Bayesian tip dating reveals heterogeneous morphological clocks in Mesozoic birds

Chi Zhang[1,2,*] and Min Wang[1,2]

[1]Key Laboratory of Vertebrate Evolution and Human Origins, Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences, Beijing 100044, China

[2]Center for Excellence in Life and Paleoenvironment, Chinese Academy of Sciences, Beijing 100044, China

∗Corresponding author: E-mail: zhangchi@ivpp.ac.cn

June 19, 2018

Keywords: Mesozoic birds, tip dating, morphological clock

Birds are one of the most speciose (over 10,000 recognized species) and ecological diverse living vertebrates (Gill 2007). It has been held that birds underwent two large-scale radiations during their 160 million years evolution, one for the stem groups in the Cretaceous, and the other for the crown groups in the Paleogene (Mayr 2009; Prum et al. 2015; Wang and Zhonghe 2017). Over the last few years, numerous well-preserved Mesozoic bird fossils have been described (Chiappe and Meng 2016; Wang and Zhonghe 2017), and consensus has been approached regarding their systematic relationships (O'Connor and Zhou 2013; Wang et al. 2017, 2018). These wealthy data have significantly bridged the large morphological gap between birds and their non-avian theropod predecessors (O'Connor and Zhou 2015; Wang and Zhonghe 2017), and demonstrated their evolutionary success related to novel traits. Expanded morphological characters that cover the major Mesozoic avian groups with chronological data become accessibly recently (Wang and Lloyd 2016), making it possible to trace the early avian evolution more quantitively by addressing important questions such as

the divergence times of the major clades and comparative study of the pattern of morphological character changes in and between lineages (Brusatte et al. 2014; Lloyd 2016).

In this study, we reinvestigate the dataset which has been analyzed previously to explore evolutionary rate heterogeneity under maximum parsimony (Wang and Lloyd 2016). The approach was stepwise: first to infer the most parsimonious trees, then to inform the internal node ages using certain *ad hoc* measures, and last to obtain the rates from dividing the number of parsimonious changes by the time durations. It did not account for the uncertainties of tree topology, times and rates statistically, and was unable to model the evolutionary process explicitly. Furthermore, we expand the original data with 18 morphological characters and ten Mesozoic birds (Wang and Zhou, in press, see also Supplementary Information). The dataset now consists of 280 morphological characters and 68 operational taxonomic units (Dromaeosauridae as the outgroup, 65 Mesozoic and 2 extant birds *Anas* and *Gallus*).

Since morphological data only provide information of evolutionary distances among taxa in unit of substitutions per character, which are the product of geological time duration (in myr) and evolutionary rate (in substitutions per character per myr), we need to incorporate the stratigraphic age of each fossil and assume a relaxed clock model (Thorne et al. 1998; Huelsenbeck et al. 2000; Yoder and Yang 2000; Kishino et al. 2001; Thorne and Kishino 2002; Drummond et al. 2006; Rannala and Yang 2007; Lepage et al. 2007; Drummond and Suchard 2010) to date the phylogeny and co-estimate the times and rates. This technique was originally developed for analyses combining both morphological and molecular data and has been termed as “total-evidence dating” (Pyron 2011; Ronquist et al. 2012; 2016; Zhang et al. 2016; Lee 2016; Gavryushkina et al. 2016). Obviously, it is also applicable to morphological data only (Lee et al. 2014; King et al. 2017) and we use “tip dating” for that matter. The major difference from another technique termed as “node dating” (Tavaré et al. 2002; Yang and Rannala 2006; Ho and Phillips 2009; Heath 2012) is that ages are directly associated with the fossil tips instead of calibrating one or several internal nodes in the tree of extant taxa. As all but two of the taxa are fossils, tip dating naturally fits to our data while node dating is not adequate.

Here, we exploit the advantage of tip dating to infer both the divergence times and evolutionary rates while accounting for their uncertainties in a coherent Bayesian statistical

framework. The inference is based on the posterior probability distribution of the model parameters, which combines the information from the data (morphological characters) and the priors (including the distributions of the fossil ages and the other parameters in the tree model and clock model). We further investigate the evolutionary rate heterogeneity across different parts of bird body by subdividing the characters into six categories.

Material and methods

Tree model

We use the fossilized birth-death process (Stadler 2010; Heath et al. 2014; Gavryushkina et al. 2014; Zhang et al. 2016) to model macroevolution represented by the time tree. The process starts at the root with two lineages sharing the same origin. Each lineage bifurcates with a constant rate λ and goes extinct with a constant rate μ. Fossils are sampled along the lineages with a constant rate ψ. We further assume that the lineage is dead upon sampling and will not produce any decedent, meaning all fossils are at the tips of the tree. The process ends at the present (Fig. 1).

