Research Article

Molecular characterization of Glutathione S-transferase in two eel species, *Monopterus cuchia* (Hamilton) and *Monopterus albus* (Zuiew)

Papari Devi[1], Chittaranjan Baruah[2*], D.K. Sharma[3]

[1]TCRP Foundation, Guwahati, Assam, INDIA.
[2]Post Graduate Department of Zoology, Darrang College, Tezpur, Assam, INDIA.
[3]School of Biological Science, USTM, Meghalaya, INDIA.

Submission Date: 19-02-2020; Revision Date: 26-03-2020; Accepted Date: 18-04-2020

ABSTRACT

The present investigation has been carried out for DNA isolation, sequencing, sequence analysis, tertiary structure prediction and phylogenetic analysis of Glutathione S-transferase from two freshwater synbranchid eel species of the genus *Monopterus* in different states of Northeast India (Assam and Manipur). Glutathione S-transferase is potentially used as a biochemical marker of aquatic pollution and they are the decisive indicators of the toxic effect. The GST gene sequences of *M. cuchia* is more G: C rich than *M. albus* and the GST gene of *M. albus* is more A: T rich than that of *M. cuchia.* The instability index (II) value of GST enzyme of the present study ranged from 49.26 to 52.67, which classifies the GST protein as unstable. The phylogenetic analysis of GST gene revealed that, *M. cuchia* and *M. albus* formed two separate clades with high bootstap separation 98% and 100% in ML and MP tree respectively clearly indicates the possibility for existence of two sub-species in *M. cuchia*. The MP phylogeny of GST gene also revealed genetic diversity within *M. albus* at 73% bootstrap separation. Some of the samples of *M. albus* and *M. cuchia* in the GST gene phylogenetic tree revealed the possibility of interbreeding or hybridization in the *Monopterus* species. The predicted 3D structure of GST will reveal more insights on GST activity in *M. albus* and *M. cuchia* in order to reveal the detailed detoxification mechanism freshwater eels.

Key words: Glutathione S-transferase, *Monopterus.* Northeast India, Synbranchid.

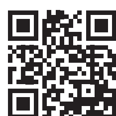
Correspondence:
Chittaranjan Baruah,
Post Graduate Department of Zoology, Darrang College, Tezpur- 784 001, Assam, INDIA.
Phone no: +91-9954294080

Email: chittaranjan_21@yahoo.co.in

INTRODUCTION

The freshwater air-breathing mud eel, *Monopterus cuchia* (Hamilton),[1] locally known as Kuchia (Order- synbranchiformes) often exhibit differences within and among population from different parts of its range (Dahanukar, 2010). *Monopterus albus* (Zuiew),[2] swamp eel, tentatively identified as belonging to the synbranchid genus *Monopterus*[3,4] termed as species complex demands taxonomic revision.[5] Both *M. cuchia* and *M. albus* bear cylindrical body with compressed tail tapering to a point. The minor difference exists, that is *M. cuchia* bears smooth, tiny cycloid and indistinct scales embedded in the skin while *M. albus* is scale less. Therefore, *Monopterus albus* and *Monopterus cuchia* are regarded as species complex and require taxonomic revision.[6] The nuclear gene encoding Glutathione S-transferase, potentially used as biomarkers has been used for molecular characterization of genus *Monopterus*. Glutathione S-transferase is potentially used as a biochemical marker of aquatic pollution and they are the decisive indicators of the toxic effect. [7]

SCAN QR CODE TO VIEW ONLINE



www.ajbls.com

DOI : 10.5530/ajbls.2020.9.6

Glutathione-S-transferases (GSTs) E.C.2.5.1.18) are a family of intracellular enzymes which play an important role in detoxification and elimination of xenobiotics.[8] GSTs catalyse the conjugation of tripeptide glutathione (GSH) with some endogenous toxic metabolites and many environmental contaminants.[9] GSTs also play an important role in phase II detoxification of lipid peroxides and demonstrate the functions such as

glutathione peroxidase activity towards reactive oxygen species in the cells under oxidative stress [10]. Therefore, in fish, GST is used primarily as a biomarker indicating aquatic environment pollution with wastewater of municipal, industrial, agricultural or mining origin.

Although, there has been availability of sequence information for Glutathione-S-transferase (GST) of different fish groups, yet species-specific structural information are lacking. Therefore, the biochemistry and molecular mechanism of their functions in fishes are still not very well understood due to lack of their structural information. Therefore, an attempt has been made for genomic DNA isolation, PCR amplification, sequencing and analysis of Glutathione-S-transferase (GST) of *Monopterus* along with sequence analysis, tertiary structure prediction [11] and phylogenetic profile of available database sequences for *M. cuchia* and *M. albus*.

MATERIALS AND METHODS

Sample collection

Field work was carried out during January, 2016 to June, 2018 in different parts of Assam and Manipur in Northeast India in certain suitable habitats like paddy-field and swamps in order to collect the *M. cuchia* and *M. albus* samples (Table 1; Figure 1). A total of 230 *Monopterus* individuals were sampled from 3 water bodies of Manipur and 18 water bodies of Assam with varying geographical co-ordinates. *Monopterus* individuals were morphologically identified, based on twenty three (23) reliably measurable morphometric characters [12]. The individuals were geographically allocated into four populations based on the proximity of the water bodies i.e. each *Monopterus* population was sampled at gap of about 100-400 km away from any other population.

DNA extraction, PCR amplification and sequencing GST gene

From each captured specimen, approximately 1 cm of tail tissue was removed with forceps and was placed it in a sterile 1.5 ml micro tube containing 95% ethanol and was stored at –20°C until processing. The eels were released at their points of capture. The samples were placed in an ice chest during transport to the laboratory. Genomic DNA was isolated from the tissue using the Chloroform-Octanol method [13,14].

Selected samples (6 each from *M. cuchia* and *M. albus*) were used for sequencing of targeted regions of GST gene. The informative regions of GST gene was PCR amplified from multiple individuals per species using taxa spe-

Table 1: GPS coordinates of the sampling sites of *Monopterus cuchia* and *Monopterus albus*.

