GLOBAL JOURNAL OF BIO-SCIENCE AND BIOTECHNOLOGY

www.scienceandnature.org

BARCODING PROFILING AND INTRA SPECIES VARIATION WITHIN THE BARCODE REGION OF TWO ESTUARINE FISHES *OREOCHROMIS MOSSAMBICUS* AND *OREOCHROMIS NILOTICUS*

G. Anbarasi*, R. Vignesh, M.P. Arulmoorthy, A.C. Rathiesh & M. Srinivasan
Centre of Advanced Study in Marine Biology, Faculty of Marine Sciences, Annamalai University, Parangipettai 608 502, Tamilnadu, India
Corresponding author's email: anbarasi.gk08@gmail.com

ABSTRACT
Classification of species is fundamental to research in biodiversity, ecology, evolutionary biology and conservation biology. As a part of this background in the present study, two fishes *O. mossambicus and O. niloticus* was sequenced for its 647bp and 645bp region of cytochrome oxidase subunit I (COI) gene to test its efficacy in identifying the species and also to demonstrate its intra species variations within the barcode region. The sequences were analyzed for their species identification using Barcode of life database (BOLD's) identification engine. The COI sequences of *O. mossambicus* and *O niloticus* from different geographical regions were extracted from NCBI for intraspecies variation analysis. While comparing the sequences in the NCBI database with the present sequences, significant alignments with maximum identity ranges of 99% to 100% was noticed. Both the sequences were aligned using Clustal W. Phylogenetic tree was constructed with bootstrap test. The optimal tree with the sum of branch length showed 0.000976 for *O. mossambicus* and 0.001556 for *O. niloticus.* In *O. mossambicus,* the maximum GC content was 46.9% whereas the GC content in all the other closely related species was found to be 46% in average. In *O. niloticus,* the maximum GC content was 47.9% and the closely related sequences exhibited 48% average GC content. The analyses revealed the molar concentration of bar-coded genes G+C content of the lion fish were analyzed and there was a little fluctuation in the G+C content of both the fishes from different geographical boundary. The evolutionary relationship was inferred with the closely related sequences obtained from the NCBI database. Sequence divergence between individuals of the same species ranged from 0 to 0.057. Both phylography and phylogeographic signals were evident from the phylogram constructed with *O. mossambicus and O. niloticus* as the same group. Even though the present results confirm that the COI could be a potential barcode gene for species level identification, further research is needed in describing the divergence of sequences in a broader sense.

KEYWORDS: Intra species, variations, barcode, cytochrome oxidase, phylogenetic.

INTRODUCTION

Fish are the largest group of vertebrates, which exhibit a remarkable diversity of morphological attributes and biological adaptations[1,2]. Species are typically circumscribed based on the presence of fixed diagnostic morphological characters which distinguish them from other species[3]. But for fishes, there are a large number of intraspecific invariants or interspecific overlapping's, so fish identification is challenging for taxonomists when facing rich biotas. The limitations inherent in morphology-based identification systems and the dwindling pool of taxonomists call for the molecular approach to species recognition [4]. Freshwater fishes show more population differentiation than marine species, although marine species can show significant differentiation [5]. Indeed, several studies have already illustrated the advances provided by the iterative processes between morphological- and DNA barcode-based studies in fishes [6-8]. DNA barcode, a short section of DNA sequence is used to identify species. Neither the idea nor the technology behind DNA Barcoding is novel. What is new and controversial is the idea of using just a small portion of a single gene to identify species from a wide taxonomic range [9]. Hebert *et al* [10] introduced the concept of a DNA barcode, and proposed a new loom to species identification which offered greater promise to counter many of the limitations. The new approach is based on the ground that the sequence analysis of a short fragment of a single gene "Cytochrome C oxidase subunit 1" enables unambiguous identification of all animals' species. Hebert *et al* [10] suggested a 650 base pair sequence of mitochondrial gene Cytochrome C oxidase subunit 1 (COI) as the reference DNA barcode for all animal life. This gene occurs in the mitochondria of all eukaryotic organisms and the initial studies revealed consistent resolving capability at the species level for many animals. DNA barcoding, which was advocated by Hebert *et al* [11, 12] seeks to facilitate identifying the increasing number of unfamiliar taxa in biological conservation and biodiversity surveys, based on sequence diversity within a short and standardized gene region[13]. For coordinating the collection data of specimens and performing data analysis with barcode data, the website Barcode of Life Data Systems (BOLD) (http://www.boldsystems.org) has been established [14].

Additionally to the mitochondrial COI gene, nuclear loci are sometimes also considered to improve assignment performance [15, 16]. This has been shown to provide species

level resolution of the vast bulk of species in a wide range of animal taxa, including ants, bats, birds, butterflies, crustaceans, fish, and spiders [17-21]. Hajibabaei *et al* [22] showed that 97.9% of 521 described species of Lepidoptera possess distinct DNA barcodes and furthermore that the few instances of sequence overlap of different species involve very similar ones. Present study concerns with barcoding of two cichlid fishes belonging to the *Oreochromis* family. The barcoding profiles are then used in investigating the intra species variation within the species.

The efficiency of DNA barcoding has been reported in the detection and description of new cryptic species [23-25]. This identification tool can clearly give support to improve classifications and to critically examine the precision of morphological traits commonly used in taxonomy [26]. For a barcoding approach to succeed, within species DNA sequences need to be more similar to one another than those between species and recent studies confirmed that the majority of species examined are well delineated by a tight cluster of very similar sequences [27-31]. The methodology requires that intra-species DNA barcode variation is substantially less than interspecies variation, allowing accurate identification of individuals [32]. Present study explores the intra species variation within the barcode region of two estuarine fishes *Oreochromis mossambicus and Oreochromis niloticus* by using DNA barcoding and computational approaches.

