Utility of mitochondrial COI gene for identification of wild ungulate species of conservational importance from Pakistan

Asif Naseem\textsuperscript{a,b}, Sajida Batool\textsuperscript{b} and Fakhar- i- Abbas\textsuperscript{c}

\textsuperscript{a}Institute of Molecular Biology and Biotechnology, The University of Lahore, Sargodha, Pakistan; \textsuperscript{b}Department of Zoology, University of Sargodha, Sargodha, Pakistan; \textsuperscript{c}Zoology, Centre for Bioresource Research (CBR), Islamabad, Pakistan

ABSTRACT

Most of the ungulates of Pakistan are either threatened or endangered species and their solitary and inaccessible life style makes them difficult to study. Therefore, estimating biodiversity, monitoring illegal trades and detecting commercial food frauds involving these species is a challenge for zoologists and conservation biologists. Here, we have attempted to exploit the discriminating power of mitochondrial COI gene to identify and to generate barcodes of the wild ungulate species of conservational importance found in Pakistan. 86 specimens of 19 wild ungulate species found in Pakistan were analyzed for their COI sequences. This is the first generated molecular data for many of these endemic and nearly endemic species. Intra and interspecific distances revealed distinct barcode gap for each species and a Neighborhood-joining tree able to discriminate all species into their respective clades. In conclusion, mtCOI is a powerful discriminatory tool for the taxonomic classification of ungulates especially for species that are inaccessible and require noninvasive sampling.

Introduction

Ungulates form the most diverse and thriving group of mammals in the world. Their numerous benefits make them a significant part of human society as most of the valuable economic domesticated animals are either artiodactyls or perissodactyls (Vaughan et al. 2011). They are well adapted to exploit variety of habitats like mountains, deserts, plains and aquatic regions (Feldhamer 2007). Landscape of Pakistan provides ideal habitat for 19 species of cetartiodactyls belonging to five representative families; Suidae, Cervidae, Moschidae, Bovidae and Platanistidae including some of the endemic and nearly endemic sub-species like Punjab urial (Ovis vignei punjabiensis), flared horned markhor (Capra falconeri falconeri), Sindh ibex (Capra aegagrus blythi) and Indus blind dolphin (Platanista gangetica minor) (Molur 2003).

The advancement in species identification research using molecular techniques, particularly DNA barcoding has been a great milestone in determining ungulate diversity worldwide. It has been applied to studies regarding biodiversity, phylogeny, food adulteration cases, illegal trades of ungulates and their manufactured products (Kumar et al. 2017; Zhong et al. 2017; Khan et al. 2018; Shukla et al. 2019; Bhaskar et al. 2020). However, most of the ungulate diversity of Pakistan has not yet been studied through mitogenome analysis and has been restricted to the pioneer morphological studies conducted by Schaller (1980), Mirza (1998), and Roberts (1997). In the present research, we have attempted to exploit the discriminating power of mitochondrial COI gene to generate reference barcodes for wild ungulates inhabiting Pakistan. This would be the first generated molecular data for many of these species. DNA barcodes of these species together with DNA sequences of other ungulate species retrieved from NCBI GenBank would also aid in resolving the long-time phylogenetic discrepancies of ungulates. The resulting data would also greatly aid in future studies involving taxonomic classification of these species.

Materials and methods

Taxon sampling and preservation

A total of 86 specimens of the 19 extant wild ungulate species were collected from different regions of Pakistan. Each species had at least three specimens except goitered gazelle (Gazella subgutturosa) and marcopolo sheep (Ovis ammon polii), both of which had one sample each due to rareness of the species and remote locations of habitat (Table 1). Specimen were morphologically identified with the help of wildlife expert (Dr. Fakhar-i-abbas and Jibran Haider) and morphological keys provided in literature (Roberts 1997; Groves and Grubb 2011).

Fecal samples were collected from zoos and breeding centers and immediately preserved in 70% ethanol, while Muscle...
tissues, skin and hair samples were obtained from cadavers from trophy hunters (Table 1). The acquired samples were preserved in 70% ethanol or kept at −20°C Celsius. The specimens were also deposited in the specimen bank of Center for Bioresource Research (CBR), Islamabad and given appropriate voucher IDs (Table 1).

