Egyptian Journal of Aquatic Biology & Fisheries
Zoology Department, Faculty of Science,
Ain Shams University, Cairo, Egypt.
ISSN 1110 – 6131
Vol. 29(2): 1101 – 1118 (2025)
www.ejabf.journals.ekb.eg

Identification and *In-Silico* Characterization of Selected Osmoregulatory and Stress-Tolerance Genes of the Hilsa Shad (*Tenualosa ilisha*)

Afroza Pervin, Md. Nazmul Hasan, Atiqur Rahman Sunny, Md Toasin Hossain Aunkor, Sorwer Alam Parvez, Asif Iqbal, Md Jahangir Alam and Md. Faruque Miah*
Department of Genetic Engineering and Biotechnology, Shahjalal University of Science and Technology, Sylhet-3114, Bangladesh
***Corresponding Author:** faruque-btc@sust.edu

ARTICLE INFO
Article History:
Received: Jan. 3, 2025
Accepted: March 6, 2025
Online: March 20, 2025

Keywords:
Tenualosa ilisha, Gene, Osmoregulation, Stress-tolerance, Characterization, Bangladesh

ABSTRACT

Osmoregulatory and stress-signaling genes are critical for fish survival and adaptation in dynamic aquatic environments. The Hilsa shad (*Tenualosa ilisha*), a migratory species of high economic and ecological importance, requires a thorough understanding of the molecular mechanisms underlying osmoregulation and stress signaling for effective management and sustainable production. In this study, we identified and characterized these genes using the draft genome of Hilsa. The identification process was based on the presence of corresponding functional domains, and gene characteristics such as coding sequence and protein length were calculated. Additionally, parameters like molecular weight (MW), isoelectric point (Ip), gene localization, gene structure, conserved motifs, and secondary/tertiary structures were analyzed, along with Gene Ontology (GO) analysis. The predicted genes were validated through PCR (Polymerase Chain Reaction) using gene-specific primers and sequencing. Sequence data confirmed the predicted genes, and blast analysis revealed high similarity with homologous genes from other fish species in the databases. Compared to other fish species, MW, Ip, gene structure, and conserved motifs exhibited notable diversity. Most of the identified genes were localized and expressed in the plasma membrane, nucleus, and cytoplasm. Secondary and 3D structural analyses showed a predominance of alpha helices with high-quality structure. GO analysis indicated that the majority of the genes are involved in ion transport, while others play roles in signal transduction, protein folding, and osteoclast differentiation. These findings provide valuable insights into the molecular functions of osmoregulatory and stress-signaling genes in Hilsa, which will contribute to a deeper understanding of their roles in osmoregulation

INTRODUCTION

Tenualosa ilisha, commonly known as the hilsa herring Hilsa or ilish, is a species of fish belonging to the herring family, *Clupeidae*, and sub-family, *Alosinae*. It is a popular and economically important fish found in the Indian subcontinent and neighboring regions, including Bangladesh, India, Pakistan, Myanmar, and the Arabian Sea **(Bhaumik, 2013)**. The body of the Hilsa fish is distinctively elongated, silver in color, and has a black patch behind its gill cover. It is an anadromous fish, which means

that it migrates to freshwater rivers from the sea to spawn. The Hilsa is renowned for its distinctive flavor and is considered a delicacy in many South Asian cuisines **(Alam *et al.*, 2012)**.

As one of three GI products, Bangladesh produces more than 75% of global Hilsa production **(Mahmud, 2020)**. In terms of nutritional value, this is a fish rich in protein (18%), fat (19.5%) and a rich source of essential micronutrients, especially phosphorus, calcium, zinc, vitamins A and E, polyunsaturated fatty acids (PUFA) and omega-3 fatty acids **(Alam *et al.*, 2012; Hossain *et al.*, 2014; Begum *et al.*, 2016)**. In Bangladesh, in 2017-2018, Hilsa contributed 517,198 tons (12% of total fish production) worth about $4 billion, of which 232,698 tons (6% of catch) were caught from inland waters and the remaining 284,500 tons (44% of marine catches) are harvested from marine waters **(Sarker *et al.*, 2019; Rahman *et al.*, 2020)**. Its non-consumptive value is estimated at approximately US$0.36 billion per annum **(Mohammed *et al.*, 2016)**. In the locality, about 2.5 million people (2% of the total population) directly or indirectly depend on Ilish production. It constitutes 12% of total fish production and contributes to more than 1% of the total GDP of Bangladesh **(Deb *et al*., 2022)**.

Hilsa is known for its ability to tolerate a wide range of salinity levels and is distributed throughout the marine, estuarine, and riverine environments of the Persian Gulf, Red Sea, Arabian Sea, Bay of Bengal, Vietnam Sea, and China Sea **(Bhaumik 2013)**. This characteristic is essential for its life cycle, as it is an anadromous fish that migrates between freshwater rivers and the marine environment for breeding purposes. During the breeding season, Hilsa migrates from the sea to the freshwater rivers to spawn. The eggs are laid in freshwater, and the larvae hatch and develop in river estuaries and brackish waters. After spending some time in these brackish waters, the young Hilsa juveniles then move back to the sea **(Hossain *et al.*, 2016)**. Because of this life cycle, Hilsa is adapted to survive and thrive in various salinity conditions, ranging from freshwater to brackish waters and full marine salinity levels. This ability allows them to navigate between different environments during their life stages and contributes to their successful reproduction and survival as a species **(Sahoo *et al.*, 2018)**. Currently, Hilsa production faces several challenges, particularly in countries where it is a significant cultural and economic resource, including overfishing, habitat loss and degradation, illegal, unreported, and unregulated fishing, climate change, lack of effective management and regulation, interference with migratory routes, pollution, and water quality, lack of data and research, socio-economic challenges and international conflicts **(Sahoo *et al.*, 2018; Mozumder *et al.*, 2019)**. Moreover, the reduction in the depth and discharge of rivers due to the construction of dams and barrages has affected the spawning, feeding, and migration of these fish **(Rana *et al.*, 2007; Ahsan *et al.*, 2014; Jahan *et al.*, 2017)**. Addressing these challenges requires a comprehensive and collaborative approach involving governments, fishing communities, conservation organizations, and researchers. Sustainable fisheries management, habitat restoration,

