ISSN 1682-296X (Print)
ISSN 1682-2978 (Online)

Bio Technology

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Biotechnology 9 (3): 373-377, 2010
ISSN 1682-296X

DNA Barcoding of *Stolephorus indicus*, *Stolephorus commersonnii* and *Terapon jarbua* of Parangipettai Coastal Waters

[1]S. Ajmal Khan, [1]P.S. Lyla, [1]B. Akbar John, [1]C. Prasanna Kuamr, [1]S. Murugan and [2]K.C.A. Jalal
[1]Centre of Advanced Study in Marine Biology, Annamalai University, Parangipettai-608 502, India
[2]International Islamic University Malaysia, Jalan Istana,
Bandar Indera Mahkota, 25200, Kuantan, Pahang, Malaysia

Abstract: Three teleost fishes, *Stolephorus indicus*, *Stolephorus commersonnii* and *Terapon jarbua* occurring in Parangipettai waters, were barcoded (sequenced) for 648 bp region of the mitochondrial cytochrome oxidase subunit I gene (COI) for phylogenetic analysis. All the possible barcode sequences of these three fishes were also extracted via FASTA format from NCBI (National Centre for Biotechnological Information). The genetic distances within the species was small compared to the genetic distances between the species i.e., the genetic distances within *S. indicus* was 7.67% and the average genetic distance between *S. indicus* and *S. commersonnii* was 9.11%. The genetic distance between the families Engraulidae and Terapontidae was 26.48% which was found higher than the genetic distance calculated within the family of Engraulidae (9.11%). This clearly showed that when organisms from distantly placed families are taken, the genetic distance increases. In the present study some phylogeographical signal was apparent in the data. In the neighbor-joining tree for all 10 sequences, two major groups were apparent: clade A comprising of *Terapon jarbua* and clade B comprising of *Stolephorus* species. The grouping pattern of clade A showed some phylogeographical signals. The barcode sequence of *Lates calcarifer* shuffled within other sequences during phylogram constructions was unambiguously placed as an out group in the phylogram. The clades after bootstrapping corresponded well with the expectations. We conclude that precise and accurate identification of *Stolephorus indicus*, *S.commersonnii* and *Terapon jarbua* could be performed using the barcode sequences of the mitochondrial DNA (in the COI gene) of these fishes.

Key words: Cytochrome oxidase subunit I, DNA barcoding, phylogeographical signal, genetic distance

INTRODUCTION

Mitochondrial DNA (mtDNA) analysis has been employed in the evolutionary study of the animal species for more than 30 years (Brown *et al*., 1979; Avise and Walker, 1999). Its higher mutational rate and lower effective population size than the nuclear DNA make mtDNA a powerful tool to probe for evolutionary studies. This fact provoked a proposal to standardize DNA-based species identification by analyzing a uniform segment of the mitochondrial genome. A library of sequences from taxonomically verified voucher specimens could be built with this approach which could serve as DNA identifiers for species, in short, DNA barcodes (Herbert *et al*., 2003). For animals, 648 bp segment of the mitochondrial gene cytochrome C oxidase I (COI), which can be readily recovered from diverse species with a limited set of primers, was declared as a DNA barcode (Kevin *et al*., 2007). In addition, by assigning specimens to known species, DNA barcoding can speed the discovery of new species, as large sequence differences in animal mtDNA generally signal species status. Since, marine fishes have been the subject of intensive taxonomic analysis, they provide an excellent opportunity to test the efficacy of barcode-based species delimitation. For this approach to be effective, it must be possible to distinguish between intraspecific and interspecific mtDNA variation. The simplest test is whether the genetic distance within the species is lesser than those between species.

