Points of View

Syst. Biol. 51(4):614–625, 2002
DOI: 10.1080/10635150290102294

Examining Basal Avian Divergences with Mitochondrial Sequences: Model Complexity, Taxon Sampling, and Sequence Length

EDWARD L. BRAUN AND REBECCA T. KIMBALL[1]

Department of Zoology, 223 Bartram Hall, P.O. Box 118525, University of Florida, Gainesville, Florida 32611-8525, USA; E-mail: ebraun@zoo.ufl.edu (E.L.B.), rkimball@zoo.ufl.edu (R.T.K.)

Traditional avian classifications have generally indicated that the paleognathous birds (ratites and tinamous) represent the earliest divergence within the extant birds and the perching birds (Passeriformes) represent one of the most recently derived lineages (reviewed by Sibley and Ahlquist, 1990; Feduccia, 1995; see Fig. 1A). However, the conclusions of several recent studies using complete mitochondrial genome sequences from birds have challenged these classifications (Mindell et al., 1997, 1999; Härlid and Arnason, 1999; Haring et al., 2001; but see Paton et al., 2002). In particular, these studies have suggested that the passeriforms are one of the most ancient groups of extant birds and that the paleognaths are derived. This conclusion has profound implications for avian evolution. An ancient origin of passerines could explain the remarkable diversity of this group, which represents over half of modern avian species, without invoking different rates of cladogenesis and/or extinction in the passerines and other groups of birds. Likewise, a derived position of paleognaths would require a neotenous origin of paleognath morphological characters (Härlid and Arnason, 1999) and a different model of avian sex chromosome evolution (Ellegren and Carmichael, 2001).

Although complete avian mitochondrial sequences provide a large number of sites, it seems unlikely that mitochondrial data have provided an accurate picture of avian evolution. The extensive DNA–DNA hybridization studies conducted by Sibley and Ahlquist (1990), which used midpoint rooting, indicate a basal position for paleognaths and derived position for passerines. In addition, studies examining either immunological distances or the sequences of individual nuclear genes are congruent with the DNA–DNA hybridization studies (Prager and Wilson, 1976; Prager et al., 1976; Stapel et al., 1984; Caspers et al., 1997; Groth and Barrowclough, 1999; Garcia-Moreno and Mindell, 2000). Studies of strings present in DNA sequences also support the basal placement of paleognaths (Edwards et al., 2002). Taken as a whole, the results of these studies indicate that avian nuclear genomes have a history consistent with traditional classifications regarding the basal placement of paleognaths.

The apparent conflict between the nuclear and mitochondrial partitions in birds may reflect historical processes, and Johnson (2001) suggested that this possibility deserves further study. Gene trees may differ from species trees for several different reasons (reviewed by Maddison, 1997), including gene duplication, lineage sorting, and hybridization. The comparison of nonorthologous sequences, which could reflect gene duplication, should not affect mitochondrial sequences. Lineage sorting requires the maintenance of ancestral polymorphisms, which is unlikely to occur for periods of much longer than a few million years (Moore, 1995). In addition, it is unlikely that the radiation of extant birds occurred rapidly enough that lineage sorting could

[1]E.L.B. and R.T.K. contributed equally to this manuscript and are listed in alphabetical order.

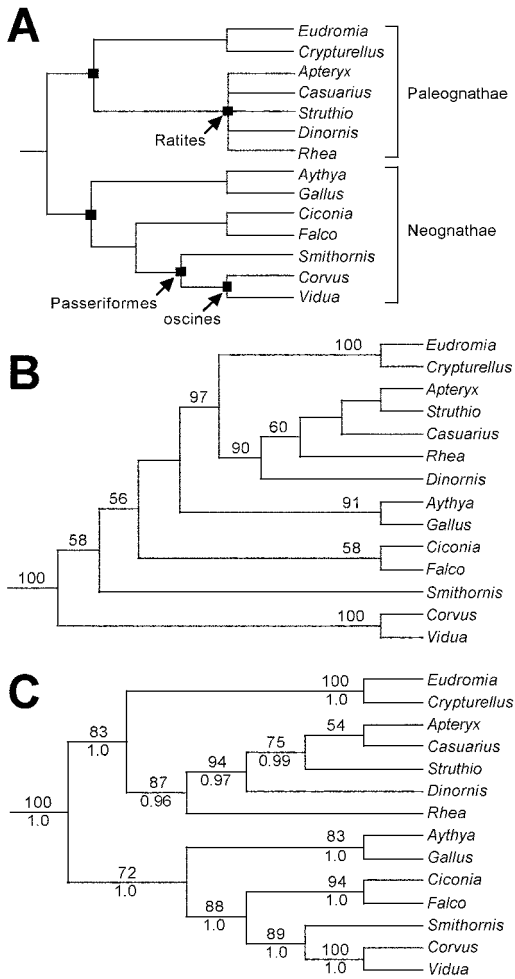

FIGURE 1. Phylogenetic hypotheses for basal avian relationships. Expected relationships were based upon various sources (A) or suggested by equally weighted parsimony (B) and transversion parsimony (C). Major nodes discussed are indicated with boxes in A. In B and C, only ingroup taxa are presented, and the same four outgroup taxa were used. Bootstrap support, estimated using 500 replicates, is indicated above the relevant branches on bootstrap consensus trees. Posterior probabilities of groups from a Bayesian analysis (using the GTR + Γ + inv model) are presented below the relevant branches for topology C.

provide an explanation for the observed conflict. If ancient hybridizations were common, some nuclear gene trees should be congruent with mitochondrial gene trees. Thus, it seems unlikely that the apparent conflict reflects historical differences between mitochondrial and nuclear sequences.

The limited sampling of complete avian mitochondria may also have led to the unexpected conclusions obtained from mitochondrial sequences (van Tuinen et al., 2000; Johnson, 2001). The complete mitochondrial sequences analyzed by Mindell et al. (1997, 1999) represented only six avian orders, about 17% of the extant orders. Two recent studies have explicitly attempted to determine whether limited taxon samples have led to unexpected conclusions regarding avian evolution. Van Tuinen et al. (2000) collected mitochondrial 12S and 16S ribosomal DNA (rDNA) sequences and nuclear 18S rDNA sequences from all avian orders and reported that both nuclear and mitochondrial data sets supported traditional classifications regarding basal placement of paleognaths. More recently, Johnson (2001) analyzed nearly 1,000 mitochondrial cytochrome *b* sequences representing a wide variety of avian taxa. Even with this dense taxon sampling, both parsimony and distance analyses could not recover a phylogeny consistent with traditional classifications. Clearly, taxon sampling does not completely explain why phylogenies inferred using mitochondrial DNA differ from those inferred using other types of data.

