THE ROYAL SOCIETY

doi 10.1098/rspb.2001.1587

Complete mitochondrial DNA genome sequences of extinct birds: ratite phylogenetics and the vicariance biogeography hypothesis

Oliver Haddrath[1]* **and Allan J. Baker**[2]

[1]*Centre for Biodiversity and Conservation Biology, Royal Ontario Museum, Toronto, Ontario, Canada M5S 2C6*
[2]*Department of Zoology, University of Toronto, Toronto, Ontario, Canada M5S 1A1*

The ratites have stimulated much debate as to how such large flightless birds came to be distributed across the southern continents, and whether they are a monophyletic group or are composed of unrelated lineages that independently lost the power of flight. Hypotheses regarding the relationships among taxa differ for morphological and molecular data sets, thus hindering attempts to test whether plate tectonic events can explain ratite biogeography. Here, we present the complete mitochondrial DNA genomes of two extinct moas from New Zealand, along with those of five extant ratites (the lesser rhea, the ostrich, the great spotted kiwi, the emu and the southern cassowary) and two tinamous from different genera. The non-stationary base composition in these sequences violates the assumptions of most tree-building methods. When this bias is corrected using neighbour-joining with log-determinant distances and non-homogeneous maximum likelihood, the ratites are found to be monophyletic, with moas basal, as in morphological trees. The avian sequences also violate a molecular clock, so we applied a non-parametric rate smoothing algorithm, which minimizes ancestor–descendant local rate changes, to date nodes in the tree. Using this method, most of the major ratite lineages fit the vicariance biogeography hypothesis, the exceptions being the ostrich and the kiwi, which require dispersal to explain their present distribution.

Keywords: phylogeny; mitochondrial DNA; palaeognaths; ratites; moas

1. INTRODUCTION

In the last three decades, studies based on morphology, immunological distances, chromosome banding and protein and DNA sequencing have indicated that the ratites are monophyletic (Cracraft 1974; Prager *et al*. 1976; De Boer 1980; Sibley & Ahlquist 1981, 1990; Stapel *et al*. 1984; Bledsoe 1988; Cooper *et al*. 1992; Casper *et al*. 1994; Cooper 1997; Lee *et al*. 1997; Van Tuinen *et al*. 1998) and that the tinamous of South America are their sister group. However, the branching order of the major ratite lineages in some of these trees has cast doubt on the role of plate tectonic events as an exclusive explanation for their current distributions because the divergences of some taxa seem to be too recent to fit with the break-up of Gondwanaland (Cooper *et al*. 1992; Cooper 1997; Van Tuinen *et al*. 1998; Härlid *et al*. 1998). Dispersal of these taxa between disjunct landmasses implies not only that ratites are descended from flying ancestors (Houde 1986) but also that loss of flight occurred several times. Ultimately, the resolution of these key issues in avian evolution depends upon the construction of a well-supported phylogeny for ratites and related taxa, and on the reliable dating of divergences.

The application of the polymerase chain reaction (PCR) has made it possible to retrieve DNA sequences from both extant and extinct ratite taxa and, thus, to construct comprehensive molecular phylogenies that can, in principle, be dated using molecular clocks. However, the promise of DNA sequences in resolving the phylogenetic relationships between the ratites has not yet been fully realized because different tree construction methods and different data partitions of mitochondrial DNA (mtDNA) sequences have produced a variety of tree topologies. For example, a molecular tree constructed from 5 kilobases (kb) of mitochondrial protein-coding and rRNA sequences placed the rheas basal among the ratites (Lee *et al*. 1997), whereas another tree based on just the rRNA partition found strong support for the ostrich as the basal taxon (Van Tuinen *et al*. 1998).

Further uncertainty about the use of DNA-based trees has come from the incongruence between these trees and a classical morphological phylogeny (Cracraft 1974; Lee *et al*. 1997) in which the kiwis and moas of New Zealand are basal sister groups to all other ratites. Although these differences have been attributed to the incorrect rooting of the molecular tree, another hypothesis of relationships based on morphological characters placed the kiwis in a more derived Australasian clade, as in the molecular trees, and separated the moas as the basal ratite lineage (Bledsoe 1988). Although this phylogeny has been criticized for the choice and polarization of characters it employs (Lee *et al*. 1997), it mirrors the major clades found in molecular phylogenies, and invites further testing with complete mtDNA sequences and denser taxon sampling.

The recently extinct moas of New Zealand are pivotal in distinguishing between competing hypotheses of phylogenetic relationships constructed from molecular and morphological characters. The only published study of ratites that included DNA sequences from the extinct moas of New Zealand employed 390 base pairs (bp) of mtDNA from the 12S rRNA gene (Cooper *et al*. 1992). This tree placed the rheas basal among ratites, with the moas branching off next. However, the stochastic error is likely to be appreciable in such short sequences and much

*Author for correspondence (oliverh@rom.on.ca).

Received 4 December 2000 *Accepted* 16 January 2001

larger amounts of sequence data are required to obtain robust estimates of relationships between major ratite lineages that may have diverged prior to or during the fragmentation of Gondwanaland (Cracraft 1973, 1974; Van Tuinen *et al*. 1998).

