DNA barcoding of *Oryx leucoryx* using the mitochondrial cytochrome C oxidase gene

K. Elmeer[1], A. Almalki[1], K.A. Mohran[2], K.N. AL-Qahtani[2] and M. Almarri[1]

[1]Genetic Engineering Department, Biotechnology Centre,
Ministry of the Environment, Doha, Qatar
[2]Molecular Biology Laboratory, Veterinary and Research Laboratory Section,
Ministry of the Environment, Doha, Qatar

Corresponding author: K. Elmeer
E-mail: elmeer@gmail.com

ABSTRACT. The massive destruction and deterioration of the habitat of *Oryx leucoryx* and illegal hunting have decimated *Oryx* populations significantly, and now these animals are almost extinct in the wild. Molecular analyses can significantly contribute to captive breeding and reintroduction strategies for the conservation of this endangered animal. A representative 32 identical sequences used for species identification through BOLD and GenBank/NCBI showed maximum homology 96.06% with *O. dammah*, which is a species of *Oryx* from Northern Africa, the next closest species 94.33% was *O. gazella*, the African antelope. DNA barcode sequences of the mitochondrial cytochrome C oxidase (COI) gene were determined for *O. leucoryx*; identification through BOLD could only recognize the genus correctly, whereas the species could not be identified. This was due to a lack of sequence data for *O. leucoryx* on BOLD. Similarly, BLAST analysis of the NCBI data base also revealed no COI sequence data for the genus *Oryx*.

Key words: Arabian oryx; White oryx; mtDNA; COI sequences

INTRODUCTION

The white oryx (*Oryx leucoryx*) or Arabian oryx is endemic to the Arabian Peninsula. It is the largest of the antelopes that once grazed the plains and deserts of the region and is uniquely adapted to the extremely arid environment. The Arabian oryx was extirpated in the wild by hunting in the early 1970s (Henderson, 1974) and classified as endangered on the International Union for Conservation of Nature (IUCN, 2010) Red List. It has been listed in Appendix 1 of the Convention on International Trade in Endangered Species since 1975. A global perspective about developing more effective captive breeding programs is necessary to maintain the genetic diversity required to save this endangered species (Iyengar et al., 2007).

Because captive breeding plays an important role in the conservation of threatened species, several *Oryx* breeds have been successfully retained in captivity after extinction in the wild. Mesochina et al. (2003) have built a captive *Oryx* population recognized as the most polymorphic of all captive herds, suggesting that no recent management-related bottleneck has occurred. Genetic analysis has suggested, however, that as much as half of the neutral genetic variation present in the pre-extinction population of the Arabian oryx may be absent from contemporary populations (Marshall et al., 1999).

The use of molecular approaches can contribute significantly to captive breeding and reintroduction strategies for the conservation of various endangered animals such as the *Oryx* (Russello and Amato, 2007). Mitochondrial DNA (mtDNA) is regarded as an important tool in the study of evolutionary relationships among various taxa owing to its conserved protein-coding regions, high variability in non-coding sequences, and lack of recombination (Olivo et al., 1983; Ingman et al., 2000). Sequence divergence accumulates more rapidly in mtDNA than in nuclear DNA owing to a faster mutation rate and lack of repair system, meaning that it often contains high levels of informative variation (Khan et al., 2008).

DNA barcoding has become a promising tool for the rapid and accurate identification of various taxa, and it has been used to reveal unrecognized species in several animal groups. Animal DNA barcodes (600- to 800-bp segments) of the mitochondrial cytochrome oxidase I (COI) gene have been proposed as a means to quantify global biodiversity. DNA barcoding has the potential to improve the way researchers relate to wild biodiversity (Janzen et al., 2005). Moreover, the introduction of DNA barcoding has highlighted the expanding use of COI as a genetic marker for species identification (Dawnay et al., 2007).

DNA barcodes consist of a standardized short sequence of DNA between 400 and 800 bp that can easily be generated and characterized for all species on the planet (Savolainen et al., 2005). These genetic barcodes can be stored in an open-access digital library that can be used to compare the DNA barcode sequences of unidentified samples from the field, garden, or market by matching them to known sequences with associated species names in the database. The Consortium for the Barcode of Life (http://www.barcoding.si.edu/) is charged with coordinating barcoding activities around the world and promoting a database of documented and vouchered reference sequences to serve as a universal DNA barcode library for all life (John Kress and Erickson, 2008).

