ISSN: 2347-5129
IJFAS 2015; 2(5): 191-197

www.fisheriesjournal.com
Received: 12-02-2015
Accepted: 15-03-2015

Dibyashree Swain
Fish Genetics and Biotechnology Division, ICAR - Central Institute of Freshwater Aquaculture, Bhubaneswar 751 002, Odisha, India.

Kiran Dashrath Rasal
Fish Genetics and Biotechnology Division, ICAR - Central Institute of Freshwater Aquaculture, Bhubaneswar 751 002, Odisha, India.

Pranati Swain
Fish Genetics and Biotechnology Division, ICAR - Central Institute of Freshwater Aquaculture, Bhubaneswar 751 002, Odisha, India.

Prajna Pradhan
Fish Genetics and Biotechnology Division, ICAR - Central Institute of Freshwater Aquaculture, Bhubaneswar 751 002, Odisha, India.

Jitendra Kumar Sundaray
Fish Genetics and Biotechnology Division, ICAR - Central Institute of Freshwater Aquaculture, Bhubaneswar 751 002, Odisha, India.

Hirak Kumar Barman
Fish Genetics and Biotechnology Division, ICAR - Central Institute of Freshwater Aquaculture, Bhubaneswar 751 002, Odisha, India.

Pallipuram Jayasankar
Fish Genetics and Biotechnology Division, ICAR - Central Institute of Freshwater Aquaculture, Bhubaneswar 751 002, Odisha, India.

Correspondence
Kiran Dashrath Rasal
Scientist Fish Genetics and Biotechnology Division, ICAR - Central Institute of Freshwater Aquaculture, Bhubaneswar 751 002, Odisha, India.

Comparative and evolutionary analysis of Growth Hormone (GH) protein of Cyprinids using computational approach

Dibyashree Swain, Kiran Dashrath Rasal, Pranati Swain, Prajna Pradhan, Jitendra Kumar Sundaray, Hirak Kumar Barman and Pallipuram Jayasankar

Abstract
Growth hormone (GH) is an important polypeptide necessary for normal growth as well as development in fishes. Even though, GH has been well characterized in several fishes, lack of information on the structure GH protein. Recent advances of computational tools are helpful for developing structure using protein sequences using either homology modelling or *de-novo* modelling. In this study, we have used computational tools for three-dimensional structure prediction and analysis evolutionary profile of GH from 12 fishes belonging to cyprinids. The sequence analysis predicted GH polypeptide of 204-210 amino-acids, including a putative signal peptide of 1-22 hydrophobic residues and mature protein containing 183 amino-acids among cyprinid fishes. We could identify well conserved region among all the species, which states, that region maintained by evolution despite speciation. The evolutionary analysis was performed using CLC genomic workbench followed by a bootstrap test (100 replicates). A phylogenetic tree was depicting the relationship of various cyprinids GH protein sequences. We have generated 3D structure of GH protein of rohu using I-TASSER, which was further validated using SAVES server. The Procheck analysis depicted the Ramachandran plot consisting of 76.8% residues lies in the most favored regions. For understanding global network of GH, STRING9.1 speculated that, this protein interacting with several proteins, strong association with two proteins such as GHRa and GHRb. The present study revealed that GH protein level tends to be highly conserved across cyprinids. This work is also useful for *in vivo* study of structural and functional analysis of GH protein.

Keywords: Growth hormone, Modelling, I-TASSER, Phylogeny, Three dimensional structure.

1. Introduction
Over the period of decades, various features of growth hormone (GH) gene physiology along with molecular aspects have been studied in several fish species. Preferably, the growth rate of fish is being an important economic trait which needs to be enhanced with help of traditional selective breeding practices as well as gene manipulation techniques for mitigating growing fish food demand. The growth rate of fish has been controlled by various genes such as growth hormone (GH), Insulin-like growth factor-I (IGF-I) and Growth hormone receptors (GHRa/b), which had explored with help of molecular techniques [1, 2, 3]. The growth hormone, which mainly regulates the growth and development as well as significant role in osmoregulation, steroidogenesis, anti-freeze protein (AFP) production, reproduction in fishes [2, 4, 5, 6, 7]. In fishes, increased attention pertains to growth has led to investigation of gene transcripts encoding GH, IGF-I, IGF-II, and their receptors along with IGF-binding proteins [3, 8]. It has been demonstrated that two sub-types of the GH receptor (GHR1 and GHR2) are present in most of the teleost [2, 9]. In addition, transgenic approach has been utilized in several fish species such as Atlantic salmon [5], catfish [10], carp [11], tilapia [12], loach [13], Arctic char [14] for improvement production and performance traits such as growth, disease resistance.
Recent knowledge of the complete sequences of some farmed carp species genes has led us to a comparative study of different sequence features among them. Currently with the wide availability of sequence data (Nucleotide/Protein), we can gain previously impossible insights into evolution of protein universe, on the domain, network, and genome levels. Network level studies allow us to understand the forces of evolution that affect changes and conservation of protein networks.

