Phylogeny of water birds inferred from mitochondrial DNA sequences of nine protein coding genes

Tsendsesmee Lkhagvajav Treutlein, Javier Gonzalez, Michael Wink

Institute of Pharmacy and Molecular Biotechnology (IPMB), Heidelberg University, Im Neuenheimer Feld 364, Germany

M. Wink (✉)
Institute of Pharmacy and Molecular Biotechnology (IPMB)
Heidelberg University
Im Neuenheimer Feld 364
69120 Heidelberg, Germany
phone: +496221-544881
e-mail: wink@uni-hd.de

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.272v1 | CC-BY 4.0 Open Access | received: 5 Mar 2014, published: 5 Mar 2014

Abstract

Background: The phylogeny of birds which are adapted to aquatic environments is controversial because of convergent evolution.

Methods: To understand water bird evolution in more detail, we sequenced the majority of mitochondrial protein coding genes (6699 nucleotides in length) of 14 water birds, and reconstructed their phylogeny in the context of other taxa across the whole class of birds for which complete mitochondrial DNA (mtDNA) sequences were available.

Results: The water bird clade, as defined by Hackett et al. (2008) based on nuclear DNA (ncDNA) sequences, was also found in our study by Bayesian Inference (BI) and Maximum Likelihood (ML) analyses. In both reconstruction methods, genera belonging to the same family generally clustered together with moderate to high statistical support. Above the family level, we identified three monophyletic groups: one clade consisting of Procellariidae, Hydrobatidae and Diomedeidae, and a second clade consisting of Sulidae, Anhingidae and Phalacrocoracidae, and a third clade consisting of Ardeidae and Threskiornithidae.

Discussion: Based on our mtDNA sequence data, we recovered a robust direct sister relationship between Ardeidae and Threskiornithidae for the first time for mtDNA. Our comprehensive phylogenetic reconstructions contribute to the knowledge of higher level relationships within the water birds and provide evolutionary hypotheses for further studies.

Keywords: Phylogeny, mitochondrial DNA, mitogenomics, water birds, Threskiornithidae, Ardeidae

Introduction

For decades, the phylogenetic placement and interrelationship of birds which are adapted to aquatic environments have been controversially discussed (Cracraft, 1981; Sibley & Ahlquist, 1990; Mindell, 1997; van Tuinen et al., 2001; Braun & Kimball, 2002; Cracraft et al., 2004; Slack et al., 2006; Watanabe et al., 2006; Ericson et al., 2006; Hackett et al., 2008; Pratt et al., 2009; Smith, 2010; Pacheco et al., 2011; Wink, 2011; Jetz et al., 2012; McCormack et al., 2013; Gibb et al., 2013). Diverse ecological and morphological specializations exist in water birds which are mainly related to feeding style, e.g. wading (storks, flamingos, and herons), foot-propelled diving (loons, grebes, cormorants) and webbed feet (pelicaniform birds). The occurrence of these traits likely involved convergent character evolution, which makes it difficult to resolve the phylogenetic relationships at the level of morphology alone (van Tuinen et al., 2001). A large scale phylogenomic study of birds (Hackett et al., 2008), which analysed ~32 kb of nuclear DNA (ncDNA) revealed that the adaptation to aquatic environments constitutes a remarkably stable evolutionary trait: the water birds (intermixed with semiaquatic species) constitute a monophyletic group (see 'group H' in Hackett et al., 2008) which is found by maximum likelihood method with relative high bootstrap support (89%). For this "water bird assemblage" the term Aequornithes has recently been introduced by Mayr (2011). However, other groups of water birds, such as flamingos, grebes, geese, ducks, swans, tropic birds, terns and gulls cluster in other independent clades (Hackett et al., 2008; Mayr, 2011). Considering phylogenetic analyses based on complete mitochondrial DNA (mtDNA), the most recent studies (Pacheco et al., 2011; Gibb et al., 2013) included only a limited number of water bird species. To analyse water bird evolution in more detail at higher taxonomic level than previously reported, we sequenced 6699 nucleotides of the majority of mitochondrial protein coding genes (i.e., *COX1, COX2, ATP8, ATP6, COX3, ND3, ND4L, ND4* and *CYTB*) of 14 species, to increase taxon number of water birds to 52 in

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.272v1 | CC-BY 4.0 Open Access | received: 5 Mar 2014, published: 5 Mar 2014

total. We reconstructed the phylogeny of these species in the context of a large number of additional taxa covering the whole bird class for which complete mtDNA data were available, using comprehensive phylogenetic reconstruction methods Bayesian Inference (BI) and Maximum Likelihood (ML).

Materials & Methods

Taxon sampling

We increased taxon density by sequencing the majority (nine) of the mitochondrial protein coding genes of further 14 water bird taxa. Sample information is given in table 1 and GenBank accession numbers are provided in table 2. Our analysed dataset comprised 100 avian taxa from which 52 were water birds. For accession numbers of the additional taxa which were retrieved from the GenBank, see Supplemental table S1. We did not handle any live vertebrate animals during this study. All animals sequenced here were sequenced from blood and feather samples provided by the Institute of Pharmacy and Molecular Biotechnology (IPMB), Heidelberg University, Germany, thus animal care approval was not required for this specific study.

