

# Complete mitochondrial DNA genome sequences show that modern birds are not descended from transitional shorebirds

Tara Paton<sup>1,2\*</sup>, Oliver Haddrath<sup>1</sup> and Allan J. Baker<sup>1,2</sup>

<sup>1</sup>Centre for Biodiversity and Conservation Biology, Royal Ontario Museum, 100 Queen's Park, Toronto, ON, Canada M5S 2C6

<sup>2</sup>Department of Zoology, University of Toronto, Toronto, ON, Canada M5S 1A1

To test the hypothesis put forward by Feduccia of the origin of modern birds from transitional birds, we sequenced the first two complete mitochondrial genomes of shorebirds (ruddy turnstone and blackish oystercatcher) and compared their sequences with those of already published avian genomes. When corrected for rate heterogeneity across sites and non-homogeneous nucleotide compositions among lineages in maximum likelihood (ML), the optimal tree places palaeognath birds as sister to the neognaths including shorebirds. This optimal topology is a re-rooting of recently published ordinal-level avian trees derived from mitochondrial sequences. Using a penalized likelihood (PL) rate-smoothing process in conjunction with dates estimated from fossils, we show that the basal splits in the bird tree are much older than the Cretaceous–Tertiary (K–T) boundary, reinforcing previous molecular studies that rejected the derivation of modern birds from transitional shorebirds. Our mean estimate for the origin of modern birds at about 123 million years ago (Myr ago) is quite close to recent estimates using both nuclear and mitochondrial genes, and supports theories of continental break-up as a driving force in avian diversification. Not only did many modern orders of birds originate well before the K–T boundary, but the radiation of major clades occurred over an extended period of at least 40 Myr ago, thus also falsifying Feduccia's rapid radiation scenario following a K–T bottleneck.

**Keywords:** phylogeny; mitochondrial DNA genomes; Charadriiformes; molecular dating

## 1. INTRODUCTION

Based on a reinterpretation of the fossil record of birds, Feduccia (1996) hypothesized that one of the few avian survivors of the mass extinction at the Cretaceous–Tertiary (K–T) boundary 65 million years ago (Myr ago) were 'transitional shorebirds', and thus they were the ancestors of all modern birds. However, this radical new hypothesis has been challenged by analyses of DNA sequences which have used molecular dating methods to estimate the ages of the splits among extant bird orders (Hedges *et al.* 1996; Cooper & Penny 1997; Waddell *et al.* 1999; Haddrath & Baker 2001; Van Tuinen & Hedges 2001). These studies suggest that most orders originated well before the K–T boundary, and thus transitional shorebirds cannot be the basal lineage to all modern birds. Nevertheless, basal avian relationships including the phylogenetic position of the shorebirds remain controversial and there is an urgent need to construct a higher-level phylogeny of birds and to provide estimates, independent of fossil evidence, of the timing of major splits in the avian tree.

Traditionally, modern birds have been divided into palaeognaths (ratites and tinamous) and neognaths (all other birds), with palaeognaths or a palaeognaths/Galliformes/Anseriformes clade placed as the basal lineage in birds (Cracraft 1988; Sibley & Ahlquist 1990; Cooper & Penny 1997; Groth & Barrowclough 1999; Van Tuinen *et al.* 2000; Haddrath & Baker 2001). By

contrast, other recent analyses using mitochondrial DNA sequences, including complete mitochondrial DNA genomes, suggest that passerines are the sister group to other modern bird lineages (Härlid & Arnason 1999; Mindell *et al.* 1999; Waddell *et al.* 1999; Johnson 2001). Because of the apparent incongruence among ordinal-level studies thus far, sequencing efforts have now shifted to a strategy that emphasizes large numbers of taxa, and smaller-sized datasets of nuclear sequences (Groth & Barrowclough 1999), mitochondrial sequences (Johnson 2001) or a combination of both (Van Tuinen *et al.* 2000). These most recent studies (using nuclear sequences) support neognath monophyly, but resolution at some phylogenetic levels is limited. Because large sequence datasets have a much higher probability of recovering the 'correct' tree (Cao *et al.* 1994; Charleston *et al.* 1994; Hillis *et al.* 1994; Cummings *et al.* 1995; Mindell & Thacker 1996; Russo *et al.* 1996; Zardoya & Meyer 1996; Naylor & Brown 1997; Rosenberg & Kumar 2001), we therefore used phylogenetic information from all of the protein-coding and ribosomal genes of the mitochondrial genome to reduce the sampling bias and improve resolution of major avian lineages.

The correct rooting of the avian tree is fundamentally important not only in resolving the evolutionary history of birds but also in understanding rates and patterns of molecular evolution in their mitochondrial DNA genomes. Incorrect rooting of vertebrate trees can be a consequence of inappropriate models of substitution that do not account for rate variation among sites (Takezaki & Gojobori 1999), and by using outgroups that are too divergent from the ingroup.

\* Author for correspondence (tara.paton@utoronto.ca).

In this paper we report the sequence of the first entire mitochondrial genomes of two shorebirds, the ruddy turnstone (*Arenaria interpres*) and the blackish oystercatcher (*Haematopus ater*), and use the concatenated sequences of 12 protein-coding genes, two rRNA genes and 19 tRNA genes from our study and other published avian genomes to re-examine relationships among modern birds. Our analyses suggest that the basal split in the avian tree is between neognaths and palaeognaths, and that shorebirds are nested within neognaths. Thus in keeping with earlier molecular studies we reject the derivation of all modern birds from transitional shorebirds.