To help to resolve the phylogenetic relationships among the ratites and to date their divergences, we sequenced the complete mtDNA genomes of seven ratites and two tinamous, and analysed them along with other published genomes. Complete mtDNA genomes have been used successfully to infer phylogenetic relationships at the ordinal and class levels (Árnason & Johnsson 1992; Cao *et al*. 1994; Russo *et al*. 1996; Janke *et al*. 1997; Nikaido *et al*. 2000) and have been shown to recover model trees more accurately than shorter sequences (Cummings *et al*. 1995; Nei *et al*. 1998). Genomic sequences also provide an opportunity to examine the efficacy of different data partitions such as protein-coding rRNA and tRNA genes in recovering well-supported trees. The inclusion of moas, along with increased taxon sampling of ratites, allows a more thorough examination of whether the apparent conflicts in the previously proposed trees arise simply from incorrect rooting or from inadequate phylogenetic signals in previous sequence data sets.

2. METHODS

Complete mtDNA genomes were obtained for the ostrich (*Struthio camelus*), emu (*Dromaius novaehollandiae*), southern cassowary (*Casuarius casuarius*), lesser rhea (*Pterocnemia pennata*), great spotted kiwi (*Apteryx haasti*), two extinct species of moas (*Anomalopteryx didiformis* and *Emeus crassus*), the elegant crested tinamou (*Eudromia elegans*) and the great tinamou (*Tinamus major*). DNA of extant species was isolated from solid tissues, except in the case of the great spotted kiwi (blood) and southern cassowary (feather), using standard methods (Sambrook *et al*. 1989). DNA of the moas was extracted and amplified from bones according to the stringent protocol of Hagelberg (1994). All moa-DNA extractions were performed in a separate room reserved for this purpose.

Segments of the extant-ratite genomes were amplified in greater than 1 kb fragments and were sequenced manually in both directions with a greater than 99% overlap with nested primers. Amplifications of the moa genes were performed separately to minimize the risk of contamination and targeted smaller segments, typically 350–500 bp in length. The amplified segments were designed to overlap by 200–300 bp to confirm sequences from separate amplifications and, thus, identify any fragments that could potentially have non-mitochondrial origins. A total of 180 primers (available on request from the authors) were used to amplify and sequence the genomes. Sequencing reactions were performed with ThermoSequenase kits (Amersham Pharmacia, Quebec, Canada).

To further guard against nuclear copies of mitochondrial sequences (numts) we scrutinized autoradiographs for evidence of multiple bands indicating more than one sequence; translated the sequences of protein-coding genes to confirm that the reading frames were free of indels that were not a multiple of three nucleotides; examined the pattern of substitution at codon positions to ensure that it followed the expected pattern of decrease from third to first to second positions; mapped the rRNA and tRNA genes onto secondary structure models; and checked for any substitutions that would be incompatible with the predicted structures. The degradation of DNA in sub-fossil moa bones means that the recovery of low-copy numts, as shown by the inability to recover single-copy nuclear genes (see Cooper 1997) is far less likely than the recovery of high-copy mtDNA genes in these extinct birds.

Amplification and sequencing of half the *E. crassus* genome and one-third of the *A. didiformis* genome were repeated for independent confirmation at the Department of Ornithology of the American Museum of Natural History. The sequences were read on an ABI 377 automated sequencer (Applied Biosystems, Foster City, CA, USA). The complete mitochondrial genomes have been deposited at GenBank with the following accession numbers: ostrich (*S. camelus*) AF338715; emu (*D. novaehollandiae*) AF338711; southern cassowary (*C. casuarius*) AF338713; lesser rhea (*P. pennata*) AF338709; great spotted kiwi (*A. haasti*) AF338708; two extinct species of moas, *A. didiformis* AF338714 and *E. crassus* AF338712; the elegant crested tinamou (*E. elegans*) AF338710; and the great tinamou (*T. major*) AF338707.

The protein-coding genes were aligned at the amino-acid level using the program Clustal X (Thompson *et al*. 1997). The rRNA and tRNA genes were aligned using models of secondary structure (Kumazawa & Nishida 1993; De Rijk *et al*. 2000). Regions of gene overlap, indels and nicotinamide adenine dinucleotide (NADH) dehydrogenase subunit 6 (ND6), which is encoded on the opposite strand to the other protein-coding genes, were excluded from all analyses. This resulted in a concatenated data set of 13 535 bp comprised 10 689 bp of protein-coding genes, 2095 bp of rRNA genes and 751 bp of tRNA genes. Phylogenetic analyses were performed on the sequences of the nine taxa sequenced here along with those from the greater rhea (*Rhea americana*) (Härlid *et al*. 1998), the chicken (*Gallus gallus*) (Desjardins & Morais 1990), the redhead duck (*Aythya americana*) (Mindell *et al*. 1999), the alligator (*Alligator mississippiensis*) (Janke & Árnason 1997) and the African side-necked turtle (*Pelomedusa subrufa*) (Zardoya & Meyer 1999). The concatenated data sets, as well as protein-coding and rRNA genes separately, were examined. Data sets containing protein-coding genes were analysed with and without third codon positions. Tests of stationarity across sequences were performed using the program PUZZLE v. 4.0 (Strimmer & Von Haeseler 1996) on all sites, on first and second codon positions and on the variable sites only. Symmetrical directional mutation pressure was estimated for the protein-coding genes using the program DMP v. 2.0 (Jermiin *et al*. 1996), which examines the G + C content at synonymous and non-synonymous sites.