DNA extraction

Collected sample material was stored following Arctander (1988). DNA was isolated from blood, pieces of muscle or ends of feather by incubation in lysis buffer [10 mM Tris (pH 7.5), 25 mM EDTA, 75 mM NaCl, 20% SDS, 20 mg /ml Proteinase K], followed by removal of proteins using a standard phenol/chloroform extraction (Sambrook et al., 1989).

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.272v1 | CC-BY 4.0 Open Access | received: 5 Mar 2014, published: 5 Mar 2014

PCR amplification and sequencing

PCR reaction mixtures contained: 100 ng DNA, 5 µL 10 × buffer [500 mM KCl, 100 mM Tris-HCl (pH 9.0 at 25°C), 1% Triton X-100], 5 µL 15 mM $MgCl_2$, 5 pmol primer, 2 µL dNTPs (1.5 mM), 0.75 U *Taq* or 1 U Red *Taq* Polymerase (Amersham Biosciences) and ddH_2O added up to a volume of 50 µL. Thermal cycling conditions consisted of an initial denaturation for 4 min at 94°C followed by 31 cycles of 45 s at 94°C, 50 s at 50–60°C and 150 s at 72°C and a final elongation of 10 min at 72°C. PCR primer sequences are documented in table 3. PCR products were sequenced directly (Cycle Sequencing Ready Reactions version 1.1 and version 3.1, Applied Biosystems). Cycle sequencing products were run on an Applied Biosystems 310 sequencer. Source and sequence of sequencing primers are shown in table 4. We obtained a total of 6699 nucleotides for the following genes: *COX1, COX2, ATP8, ATP6, COX3, ND3, ND4L, ND4* and *CYTB.*

Phylogenetic reconstruction

The sequences were aligned by BioEdit version 7.1.3.0 (Hall, 1999). We explored the model of sequence evolution which fits the data best with jModelTest version 2.1.4 (Guindon & Gascuel, 2003; Darriba et al., 2012). Bayesian inference (BI) analyses were conducted in MrBayes MPI version 3.1.2. (Ronquist & Huelsenbeck, 2003; Altekar et al., 2004). Two independent runs of 16,000,000 generations were performed along with four Markov chains and the evolutionary model GTR + Γ + I. Trees were sampled every 2000 generations and the first 800 samples were discarded as "burn-in". The "burn-in" was determined in Tracer version 1.6 (Rambaut & Drummond, 2007). Maximum likelihood (ML) tree was reconstructed using the rapid hill-climbing algorithm implemented in RaxML-HPC version 7.7.8 (Stamatakis, 2006) with the random starting tree option and the GTRGAMMA model of sequence evolution. The robustness of nodes in ML analyses was assessed by 1000 bootstrap

replicates. The both ML and BI trees were rooted with three paleognaths (Hackett et al., 2008): *Apteryx haasti*, *Dromaius novaehollandiae* and *Pterocnemia pennata*. We made use of the resources available from the CIPRES Science Gateway in order to perform our phylogenetic analyses (Miller et al., 2010).

Results and Discussion

To select for robust monophyletic clades, we run comprehensive BI and ML analyses using parallelized applications. 16,000,000 generations resulted in two chains that converged with a standard deviation of split frequencies equal to 0.006. Phylogenetic reconstructions for the dataset using BI and ML are shown in Fig. 1 and Fig. 2, respectively. The deepest bifurcation within the Neognathae constituted the split between Galloanserae (BI posterior probability [PP]: 1.0; ML bootstrap support [BS]: 100%) versus the Neoaves (PP: 1.0; BS: 93%). Passeriformes were monophyletic (PP: 1.0; BS: 100%). In BI, the water bird group as defined by 'clade H' of Hackett et al. (2008) consisting of Pelecaniformes (with Sulidae, Anhingidae, Phalacrocoracidae, Pelecanidae), Ciconiiformes (with Ardeidae, Threskiornithidae, Ciconiidae), Procellariiformes (with Procellariidae, Hydrobatidae, Diomedeidae) Sphenisciformes and Gaviiformes received a PP value of 0.78. ML also weakly suggests genetic coherence of these taxa (BS: 25%, in addition to Gruiformes). In both BI and ML reconstructions the species belonging to the same family generally clustered together with moderate to high statistical support. Above the family level, relationships were less congruent, presumably because of relative saturation of the fast evolving mtDNA at deep avian divergences. The same topologies identified across BI and ML were for a clade consisting of Procellariidae, Hydrobatidae and Diomedeidae (PP: 1.0; BS: 97%), in which the sister taxa *Calonectris* and *Puffinus* (PP: 1.0; BS: 100%) formed a monophyletic group with *Procellaria* (PP: 1.0; BS: 100%), which in turn clustered with *Pterodromia* (PP: 1.0; BS:

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.272v1 | CC-BY 4.0 Open Access | received: 5 Mar 2014, published: 5 Mar 2014

100%). This cluster is sister group to *Hydrobates* (PP: 0.98; BS: 59%). The clade consisting of *Calonectris*, *Puffinus*, *Procellaria*, *Pterodroma* and *Hydrobates* shows sister relationship to the monophylum *Pelagodroma*, *Diomedea* and *Thallasarche* (PP: 1.0; BS: 97%), the latter two genera are very closely related (PP: 1.0; BS: 100%). A further monophyletic group constitutes a clade consisting of Sulidae (genera *Morus* and *Sula*), Anhingidae (genus *Anhinga*), Phalacrocoracidae (genus *Phalacrocorax*) (PP: 1.0; BS: 100%), which is the sister group to Pelecanidae (PP: 0.99; BS: 50%). The last monophyletic assemblage consists of Ardeidae (genera *Egretta*, *Ardea*, *Nycticorax* and *Ixobrychus*) and Threskiornithidae (genera *Platalea*, *Threskiornis* and *Nipponia*) (PP: 1.0; BS: 71%).