## 2. MATERIAL AND METHODS

Purified mitochondrial DNA was isolated from liver and heart tissue of the ruddy turnstone and blackish oystercatcher using caesium chloride gradients, and total genomic DNA was extracted by a standard phenol–chloroform protocol (Sambrook 1989). PCR amplification was performed on these templates using overlapping primer sets that sampled the entire mitochondrial DNA genome. To further verify that the sequences were from mitochondrial DNA we also performed long template PCR (Boehringer Mannheim) on total genomic DNA from *A. interpres* with mitochondrial primers to yield 10 to 15 kb templates, which were used for PCR amplification and sequencing of shorter fragments. Sequencing reactions were performed with radio-labelled and DYEnamic ThermoSequenase kits (Amersham Pharmacia). Sequences generated using fluorescent-based technology were read on a LICOR 4200 bidirectional automated sequencer that simultaneously read both strands for easy confirmation.

Phylogenetic analysis was based on concatenated mitochondrial sequences of 12 protein-coding genes, two rRNA genes and 19 tRNA genes from the ruddy turnstone (*A. interpres*) AY074885, blackish oystercatcher (*H. ater*) AY074886, and the published sequences of chicken (*Gallus gallus*) X52392, greater rhea (*Rhea americana*) AF090339, lesser rhea (*Pterocnemia pennata*) AF338709, ostrich (*Struthio camelus*) AF338715, great spotted kiwi (*Apteryx haastii*) AF338708, two extinct species of moas (*Anomalopteryx didiformis*) AF338714 and (*Emeus crassus*) AF338712, emu (*Dromaius novaehollandiae*) AF338711, southern cassowary (*Casuarius casuarius*) AF338713, elegant crested tinamou (*Eudromia elegans*) AF338710, great tinamou (*Tinamus major*) AF338707, redhead duck (*Aythya americana*) AF090337, peregrine falcon (*Falco peregrinus*) AF090338, village indigobird (*Vidua chalybeata*) AF090341, grey-headed broadbill (*Smithornis sharpie*) AF090340, and rook (*Corvus frugilegus*) Y18522. The sequences of the alligator (*Alligator mississippiensis*) Y13113 and African side-necked turtle (*Pelomedusa subrufa*) NC001947 were used to root the tree. The ND6 gene was omitted because it is encoded on the opposite strand and has a very different base composition to the other genes. For the ND3 gene, base 174 was not included as this nucleotide is thought to be removed by RNA editing mechanisms before translation in birds (Härlid *et al.* 1997; Mindell *et al.* 1998b). For the concatenated sequences, indels of codons were removed from the alignment to preserve the reading frame of the proteins. Because of the problem of multiple hits in this deep branch phylogeny, the analysis did not include third codon positions in sequences. CLUSTAL X (Thompson *et al.* 1997) was used to align sequences of the rRNA and tRNA genes, and the alignments further improved using models of secondary structure (Kumazawa &

Nishida 1993; De Rijk *et al.* 2000). Gaps and loops were excluded from tRNA sequences.

Aligned DNA sequences were tested for stationarity in base composition at variable sites using TREE-PUZZLE-5.0 (Strimmer & von Haeseler 1996). Because these sequences show significant heterogeneity in base composition among taxa ( $p < 0.01$ ) we performed maximum likelihood (ML) analyses under a non-homogeneous model of substitution with a correction for rate variation among sites (HKY + G) using the program N<sub>H</sub>ML (Galtier & Gouy 1998). With the number of taxa ( $n = 20$ ) in this study, it is not possible to conduct a tree search with N<sub>H</sub>ML. Instead, we computed the log likelihood of all possible trees (105) within the constrained tree space: ((tinamous, (moas, (rheas, (ostrich, (kiwi, (emu, cassowary))))), (chicken, duck), (turnstone, oystercatcher), falcon, ((broadbill, (indigobird, rook))). This constrained tree assumes that the two species of shorebirds are sisters, and that the palaeognaths, passerines and Galloanserae are monophyletic groups. These assumptions are based on: (i) phylogenetic analyses carried out using both ML with GTR + I + G distances and neighbour-joining with log-determinant distances in the program PAUP\* v. 4.0b8 (Swofford 1998); and (ii) previous molecular studies on these groups that have corroborated their monophyly (Raikow 1982; Stapel *et al.* 1984; Sibley & Ahlquist 1990; Caspers *et al.* 1997; Groth & Barrowclough 1999; Mindell *et al.* 1999; Van Tuinen *et al.* 2000; Haddrath & Baker 2001; Irestedt *et al.* 2001; Johansson *et al.* 2001). Additionally, relationships among the Palaeognathae are based on recent findings using complete mitochondrial genomes (Haddrath & Baker 2001).

