The Updated Phylogenies of the Phasianidae Based on Combined Data of Nuclear and Mitochondrial DNA

Yong-Yi Shen¹,², Kun Dai³, Xue Cao¹,⁶, Robert W. Murphy¹,⁴, Xue-Juan Shen², Ya-Ping Zhang¹,⁵

¹ State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, the Chinese Academy of Sciences, Kunming, China, ² Joint Influenza Research Centre (SUMC/HKU), Shantou University Medical College, Shantou, China, ³ Xinjiang Institute of Ecology and Geography, the Chinese Academy of Sciences, Urumqi, China, ⁴ Department of Natural History, Royal Ontario Museum, Toronto, Canada, ⁵ Laboratory for Conservation and Utilization of Bio-resources, Yunnan University, Kunming, China, ⁶ University of the Chinese Academy of Sciences, Beijing, China

Abstract

The phylogenetic relationships of species in the Phasianidae, Order Galliformes, are the object of intensive study. However, convergent morphological evolution and rapid species radiation result in much ambiguity in the group. Further, matrilineral (mtDNA) genealogies conflict with trees based on nuclear DNA retrotransposable elements. Herein, we analyze 39 nearly complete mitochondrial genomes (three new) and up to seven nuclear DNA segments. We combine these multiple unlinked, more informative genetic markers to infer historical relationships of the major groups of phasianids. The nuclear DNA tree is largely congruent with the tree derived from mt genomes. However, branching orders of mt/nuclear trees largely conflict with those based on retrotransposons. For example, Gallus/Bambusicola/ Francolinus forms the sister-group of Coturnix/Alectoris in the nuclear/mtDNA trees, yet the tree based on retrotransposable elements roots the former at the base of the tree and not with the latter. Further, while peafowls cluster with Gallus/Coturnix in the mt tree, they root at the base of the phasianids following Gallus in the tree based on retrotransposable elements. The conflicting branch orders in nuclear/mtDNA and retrotransposons-based trees in our study reveal the complex topology of the Phasianidae.

Introduction

Rapid species radiations often result in ambiguous phylogenetic relationships because too little time is available to accrue and fix shared derived character states. This may manifest itself, in part, as incomplete lineage sorting. In such cases, different datasets, such as genes, may resolve conflicting suites of relationships. The Phasianidae, one of four families in the Galliformes, typifies this problem. Rapid radiation and convergent morphological evolution confound the resolution of relationships for many pheasants and partridges. Although the family has been the target of much phylogenetic research [1–17], not surprisingly, many unsolved nodes and much conflict remain.

Most previous molecular studies of phasianids analyze either one or a few mitochondrial (mt) genes [3,9,10,12], a single nuclear gene [1,18], or a combination of mt and a few nuclear gene sequences [2,7]. Employing complete mt genomes, Shen et al. [19] resolve a well-supported topology. The topology (Figure 1A) is largely congruent with previous molecular studies based on mt genes and nuclear segments [2,7]. However, this tree strongly conflicts with that based on retrotransposable elements (Figure 1B) [6,17,20]. For example, in the matrilineral genealogy, Gallus forms the sister-group of Coturnix, while the tree based on retrotransposable elements roots Gallus at the base of the phasianids, and Coturnix is the sister-group of the gallophoasants. Further, while Pavo is the sister-group of Gallus/Coturnix in the mt tree, it roots at the base of the phasianids following Gallus in the tree based on retrotransposable elements. Trees based on nuclear genes exhibit many ambiguous nodes (Figure 1C).

