Characterization of grass carp spleen transcriptome during GCRV infection

B.H. Xu[1], L. Zhong[1], Q.L. Liu[1], T.Y. Xiao[1], J.M. Su[2], K.J. Chen[1], H.Q. Wang[1], Y.J. Dai[1] and J. Chen[1]

[1]College of Animal Science and Technology, Hunan Agricultural University, Changsha, China
[2]College of Animal Veterinary and Medicine, Hunan Agricultural University, Changsha, China

Corresponding author: T.Y. Xiao
E-mail: tyx1128@163.com

ABSTRACT. The aim of the study was to investigate the grass carp hemorrhagic infection pathway and its key-related genes. Grass carp reovirus (GCRV) might cause hemorrhagic disease in grass carps. Healthy grass carp fingerlings (N = 60) were divided into control and infected groups. Fish in the control group were intraperitoneally (*ip*) injected with 0.6% fish physiological saline; the infected group received 5,000,000 50% tissue culture infective doses of GCRV 873 standard strain, a double-stranded RNA (dsRNA) virus strain, *ip*, in 0.5 mL. Illumina HiSeq™ 2000 was used for transcriptome sequencing, and real-time polymerase chain reaction (PCR) used to detect complement factors II (*C2*), III (*C3*), and V (*C5*); profibrinolysin (*PLG*); and coagulation factor II (*F2*) expression. A total of 2,722,223 reads were detected in the control group, and 2,751,111 in the infected group. Among 11,023 unigenes obtained after transcriptome assembly, 10,021 unigenes were significantly differentially expressed. Gene ontology and KEGG analysis, a collection of databases dealing with genomes and biological pathways, were performed to classify unigenes

into functional categories, to understand gene function and identify regulatory pathways. Real-time PCR analysis showed that *C2*, *C3*, *C5*, *PLG*, and *F2* expression levels were down-regulated, confirming results of pathway-enrichment analysis. This is the first application of high-throughput sequencing technology to investigate the *in vivo* effects of GCRV, on genes and pathways involved in the immune response to infection in grass carp.

INTRODUCTION

Grass carp (*Ctenopharyngodon idella*) is an important species of freshwater fish in China, where it accounts for 20% of the total freshwater fishery production (Jia et al., 2014). It is one of the four major freshwater aquaculture fish in China and is known for its delicious and nutrient-rich meat; the scales and skin are also used to produce glue (Liu et al., 2013). However, grass carp farming in China is severely affected by disease outbreaks caused by bacterial, viral, and protozoan infections (Zhang et al., 2013). Of these, grass carp hemorrhage results in the most serious disease, and causes significant loss of fingerlings during rearing (Ye et al., 2012). The hemorrhagic disease is highly contagious, and is caused by grass carp hemorrhagic virus (GCHV), a double-stranded RNA (dsRNA) virus also known as grass carp reovirus (GCRV) (Su et al., 2012). Grass carp reovirus infects grass carp, black carp (*Mylopharyngodon piceus*), slender top-mouth gudgeon (*Pseudorasbora parva*), and rare minnow (*Gobiocypris rarus*), resulting in significant losses in freshwater fish culture, and at least one billion US dollars in economic losses in China annually (Zeng et al., 2014). Infection results in a mortality rate of up to 85% during the summer months, when temperatures range between 24° and 30°C (Liang et al., 2014). Grass carp reovirus belongs to the family Reoviridae, genus *Aquareovirus*, and its genome consists of 11 segmented dsRNAs packaged into an unenveloped, double-layered capsid (Ye et al., 2012). The genus *Aquareovirus* is designated A to G (AQRV-A to AQRV-G), and GCRV873 belongs to AQRV-C (Wang et al., 2012). The exact mechanisms of pathogenesis of GCRV on grass carp hemorrhagic disease remain unknown. Genomic surveys, such as transcriptome analysis, have the required high resolution to provide detailed molecular information on gene expression, gene network, and cellular pathways to address specific disease-related questions (Meyer et al., 2009).

