

A multilocus phylogeny of New World jay genera

Elisa Bonaccorso *, A. Townsend Peterson

Natural History Museum and Biodiversity Research Center, Department of Ecology and Evolutionary Biology, The University of Kansas,
1345 Jayhawk Boulevard, Lawrence, KS 66045, USA

Received 22 January 2006; revised 20 June 2006; accepted 30 June 2006
Available online 18 July 2006

Abstract

We studied phylogenetic relationships of the New World Jays (NWJs) based on DNA sequences from three mitochondrial and two nuclear loci. Sampling included at least two individuals from each of the seven NWJ genera and four outgroups of closely related corvids, as well as six of the 16 *Cyanocorax* species (including two representatives of the previously recognized “*Cissilopha*”). Phylogenetic analyses were conducted using maximum parsimony, maximum likelihood, and Bayesian analyses for individual genes and a combined dataset. The combined phylogenetic analysis supports the basal position of *Cyanolyca* to all other NWJs, a (*Cyanocorax* (*Calocitta*, *Psilorhinus*)) clade, and a ((*Cyanocitta*, *Aphelocoma*) *Gymnorhinus*) clade that agrees with a novel morphological synapomorphy uniting *Cyanocitta* and *Aphelocoma*. Within *Cyanocorax*, *C. yncas* (former “*Xanthoura*”) is basal to a split among former “*Cyssilopha*” species and the rest of the *Cyanocorax* species. To explore implications for the historical biogeography of the NJWs, we used Dispersal–Vicariance Analysis, which indicated that NWJs originated either in Mesoamerica or North America + Mesoamerica, with South American NWJs dispersing three times independently from Mesoamerica.

© 2006 Elsevier Inc. All rights reserved.

Keywords: New World jays; Corvidae; Phylogeny; Biogeography

1. Introduction

The New World jays (NWJs) are an assemblage of 34 species in seven genera of corvids—*Aphelocoma*, *Cyanocitta*, *Gymnorhinus*, *Cyanolyca*, *Calocitta*, *Psilorhinus*, and *Cyanocorax*—endemic to the Americas. Monophyly of this assemblage is not in serious doubt, given unique shared morphological characters (Zusi, 1987). NWJs have radiated successfully in tropical, subtropical, and temperate habitats, and show a mix of narrow endemism (e.g., *Cyanolyca mirabilis*) and broad distributions (e.g., *Cyanocitta cristata*), including species of particular conservation concern (Bird-Life International, 2000). Furthermore, they have served as a model system for numerous analyses of the evolution of avian social systems (e.g., Brown, 1963; Fitzpatrick and Woolfenden, 1985; Edwards and Naem, 1993; Saunders

and Edwards, 2000). As such, understanding the historical underpinnings of current NWJ diversity is critical on a number of fronts.

Although a broad older literature attempted to treat NWJ history and evolution (Amadon, 1944; Hardy, 1961), the first steps towards an understanding of NWJ phylogeny began with a study based on cytochrome *b* sequences (Espinosa de los Monteros and Cracraft, 1997), which provided a basic framework, but little detail of specific relationships. Later studies (Ericson et al., 2005; Saunders and Edwards, 2000) added detail, and considered additional genes in making the NWJ tree more robust.

However, two of the previous studies (Ericson et al., 2005; Espinosa de los Monteros and Cracraft, 1997) considered only single representatives of each genus, providing little detail useful for understanding the historical biogeography of the clade and no test of monophyly of genera. The more detailed study (Saunders and Edwards, 2000), however, considered but a single gene (the mitochondrial control region) and included only two South American NWJ

* Corresponding author. Fax: +1 785 864 5335.
E-mail address: elisab@ku.edu (E. Bonaccorso).

representatives. As such, considerable room remains for clarifying the details of NWJ phylogeny.

The present paper offers a multigene view of the ‘deep’ nodes in the NWJ phylogeny. We have assembled much improved sampling of key species from each genus, particularly as relates to the older phylogeny and biogeography of the group, as well as sequence data from additional mitochondrial and nuclear loci. This study offers a key understructure to understanding the broader picture of NWJ history and evolution.

2. Methods

2.1. Sampling

We included at least two individuals from each of the seven NWJ genera—*Aphelocoma*, *Cyanocitta*, *Calocitta*, *Psilorhinus*, *Cyanocorax* (including *Cissilopha*) and *Gymnorhinus*—and four individuals representing closely related corvid genera: *Dendrocitta*, *Pica*, *Perisoreus*, and *Corvus*. Also, in a first attempt to elucidate relationships among *Cyanocorax* jays, we analyzed six of the 16 currently recognized species of *Cyanocorax*, including two representatives of the previously recognized “*Cissilopha*” assemblage, as well as Mesoamerican and South American species. Tissue samples were obtained via our own fieldwork in Mexico, El Salvador, and Paraguay, as well as from ornithological collections in the US and Mexico (Appendix A).

Genewise, based in our own sampling effort and taking advantage of that of previous studies, we 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. We obtained complete sequences for the NADH Dehydrogenase Subunit 2 (ND2), the Adenylate Kinase gene, intron 5 (AK5) and the β -Fibrinogen intron 7 (β fib7), and partial sequences for cytochrome *b* (*cytb*). Finally, we concatenated our results with published sequences of the mitochondrial control region (CR) from Saunders and Edwards (2000), ND2 from Cicero and Johnson (2001), *cytb* from Espinosa de los Monteros and Cracraft (1997), and β fib7 from Ericson et al. (2005).

