Thesis for the Degree of Master of Science

Identification of marker genes for immune-related tissues in the Siberian sturgeon *Acipenser baerii* using transcriptomic analysis

By

Mrutyunjay Kumar

Department of Fisheries Biology

The Graduate School

Pukyong National University

August 2022

Identification of marker genes for immune-related tissues in the Siberian sturgeon *Acipenser baerii* using transcriptomic analysis

Advisor: Prof. Nam Yoon Kwon

By

Mrutyunjay Kumar

A thesis submitted in fulfilment of the requirements

For the degree of

Master of Science

Department of Fisheries Biology, The Graduate School,

Pukyong National University

August 2022

Identification of marker genes for immune-related tissues in the Siberian sturgeon *Acipenser baerii* using transcriptomic analysis

A dissertation

By

Mrutyunjay Kumar

Approved by:

Chan Hee Kim (Chairman)

Yoon Kwon Nam

(Member)

Seung Pyo Gong (Member)

August 26th, 2022

Content

List of Tables

List of figures

Identification of marker genes for immune-related tissues in the Siberian sturgeon *Acipenser baerii* using transcriptomic analysis

Mrutyunjay Kumar

Department of Fisheries Biology, The Graduate School

Pukyong National University

Abstract

Immune marker genes define as genes that are highly expressed in an immune cell, tissue, and/or organ type but are lowly expressed in other organ types. These genes are key to the analysis of RNA transcription data, which make them good candidates for subsequent in-depth immunological experimental studies and developing support systems for immunological diagnosis, prognosis and monitoring. However, little is known about immune marker genes that contribute to distinct tissue types associated with immune responses in fish. In this study, we performed transcriptome analysis to identify immune marker genes for each of the kidney, liver and spleen in the Siberian sturgeon (*Acipenser baerii)* that phylogenetically occupies the most basal position below the teleost clade in ray-finned fish. The top 3 immune marker genes were chosen in each tissue i.e. Kidney, liver and spleen. The parameter for choosing the immune marker genes was based on TPM (Transcripts per million) value > 100 and a log difference of ≥5 between each tissue (Kidney, Liver and Spleen). Based on the chosen parameters we identified that in the kidney CTRP9, LRRC19 and PRG3, in the liver C1S, C8A, and C3A and in the spleen CCL25, CCL19 and MAFA genes were the top3 highly specific immune marker genes present in siberian sturgeon

I. INTRODUCTION

Sturgeons (Acipenseriformes) are a family of ray-finned, non-teleost fishes (Actinopterygii) which are also referred to as living fossils (Birstein et al., 2008). Phylogenetically sturgeons are unique as they belong to a Chondrostean fish group that has been regarded as a bridge between Chondrichthyes (Cartilaginous fish) and Osteichthyes (bony fish). Given their unique phylogenetic position, sturgeons are recognized as an invaluable comparative model for studying evolution and diversifying mechanisms in many biological systems in the vertebrate lineage (Kim et al., 2005). Female sturgeons are commercially farmed in most countries specifically for their caviar and meat but unfortunately this has resulted in overfishing and hence this has affected their population (Babaei et al., 2017) and therefore currently they are in endangered condition and as the aquaculture of sturgeons expanded due to its high value there has also been a high risk of bacterial infestations associated with it and due to which a large amount of mortality was seen in farmed sturgeons (Ping et al., 2005). This has further declined the population of this economically important farmed fish. This seems to be a major problem in sturgeon aquaculture and their overall wellbeing.

Immune marker genes are the genes which are seen to be specifically expressed or highly expressed in a specific immune tissue but are lowly/not expressed in any other organs or tissues. The expression of these marker genes is very crucial in understanding the role of that particular immune tissue with which the gene was associated. These genes expression may also change depending upon any pathogen attack for example it was seen that during the transcriptomic analysis of Rainbow trout kidney infected with *Aeromonas salmonids* the expression of IL1B was seen to be significantly higher in the

kidney as compared to other immune genes in that species (Mengqun et al., 2022) which may suggest that IL1B might be an immune marker gene for the kidney, it was also seen that in the kidney of the Atlantic salmon the expression of IL10 was highest when it was infected with the bacterial pathogen (Wang et al., 2014) having high expression of a gene in a particular immune tissue seems to tell a very different story about that gene role in the immune system of that species or how it might help the species its fight against any bacterial infection. This knowledge helps us in finding more information about the species defense mechanism, while much information is available on the teleost immune marker gene but the same information is missing in Siberian sturgeon.

Now for female sturgeons, age 5-9 is a very important phase, as at this age they are generally used for caviar retention (Chapman & Eenennaam, 2019) and thus at this stage they are very vulnerable to any type of pathogen attacks including bacterial, virus or any other diseases. The kidney, liver and spleen are considered the most vital part of the sturgeon immune system because of their high functionality against any bacterial infection (Christin et al., 2021). These vital immune organs protect sturgeons against any pathogen attack and thus help keep the sturgeon safe. However, information on how exactly these organs and the genes involved in it protect the sturgeons against bacterial infection is still very limited and thus not much is known about their function, identifying immune marker genes in these tissues will, fortunately, bridge this gap and help us in understanding how these genes help the sturgeons in its fight against pathogens and while recent advances in high-throughput technologies to survey RNA, especially microarray profiling and RNA sequencing (RNA Seq) have revolutionized the discipline and it enables the study of the fish immune system at the level of the whole transcriptome rather than individual transcripts (Qian, li et al., 2014). The transcriptome refers to the complete set of transcripts in a specific cell, tissue or organism, including all protein-coding messenger RNAs

(mRNAs) and non-coding RNAs (ncRNAs) that regulate gene expression and maintain cellular homeostasis. Earlier before there has been immune-related transcriptome analysis in several aquatic animals, such as the Yellow croaker (*Larimichthys crocea*) (Mun et al., 2014) and Rainbow trout (*Oncorhynchus mykiss*) (Long et al., 2015) which show us the presence of these immune marker genes in the teleost and how these potential immune marker genes help the species in its fight against pathogen attack. Previously, studies on the sturgeon immune system using transcriptomic analysis were done (Luo et al., 2018) and while some information has been found in the immune system of the organism and growth-related genes but not much information has been found about potential immune marker genes of a matured 6-year-old female Siberian sturgeon.

In this study with the help of our local NGS database, we have identified 9 potential immune marker genes in the Kidney, Liver and Spleen tissues respectively in a matured female 6-year-old Siberian sturgeon. From our result, we got the confirmation that CTRP9, LRRC19 and PRG3 genes had the most significant expression (According to TPM) in the kidney as compared to the liver and spleen. In the liver, we have identified C1S, C8A and C3 genes which had the most significant expression and in the end, we have identified CCL19, CCL25 and MAFA protein which had the most significant expression in spleen as compared to kidney and liver. I also found proof that some of the immune marker genes found in my study were a match with the teleost immune marker genes, for example, I found that the sturgeon liver immune marker gene C3 was a potential match with the Nile tilapia C3 immune marker gene (Chang Xu et al., 2018) we also found evidence of CCL19 and CCL20 found in the inflammatory response against the pathogen in *Rachycentron canadum* (Youlu et al., 2012) which thus suggest that sturgeon and some teleost fish might share common immune marker genes. This knowledge can thus help us understand

the immune marker genes information and how they are functioning in each of their immune tissues and thus provide us with more information on the immune marker genes. This was one of the first studies done on a 6-year-old matured female Siberian sturgeon and its immune marker genes information and thus this data thus provides us with more knowledge on their immune system at such a young age and at a very crucial stage at their life. In future studies, we can characterize the function of these genes (e.g., transcriptional responses to immune stimulation or differential expression during organ differentiation) to highlight their utility as tissue-specific immune marker genes.

II. MATERIALS AND METHODS

2.1. Transcriptome analysis

From our local transcriptomic database (AB_TD PacBio Sequel) of *Acipenser baerii* (Female, 6yr old) I found out RNA-seq data from the Kidney, Liver and Spleen. The transcriptomic data I got from the kidney, liver and spleen consisted of around 19,075 genes present in each tissue. After receiving the genes, I then aligned the 19,075 with the reference sequence of *A.baerii* which consisted of a total of 38,190 genes, this step was taken to verify the presence of our taken transcriptomic data or to find out whether it is present in our *A.baerii* reference sequence data. After the alignment, I found that around 19,075 genes matched with our reference sequence of *A.baerii* and thus state that these genes are genuinely present in the kidney, liver and spleen of Siberian sturgeon.

2.2. Functional annotation of the reference genes

For the proper understanding of the matched genes function, I did the functional annotation of my chosen *Acipenser baerii* reference sequence data. This study was done to provide me with more information about the sequences functional information and also help me understand exactly how many genes are involved in the immune function in *Acipenser baerii.* For the functional annotation report, I input around 19,075 genes into the EggNOG (Cepas et al., 2019) database. The database gave me a total in-depth summary of the genes which were involved in the immune system of *Acipenser baerii,* this step was crucial for my study of finding out immune marker genes as I needed genes which were only related to the immune system of the species.

2.3. Sorting the genes and calculation of TPM

After the functional annotation of the reference genes, I sorted the 19,075 genes into 3 tissue-specific categories basically Kidney, Liver and Spleen. This step was done to properly categorize genes based on the immune tissues and for the ease of doing the next step of my study which was identifying potential immune marker genes for each tissue. After dividing the genes I then calculated the TPM (Transcripts per million) values of around 19,075 genes in the kidney, liver and spleen using SALMON TDM software. TPM value calculation was necessary because these values give us an estimate of the expression specificity of each gene in kidney, liver and spleen tissues and thus help me identify more information about any particular immune marker gene. After the sorting, I found that the expressed kidney, liver and spleen genes were 17443, 16914 and 17526 respectively. The genes with no expression values were removed from the Kidney, Liver and Spleen tissues as they were of no value for my experiment.

