imerí, 15. Serra do Mar + Paraná, 16. Caatinga, 17. Cerrado, 18. Chaco, 19. Pampas, 20. Patagonian. (b) Montane AE: 1. Santa Marta, 2. Periján Montane, 3. Venezuelan Montane, 4. Paría Montane, 5. Meridan Montane, 6. North Andean, 7. Peruvian Andean, 8. Austral Andean, 9. Nothofagus, 10. Pantepui.

forest areas of endemism based on phylogenies of birds (Cracraft, 1988; Cracraft & Prum, 1988; Prum, 1988; Hackett, 1993; Brumfield & Capparella, 1996). By comparison, the endemic avifauna of the aridlands has seldom been studied.

Many biogeographic methods have been developed to reconstruct the history of biotas and areas of endemism (reviewed in Crisci, 2001). There is general agreement that the best evidence for area relationships is congruent phylogenetic patterns among the endemic species from multiple clades (Nelson & Platnick, 1981). This evidence is often lacking, however, especially across large spatial scales such as those characteristic of continental or oceanic biotas, and in these cases heuristic alternatives to cladistic biogeographic methods have been suggested. The most common category of alternative approach is to use raw distributional data and summarize area patterns in a branching diagram using one or more hierarchical clustering algorithms. An important problem with this set of approaches is that not all methods are historically-based, either in the way raw distributional data are collected or in the manner in which they are analysed. Compounding the problem is the fact that different conceptual approaches may be referred to by the same name. All this leads to analytical and interpretative confusion.

In this paper, we employ a method of exploring historical relationships among areas of endemism using raw distributions that was originally described (Cracraft, 1991) as an alternative to the parsimony analysis of endemism (PAE) method of Rosen (1988; also Rosen & Smith, 1988). Since 1991, however, these two methods have been lumped together under the term PAE and methods ascribed to that concept have varied widely. In order to bring some analytical and nomenclatural clarity to this situation we here name the method of Cracraft (1991): cladistic analysis of distributions and endemism (CADE). Before using this method to investigate relationships among the South American arid areas of endemism, we summarize and attempt to clarify some of the confusion surrounding the different 'PAE' methods in use.

### PAE and CADE: sorting through the nomenclatural confusion

#### *Parsimony analysis of endemism*

The paleontologist Brian Rosen (1988, 1992; Rosen & Smith, 1988) originally developed PAE to address the shortcomings of phenetic approaches used to assess area relationships of fossil

or Recent assemblages from different areas. The practice of clustering floral or faunal samples on the basis of similarity or dissimilarity measures has had a long history (summarized in Rosen, 1988; see also, for example, Sneath & Sokal, 1973; Hengeveld, 1990). Rosen clearly recognized that historical interpretations of area relationships using phenetic clustering based on similarity or dissimilarity distances were limited. Nevertheless, PAE was derivative of those previous methods in that it used shared taxic similarity (taxon samples) from point localities, or from a series of collection localities, rather than predetermined areas of endemism. Indeed, Rosen proposed examining the relationships of areas of endemism through the use of point-locality samples rather than through endemic taxa *per se* (Rosen, 1988, p. 456). Many of the coral assemblages studied by Rosen could be said to represent areas of endemism only as a consequence of their distribution on isolated islands.

Parsimony analysis of endemism begins by compiling presence/absence information for taxa at sample localities, and then uses parsimony analysis to group the localities into a cladogram. Clustering of sample localities arises as a result of the shared taxa acting as 'synapomorphies,' or 'synendemism' (Rosen, 1992). Rosen (1988, 1992) noted that the resulting cladogram can either be interpreted as a pattern of ecological similarities among the sample localities or as reflecting historical patterns that were due to geological change. If one assumes that patterns of similarity should reflect aspects of the geological history of the areas, then a historical interpretation is given to clusters of point localities. These area relationships can then be compared with evidence from the Earth sciences (e.g. Waggoner, 1999).

#### *Cladistic analysis of distributions and endemism*

Approaching the problem of area relationships from the tradition of cladistic biogeography, Cracraft (1991) reasoned that if elements of biotas were subject to a common history of vicariance over time, raw distribution patterns should contain hierarchical information about that history. This is based on a general assumption of vicariance biogeography, namely that dispersion first leads to relative degrees of cosmopolitanism in biotas, that some of the taxa then differentiate in response to vicariance events, and finally these events divide the biota into areas of endemism. Thus, a taxon shared between two areas of endemism might be taken as evidence of their close relationship relative to a third area in which the taxon is absent, if that taxon represents a failure to differentiate following vicariance of the first two areas. On the other hand, that shared taxon similarity might not be evidence of relationships if there were a long-distance dispersal event among unrelated areas, or if the similarity represented a shared 'plesiomorphic' distribution pattern (as might happen if there were extinction in other areas). If one adopts the assumption, or can empirically demonstrate through standard vicariance biogeographic methods (Nelson & Platnick, 1981), that long-distance dispersal has a minor role in structuring the areas of endemism of biotas (or the

taxa under investigation), then raw distributions should be expected to carry historical information.

In addition, Cracraft (1991) suggested that raw distributions using a single taxonomic level (e.g. species) would generally be insufficient to resolve area relationships, but that if the hierarchical information implied by accepted classification schemes, or better still, by phylogenetic relationships, were used as a framework to code distributions, historical signal about area relationships would be increased. Moreover, the expectation of recovering historical signal from raw distributions would also be sensitive to sample size; small taxon samples, as often seen in the use of PAE-related methods, would be especially susceptible to 'noisy' distributions resulting from long-distance dispersal, extinction, or taxonomic artefact. In an empirical example (Cracraft, 1991), area-cladograms derived using CADE for large sample sizes of birds and other vertebrates for Australian areas of endemism revealed considerable congruence, which also matched area relationships previously inferred from cladistic analysis of different avian taxa (Cracraft, 1986). In another use of CADE for species and subspecies of all Neotropical perching birds in 12 forest areas of endemism, Bates *et al.* (1998) found incongruence between their results and those of area relationships determined from phylogenetic studies of the birds. This latter study highlights an additional aspect of CADE. Although it is a useful first step for biogeographical analysis, it is not designed to be an alternative to historical biogeographic methods using actual phylogenetic data. Thus, only through comparisons with data from phylogenetic analyses, or cladistic biogeographical analyses, can the degree of historical information in the raw distributions be tested.

Whereas a superficial comparison might suggest CADE and PAE are similar, the former differs in significant ways (see also Cracraft, 1991, p. 215). First, in CADE, areas of endemism are predetermined, based on congruence of distributions across the target taxa, thus the purpose is to extract historical relationships among these areas, not among sample localities. The notion of areas of endemism, unlike that of collecting localities, implies a putative common history for the members of these biotic assemblages, and recognizes the influence of history on common distributional pattern. As a result, CADE would be expected to minimize the clustering of areas based solely on ecological similarity, a concern of the developer of PAE (Rosen, 1988). In fact, the use of taxon samples from point localities as surrogates for the taxa found in an area of endemism could be expected to enhance incorporation of 'ecological noise' in that point samples will reflect the composition of ecological communities in local habitats rather than of areas of endemism that are often comprised of multiple habitats. Second, cladistic information is built into the data matrix by coding distributions for more inclusive hierarchical levels (e.g. species and their subspecies, genera and their species, to the extent that systematic data are available). Thus, even if present-day taxonomic knowledge does not resolve the detailed relationships among the species in a given genus, their more recent common ancestry relative to species in other

genera implies a pattern of area relationships that can be incorporated into the analysis by coding the hierarchical level above them, in this case genus. Although this is not required in PAE analyses, and has in fact been applied in some cases (e.g. Morrone, 1998), it is essential in CADE analyses. CADE, moreover, is designed to be applied to large numbers of taxa that reflect complete or substantial portions of an area's biota.

Finally, an additional application, also called PAE, has been used with a different goal: that of identifying areas of endemism. Morrone (1994) proposed a modified version of PAE to delineate areas of endemism based on the presence or absence of taxa in predefined quadrats (see also Posadas, 1996; Linder, 2001). Although Morrone's method has some similarities to PAE, it was not proposed as a method for historical inferences of area relationships and does not use predetermined, putatively historical areas as the units of analysis. It differs from CADE, further, by lacking any hierarchical taxonomic information.

### Hybrid methods and methodological comparisons

As a result of these differences between PAE and CADE, studies abound where a hybrid methodological approach is used. The units under study may be predetermined areas of some kind, such as islands (e.g. Trejo-Torres & Ackerman, 2001), interfluvial regions (e.g. Silva & Oren, 1996), localities (Raxworthy & Nussbaum, 1996; Ron, 2000), or quadrats (e.g. García-Barros *et al.*, 2002). Moreover, the data matrices used to assess 'area' relationships rarely incorporate multiple taxonomic levels (Ron, 2000 being the only exception in the list above). Interpretations of the results also vary; the analysis may be considered to yield areas of endemism, biotic affinities, or area relationships. In some of these studies, the 'PAE' method has been further modified and even renamed.

