Mosaicism, Modules, and the Evolution of Birds: Results from a Bayesian Approach to the Study of Morphological Evolution Using Discrete Character Data

JULIA A. CLARKE[1] AND KEVIN M. MIDDLETON[2]

[1]*Department of Marine, Earth and Atmospheric Sciences, North Carolina State University, Campus Box 8208, Raleigh, NC 27695-8208, USA;*
E-mail: Julia_Clarke@ncsu.edu
[2]*Department of Biology, California State University, San Bernardino, 5500 University Parkway, San Bernardino, CA 92407, USA;*
E-mail: kmm@csusb.edu

Abstract.—The study of morphological evolution after the inferred origin of active flight homologous with that in Aves has historically been characterized by an emphasis on anatomically disjunct, mosaic patterns of change. Relatively few prior studies have used discrete morphological character data in a phylogenetic context to quantitatively investigate morphological evolution or mosaic evolution in particular. One such previously employed method, which used summed unambiguously optimized synapomorphies, has been the basis for proposing disassociated and sequential "modernizing" or "fine-tuning" of pectoral and then pelvic locomotor systems after the origin of flight ("pectoral early–pelvic late" hypothesis). We use one of the most inclusive phylogenetic data sets of basal birds to investigate properties of this method and to consider the application of a Bayesian phylogenetic approach. Bayes factor and statistical comparisons of branch length estimates were used to evaluate support for a mosaic pattern of character change and the specific pectoral early–pelvic late hypothesis. Partitions were defined a priori based on anatomical subregion (e.g., pelvic, pectoral) and were based on those hypothesized using the summed synapomorphy approach. We compare 80 models all implementing the M_k model for morphological data but varying in the number of anatomical subregion partitions, the models for among-partition rate variation and among-character rate variation, as well as the branch length prior. Statistical analysis reveals that partitioning data by anatomical subregion, independently estimating branch lengths for partitioned data, and use of shared or per partition gamma-shaped among-character rate distribution significantly increases estimated model likelihoods. Simulation studies reveal that partitioned models where characters are randomly assigned perform significantly worse than both the observed model and the single-partition equal-rate model, suggesting that only partitioning by anatomical subregion increases model performance. The preference for models with partitions defined a priori by anatomical subregion is consistent with a disjunctive pattern of character change for the data set investigated and may have implications for parameterization of Bayesian analyses of morphological data more generally. Statistical tests of differences in estimated branch lengths from the pectoral and pelvic partitions do not support the specific pectoral early–pelvic late hypothesis proposed from the summed synapomorphy approach; however, results suggest limited support for some pectoral branch lengths being significantly longer only early at/after the origin of flight. [Bayesian analyses; birds; flight; fossils; morphological data; mosaic evolution; quantitative methods.]

Few behaviors are as broadly considered to be key innovations in the context of evolutionary biology as avian flight. However, in contrast to the extensive literature on the origin of avian flight (evolution of flight in derived maniraptoran theropod dinosaurs), the study of the evolution of flight after its origin but before the radiation of extant birds (crown clade Aves, following Gauthier and de Queiroz, 2001) continues to be comparatively little explored. Relevant studies have focused on the evolution of single characters or complexes (e.g., Gatesy and Middleton, 1997; Poore et al., 1997; Ostrom et al., 1999; Hutchinson, 2001, 2002) or have proposed ecological or aerodynamic scenarios (e.g., Chatterjee, 1997; Garner et al., 1999; Padian, 2001), a disciplinary pattern earlier noted by Gatesy and Dial (1996a). Here we investigate a distinct and recurring trend in the study of morphological evolution in birds, with emphasis on proposed mosaic evolution or anatomically localized change.

De Beer (1954) coined the term and proposed a concept of mosaic evolution with reference to the urvogel *Archaeopteryx*, writing, "if the transition from reptiles to birds was characterized by what I propose to call *mosaic evolution* it becomes of interest whether the same mode of evolution was followed in the transition between other classes of vertebrates" (1954:163, emphasis in original). The presence of primitive and derived characters in distinct anatomical regions of the skeleton was considered indicative of a new mode of organic evolution (de Beer, 1954). The North American "toothed birds" (*Hesperornis* and *Ichthyornis*) with the combination of teeth and a recognizably "avian" postcranium were noted by Darwin as key support for ideas put forth in his 1859 publication (in litt., C. Darwin to O. C. Marsh, August 31, 1880). Cracraft (1970) similarly noted a mosaic pattern of "reptilian" and "avian" characteristics in the mandible of *Archaeopteryx*. In more recent treatments of the assembly of extant avian flight, compartmentalized anatomical change has again been central. Gatesy and Dial (1996a, 1996b) proposed autonomous, but highly integrated, locomotor units, termed modules: the avian pectoral girdle and limbs, the pelvic girdle and limbs, and the tail. These modules had different patterns of evolution and functional linkages across Dinosauria, including birds (Gatesy and Dial, 1996a, 1996b).

As in studies of other groups of vertebrates, these hypotheses of disjunct or compartmentalized change may be contrasted with a focus on gradualistic shifts in shape and form across the skeleton of birds. The few well-known fossil birds from the 19th century were recognized to represent transitional morphologies, although there were acknowledged needs to explain the changes occurring in perceived gaps between these taxa. In *On Growth and Form*, Thompson (1917; reprinted, 1961) illustrated hypothesized gradual transformations

using gridded representations of anatomical regions for two of the best-known Mesozoic birds (figs. 161 to 166 in Thompson, 1917; reprint 1961:306–308). The pelvis of *Archaeopteryx* was transformed into that of an ornithurine "toothed bird" (YPM 1734; Marsh, 1880; Clarke, 2004), the postcranium of which was considered nearly "modern." Ostrom (1976) also emphasized a series of stepwise changes hypothetically bridging the morphologies present among the few known extinct taxa and living birds (figs. 7 and 13 in Ostrom, 1976).

New fossil data and phylogenetic analyses have, for the first time, allowed more than a hypothetical examination of the transition from the first form of bird flight to that present in extant birds. Discrete morphological character data from these analyses were used to study morphological evolution, and patterns of compartmentalized anatomical change were once again emphasized. Unambiguous synapomorphies for two sets of anatomical characters, forelimb and hind limb, were summed for the path consisting of the subset of the cladogram's internal internodes, from the divergence of *Archaeopteryx* to the terminal representing extant birds (Chiappe, 1991, 1995a, 1995b, 1996, 2002). Differential modifications first modernizing, fine-tuning, or enhancing the forelimb (Chiappe, 2002:460) early after the evolution of flight were inferred from a peak in the number of unambiguous synapomorphies from the forelimb set phylogenetically earlier than a similar peak in hind limb synapomorphies (e.g., fig. 5 of Chiappe, 1995a). Based on analyses using this method, here referred to as the summed synapomorphy approach or "SSA," it was hypothesized that structures correlated with enhanced flight capabilities evolved early after the evolution of flight, followed phylogenetically later by those correlated with enhanced terrestrial locomotion. For simplicity, we will refer to this hypothesis as "pectoral early–pelvic late."

Mosaic evolution has been most often discussed as disjunct sequential change in one anatomical subregion relative to another in the evolution of a taxon of interest, consistent with de Beer's (1954) original concept. What is in fact at first glance surprising is the persistence of this notion after the recognition that all organisms represent combinations of plesiomorphic and derived character states. What remains common to contemporary investigations citing mosaic evolution is the proposal of disjunct change in one anatomical subregion versus another rather than among topologically closely adjacent characters. Most studies have scrutinized the optimizations of single or several characters from different subregions for such proposed patterns (e.g., Ji et al., 1999b; Rae, 1999; Cavalier-Smith, 2002; Kearney and Stuart, 2004) or have deemphasized a phylogenetic perspective (e.g., Barton and Harvey, 2000). SSA compares subsets of summed changes (unambiguous synapomorphies) for comparatively large sets of morphological characters from proposed subregions of interest without individual character-by-character analysis, but assuming these changes are proxies for key novelties modernizing or fine-tuning the focal regions (e.g., Chiappe, 1991, 2002). Exemplifying the more common approach, Rae (1999)

analyzed relative first appearances of select characters in a small data set and found that facial characters changed earlier than postcranial characters in hominid evolution. Rae (1999), however, also tested for congruence between two of the character sets (cranial and postcranial) to evaluate degree of conflict in phylogenetic signal among sets, arguably only one possible effect of disjunctive anatomical change.

Herein we use the most-inclusive available data set to investigate proposed disjunctive patterns of avian locomotor evolution after the origin of flight. Taxonomic sampling is focused on basal, or phylogenetically early, divergences within the clade Avialae ($\sim$ flighted theropod dinosaurs; sensu Gauthier and de Queiroz, 2001). Morphological evolution from the recent common ancestor of *Archaeopteryx* and crown clade Aves through the most basal nodes within Aves is considered (Fig. 1). This part of avialan evolution extends from the presently inferred minimum origin of active flight unambiguously optimized as homologous with flight in extant birds through the phylogenetically earliest node for which all of the aspects of musculoskeletal function in extant avian flight can be minimally inferred to be present; i.e., Aves (Gauthier and de Queiroz, 2001). The lineage along which this proposed sequential change may be analyzed under SSA is composed of the internodes or path segments from the basal-most ingroup node (or any internal node) to the terminal taxon of interest. In the case presented here, where Aves is the clade of interest, the path addressed is equivalent to the sum of the internodes from the most recent common ancestor of *Archaeopteryx* and Aves to the sampled crown clade taxa (e.g., Chiappe 1995a; Fig. 1). Herein, we refer to these path segments as "spine" internodes, and, for these internodes, synapomorphies are summed in SSA.

