Phylogeny and Biogeography of Ratite Birds Inferred from DNA Sequences of the Mitochondrial Ribosomal Genes

Marcel van Tuinen,* Charles G. Sibley,† and S. Blair Hedges*

*Department of Biology and Institute of Molecular Evolutionary Genetics, Pennsylvania State University; and †Santa Rosa, California

The origin of the flightless ratite birds of the southern continents has been debated for over a century. Whether dispersal or vicariance (continental breakup) best explains their origin depends largely on their phylogenetic relationships. No consensus has been reached on this issue despite many morphological and molecular studies. To address this question further we sequenced a 2.8-kb region of mitochondrial DNA containing the ribosomal genes in representative ratites and a tinamou. Phylogenetic analyses indicate that Struthio (Africa) is basal and Rhea (South America) clusters with living Australasian ratites. This phylogeny agrees with transferrin and DNA hybridization studies but not with sequence analyses of some protein-coding genes. These results also require reevaluation of the phylogenetic position of the extinct moas of New Zealand. We propose a new hypothesis for the origin of ratites that combines elements of dispersal and vicariance.

Introduction

The living ratites include two species of ostriches (Struthio) in Africa and formerly in Asia, the Australian emu (Dromaius), three species of cassowaries (Casuarius) in New Guinea and northeastern Australia, three species of forest-dwelling kiwis (Apteryx) in New Zealand, and two rheas (Rhea) in South America (Sibley 1996). All lack a keel on the sternum, a character associated with flightlessness. Based on anatomical analyses, ratite phylogeny has been controversial for over a century. This stems from subjective interpretations of anatomical characters and from difficulties in determining polarity in characters shared by ratite birds (Kurochkin 1995). Sibley and Ahlquist (1990) provide a historical review of ratite systematics.

Over the last two decades, several molecular studies of ratites have clarified some aspects of ratite phylogeny. There is general agreement that living ratites are monophyletic and that the weakly flying tinamous are their closest living relatives (Prager et al. 1976; Sibley and Ahlquist 1981, 1990; Caspers, Wattel, and De Jong 1994). Monophyly of the living Australasian ratites is also supported by molecular data (Prager et al. 1976; Sibley and Ahlquist 1981, 1990; Cooper et al. 1992), although consensus on this issue has not been reached with anatomical data (Cracraft 1974; Bledsoe 1988; Kurochkin 1995).

The relationships of the three major lineages, rheas, ostriches, and Australasian ratites, have a direct bearing on the biogeographic history of the ratites (Cracraft 1974), but these relationships remain unclear. Analysis of transferrin immunological data grouped the Rhea with the Australasian clade and thus identified Struthio as the most basal living ratite (Prager et al. 1976). The relationships of these three lineages with DNA-DNA hybridization data were considered to be unresolved because a UPGMA tree joined Rhea and Struthio, whereas trees constructed with the Fitch-Margoliash (1967) algorithm joined Rhea with the Australasian clade (Sibley and Ahlquist 1990). In contrast, mitochondrial DNA sequence data (12S rRNA; 400 bp) supported a basal position for Rhea (Cooper et al. 1992). However, reanalysis of those sequence data by Rzhetsky, Kumar, and Nei (1995) and analysis of an additional 600 bp from a nuclear gene (Cooper and Penny 1997) indicate that such small data sets are unlikely to resolve the early history of ratites. Therefore, we sequenced the complete mitochondrial ribosomal genes (2.8 kb) from representative ratite species to address the question of the nearest living relative of the Australasian ratites. These genes have proven useful for addressing higher-level phylogenetic questions in birds and other vertebrates (e.g., Hedges 1994; Hedges and Sibley 1994; Springer et al. 1997).

Materials and Methods

A 2.8-kb region of mitochondrial DNA was sequenced for each of three ratite species (Rhea americana, Struthio camelus, and Dromaius novaehollandiae) and a gray tinamou (Tinamus tao). We assumed monophyly of the living Australasian ratites based on molecular evidence (Prager et al. 1976; Sibley and Ahlquist 1990; Cooper et al. 1992), with Dromaius as the representative of the living Australasian ratite clade. The sequenced region includes the entire 12S rRNA, tRNAVal, and 16S rRNA genes. The corresponding sequences from a domestic fowl (Gallus gallus; accession number X52392, sites 1274–3994 in Desjardins and Morais 1990) and an American alligator (Alligator mississippiensis; accession number L28074) were obtained from GenBank for comparison. The new sequences reported here have been deposited in EMBL with accession numbers AJ002921–AJ002924.

