

Mapana – Journal of Sciences 2024, Vol. 23, No. 1, 133-147 ISSN 0975-3303 | https://doi.org/10.12723/mjs.68.7

DNA Barcoding of Fish Fauna Using Mitochondrial CO1 Gene

Anisha S Menon*, Nandini N.J.[†], Manoj Mathews[‡] and Jisha Jacob*

Abstract

This study aimed to investigate the quantitative relationships between four fish species from three genera based on molecular analyses (barcoding) of nine species from five genera utilizing the mitochondrial COI gene. Species within the same genus showed more transitional incompatibilities than transversional mismatches. The samples were divided into four main groups by a phylogenetic tree built from the sequencing data (cytochrome COI) of samples from the two populations using the neighbour-joining method. As dissimilar species were clustered under separate nodes and similar species were clustered under the same nodes, the neighbourjoining tree revealed various clusters corresponding to the taxonomic status of the species. In conclusion, the mitochondrial CO1 gene is a useful molecular marker for DNA barcoding.

Keywords: DNA barcoding, mitochondrial CO1 gene, molecular taxonomy, biodiversity assessment, DNA sequencing.

^{*} Molecular Microbial Ecology Lab, PG and Research Department of Zoology, St. Joseph's College (Autonomous), Devagiri, University of Calicut 680 555, Kerala, India; anishasumith13@gmail.com, jishajacob@ devagiricollege.org

[†] Department of Zoology, University College, Trivandrum; nandininnj@ gmail.com

[‡] PG and Research Department of Chemistry, St. Joseph's College (Autonomous), Devagiri, University of Calicut 680 555, Kerala, India; mathewsmanoj@gmail.com

Introduction

Fish is high in protein and micro- and macronutrient supplements, as well as vitamins A, B, and D. Omega-3 fatty acids, which are vital for the body and brain, are also abundant in it [1]. Fish are among the healthiest foods on the earth and are consumed in a broad variety due to their contribution to fitness and well-being [2]. In addition to being consumed as food, fish is becoming more and more popular as feed. [3].

Among the fishes in India, our study focuses on 9 species of fishes belonging to 5 genera: Etroplus suratensis, Etroplus maculatus, Nematalosa nausus, Sardinella longiceps Thunnus albacares, Auxis rochei, Oreochromis niloticus, Oreochromis urolepis, and Oreochromis mossambicus. Etroplus is the state fish of Kerala [4]. The largest of the native cichlids is the *Etroplus suratensis*, often known as the Banded Pearl spot. It is a euryhaline fish that is typically found in riverine estuaries, coastal lagoons, and both natural and artificial freshwater environments in peninsular India, mostly in Kerala and Sri Lanka. [4] [5]. Orange chromide, or *E. maculatus*, is found in Kerala's backwaters. Kerala's main fish, sardine, is much sought-after for the state's seafood food security and is said to be an inexpensive treat that Keralites can add to their diets to increase their protein intake [6]. One of the most significant and economically significant fisheries on India's west coast is the oil sardine, or Sardinella longiceps [7]. Nematalosa nausus is found in fresh, brackish, and marine along Mumbai, Karnataka, and Kerala coasts [8]. Auxis rochei, known as Bullet tuna, is found exclusively in tropical oceans, including the Mediterranean Sea, and can grow up to 50cm [9]. Thunnus albacares known as yellowfin tuna is a near-threatened, highly migratory species, seen except in the Mediterranean [10]. Oreochromis species (Oreochromis mossambicus (Tilapia), Oreochromis niloticus (Nile tilapia, cultured since 2012) Oreochromis urolepis (Wami tilapia) are used in aquaculture wherein utilizes their high adaptability, fecundity and ability to attain enormous size in a short time to culture and feed the rising population [11].

Numerous techniques exist for identifying species. Since fish are thought to be more phenotypically variable than most other vertebrates and have comparatively larger within-population coefficients of

variation of phenotypic traits, identification by morphometric measurement is frequently inconsistent [12]. In this work, an attempt was made to identify these fish species using molecular taxonomy since it seems to be the most effective tool for species identification and has advantages over other taxonomy methods [9]. The portion of the mitochondrial COI gene that is utilized as a "species barcode" is excellent and has a high degree of species identification efficiency [13]. Thus, this tool is employed for quick commercial analysis, particularly species confirmation [14]. Following the molecular taxonomy-based identification of species, the evolutionary relationship between organisms can be evaluated by the construction of a phylogenetic tree [14]. The importance of understanding the phylogenetic relationship existing within and between species in a population cannot be overemphasized as it has become a handy method used in tracing the origin and evolution of species from the sequences obtained from organisms [15].

Materials and Methods

Fish Collection and Identification

9 fish were collected from fish markets in and around Kozhikode district. A preliminary analysis of species of *Etroplus, Sardinella, Nematalosa, Thunnus, Auxis,* and *Oreochromis* was done with the help of fishermen and based on their taxonomic features such as length, width, the structure of the body, the color of the body, fin structure, etc, their morphological appearances were deeply studied.

Molecular Analysis

DNA Isolation, Amplification, Electrophoresis, And Gel Elution

Total genomic DNA was extracted using a genomic DNA extraction kit following the manufacturer's protocol using 25 mg abdominal muscle tissue from samples

The section of the mitochondrial DNA genome from the COI gene was amplified using the following primers;

Forward primer, F1:5 'TCAACCAACCACAAGACATTGGCAC 3'

Reverse primer, R1:5 TAGACTTCTGGGTGGCCAAAGAATCA 3' The DNA samples were subjected to the PCR reaction.

The PCR product was confirmed by Agarose gel electrophoresis using a 1.5% gel prepared in 1X TAE buffer and visualization of the bands was observed under a UV trans-illuminator. The size of the product can be estimated with the help of a 1kb base pair DNA ladder. Gel elution is performed and stored the purified DNA at -20°C.

Amplification

Using the forward primer 5'3' and the reverse primer 5'3', the COI gene-containing region of the mitochondrial DNA genome was amplified. In a PCR tube, 25µL of Emerald Amp GT Master mix (Takara), 2µL of COI forward and reverse primer, 20µL sterile water, and 1µL DNA sample (50ng/ml) were combined to create a 50µl reaction mixture under cold conditions. The PCR reaction was applied to the DNA samples. There was a five-minute initial denaturation at 95°C, and then 35 cycles of denaturation at 95°C for ten seconds each. The cycle was finished with primer annealing at 54°C for 45 seconds, primer extension at 72°C for 45 seconds, and final elongation at 72°C for three minutes. The temperature was maintained at 4°C.

