MITOCHONDRIAL CYTOCHROME C OXIDASE SUBUNIT I (COI)-BASED MOLECULAR CHARACTERIZATION OF FRESHWATER TURTLES OF GENUS *PANGSHURA*

Papari Devi[1] and Chittaranjan Baruah*[2,]
[1]Dimension Academy, Tezpur-784001, India
[2]Department of Zoology, Darrang College, Tezpur-784001, India
*Author for Correspondence: chittaranjan_21@yahoo.co.in

Abstract
The present study has been performed for sequencing and analysis, tertiary structure prediction along with evolutionary characterization of Cytochrome C Oxidase Subunit I (COI) from the highly endangered Southeast Asian turtle species under the genus Pangshura and compared with the sequence data of Batagur. The study represents the application of comparative modelling method for protein 3D structure prediction. The evolutionary analyses showed that the P. smithii and P. tentoria are the sister species followed by P.tecta and P. sylhetensis as their successive sister-taxa. The overall Quality factors for the predicted 3D structures are more than 95% as predicted by ERRAT verification programme. After successful verification of the coordinate files, the structures were deposited to PMDB Protein Model Database of University of Rome and available for further uses. The present study also provides an indispensable groundwork for future conservation and genetic analysis of other turtle genera.

Keywords: Chelonian, Conservation, Geoemydidae, Homology modeling, in-silico

INTRODUCTION

The turtle family Geoemydidae, mostly the freshwater turtles comprises of several highly endangered Southeast Asian turtle species (Van Dijk *et al.,* 2000; IUCN, 2016). This family occupies wide range of habitats from highly aquatic (*Pangshura*) to terrestrial (*Geoemyda*). The genus *Pangshura* having maximum shell length of 20–26.5 cm (Das *et al.,* 2010), comparatively smaller than *Batagur* (maximum shell lengths 48–58 cm) is characterized by sexual dimorphism (Ernst *et al.,* 2000; Sarma, 2007), now placed in two distinct genera (Das, 2001). It is quite relevant that molecular understanding at species and sub-species level of the *Pangshura* group in particular and the turtle group in general to formulate conservation strategies for many of the endangered turtle group.

The present study has been performed for sequencing of mitochondrial COI gene, analysis, tertiary structure prediction and evolutionary characterization of Cytochrome c oxidase subunit-I proteins from the highly endangered Southeast Asian turtle species under the genus *Pangshura* and its closely related genus *Batagur* (Testudines: Geoemydidae: *Pangshura* and *Batagur*).

Although there are translated sequence information in the Protein Knowledgebase (UniProtKB/Swiss-Prot) from the genus *Pangshura,* yet structural, functional and evolutionary information are still lacking and the molecular mechanism of their functions are yet to be fully understood. The present study has been conducted to identify the structural and

evolutionary properties of Cytochrome c oxidase subunit-I and to test whether the sequence polymorphism and evolutionary profile of COI within the genus *Pangshura* correspond well with distinct clades. Understanding of differentiation of these turtles at molecular level would significantly contribute to develop more powerful conservation formulation. The present analysis was performed using translated coding sequences generated from first hand sequencing data along with the sequence data from Protein Knowledgebase (UniProtKB) (Table 1).

Table 1. Currently recognized species and subspecies of *Pangshura and Batagur* analyzed in the present study (According to Fritz and Havas, 2007)

Genus	*Species*	*Translated* GenBank *sequence code/ UniProtKB Accession Number*
Pangshura Gray1856	*P. tecta* (Gray 1831)	E5L5T5
	P. sylhetensis (Jerdon 1870)	P12
	P. smithii (Gray 1863)	E5L5T4
	P. tentoria (Gray 1834)	E5L5T6
Batagur Gray 1856	*B. dhongoka* (Gray1835)	E5L5R3
	B. kachuga (Gray 1831)	E5L5R4
	B. baska (Gray 1831)	E5L5R1
Out group	-	*Cyclemys dantata* (E5L5R6)

MATERIALS AND METHODS

PCR amplification and sequencing

PCR amplification of the proteins coding COI gene was performed using primer pairs as listed in Table 2. The primers were synthesized (Le *et al.*, 2007; Spinks and Shaffer, 2007) and the PCRs were carried out in 25 µL of reaction volume containing, with 1.5 U *Taq polymerase* (Bangalore Genei, Bangalore, India), 0.25 mM of dNTP's (Bangalore Genei), 2.0 mM of MgCl2, 1 µL of 0.5 mg/ml of BSA, 0.1 M (Sigma) of each primer and 20 ng of genomic DNA. The condition for amplification was an initial denaturation temperature 94 $^{\circ}$C for five mins., followed by 35 cycles of 45 seconds at 94 $^{\circ}$C, then by 45 seconds at appropriate annealing temperature (Table 2) followed by 2 min at 72 $^{\circ}$C and lastly by a final extension step for 7 min at 72 $^{\circ}$C. PCR products were purified using QIAquick PCR Purification kit (Qiagen) and sequences were obtained commercially from MWG Biotech Pvt. Ltd. (Bangalore).