The summed synapomorphies approach is investigated with the new data set to consider what its analytical limitations may be and the significance for proposed patterns of avialan evolution if the "pectoral early–pelvic late" hypothesis were confirmed with a larger character and taxon sample. The potential for quantitative investigation of morphological data using Bayesian methods is also explored (Li, 1996; Mau, 1996; Rannala and Yang, 1996; Huelsenbeck and Ronquist, 2001; Lewis, 2001; Huelsenbeck et al., 2002). A small but increasing number of studies have employed Bayesian phylogenetic methods with morphological data (Lewis, 2001; Glenner et al., 2004; Nylander et al., 2004; Lee, 2005; Wiens et al., 2005; Müller and Reisz, 2006). The Bayesian approach presented here expands on basic phylogenetic analysis and proposes a new method for evaluating hypotheses of differential rates of morphological evolution among anatomically based sets of characters. Although in need of further investigation, appealing properties of the methodology presented here and other parametric approaches include the potential for avoiding some of the confounding effects of missing data (e.g., Wiens, 2005), as well as the potential for statistical comparisons of branch lengths. Bayesian methods allow a variety of questions analogous to those concerning
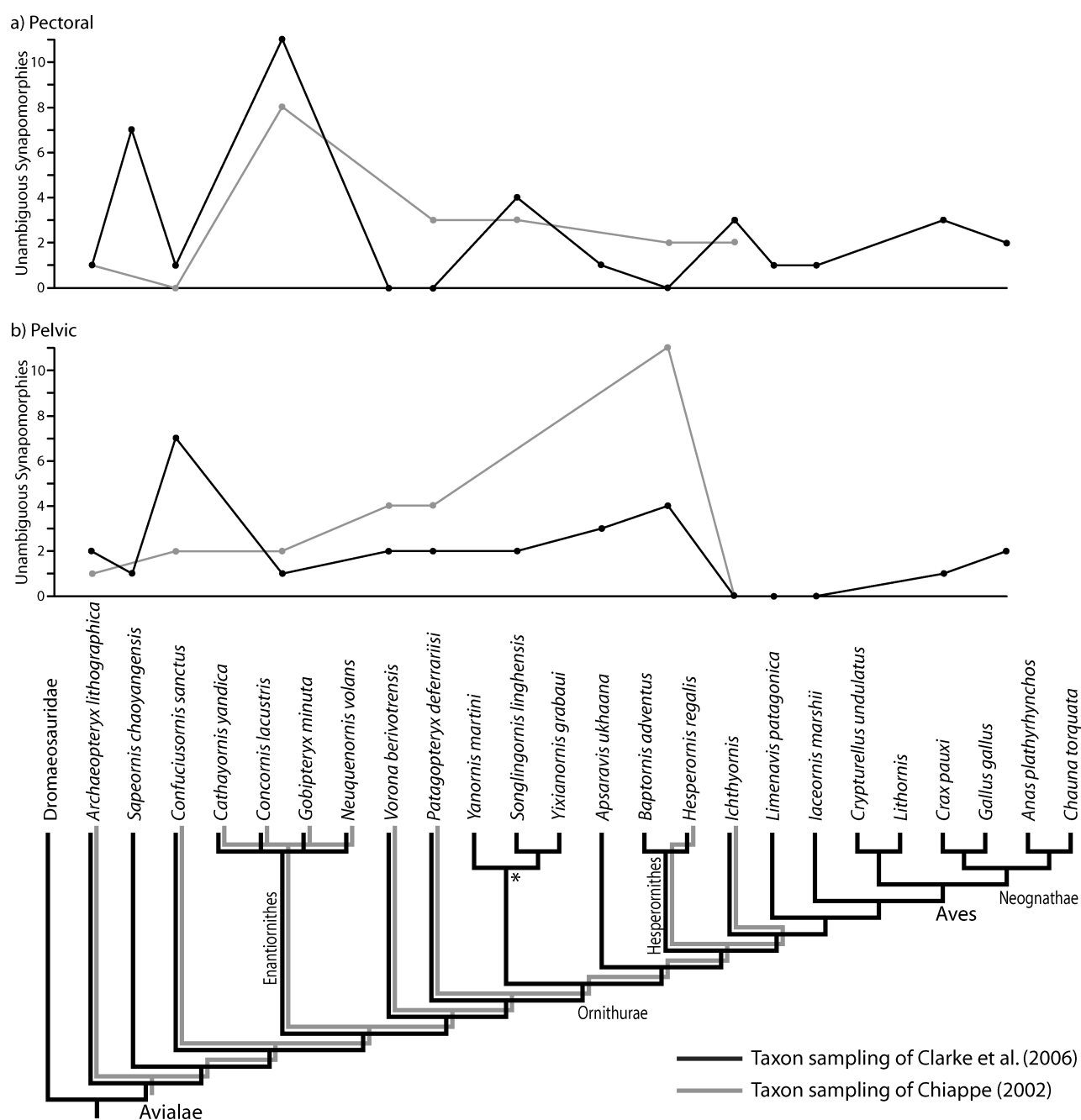
FIGURE 1. Summed unambiguous synapomorphies per internode for (a) pectoral girdle and limb characters and (b) pelvic girdle and limb characters are plotted for the segments along the "spine" of the avialan tree. The spine comprises the path from the most recent common ancestor of *Archaeopteryx* and Aves through basal divergences in Aves. The maximum parsimony tree from Clarke et al. (2006) is shown in black, and the tree approximating the taxon sampling of Chiappe (2002) is shown in gray. In a and b, results from complete taxonomic sampling in that analysis (black lines) and with the taxonomic sample trimmed to that in common with Chiappe (2002) (gray lines) are shown. Under the Chiappe (2002) sampling, a peak in pectoral characters is observed early in avialan evolution, whereas a peak in pelvic changes is observed phylogenetically later. With increased taxonomic sampling, accessory peaks in pectoral character change are observed, and the highest pelvic character peak moves phylogenetically earlier than the peak in pectoral character changes. The clade of Chinese ornithurines is indicated with an asterisk. In Figure 2, this clade is referred to as the "*Yixianornis* clade." Note that for clades, the data point is aligned with their divergence from the "spine." For instance, the data point for Hesperornithes is placed between *Baptornis* and *Hesperornis*, aligned with the node representing the most recent common ancestor of Hesperornithes and Aves.

heterogeneous evolution among genes at the molecular level (e.g., Yokoyama and Yokoyama, 1989; May et al., 1999) to be investigated in the context of morphological evolution.

Summed Unambiguous Synapomorphies

Methods

The character set and taxonomic sample used were those of Clarke et al. (2006). Two hundred and five total characters were evaluated for 23 ingroup and 2 outgroup taxa (38 ordered; 191 parsimony informative characters [PIC]). Parsimony analysis was conducted using PAUP* 4.08b (Swofford, 2002; branch and bound search, amb- [collapsing minimum 0 length branches], polymorphism differentiated from ambiguity). Two most parsimonious trees resulted (Length: 422, CI: 0.63, RI: 0.82, RC: 0:51 [parsimony informative characters only]). The strict consensus cladogram of these two trees is shown in Fig. 1 (reported in Clarke et al., 2006; see original publication for data set, other details of analysis as well as bootstrap and Bremer support values). The recovered relationships are consistent with previous analyses of avialan relationships for the taxa in common among these analyses (Chiappe, 2002; Clarke, 2004; Clarke et al., 2005).

Characters were categorized into the two anatomical subregions proposed to show disjunctive patterns of evolution, pectoral and pelvic (girdles and limbs; Chiappe, 2002). The distribution of characters across the skeleton is similar to that of Chiappe (2002). In the current analysis, pectoral girdle and limb characters (Clarke et al., 2006: characters 72 to 154) represent 40% of the data set total (38% in Chiappe, 2002), and pelvic girdle and limb characters (Clarke et al., 2006: characters 155 to 205) represent 25% of the total (30% in Chiappe, 2002). Cranial characters comprise 24% of the characters in Clarke et al. (2006: characters 1 to 52) and 17% in Chiappe (2002). Unambiguous synapomorphies were identified in MacClade (Version 4.06; Maddison and Maddison, 1997). Only characters for nodes recovered in the strict consensus cladogram unambiguously optimized in each of the two most-parsimonious trees were summed. Cranial characters were also summed to consider the results of applying the method to another anatomical subregion that was well represented in the data set but with a distinctly greater amount of missing data. Unlike the forelimb and hind limb characters, which could not be evaluated (i.e., "missing" and scored "?") for less than about one third of the included taxa (means: 32% and 30%, respectively), cranial characters could not be evaluated for a little over half of these taxa (mean: 53%).

We used three taxonomic sampling regimes. First, using a taxonomic sample approximating that of Chiappe (2002), we assessed whether the proposed pattern of mosaic evolution would be recovered from summed apomorphies with a larger, partially overlapping character set (from Clarke et al., 2006). Second, unambiguous synapomorphies were summed from analysis including the complete taxonomic sample of Clarke et al. (2006). Third, we excluded a clade of Chinese species, to con-

sider the impact of these taxa, which were included by Clarke et al. (2006), on character optimization. These species form a clade in the maximum parsimony analysis (Fig. 1: *Yixianornis grabaui*, *Yanornis martini*, and *Songlingornis linghensis*; Clarke et al., 2006) and fall close to the proposed shift point from forelimb to hind limb dominated change in SSA studies (e.g., Chiappe, 2002). Also striking in *Yixianornis* and *Yanornis* (Zhou and Zhang, 2001) is the simultaneous presence of morphologies of the pectoral girdle and forelimb often considered as touchstones of modern or otherwise fully avian flight, with plesiomorphic pelvic morphologies fitting a nonavialan dinosaurian gestalt (e.g., pubic symphysis and full set of gastralia; Zhou and Zhang, 2001; Zhou, 2004; Clarke et al., 2006). This combination of morphologies and the phylogenetic position of these new taxa (Clarke et al., 2006; Fig. 1) appeared to fit the proposed pectoral early–pelvic late hypothesis of morphological change across Avialae with early modernization of the pectoral girdle and limb prior to the pelvic.

A graphical representation of how taxon sampling differs relative to the sampling in Chiappe (2002) is shown in Fig. 1 (bottom, gray lines). We sample seven additional Mesozoic ornithurine taxa: *Apsaravis* (Norell and Clarke, 2001), *Baptornis* (Marsh, 1877), *Iaceornis* (Clarke, 2004), *Limenavis* (Clarke and Chiappe, 2001), *Songlingornis* (Hou et al., 1996), *Yanornis* (Zhou and Zhang, 2001), *Yixianornis* (Zhou and Zhang, 2001), and we exclude one poorly known ornithurine taxon included in that study, *Ambiortus* (Kurochkin, 1988). Of more basal avialans and nonavialan theropods included, we sample one additional taxon, *Sapeornis* (Zhou and Zhang, 2002), and exclude five taxa that have been indicated as non-avialan theropods (e.g., Chiappe, 2002; Hwang et al., 2002; Novas and Pol, 2005): four alvarezsaurids (Chiappe, 2002; Clark et al., 2002) and *Rahonavis* (Forster at al., 1998; Clark et al., 2002). We also exclude three Enantiornithes that were included in previous analyses: *Eoalulavis* (Sanz et al., 1996), *Iberomesornis* (Sanz and Bonaparte, 1992), and *Noguerornis* (Lacasa-Ruiz, 1989), as well as one part of Confuciusornithidae (*Changchengornis*; Ji et al., 1999a). In our sample, Aves is represented by six exemplars rather than one (*Anas* in Chiappe, 2002). The rationale for exemplar choice in Aves as well as for basal Avialae is detailed in Clarke (2004).

Results

When the summed unambiguous synapomorphy approach is applied to the current data set but using the taxonomic sample approximating that of the most recent analysis used to support the mosaic pattern (e.g., Chiappe, 2002), the two sequential peaks of the pectoral early–pelvic late hypothesis are recovered (Chiappe, 1991, 1995a, 1995b, 1996, 2002). A single peak in pectoral characters is located at the divergence of Enantiornithes (Fig. 1a, gray lines), and a single large peak in pelvic characters is located at the Hesperornithes divergence, phylogenetically later (i.e., closer to crown clade Aves; Fig. 1b, gray lines). Thus, the proposed pattern is recovered

using more characters and the Clarke et al. (2006) data set with a taxon sampling similar to Chiappe (2002).

Using the complete taxonomic sample of Clarke et al. (2006), two peaks in pectoral changes are present early after the origin of flight, still consistent with the hypothesis (Fig. 1a, black lines). However, the single pronounced pelvic peak at Hesperornithes is absent (Fig. 1b, black lines). Rather, two smaller peaks are present; the higher of these is now also recovered early at the *Confuciusornis* node, and a second low peak is present at the Hesperornithes node.

Sensitivity to taxonomic sampling and missing character data was investigated by excluding one clade and by applying the methodology to the cranial characters. The exclusion of the *Yixianornis*, *Yanornis*, and *Songlingornis* clade, although a subtraction of a single spine node ("*Yixianornis* clade" of Fig. 2a, b), alters the perceived distribution of morphological change. When these taxa were excluded, several pectoral changes are perceived as occurring later, as seen in an increase in the number of summed synapomorphies at the divergence of *Apsaravis* (Fig. 2a). Also with exclusion of these taxa a pronounced peak in pelvic change is seen at the *Apsaravis* node (Fig. 2b). This peak is absent when the full Clarke et al. (2006) taxonomic sample is considered. Thus, including these taxa diminishes the "pelvic late" portion of the proposed pectoral early–pelvic late pattern that the Gestalt combination of characters in these species appeared superficially to support (see above; Clarke et. al., 2006).

