The Israeli Journal of Aquaculture - Bamidgeh, IJA_72.2020.1118385, 12 pages

The *IJA* is a peer-reviewed open-access, electronic journal, freely available without charge to users
-
Produced by the AquacultureHub non-profit Foundation
Sale of *IJA* papers is strictly forbidden

Evaluation of two extant population sources of *Clarias macrocephalus* using liver transcriptomic profiles towards reintroduction

July B. Ariñez[1], Apple Gray Deallo[2], Barry Leonard M. Tumbokon[2], Augusto E. Serrano, Jr.[1,2]*

[1] *Institute of Aquaculture, College of Fisheries and Ocean Sciences, University of the Philippines Visayas, Miagao, Iloilo, Philippines*

[2] *UPV-National Institute of Molecular Biology and Biotechnology, University of the Philippines Visayas, Miagao, Iloilo, Philippines*

Key words: Philippine native catfish, *Clarias macrocephalus*, population source for reintroduction, RNA seq, phenotypic variation

Abstract

There is a need to evaluate and choose a possible population source of the threatened catfish *Clarias macrocephalus* for possible reintroduction into aquaculture. Identifying gene transcription profiles that provide fitness benefits in specific environments would promote more effective species reintroduction and conservation practices. In this study, comparison of hepatic transcriptome profiles was made in terms of expression patterns of various growth-, immune- and reproduction-related genes of two *Clarias macrocephalus* strains from Cagayan (CmC) and Agusan del Sur (CmA), located in the Northern and Southern Philippines, respectively. RNA sequencing results showed that there were 13,138 differentially expressed genes (DEGs) identified in CmA vs. CmC, of which 6,795 DEGs exhibited significant upregulation while 6,343 DEGs displayed significant downregulation. From this pool of DEGs, numerous immune- (complements *C3*, *C4*, *C7*), growth- (glutamine synthetase, *GLUL*; glycogen synthase, *GYS*; elongation of very long chain fatty acids protein 6, *ELOVL6*, lipoprotein lipase, *LPL*) and reproduction-related (vitellogenin, *VTG*; cytochrome p450 19b-like, *CYP19b*; progesterone receptor, *PGR*) genes were identified. Pathway enrichment analysis revealed the enrichment of the complement component and coagulation cascades, protein digestion and absorption and lipid metabolism pathways. Results of this study indicated that when reintroduction was to be done, *C. macrocephalus* population from Agusan Marsh is predicted to grow much faster because of up-regulated growth-related genes, to be more resistant to pathogens/environmental stressor as it manifested up-regulated immune-related genes and to reproduce better as reflected by its up-regulated reproduction-related genes than its counterpart catfish from Cagayan.

* Corresponding author., e-mail: aeserrano@up.edu.ph

Introduction

The gradual disappearance of the native catfish *Clarias macrocephalus* in the Philippine aquaculture industry over the last three decades has presented concern among freshwater fish farmers. Initially it was displaced by the Thai catfish *Clarias batrachus*, and later by the exotic bigger African catfish *Clarias gariepinus* (NaNakorn et al., 2004; Senanan et al., 2004; Vidthayanon and Allen 2013; Tan et al., 2016). Currently, *Clarias macrocephalus* is listed as a threatened species in the Philippines (Vidthayanon & Allen 2013) but elder fish consumers who experienced its abundance about 3 decades ago regard it as more tasty and more acceptable than the African catfish; there is also a concern that total extinction might happen soon which would mean loss of heritage.

Reintroduction is the release of a species collected from captive or wild sources into its historical habitat where it has been locally extirpated with the aim to re-establish a self-sustaining population (IUCN/SSC 2013). Populations can differ in reintroduction performance because of their genetic background which may limit their scope for adapting to novel environments. Thus, selecting an appropriate population source is very important for conservation related applications including reintroduction. Many factors can influence the outcome of a reintroduction: the quality of the receiving habitat (Wanless et al. 2002; Moorhouse et al. 2009); the developmental stage at release (Coghlan & Ringler 2004); and source population selection. This third factor-is a key determinant for reintroduction success or failure (Forsman 2014) and is the focus of this study.

Transcriptomics is the study of total RNA transcripts in a cell, tissue or organism at a specific physiological condition or developmental stage (Qian et al., 2014). RNA-sequencing technology can be used to conduct functional genomic assessments on species for which there is no prior molecular information available and can be used to investigate sequence variation between individuals or populations to identify specific genotypes involved in a phenotypic response (e.g. Manousaki et al., 2013; Zhang et al., 2015). Transcriptomics is a tool in conservation physiology as a systems-wide approach to understand mechanisms of tolerance and adaptation to environmental change. Connon et al., (2018) emphasize that while traditional physiological and ecological metrics such as tolerance, growth, reproductive output and survival provide crucial information on the capacity of populations to persist under different environmental conditions, they do not provide insight into the mechanism of how the organisms respond the way they do and transcriptomics is the approach that could provide this insight.

Wild *Clarias macrocephalus* display very low haplotype and nucleotide diversity and are thus suffering from a genetic bottleneck (Tan et al., 2016) in 2 extant populations in the country. Catfish from Agusan del Sur (South of the Philippines) are bigger than that in Cagayan (North of the Philippines) and is perhaps just one of the many phenotypic differences in these two populations. There is a need to compare phenotypic characteristics of these populations from which a population source is selected for purposes of reintroduction. Specifically, this study aimed to 1) compare and characterize prominent genes related to growth-, immune-, and reproduction-related traits and their transcriptional expression levels in the hepatic tissues of the two catfish strains; and 2) to compare and rank differentially expressed genes (DEGs) in the transcriptome of the strains.

Materials and Methods

Through purposive sampling, 18 sexually immature female *C. macrocephalus* (SL: ≥ 15 cm) were collected from Aparri, Cagayan and Bunawan, Agusan del Sur (**Figure 1**) based on the descriptions provided by Ali (1993). The fish from the two sites were collected on the same time of the day, transferred into two oxygenated plastic bags at a stocking density of 9 fish bag^{-1} and transported to the nearest municipal hatchery station for acclimatization prior to final examination and liver extraction.