Despite strong efforts in the past to identify statistically robust clades from complete mtDNA data sets, many of the interrelationships above the family level in birds are still ambiguous in some studies. One complicating factor for bird phylogenetic reconstruction is that groupings with high metabolic rates tend to show increased substitution rates (Gillooly , McCoy & Allen, 2007), e.g. passerines. As the waterbird assemblage comprises taxa which are of comparatively uniform size and physiology, statistically robust resolution can be expected within this group. Despite this benefit, resolution of the relationships in this group from DNA data is hampered by the occurrence of short internodes at this taxonomic level from analyses of ncDNA (Hackett et al., 2008), which are also among the shortest within the mtDNA topologies. The use of morphology for phylogenetic reconstruction of the Aequornithes clade is restricted by the limited number of informative characters (Livezey & Zusi, 2007; Mayr, 2011). Despite this limitation, our analyses of a large protein-coding mtDNA dataset identified three robust groupings, which delineate monophyletic entities of specialized morphological and ecological adaptations.

The first of these groupings covers three families of the traditional order Procellariiformes (Nunn & Stanley, 1998; Penhallurick & Wink, 2004), which share the characteristic of a

tubular nasal passage (Lequette et al.,1998) and an almost exclusively pelagic feeding life style. In congruence with the topology from the cytochrome *b* gene reported by Nunn & Stanley (1998), the genera *Calonectris*, *Puffinus*, *Procellaria* and *Pterodroma* are sister taxa. Compared to the eight tubenose birds of the present study, previous studies included less number of two (Slack et al., 2006), two (Pacheco et al., 2011), five (Gibb et al., 2013), five (Hackett et al., 2008) and six (Ericsson et al., 2006) taxa of this group in their analyses. Consistently over BI and ML reconstructions, we observed a placement of *Pelagodroma* as the closest related taxon to *Diomedea* and *Thallasarche*, which differed from that found by Nunn & Stanley (1998); they placed *Pelagodroma* basal to a monophylum containing *Diomedea*, *Procellaria*, *Puffinus* and *Calonectris*. We further observed a different placement of *Hydrobates*, which is placed basal to all above mentioned genera of Procellariiformes in Nunn & Stanley (1998), but in Penhallurick & Wink (2004) *Hydrobates* is closely related to *Pelagodroma*. In our mtDNA reconstructions in contrast, we found *Hydrobates* well supported as a sister taxon to the monophylum containing *Calonectris*, *Puffinus*, *Procellaria* and *Pterodroma.* Limited support of our placement of the latter two taxa is given in the mitogenomic analysis of Gibb et al. (2013) which detected a direct sister relationship between *Procellaria* and *Pterodroma*; however, only the expanded taxa number of tubenose birds in the present study, which generated the additional sequences of *Calonectris* and *Puffinus,* led to an yet unprecedented resolution of the internal structure of this clade.

The second monophyletic group shares a gular sac and four webbed toes as common morphological characters, i.e., Pelecanidae, Sulidae, Anhingidae and Phalacrocoracidae. Affinities of the latter three families had earlier been suggested by DNA-DNA hybridization analysis (Sibley & Ahlquist, 1990), cladistic analysis of a large number of morphological characters (Mayr, 2005; Livezey & Zusi, 2007) and by topologies from ncDNA (Ericsson et al., 2006; Hackett et al., 2008) and mtDNA sequence data (Gibb et al., 2013) and from

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.272v1 | CC-BY 4.0 Open Access | received: 5 Mar 2014, published: 5 Mar 2014

analyses of morphological and genetic data combined (Cracraft et al., 2004). The third grouping concerns the families Threskiornithidae and Ardeidae; their affinity was previously identified by ncDNA analyses (Ericson et al., 2006, Hackett et al., 2008). To our knowledge, the present study for the first time indicates a robust direct sister relationship of Threskiornithidae and Ardeidae based on almost complete mtDNA data, with inclusion of ibises and spoonbills of the genera *Nipponia*, *Platelea*, and *Threskiornis* in a common family Threskiornithidae and of herons, egrets and bitterns of the genera *Ixobrychus*, *Ardea*, *Egretta*, and *Nycticorax* in a common family Ardeidae. In contrast, formerly published mitogenomic topologies found Threskiornithidae and Ardeidae clustered together either with the Pelecanidae (Pacheco et al., 2011), or with the Ciconiidae (Gibb et al., 2013). Therefore, a consensus topology from ncDNA data of Hackett et al. (2008) and mtDNA data from Gibb et al. (2013) showed herons and ibises related but ungrouped, placed in a polytomy of a cluster containing additional pelecaniform and ciconiiform birds (Gibb et al., 2013).