When SSA is applied to the set of cranial characters that are missing (i.e., unable to be scored) for taxa at higher rates compared to the pectoral and pelvic characters, no unambiguous cranial synapomorphies are observed for spine internodes between *Confuciusornis* and Neognathae (Fig. 2c). If SSA is used, and summed synapomorphies are interpreted as approximating the pattern of morphological evolution, then it would be a logical but absurd conclusion that no cranial evolution occurred for most of basal avialan history. Exclusion of the *Yixianornis* clade does not impact the pattern of perceived change in these characters.

BAYESIAN ANALYSES

Methods

Drawing on the methods of Nylander et al. (2004), we explored how Bayesian phylogenetic methods might be used to quantitatively investigate morphological evolution in birds and to test the proposed mosaic pattern of character evolution. If the number of state changes in the pectoral characters across branches or internodes of the resultant cladograms was distinct from that for the pelvic characters, then analyses allowing branch lengths and, by proxy, evolutionary change, for these character partitions to vary independently would be preferred over simpler models where branch lengths were jointly estimated for the entire data set.

Bayesian analyses were performed in MrBayes v. 3.1.2 (Ronquist and Huelsenbeck, 2003), using the character matrix from Clarke et al. (2006). We employed the M_k
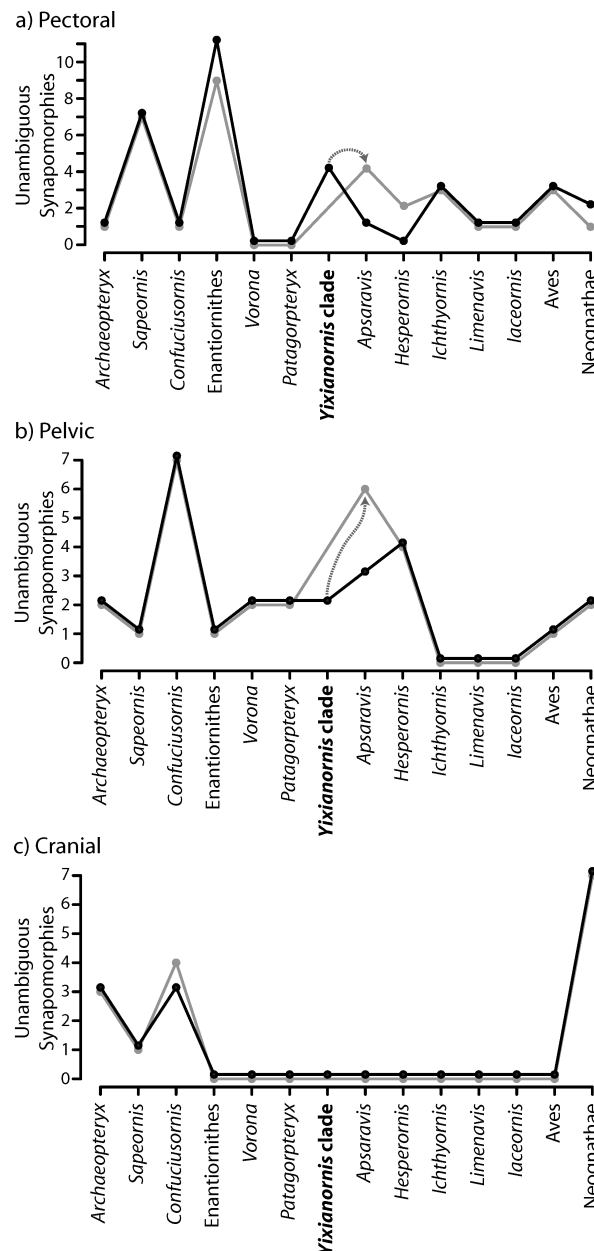

FIGURE 2.   Sensitivity of the summed synapomorphy approach to missing character data and taxonomic sampling. Summed unambiguous synapomorphies per internode in three character partitions are plotted for the segments along the spine of the avialan tree. Results using two different taxonomic samples are represented: the complete Clarke et al. (2006) sample (black lines) and with the clade of Chinese ornithurines, *Songlingornis*, *Yanornis*, and *Yixianornis*, excluded (gray lines). The removal of a single new divergence and the associated internode along the spine alters the pattern of unambiguous synapomorphies. (a) A peak in pectoral synapomorphies associated with the Chinese ornithurine clade is shifted phylogenetically later to *Apsaravis* (dotted arrow) and (b) a relatively high peak in pelvic synapomorphies is created at *Apsaravis* from synapomorphies formerly spread out across multiple nodes (dotted arrow). Finally, for the cranial characters (c), which are missing data at a higher rate across included taxa, no synapomorphies are unambiguously optimized for any spine internodes between *Confuciusornis* and Neognathae. Thus, using SSA, these data would be consistent with the conclusion that no cranial evolution took place for most of basal avialan evolution. Note that overlapping lines have been offset slightly for clarity.

model for morphology (Lewis, 2001), with character type "standard" for discrete morphological characters and coding set to "variable" in MrBayes. Because the character type in MrBayes was set to variable, four operationally constant characters were automatically excluded (characters 46, 131, 132, 172). All analyses were run for 5×10^6 repetitions MCMC (sampled every 100 repetitions; four chains and two simultaneous runs; heated chain temperature of 0.2), which was sufficient for convergence. Convergence was judged by the stabilization of the standard deviation of the split frequencies below 0.01. The initial 25% of samples (12,500 total) were discarded for burn-in. The Bayesian topology and posterior probabilities for recovered clades were compared to results obtained via maximum parsimony (shown in Fig. 1).

Models with between one and four anatomical subregion partitions were evaluated. Single-partition models were no-partition models with all characters jointly analyzed. The two-partition models consisted of pectoral (82 characters) and a partition that included all other characters (pelvic, cranial, axial; 122 total characters). Investigation of this set of models was motivated by the three-partition results (see below). The three-partition models consisted of pectoral (82 characters), pelvic (51 characters), and a partition that included all other characters (cranial and axial; 71 total characters). Finally, the four-partition models consisted of pectoral (82 characters), pelvic (51 characters), cranial (52 characters), and axial (19 characters). The characters comprising these pectoral, pelvic, and cranial partitions used were identical to those defined for SSA.

In multiple-partition analyses, branch lengths were permitted to differ among partitions using two different models of among-partition rate variation. In the terminology of Marshall et al. (2006), these models either (1) utilized a linked rate multiplier (MrBayes command: prset ratepr = variable), which allows branch lengths to vary among partitions under the constraint that the mean of all rate multipliers equals one, such that effective branch lengths may be different but must be proportional across partitions (Nylander et al., 2004), or (2) unlinked all branch lengths (MrBayes command: unlink brlen = (all)), which allows branch lengths for each partition to vary completely independently of one another. The latter model is much more parameterized than the first. Models with a gamma-shaped distribution (Γ) for among-character rate variation ($M_k + \Gamma$; shape parameter uniformly distributed from 0 to 200, approximated with four categories) were also investigated, as they had been recovered to have higher model likelihoods in the analyses of morphological data sets (Wiens et al., 2005; Müller and Reisz, 2006). First, all rates were modeled as drawn from a single gamma ("shared gamma" in Table 1), and in the second case, each partition was assigned its own gamma distribution from which rates were drawn ("per-partition gamma"). Finally, the effect of different branch length priors was investigated following the suggestions of Marshall et al. (2006). Four different branch length priors were tested

in which mean branch length varied from 0.2 to 0.025: exp(5), exp(10), exp(20), and exp(40). All permutations of these investigated variables were tested, for a total of 80 models.

The model likelihood, as estimated using the log of the harmonic mean of sampled likelihoods after discarding 25% burn-in (Newton and Raftery, 1994; Nylander et al. 2004; harmonic mean estimator sensu Lartillot and Philippe, 2006), was recorded for each analysis (Table 1). Bayes factor comparisons were calculated made from these means, using the methods described in Nylander et al. (2004; though see Lartillot and Philippe, 2006). The criteria for model choice are those of Kass and Raftery (1995), although the implications of higher cutoff values for strong model preference were also explored. Because of the large number of models investigated, only pairwise Bayes factor comparisons between the most preferred model (highest model likelihood) and all other models are presented in Table 1. To assess potential variation in estimated model likelihoods, we both (1) selected 10 models with varying levels of parameterization, which were reanalyzed five times each, and (2) analyzed the highest likelihood model for two and four times as many iterations as the original analysis (10 $\times 10^6$ and 20 $\times 10^6$ iterations, respectively). The likelihood for each reanalysis was recorded. The latter analysis was performed because the harmonic mean estimator in higher-dimensionality models has been shown to overestimate the marginal likelihood (Lartillot and Philippe, 2006), and longer iteration lengths have been suggested as one potential approach to more reliable estimates. Thermodynamic integration has been proposed as an alternative methodology (Lartillot and Philippe, 2006), but, unfortunately, this method cannot currently be used with the standard output from Bayesian analyses as implemented in MrBayes.

Based on the results of these analyses, we used randomizations to test whether models estimating branch lengths independently (unlinked) for multiple partitions may be favored over a single-partition model simply because they are more complex and parameter rich. Specifically, we compared the three-partition models with anatomical subregion–based partitions to models that were otherwise identical but with three partitions of randomly selected characters. For each of 1000 iterations, characters from the total data set were randomly sampled without replacement such that the sizes of the three partitions (number of characters in each) were equal to those used in the observed data three-partition analysis (82, 51, and 71 characters). Data set assembly and analyses identical to those used with the original, observed model (three-partition, independent among-partition rate variation, equal rates, branch length prior exp [10]) were carried out using MrBayes v. 3.1.2 (Ronquist and Huelsenbeck, 2003) coupled to R 2.5.1 (R Development Core Team, 2007), with custom shell scripts (available from the authors). As in the main analyses (Table 1), randomized-partition models were run for 5 $\times 10^6$ iterations, and chains were sampled every 100 iterations with the first 25% of samples discarded for burn-in.

TABLE 1. Summary of 80 models tested (nonrandomized data) using the Clarke et al. (2006) data set. Models are sorted in order of decreasing complexity. However, models that differ only in branch length prior but are otherwise identical have the same number of parameters. Bayes factors were calculated relative the highest estimated likelihood model (denoted by **). Models that differ from this model by less than 10 in Bayes factor comparison are denoted by *. The simplest model with the default exp(10) branch length prior is denoted by ††. More parameterized models relative to which it would be strongly preferred in Bayes factor comparison (difference >10) are denoted by †.

