5. Pornprasert, S., Kasemrad, C. and Sukunthamala, K., Diagnosis of thalassemia on dried blood spot samples by high performance liquid chaomatography. *Hemoglobin*, 2010, **34**, 486–494.
6. Pornprasert, S., Kaewbundit, A., Phusua, A., Suanta, S., Saetung, R. and Sanguansermsri, T., Comparison of real-time polymerase chain reaction SYBR Green1with high resolution melting analysis and Taqman MGB probes for detection of alpha-thalassemia-1 South-East Asian type on dried blood spots. *Eur. J. Haematol.*, 2008, **81**, 405–407.
7. Saiyed, Z. M., Bochiwal, C., Gorasia, H., Telang, S. D. and Ramchand, C. N., Application of magnetic particles (Fe_3O_4) for isolation of genomic DNA from mammalian cells. *Anal. Biochem.*, 2006, **356**, 306–308.
8. Berensmeier, S., Magnetic particles for the separation and purification of nucleic acids. *Appl. Microbiol. Biotechnol.*, 2006, **73**, 495–504.
9. Bandyopadhyay, A., Chatterjee, S. and Sarkar, K., Rapid isolation of genomic DNA from *E. coli* XL1 Blue strain approaching bare magnetic nanoparticles. *Curr. Sci.*, 2011, **101**, 210–214.
10. Zhang, Z., Zhang, L., Chen, L., Chen, L. and Wan, Q. H., Synthesis of novel porous magnetic silica microspheres as adsorbents for isolation of genomic DNA. *Biotechnol. Prog.*, 2006, **22**, 514–518.
11. Yoza, B., Matsumoto, M. and Matsunaga, T., DNA extraction using modified bacterial magnetic particles in the presence of amino silane compound. *J. Biotechnol.*, 2002, **94**, 217–224.
12. Chiang, C. L., Sung, C. S., Wu, T. F., Chen, C. Y. and Hsu, C. Y., Application of superparamagnetic nanoparticles in purification of plasmid DNA from bacterial cells. *J. Chromatogr. B*, 2005, **822**, 54–60.
13. Intorasoot, S., Srirung, R., Intorasoot, A. and Ngamratanapaiboon, S., Application of gelatin-coated magnetic particles for isolation of genomic DNA from bacterial cells. *Anal. Biochem.*, 2009, **386**, 291–292.
14. Abe, K. and Konomi, N., Hepatitis C virus RNA in dried serum spotted onto filter paper is stable at room temperature. *J. Clin. Microbiol.*, 1998, **36**, 3070–3072.
15. Davies, M. J., Taylor, J. I., Sachsinger, N. and Bruce, I. J., Isolation of plasmid DNA using magnetite as a solid phase adsorbent. *Anal. Biochem.*, 1998, **262**, 92–94.
16. Taylor, J. I., Hurst, C. D., Davies, M. J., Sachsinger, N. and Bruce, I. J., Application of magnetite and silica–magnetite composites to the isolation of genomic DNA. *J. Chromatogr. A*, 2000, **890**, 159–166.
17. Tangvarasittichai, O., Jongjitwimol, J. and Tangvarasittichai, S., Whatman 3 MM dried blood spots for identifying alpha-thalassemia-1. *Clin. Lab.*, 2008, **54**, 281–283.
18. Boppana, S. B. *et al.*, Dried blood spot real-time polymerase chain reaction assays to screen newborns for congenital cytomegalovirus infection. *JAMA*, 2010, **303**, 1375–1382.
19. Marconi, A. *et al.*, Evaluation of the Abbott real-time HIV-1 quantitative assay with dried blood spot specimens. *Clin. Microbiol. Infect.*, 2009, **15**, 93–97.

ACKNOWLEDGEMENT. This work was supported by a research grant from Faculty of Associated Medical Sciences, Chiang Mai University, Thailand.

Received 5 February 2013; revised accepted 2 May 2013

DNA barcodes of some threatened freshwater indigenous fishes in India

Subhasree Sengupta[1,2] and Sumit Homechaudhuri[1,*]

[1]Aquatic Bioresource Research Laboratory, Department of Zoology, University of Calcutta, 35 Ballygunge Circular Road, Kolkata 700 019, India
[2]Department of Zoology, Bethune College, 181, Bidhan Sarani, Kolkata 700 006, India

This study was directed to identify early life-history stages of eight culturable species currently under threat through DNA barcodes. A total of 22 sequences of cytochrome *c* oxidase I gene were generated from eight species using multiple specimens producing 655 nucleotide base pairs per taxon. The neighbour-joining tree showed three major clusters for three different orders. Similar species were clustered under the same nodes and dissimilar species were clustered under separate nodes signifying species specificity. This can be utilized for accurate identification of threatened fish species and in turn assist in conservation of fish stock either inside freshwater protected areas or in the river systems.

Keywords: Barcoding, freshwater fish, life-history stages, neighbour-joining tree.

