www.sciencemag.org/cgi/content/full/320/5884/1763/DC1

Supporting Online Material for

A Phylogenomic Study of Birds Reveals Their Evolutionary History

Shannon J. Hackett, Rebecca T. Kimball,* Sushma Reddy, Rauri C. K. Bowie, Edward L. Braun, Michael J. Braun, Jena L. Chojnowski, W. Andrew Cox, Kin-Lan Han, John Harshman, Christopher J. Huddleston, Ben D. Marks, Kathleen J. Miglia, William S. Moore, Frederick H. Sheldon, David W. Steadman, Christopher C. Witt, Tamaki Yuri

*To whom correspondence should be addressed. E-mail: rkimball@ufl.edu

Published 27 June 2008, *Science* **320**, 1763 (2008)

DOI: 10.1126/science.1157704

This PDF file includes:

Supporting online materials

Materials and Methods

DNA Extraction, PCR Amplification, Sequencing

DNA for 171 samples (Table S1) was isolated by proteinase K digestion followed by phenol-chloroform extraction with standard protocols or with DNeasy kits (Qiagen, Valencia, CA) as directed by the manufacturer. Each sample was isolated in one of three laboratories (NMNH, FMNH, LSU) before being distributed to all five laboratories (NMNH, FMNH, LSU, UF, WSU) for amplification and sequencing of individual loci (Table S2).

PCR amplification of selected loci (Table S2) was performed with locus-specific primers. Published primers were used to amplify BDNF (S*1*), CLTC (S*2*), NGF (S*1*), EGR1 (S*3*), GH1 (S*4*), MB (S*5, S6*), MUSK (S*7*), NTF3 (S*1*), TGFB2 (S*8*), TPM1 (S*8*). For some loci, primers to amplify part of the locus were previously published. This included FGB, exons 4 to 5 (S*9*), FGB, exons 5 to 6 (S*10*), HMGN2, exons 2 to 4 (S*11*), and RHO exons 1 to 2 (S*11*); some primers for MYC were also published (S*12*). Unpublished primers and amplification conditions can be found in Table S3. Novel primers were designed by manual comparison of conserved regions or by using Primer 3 (S*13*).

Cycle sequencing of PCR products was performed with BigDye 3.1 chemistry and capillary sequencing instruments (Applied Biosystems, Foster City, CA). Allelic length polymorphisms due to indels made it necessary to clone PCR products of some individuals at many loci to obtain clean sequences. Contigs were assembled with Sequencher (Genecodes, Ann Arbor, MI), with more than 95% of all nucleotide positions determined on both strands of DNA.

Alignment and Vetting of Data

Sequences were aligned with a combination of ClustalX (S*14*) and manual alignment, with criteria of maximizing similarity and minimizing the number of inferred evolutionary events (S*15*). After an initial alignment of 70 taxa was assembled, a consortium consisting of representatives from multiple labs met to evaluate all alignments and ensure all labs were using similar alignment criteria. As a group, we identified regions that were ambiguously aligned, such as homopolymer stretches, and were later excluded from most analyses. For many intron regions, homology between the crocodilian outgroups and birds could not be reliably assessed, therefore the crocodilian data for these regions were not included in the final alignment. After sequencing of the total taxon set (N=171), individual genes were vetted (see below) and aligned by their respective labs following the guidelines established after 70 taxa had been aligned. Each alignment was then sent to a second lab for review and modification and then re-assessed by the original lab. Once the final data set was assembled, members of two different labs reviewed all alignments to ensure that problematic regions were correctly identified, and that intron/exon boundaries had been correctly identified. In total, alignments were examined at least four times by multiple individuals prior to concatenation.

To ensure that errors were identified and removed from the data, all data was subject to several different vetting methods. Our taxon-sampling strategy ensured that

most taxa had a close relative with which to verify sequences. MEGABLAST (S*16*) was used to compare sequences from the alignments to each other and with the original contigs to verify a match to the original chromatogram and close relatives and to help identify incorrectly labeled sequences, chimeras, or errors introduced during alignment. For each gene region, a UPGMA or NJ tree was generated and evolutionary distances were compared to identify mis-labeled sequences or to identify identical sequences that might represent amplification errors. In addition, NJ trees were used to identify sequences with excessive branch lengths that might represent problems such as pseudogenes. Sequences that were problematic were re-amplified and re-sequenced for confirmation, or else removed from the dataset.

Phylogenetic Analysis

The aligned dataset consisted of 52,383 base pairs for 169 avian and 2 outgroup taxa (Treebase Accession S2079; Table S1). Before conducting phylogenetic analyses, we excluded sites that were present in less than three taxa, which mainly excluded large autapomorphic insertions that greatly contributed to alignment length (and computational time for some analyses) but were not informative for reconstructing phylogenetic relationships. In addition, we excluded inversions and sites where we felt the alignment was ambiguous (e.g., homopolymer runs). The reduced dataset was 31,951 aligned characters. We analyzed both the complete (52kb) and the reduced (32kb) datasets with parsimony (see below), and found few differences in the estimated phylogeny (see Fig. 1). Thus, we restricted the remainder of our analyses to the 32kb alignment, in which 69% of sites were variable and of these 84% were parsimony-informative.

Analyses with the entire dataset (32kb or 52kb alignments) were done with all taxa, as well as with runs excluding the crocodilian outgroups [rooting to the split between Paleognathes and Neognathes (Fig. S2)]. Analyses with subsets of the data matrix were conducted without the crocodilian outgroups, since these were not present (or confidently aligned) for many data partitions. We conducted phylogenetic analyses using maximum parsimony (MP), maximum likelihood (ML), and Bayesian inference (BI).

We used PAUP* v.4.0b10 (S*17*) for MP analyses. We analyzed the 52kb and 32kb datasets by conducting a heuristic search of 1000 random-addition replicates with TBR branch-swapping. In addition, we also analyzed various partitioned sets (i.e. each gene region, functional regions [i.e. introns, coding exons, and UTRs], and removing one-gene at time from complete alignment) using heuristic searches of 500 random-addition replicates. Bootstrap support was evaluated with 500 pseudo-replicates of 5 random-additions each. Additional analyses using methods aimed at analyzing large taxonomic datasets such as the Parsimony-Ratchet (S*18*) and TNT (S*19*) did not yield more optimal trees.