This confusion has also prompted criticisms of CADE and PAE together under the PAE label that do not acknowledge the differences among each method. Brooks & van Veller (2003), for example, argue that 'PAE' is deficient as an *a priori* method of historical biogeography. Their main four criticisms are: (1) it will only recover correct historical relationships when species undergo a particular combination of vicariance and non-response to vicariance; (2) it does not use any phylogenetic information; (3) areas may be grouped together by shared absences but character change is not examined; and (4) it is susceptible to being misled by shared episodes of postspeciation dispersal. Some of these criticisms, such as (1) and (4) are problems shared by all methods of historical biogeographical analysis. Brooks parsimony analysis (BPA; Wiley, 1988; Brooks *et al.*, 2001), for example, will also recover completely correct historical relationships only under particular speciation scenarios. Deviations from such expectations can only be detected *a posteriori*, as in CADE, component analysis, BPA, or perhaps other methods. The lack of phylogenetic information (criticism 2, above) has been the basis for other critiques of PAE (Humphries, 1989; Humphries & Parenti, 1999), but as suggested in the earlier comparison, this problem applies much

more to PAE than to CADE. Criticism 3, above, identifies a problem that can be encountered with any parsimony method that might treat shared absences as a reversal (derived loss). It is not correct to assume, however, that groups based on shared absences will necessarily go undetected; examination of character optimizations is always advisable in any parsimony method.

Given that detailed phylogenetic information is lacking, CADE is an appropriate method for examining biogeographical signal, provided hierarchical information is incorporated and sample size is large. Many studies using methods such as PAE or CADE, as well as critiques of those methods (Brooks & van Veller, 2003), employ sample sizes that are inadequate, and cannot be expected, to reveal well-supported area relationships (Cracraft, 1991, this paper). Furthermore, if character change is carefully examined, the informative and 'homoplastic' taxa for any given area relationship can be identified. Thus CADE can provide a preliminary understanding of area relationships and thereby serve as a framework for future phylogenetic studies.

### Biogeographic relationships of South American arid areas of endemism using CADE

Previous biogeographical studies of Neotropical areas, including PAE and CADE studies, have largely focused on the forest areas of endemism, and considerable uncertainty remains about the relationships of the Neotropical arid areas of endemism. In particular, two or more areas in the 'arid diagonal' of South America (Ab'Sáber, 1977), namely the Caatinga, Cerrado, and Chaco, are generally considered to comprise a biogeographic unit, harbouring presumably related avifaunas (Haffer, 1967; Cabrera & Willink, 1973; Morrone, 1999, 2000). A number of different biogeographic scenarios that could translate into different area relationships have been suggested for these areas (Haffer, 1967, 1985; Short, 1975; Nores, 1992; Silva, 1995a, 1997), but none was based on a comparative historical analysis of robust phylogenies. Our goal here is to use published distributional information for Neotropical birds (Parker *et al.*, 1996) and apply CADE to generate a hypothesis of area-relationships for the arid areas of endemism, particularly within Central South America (CSA) (the 'arid diagonal'), that will be tested by phylogenetic studies (Porzecanski, 2003). We collected data from all bird families, and this allowed us to compare the type of signal provided by different partitions of the data. We assess the congruence of the resulting hypotheses with previous biogeographic studies.

## MATERIALS AND METHODS

### Areas

Areas of endemism for Neotropical birds have been described by Cracraft (1985) and Haffer (1985), but unfortunately there is no data base of all avian distributions across these areas. Such a data base is available, however, for the 22

zoogeographic regions defined by Parker *et al.* (1996, map 1). Importantly, these regions correspond to the avian areas of endemism in most cases, including the arid areas of endemism that are the main focus of this study. A few zoogeographic regions contain more than one area of endemism and thus can be taken to represent areas of endemism at a larger scale. Thus, for example, the Amazonia North (AMN) and Amazonia South (AMS) zoogeographic regions include three and four avian areas of endemism each. Equivalencies between the zoogeographic regions of Parker *et al.* (1996) and areas of endemism (Cracraft, 1985; Haffer, 1985) are listed in Table 1 and illustrated in Fig. 2.

Most previous PAE or CADE studies included only subsets of Neotropical areas, for example only Amazonian areas, therefore making an assumption of 'monophyly' for the areas of interest. Here, no area is excluded *a priori*, and thus all Neotropical areas are available as potential relatives to the areas of CSA. For the primary area of interest, CSA, we divided the region into its three constituent subregions, the Caatinga, Cerrado, and Chaco areas (Parker *et al.*, 1996, map 3) so as to obtain an exact match with previously defined areas of endemism for birds (Cracraft, 1985; Haffer, 1985). We did not treat other regions such as AMN or AMS in the same manner because others have already explored these areas in detail (Cracraft & Prum, 1988; Silva & Oren, 1996; Bates *et al.*, 1998; Ron, 2000). Our approach resulted in a total of 24 terminal units (areas). The outgroup used to root the resulting network was an all-zero hypothetical area (Rosen, 1988; Cracraft, 1991).

## Taxa

Our approach for selecting taxa for this analysis is equivalent to asking a two part question: 'Where are the CSA taxa distributed, in addition to one or more CSA areas? And, if these taxa are endemic to CSA, where are their relatives (congeners) found?' Thus, the taxa used were those distributed in at least one of the three CSA areas (Caatinga, Chaco, or Cerrado) in Data base A of Parker *et al.* (1996, pp. 132–291). Taxa distributed exclusively in non-CSA regions, such as forest endemics, are not included in the matrix. 'Autapomorphic' taxa, those found in only one CSA area of endemism and nowhere else, and cosmopolitan taxa found in all areas, are uninformative about area relationships when in a method such as CADE and were also excluded. Distributional information was entered as presence (1) or absence (0) in a matrix using MacClade version 3.07 (Maddison & Maddison, 1992), with each column corresponding to a taxon. Each of these 'taxon-characters' therefore, provides signal for a connection between a CSA area and other areas.

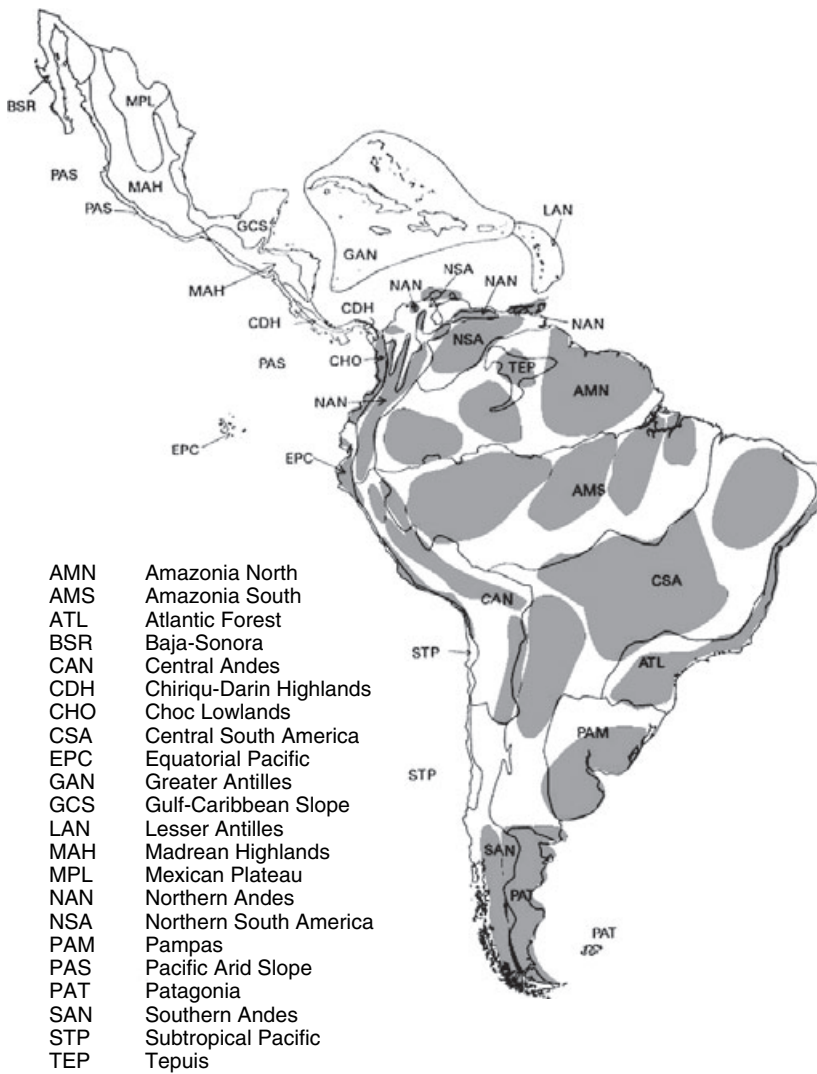
Three taxonomic levels were included: subspecies, species, and when they contained a taxon endemic to a CSA area, genera. By including genera in these instances the area relationships of narrow endemics or 'autapomorphic taxa' were incorporated into the analysis. A step-by-step example illustrating these coding procedures is provided in Fig. 3.

