Phylogenetic and evolutionary revision of some pheasants of northern Pakistan

Izaz Ali, Aman Khan, Talmiz Ur Rehman, Minhas Naseer, Shaukat Ali Khan and Humayoun Khan

DOI: https://doi.org/10.33545/27080013.2020.v1.i1a.4

Abstract
Present study describes the phylogenetic reconstruction and evolutionary relations of seven local pheasant species on the basis of cytochrome b and cytochrome c oxidase I gene sequences. DNA was isolated from blood samples, partial sequences of cytochrome c oxidase (850 bp) and cytochrome b gene (450 bp) were PCR amplified using bird specific universal primers. The nucleotide sequences obtained were compared with that of several members of Phasianidae and other bird species having above 70% homology and phylogenetic trees were constructed using Maximum Likelihood Analysis. The gene sequences obtained from local species have shown multiple conserved regions indicating a common ancestor of these species. The results have suggested a revised systematic position of local pheasants and their phylogenetic distance with related species. The investigated species have been placed in two clades and five sub-clades on the basis of COI gene sequence. The cytochrome b gene sequence has differentiated the species into three clades. Our findings suggest that COI gene can discriminate between similar bird species more efficiently than cytochrome b gene.

Keywords: Phylogenetics, Pheasants, Cytochrome b, Cytochrome c oxidase I

Introduction
Phasianidae includes one of the most important groups of birds for both human society and research purposes. More than 25% species of this family are threatened (vulnerable, critically endangered) worldwide (IUCN red list, 2010). There is curiosity in the use of phylogenetic information to update the conservation agencies (Moore et al., 2003) [28]. The group is also being used as a model to establish conservation priorities from phylogenetic information, making it more critical to have an established knowledge of its Phylogenetics (Khai and Yabe, 2014) [19]. However, the relationships among the members of Phasianidae are still very unclear (Kimball et al., 1999; Dyke et al., 2003; Kimball et al., 2011) [21, 7, 22]. Therefore; the investigations about the phylogeny of pheasants are used to limit the gaps in the evolutionary relationships and in making conservation strategies.

The cytochrome c oxidase gene I and cytochrome b gene of mitochondrial DNA have been extensively used for the phylogenetic and evolutionary analysis of species (Wink et al., 2009; Pacheco et al., 2011; Xiang et al., 2014; Prum et al., 2015) [35, 28, 36, 30]. Cytochrome b gene is a useful marker for the identification of species (Bensch et al., 2016) [5]. However, DNA bar-coding which makes the use of mitochondrial cytochrome c oxidase gene I (COI) is considered as one of the most efficient and reliable method. Studies have shown that more than 95% of species possess unique COI sequences (Hajibabaei et al., 2005; Hickerson et al., 2006; Khan et al., 2010; Arif et al., 2011) [10, 12, 20, 1]. Major objective of DNA bar-coding is the identification of unknown specimens to improve the chances of discovery of new species and to find out the specific lineages of organisms. After identification, the DNA barcodes can use the available data to find out genetic variation at the global framework by using the sequences available in the genbank (Hebert et al., 2003; Bilgin et al., 2016) [11, 3]. Mitochondrial DNA based documentation of birds has been limited to North America, Korea, Argentina and Scandinavia in the past (Yoo et al., 2006; Kerr et al., 2009; Johnsen et al., 2010) [37, 18, 15]. Present study reveals the phylogeny of local pheasants of Northern areas of Pakistan on the basis of cytochrome b and COI gene sequences. There are several breeding centers of pheasants in the country, one of them located at Dodhial (Mansehra) shelters about 4000 birds. Most of these birds include the declared endangered/threatened species.
Materials and Methods

Materials
Molecule biology grade DNA isolation, PCR and electrophoresis chemicals and reagents were obtained from Sigma-Aldrich, Thermo Fischer and local biochemical companies.

