PUBLICATIONS

PUBLICATIONS IN PEER-REVIEWED JOURNALS

1. **Dhar, B.**, Ghosh S K (2015). Genetic assessment of ornamental fish species from North East India. *Gene.* **555**(2):382-92.

2. **Dhar, B.**, Ghosh, S. K (2015). Identifying Ornamental fishes of North-east India through DNA Barcoding. *Sci and Tech Jour.* **2**(2): 43-49

3. Khomdram B., **Dhar, B.**, Ghosh S K (2014). Jiribam, the Ornamental Fishes' Hot Spot Zone Of Manipur, India. *IOSR-JAVS. 7(1): PP 85-91.*

4. Trivedi, S., Affan, R., Alessa, A. H. A., Ansari, A. A,. **Dhar, B.,** Mahadani. P., Ghosh, S. K. (2014). DNA Barcoding of Red Sea Fishes from Saudi Arabia – The first approach. *DNA Barcodes* .2: 17–20.

5. Laskar B.A, Bhattacharjee M.J., **Dhar, B.**, Mahadani P, Kundu S, Ghosh S. K. (2013). The Species Dilemma of Northeast Indian Mahseer (Actinopterygii: Cyprinidae): DNA Barcoding in Clarifying the Riddle. *PLoS ONE* **8**(1): e53704. doi:10.1371/journal.pone.0053704.

6. Bhattacharjee M.J., Laskar B.A., **Dhar, B.**, Ghosh SK (2012). Identification and Evaluation of Freshwater Catfishes through DNA Barcoding. *PLoS ONE* **7**(11): e49950. doi:10.1371/journal.pone.0049950.

7. Khomdram, B**., Dhar, B.,** Laskar BA, Ghosh S K (2013). Exploring Ornamental Fishes of Manipur through DNA Barcoding. *J. Environ. & Sociobiol*. **10**(1):83-88

8. Bhattacharjee M.J., Laskar B.A., **Dhar, B.**, Ghosh S.K. (2013). Role of transversional substitution on DNA Barcode based species differentiation: catfish as a case study. *J. Environ. & Sociobiol*. **10**(1): 65-75

9. Chakraborty, M., Ghosh, S.K., **Dhar, B.**, Khomdram, B**.,** (2013). Revealing genetic diversity of *Clarias batrachus* using DNA Barcode. *J. Environ. & Sociobiol*. **10**(1): 25-32

10. Talukdar, F.R., Ghosh, S.K., Laskar, R.S., Mahadani, P., Chakraborty, M., **Dhar, B.** and Bhattacharjee, M.J. (2013). Implication of nucleotide substitution at third codon position of the DNA barcode sequences. *J. Environ & Sociobiol.* **10**(1):55-63.

11. Ghosh, S.K., Bhattacharjee, M.J., Devi, M.K., Ahanthem, M., Kundu, S., Mahadani, P., **Dhar, B**., Khondram, B., Chakraborty, M., Rahman, F., Mondal, R., Hansa, J., Laskar, R., Mazumdar, Sarkar, P., Rajbonshi, S., Chakraborty,

A., Das, M., Ghosh, P.R., Das, K.C and Laskar, B.A. (2011). DNA Barcoding: Digital Taxonomy of Biorsources. *Strategic Physiological Research for Sustainable Animal Biodiversity. SAPIKOL*:268-275.

Manuscript communicated

Design of Character-based DNA Barcode Motif for species identification: A computational approach and its validation in fishes

Authors: Chakraborty, M., Dhar, B. and Ghosh, S.K.

Communicated in: Molecular Ecology **(Under review).**

SEQUENCE SUBMITTED:

1. NCBI: **146 Barcode sequence so far [Genbank Accession: JN628880-JN628897, JN815267- JN815311, JN245989- JN245992, JQ713844- JQ713859, JX105464- JX105485, N673954- N673955, etc]**

2. BOLD: **150 sequences so far [under Project CODE "OFISH- and "CFISH"].**

OTHER PUBLICATIONS

1. Singh, S. A., Choudhury&, J. H., Kapfo, W., Kundu, S., **Dhar, B**., Laskar, S., & Ghosh, S. K. (2015). Influence of the CYP1A1 T3801C Polymorphism on Tobacco and Alcohol-Associated Head and Neck Cancer Susceptibility in Northeast India. *Asian Pacific Journal of Cancer Prevention*, **16**(16): 6953-6961.

2. Choudhury, J.H., Singh, S.A., Kundu, S., Choudhury, B., Talukdar, F.R., Srivasta, S., Laskar, R.S., **Dhar, B.,** Das, R., Laskar, S. and Kumar, M., (2015). Tobacco carcinogen-metabolizing genes CYP1A1, GSTM1, and GSTT1 polymorphisms and their interaction with tobacco exposure influence the risk of head and neck cancer in Northeast Indian population. *Tumor Biology*, **36**(8): 5773-5783.

3. Laskar, R. S., Talukdar, F. R., Choudhury, J. H., Singh, S. A., Kundu, S., **Dhar, B.**, Mondal, R. & Ghosh, S. K. (2015). Association of HPV with genetic and epigenetic alterations in colorectal adenocarcinoma from Indian population. *Tumor Biology*, **36**(6): 4661-4670.

4. Mondal, R., Ghosh, S. K., Choudhury, J. H., Seram, A., Sinha, K., Hussain, M., Nath Choudhury, M., Talukdar, F.R., Chaudhuri, B., **Dhar, B.** (2013). Mitochondrial DNA copy number and risk of oral cancer: a report from Northeast India*.* *PloS ONE*, **8**(3), e57771.

BOOK CHAPTER CONTRIBUTION

1. Laskar, B.A., Das, K.C., **Dhar, B**., Bhattacharjee, M.J., Ghosh, S.K. (2012) Biodiversity: Inventorying and monitoring. A text book on DNA Barcoding. ISBN: 978-81-922989-4-8. **(Chapter-1)**

2. Bhattacharjee, M.J., Kundu, S., Devi, Ksh.M., **Dhar, B**., Mondal, R., Mahadani, P., Ahanthem, M., Ghosh, S.K. (2012) Molecular Biology in DNA Barcoding. A text book on DNA Barcoding. ISBN: 978- 81-922989-4-8. **(Chapter-4)**

3. **Dhar, B.,** Ghose, A., Kundu, S.,Nath, J., Choudhury, A., Ghorai, S., Trivedi, S. and Ghosh, S.K. (2016) DNA Barcoding of primitive species-nemertine from sundarbans marine bio-resource. DNA Barcoding in Marine Perspectives. Spinger-ISBN: 978-3-319-41838-4 **(Chapter-12)**

4. **Dhar, B.,** Ghose, A., Kundu, S., Malvika, S., Devi, N.N., Choudhury, A., Ghorai, S., Trivedi, S. and Ghosh, S.K. (2016) DNA Barcoding: Molecular positioning of living fossils (Horseshoe crab). DNA Barcoding in Marine Perspectives. Spinger-ISBN: 978-3-319-41838-4 **(Chapter-10)**

NEWS HIGHLIGHTS from PhD works in Media

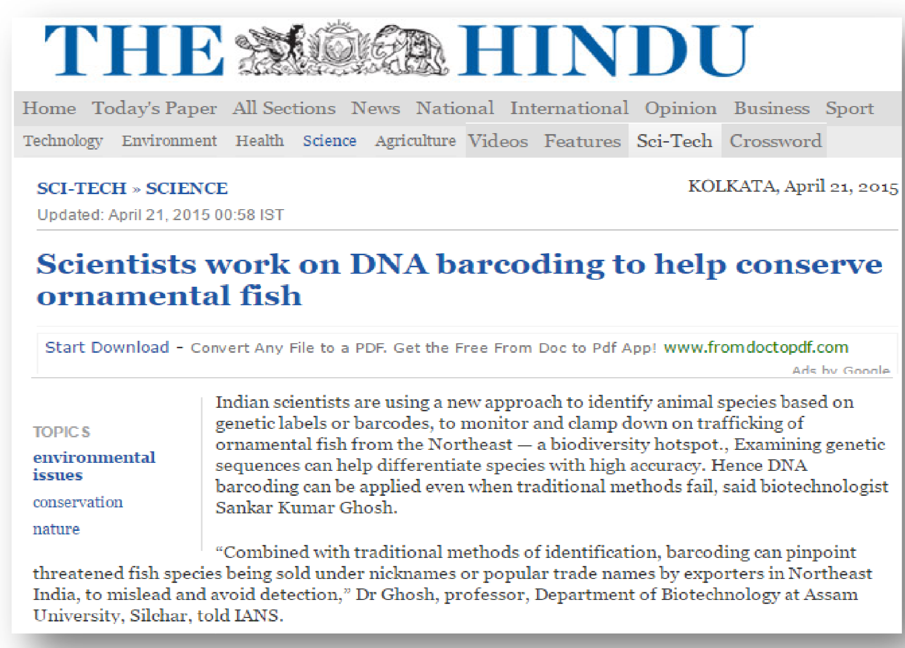
DNA barcoding to the rescue of India's ornamental fish

Last Updated: Monday, April 20, 2015 - 13:52

37 SHARES Facebook Twitter Share 2

Kolkata: Indian scientists are using a new approach to identify animal species based on genetic labels or barcodes, that can help monitor and clamp down on trafficking of ornamental fish from northeast India - a biodiversity hotspot - and aid conservation.

Just as shopkeepers scan the similar-yet-different zebra stripes (barcodes) on products to keep track of what they sell and what is in stock, examining certain ubiquitous genetic sequences can differentiate one species from the other with



THE HINDU

Home Today's Paper All Sections News National International Opinion Business Sport
Technology Environment Health Science Agriculture Videos Features Sci-Tech Crossword

SCI-TECH » SCIENCE

KOLKATA, April 21, 2015

Updated: April 21, 2015 00:58 IST

Scientists work on DNA barcoding to help conserve ornamental fish

Start Download - Convert Any File to a PDF. Get the Free From Doc to Pdf App! www.fromdoctopdf.com
Ads by Google

TOPICS
environmental issues
conservation
nature

Indian scientists are using a new approach to identify animal species based on genetic labels or barcodes, to monitor and clamp down on trafficking of ornamental fish from the Northeast — a biodiversity hotspot., Examining genetic sequences can help differentiate species with high accuracy. Hence DNA barcoding can be applied even when traditional methods fail, said biotechnologist Sankar Kumar Ghosh.

"Combined with traditional methods of identification, barcoding can pinpoint threatened fish species being sold under nicknames or popular trade names by exporters in Northeast India, to mislead and avoid detection," Dr Ghosh, professor, Department of Biotechnology at Assam University, Silchar, told IANS.

Sunday, 24 July, 2016

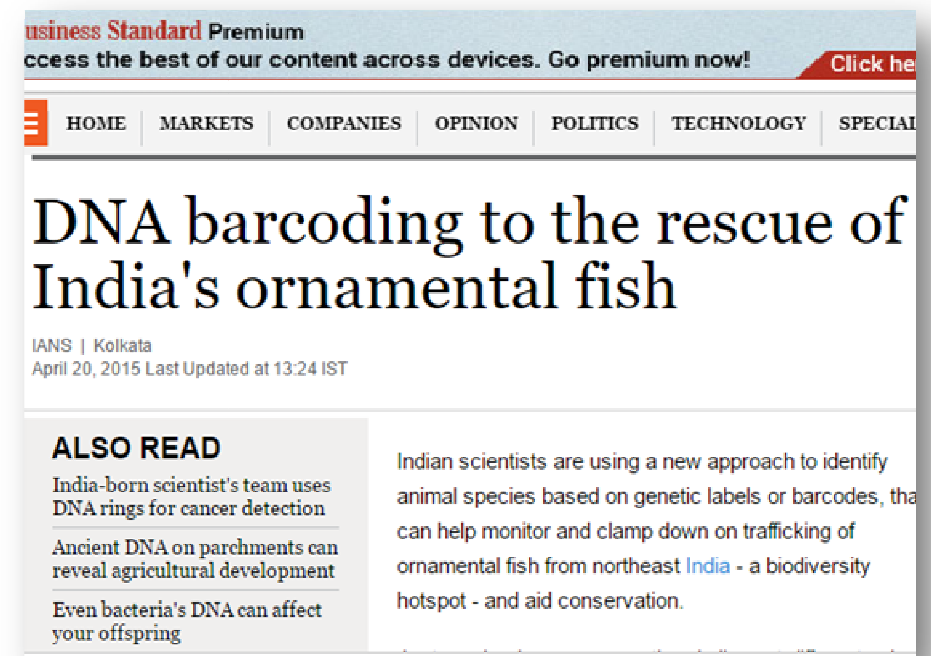
The Statesman

DNA barcoding for ornamental fishes

IANS

Posted at: Apr 20 2015 2:00PM

Indian scientists are using a new approach to identify animal species bas
biodiversity hotspot and aid conservation.
Just as shopkeepers scan the similar yet different zebra stripes (barcodes
one species from the other with high accuracy.
The upshot, says biotechnologist Sankar Kumar Ghosh, is that DNA barc
"Combined with traditional methods of identification, barcoding can pi
avoid detection," Ghosh, professor, department of biotechnology at Assa
To lure hobbyists and enthusiasts, dealers in the northeastern states also



usiness Standard Premium
ccess the best of our content across devices. Go premium now! Click he

HOME | MARKETS | COMPANIES | OPINION | POLITICS | TECHNOLOGY | SPECIAL

DNA barcoding to the rescue of India's ornamental fish

IANS | Kolkata
April 20, 2015 Last Updated at 13:24 IST

ALSO READ

India-born scientist's team uses DNA rings for cancer detection

Ancient DNA on parchments can reveal agricultural development

Even bacteria's DNA can affect your offspring

Indian scientists are using a new approach to identify animal species based on genetic labels or barcodes, tha can help monitor and clamp down on trafficking of ornamental fish from northeast India - a biodiversity hotspot - and aid conservation.

National Fisheries Development Board

(Department of Animal Husbandry, Dairying & Fisheries, Ministry of Agriculture, Govt. of India)

National Fisheries Development Board

(Department of Animal Husbandry, Dairying & Fisheries, Ministry of Agriculture, Govt. of India)

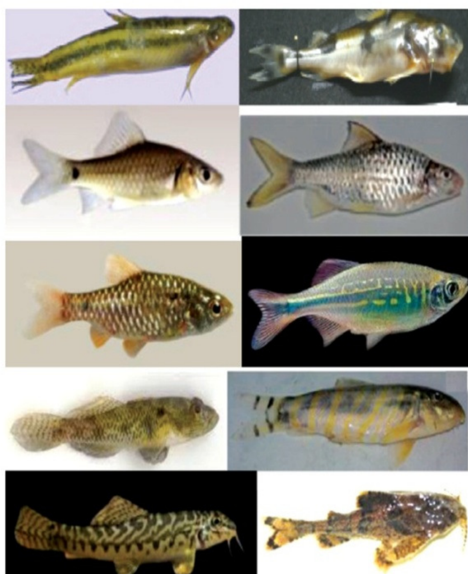
Fish Building, Pillar No. 235, P.V. Narsimha Rao Expressway
Sardar Vallabhai Patel National Police Academy (SVP NPA) Post, HYDERABAD – 500 052
Website: http://nfdb.gov.in E-mail: info.nfdb@nic.in Tel: +91-40-24000201 Fax: +91-40-24015552 / 68

Toll Free Telephone Number: 1800-425-1660

5.3 DNA Barcoding to help regulate ornamental fish trade and conserving biodiversity in North-Eastern States

North-East India is considered as one of the hot spots of freshwater fish biodiversity in the world. An inventory of fish species of North-East India showed that the highest number is recorded from Assam (187), followed by Arunachal Pradesh (165), Meghalaya (159), Manipur (139), Tripura (103), Nagaland (71), Mizoram (46), and Sikkim (29). Out of these 899 species, 250 were of potential ornamental value.

Ornamental fishes are traded with multiple names around the world, including North-East India. Most are collected from the wild, due to lack of species-specific culture or breeding, and therefore, such unmanaged collection of the wild and endemic species could lead to severe threats to their biodiversity. Development of species-specific mtDNA marker on CO1 gene that is bar-coded helps in species identification. Through DNA Barcoding and morphological assessment, identification of 128 ornamental fish specimens exported from NE India was done. The analysis helped in straightforward identification of 84 specimens into 35 species, while 44 specimens were difficult to distinguish based on barcode alone. However, these cases were resolved through morphology, neighbour-joining and distance based approach and found to be belonging to 16 species. Among the 51 identified species (35+16), 14 species had multiple trade names and 17 species belonged to threatened category. Species-level identification through DNA Barcoding, along with traditional morpho-taxonomy, reflects its efficacy in regulating ornamental fish trade and their conservation in nature. The use of trade names rather than the zoological name created the passage for trafficking of the threatened species and demands immediate attention for sustaining wildlife conservation. [Source: *Prof. Sankar Kumar Ghosh, Assam University*]

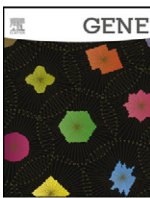


Gene 555 (2015) 382–392

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene



Genetic assessment of ornamental fish species from North East India

Bishal Dhar, Sankar Kumar Ghosh *

Department of Biotechnology, Assam University, Silchar 788011, Assam, India

ARTICLE INFO

Article history:
Received 4 August 2014
Received in revised form 4 November 2014
Accepted 15 November 2014
Available online 18 November 2014

Keywords:
Ornamental fish
Northeast India
DNA barcode
Unlawful fish trading
IUCN
CAMP

ABSTRACT

Ornamental fishes are traded with multiple names from various parts around the world, including North East India. Most are collected from the wild, due to lack of species-specific culture or breeding, and therefore, such unmanaged collection of the wild and endemic species could lead to severe threats to biodiversity. Despite many regulatory policies, trade of threatened species, including the IUCN listed species have been largely uncontrolled, due to species identification problems arising from the utilization of multiple trade names. So, the development of species-specific DNA marker is indispensable where DNA Barcoding is proved to be helpful in species identification. Here, we investigated, through DNA Barcoding and morphological assessment, the identification of 128 ornamental fish specimens exported from NE India from different exporters. The generated sequences were subjected to similarity match in BOLD-IDS as well as BLASTN, and analysed using MEGA5.2 for species identification through Neighbour-Joining (NJ) clustering, and K2P distance based approach. The analysis revealed straightforward identification of 84 specimens into 35 species, while 44 specimens were difficult to distinguish based on *CO1* barcode alone. However, these cases were resolved through morphology, NJ and distanced based method and found to be belonging to 16 species. Among the 51 identified species, 14 species represented multiple trade names; 17 species belonged to threatened category. Species-level identification through DNA Barcoding along with traditional morphotaxonomy reflects its efficacy in regulating ornamental fish trade and therefore, appeals for their conservation in nature. The use of trade names rather than the zoological name created the passage for trafficking of the threatened species and demands immediate attention for sustaining wildlife conservation.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Aquarium fish keeping has become one of the most popular hobbies and fishes suitable for keeping in the aquarium fetch a high economic value across the globe contributing to increased growth in international ornamental fish trade. The worldwide export has increased from US$ 176 million to US$ 251 million, and the import has increased from US $ 257 million to US$ 303 million during the year 1998 to 2004 (FAO Fisheries and Aquaculture Dept. Fishery and Aquaculture Information and Statistics Service, 2008). The Indian ornamental fish export represents a small portion of the global share but it is increasing with 20% rise per annum with annual income of US$ 1.2 million (Marine Products Export Development Authority, MPEDA). The North-eastern (NE) region of India contributes the lion's share of Indian aquarium fish trade (Bhattacharya and Choudhury, 2004). This region of the country is important in view of large-scale production from capture fishery underlined by the existence of innumerable rivers, rivulets and lentic water bodies, which harbour a diverse fish fauna. There are about 267 species belonging to 136 genera of fresh water fishes inhabiting in NE India (Ponniah and Sarkar, 2000). Out of which, 54.32% possess either of the three values as food for human, component in angling tourism or aquarium fish trade, and are thus potential resources for the growth of economy. About 80% of the total ornamental fish trade is rooted from wild catch and is contributed by this region of India via Kolkata Airport (Das and Biswas, 2009; Kalita and Deka, 2013). Aquarium fishes are categorized based on vivid colour pattern (colourful), morphological uniqueness (special), and behaviourally charismatic (semi-aggressive, community, non-community) (Ponniah and Sarkar, 2000). Besides, ornamental fish traders always intend to publicize uncommon or unseen species in the trade in order to attract hobbyists. In the process, some traders adopt an unfair practice, like, use of synthetic dyes to develop colour in those species which are originally colourless thus claiming them to be ornamental and traded by different names to the clients. On the other hand, many endemic fish species are being traded from the wild harvest, due to lack of established species-specific culture or breeding, and serve as a threat to biodiversity from harvest pressure. Despite several regulatory enactments in India, aquarium fishes are

Abbreviations: K2P, Kimura 2 parameter; COI, cytochrome C oxidase 1; NJ, Neighbour Joining; ML, maximum likelihood; S.E., standard error; NUMTs, nuclear DNA originating from mitochondrial DNA sequences; EDTA, ethylenediaminetetraacetic acid; PCR, polymerase chain reaction; dNTPs, deoxynucleotide triphosphates; MEGA, Molecular Evolutionary Genetic Analysis; DNA, deoxyribonucleic acid; BOLD, Barcode of Life Datasystem; BLAST, Basic Local Alignment Search Tool; IUCN, International Union for Conservation of Nature.

* Corresponding author at: Department of Biotechnology, Assam University, Silchar-788011, India.
E-mail address: drsankarghosh@gmail.com (S.K. Ghosh).

http://dx.doi.org/10.1016/j.gene.2014.11.037
0378-1119/© 2014 Elsevier B.V. All rights reserved.

Table 1
Straight forward identification of the studied ornamental fishes based on similarity match with BOLD Identification System and GenBank. The match is exclusively based on similarity of developed sequences with database sequences. 84 traded fishes identified as 35 species, i.e. in many cases; same species is traded with multiple trade names.

Fish specimen with trade name	Sequences generated (Acc. no GenBank) (size in bp)	Closest match with species in BOLD-IDS		Close match in GenBank (BLASTN similarity in %)	Identified as species (number)
		Species level barcode records (process ID)	Public record barcode database (process ID)		
Gangetic latia	JN815299 (626 bp), JX105481 (618 bp), KF511500 (624 bp), KF511501 (624 bp), KF511502 (627 bp)	*Crossocheilus latius* (ANGBF9951-12, ANGBF9956-12)	*Crossocheilus latius* (ANGBF9951-12, ANGBF9956-12)	*Crossocheilus latius (99)*	*Crossocheilus latius (1)*
Neon Hatchet	JN815300 (631 bp), JN815301 (634 bp)	*Chela cachius*	*Chela cachius* (GBGC4871-08)	*Chela cachius (97)*	*Chela cachius (2)*
Tank goby	JN815293 (537 bp), JN815296 (623 bp), JQ713857 (632 bp), JN815294 (604 bp), JN815295 (619 bp)	*Glossogobius guiris*	No match	*No match (99)*	*Glossogobius guiris (3)*
Glass fish	JN815274 (534 bp), JN815275 (584 bp), JN815276 (607 bp), JN815283 (639 bp)	*Parambassis ranga*	No match	*Pseudoambassis ranga (99)*	*Parambassis ranga (4)*
Barred Spiny eel	JX105465 (602 bp), JN815289 (631 bp)	*Macrognathus pancalus*	*Macrognathus pancalus* (GBGC4234-08, GBGC4233-08)	*Macrognathus pancalus (99)*	*Macrognathus pancalus (5)*
Labeo	JQ713848 (624 bp)	*Labeo calbasu* (GBGC4252-08, ANGBF7333-12)	*Labeo calbasu* (GBGC4252-08, ANGBF7333-12)	*Labeo calbasu (100)*	*Labeo calbasu (6)*
Flying barb	JN673955 (655 bp), KF511504 (622 bp), KF511505 (622 bp), KF511506 (622 bp)	*Esomus danricus*	*Esomus danricus* (ANGBF6124-12, ANGBF6125-12)	*Esomus danricus (99)*	*Esomus danricus (7)*
Silver barb	JQ713846 (655 bp)	*Barbonymus gonionotus* (GBGC6590-09, GBGC6591-09, GBGC6592-09, ANGBF5730-12)	*Barbonymus gonionotus* (GBGC6590-09, GBGC6591-09, GBGC6592-09, ANGBF5730-12)	*Barbonymus gonionotus (100)*	*Barbonymus gonionotus (8)*
Corsula mullet	JX105471 (623 bp)	*Rhinomugil corsula*	*Rhinomugil corsula*	*Rhinomugil corsula (98)*	*Rhinomugil corsula (9)*
Giant Danio	KF511497 (628 bp), KF511498 (628 bp), KF511499 (626 bp)	*Devario aequipinnatus*	*Devario aequipinnatus* (RCYY279-11, RCYY280-11, RCYY334-11, ANGBF6126-12)	*Devario aequipinnatus (98–99)*	*Devario aequipinnatus (10)*
Clown knifefish	KF511511 (558 bp)	*Chitala chitala* (ANGBF6042-12, ANGBF6132-12, ANGBF6043-12)	*Chitala chitala* (ANGBF6042-12, ANGBF6132-12, ANGBF6043-12)	*Chitala chitala (99–100)*	*Chitala chitala (11)*
Suckerthroat	KF511525 (622 bp), KF511526 (622 bp), KF511527 (603 bp)	*Pseudecheneis sulcata*	*Pseudecheneis sulcata* (GBGC8471-09)	*Pseudecheneis sulcata (99)*	*Pseudecheneis sulcata (12)*
Indian Hill Trout	KF511547 (629 bp), JN815290 (661 bp), JN815291 (660 bp), KF511548 (629 bp), KF511549 (628 bp)	*Barilius bendelisis* (CYTC3711-12, CYTC4266-12)	*Barilius bendelisis* (CYTC3711-12, CYTC4266-12)	*Barilius bendelisis (100)*	*Barilius bendelisis (13)*
Indian Whiptail catfish	KF511561 (625 bp)	*Sisor rabdophorus*	*Sisor rabdophorus* (CFISH028-12, CFISH027-12)	*Sisor rabdophorus (99)*	*Sisor rabdophorus (14)*
Long Whiskers Catfish	KF511564 (620 bp)	*Mystus gulio*	*Mystus gulio* (GBGCA2888-13)	*Mystus gulio (99)*	*Mystus gulio (15)*
Striped dwarf catfish	KF511563 (620 bp)	*Mystus vittatus*	*Mystus vittatus* (CFISH009-12)	*Mystus vittatus (99)*	*Mystus vittatus (16)*
Butter catfish	KF511565 (620 bp)	*Ompok bimaculatus*	*Ompok bimaculatus*	*Ompok bimaculatus (98)*	*Ompok bimaculatus (17)*
Gangetic ailia	KF511566 (620 bp)	*Ailia coila* (ANGBF6053-12, ANGBF6054-12, GBGC4011-08)	*Ailia coila* (ANGBF6053-12, ANGBF6054-12, GBGC4011-08)	*Ailia coila (100)*	*Ailia coila (18)*

(continued on next page)

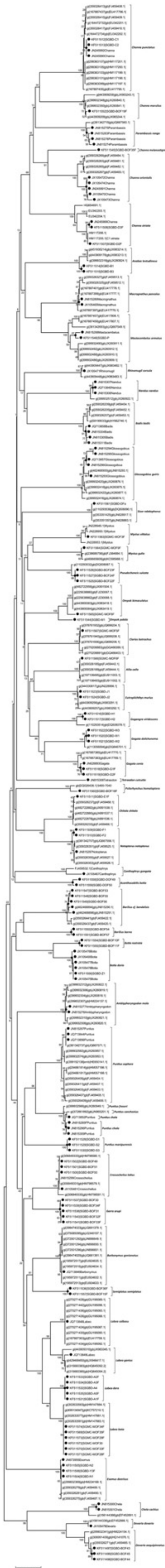
Table 1 *(continued)*

Fish specimen with trade name	Sequences generated (Acc. no GenBank) (size in bp)	Closest match with species in BOLD-IDS		Close match in GenBank (BLASTN similarity in %)	Identified as species (number)
		Species level barcode records (process ID)	Public record barcode database (process ID)		
Walking catfish	KF511567 (620 bp)	*Clarias batrachus*	*Clarias batrachus* (ANGBF2196-12, ANGBF2200-12, CFISH068-12)	*Clarias batrachus (100)*	*Clarias batrachus (19)*
Gagata	KF511520 (627 bp), KF511521 (627 bp), KF511522 (627 bp)	*Gagata dolichonema* (GBGC9465-09)	*Gagata dolichonema* (GBGC9465-09)	*Gagata dolichonema (99)*	*Gagata dolichonema (20)*
River Catfish	KF511523 (629 bp), KF511524 (629 bp)	*Eutropiichthys murius* (CFISH012-12, CFISH013-12 DBFN023-11)	*Eutropiichthys murius* (CFISH012-12, CFISH013-12 DBFN023-11)	*Eutropiichthys murius (99–100)*	*Eutropiichthys murius (21)*
Turquoise Danio	JX105479 (624 bp)	*Devario devario*	*Devario devario* (GBGC4896-08)	*Devario devario (100)*	*Devario devario (22)*
Tire track eel	JN815288 (654 bp), KF511546 (627 bp)	*Mastacembelus armatus*	No match	*Mastacembelus armatus (96–97)*	*Mastacembelus armatus (23)*
Bullseye snakehead	KF511552 (582 bp)	*Channa marulius*	*Channa marulius* (DBFN124-11)	*Channa marulius (99)*	*Channa marulius (24)*
Pool barb	JN815267 (654 bp)	*Puntius sophore*	*Puntius sophore* (RCYY481-11)	*Puntius sophore (97–99)*	*Puntius sophore (25)*
Spot fin barb	JQ713844 (624 bp)	*Puntius sophore*	*Puntius sophore* (RCYY481-11)	*Puntius sophore (97–99)*	*Puntius sophore (25)*
Mola carplet	JN815277 (617 bp)	*Amblypharyngodon mola*	*Amblypharyngodon mola* (DBFN339-12)	*Amblypharyngodon mola (99)*	*Amblypharyngodon mola (26)*
Pale carplet	JN815278 (600 bp)	*Amblypharyngodon mola*	*Amblypharyngodon mola* (DBFN339-12)	*Amblypharyngodon mola (99)*	*Amblypharyngodon mola (26)*
Mud perch	JN815306 (643 bp), JN815307 (643 bp)	*Nandus nandus* (DBFN050-11)	*Nandus nandus* (DBFN050-11)	*No match*	*Nandus nandus (27)*
Leaf fish	JQ713845 (614 bp)	*Nandus nandus* (DBFN050-11)	*Nandus nandus* (DBFN050-11)	*Nandus nandus* (DBFN050-11)	*Nandus nandus (27)*
Spotted snakehead	JN245992 (624 bp)	*Channa punctatus*	*Channa punctatus* (DSCHA078-13, ANGBF2441-12)	*Channa punctatus (99)*	*Channa punctatus (28)*
Green snakehead	JN245990 (655 bp)	*Channa punctatus*	*Channa punctatus* (DSCHA078-13, ANGBF2441-12)	*Channa punctatus (99)*	*Channa punctatus (28)*
Checkered snake fish	KF511512 (622 bp), KF511513 (622 bp)	*Channa punctatus*	*Channa punctatus* (DSCHA078-13, ANGBF2441-12)	*Channa punctatus (99)*	*Channa punctatus (28)*
Blue badis	JN815304 (631 bp), JN815305 (622 bp)	*Badis badis*	*Badis badis* (ANGBF6048-12, ANGBF6138-12)	*Badis badis (98–99)*	*Badis badis (29)*
Chameleon fish	JN815311 (628 bp), JQ713858 (569 bp)	*Badis badis*	*Badis badis* (ANGBF6048-12, ANGBF6138-12)	*Badis badis (98–99)*	*Badis badis (29)*
Grey featherback	JN815287 (624 bp)	*Notopterus notopterus* (ANGBF6016-12)	*Notopterus notopterus* (ANGBF6016-12)	*Notopterus notopterus (99)*	*Notopterus notopterus (30)*
Bronze featherback	KF511509 (619 bp), KF511510 (626 bp)	*Notopterus notopterus* (ANGBF6016-12)	*Notopterus notopterus* (ANGBF6016-12)	*Notopterus notopterus (99)*	*Notopterus notopterus (30)*
Climbing perch	KF511514 (625 bp)	*Anabas testudineus*	No match	*Anabas testudineus (98)*	*Anabas testudineus (31)*
Climbing gourami	KF511515 (629 bp)	*Anabas testudineus*	No match	*Anabas testudineus (98)*	*Anabas testudineus (31)*
Barb	KF511528 (625 bp), KF511529 (625 bp), KF511530 (625 bp)	*Puntius manipurensis*	*Puntius manipurensis* (RCYY470-11, RCYY471-11)	*Puntius padamya (97)*	*Puntius manipurensis (32)*
Snakehead murrel	JN245989 (655 bp)	*Channa striata*	*Channa striata* (ANGBF2417-12, ANGBF2419-12)	*Channa striata (96–99)*	*Channa striata (33)*
Banded snakehead	KF511507 (622 bp)	*Channa striata*	*Channa striata* (ANGBF2417-12, ANGBF2419-12)	*Channa striata (96–99)*	*Channa striata (33)*
Striped snakehead	KF511508 (655 bp)	*Channa striata*	*Channa striata* (ANGBF2417-12, ANGBF2419-12)	*Channa striata*	*Channa striata (33)*
Walking snakehead	JN245991 (655 bp), JX105470 (623 bp)	*Channa orientalis*	*Channa orientalis* (DSCHA009-07, ANGBF2437-12)	*Channa orientalis (96–99)*	*Channa orientalis (34)*
Ceylone snakehead	JX105472 (635 bp), JX105473 (570 bp) JX105474 (637 bp)	*Channa orientalis*	*Channa orientalis* (DSCHA009-07, ANGBF2437-12)	*Channa orientalis (96–99)*	*Channa orientalis (34)*
Ocellated pufferfish	JN815308 (637 bp)	*Tetraodon cutcutia* (GBGCA5150-13, CYTC3695-120)	*Tetraodon cutcutia* (GBGCA5150-13, CYTC4343-12)	*Tetraodon cutcutia (100)*	*Tetraodon cutcutia (35)*

traded largely without endowment to the government and mostly from wild capture. In the process, many threatened species, (IUCN listed) are being sold due to the problem in identification that arises due to the utilization of multiple commercial names and thereby enhancing the risk of jeopardising of those species from the wild. The use of such trade names rather than scientific names is commonly practiced in ornamental fish market, and this has proven to be always misleading in identification as there may exist various trade names for a single species. So, the development of species-specific molecular marker with respect to the morphological characteristics of those zoologically named species which trades globally is essential in order to monitor the trade of threatened or restricted species. Nevertheless, the inclusive documentation of ornamental fish resources from the north-eastern region of India has been lagging behind due to the lack of quantitative data and shortage of expert taxonomists associated with other factors too. Since, ornamental fish trade has gone global; hence, there is a paramount need of inventorying the regional/national wealth of such bio-resources for recognising proper equity sharing of their trade. Until now, ample efforts of inventorying the ornamental fishes have been accomplished through morphotaxonomy; however, the conventional taxonomy adheres with impediments in distinguishing either closely related species or those specimens falsified for trade interest.