Presence or absence was coded as specified in Parker *et al.* (1996). When these authors did not specify subregional distributions within CSA, presence/absence relative to the

**Table 1** Equivalence between the Neotropical zoogeographic regions of Parker *et al.* (1996) used in this study and the avian areas of endemism described for birds (Cracraft, 1985; Haffer, 1985). Names of the areas of endemism follow Cracraft (1985)

Zoogeographic regions	Areas of endemism
Patagonia (PAT)	Patagonian
Southern Andes (SAN)	Nothofagus
Pampas (PAM)	Pampas
Central South America (Chaco, CHA)*	Chaco
Central South America (Cerrado, CER)	Cerrado
Central South America (Caatinga, CAA)	Caatinga
Atlantic Forest (ATL)	Paraná + Serra do Mar
Subtropical Pacific (STP)	Peruvian Coastal
Equatorial Pacific (EPC)	Tumbesian
Central Andes (CAN)	Austral Andean + Peruvian Andean + Marañón
Northern Andes (NAN)	North Andean + Santa Marta + Perijan Montane + Venezuelan Montane + Parian Montane + Meridan Montane
Chocó (CHO)	Chocó
Tepuis (TEP)	Pantepui
Northern South America (NSA)	Nechí + Guajira + Magdalena + Llanos
Amazonia North (AMN)	Guyanan + Imerí + Napo
Amazonia South (AMS)	Belém + Pará + Rondónian + Tapajós + Inambari

\*Also includes the Argentine Monte and Espinal, which do not form part of the Chaco area of endemism as originally defined.



**Figure 2** Graphic comparison of the zoogeographic regions of Stotz *et al.* (1996) and the areas of endemism described for birds by Cracraft (1985) and Haffer (1985), shown in grey. Names of the areas of endemism can be found in Fig. 1. Figure modified from Stotz *et al.* (1996, p. 479, map 1).

Caatinga, Cerrado, and Chaco was determined from data in secondary references: Blake (1977) when applicable, Pacheco & Bauer (2000) and Olmos (1993) for the Caatinga, Silva (1995b) for the Cerrado, and Short (1975) and Hayes (1995) for the Chaco. When there was ambiguity regarding a taxon’s distribution in the Parker *et al.* (1996) data base and any of the secondary references, it was coded as unknown. Taxonomy, including subspecies, followed Parker *et al.* (1996) in all instances. This increased coding efficiency and provided clear guidelines for replicating the analysis. The resulting matrix can be downloaded from Appendix S1 (see Supplementary Material), or obtained from the authors.

**Analyses**

Three analyses were performed, each with a different goal:

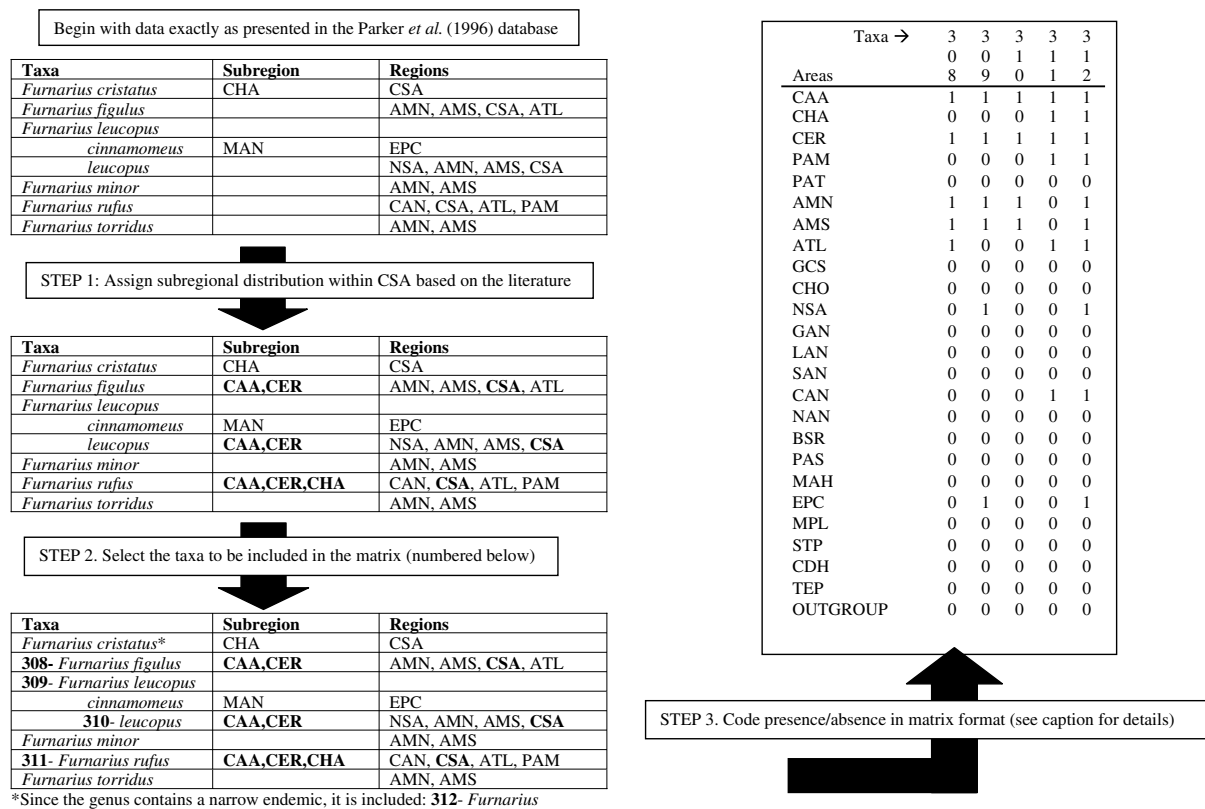
*Analysis 1*

The goal of this analysis was to examine the influence of hierarchical information on area relationships and tree

support. For this purpose, we analysed three data sets, which included gradually greater amounts of hierarchical information: species only (521 taxon-characters), species and genera (560 taxon-characters), and the total evidence data set, composed of 592 taxon-characters at the taxonomic level of species (521), genera (39), and subspecies (32).

*Analysis 2*

The goal of the second analysis was to investigate the influence of the forest-dependent Cerrado taxa on arid-area relationships. The Cerrado is known to be a mosaic of arid and humid vegetation, the latter mostly in the form of gallery and riverine forests, which contain many birds of primarily Amazonian or Atlantic Forest affinity (Silva & Bates, 2002). Do these taxa bias the results by grouping the Cerrado with forest areas? To examine this, we analysed a dataset where Cerrado taxa considered to be ‘forest-dependent’ by Silva (1995b) were re-coded as absent from the Cerrado. If this resulted in the taxon having no representatives in the CSA region, then the whole taxon was deleted



**Figure 3** Step-by-step example of the coding rules used in this study using the genus *Furnarius*. Begin with data exactly as presented in the Parker *et al.* (1996) data base. 'MAN' corresponds to the Manabí-Tumbes subregion within EPC of Parker *et al.* (1996) and is not relevant to our study. STEP 1. For those taxa that have only CSA specified for distribution, assign subregional distribution to either Caatinga, Cerrado, and/or Chaco based on the literature (see Methods). STEP 2. Select the taxa to be included in the matrix (shown numbered). STEP 3. Presence/absence is coded in matrix format. *Furnarius cristatus* is not included due to its endemism. To capture the fact that it shares a common ancestor with the rest of its congeners, *Furnarius*, the genus, is also included. Both *F. leucopus leucopus* and *F. leucopus* are included. In this case, the species *F. leucopus* is functioning as a hierarchical taxon-character. Numbers in the matrix correspond to the taxon-characters as follows: 308, *F. figulus*; 309: *F. leucopus*; 310: *F. l. leucopus*; 311: *F. rufus*; 312: *Furnarius*.