The observed conflict between mitochondrial sequences and the nuclear genome may reflect, at least in part, difficulties in the analysis of mitochondrial data. Mitochondrial sequences have highly biased base compositions, particularly at the third codon positions (e.g., Voelker and Edwards, 1998). Although transitions accumulate more rapidly than transversions, the two classes of transitions occur at very different rates, as do the four classes of transversions (Mindell et al., 1997). In addition, mitochondrial sequences evolve very rapidly, particularly at synonymous positions (Braun and Kimball, 2001). However, to maintain the function of mitochondrial genes, nonsynonymous sites show complex patterns of constraint (e.g., Naylor et al., 1995). Thus, over the evolutionary history of birds, there are large differences in the expected number of changes at different sites. In the analyses conducted by Mindell et al. (1999), more complex models supported a phylogeny compatible with traditional avian classification when certain sets of taxa were examined (Mindell et al., 1999: tables 1, 2, 5). However, a phylogeny with passerines at a basal position could not be rejected by these analyses, and Mindell et al. (1999) concluded that the basal passerine phylogeny was the most highly corroborated hypothesis.

Recently, Paton et al. (2002) recovered a tree compatible with traditional avian classification by excluding third codon positions and analyzing the data with complex models. However, the focus of Paton et al. (2002) was estimating divergence times, and they did not examine support for their topology.

We further examined the ability of mitochondrial coding sequences to address basal divergences in birds by evaluating the effect of taxon sampling, sequence length, and different weighting schemes or evolutionary models on the inferrence of basal avian relationships. We compiled a data set containing almost 11 kilobases (kb) of mitochondrial data (all light-strand coding regions) for 14 avian and 4 outgroup taxa and examined the impact of different types of analysis, taxon samples, and sequence length on the inference of basal avian relationships. To extend these results, we compiled a data set of 130 cytochrome *b* sequences chosen to provide a balanced representation for the majority of avian orders and outgroup taxa and analyzed these data using several different weighting schemes. Our analyses provided evidence that there is strong signal in mitochondrial genomes that is compatible with traditional views of avian evolution, although these results are dependent upon taxon sampling, sequence length, and analytical approach.

METHODS

All sequences were obtained from GenBank. For the 18-taxon mitochondrial data set, only the light-strand coding regions (all coding regions except ND6) were aligned. We analyzed data from 14 avian taxa: seven paleognath species: *Crypturellus tataupa* (Tataupa Tinamou; AY016012), *Eudromia elegans* (Elegant Crested-Tinamou; AY016016), *Apteryx mantelli* (Brown Kiwi; AY016010), *Casuarius casuarius* (Southern Cassowary; NC_002778), *Struthio camelus* (Ostrich; NC_002785), *Dinornis giganteus* (Giant Moa; NC_002672), and *Rhea americana* (Greater Rhea; NC_000846); and seven neognath species: *Aythya americana* (Redhead; NC_000877), *Gallus gallus* (Chicken; NC_001323), *Ciconia boyciana* (Oriental Stork; NC_002196), *Falco peregrinus* (Peregrine Falcon; NC_000878), *Smithornis sharpei* (Grey-headed Broadbill; NC_000879), *Corvus frugilegus* (Rook; NC_002069), and *Vidua chalbeata* (Village Indigobird; NC_000880). Four reptilian taxa were used as outgroups: *Alligator mississippiensis* (American Alligator; NC_001922), *Caiman crocodilus* (Spectacled Caiman; NC_002744), *Chelonia mydas* (Green Seaturtle; NC_000886), and *Chrysemys picta* (Painted Turtle; NC_002073).

For the 130-taxon cytochrome *b* alignment, only sequences that were at least 1 kb in length were included. Taxa in this data set were selected to provide a balanced sampling of most avian orders, breaking up branches by analyzing representatives from multiple families without overrepresenting well-studied lineages. Therefore, the data set was limited to a maximum of one sequence per genus, four per family, and eight per order (using the taxonomy provided by the taxonomy browser at GenBank). The final data set represented 126 avian species from 34 avian orders (6 paleognath and 28 neognath orders), and the same 4 outgroup taxa used in the 18-taxon mitochondrial alignment.

All data were analyzed using PAUP* 4.b8 (Swofford, 1999) unless otherwise noted. For all alignments, including the 130-taxon cytochrome *b* data set, the most parsimonious (MP) tree(s) was obtained using a heuristic search with 100 random addition sequence replicates. Separate analyses were performed using eight weighting schemes: (1) equal weighting with all nucleotide data, (2) transversions only, (3) transversions at third positions (but both transitions and transversions at first and second positions), (4) first and second positions only, (5) second positions only, (6) nonsynonymous coding, which excludes synonymous changes, (7) six-parameter inverse weighting as described by Cunningham (1997), and (8) a modified trilevel weighting (Benabib et al., 1997; Flores-Villela et al., 2000). Modified trilevel weighting involved three matrices: (1) the original data, (2) transversions only, and (3) nonsynonymous coded sites (rather than amino acids as described by Benabib et al. [1997] and Flores-Villela et al. [2000]). Modified trilevel weighting facilitated bootstrap analysis.

In all cases, bootstrap analyses were performed using 500 replicates and a heuristic search with 10 random addition sequence replicates. Bootstrap analysis of trilevel weighted data is problematic, because the weights reflect duplication of individual nucleotides in the data matrix but a standard

bootstrap analysis would treat these recoded nucleotides as independent data. To overcome this limitation, a C++ program was written to bootstrap trilevel weighted data, sampling elements in each of the matrices in a linked fashion (details available from E.L.B.).

For the 18-taxon mitochondrial alignment, maximum likelihood (ML) scores were calculated for the MP topologies we obtained. ML scores were calculated using six different models of DNA sequence evolution: Jukes–Cantor (Jukes and Cantor, 1969), F81 (Felsenstein, 1981), HKY85 (Hasegawa et al., 1985), general time reversible (GTR; Yang, 1994) GTR + Γ, and GTR + Γ + invariant sites. For transversion (two state) data, four different models of evolution were examined: two states with equal state frequencies (the CF model), CF with unequal state frequencies (CFu), CFu with Γ-distributed rates across sites (CFu + Γ), and CFu + Γ with invariant sites (CFu + Γ + inv). Parameter values were obtained by ML estimation, calculating parameters independently for each topology and model. The model exhibiting the best fit to each data set was determined using the likelihood ratio test (Huelsenbeck and Crandall, 1997).