The branching order in trees based on the mt genome conflicts with those derived from nuclear retrotransposons. This incongruence requires a reassessment of the phylogeny of the Phasianidae. Mitochondrial DNA markers reflect the matrilineral genealogy only; they do not provide information on paternal contributions. For retrotransposons, only a few phylogenetic approaches use indels as characters. Few genomes are available from which to design conservative retrotransposon primers for phasianids; only the chicken and a limited number of other avian genomes are available. This situation makes it difficult to obtain a sufficient number of phylogenetically informative characters. Considering the shortcomings of the mt and retrotransposon approaches, herein we report the sequencing of up to seven independent nuclear segments for 20 species and the complete mt genomes of three phasianids. We obtain other mt genomes from GenBank (Table S1 in File S1) and then combine these unlinked markers for the major groups of the Phasianidae to infer phylogenetic relationships.
Materials and Methods

Specimens Sampling

The Animal Use Ethics Committee of the Kunming Institute of Zoology, the Chinese Academy of Sciences approved the study. *Argusianus argus*, *Crossoptilon crossoptilon* and *Ithaginis cruentus* were used for mt genome sequencing. A total of 23 species were used for nuclear gene sequencing (Table S1 File S1). Feather samples of *Argusianus argus* were provided by Beijing Zoo and the Museum of the Kunming Institute of Zoology provided muscle tissue for all other samples. Additional complete mt genomes and nuclear segments were obtained from GenBank (Table S1 in File S1).

DNA Extraction, PCR Amplification, and Sequencing

Total genomic DNA was extracted using standard 3-step phenol/chloroform extraction methods [21]. For mitochondrial genomes, primers were described in our previous study [19]. For seven nuclear segments (*BDNF*, *CMOS*, *FIB4*, *NGFB*, *NTF3*, *OVOG*, and *ZENK*), primers were described in Table S2 in File S1. PCR amplifications were conducted in a 50 μl volume containing 5 μl of 10×reaction buffer, 0.2 mM dNTPs, 0.2 μM each primer, 1.5 U Taq DNA polymerase (TaKaRa Biosystems), and approximately 2 ng total DNA. PCR amplifications were carried out using the following parameters: 95°C 4 min, 20 cycles of denaturation at 94°C for 1 min, annealing at 60–50°C (1 min; 0.5°C/cycle), extension at 72°C for 1 min, and finally 15 cycles of 94°C 1 min, 50°C 1 min, 72°C 1 min. PCR products were cleaned using Watson RCR Purification Kits (Watson BioTechnologies, Shanghai). PCR products were sequenced at least three times in both directions on an ABI 3730 Sequencer (Applied Biosystems, Foster, CA, USA) using the ABI PRISM BigDye Terminator v3.0 sequencing kit. DNA sequences were edited using DNAstar Seqman software (DNASTAR Inc., Madison, WI, USA). The newly determined sequences were deposited in GenBank (GenBank accession numbers: JQ713766–JQ713768; JQ713656–JQ713765).

Phylogenetic Reconstruction

The nucleotide sequence data sets were initially aligned using ClustalX 1.81 [22] with default parameters. The combined and individual 13 mitochondrial protein coding genes, and the combined data of seven nuclear segments were analyzed separately using maximum likelihood (ML) implemented in PAUP* 4.0b10 [23]. Modeltest 3.7 [24] was used to select the preferred models of evolution under the Akaike Information Criterion. ML heuristic searches used TBR branch swapping executed in 100 replicates with the selected models. Because heuristic searches in PAUP* were very slow, we used two additional fast ML-based inference packages using 1,000 replicates each: RAxML [25] and PHYML [26]. Because their topologies were identical, and only a few bootstrap values slightly differed, we only presented trees with bootstrap values from PAUP*.

Figure 1. Phylogenetic hypotheses from the mitochondrial (mt) genome and retrotransposable elements for the Phasianidae. (a) Topology based on mt genomes (Shen et al. 2010); (b) topology based on insertion events of CR1 retrotransposable nuclear DNA elements [6,17,20]; (C) topology based on nuclear DNA segments (Crowe et al., 2006).

doi:10.1371/journal.pone.0095786.g001
Two separate runs were performed with four Markov chains. Each run was conducted with $3 \times 10^6$ generations and sampled every 100 generations. When the log-likelihood scores were found to stabilize, a consensus tree was calculated after omitting the first 25% trees as burn-in. In all these topology reconstruction, *Alectura lathami* was set as the outgroup according a previous study [19].

