Genetics and Molecular Research

# DNA barcoding and phylogenetic relationships of Ardeidae (Aves: Ciconiiformes)

**Z.H. Huang, M.F. Li and J.W. Qin**

School of Life Sciences, Jinggangshan University, Ji'an, Jiangxi Province, China

Corresponding author: Z.H. Huang
E-mail: hzhow@163.com

Genet. Mol. Res. 15 (3): gmr.15038270
Received December 8, 2015
Accepted January 29, 2016
Published August 18, 2016
DOI http://dx.doi.org/10.4238/gmr.15038270



**ABSTRACT.** The avian family Ardeidae comprises long-legged freshwater and coastal birds. There has been considerable disagreement concerning the intrafamilial relationships of Ardeidae. Mitochondrial cytochrome c oxidase subunit I (*COI*) was used as a marker for the identification and phylogenetic analysis of avian species. In the present study, we analyzed the *COI* barcodes of 32 species from 17 genera belonging to the family Ardeidae. Each bird species possessed a barcode distinct from that of other bird species except for *Egretta thula* and *E. garzetta*, which shared one barcoding sequence. Kimura two-parameter distances were calculated between barcodes. The average genetic distance between species was 34-fold higher than the average genetic distance within species. Neighbor-joining and maximum likelihood methods were used to construct phylogenetic trees. Most species could be discriminated by their distinct clades in the phylogenetic tree. Both methods of phylogenetic reconstruction suggested that *Zebrilus*, *Tigrisoma*, and *Cochlearius* were an offshoot of the primitive herons. *COI* gene analysis suggested that the other herons could be divided

into two clades: Botaurinae and Ardeinae. Our results support the Great Egret and Intermediate Egret being in separate genera, *Casmerodius* and *Mesophoyx*, respectively.

**Key words:** DNA barcoding; Cytochrome c oxidase I; Ardeidae; Phylogeny

## INTRODUCTION

The avian family Ardeidae, generally known as herons, comprises long-legged freshwater and coastal birds, which exhibit very little sexual dimorphism in their size; however, the systematics of herons have long been contended (Peters, 1931; Bock, 1956; Payne and Risley, 1976; Sheldon, 1987, 2000; Sibley and Monroe, 1990). Bock (1956) conducted a systematic review of the group and defined 64 species, 15 genera, and 2 subfamilies (Ardeinae and Botaurinae). Payne and Risley's (1976) comparison of 33 osteological characters of herons was the first cladistic estimate of heron phylogeny. Based on their findings, herons were divided into two major clades: 1) boat-billed heron (*Cochlearius cochlearius*), night-herons, and bitterns; 2) tiger-herons and day-herons. From skeletal analyses, which focused more on the bones of the body and limbs, this grouping was revealed to be incorrect (McCracken and Sheldon, 1998).

Molecular phylogeny challenges the traditional classification of the Ardeidae (Sheldon, 1987; Sibley and Ahlquist, 1990; Sheldon et al., 1995). Sibley and Ahlquist (1990) first attempted to clarify the phylogeny of the Ardeidae using DNA-DNA hybridization data, and added three genera, which were *Mesophoyx*, *Casmerodius*, *Nyctanassa*. Sheldon et al. (1995) studied the systematic relationships of herons using mitochondrial DNA sequences, and distinguished three major groups: tiger herons and the boatbill (Tigriornithinae), bitterns (Botaurinae), day-herons, egrets, and night-herons (Ardeinae). The classification of individual heron species was still fraught with difficulty, and there was no clear consensus about the correct placement of many species into any of the three major groups. Thus, relationships between genera in this family were not completely resolved. Sheldon et al. (1995) noted that the uncertainties remaining in higher-level heron phylogeny were: 1) the position and composition of some genera, and 2) the identification of the basal heron lineage.

DNA barcoding employs sequences from a short standardized gene region to identify species (Hebert et al., 2003a,b), and has been used successfully to analyze the phylogeny of many animal groups (Huang and Ke, 2015). DNA barcoding studies on Ardeidae birds remain very limited (Päckert et al., 2014). Therefore, in the present study, we examined 652 bp of the mitochondrial cytochrome c oxidase subunit I (*COI*) gene of Ardeidae birds and then conducted phylogenetic analyses within Ardeidae based on these sequences. Our main aims were to clarify the phylogeny of the herons and to show that the *COI* gene is an efficient marker for their identification.