DNA amplification was performed with the following primer pairs: 12L9/12H2, 12L9/12H5, 12L10/12H5, 12L1/12H4, 12L7/16H11, 16L11/16H5, 16L17/16H17, 16L10/16H10, 16L9/16H3, 16L8/16H13, and 16L4/16H12 (Hedges 1994; Hedges and Sibley 1994; Hedges et al. 1995). Primers not previously described are
[IUPAC]: 12L7 (GAA GGW GGA TTT AGY AGT AAA), 12L9 (AAA GCA HRK CA C TGA ARA TGY YDA GA), 12L10 (CMC AMG GGA MWC AGT GAT WAA HAT T), 12H5 (TTA GAG GAG CCT GTC CTA TAA TCG), 16L17 (CCW AMC GAR CYT RGT GAT AGC TGG TT), and 16H17 (TGT TTA CCA AAA ACA TMY CCY YYS GC).

DNA was PCR-amplified as follows: 30–40 cycles (94°C for 15 s, 50°C or 55°C for 15 s, 72°C for 45 s). Double-stranded DNA was gel-purified and served as template for another PCR run under slightly different conditions: 25–35 cycles (94°C for 15 s, 55°C or 60°C for 15 s, 72°C for 45 s). A hot start was performed at 80°C for all PCR runs. Prior to sequencing, double-stranded DNA was filtered as described earlier (Hedges and Sibley 1994). Both complementary L- and H-strands were sequenced for all primer sets. Cycle sequencing reactions were performed using 3'-dye-labeled dideoxynucleotide triphosphates (fluorescent dye terminators) and run on an ABI PRISM 377 DNA Sequencer (Perkin-Elmer ABI, Foster City, Calif.).

Alignments were performed with ESEE (Cabot and Beckenbach 1989). For phylogenetic analysis, neighbor-joining (Saitou and Nei 1987) analyses were performed in MEGA (Kumar, Tamura, and Nei 1994) using a Kimura (1980) two-parameter distance with transitions and transversions included unless otherwise noted. Maximum-parsimony analyses were conducted with PAUP (Swoford 1993), and maximum-likelihood analyses were performed with Nucml (HKY) in MOLPHY (Adachi and Hasegawa 1996). In MOLPHY, we used the default transition: transversion ratio of 4:1, which is similar to that estimated for mitochondrial DNA (Kumar 1996), and 100:1 (transversions only). Sites corresponding to alignment gaps were excluded. The interior-branch (SE) test was performed to assess confidence in the reliability of the neighbor-joining trees by testing if interior branch lengths deviate significantly from zero. In PAUP, the bootstrap method was applied (Felsenstein 1985) with 2,000 replications.

Results

Out of 2,900 sites in the initial alignment of six taxa, 2,429 were analyzed after elimination of gaps and sites showing ambiguous alignment. Of those sites, there were 1,173 varied sites and 409 sites informative for the parsimony method. The phylogenetic tree (fig. 1), rooted with Alligator, shows high confidence for a Rhea-Dromaius clade. Confidence values are significant (99%) for the interior-branch test and almost significant when only transversions are used (92%). There is significant confidence for ratite monophyly, using transversions plus transversions or only transversions. Maximum-likelihood analyses resulted in 98%–99% bootstrap support for the Rhea-Dromaius clade.

When treated separately, the tRNAVal, 12S, and 16S rRNA genes support the same tree, but with lower confidence values than the combined data set (table 1). The 16S rRNA data give the strongest support for a Rhea-Dromaius clade (99%, interior branch test), which is in accordance with the number of varied sites and those informative for parsimony for the three genes (16S: 705 varied sites, 262 sites informative for parsimony; 12S: 440 varied sites, 135 sites informative for parsimony; tRNAVal: 28 varied sites, 12 sites informative for parsimony). The 12S rRNA and tRNAVal data support the Rhea-Dromaius clade, with 71% and 70% confidence values, respectively, for the interior-branch analysis.

After this study was completed, a similar study by Lee, Feinstein, and Cracraft (1997) was published on ratite relationships inferred from DNA sequences of several noncoding and protein-coding mitochondrial genes, including a reexamination of the morphological evidence. There was partial overlap between the sequenced regions (rRNA genes) in their study and ours. Comparison of our Struthio sequence with their sequence and that of Härild, Janke, and Arnason (1997) indicated 11–12 nucleotide differences between each of the three sequences. All differences between our sequence and that of Härild, Janke, and Arnason (1997) were transitions, whereas both of those sequences shared six transversion differences with the Struthio sequence of Lee, Feinstein, and Cracraft (1997). These differences did not affect the phylogeny as shown in figure 1. However, the sequence analysis (combined genes) of Lee, Feinstein, and Cracraft (1997) supported a basal Rhea (rather than Struthio) lineage, an unexpected discordance that we investigated here.

