Research Journal of Biological Sciences 5 (6): 414-419, 2010
ISSN: 1815-8846

DNA Barcoding of *Lates calcarifer* (Bloch, 1970)

[1]Akbar John, [1]C. Prasannakumar, [1]P.S. Lyla, [1]S. Ajmal Khan and [2]K.C.A. Jalal
[1]Centre of Advanced Study in Marine Biology, Annamalai University, Parangipettai, India
[2]International Islamic University Malaysia, Jalan Sultan Ahmad Shah,
Bandar Indera Mahkota, 25200 Kuantan, Pahang, Malaysia

Abstract: An effort to assess the utility of 650 bp Cytochrome oxidase subunit I (DNA barcode) gene in delineating the members closely resembled marine fin fishes belonging to the family Latidae was made. To address the issue, all 40 barcode sequences of fishes belonging to Latidae family available at NCBI (National Centre for Biotechnological Information) along with the barcode data generated from commonly available fish Asian sea bass (*Lates calcarifer*) from Parangipettai coastal waters was used. The average GC content of selected three fishes belongs to Latidae was found to be 48.87%. *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. Compare to Australian, Indian and Singapore (48.1-48.89%) water *L. calcarifer*, Myanmar and South China species shows high average GC content (49.38-49.41%). The phylogenetic and genetic distance data showed that the maximum genetic distance is present in 3rd codon position in all the selected species. It is apparent from the genetic distance that the GD between *L. calcarifer* and *L. niloticus* was observed to be higher than the GD between *P. waigiensis* and *L. niloticus* in all the codon positions. The expected clustering of the species in separate internal branches was well observed in the constructed phylogram with high bootstraps. We conclude that COI sequencing (barcoding) in identifying the morphologically similar species of the family Latidae.

Key words: Cytochrome oxidase subunit I, COI, Latidae, codon position, India, Malaysia

INTRODUCTION

About 31800 species of fin fishes have been identified from the aquatic environments which constitute >50% of all vertebrate species. The usefulness of DNA level polymorphism as a tool in fisheries science has been recognized for some time (Hallerman and Beckmann, 1988). It has long been recognized that DNA sequencing technology can be used to identify organisms up to the species (Tautz *et al*., 2002, 2003). DNA sequence analysis has been used for 30 years to assist species identification but different loci have been used for different taxonomic groups and in different laboratories. It was proposed that a single gene sequence would be sufficient to differentiate all or at least the vast majority of animal species and proposed the use of the mitochondrial DNA gene cytochrome oxidase subunit I (coxI) as a global bio identification system for animals (Hebert *et al*., 2003). Empirical support for the barcoding concept ranges from studies of invertebrates (e.g., spring tails and butterflies) to birds (Hebert *et al*., 2004a, b; Hogg and Hebert, 2004). However, the approach is not without controversy (Lipscomb *et al*., 2003; Moritz and Cicero, 2004). For the barcoding approach to species identification to succeed within species DNA sequence need to be more similar to one another than to the different species. In fact, the hybridization among species would create taxonomic uncertainty and since, mitochondrial DNA is maternally inherited any hybrid would have the maternal species DNA only.

Lates calcarifer, one of the nine *Lates* species of the family Latidae is widely distributed in the coastal and freshwater of the tropical Indo-West pacific waters (Nelson, 1994; Berra, 2001). *Lates calcarifer* (Order: Perciformes; Family: Latidae), a diadromous fish is found in coastal waters, estuaries and lagoons in clear to turbid waters of eastern edge of the Indo-West pacific region from the Arabian gulf to China, Taiwan province of China, Papua New Guinea and Northern Australia usually occurring at the depth of 10-40 m. A very popular fish with considerable economic importance is presently used for aquaculture in India, Thailand, Indonesia and Australia. This fish is becoming an important farmed marine food fish species with global annual production of nearly 400,000 metric ton according to Food and Agriculture Organization of the United Nations statistics.

Corresponding Author: Akbar John, Centre of Advanced Study in Marine Biology, Annamalai University, Parangipettai, India

They reach 1500-3000 g in one year in ponds under optimum conditions. It is more commonly confused with *Psammoperca waigiensis* in identification due to many overlapping morphological characters in both the fishes. 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).