MATERIALS & METHODS

Study species

Oreochromis mossambicus

The *Mozambique Tilapia* is a deep bodied cichlid fish native to the eastward-flowing rivers of central and southern Africa. Mozambique *Tilapia* generally prefer slow moving water bodies such as lagoons, rivers and impoundments, but can also colonise faster-flowing creeks and streams. In addition to fresh waters, this type of *Tilapia* can also live in habitats influenced by tides, such as the upper reaches of estuaries and coastal lagoons. *Mozambique Tilapia*s are often the most abundant species in disturbed habitats like urban drainages, since they can tolerate a broad range of conditions. *Mozambique Tilapia* is thought to be one of the most salt-tolerant of all the *Tilapia* species, tolerating salt concentrations of 0–120 ppt.

Oreochromis niloticus

Nile Tilapia is a tropical species that prefers to live in shallow water. It exists in a variety of freshwater and brackish habitats. It is a Diurnal species. The species *O. niloticus* is a euryhaline species characterized by fast growth and resistance to pathogens. The lower and upper lethal temperatures for Nile *Tilapia* are 11-12 °C and 42 °C, respectively, while the preferred temperature ranges from 31 to 36 °C. It is an omnivorous grazer that feeds on phytoplankton, periphyton, aquatic plants, small invertebrates, benthic fauna, detritus and bacterial films associated with detritus. *Nile Tilapia* can filter feed by entrapping suspended particles, including phytoplankton and bacteria, on mucous in the buccal cavity, although its main source of nutrition is obtained by surface grazing on periphyton mats [33-36].

DNA extraction

DNA isolation Genomic DNA was extracted from the stored muscle tissue samples by the standard Proteinase-K/Phenol–Chloroform– ethanol method[37]. The concentration of DNA was estimated using a UV spectrophotometer. The COI gene located in the mitochondrial genome was amplified using two sets of primers [38] synthesized by Bio-Serve India, Ltd., Hyderabad, India.

MAB Fw:5’-TCAACCAACCACAAAGACATTGGCA C-3’
MAB Rw: 5’-TAGACTTCTGGGTGGCCAAAGAATCA-3

QIAGEN kit was used for sequencing reaction. The sequencing PCR was done to amplify one strand of barcode gene employing the primer FISH F1only under standard PCR conditions. The samples were precipitated and suspended in 40µl of loading solution provided with the kit. Sequencing was done with MegaBace sequencer-Bioserve India, Hyderabad.

BOLD’s identification

BOLD (Barcoding of life database) is an online workbench that aids in collection, management, analysis, and use of DNA barcodes. Identification engine is the one of the important components of BOLD database which consists of large volume of barcode sequences for both plants (intranuclear spacer gene) and animals (cytochrome c oxidase subunit gene). BOLD-IDS provide a species identification tool that accepts DNA sequences from the barcode region and returns a taxonomic assignment to the species level when possible. The BOLD identification system (IDS) accepts sequences from the 5' region of the mitochondrial gene cytochrome oxidase subunit I and returns species-level identification when one is possible. This identification engine was accessible online through http://www.barcodinglife.org/views/idrequest.php. The sequences were given in FASTA file format in the query box and results were obtained similar to that of BLAST search. This system of species identification was developed by Herbert *et al* [39].

Profiling of the barcode region

Barcode profiling is the phenomenon of extrapolating the barcode DNA sequence information such as molecular weight of the barcode strand and molar concentration of the individual nucleotides in the barcode region. Barcode profiling could be done using the software Bioedit developed [40]. The molecular weight of the single stranded barcode DNA was calculated as the sum of the monophosphate forms of each deoxyribonucleotide minus one water molecule each. One water (18 Da) was added at the end to represent the 3’ hydroxyl at the end of the chain and one more hydrogen atom at the 5’ phosphate end. Nucleotide composition summaries and plots were obtained by choosing “Nucleotide Composition” form the “Nucleic Acid” submenu of the “Sequence” menu.

Clustal W

ClustalW is a general purpose global multiple sequence alignment program for DNA or proteins. It produces biologically meaningful multiple sequence alignments of divergent sequences. It calculates the best match for the selected sequences, and lines them up so that the identities, similarities and differences can be seen. Evolutionary relationships can be seen via viewing Cladograms or Phylogram

Phylogenetic tree construction using MEGA

Phylogenetic relationships of genes or organisms are presented in a tree like form called phylogenetic tree. The branching pattern of a tree is called a topology. There are numerous methods for constructing phylogenetic trees from molecular data. Neighborhood joining (NJ) method of phylogenetic tree construction was preferred for accurate establishment of phylogenetic relationship and to trace out the presence of phylogenetic signals in the DNA sequences [41]. The distance was calculated between every pair of sequences and these were used to construct the phylogenetic tree which guided the final multiple alignment. The scores were calculated from separate pairwise alignments.

MEGA (Molecular Evolutionary Genetic Analysis)

MEGA is an integrated tool for conducting automatic and manual sequence alignment, inferring phylogenetic trees, mining the web base data bases, estimating the rates of molecular evolution, and testing evolutionary hypothesis [42].

Bootstrapping and Clustal W

One of the most commonly used tests of the reliability of an inferred tree is Felsenstein's [43] bootstrap test which is evaluated using Efron's [44] bootstrap resampling technique. If there are *m* sequences, each with *n* nucleotides (or codons or amino acids) a phylogenetic tree can be reconstructed using the same tree building method. The multiple aligned sequences from Clustal W were loaded into MEGA through Create New Alignment option in Alignment menu. The sequences were trimmed for their conserved regions and saved in MEGA format for phylogram construction. Bootstrap test for phylogeny was preferred to detect the reliability of each branch in phylogram. As a general rule if the bootstrap value for a given interior branch is 95% or higher than the topology of that branch then the value is considered "correct" [41].

RESULTS

Quantitation of DNA by electrophoresis

Quantitation of DNA was estimated in two different absorbances and the average was taken into consideration. The isolated genomic DNA of *O.mossambicus* and *O.niloticus* is shown in Fig 1. The PCR Products were separated by electrophoresis in 2% Agarose gels containing ethidium bromide (1µg ml^{-1}). Electropherogram obtained after electrophoresis of the PCR amplicons is shown in Fig 2. In the electropherogram, bands of the size ~642bp (MAB04-*O.mossambicus*), 645bp (MAB07-*O.niloticus*) was observed against 100bp DNA ladder. No overlapping of the bands in the case of test organisms was observed and so the bands were clear.