The sister relationship of flamingos and grebes (Cracraft et al., 2004; Ericsson et al., 2006; Hackett et al., 2008; Pacheco et al., 2011; Gibb et al., 2013; McCormack et al., 2013) which clustered outside the Aequornithes clade, was also evident in our analysis. The data clearly show that the superficial ressemblance of grebes and loons, which led to the affinity in traditional considerations of relationships between these two taxa (van Tuinen et al., 2001), is due to convergent evolution. The position of Gruiformes including cranes and rails is outside the Aequornithes clade in the BI which is consistent with their placement in the ncDNA topologies of Ericson et al. (2006) and Hackett et al. (2008). Therefore, the inclusion of the Gruiformes within the Aequornithes clade in the ML tree is very likely erroneous, as this position furthermore is not supported by high bootstrap values.

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.272v1 | CC-BY 4.0 Open Access | received: 5 Mar 2014, published: 5 Mar 2014

Conclusions

Our analyses, based on an enlarged dataset of 52 water bird taxa in total and complementary phylogenetic reconstruction methods, have identified three monophyletic clades of the aquatic birds above the family level and provide a robust hypothesis for further evolutionary and ecological studies.

Acknowledgements

The authors thank various colleagues, mentioned in table 1, for the provision of samples for this study.

References

Altekar G, Dwarkadas S, Huelsenbeck JP, Ronquist F. 2004. Parallel metropolis-coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. *Bioinformatics* 20:407–415.

Arctander P. 1988. Comparative studies of avian DNA by restriction fragment length polymorphism analysis: convenient procedures based on blood samples from live birds. *Journal of Ornithology* 129:205–216.

Braun EL & Kimball RT. 2002. Examining basal avian divergences with mitochondrial sequences: model complexity, taxon sampling, and sequence length. *Systematic Biology* 51:614–625.

Cracraft J. 1981. Towards a phylogenetic classification of the recent birds of the world (class: Aves). *Auk* 98:681–714.

Cracraft J, Barker FK, Braun M, Harshman J, Dyke GJ, Feinstein J, Stanley S, Cibois A, Schikler P, Beresford P, Garcia-Moreno J, Sorenson MD, Yuri T, Mindell DP. 2004. Phylogenetic relationships among modern birds (Neornithes). Toward an Avian Tree of Life. In: Cracraft J & Donoghue MJ, ed. *Assembling the tree of life*. New York: Oxford University Press, 468-489.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* 9:772.

Ericson PGP, Anderson CL, Britton T, Elzanowski A, Johansson US, llersjo MK, Ohlson JI, Parsons TJ, Zuccon D, Mayr G. 2006. Diversification of Neoaves: integration of molecular sequence data and fossils. *Biology Letters* 2:543–547.

Gibb GC, Kennedy M & Penny D. 2013. Beyond phylogeny: pelecaniform and ciconiiform birds, and long-term niche stability. *Molecular Phylogenetic Evolution* 68:229‑238.

259 Gillooly JF, McCoy MW & Allen AP. 2007. Effects of metabolic rate on protein evolution.
260 *Biology Letters* 3:655-659.
261 Guindon S & Gascuel O. 2003. A simple, fast and accurate method to estimate large
262 phylogenies by maximum-likelihood. *Systematic Biology* 52:696–704.
263 Hackett SJ, Kimball RT, Reddy S, Bowie RC, Braun EL, Braun MJ, Chojnowski JL, Cox
264 WA, Han KL, Harshman J, Huddleston CJ, Marks BD, Miglia KJ, Moore WS,
265 Sheldon FH, Steadman DW, Witt CC, Yuri T. 2008. A phylogenomic study of birds
266 reveals their evolutionary history. *Science* 320:1763–1768.
267 Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis
268 program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41:95–98.
269 Haring E, Kruckenhauser L, Gamauf A, Riesing MJ, Pinsker W. 2001. The complete
270 sequence of the mitochondrial genome of *Buteo buteo* (Aves, Accipitridae) indicates
271 an early split in the phylogeny of raptors. *Molecular Biological Evolution* 18:1892–
272 904.
273 Jetz W, Thomas GH, Joy JB, Hartmann K, Mooers AO. 2012. The global diversity of birds in
274 space and time. *Nature* 491:444–448.
275 Lequette B, Verheyden C & Jowentin P. 1989. Olfaction in subantarctic seabirds: its
276 phylogenetic and ecological significance. *Condor* 91:732–735.
277 Livezey BC & Zusi RL. 2007. Higher-order phylogeny of modern birds (Theropoda, Aves:
278 Neornithes) based on comparative anatomy. II. Analysis and discussion. *Zoological*
279 *Journal of Linnean Society* 149:1–95.
280 Mayr G. 2005. Tertiary plotopterids (Aves, Plotopteridae) and a novel hypothesis on the
281 phylogenetic relationships of penguins (Spheniscidae). *Journal of Zoological*
282 *Systematics and Evolutionary Research* 43:61–71.