Table 2
Confused species status of the studied ornamental fishes based on similarity match with the database. The developed sequences of the specimen that revealed similarity with the sequences of multiple species in BOLD-IDS and GenBank as Category 1 (Sl. Nos. 1–6) and the remaining samples with the lack of conspecific sequences in the database are grouped as Category 2 (7–21). Further analysed with morphology, NJ clustering approach and K2P distance based method for confirmation of those species.

Sl. no.	Specimen with trade name	Sequences generated (GenBank Acc. no) (size in bp)	Close match in with species in BOLD-IDS		Close match in GenBank (BLASTN similarity %)
			Species level barcode records (process ID)	Public record barcode database (process ID)	
1	Green Swamp barb	JN815309 (639 bp), JQ713852 (655 bp)	*Puntius conchonius* *P. chola*	*Puntius conchonius* (RCYY532-11) *P. chola* (SRFBI033-11)	*Puntius conchonius (99)* *P. chola (99)* *P. fraseri (98)*
2	Swamp barb	JN815285 (620 bp), JN815286 (624 bp)	*Puntius conchonius* *P. chola*	*Puntius conchonius* (RCYY532-11) *P. chola* (SRFBI033-11)	*No match (0)*
3	Gagata	KF511518 (624 bp)	*Gagata cenia*, *G. gagata*, *Nemapteryx macronotacantha*	*Gagata cenia* (CFISH035-12, CFISH037-12, GBGC9466-09), *G. gagata* (ANGBF6120-12)	*Gagata cenia (99–100)*
4	Indian Gagata	KF511519 (629 bp)	*Gagata cenia*, *G. gagata*, *Nemapteryx macronotacantha*	*Gagata cenia* (CFISH035-12, CFISH037-12, GBGC9466-09), *G. gagata* (ANGBF6120-12)	*Gagata cenia (99–100)*
5	Miniscale shark	JQ713849 (576 bp)	*Labeo gonius*, *L. fimbriatus*	*Labeo gonius* (GBGC4209-08, GBGC4210-08) *L. fimbriatus* (GBGCA2387-13)	*Labeo gonius (99–100)*
6	Minor carp	KF511569 (620 bp), KF511570 (620 bp), KF511571 (620 bp), KF511572 (620 bp), KF511573 (620 bp), KF511574 (620 bp)	*Labeo bata*, *Labeo boga*	*Labeo bata*, *Labeo boga*	*Labeo bata (100)*, *Labeo boga (100)*
7	Queen loach	JX105468 (627 bp), JX105475 (633 bp), JX105478 (618 bp), KF511556 (620 bp)	No match	No match	*Botia almorhae (99)*
8	Bengal loach	JX105476 (637 bp), JX105477 (618 bp)	No match	No match	*Botia almorhae (99)*
9	Moosefaced loach	JX105467 (565 bp)	No match	No match	*Canthophrys gongota (96)*
10	Gagata	KF511516 (629 bp), KF511517 (629 bp), KF511553 (552 bp)	No match	No match	*Gogangra viridescens (97)*
11	Kalabans	KF511531 (620 bp), KF511532 (624 bp), KF511533 (624 bp), KF511534 (624 bp), KF511535 (626 bp)	No match	No match	*Bangana sp. (100)*
12	Kingfish	KF511536 (617 bp), KF511557 (565 bp)	No match	No match	*Capoeta antalyensis (91)*, *Schizothorax sinensis (91)*
13	Garra	KF511537 (629 bp), KF511538 (616 bp), KF511539 (631 bp), KF511540 (620 bp), KF511541 (557 bp)	No match	No match	*Garra tengchongensis*, *Garra orientalis (92)*
14	Two stripe gulper catfish	KF511544 (629 bp)	No match	No match	*Ompok bimaculatus (95)*
15	Black spot grouper	KF511545 (626 bp)	No match	No match	*Channa aurantimaculata (93)*
16	Indian Hill trout	KF511550 (624 bp), KF511551 (624 bp)	No match	No match	*Raiamas bola (87)*
17	Ladder loach	KF511554 (568 bp)	No match	No match	*Botia dario (92)*
18	Gangetic loach	KF511555 (568 bp)	No match	No match	*Botia dario (92)*
19	Zipper loach	KF511558 (601 bp)	No match	No match	*Acanthocobitis botia (87)*
20	Striped loach	KF511559 (588 bp)	No match	No match	*Acanthocobitis botia (87)*
21	Torrent stone carp	KF511560 (594 bp)	No match	No match	*Psilorhynchus homaloptera (98)*



DNA barcode technology is proving to be instrumental in species identification (Hebert et al., 2003; Hajibabaei et al., 2007a, 2007b) even in case of any uncertainty in identification of the specimens and studying biodiversity, like fish (April et al., 2011), Lepidoptera (Hausmann et al., 2011), coral reef fish (Hubert et al., 2010), medicinal plants (Chen et al., 2010), spruce budworm food web (Smith et al., 2011), parasitoid flies (Diptera) (Smith et al., 2006), fresh water Fish (Ward et al., 2005; Valdez-Moreno et al., 2009; Ward et al., 2009), aquarium fish (Steinke et al., 2009; Collins et al., 2012), catfish (Wong et al., 2011; Bhattacharjee et al., 2012), overlooked marine fish (Zemlak et al., 2009), fish (Zhang, 2011; Zhang and Hanner, 2012), Mahseer fish (Laskar et al., 2013), and Indian fresh water fish (Chakraborty and Ghosh, 2014a, 2014b). There are also evidences where *CO1* barcode sequence is proved to be useful in authentication of commercial biological items used in trade like tea (Stoeckle et al., 2011), tuna sushi (Lowenstein et al., 2009), frozen fish items (Cutarelli et al., 2014), herbal food supplement (Little and Jeanson, 2013), etc. and in addition for trade monitoring for threatened animals (Luo et al., 2013). Here, we employed the combined approach of morphotaxonomy and *CO1* barcode sequence as a reliable species identification tool to categorize the ornamental fish species being traded from NE region of India with maximum precision where there is a hurdle in distinguishing species even if the specimen is falsified using multiple trade names and fallaciously catalogued.

2. Materials and methods

2.1. Sampling

The ornamental fish specimens of Northeast India were collected from different fish collectors and aquarium fish traders located within the regions. A total of 128 specimens were maintained in the Department of Biotechnology, Assam University Silchar based on preliminary identification by taxonomic keys. The details of the sample specimens with their trade names and voucher IDs are given in the Supplementary Table S1. The studied fishes are routinely caught by the expert fish collectors and subjected to trade and therefore, the study was carried out as per institutional ethical guidelines.

2.2. CO1 amplification and sequencing

The genomic DNA was extracted from the tissue samples from caudal fin region collected aseptically from each of the specimens in TES buffer (50 mM Tris–HCl, 25 mM EDTA and 150 mM NaCl) and DNA was extracted by Phenol-Chloroform-Isoamyl alcohol method (Sambrook and Russell, 2006). The amplification of barcode region of *CO1* gene was done with 96 well Veriti PCR (Applied Biosystem, Inc., USA) using published primers Fish F1–5′ TCAACCAACCACAAAGACAT TGGCAC 3′ and Fish R1–5′ TAGACTTCTGGGTGGCCAAAGAATCA 3′ (Ward et al., 2005). PCR was carried out using 1 μl of 20 pmol of each primer, 1 μl of 20–50 ng of total genomic DNA, 3 μl of 10× PCR buffer, 3 μl of 10 mM dNTPs mix (Promega, USA) and 1 U of High Fidelity DNA polymerase (Fermentas, Canada) in a final reaction volume of 30 μl. The thermo-cycling condition was 94 °C for 5 min as initial denaturation followed by 30 cycles of 1 min at 94 °C, 45 s at 50 °C and 1 min at 72 °C. A final 8 min extension step was done at 72 °C. The PCR-amplified products were visualized in 1.5% agarose gel in 1× TAE buffer (Tris–acetic acid–EDTA) with ethidium bromide staining (10 mg/ml).

Fig. 1. Neighbour-Joining tree of *CO1* gene sequences derived from ornamental fish species under trade using K2P distance parameter. The numbers at the nodes are the bootstrap values based on 1000 replicates. The GenBank accession number and the species names are shown for each of the taxa. The black dot marks on the tree correspond to the generated sequence of 128 fish specimens for the study.

Table 3
Minimum Interspecific and Maximum Intraspecific K2P genetic distance, NJ-Cluster with bootstrap values of the specimens that revealed similarity in the range of 96%–100% with the database sequences.

Specimens	Identified to be	NJ-Cluster of the developed and database sequences (bootstrap) from Fig. 1	Max K2P Distance within the same species	Min K2P Distance with other species
Striped snakehead Banded snakehead Snakehead murrel	*Channa striata*	Clustered as a three groups (82, 99) under a single node (99)	0.044	0.191
Walking snakehead Ceylone snakehead	*Channa orientalis*	Formed two sub-clusters (99, 99) under a single node (100)	0.042	0.172
Tire track eel	*Mastacembalus armatus*	Formed three sub-clusters (50, 70) under a single node (100)	0.037	0.152
Moosefaced loach Gongota loach	*Canthophrys gongota*	Clustered under single node (99)	0.043	0.236

The amplified products were purified and sequenced bi-directionally using automated DNA Sequencer (ABI 3500 Genetic Analyzer; Applied Biosystem, Inc., USA). The sequencing reaction was performed using 1 μl of BDTv3.1, 1.5 μl of 5× sequencing buffer; a final concentration of 0.5 pmol of each of the primers was maintained in separated reaction. The chain termination reaction was carried out at 25 cycles of 1 min at 94 °C, 5 s at 50 °C and 4 min at 60 °C. The fragments were then purified by sodium acetate/EDTA/ethanol method prior to run in the 3500 Genetic Analyzer.

2.3. Data analysis

The generated sequences were checked by Sequence Scanner v1 and Seq Scape v2.7 (Applied Biosystems) and further analysed by nucleotide BLAST (Altschul et al., 1990) search tool at the National Centre for Biotechnology Information (http://www.ncbi.nlm.nih.gov/blast) to check percentage similarity of the developed sequences with the database sequences. High similarity of the developed sequences with the database *COI* barcode sequences without any indels and coherent amino acid codes with a partial fragment of mitochondrial *COI* gene confirmed the sequences being correct and no NUMT being amplified (Zhang and Hewitt, 1996). The sequences were also aligned by using CLUSTALX. All the sequences were submitted in BOLD under the project "OFISH-DNA barcoding of ornamental fishes of North-East India". The sequences were also submitted in GenBank and received valid accession numbers as described in Supplementary Table S1.

The identification of the specimens was mainly done by the combined approaches of similarity match, morphology based assessment and conventional method of NJ and K2P distance based approach. First, the sequences were subjected to similarity match in the BOLD species identification system (BOLD-IDS, www.barcodinglife.org) (Ratnasingham and Hebert, 2007) using both Species Level Barcode Records and Public Record Barcode Database, where the generated or query sequences are searched for a closest match with a minimum of three representatives and a maximum conspecific divergence of 2% of a respective species. All the search results are further confirmed by morphological keys.

Furthermore, the sequences are also subjected to similarity match in GenBank database through the Mega-BLAST program which is optimized for higher similarity with pre-existing database sequences with the default settings, where match/mismatch score is 1, −2 for identification of the samples at the species level (Chen et al., 2010; Bhattacharjee et al., 2012). The range of 97–100% similarity match of the query sequence with the reference sequence having E-value lower than the cut-off was expressed as significant to categorize the query sequences into their respective species. Secondly, morphological parameters were extensively used for some cases, which showed ambiguity in similarity approach. In other words, it could not be determined through *CO1* barcodes alone. Third, conventional method of Neighbour Joining (NJ) clustering (with 1000 bootstrap support) and distance matrices was performed using the Kimura 2-parameter model (Hebert et al., 2003) through MEGA5.2 (Tamura et al., 2011) by taking the closely matched conspecific sequences from the database as

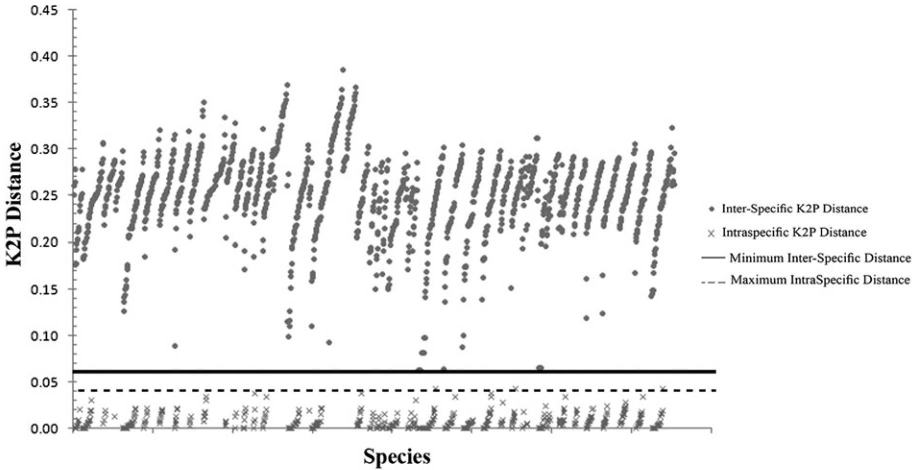


Fig. 2. The scatter plot of genetic distances within species against genetic distances between the species, the red cross marks showing intraspecific K2P distance and the blue dots showing interspecific K2P distance. A straight line is at minimum interspecific K2P distance with a dotted line at maximum intraspecies distance. The barcode gap, taken as the minimum interspecific distance was straightforward in delineating all the species taking on the study.

Identified Species	Species Status												
	International Union for Conservation of Nature, IUCN (2013)						Conservation Assessment and Management Plan, CAMP (1998)						
	Threatened Category			LC	DD	NA	Threatened Category			LRnt	LRlc	DD	NA
	EN	VU	NT				EN	VU	NT				
Puntius sophore				■						■			
Amblypharyngodon mola				■							■		
Crossocheilus latius				■								■	
Chela cachius				■									■
Tetraodon cutcutia				■						■			
Glossogobius guiris		■											■
Nandus nandus				■						■			
Parambassis ranga				■									■
Channa punctatus						■				■			
Badis badis				■									■
Notopterus notopterus				■						■			
Macrognathus pancalus						■				■			
Mastacembelus armatus				■									■
Barbonymus gonionotus				■									■
Labeo calbasu)				■						■			
Labeo gonius				■						■			
Esomus danricus				■							■		
Rhinomugil corsula				■				■					
Devario aequipinnatus				■						■			
Chitala chitala			■				■						
Anabas testudineus					■			■					
Gagata dolichonema				■									■
Eutropiichthys murius				■						■			
Pseudecheneis sulcata				■				■					
Puntius manipurensis		■											■
Barilius bendelisis				■						■			
Sisor rabdophorus				■			■						
Mystus vittatus				■				■					
Mystus gulio				■									■
Ompok bimaculatus			■				■						
Ailia coila			■					■					
Clarias batrachus				■				■					
Puntius chola				■				■					
Botia dario				■									■
Canthophrys gongota				■									■
Channa striata				■							■		
Channa orientalis				■				■					
Devario devario				■									■
Gogangra viridescens				■									■
Gagata cenia				■									■
Labeo dero				■				■					
Semiplotus semiplotus		■											■
Garra arupi						■							■
Ompok pabda			■				■						
Channa melanostigma						■							■
Barilius barna				■						■			
Channa marulius				■						■			
Botia rostrata		■								■			
Acanthocobitis botia				■									■
Psilorhynchus homaloptera													
Labeo bata				■						■			

replicates. The database sequences taken in the study are summarized in Table S2.

2.4. Morphology of the specimens

The specimens which could not be resolved by similarity match approach were categorized based on the taxonomic characters that were available from the original description and subsequent re-description. For identification, the diagnostic morphology keys are recorded as per the standard literature (Hamilton, 1822; Kottelat, 1990; Talwar and Jhingran, 1992; Roberts and Ferraris, 1998; Vishwanath and Kosygin, 2000; Vishwanath and Laisram, 2004; Heok Hee, 2007; Nebeshwar et al., 2009; Geetakumari and Vishwanath, 2011; Dishma and Vishwanath, 2012; Shangningam et al., 2013). All the measurements were recorded using a digital slide calliper (0.01 mm). Those nonquantitative morphological characters, e.g.; colour, blotch, and bands/stripes were also recorded. All the observed keys were summarized in the Supplementary Table S3.

3. Result

3.1. Species identification by pre-existing sequences in the database

The collected ornamental fishes were barcoded and submitted in the GenBank as well as in BOLD under the project, "DNA barcoding of ornamental fishes of North-East India [OFISH]". The DNA barcode database provides a system of species identification based upon the finding of the closest match of the query sequences with database sequences. The extensive species identification based on the consensus of the similarity match on BOLD based on Species Level Barcode Records and Public Record Barcode Database and GenBank for the studied ornamental fish specimen revealed straightforward identification for 84 specimens belonging to 35 species (Table 1) which are further confirmed by morphologic keys. Among them, the samples bearing single trade names like Gangetic latia, Neon Hatchet, Tank goby, Glass fish, Barred Spiny eel, Labeo, Miniscale shark, Flying barb, Silver barb and Corsula mullet, Giant Danio, Clown knifefish, Suckerthroat, Indian Hill Trout, Indian Whiptail catfish, Long Whiskers Catfish, Striped dwarf catfish, Butter catfish, Gangetic ailia, Walking catfish, Gagata, River catfish, Turquoise Danio, Tire track eel, and bullseye snakehead each showed significant similarity with each of the distinct species. However, in a few instances, the samples with multiple trade names showed the closest match with a particular species. For example, all the sequences of Pool barb, Spot fin barb showed 97–99% similarity with *Puntius sophore*, Mola carplet, Pale carplet showed 99% similarity with *Amblypharyngodon mola*, Mud perch, Leaf perch showed 99% similarity with *Nandus nandus*, Blue badis, Chameleon fish showed 98–99% similarity with *Badis badis*, Spotted snakehead, Green snakehead fish or Checkered snake fish showed 99% similarity with *Chana punctatus*, Grey featherback, Bronze Featherback showed 99% similarity with *Notopterus notopterus*, Climbing perch, Climbing gourami showed 98% similarity with *Anabas testudineus*. Some of the specimens showed a range of 96–99% similarity with the database sequence like Tire track eel with *Mastacembelus armatus* (96–97%), Snakehead murrel, Striped snakehead, Banded snakehead with *Channa striata* (96–99%), Walking snakehead, Ceylone snakehead with *Channa orientalis* (96–99%).

3.2. Identification of ambiguous cases by morphology

The remaining 44 specimens could not be identified to species based on barcoding alone and showed contradictory results in the respective database and grouped as Category 1 for the 12 specimens who could not be distinguished between several highly similar species. The Category 2 for the leftover 32 specimens which do not have vouchered pre-existing conspecific sequences in the database with which similarity match is to be made was presented (Table 2). All those cases were diagnosed using morphological parameters described in the available literature to determine the correct species. All the identified specimens along with the diagnostics keys were listed in Supplementary Table S3.

3.2.1. Category 1: higher-level similarity with multiple species

The 13 specimens under six trade names in Category 1 were identified into four species, where the specimens of Swamp barb, Green Swamp barb were identified to be *Puntius chola* (Hamilton, 1822; Vishwanath and Laisram, 2004), Gagata and Indian Gagata were identified to be *Gagata cenia* (Supplementary Fig. S1. A1) (Hamilton, 1822; Roberts and Ferraris, 1998) and Minor carp, Miniscale shark was identified as *Labeo bata* and *Labeo gonius* respectively (Talwar and Jhingran, 1992).

3.2.2. Category 2: newly generated sequences

The remaining 31 specimens under Category 2 with 15 trade names which are the newly generated sequences were identified to 12 species. The specimens, like Gagata is identified to *Gogangra viridescens* (Supplementary Fig. S1. A2). The dorsal fin rays were 6; Pectoral fin rays were 8; Anal fin rays were 8 and most importantly, outer and inner mental barbles were found widely separated from the origin of inner barbles anterior to the origin of the outer barbles while the maxillary barble extending to the pectoral fin base (Roberts and Ferraris, 1998). Queen loach, Bengal loach were identified to be *Botia dario*. The specimen had 7 vertical bars across the body with narrower interspaces (Supplementary Fig. S1. B1) and lacks any distinct markings on the interspaces unlike *Botia udomritthiruji*, a species described from Southern Myanmar, which has 5 vertical bar and have markings on the interspaces. Two specimens viz. Ladder loach and Gangetic loach were identified to be *Botia rostrata* by morphological parameters (Heok Hee, 2007). The studied specimens had dorsal fin with i9 rays, Anal fin with i5 rays, Pectoral fin rays was ii10; Pelvic fin with i7 rays; Caudal fin rays is found to be i,9 + 8,i. Lateral line scale is complete, and there are pale spots within the dark vertical bars and all the bars are interconnected like a network (Supplementary Fig. S1. B2). The specimens of Garra are identified to be *Garra arupi* (Supplementary Fig. S1. C), a new species described from Arunachal Pradesh of India (Nebeshwar et al., 2009). Kingfish is identified as *Semiplotus semiplotus*, the meristic counts revealed 25 dorsal soft rays, anal spines were 2, anal soft rays were 7, the last simple dorsal fin ray spine was not serrated and transverse row of 10–12 open pores (5–6 on each side) across the snout directed posterior toward the middle of orbit were observed (Supplementary Fig. S1. D) (Vishwanath and Kosygin, 2000). In case of Indian Hill trout, the morphological parameters were recorded and found to be within the range of description (Dishma and Vishwanath, 2012), among the other characters the barbles were absent (Supplementary Fig. S1. E1, E2), which confirmed our specimens to be *Barilius barna*. Black spot grouper is analysed by morphology and identified as *Channa melanostigma*, a new snakehead reported from NE India as compared with original description (Geetakumari and Vishwanath, 2011). Similarly, Zipper Loach, Striped loach is identified as *Acanthocobitis botia* (Kottelat, 1990), Two Stripe Gulper Catfish is identified as *Ompok pabda* (Hamilton, 1822; Talwar and Jhingran, 1992), Torrent stone carp as *Psilorhynchus homaloptera* based on described morphologic keys (Shangningam et al., 2013), Moosefaced loach is identified as

Fig. 3. The list of the identified fish species is being traded from Northeast India as ornamental fish and their status in both IUCN and CAMP. Among the 128 fish samples 17 species belong to threatened category as compared to IUCN and CAMP. EN: endangered, VU: vulnerable, LC: least concern, LRnt: lower risk near threatened, LRlc: lower risk least concern, DD: data deficient, NT: near threatened, NA: not available.

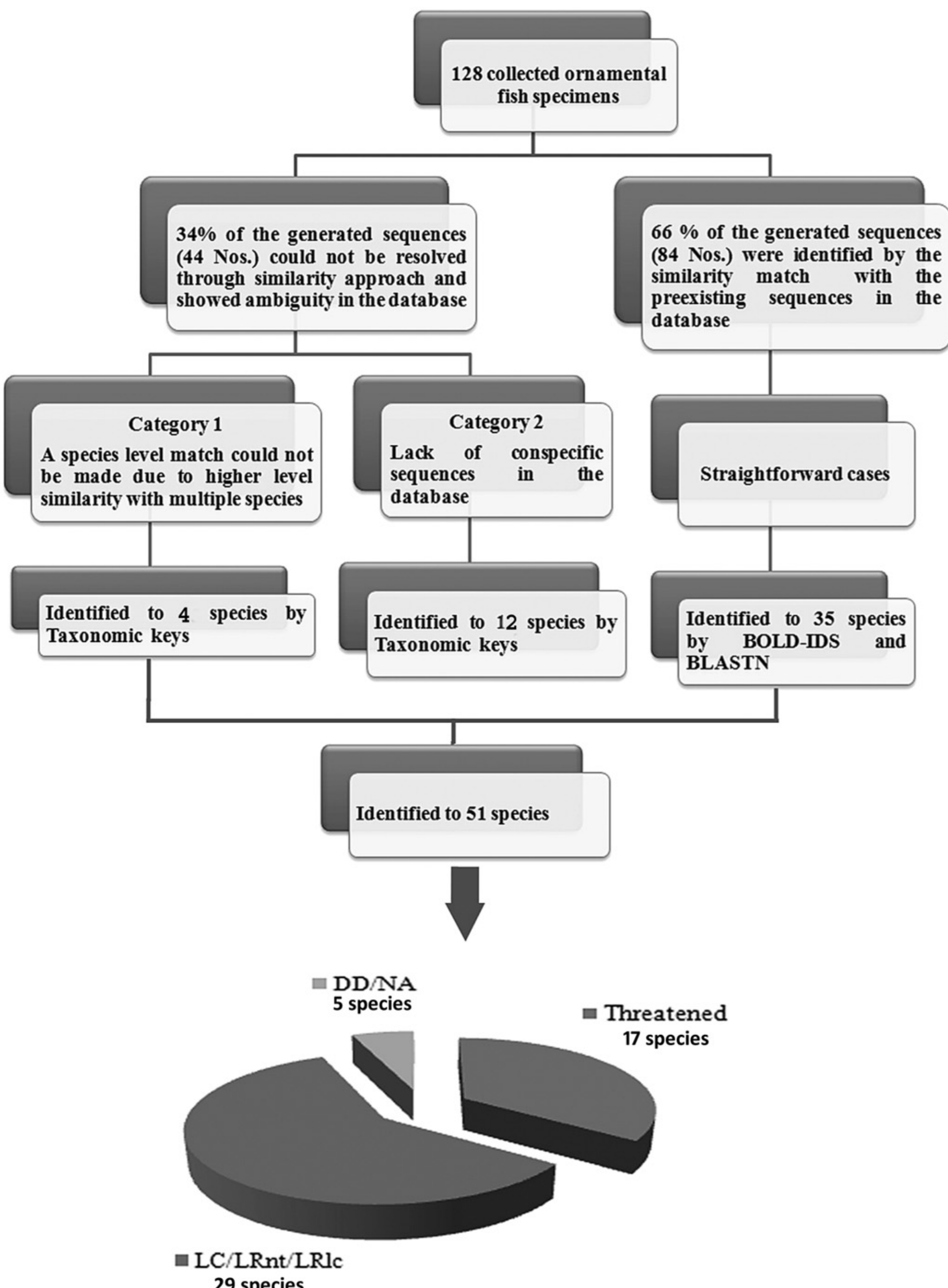


Fig. 4. A total of 128 sequences were generated for the study out of which 66% of the cases were found to be straightforward to identify 35 respective species by similarity match approach. The remaining 34% ambiguous cases were solved by morphotaxonomy and found to represent 16 species. Among the total 51 identified, 17 species were in a threatened category.

Canthophrys gongota based on original description (Hamilton, 1822), the specimen of Kalabans is recognised as *Labeo dero* (Hamilton, 1822; Talwar and Jhingran, 1992) the morphological characters were detailed in Supplementary Table S3.

3.3. Neighbour-Joining and K2P distance

The 84 specimens which showed significant similarity with 35 species clustered cohesively with the matched species and distinctly with respect to others in the NJ-tree. Our data, DNA barcode based studies showed intraspecies individuals that showed similarity in the range of 97–100% clustered cohesively with each other while distinct with respect to interspecies (Bhattacharjee et al., 2012). The remaining 44 specimens which were identified into 16 species by morphology, also showed cohesive clusters with the individuals of the same species and distinctly with the other identified species (Fig. 1). The specimens of Green Swamp barb/Swamp barb, showed higher similarity with the multiple species like *P. chola*, *Puntius conchonius*, and *Puntius fraseri* in the database. The morphology revealed our specimens to be *P. chola*. This is also a consensus with NJ clustering where the database sequence of *P. chola*, *P. conchonius*, and *P. fraseri* clustered with the generated sequences of *P. chola* as a cohesive unit. Similarly, Gagata/Indian Gagata, and Minor carp are confirmed to be *G. cenia* and *L. bata* respectively. Such anomalies may arise due to the presence of mislabelled sequence

in the database (Meier et al., 2006). Within the clusters, the mean K2P distance was found to be 0.0084 ± 0.002, and maximum K2P distance was 0.045. While, between the clusters, mean K2P distance was observed as 0.252 ± 0.024 with Minimum K2P distance of 0.056. Among them, 11 specimens representing four species showed deep conspecific divergence; concordant with the similarity match and are detailed in Table 3. The barcode gap, taken as the least inter-specific distance (Meier et al., 2008; Bhattacharjee et al., 2012) between the distinctly clustered congeners was straight forward in comparison to the cohesively clustered conspecies (Fig. 2), as we are only identifying the specimen in this case. However, a detailed study on the divergent pattern is still required. As the above facts were evident to the 128 sequences, therefore, all the specimens were identified belonging to 51 species.