**DNA extraction, COI amplification and sequencing**

Total genomic DNA of the collected specimen was extracted through a modified organic DNA extraction method (Sambrook et al. 2001). COI gene was amplified using universal vertebrate primers, VF1d (5′-TCTCAACCAACCAACAAAAA (Y)GAT(Y)GG-3′) and VR1d (5′-TAGACTTCTGGGTGCRAARAA (Y)CA-3′) (Cai et al. 2011). All PCR amplifications included a negative control reaction which lacked template DNA and were conducted in a Veriti Thermal Cycler (Applied Biosystems).

The PCR products were analyzed using 1.5% agarose gel electrophoresis and quantified by Nanodrop Spectrophotometer (ND-2000). After confirmation of successful PCR by measuring absorbance ratio at 260/280 nm, amplicons were sequenced by ABI sequencer using facility of Center of Excellence in Molecular Biology (CEMB) Lahore.

**Sequence analysis**

The generated sequences were analyzed by a sequence alignment editing software BioEdit v7.0.5 (Hall 1999) with manual proofreading, sequences were analyzed to minimize any chances of polymorphic bases before any further analysis. All generated sequences were aligned by ClustalW using any changes of polymorphic bases before any further analysis, sequences were analyzed to minimize manual proofreading, sequences were analyzed to minimize any changes of polymorphic bases before any further analysis. All generated sequences were aligned by ClustalW using MEGAX (Kumar et al. 2018) to study parsimony informative sites. Sequences were further identified by using species identification tools available in BOLD (Barcode of Life Data system) and through BLAST GenBank databases to validate their identifications. Furthermore, Sequence analysis tools of BOLD was also utilized for estimating intra and interspecific divergences, nearest neighbor (NN) distances as well as barcode gap. A neighbor-joining phylogenetic tree was generated using Jukes-Cantor model (Jukes and Cantor 1969) by BOLD analysis system (Taxon ID Tree).

The barcode data generated from the sequences of Pakistani species were submitted to NCBI GenBank through the barcode submission tool and their accession numbers are given in Table 1. These sequences were also deposited online in a dataset named UNGPK in Barcode of Life Data System.

**Results**

Successful sequencing resulted in 731-bp sequences of COI gene which were then trimmed to 656-bp readable fragments for further analyses. The comparison of newly generated barcode sequences of Pakistani ungulates against already reported sequences of NCBI GenBank revealed that almost all the sequences were 98–100% identical with their respective species. The details of BLAST results are given in Table S2. However, the sequence of musk deer showed 93% similarity to other musk deer species (KY792714.1) indicating its novelty in both the GenBank database. The sequence of gray goral (Naemorhedus goral) found in Pakistan is recognized as a sub-specie named Naemorhedus goral bedfordi and had 97% similarity to Naemorhedus goral (KT878720.1) reported in NCBI. Moreover, six of our sequences, Capra aegagrus blythi, Moschus cupreus, Naemorhedus goral bedfordi, Capra falconeri megaceros, Ovis vignei blanfordi and Ovis vignei punjabiensis were new to the database with no previous record available in databases.

The intraspecific and interspecific divergences were calculated using MEGAX. Intraspecific divergence ranged from 0 within sheep (Ovis aries) and straight horned markhor (Capra falconeri megaceros) to 0.0246 within marcopolo sheep (Ovis ammon). The interspecific divergence ranged from 0.0025 between bearded pig (Sus barbatus) and javan warty pig (Sus verrucosus) to 0.229 between Camelus dromedaries (Arabian camel) and Sus cebifrons (Visayan warty pig). According to our neighbor-joining tree (Figure 1), the overall phylogeny of ungulates remained same as previous studies and it supported all the benchmark clades of ungulates forming similar split ups and clades (Hassannin et al. 2012; Wang and Yang, 2013).

**Discussion**

The present study was designed to develop DNA barcode data for the ungulates of Pakistan, most of which are either threatened or endangered. The data generated through this study will provide a reference to future studies related to ungulate species identification in Pakistan such as food fraudulent adulteration cases, illegal trade of wildlife and animal products, biodiversity and population studies. Several of the species are unique and have no previous records at Genbank or BOLD.