Let m be the number of tips (including extinct and extant), then $m - 1$ is the number of internal nodes. Let $\boldsymbol{x} = (x_1, \ldots, x_{m-1})$ be the vector of bifurcation ages ($x_1 > \ldots > x_{m-1}$) and $\boldsymbol{y} = (y_1, \ldots, y_m)$ be the vector of tip ages ($y_1 > \ldots \geq y_m$) ($\boldsymbol{x}$ and $\boldsymbol{y}$ are in unit of myr). Conditioned on the root age ($t_{\mathrm{mrca}} = x_1$) and sampling at least one individual in both subtrees leading to the root, the joint probability density of the tree topology and node ages ($\mathcal{T}$) is,

$$f(\mathcal{T}|t_{\mathrm{mrca}} = x_1, \lambda, \mu, \psi) = \frac{p_1^2(x_1)}{m![1-p_0(x_1)]^2} \prod_{i=2}^{m-1} 2\lambda p_1(x_i) \prod_{i=1}^{m} \frac{\psi}{p_1(y_i)}, \quad (1)$$

where $p_0(t)$ is the probability that an individual at time t has no sampled descendants, and $p_1(t)$ is the probability that an individual at time t has one sampled extant and no extinct descendants, so that

$$p_0(t) = \frac{\lambda+\mu+\psi+a_1\frac{e^{-a_1 t}(1-a_2)-(1+a_2)}{e^{-a_1 t}(1-a_2)+(1+a_2)}}{2\lambda}, \quad (2)$$

$$p_1(t) = \frac{4}{2(1-a_2^2)+e^{-a_1 t}(1-a_2)^2+e^{a_1 t}(1+a_2)^2}, \quad (3)$$

where $a_1 = \left|\sqrt{(\lambda-\mu-\psi)^2+4\lambda\psi}\right|$ and $a_2 = -\frac{\lambda-\mu-\psi}{a_1}$.

The probability density here in equation 1 is analogous to that in (Gavryushkina et al. 2014) where the process is conditioned on the origin and the removal probability upon sampling can be less than 1.

Clock model

To investigate the rate heterogeneity of morphological evolution, we apply the independent gamma (white noise) relaxed clock model (Lepage et al. 2007) to accommodate the rate variation across branches. The model assumes that the substitution rate (clock rate, in unit of substitutions per character per myr) of branch i in the tree, c_i, is a product of the mean rate c and the relative rate r_i (i.e., $c_i = cr_i$), and r_i is gamma distributed with mean 1.0 and variance $\sigma/(t_i c)$, where t_i is the geological time duration of branch i (in myr). This is a reparameterization of the original model (Lepage et al. 2007) aiming to articulate the relative rates. The branch rates ($c_1, \ldots, c_{2m-2}$) are independent of each other with the same mean c. The relative rate r_i serves as a multiplier to the mean rate. Large deviation of r_i from 1.0 indicates severe heterogeneity of the morphological clock, while r_i's all being similar to 1.0 models a somewhat strict clock. The branch length (distance, in unit of substitutions per character) b_i is the product of time duration t_i and clock rate c_i (i.e., $b_i = t_i c_i$; $i = 1, \ldots, 2m - 2$).

Tip dating

In Bayesian tip dating framework, we infer the posterior distribution of the model parameters, which is proportional to the likelihood and priors,

$$f(\mathcal{T}, \lambda, \mu, \psi, c, \sigma, \alpha | \text{data}) \propto f(\text{data}|\mathcal{T}, c, \sigma, \alpha) f(\mathcal{T}|\lambda, \mu, \psi) f(\lambda, \mu, \psi) f(c, \sigma, \alpha). \quad (4)$$

For the likelihood, the Mkv model (Lewis 2001) is used for the character state substitution with gamma rate variation across characters. The gamma distribution has mean 1 ($\alpha = \beta$) and is approximated with 4 discrete categories (Yang 1994). 33 characters are defined as ordered (Supplementary Information), which means instantaneous change is only allowed between adjacent states, the rest of the characters are thus unordered.

We also need to assign prior distributions for the time tree ($\mathcal{T}$) with root age (t_{mrca}) and fossil tip ages, the rates of speciation (λ), extinction (μ) and sampling (ψ), the mean rate (c)

and variance parameter (σ) of the morphological clock, and the gamma shape (α) of character rate variation. The fossilized birth-death process described above gives rise to the prior for the time tree ($\mathcal{T}$). The age of each fossil bird is assigned a uniform prior with lower and upper bounds from the corresponding stratigraphic range (Supplementary Information). The root age is assigned a truncated exponential prior with mean 169 (slightly older than the first appearance datum of Dromaeosauridae) and minimum 153 (slightly older than the first appearance datum of *Archaeopteryx*). For convenience, we reparametrize the speciation, extinction and sampling rates as $d = \lambda - \mu - \psi$, $r = (\mu + \psi)/\lambda$ and $s = \psi/(\mu + \psi)$, so that the latter two parameters range from 0 to 1. The prior for d is exponential(100) with mean 0.01, and uniform(0, 1) for r and s. The prior used for the mean clock rate c is gamma(2, 200) with mean 0.01 and standard deviation 0.007, and that for σ is exponential(10). Lastly, the prior for gamma shape α is exponential(1).