Understanding the evolutionary pressures that affect protein networks provides insight into their functional importance for various organisms, including how they affect the organism's competitiveness. The study of protein evolution had also become a key area of scientific research with discoveries such as the coordinated changes of key residues and substitution led to a phenotype change [15]. The theory and practice of phylogenetic tree construction matured into the PHYLIP program, and phylogenetic analysis yielded significant discoveries in genome evolution, such as the relationships between life forms and the dynamics of genome structure [16].
Therefore, this wealth of DNA/protein sequence data in the database provided the perfect resources to gain further information into protein composition and their function, protein-protein interaction network and evolution. The improvement in the computational algorithms/tools made new avenue for studying the genome/proteome at wide scale.
In this study, although, there is availability of GH sequences and its encoded protein sequences in the database for cyprinids, but lack of species-specific structural information. This led to absence of knowledge of protein network and evolution in fishes. In this study, we have retrieved protein sequences of cyprinids, used for comparative as well as evolutionary analysis. We have predicted the three dimensional structure and its protein-interaction network. Also, we have used molecular docking approach to investigate receptor binding activities. Our results provide meaningful information for further studies on the cyprinids and their phylogeny.

2.0 Material and Methods

We have used different bioinformatics tools for studying the growth hormone protein, which listed along with specified purpose**.**

2.1 Retrival of GH protein sequences

The UniProt is easily accessible database of protein sequence (http://www.uniprot.org/). We have retrieved total 12 protein sequences from different fishes belongs to cyprinids for our study (Table1). We have retrieved all protein sequences in a FASTA format with accession numbers.

Table 1: Protein statistics of GH of cyprinids

Taxon	UniProtKB accession number	Number of amino acid
Danio rerio	Q571R1	210
Labeo rohita	Q9W6J7	207
Labeo calbasu	Q5Y4C2	210
Labeo bata	Q5Y4C1	210
Labeo fimbriatus	Q5Y4C8	210
Labeo kontius	Q5Y4C6	210
Cyprinus carpio	P10298	210
Carassius auratus	O93359	210
Hypophthalmichthys molitrix	P69159	210
Ctenopharyngodon idella	P69158	210
Catla catla	Q90WV7	210
Cirrhinus mrigala	Q90W30	210
Cyprinus tinca	D7RPP4	210

2.2 Physico-chemical Characterization and domain, motif analysis of GH proteins

The properties of GH proteins was determined and characterized from the protein sequences. Thus, ProtParam (http://web.expasy.org/protoparam/) expasy tool, which is useful for computation of physical and chemical parameters of given protein based on sequence information. We have calculated several physic-chemical properties such as theoretical isoelectric point (pI), molecular weight, and total number of positive and negative residues, extinction coefficient, half-life, instability index, aliphatic index and grand average hydropathy (GRAVY) of all 12 retrieved protein sequences. The PROSITE server was used to determine specific profiles as well as patterns related to the GH proteins. The signal peptide site were predicted using SignalP server (version 4.0), which is based on neural network method (http://www.cbs.dtu.dk/services/SignalP/).

2.3 Multiple Sequence Alignment and Phylogenetic study of GH

In order to determine comparison among different protein sequences, we have used global multiple sequence alignment (MSA) program for analysis of GH protein sequences from different fishes. Now a day, multiple sequence alignment (MSA) method is widely used for assessing sequence conservation and conservation of protein domains in protein study. In this step, CLC genomic workbench 7.5.1 tool was used for MSA analysis.
Understanding phylogenetic relationship among different protein sequences, we have observed the evolutionary relationship of these sequences by cladogram. The Evolutionary analysis of GH protein among 12 species was conducted CLC genomic workbench 7.5.1. The NJplot software was utilized for obtaining a graphical view of the phylogenetic tree of GH sequences.

2.4 Analysis of Gene ontology and protein-protein interaction network of GH

We further studied the gene ontology of GH protein obtaining biological, molecular and cellular functions, which recognized using Uniprot server (http://www.uniprot.org/). The STRING algorithm (Search Tool for the Retrieval of Interacting Proteins) was used for studying the protein-protein interaction network of the GH (http://string-db.org/).