Blood samples were obtained from the Dodhial Pheasant Centre established under World Wildlife Fund (WWF). A small amount of blood was collected by an expert doctor through the puncture of main wing or leg vein. The blood samples were stored in EDTA tubes at -20°C and processed for DNA extraction. The blood sample (50uL) was mixed with 3 volumes of dilution buffer (100mM sodium phosphate buffer pH 7.3 containing 20mM NaCl and 5 ug per micro liter proteinase K) and incubated at 56°C for 30 min. Lysis solution (200uL containing 0.5% SDS, 2mM DTT prepared in phosphate buffer pH 7.5) and incubated at 60°C for 3 hours. Phenol-chloroform (400uL) was added to the lysed sample, mixed by inverting the tubes, incubated at room temperature for 2-3 min and centrifuged at 12000 g and 4°C for 15 min. The upper aqueous layer was carefully transferred into to a fresh Eppendorf tube, mixed with equal volume of ice cold isopropanol, incubated at -20°C for 10 min and centrifuged as above. The DNA pellet was washed with 75% cold ethanol and air dried. The DNA was dissolved in 50uL of double distilled autoclaved water.

The purified DNA was used for the PCR amplification of COI and cytochrome b genes. Universal bird specific primers were used for PCR (Hebert et al., 2004) [29]. The sequences of primers used are as follows: Cyt-bF cctcctcaacctactcagatgatgatgattgcctctat COIF cggcaagtatgatgtgccatttag, Cyt-bR cctcctcaacctactcagatgatgatgattgcctct COIR atttcatgtgagccctgtcctctca. The PCR amplification of both genes was carried out in the reaction mixture containing 2 mM MgCl$_2$, 0.8 mM dNTPs, 1X taqbuffer, 20-40ng of template DNA, 40 picomoles forward and reverse primers, 2.5 U of Taq polymerase with final volume of 25uL. The nucleotide sequence of Cyt-b and Cyt-c gene from target species/varities were subjected to BLAST analysis in NCBI and related sequences available in the gene bank were used for the alignmant by different online softwares. After detailed analysis of cytochrome c oxidase I and cytochrome b gene sequences the phylogenetic trees were constructed to indicate the phylogenetic affinities of pheasants under investigation with that of related bird species. Evolutionary analyses were made by using MEGA6 (Tamura et al., 2013). The nucleotides sequences coding for cytochrome b and COI genes in the species investigated were aligned and the conserved sequences were highlighted to indicate the level of homology and to find out the chances of common origin of these species.

Results
A good quality DNA obtained from blood samples was used for the PCR amplification of partial cyt-b and COI gene sequences. For COI 850 bp and for cytochrome b 450 bp sequences were amplified. Purified PCR products were commercially analyzed for the nucleotide sequences (Macrogen Inc. South Korea).

Phylogenetic relationship of pheasants from Northern Pakistan with 22 other related species and genera of family Phasianidae were retrieved from Genbank and aligned using MUSCLE alignment. The aligned data sheet has shown a maximum 2834 genetic characters and after trimming, the extra and ambiguously aligned fragments from the both 5’ and 3’ ends of alignment data sheet remaining 753 genetic characters were used for further analysis. Codon positions included were 1st+2nd+3rd+Noncoding. Mean sequence divergence of COI gene within each lowest level taxonomic unit (species) was from 0.044% to 67.10%. A discrete Gamma distribution was used to model evolutionary rate differences among sites with two categories (+G+I). The estimated value of the shape parameter for the discrete Gamma distribution (+G) was observed 200.0000. The proportion of sites (+I) was estimated 0.0000% to be invariant. Substitution pattern and rates were estimated under the (JC+G+I) model (Jukes-Cantor, 1969) [16]. Mean evolutionary rates in these categories were 0.94 and 1.06 substitutions per site. The nucleotide frequencies were A = 25.00%, T/U = 25.00%, C = 25.00%, and G = 25.00%.