community engagement, and regional cooperation are essential components to ensure the conservation and sustainable production of Hilsa **(Murshed-e-Jahan *et al.*, 2014)**. There have been various attempts to research and promote the artificial breeding of Hilsa fish due to its economic and ecological significance in many South Asian countries, particularly in Bangladesh, India, and Myanmar **(Sahoo *et al.*, 2018)**. Artificial breeding and hatchery techniques are being explored as a means to replenish and sustain Hilsa stocks. It's important to note that success in Hilsa artificial breeding is a complex challenge due to the fish's unique biology and migratory behavior. Therefore, it is crucial to investigate the genetic basis of Hilsa's distinctive biology and behavior. Scientists have just mapped the whole genome of the Hilsa **(Mohindra *et al.*, 2019)**, however little is known about its osmoregulatory and stress-responsive genes. In addition to whole-genome sequencing, a transcriptomic catalog of Hilsa was generated by assembling RNA-Seq reads from various tissues, including the brain, gill, kidney, liver, and muscle. This analysis provided limited insights into osmoregulatory and stress-signaling genes **(Chowdhury *et al.*, 2023)**. Recently, large-scale gill transcriptome analyses have been conducted across three salinity regions freshwater, brackish water, and marine water yielding valuable expression profiles of osmoregulatory genes **(Mohindra *et al.*, 2023)**. Osmoregulation and stress response are critical physiological processes in fish that help them adapt to changes in their aquatic environments **(Kültz, 2015; Seale & Breves 2022)**. Osmoregulation involves the regulation of water and ion balance within the fish's body, allowing them to maintain internal stability in varying salinity conditions **(Malakpour Kolbadinezhad *et al.*, 2018)**. Stress responses help fish cope with various environmental stressors, such as temperature fluctuations, pollutants, and physical disturbances **(Barton, 2002)**. In fish, these processes are controlled by the expression of specific genes that encode for proteins involved in osmoregulation and stress responses which include Na^+/K^+-ATPase (NKA), Aquaporin, Chloride Channel, Cystic Fibrosis Transmembrane Conductance Regulator (CFTR), Heat Shock Proteins (HSPs), Glutathione-S-transferase (GST), Cortisol Receptor, Protein Kinase, Apoptosis-Related, Catecholamines and Adrenergic Receptor and Oxidative Stress-Related Genes **(Ackerman *et al.*, 2000; Eissa & Wang, 2016; Islam *et al.*, 2022)**. The expression of these genes is often regulated by complex molecular pathways, involving various transcription factors and signaling molecules **(Iwama *et al.*, 1999; Regoli & Giuliani, 2014)**. Research in this field continues to uncover the specific mechanisms through which fish respond to osmotic and stress challenges, contributing to our understanding of their physiological adaptations and potential implications for conservation and aquaculture. This study presents a curated list of candidate genes associated with key physiological processes, including osmoregulation and stress responsiveness. This resource serves as a foundation for future functional studies to uncover the specific roles of these genes. Additionally, the findings provide baseline data for comparing gene expression profiles

across varying environmental gradients, such as the salinity changes encountered during Hilsa migration.

MATERIALS AND METHODS

Prediction and identification of the osmoregulatory and stress-responsive genes

Through a systematic review of the literature, a list of potential osmoregulatory and stress genes for Hilsa's were prepared. An anadromous model fish called the three-spined stickleback (*Gasterosteus aculeatus*) was used as a search species for acquiring the protein sequences of the selected genes from the ENSEMBL database (https://asia.ensembl.org/index.html). Then, the protein sequences of the enlisted genes retrieved from the ENSEMBL were used as query sequences for the tBLASTn program against the whole-genome shotgun contig database of NCBI, and the search were kept limited by *Tenualosa ilisha* which provided numerous contig sequences of our enlisted genes. We finalized the contigs based on the highest query coverage and identity. Afterward, the FASTA format of the identified contig sequences was used to predict the gene using FGENESH, an HMM-based gene structure prediction tool (http://www.softberry.com/). Here, the *Dicentrarchus labrax* was selected as the organism-specific gene-finding parameter to predict genes. Further, the existence of the predicted genes was confirmed by tBLASTn similarity searching program against the nucleotide collection database of NCBI.

In-silico characterization of the predicted genes

The Gene Ontology of the predicted genes was annotated by InterPro scan and their enrichment analysis was conducted by ShinyGO 0.77 selecting stickleback as best matching species. Additionally, the Pfam database (http://pfam.xfam.org/) was used to determine the protein families of the predicted proteins encoded by the target genes. The physico-chemical properties i.e., the isoelectric point and molecular weight of the genes were evaluated with the help of ProtParam (http://web.expasy.org/protparam/). The subcellular localization of the encoded proteins was measured by using the Cello web tool (http://cello.life.nctu.edu.tw/cello2go/). Moreover, the length of protein and CDS as well as the number of exons and introns of all genes were estimated by FGENESH tool.

Analysis of conserved motifs and gene structure

Multiple Em for Motif Elicitation (MEME), version 5.3.0 (http://meme-suite.org) was used in order to explore the conserved motifs on the predicted proteins. Each motif's highest and lowest sites were set to 600 and 2, respectively, with a minimum width of 6 and a maximum of 50. The motifs were positioned next to their respective genes. To evaluate gene structure, CDS sequences were compared to corresponding genomic

sequences. Gene structure was shown using the GSDS 2.0 (Gene Structure Display Server) tool (http://gsds.gao-lab.org/) **(Hu *et al.*, 2015)**.

Secondary and tertiary structure prediction of the selected genes

The secondary structures of the selected proteins were built using the online tool NPS@SOPMA (https://npsa-prabi.ibcp.fr/) with default parameters. The structure investigation comprised the quantity and percentage of alpha helixes, extended strands, and random coils presented in various colors. Using Swiss-Model (https://www.swissmodel.expasy.org/), which is based on the automated ExPASy (Expert Protein Analysis System) web server, 3D models of tertiary structures and homologous modeling of proteins were simulated. The Biovia Discovery Studio Visualizer client 2016 software was used to visualize the protein structure's PDB data.

Sample collection, DNA isolation, and PCR amplification of predicted genes

To validate the predicted genes, we have collected Hilsa samples (gills) from four locations in Bangladesh (Cox's Bazar, Chandpur, Bhola, and Sunamganj districts). From the samples (3 from each location), we extracted DNA using the DNA isolation kit (Favorgen, USA) and quantified using (Nanodrop, Thermo Scientific™ NanoDrop™ One Spectrophotometer). For PCR amplification of the predicted genes, corresponding primers (Table 1) were designed using the Primer BLAST tool. PCR profiles were 10 minutes of initial denaturation at 94°C (1 cycle), 30 seconds of denaturation at 94°C (35 cycles), 30 seconds of annealing at 50-60°C (based on primer pairs) (35 cycles), 45 seconds of extension at 72°C (35 cycles) and 5 minutes of final extension at 72°C (1 cycle). To visualize the PCR products, agarose gel electrophoresis was performed (1.2% agarose) and documented under UV Transilluminator.

Table 1. List of gene-specific primers used for PCR amplification

S. No.	Gene name	Sequence	Annealing temp. (°C)	Amplicon Size (bp)
1	*TiAQP3a*	Forward Primer-5'-AAGAGCGTGGTGTTCAAAGC-3' Reverse Primer-5'-CTCCCCTTTAATCCCGTTGT-3'	58.4	482
2	*TiATP6V1a*	Forward Primer-5'-AAATGTGTTGGCGCTTGAGG-3' Reverse Primer-5'-AACCAAAGAGCAAGGGGTGT-3'	58.4	557
3	*TiBMP3*	Forward Primer-5'-GGTGGTGGTCCATTATCTGC-3' Reverse Primer-5'-CCTCTGTGTGCTGAGAGCTG-3'	61	415
4	*TiCFTR*	Forward Primer-5'- CATGTCCCCACTGAAGTCAA-3' Reverse Primer-5'- TGACAACATCAAGGCAGAGC -3'	58.4	559
5	*TiCLDN10*	Forward Primer-5'-CCAAATGACGCACTCGCTTC-3' Reverse Primer-5'-AGTCCTTGCAGTTGGAGACC-3'	60.5	509
6	*TiFOXI3b*	Forward Primer-5'-CGGGGCACCTTCTTGAAACA-3' Reverse Primer-5'-CCAAAGCTTGGCACAGAGTC-3'	60.5	667