An overview of the present molecular evidence (table 1) shows an incongruent pattern regarding the closest relative of Australasian ratites. At the nucleotide level, the available sequences from protein-coding genes suggest that Struthio is the closest relative. However, this topology is significant only in the cytochrome b and c-mos genes and, in those cases, only in a transversion analysis. Amino acid analysis of these genes (separately and combined) does not resolve the question of the position of Rhea and Struthio within the ratites (results from combined genes shown in table 1).
Molecular Evidence for the Closest Living Relatives of the Australasian Ratites

As opposed to a *Struthio*-Australasian signal in the cytochrome *b* and nuclear *c-mos* genes, the noncoding genes indicate a *Rhea*-Australasian clade. This result agrees with the transferrin data (Prager et al. 1976) and with one analysis of the DNA hybridization data (Sibley and Ahlquist 1990, fig. 326). The transferrin and DNA hybridization data sets were analyzed with the Fitch-Margoliash (1967) and UPGMA (Sneath and Sokal 1973) methods. We reanalyzed these data sets and constructed neighbor-joining trees. The transferrin tree (fig. 2A) and the DNA hybridization distance tree (fig. 2B), are both concordant with the results from noncoding genes (fig. 1). *Apteryx* of New Zealand is the closest living relative of a clade containing *Casuarius* of New Guinea and Australia and *Dromaius* of Australia. Together, they form a monophyletic Australasian clade. The South American *Rhea* is the closest relative of the Australasian clade, and the African *Struthio* is the basal lineage of living ratites.

Our tree (fig. 1) also indicates that tinamous are the closest relatives of ratites (99% interior branch test). This conclusion already was reached from DNA hybridization results (Sibley and Ahlquist 1990) and protein data (alpha-crystallin A, Stapel et al. 1984; Caspers, Wattel, and De Jong 1994) using larger numbers of taxa.

Discussion

Although the relationships of the ratites in our sequence analysis (fig. 1) agree with transferrin and DNA hybridization results (fig. 2), the significant disagreement with other sequence analyses was unexpected. Although it might be tempting to consider the larger (5.2 kb) study of Lee, Feinstein, and Cracraft (1997) more representative of the signal from sequence data than our smaller (2.8 kb) data set, especially because of overlap in sequenced regions, our separate gene analysis (table 1) shows that assumption to be incorrect. The signal for a close relationship of the African (*Struthio*) and Australasian ratites is only significant in the transversion analyses of cytochrome *b* and *c-mos*. The results of most individual gene analyses are not significant, but a difference between coding versus noncoding genes is evident, and that difference is significant when genes from each group are combined (table 1).

Lee, Feinstein, and Cracraft (1997) found that their molecular tree was sensitive to the outgroup used. When *Gallus* was omitted, and only the tinamous were included as outgroup, the statistical confidence dropped considerably. Moreover, when 58 morphological characters were added to the molecular data, the resulting tree (*Apteryx* basal) was identical to the tree from only morphological data, indicating that the molecular “signal” was not strong. However, they pointed out that the difference was a rooting problem and that their unrooted morphological and molecular trees were identical. They also noted that the internodal branch lengths of their tree were short compared with the terminal branches, suggesting (under the assumption of a molecular clock) that early divergences within the ratites occurred during a relatively short period. This was also found by Sibley and Ahlquist (1990) and is evident in our tree of the ribosomal genes (fig. 1). Such short internodal distances may explain the difficulties in resolving ratite relationships.

Without collecting additional sequence data, it is not possible to reconcile the significant differences be-
Phylogeny of Ratites

Fig. 2.—Neighbor-joining trees of ratite birds based on two different distance data sets. A, transferrin immunological distances (I.D.; Prager et al. 1976). B, DNA hybridization (Tm) distances (Sibley and Ahlquist 1990). Reciprocal values were averaged and analyzed in MEGA. A tinamou (Eudromia elegans, Nothoprocta perdicaria) and/or domestic fowl (Gallus gallus) were included for comparison.

Between analyses of coding and noncoding genes, except to point out concordant patterns among other independent data sets. The agreement between transferrin data, DNA hybridization, and the sequence results from noncoding genes suggests that the topology obtained by the coding genes (Rhea basal) may be incorrect. Further support for that assumption comes from recent criticism of the use of cytochrome b for divergences earlier than the Miocene (Moore and DeFilippis 1997). Moore and DeFilippis suggest that transversion saturation, base composition bias, and rate variation among lineages may contribute to problems with resolving avian relationships above the family level. Whatever the reasons, it may be a more general phenomenon, because concatenated sequences from all mitochondrial protein-coding genes have produced statistically significant but incorrect topologies when moderately distant taxa are included (Nei 1996; Naylor and Brown 1997). Although the reason why a topology is incorrect (e.g., frog and bird clustering with fish rather than with other tetrapods) may be complex, statistical significance of that topology probably is the result of having a large number of sites, which increases the power of the test.