In order to root the tree properly and infer the rates more reliably, we apply five topology constraints as Aves, Pygostylia, Ornithothoraces, Enantiornithines and Ornithuromorpha (Fig. 2), each of which forms a monophyletic clade. The posterior distribution is estimated using Markov chain Monte Carlo (MCMC). We execute two independent runs with four chains (one cold and three hot) per run for 10 million iterations and sample every 1000 iterations. The first 25% samples are discarded as burn-in for each run, and the remaining samples from the two runs are combined.

Partitioned analysis

We further investigate whether different body parts evolve at different rates across the major transitions such as the branches subtending to Pygostylia, Ornithothoraces, Enantiornithes and Ornithuromorpha. Enantiornithes and Ornithuromorpha, as sister groups, are the most successful clades of Mesozoic birds that demonstrate very different bauplan. As such, they provide a unique opportunity to examine evolutionary rate heterogeneity between phylogenetically close clades. The data is partitioned into six categories: skull (53 characters), axial skeleton (36 characters), pectoral girdle and sternum (48 characters), forelimb (65 characters), pelvic girdle (23 characters), and hindlimb (55 characters).

We unlink the clock model such that the substitution rate (clock rate) of each partition varies along the branches in the tree with its own variance parameter but share the same mean rate. Specifically, let r_{ij} be the relative clock rate of branch i in partition j, and r_{ij} is gamma distributed with mean 1.0 and variance $\sigma_j/(t_i c)$ ($i = 1, \ldots, 2m - 2, j = 1, \ldots, n$). Thus, the clock model now has $n + 1$ parameters ($c, \sigma_1, \ldots, \sigma_n$) instead of two ($c$ and σ), and there are $(2m - 2) \times n$ independent rates in the tree.

In the MCMC, the chain length is doubled to 20 million iterations to obtain sufficient effective sample size (ESS). The prior for σ_j is exponential (10). The other settings are unchanged.

Software availability

The models described above are implemented in MrBayes version 3.2.7 (Ronquist et al. 2012, https://github.com/NBISweden/MrBayes; last accessed June 19, 2018). The commands carrying out the analyses are in Supplementary Information.

Results and Discussion

The phylogeny estimated from tip dating without partitioning the data is shown in Figure 2. The tree is well resolved, with a few polytomies within Enantiornithines and Ornithuromorpha representing the uncertainty of the taxa relationships (rather than simultaneous speciation). The root age is estimated at 164.33 (154.47, 174.05) Ma (see also Table 1), which is consistent with the fixed age of 168.7 Ma in the previous study (1 myr older than the first appearance datum of Dromaeosauridae using the minimum branch length method, Wang and Lloyd 2016). The posterior age of Dromaeosauridae is 156.01 (142.55, 167.67) Ma, mainly within Upper Jurassic, while the prior range expands the whole Cretaceous (66.0, 167.7). As the mean relative clock rate at the branch of Dromaeosauridae is 1.17 (closed to 1.0), the similarity of the morphological characters informs a short time span. The mean ages of the divergences of Pygostylia, Ornithothoraces and Enantiornithines are about 7–9 myr older than the fixed ages in the previous study (Wang and Lloyd 2016). We emphasize that the age estimates using tip dating integrates all available sources of

information but the minimum branch length method only used the first appearance datum of the oldest taxa thus might underestimate the ages.

The mean clock rate (c) is estimated as 0.011 (0.0086, 0.014) substitutions per character per myr (i.e., approximately one character state change per 100 million years) (Table 1). The relative clock rate at each branch (r_i, with mean 1.0) represents the deviation from the mean clock rate and is also shown in Figure 2. We observe extremely high rates during early avian evolution. The relative rates of the branches subtending to Aves and to Pygostylia are 6.57 (1.18, 15.73) and 5.17 (0.00, 15.52) respectively, and are even higher at the branch subtending to Ornithothoraces as 11.92 (0.60, 36.58) and to Enantiornithines as 11.78 (1.57, 33.88) which are one order of magnitude higher than 1.0. The rates are also slightly high at the beginning of the Ornithuromorpha clade leading to the extant birds. These observations confirm the fast radiation of the early bird evolution in the Jurassic and Cretaceous.

