EVOLUTION OF THE NEW WORLD JAYS (CORVIDAE): PHYLOGENY, BIOGEOGRAPHY, AND ECOLOGY

by

Elisa A. Bonaccorso

Licenciada en Biología, Universidad Simón Bolívar, Venezuela, 1999

Submitted to the Department of Ecology and Evolutionary Biology and the Faculty of the Graduate School of The University of Kansas in partial fulfillment of the requirements for the degree of Doctor in Philosophy

_____

Chairperson

Committee members

_____

_____

_____

_____

Date defended: _____

The Dissertation Committee for Elisa A. Bonaccorso certifies
that this is the approved version of the following dissertation:

EVOLUTION OF THE NEW WORLD JAYS (CORVIDAE): PHYLOGENY, BIOGEOGRAPHY, AND ECOLOGY

Committee:

_____
Chairperson

_____

_____

_____

_____

Date approved: _____

ABSTRACT

The evolutionary relationships among the New World jays are investigated in the context of their systematics, biogeography, and ecology. Molecular phylogenetic analyses of mitochondrial and nuclear genes reconstructed a basic phylogenetic structure in which all New World jay (NWJ) genera form a monophyletic group; *Cyanolyca* is reconstructed as sister to all other NWJ genera, and the remaining genera divide in two lineages: one of *Cyanocorax, Calocitta,* and *Psilorhinus,* and another of *Aphelocoma, Cyanocitta,* and *Gymnorhinus.* Phylogenetic analysis of *Cyanolyca* produced a well supported hypothesis of relationships, and showed that the genus *Cyanolyca* comprises two major clades: one contains the Mesoamerican "dwarf" jays, and the other consist of two main groups—one containing *C. cucullata* + *C. pulchra*, and the other containing the "core" South American species. High levels of genetic differentiation within *Cyanolyca* contrast with those observed in other Andean montane forest lineages studied to date. With only one exception, diverging lineages (sister species, as well as sister clades) are distributed on either side of potentially effective barriers to gene flow, suggesting the importance of allopatry in the diversification of the group. Phylogenetic analyses of *Cyanocorax, Psilorhinus,* and *Calocitta,* indicate monophyly of the group, but paraphyly of the current *Cyanocorax*. Ingroup taxa divided into two groups: Clade A, consisting of *Psilorhinus, Calocitta, Cyanocorax violaceus, C. caeruleus, C. cristatellus,* and *C. cyanomelas,* and Clade B, formed by the remaining *Cyanocorax* species; however, relationships within these groups are not completely resolved. Based on the phylogenetic results lumping of *Calocitta* and *Psilorhinus* into *Cyanocorax* is recommended. Finally, combination of

phylogenetic information and ecological niche modeling analyses revealed ecological patterns between sister taxa, as well as ecological trends in the group. In tree out of five pairs of sister species, a pattern of ecological niche conservatism was recovered; in one of those cases such conservatism is probably associated to an allopatric mode of speciation. The remaining two species show ecological and biographic patterns that may be attributed to complex evolutionary histories. Analysis of ecological distances showed that most species are more similar to species other than their sister species, however, when those distances are analyzed in a phylogenetic context, shared ecological trends emerge among some closely related species. Finally, when species ecological niches and niche dimensions where reconstructed into the phylogeny of the assemblage, evolutionary trends pointing to both, niche conservatism and niche diversification, were recovered.

A mi negris Juan Manuel

ACKNOWLEDGMENTS

Many people and institutions helped to make this dissertation possible. First, I want to thank Linda Trueb for opening the doors of the Herpetology Division back in January 2002. Without her encouragement and trust, I would not have made it to graduate school at KU it in the first place. Since then, she has been always supportive and caring. Town Peterson took me as a student when he had a thousand other students and commitments. For his great mentoring and respect as a colleague, I will be always grateful. I am also thankful to the members of my committee, Ed Wiley and Rafe Brown, for the privilege of being around their ideas and philosophy, and Stephen Benedict for being a genuine enthusiast of biology.

This dissertation benefited, one way or another, from discussions held with Jeet Sukumaran, Charles Linkem, Mark Robbins, and Juan Guayasamin. I am also thankful to Jeet for his ideas, help, and time put into running the bird cluster. Robert Fleisher (Smithsonian Institution) generously took the time and effort to extract DNA from my museum skin samples. Shannon DeVaney and Rob Moyle were always willing to give advice on molecular techniques, primers, and protocols. Arpi Nyari patiently taught me how to prepare bird skins, and Monica Papes was always there to answer questions about GIS and niche modeling. Roberto Sosa, Erick García-Trejo, and Carlos Rengifo, and specially, Juan Guayasamin, provided invaluable assistance in the field. Patricio and María Aigage offered their help and hospitality while conducting fieldwork in Oyacachi, Ecuador. Jorge Pérez-Emán, Paola Albuja, Luis Coloma, and Adolfo Navarro-Sigüenza provided institutional support for obtaining collecting permits in Venezuela, Ecuador, and Mexico.

The following individuals and institutions provided tissue samples under their care: Adolfo Navarro-Sigüenza (Museo de Zoología, Universidad Nacional Autónoma de México); Shannon Hackett and David Willard (Field Museum); Sharon Birks (Burke Museum of Natural History and Culture); Joel Craraft and Paul Sweet (American Museum of Natural History); Donna Dittman, Rob Brumfield, and James Remsen (Louisiana State University Museum of Natural Science); Nate Rice (Academy of Natural Science of Philadelphia); James Dean (National Museum of Natural History; Kim Bostwick (Cornell University Museum of Zoology); Kristof Zyskowski (Yale Peabody Museum); and Mark Robbins (University of Kansas Natural History Museum). Also, I am grateful to numerous collectors for their endeavors in amassing museum specimens and tissues.

Nate Rice was always responsive of my questions, and provided information on the plumage of *C. violaceus* from Ecuador and Peru. Paul Sweet and Joel Cracraft granted me a loan of their best *Cyanolyca* study skins. Adolfo Navarro-Sigüenza and Town Peterson provided access to the locality data in the Atlas of the Mexican Bird Distributions project. James Remsen, Nate Rice, Alexandre Alexio (Museu Paraense Emílio Goeldi, Brazil), Luis F. Silveira (Universidade de São Paulo, Brazil), Miguel Lentino (Colección Ornitológica Phelps, Venezuela), Celsa Señaris (Museo de Historia Natural La Salle, Venezuela), and A. Lander (Museo Estación Biológica Rancho Grande, Venezuela), provided locality data for particular areas, and Juan Freile went along the Museo Ecuatoriano de Ciencias Naturales getting me locality data from labels. The remaining collection locality data were downloaded from ORNIS.

This study was funded by grants from the National Science Foundation (Dissertation Improvement Grant DEB-0508910), the AMNH Frank Chapman Memorial Fund, and the KU Natural History Museum Panorama Society, and by the University of Kansas Ida Hyde Fellowship for Women in Science, and the University of Kansas Graduate School Dissertation Fellowship.

I am greatly indebted to Monica Papes, Arpi Nyari, Liza Holeski, Jeet Sukumaran, Charles Linkem, David and Heather McLeod, and Omar Torres for their friendship and for making life fun during all these years. I am grateful to my parents, who in spite of the distance have always been there for me. Finally, I want to thank Juan Manuel Guayasamin for his unconditional support and love, and for being my comrade, and my best friend.

TABLE OF CONTENTS

INTRODUCTION

The New World jays (NWJs) are a diverse group of corvids endemic to the Americas. Consisting of 34 biological species in seven genera—*Aphelocoma*, *Cyanocitta*, *Gymnorhinus*, *Cyanolyca*, *Calocitta*, *Psilorhinus*, and *Cyanocorax*—NWJ constitute species-rich assemblages in tropical, subtropical, and temperate habitats, and show a mix of narrow endemism (e.g., *Cyanocorax dickeyi*) and broad distributions (e.g., *Cyanocitta cristata*) (Madge and Burn, 1994). NWJ species range from solitary pair-breeders (e.g., *Cyanocitta* spp.) to fully cooperative breeders (e.g., *Aphelocoma unicolor*). Because of this broad spectrum of reproductive behaviors, they have served as model organisms for numerous analyses of the evolution of avian social systems (e.g., Brown, 1963; Fitzpatrick and Woolfenden, 1985; Saunders and Edwards, 2000).

As a consequence of their high diversity in plumage and morphology, the systematics of the NWJ genera has been controversial since the first comprehensive taxonomic treatments (e.g., Amadon 1944; Hardy, 1969). Two genera, *Gymnorhinus* and *Psilorhinus*, are monotypic, and another two, *Calocitta* and *Cyanocitta*, are represented by only two species. The genera *Aphelocoma*, *Cyanolyca,* and *Cyanocorax* consist of five, nine, and sixteen species, respectively. Many of these biological species, however, may represent complexes of multiple evolutionary species (Navarro-Sigüenza and Peterson, 2004).

To date, phylogenetic studies have focused mainly on species of *Aphelocoma* (Pitelka, 1951a; Peterson, 1992a, b; Brown, 1994; Rice et al., 2003; McCormack, unpub). Also, extensive research has been conducted on the ecology and behavior of *Gymnorhinus cyanocephalus* (e.g., Balda, 1971; Marzluff, 1988; Templeton et al., 1999),

Psilorhinus morio (e.g., Williams, 2004; Williams and Hale, 2006), *Calocitta* spp. (e.g., Langen, 1996a,b), and the Mesoamerican representatives of the *Cyanocorax* radiation (e.g., Raitt and Hardy, 1979; Winterstein and Raitt, 1983). In contrast, little is known about the phylogeny, ecology, and behavior of the South American *Cyanocorax,* and species of *Cyanolyca* in general.

Cyanolyca and *Cyanocorax* are the only corvid lineages that reached South America. They represent the most broadly distributed NWJ lineages, offering opportunities for understanding the origins of biogeographic patterns at continental scales. The genus *Cyanolyca* is a group of similar, blue jays that inhabit humid montane forest from Mexico south to Bolivia. Their fragmented distributions along the Neotropical mountains, makes them a model system for studying the relationship between allopatry and speciation. In contrast, jays in the genus *Cyanocorax,* together with species in the closely related genera *Calocitta* and *Psilorhinus*, show relatively high diversity in plumage patterns, and occupy a broad diversity of habitats (Hardy, 1969), but exist largely in parapatry and sympatry. As such, their study is relevant for understanding evolutionary patterns of plumage morphology and ecological niches.

Herein, I study the systematics and evolution of the NWJs, with emphasis *Cyanolyca* and *Cyanocorax*. My approach is based on a molecular phylogenetic perspective, which is discussed in the context of the morphological evidence from previous systematic studies (e.g., Zusi, 1987; Hope, 1989). In Chapter 1, I analyze molecular data to provide a general framework for relationships among NWJ genera. In Chapter 2, I generate a phylogenetic hypothesis of relationships for the genus *Cyanolyca*, test previous hypotheses of relationships, and reconstruct biogeographic scenarios for the

genus based on the phylogeny. In Chapter 3, I generate a phylogenetic hypothesis for relationships in *Cyanocorax* jays, and discuss it in terms of its systematic and biogeographic implications. Finally, in Chapter 4, I develop a comparative analysis of the ecological niches of species in *Cyanocorax*; these results are discussed in the context of their implications for the speciation of *Cyanocorax* in particular, and the mechanisms that may promote ecological diversification in general.

CHAPTER 1

A PHYLOGENETIC FRAMEWORK FOR RELATIONSHIPS AMONG NEW WORLD JAY GENERA

The study of relationships among the American jays, or so-called New World jays (NWJ), has occupied an important place in the ornithological taxonomic literature (e.g., Ridgway, 1904; Amadon, 1944; Hardy, 1961; Goodwin, 1976). However, early revisions, based on subjective analysis of morphological and behavioral characters, offered little insight on the evolutionary history of the group.

The first steps towards understanding the systematics of the NWJs began with an osteological revision recognizing synapomorphies that support the monophyly of the assemblage (Zusi, 1987). Subsequent phylogenetic analyses of morphological characters (Hope, 1989) and cytochrome *b* sequences (Espinosa de los Monteros and Cracraft, 1997) provided a basic structure for the NWJ tree, but most relationships afforded low statistical support; also, the molecular analysis did not include the monotypic genus *Psilorhinus*, and considered only one species per genus. A more detailed study based on the mitochondrial control region (Saunders and Edwards, 2000) and a phylogeny of Corvidae (Ericson et al., 2005) based on the β-Fibrinogen intron 7 (including five genera of NWJ), recovered different and novel relationships among taxa, which raised doubts about the true phylogeny of NWJ genera.

The present chapter offers a general view of the "deep" nodes of the NWJ phylogeny, based on reanalysis of existent, as well as novel sequence data from mitochondrial and nuclear loci, and much improved sampling of key species from each genus. This study lays the foundation for understanding the broad picture of NWJ systematics.

METHODS

Taxon and Gene Sampling

I analyzed at least two samples from each of the seven NWJ genera—*Aphelocoma*, *Cyanocitta*, *Calocitta*, *Cyanolyca, Cyanocorax* (including *Cissilopha*), *Psilorhinus*, and *Gymnorhinus*—and samples representing closely related corvid genera: *Dendrocitta*, *Pica*, *Perisoreus*, and *Corvus.* Also, in a first attempt to elucidate relationships among *Cyanocorax* jays, I analyzed six of the 16 currently recognized species of *Cyanocorax*, including two representatives of the previously recognized genus "*Cissilopha*", as well as Mesoamerican and South American species. Tissue samples were obtained from ornithological collections in the US and Mexico (Table 1.1).

Based on my own sequencing effort, and taking advantage of that of previous studies, I incorporated relatively fast evolving mitochondrial loci for resolution of branching patterns at the tips of the NWJ tree, as well as more slowly evolving nuclear loci, to illuminate ambiguities in phylogenetic position of some genera in the group. Complete sequences of the NADH Dehydrogenase Subunit 2 (ND2), the Adenylate Kinase gene intron 5 (AK5), the β-Fibrinogen intron 7 (βfib7), and the Transforming Growth Factor β-2 intron 5 (TGFβ2.5), and partial sequences of Cytochrome *b* (cyt*b*) were obtained for all genera of NWJ. These sequences were concatenated with published

sequences of the mitochondrial control region (CR) from Saunders and Edwards (2000), ND2 from Cicero and Johnson (2001), cyt*b* from Espinosa de los Monteros and Cracraft (1997), and βfib7 from Ericson et al. (2005).

DNA Amplification and Sequencing

Genomic DNA was extracted from frozen tissue with the DNeasyTissue extraction kit (Qiagen Inc.). PCR amplification was conducted using published primers (Appendix I). For amplifying mitochondrial genes, I used a standard protocol (94˚C/5 min; 35 cycles of 93˚C/1 min, 52˚C/1 min, 72˚C/2 min; and 72˚C/10 min). For amplification of βfib7, I followed Prychitko and Moore (1997) and for amplifying AK5, I modified the protocol of Shapiro and Dumbacher (2001), raising annealing temperatures to 64˚C in 2˚C intervals, to increase specificity of primer annealing and obtain single PCR products. TGFβ2.5 was amplified with a touchdown protocol (94˚C/ 3min; 5 cycles of 94˚C/30 sec, 60˚C/30 sec, 72˚C/40 sec; 5 cycles of 94˚C/30 sec, 56˚C/30 sec, 72˚C/40; 35 cycles of 96˚C/30 sec, 52˚C/30 sec, 72˚C/40 sec, and 72˚C/10 min; R. Moyle, pers. comm.). When multiple bands persisted, PCR products were purified using the QIAquick Gel Extraction Kit (Quiagen Inc.). Single PCR products were visualized in agarose gel, and unincorporated primers and DNTPs were removed with ExoSap-it (GE Health Care).

Cycle sequencing reactions were completed using the corresponding PCR primers and BigDye Terminator 3.1 chemistry (Applied Biosystems). I used a standard cycle sequencing profile (96˚C/3 min; 35 cycles of 96˚C/10 s, 50˚C/15 s, 60˚C/3 min; and 72˚C/7 min). Reaction products were purified with CleanSEQ magnetic beads (Agencourt) and run in an ABI Prism 3100 Genetic Analyzer (Applied Biosystems). Data

from heavy and light strand were spliced together to arrive at a consensus sequence for each sample, using Sequencher 4.1 (Gene Codes Corp., 2000). Precautions against amplifying nuclear pseudogenes of mitochondrial origin (Sorenson and Quinn, 1998) included sequencing both DNA strands and checking that amino acid translation was possible.

Alignment and Phylogenetic Analyses

Sequences were aligned in CLUSTAL_X (Thompson et al., 1997) and corrected by eye in MacClade ver. 4.0 (Maddison and Maddison, 2000). The nuclear introns showed indels of variable size, but could be aligned with minor adjustments. For the CR, however, I excluded a small fragment of CR (17–43 bp) that could not be aligned unambiguously.

Evolutionary rate heterogeneity across lineages was tested via the likelihood ratio (LR) test (Felsenstein, 1988). Significance was assessed by comparing $\Lambda = -2\log LR$, where LR is the difference between the –ln likelihood of the tree with and without enforcing a molecular clock, with a χ^2 distribution (n–2 degrees of freedom, where n is the number of taxa; Soltis et al., 2002). Departure from homogeneity in base frequencies among lineages was assessed with a χ^2 test. Both tests were conducted using PAUP ver 4.0b (Swofford, 2000). To evaluate possible saturation of cyt*b* and ND2 at high levels of sequence divergence, uncorrected and ML distances of these two genes were plotted against βfib7 distances.

Phylogenetic analyses were conducted using Maximum Parsimony (MP), Maximum Likelihood (ML), and Bayesian analyses (BA) for individual genes, as well as

for a combined nuclear and mitochondrial dataset. Parsimony analyses of the mitochondrial genes and the combined analysis were performed in PAUP as heuristic searches (10,000 stepwise random additions with TBR branch-swapping) and clade support was estimated via 1000 bootstrap pseudo-replicates (Felsenstein, 1985) with 100 random additions. Analyses of nuclear genes followed similar methods, but trees and clade support were obtained using branch-and-bound searches. For nuclear genes, multiple base indels were coded as missing data, and new binary characters for each unique gap (0 = absent, 1 = present) were added to the end of the data matrix. To evaluate whether such indels supported the same branches as did single-nucleotide substitutions, phylogenetic trees were constructed also from sequence matrices in which alignment gaps were coded as missing data. Double picks in nuclear gene sequences, reflecting heterozygous positions, were coded with IUPAC degeneracy codes and treated as polymorphisms.

Previous to ML and BA analyses, the best-fit models of nucleotide substitution for each gene and the combined dataset (ML) were selected using ModelTest ver 3.7 (Posada and Crandall, 1998) under the Akaike information criterium (AIC), following recent recommendations (Posada and Buckley, 2004). Maximum likelihood analyses were run in PAUP under the appropriate model and model parameter values with 100 random additions. Node support was assessed via 100 bootstrap replicates, with an initial tree generated by neighbor joining.

Bayesian analyses were performed in Mr Bayes 3.1 (Ronquist and Huelsenbeck, 2003). Mitochondrial genes were analyzed individually, partitioned by codon position for cyt*b* and ND2, and by domain (domain I, central domain, and domain II) for CR. Each

analysis consisted of 2×10^6 generations and four Markov chains with default heating values. Model parameter values were estimated from the data and initiated with flat priors. Trees were sampled every 1000 generations, resulting in 2000 saved trees per analysis, of which 500 were discarded as "burn-in." Stationarity was confirmed by plotting the –ln L per generation. Additionally, I confirmed that the average standard deviation of split frequencies approached zero.

The BA for the combined mitochondrial-nuclear dataset followed a similar scheme with 12 partitions: cyt*b* and ND2 by codon, CR by domain, AK5, βfib7, and TGFβ2.5. Each partition was assigned its best-fit model with all parameters unlinked except for topology and branch lengths (i.e., model parameters estimated separately for each partition). To reduce the chance of converging on local optima, four independent runs of 2×10^6generations were performed. After confirming that all analyses reached stationarity at similar likelihood scores and that the topologies were similar, the resulting 6000 trees (8000 minus burn-in) were used to calculate posterior probabilities (PP) in a 50% majority-rule consensus tree.

For combined analyses, sequences of *Perisoreus canadensis* were concatenated with the published CR sequence of *Perisoreus infaustus*, and sequences of *Pica hudsonia* were concatenated with that of *Pica nuttallii*. Although both the *Perisoreus* and *Pica* combined sequence datasets are technically chimeric, their appropriateness to root the tree is justified because the two pairs of the chimera are in all probability more closely related to one another than to those of other species in the all-species dataset.

Topological Congruence and Combinability

Each dataset produced a different topology regarding the position of *Aphelocoma*, *Cyanocitta,* and *Gymnorhinus* (hereafter "ACG"). To assess potential causes of these differences, I compared the support (bootstrap and/or posterior probabilities) of conflicting topologies obtained from analysis of individual genes, and evaluated the consistency of characters relative to trees of different topologies via consistency indices (CI) and rescaled consistency indices (RC).

I performed series of Shimodaira-Hasegawa tests of hypothesis (SH; Shimodaira and Hasegawa, 1999) to test whether the individual-gene ML topologies were statistically better than alternate topologies generated by all other genes. For each gene, I obtained the ML tree using the best-fit model of evolution and estimated parameter values, and used the same data to perform constraint ML searches under the topologies generated by other genes (i.e., set of null hypotheses). Then, likelihood values of the ML and the constrained trees were compared using the SH test, as implemented in PAUP (rell optimization, 1000 bootstrap replicates). Strong nodal support (>70%; Hillis and Bull, 1993) and significance of likelihood differences (SH test) were taken as indicatives of conflict on phylogenetic signal. I avoided using the incongruence length difference test (ILD test; Farris et al., 1994) as test of dataset homogeneity given recent criticisms (e.g., Baker and Lutzoni, 2002; Darlu and Lecointre, 2002).

Historical Biogeography of the NWJs

To explore implications of the phylogeny for the historical biogeography of the NJWs, I optimized the known distribution of each species onto the final tree using Dispersal-Vicariance Analysis in DIVA 1.1 (Ronquist 1996, 1997). This analysis uses a three-dimensional step matrix based on a simple biogeographic model to reconstruct ancestral distributions in a given phylogeny. Each species was coded as preset/absent on each of three regions: North America, Mesoamerica, and South America.

Although the closest relative of the NWJ has not been determined, molecular and morphological evidence point towards an Asian ancestor. As such, the ancestral area occupied by the most recent ancestor would be uninformative in the present analysis. Therefore, I optimized all areas on the ingroup tree, using up to two possible ancestral areas (max-areas command = 2). To improve optimization of ancestral areas, I expanded taxon representation based on previous analyses of the phylogeny of *Aphelocoma* (Rice et al., 2003) and preliminary ND2 sequence data on *Cyanolyca* (E. Bonaccorso, unpub.).

RESULTS

Sequence Attributes

DNA sequence lengths and general characteristics for each gene are summarized in Table 1.2. As expected, sequence variation was comparable among the mitochondrial genes, and substantially higher than that of the nuclear introns. Pairwise distances for ND2, cyt*b*, and βfib7 are summarized in Figure 1.1. Among interesting features of divergence is that when plotting uncorrected distances, ND2 and cyt*b* saturate early compared to βfib7; however, saturation is corrected when using ML distances for both

genes. To reduce possible effects of saturation, parsimony analyses for ND2 and cyt*b* were performed including only outgroup taxa most closely related to the NWJs (i.e., *Perisoreus* and *Dendrocitta*).

Models of nucleotide substitution selected by MODELTEST, as well as model parameter values, are displayed in Table 1.2. For the effects of the partitioned Bayesian analyses, further MODELTEST analyses were performed for codon positions (ND2 and cyt*b*) and domains (CR). The GTR + I + Γ model was estimated as the best-fit model of substitution for all partitions, with exception of Domain I (for which GTR + Γ was better fit). Tests of homogeneity of base frequencies across taxa were not significant for any gene, even when codons and domains were considered separately ($P > 0.90$). Rate heterogeneity among lineages was detected for ND2 ($P < 0.05$), and more dramatically for cyt*b* and CR ($P < 0.001$; Table 2.2), but not for the nuclear loci.

The nuclear introns showed a number of indels, the most interesting being: a synapomorphic indel of 1 bp in *Cyanolyca* (βfib7); a long 78-bp indel in *Dendrocitta formosae* (among other 8 indels, length 2–5) and a 5-bp indel in all species of *Cyanocorax*, *Calocitta*, and *Psilorhinus* (AK5); and a 2-bp indel uniting *Calocitta* and *Psilorhinus* and a 4-bp indel uniting *Perisoreus* and *Dendrocitta* (TGFβ2.5). Three samples/genes could not be sequenced—*Cyanocorax yncas* for βfib7, and *Calocitta coillei*, and *Cyanocitta stelleri* for AK5. Finally, because sequence variation detected for TGFβ2.5 was lower than for the other nuclear genes, fewer representatives (per genus) were included in analyses.

Phylogenetic Analyses of Individual Genes

All analyses recovered the monophyly of NWJs with relatively high support. For ND2 (Fig. 1.2), ML and BA trees confirmed the position of *Cyanolyca* as sister to all other NWJs, which divide further into two clades: one including *Cyanocorax*, *Calocitta*, and *Psilorhinus*, and another including *Aphelocoma*, *Cyanocitta*, and *Gymnorhinus* (ACG). Within the first clade, *Calocitta* and *Psilorhinus* show as reciprocally monophyletic and sister to *Cyanocorax*. Within *Cyanocorax*, *C. yncas* is sister to a clade in which *C. melanocyaneus* and *C. yucatanicus* (two of the four "*Cissilophas*") are monophyletic and sister to the remaining *Cyanocorax* species. In the second clade, ML and BA trees place *Gymnorhinus* as basal to *Cyanocitta* and *Aphelocoma*, whereas the MP tree places *Cyanocitta* as basal; the latter arrangement represented the only difference between the MP 50% majority rule consensus and the ML and BA trees.

For cyt*b* (tree not shown), the MP 50% majority rule consensus tree recovers the same topology and similar nodal support as the ND2 MP consensus tree. However, the ML and BA trees show *Aphelocoma* and *Gymnorhinus* as sisters, but fail to recover the ACG clade, creating a polytomy among *Cyanocitta* and the clades formed by *Aphelocoma* + *Gymnorhinus*, and *Cyanocorax* + *Calocitta* + *Psilorhinus*. Topologies resulting from all CR analyses are identical to those obtained by Saunders and Edwards (2000); *Cyanocitta* and *Aphelocoma* show as monophyletic and sister to a clade containing *Gymnorhinus*, *Calocitta*, *Psilorhinus*, and *Cyanocorax*, with *Gymnorninus* being at the base of this latter group. Nodal support for major clades obtained in all analyses is summarized in Table 1.3.

The nuclear genes provided less resolved structures. Analyses of βfib7 and TGFβ2.5 produced identical topologies with or without indels included, whereas in the AK5 tree, the *Cyanocorax* + *Calocitta* + *Psilorhinus* clade was not supported by nucleotide variation, but by the 5-bp indel described before (Fig. 1.3). All nuclear genes reconstructed a monophyletic *Calocitta* + *Psilorhinus*, and TGFβ2.5 supported a monophyletic *Cyanocorax*. The basal position of *Cyanolyca* among NWJs was recovered by AK5 and TGFβ2.5, but not by βfib7, and the clade ACG was reconstructed by βfib7 and TGFβ2.5, but not by AK5. Relationships among *Cyanocorax* species were inconsistent across genes.

Regarding levels of homoplasy among datasets, all mitochondrial genes showed similar levels of character consistency as measured by their CI and RC (Table 1.3). The three nuclear genes had similar levels of consistency, and were about half homoplastic than mitochondrial genes.

Topological Congruence and Combined Analysis

In summary, most phylogenetic conflict centered on the positions of ACG (Table 1.3). Whereas data from ND2, cyt*b*, βfib7, and TGFβ2.5 generated an ACG clade with variable degrees of support, CR placed *Gymnnorhinus* as sister to *Cyanocorax* + *Calocitta* + *Psilorhinus* with relatively high support. However, because CR did not reject the ACG grouping (Shimodaira-Hasegawa test, $P = 0.275$), I explored the implications of an ACG clade further. Relationships among these three genera varied among analyses, as follows: whereas ND2 (ML and BA), AK5, and CR supported *Aphelocoma* + *Cyanocitta* (AC), cyt*b* and ND2 (MP) supported *Aphelocoma* + *Gymonorhinus* (AG), and βfib7

supported *Cyanocitta* + *Gymnorhinus* (CG); TGFβ2.5 recovered a polytomy among these three taxa.

Table 1.4 shows synapomorphic characters supporting each of the possible arrangements among ACG, as recovered by the MP 50% majority rule consensus trees. Parsimony analyses showed that CR has the higher average CI, with synapomorphic characters uniformly distributed across the fast-evolving domain I, and II, and the more conserved central domain. In ND2 and cyt*b*, all changes except for one are in third-codon positions, and at least half are transitions. On the other hand, AK5 and βfib7 support their topologies with 1 and 3 unreversed changes respectively. Finally, from a total of 41 statistical comparisons performed via the SH test (6 sets, each containing 5–7 null hypotheses) topological congruence was rejected only when CR was constrained to the arrangements implying AG and GC ($P < 0.05$), indicating that CR data support only the AC arrangement. Given that according to the SH test, topological conflict was restricted to AGC, I combined all datasets in a single analysis, which produced the same topology as the ND2 ML and BA trees (Fig.1.4).

Historical Biogeography of the NWJs

Optimization of ancestral areas using DIVA produced an exact solution that required 6 dispersal events. According to the analysis, the ancestral distributional area of the NWJs is restricted either to Mesoamerica or Mesoamerica + North America (Fig. 1.5). The *Cyanolyca* radiation originated in Mesoamerica and dispersed into South America. Also, the *Cyanocorax* + *Calocitta* + *Psilorhinus* clade originated in Mesoamerica and dispersed into South America in the ancestor of *Cyanocorax cayanus* +

C. chrysops, as well as independently in *C. yncas*. Finally, the origin of ACG was reconstructed as ambiguous between North America and Mesoamerica.

DISCUSSION

The phylogenetic tree of the combined analysis (Fig. 1.4) depicts intergeneric relationships among NWJs. Relatively weak nodes of the tree are restricted to one clade only—that including *Aphelocoma*, *Cyanocitta*, and *Gymnorhinus*. The low support for the monophyly of ACG is a consequence of CR supporting a topology that differs from those of the ND2, cyt*b*, βfib7, and TGFβ2.5. Because CR is physically linked to other mitochondrial loci, these topological differences should not result from independent phylogenetic history. Also, in spite of the results of the SH test, which renders the CR topology as not significantly different from the alternate topology, the high bootstrap support afforded by the CR topology in all analyses, suggests that these differences are not caused by stochastic error (Bull et al., 1993; de Queiroz et al., 1995).

Other possible causes of phylogenetic incongruence include: nucleotide base-composition bias across taxa (reviewed in Van Den Bussche et al., 1998), rate heterogeneity among lineages (Wendel and Doyle, 1998; Soltis et al., 2002), and, of course, homoplasy. Compositional bias was not significant for any loci analyzed and, although CR sequences deviated from the clock assumption, the same phenomenon is observed in both ND2 and cyt*b*. Also, its levels of homoplasy (measured as CI, and RC) are similar to those observed for the other mitochondrial genes (Tables 1.3 and 1.4). Thus, it is difficult to assess the origin of these topological differences.

Despite of the results obtained from the analyses of CR, different lines of evidence support the idea of ACG forming a monophyletic group. In addition to the nuclear and mitochondrial data presented herein, morphological synapomorphies in the squamosal-quadrate articulation and the associated temporal fossa crests, as well as the overall shape of the cranium (Hope, 1989), support a single evolutionary origin of the group.

With regard to relationships within the ACG clade, conflict in phylogenetic signal seems to be the result of stochastic error and/or loci reflecting different phylogenetic histories. For all genes, except for βfib7, bootstrap support for relationships among these genera was relatively low, suggesting that differences may be caused by stochastic error (Bull et al., 1993; de Queiroz et al., 1995). In the case of βfib7, topological discrepancies could be caused by an idiosyncratic evolutionary path. In all cases, no statistical differences were found among alternate topologies, suggesting that they are all equally good explanations of the data. Still, careful interpretation of these results should derive from the highly conservative nature of the SH test (e.g., Goldman et al., 2000; Shi et al., 2005).

Interestingly, the *Aphelocoma* + *Cyanocitta* arrangement coincides with a unique morphological trait: a lateral bar to the sclerotic ring in *Cyanocitta* (Curtis and Miller, 1938) reported herein for *Aphelocoma*. This feature was present in all species of *Cyanocitta* (>200 specimens) and *Aphelocoma* (>600 specimens), but absent in all other NWJ genera (> 100 specimens) examined by A. T. Peterson (unpubl. data). This morphological novelty, unique among birds, may serve as a synapomorphy uniting *Aphelocoma* and *Cyanocitta*, to the exclusion of *Gymnorhinus*. From the molecular point

of view, however, resolving this polytomy may require analysis of relatively large amounts of data (e.g., McCracken and Sorenson, 2005; but see Jeffroy et al., 2006).

Most other NWJ relationships reconstructed are well-supported, and confirm some of the results of previous studies. *Cyanolyca* is sister to all other NWJ, and ACG and *Cyanocorax* + *Calocitta* + *Psilorhinus* show as reciprocally monophyletic. As in previous analyses based on CR sequences (Saunders and Edwards, 2000), ND2 and cyt*b* data support *Psilorhinus* as a valid entity and not as nested within *Cyanocorax* (as suggested by A.O.U., 1983). Consistent with these results, Sutton and Gilberg (1942) described a unique morphological character in *Psilorhinus*: the "furcular pouch," a structure formed by the hyperthrophy of the cleido-traquialis muscles that creates a median, non-paired extra interclavicular diverticulum. This structure is not present in any other corvid, and constitutes a clear, discrete morphological character diagnosing *Psilorhinus*. Analysis of more *Cyanocorax* species is crucial to confirm the monophyly of *Cyanocorax* and the taxonomic validity of the genus *Psilorhinus*.

Finally, several phylogenetic hypotheses of relationships were recovered within *Cyanocorax*. First, *C. melanocyaneus* and *C. yucatanicus* were reconstructed as most closely related to each other, which is coherent with their previous recognition as part of the subgenus *Cissilopha*. Second, the Mesoamerican species *Cyanocorax dickeyi* was placed as most closely related to the South American *C. chrysops* and *C. cayanus*. Finally, *C. yncas* was placed at the base of the *Cyanocorax* tree, consistent with its former recognition as a monotypic genus (*Xanthoura*); given that this latter species is the only green jay in the assemblage of American jays, its position *Cyanocorax* tree suggests interesting implications for the evolution of plumage color.

Only the *Cyanocorax dickeyi* + *C. chrysops* and the *C. dickeyi* + *C. chrysops* + *C. cayanus* clades were recovered by AK5 and βfib7, respectively. This result is perhaps understandable, given the apparently short internodal branches in the *Cyanocorax* tree. Given that the faster-evolving mitochondrial genes have a better chance of tracking the species tree through short internodes than nuclear genes (Moore, 1995), this preliminary analysis may be a correct approximation of the *Cyanocorax* tree. Of course, these initial hypotheses need to be confirmed in a more comprehensive analysis, adding the remaining two "*Cissilopha*" species (*C. sanblasianus* and *C. beecheii*) and the rest of the South American species, as well as multiple individuals of all species.

Historical Biogeography of the NWJs

According to the DIVA analysis (Fig. 1.5), the NWJs originated either in Mesoamerica or North America + Mesoamerica, and ACG originated in North America or North America + Mesoamerica. The genus *Cyanolyca*, and the core *Cyanocorax* + *Calocitta* + *Psilorhinus* clade originated in Mesoamerica, whereas *Cyanocorax yncas*, the South American *Cyanolyca*, and the South American *Cyanocorax* each represent independent invasions of South America. Whether these radiations occurred independently in time or resulted from the same biogeographic event (e.g., the formation of the Panama Isthmus) is a difficult question still in need of answer. Given the high rate heterogeneity known in this group (Peterson, 1992) and demonstrated here, we reframed from the usual exercise of assigning dates to splits or applying a relative-time scheme using a molecular "clock".

Another aspect that remains unresolved regards the closest corvid relative of the NWJ radiation. Although both nuclear introns placed *Perisoreus* and *Dendrocitta* as most closely related to the NWJs (Fig. 1.3), more corvid outgroups need to be included in the analysis. Based on the levels of saturation observed in the mitochondrial genes at high sequence divergences, this question needs to be approached using multiple, fast-evolving nuclear loci. In any case, as proposed before (Pitelka, 1951a; Hope, 1989; Espinoza de los Monteros and Cracraft, 1997) the origin of the NWJs probably looks back to an Asian ancestor.

Conclusions

Combined analysis of three mitochondrial and three nuclear genes agree with morphological synapomorphies supporting the monophyly of the NWJ and the early divergence of *Cyanolyca*. The general structure of the phylogenetic tree indicates a deep separation of the remaining species into two clades, one containing *Aphelocoma*, *Cyanocitta*, and *Gymnorninus*, and another containing *Cyanocorax*, *Calocitta*, and *Psilorhinus*. Although robust topological differences were observed from the analyses of the Control Region, morphological and most of the molecular data available are consistent with this basic phylogenetic structure.

Branching order in the *Aphelocoma* + *Cyanocitta* + *Gymnorninus* assemblage is not clear based on the molecular data, although a unique structure in the sclerotic ring supports a closer relationship between *Cyanocitta* and *Aphelocoma*. Details on the relationships between *Cyanocorax* and allied genera *Psilorhinus* and *Calocitta*, as well as the taxonomic validity of the genus *Psilorhinus* need to be assessed with improved sampling of *Cyanocorax* species. Finally, although limited by the data at hand, the biogeographic reconstruction revealed complex patterns of colonization in which at least three independent lineages have invaded South America from Mesoamerica.

CHAPTER 2

A MOLECULAR PHYLOGENY OF THE *CYANOLYCA* JAYS AND ITS IMPLICATIONS FOR HISTORICAL BIOGEOGRAPHY AND SPECIATION

Neotropical montane regions hold the world's highest diversity of birds, as well as that of many other organisms (Churchill et al., 1995; Stattersfield, 1998). These mountain chains stretch from Mexico south to Argentina and Chile, in a fragmented, complex mosaic of topographic units of diverse geologic origins (Simpson, 1975; Ferrusquía-Villafranca, 1993; Coates and Obando, 1996; Gregory-Wodzicki, 2000). Regardless of their origins, these regions sustain extensive tropical montane forests (Churchill et al., 1995) and numerous avian lineages that overlap broadly in areas with similar environmental conditions (Chapman, 1926; Hernández-Baños et al., 1995; Peterson et al., 1999).

Early distributional studies and more recent empirical work suggest that Andean avifaunas are derived, at least partially, from lineages that have moved from lower to higher montane elevations (Chapman, 1926; Gerwin and Zink, 1989; Bates and Zink, 1994; García-Moreno et al., 1999a; Pérez-Emán, 2005; Brumfield and Edwards, 2007) and from lineages that have expanded their distributions via the Panama Land Bridge (Chapman, 1917; Haffer, 1974; Hackett, 1995). Although these hypotheses originally were formulated to explain biogeographic patterns of Andean taxa, they may be applied to Mesoamerican montane taxa as well. Moreover, the complex topography and fragmented nature of Neotropical montane forests suggest that diversification in situ after initial biological interchange might play a decisive role in shaping distributions of largely

overlapping lineages (Chapman, 1926; Remsen, 1984; Cracraft, 1985; Hernández-Baños et al., 1995; García-Moreno and Fjeldså, 2000).