4. Discussion

This study is undertaken to identify the ornamental fishes traded from NE India through DNA Barcoding approach, since, it is the harbour of many endemic and diverse fish species (Ponniah and Sarkar, 2000; Allen et al., 2010). However, the effective monitoring of ornamental fish species traded is constrained by lack of accurate, quantitative and unbiased information (Murray et al., 2012; Raghavan et al., 2013) and shortage of taxonomic intervention. Since, ornamental fish trade is one of the large global industries (Tlusty et al., 2013) hence, there is a vital need of inventorying the wealth of such bio-resources within the region. The combined approach of *CO1* DNA Barcoding and morphotaxonomy was employed to identify the collected 128 specimens into respective species, as *CO1* barcode region is already proven to be gold standard for discriminating animal species. Its short length can be sequenced quickly, easily and enough to distinguish species, due to its high sequence variability between the species as compared to within the species (Hebert et al., 2003, 2004). The barcode gap which was calculated as the minimum interspecific distance readily delineates 128 specimens to 51 species. Among the certain cases, the similarity search result showed a range of 96–99% similarity with the database sequences viz. *M. armatus*, *C. striata*, *C. orientalis*, and *C. gongota*. Besides, these sequences clustered as 2 or 3 cohesive units under a single node with respect to the species with which they showed similarity (Fig. 1) and the K2P distance within, and between the clusters was slightly higher than the divergence that appeared for the other straightforward cases (Table 3). The sequences of *C. orientalis* clustered separately from the database sequences forming two distinct clusters originating from the single node, the K2P distance between the two sub-clusters, although showed a higher divergence of 4.2%, but it is below the minimum interspecific K2P distance as barcode gap. Similarly, *C. striata* clustered as a three cohesive unit under a unique node with a bootstrap support of 99%. The maximum K2P distance between the sub-cluster showed a higher divergence of 4.4%, but it is also below the barcode gap. This may arise due to geographic isolation of the population as there are few database sequences from Philippines (HQ654691), Tamil Nadu, southern part of India (EU342203, EU342204), etc. Such high conspecific divergence was also previously reported in *C. striata* due to geographical isolation and substantial habitat re-organization (Jamsari et al., 2011), similar may be the case with *M. armatus* and *C. gongota*. Thus, in congruence to the earlier studies, we tentatively consider them as species having deep intraspecific divergence. The minimum interspecific K2P distance as threshold for species delineation is put forwarded by many workers (Meier et al., 2008; Bhattacharjee et al., 2012; Chakraborty and Ghosh, 2014a) and is used in the current study to identify species that is high conspecific divergent so as to eliminate the possible misidentification induced. However, the misidentification can only be identified and eliminated effectively by the combined approach of morphology and distance based algorithms such as setting up threshold level for species delineation as well as clustering method. Moreover, a detailed study on the species diversity covering the entire taxa is further required to determine the presence of possible crypticism, overlooked species or possibility of misidentification.

Among, all the above cases, we identified some interesting instances. In a few cases, although the specimens have different trade names, they represented same species. Many ornamental fish species, although having threatened status in IUCN, were marketed based on either multiple trade names or generic label of 'live ornamental fish' or 'live aquarium fish' or sometimes by the group label of 'Snakeheads', 'Barbs' rather than zoological nomenclature and thereby escapes regulations (Raghavan et al., 2013). The trade could still be regulated if each of the species was marketed with a single trade name with the corresponding zoological nomenclature. On the other hand, as DNA barcode is dependent on gene sequence of *CO1*, it will not change even if the species were misidentified or traded by multiple trade names. Hence, barcoding technology has an important role in regulating such case. Furthermore, it was evident throughout this study that, many threatened species were traded and for a few species, the IUCN (International Union for Conservation of Nature) and CAMP (Conservation Assessment and Management plan), which followed IUCN Red list Criteria for the assessment of species status in the wild, revealed contradictory species status (Fig. 3). Among them, the species status of *Rhinomugil corsula*, *Pseudecheneis sulcata*, and *Mystus vittatus* in IUCN is 'Least Concern' but the same is 'Vulnerable' as per CAMP, *A. testudineus*, the species status in IUCN is 'Data Deficient' but the same is 'Vulnerable' as per CAMP, similarly, the species status of *Puntius manipurensis* and *Glossogobius guiris* is 'Vulnerable' in IUCN, but the status of the same is 'Not Available' in CAMP. Therefore, species status of many fish appeared to be not accurately evaluated and needs revision. In this study, 51 species were identified from the traded samples. Among them, 14 species had multiple trade name, two species had single trade names. Importantly, 17 species belong to the threatened category and 29 species belong to lower risk near threatened/least concern and the rest are deficient in data (Fig. 4). Nevertheless, among the 14 species, four species having multiple trade names fall within the threatened category, and the species have been facing tremendous risks of jeopardising due to the ongoing uncontrolled exploitation of resources through trade with them.

5. Conclusion

This study focuses the imprudent exploitation scenario of high germplasm in nature and demands immediate attention. Furthermore, the study suggests enacting regulations that would resist the use of trade names or generic names while exporting a biological resource and recommend the use of proper zoological names of the species being traded. As anticipation, the survey highlighted the usefulness of DNA Barcoding technique in monitoring the trade of threatened species and demand conservation strategy for sustaining wildlife.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.gene.2014.11.037.

Acknowledgements

We sincerely acknowledge the Department of Biotechnology, Government of India for providing infrastructure facilities (BT/HRD/01/002/2007). We are thankful to Dr. Amalesh Choudhury, S.D. Marine Biological Research Institute, West Bengal for collection and identification of fish specimens, and to Dr Lalramliana and his team, Department of Zoology, Pucchunga University College, Mizoram for taxonomic support. We also thank Dr. Boni Amin Laskar, Bijoya Khondram and Mohua Chakraborty, Department of Biotechnology, Assam University for collection of fish samples from different parts of Northeast India which is highly acknowledged.

References

Allen, D.J., Molur, S., Daniel, B.A., IUCN—The World Conservation Union, IUCN Species Survival Commission, Zoo Outreach Organisation, 2010. The Status and Distribution of Freshwater Biodiversity in the Eastern Himalaya. IUCN, Gland, Switzerland.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

April, J., Mayden, R.L., Hanner, R.H., Bernatchez, L., 2011. Genetic calibration of species diversity among North America's freshwater fishes. Proc. Natl. Acad. Sci. U. S. A. 108, 10602–10607.

Bhattacharjee, M.J., Laskar, B.A., Dhar, B., Ghosh, S.K., 2012. Identification and re-evaluation of freshwater catfishes through DNA barcoding. PLoS One 7, e49950.

Bhattacharya, B.K., Choudhury, M., 2004. Ornamental Fish Culture and Trade in Northeastern India. Central Inland Fisheries Research Institute (CIFRI), Kolkata.

Chakraborty, M., Ghosh, S.K., 2014a. An assessment of the DNA barcodes of Indian freshwater fishes. Gene 537, 20–28.

Chakraborty, M., Ghosh, S.K., 2014b. Unraveling the sequence information in COI barcode to achieve higher taxon assignment based on Indian freshwater fishes. Mitochondrial DNA 10, 1–3.

Chen, S., Yao, H., Han, J., Liu, C., Song, J., Shi, L., Zhu, Y., Ma, X., Gao, T., Pang, X., Luo, K., Li, Y., Li, X., Jia, X., Lin, Y., Leon, C., 2010. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. PLoS One 5, e8613.

Collins, R.A., Armstrong, K.F., Meier, R., Yi, Y., Brown, S.D., Cruickshank, R.H., Keeling, S., Johnston, C., 2012. Barcoding and border biosecurity: identifying cyprinid fishes in the aquarium trade. PLoS One 7, e28381.

Cutarelli, A., Amoroso, M.G., De Roma, A., Girardi, S., Galiero, G., Guarino, A., Corrado, F., 2014. Italian market fish species identification and commercial frauds revealing by DNA sequencing. Food Control 37, 46–50.

Das, S.P., Biswas, J.N., 2009. A Handbook of Ornamental Fishes of the Brahmaputra Basin. Eastern Book House.

Dishma, M., Vishwanath, W., 2012. *Barilius profundus*, a new cyprinid fish (Teleostei: Cyprinidae) from the Koladyne basin, India. J. Threat. Taxa 4, 2363–2369.

FAO Fisheries and Aquaculture Dept. Fishery and Aquaculture Information and Statistics Service, 2008. FAO yearbook. Fishery and Aquaculture Statistics = FAO annuaire. Statistiques des pêches et de l'aquaculture = FAO anuario. Estadísticas de pesca y acuicultura. FAO, Rome, p. 72 v.

Geetakumari, K., Vishwanath, W., 2011. *Channa melanostigma*, a new species of freshwater snakehead from north-east India (Teleostei: Channidae). J. Bombay Nat. Hist. Soc. 107, 231–235.

Hajibabaei, M., Singer, G.A., Clare, E.L., Hebert, P.D., 2007a. Design and applicability of DNA arrays and DNA barcodes in biodiversity monitoring. BMC Biol. 5, 24.

Hajibabaei, M., Singer, G.A., Hebert, P.D., Hickey, D.A., 2007b. DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. Trends Genet. 23, 167–172.

Hamilton, F., 1822. An Account of the Fishes Found in the River Ganges and Its Branches, (Printed for A. Constable and company; etc., Edinburgh).

Hausmann, A., Haszprunar, G., Hebert, P.D., 2011. DNA barcoding the geometrid fauna of Bavaria (Lepidoptera): successes, surprises, and questions. PLoS One 6, e17134.

Hebert, P.D., Cywinska, A., Ball, S.L., deWaard, J.R., 2003. Biological identifications through DNA barcodes. Proc. Biol. Sci. 270, 313–321.

Hebert, P.D., Stoeckle, M.Y., Zemlak, T.S., Francis, C.M., 2004. Identification of birds through DNA barcodes. PLoS Biol. 2, e312.

Heok Hee, N., 2007. *Botia udomritthiruji*, a new species of botiid loach from southern Myanmar (Teleostei: Botiidae). Zootaxa 1608, 41–49.

Hubert, N., Delrieu-Trottin, E., Irisson, J.O., Meyer, C., Planes, S., 2010. Identifying coral reef fish larvae through DNA barcoding: a test case with the families Acanthuridae and Holocentridae. Mol. Phylogenet. Evol. 55, 1195–1203.

Jamsari, A.F., Pau, T.M., Siti-Azizah, M.N., 2011. Genetic structure of the snakehead murrel, *Channa striata* (channidae) based on the cytochrome c oxidase subunit I gene: influence of historical and geomorphological factors. Genet. Mol. Biol. 34, 152–160.

Kalita, T., Deka, K., 2013. Ornamental fish conservation in the flood plain wetlands of lower Brahmaputra Basin. Adv. Appl. Sci. Res. 4, 99–106.

Kottelat, M., 1990. Indochinese Nemacheilines: A Revision of Nemacheiline Loaches (Pisces:Cypriniformes) of Thailand, Burma, Laos, Cambodia, and Southern Viet Nam. Pfeil, München.

Laskar, B.A., Bhattacharjee, M.J., Dhar, B., Mahadani, P., Kundu, S., Ghosh, S.K., 2013. The species dilemma of Northeast Indian Mahseer (Actinopterygii: Cyprinidae): DNA barcoding in clarifying the riddle. PLoS One 8, e53704.

Little, D.P., Jeanson, M.L., 2013. DNA barcode authentication of saw palmetto herbal dietary supplements. Sci. Rep. 3, 3518.

Lowenstein, J.H., Amato, G., Kolokotronis, S.O., 2009. The real maccoyii: identifying tuna sushi with DNA barcodes—contrasting characteristic attributes and genetic distances. PLoS One 4, e7866.

Luo, J.Y., Yan, D., Song, J.Y., Zhang, D., Xing, X.Y., Han, Y.M., Yang, M.H., Dong, X.P., Peng, C., Chen, S.L., Xiao, X.H., 2013. A strategy for trade monitoring and substitution of the organs of threatened animals. Sci. Rep. 3, 3108.

Meier, R., Shiyang, K., Vaidya, G., Ng, P.K., 2006. DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. Syst. Biol. 55, 715–728.

Meier, R., Zhang, G., Ali, F., 2008. The use of mean instead of smallest interspecific distances exaggerates the size of the "barcoding gap" and leads to misidentification. Syst. Biol. 57, 809–813.

Murray, J.M., Watson, G.J., Giangrande, A., Licciano, M., Bentley, M.G., 2012. Managing the marine aquarium trade: revealing the data gaps using ornamental polychaetes. PLoS One 7, e29543.

Nebeshwar, K., Vishwanath, W., Das, D., 2009. *Garra arupi*, a new cyprinid fish species (Cypriniformes:Cyprinidae) from upper Brahmaputra basin in Arunachal Pradesh, India. J. Threat. Taxa 1, 197–202.

Ponniah, A.G., Sarkar, U.K., 2000. Fish biodiversity of North East India. NBFGR's NATP Special Publication No. 2. National Bureau of Fish Genetic, Resources, pp. 1–347.

Raghavan, R., Dahanukar, N., Tlusty, M.F., Rhyne, A.L., Krishna Kumar, K., Molur, S., Rosser, A.M., 2013. Uncovering an obscure trade: threatened freshwater fishes and the aquarium pet markets. Biol. Conserv. 164, 158–169.

Ratnasingham, S., Hebert, P.D., 2007. bold: The Barcode of Life Data Systemhttp://www.barcodinglife.org Mol. Ecol. Notes 7, 355–364.

Roberts, T.R., Ferraris, C.J., 1998. Review of South Asian Sisorid Catfish Genera Gagata and Nangra: With Descriptions of a New Genus and Five New Species. California Academy of Sciences, San Francisco, Calif.

Sambrook, J., Russell, D.W., 2006. The Condensed Protocols From Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Shangningam, B., Kosygin, L., Vishwanath, W., 2013. Redescription of *Psilorhynchus rowleyi* Hora and Misra 1941 (Cypriniformes: Psilorhynchidae). Ichthyol. Res. 60, 249–255.

Smith, M.A., Woodley, N.E., Janzen, D.H., Hallwachs, W., Hebert, P.D., 2006. DNA barcodes reveal cryptic host-specificity within the presumed polyphagous members of a genus of parasitoid flies (Diptera: Tachinidae). Proc. Natl. Acad. Sci. U. S. A. 103, 3657–3662.

Smith, M.A., Eveleigh, E.S., McCann, K.S., Merilo, M.T., McCarthy, P.C., Van Rooyen, K.I., 2011. Barcoding a quantified food web: crypsis, concepts, ecology and hypotheses. PLoS One 6, e14424.

Steinke, D., Zemlak, T.S., Hebert, P.D., 2009. Barcoding nemo: DNA-based identifications for the ornamental fish trade. PLoS One 4, e6300.

Stoeckle, M.Y., Gamble, C.C., Kirpekar, R., Young, G., Ahmed, S., Little, D.P., 2011. Commercial teas highlight plant DNA barcode identification successes and obstacles. Sci. Rep. 1, 42.

Talwar, P.K., Jhingran, A.G., 1992. Inland Fishes of India and Adjacent Countries vols. 1–2, (Balkema).

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739.

Tlusty, M.F., Rhyne, A.L., Kaufman, L., Hutchins, M., Reid, G.M., Andrews, C., Boyle, P., Hemdal, J., McGilvray, F., Dowd, S., 2013. Opportunities for public aquariums to increase the sustainability of the aquatic animal trade. Zoo Biol. 32, 1–12.

Valdez-Moreno, M., Ivanova, N.V., Elías-Gutierrez, M., Contreras-Balderas, S., Hebert, P.D., 2009. Probing diversity in freshwater fishes from Mexico and Guatemala with DNA barcodes. J. Fish Biol. 74, 377–402.

Vishwanath, W., Kosygin, L., 2000. Fishes of the cyprinid genus Semiplotus Bleeker 1859, with description of a new species from Manipur, India. J. Bombay Nat. Hist. Soc. 97, 92–102.

Vishwanath, W., Laisram, J., 2004. Two new species of Puntius Hamilton–Buchanan (Cypriniformes: Cyprinidae) from Manipur, India, with an account of Puntius species from the state. J. Bombay Nat. Hist. Soc. 101, 130–137.

Ward, R.D., Zemlak, T.S., Innes, B.H., Last, P.R., Hebert, P.D., 2005. DNA barcoding Australia's fish species. Philos. Trans. R. Soc. Lond. B Biol. Sci. 360, 1847–1857.

Ward, R.D., Hanner, R., Hebert, P.D., 2009. The campaign to DNA barcode all fishes, FISH-BOL. J. Fish Biol. 74, 329–356.

Wong, L.L., Peatman, E., Lu, J., Kucuktas, H., He, S., Zhou, C., Na-nakorn, U., Liu, Z., 2011. DNA barcoding of catfish: species authentication and phylogenetic assessment. PLoS One 6, e17812.

Zemlak, T.S., Ward, R.D., Connell, A.D., Holmes, B.H., Hebert, P.D., 2009. DNA barcoding reveals overlooked marine fishes. Mol. Ecol. Resour. 9 (Suppl. s1), 237–242.

Zhang, J., 2011. Species identification of marine fishes in China with DNA barcoding. Evid. Based Complement. Altern. Med. 2011, 978253.

Zhang, J., Hanner, R., 2012. Molecular approach to the identification of fish in the South China Sea. PLoS One 7, e30621.

Zhang, D.X., Hewitt, G.M., 1996. Nuclear integrations: challenges for mitochondrial DNA markers. Trends Ecol. Evol. 11, 247–251.

Identifying Ornamental Fishes of North-east India Through DNA Barcoding

Bishal Dhar and Sankar Kumar Ghosh

DNA barcoding and Genomic diversity Laboratory Department of Biotechnology
Assam University, Silchar-788011, Assam, INDIA
**Corresponding author's email : drsankarghosh@gmail.com*

Received: July 19, 2014: Accepted: November 12, 2014

Abstract

Aquarium fishes are traded from different parts of the world, including North-East India, and most of them are collected from the wild due to lack of established species-specific culture. The topographical and climatic diversity of NE India makes this region rich in fish resources. Most of the small fish which are treated as unwanted for conventional farming have good potency as ornamental fishes and are popularly known as Aquarium fishes. In the present study, identification of ornamental fish species exported from NE India were investigated through DNA barcoding approach. Altogether, 117 fish specimens were collected from different exporters of NE India, and their CO1 barcode sequences were developed. The sequences were subjected to similarity match in BOLD-IDS and analyzed using MEGA 5.2 for species identification through NJ clustering, and K2P distance based approach. The analysis revealed straightforward identification of 113 specimens into 53 Species. Thus, species-level identification through DNA barcoding reflects the efficacy of the technique in identifying ornamental fish trade. The DNA bases identifier along with the zoological name will helps to inventorying the trade of ornamental fishes and therefore the present study will help future researchers and others for easy access of the ornamental fishes of this region and will be of great help to conservationist and aquarists.

Keywords: DNA Barcoding, Ornamental Fish, CO1, K2P

Introduction

Aquarium fish keeping has become one of the most popular hobbies and fishes suitable for keeping in the aquarium fetch a high economic value across the globe contributing to increased growth in international ornamental fish trade. The global ornamental fish trade has increased dramatically in the last few years [1]. The Indian ornamental fish export is scanty in global share but vibrant with 20% rise per annum with annual income of US$ 1.2 million (Marine Products Export Development Authority, MPEDA).

Majority of the total Indian ornamental fish trade is rooted from wild catch and is contributed by this region of India [2,3]. Due to the diversity of topographic and climatic features of NE India, this region is rich in endemic fish. Most of the small food fish which are treated as unwanted for conventional farming have good potency as ornamental fishes and are popularly known as Aquarium fishes. These species attract hobbyists both locally and globally. This region of the country is important in view of large-scale production from capture fishery underlined by the existence of innumerable rivers, rivulets and lentic water bodies, which harbor plenty of diverse fish fauna. There are about 267 species belonging to 136 genera of fresh water fishes inhabiting in Northeast India [4]. Out of which, 54.32% possess either of the three values as food for human, component in angling tourism or aquarium fish trade, and are thus potential resources for the growth of economy. Aquarium fishes are categorized based on colour pattern (colourful), morphologically uniqueness (special), and behaviorally charismatic (semi-aggressive, community, non-community). The current checklist of fishes of North East India showed 250 potential ornamental fish species, Out of this, the highest no. recorded from Assam (187), followed by Arunachal Pradesh (165), Meghalaya (159), Manipur (139), Tripura(103), Nagaland(71), Mizoram(46), and Sikkim(29). Many endemic fish species are being traded from the wild harvest, due to lack of established species-specific culture or breeding, and serve as a threat to biodiversity from harvest pressure. Despite several regulatory enactments in India, aquarium fishes are traded largely without endowment to the government.

© *Mizoram University, Aizawl (India)*

Nevertheless, the inclusive documentation of ornamental fish resources from the North Eastern region of India has been lagging behind due to the lack of quantitative data and shortage of expert taxonomists associated with other factors too. Since, ornamental fish trade has gone global; hence, there is a paramount need of inventorying the regional/national wealth of such bio resources for recognizing proper equity sharing of their trade. Until now, ample efforts of inventorying the ornamental fishes have been accomplished through morpho-taxonomy; however, the conventional taxonomy adheres with impediments in distinguishing either closely related species. Currently, DNA Barcode technology is proving to be instrumental in species identification [5-7] even in case of any uncertainty in identification of the specimens and studying biodiversity, like Fish [8], Diptera [9], Lepidoptera [10], Mirids [11], Butterfly [12], Coral reef fish [13], Medicinal plants [14], spruce budworm food web [15], Parasitoid flies (Diptera) [16], Fresh water Fish [17-19], Catfish [20, 21], overlooked marine fish [22], Fish [23, 24], Mahseer fish [25]. There are also evidences where CO1 barcode sequence is proved to be useful in authentication of commercial food items used in trade like Tea [26], Tuna sushi [27], Frozen fish items [28], Herbal food supplement [29], etc. and also for trade monitoring for threatened animals [30]. Therefore, in the present study, we employed CO1 barcode sequence as a reliable species identification tool to categorize species with maximum precision where there is hurdle in distinguishing species.

Materials and Methods

Sampling and CO1 Amplification and Sequencing

The ornamental fish specimens of Northeast India were collected from different fish collectors located in the region. A total of 113 specimens were vouchered in the Department of Biotechnology, Assam University Silchar based on preliminary identification by taxonomic keys. The Genomic DNA was extracted from the tissue samples collected aseptically from each of the specimens using TES buffer (50mM Tris-HCl, 25mM EDTA and 150mM NaCl) and DNA was extracted by Phenol-Chloroform-Isoamyl alcohol method [31]. The amplification of *CO1* gene (655 bp) was done with 96 well Veriti PCR (Applied Biosystem, Inc. USA) using published primers Fish F1- 5′ TCAACCAACCACAAAGACATTGGCAC 3′ and Fish R1- 5′ TAGACTTCTGGGTGGCCAAAGAATCA 3′ [19]. PCR was carried out using 1µl of 20pmol of each primer, 1µl of 20–50ng of total genomic DNA, 3µl of 10x PCR buffer, 3µl of 10mM of dNTPs mix (Promega, USA) and 1U of High Fidelity DNA polymerase (Fermentas, Canada) in a final reaction volume of 30µl. The thermo-cycling condition were 94°C for 5min as initial denaturation followed by 30 cycles of 1 min at 94°C, 45sec at 50°C and 1min at 72°C. A final 8minutes extension step was done at 72°C. The PCR-amplified products were visualized in 1.5% agarose gels in 1x TAE buffer (Tris-Acetic acid-EDTA) with Ethidium Bromide staining (10 mgml). The amplified products were sequenced bi-directionally using automated DNA Sequencer (ABI 3500 Genetic Analyzer; Applied Biosystem, Inc. USA). The sequencing reaction was performed using manufacturer protocol.

Data Analysis

The generated sequence chromatograms were checked by Sequence Scanner v1 and SeqScape v2.7 (Applied Biosystems) and further analyzed by BLASTN [32] search tool at the National Centre for Biotechnology Information (http://www.ncbi.nlm.nih.gov/blast) to check percentage similarity of the developed sequences with the database sequences. High similarity of the developed sequences with the database *COI* barcode sequences without any indels and coherent amino acid codes with a partial fragment of mitochondrial *COI* gene confirmed the sequences being correct. The sequences were also aligned by using CLUSTALX. Moreover, all the sequences were greater than 600bp and without any indels and coherent amino acid codes with a partial fragment of mitochondrial *COI* gene confirmed the sequences being correct and no NUMT being amplified that confirmed no NUMT being amplified as the limits of NUMT in fishes are less than 600 bp [33]. All the sequences were submitted in BOLD under the project "OFISH-DNA barcoding of ornamental fishes of North-East India." The sequences were also submitted in GenBank and received valid accession numbers as described in supple

The identification of the specimens was mainly done by similarity match in the BOLD species identification system (BOLD-IDS, www. barcoding-life.org) [34,14,21] as well as conventional method of Neighbour Joining (NJ) clustering (with 1000 bootstrap support) and distance matrix were performed using the Kimura 2-Parameter model [5] through MEGA 5.2 [35] by taking the atleast 3 conspecific sequences from the database as replicates where only singleton specimen collected for the study.

Table-1: Straight forward Identification of the studied ornamental fishes based on similarity match with BOLD Identification System. The match is exclusively based on similarity of developed sequences with database sequences. 113 traded fishes identified as 53 species.

Sl No	Sequences generated (Acc. No Genbank)	Identified as species (Number)
1	JN815299, KF511500, KF511501, KF511502	*Crossocheilus latius* (1)
2	JN815300, JN815301	*Chela cachius* (2)
3	JN815296 , JQ713857, JN815294 , JN815295	*Glossogobius guiris* (3)
4	JN815274, JN815276, JN815283	*Parambassis ranga* (4)
5	JX105465, JN815289	*Macrognathus pancalus* (5)
6	JQ713848	*Labeo calbasu* (6)
7	JN673955, KF511504, KF511505, KF511506	*Esomus danricus* (7)
8	JQ713846	*Barbonymus gonionotus* (8)
9	JX105471	*Rhinomugil corsula* (9)
10	KF511497, KF511498, KF511499	*Devario aequipinnatus* (10)
11	KF511511	*Chitala chitala* (11)
12	KF511525, KF511526, KF511527	*Pseudecheneis sulcata* (12)
13	KF511547, KF511548, KF511549	*Barilius bendelisis* (13)
14	KF511561	*Sisor rabdophorus* (14)
15	KF511564	*Mystus gulio* (15)
16	KF511563,	*Mystus vittatus* (16)
17	KF511565	*Ompok bimaculatus* (17)
18	KF511566	*Ailia coila* (18)
19	KF511567	*Clarias batrachus* (19)
20	KF511520, KF511521, KF511522	*Gagata dolichonema* (20)
21	KF511523, KF511524	*Eutropiichthys murius* (21)
22	JX105479	*Devario devario* (22)
23	JN815288, KF511546	*Mastacembelus armatus* (23)
24	KF511552	*Channa marulius* (24)
25	JN815267, JQ713856, JQ713844	*Puntius sophore* (25)
26	JN815277, JN815278	*Amblypharyngodon mola* (26)
27	JN815306, JN815307, JQ713845	*Nandus nandus* (27)
28	JN245992, JN245990, KF511512, KF511513	*Channa punctatus* (28)
29	JN815304, JN815305, JN815311, JQ713858	*Badis badis* (29)
30	JN815287, KF511509, KF511510	*Notopterus notopterus* (30)
31	KF511514, KF511515	*Anabas testudineus* (31)
32	KF511528, KF511529, KF511530	*Puntius manipurensis* (32)
33	KF11542, KF11543	*Puntius Khugee*
34	JN245989, KF511507, KF511508	*Channa striata* (33)
35	JN245991, JX105472, JX105474	*Channa orientalis* (34)
36	JN815308	*Tetraodon cutcutia* (35)

37	KF511503	*Oreochromis mosambicus* (36)
38	KF511518, KF511519, JN628893	*Puntius chola*
39	JQ713849	*Labeo gonius*
40	KF511569, KF511571, KF511573, KF511574	*Labeo bata*
41	JX105468, JX105478, JX105476, JX105477	*Botia dario*
42	JX105467	*Canthophrys gongota*
43	KF511516, KF511517	*Gogangra viridescens*
44	KF511531, KF511535	*Labeo dero*
45	KF511536, KF511557	*Semiplotus semiplotus*
46	KF511539 KF511540, KF511541	*Garra arupi*
47	KF511544	*Ompok bimaculatus*
48	KF511545	*Channa aurantimaculata*
49	KF511550 , KF511551	*Barilius barna*
50	KF511554, KF511555	*Botia rostrata*
51	KF511558, KF511559	*Acanthocobitis botia*
52	KF511560	*Psilorhynchus homaloptera*
53	KF511575	*Cirrhinus cirrhosus*

Results and Discussion

Species identification by BOLD-IDS

The extensive species identification based on similarity match on BOLD-IDS for the studied ornamental fish specimens' revealed straightforward identification for 113 specimens belonging to 53 species (Table-1).

Neighbour-Joining and K2P Distance

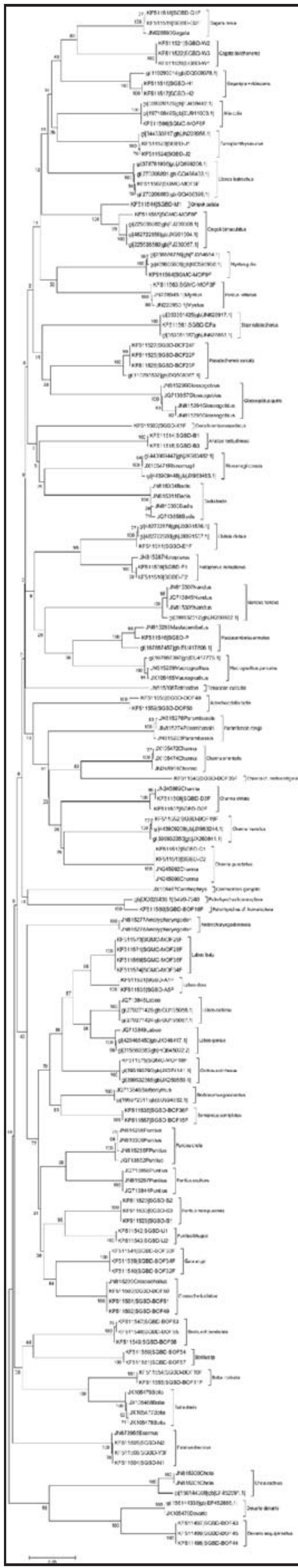
The 113 specimens which showed significant similarity with 53 species in the similarity match clustered cohesively with the matched species and distinctly with respect to others in the NJ-tree. The identified species also showed cohesive clusters with the conspecies and distinctly with the other identified species as illustrated in Figure 1. Within the clusters, the mean K2P distance was found to be 0.004 + 0.002, and maximum K2P distance was 0.007. While, between the cluster, mean K2P distance was observed as 0.250 + 0.022 with Minimum K2P distance of 0.056.

This study is undertaken to identify the ornamental fishes traded from NE India through DNA Barcoding approach, since, it is the harbor of many endemic and diverse fish species [4,36]. But the effective monitoring of ornamental fish species traded is constrained by lack of accurate, quantitative and un-biased information [37,38] and shortage of taxonomic intervention. Since, ornamental fish trade is one of the large global industries hence, there is a vital need of inventorying the wealth of such bio resources in the region.

The DNA barcode database provides a system of species identification based on the finding of the closest match of the query sequences with database sequences. Of the total 113 *CO1* barcode sequences of ornamental fishes developed in this study, all of them showed closest match with the pre-existing sequences of 53 species in the database. Also these sequences clustered cohesively with the same species sequences and disctictly with other species in the NJ-cluster (Fig.1) with well supported bootstrap value ($\geq$99%). The maximum K2P distance within the species and the minimum K2P distances between the species were 0.045 and 0.056 respectively. Furthermore, the barcode gap, taken as the minimum interspecific distance [21,39] between the distinctly clustered congeners was straightforward in comparison to the cohesively clustered conspecies (Fig.2). As all the above facts were evident, among the 113 sequences, 53 species were identified.

Conclusions

Therefore, this study focuses the application of DNA barcoding in cataloguing the traded ornamental fishes from North East India. Furthermore, the study suggests enacting regulations that would recommend the use of proper zoological names along with DNA identification system of the species being traded. As anticipation, the survey highlighted the usefulness of DNA Barcoding technique in monitoring the trafficking of threatened species and demand conservation strategy for sustaining wildlife.



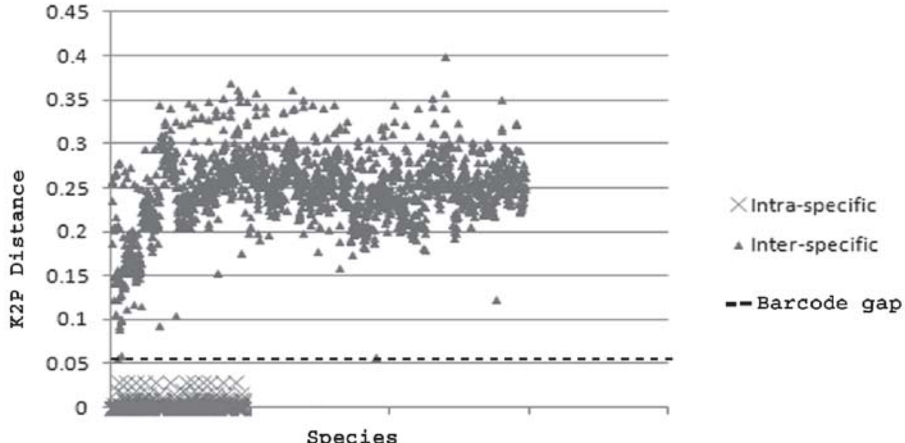


Fig. 2: The scatter plot of genetic distances within species against genetic distances to nearest-neighbor, where the red-cross marks showing intraspecific K2P distance and the blue dots showing interspecific K2P distance. A dotted line is draw at minimum interspecific K2P distance, the barcode gap, taken as the minimum interspecific distance between the distinctly clustered congeners was straightforward in delineating all the species taking in the study.