The evolutionary history was inferred by using the Maximum Likelihood (ML) method based on the Jukes-Cantor model. The tree with the highest log likelihood (-13254.1986) is shown (Figure 1). Initial tree(s) for the heuristic search were obtained automatically by applying the Maximum Parsimony method. There were 753 positions/characters were used for further analysis. Out of analyzed trait, 1 was conserved, 752 were variable, 751 were parsimony informative and 2 were singleton sites. Bootstrap supported values are presented above the branches. Maximum Parsimony analysis was performed to evaluate the evolutionary history. The most parsimonious tree with length = 4273 is shown (Figure 1). The consistency index (CI) was 0.489352 (0.489352), the retention index (RI) was 0.662751 (0.662751), and the composite index is 0.324318 (0.324318) for all sites and parsimony-informative sites (in parentheses). The MP tree was obtained using the Tree-Bisection-Regrafting (TBR) algorithm (Nei and Kumar, 2000) with heuristic search of 1000 replicates with random stepwise addition and the bootstrap 50% majority-rule consensus. For the phylogenetic analysis through cytochrome b gene sequence, 25 sequences including 7 sequences from present study were used for the construction of phylogenetic tree. Mean evolutionary rates in these categories were 0.94, 1.06 substitutions per site. The nucleotide frequencies are A = 25.00%, T/U = 25.00%, C = 25.00%, and G = 25.00%. The evolutionary history was inferred by using the Maximum Likelihood (ML) method based on the Jukes-Cantor model (Jukes-Cantor, 1969) [16]. The tree with the highest log likelihood (-7044.4538) is shown (Figure 2). Initial tree(s) for the heuristic search were obtained automatically by applying the Maximum Parsimony method. There were 414 positions/characters were used for further analysis. Out of analyzed traits, 1 was conserved, 413 were variable, 413 were parsimony informative and 1was single tone sites. The most parsimonious tree with length = 2311 is shown (Fig.2).

The consistency index (CI) was 0.396365 (0.396365), the retention index (RI) was 0.582960 (0.582960), and the composite index is 0.231065 (0.231065) for all sites and parsimony-informative sites (in parentheses). The MP tree was obtained using the Tree-Bisection-Regrafting (TBR) algorithm (Nei and Kumar, 2000) with heuristic search of 1000 replicates with random stepwise addition and the bootstrap 50% majority-rule consensus. Bootstrap supported values are presented above the branches. Phylogenetic tree consisted of three major clades. The clade I was comprised of twelve species of pheasant group, our local species lady,
Amherst pheasant, golden pheasant and black shoulder peacock. These species were clustered and formed sub group and showed genetic similarity with *Scleropitilastreptophorus* and *Symmaticuselsoni* with supported parsimony bootstrap values (MPB=88 and MPB=85 whereas branch length 0.00000 and 1.14353) respectively. Clade II included seven sequences, our local species cheer pheasant, and golden pheasant from Pakistan were clustered into this clade and showed close genetic relationship with *Chrysolophus* group with maximum parsimony bootstrap supported value(MPB=87 and branch length=0.22268). Clade III comprised of six sequences, pied peacock and local hen sequences from Pakistan fell into this clade showing highest homology with sequences of this clade along with supported bootstrap value (MPB=84, branch length=0.29300).

The alignment of nucleotide sequences has revealed the presence of several conserved patches in the cytochrome b and COI gene sequences of species under investigation (Figure 3 and 4) indicating the chances of common ancestor of these species.