Transcriptome analysis provides a comprehensive view of coding and non-coding RNAs, including messenger RNA (mRNA) profiles (Endo et al., 2012). *De novo* assembly of transcriptome can facilitate increased coverage depth, due to significantly fewer nucleotides in the transcriptome than in the whole genome (Parchman et al., 2010). In recent years, transcriptome analysis has been reported for many fish species, including zebrafish, fugu, Atlantic salmon, sea bass, rainbow trout, Atlantic halibut, bluefin tuna, turbot, and Senegal sole fish (Chen et al., 2012).

The immune system is a network of lymphoid organs, diverse immune cells and soluble cytokine mediators and immune components that are highly sensitive to toxic agents and viruses (Wen et al., 2005; Huang et al., 2010). The spleen is a major peripheral lymphoid organ, and the main immune organ of grass carp. We investigated the grass carp hemorrhagic

pathway and its key genes, by comparing gene expression profiles of GCRV-infected and uninfected carp, through transcriptome analysis. We constructed a cDNA library of grass carp spleen and identified the gene expression profile associated with GCRV infection, to enrich related pathways involved in immune responses against GCRV.

MATERIAL AND METHODS

Animals, viral infections and grouping

Healthy grass carp fingerlings with intact skin and normal feeding behavior were obtained from the Fisheries Research Institute of Hunan Province (Yuanjiang, Yiyang, China) fish-breeding center in July 2014, and cultured in the aquaculture base of Hunan Agricultural University. Fish (N = 200) weighing 100 to 150 g were acclimatized for 2 weeks prior to the commencement of experiments, to laboratory conditions of 28°C water temperature and twice-a-day feeding. In a preliminary dose range study, 30 grass carp were randomly selected, divided into three groups and given either 1×10^7, 3×10^7, or 6×10^7 50% tissue culture infective doses per mL ($TCID_{50}$/mL) GCRV873 standard strain (obtained from the Institute of Hydrobiology, Chinese Academy of Sciences) via intraperitoneal injection. Based on results from this study, 1×10^7 $TCID_{50}$/mL was selected as a suitable viral dose for infection. Sixty grass carp were then selected from the remaining fish and randomly allocated to control and infected groups (each N = 30). The control group was intraperitoneally injected with 0.6% fish physiological saline, and the infected group with 0.5 mL GCRV873 standard strain with a viral dose of 1×10^7 $TCID_{50}$/mL. Water was maintained at 28°C and intermittently oxygenated using an oxygenation stick.

Pathological features of grass carp infected with GCRV

Following inoculation of fish with either saline (controls) or GCRV (infection group), three fish at each of the time points 0, 24, 48, 72, 96, and 138 h were selected and anesthetized with tricaine methanesulfonate (MS222, Hunan Tiancheng Macromolecular Material Co., Ltd., Shaoyang, Hunan, China). Spleen tissues were collected and stored in liquid nitrogen until further use. Fish in the control group showed intact epithelial surface at each time point. Fish in the infected group were normal at 24 h, but at 48 h, were observed to swim alone; had visible red spots in the operculum, fin base, abdomen, and other parts; and had prominent, congested eyes. One fish that died at 48 h had punctate erythema in the intestine, and light red fluid inside the abdominal cavity upon anatomic examination. At 72 h, infected fish had deep body coloration and were powerless and anorectic; two dead fish were found at this time point. Upon dissection of these animals, dark red fluid was released out of abdominal cavity with some force and fish had swelling of the anus; dark red milt muscle; enlargement of the liver and pancreas; and congestion and swelling of the spleen. At 96 h, three more fish died, and the symptoms were as described for 72 h. By 138 h, the rate of death decreased; only one dead fish was found. Redness of fin strip base was observed, which gradually returned to normal.