Barcoding provides researchers with vast opportunities to expand the taxonomic research for extant species that are not correctly identified by morphological techniques. Phenotypic similarities among closely related species or subspecies can be problem in assessing the identity of an organism on just morphological basis (Lorenz et al. 2005). Moreover, identification only through morphological means is often compromised in cases of commercial food frauds and illegal trade of wildlife and their products due to the extreme processing of these illegal products (Baker 2008).

The instances of utilizing DNA barcoding as a species identification tool are numerous. Bitanyi et al. (2011) used a 470-bp fragment of COI gene to identify several species of Tanzanian antelopes. Yang et al. (2015) rectified an issue of falsely identified musk deer population of Shanxi province in China using 672-bp region of COI gene. According to that study, a musk deer population identified as Moschus moschiferus in previous studies was actually M. berezovskii.

The overall sequence divergence patterns were consistent with many previous studies (Kumar et al. 2017; Bhaskar et al. 2020). Sequence divergence within each ungulate species was <1. Within family Moschidae, the lowest evolutionary distance was 0.01 between Anhui Musk Deer (Moschus
Table 1. Sampling details of Pakistani ungulate species used in this study, including accession numbers, source, and conservation status.

| S# | Species name | Voucher IDs | No. of samples | Sample type | Accession no. | Source | Conservation status | Coordinates |
|----|--------------|-------------|----------------|-------------|---------------|--------|---------------------|-------------|
| 1  | Moschus cupreus (Kashmir Musk deer) | KMD-BRC | 5 | Skin/hair follicles | MG742692, MT251406, MT251407, MT251408, MT251409 | i) Pir Chinasi (Muzaffarabad, AJK) ii) Gurase Valley, AJK | Endangered | 34° 23'16"N 73° 32'51.9"E |
| 2  | Muntiacus muntjac (Barking Deer) | BD-BRC | 5 | Tissue | MG722903, MT251410, MT251411, MT251412, MT251413 | Margalla Hills National Park, Pakistan | Endangered | 33°43'52"N 72°56'13"E |
| 3  | Axis procinus (Hog Deer) | HD-L | 5 | Fecal | MG724969, MT251374, MT251375, MT251376, MT251377 | Lahore Zoo | Vulnerable | 31.556006°N 74.325959°E |
| 4  | Axis axis (Chital Deer) | CHT-L | 4 | Skin/hair follicles | MG742689, MT251371, MT251372, MT251373 | Lahore Zoo | Data Deficient | 31.556006°N 74.325959°E |
| 5  | Antilope cervicapra (Black Buck) | BB-BRC | 5 | Tissue | MG722904, MT251367, MT251368, MT251369, MT251370 | Lal-Suhanna National Park, Bahawalpur | Critical Endangered | 29°19'N 71°35'E |
| 6  | Gazella subgutturosa (Goitered Gazelle) | GB-BRC | 1 | Skin/hair follicles | MH261362 | Kalabagh game reserve | Vulnerable | 32°57'57.6"N 71°33'10.8"E |
| 7  | Gazella bennetti (Chinkara) | BT-J | 6 | Skin/hair follicles/fecal | MG742691, MT251378, MT251379, MT251380, MT251381, MT251382 | i) Pakistan Museum of Natural History, Islamabad; ii) Lahore Zoo; iii) Gakwal Wildlife Breeding Center, Faisalabad. | Endangered | i) 33°66'38.9"N 73°07.6389"E; ii) 31°59.0606°N 74°32.9596°E; iii) 31°28'41"N 72°12'20"E |
| 9  | Naemorhedus goral bedfordi (Gray Goral) | NGB-BRC | 3 | Hair follicles | MG742694, MT251414, MT251415 | Machhia National Park, AJK | Vulnerable | 34°30'18"N 73°33'45"E |