Number of partitions	Among-partition rate variation model	Among-character rate variation model	Branch length prior (exp[x])	Model likelihood	Bayes Factor
4	Unlinked branch lengths	Per partition gamma	5	−1502.9	27.3
4	Unlinked branch lengths	Per partition gamma	10	−1492.1	5.6*
4	Unlinked branch lengths	Per partition gamma	20	−1498.8	19.1
4	Unlinked branch lengths	Per partition gamma	40	−1499	19.4
4	Unlinked branch lengths	Shared gamma	5	−1503.3	28.1
4	Unlinked branch lengths	Shared gamma	10	−1489.2**	
4	Unlinked branch lengths	Shared gamma	20	−1494.5	10.5*
4	Unlinked branch lengths	Shared gamma	40	−1499.1	19.7
4	Unlinked branch lengths	Equal	5	−1509.9	41.3
4	Unlinked branch lengths	Equal	10	−1496.5	14.6
4	Unlinked branch lengths	Equal	20	−1499.6	20.8
4	Unlinked branch lengths	Equal	40	−1537.3	96.2†
4	Linked rate multiplier	Per partition gamma	5	−1529	79.4
4	Linked rate multiplier	Per partition gamma	10	−1522.9	67.4
4	Linked rate multiplier	Per partition gamma	20	−1522.4	66.2
4	Linked rate multiplier	Per partition gamma	40	−1524.9	71.4
4	Linked rate multiplier	Shared gamma	5	−1526.1	73.8
4	Linked rate multiplier	Shared gamma	10	−1522.5	66.5
4	Linked rate multiplier	Shared gamma	20	−1522.4	66.3
4	Linked rate multiplier	Shared gamma	40	−1527.2	75.9
4	Linked rate multiplier	Equal	5	−1529.7	80.9
4	Linked rate multiplier	Equal	10	−1527.3	76.2
4	Linked rate multiplier	Equal	20	−1525.6	72.8
4	Linked rate multiplier	Equal	40	−1537	95.6†
3	Unlinked branch lengths	Per partition gamma	5	−1500.4	22.4
3	Unlinked branch lengths	Per partition gamma	10	−1493.8	9.1*
3	Unlinked branch lengths	Per partition gamma	20	−1501.3	24.2
3	Unlinked branch lengths	Per partition gamma	40	−1496.3	14.1
3	Unlinked branch lengths	Shared gamma	5	−1503.5	28.6
3	Unlinked branch lengths	Shared gamma	10	−1492.9	7.3*
3	Unlinked branch lengths	Shared gamma	20	−1501	23.6
3	Unlinked branch lengths	Shared gamma	40	−1494.6	10.8*
3	Unlinked branch lengths	Equal	5	−1502.4	26.3
3	Unlinked branch lengths	Equal	10	−1494.2	9.9*
3	Unlinked branch lengths	Equal	20	−1498.6	18.7
3	Unlinked branch lengths	Equal	40	−1535	91.5†
3	Linked rate multiplier	Per partition gamma	5	−1527	75.6
3	Linked rate multiplier	Per partition gamma	10	−1526.5	74.4
3	Linked rate multiplier	Per partition gamma	20	−1525.2	71.9
3	Linked rate multiplier	Per partition gamma	40	−1526.1	73.8
3	Linked rate multiplier	Shared gamma	5	−1528.9	79.3
3	Linked rate multiplier	Shared gamma	10	−1521.8	65.2#
3	Linked rate multiplier	Shared gamma	20	−1523.6	68.7
3	Linked rate multiplier	Shared gamma	40	−1525.9	73.2
3	Linked rate multiplier	Equal	5	−1529.6	80.7
3	Linked rate multiplier	Equal	10	−1527.9	77.2
3	Linked rate multiplier	Equal	20	−1526.3	74.2
3	Linked rate multiplier	Equal	40	−1537.9	97.3†
2	Unlinked branch lengths	Per partition gamma	5	−1501.6	24.8
2	Unlinked branch lengths	Per partition gamma	10	−1495.9	13.2
2	Unlinked branch lengths	Per partition gamma	20	−1497.8	17
2	Unlinked branch lengths	Per partition gamma	40	−1499.4	20.3
2	Unlinked branch lengths	Shared gamma	5	−1503	27.5
2	Unlinked branch lengths	Shared gamma	10	−1497.9	17.3
2	Unlinked branch lengths	Shared gamma	20	−1499.6	20.8
2	Unlinked branch lengths	Shared gamma	40	−1501.4	24.3
2	Unlinked branch lengths	Equal	5	−1505.8	33.1
2	Unlinked branch lengths	Equal	10	−1500.1	21.6
2	Unlinked branch lengths	Equal	20	−1505.4	32.2
2	Unlinked branch lengths	Equal	40	−1528.4	78.3
2	Linked rate multiplier	Per partition gamma	5	−1526.4	74.3
2	Linked rate multiplier	Per partition gamma	10	−1523.6	68.8
2	Linked rate multiplier	Per partition gamma	20	−1524.6	70.6

(Continued on next page)

TABLE 1.   Summary of 80 models tested (nonrandomized data) using the Clarke et al. (2006) data set. Models are sorted in order of decreasing complexity. However, models that differ only in branch length prior but are otherwise identical have the same number of parameters. Bayes factors were calculated relative the highest estimated likelihood model (denoted by **). Models that differ from this model by less than 10 in Bayes factor comparison are denoted by *. The simplest model with the default exp(10) branch length prior is denoted by[††]. More parameterized models relative to which it would be strongly preferred in Bayes factor comparison (difference >10) are denoted by[†]. *(Continued)*

Number of partitions	Among-partition rate variation model	Among-character rate variation model	Branch length prior (exp[x])	Model likelihood	Bayes Factor
2	Linked rate multiplier	Per partition gamma	40	−1526.5	74.5
2	Linked rate multiplier	Shared gamma	5	−1526.8	75.1
2	Linked rate multiplier	Shared gamma	10	−1532.2	86[†]
2	Linked rate multiplier	Shared gamma	20	−1522.9	67.3
2	Linked rate multiplier	Shared gamma	40	−1528.4	78.3
2	Linked rate multiplier	Equal	5	−1531.7	84.9[†]
2	Linked rate multiplier	Equal	10	−1529.9	81.2
2	Linked rate multiplier	Equal	20	−1527	75.5
2	Linked rate multiplier	Equal	40	−1542.3	106.1[†]
1	n/a	Shared gamma	5	−1526.9	75.3
1	n/a	Shared gamma	10	−1521.4	64.4
1	n/a	Shared gamma	20	−1521.9	65.4
1	n/a	Shared gamma	40	−1525.1	71.8
1	n/a	Equal	5	−1529.3	80.1
1	n/a	Equal	10	−1525.3	72.1[††]
1	n/a	Equal	20	−1527.2	75.9[†]
1	n/a	Equal	40	−1537	95.4[†]

Results

Bayesian analysis yielded clades with posterior probabilities with only minor differences from topologies obtained via maximum parsimony (compare Fig. 1 to Fig. 3). Interrelationships among enantiornithine taxa were variable, whereas posterior probabilities for the whole clade were high (0.97 to 0.99). Some analyses recovered *Cathayornis* and *Concornis* as sister taxa. Others additionally or exclusively identified a *Neuquenornis* and *Gobipteryx* relationship (e.g., Fig. 3; Supplemental Fig. 1, available online at www.systematicbiology.org). Both of these clades had low posterior probabilities. No analyses placed *Lithornis* as the sister taxon to *Crypturellus* (Fig. 3), a relationship that was recovered in the maximum parsimony analyses (Fig. 1; Clarke et al., 2006). However, the *Lithornis* placement in the Bayesian analyses was previously recovered with a different data set and taxon sample (Clarke and Chiappe, 2001). Additionally, in the maximum parsimony analyses, *Yanornis* was the sister taxon to a clade composed of *Yixianornis* and *Songlingornis* (Fig. 1). None of the Bayesian analyses recovered these three taxa as a clade or the sister taxon relationship between *Yixianornis* and *Songlingornis*; rather, these taxa were placed as successive divergences or in unresolved polytomies with the lineage leading to Aves (Fig. 3). Lack of support for this clade may be related to its position near the center of the tree, which is a region exhibiting lower clade prior probabilities with implementation of a flat topological prior (see review in Brandley et al., 2006) and is also recognized as problematic in Bayesian analyses including missing data (Goloboff and Pol, 2005) for related reasons (see also Discussion).

The estimated model likelihoods for all 80 analyses and Bayes factor comparisons of each to the most preferred model are summarized in Table 1 and Figure 4. Models were distributed in two clusters (Fig. 4). The group of high model likelihoods (higher than −1510) consisted exclusively of multipartition models (two to four partitions) with independent among-partition rate variation (unlinked branch lengths). The group of lower model likelihoods (lower than −1521) consisted of models with a single partition, all multipartition models with among-partition rate variation proportional, as well as three models with unlinked branch lengths and an exp(40) branch length distribution prior. The highest model likelihood (−1489.2) was found in the four-partition model with unlinked branch lengths, a single shared gamma parameter, and a branch length prior distribution of exp(10) (Supplemental Fig. 1). The simplest model (single partition, no gamma, exp[10] branch length prior) had a likelihood of −1525.3, in the middle of the lower group of model likelihoods. More complex models were not always preferred over less complex models, when compared either by model likelihood or by Bayes factor comparison.

Absolute variation in estimated model likelihoods from repeated analyses of the same model (equal number of iterations) was relatively low (coefficient of variation <0.5%). However, in our most extreme example, estimated likelihoods for successive analyses of the same model were sometimes found to differ by more than 10 log likelihood units (range in model likelihood differences: 3.4 to 12.8; Bayes factor values between 6.8 and 25.6). When the iteration length of the highest likelihood model was extended to 10×10^6 and 20×10^6 iterations, the likelihood estimates varied by less than 5 log likelihood units when compared to the original analysis (Bayes factor values: 9.4 and 8.8, respectively; 5×10^6 iterations = −1491.6; 10×10^6 iterations = −1496.3; 20×10^6 iterations = −1496.0).