Isolation of spleen total RNA

Fish sampled at the 72 h time point were selected for transcriptome analysis, due

to the clear presentation of grass carp hemorrhagic symptoms by this time. Total RNA was extracted from spleen tissue (100 mg/fish) using Trizol (Invitrogen, Carlsbad, CA, USA). The detailed procedure is as follows: a mortar was baked at 180°C for 3 to 4 h, cooled to room temperature, flash-cooled with liquid nitrogen, then used to grind 100 mg spleen to powder in liquid nitrogen. Trizol reagent (1 mL/100 mg tissue) was added to fully homogenize spleen tissue samples, which were then transferred to 1.5 mL diethylpyrocarbonate (DEPC)-treated centrifuge tubes and centrifuged at 12,000 rpm and 4°C for 10 min. Supernatant was transferred to fresh tubes, avoiding the precipitate, and mixed thoroughly with 0.2 mL chloroform/mL Trizol, inverted for 15 s. Phases were allowed to separate at room temperature for 8 min, and centrifuged at 12,000 rpm and 4°C for 15 min. The aqueous supernatant layer was carefully transferred to a fresh centrifuge tube and an equal volume of isopropanol added and mixed. Phases were allowed to separate at room temperature for 10 min before centrifuging at 2000 rpm and 4°C for 10 min. Supernatant was collected, 75% ethanol (v/v) added to precipitate total RNA, and samples centrifuged at 7500 rpm and 4°C for 5 min. Ethanol was discarded and precipitate, containing total RNA, was dried at room temperature for 30 min. Total RNA was dissolved at 55°C in RNase-free DEPC-treated water (Roche Diagnostics, Milan, Italy).

RNA purity and concentration detection

Concentration and purity of total RNA were measured by ultraviolet spectroscopy using the following formula: RNA = OD_{260} x D x 40 ng/μL (D), where D is the dilution multiple.

RNA purity was determined by OD_{260}/OD_{280} and the content and integrity of total RNA verified by electrophoresis on a 0.7% agarose (w/v) gel. The OD_{260}/OD_{280} of extracted RNA was 2; RNA 28S:18S was 8S:18; and RNA integrity number (RIN) was >8, as detected by Agilent Bioanalyzer. This indicated pure RNA with very high integrity without degradation, and met the sequencing criteria. Extracted RNA samples were shipped on dry ice to Beijing Genomics Institute (BGI) Company (Shenzhen, China) for transcriptome sequencing.

cDNA library construction and sequencing

mRNA was enriched using oligo (dT)-attached magnetic beads (Search Biotech Co., Ltd., Beijing, China), and purified mRNAs were fragmented into short segments using fragmentation buffer (Shanghai Bioleaf Biotech Co., Ltd., Shanghai, China). Fragmented mRNAs were copied into first-strand cDNA using reverse transcriptase and random primers (Takara Co., Dalian, China). Next, buffer (Nanjing Senbeijia Biological Co., Ltd., Nanjing, China), dNTPs (SBS Genetech Co., Ltd., Beijing, China), RNase H (BioChiron, Shenzhen, China), and DNA polymerase I (Takara Co.) were used to synthesize second-strand cDNA. cDNAs were purified using a QiaQuick PCR kit (Wuhan Kehaojia Biological Technology Co., Ltd.) and eluted with ethidium bromide (EB) buffer (Nanjing Senbeijia Biological Co., Ltd.). The ends of the cDNAs were repaired by adding a single 'A' base (Shanghai Baoman Biological Technology Co., Ltd., Shanghai, China) and ligated to Illumina adapters. cDNA fragments of 200 bp were purified by agarose gel electrophoresis, and used for further polymerase chain reaction (PCR) amplification to construct a fragmented cDNA library. Illumina HiSeq™ 2000 (Shanghai Sangon Biotech Co., Ltd., Shanghai, China) was used to perform transcriptome sequencing. The flow chart of cDNA library construction is shown in Figure 1.
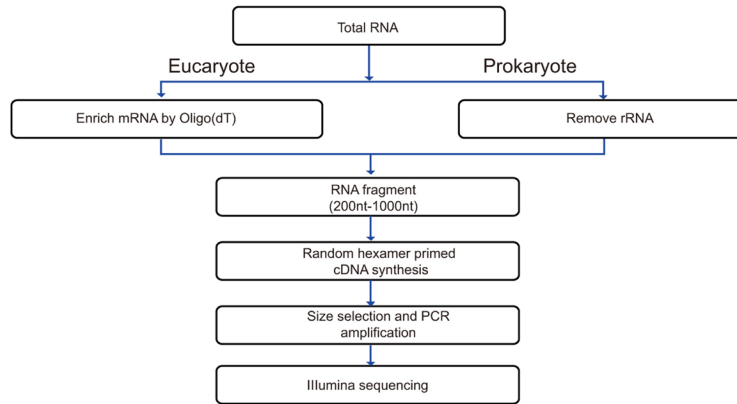
Figure 1. Flow chart of cDNA synthesis and library preparation procedures.