**Results**

**Phylogenetic Analyses of the Mitochondrial DNA Dataset**

We evaluated 39 mt genomes including those from GenBank. The 13 protein-coding genes consisted of 11,359 aligned nucleotide positions and the best-fit model of evolution was GTR+I+G. ML and BI analyses involving equal weight for each position resolved a single, robust tree (Figure 2A). Eight lineages, each with very high BSPs and BPPs, were resolved as follows: Group 1, *Arborophila*; Group 2, *Tragopan, Lophophorus, and Tetraophasis*; Group 3, *Chrysolophus, Phasianus, Lophura, and Symaticus*; Group 4, *Perdix*; Group 5, *Pucrasia*; Group 6, *Gallus, Bambusicola, and Francolinus*; Group 7, *Coturnix and Alectoris*; and Group 8, the peacocks.

Three additional weighting strategies were applied to the analysis of combined 13 protein-coding genes to avoid possible bias of nucleotide composition and saturation: (1) excluding the 3rd codon positions (Figure S1), (2) reencoding the 3rd codon position nucleotides to two-state categories, R (purine) and Y (pyrimidine) (Figure S2), and (3) reencoding the 1st and the 3rd codon position nucleotides to RY categories (Figure S3). The major topologies based on these weighting strategies were the same as evaluating all positions equally (Figure 2A).

Individual mt gene trees (Figure S4) were largely congruent with the mt genomic tree. As expected because of a low number of potentially phylogenetically informative characters for individual genes, nodes were supported by lower BSPs.

**Phylogenetic Analyses of Nuclear Dataset**

Segments of seven nuclear genes (BDNF, CMOS, FIB4, NGFB, NTF3, OVOG, and ZENK) were sequenced (Table S1 in File S1). Combined, these data consisted of 4,604 nucleotide positions. The best-fit model of nucleotide substitution was TrN+I+G. ML and BI analyses of the combined data resolved a single tree (Figure 2B). The nuclear tree was largely congruent with mitochondrial tree although some nodes conflicted. For example, *Ithaginis cruentus* rooted at the base of Galliopheasants/tragopans in mt tree but the nuclear tree resolved it as the sister group of *Pucrasia*. The position

![Figure 2. Phylogenetic hypotheses based on the complete mitochondrial genomes and nuclear segments for the Phasianidae.](https://example.com/figure2.png)
of peafowls was unstable in nuclear tree, while in mt tree they rooted at the base of Gallus/Old World quails.

**Phylogenetic Analyses of the Combined Mitochondrial and Nuclear DNA Dataset**

The mt and nuclear datasets shared 23 species. The combined dataset consisted of 15,972 aligned nucleotide positions. The best-fit model of nucleotide substitution was GTR+I+G. The topology (Figure S5) based on the combined dataset was nearly identical to that of mt tree (Figure 2A); they differed in the position of Phasianus. The mt genome tree clustered Phasianus with Chrysolophus with low support values (BPP = 84, and BSP = 60). In turn, this group clustered as the sister group of Lophura/Crossoptilon (BPP = 100, and BSP = 100). Trees based on the combined dataset clustered Chrysolophus with Lophura/Crossoptilon (BPP = 99), then with Phasianus (BPP = 100). Trees derived from combined dataset were largely congruent with the nuclear gene phylogeny. However, a few conflicts occurred. For example, the nuclear gene tree resolved Ithaginis and Pavo as sister taxa, while in combined and mt genome trees separated them far apart.

**Discussion**

Our mt genome tree depicts Archaeophila as the sister-group to all other phasianids plus the Meleagrididae and Tetraonidae. Tetraoninae clusters independently with Lophophorus and their sister-group is Tragopan. Perdix and Archaeophila do not cluster with other partridges. The non-monophyly of the pheasants and partridges is more common than not and this resolution involves a strongly supported association of Gallus, Bambusicaula, and Francolinus. The previous matrilineal genealogy did not cluster Polyplectron with Pavo [16,19]. Herein, we add a new peacock–Argusianus argus. Polyplectron forms the sister-group of Gallus/Bambusicaula/Coturnix albeit with relatively low support. This group clusters with other peacocks (Pavo and Argusianus). The phylogenetic position of Polyplectron remains unstable. More data involving new peacocks may further resolve this group’s position.