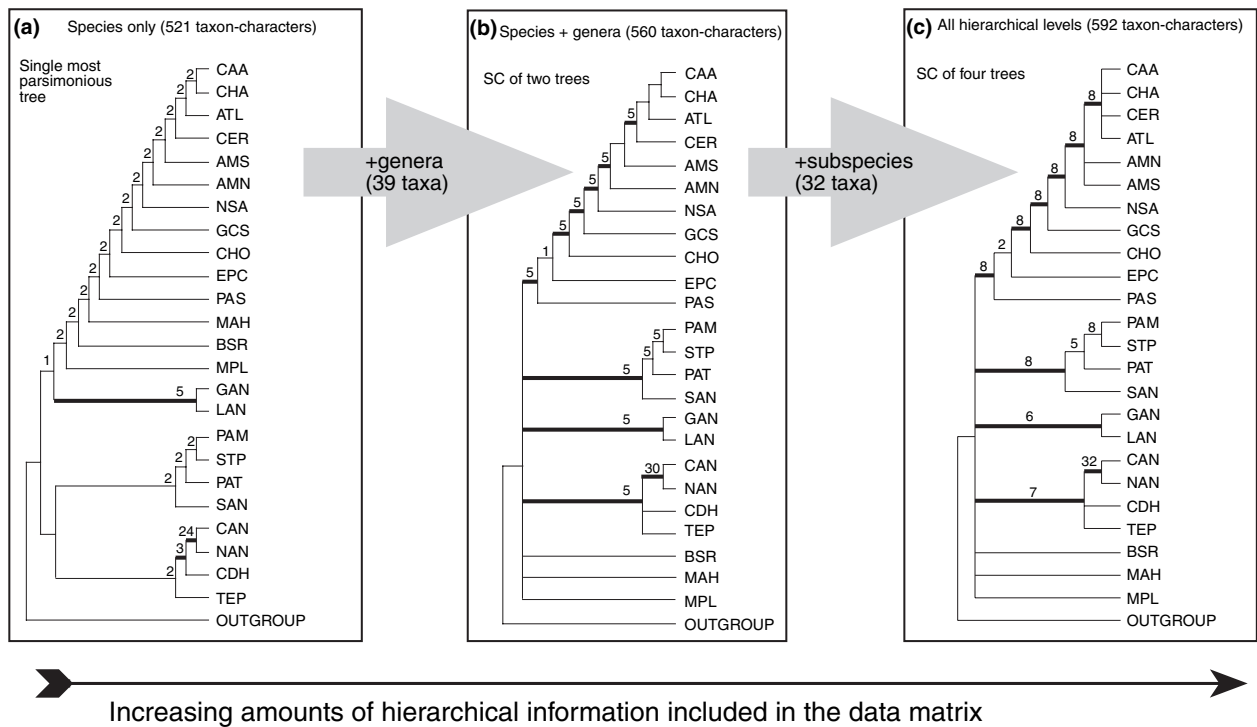
from the matrix. This procedure ensured consistency with the criterion used to sample the taxa initially, namely that all taxa included in the matrix must have at least one representative in one of the CSA subregions. The data set used for this analysis had a total of 541 taxon-characters, and is given in Appendix S2.

### Analysis 3

The third and final analysis investigated whether different taxonomic groups provide different signals. Bates *et al.* (1998) investigated this question in their study of passerine distributions and found that different taxonomic subsets yielded different, sometimes conflicting cladograms. To examine the signal carried by different taxonomic groups, parsimony analyses were performed for 12 data sets each composed of a different taxonomic subset. Subsets analysed and compared were: all taxa excluding 'water birds' (Pelecaniformes, Podicipediformes, Ciconiiformes, Anseriformes, Gruiformes, and Charadriiformes), Columbiformes, Psittaciformes, Piciformes, Dendrocolaptidae, Furnariidae, Tyrannidae, Emberizidae,

Suboscines, Oscines, Passeriformes, and non-Passeriformes. The character exclusion sets necessary to perform all of the different analyses described above are already specified in the available matrix files.

The total evidence matrix and all its partitions were analysed by maximum parsimony as implemented by PAUP\* (version 4.0b 10, Swofford, 1998), using heuristic searches with 100 random addition sequences and a TBR branch-swapping algorithm. Nodal support was assessed by performing bootstrap analyses with 100 pseudoreplicates (with 10 random addition sequences and TBR branch-swapping) and by calculating decay indices using TreeRot. version 2b (Sanderson, 1999). Decay indices are the tree-length differences, in number of steps, between the most parsimonious (shortest) tree and the (longer) trees in which a given node is no longer supported (Bremer, 1988). The higher the decay index for a given node the more steps that have to be added for that node to collapse, and thus high indices reflect higher support for nodes. State optimizations were visualized using MacClade version 3.07 (Maddison & Maddison, 1992).



**Figure 4** Results of the first analysis performed in this study, designed to investigate the influence of hierarchical information on tree topology and support. Three data sets were used, including: (a) species only, (b) species and genera, and (c) species, genera, and subspecies. A comparison of trees obtained with each dataset shows that when hierarchical information is lacking, even in a large data set, nodal support is very low throughout the tree (tree A). A gradual increase in the amount of hierarchical information that is included in the data matrix results in progressively higher values of support (trees B and C). Trees presented are strict consensus trees if multiple parsimonious solutions were obtained, and the number of most parsimonious trees obtained and included taxon-characters are indicated next to each tree. Support is indicated by decay indices (above branches) and by bootstrap values, which are indicated as bold branches when above 70%.

**RESULTS**

*Analysis 1: The effect of hierarchical information*

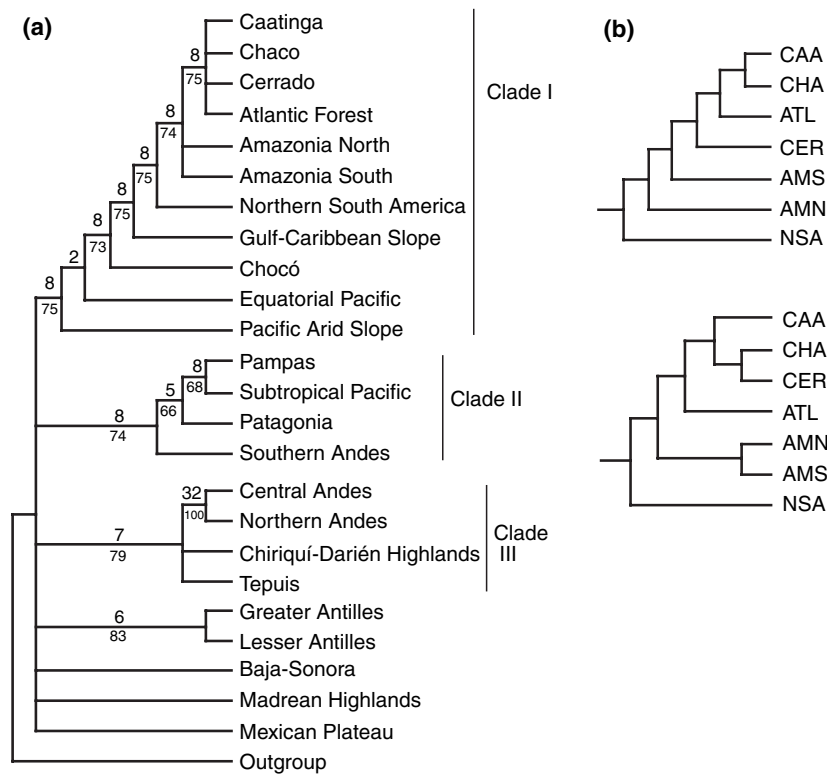
Results show that when hierarchical information is lacking, even in a large data set, nodal support is very low throughout the tree (Fig. 4, tree A).

As the amount of hierarchical information included in the data matrix increases, however, progressively higher values of support are obtained (Fig. 4, trees B and C), in this case with minor changes in relationships among the areas. Thus, it is only by adding hierarchical information that we are able to obtain robust results for the areas of interest.

Parsimony analysis of the total-evidence matrix (592 taxon-characters) results in four most-parsimonious trees of 1721 steps, CI = 0.34, and HI = 0.66. The strict consensus tree corresponds to tree C in Fig. 4 and is presented in more detail in Fig. 5. Despite the lack of resolution at the base of the tree, relationships of most areas are resolved and nodal support is moderate to high for most branches. The first major clade contains the CSA areas and all of the tropical forest areas. There is high support for a cluster comprised of the Caatinga, Cerrado, Chaco, and the Atlantic Forest areas,

but the data provide conflicting patterns of relationships among these areas. Two alternatives are found among the four most parsimonious trees (Fig. 5b), and in one of them the Cerrado area does not cluster with the other CSA areas but is instead basal to CSA + Atlantic Forest. Also unresolved is whether the two groups of Amazonian areas, AMN and AMS, are related to each other or are instead sequentially related to the CSA + Atlantic Forest group. The next nodes in this first large clade establish relationships between the CSA + Amazonian areas and several other arid and forest areas. These are, sequentially more distant: Northern South America (NSA) (Nechí + Magdalena + Guajira + Llanos areas of endemism), Gulf Caribbean Slope, Chocó, Equatorial Pacific (Tumbesian area of endemism), and as sister to all areas in this clade, the Pacific Arid Slope of Central America. The Pampas area is not part of the first clade but forms part of a second cluster that includes Patagonia, the Southern Andes, and the Subtropical Pacific. A third clade contains the Northern and Central Andes related to the Tepuis and the Chiriquí-Darién Highlands. Finally, the Greater and Lesser Antilles cluster together and their affinities are unresolved, as are those of Baja Sonora, the Madrean Highlands, and the Mexican Plateau.