However, the analysis above only inferred the branch rates averaged across all morphological characters. Further partitioning the data into six anatomical regions make it feasible for us to estimate refined evolutionary rates along branches. Different partitions have their own independent rate variations while sharing the same tree topology and geological time duration (i.e., the same time tree). In this case, the major clades inferred in the tree are unchanged, although a few taxa with large uncertainties shuffle a bit (Supplementary Figs S1–S6). The node ages closed to the root become slightly older, e.g., the root age is 171.87 (162.38, 181.40) Ma and the age of Aves is 168.17 (158.45, 177.85) Ma, while the younger nodes have similar ages as in Figure 2.

The rates of morphological character changes of the six body parts are quite heterogeneous during early avian evolution (Fig. 3), although the mean rate (c) estimated is almost identical as before (Table 1). The branch subtending to Pygostylia shows extremely high rate in the axial skeleton (see also Supplementary Fig. S2), which is one order of magnitude higher than in the rest five body parts. Clearly, the high rate observed here indicates extensive morphological changes in the vertebral column, and the most recognizable change is that a long tail consisting of over 20 caudal vertebrae in *Archaeopteryx* and *Jeholornis* is replaced by a short element named pygostyle (Wang and Zhou 2017). The transition from long to short tail is functionally important for the evolution of powered flight in birds: a short tail could

forward the gravitational center, and with attached feathers become indispensable for the avian flight apparatus (Gatesy and Dial 1996).

The branches subtending to Ornithothoraces and Enantiornithines show rapid morphological changes in the pectoral girdle and sternum, evidenced by the one order of magnitude higher relative rates along the corresponding branches. They also show relatively high rates of change in forelimb. In comparison, the corresponding rates in other body parts are closed to 1.0 with slight variation (Fig. 3, Supplementary Figs S3&S4). The relative rates of the six body parts at the branch subtending to Ornithuromorpha are more or less the same. These results suggest that most of the changes relate to the modification of the shoulder and forelimb close the origin of Ornithothoraces, e.g., ossified sternum with keel, further elongated forelimb (O'Connor and Zhou 2015), contributing to the refinement of the flight capability. The rapid morphological changes at the Enantiornithes branch correspond to their unique shoulder morphology relative to the common ancestor of the Ornithothoraces, such as the sternum with a caudally restricted keel, and the acromion of the scapula elongate (Chiappe and Walker 2002; Wang and Zhonghe 2017). Precious study suggested that Enantiornithes have different flight style compared with other basal birds using limb proportion (Dyke and Nudds 2009). The rest of the relatively slow and even rates suggest that there is no dominant selective pressure in favoring of modifications in certain parts of the body.

When we compare the evolutionary rates of different body parts between Enantiornithes and Ornithuromorpha (two sister groups) by summarizing the mean relative rates within each clade (excluding *Anas*, *Gallus*, and their common ancestral branch), the differences are not striking, with medians uniformly close to 1.0 (Fig. 4). However, significantly high rates (outliers) are detected along some basal branches within these two clades, and a slowdown in more derived branches (Fig. 4, Supplementary Figs S1–S6), suggesting adaptive radiation for these two avian groups as proposed in previous study (Wang and Lloyd 2016).

The gamma shape parameter (α) of character rate variation is larger than 1.0 (Table 1), indicating that the evolutionary rates are fairly homogeneous across characters. Note this gamma distribution models rate variation across characters and is independent of the gamma distribution modeling rate variation across branches (relaxed clock) in the tree.

Partitioning the data is a common practice in molecular phylogenetic analyses (Nylander et al. 2004; Brown and Lemmon 2007). The different partitions may correspond to different genes and may also correspond to different codon positions in a protein-coding gene. Nowadays, generating genomic-scale sequences is effortless due to fast and automatic sequencing machines, however, coding morphological characters still heavily relies on experienced paleontologists. Thus, morphological characters are typically treated as a single partition unless sufficient characters are available (Lee 2016). While successfully demonstrating the different evolutionary dynamics modeled by independent rate parameters in the six partitions, we note that the variance (width of the error bar) is very large (Fig. 3) due to limited number of characters in each partition. Further effort of coding more characters would refine the resolution.

In summary, the Bayesian tip dating approach implemented in MrBayes is a powerful and flexible tool to simultaneously estimate the tree topology, divergence times, evolutionary rates, and the other parameters of interest while accounting for their uncertainties. In the priors, we are able to incorporate the uncertainty of each fossil age, model the speciation, extinction, fossilization and sampling process explicitly, and take advantage of the relaxed clock model to investigate rate variation cross branches and partitions. It is feasible to integrate all available sources of information in the analysis, rather than discarding certain information or uncertainties in the parsimony and stepwise approach. Tip dating is a general framework applicable to wide range of taxa with potential future extensions to the theoretical model and practical implementation.