## MATERIAL AND METHODS

Ninety-eight *COI* sequences were obtained from GenBank. Thirty-two species from 17 genera belonging to the family Ardeidae were analyzed (**Table S1**).

Sequences were aligned by the Clustal X procedure (Thompson et al., 1997). A 652-bp region of the mtDNA *COI* gene was analyzed. DnaSP v5.0 (Librado and Rozas, 2009) was used to define the variable sites. Sequence divergence among species and genera was calculated using the Kimura two-parameter (K2P, Kimura, 1980) distance model in MEGA 6.0 (Tamura et al., 2013). The Neighbor-Joining method (NJ, Saitou and Nei, 1987) was used to reconstruct the phylogenetic tree based on K2P from MEGA6.0 (Tamura et al., 2013). Modeltest 3.0 (Posada and Crandall, 1998) and the Akaike information criterion (AIC, Posada and Buckley, 2004) were used to identify the appropriate nucleotide substitution models. A maximum likelihood tree (ML, Strimmer and Haeseler, 1996) was obtained using heuristic searches, based on the substitution model proposed by Modeltest 3.0 (Posada and Crandall, 1998). We calculated the ML tree using PAUP4.0 (Swofford, 2002). Statistical support for the internodes in phylogenetic tree was tested by bootstrap percentages with 1000 replicates (Felsenstein, 1985).

## RESULTS

### Barcoding analysis

The 652-bp *COI* sequences were aligned, and the gene fragment was found to correspond to the *Egretta garzetta* mitochondrial gene, which started at position 5476 and ending at position 6127 (GenBank accession No. NC023981, Zou et al., 2015). A total of 251 variable sites were identified, of which 230 were parsimoniously informative (35.27% of the entire sequence). Most species had distinct *COI* sequences, with the exception of *Egretta thula* and *E. garzetta,* which shared one barcoding sequence. The average nucleotide composition was 25.65% T, 31.08% C, 26.91% A, and 16.36% G.

K2P genetic distances within-species had a small range (0 to 3.35%), with more than 87.50% of the observations having a genetic distance less than 1% (Figure 1). Pair-wise comparisons among-species were distributed from 0.83% (between *Egretta novaehollandiae* and *Egretta sacra*) to 17.52% (between *Ixobrychus cinnamomeus* and *Tigrisoma fasciatum*) with most of the comparisons observed between K2P genetic distances of 10 and 18%, with values up to 92.23%. The average difference in the *COI* sequence between species (13.08%) was 34-fold higher compared to the average difference within species (0.40%).

### Phylogenetic relationships

On the basis of hierarchical likelihood-ratio tests as implemented in Modeltest 3.0, the model General Time Reversible (GTR) model + Gamma distribution + invariable sites was used (GTR + G + I, -lnL = 5439.66, $P < 0.001$, AIC = 11,286.63, BIC = 13,125.53). The gamma distribution and proportion of invariant sites were set as 1.46 and 0.57 (estimated by Modeltest), respectively.

The NJ method was also used to reconstruct the phylogenetic trees based on the K2P model. Meanwhile, the maximum likelihood phylogenetic tree was estimated with the best-fit model GTR + G + I. The phylogenetic trees obtained by NJ (**Figure S1**) and maximum likelihood (not shown) were very similar. All species could be discriminated by their distinct clades in the phylogenetic tree except for *E. thula* and *E. garzetta* (**Figure S1**). *Zebrilus* species were the first to split from the Ardeidae lineage. The next clades included members

