

Character-based identification system of scombrids from Indian waters for authentication and conservation purposes

Sonalismita Mahapatra, Rathipriya A., Priyadarshini Dwivedy, Suresh E., Shanmugam S. A. and Kathivelpandian A.

Institute of Fisheries Postgraduate Studies, TNJFU OMR Campus, Vaniyanchavadi, Chennai

ABSTRACT

Scombrids are the important component of pelagic fishery resources which include 54 species under 15 genera commonly known as mackerels, bonitos, and tunas. Due to the high commercial value attained, there are real chances of fraudulent substitution by species of inferior value. DNA based species identification methods can be applied to detect product adulteration, as well as to better contribute to the conservation and management of these species by providing accurate species identification independently of the age of the individuals or the tissue processed. In this study, a total of 15 commercially important scombrid species from Indian waters were analyzed. Due to the inadequacy of mitochondrial COI barcoding gene in discriminating between some *Thunnus* species, cytochrome b sequences were used instead. For all the 15 species, we propose a DNA character-based keys which uses a diagnostic combination of nucleotides and respective probes, including the first character-based keys and probes to differentiate between *Thunnus albacares* and *T. obsesus*.

ARTICLE HISTORY

Received 3 May 2020
Accepted 11 August 2020

KEYWORDS

Scombrids; CBIS; species-specific probes; BLOG 2.0; authentication

Introduction

Seafood is regarded as both an important protein source and a delicacy for human populations across the globe, and is mainly obtained from pelagic fishery resources. Among pelagic fishes, scombrids, commonly known as mackerels, bonitos, and tunas are the major component with 54 species under 15 genera. The total marine fish landings of Indian waters is 3.49 million tonnes to which the pelagic ecosystem contributes 54% (FRAD, CMFRI 2018). Out of the 21 species reported in Indian waters (Venkataraman and Sivaperuman 2014), eight species alone contribute to over 99% of the fisheries (CMFRI 2015). Scombrids contribute 10% of the total marine landings of Indian waters (FRAD, CMFRI 2018). Among scombrids, seer fishes are one of the important exploited resources in the country yielding a total 49,209 tonnes and comprising 24.8% of the total large pelagic fishery with high domestic and exports value (CMFRI 2015).

Due to the high price attained, scombrids are prone to get adulterated with other fishes (*Pangasius* species), in some restaurants and processing industries (Danni 2015). In these conditions, fish identification through the traditional methods such as morpho-meristic measurements (Strauss and Bond 1990) and otoliths (Granadeiro and Silva 2000) is difficult. Further, eggs and larvae (especially those of the genus *Thunnus*) generally cannot be distinguished easily by morphological characters (Richards et al. 1990) for the purpose of conservation and management of these important fishery resources. DNA-based identification methods have been

developed to address these issues. In most cases, the mitochondrial COI gene is used for fish identification and authentication purposes (Lakra et al. 2011; Bingpeng et al. 2018). Several studies have been carried out to utilize the mitochondrial genes for species identification purposes. There were studies using molecular tools for different groups of scombrid fishes. Molecular identification of scombrids has been ascertained by Miya et al. (2013) using complete whole mitogenome analysis. Similarly, the complete mitochondrial genome sequence of *R. kanagurta* is also studied by Chen et al. (2013) for systematic studies. The differentiation of tuna products using 464 bp of *Cyt b* sequences was reported by Abdullah and Rehbein (2015a). Wherein, speed and accuracy for authentication purposes demands alternate methods than PCR sequencing.

Furthermore, the limitations of using full-length sequences which might require technical expertise, and due to lack of a standard short-barcode fragment, most of the character-based studies have targeted the full-length COI barcode (Mutanen et al. 2015), although extraction of long DNA fragment is difficult in cases of archival specimen and processed biological material (Wandeler et al. 2007). Due to the inadequacy of mitochondrial COI gene in discriminating *Thunnus* species (Puncher et al. 2015), it is needed to supplement alternative mitochondrial genes such as cytochrome b (*Cyt b*) which is used commonly for authentication purposes (Cutarelli et al. 2018). Further, it is also necessary to develop alternative methods such as character-based identification

systems as effective tools for authentication and conservation purposes. Character-based identification system (CBIS) is a very useful tool in the identification of fish species using specific locations of nucleotide sequences of the species under study. The probes developed through CBIS will be of immense use for authenticity purposes to detect the substitution of low-value fishes with highly priced ones.

