scientific reports

Check for updates

OPEN

An insight into comparative sex based proteomic profiling of Hilsa *Tenualosa ilisha* towards domestication and aquaculture

Hena Chakraborty[1], Hirak Jyoti Chakraborty[1], Basanta Kumar Das[1]✉ & Joydev Maity[2]

This study investigates the serum proteome dynamics of male Hilsa (*Tenualosa ilisha*), a species of immense ecological and economic value, to support its conservation, aquaculture, and long-term sustainability. Using advanced LC-MS/MS-based proteomic techniques, we identified proteins across six weight groups, with the highest number (817) found in 201–300g individuals. Interestingly, protein counts did not follow a linear trend with body weight; the 401–500g group had the lowest number (84), suggesting complex physiological transitions during development. The proteome data were submitted to the ProteomeXchange Consortium via PRIDE. Bioinformatic analysis revealed proteins associated with key immune-related gene ontology (GO) terms and enriched KEGG pathways, including Neuroactive ligand-receptor interaction, phagosome, FoxO signaling, Apelin signaling, and herpes simplex virus interaction. Several immune and inflammation-related proteins showed age- and sex-dependent expression patterns. Comparative analysis of male and female Hilsa at various developmental stages revealed commonly expressed proteins like Alpha Macroglobulin, Fibulin, MEF2C, and complement proteins. Male-dominant expression was observed for MADS-box, anaphylatoxin, and cytosolic non-specific dipeptidase. These findings provide critical insights into sex-specific proteomic adaptations and immune function during freshwater migration and offer a valuable foundation for future Hilsa, domestication, and aquaculture.

Keywords Hilsa, Proteomics, Protein, STRING, LC-MS/MS, Network analysis, KEGG

Migratory fish undertake movements, either individually or in schools, from their natal habitats to growth regions, optimizing reproductive success. Typically, these fish return to their place of origin to spawn. Such migrations play a critical role in regulating population density, expanding species distribution, exploiting diverse food sources, and adapting to changing environmental conditions[1]. Among the significant migratory fish species in the Indo-Pacific region is *Tenualosa ilisha* (Hilsa), a member of the class Actinopterygii, order Clupeiformes, sub-family Alosinae, and family Clupeidae. The Hilsa is particularly abundant in the waters of Bangladesh and the Indian state of West Bengal, notably in the Hooghly River—part of the Ganges system—and its coastal zones, as well as in the Padma River, the primary distributary of the Ganges in Bangladesh. Esteemed for its taste, flavour, and culinary qualities, the Hilsa holds significant economic value in these regions[2].

Proteins are acknowledged as a key and comprehensive tool for understanding biological systems and are responsible for producing energy, actions, relationships, structure, and cell division[3]. Proteins in a particular biological system are assessed on a large scale using proteomic analysis. In addition to figuring out the structure and function of proteins, this study also involves quantifying them and identifying their intracellular location, changes, interactions, and abundance[4]. It is commonly known that the majority of biological processes are controlled by protein complexes rather than single proteins and that the way these proteins interact may determine a specific biological function[5]. The molecular mechanisms behind these changed proteins or their expressions can be easily interpreted by using pathway analysis to compact a large number of proteins onto a small list of pathways knowledge[6]. Protein-protein interactions are the subject of all of these and experimentally determined or anticipated data regarding protein interactions is placed in a number of interaction databases. Fish is a major part of the regular diet in many countries, accounting for almost two-thirds of the world's population and providing roughly 40% of the total protein consumed. Although all fish species are relevant, only a few have had their physiology thoroughly examined. Fish serum protein molecular characterisation is now regarded as

[1]ICAR-Central Inland Fisheries Research Institute, Barrackpore 700120, West Bengal, India. [2]Department of Fisheries Science, Vidyasagar University, Midnapore 721102, West Bengal, India. ✉email: basantakumard@gmail.com

essential in understanding the function of particular proteins associate in a range of physiological and metabolic processes[7,8]. A vital component of an organism is its proteome, which includes post-transcriptional and post-translational protein expression regulation and provides information on overall gene expression[9]. Thus, proteomics can map an organism's adaptive potential and provide an image of its present state of being[10].

The application of proteomic technologies in fish biology has increased dramatically in the last decade. The relevance of signaling events during development in a rapidly changing environment is highlighted by proteomic alterations that occur during distinct life stages of fish species or by studying biological systems and their behaviour under diverse conditions[11]. Over the past several decades, sex-specific proteomic studies in fish have been widely explored, particularly through the use of two-dimensional gel electrophoresis, which was a popular technique at the time. For instance, conducted a comparative proteomic analysis of male and female gonads in Persian sturgeon (*Acipenser persicus*)[12]. Similarly, as early as four decades ago, the immunochemical identification of the female-specific serum protein, vitellogenin, in medaka fish[13]. Over the past ten years, tremendous advancements in technology have been made in proteomic platforms based on liquid chromatography-tandem mass spectrometry (LC-MS/MS), which provides the ability to identify and quantify thousands of proteins from any biological material, including fish species[14,15]. Recent studies employing serum and plasma proteomics in teleosts have revealed sex- and stage-specific protein expression patterns linked to reproductive physiology and biomarker discovery[8,16–19]. Additionally, sample pooling techniques are used to rationalize the overall number of samples in large trials or to lessen biological variance when the number of samples is limited[20,21]. Nevertheless, very few published studies on the protein profiling of fishes in the Clupeidae family[22]. Remarkably, no extensive protein profile of this economically significant fish species has been reported. This may be related to the fact that many fish species in the Clupeidae family lack entire genome sequences. Additionally, author has already published a detailed analysis of female Hilsa serum proteome sequence[19], which supports us in developing a successful Hilsa protein sequence database submitted to ProteomeXchange Consortium via PRIDE. Therefore, the primary goal of this research was to generate a proteome profile of Hilsa male serum in different weight groups and a comparative functional profile of male and female hilsa[19]. This study offers the first comprehensive shotgun proteomic analysis of Hilsa serum, revealing key insights into its immune system, metabolism, and physiological adaptations. It highlights significant variations in proteome functions based on weight, sex, and growth stages, contributing to a better understanding of the species' migration mechanisms. These findings pave the way for future conservation and domestication efforts of Hilsa.

Materials and methods

Ethics statement

The Central Inland Fisheries Research Institute (ICAR) ethical committee at Barrackpore, Kolkata 700120, West Bengal, India, authorised this work. The Committee on Animal Ethics approved this experiment. (Approval date: 10/09/2019; approval code: CIFRI/EC/2019/62) The Institute's defined protocols were followed when handling the fish used in the experiment. The experimental design, sample collection, and fish sacrifice were conducted in accordance with ethical standards and relevant national guidelines and regulations. Furthermore, all methods reported in this study comply with the ARRIVE guidelines.

Sample collection and processing

Wild male Hilsa (*T ilisha*) samples of different size groups (Table 1) were collected from the Ganga River at Farakka, Murshidabad, West Bengal (24°48′1.34″ N; 87°55′21.53″ E) during August to November 2021 using a specially designed lift net. Fish were anaesthetised (No euthanasia) with MS-222 (120 mg/L water) following the protocol[23]. A total of 18 (n= 3 each weight group) male Hilsa samples from various weight groups were selected for the study. Males, typically smaller than females during the intersex stage, were identified by examining abdominal shape or through dissection and observation. During dissection, a thin, thread-like transparent gonad was observed in smaller individuals, that was identified as testes, confirming them as males. In contrast, females exhibited slightly broader, yellowish lobes, characteristic of developing ovaries. Blood was drawn from the caudal vasculature using a 2 mL sterile syringe, following the standard protocol outlined by[24]. The sampling procedure and year classes were adopted as per our previously published paper[19,25]. Briefly, the Hilsa blood was transferred to 1.5 ml Eppendorf tubes and coagulated for 2–3 hours at room temperature. Serum samples were then transported to laboratory in an ice box and stored at −40°C for further analysis. Three samples from the same weight group were pooled and processed for proteomic analysis using the LC-MS/MS method. For comparative purposes, protein data from female Hilsa, previously reported by[19], were utilised to examine differences between male and female proteomes with the dataset identifiers PXD054275, PXD054224, PXD054225, and PXD054228,

Hilsa Sample	Weight range (g)	Avg ± SD	Length range (cm)	Avg ± SD	Total proteome (no.)	UniProt mapping (no.)
H4	501-above	659.2 ± 85.3	42.5–47.9	45.2 ± 2.65	64,248	691
H22	401–500	455 ± 38.6	39.8–40.8	40.2 ± 0.50	2293	84
H37	301–400	354.1 ± 41.1	33.7–34.9	34.3 ± 0.61	25,746	336
H35	201–300	242.3 ± 31.9	25.4–27.8	26.6 ± 1.20	81,333	817
H42	101–200	155.6 ± 12.0	18.9–21.1	20.1 ± 1.11	22,915	301
H80	50–100	76.3 ± 10.2	15.3–12.8	13.8 ± 1.30	20,171	326

Table 1. LC-MS and UniProt summery.

the mass spectrometry proteomics data of female Hilsa (H3, H28, H29, and H34) have been uploaded to the ProteomeXchange Consortium through the PRIDE.

