De novo Assembly and Analysis of Amur Sturgeon (*Acipenser schrenckii*) Transcriptome in Response to *Mycobacterium Marinum* Infection to Identify Putative Genes Involved in Immunity

Qianqian Zhang¹, Xiehao Wang^{1,5}, Defeng Zhang², Meng Long^{1,3,4}, Zhenbing Wu^{1,5}, Yuqing Feng^{1,5}, Jingwen Hao^{1,5}, Shuyi Wang^{1,5}, Qian Liao^{1,5}, and Aihua Li^{1*}

¹Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, P.R. China

²Pearl River Fisheries Research Institute, Chinese Academy of Fishery Sciences, Guangzhou, P.R. China

³Shenzhen Institute of Guangdong Ocean University, Shenzhen, 518120, P.R. China

⁴Shenzhen Dapeng New District Science and Technology Innovation Service Center, Shenzhen, 518120, P.R. China

⁵University of Chinese Academy of Sciences, Beijing 100049, P.R. China

Received: March 15, 2019

Revised: July 10, 2019

Accepted: July 31, 2019

First published online

August 1, 2019

*Corresponding author

Phone: +86-27-68780053;

Fax: +86-27-68780053;

E-mail: liaihua@ihb.ac.cn

pISSN 1017-7825, eISSN 1738-8872

Copyright© 2019 by
The Korean Society for Microbiology
and Biotechnology

Fish mycobacteriosis is a common bacterial disease in many species of freshwater and marine fish and has caused severe loss of fish production. *Mycobacterium marinum* has been the most prevalent pathogen observed in several outbreaks of mycobacteriosis of farmed sturgeons in China. However, the immune responses and pathology of sturgeons in mycobacterial infection are rarely studied. Therefore, we used the Illumina RNA-seq method to analyze the transcriptome profile of *Acipenser schrenckii* challenged with *Mycobacterium marinum*. To begin, 168,220 non-redundant contigs were acquired from the infection and control groups, and among these, 33,225 contigs have acquired annotations. A total of 4,043 differently expressed (DE) contigs between the two groups were identified, and among these, 2479 were up-regulated and 1564 were down-regulated in the infected fish. A total of 1,340 DE contigs with acquired annotations in KEGG were enriched for 124 pathways including the TNF signaling pathway, and the Toll-like receptor signaling pathway. The roles of DE genes involved in significant pathways and other processes were discussed. The 2,209 DE contigs that have yet to acquire proper annotation may represent candidate genes associated with infection in sturgeons and are expected to serve as immunogenetic resources for further study. To our best knowledge, this is the first transcriptome study on sturgeons under bacterial infection.

Keywords: Amur sturgeon, transcriptome, *Mycobacterium marinum*, infection, immune mechanism

Introduction

Mycobacterium marinum is a Gram-positive, aerobic, acid-fast pathogenic bacteria first identified in 1926 [1] which forms the largest proportion of all mycobacteria isolated from fish, and is a close genetic relative of *M. tuberculosis*. It naturally infects fish in which it produces a systemic tuberculosis-like disease. The external symptoms of fish in advancing infection include scale loss, nodules, ulcers, and

fin necrosis, and the typical histological exhibition is the granulomas located in multiple organs. Basically, granulomas are organized aggregates of mature macrophages with an increase in cytoplasmic organelles and ruffled cell membranes [2]. However, the advanced structure of granulomas differs across hosts. The severity of the disease is also dependent on both the bacterial dose and strain [3]. When acute and fulminant infection occurs, fish can die within 10 days with extensive symptoms and high bacterial

burdens, while during latent infection, the majority of bacteria are in a dormant state with only a small bacterial population remaining active, and the dormancy is mediated by the adaptive immune responses of the host (such as zebrafish) [4]. Therefore, the characteristics of mycobacterial infection indicate the complex interplay between mycobacteria and hosts.

Sturgeons are ancient, freshwater, chondrocyte species and are called living fossils for their evolutionary characteristics placing them between Chondrichthyes and Osteichthyes. However, the current knowledge about the profile of the immune defense mechanism of sturgeons is still deficient. *Acipenser schrenckii* (Amur sturgeon), belonging to *Acipenseriformes*, is one of the most primitive *Actinopterygii* species. It is originally distributed in the Amur river in China and Russia and has become one of the major farmed sturgeons in China. Nowadays, diseases have increasingly threatened the sturgeon culture industry, and mycobacteriosis was identified to be involved in causing serious mortality, and *M. marinum* was the only species identified that can cause mycobacteriosis on sturgeons alone [5]. Therefore, it is of great importance to study the pathology and immune response of sturgeons to *M. marinum* infection.

RNA-Seq is a powerful tool used to acquire the transcriptome profile, especially that of the immune system of fish infected by bacterial pathogens. It provides deep insight with its unbiased nature, high accuracy, and ability to obtain novel genes. Utilizing the Trinity algorithm, we recovered transcripts of *Acipenser schrenckii* with high sensitivity by *de novo* assembly.