Sequencing of the Gene, Analysis, and Construction of Phylogenetic Tree

Sequences of the sample were done at using COI gene-specific forward and reverse primers of fish. The ends of COI gene sequences were trimmed and multiple sequence alignment is done using MEGA 10 (Molecular Evolutionary Genetic Analysis) software. The sequence was analyzed using the BLAST Bioinformatics tool for comparing species. The sequences are applied to Mega 10 (Molecular Evolutionary Genetic Analysis) software by trimming the ends for the construction of a phylogenetic tree by the neighbour-joining method.

Results and Discussion

The present study was undertaken with 9 different species of 6 genera of fishes (Fig 1)



Fig 1- Morphological appearances of 9 different species of fishes

Table 1- Scientific nomenclature of the above-mentioned 9 different

 species of fishes

| 1. | Etroplus suratensis | 2. | Etroplus maculatus | 3. | Nematalosa nausus |
|----|-----------------------|----|----------------------|----|-------------------------|
| 4. | Sardinella longiceps | 5. | Thunnus albacares | 6. | Auxis rochei |
| 7. | Oreochromis niloticus | 8. | Oreochromis urolepis | 9. | Oreochromis mossambicus |

Molecular Analysis

To verify the species identity, the sequences from the current investigation were matched using BLAST to sequences published in publicly accessible databases like GenBank. Table 2 lists the GenBank Accession numbers for the COI sequences of the investigated species.

Table 2: Accession numbers for the CO1 sequences of samples

| | Denoted by | GenBank Accession number |
|-------------------------|------------|--------------------------|
| Etroplus suratensis | D3 | MG923359 |
| Etroplus maculatus | D4 | AP009505 |
| Sardinella longiceps | D1 | MG251979 |
| Thunnus albacares | T4 | KT719291 |
| Auxis rochei | T5 | MH638691 |
| Oreochromis niloticus | T1 | LC487084 |
| Oreochromis urolepis | T2 | MF509598 |
| Oreochromis mossambicus | Т3 | MK210574 |

The pairwise BLAST alignment of D3 *Etroplus suratensis* COI sequences and with D4 *Etroplus maculatus* sequence shows 83 mismatches

which include 53 transitions and 30 transversions present. The pairwise BLAST alignment of D1 Sardinella longiceps CO I sequences and with D2 *Nematalosa nausus* sequence shows dissimilarities and there are 99 mismatches. There are 59 transitions and 40 transversions found among those mismatches. The pairwise BLAST alignment of Thunnus albacores and Auxis rochei COI sequences shows 54 mismatches and include 35 Transition mismatches and 18 transversion mismatches. Sequences between T1 Oreochromis niloticus COI sequences with T2 Oreochromis urolepis COI sequence gave 33 mismatches including 29 transitions and 4 transversions. Sequences between T1 Oreochromis niloticus COI sequences with T3 Oreochromis mossambicus COI sequence gave 31 mismatches and within them, there were 25 Transitions and 6 transversions. The sequence between T2 *Oreochromis urolepis* COI sequences with T3 *Oreochromis mossambicus* COI sequence had 25 mismatches and out of them 21 were transition and 4 were transversion. So this pairwise alignment shows the molecular discrimination of different species based on the COI gene.

Sequence Analysis

The amplified DNA was eluted using Gene JET elution kit and checked using Agarose gel electrophoresis. The eluted DNA along with its specific primers was send to SciGenome, Kakkanad for sequencing. The sequenced sample was then edited using MEGA 10 (Molecular Evolutionary Genetic Analysis) software. The aligned sequence was submitted to BLAST in order to authenticate the similar sequences in the database.



Fig.2- The chromatogram of CO1 forward sequence of T1 Oreochromis niloticus

>T1_Forward_23024-1_P3205, Trimmed Sequence (643 bp)

Fig.3 - Consensus sequence of CO1 forward sequence of T1 Oreochromis niloticus

| LC487084.1:27-669 Ore | ochromis n | iloticus Onil5 mitochondrial COI gene for cytochrome c oxidase subunit 1, partial cds |
|--------------------------|-------------|---|
| Sequence ID: Query_12689 | Length: 643 | Number of Matches: 1 |