**Discussion**

Short nucleotide sequences have been successfully used to make standard identification of organisms under the terms of DNA bar-coding or DNA taxonomy (Floyd et al., 2002; Tautz et al., 2003) [33]. The most promising application of this technology are the assignment of unknown life-history stages to adult organisms, the large scale identification of organisms in ecological investigation of phylogeny and potential description of unidentified “candidate” species from the ecosystem. It is new and technically growing procedure in modern Phylogenetics and systematic (Moritz and Cicero, 2004) [30]. The use of mitochondrial DNA as a genetic marker for recognition and identification and evaluation of genetic diversity and phylogeny to reveal evolutionary status of birds has a crucial importance (Kimball et al., 1999; Hackett et al. 2008; Czyzowski et al., 2008; Wang et al., 2013; Shen et al., 2014; Huang and Ke, 2014a) [21, 9, 6, 34, 32, 13]. In past, many traditional markers i.e. morphological, cytological and biochemical have used to classify and authenticate birds but we are reporting utilization of modern molecular technique to determine phylogenetic and evolutionary status of pheasants from Northern Pakistan for the first time. In this study, mitochondrial segments of genes (cytochrome-b and COI) were utilized to estimate genetic diversity and phylogeny of local pheasants, which are considered to threatened or endangered in this area, for conserving and improving the stock of the species. For phylogenetic analysis of birds, the nucleotide sequences of cytochrome b (cyt-b) gene have been extensively used more than any other mitochondrial gene (Kimbell et al., 1999; Czyzowski et al., 2008) [21, 6]. Keeping in view, the partial nucleotide sequence of mitochondrial cyt-b gene was obtained from the genome of selected 07 pheasant collected during 2013 and 2014 from the Dodhial Pheasant Centre established under World Wildlife Fund (WWF). After successful PCR amplification and nucleotide sequence analysis of partial cyt-b gene, the BLAST analysis was performed to obtain closely related already reported sequences from GenBank. Total 28 partial sequences along with the investigated were aligned together and the phylogenetic tree was constructed using Maximum Parsimony (MP). Based on the figure of the MP phylogenetic tree (Figure 2), three clades with mixed grouping of already reported and our 7 pheasants were observed. Previously, partial sequence of cyt-b gene was utilized to establish molecular phylogeny and lineage of pheasants (Kimbell et al., 1999; Czyzowski et al., 2008) [21, 6]. Recently, because of simplicity and robustness of cytochrome C oxidase I (COI), it has gained great attention by researchers for species authentication and molecular phylogeny of birds (Khan et al., 2010; Huang et al., 2014; Huang and Ke, 2014b; Schneider et al., 2016) [20, 13, 31]. For further analysis at molecular, another mitochondrial gene COI was sequenced to investigate phylogeny and lineage of selected species of pheasants from Northern Pakistan. The isolated partial nucleotide sequences of COI were compared with reported sequences in GenBank by doing BLAST analysis at NCBI and all the sequences were aligned and compared by Maximum Parsimony (MP) method. Furthermore, all aligned sequences were subjected to draw phenogram/tree using Maximum Likelihood (ML) method. The phylogenetic analysis of COI isolated from local pheasants and other 22 other species and genera of family Phasianidae data from GenBank were performed and two clades with subclades were observed in the ML based tree (Figure 1). Similar findings were reported in Chinese Phasianidae species and it has been concluded by various studies that COI could be very useful for DNA bar-coding to identify species and phylogenetic inference (Huang and Ke, 2014b; Schneider et al., 2016) [31].

So, we have successfully analyzed and compared the selected species of Northern Pakistan and other reported data to understand phylogenetic interference of pheasants. The species of this family (Phasianidae) have close relation with human society and that’s why facing various issues of phylogeny, geographic origin of lineage and evolution (Crowe et al., 2006; Kimball et al., 2011) [5, 22]. According to red list of IUCN, more 25% of them are threatened (IUCN, 2013) [14]. To address and resolve these issues of local pheasants, this study is helpful for conservation of these species by redraw of phylogenetic lineage using modern and reliable molecular technique (Kimball et al., 2001; Mooers et al., 2005; Wang et al., 2014; Schneider et al., 2016) [31]. In the past, many reports have been published to assess and reveal phylogeny of species/genera of family Phasianidae distributed various parts of world except our studied area using mitochondrial and other DNA based markers (Bush and Strobeck, 2003; Czyzowski et al., 2008; Shen et al., 2014; Wnag et al., 2014; Huang and Ke, 2014a; Huang and Ke, 2014b) [4, 6, 32, 13]. While some of these species have been investigated by using ISSR (Mehood et al., 2014) but this study is based on widely used molecular technique with elaborated sample size.
Fig 1: Phylogenetic tree resulting from the Maximum likelihood analysis of CO1 gene sequence. Parsimony bootstrap 50% majority-rule consensus values were given above the branches and branch length were presented below the branches. Scale bar presented 5 changes per 100 characters. (Sequences of pheasants from Northern Pakistan)
Fig 2: Phylogenetic tree resulting from the Maximum likelihood analysis of cytochrome b gene sequence. Parsimony bootstrap 50% majority-rule consensus values were given above the branches and branch lengths were presented below the branches. Scale bar presented 5 changes per 100 characters. (Sequences of the current study from Pakistan).
Fig 3: The conserved areas of cytochrome c oxidase gene I in different pheasants included in the present study.
Fig 4: The conserved areas of cytochrome b gene of different pheasants included in the present study.
Acknowledgements

The authors are thankful to the staff of WWF peasantry Mansehra Pakistan for providing samples. We are also grateful to HEC Pakistan and DSR at KAU for supporting this research.

References