Table 2. Primers used in this study

Primer name	Primer sequence ($5^{/}$–$3^{/}$)	Gene name	Temperature	Reference
L-turtCOIc	5'-TACCTGTGATTTTAACCCGTTGAT-3'	COI	59°C	Spinks and Shaffer, 2007
H-turtCOIc	5'-TGGTGGGCTCATACAATAAAGC-3'	COI	59 $^{\circ}$C	Spinks and Shaffer, 2007

Acquisition and alignment of sequences

The coding region of COI gene sequences and available GenBank sequences for closely related genus *Batagur* are included in the present study. The sequences were translated into amino acids prior to analysis using EMBOSS Transeq programme (Mackey *et al.,* 2002) and they could not show any stop codons suggesting that all are functional copies. Further, the analysis included the sequence information of the out group sequences for respective protein for molecular evolutionary studies, extracted from Protein Knowledgebase (UniProtKB) which were acquired both by database keyword search and by BLASTp (Altschul *et al.,* 1997) and FASTA (Pearson, 1991) searches (Table 1). The sequences were simultaneously aligned using ClustalW (Higgins *et al.,* 1994) and Modeller (Fiser *et al.,* 2000) programs.

Data mining and Sequence analysis

The study was extended to data mining and sequence analysis of Cytochrome c oxidase subunit-I from the genus *Pangshura* and its closely related *Batagur*. The statistical analyses were performed using the CLC Genomics Workbench 4.0 (CLC Bio, Hyderabad, India) and ProtParam (Gasteiger *et al.,* 2005). The important calculations for the amino acid composition, atomic composition, theoretical pI, molecular weight, formula, Extinction coefficients, half-life, Instability index, aliphatic index, hydrophobicity, charge vs pH, proteolytic cleavage pattern were carried out under sequence analysis.

Three-dimensional structure prediction

Comparative (Homology) modelling was conducted by using Modeller9v2 program (Marti-Renom *et al.*, *2000*). The final 3D structures of each protein were evaluated (Giorgetti *et al.*, *2005*) by ERRAT (Colovos and Yeates, 1993) and ProCheck (Laskowski *et al.,* 2003). All the graphic presentations of the 3D structure were prepared using Chimera (Pittersen *et al.*, *2004*).

Functional annotation

Preliminary investigations of function of the modelled proteins were performed from the 3D structure using ProFunc (Laskowski *et al.,* 2005). A number of databases like Pfam, PROSITE, PRINTS, ProDom, InterProScan (Zdobnov and Apweiler, 2001) were used for functional characterization.

Molecular evolutionary analysis

Evolutionary analyses were conducted in MEGA5 (Tamura *et al.,* 2011). The evolutionary history was inferred by using Maximum Likelihood (Jones *et al.,* 1992) and Neighbor-Joining (Saitou and Nei, 1987) methods. The tree is predicted to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches (Felsenstein, 1985).

RESULTS

Data mining and Sequence analysis *of Cytochrome c oxidase subunit I*

The Cytochrome c oxidase subunit I sequence analyzed in the present study has 216 amino acid residues for all the *Pangshura* species. However, The Cytochrome c oxidase subunits I of *P. tecta* and *P. sylhetensis* have a molecular weight of 23.014 kDa where as *P. smithii* and *P. tentoria* have 23.028 kDa. The aliphatic index is 120.602 in *P. tecta* and *P. sylhetensis* and 121.065 in *P. smithii* and *P. tentoria.* (Table 3, Figures 1-4).