ML estimates of phylogeny for the 18-taxon mitochondrial alignment were obtained using either the GTR + Γ + inv model or the CFu + Γ + inv model, because these models exhibited the best fit to the data. A heuristic search with 10 random addition sequence replicates and TBR branch swapping was performed. Parameters were initially estimated from the most likely of the MP trees identified using equally weighted parsimony. Parameters were then reoptimized using the topology found in the initial search, and a second heuristic search was performed. The one-tailed SH test (Shimodaira and Hasegawa, 1999) was conducted by estimating parameters independently for each topology, which had been identified by parsimony analyses, and using 1,000 RELL bootstrap replicates.

Bayesian analyses were conducted in MrBayes 2.01 (Huelsenbeck and Ronquist, 2001) using the GTR + Γ + invariant sites model. Default priors were used for the base frequencies, rate matrix, branch lengths, shape parameter for the Γ distribution, and proportion of invariant sites. An uninformative prior was used for the topology. The analysis of the full data set was conducted using 10^6 generations with four chains (with three chains "heated" using the default parameters) and sampling from the Markov chain every 100th cycle. Comparisons of five independent runs, each using four heated chains started at random points in parameter space using the full mitochondrial alignment, suggested that the chains converged rapidly (in the first 2.5×10^4 generations). To ensure sampling of topologies after chain convergence, we discarded the first 500 trees (5×10^4 generations) as "burn-in."

To examine the sensitivity of the conclusions to taxon sampling, we performed single-taxon deletions of the 14 avian taxa and 4 outgroup taxa. For each data set (after one taxon had been removed), we performed both a heuristic search for the MP tree(s) and bootstrap using equally weighted and transversion parsimony data sets. For the taxon deletion analyses, we also performed ML estimation of phylogeny using both all nucleotide data and transversion-only data in a heuristic search with a single random addition and TBR branch swapping. Parameter estimates were obtained from the MP tree (using the first tree in memory for multiple MP trees). Bayesian analyses for different taxon samples were conducted as described above but using 10^5 generations, because the Markov chain for the complete alignment appeared to converge rapidly.

We performed additional taxon deletions, creating data sets of various sizes that contained the four outgroup taxa. Deletions of outgroups were then performed. Phylogenetic analyses of these data sets were conducted as described above.

To examine the effect of number of sites on the resulting topology, we performed jackknife deletions of sequence data for the 18-taxon mitochondrial alignment and for a smaller 11-taxon alignment. The deletion proportions were set to provide random data sets of 1,000, 2,000, 4,000, or 8,000 base pairs. For each data set, we examined 1,000 jackknife replicates for both equally weighted and transversion parsimony, using heuristic searches with 10 random addition sequence replicates per jackknife replicate. These analyses were performed using all 18 taxa, the 11 taxa (which included the 7 avian taxa used by Mindell et al. [1999], except that *Corvus* was substituted for *Vidua*).

Stationarity of base composition for mitochondrial DNA open reading frames was examined by calculating the δ_{bf} statistic proposed by Gillespie (1986) for each pair of sequences (designated sequences *a* and *b*) using a C++ program written by E.L.B.:

$$\hat{\delta}_{bf} = \sum_{i=1}^{4}(\hat{\pi}_{ai} - \hat{\pi}_{bi})^2 \Big/ \sum_{i=1}^{4}[2\pi_i(1-\pi_i)/n]$$

where n is the sequence length and $\hat{\pi}_i$ refers to the observed proportion of the ith nucleotide in sequence a or sequence b. Following Gillespie (1986), we assume that $\pi_i = (\hat{\pi}_{ai} - \hat{\pi}_{bi})/2$. A similar value, designated δ_{tv}, can be calculated using transversion data by modifying the equation to sum over the two character states (purines and pyrimidines) possible. When sequences diverge under a stationary model of sequence evolution, the expected value of δ_{bf} or δ_{tv} changes from 0 to 1 with values >1 providing evidence against stationarity.

Trees were estimated from δ_{bf} and δ_{tv} matrices using least squares (with power = 2; Fitch and Margoliash, 1967). The optimal topology was identified by TBR branch swapping of an initial neighbor joining tree. Least squares was also used with LogDet/Paralinear distances (Lake, 1994; Lockhart et al., 1994). Among-sites rate variation was accommodated by assuming that a specific proportion of sites was invariant (Swofford et al., 1996). This proportion was estimated by ML using the best fitting models.

RESULTS

Mitochondrial Phylogenies Are Sensitive to the Type of Analysis Conducted

Analyses of the 18-taxon mitochondrial alignment indicated that the estimate of avian phylogeny obtained was sensitive to type of analysis. Using equally weighted parsimony of all nucleotides, the passerine taxa were basal and paleognaths were derived (Fig. 1B). Similar topologies were found in parsimony analyses using third transversion coding, six-parameter inverse weighting, first and second positions only, second positions only, nonsynonymous sites, and trilevel weighting (data not shown).

In sharp contrast, the results of transversion parsimony (Fig. 1C) were largely consistent with expectations (Fig. 1A) based upon traditional avian classifications (e.g., Mayr and Amadon, 1951; Wetmore, 1960), DNA–DNA hybridization analyses (Sibley and Ahlquist, 1990), and analysis of nuclear gene sequences or immunological data (Prager and Wilson, 1976; Prager et al., 1976; Stapel et al., 1984; Caspers et al., 1997; Groth and Barrowclough, 1999; Garcia-Moreno and Mindell, 2000; van Tuinen et al., 2000). The high degree of consistency found between the transversion parsimony tree (Fig. 1C) and other phylogenetic studies of basal avian relationships (Fig. 1A) suggested that this tree is more likely to be correct than the topologies identified using other weighting schemes (e.g., Fig. 1B). The analyses of 12S and 16S mitochondrial rDNA sequences performed by van Tuinen et al. (2000), which placed paleognaths basal and supported neognath monophyly, used transversion distances.