Currently, the transcriptomic studies on *M. marinum* infection of fish are restricted to zebrafish, however, sturgeons possess special evolutionary status and the development of its immune system should be less matured than teleost fish. This phenomenon can be proved by the fact that sturgeons are more susceptible to mycobacterial infection than other fishes reared in the same water body. So, the results obtained from zebrafish studies cannot fully explain what happens on sturgeons. To our best knowledge, this is the first transcriptomic study on sturgeons in bacterial infection. The aim of this study was to discover and determine the pathology and immune responses, including significantly changed genes and signal pathways of *Acipenser schrenckii* involved in *M. marinum* infection. The dataset provided is expected to represent a valid resource for molecular studies of both the immunity of sturgeons and mycobacterial pathogenesis.

Materials and Methods

Fish and Infection

The *Acipenser schrenckii* (weight, 26 to 30 g) were obtained from a fish farm in Hubei Province, China and acclimated for 2 months at a water temperature of 25°C before the bacterial challenge in the facilities of the Institute of Hydrobiology, Chinese Academy of Sciences. The Amur sturgeons were divided into the infection group and control group with 9 fish in each. The infection group was then intraperitoneally challenged with 10^7 CFU of *M. marinum* strain ASCy-1.0[5] per fish. The bacterial growth and preparation for infection were done using the standard method. In brief, *M. marinum* was grown in Middlebrook 7H9 broth (Difco) supplemented with 0.4% glycerol, 0.05% Tween 80, and 10% ADC (albumin–dextrose–catalase, BD Biosciences). After being incubated to the mid-logarithmic stage at 28°C, bacteria were collected by centrifugation and resuspended to a final concentration of 5×10^7 CFU/ml in sterile phosphate-buffered saline (PBS). The control group was injected with 0.2 ml sterile PBS.

Sampling and RNA Extraction

The fish were sacrificed at 22 dpi when they showed clear signs of mycobacteriosis and death occurred. Four fish from the infection group and control group each were anesthetized with tricaine methanesulfonate (MS-222, Sigma, USA), and histological examination showed that the liver was the most heavily infected organ. Therefore, we collected livers aseptically, and immediately submerged the liver samples with dimensions less than 0.5 cm into RNA Safer Reagent RNA Stabilizer (Omega-Biotek). The samples were kept in dry ice until proceeding extraction. Total RNA was extracted from the tissues with TRIzol Reagent (Invitrogen, USA) according to the manufacturer's instructions. The RNA integrity value (RIN) was checked. The RNA samples were treated with DNase I under 37°C for 1 h to remove DNA residue, and mRNAs were purified using Micropoly (A) Purist mRNA Purification Kit (Ambion, USA) according to the manufacturer's instructions. The concentration of mRNA was determined by NanoDrop (Thermo Scientific, USA). The RNA samples from each group were pooled before library construction.

Sequencing

Double-stranded cDNA was synthesized following the previous protocol [6]. In brief, first-strand cDNA synthesis was primed by a GsuI-oligodT primer, 10 mg of mRNA, and 1000 units of Superscript II reverse transcriptase (Invitrogen) with incubation at 42°C for 1 h. Thereafter, the mRNA's CAP structure was oxidized by NaIO_4 (Sigma) and ligated to biotin hydrazide, which was used to link complete mRNA/cDNA hybrids to Dynabeads M-280 (Invitrogen) for purification. The mRNAs in hybrids were removed by alkaline treatment. The 5'-end specific adaptors were added to synthesize the second strand of cDNA. Then the poly(A) tails were removed by GsuI digestion. cDNAs were then

fragmented into the length of 300~500 bp, and after purification with Ampure beads (Agencourt, USA), sequencing adaptors were ligated to the fragments. The cDNA libraries were constructed with Truseq DNA Sample Prep Kit (Illumina, USA), and a subsequent sequencing run was performed on Illumina platform.

Sequence Filtering and *de novo* Assembly

Sequences were prepared for assembly by removal of library adapter sequences and quality filtering using FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). The clean reads obtained after filtering with a length of less than 40 bp were discarded. All high-quality reads were subjected to *de novo* assembly conducted by Trinity (<https://github.com/trinityrnaseq/trinityrnaseq/wiki>). As a result, non-redundant contigs were produced for protein prediction.

Functional Annotation

ORFs were acquired with the "GetORF" function of the EMBOSS package (<http://emboss.sourceforge.net/>) and were translated into predicted protein sequences. The protein sequences were submitted to search against the NCBI Non-redundant databases (nr), UniProt databases, Gene Ontology (GO), and the Kyoto Encyclopedia of Genes and Genomes (KEGG) by BLASTP with thresholds of E-value < $1e^{-5}$.

GO annotations for protein sequences were acquired using GoPipe package using a BLASTP search against the Swiss-Prot and TrEMBL databases with an E-value < $1e^{-5}$. KEGG pathways mapping was conducted by KEGG Automatic Annotation Server (KAAS) with the single-directional best-hit method and seven fish species were selected as reference organisms. KEGG annotation assigns KEGG Orthology (KO) identifiers that represent specific orthologous groups of genes to each submitted matched sequence, and then the identifiers linked to query sequences were mapped to objects in KO maps.