The branch lengths of the ML tree were estimated using a non-homogeneous model of substitution (HKY + G) in the program PAML v. 3.0a (Yang 2000). The divergence time of each node in this tree was then estimated in the program R8s v. 1.01b (Sanderson 2002), using both penalized likelihood (PL) and non-parametric rate smoothing (NPRS) of variable substitution rates among branches in the tree. PL uses a parametric model that has a different substitution rate on each branch, and a non-parametric roughness penalty that costs the model more if rates change too quickly between branches (Sanderson 2002). Cross-validation analysis (Sanderson 2002) was performed on the ML tree to determine the optimal smoothing parameter for the data. The nodes were calibrated using: (i) the emu–cassowary split estimated at approximately 35 Myr ago based on the 25 million year old fossil *Emuaris gidju* from Australia (from the lineage leading to emu) (Boles 1992); (ii) the alligator–bird split estimated at 245 Myr ago based on the oldest fossil members of the crocodylian and bird lines (Benton 1990; M. Benton, personal communication); and (iii) the Galloanserae divergence time of 85 Myr ago based on previous estimates (Haddrath & Baker 2001, Van Tuinen & Hedges 2001). The standard errors of the divergence times were calculated by performing branch-length estimation and time of divergence calculations on 100 non-parametric replicate datasets generated in PHYLIP v. 3.5c (Felsenstein 1993). The bootstrap distribution of the age of each node was compiled, and the 2.5th and 97.5th percentiles of the distribution formed the limits for the bootstrap replicate percentile confidence interval (C.I.) (Sanderson & Doyle 2001).

## 3. RESULTS

### (a) *Sequence and characterization of mitochondrial genes in Arenaria interpres and Haematopus ater*

The complete mitochondrial genomes of the Charadriiformes examined here comprise *ca.* 16 700 bp. The

mitochondrial genome consists of 13 protein-coding genes, two rRNA genes, 22 tRNA genes and an origin of heavy strand replication (the control region). The gene order is consistent with that in the genomes of Galliformes, Anseriformes, Passeriformes and palaeognath genomes sequenced to date (Desjardins & Morais 1990; Härlid & Arnason 1999; Mindell *et al.* 1999; Haddrath & Baker 2001) and that found in another member of the Charadriiformes, the American woodcock (*Scolopax minor*) (Mindell *et al.* 1998a).

Examination of the nucleotide composition of protein-coding genes revealed significant departures from stationarity ( $p < 0.05$ ) in the tinamous, rheas, passerines, falcon and redhead duck. Among the sequences of the rRNA genes, only the passerines and the falcon sequences had nucleotide compositions that were significantly different from the other taxa examined here.

### (b) Phylogenetic analysis

ML analysis (Felsenstein 1981) was conducted in PAUP\* v. 4.068 (Swofford 1998) using the GTR + I + G model of substitution with  $I = 0.281$  and  $\alpha = 0.443$ . When the alligator and turtle were used as outgroups to root the avian tree, the optimal topology placed the Palaeognathae as the sister group of all other birds. The Galloanserae form the oldest lineage of the Neognathae, followed by the Passeriformes, Falconiformes and Charadriiformes. This topology is quite different if mammalian genomes are used as an additional outgroup, as this rooting reconstructs a basal avian divergence between Passeriformes and remaining avian orders (Mindell *et al.* 1999). Because there are significant deviations from stationarity in these data, ML analyses were also performed under non-homogeneous models of substitution for all possible topologies (105) within the defined tree space (see § 2). The tree with the best likelihood (figure 1) confirms the split between the Palaeognathae and Neognathae, and places the Galloanserae as the sister group to all other neognath birds. There is no evidence from these data that the Charadriiformes or Passeriformes lineages are ancestral to other birds, but rather appear to be more recently derived groups.

### (c) Dating the divergence of avian orders

To determine if a molecular clock can be applied to date divergences between avian lineages, log-likelihood ratio tests of the clock versus non-clock tree were also performed in PAML v. 3.0a (Yang 2000). The hypothesis of a molecular clock was rejected for these sequences. Relative rate tests revealed that the falcon, passerine and tinamou lineages are evolving significantly faster than other birds, and the ratites are evolving significantly slower. The cross-validation analysis in R8s (Sanderson 2002) yielded an optimal smoothing parameter between 0.1 and unity, which reflects the considerable rate variation among lineages. The smoothing parameter was therefore set at a value of unity, as there was very little difference in the estimates of divergence times using 0.1 or unity. When rate smoothing was applied with a time-calibration of 245 Myr ago for the bird–alligator node, and with two additional time constraints at the Galloanserae node (85 Myr ago) and the emu–cassowary node (35 Myr ago), the origin of modern birds was dated between 108.2 and 155.5 Myr ago with a mean estimate of 123 Myr ago

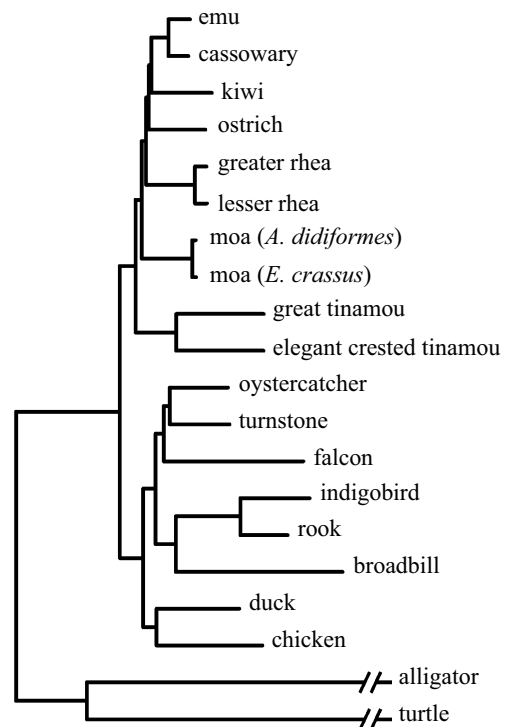


Figure 1. Non-homogeneous maximum likelihood tree topology from the concatenated sequences of protein-coding, ribosomal and tRNA genes.