| Some 1188 bhs(643) Eugent 0.0 Eugent 643/643(100%) Gass 0/643(0%) Strand Plus/Plus 1188 bhs(643) 0.0 643/643(100%) 0/643(0%) Plus/Plus 1188 bhs(643) 0.0 643/643(100%) 0/643(0%) Plus/Plus 0.0 1188 bhs(643) 0.0 643/643(100%) 0/643(0%) Plus/Plus 0.0 1188 bhs(643) 0.0 643/643(100%) 0/643(0%) Plus/Plus 0.0 1 GCCCCGTATAGTAGGAACTGCATTAGCTACCAATCTCSGAGCAACTAGACTAAGCCAAGCC | nge 1: 1 t | t to 643 Graphics | Match & Previous Match |
|---|-------------------|---|------------------------|
| guery 1 GAGECGGATAGTAGGAACTGCACTAAGCCCCCCTAATTCGGACAACTAAGCCAGCC | ore 88 bits(64 | Expect Identities Gaps Strand 643) 0.0 643/643(100%) 0/643(0%) Plus/Plu | us |
| Sbjet 1 TGAGCCGGATAGTAGGAACTGCACTAAGCCTCATATCGGCAGAACTAAGCCAGCC | ery 1 | TGAGCCGGATAGTAGGAACTGCACTAAGCCTCCTAATTC5GGCAGAACTAAGCCAGCCCG | 60 |
| guery 61 GCTECTITETEGBAGACGACAGACCATATETATAGTATGTACKGACATGCTTTCTATA 128 bbjet 61 GCTECTITETEGBAGACGACAGATETATAGTAGTGATAGTACGACATGCTTTTCATA 128 bbjet 121 TAATTTICTITATAGTAATACCAATATGATTGGAGCTTTGGAAACTGACTAGTACCCC 188 bbjet 121 TAATTTICTITATAGTAATACCAATATGATTGGAGCTTTGGAAACTGACTAGTACCCC 188 sbjet 121 TAATTTICTITATAGTAATACCAATATGATTGGAGCTTTGGAAACTGACTAGTACCCC 188 guery 121 TCATGATTGGTGCACCAGACATGGCCTTCCCCTGGAATAAATA | jct 1 | TGAGCCGGATAGTAGGAACTGCACTAAGCCTCCTAATTCSGGCAGAACTAAGCCAGCCCG | 60 |
| Sbjet 61 GCTCTTCTCCCGAGGACCAGCAGACTAGATTAGATTGAATIGTAACAGCACATACTAGCTTTCGAAA 128 Query 121 TAATTTTCTTTATAGTAATACCAATTAGGATGGAGGCTTTGGAAACTGACTAGCTAG | ery 61 | GCTCTCTTCTCGGAGACGACCAAATCTATAATGTAATTGTTACAGCACATGCTTTCGTAA | 120 |
| guery 121 AATTTICTITATAGTAATACCAATTATGATTGGAGGCTTTGGAAACTGACTAGTAGCCCC 180 Sbjet 121 TAATTTICTITATAGTAATACCAATTATGATTGGAGGCTTTGGAAACTGACTAGTACCCCC 180 Sbjet 121 TAATTTICTITATAGTAATACCAATTATGATTGGAGGCTTGGAAACTGACTG | jct 61 | GCTCTCTTCTCGGAGACGACCAAATCTATAATGTAATTGTTACAGCACATGCTTTCGTAA | 120 |
| Sbjet 121 TÄÄTTTTCTTTÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ | ery 121 | 1 TAATTTTCTTTATAGTAATACCAATTATGATTGGAGGCTTTGGAAACTGACTAGTACCCC | 180 |
| Query 181 CATGATTGGTGCACCAGACATGGCCTTCCCTGGAATAAATA | jct 121 | 1 TAATTTTCTTTATAGTAATACCAATTATGATTGGAGGCTTTGGAAACTGACTAGTACCCC | 180 |
| Sbjet 181 fcArGATTGGTGCACCAGACATGGCCTTCCCCCCAATAAATAACAACAGACGCTTTGACTTC 240 Query 241 TcccccccTCATTTCTTCTTCTTCTCGCCCCAATACTGGAGTCGAAGCAGGTGCCGGCAACAG 300 Sbjet 241 TcccccccCTCATTTCTTCTTCTCCTCGCCCCAATACTGGAGTCGAAGCAGGTGCCGGCAACAG 300 Sbjet 241 TcccccccCTCATTTCTTCTTCTCCCCGCCCCAATACTGGAGTCGAAGCAGGTGCCGGCCAACG 300 Sbjet 241 TccccccCCTCATTTCTTCTTCTCCCCGCCCGCGCGCCGCCGCGCGGCCGCC | ery 181 | 1 TCATGATTGGTGCACCAGACATGGCCTTCCCTCGAATAAATA | 240 |
| Query 241 TCCCCCCCTCATTICTTCTTCTCCCCCCCTCATCTGGAGCCAGGTCGAAGCAGGTGCGGACAGA 300 Sbjct 241 TCCCCCCCTCATTICTTCTTCTCCCCGCCTCATCTGGAGCAGGTGGAAGCAGGTGCGGACAGAG 300 Sbjct 301 GATGGACTGTTTATCCCCCGCTGGCAGGCAATCTTGGCCACGCTGGACGTCGGCAAGAG 300 Sbjct 301 GATGGACTGTTTATCCCCCGCTGGCAGGCAATCTTGCCCACGCTGGACCTTCTGTTGACCT 360 Sbjct 301 GATGGACTGTTTATCCCCCGCTGGAGGCAATCTGCCCACGCTGGACCTTCTGTTGACCT 360 Sbjct 301 GATGGACTGTTTATCCCCCGCTGGAGGCAATCTGCCCACGCTGGCCATTGTGTGACTT 360 Sbjct 10 AACCATCTTCTCCCCCCACTTGGCCGAGGCAATCTATTTTAGGTGCAATTAATT | jct 181 | 1 TCATGATTGGTGCACCAGACATGGCCTTCCCTCGAATAAATA | 240 |
| Sbjet 241 fccccccctcattfctfctfctfctcccccccccccccc | ery 241 | 1 TCCCCCCCTCATTTCTTCTTCTTCTCCGCCTCATCTGGAGCCGAAGCAGGTGCCGGCACAG | 300 |
| guery 381 GATGGACTGTTTATCCCCCCGCTCGCAGGCAATCTTGCCCCCGCTGGACCTTCTGTTGACT 368 sbjct 381 GATGGACTGTTTATCCCCCCGCTCGCAGGCAATCTTGCCCCCCCGCGGACCTTCTGTTGACT 368 guery 361 TAACCATCTTCTCCCCCCACTTGGCCGAGGCAATCTATCT | jct 241 | 1 TCCCCCCCTCATTTCTTCTTCTCCGCCTCATCTGGAGTCGAAGCAGGTGCCGGCACAG | 300 |
| sbjct 301 GATGGACTGTTTATCCCCCGCGTCGGAGGCAATCTTGCCCACGCTGGACCTTCTGTTGACT 360 Query 361 TAACCATCTTCTCCCCCCACTTGGCCGGAGTGTCATCTATTTTAGGTGCAATTATTTTA 420 sbjct 361 TAACCATCTTCTCCCCCCACTTGGCCGGAGTGTCATCTATTTTAGGTGCAATTATTTTA 420 sbjct 361 TAACCATCTTCTCCCCCCACTTGGCCGGAGTGTCATCTATTTTAGGTGCAATTATTTTA 420 guery 421 TCACAACCACTTATTAACATGAAACCCCCCTGTCCCCAATTACAAACACCCCTATTTG 480 sbjct 421 TCACAACCACTTATTAACATGAAAACCCCCCTGTCCCCAATTACCAAACACCCCTATTTG 480 yuery 481 TGTGATCCGTCCTAATTACCGCGCGAGTACTACTCCTTCTTCTCCCGCGCCGTCTTGCCGCCG 540 sbjct 481 TGTGATCCGTCCTAATTACCGGAGTACTACTACTCCTTCTATCCCTGCCCGCGTCTTGCCCGCCG 540 | ery 301 | 1 GATGGACTGTTTATCCCCCGCTCGCAGGCAATCTTGCCCACGCTGGACCTTCTGTTGACT | 360 |
| Query 361 TAACCATCTTCTCCCCCCCACTTGSCCGGAGTGTCATCTATTTTAGGTGCAATTAATTTTA 420 sbjct 361 TAACCATCTTCTCCCCCCCACTTGSCCGGAGTGTCATCTATTTTAGGTGCAATTAATTTTA 420 guery 421 TCACAACCATTATTAACATGAAACCCCCTGTCTCCCAATTCAACCAAC | jct 301 | 1 GATGGACTGTTTATCCCCCGCTCGCAGGCAATCTTGCCCACGCTGGACCTTCTGTTGACT | 360 |
| sbjct 361 TAACCATCTTCTCCCCCCACTTGGCCGGAGTGTCATCTATCT | ery 361 | 1 TAACCATCTTCTCCCCCCCACTTGGCCGGAGTGTCATCTATTTAGGTGCAATTAATT | 420 |
| Query 421 TCACAACCATTATTAACATGAAACCCCCTGCCATTCCCAATATCAAACACCCCTATTTG 480 Sbjct 421 TCACAACCATTATTAACATGAAACCCCCCTGCCATTCCCCAATATCAAACACCCCTATTTG 480 yuery 481 TGTGATCGGTCCTAATTACCGCGAGTACTACTCCTTCTTCCCGGTCCTGCCGCCG 540 ybjct 481 TGTGATCCGTCCTAATTACCGCGCGAGTACTACTCCCTTCTTCTCCCGGTCCTGCCGCCG 540 ybjct 481 TGTGATCCGTCCTAATTACCGCGCGAGTACTACTCCTTCTTCTGCCGCGCCG 540 | jct 361 | 1 TAACCATCTTCTCCCTCCACTTGGCCGGAGTGTCATCTATTTTAGGTGCAATTAATT | 420 |
| Sbjet 421 TCACAACCATTATTAACATGAAACCCCCTGCCATTCCCCAATCAAACACCCCTATTTG 480 Query 481 TGTGATCCGTCCTAATTACCGCAGTACTACTCCTTCTTTCCCTGCCCGTCTTGCCGCCG 540 Sbjet 481 TGTGATCCGTCCTAATTACCGCAGTACTACTCCTTCTTTCCCTGCCCGTCCTTGCCGCCG 540 | ery 421 | 1 TCACAACCATTATTAACATGAAACCCCCTGCCATCTCCCAATATCAAACACCCCCTATTTG | 480 |
| Query 481 TGFGATCCGTCCTAATTACCGCAGTACTACTCCTTCTATCCCTGCCCGTTCTTGCCGCG 540 bbjct 481 TGFGATCCGTCCTAATTACCGCAGTACTACTCCTTCTATCCCTGCCCGTTCTTGCCGCCG 540 | jct 421 | 1 TCACAACCATTATTAACATGAAACCCCCTGCCATCTCCCAATATCAAACACCCCCTATTTG | 480 |
| Sbjct 481 TGTGATCCGTCCTAATTACCGCAGTACTACTCCTTCTATCCCTGCCCGTTCTTGCCGCCG 540 | ery 481 | 1 TGTGATCCGTCCTAATTACCGCAGTACTACTCCTTCTATCCCTGCCCGCTTCTTGCCGCCG | 540 |
| BUARU EAS CENTENEARTHETTETANENCOLECANNEETTANAECTETTETTENEEEEEEAE COO | jct 481 | 1 TGTGATCCGTCCTAATTACCGCAGTACTACTCCTTCTATCCCTGCCCGTTCTTGCCGCCG | 540 |
| QUELA 241 GENTEREN DELEGRANCE MARCHARCELE CONTRACE CONTRACES CONTRACES CONTRACES CONTRACES CONTRACES CONTRACES | ery 541 | 1 GCATCACAATACTTCTAACAGACCGAAACCTAAACACAACCTTCTTTGACCCTGCCGGAG | 600 |
| SDJCT 541 GCATCACAATACTTCTAACAGACCGAAACCTAAACACAACCTTCTTTGACCCTGCCGGAG 600 | jct 541 | 1 GCATCACAATACTTCTAACAGACCGAAACCTAAACACAACCTTCTTTGACCCTGCCGGAG | 600 |
| QUERY 601 GAGGAGACCCCATCCTATACCAACACTTATTCTGATTCTTCGG 643 | ery 601 | 1 GAGGAGACCCCATCCTATACCAACACTTATTCTGATTCTTCGG 643 | |