Most models that attempt to explain geographic variation and speciation in situ are based on: (1) effects of deep river valleys as barriers to gene flow and consequent evolution of distinctive geographic forms (Chapman, 1926; Vuilleumier, 1969; Remsen, 1984; Cracraft, 1985; García-Moreno and Fjeldså, 2000); (2) the linearity of the Andes, which results in elongate geographical ranges and reduces potential contact and gene flow among parapatric forms (Remsen, 1984; Graves 1985, 1988); and (3) effects of Pleistocene glaciations on the cyclic fragmentation, isolation, and reconnection of montane forests (Hooghiemstra et al., 2000) and their avifaunas (Vuilleumier, 1969; Haffer, 1974; Hackett, 1995). Clearly, these propositions are not mutually exclusive, and all could operate across various temporal, spatial, and taxonomic scales.

Cyanolyca jays are model organisms for testing hypotheses of diversification across the Neotropical montane forests. It is a relatively small assemblage that represents one of the two New World jay (NWJ) lineages that reached South America. *Cyanolyca* jays are sedentary, and inhabit humid montane forests from Mexico south to Bolivia (Fig. 2.1). Most are allopatric, and their ranges are highly subdivided, creating (putatively) isolated and morphologically distinct populations (Hellmayr, 1934). Thus, study of the relationships among *Cyanolyca* jays not only provides fertile grounds for testing biogeographic hypotheses, but also opportunities to observe morphological and molecular evolution in progress.

Current taxonomic treatments (e.g., Madge and Burn, 1994; Sibley and Monroe, 1990; Dickinson, 2003) recognize nine species: *Cyanolyca mirabilis*, *C. nana*, *C. pumilo*,

C. argentigula, *C. pulchra*, *C. cucullata*, *C. armillata*, *C. turcosa*, and *C. viridicyanus*. The first four species, the so-called "dwarf jays," are allopatric and have been recognized as full species since early treatments (e.g., Hellmayr, 1934; Blake and Vaurie, 1962; Goodwin, 1976). The taxonomic status of the remaining species is more problematic. Hellmayr (1934) treated the Mesoamerican *C. cucullata* and the South American *C. pulchra* as conspecific. Blake and Vaurie (1962) considered *C. turcosa* and *C. armillata* as subspecies of *C. viridicyanus*, whereas others (Hellmayr, 1934; Fjeldså and Krabbe, 1990) recognized *C. turcosa* as a valid species, but lumped *C. armillata* and *C. viridicyanus* into a single species. Other authors have recognized all five forms as full species, based on discrete plumage differences (Ridgely and Tudor, 1994), vocalizations (Goodwin, 1976), and geographic ranges (Zimmer, 1953).

The only (non-phylognetic) hypothesis of relationships within *Cyanolyca* is that of Goodwin (1976), who proposed that dwarf jays arose from a single ancestor, with *C. mirabilis* and *C. argentigula* as sister species. In his arrangement, *C. cucullata* and the South American forms are closely related, with *C. cucullata* and *C. pulchra* as sister species. With regard to the remaining South American taxa, he suggested that *C. viridicyanus* and *C. armillata* he may form a superspecies, which is the sister clade to *C. turcosa*. This arrangement was based on subjective summary of overall plumage similarity, and therefore is subject to observation that plumage characters are extremely labile among birds (Hackett and Rosenberg, 1990; Burns, 1998; Omland and Lanyon, 2000).

To date, neither the relationships among species, nor the validation of current species as independent historical entities (sensu Wiley, 1978) have been approached in a phylogenetic context. A higher-level phylogenetic analysis of the mitochondrial Control Region revealed high sequence divergence among *Cyanolyca* species compared to divergences in related genera (Saunders and Edwards, 2000), which may indicate that speciation in *Cyanolyca* occurred deeper in history, but without major morphological change. Therefore, assessments of relationships based on overall morphological similarity may not reflect the complexity and evolutionary history of lineages in the group.

Herein, I study phylogenetic relationships among *Cyanolyca* jays based on the analysis of three mitochondrial and two nuclear genes. Based on the phylogenetic results, I test previous hypotheses of relationships among recognized species, and identify potential independent lineages that might represent different evolutionary species (Wiley, 1978). Finally, I reconstruct the ancestral distributional areas for *Cyanolyca*, and discuss speciation scenarios in the context of the biogeography of Neotropical montane avifaunas.

METHODS

Taxon and Gene Sampling

I analyzed a total of 40 samples of *Cyanolyca*, including at least one representative of each species and representatives of most subspecies. For geographically widespread or polytypic taxa (e.g., *C. turcosa, C. cucullata, C. armillata,* and *C. viridicyanus*), sampling spanned geographic populations to encompass genetic variation

among the extremes of their distributions. For *C. turcosa* and *C. viridicyanus,* sampling was relatively dense, covering most of the geographic range. Identification of samples to subspecies was based on locality data. Tissue samples were obtained from ornithological collections in the U.S. and Mexico, as well as from my own collecting efforts in Mexico, Ecuador, and Venezuela.

DNA sequences used for outgroup comparisons were obtained from previous studies (Espinosa de los Monteros and Cracraft, 1997; Cicero and Johnson, 2001; Helm-Bychowski and Cracraft, 1993; Ericson et al., 2005; Bonaccorso and Peterson, 2007), and included species representing all New World jay genera (*Aphelocoma*, *Cyanocitta*, *Gymnorhinus*, *Calocitta*, and *Cyanocorax*), and more distantly related corvid genera (*Dendrocitta* and *Perisoreus*); a novel Control Region sequence for *Dendrocitta formosae* was generated to complete the outgroup dataset (GenBank Accession number XXX-to be added upon acceptance of the paper).

A combination of mitochondrial and nuclear markers was used to resolve relationships at both terminal nodes and deep branches in the phylogeny of *Cyanolyca* jays. I obtained sequences of the mitochondrial Control Region (CR) for the full dataset, and of the NADH dehydrogenase subunit 2 (ND2) for 38 individuals. Further analyses were directed towards obtaining sequences of representative individuals for Cytochrome b (cyt*b*), and the nuclear genes Adenylate Kinase intron 5 (AK5), and β-Fibrinogen intron 7 (βfb7). Information on genes sequenced and GenBank accession numbers, as well as voucher specimen information and associated locality data, is summarized in Table 1.1. For the combined analyses (described below), sequences of *Perisoreus canadensis* for cyt*b*, ND2, and the two introns were concatenated with the published CR sequence of

Perisoreus infaustus. Although combined sequence data is technically chimeric, it appropriateness to root the tree is justified because the two pairs of the chimera are in all probability more closely related to one another than to those of other species in the dataset.

DNA Amplification and Sequencing

Genomic DNA was extracted from frozen tissue with the DNeasyTissue extraction kit (Qiagen Inc.) or a modified salt precipitation method (M. Fujita, unpubl.) based on the Puregene DNA purification kit (Gentra Systems). Because frozen tissue samples were not available for *Cyanolyca nana*, DNA for this species was obtained from a museum skin sample (KU 106856) courtesy of R. Fleischer in the laboratories of the National Museum of Natural History and National Zoological Park, using established protocols (Fleischer et al., 2000, 2001).

DNA amplification was completed using the following primer pairs: L5216 and H6313 (Sorenson et al., 1999) for ND2; JCR-13 (Saunders and Edwards, 2000) and H1248 (Tarr, 1995) for CR; L14990 (Kocher et al., 1989) and H16065 (T. Birt, unpubl.) for cyt*b*; FIB-B17U and FIB-B17L (Prychitko and Moore, 1997) for βfb7; and AK5b+ and AK6c- (Shapiro and Dumbacher, 2001) for AK5. Given that DNA extracted from museum skins usually is degraded into small fragments, amplification of DNA from the skin sample of *Cyanolyca nana* was performed using primers that were in close proximity to each other. Primer pairs used are summarized the Appendix.

I used a common PCR protocol (94˚C/5 min; 35 cycles of 93˚C/1 min, 52˚C/1 min, 72˚C/2 min; and 72˚C/10 min) for all mitochondrial genes. βfb7 was amplified

following protocols in Prychitko and Moore (1997), and AK5 was amplified using a touchdown protocol (94˚C/ 3min; 5 cycles of 94˚C/30 sec, 60˚C/30 sec, 72˚C/40 sec; 5 cycles of 94˚C/30 sec, 56˚C/30 sec, 72˚C/40; 35 cycles of 96˚C/30 sec, 52˚C/30 sec, 72˚C/40 sec, and 72˚C/10 min; R. Moyle pers. comm.). When multiple bands persisted, PCR products were purified using the QIAquick Gel Extraction Kit (Quiagen, Inc.). Single PCR products were visualized in 7 % agarose gel, and unincorporated primers and DNTPs were removed from PCR products with ExoSap (ExoSap-it, GE Health Care).

Cycle sequencing was completed with the corresponding PCR primers and BigDye Terminator 3.1 chemistry (Applied Biosystems) using a standard cycle sequencing profile (96˚C/3 min; 35 cycles of 96˚C/10 s, 50˚C/15 s, 60˚C/3 min; and 72˚C/7 min). Sequencing reaction products were purified with CleanSEQ magnetic beads (Agencourt) and resolved in an ABI Prism 3100 Genetic Analyzer (Applied Biosystems). Data from heavy and light strands were spliced together to arrive at a consensus sequence for each sample, using Sequencher 4.1 (Gene Codes Corp., 2000).

Aligning and Phylogenetic Analyses

Nucleotide sequences were aligned in CLUSTAL X (Thompson et al., 1997). MacClade ver. 4.0 (Maddison and Maddison, 2000) was used to adjust alignments by eye, as well as to translate nucleotide sequences into amino acids. To explore data prior to further analyses, best-fit models of molecular evolution were estimated in MODELTEST v.3.7 (Posada and Crandall, 2001) via the Akaike Information Criterion (AIC), for each gene and for a combined mitochondrial and nuclear dataset. Model parameters estimated from MODELTEST were used in further maximum likelihood

analyses (ML, described below), and model parameter estimates for ND2 were used for obtaining pairwise ML-corrected distances in PAUP* (Swofford, 2000) to compare levels of divergence between individual samples.

Phylogenetic reconstructions were conducted using maximum parsimony (MP) and ML for individual genes, and MP, ML, and Bayesian inference for the combined dataset. Comparisons of individual gene trees and their non-parametric bootstrap support were used as a measure of congruence of phylogenetic signal among datasets (Wiens, 1998). This approach relies on the argument that clades that are strongly supported and in conflict between datasets may be indicative of differences in underlying phylogenetic histories, whereas weakly supported conflicts may simply result from stochastic error (Bull et al., 1993; de Queiroz et al., 1995). "Strong" incongruence was identified by the presence of conflicting nodes showing 70% or more non-parametric bootstrap support (Felsenstein, 1985) or 0.95 Bayesian posterior probabilities.

Also, given that base composition and evolutionary rate heterogeneity across lineages may be important sources of phylogenetic incongruence (Sidow and Wilson, 1990; Loomis and Smith, 1992; Lockhart et al., 1994; Wendel and Doyle, 1998; Soltis et al., 2002), datasets were further explored to detect departures from base and rate homogeneity. Base frequencies were examined for nucleotide bias among taxa, using the χ^2 test of homogeneity in PAUP* (Swofford 2000). Evolutionary rate heterogeneity was tested via a likelihood ratio test (Felsenstein, 1981) with significance assessed by comparing $\Lambda = -2\log LR$, where LR is the difference between the –ln likelihood of the tree with and without enforcing a molecular clock, with a χ^2 distribution ($n - 2$ degrees of freedom, where n is the number of taxa).

Parsimony analyses were performed coding gaps as missing data; double picks in nuclear gene sequences, reflecting heterozygous positions, were coded with IUPAC degeneracy codes and treated as polymorphisms. Trees were obtained through heuristic searches in PAUP* (Swofford, 2000) using 10000 stepwise random additions (TBR branch-swapping). Clade support was estimated with heuristic searches using 1000 bootstrap pseudoreplicates (Felsenstein, 1985), each pseudoreplicate consisting of 100 stepwise random additions.

Maximum likelihood trees were estimated using GARLI (Genetic Algorithm for Rapid Likelihood Inference, ver. 0.951; Zwickl, 2006), which provides considerable advantages over PAUP* in terms of computational efficiency. It uses a genetic algorithm that finds the tree topology, branch lengths, and model parameters that maximize lnL simultaneously. This process involves evolution of a population of solutions termed "individuals," with each individual encoding a tree topology, a set of branch lengths, and a set of model parameters. Each individual is assigned a fitness based on its lnL score. Each generation, random mutations are applied to some of the components of the individuals, and their fitnesses recalculated. A subset of individuals is then chosen to be the parents of the individuals of the next generation in proportion to individual fitness values. This process is repeated several times (described below), and the population of individuals evolves toward higher fitness solutions. The highest-fitness individual is automatically maintained in the population, ensuring that it is not lost by chance (http://www.bio.utexas.edu/faculty/antisense/garli/GARLIv0.95manual.pdf).

Available models of nucleotide substitution include the General Time Reversible (GTR) model and its more common submodels, as well as less complex models. It also

accounts for gamma-distributed rate heterogeneity (Γ) with a specified number of rate categories and estimation of the proportion of invariable sites (I). Estimation of model parameter values may be optimized or fixed, and the implementation of the model is equivalent to that in PAUP*.

GARLI analyses for individual genes and the combined dataset were conducted specifying the model "family" obtained by MODELTEST, but allowing the program to estimate parameter values from the data. In cases in which a gamma distribution was implemented, the number of rate categories was set to four. Individual solutions were selected after 10000 generations with no significant improvement in likelihood, with the significant topological improvement level set at 0.01 (first condition for termination); then, the final solution was selected when the total improvement in likelihood score was lower than 0.05, compared to the last solution obtained (second condition for termination). All other GARLI settings involved in the genetic algorithm were default values, as per recommendations of the developer (Zwickl, 2006). For each dataset, I ran 10 independent analyses starting with random trees, and selected the tree with the highest ML score. Final tree score and parameter estimates were obtained by optimizing the tree and branch lengths in PAUP*. Bootstrap support was assessed via 100 and 1000 pseudoreplicates for the individual gene and the combined datasets, respectively; bootstrap searches ran under the same settings used for obtaining the best ML tree.

Bayesian analyses were performed in Mr Bayes 3.1 (Ronquist and Huelsenbeck, 2003), implementing a partition by gene and assigning to each partition its best-fit model "family". All parameters were unlinked between partitions, except topology and branch lengths. In an attempt to explore tree and parameter spaces more efficiently, analyses

consisted of six independent runs of 10×10^6 generations and 10 Markov chains (temperature set to 0.20), with trees sampled every 1000 generations. Stationarity was assessed by plotting -ln L per generation in Tracer 1.3 (Rambaut and Drummond, 2004), and plotting posterior probabilities of clades as a function of generation number using AWTY (Wilgenbusch et al., 2004). Comparison of performance of multiple runs allowed selection of only those runs that converged to the highest likelihood values and reflected stability in the posterior probabilities of clades. All six runs fulfilled these conditions and reached stationarity after 1×10^6 generations. Of the 10000 trees resulting per run, the first 2000 were discarded as "burn in." Then, the remaining 48000 trees (8000 trees × 6 runs) were combined to calculate the posterior probabilities in a 50% majority-rule consensus tree.

Hypothesis Testing

Once the "best" phylogenetic hypothesis was recovered, the topology obtained was tested against previous hypotheses of relationships among species. Statistical comparisons were conducted via two different approaches, the Shimodaira-Hasegawa test of topology (SH test; Shimodaira and Hasegawa, 1999) and the likelihood-ratio test of monophyly (Huelsenbeck et al., 1996a), via parametric bootstrapping.

SH is a non-parametric test, which assesses whether the "best" ML tree and a set of trees contingent on one or more a priori hypotheses are equally good explanations of the data (Shimodaira and Hasegawa, 1999; Goldman et al., 2000). The best ML tree was compared with a set of trees that included the ML trees under the null hypotheses, and trees under other possible realizations of the null; in doing this, I avoided breaking up

monophyletic groups that were compatible with both the null hypotheses and the ML tree (Buckey et al., 2001). Trees were compared using the SH test as implemented in PAUP, running 1000 bootstrap pseudoreplicates and RELL optimization.

The parametric bootstrap test of monophyly (Huelsenbeck et al., 1996a) compares the likelihood between the best ML topology (T_1) and that showing the monophyly of the group of interests (T_0). Significance of likelihood difference (δ) is assessed by comparing observed differences with a null distribution obtained by means of Monte Carlo simulation (Efron 1985, Felsenstein 1988, Goldman et al. 2000; see also Huelsenbeck et al. 1996b). In this procedure, replicated datasets are simulated under the model and model parameter values drawn from the original data optimized over the null topology. Then, two different ML searches are conducted for each simulated dataset to estimate the likelihood of the best general topology and the best realization of the null topology. Calculation of likelihood differences over the simulated datasets provides the null distribution for assessing whether δ deviates from random expectations. A total of 250 simulated matrices were obtained using Batch Architect (Maddison and Maddison, 2004a) in Mesquite 1.05 (Maddison and Maddison, 2004b). Given that all species were represented only by the mitochondrial genes, the SH and the parametric bootstrap test were performed on a combined mitochondrial dataset; all ML searches were performed in GARLI.

Ancestral Area Reconstruction

To explore implications of the phylogeny for the historical biogeography of the *Cyanolyca* jays, I mapped the distribution of each species onto the final tree using parsimony and likelihood reconstructions. Advantages of using likelihood-based methods include incorporation of branch length information on the reconstruction, as well as estimation of relative probabilities of all possible ancestral states at every node on a tree, allowing quantification of uncertainty over otherwise (i. e., parsimony) unequivocal ancestral states (Schluter et al., 1997). Probabilities are determined by the model of evolution used, the distribution of the character states in terminal taxa, the rate of evolution of the character, and the lengths of internal branches on the tree (Cunningham et al., 1998).

A disadvantage of a likelihood-based method is that ML estimates of ancestral states can be wrong if the model is unrealistic (Crisp and Cook, 2005); also, inaccurate estimates may be generated whenever the number of states is relatively high compared to number of terminals (Schluter et al., 1997; Pagel, 1999). Therefore, to be able to compare parsimony and ML approaches, I used a simple dataset in which species were assigned to three main areas as character states: (1) Mesoamerica (= Mexico to Panama; *C. nana*, *C. mirabilis*, and *C. cucullata, C. pumilo*, *C. argentigula*); (2) Northern Andes (= from Venezuela south to the Huancabamba Deflexion in Peru; *C. pulchra*, *C. armillata*, and *C. turcosa*); and (4) Central Andes (= from the Huancabamba Deflexion south to Bolivia; *C. viridicyanus*). This simple scheme attempted to capture the most important biogeographic patterns in the genus. The sister relationship of *Cyanolyca* jays with all other New World Jay species is strongly supported by morphological and molecular data

(Zusi, 1987; Espinosa de los Monteros and Cracraft, 1997; Saunders and Edwards, 2000; Bonaccorso and Peterson, 2007). Also, although relationships among the remaining NWJ species are not completely resolved, broad geographic distribution of taxa in this group (from northern North America to southern Bolivia) suggests that the ancestor of the NWJ was likely broadly distributed across the Americas. As such, biogeographic optimizations were conducted based on ingroup taxa only. Also, given that sequences of *C. nana* were unavailable for the nuclear genes, optimizations were completed over the mitochondrial dataset only, to reduce potential effects of distorted branch lengths.

Parsimony and ML reconstructions were performed in Mesquite 1.05 (squared parsimony option; Maddison and Maddison, 2004) by optimizing area-characters onto the molecular tree. Maximum likelihood reconstructions were based on the Mk1 model, a *k*-state generalization of the Jukes-Cantor model, corresponding to Lewis's (2001) Mk model. The single parameter is the rate of change; thus, any particular change is equally probable. In Mesquite (Maddison and Maddison, 2004), the rate of change is estimated automatically via optimization of the character onto the tree, and relative marginal probabilities are obtained for each node.

RESULTS

Sequence Attributes

The ND2 and cyt*b* sequence fragments, including those for *Cyanolyca nana,* aligned easily with other mitochondrial genomes and translated into amino acids with no reading frameshifts. Also, they showed the typical substitution pattern of protein-coding genes, with most nucleotide substitutions being transitions at third-codon positions.

Therefore, I am confident that they represent genuine mitochondrial, coding sequences and not nuclear pseudogenes (Zhang and Hewitt, 1996; Sorenson and Quinn, 1998; Bensasson et al., 2001). CR and nuclear introns showed indels of variable size, but were aligned easily after minor adjustments.

Variable and parsimony-infomative sites within gene fragments were distributed as follows: 459/379 out of 1014 for ND2; 338/269 out of 653 for CR; 398/278 out of 1023 for cyt*b*; 111/45 out of 871 for βfb7; and 96/28 out of 607 for AK5. Primer pairs used for amplifying *C. nana* produced three overlapping fragments of ND2 (397 bp) and CR (608 bp; 648 aligned positions), and two non-overlapping fragments for cyt*b* (218 bp and 183 bp). According to the AIC, MODELTEST selected the GTR + Γ + I model for cyt*b*, and CR; the TrN + Γ + I model for ND2; the GTR + Γ model for βfb7; and the HKY + I model for AK5; parameter values estimated based on ML trees are listed in Table 2. Nucleotide composition bias across lineages was non-significant for all datasets ($P = 0.999$); evolutionary rate heterogeneity was detected for all mitochondrial genes ($P < 0.001$), but not for the nuclear genes ($P > 0.05$).

Phylogenetic Congruence and Combined Analyses

Phylogenetic trees based on ML analysis of individual mitochondrial and nuclear trees were largely congruent (Fig. 2.2). *Cyanolyca* consists of two major clades, each with several geographically defined subclades. According to the mitochondrial genes, the first clade consists of Mesoamerican "dwarf" jays, with *C. pumilo* and *C. argentigula* are sisters and reciprocally monophyletic with respect to *C. nana* + *C. mirabilis* (although this relationship was poorly supported in some cases). The second clade is divided into

two main groups—one composed of *C. cucullata* + *C. pulchra*, and the other containing the "core" South American species, with *C. armillata* as sister of *C. turcosa* + *C. viridicyanus*. The ND2 tree showed the best support for all relationships with exception of the node uniting *C. nana* and *C. mirabilis*. Bootstrap analyses based on CR and cyt*b* were equivocal in their support of the sister relationship between the core South American jays and *C. cucullata* + *C. pulchra*, and cyt*b* did not demonstrate the monophyly of *Cyanolyca*. In general, MP reconstructions of mitochondrial genes were highly homoplastic, having low consistency and rescaled consistency indices (ND2: CI = 0.46, RC = 0.35; CR: CI = 0.52, RC = 0.41; cyt*b*: CI = 0.48, RC = 0.21); still, in most cases, they recovered the same major groupings as the ML trees (see MP bootstrap values in Fig. 2.2).

Parsimony and ML analyses of the nuclear introns produced results that were in general agreement with those of the mitochondrial genes. AK5 supported the monophyly of *Cyanolyca* and the sister relationships between *C. pulchra* and *C. cucullata*, *C. viridicyanus* and *C. turcosa* (ML and MP), and *C. argentigula* and *C. pumilo* (ML only). βfb7 reconstructed most of the same relationships as the mitochondrial genes, but failed to recover the sister relationship between *C. pulchra* and *C. cucullata*, and the monoplyly of *C. viridicyanus*; also, a sister relationship between *C. mirabilis* and *C. pumilo* was recovered with high bootstrap value. This latter relationship seemed odd considering that most analyses of the mitochondrial genes coincided in placing it as sister of *C. mirabilis*, and recovered *C. pumilo* + *C. argentigula* with high support. Even though mitochondrial genes are part of the same locus and, as such, can not be considered as independent assessments of relationships, it is possible that these discrepancies are caused by βfb7

not having achieved complete lineage sorting (Moore, 1995). Given that conflict among datasets was limited to a single terminal branch and species paraphyly, I was confident in combining all genes in further analyses.

Consistent with the general agreement among genes and the relatively high support at most nodes, the combined analysis recovered most of the same hypothesis of relationships as the mitochondrial genes, as well as highly robust trees under all optimization criteria. The MP analysis generated an immense number of most parsimonious trees (520,052 trees, 3695 steps, CI = 0.5009, RC = 3645), but all disagreements were caused by alternate arrangements among outgroup sequences, or within the same species or subspecies.

Ten independent ML-GARLI analyses produced highly consistent results, and the final likelihood values for the ML-GARLI and 50% majority rule consensus tree produced by Bayesian analyses were closely similar, considering that GARLI produced *the* ML tree (joint estimation), whereas Bayesian trees are produced by integrating across the parameter space (marginal estimation). This similarity may be explained by the fact that when there are relatively few parameters and a large amount of data, the ML estimate is a good predictor of the integral over the whole likelihood surface (Holder and Lewis, 2003). Figure 2.3 shows the 50% majority rule consensus tree drawn from the Bayesian analysis, indicating the level of node support recovered from Bayesian, ML, and MP analyses.

Hypothesis Testing

The final ML tree topology did not include some of the relationships proposed by Goodwin (1976) and others (Hellmayr, 1934; Fjeldså and Krabbe, 1990), being these the sister relationships between (1) *Cyanolyca argentigula* and *C. mirabilis* and (2) *C. viridicyanus* and *C. armillata*. Null hypotheses trees were generated under both independent and joint scenarios, and a set of compatible trees was built by incorporating relationships supported by Goodwin (1976) and the ML tree: i.e., monophyly of *Cyanolyca*, monophyly of the dwarf jays, and ((*C. pulchra* + *C. cucullata*) South American jays). All possible combinations of these constraints produced 27 different topologies that were compared with the "best" tree, using the SH test.

The SH test rejected topologies containing *Cyanolyca argentigula* + *C. mirabilis* ($P < 0.01$) as equally good explanations of the data. However, it was unable to reject *C. viridicyanus* + *C. armillata*. Given that the SH test is considered a more conservative test than parametric bootstrapping (e.g., Buckey 2000), further effort focused on exploring whether the *C. viridicyanus* + *C. turcosa* topology (recovered herein) and the *C. viridicyanus* + *C. armillata* topology (Goodwin 1976), produced statistically different likelihood values.

Differences in statistical power between the SH (nonparametric) and parametric bootstrapping test are expected because knowledge on the underlying distribution is available in the second, but not to the first approach (Goldman 2000). Although the cost of this power is an increased reliance on the model assumed, parametric bootstrapping has proved robust to deviations from the assumed model, as long as a reasonably

complex model encompasses the major features of the distribution of the data (Hillis et al., 1996, Huelsenbeck and Crandall, 1997; Goldman et al., 2000).

Figure 2.4 shows the distribution of δ drawn from the simulated datasets as well as the empiric δ value. Given that (1) the *Cyanolyca viridicyanus* + *C. turcosa* topology had a statistically better likelihood than the *C. viridicyanus* + *C. armillata* topology ($P < 0.001$, $\delta = 65.215$) and (2) the GTR + I + Γ model implemented herein seemed to fit well the combined mitochondrial dataset (i.e., MODELTEST), I assumed that the *Cyanolyca viridicyanus* + *C. turcosa* was the true topology, and that this result was unlikely to be an artifact of model misspecification.

Geographic Structure within Cyanolyca *Species*

On finer scales, individuals in polytypic species segregated clearly into geographic groups (Fig. 2.3). Within *Cyanolyca cucullata*, samples from the Sierra Madre Oriental of Mexico (*C. c. mitrata*) are distinct from those of the southern Central American highlands of Costa Rica (*C. c. cucullata*). In South America, individuals of *C. armillata* from the Andes of Venezuela (*C. a. meridana*) are distinct from those of the eastern Andes of Ecuador (*C. a. quindiuna*), whereas samples corresponding to *C. viridicyanus* split in two groups (Fig. 2.5), one corresponding to populations of the northern (*C. v. jolyaea*) and southern portions of the eastern Andes of Peru (*C. v. cyanolaema*), and another from western Bolivia (*C. v. viridicyanus*). Interestingly, the tree shows an additional split between individuals of *C. turcosa* from northern Ecuador versus those from southern Ecuador and northern Peru, a distinction that has not been noted on morphological grounds.

These separations translated into relatively high percent sequence divergence in ND2 (ML-corrected pair-wise distances; Appendix II); although both, uncorrected-p and ML-corrected distances showed similar values between closely related species, I present corrected distances because they are more appropriate for comparing distantly related taxa. Sequence divergence between geographically segregated groups can be summarized as follows: *C. turcosa* of northern Ecuador vs. *C. turcosa* of southern Ecuador and northern Peru, 1.3–1.5%; *C. armillata meridana* vs. *C. a. quindiuna*, 2.9–3.0%; *C. cucullata cucullata* vs. *C. c. mitrata*, 4.8–5.0%; and *C. viridicyanus jolyaea* vs. *C. v. cyanolaema* and *C.v. viridicyanus* 8.0–8.4%. Although not included in the analyses, unreversed CR indels of variable size provided information as per the segregation of some geographic groups (Fig. 2.2).

Nuclear genes also show differences among populations of *C. viridicyanus*. Individuals of *C. v. jolyaea* are united by five nucleotide synapomorphies (4 with CI = 1.000; 1 with CI = 0.3333) for βfb7, whereas individuals of *C. v. cyanolaema* and *C. v. viridicyanus* show an unreversed indel (2 bp) for the same gene, and a nucleotide synapomorphy (CI = 0.3333) in AK5. In contrast, nuclear divergence between individuals of *C. viridicyanus* and *C. turcosa* is in the form of only one nucleotide substitution (CI = 0.250 for AK5). Although nuclear sequence variation is observed among individuals within species of *C. cucullata*, *C. armillata*, and *C. turcosa*, no structure is apparent.

Ancestral Area Reconstruction

Maximum parsimony and ML optimizations produced similar results for reconstruction of ancestral areas for the *Cyanolyca* jays. Although both methods reconstruct virtually the same ancestral states, the ML reconstructions provide additional information regarding marginal probabilities associated with each character state. This probabilistic perspective is especially useful in determining ancestral areas for nodes where reconstructions are equivocal under MP optimization (Fig. 2.6). For example, according to the MP optimization, the ancestral area for *Cyanolyca* could be Mesoamerica or the northern Andes; however, the ML reconstruction indicates a higher probability for Mesoamerica as the ancestral area for the genus. According to both reconstructions, the origin of the dwarf jays is in Mesoamerica and the origin of the "core" South American species is the northern Andes, with *C. viridicyanus* representing a single dispersal to the Central Andes. Equivocal reconstructions were obtained at the node uniting *C. pulchra* + *C. cucullata* and the one uniting these to species to the core South American jays.

DISCUSSION

The inferred hypothesis of relationships for the *Cyanolyca* jays is highly robust, and is in general agreement with previous ideas (i.e., Goodwin, 1976). The robustness of the combined phylogenetic tree seems to be a product of having genes that (1) support nodes at different branching levels and (2) overall, are congruent with one another. Simultaneously, robustness may be related to the considerable number of characters involved in the combined analysis. It has been proposed that weak, but true signals may

be present in different datasets, but that the signal within any single dataset may be masked by noise (Barret et al., 1991). Thus, by increasing numbers of characters, phylogenetic signal is more likely to assert itself over stochastic error, resulting in a more accurate estimate of the true phylogeny (Huelsenbeck and Hillis, 1993; de Queiroz et al., 1995). Although these results are conditional with the data at hand, it seems that it would take a large amount of contrary evidence to recover alternate topological rearrangements among taxa.

Phylogeny

The monophyly of *Cyanolyca* has not been seriously questioned, particularly since Zusi's (1987) work. All New World jay species share a synapomorphy that involves modifications in the lower jaw and the quadrate (i.e., the "buttress complex"; Zusi, 1987); in all *Cyanolyca* examined by Zusi (1987) this complex shows a shared state. However, given that no previous morphological or molecular study of the phylogeny of NWJs (Zusi, 1987; Espinosa de los Monteros and Cracraft, 1997; Saunders and Edwards, 2000; Bonaccorso and Peterson, 2007) considered more than three of the nine species in the genus, having complete sampling provides a solid basis for inferences about speciation and evolutionary processes in the group.

Common ancestry of the dwarf jays is consistent with their smaller size and allopatric distributions across the highlands of Mesoamerica. Regardless of the morphological similarities between *Cyanolyca argentigula* and *C. mirabilis* (Goodwin, 1976), the sister relationships reconstructed herein between *C. argentigula* and *C. pumilo*, and *C. mirabilis* and *C. nana* are more congruent with their geographic distributions.

Cyanolyca pumilo inhabits montane forests between the Isthmus of Tehuantepec and the lowlands of Lake Nicaragua (hereafter "northern Central America"), and is replaced by *C. argentigula* in the mountains of Costa Rica and western Panama. Molecularly and morphologically divergent populations in these two regions have been documented for other bird taxa (*Lampornis* hummingbirds, García-Moreno et al., 2006; *Aulacorhynchus* toucanets, Puebla-Olivares et al. *in press*; *Chlorospingus* bush-tanagers, Bonaccorso et al. unpub.). As a whole, the dwarf jays form two reciprocally monophyletic groups—*C. pumilo* + *C. argentigula* and *C. nana* + *C. mirabilis*—distributed east and west of the Isthmus of Tehuantepec, respectively. Similar patterns of differentiation have been recovered in previous studies of other montane forest vertebrates (e.g., Pérez-Emán, 2005; García-Moreno et al., 2004; García-Moreno et al., 2006; Cadena et al., *in press*; Sullivan et al., 2000).

Morphological similarity between *Cyanolyca argentigula* and *C. mirabilis* is an example of independent acquisition (or conservatism) of plumage character combinations (i.e., leapfrog pattern: Remsen, 1984; Maijer and Fjeldså, 1997; Johnson, 2000). Interestingly, examples of variation in pairs of species and subspecies distributed along the Neotropics (e.g., *C. argentigula* and *C. mirabilis*, this study; *Aulacorhynchus prasinus caeruleogularis* vs. *A. cyanolaemus*, Puebla-Olivares et al. *in press*.) indicate that this phenomenon may be more geographically and taxonomically widespread than originally appreciated. Lack of genomic information linking plumage to genomic information is problematic in interpreting these patterns. Still, phylogenetic evidence points toward the hypothesis that much of the phenotypic differentiation in plumage characters might be produced by stochastic factors and evolutionary constraints (e.g.,

limited character states: Wake, 1991; high character correlation: Omland and Lanyon, 2000), rather than more predictable phylogenetic and environmentally induced factors (Remsen, 1984).

With regard to the remaining species, the close relationship between *C. pulchra* and *C. cucullata* is consistent with their distributions separated by the Panamanian and Colombian lowlands (Fig. 2.1), as well as their inhabiting lower elevational ranges than all other species in the genus (Stiles and Skutch, 1989; Ridgely and Tudor, 1994; Howell and Webb, 2004). Geographic separation of this clade from the main northern Andean lineage might have been caused by the Río Cauca Valley in western Colombia, whereas sympatry of *C. pulchra* with *C. turcosa* in western Ecuador might represent a post-speciation range expansion of the second species.

The core South American taxa form a monophyletic group, with *Cyanolyca armillata* sister to *C. turcosa* + *C. viridicyanus* (*contra* Hellmayr, 1934; Meyer de Schauensee, 1966; Goodwin, 1976; Fjeldså and Krabbe, 1990; parametric bootstrap with $P < 0.001$). This arrangement is intriguing, given that no clear geographic break is evident between these two clades (i.e., *C. armillata* vs. *C. turcosa* + *C. viridicyanus*). In fact, the eastern cordillera of Ecuador and the central Andes of Colombia—where *C. armillata* and *C. turcosa* overlap—are considered part of the same geomorphological unit (Simpson, 1975). Given that *C. turcosa* is closely related to *C. viridicyanus*, sympatry between *C. armillata* and *C. turcosa* probably resulted from secondary contact. Unfortunately, details of the original barrier separating these two clades, as well as the potential causes of cladogenesis, remain obscure.

On the other hand, the sister relationship between *C. turcosa* (Andes of Ecuador) and *C. viridicyanus* (Andes of Bolivia and Peru), makes biogeographic sense, because their ranges are adjacent to and separated by the Río Marañon Valley in northern Peru. This geographic break is considered a classic example of a barrier for allopatric speciation in the Andes (e.g., Vuilleumier, 1969; Parker et al., 1985). Distribution of sister species on either side of this valley has been documented also in *Leptopogon* flycatchers (Bates and Zink, 1994).

Geographic structure and species limits

Extended sampling of widely distributed species provided a preliminary assessment of molecular differentiation of populations, as well as identification of potential independent lineages within species. Among the least expected results was the segregation of individuals of *Cyanolyca turcosa* into two genetically divergent groups. *Cyanolyca turcosa* is distributed mainly along the two Andean cordilleras of Ecuador separated by a dry valley; however, the two cordilleras meet in southern Ecuador, as do some bird populations distributed along them (Ridgely and Greenfield, 2001). The pattern observed may indicate an abrupt disruption of gene flow between populations of the main Andes of Ecuador, and those at the southern tip of the Andes of Ecuador and northern Peru. Such a disruption, albeit coinciding with a geographic break in southern Ecuador around the Río Zamora Valley (Fig. 2.5), has not been documented in other groups of highland birds. Further and denser sampling at the population level is needed to assess the effectiveness of this geographic break as a barrier to gene flow, as well as historical signals of population expansion and fragmentation.

Samples of *Cyanolyca armillata* from the Andes of Venezuela (*C. a. meridana*) and the eastern Andes of Ecuador (*C. a. quindiuna*, also in the central Andes of Colombia), segregate into two clearly differentiated lineages. Unfortunately, individuals from the geographically intermediate populations along the eastern Andes of Colombia (*C. a. armillata*) were not available for inclusion in this study. However, a phylogeographic analysis of Emerald Toucanets revealed a close relationship between a sample from the eastern Andes of Colombia and those from the Andes of Venezuela (Puebla-Olivares et al., *in press*), compared with those inhabiting the central Andes of Colombia and the eastern Andes of Ecuador. Thus, it is likely that the observed differences are not just reflecting isolation by distance, but, instead, a signature of restricted gene flow between populations east and north of the Río Magdalena Valley in central Colombia (see also Cadena et al., *in press*). Still, the relatively weak genetic differentiation between individuals from the extreme populations couples with slight, clinal variation in coloration and size. Analysis of samples from the eastern and central Andes of Colombia, as well as careful documentation of morphological variation is crucial in assessing potential geographic structure in this species.

Divergence between samples of *Cyanolyca cucullata* from the Sierra Madre Oriental of Mexico (*C. c. mitrata*) and samples from the mountains of Costa Rica and western Panama (*C. c. cucullata*), coincides with a discrete morphological difference: *C. c. mitrata* has a well-defined white line on the crown that extends to the loral region, whereas in *C. c. cucullata*, the line is lacking, but a light, whitish shadow on the crown is present. Unfortunately, samples representing intermediate populations from northern Central America (*C. c. guatemalae* and *C. c. hondurensis* of Pitelka 1951b) were not

available for inclusion in this study. Nonetheless, in both measurements and plumage, these forms are almost identical to *C. c. mitrata*.

Morphological differences and nucleotide-based synapomorphies separating allopatric populations of *Cyanolyca cucullata* south and north of the lowlands of Lake Nicaragua, indicate that they may represent independent evolutionary lineages (sensu Wiley, 1978). From the point of view of the Phylogenetic Species Concept, study of museum skins shows that the criterion of diagnosability (Cracraft, 1983) is met for populations on both sides of the geographic break. However, further population-level studies, including more samples across the intervening areas of northern Central America, are needed to confirm the reciprocal monophyly of *C. c. cucullata* and *C. c. mitrata*.

The most dramatic molecular differences were observed between populations of *Cyanolyca viridicyanus* north (*C. v. jolyaea*) and south (*C. v. cyanolaema* and *C. v. viridicyanus*) of the Río Apurimac Valley in southeastern Peru. This valley is considered one of the most important biogeographic boundaries in the humid Andes (Vuilleumier, 1969; Haffer, 1974; Remsen and Brumfield, 1996), defining the limits of several avian taxa (Ridgely and Tudor, 1994).