INDIA is endowed with a rich fish genetic biodiversity (2200 fish species) and ranks ninth in terms of freshwater mega biodiversity[1,2]. The small, shallow rivers provide refuge to a large number of varied fish population and are also undergoing major habitat alterations due to anthropogenic stresses. The Mahananda–Tangon–Punarbhaba river system (MTPRS) is one such shallow interconnected river system in Malda district, West Bengal, India, harbouring a rich fish biodiversity throughout the year. Out of this rich fish biota, eight most economically important, culturable species were selected for the present study. They were *Chitala chitala* Hamilton, 1822; *Notopterus notopterus* Pallas, 1769; *Ompok bimaculatus* Hamilton, 1822; *Eutrophichthys vacha* Hamilton, 1822; *Ailia coila* Hamilton, 1822; *Neotropius atherinoides* Bloch, 1794; *Pangasius pangasius* Hamilton, 1822 and *Gudusia chapra* Hamilton, 1822. The populations of these fish fauna are presently in decline[3] and require efforts towards conservation through culture or introduction in fish sanctuary. Propagation of stock in controlled environment requires pure seed collected from nature and hence, precise identification of early life-history stages of these fishes is important.

Appropriate identification tool for early life-history stages is essential to confirm purity of seed for introduction

*For correspondence. (e-mail: sumithomec@yahoo.com)

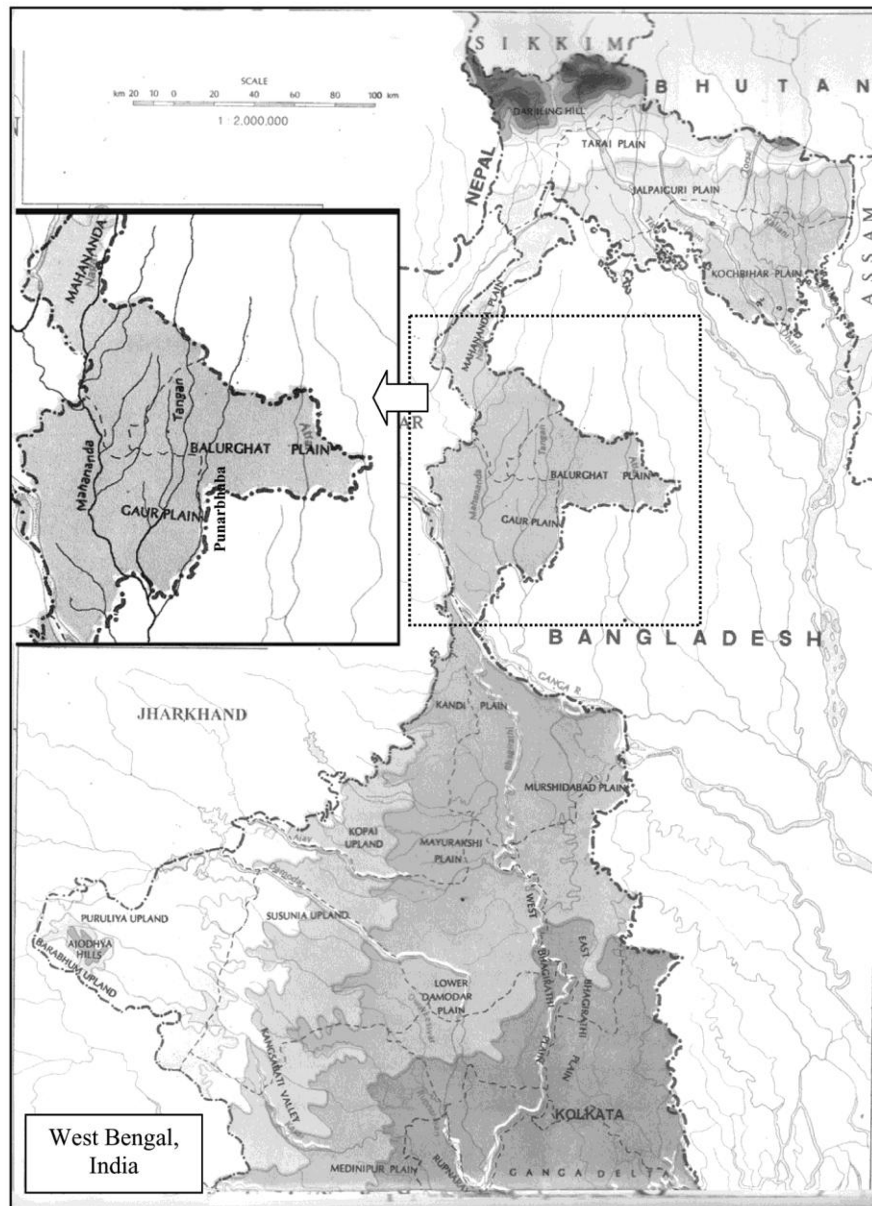


Figure 1. Map showing location of the study site.

and culture in some protected habitat since: (i) the selected fish species have a similar breeding ground, with monsoon being the breeding season; (ii) artificial propagation research efforts and limited culture activity are dependent upon fish stocked from natural collections; (iii) there is extermination of juveniles of numerous non-target fishes due to lack of proper identification and (iv) there exists differential market potential of various fishes.