To identify the best fitting model of sequence evolution in an ML framework, we compared likelihood scores for a set of standard nested models using the MP trees identified with equally weighted and transversion parsimony (Table 1). Using all data, simpler models (Jukes-Cantor and F81) favored the MP tree identified using equally weighted parsimony, whereas more complex models favored the transversion parsimony tree (Table 1). All transversion likelihood analyses favored the topology obtained using transversion parsimony. The increase in likelihood scores associated with the addition of parameters was highly significant

TABLE 1. Maximum likelihood scores for the large mitochondrial alignment. Values presented are for the most likely of the two MP trees obtained using equally weighted parsimony and for the MP tree obtained using transversion parsimony. The underlined values represent the most likely topology.

Model	Equally weighted parsimony	Transversion parsimony
Jukes-Cantor	−130428.60	−130447.20
F81	−127753.29	−127784.98
HKY85	−125764.08	−125727.74
GTR	−122904.09	−122865.26
GTR + Γ	−112793.31	−112758.78
GTR + Γ + inv	−112657.31	−112622.20
Transversion likelihood		
CF	−57183.63	−56968.97
CFu	−56841.77	−56671.66
CFu + Γ	−50644.47	−50598.48
CFu + Γ + inv	−50608.16	−50562.69

when the likelihood ratio test (Huelsenbeck and Crandall, 1997) was used, and the most parameter-rich models examined (GTR + Γ + inv and CFu + Γ + inv) provided best fits to the data.

Using the best fitting models, the SH test was not significant when all data were analyzed ($P = 0.069$) but strongly rejected the equally weighted parsimony topology when transversion data were analyzed ($P = 0.001$). The SH test can be conservative (prone to type II error) even when only two topologies are compared (Buckley, 2002). Given the limited power of the SH test, our ability to reject the equally weighted parsimony topology using transversion likelihood is likely to be biologically meaningful.

The best fitting models identified were used in heuristic searches under the ML criterion. The most likely topologies identified using either model (GTR + Γ + inv for all data or CFu + Γ + inv for transversion data) of sequence evolution were quite similar to the topology identified using transversion parsimony (Fig. 1C), differing only by rearrangements within the ratites (data not shown). High posterior probabilities were observed for the basal branches in a Bayesian analysis using the GTR + Γ + inv model of evolution (Fig. 1C).

Seven of the parsimony weighting schemes placed passerines basal, but only transversion parsimony placed passerines in a derived position; the conflict between the mitochondrial and nuclear genomes of birds is likely to reflect differences in the signal present in transitions and transversions. These results demonstrate that there is signal in mitochondrial genomes that is consistent with traditional avian classifications (Fig. 1A). However, these analyses did not examine the impact of taxon sampling or sequence length upon the estimation of basal avian relationships. To examine these issues, we restricted further consideration to parsimony analyses using the simplest weighting scheme (equally weighted parsimony) and transversion parsimony and to ML analyses using the best fitting models identified above.

Impact of Taxon Sampling

We performed single-taxon deletions and examined the data using equally weighted parsimony, transversion parsimony, and ML using the best fitting models. Equally weighted parsimony produced fewer expected nodes (Fig. 1A) and more unexpected nodes than did transversion parsimony (Table 2). However, both types of analyses were sensitive to taxon sampling. In several cases, removing one member of a lineage with two taxa (e.g., removal of *Crypturellus*,

TABLE 2. Results of single-taxon deletions for parsimony analyses. Numbers are the percentage of 500 bootstrap replicates supporting a particular relationship; a dash indicates that the node was not present in the bootstrap consensus tree. The first number is from analysis of transversions only, and the second number is from analysis of all data.

	Expected relationships			Unexpected relationships	
Taxon deleted	Paleognath monophyly	Neognath monophyly	Passerine monophyly	Oscines basal	Ratites paraphyletic
None	83/98	72/—	89/—	—/65	—/—
Crypturellus	—/74	82/—	90/—	—/—	—/—
Eudromia	—/—	88/51	94/—	—/—	—/—
Apteryx	87/98	80/—	94/—	—/62	—/—
Casuarius	87/97	73/—	92/—	—/63	—/—
Struthio	86/96	77/—	92/—	—/58	—/—
Dinornis	81/99	76/—	89/—	—/62	—/—
Rhea	75/90	70/—	89/—	—/52	—/—
Aythya	94/100	—/—	91/55	—/—	—/—
Gallus	99/97	—/—	84/—	—/70	—/—
Ciconia	96/99	62/—	73/—	—/72	—/—
Falco	78/98	72/—	91/—	—/—	—/—
Smithornis	95/100	50/—	100/100[a]	—/94	—/—
Corvus	92/99	—/—	—/—	65/93	—/—
Vidua	86/95	70/—	81/—	—/68	—/—

[a] Indicates oscine monophyly only.

TABLE 3. Results of single-taxon deletions for ML and Bayesian analyses. For ML analyses (transversion likelihood on the left and GTR + Γ + inv in the middle), a "Y" indicates the relevant node was found and a dash indicates the node was not present. For Bayesian analyses, the posterior probability of the relevant bipartition is reported on the right.

	Expected relationships			Unexpected relationships	
Taxon deleted	Paleognath monophyly	Neognath monophyly	Passerine monophyly	Oscines basal	Ratites paraphyletic
None	Y/Y/1.0	Y/Y/1.0	Y/Y/1.0	—/—/—	Y/—/—
Crypturellus	Y/Y/1.0	Y/Y/0.95	Y/Y/0.95	—/—/—	Y/Y/0.97
Eudromia	Y/Y/1.0	Y/Y/0.97	Y/Y/0.97	—/—/—	Y/—/—
Apteryx	Y/Y/1.0	Y/Y/1.0	Y/Y/1.0	—/—/—	Y/Y/1.0
Casuarius	Y/Y/1.0	Y/Y/0.99	Y/Y/0.99	—/—/—	Y/—/—
Struthio	Y/Y/1.0	Y/Y/1.0	Y/Y/1.0	—/—/—	Y/—/—
Dinornis	Y/Y/1.0	Y/Y/1.0	Y/Y/1.0	—/—/—	Y/—/—
Rhea	Y/Y/1.0	Y/Y/0.97	Y/Y/0.98	—/—/—	Y/—/—
Aythya	Y/Y/1.0	Y/Y/0.99	Y/Y/1.0	—/—/—	Y/—/—
Gallus	Y/Y/1.0	Y/Y/1.0	Y/Y/1.0	—/—/—	Y/—/—
Ciconia	Y/Y/1.0	Y/Y/0.67	Y/Y/0.78	—/—/—	Y/—/—
Falco	Y/Y/1.0	Y/Y/1.0	Y/Y/1.0	—/—/—	Y/—/—
Smithornis	Y/Y/1.0	Y/—/—	Y/Y/1.0[a]	—/Y/0.96	Y/—/—
Corvus	Y/Y/1.0	Y/Y/0.72	Y/Y/0.72	—/—/—	Y/Y/—
Vidua	Y/Y/1.0	Y/Y/1.0	Y/Y/1.0	—/—/—	Y/—/—

[a] Indicates oscine monophyly only.

