! "#$%&%'(%&)

*!"#$%&'( ) *%+&', -&., +&/0-#-0, '&"(+", 1%12*

B

The study of avian phylogeny abounds with examples of both unstable taxonomy and apparent convergent evolution. Indeed, in many cases the incidence of these conditions is correlated because taxonomists are at times misled by convergence in their systematic inferences. Progress with resolving avian phylogeny and thus identification of morphological/behavioral convergence has been enhanced by the use of molecular data that provides a source of characters independent of morphology. In particular, the use of large scale DNA, and in particular mitochondrial, sequence data has proved beneficial in stabilizing the avian tree [1,2].

In contrast, a very different perspective on the phylogeny [illegible]). To these five we add two Metaves mt genomes from Genbank; great crested grebe (*P*[illegible]) and tropicbird (*P*[illegible]), reported by Watanabe et al. [10] and Yamamoto et al. (unpublished, AP009043) respectively. In addition, we report a New Zealand rail (takahe, *P*[illegible]) because it allows us to distinguish between the traditional view of kagu and rails as being members of the same order (Gruiformes [11,12]) and the novel view of kagu in Metaves and rails in Coronaves [3].

In the process of testing the Metaves/Coronaves division, the mitochondrial genomes of the new taxa reported allow us to appraise other hypotheses about the phylogeny of modern birds. Irrespective of the validity of the Metaves/Coronaves division, we expect the hummingbird and swift to be sister taxa in this dataset (Apodiformes [11,9]). The position of the Gruiformes (including rails and kagu) is more equivocal; Livezey and Zusi [9] placed the rails very deep in the Neoavian tree and sister to the an equally good alternative alignment. We then reversed the direction of the sequences for alignment which is a proven method for detecting cases where alignment is relatively arbitrary [30]. Finally, we tested whether there was any evidence of paralogous copies by amplifying, cloning and sequencing FGB-int7 from two bird species.

R[illegible]

Complete mitochondrial genomes

The six new mitochondrial genome sequences have been deposited in GenBank under the following accession numbers: ruby-throated hummingbird (*A* [illegible]: EF532935, > 16,356 bp [incomplete due to repeats in the control region]), common swift (*A* [illegible]: AM237310

3 4 5&67,8(-0, '#+9&30,8,$9%)**+,%!

The seven birds in our mitochondrial sequence dataset that represent elements of the Metaves group do not form a monophyletic clade. Ericson and co-workers [7] also failed to obtain the Metaves clade using four nuclear genes, although they could with the inclusion of FGB-int7. No other dataset, molecular nor morphological, has found direct evidence for the division of the Neoaves into Metaves and Coronaves [8,9,17]. Nor could Fain and

This may well form a sister clade to the Charadriiformes (shorebirds); our site-stripping technique showed this relationship to be stable, from which we infer a common evolutionary history. In keeping with this idea, morphological studies of extant and fossil species reveal that early flamingos resembled more typical wading birds and place the Phoenicopteridae within the order Charadriiformes [16].

DQ780881), American kestral (*F* [illegible], DQ780880), Eurasian buzzard (*B* [illegible]; AF380305), osprey (*P* [illegible], DQ780884), Blyth's hawk eagle ([illegible], AP008239), turkey vulture (*C* [illegible], AY463690), blackish oystercatcher (*H* [illegible]; AY074886), ruddy turnstone (*A* [illegible]; AY074885), southern black-backed gull (*L* [illegible], AY293619), Oriental stork (*C* [illegible]; AB026193), red-throated loon(*G* [illegible]; AY293618), little blue penguin(*E* [illegible]; AF362763), black-browed albatross(*D* [illegible]; AY158677) and Kerguelen petrel (*P* [illegible]; AY158678), white-faced heron (*A* [illegible]; DQ780878), rockhopper penguin (*E* [illegible]; NC 008138), great crested grebe (*P* [illegible]; NC 008140), frigatebird (*F* [illegible]; AP009192), Australian pelican (*P* [illegible], DQ780883), red-tailed tropicbird (*P* [illegible]; AP009043). Paleognath taxa were not included because the paleo/neognath division has been well established for mitochondrial genomes[1,2]. Thus we rooted our Neoaves trees with the six Galloanserae sequences.