Even when two species are almost indistinguishable morphologically, they are likely to be easily distinguished genetically with the help of molecular markers. The cytochrome *c* oxidase I (COI) region in the mitochondrial genome was demonstrated as an appropriate type-I molecular marker for discriminating closely related species[4,5] across diverse animal phyla, and this was used for marine and freshwater fishes[6–10]. However, it is inappropriate to assume that molecular markers can provide the final answer to species identification; they are instead an additional marker system[11]. In this context, the present study was directed to formulate precise identification manual for threatened species through introspecting a small segment of the genome, referred to as genetic 'barcodes' and provides information about the COI sequence data of the selected fish species.

The fishes were collected from MTPRS near the Indo-Bangladesh border. The depth of the river varies from a maximum of 3–3.5 m during monsoon, to a minimum of less than 1 m during summer with a silty–clayey

Table 1. DNA barcoded species along with GenBank accession numbers

Order	Family	Genus	Species	Threat status (according to BCPP-CAMP, 1998)	No. of individuals	GenBank accession no.
Osteoglossiformes	Notopteridae	*Chitala*	*chitala*	Endangered	3	JX891536-JX891538
Osteoglossiformes	Notopteridae	*Notopterus*	*notopterus*	Lower risk, near-threatened	2	JX901490-JX901491
Siluriformes	Siluridae	*Ompok*	*bimaculatus*	Endangered	5	JX901502-JX901505
Siluriformes	Schilbeidae	*Eutropiichthys*	*vacha*	Endangered	2	JX901496-JX901497
Siluriformes	Schilbeidae	*Ailia*	*coila*	Vulnerable	4	JX901492-JX901495
Siluriformes	Schilbeidae	*Neotropius*	*atherinoides*	Endangered	2	JX901498-JX901501
Siluriformes	Pangasiidae	*Pangasius*	*pangasius*	Critically endangered	1	JX997836
Cluepeiformes	Clupeidae	*Gudusia*	*chapra*	Endangered	2	KC013543-KC013544

substratum. The basin consists of recent alluvium soil type flanked by arable land, especially graminoid and scrub vegetation, with sporadic rural settlements. Throughout the catchment, the stream waters are heavily used for cultivation of crops (Figure 1).

Sampling was done at 20 sites along MTPRS (24°57′N, 88°20′E) on the Indian side with the help of local fishermen. The specimens were retrieved from the net and identified morphologically to the lowest taxonomic level following standard references[12–14]. All species names adhere to Fishbase[15]. Approximately 100 mg of white muscle tissue from 2 to 5 individuals of each species was preserved in 95% ethanol and brought to the laboratory under chilled condition. The samples were stored at –80°C until further use. Specimen details and GenBank accession numbers are shown in Table 1.

DNA was isolated[16] and its quality was checked in agarose gel. The concentration was estimated using a UV spectrophotometer. The DNA was diluted to a final concentration of 100 ng/μl.

The COI gene was amplified in 50 μl volume with 5 μl of 10× *Taq* polymerase buffer (Sibenzyme), 2 μl of $MgCl_2$ (50 mM), 2 μl of dNTP (10 mM) mixture (Sibenzyme), 0.5 μl of each primer (0.01 mM), 0.5 U of *Taq* polymerase and 100 ng of genomic DNA. The primers used for the amplification of the COI gene were[8,17]: F1 – 5′TCAACCAACCACAAAGACATTGGCAC3′ and R1 – 5′TAGACTTCTGGGTGGCCAAAGAATCA3′.

The thermal regime consisted of an initial denaturation for 2 min at 95°C followed by 35 cycles of 94°C for 40 sec, 54°C for 40 sec, 72°C for 1 min 10 sec, a final chain extension of 10 min at 72°C and held at a temperature of 4°C. The PCR products were visualized on 1.2% agarose gels, and the most intense products were selected for sequencing after quantification through spectrophotometry. The products were purified using Qiagen gel extraction kit. Sequencing reactions were performed bidirectionally using 96 capillary high throughput sequencer ABI 3730 XL to generate sequences with accurate DNA base calling following manufacturer's instructions.

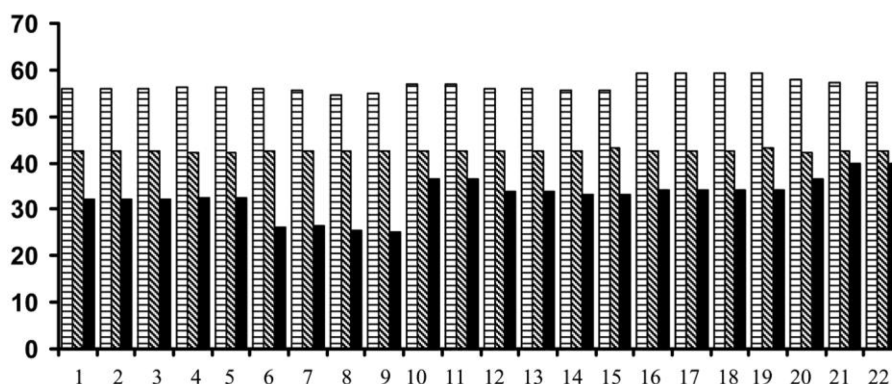
Sequences were aligned using CLUSTAL W[18] and submitted to GenBank (Table 1). The extent of sequence difference between species was calculated by averaging pairwise comparisons across all individuals. The COI sequences of the conspecific individuals were aligned to yield a final sequence of 655 bp. The average nucleotide frequencies along with standard error were calculated. Pairwise evolutionary distance among haplotypes was determined by the Kimura two-parameter method[19] using the software program MEGA (version 4) (Molecular Evolutionary Genetics Analysis)[20]. The mean nucleotide diversity was calculated[21]. The neighbour-joining (NJ) tree was constructed using MEGA 4 and to verify the robustness of the internal nodes of NJ tree, bootstrap analysis was carried out using 1000 pseudo replications.