Aythya, or *Corvus*) was disruptive (Table 2). These results are consistent with the well-established notion that addition of taxa that bisect long branches improves phylogenetic inference (e.g., Hillis, 1998; Poe and Swofford, 1999).

In these analyses, ML showed greater stability with different taxon compositions, although ML analyses were sensitive to removal of specific taxa (Table 3). ML analyses of all data using the GTR + Γ + inv model produced a paraphyletic ratite group when certain taxa were removed (Table 3), although most other clades remained intact. Bayesian analyses using the GTR + Γ + inv model of sequence evolution indicated the posterior probabilities of the relevant groups were high (Table 3). ML estimates of phylogeny using transversion data were also stable, although ratite paraphyly was observed regardless of the taxa included (Table 3).

To determine whether specific outgroup taxa were problematic, we conducted single-taxon deletions for each outgroup sequence followed by deletion of both crocodilians and both turtles. Single-taxon outgroup deletions had little impact on topology or level of support for any analysis (data not shown). The inferred position of the avian root did shift in transversion parsimony analyses to *Aythya* when both crocodilians were excluded and to the Passeriformes when both turtles were excluded. The inferred position of the root was unaltered in all outgroup deletions when ML was used (GTR + Γ + inv and CFu + Γ + inv), and posterior probabilities of the critical branches were >0.95 in most cases (the single exception was support for neognath monophyly [$P = 0.6$] when both turtles were excluded).

To extend these results, we examined specific alignments of fewer taxa. Equally weighted parsimony analyses of these smaller data sets consistently resulted in a topology that is likely to be incorrect. Transversion parsimony was able in some instances to recover phylogenies that are likely to be correct. For example, use of the seven avian taxa studied by Mindell et al. (1997, 1999) resulted in a topology in which paleognaths are derived and neognaths are paraphyletic (Fig. 2A). However, substituting *Corvus* for *Vidua* in transversion parsimony analysis resulted in the recovery of a topology likely to reflect true avian relationships (Fig. 2B). Analysis of even fewer taxa (e.g., four avian taxa) could result in an expected topology (Fig. 2C), although these results are highly dependent upon the specific taxa included. As predicted by Poe and Swofford (1999), it was possible to find examples of taxon addition that reduced the accuracy of phylogenetic analyses. For example, addition of *Rhea* to the taxa analyzed

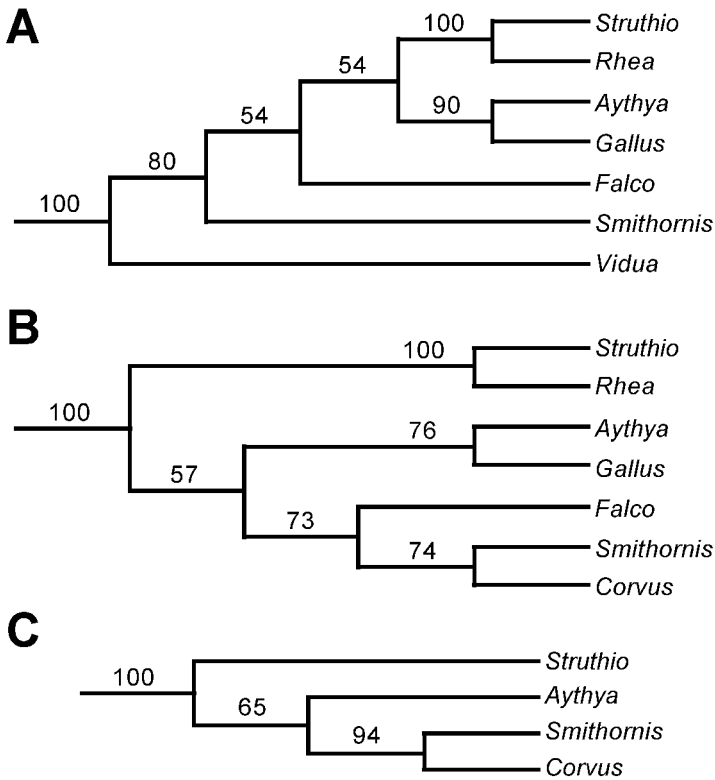

FIGURE 2. The impact of taxon sampling on transversion parsimony. Differences were observed when *Vidua* (A) or *Corvus* (B) were included. Inclusion of as few as four avian taxa (C) could support neognath and passerine monophyly. For all trees, only the ingroup taxa are presented, and the same four outgroup taxa were used in all analyses. Bootstrap support, estimated using 500 replicates, is indicated above the relevant branches.

by transversion parsimony (Fig. 2C) caused *Aythya* to shift to a basal position. Although these results depend upon the inclusion of specific taxa, they demonstrate that signal in avian mitochondria is consistent with traditional classifications, and this information can be recovered without the inclusion of large numbers of avian mitochondrial sequences, contrary to the assumption of Johnson (2001).

Our results have demonstrated that mitochondrial data have signal that supports neognath and passerine monophyly. Larger taxon samples generally performed better, so it may seem surprising that the set of more than 900 cytochrome *b* sequences analyzed by Johnson (2001) generally placed passerines in a more basal position and could not recover a monophyletic Neognathae. One potential explanation for these discrepancies is Johnson's use of sequences of unequal lengths or unbalanced taxon composition. However, in our analyses of 130 cytochrome *b* sequences, all at least 1 kb in length and balanced for taxon composition, we were unable to recover a monophyletic neognath group using any of the eight parsimony weighting schemes. Overall, our 130-taxon cytochrome *b* alignment provided little bootstrap support (< 50%) for any basal avian relationship (data not shown). This finding suggests that cytochrome *b* genes are not suitable for reconstructing deeper nodes in avian phylogeny, either because of the limited number of sites or because of specific aspects of cytochrome *b* evolution.