Identification of Differentially Expressed Genes

To identify genes in the Amur sturgeons that are differentially expressed in response to *M. marinum* infection, the numbers of trimmed reads mapped back to every contig from the infection group and control group were counted. Thereafter, by using MARS (MA-plot-based method with Random Sampling model) in DEGseq package (<http://www.bioconductor.org/packages/release/bioc/html/DEGseq.html>), the difference in expression of genes between the two groups was calculated and the differentially expressed genes were figured out. Contigs were screened using the criteria $q \text{ value} \leq 0.001$ and $|\log_2 \text{Ratio}| \geq 1$ to identify the DEGs (differentially expressed genes).

DEG Analysis

GO categorization of obtained DEGs according to GO annotations were conducted utilizing Blast2go software (<https://www.blast2go.com/>) using a BLASTP search against the Non-redundant database with an E-value < $1e^{-5}$. The enriched KEGG signaling pathways

were identified by the "KEGG profile" package (<http://www.bioconductor.org/packages/release/bioc/html/KEGGprofile.html>) based on hypergeometric tests with an adjusted p-value < 0.01, and the alteration of the expression of DEGs was marked in KEGG map using the Pathview package (<http://www.bioconductor.org/packages/release/bioc/html/pathview.html>).

qRT-PCR Confirmation

We selected seven differentially expressed contigs (elastase-like serine protease, ceruloplasmin, transferrin, peroxiredoxin-5 mitochondrial precursor, phosphorylase kinase gamma subunit 1, phenylalanyl-tRNA synthetase beta chain, asparagine synthetase) for quantitative RT-PCR analysis to confirm the RNA-seq results, with 18S rRNA used as the endogenous reference gene. The primers were designed using the GenScript online tool (<https://www.genscript.com/ssl-bin/app/primer>) based on the contig sequences we assembled. The first-strand cDNA was synthesized using PrimeScript RT Reagent Kit (Takara) with gDNA eraser according to the manufacturer's instructions. The qPCR was performed with SYBR Premix Ex Taq II (Takara) in a CFX96 Real-time PCR Cycler (Bio-Rad). The total reaction volume of 20 μ l contained 2 μ l of cDNA template, 10 μ l of SYBR Premix Ex Taq II, 0.8 μ l of each primer (10 μ M), and 6.4 μ l of dH₂O. The thermal cycling protocol was as follows: denaturation at 95°C for 30 sec, followed by 40 cycles of 95°C for 5 sec, 60°C for 30 sec, and finally melt curve was performed to examine the specificity. All reactions were performed in triplicate. The relative expression of the target genes were calculated using the Ct method (expressed as $2^{-\Delta\Delta C_t}$) [7].

Results and Discussion

Transcriptome Sequencing and Assembly

Two non-normalized cDNA libraries were constructed, and 53.77 and 52.84 million paired-end raw reads were generated by Illumina RNA-seq sequencing platform from the liver tissues of fish in infection group and control group respectively. All reads were deposited on the Short Read Archive (SRA) of the National Center for Biotechnology Information (NCBI) and can be accessed under the accession number of SRR2912926 for infection group and SRR2912929 for the control group. After filtering to remove low-quality reads, a total of 50.72 (94.3%) and 48.92 (92.5%) million clean reads with an average length of 100 bp were obtained from these two libraries. Using Trinity, clean reads from the infection group and control group were assembled into 98,056 and 117,986 contigs with a mean length of 648 bp and 613 bp and N₅₀ values of 1,026 bp and 905 bp respectively (Table 1). Subsequently, contigs from two libraries were merged for digital gene expression analysis. Removal of partial overlapping sequences generated

Table 1. Summary of the transcriptome sequencing of the control and infection groups.

Terms	Control	<i>M.marinum</i> infected
Number of raw reads	52839360	53769228
Number of clean reads	48925334	50725686
Ratio	92.5%	94.3%
Average length of reads (bp)	100	100
Number of contigs	117986	98056
Average length of contigs (bp)	612	648
Length range of contigs	201~11939	201~15355
N50 of contigs	905	1026

168,220 non-redundant contigs. It must be noted that the study used an IP infection route rather than a natural infection route or a naturally infected fish, and due to this methodological limit, the results or conclusions of this study may be somewhat biased from the actual situation.

Annotation

To annotate the assembled contigs, the predicted protein sequences of 168,220 contigs were used as queries to search against public protein databases using BLASTP with a threshold of sequence similarity $\geq 30\%$ and E-value $\leq 1e^{-5}$. In total, 33,225 (19.75%) contigs had significant hits in

NCBI non-redundant databases or GO, KEGG and UniProt databases (Fig. 1). The number of annotated sequences is similar to that of *A. sinensis* in a previous study [8], in which 36,157 unigenes were annotated. The unannotated contigs probably represent the specific genes of *Acipenser schrenckii*, which were too dissimilar to be annotated by an orthologue in other reference species, in addition to the non-protein coding genes and UTRs.

