Phylogenomic evidence for multiple losses of flight in ratite birds

John Harshman[a,b,c], Edward L. Braun[c,d,e], Michael J. Braun[c,f,g], Christopher J. Huddleston[f], Rauri C. K. Bowie[a,h,i], Jena L. Chojnowski[d], Shannon J. Hackett[a], Kin-Lan Han[d,f,g], Rebecca T. Kimball[d], Ben D. Marks[j], Kathleen J. Miglia[k], William S. Moore[k], Sushma Reddy[a], Frederick H. Sheldon[j], David W. Steadman[l], Scott J. Steppan[m], Christopher C. Witt[j,n], and Tamaki Yuri[d,f]

[a]Zoology Department, Field Museum of Natural History, 1400 South Lakeshore Drive, Chicago, IL 60605; [b]4869 Pepperwood Way, San Jose, CA 95124; [d]Department of Zoology, University of Florida, Gainesville, FL 32611; [f]Department of Vertebrate Zoology, National Museum of Natural History, Smithsonian Institution, 4210 Silver Hill Road, Suitland, MD 20746; [g]Behavior, Ecology, Evolution, and Systematics Program, University of Maryland, College Park, MD 20742; [h]Museum of Vertebrate Zoology and Department of Integrative Biology, University of California, Berkeley, CA 94720; [i]Department of Science and Technology/National Resource Foundation Centre of Excellence at the Percy FitzPatrick Institute, Department of Botany and Zoology, Stellenbosch University, Matieland 7602, South Africa; [j]Museum of Natural Science, 119 Foster Hall, Louisiana State University, Baton Rouge, LA 70803; [k]Department of Biological Sciences, Wayne State University, 5047 Gullen Mall, Detroit, MI 48202; [l]Florida Museum of Natural History, University of Florida, Gainesville, FL 32611; [m]Department of Biological Science, Florida State University, Tallahassee, FL 32306; and [n]Department of Biology and Museum of Southwestern Biology, University of New Mexico, Albuquerque, NM 87131

Edited by Morris Goodman, Wayne State University School of Medicine, Detroit, MI, and approved July 17, 2008 (received for review April 2, 2008)

Ratites (ostriches, emus, rheas, cassowaries, and kiwis) are large, flightless birds that have long fascinated biologists. Their current distribution on isolated southern land masses is believed to reflect the breakup of the paleocontinent of Gondwana. The prevailing view is that ratites are monophyletic, with the flighted tinamous as their sister group, suggesting a single loss of flight in the common ancestry of ratites. However, phylogenetic analyses of 20 unlinked nuclear genes reveal a genome-wide signal that unequivocally places tinamous within ratites, making ratites polyphyletic and suggesting multiple losses of flight. Phenomena that can mislead phylogenetic analyses, including long branch attraction, base compositional bias, discordance between gene trees and species trees, and sequence alignment errors, have been eliminated as explanations for this result. The most plausible hypothesis requires at least three losses of flight and explains the many morphological and behavioral similarities among ratites by parallel or convergent evolution. Finally, this phylogeny demands fundamental reconsideration of proposals that relate ratite evolution to continental drift.

convergence | flightlessness | Paleognath | homoplasy | vicariance biogeography

Living birds are divided into two major groups, Palaeognathae and Neognathae (1, 2), a classification based originally on bony palate structure (3, 4). Palaeognathae also is traditionally divided into two groups, the flightless ratites (defined by absence of a keel on the sternum) and the volant tinamous. Although they include fewer than 1% of extant avian species, paleognaths have long been viewed as central to understanding the early evolution of birds. The many morphological and behavioral similarities of ratites suggest common ancestry, but some have proposed that they instead reflect convergent adaptation to a flightless, cursorial lifestyle. The distribution of ratites is also remarkable; ostriches live in Africa, rheas in South America, emus and cassowaries in Australasia, kiwis and moas (now extinct) in New Zealand, and elephant birds (also now extinct) in Madagascar. How did these flightless birds get to these far-flung southern landmasses?

Paleognath relationships have been controversial since the earliest days of evolutionary biology. When Huxley defined the paleognathous palate he also stated that extant ratites "are but waifs and strays of what was once a very large and important group" (3). Nevertheless, the notion that ratites have independent origins arose around the same time, when Owen suggested that they have closer affinities to various volant groups while being united by the "arrested development of wings unfitting them for flight" (5). Ratite monophyly was debated throughout much of the last century, with Mayr and Amadon (6) stating in 1951 that the "present consensus is that the main groups of these birds are of independent origin." DeBeer provided a developmental explanation for the similarities among ratites when he interpreted the paleognathous palate and other features of extant ratites as neotenic (7). Paleognath monophyly was questioned as late as the 1980s (8), but it has been confirmed by many recent morphological and molecular studies (9–13).

Most recent studies have also strongly supported ratite monophyly (9–12, 14), suggesting a single loss of flight in their common ancestor. This puzzled biogeographers for more than a century, because ratites would be unable to achieve their current distribution on southern land masses if their common ancestor was flightless. Continental drift provided a compelling solution. No longer was it necessary to imagine giant flightless birds crossing vast oceans; they could have rafted to their current distributions on fragments of the Earth's crust (15). Although the proposed phyletic branching patterns for ratites do not correspond perfectly to the order of separation of land masses during the breakup of Gondwana, the convenient serendipity of continental drift as a mechanistic explanation for ratite distribution proved irresistible (10, 11, 14, 16), and it stands today as a textbook example of vicariance biogeography (17, 18).