Closer relationship between the two forms south the Río Apurimac Valley (*Cyanolyca viridicyanus viridicyanus* and *C. v. cyanolaema*) is congruent with revision of museum skins, as well as prior assessments of plumage similarity (Bond, 1956). Interestingly, one of the plumage characters that differs in this complex, seems to be homologous to the one that distinguishes populations of *C. cucullata.* Specimens from north of the Río Apurimac have a white line at the level of the crown, whereas those from

the south show no line, but have a light, whitish shadow; in addition, *C. v. jolyaea* is divergent in color (overall darker purplish blue instead of greenish blue) and has a narrower white band on the throat. Regardless of the relatively limited number of samples included in this study, it is clear that, at all levels of divergence, morphological, mitochondrial, and nuclear synapomophies support the reciprocal monophyly (McKitrick and Zink, 1988; Zink and McKitrick, 1995; Edwards et al., 2005) and diagnosability (Cracraft, 1983), of two historical and evolutionarily independent (Wiley, 1978) entities: *C. jolyaea* and *C. viridicyanus*.

Ancestral Area Reconstruction

Optimization of geographic areas on the tree provided simple, analytical means of exploring the directionality of range shifts in *Cyanolyca*. The whole assemblage probably originated in Mesoamerica, and so did the dwarf jays. Uncertainty in the reconstructions of the nodes uniting *C.* pulchra and *C. cucullata* and these species to the core South American species poses two possible scenarios. The first one involves an ancestor of (*C. pulchra* + *C. cucullata*) + core South American species in Mesoamerica, implying two independent dispersals to the Northern Andes in the ancestors of *C. pulchra* and the "core" South American species. The second scenario involves that the ancestor of (*C. pulchra* + *C. cucullata*) + core South American species lived in the Northern Andes, which indicates one dispersal to the Northern Andes (from a Mesoamerican ancestor for *Cyanolyca*), and then a dispersal of *C. cucullata* into Mesoamerica. As such, either scenario implies two independent dispersals across the Isthmus of Panama.

Presence of *Cyanolyca viridicyanus* in the Andes of Peru represents a range expansion from the Northern Andes to the Central Andes. Distributional patterns of bird species on both sides of the Río Marañon Valley have been interpreted as product of different scenarios of allopatric speciation: (1) range expansion of an ancestral form before the valley was in place, and posterior vicariance and speciation caused by the barrier (Bates and Zink, 1994), or (2) dispersal of a small founder population across the valley at a time when its effectiveness as a barrier was weaker than at present, with posterior allopatric speciation (Vuilleumier, 1969; Johnson, 2000). Unfortunately, both scenarios would leave the same signature in the biogeographic reconstruction.

Speciation in the Cyanolyca *Jays*

Phylogenetic analysis of *Cyanolyca* species provides evidence for discussing modes of speciation. Observed patterns of relationship do not support models involving recent speciation along altitudinal gradients, since species that replace each other in altitude are not each other's closest relatives. On the other hand, biogeographic reconstructions emphasize, once again, the importance of periodical establishment and disruption of gene flow across the Panamanian Land Bridge in shaping the avifaunas of both South American and Mesoamerican montane forests (Hackett, 1995; Pérez-Emán, 2005; Cadena et al., *in press*, Puebla-Olivares et al., *in press*).

Also, the observed differentiation provides support for hypothesis based on the effect of the linearity of the Andes in limiting contact among parapatric forms (especially in *Cyanolyca viridicyanus*) and the disruptive effect of geographic barriers in promoting speciation, not only in the Andes, but along the full range of the species. With only one

exception (i.e., the clade formed by *C. armillata* and *C. turcosa* + *C. viridicyanus*), diverging lineages (sister species, as well as sister clades) are distributed on either side of potentially effective barriers to gene flow (Fig. 2.1). In the case of *Cyanolyca armillata* vs. *C. turcosa* + *C. viridicyanus*, original species distributions are obscured by likely subsequent range expansion of *C. turcosa* in both the western and the eastern Andes of Ecuador, and by the fact that no current barriers separate these two groups.

Although allopatry is widely accepted as the most important cause of speciation among bird species (Mayr, 1942, 1963; Chesser and Zink, 1994; Barraclough et al., 1998), empirical evidence is scarce regarding Andean montane forest birds (but see Bates and Zink, 1994; Pérez-Emán, 2005 for specific examples). In most studies, adjacent, allopatric species and populations have turned to be more genetically similar than originally predicted (e.g., *Cranioleuca* Spinetails: García-Moreno et al., 1999b; *Buarremon bruneinucha* complex: Cadena et al., *in press*; *Aulacorhynchus prasinus* complex: Puebla-Olivares et al., *in press*; and other *Aulacorhynchus* toucanets: Bonaccorso, unpub.). Clearly, levels of divergence between allopatric species and populations are expected to vary according to several factors (e.g., evolutionary rates, time since isolation, dispersal abilities, and demography of populations [Mindell et al., 1990; Barraclough et al., 1998, 2001; Knowles and Maddison, 2002]). From the high levels of differentiation observed in *Cyanolyca* (Appendix II), it is likely that evolutionary rates are relatively high, or geographic isolation occurred before (or has been more effective) than in other lineages studied to date. Careful study of populations across *Cyanolyca* offers considerable potential for comparative studies linking historical factors and population parameters, with molecular and morphological evolution.

Finally, testing hypothesis of vicariance versus dispersal or the effects of Pleistocene glaciations, is not trivial, because it involves placing the phylogeny in an absolute-time framework, and associating divergence dates with geological and climatic events. Unfortunately, fossils available are not useful because they are too young (i.e., Late Pleistocene: Brodkorb, 1957; Weigel, 1967; Holman, 1959), ambiguous as per their identification as New World jays, or dated imprecisely (e.g., Late Miocene; Brodkorb, 1972). Additionally, given that widespread evolutionary rate heterogeneity has been detected for several lineages in the group (i.e., NWJs as a whole: Espinoza de los Monteros and Cracraft, 1997; Bonaccorso and Peterson, 2007; *Aphelocoma* jays: Peterson, 1992; J. McCormack, pers. comm.; and *Cyanolyca* jays: this study), using standard estimates of evolutionary rates drawn from other avian lineages (e.g., Fleischer et al., 1998; Arbogast et al., 2006), would be both arbitrary and misleading. On the other hand, considering the wide spectrum of sequence divergence recorded among different populations of *Cyanolyca*, development of extensive population-level studies for estimating coalescent-based divergence times seems promising.

In summary, study of *Cyanolyca* jays provided a unique perspective on phenomena responsible for the unparalleled biodiversity of the Neotropical mountain avifaunas. It also pointed out the value of analyzing both inter and intraspecific divergence in understanding distributional patterns and directionality of range expansions, as well as discovering the potential for answering broader questions in evolutionary biology. Future, collaborative efforts, directed towards assembling complete sampling for other taxa across the Neotropical mountains, especially in the Andes, are critical in assessing the generality of patterns found herein.

CHAPTER 3

MOLECULAR SYSTEMATICS AND EVOLUTION OF THE *CYANOCORAX* JAYS

Species in the genus *Cyanocorax* and the allied genera *Psilorhinus* and *Calocitta* have been studied extensively in terms of their complex social behaviors (e.g., Crossin, 1967; Hardy 1974; Raitt and Hardy, 1976, 1979; Langen, 1996a,b; Williams and Hale, 2006), delayed soft-part color development (Hardy, 1973; Peterson, 1991), vocal repertoires (Hardy, 1961, 1979), and habitat preferences (e.g., Hardy, 1969; Raitt and Hardy, 1979; Amaral and Macedo, 2006). However, the lack of a robust hypothesis of relationships, as well as the paucity of detailed natural history studies for several South American species, has precluded rigorous interpretation of these characteristics in an evolutionary and quantitative framework.

In addition to implications for the evolution of the group, the resolution of relationships among *Cyanocorax* species and allied genera poses interesting systematic challenges. Owing to the diverse combinations of plumage coloration, size, and morphology found in *Cyanocorax* species, the genus has been described as "a pigeon-hole for a heterogeneous assemblage of jays" (Moore, 1935), apropos to the widely held perception that no characters support a coherent, natural group (e.g., Amadon, 1944; Goodwin, 1976). As a result, *Cyanocorax* has been subjected to several taxonomic overhauls involving at one time or another large-scale splitting, lumping, and reallocation of taxa. I review these studies below.

Taxonomic History

Hellmayr (1934) presented the first comprehensive treatment of *Cyanocorax* and allied genera (but see Ridgway [1904] for treatment of North American and Mesoamerican species). He divided the current *Cyanocorax* into four genera: (1) *Xanthoura,* consisting solely of *X. yncas*, the only green-colored jay in the group (southern Texas to northern Bolivia); (2) *Cissilopha,* consisting of four Mesoamerican jays (*C. beechei, C. sanblasiana, C. yucatanica,* and *C. melanocyanea*) characterized by black heads and blue body plumage; (3) *Uroleuca,* represented by *U. cristatellus,* a distinctive species of central South America; and (4) *Cyanocorax,* containing one Mesoamerican–South American and seven South American species, including *C. affinis*, *C. caeruleus*, *C. violaceus*, *C. heilprini*, *C. cayanus*, *C. chrysops* (including *C. cyanopogon*), *C. cyanomelas*, and *C. mystacalis*. Among other New World jay (NWJs) taxa, he recognized two species in the (currently) monotypic genus *Psilorhinus* (*P. morio* and *P. mexicanus*), and one species of *Calocitta,* including the current *C. formosa* and *C. colliei*.

Based on subjective analysis of general morphology and plumage patterns, Amadon (1944) proposed a classification in which he included *Xanthoura* and *Uroleuca* in *Cyanocorax,* and placed *Cissilopha* into *Cyanocitta*, along with members of the current genera *Cyanocitta, Cyanolyca,* and *Aphelocoma.* Also, Amadon (1944) recognized *Psilorhinus* and *Calocitta* as valid genera, and divided *Cyanocorax* into four "sections": (1) the "*Coronideus*" group, containing *C. caeruleus, C. cyanomelas,* and *C. violaceus;* (2) "*Uroleuca*"*;* (3) "*Xanthoura*"*;* and (4) a more restricted "*Cyanocorax*" that included all other taxa, in addition to the newly described *Cyanocorax dickeyi* (Moore, 1935).

Blake and Vaurie (1962) "resurrected" *Cissilopha,* and recognized *Cyanocorax* (including *Xanthoura* and *Uroleuca*), *Calocitta,* and *Psilorhinus.* Hardy (1969), in contrast, lumped genera into a broad *Cyanocorax* with five subgenera: *Uroleuca, Calocitta, Psilorhinus, Cissilopha,* and *Cyanocorax* (including *Xanthoura*), and recognized *C. cyanopogon* as full species separate from *C. chrysops.* (See also Meyer de Schauensee, 1966). Goodwin (1976) followed Blake and Vaurie (1962) in conferring generic status to *Cissilopha, Psilorhinus,* and *Calocitta.* Despite much work and complex taxonomic history, the taxonomy of the group is still in debate. Monroe and Sibley (1993) and Madge and Burn (1994) recognized *Cyanocorax* (including *Xanthoura, Uroleuca,* and *Cissilopha*), *Calocitta,* and *Psilorhinus,* while others (AOU 1983, 1998, Dickinson 2003) recognized only *Cyanocorax* (including *Xanthoura, Uroleuca, Cissilopha,* and *Psilorhinus*) and *Calocitta.* In conclusion, the taxonomy of the group has been unstable, reflecting the lack of a detailed, robust evolutionary framework on which to base a classification.

The origin of the Tufted Jay, Cyanocorax dickeyi

Within the *Cyanocorax* assemblage, morphologically similar species can have highly complex and discontinuous geographic distributions (Hardy 1961, Goodwin 1976). The most notorious example is the similarity between *Cyanocorax dickeyi* and *C. mystacalis*, which live ~4000 km apart. *Cyanocorax dickeyi* inhabits oak and mixed oak-pine forest in a minute area on the Pacific slope of the Sierra Madre Occidental of Mexico (Moore 1935, Crossin 1967), whereas *C. mystacalis* is endemic to the dry forests of southwestern Ecuador and northwestern Peru (Ridgely and Tudor 1989). Amadon

(1944) hypothesized that *C. dickeyi* is a relict of a more widely distributed *Cyanocorax* assemblage that was out-competed by other Mesoamerican jays (See also Moore, 1935). In contrast, others have proposed that a flock of *C. mystacalis* was storm-blown from the Pacific coast of South America (*in* Hardy, 1969) or that *C. dickeyi* was a descendant of *C. mystacalis* brought to Mexico by Native Americans (Haemig, 1979). Still, the origin of *C. dickeyi* and its restricted distribution remain unclear.

Phylogenetic Relationships

The first phylogenetic treatment of *Cyanocorax* jays (Hope 1989) was based on discrete and mensural osteological characters of 12 of the 16 species in *Cyanocorax,* in addition to *Psilorhinus morio* and *Calocitta formosa.* Major patterns recovered included *"Cissilopha"* as monophyletic (in 2 of 3 analyses), *C. mystacalis* and *C. dickeyi* as sister species, and *C. violaceus, C. caeruleus,* and *C. cyanomelas* as a monophyletic clade. *Psilorhinus* was placed variably as sister to *Calocitta* or closely related to *C. violaceus, C. caeruleus,* and *C. cyanomelas.* More generally, relationships among other species were unstable depending on outgroup selection; the only clade supported by a discrete, unreversed synapomorphy was *C. violaceus* + *C. caeruleus* + *C. cyanomelas* (i.e., Amadon's *"Coronideus"* group).

Molecular analyses of relationships among NWJs (Saunders and Edwards, 2000; Ericson et al., 2005; Bonaccorso and Peterson, 2007[Chapter 1]) have supported the monophyly of *Cyanocorax* + *Psilorhinus* + *Calocitta* (also in Espinosa de los Monteros and Cracraft, 1997, although *Psilorhinus* was not included). However, because the most complete study considered only six of the 16 species of *Cyanocorax*, the question of the

monophyly of the genus, as well as the relationships between species, is unresolved. Herein, I expand the taxonomic sampling to include all species of *Cyanocorax*, in the context of other NWJ genera, and a good sampling of mitochondrial and nuclear markers, to provide a phylogenetic framework for understanding the evolution, systematics, and biogeography of the group.

METHODS

Taxon and Gene Sampling

I analyzed samples from 54 individuals, including all species in *Cyanocorax* and the allied genera *Psilorhinus* and *Calocitta*, in addition to representatives of all other NWJ genera (*Cyanocitta, Aphelocoma, Gymnorhinus,* and *Cyanolyca*). For widely distributed species in *Cyanocorax* (i.e., *C. yncas, C. cayanus, C. cyanomelas, C. chrysops, C. violaceus, C. cyanopogon,* and *C. cristatellus*), samples from distributional extremes were included whenever possible. Tissue samples were obtained from museum collections in the U.S. and Mexico (Table 3.1), and a subset of the sequences analyzed was drawn from previous studies (Espinosa de los Monteros and Cracraft, 1997; Cicero and Johnson, 2001; Helm-Bychowski and Cracraft, 1993; Ericson et al., 2005; Bonaccorso and Peterson, 2007). Novel Transforming Growth Factor β-2 intron 5 sequences were generated for all outgroup taxa (GenBank Accession numbers to be added upon acceptance of the paper).

The mitochondrial genes NADH dehydrogenase subunit 2 (ND2) and cytochrome *b* (cyt*b*) were analyzed as fast-evolving makers that could provide resolution at the tips of the tree. To obtain independent estimates of relationships, as well as

information about deeper divergence, I incorporated three relatively fast-evolving nuclear loci: Adenylate Kinase intron 5 (AK5), β-Fibrinogen intron 7 (βfb7), and the Transforming Growth Factor β-2 intron 5 (TGFβ2.5). Information on genes sequenced, GenBank accession numbers, and voucher specimens and associated locality data, is summarized in Table 3.1.

Genomic DNA was extracted from frozen tissue with a modified salt precipitation protocol (M. Fujita, unpubl.), based on methods provided in the Puregene DNA purification kit (Gentra Systems). Because frozen tissue samples were not available for *Cyanocorax sanblasianus*, *C. beecheii*, *C. heilprini,* and *C. caeruleus,* DNA extracts for these species were obtained from museum skin samples (Table 3.1), courtesy of R. Fleischer (National Museum of Natural History and National Zoological Park) using established protocols (Fleischer et al., 2000, 2001). Amplification of DNA extracted from frozen tissue was carried out using the following primers: L5143 or L5216, and H6313 (Sorenson et al., 1999) for ND2; L14990 (Kocher et al., 1989) and H16065 (Tim Birt, unpubl.) for cyt*b*; FIB-B17U and FIB-B17L (Prychitko and Moore, 1997) for βFb7; AK5b+ and AK6c- (Shapiro and Dumbacher, 2001) for AK5; and TGFb2.5F and TGFb2.6R for TGFb2.5 (Sorenson et al., 2004).

Given that genomic DNA extracted from museum skins often is highly degraded, amplification from skin-extracted DNA was attempted only for mitochondrial genes. I designed several internal primers (Appendix III) which, in conjunction with published primers, allowed amplifying and sequencing of short DNA fragments (200–350 bp). Independent laboratory facilities were used to separate pre- and post-PCR manipulations,

and special care (e.g., multiple controls, use of fresh lab supplies) was taken to reduce risk of contamination.

I used a standard PCR protocol (94˚C/5 min; 35 cycles of 93˚C/1 min, 52˚C/1 min, 72˚C/2 min; and 72˚C/10 min) for the mitochondrial genes, and a touchdown protocol (94˚C/ 3 min; 5 cycles of 94˚C/30 sec, 60˚C/30 sec, 72˚C/40 sec; 5 cycles of 94˚C/30 sec, 56˚C/30 sec, 72˚C/40 sec; 35 cycles of 96˚C/30 sec, 52˚C/30 sec, 72˚C/40 sec, and 72˚C/10 min; R. Moyle pers. comm.) for nuclear genes. When multiple bands persisted, PCR products were purified using the QIAquick Gel Extraction Kit (Quiagen, Inc.). Single PCR products were visualized on agarose gel, and unincorporated primers and DNTPs removed with ExoSap (ExoSap-it, GE Health Care).

Cycle sequencing was completed with the corresponding PCR primers and BigDye Terminator 3.1 chemistry (Applied Biosystems) using a standard cycle sequencing profile (96˚C/3 min; 35 cycles of 96˚C/10 sec, 50˚C/15 sec, 60˚C/3 min; and 72˚C/7 min). Sequencing reaction products were purified with CleanSEQ magnetic beads (Agencourt) and resolved in an ABI Prism 3100 Genetic Analyzer (Applied Biosystems). Data from heavy and light strands were spliced together to arrive at a consensus sequence for each sample using Sequencher 4.1 (Gene Codes Corp. 2000).

Sequence Aligning and Phylogenetic Analyses

Sequences were aligned in CLUSTAL_X (Thompson et al., 1997), and adjusted by eye and translated into amino acid sequences (coding genes) in MacClade ver. 4.0 (Maddison and Maddison 2000). Best-fit models of nucleotide substitution were estimated in MODELTEST v. 3.7 (Posada and Crandall, 1998, 2001) the Akaike

Information Criterion (AIC) for each gene and for combined mitochondrial and mitochondrial + nuclear datasets. Best fit model "families" estimated from MODELTEST were used in further maximum likelihood analyses.

Phylogenetic trees were estimated using maximum parsimony (MP), maximum likelihood (ML), and Bayesian criteria. I compared individual gene trees and their clade support as a measure of congruence in phylogenetic signal (Bull et al., 1993; de Queiroz et al., 1995; Wiens, 1998). Strong incongruence was identified by the presence of conflicting nodes showing ≥ 70% non-parametric bootstrap support (Hillis and Bull, 1993) or ≥ 0.95 Bayesian posterior probabilities. I performed combined analyses to amplify potentially congruent signals, but questioned weakly supported nodes that reflected conflicting signals (Wiens, 1998). To explore potential sources of phylogenetic incongruence, I tested the data for stationarity in base frequencies (χ^2 test of homogeneity in PAUP*; Swofford, 2000) and rate homogeneity among lineages (likelihood ratio test; significance assessed by comparing $\Lambda = -2\log LR$, where LR is the difference between the –ln likelihood of the tree with and without enforcing a molecular clock, with a χ^2 distribution with $n - 2$ degrees of freedom, where n is the number of taxa).

Parsimony analyses were performed on each individual gene, a combined mitochondrial dataset (ND2 and cyt*b*), and all data (ND2, cyt*b*, AK5, βfb7, and TGFβ2.5). In parsimony analysis, gaps were treated as missing data, and heterozygous positions (nuclear genes) as polymorphisms, using IUPAC degeneracy codes. Trees were obtained through heuristic searches in PAUP* (Swofford, 2000) with 10,000 stepwise random additions (TBR branch-swapping). Clade support was estimated via heuristic

searches with 1000 bootstrap pseudoreplicates (Felsenstein, 1985), each pseudoreplicate consisting of 100 stepwise random additions.

Maximum likelihood trees were estimated using GARLI (Genetic Algorithm for Rapid Likelihood Inference, ver. 0.951; Zwickl, 2006) and RaxML-VI-HPC (Randomized Axelerated Maximum Likelihood for high performance computing; Stamatakis, 2006a, b), as they provide considerable advantages over PAUP* in terms of computational efficiency. Both programs are based on algorithms that estimate the tree topology, branch lengths, and model parameters that maximize the –ln likelihood, in a simultaneous approach.

GARLI is based on a genetic algorithm involves the evolution of a population of solutions termed "individuals," with each individual encoding a tree topology, a set of branch lengths, and a set of model parameters. Each individual is assigned a fitness based on its lnL score. Each generation, random mutations are applied to some of the components of the individuals, and their fitnesses recalculated. A subset of individuals is then chosen to be the parents of the individuals of the next generation in proportion to individual fitness values. This process is repeated for a number of iterations (discussed below), and the population of individuals evolves toward higher fitness solutions. The highest-fitness individual is automatically maintained in the population, ensuring that it is not lost by chance (http://www.bio.utexas.edu/faculty/antisense/garli/GARLIv0.95manual.pdf). Models of nucleotide substitution available include the General Time Reversible (GTR) model and its more common submodels, as well as less complex models. It also accounts for gamma-distributed rate heterogeneity (Γ) with a specified number of rate categories and

estimation of the proportion of invariable sites (I). Estimation of model parameter values may be optimized or fixed, and the implementation of the model is equivalent to that in PAUP*.

GARLI analyses for individual genes were conducted specifying the model "family" obtained by MODELTEST, but allowing the program to estimate parameter values from the data. In cases in which a gamma distribution was implemented, the number of rate categories was set to four. Individual solutions were selected after 10,000 generations with no significant improvement in likelihood, with the significant topological improvement level set at 0.01 (first condition for termination); then, the final solution was selected when the total improvement in likelihood score was lower than 0.05, compared to the last solution obtained (second condition for termination). All other GARLI settings involved on the genetic algorithm were left at default values as per recommendations of the developer (Zwickl, 2006). For each dataset, I ran 10 independent analyses starting with random trees, and selected the tree with the highest ML score. Final tree score and parameter estimates were obtained by optimizing tree and branch lengths in PAUP*. Bootstrap support was assessed via 100 and 1000 pseudoreplicates for the individual gene and the combined datasets, respectively; bootstrap searches ran under the same settings used for obtaining the best ML tree.

RaxML-VI-HPC (Stamatakis, 2006b) operates by building an initial maximum parsimony or random tree and performing standard subtree rearrangements by subsequently removing all possible subtreees from the current best tree and re-inserting them into neighboring branches up to a specified distance of nodes. Details involved in the hill-climbing algorithm, as well as in efficient storage and scoring of alternate

topologies, are described in Stamatakis et al. (2005) and Stamatakis (2006a). RaxML was used herein because it allows analyzing different data partitions in the same ML analysis; although searches are performed on a "general" model, model parameter values can be optimized to fit every independent partition.

I estimated ML trees on the combined mitochondrial and total evidence datasets using the GTR + Γ model for all data partitions. Using this model for partitions in which more complex (GTR + Γ + I) or simpler (HKY + Γ) models were indicated by MODELTEST was not expected to have strong repercussions for the final results because: (1) thanks to its computational efficiency, RaxML can account for 25 or more gamma-rate categories, an alternate means of accounting for the proportion of invariable sites; and (2) although less complex models might be indicated for some partitions and over-parameterization errors might arise, empirical evidence using RaxML suggests that GTR tends to yield slightly better likelihoods than simpler models (http://icwww.epfl.ch/~stamatak/index-Dateien/software/RAxML-Manual.2.2.3.pdf). Other settings included starting the heuristic search with a parsimony tree and using 0.1 log likelihood units as threshold for deciding among competing topologies.

Bayesian analyses of the combined datasets were performed in Mr Bayes 3.1 (Ronquist and Huelsenbeck, 2003), implementing partitions by gene and by codon position, assigning to each partition its best-fit model family of nucleotide substitution. All parameters were unlinked between partitions, except topology and branch lengths. To explore tree and parameter spaces more efficiently, analyses consisted of six independent runs of 5×10^6 generations and 10 Markov chains (instead of the default 4 chains, temperature set to 0.20), with trees sampled every 1000 generations. Stationarity was

assessed by plotting -lnL per generation in Tracer 1.3 (Rambaut and Drummond, 2004) and plotting posterior probabilities of clades as a function of generation number using AWTY (Wilgenbusch et al., 2004). Comparison of performance of multiple runs allowed selection of only those runs that converged to the highest likelihood values and reflected stability in the posterior probabilities of clades. All six runs fulfilled these conditions and reached stationarity after 500,000 generations. From the 5000 resulting trees per run, the first 1000 were discarded as "burn in." Then, a total of 240,000 trees (4000 trees × 6 runs) was combined to calculate posterior probabilities in a 50% majority-rule consensus tree.

Hypothesis Testing

Statistical comparison between the "preferred" hypothesis of relationships and a topology enforcing the monophyly of the genus *Cyanocorax* and the monophyly of *C. mystacalis* + *C. dickeyi* were conducted via the likelihood-ratio test of monophyly (Huelsenbeck et al., 1996a). This test compares the likelihood between the best ML topology (T_1) and that showing the monophyly of the group of interests (T_0). Significance of likelihood difference δ is assessed by comparing it to a null distribution obtained by means of parametric bootstrapping (Effron, 1985; Felsenstein 1988; also Huelsenbeck et al., 1996b). In short, replicated datasets are created by simulation under the model and model parameter values empirically estimated from the original data optimized over the null topology. Then, two different ML searches are conducted for each simulated dataset to estimate the likelihood of the best general topology and the best realization of the null topology. Calculation of likelihood difference over the simulated datasets provides the

null distribution for assessing whether δ deviates from random expectations. Given that two independent test were performed (monophyly of *Cyanocorax* and monophyly of *C. mystacalis* + *C. dickeyi*), significance of δ-values was assessed after the level was adjusted using a Bonferroni correction (Sokal and Rohlf 1995).

To avoid potential problems related to missing data (e.g., potential distortion of branch lengths), parametric bootstrapping was implemented over the mitochondrial dataset only. A total of 100 matrices was simulated for each test (i.e., monophyly of *Cyanocorax* and *C. mystacalis* + *C. dickeyi*) with Batch Architect (Maddison and Maddison, 2004a) in Mesquite 1.05 (Maddison and Maddison, 2004b). I used GARLI for conducting all ML searches (i.e., "preferred" ML trees and constraint trees under the real and the simulated datasets).

RESULTS

Sequence Attributes and Model Selection

Use of *Cyanocorax*-specific internal primers allowed amplification and sequencing of samples for which fresh tissues were not available; species and numbers of base-pairs sequenced are as follows: *C. heilprini* (1002 bp for ND2, 486 bp for cyt*b*), *C. beecheii* (675 bp for ND2, 411 bp for cyt*b*), *C. caeruleus* (960 bp for ND2, 486 bp for cyt*b*), and two individuals of *C. sanblasianus* (1002 bp for ND2, 411 bp for cyt*b* and 867 bp for ND2, 411 bp for cyt*b*, respectively). Without exception, all sequences aligned easily with outgroup sequences, translated into amino acid sequences without frameshifts or unexpected stop codons, and showed the typical substitution pattern of protein-coding genes (most substitutions being synonymous transitions at third-codon positions),

suggesting that they were genuine mitochondrial sequences and not nuclear pseudogenes (Sorenson and Quinn, 1998; Benssason et al., 2001). Among nuclear introns, apart from specific length variation observed in outgroup sequences, indels were limited to 1–16 base pairs, allowing unambiguous alignment in all cases.

Variable and parsimony-informative positions across genes were distributed as follows: 457/375 out of 1002 for ND2, 394/311 out of 999 for cyt*b*, 110/19 out of 603 for AK5, 116/34 out of 872 for βfb7, and 53/18 out of 548 for TGFβ2.5. According to the AIC, MODELTEST selected models nested within the GTR + I + G model "family" for all but one data partition (HKY + Γ was selected for cyt*b* second-codon positions), and for the individual mitochondrial genes and combined mitochondrial dataset. The best-fit model families for the nuclear genes were HKY + Γ for AK5, and GTR + Γ for βfb7 and TGFβ2.5. Model parameter values estimated via ML (GARLI) analysis of individual genes, are summarized in Table 3.2. Nucleotide composition bias across lineages, considering codon positions (mitochondrial genes) and complete gene fragments, was non-significant in all cases ($P > 0.05$). Evolutionary rate heterogeneity was detected only for cyt*b* ($P < 0.01$), but all other genes could not reject the assumptions of a clock-like mode of evolution.

Phylogenetic Analyses

Individual MP and ML (GARLI) analyses of ND2 and cyt*b* reveal general congruence in phylogenetic signal. Although specific details of relationships differ among genes and among optimization criteria, conflicting nodes are relatively weakly supported (bootstrap value <64%); thus, I combined both genes in a single dataset. Figure

3.1 shows the Bayesian tree, including posterior probabilities, and ML (RaxML) and MP bootstrap support. The topology of the MP 50% majority rule consensus tree (120 trees, 2276 steps, CI = 0.443, RC = 0.3541, not shown) is in general agreement with those of ML and BA, unless otherwise specified.

Under all optimization criteria, the assemblage of *Cyanocorax* + *Psilorhinus* + *Calocitta* is recovered as monophyletic, and species divided further into two main clades: Clade A, containing *Calocitta*, *Psilorhinus, Cyanocorax violaceus*, *C. caeruleus*, *C. cristatellus*, and *C. cyanomelas*; and Clade B, including the remaining species of *Cyanocorax*. Within Clade A, *Calocitta* is sister to a group containing all other species; however, nodal support for this relationship is inconsistent between model based analyses (0.99 Bayesian pp, 66% ML bootstrap support, Fig. 3.1) and is not recovered by the MP 50% majority rule consensus tree. All other species in Clade A are represented by three groups: *Psilorhinus morio* + *C. caeruleus*, *C. violaceus* from Ecuador, and *C. violaceus* from Peru + (*C. cristatellus* + *C. cyanomelas*) ("Group I"), with this latter arrangement rendering *C. violaceus* as paraphyletic.

Within Clade A, however, different (relatively weakly supported) topologies are obtained from individual-gene analyses, with the positions of *Psilorhinus, Cyanocorax caeruleus,* and *C. violaceus* from Ecuador being unstable across analyses and datasets. For example, the ND2 MP tree shows the same relationships as the mitochondrial tree (Fig. 3.1), but the ND2 ML tree reconstructs *Psilorhinus* + *C. caeruleus* as sister to *C. violaceus* from Ecuador (54% boostrap support); the cyt*b* MP tree shows *C. violaceus* from Ecuador as sister to all other species in Clade A (64% boostrap support), and the cyt*b* ML tree recovers *C. caeruleus* at the base of Clade A and *Psilorhinus* + *Calocitta* as

sister to all other species in the clade (but with no nodal support). These discrepancies reflect in low support for nodes in Clade A, except for that of Group I (1.00 bayesian pp, 100% ML and MP boot).

Clade B is formed by a polytomy among *C. yncas, C. mystacalis,* and a well-supported clade containing "*Cissilopha*" jays + all other *Cyanocorax* species. Within "*Cissilopha,*" *C. melanocyaneus* is at the base, and *C. yucatanicus* is sister to *C. sanblasianus* + *C. beecheii;* however, *C. sanblasianus* showed as paraphyletic with respect to *C. beecheii.* Among the remaining species, *C. dickeyi* is sister to a clade ("Group II"), in which *C. affinis* + *C. heilprini* are reciprocally monophyletic to *C. cayanus* (*C. chrysops* + *C. cyanopogon*). Relationships within Clade B are consistent among MP, ML, and BA of the combined dataset, with two exceptions. ML shows *C. mystacalis* and *C. yncas* as monophyletic and sister to all other species (58% ML bootstrap support) and MP does not recover the node of *C. sanblasianus* + *C. beecheii.*

In general, individual analyses of the three nuclear introns were less informative regarding relationships within *Cyanocorax* (Fig. 3.2). AK5 recovers Clade A with high support. βfb7 recovers Clade A, the monophyly of *Cyanocorax* + *Calocitta* + *Psilorhinu*s, and a sister relationship between *Psilorhinus* and *Calocitta;* other relationships disagree with those recovered by the mitochondrial genes, but with low bootstrap support. TGFβ2.5 provided the most information, supporting the monophyly of *Cyanocorax* + *Calocitta* + *Psilorhinus,* Clades A and B, *Calocitta* + *Psilorhinus,* and *Cyanocorax melanocyaneus* + *C. yucatanicus*. In all cases, both ML and MP analyses are consistent across genes.

Analyses of the combined, total evidence dataset produces similar topologies to those of the mitochondrial dataset with few exceptions (Fig. 3.3). Within Clade A, *Calocitta* and *Psilorhinus* are placed as sister taxa (as they are in analyses of βfb7 and TGFβ2.5), but nodal support indicates a virtual polytomy among *Calocitta, Psilorhinus, Cyanocorax caeruleus, C. violaceus* from Ecuador, and Group I (*C. violaceus* from Peru + *C. cristatellus* + *C. cyanomelas*); also, *C. cristatellus* + *C. cyanomelas* was not recovered. Within Clade B, most relationships are well supported, and a sister relationship between *C. yncas* and all other *Cyanocorax* species is recovered by BA, but not by ML or MP. The MP 50% majority rule consensus tree (1050 trees, 2421 steps, CI = 0.4643, RC = 0.3707; not shown) produces the same general topology as ML and BA; the few relationships that disagree among criteria are weakly supported.

Hypothesis Testing

According to the parametric boostrapping tests, the null hypotheses of monophyly of the current concept of *Cyanocorax* and *C. mystacalis* + *C. dickeyi* are rejected with high significance. In testing the monophyly of *Cyanocorax* the difference between the ML tree and the constraint ML tree drawn from the real dataset (–ln likelihood = 15669.408 vs. 15729.410; δ = 60.002 ln units) is significantly different from δ values obtained by simulation ($P < 0.01$). Therefore, *Calocitta* and *Psilorhinus* must be considered as integral parts of the *Cyanocorax* assemblage. Also, the difference between the ML tree and the constraint ML tree enforcing the monophyly of *C. mystacalis* + *C. dickeyi*, was highly significant (–ln likelihood = 15669.408 vs. 15688.984; δ = 19.576 ln units; $P < 0.01$).

DISCUSSION

Phylogenetic analyses based on mitochondrial and nuclear genes, as well as on different optimization criteria, produce a basic structure for species-level relationships within *Cyanocorax* and allied genera. Disagreements among genes and datasets involve only on weakly supported subclades in combined (mitochondrial and total evidence) analyses. Given that stationarity in base composition can not be rejected for any gene, and evolutionary rate heterogeneity is detected only for cyt*b* (for which results are generally congruent with those of ND2), I discard the possibility of these differences being caused by heterogeneities in base composition or evolutionary rate. Regardless of the inconsistencies encountered, clear patterns emerge concerning the relationships among *Cyanocorax* and related jays.

The most important result is the paraphyly of *Cyanocorax* with respect to *Psilorhinus* and *Calocitta*. Consistent with the high bootstrap support observed for these relationships, the parametric bootstrapping test of rejects the monophyly of the current *Cyanocorax*. A second major result is the division of all species into two well-supported groups—viz., Clades A and B. Clade A includes *Psilorhinus, Calocitta, C. caeruleus*, *Cyanocorax violaceus* from Ecuador, and samples in Group I (*C. cyanomelas* + *C. cristatellus* + *C. violaceus* from Peru). Clade B contains *C. yncas, C. mystacalis,* the "*Cissilopha*" jays*, C. dickeyi,* and species in Group II (*C. heilprini, C. affinis, C. cayanus, C. chrysops,* and *C. cyanopogon*).

Phylogenetic Relationships within Clade A

Species relationships within this clade are particularly unstable with respect to the position of the Brown Jay, *Psilorhinus morio,* and its relationships with *C. caeruleus* and *C. violaceus* from Ecuador (individual gene and combined mitochondrial analyses), whereas a relatively well-supported sister relationship between *Calocitta* and *Psilorhinus* is obtained from analyses of independent nuclear loci (βf7 and TGFβ2.5, Fig. 3.2) and recovered, albeit with low support, in the combined analyses (Fig. 3.3). Similar problems in placement of *Psilorhinus* were encountered in the morphological study of Hope (1989), in which this species was recovered as sister to *Calocitta*, or to *C. violaceus* + *C. caeruleus* + *C. cyanomelas,* depending on outgroup selection.

Uncertainty in the position of *Psilorhinus* and *Cyanocorax caeruleus* and *C. violaceus* from Ecuador contrasts with the high support obtained for Group I (*C. cyanomelas*, *C. cristatellus*, and *C. violaceus* from Peru) in all combined analyses. Common ancestry among *C. cyanomelas, C. violaceus,* and *C. caeruleus* was expected based on general morphology (i.e., "*Coronideus*"group, Amadon 1944), as well as two osteological synapomorphies (one unreversed, another present also independently in *Cyanolyca;* Hope 1989). *Cyanocorax cristatellus* was not included in Hope (1989); thus, it is not possible to assess whether its molecular affinities with *C. violaceus* and *C. cyanomelas* coincide with osteological synapomorphies. Additionally, the unstable position of *C. caeruleus* and paraphyly of *C. violaceus* prevent clear conclusions about relationships among species in Clade A.

With regard to the paraphyly of *C. violaceus,* examination of the study from skins from which tissue was sequenced did not reveal marked differences this representatives of the Peruvian population other than slight color variation, attributable to preservation artifacts (N. Rice, pers. comm.). Given the long, well-supported internodal branch leading to Group I (*C. violaceus* from Peru + *C. cyanomelas* + *C. cristatellus*), it is improbable that this structure is caused by stochastic phylogenetic error (Funk and Omland, 2003), hybridization, or incomplete lineage sorting (McCracken and Sorenson, 2005). Potential for amplification paralogous genes (i.e., nuclear pseudogenes, Benssason et al., 2001; or paternal mitochondrial gene copies, Brommham et al., 2003; Hoarau et al., 2003) does not seem a conceivable explanation, because all sequences obtained showed the substitution patterns of genuine mitochondrial coding genes, and no ambiguous base callings were detected in chromatograms. Alternate explanations include the possibilities that samples from Ecuador: (1) represent a cryptic species north of the Amazon River that have retained an ancestral plumage or (2) carry ancestrally polymorphic mitochondrial haplotypes. Distinguishing between these two scenarios, however, may require extensive sampling of populations of *C. violaceus* and allied species across their ranges, as well as incorporation of faster-evolving nuclear markers.