The contigs were further assigned Gene Ontology (GO) terms using the GoPipe package. A total of 23,691 predicted proteins were assigned 190,111 GO annotations in three ontologies according to the GO database: “biological process,” “cellular component,” and “molecular function” (Fig. 2).

The analysis based on KEGG maps plays a key role in comprehending the immune response of *Acipenser schrenckii*

Fig. 1. Venn diagram of annotations from four different databases.

Number of contigs annotated by four different databases: GenBank non-redundant (NR, total annotated = 32,338), KEGG (total annotated = 29,939), UniProt (total annotated = 27,329), and GO (total annotated = 23775).

Fig. 2. Assembled contigs were classified into GO sub-categories under the biological process, cellular component, and molecular function ontologies.

Table 2. The immune-related KEGG signaling pathways mapped onto contigs assembled in this study.

Pathway	Number of contigs
Hematopoietic cell lineage	102
Complement and coagulation cascades	128
Platelet activation	334
Toll-like receptor signaling pathway	175
NOD-like receptor signaling pathway	164
RIG-I-like receptor signaling pathway	140
Cytosolic DNA-sensing pathway	97
Natural killer cell mediated cytotoxicity	216
Antigen processing and presentation	113
T cell receptor signaling pathway	246
B cell receptor signaling pathway	208
Fc epsilon RI signaling pathway	144
Fc gamma R-mediated phagocytosis	235
Leukocyte transendothelial migration	252
Intestinal immune network for IgA production	46
Chemokine signaling pathway	320

to mycobacterial infection. The contigs were mapped onto KEGG signal pathway through KEGG Automatic Annotation Server, the seven fish species (*Danio rerio*, *Takifugu rubripes*, *Maylandia zebra*, *Oryzias latipes*, *Xiphophorus maculatus*, *Latimeria chalumnae*, and *Callorhinchus milii*) which are manually annotated organisms in KEGG database were selected as a reference. A total of 19,624 contigs were assigned 6476 KO (KEGG Orthology) identifiers and were mapped onto 332 pathways involving 6 principal categories. The most predominant KEGG category is "Human Diseases," containing 12,247 (62.4%) contigs. Among the sub-categories, "Signal transduction" contained the greatest number of contigs (7020 contigs, 35.8%), followed by "Cancers: Overview" (3074 contigs, 15.7%) and "Immune system" (2920 contigs, 14.9%). Moreover, the three most abundant pathways were "Pathways in cancer" (918 contigs, ko05200), "PI3K-Akt signaling pathway" (684 contigs, ko04151), "HTLV-I infection" (538 contigs, ko05166). Also, 16 immune-related pathways were found mapped from contigs (Table 2). These data suggest that the infection triggered a vast variety of immune responses, and provide a valuable basis for further study on the specific process.

Identification of Differentially Expressed Genes

Based on M-A method with a random sampling model described in the Materials and Methods section, a total of 4043 DEGs were identified from the comparisons of the

Fig. 3. Comparison of the expression profiles of selected genes determined by RNA-seq and qRT-PCR.

Target gene abbreviations are as follows: ESP: elastase-like serine protease, CP: Ceruloplasmin, Tfn: transferrin, PRDX5: Peroxiredoxin-5 mitochondrial precursor, PHKG1: phosphorylase kinase gamma subunit 1, FARSF: Phenylalanyl-tRNA synthetase beta chain, AS: asparagine synthetase.

infection group and the control group. A total of 2479 DEGs were up-regulated (61.3%) in the infection group, while 1564 DEGs were down-regulated (38.7%). Moreover, 2233 (55.2%) differentially expressed contigs did not acquire annotation in selected bases and this is likely due to the absence of clear information on sturgeons. The function of these unidentified genes remains to be clarified in future studies.

Validation of gene expression profile by quantitative-PCR was performed, and we randomly selected 7 DEGs and designed confirmed primers based on the assembled contigs sequence. The results showed the direction of expression alteration determined by two independent methods were the same, and the extension was close (Fig. 3), providing evidence that the *de novo* assembly and read counts in this study were reliable.

Pathway Enrichment of Differentially Expressed Genes

A total of 1,340 differentially expressed contigs acquired 929 unique KO identifiers, which mapped onto 289 pathways and were significantly enriched for 124 pathways among these as determined by hypergeometric tests (p value < 0.01) using the KEGG profile package. Some statistically significant pathways related to immune system or current knowledge of mycobacterial infection are shown in Table 3. Among them, the TNF signaling pathway has the highest degree of enrichment (31 DEGs, such as PIK3C, PIK3R, MMP9, CASP3, NFKB1), followed by Toll-like receptor signaling pathway (22 DEGs, such as RIPK1,

Table 3. Some enriched KEGG signaling pathways involved in immunity and tuberculosis.