Despite the current consensus, some continue to question ratite monophyly (13, 19) and the role of Gondwana in ratite distribution (20). It has long been recognized that adaptation to a flightless, cursorial lifestyle can result in morphological convergence or parallelism, especially in the postcranial skeleton (1, 19, 21). Such convergent adaptations might mislead phylogenetic inference based on morphology. In fact, one study based solely

Author contributions: J.H., E.L.B., M.J.B., S.J.H., R.T.K., W.S.M., F.H.S., and D.W.S. designed research; J.H., E.L.B., M.J.B., C.J.H., R.C.K.B., J.L.C., S.J.H., K.-L.H., R.T.K., B.D.M., K.J.M., S.R., S.J.S., C.C.W., and T.Y. performed research; J.H., E.L.B., M.J.B., and C.J.H. analyzed data; and J.H., E.L.B., and M.J.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: New DNA sequences are deposited in GenBank (accession nos. EU805776–EU805796, and EU822937). Alignments and trees have been deposited in TreeBase (study accession no. S2138).

[c]J.H., E.L.B., and M.J.B. contributed equally to this work.

[e]To whom correspondence should be addressed at: Department of Zoology, P.O. Box 118525, University of Florida, Gainesville, FL 32611. E-mail: ebraun68@ufl.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0803242105/DCSupplemental.

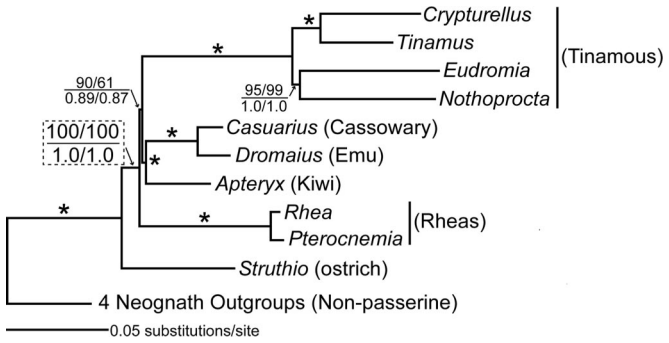

Fig. 1. Phylogenetic analyses of a 20-gene, 24-kb nuclear DNA dataset strongly supporting ratite polyphyly. All analyses used *Anas*, *Gallus*, *Buteo*, and *Ciconia* as outgroups. Branches for which all support measures were 100% or 1.0 are indicated with an asterisk; support for ratite polyphyly is highlighted. Topology obtained by using both partitioned (by locus) and unpartitioned ML and Bayesian analyses. Branch lengths reflect the unpartitioned ML analysis. Support measures are partitioned RAxML bootstrap (*Upper Left*), unpartitioned ML bootstrap (*Upper Right*), unpartitioned Bayesian posterior probability (*Lower Left*), and partitioned Bayesian posterior probability (*Lower Right*).

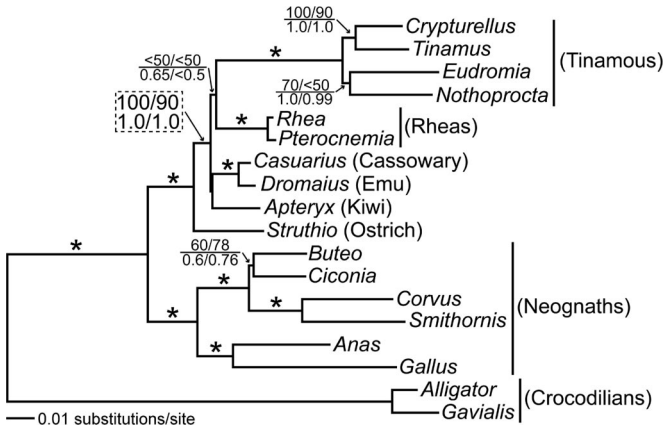

Fig. 2. Phylogenetic analyses including crocodilian outgroups and two passerine birds (*Corvus* and *Smithornis*) strongly support the conventional position of the avian root and ratite polyphyly. Analyses were conducted by using all sequences that could be aligned between crocodilians and birds (4,668 bp). Support measures are unpartitioned ML bootstrap (*Upper Left*), MP bootstrap (*Upper Right*), unpartitioned Bayesian posterior probability (*Lower Left*), and partitioned Bayesian posterior probability (*Lower Right*). Branch lengths shown reflect the unpartitioned ML analysis. Branches for which all support measures were 100% or 1.0 are indicated with an asterisk; the branch with no support values had <50% bootstrap support and <0.5 Bayesian posterior probability in all analyses. MP and ML analyses conducted after Y coding produced similar results (not shown).

on cranial characters, where convergence may be less likely, suggested that ratites are not monophyletic (13). These morphological results are provocative in light of recent sophisticated analyses of mitochondrial DNA (mtDNA) that show equivocal support for ratite monophyly (22, 23). Given the profound implications of this group for Gondwanan biogeography and the evolution of flightlessness, determining the true phylogeny of ratites is a key question in avian systematics.

Phylogenomic studies, which combine data from many genetic loci sampled to represent the genome, are proving useful in resolving difficult phylogenetic problems (24, 25). We assembled a dataset of 20 nuclear loci widely dispersed in the avian genome [supporting information (SI) Table S1] to examine ratite monophyly. It contains ≈30% protein-coding and 70% noncoding sequence (Table S1), taking advantage of the phylogenetic signal in archosaur noncoding sequences (26–28). The dataset comprises 18 taxa, including all extant ratite genera, four tinamou genera, and eight outgroup taxa (Table S2). Analyses of this dataset support a phylogeny in which paleognaths are monophyletic but ratites are not.