Acknowledgments

This research is supported by the 100 Young Talents Program of Chinese Academy of Sciences (to C.Z.). M.W. is supported by National Natural Science Foundation of China (41722202).

References

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

Brusatte S.L., Lloyd G.T., Wang S.C., Norell M.A. 2014. Gradual assembly of avian body plan culminated in rapid rates of evolution across the dinosaur-bird transition. Curr. Biol. 24:2386–2392.

Chiappe L.M., Meng Q.J. 2016. Birds of Stone: Chinese Avian Fossils from the Age of Dinosaurs. Johns Hopkins University Press.

Chiappe L.M., Walker C.A. 2002. Skeletal morphology and systematics of the Cretaceous Euenantiornithes (Ornithothoraces: Enantiornithes). In: Chiappe L.M., Witmer L.M. editors. Mesozoic birds: above the heads of dinosaurs. California: University of California Press. 240–267.

Drummond A.J., Ho S.Y.W., Phillips M.J., Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol. 4:e88.

Drummond A.J., Suchard M.A. 2010. Bayesian random local clocks, or one rate to rule them all. BMC Biology. 8:114.

Dyke G.J., Nudds R.L. 2009. The fossil record and limb disparity of enantiornithines, the dominant flying birds of the Cretaceous. Lethaia 42:248–254.

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

Gavryushkina A., Heath T.A., Ksepka D.T., Stadler T., Welch D., Drummond A.J. 2016. Bayesian total-evidence dating reveals the recent crown radiation of penguins. Syst. Biol. 66:57–73.

Gavryushkina A., Welch D., Stadler T., Drummond A.J. 2014. Bayesian inference of sampled ancestor trees for epidemiology and fossil calibration. PLoS Comput. Biol. 10:e1003919.

Gill FB. 2007. Ornithology. 3rd edition. New York, W.H. Freeman.

Heath T.A. 2012. A hierarchical Bayesian model for calibrating estimates of species divergence times. Syst. Biol. 61:793–809.

Heath T.A., Huelsenbeck J.P., Stadler T. 2014. The fossilized birth-death process for coherent

calibration of divergence-time estimates. Proc. Natl. Acad. Sci. USA. 111:E2957–66.

Ho S.Y.W., Phillips M.J. 2009. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. Syst. Biol. 58:367–380.

Huelsenbeck J.P., Larget B., Swofford D.L. 2000. A compound poisson process for relaxing the molecular clock. Genetics. 154:1879–1892.

King B., Qiao T., Lee M.S.Y., Zhu M., Long J.A. 2017. Bayesian Morphological Clock Methods Resurrect Placoderm Monophyly and Reveal Rapid Early Evolution in Jawed Vertebrates. Syst. Biol. 66:499–516.

Kishino H., Thorne J.L., Bruno W.J. 2001. Performance of a divergence time estimation method under a probabilistic model of rate evolution. Mol. Biol. Evol. 18:352–361.

Lee M.S.Y. 2016. Multiple morphological clocks and total-evidence tip-dating in mammals. Biol. Lett. 12:20160033.

Lee M.S.Y., Cau A., Naish D., Dyke G.J. 2014. Morphological clocks in paleontology, and a mid-Cretaceous origin of crown Aves. Syst. Biol. 63:442–449.

Lepage T., Bryant D., Philippe H., Lartillot N. 2007. A general comparison of relaxed molecular clock models. Mol. Biol. Evol. 24:2669–2680.

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

Lloyd G.T. 2016. Estimating morphological diversity and tempo with discrete character-taxon matrices: implementation, challenges, progress, and future directions. Biol. J. Linn. Soc. 118:131–151.

Mayr G. 2009. Paleogene fossil birds. Berlin, Springer Berlin Heidelberg.

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

O'Connor J.K., Zhou Z. 2015. Early evolution of the biological bird: perspectives from new fossil discoveries in China. J. Ornithol. 156:333–342.

O'Connor J.K., Zhou Z. 2013. A redescription of *Chaoyangia beishanensis* (Aves) and a comprehensive phylogeny of Mesozoic birds. J. Syst. Palaeontol. 11:889–906.

Prum R.O., Berv J.S., Dornburg A., Field D.J., Townsend J.P., Lemmon E.M., Lemmon A.R. 2015. A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA

sequencing. Nature 526:569–573.

Pyron R.A. 2011. Divergence time estimation using fossils as terminal taxa and the origins of Lissamphibia. Syst. Biol. 60:466–481.

Rannala B., Yang Z. 2007. Inferring speciation times under an episodic molecular clock. Syst. Biol. 56:453–466.

Ronquist, F. et al. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542.

Ronquist F., Klopfstein S., Vilhelmsen L., Schulmeister S., Murray D.L., Rasnitsyn A.P. 2012. A total-evidence approach to dating with fossils, applied to the early radiation of the Hymenoptera. Syst. Biol. 61:973–999.

