PNAS RESEARCH ARTICLE | EVOLUTION OPEN ACCESS

Genomes, fossils, and the concurrent rise of modern birds and flowering plants in the Late Cretaceous

Shaoyuan Wu[a,1], Frank E. Rheindt[b], Jin Zhang[c], Jiajia Wang[a], Lei Zhang[a], Cheng Quan[d], Zhiheng Li[e], Min Wang[e], Feixiang Wu[e], Yanhua Qu[f,1], Scott V. Edwards[g,1], Zhonghe Zhou[e,1], and Liang Liu[h,1]

Edited by Douglas Futuyma, Stony Brook University, Stony Brook, NY; received November 17, 2023; accepted December 29, 2023

The phylogeny and divergence timing of the Neoavian radiation remain controversial despite recent progress. We analyzed the genomes of 124 species across all Neoavian orders, using data from 25,460 loci spanning four DNA classes, including 5,756 coding sequences, 12,449 conserved nonexonic elements, 4,871 introns, and 2,384 intergenic segments. We conducted a comprehensive sensitivity analysis to account for the heterogeneity across different DNA classes, leading to an optimal tree of Neoaves with high resolution. This phylogeny features a novel Neoavian dichotomy comprising two monophyletic clades: a previously recognized Telluraves (land birds) and a newly circumscribed Aquaterraves (waterbirds and relatives). Molecular dating analyses with 20 fossil calibrations indicate that the diversification of modern birds began in the Late Cretaceous and underwent a constant and steady radiation across the KPg boundary, concurrent with the rise of angiosperms as well as other major Cenozoic animal groups including placental and multituberculate mammals. The KPg catastrophe had a limited impact on avian evolution compared to the Paleocene–Eocene Thermal Maximum, which triggered a rapid diversification of seabirds. Our findings suggest that the evolution of modern birds followed a slow process of gradualism rather than a rapid process of punctuated equilibrium, with limited interruption by the KPg catastrophe. This study places bird evolution into a new context within vertebrates, with ramifications for the evolution of the Earth's biota.

Neoavian phylogeny | molecular clock | modern bird divergence | gradualism | punctuated evolution

Significance

Despite modern DNA advances, scientists still know little about how and when early bird groups evolved. Using new approaches to mine genomic information among 124 species covering most of modern bird diversity, we found that the main lineages of birds first divided into two groups: one mostly land-based and the other containing water-associated species. We demonstrate that modern birds date back further than previously assumed, much earlier than the dinosaurian extinction event, which seems to have had a limited impact on birds' evolution. Instead, a warming event around ~55 Mya appears to have triggered the diversification of modern seabirds. Our study indicates that the radiation of modern birds was in remarkable lockstep with that of flowering plants and other organisms.

The phylogenetic relationships among modern birds (crown Aves) have been intensively studied in the past decade (1–3). This research has corroborated the traditional basal division of modern birds into two monophyletic clades, Palaeognathae (flightless birds) and Neognathae (the majority of living birds), and the division of Neognathae into Galloanserae (landfowl and waterfowl) and all remaining birds (Neoaves, ~95% of living birds) (4, 5). Despite the increasing size of genomic datasets, however, phylogenetic relationships among Neoaves have remained in conflict across this new crop of genome-scale studies (2, 3). The intractability of early avian relationships has been attributed to a "Neoavian explosion" after the KPg boundary extinction event, when the demise of nonavian dinosaurs and other megafauna would have led to a global vacuum with subsequent ecological release, fueling the rapid diversification of Neoaves and other major animal lineages (2, 3).

Meanwhile, initial beliefs that more data, either in terms of genomic or taxon sampling, or both, will ultimately resolve phylogenetic uncertainties in the Tree of Life have not yet been borne out (6–10). The successful reconstruction of the Tree of Life is subject to multiple factors, such as comprehensive taxon sampling, rigorous locus sampling, sufficient coverage of different types of DNA loci, and the use of appropriate tree-building models. To resolve the controversies surrounding phylogenetic relationships and diversification timing of Neoaves, it is necessary to strike a balance among these critical factors (8–11).

In this study, we assembled genome-scale data from 118 species representing all 35 orders of Neognathae. We additionally sampled seven outgroup taxa, including six palaeognaths and the American alligator (Dataset S1). Our genomic data comprise four different classes of DNA loci—5,756 coding sequences (CDS), 12,449 conserved nonexonic elements (CNEE), 4,871 introns, and 2,384 intergenic segments, with a total of 20,652,290 aligned base pairs (*SI Appendix*, *Materials and Methods*). We applied NJst, a fast, accurate "two-step" species tree method based on the neighbor-joining algorithm and derived from the multispecies coalescent model, which can infer a species tree from unrooted gene trees without the need to commit to an outgroup and can handle missing data (12). Moreover, we applied concatenation methods rooted in maximum likelihood (ML) approaches as implemented in RAxML (13).

Author contributions: S.W., S.V.E., Z.Z., and L.L. designed research; S.W., L.Z., Y.Q., S.V.E., Z.Z., and L.L. performed research; Y.Q. contributed new reagents/analytic tools; S.W., F.E.R., J.Z., J.W., C.Q., Z.L., M.W., F.W., Y.Q., S.V.E., Z.Z., and L.L. analyzed data; S.W. conceived of the idea; F.E.R., S.V.E., Z.Z., and L.L. contributed to the writing and discussion; and S.W. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: shaoyuan5@gmail.com, quyh@ioz.ac.cn, sedwards@fas.harvard.edu, zhouzhonghe@ivpp.ac.cn, or lliu@uga.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2319696121/-/DCSupplemental.

Published February 12, 2024.

Results

Effect of DNA Classes on Tree Inference. We performed intensive sensitivity analysis using both NJst and RAxML to examine the phylogenomic utility of each of the four DNA classes on species tree inference. NJst trees estimated from CDS, CNEE, and intergenic segments, respectively, were consistent with previous studies (1–3) regarding basal divisions in birds long considered resolved (*SI Appendix*, Figs. S1 and S2*A*). In contrast, the intron-based NJst tree grouped Galloanserae with Palaeognathae, not with Neoaves (*SI Appendix*, Fig. S2*B*). The NJst trees based on the four individual DNA classes included 5, 11, 7, and 7 low-confidence nodes (bootstrap support (BS) < 70%) with a minimum bootstrap confidence of 51%, 14%, 37%, and 45% for CDS, CNEE, introns, and intergenic segments, respectively (*SI Appendix*, Figs. S1–S3). All these low-confidence nodes were located deep within Neoaves, indicating that the maximum likelihood gene trees built from each of the four DNA classes alone do not contain sufficient phylogenetic signal to fully resolve the relationships within Neoaves.

Applying a concatenation approach using RAxML to each of the four DNA classes produces trees with higher putative node support compared to corresponding NJst trees. All resultant concatenation trees supported the well-established basal divisions in birds, with the exception of intergenic segments, which separated the tinamous (Tinamiformes) from Palaeognathae to form the sister clade to Neognathae (*SI Appendix*, Figs. S4 and S5). The four concatenation trees based on individual DNA classes comprised 0, 4, 1, and 3 low-confidence nodes (BS < 70%) with a minimum bootstrap confidence of 85%, 58%, 57%, and 55% for CDS, CNEE, introns, and intergenic segments, respectively (*SI Appendix*, Fig. S3). The low-confidence nodes in the CNEE concatenation tree failed to resolve four key nodes deep within Neoaves (*SI Appendix*, Fig. S4*B*). In contrast, the CDS, intronic, and intergenic concatenation trees were each well-resolved at the ordinal level, but exhibited substantial incongruences with one another, indicating that different DNA classes contain conflicting phylogenetic signal (*SI Appendix*, Figs. S4 and S5).