LC-MS/MS

The ACQUITY UPLC apparatus from Waters Corporation in the US was used for liquid chromatography. Following dilution with 100 mM NH4HCO3, the sample (100 µg) was subjected to a 1-hour treatment with 250 mM DTT (dithiothreitol) at 95 °C and a 45-minute dark treatment with 250 mM IDA (iminodiacetic acid) at room temperature. Samples were digested in trypsin, and a Waters-made ACQUITY UPLC BEH C18 column (150mm X 2.1mm X 1.7m) was used to separate the trypsinised peptide solution. The chromatographic separation has been completed using a gradient elution procedure with mobile phases A and B (0.1% formic acid in water and acetonitrile, respectively). A SYNAPT G2 QTOF (Waters, USA) equipped with an ESI source was utilised for mass spectrometric detection. The starting mass is 300,000 Da, and the final mass is 1700,000 Da for the acquisition mass range. Protein Lynx Global Server (PLGS) 3.0.2, a database search engine (Fig. 1), was used to handle the raw mass spectrometry data before being sent via PRIDE to the ProteomeXchange Consortium[26–28].

Proteomics data interpretation and functional profiling

Advanced techniques for interpreting high-throughput proteomics data were employed to map and annotate mass spectrometry-based proteomics identifiers using the UniProtKB fish proteome sequence database (www.uniprot.org). STRING (https://string-db.org/) was used to analyze protein interaction networks., allowing insights into the functional dynamics of enriched proteins. The Danio rerio STRING background database was used for comparative analysis with the in-house Hilsa proteome dataset interactions with a confidence score ≥0.7 (high confidence) were included. Network edges were derived from experimental and co-expression evidence, using Danio rerio as the reference organism. Proteins detected in at least two biological replicates per group were considered reliable nodes in the PPI network. Network topology was examined using Cytoscape (v3.10.2), enabling visualisation and exploration of protein interactions. Gene ontology (GO) annotations—covering "molecular functions," "cellular components," and "biological processes" and additionally KEGG pathway analysis was also conducted with significant terms identified at an adjusted p-value <0.05 (Benjamini–Hochberg correction) and comparative study representation through butterfly graph—were performed using R packages (version 4.5.0). To visualise pathway-level comparisons were illustrated using a butterfly graph, representing the relative enrichment of pathways between male and female Hilsa datasets.

Comparative statistical analysis between male and female

Statistical studies were conducted using the R packages (version 4.5.0) to generate violin plots, scatter plots, and Spearman's rank correlation. All correlations and visualizations were created to evaluate the connections between distributions (the top 15 frequently occurring proteins) and variables (male and female hilsa).

Results

Across six male groups, a total of 2,555 non-redundant proteins were identified. Although the number of detected proteins varied slightly by weight class (ranging ≈ 84–691), the overall qualitative composition remained highly conserved. A total of 81333 (maximum) peptide-matched spectra were found in 201-300g male hilsa, these data were further filtered (after removing duplicates) in UniProt mapping and the maximum number of annotated proteomes quantified were 817 (maximum). The lowest nos. of annotated proteome (84) were found in 401-500g hilsa. These proteome profiles (Supplementary file. 1) were analysed based on their PLGS score, peptides, coverage (%), molecular weight, etc. The major identified proteins were Alpha Macroglobin, fibulin, MEF2C, complement proteins (C3, C4, C5), albumin, serum response factor (SRF), anaphylatoxin, CD109, MADS-box domain and cytosolic non-specific dipeptidase protein found in male Hilsa. Fibulin, C3, serum response factor and Alpha Macroglobin were primarily dominant. The core male serum proteome, shared across all groups including H22, comprised 142 proteins, dominated by immune and carrier molecules such as Alpha-2-Macroglobulin, Apolipoproteins A-I/A-IV, Complement C3–C5, Fibulin, and Serum Response Factor. These are typical of teleost serum and reflect the essential metabolic and immunological framework. A small number of proteins showed variability in observation frequency, notably in male- MADS-box, cytosolic, and anaphylatoxin which were indicating stage-dependent immune and developmental regulation, also called variable proteins.

A Venn diagram (Supplementary Fig. S1) visualises shared and stage-specific proteins. The smaller juvenile groups (H80–H42) expressed Cytosolic Non-Specific Dipeptidase and MADS-box proteins, possibly linked to early growth and cellular differentiation. Intermediate groups (H35–H37) showed a higher frequency of Fibulin and CD109, related to tissue remodelling, while mature males (H4) displayed enrichment of Complement C3–C5 and Alpha-2-Macroglobulin, consistent with immune maturation.

Functional annotation of proteomics and gene ontology

In this study, Gene Ontology (GO) served as a functional classification system, providing detailed insights into the properties of genes and gene products in Hilsa. This GO technique is categorised into three forms, viz. biological processes, molecular function and cellular components, to complete an understanding of the functionality of regulatory processes. The top-hit categories from the assembly annotation were filtered and are illustrated in Figures 3, 4, and 5. The enrichment values, nGenes (No. of genes) and -$\log_{10}$ (FDR) were also calculated and analysed.

The top 20 biological functions were presented in the male Hilsa samples (Fig. 2). Complement activation showed the highest –$\log_{10}$ FDR score in most of the samples, except for H37 and H80, and was associated with the maximum number of genes. In contrast, proteolysis, protein metabolic processes, and regulation of

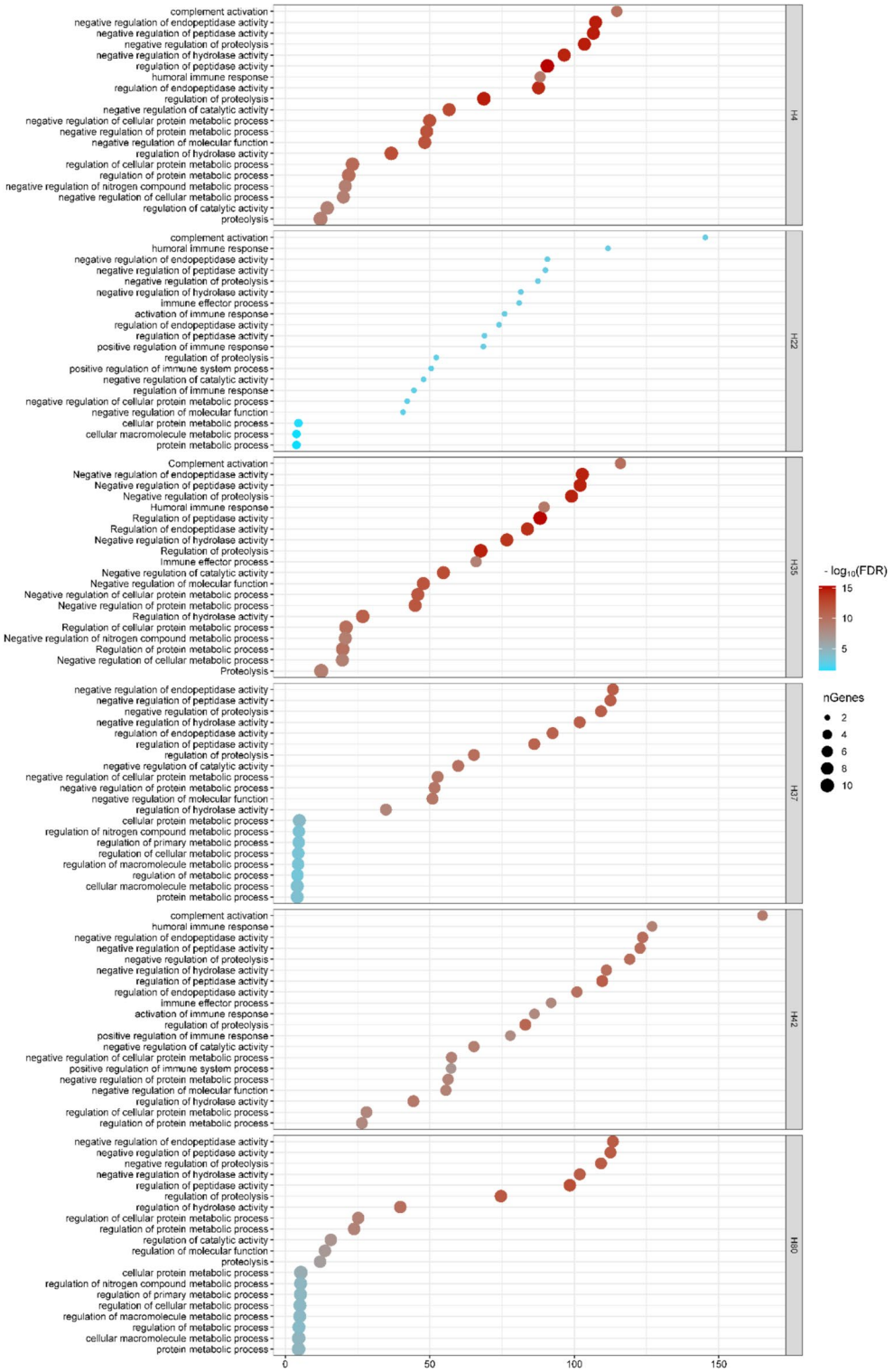

Fig. 1. Biological function of male hilsa (H4- 501g- above, H22- 401-500g, H37- 301-400g, H35-201-300g, H42-101-200g, H80-50-100g). The red color defined highest -log10 FDR score and blue defined the lowest. Size of the circles showed no of genes.