Table 3. Protein statistics of Cytochrome c oxidase subunit I

Protein statistics	*P. tecta*	*P. sylhetensis*	*P. smithii*	*P. tentoria*
Number of amino acids	216	216	216	216
Molecular weight	23.014 kDa	23.014 kDa	23.028 kDa	23.028 kDa
Isoelectric point	5.26	5.26	5.26	5.26
Total number of negatively charged residues (Asp + Glu)	9	9	9	9
Total number of positively charged residues (Arg + Lys)	4	4	4	4
Aliphatic index	120.602	120.602	121.065	121.065
Total number of Hydrophobic residues (A,F,G,I,L,M,P,V,W)	150	150	150	150

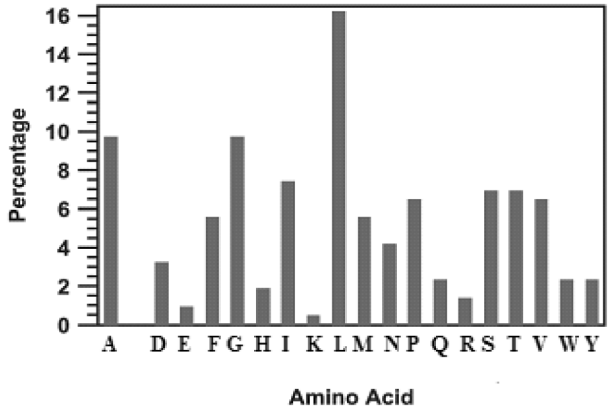


Figure 1. Distribution of amino acids in Cytochrome c oxidase subunit I in *Pangshura* (*P. sylhetensis*)

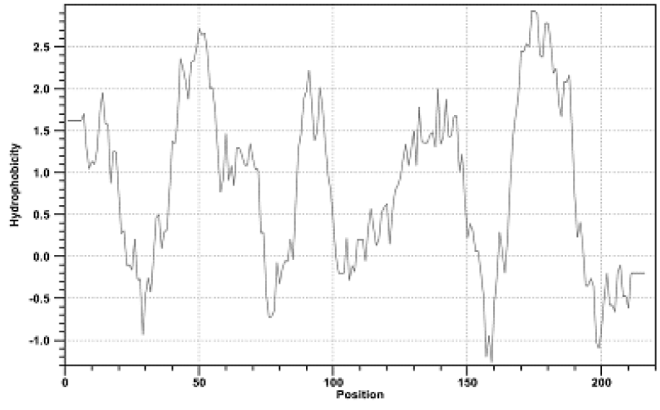


Figure 2. Plot of local Hydropathy for Cytochrome c oxidase subunit I of *Pangshura* (*P. sylhetensis*)

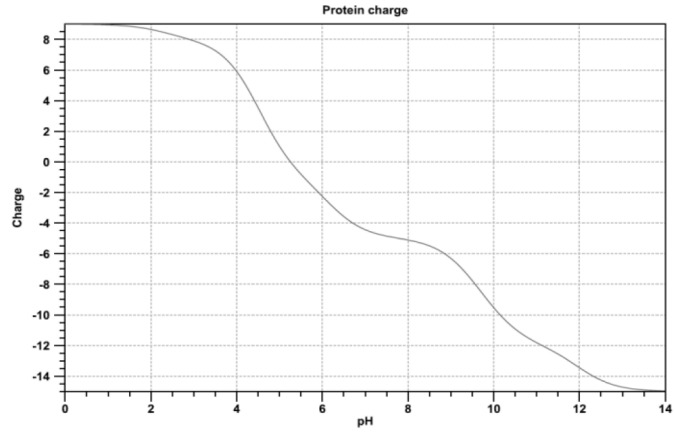


Figure 3**.** Electrical charge as a function of pH for Cytochrome c oxidase subunit I of (*P. sylhetensis*)