We next examined the number of significantly discordant nodes (with BS >= 80%) across trees constructed from the four DNA classes, testing for the significance of these differences between NJst and RAxML trees using a binomial test. In pairwise comparisons between NJst trees constructed from CDS and CNEE, CDS and intergenic segments, CDS and introns, CNEE and intergenic segments, CNEE and introns, and intergenic segments and introns, there were 13, 21, 13, 14, 2, and 17 significantly discordant nodes, respectively (Fig. 1*A*). The same pairwise comparisons between RAxML trees yielded 20, 30, 20, 25, 5, and 27 significantly discordant nodes (Fig. 1*A*), revealing that comparisons of concatenated RAxML trees based on different DNA classes resulted in a much higher number of significantly discordant nodes than comparisons of coalescent NJst trees, and also underscoring the substantial conflict in phylogenetic signal among the four DNA classes.

Compatibility of Phylogenetic Signal across DNA Classes. To examine the goodness of fit of the multispecies coalescent model (MSC) to the four DNA classes, we calculated triplet frequencies of the ML gene trees built from each DNA class, respectively. According to the MSC, the frequencies of two minority triplets should be roughly equal if gene tree heterogeneity is caused mainly by incomplete lineage sorting (14, 15). We tested for significance in the difference between frequencies of two minority triplets using a binomial test with a Bonferroni correction for multiple comparisons. In addition to the gene tree heterogeneity introduced by estimation error (16), the MSC accounts for 76.9%, 96.5%, 95.3%, and 87.7% of gene tree heterogeneity in CDS, CNEE, introns, and intergenic segments, respectively, indicating that it has a good overall fit (Fig. 1*B*).

To better assess the conflicting phylogenetic signals among the four DNA classes, we calculated the proportion of species tree triplets for which significant majority gene tree triplets are topologically discordant across the four DNA classes, meaning their underlying alignments contain significantly conflicting signals as measured by the bootstrap. A binomial test of significance was performed, with *P*-values adjusted using a Bonferroni correction for multiple comparisons. The resultant proportions of discordant majority triplets were less than 1% in all six pairwise comparisons among DNA classes (Fig. 1*C*), suggesting that the ML gene trees built from the four DNA classes are compatible and can be combined to build a species tree using coalescence-based methods.

Effects of DNA Class Combinations on Species Tree Inference. To obtain a consistent and well-resolved phylogeny of Neoaves, we assessed the phylogenomic suitability of all 11 combinations involving two or more types of DNA classes across the four DNA classes using NJst. Two criteria were adopted to find the best combination of DNA classes based on the assumption that such a combination should maximize branch support in the corresponding NJst tree: 1) the minimum node support of the tree constructed from the combined DNA classes should be higher than the minimum support values of the trees built from each of the individual DNA classes, and 2) the number of nodes with a bootstrap support smaller than 50% and 70% of the tree built from the combined DNA classes should be less than that of the trees built from each of the individual DNA classes. Based on the two criteria, we determined that the combination of CDS, introns, and intergenic segments performs better than all other combinations and is most suitable for constructing a robust species tree (*SI Appendix*, Fig. S6). As a result, we eliminated CNEE from subsequent analysis. The resultant NJst tree involving CDS, introns, and intergenic segments was well resolved, with all but seven nodes receiving high branch support (BS >= 80%) (*SI Appendix*, Fig. S7). A RAxML tree based on the concatenated CDS, introns, and intergenic segments was largely congruent in topology with the corresponding NJst tree (*SI Appendix*, Fig. S8*A*). This phylogeny identified many novel clades in Neoaves with high confidence, foremost a basal division into two major lineages, one mainly containing land birds (Telluraves) and another mainly containing waterbirds and relatives (which we designate Aquaterraves).

Robustness of Novel Neoavian Relationships. The new Neoavian phylogeny built from the combination of CDS, introns, and intergenic segments uncovered a novel basal clade, Aquaterraves, uniting a wide array of waterbirds and their relatives. All seven low-confidence nodes in this new phylogeny (BS < 80%) were located within Aquaterraves, and this clade itself received a relatively low bootstrap support (BS = 78%) (*SI Appendix*, Fig. S7). The accuracy of inferring species trees can be compromised by gene trees that deviate significantly from gene trees predicted by the coalescent model. Recent studies suggest that excluding these outlier gene trees can potentially improve the fit of the coalescent model and ultimately enhance the accuracy of species tree inference (17, 18). To evaluate the robustness of the new phylogeny of Neoaves, we compared the species tree constructed using all gene trees to trees constructed by excluding outlier gene trees.

We calculated quartet distances between the new phylogeny and each of the ML gene trees estimated from CDS, introns, and intergenic segments, respectively (*SI Appendix*, *Materials and*