The results were presented for eight species representing five families and four orders (Table 1). A total of 22 sequences were generated from the eight species using multiple specimens for all species. Sequencing of the COI gene produced 655 nucleotide base pairs per taxon. Sequence analysis revealed average nucleotide frequencies as T = 29.7%, C = 26.4%, A = 26.2%, G = 17.7% (Table 2). All the fish species studied showed relatively higher GC (45.1%) content; the highest GC content was found at the first codon (56.4%) position with progressive declining trend towards second (42.7%) and third (32.9%) codon positions (Figure 2). The average transitional ($si = 62$) pairs were more frequent than average transversional ($sv = 47$) pairs, with an average ratio ($R = si/sv$) of 1.3. The mean nucleotide diversity[20] was found to be 0.167. The average K2P distance[18] in percentage within different taxonomic levels was calculated (Table 3). The average genetic distance within species, family and order was found to be 0.43%, 16.8% and 18% respectively. The NJ tree is shown in Figure 3. The details are discussed below.

Five individual fishes belonging to two genera and family Notopteridae under the order Osteoglossiformes were analysed. The sequence analysis revealed average nucleotide frequencies as T = 28.9%, C = 26.4%, A = 27.6%, G = 17.1% for *C. chitala* and T = 29.05%, C = 26.25%, A = 27.2%, G = 17.6% for *N. notopterus* (Table 2). The highest GC content was found at the first codon (56% and 56.4%) position with progressive declining

Table 2. Mean percentage base composition (with SE) comparing cytochrome *c* oxidase 1 (COI) sequence of prioritized fishes

	Average nucleotide composition (%)			
Fish	T	C	A	G
Chitala chitala	28.9 ± 0	26.4 ± 0	27.6 ± 0	17.1 ± 0
Notopterus notopterus	29.05 ± 0.21	26.25 ± 0.21	27.2 ± 0	17.6 ± 0
Ompok bimaculatus	28.65 ± 0.44	27.05 ± 0.5	26.13 ± 0.23	18.15 ± 0.3
Eutropiichthys vacha	29.3 ± 0	26.6 ± 0	25.3 ± 0	18.8 ± 0
Ailia coila	31.4 ± 0.76	24.45 ± 0.52	27.38 ± 0.62	16.75 ± 0.17
Neotropius atherinoides	29.73 ± 0.15	26.75 ± 0.17	26.3 ± 0	17.32 ± 0.05
Pangasius pangasius	29.9 ± 0	26.4 ± 0	24.6 ± 0	19.1 ± 0
Gudusia chapra	30.4 ± 0	27.6 ± 0	23.1 ± 0	18.9 ± 0



Key: 1, *Chitala chitala* 1; 2, *Chitala chitala* 2; 3, *Chitala chitala* 3; 4, *Notopterus notopterus*1; 5, *Notopterus notopterus* 4; 6, *Ailia coila* 1; 7, *Ailia coila* 2; 8, *Ailia coila* 3; 9, *Ailia coila* 5; 10, *Eutrophichthys vacha* 3; 11, *Eutrophichthys vacha* 4; 12, *Neotropius atherinoides* 2; 13, *Neotropius atherinoides* 3; 14, *Neotropius atherinoides* 1; 15, *Neotropius atherinoides* 4; 16, *Ompok bimaculatus* 2; 17, *Ompok bimaculatus* 3; 18, *Ompok bimaculatus* 4; 19, *Ompok bimaculatus* 5; 20, *Pangasius pangasius*; 21, *Gudusia chapra* 1; 22, *Gudusia chapra* 2.

Figure 2. Variation in GC content for cytochrome *c* oxidase 1 (COI) gene sequence at first, second and third codon positions respectively, among the fishes under study.

trend towards second (42.7% and 42.2%) and third (32% and 32.4%) codon positions (Figure 2) for *C. chitala* and *N. notopterus* respectively. The average transitional ($si = 28$) pairs were more frequent than average transversional ($sv = 11$) pairs, with an average ratio ($R = si/sv$) of 2.6. The average genetic distance within species was 0% whereas the average genetic distance between species was 11%. The NJ tree revealed distinct clusters shared by species of the same genera (Figure 3). All assemblages of conspecific individuals had 100% bootstrap values and the congeneric species formed the same clade.