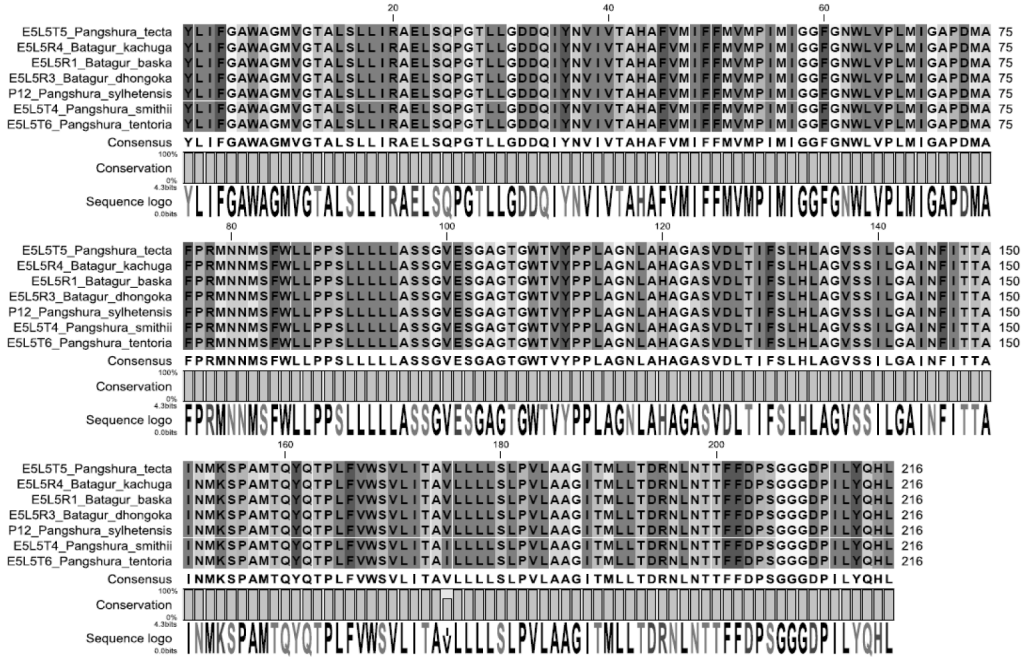


Figure 4. Sequence alignment Cytochrome c oxidase subunit I between the genera *Pangshura* and *Batagur*. A variation observed at position 175.

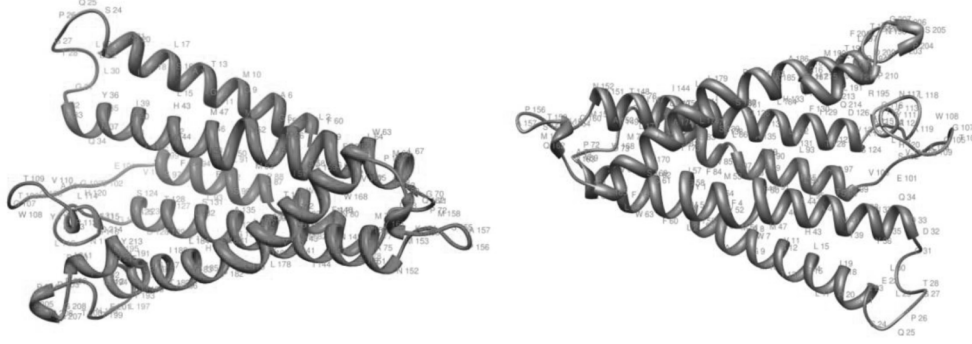
Protein tertiary structures

A. The predicted 3D Structure of *Cytochrome c oxidase subunit-I*

This model is based upon the coordinates of PDB entry 1v54.pdb chain A (oxidoreductase). The motif analysis by ProMotif revealed that the 3D structure Cytochrome c oxidase subunit I has 1 sheet, 1 beta hairpin, 2 strands, 11 helices, 20 helix-helix interacs and 8 beta turns both in *P. sylhetensis* and *P. tecta.* However, *P. tentoria* and *P. smithii* has 11 helices, 20 helix-helix interacs, 7 beta turns. Cytochrome c oxidase subunit I in *Batagur kachuga* and *B. dhongoka* has 1 sheet, 1 beta hairpin, 2 strands, 11 helices, 20 helix- helix interacs and 7 beta turns (Figures 5 and 6).

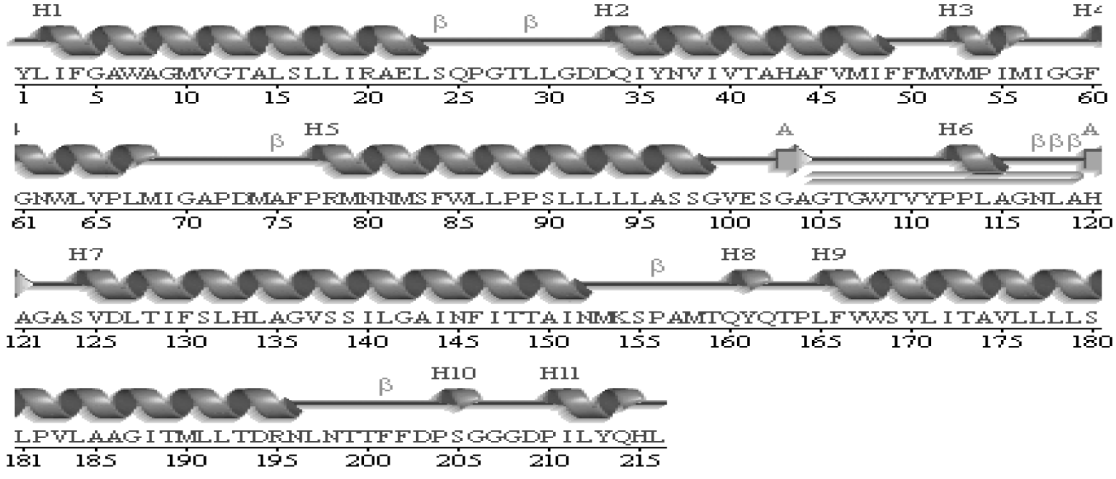
Table 4. The PMDB Protein Model DataBase ID assigned to the submitted structures of Cytochrome c oxidase subunit I