Transcriptome sequencing data analysis and *de novo* assembly

Raw reads were obtained by translating original image data into sequence data by base calling. Adaptor sequences duplicated or low-quality reads were removed to eliminate the potential effects of these reads on *de novo* assembly, and clean reads were achieved. Processing steps were as follows: raw reads of adaptors were removed; reads with unknown nucleotide N greater than 5% of the raw reads were removed; and low-quality raw reads were removed (number of bases with $Q \leq 5$ accounted for more than 50% of the entire read). Overlapping reads of certain length were linked into longer segments and compared with contigs by SOAP *de novo*. Paired-end reads were used to determine contigs from the same transcript and the distance between contigs. SOAP *de novo* was used to link these contigs together, and unknown sequences between contigs were represented by N to achieve scaffold. Paired-end reads were further used to shorten the scaffold, by decreasing N sequences. A unigene whose both ends could not be extended was achieved (Figure 2).
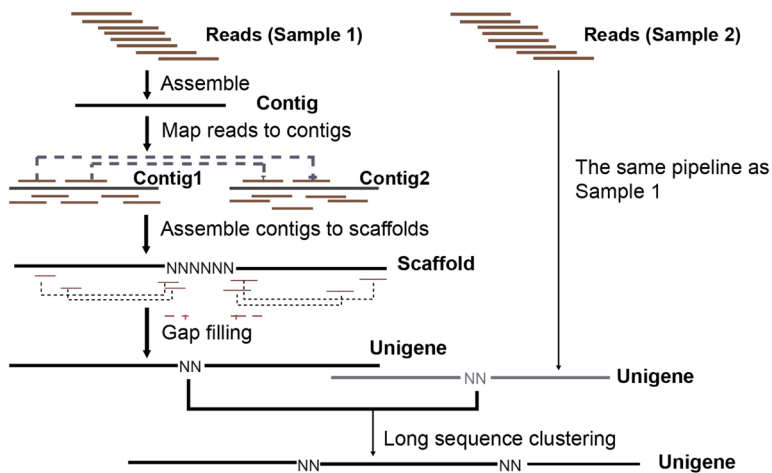

Figure 2. Grass carp transcriptome sequencing and assembly.

Functional annotation

BLASTx was performed to align assembled Unigene sequence to NCBI non-redundant (Nr) protein database, Swiss-Prot database, Kyoto Encyclopedia of Genes and Genomes (KEGG), and Cluster of Orthologous Groups of proteins (COG) databases with the threshold set to E-value <10^{-5} for functional annotation. The unigene name was determined based on the top BLASTx hit with the highest score. Gene ontology (GO) and KEGG pathway analyses were used to determine differences in gene expression between the control and infection group. GO analysis was used to compare differentially expressed genes with all genes in the GO database (http://www.geneontology.org/), and determine the main biological function of our studied genes. Base on the Nr annotation information, the Blast2GO software was used to achieve the GO annotation information of unigenes, and the WEGO software was used to analyze unigene GO functional classification and understand the gene function distribution of the certain species (http://wego.genomics.org.cn/cgi-bin/wego/index.pl). The KEGG database was used to analyze the metabolic pathways of gene products in cells, and functions of these gene products. The most important biochemical pathway and signal transduction pathway of differentially expressed genes was determined by pathway-enrichment analysis.