283 Mayr G. 2011. Metaves, Mirandornithes, Strisores and other novelties – a critical review of
284 the higher-level phylogeny of neornithine birds. *Journal of Zoological Systematics and*
285 *Evolutionary Research* 49:58–76.
286 Miller MA, Pfeiffer W & Schwartz T. 2010. Creating the CIPRES Science Gateway for
287 inference of large phylogenetic trees. *Proceedings of the Gateway Computing*
288 *Environments Workshop (GCE)*, 14 Nov. 2010, New Orleans, LA pp 1–8.
289 Mindell DP, Sorenson MD, Huddleston CJ, Miranda HC, Knoght A, Sawchuk SJ, Yuri T.
290 1997. Phylogenetic relationships among and within select avian orders based on
291 mitochondrial DNA. In: Mindell DP, ed. *Avian molecular evolution and systematics*.
292 San Diego: Academic Press, 213-247.
293 McCormack JE, Harvey MG, Faircloth BC, Crawford NG, Glenn TC, Brumfield RT. 2013. A
294 phylogeny of birds based on over 1,500 loci collected by target enrichment and high-
295 throughput sequencing. *PLoS ONE* 8(1):e54848.
296 Nunn GB & Stanley SE. 1998. Body size effects and rates of cytochrome b evolution in tube-
297 nosed seabirds. *Molecular Biology and Evolution* 15:1360-1371.
298 Pacheco MA, Battistuzzi FU, Lentino M, Aguilar RF, Kumar S, Escalante AA. 2011.
299 Evolution of modern birds revealed by mitogenomics: timing the radiation and origin
300 of major orders. *Molecular Biology and Evolution* 28:1927–42.
301 Penhallurick J & Wink M. 2004. Analysis of the taxonomy and nomenclature of the
302 Procellariiformes based on complete nucleotide sequences of the mitochondrial
303 cytochrome b. *Emu* 104:124–147.
304 Pratt RC, Gibb GC, Morgan-Richards M, Phillips MJ, Hendy MD, Penny D. 2009. Toward
305 resolving deep Neoaves phylogeny: data, signal enhancement, and priors. *Molecular*
306 *Biology and Evolution* 26:313–326.

307 Rambaut A & Drummond AJ. 2007. Tracer v1.4. Available at http://beast.bio.ed.ac.uk/Tracer
308 (accessed 20 November 2013).
309 Ronquist F & Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under
310 mixed models. *Bioinformatics* 19:1572–1574.
311 Sambrook J, Fritsch EF & Maniatis T. 1989. *Molecular cloning: a laboratory manual*. New
312 York: Cold Spring Harbor Laboratory Press.
313 Sibley CG & Ahlquist JE. 1990. *Phylogeny and classification of birds*. New Haven &
314 London: Yale University Press.
315 Slack KE, Jones CM, Ando T, Harrison GL, Fordyce RE, Arnason U, Penny D. 2006. Early
316 penguin fossils, plus mitochondrial genomes, calibrate avian evolution. *Molecular*
317 *Biology and Evolution* 23:1144–1155.
318 Smith ND. 2010. Phylogenetic analysis of Pelecaniformes (Aves) based on osteological data:
319 implications for waterbird phylogeny and fossil calibration studies. *PLoS ONE* 5(10):
320 e13354.
321 Sorenson MD, Ast JC, Dimcheff DE, Yuri T, Mindell D. 1999. Primers for a PCR-based
322 approach to mitochondrial genome sequencing in birds and other vertebrates.
323 *Molecular Phylogenetics and Evolution* 12:105–114.
324 Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses
325 with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690.
326 van Tuinen M, Butvill DB, Kirsch JAW, Hedges SB. 2001. Convergence and divergence in
327 the evolution of aquatic birds. *Proceedings of the Royal Society of London B*
328 268:1345–1350.
329 Watanabe M, Nikaido M, Tsuda TT, Kobayashi T, Mindell D, Cao Y, Okada N, Hasegawa
330 M. 2006. New candidate species most closely related to penguins. *Gene* 378:65-73.

Wink M. 1995. Phylogeny of old and new world vultures (Aves: Accipitridae and Cathartidae) inferred from nucleotide sequences of the mitochondrial cytochrome b gene. *Zeitschrift für Naturforschung* 50c:868–882.

Wink M. 2011. Evolution und Phylogenie der Vögel - Taxonomische Konsequenzen. *Vogelwarte* 49:17–24.

Legends of the tables

Table 1 Origin, collector information and voucher numbers of samples analyzed in this study

Table 2 GenBank accession numbers of taxa sequenced in the present study

Table 3 Primers used for PCR amplifications

Table 4 Primers used for sequencing

Legends of the figures

Fig. 1 Bayesian inference tree of waterbirds in relation to other bird orders based on 6699 nucleotides of protein coding mtDNA genes, reconstructed under the GTR + Γ + I nucleotide substitution model and 16,000,000 generations. Posterior probabiliy values are indicated for each node.

Fig. 2 Maximum likelihood tree reconstructed under the GTRGAMMA nucleotide substitution model (ln L = –207933.05). Bootstrap support (1000 replicates) is indicated for each node.