Taxon	Protein name with PMDB Protein Model DataBase ID
P. tecta	PM0077614
P. sylhetensis	PM0077616
P. smithii	PM0077615
P. tentoria	PM0077613
B. dhongoka	PM0077634
B. kachuga	PM0077635
B. baska	PM0077633



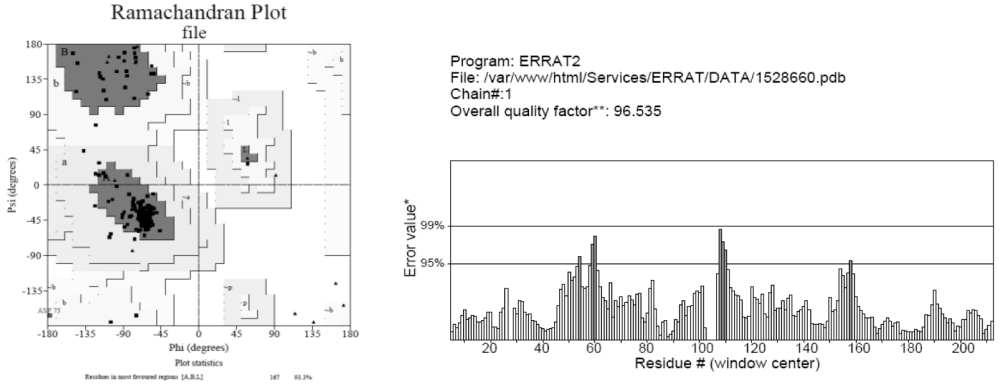
A. B.

Figure 5. 3D structure of Cytochrome c oxidase subunit I (A. *Pangshura*, B. *Batagur*)



Key: Sec. struc:		Helices labelled H1, H2, ... and strands by their sheets A, B, ...
	Helix Strand; Motifs: β beta turn beta hairpin	

Figure 6. Secondary structure assignment of the predicted three-dimensional model of Cytochrome c oxidase subunit I (*P. sylhetensis*)**.**



A. B.

Figure 7. 3D structure verification of Cytochrome c oxidase. A. Ramachandran analysis of the backbone dihedral angles PSI (Ψ) and PHI (ϕ) for the final structure ; B. ERRAT verification for the overall quality factors of the predicted 3D structure

Verification of protein structures

Procheck verification proved that the models are of good quality as judged by Ramachandran Plot (Figure 7A; Ramachandran and Sasisekharan, 1968). The overall quality factors predicted by ERRAT verification programme for the predicted 3D structures are around 95% (Figure 7B). After fruitful verification of the coordinate files, the structures were successfully deposited to PMDB Protein Model Database (Tiziana *et al.,* 2006) of University of Rome and now available for download. Each 3D structure has been assigned a unique PMDB ID for the coordinate entry (http://mi.caspur.it/PMDB) (Table 4).

Functional annotation

The preliminary functional annotation results are listed in the Table 5.

Table 5. Summary of predicted function with ProFunc score (shown within parenthesis)

Name of the Protein	Protein name terms	Gene Ontology (GO) terms		
		Cellular component	*Biological process*	*Biochemical function*
Cytochrome c oxidase subunit I	oxidase (21.73) cytochrome (21.21) cytochrome oxidase (20.42) fragment (9.63) reduced state (2.37)	integral to membrane (0.80) cell (0.80) cell part (0.80) membrane (0.80)	aerobic respiration (0.80) oxidation reduction (0.80) cellular metabolic process (0.80)	cytochrome\-c oxidase activity (0.80) iron ion binding (0.80) electron carrier activity (0.80) heme binding (0.80)

Molecular evolution in Pangshura and Batagur

The evolutionary tree of COI protein supports the fact that *P.tecta*, *P. smithii* and *P. tentoria* are sister groups, while *P. sylhetensis* is their successive sister taxon. *Batagur dhongoka* is found to be an intermediate species of the two genera that belongs to genus *Batagur*.

