Syst. Biol. 56(4):656–672, 2007

ISSN: 1063-5157 print / 1076-836X online
DOI: 10.1080/10635150701549672

Mitochondrial and Nuclear DNA Sequences Support a Cretaceous Origin of Columbiformes and a Dispersal-Driven Radiation in the Paleogene

SERGIO L. PEREIRA,[1] KEVIN P. JOHNSON,[2] DALE H. CLAYTON,[3] AND ALLAN J. BAKER[1,4]

[1]*Department of Natural History, Royal Ontario Museum, Toronto, Ontario M5S 2C6, Canada; E-mail: sergio.pereira@utoronto.ca*
[2]*Illinois Natural History Survey, Champaign, Illinois 61820, USA*
[3]*Department of Biology, University of Utah, Salt Lake City, Utah 84112, USA*
[4]*Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Ontario M5S 3B2, Canada*

Abstract.— Phylogenetic relationships among genera of pigeons and doves (Aves, Columbiformes) have not been fully resolved because of limited sampling of taxa and characters in previous studies. We therefore sequenced multiple nuclear and mitochondrial DNA genes totaling over 9000 bp from 33 of 41 genera plus 8 outgroup taxa, and, together with sequences from 5 other pigeon genera retrieved from GenBank, recovered a strong phylogenetic hypothesis for the Columbiformes. Three major clades were recovered with the combined data set, comprising the basally branching New World pigeons and allies (clade A) that are sister to Neotropical ground doves (clade B), and the Afro-Eurasian and Australasian taxa (clade C). None of these clades supports the monophyly of current families and subfamilies. The extinct, flightless dodo and solitaires (Raphidae) were embedded within pigeons and doves (Columbidae) in clade C, and monophyly of the subfamily Columbinae was refuted because the remaining subfamilies were nested within it. Divergence times estimated using a Bayesian framework suggest that Columbiformes diverged from outgroups such as Apodiformes and Caprimulgiformes in the Cretaceous before the mass extinction that marks the end of this period. Bayesian and maximum likelihood inferences of ancestral areas, accounting for phylogenetic uncertainty and divergence times, respectively, favor an ancient origin of Columbiformes in the Neotropical portion of what was then Gondwana. The radiation of modern genera of Columbiformes started in the Early Eocene to the Middle Miocene, as previously estimated for other avian groups such as ratites, tinamous, galliform birds, penguins, shorebirds, parrots, passerine birds, and toucans. Multiple dispersals of more derived Columbiformes between Australasian and Afro-Eurasian regions are required to explain current distributions. [Columbiformes; dispersal; divergence times; doves; fossil; K/T boundary; molecular phylogeny; pigeons.]

Strongly supported phylogenies inferred from multiple genes, the fossil record, and modern methods of molecular dating are key elements to interpret patterns of macroevolution and biogeography within a temporal framework. For birds the fossil record has revealed that many lineages became extinct about 65.5 million years ago (Mya) during the transition between the Mesozoic and Cenozoic Periods, and most extant lineages radiated shortly after (Brodkorb, 1964; Chiappe and Dyke, 2002; Olson, 1985; Slack et al., 2006). Conversely, inferences from molecular data have suggested that many orders of extant birds are older than 65.5 Mya (e.g., Cooper and Penny, 1997; Hedges et al., 1996; Pereira and Baker, 2006a; Slack et al., 2006; van Tuinen and Hedges, 2001). Additionally, molecular time estimates place the origin of extant avian families at the end of the Mesozoic, and the radiation of modern genera in the Paleocene and Eocene, at least for ratites, tinamous, galliform birds, penguins, shorebirds, parrots, passerine birds, and toucans (e.g., Baker et al., 2004, 2006; Nahum et al., 2003; Paton et al., 2003; Pereira and Baker, 2006a, 2006b; Pereira et al., 2002; Tavares et al., 2006). For many other avian groups, such as the Columbiformes, the lack of a strong phylogenetic hypothesis at and above the genus level based on multiple genes has precluded the interpretation of their evolution within a timeframe.

Columbiformes is one of most easily recognized avian orders worldwide and is traditionally subdivided in two families, the Columbidae and the Raphidae. The Columbidae is represented today by over 300 living species of pigeons and doves. Members of the Columbidae have sturdy bodies, short necks, short slender bills, and a fleshy cere. They vary greatly in body size, from the sparrow-sized ground doves of the genus *Columbina* to the fowl-sized crowned pigeons of the genus *Goura*, as well as in coloration, from dull brown or grey to glossy orange or green plumage, especially around the head and wings. Columbidae is generally subdivided into the subfamilies Columbinae (typical pigeons and doves), Treroninae (fruit and green doves), Gourinae (crowned pigeons), and the monotypic Otidiphabinae (pheasant pigeon) and Didunculinae (tooth-billed pigeon). The Raphidae was endemic to the Mascarene Islands in the Indian Ocean, becoming extinct in the 17th and 18th centuries. The family contained only two species of *Raphus* and the monotypic *Pezophaps*, known as dodo and solitaires, respectively. Dodo and solitaires were flightless birds of about one meter high, likely the outcome of adaptation to a terrestrial and predator-free lifestyle that led to little morphological resemblance to extant Columbidae (Gibbs et al., 2001).

Morphological differences among members of these two families have guaranteed separate taxonomic status for the Columbidae and the Raphidae in most historical and current classifications (Gibbs et al., 2001; Goodwin, 1983; Sibley and Ahlquist, 1990). However, the only molecular phylogeny of high-level columbiform taxa to include the Raphidae suggested otherwise (Shapiro et al., 2002). Maximum likelihood and Bayesian analyses of 1.4 kb of mitochondrial cytochrome *b* and the ribosomal 12S rDNA sequences revealed a close phylogenetic relationship between the Raphidae and the monotypic *Caloenas* (Columbinae) in a well-supported sister clade to the subfamilies Gourinae and Didunculinae (Shapiro et al., 2002). The more derived position of the Raphidae within

the Columbiformes conflicts with suggestions that they became isolated before modern lineages evolved or were descendants of unknown extinct lineages (Gibbs et al., 2001).

Furthermore, the monophyly of subfamilies within Columbidae is also questionable as judged from molecular phylogenies estimated with mitochondrial and nuclear DNA sequences (Johnson, 2004; Johnson and Clayton, 2000; Shapiro et al., 2002). Monophyly of the subfamily Columbinae is usually disrupted by the inclusion of all other subfamilies within it, as previously suggested by an unrooted representation of the relationships of genera (Goodwin, 1983). Likewise, molecular data tend to place the genus *Treron* apart from other members of the Treroninae, and there is disagreement among phylogenetic hypotheses regarding which other genera are closely related to *Treron* (Johnson, 2004; Shapiro et al., 2002). Nonetheless, these molecular hypotheses have to be interpreted with caution because they were obtained from analyses using limited character and taxa sampling, and recovered many nodes with low branch support.

The origin of modern Columbiformes is also disputed. The Oriental and Australasia regions harbor the highest diversity of species, and hence these areas would be inferred to be the center of origin and dispersal according to classical biogeographic reasoning (Darwin, 1859; Wallace, 1876). On the other hand, molecular phylogenetic analyses suggested that the Neotropics are likely the center of origin of Columbiformes because Neotropical genera were placed closer to the root of the trees obtained in two independent studies (Johnson and Clayton, 2000; Shapiro et al., 2002). However, the two phylogenetic hypotheses proposed for Columbiformes based on modern cladistic methods disagree as to whether the mourning and quail doves (Johnson and Clayton, 2000) or ground doves and allies (Shapiro et al., 2002) are the sister group to the remaining Columbiformes. It is not clear if the topological differences observed in these two studies are due to the choice of outgroup used (Apodiformes in Johnson and Clayton, 2000, and Charadriiformes in Shapiro et al., 2002), bias in base composition, or conflicting phylogenetic signal among the sequence data used. Additionally, both studies did not find strong support for the sequence of branching events closer to the root of the tree, and thus it has not been possible to test biogeographic hypotheses of the origin and radiation of the Columbiformes.

We inferred the phylogenetic relationships within Columbiformes using multiple approaches to clarify the high-level systematics of pigeons and doves. Our phylogeny is based on the largest data set gathered for this group so far, including 9 kilobase pairs (kb) of aligned mitochondrial and nuclear DNA sequences for most of the extant genera. We use the strongly supported phylogeny to estimate divergence times and infer ancestral areas, and propose a temporal-biogeographic hypothesis for dispersal of the Columbiformes across the globe, which seems to have been facilitated by plate tectonics during the Cenozoic.

Material and Methods

Taxon Sampling

We sampled species from 33 out of 41 Columbiformes genera and eight outgroup species (Table 1). Apodiformes and Caprimulgiformes were chosen as closer outgroups to Columbiformes based on molecular and anatomical evidence (Ericson et al., 2006; Livezey and Zusi, 2007). Anseriformes, Galliformes, and Struthioniformes outgroups were also included in the analysis for calibration purposes (see Temporal Constraints in Molecular Dating). DNA or tissue samples from species from five other genera (the extinct *Pezophaps* and *Raphus*; the rare or endemic *Alectroenas*, *Didunculus*, and *Drepanoptila*) were not available to us but we included 1.4 kb of DNA sequences from a previous study (Shapiro et al., 2002) in our analyses. Hence, only specimens from the extinct *Microgoura* and the monotypic *Cryptophaps* and *Starnoenas* were missing from our taxon sampling. We considered the monotypic *Nesoenas mayeri* a synonym of *Streptopelia* following Johnson et al. (2001), and the New and Old World *Columba* pigeons as separate genera (*Patagioenas* and *Columba*, respectively), following Banks et al. (2003) and Johnson et al. (2001).

DNA Amplification, Sequencing, and Sequence Alignments

Building on the previously published gene segments of Johnson (2004) and Johnson and Clayton (2000), we selected the following mitochondrial regions for amplification: small ribosomal subunit (12S rDNA), the ATP synthase F0 subunits 8 and 6 plus part of cytochrome *c* oxidase subunit III (ATPase8/6–COIII), NADH dehydrogenase subunit 2 (ND2), cytochrome *b* (cyt *b*) (primers designed by O. Haddrath and described in Pereira and Baker, 2004), and cytochrome *c* oxidase subunit I (COI) (COIaR: AAC YAA CCA CAA AGA CAT YGG and COIbR: GAN AGG ACA TAG TGG AAG TGG GC; O. Haddrath, personal communication). We chose the following nuclear DNA regions for amplification: recombination activating protein (RAG-1) gene (primers R13, R18, R17, R22, R21, and R2b described in Groth and Barrowclough, 1999), interphotoreceptor retinoid-binding protein (IRBP) gene (IRBP217GG: GAC ATG GCT AAR GTN CTA CTG GAT AAC TAC TG and IRBP1095H: CTC TGC GCA CTG CCA GGA TGT C; O. Haddrath, personal communication), and β fibrinogen intron 7 (FIB7) (primers FIB-BI7U and FIB-BI7L from Prychitko and Moore, 1997). We performed mitochondrial and nuclear DNA amplifications as previously described in Pereira and Baker (2005) and Groth and Barrowclough (1999), respectively. PCR products were recovered from 1% agarose gels, purified by centrifuging each through a filter tip, and cycle-sequenced and run on an ABI 3100 automated DNA sequencer (IRBP and 12S rDNA) or Li-Cor 4200 bidirectional automated DNA sequencer (all other genes) according to the manufacturer's suggested protocols. Both L- and H-strands sequences were checked for ambiguities and the final consensus sequence was created for each fragment in Sequencher 4.1.2 (GeneCodes, Ann Arbor, Michigan). All consensus

TABLE 1. Species sampled and GenBank accession numbers. Genes sequenced were the mitochondrial small ribosomal subunit (12S rDNA), the ATP synthase F0 subunits 8 and 6 plus part of cytochrome *c* oxidase subunit III (ATPCO3), NADH dehydrogenase subunit 2 (ND2), cytochrome *b* (cyt *b*), cytochrome *c* oxidase subunit I (COI), the nuclear recombination activating protein (RAG-1), interphotoreceptor retinoid-binding protein (IRBP), and β fibrinogen intron 7 (FIB7).