Real-time (RT) PCR primer design

β-actin was used as a reference gene and profibrinolysin (*PLG*), coagulation factor II (*F2*), and complement factors V (*C5*), III (*C3*) and II (*C2*) cDNA selected from the transcriptome data for validation by RT-PCR. Primers were designed using the Primer Premier 5.0 software (PREMIER Biosoft International, USA) and synthesized by Shanghai Sangon Biotech Co., Ltd. Primer names and sequences for each unigene are shown in Table 1. Isolated RNA from spleen tissue 72 h after GCRV infection (N = 3 from each group) was used for first-strand cDNA synthesis, following PrimeScript Reverse Transcriptase kit instructions (TaKaRa BIO, Shiga, Japan).

Table 1. Primers for *PLG*, *C2*, *C3*, *C5*, and *F2* used in RT-PCR assays.

Unigene name	Primer name	Primer sequence (5'-3')	Amplification size (bp)
Unigene 69241	PLG Forward	AAGTCCCTGCCTGTCAATG	110
	PLG Reverse	GAACATCTGTGCTTTGGGTC	
Unigene 65489	F2 Forward	TGGGCAGACAGACAGGA	135
	F2 Reverse	CCCTAAATACAACTGGAAGGAA	
Unigene 72231	C5 Forward	CAAAGCCCATCCTGAACT	119
	C5 Reverse	CAAAGTAAGAATGCTGTCCTGT	
AY374472.1	C3 Forward	AAACCTGGAATGCCCTTC	150
	C3 Reverse	CTCCTTGAGTGTTGACGGT	
Unigene 72241	C2 Forward	GTGATAGGTCAGCACGCTT	107
	C2 Reverse	GGAGATTATGAGGAACCACAG	
HM045420	β-actin Forward	GCCGTGACCTGACTGACTA	103
	β-actin Reverse	TCAAGAGCCACATAGCAGAG	

Statistical analysis

The SPSS 20.0 statistical software (SPSS, Inc., Chicago, IL, USA) was used for data analysis and enumeration data are reported as ratio or rate. The chi-square test was conducted to analyze differences between two groups.

RESULTS

Statistics of grass carp transcriptome sequences

Clean reads from grass carp transcriptome sequencing were achieved after removing ineffective raw reads. The total number of reads was 2,722,223 from the control group and 2,751,111 from the infected group. Statistical analysis of grass carp transcriptome sequences are shown in Table 2.

Table 2. Statistics of grass carp transcriptome sequences.

| Group	Total reads	Total nucleotides	Q_{20} percentage	N percentage	Percentage of G/C
Infected	2,751,111	2,276,014	95.97%	0.02%	46.89%
Control	2,722,223	2,350,021	93.34%	0.89%	45.64%

Statistics of transcriptome assembly

SOAP *de novo* produced 65,120 contigs with an average length of 142 bp in the infected group, and 70,471 contigs with an average length of 124 bp in the control group (Table 3). At the next step of assembly, 14,015 scaffolds (average length, 361 bp; proportion larger than 2000 nt, 16.6%) in the infected group and 20,154 scaffolds (average length, 250 bp; proportion larger than 2000 nt, 0.09%) in the control group were produced (Figure 3). A total of 9631 unigenes were assembled in the infected group, and 13,150 in the control group. Unigenes [11,023 (71.60% of these were less than 500 nt in length)] were obtained in all the samples; of these, sequence direction was determined in 4784 unigenes, and 6239 unigenes had no clear sequence direction (Figure 4).