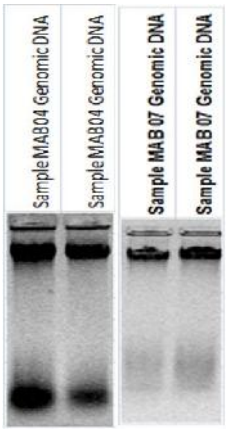

FIGURE 1: The isolated genomic DNA of *O.mossambicus (MAB04)* and *O. niloticus (MAB07)*

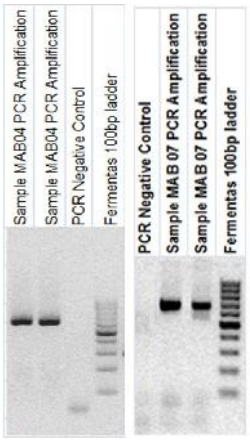

FIGURE 2: Electropherogram obtained after electrophoresis of the PCR amplicons

Top 10 Sequences Producing Significant Alignments from NCBI

The top sequences producing similar and significant alignments were identified for the study species from NCBI and listed. In *O. mossambicus*, 10 similar hits were found in NCBI with a maximum identity of 100% and with a maximum score of 1166 (Fig 3). For *O. niloticus*, the hits were found with a maximum identity of 99% with a score of 1155 (Fig 4). As a general rule, a top match with a sequence similarity of at least 98% was used as a criterion to designate potential species identifications [45]. The distance trees showed that, *O. mossambicus* (MAB04) was closely related to *O. niloticus*, *O. aureus* and *Perciformes* sp (Fig 5). Whereas, the distance tree for *O. niloticus*, showed close evolutionary relationship with *O. aureus* and similar relationship to bony fishes (Fig 6) similarly in NCBI.

Accession	Description	Max score	Total score	Query coverage	E value	Max ident
GU477631.1	Oreochromis sp. 'red tilapia' mitochondrion, complete genome	1165	1165	100%	0.0	100%
GU477628.1	Oreochromis niloticus strain 'American' mitochondrion, complete genome	1165	1165	100%	0.0	100%
GU477626.1	Oreochromis niloticus strain 'Philippines' mitochondrion, complete genome	1165	1165	100%	0.0	100%
GU477625.1	Oreochromis niloticus strain 'Egypt' mitochondrion, complete genome	1165	1165	100%	0.0	100%
EU732146.1	Oreochromis niloticus voucher OREONILO-I96-002 'BERT' cytochrome	1165	1165	100%	0.0	100%
GU673930.1	Perciformes sp. BOLD:AAA6538 voucher BW-A6975 cytochrome oxi	1165	1165	100%	0.0	100%
GU370126.1	Oreochromis niloticus mitochondrion, complete genome	1165	1165	100%	0.0	100%
EU417781.1	Oreochromis mossambicus voucher WL-F30 cytochrome c oxidase s	1165	1165	100%	0.0	100%
EU417782.1	Oreochromis mossambicus voucher WL-F31 cytochrome c oxidase s	1165	1165	100%	0.0	100%
DQ426567.1	Oreochromis niloticus cytochrome oxidase subunit I (COI) gene, pa	1165	1165	100%	0.0	100%

FIGURE 3: Top 10 Sequences Producing Significant Alignments *O.mossambicus*

Accession	Description	Max score	Total score	Query coverage	E value	Max ident
EU751833.1	Oreochromis niloticus voucher MX0013 cytochrome oxidase subunit 1	1155	1155	100%	0.0	99%
EU751831.1	Oreochromis niloticus voucher MX0014 cytochrome oxicase subunit 1	1155	1155	100%	0.0	99%
GU477527.1	Oreochromis niloticus strain Guangdong mitochondrion, complete gen	1146	1146	100%	0.0	99%
HM882785.1	Oreochromis niloticus voucher BNF 101 cytochrome oxidase subunit 1	1146	1146	100%	0.0	99%
HM882787.1	Oreochromis niloticus voucher BNF 103 cytochrome oxidase subunit 1	1146	1146	100%	0.0	99%
GU477530.1	Oreochromis aureus strain 'America' mitochondrion, complete genome	1146	1146	100%	0.0	99%
GU477529.1	Oreochromis aureus strain 'Guangdong' mitochondrion, complete genc	1146	1146	100%	0.0	99%
GU370125.1	Oreochromis aureus mitochondrion, complete genome	1146	1146	100%	0.0	99%
EU751832.1	Oreochromis niloticus voucher MX0010 cytochrome oxicase subunit 1	1146	1146	100%	0.0	99%
EU751830.1	Oreochromis niloticus voucher MX0012 cytochrome oxicase subunit 1	1146	1146	100%	0.0	99%
DQ426565.1	Oreochromis aureus cytochrome oxidase subunit I (COI) gene, partia	1141	1141	100%	0.0	99%