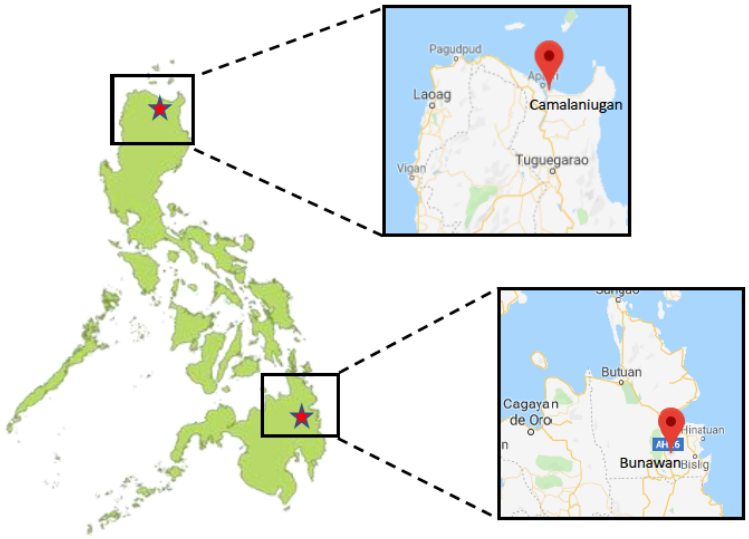

Figure 1 Location of sampling sites

On the day of sample selection for liver extraction, the collected specimens from each site were first distributed into 6 plastic basins at a stocking density of 3 fish container^{-1}; random selection was done through lottery. The catfish samples were then placed in a cold bath for 15 min prior to pithing. The length and weight of each fish sample were measured, dissected, liver excised and appropriate amount of liver were transferred into separate microtubes that contained 500 µL of RNA later^{TM}. The extracted tissue samples were kept at ice water for 24 h prior to -20 °C storage for subsequent RNA extraction.

RNA Extraction, Gel Electrophoresis and Spectrophotometry
Extraction of total RNA was done using Trizol plus reagent kit (Invitrogen, USA) following manufacturer's instructions. Liver tissue (100 mg of liver tissue of the catfish samples) were used for the total RNA extraction. RNA quality and quantity were determined through gel electrophoresis (1.2% agarose in 40 mL TAE buffer) and spectrophotometry (OD 260/280), respectively. RNA integrity was evaluated by the distinct gel red-stained 28s and 18s ribosomal RNA (rRNA) bands. Optical density ratios close or equal to 2.0 were selected for subsequent analysis. The RNA samples were then dessicated in solution of RNAstable® (Biomatrica, San Diego, CA) for 24-h before sending the samples to Novogene Bioinformatics in Beijing, China for cDNA library construction, RNA sequencing and bioinformatic analyses.

cDNA Library Construction and Sequencing
cDNA libraries were constructed (Sun et al., 2016) in which messenger RNA (mRNA) was enriched using oligo(dT) beads, cut into 150 bp and were used to synthesize the first strand using random hexamers and reverse transcriptase. The second strand was generated by the addition of a custom synthesis buffer (Illumina) with dNTPs, RNase H and *Escherichia coli* polymerase I via nick translation. cDNA was purified using AMPure XP beads (Illumina, San Diego, USA) and subjected to end repair, A-tailing, ligation of adapter sequences and polymerase chain reaction (PCR) enrichment of the purified products to yield the final cDNA library. The cDNA libraries were generated from three biological replicates of each catfish strain for sequencing on Hiseq Illumina 4000 platform generating ~150 bp paired-end reads.

Bioinformatic Analysis
The following analysis were done: (1) De Novo Transcriptome Assembly using Trinity software. Corset was then further used to perform hierarchical clustering to remove redundancy. The longest transcripts in each cluster were selected to represent a particular unigene. (2) Gene Functional Annotation in which annotation of contigs or genes was done using 7 databases, namely Nr (NCBI non-redundant protein sequences), Nt (NCBI nucleotide sequences), Pfam (Protein family), Swiss-Prot, KOG/COG (Cluster of Orthologous Groups of Proteins, COG; and eukaryotic Orthologous Groups, KOG), GO (Gene Ontology), and KEGG (Kyoto Encyclopedia of Genes and Genome). (3)

Gene Expression Analysis*,* in which *de novo* assembled transcriptome was used as the reference sequence for gene expression analysis. Bowtie software was initially used to map out reads back to the assembled transcriptome. RSEM software was used to analyze the mapping results of Bowtie and quantify gene expression levels. The reads were also converted into fragments per kilobase of transcript sequence per million mapped reads (FPKM) values. (4) Analysis of Differentially Expressed Genes (DEGs), to identify DEGs using DESeq software which is based on a negative binomial distribution wherein the *p*-values were normalized first using the *q*-value. A q-value <0.05 and |log2 (fold-change)| >1 were set as threshold to classify significant differential expression; (5) Enrichment Analysis - GO enrichment of DEGs was done using GOseq R packages and KOBAS software for KEGG Pathway Enrichment Analysis.

RESULTS

Sample Collection and Identification

Juvenile *C. macrocephalus* (standard length, SL: 16.3-20 cm) were collected on the same season from Cagayan and Agusan del Sur (north and south of the Philippines, respectively) for comparative transcriptome analysis. A total of 18 sexually immature catfish juveniles were selected based on the rubrics of Ali (1993) for sexually maturity such as the presence of firm and translucent ovaries with no evident presence of oocytes. Catfish from Agusan del Sur were generally bigger than that from Cagayan (**Table 1**).

Morphological features that distinguish *C. macrocephalus* from the other *Clarias* species were observed such as (1) the presence of rounded occipital process, (2) distinct white spots on the sides of the body, (3) number of gill rakers on the first branchial arch, (4) long dorsal fins with specified number of soft rays (**Table 1**) as described (Conlu (1986; Teugels et al. (1999); Sudarto and Pouyad (2005).

Table 1 Morphological measurements of Clarias macrocephalus collected from Agusan del Sur (CmA) or Cagayan (CmC).

Sample	*OP*	*ASL (cm)*	*ABW (cm)*	*ADF*	*AAF*	*APF*	*GRBA*
CmA	rounded	19.5	102.5	68	48	I, 8	32
CmC	rounded	16.9	63.5	68	48	I, 8	31

CmA, *C. macrocephalus* samples from Agusan del Sur; *CmC*, *C. macrocephalus* samples from Cagayan; *CgC*;*OP*, occipital process; A*SL*, average standard length; A*BW*, average body weight; *ADF*, average number of dorsal fin rays; *AAF*, average number of anal fin rays; *APF*, average number of pectoral fin rays (numbers expressed in roman numerals denote counts of spiny rays); *GRBA*, number of gill rakers on the first branchial arch.