2.5 Three dimensional structure analysis

For getting homologous 3D structure for GH protein, BLASTp (Basic local alignment) searches were performed against protein databank (PDB). But, due to absence of homologous 3D structure/template, we have used I-TASSER (*I*terative *T*hreading *ASSE*mbly *R*efinement)) server, which is an on-line platform for protein structure -function prediction (http://zhanglab.ccmb.med.umich.edu/I-TASSER/). It is based on multiple threading alignments and iterative structural assembly simulations. The prediction of accuracy of the model depend upon a confidence score (C-score) based on the quality of the threading alignments and structural assembly refinement simulations [17]. In this server, we have submitted as a query sequence (FASTA) of GH protein for obtaining 3D model with academic id.
The best/top 3D model was selected according rank based on C-score, TM score and RMSD value. The 3D structure of GH was visualized by PyMOL (http://www.pymol.org/) and UCSF CHIMERA (https://www.cgl.ucsf.edu/chimera/). The PyMoL is open source molecular visualization tool for interactive visualization and analysis of molecular structures.
The structural evaluation along with stereo-chemical quality assessment of predicted model were carried out by using the SAVES (Structural Analysis and Verification Server), which is integrated server (http://nihserver.mbi.ucla.edu/SAVES/). The

ProSA (Protein Structure Analysis) web server (https://prosa.services.came.sbg.ac.at/prosa) was used for refinement and validation of protein structure [18]. The ProSA was used for checking model structural quality with potential errors and program shows a plot of its residue energies and Z-scores which determine overall quality of the model.

3. Results and Discussion

Recently, several bioinformatics tools/algorithms are being used in diverse biological fields, alternatively, improvement in the precision of experiment [19, 20, 21, 22]. In this work, in order to understand evolutionary as well as comparative analysis of the GH protein of fishes belonging to cyprinids, we have used several bioinformatics algorithms. The precise annotations as well as classification of proteins, which deduced from nucleotide sequences, are important for the perfect understanding and study of protein sequence.

Thus, information available via the protein- database can be utilized for evaluation and/or prediction of specific characters about a protein of interest. Although, the number of aa residues showed 171 to 217 in human, bovine, and in some fishes like ornamental fish group having in the range of 204 to 208, the present study also shown aa residues in the range of 207-210. The most of the proteins possessed putative signal peptide length of 1-22 hydrophobic residues.

The various physico-chemical characteristics were determined using ProtParam tool on the ExPASY proteomic server [23]. We have depicted that, the physico-chemical properties were almost same in all the species of cyprinids such as length, theoretical pI, positive residues, negative residues and aliphatic index (Table 2). The primary structures of GH protein are shown to be comprised of 207 to 210 amino acid residues. We inferred that, negatively charged and positively charged aa residues were 25-26, 23-25 respectively, in the GH sequence. The amino acid leucine (L=16%) and serine (S=12.3%) has been found mainly abundant in the GH of these 12 fish species. The aliphatic index (AI) for GH was calculated in the range of 97 to 105 and predicted as stable. Because, if the instability index of a given protein is < 40, it is predicted as a stable, while > 40 indicated as unstable protein [24]. In our case, the value of an instability index of all protein was above 40 which indicate that all 12 proteins were unstable. The Extinction coefficients (EC) value of GH was calculated, which helpful in the protein- ligand and protein-protein interaction study**.** The pI values below 7 of all proteins were acidic in character. The GRAVY value or index for given proteins is calculated as the sum of the hydropathy values of all the aa divided by the number of residues in the protein sequence [25]. GRAVY index > 0 indicates a protein is hydrophobic and reveals hydrophobicity of the protein. All the GH protein sequences of 12 different fishes having hydrophilic in nature. We also analyzed the presence of cysteine residues; which form disulfide bonds, significant for protein folding and stability [26]. The earlier report predicted that, the GH amino acid sequence has around four cysteine residues with highly conserved positions in the all vertebrates [27]. Interestingly, we have identified the potentially evolutionarily significant difference in the number of cysteine residues in the mature peptide of GH. It was found that, the large disulfide bonding in the loop is crucial for growth stimulating/promoting activity [11]. The two disulfide bonds, supposed to contribute for forming 3D structure of growth hormone molecule and a site for N-linked glycosylation [2].