We add the nearly complete mt genome of Crossoptilon crossoptilon and analyses involving this species clusters it with Lophura (BPP = 100; BSP = 100). Analyses of the new mt genome of Ithaginis cruentus strongly unites it with the galloansate Perdix/tragopans/Tetraoninae/Meleagrididae (BPP = 100; BSP = 100). Tragopan, Ithaginis, Pavo, Lophophorus form the tragopans tribe [23]. Our resolution of a sister relationship for Tragopan and Lophophorus is well supported, but Ithaginis and Pavo do not cluster with them. The mt genome tree also does not support the morphological and behavioral placement of Ithaginis as the sister-group of New World and Old World quails [4]. Previous molecular studies did not conclusively resolve the phyletic position of Ithaginis [2,10].

The mt genomes provide a greater abundance of information, thus, have a greater likelihood of fully resolving a tree than individual protein-coding genes (Figure S4). In our analyses, almost all nodes receive very high support. Thus, the rapid rate of mutation renders the mt genome phylogenetically informative at the levels of genera and species for the phasianids.

Reliance on mt data for phylogenetic reconstruction may be fraught with problems. Functionally, in most animals the mt genome serves as a single, large genetic locus and it provides a matrilineal perspective only on the evolutionary history of a group [29,30]. Paternal contributions are not considered. Thus, mtDNA data alone are often inadequate for macroevolutionary phylogenetic analyses, especially in the face of complex evolutionary scenarios such as gene introgression, hybridization, and/or selection [31]. Our seven nuclear segments address this concern. BSIs tend to be lower in nuclear tree compared to the mt tree. The relatively slow rate of mutation rate of nuclear DNA compared to mtDNA generally results in relatively poorly resolved nuclear gene trees.

Previous studies based on a single nuclear gene failed to solve many nodes [1,18]. In contrast, our multi-gene analyses resolve many nodes with very high levels of support. This result indicates that additional informative sites greatly help to resolve ambiguous relationships. Our nuclear phylogeny is largely congruent with trees derived from mt genomes. However, our nuclear tree resolves Ithaginis and Pavo as sister groups, but the mt tree depicts divergent relationships. The position of Ithaginis was ambiguous in the previous morphological-behavioral parsimony cladogram of Dyke et al. [2003] and molecular studies [2,9,10]. The position of Ithaginis received high BSIs in our mt and nuclear trees, yet further explorations into the nature of conflicting trees is necessary. The positions of peafowls are unstable in nuclear tree. Pavo and Polyplectron do not cluster together, including in our mt tree.

The tree based on combined mitochondrial and nuclear dataset results a well-supported tree (Figure S5). In mt genome tree (Figure 2A), the position of genus Phasianus is not well supported; it clusters with Chrysolophus with low support values (BPP = 84; BSP = 60). In contrast, the trees of the combined datasets cluster Phasianus with (Chrysolophus, Lophura/Crossoptilon), and with strong support. The positions of peafowls are unstable in the mt genome, nuclear data, and combined data trees. More species of peafowl and additional markers may resolve the position of the peafowl.

Retrotransposons-based trees (Figure 1B) strongly conflict with nuclear/mtDNA trees (Figure 2) at the level of genus. For example, in the latter case, Gallus/Bambusicaula/Francolinus forms the sister-group of Coturnix/Alectoris, while the former analyses root Gallus/Bambusicaula/Francolinus at the base of the phasianids, and Coturnix/Alectoris formed the sister-group of galloansate/tragopans. Further, while peafowls cluster with Gallus/Coturnix in the mt tree, they root at the base of the phasianids following Gallus in the tree based on retrotransposon-sasible elements.