**Figure 5** (a) Strict consensus tree of the four most parsimonious trees obtained with the complete data set (total evidence tree, equal to tree C in Fig. 4). Decay indices are above branches and bootstrap support values are indicated below branches. (b) The two alternative sets of relationships found for the CSA areas among those four most parsimonious trees.

The number of shared absences was found to be extremely low in this analysis, never above 9% of the total number of apomorphies along a branch, and typically much lower than that. The only exception was the branch joining the two Antillean areas, where shared absences accounted for 25% of the apomorphies.

#### *Analysis 2: The influence of forest-dependent Cerrado taxa*

A total of 127 taxa present in the Cerrado are considered forest-dependent by Silva (1995b). These were re-coded as being absent from that area, which resulted in 51 being absent from CSA altogether. When this new data set of 541 taxon-characters is analysed, a new tree is obtained in which the CSA areas form a cluster to the exclusion of any forest area and a strongly supported node joins the Chaco and Cerrado areas (Fig. 6). The number of shared absences increases only slightly to 15% of the total apomorphies along the branch joining the Chaco and Cerrado.

#### *Analysis 3: Do different taxonomic subsets provide different signals?*

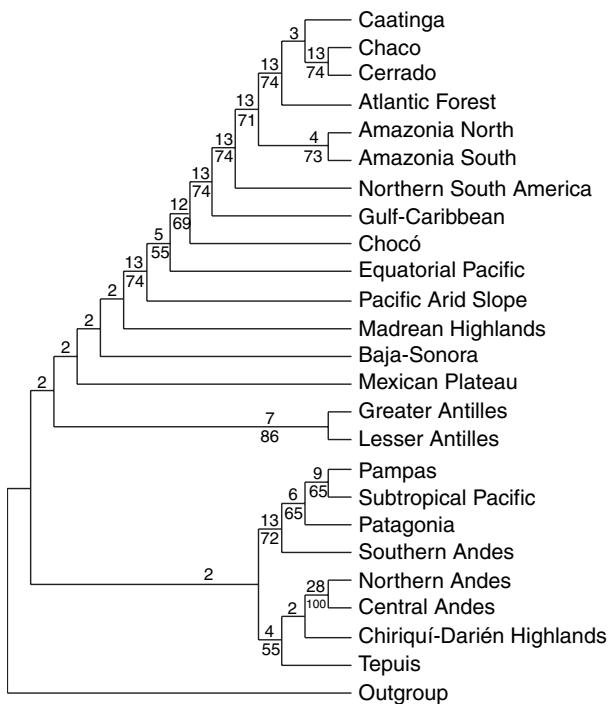
Trees obtained with different taxonomic subsets can be found in Fig. 7. When the 71 water birds are excluded the resulting tree is compatible with the total evidence consensus tree (Fig. 5a), and in addition, contains an exclusive cluster of CSA

areas (Fig. 7, tree 1). Most of the remaining taxonomic subsets, such as Columbiformes, Psittaciformes, Dendrocolaptidae, Tyrannidae, and Emberizidae, have very few shared taxon-characters across the large number of areas being analysed. Thus, they result in trees with very low support and resolution. Piciformes and Furnariidae determine a few robust nodes, but these are contradictory about CSA area relationships. Finally, the non-Passeriformes and the suboscines recover trees compatible with the total evidence consensus tree, but their complementary data sets, the Passeriformes and the oscines, recover trees that are in conflict with the total evidence solution in terms of the position of the Pampas area (Fig. 7, trees 10 and 11). In these trees, the Pampas and Chaco areas form a cluster with high support.

## DISCUSSION

### Central South America area-relationships

The first analysis discussed earlier demonstrates that a gradual increase in the amount of hierarchical information can result in progressively higher values of support, even when the relative number of hierarchical taxon-characters that is being added is low (analysis 1; Figs 4 and 5). Furthermore, the addition of hierarchical information causes branch support to increase substantially, and allows us to conclude that the CSA areas form a clade together with the Atlantic Forest area.



**Figure 6** Tree obtained in the second analysis performed in this study, when the forest-dependent Cerrado taxa are coded as absent from the Cerrado area, and total of 541 taxon-characters are included in the analyses (see methods for details). The three CSA areas now form a clade, with relationships among them fully resolved. Decay indices are above branches and bootstrap support values are indicated below.

Relationships among these areas, however, are not completely resolved, even in the total evidence solution (Fig. 5). This might be attributed to numerous factors that can introduce noise into the analyses, which can relate both to the quality of the data, such as incorrect distributions or taxonomy, and to biological processes, such as extinction and dispersal. Incorrect distributional data are unlikely to be a significant source of noise for this data set; birds are probably the most comprehensively surveyed taxonomic group for the Neotropics, especially at the broad scale of areas of endemism (Ridgely & Tudor, 1989; Parker *et al.*, 1996; Brooks *et al.*, 1999). The main limitation of the taxonomy used here is that most species and subspecies are based on the biological species concept (Mayr, 1942) and are in need of revision. Thus, this data set is surely an underestimate of the information present in raw distributions. Taxonomic inaccuracy could be a source of noise, but is also unlikely to generate congruent misleading signal toward a particular hypothesis of area-relationships. It is more likely that the total evidence solution lacks support due to noise introduced by extinctions and/or dispersal events.

Relationships among areas are resolved, in some instances with high support, when forest-dependent taxa are removed from the Cerrado (analysis 2; Fig. 6). Because these taxa are thought to have dispersed into the Cerrado relatively recently (Silva, 1995b), this result suggests that long-distance dispersal

and/or dispersion are introducing noise in the total evidence solution. The tree obtained in analysis 2 resolves relationships among CSA areas so that the Cerrado and Chaco are clustered together with high support, and related in turn to the Caatinga. One interpretation of these relationships is based on shared numbers alone: the Cerrado and Chaco share more species with each other than either does with the Caatinga. An alternative interpretation, given that hierarchical data were also included, invokes shared history: elements of the Caatinga avifauna were isolated from their Cerrado and/or Chaco relatives and can be expected to be basal in a phylogenetic context, and there was a subsequent separation between the Cerrado and Chaco.

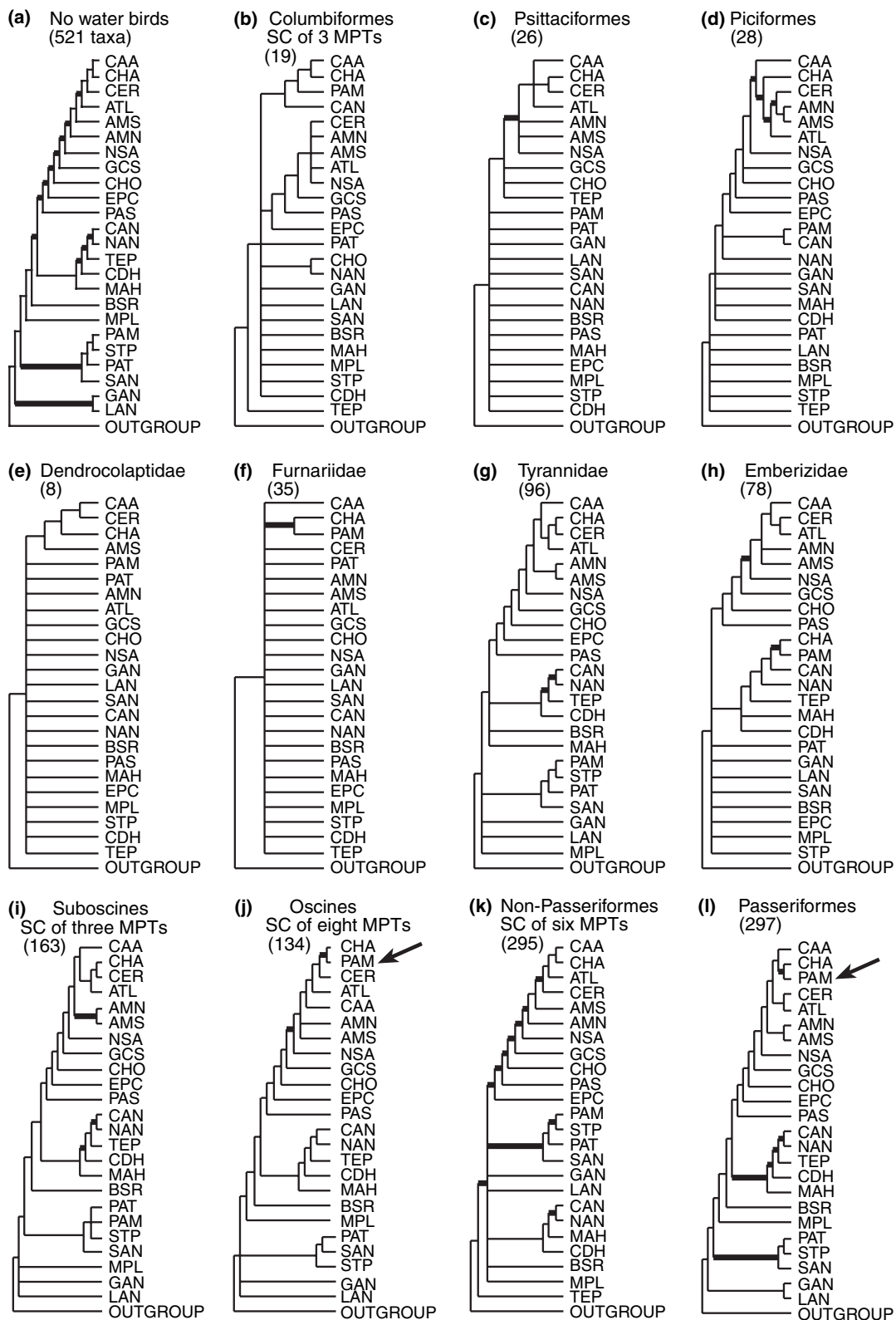
This hypothesis of CSA relationships is congruent with a previously proposed vicariance scenario for the diversification of the Cerrado avifauna. Based on an analysis of the available taxonomic information for 16 Cerrado endemic bird species, Silva (1995a) proposed that several groups display congruent relationships reflecting a strong area-relationship between the Cerrado and Chaco avifauna, but not between the Cerrado and Caatinga. He postulated that this pattern could have been caused by a common vicariant event between the Cerrado and other southern areas, mainly the Chaco, Patagonia, Pampas, and adjacent sub-Andean zone. More specifically, he postulated that the vicariant event involved was the Late Pliocene–Early Pleistocene (2–3 Ma) generalized uplift of the Brazilian Plateau (along with the Espinhaço Range, Serras do Mar and Mantiqueira) and the correlated subsidence of the Chaco, Pantanal, and Llano regions due to an intense phase of uplift of the Andes. Silva (1995b) further suggested that events of different ages may be relevant for different elements of the endemic avifauna.