Table 2: Physico-chemical characteristics of growth hormone protein of 13 fishes of cyprinids

Taxon	MW (Da)	pI	Negative charged residues	Positive charged residues	Formula	AI	GRAVY
Danio rerio	23768.4	6.32	25	24	$C_{1047}H_{1703}N_{287}O_{317}S_{12}$	100.24	-0.139
Labeo rohita	23521.1	6.31	26	25	$C_{1030}H_{1682}N_{290}O_{313}S_{12}$	105.02	-0.142
Labeo calbasu	23804.4	6.31	26	25	$C_{1043}H_{1703}N_{293}O_{317}S_{12}$	104.90	-0.140
Labeo bata	23805.5	6.31	26	25	$C_{1044}H_{1706}N_{292}O_{317}S_{12}$	105.38	-0.125
Labeo fimbriatus	23790.4	6.31	26	25	$C_{1042}H_{1701}N_{293}O_{317}S_{12}$	104.43	-0.141
Labeo kontius	23790.4	6.31	26	25	$C_{1042}H_{1701}N_{293}O_{317}S_{12}$	104.43	-0.141
Cyprinus carpio	23765.3	5.82	25	22	$C_{1045}H_{1682}N_{290}O_{317}S_{12}$	99.33	-0.159
Carassius auratus	23759.3	5.95	25	23	$C_{1043}H_{1684}N_{290}O_{318}S_{12}$	97.48	-0.173
Hypophthalmichthys molitrix	23580.1	5.96	25	23	$C_{1036}H_{1679}N_{287}O_{317}S_{11}$	99.33	-0.140
Ctenopharyngodon idella	23580.1	5.96	25	23	$C_{1036}H_{1679}N_{287}O_{317}S_{11}$	99.33	-0.140
Catla catla	23790.4	6.31	26	25	$C_{1042}H_{1701}N_{293}O_{317}S_{12}$	104.43	-0.141
Cirrhinus mrigala	23791.4	6.31	26	25	$C_{1043}H_{1704}N_{292}O_{317}S_{12}$	104.90	-0.127
Cyprinus tinca	23699.2	6.32	25	24	$C_{1042}H_{1684}N_{290}O_{317}S_{11}$	97.00	-0.175

The PROSITE server is one of the most widely utilized for abundant pattern and profile analysis of a given protein [28]. Thus, functions of the GH proteins were analyzed by submitting the aa sequences to the PROSITE server.

In subsequent analysis, Multiple Sequence Alignment (MSA) can give insight into sequence conservation across several species and thus allow identification of those sections of the sequence most critical to protein function. For this, several advanced and accurate algorithms/tools have been developed such as CLUSTAL-W, T-Coffee, MAFFT, based on the fast Fourier transformation, and MEGA [29, 30, 31, 32]. After performing MSA, we have seen that "FRLLACFKKDMHKVETYLRVANCRRSLDSNCTL" is the conserved region (identical region) in all 12 different protein sequences of GH which specifies that, this peptide sequence may have been maintained by evolution despite speciation (Fig 1). Again, we have made BLASTp of this conserved sequence in NCBI and revealed that, sequence is conserved among all the fishes with more than 92% identity.

In our study, phylogenetic tree was constructed CLC genomic workbench UPMGA and NJ methods (Fig 2). The earlier study demonstrated that, bootstrapping method generates reliable trees using re-sampling the given dataset several times [33, 34]. We observed that, *Labeo* species fishes are more related to catla than *cyprinus* species at protein level. The previous studies depicted that, phylogenetic analysis and comparative modelling for depiction of growth hormone gene of ornamental fishes [35]. The earlier study revealed, GH sequences cDNAs, and protein sequences have been used as a phylogenetic marker for studies of evolutionary genetics of various fishes [36, 37].