7	*TiPLCB4*	Forward Primer-5'-CCTCCATCTGGTGGTGAATC-3' Reverse Primer-5'-GACAATGCTGTGCAGGGTAA-3'	59	683
8	*TiHSP70*	Forward Primer: 5'- ACTTATCCAGCTTGGCATCTC-3' Reverse Primer: 5'- CATCACTGTTCCTGCCTACTT -3'	59.4	557
9	*TiHSPD1*	Forward Primer: 5'- CTGGCATCCTTAGCACCTTT -3' Reverse Primer: 5'- GTACGGCCCACTTTCTTCAT -3'	58.4	727

Confirmation of genes through sequencing

After PCR and Gel Electrophoresis, gel purification of the PCR products was accomplished with using the FavorPrepTM Gel Purification Kit (FAVORGEN Biotech Corporation) following the instructions. A Nanodrop spectrophotometer (Thermo Scientific™) was used to determine the gel-purified DNA quality. The absorbance ratios (A260/280) of the samples ranged from 1.8 -2.0 which are optimal for good-quality DNA. Then, 20µl purified DNA samples with corresponding forward and reverse primers were sent to Macrogen (South Korea) for sequencing. Finally, using the sequence data, we performed the BLASTn for similarity search against the Hilsa genome and confirmed our target genes.

RESULTS

1. Prediction and identification of the osmoregulatory and stress-responsive genes

A total number of 9 genes were listed through the extensive literature study to predict the target genes of *T. ilisha*. All genes play a vital role in the homeostasis of anadromous fishes. Among the selected genes, 7 of them were responsible for osmoregulation, whereas the others were crucial for stress tolerance. The contigs containing the target genes were selected based on the query coverage and identity (Table 2).

Table 2. Prediction of the genes in the contig and identification of the target genes

SL No.	Accession ID (Query sequence)	Contig ID	Predicted Target Gene	Position		Top BLAST Hit	Organism	Identity (%)
				Start	End			
1	ENSGACG00000010315	QYSC01119003.1	*TiAQP3a*	204908	211015	XM_012815871.2	*Clupea harengus*	87.62
2	ENSGACG00000003211	QYSC01119162.1	*TiATP6V1a*	*193660*	*203586*	XM_018761150.2	*Scleropages formosus*	*94.49*
3	ENSGACG00000001853	QYSC01001252.1	*TiBMP3*	*62921*	*68912*	XM_012835121.2	*Clupea harengus*	*82.53*
4	ENSGACG00000009039	QYSC01123815.1	*TiCFTR*	*10510*	*41014*	XM_012815624.2	*Clupea harengus*	*85.35*
5	ENSGACG00000004162	QYSC01123208.1	*TiCLDN10*	*888069*	*893292*	XM_012825117.2	*Clupea harengus*	*54.55*
6	ENSGACG00000018474	QYSC01118910.1	*TiFOXI3b*	*528861*	*530806*	XM_012835758.1	*Clupea harengus*	*79.58*

7	ENSGACG0000 0012490	QYSC011240 85.1	*TiPLCB 4*	*2226 5*	*4950 1*	XM_0128380 55.2	*Clupea harengus*	*76.76*
8	ENSGACG0000 0013048	QYSC011223 17.1	*TiHSP70*	5126 2	5467 5	XM_0315635 57.1	*Clupea harengus*	89.49
9	ENSGACG0000 0008860	QYSC011220 08.1	*TiHSPD1*	3353 88	3423 26	XM_0187391 88.1	*Scleropages formosus*	83.66

The contig sequences were further proceeded for the prediction of the genes and the target genes were then identified through BLAST. Furthermore, a total GO terms was found for the predicted 9 genes for the biological processes, molecular functions, and cellular components, as shown in Table (S1). Additionally, the protein family of all genes was also analyzed and the InterPro ID of their family was included in the table. Any kind of GO term was not found for predicted HSP70 gene of *T. ilisha*. However, Fig. (1) represents GO term enrichment analysis. The GO term enrichment analysis for all available genes to elucidate the statistically significant gene set. Here, the x-axis values display the FDR values for the enrichment of related pathways. The color chart exhibits fold enrichment for each pathway. The size of the dots is attributed to the number of genes assigned to each pathway. The GO enrichment analysis showed the engagement of the annotated genes in several specific pathways. The higher "-10 log (FDR)" value indicates a more significant enrichment of the gene set concerning the corresponding GO term. In this study, genes responsible for ion transporter activity have a significant enrichment of the gene set with respect to other GO terms.

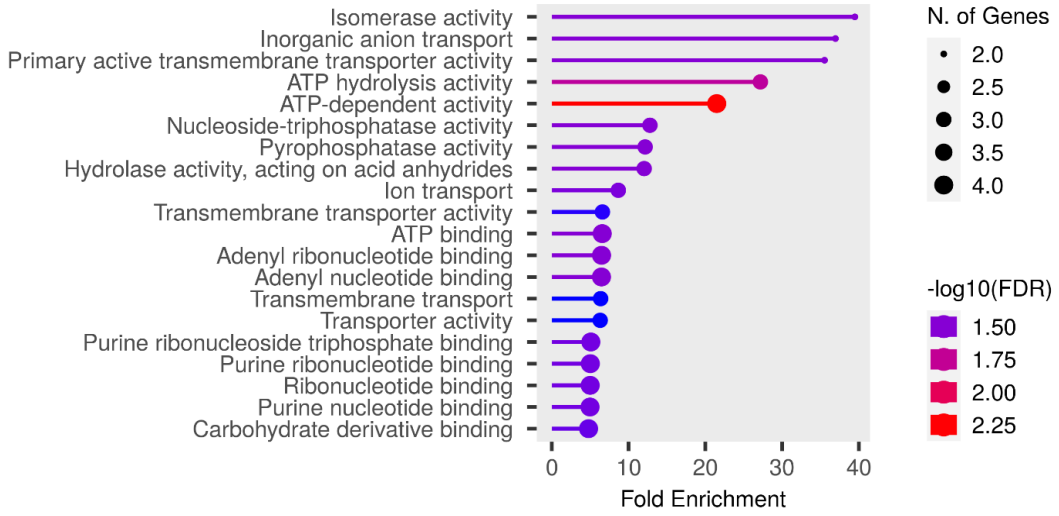

Fig. 1. Representation of the GO term enrichment analysis for all available genes to elucidate the statistically significant gene set. Here, the x-axis values display the FDR values for the enrichment of related pathways. The color chart exhibits fold enrichment for each pathway. The size of the dots is attributed to the number of genes assigned to each pathway

2. *In-silico* characterization of the predicted genes

Table (3) demonstrates the characteristic features of our predicted genes. Briefly, the gene length ranged from 30504bp (*TiCFTR*) to 1945bp (*TiFOXI3b*), the CDS length ranged from 4428bp (*TiCFTR*) to 783bp (*TiCLDN10*), and the amino acid length ranged from 1475 (*TiCFRT*) to 260 (*TiCLDN10*) amino acids. Furthermore, the PI range and molecular weight varied from 5.40 to 9.38 and 167.27 to 32.82K daltons. The *in-silico* subcellular localization analysis claimed that most of the proteins are located in the plasma membrane (3), whereas the rest are located in the cytoplasm (2), nucleus (2), extracellular areas (1), and mitochondria (1).