Species	12S rDNA	ATPCO3	ND2	cyt b	COI	RAG-1	IRBP	FIB7
Pezophaps solitaria	AF483300	—	—	AF483337	—	—	—	—
Raphus cucullatus	AF483301	—	—	AF483338	—	—	—	—
Alectroenas madagascariensis	AF483307	—	—	AF483344	—	—	—	—
Caloenas nicobarica	EF373289	EF373439	EF373326	AF483336	EF373363	EF373493	EF373400	EF373477
Chalcophaps stephani	EF373293	EF373444	EF373328	AY443673	EF373365	EF373498	EF373405	AY443695
Claravis pretiosa	EF373294	EF373445	EF373329	AF182682	EF373366	EF373499	EF373406	AF182649
Columba livia	EF373295	EF373446	AF353433	AF182694	EF373367	EF373500	EF373407	AF182661
Columbina squammata	EF373296	EF373447	EF373330	AF182684	EF373368	EF373501	EF373408	AF182651
Didunculus strigirostris	AF483306	—	—	AF483343	—	—	—	—
Drepanoptila holosericea	AF483308	—	—	AF483345	—	—	—	—
Ducula rufigaster	EF373297	EF373448	EF373331	EF373277	EF373369	EF373502	EF373409	EF373479
Ectopistes migratorius	AF483314	—	—	AF483351	—	—	—	—
Gallicolumba jobiensis	EF373298	EF373449	EF373332	EF373278	EF373370	EF373503	EF373410	EF373480
Geopelia striata	EF373299	EF373450	EF373333	EF373279	EF373371	EF373504	EF373411	EF373481
Geophaps lophotes	EF373300	EF373451	EF373334	AF483323	EF373372	EF373505	EF373412	EF373482
Geotrygon montana	EF373301	EF373452	EF373335	AF182696	EF373373	EF373506	EF373413	AF182663
Goura cristata	EF373302	EF373453	EF373336	AF182709	EF373374	EF373507	EF373414	AF182676
Gymnophaps albertsii	EF373303	EF373454	EF373337	EF373280	EF373375	EF373508	EF373415	EF373483
Hemiphaga novaeseelandiae	EF373304	EF373455	EF373338	AY443666	EF373376	EF373509	EF373416	AY443688
Henicophaps albifrons	EF373305	EF373456	EF373339	EF373281	EF373377	EF373510	EF373417	EF373484
Leptotila rufaxilla	EF373306	EF373457	EF373340	AF182698	EF373378	EF373511	EF373418	AF182665
Leucosarcia melanoleuca	EF373307	EF373458	EF373341	AF182712	EF373379	EF373512	EF373419	AF182679
Lopholaimus antarcticus	EF373308	EF373459	EF373342	EF373282	EF373380	EF373513	EF373420	EF373485
Macropygia amboinensis	EF373309	EF373460	EF373343	EF373283	EF373381	EF373514	EF373421	EF373486
Metriopelia morenoi	EF373310	EF373461	EF373344	AY443677	EF373382	EF373515	EF373422	AY443699
Oena capensis	EF373311	EF373462	EF373345	AF182707	EF373383	EF373516	EF373423	AF182674
Otidiphaps nobilis	EF373312	EF373463	EF373346	AF483352	EF373384	EF373517	EF373424	EF373487
Patagioenas speciosa	EF373313	EF373464	EF373347	AF279711	EF373385	EF373518	EF373425	AF279721
Petrophassa albipennis	EF373314	EF373465	EF373348	EF373284	EF373386	EF373519	EF373426	EF373488
Phapitreron amesthystina	EF373315	EF373466	EF373349	AF182706	EF373387	EF373520	EF373427	AF182673
Phaps chalcoptera	EF373316	EF373467	EF373350	AF182713	EF373388	EF373521	EF373428	AF182680
Ptilinopus pulchellus	EF373317	EF373468	EF373351	EF373285	EF373389	EF373522	EF373429	EF373489
Reinwardtoena browni	EF373318	EF373469	EF373332	AF353417	EF373390	EF373523	EF373430	AF353468
Streptopelia capicola	EF373319	EF373470	EF373333	AF279709	EF373391	EF373524	EF373431	AF279719
Treron calva	EF373320	EF373471	EF373354	AY443674	EF373392	EF373525	EF373432	AY443696
Trugon terrestris	EF373321	EF373472	EF373355	EF373286	EF373393	EF373526	EF373433	EF373490
Turacoena manadensis	EF373322	EF373473	EF373356	EF373287	EF373394	EF373527	EF373434	EF373491
Turtur chalcospilos	EF373323	EF373474	EF373357	AY443671	EF373395	EF373528	EF373435	AY443693
Uropelia campestris	EF373324	EF373475	EF373358	EF373288	EF373396	EF373529	EF373436	EF373492
Zenaida macroura	EF373325	EF373476	EF373359	AF182703	EF373397	EF373530	EF373437	AF258321
Caprimulgus vociferus	EF373292	EF373443	EF373327	U89194	EF373364	EF373497	EF373404	AY695136
Podargus strigoides	—	EF373442	—	EF373276	EF373362	EF373496	EF373403	AY082408
Chaetura pelagica	EF373291	EF373441	AY294537	AY294475	EF373361	EF373495	EF373402	AY830606
Hirundapus caudacutus	EF373290	EF373440	AY294536	AY294474	EF373360	EF373494	EF373401	EF373478
Anhima cornuta	AY140699	EF373438	AY140737	AY140735	AY140729	AY140765	EF373399	AY140701
Crax blumenbachii	AF165444	AY143682	AY140747	AF165468	AF165492	AY140775	EF373398	AY140711
Gallus gallus	NC_001323	NC_001323	NC_001323	NC_001323	NC_001323	AF143730	AY994153	AY082425
Struthio camelus	NC_002785	NC_002785	NC_002785	NC_002785	NC_002785	AF143727	—	AY082424

fragments were aligned visually in MacClade 4.0 (Maddison and Maddison, 2000) for each gene, and then concatenated in a matrix of 9083 base pairs (bp), including gaps. Ambiguously aligned regions (1.1% of total alignment) were excluded from the analysis. All sequences obtained in this study were deposited in GenBank (accession numbers EF373276 to EF373530). The final matrix and all trees inferred in this study were deposited in TreeBASE (Study Number S1808).

Data Partition, Model Selection, and Bias in Base Composition

We partitioned the data according to the fragments amplified. Thus, ATPase 8, ATPase 6, and COIII were analyzed as a single fragment. This partition is statistically more plausible because it reduces stochastic error in the estimation of the model parameters if the 168 bp for ATPase 8 and 173 bp for COIII were to be considered separate partitions. Model selection for each partition was made according to the Akaike information criteria (AIC) as implemented in MrModelTest 2.0 (Nylander, 2004). AIC was chosen over hierarchical likelihood ratio tests because AIC is not influenced by the order in which models are compared, and the AIC allows quantification of model uncertainty through Akaike weights. The best fitting model for each partition was used in subsequent Bayesian phylogenetic analyses, with parameters estimated during the analyses. We also performed model

selection for all genes combined to carry out a nonpartitioned maximum likelihood analysis. Bias in base composition can negatively influence phylogenetic inference (Haddrath and Baker, 2001) and, hence, we tested it by a chi-square test as implemented in TREEPUZZLE 5.0 (Strimmer and von Haeseler, 1996).

Phylogenetic Bayesian Inference

We applied a Metroplis-coupled Markov chain Monte Carlo (MCMC) approach as implemented in MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003) to infer Bayesian phylogenetic relationships under two simultaneous independent runs, each starting with a different random tree. We set each run to have one cold and five heated chains to allow better mixing of the MCMC chain and minimize the chance of being trapped in local optima. The addition of more heated chains allows for quicker convergence of large and complex data sets such as ours. We considered runs to have reached convergence when the average standard deviation of the split frequencies between both simultaneous runs was smaller than 0.01. We performed runs assuming the same topology to be shared among all partitions but model parameters (substitution matrix, state frequency, proportion of invariable sites, and shape parameter of the gamma distribution) were unlinked and each partition allowed to have its own model of DNA evolution as chosen by AIC. We also assumed a priori among-partition rate variation. All trees were considered equally likely. The tree was rooted with sequences from the Ostrich (*Struthio camelus*). Other priors applied for all partitions were unconstrained: exponential (10.0) for branch lengths, flat Dirichlet (1,1,1,1) for stationary base frequencies, flat Dirichlet (1,1,1,1,1,1) for the nucleotide substitution ratio, uniform distribution (0,200) for the shape parameter of the gamma distribution of rate variation, and uniform distribution (0,1) for the proportion of invariable sites. MCMC samples were taken in every 1000th cycle. We plotted the log likelihood of sampled topologies to determine the burn-in period in which the MCMC chain had reached a stationary status. Post–burn-in samples from both simultaneous and independent runs were used to construct a 50% majority-rule consensus tree. The proportion of trees in which nodes were recovered after the burnin period is interpreted as the posterior probability (PP) of that node, or the probability that that node is true (Ronquist and Huelsenbeck, 2003). Nodes receiving PP $\geq$ 0.95 were considered to be strongly supported.

Maximum Likelihood Analysis

Maximum likelihood (ML) inference of phylogeny was executed in GARLI v 0.942 (Zwickl, 2006), assuming the GTR substitution model with a proportion of invariable sites and four-category gamma-distributed rate variation estimated from the data set for 100 replicates. GARLI uses a genetic algorithm to search for the topology, branch lengths, and the substitution model parameters (rate of DNA substitution, base frequency, proportion of invariable sites, and shape parameter of the gamma distribution) that maximize the likelihood of generating the observed set of DNA sequences. Briefly, the genetic algorithm will create a sample of individual solutions at a given time, each solution with its own topology, branch lengths, and model parameters representing the phylogenetic relationships of the given DNA sequences; next, each individual solution is ascribed a fitness likelihood score, which determines the proportion of "offspring" solutions that each individual solution will contribute to the next generation; the topology, branch lengths, or one of the substitution model parameters of each offspring solution is then changed to a new state and the others are reestimated to maximize the likelihood; these steps are repeated for a number of generations until an arbitrary stop condition is reached. For the Columbiformes data set, the stop condition was set to automatic termination, which is reached when three conditions are met: (1) changes in topology do not significantly increase the offspring fitness; (2) the total increase in the likelihood score over a number of generations is less than a specified amount; and (3) the parameter controlling the degree of branch length optimization has reached its minimum value. Two independent GARLI runs, each starting with a different random tree topology, were performed. Judging from the similar results compared across runs, GARLI is unlikely to have been trapped in local optima.

Maximum Parsimony Analysis

Maximum parsimony (MP) inference of phylogeny was conducted in PAUP 4 beta 10 (Swofford, 2001). Gaps were treated as missing characters and multistate characters treated as uncertain. A heuristic tree search was performed by stepwise addition and sequences were added randomly in 100 replicates. The "MulTrees" option was in effect. Branch support was estimated in PAUP 4 by nonparametric bootstrapping with the same parameters as the heuristic search. In this case, for each bootstrap replicate 100 random sequence additions was in effect.

Tree Selection

To compare different tree topologies obtained by different methods of phylogenetic inference, we employed the approximately unbiased (AU) test (Shimodaira, 2002) as implemented in the program CONSEL version 0.1f (Shimodaira and Hasegawa, 2001). The AU test uses a multiscale bootstrap technique and site log-likelihoods when computing *P*-values for the topologies being tested. The multiscale bootstrap technique removes the bias associated with more conservative tests that will result in more trees in the confidence set when the number of trees being compared increases. The distribution of posterior probabilities for each alternative topology is obtained by counting the number of times the hypothesis is supported by the replicates.

Exclusion of Taxa with Missing Data

To evaluate whether the inclusion of six genera (*Alectroenas*, *Drepanoptila*, *Didunculus*, *Ectopistes*, *Pezophaps*,

and *Raphus*) with about 84% of missing data would affect tree topology we repeated all the phylogenetic analysis as described above excluding these six taxa from the data set. We also performed a Bayesian analysis including all taxa but reducing the data set to 1590 bp for which we had 12S rDNA and cyt *b* sequences for all taxa. The outgroup *Podargus* was also excluded from this last analysis because we could not obtain 12S rDNA sequences for this bird.