Population Name/ Sampling area	Population Code	Sampling Site(s)	Coordinates
Manipur	1	Loktak Lake	24.3°N, 93.5°E
		Ijei River, Longmai, Tamenglong	24.5°N, 93.4°E
		Tuivel River, Manipur	24.08°N , 93.2°E
Assam	2	Wetlands, Hallawgaon, Sadiya, Tinsukia	27.50°N, 95.45°E
		Rice field, Mariani, Jorhat	27.52°N, 95.37°E
		Bhogdoi River, Jorhat	26.77°N, 94.22°E
		Kakodonga River, Golaghat	26.43°N, 94.3 °E
		Dhansiri River, Golaghat	26.35°N, 93.35°E
	3	Kolong River, Nagaon	26.36°N, 92.69°E
		Jia Bhorali River, Sonitpur	26.69°N, 92.87°E
		Bishwanath Ghat, Sonitpur	26.66°N, 93.17°E
	4	Chandubi beel, Kamrup	25.51°N; 91.21° E
		Kulshi river, Kamrup	26.03°N; 91.26° E
		Rice field, Nalbari	26.14°N, 91.08°E
		Puthimari river, Kamrup	26.22°N, 91.4°E
		Rice field, Hajo, Kamrup	26.14°N, 91.32°E
		Wetland, Manikpur, Bangaigaon	26.45°N, 90.80°E
		Rice field, Bilaishipara, Dhubri	26.11°N, 90.16°E
		Rice field, Dhubri	26.01°N, 89.6°E
		Urpad Beel, Goalpara	26.06°N, 90.35°E
		Rice field, Dudhnoi, Goalpara	25.98°N, 90.79°E

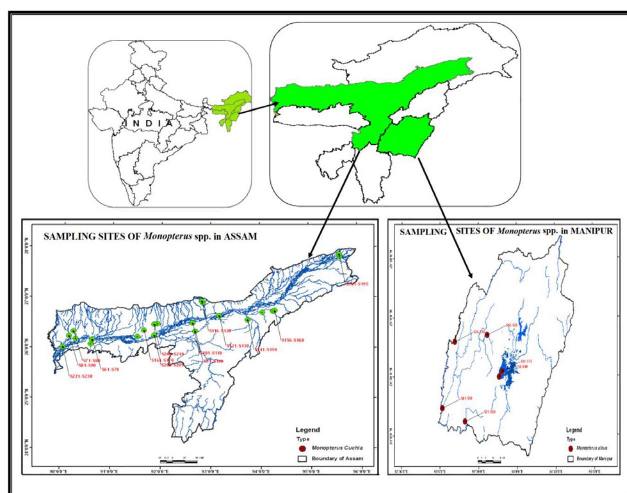


Figure 1: Map of study area showing the sampling sites.

cific primers. The GST gene was PCR amplified using primer pairs GST1-F-GATGTGTTCCTGCCTTCTC-CACA and GST1-R-GAAGACCAATCACCA-CAAAGGGAA (Fu and Xie, 2006). PCR amplifications were carried out in 25 µL reaction volume, with 1.5 units of Taq DNA Polymerase (Bangalore Genei, Bangalore, India), 0.25 mM of dNTPs (Bangalore Genei), 2.0 mM of $MgCl_2$, 0.1 M (Sigma) of each primer and 20 ng of genomic DNA. The condition for amplification was an initial denaturation temperature 94°C for 5 min, followed by 35 cycles of 94°C denaturing temperature for 50 sec, then by 45 sec at 55°C annealing temperature followed by extension temperature of 72°C for 90 sec and then by a final extension at 72°C for 10 min.

Agarose gel electrophoresis of amplified PCR products

The amplified products were separated in 2% agarose gel (i.e. 2g Agarose in 100 ml 1X TBE or 0.6 g agarose in 30 ml 1x TBE) by electrophoresis. Low Range DNA Ruler or 1 kb ladder was used as MW markers. 4 µl PCR products with 1µl of 6X Loading dye were loaded in each of the wells. Electrophoresis was performed at 90 V till the tracking dye; bromophenol blue reached the anodal end of the gel. The results of electrophoresis was observed and recorded in a gel documentation system (UVIdoc). The amplified PCR products were purified using QIAquick PCR Purification kit (Qiagen).

Sequencing GST gene

Sequencing of GST gene were performed by ABI PRISM® 377 DNA sequence (at BioAxis DNA Research Centre, Hyderabad). After verification the nucleotide sequences were deposited to GenBank (NCBI) public sequence repository [15].

Acquisition of additional nucleotide sequences Translation of coding region

The study also included the sequences of GST extracted from GenBank (NCBI)[15] by database keyword search and by BLASTp[16] and FASTA[17] searches. The existing GenBank sequences of closely related species was used as outgroup of phylogenetic analysis. The coding regions for GST gene were translated into amino acids of Glutathione-s-tranferese proteins (using EMBOSS Transeq; [18] for structural and evolutionary studies. Apart from these, amino acid sequences of Glutathione-s-tranferese were also extracted from Protein Knowledgebase (UniProtKB, [19] both by database keyword search and by BLASTp [16] and FASTA [17] searches.

Sequence Alignment and Comparative sequence analysis

Data mining and sequence analyses GST gene along with Glutathione-S-tranferese protein were performed using the CLC Genomics Workbench 7.0.3 (CLC Bio, Hyderabad, India). The nucleotide and protein sequences were simultaneously aligned using CLUSTAL-W[20] and Modeller version 9.12 [21] programs.

The physico-chemical parameters of Glutathione-S-tranferese proteins were computed using CLC Genomics Workbench 7.0.3 (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 were carried out under sequence analysis.

Molecular Phylogenetic analysis

The sequences for the GST gene were aligned using ClustalW 1.6[22] integrated in software MEGA.[23] The evolutionary history was inferred by using two different methods namely the Maximum Parsimony[24] and Maximum Likelihood[25] methods. Maximum parsimony (MP) tree was estimated using the Close-Neighbor-Interchange algorithm.[26] The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates).[27] The branch lengths calculated using the average pathway method.[26]

Three-dimensional structure prediction

Comparative modeling of Glutathione-S-tranferese was conducted by using Modeller9v12 program [28]. BlastP and FASTA searches were performed independently with PDB for obtaining a suitable template. The significance of the BLAST results was assessed by expect

values (e-value) generated by BLAST family of search algorithm. The target-template alignment was carried out using ClustalW version 2.1 [20] and Modeller version 9.12 [21] programmes. Comparative modelling was conducted by the Modeller version 9.12 programme [28]. The loop regions were modeled using MODLOOP server [21]. The final 3D structures with all the coordinates for GST were obtained by optimization of a molecular probability density function (pdf) of Modeller [29]. The molecular pdf for homology modelling was optimized with the variable target function procedure in Cartesian space that employed the method of conjugate gradients and molecular dynamics with simulated annealing [30].