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.272v1 | CC-BY 4.0 Open Access | received: 5 Mar 2014, published: 5 Mar 2014

381 **Table 1**

Taxon	Voucher number[a]	Sample origin, collector
Aechmophorus occidentalis	3768	Mexico, H. Witt
Gavia immer	3770	Mexico, H. Witt
Gavia pacifica	3774	Mexico, H. Witt
Spheniscus humboldti	477	Zoo Leipzig, Germany, G. Ehlers
Spheniscus demersus	8128	Zoo Leipzig, Germany, G. Ehlers
Puffinus yelkouan	3163	Naxos, Greece, D. Ristow
Calonectris diomedea	15872	Crete, Greece, D. Ristow
Hydrobates pelagicus	4073	Malta, Italy, J. Borg
Morus bassanus	3288	Saltees, Ireland, M. Wink
Pelecanus occidentalis	3747	Mexico, H. Witt
Phoenicopterus ruber	6927	Namibia, J. Osborne
Ardea cinerea	2822	Bünder Natur-Museum, Switzerland
Ciconia ciconia	468	Zoo Leipzig, Germany, G. Ehlers
Ciconia nigra	3848	Spain, U. Höfle

382

383 [a] Voucher numbers of the Institute of Pharmacy and Molecular Biotechnology (IPMB),

384 University Heidelberg, Germany

385 **Table 2**

Taxon	*COX1*	*COX2*	*ATP8*	*ATP6*	*COX3*	*ND3*	*ND4L*	*ND4*	*CYTB*
Aechmophorus occidentalis	AY567891	AY567905	AY567863	AY567849	AY567877	AY567933	AY567961	AY567947	AY567919
Gavia immer	AY567890	AY567904	AY567862	AY567848	AY567876	AY567932	AY567960	AY567946	AY567918
Gavia pacifica	AY567889	AY567903	AY567861	AY567847	AY567875	AY567931	AY567959	AY567945	AY567917
Spheniscus humboldti	AY567888	AY567902	AY567860	AY567846	AY567874	AY567930	AY567958	AY567944	AY567916
Spheniscus demersus	AY567887	AY567901	AY567859	AY567845	AY567873	AY567929	AY567957	AY567943	AY567915
Puffinus yelkouan	AY567884	AY567898	AY567856	AY567842	AY567870	AY567926	AY567954	AY567940	AY567912
Calonectris diomedea	AY567883	AY567897	AY567855	AY567841	AY567869	AY567925	AY567953	AY567939	AY567911
Hydrobates pelagicus	AY567885	AY567899	AY567857	AY567843	AY567871	AY567927	AY567955	AY567941	AY567913
Morus bassanus	AY567893	AY567907	AY567865	AY567851	AY567879	AY567935	AY567963	AY567949	AY567921
Pelecanus occidentalis	AY567892	AY567906	AY567864	AY567850	AY567878	AY567934	AY567962	AY567948	AY567920
Phoenicopterus ruber	AY567894	AY567908	AY567866	AY567852	AY567880	AY567936	AY567964	AY567950	AY567922
Ardea cinerea	AY567886	AY567900	AY567858	AY567844	AY567872	AY567928	AY567956	AY567942	AY567914
Ciconia ciconia	AY567881	AY567895	AY567853	AY567839	AY567867	AY567923	AY567951	AY567937	AY567909
Ciconia nigra	AY567882	AY567896	AY567854	AY567840	AY567868	AY567924	AY567952	AY567938	AY567910

386

387

388

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.272v1 | CC-BY 4.0 Open Access | received: 5 Mar 2014, published: 5 Mar 2014

389 **Table 3**

Gene	Primer pair	Sequence (5′ to 3′)	Reference
CYTB	L14995	CTCCCAGCCCCATCCAACATCTCAGCATGATGAAACTTCG	Wink (1995)
	H16065	CTAAGAAGGGTGGAGTCTTCAGTTTTTGGTTTACAAGAC	
COX1	L6615	CCTCTGTAAAAAGGACTACAGCC	Sorenson et al. (1999)
COX2	H9233	AAGAAGCTTAGGTTCATGGTCAGG	
ATP8			
ATP6	L9034	CAGCACTAGCCTTTTAAGCTA	Sorenson et al. (1999)
COX3	H12976	CAGATGCAGGAATTAGCAGTTCTTG	
ND3	nd4-3+	ACCAACTACGAGCGGACACACAG	Haring et al. (2001)
ND4L	nd5-2-	ATGATTCCCACTCCTTCTCAGCC	
ND4	L9034	CAGCACTAGCCTTTTAAGCTA	Sorenson et al. (1999)
	nd4-4-	GCTTTCTAGGCATAGTAGGGC	Haring et al. (2001)

390

391

392

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.272v1 | CC-BY 4.0 Open Access | received: 5 Mar 2014, published: 5 Mar 2014