2.4. Immune marker genes selection

The next step of our study was to remove the genes which had TPM values less than 100 in the kidney, liver and spleen tissues as the values less than 100 seem not to be highly expressed genes and thus these genes are not required for my study. After removing the genes which had TPM values less than 100, I got the end result that almost 998 genes were present in the kidney, 805 genes in the Liver and about 905 genes in Spleen which had a TPM value higher than 100.

My next goal for my experiment was to get rid of the genes which were present commonly among all 3 of these tissues or in the Kidney, Liver and Spleen. This step was taken to identify the DEG genes or differentially expressed genes among all 3 tissues. After removing the common genes across the kidney, liver and spleen I got the final result that the kidney had almost 111 genes which were DEG, the liver had 161 and the spleen had almost 141 genes respectively.

Among the DEG genes that I found in each tissue in the Kidney, Liver and Spleen I did functional annotation of all of the found genes to help me understand which genes were involved in the immune function and I also discarded or did not take the genes which were non-immune specific, this step was taken to identify specific immune marker genes in each chosen immune tissue.

The parameters which were taken for finding out immune marker genes were based on 3 things, first, the genes should be immune specific, secondly, the genes should have a TPM value greater than 100 as compared to the other 2 tissues and third the log value/ difference in expression value should be ≥5. These parameters were taken to identify the most specific immune marker genes in each tissue (Kidney, liver and spleen). Based on these parameters I found the 9 most specific immune marker genes in the Kidney, Liver and Spleen.

2.5. Sequence and structural analysis

The chosen immune marker genes in kidney, liver and spleen sequences were then analyzed for conserved domain or functional part of the sequence using NCBI-CDD search for each immune marker gene. This step was necessary to find out the particular functional domain in the sequence and to verify if the protein contains immune functionality. After the conserved domain analysis, the sequence 3D structure prediction was done using the SWISS-MODEL web server to find out the best structural match for our immune marker gene and to properly visualize the structure of our immune marker genes. For finding out the quality report of the structure, we used the Q-MEAN value to determine the best structure for the protein.

III. RESULTS AND DISCUSSION

3.1. Transcriptional analysis and annotation report

A total of 19,075 genes were taken signaling, metabolism and some of them were poorly characterized. After the sequence functional analysis I did the total analysis of the length of the sequences of our reference sequence. We found that the average length of our sequence was around 2581 bp for a total of 49249763 bp length of all sequences (Fig. 1). When looking at the orthologous group's distribution I saw that around 9020/12.73% of the genes belonged to the eukaryote and the lowest number of genes around 272/0.38% belonged to the Euarchontoglires (Fig. 2).

During the transcriptomic analysis, we found that around 607 of the contigs were involved in the transcription and around 449 of the contigs were involved in the RNA processing and modification and around 283 of the contigs were involved in the translation, ribosomal structure and biogenesis, around 153 of the contigs were involved in the replication, recombination and repair and around 95 of the contigs were involved in the chromatin structure and dynamics (Fig. 3).

In the cellular processes and signalling, we found a total of 1195 contigs related to the signal transduction mechanism and around 814 contigs related to the post-translation modification, whereas we got around 40 genes which were related to the defense mechanism and which were functioning in the immune system in the *Acipenser baerii* (Fig. 4). This gave us prior knowledge of the sturgeon immune system and its position. In the Metabolism function, we found that around 222 contigs were related to energy production and 201 to

carbohydrate transport and metabolism (Fig. 5). I also found out that around 2501 of the contigs had function unknown and 0 of them had almost zero general function prediction only (Fig. 6)

EggNOG Annotation Report

Input Data

AB_TD.corrected_LR.fasta

General Information

Total amount of input sequences:	19075
Average length:	2581.0
Number of GO annotated sequences:	7197 / 37.73%
Number of GO annotations:	140794
Average GOs per sequence:	19.56

COG Categories Distribution

Information Storage and Processing:	
Transcription (K):	607
RNA processing and modification (A):	448
Translation, ribosomal structure and biogenesis (J):	283
Replication, recombination and repair (L):	152
Chromatin structure and dynamics (B):	95
Total:	**1585 / 17.82%**
Cellular Processes and Signaling:	
Signal transduction mechanisms (T):	1195
Posttranslational modification, protein turnover, chaperones (O):	814
Intracellular trafficking, secretion, and vesicular transport (U):	462
Cytoskeleton (Z):	327
Cell cycle control, cell division, chromosome partitioning (D):	190
Extracellular structures (W):	46
Defense mechanisms (V):	40
Cell wall/membrane/envelope biogenesis (M):	27
Nuclear structure (Y):	4
Cell motility (N):	2
Total:	**3107 / 34.93%**
Metabolism:	
Energy production and conversion (C):	222
Carbohydrate transport and metabolism (G):	201
Lipid transport and metabolism (I):	156
Inorganic ion transport and metabolism (P):	148
Amino acid transport and metabolism (E):	140
Nucleotide transport and metabolism (F):	82
Secondary metabolites biosynthesis, transport and catabolism (Q):	64
Coenzyme transport and metabolism (H):	37
Total:	**1050 / 11.81%**
Poorly Characterized:	
Function unknown (S):	2501
General function prediction only (R):	0
Total:	**2501 / 28.12%**

Figure 1. Functional annotation report of the reference genes (19,075).

Top 10 Orthologous Groups Distribution

root:	9034 / 12.75%
Eukaryota:	9020 / 12.73%
Opisthokonta:	8833 / 12.46%
Metazoa:	8812 / 12.43%
Bilateria:	8797 / 12.41%
Chordata:	8757 / 12.36%
Vertebrata <vertebrates>:	8746 / 12.34%
Actinopterygii:	6798 / 9.59%
Mammalia:	590 / 0.83%
Testudines:	272 / 0.38%
Euarchontoglires:	272 / 0.38%