ML analyses were conducted with GARLI v.0.94 (S*20*), RAxML VIHPC (S*21*), and PhyML (S*22*) for the 32kb dataset and each gene partition. In GARLI and RAxML, we conducted a minimum of five runs until at least two searches from different starting points converged within two likelihood units of the best tree. If, convergence was not achieved we conducted a minimum of 20 runs. For each analysis, we chose the results from the best run. Both GARLI and RAxML were used in order to compare the results from these two programs as they use different approaches to speed up ML searches.

Analyses from both ML programs resulted in identical to highly similar topologies. We also tested PhyML but did not use the results because we found that GARLI and RAxML consistently found trees with better likelihood scores. ML nodal support was calculated by analyzing 100 bootstrap replicates in GARLI and RAxML. Partitioned likelihood and RY-coded analyses were conducted in RAxML, both to obtain the best tree and for 100 bootstrap replicates. For the partitioned analysis, each locus was defined as a separate partition whose parameters were allowed to vary independently. In RY-coded analyses, nucleotide bases were translated into purines or pyrimidines.

MrBayes v.3.11 (S*23*) was used for BI searches. We attempted both partitioned (with unlinked rate parameters across the 19 gene regions, while branch lengths remained linked) and unpartitioned analysis. We used two different strategies for the partitioned analysis, both of which were unsuccessful in approaching stationarity. First we tried two simultaneous runs with six heated chains each for 10 million generations, saving each 500th generation (Analysis A). Next we conducted a search with six runs of two chains each, one cold and one heated, for 10 million generations (Analysis B). Both searches, performed at the supercomputing facilities at the Illinois Bio-grid at DePaul University, took over 2 months each. In the end, none of the different runs had either converged or reached stationarity (Fig. S3). Failure of Bayesian analyses to satisfactorily complete a phylogenetic analysis of large datasets has also been demonstrated recently by (S*24*), though that study included more taxa and fewer gene partitions. Unpartitioned analyses of our data set always immediately crashed, regardless of the memory capacity of the computers used. Undoubtedly, Bayesian analysis of large datasets is a complex problem such that the number of rearrangements needed to find optimal trees may be so large (on the order of hundreds of millions to billions of generations for a dataset of our size) (S*25*) that is it not currently computationally feasible.

We compared the concordance of the relationships estimated by the different phylogenetic methods and data partitions to each other in order to quantify the level of congruence among these analyses. We tallied the presence of nodes found in the ML analysis that were also found in the MP analysis and vice versa (see Fig. 1 in main text). Additionally, we also assessed the nodes that the various gene partitions had in common with the combined analysis. There was also a high degree of concordance between nodes found in the combined analyses that were also present in the analyses of the separate gene partitions (see Fig. 1 in main text).

Supporting Online Text

<u>Comparison to other studies</u>

Using ordinal-level classifications allowed a comparison of well-defined groups but did not allow comparisons of higher-level groupings across studies. For this, we calculated the likelihood of alternative phylogenetic topologies using our data. We had to limit our comparisons to (S*26*) and (S*27*) since this analysis required well-resolved and taxonomically comparable phylogenies. To obtain the difference in likelihood, we looked for taxonomic equivalents between our dataset and the published topologies. We then arranged our taxa to match the published topology. Where genera differed, we used the literature to determine the closest match to our taxa. In the few cases where a clear match was not available, we left those taxa as unresolved in order to prevent biasing our results.

The likelihood of the best topology we obtained as well as the alternative topologies was determined using PAUP* using the GTR+I+G model with parameters estimated independently for each tree.

Our phylogenomic dataset showed a substantially worse fit to the DNA-DNA hybridization (Δ *ln* likelihood = 14838.26) and morphological (Δ *ln* likelihood = 13070.54) topologies relative to our ML tree. There were striking differences in the placement of many groups among studies. For example, the base of Neoaves includes some of our landbirds in DNA-DNA hybridization (S*26*) but taxa from our waterbird and shorebird clades in morphology (S*27*).

Evolutionary rates may represent one reason why DNA-DNA hybridization (S*26*) exhibited major differences from our phylogeny. We observed substantial evolutionary rate heterogeneity within and across lineages of birds (Figs. 3, S1). For example, the rapidly evolving shorebird *Turnix* (Fig. S1) (S*28*-*S30*) was recovered at the base of Neoaves with DNA-DNA hybridization. Other groups with accelerated rates of evolution [e.g., Odontophoridae, or New World quail, in Galliformes (*Colinus* in our study), Piciformes and Coraciiformes] are also basal in the DNA-DNA hybridization tree. Analyses of sequence data accommodate heterogeneity of evolutionary rates more effectively than DNA-DNA hybridization, which can be misleading when substantial differences exist (S*31*).

Supporting Figures and Tables:

Figure S1. ML phylogram with species names (same as Fig. 2 in main text) rooted with the crocodilian outgroups.

Figure S2. Avian phylogeny with maximum likelihood (ML) on the 19-locus data set rooted to paleognaths rather than crocodilian taxa. All nodes recovered were identical to the analysis including the non-avian outgroups. Thick branches represent nodes where ML bootstrap support is ≥ 65%.

Figure S3. Lack of convergence in Bayesian chains. Likelihood scores per 500th generation (after discarding the first 2 million generations as burnin) from each different run of MrBayes analyses (see above).

Table S1. List of samples and voucher information.

Table S2. Gene regions sequenced for this project, organized by HUGO gene codes.

Table S3. Unpublished primers and amplification conditions.