Molecular Dating

For each data partition used in the Bayesian inference of phylogeny we obtained maximum likelihood estimates of the transition/transversion ratio, and nucleotide frequencies in PAML 3.14 (Yang, 1997) under the F84 model of DNA substitution assuming rate variation across sites to follow a gamma distribution with five discrete rate categories (Hasegawa et al., 1985). These parameters were used to estimate branch lengths for each data partition and their approximate variance-covariance matrix to derived estimates of divergence times and 95% credibility intervals (95% CrI) based on all data partitions in a Bayesian framework (Thorne and Kishino, 2002; Thorne et al., 1998). These methods are implemented in the software ESTBRANCHES and MULTIDIVTIME from the MULTIDISTRIBUTE package, freely available from J. Thorne's website: http://statgen.ncsu.edu/thorne/multidivtime.html. The method requires an outgroup to root the tree and imposes the condition that the rate of change in the rate of DNA substitution at the root node is the same at the beginning and at the end of that branch (Kishino et al., 2001; Thorne et al., 1998).

Bayesian dating was run assuming a burn-in period = 3000, sample frequency = 100, number of samples = 10,000. We set the following gamma priors: expected time between tip and root (rttm) = 122.2 Mya based on the split between Galloanserae versus Neoaves estimated from mitogenomic data and 13 fossil constraints equally spread across the tree (Pereira and Baker, 2006a) with standard deviation (SD) = 20 Mya, rate of the root node (rtrate) and its SD = 0.00306 substitutions per site per million years as estimated from the median of the tip-to-root branch lengths for all genes. However, it seems of little practical importance to specify these priors because they do not appear to have any appreciable effect on the Bayesian posterior distribution of node ages and rates of evolution, and because sequence data and time constraints should determine the overall rate and the age of the root (Yang and Yoder, 2003; Wiegmann et al., 2003). We also set the prior for the rate change between ancestral (brownmean) and descendant nodes = 0.00818 (SD = 0.00818) substitutions per site per million of years, so that rttm × brownmean = 1. This later prior follows the suggestion that this is a meaningful value for real and simulated data sets (Wiegmann et al., 2003). Because a priori information for rate change is unknown, a large SD value was chosen as suggested (Thorne and Kishino, 2002), which allows a gene to have a priori a large variation in rate change over time. We assessed the convergence of the MCMC algorithm by running multiple analyses (each one starting with a different randomly selected initial state) and comparing the posterior distribution of divergence times, branch lengths, and the proportion of successful changes of those parameters along the Markov chain.

Temporal Constraints in Molecular Dating

The fossil record of the Columbiformes is scanty and made up mostly of specimens from the Pliocene and Pleistocene, which have not been placed cladistically within the order. The oldest fossil attributed to the Columbiformes is *Gerandia calcaria* (Olson, 1985) from the Early Miocene (Aquitanian) of France between 20.4 and 23.0 Mya (Gradstein et al., 2004). Hence, we assumed a conservative minimum age of 20.4 Mya for Columbiformes. The fossil record suggests that Apodiformes and Caprimulgiformes were present in the Early Eocene (Dyke and Van Tuinen, 2004). We assumed a conservative minimum age of 48.6 Mya for the separation between these two orders. In previous work (Pereira and Baker, 2006a), the divergence time between Galloanserae (Galliformes and Anseriformes) and Neoaves (all other Neognathae) was estimated at 122.2 Mya (95% CrI 110, 135 Mya) and between Galliformes and Anseriformes at 101.7 Mya (95% CrI 92, 113 Mya), based on a Bayesian analysis of complete mitochondrial genomes and assuming several independent time constraints from the fossil record. Hence, we set minimum and maximum ages for the separation of Galloanserae and Neoaves and for Galliformes and Anseriformes based on the 95% CrI. Similarly, we assumed a minimum and maximum age for the split between *Crax* and *Gallus* at 71 and 114 Mya based on the Bayesian 95% CrI of age inferences for major Galliformes lineages using mitochondrial genes and several independent fossils as time constraints (Pereira and Baker, 2006b).

Reconstruction of Ancestral Area States

To infer ancestral areas where Columbiformes may have originated we applied Bayesian and maximum likelihood approaches. A Bayesian stochastic mapping approach was derived for discrete morphological characters, which accounts for uncertainties in model parameters and assumes that character change follows a continuous-time Markov process (Huelsenbeck et al., 2003; Lewis, 2001). The Bayesian method is implemented in SIMMAP 1.0 Beta 2.0.8 Build 10042006 (Bollback, 2006). Biogeographic regions of distribution of Columbiformes were treated as discrete characters, in an analogous way to the treatment of morphological characters, and specific character histories that are consistent with the observations are sampled according to their probability under the model (Huelsenbeck et al., 2003). We accounted for phylogenetic and branch length uncertainties by using the trees and their respective branch lengths included in the post-burn-in portion of the Bayesian posterior distribution of trees obtained in MrBayes 3.1.2. The

maximum likelihood approach for ancestral reconstruction maximizes the probability that the observed discrete states evolved under a stochastic Markov model of evolution (Lewis, 2001). The method is implemented in Mesquite 1.1 (build h61) (Maddison and Maddison, 2005a, 2005b). For each internal node, the state that maximizes the probability of observing the states at the terminal nodes is found, while allowing states at all other nodes to vary independently. We reconstructed ancestral area states based on the consensus Bayesian tree topology assuming branch lengths to be scaled to divergence times as estimated in MULTIDIVTIME.

Both the Bayesian and the maximum likelihood approaches overcome the problems associated with reconstruction of ancestral states by a maximum parsimony approach. The maximum parsimony method does not allow multiple changes at a site to occur along a branch, underestimates the variance in ancestral states assignments by considering only those reconstructions that minimize the number of changes, imposes the same probability of change in both long and short nodes, and does not accommodate phylogenetic and parameter uncertainties.

We categorized the current distribution of genera (Goodwin, 1983) as discrete character states based on biogeographic regions (Sclater, 1858; Wallace, 1876) as: (0) Neotropical, (1) Afrotropical, (2) Oriental, and (3) Australasian. Because both Bayesian and maximum likelihood methods require unique character states to be assigned to taxa, and as *Caloenas*, *Chalcophaps*, *Columba*, *Ducula*, *Gallicolumba*, *Geopelia*, *Macropygia*, *Otidiphaps*, *Ptilinopus*, *Streptopelia*, and *Treron* occur in more than one of the defined biogeographic areas, we also defined the following composite area states: (4) Australasian plus Oriental, (5) Afrotropical plus Oriental, and (6) Afrotropical plus Oriental plus Australasian.

RESULTS

Sequence Variability

The combined data set had a total of 2973 parsimony-informative characters, and uncorrected sequence divergence between taxa varied from 4% to 15% (Table 2). IRBP and ND2 have the least and highest number of parsimony-informative characters, respectively (Table 2). IRBP has the narrowest and ND2 the widest range of uncorrected distances among Columbiformes (Table 2). A Bonferroni-corrected chi-square test indicated that no taxon has a significant deviation in base composition. FIB7 for *Claravis pretiosa* was only 395 bp in length, considerably shorter than the average 1034 bp across Columbiformes (range 880 to 1071 bp).

Model Selection

Table 2 also shows the results of model selection for each data partition. For most partitions, the uncertainty in model selection was very low (Akaike weight for best fitting model > 0.9). For FIB7, the second best model added a proportion of invariable sites to the GTR+g model and had an Akaike weight of 0.245. Together, these two models made up 91.2% of the cumulative Akaike weight for FIB7. For IRBP, the second best model dropped the proportion of invariable sites from the best model, GTR+i+g, and had an Akaike weight of 0.217. The first and second best model made up 78.0% of the cumulative Akaike weight.

Tree Inference

The 50% majority-rule consensus Bayesian (BA) topology was obtained from 24,000 post–burn-in trees sampled from two simultaneous and independent runs (Fig. 1). The Columbiformes was recovered as a monophyletic group, with posterior probability of 1.0. Three highly supported major clades were identified within the Columbiformes and will be referred to as A, B, and C (Fig. 1). Clade A is subdivided into two clades with highly supported relationships; one of the subclades contains genera exclusively distributed in the Americas and the other shows a basal polytomy including Old and New World pigeons and doves, including the extinct passenger pigeon *Ectopistes migratorius*. Clade B is well resolved and groups the small New World ground doves together. Clade C includes mostly genera found in Africa, Asia, Australia, the East Indies, and New Zealand. Several internal nodes have low posterior probabilities within Clade C. The posterior distribution

TABLE 2. Sequence variability and model parameters for each data partition. Gene abbreviations as in Table 1.

Gene	Number of aligned sites	% Gapped sites	Parsimony informative sites	% Variable	Range of ingroup uncorrected sequence divergence	Best model	Proportion of invariable sites	Gamma	Akaike weight
ATPCO3	1,015	4.4	489	56.65	0.07–0.19	GTR+i+g	0.37	0.56	1.0
cyt *b*	1,044	5.5	453	48.40	0.07–0.17	GTR+i+g	0.48	0.57	1.0
ND2	951	0	507	63.51	0.08–0.22	GTR+i+g	0.29	0.56	1.0
COI	645	12.9	235	41.71	0–0.16	GTR+i+g	0.56	0.75	0.99
12S rDNA	756	14.8	290	50.41	0.01–0.16	GTR+i+g	0.35	0.52	1.0
FIB7	1,129	79.0*	416	65.99	0.01–0.08	GTR+g	—	2.49	0.67
RAG-1	2,739	7.3	457	35.96	0.01–0.04	GTR+i+g	0.34	0.96	1.0
IRBP	804	0	126	31.72	0.01–0.05	GTR+i+g	0.4	0.89	0.56
All genes	9,083	11.4	2973	47.55	0.04–0.15	GTR+i+g	0.31	0.43	1.0

*Including *Claravis pretiosa* for which FIB7 is only 395 bp compared to the average 1034 bp across Columbiformes. Excluding this taxon, gaps correspond to 48.4% of the FIB7 alignment.

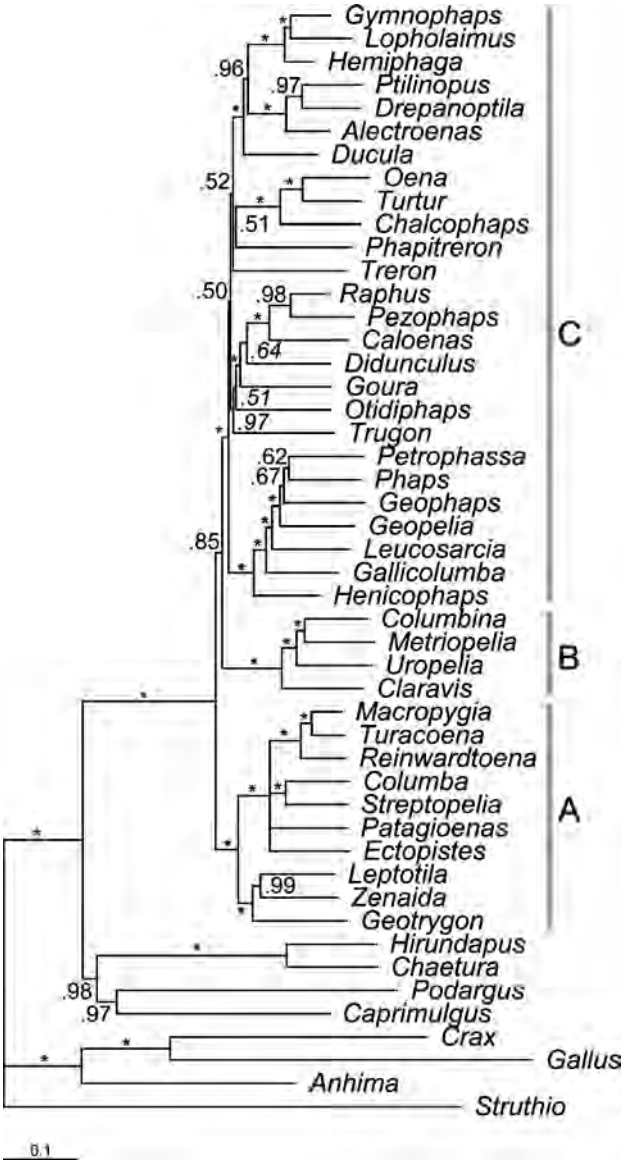

FIGURE 1. Bayesian consensus tree. Major clades within Columbiformes are indicated as A, B, and C. Numbers at nodes are posterior probabilities, which are indicated with an * when = 1.0. Scale bar represents expected number of substitutions per site.