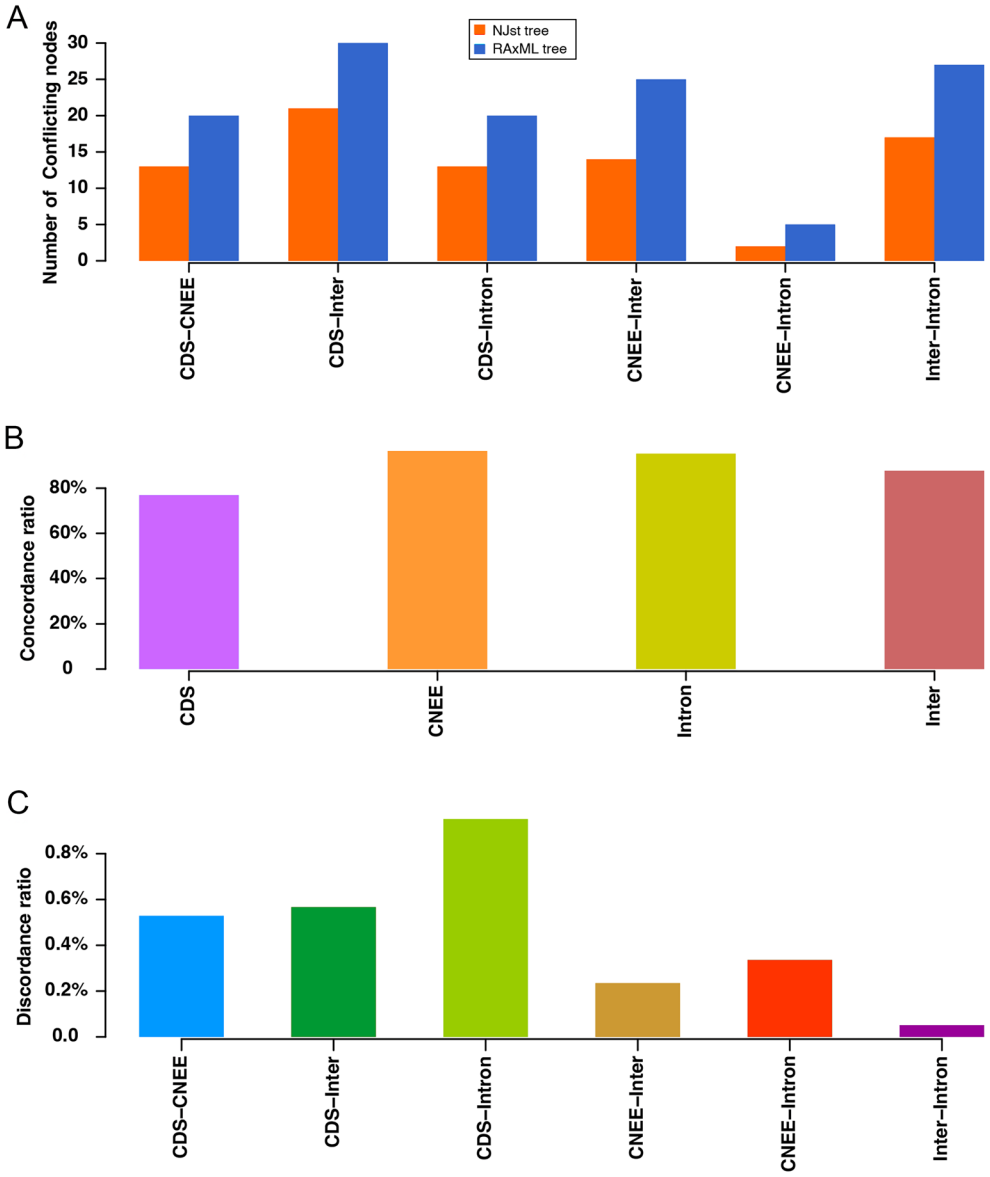

Fig. 1. Effects of the four individual DNA classes on species tree inference. (*A*) Comparison of the number of significantly discordant nodes for trees estimated using NJst and RAxML methods across all six pairwise combinations of DNA classes. The significance of differences between NJst and RAxML trees was examined using a binomial test. Note that the number of significantly discordant nodes (bootstrap support >= 80%) for trees built using RAxML is significantly greater than for trees built using NJst. (*B*) Consistency of the four DNA classes with the multispecies coalescent model. The concordant ratio indicates the proportion of gene tree heterogeneity that can be explained by the multispecies coalescent model. (*C*) Compatibility of maximum likelihood gene trees built from the four DNA classes. The discordant ratio indicates the proportion of significantly discordant majority triplets across all pairwise comparisons among DNA classes. Abbreviation: Inter, intergenic.

Methods). We sorted ML gene trees by quartet distances from the highest to the lowest and then created eight subsets of data by successively removing the highest 5%, 10%, 15%, 20%, 25%, 30%, 35%, and 40% of the sorted ML gene trees built from CDS, introns, and intergenic segments, respectively. We then constructed species trees using NJst from the eight subsets. The results show that the topology of all highly supported nodes (BS ≥ 80%) remains unchanged, whereas five low-confidence nodes (BS < 80%) experienced topological changes across trees built from the eight subsets (Fig. 2 and *SI Appendix*, Figs. S7, S9, and S10). Storks (Ciconiiformes) were grouped with penguins (Sphenisciformes) and tubenoses (Procellariiformes) with relatively low branch support (BS = 74%) in the tree built from the full dataset but emerged as the sister taxon to a lineage containing pelicans and allies (Pelecaniformes) and gannets (Suliformes) with high branch support (BS = 99%) in the tree built from the subset eliminating the highest 30% of outlier gene trees (Fig. 2 and *SI Appendix*, Figs. S7, S9, and S10). The other four low-confidence nodes were characterized by roughly equally low branch support (BS < 80%) across trees built from the eight subsets. The bootstrap value supporting the monophyly of Aquaterraves increased with the removal of outlier gene trees, reaching 100% after the removal of 25% of top outlier gene trees (Fig. 2 and *SI Appendix*, Figs. S7, S9, and S10). Moreover, the bootstrap support value for the subclade grouping Charadriiformes, Gruiformes, Apodiformes, and Caprimulgiformes increased from 80% in the tree built from the full dataset to 92% in the tree built from the subset eliminating 35% of gene trees (Fig. 2 and *SI Appendix*, Figs. S7, S9, and S10). Notably, while node support remained below 80%, the sister taxon relationship between Charadriiformes and Gruiformes was stable across all trees (i.e., full dataset and each of the eight subsets) (Fig. 2 and *SI Appendix*, Figs. S7, S9, and S10).

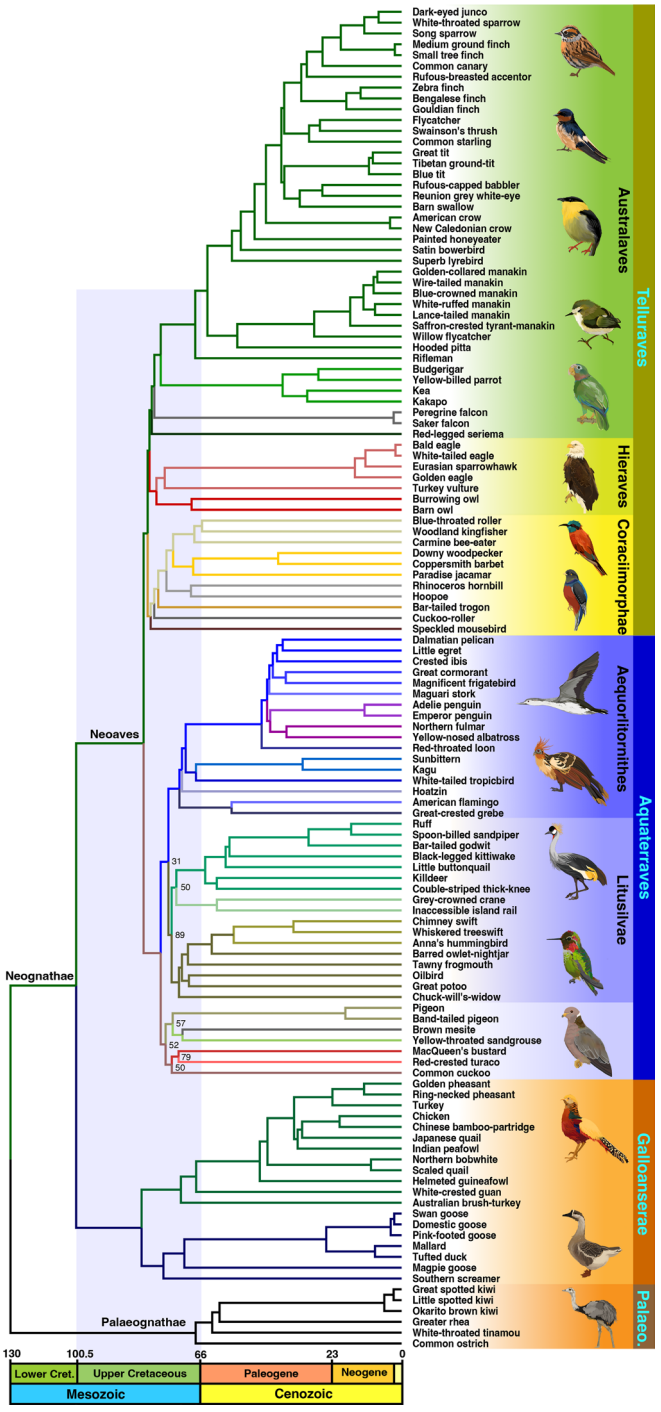

Fig. 2. Phylogenetic relationships of modern birds. The time-calibrated tree of 124 avian species was constructed using the coalescence-based method NJst from a combination of maximum likelihood gene trees of CDS, intron, and intergenic segments after removing 30% outliers. Numbers on nodes indicate bootstrap support values, with numbers >90% not shown. Background shading indicates the Late Cretaceous era.

We adopted three criteria to identify the best topology among trees built from the eight subsets: 1) maximizing bootstrap support for the novel clade Aquaterraves within Neoaves; 2) maximizing bootstrap support for other unchanged nodes; and 3) retaining as many gene trees as possible. Based on these criteria, the tree built from the subset containing 70% of ML gene trees, i.e., after removal of 30% of outlier loci, produced the most reliable topology for crown Aves, hereafter referred to as the "optimum tree" (Fig. 2). We also built a concatenation tree using RAxML based on the subset containing 70% of loci. This concatenation tree was largely consistent with the optimum tree in topology, supporting the division of Neoaves into Telluraves and Aquaterraves (*SI Appendix*, Fig. S8*B*).

Overall, the major relationships in the species tree constructed using all gene trees were found to be stable and very similar to those obtained when a certain percentage of the outlier gene trees removed. The discrepancies were mostly observed in relationships that were not strongly supported by the data. The analysis of outlier gene trees indicates that removing atypical gene trees can considerably improve the resolution of the inferred species tree.