The computational protein structures were evaluated [31] by ERRAT [32] and ProCheck [33] programmes. After fruitful verification, the coordinate files were successfully deposited to PMDB (Protein Model Database)[34]. All the graphic presentations of the 3D structures were prepared using Chimera version 1.8.1 [35].

Proteomics analysis and functional annotation

The proteomics analyses were carried out using CLC Genomics Workbench 7.0.3 (CLC Bio, Hyderabad, India). ProFunc server [33] was used to identify the likely biochemical function of proteins from the predicted three-dimensional structure. Databases like PFam, PROSITE, PRINTS, ProDom, InterProScan were used for functional characterization.

RESULTS

Comparative sequence analysis of GST gene

The Glutathione S-transferase gene (GST genes) sequenced in the present study was of 669 nucleotide base pair long (Figure 2) with molecular weights 216.442 kDa in *M. cuchia* and 216.173 kDa in *M. albus*. The melting temperature range was between 86.78 (in *M. albus*) to 87.02 (in *M. cuchia*) at 0.1M salt concentration (Table 2). The frequency of GC ranged from 0.534 (in *M. albus*) to 0.540 (in *M. albus*). On the other hand frequency of AT in GST mRNA (cDNA) sequence in *M. cuchia* and *M. albus* was found to be 0.460 and 0.466 respectively (Table 2). The GST gene sequences of *M. cuchia* and *M. albus* were found to be G: C rich. The GST gene of *M. cuchia* is more G:C rich than that of *M. albus* (Table 2; Figure 3).

Comparative sequence analysis of Glutathione-S-transferase (GST) enzyme

The sizes of protein sequences of Glutathione S-transferase enzyme in the present study ranged at 231 (in *M. cuchia*) to 233 (in *M. albus*) amino acids. The amino acid Lucine (L) and Serine (S) have been found to be predominantly rich in the Glutathione S-transferase of *M. cuchia* and *M. albus*. However, there is more Leucine (L) frequency in GST of *M. cuchia* (frequency= 0.117) than in *M. albus* (frequency= 0.116). On the other hand, Serine (S) frequency in the GST of *M. albus* (frequency= 0.107) was more than that of *M. cuchia* (frequency= 0.069) (Table 3, Figure 4a). Sequence analysis of Glutathione S-transferase revealed –ve hydropathy on average (-0.210 in *M. cuchia* and -0.264 in *M. albus*). In the hydrophobicity plot, more –ve hydropathicity has been observed in *M. albus* than that of *M. cuchia* (Table 3; Figure 4b). The molecular weight of GST in *Monopterus* ranged

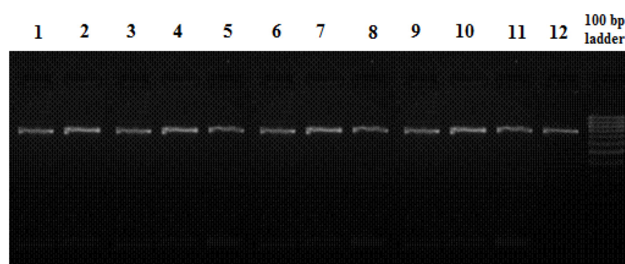


Figure 2. PCR amplification profile of GST gene (~670 bp)

Table 2: Statistical analysis of nucleotide sequence of the GST cDNA sequence in *M. cuchia* and *M. albus*.

Statistical parameter	GST	
	M. cuchia	*M. albus*
Sequence source/ GenBank Accession Number	KR705879 (This study)	KR705885 (This study)
Length (bp)	669	669
MW in single stranded condition (kDa)	216.442	216.173
Melting temperature (^{0}C) [salt] = 0.1M	87.02	86.78
Frequency of A + T	0.460	0.466
Frequency of C + G	0.540	0.534

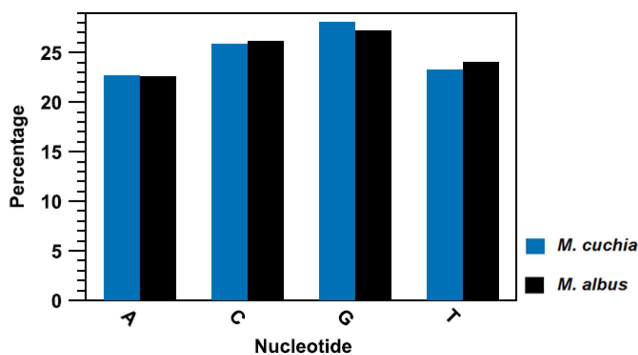


Figure 3. Nucleotide composition of the GST cDNA sequence of *M. cuchia* and *M. albus*

Table 3: Physico-chemical parameters of Glutathione S-transferase protein

Statistical parameter	GST	
	M. cuchia	*M. albus*
UniProtKB Accession number		
No. of amino acids	231	233
MW (kDa)	26.803	25.657
pI	8.95	9.61
-ve charged residues	24	14
+ve charged residues	28	24
Formula	$C_{1221}H_{1880}N_{320}O_{338}S_{11}$	$C_{1126}H_{1780}N_{328}O_{327}S_{16}$
II	49.26	56.67
AI	87.84	75.79
GRAVY	-0.210	-0.264

MW: Molecular weight; pI: Isoelectric point; II: Instability index; AI: Aliphatic index; GRAVY: Grand average of hydropathicity.