protein metabolic processes exhibited relatively low $-\log_{10}$ FDR scores (False Discovery Rate = 5.0, respectively). Other significant biological processes included negative regulation of endopeptidase activity, included negative regulation of peptidase activity, negative regulation of proteolysis, Humoral immune response etc. (Fig 1)

In the molecular function (MF), the male with the dominant -log10 FDR score was negative regulation of endopeptidase activity, negative regulation of peptidase activity, negative regulation of proteolytic activity,

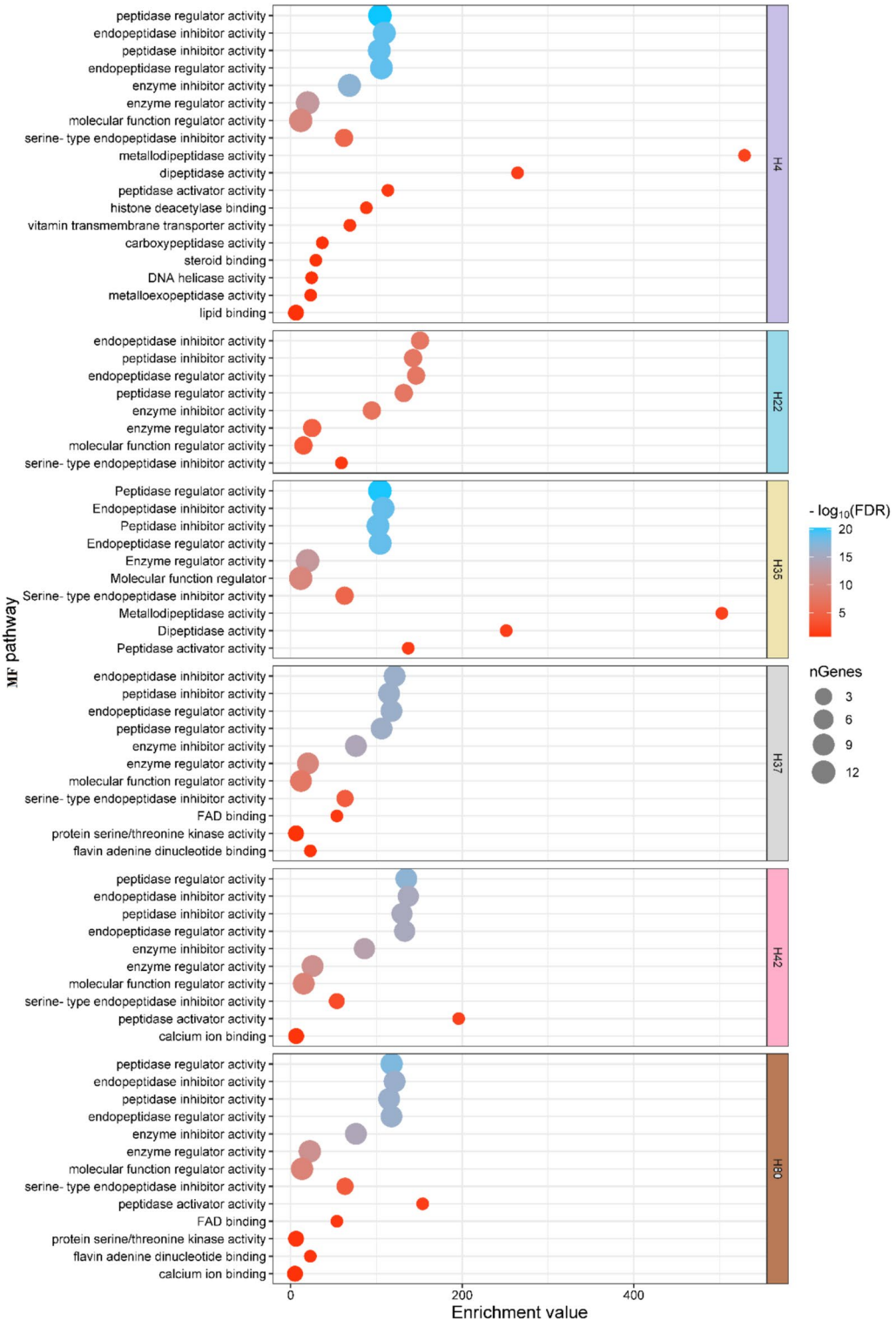

Fig. 2. Molecular function (MF) of male Hilsa (H4- 501g- above, H22- 401-500g, H37- 301-400g, H35-201-300g, H42-101-200g, H80-50-100g). The blue color defined highest -$\log_{10}$ FDR score and red defined the lowest. Size of the circles showed no of genes.

negative regulation of hydrolase, regulation of peptidase activity, regulation of proteolysis, protein metabolic process, cellular macromolecule metabolic process, regulation of the primary and metabolic process, regulation of the cellular metabolic process, regulation of nitrogen compound metabolic process, complement activation etc. In overall sample, the first four mentioned above were the highest-$\log_{10}$ FDR score, and a maximum number of nGenes was noticeable. In lower (H80) and middle weight groups (H22 and H37) of male Hilsa expressed

lower $\log_{10}$ FDR score in metabolic processes and functions (Fig. 1). In this study, 400 g Hilsa(H22) showed lower numbers of genes and lower -$\log_{10}$ FDR score

In the case of the cellular component (CC), only five samples were found, and extracellular space and extracellular region were enriched and mostly found to have the highest -log10 FDR value and a maximum number of nGenes. Except for sample H4, which had a significantly higher weight, and H35 (Fig. 3), with a medium weight, three additional GO terms—collagen-containing extracellular matrix, external encapsulating structure, and extracellular matrix—exhibited low -log10 FDR values and a minimal number of nGenes in the other three samples.

In the KEGG pathway study, there were nine pathways were enriched. Among them, four common pathways were enriched in all the samples of Hilsa male, which were Herpes simplex virus, Phagosome, FoxO and Neuroactive ligand-receptor interaction pathway (Fig. 4). In the larger fish, all the pathways were enriched. The