2.2. DNA amplification and sequencing

Genomic DNA was extracted from frozen tissue with the DNeasyTissue extraction kit (Qiagen). PCR amplification was conducted using published primers (Appendix B). Amplification of mitochondrial genes followed 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); amplification of β fib7 followed Pritchitko and Moore (1997). For amplifying AK5, we applied a modified protocol based on that of Shapiro and Dumbacher (2001), where annealing temperature was raised to 64 °C, increasing specificity of primer annealing to obtain single PCR products.

PCR products were visualized in agarose gel, and unincorporated primers and DNTPs were removed from PCR products using ExoSap purification (ExoSap-it, GE Health Care). Cycle sequencing reactions were completed using the corresponding PCR primers and BigDye Terminator 3.1 chemistry (Applied Biosciences). For mitochondrial genes and most nuclear samples, we 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); for difficult samples, we lowered extension temperature to 52 °C to increase the length of the sequence reading.

Reaction products were purified with CleanSEQ magnetic beads (Agencourt) and run in an ABI Prism 3100 Genetic Analyzer (Applied Biosciences). Data from heavy and light strand were spliced together to arrive at a consensus sequence for each taxon, using Sequencher 4.1 (Gene Codes Corporation, 2000). Precautions against potential nuclear pseudogenes of mitochondrial origin (Sorensen and Quinn, 1998) included sequencing both DNA strands and checking that amino acid translation was possible. Sequencing of *Cyanocorax dickeyi* was performed at Universidad Nacional Autónoma de México (UNAM).

2.3. Alignment and phylogenetic analyses

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

Evolutionary rate heterogeneity across lineages was tested by using the likelihood ratio (LR) test (Felsenstein, 1988). Significance was assessed by comparing $A = -2 \log LR$, 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 *cytb* and ND2 at high levels of sequence divergence, we plotted uncorrected and ML distances of these two genes against β fib7 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 with 100 random additions (Felsenstein, 1985). Analysis 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 see whether such indels supported the same branches as did the single-nucleotide variation, phylogenetic trees were also constructed 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 evolution 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). Then, ML 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). We analyzed mitochondrial genes individually, partitioning by codon position in the case of *cytb* and ND2, and by domain (domain I, central domain, and domain II; Saunders and Edwards, 2000) for CR. Each analysis consisted of 2×10^6 generations and four Markov chains with default heating values. Parameter values for the model 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. We also made sure that the potential scale reduction factor (PSRF) was around 1.00 for all parameters and 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: *cytb* and ND2 by codon, CR by domain, AK5, and β fib7. Each partition was assigned its best-fit model of evolution 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 analyses of 2×10^6 generations were performed. After confirming that the four analyses reached stationarity at a similar likelihood scores and that the topologies were similar, the resulting 6000 trees were used to calculate posterior probabilities (PP) in a 50% majority-rule consensus tree.

For combined analysis, our sequences of *Perisoreus canadensis* were concatenated with the published CR sequence of *Perisoreus infaustus*, and our sequences of *Pica hudsonia* were concatenated with that of *Pica nuttallii*. Although both the *Perisoreus* and *Pica* combined sequence data are technically chimeric, their appropriateness to root the tree is justified because the two pairs of the chimera are in all probability much more closely related to one another than to those of other species in the dataset (Saunders and Edwards, 2000).

2.4. Topological congruence and combinability

Each dataset produced a different topology regarding the position of *Aphelocoma*, *Cyanocitta*, and *Gymnorhinus* (hereafter the “ACG” clade). To assess potential causes of these differences, we 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).

We also performed Shimodaira–Hasegawa tests of hypothesis (Shimodaira and Hasegawa, 1999) to assess whether each individual dataset rejected a particular topology for the ACG clade, when compared with the ML topology generated by that dataset. For each dataset, we obtained the ML tree using the best-fit model of evolution and estimated parameter values, and used the same data to obtain a ML tree under the constrained topology generated by a different dataset as the null hypothesis. Then, likelihood values of both trees were compared using the Shimodaira–Hasegawa test, as implemented in PAUP (full optimization, 1000 bootstrap replicates). Clade support and significance of likelihood differences using the Shimodaira–Hasegawa test were taken as measurements of potential combinability of datasets. We avoided using the incongruence length difference test (IDL test; Farris et al., 1994), as test of combinability given recent criticisms (e.g., Cunningham, 1997; Barker and Lutzoni, 2002).

2.5. Historical biogeography of the NJWs

To explore the implications of our results for the historical biogeography of the NJWs, we optimized the known distribution of each species onto our combined 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.

Since the closest relative of the NWJ has not been determined, we excluded outgroups, and optimized areas on the ingroup tree, using default settings and up to two possible ancestral areas (max-areas command = 2). To improve optimization of ancestral areas, we expanded taxon representation on our tree based on previous analyses of the phylogeny of *Aphelocoma* (Rice et al., 2003) and preliminary ND2 sequence data on *Cyanolyca* (Bonaccorso, unpublished data).