Major enriched KEGG pathways	Number of DEGs	Number of genes	Percentage of DEGs	Adjusted P value
TNF signaling pathway	31	87	36%	7.25E-16
Toll-like receptor signaling pathway	22	76	29%	1.39E-11
Tuberculosis	27	131	21%	6.66E-10
NOD-like receptor signaling pathway	14	51	27%	8.10E-08
NF-kappa B signaling pathway	18	86	21%	2.31E-07
Complement and coagulation cascades	14	69	20%	5.44E-06
RIG-I-like receptor signaling pathway	12	53	23%	5.68E-06
Natural killer cell mediated cytotoxicity	14	79	18%	3.02E-05
Phagosome	15	95	16%	7.53E-05
Chemokine signaling pathway	19	144	13%	1.57E-04
B cell receptor signaling pathway	10	54	19%	1.82E-04
T cell receptor signaling pathway	13	84	15%	2.37E-04
Fc gamma R-mediated phagocytosis	10	56	18%	2.56E-04
Cytosolic DNA-sensing pathway	9	51	18%	4.94E-04
Platelet activation	12	85	14%	9.12E-04
Antigen processing and presentation	7	41	17%	1.86E-03
Fc epsilon RI signaling pathway	7	44	16%	2.96E-03

Note: The significance of each pathway was estimated by hypergeometric tests (adjusted p-value for multiple tests <0.01).

CD40, TRAF3, FOS, RAC1), tuberculosis (27 DEGs, such as ATP6D, C3, dnaK, CASP8, P38), NOD-like receptor signaling pathway (14 DEGs, such as CASP1, HSP90A, TAB2, IL1B, RIPK2), and NF-kappa B signaling pathway (18 DEGs, such as TRAF2, TRAF3, MYD88, IRAK4, CCL19). There were also some pathways involved in other diseases found perturbed, including influenza A, Epstein-Barr virus infection, hepatitis B, herpes simplex, pertussis, salmonella infection and amoebiasis. Many of these are caused by intracellular pathogens, indicating the conservation of host defense mechanisms against the intracellular survival of these pathogens and *M. marinum*. Pathways involved in some other biological processes that may indirectly participate in immune response were found enriched too, such as protein folding, sorting and degradation (the pathway of protein processing in endoplasmic reticulum, proteasome, and protein export), cell growth and death (apoptosis, p53 signaling pathway), cellular community (focal adhesion, tight junction), lipid metabolism (biosynthesis of unsaturated fatty acids, glycerolipid metabolism, arachidonic acid metabolism), cell motility (regulation of actin cytoskeleton), and environmental adaptation (circadian entrainment).

Analysis of the Canonical Tuberculosis Pathway

Ninety-six genes in our transcriptome data were found mapped onto the tuberculosis-related pathway, and among

them, 27 genes were differentially expressed, indicating some extent of conservation of immune responses to mycobacteria infection between mammals and sturgeons (Fig. 4).

Analysis of TNF Signaling Pathway

Tumor necrosis factor (TNF) signaling pathway appeared to be the pathway with the highest degree of enrichment. Although there was no homolog of TNF molecule found in the transcriptome of *A. schrenckii*, which may be attributable to the species specificity, our data revealed 71 genes involved in TNF cascades based on KEGG and 31 of these were differentially expressed, indicating the significance of TNF related pathway (Fig. 5).

TNF is a cytokine produced primarily by macrophages in response to stimuli that activate Toll-like receptors, as well as by activated T cells, B cells, and NK cells. It is well established that TNF is critical for protection against mycobacteria and plays a central role in granuloma formation [9, 10]. On the one side, TNF can mediate apoptosis, which is associated with diminished pathogen viability, on the other side, TNF excess induces necroptosis, resulting in susceptibility to mycobacteria [11]. These two aspects of cell death, in this case, will be discussed in the corresponding section.

TNF receptor-associated factor 2/5 (TRAF2/5) and

Fig. 4. Tuberculosis signaling pathway map generated by KEGG. Red indicates significantly up-regulated genes, while green indicates significantly down-regulated genes.

receptor interacting protein 1(RIP1), two components in an active TNF signaling complex, in company with a TNF receptor superfamily member 11B precursor, were significantly elevated, which is indicative of the activation of TNF signaling. Considering that virulent mycobacteria adopt a variety of mechanisms to inhibit TNF-mediated host cell apoptosis for their survival, such as activating the release of membrane-bound TNFR2 as a soluble form, and increasing the expression of anti-apoptotic proteins [12], notably, the up-regulation of the silencer of death domains

(SODD) might be a hint of another mechanism of inhibition of apoptosis by *M. marinum*, since SODD binds to the death domain in TNFR1 and prevents DD-interacting activator proteins from taking hold, dampening TNFR-1 stimulated signaling [13].

The p38 group of MAP kinases plays a vital role in numerous biological processes. A study on cell infection reported that p38 was activated in macrophages by both pathogenic and non-pathogenic mycobacteria in the early stage of infection, but its phosphorylation was lost later

Fig. 5. TNF signaling pathway map generated by KEGG.

Red indicates significantly up-regulated genes, while green indicates significantly down-regulated genes.

only in infection of the pathogenic *Mycobacterium avium*, suggesting that p38 may be a point of immune intervention by pathogenic mycobacteria [14]. However, as the p38 was up-regulated in this study on individual and sustaining infection, the possible inhibition effect of *M. marinum* failed to impact the p38 consistently.