The seven highest model likelihoods (and 13 of the highest 16) were recovered for three- or four-partition
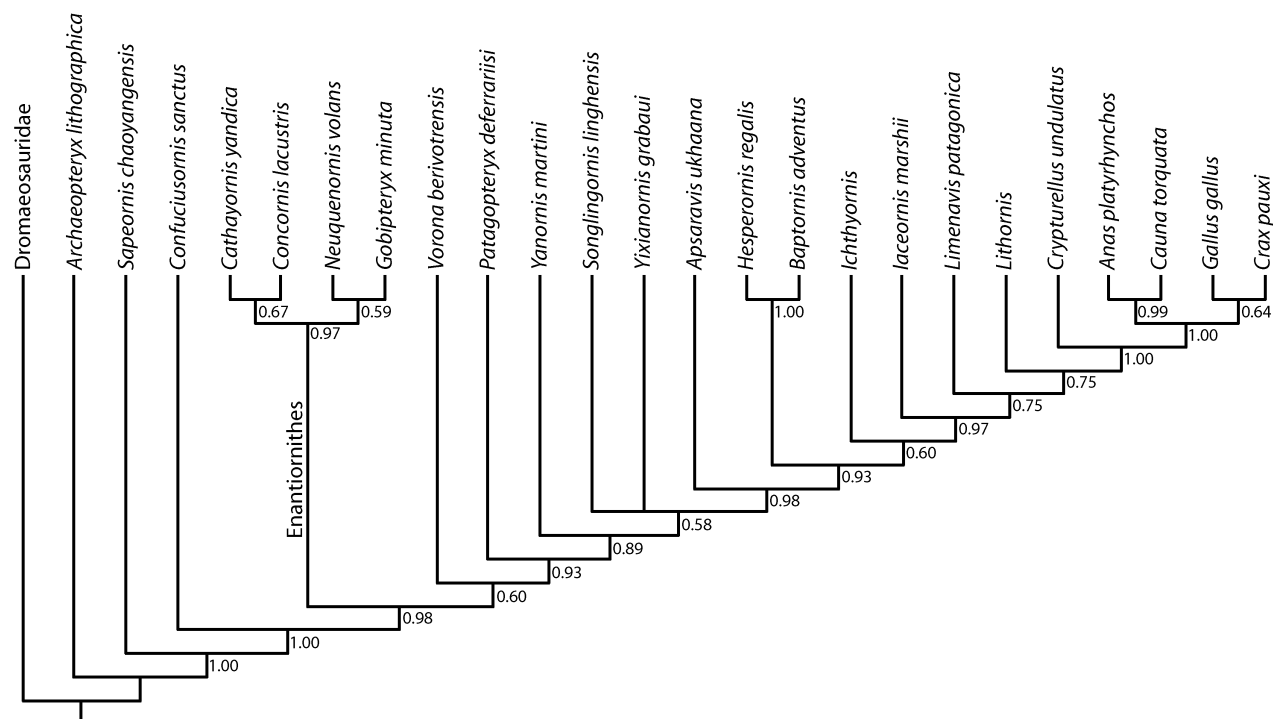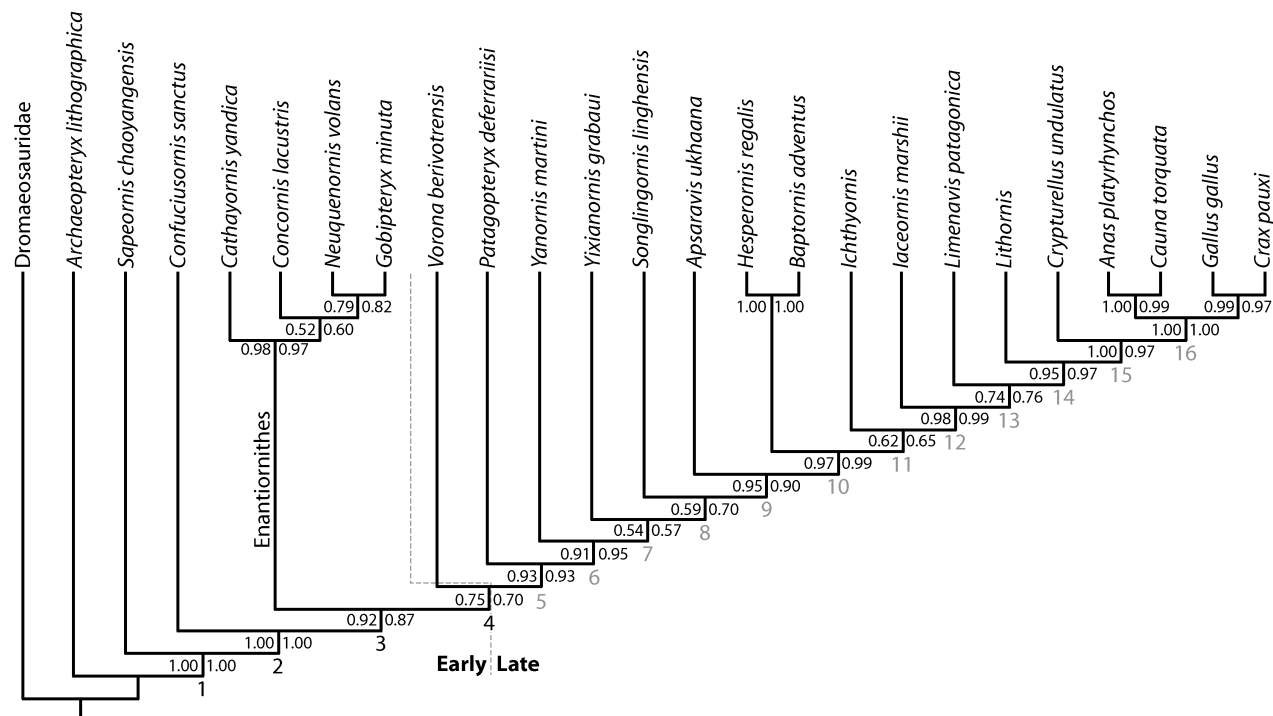
FIGURE 3.   Comparison of topologies from Bayesian phylogenetic analyses utilizing (a) a single partition of all characters and (b) three- and four-partition models. Tree topology is similar, except for differences in the interrelationships of Enantiornithes and the positions of *Songlingornis*, *Yanornis*, and *Yixianornis* relative to one another. Posterior probabilities of clades are listed next to the nodes. In (b), the posterior probability to the left of the branch is for the three-partition model, and the posterior probability to the right of the branch is for the four-partition model. The branch numbers referenced in Figure 6 are shown below the posterior probability (black and gray corresponding to Fig. 6), and the location of the hypothetical transition point from pectoral-dominated evolutionary change to pelvic-dominated change (e.g., Chiappe, 2002) is noted by the gray dashed line.
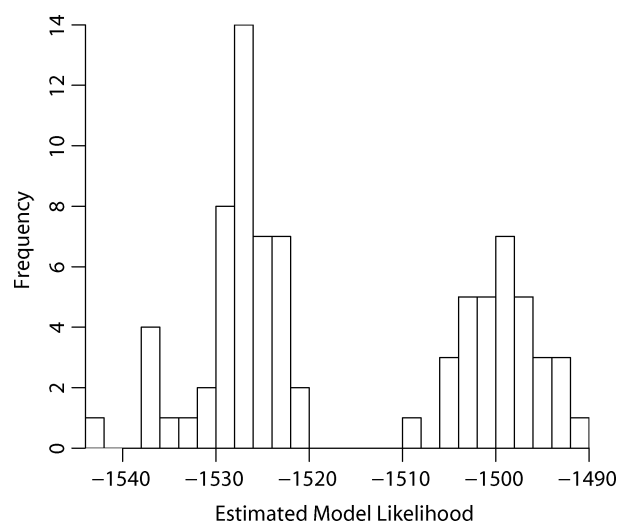
FIGURE 4. Distribution of estimated likelihoods for 80 different models. Two clusters are present. The cluster with higher model likelihoods includes only multipartition models with unlinked branch lengths. The cluster of lower model likelihoods includes all single-partition models, all models with among-partition rate variation modeled as linked or proportional, and several multipartition models with unlinked branch lengths and an exp(40) branch length distribution prior.
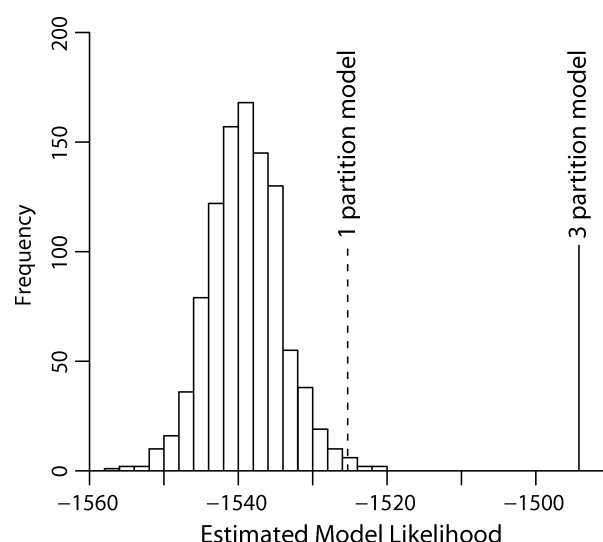

FIGURE 5. Distribution of estimated model likelihoods from Bayesian analyses of three partition models with characters randomly assigned to partitions. Model likelihoods from these models were significantly lower than for the same three-partition model with anatomically based partitions (solid line = −1494.2; $P < 0.001$). Estimated likelihoods from the randomized character partition models were also significantly lower than even single-partition models (dashed line = −1525.3; $P = 0.006$).

models. Although the two highest model likelihoods were found for four-partition models, the Bayes factor between the best three-partition model and the best overall model was only 7.3, indicating that the four-partition model is not very strongly preferred relative to the less-parameterized model (sensu Nylander et al., 2004). By contrast, the Bayes factor (calculated relative to the highest model likelihood) of the best two-partition (pectoral and all other characters) model is 13.2 and of the best single-partition model, 64.4. Bayes factor comparison indicates that, to different degrees, the four-partition model would be very strongly preferred relative to either of these models. Modeling among-partition rate variation as variable but proportional never improved model likelihood above that of the best single-partition model regardless of partition number and other varied parameters. The most preferred model with proportional among-partition rate variation was a three-partition analysis with a shared gamma and a branch length prior of exp(10) that had model likelihood of −1521.8 (indicated with a "#" in Table 1); however, the Bayes factor comparison between this model and the highest-likelihood model was 65.2. The mean estimated model likelihood from the 1000 analyses of partitioned, randomly selected characters was −1539.0, in all iterations lower than the observed model likelihood (−1494.2; $P \leq$ 0.001; Fig. 5). The models with partitions of randomly selected characters were also significantly poorer than the single-partition model (exp[10]; "††" in Table 1; model likelihood = −1525.3; $P = 0.006$; Fig. 5).

The parameters that most influenced the model likelihood were then analyzed using ANOVA. We used the stepAIC function (Venables and Ripley, 2002) in R (version 2.5.1; R Development Core Team, 2007) to deter-

mine which of the four factors was most important: number of partitions, type of among-partition rate variation (proportional, independent), model for among character rate variation (equal, gamma distributed, per-partition gamma distributed), or the prior probability distribution on branch lengths. Starting with the simplest model (likelihood modeled by partition number only), successively more complex models were iteratively fit. The optimal combination of factors included only partition number, type of among-partition rate variation, and among-character rate distribution, but not the branch length prior. Based on the results of the stepAIC analysis, we fit an ANOVA with just the three relevant factors and their interactions. All three factors were highly significant predictors of model likelihood, with type of among-partition rate variation most influential ($F_{1,60} = 205.66$, $P \ll 0.0001$), followed by among-character rate variation (equal, gamma, per-partition gamma; $F_{2,60} = 10.27$; $P = 0.0001$), and number of partition ($F_{3,60} = 6.44$; $P = 0.0007$). No interactions among the factors were found to be significant at $P < 0.05$. These statistical findings are consistent with the qualitative observations of model performance from results presented in Table 1.

BRANCH LENGTH COMPARISONS

Methods

In the partitioned Bayesian analyses, branch lengths could be compared both across the tree and among partitions. Statistical tests for differences among mean posterior branch length estimates were formulated based on the expectations of the pectoral early–pelvic late hypothesis, using the "shift point" proposed by Chiappe (2002)

and the branch lengths from a three-partition, unlinked, equal-rate model (exp[10]). Branches basal to this shift point, at the *Vorona* divergence (Fig. 3b; Supplemental Fig. 1), were included in one group (here referred to as the "early" group). The *Vorona* branch, as well as all branches in the tree that were closer to or included those within Aves, were placed in the second (the "late" group). We first examined a close restatement of the original pectoral early–pelvic late hypothesis (Chiappe, 1995a; 2002), involving a differential frequency of character-state change in pectoral and pelvic characters only along the spine internodes that are treated in SSA and excluding branches leading to terminal taxa off of this path. Further comparisons of all branch lengths (spine and all terminal branches) were then made between the early and late groups for these two partitions.

Using *t*-tests (assuming equal variances among samples), we compared mean posterior estimates of branch lengths in the early and late groups. First, only the means of the spine branch lengths in the early and late groups were compared. This sample is comprised only of those internodes treated in SSA. Second, the means of all estimated branch lengths from the early and late groups were compared. Because of the small sample sizes (n = 4–14 in the early group and n = 11–33 in the late group), we were unable to determine if the branch lengths were normally distributed. To avoid potential bias due to a non-normal distribution, *P* values were determined analytically by randomly reassigning branch lengths to the early and late groups (10,000 permutations, sampling without replacement; following the methods described in Roff, 2006). One-tailed *P* values are presented for hypotheses for which we had a priori predictions based on the hypotheses under evaluation (longer mean branch lengths in pectoral partition), and all other *P* values are two-tailed. All statistical analyses were carried out using R (version 2.5.1; R Development Core Team, 2007).

Results

For the most preferred model (** in Table 1), mean branch lengths for the pectoral, pelvic, cranial and axial partitions were 0.09, 0.08, 0.09, and 0.10, respectively. Results of the comparison of branch lengths in two-, three-, and four-partition models are summarized in Table 2. Only one comparison of early and late branch lengths for a given character partition and branch length set (i.e., all or spine only) was statistically significant: the mean estimated pectoral branch lengths for only the spine internodes of the early and late groups for the three- and four-partition models. In this comparison, the spine pectoral branch lengths in the early group are significantly longer than in those in the late group (*P* = 0.04, one-tailed). A similar but nonsignificant result was found for a two-partition, unlinked, equal-rate (exp[10]) model; spine pectoral branch lengths were longer in the early group (*P* = 0.08, one-tailed; see Table 2). However, when all branches were included (not spine only), no statistically significant difference was found between early ver-

TABLE 2. Results of *t*-tests comparing branch lengths below and above the divergence of *Vorona* for two-, three-, and four-partition models. The four-partition model was the most preferred model in Bayes factor comparison (Table 1: denoted with **). One-tailed *P* values are presented for hypotheses about which we had a priori branch length predictions are denoted by [†]. Low *P* values indicate a significant difference in mean branch lengths for that character partition when all branch lengths are divided into two groupings (below the divergence of *Vorona* and above the divergence of *Vorona*). By contrast, high *P* values indicate no significant difference in mean branch lengths between the two groupings. All *P* values were determined analytically by randomization (see text for details).