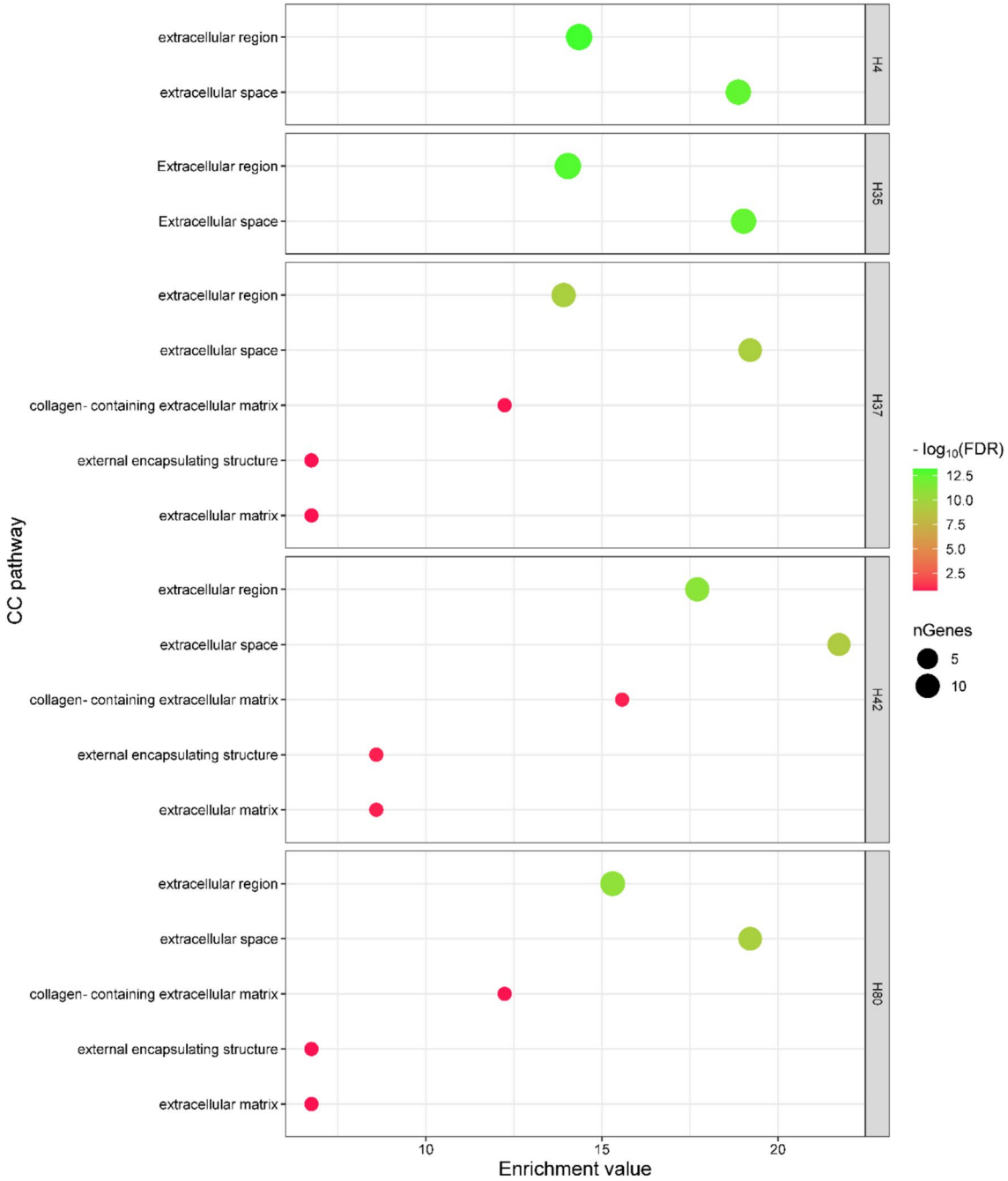

Fig. 3. Cellular component (CC) of male Hilsa (H4- 501g- above, H22- 401-500g, H37- 301-400g, H35-201-300g, H42-101-200g, H80-50-100g). The green color defined highest -$\log_{10}$ FDR score and pink defined the lowest. Size of the circles showed no of genes.

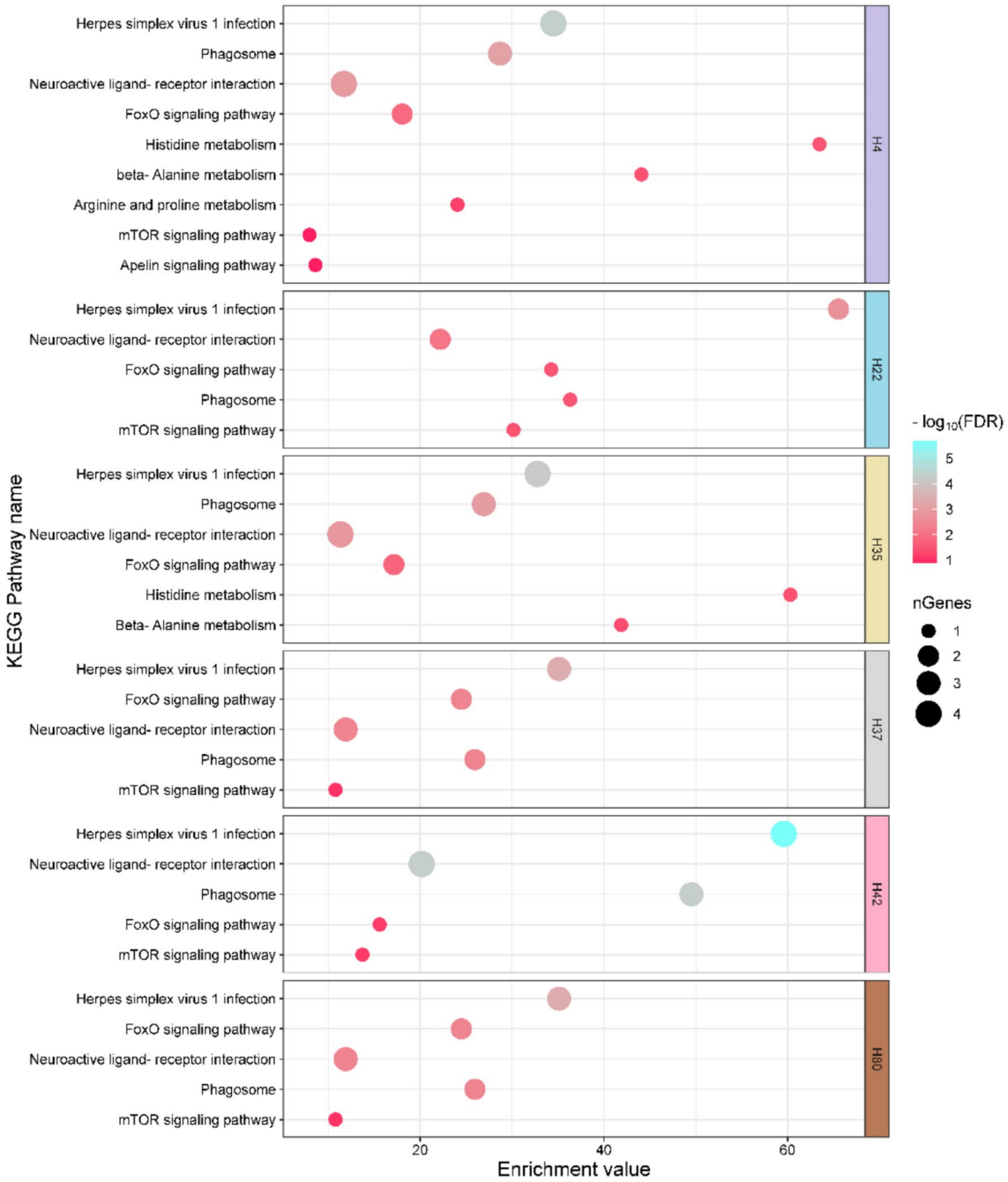

Fig. 4. KEGG pathway analysis of male Hilsa (H4- 501g- above, H22- 401-500g, H37- 301-400g, H35-201-300g, H42-101-200g, H80-50-100g); color showed -$\log_{10}$(FDR) score and circle size showed no. of genes. Blue color showed highest -$\log_{10}$ (FDR) score and pink denoted lowest.

Herpes simplex virus had the maximum -$\log_{10}$ FDR value and nGenes. Among all mTOR signaling pathway had lowest -log10 FDR value and nGenes.

Comparative functional profiling of male and female proteome

For the purpose of comprehending the molecular mechanisms of functions, the proteome approach was used to identify and thoroughly characterize the protein content of the serum of male and female Hilsa fish. LC-MS/MS was applied to map the protein content and compare proteins of both genders. Male datasets (n = 6) and female datasets (n = 4) were kept as separate groups during analysis to preserve biological variability. Proteins observed in at least two datasets from each group were retained for comparative evaluation. The results showed that the female (952 nos.) (using the data from Chakraborty et al., 2024), have more frequently observed serum protein than male (817 nos.) through UniProt mapping. In the case of males, two more weight groups were added (Below 100 g and above 500 g) to cover a more exhaustive study area than before in females[19]. In this gender

comparative serum proteome analysis, maximum revealed proteins were common, but in the case of male MADS box, anaphylatoxin and Cytosolic non-specific dipeptidase, CD109, MEF2C-like proteins, which were very less quantity found in female, were dominant in male. Comparison between pooled male and female datasets identified no statistically significant qualitative differences after correction, indicating that both sexes share a highly conserved immune-dominant serum proteome. Nevertheless, several proteins differed in observation frequency: MADS-box, Anaphylatoxin, and CD109 were more frequent in males, whereas Vitellogenin-like and lipid-transport proteins were relatively more frequent in females.

In the comparative KEGG pathway study, the highest weight group of Hilsa males comprised all nine pathways. In contrast, females with the same weight group comprised only six and also, the $-\log_{10}$ FDR score and fold enrichment value were highest in males. Like the Herpes simplex virus 1 infection, the $\log_{10}$ FDR score and fold enrichment value of males was 4.84E-05 and 34.5, but for females, it showed 0.0082 and 25.211, respectively. In the same weight groups of males and females, the same pathways, like Histidine metabolism and beta-alanine metabolism, were absent in 300-400g in Hilsa males, which is present in females, but mTOR was absent in females. In the weight range 200-300g, females (three) comprised half of the KEGG pathway than males (six), and the common pathways were Herpes simplex virus 1 infection, Phagosome, and Neuroactive ligand-receptor interaction. Most interestingly, it was found that in the lower weight group of Hilsa there was no common pathways were found (Fig 5).

Comparative analysis of PPI networks

Protein-protein interaction (PPI) networks provide critical insights into the molecular mechanisms underlying biological processes. The comparative analysis of male and female serum protein networks in Hilsa (*Tenualosa ilisha*) reveals notable differences in protein-protein interaction clusters, functional annotations, and biological implications.