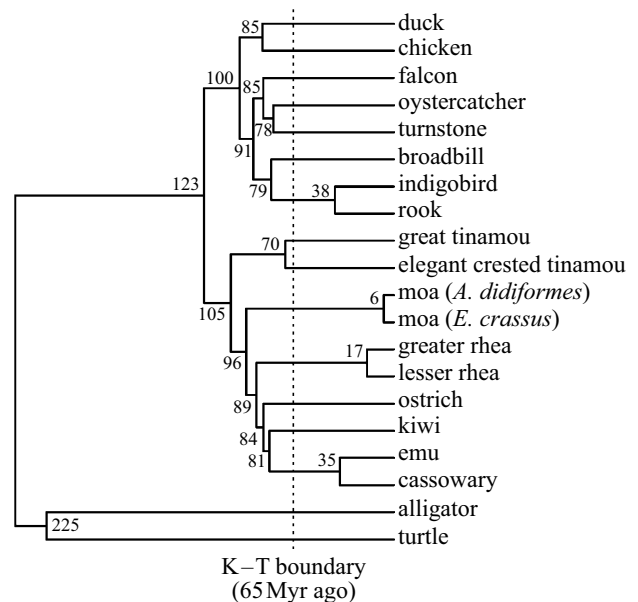


Figure 2. Chronogram showing estimated dates of divergence among avian lineages (in millions of years). 95% bootstrap percentile confidence intervals are given in table 1. (K–T, Cretaceous–Tertiary boundary).

(figure 2). The divergence of the Galloanserae from other neognath birds was estimated to be 88.8–112.4 Myr ago (mean estimate, 100 Myr ago); the radiation of the Passeriformes between 82.7 and 104.9 Myr ago (mean estimate, 91 Myr ago) and the divergence of shorebirds, as represented by the ruddy turnstone and blackish oystercatcher, from the Falconiformes between 79 and 99.5 Myr ago (mean estimate, 85 Myr ago). Divergence

Table 1. Estimated dates of divergence of avian lineages along with 95% confidence limits of bootstrapped replicate datasets for two rate-smoothing methods.

(NA denotes nodes that were constrained during the estimation of time of branching events.)

node	penalized likelihood		non-parametric	
	divergence date	95% C.I.	divergence date	95% C.I.
palaeognaths/neognaths	123	(108.2, 155.5)	121	(108.6, 154.6)
neognaths	100	(88.8, 112.4)	99	(90.8, 110.8)
Galloanserae	85	NA	85	NA
passerines	91	(82.7, 104.9)	90	(84.2, 104.0)
falcon/shorebirds	85	(79.0, 99.5)	84	(78.9, 98.2)
shorebirds	78	(71.0, 94.0)	77	(71.5, 93.3)
oscine/suboscine	79	(71.6, 90.8)	78	(71.7, 90.2)
indigo/rook	38	(32.6, 42.0)	38	(32.4, 41.6)
palaeognaths	105	(88.2, 137.5)	104	(87.7, 135.4)
tinamous	70	(58.5, 91.7)	69	(58.6, 90.5)
moas	96	(87.0, 133.7)	94	(86.4, 133.2)
between moas	6	(4.1, 9.8)	6	(3.9, 9.8)
rheas	89	(83.0, 127.4)	88	(82.4, 127.2)
between rheas	17	(14.1, 25.2)	17	(14.5, 25.1)
ostrich	84	(81.6, 120.6)	84	(80.8, 121.5)
kiwi	81	(75.8, 116.3)	80	(76.5, 116.2)
emu/cassowary	35	NA	35	NA
alligator/turtle	225	(205.4, 238.3)	225	(205.3, 238.3)

dates estimated using both PL and NPRS in R8s (Sanderson 2002) were very similar, as were the 95% C.I. of the bootstrapped replicate datasets (table 1).

#### 4. DISCUSSION

##### (a) *Phylogenetic reconstruction and rooting of the bird tree*

Sequence similarity between Charadriiformes and other birds in this study ranges from 1% to 21% for amino acids, 3% to 13% for nucleotides of protein-coding genes (first and second codon positions), 1% to 20% for ribosomal genes and 1% to 16% for tRNA genes, which is in agreement with other demonstrations of rate heterogeneity among genes in mitochondrial DNA (Kumar 1996). Rate variation among nucleotide sites is substantial, requiring a gamma parameter in the model of substitution for better phylogenetic estimates (Yang 1996). Additionally, there are unequal nucleotide compositions among the avian lineages examined here, which violate basic assumptions of many tree-building methods and can bias phylogeny reconstruction (Lockhart *et al.* 1994; Galtier & Gouy 1995; Eyre-Walker 1998; Foster & Hickey 1999; Chang & Cambell 2000; Mooers & Holmes 2000; Haddrath & Baker 2001). These features of sequence evolution are further complicated by differing rates of evolution among the bird lineages, most notably revealed in rates of substitution of protein-coding genes, which may result in construction of a 'rate tree' rather than a species tree because of attraction to faster-evolving outgroups. Despite the fact that ribosomal and tRNA genes appear to be less sensitive to rate variation among lineages, the total evidence tree is not recovered using only sequences of these RNA genes. This result is not unexpected because many studies have established that larger amounts of sequence data have a higher probability of recovering the 'correct' tree (Charleston *et al.* 1994; Hillis *et al.* 1994; Cummings *et*

*al.* 1995; Mindell & Thacker 1996; Russo *et al.* 1996; Naylor & Brown 1997; Rosenberg & Kumar 2001). Therefore, it is preferable to use a dataset containing concatenated mitochondrial DNA sequences to minimize sampling error across sites in smaller datasets of a few genes, and to use methods of tree construction that correct for sequence heterogeneity among taxa.