L. calcarifer consists of smallest genome (~700 Mb) among marine food fish species (Wang *et al*., 2007). Loss of genetic diversity due to hatchery culture practice in barramundi, *L. calcarifer* was also reported recently (Frost *et al*., 2007). Chenowith *et al*. (1998) extensively studied the effect of environment changes on the mtDNA geneology and geography of *L. calcarifer* was evident. In the present study, the efficiency of mitochondrial gene, coxI in delineating closely resembling species belonging to the family Latidae has been examined in addition to the presence of geographical signals.

MATERIALS AND METHODS

Sampling, DNA extraction, amplification and sequencing: *Lates calcarifer* from Parangipettai coastal waters was colleted alive and transported to the laboratory where the right side of the fish was photographed and a cube of lateral muscle (5-7 mm) from left side of the fish was exercised for DNA isolation. The fish was preserved in 95% ethanol. Salting out protocol was adopted for precise and quick DNA isolation from the fish tissue (Miller *et al*., 1988).

The tissue was placed in 1.5 mL eppendorf tube and 500 μL of solution I (50 mM Tris-HCl pH 8, 20 mM EDTA pH8 and 2% SDS) was added. The tissue was homogenized with sterile homogenizer and 5 μL of Proteinase K (20 mg mL^{-1}) was added and quick vortexed. The sample was incubated at 55°C in water bath for 2 h with occasional mixing. Following incubation, the sample was chilled over ice for 10 min and 250 μL of solution II (6M NaCl) was added and inverted several times for thorough mixing.

The tube was chilled on ice for 5 min and centrifuged at 8000 rpm for 15 min. About 500 μL of supernatant was carefully collected in to new labeled 1.5 mL tube and twice the volume (i.e., 1 mL) of 100% AR grade ethanol was added to precipitate the DNA. The precipitate was pellet down at 8000 rpm for 5 min and the supernatant was removed without touching the pellet. The DNA pellet was rinsed with 500 μL of cold ethanol and centrifuged at 11000 rpm for 5 min. The supernatant was carefully removed and the excess liquid was drained using pipette.

The pellet was partially dried (devoid of ethanol) with lid off at 55°C on heating block. The pellet was re-suspended with 50-200 μL of fresh sterile H_2O depending on size of pellet (100 μL average) by gently pipetting sample with wide-bore filter tip until dissolved. This dissolved DNA acted as a template for Polymerase Chain Reaction (PCR). The fragment of COI was amplified by GeneAmp PCR system 9700. PCR was carried out in 25 μL volumes (2.5 μL of 10X PCR buffer, 1.5 μL of $MgCl_2$ (2 mM μL^{-1}), 1 μL of DNA template, 1 μL of each primer (10 p moles μL^{-1}), 2 dNTPs (1 mM μL^{-1}), 10 U of 1 μL of Taq polymerase (Bioserve Biotechnologies Pvt., Ltd., Hyderabad, India) and 15 μL of sterile Mill Q water).

Fish F1 [5'- TCAACCAACCACAAAGACATTGGC AC-3'] and Fish R1 [5'-TAGACTTCTGGGTGGCCAAAGA ATCA-3'] primers were employed for COI amplification. The thermocyclic conditions for PCR included the initial denaturation at 94°C for 1 min, 5 cycles of 94°C for 30 sec, annealing at 45°C for 40 sec and extension at 72°C for 1 min with a final extension at 72°C for 10 min followed by indefinite hold at 4°C.

Following PCR, about 10 μL of PCR product with 2 μL of bromo thymol blue were added to 2% agarose gel, prepared with 2.5 μL of 1% ethidium bromide and electrophorized at 90 V until the dye moved for 6 cm in the gel. The gel was moved to gel doc system for viewing the amplicons with the aid of UV trans-illuminator. Sequencing PCR was carried out using dye terminator mix (v3.1) and quantified in Euro bio-agarose gel. The samples were loaded onto MegaBace sequencer at Bioserve Biotechnologies, Pvt. Ltd. Hyderabad, India.

Sequence data analysis: The electrophenerogram generated by automated DNA sequencer was read by Chromas Pro (v1.42) and the sequences were carefully checked for mis-calls and base spacing. About 31 barcode sequences of *Lates calcarifer*, 5 sequences of *Psammoperca waigiensis* and same number of sequences of *Lates niloticus* were extracted via FASTA format from NCBI.