Phylogenetic Relationships within Clade B

In general, phylogenetic structure within Clade B was stable across analyses, although the positions of *Cyanocorax mystacalis* and *C. yncas* are ambiguous. Clearly, independent of its position in the tree, *C. yncas* sequences separated in distinct groups in the North and South American portions of the range of the species (Fig. 3.4). Differences

in plumage, habitat preferences (e.g., Meyer de Shauensee, 1966; Goodwin, 1976), social behavior (Alvarez, 1975; Gayou, 1986), and vocalizations (Ridgely and Greenfield, 2001) suggest that these populations might represent different species (Ridgely and Greenfield, 2001; Hilty, 2003): *C. yncas* (from northern Colombia and Venezuela to northern Bolivia) and *C. luxuosus* (from Texas to northern Honduras). Further analysis of populations from across the range of the species, particularly in Central America and northern South America, will be crucial in assessing their validity as independent evolutionary lineages (sensu Wiley, 1978).

The monophyly of the *"Cissilopha"* jays is consistent with their shared plumage characters, as well as their geographic distribution in Mexico and Central America, fitting previous assessments of relationships (e.g., Amadon, 1944; Goodwin, 1976; Hardy, 1969; Hope, 1989). Although paraphyly of *C. sanblasianus* with respect to *C. beecheii* is unexpected given their differences in plumage and size, the short branches separating sequences of these species suggest that speciation might have occurred relatively recently in evolutionary time; if such were the case, a pattern of paraphyly would be expected under coalescent theory (Avise and Ball 1990, Hudson 1990, Avise 1994). Empirical support for this scenario comes from studies documenting rapid evolution of morphological characters coupled with low levels of molecular divergence and species paraphyly (e.g., *Spizella taverneri* and *S. breweri*, Klicka et al., 1999; Galapagos ground finches, Sato et al., 1999; *Icterus galbula* and *I. abeillei*, Kondo et al., 2004). Interestingly, although low levels of molecular divergence and species paraphyly were once seen as exceptions, rather than rules (e.g., Klicka and Zink, 1997; Klicka et al., 1999), in recent years, the number of studies showing these phenomena has been directly

related to deep scrutiny of sister-species relationships (e.g., Johnson and Cicero, 2004) and analyses including multiple samples per species (Funk and Omland, 2003; for an example, see Kondo et al., 2004).

An alternate explanation would be introgression in the area of overlap between *Cyanocorax beecheii* and *C. sanblasianus* in western Nayarit (Fig. 3.4). Although examination of the specimens involved did not reveal any deviations from their expected morphology, hybridization and introgression could have occurred in the history of these species. Occasional reports of hybrids between NWJ species (e.g., *Psilorhinus* × *Calocitta* from western Chiapas, Mexico; Pitelka et al., 1956) suggest that hybridization may occur in the wild. Moreover, mitochondrial DNA is particularly susceptible to introgression because of the lack of recombination (Smith, 1992), being expected to introgress farther than nuclear loci if their persistence in a foreign gene pool is less constrained by linkage to selected loci than are alleles of nuclear genes (Harrison, 1990; Funk and Omland, 2003). However, as heterogametic taxa following Haldane's rule, birds are predicted to have reduced viability of female hybrids, which might restrict the introgression of maternally inherited DNA (Funk and Omland, 2003). This constraint explains the low mitochondrial DNA introgression observed in several avian hybrid zones (e.g., Brumfield et al., 2001; Sattler and Braun, 2000). However, discerning among incomplete lineage sorting and occasional hybridization/introgression in the present case is not possible with the data currently available.

Regarding the South American jays in Group II, the sister relationship of *Cyanocorax chrysops* and *C. cyanopogon* was expected given their close similarity in plumage and their early placement in a single species (Hellmayr, 1934). A close

relationship with *C. cayanus* is consistent with their similar morphology and geographic distribution north (*C. cayanus*) and south (*C. chrysops* + *C. cyanopogon*) of the Amazon River (Fig. 3.4). On the other hand, the sister relationship between *C. heilprini* and *C. affinis* is interesting on morphological and ecological grounds. *Cyanocorax heilprini* differs from *C. affinis* and all other species in its group in having a blue venter and being specialized on white-sand habitats around the black-water rivers of the Río Negro Basin. Finally, the position of *C. dickeyi* as sister to Group II and the significance of the parametric bootstrapping test, contradict previous hypothesis regarding its close relationship with *C. mystacalis* (Haemig, 1979; Hope 1989). Rather, this result supports the hypothesis of retention of ancestral morphology in *C. dickeyi*.

Biogeography

The phylogenetic relationships reconstructed herein illustrate why past attempts to understand the biogeography of *Cyanocorax* and allied genera have been unfruitful and puzzling (e.g., Hardy, 1961; Hardy, 1969; Goodwin, 1976). The general pattern that emerges is one in which relatively recent radiations (e.g., the Group II and the "*Cissilopha*" jays) seem to fit general expectations based on biogeography and morphology. However, farther back in evolutionary time, traces of ancestral biogeographic history seem to be lost from the phylogenetic record. Two clear cases of discontinuous geographic distributions are recovered by the phylogeny—*Calocitta* and *Psilorhinus* from North and Central America being well supported as most closely related to species in the Amazon Basin and central South America (Fig. 3.5), and *C. dickeyi* from northwestern Mexico and the *"Cissilopha"* jays from western Mexico and northern

Central America being placed within a clade of mainly South American species (Fig. 3.4).

Repeated instances of discontinuous distributions across the phylogeny could be explained by two different processes—i.e., long-distance dispersal or local extinction of widely distributed ancestors. Some species in this group are known to stray from their customary distributional areas (e.g., *C. yncas* [Hilty, 2003]), occasionally by hundreds of kilometers (*C. sanblasianus* in Arizona [Phillips, 1950]; "*Cissilopha*" sp. in Texas, [J. Eitniear, pers. comm.]). However, the idea that these occasional events could be responsible for these continental-scale distributional gaps seems unlikely.

Local extinction of a broadly distributed ancestor has been invoked to explain the peculiar distribution of *C. dickeyi* (Amadon, 1944). Interestingly, extinction seems to be the only plausible explanation for broad distributional gaps observed in (non-migratory) avian taxa, for which reasonable geographic sampling has been accomplished: e.g., azure-winged magpie, *Cyanopica cyanus*, disjunct distribution in the Iberian peninsula and Asia, (Fok et al., 2002); wrentits, *Chamaea fasciata*, from western North America, more closely related to the Old World genus *Sylvia* (Barhoum and Burns, 2002); South American and Asian piculets, *Picumnus*, most closely related to each other (Benz et al., 2006). For example, the discontinuous distribution of *C. yncas,* from northern Texas to northern Honduras, and from northern Colombia to Bolivia (Fig. 3.4), could represent the early stage of a geographic gap produced by local extinction. Given a high potential for extinction, the usual exercise of reconstructing ancestral areas would be misleading in this case, and could produce simply erroneous results.

Systematic Implications

Phylogenetic analysis of species in *Cyanocorax* and allied genera reveals that *Cyanocorax* is paraphyletic, as defined currently. This result is supported by the preferred phylogenetic hypothesis and by the rejection of the null hypothesis of monophyly of the genus. Thus, if phylogenetic relationships are to be reflected in systematic classification, and if taxa are to be natural groups, the genus *Cyanocorax* must be redefined. From the relationships recovered for the overall assemblage, two different approaches could be taken.

The first option involves maintaining the genus *Cyanocorax* specifically for species in Clade B—*C. chrysops* Boie 1826 is the type species (Hellmayr, 1934)—and naming species in Clade A with the next available name, which would be *Psilorhinus* Ruppell 1838 (Hellmayr, 1934). Given the relatively large amount of data presented herein, I do not foresee an unambiguous resolution of these relationships in the short term, unless a greater amount of faster-evolving nuclear markers can be incorporated in analyses. The second approach would be to include all taxa (including *Psilorhinus* and *Calocitta*) into a broader concept of *Cyanocorax*, given its taxonomic priority. Some current classifications (AOU, 1983, 1998; Dickinson, 2003) already submerge *Psilorhinus morio* in *Cyanocorax*. Placing the contents of *Calocitta* into *Cyanocorax* would be a minimal change in taxonomy.

Based on the phylogenetic results, the monophyletic group including *Cyanocorax, Calocitta,* and *Psilorhinus,* is highly diverse from morphological (e.g., Moore, 1935; Amadon 1944; Hardy 1969; Hope, 1989), behavioral (Alvarez, 1975; Gayou, 1986; Peterson, 1991), and ecological (e.g., *C. heilprini*) points of view. Although a previous

study (Bonaccorso and Peterson, 2007; Chapter 1) discussed the uniqueness of *Psilorhinus morio* in terms of its possession of a "furcular pouch" (Sutton and Gilberg 1942) reflecting the actual morphological variation within the overall assemblage would require recognizing a multiplicity of genera: *Xanthoura* (if *Cyanocorax yncas* were sister to all other species in Clade B), *Cyanocorax* (the remaining species in Clade B), *Calocitta, Psilorhinus* (if sister to *Calocitta*), and *Uroleuca* Bonaparte 1859 (including *Cyanocorax cristatellus, C. violaceus, C. cyanomelas, C. caeruleus*). On the contrary, merging the contents of *Psilorhinus* and *Calocitta* into *Cyanocorax* is a step toward recognizing the uniqueness of group as a whole, and, as a consequence, focuses attention on understanding the evolutionary mechanisms that have produced this remarkable diversity.

Conclusions

Analyses of phylogenetic relationships among *Cyanocorax, Psilorhinus,* and *Calocitta,* revealed consistent patterns indicating overall monophyly of the group, but paraphyly of the current *Cyanocorax*. Also, division of ingroup taxa into two reciprocally monophyletic groups (Clades A and B) is highly supported. Within Clade A, a robust node indicates the monophyly of *C. violaceus* from Peru + *C. cristatellus* + *C. cyanomelas* (Group I). Based on previous morphological studies (Hope 1989), it is possible that this clade includes also *C. caeruleus;* however, the ambiguous phylogenetic position of this species and *Psilorhinus morio,* and the paraphyly of *C. violaceus* prevent major conclusions on relationships within this clade. Within Clade B, ambiguous relationships were limited to resolving the positions of *C. yncas* and *C. mystacalis.* The

"*Cissilopha*" jays form a monophyletic group sister to *C. dickeyi* + Group II. I suggest that the paraphyly of *C. sanblasianus* is most likely the result of incomplete lineage sorting.

Further resolution of problematic, weakly supported relationships associated with relatively short internodal branches (e.g., the positions of *C. mystacalis*, *C. yncas*, and *Psilorhinus*) seems difficult at this point. Theoretical models indicate that when the time of shared ancestry is short compared to that of independent ancestry (e.g., short internodal branches leading to long terminal branches), the number of informative characters is so small that the probabilities of recovering a bifurcating pattern is low (Lanyon, 1988). Moreover, signals of common ancestry (i.e., synapomorphies) are most likely to be overwritten by homoplasy (McCracken and Sorenson, 2005). Thus, resolving these relationships might require considerably amounts of mitochondrial and nuclear characters.

The phylogenetic complexity recovered in this study suggests several complications for the taxonomy of this group. Although high diversity in morphology indicates that multiple taxonomic arrangements might be possible, my recommendations are for recognizing one genus only, *Cyanocorax*, including species currently assigned to both *Psilorhinus* and *Calocitta*. This proposition maximizes stability of a systematic classification consistent with phylogeny.

The genealogical patterns recovered in this study highlight the importance analyzing multiple samples per taxa, particularly as it concerns to the detection of species paraphyly. However, more extensive population-level, multi-locus analyses will be necessary to understand the origin of these patterns. Finally, these results emphasize the

importance of collecting and preserving proper voucher specimens when trying to detect coupled genetic and morphological change, as well as traces of hybridization.

CHAPTER 4

ECOLOGICAL DIVERSIFICATION IN THE *CYANOCORAX* JAYS

The study of historical and ecological biogeography has experienced considerable development and transformation in recent years. Integration of analytical tools from geographic information systems (GIS) and phylogenetic analysis has provided enormous potential for understanding the geographic, ecological, and historical underpinnings of organismal diversity (e.g., Rode and Lieberman, 2005; Smith et al., 2005; Wiens et al., 2006). Furthermore, the analysis and characterization of ecological niches have become integral components of studies of species' diversification and evolution (e.g., Rice et al., 2003; Graham et al., 2004; Martínez-Meyer et al., 2004; Kozak and Wiens, 2006; Knouft et al., 2006; Carstens and Richards, 2007).

Relative stasis of ecological niches in evolutionary time has been implicated in promoting speciation between allopatric species (Kozak and Wiens, 2006). In sympatry, however, niche conservatism may be countered by natural selection favoring ecological divergence to minimize interspecific interactions (Losos et al., 2003). Evolutionary history of allopatric, parapatric, and marginally sympatric lineages, in which both niches and geographic distributions might have changed in evolutionary time, involve even more complex and dynamic scenarios (e.g., Losos and Glor, 2003). Allopatry may result from geographic splitting of populations followed by either stasis (Peterson et al., 1999) or evolution in the ecological realm (Vanzolini and Williams, 1981; Graham et al., 2004). On the other hand, parapatry may result from speciation across ecological gradients (i.e., divergence with gene flow; e.g., Endler, 1982), or from secondary contact after allopatric (non-ecologically mediated) speciation. However, assessing the role of ecological

differentiation in “lineage splitting” vs. “lineage diversification” (Wiens, 2004) requires determination of whether niches have evolved in the first place, and exploring the potential mechanisms behind their evolution.

Species in the genus *Cyanocorax* and the allied genera *Psilorhinus* and *Calocitta,* a monophyletic assemblage of jays endemic to tropical and subtropical America, offer opportunities for studying evolutionary changes in ecological niches and their relationship to speciation. *Cyanocorax* jays occur from the extreme southern United States south to northern Argentina and occupy diverse habitats, including cloud forest (*Cyanocorax melanocyaneus*), humid foothill forest (*C. affinis*), mixed pine-oak forest (*C. dickeyi*), arid scrub (*Calocitta* spp.), dry forest, cactus-dominated desert (*C. mystacalis*), Amazon “varzea” forest (*C. violaceus*), Brazilian “cerrado” (*C. cristatellus* and *C. cyanopogon*), and white-sand river gallery forest (*C. heilprini*) (Moore, 1935; Monroe, 1968; Hardy, 1969; Ridgely and Tudor, 1989).

Molecular phylogenetic analyses of the *Cyanocorax* assemblage (Chapter 3; Fig. 4.1) confirm that species that overlap broadly in their geographic ranges are not closely related within the overall group. Two pairs of sister taxa (i.e., *C. affinis* and *C. heilprini,* and the northern and southern isolates of *C. yncas*) are allopatric, and three pairs of sister species are marginally sympatric (i.e., *C. beecheii* and *C. sanblasianus*, *C. chrysops* and *C. cyanopogon*, *Calocitta formosa* and *C. colliei*). Three other species show marginal sympatry (*Cyanocorax cyanomelas* with *C. cristatellus* and *C. violaceus*), but their phylogenetic positions remain poorly resolved.

Herein, I characterize ecological niches of species in the *Cyanocorax* assemblage in terms of a suite of environmental variables, using tools from ecological niche

modeling. Once characterized, a molecular-based phylogenetic hypothesis (Chapter 3) provides a framework for analyzing ecological niches in an historical evolutionary perspective. First, I test for ecological niche conservatism among sister species. The observation of currently allopatric sister taxa having high niche conservatism would be consistent with the predictions of an allopatric mode of speciation (Peterson, 1999; Kozak and Wiens, 2006); on the other hand, currently parapatric, or marginally sympatric, sister taxa separated by sharp ecological gradients, would be expected from a parapatric mode of speciation (Endler, 1982) associated to ecological niche differentiation.

Second, ecological distances (derived from the ecological niche models) are compared across the *Cyanocorax* assemblage to visualize differences in ecological niches across lineages. Finally, ecological niches are reconstructed over the *Cyanocorax* tree to explore their associations with phylogenetic structure (i.e., phylogenetic inertia), as well as their evolutionary trends. Results are discussed in terms of their implications for understanding the diversification of the *Cyanocorax* jays.

METHODS

Cyanocorax Phylogeny

Molecular phylogenetic analyses (Chapter 3) based on complete taxonomic sampling and nuclear and mitochondrial gene sequences confirm the monophyly of *Cyanocorax* + *Calocitta* + *Psilorhinus* and indicate a deep division into two groups—viz., Clades A and B (Fig. 4.1). Because phylogenetic relationships in Clade A remain poorly resolved, I focus on Clade B, for which all nodes in the phylogeny are robust,

except for the node defining the position of *C. mystacalis*. Also, the monophyly of *C. beecheii* + *C. sanblasianus* is weakly supported, possibly because of the few data available for these two species, for which DNA had to be extracted from museum study skins.

Analyses involving sister-species comparisons include two species from Clade B, *Calocitta formosa* and *C. colliei*, which were recovered as sister taxa in all analyses. Also, given the molecular (Chapter 3), morphological (Goodwin, 1976), and behavioral (Alvarez, 1975; Gayou, 1986) differences between the northern and southern isolates of *Cyanocorax yncas* (i.e., *C. y. luxuosus* and *C. y. yncas*), they are considered as sister taxa for effects of analyses and their ecological niches are analyzed separately.

For the present study, I estimated a pruned ML tree that included one sample per species. To avoid potential distortion of branch lengths, phylogenetic analyses are based only on ND2 sequences for all species, given that more limited information is available for other loci (cyt *b* and nuclear introns) in the original analyses. This phylogenetic tree was estimated in RaxML (Stamatakis, 2006), using a partition by codon-position and the same settings and parameters used in Chapter 3.

Occurrence and Environmental Data

Distributional information for *Cyanocorax* species was based on locality data associated with specimens housed in natural history museums drawn from ORNIS and individual institutions (cited in Acknowledgments). All occurrence records were

georeferenced to the nearest 1' using on-line gazetteers[1]. This information was complemented with georeferenced occurrence records from the Atlas of Mexican Bird Distributions (Peterson et al., 1998; Navarro-Sigüenza et al., 2002, 2003), a compendium of label data from natural history museum specimens of birds of Mexico. Despite the limited data available for several species—especially those inhabiting vast areas of South America—a reasonable amount of spatially unique localities (≥45) was obtained for all taxa, except for *C. heilprini*, a rare species of northern Amazonia, for which only 10 unique records were available.

I used 2.5' resolution GIS coverages summarizing aspects of temperature and precipitation (WorldClim[2]) to describe environmental characteristics across the range of the assemblage. A subset of seven coverages (of the 19 available) was chosen based on their utility in describing global climate patterns (i.e., annual mean temperature, annual precipitation), as well as extreme climatic conditions (i.e., mean diurnal range, maximum temperature of the warmest month, minimum temperature of the coldest month, precipitation of the wettest month, precipitation of the driest month). I avoided inclusion of too many environmental dimensions to prevent (1) limiting predictability caused by over parameterization of models (Peterson and Nakazawa, *in press*), and (2) problems related to broad seasonal differences between Northern and Southern hemispheres.

Ecological Niche Modeling

[1] http://www.fallingrain.com; http://gnswww.nga.mil; http://webclient.alexandria.ucsb.edu/client/gaz/adl/index.jsp
[2] http://biogeo.berkeley.edu/worldclim/worldclim.htm

The present approach focuses on approximation of fundamental niches (Peterson, 2001), the conjunction of environmental combinations in which species can survive and maintain viable populations (Hutchinson, 1959). Although several analytical tools have been developed to model ecological niches (Elith et al., 2006), one such tools, the Genetic Algorithm for Rule-set Prediction (GARP), has been tested under extensive sensitivity analysis and validation (Peterson and Cohoon, 1999; Peterson, 2001; Peterson and Shaw, 2003; Stockwell and Peterson, 2002a, b, 2003; Anderson et al. 2003; Levine et al., 2004). Additionally, it has proved robust in modeling niches and projecting them over broad, unsampled regions (Peterson et al., 2007). Therefore, I employed GARP to investigate ecological niches of the *Cyanocorax* jays.

GARP searches for non-random associations among environmental characteristics of localities in which species are known to occur versus environmental features characterizing the entire study region (Stockwell and Peters, 1999). It works in an iterative process of rule selection, evaluation, testing, and incorporation or rejection to produce a heterogeneous rule set summarizing the species' ecological requirements. First, a method chosen from a set of possibilities (e.g., logistic regression, bioclimatic envelope rules) is applied to the data, and a rule is developed. Then, predictive accuracy (for *intrinsic* model refinement) is evaluated based on points resampled both from the known distribution and from the study region as a whole. Change in predictive accuracy from one iteration to the next is used to evaluate whether a particular rule should be incorporated into the model. The algorithm runs for a number of iterations or until convergence.

Because of the stochastic nature of the GARP modeling process, predictions vary among individual replicate models (Anderson et al., 2003). Therefore, a number of models are generated, and final predictions are based on consensus among a subset of models, chosen following specific criteria (e.g., models able to predict known occurrences—measured as omission error—and sufficiently broad to predict areas from where not data is available—measured as commission error). Finally, model predictions are projected into geographic space as digital (GIS) maps of areas with environmental characteristics suitable for the species.

Model Development and Testing

A three-step procedure was used, including preliminary model development, independent testing, and final model development. First, a random subset (~50%) of the occurrence points was used to develop models that were trained across the extent of known range of each species, based on locality records. The area of analysis for each species was defined by a buffer around the occurrence points; this buffer had a radius equal to the maximum distance between occurrence points of the species. In all, 100 models were generated per testing dataset, with GARP experiments running for 1000 iterations or until convergence (convergence parameter set at 0.01). Preliminary niche predictions were based on a best subset of models (Anderson et al., 2003) with omission error $\leq 5\%$, which were intersected (in the form of Arc grids) using grid multiplication in ArcView 3.1 (ESRI, 1998). This operation permitted summarizing commonalities among low-omission models.

Second, significance and accuracy of preliminary predictions were tested via overlay of the remaining 50% of the occurrence points. Significance was defined as the ability to predict testing points better than expected at random (Anderson et al., 2002). Because expected values were not always large enough to permit use of a χ^2 statistic (Zar, 1996), exact one-tailed binomial probabilities of predicting actual presences (test points) were calculated given overall proportion of pixels predicted present (= probability of success). The percentage of independent test points falling into regions of predicted presence was taken as a measure of model accuracy (Anderson et al., 2002). It is possible to achieve statistical significance without predicting an acceptable proportion of test points, if the predicted area is a small proportion of the study area; conversely, models can attain high accuracy without achieving significance by including a large proportion of the study area (Anderson et al., 2002; Peterson, 2005); thus, overall model quality was assessed based on both significance and accuracy.

The small number of occurrence points available for *Cyanocorax heilprini* (n = 10), required a different approach for assessing model quality. In this case, a jackknife procedure developed recently (Pearson et al., 2007) was implemented. Each observed locality was removed once from the dataset and models were built using the remaining n –1 localities. Significance was assessed based on the ability of each model to predict the single locality excluded from the training dataset, according to the following criterion. If p_i is the proportion of the study area predicted present having deleted the i^{th} point, and X_i a binary success-failure variable to indicate whether the ith point is included in the ith predicted area (X_i = 1 if such point is included in the prediction, and X_i = 0 if otherwise). Let H denote the assumption of a completely random assignment; under H, X_i is a

random trial with probability of success $P_{i.}$ To examine whether observed successes constitute evidence against H (in the direction of better than random), a P value is obtained based on the test statistic: $D = \Sigma X_i (1 - P_i)$, in which a success ($X_i = 1$) carries a greater weight ($1 - P_i$) if it has occurred under a small assumed probability (P_i). Thus, D ranks possible values of Xs according to the evidence they provide against H. If d denotes the observed value of D, the corresponding P value is computed as the probability under H that $D \geq d$. Because the distribution of D is dependent on the proportional areas predicted in the n runs, a universal table of P values cannot be created and the probability calculation must be achieved by exhaustive case-by-case summation; thus, a program for calculating this P value is available (Pearson et al., 2007, Supplementary Appendix S1[3]). An implicit assumption in the calculation of D is that the jackknife trials are independent. In a strict, sense this assumption is not true because jackknife trials share most of the data; however, when sample sizes are small, each locality is expected to have a large influence on model prediction, giving varying results from different combinations of available localities (Pearson et al., 2007).

Final models were developed using all occurrence points available per species, based on the same GARP settings described previously. Final predictions where obtained according to the same criteria used for building preliminary predictions, and then were projected across the entire distributional area of the *Cyanocorax* assemblage. Projecting onto this broader area, which covers most of the tropics and subtropics of the Americas (i.e., southern United States south to northern Argentina), allowed comparisons of

[3] http://www.blackwell synergy.com/doi/abs/

ecological niches among species in ecological and geographic dimensions, beyond the limits of distributional area of each species.

Testing for Ecological Differentiation

Expectations of ecological niche conservatism can be tested in the geographic realm by asking whether pairs of sister taxa can predict each other's ecological niches (Peterson et al., 1999); this amounts to a test of a null hypothesis of niche identity. Niche "interpredictivity" was assessed for the following pairs of sister taxa: *Calocitta formosa* and. *Calocitta colliei, C. sanblasianus* and. *C. beecheii, Cyanocorax heiprini* and. *C. affinis*, *C. chrysops* and. *C. cyanopogon*, and *C. y. yncas* and *C. y. luxuosus*. Reciprocal predictivity was tested following similar procedures to those used in preliminary model testing. After final niche predictions were projected over the entire area of analysis, the prediction of Species A (predicting species) was clipped using the buffered area around the occurrence points of Species B (predicted species). Then, one-tailed binomial probabilities of predicting occurrences of Species B were calculated given overall proportion of pixels predicted present by Species A (in the buffer area of B). The reciprocal test (B predicts A) was performed in the same fashion. In summary, this procedure tests whether sister species can predict each other's occurrence points better than random.

Statistical tests of species predictions may suffer from effects of spatial autocorrelation. This situation arises because grid values associated with species' model predictions are not independent from one another, given that contiguous grid cells are

expected to be more similar to each other than to other cells in the landscape. In the light of these limitations, only highly significant results ($P < 0.001$) are discussed.

Visualizing Niche Differentiation in Ecological Space

To complement the geographic perspective of niche differentiation with an ecological perspective, visualization of occurrence points and niche availability was achieved as follows. In ArcView the "combine grid" option was used to create a composite grid of input environmental grids in each species' area of analysis (buffer area). Then, the attributes table associated with this grid was exported in text format. This table was imported into SPSS (SPSS Inc., 2006) and data density was reduced by selecting a random to 10% of records. Original dimensionality (7 environmental variables) associated to buffer areas and occurrence points, was reduced by performing a canonical discriminant function analysis (DFA). Apart from observing general trends of niches in ecological space, this visualization aimed to determine the roll of niche availability in observed niche differentiation among species (Kambhampati and Peterson, 2007). Such consideration is important because niche differentiation may be real if Species A has access to, but does not inhabit, the environment occupied by Species B; in contrast, niche differentiation could be apparent if Species A does not inhabit the environment occupied by Species B, simply because those conditions do not exist near its distributional area (Kambhampati and Peterson, 2007).

Comparing Ecological Niche Distances

To visualize overall ecological distances among species, distance matrices were generated and analyzed as follows. First, each ecological variable derived from the niche predictions (Arc composite grid) was *z*-standardized (i.e., subtracting the mean and dividing by the standard deviation), to produce a standard normal variable with mean = 0 and variance = 1 (Abdi, 2007). Then, a distance matrix was calculated based on mean Euclidean distances among centroids of the *z*-standarized variables (Martínez-Meyer, 2002).

The centroid-distance matrix was used to build an unrooted distance tree via least-squares calculation implemented in Fitch (PHYLIP[4]; Felsenstein, 1982); this procedure was used for visualization of the distance-matrix only, and not for making assessments of the evolution of ecological characteristics of species. To visualize ecological differences in a phylogenetic framework, the ecological distance matrix was fitted to the *Cyanocorax* phylogeny, using the least-squares fitting and user-defined tree options in Fitch. Branch lengths resulting from this procedure were taken as a measure of evolutionary change in ecological niche characteristics along each branch in the tree (Martínez-Meyer et al., 2004). To test for potential correlation between ecological and molecular distances (i.e., patristic distances based on the ML tree), a Mantel test was performed in Arlequin 3.11 (Excoffier, 2005) with significance assessed via 1000 random permutations.

Testing for Phylogenetic Inertia

To test the effects of phylogeny on evolution of ecological characteristics, individual niche variables were reconstructed directly into on the phylogeny.

[4] http://evolution.genetics.washington.edu/phylip.html

Phylogenetic reconstructions were performed in the ML continuous character evolution module implemented in BayesTraits (Pagel, 1997, 1999). This procedure uses the generalized least-squares model (GLS) to evaluate whether a trait covaries with evolutionary divergence. Two models of evolution, A and B, are available: Model A corresponds to the constant-variance random-walk model, which has as single parameter, the instantaneous variance of evolution, whereas Model B is a directional random-walk with two parameters, the variance of evolution (as in model A) and a directional change parameter (β). This last parameter measures the regression of trait values across species against the total path length of the tree, from roots to tips; it detects any general trends towards a dominant direction of evolutionary change (e.g., general increase in temperature or precipitation). The fit of models A and B to the data can be assessed through a likelihood-ratio test (LRT), in which likelihood differences are χ^2-distributed (given that Model A is a particular case of Model B), with degrees of freedom equal to the difference in numbers of parameters between the two models. If Model B fits the data better, a directional trend exists. The ability to detect directionality is an advantage of this method over Felsenstein's (1985) independent contrast (Pagel, 1997). Apart from estimating α (the root value), GLS does not calculate values explicitly at internal nodes, avoiding some of the pitfalls inherent in ancestral character reconstruction (Pagel, 1997; Schluter et al., 1997).

Additionally, this procedure allows estimation of three scaling parameters: lambda (λ), kappa (κ), and delta (δ). The λ parameter reveals whether the phylogeny predicts the patterns of covariance among species on a given trait; if λ is significantly < 1, traits are evolving independently of the phylogeny. Kappa scales individual branch lengths, and

can be used to test punctual ($\kappa = 0$) vs. gradual ($\kappa > 0$) modes of trait evolution. Finally, δ scales overall path lengths in the phylogeny (i.e., distance from the root to the species and shared path lengths), and can detect signals of adaptative radiation ($\delta < 1$) or species-specific adaptation ($\delta > 1$). As in the comparison between models A and B, a likelihood-ratio test is indicated for comparing (simpler) models in which parameters take a fixed value, with models in which parameters are estimated from the data; the degrees of freedom are determined by the difference in numbers of paramenters estimated.

In this context, two hypotheses were tested for each ecological variable using the likelihood-ratio test: (1) likelihood of Model B vs. that of Model A (directional random-walk vs. random-walk); and (2) after selecting Model A or B, likelihood of the best model (all parameters estimated from the data) vs. that of a model in which $\lambda = 1$ (i.e., phylogenetic inertia enforced) and all other parameters are estimated from the data. Values of κ and δ were recorded for descriptive purposes, given that no trends were expected a priori to be dominant in the phylogeny. Ecological traits reconstructed on the phylogeny included the mean of each environmental variable and the niche centroid (calculated for the distance comparisons).

RESULTS

Phylogeny and Niche Models

The phylogeny that results from the analysis of the pruned ND2 dataset recovers the same topology obtained from the analysis of the combined mitochondrial and nuclear datasets (Chapter 3; Fig. 4.1). Tests of model predictivity developed for preliminary models are significant for all species ($P < 0.05$; Appendix I). In the five cases in which

model accuracy was lower for one independent (50% point split) trial (i.e., ~60%; for *C. affinis*, *C. cyanopogon*, *C. melanocyaneus*, *C. mystacalis*, *C. yncas yncas*), the reciprocal trial showed high accuracy (78–92%). Also, jackknife tests for *C. heilprini* were highly significant ($D = 0.7$; $P < 0.001$), with seven of the ten jackknife trials correctly predicting localities excluded from the training dataset. Based on these results, I assumed that final models built with all localities available represented fair approximations of the species' ecological niches given the variables selected and the data at hand.

Tests of Ecological Differentiation and Visualization of niches

Three pairs of sister taxa—*Calocitta formosa* and *C. colliei* (Fig. 4.2A), *Cyanocorax beecheii* and *C. sanblasianus* (Fig. 4.2B), and *C. yncas luxuosus* and *C. y. yncas* (Fig. 4.3A, B)—show niche interpredictivity with high statistical significance ($P < 0.001$; Table 1). In contrast, the remaining two pairs of sister species—*C. heilprini* and *C. affinis* (Fig. 4.4A) and *C. chrysops* and *C. cyanopogon* (Fig. 4.4B)—show low niche interpredictivity. Accuracy of predictions varies from comparison to comparison (Table 4.1). In general, species pairs show relatively low reciprocal predictive accuracy with the exception of *Calocitta* species (50–79% accuracy). Also, strong predictive asymmetry is observed between *C. y. luxuosus* (69%) and *C. y. yncas* (14%).

Visualization of species occurrence points in DFA space shows that niche differentiation (indicated by means of low-interpredictivity) is a real phenomenon, rather of an artifact of niche availability. Environmental conditions available to *C. heilprini* are contained within the environmental conditions available to *C. affinis* (Fig. 4.5A) and still,

C. heilprini separates from *C. affinis* in DFA space. The same phenomenon is observed between *C. cyanopogon* and *C. chrysops* (Fig. 4.5B).

Ecological Niche Distances

Distance matrices based on niche centroids show that most species are more similar to species other than their sister species (Fig. 4.6A; Appendix II). For example, among the four sister species pairs included in the analysis, only *Cyanocorax sanblasianus* and *C. beecheii* are placed as most similar to each other in overall ecological dimensions. Also, no significant correlation is found between ecological and molecular distances (Mantel Test; $R^2 = –0.130$, $P = 0.802$). However, when mean centroid distances are applied to the branching patterns of the phylogeny (Fig. 4.6B), shared ecological characteristics among lineages are recovered as internodal branches of various lengths. A relatively long internodal branch is observed for Group II + *C. dickeyi*, indicating that common ecological characteristics are shared by the clade; shared trends are detected also for Group II, albeit being represented by a very short branch. Relatively long branches are recovered for *C. sanblasianus* + *C. becheii* and *C. heilprini* + *C. affinis*, although independent divergence is observed for all species, especially for *C. heilprini*. No internodal branches are recovered for *C. chrysops* + *C. cyanopogon*, *C. y. yncas* + *C. y. luxuosus*, *C. yucatanicus* + (*C. beecheii*, *C. sanblasianus*) or for "*Cissilopha*" as a whole. Note that because centroid distances maximize ecological differences among species and do not incorporate data dispersion, they cannot detect the effect of "nested" niches (e.g., the niche of Species A being nested within the niche of Species B).

Phylogenetic Inertia

Phylogenetic reconstructions provide a formal analytical perspective on the evolution of ecological niches in the whole *Cyanocorax* assemblage. In seven of eight cases, Model A (random-walk) can not be rejected as a better fit to the data (Table 4.2). No phylogenetic signal is detected for the niche centroid, as well as for three variables (i.e., precipitation of the wettest month, maximum temperature of the warmest month, and minimum temperature of the coldest month; all $\lambda = 0$; Model A, LRT $P > 0.05$). For annual precipitation and mean diurnal range, strong phylogenetic signal is indicated, and for annual mean temperature, the signal is weak but not significantly different from 1 (LRT; $P = 0.214$). Analysis of precipitation of the driest month rejects Model A favor of Model B ($P < 0.05$), indicating a decreasing trend ($\beta < 0$) on the trait. Furthermore, the evolution of this trait is highly correlated with phylogenetic structure ($\lambda = 1$), indicating high phylogenetic inertia in the overall decrease in precipitation of the driest month across the tree. Finally, for most environmental variables the κ and δ parameters indicated a gradual mode of evolution ($\kappa > 0$), mostly driven by species-specific divergence ($\delta > 1$).

DISCUSSION

Ecological Niche Differentiation

Comparisons of ecological niches the geographic dimension provided a preliminary perspective on the evolution of ecological niches among sister species. Interpredictivity tests were significant in three pairs of sister species (*Calocitta formosa* and *C. colliei*, *Cyanocorax beecheii* and *C. sanblasianus*, and *C. yncas luxuosus* and *C. y. yncas*), fitting the overall expectation of ecological niche conservatism at least at this

scale of analysis. However, accuracy of interpredictions (i.e., percentage of occurrence points predicted correctly) for these species pairs reveals some degree of niche differentiation, especially for *C. beecheii* and *C. sanblasianus*, and *C. yncas luxuosus* and *C. y. yncas*. In these cases, more rigorous comparisons (e.g., Turelli et al., in press) may be more appropriate for assessing significant shifts in ecological niches.

Relatively low differentiation between *Calocitta formosa* and *C. colliei* suggest that in spite of their morphological and genetic differences (5.5% ND2 p-distance), *Calocitta* species show high niche conservatism. Interestingly, no physical barrier separates these species, which overlap marginally in Jalisco (Fig. 4.2A). *Cyanocorax beecheii* and *C. sanblasianus* overlap around the same region as *Calocitta* species, but in western Nayarit (Fig. 4.2B). In contrast, however, *C. beecheii* and *C. sanblasianus* show at least some degree of niche differentiation measured as accuracy of niche prediction.

Both, niche conservatism and limited geographic overlap among *Calocitta* species, suggest scenarios of allopatric speciation followed by secondary contact (Barraclough et al., 1998, 1999). On the other hand, the observed degree of ecological differentiation between *C. beecheii* and *C. sanblasianus* deserves further investigation, especially given the relatively recent divergence of these two lineages (Chapter 3). Detailed study of contact areas, as well as study of other co-distributed populations across the Pacific slopes of Mexico, may prove useful in understanding biogeographic patterns and uncovering speciation mechanisms in the region.

Interpredictivity tests were significant also for the geographic isolates of *C. yncas*, although asymmetry in predictive accuracy may imply that *C. y. yncas* is occupying a portion of the ecological niche of *C. luxuosus* (Fig. 4.3A, B; Table 4.1). Intriguingly, the

ecological niche of *C. y. luxuosus* extends south to central Panama, but the southern limit of the range of this population is northern Honduras. Unfortunately, it is difficult to establish whether the absence of *C. y. luxuosus* south of northern Honduras results from local extinction or recent changes in niche availability.

The remaining two pairs of sister species show signs of ecological niche evolution, independent of their allopatric (*Cyanocorax affinis* and *C. heilprini*) or marginally sympatric (*C. chrysops* and *C. cyanopogon*) distributions. However, because the biogeographic and ecological patterns in these species are highly complex, the potential role of ecological diversification in the original lineage splitting must remain an open question. *Cyanocorax affinis* and *C. heilprini* are separated (physically and ecologically) by the highlands of the Andes and the Cordillera de la Costa of Venezuela (Fig. 4.4A). Although there seems to be a narrow "ecological corridor" connecting the distribution of these two species, such a pattern could result from the low spatial resolution used in the present analysis.

As per *Cyanocorax chrysops* and *C. cyanopogon*, the relatively large area of niche overlap between both species (Fig. 4.4B) contrasts with their narrow area of sympatry. Given that no physical barrier or sharp ecological gradient separates these species, the observed pattern seems to indicate a complex biogeographic history not reflected by the current distribution and ecological conditions (e.g., Losos and Glor, 2003; Lynch, 1989). Although the fact that no hybrids have been observed in the contact zone might support a scenario of parapatric speciation, the same situation is observed in *Calocitta* spp., in which stasis in ecological niches and current distributions suggest an allopatric mode of speciation. Clearly, further ecological, behavioral, and genetic studies are crucial in

understanding the origin of these patterns, as well as the maintenance of species identity across the areas of contact.