Because of variation in the rate of evolution in different lineages, phylogenetic relationships among the birds in this study are also significantly affected by the placement of the root. If mammals are selected as outgroups to birds, the tree is rooted so that passerines are basal to all other birds, as found in other studies using mitochondrial DNA (Härlid *et al.* 1998; Härlid & Arnason 1999; Mindell *et al.* 1999; Waddell *et al.* 1999). In previous arrangements, the falcon is also relatively basal and groups with the suboscine taxon, strongly suggesting attraction between faster-evolving outgroups and these avian lineages that are significantly accelerated in rate compared with other birds. When mammals are excluded and the avian tree is rooted with only reptiles, this apparent long-branch attraction is much reduced and the topology we obtained (figure 1) is the most strongly supported. A split between palaeognath and neognath birds is consistent with several traditional classifications and lines of evidence including morphological, chromosomal and biochemical data (De Boer 1980; Stapel *et al.* 1984; Cracraft & Mindell 1989; Van Tuinen *et al.* 1998; Groth & Barrowclough 1999). The Galloanserae is assigned to the Neognathae rather than the Palaeognathae, contrary to evidence using DNA-DNA hybridization distances (Sibley & Ahlquist 1990). The optimal phylogeny of figure 1 supports a derived position for passerines, contrary to recent evidence suggesting their basal origins (Mindell *et al.* 1997, 1999; Härlid *et al.* 1998; Waddell *et al.* 1999). There is very weak support for the alliance of the falcon and passerine lineages (Mindell *et al.* 1999), as the optimal tree places the falcon as sister to the

shorebird lineage. Furthermore, the shorebird lineage is not supported as basal in the avian tree.

### (b) *Dating avian divergences*

The difficulty of dating avian branching events arises from both the significant rate variation among lineages of birds, and from using the date of divergence of alligators from birds as a reference when bird mitochondrial DNA is evolving at about half the rate of that of the alligator. Calibrating avian trees using fossils or molecular dates outside of birds can result in underestimates of divergence times within birds (Hedges *et al.* 1996; H  rlid *et al.* 1998; H  rlid & Arnason 1999), and thus it is necessary to use a dating method that does not assume a single rate of evolution for birds or their outgroups. Previous methods have dealt with rate variation among lineages by eliminating taxa or genes that do not evolve in a clock-like manner, or by using local molecular clocks in a phylogeny (Hedges *et al.* 1996; Cooper & Penny 1997; Kumar & Hedges 1998; Rambaut & Bromham 1998; Waddell *et al.* 1999; Van Tuinen & Hedges 2001). The rate-smoothing process estimates the rate of evolution for each branch in the tree and minimizes rate changes between ancestral and descendant lineages (Sanderson 1997), no rate-constancy assumptions are needed and all genetic information and lineages can be retained in the analysis. This method was recently used to date speciation events among the ratites (Haddrath & Baker 2001), because they were shown to evolve significantly slower than other modern lineages of birds examined. In this study, we further attempt to improve upon estimates of divergence times by using a calibration point at the root, and by applying constraints in both the palaeognath and neognath clades of the tree. Because of the disparity in rates of evolution within birds, these internal points allow more reasonable approximations of rates for individual clades (Sanderson & Doyle 2001). NPRS has been shown in plants to allow too much rate-smoothing relative to the optimal level calculated with cross-validation in PL (Sanderson 2002). However, in our dataset, both methods give nearly identical estimates of rates of evolution and mean times of divergence because of the accelerated rate in the outgroup alligator and turtle sequences. Cross-validation thus selects extreme rate-smoothing as optimal in PL, similar to that used in NPRS.

The mean estimate of the palaeognath–neognath split is 123 Myr ago (C.I. 108, 156), placing the divergence times of modern birds as a group as well as many lineages in this study well before the K–T boundary. This date for the origin of modern birds is close to recent estimates using both nuclear (119 Myr ago; Van Tuinen & Hedges 2001) and mitochondrial genes (110 Myr ago; Haddrath & Baker 2001), supporting theories of continental break-up as a driving force in avian speciation events (Cracraft 2001; Haddrath & Baker 2001; Van Tuinen & Hedges 2001). These findings indicate that many more avian lineages have survived this extinction event than hypothesized by Feduccia (1995). With the broader range of mitochondrial DNA genomes included in this study, the rate-smoothing methods provide mean estimates for the divergence times of major lineages of the ratites in excess of 80 Myr ago. Because the C.I. include continental separation dates, vicariance cannot be excluded for most lineage splits in this group. Thus both the kiwi and ostrich

lineages appear to conform with the vicariance biogeography hypothesis for Gondwana origins of ratites (see Cooper *et al.* 2001; Haddrath & Baker 2001).