Retrotransposon data often consist of insertion/deletion (indel) events. Only a few phylogenetic approaches use indels as characters. Most researchers either delete them or treat the gaps as missing data. Indels cannot resolve relationships of clades branching off the focal clade–the lineage leading to the species in which the markers are originally identified–either before or after the insertion event [6]. Further, few genomes are available for identifying retrotransposon markers. These consist of the chicken and a limited number of other avian genomes. This paucity not only limits the design of conservative PCR primers for the target group, it also limits identification of an adequate number of informative characters. Conclusions based upon a few markers may be lead to inaccurate findings [17,32]. Thus, retrotransposons appear to be severely limited in their ability to resolve relationships at the hierarchical levels of genus and species, especially in cases of rapid radiations of species. This may explain the conflicting branching orders.

In conclusion, we combine mt genomes and segments of seven nuclear genes to reassess the phylogenetic relationships of phasianids. These multiple unlinked and informative genetic markers provide an updated topology. Our nuclear gene phylogeny is largely congruent with trees derived from mt genomes. However, our mt and nuclear topology largely conflict retrotransposon-based trees.
Supporting Information

Figure S1  Bayesian inference analyses of 13 mt genes that excluding the 3rd codon position. (TIF)

Figure S2  Bayesian inference analyses of 13 mt genes that recoding the 3rd codon position nucleotides to two-state categories, R (purine) and Y (pyrimidine). (PDF)

Figure S3  Bayesian inference analyses of 13 mt genes that recoding the 1st and the 3rd codon position nucleotides to two-state categories, R (purine) and Y (pyrimidine). (TIF)

Figure S4  Bayesian inference analyses of individual mt genes and control region (CR). Each run was conducted with 5,000,000 generations and sampled every 100 generations. Bayesian Posterior Probabilities >70% were indicated on the branches. (A) 12S, 1,036 aligned sites; (B) 16S, 1,702 aligned sites; (C) ATP6, 681 aligned sites; (D) ATP8, 165 aligned sites; (E) CoxI, 1,548 aligned sites; (F) CoxII, 681 aligned sites; (G) CoxIII, 783 aligned sites; (H) CR, 1,352 aligned sites; (I) ND1, 972 aligned sites; (J) ND2, 1,038 aligned sites; (K) ND3, 548 aligned sites; (L) ND4, 1,377 aligned sites; (M) ND5L, 291 aligned sites; (N) ND5, 1,818 aligned sites; (O) ND6, 519 aligned sites; (P) GypB, 1,137 aligned sites.

Figure S5  Bayesian phylogenetic tree based on the combined dataset of 13 mt protein-coding genes and seven nuclear segments. (TIF)