Ten individual fishes belonging to three genera and family Schilbeidae under the order Siluriformes were analysed. The sequence analysis revealed average nucleotide frequencies as T = 31.4%, C = 24.45%, A = 27.38%, G = 16.75% for *A. coila*, T = 29.73%, C = 26.75%, A = 26.3%, G = 17.32% for *N. atherinoides* and T = 29.3%, C = 26.6%, A = 25.3%, G = 18.8% for *E. vacha* (Table 2). The highest GC content was found at the first codon (55.28%, 56.9% and 55.8%) position with progressive declining trend towards the second (42.6%, 42.6% and 42.73%) and third (25.83%, 36.5% and 33.55%) codon positions (Figure 2) for *A. coila*, *E. vacha* and *N. atherinoides* respectively. The average transitional ($si = 53$) pairs were more frequent than average transversional ($sv = 25$) pairs, with an average ratio ($R = si/sv$) of 2.1. The average genetic distance within species was 1.33%, whereas the average genetic distance between species was 19.3%. The NJ tree revealed distinct clusters shared by the species of the same family (Figure 3). All assemblages of conspecific individuals had high bootstrap values and the confamilial species formed the same clade.

Four individual fishes of *O. bimaculatus* belonging to the family Siluridae under the order Siluriformes were analysed. The sequence analysis revealed average nucleotide frequencies as T = 28.65%, C = 27.05%, A = 26.13%, G = 18.15% (Table 2). The highest GC content was found at the first codon (59.2%) position with progressive declining trend towards the second (42.6%) and third (34.3%) codon positions (Figure 2). The average genetic

Table 3. Summary of genetic divergence (K2P percentage) within various taxonomic levels calculated on the basis of COI sequence of prioritized fishes

Comparison within	Minimum	Maximum	Average	Standard error
Species	0.0	2.0	0.43	0.001
Families	11	19	16.8	0.039
Orders	11	21	18.0	0.03

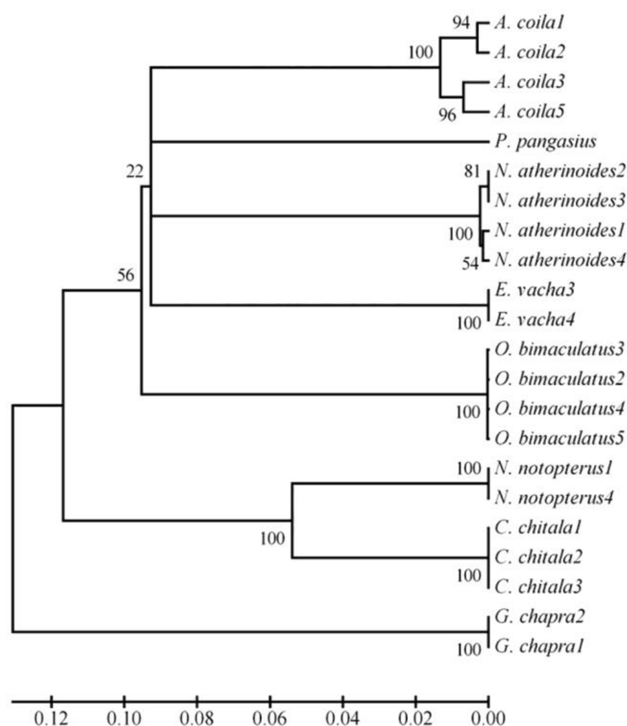


Figure 3. Neighbour-joining tree constructed with the sampled fishes on the basis of COI gene sequence with bootstrap values.

distance within species was 0%. The average genetic distance of individuals belonging to *O. bimaculatus* with other fishes under study (i.e. between species) was 20.3, with value ranging from 17% to 23%. The NJ tree revealed distinct clusters shared by the conspecific individuals having high bootstrap values (Figure 3).

A single individual fish of *P. pangasius* belonging to the family Pangasidae under the order Siluriformes was analysed. The sequence analysis revealed average nucleotide frequencies as T = 29.9%, C = 26.4%, A = 24.6%, G = 19.1% (Table 2). The highest GC content was found at the first codon (57.8%) position with progressive declining trend towards the second (42.2%) and third (36.5%) codon positions (Figure 2). The average genetic distance of individuals belonging to *P. pangasius* with other fishes under study (i.e. between species) was 21.16, with value ranging from 18% to 26%. The NJ tree revealed distinct clusters shared by *P. pangasius* with the members of order Siluriformes having high bootstrap values (Figure 3).

Two individual fishes of *G. chapra* belonging to the family Clupeidae under the order Clupeiformes were analysed. The sequence analysis revealed average nucleotide frequencies as T = 30.4%, C = 27.6%, A = 23.1%, G = 19.1% (Table 2). The highest GC content was found at the first codon (57.4%) position with progressive declining trend towards the second (42.7%) and third (39.8%) codon positions (Figure 2). The average genetic distance within species was 0%. The average genetic distance of individuals belonging to *G. chapra* with other fishes under study (i.e. between species) was 26.14, with value ranging from 25% to 28%. The NJ tree revealed distinct clusters shared by the conspecific individuals having 100% bootstrap value and formed an entirely different clade in the NJ tree (Figure 3) from the members of order Siluriformes and Osteoglossiformes.