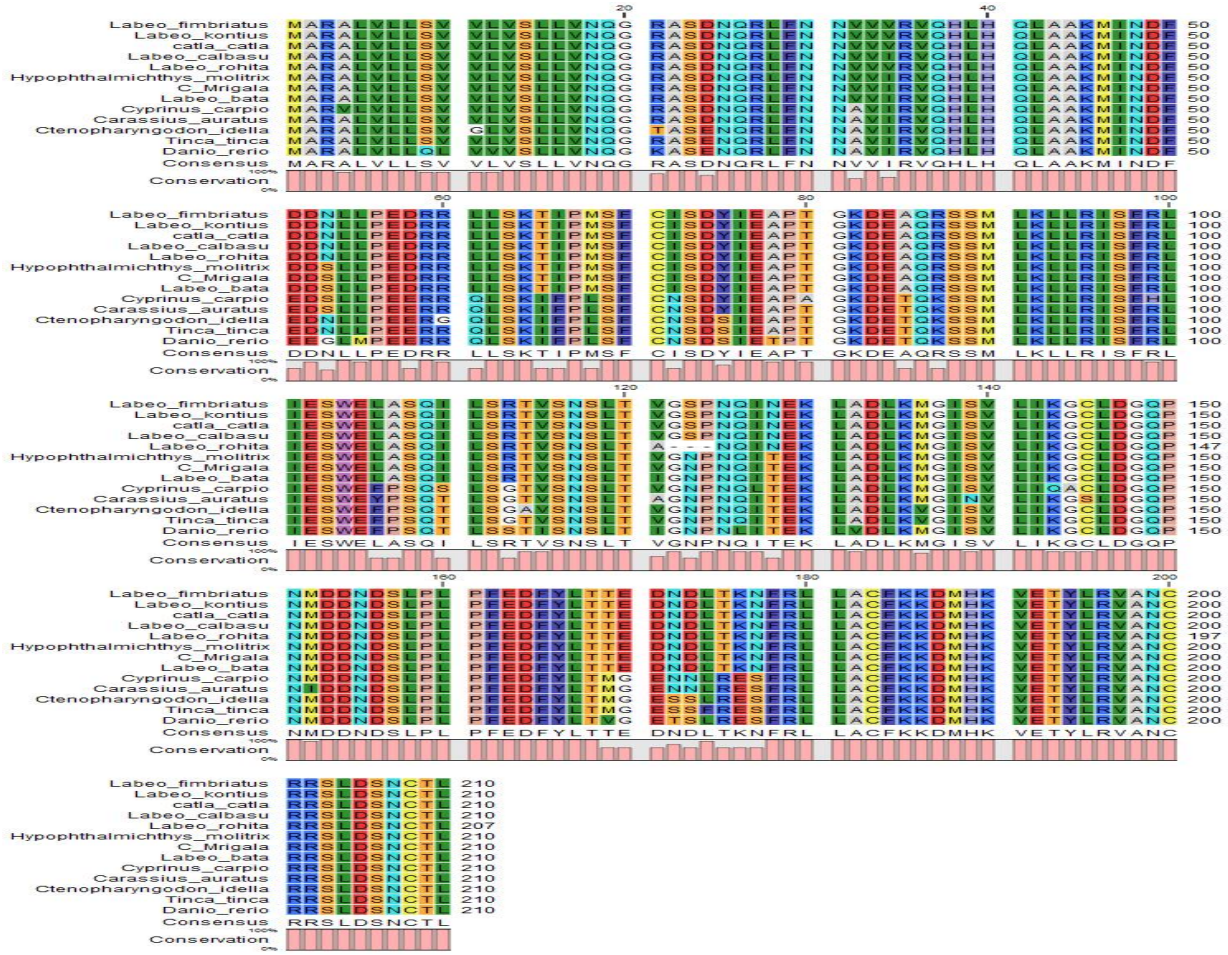

Fig 1: Comparison of GH peptide sequence of cyprinids and showing conserved residues among sequences found.

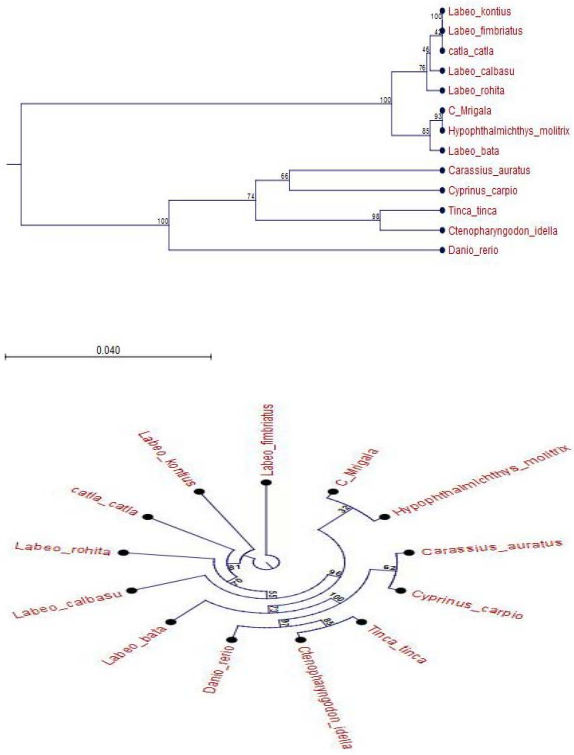

Fig 2: Phylogenetic tree and Cladogram for growth hormone (GH) constructed by the UPMGA and neighbor-joining method with 100 replicates of bootstrap.

Further, in order to understand the structure level of GH protein, the modeling of protein of rohu was done by I-TASSER an automated protein structure prediction algorithm. The best 3D structure with high confidence score (C-score) was selected and used for further investigations. The top model of C-score of 0.46 was selected with estimated TM-score was 0.65±0.13 and estimated RMSD value 7.7±4.3Å (Fig 3). The selected model validated and Ramachandran plot showed good quality of model with most favored region 76.8% (Fig 4). The ProSA analysis showed that, Z score value, which indicates its quality assessment.