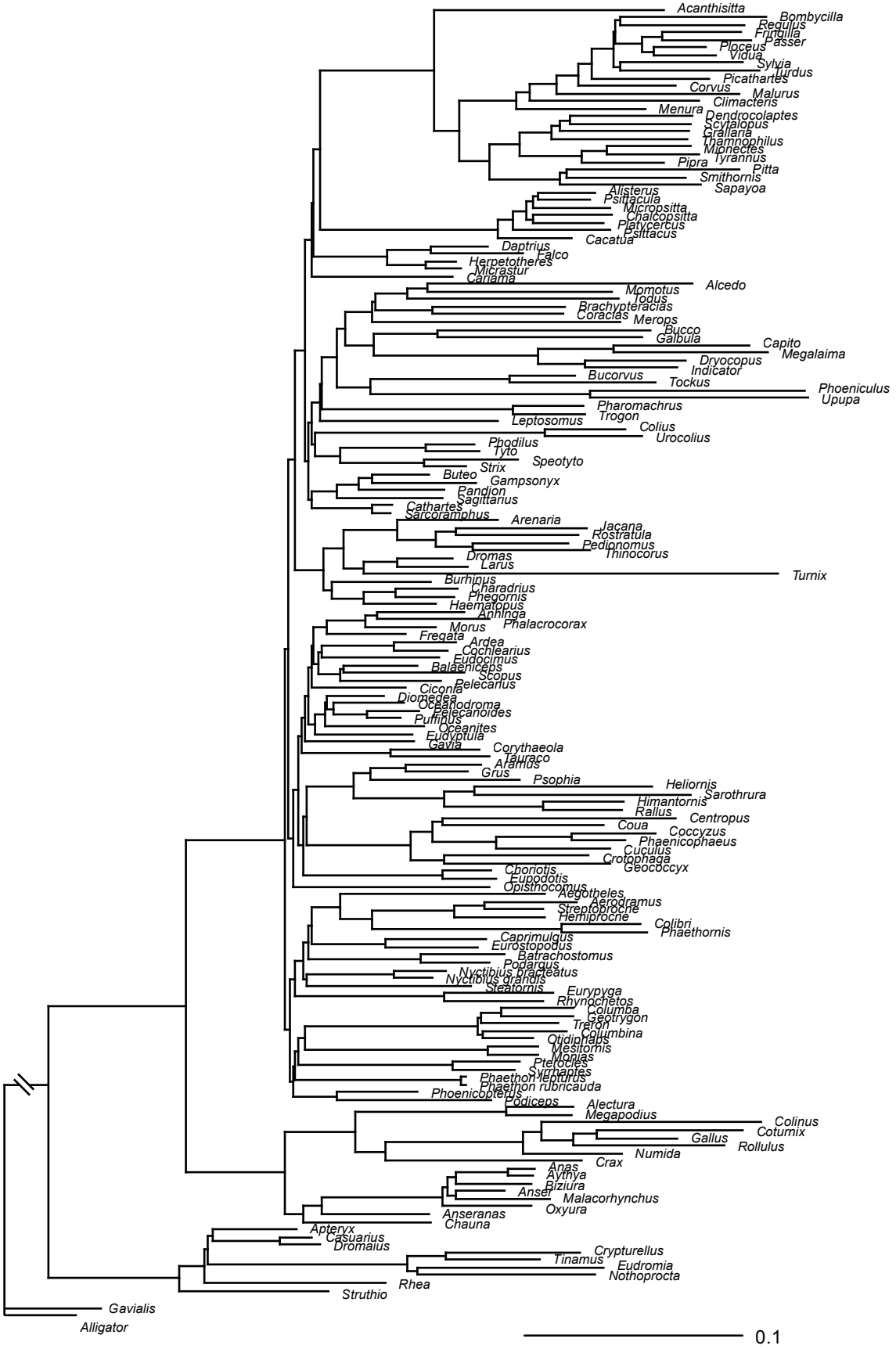

Figure S1. ML phylogram with species names (same as Fig. 2 in main text) rooted with the crocodilian outgroups.