3. Results

3.1. Sequence attributes

DNA sequence lengths and general characteristics for each gene are summarized in Table 1. As expected,

Table 1
Number of aligned and informative positions, base frequencies, best-fit models, $-\ln$ likelihood with and without enforcing a molecular clock, and test of heterogeneity among lineages for all loci analyzed

Gene	Alignment positions		Base frequency				Best-fit model (AIC)	$-\ln$ likelihood		$-2 \log LR$	χ^2 crit	df	P
	Total	Inform.	A	C	G	T		no molecular clock	molecular clock				
ND2	1041	367	0.33	0.35	0.09	0.23	Trn + I + Γ	7331.8	7348.0	32.4	31.41	20	<0.05
cytb	999	301	0.30	0.36	0.13	0.21	GTR + I + Γ	6358.6	6385.8	54.4	45.32	20	<0.001
CR	1299	425	0.29	0.25	0.14	0.31	GTR + I + Γ	8184.6	8226.5	83.8	45.32	20	<0.001
AK5	602	23	0.22	0.29	0.31	0.17	HKY + Γ	1466.5	1476.7	20.4	28.87	18	>0.05
β Fb7	871	40	0.32	0.18	0.19	0.32	GTR + Γ	1995.9	2007.7	23.6	30.14	19	>0.05

sequence variation was comparable among the mitochondrial genes, and substantially higher than in the nuclear introns, with β fib7 evolving faster than AK5. Pairwise distances among the 18 species of jays and 4 outgroups for ND2, cytb, and β fib7, are summarized in Fig. 1. Among interesting features of divergence is that when plotting uncorrected distances, ND2 and cytb saturate early compared to β fib7; saturation is corrected using ML distances. Giving the saturation observed, parsimony analyses for ND2 and cytb were performed including only the outgroups most closely related to the NWJs (i.e., *Perisoreus* and *Dendrocitta*, see below). Tests of homogeneity of base frequencies across taxa were not significant for any gene ($P > 0.90$), and rate heterogeneity was detected marginally in the case of ND2 ($P < 0.05$), and more dramatically in cytb and CR ($P < 0.001$; Table 1).

Both introns showed a number of indels, the most interesting being a synapomorphic indel of 1-bp in *Cyanolyca*

for β fib7, 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* for AK5. Three samples/genes could not be sequenced—*Cyanocorax yncas* for β fib7, and *Calocitta coillei*, and *Cyanocitta stelleri* for AK5. In the case of *Dendrocitta formosae*, we obtained a partial sequence for ND2 (999 bp).

3.2. Phylogenetic analyses

In all analyses, we recovered the monophyly of NWJs with relatively high support. For ND2 (Fig. 2), ML and BA trees confirmed the basal position of *Cyanolyca* to all other NWJs, as well of a strongly supported clade that includes *Cyanocorax*, *Calocitta*, and *Psilorhinus*, of which the later two are reciprocally monophyletic and sister to *Cyanocorax*. Inside *Cyanocorax*, *C. yncas* appears as basal to a clade in which *C. melanocyaneus* and *C. yucatanicus* (two of the four “*Cissilophas*”) are monophyletic and sister to the rest of the *Cyanocorax*. The MP tree differed only in placing *Cyanocitta* basal to *Aphelocoma* + *Gymnorhinus* (as opposed to *Gymnorhinus* basal as in the ML and BA trees).

For cytb, the MP tree had the same topology and similar bootstrap supports as the ND2 MP tree. The ML and BA trees showed *Aphelocoma* and *Gymnorhinus* as sisters, but failed to recover the ACG clade, creating a polytomy among *Cyanocitta* and the clades formed by *Aphelocoma* + *Gymnorhinus*, and *Cyanocorax* + *Calocitta* + *Psilorhinus*.

The two nuclear genes provided a less resolved picture. For β fib7, we obtained an identical topology with or without indels included (Fig. 3A). However, for AK5, the *Cyanocorax* + *Calocitta* + *Psilorhinus* clade was not supported by nucleotide variation, but was by the 5-bp indel described before (Fig. 3B). Both nuclear genes reconstructed a monophyletic *Calocitta* + *Psilorhinus*, but were unable to provide more resolution within the assemblage *Cyanocorax* + *Calocitta* + *Psilorhinus*. The basal position of *Cyanolyca* among NWJs was recovered by AK5, but not by β fib7, and the clade ACG appeared in the β fib7 but not in the AK5 tree.

Regarding levels of homoplasy among datasets, all mitochondrial genes showed similar levels of character consistency as measured by CI and RC (Table 2). Both nuclear genes had similar levels of consistency, and were about half homoplasious than mitochondrial genes.

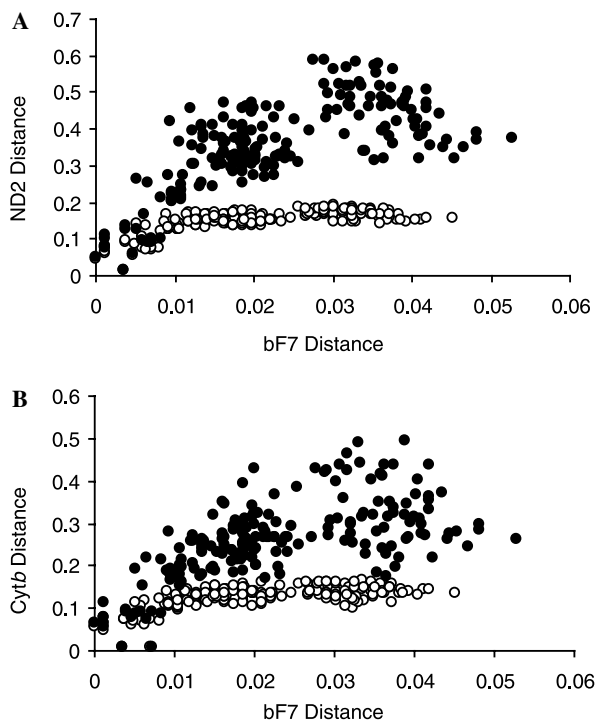


Fig. 1. Nuclear versus mitochondrial sequence divergence among species studied. β fib7 distances are plotted on the x-axis and cytb and ND2 distances on the y-axis. (A) Uncorrected proportional distances. (B) Maximum likelihood distances determined using models chosen by Modeltest under the AIC.