Fig.4- BLAST alignment of Oreochromis niloticus with database sequences LC487084

GTGCTTG /GCCGG /TAGTAGGA/CC GC/CTA/AGCCTCCTA/IT/CGGGC/GA/CTA/AGCC/GCCCGGCTCTCTCCGG/GA/CG/CC/GA/TTATA/IGTA/IT/GTA/AGC/A/AGCC/GC/CTCCGG/GA/CG/CC/GA/TTATA/IGTA/IT/GTA/AGC/A/AGCC/GA/AGCC/GA/AGC/GAGC/GA/AGC/GA/AGC/GA/AGC/GA/AGC/GA/AGC/GA/ иттептитиятиятиятиятиятиванаетитериистовие и посиетиление и терескосивие и посиетование и посиетие и посиетование и пос n alter the manuscript and walk and be manustral to be made the description and a new terms of the second second IGGGCCTTCTGTTG./CTT/ line desentand introducer the enternation of the constitution of the СТА АСАБИТ СВСА ИССТАА АСАСА АССТ ТСТТТБАСССТВСОВ АВ В В АССССАТССТТТ ИССА АСАСТ Т ИТТСБИТТС 560 570 180 590 600 610 620 650 640 and and an and a second a s

Fig.5- The chromatogram of COI forward sequence of T2 Oreochromis urolepis

>T2_Forward_23024-4_P3205, Trimmed Sequence (645 bp)