Causes of Low-Confidence Nodes in Aquaterraves. Our analysis of branch robustness across datasets revealed five nodes within the new Aquaterraves with low branch support across trees built from the full dataset and the eight subsets of outlier removal. To verify whether these low branch support values were caused by conflicting phylogenetic signal among DNA classes or by a lack of sufficient phylogenetic signal, we quantified triplet frequencies in ML gene trees built from CDS, introns, and intergenic segments, respectively. We then measured the proportion of significant majority triplets of ML gene trees, which amounted to 88.1%, 83.6%, and 76% for ML gene trees built from CDS, introns, and intergenic segments, respectively, for the entire tree. When restricting this calculation to the clade comprising the new Aquaterraves, however, the proportion of significant majority triplets decreased to 55.8%, 36%, and 38% for the same DNA classes (*SI Appendix*, Fig. S11*A*). In the meantime, the proportions of majority triplets exhibiting significant conflict among one another was below 1% for both the whole tree and the clade comprising Aquaterraves across ML gene trees built from all three DNA classes (*SI Appendix*, Fig. S11*B*). These results suggest that low-confidence nodes within Aquaterraves are not caused by conflicting phylogenetic signal among DNA classes, but rather by a lack of sufficient phylogenetic signal in the dataset.

Dating of Nodes within the Avian Tree. We applied 20 conservatively selected fossil constraints to date the optimum tree (*SI Appendix*, *Materials and Methods*). For the divergence of crown Aves, we assigned the age of the earliest known member of crown Aves, *Asteriornis maastrichtensis* (66.7 Ma) (19), as the minimum bound and the age of the earliest known member of Ichthyornithes, *Ichthyornis dispar* (94.3 Ma) (20), as the maximum bound. The clade Ichthyornithes has played a substantial role in biologists' understanding of the origin of crown Aves, given their consistent phylogenetic placement close to crown Aves across different studies (20, 21). Ichthyornithes are arguably the member of the Ornithurae exhibiting the most complete fossil preservation and a well-known stratigraphical duration, with fossil representation across 10 Ma, extending from the Cenomanian to Campanian (~95 to 83.5 Ma) (22). Therefore, it is reasonable to use the age of the oldest known fossil of Ichthyornithes, *I. dispar*, as a maximal bound of crown Aves. In addition, we conducted sensitivity analyses to evaluate the impact of even earlier fossil Ornithurae taxa on dating, assigning *Enaliornis barretti* (105.3 Ma)—the earliest known Hesperornithes—and *Gansus yumenensis* (122.8 Ma)—the earliest known *Gansus*—as maximum bounds for the divergence of crown Aves (21, 23), along with the remaining 19 fossil calibrations.

We performed dating analysis on the optimum tree using MCMCTree (24) with CDS data and the full set of 20 fossil constraints (*SI Appendix*, *Materials and Methods*). Our results show that interordinal diversification of modern birds largely occurred in the Late Cretaceous and underwent a continuous radiation across the KPg boundary. Based on the lower level of 95% CIs, there are approximately 15 avian orders that likely originated in

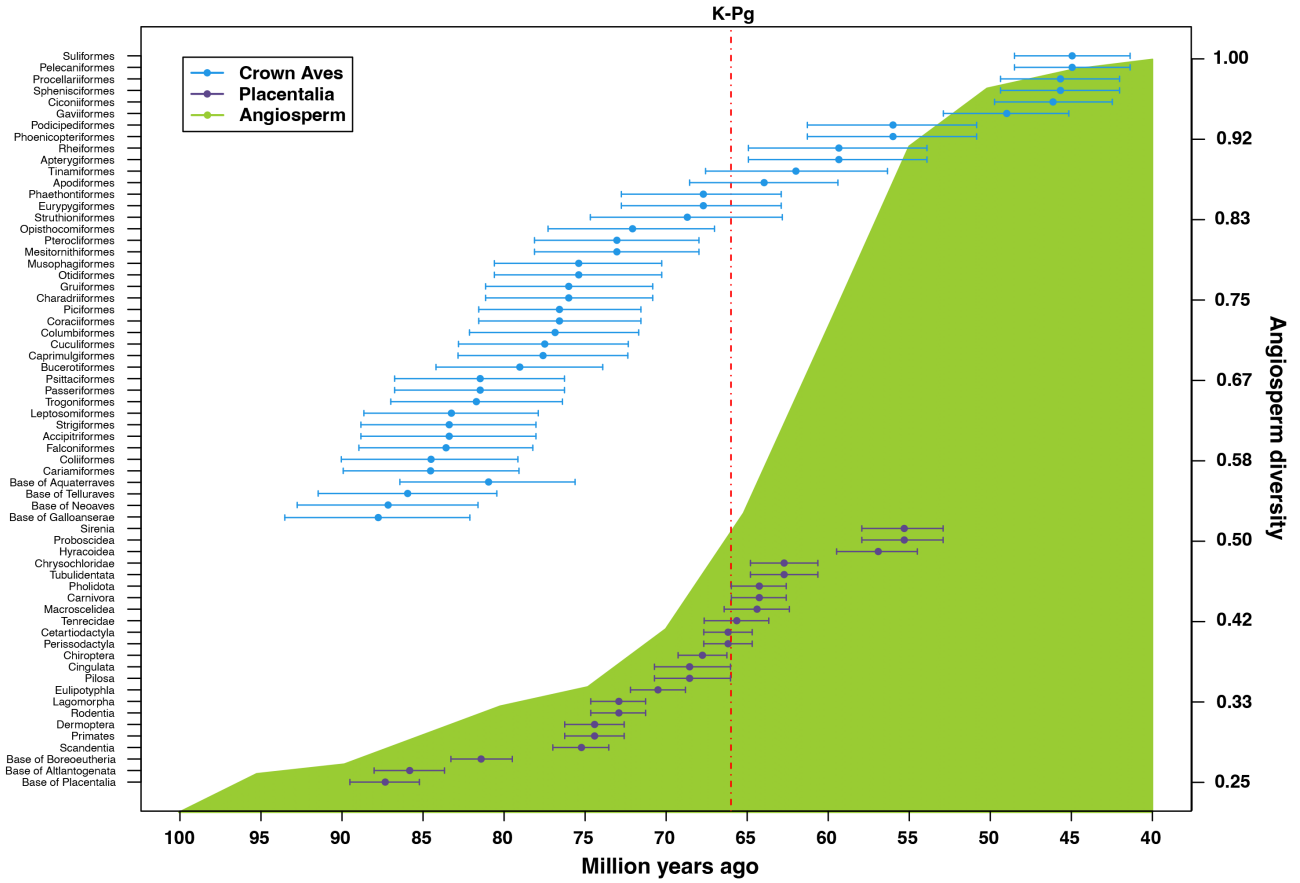

Fig. 3. Interordinal diversification of modern birds and placental mammals from the Late Cretaceous to the Paleogene. Note that both crown Aves and placental mammals underwent a constant and steady differentiation across the KPg boundary, without apparent interruption by the KPg mass extinction event. The green background indicates the increase of taxonomic diversity of flowering plants. The data on placental mammal and angiosperm diversity are based on Liu et al. (25) and Condamine et al. (26), respectively.

the Paleogene, whereas the origin of the remaining 24 orders falls within the Late Cretaceous (Fig. 3). Similar dating results were obtained when constrained with the fossils of *E. barretti* and *G. yumenensis* as the maximum bound for the divergence of crown Aves, respectively (*SI Appendix*, Fig. S12 and Datasets S2–S4).