References

1. Armstrong MH, Braun EL, Kimball RT (2001) Phylogenetic utility of avian ovomucoid intron G: A comparison of nuclear and mitochondrial phylogenies in Galliformes. Auk 118: 799–804.
2. Crowe TM, Bowie RCK, Blommer P, Mandhiana TG, Hedderson TAJ, et al. (2006) Phylogenetics, biogeography and classification of, and character evolution in, gamebirds (Aves: Galliformes): effects of character exclusion, data partitioning and missing data. Cladistics 22: 495–532.
3. Dinhcheff DE, Drovetski SV, Mindell DP (2002) Phylogeny of Tetraoninae and other galliform birds using mitochondrial 12S and ND2 genes. Mol Phylogenet Evol 24: 283–293.
4. Dyke GJ, Gulas BE, Crowe TM (2003) Suprageneric relationships of galliform birds (Aves, Galliformes): a cladistic analysis of morphological characters. Zool J Linn Soc 137: 237–244.
5. Fumihito H, Miyake T, Takada M, Ohtsu S, Kondo N (1995) The genetic link between the chinese bamboo partridge (Bambusicola fytchae) and the chicken and junglefowl of the genus Gallus. Proc Nat Acad Sci USA 92: 11053–11056.
6. Kaiser VB, van Tuinen M, Ellegren H (2007) Insertion events of CR1 retrotransposable elements elucidate the phylogenetic branching order in galliform birds. Mol Biol Evol 24: 338–347.
7. Kimball RK, Braun EL (2000) A multigene phylogeny of Galliformes supports a single origin of crested ability in non-feathered facial traits. J Avian Biol 39: 403–445.
8. Kimball RK, Braun EL, Ligon JD (1997) Resolution of the phylogenetic position of the Congo peafowl, Afropavo congensis: a biogeographic and evolutionary enigma. Proc R Soc Lond B 264: 1517–1523.
9. Kimball RK, Braun EL, Ligon JD, Lucchini V, Randi E (2001) A molecular phylogeny of the peacock-pheasants (Galliformes: Phasianidae spp.) indicates loss and reduction of ornamental traits and display behaviours. Biol J Linn Soc 73: 187–198.
10. Kimball RK, Braun EL, Zwarich PW, Crowe TM, Ligon JD (1999) A molecular phylogeny of the peafowl and partridges suggests that these lineages are not monophyletic. Mol Phylogenet Evol 11: 38–54.
11. Meng Y, Dai B, Ran J, Li J, Yue B (2008) Phylogenetic position of the genus Tetraonias (Aves, Galliformes, Pheasidae) as inferred from mitochondrial and nuclear sequences. Biochem Syst Ecol 36: 626–637.
12. Pereira SL, Baker AJ (2006) A molecular timescale for galliform birds accounting for uncertainty in time estimates and heterogeneity of rates of DNA substitutions across lineages and mt. Mol Phylogenet Evol 38: 499–509.
13. Randi E, Lucchini V, Armijo-Prewitt T, Kimball RK, Braun EL, et al. (2000) Mitochondrial DNA phylogeny and speciation in the tragopans. Auk 117: 1003–1013.
14. vanTuinen M, Hedges SB (2004) The effect of external and internal fossil calibrations on the avian evolutionary timescale. J Paleont 78: 45–50.
15. Kan XZ, Yang JK, Li XF, Chen L, Lei ZP, et al. (2010) Phylogeny of major lineages of galliform birds (Aves: Galliformes) based on complete mitochondrial genomes. Genet Mol Res 9: 1625–1633.
16. Bonilla AJ, Braun EL, Kimball RT (2010) Comparative molecular evolution and phylogenetic utility of 7–UTRs and introns in Galliformes. Mol Phylogenet Evol 56: 536–542.
17. Liu Z, He L, Yuan H, Yue B, Li J (2012) CR1 retroposons provide a new insight into the phylogeny of Phasianidae species (Aves: Galliformes). Gene 502: 125–132.
18. Smith EJ, Shi L, Tu ZJ (2005) Gallus gallus aggrecan gene-based phylogenetic analysis of selected avian taxonomic groups. Genetica 124: 23–32.
19. Shen YY, Liang L, Sun YB, Yue BS, Yang XJ, 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.
20. Kriegs JO, Matzke A, Churakov G, Kuritzin A, Mayr G, et al. (2007) Waves of archaeal transposable element insertions homoplasy free?: An examination using the Akaike information criterion and Bayesian model-averaged reagents/materials/analysis tools: YPZ. Wrote the paper: YYS YPZ RWM.
21. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. New York: Cold Spring Harbor Laboratory Press.
22. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Version 1.8. Nucleic Acids Res 24: 4676–4682.
23. Swoford DL (2003) PAUP*: 4.06. 10: Phylogenetic analysis using parsimony. Sunderland (MA): Sinauer Associates.
24. Posada D, 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.
25. Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22: 2688–2690.
26. Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52: 696–704.
27. Rouquist F, Hueslenbeck JP (2003) MrBayes: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572–1574.
28. Johansgard PA (1986) The pheasants of the world. Oxford: University Press New York.
29. Zhang DX, Hewitt GM (2003) Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. Mol Ecol 12: 563–584.
30. Ballard JWO, Whitlock MC (2004) The incomplete natural history of mitochondria. Mol Ecol 13: 729–744.
31. Chan KMA, Levin SA, Gavrilets S (2005) Leaky prezygotic isolation and porous genomes: rapid introgression of maternally inherited DNA. Evolution 59: 720–729.
32. Han KL, Braun EL, Kimball RT, Reddy S, Bowie RCK, et al. (2011) Are transposable element insertions homoplasy free?: An examination using the avian tree of life. Syst Biol 60: 375–386.