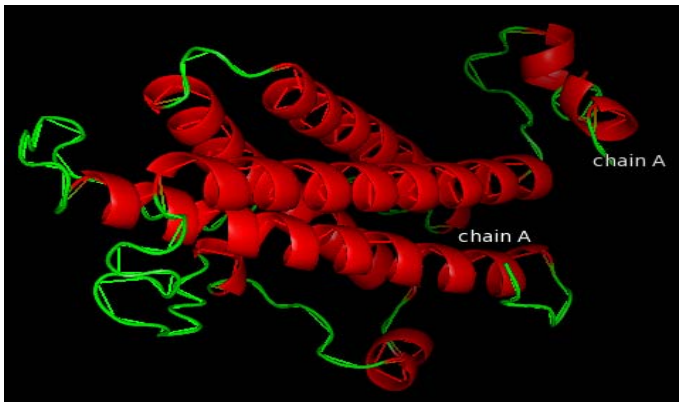

Fig 3: Growth hormones of rohu consensus three dimensional structure generated using I-TASSER

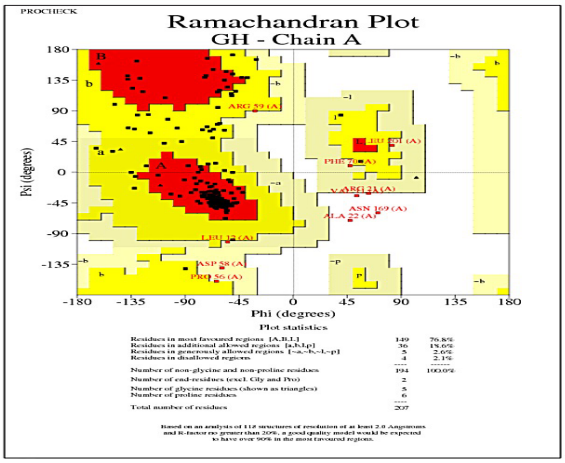

Fig 4: Ramachandran Plot of GH protein showing residues in most favored (red), additionally allowed (yellow), generously allowed (pale yellow) and disallowed regions(white) generated by PROCHECK (SAVES, server).

The protein-protein interaction exploration is a comprehensive approach to identify the organization of desire proteome. The functional network of protein study will be helpful for drug discovery, to understand metabolic pathways and to predict or develop genotype-phenotype associations [38, 39]. In order to understand network of GH protein, we performed analysis with submitting FASTA sequences to STRING 9.1. In STRING, the functional interaction was analyzed by using confidence score. Interactions with score < 0.3 are considered as low confidence, scores ranging from 0.3 to 0.7 are classified as medium confidence and scores > 0.7 yield high confidence [40, 41]. The GH protein network showing functional association with 10 proteins such as GHRa, GHRb, prolactin receptor a, prolactin receptor b, Janus kinase 2, Janus kinase 2a, prolactin, signal transducer and activator of transcription 5.1, suppressor of cytokine signaling 2, and signal transduction and activation of transcription 3 with confidence score 0.975, 0.975, 0.953, 0.952, 0.943, 0.943, 0.943, 0.932, 0.919, 0.913, and 0.911, respectively (Fig 5). We have depicted that, the receptor proteins are showing interacting more with the high confidence score as compare to other protein. The earlier study depicted, growth hormone gene structure is not conserved among teleosts, presenting a mammalian like organization (5 exons and 4 introns) in carps [42]. We could identify that, GH is more conserved at the protein level in fish species belonging to cyprinids. But, still questions remains, why their growth is differ from each other.

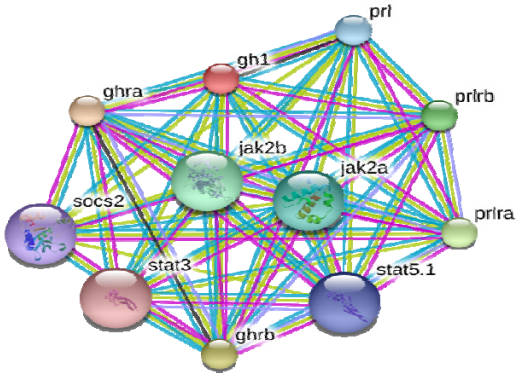

Fig 5: Protein interaction network of GH protein of rohu depicted using STRING

4. Conclusion

This is first comprehensive study on comparative as well as evolutionary study of growth hormone protein sequences among 12 different important fish species of cyprinids. In this study, we have investigated that most of physico-chemical properties were almost same in all the protein sequences. After performing multiple sequence alignment, we have depicted that "FRLLACFKKDMHKVETYLRVANCRRSLDSNCTL" is the conserved (identical) motif which present in all 12 protein sequences of GH. We also analyzed overall 100 hit from the NCBI protein database, which indicated the importance of this region among a diverse group of fish species with more than 91-100% identity. Moreover, protein-protein interaction pathway of this GH protein has helped us to understand the roles and associated proteins in various cellular pathways. We also generated 3D structure of GH protein using *de novo* modelling and subsequently validated using Ramachandran plot. The GH at protein level tends to be highly conserved across cyprinids. Thus finally, present work will support to understand more about GH proteins of fish species among cyprinids and their network.