DNA barcoding allows users to recognize known species and retrieve information about them quickly and cheaply. It may also speed the discovery of the thousands of species yet to be named. Barcoding, if developed sufficiently, will be a vital new tool for appreciating and managing the immense and changing biodiversity on earth (Cowan et al., 2006).

A DNA barcode, in its simplest definition, is one or more short gene sequences taken from a standardized portion of the genome used to identify species. The use of such short DNA

sequences for biological identification with the ultimate goal of quick and reliable species-level identifications applies to all forms of life, including animals, plants, and microorganisms. Currently, the concept of a universally recoverable segment of DNA that can be applied as an identification marker across species has been most successfully applied to animals (Hebert et al., 2004). At a minimum, three criteria must be met to identify a gene region as appropriate for a DNA barcode: 1) significant species-level genetic variability and divergence; 2) short sequence length to facilitate DNA extraction and amplification, and 3) universal PCR primers. For most groups of animals, a portion of the mitochondrial gene for COI has been identified as a species-level barcode. COI has been shown to fit the three criteria in the great majority of animal taxa to which it has been applied (John Kress and Erickson, 2008).

In this study, the DNA barcoding for white oryx (*O. leucoryx*) was carried out using the COI gene. The DNA barcode was determined for the species using blood samples from 32 individuals. The barcode was compared with the Barcode of Life Data Systems (BOLD; www.barcodinglife.org) and GenBank/National Center for Biotechnology Information (NCBI; www.ncbi.nlm.nih.gov) databases. A phylogenetic relationship with some closely related species was constructed using the DNA barcode sequence.

MATERIAL AND METHODS

Blood samples from 32 male and female *Oryx* were collected from three conserved locations in the State of Qatar (Um Qarn Park, Mazhabia Park, and USharige Park). DNA was purified from 200 µL ethylenediaminetetraacetic acid-anticoagulated blood with a QIAamp Blood Mini Kit (Qiagen, Basel, Switzerland). The isolated DNA was quantified and qualified using a NanoDrop® ND-1000 spectrophotometer. For further estimation of DNA quantity, 2 µL was loaded onto 0.85% agarose gel at 100 V for 30 min. The gels were stained in ethidium bromide and visualized under ultraviolet light.

PCR amplification of a COI gene fragment using the Universal Animal Barcoding primer recommended by the Consortium for Barcoding of Life was carried out according to the procedure of Folmer et al. (1994). The following sequence was used: BLCO1490F 5'-GGT CAA CAA ATC ATA AAG ATA TTG G-3' and BHCO2198R 5'-TAA ACT TCA GGG TGA CCA AAA AAT CA-3'.

PCR was performed in a total reaction mixture of 20 µL containing 1 µL (5 ng) DNA template, 10 µL AmpliTaq Gold® 360 Mastermix (Applied Biosystems), 0.25 µL (10 pmol/µL) each of the forward and reverse primers, and 8.5 µL nuclease-free water. Amplification was carried out in a Veriti 96-Well Fast Thermal Cycler (Applied Biosystems) according to the procedure of Folmer et al. (1994), which consisted of an initial denaturation at 95°C for 10 min followed by 35 cycles of denaturation, annealing, and chain extension at 95°C (60 s), 40°C (60 s), and 72°C (90 s), respectively. The final chain extension step was 7 min at 72°C, and a final hold was carried out at 4°C. The PCR amplifications were visualized on 2.0% agarose gel using ethidium bromide staining, and the images were captured using a gel documentation system. The PCR products were then purified using ExoSap-IT.

DNA sequencing was carried out with forward as well as reverse primers of the universal primer according to a standard protocol provided with a Big Dye Terminator Kit® V 3.1 (Applied Biosystems) using an ABI 3130 genetic analyzer. One microliter cleaned PCR product was used for each 10-µL reaction. The DNA sequence data were analyzed and edited with

ABI Sequencing Analysis V 5.2. Sequence edition; multiple sequence alignment and study of intraspecific variation were carried out with Molecular Genetics Evolution Analysis (MEGA) version 4 (Tamura et al., 2007).