| 10 | Capra aegagrus blythi (Sind wild goat) | SI-J | 6 | Tissue | MG742690, MT251383, MT251384, MT251385, MT251386, MT251387 | Kirthar National Park | Near Threatened | 25°42'N 67°35'E |
| 11 | Capra sibirica (Siberian ibex) | HI-BRC | 4 | Skin/hair follicles | MH261363, MT251398, MT251399, MT251400 | Central Karakoram National Park | Least Concern | 35°54'0.33"N 75°31'39.62"E |
| 12 | Capra falconeri falconeri (Flared horned Markhor) | AMK-GB | 7 | Skin/hair follicles | MG724968, MT251388, MT251389, MT251390, MT251391, MT251392, MT251393 | Chitral Gol National Park | Endangered | 35°56'N 71°40'E |
| 13 | Capra falconeri megacerus (Straight Horned Markhor) | SMK-BL | 5 | Skin/hair follicles | MG742697, MT251394, MT251395, MT251396, MT251397 | Torkhar Hills, Baluchistan | Near Threatened | 34°36'49"N 72°47'18"E |
| 14 | Pseudois nayaur nayaur (Bharal) | BS-J | 5 | Skin/hair follicles | MG742688, MT251427, MT251428, MT251429, MT251430 | Khunjerab National Park | Endangered | 36°35'13.21"N 75°23'59.5"E |
| 15 | Ovis ammon polii (Marcopolo Sheep) | MS-K | 1 | Skin/hair follicles | MH261364 | i) Pakistan Museum of Natural History; ii) Khunjerab National Park | Critically Endangered | i) 33°66'38.9"N 73°07.6389"E; ii) 31°59.0606°N 74°32.9596°E |
| 16 | Ovis vignei punjabiensis (Punjab Urial) | PU-BRC | 5 | Skin/hair follicles/fecal | MG735444, MT251419, MT251420, MT251421, MT251422, MT251416, MT251417, MT251418 | i) Kirthar National Park; ii) Khunjerab Game Reserve | Endangered | 32°57'57.6"N 71°33'10.8"E |
| 17 | Ovis vignei blanfordi (Blandford Urial) | UB-J | 4 | Tissue | MG735443, MT251416, MT251417, MT251418 | Kirthar National Park | Vulnerable | 25°42'N 67°35'E |
| 18 | Ovis vignei vignei (Ladakh Urial) | LU-J | 5 | Skin/hair follicles | MG735445, MT251423, MT251424, MT251425, MT251426 | Chitral Gol National Park | Endangered | 35°56'N 71°40'E |
| 19 | Sus scrofa cristatus (Wild Boar) | WB-SGD | 4 | Skin/hair follicles | MF125268, MT251431, MT251432, MT251433 | i) Kirthar Hills, Baluchistan; ii) Lahore Zoo | Least concern | i) 31°57'15"N 72°42'26"E; ii) 31°55'06.0606°N 74°32.9596°E |

The accession numbers represent the newly generated sequences of Pakistani ungulates submitted to NCBI GenBank.

*Animal was found dead on Sargodha to Sillanwali road (near Kirana hills).*

**Conservation status according to data collected from Pakistan (Molur 2003).**
anhuiensis) and Forest Musk Deer (Moschus berezovskii) while, highest sequence divergence 0.07 was recorded of Kashmir musk deer with Forest Musk Deer and Siberian Musk Deer (Moschus moschiferus).

Our study eliminated many misperceptions regarding the taxonomy of Pakistani ungulates. Built on the basis of morphological studies is a common perception that three subspecies of Markhor (Capra falconeri) inhabit in Pakistan, including Kashmir markhor (Capra falconeri cashmiiriensis), flare horned markhor (Capra falconeri falconeri) and straight horned markhor (Capra f. megaceros) (Ali 2008; Ashraf et al. 2014). The percentage identity of Capra falconeri cashmiiriensis was 99% to Capra falconeri falconeri and thus, it was inferred that both sub-species are in fact one which is also supported by morphological studies (Schaller and Khan 1975).