Sequencing and De Novo Assembled Transcriptome

A total of 88,199,051 raw reads were generated from the pooled RNA samples extracted from the hepatic tissues of the 2 catfish strains. After removing low quality, over-represented and adapter sequences, 58,877,141 clean reads with Q20 percentage of 97.22-97.88 % were retained (**Table 2**). The *de novo* assembled transcript yielded 86,011 unigenes with an average length of 1,155 bp and an N50 of 1,741 bp (data not shown). Distribution analysis of unigenes with various lengths showed that 200-500 bp unigenes were the most abundant (29.97%) followed by 500-1,000 bp, 1000-2,000 bp and >2,000 bp unigenes which comprised 29.75%, 23.92%, and 16.36% of the total unigene counts, respectively (data not shown).

Table 2 Statistics of the assembled reads and unigenes generated from sequencing the cDNA libraries of *Clarias macrocephalus* from Agusan (CmA) or Cagayan (CmC).

Sample	*Raw reads*	*Clean reads*	*Q20 (%)*	*GC content (%)*
CmA	31,608,848	31,193,690	97.88	47.31
CmC	28,057,443	27,683,451	97.22	46.80

Functional Annotation and Gene Identification

Of the 86,011 unigenes, 6.44% (5,543 unigenes) were annotated in all databases while 61.06% (52,519 unigenes) resulted in significant hits in at least one database only. Results of the functional annotation also showed that Nt obtained the highest number of annotated unigenes which comprised 55.52% (47,750 unigenes) of the total number of assembled unigenes. In contrast, GO exhibited the lowest annotation percentage which annotated a total of 11,239 unigenes.

BLASTx top-hit species distribution revealed that 74.5% of the unigenes exhibited highest similarity to the genes sequenced from *I. punctatus* followed by *Pygocentrus nattereri* (5.3%), *Astyanax mexicanus* (3.5%), *Cyprinus carpio* (1.4%), *Danio rerio* (1.1%) and others (14.2%) (Figure 2). These results suggested that *C. macrocephalus* was closely related to *I. punctatus*; both belong to one taxonomic order. The similarity of sequences to those of other teleosts indicated that numerous genes are phylogenetically conserved among fishes.

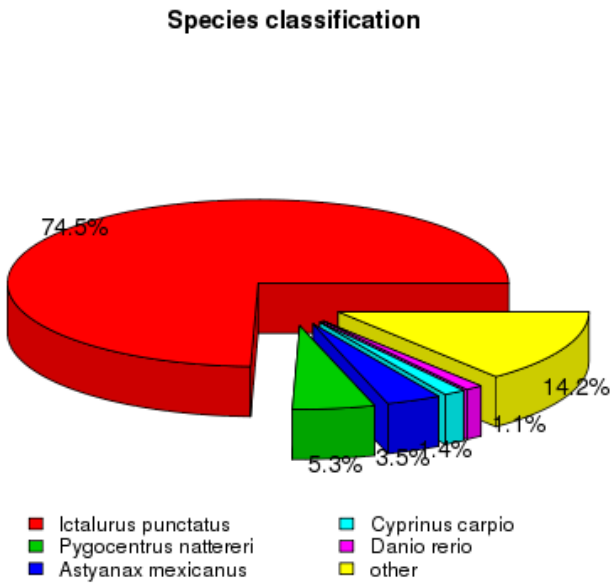

Figure 2 Homology of *C. macrocephalus* unigene sequences to other fish species.

Gene Expression Analysis

Normalized FPKM values showed that more than half (59.81-68.44%) of the genes from the 2 catfish strains were either unexpressed or expressed at very low levels (0 < FPKM < 1) while only less than two percent (1.73-1.85%) were highly expressed (FPKM > 60) (data not shown).

A co-expression analysis (RSEM software) was done to determine the overall similarity of gene expression between the 2 catfish strains. A total of 44,125 genes were expressed in the hepatic tissues of the two *C. macrocephalus* strains from Cagayan and Agusan del Sur of which 8,669 and 6,948 genes were identified to be specifically expressed in CmA and CmC tissues, respectively (**Figure 3**).

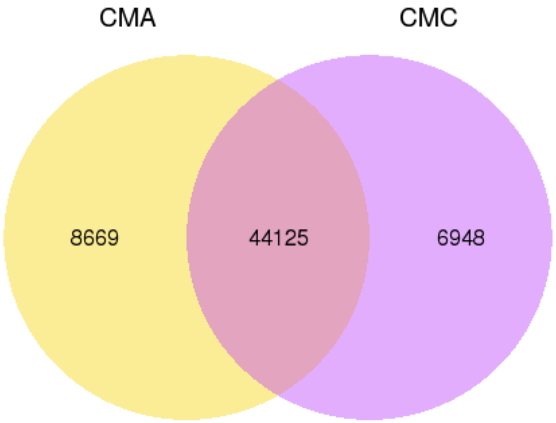

Figure 3 Venn diagram of co-expressed and specifically-expressed genes in CmA and CmC.

Differentially Expressed Genes (DEGs)

The number and overall distribution of upregulated and downregulated DEGs were determined and fitted to a volcano plot (**Figure 4**). From a total of 13,138 DEGs that were identified in CmA vs. CmC, 6,795 DEGs were significantly upregulated while 6,343 DEGs were downregulated.