Ronquist F., Lartillot N., Phillips M.J. 2016. Closing the gap between rocks and clocks using total-evidence dating. Philos. Trans. R. Soc. Lond. B Biol. Sci. 371.

Stadler T. 2010. Sampling-through-time in birth-death trees. J. Theor. Biol. 267:396–404.

Tavaré S., Marshall C.R., Will O., Soligo C., Martin R.D. 2002. Using the fossil record to estimate the age of the last common ancestor of extant primates. Nature. 416:726–729.

Thorne J.L., Kishino H. 2002. Divergence time and evolutionary rate estimation with multilocus data. Syst. Biol. 51:689–702.

Thorne J.L., Kishino H., Painter I.S. 1998. Estimating the rate of evolution of the rate of molecular evolution. Mol. Biol. Evol. 15:1647–1657.

Wang M., Lloyd G.T. 2016. Rates of morphological evolution are heterogeneous in Early Cretaceous birds. Proc. Biol. Sci. 283:20160214.

Wang M., O'Connor J.K., Pan Y., Zhou Z. 2017. A bizarre Early Cretaceous enantiornithine bird with unique crural feathers and an ornithuromorph plough-shaped pygostyle. Nat. Commun. 8:14141.

Wang M., O'Connor J.K., Zhou Z. 2018. A taxonomical revision of the Confuciusornithiformes (Aves: Pygostylia). Vertebrate PalAsiatica.

Wang M., Zhonghe Z. 2017. The evolution of birds with implications from new fossil evidences. In: Maina N.J. editor. The biology of the avian respiratory system, Springer International Publishing. 1–26.

Yang Z. 1994. Maximum likelihood phylogenetic estimation from DNA sequences with

variable rates over sites: approximate methods. J. Mol. Evol. 39:306–314.

Yang Z., Rannala B. 2006. Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. Mol. Biol. Evol. 23:212–226.

Yoder A.D., Yang Z. 2000. Estimation of primate speciation dates using local molecular clocks. Mol. Biol. Evol. 17:1081–1090.

Zhang C., Stadler T., Klopfstein S., Heath T.A., Ronquist F. 2016. Total-evidence dating under the fossilized birth-death process. Syst. Biol. 65:228–249.

Table 1. Posterior distributions (mean and 95% HPD interval) of the model parameters

	single partition	six partitions
α	1.843 (1.382, 2.356)	1.809 (1.343, 2.306)
$d = \lambda - \mu - \psi$	0.00651 (0.000085, 0.0150)	0.00622 (0.000211, 0.0142)
$r = (\mu + \psi)/\lambda$	0.927 (0.829, 0.999)	0.925 (0.823, 0.998)
$s = \psi/(\mu + \psi)$	0.614 (0.172, 0.999)	0.603 (0.151, 0.999)
t_{mrca}	164.330 (154.467, 174.054)	171.865 (162.382, 181.404)
c	0.0111 (0.00856, 0.0140)	0.0105 (0.00825, 0.0130)
σ	0.0604 (0.0373, 0.0852)	0.0635 (0.0230, 0.109)
		0.0707 (0.0302, 0.114)
		0.0584 (0.0279, 0.0917)
		0.0564 (0.0273, 0.0889)
		0.0436 (0.00979, 0.0829)
		0.0629 (0.0287, 0.0995)

Figures

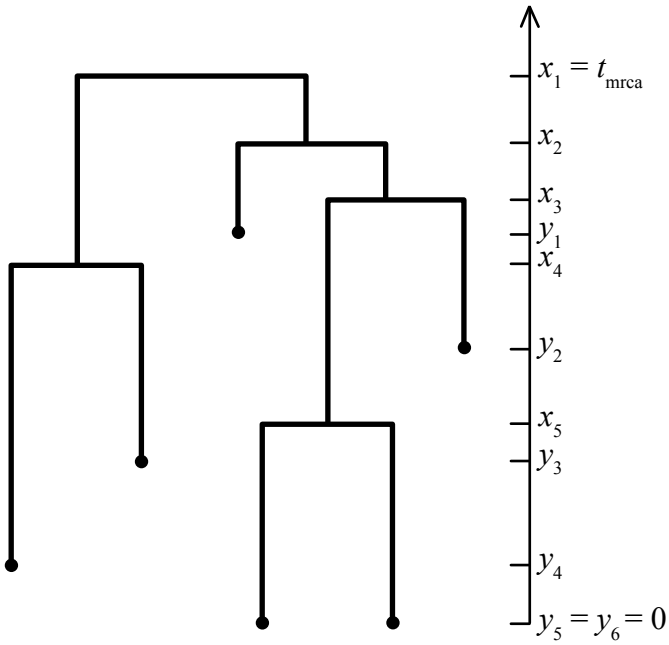

Figure 1. A sampled tree under the fossilized birth-death process. The bifurcation ages are x_1 to x_5 with $x_1 = t_{mrca}$ being the root age. The tip ages are y_1 to y_6.