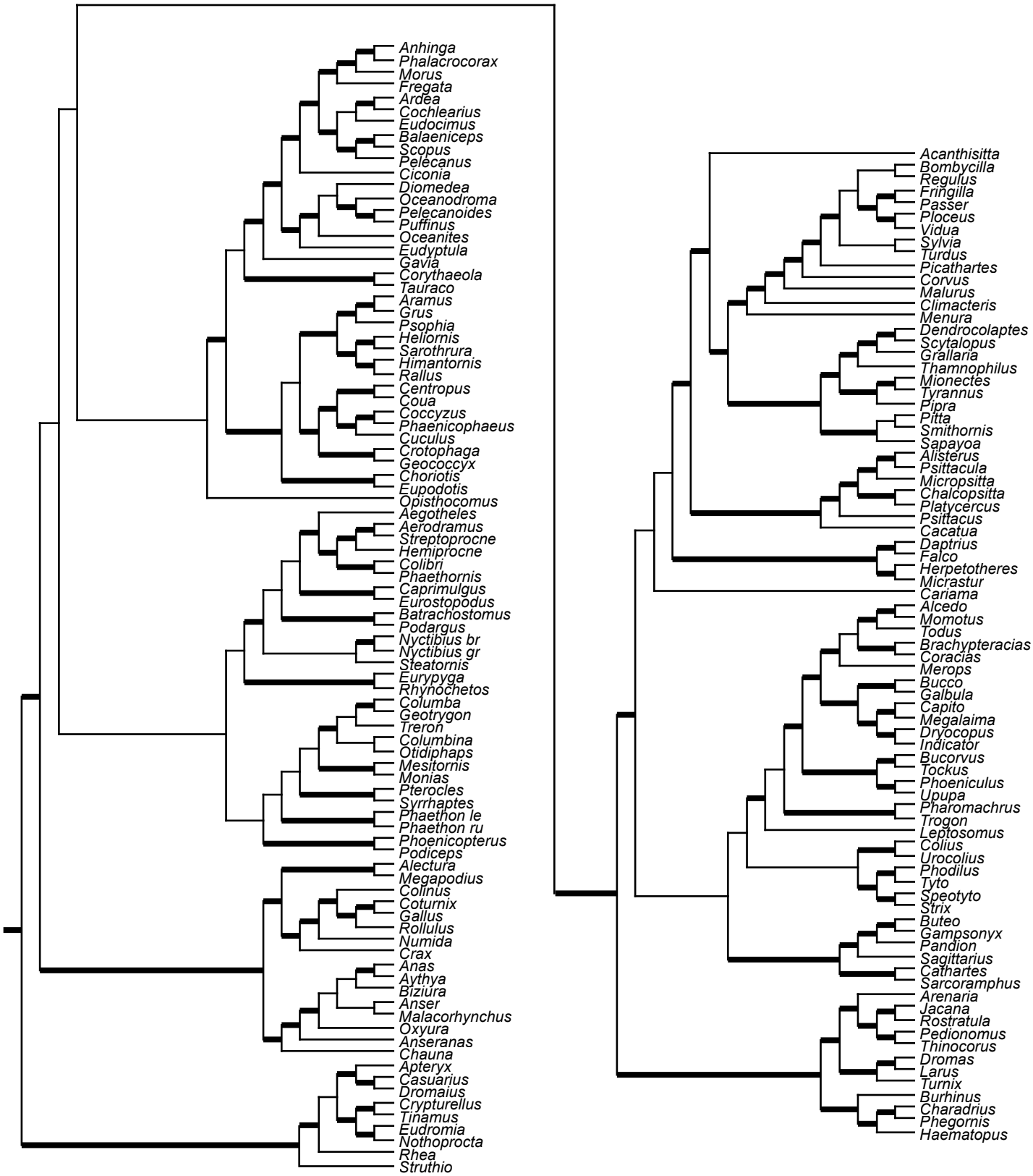

Figure S2. Avian phylogeny with maximum likelihood (ML) on the 19-locus data set rooted to paleognaths rather than crocodilian taxa. All nodes recovered were identical to the analysis including the non-avian outgroups. Thick branches represent nodes where ML bootstrap support is ≥ 65%.

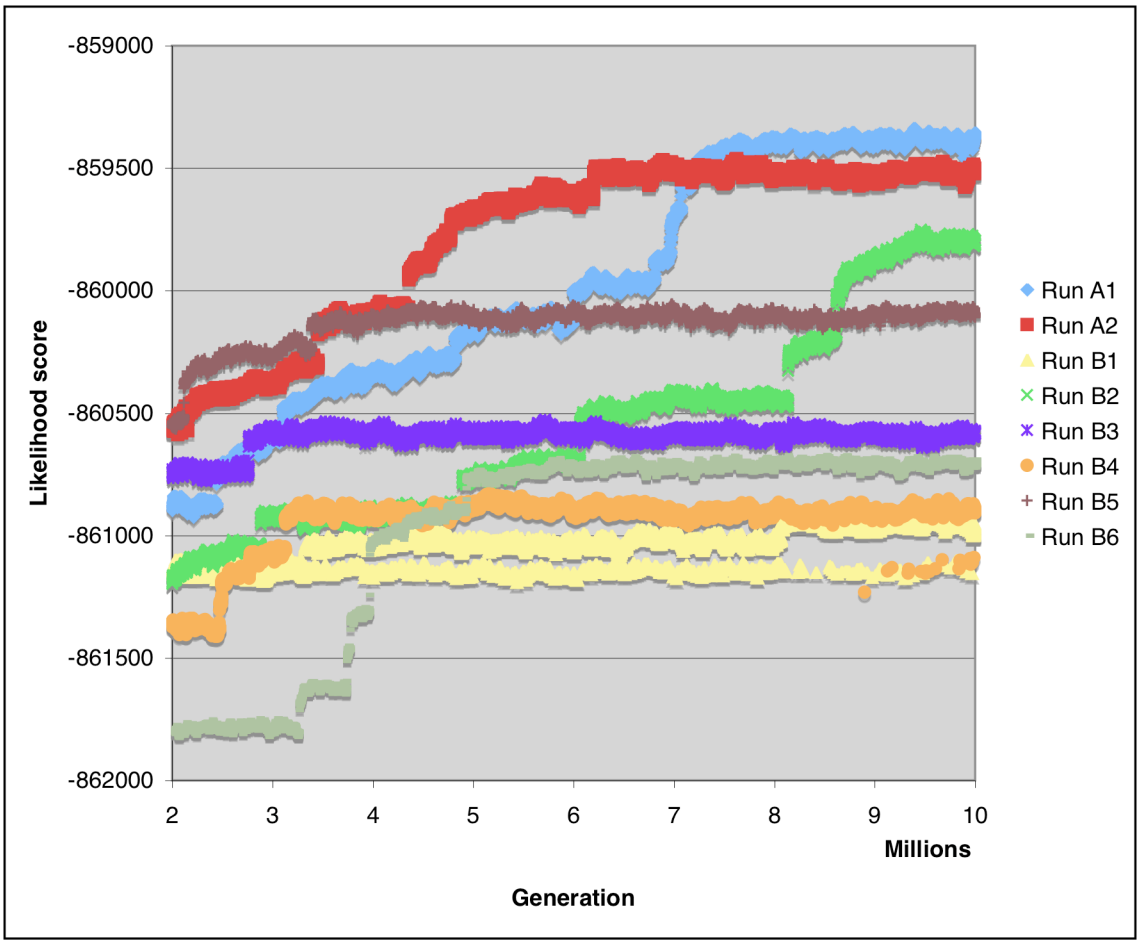

Figure S3. Lack of convergence in Bayesian chains. Likelihood scores per 500th generation (after discarding the first 2 million generations as burnin) from each different run of MrBayes analyses (see above).

Supplementary Table 1. List of samples and voucher information.