Male Hilsa exhibits eight clusters, including a larger complement activation cluster (24 genes), alpha-2-macroglobulin family proteins, and NGF-stimulated transcription, indicating heightened immune activity

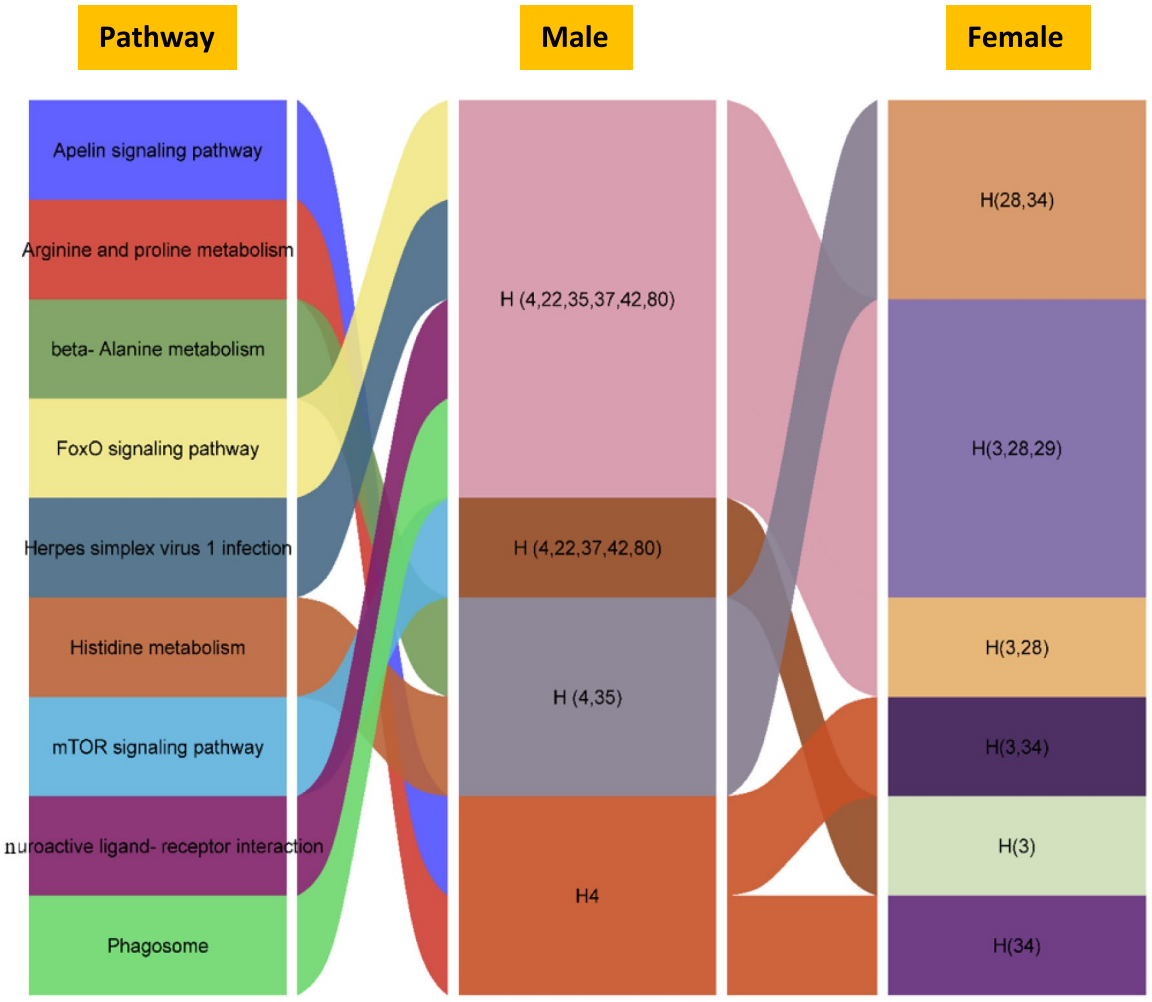

Fig. 5. Comparative KEGG pathway analysis of male (H4- 501g- above, H22- 401-500g, H37- 301-400g, H35-201-300g, H42-101-200g, H80-50-100g) and female H3- 401-500g, H28- 301-400g, H29- 201-300g, H34- 100-200g[19].

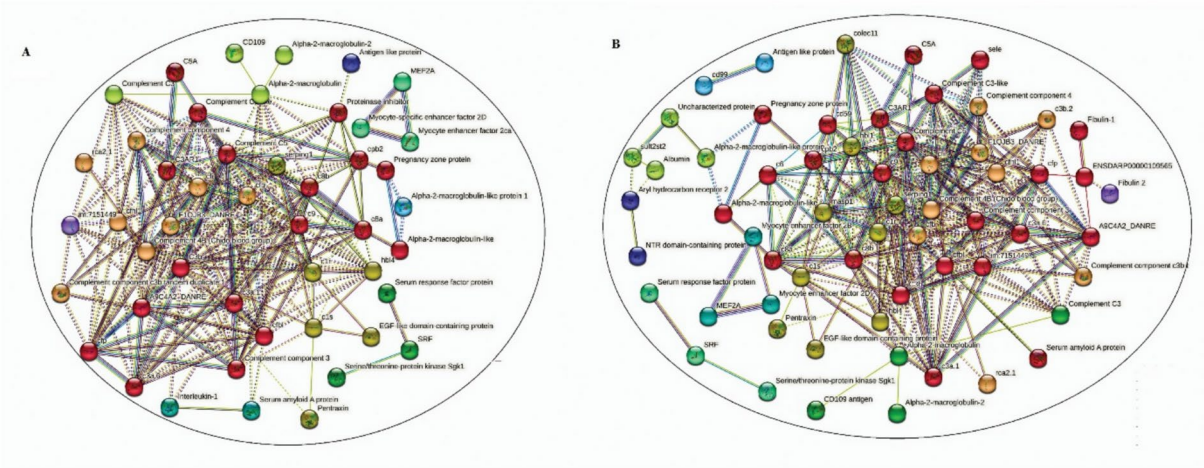

Fig. 6. Protein-Protein Interaction of Hilsa serum proteome -A. Male, B. Female [19]; each color showed one cluster of the network.

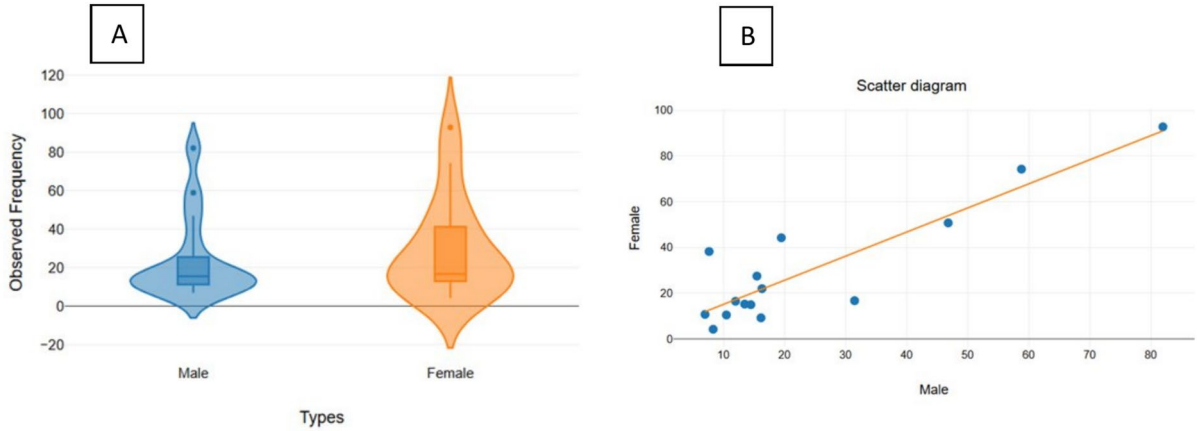

Fig. 7. (**A**) Violin plot (**B**) scatter diagram- comparing the observer frequency distribution between male and female[19].

and stress response mechanisms (Fig. 6A). Additionally, the number of nodes was 76, average local clustering coefficient – 0.508. In contrast, female Hilsa[19] (Fig. 6B), displays ten clusters, highlighting acute phase responses, pentraxin-related proteins, and interleukin pathways, reflecting a broader serum protein repertoire and readiness for systemic inflammation, additionally, the number of nodes was 45, average local clustering coefficient – 0.421 and both the cases PPI enrichment p-value was same (<1.0e-16). Both sexes share key proteins such as C3, C5, and alpha-2-macroglobulin but differ in interaction density and clustering. Females exhibit additional antigen-like proteins and more diverse interaction networks. These findings underscore sex-specific immune strategies, with males emphasizing complement cascade activation and females focusing on acute phase proteins and broader immune modulation. Further experimental validations and environmental impact studies are warranted to elucidate these molecular mechanisms and their adaptive significance in *T. ilisha*.

Comparative statistical analysis of male and female Hilsa

The violin plot (Fig. 7A) provides information on distribution and variability by combining a box plot with a density trace. The distribution is more concentrated around lower observed frequency values and was comparatively narrower for male Hilsa (blue plot). There was less variety since the density was more compressed. The distribution was wider and more dispersed for females (orange plot), indicating higher variability in observed frequencies. There were some severe values and a larger range in the density trace. In comparison to males and females had more diverse and expansive distribution of observed protein frequencies. This result implied that whereas males exhibit a more consistent pattern but females had more individual variability in the assessed characteristic. A linear regression analysis (Fig. 7B) was performed to examine the influence of the variable male on the variable female.