FIGURE 4: Top 10 Sequences Producing Significant Alignments *O.niloticus*

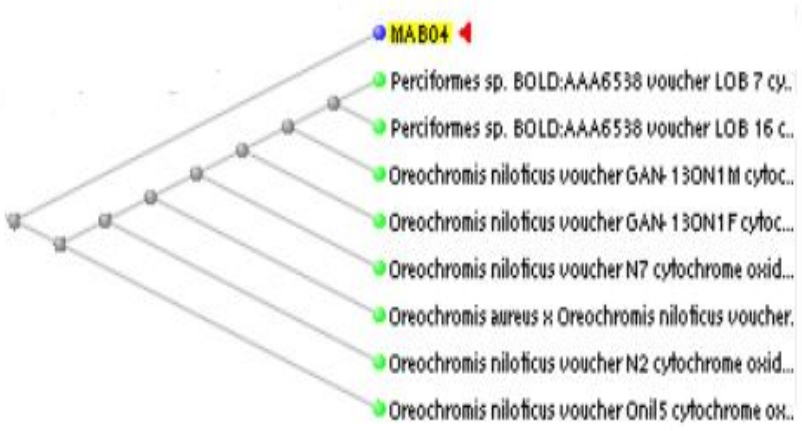

FIGURE 5: The distance trees *O.mossambicus*

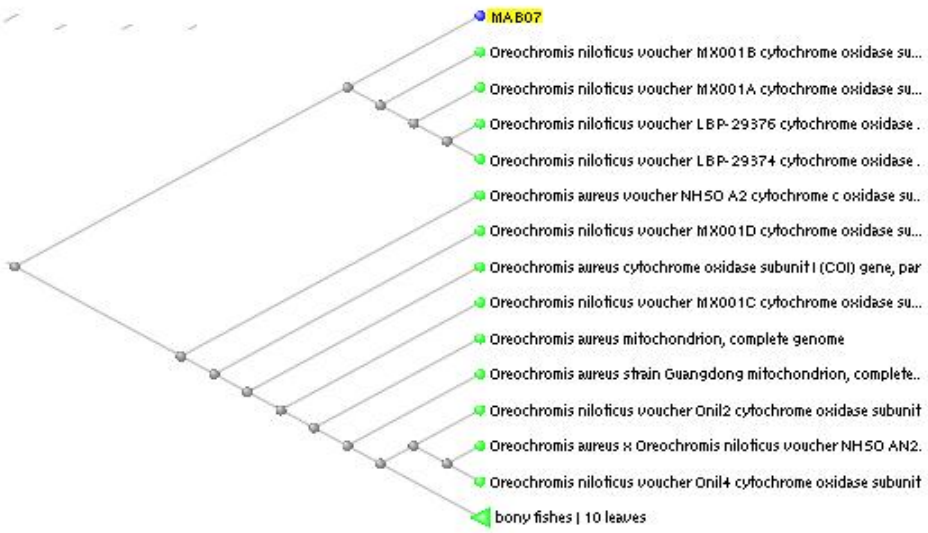

FIGURE 6: The distance trees *O.niloticus*

BOLD's search

For coordinating the collection data of specimens and performing data analysis with barcode data, the website Barcode of Life Data Systems (BOLD) (http://www.boldsystems.org) has been established [14]. BOLD is an accessible database that aids in management, analysis, dissemination, and searching of DNA barcodes.

BOLD's search results showed probability placement percentage of 100% identification for *O.mossambicus* (Fig 7) *and O.niloticus* (Fig 8).The COI species database tree for similarity and close relationship of the study animals, *O.mossambicus* (Fig 9) and *O.niloticus* (Fig 10) was compared.