**Table 3.** Statistics of grass carp transcriptome assembly.

Group	Contig	75-100 nt	100-300 nt	300-500 nt	≥500 nt	N50	Mean length	Total
Infected	Number	49,784	10,732	2325	2279	105	142	65,120
	Percent	76.45%	16.48%	3.57%	3.50%			
Control	Number	51,550	14,122	2614	2185	90	124	70,471
	Percent	73.15%	20.04%	3.71%	3.10%			

nt = nucleotide.
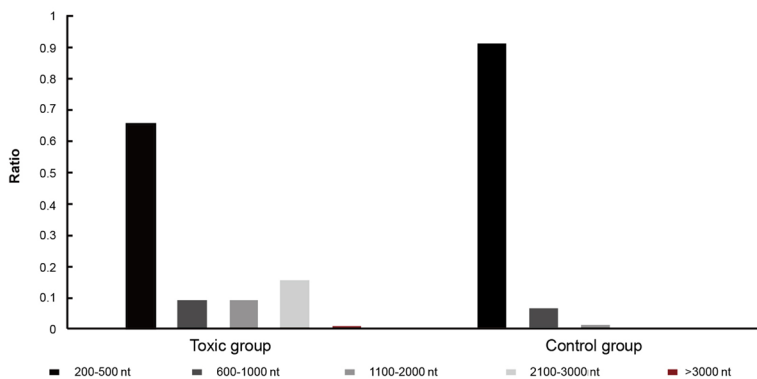
Figure 3. Length distribution of scaffolds from grass carp transcriptome assembly in infected and control groups.

1.59%   0.69%

7.63%

18.49%

71.60%

■ 300-500 nt ■ 600-1000 nt □ 1100-2000 nt □ 2100-3000 nt ■ >3000 nt
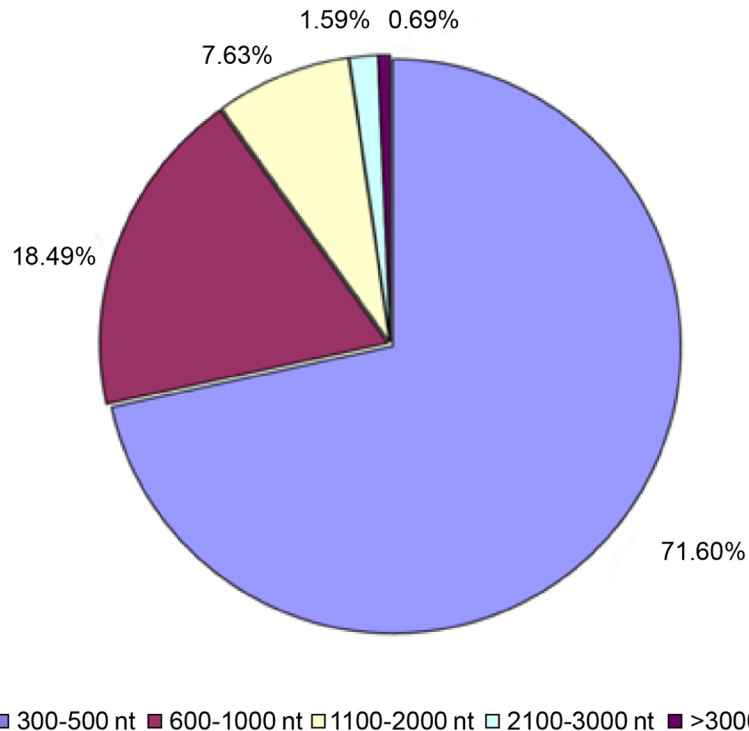
**Figure 4.** Length distribution of unigenes from grass carp transcriptome assembly.