Phylogenetic analyses were carried out using neighbour-joining with log-determinant distances (Lockhart *et al*. 1994) and with the removal of a proportion of the invariant sites as implemented in the program PAUP* v. 4.0b4 (Swofford 1998). For analyses using maximum likelihood, likelihood scores for all of the proposed ratite phylogenies (Cracraft 1974; Prager *et al*. 1976; Sibley & Ahlquist 1981, 1990; Bledsoe 1988; Cooper *et al*. 1992; Van Tuinen *et al*. 1998) were examined, along with all the other possible trees within a constrained ratite tree space: [(tinamous), ((moas), (rheas), ostrich, kiwi, (emu, cassowary))]. Maximum-likelihood analyses were performed with non-homogenous models of substitution and a correction for rate variation between sites using the programs NHML (Galtier *et al*. 2000) and PAML v. 3.0c (Yang & Roberts 1995; Yang 2000).

Rate constancy across taxa was violated for the bird sequences, so we employed a non-parametric rate smoothing process that relies on evolutionary rates being autocorrelated in time and minimizes ancestor–descendant local rate changes,

Table 1. *Symmetrical directional mutation pressure (μ_D) for the 12 H-strand protein-coding genes*

(Estimates were made using the program DMP (Jermiin *et al*. 1996), which examines the G + C content at synonymous (P_{syn}) and non-synonymous (P_{non}) sites. μ_D less than 0.5 has greater A + T content and μ_D greater than 0.5 has greater G + C content.)

taxon	P_{obs}	μ_D	P_{non}	P_{syn}
turtle	0.394	0.385***	0.393	0.396
alligator	0.435	0.467***	0.406	0.48
duck	0.495	0.585***	0.43	0.597
chicken	0.47	0.536**	0.418	0.549
tinamou	0.442	0.477***	0.41	0.491
moas	0.46	0.509	0.419	0.524
rheas	0.486	0.567***	0.425	0.58
ostrich	0.455	0.493*	0.42	0.507
kiwi	0.434	0.454***	0.413	0.468
emu	0.435	0.447***	0.419	0.46

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

using the program r8s (Sanderson 1997). The topology and branch lengths of the maximum-likelihood tree were specified as the user tree. The nodes were calibrated internally using the fossil *Emuarius gidju* from Australia, which dates back to the late Oligocene, 25 million years (Myr) ago. This species, while appearing close to the common ancestor of the emu and the cassowaries, was clearly on the lineage leading to the emu and was included in the subfamily Dromaiinae (Boles 1992). As this fossil post-dates the divergence between the emu and the cassowaries, we added an estimate of 10 Myr, so that their common ancestor was dated at 35 Myr.

3. RESULTS

(a) *Genome characteristics*

The mtDNA genomes of the palaeognathous birds are similar in size (*ca*. 16 700 bp) to other avian mtDNA genomes sequenced to date (Desjardins & Morais 1990; Härlid *et al*. 1997, 1998; Mindell *et al*. 1999). We compared the ostrich genome sequence we obtained to other published ostrich mtDNA sequences and found that our sequence was virtually identical to the 10 400 bp sequence of Mindell *et al*. (1999) (99.96%) and the complete sequence of Härlid *et al*. (1997) (99.69%). The lesser-rhea genome had 93.3% sequence identity with the greater-rhea sequence in Härlid *et al*. (1998). The 12S rRNA sequences we obtained for the two moas were identical to those published over the 390 bp obtained for the same taxa by Cooper *et al*. (1992). All of the ratites and the great tinamou had the same gene order as the chicken (Desjardins & Morais 1990). The elegant crested tinamous, however, had a different gene order, with ND6 having moved from in front of the control region to behind it. Similar rearrangements have been described in other avian species (Mindell *et al*. 1998*a*; Bensch & Härlid 2000) representing a wide sampling of groups. As these species do not share a direct common ancestry, the rearrangement has been assumed to have arisen independently on multiple occasions. The presence of an 'extra' nucleotide 174 bases in from the start of gene ND3 was found in all nine of the genomes, as in many other avian genomes (Härlid *et al*. 1997, 1998; Mindell *et al*. 1998*b*). This extra base would be expected to produce a truncated protein product for this gene but is presumed not to be translated (Mindell *et al*. 1998*b*).

Table 2. *Estimated dates of avian divergences along with, in parentheses, bootstrap estimates of their standard errors. These dates are compared, for the ratites, with the Gondwanan fragmentation dates of the land masses on which they are currently distributed*

divergence (node)	divergence date (Myr)	fragmentation date (Myr)	fragment separation
palaeognaths–Galloanserae (1)	110.1 (±19.8)	—	—
chicken–duck (2)	85.4 (±16.9)	—	—
tinamous (3)	92.2 (±18.9)	—	—
between tinamous (4)	62.8 (±14.1)	—	—
moas (5)	78.9 (±14.8)	82	New Zealand
between moas (6)	5.3 (±1.6)	—	—
kiwis (10)	62.4 (±7.0)	—	—
rheas (7)	69.3 (±12.4)	35	South America
between rheas (8)	13.7 (±3.2)	—	—
ostrich (9)	65.3 (±6.9)	100	Africa
emu (11)	35 (30–35)	55	Australia

Examination of the mtDNA genomes sequenced here revealed significant variation ($p < 0.05$) in nucleotide composition, both among ratites and between most of the ratites and the Galloanserae outgroup (chicken and redhead duck). This was most evident in the protein-coding genes, where the rheas, chicken and redhead duck were shown to be experiencing symmetrical directional mutation pressure driving them towards greater G + C content, while the other extant ratites and the tinamous were being driven toward greater A + T content (table 1).