The classification of Urial (three sub-species in Pakistan) has also been controversial as some researchers report it as sub-species of Mouflon (Ovis orientalis) (Frisina et al. 2007; Ayaz et al. 2012), while some as Urial (Ovis vignei) (Awan et al. 2006; Siraj-ud-Din et al. 2018). Comparison of newly generated sequences of Urial with sequences already reported in GenBank and BOLD shows its high sequence similarity (99%) with Ovis vignei (KF938361.1) while only 97% similarity to Ovis orientalis (KF938360.1), therefore, we suggest that the sub-species of Urial found in Pakistan are sub-species of Ovis vignei rather than Asiatic mouflon (Ovis orientalis).

The Sindh ibex (Capra aegagrus blythi), locally known as Sara, was recognized as a sub-species of wild goat (Capra aegagrus aegagrus) by Hume (1875) on the basis of morphological characteristics. Comparative analysis of our Sindh ibex sequences against Genbank databases showed same similarity index with three different species of genus Capra viz. (98.34%) domestic goat (Capra hircus: AB736134.1), (98.33%) Markhor (Capra falconeri: MG742698.1) and (97.74%) bezoar (Capra aegagrus: KR059222.1). Neighbor-joining tree placed Sindh ibex as a sister clade to markhor (Capra falconeri falconeri) (Figure 1) although this relationship was not supported by high bootstrap value (49%). Similar results were observed in a previous study conducted by (Sultana et al. 2003) on the basis of mitochondrial Cytochrome b and D-loop genes.

**Conclusion**

We conclude that mitochondrial COI is an efficient marker for species identification as well as for studying the phylogenetic relationships of ungulates. However, our study also specifies the requirement of increased taxon sampling as well as studying more genes in order to resolve the taxonomic and phylogenetic issues that surround ungulates. In addition, because DNA barcoding contributes a lot in the conservation of endangered species and monitoring their illegal trades, we promote the utilization of DNA barcoding to devise suitable conservation strategies for the endangered ungulates of Pakistan.
Acknowledgements

This article is core part of doctoral thesis of A. Naseem. It is bilateral project between Center for Bioresource Research (CBR), Islamabad and Department of Zoology, University of Sargodha. We are thankful to Jibran Haider, District Forest Officer, Gilgit Baltistan for his help in provision of several specimens and their morphological identification.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Asif Naseem http://orcid.org/0000-0002-4784-4812

References

Ali S. 2008. Conservation and status of markhor (Capra falconeri) in the northern parts of North West Frontier Province, Pakistan [MS thesis]. Missoula: College of Forestry and Conservation, University of Montana.

Ashraf N, Anwar M, Hussain I, Nawaz MA. 2014. Competition for food between the markhor and domestic goat in Chitral, Pakistan. Turk J Zool. 38(2):191–198.

Awan GA, Festa-Bianchet M, Ahmad T. 2006. Poaching, recruitment and conservation of Punjab urial Ovis vignei punjabiensis. Wildlife Biol. 12(4):443–449.

Ayaz S, Muhibullah AA, Jamil M, Khan MA, Qamar MF. 2012. Behaviour and biology of Ovis orientalis (urial) in Kotal Wildlife Park and Borraka Wildlife Sanctuary in Kohat. J. Anim. Pl. Sci. 22:29–31.

Baker CS. 2008. A truer measure of the market: the molecular ecology of fisheries and wildlife trade. Mol Ecol. 17(18):3985–3998.

Bhaskar R, Kanaparthi P, Sakthivel R. 2020. DNA barcode approaches to reveal interspecies genetic variation of Indian ungulates. Mitochondrial DNA B. 5(1):938–944.

Bitanyi S, Bjørnstad GRO, Ernest EM, Nesje M, Kusuikupa LJ, Keyyu JD, Mdegha RH, Reed KH. 2011. Species identification of Tanzanian antelopes using DNA barcoding. Mol Ecol Resour. 11(3):442–449.

Cai Y, Zhang L, Shen F, Zhang W, Hou B, Yue B, Li J, Zhang Z. 2011. DNA barcoding of 18 species of Bovidae. Chin Sci Bull. 56(2):164–168.

Feldhammer GA. 2007. Mammalogy: adaptation, diversity, ecology. Baltimore: JHU Press.