The regression model showed that the variable female explained 79.83% of the variance from the variable male Hilsa. ANOVA found that the effect was significantly different from zero, F=51.44, p = <.001, R2 = 0.8.

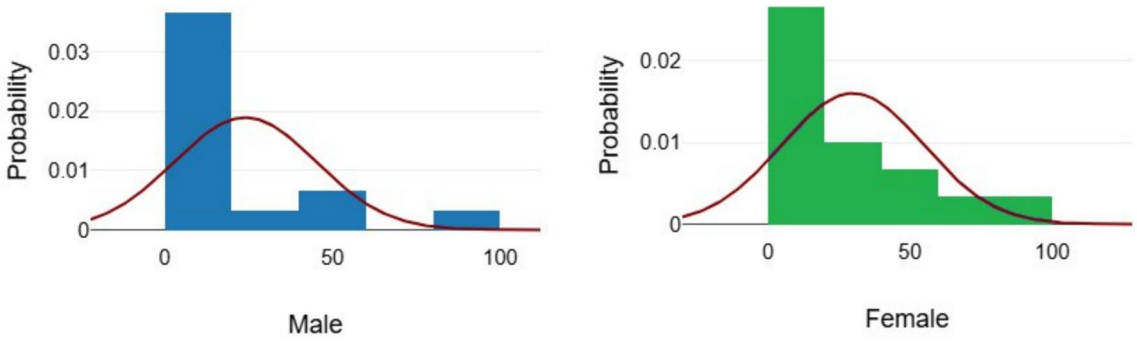

Fig. 8. Distribution of Hilsa proteins observed frequencies in the male and female.

Overall, the scatter plot shows that there was a general correlation between higher observed frequencies in females and higher observed frequencies in males. This favorable correlation may indicate that both groups are influenced by a common biological or environmental source.

A Spearman correlation presents (Fig. 8) two probability histograms, one for males (left, blue) and one for females (right, green), showing the distribution of observed frequencies. A fitted density curve (red line) is superimposed on each histogram to illustrate the overall distribution trend. The result of the Spearman correlation showed that there was a significant correlation between male and Female, $r(13) = 0.69$, $p = .004$. The distributions of observed frequencies for both males and females are positively skewed, with most observations taking place at lower frequency levels. However, females have a wider and more varied distribution, indicating that the assessed attribute is more heterogeneous among females.

Discussion

During this current study, proteomic profile of different weight groups of males and compared male-females of *T. ilisha* were presented here. Due to a paucity of resources in the fish proteins functional annotation database, the functional analysis of the Hilsa proteome collection was cross-validated by two different bioinformatics platforms utilising the Zebrafish gene reference as a thoroughly studied model fish organism[29]. This study identified a wide range of proteins with diverse functions, indicating the complexity of Hilsa serum proteome and its role in physiological, metabolic, and immune regulation. Interestingly, the maximum number of peptide-matched spectra and the proteomics study reveals that the male Hilsa had a maximum of 817 annotated proteins, indicating higher protein levels than the female (691). This is a remarkable finding that the 200 g Hilsa had a maximum of annotated proteins, whereas females with 500 g Hilsa contained a maximum[19]. As a result, the number of novel proteins did not follow a notable trend as weight increased. Several common key proteins, including Alpha Macroglobulin, Fibulin, Serum Response Factor (SRF), Complement components (C3, C4, C5), and CD109, were consistently dominant. The top 10 frequent proteins found in the male-female[19] were the same and complement proteins were dominant[19]. These proteins are widely recognized for their roles in complement activation, immunological defense, and extracellular matrix organization, are essential for surviving in dynamic aquatic environments. These represent the core serum proteome, crucial for metabolic regulation and immune defense, confirming their essential role in basic physiological processes. However, in the case of males, three more proteins were predominant: MADS-box, cytosolic, and anaphylatoxin. Three different processes in the mammalian system carry out target recognition and creating a protease complex for C3 activation. The lectin pathway is similar to the classical pathway, but C1 is replaced by a lectin–protease complex containing the mannose-binding lectin (MBL) or ficolin (FCN). Alternative pathway activation depends on complex interactions of complementary regulatory molecules like factor H and properdin (P). The classical pathway, which comprises C1, C4, and C2 components reacting in this order, primarily recognises antibodies in immune complexes. According to[30], this pathway is also in charge of amplifying C3-activation that is brought on by the other two pathways. To activate C5 into C5a and C5b, the target-bound C3b subsequently forms a protease complex. This serves as a scaffold for C6–C9 to create a membrane-attack complex (MAC) and exhibit cytolytic activity. Even though the aforementioned effector actions are essential for effective host defense, excessive complement activation and host cell stimulation may seriously harm host tissue, resulting in anaphylaxis and cell damage[31]. The core sex-shared proteins, Complement C3–C5, Fibulin, Alpha-2-Macroglobulin, Apolipoproteins, and Serum Response Factor—represent proteins essential for cellular defence and transport. Their ubiquitous presence suggests fundamental roles in survival, while frequency variation indicates physiological fine-tuning between sexes.

KEGG and GO analyses revealed predominant enrichment of immune-related pathways, including Complement and Coagulation Cascades, Acute-Phase Response, and Lipid-Transport processes. Smaller male groups (H80, H42) showed fewer enriched pathways, reflecting immature metabolic and immune development, whereas larger groups displayed stronger complement activation and protease-inhibition networks. The absence of common pathways in smaller males corresponds with limited hepatic maturity and reduced serum protein complexity at juvenile stages showed that developmental and physiological changes; complexity increases with maturity and stress response[32], additionally, smaller fish have limited exposure to pathogens and less-developed complement and metabolic systems, resulting in fewer annotated immune pathways. Insights from the advanced

LC-MS/MS technique method can be used to identify proteins that are part of a particular biological pathway or alter their expression in response to disease, toxicants, or other stressors[33]. Thus, there is no current information on this MAD Box protein in fish, but in male Hilsa, it was identified and reported for the first time. The consensus CTA(T/A)4TAG/A sequence, found in the regulatory areas of many muscle-specific and growth-inducible genes, is bound by this protein, which is a family of transcription factors of MEF2 genes and its encoded proteins, in both homo- and heterodimers[34]. There is limited research on the roles of MADS-box genes in stress resistance in animals, and none have been published in fish-related studies. The term "MADS-box" refers to four of the initially recognized members of this conserved motif, which is present in the DNA-binding domains of these proteins: MCM1, AG, DEFA, and SRF[35]. Furthermore, fish development is significantly influenced by MADS-box transcription factors, especially the Myocyte Enhancer Factor 2 (MEF2) family. MEF2 proteins are necessary for the development of skeletal muscles and the preservation of myofibril structure in zebrafish (*Danio rerio*)[36]. According to these results, MADS-box genes are essential for teleost species' growth, reproductive development, and adaption processes. This information is particularly pertinent for comprehending the physiology of migration in species such as Hilsa (*Tenualosa ilisha*). Notably, Alpha macroglobin (A2m) was dominant in Hilsa, which may work in disease resistance. According to[37], a potential marker found in A2M protein could be used as a molecular marker to diagnose fish resistance or susceptibility to EUS disease, and they might even be able to be included in the marker panel for pilot research and this marker also anticipated in *T. Ilisha* also. Fibulin and Alpha Macroglobulin in particular were dominant, that play a part in controlling innate immunity and preserving structural integrity under stressful condition during migration.

Proteomics study at the molecular level can reveal crucial information on the signaling proteins and transcription factors that control fish growth patterns. Nine KEGG pathways were identified via pathway enrichment analysis, with neuroactive ligand-receptor interactions, phagosomes, FoxO signaling, and Herpes simplex virus infection being prevalent in male groups. As observed in proteomic studies of other teleosts[38], the steady enrichment of pathways linked to viral infections (e.g., Herpes simplex virus 1) is intriguing because it was not reflected actual viral exposure but rather emphasize conserved immune and stress-related protein networks in fish. In this study, Apelin signaling pathways found in both adult male and female (stage I and IV) Hilsa, which was also found in Salmon fish kidney cell line[39] were important in fish many biological functions, work as energy balance regulation, intake of food, immunity regulation[40], the FoxO signaling pathway was present and plays significant roles during stress associated with migration in Hilsa[41–43]claimed that although FoxO is necessary for numerous unique cellular processes, such as differentiation, apoptosis, autophagy, metabolism, inflammation, and stress tolerance, it also contributes to heat stress. This process was notably intensified in Hilsa at elevated temperatures, suggesting a potential link between modified energy metabolism and thermal adaptation during sea-to-river migration. The mTOR (mammalian target of rapamycin) signaling pathway was found in a previous study of female adult Hilsa, which helps in immunology, but in males, it showed in each weight group, that in different life stages of Hilsa, it might work in the development and growth[43]. studied n triploid crucian carp, and it showed that the mTOR signaling pathway helps to regulate embryonic development to rapid growth. Additionally, in the realm of cellular metabolism[44,45], regulation of essential biological processes such as cell proliferation, metabolism, differentiation, and survival of fish[46,47] this signaling pathway has received a lot of attention. In KEGG pathways analysis, a few genes were revealed which was associated with specific pathways like metabolism pathways were regulated by *cndp2* genes, *sgk1* and *sgk3* in signaling pathway (mTOR and FOXo), mef2 in Apeline signaling pathway and *c3a.1, c3a.2, c3a.3, c5* was related to immunology like phagosome, Herpes simplex virus 1 infection and Neuroactive ligand-receptor interaction. These cndp2 genes are found in mainly human serum, which is related to carnosinase[48] but also identified for the first time in the tissues of Japanese eel[49]. Therefore, understanding the genes involved in the immune system is crucial for developing effective disease management strategies for wild Hilsa.