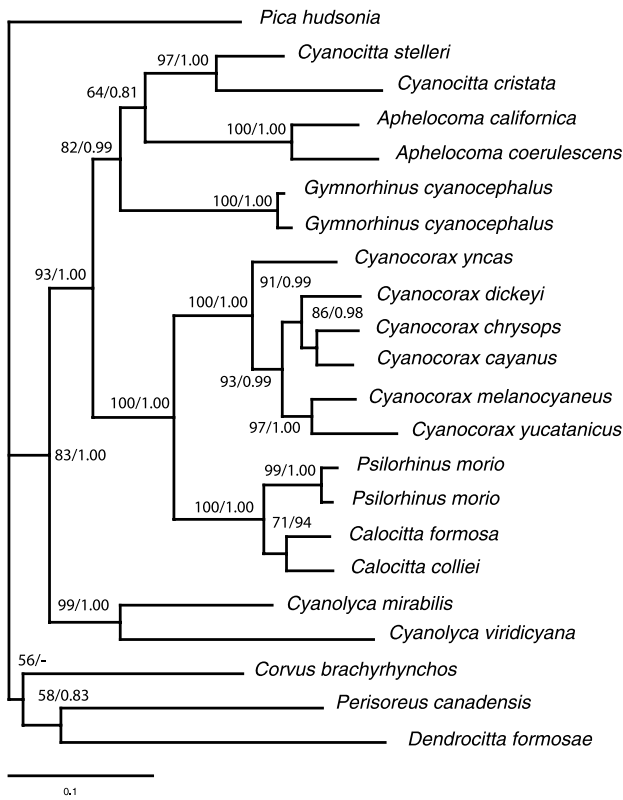


Fig. 2. Maximum likelihood tree obtained from ND2 sequences. Numbers over nodes indicate ML bootstrap support and posterior probabilities obtained for the Bayesian analysis (BA) 50% majority rule consensus tree; “-” indicates nodes not supported by BA.

3.3. Topological congruence and combined analysis

Most conflict in our results centered on the arrangement of ACG (Table 2). Whereas data from ND2, *cytb*, and *βfib7* generated an ACG clade with variable degrees of support, CR placed *Gymnorhinus* as sister to *Cyanocorax* + *Calocitta* + *Psilorhinus* with relatively high support; however, this dataset did not reject the ACG grouping (Shimodaira–Hasegawa test, $P=0.08$), so we are comfortable with exploring the implications of an ACG clade further. Variation existed on how the ACG taxa themselves grouped: whereas ND2, AK5, and CR supported *Aphelocoma* + *Cyanocitta* (AC), *cytb* supported *Aphelocoma* + *Gymnorhinus* (AG), and *βfib7* supported *Cyanocitta* + *Gymnorhinus* (CG). Topological congruence was rejected only when CR was constrained to the arrangements (AG) and (GC) indicating that CR data only support an (AC) arrangement ($P < 0.05$).

Table 3 shows synapomorphic characters supporting each of the possible arrangements within ACG. Among mitochondrial genes, 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 *cytb*, all but changes but one are in third positions, and at least half are transitions, which is congruent with their lower CIs. In ND2,