Ecological Distances and Phylogenetic Inertia

Visualization of ecological niches based on ecological distances (Fig. 4.6A) indicates that *Cyanocorax* species tend to be more similar to non-related species in other clades in ecological space; also, no significant relation exists between ecological and molecular distances (Mantel test). Such phenomena have been reported previously for *Aphelocoma* jays (Rice et al., 2003) and *Anolis* lizards (Knouft et al., 2006). However, when ecological distances are examined in the context of the phylogeny (Fig. 4.6B), it is clear that certain clades (e.g., Group II + *C. dickeyi*) do share general ecological characteristics, whereas others (e.g., *"Cissilopha"*) do not. Comparing these distance-based (Fig. 4.6A) and "pseudo-phylogenetic" (Fig. 4.6B) perspectives illustrates why evolution of ecological niches above the species level needs to be interpreted in a phylogenetic framework. Although the non-directional approach advocated by Knouft et al. (2006) and Lovette and Hochachka (2006) is appropriate for detecting changes at the species level and for exploring general evolutionary trends, the use of distance ignores the importance of shared ancestral characters.

Formal phylogenetic reconstruction of ecological niches can be achieved by methods that do not relay on explicit, unequivocal reconstruction ancestral character states at each node (Pagel, 1997). The GLS approach used herein indicates that overall, ecological niches in *Cyanocorax* are not subject to phylogenetic inertia. However, when niches are partitioned into different components, strong phylogenetic inertia is detected in

some components. The clearest case is precipitation of the driest month, in which I found not only strong phylogenetic inertia, but also a statistically significant decreasing trend across the phylogeny. With few exceptions, both conserved and labile characteristics are evolving in a gradual fashion that is biased toward species-specific changes, rather than adaptative radiation.

These results indicate that species-specific niche evolution may result from the combination of both conserved and labile ecological characteristics. For example, species may be able to shift their distributions toward areas with ecological combinations bearing the same values for the conserved characteristic, but different values for the more labile ones. Dynamic interaction among conserved and labile ecological traits could promote: (1) ecological segregation of populations into contiguous areas bearing different ecological combinations; (2) successful dispersal and establishment of populations into non-contiguous areas bearing different ecological combinations; or (3) adaptation of populations to changing environmental conditions (e.g., changes promoted by Pleistocene-related climate events). More detailed studies, especially those focused on species diverging rapidly from each other in ecological dimensions (e.g., "*Cissilopha*" jays) are promising in understanding evolution of ecological differentiation. Also, population-level genetic sampling of relatively recent lineages may prove informative in detecting the signatures of adaptation by natural selection.

Conclusions, limitations, and prospects

Analyses of the evolution of ecological niches in *Cyanocorax* jays indicate both ecological niche conservatism and ecological niche diversification at different

phylogenetic and analytical levels; this result coincides with previous studies focusing on analysis of niche evolution among closely related species (e.g., Knouft et al., 2006; Lovette and Hochachka, 2006). Whereas comparisons among sister species provided preliminary information for exploring potential speciation mechanisms in *Cyanocorax,* both mapping of niche distances and phylogenetic reconstructions of niche characteristics provided a perspective on the evolutionary dimensions of ecological trends in the group.

Certainly, several limitations are implied in the analyses presented. First, among other burdens of character reconstruction methods (e.g., Schluter, 1997), the inability to incorporate data dispersion is among the most problematic when reconstructing ecological data, because only punctual estimates are used for highly variable characters. Second, the present approach does not account for phylogenetic uncertainty on the phylogenetic reconstruction (Schultz and Churchill, 1999). However, in the present case, analyses of the alternate hypothesis of relationships (*C. mystacalis* + *C. yncas* assigning a branch length = 0; not shown) reconstructed the same general evolutionary trends. Third, comparisons of ecological niches can be performed using more rigorous statistical methods (Turelli et al., in press). Last, as in other studies involving ecological niche modeling, this study does not take into account all environmental variables that may be important in determining the ecological niches of species, and it does not incorporate non-environmental factors (e.g., biotic interactions) that may affect the realized niche of the species, from which occurrence data are derived in the first place.

Regardless of the analytical limitations, general trends recovered in this study will be useful for generating specific hypothesis that can be tested through subsequent detailed ecological and population genetic analyses. Also, an advantage of the present

study in contrast to previous contributions (e.g., Graham et al., 2004; Knouft et al., 2006; Kozak and Wiens, 2006) stands on the fact that, because complete sampling is available for the *Cyanocorax* assemblage, species' pairs analyzed are most likely to be each other's closest relatives. Resolution of phylogenetic relationships among the remaining species in *Cyanocorax* (Clade B) will offer interesting possibilities in terms of understanding evolutionary trends in the ecological niches of species within the overall assemblage.

CONCLUSIONS

The present study provides novel phylogenetic information on the relationships among New World jay genera, as well as the first molecular systematic treatments for the genera *Cyanolyca* and *Cyanocorax*. The molecular phylogeny of the American jays agrees with morphological synapomorphies supporting the monophyly of the NWJs and the early divergence of *Cyanolyca*. The general structure of the phylogenetic tree indicates a deep separation of the remaining species into two clades, one containing *Aphelocoma*, *Cyanocitta*, and *Gymnorhinus*, and another containing *Cyanocorax*, *Calocitta*, and *Psilorhinus*.

Phylogenetic analyses of *Cyanolyca* offered a highly robust hypothesis of relationships among species in the genus, as well as a general overview of genetic differentiation among some populations. Also, the phylogeny of *Cyanolyca* highlighted the importance of allopatric speciation along the Neotropical montane forests. Although this result is concordant with theoretical expectations, the present study constitutes the first complete sampling and multi-gene analyses showing evolutionary patterns that are highly consistent with an allopatric mode of speciation in this largely fragmented region.

Analyses of phylogenetic relationships among *Cyanocorax, Psilorhinus,* and *Calocitta,* revealed consistent patterns indicating overall monophyly of the group, but paraphyly of the current *Cyanocorax*. Also, division of ingroup taxa into two reciprocally monophyletic groups is highly supported, although relationships within those clades are not completely resolved. The results suggest that retention of ancestral polymorphisms, incomplete lineage sorting, and local extinction of broadly distributed

ancestors, likely have contributed to the complex biogeographic patterns observed. Also, the hypothesis of relationships recovered herein implies several complications for the taxonomy of this group. Although high diversity in morphology indicates that multiple taxonomic arrangements are possible, recognizing a broader *Cyanocorax*, including *Calocitta* and *Psilorhinus*, would maximize stability of a systematic classification consistent with phylogeny.

Analyses of the evolution of ecological niches in *Cyanocorax* jays indicated both ecological niche conservatism and ecological niche diversification at different phylogenetic and analytical levels. Whereas comparisons among sister species provided preliminary information for exploring potential speciation mechanisms in *Cyanocorax*, mapping of niche distances and phylogenetic reconstructions of niche characteristics provided a perspective on the evolutionary dimensions of ecological trends in the group. These results will be useful for generating specific hypothesis that can be tested through subsequent detailed ecological and population genetic analyses.

Finally, this study contributes significantly to the understanding of systematics, evolution, biogeography, and ecology of two poorly known Neotropical lineages. The robust phylogeny of *Cyanolyca* provides an evolutionary framework for further ecological and behavioral studies, as well as solid ground for more detailed phylogeographic analyses. The phylogeny and ecological niche analysis of *Cyanocorax* revealed complex evolutionary patterns that need to be scrutinized more rigorously from the genetic and ecological points of view. In a broader scale, these results contribute to our understanding of patterns and processes involved in the evolution of the New World jay lineage and Neotropical birds in general.

LITERATURE CITED

Abdi, H. 2007. Z-scores. In N.J. Salkind Ed., Encyclopedia of Measurement and Statistics. Thousand Oaks, California.

Alvarez, H. 1975. The social system of the Green Jay in Colombia. Living Bird 14:15–43.

Amadon, D. 1944. The genera of Corvidae and their relationships. American Museum Novitates 1251:1–21.

Amaral, M. F. and R. H. F. Macedo. 2003. Breeding patterns and habitat use in the endemic Curl-crested Jay of central Brazil. Journal of Field Ornithology 74:331–340.

Anderson, R. P., M. Gómez-Laverde, and A. T. Peterson. 2002. Geographical distributions of spiny pocket mice in South America: Insights from predictive models. Global Ecology and Biogeography 11:131–141.

Anderson, R. P., D. Lew, and A. T. Peterson. 2003. Evaluating predictive models of species' distributions: Criteria for selecting optimal models. Ecological Modelling 162:211–232.

A.O.U. 1983. Check-list of North American birds, 6th ed. [with supplements through 1993]. American Ornithologists Union, Washington DC.

AOU. 1998. Check-list of North American birds, 7th ed. American Ornithologists' Union, Washington, D.C.

Arbogast, B. S., S. V. Drovetski, R. L. Curry, P. T. Boag, G. Seutin, P. R. Grant, B. R. Grant, and D. J. Anderson. 2006. The origin and diversification of Galapagos mockingbirds. Evolution 60:370–382.

Avise, J. C. and R. M. Ball. 1990. Principles of genealogical concordance in species concepts and biological taxonomy. Oxford Surveys in Evolutionary Biology 7:45–67.

Avise, J. C. 1994. Molecular markers, natural history, and evolution. Chapman and Hall, New York.

Baker, F.K. and F. M. Lutzoni. 2002. The utility of the incongruence length difference test. Systematic Biology 51:625–637.

Barhoum, D. N. and K. J. Burns. 2002. Phylogenetic relationships of the wrentit based on mitochondrial cytochrome b sequences. Condor 104:740–749.

Balda, R. P. and G. C. Bateman. 1971. Flocking and annual cycle of the Piñon Jay, *Gymnorhinus cyanocephalus*. Condor 73:287–302.

Barraclough, T. G., A. P. Vogler, and P. H. Harvey. 1998. Revealing the factors that promote speciation. Philosophical Transactions of the Royal Society of London 353:241–249.

Barraclough, T. G., A. P. Vogler, and P. H. Harvey. 1999. Revealing the factors that promote speciation. Pages 202–219 in A. E. Magurran and R. M. May, Eds. The evolution of biological diversity. Oxford University Press, Oxford.

Barraclough, T. G., and V. Savolainen. 2001. Evolutionary rates and species diversity in flowering plants. Evolution 55:677–683.

Barret, M., M. J. Donohue, and E. Sober. 1991. Against consensus. Systematic Zoology 40:486–493.

Bates, J. M. and R. M. Zink. 1994. Evolution into the Andes: Molecular evidence for species relationships in the genus *Leptopogon*. Auk 111:507–515.

Benz, B. W., M. B. Robbins, and A. T. Peterson. 2006. Evolutionary history of woodpeckers and allies (Aves: Picidae): Placing key taxa on the phylogenetic tree. Molecular Phylogenetics and Evolution 40:389–399.

Bensasson, S., Z. De-Xing, D. L. Hartl, and G. H. Hewitt. 2001. Mitochondrial pseudogenes: Evolution's misplaced witnesses. Trends in Ecology and Evolution 16:314–321.

Blake, E. R. and C. Vaurie. 1962. Family Corvidae. Pages 204–282 *in* Check-list of birds of the World, Vol. 15 (E. Mayr and J. C. Greenway, Jr., Eds.). Museum of Comparative Zoology, Cambridge, Massachusetts.

Bonaccorso, E. and A. T. Peterson. 2007. A multilocus phylogeny of New World Jay genera. Molecular Phylogenetics and Evolution 42:467–476.

Bond, J. 1956. Additional notes on Peruvian birds II. Proceedings of the Academy of Natural Sciences of Philadelphia 108: 227–247.

Brodkorb, P. 1957. New passerine birds from the Pleistocene of Reddick, Florida. Journal of Paleontology 31:129–138.

Brodkorb, P. 1972. Neogene fossil jays from the Great Pains. Condor. 74:347–349.

Brommham, L., A. Eyre-Walker, N. H. Smith, and J. M. Smith. 2003. Mitochondrial Steve: Parental inheritance of mitochondrial DNA in humans. Trends in Ecology and Evolution 18:2–4.

Brumfield, R. T., R. Jernigan, D. B. McDonald, and M. J. Braun. 2001. Evolutionary implications of divergent clines in an avian (*Manacus:* Aves) hybrid zone. Evolution 55:2070–2087.

Brown, J.L. 1963. Social organization and behavior of the Mexican Jay. Condor 65:126–153.

Brown, J. L. 1994 Mexican jay (*Aphelocoma ultramarina*). In The Birds of North America, No. 118 (ed. A. Poole and F. Gill). The Academy of Natural Sciences and The American Ornithologists' Union, Philadelphia, PA & Washington, D.C.

Brumfield, R. T., S. V. Edwards. 2007. Evolution into and out of the Andes: A Bayesian analysis of historical diversification in *Thamnophilus* antshrikes. Evolution 61:346–367.

Buckey, T. R., C. Simon, H. Shimodaira, andG. K. Chambers. Evaluation hypotheses on the origin and evolution of New Zealand alpine cicadas (Maoricicada) using multiple-comparison tests of tree topology. Molecular Biology and Evolution 18:223–234.

Bull, J.J., J. P. Huelsenbeck, C. W. Cunninham, D. L. Swofford, and P. J. Wandell. 1993. Partitioning and combining data in phylogenetic analysis. Systematic Biology 42:384–497.

Burns, K. J. 1998. Molecular phylogenetics of the genus *Piranga*: Implications for biogeography and the evolution of morphology. Auk 115:621–634.

Cadena, C.D., J. Klicka, and R. E. Ricklefs. *in press*. Evolutionary differentiation in the Neotropical montane region: Molecular phylogenetics and phylogeography of *Buarremon* brush-finches (Aves, Emberizinae). Molecular Phylogenetics and Evolution.

Carpenter, G., A. N. Gillison, and J. Winter. 1993. DOMAIN: A flexible modeling procedure for mapping potential distributions of animals and plants. Biodiversity and Conservation 2:667–680.

Carstens, B. C. and C. L. Richards. 2007. Integrating coalescent and ecological niche modeling in comparative phylogeography. Evolution 61:1439–1454.

Chapman, F.M., 1917. The distribution of bird life in Colombia. Bulletin of the American Museum of Natural History 36–169.

Chapman, F.M., 1926. The distribution of bird life in Ecuador. Bulletin of the American Museum of Natural History 55:1–784.

Chesser, R. T. and R. M. Zink. 1994. Modes of speciation in birds: A test of Lynch's method. Evolution 48:490–497.

Churchill S. P., Balslev H., Forero E., Luteyn J. L (Eds.). 1995. Biodiversity and conservation of Neotropical montane forests. Pages xi–xiv *in* Proceedings of the Neotropical Montane Forest Biodiversity and Conservation Symposium. The New York Botanical Garden, Bronx, New York.

Cicero, C. and N. K. Johnson. 2001. Higher level phylogeny of vireos (Aves: Vireonidae) based on sequences of multiple mtDNA genes. Molecular Phylogenetics and Evolution 20:27–40.

Coates, A.G. and J. A. Obando. 1996. The geologic evolution of the Central American Isthmus. Pages 21–56 *in* Evolution and Environment in Tropical America (J. Jackson, A. F. Budd, and A. G. Coates Eds.). University of Chicago Press, Chicago, Illinois.

Cracraft, J. 1983. Species concepts and speciation analysis. Current Ornithology 1:189-187.

Cracraft, J. 1985. Historical biogeography and patterns of differentiation within the South American avifauna: Areas of endemism. Ornithological Monographs 36:49–84.

Crisp, M. D. and L. G. Cook. 2005. Do early branching lineages signify ancestral traits? Trends in Ecology and Evolution 13:361–366.

Crossin, R. S. 1967. The breeding biology of the Tufted Jay. Proceedings of the Western Foundation of Vertebrate Zoology 1:265–299.

Cunningham, C. W., K. E. Omland, and T. H. Oakley. 1998. Reconstructing ancestral character states: A critical reappraisal. Trends in Ecology and Evolution 13:361–366.

Curtis, E. L. and R. C. Miller. 1938. The sclerotic ring in North American birds. Auk 55:225–243.

Darlu, P. and G. Lecointre. 2002. When Does the Incongruence Length Difference Test Fail? Molecular Biology and Evolution 19:432–437.

de Queiroz, A. 1993. For consensus (sometimes). Systematic Biology 42:368–372.

de Queiroz, A., M. J. Donoghue, and K. Junhyong. 1995. Separate versus combined analysis of phylogenetic evidence. Annual Review of Ecology and Systematics 26:657–681.

Desjardins, P. and R. Morais. 1990. Sequence and gene organization of the chicken mitochondrial genome. Journal of Molecular Biology 212:599–634.

Dickinson, E.C. 2003. The Howard & Moore Complete Checklist of the Birds of the World, 3d ed. Princeton Univ. Press, Princeton, New Jersey.

Edwards, S. V., S. B. Kingan, J. D. Calkins, C. N. Balakrishnan, W. B. Jennings, W. J. Swanson, and M. D. Sorenson. (2005). Speciation in birds: Genes, geography and sexual selection. Proceedings of the National Academy of Sciences USA 102:6550–6557.

Efron, B. 1985. Bootstrap confidence intervals for a class of parametric problems. Biometrika 73:45–58.

Elith, J. and M. Burgman. 2002. Predictions and their validation: Rare plants in the central highlands, Victoria. Pp. 303–314. In: Scott, J. M., P. J. Heglund, M. L. Morrison. Eds. Predicting species occurrences: Issues of scale and accuracy. Island Press, Washington D. C.

Elith, J., C. H. Graham, R. P. Anderson, M. Dudı´k, S. Ferrier, A. Guisan, R. J. Hijmans, F. Huettmann, J. R. Leathwick, A. Lehmann, J. Li, L. G. Lohmann, B. A. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J. McC. Overton, A. T. Peterson, S. J. Phillips, K. Richardson, R. Scachetti-Pereira, R. E. Schapire, J. Soberón, S. Williams, M. S. Wisz, and N. E. Zimmermann. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29:129–151.

Endler, J. A.1982. Pleistocene forest refuges: Fact or fancy. Pp. 641–657 in G. T. Prance, ed. Biological Diversification in the Tropics. Columbia Univ. Press, New York.

Ericson, P. G. P., A. L. Jansén, U. S. Johansson, and J. Ekman. 2005. Inter-generic relationships of the crows, jays, magpies and allied groups (Aves: Corvidae) based on nucleotide sequence data. Journal of Avian Biology 36:222–234.

Espinosa de los Monteros, A. and J. Cracraft.1997. Intergeneric relationships of the New World Jays inferred from cytochrome b gene sequences. Condor 99:490–502.

ESRI, 1998. ArcView GIS, version 3.1. Environmental Systems Research Institute Inc., Redlands, California.

Excoffier, L. G. Laval, and S. Schneider. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1:47-50.

Farris, J. S., M. Källersjö, Kluge, A. G., and C. Bult. 1994. Testing significance of congruence. Cladistics 10:315–319.

Felsenstein, J. 1981. Evolutionary trees from DNA sequences: A maximum likelihood approach. Journal of Molecular Evolution 17:368–376.

Felsenstein, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39:783–791.

Felsenstein, J. 1985. Phylogenies and the comparative method. American Naturalist 125:1–15.

Felsenstein, J. 1988. Phylogenies from molecular sequences: Inference and reliability. Annual Review of Genetics 22:521–565.

Ferrusquía-Villafranca, I. 1993. Geology of Mexico: A synopsis. Pages 3–107 *in* Biological Diversity of Mexico: Origins and Distribution (T. P. Ramamoorthy, R. Bye, A. Lot, J. Fa, Eds.). Oxford University Press, Oxford, UK.

Fjeldså, J., and N. Krabbe. 1990. Birds of the High Andes. Zoological Museum, University of Copenhagen, Denmark.

Fitzpatrick, J.W. and G. E. Woolfenden. 1985. The Florida Scrub Jay: Demography of a Cooperative-breeding Bird. Princeton University Press, Princeton.

Fleischer, R. C., C. E. McIntosh, and C. L. Tarr. 1998. Evolution on a volcanic conveyor belt: Using phylogeographic reconstructions and K-Ar-based ages of the Hawaiian Islands to estimate molecular evolutionary rates. Molecular Ecology 7:533–545.

Fleischer, R.C., S. Olson, H. F. James, and A. C. Cooper. 2000. The identity of the extinct Hawaiian eagle (*Haliaeetus*) as determined by mitochondrial DNA sequence. Auk 117:1051–1056.

Fleischer, R.C., C. L. Tarr, H. F. James, B. Slikas, and C. E. McIntosh. 2001. Phylogenetic placement of the po`o-uli (*Melamprosops phaeosoma*) based on mitochondrial DNA sequence and 466 osteological characters. Studies in Avian Biology 22:98–103.

Fok, K. W., C. M. Wade, and D. T. Parkin. 2002. Inferring the phylogeny of disjunct populations of the azure-winged magpie *Cyanopica cyanus* from mitochondrial control region sequences. Proceedings of the Royal Society of London B 269:1671–1679.

Funk, D. J. and K. E. Omland. 2003. Species-level paraphyly and polyphyly: Frequency, causes, and consequences, with insights from animal mitochondrial DNA. Annual Review of Ecology, Evolution, and Systematics 34:397–423.

García-Moreno, J., P. Arctander, and J. Fjeldså. 1999a. Strong diversification at the treeline among *Metallura* hummingbirds. Auk 116:702–711.

García-Moreno, J., P. Arctander, and J. Fjeldså. 1999b. A case of rapid diversification in the Neotropics: Phylogenetic relationships among *Cranioleuca* Spinetails (Aves, Furnaridae). Molecular Phylogenetics and Evolution 12:273–281.

García-Moreno, J., and J. Fjeldså. 2000. Chronology and mode of speciation in the Andean avifauna. Bonner Zoologische Monographien 46:25–46.

García-Moreno, J., A. G. Navarro-Sigüenza, A. T. Peterson, and L. A. Sánchez-González. 2004. Genetic variation coincides with geographic structure in the common bush-tanager (*Chlorospingus ophthalmicus*). Molecular Phylogenetics and Evolution 33:186–196.

García-Moreno, J., N. Cortés, G. M. García-Deras, and B. A. Hernández- Baños. 2006. Local origin and diversification among *Lampornis* hummingbirds: A Mesoamerican taxon. Molecular Phylogenetics and Evolution 38:488–498.

Gayou, D. C. 1986. The social system of the Texas Green Jay. Auk 103:540–547.

Gene Codes Corporation, 2000. Sequencher Version 4.1. Gene Codes Corporation, Ann Arbor, Michigan.

Gerwin, J. A. and R. M. Zink. 1989. Phylogenetic patterns in the genus *Heliodoxa* (Aves; Trochilidae): An allozymic perspective. Willson Bulletin 101:525–554.

Goldman, N., J. P. Anderson, and A. G. Rodrigo. 2000. Likelihood-based tests of topologies in phylogenetics. Systematic Biology 49:652–670.

Goodwin, D. 1976. Crows of the World. Cornell University Press, Ithaca, New York.

Graham, C. H., S. R. Ron, J. C. Santos, C. J. Schneider, and C. Moritz. 2004. Integrating phylogenetics and environmental niche models to explore specieation mechanisms in dendrobatid frogs. Evolution 58:1781–1793.

Graves, G.R. 1985. Elevational correlates of speciation and intraspecific geographic variation in plumage in Andean forest birds. Auk 102:556–579.

Graves, G.R. 1988. Linearity of geographic range and its possible effect on the population structure of Andean birds. Auk 105:47–52.

Gregory-Wodzicki, K. M. 2000. Uplift history of Central and Northern Andes: A review. Geological Society of America Bulletin 112:1091–1105.

Haemig, P. D. 1979. Secret of the Painted Jay. Biotropica 11:81–87.

Hackett, S. J., and K. V. Rosenberg. 1990. Evolution of South American antwrens (Formicariidae): Comparison of phenotypic and genetic differentiation. Auk 107:473–489.

Hackett, S. J. 1995. Molecular systematics and zoogeography of flowerpiercers in the *Diglossa baritula* complex. Auk 112:156–170.

Haffer, J. 1974. Avian speciation in tropical South America. Publications of the Nuttall Ornithological Club No. 14:1–390.

Hardy, J. W. 1961. Studies in behavior and phylogeny of certain New World Jays (Garrulinae). University of Kansas Science Bulletin 42:13–149.

Hardy, J. W. 1973. Age and sex differences in the black-and-blue jays of Middle America. Bird Banding 44:81–90.

Hardy, J. W. 1974. Behavior and its evolution in Neotropical jays (*Cissilopha*) Bird Banding 45:253–268.

Hardy, J. W. 1979. Vocal repertoire and its possible evolution in the black and blue Jays (*Cissilopha*). Willson Bulletin 91:187–201.

Harrison R. G. 1990. Hybrid zones: Windows on evolutionary process. Oxford Surveys in Evolutionary Biology 7:69–128.

Hellmayr, C. E. 1934. Catalogue of the Birds of the Americas and the Adjacent Islands. Field Museum of Natural History, Chicago.

Helm-Bychowski, K. and J. Cracraft. 1993. Recovering phylogenetic signal from DNA sequences: relationships within the corvine assemblage (Class Aves) as inferred from complete sequences of the mitochondrial DNA cytochrome-b gene. Molecular Biology and Evolution 10:1196–1214.

Hernández-Baños, B., A. T. Peterson, A.G. Navarro-Sigüenza, and B. P. Escalante-Pliego. 1995. Bird faunas of the humid montane forests of Mesoamerica: Biogeographic patterns and priorities for conservation. Bird Conservation International 5:251–277.

Hillis, D. M. and J. J. Bull. 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Systematic Biology 42:182–192.

Hillis, D. M., B. K. Mable, and C. Moritz. 1996. Applications of molecular systematics: The state of the field and a look to the future. Pages 515–543 in Molecular systematics (D. M. Hillis, C. Moritz, and B. K. Mable, eds.). Sinauer, Sunderland, Massachusetts.

Hilty, S. L. 2003. Birds of Venezuela. Princeton University Press, Princeton, New Jersey.

Hoarau G., S. Holla, R. Lescasse, W. T. Stam, and J. L. Olsen. 2003. Heteroplasmy and evidence for recombination in the mitochondrial control region of the flatfish *Platichthys flesus*. Molecular Biology and Evolution 19:2261–2264.

Holder, M. and P. O. Lewis. 2003. Phylogeny estimation: Traditional and Bayesian approaches. Nature Reviews Genetics 4:275–284.

Holman, J. A. 1959. Birds and mammals from the Pleistocene of Willinston, Florida. Bulletin of the Florida State Museum 5:1–24.

Hooghiemstra H, T. Van der Hammen, and A. Cleef. 2000. Evolution of forests in the northern Andes and Amazonian lowlands during the Tertiary and Quaternary *in* Ecology of Neotropical Rainforests (M. Guariguata and G. Kattan, Eds.). Ediciones LUR, Cartago, Costa Rica.

Hope, S. 1989. Phylogeny of the avian family Corvidae. Ph. D. dissertation, City University of New York, New York, New York.

Howell, S. N. G. and S. Webb. 2004. A guide to the Birds of Mexico and Northern Central America. Oxford University Press, New York, New York.

Hudson, R. R. 1990. Gene genealogies and the coalescent process. Oxford Surveys in Evolutionary Biology 7:1–44.

Huelsenbeck, J. P. and D. M. Hillis. 1993. Success of phylogenetic methods in the four-taxon case. Systematic Zoology 40:247–264.

Huelsenbeck, J. P., D. M. Hillis, and R. Nielsen. 1996a. A likelihood-ratio test of monophyly. Systematic Biology 45:546–558.

Huelsenbeck, J. P., D. M. Hillis, and R. Jones. 1996b. Parametric bootstrapping in molecular phylogenetics: Applications and performance (Pp. 19–45). J. D. Ferraris and S. R. Palumbi, Molecular Zoology: Advances, Strategies, and Protocols. Wiley-Liss, New York.

Huelsenbeck, J. P, and K. A. Crandal. 1997. Phylogeny estimation and hypothesis testing using maximum likelihood. Annual Review in Ecology and Systematics 28:437–466.

Hutchinson, G. E. 1959. Homage to Santa Rosalia, or why are there so many kinds of animals? American Naturalist 93:145–159.

Jeffroy O., H. Brinkmann, F. Delsuc, and H. Philippe. 2006. Phylogenomics: The beginning of incongruence? Trends in Genetics 22:225–231.

Johnson, N. K. 2000. Leapfrogging revisited in Andean bird: Geographical variation in the Tody-tyrant superspecies *Poecilotriccus ruficeps* and *P. luluae*. Ibis 144:69–84.

Johnson N. K. and C. Cicero. 2004. New mitochondrial DNA data affirm the importance of Pleistocene speciation in North American birds. Evolution 58:1122–1130.

Klicka J. and R. M. Zink. 1997. The importance of recent ice ages in speciation: A failed paradigm. Science. 277:1666–1669.

Klicka, J., R. M. Zink, J. C. Barlow, W. B. McGillivray, and T. J. Doyle. 1999. Evidence supporting the recent origin and species status of the timberline sparrow. Condor 101:577–588.

Knouft, J. H., J. B. Losos, R. E. Glor, and J. J. Kolbe. 2006. Phylogenetic analysis of the evolution of the niche in lizards of the *Anolis sagrei* group. Ecology 87:S29–38.

Knowles, L. L. and W. P. Maddison. 2002. Statistical phylogeography. Molecular Ecology 11:2623–2635.

Kocher, T. D., W. K. Thomas, A. Meyer, S. V. Edwards, S. Paabo, F. X. Villablanca, A. C. Wilson.1989. Dynamics of mitochondrial DNA evolution in animals:

Amplification and sequencing with conserved primers. Proceedings of the National Academy of Sciences USA 86:6196–6200.

Kozack, K. H. and J. J. Wiens. 2006. Does niche conservatism promote speciation? A case study in North American salamanders. Evolution 60:2604–2621.

Langen, T. A.1996a. The mating system of the white-throated magpie-jay *Calocitta formosa* and Greenwood's hypothesis for sex-biased dispersal. Ibis 138:506–513.

Langen, T. A. 1996b. Skill acquisition and the timing of natal dispersal in the White-throated magpie-jay, *Calocitta formosa*. Animal Behaviour 51:575–588.

Lanyon, S. M. 1988. The stochastic mode of molecular evolution: What consequences for systematic investigations? Auk 105:565–573.

Levine, R. S., A. T. Peterson, and M. Q. Benedict. Distributions of the *Anopheles gambiae* complex modeled using a genetic algorithm, with implications for malaria history in Brazil. American Journal of Tropical Medical and Hygiene 70:105–109.

Lewis, P. O. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. Systematic Biology 50:913–925.

Lockhart, P. J., M. A. Steel, M. D. Hendy, and D. Penny. 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. Molecular Biology and Evolution 11: 605–612.

Loomis, W. F. and D. W. Smith. 1992. Molecular phylogeny of *Dictyostelium discodeum* by protein sequence comparison. Proceedings of the National Academy of Sciences USA 87:9093–9097.

Losos, J. B and R. E. Glor. 2003. Phylogenetic comparative methods and the geography of speciation. Trends in Ecology and Evolution 18:220–227.

Losos J. B., M. Leal, R. E. Glor, K. de Queiroz , P. E. Hertz, L. Rodríguez-Schettino, A. Chamizo-Lara, T. R. Jackman, and A. Larson. 2003. Niche lability in the evolution of a Caribbean lizard community. Nature 424:542–50.

Lynch, J. D. 1989. The gauge of speciation: On the frequency of modes of speciation. Pp. 527–553 in D. Otte and J. A. Endler, eds. Speciation and its consequences. Sinauer Associates, Sunderland, MA.

Maddison, D. R. and W. P. Maddison. 2000. MacClade: Analysis of phylogeny and character evolution. Vers. 4.0. Sinauer, Sunderland, MA.

Maddison, D. R., and W. P. Maddison. 2004a. Batch Architect: Automation of Simulations and Replicated Analyses. A package of modules for Mesquite. Version 1.01. Available from: http://mesquiteproject.org/mesquite/download/download.html

Maddison, W. P. and D. R. Maddison. 2004b. Mesquite: A modular system for evolutionary analysis ver 1.05. Available from: http://mesquiteproject.org/mesquite/download/download.html

Madge, S. and H. Burn. 1994. Crows and jays: A guide to the crows, jays, and magpies of the World. Houghton Mifflin Co., Boston and New York.

Maijer, S. and J. Fjeldså. 1997. Description of a new *Cranioleuca* spinetail from Bolivia and a “leapfrog pattern” of geographic variation in the genus. Ibis 139:606–616.

McKitrick, M. C. and R. M. Zink. 1988. Species concepts in ornithology. Condor 90:1–14.

Martínez-Meyer E. 2002. Evolutionary trends in ecological niches of species. PhD dissertation, University of Kansas, Lawrence, Kansas

Martínez-Meyer, E., A. T. Peterson, and A. G. Navarro-Sigüenza. 2004. Evolution of seasonal ecological niches in the *Passerina* buntings (Aves: Cardinalidae). Proceedings of the Royal Society B 271:1151–1157.

Marzluff, J. M. 1988. Do Pinyon Jays alter nest placement based on prior experience? Animal Behaviour 36:1–10.

McCracken, K. G. and M. D. Sorenson. 2005. Is homoplasy or lineage sorting the source of incongruent mtDNA and nuclear gene trees in the stiff-tailed ducks (*Nomonyx-Oxyura*)? Systematic Biology 54 35–55.

Mayr, E. 1942. Systematics and the Origin of the Species from the Point of View of a Zoologist. Columbia University Press, New York.

Mayr, E. 1963. Avian Species and Evolution. Harvard University Press, Cambridge, Massachusetts.

Meyer de Schauensee, R. 1966. The species of birds of South America and their distribution. Livingston Publishing Co., Narberth, Pennsylvania.

Mindell, D. P., J. W. Sites Jr., and D. Graur. 1990. Mode of allozyme evolution: Increased genetic distance associated with speciation events. Journal of Evolutionary Biology 3:125–131.

Monroe, B. L. 1968. A distributional survey of the birds of Honduras. Ornithological Monographs 7:1–458.

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

Moore, R. T. 1935. A new jay of the genus *Cyanocorax* from Sinaloa, Mexico. Auk 52:274–277.

Moore, W. S. 1995. Inferring phylogenies from mtDNA variation: Mitochondrial-gene trees versus nuclear-gene trees. Evolution 49:718–726.

Navarro-Sigüenza, A. G., A. T. Peterson, and A. Gordillo-Martínez. 2002. A Mexican case study on a centralized database from world natural history museums. CODATA Journal1:45–53.

Navarro-Sigüenza, A. G., A. T. Peterson, and A. Gordillo-Martínez. 2003. Museums working together: The atlas of the birds of Mexico. Bulletin of the British Ornithologists' Club 123A:207–225.

Nepokroeff, M., K. J. Sytsma, W. L. Wagner, and E. A. Zimmer. 2003. Reconstructing ancestral patterns of colonization and dispersal in the Hawaiian understorey tree genus *Psychotria* (Rubiaceae): A comparison of parsimony and likelihood approaches. Systematic Biology 52:820–838.

Omland, K. E. and S. M. Lanyon. 2000. Reconstructing plumage evolution in orioles (*Icterus*): Repeated convergence and reversal in patterns. Evolution 54:2119–2133.

Pagel, M. 1997. Inferring evolutionary processes from phylogenies. Zoologica Scripta 26:331–348.

Pagel, M. 1999. The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. Systematic Biology 48:612–622.

Parker, T. A., T.S. Schulenberg, G. R. Graves, and M. J. Braun. 1985. The avifauna of the Huancabamba region, northern Peru. Ornithological Monographs 36:169–197.

Pearson, R. G., C. J. Raxwhorthy, M. Nakamura, and A. T. Peterson. 2007. Predicting species distributions from small numbers of occurrence records: A test case using cryptic geckos in Madagascar. Journal of Biogeography 34:102–117.

Pérez-Emán, J. 2005. Molecular phylogenetics and biogeography of the Neotropical redstarts (*Myioborus*: Aves, Parulinae). Molecular Phylogenetics and Evolution 37:511–528.

Peterson, A. T. 1991. Sociality and ontogeny of coloration in the blue-and-black jays (*Cyanocorax*). Wilson Bulletin 103:59–67.

Peterson, A. T. 1992a. Philopatry and genetic differentiation in the *Aphelocoma* jays. Biological Journal of the Linnaean Society 47:249–260.

Peterson, A. T. 1992b. Phylogeny and rates of molecular evolution in the jays of the genus *Aphelocoma* (Corvidae). Auk 109:134–148.

Peterson, A. T. 2001. Predicting species' geographic distributions based on ecological niche modeling. Condor 103:599–605.

Peterson, A. T. 2005. Kansas Gap Analysis: The importance of validating distributional models before using them. Southwestern Naturalist 50:230–236.

Peterson, A. T., A. G. Navarro-Sigüenza, and H. Benítez-Díaz.1998. The need for continued scientific collecting: A geographic analysis of Mexican bird specimens. Ibis 140:288–294.

Peterson, A.T. and K. C. Cohoon. 1999. Sensitivity of distributional prediction algorithms to geographic data completeness. Ecological Modelling 117:159–164.

Peterson, A. T., J. Soberón, and V. Sánchez-Cordero. 1999. Conservatism of ecological niches in evolutionary time. Science 285:1265–1267.

Peterson, A. T., D. R. B. Stockwell, and D. A. Kluza. 2002. Distributional prediction based on ecological niche modeling of primary occurrence data. *in* J. M. Scott, P. J. Heglund and M. L. Morrison (eds.), Predicting Species Occurrences: Issues of Scale and Accuracy. Island Press, Washington, D.C.

Peterson, A. T. and D. A. Kluza. 2003. New distributional modeling approaches for gap analysis. Animal Conservation 6:47–54.

Peterson, A. T. and J. J. Shaw. 2003. *Lutzomya* vectors for cutaneous leishmaniasis in southern Brazil: Ecological niche models, predicted geographic distributions, and climate change effects. Internaltional Journal of Parasitology 33: 919–931.

Peterson, A. T., M. Eaton, and M. Papes. 2007. Transferability and model evaluation in ecological niche modeling: A comparison of GARP and MAXENT. Ecography OE:1–10.

Peterson, A. T. and Y. Nakazawa. In press. Environmental data sets matter in ecological niche modeling: An example with *Solenopsis invicta* and *Solenopsis richteri*. Global Ecology and Biogeography.

Phillips, A. R. 1950. The San Blas Jay in Arizona. Condor 52:86.

Pitelka, F. A. 1951a. Speciation and ecological distribution in American jays of the genus *Aphelocoma*. University of California Press, Berkeley and Los Angeles, California.

Pitelka, F. A. 1951b. Central American races of *Cyanolyca mitrata*. Condor 53:97–98.

Pitelka, F. A., R. K. Selander, and M. Alvarez Del Toro. 1956. A hybrid jay from Chiapas, Mexico. Condor 58:98–106.

Posada, D. and K. A. Crandall. 1998. Modeltest: Testing the model of DNA substitution. Bioinformatics 14:817–818.

Posada, D. and K. A. Crandall. 2001. Selecting the best-fit model of nucleotide substitution. Systematic Biology 50:580–601.

Posada, D. and T. R. Buckley. 2004. Model selection and model averaging in phylogenetics: Advantages of akaike information criterion and Bayesian approaches over likelihood ration tests. Systematic Biology 53:793–808.

Prychitko, T. M. and W. S. Moore. 1997. The utility of DNA sequences of an intron from the β-fibrinogen gene in phylogenetic analysis of woodpeckers (Aves: Picidae). Molecular Phylogenetics and Evolution 8:193–204.

Puebla-Olivares, F., E. Bonaccorso, A. Espinosa de los Monteros, K. E. Omland, Jorge E. Llorente-Bousquets, A. T. Peterson, and A. G. Navarro-Sigüenza. *In press*. Speciation in the Emerald Toucanet (*Aulacorhynchus prasinus*) complex. Auk.

Raitt, R. J. and J. W. Hardy. 1976. Behavioral ecology of the Yucatan Jay. Wilson Bulletin 88:529–721.

Raitt, R. J. and J.W. Hardy. 1979. Social behavior, habitat, and food of the Beechey Jay. Wilson Bull. 91:1–15.