Results and Discussion

Nuclear DNA Sequences Strongly Support Ratite Polyphyly. All phylogenetic analyses of our 20-locus dataset revealed strong support for the ostrich as the sister group of all other paleognaths, placing the volant tinamous within ratites and making ratites polyphyletic.[o] When all loci were combined, the branch uniting tinamous with all ratites except the ostrich received 100% bootstrap support in maximum-likelihood (ML) analyses (Fig. 1), whether analyses were partitioned by locus (each locus assigned its own best-fit evolutionary model) or unpartitioned (a single model for all loci). Maximum-parsimony (MP) analyses produced similar results (Fig. S1*A*). The critical branch uniting non-ostrich paleognaths had a posterior probability of 1.00 in both partitioned and unpartitioned Bayesian analyses (Fig. 1) and highly significant support ($P < 0.001$) in the Shimodaira–Hasegawa (SH) test (see *SI Methods* and Fig. S2*C*).

Support for ratite polyphyly is robust to the assumptions of specific analyses. The critical branch was strongly supported in analyses using purine/pyrimidine (RY) coding (Fig. S1*A*), only protein-coding exons (Fig. S3) and different data partitioning schemes (see *SI Methods* and Fig. S1*B*). Furthermore, our conclusions do not reflect the specific set of outgroups used; analyses including crocodilians (Fig. 2) and/or up to 150 additional neognaths, representing all major living avian lineages, also support ratite polyphyly (25, 26, 30).

Separate analyses of individual loci show that 19 of 20 support paleognath monophyly in all analytical approaches (data not shown). The one that does not, *BDNF*, has a severe base compositional bias (see below). The great majority of single-locus trees (17 for ML and 15 for MP) support the ostrich as the sister group of all other paleognaths (Table 1). The probability of 15 or more of 20 independent gene trees agreeing by chance is extremely low ($P = 2 \times 10^{-9}$; binomial test using equiprobable trees null model). A number of loci that fail to support ratite polyphyly in individual analyses show hidden support (31, 32) in combined analyses (Table S3). Thus, the phylogenetic signal is widespread in the nuclear genome and any attempt to explain ratite polyphyly as an artifact must invoke a genome-wide systematic bias.

Rare Genomic Changes also Support Ratite Polyphyly. Beyond the strong signal present in nucleotide substitutions, three insertion/deletion events (indels) provide additional information regarding ratite phylogeny. Rare genomic changes like indels are thought to be valuable phylogenetic markers that may be free from a number of caveats that apply to nucleotide substitutions (33). The ostrich shares the ancestral character state with neognaths for two indels, an 8-bp deletion in *ALDOB* (Fig. 3) and a 9-bp insertion in *MYC* (Fig. S4). Tinamous share the derived character state with all other ratites, so both indels support the optimal tree found here. A single 1-bp indel in *CLTC* (Fig. S5) is found in ratites (including ostrich) but not tinamous, thus conflicting with the other indels and analyses of nucleotide substitutions. Vertebrate indels consistently exhibit less homoplasy than nucleotide substitutions (26, 34), and an exami-

[o]A group is polyphyletic if its defining characters are convergent (29). Ratites have long been defined by the absence of a keel on their sternum (e.g., ref. 3), a character related to flightlessness. Our analyses (see below) indicate the common ancestor of ratites was likely capable of flight and thus had a keeled sternum and was not a ratite.

EVOLUTION

Table 1. Bootstrap support for the crucial branch uniting a non-ostrich paleognath clade from ML and MP analyses of individual loci

	ML		MP	
Gene	Support	Conflict	Support	Conflict
ALDOB	64	—	63	—
BDNF	—	57*	—	91
CLTC	42	—	—	59
CLTCL1†	68	—	—	—
CRYAA	66	—	54	—
EEF2	99	—	88	—
EGR1	77	—	56	—
FGB	—	42	62	—
GH1	71	—	77	—
HMGN2	81	—	47	—
IRF2	60	—	71	—
MB	97	—	88	—
MUSK	95	—	91	—
MYC†	85	—	—	—
NGF	78	—	74	—
NTF3	83	—	52	—
PCBD1	92	—	76	—
RHO	78	—	65	—
TGFB2†	—	52	—	—
TPM1	47	—	69	—
Number of loci	17‡	3	15§	2

*ML analysis of *BDNF* does not support paleognath monophyly, instead placing the tinamous within neognaths.

†These genes have multiple MP trees, some of which support and some of which conflict with ratite polyphyly.

‡Binomial test, $P = 8 \times 10^{-15}$ (pure birth model) or 3×10^{-12} (equiprobable model).

§Binomial test, $P = 9 \times 10^{-12}$ (pure birth model) or 2×10^{-9} (equiprobable model).

nation of avian *FGB* intron 7 indels (34) revealed that 1-bp indels are more than twice as likely to exhibit homoplasy as longer (≥5-bp) indels. Thus, the likelihood of homoplasy in the 1-bp indel is much higher than the combined likelihood of homoplasy in the two longer indels that support ratite polyphyly.

Known Phylogenetic Artifacts Do Not Explain Ratite Polyphyly. Several distinct systematic biases that can mislead phylogenetic analysis have been characterized (35–38), and we determined whether any of these artifacts could explain ratite polyphyly.

Inconsistent phylogenetic estimation due to long-branch attraction (35) is a concern in phylogenomics. Superficially, it might appear problematic for paleognath phylogeny because tinamous have a high rate of molecular evolution (Table S4). Although the branch leading to tinamous is long (e.g., Fig. 1), it is not united with the other long branch on the tree, which leads to the outgroup. In fact, long-branch attraction would be expected to favor the conventional hypothesis of ratite monophyly rather than the surprising alternative of ratite polyphyly that we observed. We confirmed this expectation in two ways. First, a parametric bootstrap (Swofford–Olson–Waddell–Hillis, SOWH) test (39) rejected a null hypothesis of ratite monophyly ($P < 0.001$). Second, a series of 1,000 MP analyses in which the outgroup was replaced with a random sequence of similar base composition, simulating the longest possible branch (40), all rooted the tree within tinamous or along the branch leading to them. This confirmed our expectations and rejected long-branch attraction as an explanation for ratite polyphyly ($P < 0.001$).