Table 3. Features of identified osmoregulatory and stress-tolerance proteins *Tenualosa ilisha*

Sl No	Query Gene	Position		Gene Length (bp)	Protein length	CDS length (bp)	pI*	MW (kdaltons)	Exon number	Localization*
		Start	End							
1	*TiAQP3a*	204908	211015	6107	304aa	915	6.42	32.83936	6	Plasma membrane
2	*TiATP6V1a*	193660	203586	9926	615 aa	1848	5.61	68.00986	14	Cytoplasmic
3	*TiBMP3*	62921	68912	5991	458 aa	1377	9.38	52.56404	3	Extracellular
4	*TiCFTR*	10510	41014	30504	1475 aa	4428	8.62	167.26908	27	Plasma membrane
5	*TiCLDN10*	881776	887825	6049	260 aa	783	8.99	27.90498	3	Plasma membrane
6	*TiFOXI3b*	528861	530806	1945	382 aa	1149	5.4	41.14949	2	Nuclear
7	*TiPLCB4*	22265	49501	27236	1170 aa	3456	7.69	130.99076	35	Nuclear
8	*TiHSP70*	51262	54675	3413	647 aa	1944	5.33	71.01525	1	Cytoplasmic
9	*TiHSPD1*	335388	342326	6938	580 aa	1743	5.40	61.41179	11	Mitochondrial

*pI- Isoelectric point, MW- Molecular Weight, and Localization of the selected genes predicted using the online tool CELLO: Subcellular Localization Predictive System and during localization prediction, only statistically significant results were considered.

3. Gene structure and conserved motif analysis of predicted genes

The exon/intron position of a gene family may play a major role in the process of evolution **(Xu *et al.*, 2012)**. The exon/intron position and diversity of exon/intron patterns of all predicted genes were determined by comparing the transcript sequences with the respective genomic DNA sequences using GSDS (Gene Structure Display Server) (Fig. 2).

The conserved motif analysis of nine genes in Hilsa (*T. ilisha*) revealed distinct yet highly conserved motif patterns across multiple fish species. Each gene exhibited unique motif arrangements, with color-coded motifs indicating evolutionary conservation. The p-values confirmed significant motif conservation, suggesting functional importance. Genes such as *TiAQP3a*, *TiATP6V1a*, *TiPBCB4*, *TiHSP70*, and *TiHSPD1* displayed highly conserved motif distributions across species, indicating their evolutionary stability. The motif consensus sequences further supported conservation patterns, highlighting potential regulatory and functional significance. Overall, the findings suggest that these genes have been evolutionarily conserved, likely playing crucial roles in physiological processes in hilsa and related species (Fig. 3).

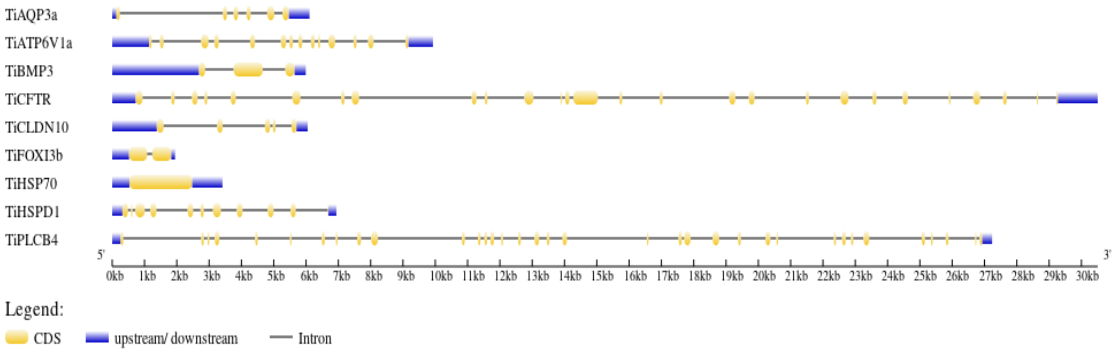

Fig. 2. Exon-intron structure of the identified osmoregulatory and stress-tolerance genes. Exons are indicated by yellow boxes; untranslated upstream and downstream areas are indicated by blue boxes, and introns are indicated by black lines

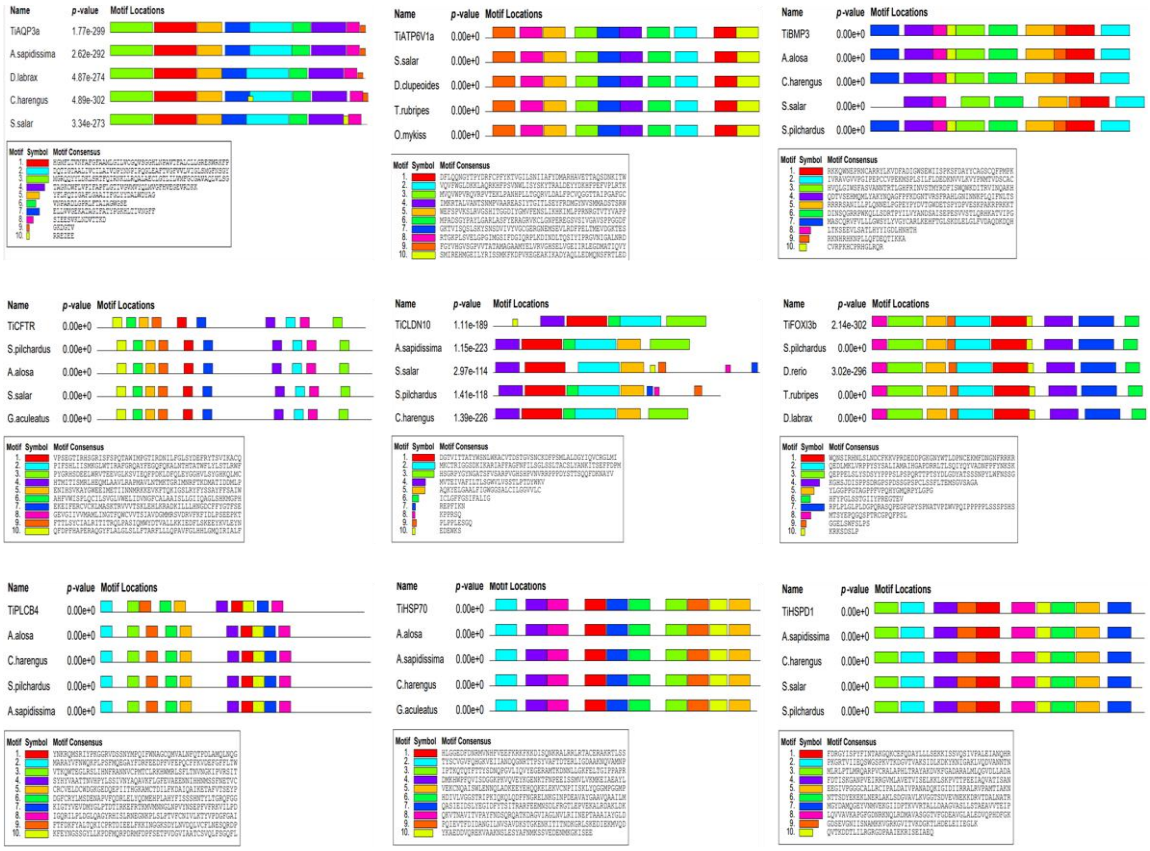

Fig. 3. The identified motifs of osmoregulatory and stress-tolerance genes in *T. ilisha* and comparison with other fish species. The motif 1 to motif 10, are displayed in different colored boxes