Species name	Museum	Catalog/ collection#	Locality	Collector
Acanthisitta chloris	MVZ	No voucher	New Zealand	C. Sibley collection (#1023)
Aegotheles insignis	KUMNH	95997	Papua New Guinea	
Aerodramus vanikorensis	USNM	608672	Papua New Guinea	J. P. Angle & B. M. Beehler
Alcedo leucogaster	FMNH	429397	Central African Republic	D. E. Willard
Alectura lathami	USNM	612658	Australia	V. Davies
Alisterus scapularis	USNM	612684	Australia	A. Hiller
Anas platyrhynchos	FMNH	398624	Illinois, USA	T. Valentino
Anhinga anhinga	USNM	622528	Florida, USA	M. E. Kennedy
Anser erythropus	LSUMNS	19457	Captive	
Anseranas semipalmata	USNM	621019	Captive	
Apteryx australis	LSUMNS	8606	Captive	
Aramus guaranus	FMNH	385742	Florida, USA	
Ardea herodias	FMNH	385673	Florida, USA	
Arenaria interpres	USNM	607650	Panama	T. J. Parsons
Aythya americana	LSUMNS	2651	USA	D. L. Dittman
Balaeniceps rex	LSUMNS	13372	Captive	
Batrachostomus septimus	FMNH	429205	Philippines	
Biziura lobata	ANWC	50099	Australia	
Bombycilla garrulus	USNM	621231	Russia (USSR)	A. B. Sokolov
Brachypteracias squamigera	FMNH	384758	Madagascar	S. M. Goodman
Bucco macrodactylus	FMNH	433243	Peru	T. Pequeno
Bucorvus abyssinicus	SFZG	201050	Captive (live blood)	San Francisco Zoological Garden
Burhinus bistriatus	FMNH	389169	Brazil	D. F. Stotz
Buteo jamaicensis	LSUMNS	33264	Louisiana, USA	
Cacatua sulphurea	USNM	542225	Captive	
Capito niger	KUMNH	88938	Guyana	M. B. Robbins
Caprimulgus longirostris	LSUMNS	32361	Peru	D. G. Christian
Cariama cristata	LSUMNS	8656	Captive	
Casuarius casuarius	LSUMNS	10202	Captive	P. P. Marra
Cathartes aura	USNM	625337	Guyana	B. K. Schmidt
Centropus viridis	USNM	607313	Philippines	C. A. Ross, R. W. Dickerman
Chalcopsitta cardinalis	UWBM	76322	Solomon Islands	C. E. Smith
Charadrius vociferus	FMNH	430574	Wisconsin, USA	Bay Beach Wildlife Sanctuary
Chauna torquata	USNM	614546	Argentina	J. P. Angle
Choriotis (Ardeotis) kori	FMNH	380020	Tanzania	Northland Wildlife
Ciconia ciconia	KUMNH	90088		
Climacteris erythrops	ANWC	46255	Australia	
Coccyzus americanus	USNM	626461	Florida, USA	C. A. Gebhard
Cochlearius cochlearius	USNM	621016	Captive	
Colibri coruscans	LSUMNS	5574	Peru	T. J. Davis
Colinus cristatus	USNM	626183		
Colius colius	USNM	622447	Captive	
Columba livia	FMNH	428785	Wisconsin, USA	Bay Beach Wildlife Sanctuary
Columbina passerina	USNM	626037	Guyana	C. M. Milensky
Coracias caudata	USNM	621662	Captive	
Corvus corone	USNM	621224	Russia	B. K. Schmidt
Corythaeola cristata	FMNH	357942	Brundi	P. K. Austin & J. C. Kerbis
Coturnix coturnix	USNM	621265	United Kingdom	T. West
Coua cristata	FMNH	352800	Madagascar	S. M. Goodman
Crax alector	USNM	625104	Guyana	C. M. Milensky
Crotophaga sulcirostris	USNM	607659	Panama	C. Smith
Crypturellus soui	USNM	586295	Guyana	K. S. Bostwick
Cuculus canorus	USNM	621533	Mongolia	D. A. Banin
Daptrius ater	KUMNH	88118	Guyana	C. M. Milensky
Dendrocolaptes certhia	KUMNH	88158	Guyana	M. B. Robbins
Diomedea nigripes	USNM	622480	USA	B. Callahan
Dromaius novaehollandiae	LSUMNS	5895	Captive	S. M. Lanyon
Dromas ardeola	MVZ	No voucher		C. Sibley collection (#799)
Dryocopus pileatus	LSUMNS	20929	Louisiana, USA	T. S. Sillett
Eudocimus albus	FMNH	387623	Florida, USA	
Eudromia elegans	LSUMNS	5893	Captive	D. L. Dittmann
Eudyptula minor	MV	I224	Australia	
Eupodotis ruficrista	MBM	8090	South Africa	J. Klicka

Supplementary Table 1 cont'd. (2 of 3)