Fig. 1: Summary of Neighbour-Joining tree of CO1 gene sequences derived from ornamental fish species under trade using K2P distance paramenter. The numbers at the nodes are the bootstrap values based on 1000 replicates. The Genbank accession number and the species names are shown for each of the taxa.

Acknowledgements

We sincerely acknowledge Department of Biotechnology, Government of India for providing infrastructure facilities (BT/HRD/01/002/2007) and UGC-BSR fellowship to B. Dhar. We are thankful to Dr. Amalesh Choudhury, S.D. Marine Biological Research Institute, West Bengal for collection and identification of fish specimens, Bijoya Khondram, Department of Biotechnology, Assam University for providing samples from Manipur, Mohua Chakraborty, Department of Biotechnology, Assam University for collection of samples from Southern Assam.

Reference

[1] FAO Fisheries and Aquaculture Dept. Fishery and Aquaculture Information and Statistics Service. 2008. FAO yearbook. Fishery and aquaculture statistics = FAO annuaire. Statistiques des peches et de l'aquaculture = FAO anuario. Estadìsticas de pesca y acuicultura. Rome: FAO, p.v.

[2] Kalita, T., and Deka, K. 2013. Ornamental fish conservation in the flood plain wetlands of lower Brahmaputra Basin. Advances in Applied Science Research. 4:99-106.

[3] Das, S.P., and Biswas, J.N. 2009. A Handbook of Ornamental Fishes of the Brahmaputra Basin. Eastern Book House.

[4] Ponniah, A.G., and Sarkar, U.K. 2000. Fish biodiversity of North East India. NBFGR's NATP special publication no.2. National Bureau of Fish Genetic Resources:1-347.

[5] Hebert, P.D., Cywinska, A., Ball, S.L., and deWaard, J.R. 2003. Biological identifications through DNA barcodes. Proc. Biol. Sci. 270:313-321.

[6] Hajibabaei, M., Singer, G.A., Clare, E.L., and Hebert, P.D. 2007. Design and applicability of DNA arrays and DNA barcodes in biodiversity monitoring. BMC Biol. 5:24.

[7] Hajibabaei, M., Singer, G.A., Hebert, P.D., and Hickey, D.A. 2007. DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. Trends Genet. 23:167-172.

[8] April, J., Mayden, R.L., Hanner, R.H., and Bernatchez, L. 2011. Genetic calibration of species diversity among North America's freshwater fishes. Proc. Natl. Acad. Sci. U S A. 108: 10602-10607.

[9] Pramual, P., Wongpakam, K., and Adler, P.H. 2011. Cryptic biodiversity and phylogenetic relationships revealed by DNA barcoding of Oriental black flies in the subgenus Gomphostilbia (Diptera: Simuliidae). Genome. 54:1-9.

[10] Hausmann, A., Haszprunar, G., and Hebert, P.D. 2011. DNA barcoding the geometrid fauna of Bavaria (Lepidoptera): successes, surprises, and questions. PLoS One. 6:e17134.

[11] Rebijith, K.B., Asokan, R., Kumar, N.K., Srikumar, K.K., Ramamurthy, V.V., and Bhat, P.S. 2012. DNA barcoding and development of species-specific markers for the identification of tea mosquito bugs (Miridae: Heteroptera) in India. Environ. Entomol. 41:1239-1245.

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

[13] Hubert, N., Delrieu-Trottin, E., Irisson, J.O., Meyer, C., and Planes, S. 2010. Identifying coral reef fish larvae through DNA barcoding: a test case with the families Acanthuridae and Holocentridae. Mol. Phylogenet. Evol. 55:1195-1203.

[14] Chen, S., Yao, H., Han, J., Liu, C., Song, J., Shi, L., Zhu, Y., Ma, X., Gao, T., Pang, X., Luo, K., Li, Y., Li, X., Jia, X., Lin, Y., and Leon, C. 2010. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. PLoS One. 5:e8613.

[15] Smith, M.A., Eveleigh, E.S., McCann, K.S., Merilo, M.T., McCarthy, P.C., and Van Rooyen, K.I. 2011. Barcoding a quantified food web: crypsis, concepts, ecology and hypotheses. PLoS One. 6:e14424.

[16] Smith, M.A., Woodley, N.E., Janzen, D.H., Hallwachs, W., and Hebert, P.D. 2006. DNA barcodes reveal cryptic host-specificity within the presumed polyphagous members of a genus of parasitoid flies (Diptera: Tachinidae). Proc. Natl. Acad. Sci. U S A. 103:3657-3662.

[17] Valdez-Moreno, M., Ivanova, N.V., Elias-Gutierrez, M., Contreras-Balderas, S., and Hebert, P.D. 2009. Probing diversity in freshwater fishes from Mexico and Guatemala with DNA barcodes. J. Fish Biol. 74:377-402.

[18] Ward, R.D., Hanner, R., and Hebert, P.D. 2009. The campaign to DNA barcode all fishes, FISH-BOL. J Fish Biol. 74:329-356.

[19] Ward, R.D., Zemlak, T.S., Innes, B.H., Last, P.R., and Hebert, P.D. 2005. DNA barcoding Australia's fish species. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 360:1847-1857.

[20] Wong, L.L., Peatman, E., Lu, J., Kucuktas, H., He, S., Zhou, C., Na-nakorn, U., and Liu, Z. 2011. DNA barcoding of catfish: species authentication and phylogenetic assessment. PLoS One. 6:e17812.

[21] Bhattacharjee, M.J., Laskar, B.A., Dhar, B., and Ghosh, S.K. 2012. Identification and re-evaluation of freshwater catfishes through DNA barcoding. PLoS One. 7:e49950.

[22] Zemlak, T.S., Ward, R.D., Connell, A.D., Holmes, B.H., and Hebert, P.D. 2009. DNA barcoding reveals overlooked marine fishes. Mol. Ecol. Resour. 9 Suppl s1:237-242.

[23] Zhang, J. 2011. Species identification of marine fishes in china with DNA barcoding. Evid Based Complement. Alternat. Med. 2011:978253.

[24] Zhang, J., and Hanner, R. 2012. Molecular approach to the identification of fish in the South China Sea. PLoS One. 7:e30621.

[25] Laskar, B.A., Bhattacharjee, M.J., Dhar, B., Mahadani, P., Kundu, S., and Ghosh, S.K. 2013. The species dilemma of Northeast Indian Mahseer (Actinopterygii: Cyprinidae): DNA barcoding in clarifying the riddle. PLoS One. 8:e53704.

[26] Stoeckle, M.Y., Gamble, C.C., Kirpekar, R., Young, G., Ahmed, S., and Little, D.P. 2011. Commercial teas highlight plant DNA barcode identification successes and obstacles. Sci. Rep. 1:42.

[27] Lowenstein, J.H., Amato, G., and Kolokotronis, S.O. 2009. The real maccoyii: identifying tuna sushi with DNA barcodes—contrasting characteristic attributes and genetic distances. PLoS One. 4:e7866.

[28] Cutarelli, A., Amoroso, M.G., De Roma, A., Girardi, S., Galiero, G., Guarino, A., and Corrado, F. 2014. Italian market fish species identification and commercial frauds revealing by DNA sequencing. Food Cont. 37:46-50.

[29] Little, D.P., and Jeanson, M.L. 2013. DNA barcode authentication of saw palmetto herbal dietary supplements. Sci. Rep. 3:3518.

[30] Luo, J.Y., Yan, D., Song, J.Y., Zhang, D., Xing, X.Y., Han, Y.M., Yang, M.H., Dong, X.P., Peng, C., Chen, S.L., and Xiao, X.H. 2013. A strategy for trade monitoring and substitution of the organs of threatened animals. Sci. Rep. 3:3108.

[31] Sambrook, J., and Russell, D.W. 2006. The condensed protocols from Molecular cloning : a laboratory manual. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.

[32] Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. 1990. Basic local alignment search tool. J. Mol. Biol. 215:403-410.

[33] Zhang, D.X., and Hewitt, G.M. 1996. Nuclear integrations: challenges for mitochondrial DNA markers. Trends Ecol Evol. 11:247-251.

[34] Ratnasingham, S., and Hebert, P.D. 2007. bold: The Barcode of Life Data System (http://www.barcodinglife.org). Mol. Ecol. Notes. 7:355-364.

[35] Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28:2731-2739.

[36] Allen, D.J., Molur, S., Daniel, B.A., IUCN—The World Conservation Union., IUCN Species Survival Commission., and Zoo Outreach Organisation. 2010. The status and distribution of freshwater biodiversity in the eastern Himalaya. Gland, Switzerland: IUCN.

[37] Raghavan, R., Dahanukar, N., Tlusty, M.F., Rhyne, A.L., Krishna Kumar, K., Molur, S., and Rosser, A.M. 2013. Uncovering an obscure trade: Threatened freshwater fishes and the aquarium pet markets. Biological Conservation. 164:158-169.

[38] Murray, J.M., Watson, G.J., Giangrande, A., Licciano, M., and Bentley, M.G. 2012. Managing the marine aquarium trade: revealing the data gaps using ornamental polychaetes. PLoS One. 7:e29543.

[39] Meier, R., Zhang, G., and Ali, F. 2008. The use of mean instead of smallest interspecific distances exaggerates the size of the “barcoding gap” and leads to misidentification. Syst. Biol. 57:809-813.

IOSR Journal of Agriculture and Veterinary Science (IOSR-JAVS)
e-ISSN: 2319-2380, p-ISSN: 2319-2372.Volume 7, Issue 1, Ver. IV (Feb. 2014), PP 85-91
www.iosrjournals.org

Jiribam, the Ornamental Fishes' Hot Spot Zone Of Manipur, India

Bijoya Khomdram, Bishal Dhar and Sankar Kumar Ghosh*
Department of Biotechnology, Assam University, Silchar-788011, Assam, India

Abstract: *An investigation on the ornamental fish species availability was carried out in the Jiribam sub division, Imphal east district, Manipur, India. All the possible areas were surveyed and many experts were interacted. Out of the total 139 ornamental fishes found in the state of Manipur, 61 were recorded from Jiribam alone which comprise ~44 %. Therefore, we can categorize Jiribam as one of the hot spots of ornamental fishes in Manipur. The total 61 species belonged to 22 families and 7 orders. 42 species were recorded as threatened species and 3 species are endemic.*
Keywords: *Jiribam, ornamental fish, hot-spot, endemic, threatened*

I. Introduction

Due to the diversity of topographic and climatic features of NE India, this region is rich in endemic fish. Most of the small food fish which are treated as unwanted for conventional farming have good potency as ornamental fishes and are popularly known as Aquarium fishes. These species are attracting hobbyists both locally and globally.

North East India is considered as one of the hot spots of freshwater fish biodiversity in the world (Kottelat and Whitten, 1996). It is a well-recognized fact that there has been drastic reduction in abundance of the fresh water fishes in this region due to destruction of the habitat, overexploitation and other anthropogenic effects. Review of literature indicates that only limited information is available on fish germplasm resources of north east India with special reference to its potential as cultivable, sport and ornamental fishes. There has been a wide variation in the number of fishes reported from this region ranging from 172 (Ghosh and Lipton, 1982) to 267 (Sen, 2000). Also detailed drainage wise distribution, seasonal abundance, endemism and preferred microhabitats have not been critically examined.

The up to date inventory of the fish species of North East India showed 250 potential ornamental fish species. Out of this, the highest no. recorded from Assam (187), followed by Arunachal Pradesh (165), Meghalaya (159), Manipur (139), Tripura(103), Nagaland(71), Mizoram(46), and Sikkim(29). Conservation status of native ornamental fishes have shown that out of 250 sp., 10 are critically endangered, 28 are endangered, 49 are vulnerable, 45 are lower risk near threatened, 8 are lower risk least concern, 3 are data deficient and 107 are not evaluated (Ponniah, A. G. *et al.,* 2006). NE harbors diversified native ornamental fish species. These include both classified and non-classified types of aquarium fishes (Mahapatra *et al*., 2004). Small fishes like *Botia derio*, *Danio dangila*, *Puntius shalynius* etc. are classified types on the other hand larger food fishes like *Labeo gonius*, *Rita rita* etc. are termed as non classified ones. There are two major river basins within the state of Manipur, viz. the Barak River Basin and the Manipur River Basin. After the world famous Shiroi Lily and the Sangai, matter has now come to light that Manipur is also home to a number of fish species which are highly prized for their ornamental values in Europe and the United States.

Diversity of fish fauna in Jiribam, Manipur

The north east region shares its fish fauna predominantly with that of the Indo Gangetic fauna and to a little extent with the Burmese and South China fish fauna (Yadav and Chandra, 1994). Exploring the literature shows that 172 fish species with reference to their economic importance were recorded by Ghosh and Lipton (1982) while Sen (1985) and Mahanta *et al.* (1998) recorded altogether 187 fish species from Assam and the neighboring north eastern states of India. Compilation of Yadav and Chandra (1994) listed a total of 129 species. Sinha (1996) in his comprehensive review gave a list of 230 species of fishes as available from north eastern region. Sen (2000) has indicated that more number of species (267) has been reported from north east India. The various reports show a wide variation in the total number of fishes reported. Since Manipur is located in the extreme east zone of India therefore there is a greater chance in the available fish fauna being influenced by Burmese fish fauna. According to recent reports, a total of 139 ornamental fishes are found in Manipur. Jiribam is a small sub division in the westernmost part of Manipur where it borders with the state of Assam. It is drained by a single river, Jiri River and many small canals. The Jiri River joins the Barak River in Tipaimukh and hence becomes the main reason of harboring a large number and variety of fishes. The literature reveals that no efforts have been made to explore the rich ichthyodiversity of this region and also no work have been done to

assess the rich fish resources available in this region with respect to commercial utilization. With the growing demand for consumption and aquarium, it is necessary to evaluate potential species on the basis of different criteria.

In the present paper, an attempt has been made to prioritize among the fishes of this region the potential ornamental fishes along with their endemic status and status of threat on the basis of available literature as well as enquiry and interaction with the local fishery experts. Potential aquarium fishes have been identified based on actual present demand, bright coloration, uncommon look and uniqueness and following the records as mentioned in the literature. Recent estimates suggested that worldwide 20% of all freshwater fish species are extinct, endangered or vulnerable (Maclean and Jones, 1995). As a result fish stocks particularly those dwelling in inland open water areas, have gradually become endangered.

Extensive field survey conducted from September 2009 to December 2010 in Jiribam sub division of Manipur revealed the occurrence of bewildering diversity of ornamental fishes. Study about the species availability helps to know the present status of species variety and their relative abundance in the respective water bodies.

II. Materials And Methods

Study area

The survey work was carried out in Jiribam sub division (Imphal east district) of Manipur, India. Data were collected from all the major fish landing centers and interaction with the fishermen and local people.

Data collection and analysis

In order to collect data field visit was made every month and sometimes daily during the study period according to information and preference in the respective areas. In addition relevant information was also collected from various sources. The data were assembled through field survey using appropriate questionnaire. The questionnaire form was filled in by interviewing the fishermen directly from the field and local fish experts and also the local people. All the collected data were analyzed and the species observed were grouped in different categories.

III. Results And Discussion

Species composition

A total of 61 species were found from the surveyed area out of the total 139 species of ornamental fishes found in the state of Manipur. All the species were Freshwater fishes. They were belonged to 22 families and 7 orders. They are serially depicted in the tables 1, 2 and 3. Out of the 61 species recorded 21 species belonged to the family, Cyprinidae, only 1 sp. belonged to Anabantidae, 2 species belonged to Anguillidae, 1 species belonged to Badidae, 2 species from Balitoridae, 3 species from Ambassidae, 3 species from Channidae, 1 species from Clariidae, 3 species from Osphronemidae, 1 species from Erethistidae, 1 species from Schilbeidae, 3 species from Sisoridae, 4 species from Cobtidae, 5 species from Bagiridae, 1 species from Nandidae, 2 species from Notopteridae, 1 species from Mastacembelidae, 1 species from Chachidae , 1 species from Synbranchidae and 2 species from siluridae, 1 species from Heteropneustidae and 1 from the family Belonidae. Among the 7 orders of fishes found, the order Cypriniformes dominated others with a total number of 27 species then comes Siluriformes (15) and Perciformes ranked third with a total number of 12 species.

Species variation in different season

During the survey it was observed that not all species were available in all season. A total of 14 species were more available in winter season, 12 in summer and 35 fish species were available throughout the year. There were some species which were more available in summer but not in winter. And some species were available only in winter season. The species which have high ornamental value are mostly available during the months of October to December.

Status of the fishes

According to the IUCN (2008) Red List of all life forms, 16,928 species are threatened globally, and of these 1275 specie are fishes. There are 9 categories in the IUCN Red List namely, Ex- Extinct, EW- Extinct in the Wild, CR- Critically Endangered, EN- Endangered, VU- Vulnerable, LR/cd- Lower risk/ conservation dependent, NT- Near Threatened (includes LR/nt- Lower Risk/ near threatened), DD- Data deficient, LC-Least Concern (includes LC/lc- Lower Risk/ least concern). Species may move between categories for a variety of reasons, including genuine improvement or deterioration in status, new information being available about the species that was not known at the time of previous assessment, taxonomic changes, or mistakes being made in previous assessment (eg., incorrect information used previously, misapplication of the IUCN Red List criteria, etc,). Out of the 61 species of ornamental fishes recorded from Jiribam, 42 species are found to be in the list of

threats according to the report of CAMP workshop on freshwater fishes of India organized by NBFGR, 1997 and also following the records in the recent NBFGR publication (Lakra, W. S. *et al.,* 2010), after exercising all the related records and publications. These fishes are shown in Table 4 with their category of threat. And we observed 3 endemic fishes in jiribam namely, *Devario acuticephala, Schistura manipurensis* and *Garra manipurensis.* Except these 3 species, others are native.

In the present study, the percentage contribution of Cyprinids are found to be 33.3% being the dominant family. Some fishes are found to be surprisingly in the verge of extinction. The noted ornamental fish *Chaca chaca* is reported to be found in some restricted area of Jiribam but due to some reasons we are facing a great problem in getting this fish. The major reason behind the threatening status of this particular fish is the construction of Tipaimukh dam in the water body which is the sole and native home for this very species of ornamental fish. *Bagarius bagarius* is almost extinct in Jiribam nowadays, during the whole survey period it was recorded to catch only once in the month of December. During the last few years many exotic fishes are also introduced by the fishermen so that they could get a greater profit in their business without a second thought of affecting the local and indigenous fish diversity of the area. Freshwater fish are not only the most diverse group of vertebrates but they also represent and feature the greatest proportion of threatened species (Bruton, 1995; Leidy and Moyle, 1998; Duncan and Lockword, 2001). The principal threats to freshwater fish are the deterioration or destruction of habitats, both by pollution and intense modifications (like damming, channelization and so on.) and introduction of exotic species (Moyle,1986; Allan and Flecker, 1993). Though most of the fish resources of Manipur are already explored by Vishwanath and his team (Vishwanath and Sarojnalini, 1986, Vishwanath *et al.,* 2007, Vishwanath and Dishma, 2012), the fish species of Jiribam region remain untouched. The present study will help future researchers and others for easy access of the ornamental fishes of this region and will be of great help to conservationist and aquarists.

Acknowledgement:

We would like to thank to the Department of Biotechnology (DBT), Government of India, for providing the infrastructural fund. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. No additional external funding received for this study.

References

[1]. Allan, J.D. and Flecker, A.S. (1993). Biodiversity conservation in running waters. Bioscience. 43: 32-43.
[2]. Bruton, M.N. (1995). Have fish had their chips? The dilemma of threatened fishes. Environmental Biology of Fishes. 43: 1-27.
[3]. Duncan, J.R. and Lockwood, J.L. (2001). Extinction in a field of bullets: a search for causes in the decline of the world's freshwater fishes. Biological Conservation. 102: 97-105.
[4]. Ghosh, S.K. and Lipton, A.P. (1982). Ichthyofauna of the NEH Region with Special Reference to Their Economic Importance. ICAR Spl. Bulletin No.1, 119-126.
[5]. Kottelat, M. and Whitten, T. (1996). Freshwater Biodiversity in Asia with special reference to fish. World Bank Tech. Paper No. 343. The World Bank, Washington DC. 17-22.
[6]. Lakra, W.S., Sarkar, U.K., Gopalakrishnan, A. and Kathivelpandian, A. (2010). Threatened freshwater fishes of India. NBFGR publication. Lucknow.
[7]. Leidy, R.A. and Moyle, P.B. (1998). Conservation status of the world's freshwater fish fauna: an overview. Conservation Biology: For the Coming Decade, 2nd edition. Fielder, P.L. Karieva, P.M. (eds). Chapman and Hall: New York. 187-227.
[8]. Maclean, R.H. and Jones, R.W. (1995). Aquatic Biodiversity Conservation: a review of current issues and efforts. Ottawa. ON. SIFR.
[9]. Mahapatra, B.K., Vinod, K. and Manda,l B.K. (2004). Ornamental fish of North Eastern
[10]. India – Its distribution and conservation status. *Environment Ecol.* 22(3): 674-683.
[11]. Moyle, P.B. (1986). Fish introductions into North America: patterns and ecological impact. Ecology of Biological Invasions of North America and Hawaii. Springer –Verlag: Berlin. 27-43.
[12]. Ponniah, A.G. and Sarkar, U.K. (2006). Evaluation of North East Indian Fishes for their Potential as Cultivable, Sport and Ornamental Fishes along with their Conservation and Endemic Status. NBFGR, Lucknow.
[13]. Sen, N. (2000). Occurrence, distribution and status of diversified fish fauna of North East India. Occasional paper. Zoological survey of India.
[14]. Sen, T.K. (1985). The fish fauna of Assam and the neighbouring north eastern states of India. Occasional paper No. 64, Records of Zoological Survey of India. 1-216.
[15]. Sinha, M. (1994a). Threatened Coldwater Fishes of North Eastern India. *Threatened Fishes of India.* Natcon Publication, No. 4, UP. 173-176.
[16]. Vishwanath, W. and Dishma, M. (2012). *Barilius profundus*, a new cyprinid fish (Teleostei: Cyprinidae) from Koladyne basin, India. *Journal of Threatened Taxa* 4(2): 2363–2369.
[17]. Vishwanath, W, Lakra, W.S. and Sarkar, U.K. (2007). *Fishes of North East India.* National Bureau of Fish Genetic Resources, Lucknow, UP, India, 264.
[18]. Viswanath, W. and Sarojnalini, Ch. (1986). A new cyprinid fish, *Garra manipurensis*, from Manipur, India. *Japanese Journal of Ichthyology.* 35(2): 124-126p.
[19]. Yadava, Y.S. and Chandra, R. (1994). Some threatened carps and catfishes of Brahmaputra river system. *Threatened Fishes of India.* Natcon Publication No.4. UP. 45-55.

Table 1:- List of the ornamental fishes being collected from Jiribam. (A-abundant,LA- less abundant, MA-moderately abundant, R-rare, VR-very rare)

Sl no	Local name	Scientific name	Ornamental value	Abundance
1	*Muka nga*	*Amblypharyngodon mola*	Medium	A
2	*Ngawa*	*Barillius ngawa*	High	R
3	*Ngawa phuri thungbi*	*Barillius bendelisis*	High	R
4	*Muka nga macha*	*Devario acuticephala*	High	A
5	*Ngasang*	*Esomus dbnricus*	High	A
6	*Ngasang macha*	*Rasbora rasbora*	High	MA
7	*Ngathi*	*Labeo calbasu*	Medium	R
8	*Khabak*	*Labeo gonius*	Medium	MA
9	*Phabou*	*Puntius manipurensis*	High	R
10	*Phabou*	*Puntius punctata*	Medium	MA
11	*Phabou*	*Puntius sarana*	High	MA
12	*Phabou*	*Puntius conchonius*	Medium	MA
13	*Phabou*	*Puntius vittatus*	High	MA
14	*Mirga*	*Cirhinus mrigala*	Medium	A
15	*Ngara*	*Tor tor*	High	VR
16	*Rou*	*Labeo rohita*	Low	A
17	*Ukabi*	*Anabus testudineus*	Medium	MA
18	*Ngaril*	*Anguilla bengalensis*	High	R
19	*Ngaril leisna*	*Anguilla bengalensis Gray*	High	VR
20	*Napet nga*	*Badis badis*	High	MA
21	*Ngatup*	*Schistura sp.*	Medium	R
22	*Ngatup*	*Schistura manipurensis*	High	VR
23	*Ngamhai akoiba*	*Chanda nama*	High	MA
24	*Ngamhai asangba*	*Chanda nama*	High	MA
25	*Ngamhai anganba*	*Chanda nama*	High	MA
26	*Porom*	*Channa marulia*	High	MA
27	*Ngamu*	*Channa punctata*	Medium	A
28	*Meitei ngamu*	*Channa orientalis*	High	MA
29	*Ngakra*	*Clarias batrachus*	High	A
30	*Ngabemma*	*Colisa chuna*	High	MA
31	*Ngabemma*	*Colisa fasciata*	High	MA
32	*Ngabemma*	*Colisa lalia*	High	MA
33	*Samu khongpak*	*Erithistes hara*	High	VR
34	*Basa*	*Eutropichthys vacha*	Medium	MA
35	*Leingoi chabi*	*Gogangra viridescens*	High	R
36	*Hangoi nga*	*Sisor rabdophorus*	High	VR
37	*Ngakijou angangba*	*Lepidocephalichthys anandalei*	High	MA
38	*Ngakijou amuba*	*Lepidocephalichthys guntea*	High	MA
39	*Ngakijou awaoba*	*Lepidocephalichthys berdmorei*	Medium	MA
40	*Sarengkhoibi*	*Botia derio*	High	MA
41	*Ngasep*	*Mystus tengara*	Medium	A
42	*Ngasep*	*Mystus vittatus*	Medium	MA
43	*Ngachou*	*Sperata singhala*	Medium	R
44	*Ngarel*	*Bagarius bagarius*	High	VR
45	*Litha*	*Rita rita*	Medium	MA
46	*Kharaobi*	*Nandus nandus*	High	MA
47	*Kandla*	*Notopterus notopterus*	Medium	MA
48	*Ngapai*	*Notopterus chitala*	High	MA
49	*Ngamoi*	*Mastacembalus armatus*	High	MA
50	*Ngaprum*	*Monopterus cuchia*	Medium	MA
51	*Ngaseksha*	*Ompok pabda*	High	R
52	*Gajeb bakau*	*Chaca chaca*	High	VR
53	*Nunga amuba*	*Garra manipurensis*	High	VR
54	*Nunga awaoba*	*Schizothorax richardsoni*	High	R
55	*Ngaching*	*Gagata sp.*	High	R
56	*Nga cheklaobi*	*Xenentodon cancila*	High	MA
57	*Bata*	*Labeo bata*	Low	MA
58	*Mitlangbi*	*garra sp.*	Medium	R
59	*Ngamu sengum*	*Garra gotyla*	Medium	A
60	*Ngachik*	*Heteropneustes fossilis*	Medium	MA
61	*Sareng*	*Wallago attu*	Medium	MA

Table 2:- Families of the ornamental fishes found in Jiribam along with the number of species

Sl no.	Family	No. of species
1	Cyprinidae	21
2	Anabantidae	1
3	Anguillidae	2
4	Badidae	1
5	balitoridae	2
6	Ambassidae	3
7	Channidae	3
8	Clariidae	1
9	Osphronemidae	3
10	Erethistidae	1
11	Schilbeidae	1
12	Sisoridae	3
13	Cobtidae	4
14	Bagaridae	5
15	Nandidae	1
16	Notopteridae	2
17	Mastacembelidae	1
18	Chachidae	1
19	Synbranchidae	1
20	Siluridae	2
21	Heteropneustidae	1
22	Belonidae	1

Table 3:- Orders of the ornamental fishes found in Jiribam along with the number of species

Sl no.	Order	No. of species
1	Cypriniformes	27
2	Perciformes	12
3	Anguilliformes	2
4	Siluriformes	15
5	Osteoglossiformes	2
6	beloniformes	1
7	Synbranchiformes	2

Table 4:- Threatened species being detected in Jiribam sub division of Manipur based on NBFGR (National Bureau of Fish genetic Resources) data and report of CAMP (Conservation Assessment and Management Plan) workshop organized by NBFGR (EN- Endangered, VU- Vulnerable, CR-Critically Endangered, LRnt- Lower Risk near threatened, LRlc- Lower Risk least concern).

Sl no.	Species name	category
1	*Puntius manipurensis*	EN
2	*Sisor rabdophorus*	EN
3	*Tor tor*	EN
4	*Badis badis*	VU
5	*bagarius bagarius*	VU
6	*Botia derio*	VU
7	*Eutropiichthys vacha*	VU
8	*garra gotyla*	VU
9	*Heteropneustes fossilis*	VU
10	*Ompok pabda*	VU
11	*Puntius sarana*	VU
12	*Puntius vittatus*	VU

13	*Schizothorax richardsoni*	VU
14	*Notopterus chitala*	EN
15	*Amblypharyngodon mola*	LRlc
16	*Anabus testudineus*	VU
17	*Anguilla bengalensis Gray*	EN
18	*Channa marulia*	LRnt
19	*Channa orientalis*	VU
20	*Channa punctata*	LRnt
21	*Cirhinus mrigala*	LRnt
22	*Clarias batrachus*	VU
23	*Colisa fasciata*	LRnt
24	*Esomus danricus*	LRlc
25	*Garra manipurensis*	CR
26	*Rita rita*	LRnt
27	*Schistura manipurensis*	VU
28	*Wallago attu*	LRnt
29	*Xenentodon cancila*	LRnt
30	*Labeo bata*	LRnt
31	*Labeo calbasu*	LRnt
32	*Labeo rohita*	LRnt
33	*Lepidocephalus anandalei*	LRnt
34	*Lepidocephalus berdmorei*	EN
35	*Monopterus cuchia*	LRnt
36	*Mystus vittatus*	VU
37	*Nandus nandus*	LRnt
38	*Gogangra viridescens*	LRnt
39	*Notopterus notopterus*	LRnt
40	*Puntius conchonius*	VU
41	*Puntius vittatus*	VU
42	*Barilius bendelisis*	LRnt

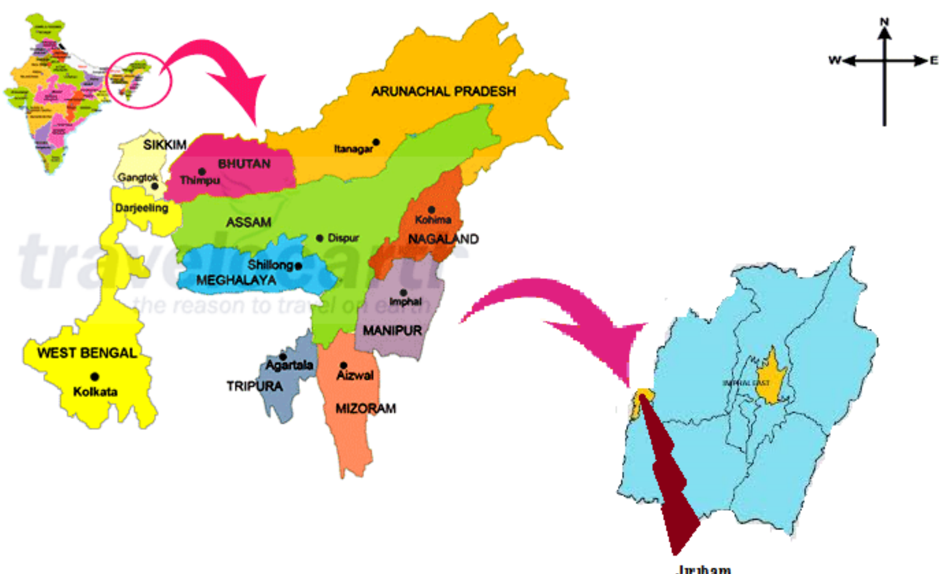


Figure 1. Map of India, Manipur, showing Jiribam, the study area (www.mapsofindia.com).