Rambaut A. and A. J. Drummond. 2004. Tracer 1.3 [computer program] Available at: http://evolve.zoo.ox.ac.uk/software.html

Remsen, J. V. 1984. High incidence of "leapfrog" pattern of geographic variation in Andean birds: Implications for the speciation process. Science 224:171–173.

Remsen, J.V. and R. T. Brumfield. 1996. Geographic variation and species limits of the *Cinnycerthia* wrens in the Andes. Wilson Bulletin 108:205 227.

Rice, N. H., E. Martínez-Meyer, and A. T. Peterson. 2003. Ecological niche differentiation in the *Aphelocoma* jays: A phylogenetic perspective. Biological Journal of the Linnean Society 80:369–383.

Ridgely, R. S. and G. Tudor. 1994. The birds of South America, vol. 2. University of Texas Press, Austin.

Ridgely, R. S. and P. J. Greenfield. 2001. The birds of Ecuador. Vol. I. Status, distribution, and taxonomy. Cornell University Press, Ithaca, New York.

Ridgway, R. 1904. The birds of North and middle America. Part III. Bulletin of the U. S. National Museum no. 50.

Rode, A. L. and B. S. Lieberman. 2005. Integrating evolution and biogeography: A case study involving Devonian crustaceans. Journal of Paleontology 79:267–276.

Ronquist, F. 1996. DIVA, ver. 1.1. Computer program available at http://www.ebc.uu.se/systzoo/research/diva/diva.html. Uppsala University, Uppsala.

Ronquist, F. 1997. Dispersal-vicariance analysis: A new approach to the quantification of historical biogeography. Systematic Biology 49:195–203.

Ronquist F., and J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.

Sato, A., C. O'huigin, F. Figueroa, P. R. Grant, B., R. Grant, and H. Tichy. 1999. Phylogeny of Darwin's finches as revealed by mtDNA sequences. Proceedings of the National Academy of Sciences USA 96:5101–5106.

Sattler, G. D. and M. J. Braun. 2000. Morphometric variation as an indicator of genetic interactions between Black-capped and Carolina Chickadees at a contact zone in the Appalachian mountains. Auk 117:427–444.

Saunders, M.A. and S. V. Edwards. 2000. Dynamics and phylogenetic implications of mtDNA control region sequences in New World Jays (Aves: Corvidae). Journal of Molecular Evolution 51:97–109.

Schluter, D., T. Price, A. Ø. Mooers, and D. Ludwing. 1997. Likelihood of ancestor states in adaptative radiation. Evolution 51:1699–1711.

Schultz, T. R. and G. A. Churchill. 1999. The role of subjectivity in reconstructing ancestral character states: A Bayesian approach to unknown rates, states, and transformation asymmetries. Systematic Biology 48:651–664.

Shapiro, L. H. and J. P. Dumbacher. 2001. Adenylate kinase intron 5: A new nuclear locus for avian systematics. Auk 118:248–255.

Shi, X., H. Gu, E. Susko, and C. Field. 2005. The comparison of confidence regions in phylogeny. Molecular Biology and Evolution 22:2285–2296.

Shimodaira, H. and M. Hasegawa. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Molecular Biology and Evoluation 16:1114–1116.

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

Sidow, A. and A. C. Wilson. 1990. Compositional statistics: An improvement of evolutionary parsimony and its application to deep branches in the tree of life. Journal of Molecular Evolution 31:51–68.

Simpson, B. B. 1975. Pleistocene changes in the flora of the high tropical Andes. Paleobiology 1:273–294.

Smith, G. R. 1992. Introgression in fishes: Significance for paleontology, cladistics, and evolutionary rates. Systematic Biology 41:41–57.

Smith, S. A., P. R. Sthephens, and J. J. Wiens. 2005. Replicate patterns of species richness, historical biogeography, and phylogeny in Holartic treefrogs. Evolution 59:2433–2450.

Sokal, RR and F. J. Rohlf. 1995. Biometry: the Principles and Practice of Statistics in Biological Research, 3rd edn. W.H. Freeman and Company, New York.

Soltis, P. S., D. E. Soltis, V. Savolalnen, P. R. Crane, and T. G. Barraclough. 2002. Rate heterogeneity among lineages of tracheophytes: Integration of molecular and fossil data and evidence for molecular living fossils. Proceedings of the National Academy of Science USA 99:4430–4435.

SPSS Inc. 2006. SPSS 15.0 for windows evaluation version. SpSS Inc., Chicago, Illinois.

Sorenson, M.D. and T. W. Quinn. 1998. Numts: A challenge for avian systematics and population biology. Auk 115:214–221.

Sorenson, M. D., J. C. Ast, D. E. Dimcheff, T. Yuri, and D. P Mindell. 1999. Primers for a PCR-based approach to mitochondrial genome sequencing in birds and other vertebrates. Molecular Phylogenetics and Evolution 12:105–114.

Sorenson, M. D., C. N. Balakrishnan, and R. B. Payne. 2004. Clade-limited colonization in brood parasitic finches (*Vidua* spp.). Systematic Biology 53:140–153.

Stamatakis, A. T. Ludwig, and H. Meier. 2005. RaxML-III: A fast program for maximum likelihood-based inference of large phylogenetic trees. Bioinformatics 21:456–463.

Stamatakis, A. 2006a. The RAxML-VI-HPC Version 2.2.3 Manual http://icwww.epfl.ch/~stamatak/index-Dateien/software/RAxML-Manual.2.2.3.pdf

Stamatakis, A. 2006b. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688–2690.

Stattersfield, A. J., M. J. Crosby, A. J. Long, and D. C. Wege. 1998. Endemic Bird Areas of the World: Priorities for Biodiversity Conservation. Birdlife International, Cambridge, UK.

Stiles, G. F. and A. F. Skutch. 1989. A Guide to the Birds of Costa Rica. Cornell University Press, Ithaca, New York.

Stockwell, D.R.B. and D. P Peters. 1999. The GARP modelling system: Problems and solutions to automated spatial prediction. International Journal of Geographic Information Systems 13:143–15.

Stockwell, D.R.B. and A. T. Peterson. 2002a. Controlling bias in biodiversity data. Pp. 537–546. In: Scott, J. M., P. J. Heglund, M. L. Morrison. Eds. Predicting species occurrences: Issues of scale and accuracy. Island Press, Washington D. C.

Stockwell, D. R. B. and A. T. Peterson. 2002b. Effects of sample size on accuracy of species distribution models. Ecological Modelling 148:1–13.

Stockwell, D. R. B. and A. T. Peterson. 2003. Comparison of resolution of methods used in mapping biodiversity patterns from point occurrence data. Ecological Indicators 3:213–221.

Sullivan, J., E. Arellano, and D. S. Rogers. Comparative phylogeography of Mesoamerican highland rodents: Concerted versus independent response to past climatic fluctuations. Amer. Nat. 155:755–768.

Sutton, J. M. and P. W. Gilbert. 1947. The Brown Jay's furcular pouch. Condor 44:160–165.

Swofford, D. L. 2000. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Ver 4.0b Sinauer Associates, Sunderland, MA.

Tarr, C. L. 1995. Primers for amplification and determination of mitochondrial control-region sequences in oscine passerines. Molecular Ecology 4:527–529.

Templeton J. J., AC Kamil , and R. P. Balda. 1999. Sociality and social learning in two species of corvids: The pinyon jay (*Gymnorhinus cyanocephalus*) and the Clark's nutcracker (*Nucifraga columbiana*). Journal of Comparative Psychology 113:450–455.

Thompson, J. D., T. J. Gibson, E. Plewniak, E. Jean-Mougin, and D. G. Higgins. 1997. The Clustal X window interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research 24:4876–4882.

Vanzolini, P. E. and E. E. Williams. 1981. The vanishing refuge: A mechanism for ecogeographic speciation. Papéis Avulsos de Zoologia 34:251–255.

Vuilleumier, F. 1969. Systematics and evolution in *Diglossa* (Aves, Coerebidae). American Museum Novitates 2381:1–44.

Wake, D. B. 1991. Homoplasy: The result of natural selection, or evidence of design limitations? American Naturalist 138:543–567.

Wendel, J. F and J. J. Doyle. 1998. Phylogenetic incongruence: A window into genome history and molecular evolution. In: Soltis, D., Soltis P., Doyle J. (Eds.), Molecular Systematics of Plants II: DNA Sequencing. Kluwer Academic, Boston, MA.

Weigel, R. D. 1967. Fossil birds from Miller's Cave, Llano Co., Texas. Texas Journal of Science 19:107–109.

Wiens, J. J. 1998. Combining datasets with different phylogenetic histories. Systematic Biology 47:568–58.

Wiens, J. J. 2004. What is speciation and how should we study it? American Naturalist 163:914–923.

Wiens, J. J., C. H. Graham, D. Moen, S. A. Smith, and T. W. Reeder. 2006. Evolutionary and ecological causes of the latitudinal diversity gradient in hylid frogs: Treeforg trees unearth the roots of high tropical diversity. American Naturalist 168:579–596.

Wiley, E.O. 1978. The evolutionary species concept reconsidered. Systematic Zoology 27:17–26.

Williams, D. A. 2004. Female control of reproductive skew in cooperatively breeding brown jays. Behavioral Ecology and Sociobiology 55:370–380.

Wilgenbusch, J. C., D. L. Warren, and D. L. Swofford. 2004. AWTY: A system for graphical exploration of MCMC convergence in Bayesian phylogenetic inference. http://ceb.csit.fsu.edu/awty.

Williams, D. A. and A. M. Hale. 2006. Helper effects on offspring production in cooperatively breeding Brown Jays. Auk 123:847–857.

Winterstein, S. R. and R. J. Raitt. 1983. Nestling growth and development and the breeding ecology of the Beechey jay. Wilson Bulletin 95:256–268.

Zar, J.H. 1996. Biostatistical Analysis, 3rd ed. Prentice Hall, Upper Saddle River, New Jersey.

Zhang, D.-X. and G. M. Hewitt. 1996. Nuclear integrations: Challenges for mitochondrial DNA markers. Trends in Ecology and Evolution 11:247–251.

Zimmer, J. T. 1953. Studies of Peruvian birds. No. 65. The jays (Corvidae) and the pipits (Motacillidae). American Museum Novitates 1649:1–27.

Zink, R. M. and M. C. McKitrick. 1995. The debate over species concepts and its implications for ornithology 112:701–719.

Zusi, R. L. 1987. A feeding adaptation of the jaw Articulation in New-World Jays (Corvidae). Auk 104:665–680.

Zwickl, D. J. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Ph.D. dissertation, University of Texas, Austin. www.bio.utexas.edu/faculty/antisense/garli/Garli.htm (accessed 1 November 2006)

TABLES

Table 1.1. List of samples and gene sequences (and GenBank Accession Numbers) included in this study. UNAM (Universidad Nacional Autónoma de México); KUNHM (University of Kansas Natural History Museum); FMNH (Field Museum); AMNH (American Museum of Natural History);UWBM: University of Washington, Burke Museum of Natural History and Culture. Tissue numbers and localities correspond only to specimens sequenced herein (DQ genbank series). [a]Espinosa de los Monteros and Cracraft (1997), [b]Cicero and Johnson (2001), [c]Helm-Byshowski and Craft (1993), [d]Ericson *et al.* (2005), [e]Saunders and Edwards (2000). **GenBank Accession numbers pending.

Genus/Species	Tissue #	Country, State	cyt*b*	ND2	AK5	βFb7	TGFβ2.5	CR
Aphelocoma californica	UWBM 78049	USA, Arizona	AY030116[b]	AY030142[b]	DQ912618	DQ912638		AF218919[e]
A. coerulescens	FMNH 396259	USA, Florida	U77335[a]	DQ912601	DQ912619	AY395598[d]	**	AF218918[e]
Calocitta formosa	KUNHM 9352	El Salvador, Usulutan	U77336[a]	DQ912602	DQ912620	DQ912639	**	AF218925[e]
C. colliei	FMNH 343602	Mexico, Sinaloa	DQ912591	DQ912603	*	DQ912640		AF218926[e]
Cyanocitta cristata	KUNHM 4390	USA, Kansas	X74258[c]	DQ912604	DQ912621	DQ912641	**	AF218921[e]
C. stelleri	UWBM 58631	USA, Washington	AY030113[b]	AY030139[b]	*	DQ912642		AF218922[e]
Cyanolyca viridicyanus	AMNH CBF 35	Bolivia, La Paz	U77333[a]	DQ912605	DQ912622	DQ912643	**	AF218933[e]
C. mirabilis	FMNH 343601	Mexico , Guerrero	DQ912592	DQ912606	DQ912623	DQ912644		AF218934[e]
Psilorhinus morio	KUNHM B-1896	Mexico, Campeche	DQ912593	DQ912607	DQ912624	DQ912645	**	AF218927[e]
P. morio	KUNHM B-2169	Mexico, Campeche	DQ912594	DQ912608	DQ912625	–		–
Cyanocorax chrysops	KUNHM 171	Paraguay, Concepción	U77334[a]	DQ912609	DQ912626	DQ912646	**	AF218928[e]
C. yncas	UNAM 15722	Mexico, Queretaro	DQ912595	DQ912610	DQ912627	*	**	AF218930[e]
C. dickeyi	UNAM 15315	Mexico, Sinaloa	DQ912596	DQ912611	DQ912628	DQ912647	**	–
C. melanocyaneus	KUNHM 7657	El Salvador, San Vicente	DQ912597	DQ912612	DQ912629	DQ912648	**	AF218929[e]
C. yucatanicus	UNAM B1661	Mexico, Yucatán	DQ912598	DQ912613	DQ912630	DQ912649	**	–
C. cayanus	KUNHM 5817	Guyana, Barima-Waini	DQ912599	DQ912614	DQ912631	DQ912650	**	–
Gymnorhinus cyanocephalus	FMNH 334283	USA, Nevada	U77332[a]	DQ912615	DQ912632	AY395607[d]		AF218931[e]
G. cyanocephalus	UWBM 77532	USA, New Mexico	AY030115[b]	AY030141[b]	DQ912633	DQ912651	**	AF218932[e]
Corvus brachyrhynchos	KUNHM 6518	USA, Kansas	AY030112[b]	AY030138[b]	DQ912634	DQ912652		AF218937[e]
Perisoreus Canadensis	UWBM 64890	USA, Washington	U77331[a]	DQ912616	DQ912635	DQ912653	**	–
Perisoreus infaustus			–	–	–	–		AF218935[e]
Pica hudsonia	KUNHM 1663	USA, Kansas	AY030114[b]	AY030140[b]	DQ912636	DQ912654	**	–
Pica nuttallii			–	–	–	–		AF218936[e]
Dendrocitta formosae	KUNHM B-6699	China, Hunan	DQ912600	DQ912617	DQ912637	DQ912655	**	–

Table 1.2. Number of aligned, variable/parsimony informative positions, number of trees produced by the maximum parsimony analyses, base frequencies, best-fit models and parameters, –ln likelihood with and without enforcing a molecular clock, and *P* value for rejecting the assumption of a clock-like behavior for all loci analyzed.

Gene	**Alignment positions**		**Trees**	**Base frequencies**				**Model**	**K**	**α**	**p-invar**	**-ln likelihood**		***P***
	Total	Var/Infor		A	C	G	T					no clock	clock	
ND2	1041	457/369	7	0.328	0.349	0.089	0.234	TrN + I + Γ	–	1.133	0.492	7331.8	7348.0	< 0.05
Cyt*b*	999	394/303	41	0.300	0.363	0.127	0.210	GTR + I + Γ	–	1.704	0.551	6358.6	6385.8	< 0.001
CR	1299	544/388	2	0.296	0.253	0.138	0.313	GTR + I + Γ	–	0.9487	0.398	8184.6	8226.5	< 0.001
AK5	602	101/23	2	0.236	0.291	0.301	0.172	HKY + Γ	2.833	1.129	0	1466.5	1476.7	> 0.05
βFb7	871	106/40	2	0.319	0.176	0.188	0.317	GTR + Γ	–	0.346	0	1995.9	2007.7	> 0.05
TGFβ2.5	601	68/18	5	0.240	0.241	0.214	0.305	TrN + Γ	–	0.241	0	1213.9	1221.2	> 0.05

Table 1.3. Consistency indexes and nodal support for major groupings among *Aphelocoma* (A), *Cyanocitta* (C), *Gymnorhinus* (G), *Cyanocorax* (Cx), *Psilorhinus* (P), and *Calocitta* (Ca), obtained when analyzing individual genes using MP, ML, and BA.

Dataset	CI	RC	Analysis	(A, C, G)	(G, (Cx, P, Ca))	(A, C)	(A, G)	(G, C)
			MP	33	—	—	48	—
ND2	0.45	0.24	ML	82	—	64	—	—
			BA	0.99	—	0.81	—	—
			MP	54	—	—	54	—
cyt*b*	0.49	0.27	ML	—	—	—	41	—
			BA	—	—	—	0.65	—
			MP	—	78	60	—	—
CR	0.50	0.26	ML	—	87	62	—	—
			BA	—	0.91	0.86	—	—
βFb7	0.87	0.73	MP	77	—	—	—	86
			ML	83	—	—	—	80
AK5	0.91	0.74	MP	—	—	65	—	—
			ML	—	—	70	—	—
TGFβ2.5	0.94	0.84	MP	44	—	—	22	—
			ML	51	—	—	—	—

Table 1.4. Synapomorphic character state changes among the genera *Aphelocoma* (A), *Cyanocitta* (C), and *Gymnorhinus* (G) supporting the three possible branching patterns drawn from the maximum parsimony 50% majority rule consensus trees. TI = transitions, TV = transversions, DI = domain I, DII = domain II, CD = central domain. Position refers to that on the single loci alignment.

Topology	Locus	Position	Change	Type	CI
A,G	ND2	42	T → C	3rd pos. TI	0.333
		210	G → C	3rd pos. TV	0.429
		327	A → T	3rd pos. TV	0.333
		355	T → C	1st pos. TI	1.000
		426	A → C	3rd pos. TV	0.600
		453	C → T	3rd pos. TI	0.200
		627	T → C	3rd pos. TI	0.400
		855	G → A	3rd pos. TI	0.333
		882	T → A	3rd pos. TV	0.500
		915	C → T	3rd pos. TI	0.500
		993	C → A	3rd pos. TV	0.750
		1020	C → T	3rd pos. TI	0.333
				Mean ± SD	0.48 ± 0.22
	cyt*b*	57	A → T	3rd pos. TV	0.375
		60	C → T	3rd pos. TI	0.250
		81	A → G	3rd pos. TI	0.250
		120	C → A	3rd pos. TV	0.500
		135	C → T	3rd pos. TI	0.333
		255	C → T	3rd pos. TI	0.250
		261	C → A	3rd pos. TV	0.600
		273	A → G	3rd pos. TI	0.200
		324	C → T	3rd pos. TI	0.500
		411	C → T	3rd pos. TI	0.200
		477	C → T	3rd pos. TI	0.200
		597	C → T	3rd pos. TI	0.250
		759	C → T	3rd pos. TI	0.500
		804	T → C	3rd pos. TI	0.333
		921	C → T	3rd pos. TI	1.000
		939	T → C	3rd pos. TI	0.333
		964	A → G	1st pos. TI	0.333
		981	G → A	3rd pos. TI	0.200
				Mean ± SD	0.37 ± 0.20
A,C	CR	78	C → A	D I, TV	0.333
		161	A → T	CD, TV	0.500
		166	C → A	CD, TV	1.000
		358	C → A	CD, TV	1.000
		554	T → A	D II, TV	0.400
		1037	G → A	D II, TI	0.250
		1040	T → C	D II, TI	0.667
		1081	A → G	D II, TI	1.000
		1140	T → C	D II, TI	0.250
		1165	A → T	D II, TV	0.333
		1215	T → C	D II, TI	0.333
		1238	A → G	D II, TI	0.250
		1279	C → T	D II, TI	0.333
				Mean ± SD	0.51 ± 0.30
	AK5	398	C → G	TV	1.000
C,G	βfb7	16	T → G	TV	1.000
		249	T → C	TI	1.000
		602	A → G	TI	1.000

Table 2.1. List of tissue samples and GenBank accession numbers for sequences of *Cyanolyca* included in the present study. Acronyms: AMNH, American Museum of Natural History; FMNH, Field Museum of Natural History; LSUMZ, Louisiana State University Museum of Natural Science; QCAZ, Museo de Zoología, Pontificia Universidad Católica del Ecuador, Quito; EBGR, Museo Estación Biológica Rancho Grande, Venezuela; CVULA, Colección de Vertebrados, Universidad de los Andes, Venezuela; NMNH, National Museum of Natural History, Smithsonian Institution; ANSP, Academy of Natural Sciences; MZFC, Museo de Zoología, Facultad de Ciencias, Universidad Nacional Autónoma de México; KUNHM, University of Kansas Natural History Museum. [1]DNA extracted from museum skin; [2]Sequences from Saunders and Edwards (2000). ***NOTE: accession numbers will be added upon acceptance of the paper***

Sample	**Collection**	**Voucher**	**Locality**	**GenBank Accession Numbers**				
				ND2	CR	cyt*b*	βfib7	AK5
C.v. viridicyanus 1	AMNH	CJV 29	Bolivia, La Paz: Piara, near Pelechuco	X	X			
C.v. viridicyanus 2	AMNH	CBF MH 35	Bolivia, La Paz: Piara, near Pelechuco	X	X		X	X
C. v. viridicyanus 3	LSUMZ	B1268	Bolivia, La Paz: Ca 1km S Chuspipata	X	X[2]			
C. v. viridicyanus 4	LSUMZ	B22738	Bolivia, La Paz: Saavedra, 12 km E Charazani	X	X			
C. v. cyanolaema 5	FMNH	430148	Cuzco Peru: Paucartambo, Pillahuata	X	X	X	X	X
C. v. jolyaea 6	LSUMZ	B3501	Peru, Huanuco: ca. 14 km W Panao	X	X		X	
C. v. jolyaea 7	LSUMZ	B8249	Peru, Pasco: Millpo	X	X	X	X	X
C. v. jolyaea 8	LSUMZ	B8412	Peru, Pasco: Millpo	X	X		X	X
C. v. jolyaea 9	LSUMZ	B43820	Peru, San Martin: Ca 22km ENE Florida	X	X			
C. v. jolyaea 10	LSUMZ	B44528	Peru, San Martin: Ca 22km ENE Florida	X	X			
C. armillata quindiuna 1	QCAZ	2955	Ecuador, Napo: Oyacachi	X	X		X	
C. a. quindiuna 2	QCAZ	2956	Ecuador, Napo: Oyacachi	X	X		X	X
C. a. quindiuna 3	QCAZ	2957	Ecuador, Napo: Oyacachi	X	X	X		
C. a. meridana 1	EBRG	12238	Venezuela, Mérida: La Mucuy	X	X		X	X
C. a. meridana 2	CVULA	563	Venezuela, Mérida: La Mucuy	X	X		X	
C. a. meridana 3	CVULA	564	Venezuela, Mérida: La Mucuy	X	X		X	
C. turcosa 1	NMNH	B03153	Ecuador, Sucumbíos: Cocha Seca	-	X			
C. turcosa 2	ANSP	512	Ecuador, Carchi: Between Maldonado and Tulcán	X	X			
C. turcosa 3	NMNH	B03154	Ecuador, Sucumbíos: Cocha Seca	-	X			
C. turcosa 4	QCAZ	2958	Ecuador, Pichincha: Palmeras, 35 km NW Quito	X	X	X	X	X
C. turcosa 5	LSUMZ	B7770	Ecuador, Pichincha: SW side Cerro Pichincha	X	X			
C. turcosa 6	LSUMZ	B7784	Ecuador, Pichincha: SW side Cerro Pichincha	X	X			
C. turcosa 7	ANSP	4055	Ecuador, Loja: 7 km SE Saraguro	X	X			
C. turcosa 8	ANSP	5046	Ecuador, Zamora-Chinchipe: 6 km NW San Andres	X	X			
C. turcosa 9	LSUMZ	B31759	Peru, Cajamarca: Quebrada Lanchal	X	X		X	X
C. turcosa 10	LSUMZ	B31823	Peru, Cajamarca: Quebrada Lanchal	X	X		X	X

Table 2.1. Continued.

Sample	Collection	Voucher	Locality	GenBank Accession Numbers				
				ND2	CR	cyt*b*	βfib7	AK5
C. turcosa 11	LSUMZ	B31834	Peru, Cajamarca: Quebrada Lanchal	X	X		X	X
C. cucullata cucullata 1	NMNH	B05557	Panama, Chiriquí: Los Planes	X	X	X	X	X
C. c. cucullata 2	LSUMZ	B26406	Panama, Chiriquí: Cordillera Central	X	X			
C. c. cucullata 3	LSUMZ	B26418	Panama, Chiriquí: Cordillera Central	X	X		X	X
C. c. mitrata 1	FMNH	394011	Mexico, Hidalgo: 5 km E Tlanchinol	X	X		X	X
C. c. mitrata 2	FMNH	343730	Mexico, Hidalgo: 5 km E Tlanchinol	X	X	X	X	
C. c. mitrata 3	MZFC	11233	Mexico, Oaxaca: San Martín Caballero	X	X		X	
C. pulchra	QCAZ	3000	Ecuador, Pichincha: Palmeras, 35 km NW Quito	X	X	X	X	X
C. nana	KUNHM	106856*	Oaxaca, Mexico: Totontepec	X[1]	X[1]	X[1]		
C. mirabilis	FMNH	343601	Mexico, Guerrero: El Iris, Sierra de Atoyac	X	X[2]	X	X	X
C. argentigula 1	LSUMZ	B19770	Costa Rica, San José: Villa Mills	X	X	X	X	
C. argentigula 2	LSUMZ	B19790	Costa Rica, San José: Villa Mills	X	X			X
C. argentigula 3	LSUMZ	B19819	Costa Rica, San José: Villa Mills	X	X		X	
C. pumilo	MZFC	B19493	Mexico, Chiapas: 5 km N Coapilla	X	X	X	X	X

Table 2.2. Summary of model parameters and tree scores estimated from maximum likelihood and Bayesian inference. Parameters and tree scores were estimated in PAUP based on the GARLI ML trees.

	Base frequencies				Rate matrix							
Gene	A	C	G	T	AC	AG	AT	CG	CT	α	Pinv	-ln likelihood
CR	0.345	0.228	0.103	0.324	4.463	17.154	2.058	0.703	13.753	1.089	0.343	5118.106
ND2	0.333	0.346	0.095	0.226	0.768	17.930	0.914	0.665	13.431	1.452	0.478	7142.754
cyt*b*	0.298	0.367	0.126	0.210	1.321	11.204	1.687	0.462	22.000	1.004	0.518	6066.965
AK5	0.221	0.289	0.313	0.178	Kappa = 6.425560					-	0.382	1444.680
βfb7	0.319	0.174	0.191	0.316	0.872	5.556	0.351	2.093	3.293	0.645	-	1852.228
Combined	0.308	0.285	0.155	0.252	1.898	11.224	1.705	0.640	15.798	0.585	0.455	22375.786

Table 2.3. Comparison of ancestral area reconstructions showing estimated ancestral areas and relative probabilities of reconstructed states, under maximum parsimony and maximum likelihood models, respectively. Acronyms: Mexico (Mex); Central America (Cent); Northern Andes (NA); and Central Andes (CA). Higher relative (marginal) probabilities are in bold. N = correspond to nodes in Figure 2.6.

				Maximum Likelihood							
	Maximum Parsimony			Mk1 model relative probabilities				GTR model relative probabilities			
N	Mex	Cent	NA	Mex	Cent	NA	CA	Mex	Cent	NA	CA
1	X	X	X	**0.309**	0.252	0.262	0.177	**0.366**	0.324	0.311	0.000
2	X	X	X	**0.403**	0.287	0.159	0.130	**0.544**	0.399	0.058	0.000
3		X		0.031	**0.923**	0.023	0.026	0.056	**0.942**	0.003	0.000
4	X			**0.670**	0.143	0.099	0.089	**0.783**	0.203	0.014	0.000
5	X	X	X	0.230	0.216	**0.377**	0.178	0.193	0.224	**0.583**	0.000
6			X	0.181	0.174	**0.444**	0.201	0.058	0.076	**0.877**	0.000
7			X	0.157	0.154	**0.403**	0.287	0.014	0.021	**0.959**	0.006
8	X	X	X	0.212	0.241	**0.408**	0.132	0.200	0.264	**0.536**	0.000
9	X	X	X	0.353	**0.463**	0.113	0.071	0.390	**0.565**	0.045	0.000

Table 3.1. List of tissue samples and GenBank accession numbers for sequences of species included in the present study. Acronyms: AMNH, American Museum of Natural History; ANSP, Academy of Natural Sciences; CUMZ, Cornell University Museum of Zoology; FMNH, Field Museum of Natural History; KUNHM, University of Kansas Natural History Museum; LSUMZ, Louisiana State University Museum of Natural Science; MZFC, Museo de Zoología, Facultad de Ciencias, Universidad Nacional Autónoma de México; USMNH, National Museum of Natural History, Smithsonian Institution; YPM, Yale Peabody Museum. [1]DNA extracted from museum study skin. ***NOTE: genbank accession numbers will be added upon acceptance of the paper***

Species	#	Collection	Catalog #	Locality	ND2	cyt*b*	AK5	βfib7	TGFβ2.5
C. cayanus	1	KUNHM	5817	Guyana, E. Barima River	DQ912614[a]	DQ912599[a]	DQ912631[a]	DQ912650[a]	X
	2	KUNHM	5819	Guyana, E. Barima River	X	X			
	3	AMNH	ROP 252	Venezuela, Bolivar, 40 km E Tumeremo	X	X			
	4	LSUMZ	25552	Brazil , Amapa	X	X			
	5	FMNH	391652	Brazil , Amapa	X	X			
C. cyanomelas	6	KUNHM	134	Paraguay, Concepcion, San Luis	X	X	X	X	X
	7	KUNHM	3171	Paraguay, Alto Paraguay, Estancia Triunfo	X	X			
	8	FMNH	324102	Peru, Madre de Dios, Hacienda Amazonia	X	X			
	9	FMNH	334725	Bolivia, Santa Cruz, Chiquitos	X	X			
	10	AMNH	2279	Bolivia, Santa Cruz, Velazco, near El Tuna	X	X			
C. chrysops	13	KUNHM	171	Paraguay, Concepción, San Luis	DQ912609[a]	U77334[a]	DQ912626[a]	DQ912646[a]	X
	14	KUNHM	3667	Paraguay, Itapua, San Rafael N.P., San Pedro Mi	X	X			
	15	AMNH	2249	Bolivia, Santa Cruz, Comunidad Karapari	X	X			
	17	LSUMZ	18785	Bolivia, Santa Cruz	X	X			
	16	CUMV	52152	Argentina, Jujuy	X	X			
C. affinis	18	LSUMZ	28539	Panama, Colón	X	X			
	19	LSUMZ	28602	Panama, Colón	X	X			
	20	USMNH	BO 1960	Panama, Bocas del Toro, 4 km W Chriquí Grande	X	X	X	X	X
	21	USMNH	BO 1980	Panama, Bocas del Toro, 4 km W Chriquí Grande	X	X			
C. mystacalis	22	ASNP	1813	Ecuador, Loja, SE Celica, along Río Catamayo	X	X			
	24	ASNP	4153	Ecuador, Loja, 4 km N of Zapotillo	X	X			
	25	ASNP	4596	Ecuador, Loja, 10 km E Mangahurcu	X	X			
	26	LSUMZ	5160	Peru, Lambayeque	X	X	X	X	X
C. heilprini		LSUMZ	48619	Brazil, Amazonas, Rio Uaupés, Jauareté	X	X			
C. cyanopogon	27	FMNH	392998	Brazil, Alagoas, Pirañas, Fazenda Mecejana	X	X			
	28	FMNH	392999	Brazil, Alagoas, Pirañas, Fazenda Mecejana	X	X	X	X	X
C. caeruleus		YPM	80868	Brasil, Sao Paulo, Riberao Onca Parda	X	X			
C. violaceus	29	FMNH	324104	Peru, Madre de Dios, Hacienda Amazonia	X	X			
	30	FMNH	398598	Peru, Madre de Dios, 2.75 km E Shintuya	—	X	X	X	X
	31	ANSP	5697	Ecuador, Sucumbios, Ca. 14 km N Tigre Playa	X	X			
	32	ANSP	5940	Ecuador, Sucumbíos, Ca. 20 km NE Lumbaqui	X	X			
C. cristatellus	33	LSUMZ	13888	Bolivia, Santa Cruz	X	X	X	X	X
	33b	LSUMZ	13889	Bolivia, Santa Cruz	X	X			
	33c	LSUMZ	13915	Bolivia, Santa Cruz	X	X			

Table 3.1. Continued.

Species	#	Collection	Catalog #	Locality	ND2	cyt*b*	AK5	βfib7	TGFβ2.5
C. dickeyi	33s	MZFC	15315	Mexico, Sinaloa	DQ912611[a]	DQ912596[a]	DQ912628[a]	DQ912647[a]	X
	34s	MZFC	SIN 34	Mexico, Sinaloa	X	X			
	35s	MZFC	SIN 35	Mexico, Sinaloa	X	X			
	36s	MZFC	SIN 36	Mexico, Sinaloa	X	X			
C. yncas	35	LSUMZ	30899	USA, Texas	X	X			
	36	LSUMZ	43650	Peru, San Martin	X	X			
	37	LSUMZ	6114	Ecuador, Morona-Santiago	X	X			
	39	MZFC	15927/CON 442	Mexico, San Luis Potosí, San Nicolás de los Montes	X	X			
	40	MZFC	15722/CON 1299	Mexico, Querétaro, Laguna de la Cruz	DQ912610[a]	DQ912595[a]	DQ912627[a]	X	X

Table 3.2. Summary of model parameters and tree scores estimated from maximum likelihood and Bayesian inference. Parameters and tree scores were estimated in PAUP based on the best ML or Bayesian tree.

	Base frequencies				Rate matrix							
Gene	A	C	G	T	AC	AG	AT	CG	CT	α	Pinv	-ln l
ND2	0.330	0.348	0.087	0.236	0.880	23.122	0.669	0.657	11.500	1.146	0.466	-8645.9573
cyt*b*	0.309	0.368	0.120	0.204	0.953	10.132	1.082	0.257	15.627	1.482	0.553	-7973.7999
AK5	0.220	0.292	0.313	0.825	Transition/transversion ratio = 3.276					0.651	N/A	-1415.9265
βfb7	0.322	0.175	0.189	0.315	1.501	8.530	0.655	2.484	0.654	0.485	N/A	-1926.3864
TGFβ2.5	0.241	0.239	0.215	0.305	0.826	7.842	0.649	1.176	4.608	0.332	N/A	-1231.0593

Table 4.1. Interpredictivity tests developed among sister species. Statistical significance was assessed using a binomial test (i.e., calculating the exact binomial probability of predicting a number of occurrence points or fewer). Sister species' pairs are in gray. *denotes cases in species can predict each other significantly better than random ($P < 0.001$).

Comparison	N	Prop. Area	% correct
Calocitta colliei predicts *C. formosa*	340	0.076*	50
C. formosa predicts *C. colliei*	117	0.223*	79
Cyanocorax beecheii predicts *C. sanblasianus*	83	0.123*	41
C. sanblasianus predicts *C. beecheii*	51	0.065*	55
C. affinis predicts *C. heilprini*	10	0.440	10
C. heilprini predicts *C. affinis*	108	0.067	6
C. chrysops predicts *C. cyanopogon*	49	0.405	32
C. cyanopogon predicts *C. chrysops*	112	0.196	10
C. yncas yncas predicts *C. y. luxuosus*	468	0.087*	14
C. y. luxuosus predicts *C. yncas yncas*	127	0.500*	69

Table 4.2. Model parameters for the reconstruction of ecological characters on the *Cyanocorax* phylogeny. Variables: AMT = annual mean temperature; AP = annual precipitation; PWM = precipitation of the wettest month; PDM = precipitation of the driest month; MDR = mean diurnal range; MTWM = maximum temperature of the warmest month; MTCM = minimum temperature of the coldest month. *rejected Model A ($\chi^2_{df=1}$; $P < 0.05$); **not significantly different from $\lambda = 1$ ($\chi^2_{df=1;}$ $P > 0.05$).

					Scaling parameters		
Variable	Ln likelihood	α	Instantaneous variance	β	λ	K	δ
Centroid	-0.054	0.023	0.0181		0	0	0.851
AMT	-1.428	0.917	1.023		0.333**	0.800	1.529
AP	-8.923	0.111	58.155		0.915**	0.833	3
PWM	-6.220	0.106	11.154		0	0.899	2.084
PDM	-6.138*	0.474	248.883	-711.216	1	0.997	3
MDR	-7.524	-0.097	0.042		1	0.215	2.807
MTWM	-5.548	0.065	1.634		0	3	0.247
MTCM	-9.106	0.065	0.048		0	0	1.133

APPENDICES

Appendix I. Published primers used in this study (for both amplification and sequencing).

Gene	Primer	Sequence (5' to 3')	Source
ND2	L5216	GGCCCATACCCCGRAAATG	Sorenson, et al. (1999)
	H6313	ACTCTTRTTTAAGGCTTTGAAGGC	Sorenson, et al. (1999)
cyt*b*	L-14990	CCATCCAACATCTCAGCATGATGAAA	Kocher et al. (1989)
	H16065	GTCTTCAGTTTTTGGTTTACAAGAC	Tim Birt, unpublished
AK5	AK5b+	ATTGACGGCTACCCTCGCGAGGTG	Shapiro and Dumbacher (2001)
	AK6c-	CACCCGCCCGCTGGTCTCTCC	Shapiro and Dumbacher (2001)
βfib7	FIB-B17U	GGAGAAAACAGGACAATGACAATTCAC	Prychitko and Moore (1997)
	FIB-B17L	TCCCCAGTAGTATCTGCCATTAGGGTT	Prychitko and Moore (1997)
TGFβ2.5	TGFb2.5F	TTGTTACCCTCCTACAGACTTGAGTC	Sorenson et al. (2004)
	TGFb2.6R	GACGCAGGCAGCAATTATCC	Sorenson et al. (2004)

Appendix II. ND2 maximum likelihood-corrected pair-wise distances within (diagonal) and among taxa in *Cyanolyca*. Distances for *C. nana* are not included, given the considerable amount of missing data for this species could inflate distance values.

		1	2	3	4	5	6	7	8	9	10	11	12
1	*C. viridicyanus viridicyanus*	0											
2	*C. v. cyanolaema*	0.001	-										
3	*C. v. jolyaea*	0.080-0.083	0.080-0.083	0.001-0.010									
4	*C. armillata quindiuna*	0.149-0.121	0.147-1.149	0.150-0.165	0-0.001								
5	*C. a. meridana*	0.145-0.158	0.143-0.156	0.159-0.177	0.03	0							
6	*C. turcosa*	0.116-0.124	0.111-0.122	0.117-0.134	0.164-0.172	0.170-0.191	0-0.017						
7	*C. cucullata cucullata*	0.18	0.183	0.190-0.195	0.219-0.221	0.229-0.235	0.180-0.188	0					
8	*C. c. mitrata*	0.178-0.182	0.176-0.180	0.201-0.211	0.211-0.217	0.231-0.239	0.177-0.196	0.048-0.05	0-0.003				
9	*C. pulchra*	0.181	0.179	0.190-0.195	0.186-0.189	0.209	0.170-0.186	0.120	0.138-0.142	-			
10	*C. mirabilis*	0.239	0.236	0.258-0.263	0.258-0.260	0.249-0.271	0.236-0.244	0.197	0.194-0.204	0.204	-		
11	*C. pumilo*	0.257	0.254	0.272-0.285	0.262-0.265	0.254-0.275	0.261-0.266	0.234	0.229-0.241	0.218	0.140	-	
12	*C. argentigula*	0.255-0.260	0.253-0.257	0.288-0.295	0.277-0.282	0.272-0.299	0.251-0.268	0.238-0.242	0.219-0.235	0.230	0.132	0.051	0-0.001

Appendix III. Internal primers used to amplify sequences from museum study skin samples.