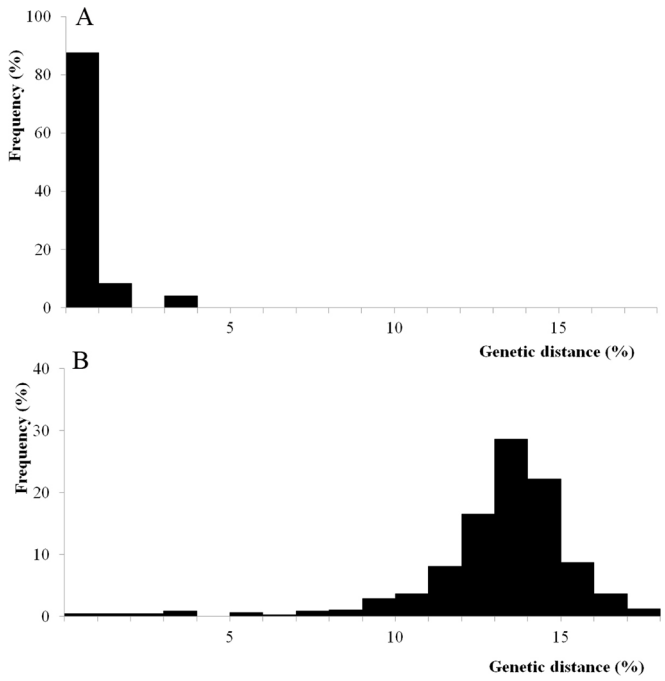


**Figure 1.** Frequency distribution of genetic distances within (**A**) and among (**B**) species of Ardeidae.

of *Tigrisoma* and *Cochlearius*. Both phylogenetic analyses suggested that *Pilherodius* was the sister genus to *Syrigma*. *Agamia* was the sibling taxon to *Ardeola*. Analysis of *COI* genes supported the inclusion of *Nycticorax*, *Gorsachius*, and *Butorides* in a clade (**Figure S1**).

The genus *Ixobrychus* was divided into two subclades: *involucris* (*sinensis* + *miutus* + *flavicollis*, *cinnamomeus* + *eurhythmus*). Species from *Botaurus* were the sister group to *Ixobrychus involucris*. The genus *Ardea* was also divided into two subclades: *purpurea*, *cinerea* + *herodias* + *cocoa*. Members of *Egretta* grouped into one clade: *thula* + *garzetta*, *sacra* + *eulophotes*, *novaehollandiae*. However, *Casmerodius alba* grouped with *Bubulcus ibis* and *Mesophoyx intermedia* into one clade (**Figure S1**).

## DISCUSSION

The Ardeidae are a well-studied group of birds; however, there has been considerable disagreement concerning the intrafamilial relationships of these birds (Sheldon, 1987). The results of the present study provide the first DNA barcoding and phylogenetic analysis of the family using the *COI* gene. These results demonstrate the discriminatory power of barcodes for the identification of herons. *COI* data showed a well-supported phylogeny for Ardeidae species.

The zigzag heron (*Zebrilus undulatus*) was the most enigmatic of the herons (Payne and Risley, 1976). The zigzag heron resembles a tiger-heron in its barred plumage and forest habitat (Ridgway, 1878; Bock, 1956). However, Sharpe (1898) and Peters (1931) considered the zigzag heron to be a bittern. Payne and Risley (1976) also found that the zigzag heron

was associated with the bitterns rather than with the tiger herons based on skeletal characters. Molecular data have also suggested that the zigzag heron is most closely related to bitterns (Sheldon et al., 1995; McCracken and Sheldon, 1998). Our results supported that the zigzag heron was sister species of the outgroup (Figure 1). It is likely that the heron is a basal ancestor of herons rather than a tiger-heron or bittern. Some of its characteristics resemble those of the bitterns, while others resemble those of tiger herons (Payne and Risley, 1976). *Zebrilus* may be best considered a highly aberrant heron, which has retained several primitive characteristics of the heron ancestors. More taxon sampling and different markers are needed to resolve the taxonomic status of the zigzag heron.

*COI* analysis supported the next clades being formed by the genera *Tigrisoma* and *Cochlearius*. The taxonomic position of the boat-billed heron (*Cochlearius cochlearius*) has been much disputed. Members of *Cochlearius* were previously considered to be night-herons based on morphological and behavioral characters (Payne and Risley, 1976). It has also been proposed that the boat-billed heron is a separate group (Bock, 1956). Sheldon et al. (1995) considered *Cochlearius* to be the sister taxon of *Tigrisoma* based on DNA-DNA hybridization. Analysis of *COI* genes suggested the *Cochlearius* was related to *Tigrisoma*, but supported these herons to be in a monophyletic group. Within Ardeidae, with the exception of the three aforementioned genera, analysis of *COI* genes supported the conventional division between Botaurinae and Ardeinae.