Convergence in base composition can also create artifactual relationships (23). If the observed topology reflects compositional convergence, we would expect a tree derived from base-compositional information alone to have tinamous nested within ratites. When we clustered taxa using base-compositional distances, however, tinamous fell within ratites for few loci (Table S5). Instead, the most common grouping (10 of 20 loci) clustered ratites, a signal expected to reinforce the conventional topology. Furthermore, ratite polyphyly was supported both by analyses conducted after RY coding (Fig. S1*A*), which increases historical signal relative to compositional bias (37, 41), and ML analyses using a model allowing base compositional change (Table S6 and Fig. S1*B*). Base compositional convergence therefore cannot explain our results.

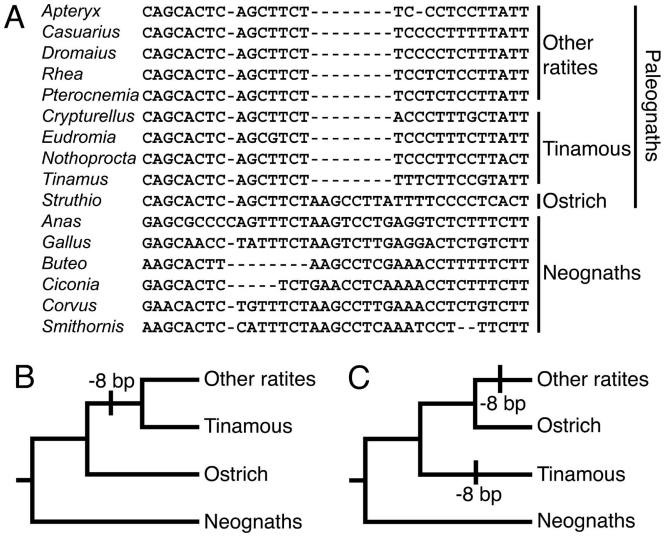

Fig. 3. An 8-bp deletion in *ALDOB* supports ratite polyphyly. (*A*) Alignment of the region around the informative deletion in *ALDOB* (positions 3213–3220). The ostrich shares its character state (+8 bp) with neognaths, whereas tinamous share the character state of all other ratites (−8 bp). (*B*) The distribution of character states can be mapped as a single deletion on the optimal topology found in this study. (*C*) The distribution requires at least two steps on the traditional topology (one possible reconstruction shown).

The one locus that failed to support paleognath monophyly was *BDNF*, which places tinamous within neognaths in ML analyses. However, *BDNF* has the greatest base-compositional variation and longest tinamou branch of all loci studied (Table S4). ML analyses of *BDNF* after RY coding do not place tinamous within neognaths (data not shown), suggesting that the anomalous results for that locus are artifactual.

Another potential problem occurs when individual gene trees differ from the species tree. Under some conditions, stochastic lineage sorting can produce a bias toward gene trees more symmetrical than the species tree (38), and the gene tree signal can predominate when many loci are analyzed (42). However, this is unlikely to occur unless the relevant branches are short relative to the coalescence time for the genes examined. The expected coalescence time for nuclear genes is twice the effective population size in generations. Because the relevant branch length and the ancestral paleognath population size are unknown, we assessed the potential impact of lineage sorting empirically. The gene-tree bias is weak (42), so it is unlikely to produce the consistency of single-gene trees we observed (Table 1). The aggregate probability of all trees with non-ostrich paleognath monophyly is 1/10, given the pure birth model for gene trees (42), so the probability of finding this result in 15 or more of 20 single-gene analyses is very small ($P = 9 \times 10^{-12}$). This calculation is conservative; it reflects the minimum number of loci supporting the relevant branch in single-gene MP analyses. More loci support non-ostrich paleognath monophyly using

ML (Table 1 and Table S6), and all loci that fail to support this group in single-gene ML analyses have positive partitioned hidden branch support in combined analyses (Table S3).

Phylogenetic analyses can be influenced by alignment (43), although our alignment protocol was conservative, and we excluded ambiguously aligned sites. We further tested the possibility of an alignment artifact by performing two analyses using only length-conservative regions: first, those regions that could be reliably aligned with distantly related crocodilian sequences (Fig. 2) and second, protein-coding exon sequences (Fig. S3). Both analyses strongly supported the non-ostrich paleognath clade, like the full data. In sum, our analyses indicate that the signal evident in the 20-locus dataset does not reflect any known phylogenetic artifact.

Monophyly of Australasian Ratites, Placement of Tinamous, and the Root of the Avian Tree. Other relationships strongly supported by our results include monophyly of rheas, tinamous, and an emu–cassowary clade (e.g., Fig. 1), in agreement with previous studies. Monophyly of extant Australasian ratites (kiwis, emus, and cassowaries) is also strongly supported (Fig. 1, Fig. S1), in agreement with previous molecular analyses (9–11, 14) but contrary to some morphological analyses (9, 12). Kiwis may be difficult to place in morphological phylogenies because of the striking differences between kiwis and other ratites in life history, behavior, and size (44). The molecular support for nesting kiwis within ratites is especially interesting given suggestions that kiwis were derived directly from volant paleognaths (45).