Table 6. Maximum-likelihood model parameters for data sets as estimated in Modeltest (Posada and Crandall, 1998)

Parameter	Data set of Cytochrome c oxidase subunit I
Amino acids	216
Model	MtREV24
Parameter	13
Bayesian Information Criterion (BIC) scores	1354.4
Akaike Information Criterion, corrected (AICc) value	1283.7
Maximum Likelihood value (*lnL*)	-628.75
Gamma distribution (*G*)	n/a
invariable (*I*)	n/a
Total positions in the final dataset	216

I. Evolution of Cytochrome c oxidase subunit I in Pangshura and Batagur:

The analysis involved 8 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 216 positions in the final dataset. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches (Felsenstein, 1985) (Figure 8).

A. The evolutionary history was inferred by using the Maximum Likelihood method based on the General Reversible Mitochondrial model (Adachi and Hasegawa, 1996) as stated above (Table 6). The evolutionary rate differences among sites [5 categories (+*G*, parameter = 38.9859)]. The rate variation model allowed for some sites to be evolutionarily invariable [(+*I*), 0.0000% sites].

B. The evolutionary history was inferred using the Neighbor-Joining method and the optimal tree is with the sum of branch length = 0.02777778.

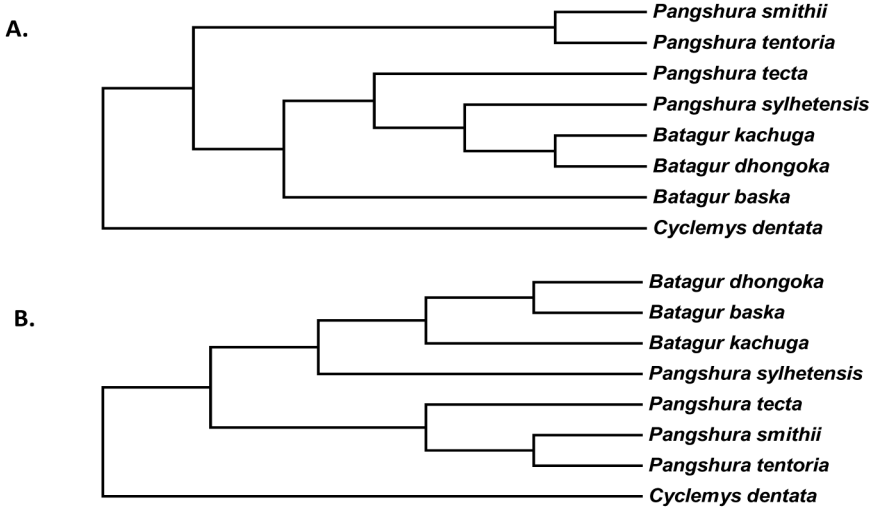


Figure 8 (A-B). Evolutionary relationship of Cytochrome c oxidase subunit I among the species of *Pangshura* and *Batagur.*

All the results presented above strongly suggests that *P. smithii* and *P. tentoria* as sister species followed by *P. tecta* and *P. sylhetensis* as their successive sister-taxa. *Batagur* is a distinctly separate genus from *Pangshura.* Sequence analysis revealed highly identical and conserved regions among *P. sylhetensis* and *P.tecta.* Similarly, *P. tentoria* and *P. smithii* shares high sequence identity. The molecular evolutionary analysis results for all the proteins also strongly supports the fact that *P. smithii* and *P. tentoria* as sister species, followed by *P. tecta* and *P. sylhetensis* as their successive sister-taxa.

DISCUSSION

The results are highly consistent with the earlier molecular genetics study involving 12S rRNA genes (Shaffer *et al.,* 1997) while compared all species of *Batagur* and *Pangshura,* with well-supported clades. The 3D structures generated under the present study could be helpful in structural biology for further investigations on allocation of amino acid residues in each fold, prediction of active sites, molecular mechanism of function and structure based phylogeny. The structures of all the proteins were found to be statistically significant by the structure verification programs. The present analysis corroborate that the genus *Pangshura* is monophyletic. The modeling of proteins from the genus *Pangshura* gains importance for the structural biology and even to the conservation genetic research from several angles.