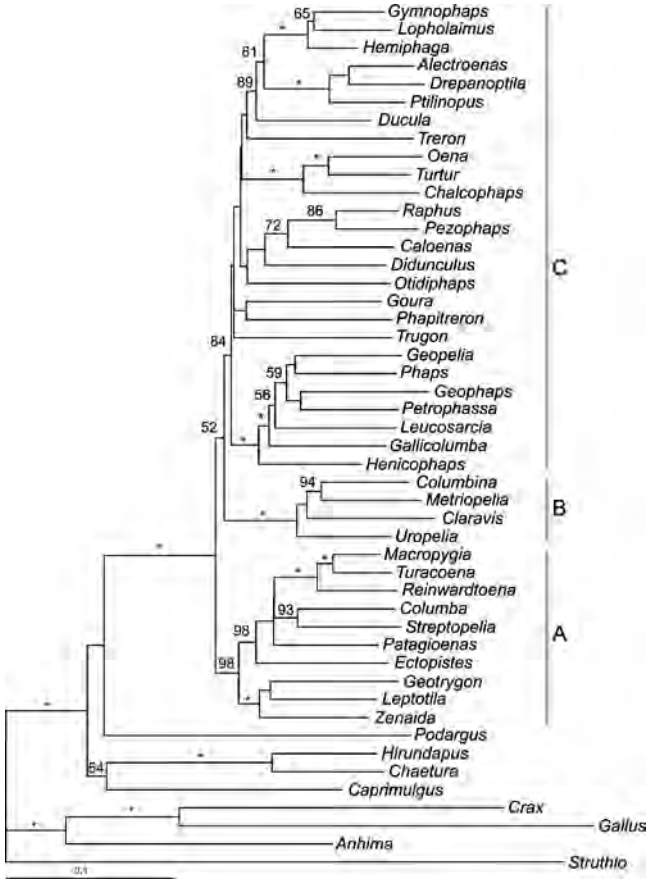

FIGURE 2. Maximum likelihood tree. Major clades within Columbiformes are indicated as A, B, and C. Numbers at nodes are bootstrap proportions and are indicated with an * when = 100 and not indicated when <50%. Scale bar represents expected number of substitutions per site.

of trees contained 8175 alternative topologies that differ mostly in the position of some ingroup taxa with low posterior probability, especially within Clade C. Among these alternative topologies, the three major clades are related to each other in three possible alternative ways: (A(B,C), (A,C)B), and (A,B)C), with cumulative posterior probabilities of 0.83, 0.09, and 0.08, respectively.

The best ML tree obtained in GARLI version 0.942 had a very similar topology to the BA tree, and strongly supported the monophyly of clades A, B, and C (i.e., bootstrap proportion [BP] > 80%) as measured by 100 bootstrap replicates (Fig. 2). The placement of B and C as sister clades received 52% bootstrap support. Relationships within clades A and B are very similar in the ML and the BA topologies, except that *Zenaida* and *Uropelia* were placed more basally in clades A and B in the ML tree, respectively, but without strong support. Additionally, many relationships within clade C were more strongly supported by bootstrapping (i.e., BP > 70%) and were similar to those recovered in the BA analysis, except for the placement of *Alectroenas*, *Drepanoptila*, *Goura*, *Phapitreron*, *Treron,* and *Trugon.* These taxa received less than 50% bootstrap support in the ML topology, but high posterior probabilities in the BA topology. The ML analysis failed to recover Caprimulgiformes as a monophyletic assemblage in the best sampled ML tree, but *Caprimulgus* and *Podargus* were sister genera in 60% of the bootstrapped trees.

One most parsimonious (MP) tree of 19,082 steps was recovered (Fig. 3). Monophyly of clades A, B, and C was well supported by MP bootstrap proportions. Most phylogenetic relationships within clade C were not supported by bootstrapping. The most parsimonious tree also differed from the consensus BA and ML topologies by positioning clade B as a sister clade to A and C. However, the AU test rejected the MP topology ($P = 0.019$) in favor of the consensus BA ($P = 0.896$) and ML topologies

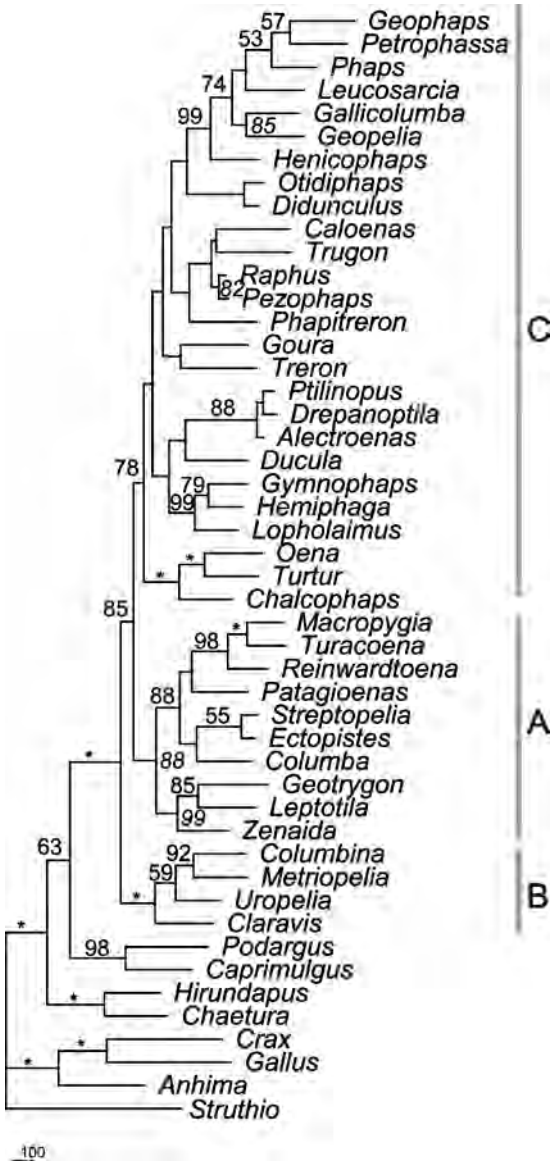

FIGURE 3. Maximum parsimony tree. Major clades within Columbiformes are indicated as A, B, and C. Numbers at nodes are bootstrap proportions and are indicated with an * when = 100 and not indicated when <50%. Scale bar represents expected number of substitutions per site.

($P = 0.161$) as good representations of the phylogenetic relationships among Columbiformes.

Exclusion of Taxa with Missing Data

Phylogenetic analyses of the 9.0-kb data set, which excludes six Columbiformes with significant amount of missing data (~84%), did not show any appreciable topological changes in any method of tree inference (Fig. 4a–c). Furthermore, a Bayesian analysis of 1590 bp of aligned sequences from 12S rDNA and cyt *b* for all taxa (except *Podargus* for which we could not obtain 12S rDNA sequences) resulted in a less resolved tree topology (Fig. 4d), but the placement of the six genera with significant missing data (Fig. 4d) was similar to that in the topologies derived from the 9.0 kb data set (Figs. 1 to 3).

Molecular Dating

We approximated the prior and posterior distribution of divergence times based on a parametric Bayesian approach (Kishino et al., 2001; Thorne and Kishino, 2002). The size of the 95% CrI of the prior distribution for node ages is considerably larger than the size of the 95% CrI of the posterior distribution (Table 3). These differences in the size of credible intervals between the prior and posterior distribution indicate that the prior specification has little influence on the posterior distribution and that most of the information about divergence time is retrieved from the sequence data (Thorne and Kishino, 2002). The prior distribution ignores the information contained in the sequence data; hence, it is expected that there will be a larger amount of uncertainty in prior divergence time estimates. Based on the 95% CrI of the posterior distribution, Columbiformes diverged from outgroups such as Caprimulgiformes and Apodiformes in the Early to Late Cretaceous between 87 to 110 Mya, and Caprimulgiformes and Apodiformes last shared a common ancestor between 83 to 107 Mya. Radiation of modern Columbiform genera occurred from the Early Eocene to Middle Miocene.

Reconstruction of Ancestral Area States

Bayesian reconstruction of ancestral areas implemented in SIMMAP and accounting for phylogenetic uncertainty indicated that the most recent common ancestor of all extant Columbiformes lived in Gondwanaland, most likely in what is now the Neotropical region (posterior probability = 0.977). The Australasian and the composite Australasian plus Oriental areas were the only other regions to receive non-zero but low posterior probabilities (<0.01 and 0.02, respectively). The Markov-ML reconstruction of ancestral areas implemented in Mesquite and accounting for divergence times provided similar conclusions to those from the Bayesian approach. The proportions of the maximum likelihood attributed to each biogeographic region were Neotropical region = 0.85, Australasian = 0.14, remaining areas and composite areas = 0.002 each. The results of reconstruction of ancestral areas favor an ancestral area of modern Columbiformes in Gondwanaland (Fig. 5), when Africa and India had already drifted apart from the Antarctica–Australia–South America landmass. By counting the number of significant changes in the estimates of ancestral areas (Fig. 5), it seems that members of clade A dispersed at least twice to the Oriental biogeographic zone and at least once to the Afrotropical zone, and members of clade C dispersed at least eight times to the Oriental and four times to the Afrotropical zones. No intercontinental dispersals of members of Clade B seem to have occurred according to the reconstruction of ancestral areas (Fig. 5).

DISCUSSION

High-Level Columbiformes Systematics

We gathered the largest mitochondrial and nuclear DNA data set so far to infer the phylogenetic relationships among Columbiformes. Phylogenetic analyses of these sequences recovered competing phylogenetic trees that confirm the monophyly of Columbiformes (Figs. 1 to 3) and refute the monophyly of the family Columbidae with the placement of Raphidae within it, and also contest subfamily classifications proposed for Columbiformes (reviewed in Sibley and Ahlquist, 1990). Despite some topological differences, all the alternative methods of phylogenetic reconstruction identified three major clades within the Columbiformes (Figs. 1 to 3). Clade A comprises *Geotrygon*, *Leptotila*, and *Zenaida* of the New World as a sister group to a clade containing typical New World pigeons, including the extinct passenger pigeon *Ectopistes migratorius*, and Old World cuckoo and turtle doves plus typical pigeons. Clade B includes the small New World ground doves. Clade C contains genera found in Africa, Asia, Australia, the East Indies, and New Zealand, including mountain pigeons, fruit doves, emerald doves, bronzewings, crowned pigeons, Old World ground doves, and many monotypic forms such as the Nicobar, the thick-billed ground, pheasant, wonga,