Although placement of tinamous within the ratites is strongly supported, the sister group of tinamous is unclear. ML and Bayesian analyses place tinamous sister to Australasian ratites (Fig. 1, Fig. S1*B*), whereas MP and RY-coded ML analyses place tinamous sister to rheas (Fig. S1*A*). The alternative trees do not differ significantly when the SH test is used (Fig. S2), suggesting that the topology is sensitive to the model of sequence evolution applied. The second alternative is most parsimonious from a biogeographic standpoint, because both tinamous and rheas are exclusively Neotropical.

Our large nuclear dataset allowed us to address a controversy regarding the root of the avian tree and examine the impact of the root on our conclusions. Although the traditional view is that the root lies between paleognath and neognath clades (12, 27, 46), early analyses of mtDNA placed the root either between passerines and all other birds (47) or within passerines (48, 49), contradicting neognath monophyly. More sophisticated analyses (e.g., RY-coding) of mtDNA data strongly support the traditional rooting (22, 23, 50, 51), unlike the early analyses. Some morphological studies also suggest nonmonophyly of paleognaths (8, 45). Given these questions, we analyzed our dataset using several different methods and confirmed that the position of the root lies between paleognath and neognath clades (Fig. 2).

Contrasts with Previous Studies. The strong support for ratite polyphyly in our study raises an important question: Why have most modern analyses supported ratite monophyly? We cannot fully answer this question, but we can offer plausible hypotheses. A study based on a single, short nuclear intron (52) showed limited support for ratite monophyly, probably reflecting limited power. DNA–DNA hybridization (14) required extrapolation of distances far beyond the useful range of the method. Early studies using mtDNA strongly supported ratite monophyly (9–11), but recent analyses using more sophisticated methods revealed that support for a ratite clade is weak (22, 23). Avian mtDNA evolves rapidly and exhibits high among-site rate variation (10, 11), making analyses sensitive to the evolutionary model and taxon sample used (23). Furthermore, tinamous have a much higher mtDNA evolutionary rate than other paleognaths (11), so long-branch attraction uniting tinamous with the long branch leading to the outgroup is likely (as described above). Thus, we do not consider mtDNA to be in strong conflict with ratite polyphyly; indeed, the ambivalent signal evident in recent analyses (22, 23) suggests more sophisticated analyses of mtDNA may ultimately support polyphyly.

Several morphological studies strongly support ratite monophyly (9, 12) in conflict with our nuclear genetic data. Simply collecting more data or combining molecular and morphological data are unlikely to resolve the impasse generated by these incongruent signals (53). Both signals cannot represent evolutionary relationships; at least one must be nonphylogenetic. Cranial morphology offers a potential solution, because two studies using cranial characters (13, 19) agree with our data. There is no reason to expect convergence in the cranial characters of tinamous and any particular subset of ratites. However, convergence is common in the postcranial skeletons of other flightless birds (54), and morphological convergence certainly can mislead phylogenetic analysis (55). Flightlessness is expected to result in reduction or loss of the sternal keel; reduction in size, complexity, and number of wing bones; increase in size of leg bones; and nonaerodynamic changes in plumage structure. Because the volant ancestors of each ratite lineage may have been morphologically similar, parallel evolution could have produced some ratite traits identical in state but not by descent.

Evolution of Flightlessness. Any topology that nests the volant tinamous within the flightless ratites requires either multiple losses of flight or a loss of flight in the ancestral paleognath and a regain in tinamous. Although loss and regain is more parsimonious if both transitions are equally probable, multiple losses of flight are more likely. Flight has been lost in members of 18 extant bird families, many more times in extinct groups, and hundreds of times in the family Rallidae alone (21, 54, 56). Thus, the loss of flight is much more probable than gain. Given the position of tinamous in either optimal tree based on the complete dataset (Fig. 1, Fig. S1*A*), flight must have been lost independently at least three times, in ostriches, rheas, and Australasian ratites. A scenario in which tinamous regained flight would be even more interesting, but there are no examples of avian lineages that have lost and regained flight.

Biogeographic Implications. There have been many proposals relating divergences among living ratites to the breakup of Gondwana (10, 11, 14, 16), but all assume ratite monophyly. Our phylogeny actually fits a strictly vicariant hypothesis better than previous phylogenies. Africa was the first piece of Gondwana to separate (16), and the African ostriches are the first clade to diverge (e.g., Fig. 1). Nevertheless, no proposed phylogeny, ours included, can be explained entirely by the order of separation of Gondwanan fragments.

Multiple losses of flight, with the implication of greater dispersal capability for ancestral paleognaths, make a strictly vicariant model less compelling. The existence of volant paleognaths in the Paleogene of Europe and North America also suggests that dispersal must be considered (45, 57). Dispersal of ratites is further suggested by phylogenies in which the extinct moas of New Zealand are not sister to the extant kiwis (2, 10, 11), as would be predicted by strict vicariance. Thus, fossil data confirm that simple vicariant models can be rejected. It may be possible to distinguish among evolutionary scenarios with time-calibrated trees, but the large differences in evolutionary rates evident in our analyses (e.g., Fig. 1) combined with the paucity of good fossil calibration points for paleognaths render this task difficult.