(b) *Phylogenetic analysis*

Most methods of phylogenetic reconstruction assume a stationary Markov process of nucleotide substitution, with the nucleotide frequencies remaining constant across sequences. Non-stationarity can bias tree construction methods and result in erroneous tree topologies (Lockhart *et al*. 1994; Foster & Hickey 1999; Galtier *et al*. 2000). Maximum parsimony and homogeneous maximum likelihood either placed the tinamous within the ratite clade or, alternatively, had the G + C-rich rheas at the base of the ratite tree. Each outcome could be the result of the rheas being attracted towards the neognath outgroups that also have a high G + C content in their sequences. To correct for this compositional bias, we analysed the concatenated protein-coding rRNA and tRNA genes using the neighbour-joining method on log-determinant distances and non-homogeneous models of evolution in maximum likelihood, which do not assume base-compositional stationarity. The resulting tree topology for ratites (figure 1) is well supported, with relatively high bootstrap values at all nodes except for that defining whether the ostrich or the rheas are basal among extant ratites. Inclusion or exclusion of third

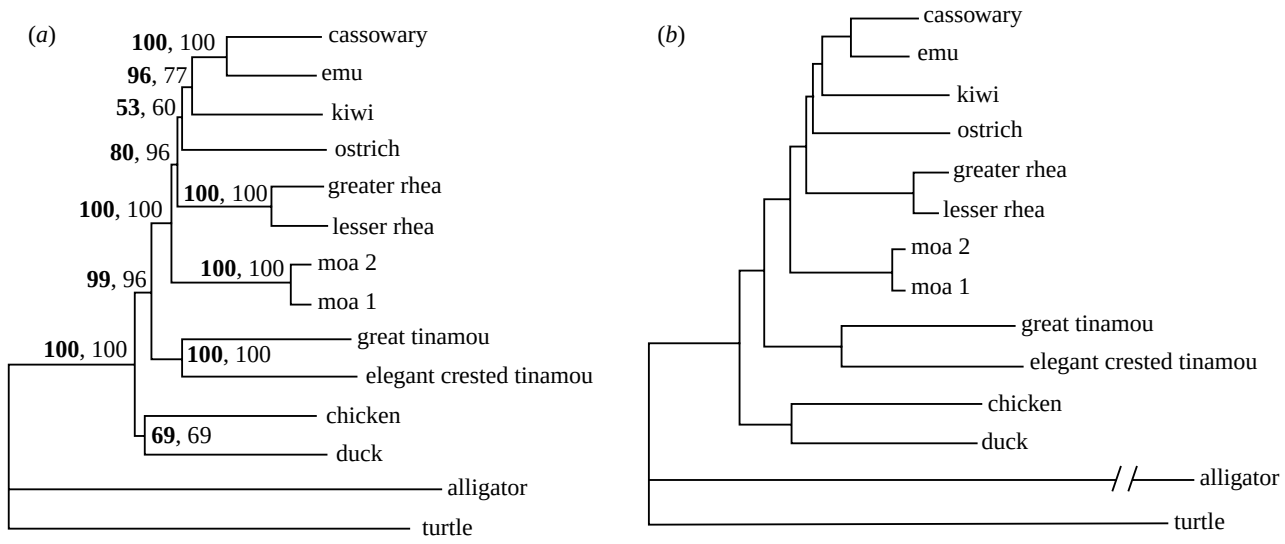

Figure 1. (*a*) Neighbour-joining (log-determinant distances) and (*b*) maximum-likelihood (non-homogeneous model) tree topologies from the combined mtDNA protein-coding rRNA and tRNA genes. The bootstrap values shown are from the neighbour-joining analyses with (first number, in bold type) and without (second number) third codon positions. Moa 1 is *Anomalopteryx didiformis* and moa 2 is *Emeus crassus*.

codon positions did not alter this tree topology. The ratite topology was also unaffected by increased taxon sampling and choice of outgroup (non-avian, tinamous or Galloanserae). The moas form the basal lineage of the included ratites, whereas the great spotted kiwi is placed in a well-supported and derived clade with the emu and the cassowary. Tinamous are sister to the ratites, which in turn are sister to the Galloanserae.

(c) *Molecular dating of divergences*

A noticeable feature of the non-homogeneous maximum-likelihood tree (figure 1) is that the branch lengths of the different lineages vary considerably. In particular, the tinamous and the Galloanserae have very long branch lengths while the moas, the emu and the cassowary have much shorter branch lengths. Thus, striking rate variation between lineages is apparent both among birds and within ratites. Rate constancy between lineages was rejected statistically; likelihood-ratio tests with and without a molecular clock were significant among all birds and among ratites ($p < 0.001$). A method of dating is required that incorporates rate changes between taxa, so we used a non-parametric rate smoothing approach in the program r8s (Sanderson 1997). The divergence dates for the avian taxa included in this study, along with estimates of the standard errors for these dates, are given in table 2 (see figure 2). Within the major lineages of ratites, the two genera of moas had a recent common ancestry at 5.3 ± 1.6 Myr, while the two rhea genera diverged 13.7 ± 3.2 Myr ago. Divergences between the remaining lineages are correspondingly older; the great spotted kiwi diverged from the emu–cassowary lineage 62.4 ± 7.0 Myr ago, while the ostrich, the rheas and the moas branched off 65.3 ± 6.9 Myr, 69.3 ± 12.4 Myr and 78.9 ± 14.8 Myr ago, respectively. Palaeognaths first diverged about 92.2 ± 18.9 Myr ago, while the last Galloanserae common ancestor is dated at 85.4 ± 16.9 Myr. The common ancestor of these groups was calculated to have lived 110.1 ± 19.8 Myr ago.