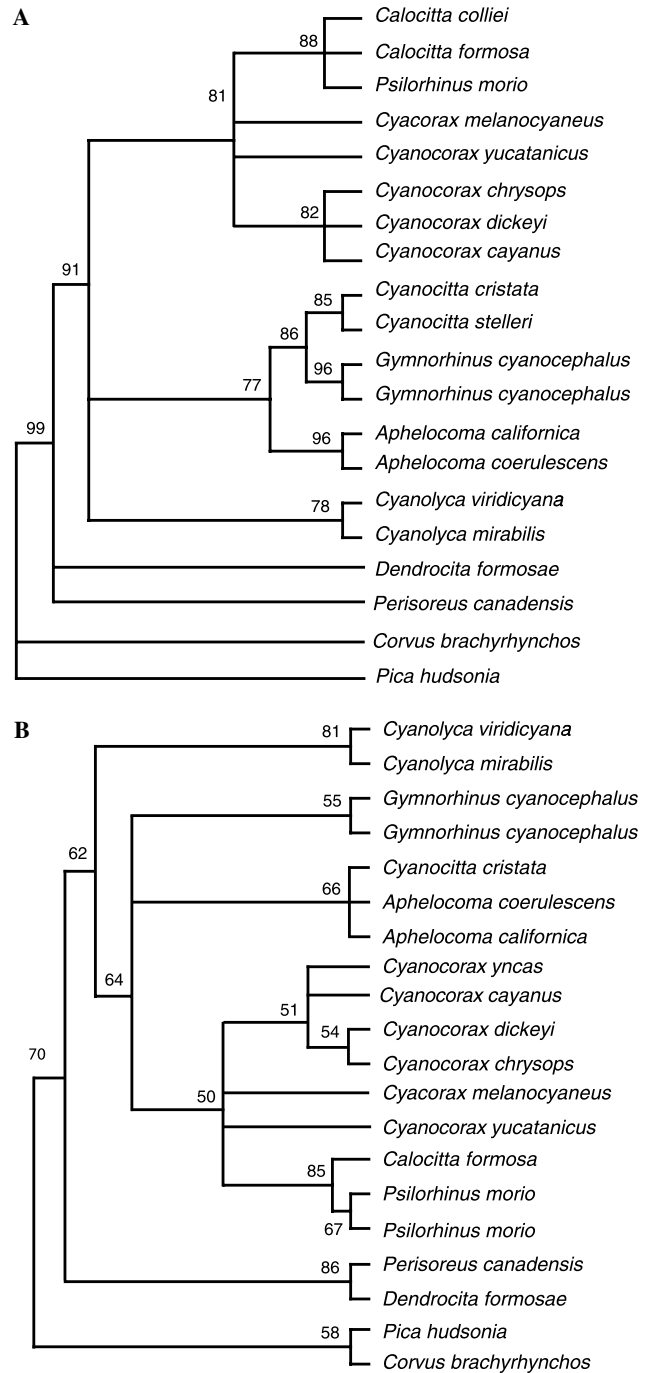


Fig. 3. Maximum parsimony (MP) trees obtained from *βfib7* (strict consensus; A) and AK5 (50% majority rule consensus; B). Numbers indicate MP bootstrap support.

ML and BA produced the (AC) topology probably because of correction of saturation at high levels of sequence divergence. In contrast, AK5 and *βfib7* support their topologies with 1 and 3 unreversed changes, respectively. Given that topological incongruence was restricted to ACG, we were confident of the appropriateness of combining all datasets in a single analysis. The combined analysis produced the same topology as the ND2 ML and BA trees (Fig. 4).

Table 2

Consistency indexes and nodal support for groupings among *Aphelocoma* (A), *Cyanocitta* (C), *Gymnorhinus* (G) and *Cyanocorax* (Cx), obtained when analyzing individual genes using MP, ML, and BA

Dataset	CI	RC	Analysis	(G, A, C)	(G, Cx)	(A, C)	(A, G)	(G, A)
ND2	0.45	0.24	MP	33	—	—	48	—
			ML	82	—	64	—	—
			BA	0.99	—	0.81	—	—
cytb	0.49	0.27	MP	54	—	—	54	—
			ML	—	—	—	41	—
			BA	—	—	—	0.65	—
CR	0.50	0.26	MP	—	78	60	—	—
			ML	—	87	62	—	—
			BA	—	0.91	0.86	—	—
βFb7	0.87	0.73	MP	77	—	—	—	86
AK5	0.91	0.74	MP	30	—	65	—	—

3.4. 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. 5). 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.

4. Discussion

4.1. Phylogenetic analyses

Our combined phylogenetic tree (Fig. 5) depicts intergeneric relationships among NWJs. Relatively weak nodes of the tree are restricted to one clade—that including *Aphelocoma*, *Cyanocitta*, and *Gymnorhinus* (the “ACG” clade). The low support for the monophyly of ACG is a consequence of CR, which supports a topology that differs from those of the βf1b7, ND2, and cytb trees. Since CR is physically linked to the other mitochondrial loci, its different topology should not be the result of independent phylogenetic history, and does not seem attributable to homoplasy, given that characters supporting the CR topology have similar levels of homoplasy (average CI = 0.41, range = 0.167–1.000) as those supporting ACG in ND2 and cytb (average CIs = 0.46 and 0.42, respectively, ranges = 0.2–1.000 and 0.167–0.667, respectively). One possible explanation is that CR shows high rate heterogeneity across lineages, which may obscure phylogenetic signal (Wendel and Doyle, 1998).

The second weak node appears to be an example of loci with different phylogenetic histories. Whereas combined analysis of the mitochondrial dataset (not shown) and AK5 recovered *Aphelocoma* + *Cyanocitta*, βf1b7 recovered *Cyanocitta* + *Gymnorhinus*. Differences in underlying phylogenetic histories between these two sets

Table 3

Synapomorphic character state changes among the genera *Aphelocoma* (A), *Cyanocitta* (C), and *Gymnorhinus* (G) supporting the three possible branching patterns

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
		cytb	cytb	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		
C,G	AK5	398	C → G	TV	1.000		
		Mean ± SD					1.000
C,G	βf1b7	16	T → G	TV	1.000		
		249	T → C	TI	1.000		
		602	A → G	TI	1.000		

TI, transitions; TV, transversions; DI, domain I; DII, domain II; CD, central domain.

of characters may be the result of deviation of the gene tree from the species tree, which may be caused by a variety of processes (e.g., paralogy, lineage sorting of ancestral polymorphisms; de Queiroz, 1993; Wendel and Doyle, 1998).