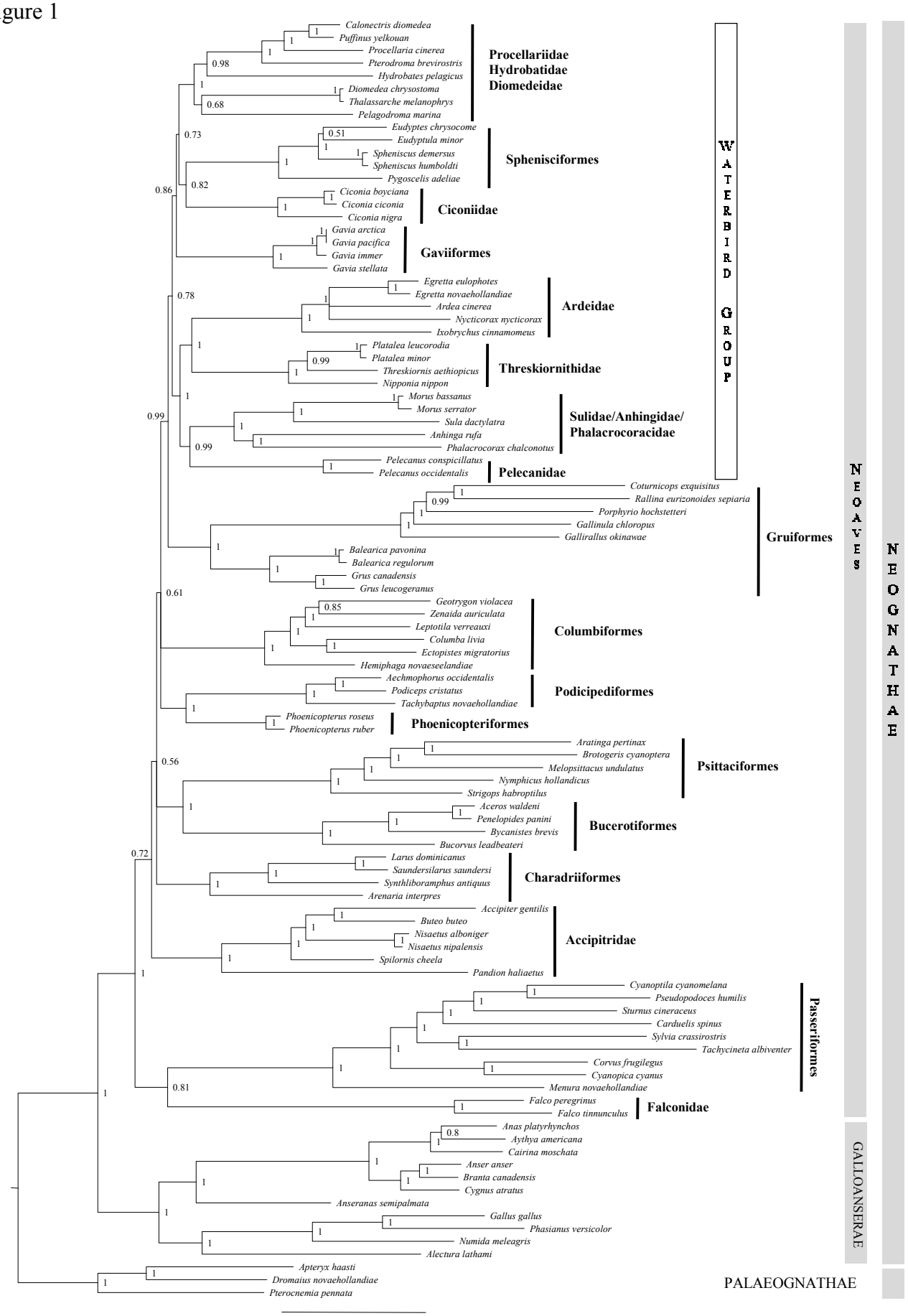
393 **Table 4**

Gene	Nucleotides	Primer	Sequence (5′ to 3′)	Reference
COX1	1530	L6615	CCTCTGTAAAAAGGACTACAGCC	Sorenson et al. (1999)
COX2	675	L6615+1	GC(CT)A(CT)(CT)AACATAAAACCCCCAGCC	Present study
ATP8	150	L6615-1N	CAAC(CT)GC(CT)ATCAACATAAA	Present study
		L6615+2	ATAGA(CT)GTAGA(CT)AC(CT)CGAGC	Present study
		L6615+3	GACGATACTCCGACTA(CT)CCAGA	Present study
		L6615+4N	TCTCCTATCATAGAAGA	Present study
		L6615-4	TCCCC(CT)AT(CT)ATAGAAGA	Present study
		H9233	AAGAAGCTTAGGTTCATGGTCAGG	Sorenson et al. (1999)
ATP6	666	L9034	CAGCACTAGCCTTTTAAGCTA	Sorenson et al. (1999)
COX3	780	L9034-1	CAACTATC(AGC)AT(AG)AACAT	Present study
ND3	342	L9034-1n	CTAGGCCTACTACCATACACAT	Present study
ND4L	285	L9034+1N	CAA(CT)TATCAATAAACATA	Present study
ND4	1359	L9034+2	CAATGATGACGAGACAT(CT)GTACG	Present study
		L9034+2N	CAATGATGACGTGATATTGTACG	Present study
		L9034+3	TTCCA(CT)GGACT(GCT)CACGTAAT	Present study
		L9034+3N	GTATA(CT)GGCTCAACCTTCTT	Present study
		L9034+4	TTCCTCAGTAGCAAT(CT)(CT)TATTCCT	Present study
		L9034+5	ACTCATACT(ACT)ACATTCTC	Present study
		L9034+5D	CCAAACAACAACATTCACT	Present study
		L9034+5N5	ATCGAAAACCAAGCAGCA	Present study
		L9034+6	AGCTCCCTACC(CT)TTACT	Present study
		L9034+6N	TACCTACTATTCTA(CT)ACCCT	Present study
		Nnd4-3+	ACAAACTA(CT)GAACG(AC)ACACACAG	Present study
		nd4-3+	ACCAACTACGAGCGGACACACAG	Haring et al. (2001)
CYTB	912	mt-A(L14995)	CTCCCAGCCCCATCCAACATCTCAGCATGAT GAAACTTCG	Wink (1995)
		mt-C (L15320)	TA(CT)GTCCTACCATGAGGACAAATATCATT CTGAGG	Wink (1995)