Functional annotation of the transcriptome

After successfully aligning 11,023 assembled unigenes to the COG database, total sequenced unigenes were classified into 25 gene function categories, such as modified processing of RNA, signal transduction mechanisms, and defense mechanisms. A total of 1455 unigenes were categorized as immune signals, while no known functions could be assigned to 987 unigenes (Figure 5). The assembled unigenes were matched with three main annotations of the GO database - biological processes, cellular components and molecular functions - containing 47 functional groups. From such matching, we were able to assign 2886 unigenes into 23 subcategories of biological processes (including anatomical structure formation, growth, and development process); 4268 to 13 subcategories of cell components (such as envelope and synapse); and 3869 to 11 subcategories of molecular function (including antioxidant activity and auxiliary transport protein activity) (Figure 6). Among the 11,023 unigenes, 2770 participated in 217 corresponding pathways. Notably, 32 pathways related to immunity and disease resistance, including complement and coagulation cascade, phagosome, intestinal IgA, chemokine signaling, and T-cell receptor signaling pathways. The most unigenes in a single category (1390) were associated with metabolic pathways and accounted for 12.61% of total unigenes; the category containing the second most unigenes (556; 5.04%) was related to cancer pathways.
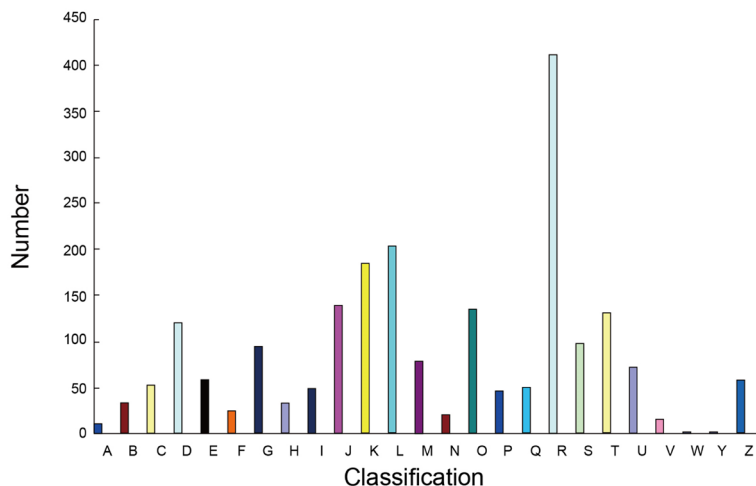
Figure 5. COG functional classification of grass carp unigenes. A, RNA processing and modification; B, chromatin structure and dynamics; C, energy production and conversion; D, cell cycle control, cell division, and chromosome partitioning; E, amino acid transport and metabolism; F, nucleotide transport and metabolism; G, carbohydrate transport and metabolism; H, coenzyme transport and metabolism; I, lipid transport and metabolism; J, translation, ribosomal structure, and biogenesis; K, transcription; L, replication, recombination, and repair; M, cell wall/membrane/envelope biogenesis; N, cell motility; O, posttranslational modification, protein turnover, and chaperone; P, inorganic ion transport and metabolism; Q, secondary metabolite biosynthesis, transport, and catabolism; R, general function prediction only; S, function unknown; T, signal transduction mechanisms; U: intracellular trafficking, secretion, and vesicular transport; V, defense mechanisms; W, extracellular structures; Y, nuclear structure; Z, cytoskeleton.
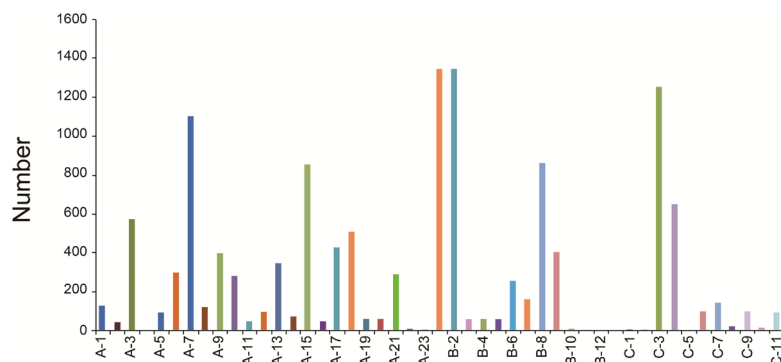
Figure 6. Gene ontology classification of grass carp unigenes. A, biological process; B, cellular component; C, molecular function. A-1, anatomical structure formation; A-2, biological adhesion; A-3, biological regulation; A-4, cell killing; A-5, cellular component biogenesis; A-6, cellular component organization; A-7, cellular process; A-8, death; A-9, developmental process; A-10, establishment of localization; A-11, growth; A-12, immune system process; A-13, localization; A-14, locomotion; A-15, metabolic process; A-16, multi-organism process; A-17, multicellular organismal process; A-18, pigmentation; A-19, reproduction; A-20, reproductive process; A-21, response to stimulus; A-22, rhythmic process; A-23, viral reproduction; B-1, cell; B-2, cell part; B-3, envelope; B-4, extracellular region; B-5, extracellular region part; B-6, macromolecular complex; B-7, membrane-enclosed lumen; B-8, organelle; B-9, organelle part; B-10, synapse; B-11, synapse part; C-1, antioxidant activity; C-2, auxiliary transport protein activity; C-3, binding; C-4, catalytic activity; C-5, electron carrier activity; C-6, enzyme regulator activity; C-7, molecular transducer activity; C-8, structural molecule activity; C-9, transcription regulator activity; C-10, translation regulator activity; C-11, transporter activity.