4. Secondary and tertiary structure of the genes

The secondary structure analysis identified the positions of alpha-helices, beta-sheets, isolated beta bridges, turns, coils, 310-helices, and transmembrane helices (Fig. 4). Their percentages and the number of the transmembrane helix are detailed in Table (S2). The analysis revealed that in most cases alpha-helices were the most dominant over all the other secondary structures, followed by turns, beta-sheets, coils, and 310-helices. Alpha-helices are a major structural component, stabilizing proteins through hydrogen bonding. The highest alpha-helix percentage is observed in *TiHSPD1* (53.62%), followed by *TiCFTR* (49.69%), *TiPLCB4* (45.87%), and *TiATP6V1a* (44.72%), suggesting these proteins have a well-defined structure. Membrane-associated proteins like *TiCFTR* (49.69%) and *TiAQP3a* (40.59%) also exhibit a high alpha-helix content, which is crucial for maintaining stability in the lipid bilayer. In contrast, *TiFOXI3b* (6.02%) has the lowest alpha-helix content. Beta-sheets contribute to protein stability, often forming rigid, sheet-like structures. The highest beta-sheet content is found in *TiAQP3a* (20.13%) and *TiHSP70* (18.70%), suggesting a role in structural support and protein-protein interactions. *TiCFTR* (10.31%), *TiPLCB4* (9.90%), and *TiHSPD1* (10.69%) have relatively lower beta-sheet content, while *TiFOXI3b* (3.93%) has the lowest, indicating minimal contribution to its structure. Coils represent flexible, unstructured regions that contribute to dynamic functions such as protein-protein interactions and regulatory roles. *TiFOXI3b* (90.05%) is highly disordered, suggesting it functions as a flexible regulatory protein. *TiBMP3* (64.85%) also has a high coil percentage, indicating a largely disordered nature. Conversely, *TiHSPD1* (35.69%), *TiCFTR* (40.00%), and *TiPLCB4* (44.22%) have lower coil percentages, indicating they adopt a more ordered structure.

Certain proteins contain membrane-spanning motifs (MSMs), which are transmembrane helices enabling their integration into cell membranes. *TiCFTR* (12 MSMs) and *TiAQP3a* (6 MSMs) contain multiple transmembrane helices, supporting their roles in transport across membranes. *TiCLDN10* (4 MSMs) also contains transmembrane segments, aligning with its role in cell junctions.

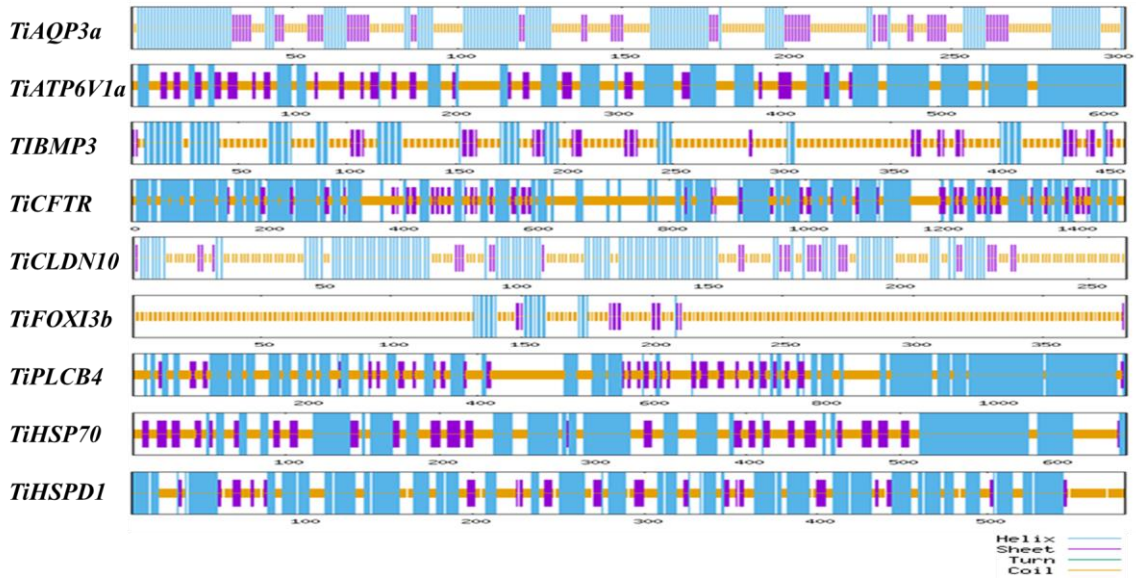

Fig. 4. Graphical representation of secondary structure elements in nine proteins (blue: helix, purple: sheet, yellow: coil)

SWISS-MODEL, an automated protein structure homology-modeling server, was applied to construct 3D models of proteins. GalaxyRefine further refined the generated models, and then the energy minimization was done via Swiss-PdbViewer. The tertiary structures of all the proteins are modeled and visualized using Discovery Studio, as shown in Fig. (5). The PROCHECK, ERRAT, and ProSA-web servers were then used to do a validation study on these constructed structures. Protein quality was assessed using PROCHECK's Ramachandran plot analysis (Fig. 6). Over 90% of the residues (excluding *TiFOXI3b*; 86.9%) were found in the favored and additional allowed regions, while only <1.0% were found in the banned regions (excluding *TiFOXI3b*; 3.1%), indicating that the projected models were of high quality. Based on ERRAT analysis, all proteins showed an overall quality factor of >90, with the exception of TiBMP3. The Z-score, the degree of nativeness of the designed models, and the energy plot, which displayed the models' local quality, were all made available by ProSA-web. The models were well within the range of X-ray crystal structure, and the Z-score was found to be in the spectrum of values often reported for natural proteins, suggesting higher quality of the structures developed (Table S3). Every residue in the simulated structure had a lower energy value, according to the energy plot. These results demonstrated the superior quality of the modeled proteins' tertiary structures.