### (c) *Early bird origins and evidence for 'transitional shorebirds'*

If modern birds originated from a founder lineage of 'transitional shorebirds' (Feduccia 1995), this should be evidenced in the DNA sequences of extant avian lineages and current interpretations of the fossil record. Our phylogenetic analysis and estimation of avian divergences allows the testing of specific assumptions based on the hypotheses of Feduccia (1995, 1996). First, if the radiation of modern birds occurred over a time-frame of 5 to 10 Myr around the K–T extinction event, the phylogenetic relationships among modern groups of birds should be very difficult to resolve. However, we have demonstrated that long mitochondrial DNA sequences provide good resolution at the ordinal level in birds, and do not show a 'star phylogeny' typical of a rapid radiation at the K–T boundary. Instead, ordinal diversification among the neognath birds in this study has occurred over a period of close to 44 Myr (longer among Palaeognaths).

Second, Feduccia (1996) maintained that the orders Anseriformes, Gruiformes, Procellariiformes, Podicipediformes, Pelecaniformes along with the Psittaciformes and Columbiformes were descended from 'ancient shorebird stock', suggesting that these orders should be closely related and have divergence times that predate most other avian orders. However, there is now considerable evidence that Anseriformes and Galliformes are sister groups (Sibley & Ahlquist 1990; Caspers *et al.* 1997; Livezey 1997; Mindell *et al.* 1997, 1999; Groth & Barrowclough 1999; Van Tuinen *et al.* 2000), a finding also corroborated in this study. Although representatives of Gruiformes, Procellariiformes, Podicipediformes and Pelecaniformes have not been examined here, other studies have not found these lineages to be reciprocally monophyletic (Sibley & Ahlquist 1990; Mindell *et al.* 1997; Groth & Barrowclough 1999; Van Tuinen *et al.* 2000, 2001). Further, the divergence dates of Palaeognaths, Galliformes and other birds in this study predate that of Charadriiformes.

Third, the 'transitional shorebird' hypothesis is based on the interpretation of early avian fossils such as *Presbyornis*, *Juncitarsus* and *Rhynchaeites* specimens as shorebird-modern order mosaics, suggesting that Charadriiformes should have close phylogenetic affinities to Anseriformes, Phoenicopteriformes and Ciconiiformes, and that these groups should all be relatively basal in the avian tree. *Presbyornithidae* fossils have been studied in a phylogenetic context with modern Anseriformes and have been unequivocally placed as a sister to the ducks, geese and swans (Anatidae) (Ericson 1997; Livezey 1997). Therefore, the idea of *Presbyornis* as an ancient Charadriiform is not supported. Further, contrary to osteological studies (Feduccia 1976; Ericson 1997), there is no molecular evidence that unites the Phoenicopteriformes (flamingos) with the Charadriiformes, but instead places flamingos as a sister clade to grebes (Podicipediformes) (Van Tuinen *et al.* 2001) or alternatively to Ciconiiformes and Pelecaniformes (Sibley & Ahlquist 1990). Similarly, DNA evidence does not support Ciconiiformes (defined

as the families Ardeidae, Ciconiidae and Threskiornithidae) as closer to Charadriiformes than any other avian group (Sibley & Ahlquist 1990; Van Tuinen *et al.* 2001). The interpretation of *Presbyornis*, *Juncitarsus* and *Rhynchaetites* as shorebird-modern order mosaics is erroneous and is not supported by current palaeontological studies or examination of genealogical relationships among modern bird orders.

Finally, a radiation of modern bird lineages after the K–T boundary should result in dates of divergence of birds that are no older than 65 Myr. A second radiation during the late Oligocene (33.7–23.8 Myr ago) of passerine birds (Feduccia 1996) should also be apparent in the divergence dates of this group. Our mean estimate of the time that passerines diverged from other neognath birds was 91 Myr ago (C.I. 83, 105), and the split between oscine and suboscine clades was 79 Myr ago (C.I. 72, 91). The oscine–suboscine divergence was found to be 77 Myr ago using other molecular markers (Sibley & Ahlquist 1990; Van Tuinen & Hedges 2001) and thus serves as an independent check of our date that is based on complete mitochondrial DNA genomes. The oldest passerine fossils from Australia date to the early Eocene (*ca.* 55 Myr ago) (Boles 1995) but cannot be classified in any specific passerine family at this time, and thus are not useful to date divergences within Passeriformes. However, these specimens demonstrate that this group is much older than previously believed and that the fossil record of birds remains incomplete.

Our re-examination of the rooting of the avian tree has recovered the basal split between the Palaeognathae and Neognathae clades reflected in historical classifications, and casts considerable doubt on the putative basal positions of either Charadriiformes or Passeriformes relative to other birds. We contend that appropriate models of sequence evolution combined with more appropriate outgroups increases the likelihood of recovering the correct topology of the avian tree, and thus will allow better estimates of divergence times. Avian fossils may assist in estimating the minimum age of branching events, following a careful examination of their relationships to modern lineages. Presbyornithidae fossils are known from the Maastriichtian (64–75 Myr ago), with possible specimens from the Campanian (74–83 Myr ago) (Unwin 1993), and therefore the molecular estimates of approximately 85 Myr for the divergence of Galliformes from Anseriformes (Haddrath & Baker 2001; Van Tuinen and Hedges 2001) are reasonable to use as a calibration point to date other divergences among bird lineages. The history of modern bird lineages will be better understood when a well-resolved phylogeny with much broader taxon sampling is used in conjunction with the growing body of palaeontological evidence.

Funding was provided by grants to A.J.B. from the Natural Sciences and Engineering Research Council of Canada and the Royal Ontario Museum Foundation.