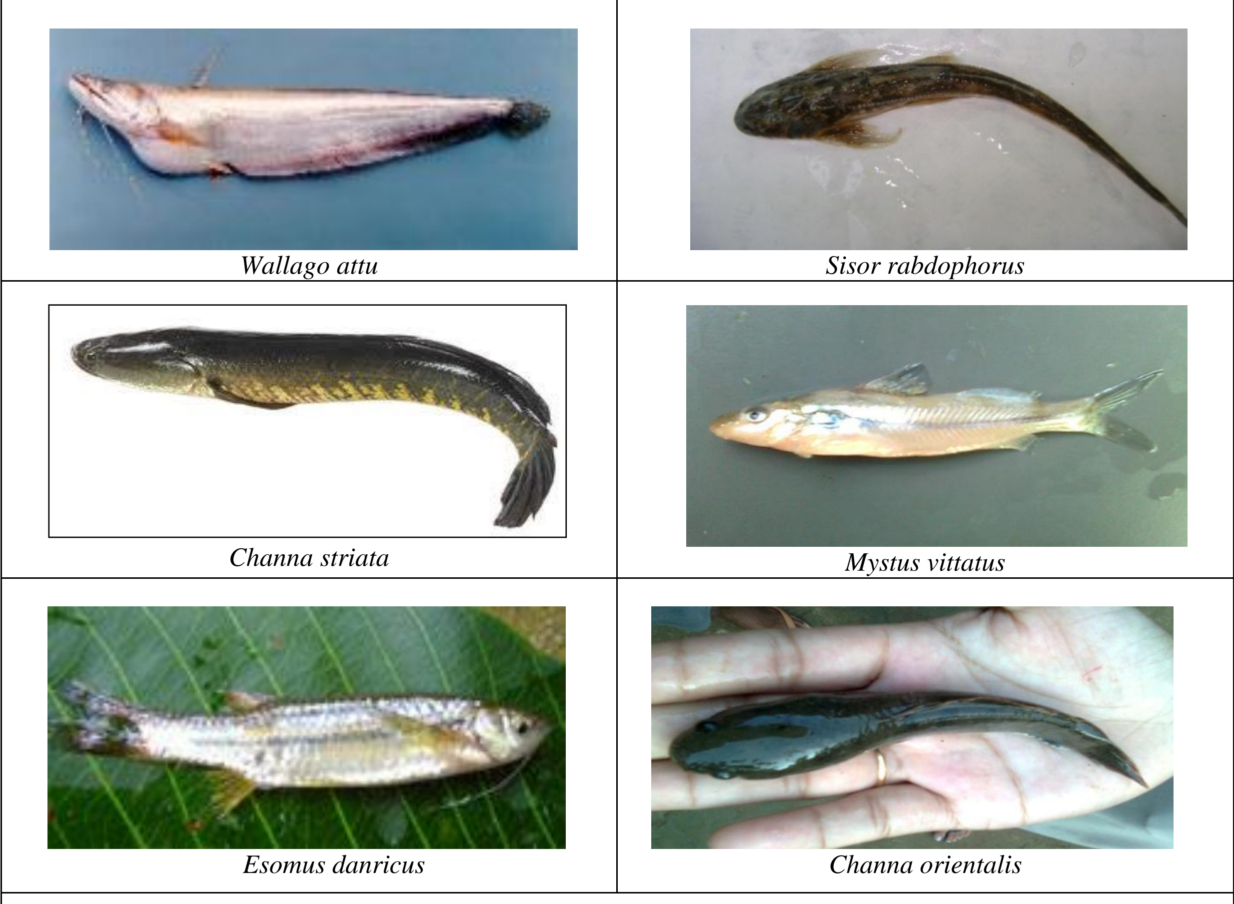


Figure 2. Some pictures of the ornamental fishes found in Jiribam (Photograph by Bijoya Khomdram)

The Species Dilemma of Northeast Indian Mahseer (Actinopterygii: Cyprinidae): DNA Barcoding in Clarifying the Riddle

Boni A. Laskar, Maloyjo J. Bhattacharjee, Bishal Dhar, Pradosh Mahadani, Shantanu Kundu, Sankar K. Ghosh*

Department of Biotechnology, Assam University, Silchar, Assam, India

Abstract

Background: The taxonomic validity of Northeast Indian endemic Mahseer species, *Tor progeneius* and *Neolissochilus hexastichus*, has been argued repeatedly. This is mainly due to disagreements in recognizing the species based on morphological characters. Consequently, both the species have been concealed for many decades. DNA barcoding has become a promising and an independent technique for accurate species level identification. Therefore, utilization of such technique in association with the traditional morphotaxonomic description can resolve the species dilemma of this important group of sport fishes.

Methodology/Principal Findings: Altogether, 28 mahseer specimens including paratypes were studied from different locations in Northeast India, and 24 morphometric characters were measured invariably. The Principal Component Analysis with morphometric data revealed five distinct groups of sample that were taxonomically categorized into 4 species, viz., *Tor putitora*, *T. progeneius*, *Neolissochilus hexagonolepis* and *N. hexastichus*. Analysis with a dataset of 76 DNA barcode sequences of different mahseer species exhibited that the queries of *T. putitora* and *N. hexagonolepis* clustered cohesively with the respective conspecific database sequences maintaining 0.8% maximum K2P divergence. The closest congeneric divergence was 3 times higher than the mean conspecific divergence and was considered as barcode gap. The maximum divergence among the samples of *T. progeneius* and *T. putitora* was 0.8% that was much below the barcode gap, indicating them being synonymous. The query sequences of *N. hexastichus* invariably formed a discrete and a congeneric clade with the database sequences and maintained the interspecific divergence that supported its distinct species status. Notably, *N. hexastichus* was encountered in a single site and seemed to be under threat.

Conclusion: This study substantiated the identification of *N. hexastichus* to be a true species, and tentatively regarded *T. progeneius* to be a synonym of *T. putitora*. It would guide the conservationists to initiate priority conservation of *N. hexastichus* and *T. putitora*.

Citation: Laskar BA, Bhattacharjee MJ, Dhar B, Mahadani P, Kundu S, et al. (2013) The Species Dilemma of Northeast Indian Mahseer (Actinopterygii: Cyprinidae): DNA Barcoding in Clarifying the Riddle. PLoS ONE 8(1): e53704. doi:10.1371/journal.pone.0053704

Editor: Vincent Laudet, Ecole Normale Supérieure de Lyon, France

Received August 16, 2012; **Accepted** December 3, 2012; **Published** January 16, 2013

Copyright: © 2013 Laskar et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: BAL has received a fellowship (DBT-RA) from the Department of Biotechnology, Government of India. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: drsankarghosh@gmail.com

Introduction

The term 'mahseer' refers to a group of freshwater cyprinid fishes easily distinguishable by relatively larger size of scales on their body compared to the other cyprinid fishes [1,2]. The members of mahseer belong to two genera, viz., *Tor* and *Neolissochilus*. These two genera are distinguished by the presence of a continuous labial groove in *Tor* but interrupted in *Neolissochilus,* and 10–14 gill rakers on the lower arm of first gill arch in the former and 6–9 in the latter [3,4]. They inhabit in the mountain streams and distributed in the range from Pakistan throughout Southern Asia to Southeast Asia up to the Malay Peninsula and the larger Indonesian islands across Sumatra, Borneo and Java [5,6]. However, species composition within each genus varies in different locations, like Southeast Asian species are different from Southern Asian species. Furthermore, within India, many species of mahseer are discontinuously distributed and mostly endemic in the South, Central and Northeast India. Among the mahseer of the Indian subcontinent, *Tor putitora* is widely distributed in Pakistan, India, Nepal and Bhutan; while *Neolissochilus hexagonolepis* is distributed in Nepal, Bhutan, North India and Northeast (NE) India [7,8]. A few studies suggest that the angling of mahseer provides superlative thrills than any other sport fishes except European Salmon [9,10]. They are highly sought-after because of great attraction to recreational anglers and are important components of the Angling-tourism pursuit [11]. In developing countries, there are many instances where the tourism industry has added recreational fishing to their attractions [12]. Owing to the growing value, the mahseer has become popular and considered as a cultural icon of diverse economic, recreation, and conservation

standpoint in rivers of eleven Asian nations [13]. Above all, the mahseer is an integral component of the aquatic ecosystem, serves as an important indicator of its health and supports the livelihood of many rural and indigenous ethnic groups in Asia [14]. However, the important mahseer fishes are threatened in the NE India as well as other distribution areas due to the growing harvest pressure as well as anthropogenic effects [15,16]. The two most threatened species, viz., *Tor putitora* and *Neolissochilus hexagonolepis* are regarded as the flagship species in NE India (http://www.nbfgr.res.in/). The conservation of mahseer has been hampered because the taxonomy of mahseer is most confusing due to the morphological variations they exhibit [17] that poised the understanding of actual species composition, distribution, autecology and biology at large.

Historically, with the pioneering work of Hamilton-Buchanan (1822) [1], many new descriptions of different species of mahseer have been proposed from Indian waters by distinguished naturalists. McClelland (1839) [18] recorded 4 new species from NE India, viz., *Tor progeneius, T. macrocephalus, Neolissochilus hexagonolepis* and *N. hexastichus*. McClelland, however, admitted difficulty in identifying Hamilton's *Cyprinus* (now *Tor*) *putitora* and particularly emphasized on a large cellular appendage to the apex of the lower jaw for *T. progeneius*, and the color gray on the back and reddish yellow on rest of the body for *N. hexastichus* [18]. The taxonomy of *T. progeneius* had long been in doubtful status, and it has been considered as a junior synonym of *T. putitora* [19]. Sen and Jayaram (1982) [20] characterized *T. progeneius* and elucidated with some new characteristics. Later, Rainboth (1985) [3] noted that *T. progeneius* is confusing to be classified whether within the genus *Neolissochilus* or *Tor*. It was further noted that most of the McClelland's type specimens were misplaced and some constituted curatorial nightmare [3]. Yet, McClelland's descriptions of two distinct species, viz., *Neolissochilus hexagonolepis* and *Tor progeneius* are recognized to be valid; while *T. macrocephalus* and *N. hexastichus* have been considered to be not valid rather the former was synonymized with *T. putitora* and the latter with *T. tor* [5,21].

Thus, the traditional taxonomy of mahseer in NE India has been facing several problems due to (1) lack of morphometric details in original description, (2) presence of very few holotypes of mahseer species, (3) indiscernible morphological nuances in them, and (4) disagreements in recognizing specific morphological characters. Consequently, the taxonomy of a few mahseer species has been extremely chaotic and described severally [2,4,5,20,21,22,23]. The mahseer species composition in the region is poorly understood and the identification of two species, viz., *T. progeneius* and *N. hexastichus*, has been difficult due to inconsistent taxonomic descriptions. Therefore, species level identification of mahseer is needed to be strengthened to facilitate the autecological study of mahseer and to develop conservation strategy for sustainable utilization in recreational fishing based tourism. Genomic approaches of taxon diagnosis have been found to be resourceful to aid traditional taxonomy [24,25]. In this context, the mitochondrial genome is a better target than nuclear genome because it evolved faster and can thus give more information to discriminate close species.Lately, a partial fragment of mitochondrial cytochrome oxidase c subunit I (COI) gene has been proposed to be sufficient singly to differentiate all, or at least the vast majority of animal species [26]. As such, this partial locus (COI) has been extensively tested for its efficacy in fish species identification and recognized as a unique marker of species identification with high confidence and called as ''DNA barcode'' [27,28,29]. The concept of DNA barcode based species identification is easy, rapid and accurate for being sequencing and web based; as such it has gained great attention worldwide [30,31,32]. Recently, the catfish diversity in NE India has been re-evaluated through DNA barcoding [33]. Therefore, morphological and DNA barcode data in combination can help to resolve the species dilemma of Northeast Indian mahseer, particularly *T. progeneius* and *N. hexastichus*, for effective conservation and management of the species.

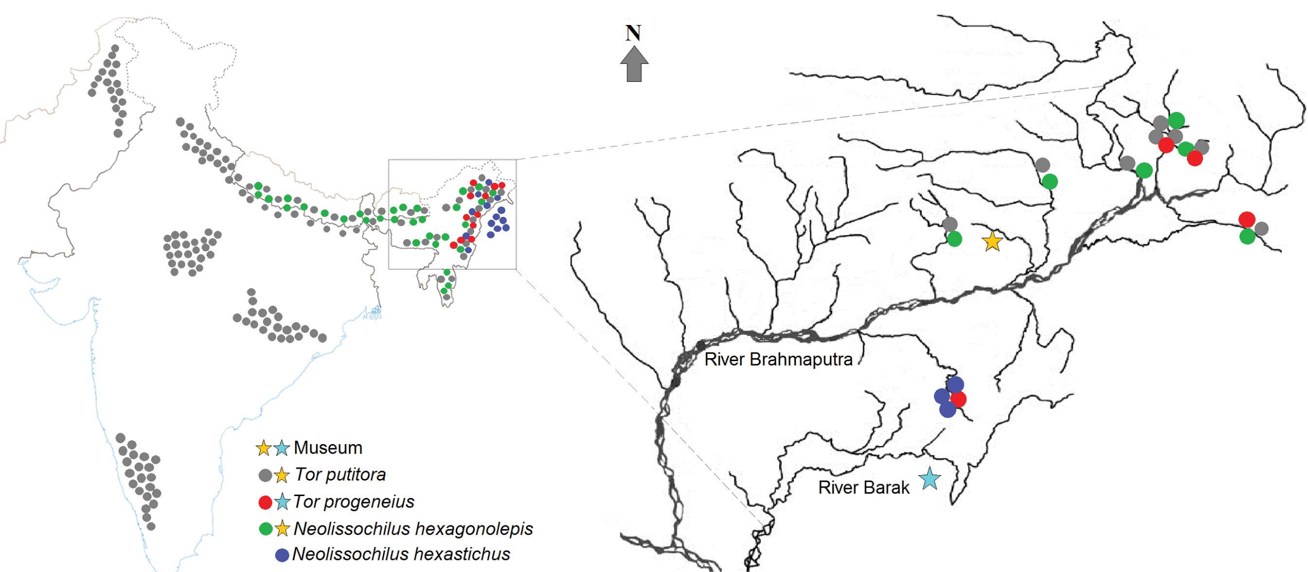


Figure 1. Map of the study site showing the known distribution of the studied species and the collection sites in different river drainages. The figure shows that the Northeastern region of India is drained mostly by River Brahmaputra and partly by River Barak. The studied specimens were collected from the drainages of River Brahmaputra. The topography of the region restricts the convergence of Southeast Asian fish composition with this region.
doi:10.1371/journal.pone.0053704.g001

Table 1. Morphological grouping of the studied organisms along with the corresponding codes.

Group	Nomenclature in practice	Sequence accession number used in molecular analysis/ catalogue number of paratypes in museum	Sample code used in morphological analysis (PCA)
N_2	*Neolissochilus hexastichus*	SGBL-BMF35	A
		JX127237	B
		JX127239	C
		JX127235	D
		JX127236	E
		JX127238	F
		SGBL-BMF36	*
N_1	*N. hexagonolepis*	JX127232	G
		JX127234	H
		JX127231	I
		JX127233	*
		** RGUMF-0036	V
		** RGUMF-0037	W
		** RGUMF-0038	X
T_2	*Tor progeneius*	JX127229	J
		***	K
		JX127228	L
		***	M
		JX127230	N
T_3	*T. putitora*	** RGUMF-0034	Aa
		JX127240	O
		JX127224	P
		JX127241	Q
		JX127242	U
T_1	*T. putitora*	JX127227	R
		JX127226	T
		JX127225	S
		** RGUMF-0035	Y
		***	Z
		***	Ab

The grouping was done based on scatter plot from Principal Component Analysis (PCA) as well as following the authoritative taxonomic keys. Sequence accession numbers in GenBank are used in the presentation of molecular analysis and the sample codes in PCA. Alphabetic sample codes replacing the full name of organisms are ascribed for ease of presentation those however clearly mentioned in Table S2.
*big specimen from market whose morphometric not done.
**paratypes from museum preserved in formaline whose sequencing not done.
***previously identified specimens preserved in formaline whose sequencing not done.
doi:10.1371/journal.pone.0053704.t001

Materials and Methods

Sample Collection

Fish specimens belonging to the group mahseer in the range of sub-adult to adult size were collected through participatory sampling with the marginal fishers engaged in commercial fishing. The specimens were from various locations in the hills and foothills across the Northeast India, particularly in the drainages of River Brahmaputra (Figure 1). The method of sample collection was approved by the Ministry of Science and Technology, Department of Biotechnology, Government of India (vide No. BT/HRD/01/002/2007). Some known voucher specimens within the genera *Tor* and *Neolissochilus* were examined from the Museum of Biodiversity in Rajiv Gandhi University, Arunachal Pradesh (voucher numbers are given in Table 1). The morphometrics of previously identified specimens from collection of *T. putitora* and *T. progeneius*, as well as the type specimens of *T. putitora* and *N. hexagonolepis* were included in the analysis. The type specimens of *T. progeneius* and *N. hexastichus* are not available in the museum. In lieu of examining type specimens of *T. progeneius* a small review on the existing contradictions among the taxonomists regarding the taxonomic descriptions and opinions on the status of the species is given in Supporting Information S1. Concerning the identification of *T. progeneius* and *N. hexastichus*, the original descriptions were emphasized. A total of 19 fresh specimens belonging to 4 species were studied in association with 5 paratypes and 4 previously collected specimens. Muscle tissue samples were invariably collected aseptically from behind of dorsal fin of the fresh specimens and taken in 500 µL of TES buffer (50 mM Tris HCl, 25 mM EDTA and 150 mM NaCl). The whole body

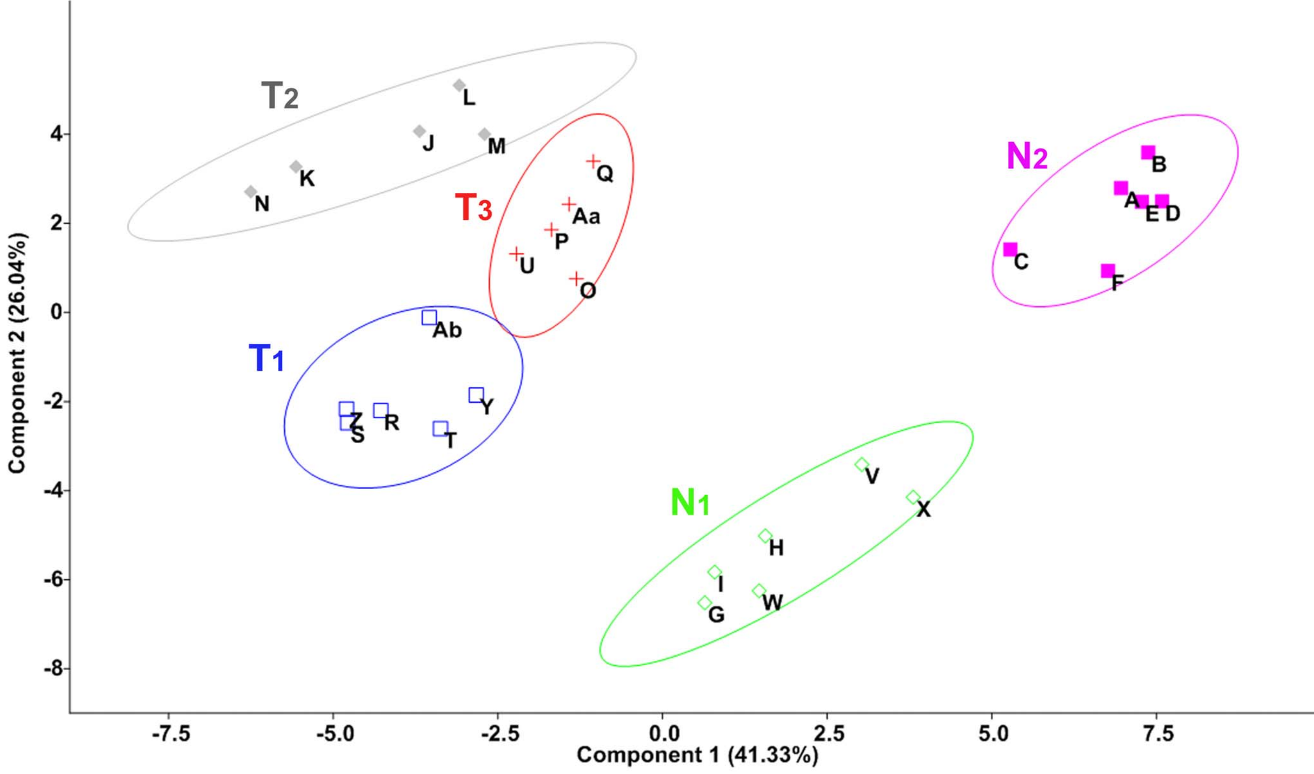


Figure 2. Principal Component Analysis (PCA) on 24 morphometric variables of the study samples including paratypes. The clusters of samples obtained from PCA were assigned to respective taxa based on meristic counts as well as non-quantitative characters of samples following authoritative taxonomic keys. The groups are like T_1, T_2 and T_3 comprising *Tor* congener, and N_1 and N_2 comprising *Neolissochilus* congener.
doi:10.1371/journal.pone.0053704.g002

specimens are preserved and stored at the Department of Biotechnology of Assam (Central) University, Silchar, Assam, India, for frequent examination and record of vouchers (vouchers' details are provided in Table S1).

Taxonomic Identification and Nomenclature

Specimens were categorized systematically based on the taxonomic characters available from the original description as well as subsequent re-descriptions and taxonomic reviews. Altogether 24 morphometric variables along with 6 important meristic counts were measured following standard literatures [23,34] (Figure S1) and the measurements were recorded using digital slide caliper (0.01 mm). The morphological characters those are non-quantitative yet taxonomically relevant, e.g. color pattern on the body and fins, presence or absence of tubercles, appearance and diagonal shape of mouth, etc. were also recorded from all the specimens. The measurements were taken at least three times independently and mode of each parameter was finally considered to minimize the error. The samples were designated into the respective species as per the authoritative taxonomic keys [4,20,23] and the species nomenclature was adopted as per the updated catalogue [8].

PCR Assay and Purification

DNA was extracted with standardized Phenol-Chloroform-Isoamyl alcohol method [35]. COI gene fragment (~655 bp) was amplified using the set of published primers: FishF1-5′TCAACCAACCACAAAGACATTGGCAC 3′ and FishR1-5′TAGACTTCTGGGTGGCCAAAGAATCA 3′[27]. The amplification was performed in 25 µl reaction mixture of 1X PCR buffer, 2 mM $MgCl_2$, 10 pmol of each primer, 0.25 mM of each dNTPs, 0.25 U high-fidelity polymerase and 100 ng of DNA template. PCR conditions were: initial denaturation at 94°C (2 minutes) followed by 30 cycles at 94°C (45 seconds), 50°C (45 seconds) and 72°C (1 minute), and a final elongation at 72°C (8 minutes). The PCR-amplified products were checked in 1% agarose gels containing ethidium bromide (10 mg/ml) and the single uniform band was then purified using QIAquick^{R} Gel extraction kit (QIAGEN, USA). The amplicons were bi-directionally sequenced in an automated DNA sequencer (ABI 3500, Applied Biosystems Inc., CA, USA).

Sequence Quality Control Measures

Two chromatograms that represent sequences of both the DNA strands were obtained for each sample. The PCR amplified products as well as their corresponding DNA sequences were larger than 600 bp that assured the sequences being not Numts as the limit of Numt hardly reaches 600 bp [36]. The noisy sequences were trimmed at both end and greater than 2% ambiguous bases were discarded, using quality value of >40 for bidirectional reads. BLASTN program was used to compare the sequences retrieved from the two chromatograms [37], and the fragment showing 100% alignment with no gap or indel (insertion/deletions) was selected. In some cases of discrepancy, both the sequences were reviewed and quality value of the sequences were considered to determine the most likely nucleotide using the software SeqScanner Version 1.0 (Applied Biosystems Inc., CA, USA). The selected fragments of the sequence were aligned using ClustalX software [38]. Finally, each of the sequences was compared in NCBI through BLASTN to examine

Table 2. Summary of PCA on 24 morphometric measurements of 28 samples within 4 species.

	PC 1	PC 2
% variance	41.336	26.04
Eigen value	19.4715	12.2662
Variable	Loadings	
SL	−0.3807	−0.1508
PrDL	−0.1599	0.2042
PoDL	0.1929	−0.1689
HtCF	0.298	0.1092
HL	−0.2624	0.3204
HtPF	0.1026	0.09338
HtDF	0.2399	0.3027
HtAF	0.1577	0.1663
HtDS	0.1129	0.1021
DP&V	−0.1125	−0.4824
LnCP	−0.1836	−0.1738
BDdf	0.4392	−0.1007
HDop	0.2464	−0.1186
HDe	0.1665	−0.03199
BWdf	0.1187	−0.1101
HWe	0.1814	0.00083
SnL	−0.1136	0.1563
ED	0.0623	0.07016
LnLF	−0.182	0.363
LHtCP	0.1185	−0.00832
HtVF	0.1842	0.07278
DVF&AF	−0.08024	−0.4201
LnBDF	0.2209	−0.09064
LnBAF	0.03401	−0.05084

Proportion of variance, Eigen values and coefficients (loadings) of the first two principal components (PC1 and PC2) for the % total length of the morphometrics of studied mahseer species.
doi:10.1371/journal.pone.0053704.t002

the complete alignment with the partial coding sequence of fish mitochondrial COI gene. The sequences were translated using the online software ORF finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) and aligned through BLASTP to examine whether the partial amino acid codes were coherent with the fish mitochondrial COI gene frame and without any stop codon. In this way, the generated sequences were confirmed to be the fragments of mitochondrial COI gene. All the analyzed sequences were then deposited in GenBank (details of accession numbers are given in Table S1). The sequences were also submitted in a FISH-BOL project entitled "DNA barcoding of Mahseer fishes from Northeast India" and the code name 'MFISH'.

Data Analysis

Morphometry. Principal Component Analysis (PCA), a multivariate statistical procedure commonly used to reveal patterns in measured correlated variables, was used to differentiate the samples into possible groups and any variation among the samples of same species and the paratypes. The morphometric measurements were transformed into percentage of the total body length to develop the relative data of each variable for the samples of different size and species. The analysis was performed using PAST version 2.17 b (http://folk.uio.no/ohammer/past). The PCA output is presented as scatter plot showing the groups of the samples with designated codes.

COI sequence data analysis. The sampled specimens were invariably sequenced and their congeneric sequences were acquired from the databases (GenBank and BOLD) to examine the level of intraspecific variation. Most of the database sequences lack geographical information yet they were assumed to be at least from distant locations. The analysis was based on a total data set of 76 COI barcode sequences of mahseer containing 21 *denovo* sequences and 55 database sequences. Additionally, 2 sequences of *Hypsibarbus wetmorei* and 3 sequences of *Puntius sarana* were acquired from GenBank to represent the out-group in the study. Geographical information and GenBank accession numbers of the developed as well as acquired sequences are given in Table S1. The calculation of Kimura 2-parameter (K2P) congeneric and conspecific distance [39] as well as phylogenetic analysis through Neighbor Joining (NJ) method were performed using MEGA Version 5.1 [40]. The tree topology obtained through NJ method was double-checked by Maximum Likelihood (using MEGA Version 5.1) and Bayesian approach (using MrBayes 3.2.0) [41].

Results

Morphological Characteristics

The PCA yielded 24 components which correspond to the 24 morphometric measurements. Projection of the morphometric data of studied mahseer species on first 2 principal axes showed the separation of the samples into 5 groups at 75% concentration ellipse level (Figure 2). The first 2 principal components contributed to 67.37% of total variance (PC1 = 41.33% and PC2 = 26.04%) (Table 2). The third, fourth and fifth components contributed to 8.57%, 4.93% and 3.55%, respectively, but did not improve the separation of the samples. These 5 groups were categorized into 2 broad groups and each corresponds to a genus, as per the authoritative taxonomic keys. The meristic count of the samples is presented in Table 3 which depicts a prominent difference in number of gill rakers on the lower arm of first arch between the two genera. The rakers were 8–9 in *Neolissochilus* and 13–14 in *Tor*. The other meristics were almost similar in all the samples. In the PCA scatter plot, the samples within the genus *Neolissochilus* further formed two distinct groups, one of which grouped with the paratypes of *N. hexagonolepis* but the other group stood distant indicating both the groups belonging to different species. The samples within the genus *Tor* appeared to be in a single but very stretched out group indicating a wide range of variation. In this group, some samples formed two slightly distant groups, yet each of the groups assembled with at least one of the paratypes of *T. putitora* while the rest few samples formed a slightly separate group and remained away from the paratypes. The non-quantitative characters of samples within *Tor* and the prevailing taxonomic descriptions suggested two possible species name. The groups of samples appeared in PCA were designated as T_1, T_2 and T_3 comprising *Tor* congener, and N_1 and N_2 comprising *Neolissochilus* congener. The constituent samples within each group were given the alphabetic sample code (Table 1), like S, R, T, Y, Z and Ab fall within T_1; J, K, L, M, and N fall within T_2; Aa, O, P, Q and U fall within T_3; G, H, I, V, W and X fall within N_1; and A, B, C, D, E and F fall within N_2. The meristic counts and morphometric data are given in Table 3 and supplementary Table S2 respectively.

Table 3. Important meristic counts of three specimens in each species.

Organism name (Species)	Replicates	Parameters					
		Gill rakers on first arch (upper arm+lower arm)	Scales on lateral line	Dorsal fin rays	Ventral fin rays	Pectoral fin rays	Anal fin rays
Tor putitora	a	2+14	25	9+ii	8+i	15+i	6+i
	b	2+14	25	9+ii	8+i	15+i	6+i
	c	2+14	25	9+ii	8+i	15+i	6+i
Tor progeneius	a	2+14	26	9+ii	8+i	15+i	6+i
	b	2+13	26	9+ii	8+i	15+i	6+ii
	c	2+14	26	9+ii	8+i	15+i	6+i
Neolissochilus hexagonolepis	a	2+8	27	9+ii	8+i	14+i	6+i
	b	2+8	27	9+ii	8+i	14+i	6+i
	c	2+8	27	9+ii	8+i	14+i	6+i
Neolissochilus hexastichus	a	2+9	24	10+ii	7+i	14+i	7+i
	b	2+9	24	10+ii	7+ii	14+i	7+i
	c	2+9	24	10+ii	8+i	14+i	7+i

The table shows that the number of gill rakers on the lower arm of first arch is a very important distinguishing character between the two genera. This character is very easily identifiable and based on this character the first hand classification of mahseer in to respective genera can be easily done.
• Lowercase roman numerals are used to denote the simple rays in fin ray count.
doi:10.1371/journal.pone.0053704.t003

Figure 3. Neighbor Joining (NJ) tree developed using K2P distance among 76 COI sequences of mahseer. In most cases the studied samples (marked as red dots) showed cohesive clustering with their conspecific database sequences. The samples morphologically identified as *Tor progeneius* (accession numbers JX127229, JX127230 and JX127228) clustered conspecific with developed sequences as well as database sequences of *T. putitora.* All samples of *N. hexastichus* clustered cohesive as the same species and distinct from *N. hexagonolepis* sequences. Few database sequences of 3 species revealed aberrant clustering {*Neolissochilus stracheyi* (accession number HM536922), *Tor mosal mahanadicus* (accession numbers HQ609722, GQ469780), *Tor macrolepis* (accession numbers GQ469827-29)}. ● The numbers at the nodes are bootstrap values based on 1000 replications. ● The specimens' GenBank accession number and species name are shown for each taxon. ● Red dots and black triangles correspond to the sequences developed in this study. Black triangles also correspond to the sequences of samples, although morphologically identified as *Tor progeneius* but were found conspecific with *Tor putitora* based on COI sequence data analysis. Black dots correspond to the cases of abnormal clustering.
doi:10.1371/journal.pone.0053704.g003

Tor Congener

The first taxonomic key to differentiate species within *Tor* is based on the relative head length to the body depth. In this study, all *Tor* congener possessed slightly longer head than body depth. There was no stringent variation in meristic counts among the *Tor* congener (Table 3); and both T_1 and T_2 samples were similar in most of the other taxonomic features (Table S2). But, T_3 differed from T_1 and T_2; firstly on having long (up to the margin of maxilla) mental lobe (also called lower labial flap) vs. short/absent, and secondly on having longer upper jaw and with skin like flap extending behind upper lip vs. both the jaws equal and upper lip without a flap. The mental lobe length was 4.29% to 5.91% of total length in T_3 samples vs. 1.47% to 1.99% in T_1 and T_2 samples. The longer upper jaw in T_3 samples correspondingly shared to greater head length and snout length than in T_1 and T_2 samples. It appeared that the presence of both upper and lower lips as being relatively more fleshy and the mental lobe being prominent and long in all the samples of T_3 differentiate them from T_1 and T_2 samples. The observed morphological features of all the samples within T_1 and T_2 bear close affinity with the described features of *Tor putitora*. Therefore, despite minor differences, T_1 and T_2 samples were considered to be belonging to the same species and named accordingly. The particular lip character in T_3 samples resemble with the original descriptions of *Tor progeneius*. It was observed that the three groups though bear minor variation in morphometrics but they are not discernible except the particular differentiating features of *Tor progeneius* that appeared to be unique and very much noticeable. Thus, T_3, T_1 and T_2 samples are tentative considered as morphs and the former is designated as long mental lobed while the latter 2 as short mental lobed.