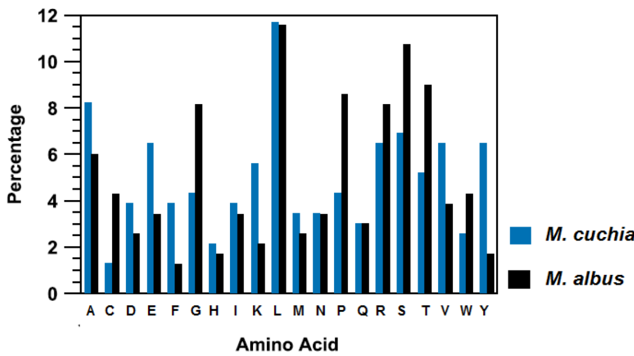


Figure 4a. Amino acid distribution histogram of for Glutathione S-transferase in *M. cuchia* and *M. albus*

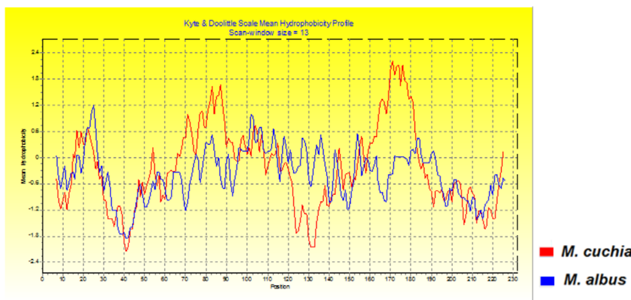


Figure 4b. Hydropathicity plot for Glutathione S-transferase protein in *M. cuchia* and *M. albus* (Kyte-Doolittle scale, Kyte and Doolittle, 1982).

from 25.657 kDa (in *Monopterus albus*) to 26.803 kDa (in *M. cuchia*). The Isoelectric point of the GST was found to be 8.95 (in *M. cuchia*) and 9.61(in *M. albus*) (Table 3; Figure 4c). The Instability index of GST of the present study was found to be 49.26 and 56.67 in *M. cuchia* and *M. albus* respectively (Table 3). Multiple amino acid sequence alignment of Glutathione S-transferase enzyme of *M. cuchia* and *M. albus* showed higher degree of conservation of respective amino acid in each alignment position (Figure 4d).

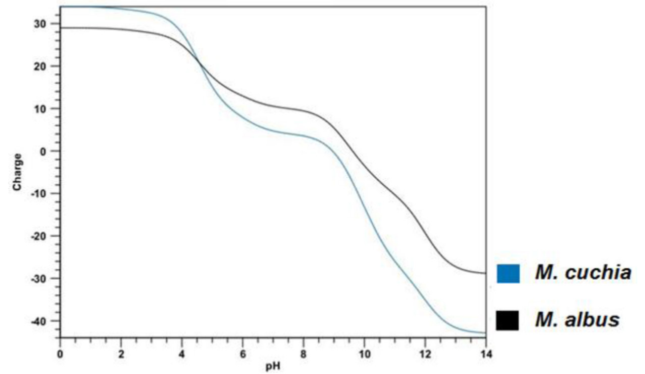


Figure 4c. Electrical *vs* pH graph for Glutathione S-transferase in *M. cuchia* and *M. albus*.

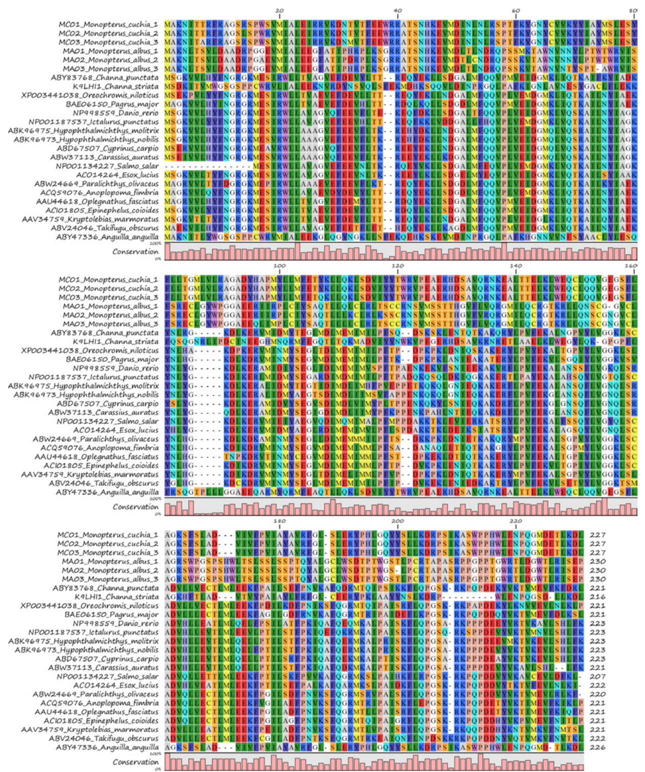


Figure 4d. Multiple sequence alignment of Glutathione S-transferase protein sequence in *M. cuchia* and *M. albus*. The sizes of the letter in the sequence logo represent the degree of conservation of respective amino acid in each alignment position. '-' represent sequence not conserved.

Molecular evolution of Glutathione S-transferase (GST) gene

The Maximum-likelihood model parameters for GST gene data sets as estimated in Modeltest (Posada and

Table 4: Model selected for Maximum-likelihood phylogeny of *gs* data sets estimated in Model test[36]

Parameter	GST gene
Model	K2+G
Bayesian Information Criterion (BIC) scores	15974
Akaike Information Criterion, corrected (AICc) value	15526
Maximum Likelihood value (*lnL*)	-7703.8
Gamma distribution (*G*)	2.10798
invariable (*I*)	n/a
Transition/Transversion bias (*R*)	1.0577
Total positions in the final dataset	490

Crandall, 1998) are listed in Table 4. There were a total of 490 positions in GST gene final dataset.

Glutathione S-transferase gene (GST) phylogeny

The Glutathione S-transferase (GST) gene phylogenetic analysis involved 30 nucleotide sequences. There were a total of 490 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

A. The evolutionary history was inferred by using the Maximum Likelihood method based on the Kimura 2-parameter model (Kimura, 1980). The tree with the highest log likelihood (-7678.3102) is shown. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+*G*, parameter = 2.1474)). The analysis involved 30 nucleotide sequences (Figure 5A).

B. The evolutionary history was inferred using the Maximum Parsimony method. Tree-1 out of 3 most parsimonious trees (length = 1803) is shown. The consistency index is (0.550223), the retention index is (0.803223), and the composite index is 0.444156 (0.441952) for all sites and parsimony-informative sites (in parentheses). The MP tree was obtained using the Close-Neighbor-Interchange algorithm (Nei and Kumar, 2000) (Figure 5 B).