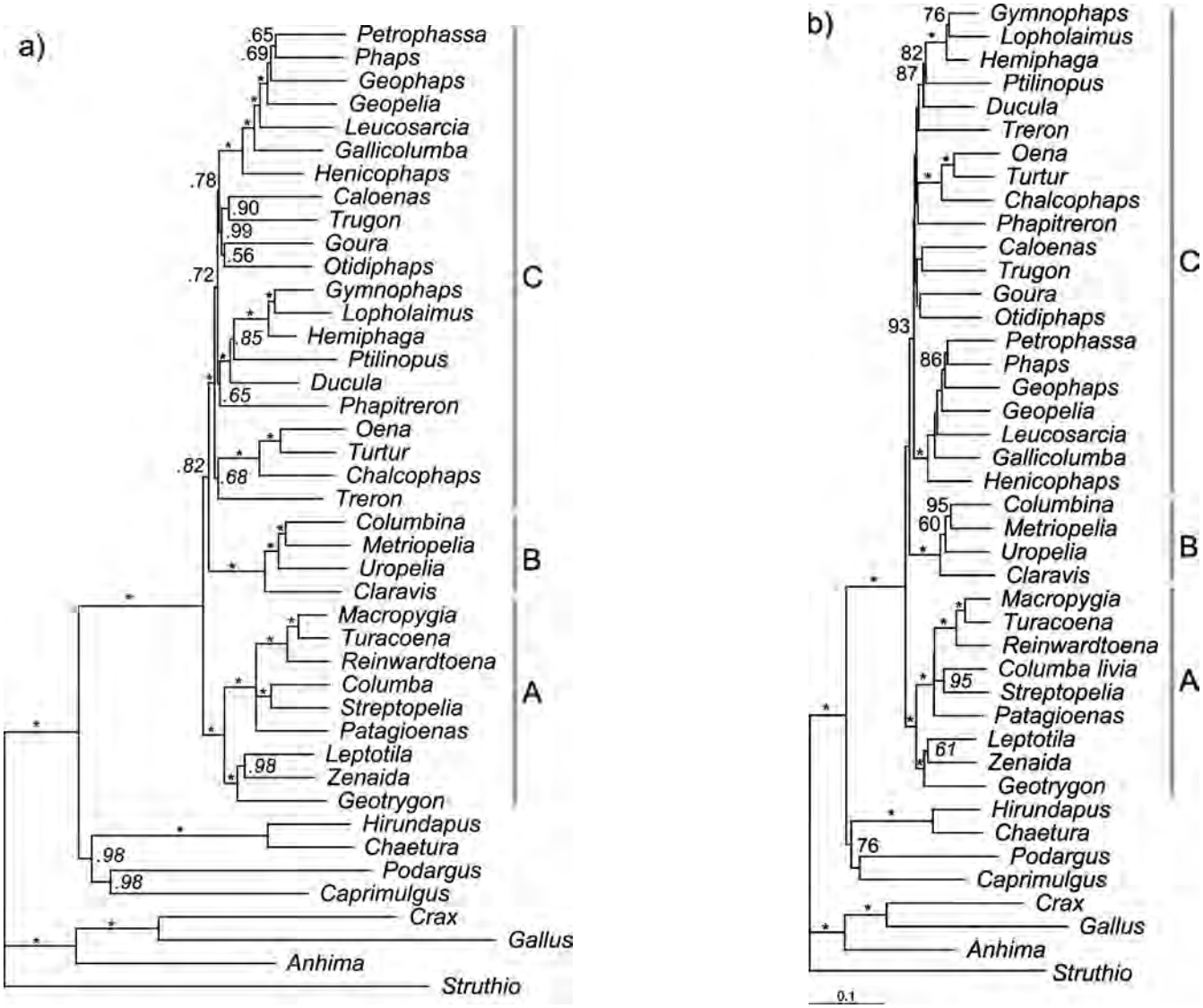

FIGURE 4. Phylogenetic reconstruction obtained when six taxa with significant amount of missing data are excluded from the data set (a, b, and c) and when only 12S rDNA and cyt *b* sequences were used, including GenBank sequences for genera not sampled in this study (d). Reconstructions were based on Bayesian analysis (a and d), maximum likelihood (b), and maximum parsimony (c). Major clades within Columbiformes are indicated as A, B, and C. Numbers at nodes are Bayesian posterior probabilities (a, d) or bootstrap proportions (b, c), indicated with an * when 1.0 or 100, respectively. Scale bar represents expected number of substitutions per site. *(Continued)*

topknot, and New Zealand pigeons, and also the extinct dodo and solitaires usually placed in their own subfamily, the Raphidae. The internal phylogenetic relationships within clades A and B were well supported despite the method of phylogenetic inference used. Conversely, basal phylogenetic relationships within clade C were difficult to resolve and were not well supported by Bayesian posterior probabilities or bootstrap proportions in maximum parsimony and maximum likelihood, even with the combined analysis of about 9.0 kb of mitochondrial and nuclear sequences and appropriate modeling of the DNA substitution process. Clade C may be approaching a hard polytomy, in which case additional DNA sequences are unlikely to improve resolution because many of its members are endemic forms that may have colonized several islands of the Oriental–Australasian biogeographic regions almost simultaneously.

Maximum parsimony analysis indicated clade B to be sister to A and C, as also found previously with maximum parsimony and maximum likelihood analyses of combined cyt *b* and FIB7 sequences (Johnson and Clayton, 2000). Conversely, Bayesian and maximum likelihood analysis inferred a sister relationship between clades B and C, and in turn as a sister group to A. By applying a topological test, the former topology was strongly rejected in favor of the consensus Bayesian and maximum likelihood topology. Given the results of the topological test, the Bayesian posterior distribution of the alternative topological placements of clades A, B, and C, and the use of different models of DNA substitution for different gene partitions in the Bayesian analysis to

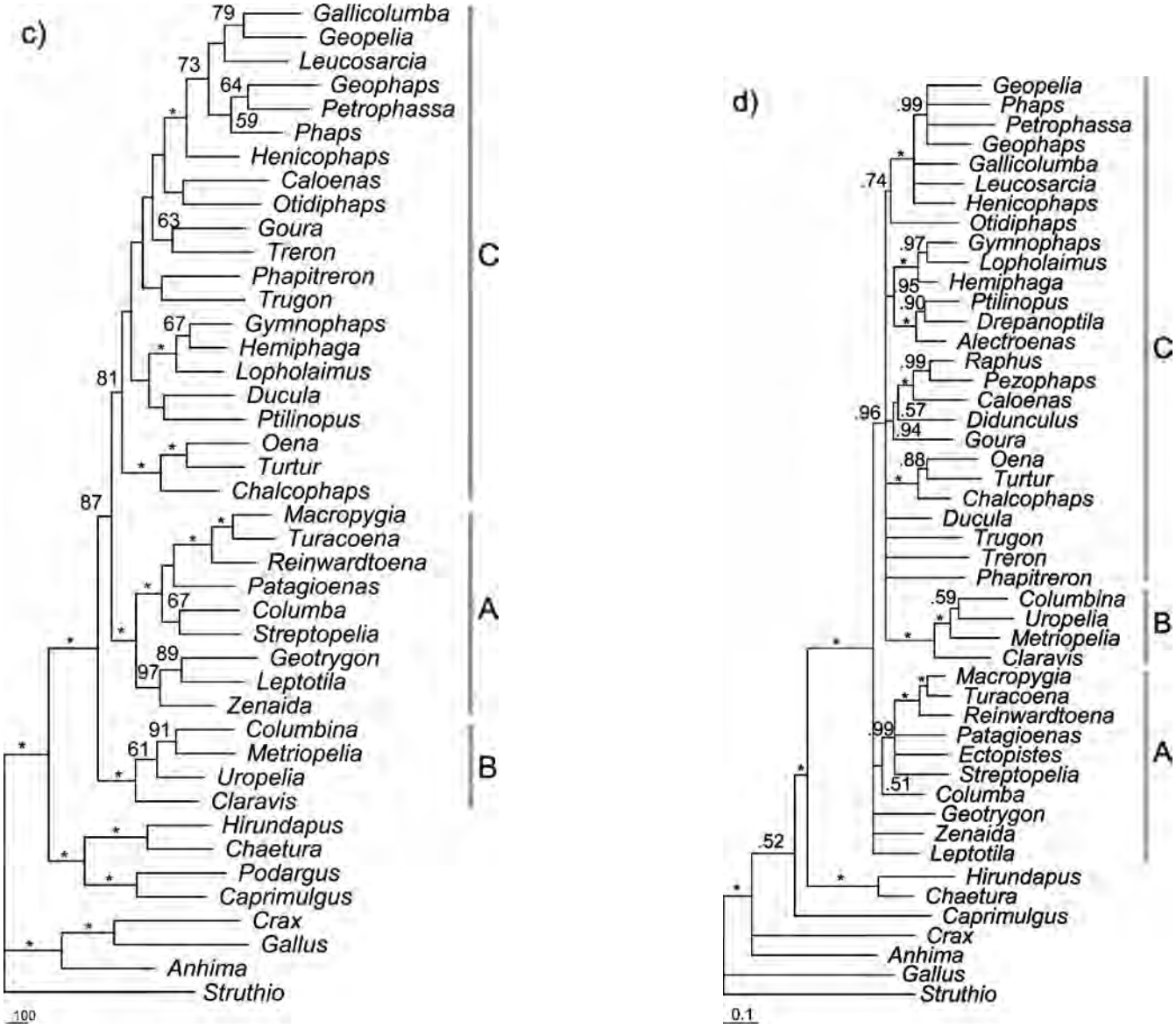

FIGURE 4. (Continued)

TABLE 3. Prior and posterior distribution of divergence times using a Bayesian multigene approach.

	Prior			Posterior		
Node	Mean	SD	95% credible interval	Mean	SD	95% credible interval
1	112.6	8.7	89.6, 122.0	98.0	6.1	87.1, 110.0
2	102.5	12.2	73.9, 119.5	54.4	4.5	46.1, 63.6
3	82.5	19.3	39.7, 113.0	46.9	4.1	39.3, 55.3
4	55.4	23.8	11.7, 99.7	40.2	3.8	33.3, 48.0
5	27.7	20.9	0.9, 76.9	36.9	3.7	30.2, 44.4
6	61.6	22.2	18.1, 101.8	34.6	3.4	28.6, 41.5
7	30.8	22.2	1.2, 80.8	29.2	3.1	23.6, 35.6
8	41.6	21.4	6.7, 86.1	20.0	2.4	15.7, 24.8
9	20.9	17.4	0.6, 64.7	14.4	1.9	11.0, 18.4
10	91.9	14.5	60.2, 115.3	51.9	4.4	43.8, 60.8
11	68.6	21.3	24.6, 104.9	33.4	3.4	27.3, 40.5
12	45.7	22.0	8.6, 90.6	29.8	3.2	24.0, 36.3
13	23.1	18.4	0.9, 67.2	26.9	3.0	21.5, 33.0
14	81.7	15.5	50.0, 108.9	48.5	4.2	40.9, 57.1
15	70.0	16.7	36.4, 100.4	40.0	3.7	33.3, 47.6
16	58.3	17.2	25.9, 91.5	36.1	3.4	29.9, 43.0
17	46.7	17.0	16.6, 81.7	34.5	3.3	28.5, 41.3
18	35.0	16.0	8.8, 70.0	32.3	3.2	26.4, 38.9
19	23.3	14.0	3.3, 56.0	31.1	3.1	25.3, 37.6
20	11.5	10.5	0.3, 39.1	29.3	3.1	23.6, 35.8
21	71.5	16.4	38.8, 101.6	47.3	4.1	39.8, 55.9
22	61.5	16.8	29.6, 94.1	46.1	4.1	38.7, 54.4
23	51.2	16.8	20.8, 85.9	44.3	4.0	37.2, 52.5
24	41.1	16.1	13.8, 75.6	42.5	4.0	35.3, 50.7
25	30.8	14.8	7.1, 63.4	40.8	4.0	33.6, 48.9
26	20.6	12.9	2.6, 51.0	33.6	4.3	25.5, 42.5
27	10.5	9.6	0.3, 35.4	22.8	4.1	15.4, 31.4
28	59.8	16.9	28.2, 92.9	46.1	4.1	38.7, 54.4
29	44.8	17.7	13.8, 81.8	45.1	4.0	37.7, 53.5
30	29.9	16.5	4.7, 67.4	31.4	3.3	25.5, 38.2
31	14.8	13.0	0.4, 47.8	23.7	2.8	18.6, 29.6
32	48.0	17.0	18.4, 83.0	41.0	3.8	34.1, 48.7
33	36.1	16.0	10.3, 71.0	38.4	3.7	31.6, 45.9
34	24.3	14.2	3.9, 57.6	25.9	3.6	19.5, 33.4
35	12.2	10.9	0.4, 40.3	22.6	3.5	16.2, 29.9
36	24.2	14.2	3.7, 57.5	24.0	2.7	19.2, 29.6
37	12.1	10.8	0.3, 40.2	21.4	2.6	16.9, 26.8
38	85.0	18.8	51.2, 116.4	94.6	6.1	83.6, 106.7
39	43.4	26.8	2.3, 98.1	89.4	6.2	78.2, 101.7
40	42.4	27.0	1.8, 98.0	33.6	3.8	26.8, 41.5

obtain a better fit to the evolutionary process in which these sequences have evolved, we conclude that the Bayesian consensus topology is the best phylogenetic hypothesis among Columbiformes genera and will be the only one considered herein for further discussion.

Our Bayesian consensus phylogenetic hypothesis differs from hypotheses in two other studies that employed limited taxon sampling at the genus level (Johnson and Clayton, 2000; Shapiro et al., 2002). Using cyt *b* and FIB7 sequences, Johnson and Clayton (2000) recovered congruent maximum parsimony and maximum likelihood trees where clade B (as defined here) was a sister group to clades A and C, and clade C was not monophyletic. The only relationship of genera within a clade that differed in our phylogeny was the placement of *Treron* as sister to *Goura*. However, Johnson and Clayton (2000) did not sample *Turtur* and *Chalcophaps*, which, together with *Oena*, were more closely related to *Treron* than to *Goura* in our study. Additionally, Johnson and Clayton (2000) recovered a phylogenetic tree that had many nodes with low bootstrap support, especially within clade C and for the sister group relationships among the three major clades.