ClustalX 2.0.6 was used to align the nucleotide sequences (Thomson, 1997). The GC content of all 41 barcodes was estimated by BioEdit sequence alignment editor (Hall, 1999). MEGA 4.1 β 3 was used to construct phylogenetic trees via Neighbourhood joining method using Kimura 2 parameter and to calculate genetic distance of the given set of sequences in each codon position.

Barcode sequence of *Terapon jarbua* (FJ384681) sampled from Parangipettai coastal waters was used as an out group in Phylogenetic tree construction (Tamura *et al*., 2007).

RESULTS AND DISCUSSION

Sequence features: Totally 41 sequences (collected from NCBI) of *Lates calcarifer*, *Psammoperca waigiensis* and *L. niloticus* belonging to the family Latidae (Table 1) distributed from Tanzania to Northern Australian waters were collected from NCBI.

These along with the barcode data of *L. calcarifer* (540 bp) generated from this process were considered for the analysis (Fig. 1). Simultaneously, GC content in the barcode region (5' cytochrome oxidase C subunit I) of the selected fishes was noted. It ranged between 47.7 and 49.61%.

The maximum GC content was observed in fishes collected from the South China sea and minimum among the fishes of Tanzanian waters. The average GC content of *Lates calcarifer*, *Psammoperca waigiensis* and *L.niloticus* was 49.18, 48.03 and 47.70%, respectively. The GC content of *Lates calcarifer* from South China sea and Myanmar was found to be higher (49.41 and 49.38%, respectively) than the same species from other waters (Australia, India and Singapore). The average GC content of selected fishes was found to be 48.87%.

Table 1: Sequences of Lates calcarifer, Psammoperca waigiensis and L. niloticus from different countries used in the present study

Species	Countries	No. of sequences used	Sequence length
Lates calcarifer	South China	16	654
	Myanmar	5	657
	India	1	540
	Australia	9	612
	Singapore	1	1503
Psammoperca waigiensis	India	5	657
Lates niloticus	Tanzania	5	654
Terapon jarbua (Out group)	India	1	516

Phylogenetic analysis: A phylogenetic tree was constructed to verify the efficiency of coxI gene in delineating closely related species. Barcode sequence of *Terapon jarbua* from Parangipettai coastal waters was used as an out group and this has been clearly distinguished as an out group in the phylogenetic tree (Fig. 2). The constructed phylogram showed two distinct Clads (A and B).

Genetic distance: The genetic distance within and between the species and genera of family Latidae was calculated using pair-wise distance analysis via Kimura 2 parameter (Table 2). The intra species genetic distance within the members of *L. calcarifer* in 1-3rd codon position was found to be very low (0.01, 0.018, 0.082, respectively). The intra species variation of genus *P.waigiensis* in 1-3rd codon position was found to be negligible (0, 0.0040, respectively). The genetic distance within the genus *Lates* was found to be higher than between the same species.

L. calcarifer from South China sea showed more genetic relatedness with the same species from Singapore and Australia as these sequences were in Clad A. In the second internal, branch of Clad A, *L. calcarifer* (FJ384689 or IOBML45) from Parangipettai water showed more genetic relatedness with same species of Myanmar waters as these areas are present in the Indian ocean basin. That way Clad A showed significant phylogeographical cues. Clad B consisting of the members of *P. waigiensis* and