About 510 species of marine fishes had been documented so far in and around Parangipettai coastal waters. In this study we have examined the phylogeography, intraspecific and interspecific variations in the genus *Stolephorus* and intraspecific/ phylogeography of *Terapon jarbua* (Forskal, 1775). *Stolephorus* species belong to Engraulidae family and are mostly distributed in the Indo-Pacific waters. Schools of these species are found in rivers and estuaries and

Corresponding Authors: S. Ajmal Khan, Centre of Advanced Study in Marine Biology, Annamalai University, Parangipettai-608 502, India
K.C.A. Jalal, Kulliyyah of Science, International Islamic University Malaysia, Jalan Sultan Ahmad Shah, Bandar Indera Mahkota, 25200, Kuantan, Malaysia

migrate out into deeper and more saline water for spawning. Being pelagic shoaling fishes they contribute their mite to the fishery wealth of India. *Terapon jarbua* belongs to Terapontidaefamily and is mostly distributed along Indo-Pacific region. They occur over shallow sandy bottoms and migrate to sea for spawning (www.fishbase.org).

Accurate and unambiguous identification of fish and fish products from eggs to adults is important in many areas. A wide variety of methods including protein, DNA based methods have been used for the generic identification of fishes (Sotelo *et al*., 2001). Conventional taxonomist differentiate *S. indicus* and *S. commorsonnii* by observing the maxilla tip which is pointed and reaching to or a little beyond the border of pre-operculum in former and extend beyond the hind border of pre-operculum in later which makes difficulty in accurate identification of these fishes among the non exerts. Differentiating *T. jarbua* from *T. puta* is also lead to ambiguity among the non experts which are differentiated conventionally by observing the longitudinal body strips.

As, COI has been proposed as a barcode gene for most of the eukaryotes (Herbert *et al*., 2003). in the present study the COI diversity within and among the marine fishes *Stolephorus indicus* (Van Hasselt, 1823), *Stolephorus commersonni* (Lacepede, 1803) and *Terapon jarbua* (Forskal, 1775) collected from Parangipettai waters with that of available sequences from NCBI (National Center for Biotechnological information) was compared to study the phylogeographical signals, in the barcode region and to find out the efficiency of COI in identifying the selected species without ambiguity.

MATERIALS AND METHODS

Two specimens of *S. indicus* were collected alive from two stations, viz., Vellar estuary (lat. 11°30'N long. 79°47'E) and Cuddalore inshore waters (lat. 11°41'N long.79°50'E). *S.commersonii* was collected alive from Cuddalore inshore waters and *T. jarbua* was collected alive from Vellar estuary on June 2009.

Tissue sub-samples were exercised from the lateral side of the fishes and stored in 95% absolute alcohol. In order to optimize the DNA preservation, the tissue samples were re-suspended in fresh 95% absolute alcohol after few days of storage. DNA isolation was carried out using the DNA isolation kit provided by Bioserve Biotechnologies Pvt. Ltd., Hyderabad, India.

Approximately 450-bp was amplified from the 5' region of COI gene from the mitochondrial DNA using the following primers (Herbert *et al*., 2003).

- Fish F1- 5' TCAACCAACCACAAAGACATTG GC AC 3'
- Fish R1- 5' TAGACTTCTGGGTGGCCAAAGAATCA 3'

The 25 μL PCR reaction mixes included 18.75 μL of ultrapure water, 2.25 μL of 10X PCR buffer, 1.25 μL of $MgCl_2$ (50 mM), 0.25 μL of each primer (0.01 mM), 1.25 μL of each dNTP (0.05 mM), 0.625 U of Taq polymerase and 0.5-2 μL of DNA template. Amplification was performed using Beckman's PCR workstation (Bioserve Biotechnologies Pvt. Ltd. Hyderabad, India). The PCR program consisted of an initial step of 2 min at 95°C followed by 35 cycles of 0.5 min at 94°C, 0.5 min at 54°C and 1 min at 72°C, followed in turn by 10 min at 72°C and then held at 4°C. PCR products were visualized on 2% China agarose gel. Products were labeled using Qiagen sequencing kit and sequenced unidirectionally using a MegaBace capillary sequencer. The barcode sequences of all three fishes from other waters were extracted from NCBI (National Centre for Biotechnological Information) for interspecific and intraspecific analysis.