Within Botaurinae, the present analysis succeeded in producing a completely dichotomy phylogenetic tree. Members of *Botaurus* were the sibling group to *I. involucris*. The other species of *Ixobrychus* were divided into two subclades. *Flavicollis* was the sister species of *miutus/sinensis,* which together formed the sister group of *eurhythmus/cinnamomeus*. The species *Ixobrychus flavicollis* has often been recognized as a monotypic genus *Dupetor* (Peters, 1931). However, *flavicollis* was somewhat aberrant in size and there were no obvious morphological differences between it and other species of small bitterns; therefore, it appeared to be best recognized as *Ixobrychus* (Payne and Risley, 1976). *COI* gene data well supported the placement of *flavicollis* in the genus *Ixobrychus*.

Herons have traditionally been divided into four or five ecological groups: day herons, night herons, bitterns, tiger herons, and boat-billed herons (Hancock and Kushlan, 1984). *COI* gene data suggested that day- and night-herons form a clade with bitterns as their sister group. Some studies also found day- and night-herons to be monophyletic (Sheldon et al., 1995). Species have been moved between different genera. Within Ardeinae, the taxonomic position of the great egret (*C. alba*) and the intermediate egret (*M. intermedia*) has been disputed. Traditionally, the two egrets were placed in the genus *Egretta* based on morphology (Bock, 1956). Payne and Risley (1976) proposed that the great egret should be included in the genus *Ardea* rather thank *Egretta* based on its skeletal characteristics (Sheldon, 1987). Analysis of the *COI* gene clearly demonstrated that the great egret and the intermediate egret are closer to the cattle egret (*B. ibis*) than to the little egret (*E. garzetta*) (Figure 1). Our results support Sibley and Monroe's (1990) classification of the great egret and the intermediate egret in separate genera, *Casmerodius* and *Mesophoyx*, respectively.

*Egretta* and *Ardea* have been associated because they share many similar characters and because the great egret and the intermediate egret appeared to bridge the gap between them (Sheldon 1987). The phonetic data of Payne and Risley (1976) and the DNA data of Sheldon (1987) also revealed a significant difference between *Egretta* and *Ardea*. Morphological

comparisons showed that *Syrigma* was the sister group of day- and tiger herons and that *Bubulcus* was closer to *Egretta* than to *Ardea* and *Casmerodius* (Payne and Risley, 1976). *COI* gene data support a different phylogeny to that supported by the morphological characters. The molecular data demonstrated: 1) *Zebrilus*, *Tigrisoma*, and *Cochlearius* branched from the base of the heron tree, 2) day- and night-herons formed a clade, 3) *Syrigma* belonged in the egret clade, and *Bubulcus* was part of the *Casmerodius/Mesophoyx/Ardea* clade. Our results indicate that *Egretta* and *Ardea* are not necessarily sister taxa.

## Conflicts of interest

The authors declare no conflict of interest.

## ACKNOWLEDGMENTS

Research supported by the National Natural Science Foundation of China (#31260088, #31560590), the Jiangxi Province Talent Project "555", the Jiangxi Province Major Disciplines Academic Leaders (#20133BCB22010), the Natural Science Foundation of Jiangxi Province (#20132BAB204022 and #20152ACB21006), and the Science and Technology Foundation of Jiangxi Provincial Department of Education (#GJJ150768).

## REFERENCES

Aliabadian M, Beentjes KK, Roselaar CS, van Brandwijk H, et al. (2013). DNA barcoding of Dutch birds. *ZooKeys* 365: 25-48. http://dx.doi.org/10.3897/zookeys.365.6287

Bock WJ (1956). A generic review of the family Ardeidae (Aves). *Amer Mus Novitates* 1779.