	Spine only	All branches
4-Partition		
Pectoral	0.04[†]	0.47[†]
Pelvic	0.42[†]	0.32[†]
Cranial	0.57	0.82
Axial	0.26	0.15
3-Partition		
Pectoral	0.04[†]	0.43[†]
Pelvic	0.42[†]	0.31[†]
All other characters	0.55	0.88
2-Partition		
Pectoral	0.09[†]	0.44[†]
All other characters	0.78	0.59

sus late mean pectoral, pelvic, or all other characters' branch lengths.

When mean estimated branch lengths across the whole tree were compared among the three anatomical partitions, no statistically significant differences were found (pectoral versus pelvic: $t = 0.66$, $df = 94$, $P = 0.51$; pectoral versus all other characters: $t = -0.27$, $df = 94$, $P = 0.79$; pelvic versus all other characters: $t = -1.1$, $df = 94$, $P = 0.27$). Visual inspection of plots of estimated pectoral branch lengths relative to pelvic and all other characters' branch lengths by branch (Fig. 6) shows individual early pectoral branches are in some cases relatively longer when compared to pelvic and all other character lengths for the same branch as well as shorter (numbers and points in black, Fig. 6). Similarly, for late branches (numbers and points in grey, Fig. 6), there is no systematic deviation towards longer relatively pelvic branch lengths.

We compared the distribution of the posterior estimates of branch lengths for the pectoral, pelvic, cranial, and axial partitions for the highest likelihood model (four partitions, unlinked branch lengths, single shared gamma parameter, branch length prior distribution of exp[10]) to the prior distribution of branch lengths (exp[10]) using Komolgorov-Smirnov tests. In all but one partition, the difference between the posterior distribution of branch lengths and the prior distribution was statistically significant. Thus, the data support the hypothesis that the partitions contain sufficient information to inform estimates of the branch length posteriors.

DISCUSSION

Evaluation of the Pectoral Early–Pelvic Late Hypothesis

Applying SSA to the present data set, we do not confirm the originally proposed pectoral early–pelvic late pattern; both peaks in forelimb and hind limb
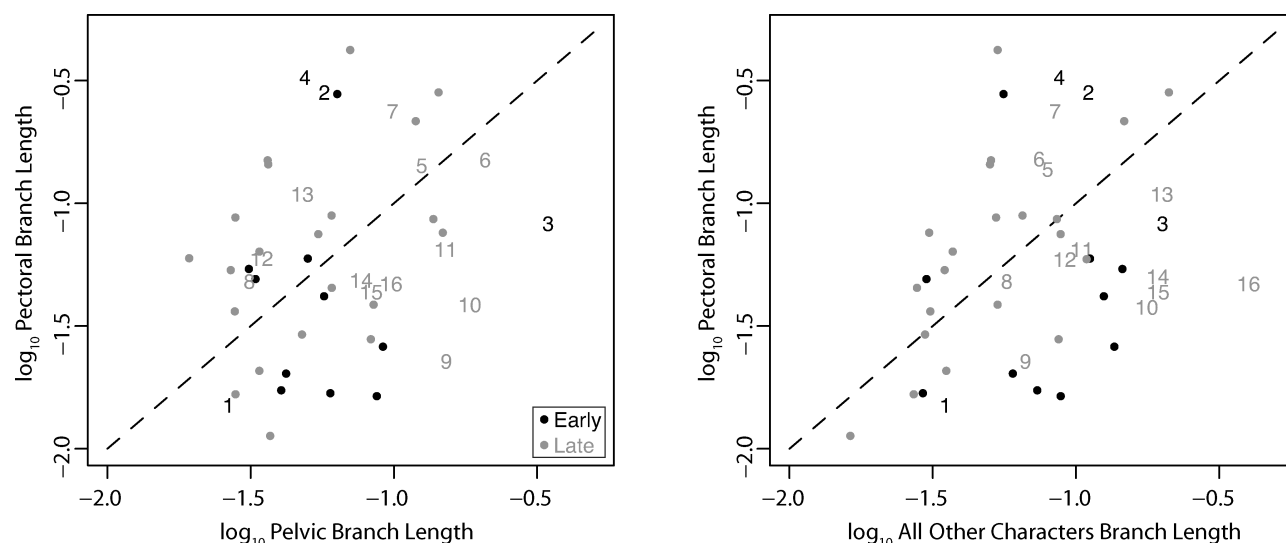
FIGURE 6.   Pectoral branch lengths are plotted against (left) pelvic and (right) all other branch lengths for spine internodes only (numbers corresponding to nodes labeled in Fig. 3b) and all branches (numbers and circles) from analysis of the three partition-unlinked equal-rates model (exp[10]). The dashed line represents equality of branch lengths. Points falling above the line indicate a longer pectoral branch, and those falling below the line indicate a shorter pectoral branch relative to pelvic or other. "Early" branch lengths, those below the SSA identified shift point at the divergence of *Vorona* (see Fig. 3b), are shown in black numbers and points, and "late" branch lengths are shown in gray.

apomorphies are now located phylogenetically early in avialan evolution (Fig. 1; black lines). From the Bayesian results, statistical tests do support the hypothesis that pectoral branch lengths for spine internodes are longer in the early group than in the late group for three- and four-partition models. Visual comparisons of individual branches locate the main differences in pectoral versus pelvic branch lengths in two early spine internodes for pectoral characters (numbered 2 and 4, Fig. 6). These internodes correspond to two early peaks recovered with SSA with the Clarke et al. (2006) complete taxonomic sample (Fig. 1a). This pattern may identify key sampling biases or could have a biological explanation; it appears driven by only these two branches. However, because the late apomorphy peak in pelvic characters is not recovered in SSA with increased taxon sampling (Fig. 1b) and because late pelvic branch lengths for spine internodes are not significantly longer in the Bayesian analyses (Table 2, Fig. 6), the hypothesized sequential development of wing and then leg morphologies is not supported using either method.

Functional anatomical studies suggesting that major reorganization of the hind limb had in fact taken place phylogenetically early after the origin of flight are consistent with the disappearance of the pelvic late pattern using SSA with increased taxon sampling as well as with the Bayesian results. Avian hind limb locomotor function, including the strongly flexed hip and knee joints, more vertical orientation of the metatarsus, and the derived limb control mechanisms of crown-group birds have been inferred to be largely present phylogenetically before the divergence of Hesperornithes (Farlow et al., 2000; Hutchinson and Gatesy, 2000; Hutchinson, 2002). Further, coupled evolution of fore- and hind limb

modules early after the evolution of elongate wing feathers may in fact be predicted by wing-assisted incline running (WAIR) as a proposed fundamental behavioral/aerodynamic organizing principle in the origin of flight (Dial, 2003; Tobalske and Dial, 2007).

Limitations of the Summed Synapomorphy Approach

We find that SSA is sensitive to both taxonomic sampling and missing character data (Figs. 1, 2). Fundamental drawbacks of the existing formulations of this approach arise from two sources: (1) consideration of only unambiguously optimized character changes and (2) consideration of such changes optimized for spine internodes (i.e., along a path from the most recent common ancestor of the ingroup to the taxon of interest) as representing character changes towards the morphologies of the taxon of interest or, as in the present example, modernizing anatomical subregions (e.g., Chiappe, 2002:460) to reach character states in crown Aves. No study applying these methods has yet followed the character-by-character changes to establish if the summed apomorphies on the path or spine are retained in the taxon of interest. These states may be reversed either before or at the origin of this clade. Without investigation of the retention (possibly using Retention Index) of these characters, neither the early pectoral apomorphy peaks (Fig. 1) nor the differentially long branches in the pectoral set for these spine internodes in the Bayesian analyses (Table 2, Fig. 6) can be interpreted as necessarily representing rapid, or increased, pectoral change *toward* the crown clade avian flight apparatus. If the pectoral character peaks and differentially long pectoral branches early after the evolution of flight were a result of differential and

disjunctive evolution in this anatomical subregion, then the pattern should be seen in the total pectoral branch length set, including terminal branches as well.

In the Bayesian analyses, however, early and late pectoral branch lengths did not differ significantly when terminal branches were included with the spine internodes in these groups (Table 2), and their mean lengths were nearly identical. When the means of all branch lengths (whole tree) were compared, pectoral and pelvic sets also did not differ significantly. The appropriate comparisons, if this approach were to be applied more generally, would appear to be among complete branch length sets ideally calibrated such that comparisons may be among branch lengths reflecting the amount of character change per time interval.

Substantial morphological diversity among taxa representing phylogenetically early divergences after the origin of flight (e.g., taxa reviewed in Zhou, 2004) also undermines the SSA assumption of teleological progress towards the extant avian conformations of the hind limb and forelimb modules. If there is significant morphological diversity within a particular anatomical subregion, a hypothesized pattern for extant birds that may be related to locomotor modularity itself (Gatesy and Dial, 1996b), ambiguous character optimizations would be expected. Characters with a complex distribution in comparatively closely related taxa, perhaps because they are actually rapidly evolving, would be interpreted as evolutionarily static, and the perceived importance of morphological evolution happening in that subregion would be underestimated, given the way results from this method have been previously interpreted (see caudal region characters in Chiappe, 2002).

SSA can conceivably be reformulated to avoid referencing fine-tuning or modernization. However, patterns inferred from SSA more effectively measure which taxa are known from complete fossils, rather than morphological evolution in the clade under consideration. Any preservational biases toward or against recovery or completeness of certain elements (e.g., the skull; Fig. 2c) or influencing taxonomic sampling will directly influence the number of unambiguously optimized synapomorphies and therefore alter the perceived pattern of relevant subregion evolution. Virtually all analyses utilizing variants of SSA in attempting to compensate for these missing data effects will suffer from a common suite of drawbacks. Removing taxa or characters with missing data would result in significantly limiting the generality of any recovered pattern. Any hypothesis would then be only applicable to well-preserved taxa or a potentially skewed set of commonly preserved characters. Further, if SSA is used in conjunction with phylogenetic analysis, the latter would also be potentially impacted by exclusion of critically located specimens lacking one or more anatomical subregions of interest (Fig. 2a, b).

Utilizing data from ambiguously optimized characters (e.g., via ACCTRAN and DELTRAN optimizations) to give ranges of apomorphies at a single node is also problematic. For example, we recovered zero unambiguously optimized synapomorphies among cranial characters throughout early bird evolution (Fig. 2c). Investigating alternative optimizations of ambiguous cranial characters would simply place nearly all apomorphies at one of two nodes. As a result, and depending on which method was used, either early at the divergence of *Confuciusornis* or late at the origin of Neognathae, there would be exceptionally large number of optimized synapomorphies. Missing data issues impact the Bayesian branch length estimates as well, although these approaches would be expected to be more robust because of fundamental differences in the way branch length estimates are calculated in the M_k model as structural parameters (Lewis, 2001; and see further discussion below).

Implications for Bayesian Analyses of Morphological Data and Investigation of Mosaic or Disjunctive Patterns of Change

If model likelihoods and Bayes factor comparisons are judged a reliable arbiter of model performance (see Brown and Lemmon, 2007), then our results may guide the design of Bayesian analyses addressing morphological data. We also find significant opportunities for data exploration and for novel quantitative studies of morphological evolution with Bayesian approaches. Bayesian phylogenetic methods have remained so far comparatively little explored with morphology-only data sets, perhaps because of concerns over the model for morphological evolution proposed by Lewis (2001) and implemented in MrBayes (e.g., Goloboff and Pol, 2005, and references therein).