Cytochrome c oxidase subunit I of *P. tentoria* and *P. smithii* contains 11 helices and 7 beta turns while the *P. tecta* and *P. sylhetensis* have 11 helices and 8 beta turns. The tertiary structure of the proteins from the present study clarify that within the genus *Pangshura*, The species *P. tecta* and *P. sylhetensis* are most closely related followed by *P. smithii* and *P. tentoria* as their successive sister taxa.

The computational models generated from the present study have been deposited to the Protein Model Database (PMDB) and each of them has been assigned a unique PMDB ID. The models can be downloaded for further understanding of various functional and evolutionary aspects.

The molecular evolutionary analysis of Cytochrome c oxidase-I provides a stable phylogenetic hypothesis for all *Pangshura* species, with the suggestion that *P.smithii* and *P.tentoria* as sister species followed by *P. tecta* and *P. sylhetensis* as their successive sister-taxa. The present study supports that *Batagur* is an outgroup and a distinctly separate genus from *Pangshura.* The present study provides an indispensable groundwork for future molecular analysis at the protein level. The choice of molecular data is crucial for phylogenetic analysis and molecular studies can now be tailored specifically for particular phylogenetic groups and or questions (Lamb and Lydeard, 1994).

Using diverse molecular markers, the present study has been able to generate a robust phylogeny with high statistical support values for all nodes, regardless of analysis methods employed, except for the uncertainty in the relationship between the two sub-species of *P. smithii*. The present analysis confirm that the sister relationship between *Batagur* and *Pangshura* although weakly supported in Diesmos *et al.* (2005).The results are highly

consistent with earlier molecular genetics study involving 12S rRNA genes (Shaffer *et al.*, 1997) that within *Batagur* and *Pangshura,* all species correspond with well-supported clades.

The structures of all the proteins were found to be statistically significant by the structure verification programs. The present analysis corroborate that the genus *Pangshura* is monophyletic. The modeling of proteins from the genus *Pangshura* gains importance for the structural biology and even to the conservation genetic research from several angles.

CONCLUSION

The present study can be used as an additional method for identification of species as well as for identification of unknown samples with unusual appearances and could be made available for the identification of confiscated specimens. The predicted 3D structures presented here can serve as a guide for the allocation of amino acid residues involved in each fold, which is important for further investigations on molecular mechanism of functions. The molecular evolutionary analysis underline that further sampling is in dire need for developing effective conservation strategies. *Pangshura* represents distinct genera with four well supported species. Much is still to be learned about how the protein can manipulate a sequence of base pairs in such a peculiar way that results in a fully functional organism.

ACKNOWLEDGEMENTS

The authors are thankful to the Ministry of Environment and Forests, Govt. of India for necessary approval letter and the state forest departments of Assam, West Bengal and Uttar Pradesh for necessary permission to carry out the research.

REFERENCES

Adachi J and Hasegawa M (1996): Model of amino acid substitution in proteins encoded by mitochondrial DNA. *J. Molecular Evolution 42*: 459-468.

Altschul S F, Madden T *et al*. (1997): Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research* 25(17*)*: 3389-3402.

Colovos C and Yeates T O (1993): Verification of protein structures: Patterns of nonbounded atomic interections. *Protein Science* 2:1511-1519.

Das I (2001): Die Schildkröten des Indischen Subkontinents. *Frankfurt am Main: Chimaira* pp.181

Das I, Sengupta S, Praschag P (2010): *Pangshura sylhetensis* Jerdon 1870 – Assam roofed turtle. In: Rhodin AGJ, Pritchard PCH, van Dijk PP, Saumure RA, Buhlmann KA, Iverson JB, and Mittermeier RA (Eds). Conservation Biology of Freshwater Turtles and Tortoises: A Compilation Project of the IUCN/ SSC Tortoise and Freshwater Turtle Specialist Group. *Chelonian Research Monographs* 5: 046.1–046.6.

Diesmos A C, Parham J F, Stuart B L, Brown R (2005): The phylogenetic position of the recently rediscovered Philippine forest turtle (Bataguridae: Heosemysleytensis). *Proceedings of the California Academy of Sciences.* 56(3): 31-41.