Although our phylogenetic hypothesis was not based on extensive sampling at the species level, it agrees with previous suggestions based on molecular (Johnson, 2004; Johnson and Clayton, 2000; Johnson et al., 2001) and nonmolecular data (Goodwin, 1983; Peters, 1937) that the New World *Patagioenas* (previously considered part of *Columba*) and the Old World *Columba* are not sister taxa. *Columba* was recovered in our analyses as sister to *Streptopelia* in agreement with other molecular phylogenetic analyses (Johnson, 2004; Johnson et al., 2001), but the closest sister genus to *Patagioenas* is still uncertain (this study; Johnson, 2004; Johnson et al., 2001) because it is part of a polytomous subclade within clade A (Figs. 1 to 3).

Shapiro et al. (2002) recovered a maximum likelihood tree based on 12s rDNA and cyt *b* sequences, with a well-supported sister relationship between clades B and C. Clade A was not recovered as a monophyletic assemblage in their analysis and was not well supported

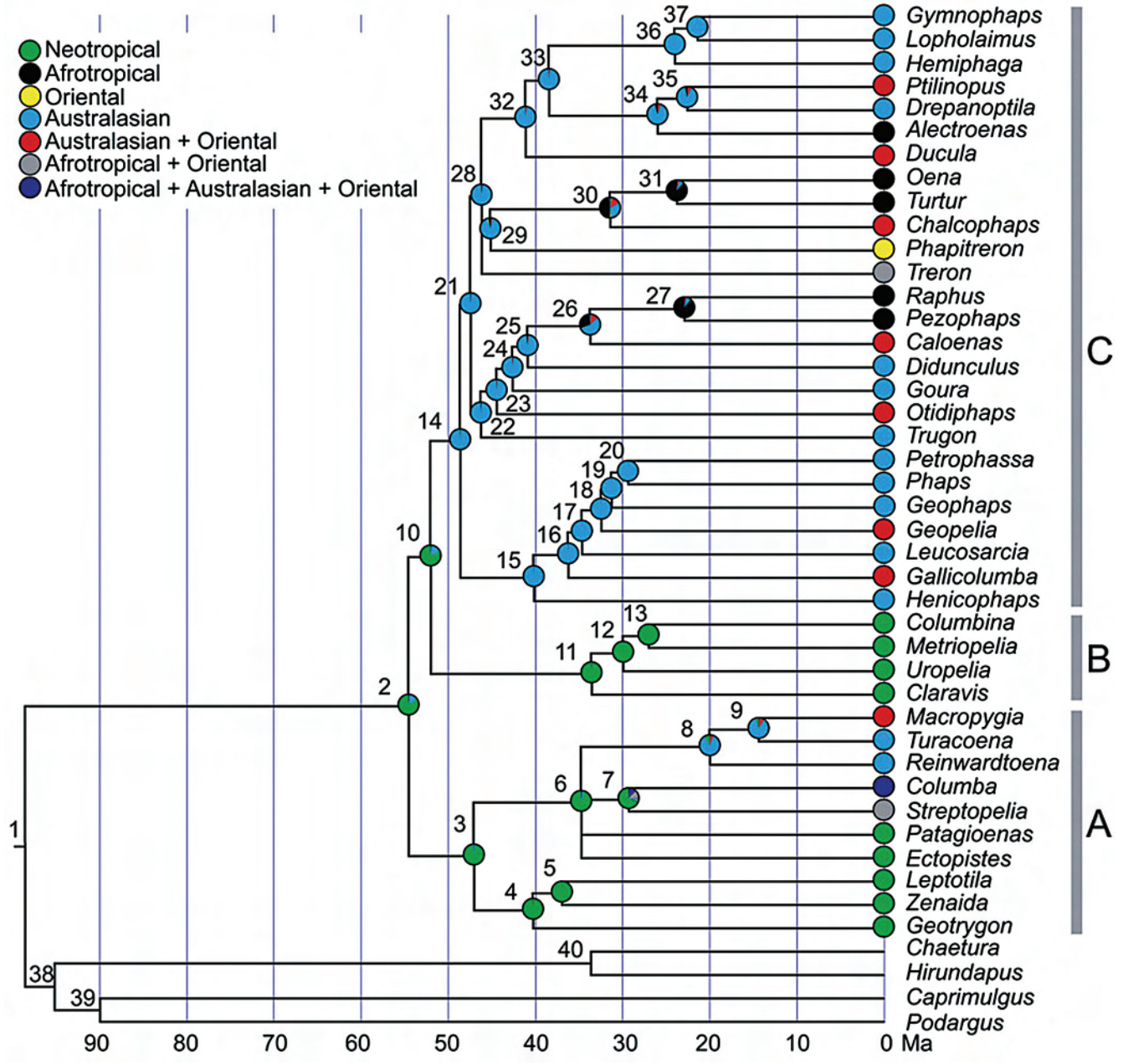

FIGURE 5. Chronogram of Columbiformes diversification and reconstruction of ancestral areas. Major clades within Columbiformes are indicated as A, B, and C. Branch lengths are given in Mya following the posterior distribution of divergence times shown in Table 3. Present biogeographic distribution is indicated as colored circles at the tips, and the proportion of the total likelihood received by each biogeographic region as the ancestral area of a given clade is represented by pie charts at internal nodes. Node numbers as in Table 3.

either by maximum likelihood bootstrapping or Bayesian posterior probability. Additionally, the differences observed between our phylogenetic hypothesis and that of Shapiro et al. (2002) related to those nodes that received low support in their tree, suggesting that the increased character and taxon sampling in our study largely improved tree resolution as expected (e.g., Pereira et al., 2002; Poe and Swofford, 1999). Furthermore, Shapiro et al. (2002) included three endemic (*Alectroenas* and the monotypic *Drepanoptila* and *Didunculus*) and three extinct genera (*Ectopistes*, *Pezophaps*, and *Raphus*) that we were not able to obtain samples of for further DNA sequencing. However, we included the 12S rDNA and cyt *b* sequences from Shapiro et al. (2002) in our Bayesian analysis and recovered these genera in similar positions. We also ran a second Bayesian analysis reducing our data set to match in length the available sequences for these five genera. The tree recovered (Fig. 4d) placed *Pezophaps* and *Raphus* (posterior probability [PP] = 0.99) as a sister clade to *Caloenas* (PP = 1.0), as

in Shapiro et al. (2002), but our tree recovered *Alectroenas* as a sister (PP = 1.0) to *Drepanoptila* and *Ptilinopus* (PP = 0.90), instead of placing *Alectroenas* and *Drepanoptila* as sister genera. These results have to be taken cautiously because we only sampled one species per genus and Shapiro et al. (2002) included three species of *Ptilinopus*, which were not monophyletic and did not receive strong support from either Bayesian posterior probabilities or maximum likelihood nonparametric bootstrapping in their analysis. Moreover, the Bayesian method of tree inference does not seem to be affected adversely by missing data in the complete data set, as predicted by simulation studies (Wiens, 2005), because most of the relationships within clades A, B, and C are recovered with moderate to high nodal support despite the inclusion of those taxa or not.

None of the several published classifications of Columbiformes based on nonmolecular data (revised in Sibley and Ahlquist, 1990) are congruent with the phylogenetic relationships recovered in our analyses. However, Goodwin (1983) presented a noncladistic hypothesis of relationships among genera based on behavior and plumage patterns, where clades A and C were not depicted as monophyletic assemblages, but some of his presumed genera associations were similar to the relationships recovered in the present study. Behavioral characters seem to be more phylogenetically conserved than morphological and osteological characters (Baker et al., 2006; Pereira and Baker, 2005). Hence, we might predict that that the rare, endemic genera *Cryptophaps* and *Starnoenas* and the extinct *Microgoura*, for which we were not able to obtain samples, are closely related to genera suggested by Goodwin (1983). Future sequencing and phylogenetic analysis will be necessary to formally place them in a phylogenetic framework and evaluate these predictions. At this point, we refrain from proposing a new classification at the family level and below because we feel that it is essential to integrate our data with revised nonmolecular data to forward a new classification for the order Columbiformes.

Divergence Times and Historical Biogeography of Columbiformes

We applied a Bayesian approach to estimate divergence times among Columbiformes genera that allowed data to be partitioned by gene and accounted for uncertainty in branch lengths and time constraints (Kishino et al., 2001; Thorne and Kishino, 2002; Thorne et al., 1998). Based on the 95% credible interval (CrI) of the Bayesian posterior distribution of age estimates, our molecular time estimates add the Columbiformes, Apodiformes, and Caprimulgiformes to the group of avian orders that originated in the Cretaceous and survived the Cretaceous-Tertiary mass extinction (Baker et al., 2006; Cooper and Penny, 1997; Nahum et al., 2003; Pereira and Baker, 2006a). Avian fossil evidence supports these molecular time estimates by clearly indicating that Palaeognathae and Galloanserae (Anseriformes and Galliformes) were already independent lineages in the Cretaceous (e.g., Clarke et al., 2005; Livezey, 1997). Some lineages within Neoaves, which is the sister group to Galloanserae, may also have been present in the Cretaceous, although the fossil record is too scanty and fragmentary to be unambiguously attributed to specific modern Neoaves (Cracraft, 2001; Dyke, 2001; Padian and Chiappe, 1998). Furthermore, the trans-Antarctic distributional pattern of Columbiformes and several other groups of birds (e.g., Neotropical and Australasian parrots, cracids and megapodes, and Apodiformes and Caprimulgiformes) suggest a Cretaceous age for these clades that became isolated in South America and Australia with the breakup of Gondwanaland and glaciation of Antarctica (Cracraft, 2001).

Considering the estimated Cretaceous age for modern Columbiformes lineages and their trans-Antarctic distribution, we tested whether the ancestors of the group may have lived in Gondwana by applying a maximum likelihood and a Bayesian framework to map the present-day distribution of genera onto the well-supported consensus Bayesian tree with time constraints and reconstructing ancestral areas (Fig. 5). Our results strongly support a Gondwanaland origin for Columbiformes (as suggested by Johnson and Clayton, 2000; Shapiro et al., 2002), implying diversification by vicariance due to continental drift and dispersal to the Afrotropical and Oriental biogeographic areas by members of clades A and C. Based on 95% CrI of the Bayesian posterior distribution of divergence times, the generic radiation within clade A, whose most recent common ancestor likely lived in the Neotropics (node 3 in Fig. 5), occurred between 39 and 55 Mya when the lineages leading to Neotropical genera became independent from those lineages that would lead to non-Neotropical genera. The ancestors of clade C were more likely to have been present in the Australasian region (node 14 in Fig. 5) before the group started to radiate around 41 to 57 Mya. Mapping ancestral areas onto the time-constrained Bayesian consensus tree (Fig. 5) suggests that the Oriental biogeographic zone seemed to have been colonized at least once by members of clade A and at least eight times by members of clade C, and the Afrotropical biogeographic zone at least once by members of clade A, and at least four times by members of clade C. It is very likely that the colonization route of both biogeographic zones would have been via Australasia judging from the reconstructed ancestral areas for the last common ancestor of Afrotropical and Oriental taxa (Fig. 5). The hypothesized dispersal events are summarized in Figure 6.

Considering the timing of diversification of Oriental and Afrotropical taxa (50 Mya and younger) and the geological history of the Oriental and Australasian biogeographic zones helps to understand why phylogenetic analyses of over 9.0 kb of mitochondrial and nuclear DNA sequences could not provide stronger support for some of the phylogenetic relationships within clade C as it did for clades A and B. In the Early Eocene, Africa and India were separate landmasses drifting away from the Antarctica–Australia–South America supercontinent (Li and Powell, 2001). This implies that dispersal

to Africa and India from Antarctican–Australia–South America would be unlikely because they were separated by large oceans at that time. Conversely, Australia separated from Antarctica around 40 Mya and started to drift north towards southeast Asia (Li and Powell, 2001), facilitating the exchange of biota among the Oriental and Australasian biogeographic zones (Hall, 1998, 2001; Li and Powell, 2001; Moss and Wilson, 1998; Pubellier et al., 2003). The exchange process may have been also favored by a considerable decrease in sea level (Haq et al., 1987) that likely resulted in greater land exposure between southeastern Asia and northern Australia, including several intermittent emergent Pacific Islands (Hall, 1998).