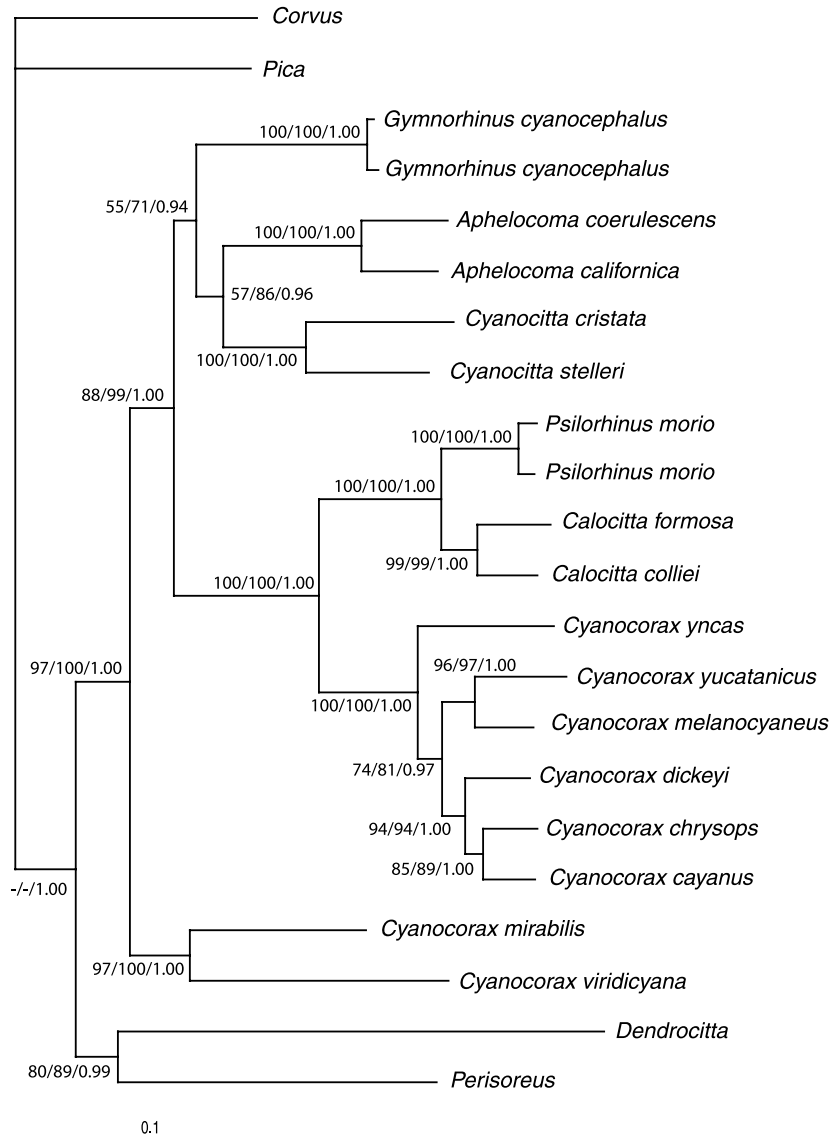


Fig. 4. Maximum likelihood tree that resulted from the combined analysis of mitochondrial (ND2, *cytb*, and CR) and nuclear (*βfib7* and AK5) genes. Numbers over nodes indicate maximum parsimony (MP) bootstrap support of the 50% majority rule consensus tree/maximum likelihood (ML) bootstrap support/and posterior probability values obtained for the Bayesian analysis 50% majority rule consensus tree; “-” indicates nodes not supported by MP and ML.

Interestingly, the (*Aphelocoma*, *Cyanocitta*) arrangement coincides with a novel morphological trait. Curtis and Miller (1938) described a unique bar lateral to the sclerotic ring in *Cyanocitta*. This feature was present in >200 *Cyanocitta* and >600 *Aphelocoma* dissected, but absent in all other NWJs ($N > 100$) examined (A.T. Peterson, unpubl. data). Therefore, this morphological novelty unique among birds serves as a synapomorphy uniting *Aphelocoma* and *Cyanocitta*, to the exclusion of *Gymnorhinus*.

Most other NWJ relationships reconstructed are well-supported, and confirm results of previous phylogenetic studies based on individual genes. *Cyanolyca* is basal on the NWJ tree, and *Cyanocorax* and *Calocitta* + *Psilorhinus* are reciprocally monophyletic. As in previous analyses based on CR sequences (Saunders and Edwards, 2000), ND2 and *cytb* data support *Psilorhinus* as a valid taxon. Consistent

with these results, Sutton and Gilbert (1947) described a unique morphological character in *Psilorhinus*: the “furcular pouch,” a structure formed by the hypertrophy 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 that coincides with the molecular results, diagnosing the genus *Psilorhinus* (Sutton and Gilbert, 1947). For these reasons, *Psilorhinus* should be consistently recognized as a valid genus, rather than submerged in *Cyanocorax* (as suggested by A.O.U., 1983).