## REFERENCES

- Benton, M. J. 1990 Phylogeny of the major tetrapod groups: morphological data and divergence dates. *J. Mol. Evol.* **30**, 409–424.
- Boles, W. E. 1992 Revision of *Dromaius gidju* Patterson and Rich 1987 for Riversleigh, Northwestern Queensland, Australia, with a reassessment of its generic position. *Nat. Hist. Mus. Los Angeles County, Sci. Series* **36**, 195–208.
- Boles, W. E. 1995 The world's oldest songbird. *Nature* **374**, 21–22.
- Cao, Y., Adachi, J., Janke, A., Paabo, 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.
- Caspers, G.-J., de Weerd, D., Wattel, J. & de Jong, W. W. 1997 alpha-Crystallin sequences support a galliform/anseriform clade. *Mol. Phylogenet. Evol.* **7**, 185–188.
- Chang, B. S. & Campbell, D. L. 2000 Bias in phylogenetic reconstruction of vertebrate rhodopsin sequences. *Mol. Biol. Evol.* **17**, 1220–1231.
- Charleston, M. A., Hendy, M. D. & Penny, D. 1994 The effects of sequence length, tree topology and number of taxa on the performance of phylogenetic methods. *J. Comput. Biol.* **1**, 133–151.
- Cooper, A. & Penny, D. 1997 Mass survival of birds across the Cretaceous–Tertiary boundary: molecular evidence. *Science* **275**, 1109–1113.
- Cooper, A., Lalueza-Fox, C., Anderson, S., Rambaut, A., Austin, J. & Ward, R. 2001 Complete mitochondrial genome sequences of two extinct moas clarify ratite evolution. *Nature* **409**, 704–707.
- Cracraft, J. 1988 The major clades of birds. In *The phylogeny and classification of the tetrapods*, vol. 1 (amphibians, reptiles, birds) (ed. M. J. Benton), pp. 339–361. Systematics Association Special Volume No. 35A. Oxford: Clarendon Press.
- Cracraft, J. 2001 Avian evolution, Gondwana biogeography and the Cretaceous–Tertiary mass extinction event. *Proc. R. Soc. Lond. B* **268**, 459–469. (DOI 10.1098/rspb.2000.1368.)
- Cracraft, J. & Mindell, D. P. 1989 The early history of modern birds: a comparison of molecular and morphological evidence. In *The hierarchy of life* (ed. B. Fernholm, K. Bremer & H. Jornvall), pp. 389–403. Amsterdam: Elsevier.
- 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 Wachter, 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.
- Ericson, P. G. P. 1997 Systematic relationships of the palaeogene family Presbyornithidae (Aves: Anseriformes). *Zool. J. Linn. Soc.* **21**, 429–483.
- Eyre-Walker, A. 1998 Problems with parsimony in sequences of biased base composition. *J. Mol. Evol.* **47**, 686–690.
- Feduccia, A. 1976 Osteological evidence for shorebird affinities of the flamingos. *Auk* **93**, 587–601.
- Feduccia, A. 1995 Explosive evolution in Tertiary birds and mammals. *Science* **267**, 637–638.
- Feduccia, A. 1996 *The origin and evolution of birds*. New Haven, CT: Yale University Press.
- Felsenstein, J. 1981 Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.* **17**, 368–376.
- Felsenstein, J. 1993 PHYLIP (*Phylogeny inference package*) v. 3.5c. Distributed by the author. Seattle: University of Washington.
- 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. & Gouy, M. 1995 Inferring phylogenies from DNA sequences of unequal base compositions. *Proc. Natl Acad. Sci. USA* **92**, 11 317–11 321.
- Galtier, N. & Gouy, M. 1998 Inferring pattern and process: maximum likelihood implementation of a nonhomogenous model of DNA sequence evolution for phylogenetic analysis. *Mol. Biol. Evol.* **15**, 871–879.
- Groth, J. G. & Barrowclough, G. F. 1999 Basal divergences in birds and the phylogenetic utility of the nuclear RAG-1 gene. *Mol. Phylogenet. Evol.* **12**, 115–123.
- Haddrath, O. P. & Baker, A. J. 2001 Complete mitochondrial DNA genome sequences of extinct birds: ratite phylogenetics and the vicariance biogeography hypothesis. *Proc. R. Soc. Lond. B* **268**, 939–945. (DOI 10.1098/rspb.2001.1587.)
- Härlid, A. & Arnason, U. 1999 Analyses of mitochondrial DNA nest ratite birds within the Neognathae: supporting a neotenus origin of ratite morphological characters. *Proc. R. Soc. Lond. B* **266**, 305–309. (DOI 10.1098/rspb.1999.0638.)
- Härlid, A., Janke, A. & Arnason, U. 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. & Arnason, U. 1998 The complete mitochondrial genome of *Rhea americana* and early avian divergences. *J. Mol. Evol.* **46**, 669–679.
- Hedges, S. B., Parker, P. H., Sibley, C. G. & Kumar, S. 1996 Continental breakup and the ordinal diversification of birds and mammals. *Nature* **381**, 226–229.
- Hillis, D. M., Huelsenbeck, J. P. & Cunningham, C. W. 1994 Application and accuracy of molecular phylogenies. *Science* **264**, 671–677.
- Irestedt, M., Johansson, U. S., Parsons, T. J. & Ericson, P. G. P. 2001 Phylogeny of major lineages of suboscines (Passeriformes) analysed by nuclear DNA sequence data. *J. Avian Biol.* **32**, 15–25.
- Johansson, U. S., Parsons, T. J., Irestedt, M. & Ericson, P. G. P. 2001 Clades within the 'higher land birds', evaluated by nuclear DNA sequences. *J. Zool. Syst. Evol. Res.* **39**, 37–51.