Sequences were aligned using Clustal X v.2.0 software (Thompson *et al*., 1997). Sequence divergences were calculated using the Kimura two parameter (K2P) distance model (Kimura, 1980). Neighbour-Joining (NJ) trees of K2P distances were created to provide a graphical representation of the patterning of divergence between species (Saitou and Nei, 1987). Phylogram was constructed with bootstrapping of 1000 replication through MEGA 4.0 (Tamura *et al*., 2007).

All four species of fishes gave perfect amplification with the primers used. The concentrations of amplicon were ideal for the sequencing reaction.

RESULTS

A total of 4 sequences (two sequences of *Stolephorus indicus*, one each of *S. commersonii* and *Terapon jarbua*) were obtained. By scanning through NCBI (National Centre for Biotechnological Information), a barcode sequence each for *S. indicus* (EU541321) and *S. commersonnii* (EU541323) was extracted. About 4 barcode sequences of *T. jarbua* were extracted from NCBI generating totally of 10 barcode sequences for analysis.

(A) Genetic distance: Read length (assembled short sequences in blocks of about 500 bp size) were all about 450 nt long. No insertion, deletions or stop codons were observed in any sequence. The average genetic distance (K2P distance) within *S. indicus* was 7.67% and average

Table 1: Distribution of K2P distance (percent) for COI within the three marine fishes barcoded

Organism name	1	2	3	4	5	6	7	8	9
S. commersonnii 1									
EU541323									
S. comersonnii 2									
IOBML123	22.1								
S. indicus 3									
EU542321	2.0	1.4							
S. indicus 4									
IOBML126	24.9	7.5	5.9						
S.indicus 5									
IOBML48	20.0	1.4	0	5.9					
T. jarbua 6									
EF607573	31.6	25.3	25.3	25.4	25.3				
T. jarbua 7									
EF607574	26.2	25.2	25.2	28.2	25.2	22.4			
T. jarbua 8									
EF607580	26.2	25.2	25.2	28.2	25.2	22.4	0.7		
T. jarbua 9									
FJ265859	29.3	26.3	26.3	26.4	26.3	3.7	22.4	22.4	
T. jarbua 10									
IOBML7	29.3	26.3	26.3	26.4	26.3	3.7	22.4	22.4	0

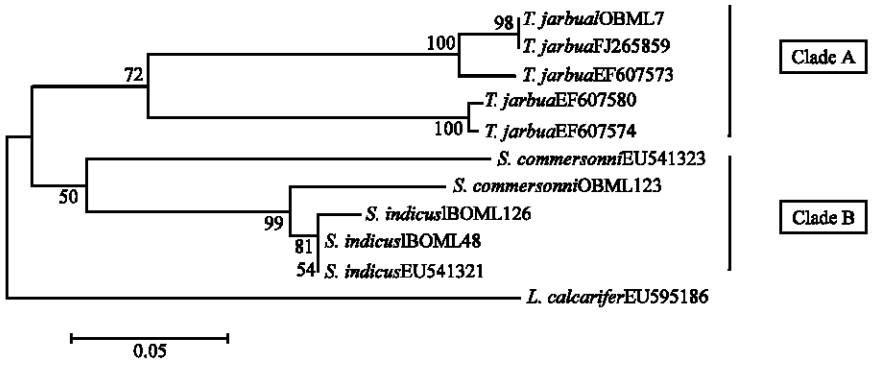

Fig. 1: Neighbour-joining tree (with bootstrap values; 500 replicates) of 11 COI sequences from 4 marine fishes, constructed using K2P distance. The specimen code IOBML corresponds to Indian Ocean Barcoding of Life Database (BoLD, www.barcodinglife.org). *Lates calcarifer* was selected as an out group of the phylogram. Scale bar: 0.05 substitutions per site

genetic distance within *S. commersonii* was 11.05%. The average of K2P distance between *S. indicus* and *S. commersonni* was 9.11% (Table 1).

Intra-specific genetic distance within *Terapon jarbua* was 14.25%. The average of genetic distance between the species of genera *Stolephorus* and *Terapon* was 26.48%.