Multiple histories remain possible for the CSA avifauna, but in this study, in which the endemic arid avifauna is analysed as a whole, congruent patterns seem to be emerging.

### Relationships of other areas

In our results the Atlantic Forest is never positioned as a sister area to the Amazonian areas, instead showing a close relationship to the CSA areas (Figs 5 and 6). Ambiguity with respect to the exact relationships of the Atlantic Forest could be due to its 'composite nature', a feature that was proposed for this area by numerous authors based on the complex phylogenetic relationships of Atlantic Forest species with respect to Amazonian species (Cracraft & Prum, 1988; Prum, 1988; Hackett & Lehn, 1997). Bates *et al.* (1998) found that the two Atlantic Forest areas were grouped together, but their relationship to Amazonia was the weakest node on their tree. Our data are far from ideal for testing Amazonian or Atlantic Forest area-relationships, but the lack of support for a monophyletic assemblage of these forest areas highlights the need for more comprehensive studies that are continental in scope, and include both forest and arid areas.

The relationships of the remaining areas in the large CSA clade found in the total evidence tree (Fig. 5a, clade I) are fully



**Figure 7** Results obtained under 12 data partitions corresponding to different taxonomic subsets. Trees presented are single most parsimonious trees or strict consensus trees when indicated, and number of most parsimonious trees and included taxa are indicated next to each tree. Nodal support is shown by bootstrap values, which are indicated as bold branches when above 70%. Arrows in trees 10 and 12 highlight the novel position that is obtained for the Pampas with the Passeriformes and the oscine passerines.

compatible with the area-relationships obtained in our analysis 2 (Fig. 6). They are also congruent with previous biogeographic hypotheses. Northern South America and the Chocó have been proposed to be related to CSA and Amazonia, respectively. Haffer (1967) studied the distribution and traditional taxonomy of 101 species characteristic of NSA and found an almost complete lack of endemic genera and a low number of endemic species, the latter of which he considered to be strongly differentiated representatives of CSA taxa. He concluded that the closest arid area to the CSA areas was NSA (equivalent to the Nechí, Magdalena, Guajira, and Llanos areas of endemism). In his review of the Chaco avifauna Short (1975, p. 340) remarked that for 'Chaco-dwelling species, disjunctions involving NSA are the most numerous', and based on the small number of mostly intraspecific disjunctions, he suggested that the Amazonian forests could have been a barrier between populations of these areas. In Short's view, these disjunct distributions reflect a past connection by means of former favourable habitat either around Amazonia coastally to the East, through Amazonia, or along the Andes to the West. The first route has been proposed to be the source of connections between the NSA arid areas and the Caatinga for scorpions (Lourenço, 1994) and between NSA and the Cerrado for birds (Silva, 1995b). Although both Short and Haffer considered these connections to have occurred during the Pleistocene, they could be older; the Amazon river and its basin have been a major barrier since late in the Miocene, when the Purus Arch was breached and the continuous Amazon was formed (Dobson *et al.*, 1997; Lundberg *et al.*, 1998). The basal position of NSA in our cladogram, in relation to the divergences among the CSA areas suggests a divergence time older than the Late Pliocene.

In agreement with previous studies (Cracraft & Prum, 1988; Bates *et al.*, 1998) our trees also support the close relationship of the Chocó and Central American areas relative to all Amazonian areas (Figs 5 and 6). This separation is accepted to be old and some postulate it could have been initiated as early as 25 Ma (Bush, 1994; Brumfield & Capparella, 1996; Bates *et al.*, 1998). In our trees the Chocó and Central American areas are sequentially related to the Amazonia + CSA clade. Since previous studies did not include arid areas, our result is compatible with previous findings.

The remaining South American areas in this clade (Figs 5a and 6), basal to all of the above, are the Equatorial Pacific and the Pacific Arid Slope. The Pacific Arid Slope was postulated by previous authors (Haffer, 1967; Campbell, 1982) to have been the source for part of the Equatorial Pacific endemic avifauna in a 'pre-Andean' stage before the development of the dense Chocó forests, but the former two areas are not grouped as sisters here. While Haffer (1967) postulated the timing of these events to be Pleistocene, the relative position of these nodes on our tree suggests again an older time-frame.

It is interesting to note that joint analysis of arid and forest areas results in an arrangement that has the arid and forest areas interspersed, but is still compatible in terms of relationships and relative temporal order with previous hypotheses

that included either only arid or only forest areas. One interpretation of this result is that these areas are sequentially related, and the nodes reflect vicariance events that shaped the avifaunas of these areas. This would imply that the evolutionary relationships of forest and aridland biotas could have been more dynamic and interconnected than previously thought (García-Moreno & Silva, 1997). Alternatively, and instead of vicariant history, the signal uniting the areas in this particular fashion could have been created by recent dispersion of species among the arid and neighbouring forest areas. Silva (1995b), for example, has reported that 36% of the breeding avifauna of the Cerrado may have expanded, or may be expanding, their ranges from Amazonian or Atlantic Forest areas by following mesic gallery forests. Thus, the close relationships between the CSA and the Amazonian and Atlantic Forest areas could be emerging from dispersion between these areas after the CSA–NSA divergence. This receives support, at least for the case of the Cerrado and its neighbouring forest areas, from the results of the second analysis performed here (Fig. 6).

The areas in the second clade of the total evidence tree (clade II in Fig. 5), with the exception of the Pampas, have been proposed by Morrone (2000) to be part of the Andean Region or Andean biogeographic realm. This realm is postulated to be related to other austral temperate areas in Southern Africa and Australasia, as well as the Northern and Central Andes (Takhtajan, 1986; Crisci *et al.*, 1991). In addition, their relationship to the Pampas is unconventional. The Pampas area is generally considered to be associated with the CSA areas, as in Morrone's (2000) Chacoan biogeographic region. Some signal for a Pampas–CSA relationship exists in the total evidence data set, because a number of taxa display a 'homoplastic' distribution of their species in both CSA and the Pampas to the exclusion of other areas (not shown), and in fact, a CSA + Pampas relationship emerges in some taxonomic subsets (Passeriformes, oscines, Emberizidae, Columbiformes; see Fig. 7). This pattern could indicate mixed relationships for the Pampas area, due to multiple vicariance histories and/or dispersal events between this area and surrounding areas. The Pampas habitat is indeed a mosaic of Monte, Chaco, and forest elements (Paraenense and Atlantic) that could be very hospitable to dispersing populations. It has been suggested that the avifauna from different habitats within the Pampas could have different biogeographic relationships (Arballo & Cravino, 1999). While our data do not allow further investigation of this question, it can be further addressed with phylogenetic hypotheses for the taxa involved.