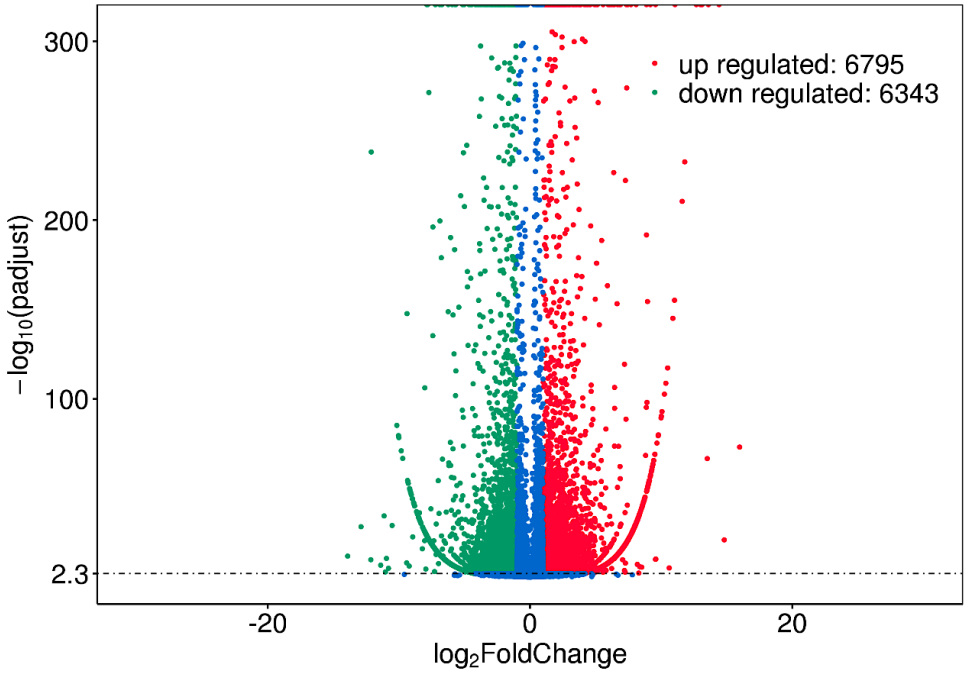

Figure 4 Volcano plot showing the distribution of DEGs in CmA relative to CmC. The red dots represent significantly up-regulated genes while the green dots the down-regulated genes. Blue dots represent the unaffected genes.

DEGs were ranked to determine the underlying genetic factors that are responsible for the desirable traits exhibited by CmA. Based on the top 30 upregulated or downregulated DEGs, many genes such as those associated with growth, immunity and reproduction were found in CmA relative to CmC. Glutamine synthetase (*GLUL*) (**Table 3**) and G0/G1 switch protein 2 (*G0S2*) (**Table 4**) represented the most significantly upregulated and downregulated genes in CmA vs. CmC, respectively.

Majority of the immune-related genes with the highest differential expression in CmA were involved in the complement component system, pathogen recognition, iron storage and transport and regulation of cytokine production. Complement component C7 (*C7*), complement C4-A-like (*C4*), complement C3-like isoform X2 (*C3*), C-type lectin domain family 4 member C (*CLEC4C*), C-type lectin domain family 4 member E-like (*CLECL4E*), pentraxin-related protein PTX3 (*PTX3*), transferrin receptor protein 1-like (*TFRC*) (Table 3). In contrast, interferon-induced protein 44-like (*IFI44L*) was significantly downregulated (**Table 4**).

In terms of growth, 3 prominent gene groups involved in protein synthesis, glycogenesis and lipid metabolism were also detected in CmA. Glycogen synthase (*GYS*), elongation of very long chain fatty acids protein 6-like (*ELOVL6*) and lipoprotein lipase (*LPL*) were upregulated (**Table 3**) while mitochondrial carnitine/acylcarnitine carrier protein (*CACT*), carnitine opalmitoyltransferase 1, liver isoform-like (*CPT1*) and peroxisome proliferator-activated receptor alpha-like (*PPAR-α*) were downregulated (**Table 4**).

Among reproduction-related DEGs, a small number of genes were identified in CmA vs. CmC. Those that are primarily involved in steroidogenesis and oocyte development such as progesterone receptor (*PGR*), vitellogenin (VTG), bone morphogenetic proteins (BMP) and activin membrane-bound inhibitor homolog (*BAMBI*) and cytochrome p450 19b (*CYP19B*) and 3-hydorxy-3-methyglutaryl-coenzyme A reductase (*HMGCR*) were upregulated (Table 3). However, those that are associated with placenta development and inhibition of steroid hormone activities such as *r*-spondin-3 (*RSPO3*), scaffold attachment factor B2 isoform X1 (*SAFB2*) and relaxin-3-like isoform X2 (*RLN3*) were downregulated (**Table 4**).

Enrichment Analysis

In the GO enrichment analysis, the extracellular region term under the cellular component (CC) domain was the only significantly enriched term in CmA vs. CmC (Figure 5). Many DEGs were related to catalytic (enzymatic), metabolic and immunological processes in CmA. Apart from extracellular region, the extracellular space term was also highly enriched in the CC domain. Enzymatic activities related to hydrolase, peptidase, oxidoreductase, enzyme binding and enzyme regulator activity composed the most enriched terms under the molecular function (MF) domain. Additionally, two major term groups associated with metabolic and immunological processes comprised the biological process (BP) domain. The metabolism-related group consisted of lipid, cellular amino acid and small molecule metabolic process terms while the immune-related group consisted of immune system process, transmembrane transport, cell proliferation, cell adhesion and homeostatic process terms.

Table 3 Top 30 significantly upregulated genes in CmA relative to CmC.

Description	Padj	Pvalue	Log2 FC	Species
Glutamine synthetase	0	0	13.12320534	*Clarias magur*
Progesterone receptor	0	0	8.882846371	*Ictalurus punctatus*
Complement c4-A-like	5.29E-14	1.11E-14	6.716785398	*Ictalurus punctatus*
Complement component C7	7.38E-17	1.31E-17	7.031864826	*Ictalurus punctatus*
Vitellogenin	0	0	6.825938905	*Clarias macrocephalus*
C-type lectin domain family 4 member C	0	0	6.058854264	*Ictalurus punctatus*
Lipoprotein-lipase-like	0	0	5.955850391	*Ictalurus punctatus*
Transferrin receptor protein 1-like	0	0	5.054366349	*Ictalurus punctatus*
Alanine aminotransferase 2-like	0.00012	6.83E-05	4.81400989	*Ictalurus punctatus*
E3 ubiquitin-protein ligase RNF144B-like	0.00013	7.11E-05	4.807743131	*Ictalurus punctatus*
Tax1-binding protein 1 isoform X2	0.00013	7.57E-05	4.797840037	*Ictalurus punctatus*
Isocitrate dehydrogenase [NAD] subunit γ, mitochondrial-like	0.00018	0.00011	4.741653941	*Ictalurus punctatus*
Reticulun-4 isoform X4	0.00019	0.00011	4.733175787	*Ictalurus punctatus*
Glutaminase liver isoform, mitochondrial	0.0002	0.00012	4.727495878	*Ictalurus punctatus*
C-type lectin domain family 4 member	0.00021	0.00012	4.717024274	*Ictalurus punctatus*
Fumarate hydratase, mitochondrial	0.00022	0.00013	4.707438225	*Scleropages formosus*
Pentraxin-related protein PTX3	0.00022	0.00013	4.707438225	*Ictalurus punctatus*
Plexin-B2-like isoform X1	0.00023	0.00013	4.705513349	*Ictalurus punctatus*
Integrator complex subunit 6	0.00025	0.00015	4.688072901	*Ictalurus punctatus*
Alpha-2 macroglobulin-like	0.00027	0.00016	4.673376389	*Ictalurus punctatus*
Elongation of very long chain fatty acids protein	0.00027	0.00016	4.670419036	*Ictalurus punctatus*
BMP and activin membrane-bound inhibitor homolog	0.00028	0.00017	4.665476602	*Ictalurus punctatus*
Uncharacterized protein LOC108270446	0.00029	0.00018	4.658528625	*Ictalurus punctatus*
Lamin-A	0.00032	0.0002	4.637481641	*Astyanax mexicanus*
Tyrosine-protein kinase Tec-like, partial	0.00033	0.0002	4.63546106	*Sinocyclocheilus rhinoceros*
Actin-related protein 6 isoform X1	0.00035	0.00022	4.623277778	*Ictalurus punctatus*
Glycogen synthase, liver	0.00035	0.00022	4.62123719	*Pygocentrus nattereri*
Cytochrome P450 19b-like	0.00037	0.00023	4.613045855	*Scleropages formosus*
Complement C3-like isoform X2	0.00038	0.00024	4.607902552	*Ictalurus punctatus*
CC chemokine SCYA101	0.00051	0.00033	4.547935029	*Ictalurus punctatus*