The ML phylogenetic tree of GST gene revealed that, *M. cuchia* and *M. albus* formed two separate clades with bootstrap separation 98%. Within the species *M. cuchia* phylogenetic diversity has been observed. The sequence of *M. cuchia* 1, 2, 4 formed a sister sub-clades with *M. cuchia* 3, 5, 6 with a bootstrap support 74%. *Anguilla anguilla* was found to be subsequent taxa, which is separated from the genus *Monopterus* with bootstrap value 98%. The MP phylogeny of GST gene clearly showed the phylogenetic diversity of *M. cuchia* and *M. albus* (bootstrap support100%). Both *M. cuchia* and *M. albus* showed two

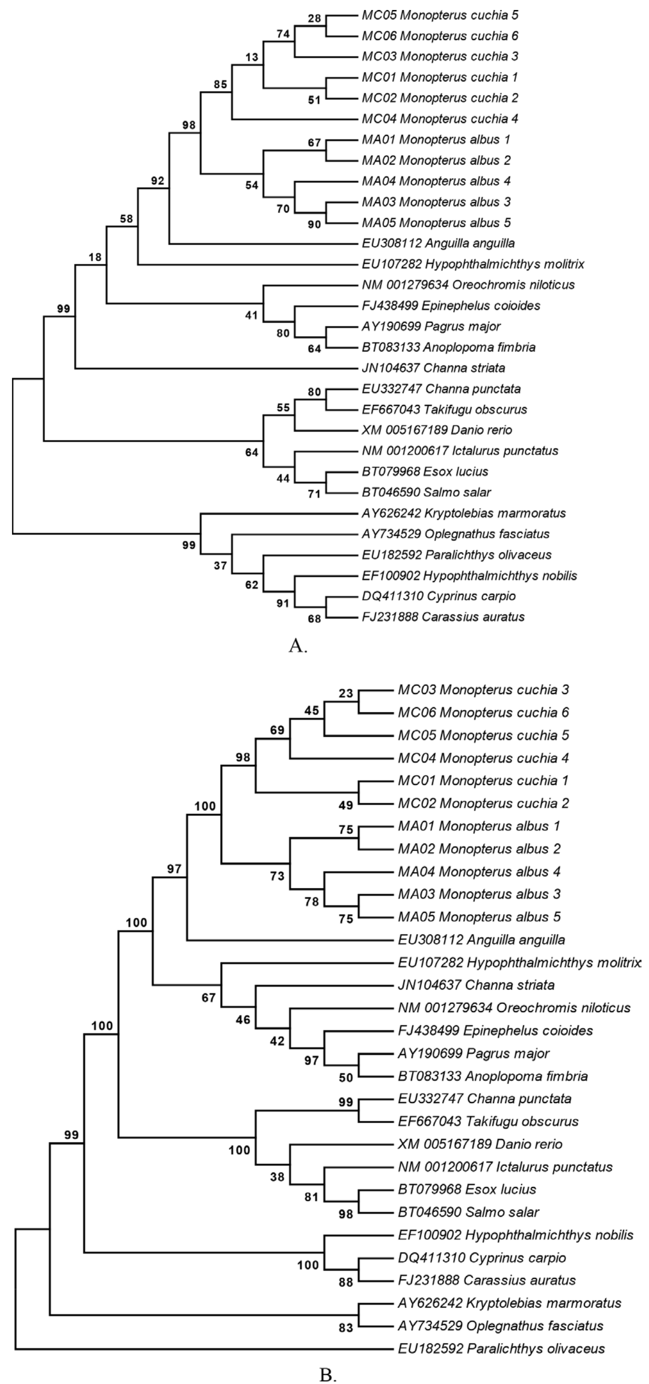


Figure 5. Phylogenetic tree of Glutathione S-transferase (GST) gene of *M. cuchia* and *M. albus* along with other eel species. (A) Maximum Likelihood method, (B) Maximum Parsimony method. The bootstrap percentage is shown next to the branches.

sub-clades in the MP tree at 98% and 73% bootstrap separation respectively (Figure 5A &5B).

Tertiary structures of Glutathione S-transferase enzyme

The 3D structure of GST *of M cuchia* (231 amino acid residues) has 1 sheet, 1 beta-alpha unit, 1 beta hairpin,

Table 5: Structural properties of predicted tertiary structures of GST

Protein name	Taxon	Number of helices	Number of helix- helix interacts	Number of sheets	Number of beta hairpins	Number of beta turns	Number of gamma turns
GST	*Monopterus cuchia*	11	11	1	1	15	7
	Monopterus albus	11	9	1	-	34	5

Table 6: Predicted functions of GST enzyme with respective ProFunc score (shown within parenthesis)

Taxon	Protein name terms	Gene Ontology (GO) terms		
		Cellular component	*Biological process*	*Biochemical function*
M. cuchia	glutathione (85.41), s-transferase (37.37), transferase (37.09), glutathione transferase (35.69), glutathione s-transferase (35.65)	cell (48.07), cell part (48.07), intracellular (48.07) ,intracellular part (48.07)	metabolic process (103.91), cellular process (50.68), cellular metabolic process (50.68), response to stimulus (31.61)	catalytic activity (102.66), transferase activity (95.85), glutathione transferase activity (61.57)
M. albus	glutathione (6.44), uncharacterized (5.61), fragment (3.90), glutathione transferase (2.25)	cell (8.99), cell part (8.36), intracellular (7.45), intracellular part (6.54)	metabolic process (18.00), cellular process (10.08), cellular metabolic process (8.32)	catalytic activity (16.49), transferase activity (10.14), binding (7.51)

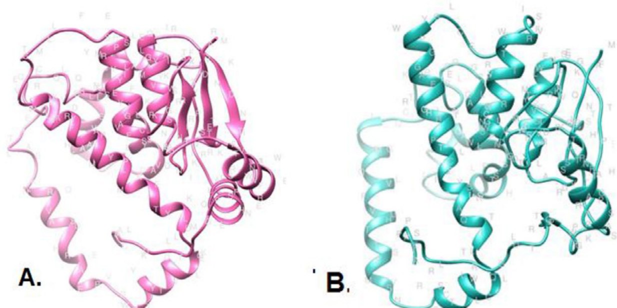


Figure 6. The tertiary structures of Glutathione S-transferase (GST), displayed in UCSF Chimera. (A) GST of *M. cuchia* (PMDB ID: PM0079982); (B) GST of *M. albus* (PMDB ID: PM0079983).