The final clade in the total evidence tree (clade III in Fig. 5) joins the Central and Northern Andes with the Tepuis and the Chiriquí-Darién highlands of Central America. This is similar to the results of Mayr & Phelps (1967), which led them to conclude that 43 of 48 Pantepui taxa were probably of northern Andean origin. In addition, the incorporation of the Central American Chiriquí-Darién highlands into this clade is in agreement with the connections among South America and the Central American highlands that have been proposed for some bird taxa based on molecular data (Hackett, 1993).

Relationships of the last three areas in our matrix, Baja Sonora, the Madrean Highlands, and the Mexican Plateau, are unresolved in most analyses. This is not surprising, for if a taxon is found in any of these areas and in a CSA area, it is likely to be very widespread and minimally informative as well. The total evidence dataset nevertheless provided resolution throughout most of the tree.

## CONCLUSIONS

Our analyses suggest that at the scale of continental areas of endemism, and with a dataset encompassing hierarchical information from all Neotropical bird families, extinction, dispersal, and dispersion do not erase the historical signal present in raw distributional information. Our results reveal that in spite of the limitations presented by dated taxonomic classifications and a coarse-grained representation of areas of endemism, a remarkable degree of resolution can be obtained for most areas in the matrix. The most parsimonious solutions cannot completely resolve the relationships of the Cerrado, Chaco, and Caatinga, but after coding forest-dependent Cerrado taxa as absent from the Cerrado we find support for CSA as an exclusive clade, with the Chaco and Cerrado as sister areas, as well as considerable congruence with previous studies and proposed biogeographic scenarios. This study also suggests that the most important sources of noise for these types of analyses might come from species that have recently expanded their ranges into neighbouring areas, and illustrates how such ecological noise and historical information are more easily distinguished in CADE analyses than they would be in PAE analyses.

We agree with Bates *et al.* (1998) that this approach is not a substitute for comparative phylogenetic studies, but as illustrated here, CADE can yield much useful information when sample sizes are large: it can provide a testable hypothesis of area-relationships, it allows exploration of the different relationships within the avifauna of a given area, and it can inform taxonomic sampling for future phylogenetic and historical biogeographic studies.

## ACKNOWLEDGEMENTS

We are grateful to Alice Cibois, Keith Barker, Liliana Dávalos, and Leo Joseph for comments on earlier versions of this manuscript, to John Bates for suggestions on how to analyse the Cerrado avifauna, and to Dan Brooks and Marco van Veller for allowing us to see a draft of their manuscript. ALP received generous support from the Frank M. Chapman Memorial Fund at the AMNH, the Center for Environmental Research and Conservation and the Department of Ecology, Evolution and Environmental Biology at Columbia University, and the AMNH's Center for Biodiversity and Conservation (NASA Grant no. NAG5-8543). This paper is a contribution from the Cullman Research Facility in the Department of Ornithology, American Museum of Natural History, and has

received generous support from the Lewis B. and Dorothy Cullman Program for Molecular Systematics Studies, a joint initiative of The New York Botanical Garden and The American Museum of Natural History.

## SUPPLEMENTARY MATERIAL

The following material is available from <http://www.blackwellpublishing.com/products/journals/suppmat/JBI/JBI1138/JBI1138sm.htm>

**Appendix S1** Complete matrix.

**Appendix S2** Modified matrix.

## REFERENCES

- Ab'Sáber, A.N. (1977) Os domínios morfoclimáticos da América do Sul. Primeira Aproximação. *Geomorfologia*, **52**, 1–21.
- Arballo, E. & Cravino, J.L. (1999) *Handbook of the birds of Uruguay*, Vol. 1. Hemisferio Sur, Montevideo.
- Bates, J., Hackett, S. & Cracraft, J. (1998) Area-relationships in the Neotropical lowlands: an hypothesis based on raw distributions of Passerine birds. *Journal of Biogeography*, **25**, 783–793.
- Blake, E.R. (1977) *Manual of neotropical birds*. University of Chicago Press, Chicago, IL.
- Bremer, K. (1988) The limits of amino-acid sequence data in angiosperm phylogenetic reconstruction. *Evolution*, **42**, 795–803.
- Brooks, D.R. & Van Veller, M.G.P. (2003) Critique of parsimony analysis of endemism as a method of historical biogeography. *Journal of Biogeography*, **30**, 819–825.
- Brooks, T., Mehlman, D., Roca, R., Jahn, A., Limp, F., Smith, K. & Williams, P. (1999) *Worldmap for birds at risk in Latin America*. The Nature Conservancy, Washington, DC.
- Brooks, D.R., Van Veller, M.G.P. & McLennan, D.A. (2001) How to do BPA, really. *Journal of Biogeography*, **28**, 345–358.
- Brumfield, R.T. & Capparella, A.P. (1996) Historical diversification of birds in Northwestern South America: A molecular perspective on the role of vicariant events. *Evolution*, **50**, 1607–1624.
- Bush, M.B. (1994) Amazonian speciation: a necessarily complex model. *Journal of Biogeography*, **21**, 5–17.
- Cabrera, A.L. & Willink, A. (1973) *Biogeografía de América Latina*. OEA, Washington, DC.
- Campbell, K.E., Jr. (1982) Late Pleistocene events along the coastal plane of Northwestern South America. *Biological diversification in the tropics* (ed. by G. T. Prance). Columbia University Press, New York.
- Cracraft, J. (1985) Historical biogeography and patterns of differentiation within the southamerican avifauna: areas of endemism. *American Ornithological Union Ornithology Monographs*, 49–84.