In this context, this study is aimed at developing species-specific nucleotide keys through character-based identification system (CBIS) which could be used in developing a DNA chip for authentication purposes of commercially important scombrid species from Indian waters.

Materials and methods

Data collection

A total of 1203 mitochondrial cytochrome b sequences of commercially important species of scombrids in India were collected from GenBank, namely: *Acanthocybium solandri*, *Auxis thazard*, *Rastrelliger kanagurta*, *Thunnus tonggol*, *Auxis rochei*, *Euthynnus affinis*, *Katsuwonus pelamis*, *Rastrelliger brachysoma*, *Rastrelliger faughni*, *Scomberomorus commerson*, *Scomberomorus guttatus*, *Sarda orientalis*, *Thunnus alalunga*, *Thunnus obesus*, and *Thunnus albacares* (Supplemental online material). The number of sequences per species ranged from 15 to 284. A total of 776 sequences from 15 species were used for further analysis after checking quality and authenticity.

Data analysis

Alignment of downloaded *Cyt b* sequences was done in Bio edit ver. 7.0.5.2 (Hall 1999) using the ClustalW tool (Thompson et al. 1997). Trimming of aligned sequences was done for making the downloaded sequences in uniform length for further analysis. BLOG 2.0 (Weitschek et al. 2012) was used for character-based species classification of scombrid group of fishes. Classification of scombrids and the quality control of the designed probes was checked with MFEprimer 3.0 server (Wang et al. 2019).

Results

The sequences downloaded from NCBI were in the length range of 308–1141bp. The resulted length after alignment and trimming of sequences was of 640bp. BLOG analysis yielded a total of 25 diagnostic nucleotides, with two diagnostic nucleotides for nine species and a single diagnostic nucleotide for the remaining six species. A total of 25 probes (18–37bp) were designed for 15 species, having 9 species with two probes and 6 species with a single probe (Table 1).

Discussion

For commercially important group of fishes such as scombrids, it is essential to derive the method for fast and accurate identification to manage the adulteration. Scombrids

possess high degree of chance of mislabeling with other cheap fish species as morphological identification is not possible in processed fish specimens. Recent studies established the utility of CBIS as an appropriate method for identification purposes. Several studies have been carried out to generate the character-based keys (Paine et al. 2007; Lowenstein et al. 2009; Bergmann et al. 2009; Puncher et al. 2015; Vargheese et al. 2019). In order to identify the mislabeling of high-value Bluefin tuna, Lowenstein et al. (2009) developed a character-based identification system using mitochondrial COI gene for diagnosing *Thunnus* species. Similarly, Paine et al. (2007) used a character-based identification system for 17 species of scombrids from the western Atlantic Ocean. Molecular identification of Atlantic blue fin tuna larvae using *Cyt b* is also studied by Puncher et al (2015). In this study, we developed 25 positions of character-based keys for identification of scombrids from Indian waters which is in similar line with Lowenstein et al. (2009) who described 40 positions for differentiating the blue fin tuna.

CBIS could be done using several software systems such as CAOS (Bergmann et al. 2009), BLOG (Weitschek et al. 2013), etc. In this study, BLOG (Weitschek et al. 2013) is used as it has several advantages over tree-based methods and similarity methods (Van Velzen et al. 2012). Vargheese et al. (2019) used BLOG 2.0 for differentiating a total of 82 species of elasmobranchs and observed 214 diagnostic nucleotides when compared to 25 positions for scombrids from Indian waters.