4. DISCUSSION

(a) *Phylogenetic relationships between ratites*

Our large data set of 13 535 bp confirms the placement of the kiwi as sister to the emu and the cassowary in a derived Australasian clade, as has been shown in other molecular studies (Sibley & Ahlquist 1990; Cooper *et al*. 1992; Cooper 1997; Lee *et al*. 1997) and a single morphological study (Bledsoe 1988). However, this topology conflicts with two morphological studies (Cracraft 1974; Lee *et al*. 1997) that placed kiwis at the base of the ratite tree with the moas. Lee *et al*. (1997) attempted to reconcile these disparate topologies by arguing that the discrepancy was a result of improper rooting of the ratite molecular tree. This argument could only be posited because their molecular phylogeny lacked moa sequences. The inclusion of moas in our study yielded a tree that is congruent with Lee *et al*.'s (1997) molecular tree but cannot be rooted to be consistent with their morphological tree. Thus incorrect rooting is not the cause of this incongruence.

Another uncertainty in ratite systematics has been the placement of moas relative to the other ratites. Most phylogenetic studies that have included moas have relied exclusively on morphological characters due to the difficulties in extracting organic macromolecules from sub-fossil remains. These studies placed the moas at the base of the ratite tree (Cracraft 1974; Bledsoe 1988; Lee *et al*. 1997). The single molecular phylogeny that included moas (Cooper *et al*. 1992) was constructed from short (390 bp) 12S rRNA sequences, and found that rheas were basal among ratites with the moas branching off next. Our analyses, in contrast, were based on the considerably larger data set from the mtDNA genome, and placed the moas as the basal lineage among ratites, in agreement with morphological studies. This conclusion was also supported by separate analyses of both the protein-coding and the rRNA genes, and is thus represented robustly in different partitions of the mtDNA genomic sequences.

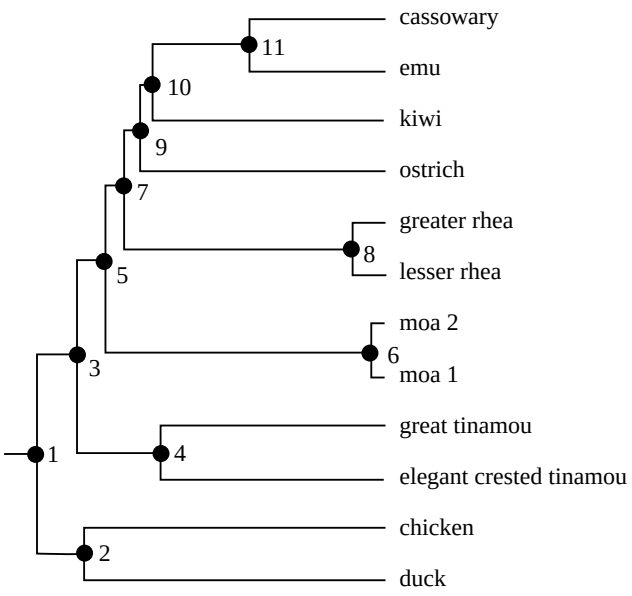

Figure 2. Tree topology with branch lengths scaled proportional to divergence times calculated with the program r8s. Nodes are numbered according to table 2.

The phylogenetic positions of the ostrich and the rheas have also been problematic. Three topologies have been proposed for their placement: the ostrich branched off first (Sibley & Ahlquist 1990; Van Tuinen *et al*. 1998); the rheas branched off first (Cooper *et al*. 1992; Cooper 1997; Lee *et al*. 1997); or the ostrich and the rheas form a clade near the base of the tree (Cracraft 1974; Sibley & Ahlquist 1990; Bledsoe 1988). The tree topologies produced in our study, using the total concatenated data set, showed the rheas branching off first among extant ratites but this was not strongly supported, as shown by the low bootstrap values (figure 1). Phylogenetic analyses using just the protein-coding or just the rRNA genes failed to resolve the branching order of the ostrich and the rheas but the tree with the highest bootstrap support (72%) for the rheas being basal was based on the rRNA genes alone. This finding contradicts that of Van Tuinen *et al*. (1998), who argued that the mitochondrial rRNA genes placed the ostrich basal while the protein-coding genes placed the rheas basal. The probable explanation for the differences between the rRNA trees in Van Tuinen *et al*. (1998) and our study is the difference in taxon sampling. The extremely short inter-nodal branch length between the ostrich and the rheas helps to explain why their relative positions in the ratite tree have been difficult to resolve.

(b) *The vicariance biogeography hypothesis*

The vicariance biogeography hypothesis proposes that the major lineages of ratites are descended from a flightless ancestor that was widespread in Gondwanaland. As the landmass fragmented, ancestral ratites rafted passively with their respective continental fragments to their current positions (Cracraft 1973, 1974). This hypothesis has been questioned by recent phylogenetic analyses that have placed kiwis in a far more derived position in the ratite tree than expected under vicariance (Sibley & Ahlquist 1981, 1990; Bledsoe 1988; Cooper *et al*. 1992; Cooper 1997; Lee *et al*. 1997). This suggested that the divergence of kiwis may have occurred too recently for them to have been present when New Zealand separated from the rest of Gondwanaland 82 Myr ago (Cooper & Millener 1993; L. Lawver, personal communication). Thus, their presence in New Zealand could have resulted from a secondary invasion, by swimming, island hopping or flight, after the landmass had rafted away (Cooper *et al*. 1992). The discovery of volant birds with archaic palates in North America and Europe dating from *ca*. 50–60 Myr ago (Houde 1986) suggests that flightlessness may have arisen independently among the ratite lineages and that kiwis may have flown to New Zealand.