Several interesting phylogenetic hypotheses were recovered within *Cyanocorax*. First, *C. melanocyaneus* and *C. yucatanicus* were reconstructed as sister species, which is coherent with their previous recognition as a separate genus, *Cissilopha* (which may be merited once more species

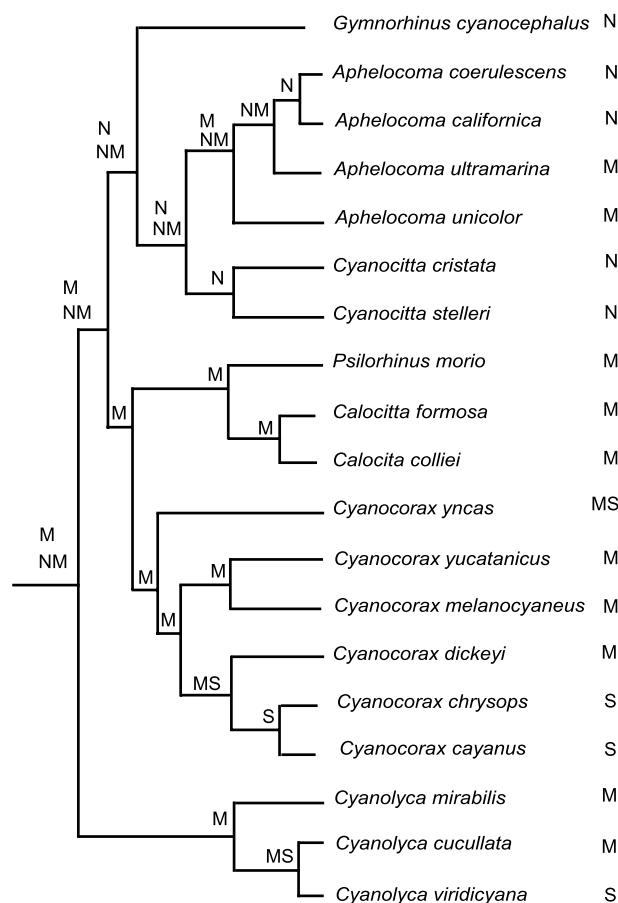


Fig. 5. Optimization of ancestral areas—North America (N), Mesoamerica (M), and South America (S)—using DIVA. Taxon representation over our combined tree was expanded based on previous analyses of *Aphelocoma* (Rice et al., 2003) and ND2 sequence data on *Cyanolyca* (E. Bonaccorso, unpublished data).

are added to the tree). 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*), which has interesting implications for morphological evolution in the clade.

Of these relationships, only *Cyanocorax dickeyi* + *C. chrysops* and the *C. dickeyi* + *C. chrysops* + *C. cayanus* were recovered by AK5 and β fib7, respectively. This result is perhaps understandable, given the relatively short internodal branches following speciation events 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 indeed a correct approximation of the *Cyanocorax* tree. Of course, these initial hypothesis need to be confirmed in a more comprehensive analysis, adding the remaining

two “*Cissilopha*” species (*C. sanblasianus* and *C. beechei*) and the rest of the South American species, as well as multiple individuals of all species.

4.2. Historical biogeography of the NWJs

According to the DIVA analysis, our phylogenetic reconstruction makes biogeographic sense. 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 South American *Cyanocorax* each represents 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. Nonetheless, that two radiations (*Cyanocorax* and *Cyanolyca*) versus some populations of a single species (*Cyanocorax yncas*) would result from the same event might seem unlikely. Given the rate heterogeneity known in this group (Peterson, 1992) and demonstrated here, we reframed from the usual exercise of assigning dates to splits 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, more corvid outgroups should be included in the analysis. Given the levels of saturation observed in our mitochondrial genes at high sequence divergences, this question needs to be approached using multiple, fast-evolving nuclear loci. In any case, it seems that the origin of the NWJs probably looks back to an Asian ancestor.

Acknowledgments

We are grateful to A.G. Navarro-Sigüenza (Museo de Zoología, Universidad Nacional Autónoma de México), S. Hackett and D.E. Willard (Field Museum), S. Birks (Burke Museum of Natural History and Culture), and J. Crraft and P. Sweet (American Museum of Natural History), and M.B. Robbins (University of Kansas Natural History Museum) for providing tissue samples under their care. A.G. Navarro-Sigüenza and B. Hernández-Baños made possible sequencing of *Cyanocorax dickeyi* in Mexico. Two anonymous reviewers made useful comments. This study was funded by grants from the National Science Foundation Dissertation Improvement Grant program (to Bonaccorso; DEB-0508910) and the University of Kansas General Research Fund (to Peterson).

Appendix A

List of samples and gene sequences included in this study

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

UNAM (Facultad de Ciencias, 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 those specimens sequenced for this study. References for published sequences: ^aEspinosa de los Monteros and Cracraft (1997), ^bCicero and Johnson (2001), ^cHelm-Bychowski and Cracraft (1993), ^dEricson et al. (2005), ^eSaunders and Edwards (2000). *Could not be sequenced.

Appendix B

Primers used in this study

Gene	Primer	Sequence (5' to 3')	Source
ND2	L5216	GGCCCATACCCCCGRAAATG	Sorenson et al. (1999)
	H6313	ACTCTTRTTAAGGCTTGAAGGC	Sorenson et al. (1999)
cytb	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)
βf7	FIB-B17U	GGAGAAAACAGGACAATGACAATTCAC	Prychitko and Moore (1997)
	FIB-B17L	TCCCCAGTAGTATCTGCCATTAGGGTT	Prychitko and Moore (1997)