Neolissochilus Congener

The N_1 samples were distinguished from N_2 due to absence of mental lobe vs. prominent, and interrupted groove behind the lower lip vs. continuous groove. Both these features of N_2 samples resembled with the *Tor* congener. But, the gill rakers in them were 9 vs. 13–14 in *Tor* congener. Tubercles were mostly present on the cheeks in N_1 but entirely absent in N_2. Mouth smoothly rounded in N_2 vs. truncate in N_1, edge of lower jaw blunts in N_2 vs. sharp in N_1. The color of the back, bases of caudal and dorsal as well as the upper part of the head in N_2 samples was greenish gray, reddish yellow on rest of the body, and the tips of the fins red. These observed features in N_2 samples have been originally emphasized

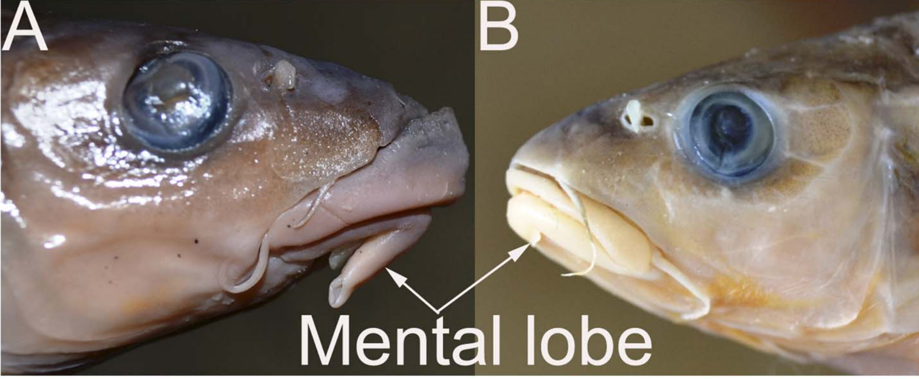


Figure 4. Illustrating the counter appearances of mouth in (A) *T. progeneius* and (B) *T. putitora*. Showing fleshy lips, a semicircular flap extending behind the upper lip, and a fleshy appendix extending from the lower lip up to the margin of maxilla (very long mental lobe) in (A), but absent in (B).
doi:10.1371/journal.pone.0053704.g004

to describe *N. hexastichus* as a distinct species. Thus, the N_2 samples were named as *N. hexastichus* according to the authoritative descriptions. The morphometrics of N_1 samples were mostly similar to both *N. hexagonolepis* and *N. stracheyi*. The N_1 samples were uniquely identified to be *N. hexagonolepis*, based on color pattern having scales coppery colored with a tinge of red above lateral line and fins deep slate paling towards their margins. As per the prevailing taxonomic description, *N. hexagonolepis* is different from *N. stracheyi* due to the absence of a lateral black stripe as in N_1 samples. Thus, following the taxonomic keys, the N_1 samples were named as *N. hexagonolepis*.

DNA Barcoding Analyses

The K2P divergence matrix of the dataset (as shown in Table S3) revealed that the congener of *Tor* maintained divergences in the range of 3.5% to 7.4% with the congener of *Neolissochilus*. The maximum K2P divergence among T_1 and T_2 samples was 0.2% and the comparison of both T_1 and T_2 samples with T_3 samples also revealed a maximum K2P divergence of 0.2%. The maximum divergence of all the samples belonging to T_1, T_2 and T_3 with the closest database sequences of *Tor putitora* was 0.8%. The divergence matrix suggested that all samples of T_1, T_2 and T_3 are conspecific of *T. putitora* in the absence of any database sequence of *T. progeneius*. Therefore, COI gene sequences of these 3 groups were submitted to both GenBank and BOLD under the putative species *T. putitora*. The maximum divergence within N_1 samples was 0.6% while their divergences with the conspecific database sequences were in the range of 0.4% to 0.8%. The divergence matrix suggested that N_1 samples are conspecific of *Neolissochilus hexagonolepis*. The within group divergences of N_2 samples were in the high range up to 0.9% possibly due to a particular sequence. Excluding the particular sequence (accession number JX127239), the within group divergence of N_2 samples remained nil in the absence of any conspecific sequence in the database.

The averages of conspecific and congeneric divergences were determined from the matrix to be 0.5% ±0.2% and 2.8% ±0.7% respectively. In the dataset, the minimum distance between the closest species (closest congener) was 1.5%. Therefore, the closest congeneric divergence among mahseer species was calculated to be 3 times higher than the mean conspecific divergence, which is called as the 'barcode gap'. Based on the barcode gap, T_1, T_2 and T_3 samples were found to be conspecific with *T. putitora*; N_1 samples were conspecific with *N. hexagonolepis*; and N_2 samples including JX127239 were discrete in the absence of any database sequence of *N. hexastichus*.

The NJ as well as Maximum Likelihood (ML) and Bayesian tree based cluster revealed that the congener of *Tor* and *Neolissochilus* formed two related clades while *Puntius sarana* and *Hypsibarbus wetmorei* remained as out-group (Figure 3, Figure S2, and Figure S3). This also revealed that *T. putitora*, *T. tor*, *T. khudree*, *T. sinensis*, *T. mussullah*, *T. mosal*, *T. malabaricus*, *T. douronensis*, *T. tambroides*, *N. hexagonolepis* and *N. stracheyi* clustered separately and are distinct species. All the samples of T_1, T_2 and T_3 clustered in the same clade and nearest to *T. putitora*; N_1 samples clustered with *N. hexagonolepis*; and the 6 N_2 samples clustered in the same clade while the other N_2 sample remained a bit distant. In addition, some database sequences reflected aberrant clustering like, 1) all sequences of *T. mosal mahanadicus* and *T. macrolepis* clustered with *T. putitora* and 2) a single sequence of *N. stracheyi* (accession number HM536922) clustered with *N. Hexagonolepis*.

Discussion

In this study, all the possible mahseer habitats across the Northeast India were surveyed. Altogether three morphologically distinct groups of mahseer within the genus *Tor* and two within *Neolissochilus* were identified from the study site. DNA barcoding analyses however recognized all the three groups belonging to a single species within the genus *Tor* and conspecific of *T. putitora*. The *T. putitora* is a widely distributed species and it has been reported to be exhibiting polymorphism in geographically isolated populations [42]. Among the study samples, T_3 samples possessed long fleshy appendage to the lower lip (mental lobe) while the others lack this feature. This feature corresponds to the original description of *T. progeneius* where this particular feature was specially emphasized for nomenclature [18,21]. This species had been also considered closely allied to *T. tor* in view of its lower lip character; consequently these two species have been synonymized very often [5,19]. *T. progeneius* was however differently described after its original description probably due to lack of original holotype [3] and non-availability of fresh specimens [21]. It was identified to be distinct from *T. tor* due to length of head almost equal to depth of the body in the former vs. length of head considerably shorter than depth of the body in the latter [20]. Subsequently, based on archival specimens (Zoological survey of India, Kolkata; specimens' catalogue details not mentioned), it was characterized to be having 8–10 rakers on the lower arm of first gill arch, tubercles on the cheek and lacking completely a mental

lobe. Based on such characters this species was remarked to be doubtful to place in either in *Tor* or *Neolissochilus* [3]. According to one proposition, there are two types among the yellow finned mahseer: i) the lips are fleshy and the lower one is produced backwards into a long fleshy appendage, and ii) the lips are of normal type and the lower lip does not form an appendage [21]. Based on such descriptions, Hamilton's *Cyprinus* (present *Tor*) *putitora* and *C. mosal* have been stated to be the same species and the nature of their lips was stated to be adaptive characters [5]. Besides, the description of a fan-shaped structure behind upper jaw in *T. progeneius* [21] was stated to be an abnormal formation based on archival specimens (Zoological Survey of India, Kolkata; specimens' catalogue details not mentioned) [5]. In contrary, Menon (1992) [5] described *T. progeneius* to be possessing of 27–31 numbers of lateral line scales on the body. It seems that Menon (1992) was so influenced by this feature of *T. progeneius* that he used it as a taxonomic key to species. Secondly, in contrary to all previous descriptions except Rainboth (1985) [3], Menon (1992) noted the presence of cheek tubercles in *T. progeneius*. On the other hand, according to original description as well as the prevailing adoption of taxonomic character for this species indicate that the number of scales on the lateral line was never more than 26, and the extension of singular appendage from lower lip has been largely emphasized. This species has long been remained unreported, that might be due to the above mentioned morpho-taxonomic perplexity arising from vague and varied presentation of its specific characters incongruent to the original description [18]. All the T_3 samples were observed to be possessing of maximum of 26 lateral line scales, 13–14 gill rakers, the slightly longer head than body depth and particularly the fleshy lips with long angular appendage to the lower jaw (long mental lobe) that is in contrast to the short mental lobe in both T_1 and T_2 samples (Figure 4). The different lower lip structure in T_3 samples could be an adaptive [5,21] or a sexually dimorphic feature [43]. Moreover, different geographical populations of *T. putitora* have been reported for significant Nuclear Organiser Region polymorphism [42] that indicates the possibility of the presence of a polymorphic form of this species in northeast India. Because, the collection site of T_3 samples in the drainages of river Brahmaputra is phylogeographically poorly connected with the other Himalayan streams such as Ganga. Therefore, notwithstanding such noticeable differences in mouth structure, following DNA barcoding results, we conclude that *T. progeneius* is a synonymous species of *T. putitora*. This study contributed 10 replica barcode sequences in GenBank of *T. putitora*. In elsewhere, DNA barcoding approach has been successful in describing different nominal species in one [44]. This study would guide the conservationists to turn away the focus of conservation endeavor from *T. progeneius* to *T. putitora*.

The present study recognized two morphologically distinct groups of mahseer within the genus *Neolissochilus*. Among them, the N_1 and N_2 samples were identified to be *Neolissochilus hexagonolepis* and *N. hexastichus* respectively. DNA barcoding also differentiated both the species with considerable barcode gap and hence their identifications were confirmed. This study added in GenBank 3 replica barcode sequences of *N. hexagonolepis* and 7 new barcode sequences of *N. hexastichus*. The latter species has long been concealed since its first description in around 175 year back [18] due to lack of its morpho-taxonomic details and mis-identification with *T. tor*. The species *N. hexastichus* is though reported from other locations in the Salween basin [45] and Myanmar [46] but there are almost no biological data available on this species. Yet, it was first categorized into 'Vulnerable' [15] and subsequently to 'Near Threatened' status [16]. In this study, two species of mahseer, viz., *T. putitora* and *N. hexagonolepis* were found frequently in all the mahseer habitats in the study area. On the other hand, the species *N. hexastichus* was absent in all the surveyed habitats except a particular river (25.420 N 92.993 E) in the entire study area that raises a serious concern about the future sustainability of this species. Although this river also harbors the other two most common species of mahseer but we observed illegal harvest of fishes through destructive fishing in the river. Thus, the mahseer species in this river are assumed to have been threatened from anthropogenic activities that demands mass awareness. This study would provide benefit to generate life history parameters of *N. hexastichus* for its conservation standpoint and development of aquaculture package of practice for sustainable utilization. Therefore, this study suggests to initiate priority conservation of *N. hexastichus*.

One of the differentiating characters of two genera *Neolissochilus* and *Tor* is based on the presence of labial groove interrupted in the former and continuous in the latter. This generic character was found to be confusing because this difference was not evident in *N. hexastichus*. Therefore, we consider that interrupted labial groove would be confusing to treat as the generic character of *Neolissochilus*. On the other hand, the characteristic difference of the number of gill rakers on the first arm of gill arch was found to be a very pronounced generic character of the two genera that may be emphasized in genus categorization.

The NJ, ML and Bayesian cluster showed that the genera *Puntius* and *Hypsibarbus* remained as out-group with respect to the two genera *Tor* and *Neolissochilus* of mahseer. In another study the two genera of mahseer have been proposed to be in a distinct clade compared to other six different clades within the subfamily Cyprininae [47]. So, the grouping of mahseer in a separate tribe [34,48] appeared justified, but, the particular tribe name is contentious. In the NJ phylogenetic analysis, some sequences, e.g., *T. macrolepis* (2 sequences) and *T. mosal mahanadicus* (3 sequences), though carried distinct names in the database but clustered cohesively with a popular species *T. putitora*. Such a wrong clustering of sequences may arise either due to misidentification or due to the occurrence of synonymous species, such as *T. macrolepis* has been stated to be a synonym of *T. putitora* [11,17]. Besides, the two samples of *Neolissochilus stracheyi* did not cluster with each other and have been possibly misidentified in the database.

In the history of taxonomy, the dawn of DNA barcoding technique has sufficiently helped in troubleshooting of many species identification where morphological characters were overlooked or overemphasized [29]. Yet, the reference database is found to be lacking of information on many extant species of mahseer. Hence, development of both new barcodes and replica barcodes from wide spatial scale would be important to enrich the DNA barcode reference library. New barcodes are particularly essential to achieve the objective of DNA barcoding to complete the digital taxonomic guide of earth's biota, while the replica barcodes from wide geographical ranges would substantiate the range distribution of the extant species.

Supporting Information

Figure S1 Scheme of measurement of morphometric variables on Fish. (adopted from Jayaram (1999) [23].
(TIF)

Figure S2 ML phylogeny. The tagging of the sequences with red and black dots as well as black triangles follow the same description as given for NJ phylogenetic tree in Figure 3.
(TIF)

Figure S3 Bayesian phylogeny. The specimens' GenBank accession number and species name are shown for each taxon. The

sequences highlighted with red and blue colour correspond to the sequences developed in this study while blue coloured sequences alone correspond to the sequences of samples morphologically identified as *Tor progeneius*, but are found conspecific with *Tor putitora* in this study hence, marked as *Tor putitora.* The green coloured sequences correspond to the cases of abnormal clustering.
(TIF)

Table S1 List of the studied species, GenBank Accession of the analyzed sequences and the geographical positions of the sample.
(DOC)

Table S2 Morphometric details of the studied species.
(DOC)

Table S3 Pairwise K2P divergence matrix between the sequences.
(XLS)

Supporting Information S1 Comparison of taxonomic descriptions based on morphology of *T. Progeneius* from time to time.
(DOC)

Acknowledgments

Our humble thanks go to the Department of Biotechnology, Government of India, for providing infrastructure facilities (vide No. BT/HRD/01/002/2007).

Author Contributions

Conceived and designed the experiments: SKG BAL MJB. Performed the experiments: BAL MJB BD PM SK. Analyzed the data: BAL MJB PM SK. Contributed reagents/materials/analysis tools: BD PM SK. Wrote the paper: BAL MJB SKG.

References

1. Hamilton F (1822) An Account of the Fishes found in the River Ganges and its Branches. Edinburg. 405 p.
2. Desai VR (2003) Synopsis of biological data on the *Tor* mahseer *Tor tor* (Hamilton, 1822). FAO Fisheries synopsis, Rome, FAO. 158: 36 p.
3. Rainboth WJ (1985) *Neolissochilus*, a new genus of south Asian Cyprinid fishes. Beaufortia. 35: 25–35.
4. Talwar PK, Jhingran A (1991) Inland fishes of India and adjacent countries. Oxford and IBH publishing, New Delhi. 1 & 2: 1158 p.
5. Menon AGK (1992) Taxonomy of mahseer fishes on the genus *Tor* of Gray with description of new species from Deccan. J. Bombay Nat. Hist. Soc. India. 89: 244–258.
6. Roberts TR (1999) Fishes of the cyprinid genus Tor in the Nam Theun watershed, Mekong Basin of Laos, with description of a new species. Raffles Bulletin of Zoology. 47: 225–236.
7. Bagra K, Kadu K, Sharma KN, Laskar BA, Sarkar UK, et al. (2009) Ichthyological survey and review of the checklist of fish fauna of Arunachal Pradesh, India. Check List. 5(2): 330–350.
8. Eschmeyer WN (ed) (2012) Catalog of Fishes. California Academy of Sciences. Available: http://research.calacademy.org/research/ichthyology/catalog/fishcatmain.asp. Accessed 2012 Aug 04.
9. Thomas HS (1897) The Rod in India. Thacker and Company, London. 317 p.
10. McDonald JA (1948) Circumventing the Mahseer and other sporting fish in India and Burma. Natraj publication, Dehradun. 306 p.
11. Nguyen TTT (2008) Population structure in the highly fragmented range of *Tor douronensis* (Cyprinidae) in Sarawak, Malaysia, revealed by microsatellite DNA markers. Freshwater Biology. 53: 924–934.
12. FAO (2005) Fisheries and Aquaculture topics. Recreational fisheries. Topics Fact Sheets. Text by Andrew Smith. In: FAO Fisheries and Aquaculture Department [online]. Rome. Updated 27 May 2005. Accessed on 19 July 2012. http://www.fao.org/fishery/topic/14831/en.
13. Nautiyal P (2006) Rising awareness and efforts to conserve the Indian mahseers. Current Science. 91: 1604.
14. FAO (2006) FAO newsletter No.35. http://www.fao.org/docrep/013/a0595e/a0595e00.htm.
15. Molur S, Walker S (1998) Report of the workshop "Conservation, Assessment and Management plan (CAMP) for freshwater fishes of India. Zoo Outreach Organization, Conservation Breeding specialist group, Coimbatore. 156 p.
16. IUCN (2012) IUCN Red List of Threatened Species. Version 2012.1. <www.iucnredlist.org>. accessed on 04 August 2012.
17. Mohindra V, Khare, Praveen Lal KK, Punia P, Singh RK, et al. (2007) Molecular discrimination of five Mahseer species from Indian peninsula using RAPD analysis. Acta Zoologica Sinica 53: 725–732.
18. McClelland J (1839) Indian Cyprinidae. Asiatic Researches Calcutta, Bishop College Press. 19(2): 217–468.
19. Day F (1878) The fishes of India: being a natural history of the fishes known to inhabit the seas and fresh waters of India, Burma and Ceylon. Vols. I & II. Williams and Norgate, London. 778 p.
20. Sen TK, Jayaram KC (1982) The Mahseer Fishes of India: A review. Rec. Zool. Surv. India, Occ. Paper 39: 1–38.
21. Hora SL (1941) The game fishes of India: VIII. Mahseers of the large scaled barbels of India, 6. The Jungha of Assamese, *Barbus* (*Tor*) *progeneius* McClelland. Journal of Bombay Natural History Society 12: 526–532.
22. Hora SL, Mukerji DD (1936) Fish of the Naga Hills, Assam. Records of Indian Museum 38: 328–330.
23. Jayaram KC (1999) The freshwater fishes of Indian region. Narendra publishing house, Delhi. 551p.
24. Avise JC (1994) Molecular markers, natural history and evolution. New York: Chapman & Hall.
25. Tautz D, Arctander P, Minelli A, Thomas RH, Vogler AP (2003) A plea for DNA taxonomy. Trends Ecol. Evol. 18: 70–74.
26. Hebert PDN, Stoekle M, Zemlak T, Francis CM (2004) Identification of birds through DNA barcodes. Plos Biology 2: 1657–1668.
27. Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN (2005) DNA barcoding of Australia's fish species. Philos Trans R Soc Lond B Biol Sci 360: 1847–1857.
28. Ward RD, Hanner R, Hebert PDN (2009) The campaign to DNA barcode all fishes, FISH-BOL. Journal of Fish Biology 74: 329–356.
29. April J, Mayden RL, Hanner RH, Bernatchez L (2011) Genetic calibration of species diversity among North America's freshwater fishes. Proc Natl Acad Sci USA 108: 10602–10607.
30. Clare EB, Lim BK, Engstrom MD, Eger JL, Hebert PDN (2006) DNA barcoding of Neotropical bats: species identification and discovery within Guyana. Molecular Ecology Notes 7: 184–190.
31. Smith MA (2008) Extreme diversity of tropical parasitoid wasps exposed by iterative integration of natural history, DNA barcoding, morphology, and collections. Proc Natl Acad Sci USA 105: 12359–12364.
32. Frezal L, Leblois R (2008) Four years of DNA barcoding: current advances and prospects. Infection, genetics and evolution. Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases 8: 727–736.
33. Bhattacharjee MJ, Laskar BA, Dhar B, Ghosh SK (2012) Identification and re-evaluation of freshwater catfishes through DNA barcoding. PLoS ONE 7(11): e49950. Doi:10.1371/journal.pone.0049950.
34. Rainboth WJ (1996) FAO species identification field guide for fishery purposes, fishes of the Cambodia Mekong. FAO, Rome. 265 p.
35. Sambrook J, Russell DW (2001) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, New York.
36. Zhang DX, Hewitt GM (1996) Nuclear integrations: challenge for mitochondrial DNA markers. Trends in Ecology and Evolution 11: 247–251.
37. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. Journal of Molecular Biology 215: 403–410.
38. Thompson JD, Gibson TJ, Higgins DG (2002) Multiple sequence alignment using ClustalW and ClustalX. Current protocols in Bioinformatics/editorial board, Andreas D Baxevanis [et al] Chapter 2: Unit 2.3.
39. Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution 16: 111–120.
40. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Molecular Biology and Evolution 10: 2731–2739.
41. Ronquist F, Teslenko M, van der Mark P, Ayres D, Darling A, et al. (2012) MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology 61: 539–542.
42. Singh M, Kumar R, Nagpure NS, Kushwaha B, Mani I, et al. (2009) Extensive NOR site polymorphism in geographically isolated populations of Golden mahseer, Tor putitora. Genome 52: 783–789.
43. Byrkjedal I, Rees DJ, Willassen E (2007) Lumping lumpsuckers: Molecular and morphological insights into the taxonomic status of Eumicrotremus spinosus (Fabricius, 1776) and E. eggvinii Koefoed, 1956 (Teleostei: Cyclopteridae). J. Fish Biol. 71: 111–13.
44. Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proceedings of the National Academy of Sciences, USA 101: 14812–14817.
45. Doi A (1997) A review of taxonomic studies of cypriniform fishes in Southeast Asia. Jap. J. Ichthyol. 44: 1–33.

46. Oo W (2002) Inland fisheries of the Union of Myanmar. In T. Petr and DB Swar (eds.) Cold Water Fisheries in the Trans-Himalayan Countries. FAO Fish. Tech. Pap. 431: 70–76.
47. Yang L, Mayden RL, Sado T, He S, Saitoh K, et al. (2010) Molecular phylogeny of the fishes traditionally referred to Cyprinini sensu stricto (Teleostei: Cypriniformes) Zoologica Scripta 39: 527–550.
48. Devi RK, Indra TJ (1997) Check List of the Native Freshwater Fishes of India. Available: zsi.gov.in/checklist/Native%20freshwater%20Fishes%20of%20India. Accessed 2012 Aug 04.

OPEN ACCESS Freely available online

Identification and Re-Evaluation of Freshwater Catfishes through DNA Barcoding

Maloyjo J. Bhattacharjee, Boni A. Laskar, Bishal Dhar, Sankar K. Ghosh*

Department of Biotechnology, Assam University, Assam, India

Abstract

Background: Catfishes are globally demanded as human food, angling sport and aquariums keeping thus are highly exploited all over the world. North-East India possess high abundance of catfishes and are equally exploited through decades. The strategies for conservation necessitate understanding the actual species composition, which is hampered due to sporadic descriptions of the species through traditional taxonomy. Therefore, actual catfish diversity in this region is important to be studied through the combined approach of morphological and molecular technique of DNA barcoding.

Methodology/Principal Findings: Altogether 75 native catfish specimens were collected from across the North-East India and their morphological features were compared with the taxonomic keys. The detailed taxonomic study identified 25 species belonging to 17 genera and 9 families. The cytochrome oxidase c subunit-I gene fragment were then sequenced from the samples in accordance with the standard DNA barcoding protocols. The sequences were compared with public databases, viz., GenBank and BOLD. Sequences developed in the current study and from databases of the same and related taxa were analyzed to calculate the congeneric and conspecific genetic divergences using Kimura 2-parameter distance model, and a Neighbor Joining tree was created using software MEGA5.1. The DNA barcoding approach delineated 21 distinct species showing 4.33 folds of difference between the nearest congeners. Four species, viz., *Amblyceps apangi*, *Glyptothorax telchitta*, *G. trilineatus* and *Erethistes pusillus*, showed high conspecific divergence; hence their identification through molecular approach remained inconclusive. On the other hand, the database sequences for three species, viz., *Mystus horai, Bagarius yarrelli* and *Clarias batrachus*, appeared mislabeled.

Conclusion: The efficiency of DNA barcoding was reaffirmed from its success by easily identifying the major share (84%) of the studied catfish into 21 distinct species. The study contributed 27 new barcodes for 7 species and confirmed the range expansion of 2 important species in NE India.

Citation: Bhattacharjee MJ, Laskar BA, Dhar B, Ghosh SK (2012) Identification and Re-Evaluation of Freshwater Catfishes through DNA Barcoding. PLoS ONE 7(11): e49950. doi:10.1371/journal.pone.0049950

Editor: Bernd Schierwater, University of Veterinary Medicine Hanover, Germany

Received November 29, 2011; **Accepted** October 19, 2012; **Published** November 15, 2012

Copyright: © 2012 Bhattacharjee et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The Department of Biotechnology, Government of India, provided the instrumental facilities. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. No additional external funding received for this study.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: drsankarghosh@gmail.com

Introduction

Catfishes are members of the order Siluriformes (Actinopterygii) and inhabit inland and marine ecosystems. They are generally bottom dwellers and feed upon almost any kind of plant or animal matters, hence, play an important role in transferring energy throughout the food web [1]. Most of the species hold demand all over the world, including North-East (NE) India, as human food and aquarium keeping [2]. However, due to pressure from unregulated harvest for commercial sale along with other anthropogenic and environmental threats in NE India, some of the native catfish species have become threatened (www.iucnredlist.org). NE India is rich in biodiversity and shares two of the Biodiversity Hotspots in the world, viz., the Eastern Himalaya and the Indo-Burma [3]. The region is bestowed with numerous water bodies of diverse nature and is home to around 267 species of fishes, including many endemic catfishes such as *Amblyceps apangi*, *Amblyceps arunachalensis*, *Glyptothorax striatus* [4]. The inventories of fishes from this region were plethoric and entirely based on conventional taxonomy. Due to adherent impediments with traditional taxonomy [5], a few species once claimed new are remarked to be not valid. Instead, many cases of synonym species have been uncovered and await taxonomic revision. For example, the occurrence of *M. vittatus* in NE has been debated repeatedly [6,7,8] and is a great concern for the systematic. Moreover, taxonomic confusions exist with some other species of the genera *Sperata, Ompok, Eutropiichthys, Clupisoma, Gagata* and *Nangra*. The congeners of *Sperata* are distinguishable by either round or spatula-shaped snout or length of maxillary barbells that either extends to base of caudal fin or no further than pelvic fins. The congeners of *Eutropiichthys* are differentiable based on length of maxillary barbells and number of fin rays. Since all such characters are prominent only in adults hence the specimens at early stage are difficult to identify [6]. *Ompok bimaculatus* has been considered ''restricted in southern India'' and its conspecificity throughout the Indian subcontinent was remarked to be doubtful. Rather, the populations of the species from different areas of the subcontinent were assumed to be representing different species [9,10]. Nevertheless, the congeners of *Amblyceps* from NE India were often synonymized with *Amblyceps mangois* [11]. *Pterocryptus indicus*

was often described endemic in NE India but mostly remarked doubtful about its validity [12]. Many such cases are due to unresolved issues related to proper documentation of catfish diversity in NE India. Indeed, the earlier inventories seem to be non-exclusive and sporadic and a correct checklist of catfish diversity in the region is unavailable. Therefore, evaluation of actual catfish diversity using molecular tools is important to resolve the species and develop strategies for conservation of threatened taxa in the region. It is expected that the perplexity in identification of many catfish species of the region can be resolved and made easier through the intervention of the advanced DNA barcode based species identification technique.

The DNA barcoding concept has been launched as a rapid, accurate, automatable, and globally accessible procedure for species delimitation and identification [13]. The effectiveness of this method relies on the relatively conserved stretch of approximately 655 nucleotides of the mitochondrial cytochrome oxidase c subunit-I (COI) gene. Based on the nucleotide sequences, accurate identification of organisms at the species level is reasonably straightforward and has been applied to numerous animal taxa [14,15,16,17]. The DNA barcode reference library is rapidly growing by the contributions of the global community in the Barcode of Life Data Systems (BOLD) [18]. With the glory as an attractive species identifier in fish biodiversity research [19], the application of DNA barcoding has recently reflected 28% increase of North American freshwater fish diversity [20]. The technique further bears application in monitoring fish products for health safety [21,22] and in regulating the exploitation of fish species under aquarium trade [23,24,25].

DNA barcoding technique was adopted in NE India to study the actual diversity of catfishes inhabiting in the region. This will also enrich the database by contributing both new barcodes and

Table 1. Summary of identification based on each species consensus barcoded sequence using BLASTN search from GenBank and BOLD Identification System (BOLD-IDS).