Interaction of Sequence Length and Taxon Sampling

To determine whether the inability of the 130-taxon cytochrome *b* data set to support expected nodes was specific to cytochrome *b* or simply reflected the paucity of sites in cytochrome *b* genes relative to the larger alignment (a maximum of 1,143 sites versus 10,860 sites), we used jackknife deletion analysis to examine the performance of phylogenetic

TABLE 4. Percentage of 1,000 jackknife data sets supporting monophyly of selected groups.

		Transversion parsimony		Equally weighted parsimony	
	No. base pairs	18-Taxon data set	11-Taxon data set	18-Taxon data set	11-Taxon data set
Paleognath monophyly	1,000	73.4	94.7	75.4	95.3
	2,000	83.2	99.0	88.8	100.0
	4,000	82.0	100.0	94.3	100.0
	8,000	91.1	100.0	98.3	100.0
Neognath monophyly	1,000	34.2	23.6	14.7	5.2
	2,000	39.9	30.2	12.5	<5
	4,000	62.2	45.8	8.2	<5
	8,000	92.8	76.0	<5	<5
Passerine monophyly	1,000	39.4	40.7	21.7	24.5
	2,000	57.3	54.4	23.0	23.6
	4,000	82.5	65.5	15.5	16.2
	8,000	98.3	80.1	<5	<5

analyses using randomly selected data sets of various lengths (1,000, 2,000, 4,000, and 8,000 base pairs). We examined both the complete large alignment (18 taxa) and a smaller alignment containing 7 avian taxa plus 4 outgroups (see Fig. 2B for taxa).

Using transversion parsimony, the addition of sites increased the proportion of replicate data sets that supported paleognath, neognath, and passerine monophyly (Table 4). In contrast, when using equally weighted parsimony, the addition of sites decreased the proportion of replicates that recovered a monophyletic neognath or passerine group (Table 4). These observations were true for both alignments, although the 18-taxon alignment generally performed better than did the 11-taxon alignment.

Deviation from Stationarity and Phylogenetic Estimation

An often-untested aspect of molecular evolution for phylogenetic studies is the possibility that nucleotide composition has changed during evolution. The incorrect inference of some animal relationships, including those among birds, using mitochondrial sequences has been suggested to reflect compositional convergence (Foster and Hickey, 1999; Haddrath and Baker, 2001). Although Conant and Lewis (2001) suggested that the impact of changes in nucleotide composition may not be as problematic as previously believed, we examined the impact of deviations from stationarity upon the results of our analyses.

We found evidence for strong deviation from stationarity for the sequences examined in this study; many values of Gillespie's (1986) δ_{bf} statistic were substantially >1 (Fig. 3A). The value for δ_{bf} seldom exceeded 1 when data were simulated using a complex but stationary model (GTR + Γ + inv; Braun and Kimball, pers. obs.), suggesting values substantially >1 are unlikely to reflect bias in the estimation of δ_{bf}. Haddrath and Baker (2001) also found evidence for deviation from stationarity using a subset of avian and reptilian mitochondrial DNA sequences, although they only examined differences in the proportion of strong (G + C) and weak (A + T) nucleotides. In principle, changes in the proportions of strong and weak nucleotides might not affect proportions of purine and pyrimidine nucleotides. However, estimates of δ_{tv} also indicated substantial deviation from stationarity for transversion data (Fig. 3B).

To determine whether the observed deviations from stationarity affected inference of basal avian relationships using mitochondrial sequences, we examined a least squares tree of δ_{bf} values (Fig. 3C). This tree showed limited clustering of related organisms, emphasizing the fact that archosaurian mitochondrial sequence composition per se has limited historical signal. More complex sequence composition data, such as the DNA strings analyzed by Edwards et al. (2002), appear to contain stronger historical signal given the high level of bootstrap support (100%) for neognath monophyly in that study. In contrast, few groups were supported by bootstrap analysis in the δ_{bf} tree, and the optimal tree does not contain any of the unexpected groups present in topologies identified by various parsimony weighting

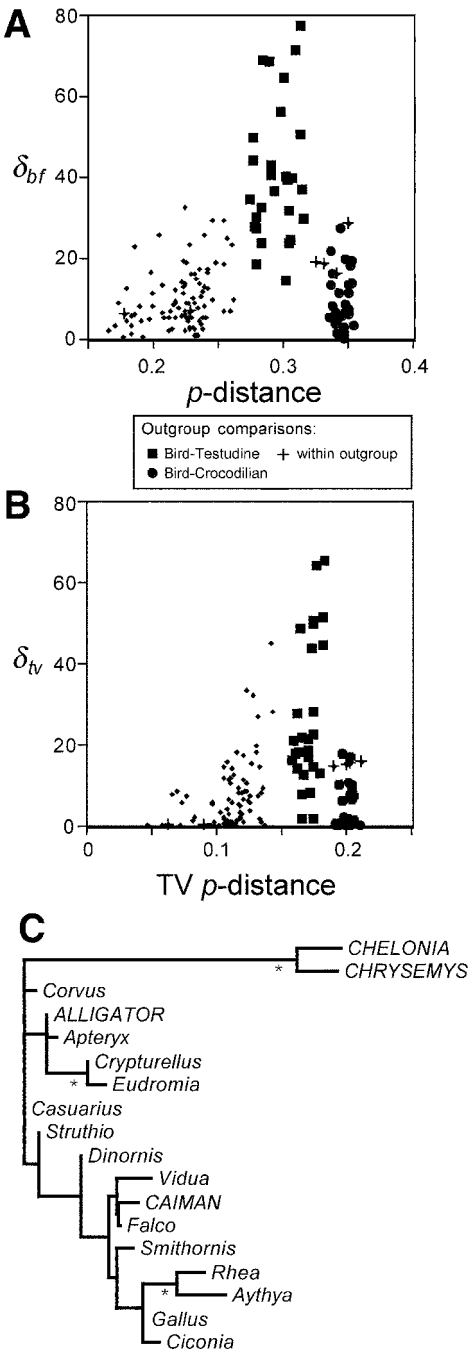

FIGURE 3. Estimates of the δ_{bf} (A) and δ_{tv} (B) statistic, which should not exceed 1 when the sequences have evolved under a stationary model, were plotted against p distances. Clustering of δ_{bf} values by least-squares (C) is inconsistent with both organismal phylogeny and the incorrect estimates of mitochondrial phylogeny obtained using equally weighted parsimony. Branches in the least-squares tree that are marked with asterisks were supported by bootstrap values >70%, and outgroup taxa are presented in capital letters.

schemes (e.g., Fig. 1B), suggesting that convergence in overall base composition is not driving the observed relationships.