The species were identified from the representative DNA sequence of the DNA samples using the BOLD search engine. In addition, species identification was performed from the DNA sequence using the Basic Local Alignment Search Tool (BLAST) of GenBank/NCBI. Based on the BOLD identification engine and BLAST analysis, COI sequences of the 20 species most closely related to *Oryx* were obtained from GenBank/NCBI. These sequences were aligned and compared with COI sequences generated for the 32 *Oryx* samples using MEGA4 (Tamura et al, 2007). Phylogenetic and molecular evolutionary analyses were carried out with MEGA. The distance matrix was calculated using the Kimura 2-parameter, and the neighbor-joining tree was plotted using the Kimura 2-parameter.

RESULTS AND DISCUSSION

DNA from the 32 samples visualized with 2.0% agarose gel electrophoresis and ethidium bromide staining was successfully amplified using a standard protocol (Figures 1 and 2). All samples were successfully sequenced using the forward and reverse primers to obtain robust forward and reverse sequences of approximately 687 bp. The introduction of DNA barcoding has highlighted the expanding use of COI as a genetic marker for species identification (Dawnay et al., 2007).
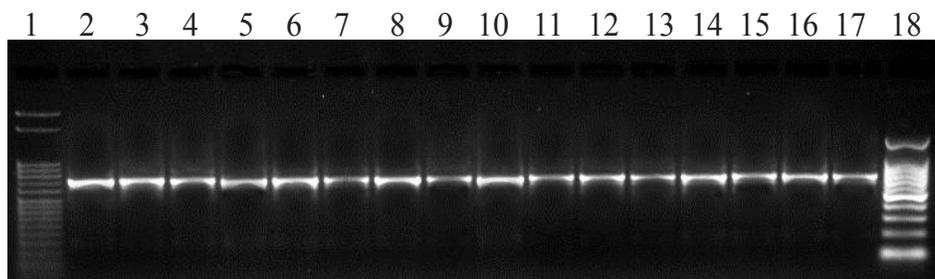
Figure 1. Amplified DNA product visualized by ethidium bromide staining. *Lanes 2-17* = female *Oryx* samples; *lane 1* = 50-bp ladder; *lane 18* = 100-bp ladder.
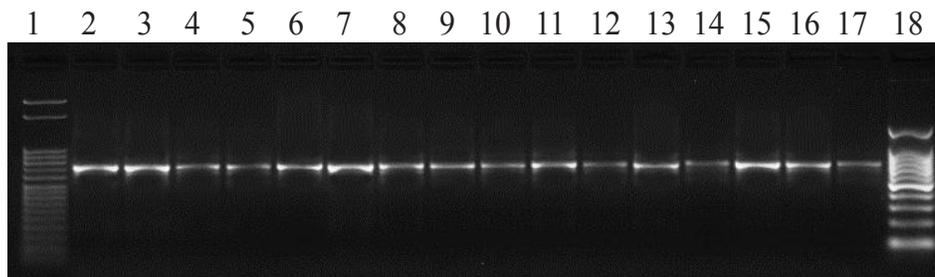
Figure 2. Amplified DNA product visualized by ethidium bromide staining. *Lanes 2-17* = male *Oryx* samples; *lane 1* = 50-bp ladder; *lane 18* = 100-bp ladder.