Figure 2. Selected genes orthologous groups distribution

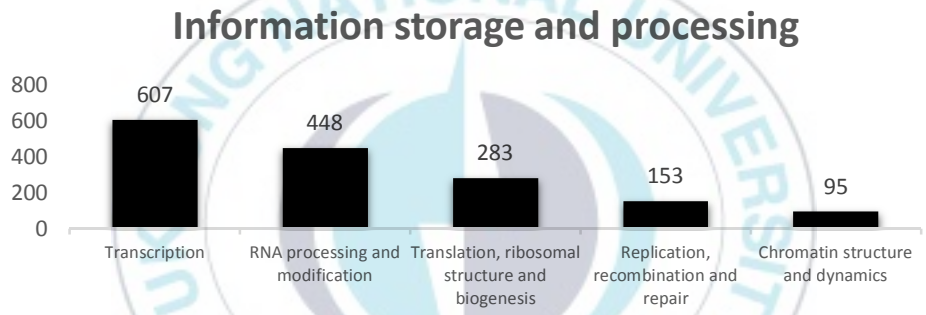


Figure 3. Information storage and processing total contigs involved

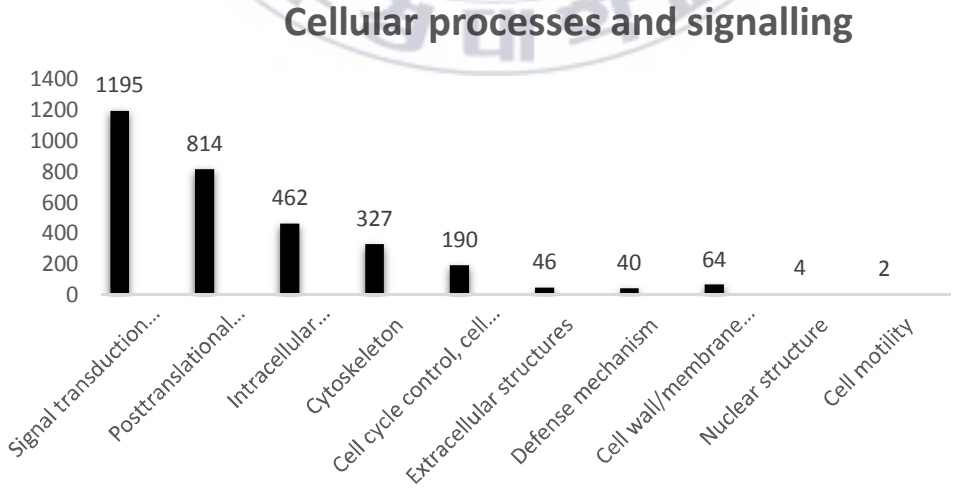


Figure 4. Cellular processes and signaling total contigs involved

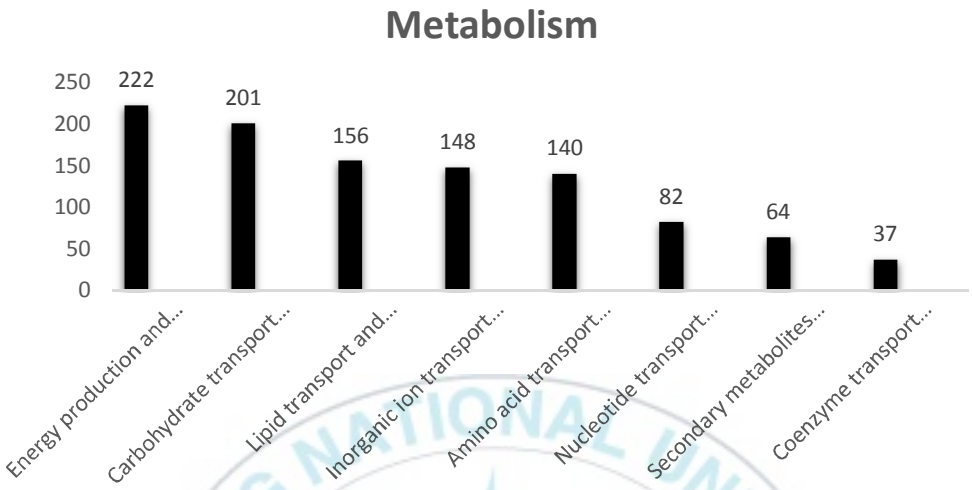


Figure 5. Total contigs involved in metabolism process

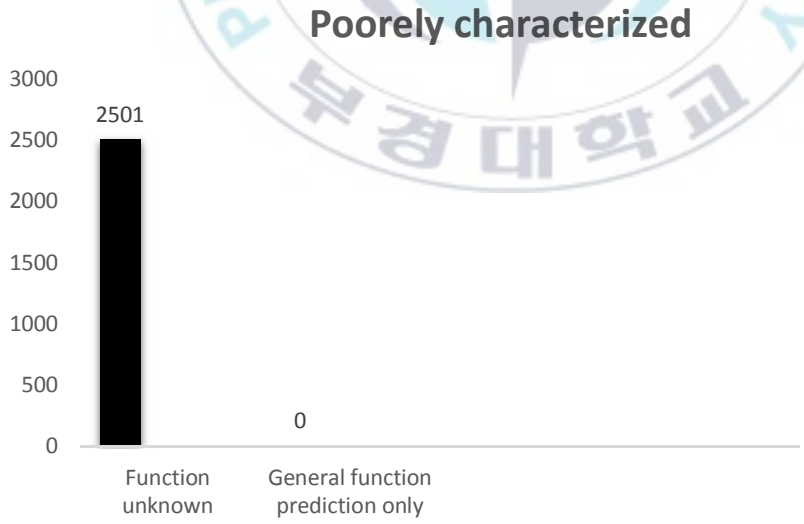


Figure 6. Total contigs which were poorly characterized

3.2. Sorting for immune marker genes

I collected a total of 19,075 genes from the kidney, liver and spleen immune tissues respectively. Our next goal was to sort out the genes based on the expression parameters or based upon TPM values, for this purpose we divided the genes into 3 sections, broadly categorizing them into 3 parts such as kidney, liver and spleen with each containing 19,075 genes respectively (Table 1). From the selected immune genes we had to trim down the genes with low expression values or values below 100 TPM (Transcripts per million), this step was taken to find the genes which had higher expression values in the kidney, liver and spleen respectively. Upon removing the genes which had TPM<100 we left out around 997 genes in the kidney, 804 genes in the liver and 904 in the spleen respectively (Table 2). After we found this data we then again arranged them in another table to properly sorting into a better way. Our next goal was to remove the similar or common genes which were present among all tissues i.e. kidney, liver and spleen respectively. This step was taken to find out the DEG (Differentially expressed genes) in each immune tissue. We got a total of 111 genes in the Kidney, 160 in the liver and 140 in the spleen respectively (Table 2). These genes were genes which were DEG in all of the 3 immune tissues and were expressed significantly in each immune tissue. After gathering all of the DEG genes in the kidney, liver and spleen we did the functional annotation of the genes and sorted the genes based on immune functionality. We did this step to find out genes which were related to immune function and discarded the genes which were non-immune to properly find out the immune marker genes. Upon removing the non-immune genes we found 3 immune marker genes in each tissue i.e. kidney, liver and spleen respectively. These immune marker genes were picked based on immune function, TPM value and specificity regarding immune tissue.

Table 1.Total number of genes taken for the study

Kidney total number of genes	Liver total number of genes	Spleen total number of genes
19,075	19,075	19,075

Table 2. Immune marker gene mining process

Filtering parameters	Kidney tissue genes	Liver tissue genes	Spleen tissue genes
Total genes present in the tissues	19,075	19,075	19,075
Total number of expressed genes	17,443	16,914	17,526
After removing less expressed genes in each tissue (Below 100 TPM)	997	804	904
After removing similar genes among all tissues (DEG Genes)	111	160	140
Finding immune marker genes	3	3	3

3.3. Functional annotation of DEG

3.3.1. Kidney DEGs

During the functional annotation of the total number of kidney DEG genes (111) we found that around 12 genes were related to the function (carbohydrate metabolism and transport), 25 in the (post-translational modification, protein turnover and chaperone function), 3 of the genes in the (rna processing and modification), 7 of them related to the (lipid metabolism), 24 of the genes related to the (energy production and conversion), 3 of them were related to (transcription process), 3 in the (rna processing and modification), 5 of them functioning in the (amino acid metabolism and transport), 1 in the trafficking and secretion, 10 in the (signal transduction) process, 2 in the (secondary structure), 6 in the (cytoskeleton), 2 in the (defence mechanism), 3 in the (extracellular structure), 1 in the (nucleotide metabolism and transport), 8 in the (inorganic ion transport) and metabolism) and in the end 24 of the genes whose functions were unknown (Fig. 7). After the functional annotation, we then removed the genes which had not related to the immune system and found 3 potential immune marker genes in the kidney.

3.3.2. Liver DEGs

During our Liver DEG genes functional annotation, we found that around 22 of the genes were related to the (signal transduction), 10 functioned in the (Inorganic ion transport and metabolism), 42 of the genes were the genes which had their (function unknown), 6 of the genes were present in the (defense mechanism),14 of the genes were related to the (amino acid metabolism) and (transport) 3 of them related to the (coenzyme metabolism),47 of the genes were related to the (post-translational modification), (protein turnover and chaperone functions), 8 of the genes were related to the secondary structure and 4 of the

genes were related to the (carbohydrate metabolism and transport), 5 of the genes were related to the (extracellular structures), 1 gene related to the (cytoskeleton), 2 in the (translation process), 1 in the (transcription), 1 in the (cell wall/membrane/envelop biogenesis), 7 in the (energy production and conversion) and 1 of them related to the (nucleotide metabolism and transport) (Fig. 8). From these genes, we identified the 3 most suitable immune marker genes in the Siberian sturgeon liver.

3.3.3. Spleen DEGs

During our functional annotation of the Spleen DEG genes, we found that 2 genes were related to the chromatin structure and dynamics, 1 in the Nucleotide transport and metabolism, 2 in the Energy production and conversion, 9 in the (cytoskeleton), 5 in the (post-translational modification), (protein turnover and chaperone), 1 in the (lipid transport and metabolism), 14 in the (signal transduction mechanism), whereas 25 of the genes had the (function unknown), 10 of the genes were related to the (RNA processing and modification), 7 in the (transcription), 13 in the (translation, ribosomal structure and biogenesis), 2 in the (defence mechanism), 4 in the (extracellular structures) for a total of 95 of the genes (Fig. 9). After the functional annotation, we selected 3 genes based on the functional property and also looked for any possible important information for the genes.