Gene	Primers
ND2	**L155** CAT CGA AGC AGC YAC TAA RTA YTT CC
	L350 AGG TCA TGC AAG GCT CAT CTC TCA
	L449 TGA ATA GGA CTA AAY CAA ACA C
	L493 GGA GGA TGA ATA GGA CTA AAC
	L664 ATG AAA GTY CTA AAA CTA TCA ACR C
	L835 GCA ATA ATC ATY TCA CTT CTG TC
	H200 GAG GCA GCT TGT ACY ARG
	H416 ATT GGT GGG AAT TTY ATY ACY GTG G
	H574 TAG CTA TTC AGC CCA GGT GAG CAA
	H727 TTG TGC TAA GTG AAG GTG
	H903 TGT TGC RCA GTA TGC TAG GCG AAG
cyt*b*	**L691** CTA GGA TTY GCA CTA ATR CTA RTC CTA C
	L939 ACG CTC AAT AAC CTT CCG TC
	H998 AGG TCT GCA ACT AGK GTT CAR AAT AG
	cyt*b*2.RC and **cyt*b*2.wow** (Dumbacher et al. 2003)

Appendix IV. Summary of results of preliminary model tests based on the 50% split of occurrence points. For each species, the proportional area predicted present and percentage of test points correctly predicted is given, as well as the total test sample size.*indicates that all binomial tests were significant ($P < 0.05$).

Species	A predicts B			B predicts A		
	N_B	Prop. Area*	% correct	N_A	Prop. Area*	% correct
C. affinis	56	0.527	82	52	0.317	62
C. beecheii	26	0.127	89	25	0.095	88
C. cayanus	23	0.317	83	23	0.330	87
C. chrysops	56	0.233	89	56	0.278	82
C. cyanopogon	25	0.198	60	24	0.209	79
C. dickeyi	17	0.069	71	16	70.589	88
C. melanocyaneus	14	0.267	93	11	0.145	64
C. mystacalis	18	0.069	78	18	0.048	61
C. sanblasianus	41	0.086	93	42	0.079	90
C. yucatanicus	81	0.213	94	80	0.294	94
C. yncas yncas	63	0.113	83	64	0.100	66
C. y. luxuosus	234	0.209	96	234	0.208	97
Calocitta colliei	60	0.117	90	59	0.121	90
C. formosa	171	0.115	97	170	0.101	94

Appendix V. Niche centroid distances based on the ecological niche models for *Cyanocorax* species.

Species	1	2	3	4	5	6	7	8	9	10	11	12
1 *Cyanocorax cayanus*												
2 *C. chrysops*	0.990											
3 *C. cyanopogon*	0.985	0.796										
4 *C. heilprini*	0.898	1.000	1.000									
5 *C. affinis*	0.500	0.954	0.850	0.925								
6 *C. dickeyi*	1.000	0.927	0.848	1.000	1.000							
7 *C. sanblasianus*	1.000	0.904	0.400	1.000	0.947	0.898						
8 *C. yucatanicus*	0.929	0.988	0.795	1.000	0.856	0.996	0.932					
9 *C. beecheii*	1.000	0.787	0.702	1.000	0.986	0.738	0.670	0.998				
10 *C. melanocyaneus*	0.985	0.752	0.769	1.000	0.943	0.798	0.917	0.951	0.792			
11 *C. mystacalis*	1.000	0.959	0.838	1.000	0.947	0.910	0.847	0.978	0.854	0.924		
12 *C. yncas yncas*	0.915	0.835	0.950	1.000	0.844	0.951	0.994	1.000	0.959	0.701	0.983	
13 *C. y. luxuosus*	0.774	0.646	0.534	0.984	0.648	0.947	0.795	0.687	0.857	0.819	0.919	0.845

FIGURES

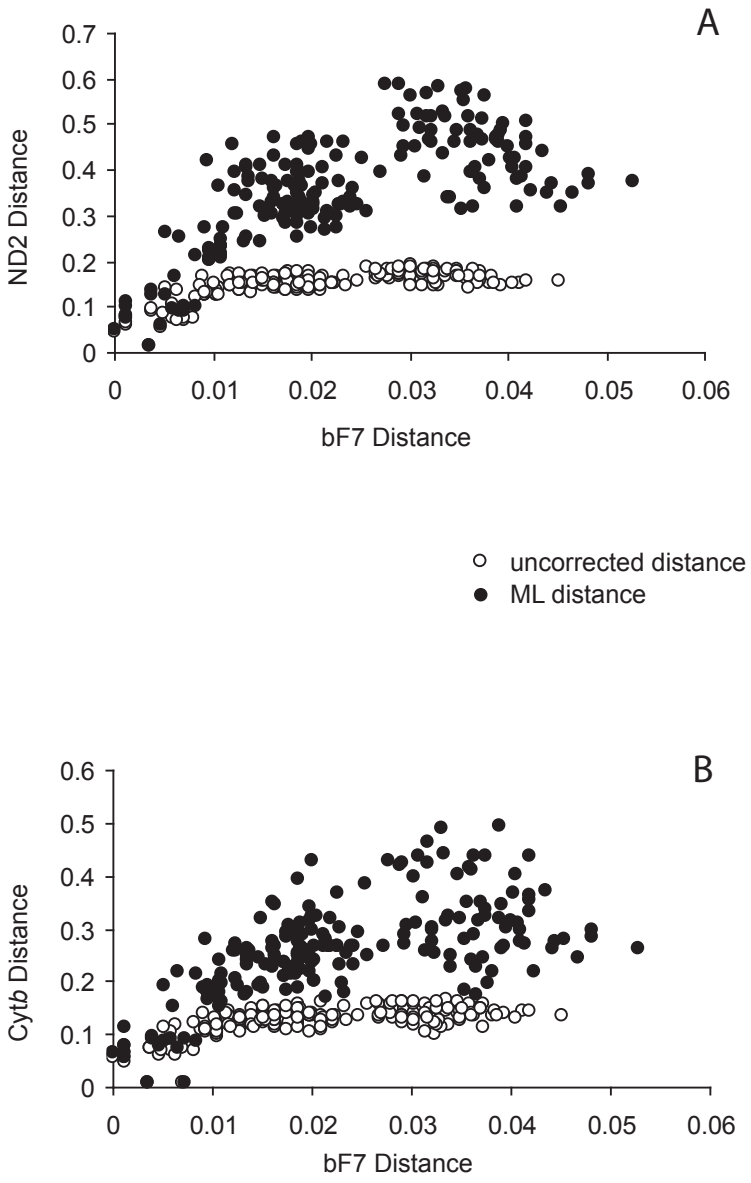

Fig.1.1. Nuclear versus mitochondrial sequence divergence among species studied. βfib distances are plotted on the x-axis and ND2 (A) and cytb (B) distances on the y-axis. Clear circles indicate points derived from uncorrected pairwise distances; solid circles indicate points derived from maximum likelihood distances determined using model and parameter values estimated by Modeltest under the AIC.

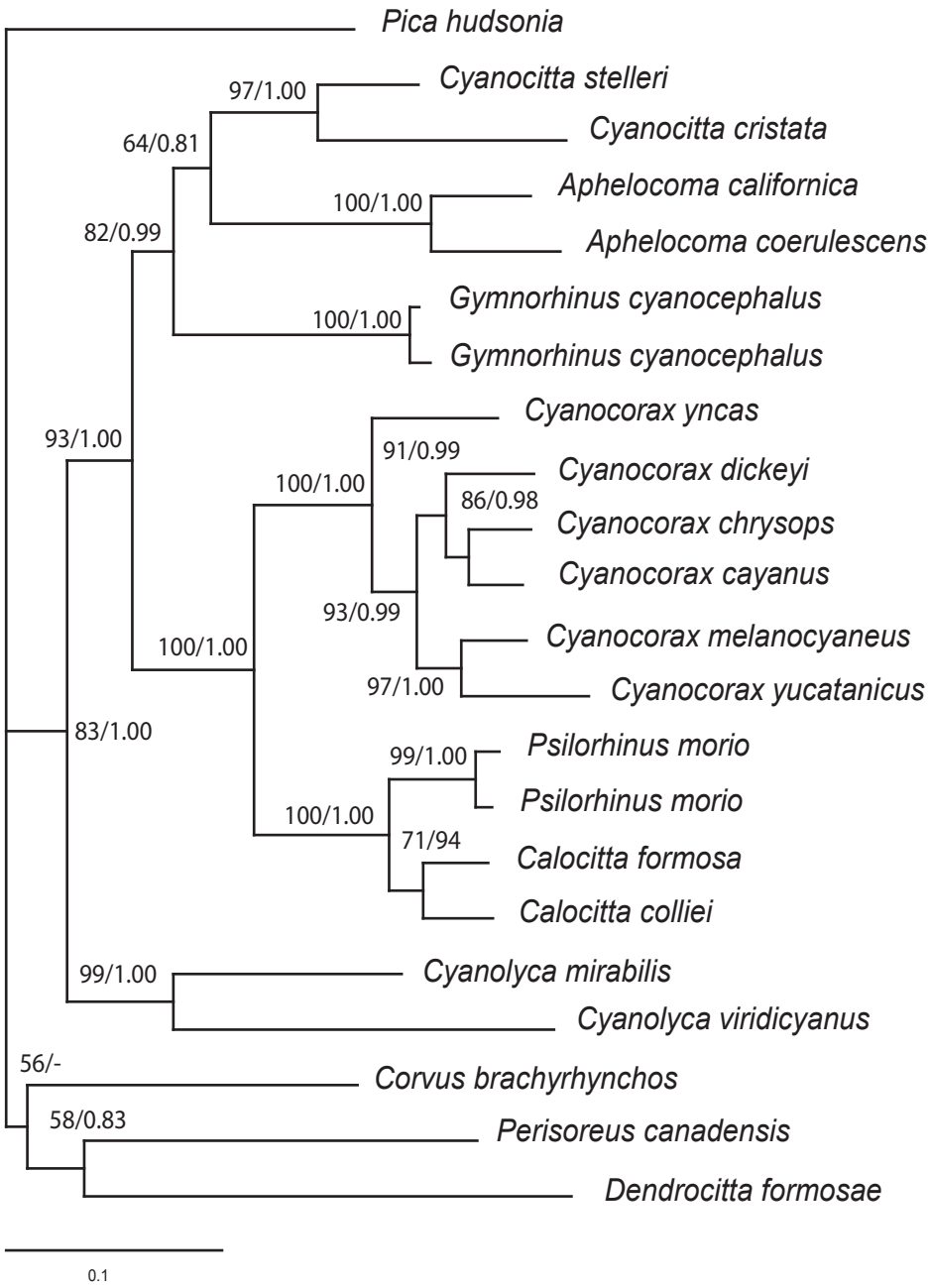

Fig. 1.2. Maximum likelihood tree obtained from ND2 sequences. Numbers over nodes indicate ML bootstrap support and posterior probabilities obtained for the Bayesian analysis (BA) 50% majority rule consensus tree; "-" indicates nodes not supported by BA.

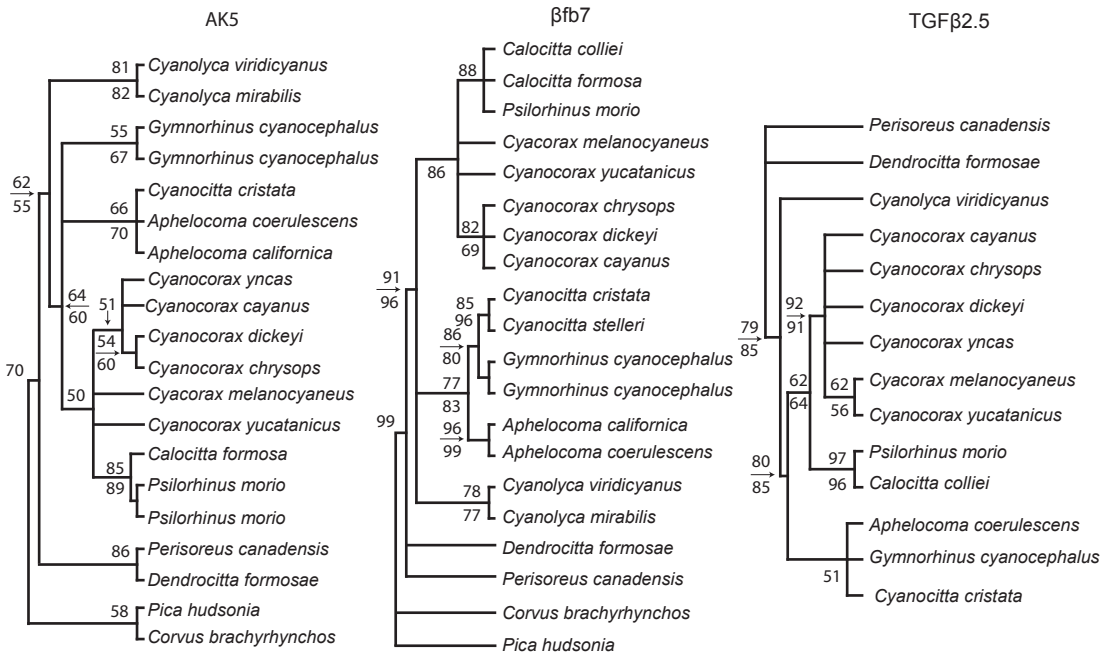

Fig. 1.3. Maximum parsimony 50% majority rule consensus trees obtained from the analyses of AK5, βfib7, and TGFβ2.5. Numbers over and under nodes indicate MP and ML non-parametric bootstrap values >50% respectively.

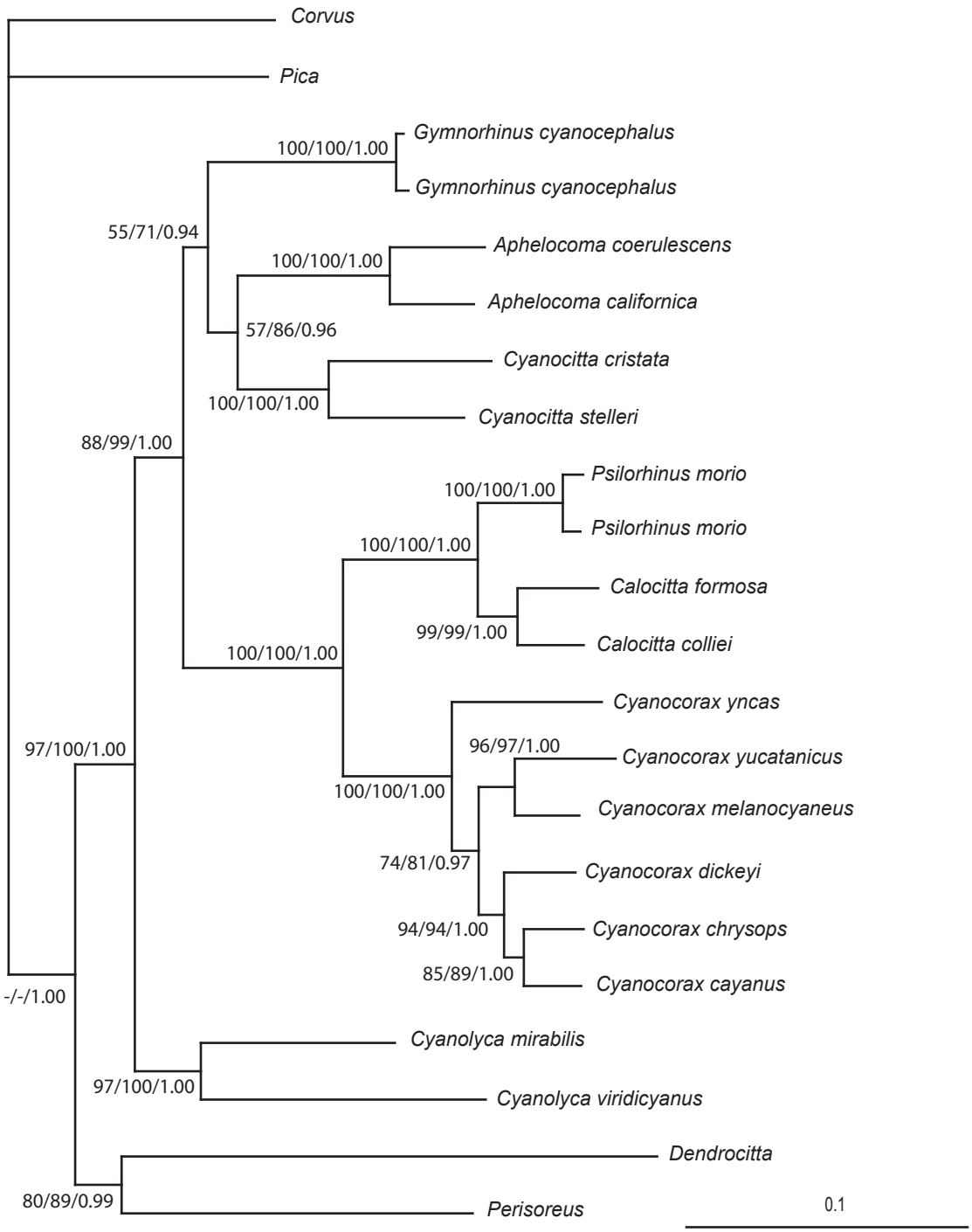

Fig. 1.4. Maximum likelihood tree that resulted from the combined analysis of mitochondrial (ND2, cytb, and CR) and nuclear (AK5, βfib7, and TGFβ2.5) genes. Numbers over nodes indicate maximum parsimony bootstrap support of the 50% majority rule consensus. tree/maximum likelihood bootstrap support/Bayesian posterior probabilities; "-" indicates nodes not supported.

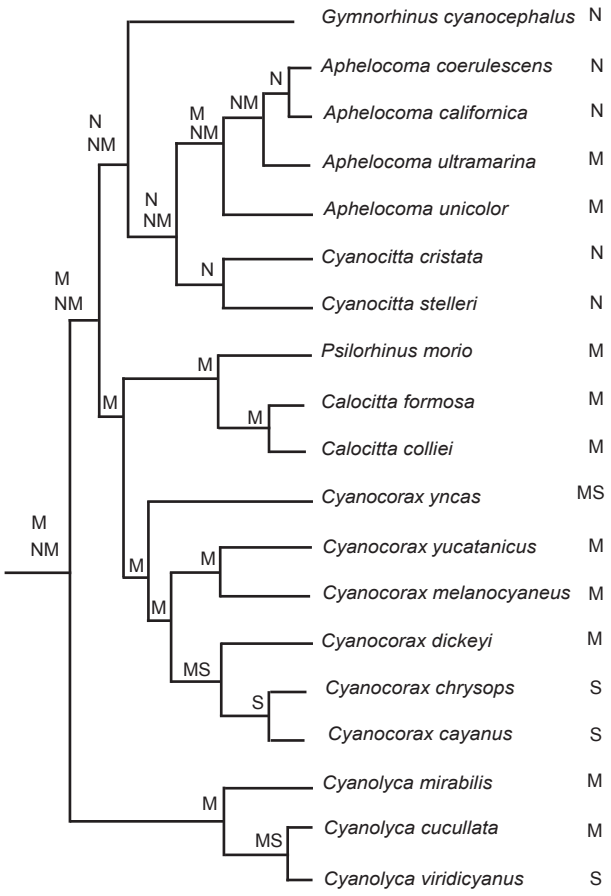

Fig. 1.5. Optimization of ancestral areas—North America (N), Mesoamerica (M), and South America (S)—using DIVA. Taxon representation over our combined tree was expanded based on previous analyses of *Aphelocoma* (Rice et al., 2003) and ND2 sequence data on *Cyanolyca* (E. Bonaccorso, unpub.).

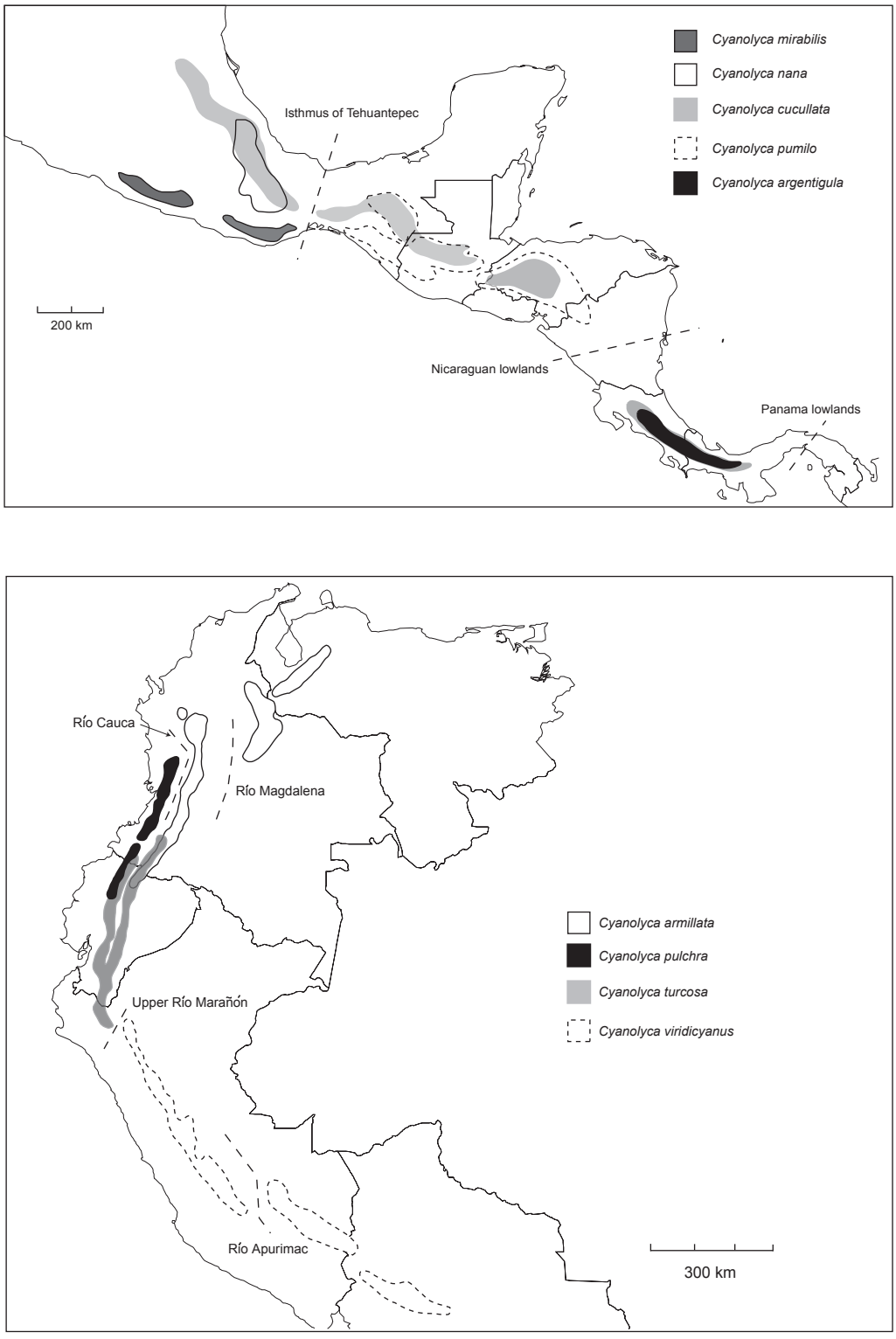

Fig. 2.1. Maps of Mesoamerica and northern South America showing the distribution of the nine species of *Cyanolyca* jays.

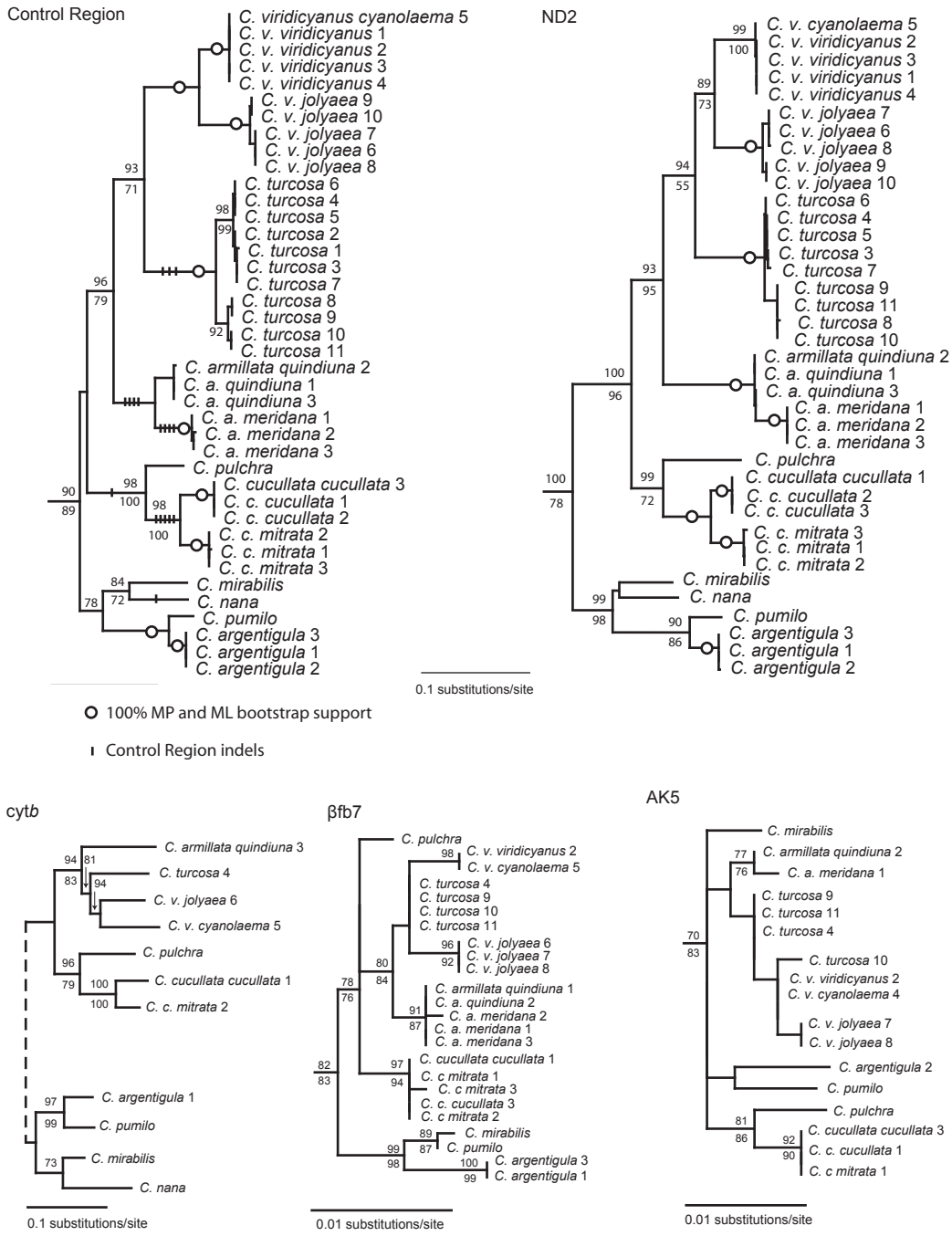

Fig. 2.2. Maximum likelihood (ML) trees estimated for the individual mitochondrial (Control Region, ND2, cytb) and nuclear genes (βfib7 and AK5). Values on nodes indicate ML (above) and maximum parsimony (below) bootstrap proportions. Dashed line among main clades in the cyt*b* tree indicates that the monophyly of *Cyanolyca* was not recovered.

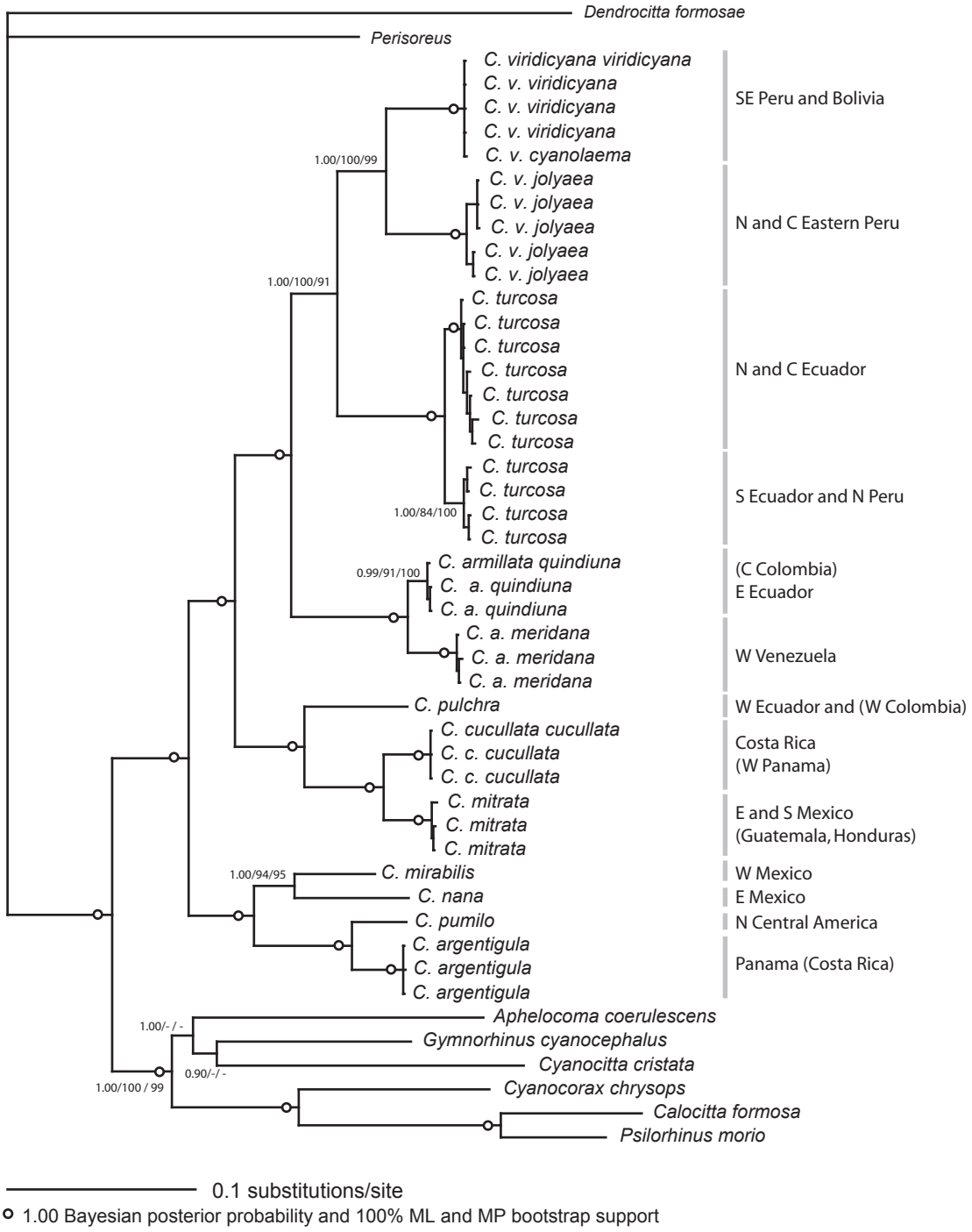

Fig. 2.3. Bayesian majority rule consensus tree (50%) estimated from the combined analysis of the Control Region, ND2, cytb, βfib7 and AK5. Bayesian posterior probabilities/maximum likelihood bootstrap values/and maximum parsimony bootstrap values are indicated whenever nodes were recovered with less than 1.00 posterior probability or 100% bootstrap support.

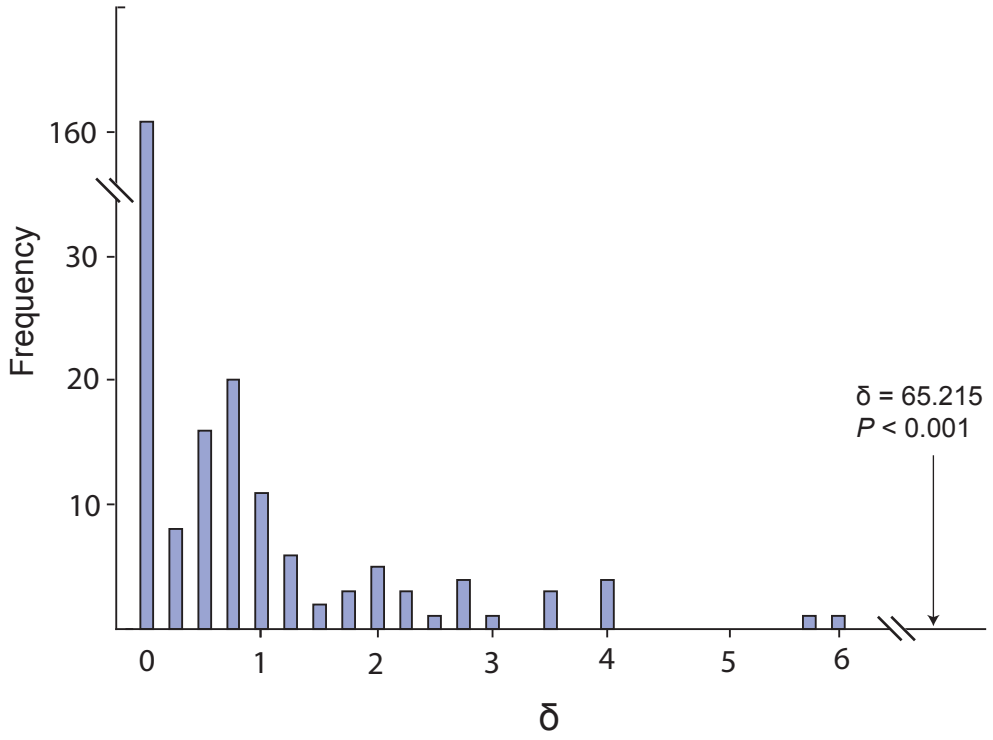

Fig. 2.4. Null distribution of δ drawn from the simulated datasets. The arrow indicates the δ calculated from the actual dataset and the level of significance used to reject a sister relationship between *Cyanolyca viridicyanus* and *C. armillata.*

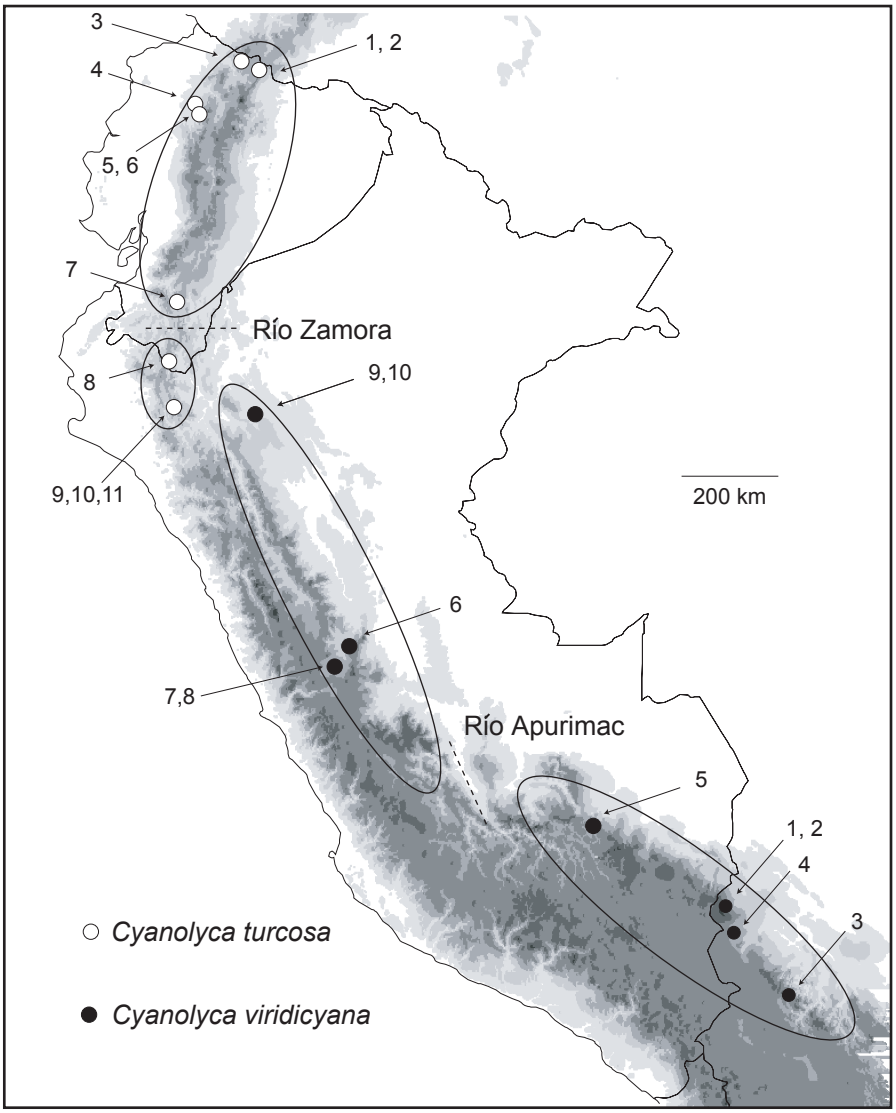

Fig. 2.5. Sampling localities in Ecuador and Peru for *Cyanolyca turcosa* and *C. viridicyanus*. Ovals indicate different geographic groups identified in the phylogenetic analysis of the Control Region and ND2.

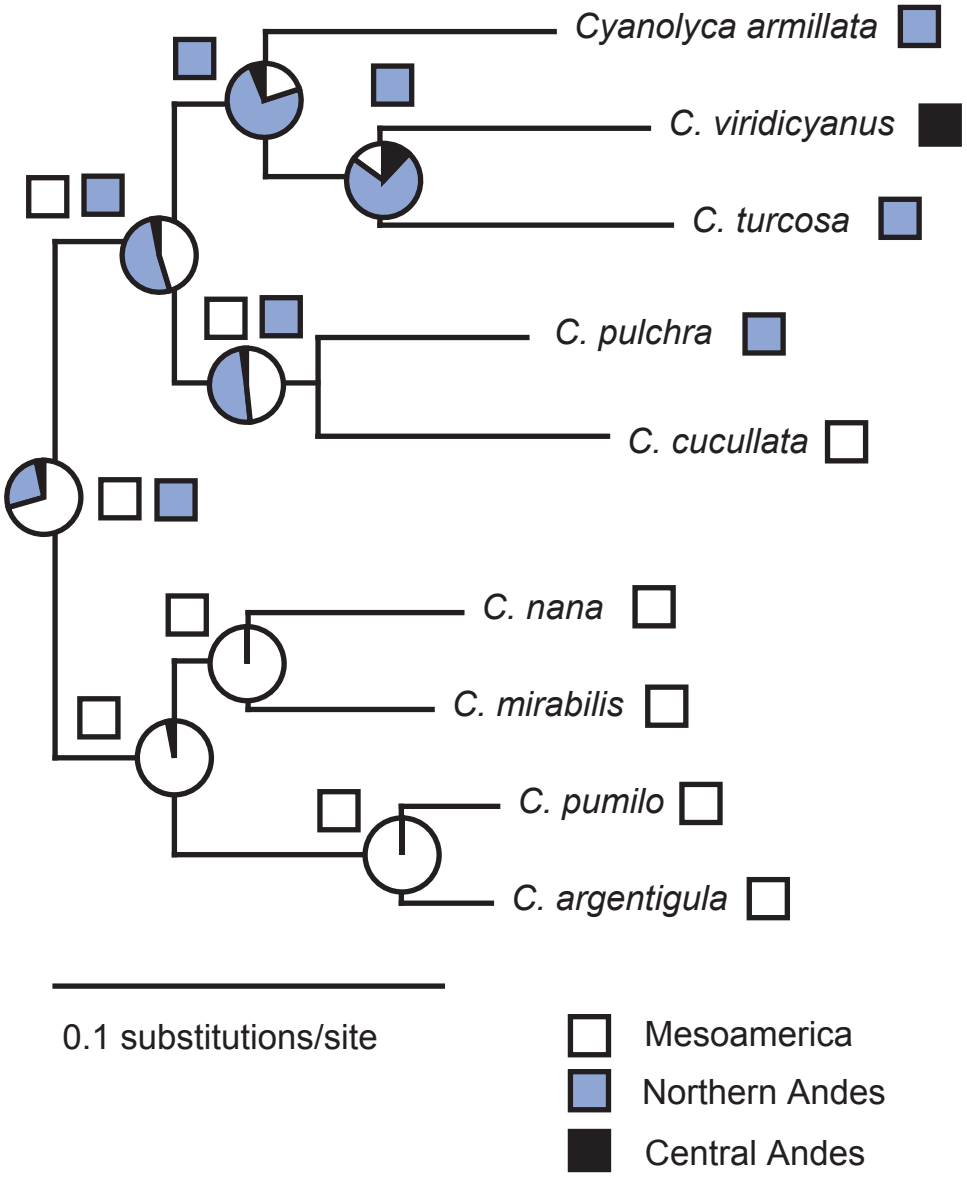

Fig. 2.6. Ancestral area reconstructions for species in *Cyanolyca.* Relative probabilities of each character state are indicated by the pie chart at every node (maximum likelihood), and discrete ancestral character states (Parsimony) are indicated by rectangles.