The multiple sequence alignment from the 32 samples showed no intraspecific variation on the similarity matrix based on the pairwise analysis obtained via a bootstrap procedure (1000 replicates). Analyses were conducted in MEGA4, which showed that no single difference existed between the 32 sequences of *O. leucoryx* given as follows:

CCCGAGATATCTTTTCTATATTCGGTGCTTGAGCTGGCATAGTGGGGACC
GCCCTAGGCGTACTAATTCGCGCTGAATTGGGTCAACCTGGAACTTTACTTGGA
GATGATCAAATCTACAACGTAGTCGTAACCGCACATGCATTCGTAATAATTTTCTT
TATAGTGATACCTATTATGATTGGAGGGTTTGGCAACTGACTAGTCCCTCTA
ATAATTGGAGCCCCCGACATAGCATTCCCTCGAATAAACAATATAAGCTTTT
GACTACTTCCTCCTTCTTTTCTACTACTCCTAGCATCTTCTATAGCTGAGGCTG
GAGCCGGAACAGGTTGAACTGTATATCCCCCTCTAGCTGGCAACCTAGCCCAT
GCAGGAGCCTCAATAGATCTCACTATTTTCTCTATACACTTAGGAGGTGTTTCCT
CAATTCTAGGAGCCATCAATTTTATCACAACAATCATTAACATAAAACCCCCTG
CAATAACACTATATCAAACTCCCTTGTTTGTATGATCTGTACTAATTACTGCTGTT
TACTACTCCTTTCACTCCCTGTGTTAGCAGCCGGCATTACAATATTAATTACAGATC
GAAACCTAAATACAACCTTCTTTGACCCGGCAGGAGGGGGGGACCCTATCT
TATATCAACATCTGTACTGATTGTGGCACCCCCGTGAGACTAA.

An ideal DNA barcode should allow fast, reliable, automatable, and cost-effective species identification by users with little or no taxonomic experience (Hebert et al., 2003; Hebert and Gregory, 2005). Representative 32 identical sequences were used for species identification through BOLD and GenBank/NCBI. Sequence identification through the BOLD Identification Engine revealed that the sequence showed maximum homology (96.06%) with *Oryx dammah* (Table 1), a species of *Oryx* from northern Africa that has been declared extinct in the wild by the IUCN (1998). Identifications are usually made by comparing unknown sequences against the DNA barcodes of known species via distance-based tree construction (Hebert et al., 2003, 2004), alignment searching (e.g., BLAST; Altschul et al., 1990, 1997), or recently proposed methods such as the characteristic attribute organization system (Kelly et al., 2007), decision theory (Abdo and Golding, 2007), and the back-propagation (BP) neural network (BP-based species identification; Zhang et al., 2008).

Table 1. Match statistics (top 20 matches) for *Oryx leucoryx* sequence generated through BOLD.

Phylum	Class	Order	Family	Genus	Species	Specimen similarity (%)
Chordata	Mammalia	Artiodactyla	Bovidae	*Oryx*	*dammah*	96.06
Chordata	Mammalia	Artiodactyla	Bovidae	*Oryx*	*gazella*	94.33
Chordata	Mammalia	Artiodactyla	Bovidae	*Capra*	*hircus*	87.56
Chordata	Mammalia	Artiodactyla	Bovidae	*Capra*	*hircus*	87.40
Chordata	Mammalia	Artiodactyla	Bovidae	*Capra*	*hircus*	87.40
Chordata	Mammalia	Artiodactyla	Bovidae	*Capra*	*hircus*	87.40
Chordata	Mammalia	Artiodactyla	Bovidae	*Ammotragus*	*lervia*	87.40
Chordata	Mammalia	Artiodactyla	Bovidae	*Ovis*	*dalli*	87.09
Chordata	Mammalia	Artiodactyla	Bovidae	*Ovis*	*dalli*	87.06
Chordata	Mammalia	Artiodactyla	Bovidae	*Ovis*	*canadensis*	86.77
Chordata	Mammalia	Artiodactyla	Bovidae	*Ovis*	*canadensis*	86.75
Chordata	Mammalia	Artiodactyla	Bovidae	*Cephalophus*	*monticola*	86.36
Chordata	Mammalia	Artiodactyla	Bovidae	*Cephalophus*	*monticola*	86.33
Chordata	Mammalia	Artiodactyla	Bovidae	*Cephalophus*	*monticola*	86.33
Chordata	Mammalia	Artiodactyla	Bovidae	*Cephalophus*	*monticola*	86.31
Chordata	Mammalia	Artiodactyla	Bovidae	*Ovis*	*canadensis*	86.21
Chordata	Mammalia	Artiodactyla	Bovidae	*Cephalophus*	*monticola*	86.20
Chordata	Mammalia	Artiodactyla	Bovidae	*Cephalophus*	*monticola*	86.20
Chordata	Mammalia	Artiodactyla	Bovidae	*Cephalophus*	*monticola*	86.20
Chordata	Mammalia	Artiodactyla	Bovidae	*Cephalophus*	*monticola*	86.17