Analysis of Changing Diversity of Modern Birds across Time. Louca and Pennell (27) recently showed that some methods for estimating diversification rates suffer from identifiability problems. But this warning does not mean that such efforts should be abandoned (28), and there is still considerable information in some extant-only methods for understanding the diversification process (29). To understand diversification dynamics over time across our optimal tree, we analyzed the time tree of modern birds using a maximum likelihood method for detecting shifts in diversification rate, implemented in the R package TreePar (30, 31). Because the method applied here aims to estimate the timing of shifts in diversification rate rather than diversification rates per se, the identifiability concerns outlined by Louca and Pennell likely do not apply. The divergence times of the time tree were fit to the birth—death models with 0, 1, 2, 3, and 4 rate shifts, respectively. Rate shifts were estimated in a grid of 1-My step between 2 and 110 Mya (30). Utilizing the likelihood ratio test, models with 0, 1, and 2 rate shifts were rejected with a P-value < 0.05. In contrast, the model with 3 shifts was not rejected, as indicated by a P-value > 0.05 (*SI Appendix*, Table S1). Thus, this likelihood ratio test favors a model with 2 rate shifts occurring at 68 My, and 81 My. In addition, 100 bootstrap samples of divergence times were generated to construct the 90% CI [0.26, 0.7] and [0.68, 0.87] for the two shift estimates. The results suggest that crown Aves underwent two shifts in diversification rate, both of which occurred in the Late Cretaceous before the KPg mass extinction event.

Causes of Incongruence with Previous Dating Estimates. Recent phylogenetic analyses of large-scale DNA matrices have supported an explosive radiation of crown Aves after the KPg boundary in the Paleogene (2, 3), contradicting our dating results. We performed sensitivity analyses to examine the cause of conflict with the results of the two most comprehensive previous studies.

Eighteen out of 20 fossil calibrations used by Jarvis et al. (2) can be applied to our optimum tree. In addition, our taxon sampling includes all 48 avian species and the outgroup, American alligator, sampled by Jarvis et al. We performed two dating analyses using the above 18 fossil calibrations on two different trees: one is our optimum tree based on 125 taxa, and another is a reduced tree including only the 49 taxa used by Jarvis et al., generated by excluding 76 species from the optimum tree but otherwise maintaining its topological structure. Our results show that the ages estimated using the reduced tree are consistent with the findings of Jarvis et al., supporting a rapid radiation of crown Aves post the KPg boundary in the Paleogene (*SI Appendix*, Fig. S13*A*). In contrast, the ages estimated using the full tree are consistent with the results of this study, supporting a continuous radiation of crown birds from the Late Cretaceous across the KPg boundary until the Paleogene (*SI Appendix*, Fig. S13*B*).

In a similar approach, 15 fossil calibrations employed in this study were derived from Prum et al. (3). To examine the cause of

conflict with the results of Prum et al., we performed dating analyses using our optimum tree based on the fossil constraint settings by Prum et al., that is, the maximum bound of all avian fossil calibrations was set to 65 Ma, and the minimum and maximum constraints for the divergence of crown Aves were set to 60.5 and 86.5 Ma, respectively. The results of this dating analysis are consistent with the findings of Prum et al., supporting a rapid radiation of crown Aves post the KPg boundary (*SI Appendix*, Fig. S14*A*). Next, we performed dating analysis by relaxing the maximum bound on fossil constraints utilized by Prum et al., leading to results that contradict the findings of Prum et al., but are consistent with the findings of this study, and support a continuous radiation of crown Aves across the KPg boundary from the Late Cretaceous until the Paleogene (*SI Appendix*, Fig. S14*B*).

Our results demonstrate that the dating results of Jarvis et al. (2) are biased toward younger age estimates due to sparse taxon sampling (*SI Appendix*, Fig. S13), whereas the dating results of Prum et al. (3) can only be obtained by imposing a strict restriction on the maximum bound of all avian fossil constraints to 65 Ma (*SI Appendix*, Fig. S14). The younger dating estimates caused by sparse taxon sampling might be due to the node density effect, that is, phylogenetic inference tends to underestimate the true amount of molecular changes along lineages when taxon sampling is sparse (32). After overcoming the aforementioned issues, the dating results of the two recent studies are consistent with this study, supporting a continuous radiation of crown Aves during the Late Cretaceous and through the KPg boundary (*SI Appendix*, Figs. S13 and S14).

Discussion

Phylogenetic Relationships of Neoaves. Our optimum tree provides consistent resolution of phylogenetic relationships within Neoaves, which comprise the bulk of modern birds and whose relationships have long been phylogenetically controversial, dividing them into two basal clades, a previously recognized Telluraves (landbirds) (33) and a newly circumscribed Aquaterraves (waterbirds and relatives) (Fig. 2).

Within Telluraves, the optimum tree supports the large and previously established monophyletic clade Coraciimorphae (3) as the most basal clade (Fig. 2). It also uncovers a novel monophyletic clade, uniting Accipitriformes (hawks) and Strigiformes (owls), which we designate Hieraves. The Hieraves are sister to Australaves (34), a known clade consisting of Passeriformes (perching birds), Psittaciformes (parrots), Falconiformes (falcons), and Cariamiformes (seriemas).

The optimum tree identifies two highly supported lineages within the newly circumscribed clade of Aquaterraves (Fig. 2). One lineage, named Aequorlitornithes (3), contains a number of aquatic groups as well as the enigmatic hoatzin (Opisthocomiformes), as successive sister taxa to Aequornithes (core waterbirds). Another lineage, which we name Litusilvanae, consists of Charadriiformes (shorebirds), Gruiformes (cranes and allies), and Caprimulgimorphae (swifts, hummingbirds, and nightjars). The optimum tree strongly supports grouping pigeons (Columbiformes) and a number of related orders within Aquaterraves; their relationships have been particularly unstable across previous genomic studies (2, 3), and their exact position within Aquaterraves remains controversial even in our dataset, requiring future attention.

Uninterrupted Radiation of Modern Birds across the KPg Boundary. The new timeline for modern bird diversification offered by this study improves our understanding of vertebrate evolution to one in which the early radiation of birds largely parallels that of other organisms, such as angiosperms, mammals, fishes, and insects (Fig. 4). Using data from diverse sources, we show that the taxonomic diversity of crown Aves, placental mammals, multituberculate mammals, ray-finned fishes and holometabolous insects (butterflies, bees, ants, etc.) is positively correlated with the rise of angiosperms during the Late Cretaceous and Paleogene (Fig. 4 and *SI Appendix*, Fig. S15), suggesting that angiosperm blooms promoted diversification in these animal groups. Naturally, the formation of angiosperm forests and other dense habitats would have had cascading effects on animals by providing fruits, seeds, and other biomass, in both terrestrial and aquatic ecosystems, driving the diversification of invertebrates, then insectivorous vertebrates feeding on them, and finally carnivores feeding on the new assemblages of smaller animals.

Galliformes and Anseriformes were the earliest modern-day neognath birds, appearing in our optimum tree at about 87.8 Ma (95% CI: 82.1 to 93.5), ~20 My prior to previously accepted estimates (2, 3) (Fig. 3). This incongruence notwithstanding, the appearance of both orders is consistent with multiple potential early fossils of crown Aves. Foremost among these putative Late Cretaceous crown Aves are *Teviornis gobiensis* (~70 Ma) and *Austinornis lentus* (~85 Ma), which are considered possible Anseriformes and Galliformes, respectively (38–41). Beyond these two fossils, Olson and Parris (42) described nine avian species from the Maastrichtian age of the Late Cretaceous under the newly established family Graculavidae and considered them primitive Charadriiformes. They also considered *Tytthostonyx glauconiticus* (~66 Ma) a primitive species of seabird, belonging either to modern Pelecaniformes or Procellariiformes. We estimated the origination time of Charadriiformes and the common ancestor of seabirds to be ~76 Ma (95% CI: 81.1 to 70.8) and ~70.8 Ma (95% CI: 76 to 65.9), respectively, agreeing with the fossil evidence presented by Olson and Parris. At the same time, the exact phylogenetic positions of these aforementioned fossils remain controversial (38–42). Future fossil discoveries are avidly awaited to clarify their phylogenetic positions and corroborate our dating results.