Cai Y, Yue B, Jiang W, Xie S, et al. (2010). DNA barcoding on subsets of three families in Aves. *Mitochondrial DNA* 21: 132-137. http://dx.doi.org/10.3109/19401736.2010.494726

Felsenstein J (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783-791. http://dx.doi.org/10.2307/2408678

Gibb GC, Kardailsky O, Kimball RT, Braun EL, et al. (2007). Mitochondrial genomes and avian phylogeny: complex characters and resolvability without explosive radiations. *Mol. Biol. Evol.* 24: 269-280. http://dx.doi.org/10.1093/molbev/msl158

Hancock J and Kushlan J (1984) The Herons handbook. Harper and Row, New York.

Hebert PDN, Ratnasingham S and deWaard JR (2003a). Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc. Biol. Sci.* 270 (Suppl 1): S96-S99. http://dx.doi.org/10.1098/rsbl.2003.0025

Hebert PDN, Cywinska A, Ball SL and deWaard JR (2003b). Biological identifications through DNA barcodes. *Proc. Biol. Sci.* 270: 313-321. http://dx.doi.org/10.1098/rspb.2002.2218

Huang ZH and Ke DH (2015). DNA barcoding and phylogenetic relationships in Timaliidae. *Genet. Mol. Res.* 14: 5943-5949. http://dx.doi.org/10.4238/2015.June.1.11

Kerr KC, Stoeckle MY, Dove CJ, Weigt LA, et al. (2007). Comprehensive DNA barcode coverage of North American birds. *Mol. Ecol. Notes* 7: 535-543. http://dx.doi.org/10.1111/j.1471-8286.2007.01670.x

Kerr KC, Birks SM, Kalyakin MV, Red'kin YA, et al. (2009a). Filling the gap - COI barcode resolution in eastern Palearctic birds. *Front. Zool.* 6: 29-42. http://dx.doi.org/10.1186/1742-9994-6-29

Kerr KC, Lijtmaer DA, Barreira AS, Hebert PD, et al. (2009b). Probing evolutionary patterns in neotropical birds through DNA barcodes. *PLoS One* 4: e4379. http://dx.doi.org/10.1371/journal.pone.0004379

Kimura M (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16: 111-120. http://dx.doi.org/10.1007/BF01731581

Librado P and Rozas J (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451-1452. http://dx.doi.org/10.1093/bioinformatics/btp187

McCracken KG and Sheldon FH (1998). Molecular and osteological heron phylogenies: sources of incongruence. *Auk*

115: 127-141. http://dx.doi.org/10.2307/4089118

Päckert M, Hering J, Fuchs E, Barthel P, et al. (2014). Genetic barcoding confirms first breeding record of the Yellow Bittern, *Ixobrychus sinensis*, (Aves: Pelecaniformes, Ardeidae) in the Western Palearctic. *Vertebr. Zool.* 64: 251-260.

Payne RB and Risley CJ (1976). Systematics and evolutionary relationships among the herons (Ardeidae). *Misc. Publ. Univ. Mich. Mus. Zool.* 150: 1-115.

Peters JL (1931). Check-list of Birds of the World. vol. 1. Cambridge, Massachusetts: Harvard University Press.

Posada D and Crandall KA (1998). MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14: 817-818. http://dx.doi.org/10.1093/bioinformatics/14.9.817

Posada D and Buckley TR (2004). Model selection and model averaging in phylogenetics: advantages of akaike information criterion and bayesian approaches over likelihood ratio tests. *Syst. Biol.* 53: 793-808. http://dx.doi.org/10.1080/10635150490522304

Ridgway R (1878). Studies of the American Herodiones. Part I. *Bull. U. S. Geol. Geogr. Surv. Territories* 4: 226-237.

Saitoh T, Sugita N, Someya S, Iwami Y, et al. (2015). DNA barcoding reveals 24 distinct lineages as cryptic bird species candidates in and around the Japanese Archipelago. *Mol. Ecol. Resour.* 15: 177-186. http://dx.doi.org/10.1111/1755-0998.12282

Saitou N and Nei M (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4: 406-425.

Schindel DE, Stoeckle MY, Milensky C, Trizna M, et al. (2011). Project description: DNA barcodes of bird species in the national museum of natural history, smithsonian institution, USA. *ZooKeys* 152: 87-92. http://dx.doi.org/10.3897/zookeys.152.2473

Sharpe RB (1898). Catalogue of the birds in the British Museum. vol. 26. London: British Museum

Sheldon FH (1987). Rates of single-copy DNA evolution in herons. *Mol. Biol. Evol.* 4: 56-69.