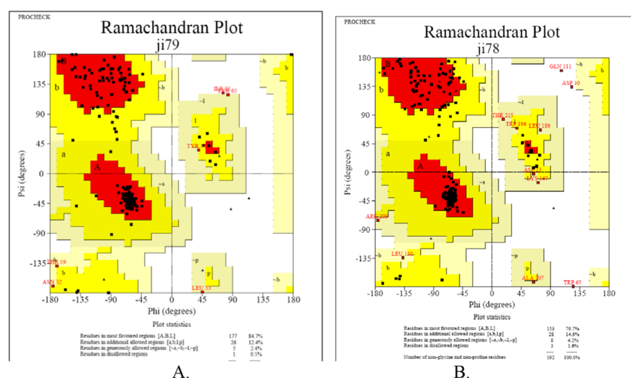


Figure 7. Ramachandran plot of backbone dihedral angles PSI (y) and PHI (f) for the tertiary structure of GST from (A) *Monopterus cuchia* (left.) and (B) *M. albus* (right). Red region represents the most favored region, yellow = allowed region, light yellow = generously allowed region, white = disallowed region [ProCheck].

1 psi loop, 4 strands, 11 helices, 11 helix-helix interacts, 15 beta turns and 7 gamma turns (Table 5; Figure 6A). On the other hand, the 3D structure of GST *of M albus* (233 amino acid residues) has 1 sheet, 1 beta-alpha-beta unit, 1 psi loop, 3 strands, 11 helices, 9 helix-helix interacts, 34 beta turns and 5 gamma turns (Table 5; Figure 6B). The predicted tertiary structure of GST were verified by the visualization of Ramachandran plot using ProCheck programme (Figure 7). The functional annotation results of GST with respective ProFunc score are listed in Table 6.

The sequences of GST gene of the present study have been successfully deposited to GenBank Database of NCBI and accession number has been obtained for each submitted sequence (Table 7). The protein structure of GST of *M. cuchia* and *M. albus* have been deposited to the Protein Model Database (PMDB) and PMDB-IDs PM0079982 and PM0079982 has been assigned to each submitted structures.

DISCUSSION

The present study had revealed an interesting point of identification that the GST gene sequences of *M. cuchia* is more G:C rich than *M. albus* and the GST gene of *M. albus* is more A:T rich than that of *M. cuchia* (Table 2). The instability index (II) value of GST enzyme of the present study ranged between 49.26 and 52.67. This classifies the GST protein as unstable. The Grand average of hydropathicity (GRAVY) value of -0.210 and -0.264 (Table 3). The high Leucine content the amino acid

Table 7: Details of the specimens for which the gst genes were sequenced in the present study and submitted to GenBank

Sample name	Voucher number	Locality	GenBank accession numbers assigned
M. cuchia	MC0001	Bhogdoi River, Jorhat	KR705879
M. cuchia	MC0002	Kakodonga River, Golaghat	KR705880
M. cuchia	MC0003	Jia Bhorali River, Sonitpur	KR705881
M. cuchia	MC0004	Bishwanath Ghat, Sonitpur	KR705882
M. cuchia	MC0005	Kulshi river, Kamrup	KR705883
M. cuchia	MC0006	Rice field, Bilaishipara, Dhubri	KR705884
M. albus	MA0001	Loktak Lake, Manipur	KR705885
M. albus	MA0002	Loktak Lake, Manipur	KR705886
M. albus	MA0003	Ijei River, Longmai, Tamenglong, Manipur	KR705887
M. albus	MA0004	Ijei River, Longmai, Tamenglong, Manipur	KR705888
M. albus	MA0005	Tuivel River, Manipur	KR705889
M. albus	MA0006	Tuivel River, Manipur	KR705890

sequence of the GST is responsible for increased kinetics of the protein synthesis and controlling protein breakdown rates [37].

The phylogenetic analysis of GST gene clearly indicates the possibility for existence of two sub-species in *M. cuchia*. The MP phylogeny of GST gene also revealed genetic diversity within *M. albus* at 73% bootstrap separation. The study on GS and GST proteins and their respective genes also supports the earlier findings on other markers used either on *M. cuchia* [38-42] or *M. albus* [42,43] that high genetic diversity/polymorphism exist within the species *Monopterus cuchia* or *M. albus*.

Efficient identification of the two eel species of the present study is critical for aquaculture management as well as for eel conservation. Thus, identification of *M. cuchia* and *M. albus* needs to be supported by molecular characterization instead of conventional methods [44]. The present study had revealed an interesting point of identification that the GST gene sequences of *M. cuchia* is more G: C rich than *M. albus* and the GST gene of *M. albus* is more A: T rich than that of *M. cuchia*. The A/T content has driven the relocation of genes to and from the nuclear lamina, in tight association with changes in expression level. The A/T-rich stretches in genomes serve as anchoring sequences that form a structural backbone of interphase chromosomes [45].

The phylogenetic analysis of GST gene revealed that, *M. cuchia* and *M. albus* formed two separate clades with high bootstrap separation 98% and 100% in ML and MP tree respectively clearly indicates the possibility for existence of two sub-species in *M. cuchia*. The MP phylogeny of GST gene also revealed genetic diversity within *M. albus* at 73% bootstrap separation, which may be due to existence of species complex. Some of the samples of *M. albus* and *M. cuchia* in the GST gene phylogenetic tree revealed the possibility of interbreeding or hybridization in the *Monopterus* species.

ProFunc analysis has revealed that GST has several functional properties which include cellular metabolic process, response to stimulus, catalytic activity and transferase activity. The predicted 3D structure of GST will reveal more insights on GST activity in *M. albus* and *M. cuchia* in order to reveal the detailed detoxification mechanism freshwater eels.

The present study provides a strong basis for further analysis concerning population structure and evolutionary relationships of two synbranchid species of genus *Monopterus*. The 3D structures generated under the present study can 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.

Rapid application of new genetic tools and analytical techniques is apparent within the literature on fishes. The present study has been able to provide better resolution to phylogenetic analysis and to discover aspects of fish biology. It is now possible to estimate levels of gene flow using a diversity of markers, to estimate migration rates and identify potential migrants and to determine whether there is sex-biased gene flow.