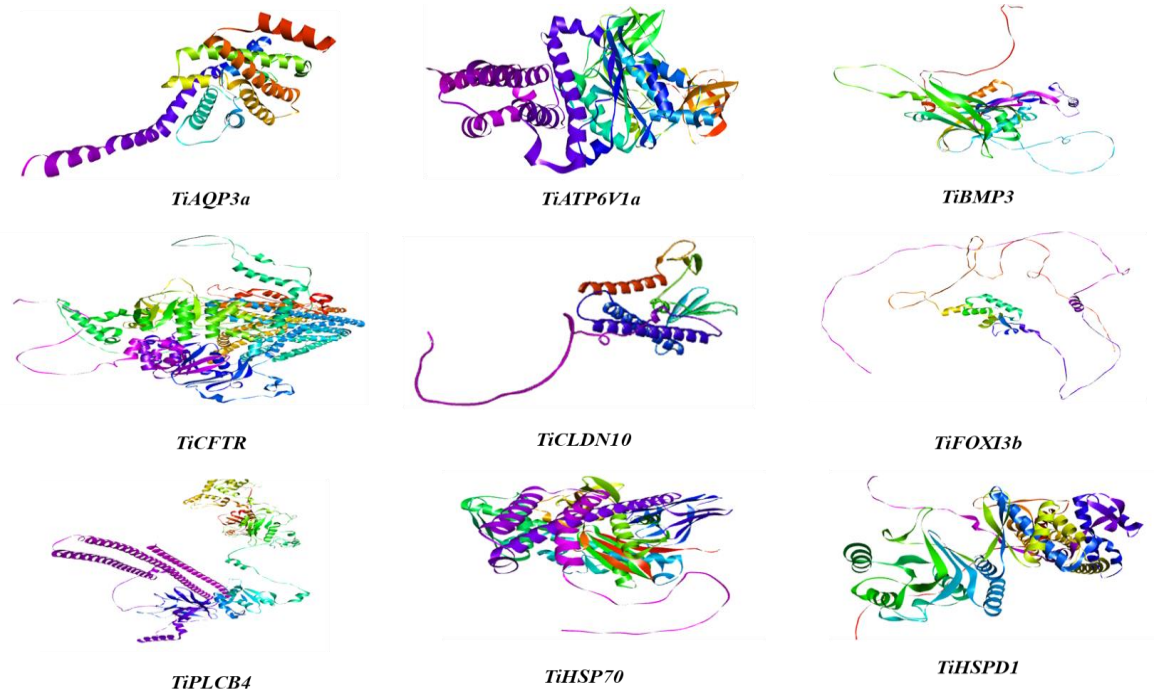

Fig. 5. Tertiary structure of the identified osmoregulatory and stress-tolerance protein in Hilsa

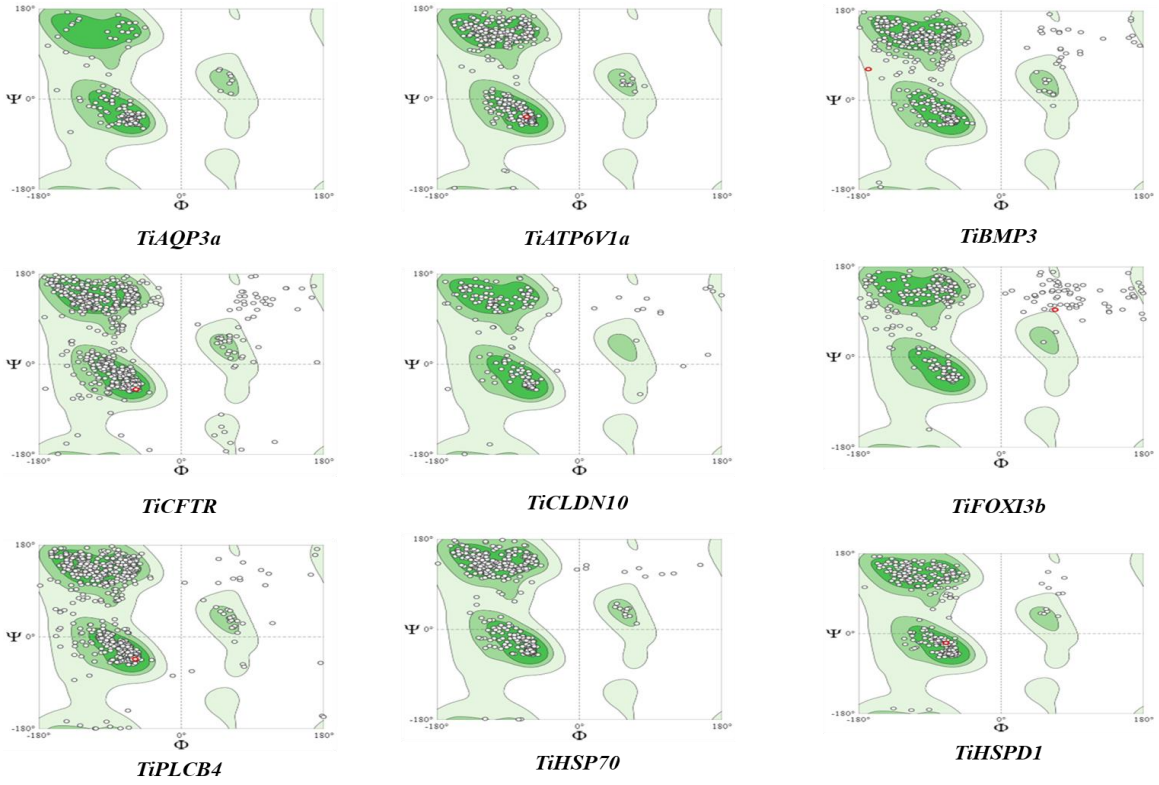

Fig. 6. Ramachandran plots presenting the structural analysis of selected proteins (*TiAQP3a*, *TiATP6V1a*, *TiBMP3*, *TiCFTR*, *TiCLDN10*, *TiFOXl3b*, *TiPLCB4*, *TiHSP70*, and *TiHSPD1*) with favorable regions highlighted

5. Confirmation of the predicted genes through PCR and sequencing

We have chosen nine genes at random for PCR amplification to verify the predicted genes. Gene-specific primers were designed and amplified. The genes from the DNA samples of Hilsa fish were collected from the different regions of Bangladesh. Analysis of the PCR results proved that the amplifications were successful. The existence of the anticipated bands (Fig. 7) supported the success of the osmoregulatory and stress-responsive genes' predictions. For further verification, all the PCR products were sequenced, and sequence data analysis (nBLAST) confirms the corresponding genes that were predicted from the draft genome of the Hilsa.

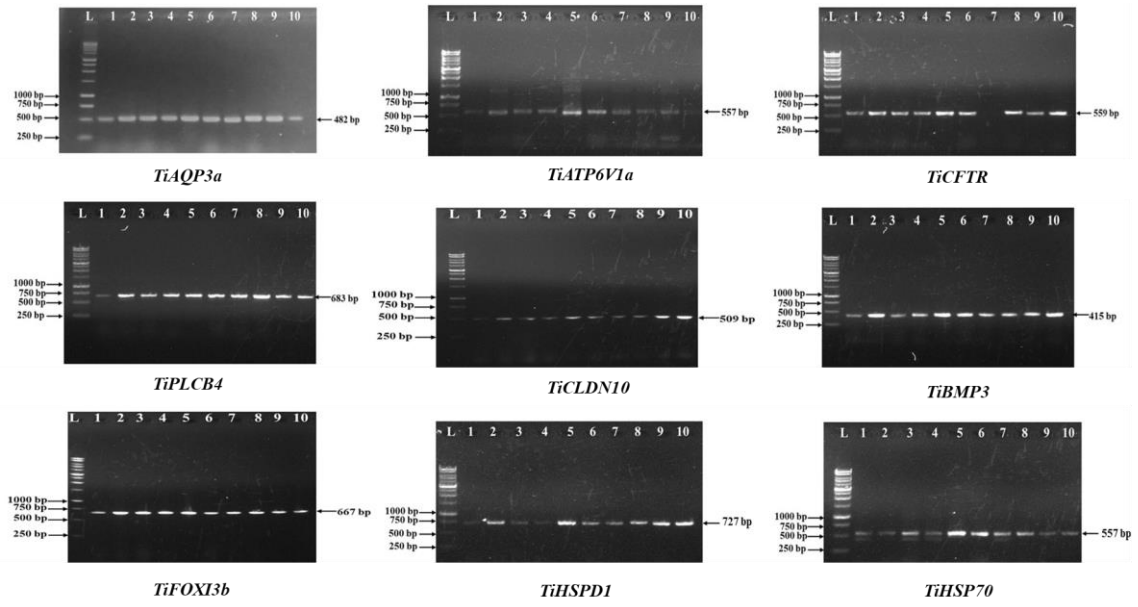

Fig. 7. Confirmation of the selected osmoregulatory and stress-tolerance genes through PCR using corresponding gene-specific primers