We used the fossil of a stem group member of the Fregatidae, *Limnofregata azygosternon*, as the minimum bound (51.58 Ma) for the divergence between Fregatidae and Phalacrocoracidae (*SI Appendix*, Fig. S19). The estimated age for the divergence between Fregatidae and Phalacrocoracidae is 40.57 Ma (95% CI: 43.99 to 37.15 Ma) (Fig. 2 and Dataset S2), which is younger than the fossil constraint placed on this node. *Limnofregata* is considered one of the best available fossils of early pelecaniform birds and shows some affinity with extant frigatebirds (43). In strong contrast to extant frigatebirds which live in oceanic environments, however, *Limnofregata* used to live in freshwater environments, suggesting a transition from freshwater to a marine lifestyle in the course of pelecaniform evolution (44, 45). Since we placed a minimum fossil constraint to date the node of Fregatidae and Phalacrocoracidae, our result is consistent with recent studies suggesting that the Eocene occurrence of *Limnofregata* preceded a long transition of early pelecaniforms into marine frigatebirds (44, 45).

The origin of placental mammals is commonly placed in the Late Cretaceous (25, 46–48), and the radiation of both placental and multituberculate mammals is thought to have been continuous across the KPg boundary (25, 35). Notably, the basal divergences of both Neoaves (~87.2 Ma, 95% CI: 81.6 to 92.8) and placental mammals (~87.3 Ma, 95% CI: 85.2 to 89.5) (25) occurred almost concurrently during the Late Cretaceous (Fig. 3). Furthermore, within Neoaves and placental mammals, the basal divergence of Telluraves (landbirds, ~86 Ma, 95% CI: 80.4 to 91.5) occurred nearly simultaneously with that of Atlantogenata (Xenarthra and Afrotheria, ~85.8 Ma, 95% CI: 83.7 to 88) (25), and the basal divergence of Aquaterraves

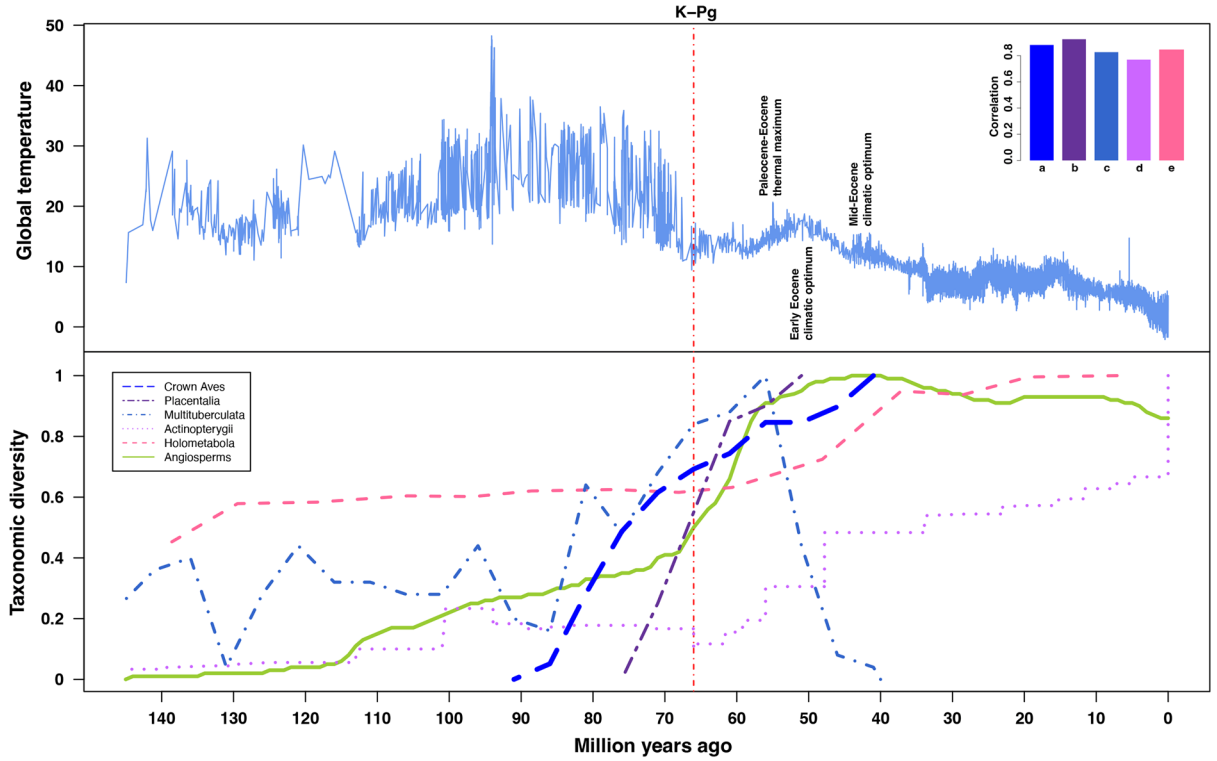

Fig. 4. Diversity changes in modern birds, placental mammals, multituberculate mammals, marine ray-finned fishes, holometabolous insects (butterflies, bees, ants, etc.), and flowering plants from the Cretaceous to the present time. The panel on the *Upper Right* shows the correlation coefficients of diversity changes between angiosperms and each of the five animal groups: a) crown Aves, b) placental mammals, c) multituberculates, d) marine ray-finned fishes, and e) holometabolous insects. The curve of average global temperature represents a proxy for climatic changes during the Cretaceous and Cenozoic. Note that among the six taxonomic groups analyzed in the figure, the KPg extinction event only had a noticeable negative impact on the biodiversity of ray-finned fishes, but not on the other five groups. The diversity curves of placental mammals, multituberculates, ray-finned fishes, holometabolous insects, and angiosperms were plotted based on data extracted from Liu et al. (25), Wilson et al. (35), Guinot and Cavin (36), Nicholson et al. (37), and Condamine et al. (26), respectively. The curve of global temperature changes was plotted based on data extracted from Condamine et al. (26).

(waterbirds and relatives, ~81 Ma, 95% CI: 75.6 to 86.4) occurred at approximately the same time as that of Boreoeutheria (Euarchontoglires and Laurasiatheria, ~81.4 Ma, 95% CI: 79.5 to 83.3) (25). The contemporaneous diversification of modern birds and mammals could well be driven by the Angiosperm Terrestrial Revolution event, the rise and increasing dominance of angiosperms during the Late Cretaceous and early Paleogene (49).

Uniquely among Neoaves, a tight assemblage of six orders forming the core waterbirds (Aequornithes) is noteworthy by virtue of their relatively young origin in the Early Eocene (~50 Ma; Figs. 2 and 3). Most members of this group are seabirds and include such lineages as penguins, albatrosses, loons and pelicans. Our results confirm their explosive origin in the wake of the Paleogene-Eocene Thermal Maximum (PETM) event (~56 Ma) when a sudden spike in atmospheric carbon led to increased temperatures and strong ocean acidification across the globe (50) (Figs. 3 and 4). After surviving the PETM crisis, the common ancestor of Aequornithes found itself in a global environment that was slowly but steadily cooling, gradually creating marine conditions more conducive to the survival and diversification of ocean fishes and the seabirds feeding on them (Fig. 4). Today, the bulk of aequornithine diversity can be found in the nutrient-rich waters of the polar and subpolar seas of the world. Under some projections, anthropogenic climate change would create PETM-like atmospheric conditions, which might lead to a repeat of the marine extinction crisis that unfolded ~56 Mya.