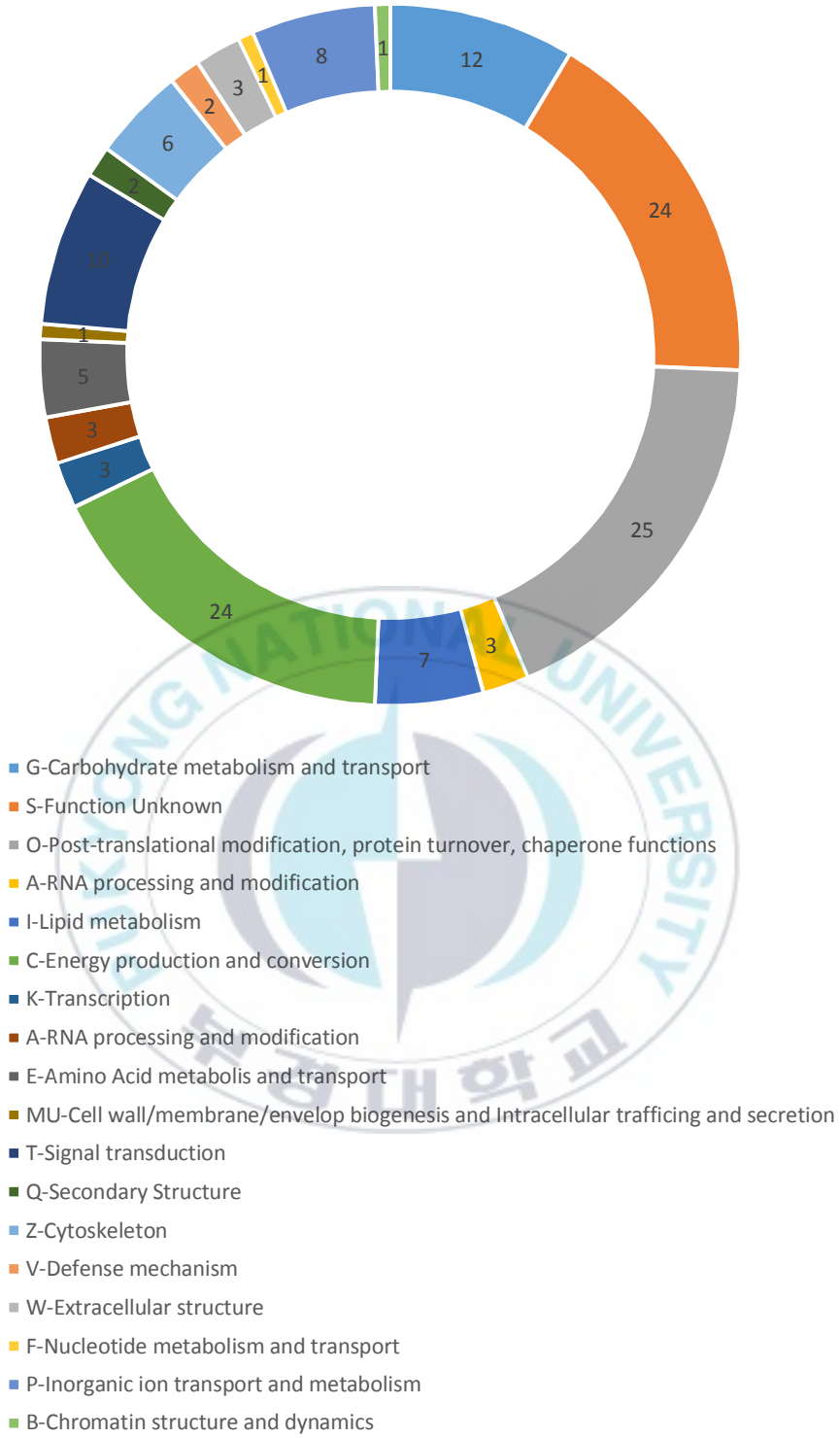


Figure 7. Functional annotation of Kidney DEGs

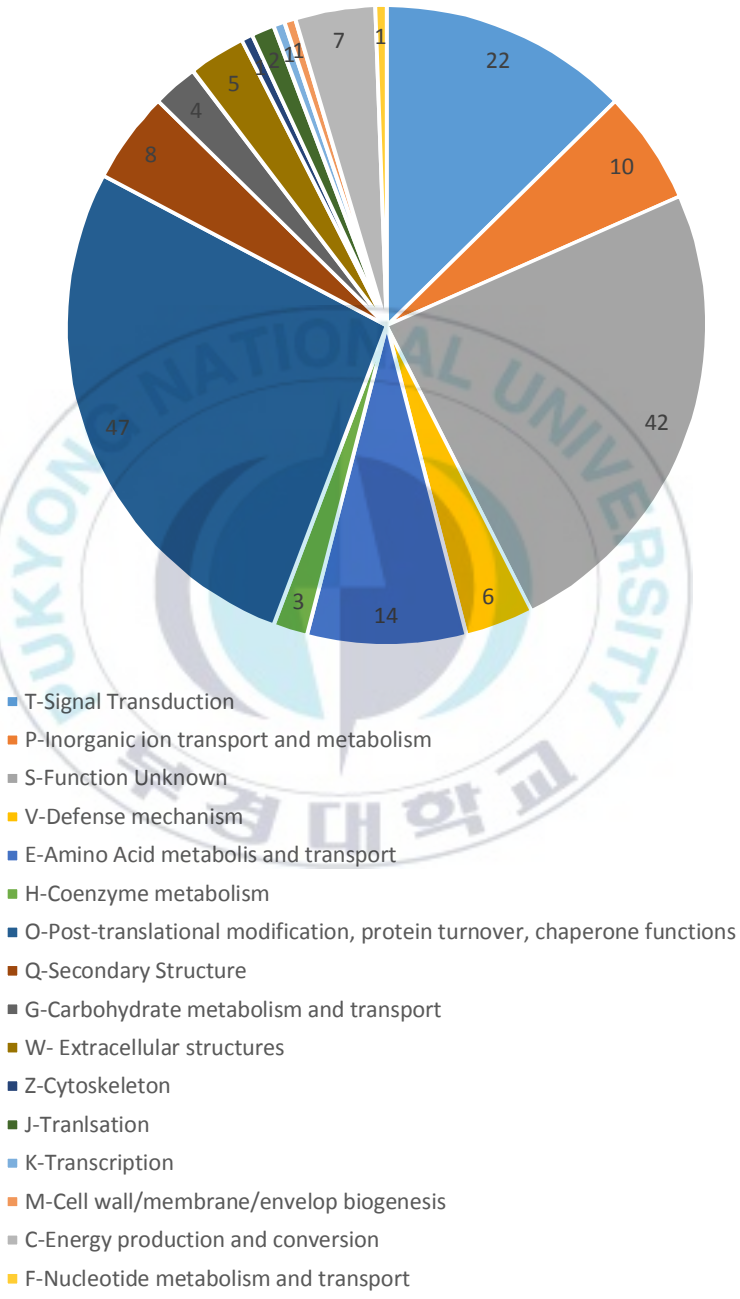


Figure 8. Functional annotation of the liver DEGs

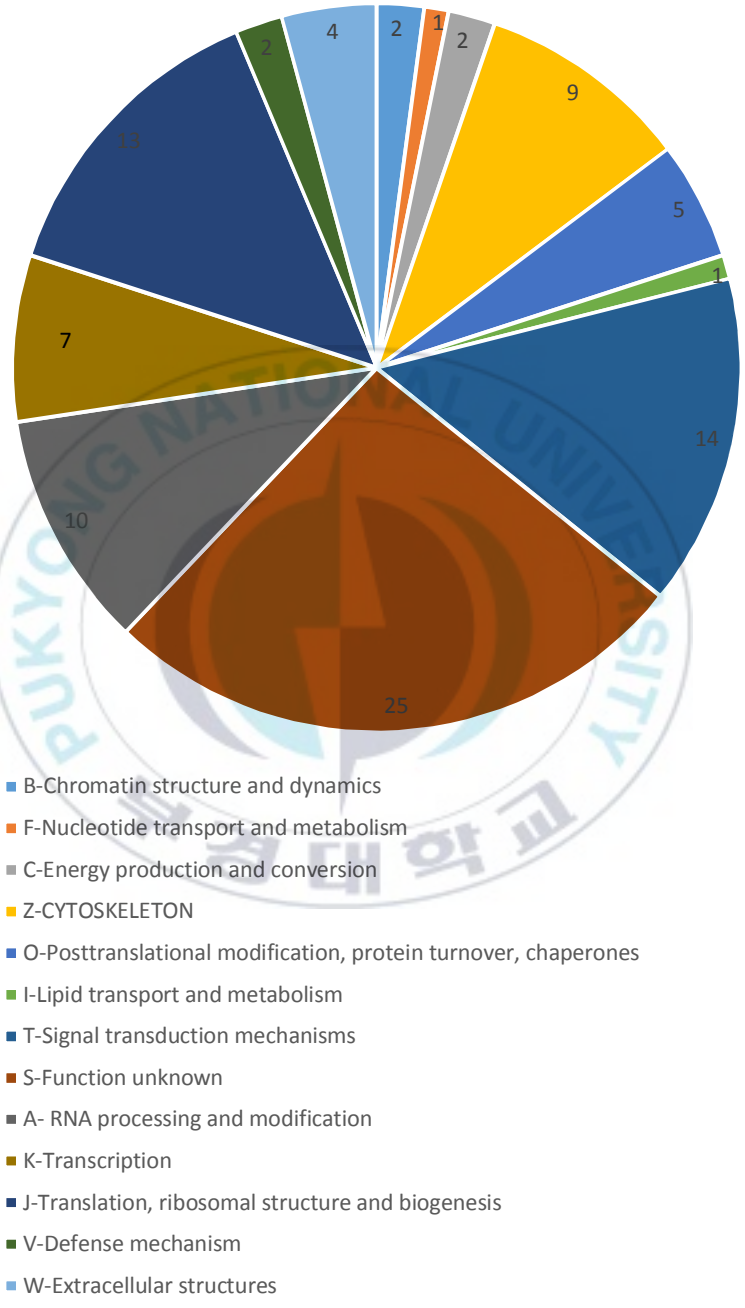


Figure 9. Functional annotation of the spleen DEGs

Table 3. Selected kidney immune marker genes

Tissue	Gene name	TPM(K)	TPM(L)	TPM(S)	LOG Difference
Kidney	1-CTRP9	320.4	0	0	8.3
	2-LRRC19	101.77	0.04	0	8.85
	3-PRG3	158.7	3.71	2.8	5.6

Table 4. Selected liver immune marker genes

Tissue	Gene name	TPM(K)	TPM(L)	TPM(S)	LOG Difference
Liver	1-C8A	0	172	0.1	11.15
	2-C3	0.2	539.6	0.7	10.28
	3-C1S	0.2	473	0.9	9.81

Table 5. Selected spleen immune marker genes

Tissue	Gene name	TPM(K)	TPM(L)	TPM(S)	LOG Difference
Spleen	1-MAFA	0	1.8	1456	10.05
	2-CCL25	2.07	2.6	105.3	5.45
	3-CCL19	38	7.6	574.8	5

After the functional annotation of the DEG genes, I then selected the 3 most specific immune marker genes which seem to have the highest expression in each particular tissue (kidney, liver and spleen). These immune marker genes were based on the 3 most important parameters, first, the genes should be related to the immune system, secondly, the gene TPM value should be significantly higher than the other 2 tissues and third the gene should have a log value difference of ≥ 5 between other 2 tissues. After the selection, we found 3 most specific immune marker genes in the kidney (CTRP9, LRRC19 and PRG3) (Table. 3), the liver (C8A, C3 and C1S) (Table. 4) and in the spleen (MAFA, CCL25 and CCL19) (Table. 5) which had the most significant expression in respective immune tissue and thus are the most potential immune marker genes found in Siberian sturgeon.