MSK1 [16] is regulated by ERK and p38 MAPK activation. It has been demonstrated that MSK1 is necessary for regulating the pathogen-driven expression of cyclooxygenase-2 acting in immunosuppression by controlling transcription factors NFkB and CREB [15]. The down-regulation here may suggest the defense of host balancing the induction by mycobacteria.

Notably, we observed the induction of many transcription factors related to TNF pathway, including ubiquitous transcription factor NFkB, CREB (cAMP response-element-binding protein), C/EBPβ (CCAAT/enhancer binding protein beta), and the components of activator protein 1,

including C-FOS, C-JUN, and JUN-B. Triggering anti-inflammatory response is a key survival mechanism adopted by mycobacteria for chronic infection [16]. It was reported that mycobacteria can switch transcriptional activity from pro-inflammatory NFkB to anti-inflammatory CREB to promote the transition from acute to chronic inflammation by phosphorylation of glycogen synthase kinase 3 (GSK3) through p38, the elevated expression of p38 and a predicted GSK3-beta interaction protein-like is a relevant observation. C/EBPβ increases a wide variety of genes, especially those associated with inflammatory response, including TNF-α, IL-1β, and C-FOS. The up-regulation of the transcript of c-Fos and c-Jun corroborates a previous report in which the protein level of c-Fos and c-Jun increased in an epithelial cell after *M. bovis* BCG infection, in addition, C-JUN and CREB can induce IFN-γ transcription by human T cells in response to mycobacterial antigen [17]. The expression elevation of these transcription

factors suggests complex inflammatory response occurring during mycobacteriosis on sturgeons, as reported for zebrafish [18].

Changes in the Connection between Metabolism and Immunity

Some differentially expressed genes act between

metabolism and immunity of the host. The expression of a precursor of leptin had strongly increased in the infection group. Leptin is produced mainly by the adipocytes and is a pleiotropic hormone proposed to link nutritional status to the development of strong Th1 immunity, favoring the Th1 response while inhibiting the secretion of Th2 cytokines. Our finding, in accordance with previous work on mouse-

Table 4. Significantly immune-related differentially expressed genes ($\log_2(\text{FC}) > 2$) in *Acipenser schrenckii*.

Category and Contig ID	NCBI Accession	Nr Annotation	$\log_2(\text{FC})$	q-value
Innate immune response				
comp128207_c0_seq1	ref NP_001005599.1	Serum amyloid A	5.91	0
comp139954_c1_seq4	gb ACT80308.1	Inducible nitric oxide synthase 2	7.65	8.55E-22
comp137209_c2_seq1	emb CAG32523.1	Chain A, Hsp90 Alpha Atpase Domain With Bound Small Molecule Fragment	4.30	0
comp135208_c0_seq2	gb AIZ71528.1	Toll-like receptor 05L, partial	3.65	2.71E-32
comp137399_c0_seq1	ref XP_006625715.1	Toll-like receptor 13-like	2.85	4.96E-38
comp137209_c1_seq3	ref NP_001167173.1	Heat shock protein 90	2.50	0
Dendritic cell differentiation				
comp138818_c0_seq1	ref XP_003216061.1	PREDICTED: basic leucine zipper transcriptional factor ATF-like 3	3.20	1.42E-06
Antigen processing and presentation				
comp125825_c0_seq1	gb ACF21830.1	MHC class I antigen	2.95	2.03E-06
Leukocyte migration involved in inflammatory response				
comp129470_c2_seq3	gb AAO25640.1	Interleukin-8	2.56	4.75E-15
comp137409_c2_seq2	ref XP_003209351.1	PREDICTED: dual oxidase	2.14	1.02E-3
Cell chemotaxis				
comp130151_c0_seq1	gb ACI67502.1	C-C motif chemokine 19 precursor	5.22	0
comp133479_c0_seq1	dbj BAB88677.1	CXC chemokine	2.87	3.91E-35
comp129470_c2_seq3	gb AAO25640.1	PREDICTED: permeability factor 2-like	2.56	4.75E-15
comp134249_c0_seq1	gb AEM98349.1	C-X-C motif chemokine 10	2.52	1.21E-29
Positive regulation of T cell activation				
comp129937_c0_seq2	ref NP_001229653.1	PREDICTED: indoleamine 2,3-dioxygenase 1	4.38	4.77E-23
comp133479_c0_seq1	dbj BAB88677.1	PREDICTED: C-X-C motif chemokine 10	2.87	3.91E-35
T cell proliferation				
comp134118_c0_seq4	ref NP_001117054.1	Interleukin-1 beta	3.65	4.53E-05
Negative regulation of B cell activation				
comp138227_c0_seq2	ref XP_692922.3	PREDICTED: tumor necrosis factor alpha-induced protein 3	2.87	5.82E-06
Positive regulation of B cell proliferation				
comp131907_c0_seq4	emb CBN82130.1	PREDICTED: ADP-ribosyl cyclase 1 isoform X1	5.86	1.04E-13
IFN-related genes				
comp124662_c3_seq1	gb ADE73873.1	Interferon-inducible 58 kDa protein	3.95	2.07E-3
comp119974_c1_seq1	ref XP_543917.3	PREDICTED: interferon-induced protein with tetratricopeptide repeats 5	3.65	1.02E-2
comp127272_c1_seq3	ref XP_003460004.1	PREDICTED: interferon-inducible GTPase 5-like	3.54	2.60E-38

Mtb model demonstrating that the deficiency of leptin or leptin receptors causes the ineffectiveness of IFN γ and disordered granulomas [19, 20], suggests leptin is an essential regulator in mycobacterial infection.