Species name	Museum	Catalog/ collection#	Locality	Collector
Eurostopodus macrotis	USNM	607328	Philippines	R. W. Dickerman et al.
Eurypyga helias	USNM	586303	Guyana	M. J. Braun
Falco mexicanus	MSB	103400	New Mexico, USA	
Fregata magnificens	USNM	613348	Panama	S. L. Olson
Fringilla montifringilla	USNM	586194	Russia (USSR)	C. S. Wood
Galbula albirostris	USNM	609136	Guyana	M. J. Braun
Gallus gallus	LSUMNS	19436	Captive	D. L. Dittmann
Gampsonyx swainsonii	LSUMNS	5174	Peru	S. W. Cardiff
Gavia immer	FMNH	363266	Minnesota, USA	Minnesota DNR
Geococcyx californianus	USNM	621303	USA	P. S. Crossin
Geotrygon montana	KUMNH	88129	Guyana	N. Rice
Grallaria varia	USNM	625538	Guyana	M. J. Braun
Grus canadensis	USNM	621384	Captive	
Haematopus ostralegus	LSUMNS	163409?		
Heliornis fulica	FMNH	391998	Brazil	J. Haffer
Hemiprocne mystacea	UWBM	68087	Solomon Islands	L. K. Wang
Herpetotheres cachinnans	USNM	612262	Panama	
Himantornis haematopus	ZMUC	No voucher	Ghana	L. Holbech & J. Fjeldsa
Indicator maculatus	FMNH	429419	Central African Republic	D. E. Willard
Jacana jacana	LSUMNS	7288	Peru	T. C. Maxwell
Larus marinus	FMNH	385745	Florida, USA	
Leptosomus discolor	FMNH	434672	Madagascar	T. S. Schulenberg
Malacorhynchus membranaceus	LSUMNS	29875	Captive	D. L. Dittmann
Malurus melanocephalus	ANWC	29907	Australia	
Megalaima virens	LSUMNS	20788	Captive	S. W. Cardiff
Megapodius eremita	MV	KNG16	Australia	
Menura novaehollandiae	ANWC	42757	Australia	
Merops nubicus	AMNH	9269*	Captive	P. R. Sweet
Mesitornis unicolor	FMNH	345610	Madagascar	S. M. Goodman
Micrastur semitorquatus	LSUMNS	11298	Peru	D. C. Schmitt
Micropsitta finschii	USNM	615016	Papua New Guinea	J. P. Angle & B. M. Beehler
Mionectes macconnelli	USNM	586411	Guyana	R. T. Brumfield
Momotus momota	USNM	609134	Guyana	M. J. Braun
Monias benschi	FMNH	438526	Madagascar	M. J. Raherilalao
Morus bassanus	LSUMNS	23898	Louisiana, USA	D. L. Dittmann
Nothoprocta perdicaria	LSUMNS	23841	Captive	D. L. Dittmann
Numida meleagris	FMNH	384667	Madagascar	S. M. Goodman
Nyctibius bracteatus	LSUMNS	4509	Peru	S. W. Cardiff
Nyctibius grandis	LSUMNS	15415	Bolivia	J. M. Bates
Oceanites oceanicus	LSUMNS	37197	Louisiana, USA	D. L. Dittmann
Oceanodroma tethys	USNM	613258	Pacific Ocean	L. Spear
Opisthocomus hoazin	LSUMNS	9660	Bolivia	D. C. Schmitt
Otidiphaps nobilis	USNM	621030	Captive	
Oxyura jamaicensis	FMNH	397634	Wisconsin, USA	Bay Beach Wildlife Sanctuary
Pandion haliaetus	FMNH	385893	Florida, USA	Archibold Biological Station
Passer montanus	FMNH	347952	Pakistan	S. M. Goodman & P. Myers
Pedionomus torquatus	ANWC	44673	Australia	
Pelecanoides urinatrix	AMNH	2583*	New Zealand	
Pelecanus occidentalis	USNM	621489	USA	
Phaenicophaeus curvirostris	LSUMNS	47001	Malaysia	R. G. Moyle
Phaethon lepturus	LSUMNS	35134	USA	D. L. Dittmann
Phaethon rubricauda	FMNH	346042	Micronesia	S. M. Lanyon
Phaethornis griseogularis	LSUMNS	39763	Peru	J. P. O'Neill
Phalacrocorax carbo	LSUMNS	45740	Kuwait	J. M. Bishop
Pharomachrus auriceps	FMNH	433224	Peru	B. J. O'Shea
Phegornis mitchelli	LSUMNS	103926	Peru	T. S. Schulenberg
Phodilus badius	AMNH	10244*	Captive (live blood)	P. R. Sweet
Phoenicopterus chilensis	USNM	614545	Argentina	J. P. Angle
Phoeniculus purpureus	LSUMNS	39521	Ghana	B. D. Marks
Picathartes gymnocephalus	LSUMNS	18002	Captive	D. C. Moyer
Pipra coronata	FMNH	433694	Peru	D. F. Stotz
Pitta guajana	LSUMNS	36368	Malaysia	R. G. Moyle
Platycercus elegans	USNM	612685	Australia	D. Davidson

Supplementary Table 1 cont'd. (3 of 3)

Species name	Museum	Catalog/ collection#	Locality	Collector
Ploceus cucullatus	FMNH	429881	Democratic Republic of Congo	J. M. Bates
Podargus strigoides	USNM	612702	Australia	F. Smith & T. Smith
Podiceps auritus	LSUMNS	19296	Florida, USA	A. Chapman-Kofron
Psittacula alexandri	LSUMNS	36555	Captive	D. L. Dittmann
Psittacus erithacus	FMNH	363153	Tanzania	Birds Haven
Psophia crepitans	USNM	621709	Guyana	M. J. Braun
Pterocles namaqua	No voucher		Republic of South Africa	
Puffinus griseus	USNM	613236	Pacific Ocean	L. Spear
Rallus limicola	FMNH	380034	Illinois, USA	D. E. Willard
Regulus calendula	USNM	601728	USA	P. R. Windler
Rhea americana	USNM	541231	Captive	
Rhynochetos jubatus	FLMNH	42714	Captive	
Rollulus rouloul	LSUMNS	24971	Captive	D. L. Dittmann
Rostratula benghalensis	FMNH	358283	Philippines	T. P. Gnoske
Sagittarius serpentarius	USNM	621021	Captive	
Sapayoa aenigma	LSUMNS	2333	Panama	S. M. Lanyon
Sarcoramphus papa	USNM	623238	Captive	C. M. Milensky
Sarothrura elegans	FMNH	346189	Burundi	J. C. Kerbis & A. J. Fisher
Scopus umbretta	LSUMNS	28330	Captive	D. L. Dittmann
Scytalopus magellanicus	LSUMNS	8346	Peru	
Smithornis rufolateralis	FMNH	429425	Central African Republic	D. E. Willard
Speotyto cunicularia	FMNH	396871	Florida, USA	
Steatornis caripensis	LSUMNS	32579	Peru	C. C. Witt
Streptoprocne zonaris	USNM	626065	Guyana	C. M. Milensky
Strix occidentalis	USNM	608652	USA	R. W. Skaggs
Struthio camelus	LSUMNS	1526	Captive	D. L. Dittmann
Sylvia nana	USNM	586676	Mongolia	B. K. Schmidt
Syrrhaptes paradoxus	USNM	586725	Mongolia	B. Shagdarsuren
Tauraco erythrolophus	LSUMNS	5354	Captive	D. L. Dittmann
Thamnophilus nigrocinereus	LSUMNS	20234	Brazil	M. Cohn-Haft
Thinocorus orbignyianus	USNM	616270	Argentina	B. K. Schmidt
Tinamus guttatus	FMNH	389673	Brazil	B. D. Patterson
Tockus camurus	USNM	616704	Gabon	B. K. Schmidt
Todus angustirostris	AMNH	6949*	Dominican Republic	N. K. Klein
Treron vernans	LSUMNS	47229	Malaysia	F. H. Sheldon
Trogon personatus	LSUMNS	7644	Peru	G. H. Rosenberg
Turdus falklandii	AMNH	10421*	Argentina	P. R. Sweet
Turnix sylvatica	AMNH	5860*	Republic of South Africa	J. G. Groth
Tyrannus tyrannus	USNM	586080	USA	B. K. Schmidt
Tyto alba	LSUMNS	19295	Louisiana, USA	J. M. Bates
Upupa epops	USNM	586147	Russia (USSR)	A. Vanushkin
Urocolius indicus	LSUMNS	34225	South Africa	R. G. Moyle
Vidua chalybeata	LSUMNS	39547	Ghana	B. D. Marks
Alligator mississippi	LSU HSC		Captive	H. Dessauer
Gavialis gangeticus	LLD	1001871		L. Densmore