We find few differences between topologies recovered from the Bayesian analyses and those recovered using a maximum parsimony estimator. From the nature of differences in the recovered topologies as well as from the 80 investigated models and randomizations, we identify the factors most strongly associated with higher model likelihoods and call attention to some potential effects from implementation of a flat topological prior and missing data that deserve further investigation for paleontological data sets in particular.

Independently estimating branch lengths for character data partitioned by anatomical subregion shows the strongest relationship with higher model likelihoods as statistically assessed. However, only allowing branch lengths to vary freely among partitions (unlinked branch lengths) dramatically improved model likelihoods. Modeling among-partition rate variation with a rate multiplier (brlen = variable) resulted in lower likelihood models than single-partition analyses. The second and third most influential factors linked to higher model likelihoods were modeling among-site rate variation as gamma distributed (shared or per partition), and increasing partition number, respectively (Table 1). Inclusion of a single or per partition gamma improves model likelihood, but models with a shared single gamma in several instances have higher likelihoods than the more parameterized per partition gamma models (compare the two highest likelihood models recovered, Table 1). Consistent with previous results (e.g., Nylander et al., 2004),

we find that more richly parameterized models are generally, but not always, preferred over simpler ones. Varying the branch length prior did not show any systematic relationship to model likelihood. However, the five highest likelihood models had the MrBayes default branch length prior distribution, exp(10).

The range in estimated model likelihoods from repeated analyses of the same model or of a single model run for more iterations was similar ($\sim$5 log likelihood units; approximately 10 units in Bayes factor comparisons). Although 10 has been used as a standard cutoff for strong or very strong model preference in Bayesian model comparisons (Kass and Raftery, 1995; Nylander et al., 2004), such values remain interpretive guides without the power of statistical tests. Given that our observed variation in estimates of model likelihood were approximately 5 log likelihood units, we agree with previous authors, who have suggested that standard cutoffs used in Bayes factor comparison may need to be revisited (e.g., Brandley at al., 2005). A more conservative approach may be to use a larger difference in Bayes factor comparisons.

Even with observed variation in estimated model likelihoods, the fundamental result of anatomically partitioned models having significantly higher likelihoods than either single-partition models or models with proportional branch lengths remains strongly supported. No unpartitioned models and no models with proportional branch lengths were found in the higher likelihood cluster (Fig. 4, Table 1), and the upper and lower clusters differed by 23 units in Bayes factor comparison. This upper cluster did include, however, less partitioned and less parameter-rich models, varying from two to four partitions as well as gamma-distributed among-character rate variation and equal rates.

Analyses of 1000 three-partition data sets assembled from randomly selected characters investigated whether simply the additional branch length parameters alone led to model preference for the partitioned analyses. In none of the randomizations, however, did model likelihoods closely approach those of the three-partition model using the previously proposed biologically based partitions ($P \leq 0.001$; Fig. 5), and nearly all had lower model likelihoods than even the simplest model investigated: the single-partition equal-rate model ($p = 0.006$). The preference for all models with anatomical subregion partitions may support a disjunctive pattern for character change for the data set investigated. Lewis (2001) recognized that rate heterogeneity was reasonably as much a factor in morphological as in molecular sequence data. Techniques proposed to address this issue for morphological data (Lewis, 2001) were similar to those used with sequence data, which utilize a gamma distribution for the relative rates (Yang, 1994). Models with gamma-distributed rates have been preferred in previous analyses of morphological data (Wiens et al., 2005; Müller and Reisz, 2006). The hidden Markov model (Felsenstein and Churchill, 1996) that allows for correlation among neighboring characters was considered unrealistic by Lewis (2001) because location in the data matrix was considered arbitrary. However, this is not the case for many morpho-

logical data sets (e.g., Chiappe, 2002; Clarke et al., 2006; Livezey and Zusi, 2006), in which characters from the same anatomical subregion are most often adjacent in the matrix with distinct, arguably nonarbitrary breaks (e.g., between skull and neck). For these data sets or subsets of these data sets, implementation of the hidden Markov model may deserve further consideration.

Further investigation of character partitions in Bayesian analyses of morphological data are needed given the results for model likelihoods and phylogenetic analyses found here, as well as to confirm the utility of the methods proposed for studying morphological evolution. These analyses should include investigation of models with different partitioning schemas. Preliminary analysis of other data sets (Boyd and Clarke, 2007) suggests that the approach may be used to locate partitions of interest. For example, Bayes factor comparisons among models with distinct partitioning strategies informed by prior hypotheses concerning morphological trends in a lineage may be evaluated. Some partitioning strategies yield small increases in estimated model likelihoods, whereas other partitioning strategies were strongly preferred in model comparison. Covarying character complexes from different parts of the avian skeleton have also been identified (Nemeschkal et al., 1992) that might be used as possible partitions. Similarly, character sets proposed to be associated with developmental and biomechanical modules might be investigated. Partition size as a potential source of systematic error also deserves exploration; very small partitions could lack sufficient information to inform posterior estimates of branch lengths. New methods for stochastic character mapping (Bollback, 2006) may also be effective tools used in conjunction with partitioned Bayesian analyses for consideration of the pattern of morphological change.

The topologies recovered from the Bayesian analyses, although similar to those recovered using a maximum parsimony estimator, do not include two clades near the center of what would be the unrooted network that are recovered in the maximum parsimony analysis. These clades included extinct terminal taxa that range in completeness from 16% to 65% (84% to 35% missing character data) but are nonetheless recovered with significant bootstrap support (e.g., the *Yixianornis* clade; 75%) and have some of the highest Bremer support values found in the maximum parsimony analysis (Clarke et al., 2006). Goloboff and Pol (2005) found taxa represented only by missing entries (no data) to be placed near the center of an unrooted network, an effect due to implied clade prior probabilities in this region (near the center of the network) from implementation of a flat topological prior. Flat topological priors confer unequal clade prior probabilities with a more pronounced effect expected in small data sets (reviewed in Brandley et al., 2006).

Although it has been indicated that missing data may be less of an issue in model-based phylogenetic approaches than maximum parsimony analyses (Wiens, 2005), it merits further scrutiny whether lower prior probabilities of clades of intermediate size (e.g., Brandley et al., 2006), combined with the possible missing data

effect noted by Goloboff and Pol (2005) and the relatively small size of many paleontological data sets, may be of concern for Bayesian analyses of morphological data that include fossil taxa. Fossil taxa often sample stem lineages of extant clades and not uncommonly comprise many successive sister taxa to these crowns. These taxa may, thus, be differentially positioned near the center of the unrooted network, in this region of slightly lower clade prior probabilities, and would be relatively less complete. Analyses of further real and simulated data sets are necessary to explore the possible generality of any effects of these factors.

Conclusions

Comparisons of model performance and posterior mean branch lengths across anatomical subregion partitions afforded by Bayesian analyses comprise nuanced quantitative methods of investigating morphological evolution with phylogenetic data sets. So far there have been few methods proposed for use with morphology-only data sets or for more than a small sample of characters. Results from the Bayesian approach used here did not support the specific pectoral early–pelvic late hypothesis tested. However, comparisons of branch lengths among partitions were able to locate the only significant difference in pectoral versus pelvic branch lengths in two early spine internodes. Although no other branch length comparisons across partitions and taxa yielded significant results, these two longer pectoral branches were identified. This pattern may identify key sampling biases or have a biological explanation as it appears to be driven by only these two branches.

The strong preference in Bayes factor comparisons for all models with anatomical subregion partitions and the negative randomization results are consistent with a disjunctive or mosaic pattern of character change for the data set investigated. Based on our analysis, the highest model likelihoods are predicted by (1) type of among-partition rate variation, specifically allowing unlinked branch lengths and utilizing anatomical subregion partitions; (2) modeling among-character rate variation as gamma distributed (single or per-partition); and (3) number of partitions. Modifying branch length prior was not correlated with higher model likelihoods.

These results have implications for model parameterization of analyses of morphological data sets as well as for the study of morphological evolution. They suggest that other nonrandom partitioning schemas deserve investigation in Bayesian analyses. These could include investigation of models partitioning by different methods from the more theory neutral (e.g., by bone) to partitioning by subregions base on proposed modules or characters with identified functional or ontogenetic linkages. The particular differences between Bayesian and maximum parsimony trees in recovered intermediate-sized clades also indicate that possible combined missing data and unequal clade prior probabilities effects deserve further scrutiny in morphological data sets including extinct taxa.

The summed synapomorphy approach to patterns of morphological evolution is sensitive to both taxonomic sampling and missing character data. However, more troubling with SSA is its consideration of character changes only on spine internodes; i.e., along a path from the most recent common ancestor of the ingroup to the taxon of interest. Synapomorphies optimized for these internodes cannot be expected to equal, as assumed, the set of character changes *towards* the morphologies of the taxon of interest or modernizing anatomical subregions to achieve, for example, crown clade character states. If SSA-like approaches are to be refined, character-by-character analysis will be necessary to assess whether or not summed apomorphies on the path or spine are retained or reversed in the taxon of interest.

References

Barton, R. A., and P. H. Harvey. 2000. Mosaic evolution of brain structure in mammals. Nature 405:1055–1058.

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

Boyd, C., and J. A. Clarke. 2007. Taxonomic chimeras and stratigraphic incongruence resolved: A new phylogeny of the basal Ornithischia. J. Vert. Paleo. 27:51A

Brandley, M. C., A. D. Leaché, D. L. Warren, and J. A. McGuire. 2006. Are unequal clade priors problematic for Bayesian phylogenetics? Syst. Biol. 55:138–146.

Brandley, M. C., A. Schmitz, and T. W. Reeder. 2005. Partitioned Bayesian analyses, partition choice, and the phylogenetic relationships of scincid lizards. Syst. Biol. 54:373–390.

Brown, J. M., and A. R. Lemmon. 2007. The importance of data partitioning and the utility of Bayes factors in Bayesian phylogenetics. Syst. Biol. 56:643–655.

Cavalier-Smith, T. 2002. The neomuran origin of archebacteria, the negibacterial root of the universal tree and bacterial megaclassification. Int. J. Syst. Evol. Evol. Microbiol. 52:7–76.

Chatterjee, S. 1997. The rise of birds: 225 million years of evolution. Johns Hopkins University Press, Baltimore, Maryland.

Chiappe, L. M. 1991. Cretaceous birds of Latin America. Cret. Res. 12:55–63.

Chiappe, L. M. 1995a. The first 85 million years of avian evolution. Nature 378:349–355.

Chiappe, L. M. 1995b. The phylogenetic position of the Cretaceous birds of Argentina: Enantiornithes and *Patagopteryx deferrariisi*. Cour. Forschungsinst. Senckenb. 181:55–63.

Chiappe, L. M. 1996. Late Cretaceous birds of southern South America: Anatomy and systematics of Enantiornithes and *Patagopteryx deferrariisi*. Münch. Geowiss. Abh. A 30:203–244.

Chiappe, L. M. 2002. Basal bird phylogeny: Problems and solutions. Pages 448–471 *in* Mesozoic birds: Above the heads of dinosaurs (L. M. Chiappe and L. M. Witmer, eds.). University of California Press, Berkeley.

Clark, J. M., M. A. Norell, and P. J. Makovicky. 2002. Cladistic approaches to the relationships of birds to other theropod dinosaurs. Pages 31–61 *in* Mesozoic birds: Above the heads of dinosaurs (L. M. Chiappe and L. M. Witmer, eds.). University of California Press, Berkeley.