Passerines were the basal group of birds in a least-squares tree of LogDet distances (data not shown), even when a proportion of sites (25.9%) were considered invariant to accommodate among-sites rate variation. In sharp contrast, least-squares trees of LogDet distances for transversion data provided strong support for neognath monophyly, with similar levels of bootstrap support (>75%) whether distances were calculated assuming all sites were variable or assuming that a subset of sites (32.6%) were invariant. The least-squares tree of transversion LogDet distances was identical to the transversion parsimony tree (Fig. 1C) except for a rearrangement of derived ratites (data not shown), suggesting basal avian relationships inferred using transversion data are largely unaffected by deviations from stationarity.

CONCLUSIONS

This study indicates that phylogenetic analyses of mitochondrial coding data are capable of supporting a phylogeny consistent with traditional classification and the likely history of the nuclear genome. However, conclusions drawn from complete mitochondria are sensitive to the type of analysis conducted, the number of sites analyzed, and taxon composition. Despite these issues, transversion parsimony and ML using standard parameter-rich models generally recover a phylogeny likely to reflect the organismal phylogeny (Tables 2 and 3). Bootstrap analyses using transversion parsimony and Bayesian analyses using a complex model (GTR + Γ + inv) provided relatively high levels of support (>70% bootstrap support and posterior probabilities > 0.95) for many clades defined by deep branches (Fig. 1C).

The results of these analyses were consistent with the generally accepted contention that adding taxa increases the likelihood of recovering the correct topology in phylogetetic analyses. However, they also indicate that the addition of taxa does not represent a panacea and suggest that the limited set of taxa used here is likely to be sufficient for exploring questions regarding basal avian phylogeny. The extremely broad taxon sampling of Johnson (2001) is clearly unnecessary, and

analyses of basal avian relationships using mitochondrial sequences are likely to require longer sequences than those of cytochrome *b* alone. Sequences of complete avian mitochondria can provide insights into basal avian relationships as long as appropriate analyses are used.

ACKNOWLEDGMENTS

This work was initiated while E.L.B. was supported by a U.S. Department of Agriculture fellowship (1999-01582) and R.T.K. was supported by an Ohio State University postdoctoral fellowship. Karl Kjer, David Mindell, Chris Simon, and an anonymous reviewer provided comments that greatly improved this manuscript. Thomas Buckley, Scott Edwards, and Tara Paton generously permitted us to cite manuscripts in press.

REFERENCES

BENABIB, M., K. M. KJER, and J. W. SITES. 1997. Mitochondrial DNA sequence-based phylogeny and the evolution of viviparity in the *Sceloporus scalaris* group (Reptilia, Squamata). Evolution 51:1262–1275.

BRAUN, E. L., and R. T. KIMBALL. 2001. Polytomies, the power of phylogenetic inference, and the stochastic nature of molecular evolution: A comment on Walsh et al. (1999). Evolution 55:1261–1263.

BUCKLEY, T. R. 2002. Model misspecification and probabilistic tests of topology: Evidence from empirical data sets. Syst. Biol. 51:509–523.

CASPERS, G.-J., D. U. DE WEERD, J. WATTEL, and W. W. DE JONG. 1997. α-Crystallin sequences support a galliform/anseriform clade. Mol. Phylogenet. Evol. 7:185–188.

CONANT, G. C., and P. O. LEWIS. 2001. Effects of nucleotide composition bias on the success of the parsimony criterion in phylogenetic inference. Mol. Biol. Evol. 18:1024–1033.

CUNNINGHAM, C. W. 1997. Is congruence between data partitions a reliable predictor of phylogenetic accuracy? Empirically testing an iterative procedure for choosing among phylogenetic methods. Syst. Biol. 46:464–478.

EDWARDS, S. V., B. FERTIL, A. GIRON, and P. J. DESCHAVANNE. 2002. A genomic schism in birds revealed by phylogenetic analysis of DNA strings. Syst. Biol. 51:599–613.

ELLEGREN, H., and A. CARMICHAEL. 2001. Multiple and independent cessation of recombination between avian sex chromosomes. Genetics 158:325–331.

FEDUCCIA, A. 1995. Explosive evolution in tertiary birds and mammals. Science 267:637–638.

FELSENSTEIN, J. 1981. Evolutionary trees from DNA sequences: A maximum likelihood approach. J. Mol. Evol. 17:368–376.

FITCH, W. M., and E. MARGOLIASH. 1967. Construction of phylogenetic trees. Science 155:279–284.

FLORES-VILLELA, O., K. M. KJER, M. BENABIB, and J. W. SITES. 2000. Multiple data sets, congruence, and hypothesis testing for the phylogeny of basal groups of the lizard genus *Sceloporus* (Squamata, Phrynosomatidae). Syst. Biol. 49:713–739.

FOSTER, T. G., and D. A. HICKEY. 1999. Compositional bias may affect both DNA-based and protein-based phylogenetic reconstructions. J. Mol. Evol. 48:284–290.

GARCIA-MORENO, J., and D. P. MINDELL. 2000. Rooting a phylogeny with homologous genes on opposite sex chromosomes (gametologs): A case study using avian CHD. Mol. Biol. Evol. 17:1826–1832.

GILLESPIE, J. H. 1986. Variability of evolutionary rates of DNA. Genetics 113:1077–1091.

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.

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.

HARING, E., L. KRUCKENHAUSER , A. GAMAUF, M. J. RIESING, and W. PINSKER. 2001. The complete sequence of the mitochondrial genome of *Buteo buteos* (Aves, Accipitridae) indicates an early split in the phylogeny of raptors. Mol. Biol. Evol. 18:1892–1904.

HÄRLID, A., and A. ARNASON. 1999. Analyses of mitochondrial DNA nest ratite birds within the Neognathae: Supporting a neotenous origin of ratite morphological characters. Proc. R. Soc. Lond. B 266:305–309.

HASEGAWA, M., H. KISHINO, AND T. YANO. 1985. Dating the human–ape spliting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22:160–174.

HILLIS, D. M. 1998. Taxonomic sampling, phylogenetic accuracy, and investigator bias. Syst. Biol. 47:3–8.

HUELSENBECK, J. P., and K. A. CRANDALL. 1997. Phylogeny estimation and hypothesis testing using maximum likelihood. Annu. Rev. Ecol. Syst. 28:437–466.