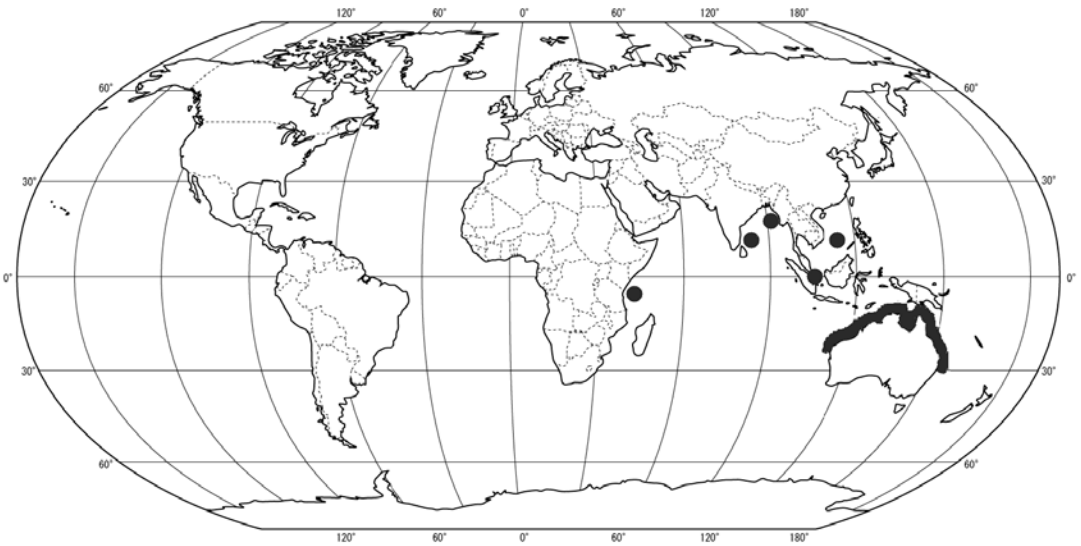

Fig. 1: Map showing the different geographical region around the world for which the barcode data are available in NCBI. The sequences in the database were largely contributed by countries like China, Australia and Myanmar

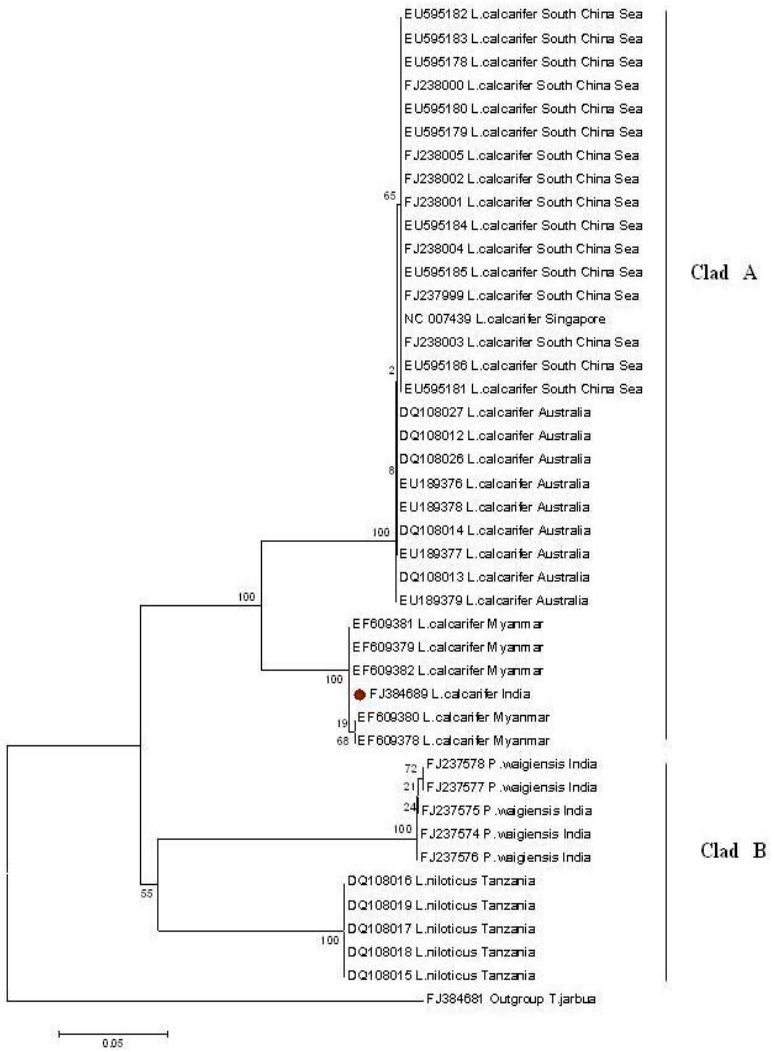

Fig. 2: Kimura 2 parameter distance neighbour joining tree of 42 barcode sequences from 3 species belonging to family Latidae. COI sequence of *Terapon jarbua* collected from Parangipettai coastal waters was used as an out group. Specimen number denote the accession number of NCBI database and *L. calcarifer* India* represents the specimens collected from Parangipettai coastal waters and the barcode sequences submitted at Barcode of Life Database (BOLD, www.barcodinglife.org)