GCTTGAGCCGGATAGTAGGAACCGCACTAAGCCTCCTAATTCGGGCAGAACTAAGCCAGC CCGGCTCTCTCCTCGGAGACGACCAGATTTATAATGTAATTGTTACAGCACATGCTTTCG TAATAATTTTCTTTATAGTAATACCAATTATAATTGGAGGTTTTGGAAACTGACTAGTAC CACTTATGATTGGTGCACCAGACATGGCCTTCCCTCGAATGAACAACATGAGTTTTTGAC TCCTTCCCCCCTCATTTCTCCTCCTCCTCGCCTCATCCGGAGTCGAAGCAGGGGCCGGTA CAGGATGAACTGTTTACCCCCCACTCGCAGGCAATCTCGCCCACGCTGGGCCTTCTGTTG TTATCACAACCATTATTAACATAAAACCCCCTGCCATCTCTCAATATCAAACACCCCCTAT TTGTGTGATCCGTTCTAATTACCGCAGTGCTGCTCCTACTATCCCTGCCCGTTCTTGCCG CCGGCATCACAATACTTCTAACAGATCGCAACCTAAACACAACCTTCTTTGACCCTGCCG GAGGAGGAGACCCCATCCTTTACCAACACTTATTCTGATTCTTCG





MF509598.1:14-659 Oreochromis urolepis voucher JeliRed04 cytochrome oxidase subuni: 1 gene, partial cds; mitochondrial





Fig.8-The chromatogram of COI forward sequence of T3 Oreochromis mossambicus

>T3_Forward_23024-6_P3205, Trimmed Sequence (646 bp)

Fig.9- Consensus sequence of COI sequence from T3 Oreochromis mossambicus

ISSN 0975-3303

MK210574.1:5-647 Oreochromis mossambicus voucher ANU OM1 cytochrome c oxidase subunit I (CO1) gene, partial cds; mitochondrial Sequence ID: Query_3905 Length: 643 Number of Matches: 1

| Range | 1:1t | o 643 Gra | aphics | | | * Next | Match & Previous Match |
|-----------------|--------|-----------|---------------|-----------------------------|---------------------|-------------------|------------------------|
| Score 1188 b | its(64 | 3) | Expect 0.0 | Identities 643/643(100%) | Gaps 0/643(0%) | Strand Plus/Pl | us |
| Query | 1 | GCTTGAG | CCGGATAGT | AGGAACTGCATTAAGCCT | CCTAATTCGGGCAGAACT | AAGCCAGC | 60 |
| Sbjct | 1 | GCTTGAG | CCGGATAGT | AGGAACTGCATTAAGCCT | CCTAATTCGGGCAGAACT | AAGCCAGC | 60 |
| Query | 61 | CCGGCTC | TCTCCTCGG | AGACGACCAGATTTATAA | TGTAATTGTTACAGCACA | TGCTTTCG | 120 |
| Sbjct | 61 | CCGGCTC | TCTCCTCGG | AGACGACCAGATTTATAA | TGTAATTGTTACAGCACA | TGCTTTCG | 120 |
| Query | 121 | TAATAAT | TTICTITAT | AGTAATGCCAATTATAAT | TGGAGGTTTTGGAAACTG | ACTAGTGC | 180 |
| Sbjct | 121 | TAATAAT | TTTCTTTAT | AGTAATGCCAATTATAAT | TGGAGGTTTTGGAAACTG | ACTAGTOC | 180 |
| Query | 181 | CACTAAT | GATTGGTGG | ACCAGACATGGCCTTCCC | TCGAATAAATAACATGAG | TTTTTGAC | 240 |
| Sbjct | 181 | CACTAAT | GATTGGTGC | ACCAGACATGGCCTTCCC | TCGAATAAATAACATGAG | TTTTTGAC | 240 |
| Query | 241 | TCCTccc | CCCCTCATT | TETECTTETECTEGECTE | ATCCGGGGTCGAAGCAGG | GGCCGGTA | 300 |
| Sbjct | 241 | TCCTCCC | CCCCTCATT | TETECTTETECTEGEETE | ATCCGGGGTCGAAGCAGG | GGCCGGTA | 300 |
| Query | 301 | CAGGATO | GACTGTTTA | TCCCCCACTCGCAGGCAA | TCTCGCCCATGCTGGGCC | TTCCGTTG | 360 |
| Sbjct | 301 | CAGGATO | GACTGTTTA | TCCCCCACTCGCAGGCAA | TCTCGCCCATGCTGGGCC | TTCCGTTG | 360 |
| Query | 361 | ACTTAAC | CATCTTCTC | CCTCCACTTGGCCGGGGT | STCATCTATTTTAGGTGC | AATTAATT | 420 |
| Sbjct | 361 | ACTTAAC | CATCTTCTC | CCTCCACTTGGCCGGGGT | STCATCTATTTTAGGTGC | AATTAATT | 420 |
| Query | 421 | TTATTAC | AACCATTAT | TAACATAAAACCCCCTGC | CATCTCCCAATATCAAAC | ACCCCTCT | 480 |
| Sbjct | 421 | TTATTAC | AACCATTAT | TAACATAAAAACCCCCTGC | CATCTCCCAATATCAAAC | ACCCCTCT | 480 |
| Query | 481 | TIGTATO | ATCCGTTCT | AATTACCGCAGTACTACT | CCTACTATCCCTACCCGT | TCTTGCCG | 540 |
| Sbjct | 481 | TTGTATO | ATCCGTTCT | AATTACCGCAGTACTACT | CCTACTATCCCTACCCGT | TCTTGCCG | 540 |
| Query | 541 | CCGGCAT | CACAATACT | TCTAACAGACCGAAACCT | ΑΑΑΕΑΕΑΑΕΕΤΤΟΤΤΤΓΑΑ | CCCTGCCG | 600 |
| Sbjct | 541 | CCGGCAT | CACAATACT | TCTAACAGACCGAAACCT | AAACACAACCTTCTTTGA | CCCTGCCG | 600 |
| Query | 601 | GAGGAGG | AGACCCCAT | CCTTTACCAACACTTATT | CTGATTCTT 643 | | |
| Sbjct | 601 | GAGGAGG | AGACCCCAT | CCTTTACCAACACTTATT | CTGATTCTT 643 | | |