Table 4 Top 30 significantly downregulated genes in CmA relative to CmC.

Description	Pvalue	Padj	Log2 FC	Species
G0/G1 switch protein 2	0	0	-5.72013775	*Ictalurus punctatus*
Uncharacterized protein LOC108272473	0	0	-5.66084485	*Ictalurus punctatus*
Telomerase Cajal body protein 1	0	0	-5.00680729	*Ictalurus punctatus*
Hypoxia induced factor 2 alpha	0.0001	0.00018	-	*Clarias batrachus*
Target of Myb protein 1-like	0.00017	0.00028	-4.84446225	*Astyanax mexicanus*
Hamartin-like	0.00018	0.00029	-4.83752620	*Ictalurus punctatus*
Cyclin-dependent kinase-12	0.00019	0.00032	-4.82121122	*Ictalurus punctatus*
Apoptosis regulator BAX-like isoform X2	0.00028	0.00044	-4.75649478	*Ictalurus punctatus*
Peptidyl-prolyl cis-trans isomerase A-like	0.00033	0.00052	-4.72175926	*Ictalurus punctatus*
Opioid growth factor receptor-like protein 1	0.00034	0.00053	-4.71798750	*Ictalurus punctatus*
Relaxin-3 like isoform X2	0.00046	0.0007	-4.66019493	*Ictalurus punctatus*
Scaffold attachment factor B2 isoform X1	0.00051	0.00076	-4.64040456	*Ictalurus punctatus*
Peripilin-2	0.00053	0.00079	-4.63241177	*Ictalurus punctatus*
OTU domain-containing protein 4 isoform X1	0.00058	0.00086	-4.61358790	*Ictalurus punctatus*
Sterol O-acyltransferase 1 isoform X1	0.00063	0.00092	-4.59725533	*Ictalurus punctatus*
Carnitine O-palmitoyltransferase 1, liver isoform-like	0.00066	0.00095	-4.58901918	*Danio rerio*

Arachidonate 15-lipoxygenase B-like	0.0007	0.001	-4.57657611	*Cyprinodon variegatus*
Heme oxygenase 1	0.00091	0.00127	-4.52138046	*Ictalurus punctatus*
Uncharacterized protein LOC108258246	0.00091	0.00127	-4.52138046	*Ictalurus punctatus*
Transmembrane protein 168	0.00147	0.00195	-4.41513096	*Ictalurus punctatus*
Cyclin-A2	0.00179	0.00233	-4.36937737	*Ictalurus punctatus*
Tripartite motif-containing protein 16-like	0.0019	0.00246	-4.35487780	*Ictalurus punctatus*
N-acylneuraminate-9-phosphatase	0.00242	0.00304	-4.29706316	*Ictalurus punctatus*
Centrosomal protein of 55 kDa	0.00261	0.00323	-4.27840371	*Ictalurus punctatus*
Period circadian protein homolog 2-like	0.00308	0.00375	-4.23683459	*Ictalurus punctatus*
Interferon-induced protein 44-like	0.00356	0.00424	-419945237	*Ictalurus punctatus*
Peroxisome proliferator-activated receptor alpha-like	0.00401	0.00471	-4.16846467	*Ictalurus punctatus*
Pyruvate dehydrogenase kinase, isozyme 4	0.00421	0.0049	-4.15550902	*Ictalurus punctatus*
Mitochondrial carnitine/acylcarnitine carrier protein	0.00427	0.00496	-4.15178593	*Astyanax mexicanus*
R-spondin-3	0.00014	0.00024	-4.00680729	*Ictalurus punctatus*

KEGG enrichment analysis resulted in 23 pathways that were significantly enriched, the top pathways were protein digestion and absorption, arginine biosynthesis, central carbon metabolism in cancer, insulin resistance, alanine, aspartate and glutamate metabolism, proximal tubule bicarbonate reclamation, complement and coagulation cascades, glycerophospholipid cascades, GABAergic synapse and pancreatic secretion comprised the top enriched pathways (Figure 6).