Sl. No.	Studied species	Species match by name		% Similarity	
		GenBank (BLASTN)	BOLD-IDS	GenBank (BLASTN)	BOLD-IDS
1.	*Rita rita* (1)	*Rita rita*	*Rita rita*	99	100
2.	*Mystus bleekeri* (8)	*Mystus bocourti*	*No match*	88	No match
3.	*M. cavasius* (4)	*Mystus oculatus*	*No match*	89	No match
4.	***M. vittatus* (7)**	*Mystus vittatus*	*Mystus vittatus*	99	99.83
		Mystus horai	***Mystus horai***	**99**	**99.64**
5.	*Sperata aor* (3)	*Sperata aor*	*Sperata aor*	100	100
6.	*Hemibagrus menoda* (2)	*Sperata aor*	*No match*	86	No match
7.	***Bagarius bagarius* (3)**	*Bagarius bagarius*	*Bagarius bagarius*	100	100
		Bagarius yarrelli	***Bagarius yarrelli***	**100**	**100**
		Bagarius yarrelli		91	
8.	*Gagata cenia* (6)	*Gagata cenia*	*Gagata cenia*	99	99.5
9.	*Gagata sexualis* (2)	*Gagata sexualis*	*Gagata sexualis*	99	99
10.	***Glyptothorax telchitta* (3)**	***Glyptothorax telchitta***	*No match*	**93**	No match
11.	*G. striatus* (1)	*Glyptothorax striatus*	*Glyptothorax striatus*	97	97.6
12.	***G. trilineatus* (1)**	***Glyptothorax trilineatus***	*No match*	**96**	No match
13.	*Sisor rabdophorus* (4)	*Sisor rabdophorus*	*Sisor rabdophorus*	100	100
14.	*Ailia coila* (3)	*Ailia coila*	*Ailia coila*	99	99.67
15.	*Clupisoma garua* (1)	*Laides hexanema*	*No match*	91	No match
16.	*Eutropiichthys murius* (3)	*Pangasius larnaudii*	*No match*	86	No match
17.	*E. vacha* (7)	*Laides hexanema*	*No match*	89	No match
18.	*Ompok bimaculatus* (2)	*Ompok bimaculatus*	*Ompok bimaculatus*	99	99.84
19.	*O. pabo* (2)	*O. pabo*	*O. pabo*	99	99.33
20.	*Wallago attu* (2)	*Wallago attu*	*Wallago attu*	100	100
21.	***Clarias batrachus* (2)**	*Clarias batrachus*	*Clarias batrachus*	98	98.64
		Clarias batrachus		**90**	
22.	*Heteropneustes fossilis* (3)	*Heteropneustes fossilis*	*No match*	100	No match
23.	*Erethistes pusillus* (2)	*Erethistes pusillus*	*No match*	93	No match
24.	***Amblyceps apangi* (1)**	*Amblyceps apangi*	*No match*	95	No match
		Amblyecps apangi		**88**	
25.	*Olyra longicaudata* (1)	*Amblyceps mucronatum*	*No match*	90	No match

• Similarity description used in the study- 97%–100%– significant, 92%–96%– moderate, ≤91%– insignificant.
• Bolded words correspond to problematic identification of species in the present study using either one or both the databases. Details are further discussed in the text.
• Numbers in brackets indicate the number of individual sequences of each species.
doi:10.1371/journal.pone.0049950.t001

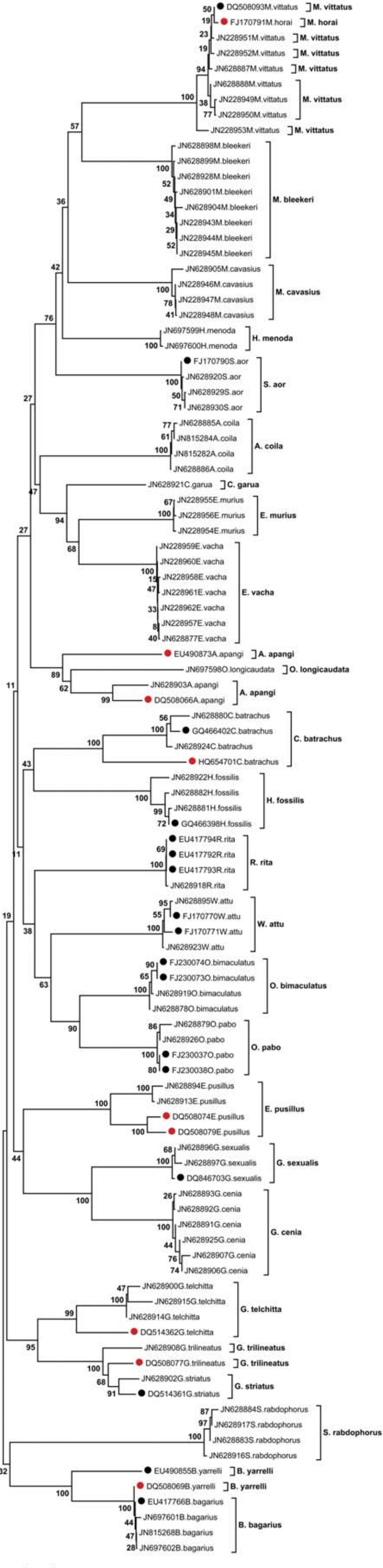


Figure 1. Neighbor joining (NJ) tree developed using K2P distance among 101 CO1 sequences. Notable anomalies in clustering are shown by 4 species {*Mystus horai* (accession number FJ170791), *Bagarius yarrelli* (accession number DQ508069), *Clarius batrachus* (accession number HQ654701) and *Amblyceps apangi* (accession number EU490873). Deep conspecific divergences are shown by 3 species (*Glyptothorax trilineatus, Glyptothorax telchitta* and *Erethistes pusillus*). • The numbers at the nodes are bootstrap values based on 1000 replications. • Specimen GenBank accession number and species name are shown for each taxon. • Red and black dots correspond to the sequences acquired from database. Red dot alone corresponds to the cases of abnormal clustering and deep conspecific divergence.
doi:10.1371/journal.pone.0049950.g001

replica of existing barcodes thereby enabling the evaluation of taxonomic status of the native catfish diversity in NE India. Here we studied the first DNA barcode based taxonomic resolution of freshwater catfishes from NE India to resolve key areas of doubt arising from morphological taxonomy. This investigation not only prove the potential use of DNA barcoding as a tool to aid traditional taxonomy of freshwater catfish but also will help further in easy identification of the studied species from any of their body parts and at any stage of life. Nevertheless the sequences generated from this study would be accessible to establish the conspecificity of NE Indian catfish with other geographical location and vice versa.

Materials and Methods

Sample collection

The native catfishes were collected from different natural water bodies during different seasons of the year from random geographical locations within the NE region of India (28°09′N 97°24′E on the East to 27°49′N 88°15′E on the West, and 29°18′N 96°04′E on the North to 92°59′E on the South). Fresh specimens were spot examined for specific morphological characters that define the catfish and sampled from different wild habitats as and when caught by the professional fishers. Each of the catch was investigated by an experienced fish taxonomist to ensure correct sampling and labeling. Upon every spate of collection, the tissue samples from each of the specimens were collected aseptically and preserved in 90% ethanol. Major taxonomic keys of each of the fish specimens were noted and measurements were taken using a digital caliper. Species level identification of the specimens was confirmed by comparing with the described characters and the taxonomic keys available in the leading taxonomic guides of the fishes in India [6,26]. However, the nomenclature of species follows the Catalogue of Fishes [27]. The comparisons of the observed characters in respect of each species with their described characters along with the particular dispute are presented in Table S1. Altogether 75 fish specimens belonging to 25 species within 17 genera and 9 families were collected and included in this study. All the voucher specimens have been deposited in the Department of Biotechnology, Assam University, Silchar. The specimen information, IUCN Red list status and distribution of the studied species are given in Table S2.

Since the studied fishes were routine caught by the professional fishers for sale hence no permission was required for their sampling.

DNA extraction

20 mg of anal fin tissue was taken aseptically and dissolved in 500 μL of TES buffer (50 mM Tris HCl, 25 mM EDTA and 150 mM NaCl) in a microcentrifuge tube. The extraction of DNA was performed with Phenol-Chloroform-Isoamylalcohol method [28].

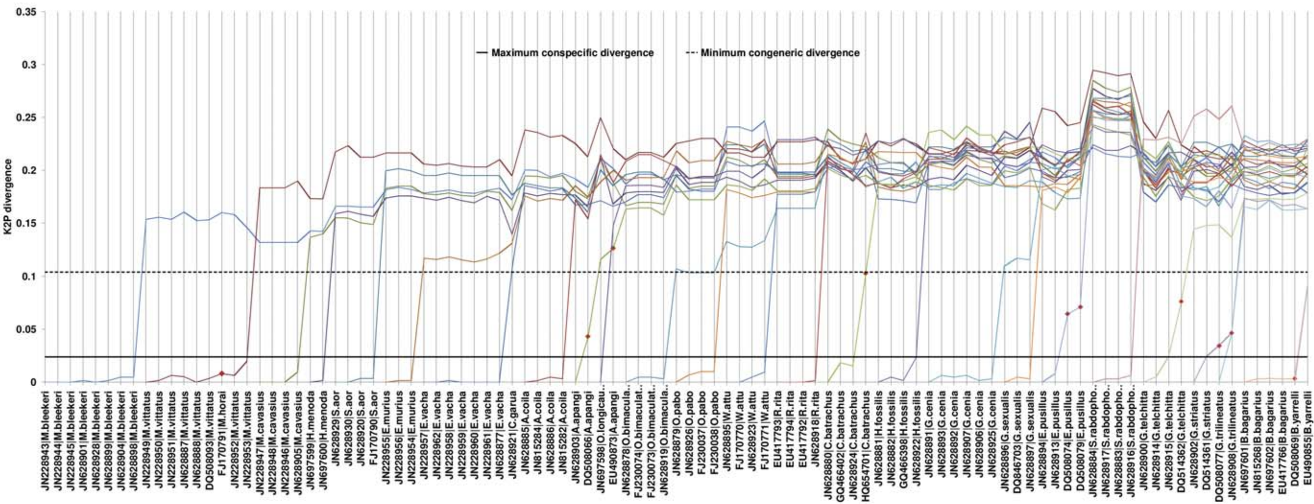


Figure 2. Congeneric and conspecific K2P divergence for 101 sequences of 27 catfish species. The maximum conspecific divergence (0.024, black solid line) and minimum congeneric divergence (0.104, black dotted line) represent the threshold level of conspecific and congeneric divergence respectively. Most of the studied species (21) obeyed the thresholds and are readily delineated showing barcoding gap of 4.33 or above. Sequences of species like *M. horai* (accession number FJ170791), *B. yarrelli* (accession number DG508069), *A. apangi* (accession numbers EU490873 and DQ508066), *C. batrachus* (accession number HQ654701), *E. pusillus* (accession numbers DQ508074 and DQ508079), *G. telchitta* (accession number DQ514362), *G. trilineatus* (accession number DQ508077) did not obey the thresholds and are thus ambiguous (shown in red dots). *G. striatus* with two sequences obeyed the threshold of maximum conspecific divergence and minimum congeneric divergence with all congeners except *G. trilineatus*. • In the X-axis the specimens involved in this study were plotted and marked as, GenBank accession number|species name.
doi:10.1371/journal.pone.0049950.g002

PCR amplification and purification

The COI gene (655 bp) was amplified using the set of published primers [29] as follows. FishF1-5′ TCAACCAACCACAAAGACATTGGCAC 3′ and FishR1-5′ TAGACTTCTGGGTGGCCAAAGAATCA 3′ in a Veriti Mastercycler (Applied Biosystems Inc., CA, USA). The amplification reactions were performed in a total volume of 25 µl comprising 1X PCR buffer, 2 mM $MgCl_2$, 10 pmol of each primer, 0.25 mM of each dNTPs, 0.25 U of high-fidelity Taq polymerase (Applied Biosystems Inc., CA, USA) and 100 ng of DNA template. The thermal profile of the PCR reaction was as follows: An initial denaturation at 94°C for 2 minutes, 30 cycles at denaturation temperature of 94°C for 45 seconds, annealing temperature of 50°C for 45 seconds and elongation temperature of 72°C for 1 minute, and concluded with a final elongation step at 72°C for 8 minutes followed by a hold at 4°C. The PCR-amplified products were analyzed in 1% agarose gels containing ethidium bromide staining (10 mg/ml) and the single uniform band was then purified using QIAquick^R Gel extraction kit (QIAGEN, USA), following manufacturer's instructions. The amplicons were bidirectionally sequenced in an automated DNA sequencer (ABI 3500, Applied Biosystems Inc., CA, USA), through the best known service of GCC Biotech India Pvt. Ltd. (Kolkata, India). The COI amplicons were recovered from all the collected specimens.

Sequence quality control measures

Both the PCR amplified products and their corresponding DNA sequences were larger than 600 bp that assured no NUMTs being amplified as the limit of NUMT hardly reaches 600 bp [30]. Ends of the noisy sequences were trimmed and more than 600 bp sequences were used for the final analysis (except in four cases, accession numbers JN697602, JN628915, JN628929 and JN628930). For each sample two chromatograms that represent sequences of both the strands of DNA were obtained. BLASTN [31] program was used to compare the sequences from the two chromatograms, and the fragment of the two sequences showing 100% alignment with no gap or indels (insertion/deletions) was selected. In case of any discrepancy, both the sequences were reviewed and quality value of the sequence was considered to determine the most likely nucleotide using the software SeqScanner Version 1.0 (Applied Biosystems Inc., CA, USA). In most cases, the sequence quality values were above 50. The selected fragments of the sequences for all the specimens were aligned using ClustalX software [32] and found no indels in any of the sequences. Finally, each of the sequences were subjected to BLASTN searches at the National Centre for Biotechnology Information [31], that showed alignment with the partial coding sequence of fish mitochondrial COI gene without any indels. The sequences were translated using the online software ORF finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) and aligned through BLASTP [31] that revealed coherent partial amino acid codes with fish mitochondrial COI gene frame without any stop codon. Therefore, it was confirmed that the generated sequences were fragments of mitochondrial COI gene. All the analyzed sequences were then deposited in GenBank and received valid accession numbers (Table S1). The sequences were also submitted by creating a FISH-BOL project in BOLD in the code name of 'CFISH' entitled "DNA barcoding of freshwater catfishes of Northeast India".

Data analysis

The BOLD provided 778 COI sequences for catfish (accessed on 17 November, 2011) among them the database sequences of the same and/or related taxa are used in association with developed sequences for evaluating the taxonomic status of our target species. The total dataset included 101 COI barcode sequences for 27 catfish species among which 75 sequences belonging to 25 species were developed de-novo, and 26 sequences representing same and related taxa were acquired from GenBank only because there were no additional sequences available in BOLD other than those mined from GenBank source. Geograph-

ical information and GenBank accession numbers of the developed as well as acquired database sequences are given in Table S2. The sequence similarity search for species identification was done in two public databases, viz., BOLD and GenBank. The highest percent pairwise identity for each sequence blasted (BLASTN) at NCBI were compared with the percent similarity scores of the same sequence within the BOLD-IDS (BOLD Identification System) [18]. The query species that matched either with the same or different species in the databases has been termed as 'specific' or 'non-specific' respectively. The similarity range of 97%–100%, 92%–96% and ≤91% between the query and the database sequence have been expressed as significant, moderate and insignificant respectively. Kimura 2-parameter (K2P) congeneric and conspecific variation [33] and Neighbor Joining (NJ) tree construction were done using the computer program MEGA Version 5 [34]. Maximum conspecific and minimum congeneric divergences have been determined considering the sequences showing cohesive NJ clustering within a species and remained distinct from other species. The number of times the minimum congeneric divergence differs from the maximum conspecific divergence is the lowest divergence between congeners and has been assumed to be the threshold level of species delineation and thereby considered as a barcoding gap in this study.

Results

Comprehensive species identification of the studied catfishes based on BOLD and GenBank databases is depicted in Table 1. The study helped in straightforward identification of 10 species that showed significant species specific similarities in both the databases. The species are *Sperata aor, Sisor rabdophorus*, *Wallago attu*, *Gagata sexualis*, *Rita rita*, *Gagata cenia*, *Glyptothorax striatus*, *Ailia coila*, *Ompok bimaculatus* and *Ompok pabo*. GenBank sequences showed moderate species specific similarity for both *Glyptothorax telchitta* and *Erethistes pusillus* at 93%, *G. trilineatus* at 96% and significant species specific similarity for *Heteropneustes fossilis* at 100%. It also showed insignificant non-specific similarity (≤91%) for seven species, viz., *Mystus bleekeri*, *M. cavasius*, *Hemibagrus menoda*, *Clupisoma garua*, *Eutropiichthys murius*, *E. vacha* and *Olyra longicaudata*. Both the databases showed significant species specific similarity (≥98%) for *M. vittatus, B. Bagarius* and *Clarias batrachus*, as well as significant non-specific similarity (≥99%) for *M. vittatus* (query) with *M. horai* (database accession number FJ170791), and *B. bagarius* (query) with *B. yarrelli* (database accession number DQ508069). GenBank alone showed insignificant species specific similarity (90%) for *C. batrachus* with a database accession number HQ654701. GenBank concurrently showed moderate species specific similarity (95%) and insignificant species specific similarity (88%) for *Amblyceps apangi* with database accession numbers DQ508066 and EU490873 respectively.

The similarity search result thereby confirmed definitive identity showing significant species specific match in GenBank and BOLD for 11 species, viz., *R. rita*, *S. aor*, *G. cenia*, *G. sexualis*, *G. striatus*, *S. rabdophorus*, *A. coila*, *O. bimaculatus*, *O. pabo, W. attu* and *H. fossilis*. *H. fossilis* latter was identified by GenBank alone. The rest of the studied species (14) showed ambiguous match categories, like, 1) significant but equally species specific and non-specific (e.g. *M. vittatus* and *B. Bagarius*), 2) species specific but equally significant and insignificant (e.g. *C. batrachus*), 3) species specific but moderate (e.g. *G. telchitta*, *G. trilineatus*, *E. pusillus* and *A. apangi*), 4) species specific but insignificant (e.g. *A. apangi* accession number EU490873), and 5) non-specific and insignificant (e.g. *M. bleekeri*, *M. cavasius*, *H. menoda*, *C. garua*, *E. murius*, *E. vacha* and *O. longicaudata*).

The Neighbor Joining (NJ) cluster analysis (Figure 1) revealed straight forward identification showing either a single or distinct cluster of individual(s) for 18 of our studied species. These include 11 accurately identified species which showed significant species specific similarity and 7 species which showed insignificant non-specific match (in parity with ambiguous match category-5). However, ambiguities persisted for 7 other species and showed three distinct patterns: 1) same and different named-species clustered together (in parity with ambiguous match category-1, e.g., all sequences of *M. vittatus* clustered with *M. horai* (accession number FJ170791) and all sequences of *B. bagarius* clustered with *B. yarrelli* (accession number DQ508069), 2) same named-species clustered both jointly and distinctly (in parity with ambiguous match category-2, e.g., *C. batrachus* of accession number GQ466402 and HQ654701), and same named-species clustered only distinctly (in parity with ambiguous match category-4, e.g., *A. apangi* of accession number EU490873), and 3) high range of clustering differences with conspecific query sequences (in parity with ambiguous match category-3, e.g., *G. telchitta* of accession number DQ514362, *G. trilineatus* of accession number DQ508077, *E. pusillus* of accession number DQ508074 and DQ508079, and *A. apangi* of accession number DQ508066).

The minimum congeneric and maximum conspecific K2P divergences were determined to be 0.104 and 0.024 respectively and presented in Figure 2. Based on these divergence values, a 4.33 folds barcoding gap was calculated. 72% (18) of the studied species identified through the NJ clustering have also been easily delineated following the barcoding gap. Few database sequences of the same and/or related species have not obeyed this gap and hence designated as ambiguous. For example, 1) a few congeneric sequences merged within the range of conspecific divergence (e.g., *M. horai* of accession number FJ170791 and *B. yarrelli* of accession number DQ506089) and vice versa (e.g., *C. batrachus* of accession number HQ654701 and *A. apangi* of accession number EU490873), 2) a few individual sequences were widely dispersed from their conspecific sequences and did not reach the congeneric threshold divergence (e.g., *G. trilineatus* of accession number DQ508077, *G. telchitta* of accession number DQ514362, *E. pusillus* of accession numbers DQ508074 and DQ508079 and *A. apangi* of accession number DQ508066), and finally 3) congeneric distances of *G. striatus* with G. *trilineatus* have remained much below the congeneric threshold divergence.

Discussion

This study of identification of catfishes from NE India was based on the morphological investigation followed by DNA barcoding approach. The morphological study of the specimens has raised a few questions on the observed features versus the described features. In a few cases, morphological species keys were difficult to discern. Moreover, disparities relating to the species keys were observed in a few cases between the two leading taxonomic guide books of fishes in India (Table S1). The DNA barcoding approach resolved some identification issues and explained the actual species composition in the region.

Among the 25 studied catfish species, the similarity search approach revealed two straightforward cases for 18 species. Firstly, 11 species, viz., *R. rita*, *S. aor*, *G. cenia*, *G. sexualis*, *G. striatus*, *S. rabdophorus*, *A. coila*, *O. bimaculatus*, *O. pabo*, *W. attu* and *H. fossilis*, showed significant species specific similarity in the range of 97%–100%, and were readily identified as true species. Secondly, seven species, viz., *M. bleekeri*, *M. cavasius*, *H. menoda*, *C. garua*, *E. murius*, *E. vacha* and *O. longicaudata*, showed insignificant non-specific similarity (≤91%) in GenBank and no match in BOLD. This has

reflected the lack of barcode reference data for these species in both the databases. However, all the sequences of the above mentioned 18 species showed conspecific NJ clustering by the specimens within each species having well supported bootstrap proportion (>95%) [35]. Further, all the 18 species were definitely delineated considering barcoding gap principle (Figure 2) and identified as true species based on the combined approaches, including 7 species whose barcode data were not available in the databases previously. So, the study contributed new barcode data in the global database for those seven species. Range expansion of *O. bimaculatus* in NE India was also evident from this study through observed high conspecificity among the queries and database sequences from the Ganga basin (FJ230073–4) in India. We concentrated to delineate the species based on the threshold level of species divergence taking the maximum conspecific versus minimum congeneric divergence into account rather than considering the conventional mean value of congeneric and conspecific divergence. This has led to the reflection of the lowest barcoding gap of 4.33 folds compared to the previous DNA barcoding studies of fishes [23,29,36,37,38].

While morphological identification were convincing up to species-level for all the studied species, DNA barcoding even remained inconclusive for 7 species (viz., *M. vittatus*, *B. bagarius*, *C. batrachus*, *A. apangi*, *G. trilineatus*, *G. telchitta*, and *E. pusillus*), at its first hand approach. Combined approach has confirmed that database sequence of *M. horai* and *M. vittatus* are conspecific. Moreover, the query sequences of *M. vittatus* showed conspecificity with a same named database sequence of accession number DQ508093. Previous morphological studies have already raised doubtfulness on the taxonomic validity of *M. horai* and the species was remarked to be not recorded from any other location than its type locality (in Indus drainage) [27,39]. Since the molecular evidences also reckon the previous morphological debate, hence the study tentatively considered *M. horai* as a synonym of *M. vittatus*. As such the study identified *M. vittatus* being a true species and recognized its range expansion in NE India. Therefore, the study resolved the debate surrounding the existence of *M. vittatus* in NE India [7,8]. Again, one of the two database sequences of *B. yarrelli* (accession number DQ508069) revealed conspecific divergence with *B. bagarius* while the other sequence (accession number EU490855) maintained congeneric divergence with the same. This reflected that the former sequence of *B. yarrelli* is mislabeled in the database. This study thus confirmed both *B. bagarius* and *B. yarrelli* being true species and met with the previous argument [40]. In case of *C. batrachus*, a single database sequence (accession number HQ654701) showed congeneric divergence with other conspecific sequences. On the other hand, the queries as well as the other database sequences of *C. batrachus* showed conspecificity. This confirmed *C. batrachus* to be a true species and indicated a clear mislabeling of the said sequence in the database. Similar cases of mislabeling have also been reported earlier [22]. Ignoring the few mislabeled database sequences, the identification of above three species was confirmed.

In another case, one of the two database sequences of *Amblyceps apangi* (accession number EU490873) was identified to be a distinct congener of the query sequence. This may be again a case of mislabeling because the two species within the genus *Amblyceps* are endemic to NE India [41] and inadequately described or frequently synonymised [11,42]. The other database sequence of *A. apangi* (accession number DQ508066) and all the database sequences of *G. trilineatus*, *G. telchitta*, and *E. pusillus* were not conspecific with the respective query sequences, and remained below the congeneric threshold thereby revealed deep conspecific divergences. This indicated possible cases of independently evolving lineages of a species from different geographical location [20] or cryptic species with low divergence, or even recently-diverged overlooked species [23,29]. As such, the identification of species under the genera, viz., *Glyptothorax, Erethistes*, and *Amblyceps* remained inconclusive due to inadequate and perplexing descriptions from conventional taxonomy. Hence, in order to develop a correct barcode reference library, there is a paramount need of extensive revision, combining morphological and DNA barcoding of the extant species under these genera.

Thus, it can be concluded that 21 species representing 84% of the studied catfish species were identified straightforward through DNA barcoding that reaffirmed the efficacy of the technique. The study resolved some cases of synonymy, clarified the range distribution and revealed the catfish diversity in NE India. Occurrence of *Pterocryptis indicus* in NE India was not evident in the study and hence holds to agree upon arguments on the doubtful status of this species [12]. However, remaining 16% of the studied species representing 3 genera remained inconclusive and warrant further evaluation. Few database sequences were observed to be bearing misidentified species caption among the species those possess either confusing morphological description or share crypticism. Given that the database is enriched with the multiple sequences for a target species and for the extant species within a target genus from the range of distribution, the species taxonomy would be rectified and assessment of biodiversity would be correct and easier.

Supporting Information

Table S1 Morphological taxonomic keys observed versus described.
(DOC)

Table S2 List of the studied species, GenBank accession numbers of the analyzed sequences, the geographical position, and IUCN status.
(DOC)

Acknowledgments

Our humble acknowledgement goes to the Department of Biotechnology, Government of India for providing infrastructure facilities (BT/HRD/01/002/2007) and post-doc fellowship to Dr. Boni A Laskar in the Department of Biotechnology, Assam University Silchar. Comments on the manuscript from reviewers have helped to enrich the revised version to present in a better way and are highly acknowledged.

Author Contributions

Conceived and designed the experiments: SKG MJB. Performed the experiments: SKG MJB BAL BD. Analyzed the data: MJB. Contributed reagents/materials/analysis tools: SKG MJB. Wrote the paper: SKG MJB BAL.

References

1. Layher WG, Maughan OE (1985) Relations between habitat variables and channel catfish populations in prairie streams. Transactions of the American Fisheries Society 114: 771–781.
2. FAO (2008–2012) Small-scale fisheries- website. Nutrition and food security. In: FAO Fisheries and Aquaculture Department (online) Rome. Available: http://www.fao.org/fishery/topic/16603/en. Accessed 2012 Mar 19.

3. Myers N (1999) What we must do to counter the biotic holocaust. International wildlife 29: 6.
4. Ponniah AG, Sarkar UK(2000) Fish biodiversity of North East India: National Bureau of Fish Genetic Resources.
5. Hoagland KE (1996) The Taxonomic Impediment and the Convention on Biodiversity. ASC News 24: 61–62, 66–67.
6. Talwar PK, Jhingran AG (1991) Inland fishes of India and adjacent countries: Oxford & IBH Pub. Co. Vol 2. pp554–1158.
7. Roberts TR (1992) Revision of the striped catfishes of Thailand misidentified as *Mystus vittatus*, with description of two new species (Pisces: Bagridae). Ichthyol. Explor. Freshwat 3: 77–88.
8. Darshan A, Anganthoibi N, Vishwanath W (2010) Redescription of the striped catfish Mystus carcio (Hamilton) (Siluriformes: Bagridae). Zootaxa 2475: 48–54.
9. Kottelat M, Lim KKP (1995) Freshwater fishes of Sarawak and Brunei Darussalam: a preliminary annotated checklist. Sarawak Museum Journal 48: 227–258.
10. Ng HH, Hadiaty RK (2009) Ompok brevirictus, new catfish (Teleostei: Siluridae) from Sumatra. Zootaxa 2232: 50–60.
11. Vishwanath W, Linthoingambi I (2007) Redescription of catfishes *Amblyceps arunachalensis* Nath and Dey and *Amblyceps apangi* Nath and Dey (Teleostei:Amblycipitidae). Zoos' Print Journal 22: 2662–2664.
12. Ferraris CJ Jr (2007) Checklist of catfishes, recent and fossil (Osteichthyes: Siluriformes) and catalogue of siluriform primary types. Zootaxa 1418: 1–628.
13. Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. Proc Biol Sci 270: 313–321.
14. Hebert PDN, Stoekle M, Zemlak T, Francis CM (2004) Identification of birds through DNA barcodes. Plos Biology 2: 1657–1668.
15. Clare EB, Lim BK, Engstrom MD, Eger JL, Hebert PDN (2006) DNA barcoding of Neotropical bats: species identification and discovery within Guyana. Molecular Ecology Notes 7: 184–190.
16. Smith MA, Rodriguez JJ, Whitfield JB, Deans AR, Janzen DH, et al. (2008) Extreme diversity of tropical parasitoid wasps exposed by iterative integration of natural history, DNA barcoding, morphology, and collections. Proc Natl Acad Sci USA 105: 12359–12364.
17. Frezal L, Leblois R (2008) Four years of DNA barcoding: current advances and prospects. Infection, genetics and evolution. Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases 8: 727–736.
18. Ratnasingham S, Hebert PDN (2007) BOLD: The Barcode of Life Data System. Molecular Ecology Notes 7: 355–364. Available: http://www.barcodinglife.org.
19. Ward RD, Hanner R, Hebert PDN (2009) The campaign to DNA barcode all fishes, FISH-BOL. Journal of Fish Biology 74: 329–356.
20. April J, Mayden RL, Hanner RH, Bernatchez L (2011) Genetic calibration of species diversity among North America's freshwater fishes. Proc Natl Acad Sci USA 108: 10602–10607.
21. Lowenstein JH, Burger J, Jeitner CW, Amato G, Kolokotronis SO, et al. (2010) DNA barcodes reveal species-specific mercury levels in tuna sushi that pose a health risk to consumers. Biology Letters 6: 692–695.
22. Wong LL, Peatman E, Lu J, Kucuktas H, He S, et al. (2011) DNA barcoding of catfish: species authentication and phylogenetic assessment. PloS ONE 6: e17812.
23. Steinke D, Zemlak TS, Hebert PDN (2009) Barcoding nemo: DNA-based identifications for the ornamental fish trade. PloS ONE 4: e6300.
24. Mohd-Shamsudin MI, Fard MZ, Mather PB, Suleiman Z, Hassan R, et al. (2011) Molecular characterization of relatedness among colour variants of Asian Arowana (Scleropages formosus). Gene 490: 47–53.
25. Collins RA, Armstrong KF, Meier R, Yi Y, Brown SD, et al. (2012) Barcoding and border biosecurity: identifying cyprinid fishes in the aquarium trade. PloS ONE 7: e28381.
26. Jayaram KC (1999) The freshwater fishes of the Indian region: Narendra Pub. House.
27. Eschmeyer WN, Frickle R (2011) Catalog of fishes. California Academy of Sciences. Available: http://research.calacademy.org/research/ichthyology/catalog/fishcatmain.asp. Accessed 2011 Nov 17.
28. Sambrook J, Russell DW (2001) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, New York.
29. Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN (2005)DNA barcoding of Australia's fish species. Philos Trans R Soc Lond B Biol Sci 360: 1847–1857.
30. Zhang DX, Hewitt GM (1996) Nuclear integrations: challenge for mitochondrial DNA markers. Trends in Ecology and Evolution 11: 247–251.
31. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. Journal of Molecular Biology 215: 403–410.
32. Thompson JD, Gibson TJ, Higgins DG (2002) Multiple sequence alignment using ClustalW and ClustalX. Current protocols in bioinformatics/editorial board, Andreas D Baxevanis [et al] Chapter 2: Unit 2.3.
33. Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution 16: 111–120.
34. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Molecular Biology and Evolution 10: 2731–2739.
35. Felsenstein J (1985) Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39: 783–791.
36. Hubert N, Hanner R, Holm E, Mandrak NE, Taylor E, et al. (2008) Identifying Canadian Freshwater Fishes through DNA Barcodes. PLoS ONE 3: e2490.
37. Lakra WS, Verma MS, Goswami M, Lal KK, Mohindra V, et al. (2011) DNA barcoding of Indian marine fishes. Molecular Ecology Resources 11: 60–71.
38. Mabragaña E, Díaz de Astarloa JM, Hanner R, Zhang J, González Castro M (2011) DNA Barcoding Identifies Argentine Fishes from Marine and Brackish Waters. PLoS ONE 6: e28655. doi:10.1371/journal.pone.0028655.
39. Talwar PK, Jhingran AG (1991) Inland fishes of India and adjacent countries: Oxford & IBH Pub. Co. Vol 2. 554–560.
40. Roberts TR (1983) Revision of south and southeast Asian sisorid catfish genus *Bagarius*, with description of a new species from the Mekong. Copeia 2: 435–445.
41. Nath P, Dey SC (1989) Two new species of the genus *Amblyceps* Blyth from Arunachal Pradesh, India. Journal of Assam science 32: 1–6.
42. Ng HH (2005) *Amblyceps carinatum*, a new species of hillstream catfish from Myanmar (Teleostei: Amblycipitidae). Raffles bulletin of Zoology 53: 243–249.

DOI: 10.2478/dna-2014-0003 — DNA Barcodes 2014; Volume 2: 17–20

Short communication Open Access

Subrata Trivedi*, Raafat Affan, Abdulrahman Hirab A Alessa, Abid A Ansari, Bishal Dhar, Pradosh Mahadani, Sankar K Ghosh

DNA Barcoding of Red Sea Fishes from Saudi Arabia – The first approach

Abstract: DNA barcoding is a worldwide initiative. Although there is a campaign to barcode all fishes, there is no information on the DNA barcoding of Red Sea fishes. With the aim of strengthening the BOLD/barcode database with the sequences of Red sea fishes this preliminary study was conducted with six different fish species collected from coastal waters of Tabuk in Saudi Arabia. Among the six fish samples collected from the same sampling site, two sequences were novel. So, it is necessary to carry out large scale sampling covering a bigger area of Red Sea because there is high possibility to detect new/cryptic species from the unexplored biodiversity rich Red Sea.

Keywords: DNA Barcoding, Fishes, Red Sea, COI gene, sequence analysis, Saudi Arabia.