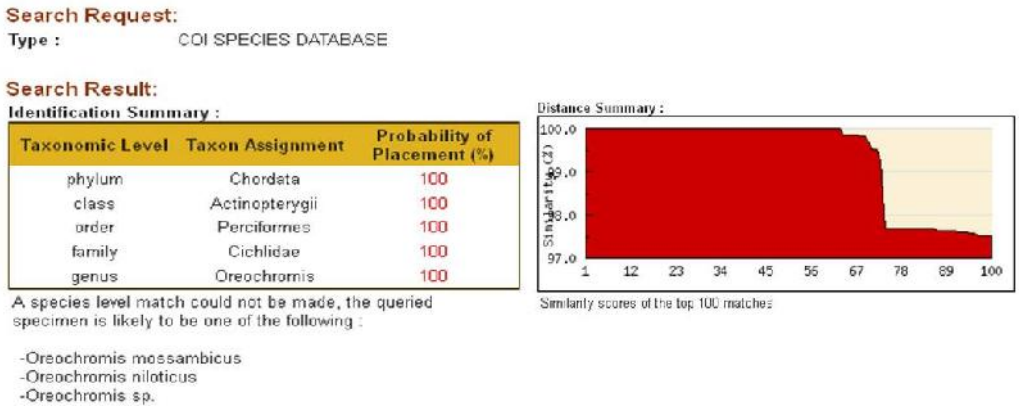

FIGURE 7: BOLD's search *O.mossambicus*

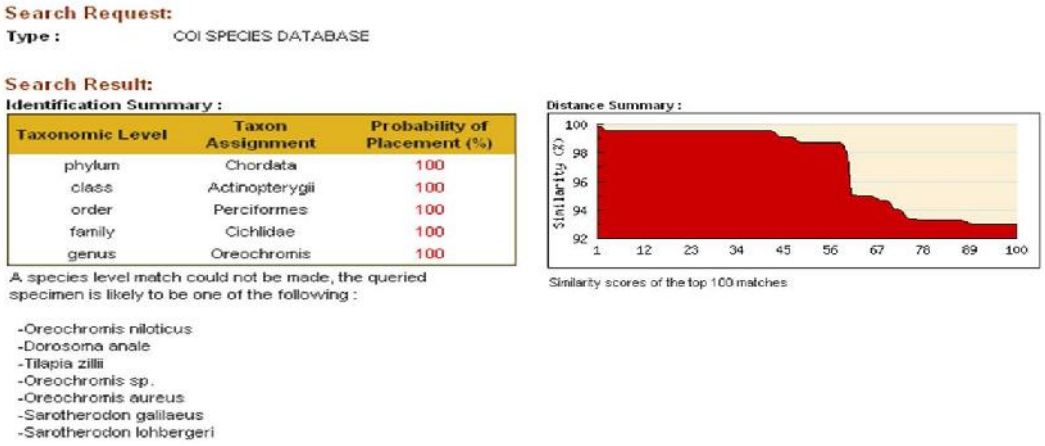

FIGURE 8: BOLD's search *O.niloticus*

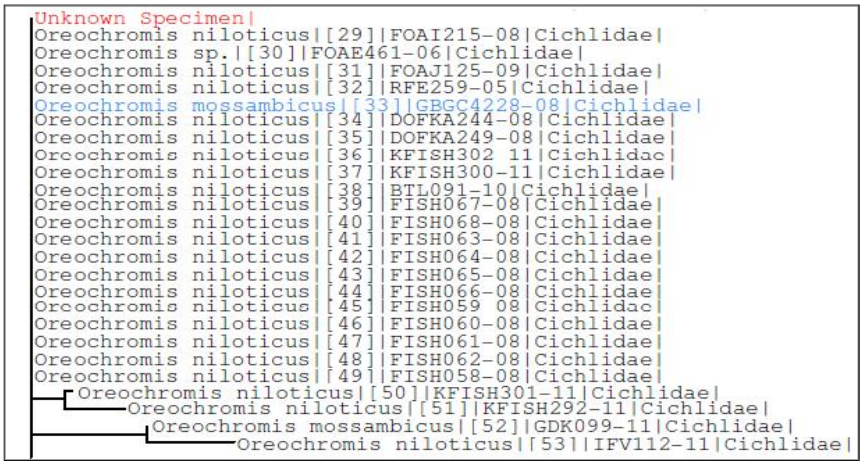

FIGURE 9: COI database tree for *O.mossambicus*

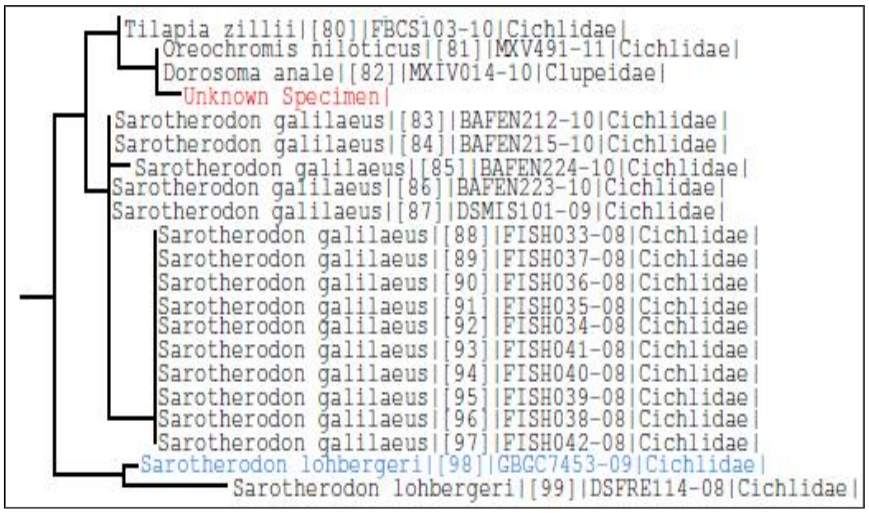

FIGURE 10: COI database tree for *O.niloticus*

Accession numbers of sequences closely related to the study species

The CO1 nucleotide sequences of the four animals were submitted in GENBANK for obtaining accession numbers. The accession numbers were received in two weeks after successful submission of the sequences. Accession numbers of sequences closely related to the test organisms used in the analysis & their locations is given in Table 1.

TABLE 1: Accession numbers of study species and sequences closely related to the Test organism used in the analysis & their locations.

Oreochromis mossambicus		
S.No.	**Accession No.**	**Country**
1	**JX173758 (MAB04)**	**India**
2	EU752146	USA
3	GU673990	Indonesia
4	EU417781	India
5	EU4177812	India
O.niloticus		
1	**JX173759(MAB07)**	**India**
2	EU751883	Mexico
3	HM882785	Nigeria
4	GU477627	China
5	EU751880	Mexico

Testing Evolutionary relationships of taxa

The close phylogenetic relationships were found within the samples. This is clear evidenced in both Multiple Sequence Alignments and Molecular Phylogenetic analysis by Maximum Likelihood and UPGMA methods. The alignment was easy and clear because no gaps were found in our barcode sequences. Read lengths were about 642 bp, 645bp, 647bp and 481bp long, although in some

instances some base calls were uncertain. No insertions, deletions or stop codons were observed in any sequence. Lack of stop codons is consistent with all amplified. The evolutionary history was inferred using the "Unweighted pair group method with arithmetic mean" [46]. The optimal tree with the sum of branch length showed 0.000976 - *O.mossambicus* (Fig 11), 0.001556-*O.niloticus* (Fig 12). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method [47] and are in the units of the number of base substitutions per site. The analysis involved 4 nucleotide sequences. All positions containing gaps and missing data were eliminated. Evolutionary analyses were conducted in MEGA5 [48].

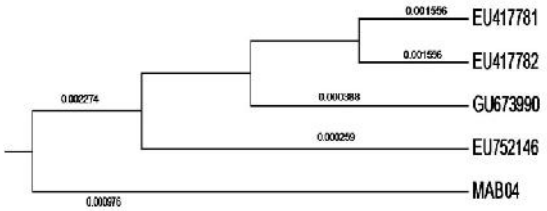

FIGURE 11: Inferred evolutionary history *O.mossambicus* by UPGMA method

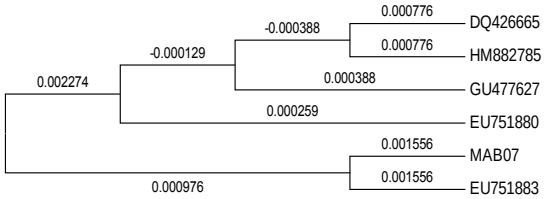

FIGURE 12: Inferred evolutionary history *O.niloticus* by UPGMA method

Profiling the barcode region of *O.mossambicus* and *O.niloticus*

In *O.mossambicus*, the maximum G+C content was 46.9% and A+T content was 53.1%. The molar concentration of ATGC was 24.5 %, 28.6 %, 17.2% and 29.7% (Table 2). In *O.niloticus*, the maximum G+C content was 47.9% and A+T content was 52.1%. The molar concentration of ATGC was 23.7 %, 28. 4 %, 17.8% and 30.1% (Table 3).