Sheldon FH, McCracken KG and Stuebing KD (1995). Phylogenetic relationships of the zigzag heron (*Zebrilus undulatus*) and white-crested bittern (*Tigriornis leucolophus*) estimated by DNA-DNA hybridization. *Auk* 112: 672-679.

Shen YY, Liang L, Sun YB, Yue BS, et al. (2010). A mitogenomic perspective on the ancient, rapid radiation in the Galliformes with an emphasis on the Phasianidae. *BMC Evol. Biol.* 10: 132. http://dx.doi.org/10.1186/1471-2148-10-132

Sibley CG and Ahlquist JE (1990). Phylogeny and Classification of Birds. New Haven, Conn. Yale University Press.

Sibley CG and Monroe BL (1990). Distribution and Taxonomy of Birds of the World. New Heaven and London: Yale University Press.

Strimmer K and Haeseler AV (1996). Quartet puzzling: a quartet maximum likelihood method for reconstructing tree topologies. *Mol. Biol. Evol.* 13: 964-969. http://dx.doi.org/10.1093/oxfordjournals.molbev.a025664

Swofford D (2002). PAUP*: Phylogenetic Analysis using Parsimony and other methods. Sunderland, MA: Sinauer Associates

Tamura K, Stecher G, Peterson D, Filipski A, et al. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* 30: 2725-2729. http://dx.doi.org/10.1093/molbev/mst197

Tavares ES, Gonçalves P, Miyaki CY and Baker AJ (2011). DNA barcode detects high genetic structure within neotropical bird species. *PLoS One* 6: e28543. http://dx.doi.org/10.1371/journal.pone.0028543

Thiemann TC, Brault AC, Ernest HB and Reisen WK (2012). Development of a high-throughput microsphere-based molecular assay to identify 15 common bloodmeal hosts of Culex mosquitoes. *Mol. Ecol. Resour.* 12: 238-246. http://dx.doi.org/10.1111/j.1755-0998.2011.03093.x

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, et al. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25: 4876-4882. http://dx.doi.org/10.1093/nar/25.24.4876

Wang Y, Zhou S, Qian L, Pan T, et al. (2016). The mitochondrial genome of *Ixobrychus flavicollis* (Ardeidae: Ixobrychus). *Mitochondrial DNA A DNA MappSeq Anal* 27: 921-922. http://dx.doi.org/10.3109/19401736.2014.926486

Zhang L, Wang L, Gowda V, Wang M, et al. (2012). The mitochondrial genome of the Cinnamon Bittern, Ixobrychus cinnamomeus (Pelecaniformes: Ardeidae): sequence, structure and phylogenetic analysis. *Mol. Biol. Rep.* 39: 8315-8326. http://dx.doi.org/10.1007/s11033-012-1681-1

Zhou X, Wang Y, Chen X, Lin Q, et al. (2008). PERMANENT GENETIC RESOURCES: A set of primer pairs for amplifying the complete mitochondrial DNA of endangered Chinese egret (Aves, Ardeidae, Egretta eulophotes). *Mol. Ecol. Resour.* 8: 412-414. http://dx.doi.org/10.1111/j.1471-8286.2007.01975.x

Zhou X, Lin Q, Fang W and Chen X (2014). The complete mitochondrial genomes of sixteen ardeid birds revealing the evolutionary process of the gene rearrangements. *BMC Genomics* 15: 573. http://dx.doi.org/10.1186/1471-2164-15-573

Zou Y, Jing MD, Bi XX, Zhang T, et al. (2015). The complete mitochondrial genome sequence of the little egret (*Egretta garzetta*). *Genet. Mol. Biol.* 38: 162-172. http://dx.doi.org/10.1590/S1415-4757382220140203

## Supplementary material

**Table S1.** Species examined in the present study and the data sources of sequences

**Figure S1.** Phylogenetic tree of Ardeidae constructed from *COI* sequences. Numbers (in internodes) represent bootstrap values (>80%) from 1000 replications. The codes are given in Table S1.