394

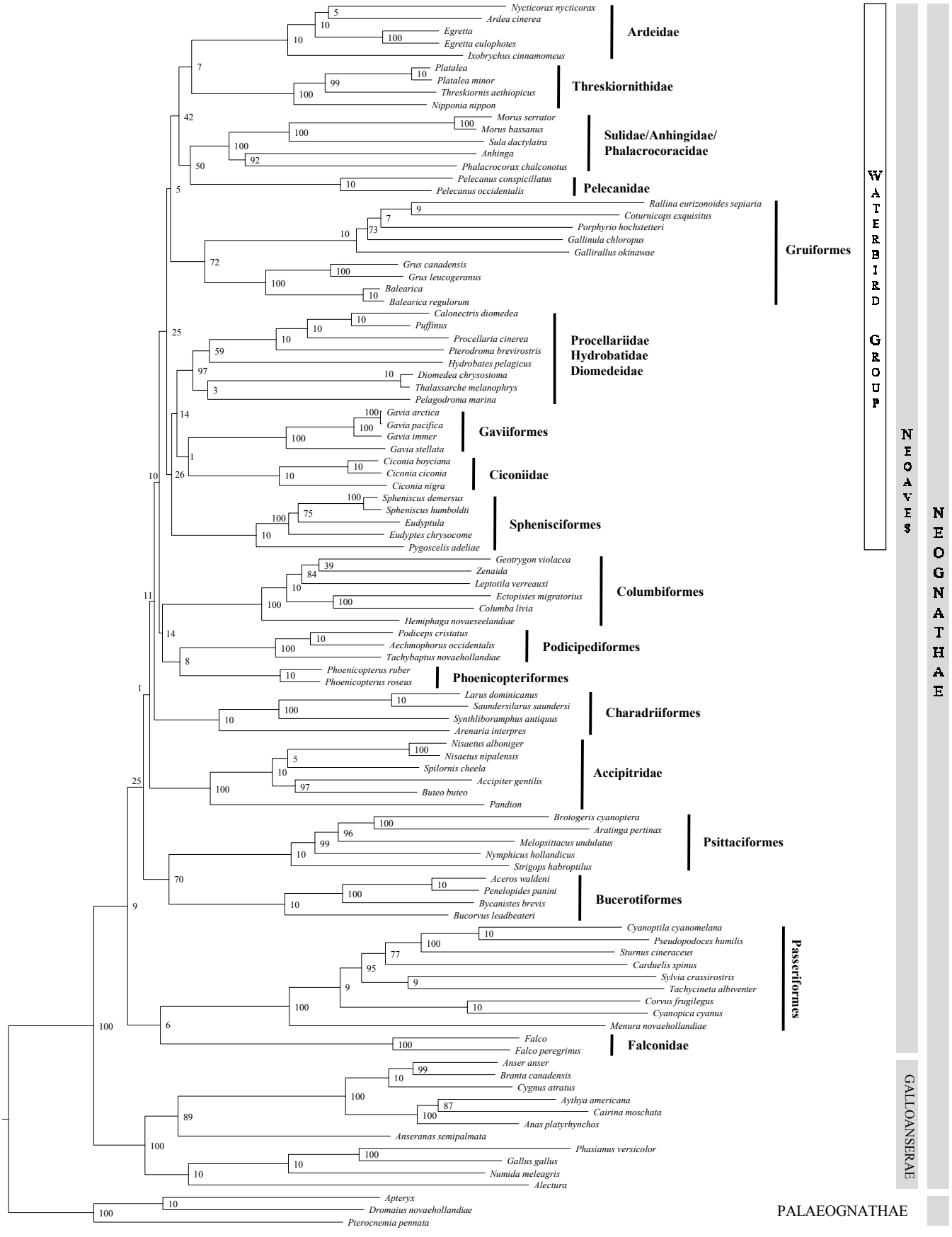
PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.272v1 | CC-BY 4.0 Open Access | received: 5 Mar 2014, published: 5 Mar 2014

Figure 1



PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.272v1 | CC-BY 4.0 Open Access | received: 5 Mar 2014, published: 5 Mar 2014

Figure 2



PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.272v1 | CC-BY 4.0 Open Access | received: 5 Mar 2014, published: 5 Mar 2014