5. Acknowledgment

We are thankful to Indian Council of Agriculture Research, New Delhi.

6. References

1. Zhu T, Goh EL, Graichen R, Ling L, Lobie PE. Signal transduction via the growth hormone receptor. Cellular signalling 2001; 13(9):599-616.
2. Sciara AA, Rubiolo JA, Somoza GM, Arranz SE. Molecular cloning, expression and immunological characterization of pejerrey (Odontesthes bonariensis) growth hormone. Comparative biochemistry and physiology Toxicology & pharmacology: CBP 2006; 142(3-4):284-92.
3. Filby AL, Tyler CR. Cloning and characterization of cDNAs for hormones and/or receptors of growth hormone, insulin-like growth factor-I, thyroid hormone, and corticosteroid and the gender-, tissue-, and developmental-specific expression of their mRNA transcripts in fathead minnow (Pimephales promelas). General and comparative endocrinology 2007; 150(1):151-63.
4. Degani G, Tzchori I, Yom-Din S, Goldberg D, Jackson K. Growth differences and growth hormone expression in male and female European eels [Anguilla anguilla (L.)]. General and comparative endocrinology 2003; 134(1):88-93.
5. Duan DS, Werner S, Williams LT. A naturally occurring secreted form of fibroblast growth factor (FGF) receptor 1 binds basic FGF in preference over acidic FGF. The Journal of biological chemistry 1992; 267(23):16076-80.
6. Gomez JM, Boujard T, Fostier A, Le Bail PY. Characterization of growth hormone nycthemeral plasma profiles in catheterized rainbow trout (Oncorhynchus mykiss). The Journal of experimental zoology 1996; 274(3):171-80.
7. Sakamoto T, Shepherd BS, Madsen SS, Nishioka RS, Siharath K, Richman NH 3rd *et al.* Osmoregulatory actions of growth hormone and prolactin in an advanced teleost. General and comparative endocrinology 1997; 106(1):95-101.