- Cracraft, J. (1986) Origin and evolution of continental biotas: speciation and historical congruence within the Australian avifauna. *Evolution*, **40**, 977–996.
- Cracraft, J. (1988) Deep-history biogeography: retrieving the historical pattern of evolving continental biotas. *Systematic Biology*, **37**, 221–236.
- Cracraft, J. (1991) Patterns of diversification within continental biotas: hierarchical congruence among the areas of endemism of Australian vertebrates. *Australian Systematic Botany*, **4**, 211–227.
- Cracraft, J. & Prum, R. (1988) Patterns and processes of diversification: speciation and historical congruence in some Neotropical birds. *Evolution*, **42**, 603–620.
- Crisci, J.V. (2001) The voice of historical biogeography. *Journal of Biogeography*, **28**, 157–168.
- Crisci, J.V., Cigliano, M.M., Morrone, J.J. & Roig-Juñent, S. (1991) Historical biogeography of southern South America. *Systematic Zoology*, **40**, 152–171.
- Dinerstein, E., Olson, D.M., Graham, D.J., Webster, A.L., Primm, S.A., Bookbinder, M.P. & Ledec, G. (1995) *A conservation assessment of the terrestrial ecoregions of Latin America and the Caribbean*. The World Bank, The World Wildlife Fund, Washington, DC.
- Dobson, D.M., Dickens, G.R. & Rea, D.K. (1997) Terrigenous sedimentation at Ceará Rise. *Proceedings of the Ocean Drilling Program Scientific Results* (ed. by N. J. Shackleton, W. B. Curry, C. Richter and T. J. Bralower), **154**, 465–473.
- García-Barros, E., Gurrea, P., Lucíañez, M.J., Cano, J.M., Munguira, M.L., Moreno, J.C., Sainz, H., Sanz, M.J. & Simón, J.C. (2002) Parsimony analysis of endemism and its application to animal and plant geographical distributions in the Ibero-Balearic region (western Mediterranean). *Journal of Biogeography*, **29**, 109–124.
- García-Moreno, J. & Silva, J.M.C. (1997) An interplay between forest and non-forest South American avifaunas suggested by a phylogeny of *Lepidocolaptes* Woodcreepers (Dendrocolaptinae). *Studies on Neotropical Fauna & Environment*, **32**, 164–173.
- Hackett, S. (1993) *Molecular phylogenies and the biogeography of Central American Birds*. Doctoral Thesis, Louisiana State University, Baton Rouge, LA.
- Hackett, S.J. & Lehn, C.A. (1997) Lack of genetic divergence in a genus (*Pteroglossus*) of Neotropical birds: the connection between life history characteristics and levels of genetic divergence. *Neotropical ornithology honoring Ted Parker* (ed. by J. J. V. Remsen), pp. 267–280. American Ornithologists' Union, Washington, DC.
- Haffer, J. (1967) Zoogeographical notes on the 'nonforest' lowland bird faunas of northwestern South America. *Hornero*, **10**, 315–333.
- Haffer, J. (1985) Avian zoogeography in the Neotropical lowlands. *Ornithological Monographs*, **36**, 113–146.
- Haffer, J. (1997) Alternative models of vertebrate speciation in Amazonia: an overview. *Biodiversity and Conservation*, **6**, 451–476.
- Hayes, F.E. (1995) *Status, distribution and the biogeography of the birds of Paraguay*. American Birding Association, Monographs in Field Ornithology, no.1 Colorado Springs, Colorado.
- Hengeveld, R. (1990) *Dynamic biogeography*. Cambridge University Press, Cambridge.
- Hooghiemstra, H. & van der Hammen, T. (1998) Neogene and Quaternary development of the Neotropical forest: the forest refugia hypothesis, and a literature overview. *Earth Science Reviews*, **44**, 147–183.
- Hueck, K. & Seibert, P. (1972) *Vegetationskarte von Sudamerika. Mapa de la vegetación de America del Sur. Erläuterungen zur Karte 1:8 Millionen*. G. Fischer, Stuttgart.
- Humphries, C. (1989) Any advance on assumption 2? *Journal of Biogeography*, **16**, 101–102.
- Humphries, C.J. & Parenti, L.R. (1999) *Cladistic biogeography*, 2nd edn. Oxford University Press, Oxford.
- Linder, H.P. (2001) On areas of endemism, with an example from the African Restionaceae. *Systematic Biology*, **50**, 892–912.
- Lourenço, W.R. (1994) Biogeographic patterns of tropical South American Scorpions. *Studies on Neotropical Fauna and Environment*, **29**, 219–231.
- Lundberg, J.G., Marshall, L.G., Guerrero, J., Horton, B., Malabarba, M.C.S.L. & Wesselingh, F. (1998) The stage for Neotropical fish diversification: a history of tropical South American rivers. *Phylogeny and classification of neotropical fishes* (ed. by L. Malabarba, R. E. Reis, R. P. Vari, Z. M. S. Lucena and C. A. S. Lucena), pp. 13–48. EDIPUCRS, Porto Alegre, Brazil.
- Maddison, W.P. & Maddison, D.R. (1992) *MacClade: analysis of phylogeny and character evolution*. Sinauer Associates, Sunderland, MA.
- Mares, M.A. (1992) Neotropical mammals and the myth of Amazonian biodiversity. *Science*, **255**, 976–979.
- Mayr, E. (1942) *Systematics and the origin of species*. Columbia University Press, New York.
- Mayr, E. & Phelps, W.J. (1967) The origin of the bird fauna of the South Venezuelan highlands. *Bulletin of the American Museum of Natural History*, **136**, 269–328.
- Morrone, J.J. (1994) On the identification of areas of endemism. *Systematic Biology*, **43**, 438–441.
- Morrone, J.J. (1998). On Udvardy's Insularctica province: a test from the weevils (Coleoptera: Curculionioidea). *Journal of Biogeography*, **25**, 947–955.
- Morrone, J.J. (1999) Presentación preliminar de un nuevo esquema biogeográfico de America del Sur. *Biogeografica*, **75**, 1–16.
- Morrone, J.J. (2000) What is the Chacoan subregion? *Neotropica*, **46**, 51–68.
- Nelson, G.J. & Platnick, N.I. (1981) *Systematics and biogeography: cladistics and vicariance*. Columbia University Press, New York.
- Nores, M. (1992) Bird speciation in subtropical South America in relation to forest expansion and retraction. *Auk*, **109**, 346–357.

- Olmos, F. (1993) Birds of Serra da Capivara National Park, in the 'caatinga' of north-eastern Brazil. *Bird Conservation International*, **3**, 21–36.
- Pacheco, J.F. & Bauer, C. (2000) As Aves da Caatinga – Apreciação histórica do processo de conhecimento. Available online at: <http://www.biodiversitas.org.br/caatinga/documentos.html>.
- Parker, T.A., III, Stotz, D.F. & Fitzpatrick, J.W. (1996) Ecological and Distributional Databases. *Neotropical birds: ecology and conservation* (ed. by D. F. Stotz, J. W. Fitzpatrick, I. T. A. Parker and D. K. Moskovits). The University of Chicago Press, Chicago, IL.
- Patton, J.L. & Silva, M.N.F.d. (1998) Rivers, Refuges, and Ridges. *Endless Forms* (ed. by D. J. Howard and S. H. Berlocher), pp. 202–213. Oxford University Press, Oxford.
- Porzecanski, A.L. (2003) *Historical biogeography of the South American aridlands: a molecular study of endemic Avian Taxa*. Doctoral Thesis, Columbia University, New York.
- Posadas, P. (1996) Distributional patterns of vascular plants in Tierra del Fuego: a study applying Parsimony Analysis of Endemicity (PAE). *Biogeographica*, **72**, 161–177.
- Prum, R.O. (1988) Historical relationships among avian forest areas of endemism in the Neotropics. *Proceedings of the International Ornithological Congress*, **19**, 2562–2572.
- Raxworthy, C.J. & Nussbaum, R.A. (1996) Patterns of endemism for terrestrial vertebrates in Eastern Madagascar. *Biogéographie de Madagascar* (ed. by W. R. Lourenco), pp. 369–384. Orstom, Paris.
- Ridgely, R.S. & Tudor, G. (1989) *The birds of South America*, Vol. I. University of Texas Press, Austin, TX.
- Ridgely, R.S. & Tudor, G. (1994) *The birds of South America*, Vol. II. University of Texas Press, Austin, TX.
- Ron, S.R. (2000) Biogeographic area relationships of lowland Neotropical rainforest based on raw distributions of vertebrate groups. *Biological Journal of the Linnean Society*, **71**, 379–402.
- Rosen, B.R. (1988) From fossils to Earth history: applied historical biogeography. *Analytical biogeography* (ed. by A. A. Myers and P. S. Giller). Chapman & Hall, London.
- Rosen, B.R. (1992) The biogeographical black box: concepts and methods in marine paleobiogeography. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **92**, 171–205.
- Rosen, B.R. & Smith, A.B. (1988) Tectonics from fossils? Analysis of reef-coral and sea-urchin distributions from late Cretaceous to Recent, using a new method. *Gondwana and Tethys* (ed. by M. G. Audley-Charles and A. Hallam), pp. 275–306. Special Publication of Geological Society, London.
- Sanderson, M.D. (1999) *TreeRot, version 2*. Boston University, Boston, MA.
- Short, L.L. (1975) A zoogeographic analysis of the South American Chaco avifauna. *Bulletin of the American Museum of Natural History*, **154**, 163–352.
- Silva, J.M.C. (1995a) Biogeographic analysis of the South American Cerrado avifauna. *Steenstrupia*, **21**, 49–67.
- Silva, J.M.C. (1995b) Birds of the Cerrado region, South America. *Steenstrupia*, **21**, 69–92.
- Silva, J.M.C. (1997) Endemic birds species and conservation in the Cerrado region, South America. *Biodiversity and Conservation*, **6**, 435–450.
- Silva, J.M.C. & Bates, J.M. (2002) Biogeographic patterns and conservation in the South American Cerrado: a tropical savanna hotspot. *Bioscience*, **52**, 225–233.
- Silva, J.M.C. & Oren, D.C. (1996) Application of parsimony analysis of endemism in Amazonian biogeography: an example with primates. *Biological Journal of the Linnean Society*, **59**, 427–437.
- Sneath, P.H.A. & Sokal, R.R. (1973) *Numerical taxonomy*. W. H. Freeman and Co., San Francisco, CA.
- Stotz, D.F., Fitzpatrick, F.W., Parker III, T.A. & Moskovits, D.K. (1996) *Neotropical birds: ecology and conservation*. University of Chicago Press, Chicago.
- Swofford, D.L. (1998) *PAUP\*. Phylogenetic Analysis Using Parsimony (\*and other methods)*. Version 4. Sinauer Associates, Sunderland, MA.
- Takhtajan, A. (1986) *Floristic regions of the world*. University of California Press, Berkeley, CA.
- Trejo-Torres, J.C. & Ackerman, J.D. (2001) Biogeography of the Antilles based on a parsimony analysis of orchid distributions. *Journal of Biogeography*, **28**, 775–794.
- Waggoner, B. (1999) Biogeographic analyses of the Ediacara biota: a conflict with paleotectonic reconstructions. *Paleobiology*, **25**, 440–458.
- Wiley, E.O. (1988) Parsimony analysis and vicariance biogeography. *Systematic Zoology*, **37**, 271–290.

## BIOSKETCHES

**Ana Luz Porzecanski** received her PhD from Columbia University, working on avian molecular systematics and the historical biogeography of the South American aridlands. She is currently at the American Museum of Natural History, where she continues to investigate systematic and biogeographic questions, and also coordinates a Latin-American network of conservation educators and professionals.

**Joel Cracraft** is Lamont Curator of Birds at the American Museum of Natural History. His research has examined speciation and diversification, systematic theory and methodology, historical biogeography, and the evolutionary history of birds.

---

Editor: Philip Stott