3.4. Bio-informatics sequence analysis

3.4.1. Kidney immune marker gene

1. CTRP9

C1q/necrosis-related protein 9 (CTRP9) is a newly discovered adipokine that is the closest paralog of adiponectin.CTRP9 is also involved in the signaling pathways and physiological functions, which helps in the understanding of the occurrence of diseases.CTRP9 can also be used in therapeutic intervention (Guan et al., 2022). CTRP9 has also attracted attention due to its role in the pathogenesis of various diseases, and it also serves as a potential biomarker or therapeutic target. We found a conserved functional domain in the sequence of the *Acipenser baerii* CTRP9. During the conserved domain analysis, we found that it almost had around 131AA (Amino acid) involved in the immune function of the C1Q domain and also it had a collagen family domain involved also. During my 3D structure analysis (Figure 10). I found that the sturgeon CTRP9

structure had the maximum similarity with the crystal structure of a (single-chain trimer of the human adiponectin globular domain). This states that this protein was thus involved in the immune system of the sturgeon and thus provides us with much knowledge on the sturgeon immune system.

2. LRRC19

LRRC19 (Leucine Rich Repeat Containing 19) is a receptor pathogen-recognition receptor which mediates the activation of TRAF2 and TRAF6 NF-kappa-B signaling pathways and induces the expression of pro-inflammatory cytokines, chemokines and antimicrobial substances (Fuchs et al., 2016). While its presence in gut, involved in host-microbiota interactions, plays a critical role in the recruitment of immune cells and intestinal inflammation. We found that in *Acipenser baerii* the presence of LRRC19 seemed to be very high in kidney tissue due to the fact that it had an almost higher significant TPM value as compared to the other two immune tissues (Liver and Spleen). During our sequence domain analysis using the NCBI CDD tool (Figure 11), we found that the sequence had an LRRC domain present and during our 3D structure analysis we found that the *Acipenser baerii* LRRC19 had the maximum similarity with the (Slit homolog 2 protein N–product). Thus it seems to us that this protein was involved in the *Acipenser baerii* immune function.

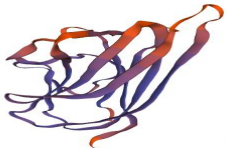
3. PRG3

PRG3 (Proteoglycan 3) is an extracellular matrix structural constituent conferring compression resistance. Involved in several processes, including granulocyte activation; histamine biosynthesis process; regulation of gene expression. It's also involved in the innate immune system and also possesses similar cytotoxic and cyto-stimulatory activities to PRG2/MBP (Stelzer et al.,

2016. While in vitro it stimulates neutrophil superoxide production and IL8 release and histamine and leukotriene C4 release from basophils) during our sequence analysis, we found that the sequence contained C-Type lectin domain present and during our 3D structure analysis we found that the structure shows similarity with the (C-Type mannose receptor-2) structure (Figure 12). We found that this protein is present in the Siberian sturgeon and its expression was seen to be highest in the Kidney as compared to the liver and spleen.

Total conserved domain – 131 AA

RFPRGKKQKNKTDKTFEKMCRVAVLLCVVLLVEMGTVLCIEEKPCQLGIPGIPGAPGPQGVPGKDGRDGPAGPRGEKGDPGAPGERGHRGPPGKAGPPGAFD
PSIHGGSAAEVCAFHVGLGAGGPAAGEPIKFTKVTYNEQNAYSPASGKFTAPVPGLYFFTYQITVFNNDMAWTLMRNKQKVQVYTYMADLKRTTQASGASLIK
LGVGEEVWLQVHGSQNGLYADADDDTTFTGFLLQQITKIAITQPVH-DSNHTACALR-QPHSLCTKIATTQPVH-DSNHTACALR-
QPHSLCTKIATTQPVH-DSNHTACALR-QPHSLCTKIATTQPVH-DSNHTACALR-QPHSLCTKIATHSLCTKIATTQPVH-DSNHTACALR-
QPHSLCTKIATTQPVH-DSNHTACALR-QPHSLCTKIATTQPVH-DSNTQPVH-DSNHTACALR-QPHSLCTKIATTQPVH-DSNHTACALR-
QPHSLCTKIATTQPVH-DSNHTACALR-QPHSLCTKIATTQPVH-DSNHTACALR-QHTACALR-QPHSLCTKIATHSLCTKIATTQPVH-DSNHTACALR-
QPHSLCTKIATTQPVH-DSNTQPVH-DSNHTACALR-QPHSLCTKIATTQPVH-DSNHTACALR-
QPHSLCTKIATHSLCTKIATTQPVHDSNHTACALRINVKHSKQISFYFFFFPLCVOCHLCLHLSVKIYMT-NV-F-KKLNEIMHNIEQ-R-INGYIDSVS-
CM-RNNQQIDRFTVNRNK-VKLPWHQKSGRTSSTVYIYYYYYYYYYYYYYYYYYYY---STVYIYVC-YSMNTNTEIR-GLSCFKFLLQNNTGVNYSYNKHRT



3D structure of CTRP9 Q-Mean score of the structure

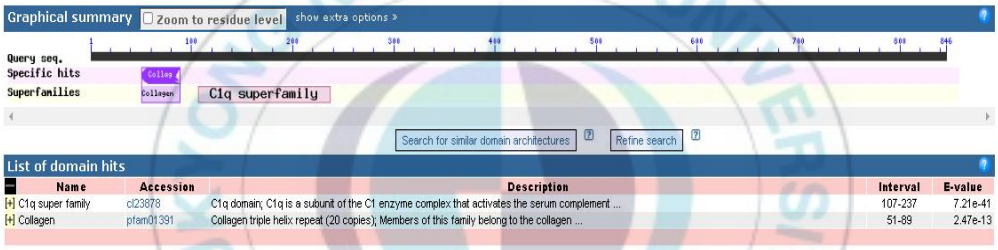
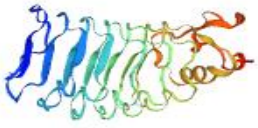
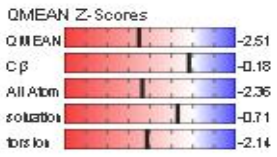


Figure 10.Function and 3D structure of CTRP9

Total conserved domain- 236 AA

KKQSSNYISHLYLFSANLMMMAACYLLLVWTASFFLASGFANDDICDYSNRNLTAIPLDLPLNTTKLILSYNKIDIYPNSIGVLNKYTNLSELYLN
HNNISDLPGHMFHKLSKLKVLNIANNIISRIEPKAFEGLGSLQDLDLCHNQINYLHPEVFANLSSLKNLTLRGNKLRTLKEGTLNLFRLTSINLQD
NPWNCTCKLLSLQKRVKDTNVTLANYNKTTCESPTELKGQCILTVLINGTQCSAGSTTTKAPQFLTIQSKGSIARDSSSVTNETKTKEQDSQPIGS
SWHFLLGVIVAALSTAMLIVCAVKCPTWYKMFFSYQHQQLQEEDLDTLENMGKFTDDEDGFIEDRYIDSVECNDNREN-
MVLDNPMQQERHVRPLISKDVNVFIRS-RPV-KWPATWRIT-LYNLSLCVLE-HPCLKPCTKILWHRYIKK-QQKDYSM-LPGLLNT-TMKAM-
VNAPQNT-I-LFIIKCVHFKYQN--HTHLIFTLAVPRFHLLVFGIFNVL-KHQFRAMTN-
MWWCYNKLYIQAQKLKPCTCLGFLPHKQILNCKSLDSHFLPELSWLNYCLYYFPSLKRLQKQENL-HISWGMLCVFLMQLI-RYQMSI-KSV-F-
LY-YHVLHKFRQCYF-YAR-YCFSTYSRLSVKMTSMLLASLSFCALS-KKKTGLTNLIHFFFSKLLLFVSSQIKSK





3D structure of LRRC19 Q-Mean score of the structure

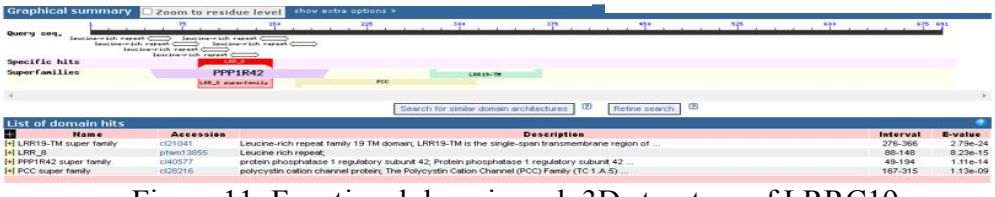
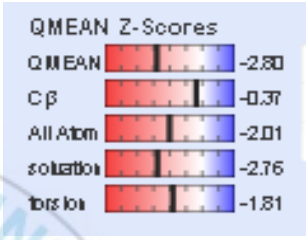