Clarke, J. A. 2004. Morphology, phylogenetic taxonomy, and systematics of *Ichthyornis* and *Apatornis* (Avialae: Ornithurae). Bull. Am. Mus. Nat. Hist. 286:1–179.

Clarke, J. A., and L. M. Chiappe. 2001. A new carinate bird from the Late Cretaceous of Patagonia (Argentina). Am. Mus. Novit. 3323:1–23.

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.

Clarke, J. A., Z. Zhou, and F. Zhang. 2006. Insight into the evolution of avian flight from a new clade of Early Cretaceous ornithurines from China and the morphology of *Yixianornis grabaui*. J. Anat. 208:287–308.

Cracraft, J. 1970. Mandible of *Archaeopteryx* provides an example of mosaic evolution. Nature 226:1268.

de Beer, G. R. 1954. *Archaeopteryx lithographica*. A study based upon the British Museum specimen. British Museum (Natural History), London.

Dial, K. P. 2003. Wing-assisted incline running and the evolution of flight. Science 299:402–404.

Farlow, J. O., S. M. Gatesy, T. R. Holtz, Jr., J. R. Hutchinson, and J. M. Robinson. 2000. Theropod locomotion. Am. Zool. 40:640–663.

Felsenstein, J., and G. A. Churchill. 1996. A hidden Markov model approach to variation among site in rate of evolution. Mol. Biol. Evol. 13:93–104.

Forster, C. A., S. D. Sampson, L. M. Chiappe, and D. W. Krause. 1998. The theropod ancestry of birds: New evidence from the Late Cretaceous of Madagascar. Science 279:1915–1919.

Garner, J. P., G. K. Taylor, and A. L. R. Thomas. 1999. On the origins of birds: The sequence of character acquisition in the evolution of avian flight. Proc. R. Soc. Lond. B 266:1259–1266.

Gatesy, S. M. 2002. Locomotor evolution on the line to modern birds. Pages 432–447 *in* Mesozoic birds: Above the heads of dinosaurs (L. M. Chiappe and L. M. Witmer, eds.). University of California Press, Berkeley.

Gatesy, S. M., and K. P. Dial. 1996a. From frond to fan: *Archaeopteryx* and the evolution of short-tailed birds. Evolution 50:2037–2048.

Gatesy, S. M., and K. P. Dial. 1996b. Locomotor modules and the evolution of avian flight. Evolution 50:331–340.

Gatesy, S. M., and K. M. Middleton. 1997. Bipedalism, flight, and the evolution of theropod locomotor diversity. J. Vert. Paleo. 17:308–329.

Gauthier, J. A., and K. de Queiroz. 2001. Feathered dinosaurs, flying dinosaurs, crown dinosaurs, and the name "Aves." Pages 7–41 *in* New perspectives on the origin and early evolution of birds: Proceedings of the International Symposium in Honor of John H. Ostrom (J. A. Gauthier and L. F. Gall, eds.). Peabody Museum of Natural History, Yale University, New Haven.

Glenner, H., A. Hansen, M. Sørensen, F. Ronquist, J. P. Huelsenbeck, and E. Willerslev. 2004. Bayesian inference of the metazoan phylogeny: A combined molecular and morphological approach. Curr. Biol. 14:1644–1649.

Goloboff, P. A., and D. Pol. 2005. Parsimony and Bayesian phylogenetics. Pages 148–159 *in* Parsimony, phylogeny, and genetics (V. A. Albert, ed.) Oxford University Press, Oxford, UK.

Hou, L., L. D. Martin, Z. Zhou, and A. Feduccia. 1996. Early adaptive radiation of birds: Evidence from fossils from northeastern China. Science 274:1164–1167.

Huelsenbeck, J. P., B. Larget, R. E. Miller, and F. Ronquist. 2002. Potential applications and pitfalls of Bayesian inference of phylogeny. Syst. Biol. 51:673–688.

Huelsenbeck, J. P., and F. Ronquist. 2001. MrBayes: Bayesian inference of phylogeny. Bioinformatics 17:754–755.

Hutchinson, J. R. 2001. The evolution of pelvic osteology and soft tissues on the line to extant birds (Neornithes). Zool. J. Linn. Soc. 131:123–168.

Hutchinson, J. R. 2002. The evolution of hind limb tendons and muscles on the line to crown-group birds. Comp. Biochem. Physiol. A 133:1051–1086.

Hutchinson, J. R., and S. M. Gatesy. 2000. Adductors, abductors, and the evolution of archosaur locomotion. Paleobiology 26:734–751.

Hwang, S. H., M. A. Norell, Q. Ji, and K. Gao. 2002. New Specimens of *Microraptor zhauianus* (Theropoda: Dromaeosauridae) from Northeastern China. Am. Mus. Novit. 3381:1–44.

Ji, Q., L. M. Chiappe, and S. a. Ji. 1999a. A new late Mesozoic confuciusornithid bird from China. J. Vert. Paleo. 19:1–7.

Ji, Q., Z.-X. Luo, and S. a. Ji. 1999b. A Chinese triconodont mammal and mosaic evolution of the mammalian skeleton. Nature 398:326–330.

Kass, R. E., and A. E. Raftery. 1995. Bayes factors. J. Am. Stat. Assoc. 90:773–795.

Kearney, M., and B. L. Stuart. 2004. Repeated evolution of limblessness and digging heads in worm lizards revealed by DNA from old bones. Proc. R. Soc. Lond. B 271:1677–1683.

Kurochkin, E. N. 1988. Cretaceous birds of Mongolia and their significance for study of phylogeny of class Aves. Trans. Joint Sov.-Mongol. Paleo. Exped. 34:33–42.

Lacasa-Ruiz, A. 1989. Nuevo genero de ave fosil del Yacimiento Neocomiense del Montsec (Provincia de Lerida, España). Estud. Geol. 45:417–425.

Lartillot, N., and H. Philippe. 2006. Computing Bayes factors using thermodynamic integration. Syst. Biol. 55:195–207.

Lee, M. S. Y. 2005. Molecular evidence and marine snake origins. Biol. Lett. 1:227–230.

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

Li, S. 1996. Phylogenetic tree construction using Markov chain Monte Carlo. PhD dissertation, Ohio State University, Columbus.

Livezey, B. C., and R. L. Zusi. 2006. Higher-order phylogenetics of modern birds (Theropoda, Aves: Neornithes) based on comparative anatomy. I. Methods and characters. Bull. Carnegie Mus. Nat. Hist. 37:1–544.

Maddison, W. P., and D. R. Maddison. 1997. MacClade: Analysis of phylogeny and character evolution. Sinauer Associates, Sunderland, Massachusetts.

Marsh, O. C. 1877. Characters of the Odontornithes, with notice of a new allied genus. Am. J. Sci. (3rd ser.) 14:85–87.

Marsh, O. C. 1880. Odontornithes: a monograph on the extinct toothed birds of North America. U. S. Government Printing Office, Washington, DC.

Marshall, D. C., C. Simon, and T. R. Buckley. 2006. Accurate branch length estimation in partitioned Bayesian analyses requires accommodation of among-partition rate variation and attention to branch length priors. Syst. Biol. 55:993–1003.

Mau, B. 1996. Bayesian phylogenetic inference via Markov chain Monte Carlo. PhD dissertation, University of Wisconsin, Madison.

May, G., F. Shaw, H. Badrane, and X. Vekemans. 1999. The signature of balancing selection: Fungal mating compatibility gene evolution. Proc. Natl. Acad. Sci. USA 96:9172–9177.

Müller, J., and R. R. Reisz. 2006. The phylogeny of early eureptiles: Comparing parsimony and Bayesian approaches in the investigation of a basal fossil clade. Syst. Biol. 55:503–511.

Nemeschkal, H. L., R. Van Den Elzen, and H. Brieschke. 1992. The morphometric extraction of character complexes accomplishing common biological roles: Avian skeletons as a case study. Z. Zool. Syst. Evol. 30:201–219.

Newton, M. A., and A. E. Raftery. 1994. Approximate Bayesian inference by the weighted likelihood bootstrap (with discussion). J. R. Stat. Soc. B 56:3–48.

Norell, M. A., and J. A. Clarke. 2001. Fossil that fills a critical gap in avian evolution. Nature 409:181–184.

Novas, F. E., and D. Pol. 2005. New evidence on deinonychosaurian dinosaurs from the Late Cretaceous of Patagonia. Nature 433:858–861.

Nylander, J. A. A., F. Ronquist, J. P. Huelsenbeck, and J. L. Nieves-Aldrey. 2004. Bayesian phylogenetic analysis of combined data. Syst. Biol. 53:47–67.

Ostrom, J. H. 1976. Some hypothetical anatomical stages in the evolution of avian flight. Smiths. Cont. Paleo. 27:1–21.

Ostrom, J. H., S. O. Poore, and G. E. Goslow, Jr. 1999. Humeral rotation and wrist supination: Important functional complex for the evolution of powered flight in birds? Smiths. Cont. Paleo. 89:301–309.

Padian, K. 2001. Cross-testing adaptive hypotheses: phylogenetic analysis and the origin of bird flight. Am. Zool. 41:598–607.

Poore, S. O., A. Sánchez-Haiman, and G. E. Goslow, Jr. 1997. Wing upstroke and the evolution of flapping flight. Nature 387:799–802.

R Development Core Team. 2007. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. http://www.R-project.org.

Rae, T. C. 1999. Mosaic evolution in the origin of the Hominoidea. Folia Primatol. 70:125–135.

Rannala, B., and Z. Yang. 1996. Probability distribution of molecular evolutionary trees: A new method of phylogenetic inference. J. Mol. Evol. 43:304–311.

Roff, D. A. 2006. Introduction to computer-intensive methods of data analysis in biology. Cambridge University Press, Cambridge, UK.

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

Sanz, J. L., and J. F. Bonaparte. 1992. A new order of birds (Class Aves) from the Lower Cretaceous of Spain. Nat. Hist. Mus. Los Angeles Co. Sci. Ser. 36:39–49.

Sanz, J. L., L. M. Chiappe, B. P. Pérez-Moreno, A. D. Buscalioni, J. J. Moratalla, F. Ortega, and F. J. Poyato-Ariza. 1996. An Early Cretaceous bird from Spain and its implications for the evolution of avian flight. Nature 382:442–445.

Swofford, D. L. 2002. PAUP*: Phylogenetic analysis using parsimony (*and other methods). Version 4.0. Sinauer Associates, Sunderland, Massachusetts.

Thompson, D. W. 1917. On growth and form, reprinted 1961 edition. Cambridge University Press, Cambridge, UK.

Tobalske, B. W. and K. P. Dial. 2007. Aerodynamics of wing-assisted incline running in birds. J. Exp. Biol. 210:1742–1751.

Venables, W. N. and B. D. Ripley. 2002. Modern applied statistics with S, 4th edition. Springer, New York.

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

Wiens, J. J., R. M. Bonett, and P. T. Chippindale. 2005. Ontogeny discombobulates phylogeny: Paedomorphosis and higher-level salamander relationships. Syst. Biol. 54:91–110.

Yang, Z. 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Approximate methods. J. Mol. Evol. 39:306–314.

Yokoyama, S., and R. Yokoyama. 1989. Molecular evolution of human visual pigment genes. Mol. Biol. Evol. 6:186–197.

Zhou, Z. 2004. The origin and early evolution of birds: Discoveries, disputes, and perspectives from fossil evidence. Naturwissenschaften 91:455–471.

Zhou, Z., and F. Zhang. 2001. Two new ornithurine birds from the Early Cretaceous of western Liaoning, China. Chin. Sci. Bull. 46:1258–1264.

Zhou, Z., and F. Zhang. 2002. Largest bird from the Early Cretaceous and its implications for the earliest avian ecological diversification. Naturwissenschaften 89:34–38.