In the recent years, increase in trade of sea foods leads to an increase in the occurrence of mislabeling and also fish adulteration is common by some industries. In Brazilian fish markets, 80% of the surubim (*Pseudoplatystoma* spp.) products were mislabeled (Carvalho et al. 2011). Similarly, in Italian fish markets commercialized shark samples got mislabeled by some local fish species (Scarano and Rao 2014). Abdullah and Rehbein (2015b) amplified parvalbumin intron by exon-primed intron-crossing PCR for differentiating scombrids and possible substituted species such as catfish, tilapia and snapper fish species. Characterization of amplicons was done through sequencing (tunas), single-strand conformation polymorphism (SSCP) (scombrid, catfish, tilapia, and snapper species) and restriction fragment length polymorphism (RFLP) (catfish). The SSCP method differentiated catfish, tilapia, snapper, and scombrid species except tunas. Tunas of the genus *Thunnus* had an unexpected low variability of intron sequences, which prevented their differentiation by sequencing or SSCP. Lin and Hwang (2007) developed RFLP with two sets of primers to amplify 126 and 146 bp of partial mitochondrial cytochrome b gene, and five restriction enzymes, Bsp1286I, HincII, RsaI, Scal, and MbolI were determined to analyze the short length fragments to authenticate 18 canned tunas. Despite several methods used for differentiating *Thunnus* sp, still there is a need for fast, reliable, and low-cost methods to identify processed tuna and to differentiate other scombrid species for supporting sustainable fisheries and tuna trade. In this context, the species-specific probes for 15 species of scombrids obtained in this study will be of immense help to address the similar mislabeling concerns in the commercially important scombrid group of fishes especially tunas. The

Table 1. Species-specific formula and probes of scombrid fish species identified through CBIS.

Sl. No.	Species name	Species formula	Probe	Ta
1	<i>Acanthocybium solandri</i>	pos81 = A	CCCAAATCCTCACAGGCTATTCC	54.5 °C
2	<i>Auxis thazard</i>	pos153 = T	GCCTATTCCTTGAATACACTACACCCC	56.4 °C
3		pos84 = C	CCCTCCAATATTTCCGCATGATGAAACT	54.8 °C
4	<i>Rastrelliger kanagurta</i>	pos177 = T	ATACACTACACTCCCGATGTTGAATCAGCA	56.6 °C
5		pos261 = T	ACGCAAATGGCGCTTCTTTCTTTT	56 °C
6	<i>Thunnus tonggol</i>	pos45 = T	CTAGTTGACCTTCTACCCCTCTAATATT	54 °C
7	<i>Auxis rochei</i>	pos66 = G	GTCATGATAACTGCATTGCTCGGCTA	54.7 °C
8	<i>Euthynnus affinis</i>	pos102 = A	CATGATGAAACTTTGGATCACTGCTTGG	54.3 °C
9		pos108 = G	TGGCAACTGCTTGGTCTCTGCCTTA	56.8 °C
10	<i>Rastrelliger brachysoma</i>	pos28 = G	ACACTCCCGATGTTGAATCAGCAT	54.1 °C
11	<i>Katsuwonus pelamis</i>	pos177 = A	TATACCCTGACGTAGAATCAGCCTT	54 °C
12		pos246 = C	ATCCGAAACCTCCATGCCAACG	55.2 °C
13	<i>Rastrelliger faughni</i>	pos168 = C	ATCCATACACCCCGACGTTGA	54.8 °C
14		pos258 = T	ACGCAAACGGTGCCTCCTTC	55.7 °C
15	<i>Scomberomorus commerson</i>	pos246 = G	GCAGATTCGATGTCGGGGTA	53.9 °C
16	<i>Scoberomorus guttatus</i>	pos246 = T	TCCGAAACCTTCAGCAAATGG	53.7 °C
17		pos258 = A	CAGCAAATGGAGCATCCTTCTTCTT	55.4 °C
18	<i>Sarda orientalis</i>	pos81 = C	CCCCTTCCAATATCTCTGCATGATGAAATTT	54.7 °C
19		pos258 = C	GCAAACGGCGCTCCTTTTTC	55.5 °C
20	<i>Thunnus alalunga</i>	pos258 = A	TCCACGCAAACGGGGCC	55.8 °C
		pos261 = T	CTGAATGATACTTCTTATTTGCTTACGCAATCT	54.8 °C
21	<i>Thunnus obesus</i>	pos261 = G	CGGATTAACACTCAAATGCAGATAAAATCTCATTCCA	55.1 °C
22	<i>Thunnus albacares</i>	pos81 = C	ACAAAGGCTTGGTCTGACTTTACTGT	55.6 °C
23		pos258 = G	ATGCTTTCTGAGGAGCTACCGTCATTA	54.5 °C