HUELSENBECK, J. P., and F. RONQUIST. 2001. MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755.

JOHNSON, K. P. 2001. Taxon sampling and the phylogenetic position of Passeriformes: Evidence from 916 avian cytochrome *b* sequences. Syst. Biol. 50:128–136.

JUKES, T. H., AND C. R. CANTOR. 1969. Evolution of protein molecules. Pages 21–132 *in* Mammalian protein metabolism (H. N. Munro, ed.). Academic Press, New York.

LAKE, J. A. 1994. Reconstructing evolutionary trees from DNA and protein sequences—Paralinear distances. Proc. Natl. Acad. Sci. USA 91:1455–1459.

LOCKHART, P. J., M. A. STEEL, M. D. HENDY, and D. PENNY. 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. Mol. Biol. Evol. 11:605–612.

MADDISON, W. P. 1997. Gene trees in species trees. Syst. Biol. 46:523–536.

MAYR, E., and D. AMADON. 1951. A classification of recent birds. Am. Mus. Novit. 1496:1–42.

MINDELL, D. P., M. D. SORENSON, D. E. DIMCHEFF, M. HASEGAWA, J. C. AST, and T. YURI. 1999. Interordinal relationships of birds and other reptiles based on whole mitochondrial genomes. Syst. Biol. 48:138–152.

MINDELL, D. P., M. D. SORENSON, C. J. HUDDLESTON, H. C. J. MIRANDA, A. KNIGHT, S. J. SAWCHUK, and T. YURI. 1997. Phylogenetic relationships among and within select avian orders based on mitochondrial DNA. Pages 214–247 *in* Avian molecular evolution and systematics (D. P. Mindell, ed.). Academic Press, San Diego.

MOORE, W. S. 1995. Inferring phylogenies from mtDNA variation: Mitochondrial-gene trees versus nuclear-gene trees. Evolution 49:718–726.
NAYLOR, G. J. P., T. M. COLLINS, and W. M. BROWN. 1995. Hydrophobicity and phylogeny. Nature 373:565–566.
PATON, T., O. HADDRATH, and A. J. BAKER. 2002. Complete mitochondrial DNA genome sequences show that modern birds are not descended from transitional shorebirds. Proc. R. Soc. Lond. B 269:839–846.
POE, S., and D. L. SWOFFORD. 1999. Taxon sampling revisited. Nature 398:299–300.
PRAGER, E. M., and A. C. WILSON. 1976. Congruencey of phylogenies derived from different proteins. J. Mol. Evol. 9:45–57.
PRAGER, E. M., A. C. WILSON, D. T. OSUGA, and R. E. FEENEY. 1976. Evolution of flightless land birds on southern continents: Transferrin comparison shows monophyletic origin of ratites. J. Mol. Evol. 8:283–294.
SHIMODAIRA, H., and M. HASEGAWA. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16:1114–1116.
SIBLEY, C. G., and J. E. AHLQUIST. 1990. Phylogeny and classification of birds: A study in molecular evolution. Yale Univ. Press, New Haven, Connecticut.
STAPEL, S. O., J. A. M. LEUNISSEN, M. VERSTEEG, J. WATTEL, and W. W. DE JONG. 1984. Ratites as oldest offshoot of avian stem—Evidence from α-crystallin A sequences. Nature 311:257–259.
SWOFFORD, D. L. 1999. PAUP*: Phylogenetic analysis using parsimony (*and other methods), version 4.0. Sinauer, Sunderland, Massachusetts.
SWOFFORD, D. L., G. J. OLSEN, P. J. WADDELL, and D. M. HILLIS. 1996. Phylogenetic inference. Pages 407–514 *in* Molecular systematics, 2nd edition (D. M. Hillis, C. Moritz, and B. K. Mable, eds.). Sinauer, Sunderland, Massachusetts.
VAN TUINEN, M., C. G. SIBLEY, and S. B. HEDGES. 2000. The early history of modern birds inferred from DNA sequences of nuclear and mitochondrial ribosomal genes. Mol. Biol. Evol. 17:451–457.
VOELKER, G., and S. V. EDWARDS. 1998. Can weighting improve bushy trees? Models of cytochrome *b* evolution and the molecular systematics of pipits and wagtails (Aves: Motacillidae). Syst. Biol. 47:589–603.
WETMORE, A. 1960. A classification for the birds of the world. Smithson. Misc. Collect. 139:1–37.
YANG, Z. 1994. Estimating the pattern of nucleotide substitution. J. Mol. Evol. 39:105–111.

First submitted 28 November 2001; reviews returned 13 March 2002; final acceptance 15 April 2002
Associate Editor: Karl Kjer

Syst. Biol. 51(4):625–637, 2002
DOI: 10.1080/10635150290102302

The Utility of the Incongruence Length Difference Test

F. KEITH BARKER[1] AND FRANÇOIS M. LUTZONI[2]

[1]*Department of Ornithology, American Museum of Natural History, Central Park West at 79th Street, New York, New York 10024, USA; E-mail: fbarker@amnh.org*
[2]*Department of Biology, Duke University, Box 90338, Durham, North Carolina 27708, USA*

Conditional combination of phylogenetic data requires definition of explicit criteria for combinability (Bull et al., 1993). In this context, combinability refers to the methodological validity of combining multiple sources of phylogenetic data, given the underlying assumptions (explicit or otherwise) of the analysis. Combinability has been evaluated by the effect of data set combination on phylogenetic accuracy: Combinable data sets increase accuracy (Bull et al., 1993; Cunningham, 1997b). When inferential methods are statistically consistent, this convergent property is guaranteed by statistical homogeneity of the data sets to be combined: Increasing sample size increases precision. In a phylogenetic context, data homogeneity can be defined as the sharing of a single history (topological pattern of ancestor–descendant relationships among terminals) and uniform probabilities of change among character states (e.g., branch lengths and relative frequencies of character state transformation). Data sets sampling the same phylogenetic history, but with drastically different evolutionary dynamics, could yield biased estimates when combined and analyzed using a model and parameters with a poor fit to at least one of the partitions. For molecular data, these requirements are explicit in the calculation of conditional probabilities based on the maximum-likelihood criterion, where the overall likelihood is the product of individual site likelihoods, under the assumption that site patterns are independent and identically distributed (Felsenstein, 1981). However, likelihood methods allow this