Phylogenetic Analysis

Nucleotide sequences for each gene were aligned separately in Se-Al v2 [45]. Protein-coding genes were aligned using translated amino acid sequences and RNA genes were aligned based on secondary structure. The resulting dataset has 12 protein-coding genes, two rRNA genes and 21 tRNAs (lacking tRNA-Phe because sequence data is missing in some taxa). Gaps, ambiguous sites adjacent to gaps, the NADH6 (light-strand encoded), and stop codons (often incomplete in the DNA sequence), were excluded from the alignment. The full analysed mtDNA dataset was 13,229 bp in length.

In previous work [46-48] we found that RY-coding of the most variable partitions of the nucleotide data (especially the 3rd codon position) was advantageous. This recoding increases the proportion of changes on internal branches of the tree (that is, 'treeness'), reduces effective differences in nucleotide composition (relative compositional variability; RCV), and was shown to increase concordance between mitochondrial and nuclear datasets. RY-coding does improve the ML scores, but because RY-coding is not strictly nested within nucleotide-coding (M.A. Steel, pers. comm.) it is not valid to compare their respective ML scores directly. However, because of the better fit of the data to the model (higher treeness, and lower RCV) this has been our preferred method of analysis of vertebrate mitochondrial data. Thus the trees reported here have the third codon positions of 12 protein-coding genes recoded as R (instead of A & G), and Y (instead of C & T). The full data set is available [49]. Analysis used standard programs including ModelTest [50] PAUP*4.0b10 [51], MrBayes 3.1.2 [52], and consensus networks [39]. We ran 1000 unconstrained ML bootstrap replicates with PAUP*4.0b10 on the Helix computing cluster [53], plus a Bayesian analysis using chains of 10^7 generations. For some runs, we constrained the seven 'Metaves' taxa to be monophyletic (see Figure 1) and used a Shimodaira-Hasegawa (SH) test [54] implemented in PAUP* to com-

3 4 5&67 , 8(-0 , '#+9&30 , 8 , $9%) * *+,%!

2. Sac KE, De 4c F, McLe ac a PA, A a U, Pe D: **R** . Mol Phyl Evol 2007, **42:**1-13.
3. Fa MG, H 4de P: **P** . Evolution 2004, **58:**2558-2573.
4. L4c e WP: **S** **M** : J Mamm Evol 1994, **2:**255-288.
5. H I, S c e -V a a MR: **A** Cladistics 2003, **19:**181-212.
6. L4 Z-X, J Q, W b e JR, Y4a CX: **A E C** . Science 2003, **302:**1934-1940.
7. E c PGP, A de CL, B T, E a A, J a US, Ke e M, O JI, Pa TJ, Z4cc D, Ma G: **D** **N** : . Biol Lett 2008, **4:**543-547.
8. C 4bb AL: **N** **ZENK.** Mol Phylogen Evol 2004, **30:**140-151.
9. L e e BC, Z4 RL: **H** **(T , A : N). II. A** . Zool J Linn Soc 2007, **149:**1-95.
10. Wa a abe M, N a d M, T 4da TT, K ba a T, M de D, Ca Y, O ada N, Ha e a a M: **N** . Gene 2006, **378:**65-73.
11. C ac af J: **A** **, G** **C** **-T** Proc Roy Soc Lond B 2001, **268:**459-469.
12. S b e CG, A 4 JE: Phylogeny and Classification of Birds: A Study in Molecular Evolution Ne Ha e , Ya e U e P e ; 1990.
13. C ac af J: **T** **(C A).** Auk 1981, **98:**681-714.
14. a T4 e M, B4 DB, K c JAW, Hed e SB: **C** . Proc R Soc Lond B 2001, **268:**1345-1350.
15. L e e C, Z4 RL: **H** **A** . Neth J Zool 2001, **51:**179-205.
16. O SL, Fed4cc a A: **R** **(: P).**