developed probes could be used to develop DNA chip for faster and accurate identification of scombrid fish species.

DNA chip is well established method to differentiate multiple species at a particular point of time for food and forensic applications (Lettieri 2006). A DNA chip was developed by Kochzius et al. (2008) for 11 commercially important fish species using 16S rDNA gene.

DNA chip could be of immense use wherein a large number of species-specific nucleotides could be spotted in small space for identifying target species accurately (Kim et al. 2011). The probes developed in this study will be used for curbing the mislabeling and adulteration of commercially important scombrids apart from conservation and management of these valuable resources from Indian waters.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Data availability statement

The availability of data used in the present study is in the NCBI database (<https://www.ncbi.nlm.nih.gov/nucleotide/>). The nucleotide information taken from NCBI were analyzed using various molecular tools and the findings are presented in the manuscript. The details are given as supplementary Table 1.

References

- Abdullah A, Rehbein H. 2015a. The differentiation of tuna (family: Scombridae) products through the PCR-based analysis of the cytochrome b gene and parvalbumin introns. *J Sci Food Agric.* 96(2): 456–464.
- Abdullah A, Rehbein H. 2015b. Authentication of closely related scombrid, catfish and tilapia species by PCR-based analysis and isoelectric focusing of parvalbumin. *Eur Food Res Technol.* 241(4):497–511.
- Bergmann T, Hadrys H, Breves G, Schierwater B. 2009. Character-based DNA barcoding: a superior tool for species classification. *Berl Munch Tierarztl Wochenschr.* 122(11-12):446–450.
- Bingpeng X, Heshan L, Zhilan Z, Chunguang W, Yanguo W, Jianjun W. 2018. DNA barcoding for identification of fish species in the Taiwan Strait. *PLoS One.* 13(6):e0198109.
- Carvalho DC, Neto DAP, Brasil BSAF, Oliveira DAA. 2011. DNA barcoding unveils a high rate of mislabeling in a commercial freshwater catfish from Brazil. *Mitochondrial DNA.* 22(sup1):97–105.
- Central Marine Fisheries Research Institute (CMFRI). 2015. CMFRI Annual Report 2014–2015. Thoothukudi, India.
- Cutarelli A, Galiero G, Capuano F, Corrado F. 2018. Species identification by means of mitochondrial cytochrome b DNA sequencing in processed anchovy, sardine and tuna products. *FNS.* 09(04):369–375.
- Chen Y, Cheng Q, Qiao H, Zhu Y, Chen W. 2013. The complete mitochondrial genome sequence of *Rastrelliger kanagurta* (Perciformes: Scombridae). *Mitochondrial DNA.* 24(2):114–116. DOI: [10.3109/19401736.2012.726624](https://doi.org/10.3109/19401736.2012.726624)
- Danni W. 2015. The use of *Pangasius* fishes in restaurants [thesis]. Florida State University, College of Human Sciences. 1–58.
- FRAD, CMFRI. 2018. Marine fish landings in India 2017. Kochi: CMFRI, Technical Report.
- Granadeiro JP, Silva MA. 2000. The use of otoliths and vertebrae in the identification and size- estimation of fish in predator-prey studies. *Cybiurn.* 24: 4:383–393.
- Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser.* 41:95–98.
- Kim S, Koo H, Kim JH, Jung JW, Hwang SY, Kim W. 2011. DNA chip for species identification of Korean freshwater fish: a case study. *BioChip J.* 5(1):72–77.
- Kochzius M, N Olte M, Weber H, Silkenbeumer N, HjOrleifsdottir S, Hreggvidsson GO, Marteinsson V, Kappel K, Planes S, Tinti F, et al. 2008. DNA microarrays for identifying fishes. *Mar Biotechnol.* 10(2): 207–217.
- Lakra WS, Verma MS, Goswami M, Lal KK, Mohindra V, Punia P, Gopalakrishnan A, Singh KV, Ward RD, Hebert P. 2011. DNA barcoding Indian marine fishes. *Mol Ecol Resour.* 11(1):60–71.
- Lettieri T. 2006. Recent applications of DNA microarray technology to toxicology and ecotoxicology. *Environ Health Perspect.* 114(1):4–9.
- Lin WF, Hwang DF. 2007. Application of PCR-RFLP analysis on species identification of canned tuna. *Food Control.* 18(9):1050–1057.
- Lowenstein JH, Amato G, Kolokotronis SO. 2009. The real *maccoyii*: identifying tuna sushi with DNA barcodes – contrasting characteristic attributes and genetic distances. *PLOS One.* 4(11):e7866.
- Miya, M., Friedman, M., Satoh, T.P., et al. 2013. Evolutionary origin of the Scombridae (tunas and mackerels): Members of a Paleogene adaptive