Once the Oriental biogeographic zone has been colonized, further dispersal into the Afrotropical and the Palearctic zones would have been a natural outcome because the physical barriers that delimit these zones (e.g., Sahara desert, Himalayas mountains) did not form until the Neogene (Kroepelin, 2006; Li and Powell, 2001). Additionally, the age of Afrotropical Columbiform taxa falls within the 95% CrI of 41 to 28 Mya for members of clade A (node 6 in Fig. 5) and 53 to 15 Mya for members of clade C (nodes 27 and 29 in Fig. 5). Dispersal into the Afrotropical and Palearctic zones may have been facilitated by the close proximity of the African and Eurasian plates (Bumby and Guiraud, 2005; Hall, 1998) and the beginning of the collision of the Indian subcontinent with Eurasia (Li and Powell, 2001; Zhu et al., 2005). This last event may have caused many islands to emerge between India and Africa during times of sea level fall, facilitating the dispersal of organisms between these landmasses. For example, a "stepping stone" route has been suggested previously for the extinct dodo and solitaires (*Raphus* and *Pezophaps*) of the Mauritius and Rodrigues Islands in the Mascarene Plateau (Shapiro et al., 2002). Our estimates for the split of dodos and solitaries from their sister genus *Caloenas* (posterior mean estimate = 33 Mya; 95% CrI 25 to 42 Mya) was younger than those of Shapiro et al. (2002; 42 Mya; 95% CI 31 to 56 Mya). On the other hand, our age estimates for the split of *Raphus* and *Pezophaps* (posterior mean estimate = 23 Mya; 95% CrI 15–31 Mya) is similar to that of Shapiro et al. (2002; 25 Mya; 95% confidence interval [CI] 17 to 36 Mya). Our analyses of nuclear and mitochondrial DNA sequences confirm suggestions that the ancestors of dodo and solitaires may have lived in other now-submerged islands in the Indian Ocean (Shapiro et al., 2002), tens of million years before Mauritius and Rodrigues emerged in the Mascarene Plateau (McDougall and Chamalaun, 1969).

In our reconstruction of ancestral areas, we did not consider Antarctica as a character state because no Columbiformes now occur in this continent. However, we cannot rule out the hypothesis that Antarctica and not the Neotropics was in fact the place where the ancestors of Columbiformes first evolved (Fig. 6). If the ancestors of Columbiformes were in Antarctica, the development of a permanent ice cap in Antarctica in the Oligocene and Late Miocene-to-present (Zachos et al., 2001) may have forced them to disperse northwards to more temperate and tropical climates (e.g., Baker et al., 2006; Tavares et al., 2006). Indeed, our estimates of divergence times indicate that the non-Neotropical genera of clade A last shared a common ancestor with Neotropical taxa about 28 to 41 Mya (node 6 in Fig. 5), coinciding with the beginning of glaciation in Antarctica (Zachos et al., 2001). However, most of the colonization events of the Oriental regions via Australia seemed to have occurred in the Eocene, prior to the Antarctic glaciation periods.

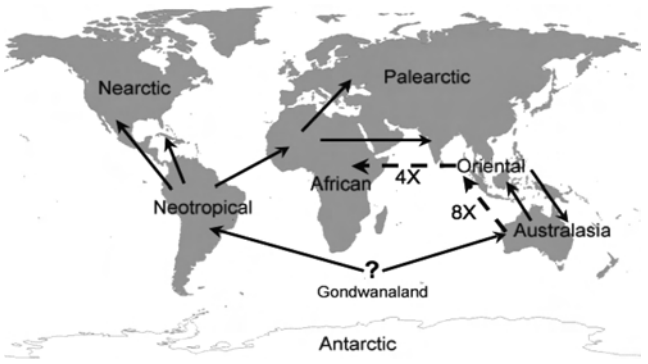

FIGURE 6. Summary of inferred dispersal events of Columbiformes mapped on present-day world. Zoogeographic zones are indicated. The question mark above Gondwanaland indicates the uncertainty about which region of the supercontinent may have harbored the ancestors of Columbiformes. Solid and dashed arrows represent dispersal events by members of clades A and C, respectively. Minimum numbers of independent dispersal events are indicated next to arrows when greater than one.

Although the breakup of Gondwanaland itself may not have been the ultimate cause of avian diversification, it created the necessary isolation for the ancestors of modern lineages to start their differentiation because of other reasons (Cracraft, 2001). As in many other groups of birds such as ratites, tinamous, galliform birds, penguins, shorebirds, parrots, passerine birds, toucans, and trogons (e.g., Baker et al., 2004, 2006; Crowe et al., 2006; Moyle, 2005; Nahum et al., 2003; Paton et al., 2003; Pereira and Baker, 2006a, 2006b; Pereira et al., 2002; Tavares et al., 2006), most cladogenetic events that led to extant genera of Columbiformes occurred throughout the Eocene and Oligocene epochs, and only a few events took place in the Miocene. These cladogenetic events were likely influenced by significant changes in Earth physiography that occurred during these epochs, like the formation of major mountain ranges, reshaping of river basins, sea level fluctuations, and climatic changes (Hall, 1998; Haq et al., 1987; Li and Powell, 2001; Lundberg et al., 1998; Zhu et al., 2005). Although it is assumed that these events caused major global changes in floral and faunal composition (MacFadden, 2000; Nores, 2004), it is a hard task to link which paleoevents are associated with particular generic divergences within Columbiformes at this moment. Future phylogenetic analysis and molecular dating at the species level might help refine the historical biogeography of the Columbiformes and evaluate the influence of specific continental paleoevents on their radiation (e.g., Baker et al., 2006; Pereira and Baker, 2004).

Our estimates of divergence time revealed a temporal gap of about 41 to 46 Mya between the origin of Columbiformes in the Late Cretaceous and diversification of genera within the order in the Eocene. Similar gaps of about 30 to 45 Mya over the same geological time span has also been observed for diversification of genera among Neotropical cracids (Pereira et al., 2002), toucans (Nahum et al., 2003), one of two major lineages of Neotropical parrots (Tavares et al., 2006), and penguins (Baker et al., 2006). Although diversification of extant genera in these clades of birds might suggest a major culling following the K/P bolide impact, other explanations are possible. Unfortunately, the fossil record for many avian groups is scanty or fragmentary, and thus inadequate to provide clues about the flowering of genera of most of the above groups of birds during the Late Cretaceous to Eocene. However, the Eocene fossil record of penguins indicates that stem penguins (reviewed in Ksepka et al., 2006) became extinct as extant penguins started to diversify (Baker et al. 2006). Baker et al. (2006) and Tavares et al. (2006) suggested that penguins and parrots dispersed from Antarctica to regions of warmer climates as Antarctica became ice-encrusted, and global climate started to cool down in the Eocene (e.g., Shevenell et al., 2004). Other birds do not show such a temporal gap, such as shorebirds, which are distributed worldwide, and a lineage of Neotropical parrots assumed to have been already in South America in the Early Eocene (Baker et al., 2007; Tavares et al., 2006). Hence, the lack of cladogenesis observed in certain avian groups prior to the Eocene may be more apparent than real, and the temporal gap may in fact represent a reshaping of the biodiversity of these groups during the early days of the Eocene. The diverse temporal pattern of avian diversification unveiled so far suggests that avian groups responded differently to the physiographic and climate changes of the Late Cretaceous and Cenozoic, and faunal turnover was greater in some groups than in others.

ACKNOWLEDGMENTS

We are grateful for Robert Blaquière for creating a database to filter and report the posterior probabilities of specific tree topologies among the thousands of alternatives in the Bayesian posterior distribution of trees. The following institutions provided tissues specimens: Australian National Wildlife Collection, Field Museum of Natural History, University of Kansas Museum of Natural History, Louisiana State University Museum of Natural Science, University of Washington Burke Museum, and U.S. National Museum of Natural History. KPJ was supported by National Science Foundation grants DEB-0107891 and DEB-0612938. DHC was supported by National Science Foundation grants DEB-0118794 and DEB-0614565. AJB was supported by an operating grant from the Natural Sciences and Engineering Research Council of Canada, the Royal Ontario Museum Foundation, and the National Science Foundation (Assembling the Tree of Life (AToL) Program—EF-0228693).

REFERENCES

Baker, A. J., S. L. Pereira, O. P. Haddrath, and E. K.-A. 2006. Multiple gene evidence for expansion of extant penguins out of Antarctica due to global cooling. Proc. R. Soc. Lond. B 273:11–17.

Baker, A. J., S. L. Pereira, and T. A. Paton. 2007. Phylogenetic relationships and divergence times of Charadriiformes genera: Multigene evidence for the Cretaceous origin of at least 14 clades of shorebirds. Biol. Lett. 3:205–209.

Banks, R. C., C. Cicero, J. L. Dunn, A. W. Kratter, P. C. Rasmussen, J. V. Remsen Jr., J. D. Rising, and D. F. Stotz. 2003. Forty-fourth supplement to the American Ornithologists' Union check-list of North American birds. Auk 120:923–931.

Barker, F. K., A. Cibois, P. Schikler, J. Feinstein, and J. Cracraft. 2004. Phylogeny and diversification of the largest avian radiation. Proc. Natl. Acad. Sci. USA 101:11040–11045.

Bollback, J. P. 2006. SIMMAP: Stochastic character mapping of discrete traits on phylogenies. BMC Bioinformatics 7:88.

Brodkorb, P. 1964. Catalogue of fossil birds, part 2 (Anseriformes through Galliformes). Bull Florida State Mus. Biol. Sci 8:195–335.

Bumby, A. J., and R. Guiraud. 2005. The geodynamic setting of the Phanerozoic basins of Africa. J. Afr. Earth Sci. 43:1–12.

Chiappe, L. M., and G. J. Dyke. 2002. The Mesozoic radiation of birds. Annu. Rev. Ecol. Syst. 33:91–124.

Clarke, J. A., C. P. Tambussi, J. I. Noriega, G. M. Erickson, and R. A. Ketcham. 2005. Definitive fossil evidence for the extant avian radiation in the Cretaceous. Nature 433:305–308.

Cooper, A., and D. Penny. 1997. Mass survival of birds across the Cretaceous-Tertiary boundary: Molecular evidence. Science 275:1109–1113.

Cracraft, J. 2001. Avian evolution, Gondwana biogeography and the Cretaceous-Tertiary mass extinction event. Proc. R. Soc. Lond. B 268:459–469.

Crowe, T. M., R. C. Bowie, P. Bloomer, T. G. Mandiwana, T. A. J. Hedderson, E. Randi, S. L. Pereira, and J. Wakeling. 2006. Phylogenetics, biogeography and classification of, and character evolution in, gamebirds (Aves: Galliformes): Effects of character exclusion, data partitioning and missing data. Cladistics 22:495–532.

Darwin, C. 1859. On the origin of species by means of natural selection or the preservation of favoured races on the struggle for life. John Murray, London.

Dyke, G. J. 2001. The evolutionary radiation of modern birds: Systematics and patterns of diversification. Geol. J. 36:305–315.

Dyke, G. J., and M. Van Tuinen. 2004. The evolutionary radiation of modern birds (Neornithes): Reconciling molecules, morphology and the fossil record. Zool. J. Linn. Soc. 141:153–177.

Ericson, P. G., C. L. Anderson, T. Britton, A. Elzanowski, U. S. Johansson, M. Kallersjo, J. I. Ohlson, T. J. Parsons, D. Zuccon, and G. Mayr. 2006. Diversification of Neoaves: Integration of molecular sequence data and fossils. Biol. Lett. 2:543–547.

Gibbs, D., E. Barnes, and J. Cox. 2001. Pigeons and Doves. A guide to the pigeons and doves of the world. Pica Press, East Sussex, UK.

Goodwin, D. 1983. Pigeons and doves of the world, 3rd. edition. Comstock Publishing Associates, Ithaca, NY.