In this study, the COI gene of eight species belonging to three orders and five families, collected from MTPRS was sequenced (Table 1). Efficient management of the threatened fish stock should start from unambiguous identification of any life-history stage of these fishes. DNA-based barcode identification system is believed to provide a simple, universal tool for identification of these fishes[16].

The task of barcode generation is of utmost importance, since the task of routine species identification based on morphological and meristic features has four significant limitations: (i) incorrect identification of species due to phenotypic plasticity and genetic variability in the characters involved[22]; (ii) inability to recognize morphologically cryptic taxa[23,24]; (iii) keys designed for a particular life-history stage or gender may not be effective for the others and (iv) misdiagnoses due to lack of expertise and the dwindling pool of taxonomists signalling the need for a new approach to taxon recognition[22].

The universal primers amplified the target region in all the fishes generating barcodes of 655 bp. The sequences were found to be conserved enough to allow amplification by a single set of forward and reverse primers in different fish families, yet diverse enough to permit unambiguous identification of fish species. No insertion, deletion or stop codons were observed in any of the sequences, which confirmed all amplified sequences being functional mitochondrial COI sequence. The 655 bp sequence length suggested that nuclear DNA sequences originating from mitochondrial DNA sequences (NUMTs) were not sequenced as vertebrate NUMTs are typically smaller than 600 bp (ref. 25).

Mitochondrial genomes show profound shifts in nucleotide usage among major taxonomic groups[26] and this

Table 4. Comparison of mutational rate at three different codon positions of COI sequence of prioritized fishes

	Transitional pairs at three codon positions			Transversional pairs at three codon positions		
Comparison within members of	First	Second	Third	First	Second	Third
Family Notopteridae	4 (14.3%)	1 (3.6%)	23 (82.1%)	0	0	11 (100%)
Family Schilbidae	9 (17%)	0	44 (83%)	0	0	25 (100%)
Order Siluriformes	10 (16.9%)	0	49 (83.1%)	0	0	30 (100%)

Figures are shown in absolute number and in percentage in parenthesis.

can have serious impacts on phylogenetic analyses[27–29]. The average GC content was found to be 45.1% (Table 2) in this study and the highest GC content was found at the first codon position with progressive declining trend towards the second and third codon positions (Figure 2). This result corresponds reasonably well with those on complete mitochondrial genomes of *Osteichthyes* species, deriving GC contents of 43.2% (ref. 25) and 47.1% (refs 8 and 16) for the COI gene sequence. The highest GC content (46.7%) was found in *G. chapra*, whereas the lowest (41.2%) was observed in *A. coila*. In a study embracing 16 animal phyla, the GC content was found to vary within a range of 22–53%, with bony fish species *Chirocentrus dorab* (52.8%) and *Salanx ariakensis* (53.4%) showing higher GC content in general among all other species studied[30]. Hence, it can be concluded that fishes in general display higher GC content.

The average transition and transversion ratio was 1.3, showing bias for transitional events. Members of the family Notopteridae, Schilbeidae and order Siluriformes demonstrated highest nucleotide changes occurring at the third codon position followed by the first and the second as evident from Table 4. This bias has been supported by observations in other fishes, like brook charr (*Salvelinus fontinalis*) control region[31], brown trout (*Salmo trutta*)[32], D-loop region of sturgeon[33] and *Poecilia reticulata*[34], as well as in coding regions of fishes[35,36]. No tranversional pairs were found at the first and second codon positions, ensuring minimal changes in the amino acid to be coded taking advantage of wobble effect. This reflects the fact that most synonymous mutations occur at the third position with a few at the first position and rarely at the second position. This observation was also in agreement with the earlier results[8,17].

The mean nucleotide diversity (*Pi*) among all the species was estimated as 0.167. The COI locus harbours a high mutational rate for mtDNA[37]. The average K2P distance of individuals within species was estimated as 0.43%, whereas it was 16.8% at the family level and 18% for orders (Table 3). Hence, there was more than 39-fold sequence difference among confamilial species than conspecific individuals. The variation was even more (41.86 fold) among the individuals belonging to different orders. Hence, mean genetic distance within species was much smaller than the average distance between individuals of different species indicating marked genetic divergence beyond species level that has been referred to as the barcoding gap[38]. All these observations were in agreement with the studies on Australian fish species[8], Alaskan skates[39] and Indian marine fishes[17].

Although barcode analysis was performed to delineate species boundaries only, there was some phylogenetic signal in the COI sequence data[8]. The NJ tree (Figure 3) revealed three distinct clusters formed by members of Siluriformes, Osteoglossiformes and Clupeiformes separately. The conspecific individuals were always found to cluster under the same node supported by high bootstrap value, whereas dissimilar species were clustered under separate nodes, enabling unambiguous identification of species. The nodes were supported by high bootstrap values. However, the true phylogeny of fishes from a 655 bp fragment of mitochondrial DNA through K2P distance and NJ method may not be conclusive enough[8], and hence more gene regions should be used (including nuclear genes) and additional analytic methods deployed, including maximum parsimony and maximum likelihood.