* AMNH DOT (Department of Ornithology Tissue) numbers

AMNH=American Museum of Natural History, ANWC=Australian National Wildlife Collection, UWBM=Burke Museum of Natural History and Culture (University of Washington), FLMNH=Florida Museum of Natural History, FMNH=Field Museum of Natural History, KUMNH=University of Kansas Natural History Museum & Biodiversity Center, LLD=Collection of Llewellyn L. Densmore, Texas Tech University, LSU HSC=Louisiana State University Health Science Center, LSUMNS=Louisiana State University Museum of Natural Science, MBM=Marjorie Barrick Museum (University of Nevada, Las Vegas), MSB=Museum of Southwestern Biology (University of New Mexico), MVZ=Museum of Vertebrate Zoology (University of California, Berkeley), MV=Museum Victoria, USNM=National Museum of Natural History, SFZG=San Francisco Zoological Garden, and ZMUC=Zoological Museum University of Copenhagen.

Table S2. Gene regions sequenced for this project, organized by HUGO gene codes.

HUGO Name[1]	Gene Description	Chr.[2]	Chicken Length[3]	Key Regions Included[4]
ALDOB[UF]	Aldolase B, fructose-bisphosphate	Z	2413	Introns 3 - 7
BDNF[FMNH]	Brain-derived Neurotrophic Factor	5	688	Exon 1
CLTC[UF]	Clathrin, Heavy Polypeptide (Hc)	19	1836	Introns 6 - 7
CRYAA[UF]	Crystallin, Alpha A	1	1193	Intron 1
EEF2[UF,LSU]	Eukaryotic Translation Elongation Factor 2	28	2976	Introns 4 - 8
EGR1[FMNH]	Early Growth Response 1 (Zenk)	13	1717	Exon 2, 3' UTR
FGB[FMNH,WSU]	Fibrinogen Beta Chain	4	2579	Introns 4-7
GH1[NMNH]	Growth Hormone 1, Somatotropin	27	1659	Introns 2 - 3
HMGN2[UF,NMNH]	Chicken Nonhistone Chromosomal Protein HMG-17	23	1667	Introns 2 - 5
IRF2[WSU]	Interferon Regulatory Factor 2	4	617	Intron 2
MB[FMNH]	Myoglobin	1	945	Intron 2
MUSK[FMNH]	Muscle, Skeletal Receptor Tyrosine Kinase	Z	642	Intron 3
MYC[NMNH]	V-Myc Myelocytomatosis Viral Oncogene Homolog (avian)	2	1225	Intron 2, 3' UTR
NGFB[FMNH]	Nerve Growth Factor, beta polypeptide	26	749	Exon 4
NTF3[FMNH]	Neurotrophin 3	1	728	Exon 3
PCBD1[UF]	Pterin-4 Alpha-Carbinolamine Dehydratase /Dimerization Cofactor of Hepatocyte Nuclear Factor 1 Alpha (DCOH)	6	1019	Introns 2 - 3
RHO[UF]	Rhodopsin	12	1555	Introns 1 - 3
TGFB2[FMNH]	Transforming Growth Factor, Beta 2	3	574	Intron 5
TPM1[LSU]	Tropomyosin 1 (alpha)	10	459	Intron 6

[1] Letters indicate specific labs where genes were amplified and contacts for additional information: FMNH (S. Reddy, S. Hackett); UF (R.T. Kimball, E.L. Braun); NMNH (M.J. Braun, C.J. Huddleston, K. Han, T. Yuri); WSU (K.J. Miglia, W.S.Moore); LSU (B. Marks, F.H. Sheldon, C.C. Witt).
[2] Chr. = chromosome in the chicken genome.
[3] Length of the *Gallus gallus* sequence in our alignment.
[4] Partitions may include small amounts of other data (e.g., exon flanking introns). For partitions that covered more than one intron, the intervening exon was also sequenced except for FGB, which forms three discontiguous segments.

Table S3. Unpublished primers and amplification conditions.