To estimate divergence dates across taxa, a calibration point, usually from the fossil record, is required. Härlid *et al*. (1998) estimated the divergence date between the ostrich and the greater rhea at 51 ± 5 Myr ago using an avian–crocodilian split of 254 Myr. Birds are known to have slower rates of molecular evolution than other tetrapods (Prager *et al*. 1974; Britten 1986; Mindell *et al*. 1996; Stanley & Harrison 1999). Härlid *et al*.'s (1998) clock calculation assumed that any slowdown in the avian rate of evolution occurred at the point of divergence between birds and crocodilians. If the slowdown occurred gradually, or further along the lineage leading to birds, then the ostrich–rhea divergence date would be underestimated. A better calibration date would be within the ratites, such as the date of the emu–cassowary common ancestor. The dates derived using non-parametric rate smoothing are more appropriate for our sequences than assuming a molecular clock because there is clear variation in the rates of evolution between taxa.

Our estimate of 62.4 ± 7.0 Myr for the divergence date of the kiwi lineage from the Australian ratites suggests that kiwis were not present when New Zealand separated from the rest of Gondwanaland, and therefore that they probably invaded secondarily. A land bridge via the Lord Howe Rise between New Zealand and Australia may have permitted faunal interchange until as recently as 75 Myr ago (Cooper & Millener 1993; L. Lawver, personal communication). This date is very close to the limit of our calculated divergence date for the kiwis, suggesting that a secondary invasion of New Zealand may not have required the kiwis to be volant. The divergence date of the moas, at 78.9 ± 14.8 Myr ago, is consistent with the separation of New Zealand from Australia. The common ancestor of the emu and the cassowaries diverged from the rest of the ratites before Australia separated from Antarctica 55 Myr ago (Ocean Drilling Program 2000). However, the divergence of the ostrich at 65.3 ± 6.9 Myr ago cannot be reconciled with the separation of Africa from South America about 100 Myr ago (Smith *et al*. 1994). Van Tuinen *et al*. (1998) recognized this and proposed that some dispersal between South America and Africa through the Northern Hemisphere must have occurred. This hypothesis is supported by the presence of an ostrich-like ratite in middle-Eocene Europe (Houde 1987). An alternative scenario, also involving vicariance and dispersal, is possible via the proposed land bridge linking Antarctica, India and Madagascar (Hay *et al*. 1999). The existence of this land link is supported by the pan-Gondwanan distribution of several Late-Cretaceous terrestrial faunas (Krause *et al*. 1997; Sampson *et al*. 1998). The common ancestor of the extinct

Malagasy elephant birds (Aepyornithidae) and the ostrich may have become isolated on Madagascar and India when the land bridge to Antarctica was submerged. This occurred close to the time calculated for the divergence of the ostrich and, thus, a vicariant origin for this lineage cannot be excluded. This ancestral stock could have invaded Africa from Madagascar and then spread to Europe, Arabia, India and Asia. The sequencing of an elephant-bird mitochondrial genome should provide additional insights into which of these scenarios is most likely, by establishing its phylogenetic placement among the other ratites and its divergence date from the ostrich. Ancient DNA techniques could provide the critical evidence in resolving these final parts of an evolutionary puzzle that has mystified biologists for over a century.

We thank Trevor Worthy and Brian Gill for providing moa bone samples, Leo Joseph for the great tinamou tissue, John and Rita Squire for the cassowary feather and Wieslawa and Owen Haw for the emu tissue. We are grateful to Chris McGowan, David Irwin, Hans Dieter Sues and Úlfar Árnason for comments on the manuscript, and Parneet Arora for assistance in computer programming. Funding was provided by grants to A.J.B. from the Natural Sciences and Engineering Research Council of Canada and the Royal Ontario Foundation. We thank George Barrowclough and Joel Cracraft for funding from the Frank M. Chapman Memorial Fund when A.J.B. was a Chapman Fellow in the Ornithology Department, and for the use of facilities in the Lewis B. and Dorothy Cullman Molecular Laboratory.

REFERENCES

Árnason, Ú. & Johnsson, E. 1992 The complete mitochondrial DNA sequence of the harbour seal, *Phoca vitulina*. *J. Mol. Evol.* **34**, 493–505.

Bensch, S. & Härlid, A. 2000 Mitochondrial genomic rearrangements in songbirds. *Mol. Biol. Evol.* **17**, 107–113.

Bledsoe, A. H. 1988 A phylogenetic analysis of postcranial skeletal characters of the ratite birds. *Annls Carnegie Mus.* **57**, 73–90.

Boles, W. E. 1992 Revision of Dromaius gidju Patterson and Rich 1987 from Riversleigh, northwestern Queensland, Australia, with a reassessment of its generic position. *Nat. Hist. Mus. LA County Sci. Ser.* **36**, 195–208.

Britten, R. J. 1986 Rates of DNA sequence evolution differ between taxonomic groups. *Science* **231**, 1393–1398.