TABLE 2: Nucleotide composition of *O.mossambicus* & closely related sequences

Name of species	Accession ID	Base pair length	G+C content (%)	A+T content (%)	Nucleotide Number and Mol%			
					A	T	G	C
O.mossambicus	JX173758 MAB04	642	46.9%	53.1%	158 24.5%	185 28.6%	111 17.2%	192 29.7%
,,	EU752146	652	46.8%	53.2%	160 24.5%	187 28.7%	111 17.0%	194 29.8%
,,	GU673990	648	46.9%	53.1%	159 24.5%	185 28.5%	111 17.1%	193 29.8%
,,	EU417781	655	46.7%	53.3%	160 24.4%	189 28.9%	111 16.9%	195 29.8%
,,	EU417782	655	46.7%	53.3%	160 24.4%	189 28.9%	111 16.9%	195 29.8%

TABLE 3: Nucleotide composition of *O.niloticus* & closely related sequences

Name of species	Accession ID	Base pair length	G+C content (%)	A+T content (%)	Nucleotide Number and Mol%			
					A	T	G	C
O.niloticus	JX173759 MAB07	645	47.9%	52.1%	153 23.7%	183 28.4%	115 17.8%	194 30.1%
,,	EU751883	652	47.9%	52.1%	155 22.8%	185 28.4%	115 17.6%	197 30.2%
,,	HM882785	651	48.1%	51.9%	154 23.7%	184 28.3%	116 17.8%	197 30.3%
,,	GU477627	645	48.2%	51.8%	153 23.7%	181 28.1%	116 18.0%	195 30.2%
,,	EU751880	652	48.2%	51.8%	154 23.6%	184 28.2%	116 17.8%	198 30.4%

DISCUSSION

In the present study, the DNA extracted from the fish samples was successfully amplified with the COI primer cocktail. From these, the resulting PCR products were sequenced to obtain full length DNA barcodes with 645bp and 647 bp in length. Similarity in the sequences of these two species was compared with the sequences that are available in database and a phylogram was constructed. No insertions, deletions or stop codons were observed in any of the COI sequences, consistent with all amplified sequences being functional mitochondrial COI sequences. From the phylogram it is clear that across geography barcodes the two species did not contain variations within the same family based on the fact that barcodes of the same species invariably get clustered in same clade, Thus CO1 gene sequences can act serve as universal DNA markers in fish barcoding studies. The COI sequences of study species from different geographical regions were extracted from NCBI for intraspecies variation analysis. The top sequences producing similar and significant alignments were identified for the study species from NCBI and listed. For *O.mossambicus*, 10 similar hits were found in NCBI with a maximum identity of 100%. For *O.niloticus*, the hits were found with a maximum identity of 99%. From the previous reports it is known that for 248 generated COI sequences, maximum species identities in the range of 98–100% were obtained in GenBank and/or BOLD. For two specimens identified in BOLD as *Coryphaena equiselis* and *Etelis coruscans*, corresponding COI sequences were not available in GenBank to permit species identifications [49]. In the present study, the GC content of the mitochondrial CO1 regions was average in both species of *Oreochromis.* In *O.mossambicus*, (JX173758) the maximum GC content was 46.9% whereas the GC content in all the other closely related species (EU752146, GU673990, EU417781 and EU417782) was found to be 46% in average. In *O.niloticus*, the maximum GC content was 47.9% (JX173759) and the closely related sequences (EU751883, HM882785, GU477627, DQ426665 and EU751880) exhibited 48% average GC content. Saccone *et al* [50] reviewed data from the complete mitochondrial genomes of nine Osteichthyes and three Chondrichthyes species, deriving GC contents of 43.2% and 38.4%, respectively. These values correspond reasonably moderate to the present results. Akbar John *et al* [51] reported the average GC content of three fishes belonging to the Latidae family. The GC content was 8.87%. *Lates niloticus* from Tanzanian waters showed less GC content (47.70%) whereas *L. calcarifer* from South China Sea showed high GC content (49.61%) among the *Latidae* species studied. Comparing to Australian, Indian and Singapore (48.1-48.89%) water *L. calcarifer*, Myanmar and South China species showed high average GC content (49.38-49.41%). According to the literature, results of the present study showed the difference of 2% to 8% in the G+C content of the study species. The phylogenetic and genetic distance data showed that the maximum genetic distance is present in 3rd codon position in all the selected species. Prasanna Kumar *et al* [52] sequenced Mysis, post larvae and adult specimens of *Penaeus monodon* for Cyctochrome C Oxidase subunit I (*COI*) gene (*DNA* barcode) to check the efficiency of *DNA* barcode in delineating species irrespective of its different life stages. Although mtCOI has been shown to contain some phylogenetic information between closely related taxa [53,54], DNA barcodes generally lack sufficient phylogenetic signal at deeper levels [55]. Closer examination of intraspecific variation is useful to reveal cryptic species and analyze geographic distribution of lineages, or phylogeography [56]. Both phylography and phylogeographic signals were evident from the phylogram constructed with *O.mossambicus* and *O.niloticus* as the same group. Even though the present results confirm that the COI could be a potential barcode gene for species level identification, further research is needed in describing the divergence of sequences in a broader sense.

ACKNOWLEDGEMENT

The authors are thankful to the University Grants Commission (UGC), Department of Biotechnology (DBT) and authorities of Annamalai University for their constant support and encouragement.

REFERENCES

[1]. Eschmeyer, W.N., Ferraris, C.J., Hoang, D. and Long, D.J. (1998) Species of fishes. California Academy of Sciences, Catalog of Fishes. San Francisco, Part I. 25-1820.

[2]. Nelson, J.S., (2006) Fishes of the World. 4th Edition. John Wiley and Sons, Inc Newyork. pp. 601.

[3]. Wiens, J.J. and Servedio, M.R. (2000) Species delimitation in systematics: Inferring diagnostic differences between species. Proc R Soc Lond B. 267, 631-636.