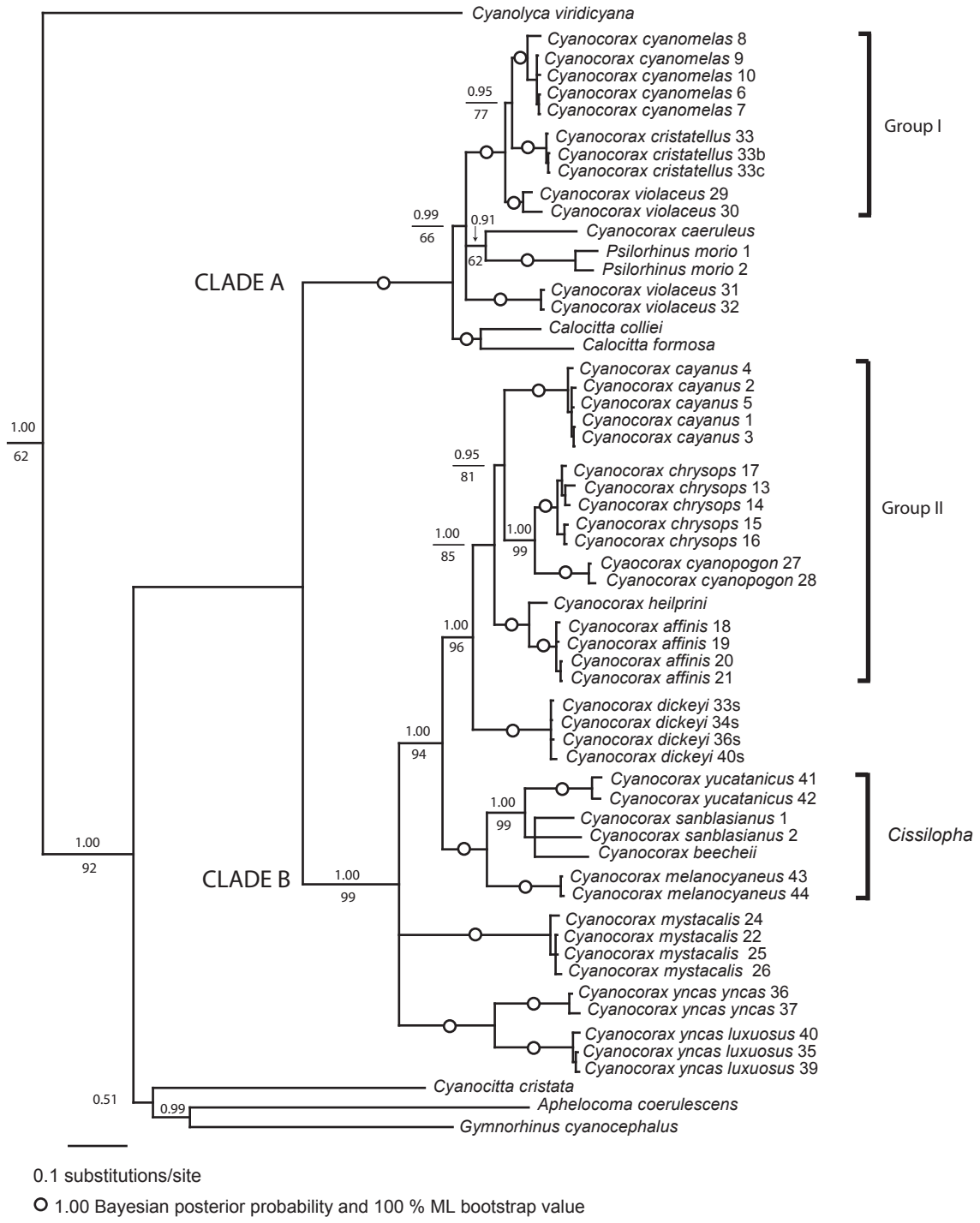

Fig. 3.1.Bayesian 50% majority rule consensus tree estimated from the mitochondrial dataset (ND2 and cytb; 2001 bp). Bayesian posterior probabilities/maximum likelihood bootstrap values are indicated whenever nodes were recovered with less than 1.00 posterior probability or 100% bootstrap support..

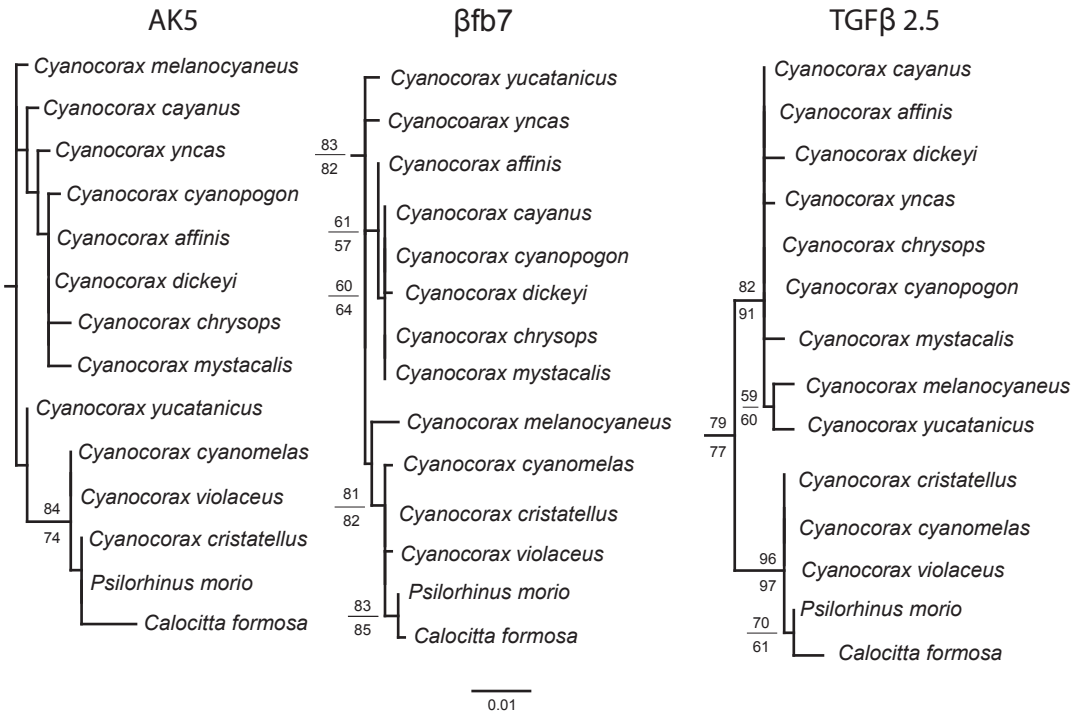

Fig.3.2 Maximum likelihood (ML) trees estimated for the individual nuclear genes (AK5, βfib7, and TGFβ2.5). Values on nodes indicate ML (above) and maximum parsimony (below) bootstrap support.

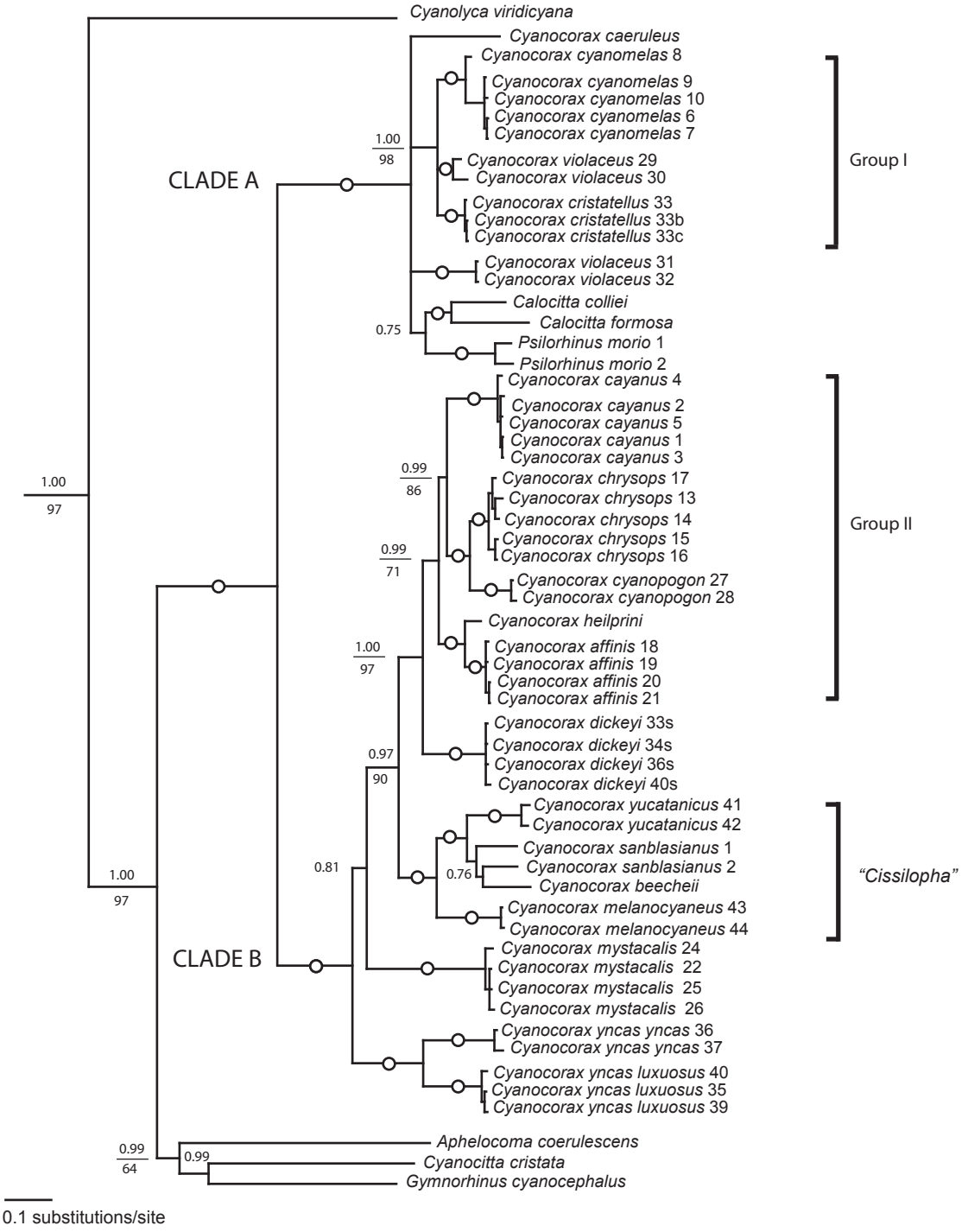

Fig. 3.3. Bayesian 50% majority rule consensus tree estimated from the total evidence dataset (ND2, cytb, AK5, βfb7, and TGFβ2.5; 4077 bp). Bayesian posterior probabilities/maximum likelihood bootstrap values are indicated whenever nodes were recovered with less than 1.00 posterior probability or 100% bootstrap support.

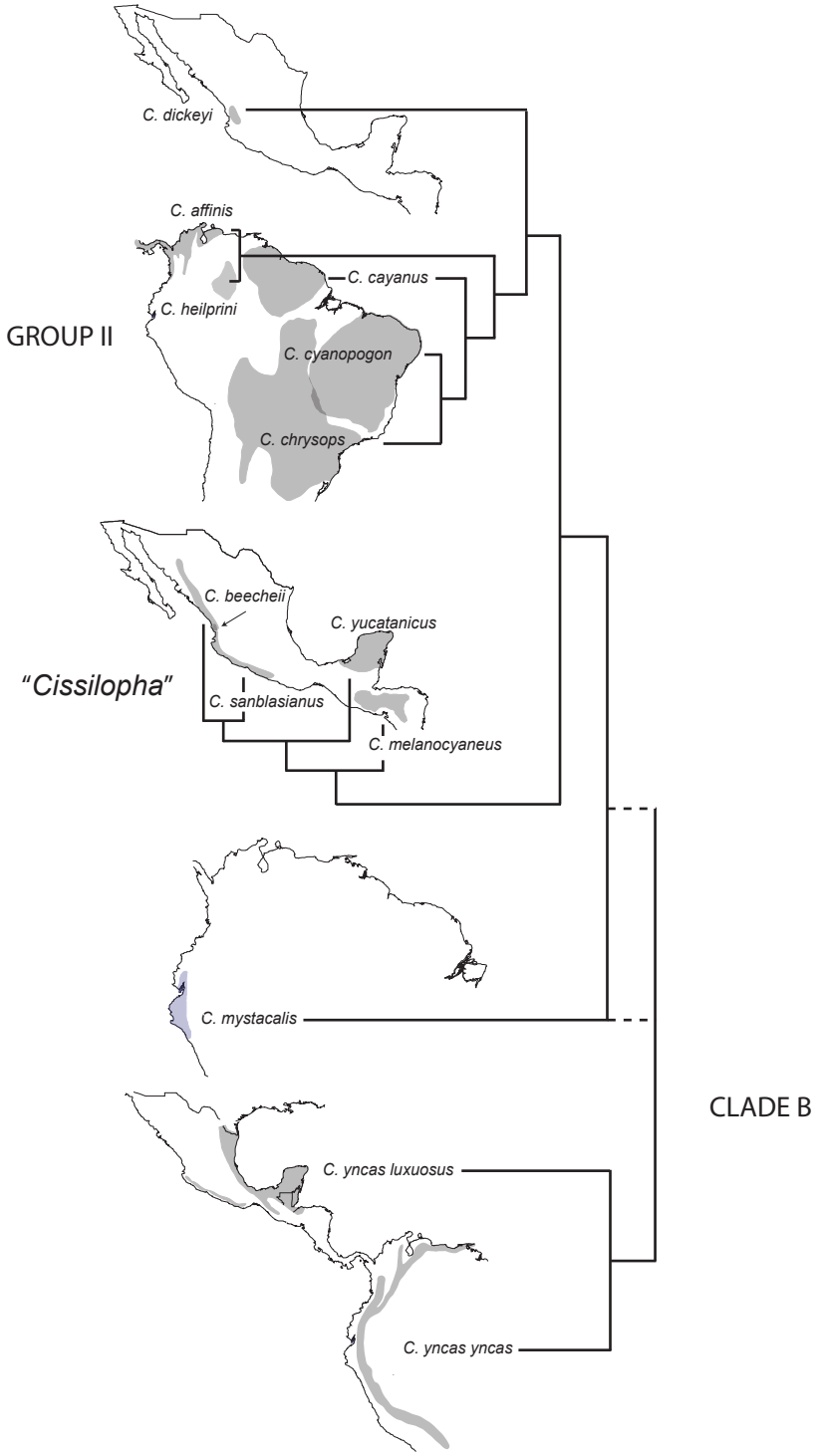

Fig. 3.4. Geographic distribution and phylogenetic relationships among species in clade B. Darker shades (and arrow in the case of *Cyanocorax beecheii* and *C. sanblasianus*) indicated areas of range overlap. Dashed lines indicate uncertainty in the position of *C. mystacalis.*

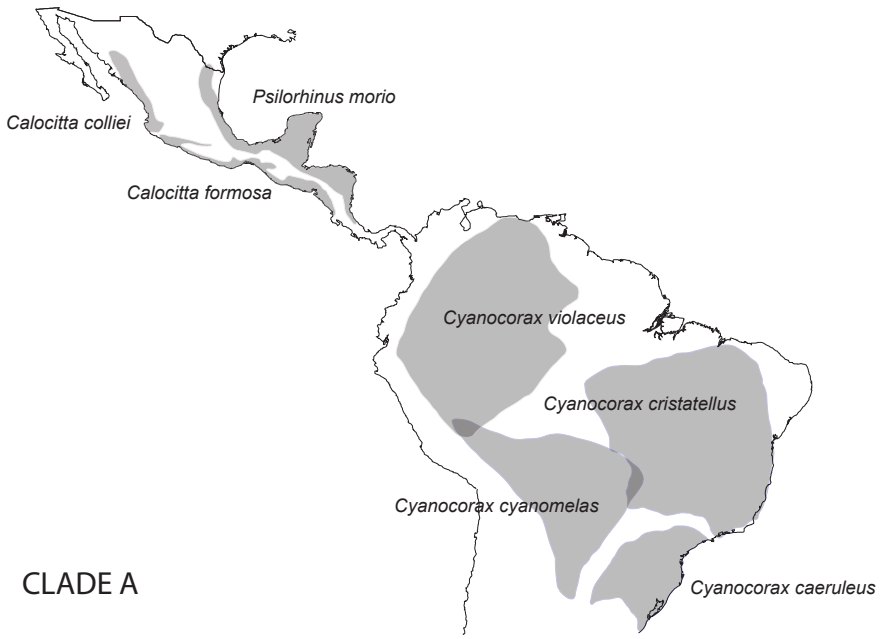

Fig. 3.5. Geographic distribution of species in clade A. Darker shades indicated areas of range overlap.

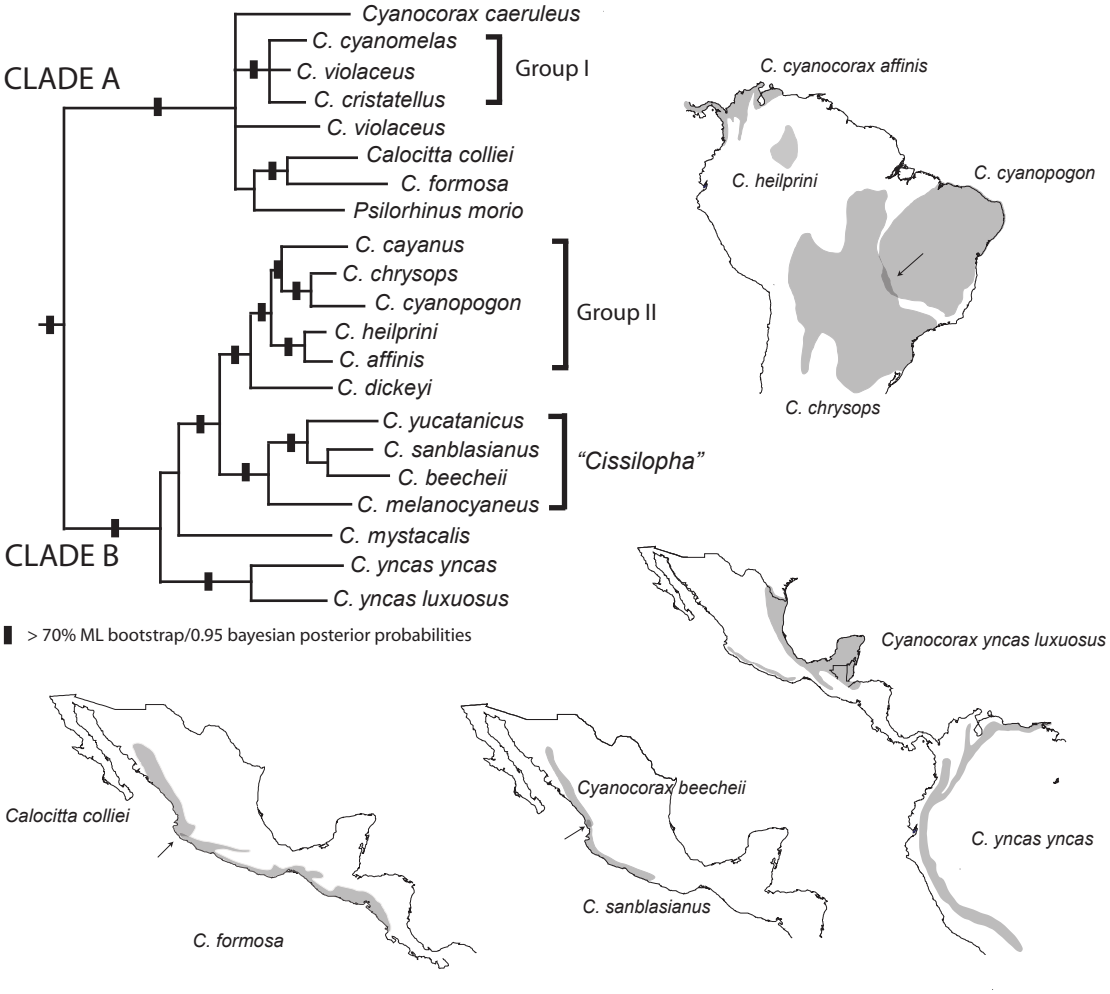

Fig. 4.1. Phylogeny and geographic distribution of the "*Cyanocorax*" assemblage. Arrows indicate areas of contact between sister species.

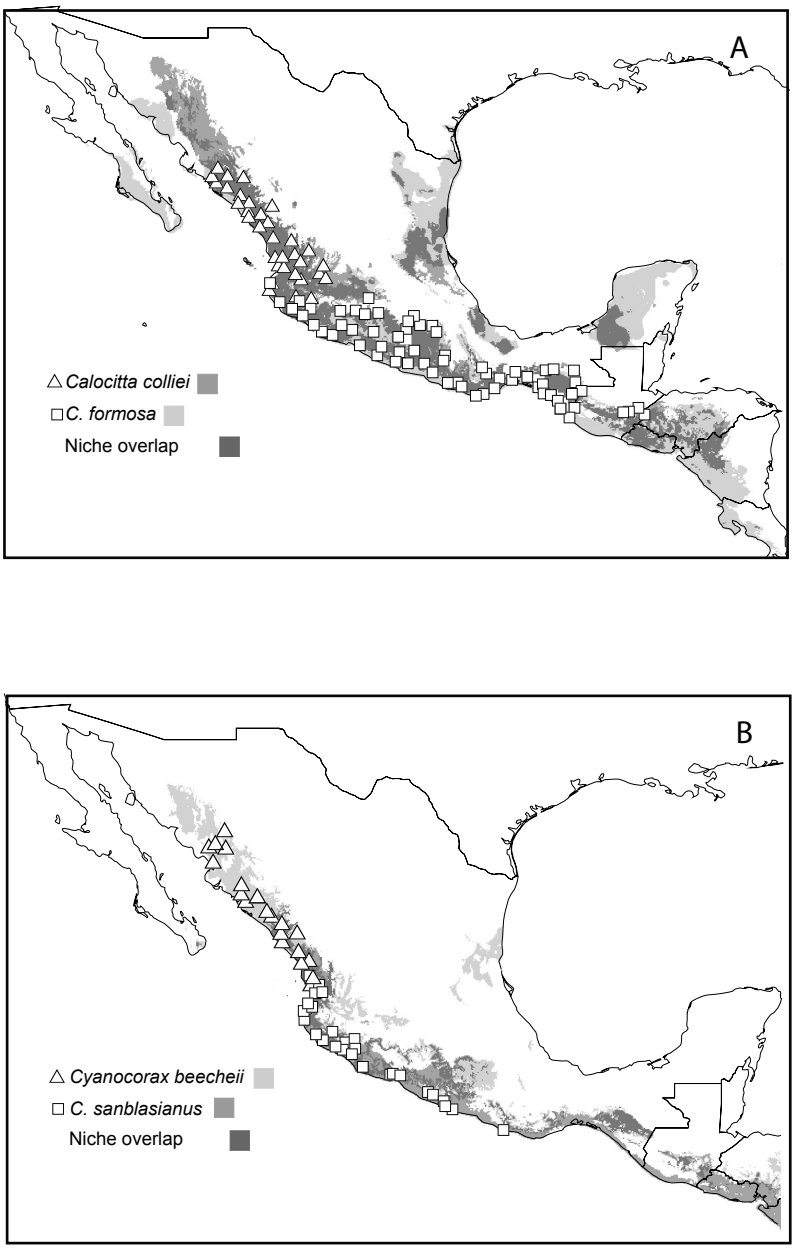

Fig. 4.2. Geographic projections of ecological niche models for two sister species pairs: *Calocitta formosa* and *C. colliei* (A), and *Cyanocorax beecheii* and *C. sanblasianus*. The darkest grey represents areas of predicted niche overlap. Density of occurrence data has been reduced to facilitate display.

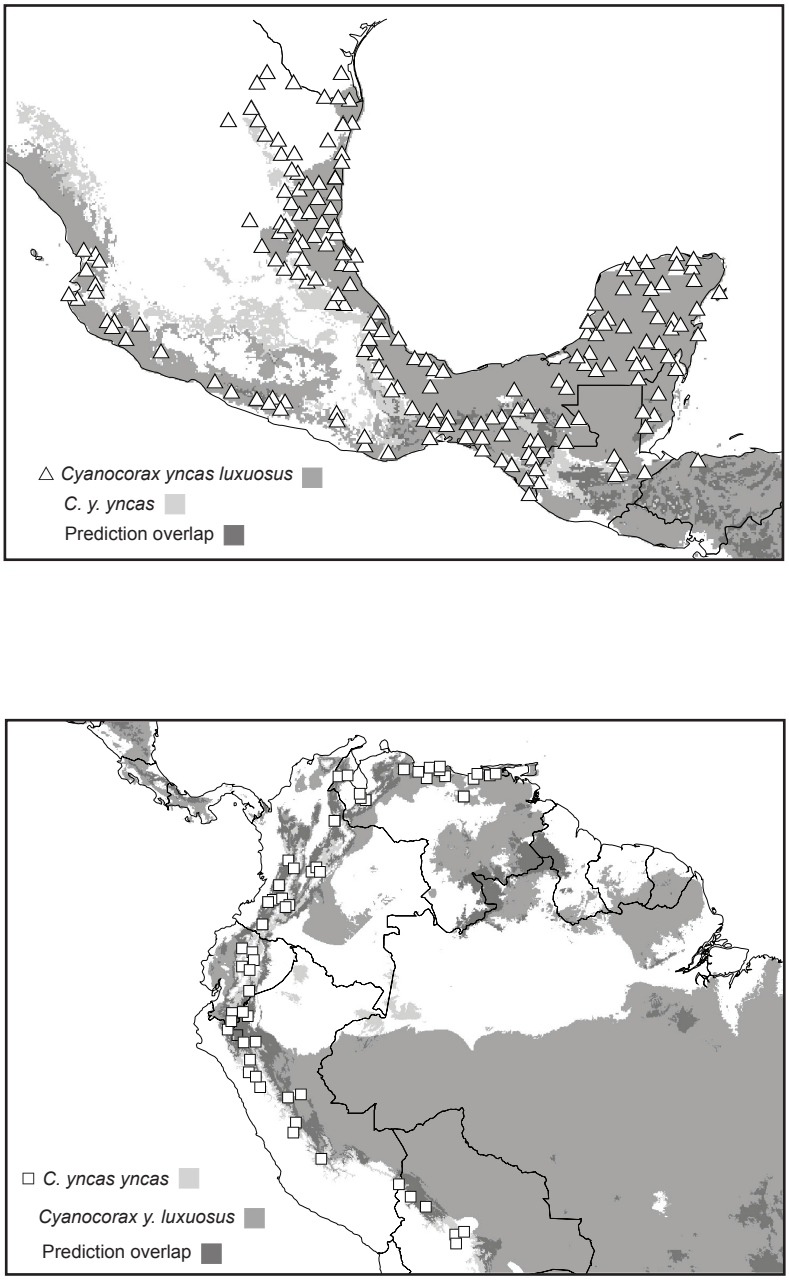

Fig. 4.3. Geographic projections of ecological niche models for northern and southern populations of *Cyanocorax yncas*. Panels depict *C. y. luxuosus* (A) and *C. y. yncas* (B) in their areas of distribution, as well as the niche prediction projections for both populations. The darkest grey represents areas of predicted niche overlap. Density of occurrence data has been reduced to facilitate display.

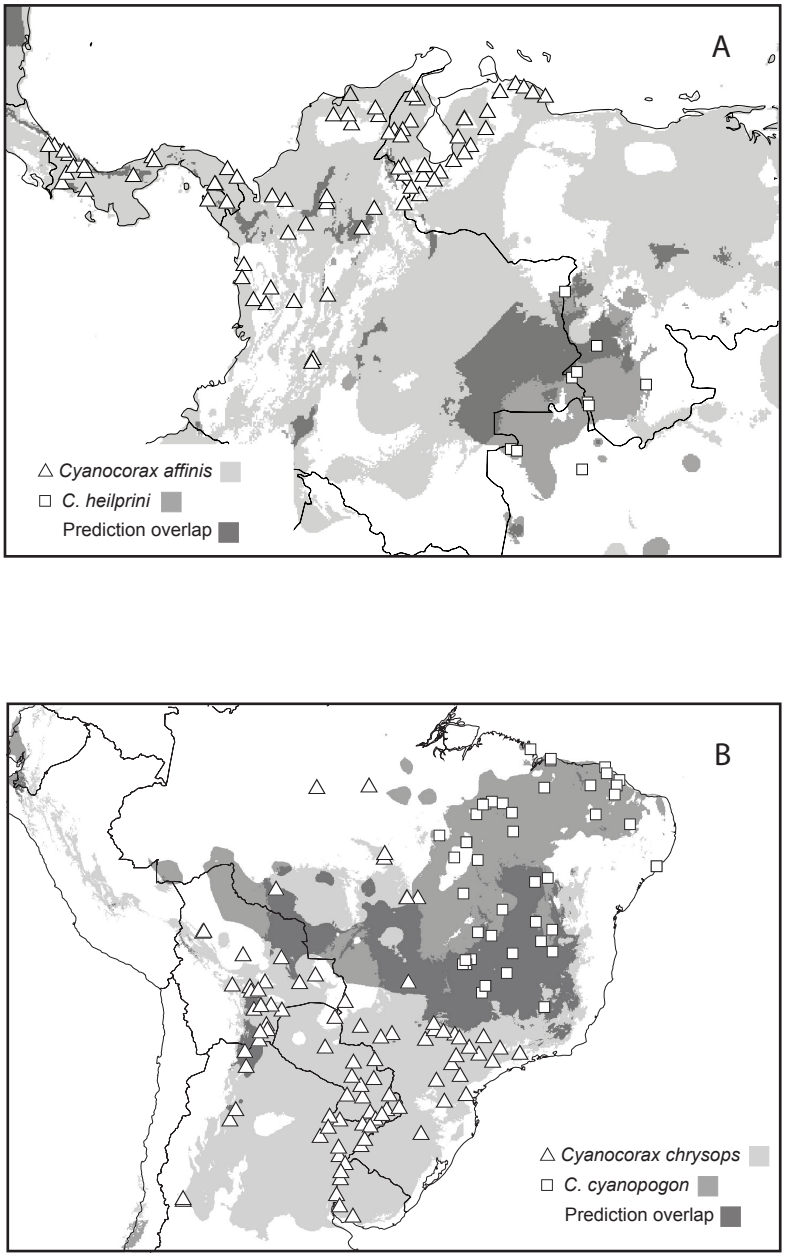

Fig. 4.4. Geographic projections of ecological niche models for two sister species pairs: *Cyanocorax affinis and C. heilprini* (A), and *C. chrysops* and *C. cyanopogon* (B). The darkest grey represents areas of predicted niche overlap. Density of occurrence data has been reduced to facilitate display.

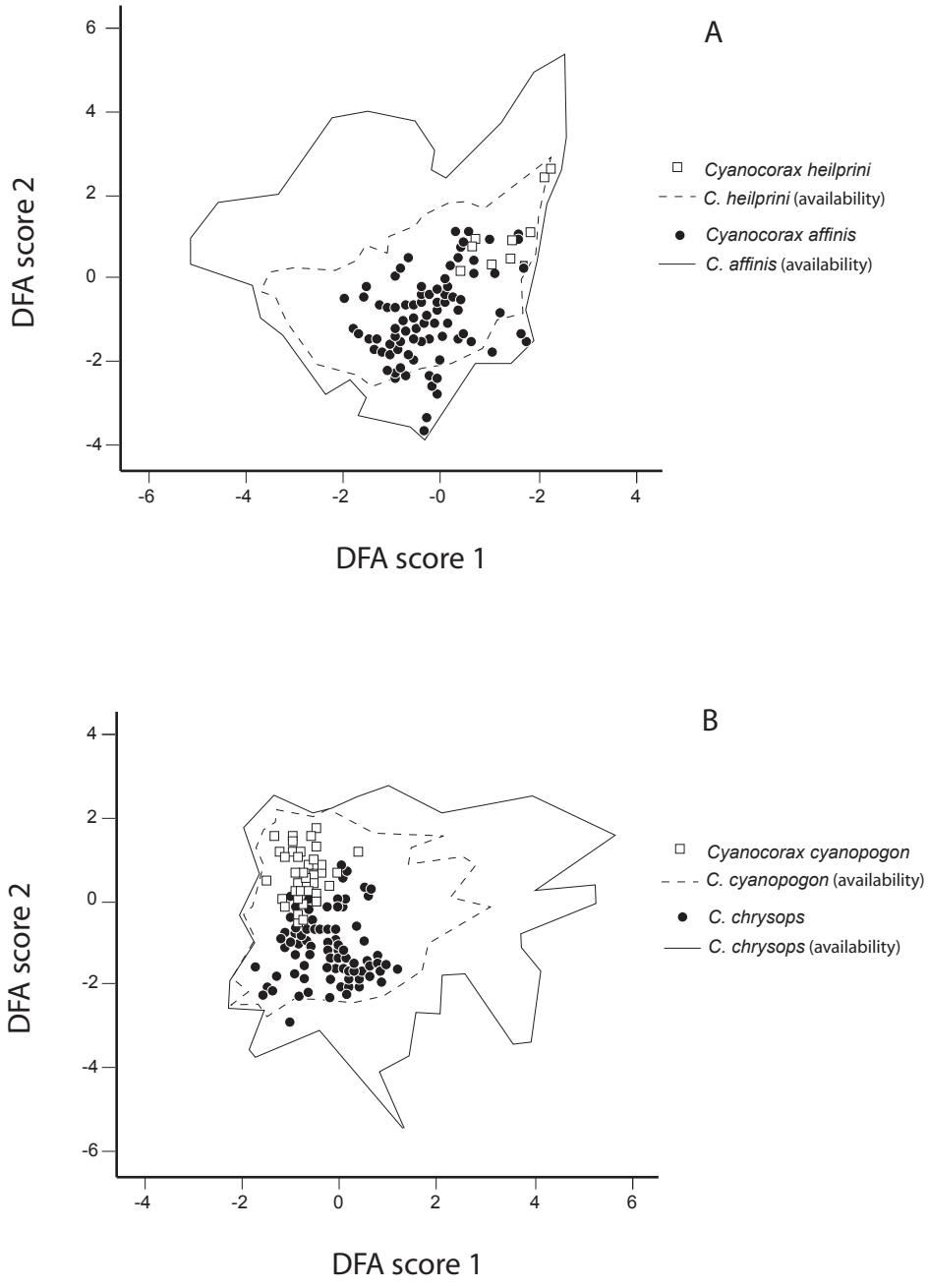

Fig. 4.5. Scaterplot of canonical discriminant analysis scores for : (A) *C. affinis* vs. *C. heilprini* and (B) *C. cyanopogon* vs. *C. chrysops*. Simbols represent occurence points; lines represent environmental conditions existing in a buffer defined around each species' distributional area.

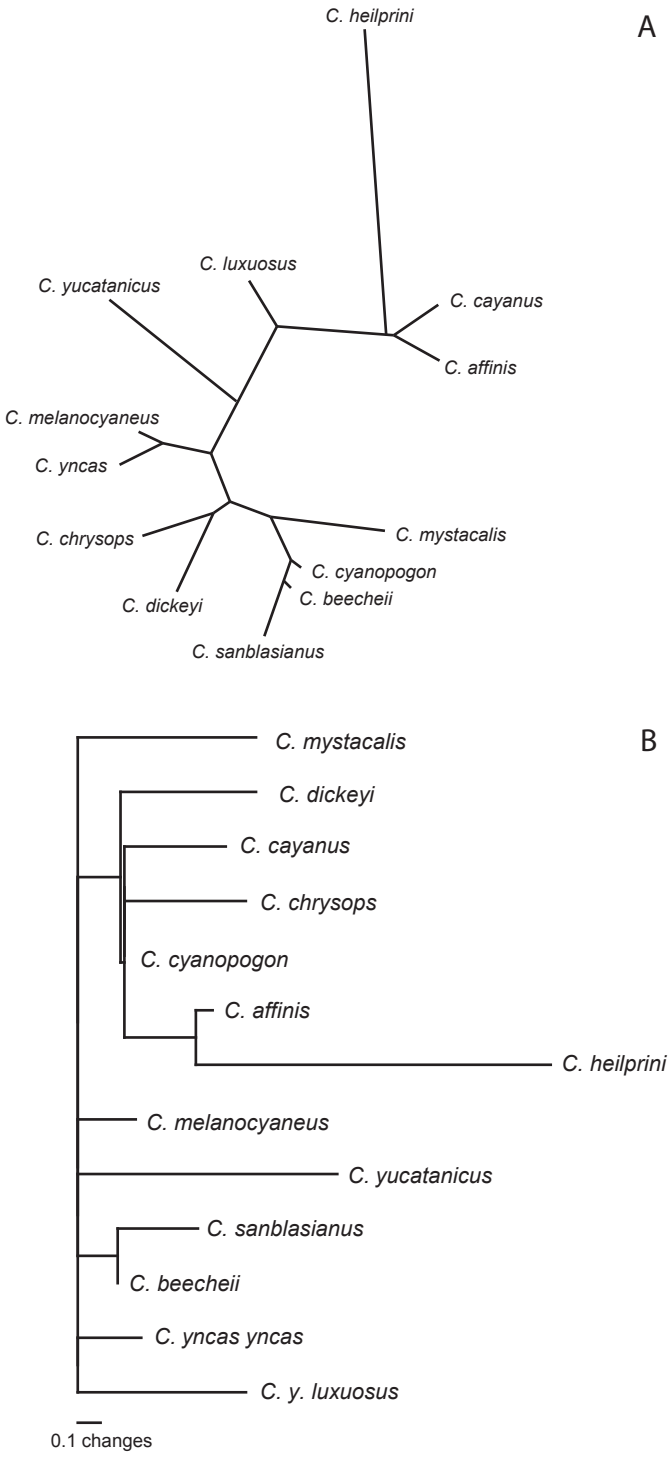

Fig. 4.6. Representation of centroid distances derived from ecological niches. Unrooted tree based on ecological niches (A) and ecological distances mapped onto the phylogeny (B). Booth trees were obtained using least-square fitting procedures.