The development of molecular tools for freshwater fishes is not complete. Obviously there is great potential in exploring and applying entirely new molecular techniques, such as sequencing entire mitochondrial genomes [46] development of additional informative nuclear markers [47], or microarrays and beyond. The emergence and development of DNA sequencing techniques and methods for the analysis of molecular

and morphometric data has led to an exponential increase in the number of papers that have included phylogenetic trees for various fish groups.

CONCLUSION

The nuclear gene encoding Glutathione S-transferase has been used for molecular characterization of genus *Monopterus*. Glutathione S-transferase is potentially used as a biochemical marker of aquatic pollution and they are the decisive indicators of the toxic effect. The gs gene sequences of *M. albus* were found to be A:T rich for all the three gs gene sequences. On the other hand, the gs genes sequences of *M. cuchia* were rich in G:C frequency. Glycine and Proline has been found predominantly rich in the GS-I of *M. cuchia*. The GST gene sequences of *M. cuchia* is more G:C rich than *M. albus* and the GST gene of *M. albus* is more A:T rich than that of *M. cuchia*. Grand average of hydropathicity (GRAVY) value of -0.210 to -0.264 of the present study revealed that GST in *Monopterus* is of hydrophilic nature and it is more hydrophilic in *M. albus* than *M. cuchia*. GST is a Lucine (L) rich protein in *Monopterus*. The high Leucine content the amino acid sequence of the GST is responsible for increased kinetics of the protein synthesis and controlling protein breakdown rates. The phylogenetic analysis of GST gene revealed that, *M. cuchia* and *M. albus* formed two separate clades with high bootstap separation 98% and 100% in ML and MP tree respectively clearly indicates the possibility for existence of two sub-species in *M. cuchia*. Some of the samples of *M. albus* and *M. cuchia* in the GST gene phylogenetic tree revealed the possibility of interbreeding or hybridization in the *Monopterus* species. The predicted 3D structure of GST will reveal more insights on GST activity in *M. albus* and *M. cuchia* in order to reveal the detailed detoxification mechanism freshwater eels. GST can be used as a potential molecular marker can establish differentiation and variation in a species complex- *M. cuchia* and *M. albus*.

REFERENCES

1. Hamilton, F. An account of the fishes found in the river Ganges and its branches. Edinburgh & London. 1822; I-VII + 1-405, Pls. 1-39.
2. Zuiew, B. Biga muraenarum, novae species descriptae. *Nova Acta Acad. Sci. Petropolitanae*. 1793; 7: 296-301.
3. Collins, T.M., Trexler, J., Nico, L. and Rawlings, T.A. Genetic diversity in a morphologically conservative invasive taxon: multiple introductions of swamp eels to the southeastern United States. Conserv Biol. 2002; 16: 1024-1035.
4. Li, D., Kang, D., Yin, Q., Sun, X. and Liang, L. Microsatellite DNA Marker Analysis of Genetic Diversity in Wild Common Carp (*Cyprinus carpio* L.) Populations. Journal of Genetics and Genomics. 2007; 34 (11): 984–993.
5. IUCN. The IUCN Red List of Threatened Species. 2020; Version 2019-3. https://www.iucnredlist.org
6. Dahanukar, N. *Monopterus cuchia*. In: IUCN 2013. IUCN Red List of Threatened Species. 2010; Version 2011.2. www.iucnredlist.org.
7. Van der Oost, R., Beyer, J. and Vermeulen, N.P.E. Fish bioaccumulation and biomarkers in environmental risk as a bioactivation reaction. Chem Biol Interact. 2003; 129: 61-76.
8. Jakoby, W. B. The glutathione S-transferases: a group of multi-functional detoxification proteins. Adv Enzymol. 1978; 46: 383-414.
9. Nimmo, I.A. The glutathione S-transferase of fish. Fish Physiol. Biochem. 1987; 3(4): 163-172.
10. Leaver, J., Scott, K. and George, S.G. Cloning and characterization of the major hepatic gluthathione-S-transferase from a marine teleost flatfish, the plaice (*Pleuronectes platessa*) with structural similarities to plant, insect and mammalian Theta class isoenzymes. Biochem. J. 1993; 292: 189-195.
11. Zemla, A., Venclovas, C., Moult J. and Fidelis, K. Processing and analysis of CASP3 protein structure predictions. Proteins Suppl. 1999; 3: 22–29.
12. Talwar, P. K. and A. G.Jhingran. Inland fishes of India and adjacent countries, Vol. 2. A. A. Balkema, Ilottcrdam. Pp. 543-1158. 1992. [Synbratlchitlae, pp. 774-780].
13. Salah, M.A. and Iciar, M. Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. Nucleic Acids Res. 1997; 25: 4692-4693.
14. Cabe, P.R., Page, R.B, Hanlon, T.J., Aldrich, M.E., Connors, L. and Marsh, D.M. Fine-scale population differentiation and gene flow in a terrestrial salamander (*Plethodon cinereus*) living in continuous habitat. Heredity. 2007; 98: 53–60.
15. Benson, D.A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J. and Sayers, E.W. GenBank. Nucleic Acids Research. 2013; 41: D1Pp. D36-D42 (doi:10.1093/nar/gks1195).
16. Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nuc. Ac. Res. 1997; 25(17): 3389-3402.
17. Pearson, W.R. Searching protein sequence libraries: comparison of the sensitivity and selectivity of the Smith-Waterman and FASTA algorithms. *Genomics.* 1991; 11(3): 635-50.
18. Rice, P., Longden, I. and Bleasby A. EMBOSS: The European Molecular Biology Open Software Suite. Trends Genet. 2000; 16:276-277.
19. Apweiler, R., Bairoch, A., Wu, C.H., Barker, W.C., Boeckmann, B., Ferro, S., Gasteiger, E., Huang, H., Lopez, R., Magrane, M., Martin, M.J., Natale, D.A., O'Donovan, C., Redaschi, N. and Yeh, L.S. UniProt: the Universal Protein knowledgebase. Nucleic Acids Research. 2004; 32:D115-119.
20. Higgins, D., Thompson,J., Gibson, T., 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. Nuc. Ac. Res. 1994; 22:4673-4680.
21. Fiser, A., Do, R.K. and Sali, A. Modeling of loops in protein structures. Protein Science. 2000; 9: 1753-1773.
22. 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. Nuc. Ac. Res. 1994; 22: 4673–4680.
23. Tamura, K., Dudley, J., Nei, M. and Kumar, S. MEGA 4: Molicular Evolutionary Genetics Analysis (MEGA) software verson 4.0. Molecular Biology and Evolution. 2007; 24: 1596-1599.
24. Eck, R. V., Dayhoff, M. O. Atlas of protein sequence and structure. National Biomedical Research Foundation, Silver Spring, Maryland, USA. 1966.
25. Jones, D.T., Taylor, W.R. and Thornton, J.M. The rapid generation of mutation data matrices from protein seque nces. *Comput. Applic. Biosci.* 1992; 8: 275-282.
26. Nei, M. and Kumar, S. Molecular evolution and phylogenetics. Oxford University Press, Oxford. 2000.
27. Felsenstein, J. Inferring phylogenies. Sunderland, Massachusetts: Sinauer Associates. pp. 664. 2003.
28. Martí-Renom, M.A., Yerkovich, B. and Sali, A. Comparative protein structure prediction. John Wiley & Sons, Inc. Curr. Protocol. Prot. Sci. 2002; 1: 2.9.1 - 2.9.22.
29. Eswar, N., Webb, B., Marti-Renom, M.A., Madhusudhan, M.S., Eramian, D., Shen, M.Y., Pieper, U. and Sali, A. Comparative Protein Structure Modelling Using MODELLER. Ed: Coligan JE, Dunn BM, Speicher DW, Wingfield PT. Current Protocols in Protein Science Unit. 2.9 1-31. 2006.