[4]. Steinke, D., Zemlak,T.S., Boutillier,J.A. and Hebert,P.D.N.(2009a) DNA barcoding Pacific Canada's fishes. Mar Biol. 156, 2641-2647.

[5]. Ward, R.D., Woodwark, M. and Skibinski, D.O.F. (1994) A comparison of genetic diversity levels in marine, freshwater and anadromous fish. J Fish Biol. 44, 213-232.

[6]. Page, L.M., Beaman, R. F., Funk, M., Jeffords, D., Lipscomb, M., Mares, S., Noble, A., Prather, D., Stevenson, J. and Wheeler, Q (2005) LINNE: Legacy Infrastructure Network for Natural Environments. Illinois Natural History Survey, Champaign, Illinois.

[7]. Carlini, R.G., Alonzo, E., Bellorin Font, E. and Weisinger, J.R. (2006) Apoptotic stress pathway activation mediated by iron on endothelial cells *in vitro*. Nephrol Dial Transplant. 21, 3055-3061.

[8]. van Velzen, R., Bakker, F.T. and van Loon, J.J.A. (2007) DNA barcoding reveals hidden species diversity in Cymothoe (Nymphalidae).Proceedings of the Netherlands Entomological Society Meeting. 18,95-103.

[9]. Meyer, C.P., Paulay, G. (2005) DNA barcoding: error rates based on comprehensive sampling. PloS Biology. 3, 2229–2238.

[10]. Hebert, P.D.N., Penton, E.H., Burns, J. M., Janzen, D. H. and Hallwachs, W. (2004) Ten species in one: DNA barcoding reveals crypticspecies in the neotropical skipper butterfly *Astrapes fulgerator*. Proc Natl Acad Sci U.S.A. 101, 14812-14817.

[11]. Hebert, P.D.N., Cywinska, A., Ball, S.L.and DeWaard, J.R. (2003a) Biological identifications through DNA barcodes. Proc R Soc Lond B Biol Sci. 270,313-321.

[12]. Hebert, P.D.N., Ratnasingham, S. and deWaard, J.R. (2003b) Barcoding animal life: cytochrome *c* oxidase subunit 1 divergences among closely related species. Philos Trans R Soc Lond B. 270,596-599.

[13]. Marshall, E. (2005) Will DNA bar codes breathe new life into classification?. Science. 307:1037.

[14]. Austerlitz, F., David, O., Schaeffer, B., Bleakley, K., Olteanu, M., Leblois, R., Veuille, M., Laredo, C. (2009) DNA barcode analysis: a comparison of phylogenetic and statistical classification methods. BMC Bioinformatics, **10,** (Suppl 14):S10.

[15]. Elias, M., Hill, R.I., Willmott, K.R., Dasmahapatra, K.K., Brower, A.V., Mallet, J., Jiggins, C.D. (2007) Limited performance of DNA barcoding in a diverse community of tropical butterflies. Proc R Soc B 274, 2881–2889.

[16]. Hebert, P.D. and Gregory, T.R. (2005) The promise of DNA barcoding for taxonomy. Syst Biol. 54, 852-859.

[17]. Barrett, R. D. H. and Hebert, P. D. N. (2005) Identifying spiders through DNA barcodes. Can J Zool. 83, 481–491.

[18]. Smith, M.A., Fisher, B. L.and Hebert, P. D.N. (2005) DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: The ants of Madagascar. Philos Trans R Soc Lond B Biol Sci. 360(1462), 1825-1834.

[19]. Clare, E.L., Lim, B.K., Engstrom, M.D., Eger. J.L., Hebert, P.D.N. (2007) DNA barcoding of Neotropical bats: species identification and discovery within Guyana. Mol Ecol Notes. 7,184–190.

[20]. Costa, F.O., DeWaard, J.R., Boutillier, J., Ratnasingham, S., Dooh, R.T., Hajibabaei, M. and Hebert, P.D.N (2007) Biological identifications through DNA barcodes: the case of the Crustacea. Can J Fish Aquat Sci, 64,272-295.

[21]. Hajibabaei, M., Janzen, D.H., Burns, J.M., Hallwachs, W. and Hebert, P.D.N. (2006) DNA barcodes distinguish species of tropical Lepidoptera. Proc Natl Acad Sci U.S.A. 103,968-971.

[22]. Anker, A., Hurt, C. and Knowlton, N. (2007) Revision of the *Alpheus nuttingi* (Schmitt) species complex (Crustacea : Decapoda : Alpheidae), with description of a new species from the tropical eastern Pacific. Zootaxa. 1577, 41-60.

[23]. Bucklin, A., Wiebe, P.H., Smolenack, S.B., Copley, N.J., Beaudet, J.G., Bonner, K.G.,Farber-Lorda, J., Pierson, J.J. (2007) DNA barcodes for species identification of euphausiids (Euphausiacea, Crustacea). J Plankton Res. 29, 483–493.

[24]. Tavares, E.S. and Baker, A.J. (2008). Single mitochondrial gene barcodes reliably identify sister-species in diverse clades of birds. BMC Evol Biol. 8:81.

[25]. Lise Frezala and Raphael Leblois. (2008) Four years of DNA barcoding: Current advances and prospects. Infection, Genetics and Evolution. 8, 727–736.

[26]. Robins, J.H., Hingston, M., Matisoo-Smith, E. and Ross, H.A. (2007). Identifying Rattus species using mitochondrial DNA. Molecular Ecology Notes. 7, 717–729.

[27]. Kerr, K.C.R., Stoeckle, M.Y., Dove, C.J., Weigt, L.A., Francis, C.M. and Hebert, P.D.N. (2007) Comprehensive DNA barcode coverage of North American birds. Mol Ecol Notes, 7,535-543.

[28]. Hubert, N., Hanner, R., Holm, E., Mandrak, N., Taylor, E., Burridge, M., Watkinson, D.A., Dumont, P., Curry, A., Bentzen, P., Zhang, J., April, J. and Bernatchez, L. (2008) Identifying Canadian freshwater fishes through DNA barcodes. PLoS.ONE. 3,2490pp.