Table 2: Average generic variations encoded in COI of selected fishes in different codon positions

Species	*Lates calcarifer*	*Psammoperca waigiensis*	*Lates niloticus*
Lates calcarifer			
(1st +2nd +3rd +Non coding)	0.034	0.244	0.210
(1st codon position)	0.010	0.096	0.091
(2nd codon position)	0.018	0.051	0.041
(3rd codon position)	0.082	0.916	0.707

L. niloticus. *P. waigiensis* (from India) showed close relatedness with *L. calcarifer* from India and Myanmar. This is probably due to the recent divergence of *P. waigiensis* from *L. calcarifer*. *L. niloticus* showed distant relatedness with *L. calcarifer* since, it was from West African Tanzanian coast. Species from different geographical locations occupied distinct internal branches in the phylogram and ultimately formed 2 distinct clads clearly showing the phylogeographical signals. The observed pattern in hologram mirrors that obtained by Hebert *et al.* (2003) using 1520 base pairs of data from portions of various mitochondrial genes. The higher bootstrap values in all the interior branches of the phylogram indicated the efficiency of cox1 gene in delineating closely related fishes from the same family.The negligible intra species variation of genus *P. waigiensis* in 1-3rd codon position might be due to very limited haplotype diversity existing within the species since all

the species of *P. waigiensis* used in this analysis was from the same geographical area in the Indian waters. The same thing was observed in *L. niloticus* where the genetic distance in all the codon position was the same (0.000). This clearly explained that on a regional scale, genetic variation at COI was low for the selected taxa.

Similar, findings were also observed in Australian and South African fish population (Ward *et al*., 2005). It was observed that the genetic distance between *L. calcarifer* and *L. niloticus* in 1-3rd codon position was 0.091, 0.041 and 0.707, respectively. The genetic distance between *L. calcarifer* and *P. waigiensis* in 1-3rd codon position was observed to be 0.096, 0.051 and 0.916, respectively. That way, it was quite clear that within species, DNA sequence was more similar to one another than to the different species (Zemlak *et al*., 2009).

It was also observed that the genetic distance between *P. waigiensis* and *L. niloticus* in 1-3rd codon position was 0.067, 0.049 and 0.662, respectively which is comparatively lower than the genetic distance between *L. calcarifer* and *L. niloticus*. So, there are more possibilities that *P. waigiensis* emerged from *L. niloticus* than from *L. calcarifer*. This exact ancestry can be further confirmed by constructing the phylogram with all the 17 representatives of Latidae family.

The low genetic divergences between *L. calcarifer* from Indian and Myanmar water would reflect gene flow mediated through short distance dispersal events and this hypothesis could be validated through studies along the coastal arc connecting these two geographical areas. Similar results with other inshore/offshore overlooked marine fishes were observed from Australia and South African coast (Simmons *et al*., 2006). None of the selected species showed markedly deeper COI variation above the divergence threshold value (3.5%). This metric was based on the recommendation by Hebert *et al*. (2004b) that an appropriate screening threshold for flagging provisional species is 10X that of the average intra species variation for a focus group.

Since, this study was conducted using only three different species of Latidae family, further study is needed to find the minimum and conservative genetic divergence threshold value of fishes belonging to Latidae family. Another interesting observation made from the genetic distance calculation was higher genetic distance in 3rd codon position than its corresponding 1st and 2nd codon positions.

Similar observation was made by Ward *et al*. (2005) while barcoding of fishes from Australian waters. Simmons *et al*. (2006) also observed that greater phylogenetic signal is often found in parsimony based analyses of 3rd codon positions of protein coding genes relative to their corresponding first and second codon positions even for early-derived basal clads.

CONCLUSION

The COI sequence in the phylogram constructed clearly clustered the selected species in individual group, proving the efficacy of COI gene in delineating the members of Ladidae to their species level. Hence, we conclude that COI sequence (DNA barcode) could be potentially used to identify the individual members of family Ladidae. Intra species genetic variations of species from same geographical area will be low compared to the inter species genetic distance. We could also conclude that the greater phylogenetic signal is often found in 3rd codon position relative to their corresponding first and second codon positions as reported by previous studies (Ward *et al*., 2005; Zemlak *et al*., 2009; Simmons *et al*., 2006).