8. Gabillard JC, Kamangar BB, Montserrat N. Coordinated regulation of the GH/IGF system genes during refeeding in rainbow trout (Oncorhynchus mykiss). The Journal of endocrinology 2006; 191(1):15-24.
9. Saera-Vila A, Calduch-Giner JA, Perez-Sanchez J. Duplication of growth hormone receptor (GHR) in fish genome: gene organization and transcriptional regulation of GHR type I and II in gilthead sea bream (Sparus aurata). General and comparative endocrinology 2005; 142(1-2):193-203.
10. Dunham RA, Ramboux AC, Duncan PL, Hayat M, Chen TT, Lin CM *et al.* Transfer, expression, and inheritance of salmonid growth hormone genes in channel catfish, Ictalurus punctatus, and effects on performance traits. Molecular marine biology and biotechnology 1992; 1(4-5):380-9.
11. Chen TT, Kight K, Lin CM, Powers DA, Hayat M, Chatakondi N *et al.* Expression and inheritance of RSVLTR-rtGH1 complementary DNA in the transgenic common carp, Cyprinus carpio. Molecular marine biology and biotechnology 1993; 2(2):88-95.
12. Martinez R, Estrada MP, Berlanga J, Guillen I, Hernandez O, Cabrera E *et al.* Growth enhancement in transgenic tilapia by ectopic expression of tilapia growth hormone. Molecular marine biology and biotechnology 1996; 5(1):62-70.
13. Nam YK, Noh JK, Cho YS, Cho HJ, Cho KN, Kim CG, *et al.* Dramatically accelerated growth and extraordinary gigantism of transgenic mud loach Misgurnus mizolepis. Transgenic research 2001; 10(4):353-62.
14. Pitkanen TI, Xie SQ, Krasnov A, Mason PS, Molsa H, Stickland NC. Changes in tissue cellularity are associated with growth enhancement in genetically modified arctic char (Salvelinus alpinus L.) carrying recombinant growth hormone gene. Marine biotechnology 2001; 3(2):188-97.
15. Reeck GR, de Haen C, Teller DC, Doolittle RF, Fitch WM, Dickerson RE, *et al.* "Homology" in proteins and nucleic acids: a terminology muddle and a way out of it. Cell 1987; 50(5):667.
16. Iwabe N, Kuma K, Hasegawa M, Osawa S, Miyata T. Evolutionary relationship of archaebacteria, eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes. Proceedings of the National Academy of Sciences of the United States of America 1989; 86(23):9355-9.
17. Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc 2010; 5(4):725-38.
18. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic acids research (Web Server issue) 2007; 35:W407-10.
19. Zamani A, Benjakul S. Trypsin from unicorn leatherjacket (Aluterus monoceros) pyloric caeca: Purification and its use for preparation of fish protein hydrolysate with antioxidative activity. Journal of the science of food and agriculture, 2015.
20. Rahbar MR, Rasooli I, Mousavi Gargari SL, Amani J, Fattahian Y. In silico analysis of antibody triggering biofilm associated protein in Acinetobacter baumannii. Journal of theoretical biology 2010; 266(2):275-90.
21. George Priya Doss C, Rajith B. Computational refinement of functional single nucleotide polymorphisms associated with ATM gene. PloS one 2012; 7(4):e34573.
22. Kamaraj B, Purohit R. In silico screening and molecular dynamics simulation of disease-associated nsSNP in TYRP1 gene and its structural consequences in OCA3. BioMed research international 2013; 2013:697051.
23. Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, de Castro E *et al.* ExPASy: SIB bioinformatics resource portal. Nucleic acids research (Web Server issue) 2012; 40:W597-603.
24. Guruprasad K, Reddy BV, Pandit MW. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting *in vivo* stability of a protein from its primary sequence. Protein engineering 1990; 4(2):155-61.
25. Gasteiger J. Chemoinformatics: a new field with a long tradition. Analytical and bioanalytical chemistry 2006; 384(1):57-64.
26. Deeds EJ, Ashenberg O, Gerardin J, Shakhnovich EI. Robust protein protein interactions in crowded cellular environments. Proceedings of the National Academy of Sciences of the United States of America 2007; 104(38):14952-7.
27. Deeds EJ, Shakhnovich EI. A structure-centric view of protein evolution, design, and adaptation. Advances in enzymology and related areas of molecular biology 2007; 75:133-91, xi-xii.
28. de Castro E, Sigrist CJ, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger E *et al.* ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic acids research 2006; 34 (Web Server issue):W362-5.
29. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W. improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic acids research 1994; 22(22):4673-80.
30. Notredame C, Higgins DG, Heringa J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. Journal of molecular biology 2000; 302(1):205-17.
31. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic acids research 2002; 30(14):3059-66.
32. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular biology and evolution 2007; 24(8):1596-9.
33. Kumar S, Stecher G, Peterson D, Tamura K. MEGA-CC: computing core of molecular evolutionary genetics analysis program for automated and iterative data analysis. Bioinformatics 2012; 28(20):2685-6.
34. Kumar S, Nei M, Dudley J, Tamura K. MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. Briefings in bioinformatics 2008; 9(4):299-306.
35. Clements MD, Bart HL, Jr., Hurley DL. Isolation and characterization of two distinct growth hormone cDNAs from the tetraploid smallmouth buffalofish (Ictiobus bubalus). General and comparative endocrinology 2004; 136(3):411-8.
36. Vaz BS, Cerqueira GM, Silva JC, Manzke VH, Moreira CG, Moreira HL. Sequence analysis of the growth hormone gene of the South American catfish Rhamdia

quelen. Genetics and molecular research: GMR 2010; 9(4):2184-90.
37. Gonzalez PN, Kristensen E, Morck DW, Boyd S, Hallgrimsson B. Effects of growth hormone on the ontogenetic allometry of craniofacial bones. Evolution & development 2013; 15(2):133-45.
38. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nature reviews Genetics 2009; 10(1):57-63.
39. Wang Z, Moult J. SNPs, protein structure, and disease. Human mutation 2001; 17(4):263-70.
40. Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Minguez P *et al*. The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. Nucleic acids research (Database issue) 2011; 39:D561-8.
41. von Mering C, Huynen M, Jaeggi D, Schmidt S, Bork P, Snel B. STRING: a database of predicted functional associations between proteins. Nucleic acids research 2003; 31(1):258-61.
42. Hong Y, Schartl M. Sequence of the growth hormone (GH) gene from the silver carp (Hypophthalmichthys molitrix) and evolution of GH genes in vertebrates. Biochimica et biophysica acta 1993; 1174(3):285-8.