References

- Amadon, D., 1944. The genera of Corvidae and their relationships. *Am. Mus. Nov.* 1251, 1–21.
- A.O.U., 1983. Check-list of North American birds, sixth ed. American Ornithologists Union, Washington DC [with supplements through 1993].
- Barker, F.K., Lutzoni, F.M., 2002. The utility of the incongruence length difference test. *Syst. Biol.* 51, 625–637.
- BirdLife International, 2000. Threatened Birds of the World, Lynx Editions, Barcelona.
- Brown, J.L., 1963. Social organization and behavior of the Mexican Jay. *Condor* 65, 126–153.
- Cicero, C., Johnson, N.K., 2001. Higher level phylogeny of vireos (Aves: Vireonidae) based on sequences of multiple mtDNA genes. *Mol. Phyl. Evol.* 20, 27–40.
- Cunningham, C.W., 1997. Can three incongruence tests predict when data should be combined? *Mol. Biol. Evol.* 14, 733–740.
- Curtis, E.L., Miller, R.C., 1938. The sclerotic ring in North American birds. *Auk* 55, 225–243.
- de Queiroz, A., 1993. For consensus (sometimes). *Syst. Biol.* 42, 368–372.
- Edwards, S.V., Naem, S., 1993. The phylogenetic component of cooperative breeding in perching birds. *Am. Nat.* 141, 754–789.
- Ericson, P.G.P., Jansén, A.L., Johansson, U.S., Ekman, J., 2005. Intergeneric relationships of the crows, jays, magpies and allied groups (Aves: Corvidae) based on nucleotide sequence data. *J. Avian Biol.* 36, 222–234.
- Espinosa de los Monteros, A., Cracraft, J., 1997. Intergeneric relationships of the New World Jays inferred from cytochrome b gene sequences. *Condor* 99, 490–502.
- Farris, J.S., Källersjö, M., Kluge, A.G., Bult, C., 1994. Testing significance of congruence. *Cladistics* 10, 315–319.

- Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39, 783–791.
- Felsenstein, J., 1988. Phylogenies from molecular sequences: inference and reliability. *Ann. Rev. Genet.* 22, 521–565.
- Fitzpatrick, J.W., Woolfenden, G.E., 1985. The Florida Scrub Jay: Demography of a Cooperative-breeding Bird. Princeton University Press, Princeton.
- Gene Codes Corporation, 2000. Sequencher Version 4.1. Gene Codes Corporation, Ann Arbor, Michigan.
- Hardy, J.W., 1961. Studies in behavior and phylogeny of certain New World Jays (Garrulinae). *Univ. of Kansas Sci. Bull.* 42, 13–149.
- Helm-Bychowski, K., Cracraft, J., 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. *Mol. Biol. Evol.* 10, 1196–1214.
- Kocher, T.D., Thomas, W.K., Meyer, A., Edwards, S.V., Paabo, S., Villablanca, F.X., Wilson, A.C., 1989. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proc. Nat. Acad. Sci. USA* 86, 6196–6200.
- Maddison, D.R., Maddison, W.P., 2000. MacClade: Analysis of Phylogeny and Character Evolution. Vers. 4.0. Sinauer, Sunderland, MA.
- Moore, W.S., 1995. Inferring phylogenies from mtDNA variation: Mitochondrial-gene trees versus nuclear-gene trees. *Evolution* 49, 718–726.
- Peterson, A.T., 1992. Phylogeny and rates of molecular evolution in the jays of the genus *Aphelocoma* (Corvidae). *Auk* 109, 134–148.
- Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics* 14, 817–818.
- Posada, D., Buckley, T.R., 2004. Model selection and model averaging in phylogenetics: advantages of akaike information criterion and Bayesian approaches over likelihood ratio tests. *Syst. Biol.* 53, 793–808.
- Prychitko, T.M., Moore, W.S., 1997. The utility of DNA sequences of an intron from the β -fibrinogen gene in phylogenetic analysis of woodpeckers (Aves: Picidae). *Mol. Phyl. Evol.* 8, 193204.
- Rice, N.H., Martı́nez-Meyer, E., Peterson, A.T., 2003. Ecological niche differentiation in the *Aphelocoma* jays: a phylogenetic perspective. *Bio. J. Linnean Soc* 80, 369–383.
- Ronquist, F., 1996. DIVA, ver. 1.1. Computer program available at <http://www.abc.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. *Syst. Biol.* 49, 195–203.
- Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572–1574.
- Saunders, M.A., Edwards, S.V., 2000. Dynamics and phylogenetic implications of mtDNA control region sequences in New World Jays (Aves: Corvidae). *J. Mol. Evol.* 51, 97–109.
- Shapiro, L.H., Dumbacher, J.P., 2001. Adenylate kinase intron 5: a new nuclear locus for avian systematics. *Auk* 118, 248–255.
- Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. and Evol.* 16, 1114–1116.
- Soltis, P.S., Soltis, D.E., Savolainen, V., Crane, P.R., Barraclough, T.G., 2002. Rate heterogeneity among lineages of tracheophytes: integration of molecular and fossil data and evidence for molecular living fossils. *Proc. Nat. Acad. Sci.* 99, 4430–4435.
- Sorensen, J.D., Quinn, T.W., 1998. Numts: A challenge for avian systematics and population biology. *Auk* 115, 214–221.
- Sorenson, M.D., Ast, J.C., Dimcheff, D.E., Yuri, T., Mindell, D.P., 1999. Primers for a PCR-based approach to mitochondrial genome sequencing in birds and other vertebrates. *Mol. Phyl. Evol.* 12, 105–114.
- Sutton, J.M., Gilbert, P.W., 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.
- Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acid Res.* 25, 4876–4882.
- Wendel, J.F., Doyle, J.J., 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.
- Zusi, R.L., 1987. A feeding adaptation of the jaw articulation in New-World Jays (Corvidae). *Auk* 104, 665–680.