Conclusions

Exhaustive analyses of DNA sequence data from 20 unlinked nuclear genes provide strong evidence that ratites are polyphyl-

etic. We have discovered a robust genome-wide signal that is not associated with any known phylogenetic artifact. We believe this phylogeny resolves a debate on ratite origins that began in the time of Huxley and Owen (3–5). Our phylogeny implies that the numerous striking similarities associated with flightlessness (1) had independent origins in various ratite lineages. Thus, the flightless ratites are living evidence of parallel evolutionary trajectories from flighted ancestors. The possibility that multiple, unique developmental genetic pathways underlie the ratite form should be tested in light of this new phylogenetic hypothesis. Finally, our phylogeny removes the need to postulate vicariance by continental drift to explain ratite distribution. Although that theory seemed to represent a consilience between evolutionary biology and geology, it was never completely consistent either with any published phylogeny or the existence of paleognath fossils in the Northern hemisphere (45, 57). Perhaps the impact of our phylogeny should be viewed as yet another example of the phenomenon that Huxley called ''the great tragedy of science—the slaying of a beautiful theory by an ugly fact.''

Methods

Sequencing and Alignment of Nuclear Loci. We amplified and sequenced 20 nuclear loci (see *SI Methods* for detailed methods). Sequences (accession numbers listed in Table S7) were aligned manually (see *SI Methods*); ambiguously aligned regions and sparsely sampled sites (those not present in at least four birds and three paleognaths) were excluded from analyses.

Phylogenetic Analyses. Four nonpasserine neognaths were used as outgroups in most analyses of the 14 taxon, 23,902-bp dataset (e.g., Fig. 1). Crocodilians and passerines were included only in analyses designed to test the position of the root of the avian tree using the subset of sequences for which crocodilian and avian sequences could be aligned (4,668 bp; Fig. 2).

Optimal MP and ML trees using unpartitioned data were identified by using PAUP* 4.0b10 (58). MP analyses used branch-and-bound searches with equally weighted characters and assessed support using 1,000 bootstrap replicates. The appropriate nucleotide substitution model (Table S8) for ML analyses was determined by using the Akaike information criterion and Modeltest 3.6 (59) or the models appropriate for RY-coded data (see *SI Methods*). ML analyses used heuristic searches with 10 random addition sequence replicates and tree bisection and reconnection (TBR) branch swapping; support was assessed by using 100 bootstrap replicates.

Three ML analyses with the data partitioned by locus were conducted. First, a partitioned RAxML (60) analysis, using the general time-reversible (GTR)+Γ model with distinct parameter values for each partition and linked branch length parameters, was conducted and support assessed by using 100 bootstrap replicates. The other two analyses used PAUP* and nhPhyML (61), allowing use of a more diverse set of nucleotide substitution models. Branch lengths were unlinked and monophyly of well established clades was assumed *a priori*, focusing on a ''plausible set'' of 315 trees (105 arrangements of the five major paleognath groups and three arrangements within tinamous; Fig. S2). The log likelihoods of all partitions for each tree were summed after being calculated using PAUP* (Table S9) or nhPhyML (not shown). PAUP* analyses used the best-fit model for each locus (Table S8) and nhPhyML analyses used the Galtier and Gouy (62) model, which allows the GC content to vary across the tree. In these analyses, support was assessed by using 1,000 resampling of estimated log-likelihoods (RELL) (63) bootstrap replicates.

Bayesian Markov chain Monte Carlo analyses were performed by using MrBayes 3.1.1 (64). We ran four chains for 10 million generations, sampling every 500 generations and discarding the first 500 trees, with other run parameters set at defaults. Partitioned Bayesian analyses used linked branch lengths.

The relationship between the probability that a clade is correct and bootstrap support is complex (65, 66), but the bootstrap is conservative under many circumstances. In contrast, Bayesian posterior probabilities can overestimate the probability that a clade is correct (e.g., 66). We also assessed branch (Bremer) support as well as hidden support and conflict for specific branches (31, 32).

Tests of Topologies, Base Composition, and Individual Gene Trees. We used two topology tests appropriate for comparisons of trees that were not specified *a priori* (67) to examine the position of the ostrich. The SH test (68) was performed by using the plausible set of 315 distinct trees (Fig. S2) because the test requires the inclusion of all trees that can be entertained as the true topology (67). The parametric bootstrap (SOWH) test (39) compares the difference in optimality scores for the empirical data on the optimal tree and a null hypothesis tree to a distribution generated by simulation. The null hypothesis topology was the most likely one with ratite monophyly (tree 19a of Fig. S2); 1,000 simulated datasets were analyzed by using MP and ML (*SI Methods*).

Base-compositional clustering was performed by using minimum evolution in PAUP* with a matrix of Euclidean distances between base-composition vectors for each taxon (see *SI Methods*). Topologies examined were limited to the plausible set (Fig. S2).

Individual gene trees may differ from the species tree. We tested whether the observed number of gene trees showing monophyly of non-ostrich paleognaths was unexpected given the null hypothesis of a completely polytomous species tree. We used binomial tests and two appropriate models of tree probability to assess this possibility (*SI Methods*).

Identification of Informative Indels. To identify low-homoplasy indels, all gaps in a 19-gene, 171-taxon dataset (25) were coded by using the simple gap coding method (69). These indels were mapped by using MP on the ML tree for the 171-taxon dataset (25) and the same tree rearranged so ratites were monophyletic. Indels mapping unambiguously on the branch of interest with a consistency index $\geq$0.5 were examined further, and ambiguously aligned regions were removed. *CLTCL1* was not included in the 171-taxon analysis (25), so it was examined independently by using a similar methodology.

ACKNOWLEDGMENTS. We thank D. L. Swofford, P. O. Lewis, and D. J. Zwickl for advice on phylogenetic analyses; D. J. Levey, K. Braun, and anonymous reviewers for helpful comments; and A. B. Thistle for a technical edit of the manuscript. For the loan of samples, we thank the Field Museum of Natural History, the University of Kansas, Louisiana State University, the United States National Museum of Natural History, L. D. Densmore, and the collectors (Table S2). This work was supported by the U.S. National Science Foundation Assembling the Tree of Life Program Grants DEB-0228675 (to S.J.H.); DEB-0228682 (to R.T.K., E.L.B., and D.W.S.); DEB-0228688 (to F.H.S.); and DEB-0228617 (to W.S.M.).