[29]. Foottit R.G., Maw, H.E.L., Havill, N.P., Ahern, R.G. and Montgomery, M.E. (2009) DNA Barcodes to Explore Diversity and Life Cycles in the Adelgidae (Insecta:Hemiptera: Aphidoidea). Molecular Ecology Resources. 9(1), 188-195.

[30]. Sheffield, C.S., Hebert, P.D.N., Kevan, P.G. and Packer, L. (2009) DNA barcoding a regional bee (Hymenoptera: Apoidea) fauna and its potential for ecological studies. Mol Ecol Res. 9,196–207.

[31]. Ward, R.D., Holmes, B.H., White, W.T., Last, P.R. and Hebert, P.D.N. (2008) DNA barcoding Australasian chondrichthyans results and potential uses in conservation. Mar Freshwater Res. 59, 57-71.

[32]. Trewavas, E. (1983) Tilapiine Fishes of the Genera *Sarotherodon,Oreochromis* and *Danakilia*, Br Mus (Nat. Hist.), London.

[33]. Randall, D. J., Wood, C.M., Perry, S.F., Bergman, H., Maloiy, G.M.O., Mommsen T.P. and Wright, P.A. (1989). Urea excretion as a strategy for survival in a fish living in a very alkaline environment. Nature. 337,165-166.

[34]. Tine, M., de Lorgeril, J., D'Cotta, H., Pepey, E., Bonhomme, F., Baroiller, J.F. and Durand, J.D. (2008) Transcriptional responses of the black-chinned tilapia *Sarotherodon melanotheron* to salinity extremes. Mar Genomics. 1(2), 37-46.

[35]. Baroiller, J.F. and Toguyeni, A. (2004) The Tilapiini tribe: environmental and social aspects of reproduction and growth. In: Fisheries and Aquaculture, Encyclopedia of Life Support Systems (EOLSS), Developed under the Auspices of the UNESCO (P. Safran). Eolss Publishers, Oxford, UK. Ed.

[36]. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

[37]. Word, R.D., Holmes, B.H. and Yearsley, G.K. (2008) DNA barcoding reveals a likely second species of Asian Sea bass (*Baramundai)(Lates calcarifer*). J Biology.72, 458-463.

[38]. Hebert, P.D. and Gregory, T.R. (2005) The promise of DNA barcoding for taxonomy. Syst Biol. 54, 852-859.

[39]. Hall Jackson, C.A., Cross, D.A., Morrice, N. and Smythe, C. (1999) ATR is a caffeine-sensitive, DNA-activated protein kinase with a substrate

specificity distinct from DNA-PK. Oncogene. 18, 6707-6713.

[40]. Nei, M. and Kumar, S. (2000) Molecular Evolution and Phylogenetics. Oxford University Press. New York, 333pp.

[41]. 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.

[42]. Felsenstein, J. (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution. 39, 783-791.

[43]. Efron, B. (1982). The jackknife, the bootstrap, and other resampling plans. Society for Industrial and Applied Mathematics, Philadelphia, 38pp.

[44]. Barbuto, M., Galimberti, A., Ferri, E., Labra, M., Malandra, R., Galli, P. and Casiraghi, M. (2010) DNA barcoding reveals fraudulent substitutions in shark seafood products. Food Res Int. 43, 376-381.

[45]. Ratnasingham, S. and Hebert, P.D. (2007) Bold: The barcode of life data system. Mol Ecol Notes. 7, 355-364.

[46]. Sneath, P.H.A. and Sokal, R.R. (1973) Numerical Taxonomy. W.H. Freeman and Company, San Francisco.

[47]. Tamura, K., Nei, M. and Kumar, S. (2004) Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc Natl Acad Sci U.S.A. 101, 1030-11035.

[48]. Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011) Molecular Evolutionary Genetics Analysis: using maximum likelihood, evolutionary distance and maximum parsimony methods. Mol Biol Evol. 28, 2731-2739.

[49]. Wong, E. and Hanner, R. (2008) DNA barcoding detects market substitution in North American seafood. Food Res Int. 41, 828-837.

[50]. Saccone, N.L., Jr. Downey, T.J., Meyer, D.J., Neuman, R.J. and Rice, J.P. (1999) Mapping genotype to phenotype for linkage analysis. Genet. Epidemiol. 17(1), 703-708.

[51]. Akbar John, B., Prasanna Kumar, C., Lyla, P.S., Ajmal Khan, S. and Jalal, K.C.A. (2010) DNA Barcoding of *Lates calcarifer* (Bloch, 1970). Res J Biol Sci. 5(6), 414-419.

[52]. Prasanna Kumar, S., Ravikumar, A., Somu, L., Vijaya Prabhu, P., Subbaiya, M. and Periyasamy Subbaraj, R. (2011) Tracheostomal myiasis: a case report and review of the literature. Case Rep Otolaryngol. 2011, 303-510.

[53]. Bucklin, A., Frost, B.W., Bradford-Grieve, J., Allen, L.D. and Copley, N.J. (2003) Molecular systematic and phylogenetic assessment of 34 calanoid copepod species of the *Calanidae* and *Clausocalanidae*. Mar Biol.142, 333-343.

[54]. Bucklin, A., Wiebe, P.H., Smolenack, S.B., Copley, N.J., Beaudet, J.G., Bonner, K.G., Farber Lorda, J. and Pierson, J.J. (2007) DNA barcodes for species identification of *euphausiids* (Euphausiacea: Crustacea). J Plank Res. 29, 483-493.

[55]. Hajibabaei, M., Singer, G.A.C., Clare, E.L. and Hebert, P.D.N. (2007a) Design and applicability of DNA arrays and DNA barcodes in biodiversity monitoring. BMC Evol Biol. 5, 24.

[56]. Avise, J.C. (2000) Phylogeography: The History and Formation of Species. Harvard University Press, Cambridge.