Cao, Y., Adachi, J., Janke, A., Pääbo, S. & Hasegawa, M. 1994 Phylogenetic relationships among eutherian orders estimated from inferred sequences of mitochondrial proteins: instability of a tree based on a single gene. *J. Mol. Evol.* **39**, 519–527.

Casper, G., Wattel, J. & de Jong, W. W. 1994 Alpha A-crystallin sequences group tinamou with ratites. *Mol. Biol. Evol.* **11**, 711–713.

Cooper, A. 1997 Studies of avian ancient DNA: from Jurassic park to modern island extinctions. In *Avian molecular evolution and systematics* (ed. D. P. Mindell), pp. 345–373. San Diego, CA: Academic Press.

Cooper, A., Mourer-Chauvire, C., Chambers, G. K., Von Haeseler, A., Wilson, A. & Pääbo, S. 1992 Independent origins of New Zealand moas and kiwis. *Proc. Natl Acad. Sci. USA* **89**, 8741–8744.

Cooper, R. A. & Millener, P. 1993 The New Zealand biota: historical background and new research. *Trends Ecol. Evol.* **8**, 429–433.

Cracraft, J. 1973 Continental drift, paleoclimatology and the evolution and biogeography of birds. *J. Zool.* **169**, 455–545.

Cracraft, J. 1974 Phylogeny and evolution of the ratite birds. *Ibis* **116**, 494–521.

Cummings, M. P., Otto, S. P. & Wakeley, J. 1995 Sampling properties of DNA sequence data in phylogenetic analysis. *Mol. Biol. Evol.* **12**, 814–822.

De Boer, L. E. M. 1980 Do the chromosomes of the kiwi provide evidence for a monophyletic origin of the ratites? *Nature* **287**, 84–85.

De Rijk, P., Wuyts, J., Van de Peer, Y., Winkelmans, T. & De Watchter, R. 2000 The European large subunit ribosomal RNA database. *Nucleic Acids Res.* **28**, 177–178.

Desjardins, P. & Morais, R. 1990 Sequence and gene organization of the chicken mitochondrial genome: a novel gene order in higher vertebrates. *J. Mol. Biol.* **212**, 599–634.

Foster, P. G. & Hickey, D. A. 1999 Compositional bias may affect both DNA-based and protein-based phylogenetic reconstructions. *J. Mol. Evol.* **48**, 284–290.

Galtier, N., Tourasse, N. & Guoy, M. 2000 A nonhyperthermophilic common ancestor to extant life forms. *Science* **283**, 220–221.

Hagelberg, E. 1994 Mitochondrial DNA from ancient bones. In *Ancient DNA* (ed. B. Herrmann & S. Hummel) pp. 195–204. New York: Springer.

Härlid, A., Janke, A. & Árnason, Ú. 1997 The mtDNA sequence of the ostrich and the divergence between paleognathous and neognathous birds. *Mol. Biol. Evol.* **14**, 754–761.

Härlid, A., Janke, A. & Árnason, Ú. 1998 The complete mitochondrial genome of *Rhea americana* and early avian divergences. *J. Mol. Evol.* **46**, 669–679.

Hay, W. W. (and 10 others) 1999 Alternative global Cretaceous paleogeography. In *The evolution of Cretaceous ocean/climate systems*, special paper 332 (ed. E. Barrera & C. Johnson), pp. 1–47. Boulder, CO: Geological Society of America.

Houde, P. 1986 Ostrich ancestors found in the northern hemisphere suggest new hypothesis of ratite origins. *Nature* **324**, 563–565.

Houde, P. 1987 *Paleotis wiegelti* restudied: a small middle Eocene ostrich (Aves: Struthioformes). *Palaeovertebrata* **17**, 27–42.

Janke, A. & Árnason, Ú. 1997 The complete mitochondrial genome of *Alligator mississippiensis* and the separation between recent archosauria (birds and crocodiles). *Mol. Biol. Evol.* **14**, 1266–1272.

Janke, A., Xu, X. & Árnason, Ú. 1997 The complete mitochondrial genome of the wallaroo (*Macropus robustus*) and the phylogenetic relationship among Monotremata, Marsupiala and Eutheris. *Proc. Natl Acad. Sci. USA* **94**, 1276–1281.

Jermiin, L. S., Foster, P. G., Graur, D., Lowe, R. M. & Crozier, R. H. 1996 Unbiased estimation of symmetrical directional mutation pressure from protein-coding DNA. *J. Mol. Evol.* **42**, 476–480.

Krause, D. W., Prasad, G. V. R., Von Koenigswald, W., Sahni, A. & Grine, F. E. 1997 Cosmopolitanism among Gondwanan late Cretaceous mammals. *Nature* **390**, 504–507.

Kumazawa, Y. & Nishida, M. 1993 Sequence evolution of mitochondrial tRNA genes and deep-branch animal phylogenetics. *J. Mol. Evol.* **37**, 380–398.

Lee, K., Feinstein, J. & Cracraft, J. 1997 The phylogeny of ratite birds: resolving conflicts between molecular and morphological data sets. In *Avian molecular evolution and systematics.* (ed. D. P. Mindell), pp. 173–209. San Diego, CA: Academic Press.

Lockhart, P. J., Stell, M. A., Hendy, M. D. & Penny, D. 1994 Recovering evolutionary trees under a more realistic model of sequence evolution. *Mol. Biol. Evol.* **11**, 605–612.