Figure 11. Functional domain and 3D structure of LRRC19

Total conserved domain-110 AA

LPAGHWTSPVSPSIFCVYFPICLTKMKSLFLLVLLCSAELISSFAVPANVEVDADSAAHADCAAHADCAAHADCAADADCAVQADSAVHADCAAHADSAEHADCTPNA
DCAADVEVKDWGAYCNATFPGLCETQSYSDGCHRYTFHGQAVSFSCAEIHCQRSHCGHLASVNNHCTNHALICLTVQCNHAVAAWIGGQMTCTGFQWTDGSCWSYQNW
YQGCPRNCGGSCVVLNWGEAGKWCNSPCSESRPFICKGTR-GMKIHPVSSTLRLAVV-
NGAPSYVRLYWNHRSVTLKYGLKTTHLDRGSSLKYHCVNAFCFTFLCVMCFWYSLPYCLKF-FYTK-RQYRPATDRGQNKALSFEDEDSRLFKV-
KHTHTISDPFSLSLL-GILVIKYN-EEGLPSIKTRLQITL-K-IDILIHIIEGALTIYKTHSLASY-
LHHWTLPLKERPSLKSILSPLISFIINIITGLFQATGGDFERYVEYNIIRNINVVFCFLKPSVTV-KKC-TFFTLCKILHALCF-MQ-NSV-S

3D structure of PRG3



Q-mean score of the structure

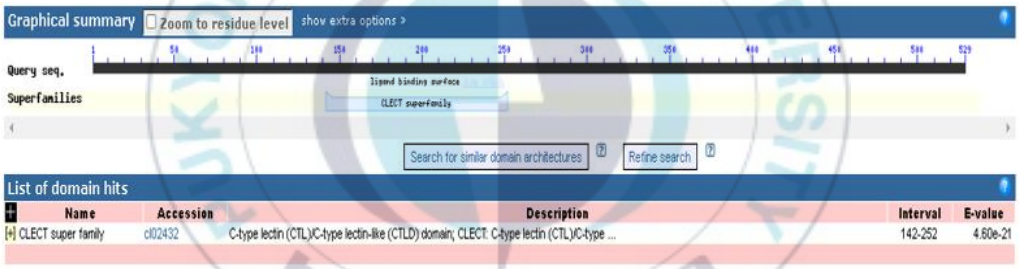


Figure 12. Functional domain and 3D structure of PRG3

3.4.2. Liver immune marker genes

1. C1S protein

This gene encodes a serine protease, which is a major constituent of the human complement subcomponent C1. C1S associates with two other complement components C1r and C1q in order to yield the first component of the serum complement system. Defects in this gene are the cause of selective C1s deficiency. Diseases associated with C1S include complement component C1s deficiency (Safran et al, 2022). Among its related pathways, it includes immune response lectin induced complement pathway and innate immune system. It was also found out that it's a paralog of C1R. During our sequence functional domain analysis, we found that the sequence had MBL (Mannose-binding lectin domain) present. During our sequence 3D structural analysis using SWISS-MODEL, we found that the sequence had the maximum similarity with the (MBL-ficolin associated protein-1) structure (Figure 13). We also found that this gene was involved in the immune system of the *Acipenser baerii* and its expression seems to express in the liver as compared to the kidney and spleen making this a liver immune marker gene.

2. C8A protein

C8 is a component of the complement system and contains three polypeptides, alpha, beta and gamma.This gene encodes the alpha subunit of C8.C8 participates in the formation of the membrane attack complex (MAC).The MAC assembles on the bacterial membranes to form a pore, permitting disruption of bacterial membrane organization. Mutations in this gene cause complement C8 alpha-gamma deficiency.C8A(Complement C8 Alpha Chain) is a protein-coding gene. Diseases associated with C8A include Complement component 8 Deficiency, Type I and Immunodeficiency due to a late component of

complement deficiency (Rappaport et al., 2016) Among its related pathways are complement cascade and immune response lectin-induced complement pathway. It's also a very important constituent of the membrane attack complex (MAC) that plays a key role in the innate and adaptive immune response by forming pores in the plasma membrane of the target cells. C8A inserts into the target membrane but does not form pores by itself. During our sequence functional domain analysis we found that the sequence had a MAC complex domain in the sequence) and during our 3D structure analysis, we found that the structure had the maximum similarity with the (Complement component C8 alpha chain cryoEM structure of 2C9-Smac) (Figure 14) .I also found that this protein was involved in the Siberian sturgeon liver tissue immune function and according to the TPM we found that this gene was highly present in the Siberian sturgeon especially the Liver as compared to kidney and spleen

3. C3 protein

Complement component C3 plays a central role in the activation of complement system. Its activation is required for both classical and alternative complement activation pathways.The encoded preprotein is proteolytically processed to generate alpha and beta subunits that form the mature protein, which is then further processed to generate numerous peptide products (Fuchs et al., 2016). The C3A peptide also known as the C3A anaphylatox, modulates inflammation and possesses antimicrobial activity. Diseases associated with often include complement component 3 deficiency, Autosomal recessive and hemolytic uremix syndrome. Among its related pathways are immune response lectin-induced complement pathway and peptide ligand-binding receptors. During the sequence analysis, we found that the sequence had the functional domain present which is the Complement C3 domain and during our 3D structural analysis we found that the structure had the maximum similarity with

the (Mammalian C3 with an intact thioester at a 3A resolution) (Figure 15). This protein was found in the *Acipenser baerii* liver tissue and was expressed significantly higher as compared to kidney and spleen, thus making it an immune marker gene in the Siberian sturgeon.

Total conserved domain- 266 AA

FDPRDCRLLKMRLSHLLSLSFSLFLAPSLSEGVSQMFGLVVSPLFPQGYPNNVLETWDIAVPPGYAIQLHLTHLDIENSENCEYD
FLKVSVDGVVLSTLCGEKSFSELRSTVNPALYSTLGSMRLTFQSDFSNTERHTGFSAHYSAVDINECDDPYTGCSHDCSNYIGGF
YCSCQPGYLLQEDNTTCAVNCSGEVFSSLLGSLSSPSYPGLYPENSACSYSLQVEEGFQLVLAFEGVFDIESGERGACVDSLRIR
AGGREWGPYCGNVAPKDPVVTNSNHAEVLFNTDGYGANKGWKIKYTTKAKTCPDFVTANSVLEPKRPVQQFKDQVRVTCNKGHEI
VIGAKAKHHFDSWCQKNGVWSDIEICEPVDCGPPKEIDNAELKFNKESSPTTYQSQIYYECEAPYYQINTDGKAQFTCEADGMWR
SSDGKLDIPHCIPVCGRPSKPLESIQRIFGGKPAQLGNVPWQVFFQSPRGGGALISDRWVLTVAHMAQVDQPYIYMGFLDARHLG
QAVRLEVEKMIIHPKYKRHGGLRTVQLRQRHRARQTQGQGEDGAQCDPHLPPSERGRRATRAGEDRVHIGMGASQTETDWQDTSV
PSTAVRDDPCSKDTELYRREAHRTRDSEGEVCDHRHVLCWDARRRQLSRRQWRGLCVPGRGQRVPVEGACVLGDRMWFLWVLHKG
RELPGLDTRHHGSKRRLNGTVEFDTAVSLLGVHSFSTFLQSLASLKLQTHSCVTLKYNPFVSVGSFTTFLLCLLSLKLHKDTVP-
FSPVCLHFFSCRDTSFLPCLPLHVRNPVCNLEHMKKMLCIINRVR

3D structure of C1S

Q-Mean score of the structure

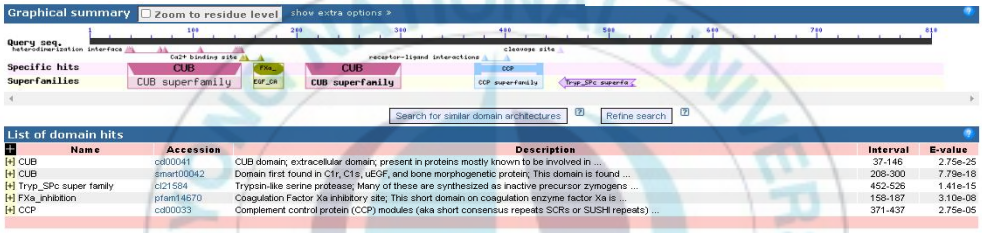
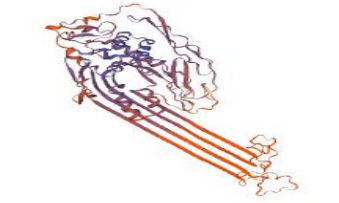
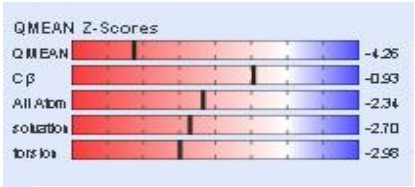


Figure-13-Functional domain and 3D structure of C1S

Total conserved domain-255 AA



3D structure of C8A



Q-mean score of the structure

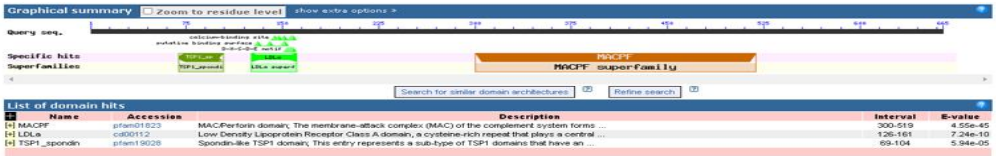
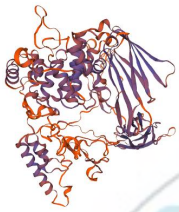


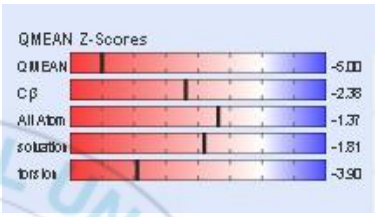
Figure 14-Functional domain and 3Dstructure of C8A protein