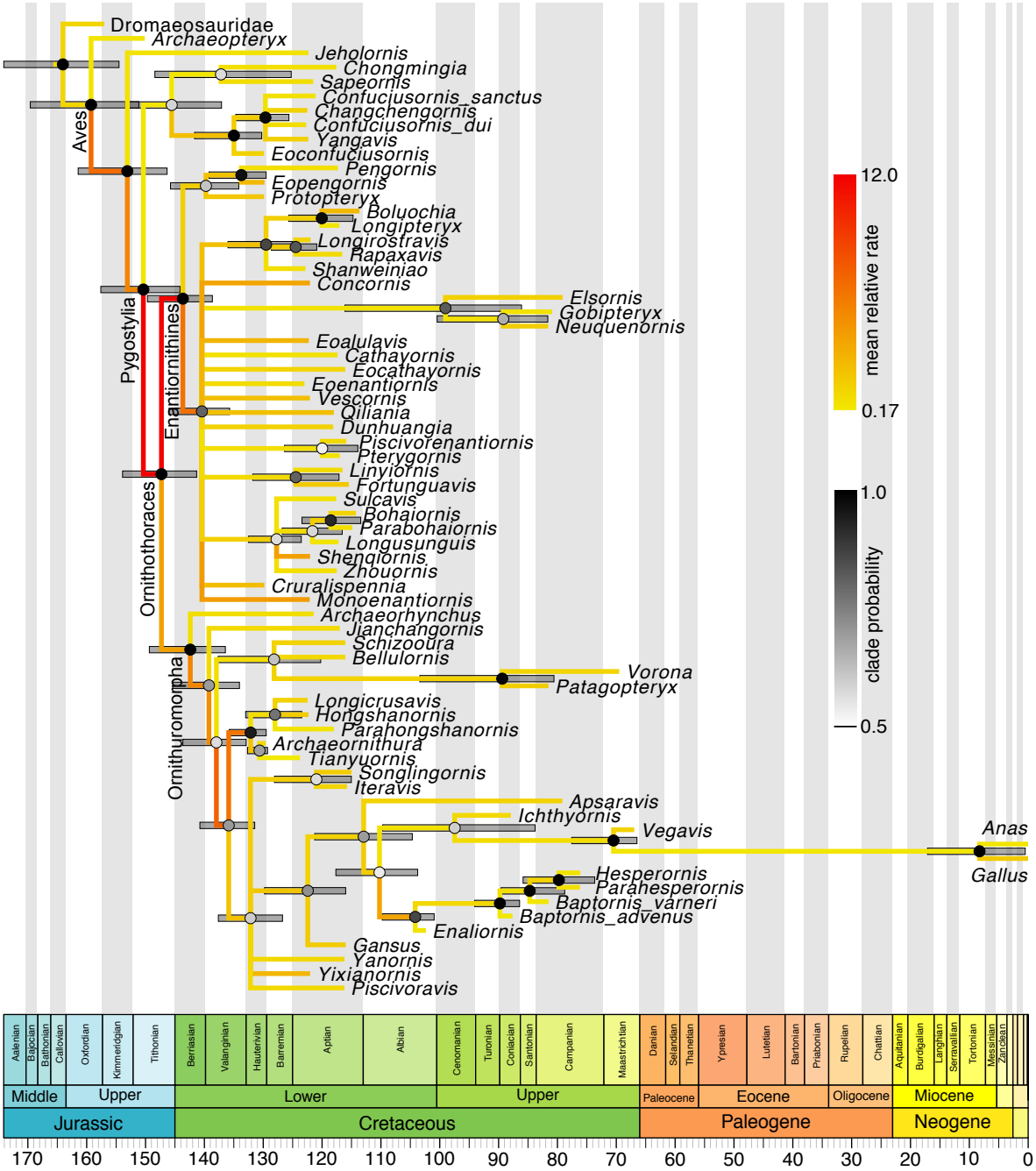

Figure 2. Dated phylogeny (time tree) of the Mesozoic birds. The node ages are the posterior medians and the error bars denote the 95% HPD intervals. The shade of the dot at an internal node represents the posterior probability of the corresponding clade. The color of the branch represents the mean relative clock rate at that branch.