(B) Inter and Intra-specific analysis: The phylogram (Fig. 1) generated through NJ method using K2P distance was highly reliable as the out group used (*Lates calcarifer*) was segregated in separate clade as expected. All the clades of the phylogram were reliable as the least bootstrap values after 500 replication was 50. The ability of COI in distinguishing two different genera was proved through the phylogram as two different clades (Clades A and B) were distinguishable. Bit of phylogenetic signals were evident in the clade A as the *T. jarbua* of Parangipettai waters was segregated from the *T. jarbua* of other waters along with few other related strains.

DISCUSSION

The usage of K2P genetic distance for analyzing the data revealed that within the species barcode variation is low compared to the sequence variation between the species. For example; the average genetic distance of *S. indicus* IOBML126, *S. indicus* IOBML48 and *S. indicus* EU541321 was 7.67%, whereas, the genetic distance between the *S. indicus* and *S. commersonnii* (IOBML123 and EU541323) was 9.11%. Further, when we moved across the family the genetic distance increased. Example: the average genetic distance within *Terapon jarbua* was 14.25% whereas, the average genetic distance within the genus *Stolephorus* was 9.11%. But the average genetic

distance between *Terapon jarbua* and *Stolephorus* sp., was 26.48% (Table 1). Thus, the genetic distance keeps increasing when the sequences of fishes belonging to different genera, families and orders are compared. This proves the efficacy of barcode region of COI gene in species identification. Similar observation was observed in various studies proved that the genetic distance within the genus is lesser than between the genus (Avise *et al*., 1987; Black *et al*., 1997; Blaxter and Floyd 2003; Borda and Siddall, 2004; Brower, 1999; Herbert *et al*., 2003).

The ability of COI in distinguishing the two different species was evident from the clade B, though the *S. commersonii* IOBML123 was misplaced inside the clade of *S. indicus*. This might be due to the usage of shorter segment (450 bp) of barcode region for phylogram construction than Herbert's recommended length (648 bp). The phylogeography was not seen in the barcode sequences of *S. indicus* in clade B as *S. indicus* from the closest geography (*S. indicus* IOBML126 and *S. indicus* IOBML48) was segregated differently. The bootstrap values of all clades corresponded well without exceptions.

Although, barcode analysis seeks only to delineate species boundaries, there are clearly some phylogeographical signals in COI sequence data. For example: clade A was by the COI data of same species (*T. jarbua*) occurring in the waters of India and China. *T. jarbua* (IOBML7, FJ265859) from India was segregated in separate group with few *T. jarbua* from China waters and the other sequences (2) of *T. jarbua* of China waters got grouped separately. An exception was noted in clade B, where the *S.commersonii* IOBML123 was misplaced in the clade of *S. indicus*. This might be due to the usage of shorter sequences (450 bp) of barcode region for phylogram construction than Herbert's recommended length (648 bp). So in such cases recovery of the entire barcode region (648 bp) for phylogram construction would be desired to avoid such kind of exception (Herbert *et al*., 2003). However, methodologies for phylogeny construction, from the DNA sequence data remain somewhat controversial, with wide variety of disparate approaches possible (Nei and Kumar, 2000).

It may not be possible to recover the true phylogeny of fishes from a 450 nt fragment of mitochondrial DNA through K2P distance and neighbour joining- rather more gene regions should be used, including nuclear genes (Ward *et al*., 2005). Barcoding discriminated all the 3 fish species we examined and would clearly be capable of unambiguously identifying individually isolated fish eggs, larvae, fillets and fins from these species. Bootstrap values of all the clades were on higher side as the minimum of 50 was noted at clade B. Bootstrapping of cent percent had been noted frequently at the nodes of clade A. This study has strongly validated the efficacy of COI barcodes for identifying *S. indicus*, *S. commersonii* and *T. jarbua* of Parangipetatai and other international waters. All three fishes were readily amplified with the primers FIshF1 and FishR1. No failure in amplification was noted in any instance.