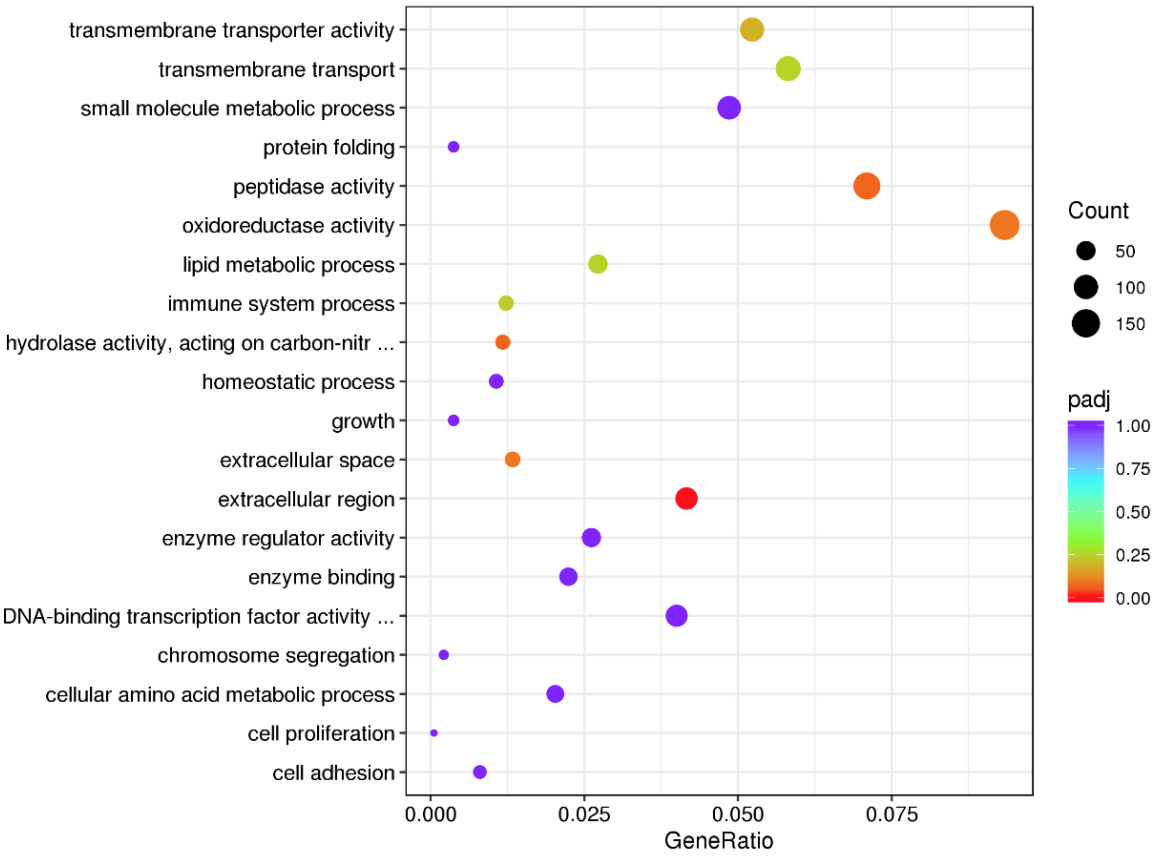

Figure 5 GO enrichment of DEGs in CmA relative to CmC.

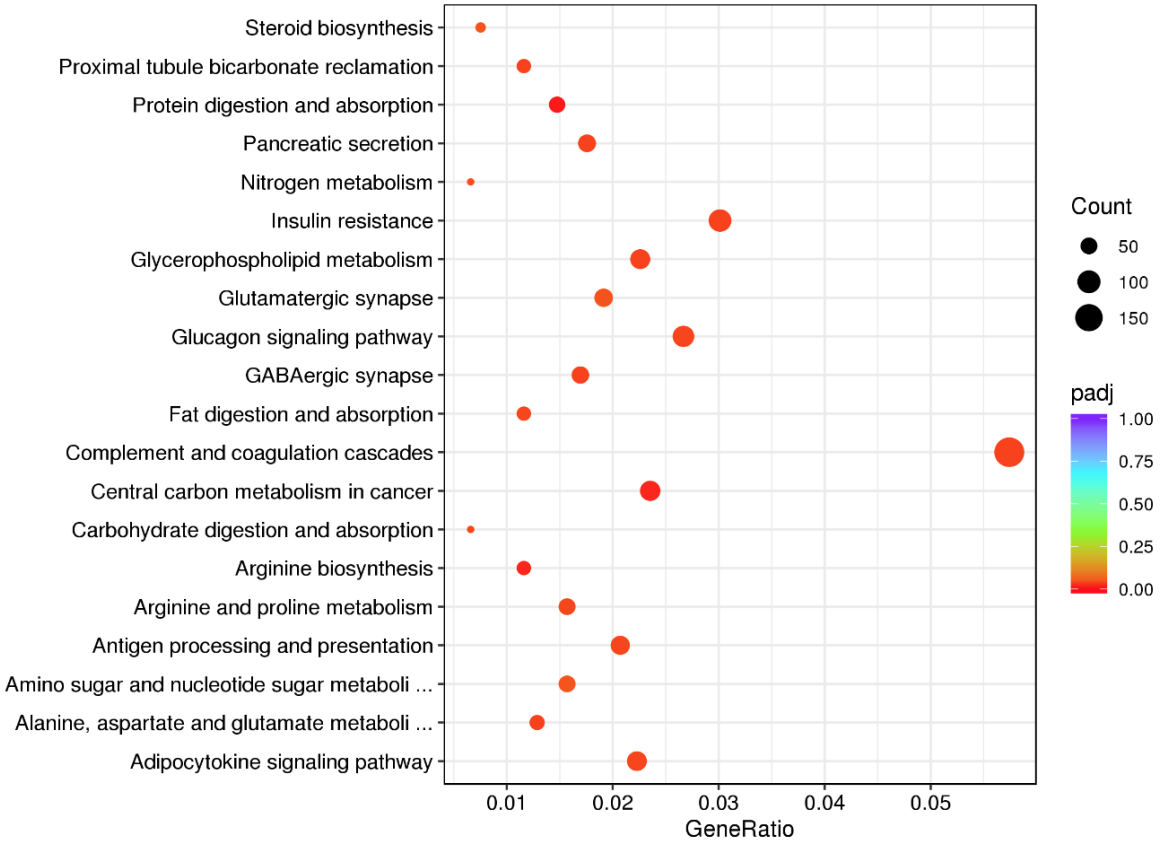

Figure 6 KEGG enrichment of DEGs in CmA relative to CmC.

Discussion

In the present study, RNA seq was used to compare two *Clarias macrocephalus* population that are geographically distant from each other and to be able to identify a source population for reintroduction of the fish into aquaculture. Liver transcriptome profiles were compared to identify the presence and degree of expression of various growth, immune- and reproduction-related genes for identification of physiological phenotype that would enable adaptation and survival to a new environment.

Immune-Related Genes

Numerous DEGs related to the innate immune system of fish have been identified in the present study. The largest group of immune-related genes in CmA relative to CmC were dominated by the complement component system. The complement components *C3*, *C4* and *C7* showed significant upregulation in CmA relative to CmC. Furthermore, the Agusan strain of *C. macrocephalus* in the present study exhibited a high expression of two members of the C-type lectin domain family, *CLEC4C* and *CLECL4E*. Also, transferrin receptor gene was upregulated in CmA relative to CmC; these transferrin receptors are involved in the regulation of inflammatory response during pathogen infection (Aisen et al., 2001). In addition, the interferon-induced protein 44 (*IFI44*) gene was downregulated in CmA. IFI44 obstructs the host immune response and promotes virus replication during an induction of viral infection (DeDiego et al., 2019).