Gradstein, F. M., J. G. Ogg, A. G. Smith, F. P. Agterberg, W. Bleeker, R. A. Cooper, V. Davydov, P. Gibbard, L. A. Hinnov, M. R. House, L. Lourens, H. P. Luterbacher, J. McArthur, M. J. Melchin, L. J. Robb, J. Shergold, M. Villeneuve, B. R. Wardlaw, J. Ali, H. Brinkhuis, F. J. Hilgen, J. Hooker, R. J. Howarth, A. H. Knoll, J. Laskar, S. Monechi, K. A. Plumb, J. Powell, I. Raffi, U. Röhl, P. Sadler, A. Sanfilippo, B. Schmitz, N. J. Shackleton, G. A. Shields, H. Strauss, J. Van Dam, T. van Kolfschoten, J. Veizer, and D. Wilson. 2004. A geologic time scale 2004. Cambridge University Press, Cambridge, UK.

Groth, J. G., and G. F. Barrowclough. 1999. Basal divergences in birds and the phylogenetic utility of the nuclear RAG-1 gene. Mol. Phylogenet. Evol. 12:115–123.

Guindon, S., and O. Gascuel. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52:696–704.

Haddrath, O., and A. J. Baker. 2001. Complete mitochondrial DNA genome sequences of extinct birds: Ratite phylogenetics and the vicariance biogeography hypothesis. Proc. R. Soc. Lond. B 268:939–945.

Hall, R. 1998. The plate tectonics of Cenozoic SE Asia and the distribution of land and sea. Pages 99–131 *in* Biogeography and geological evolution of SE Asia (R. Hall and J. D. Holloway, eds.). Backhuys Publishers, Leiden, The Netherlands.

Hall, R. 2001. Cenozoic reconstructions of SE Asia and the SW Pacific: Changing patterns of land and sea. Pages 35–56 *in* Faunal and floral migrations and evolution in SE Asia-Australasia (I. Metcalfe, J. M. B.

Smith, M. Morwood, and I. D. Davidson, eds.). Swets & Zeitlinger Publishers, Lisse.
Haq, B. U., J. Hardenbol, and P. R. Vail. 1987. Chronology of fluctuating sea levels since the Triassic. Science 235:1156–1167.
Hasegawa, M., H. Kishino, and T. Yano. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22:160–174.
Hedges, S. B., P. H. Parker, C. G. Sibley, and S. Kumar. 1996. Continental breakup and the ordinal diversification of birds and mammals. Nature 381:226–9.
Huelsenbeck, J. P., R. Nielsen, and J. P. Bollback. 2003. Stochastic mapping of morphological characters. Syst. Biol. 52:131–158.
Johnson, K. P. 2004. Deletion bias in avian introns over evolutionary timescales. Mol. Biol. Evol. 21:599–602.
Johnson, K. P., and D. H. Clayton. 2000. Nuclear and mitochondrial genes contain similar phylogenetic signal for pigeons and doves (Aves: Columbiformes). Mol. Phylogenet. Evol. 14:141–51.
Johnson, K. P., S. de Kort, K. Dinwoodey, A. C. Mateman, C. ten Cate, C. M. Lessells, and D. H. Clayton. 2001. A molecular phylogeny of the dove genera *Streptopelia* and *Columba*. Auk 118:874–887.
Kishino, H., J. L. Thorne, and W. J. Bruno. 2001. Performance of a divergence time estimation method under a probabilistic model of rate evolution. Mol. Biol. Evol. 18:352–361.
Kroepelin, S. 2006. Revisiting the age of the Sahara Desert (letters). Science 312:1138.
Ksepka, D. T., S. Bertelli, and N. P. Giannini. 2006. The phylogeny of the living and fossil Sphenisciformes (penguins). Cladistics 22:412–441.
Lewis, P. O. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. Syst. Biol. 50:913–925.
Li, Z. X., and C. M. Powell. 2001. An outline of the palaeogeographic evolution of the Australasian region since the beginning of the Neoproterozoic. Earth-Sci. Rev. 53:237–277.
Livezey, B. C. 1997. A phylogenetic analysis of basal Anseriformes, the fossil *Presbyornis*, and the interordinal relationships of waterfowl. Zool. J. Linn. Soc. 121:361–428.
Livezey, B. C., and R. L. Zusi. 2007. Higher-order phylogeny of modern birds (Theropoda, Aves: Neornithes) based on comparative anatomy. II. Analysis and discussion. Zool. J. Linn. Soc. 149:1–95.
Lundberg, J. G., L. G. Marshall, J. Guerrero, B. Horton, M. C. S. L. Malabarba, and F. Wesselingh. 1998. The stage for Neotropical fish diversification: A history of tropical South American rivers. Pages 13–48 *in* Phylogeny and classification of neotropical fishes (L. R. Malabarba, R. E. Reis, R. P. Vari, Z. M. S. Lucena, and C. A. S. Lucena, eds.). Editora Universitaria PUCRS, Porto Alegre, Brazil.
MacFadden, B. J. 2000. Cenozoic mammalian herbivores from the Americas: Reconstructing ancient diets and terrestrial communities. Annu. Rev. Ecol. Syst. 31:33–59.
Maddison, D. R., and W. P. Maddison. 2000. MacClade 4.0. Sinauer Associates, Sunderland, Massachusetts.
Maddison, W. P., and D. R. Maddison. 2005a. Mesquite: A modular system for evolutionary analysis, version 1.06. http://mesquiteproject.org.
Maddison, W. P., and D. R. Maddison. 2005b. Stochar: A package of Mesquite modules for stochastic models of character evolution, version 1.06. http://mesquiteproject.org.
McDougall, I., and F. H. Chamalaun. 1969. Isotopic dating and geomagnetic polarity studies on volcanic rocks from Mauritius, Indian Ocean. Geol. Soc. Am. Bull. 80:419–442.
Moss, S. J., and M. E. J. Wilson. 1998. Biogeographic implications of the Tertiary palaeoeographic evolution of Sulawesi and Borneo. Pages 133–163 *in* Biogeography and geological evolution of SE Asia (R. Hall and J. D. Holloway, eds.). Backhuys Publishers, Leiden, The Netherlands.
Moyle, R. 2005. Phylogeny and biogeographical history of Trogoniformes, a pantropical bird order. Biol. J. Linn. Soc. 84:725–738.
Nahum, L. A., S. L. Pereira, F. M. D. Fernandes, S. R. Matioli, and A. Wajntal. 2003. Diversification of Ramphastinae (Aves, Ramphastidae) prior to the Cretaceous/Tertiary boundary as shown by molecular clock of mtDNA sequences. Genet. Mol. Biol. 26:411–418.
Nores, M. 2004. The implications of Tertiary and Quaternary sea level rise events for avian distribution patterns in the lowlands of northern South America. Global Ecol. Biogeogr. 13:149–161.
Nylander, J. A. 2004. MrModelTest 2.0. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.
Olson, S. L. (ed.) 1985. The fossil record of birds. Academic Press, New York.
Padian, K., and L. M. Chiappe. 1998. The origin and early evolution of birds. Biol. Rev. 73:1–42.
Paton, T. A., A. J. Baker, J. G. Groth, and G. F. Barrowclough. 2003. RAG-1 sequences resolve phylogenetic relationships within Charadriiform birds. Mol. Phylogenet. Evol. 29:268–278.
Pereira, S. L., and A. J. Baker. 2004. Vicariant speciation of curassows (Aves, Cracidae): A hypothesis based on mitochondrial DNA phylogeny. Auk 121:682–694.
Pereira, S. L., and A. J. Baker. 2005. Multiple gene evidence for parallel evolution and retention of ancestral morphological states in the shanks (Charadriiformes: Scolopacidae). Condor 107:514–526.
Pereira, S. L., and A. J. Baker. 2006a. A mitogenomics timescale for birds detects variable phylogenetic rates of molecular evolution and refutes the standard molecular clock. Mol. Biol. Evol. 23:1731–1740.
Pereira, S. L., and A. J. Baker. 2006b. A molecular timescale for galliform birds accounting for uncertainty in time estimates and heterogeneity of rates of DNA substitutions across lineages and sites. Mol. Phylogenet. Evol. 38:499–509.
Pereira, S. L., A. J. Baker, and A. Wajntal. 2002. Combined nuclear and mitochondrial DNA sequences resolve generic relationships within the Cracidae (Galliformes, Aves). Syst. Biol. 51:946–958.
Peters, J. L. 1937. Check-list of birds of the world. Harvard University Press, Cambridge, Massachusetts.
Poe, S., and D. L. Swofford. 1999. Taxon sampling revisited. Nature 398:299–300.
Prychitko, T. M., and W. S. Moore. 1997. The utility of DNA sequences of an intron from the beta-fibrinogen gene in phylogenetic analysis of woodpeckers (Aves: Picidae). Mol. Phylogenet. Evol. 8:193–204.
Pubellier, M., J. Ali, and C. Monnier. 2003. Cenozoic plate interaction of the Australia and Philippine sea plates: "Hit-and-run" tectonics. Tectonophysics 363:181–199.
Ronquist, F., and J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–4.
Sclater, P. L. 1858. On the general geographic distribution of the members of the class Aves. J. Proc. Linn. Soc. Zool. 2:130–145.
Shapiro, B., D. Sibthorpe, A. Rambaut, J. Austin, G. M. Wragg, O. R. Bininda-Emonds, P. L. Lee, and A. Cooper. 2002. Flight of the dodo. Science 295:1683.
Shevenell, A. E., J. P. Kennett, and D. W. Lea. 2004. Middle Miocene Southern Ocean cooling and Antarctic cryosphere expansion. Science 305:1766–1770.
Shimodaira, H. 2002. An approximately unbiased test of phylogenetic tree selection. Syst. Biol. 51:492–508.
Shimodaira, H., and M. Hasegawa. 2001. CONSEL: For assessing the confidence of phylogenetic tree selection. Bioinformatics 17:1246–1247.
Sibley, C. G., and J. E. Ahlquist. 1990. Phylogeny and classification of Birds: A study in molecular evolution. Yale University Press, New Haven, Connecticut.
Slack, K. E., C. M. Jones, T. Ando, G. L. Harrison, R. E. Fordyce, U. Arnason, and D. Penny. 2006. Early penguin fossils, plus mitochondrial genomes, calibrate avian evolution. Mol. Biol. Evol. 23:1144–1155.
Strimmer, K., and A. von Haeseler. 1996. Quartet puzzling: A quartet maximum likelihood method for reconstructing tree topologies. Mol. Biol. Evol. 13:964–969.
Swofford, D. L. 2001. PAUP*: Phylogenetic analysis using parsimony (*and other methods), version 4.0. Sinauer Associates, Sunderland, Massachussets.
Tavares, E. S., A. J. Baker, S. L. Pereira, and C. Y. Miyaki. 2006. Phylogenetic relationships and historical biogeography of neotropical parrots (Psittaciformes: Psittacidae: Arini) inferred from mitochondrial and nuclear DNA sequences. Syst. Biol. 55:454–70.
Thorne, J. L., and H. Kishino. 2002. Divergence time and evolutionary rate estimation with multilocus data. Syst. Biol. 51:689–702.

Thorne, J. L., H. Kishino, and I. S. Painter. 1998. Estimating the rate of evolution of the rate of molecular evolution. Mol. Biol. Evol. 15:1647–57.

van Tuinen, M., and S. B. Hedges. 2001. Calibration of avian molecular clocks. Mol. Biol. Evol. 18:206–13.

Wallace, A. R. 1876. The geographic distribution of animals. Macmillan, London.

Wiegmann, B. M., D. K. Yeates, J. L. Thorne, and H. Kishino. 2003. Time flies, a new molecular time-scale for brachyceran fly evolution without a clock. Syst. Biol. 52:745–56.

Wiens, J. J. 2005. Can incomplete taxa rescue phylogenetic analyses from long-branch attraction? Syst. Biol. 54:731–742.

Yang, Z. 1997. PAML: A program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. 13:555–556.

Yang, Z., and A. D. Yoder. 2003. Comparison of likelihood and Bayesian methods for estimating divergence times using multiple gene loci and calibration points, with application to a radiation of cute-looking mouse lemur species. Syst. Biol. 52:705–716.

Zachos, J., M. Pagani, L. Sloan, E. Thomas, and K. Billups. 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. Science 292:686–693.

Zhu, B., W. S. F. Kidd, D. B. Rowley, B. S. Currie, and N. Shafique. 2005. Age of initiation of the India-Asia collision in the East-Central Himalaya. J. Geol. 113:265–285.

Zwickl, D. J. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. PhD Dissertation, The University of Texas at Austin.

First submitted 7 November 2006; reviews returned 26 February 2007; final acceptance 19 May 2007

Associate Editor: Adrian Paterson