This study through PPI networking showed several key proteins and their associate neighbour interactor protein linked to the immune system of males (Fig. 6A) and females (Fig. 6B). Hilsa were identified, offering significant advancements in understanding hilsa's self-defense mechanisms. These proteins are associated with vital immune processes, including pathogen recognition, antigen presentation, and the activation of innate and adaptive immune responses. During the spawning migration phase, Hilsa experiences starvation, which significantly impacts its immune system and overall physiology[50]. Sexual maturity and spawning are reported to influence humoral immunity through interactions between sex hormones and the immune system[51], while starvation suppresses the expression of genes associated with immune proteins[52]. In *T. ilisha*, a decline in immunological indices during migration has been observed, likely due to a combination of environmental and hormonal factors[53]. Understanding fish immune defense mechanisms in response to chemical[54] and physical[55,56], stresses, such as pathogens[57,58], is another significant use of the proteomic technique.

Our integrated proteomic and bioinformatics approach reveals key immune-related proteins in Hilsa, highlighting sex- and weight-dependent differences in response to migration, stress, and infection. These insights provide biomarkers for disease resistance, support selective breeding of robust populations, and strengthen aquaculture and conservation strategies. As seen in salmon and trout[52,59], such molecular-level understanding lays the foundation for Hilsa domestication and sustainable management. The findings were consistent with the biology of serum, which was dominated by highly abundant, conserved immune and carrier proteins that exhibited limited qualitative variability across individuals and sexes. The absence of statistically significant differences after correction indicated that sex and developmental stage did not alter the serum proteome composition at the level of protein identity in our dataset; rather, the quantitative detectability (observation frequency) of shared proteins varied modestly across groups.

Conclusion

This study represents the first comprehensive shotgun proteomic analysis of Hilsa (*Tenualosa ilisha*) serum, providing an in-depth overview of its functional proteome profile with the highest enrichment in immunological properties. Across six weight groups, the maximum number of annotated proteins (817) in 201–300 g individuals, and other groups also comparative study with female reflecting non-linear proteome dynamics during growth and sex. A core set of 142 proteins formed the conserved proteomic backbone across sexes, with only stage-specific shifts in immune-related proteins (MADS-box, Anaphylatoxin, CD109) and no major qualitative sex-specific differences. Functional annotation and network interactions revealed that defence-related genes and immune pathways—including neuroactive ligand–receptor interaction, phagosome, FoxO signaling, and Apelin signaling—varied across developmental stages. Several proteins such as Alpha Macroglobulin, Fibulin, MEF2C, and complement proteins were consistently expressed across groups, whereas MADS-box, anaphylatoxin, and cytosolic non-specific dipeptidase displayed male-biased expression, underscoring sex-dependent proteomic adaptations. Collectively, biological molecular and cellular functions of the hilsa proteomics and KEGG pathways and protein-protein interaction networks helped to understand the physiological plasticity of this ecologically and commercially important species. Considering its high consumer demand and market value, domestication of Hilsa is crucial; proteomic insights into hormonal regulation, ovulation, and migration offer a strong baseline for manipulation, artificial breeding, and sustainable aquaculture of this wild species.

Data availability

All supporting data for this study are included in the manuscript and the supplementary materials. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via PRIDE (Perez et al., 2016; 2022; Deutsch et al., 2023) partner repository of male Hilsa H4, H22, H37 H35, H42, and H80 dataset identifier [PXD063644],(http:/ebi.ac.uk/pride/archive/projects/PXD063644/private), [PXD062032](https:/www.ebi.ac.uk/pride/archive/projects/PXD062032/private), [PXD063281](https:/www.ebi.ac.uk/pride/archive/projects/PXD063281/private), [PXD062121](https:/www.ebi.ac.uk/pride/archive/projects/PXD062121/private), [PXD061789](https:/www.ebi.ac.uk/pride/archive/projects/PXD061789/private) and [PXD063282](https:/www.ebi.ac.uk/pride/archive/projects/PXD063282/private) respectively. For additional data requests, please contact the corresponding author.

Received: 3 May 2025; Accepted: 24 November 2025
Published online: 25 February 2026

References

1. Ueda, H. Physiological mechanisms of imprinting and homing migration in Pacific salmon *Oncorhynchus* spp. *J. Fish Biol.* **81**, 543–558 (2012).
2. Nath, A. K., Banerjee, B. & Ray, S. Socio-economic aspects of the fishers engaged in Hilsa fisheries in Hooghly estuary of West Bengal, India Int. *J. Innov. Stud. Aquat. Biol. Fish* **2**(20), 26 (2016).
3. Karpievitch, Y. V. et al. Liquid chromatography mass spectrometry-based proteomics: biological and technological aspects. *Ann. Appl. Stat.* **4**, 1797–1823 (2010).
4. Carrera, M. et al. Characterization of foodborne strains of *Staphylococcus aureus* by shotgun proteomics: functional networks, virulence factors and species-specific peptide biomarkers. *Front. Microbiol.* **8**, 2458 (2017).
5. Monti, M. et al. Puzzle of protein complexes in vivo: a present and future challenge for functional proteomics. *Expert Rev. proteome.* **6**(2), 159–169. https://doi.org/10.1586/epr.09.7 (2009).
6. Khatri, P., Sirota, M. & Ten Butte, A. J. years of pathway analysis: current approaches and outstanding challenges. *PLoS Comput. Biol.* **8**(2), e1002375 (2012).
7. Li, C. et al. Comprehensive and quantitative proteomic analyses of zebrafish plasma reveals conserved protein profiles between genders and between zebrafish and human. *Sci. Rep.* **6**(1), 24329 (2016).
8. Alderman, S. L. et al. Proteomic analysis of sockeye salmon serum as a tool for biomarker discovery and new insight into the sublethal toxicity of diluted bitumen. *Comp. Biochem. Physiol. Part D: Genom. Proteom.* **22**, 157–166 (2017).
9. Schwanhäusser, B. et al. Global quantification of mammalian gene expression control. *Nature* **473**(7347), 337–342. https://doi.org/10.1038/nature10098 (2011).
10. Sveinsdóttir, H., Martin, S. A., & Vilhelmsson, O. T. Application of proteomics to fish processing and quality. *Food biochemistry and food processing* 2nd (Eds) 406-424 https://doi.org/10.1002/9781118308035.ch22 (2012)
11. Jan, K. et al. LC–MS/MS based characterisation and differential expression of proteins in Himalayan snow trout, *Schizothorax labiatus* using LFQ technique. *Sci. Rep.* **13**(1), 10134 (2023).
12. Keyvanshokooh, S., Kalbassi, M. R., Hosseinkhani, S. & Vaziri, B. Comparative proteomics analysis of male and female Persian sturgeon (*Acipenser persicus*) gonads. *Anim. Reprod. Sci.* **111**(2–4), 361–368 (2009).
13. Hidematsu, H., Kazunori, T. & Akihiko, H. Immunochemical identification of female-specific serum protein, vitellogenin, in the medaka, *Oryzias latipes* (Teleosts). *Comp. Biochem. Physiol. Part A: Physiol.* **76**, 135–141 (1983).
14. Nissa, M. U. et al. Organ-based proteome and post-translational modification profiling of a widely cultivated tropical water fish *Labeo rohita*. *J. Proteom. Res.* **21**, 420–437 (2021).
15. Xiang, H. et al. Proteomics study of mitochondrial proteins in tilapia red meat and their effect on colour change during storage. *Food Chem.* **400**, 134061 (2023).
16. Bakke, F. K. et al. Plasma proteome responses in salmonid fish following immunization. *Front. Immunol.* **11**, 581070 (2020).
17. Wiens, G. D. et al. Plasma proteomic profiling of bacterial cold-water disease-resistant and-susceptible rainbow trout lines and biomarker discovery. *Front. Immunol.* **14**, 1265386 (2023).
18. He, Z. et al. Protein expression patterns and metal metabolites in a protogynous hermaphrodite fish, the rice field eel (*Monopterus albus*). *BMC Genom.* **25**(1), 500 (2024).
19. Chakraborty, H., Chakraborty, H. J., Das, B. K. & Maity, J. Age-specific changes in the serum proteome of female anadromous, hilsa *Tenualosa ilisha*: a comparative analysis across developmental stages. *Front. Immunol.* **15**, 1448627 (2024).
20. Diz, A. P., Truebano, M. & Skibinski, D. O. The consequences of sample pooling in proteomics: an empirical study. *Electrophoresis* **30**, 2967–2975 (2009).
21. Molinari, N. et al. Sample pooling and inflammation linked to the false selection of biomarkers for neurodegenerative diseases in top–down proteomics: a pilot study. *Front. Mol. Neurosci.* **11**, 477 (2018).