DISCUSSION

The present study aimed to identify and characterize osmoregulatory and stress-responsive genes in *T. ilisha*, an anadromous fish species crucial for maintaining homeostasis under varying environmental conditions. Through extensive bioinformatics analyses, including sequence prediction, gene structure analysis, motif conservation, secondary and tertiary structure modeling, and experimental validation via PCR, we successfully identified nine key genes associated with osmoregulation and stress tolerance. These findings provide significant insights into the molecular mechanisms underlying the adaptive responses of *T. ilisha* to diverse salinity conditions. Our study successfully identified nine osmoregulatory and stress-responsive genes through an extensive literature survey and bioinformatics analysis. Among these, seven genes

(*TiAQP3a, TiATP6V1a, TiBMP3, TiCFTR, TiCLDN10, TiFOXI3b,* and *TiPLCB4*) were found to be associated with osmoregulation, while the remaining two (*TiHSP70* and *TiHSPD1*) were involved in stress tolerance. The top BLAST hits for these genes suggested high sequence similarity with homologous genes from *Clupea harengus* and *Scleropages formosus*, confirming their evolutionary conservation. Notably, *TiATP6V1a* exhibited the highest identity (94.49%) with *S. formosus*, indicating its strong evolutionary stability **(Mohindra *et al.*, 2019)**. The functional annotation of the predicted genes using Gene Ontology (GO) analysis further confirmed their roles in key biological processes such as ion transport, cellular response to stress, and signal transduction pathways. The enrichment analysis revealed significant associations with pathways linked to ion transporter activity, suggesting their importance in osmoregulation. These findings are consistent with previous studies on other anadromous fishes, highlighting the significance of these genes in maintaining ionic balance during migration **(Xu *et al.*, 2012; Hossain *et al.*, 2019)**. The structural features of the identified genes were analyzed to determine their functional significance. The gene lengths varied widely, ranging from 1,945bp (*TiFOXI3b*) to 30,504bp (*TiCFTR*), indicating diverse gene architectures. The predicted protein lengths, molecular weights, and isoelectric points (pI) suggested that these proteins exhibit varied biochemical properties suited for different cellular environments. Subcellular localization analysis revealed that most of the proteins were associated with the plasma membrane (e.g., *TiAQP3a, TiCFTR, TiCLDN10*), while others were localized in the cytoplasm, nucleus, or mitochondria. These findings align with previous studies on osmoregulatory genes, which reported the involvement of membrane-associated proteins in ion transport **(Evans *et al.*, 2005)**. Gene structure analysis indicated diverse exon-intron patterns among the identified genes, supporting their evolutionary adaptations. Exon-intron organization plays a crucial role in gene regulation, and the observed variations suggest potential alternative splicing mechanisms that may contribute to functional diversity. The exon-intron structure of *TiCFTR*, a key chloride channel protein, showed a complex arrangement with 27 exons, consistent with findings in other teleost species **(Wong *et al.*, 2007)**. Motif conservation analysis revealed distinct yet highly conserved motifs among the identified genes. The presence of conserved motifs across multiple species underscores their evolutionary importance. Genes such as *TiAQP3a*, *TiATP6V1a*, and *TiHSP70* exhibited strong motif conservation, suggesting functional stability across evolutionary lineages. The statistical significance of motif conservation highlights the critical roles of these genes in physiological processes.

The secondary structure analysis revealed that alpha-helices dominated the structural composition of most proteins, followed by beta-sheets and coils. Alpha-helices are known to stabilize protein structures through hydrogen bonding, a feature particularly important for transmembrane proteins such as *TiCFTR* and *TiAQP3a*. The high alpha-helix content in these proteins suggests their role in maintaining structural integrity in lipid bilayers. Tertiary structure modeling provided further insights into the functional

conformation of the identified proteins. The structural models generated through SWISS-MODEL and refined using GalaxyRefine demonstrated high-quality predictions, as validated by Ramachandran plot analysis. More than 90% of the residues in all proteins, except *TiFOXI3b*, were located in the favored and allowed regions, indicating the reliability of the predicted structures. Additionally, ERRAT and ProSA-web validation confirmed that the constructed models exhibited high structural accuracy, suggesting their suitability for functional studies. The PCR amplification and sequencing of selected genes provided strong experimental validation for the computational predictions. The successful amplification of gene-specific fragments confirmed the presence of these genes in the *T. ilisha* genome. Subsequent sequencing and nBLAST analysis further verified their identity, reinforcing the accuracy of the bioinformatics predictions. These results underscore the reliability of *in-silico* approaches in identifying functionally relevant genes in non-model species. The identification and characterization of osmoregulatory and stress-responsive genes in *T. ilisha* provide valuable insights into the adaptive mechanisms of anadromous fish. The conservation of these genes across multiple teleost species suggests that osmoregulation and stress response mechanisms are evolutionarily conserved processes. The presence of conserved motifs and structural elements further supports the notion that these genes play critical roles in maintaining homeostasis under varying environmental conditions. The findings of this study align with previous research on teleost fish, where osmoregulatory genes such as CFTR and AQP3 have been implicated in maintaining ionic balance during salinity fluctuations **(Marshall & Grosell, 2006)**. The structural and functional similarities observed in *T. ilisha* suggest that these genes have undergone strong selective pressure to retain their essential functions in osmoregulation and stress adaptation.

CONCLUSION

In this study, *in silico* techniques were applied to predict and characterize a variety of Hilsa fish osmoregulatory and stress-responsive genes. We have identified 9 candidate genes using computational analysis and comparative genomics that are important in the osmotic regulation and stress responses of Hilsa shad. The identified genes consist of ion transporters, stress-related transcription factors, signaling molecules, and other functional categories that are important for environmental adaptations. This study will help to uncover a novel scope for subsequent experimental investigations to determine their functional roles. Additionally, it will reveal a deeper understanding of its biology and its adaptation to diverse environmental challenges.

Acknowledgment

The authors would like to acknowledge Grant of Advanced Research in Education (GARE); SUST Research Centre; and Bangladesh Ministry of Science and Technology

for financial support and the Department of Genetic Engineering and Biotechnology, Shahjalal University of Science and Technology, Sylhet-3114, Bangladesh, for providing laboratory support to conduct this research.