1. Davies SJJF (2002) *Ratites and Tinamous: Tinamidae, Rheidae, Dromaiidae, Casuariidae, Apterygidae, Struthionidae* (Oxford Univ Press, Oxford).
2. Harshman J (2007) in *Reproductive Biology and Phylogeny of Birds*, ed Jamieson BGM (Science Publishers, Enfield, NH), pp 1–35.
3. Huxley TH (1867) On the classification of birds; and on the taxonomic value of the modifications of certain of the cranial bones observable in that class. *Proc Zool Soc London* 145:415–472.
4. Pycraft WP (1900) Part II. On the morphology and phylogeny of the Palaeognathae (Ratitae and Crypturi) and Neognathae (Carinatae). *Trans Zool Soc London* 15:149–290.
5. Owen R (1866) *The Anatomy of Vertebrates, Vol. 2: Birds and Mammals* (Longmans, Green, London).
6. Mayr E, Amadon D (1951) A classification of recent birds. *Am Mus Novit* 1496: 1–42.
7. DeBeer G (1956) The evolution of ratites. *Bull Br Mus (Nat Hist) Zool* 4:57–76.
8. Olson SL (1985) The fossil record of birds. *Avian Biol* VIII:79–238.
9. Lee K, Feinstein J, Cracraft J (1997) in *Avian Molecular Evolution and Systematics*, ed Mindell DP (Academic, San Diego), pp 173–211.
10. Cooper A, *et al.* (2001) Complete mitochondrial genome sequences of two extinct moas clarify ratite evolution. *Nature* 409:704–707.
11. Haddrath O, Baker AJ (2001) Complete mitochondrial DNA genome sequences of extinct birds: Ratite phylogenetics and the vicariance biogeography hypothesis. *Proc R Soc London Ser B* 268:939–945.
12. Livezey BC, Zusi RL (2007) Higher-order phylogeny of modern birds (Theropoda, Aves: Neornithes) based on comparative anatomy. II. Analysis and discussion. *Zool J Linn Soc* 149:1–95.
13. Elzanowski A (1995) Cretaceous birds and avian phylogeny. *Cour Forsch Inst Senck* 181:37–53.
14. Sibley CG, Ahlquist JA (1990) *Phylogeny and Classification of Birds* (Yale Univ Press, New Haven, CT).
15. Cracraft J (1973) Continental drift, palaeoclimatology, and the evolution and biogeography of birds. *J Zool* 169:455–545.
16. Cracraft J (2001) Avian evolution, Gondwana biogeography and the Cretaceous-Tertiary mass extinction event. *Proc R Soc London Ser B* 268:459–469.
17. Futuyma DJ (2005) *Evolution* (Sinauer, Sunderland, MA).
18. Gill F (2007) *Ornithology* (Freeman, New York).
19. Bock WJ, Bühler P (1990) in *Proceedings of the International Centennial Meeting of the Deutsche Ornithologen-Gesellschaft*, eds van den Elzen R, Schuchmann K-L, Schmidt-Koenig K (Verlag der Deutschen Ornithologen-Gesellschaft, Bonn, Germany), pp 31–36.