22. Meneja, R. H. et al. The proteome of atlantic herring (*Clupea harengus* L.) larvae is resistant to elevated pCO2. *Mar. Pollut. Bull.* **86**, 154–160 (2014).
23. Neiffer, D. L. & Stamper, M. A. Fish sedation, anesthesia, analgesia, and euthanasia: considerations, methods, and types of drugs. *ILAR J.* **50**(4), 343–360 (2009).
24. Lawrence, M. J. et al. Best practices for non-lethal blood sampling of fish via the caudal vasculature. *J. Fish Biol.* **97**(1), 4–15 (2020).
25. Rahman, M. J. Recent advances in the biology and management of Indian shad (*Tenualosa ilisha* Ham.). *SAARC J. Agric.* **4**, 67–90 (2006).
26. Perez-Riverol, Y. et al. PRIDE inspector toolsuite: moving toward a universal visualization tool for proteomics data standard formats and quality assessment of Proteomexchange datasets. *Mol. Cell. Proteom.* **15**(1), 305–317 (2016).
27. Perez-Riverol, Y. et al. The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* **50**(D1), D543–D552 (2022).
28. Deutsch, E. W. et al. The ProteomeXchange consortium at 10 years: 2023 update. *Nucleic Acids Res.* **51**(D1), D1539–D1548 (2023).
29. Ribas, L. & Piferrer, F. The zebrafish (*Danio rerio*) as a model organism, with emphasis on applications for finfish aquaculture research. *Rev. Aquacult.* **6**, 209–240 (2014).
30. Walport, M. J. Complement. First of two parts. *N. Engl. J. Med.* **344**, 1058–1066 (2001).
31. Nakao, M. et al. The complement system in teleost fish: progress of post-homolog-hunting researches. *Dev. Comp. Immunol.* **35**(12), 1296–1308 (2011).
32. Braceland, M. et al. The serum proteome of Atlantic salmon, *Salmo salar*, during pancreas disease (PD) following infection with salmonid alphavirus subtype 3 (SAV3). *J. Proteome.* **94**, 423–436 (2013).
33. Bradley, B. P., Kalampanayil, B., & O'Neill, M. C. Protein expression profiling. In *Two-dimensional electrophoresis protocols* (455-468). Totowa NJ: Humana Press. (2009). https://doi.org/10.1007/978-1-59745-281-6_30
34. Hobson, G. M. et al. Regional chromosomal assignments for four members of the MADS domain transcription enhancer factor 2 (MEF2) gene family to human chromosomes 15q26, 19p12, 5q14, and 1q12–q23. *Genomics* **29**(3), 704–711 (1995).
35. Shore, P. & Sharrocks, A. D. The MADS-box family of transcription factors. *Eur. J. Biochem.* **229**, 1–13 (1995).
36. Hinits, Y. & Hughes, S. M. Mef2s are required for thick filament formation in nascent muscle fibres. *Development* **134**(13), 2511–2519 (2007).
37. Mohindra, V., Dangi, T., Chowdhury, L. M. & Jena, J. K. Tissue specific alpha-2-Macroglobulin (A2M) splice isoform diversity in Hilsa shad, *Tenualosa ilisha* (Hamilton, 1822). *Plos one* **14**(7), e0216144 (2019).
38. Wolf, K. & Smith, C. *Herpesvirus salmonis*: pathological changes in parenterally-infected rainbow trout, *Salmo gairdneri* Richardson, fry. *J. Fish Dis.* **4**, 445–457 (1981).
39. Gjoen, T., Ruyter, B. & Østbye, T. K. Effects of eicosapentaneoic acid on innate immune responses in an Atlantic salmon kidney cell line in vitro. *PLoS One* **19**, e0302286 (2024).
40. Yan, X. et al. The regulatory role of apelin on the appetite and growth of common carp (Cyprinus carpio L.). *Animals* **10**, 2163 (2020).
41. Gross, D. N., Van Den Heuvel, A. P. J. & Birnbaum, M. J. The role of FoxO in the regulation of metabolism. *Oncogene* **27**, 2320–2336 (2008).
42. Wang, Y., Zhou, Y. & Graves, D. T. FOXO transcription factors: their clinical significance and regulation. *Biomed. Res. Int.* **2014**, 1–13 (2014).
43. Huang, Z. et al. Transcriptome analysis and weighted gene co-expression network reveals potential genes responses to heat stress in turbot *Scophthalmus maximus*. *Comp. Biochem. Physiol. D Genom. Proteom.* **33**, 100632 (2020).
44. Saxton, R. A. & Sabatini, D. M. mTOR signaling in growth, metabolism, and disease. *Cell* **168**, 960–976 (2017).
45. Schmelzle, T. & Hall, M. N. TOR, a central controller of cell growth. *Cell* **103**, 253–262 (2000).
46. Kapahi, P. et al. With TOR, less is more: a key role for the conserved nutrient-sensing TOR pathway in aging. *Cell Metab.* **11**, 453–465 (2010).
47. Lloyd, B. A. et al. Exercise increases mTOR signaling in brain regions involved in cognition and emotional behavior. *Behav. Brain Res.* **323**, 56–67 (2017).
48. Teufel, M. et al. Sequence identification and characterization of human carnosinase and a closely related non-specific dipeptidase. *J. Biol. Chem.* **278**, 6521–6531 (2003).
49. Oku, T. et al. Purification and identification of two carnosine-cleaving enzymes, carnosine dipeptidase I and Xaa-methyl-His dipeptidase, from Japanese eel (*Anguilla japonica*). *Biochimie* **94**, 1281–1290 (2012).
50. Cuesta, A. et al. Effect of sex-steroid hormones, testosterone and estradiol, on humoral immune parameters of gilthead seabream. *Fish Shellfish Immunol.* **23**, 693–700 (2007).
51. Harris, J. & Bird, D. J. Modulation of the fish immune system by hormones. *Vet. Immunol. Immunopathol.* **77**, 163–176 (2000).
52. Martin, S. A., Douglas, A., Houlihan, D. F. & Secombes, C. J. Starvation alters the liver transcriptome of the innate immune response in Atlantic salmon (*Salmo salar*). *BMC Genom.* **11**, 1–20 (2010).
53. Ghobeishavi, A. et al. The innate immunity changes of the female anadromous Hilsa shad, *Tenualosa ilisha*, during spawning and post-spawning season. *Iran. J. Fish. Sci.* **15**, 1526–1539 (2016).
54. Lin, H. D., Hsu, L. S., Chien, C. C. & Chen, S. C. Proteomic analysis of ametryn toxicity in zebrafish embryos. *Environ. Toxicol.* **33**(5), 579–586 (2018).
55. Smith, R. W., Moccia, R. D., Mothersill, C. E. & Seymour, C. B. Irradiation of rainbow trout at early life stages results in a proteomic legacy in adult gills. Part B; the effect of a second radiation dose, after one year, on the proteomic responses in the irradiated and non-irradiated bystander fish. *Environ. Res.* **163**, 307–313. https://doi.org/10.1016/j.envres.2018.01.020 (2018).
56. Smith, R. W., Moccia, R. D., Seymour, C. B. & Mothersill, C. E. Irradiation of rainbow trout at early life stages results in a proteomic legacy in adult gills. Part A; proteomic responses in the irradiated fish and in non-irradiated bystander fish. *Environ. Res.* **163**, 297–306. https://doi.org/10.1016/j.envres.2017.12.023 (2018).
57. Kumar, G. et al. Proteome analysis reveals a role of rainbow trout lymphoid organs during *Yersinia ruckeri* infection process. *Sci. Rep.* **8**(1), 13998 (2018).
58. Lü, A. et al. iTRAQ analysis of gill proteins from the zebrafish (*Danio rerio*) infected with Aeromonas hydrophila. *Fish Shellfish Immunol.* **36**(1), 229–239 (2014).
59. Long, M. et al. Transcriptomic and proteomic analyses of splenic immune mechanisms of rainbow trout (*Oncorhynchus mykiss*) infected by *Aeromonas salmonicida* subsp. *salmonicida*. *J. Proteom.* **122**, 41–54 (2015).

Acknowledgements

The authors gratefully acknowledge the Director of ICAR-CIFRI for providing the necessary facilities to conduct this research and thanks to the National Mission for Clean Ganga for funding support. Special appreciation is given to Mr. Anjon Talukdar, Technical Assistant, for his valuable assistance with sample preparation.

Author contributions

The experiment was designed by BKD. Samples were collected, prepared and manuscript writing by HC. Proteomics analysis was done by HJC. Data interpretation by BKD, HC, HJC and JM. The manuscript is prepared

and edited by all the authors.

Declarations

Competing interest
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-30326-5.

Correspondence and requests for materials should be addressed to B.K.D.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.