The ability of the host to demonstrate a strong immune response is critical for its survival. In the present study, variation on the presence and significant expression of immune-related genes implied that the 2 catfish strains might have different levels of resistance and susceptibility to various diseases. From the results of gene expression of immune-related DEGs, the CmA strain appeared to have a robust immunological machinery that was essential for the maintenance of homeostasis and elimination of unwanted pathogenic organisms.

Growth-Related Genes

Glutamine synthase (*GLUL*) and glycogen synthase (GYS) exhibited upregulation in CmA relative to CmC. Furthermore, the up-regulated elongation of very long chain fatty acids protein 6 (*ELOVL6*)

in CmA indicated a promoted *de novo* synthesis of long chain fatty acid biosynthesis in lipogenic tissues like adipose and liver (Shi et al., 2017; Li et al., 2019). This could explain the generally bigger sizes of the CmA catfish than that of CmC. Higher growth rate is an advantage for reintroduction of the species of interest since it could immediately attain a size that could render it less vulnerable to predators and/or competitors.

Reproduction-related genes

Estrogens serve as primary activators of vitellogenin (*VTG)*, the presence of genes involved in estrogen synthesis in the present study more likely contributed to the upregulation of hepatic *VTG* in CmA. The up-regulation of other genes related to synthesis and regulation of sex steroid hormones further pointed to a predicted better reproductive performance of CmA than that of CmC. BMP and activin membrane-bound inhibitor homolog (*BAMBI*) and progesterone receptor (*PGR*) were upregulated in the present study. *BAMBI* functions as a negative regulator of transforming growth factor beta (*TGF-β*) genes which also has inhibitory effects on oocyte maturation and ovarian steroid production. *TGF-β* negatively affects the production of 17α,20β-dihydroxyprogesterone (DHP), also known as maturation inducing hormone (MIH) (Kohli et al., 2005). When MIH is stimulated, it results in increased expression of *BAMBI* (Lankford & Weber, 2010). The significant upregulation of *BAMBI* in the present study could have stimulated the production of MIH by inhibiting the expression of *TGF-β*.

Strategies for survival of a species through reproduction could either be by sheer high number of offspring or by fewer but bigger offspring. The up-regulation of the reproduction-related genes in CmA vaguely relates this to the kind of reproductive strategy of the CmA catfish, but knowing the low fecundity of this species (80-110 eggs g^{-1} body weight; Fermin et al., 1997), the strategy is more likely to produce few but bigger offspring. This is an advantageous phenotype for reintroduction and is related to the growth advantage of attaining bigger size in a shorter time, enabling the fish to evade predators and/or competitors.

Identifying differentially expressed genes in the same species but geographically distant from each other could stem from either differences in evolutionary histories and marked life history differences (He et al., 2015). Nevertheless, these differences among the two populations likely underlie adaptations to different environmental conditions. Identifying genes with transcription profiles that indicate selection-based differences among the populations is important in conservation and management (He et al., 2015).Transcriptional profiling of potential source populations can enhance reintroduction efforts in two ways: (1) gene expression comparisons can identify functional differences that are related to important physiological processes and responses to environmental stressors, and subsequently, (2) variation in individual gene transcription can be used to predict specific trait response upon reintroduction (Miller et al., 2011).

In conclusion, based on comparison of the transcriptomic profile of *Clarias macrocephalus* collected from Agusan del Sur (CmA, south of the Philippines) and that from Cagayan (CmC, north of the Philippines), CmA is predicted to be a better population source for reintroduction than CmC in terms of upregulated DEGs for growth-, immune- and reproduction-related genes.

Acknowledgements

The authors would like to thank the Philippine Council for Agriculture, Aquatic and Natural Resources Research and Development (PCAARRD) of the Department of Science and Technology (DOST) for the research grant; the Office of the Vice-Chancellor of Research and Extension of the University of the Philippines Visayas for the publication grant, and the Accelerated Science and Technology Human Resources Development program-National Science Consortium of Department of Science and Technology (DOST-ASTHRDP) for the Scholarship of Mr. July B. Arinez. We are also grateful to Mr. Vicente Nim, Ms. and Ms. Shellah Dee Canillo for their technical support in the laboratory.

References

Aisen, P., Enns, C., Wessling-Resnick, M., **2001**. Chemistry and biology of eukaryotic iron metabolism. Int. J. Biochem. Cell Biol. 33(10): 940–959. https://doi.org/10.1016/s1357-2725(01)00063-2.

Ali, A.B., 1993. Aspects of the fecundity of the feral catfish, *Clarias macrocephalus* (Gunther), population obtained from the rice fields used for rice-fish farming in Malaysia. *Hydrobiologia* 254(2): 81–89. https://doi.org/10.1007/BF00014311.