Fig.10- BLAST alignment of T3 Oreochromis mossambicus sequence MK210574

2_T6Forward_23024-4_P3205,Trimmed Sequence(645 bp)

| Score 987 bit | s(534 | Expect) 0.0 | Identities 606/642(94%) | Gaps 0/642(0%) | Strand Plus/Plu | IS |
|------------------|-------|-----------------|----------------------------|---------------------|--------------------|-----|
| Query | 1 | TGAGCCGGATAGTAG | GAAGTGCACTAAGCCTCC | TAATTCGGGCAGAACTAA | SCCAGCCCG | 60 |
| Sbjct | 4 | TGAGCCGGATAGTAG | GAAdcGCACTAAGCCTCC | TAATTCGGGCAGAACTAA | SCCAGCCCG | 63 |
| Query | 61 | GCTCTCTTCGGAG | ACGACCARATCTATAATG | TAATTGTTACAGCACATG | CTTTCGTAA | 120 |
| Sbjct | 64 | GCTCTCTCCCGGAG | ACGACCAGATTTATAATG | TAATTGTTACAGCACATG | TTTCGTAA | 123 |
| Query | 121 | TAATTTTCTTTATAG | TAATACCAATTATGATTG | SAGGETTTGGAAACTGAC | AGTACCEC | 180 |
| Sbjct | 124 | TAATTTTCTTTATAG | TAATACCAATTATAATTG | GAGGETTTTGGAAACTGAC | TAGTACCAC | 183 |
| Query | 181 | TCATGATTGGTGCAC | CAGACATGGCCTTCCCTC | GAATAAATAACATGAGCT | TTGACTTC | 240 |
| Sbjct | 184 | TEATGATTGGTGCAC | CAGACATGGCCTTCCCTC | GAATGAACAACATGAGTT | TTTGACTCC | 243 |
| Query | 241 | TECCCCCCTCATTIC | TTETTETTETCGCCTCAT | TGGAGTCGAAGCAGGTG | CGGEACAG | 300 |
| Sbjct | 244 | TECCCCCCTCATTIC | TEETEETEETEGEETEAT | GGGAGTCGAAGCAGGGG | CGGTACAG | 303 |
| Query | 301 | GATGGACTGTTTATC | CCCCGCTCGCAGGCAATC | TIGCCCACGCTGGACCTTO | TGTTGACT | 360 |
| Sbjct | 304 | GATGAACTGTTTAC | CCCCACTCGCAGGCAATC | TO GCCCACGCTGGCCTTO | TGTTGACT | 363 |
| Query | 361 | TAACCATCTTETCCC | TCCACTTGGCCGGAGTGT | CATOTATTTTAGGTGCAA | TTAATTTTA | 420 |
| Sbjct | 364 | TAACCATCTTUTCCC | TCCACTTGGCCGGGGTGT | CATOCATTTTAGGTGCAA | TAATTTTA | 423 |
| Query | 421 | TCACAACCATTATTA | ACATGAAACCCCCTGCCA | TCTOCCAATATCAAACACO | CCTATTTG | 480 |
| sbjct | 424 | TCACAACCATTATTA | ACATAAAACCCCCTGCCA | TCTOTCAATATCAAACACO | CCTATTTG | 483 |
| Query | 481 | TGTGATCCGTCCTAA | TTACCGCAGTACTACTCC | TECTATCCCTGCCCGTTC | TTGCCGCCG | 540 |
| Sbjct | 484 | TGTGATCCGTTCTAA | TTACCGCAGTGETGETCC | TACTATCCCTGCCCGTTC | TGCCGCCG | 543 |
| Query | 541 | GCATCACAATACTTC | TAACAGAECGAAACCTAA | ACACAACCTTCTTTGACCO | TGCCGGAG | 600 |
| sbjct | 544 | GCATCACAATACTTC | TAACAGATCGCAACCTAA | ACACAACCTTCTTTGACCO | TGCCGGAG | 603 |
| Query | 601 | GAGGAGACCCCATCC | TATACCAACACTTATTCT | SATTCTTCG 642 | | |
| Sbjct | 684 | GAGGAGACCCCATCC | TTTACCAACACTTATTCT | GATTCTTCG 645 | | |

Fig.11- The pairwise BLAST alignment of Oreochromis niloticus COI sequences with Oreochromis urolepis

3_T6Forward_23024-6_P3205,Trimmed Sequence(646 bp) Sequence ID: Query_40385 Length: 646 Number of Matches: 1