CONCLUSION

Present result revealed that COI barcoding will permit the unambiguous identification of the vast majority of fish species. We now intent to extend our survey to all fishes and other important groups of marine fauna of Parangipettai coastal waters. This will mean collecting sequences from at least 500 species of marine fishes occurring in Parangipettai waters. There might also be instances of supposedly distinct species that are having identical COI sequences, suggesting the possibility of species fusion. The resolution of cases of this nature will require careful morphological analysis from expert taxonomists before any final recommendations are made. Barcoding and morphological analysis should go hand-in-hand.

Once a global COI barcode database has been established for fishes, anyone with direct or indirect access to a DNA sequencer will be able to identify, to a high degree of certainty, any fish egg and larva or carcass fragment. This will be an invaluable tool for fisheries managers, fisheries ecologists and fish retailers and for those wishing to develop fish identification microarrays. Thus scientific and practical benefits of fish barcoding will be manifold.

ACKNOWLEDGMENT

The authors are thankful to Dr. T. Balasubramanian, Professor and Director, Centre of Advanced Study in Marine Biology, Parangipettai for the encouragement and the university authorities for the facilities.

REFERENCES

Avise, J.C. and D. Walker, 1999. Species realities and numbers in sexual vertebrates: perspectives from an asexually transmitted genome. Proc. National Acad. Sci. USA., 96: 992-995.

Avise, J.C., J. Arnold, R.M. Ball, E. Bermingham and T. Lamb *et al*., 1987. Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. Ann. Rev. Ecol. Syst., 18: 489-522.

Black, M.B., K.M. Halanych, P.A.Y. Maas, W.R. Hoeh and J. Hashimoto *et al*., 1997. Molecular systematics of vestimentiferan tubeworms from hydrothermal vents and cold-water seeps. Mar. Biol., 130: 141-149.

Blaxter, M. and R. Floyd, 2003. Molecular taxonomics for biodiversity surveys: Already a reality. Trends Ecol. Evol., 18: 268-269.

Borda, E. and M.E. Siddall, 2004. Arhynchobdellida (Annelida: Oligochaeta: Hirudinida): Phylogenetic relationships and evolution. Mol. Phylogenet. Evol., 30: 213-225.

Brower, A.V.Z., 1999. A critique of Davis and Nixon's population aggregation analysis. Syst. Biol., 48: 199-213.

Brown, W.G., M.Jr. George and A.C. Wilson, 1979. Rapid evolution of animal mitochondrial DNA. Proc. National Acad. Sci. USA., 76: 1967-1971.

Herbert, P.D.N., A. Cywinska, S.L. Ball and J.R. Dewaard, 2003. Biological identification through DNA barcodes. Proc. Biol. Sci., 270: 313-321.

Kevin, C., R. Kerr, M.Y. Stoeckle, C.J. Dove, L.A. Weigh, C.M. Francis and P.D.N. Hebert, 2007. Comprehensive DNA barcode coverage of North American birds. Mol. Ecol. Notes, 7: 535-543.

Kimura, M., 1980. A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. J. Mol. Evol., 16: 111-120.

Nei, M. and S. Kumar, 2000. Molecular Evolution and Phylogenetics. Oxford University Press, New York.

Saitou, N. and M. Nei, 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol. Biol. Evol., 4: 406-425.

Sotelo, C.G., P. Calo-Mata, M.J. Chapela, R.I. Perez-Martin and H. Rehbein *et al*., 2001. Identification of flatfish (Pleuronectiformes) species using DNA-based techniques. J. Agric. Food Chem., 49: 4562-4569.

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

Thompson, J.D., T.J. Gibson, F. Plewniak, F. Jeanmougin and D.G. Higgins, 1997. The ClustalX windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res., 24: 4876-4882.

Ward, R.D., T.S. Zemlak, H.I. Bronwyn, R.L. Peter and P.D.N. Hebert, 2005. DNA barcoding Australia's fish species. Philos Trans. R. Soc. Lond B Biol. Sci., 360: 1847-1857.