We also observed serious down-regulation of two contigs linked to a neutrophil elastase-like protein, and of contigs linked to chymotrypsin-like elastase 1-like, 2a, 3b, as well as a pancreatic elastase precursor, and an elastase-like serine protease respectively. The chymotrypsin-like elastase family members 1, 2a, 3b had been also identified in a lesser spotted catshark transcriptome [21]. The depression reflects that the function of host elastase was seriously inhibited in the infection. The neutrophil elastase can be inhibited by the acute-phase protein α 1-antitrypsin (A1AT), which binds almost irreversibly to the active site of elastase. Correspondingly, A1AT significantly increased by 16% in the infection group in this study with an abundant expression, which may account for the elastase depression. The mycobactericidal effects of neutrophil elastase have been revealed [22], besides, liposomal encapsulated neutrophil elastase exhibits therapeutic potential against pulmonary mycobacterial infections [22], but the connection between mycobacteria and other members of elastase are still unknown. These members were commonly regarded as digestive enzymes, on this regard, the depression of them reflects a loss of the intestinal function, in accordance with the findings in other infection studies, such as on turbot suffering enteromycosis [23].

Notably, the annotation of facilitated glucose transporter member 11-like (GLUT-11) was assigned to four strongly up-regulated contigs. The ESAT-6 of *M. tuberculosis* has been reported to induce GLUT-1-mediated enhanced glucose uptake by human macrophages in the differentiation of foamy macrophages, which accumulate and provide nutrient for the intracellular bacilli [24]. A similar increase was also found in infection of other intracellular pathogens, such as human cytomegalovirus [25], and human immunodeficiency virus [25]. The homologous GLUT-11-like protein might play a similar role in sturgeons in mycobacterial infection in favor of bacilli, and the specific mechanism remains to be further studied.

Other DEGs Underlying Innate and Adaptive Immune Responses

A total of 78 differentially expressed contigs (mapped to 68 genes) were found related to immune system processes according to GO annotation and categorization. Specifically, they were involved in innate immune response, phago-

cytosis, cytotoxicity, antigen processing and presentation, leukocyte migration, cell chemotaxis, lymphocyte activation and its regulation. Representative examples of these were listed in Table 4. It seems that the defense of *A. schrenckii* against *M. marinum* was characterized by the activation of the innate immunity.

In conclusion, we used the *M. marinum* infected *A. schrenckii* as a model to study the defense mechanisms of sturgeons against mycobacteria, and to our best knowledge, this is the first transcriptomic study on sturgeons in bacterial infection. We have annotated 33,225 assembled contigs from the liver tissue of infection and control groups and revealed 4043 differentially expressed contigs reflecting a number of signal changes in immune-related pathways. 1340 differentially expressed contigs were enriched for 124 pathways, among which TNF pathway of the highest degree of enrichment, together with the canonical tuberculosis pathway was analyzed based on current knowledge about mycobacterial infection. A total of 2209 unannotated contigs that may represent novel genes associated with infection in sturgeons were provided for future study. Our results reveal the immune mechanism and pathology of sturgeons in mycobacteriosis and provide new molecular targets for further immunological and immunogenetic study.

Acknowledgments

This work was supported by the Key Project of Scientific & Technological Innovation of Hubei Province (2018ABA101), FEBL Project (2016FBZ04), the Key Laboratory of Fishery Drug Development and Key Laboratory of Aquatic Animal Immune Technology, Ministry of Agriculture (201503), Jiangsu Marine and Fisheries Bureau Project (D2017-3).

Conflict of Interest

The authors have no financial conflicts of interest to declare.

References

1. Aronson JD. 1926. Spontaneous tuberculosis in salt water fish. *J. Infect. Dis.* **187**(Suppl 2): 315-320.
2. Ramakrishnan L. 2013. Looking within the zebrafish to understand the tuberculous granuloma, pp. 251-266. *The New Paradigm of Immunity to Tuberculosis*, Ed. Springer
3. Swaim LE, Connolly LE, Volkman HE, Humbert C, Bern DE, Ramakrishnan L. 2006. *Mycobacterium marinum*