Overall, our study adds considerable support to the narrative that the rise of modern birds and mammals followed a relatively slow process of gradual evolution. The KPg extinction event has long been thought to be the dominant factor in early bird evolution; our study shows that—as opposed to nonavian dinosaurs—birds weathered the KPg crisis relatively well, but that the more recent PETM event had a substantial impact at least on those birds that depend on the ocean. In a previous study, we had shown that the evolution of mammals likely followed a trans-KPg model of diversification at a time when modern birds were assumed to have undergone their extensive radiation much later (25). Our present phylogenomic analyses indicate that the diversification of Neoaves is in close lockstep with that of mammals, suggesting a model of constant diversification that is in keeping with Charles Darwin's original interpretation of gradualism, but at loggerheads with alternative theories involving punctuated evolution.

Materials and Methods

Taxon Sampling. We sampled 118 species of neognaths in this study, representing all 35 orders of Neognathae. In addition, we sampled six species of paleognaths and one crocodilian species as outgroups (Dataset S1). We downloaded annotated genomes of 124 species from NCBI, Ensembl, and GiGa. In addition, we conducted de novo whole genome sequencing for rufous-breasted accentor (*Prunella strophita*).

Genetic Sampling.

Identification of orthologous CDS. To identify orthologous CDS sequences across species, we applied the cDNA sequences of the chicken as a query against the cDNA sequences of the remaining 124 species using the program OrthoFinder (v2.3.12) with an E-value cutoff of 1e−10 (51). We then filtered out those candidate orthologs with more than 25 missing species and/or less than 100 bp in sequence length. As a result, a total of 5,756 CDS were retained for further analysis.

Identification of orthologous introns. We extracted intron sequences of the 5,756 genes from the genome of each sampled species according to the annotation files using a custom script written in Perl. Due to considerable variation in the intron number and length of orthologous genes across species, we performed blast searches using the intronic sequences of every chicken gene as a query to blast against the intron sequences of the same orthologous genes from each of the remaining 124 species with an E-value cutoff of 1e−10 (52). For genes having multiple introns, we ran blast searches for each of the intronic sequences against those of the same orthologous genes across species and selected the one with maximum species coverage. We filtered out those orthologous intronic sequences with more than 25 missing species and/or less than 100 bp in sequence length. A total of 4,871 orthologous introns were retained for this study.

Identification of orthologous intergenic segments. The chicken genome was used as a reference to identify new noncoding intergenic markers across species for this study. We identified intergenic markers based on autosomes and both sex chromosomes (Z and W) and excluded mitochondrial DNA. We first applied a custom script written in R to extract noncoding intergenic sequences from the chicken genome based on the annotation file. Next, we cut the extracted intergenic sequences into segments with a length of ~1,000 bp. To identify orthologs of intergenic markers across species, we performed blastn searches using the intergenic segments obtained from the chicken against the genomes of the remaining 124 species with an E-value cutoff of 1e−10. We filtered out those orthologs of intergenic segments with more than 25 missing species and/or less than 100 bp in sequence length. As a result, a total of 2,384 intergenic markers were retained for this study.

Identification of orthologous CNEEs. We downloaded orthologous CNEEs of seven avian taxa deposited in Dryad (https://doi.org/10.5061/dryad.fj02s0j) (53), all of which were incorporated in this study, including chicken, okarito brown kiwi, little spotted kiwi, great spotted kiwi, greater rhea, common ostrich, and white-throated tinamou. To identify orthologous CNEEs across species, we conducted blastn searches using chicken CNEE sequences as a query against the genome of each of the remaining 118 species with an E-value cutoff of 1e−10. In addition, we used the CNEE sequences from the little spotted kiwi and the greater rhea as a query to blast against the genomes of each of the remaining 118 species, respectively, in an effort to collect sufficient numbers of orthologous CNEE sequences from these target genomes. We also filtered out orthologous CNEEs with more than 25 missing species and/or less than 100 bp in sequence length. As a result, we identified a total of 12,582 CNEEs for subsequent analysis. In addition, we performed a blastn search to use the chicken's CNEE sequences as a query against the chicken's intergenic segments with an E-value cutoff of 1e−10. We found 133 CNEE sequences related to intergenic segments. Because intergenic segments are a primary data source for this study, we excluded these 133 CNEE sequences from the CNEE data and retained 12,449 CNEE sequences.

Data, Materials, and Software Availability. The assembly of the *Prunella strophita* genome has been deposited in the National Genomics Data Center (https://db.cngb.org/) under the accession number CNP0002307 (54). The alignment data of CDS, CNEE, introns, and intergenic segments used in this study have been deposited in Figshare under accession number https://doi.org/10.6084/m9.figshare.21499230.v1 (55). We provide additional details regarding de novo genome sequencing, phylogenomic analysis, molecular dating analysis, fossil calibration points, ancient gene flow test, diversification rate analysis, and correlation analysis of taxonomic diversity in *SI Appendix*.

ACKNOWLEDGMENTS. We thank two anonymous reviewers for their helpful comments, S. Song, Y. Zhang, and C. Cannistraci for technical assistance, and X. Cui for providing drawing. We also thank the National Supercomputer Center at Tianjin and the Georgia Advanced Computing Resource for computing support. This work was supported by funds from the National Natural Science Foundation of China (Grant 32020103005 to Y.Q. and S.W.; Grant 31772441 to S.W. and F.E.R.) and the Priority Academic Program Development of Jiangsu Higher Education Institutions.

Author affiliations: [a]Jiangsu Key Laboratory of Phylogenomics and Comparative Genomics, Jiangsu International Joint Center of Genomics, School of Life Sciences, Jiangsu Normal University, Xuzhou, Jiangsu 221116, China; [b]Department of Biological Sciences, National University of Singapore, Singapore 117543, Singapore; [c]School of Computer and Communication Engineering, Changsha University of Science and Technology, Changsha, Hunan 410114, China; [d]School of Earth Science and Resources, Chang'an University, Xi'an, Shaanxi 710054, China; [e]Key Laboratory of Vertebrate Evolution and Human Origins, Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences, Beijing 100044, China; [f]Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China; [g]Department of Organismic and Evolutionary Biology, Museum of Comparative Zoology, Harvard University, Cambridge, MA 02138; and [h]Department of Statistics, Institute of Bioinformatics, University of Georgia, Athens, GA 30606