Locus	Primer	Sequence (5' – 3')	Annealing Temp. (Mg++ conc.) [1]	Contact or Citation
ALDOB	AldB.3F	GCCATTTCCAGCTCTCATCAAAG	58° (1.5mM)	RTK and ELB
	AldB.7R	AGCAGTGTCCCTTCCAGGTASAC		RTK and ELB
	AldB.6F	GAGCCAGAAGTCTTACCTGAYGG	50° (1.5mM)	(S11)
	AldB.8R	GCTCKCCCGTATGAGAAGGTCAGYTT		RTK and ELB
CRYAA	CRY.1F	TTACTATYCAGCACCCCTGGTTCAA	63° (1.5mM)	RTK and ELB
	CRY.2R	CTGTCTTTCACTGTGCTTGCCRTGRAT		RTK and ELB
EEF2	EEF2.5F	GAAACAGTTTGCTGAGATGTATGTTGC	60° (1.5mM)	RTK and ELB
	EEF2.7R	GGTTTGCCCTCCTTGTCCTTATC		RTK and ELB
	EEF2.6F	CCTTGAYCCCATCTTYAAGGT	58° (1.5mM)	RTK and ELB
	EEF2.9R	CCATGATYCTGACTTTCARGCCAGT		RTK and ELB
FGB	Fib.6F	TTGCAAAGAGTGGAGGGAAG	53.8° (2.0mM)	KJM
	Fib.8R	CCATCCACCACCATCTTCTT		KJM
HMGN2	HMG17.3F	AACGGAGATCGGCGAGGTTATC	58° (1.5mM)	RTK and ELB
	HMG17.6R	ACTTGGCATCNCCAGCACCTTC		RTK and ELB
IRF2	IRF2.2F	ATGTCTTTGGGTCGGGTTTA	55.5° (2.0mM)	KJM
	IRF2.3R	GAAACTGGGCAATTCACACA		KJM
MYC[2]	MYC-FOR-01	TAATTAAGGGCAGCTTGAGTC	53° (1.75mM)	(S12)
	MYC-FOR-02	TGAGTCTGGGAGCTTTATTG	55° (1.7mM)	(S12)
	MYC-REV-47	CTATAAAGACTTTATTAAAGGTATTTACAT		MJB and
PCBD1	PCBD.2F	AGAGCTGTGGGGTGGAACGAGGTGGA	64° (1.5mM)	RTK and ELB
	PCBD.4R	TCRTGGGTGCTCAAGGTGATGTGAAC		RTK and ELB
RHO	Rhod.2F	GAAATTGCTCTCTGGTCRCTGGTYGT	60° (1.5mM)	RTK and ELB
	Rhod.4R	AAAGAANGCYGGGATGGTCATGAAGA		RTK and ELB

[1]Amplification conditions given after the forward primer are associated with the reverse primer listed below (except MYC).
[2]For MYC, an initial PCR was done using MYC-FOR-01 and MYC-REV-47 (using conditions for MYC-FOR-01). This product was then used in a semi-nested PCR with primers MYC-FOR-02 and MYC-REV-47 (using conditions for MYC-FOR-02).

Supporting Online References

S1. R. N. M. Sehgal, I. J. Lovette, *J. Mol. Evol.* **57**, 335 (2003).
S2. J. L. Chojnowski, R. T. Kimball, E. L. Braun, *Gene* **410**, 89 (2008).
S3. A. L. Chubb, *Mol. Phylogenet. Evol.* **30**, 140 (2004).
S4. T. Yuri, R. T. Kimball, E. L. Braun, M. J. Braun, *Mol. Biol. Evol.* **25**, 352 (2008).
S5. M. M. E. Heslewood, M.S., S. C. Tidemann, P. R. Baverstock, *Electrophoresis* **19**, 142 (1998).
S6. R. W. Slade, C. Moritz, A. Heideman, P. T. Hale, *Mol. Ecol.***2**, 359 (1993).
S7. R. M. Zink, A. Pavlova, S. Rohwer, S. V. Drovetski, *Proc. R. Soc. Lond. B. Biol. Sci.* **273**, 1245 (2006).
S8. C. R. Primmer, T. Borge, J. Lindell, G. P. Saetre, *Mol. Ecol.* **11**, 603 (2002).
S9. F. K. Barker, *Mol. Phylogenet. Evol.* **31**, 486 (2004).
S10. A. C. Driskell, L. Christidis, *Mol. Phylogenet. Evol.* **31**, 943 (2004).
S11. W. A. Cox, R. T. Kimball, E. L. Braun, *Auk* **124**, 71 (2007).
S12. J. Harshman, C. J. Huddleston, J. P. Bollback, T. J. Parsons, M. J. Braun, *Syst. Biol.* **52**, 386 (2003).
S13. S. Rozen, H. Skaletsky, in *Bioinformatics Methods and Protocols: Methods in Molecular Biology*, S. Krawetz, S. Misener, Eds. (Humana Press, Totowa, NJ, 2000), pp. 365-386.
S14. J. D. Thompson, T. J. Gibson, F. Plewniak, F. Jeanmougin, D. G. Higgins, *Nucleic Acids Res.* **24**, 4876 (1997).
S15. M. P. Simmons, *Mol. Phylogenet. Evol.* **31**, 874 (2004).
S16. Z. Zhang, S. Schwartz, L. Wagner, W. Miller, *J. Comput. Biol.* **7**, 203 (2000).
S17. D. L. Swofford, *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4.* (Sinauer Associates, Sunderland, Massachusetts, 2003)
S18. K. C. Nixon, *Cladistics* **15**, 407 (1999).
S19. P. A. Goloboff, J. S. Farris, K. Nixon, *TNT (Tree analysis using New Technology) (BETA) vers. 0.9.* (Published by the authors, Tucuman, Argentina, 2000)
S20. D. J. Zwickl, Ph.D. thesis, The University of Texas at Austin (2006).
S21. A. Stamatakis, *Bioinformatics* **22**, 2688 (2006).
S22. S. Guindon, O. Gascuel, *Syst. Biol.* **52**, 696 (2003).
S23. F. Ronquist, J. P. Huelsenbeck, *Bioinformatics* **19**, 1572 (2003).
S24. D. E. Soltis, M. A. Gitzendanner, P. S. Soltis, *Int. J. Plant Sci.* **168**, 137 (2007).
S25. P. A. Goloboff, D. Pol, in *Parsimony, Phylogeny, and Genomics*, V. A. Albert, Ed. (Oxford University Press, Oxford, 2005), pp. 148-162.
S26. C. G. Sibley, J. A. Ahlquist, *Phylogeny and Classification of Birds* (Yale University Press, New Haven, 1990).
S27. B. C. Livezey, R. L. Zusi, *Zool. J. Linn. Soc.* **149**, 1 (2007).
S28. P. G. P. Ericson *et al.*, *Biol. Lett.* **2**, 543 (2006).
S29. M. G. Fain, P. Houde, *Evolution Int. J. Org. Evolution* **58**, 2558 (2004).
S30. T. A. Paton, A. J. Baker, J. G. Groth, G. F. Barrowclough, *Mol. Phylogenet. Evol.* **29**, 268 (2003).
S31. D. P. Mindell, *Syst. Biol.* **41**, 126 (1992).