        The next closest species (94.33%; Table 1) was *O. gazella* or African antelope (gems-buck). Iyengar et al. (2006) performed a comparative study of control region (CR) sequences from several captive *Oryx* species and proposed a close grouping of *O. leucoryx* with *O. gazelle* instead of *O. dammah*. The match statistics for *Oryx* samples as derived from BOLD are given in Figure 3. The identification tree generated through BOLD is shown in Figure 4. The low similarity values shown by sequences available in BOLD reveal that no sequences for this species are yet available in the database.
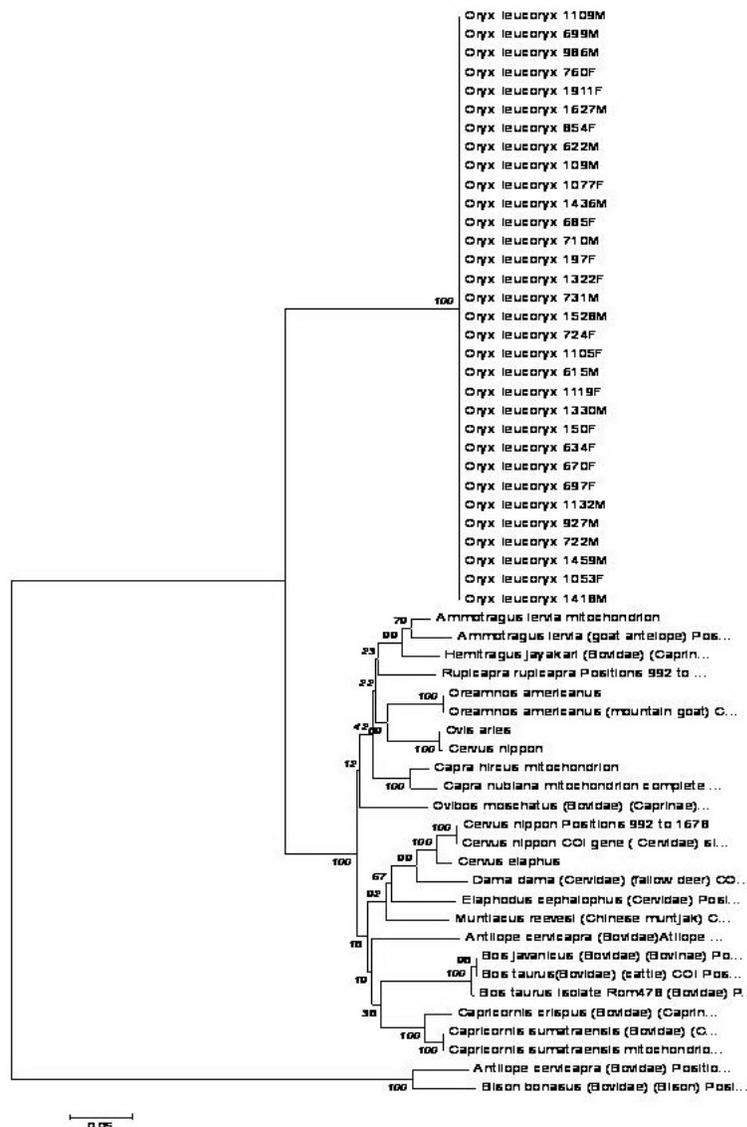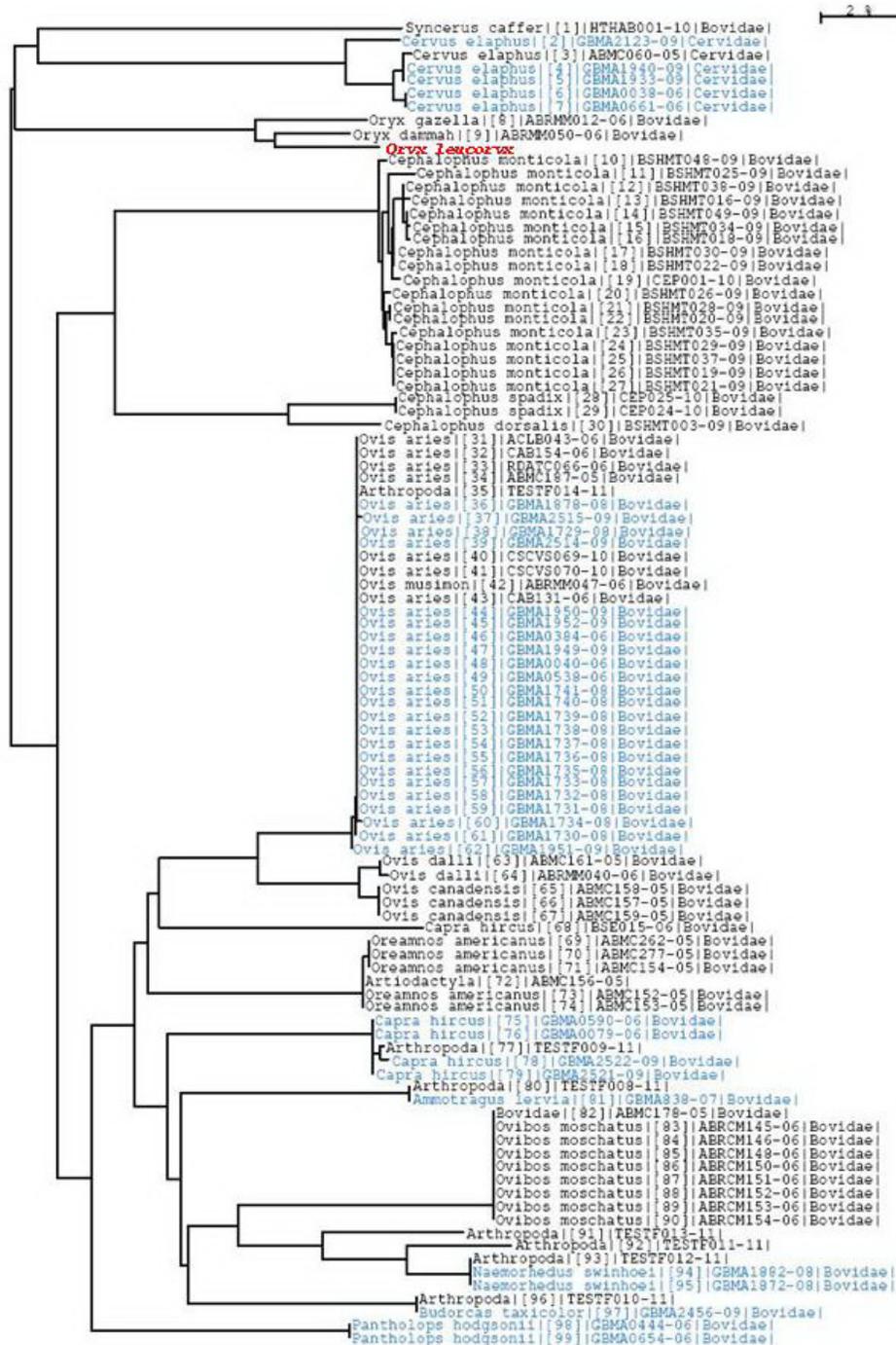