Funding

This work was financially supported by the Grant of Advanced Research in Education (GARE), Bangladesh Bureau of Educational Information & Statistics (BANBEIS), Bangladesh Ministry of Education, (Project ID: LS 2017 506); SUST Research Centre, Shahjalal University of Science and Technology, Sylhet, Bangladesh (Project ID: LS/2018/3/15); and Bangladesh Ministry of Science and Technology research project (Project ID: 39.00.0000.009.06.009.20-1331/BS-247)

Availability of data and material

The data that support the findings of this study are available from the corresponding author upon request.

REFERENCES

Ackerman, P.; R. Forsyth; C. Mazur and G. Iwama (2000). "Stress hormones and the cellular stress response in salmonids." Fish Physiology and Biochemistry 23: 327-336.

Ahsan, D.; N. Naser; U. Bhoumik; S. Hazra and S. Bhattacharya (2014). "Migration, spawning patterns and conservation of hilsa shad (Tenualosa ilisha) in Bangladesh and India."

Alam, A.; B. P. Mohanty; M. E. Hoq and S. Thilsted (2012). Nutritional values, consumption and utilization of Hilsa Tenualosa ilisha (Hamilton 1822). Proceedings of the regional workshop on hilsa: potential for aquaculture.

Barton, B. A. (2002). "Stress in fishes: a diversity of responses with particular reference to changes in circulating corticosteroids." Integrative and comparative biology 42(3): 517-525.

Begum, M.; S. Bhowmik; F. Juliana and M. S. Hossain (2016). "Nutritional profile of hilsa fish [Tenualosa ilisha (Hamilton, 1822)] in six selected regions of Bangladesh." J. Nutr. Food Sci 6(2): 567-570.

Bhaumik, U. (2013). "Decadal studies on Hilsa and its fishery in India-a review." J. Interacad 17(2): 377-405.

Chowdhury, M. A. A.; M. R. Islam; A. Amin; S. N. Mou; K. N. Ullah; A. Baten; M. Shoyaib; A. A. Ali; F. T. Chowdhury and M. L. Rahi (2023). "Integrated

transcriptome catalog of Tenualosa ilisha as a resource for gene discovery and expression profiling." Scientific Data 10(1): 214.

Deb, P.; M. M. Dey and P. Surathkal (2022). "Fish price volatility dynamics in Bangladesh." Aquaculture Economics & Management 26(4): 462-482.

Eissa, N. and H. P. Wang (2016). "Transcriptional stress responses to environmental and husbandry stressors in aquaculture species." Reviews in Aquaculture 8(1): 61-88.

Evans, D. H., Piermarini, P. M., & Choe, K. P. (2005). The multifunctional fish gill: dominant site of gas exchange, osmoregulation, acid-base regulation, and excretion of nitrogenous waste. Physiological Reviews, 85(1), 97-177.

Hossain, M.; S. Almatar and A. Al-Hazza (2014). "Proximate, fatty acid and mineral composition of hilsa, Tenualosa ilisha (Hamilton 1822) from the Bay of Bengal and Arabian Gulf." Indian J. Fish 61(2): 58-66.

Hossain, M.; S. Sharifuzzaman; S. Chowdhury and S. Sarker (2016). "Habitats across the life cycle of Hilsa shad (Tenualosa ilisha) in aquatic ecosystem of Bangladesh." Fisheries Management and Ecology 23(6): 450-462.

Hossain, M. A.; I. Das, L. Genevier; S. Hazra; M. Rahman; M. Barange and J. A. Fernandes (2019). "Biology and fisheries of Hilsa shad in Bay of Bengal." Science of the Total Environment 651: 1720-1734.

Hu, B.; J. Jin; A.-Y. Guo; H. Zhang; J. Luo and G. Gao (2015). "GSDS 2.0: an upgraded gene feature visualization server." Bioinformatics 31(8): 1296-1297.

Islam, M. J.; A. Kunzmann and M. J. Slater (2022). "Responses of aquaculture fish to climate change-induced extreme temperatures: A review." Journal of the World Aquaculture Society 53(2): 314-366.

Iwama, G. K.; M. M. Vijayan; R. B. Forsyth and P. A. Ackerman (1999). "Heat shock proteins and physiological stress in fish." American Zoologist 39(6): 901-909.

Jahan, I.; D. Ahsan and M. H. Farque (2017). "Fishers' local knowledge on impact of climate change and anthropogenic interferences on Hilsa fishery in South Asia: evidence from Bangladesh." Environment, Development and Sustainability 19: 461-478.

Kültz, D. (2015). "Physiological mechanisms used by fish to cope with salinity stress." The Journal of Experimental Biology 218(12): 1907-1914.

Mahmud, Y. (2020). Hilsa Fisheries Reseach and Development in Bangladesh, Bangladesh Fisheries Reseach Institute.

Malakpour Kolbadinezhad, S.; J. Coimbra and J. M. Wilson (2018). "Osmoregulation in the Plotosidae catfish: role of the salt secreting dendritic organ." Frontiers in Physiology 9: 761.

Marshall, W. S., & Grosell, M. (2006). Ion transport, osmoregulation, and acid-base balance. The Physiology of Fishes, 3, 177-230.

Mohammed, E. Y.; L. Ali; S. Ali; B. Hussein; M. A. Wahab and N. Sage (2016). Hilsa's non-consumptive value in Bangladesh: estimating the non-consumptive value of the hilsa fishery in Bangladesh using the contingent valuation method, International Institute for Environment and Development.

Mohindra, V.; L. M. Chowdhury; N. Chauhan; A. Paul; R. K. Singh; B. Kushwaha; R. K. Maurya; K. K. Lal and J. Jena (2023). "Transcriptome analysis revealed osmoregulation related regulatory networks and hub genes in the gills of Hilsa shad, Tenualosa ilisha, during the migratory osmotic stress." Marine Biotechnology 25(1): 161-173.

Mohindra, V.; T. Dangi; R. K. Tripathi; R. Kumar; R. K. Singh; J. Jena and T. Mohapatra (2019). "Draft genome assembly of Tenualosa ilisha, Hilsa shad, provides resource for osmoregulation studies." Scientific reports 9(1): 16511.

Mozumder, M. M. H.; A. Pyhälä, M. A. Wahab; S. Sarkki; P. Schneider and M. M. Islam (2019). "Understanding social-ecological challenges of a small-scale hilsa (Tenualosa ilisha) fishery in Bangladesh." International journal of environmental research and public health 16(23): 4814.

Murshed-e-Jahan, K.; B. Belton and K. K. Viswanathan (2014). "Communication strategies for managing coastal fisheries conflicts in Bangladesh." Ocean & Coastal Management 92: 65-73.

Rahman, M.; M. Wahab; M. Nahiduzzaman; A. Haque and P. Cohen (2020). Hilsa fishery management in Bangladesh. IOP Conference Series: Earth and Environmental Science, IOP Publishing.

Wong, C. K., Chan, D. K., & Chan, H. K. (2007). Cloning and expression of Cl(-) channel in Japanese eel gills. American Journal of Physiology-Regulatory, Integrative and Comparative Physiology, 293(1), R205-R215.

Xu, C., Guo, Y., Wang, X., & Xu, Y. (2012). Evolutionary conservation of exon-intron structure in osmoregulatory genes. BMC Evolutionary Biology, 12(1), 25.