***Corresponding author: Subrata Trivedi:** Department of Biology, Faculty of Science, University of Tabuk, Tabuk-71491, Ministry of Higher Education, Kingdom of Saudi Arabia, trivedi.subrata@gmail.com, strivedi@ut.edu.sa, Phone 00966-553669517
Raafat Affan, Abdulrahman Hirab A Alessa, Abid A Ansari: Department of Biology, Faculty of Science, University of Tabuk, Tabuk-71491, Ministry of Higher Education, Kingdom of Saudi Arabia
Bishal Dhar, Pradosh Mahadani, Sankar K Ghosh: Department of Biotechnology, Assam University, Silchar-788011, Assam, India

Introduction

Fishes are important sources of protein, vitamins and minerals for human consumption. Large amount of unrefined oil extracted from fishes is used in making paints, pesticides, soaps, medicines, etc. The marine fisheries sector contributes significantly to food security, income generation, and economic welfare. The proper identification of fish species is not only of interest for taxonomy and systematics, but also of prime importance in management of fisheries and authentication of food products. DNA barcoding is an efficient way of species identification. DNA barcoding promises fast, accurate species identifications by focusing analysis on a short standardized segment of the genome. Mitochondrial *Cytochrome c oxidase* subunit 1 (COI) is regarded as barcode region of varied animal species [1]. Several studies have now established that sequence diversity in a ~ 650 bp region near the 5′ region of COI gene provides strong species level resolution for varied animal groups including birds [2–4], springtails [5], spiders [6], and also fishes [7–10].

DNA barcoding has gained worldwide popularity as an effective tool for species identification. The Red Sea is rich in biodiversity with large varieties of fish and other marine animals. Although there is a campaign to barcode all fishes [11], there is no information on the DNA barcoding of Red Sea fishes. This is the first approach to Barcode Red Sea fishes. With the aim of strengthening the BOLD/barcode database with the sequences of Red Sea fishes, we have analyzed the COI sequences of six different fish species that were collected from the Red Sea coastal waters of Tabuk in Saudi Arabia. One of our major targets was to search for new/cryptic fish species, if any, from the Red Sea.

Methods

Sample Collection

Fish samples were collected from Red Sea coastal waters of Al-Khuraybah (28.058 N 35.171 E) located in Tabuk province of Saudi Arabia. Fishes were carried from the sample collection site to the laboratory in University of Tabuk in ice box. Small amount of muscle tissue was collected aseptically from behind of dorsal fin and stored in sterile 1.5 ml tube containing absolute alcohol. The whole body specimens were preserved and stored in the Department of Biology, University of Tabuk.

Extraction of DNA

Muscle tissue samples were washed in TES buffer containing 50mM Tris HCl (pH 8.0), 25 mM EDTA (pH 8.0), 150 mM NaCl. The sample was then homogenized in

© 2014 Subrata Trivedi, et al., licensee De Gruyter Open. This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 License.
Unauthenticated | 212.138.88.100
Download Date | 4/29/14 12:12 PM

J. Environ. & Sociobiol. : **10**(1) : 83-88, 2013

EXPLORING ORNAMENTAL FISHES OF MANIPUR THROUGH DNA BARCODING

Bijoya Khomdram Devi, Bishal Dhar, Boni Amin Laskar and Sankar Kumar Ghosh*
DNA barcode and Genomics Laboratory, Department of Biotechnology, Assam University, Silchar-788011, Assam, India

ABSTRACT

DNA barcoding is an effective technique of species identification. The large reference library of DNA barcode should be enriched from every region and may be assessed together with morphological taxonomy for species authentication. In this study, a total of 27 fish samples belonging to 21 different species having ornamental relevance were collected and morphologically identified from Manipur, Northeast India. DNA was isolated from muscle tissue of samples and then COI barcode region was amplified using universal fish primers, sequenced and submitted to NCBI (Accession nos. JQ713844 - JQ713860). The sample sequences were identified through similarity search in two databases: GenBank and BOLD. Both the databases revealed definitive identity matches in the range of 98.37% - 100% for 44% of the species. This indicated that the database still lacks reference data of ornamental fishes of Manipur and this study contributed new barcode sequences of ornamental fishes.

Key words : *Ornamental fish, Manipur, DNA barcoding, COI, sequence homology*

INTRODUCTION

Ornamental fishes usually signify attractive colorful fishes culturable in an aquarium or in a garden pool for fun and fancy. The increasing demand for aquarium fishes gradually paved the avenue towards global trade of ornamental fishes. According to the international trade in tropical aquarium fish, global market of ornamental is about US $600 million at wholesale level. Indian waters possess a rich diversity of attractive fish, with over 100 indigenous varieties and about 80% of ornamental fishes are from fresh waters (Ponniah and Sarkar, 2006). World trade of the fish is estimated to be about Rs. 2000 crores, but India's share is only Rs. 15 crores, which is very insignificant. The major shares of the country's ornamental fish export are captive collection from Northeastern states comprising about 85% of the total aquarium fish trade of India (Ponniah and Sarkar, 2006). Due to the diversity of topographic and climatic features of Manipur of North-east India, this region is rich in endemic fish. Most of the small food fish (139

* Department of Biotechnology, Assam University, Silchar-788011, Assam, India, Telephone No: +919435372338, Fax: +91 3842270802, Email ID: drsankarghosh@gmail.com

J. Environ. & Sociobiol. : **10**(1) : 65-75, 2013

ROLE OF TRANSVERSIONAL SUBSTITUTION ON DNA BARCODE BASED SPECIES DIFFERENTIATION: CATFISH AS A CASE STUDY

Maloyjo J. Bhattacharjee, Boni A. Laskar, Bishal Dhar, Sankar K. Ghosh*

Department of Biotechnology, Assam University,
Silchar-788011, Assam, India

ABSTRACT

Cytochrome c oxidase (*COI*) DNA barcode is considered to be universal species identifier for animals based on many superlative features, but none has considered the pattern of characteristic nucleotide variation for their role if any. In this study, 50 *COI* barcode sequences of 10 species of catfishes were analyzed in two steps. In the first step, accurate delineation of species was confirmed while the second step involves analysis of the nucleotide variability in terms of transition and transversion at the different taxon levels and their subsequent effect on genetic divergence. The analysis revealed that the individuals within all the species exhibit distinct barcode cluster and a straight forward divergence at the congeneric (17.4% ± 1.8%) and conspecific level (0.7% ± 0.3%) allowing their unambiguous identification. Transition was found to be significantly higher (t= 4.85 - 41.12, $p<0.0001$) than transversion at all the three taxon levels. However, as transversion showed a significant increase (0.47 ± 0.38, 29.65 ± 5.12 and 34.21 ± 5.50) with the increase in taxon comparison (conspecific, congeneric and intergeneric), transition only showed a significant rise (t= 119.6, $p<0.00001$) between the conspecific (1.64 ± 0.76) and congeneric level (54.96 ± 6.75) and then transition became saturated at the intergeneric level (54.83 ± 6.68). The study importantly revealed that the genetic divergence between the congeners at the partial fragment of *COI* was highly correlated (Pearson correlation, R= 0.935, $P<0.001$) with transversional substitution than with transition (R= 0.334, $P<0.05$). Hence, although transition is more frequent, still transversional substitution within the *COI* barcode sequence plays the key role in species delineation.

Key words : *COI, DNA barcode, nucleotide substitution, transition, transversion*

INTRODUCTION

The DNA barcoding concept launched in the year 2003 has pinpointed that a 650 nucleotide stretch of mitochondrial cytochrome oxidase c subunit 1 (*COI*) gene is the ultimate identifier of a species (Hebert *et al*., 2003). The sequence can distinguish closely related species based on the variable nucleotides and has been used successfully for a

* E-mail: drsankarghosh@gmail.com. Tel: 91 9435372338

J. Environ. & Sociobiol. : **10**(1) : 25-32, 2013

REVEALING GENETIC DIVERSITY OF *CLARIAS BATRACHUS* USING DNA BARCODE

Mohua Chakraborty, S.K.Ghosh,* Bishal Dhar and Bijoya Khomdram Devi

DNA barcode and Genomics Laboratory, Department of Biotechnology, Assam University, Silchar-788011, Assam, India

ABSTRACT

Clarias batrachus is a species of freshwater catfish widely used for human consumption. Increasing demand of this species worldwide, coupled with its ability to survive in wide range of environmental conditions, has led to its introduction in many countries. This widespread translocation and distribution of *C. batrachus* have resulted in the rise of wide variety of haplotypes of this species. Nevertheless, there has been constant decrement in their population density in the last two decades. Thus, to endeavour conservation of the species we employ molecular technique of DNA barcoding in solving the standing problem of crypticism and haplotype sharing of the species. A better knowhow of the genetic makeup of the unique identifier region, that is, the 648 basepair region of COI DNA barcode will help to differentiate among closely related species and identify endemic species. In this study, a comparative analysis of *C. batrachus* from different regions in India and other parts of world shows presence of distinct haplotypes in different geographical locations. We also present a descriptive study of the various species of *Clarias* genus that have been barcoded in India till date. Our results also solve the dilemma of considering some species as synonymy of *C. batrachus.*

Key words : *Clarias batrachus, cytochrome c oxidase subunit I (COI), genetic diversity, Kimura's two parameter (K2P), mean divergence, neighbour joining tree.*

INTRODUCTION

The walking catfish, *Clarias batrachus*, is a species of freshwater air breathing catfish so named for its ability to "walk" across dry land, to find food or suitable environments. They normally inhabit in swamps, marshy and derelict waters (Lakra and Sarkar, 2007). The walking catfish is a native of South East Asia including Malaysia, Thailand, eastern India, Sri Lanka, Bangladesh, Burma, Indonesia, Singapore and Brunei (Kuang, 1986; Shrestha, 1978) and is also found in Philippines. Owing to its capability to adapt to wide range of environmental conditions, candidates of the genus *Clarias* has extensively travelled to many continents, adapting itself successfully and is now found throughout Asia and Africa.

* Telephone No: +919435372338, Fax: +91 3842270802, Email ID: drsankarghosh@gmail.com

J. Environ. & Sociobiol. : **10**(1) : 55-63, 2013

IMPLICATION OF NUCLEOTIDE SUBSTITUTION AT THIRD CODON POSITION OF THE DNA BARCODE SEQUENCES

Fazlur Rahman Talukdar, Sankar K. Ghosh*, Ruhina S. Laskar, Pradosh Mahadani, Mohua Chakraborty, Bishal Dhar and M. Joyraj Bhattacharjee

Department of Biotechnology, Assam University, Silchar-788011, Assam, India

ABSTRACT

DNA markers (barcode) differentiate species based on their nucleotide sequence diversity among various species. In this study we analyzed the rate and pattern of nucleotide substitution and their consequent influence on the amino acid substitution patterns of the sequences used as barcode mitochondrial COI, cyt *b* and the exon 1 of nuclear IRBP gene for animals from 15 different species of vertebrates. The analysis shows unlike other codon positions, nucleotide substitution at the third codon position does not show strong correlation with the amino acid substitution, for the three gene sequences. Furthermore, COI gene shows a very low percentage of amino acid variability (15.38%) inspite of high percentage of variation in its nucleotide sequence (40.76%) as well as a significantly ($p<0.0001$) low level of amino acid sequence divergence than the other gene fragments under study. Interestingly among the compared sequences, a significantly conserved amino acid substitution pattern seems to be a unique feature of barcode region of the COI gene making it a more efficient marker for species identification. Hence, it was concluded that the property of species identification of these sequences is based upon the variable nature of third codon position.

Key words : *Species-specific marker, DNA barcoding, nucleotide substitution, codon positions, amino acid substitution*

INTRODUCTION

Diversity of animal kingdom provides a challenging task for taxonomists to identify and characterize the millions of species thriving on earth. In addition to morphology-based identification systems, recent scientific investigations have revealed that DNA-based markers are promising and might provide more accurate and easy identification and characterization of species. DNA markers distinguish species based on sequence diversity in small segments of DNA among different species and the DNA used for this purpose can be both protein coding and protein non-coding fragments (Kress and Erickson, 2012).

* E-mail: drsankarghosh@gmail.com.

DNA Barcoding: Digital Taxonomy of Bioresources

Sankar Kumar Ghosh*

Maloyjo Joyraj Bhattacharjee, Ksh. Miranda Devi, Monika Ahanthem, Shantanu Kundu, Pradosh Mahadani, Bishal Dhar, Khomdram Bijoya Devi, Mohua Chakraborty, Fazlur Rahman, Rosy Mondal, Jagdish Hansa, Ruhina S Laskar, Tarikul Huda Mazumder, Priyanka Sarkar, Shiba Rajbongshi, Arijit Chakraborty, Mridul M Das, Probal Ranjan Ghosh, Kulendra Chandra Das & Boni Amin Laskar

Department of Biotechnology, Assam University, Silchar-788011

Email: drsankarghosh@gmail.com

*Supervisor and presenting author

Biodiversity assessments and implementation of conservation actions worldwide are hampered by slow progress in taxonomic research, termed the Taxonomic Impediment[1]. The existing workforce of taxonomists cannot cope with the overwhelming need for basic field surveys, species descriptions and systematic revisions to provide basic information for conservation planning. In addition, few taxonomists are able to distinguish critically between more than 1000 taxa[2]. The reality facing taxonomy is that there may be more species that remain to be discovered on Earth than those that have already been described[3]. It is estimated that 1.4-1.8 million species have been described[4-6] out of a possible total of approximately 7-15 million6. Species and populations are going extinct at an alarming but poorly understood rate[7,8]. Many species may be going to extinct before they can be identified or described. This presents a problem for conservation planning and prioritization, obviously because species that have not been identified cannot be protected effectively and is evident throughout the world. Taxonomic expertise is lacking even for major and commercially important groups. The few taxonomists who are working in developing countries, home to more than 95% of globally described species; find it difficult to access basic taxonomic information such as species descriptions[9]. Where taxonomic keys are available, they are rarely revised and often inadequate to identify specimens unambiguously to the species level[10]. There have been strident calls on the taxonomy community to embrace new technologies and to form networks to speed up the description of biodiversity[11,12] and to improve our ability to identify species[2, 13-16]. How this could be achieved without compromising rigorous taxonomic research principles is, however, questioned[17]. There is legitimate concern that too little money is being spent on morphological taxonomy compared to molecular studies[18]. A purely DNA taxonomy approach (e.g.Blaxter[19]), not to be confused with DNA barcoding[21] is too simplistic in our view. DNA taxonomy can, however, provide additional characters for discrimination, especially in cases where other characters vary among species and are thus difficult to interpret. Therefore we do not agree with de Carvalho *et at.*[17] that a molecular approach will do little to address the real problems in taxonomy. In our experience, most traditional taxonomists welcome the opportunity to refer to molecular data, but DNA taxonomy should not replace taxonomic research that can be based on multiple character data sets. Nucleotide sequence divergence in short, standardized gene regions as DNA barcodes can be used to identify known species and facilitate the discovery of new ones[13,21]. Mitochondrial DNA is a valuable marker in population genetic or phylogeographic studies because it is maternally inherited, evolves rapidly and recombination is rare or absent[22]. Therefore, a part of the mitochondrial cytochrome oxidase subunit I (COI) has been chosen as a standard gene region for barcoding animals. Although there was no prior reason for choosing COI among the 13 protein-coding mitochondrial genes[13] it has the advantage of having robust universal primers that can recover the 5' end of COI of most animal species. Barcoding of species has become cheaper through the technological advances made by other molecular programmes, especially the human genome project[23]. Rapid barcoding and comparison with the growing database of COI sequences[24] will increase the speed of identification of newly collected or unknown specimens. This may focus taxonomists' attention on unidentified lineages, and with the addition of morphological and other taxonomically relevant data, could lead to a faster rate of species description. Kress et al. (2005)[25] suggest that the use of the COI sequence is not appropriate for most species of plants because of a much slower rate of cytochrome c oxidase I gene evolution in higher plants than in animals. A series of experiments was then conducted to find a more suitable region of the genome for use in the DNA barcoding of flowering plants (or the larger group of land plants)[26]. One 2005 proposal was the nuclear internal transcribed spacer region and the plastid trnH-psbA intergenic spacer[25] other researchers advocated other regions such as matK[26]. In 2009, a collaboration of a large group of plant DNA barcode researchers proposed two chloroplast genes, rbcL and matK, taken together,

RESEARCH ARTICLE

Influence of the *CYP1A1* T3801C Polymorphism on Tobacco and Alcohol-Associated Head and Neck Cancer Susceptibility in Northeast India

Seram Anil Singh[&], Javed Hussain Choudhury[&], Wetetsho Kapfo, Sharbadeb Kundu, Bishal Dhar, Shaheen Laskar, Raima Das, Manish Kumar, Sankar Kumar Ghosh*

Abstract

Background: Tobacco and alcohol contain or may generate carcinogenic compounds related to cancers. *CYP1A1* enzymes act upon these carcinogens before elimination from the body. The aim of this study was to investigate whether *CYP1A1* T3801C polymorphism modulates the relationship between tobacco and alcohol-associated head and neck cancer (HNC) susceptibility among the northeast Indian population. Materials and Methods: One hundred and seventy histologically confirmed HNC cases and 230 controls were included within the study. The *CYP1A1* T3801C polymorphism was determined using PCR-RFLP, and the results were confirmed by DNA sequencing. Logistic regression (LR) and multifactor dimensionality reduction (MDR) approaches were applied for statistical analysis. Results: The *CYP1A1* CC genotype was significantly associated with HNC risk (P=0.045). A significantly increased risk of HNC (OR=6.09; P<0.0001) was observed in individuals with combined habits of smoking, alcohol drinking and tobacco-betel quid chewing. Further, gene-environment interactions revealed enhanced risks of HNC among smokers, alcohol drinkers and tobacco-betel quid chewers carrying *CYP1A1* TC or CC genotypes. The highest risk of HNC was observed among smokers (OR=7.55; P=0.009) and chewers (OR=10.8; P<0.0001) carrying the *CYP1A1* CC genotype. In MDR analysis, the best model for HNC risk was the three-factor model combination of smoking, tobacco-betel quid chewing and the *CYP1A1* variant genotype (CVC=99/100; TBA=0.605; P<0.0001); whereas interaction entropy graphs showed synergistic interaction between tobacco habits and *CYP1A1*. Conclusions: Our results confirm that the *CYP1A1* T3801C polymorphism modifies the risk of HNC and further demonstrated importance of gene-environment interaction.

Keywords: Head and neck cancer - *CYP1A1* T3801C polymorphism - tobacco - alcohol - northeast India

Asian Pac J Cancer Prev, **16 (16)**, 6953-6961

Introduction

Head and neck cancers (HNC) are a heterogeneous group of cancers that include cancers of the skin of the head and neck, nasal cavity, paranasal sinuses, nasopharynx, lip, oral cavity, oropharynx, larynx, hypopharynx, trachea, neck, salivary glands and the parapharyngeal region (Siegel et al., 2014). Incidence of HNC has increased at an alarming rate for the past 10 years. It is the fifth most prevalent among all cancers throughout the world with estimated death of 355 per 100,000 (Masood et al., 2014). These cancers occur more frequently in men (male: female ratio of 4-5:1). In India, HNC acquires 30-40% cancers of all sites and is the sixth most common cause of death in males and seventh in females (Bhattacharjee et al., 2006).

Tobacco smoking and betel quid chewing with or without tobacco are major risk factors to be associated with HNC (Mondal et al., 2013; Amtha et al., 2014; Choudhury et al., 2014; Choudhury and Ghosh, 2014). Alcohol habit is another risk factor associated with HNC (Risch et al., 2003; Hashibe et al., 2009; Ghosh et al., 2014b). The magnitude of risk conferred by the interaction between tobacco, and alcohol is still unclear. Studies have implicated that many carcinogens present in tobacco, and alcohols are metabolized to active forms that have deleterious effects in our body. These activated substances can cause oxidative reactions in tissues, and initiate reactions to produce free radicals (Cury et al., 2012). The presence of reactive oxygen species (ROS) can cause damage to cellular biomolecules, including protein and DNA, consequently resulting in carcinogenesis (Mondal and Ghosh, 2013; Kumar and Muniyandi, 2015).

Metabolic enzymes that are potentially involved in either the activation (phase I) or detoxication (phase II) of chemical carcinogens have received a great deal of attention recently as possible genetic susceptibility factors

Molecular Medicine Laboratory, Department of Biotechnology, Assam University, Silchar, Assam, India [&]*Equal contributors*
**For correspondence: drsankarghosh@gmail.com*

Tumor Biol.
DOI 10.1007/s13277-015-3246-0

RESEARCH ARTICLE

Tobacco carcinogen-metabolizing genes CYP1A1, GSTM1, and GSTT1 polymorphisms and their interaction with tobacco exposure influence the risk of head and neck cancer in Northeast Indian population

Javed Hussain Choudhury · Seram Anil Singh · Sharbadeb Kundu · Biswadeep Choudhury · Fazlur R. Talukdar · Shilpee Srivasta · Ruhina S. Laskar · Bishal Dhar · Raima Das · Shaheen Laskar · Manish Kumar · Wetetsho Kapfo · Rosy Mondal · Sankar Kumar Ghosh

Received: 5 November 2014 / Accepted: 10 February 2015
© International Society of Oncology and BioMarkers (ISOBM) 2015

Abstract Genetic polymorphisms in tobacco-metabolizing genes may modulate the risk of head and neck cancer (HNC). In Northeast India, head and neck cancers and tobacco consumption remains most prevalent. The aim of the study was to investigate the combined effect of cytochrome P450 1A1 (*CYP1A1*) T3801C, glutathione S-transferases (*GSTs*) genes polymorphisms and smoking and tobacco–betel quid chewing in the risk of HNC. The study included 420 subjects (180 cases and 240 controls) from Northeast Indian population. Polymorphisms of *CYP1A1* T3801C and *GST* (*M1 & T1*) were studied by polymerase chain reaction restriction fragment length polymorphism (PCR-RFLP) and multiplex PCR, respectively. Logistic regression (LR) and multifactor dimensionality reduction (MDR) approach were applied for statistical analysis. LR analysis revealed that subjects carrying *CYP1A1* TC/CC+*GSTM1* null genotypes had 3.52-fold ($P<0.001$) increase the risk of head and neck squamous cell carcinoma (HNSCC). Smokers carrying *CYP1A1* TC/CC+ *GSTM1* null and *CYP1A1* TC/CC+*GSTT1* null genotypes showed significant association with HNC risk (odds ratio [OR]=6.42; $P<0.001$ and 3.86; $P=0.005$, respectively). Similarly, tobacco–betel quid chewers carrying *CYP1A1* TC/CC+ *GSTM1* null genotypes also had several fold increased risk of HNC ($P<0.001$). In MDR analysis, the best model for HNSCC risk was the four-factor model of tobacco–betel quid chewing, smoking, *CYP1A1* TC/CC, and *GSTM1* null genotypes (testing balance accuracy [TBA]=0.6292; cross-validation consistency [CVC]=9/10 and $P<0.0001$). These findings suggest that interaction of combined genotypes of carcinogen-metabolizing genes with environmental factors might modulate susceptibility of HNC in Northeast Indian population.

Keywords Head and neck cancer · CYP1A1 · GSTM1 · GSTT1 · Tobacco · MDR analysis

Introduction

Head and neck cancer (HNC) is one of the most common malignancy worldwide, and it refers to a group of biologically similar cancers arising from the mucous membranes of the oral cavity, nasal cavity, paranasal sinuses, pharynx, and larynx [1–3]. In India, it acquires 30–40 % cancers at all sites and is the sixth most common cause of death in males and seventh in females [4]. Tobacco chewing and smoking are given special attention in relation to HNC as they contain various types of carcinogens, including polycyclic aromatic hydrocarbons (PAHs), heterocyclic aromatic amines, and N-nitrosamine. Xenobiotics metabolizing enzymes are responsible for metabolism of many exogenous chemicals that are toxic, mutagenic, or carcinogenic [5]. Carcinogen-metabolizing enzymes include the phase I enzymes involved in the detoxification of

Javed Hussain Choudhury and Seram Anil Singh contributed equally to this work.

J. H. Choudhury · S. A. Singh · S. Kundu · B. Choudhury · F. R. Talukdar · S. Srivasta · R. S. Laskar · B. Dhar · R. Das · S. Laskar · M. Kumar · W. Kapfo · R. Mondal · S. K. Ghosh (✉)
Molecular Medicine Laboratory, Department of Biotechnology, Assam University, Silchar, Assam 788011, India
e-mail: drsankarghosh@gmail.com

Published online: 28 February 2015

Tumor Biol.
DOI 10.1007/s13277-015-3114-y

RESEARCH ARTICLE

Association of HPV with genetic and epigenetic alterations in colorectal adenocarcinoma from Indian population

Ruhina S Laskar · Fazlur R Talukdar · Javed H Choudhury · Seram Anil Singh · Sharbadeb Kundu · Bishal Dhar · Rosy Mondal · Sankar Kumar Ghosh

Received: 3 November 2014 / Accepted: 14 January 2015
© International Society of Oncology and BioMarkers (ISOBM) 2015

Abstract Several studies from developing countries have shown human papillomavirus to be associated with colorectal cancers, but the molecular characteristics of such cancers are poorly known. We studied the various genetic variations like microsatellite instability (MSI), oncogenic mutations and epigenetic deregulations like CpG island methylation in HPV associated and nonassociated colorectal cancer patients from Indian population. HPV DNA was detected by PCR using My09/My11 and Gp5+/Gp6+ consensus primers and typed using HPV16 and HPV18 specific primers. MSI was detected using BAT 25 and BAT 26 markers, and mutation of *KRAS*, *TP53* and *BRAF* V600E were detected by direct sequencing. Methyl specific polymerase chain reaction (MSP) was used to determine promoter methylation of the classical CIMP panel markers (*P16*, *hMLH1*, *MINT1*, *MINT2* and *MINT31*) and other tumour-related genes (*DAPK*, *RASSF1*, *BRCA1* and *GSTP1*). HPV DNA was detected in 34/93 (36.5 %) colorectal tumour tissues, HPV 18 being the predominant high-risk type. MSI was detected in 7.5 % cases; KRAS codon 12, 13, *BRAF* V600E and *TP53* mutations were detected in 36.5, 3.2 and 37.6 % of the cases, respectively. CIMP-high was observed in 44.08 % cases. HPV presence was not associated with age, stage or grade of tumours, MSI or mutations in *KRAS*, *TP53* or *BRAF* genes. Higher methylation frequencies of all genes/loci under study except *RASSF1*, as well as significantly higher CIMP-high characteristics were observed in HPV positive tumours as compared to negative cases. HPV in association with genetic and epigenetic features might be a potent risk factor for colorectal cancer in Indian population.

Keywords Colorectal cancer · HPV · Genetic alterations · Epigenetics · CIMP · *KRAS* · *BRAF* · *P53* · MSI · Northeast India

Introduction

Colorectal cancer (CRC) is the fourth leading cancer worldwide with approximately 1,360,602 new cases annually [1]. The genesis of CRC is thought to be a multistep process involving complex interaction of genetic, epigenetic and environmental factors or viral components in distinct geographical regions [2–5]. Although, epidemiological research have established several environmental risk factors like smoking, alcohol consumption [6] etc.; evidences of probable human papillomavirus (HPV) infection in CRC have accrued over quite some time [7–9]. However, the detection rates exhibited significant geographic variations, some studies even failing to detect HPV altogether [10–12]. This discrepancy was addressed in two comprehensive meta-analyses conducted by Damin et al. (2013) [13] and Baandrup et al. (2014) [5], where they identified two main grounds of disparity: distinct geographical variances in HPV prevalence and differences in detection techniques used. Overall, HPV presence was found to confer 6-10 folds higher risk of CRC compared to controls (Table 1). Following the well-recognised role of HPV in cervical and other anogenital carcinoma, this finding is indeed intriguing and indicative of probable HPV mediated CRC pathogenesis.

Electronic supplementary material The online version of this article (doi:10.1007/s13277-015-3114-y) contains supplementary material, which is available to authorized users.

R. S. Laskar · F. R. Talukdar · J. H. Choudhury · S. A. Singh · S. Kundu · B. Dhar · R. Mondal · S. K. Ghosh (✉)
Molecular Medicine Laboratory, Department of Biotechnology, Assam University, Silchar 788011, Assam, India
e-mail: drsankarghosh@gmail.com

Published online: 03 February 2015

OPEN ACCESS Freely available online

Mitochondrial DNA Copy Number and Risk of Oral Cancer: A Report from Northeast India

Rosy Mondal, Sankar Kumar Ghosh*, Javed Hussain Choudhury, Anil Seram, Kavita Sinha, Marine Hussain, Ruhina Shirin Laskar, Bijuli Rabha, Pradip Dey, Sabitri Ganguli, Monisha NathChoudhury, Fazlur Rahman Talukdar, Biswadeep Chaudhuri, Bishal Dhar

Department of Biotechnology, Assam University, Silchar, Assam, India

Abstract

Background: Oral squamous cell carcinoma (OSCC) is the sixth most common cancer globally. Tobacco consumption and HPV infection, both are the major risk factor for the development of oral cancer and causes mitochondrial dysfunction. Genetic polymorphisms in xenobiotic-metabolizing enzymes modify the effect of environmental exposures, thereby playing a significant role in gene–environment interactions and hence contributing to the individual susceptibility to cancer. Here, we have investigated the association of tobacco - betel quid chewing, HPV infection, *GSTM1-GSTT1* null genotypes, and tumour stages with mitochondrial DNA (mtDNA) content variation in oral cancer patients.

Methodology/Principal Findings: The study comprised of 124 cases of OSCC and 140 control subjects to PCR based detection was done for high-risk HPV using a consensus primer and multiplex PCR was done for detection of *GSTM1-GSTT1* polymorphism. A comparative ΔCt method was used for determination of mtDNA content. The risk of OSCC increased with the ceased mtDNA copy number ($P_{trend}=0.003$). The association between mtDNA copy number and OSCC risk was evident among tobacco – betel quid chewers rather than tobacco – betel quid non chewers; the interaction between mtDNA copy number and tobacco – betel quid was significant ($P=0.0005$). Significant difference was observed between *GSTM1 - GSTT1* null genotypes ($P=0.04$, $P=0.001$ respectively) and HPV infection ($P<0.001$) with mtDNA content variation in cases and controls. Positive correlation was found with decrease in mtDNA content with the increase in tumour stages ($P<0.001$). We are reporting for the first time the association of HPV infection and *GSTM1-GSTT1* null genotypes with mtDNA content in OSCC.

Conclusion: Our results indicate that the mtDNA content in tumour tissues changes with tumour stage and tobacco-betel quid chewing habits while low levels of mtDNA content suggests invasive thereby serving as a biomarker in detection of OSCC.

Citation: Mondal R, Ghosh SK, Choudhury JH, Seram A, Sinha K, et al. (2013) Mitochondrial DNA Copy Number and Risk of Oral Cancer: A Report from Northeast India. PLoS ONE 8(3): e57771. doi:10.1371/journal.pone.0057771

Editor: Syed A. Aziz, Health Canada, Canada

Received November 19, 2012; **Accepted** January 24, 2013; **Published** March 4, 2013

Copyright: © 2013 Mondal et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: Infrastructural facilities provided by Department of Biotechnology, Govt. of India. No external fund for this study. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: drsankarghosh@gmail.com

Introduction

OSCC, the most frequent tumour of oral cavity, [1] and the sixth most common cancer globally that accounts for approximately 5 per cent of all malignant tumours worldwide [2,3]. The statistical analysis by the International Agency for Research on Cancer (IARC) indicated that the lip and oral cavity is the tenth most common tumour site in the human [4]. Smokeless tobacco products and betel quid with or without tobacco are the major risk factors for oral cavity cancer in Taiwan, India, and other neighboring countries [5–7]. In Northeast India, incidence of tobacco related oral cancers is about 33% [8]. Smoking, alcohol use, smokeless tobacco products, and HPV (Human papilloma virus) infections are the major risk factors for oral cavity cancer, with smoking and alcohol having synergistic effects [9,10].

The development of carcinogenesis due to environment-gene interaction has been well illustrated by phase I and phase II enzymes that are involved in the metabolism of carcinogens. The phase I enzymes are CYPs (Cytochrome P450) that are involved in activating the environmental procarcinogens adding or exposing their functional groups whereas phase II enzyme like GST (Glutathione S-transferase) are involved in detoxication of the activated metabolites of the carcinogens [11]. Tobacco smoke is a complex mixture of carcinogenic compounds, and smokeless tobacco is rich in nitrosamines. Furthermore, the concomitant use of betel quid leads to 50-fold increase in reactive oxygen species generation (ROS) [12,13]. A structural deletion in these genes represents a null genotype and has been associated with an increased risk to oral cancer [14].

Mitochondrial defects have long been suspected to play an important role in the development and progression of cancer [15,16]. Mitochondrial respiratory activity is associated with the generation of ROS. The mitochondrial genome is susceptible to ROS and other types of genotoxic damage due to lack of