| Score 1011 b | its(54 | Exp 7) 0.0 | ect | Identities 611/643(95%) | Gaps 0/64 | 5 3(0%) P | itrand lus/Plu | JS |
|-----------------|--------|---------------|-------|----------------------------|--------------|----------------|-------------------|-----|
| Query | 1 | TGAGCCGGATAG | TAGGA | ACTGCACTAAGCCT | CCTAATTCGGGC | AGAACTAAGCCAG | CCCG | 60 |
| Sbjct | 4 | TGAGCCGGATAG | TAGGA | ACTGCATTAAGCCT | CCTAATTCGGGC | AGAACTAAGCCAG | CCCG | 63 |
| Query | 61 | GCTCTCTTTCTG | GAGAC | GACCARATCTATAA | TGTAATTGTTAC | AGCACATGCTTTC | GTAA | 126 |
| Sbjct | 64 | GCTCTCTCTCTCG | GAGAC | GACCAGATTTATAA | TGTAATTGTTAC | AGCACATGCTTTC | GTAA | 123 |
| Query | 121 | TAATTTTCTTTA | TAGTA | ATACCAATTATGAT | TGGAGGCTTTGG | AAACTGACTAGT | cocc | 186 |
| Sbjct | 124 | TAATTTTCTTTA | TAGTA | ATGCCAATTATAAT | TGGAGGTTTTGG | AAACTGACTAGT | cdar | 183 |
| Query | 181 | TCATGATTGGTG | CACCA | GACATGGCCTTCCC | TCGAATAAATAA | CATGAGCTTTTGA | CITE | 246 |
| Sbjct | 184 | TAATGATTGGTG | CACCA | GACATGGCCTTCCC | TCGAATAAATAA | CATGAGTTTTTGA | ctcc | 243 |
| Query | 241 | TCCCCCCCTCAT | TTCTT | CTTCTFCTCGCCTC | ATOTGOAGTCGA | AGCAGGTSCCGG | ACAG | 300 |
| 5bjct | 244 | TCCCCCCCTCAT | TTCTC | ETTERETESEETE | ATCCGGGGTCGA | AGCAGGGGGCCGG | ACAG | 303 |
| Query | 301 | GATGGACTGTTT | ATCCC | COGCTCGCAGGCAA | TCTTGCCCACGC | TGGACCTTCTGTT | GACT | 36 |
| Sbjct | 304 | GATGGACTGTTT | ATCCC | COACTCGCAGGCAA | TCTOGCCCATEC | TGGGCCTTCCGGTT | GACT | 36 |
| Query | 361 | TAACCATCTTCT | | CACTTGGCCGGAGT | GTCATCTATTTT | AGGTGCAATTAAT | TTTA | 426 |
| Sbjct | 364 | TAACCATCTTCT | ccctc | CACTTGGCCGGGGT | GTCATCTATTT | AGGTGCAATTAAT | TTTA | 423 |
| Query | 421 | TCACAACCATTA | TTAAC | ATGAAACCCCCTGC | CATCTCCCAATA | TCAAACACCCCTA | TTTG | 486 |
| sbjct | 424 | TACAACCATTA | TTAAC | ATAAAACCCCCTGC | CATCTCCCAATA | TCAAACACCCCTC | TTTG | 483 |
| Query | 481 | TGTGATCCGTCC | TAATT | ACCGCAGTACTACT | CETTETATCCET | GECCGTTETTGEC | GCCG | 546 |
| Sbjct | 484 | TATGATCCGTTC | TAATT | ACCGCAGTACTACT | CCTACTATCCCT | ACCOTTCTTGCC | GCCG | 543 |
| Query | 541 | GCATCACAATAC | TTCTA | ACAGACCGAAACCT | AAACACAACCTT | CTTTGACCCTGCC | GGAG | 60 |
| 5bjct | 544 | GCATCACAATAC | TTCTA | ACAGACCGAAACCT | AAACACAACCTT | CTTTGACCCTGCC | GGAG | 603 |
| Query | 601 | GAGGAGACCCCA | TCCTA | TACCAACACTTATT | CTGATTCTTCGG | 643 | | |
| Sbjct | 604 | GAGGAGACCCCA | TCCTT | TACCAACACTTATT | CTGATTCTTCGG | 646 | | |

Fig.12- The pairwise BLAST alignment of Oreochromis niloticus COI sequences with Oreochromis mossambicus

| Range 1: 1 to 645 Graphics | | | | | | | | |
|----------------------------|--------|--------------------|----------------------------|--------------------|--------------------|-----|--|--|
| Score 1053 b | its(57 | (0) Expect | Identities 620/645(96%) | Gaps 0/645(0%) | Strand Plus/Plu | IS | | |
| Query | 1 | GCTTGAGCCGGATAGT | AGGAACGCACTAAGCCT | CTAATTCGGGCAGAACT | AAGCCAGC | 60 | | |
| Sbjct | 1 | GCTTGAGCCGGATAGT | AGGAACTGCATTAAGCCT | CTAATTCGGGCAGAACT | AAGCCAGC | 60 | | |
| Query | 61 | CCGGCTCTCTCCTCGG | AGACGACCAGATTTATAA | IGTAATTGTTACAGCACA | TGCTTTCG | 120 | | |
| sbjct | 61 | cceectctctctcctcee | AGACGACCAGATTTATAA | IGTAATTGTTACAGCACA | TGCTTTCG | 120 | | |
| Query | 121 | TAATAATTTTCTTTAT | AGTAATACCAATTATAAT | GGAGGTTTTGGAAACTG | ACTAGTAC | 180 | | |
| Sbjct | 121 | taataattttctttat | AGTAATGECAATTATAAT | IGGAGGTTTTGGAAACTG | ACTAGTE | 180 | | |
| Query | 181 | CACTTATGATTGGTGC | ACCAGACATGGCCTTCCC | CGAATGAACAACATGAG | TTTTTGAC | 240 | | |
| sbjct | 181 | CACTAATGATTGGTGC | ACCAGACATGGCCTTCCC | CGAATAAATAACATGAG | TTTTTGAC | 240 | | |
| Query | 241 | TCCTTCCCCCCTCATT | τετεετετεστεσε | ATCCGGAGTCGAAGCAGG | GGCCGGTA | 300 | | |
| Sbjct | 241 | TCCTCCCCCCCCCATT | יביבכין דביבביבפבביבי | ATCCGGGGTCGAAGCAGG | GGCCGGTA | 300 | | |
| Query | 301 | CAGGATGACTGTTTA | CCCCCACTCGCAGGCAA | TCTCGCCCACGCTGGGCC | TTCTGTTG | 360 | | |
| sbjct | 301 | chochtdohctottt | teccccheteschescha | teteseccientesecce | TTCESTTG | 360 | | |
| Query | 361 | ACTTAACCATCTTTTC | CCTCCACTTGGCCGGGGT | STCATCCATTTTAGGTGC | AATTAATT | 420 | | |
| Sbjct | 361 | ACTTAACCATCTICIC | cctccActtggccggggt | STCATCHATTITAGGTGC | AATTAATT | 420 | | |
| Query | 421 | TTATCACAACCATTAT | TAACATAAAACCCCCTGC | | ACCCCTAT | 480 | | |
| sbjct | 421 | TTATTACAACCATTAT | TAACATAAAACCCCCTGC | CATCTC CAATATCAAAC | ACCCCTCT | 480 | | |
| Query | 481 | TTGTGTGATCCGTTCT | AATTACCGCAGTGCTGCT | CTACTATCCCTGCCCGT | TCTTGCCG | 540 | | |
| Sbjct | 481 | TTGTATGATCCGTTCT | AATTACCGCAGTACTACT | CTACTATCCCTACCGT | TCTTGCCG | 540 | | |
| Query | 541 | CCGGCATCACAATACT | TCTAACAGATCGCAACCT | AAACACAACCTTCTTTGA | CCCTGCCG | 600 | | |
| Sbjct | 541 | CCGGCATCACAATACT | TCTAACAGAGCGAAACCT | AACACAACCTTCTTTGA | CCCTGCCG | 600 | | |
| Query | 601 | GAGGAGGAGACCCCAT | | TGATTCTTCG 645 | | | | |
| Sbjct | 601 | GAGGAGGAGACCCCAT | CCTTTACCAACACTTATT | TGATTCTTCG 645 | | | | |

Fig.13- The pairwise BLAST alignment of Oreochromis urolepis COI sequences with Oreochromis mossambicus



Construction of Phylogenetic Tree



The samples in this study were divided into four main groups according to population using the phylogenetic tree created from the sequence data (cytochrome COI) of samples from the two populations.