- infection of adult zebrafish causes caseating granulomatous tuberculosis and is moderated by adaptive immunity. *Infect. Immun.* **74**: 6108-6117.
4. Parikka M, Hammaren MM, Harjula S-KE, Halfpenny NJ, Oksanen KE, Lahtinen MJ, et al. 2012. *Mycobacterium marinum* causes a latent infection that can be reactivated by gamma irradiation in adult zebrafish. *PLoS Pathog.* **8**: e1002944.
  5. Zhang DF, Ji C, Zhang XJ, Li TT, Li AH, Gong XN. 2015. Mixed mycobacterial infections in farmed sturgeons. *Aquac. Res.* **46**: 1914-1923.
  6. Ng P, Wei C-L, Sung W-K, Chiu KP, Lipovich L, Ang CC, et al. 2005. Gene identification signature (GIS) analysis for transcriptome characterization and genome annotation. *Nat. Methods* **2**: 105-111.
  7. Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* **25**: 402-408.
  8. Yue H, Li C, Du H, Zhang S, Wei Q. 2015. Sequencing and de novo assembly of the gonadal transcriptome of the endangered Chinese sturgeon (*Acipenser sinensis*). *PLoS One* **10**: e0127332.
  9. Kaneko H, Yamada H, Mizuno S, Udagawa T, Kazumi Y, Sekikawa K, et al. 1999. Role of tumor necrosis factor- α in *Mycobacterium*-induced granuloma formation in tumor necrosis factor- α -deficient mice. *Lab. Invest.* **79**: 379-386.
  10. Harris J, Hope JC, Keane J. 2008. Tumor necrosis factor blockers influence macrophage responses to *Mycobacterium tuberculosis*. *J. Infect. Dis.* **198**: 1842-1850.
  11. Roca FJ, Ramakrishnan L. 2013. TNF dually mediates resistance and susceptibility to mycobacteria via mitochondrial reactive oxygen species. *Cell* **153**: 521-534.
  12. Abdallah AM, Bestebroer J, Savage ND, De Punder K, Van Zon M, Wilson L, et al. 2011. Mycobacterial secretion systems ESX-1 and ESX-5 play distinct roles in host cell death and inflammasome activation. *J. Immunol.* **187**: 4744-4753.
  13. MacEwan DJ. 2002. TNF ligands and receptors—a matter of life and death. *Br. J. Pharmacol.* **135**: 855-875.
  14. Roach SK, Schorey JS. 2002. Differential regulation of the mitogen-activated protein kinases by pathogenic and nonpathogenic mycobacteria. *Infect. Immun.* **70**: 3040-3052.
  15. Pathak SK, Bhattacharyya A, Pathak S, Basak C, Mandal D, Kundu M, et al. 2004. Toll-like receptor 2 and mitogen- and stress-activated kinase 1 are effectors of *Mycobacterium avium*-induced cyclooxygenase-2 expression in macrophages. *J. Biol. Chem.* **279**: 55127-55136.
  16. Sasindran SJ, Torrelles JB. 2011. *Mycobacterium tuberculosis* infection and inflammation: what is beneficial for the host and for the bacterium? *Front. Microbiol.* **2**: 2.
  17. Samten B, Townsend JC, Weis SE, Bhounik A, Klucar P, Shams H, et al. 2008. CREB, ATF, and AP-1 transcription factors regulate IFN- γ secretion by human T cells in response to mycobacterial antigen. *J. Immunol.* **181**: 2056-2064.
  18. van der Sar AM, Spaink HP, Zakrzewska A, Bitter W, Meijer AH. 2009. Specificity of the zebrafish host transcriptome response to acute and chronic mycobacterial infection and the role of innate and adaptive immune components. *Mol. Immunol.* **46**(11-12): 2317-2332.
  19. Lemos MP, Rhee KY, McKinney JD. 2011. Expression of the leptin receptor outside of bone marrow-derived cells regulates tuberculosis control and lung macrophage MHC expression. *J. Immunol.* **187**: 3776-3784.
  20. Wieland CW, Florquin S, Chan ED, Leemans JC, Weijer S, Verbon A, et al. 2005. Pulmonary *Mycobacterium tuberculosis* infection in leptin-deficient ob/ob mice. *Int. Immunol.* **17**: 1399-1408.
  21. Mulley JF, Hargreaves AD, Hegarty MJ, Heller RS, Swain MT. 2014. Transcriptomic analysis of the lesser spotted catshark (*Scyliorhinus canicula*) pancreas, liver and brain reveals molecular level conservation of vertebrate pancreas function. *BMC Genomics* **15**: 1074.
  22. Steinwede K, Maus R, Bohling J, Voedisch S, Braun A, Ochs M, et al. 2012. Cathepsin G and neutrophil elastase contribute to lung-protective immunity against mycobacterial infections in mice. *J. Immunol.* **188**: 4476-4487.
  23. Robledo D, Ronza P, Harrison PW, Losada AP, Bermúdez R, Pardo BG, et al. 2014. RNA-seq analysis reveals significant transcriptome changes in turbot (*Scophthalmus maximus*) suffering severe enteromyxosis. *BMC Genomics* **15**: 1149.
  24. Singh V, Kaur C, Chaudhary VK, Rao KV, Chatterjee S. 2015. *M. tuberculosis* secretory protein ESAT-6 induces metabolic flux perturbations to drive foamy macrophage differentiation. *Sci. Rep.* **5**: 12906.
  25. Yu Y, Maguire TG, Alwine JC. 2011. Human cytomegalovirus activates glucose transporter 4 expression to increase glucose uptake during infection. *J. Virol.* **85**: 1573-1580.