Total conserved domain-352 AA

LESKYADKHIKKCCRDGMKEIRMDYTCEKRSQYITSSKECVDAFLHCCTEIAKRRKEGQRDDLTLARSDEDDEYMSDDEIVSRTEFPESWLWDV
VKLPSGPKDRNGLVSLDLRHFLKDSITTWEILAISVSPQKDKQPPLGICVADPYEVTVQDFFIDLKLPYSVVRNEQVEIKAVLYNYGSDDITVR
IELLENELVCSAASQKRKYRQEVPVEASSSRAVPFIIIPLKRGEVEIEVKASVKGTLASDGVRKKLLVVPEGVKVKKAVRSVVLNPSIVVGGVQ
NEYVPPVDLRSIVPDSDAQTFINVQEILGQTLESAISGDKLGHLITVPGGCGEQNMITMTPPVIATLYLDKTNQWEYAGVQRRAEAIRNIKQGY
TQQLVYRQKDSSYAAWQNNQEVPGSPRTRCSPWHTPRLWMRMCCVKPNGTSSSQTAHSERMPLLFTERWWEARDLQNLKPLSQLLSSLCWNRKK
SANSWSLPCQEALKNLQAIWRNASRAYRLRTLWPRHTLPAAVLITTPFSDSHRQIEAIGLYLHPISSPLKPQPTPSWPVRRKQYDRAGKAVKWL
TEQRFYGGGYSSTQATIIVFQAIADYMVAVPLVQDVNLDVELSLSGRSKPVKWKISNDNIYMPRSEKARIDQAFNVTAKGVGQGTLTVMTFYNA
IPEGKSTDCKNFELDVSITRPNNVTRSLGHARNKNLCNLSGYKRRHDVYRYHHANRIQTQGRDDEWSGPVFSEVDGHSSFRGLSHPLQDISQVQ
GLHLFSAQGVSGSDPASCSHCLVLRLREPLYEILPSRQRKRAAEQDLPRGSVQVCRRKLQFTEKTRSQNHCESETHCLTWNGVRVQSQTSERQT
KLNILLHHGHANYQRREKSFTQEERLHLPQQLPNSPGPAGGQVLPNHGQRSGSVETGKRDELHPRRRDLDRVVADQGMPRHARHMPGHPIIRRR
NGLRLPQLNTVCGANLALRSAESFTICSCLKSILTHRFFLLLLLLLIFVILPLLIQVMPLTHY



3D structure of C3



Q-Mean score of the structure

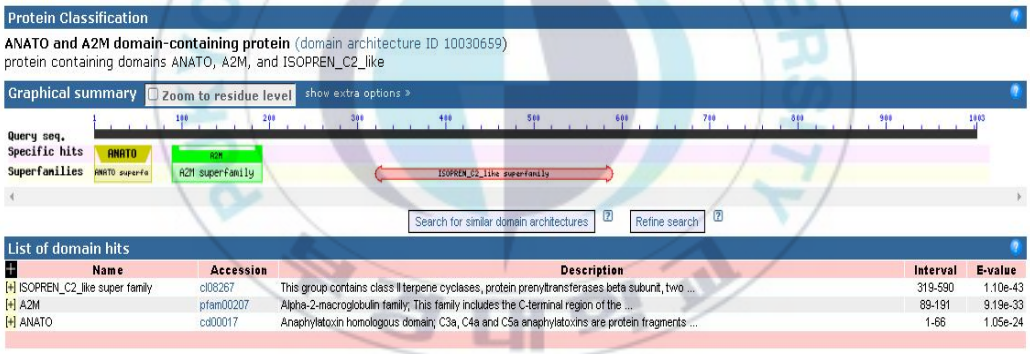


Figure 15- Functional domain and 3D structure of C3 protein

3.4.3. Spleen immune marker genes

1. CCL25 protein

This antimicrobial gene belongs to the subfamily of small cytokine CC genes.Cytokines are a family of selected proteins involved in immunoregulatory and inflammatory processes.The CC cytokines are proteins characterized by two adjacent cysteines (Belinky et al., 2015) The cytokine encoded by this gene displays chemotactic activity for dendritic cells, thymocytes, and activated macrophages but is inactive on peripheral blood lymphocytes and neutrophils. During our sequence analysis, we found that the structure had the CXC domain present and during our 3D structure analysis we found that the structure had the highest similarity with the HUMAN chemokine CCL19)(Figure 16). During our analysis we found out that this protein was expressed most significantly in the *Acipenser baerii* spleen immune tissue was seen to the immune marker gene in the spleen.

2. CCL19 protein

This antimicrobial gene is one of the several CC cytokine genes clustered on the P-arm of chromosome 9. Cytokines are a family of secreted proteins involved in immune regulatory and inflammatory processes. The CC cytokines are proteins characterized by two adjacent cysteines (Fishilevich et al., 2017). The cytokine encoded by this gene may play a role in normal lymphocyte recirculation and homing. It also plays an important role in trafficking of T cells in the thymus, and in T cell and B cell migration to secondary lymphoid organs. It specifically binds to the chemokine receptor CCR7 (Figure 17). Among its related pathways are cytokine signaling in the immune system and PD induced signaling. It was also seen that it also plays a role in normal lymphocyte

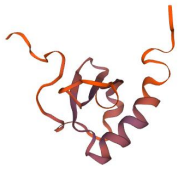
recirculation and homing and in the trafficking of T-cells in the thymus, and T-cells and B-cells migration to secondary lymphoid organs. Binds to chemokine receptor CCR7. Recombinant CCL19 shows potent chemotactic activity for T-CELLS and B-CELLS but not for granulocytes and monocytes. During our sequence functional domain analysis we found that the sequence had CXC domain present and also it was seen that the sequence 3D structure had the maximum similarity with the (Human lymphotactin structure). It was seen that this proteins maximum expression was seen in the spleen immune tissue and not in any other immune tissue, making it an immune marker gene in the Siberian sturgeon.

3. MAFA protein

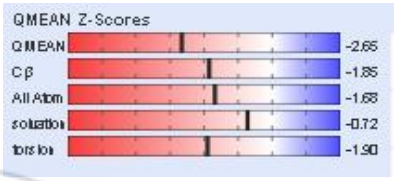
MAFA is a transcription factor that binds RIPE3b, a conserved enhancer element that regulates pancreatic beta-cell specific expression of the insulin genes (Guan et al., 2022). It's also involved in the regulation of beta-cell development and glucose/energy metabolism. Gene ontology annotations related to this gene include DNA-binding transcription factor activity and sesNA binding. During our sequence functional domain analysis we found that the sequence had a COX3 domain (Component of the cytochrome C oxidase) and during our 3D structure analysis we found that the structure had the most similarity with the (Bovine heart cytochrome c oxidase) structure (Figure 18). In our analysis it was found that this gene was expressed maximum in the Spleen in the Siberian sturgeon as regard to other immune tissues of the Siberian sturgeon or kidney and liver.

Total conserved domain- 60 AA

ASDQTDKPAHYHSEACSTIPSILRLLFLGSEFACLGSQSNREAPEPWESPSSAAMKSRALFFCLLL
ACFCLSLAQGSYENCCLSYVKSVKPSVKKAVRSYRTQETDGGCNIQAIVFKMRKGKVFCANPETSW
VQMLMKKLDAKNNKAENRKKSGMEPFQPLVSHARNEINHFQAFCISLLPVINMLAAMQAPQGILSF
FCPEKHHRASTLHPLSCYTATTKPKNQSTLDNPDCCHLTHCIQSHRSIKKNAYVHYVAIFYSFFSL
MKMYKQKLCVTIRANASNYIKIITGLHDIICLIFVKGNRLNMYMCILTSSALPLMIQDHEIEENDF
FPLVPKHALIEMYFFISFFIFSMAAILPESINWGVRGGHIVLYCTEMNIQLKRMMHSCSENHYHAI
FLYLFWIVCSRMSPCPCGVVFGRSGVHFHQKGSQWRPRIRQSKLMFPVCLRAMIVLLSLFFICLFV
LLQFFNILV-SHTHCECTL-S-DENTFYFL



3D structure of CCL25



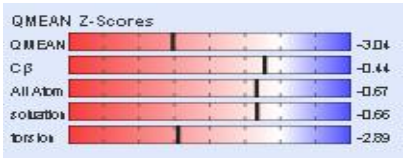
Q-Mean score of the structure

Figure 16. Functional domain and 3D structure of CCL25

Total conserved domain- 53 AA

LISDGSVLKTSRSQRLPLPRTTPSVDMAPSITALFLCSFLLYCCSKAAYGDSNLVVDCCLRVSKSPIPIKIVKNYKEQGQ
QDGCDLTAVVFTTKGNKKLCAPSNEPWVDSLRAKIDKRSQFCRNHKQKNRMCLNFARSMFICIHVEDGRSSKGFEVWQKL
SASWIQEQLKKQNCTMLFPTGFVCFMCLHGNRKSVAFYLKFSFLSLINKAASNVIFVYNMGLYFLKVFLYAIIQTSGSYD
HCFVKGPLKSSPVNLLVICSCIMLLPCLCFACTICCMSLMSVLKLCLQHEMSVALCIAWHEHPCNIVIIGLPIISALTYL
LCDNMEIALGTISCYLMHGNFVIIPMAHLYTSSHPLRPENTISYHVITMLLHSHVCPVIESATKILLAEDQFLFMSYRNA
MRSLKHVIQLPRSPGNSDHLDLEKQVLQRSELTSSRCSSFPKSPALLEEDRFAMLLHDPHYKRVLGNIYSGAPVRSLITFM
IKLYPKEITRIRLLFNGPVYSFIGMCSILVQKPFTNHILYKYSKEITVYRKLIQQPCTAHKEHSRRYSNSHSNADCSVH
HSVKGDRALSAPGVIYRTNVNTACLHCVCNSINSRGDWRIQIYKQTWEQDIFRNFMLKLYTQLYFSHIEQCTMNVFLCFC
EVVFVCLLSTLVFYFKRMSHLYQNAALSSLTFRMNAVLCIATSTYGEVPQLFLVDSHFIITATLTECQCYTIHFKAIENS
IAYSKIFLKFLFFNNVIVPMHFYTEPESFFFIYTFNKNIT