1. S. Hackett *et al.*, A phylogenomic study of birds reveals their evolutionary history. *Science* **320**, 1763–1768 (2008).
2. E. D. Jarvis *et al.*, Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science* **346**, 1320–1331 (2014).
3. R. Prum *et al.*, A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature* **526**, 569–573 (2015).
4. G. W. Cottrell *et al.*, Eds. Check-list of Birds of the World (Harvard University Press, Cambridge, MA, USA, 1931–1987).
5. G. Fain, P. Houde, Parallel radiation in the primary clades of birds. *Evolution* **58**, 2558–2573 (2004).
6. A. Graybeal, Is it better to add taxa or characters to a difficult phylogenetic problem? *Syst. Biol.* **47**, 9–17 (1998).
7. J. Townsend, F. Lopez-Giraldez, Optimal selection of gene and ingroup taxon sampling for resolving phylogenetic relationships. *Syst. Biol.* **59**, 446–457 (2010).
8. E. Braun, J. Cracraft, P. Houde, "Resolving the Avian Tree of Life from top to bottom: The promise and potential boundaries of the phylogenomic era" in *Avian Genomics in Ecology and Evolution: From the Lab into the Wild*, R. H. S. Kraus, Ed. (Springer, 2019), pp. 151–210.
9. S. Reddy *et al.*, Why do phylogenomic data sets yield conflicting trees? Data type influences the Avian Tree of Life more than taxon sampling. *Syst. Biol.* **66**, 857–879 (2017).
10. E. Braun, R. Kimball, Data types and the phylogeny of Neoaves. *Birds* **2**, 1–22 (2021).
11. G. Mayr, J. Clarke, The deep divergences of neornithine birds: A phylogenetic analysis of morphological characters. *Cladistics* **19**, 527–533 (2006).
12. L. Liu, L. Yu, Estimating species trees from unrooted gene trees. *Syst. Biol.* **60**, 661–667 (2011).
13. A. Stamatakis, RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
14. J. Degnan, N. Rosenberg, Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends Ecol. Evol.* **24**, 332–340 (2009).
15. C. Ané, "Reconstructing concordance trees and testing the coalescent model from genome-wide data sets" in *Estimating Species Trees: Practical and Theoretical Aspects*, L. Knowles, L. Kubatko, Eds. (Wiley-Blackwell, New Jersey, 2010), pp. 35–52.
16. L. Cai *et al.*, The perfect storm: Gene tree estimation error, incomplete lineage sorting, and ancient gene flow explain the most recalcitrant ancient angiosperm clade, malpighiales. *Syst. Biol.* **70**, 491–507 (2021).
17. X. Jiang, S. Edwards, L. Liu, The multispecies coalescent model outperforms concatenation across diverse phylogenomic data sets. *Syst. Biol.* **69**, 795–812 (2020).
18. J. Walker, J. Brown, S. Smith, Analyzing contentious relationships and outlier genes in phylogenomics. *Syst. Biol.* **67**, 916–924 (2018).
19. D. Field, J. Benito, A. Chen, J. Jagt, D. Ksepka, Late Cretaceous neornithine from Europe illuminates the origins of crown birds. *Nature* **579**, 397–401 (2020).
20. D. Field *et al.*, Complete Ichthyornis skull illuminates mosaic assembly of the avian head. *Nature* **557**, 96–100 (2018).
21. T. Tanaka, Y. Kobayashi, K. Kurihara, A. Fiorillo, M. Kano, The oldest Asian hesperornithiform from the Upper Cretaceous of Japan, and the phylogenetic reassessment of Hesperornithiformes. *J. Syst. Palaeontol.* **16**, 689–709 (2017).
22. J. Benito *et al.*, Forty new specimens of Ichthyornis provide unprecedented insight into the postcranial morphology of crownward stem group birds. *PeerJ* **10**, e13919 (2022).
23. H.-L. You *et al.*, A nearly modern amphibious bird from the Early Cretaceous of northwestern China. *Science* **312**, 1640–1643 (2006).
24. M. dos Reis, Z. Yang, Approximate likelihood calculation for Bayesian estimation of divergence times. *Mol. Biol. Evol.* **28**, 2161–2172 (2011).
25. L. Liu *et al.*, Genomic evidence reveals a radiation of placental mammals uninterrupted by the KPg boundary. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E7282–E7290 (2017).
26. F. Condamine, D. Silvestro, E. Koppelhusb, A. Antonellid, The rise of angiosperms pushed conifers to decline during global cooling. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 28867–28875 (2020).
27. S. Louca, M. Pennell, Extant timetrees are consistent with a myriad of diversification histories. *Nature* **580**, 502–505 (2020).
28. A. Helmstetter *et al.*, Pulled diversification rates, lineages-through-time plots, and modern macroevolutionary modeling. *Syst. Biol.* **71**, 758–773 (2022).
29. T. Vasconcelos, B. O'Meara, J. Beaulieu, A flexible method for estimating tip diversification rates across a range of speciation and extinction scenarios. *Evolution* **76**, 1420–1433 (2022).
30. T. Stadler, Mammalian phylogeny reveals recent diversification rate shifts. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 6187–6192 (2011).
31. T. Stadler, Sampling-through-time in birth-death trees. *J. Theo. Biol.* **267**, 396–404 (2010).
32. C. Venditti, A. Meade, M. Pagel, Detecting the node-density artifact in phylogeny reconstruction. *Syst. Biol.* **55**, 637–643 (2006).
33. T. Yuri *et al.*, Parsimony and model-based analyses of indels in avian nuclear genes reveal congruent and incongruent phylogenetic signals. *Biology (Basel)* **2**, 419–444 (2013).
34. P. Ericson, Evolution of terrestrial birds in three continents: Biogeography and parallel radiations. *J. Biogeogr.* **39**, 813–824 (2012).
35. G. P. Wilson *et al.*, Adaptive radiation of multituberculate mammals before the extinction of dinosaurs. *Nature* **483**, 457–460 (2012).

36. G. Guinot, L. Cavin, 'Fish' (Actinopterygii and Elasmobranchii) diversification patterns through deep time. *Biol. Rev.* **91**, 950–981 (2016).
37. D. Nicholson, A. Ross, P. Mayhew, Fossil evidence for key innovations in the evolution of insect diversity. *Proc. R. Soc. B* **281**, 20141823 (2014).
38. E. Kurochkin, G. Dyke, A. Karhu, A new presbyornithid bird (Aves, Anseriformes) from the Late Cretaceous of southern Mongolia. *Am. Mus. Novit.* **3386**, 1–11 (2002).
39. J. Clarke, M. Norell, New avialan remains and a review of the known avifauna from the Late Cretaceous Nemegt Formation of Mongolia. *Am. Mus. Novit.* **3447**, 1–12 (2004).
40. J. Clarke, Morphology, phylogenetic taxonomy, and systematics of Ichthyornis and Apatornis (Avialae, Ornithurae). *Bull. Am. Mus. Nat. Hist.* **286**, 1–179 (2004).
41. F. Agnolin, F. Novas, G. Lio, Neornithine bird coracoid from the Upper Cretaceous of Patagonia. *Ameghiniana* **43**, 245–248 (2006).
42. S. Olson, D. Parris, The cretaceous birds of New Jersey. *Smithson Contrib. Paleobiol.* **63**, 1–22 (1987).
43. G. Mayr, *Paleogene Fossil Birds* (Springer Science & Business Media, New York, 2009).
44. N. Smith, Phylogenetic analysis of Pelecaniformes (Aves) based on osteological data: Implications for waterbird phylogeny and fossil calibration studies. *Plos One* **5**, e13354 (2010).
45. T. Stidham, A new species of Limnofregata (Pelecaniformes: Fregatidae) from the Early Eocene Wasatch Formation of Wyoming: Implications for palaeoecology and palaeobiology. *Paleontology* **58**, 239–249 (2014).
46. R. W. Meredith *et al.*, Impacts of the Cretaceous terrestial revolution and KPg extinction on mammal diversification. *Science* **334**, 521–524 (2011).
47. M. J. Phillips, Geomolecular dating and the origin of placental mammals. *Syst. Biol.* **65**, 546–557 (2016).
48. N. Foley *et al.*, A genomic timescale for placental mammal evolution. *Science* **380**, eabl8189 (2023).
49. M. Benton, P. Wilf, H. Sauquet, The Angiosperm Terrestrial Revolution and the origins of modern biodiversity. *New Phytol.* **233**, 2017–2035 (2022).
50. L. Haynesa, B. Hönischa, The seawater carbon inventory at the paleocene–eocene thermal maximum. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 24088–24095 (2020).
51. D. Emms, S. Kelly, OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biol.* **20**, 238 (2019).
52. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
53. A. Cloutier *et al.*, Whole-genome analyses resolve the phylogeny of flightless birds (Palaeognathae) in the presence of an empirical anomaly zone. *Syst. Biol.* **68**, 937–955 (2019).
54. Y. Qu, The genome of Prunella strophiata. National Genomics Data Center. https://db.cngb.org/search/project/CNP0002307. Accessed 19 January 2024.
55. L. Liu, S. Wu, Genomes, fossils, and the concurrent rise of modern birds and angiosperms in the Late Cretaceous. Figshare. https://doi.org/10.6084/m9.figshare.21499230.v1. Accessed 17 January 2024.