Hence, species-specific signals in the COI sequence should help in accurate and unambiguous identification of various developmental stages of threatened fish species. Our results suggest that COI barcoding can be taken up as a pragmatic approach for resolving unambiguity in the identification of the fish fauna with applications in its management and conservation. One management option could be a collection of early life-history stages of selected fish species from MTPRS, identified through DNA barcodes in monsoon or autumn, when fish abundance is maximum followed by their introduction in some freshwater protected area (FPA). The protected stock within FPA can be utilized for rehabilitation work in degraded aquatic systems and also to demonstrate the benefits of healthier ecosystems to the community and stakeholders. Protected areas could also serve as 'banks' of organisms for the replenishment of unprotected or degraded areas[40]. This could be an invaluable tool for fisheries managers, fisheries ecologists and fish retailers, and for those wanting to develop fish identification microarrays.

1. Mittermeier, R. A. and Mittermeier, C. G., Megadiversity; Earth's biologically wealthiest nation. In *Global Freshwater Biodiversity* (eds McAlister, D. E., Hamilton, A. I. and Harvey, B.), Sea Wind, Cemex, Mexico City, 1997, vol. 11, pp. 1–140.

2. Lakra, W. S., Conservation and sustainable utilization of aquatic genetic resources. Souvenir. In 8th Indian Fisheries Forum, Kolkata, 22–26 November 2008, pp. 93–103.
3. CAMP, Report of the Workshop on Conservation Assessment and Management Plan for Freshwater Fishes of India. Zoo Outreach Organization and NBFGR, Lucknow, 22–26 September 1997, 1998, p. 156.
4. Hebert, P. D. N., Penton, E. H., Burns, J. M., Janzen, D. H. and Hallwachs, W., Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proc. Natl. Acad. Sci. USA*, 2004, **101**, 14812–14817.
5. Hebert, P. D. N., Stoeckle, M. Y., Zemlak, T. S. and Francis, C. M., Identification of birds through DNA barcodes. *PLoS Biol.*, 2004, **2**, e312; doi:10.1371/journal.pbio.0020312.
6. Hajibabaei, M. *et al.*, Critical factors for assembling a high volume of DNA barcodes. *Philos. Trans. R. Soc. London, Ser. B*, 2005, **360**, 1959–1967.
7. Steinke, D., Vences, M., Salzburger, W. and Meyer, A., TaxI: a software tool for DNA barcoding using distance methods. *Philos. Trans. R. Soc. London*, *Ser. B*, 2005, **360**, 1975–1980.
8. Ward, R. D., Zemlac, T. C., Innes, B. H., Last, P. R. and Hebert, P. D. N., DNA barcoding of Australia's fish species. *Philos. Trans. R. Soc. London, Ser. B*, 2005, **360**, 1847–1857.
9. Hubert, N. *et al.*, Identifying Canadian freshwater fishes through DNA barcodes. *PLoS One*, 2008, **3**, e2490; doi: 10.1371/journal.pone.0002490.
10. Lakra, W. S., Goswami, M. and Gopalakrishnan, A., Molecular identification and phylogenetic relationships of seven Indian Sciaenids (Pisces: perciformes, Sciaenidae) based on 16S rRNA and cytochrome oxidase subunit I mitochondrial genes. *Mol. Biol. Rep.*, 2009, **36**, 831–839.
11. Okumus, I. and Ciftci, Y., Fish population genetics and molecular markers: II – molecular markers and their applications in fisheries and aquaculture. *Turk. J. Fish. Aqcult*., 2003, **3**, 51–79.
12. Shaw, G. E. and Shebbeare, E. O. The fishes of North Bengal. *J. R. Asiat. Soc. Bengal*, 1937, **III**, 1–137.
13. Day, F., *The Fishes of India*, *Vols 1 and 2*, William Dawson, London, 1958.
14. Talwar, P. K. and Jhingran, A., *Inland Fishes of India and Adjacent Countries*, Oxford and IBH Publishing Co Pvt Ltd, New Delhi, 1991, two vols, pp. xix + 1158.
15. Froese, R. and Pauly, D. (eds), FishBase. World Wide Web electronic publication, www.fishbase.org, version (04/2013).
16. Ruzzante, D. E., Taggart, C. T., Cook, C. and Goddard, S., Genetic differentiation between inshore and offshore Atlantic cod *Gadus morhua* off Newfoundland: microsatellite DNA variation and antifreeze level. *Can. J. Fish. Aquat. Sci.*, 1996, **53**, 634–645.
17. Lakra, W. S., Sarkar, U. K., Dubey, V. K., Sani, R. and Pandey, A., River inter linking in India: status, issues, prospects and implications on aquatic ecosystems and freshwater fish diversity. *Rev. Fish Biol. Fisheries*, 2011, **21**, 463–479.