Coghlan, S.M., Ringler, N.H. 2004. A comparison of Atlantic salmon embryo and fry stocking in the Salmon River, New York. North Am. J. Fish. Manage. 24: 1385–1397. https://doi.org/10.1577/M03-253.1.
Conlu, P.V., 1986. *Guide to Philippine flora and fauna*. Vol. 9. Fishes. Quezon City: Natural Resources Management Center, University of the Philippines, 495p.
Connon, R.E., Jeffries, K.M., Komoroske, L.M., Todgham, A.E., Fangue, N.A., 2018. The utility of transcriptomics in fish conservation. *Journal of Experimental Biology*, 221, jeb148833. https://doi:10.1242/jeb.148833.
DeDiego, M.L., Nogales, A., Martinez-Sobrido, L., Topham, D.J., 2019. Interferon induced protein 44 interacts with cellular FK506-binding protein 5, negatively regulates host antiviral responses, and supports virus replication. *mBio* 10:e01839-19. https://doi.org/10.1128/mBio.01839-19.
Forsman, A., 2014. Effects of genotypic and phenotypic variation on establishment are important for conservation, invasion, and infection biology. *Proceedings of the National Academy of Sciences of the United States of America*, 111: 302–307. https://doi.org/10.1073/pnas.1317745111.
He, X., Wilson, C.C., Wellband, K.W., Houde, A.L.S., Neff, B.D., Heath, D.D., 2015. Conservation Genetics, 16:277-287. https://doi:10.1007/s10592-014-0657-1.
IUCN/SSC, 2013. *Guidelines for reintroductions and other conservation translocations*. Version 1.0. IUCN Species Survival Commission, Gland, Switzerland.
Li, Y., Pang, Y., Xiang, X., Du, J., Mai, K., Ai, Q., 2019. Molecular cloning, characterization, and nutritional regulation of elovl6 in large yellow croaker (*Larimichthys crocea*). Int. J. Mol. Sci. 20(7): 1-13. https://doi.org/10.3390/ijms20071801.
Manousaki, T., Hull, P. M., Kusche, H., Machado-Schiaffino, G., Franchini, P., Harrod, C., Elmer, K. R., Meyer, A., 2013. Parsing parallel evolution: ecological divergence and differential gene expression in the adaptive radiations of thick-lipped Midas cichlid fishes from Nicaragua. *Molecular Ecology*, 22: 650-669. https://doi:10.1111/mec.12034.
Miller, K.M., Li, S., Kaukinen, K.H., Ginther, N., Hammil, E., Curtis, J.M.R., Patterson, D.A., Sierocinski, T., Donnison, L., Pavlidis, P., Hinch, S.G., Hruska, K.A., Cooke, S.J., English, K.K., Farrell, A.P., 2011. Genomic signatures predict migration and spawning failure in wild Canadian salmon. Science 331:214–217. https://doi:10.1126/science.1196901.
Moorhouse, T.P., Gelling, M., Macdonald, D.W., 2009. Effects of habitat quality upon reintroduction success in water voles: evidence from a replicated experiment. Biol. Conserv. 142: 53–60. https://doi.org/10.1016/j.biocon.2008.09.023.
Na-Nakorn, U., Kamonrat, W., Ngamsiri, T., 2004. Genetic diversity of walking catfish, *Clarias macrocephalus*, in Thailand and evidence of genetic introgression from introduced farmed *C. gariepinus*. Aquaculture 240: 145–163. https://doi.org/10.1016/j.aquaculture.2004.08.001.
Qian, X., Ba, Y., Zhuang, Q., Zhong, G., 2014. RNA-seq technology and its application in fish transcriptomics. OMICS. 18(2): 98–110. https://doi.org/10.1089/omi.2013.0110
Senanan, W., Kapunscinski, A.R., Na-Nakorn, U., Miller, L.M., 2004. Genetic impacts of hybrid catfish farming (*Clarias macrocephalus* × *C. gariepinus*) on native catfish populations in central Thailand. Aquaculture 235: 167–184. https://doi.org/10.1016/j.aquaculture.2003.08.020.
Shi, H.B., Wu, M., Zhu, J.J., Zhang, C.H., Yao, D.W., Luo, J., Loor, J.J., 2017. Fatty acid elongase 6 plays a role in the synthesis of long-chain fatty acids in goat mammary epithelial cells. J. Dairy Sci. 100(6): 4987–4995. https://doi.org/10.3168/jds.2016-12159.
Sudarto, Pouyaud, L., 2005. Identification key based on morphological characters of the Southeast Asian species of the Clarias (PISCES: CLARHDAE) [Kunci Identifikasi Ikan Clarias (Pisces: Clariidae) Asia Tenggara Berdasarkan Ciri Morfologi. Indonesian Journal of Ichthyology, 5(2):39-47. https://doi.org/10.32491/jii.v5i2.236.
Tan, M.T.C., Jumawan, J.C., Quilang, J.P., 2016. Low genetic diversity in *Clarias macrocephalus* Günther, 1864 (Siluriformes: Clariidae) populations in the Philippines and its implications for conservation and management. J. Threat. Taxa 8(6): 8849-8859. https://doi.org/10.11609/jott.2261.8.6.8849-8859.
Tan-Fermin, J.D., Ijiri, S., Ueda, H., Adachi, S., Yamauchi, K., 1997. Ovarian development and serum steroid hormone profiles in hatchery-bred female catfish Clarias macrocephalus (Gunther) during an annual reproductive cycle. Fisheries Science, 63(6): 867-872. https://doi.org/10.2331/fishsci.63.867.
Teugels, G.G., Diego, R.C., Pouyaud, L., Legendre, M., 1999. Redescription of *Clarias macrocephalus* (Siluriformes: Clariidae) from South-east Asia. Cybium 23(3): 285–295.

Tchoudakova, A., Callard, G.V., 1998. Identification of multiple CYP19 genes encoding different cytochrome P450 aromatase isozymes in brain and ovary. Endocrinol. 139(4): 2179–2189. https://doi.org/10.1210/endo.139.4.5899.
Tong, S.K., Chiang, E.F., Hsiao, P.H., Chung, B., 2001. Phylogeny, expression and enzyme activity of zebrafish cyp19 (P450 aromatase) genes. J. Steroid Biochem. Mol. Biol. 79(1-5): 299–303. https://doi.org/10.1016/S0960-0760(01)00146-7.
Vidthayanon, C., Allen, D., 2013. [Internet]. *Clarias macrocephalus*. The IUCN Red List of Threatened Species 2013. [cited July 29, 2019]. Available from: www.iucnredlist.org.
Wanless R.M., Cunningham J., Hockey P.A.R., Wanless J., White R.W. & Wiseman R., 2002. The success of a soft-release reintroduction of the flightless Aldabra rail (*Dryolimnas [cuvieri] aldabranus*) on Aldabra Atoll, Seychelles. Biol. Conserv. 107: 203-210. https://doi.org/10.1016/S0006-3207(02)00067-8.
Yao, J., Mu, W., Liu, S., Zhang, J., Wen, H., Liu, Z., 2015. Identification, phylogeny and expression analysis of suppressors of cytokine signaling in channel catfish. Mol. Immunol. 64(2): 276–284. https://doi.org/10.1016/j.molimm.2014.12.003
Zhang, R., Ludwig, A., Zhang, C., Tong, C., Li, G., Tang, Y., Peng, Z., Zhao, K., 2015. Local adaptation of *Gymnocypris przewalskii* (Cyprinidae) on the Tibetan Plateau. Sci. Rep. 5, 9780. https://doi.org/10.1038/srep09780.
Zardakis, I., Duraj, S., Chondrou, M., 2005. Molecular cloning of the seventh component of complement in rainbow trout. Dev. Comp. Immunol. 29(2): 95–102. https://doi.org/10.1016/j.dci.2004.06.006