Mindell, D. P., Knight, A., Baer, C. & Huddleston, J. 1996 Slow rates of molecular evolution in birds and the metabolic rate and body temperature hypothesis. *Mol. Biol. Evol.* **13**, 422–426.

Mindell, D. P., Sorenson, M. D. & Dimcheff, D. E. 1998*a* Multiple independent origins of mitochondrial gene order in birds. *Proc. Natl Acad. Sci. USA* **95**, 10 693–10 697.

Mindell, D. P., Sorenson, M. D. & Dimcheff, D. E. 1998*b* An extra nucleotide is not translated in mitochondrial ND3 of some birds and turtles. *Mol. Biol. Evol.* **15**, 1568–1571.

Mindell, D. P., Sorenson, M. D., Dimcheff, D. E., Hasegawa, M., Ast, J. C. & Yuri, T. 1999 Interordinal relationships of birds and other reptiles based on whole mitochondrial genomes. *Syst. Biol.* **48**, 138–152.

Nei, M., Kumar, S. & Takahashi, K. 1998 The optimization principle in phylogenetic analysis tends to give incorrect topologies when the number of nucleotides or amino acids used is small. *Proc. Natl Acad. Sci. USA* **95**, 12 390–12 397.

Nikaido, M., Harada, M., Cao, Y., Hasegawa, M. & Okada, N. 2000 Monophyletic origin of the order Chiroptera and its phylogenetic position among Mammalia, as inferred from the complete sequence of the mitochondrial DNA of a Japanese megabat, the Ryuku flying fox. *J. Mol. Evol.* **51**, 318–328.

Ocean Drilling Program 2000 Leg 189. Preliminary report. See http://www-odp.tamu.edu/publications.

Prager, E. M., Brush, A. H., Nolan, R. A., Nakanishi, M. & Wilson, A. C. 1974 Slow evolution of transferrin and albumin in birds according to micro-complement fixation analysis. *J. Mol. Evol.* **3**, 243–262.

Prager, E. M., Wilson, A. C., Osuga, D. T. & Feeney, R. E. 1976 Evolution of flightless land birds on southern continents: transferrin comparison shows monophyletic origin of ratites. *J. Mol. Evol.* **8**, 283–294.

Russo, C. A., Takezaki, N. & Nei, M. 1996 Efficiencies of different genes and different tree-building methods in recovering a known vertebrate phylogeny. *Mol. Biol. Evol.* **13**, 525–536.

Sambrook, J., Fritsch, E. F. & Maniatis, T. 1989 *Molecular cloning: a laboratory manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Sampson, S. D., Witmer, L. M., Forster, C. A., Krause, D. W., O'Connor, P. M., Dodson, P. & Ravoavy, F. 1998 Predatory dinosaur remains from Madagascar: implications for the cretaceous biogeography of Gondwana. *Science* **280**, 1048–1051.

Sanderson, M. J. 1997 A nonparametric approach to estimating divergence times in the absence of rate constancy. *Mol. Biol. Evol.* **14**, 1218–1231.

Sibley, C. G. & Ahlquist, J. E. 1981 The phylogeny and relationships of the ratite birds as indicated by DNA–DNA hybridization. In *Evolution today: Proceedings of the Second International Congress for Systematic Evolutionary Biology* (ed. G. G. E. Scudder & L. L. Reveal), pp. 301–335. Pittsburgh, PA: Hunt Institute Botanic Document.

Sibley, C. G. & Ahlquist, J. E. 1990 *Phylogeny and classification of birds*. New Haven, CT: Yale University Press.

Smith, A. G., Smith, D. G. & Funnell, B. M. 1994 *Atlas of Mesozoic and Cenozoic coastlines*. Cambridge University Press.

Stanley, S. E. & Harrison, R. G. 1999 Cytochrome *B* evolution in birds and mammals: an evaluation of the avian constraints hypothesis. *Mol. Biol. Evol.* **16**, 1575–1585.

Stapel, S. O., Leunissen, J. A. M., Versteeg, M., Wattel, J. & de Jong, W. W. 1984 Ratites as the oldest offshoot of avian stem—evidence from αcrystallin A sequences. *Nature* **311**, 257–259.

Strimmer, K. & Von Haeseler, A. 1996 Quartet puzzling: a quartet maximum likelihood method for reconstructing tree topologies. *Mol. Biol. Evol.* **13**, 964–969.

Swofford, D. L. 1998 *PAUP**: *phylogenetic analysis using parsimony* (**and other methods*), v. 4.0.b4. Sunderland, MA: Sinauer Associates.

Thompson, J. D., Gibson, T. J., Plewiak, F., Jeanmolgin, J. & Higgins, D. G. 1997 The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **24**, 4876–4882.

Van Tuinen, M., Sibley, C. G. & Hedges, S. B. 1998 Phylogeny and biogeography of ratite birds inferred from DNA sequences of the mitochondrial ribosomal genes. *Mol. Biol. Evol.* **15**, 370–376.

Yang, Z. 2000 Phylogenetic analysis by maximum likelihood (*PAML*), v. 3.0. University College London.

Yang, Z. & Roberts, D. 1995 On the use of nucleic acid sequences to infer early branchings in the tree of life. *Mol. Biol. Evol.* **12**, 451–458.

Zardoya, R. & Meyer, A. 1999 Complete mitochondrial genome suggests diapsid affinities of turtles. *Proc. Natl Acad. Sci. USA* **95**, 14 226–14 231.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.