More DNA sequence data are needed to understand the dichotomy in phylogenetic signal between the coding and noncoding genes. However, agreement among these diverse molecular data sets (noncoding genes, DNA hybridization data, and transferrin data) warrants a discussion of the biogeographic implications of a Rhea-Australasia connection.

Biogeographic History of Ratites

The supercontinent Gondwana began to break up about 150 MYA (Smith, Smith, and Funnell 1994) and a correspondence between these geologic events and ratite phylogeny has been proposed (Cracraft 1974). However, the subsequent discovery of Laurasian fossils (Houde 1986, 1988) was used as evidence against that hypothesis (Feduccia 1996). Also, interordinal divergence times estimated from nuclear and mitochondrial genes (Hedges et al. 1996; Härild, Janke, and Arnason 1997) now suggest that divergences within ratites probably were not earlier than about 90 MYA. If the prevailing phylogenetic pattern among the molecular data sets is correct, it suggests a new biogeographic hypothesis for the early evolution of ratites compatible with these constraints. Although it includes some dispersal, a Rhea-Australasia connection agrees more with earth history than does a Struthio-Australasia or Rhea-Struthio relationship.

We propose two possible origins for the ratites: an African origin or a South American origin. Both scenarios are in agreement with the presence of early Cenozoic ratite fossils in Laurasia, and both suggest a South American origin for the Australasian ratites. They differ in the location of the earliest ratite.

Under the African-origin hypothesis, the palaeognath lineage existed on the Africa–South America land mass, and the divergence of proto-ratite and proto-tinamou lineages was caused by vicariance after the separation of Africa and South America approximately 100 MYA (Smith, Smith, and Funnell 1994). Sometime between then and the late Cretaceous, the proto-ratite reached Laurasia by dispersal, leaving behind the Struthio lineage. After reaching North America, a lineage would have crossed a proto-Antilles land connection to South America in the late Cretaceous, establishing the Rhea lineage, and would shortly thereafter have dispersed to Australasia (fig. 3A). Faunal exchanges between North America and South America from the late Cretaceous to the early Tertiary probably occurred in both directions (Hallam 1994). The earliest palaeognaths in Paleocene deposits in Laurasia (Houde 1986, 1988; Martin 1992) and Gondwana (Alvarenga 1983) suggest that the proto-Antillean land connection could have been used by ratites.

Alternatively, the palaeognath lineage may have originated in South America after separation of that continent from Africa. Following an early tinamou divergence, a Struthio lineage may have arisen by dispersal northward, across the proto-Antilles, to North America in the late Cretaceous, and subsequently to Laurasia and
A large gap in the ratite fossil record exists in Africa, Madagascar, and New Zealand. The oldest irrefutable *Struthio* has been found in Africa and dated to the Early Miocene (Mourer-Chauviré et al. 1996). Morphological (Bledsoe 1988; Kurochkin 1995) and eggshell evidence (Mourer-Chauviré et al. 1996) suggest a close relationship between elephantbirds (*Aepyornis*) and ostriches. The early Cretaceous separation of Madagascar and Africa (Smith, Smith, and Funnell 1994; Hallam 1994) was too early for vicariance to explain the origin of elephantbirds; dispersal from Africa to Madagascar is more likely.

Based on 400 bp of the 12S rRNA gene, it was proposed that New Zealand was colonized twice by ancestors of ratite birds (Cooper et al. 1992; Cooper 1997). However, because longer sequences of the same gene and adjacent genes yield a significantly different topology for the living ratites (fig.1), those results regarding the position of the moas are placed in question. A relationship between moas and kiwis based on earlier morphological studies (Mivart 1877; McDowell 1948; Cracraft 1974) must again be reconsidered.

It is believed that New Zealand separated from Antarctica in the late Cretaceous (Hallam 1994), which raises the possibility of a vicariant origin for moas and kiwi. This would be compatible with our biogeographic hypothesis, although a later dispersal from Australia to New Zealand also may have occurred. Most discussions of dispersal in flightless birds involve the loss of the ability to fly after dispersing, occasionally with reference to swimming (Cooper et al. 1992). However, most other nonflying terrestrial vertebrates disperse over long distances of open ocean by rafting on flotsam (Hedges 1996), and there is no reason to exclude this mechanism with regard to flightless birds.

**Acknowledgment**

We thank J. L. Cracraft for providing comments on the manuscript.

**LITERATURE CITED**


Naruya Saitou, reviewing editor

Accepted December 3, 1997