Figure 3. Neighbor joining tree as generated through MEGA4.

Figure 4. Identification tree generated through BOLD (*Oryx leucoryx*).

A related study using three molecular markers, 16S ribosomal RNA (rRNA), cytochrome b, and a CR, the molecular phylogeny of *Oryx* species, including the Arabian oryx (*O. leucoryx*), scimitar-horned *Oryx* (*O. dammah*), and plains *Oryx* (*O. gazella*), has indicated that 16S rRNA and cytochrome b produced similar phylogeny (*O. dammah* grouped with *O. gazella*), whereas the CR grouped *O. dammah* with *O. leucoryx* (Khan et al., 2008). Iyengar et al. (2006) have performed a comparative study of CR sequences from several captive *Oryx* species and proposed a close grouping of *O. leucoryx* with *O. gazella* instead of *O. dammah*.

Luo et al. (2011) have indicated that the 5'-end of COI, the standard barcoding region for animals, is not only representative of the entire COI gene but also the 12-mt PCGs, despite the fact that gene lengths ranged from the 216 bp of ATP8 to the 1866 bp of ND5. This finding is consistent with the conclusion of Roe and Sperling (2007) that subsections of the COI-COII region (~2.3 kb) perform similarly.

Sequence identification with the NCBI BLAST tool also revealed that no sequences for the COI gene are available for the genus *Oryx* in the database, and maximum homology was shown with *Capra hircus*, *Ammotragus lervia*, *Pseudois nayaur*, *Capra falconeri*, *Hemitragus jayakari*, *Ovis aries*, *Cephalophus monticola*, *Pantholops hodgsonii*, and others with an approximately 85 to 87% match. This result clearly reveals that no COI sequences are available for *Oryx* through NCBI. It was interesting to note that the species that showed maximum homology with *Oryx* samples thorough BLAST analysis are those from temperate zones and high altitude, especially *P. nayaur* and *P. hodgsonii*, which are found in India and China and at high altitudes in the Himalayas and the Tibetan plateau.

Min and Hickey (2007) have shown that the COI barcoding region provides a quick preview of mitochondrial genome composition. Luo et al. (2011) provide results from comparisons between the genome profile and the 13 individual gene regions indicating that the COI barcoding region is also representative of the efficacy of the mitochondrial genome as a whole of the twelve PCGs together.

DNA barcode sequences (COI) were successfully determined for *O. leucoryx*. Identification through BOLD could identify only the genus correctly. The species could not be identified owing to a lack of sequence data for *O. leucoryx* on BOLD. Hence, the database showed maximum homology with *O. dammah*, a species from northern Africa that has been declared extinct in the wild by IUCN (1998). Similarly, BLAST analysis through the NCBI database revealed no COI sequence data for the genus *Oryx*.

The sequence generated can be submitted to NCBI and BOLD such that a sequence database becomes available for future use. The availability of sequences for genes such as cytochrome b, 16S rRNA, and CRs for *O. leucoryx* on NCBI suggests that studies of these genes should be undertaken to ascertain the genetic variation among the various populations of *O. leucoryx* distributed across different countries. These surveys would help in establishing *O. leucoryx* from Qatar as a separate entity; however, an elaborate study based on cytochrome b, 16S rRNA, and CRs should be undertaken. These proposed studies can shed immense light on the phylogeny of *O. leucoryx* and help in developing a strategy for its long-term conservation.

REFERENCES

Abdo Z and Golding GB (2007). A step toward barcoding life: a model-based, decision-theoretic method to assign genes to preexisting species groups. *Syst. Biol.* 56: 44-56.
Altschul SF, Gish W, Miller W, Myers EW, et al. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215: 403-410.

Altschul SF, Madden TL, Schaffer AA, Zhang J, et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25: 3389-3402.

Cowan RS, Chase MW, Kress WJ and Savolainen V (2006). 300,000 species to identify: problems, progress, and prospects in DNA barcoding of land plants. *Taxon* 55: 611-616.

Dawnay N, Ogden R, McEwing R, Carvalho GR, et al. (2007). Validation of the barcoding gene COI for use in forensic genetic species identification. *Forensic Sci. Int.* 173: 1-6.

Folmer O, Black M, Hoeh W, Lutz R, et al. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol. Mar. Biol. Biotechnol.* 3: 294-299.

Hebert PD and Gregory TR (2005). The promise of DNA barcoding for taxonomy. *Syst. Biol.* 54: 852-859.

Hebert PD, Cywinska A, Ball SL and deWaard JR (2003). Biological identifications through DNA barcodes. *Proc. R. Soc. Lond. B Biol. Sci.* 270: 313-321.