The phylogenetic tree constructed is un-rooted. From a distant common ancestor, four branches formed represent two descendant groups; *Oreochromis* species (Clade 1) *Etroplus* species (clade 2) *Sardinella, Nematalosa* species (Clade 3), and *Auxis, Thannus* species (Clade 4). Clade 1 is further branched suggesting that *Oreochromis niloticus* has diverged from the sister taxa *Oreochromis urolepis* and *Oreochromis mossambicus*. Similar species were clustered under the same nodes in the phylogenetic tree, whereas dissimilar species were clustered under distinct nodes, revealing an identical evolutionary link among the species.

Conclusion

The current study analyzed the molecular and phylogenetic relationship between 9 different species of fishes which can be classified into *Etroplus, Sardinella, Nematalosa, Thunnus, Auxis,* and *Oreochromis.* Our result revealed that these 9 species can be differentiated by their external appearance which can be verified using mitochondrial COI sequence analysis and the relationship can be suggested by the construction of a phylogenetic tree. From the results obtained, the

mitochondrial COI gene can be named as an ideal region for "species barcode" and it has high efficiency in species identification [13]. So, this tool is used for rapid analysis for commercial purposes especially confirmation for the particular species

The knowledge of the biochemical composition of any edible organism is extremely important since the nutritive value is reflected in its biochemical contents [16]. In these experimental studies, we measured protein; carbohydrates, and lipids content in our sample. This analysis is required as the demand for protein-rich food is increasing, especially in developing countries [17]. This study it's revealed that *Sardinella longiceps* have higher protein, carbohydrate, and lipid compared with the other 3 species of *Etroplus suratensis*, *Etroplus maculatus*, and *Nematalosa nausus*.

The cooked or raw products, pickles, and sausages are readily available in tins, cans, etc. So, knowing the quality of the by-products is very important. To analyze the quality of the by-products we can design the internal primer from the already analyzed sequences. Thus, this study can be used as a tool for the identification of inferior adulteration in fish by-products.

In addition to identifying fish species, DNA barcoding using the CO1 gene can also provide valuable information about evolutionary relationships between fish species and populations, as well as biogeographical patterns of fish distribution.

References

- F. Jim et al., "Comparative analysis of nutritional value of Oreochromis niloticus (Linnaeus), Nile tilapia, meat from three different ecosystems," J. Food Qual., pp. 5-10, 2017.
- [2] Rust. BM, Nutritional Physiology, 2003, pp. 367-452.
- [3] J. H. Tidwell and G. L. Allan, "Fish as food: Aquaculture's contribution. Ecological and economic impacts and contributions of fish farming and capture fisheries," *EMBO Rep.*, vol. 2, no. 11, pp. 958-963, 2001 [doi:10.1093/embo-reports/kve236].
- [4] K. G. Padmakumar et al., "Etroplus suratensis (Bloch), the State Fish of Kerala," J. Biosci., vol. 37, no. 6, pp. 925-931, 2012 [doi:10.1007/ s12038-012-9271-x].

- [5] Gunawickrama and K. B. Suneetha, "Morphological heterogeneity and population differentiation in the Green chromide in Sri Lanka," *Ruhuna J. Sci.*, vol. 2, pp. 70-81, 2007.
- [6] Shyam et al., "Sardine economy of Kerala: Paradigms and Perspectives," *Int. J. Fish. Aquat. Stud.*, vol. 2, no. 6, pp. 351-356, 2015.
- [7] P. V. Kagwade, "The food and feeding habits of the Indian oil sardine Sardinella longiceps Valenciennes," *Indian J. Fish.*, vol. 11a, no. 1, pp. 345-370, 1964.
- [8] R. Froese and D. Pauly, 2008, "FishBase. World Wide Web electronic publication," *Version* (02/2009). Available at: http://www.fishbase.org.
- [9] E. G. Silas and P. P. Pillai, *Resources of Tunas and Related Species and Their Fisheries in the Indian Ocean*, 1982, pp. 10-15.
- [10] R. Froese, 2003, "FishBase. World Wide Web electronic publication," *Version* (02 / 2009). Available at: http://www.fishbase.org.
- [11] C. D. Webster and C. Lim, *Tilapia: Biology, Culture, and Nutrition,* 2006, pp. 5-6.
- [12] V. Sachithanandam et al., "DNA bar coding, phylogenetic diversity studies of Etroplus suratensis fish from Pooranankuppam brackish water, Puducherry," *Int. J. Curr. Res. Rev.*, vol. 04, pp. 33-42, 2012.
- [13] X. Bingpeng et al., "DNA barcoding for identification of fish species in the Taiwan Strait,", *PLoS One*, vol. 13, no. 6, pp. e0198109, 2018 [doi:10.1371/journal.pone.0198109].
- [14] J. Zhang and R. Hanner, "DNA barcoding is a useful tool for the identification of marine fishes from Japan," *Biochem. Syst. Ecol.*, vol. 39, no. 1, p. 31-42, 2011 [doi:10.1016/j.bse.2010.12.017].
- [15] J. K. Chambers et al., "AG protein-coupled receptor for UDPglucose," J. Biol. Chem., vol. 275, no. 15, p. 10767-10771, 2000 [doi:10.1074/jbc.275.15.10767].
- [16] R. Nagabhushanam and U. H. Mane, "Seasonal variation in the biochemical composition of *Mytilus edulis* at Ratnagiri on the west coast of India," *Hydrobiologia*, vol. 57, no. 1, pp. 69-72, 1978 [doi:10.1007/BF00018630].

 [17] Gopalsamy., Idayachandiran., Arumugam., Muthukumar., Kumaresan., Saravanan., Thangavel., Balasubramanian. (2014).
 Nutritional Value of Marine Bivalve, *Donax cuneatus* (Linnaeus, 1758) from Cuddalore Coastal Waters, Southeast Coast of India. *Invention*. 1-10.