3D structure of MAFA



Q-Mean score of the structure

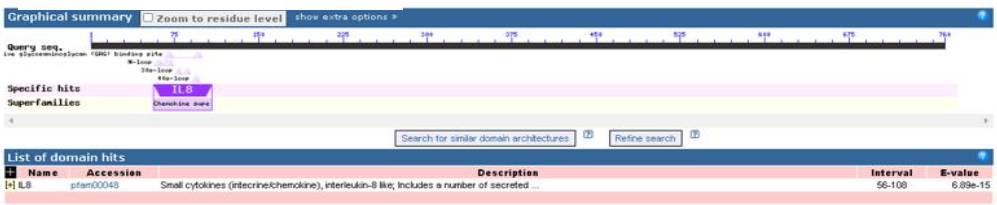
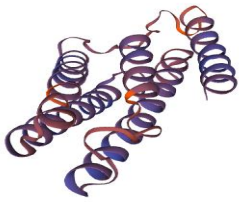


Figure 17. Functional domain of CCL19 and its 3D structure

Total conserved domain-169 AA

IILFITSEVFFFLGFFAFYHASLAPTPELGGCPPTGIITLDPFEVPLLNTAVLLASGVTVTAHHSIIERERKQTIQALALTILLGFYFTALQAI
EYYEAPFTIADGVYGSTFFVATGFHGLHVIIGSTFLAVCLLRQIQYHFTSEHHFGFEAAAYHFVDVLFLYVSIYX



Structure of the protein

Q-Mean score of the structure

Figure 18. Functional domain of MAFA like protein and 3D structure

SUMMARY

This is one of the first studies done on a 6-year-old matured female Siberian sturgeon on the presence of its immune marker genes and its possible role in the immune system of this species. I successfully identified 9 potential immune marker genes in *Acipenser baerii* Liver, kidney and spleen. These potential immune marker genes presence is confirmed by the TPM value associated with it and also the log difference between each tissue ≥5. From my study, I also got the confirmation that the sturgeons and teleost do share some common immune marker genes, for example-, I found that the sturgeon C3, CCL19 and CCL20 do have similarities with some teleost immune marker genes and thus confirms the presence of a shared immune marker gene. The found immune marker genes in this species were novel and thus no study earlier confirms the presence of these 9 potential immune marker genes in the Siberian sturgeon. This data could indeed help us in the understanding potential role of these immune marker genes in the sturgeon immune system and future studies could shed a light on how these genes are helpful in sturgeon fight against pathogens. Overall we found 9 most potential immune marker genes in the Siberian sturgeon which were not been studied before and neither found.

REFERENCE

Babaei, S., Abedian-Kenari, A., Hedayati, M. & Yazdani-Sadati. (2011).M. A. Growth response, body composition, plasma metabolites, digestive and antioxidant enzymes activities of Siberian sturgeon (*Acipenser baerii*, Brandt) fed different dietary protein and carbohydrate: Lipid ratio. Aquaculture. Research. 48:2642–2654.

Ben-Ari Fuchs S, Lieder I, Stelzer G, Mazor Y, Buzhor E, Kaplan S, Bogoch Y, Plaschkes I, Shitrit A, Rappaport N, Kohn A, Edgar R, Shenhav L, Safran M, Lancet D, Guan-Golan Y, Warshawsky D, and Strichman R.(2016). GeneAnalytics: An Integrative Gene Set Analysis Tool for Next Generation Sequencing, RNAseq and Microarray Data. OMICS. 20:39-51.

Belinky F, Nativ N, Stelzer G, Zimmerman S, Iny Stein T, Safran M, and Lancet, D. (2015).Path Cards: multi-source consolidation of human biological pathways. Oxford.15:6-13

Chang Xu, Erchao Li, Yantong Suo, Yujie Su, Minghui Lu, Qun Zhao, Jian G. Qin, Liqiao Chen. (2018). Histological and transcriptomic responses of two immune organs, the spleen and head kidney, in Nile tilapia (*Oreochromis niloticus*) to long-term hypersaline stress. Fish & Shellfish Immunology.76:48-57.

Fishilevich S, Nudel R, Rappaport N, Hadar R, Plaschkes I, Iny Stein T, Rosen N, Kohn A, Twik M, Safran M, Lancet D, and Cohen D. (2017).GeneHancer: genome-wide integration of enhancers and target genes in Gene Cards. Oxford. 28:23-35

Frank chapman and Joel van enennaam. (2008). Sturgeon aquaculture specialized techniques: Determining the stage of sexual maturity in female sturgeon for artificial spawning: the egg polarization index or pi, Electronic data information source.13:153-157

Guan H, Wang Y, Li X, Xiang A, Guo F, Fan J and Yu Q. (2022).C1q/Tumor Necrosis factor-related Protein 9: Basics and Therapeutic Potentials. Frontiers in physiology.13:18-34.

Höhne, C., Prokopov, D., Kuhl, H., Du, K., Klopp, C., Wuertz, S., Trifonov, V. and Stöck, M. (2021).The immune system of sturgeons and paddlefish (Acipenseriformes): a review with new data from a chromosome-scale sturgeon genome. Reviews in Aquaculture.13:1709-1729.

Jaime Huerta-Cepas, Damian Szklarczyk, Davide Heller, Ana Hernández-Plaza, Sofia K Forslund, Helen Cook, Daniel R Mende, Ivica Letunic, Thomas Rattei, Lars J Jensen, Christian von Mering, Peer Bork.(2019). Nucleic Acids Research. 47:309-314.

Kim, D., Nam, Y., Noh, J. et al. (2005). Karyotype of North American shortnose sturgeon *Acipenser brevirostrum* with the highest chromosome number in the Acipenseriformes. Ichthyological Research. 52:94–97.

Liu M, Yang X, Zeng C, Zhao H, Li J, Hou Z, Wen H. (2022). Transcriptional Signatures of Immune, Neural, and Endocrine Functions in the Brain and Kidney of Rainbow Trout (*Oncorhynchus mykiss*) in Response to *Aeromonas salmonicida* Infection. International Journal Molecular Science, 23:1340-1347.

M.Long, J.Zhao, T.li, C.tafalla,Q.Zhang, X.Wang, et al.(2015).Transcriptomic and proteomic analysis of the splenic immune mechanism of rainbow trout (*Oncorhynchus mykiss*) infected by *Aeromonas salmonids subsp*, Salmonidae. Journal of Proteomics.122:41-55.

Rappaport N, Twik M, Plaschkes I, Nudel R, Iny Stein T, Levitt J, Gershoni M, Morrey CP, Safran M, and Lancet D(2016) MalaCards: an amalgamated human disease compendium with diverse clinical and genetic annotation and structured search. Nucleic Acids Research.12:34-42

Safran M, Rosen N, Twik M, BarShir R, Iny Stein T, Dahary D, Fishilevich S, and Lancet D. (2022).The GeneCards Suite Chapter.Practical Guide to Life Science Databases.14:27-56.

Shun Maekawa, Pei-Chi Wang, Shih-Chu Chen. (2019). Differential expression of immune-related genes in head kidney and spleen of cobia (*Rachycentron*

canadum) having *Streptococcus dysgalactiae* infection. Fish & Shellfish Immunology.92:842-850.

Stelzer G, Plaschkes I, Oz - Levi D, Alkelai A, Olender T, Zimmerman S, Twik M., Belinky F, Fishilevich S, Nudel R, Guan - Golan Y, Warshawsky D, Dahary D, Kohn A, Mazor Y, Kaplan S, Iny Stein T, Baris H, Rappaport N, Safran M, and Lancet D.(2016).VarElect: The phenotype-based variation prioritizer of the Gene Cards suite. BMC Genomics.17:444-448.

V.J.Birstein, R.hanner, R.DeSalle.(1997).Phylogeny of the Acipenseriformes: cytogenetic and molecular approaches. Environmental Biology Fish.48:127-155.

Wang, S., Furmanek, T., Kryvi, H. (2014). Transcriptome sequencing of Atlantic salmon (Salmo salar) notochord before the development of the vertebrae provide clues to the regulation of positional fate, chord blast lineage and mineralization. BMC Genomics.15**:**141-147.

X.Qian, Y.Ba, Q.Zhuang, G.Zhong. (2014).RNA seq. technology and its application in fish Transcriptomics.OMICS.18:98-110.

Y.Mu, M.Li, F.Ding, Y.Ding, J.Ao, S.hu.(2014).De novo characterization of the spleen transcriptome of the large yellow croaker(*Pseudosciana crocea*) and analysis of the immune relevant genes and pathways involved in the antiviral response. Plos one, 9:21-32

Youlu Su, Zhixun Guo, Liwen Xu, Jingzhe Jiang, Jiangyong Wang, Juan Feng .(2012). Identification of a cobia (*Rachycentron canadum*) CC chemokine gene and its involvement in the inflammatory response, Fish & Shellfish Immunology, 32:204-210.

Yu-ping, H., Di, W. (2005). A review of sturgeon virosis. Journal of Forestry Research.16:79–82.