Hebert PD, Penton EH, Burns JM, Janzen DH, et al. (2004). Ten species in one: DNA barcoding reveals cryptic species in the Neotropical skipper butterfly *Astraptes fulgerator*. *Proc. Natl. Acad. Sci. U. S. A.* 101: 14812-14817.

Henderson DS (1974). The Arabian oryx: a desert tragedy. *Nat. Parks Conserv. Magazine* 48: 15-21.

Ingman M, Kaessmann H, Paabo S and Gyllensten U (2000). Mitochondrial genome variation and the origin of modern humans. *Nature* 408: 708-713.

IUCN (1998). IUCN Guidelines for Re-Introduction. IUCN/SSC Re-Introduction Specialist Group, Gland, Switzerland and Cambridge.

IUCN (2010). Environment Agency - Abu Dhabi E.A.D, The Coordination Committee for the Conservation of the Arabian oryx (CCCAO) IUCN/SSC Antelope Specialist Group (ASG), Arabian oryx regional conservation strategy and action plan. Environ. Agency, Abu Dhabi.

Iyengar A, Diniz FM, Gilbert T, Woodfine T, et al. (2006). Structure and evolution of the mitochondrial control region in oryx. *Mol. Phylogenet. Evol.* 40: 305-314.

Iyengar A, Gilbert T, Woodfine T, Knowles JM, et al. (2007). Remnants of ancient genetic diversity preserved within captive groups of scimitar-horned oryx (*Oryx dammah*). *Mol. Ecol.* 16: 2436-2449.

Janzen DH, Hajibabaei M, Burns JM, Hallwachs W, et al. (2005). Wedding biodiversity inventory of a large and complex Lepidoptera fauna with DNA barcoding. *Phil. Trans. R. Soc. B* 360: 1835-1845.

John Kress W and Erickson DL (2008). DNA barcoding a windfall for tropical biology. *Biotropica* 40: 405-408.

Kelly RP, Sarkar IN, Eernisse DJ and Desalle R (2007). DNA barcoding using chitons (genus *Mopalia*). *Mol. Ecol. Notes* 7: 177-183.

Khan HA, Arif IA, Al Homaidan AA and Al Farhan AH (2008). Application of 16S rRNA, cytochrome b and control region sequences for understanding the phylogenetic relationships in *Oryx* species. *Genet. Mol. Res.* 7: 1392-1397.

Luo A, Zhang A, Ho SY, Xu W, et al. (2011). Potential efficacy of mitochondrial genes for animal DNA barcoding: a case study using eutherian mammals. *BMC Genomics* 12: 84.

Marshall TC, Sunnucks P, Spalton JA and Greth A (1999). Use of genetic data for conservation management: the case of the Arabian oryx. *Anim. Conserv.* 2: 269-278.

Mesochina P, Bedin E and Ostrowski S (2003). Reintroducing antelopes into arid areas: lessons learnt from the oryx in Saudi Arabia. *C. R. Biol.* 326 (Suppl 1): S158-S165.

Min XJ and Hickey DA (2007). DNA barcodes provide a quick preview of mitochondrial genome composition. *PLoS One* 2: e325.

Olivo PD, Van de Walle MJ, Laipis PJ and Hauswirth WW (1983). Nucleotide sequence evidence for rapid genotypic shifts in the bovine mitochondrial DNA D-loop. *Nature* 306: 400-402.

Roe AD and Sperling FA (2007). Patterns of evolution of mitochondrial cytochrome c oxidase I and II DNA and implications for DNA barcoding. *Mol. Phylogenet. Evol.* 44: 325-345.

Russello MA and Amato G (2007). On the horns of a dilemma: molecular approaches refine *ex situ* conservation in crisis. *Mol. Ecol.* 16: 2405-2406.

Savolainen V, Cowan RS, Vogler AP, Roderick GK, et al. (2005). Towards writing the encyclopedia of life: an introduction to DNA barcoding. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360: 1805-1811.

Tamura K, Dudley J, Nei M and Kumar S (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24: 1596-1599.

Zhang AB, Sikes DS, Muster C and Li SQ (2008). Inferring species membership using DNA sequences with back-propagation neural networks. *Syst. Biol.* 57: 202-215.