20. Briggs JC (2003) Fishes and birds: Gondwana life rafts reconsidered. *Syst Biol* 52:548–553.
21. Feduccia A (1996) *The Origin and Evolution of Birds* (Yale Univ Press, New Haven, CT).
22. Slack KE, Delsuc F, Mclenachan PA, Arnason U, Penny D (2007) Resolving the root of the avian mitogenomic tree by breaking up long branches. *Mol Phylogenet Evol* 42:1–13.
23. Braun EL, Kimball RT (2002) Examining basal avian divergences with mitochondrial sequences: Model complexity, taxon sampling, and sequence length. *Syst Biol* 51:614–625.
24. Delsuc F, Brinkmann H, Phillippe H (2005) Phylogenomics and the reconstruction of the tree of life. *Nat Rev Genet* 6:361–375.
25. Hackett SJ, *et al.* (2008) A phylogenomic study of birds reveals their evolutionary history. *Science* 320:1763–1768.
26. Chojnowski JL, Kimball RT, Braun EL (2008) Introns outperform exons in analyses of basal avian phylogeny using clathrin heavy chain genes. *Gene* 410:89–96.
27. Harshman J, Huddleston CJ, Bollback J, Parsons TM, Braun MJ (2003) True and false gharials: A nuclear gene phylogeny of Crocodylia. *Syst Biol* 52:386–402.
28. Prychitko TM, Moore WS (1997) The utility of DNA sequences of an intron from the β-fibrinogen gene in phylogenetic analysis of woodpeckers (Aves: Picidae). *Mol Phylogenet Evol* 8:193–204.
29. Wiley EO (1981) *Phylogenetics: The Theory and Practice of Phylogenetic Systematics* (Wiley, New York).
30. Yuri T, Kimball RT, Braun EL, Braun MJ (2008) Duplication and accelerated evolution of growth hormone gene in passerine birds. *Mol Biol Evol* 25:352–361.
31. Gatesy J, O'Grady P, Baker RH (1999) Corroboration among data sets in simultaneous analysis: Hidden support for phylogenetic relationships among higher level artiodactyl taxa. *Cladistics* 15:271–313.
32. Lee MSY, Hugall AF (2003) Partitioned likelihood support and the evaluation of data set conflict. *Syst Biol* 52:15–22.
33. Rokas A, Holland PWH (2000) Rare genomic changes as a tool for phylogenetics. *Trends Ecol Evol* 15:454–459.
34. Fain MG, Houde P (2004) Parallel radiations in the primary clades of birds. *Evolution (Lawrence, Kans)* 58:2558–2573.
35. Felsenstein J (1978) Cases in which parsimony or compatibility methods will be positively misleading. *Syst Zool* 27:401–410.
36. Phillips MJ, Delsuc F, Penny D (2004) Genome-scale phylogeny and the detection of systematic biases. *Mol Biol Evol* 21:1455–1458.
37. Jeffroy O, Brinkmann H, Delsuc F, Phillippe H (2006) Phylogenomics: The beginning of incongruence? *Trends Genet* 22:225–231.
38. Degnan JH, Rosenberg NA (2006) Discordance of species trees with their most likely gene trees. *PLoS Genet* 2:e68.
39. Swofford DL, Olson GJ, Waddell PJ, Hillis DM (1996) in *Molecular Systematics*, eds Hillis DM, Moritz C, Mable BK (Sinauer, Sunderland, MA), pp 407–514.
40. Sullivan J, Swofford DL (1997) Are guinea pigs rodents? The importance of adequate models in molecular phylogenetics. *J Mamm Evol* 4:77–86.
41. Phillips MJ, Penny D (2003) The root of the mammalian tree inferred from whole mitochondrial genomes. *Mol Phylogenet Evol* 28:171–185.
42. Kubatko LS, Degnan JH (2007) Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Syst Biol* 56:17–24.
43. Lake JA (1991) The order of sequence alignment can bias the selection of tree topology. *Mol Biol Evol* 8:378–385.
44. Fuller E (1991) *Kiwis* (Swan–Hill, Shrewsbury, UK).
45. Houde P (1988) *Paleognathous Birds from the Early Tertiary of the Northern Hemisphere* (Nuttall Ornithological Club, Cambridge, MA).
46. Groth JG, Barrowclough GF (1999) Basal divergences in birds and the phylogenetic utility of the nuclear RAG-1 gene. *Mol Phylogenet Evol* 12:115–123.
47. Härlid A, Arnason U (1999) Analyses of mitochondrial DNA nest ratite birds within the Neognathae: Supporting a neotenous origin of ratite morphological characters. *Proc R Soc London Ser B* 266:305–309.
48. Mindell DP, *et al.* (1999) Interordinal relationships of birds and other reptiles based on whole mitochondrial genomes. *Syst Biol* 48:138–152.
49. Johnson KP (2001) Taxon sampling and the phylogenetic position of Passeriformes: Evidence from 916 avian cytochrome *b* sequences. *Syst Biol* 50:128–136.
50. Harrison GL, *et al.* (2004) Four new avian mitochondrial genomes help get to basic evolutionary questions in the Late Cretaceous. *Mol Biol Evol* 21:974–983.
51. Paton T, Haddrath O, Baker AJ (2002) Complete mitochondrial DNA genome sequences show that modern birds are not descended from transitional shorebirds. *Proc R Soc London Ser B* 269:839–846.
52. García-Moreno J, Mindell DP (2000) Rooting a phylogeny with homologous genes on opposite sex chromosomes (gametologs): A case study using avian CHD. *Mol Biol Evol* 17:1826–1832.
53. Bull JJ, Huelsenbeck JP, Cunningham CW, Swofford DL, Waddell PJ (1993) Partitioning and combining data in phylogenetic analysis. *Syst Biol* 42:384–397.
54. Steadman DW (2006) *Extinction and Biogeography of Tropical Pacific Birds* (Univ of Chicago Press, Chicago).
55. McCracken KG, Harshman J, McClellan DA, Afton AD (1999) Data set incongruence and correlated character evolution: An example of functional convergence in the hind-limbs of stifftail diving ducks. *Syst Biol* 48:683–714.
56. Roff DA (1994) The evolution of flightlessness: Is history important? *Evol Ecol* 8:639–657.
57. Dyke G (2003) in *Telling the Evolutionary Time: Molecular Clocks and the Fossil Record*, eds Smith P, Donoghue P (Taylor and Francis, London), pp 263–278.
58. Swofford DL (2003) *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods), Version 4* (Sinauer, Sunderland, MA).
59. Posada D, Crandall KA (1998) MODELTEST: Testing the model of DNA substitution. *Bioinformatics* 14:817–818.
60. Stamatakis A (2006) RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690.
61. Boussau B, Gouy M (2006) Efficient likelihood computations with nonreversible models of evolution. *Syst Biol* 55:756–768.
62. Galtier N, Gouy M (1998) Inferring pattern and process: Maximum-likelihood implementation of a nonhomogeneous model of DNA sequence evolution for phylogenetic analysis. *Mol Biol Evol* 15:871–879.
63. Kishino H, Miyata T, Hasegawa M (1990) Maximum likelihood inference of protein phylogeny and the origin of chloroplasts. *J Mol Evol* 31:151–160.
64. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574.
65. Felsenstein J (2004) *Inferring Phylogenies* (Sinauer, Sunderland, MA).
66. Alfaro ME, Holder MT (2007) The posterior and the prior in Bayesian phylogenetics. *Annu Rev Ecol Evol Syst* 37:19–42.
67. Goldman N, Anderson JP, Rodrigo AG (2000) Likelihood-based tests of topologies in phylogenetics. *Syst Biol* 49:652–670.
68. Shimodaira H, Hasegawa M (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol Biol Evol* 16:1114–1116.
69. Simmons MP, Ochoterena H (2000) Gaps as characters in sequence-based phylogenetic analyses. *Syst Biol* 49:369–381.