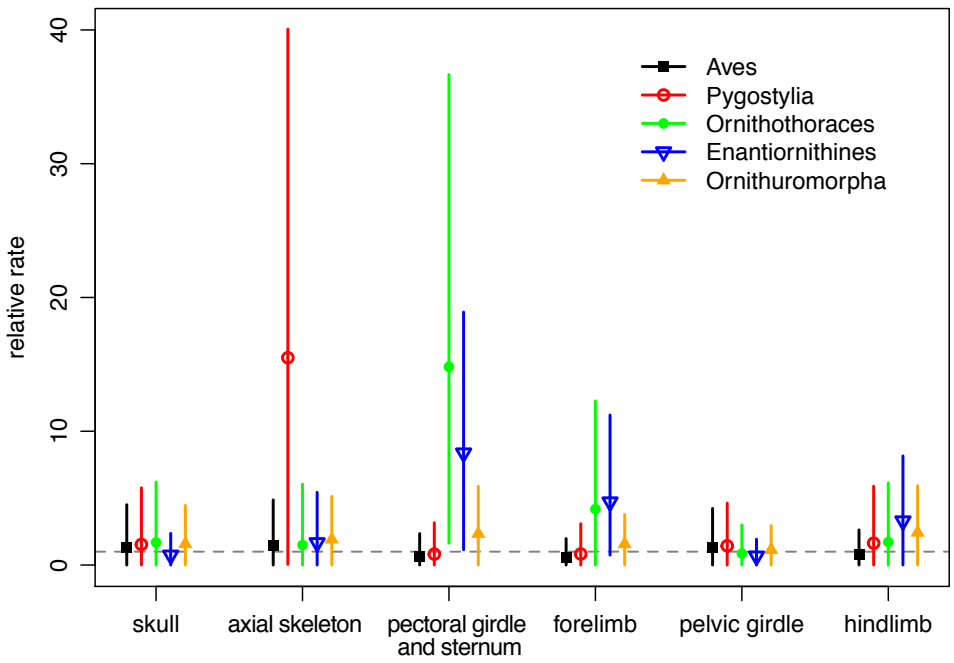

Figure 3. Posterior estimates of the relative clock rates along the branches subtending to the major transitions of early avian evolution for six parts of the bird body. The dot with error bar denote the mean and 95% HPD interval for each estimate. The horizontal dashed line indicates the mean relative rate of 1.0 in the relaxed model.

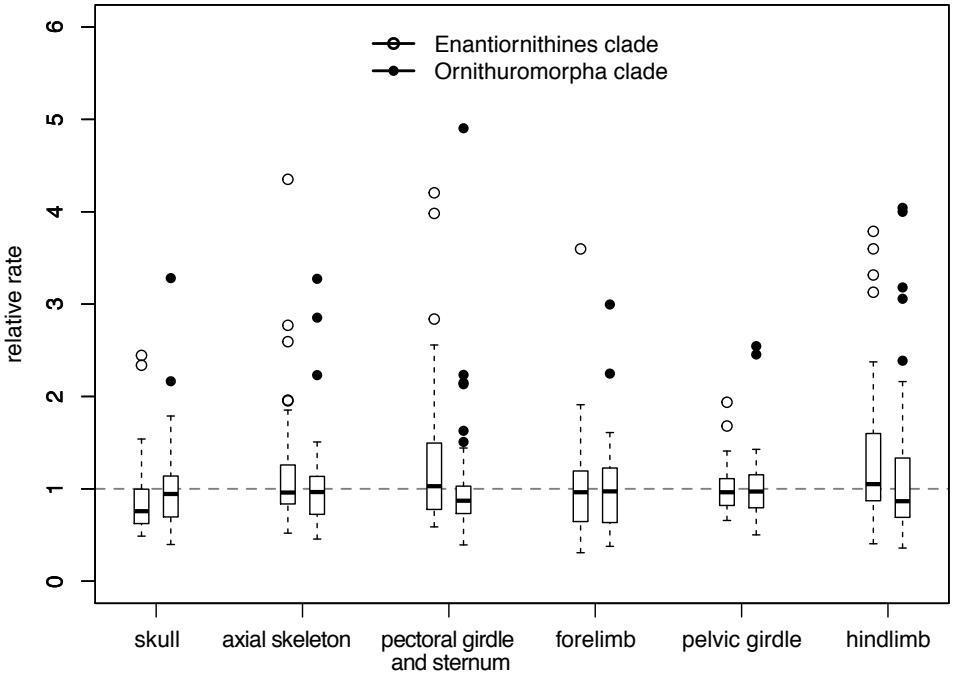

Figure 4. Boxplots summarizing the mean relative clock rates across branches in the Enantiornithines clade and Ornithuromorpha clade (excluding *Anas*, *Gallus*, and their common ancestral branch) respectively, for six parts of the bird body. The box denotes the 1st, 2nd (median) and 3rd quartiles while the dots are the outliers. The horizontal dashed line indicates the mean relative rate of 1.0 in the relaxed clock model.