Frisina MR, Awan GA, Woodford MH. 2007. Determining trophy harvest quotas through a status survey of urial (Ovis orientalis) in the Kalabagh Game Reserve, Punjab province, Pakistan. J Bombay Nat Hist Soc. 104(1):35.

Groves C, Grubb P. 2011. Ungulate taxonomy. Baltimore: JHU Press.

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In Nucleic acids symposium series (Vol. 41, No. 41, pp. 95–98). London: Information Retrieval Ltd.; p. c1997–c2000.

Hassanin A, Des luc F, Ropiquet A, Hammer C, Jansen van Vuuren B, Matthee C, Ruiz-Garcia M, Catzeflis F, Areskoug V, Nguyen TT, et al. 2012. Pattern and timing of diversification of Cetartiodactyla (Mammalia, Laurasiatheria), as revealed by a comprehensive analysis of mitochondrial genomes. CR Biol. 335(1):32–50.

Jukes TH, Cantor CR. 1969. Evolution of protein molecules. In: Munro HN, editor. Mammalian protein metabolism. New York (NY): Academic Press; p. 21–132.

Khan FM, William K, Aruge S, Janjua S, Shah SA. 2018. Illegal product manufacturing and exportation from Pakistan: revealing the factuality of highly processed wildlife skin samples via DNA mini-barcoding. Nucleosides Nucleotides Nucleic Acids. 37(3):179–185.

Kumar V, Sharma N, Sharma A. 2017. DNA barcoding of the Indian blackbuck (Antilope cervicapra) and their correlation with other closely related species. Egypt J Forensic Sci. 7(1):31.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 35(6):1547–1549.

Lorenz JG, Jackson WE, Beck JC, Hanner R. 2005. The problems and promise of DNA barcodes for species diagnosis of primate biomaterials. Phil Trans R Soc B. 360(1462):1869–1877.

Mirza ZB. 1998. Animal biodiversity of Pakistan. Pakistan: CERC.

Molur S. 2003. Status and Red List of Pakistan’s mammals. Pakistan mammal conservation assessment & management plan workshop, 18–22 August 2003.

Roberts TJ. 1997. The mammals of Pakistan. London: Ernest Benn.

Sambrook J, Russell DW, Janssen K, Argentine J. 2001. Molecular cloning: a laboratory manual on the web. New York (NY): Cold Spring Harbor Laboratory.

Schaller GB, Khan SA. 1975. Distribution and status of markhor (Capra falconeri). Biol Cons. 7(3):185–198.

Schaller GB. 1980. Stones of silence. London: Andre Deutsch Limited.

Shukla M, Joshi BD, Kumar VP, Thakur M, Mehta AK, Sathyakumar S, Goyal SP. 2019. Species dilemma of musk deer (Moschus spp) in India: molecular data on cytochrome c oxidase I suggests distinct genetic lineage in Uttarakhand compared to other Moschus species. Anim Biotechnol. 30(3):193–201.

Siraj-Ud-Din M, Minhas RA, Ali U, Khan M, Awan MS, Shafi N, Ahmad B. 2018. Conservation status of Ladakh Urrial (Ovis vignei vignei Blyth, 1841) in Gilgit Baltistan, Pakistan. PIZ. 50(1):1–400.

Sultana S, Mannen H, Tsuji S. 2003. Mitochondrial DNA diversity of Pakistani goats. Animal Genetics. 34(6):417–421. doi:10.1046/j.0306-9146.2003.01040.x.

Vaughan T, Ryan J, Czapelewski N. 2011. Mammalogy. London: Jones & Bartlett Learning.

Wang Q, Yang C. 2013. The phylogeny of the Cetartiodactyla based on complete mitochondrial genomes. JUB. 5(3):30.

Yang C, Xiao Z, Zou Y, Zhang X, Yang B, Hao Y, Meermont T, Yue B. 2015. DNA barcoding revises a misidentification on musk deer. Mitochondrial DNA. 26(4):605–612.

Zhong W, Wang F, Li B, Jiang W, Yan H. 2017. Research on the non-directional test in meat adulteration based on DNA barcode. J Food Saf Qual. 8(5):1547–1551.