Ernst CH, Altenburg RGM, Barbour RW (2000): Turtles of the World. World Biodiversity Database, CD-ROM Series, Windows, Version 1.2. Amsterdam: *Biodiversity Center* of ETI.

Felsenstein J (1985): Confidence limits on phytogenies: An approach using the bootstrap. *Evolution* 39: 783-791.

Fiser A, Do RK, Sali A (2000): Modeling of loops in protein structures, *Protein Science* 9: 1753-1773.

Fritz U and Havas P (2007): In. *Checklist of Chelonians of the World*. Bonn and Dresden: German Federal Ministry of Environment, Nature Conservation and Nuclear Safety and Museum of Zoology Dresden.

Gasteiger E, Hoogland C *et al*. (2005): *Protein Identification and Analysis Tools on the ExPASy Server;* In The Proteomics Protocols Handbook, (Eds John M. Walker) Humana Press . Pp. 571-607

Giorgetti A, Raimondo D, Miele A E, Tramontano A (2005): Evaluating the usefulness of protein structure models for molecular replacement. *Bioinformatics* 21: 72–76.

Higgins D, Thompson J *et al*. (1994): CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22: 4673-4680.

IUCN (2016): International Union for the Conservation of Nature and Natural Resources (IUCN) Red List of Threatened Species.

Jones DT, Taylor WR, Thornton JM (1992): The rapid generation of mutation data matrices from protein sequences. *Computer Applications in the Biosciences* 8: 275-282.

Lamb T and Lydeard C (1994). A molecular phylogeny of the gopher tortoises, with comments on familial relationships within the Testudinoidea. *Molecular Phylogenetics and Evolution* 3: 283-291.

Laskowski RA, Watson JD, Thornton J M (2003): From protein structure to biochemical function. *Journal of Structural and. Functional Genomics* 4**:** 167–177.

Laskowski RA, Watson JD, Thornton JM (2005): ProFunc: a server for predicting protein function from 3D structure. *Nucleic Acids Res.* 33: W89-W93.

Le M, Mccord WP, Iverson JB (2007): On the paraphyly of the genus *Kachuga* (Testudines: Geoemydidae). *Molecular Phylogenetics and Evolution* 45: 398-404.

Mackay IM, Arden KE, Nitsche A (2002): Real-time PCR in virology. *Nucleic Acids Res*. 30(6): 1292-305.

Marti-Renom M A, Stuart A *et al.* (2000): Comparative protein structure modeling of genes and genomes. *Annual Review of Biophysics and Bimolecular Structure* 29: 291-325.

Pearson WR (1991): Searching protein sequence libraries: comparison of the sensitivity and selectivity of the Smith-Waterman and FASTA algorithms. *Genomics* 11(3): 635-50.

Pittersen EF, Goddard TD *et al*. (2004): UCSF Chimera-a visualization system for exploratory research and analysis. *J.Comput.chem.* 25:1605-1612.

Posada D and Crandall KA (1998): MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14: 817–818.

Ramachandran GN and Sasisekharan V (1968): Conformation of Polypeptides and Proteins. *Advances in Protein Chemistry* 23: 283-437.

Saitou N and Nei M (1987): The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4: 406-425.

Sarma PK (2007): Habitat Ecology, population status and distribution of *Kachuga sylhetensis* (Jerdon) in certain districts of Assam. PhD. Thesis (Unpublished), Gauhati University, pp.137

Shaffer HB, Meylan P, Mcknight M L (1997): Test of turtle phylogeny: molecular, morphological, and paleontological approaches. *Systematic Biology* 46: 235–268.

Spinks PQ and Shaffer HB (2007): Conservation phylogenetics of the Asian box turtles (Geoemydidae, Cuora): mitochondrial introgression, numts, and inferences from multiple nuclear loci. *Conservation Genetics* 8: 641–657.

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. *Molecular Biology and Evolution*, doi:10.1093/molbev/msr121

Tiziana C, Paolo DM, Domenico C, Ivano GT, Anna T (2006): The PMDB Protein Model Database CASPUR, *Nucleic Acids Research*. 34 (Database Issue), D306-D309.

Van Dijk PP, Stuart BL, Rhodin AGJ (2000): Asian Turtle Trade. Proceedings of a Workshop on Conservation and Trade of Freshwater Turtles and Tortoises in Asia, Phnom Penh, Cambodia, 1–4 December 1999. *Chelonian Research Monographs* 2: 1–164.

Zdobnov EM and Apweiler R (2001): InterProScan: an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* 17: 847-848.