ACKNOWLEDGEMENTS

Researchers are thankful to Dr. T. Balasubramanian, Professor and Director of the centre for the encouragement and the authorities of our university for the facilities. We also extend our sincere thanks to Dr. Wafer, National Institute of Oceanography, Goa and Bioserve Biotechnologies (India) Pvt. Ltd. for their constant supporting throughout the research.

REFERENCES

Berra, T.M., 2001. Freshwater Fish Distribution. Academic Press, San Diego, California.

Chenowith, S.F., M.J. Hughes, C.P. Keenen and S. Lavery, 1998. When Oceans meet: A teleost shoes specific integration at an Indian-Pacific interface. Proc. Royal Soc. London Biol. Sci., 265: 415-420.

Frost, L.A., B.S. Evans and D.R. Jerry, 2007. Loss of genetic diversity due to hatchery culture practices in barramundi (*Lates calcarifer*). Aquaculture, 272: 238-321.

Hall, T.A., 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. Nucleic Acid Symp. Ser., 41: 95-98.

Hallerman, E.M. and J.S. Beckmann, 1988. DNA level polymorphism as a tool in fisheries science. Can. J. Fish. Aquat. Sci., 45: 1075-1087.

Hebert, P.D.N., A. Cywinska, S.L. Ball and J.R. de Waard, 2003. Biological identifications through DNA barcodes. Proc. R. Soc. Lond. B, 270: 313-321.

Hebert, P.D.N., M.Y. Stoeckle, T.S. Zemlak and C.M. Francis, 2004a. Identification of birds through COI DNA barcodes. Public Library Sci. Biol., 2: 1-7.

Hebert, P.D.N., E.H. Penton, J.M. Burns, D.H. Janzen and W. Hallwachs, 2004b. Ten species in one: DNA barcoding reveals cryptic species in the Neotropical skipper butterfly *Astraptes fulgerator*. Proc. Nat. Acad. Sci. USA., 101: 14812-14817.

Hogg, I.D. and P.D.N. Hebert, 2004. Biological identifications of springtails (Collembola: Hexapoda) from the Canadian arctic using mitochondrial DNA barcodes. Can. J. Zool., 82: 749-754.

Lipscomb, D., N. Platnick and Q. Wheeler, 2003. The intellectual content of taxonomy: A comment on DNA taxonomy. Trends Ecol. Evol., 18: 65-66.

Miller, S.A., D.D. Dykes and H.F. Polesky, 1988. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res., 16: 1215-1215.

Moritz, C. and C. Cicero, 2004. DNA barcoding: Promise and pitfalls. PLoS Biol., 2: 354-354.

Nelson, J.S., 1994. Fishes of the World. A Wiley Inter Science Publications. 1st Edn., John Wiley and Sons, New York, pp: 1-523.

Simmons, M.P., L.B. Zhang, C.T. Webb and A. Reeves, 2006. How can third codon positions outperform first and second codon positions in phylogenetic inference: An empirical example from the seed plants. Syst. Biol., 55: 245-258.

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.

Tautz, D., P. Arctander, A. Minelli, R.H. Thomas and A.P. Vogler, 2003. A plea for DNA taxonomy. Trends Ecol. Evolut., 18: 70-74.

Tautz, D.P., Arctander, A. Minelli, R.H. Thomas and A.P. Vogler, 2002. DNA points the way ahead in taxonomy. Nature, 418: 479-479.

Thomson, J.M., 1997. The mugilidae of the world. Mem. Queensland Mus., 41: 457-562.

Wang, C.M., Z.Y. Zhu, L.C. Lo, F. Feng and G. Lin *et al*., 2007. A microsatellite linkage map of barramundi, *Lates calcarifer*. Genetics, 175: 907-915.

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.

Zemlak, S.T., R.D. Ward, A.D. Connell, B.H. Holmes and P.D.N. Hebert, 2009. DNA barcoding reveals overlooked marine fishes. Mol. Ecol. Res., 9: 237-242.