- Johnson, K. P. 2001 Taxon sampling and the phylogenetic position of passeriformes: evidence from 916 avian cytochrome *b* sequences. *Syst. Biol.* **50**, 128–136.
- Kumar, S. 1996 Patterns of nucleotide substitution in mitochondrial protein coding genes of vertebrates. *Genetics* **143**, 537–548.
- Kumar, S. & Hedges, S. B. 1998 A molecular timescale for vertebrate evolution. *Nature* **392**, 917–920.
- Kumazawa, Y. & Nishida, M. 1993 Sequence evolution of mitochondrial tRNA genes and deep-branch animal phylogenetics. *J. Mol. Evol.* **37**, 380–398.
- Livezey, B. C. 1997 A phylogenetic analysis of basal Anseriformes, the fossil *Presbyornis*, and the interordinal relationships of waterfowl. *Zool. J. Linn. Soc.* **21**, 361–428.
- Lockhart, P. J., Steel, 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. & Thacker, C. E. 1996 Rates of molecular evolution: phylogenetic issues and applications. *A. Rev. Ecol. Syst.* **27**, 279–303.
- Mindell, D. P., Sorenson, M. D., Huddleston, C. J., Miranda Jr, H. C., Knight, A., Sawchuk, S. J. & Yuri, T. 1997 Phylogenetic relationships among and within select avian orders based on mitochondrial DNA. In *Avian molecular evolution and systematics* (ed. D. P. Mindell), pp. 214–247. New York: Academic.
- Mindell, D. P., Sorenson, M. D. & Dimcheff, D. E. 1998a 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. 1998b 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.
- Mooers, A. O. & Holmes, E. C. 2000 The evolution of base composition and phylogenetic inference. *Trends Ecol. Evol.* **15**, 365–369.
- Naylor, G. J. & Brown, W. M. 1997 Structural biology and phylogenetic estimation. *Nature* **388**, 527–528.
- Raikow, R. J. 1982 Monophyly of the Passeriformes: test of a phylogenetic hypothesis. *Auk* **99**, 431–445.
- Rambaut, A. & Bromham, L. 1998 Estimating divergence dates from molecular sequences. *Mol. Biol. Evol.* **15**, 442–448.
- Rosenberg, M. S. & Kumar, S. 2001 Incomplete taxon sampling is not a problem for phylogenetic inference. *Proc. Natl Acad. Sci. USA* **98**, 10 751–10 756.
- 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. 1989 *Molecular cloning: a laboratory manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Sanderson, M. J. 1997 A nonparametric approach to estimating divergence times in the absence of rate constancy. *Mol. Biol. Evol.* **14**, 1218–1231.
- Sanderson, M. J. 2002 Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Mol. Biol. Evol.* **19**, 101–109.
- Sanderson, M. J. & Doyle, J. A. 2001 Sources of error and confidence intervals in estimating the age of Angiosperms from *rbcL* and 18S rDNA data. *Am. J. Bot.* **88**, 1499–1516.
- Sibley, C. G. & Ahlquist, J. E. 1990 *Phylogeny and classification of birds: a study in molecular evolution*. New Haven, CT: Yale University Press.
- Stapel, S. O., Leunissen, J. A., Versteeg, M., Wattel, J. & de Jong, W. W. 1984 Ratites as oldest offshoot of avian stem: evidence from alpha-crystallin 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. 4.0b8. Sunderland, MA: Sinauer.
- Takezaki, N. & Gojobori, T. 1999 Correct and incorrect vertebrate phylogenies obtained by the entire mitochondrial DNA sequences. *Mol. Biol. Evol.* **16**, 590–601.
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, J. & Higgins, D. G. 1997 The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**, 4876–4882.
- Unwin, D. M. 1993 Aves. In *The fossil record*, vol. 2 (ed. M. J. Benton), pp. 717–738. London: Chapman & Hall.
- Van Tuinen, M. & Hedges, S. B. 2001 Calibration of avian molecular clocks. *Mol. Biol. Evol.* **18**, 206–213.
- 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.
- Van Tuinen, M., Sibley, C. G. & Hedges, S. B. 2000 The early history of modern birds inferred from DNA sequences of nuclear and mitochondrial ribosomal genes. *Mol. Biol. Evol.* **17**, 451–457.
- Van Tuinen, M., Butrill, D. B., Kirsch, J. A. W. & Hedges, S. B. 2001 Convergence and divergence in the evolution of

- aquatic birds. *Proc. R. Soc. Lond. B* **268**, 1345–1350. (DOI 10.1098/rspb.2001.1679.)
- Waddell, P. J., Cao, Y., Hasegawa, M. & Mindell, D. P. 1999 Assessing the Cretaceous superordinal divergence times within birds and placental mammals by using whole mitochondrial protein sequences and an extended statistical framework. *Syst. Biol.* **48**, 119–137.
- Yang, Z. 1996 Among-site rate variation and its impact on phylogenetic analyses. *Trends Ecol. Evol.* **11**, 367–372.
- Yang, Z. 2000 *Phylogenetic analysis by maximum likelihood (PAML), v. 3.0*. University College London.
- Zardoya, R. & Meyer, A. 1996 Phylogenetic performance of mitochondrial protein-coding genes in resolving relationships among vertebrates. *Mol. Biol. Evol.* **13**, 933–942.

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