- radiation with 14 other pelagic fish families, PLoS One 8(9):e73535: 1–19.
- Mutanen M, Kekkonen M, Prosser SW, Hebert PD, Kaila L. 2015. One species in eight: DNA barcodes from type specimens resolve a taxonomic quagmire. Mol Ecol Resour. 15(4):967–984.
- Paine MA, McDowell JR, Graves JE. 2007. Specific identification of western Atlantic Ocean scombrids using mitochondrial DNA cytochrome c oxidase subunit I (COI) gene region sequences. Bull Marine Sci. 80:353–367.
- Puncher GN, Arrizabalaga H, Alemany F, Cariani A, Oray IK, Karakulak FS, Basilone G, Cuttitta A, Mazzola S, Tinti F. 2015. Molecular identification of Atlantic Bluefin tuna (*Thunnus thynnus*, Scombridae) larvae and development of a DNA character-based identification key for Mediterranean scombrids. PLoS One. 10(7):e0130407.
- Richards LJ, Schnute JT, Hand CM. 1990. A multivariate maturity model with a comparative analysis of three lingcod (*Ophiodon elongatus*) stocks. Can J Fish Aquat Sci. 47(5):948–959.
- Scarano D, Rao R. 2014. DNA markers for food products authentication. Diversity. 6(3):579–596.
- Strauss RE, Bond CE. 1990. Taxonomic methods: morphology. Methods for fish biology. Bethesda, MD: American Fisheries Society.
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25(24):4876–4882.
- Van Velzen R, Weitschek E, Felici G, Bakker FT. 2012. DNA barcoding of recently diverged species: relative performance of matching methods. PLoS One. 7(1):e30490.
- Vargheese S, Chowdhury LM, Ameri S, A. A. m, A. K. 2019. Character based identification system for Elasmobranchs for conservation and forensic applications. Mitochondrial DNA Part A. 30(4):651–656.
- Venkataraman K, Sivaperuman C. 2014. Marine faunal diversity in India: taxonomy, ecology and conservation. San Diego, CA: Academic Press.
- Wandeler P, Hoeck PE, Keller LF. 2007. Back to the future: museum specimens in population genetics. Trends Ecol Evol (Amst). 22(12):634–642.
- Wang K, Li H, Xu Y, Shao Q, Yi J, Wang R, Cai W, Hang X, Zhang C, Cai H. 2019. MFEprimer-3.0: quality control for PCR primers. Nucleic Acids Res. 47(W1):W610–W613.
- Weitschek E, Van Velzen R, Felici G, Bertolazzi P. 2013. BLOG 2.0: a software system for character-based species classification with DNA Barcode sequences. What it does, how to use it. Mol Ecol Res. 13: 1043–1046.