18. Thompson, J. D., Higgins, D. G. and Gibson, T. J., CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, 1994, **22**, 4673–4680.
19. Kimura, M., A simple method of estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.*, 1980, **16**, 111–120.
20. Tamura, K., Dudley, J., Nei, M. and Kumar, S., MEGA4: Molecular evolutionary genetics analysis (MEGA) Software version 4.0. *Mol. Biol. Evol.*, 2007, **24**, 1596–1599.
21. Tajima, F., Statistical methods to test for nucleotide mutation hypothesis by DNA polymorphism. *Genetics*, 1989, **123**, 585–595.
22. Hebert, P. D. N., Cywinska, A., Ball, S. L. and Ward, J. R., Biological identifications through DNA barcodes. *Proc. R. Soc. London, Ser. B*, 2003, **270**, 313–322.
23. Knowlton, N., Sibling species in the sea. *Annu. Rev. Ecol. Syst.*, 1993, **24**, 189–216.
24. Jarman, S. N. and Elliott, N. G., DNA evidence for morphological and cryptic Cenozoic speciations in the Anaspididae, 'living fossils' from the Triassic. *J. Evol. Biol*., 2000, **13**, 624–633.
25. Zhang, D. X. and Hewitt, G. M., Nuclear integrations: challenges for mitochondrial DNA markers. *Trends Ecol. Evol.*, 1996, **11**, 247–251.
26. Jermin, L. S., Graur, D., Lowe, R. M. and Crozier, R. H., Analysis of directional mutation pressure and nucleotide content in mitochondrial cytochrome b genes. *J. Mol. Evol*., 1994, **39**, 160–173.
27. Philippe, H. and Laurent, J., How good are deep phylogenetic trees? *Curr. Opin. Genet. Dev.*, 1998, **8**, 616–623.
28. Foster, P. G. and Hickey, D. A., Compositional bias may affect both DNA-based and protein-based phylogenetic reconstructions. *J. Mol. Evol*., 1999, **48**, 284–290.
29. Mooers, A. O. and Holmes, E. C., The evolution of base composition and phylogenetic inference. *Trends Ecol. Evol*., 2000, **15**, 365–369.
30. Clare, E. L., Kerr, K. C. R., von Konigslow, T. E., Wilson, J. J. and Hebert, P. D. N., Diagnosing mitochondrial DNA diversity: applications of a sentinel gene approach. *J. Mol. Evol*., 2008, **66**, 362–367.
31. Bernatchez, L. and Danzmann, R. G., Congruence in control-region sequence and restriction- site variation in mitochondrial DNA of Brook Charr (*Salvelinus fontinalis* Mitchill). *Mol. Biol. Evol*., 1993, **10**, 1002–1014.
32. Bernatchez, L., Gyuomard, R. and Bonhomme, F., Sequence variation of the mitochondrial control region among geographically and morphologically remote European brown trout (*Salmo trutta*, L.) populations. *Mol. Ecol*., 1992, **1**, 161–173.
33. Brown, J. R., Beckenbach, A. T. and Smith, M. J., Intraspecific DNA sequence variation of the mitochondrial control region of white sturgeon (*Acipenser transmontanus*). *Mol. Biol. Evol*., 1993, **10**, 326–341.
34. Fajen, A. and Breden, F., Mitochondrial DNA sequence variation among natural populations of the Trinidad guppy, *Poecilia reticulata*. *Evolution*, 1992, **46**, 1457–1465.
35. Carr, S. M. and Marshall, H. D., Detection of intraspecific DNA sequence variation in the mitochondrial cytochrome *b* gene of Atlantic cod (*Gadus morhua*) by the polymerase chain reaction. *Can. J. Fish. Aquat. Sci.*, 1991, **48**, 48–52.
36. McYeigh, H. P., Bartlett, S. E. and Davidson, W. S., Polymerase chain reaction/ direct sequence analysis of the cytochrome *b* gene in *Salmo salar*. *Aquaculture*, 1991, **95**, 225–233.
37. Saccone, C., De Giorgi, C., Gissi, C., Pesole, G. and Reyes, A., Evolutionary genomics in Metazoa: the mitochondrial DNA as a model system. *Gene*, 1999, **238**, 195–209.
38. Meyer, C. P. and Paulay, G., DNA barcoding: error rates based on comprehensive sampling. *PLoS Biol.*, 2005, **3**, 2229–2238.
39. Spies, I. B., Gaichas, S., Stevenson, D. E., Orr, J. W. and Canino, M. F., DNA based identification of Alaska skates (Amblyraja, Bathyraja and Raja: Rajidae) using cytochrome *c* oxidase submit (COI) variation. *J. Fish Biol.*, 2006, **69**, 283–292.
40. Nevill, J. and Phillips, N., Representative freshwater reserves: better late than never! In Paper presented to the International River Symposium, Brisbane, 29–31 August 2001.

ACKNOWLEDGEMENTS. We thank the Head, Department of Zoology, University of Calcutta, Kolkata for providing infrastructural facilities and the Principal, Bethune College, Kolkata and Dipak Roy (Roy's Fish Farm, Malda) for support.

Received 18 February 2013; revised accepted 29 April 2013