30. Sali, A. and Blundell, TL. Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol.1993; 234: 779-815.
31. Giorgetti, A., Raimondo, D., Miele, A.E. and Tramontano, A. Evaluating the usefulness of protein structure models for molecular replacement. Bioinformatics. 2005; 21(2): 72–76.
32. Colovos, C., Yeates, T.O., Verification of protein structures: Patterns of nonbounded atomic interections. Protein Science. 1993; 2:1511-1519.
33. Laskowski, R.A., Watson, J.D. and Thornton, J.M. ProFunc: a server for predicting protein function from 3D structure. Nucleic Acids Res. 2005; 33: W89-W93.
34. Tiziana, C., Paolo D'Onorio, D.M., Domenico, C., Talamo, I.G. and Tramontano, A. The PMDB Protein Model Database CASPUR. Nuc. Ac. Res. 2006; 34(Database Issue): D306-D309.
35. Pittersen, E.F., Goddard, T.D., Huang, C.C.,Couch, G.S.,Greenblatt, D.M., Merg, E.C., Ferrin, T.E., UCSF Chimera-a visualization system for exploratory research and analysis. J.Comput.chem. 2004; 25: 1605-1612.
36. Posada, D., Crandall, K.A. MODELTEST: testing the model of DNA substitution. Bioinformatics. 1998; 14: 817–818.
37. Garlick, P.J. The role of leucine in the regulation of protein metabolism. J. Nutr. 2005; 135: 1553S-156S.
38. Alam, M.S., Islam, M.S. and Alam, M.S. DNA Fingerprinting of the Freshwater Mud Eel, Monopterus cuchia (Hamilton) by Randomly Amplified Polymorphic DNA (RAPD) Marker. International Journal of Biotechnology and Biochemistry. 2010; 6(2): 271-278.
39. Miah, G., Rafii, M.Y. and Ismail, M.R. A Review of Microsatellite Markers and Their Applications in Rice Breeding Programs to Improve Blast Disease Resistance. International Journal of Molecular Sciences. 2013; 14(11):22499-22528.
40. Devi P. , Baruah C. and Sharma D.K. Mitochondrial DNA sequence and amino acid analysis of the Cytochrome c oxidase subunit I (COI) from two eel species, *Monopterus cuchia* and *Monopterus albus.* Internat. J. Adv. Biot. Res. 2014; 5(3):283-294.
41. Devi P., Baruah C. and Sharma D.K. Sequence Analysis, Structure Prediction and Phylogeny of Glutamine Synthetase of Two Eel Species, *Monopterus cuchia* (Hamilton) and *Monopterus albus* (Zuiew). International Journal of Molecular Biology. 2015; 6(1):102-111.
42. Devi P, Laskar MA, Baruah C, Sharma D.K. Genetic Diversity and Population Structure of *Monopterus* spp. of Assam and Manipur states of Northeast India using Microsatellite Markers. International Journal of Fisheries and Aquatic Studies. 2017;5(5):457-64.
43. Lei, L., Feng, L., Jian, T.R. and Yue, G.H. Characterization and multiplex genotyping of novel microsatellites from Asian swamp eel, Monopterus albus. Conservation Genetics Resources. 2012; 4 (2): 363-365.
44. Huang, J.P., Han, Y. S. and Tzeng, W. N. Species identification of Anguillid eels by polemerase chain reaction/restriction fragment length polymorphism analysis of the gonadotrophin II-β subunit gene. Acta Zoologica Taiwanica. 2001; 12(2): 41-49.
45. Meuleman, W., Peric-Hupkes, D., Kind, J., Beaudry, J.B., Pagie, L., Kellis, M., Reinders, M., Wessels, L. and van Steensel, BConstitutive nuclear lamina–genome interactions are highly conserved and associated with A/T-rich sequence. Genome Res. 2013; 23(2): 270-280. doi: 10.1101/gr.141028.112.
46. Pollock, D.D. Eisen, J.A., Doggett, N.A. and Cummings,M.P. A Case for Evolutionary Genomics and the Comprehensive Examination of Sequence Biodiversity. Mol. Biol. Evol. 2000;17(12):1776-88.
47. Matthee, C.A. and Davis, S.K. Molecular Insights into the Evolution of the Family Bovidae: A Nuclear DNA Perspective. Mol. Biol. Evol. 2001;18(7): 1220-1230.

Cite this article: Devi P, Baruah C, Sharma DK. Molecular characterization of Glutathione S-transferase in two eel species, *Monopterus cuchia* (Hamilton) and *Monopterus albus* (Zuiew). Asian J Biol Life Sci. 2020;9(1):32-41.