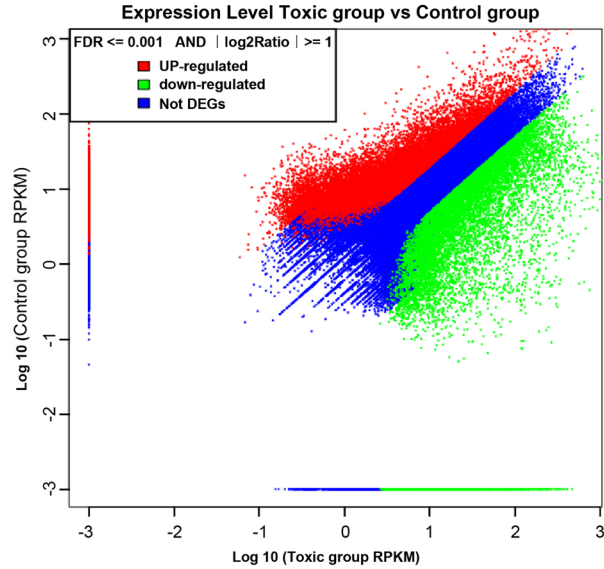## Differentially expressed gene screening and GO classification

Based on the method for screening of differentially expressed genes in two samples proposed by Audic and Claverie (1997), P value <0.05, FDR < 0.001 and expressed differences larger than 2-fold ($|\log2\text{Ratio}| \geq 1$) were considered as cutoffs for differentially expressed genes (Table 4). Among the 11,023 unigenes, 10,021 showed differential expression levels and 9145 were significantly differentially expressed, including 2083 significantly up-regulated and 7062 significantly down-regulated unigenes (Figure 7). Gene ontology classification was performed to analyze differentially expressed unigenes. Among 9145 unigenes, 7356 participated in 23 subcategories of biological processes; 1049 were found in 11 subcategories of cell component; and 751 belonged to 11 subcategories of molecular function (Figure 8).

Table 4. Gene ontology classification of significantly differentially expressed grass carp unigenes.

Class	Gene ontology term	Cluster frequency (total genes)	Corrected P value
Process ontology	Response to wounding	117 (4158), 2.8%	3.69E-12
	Response to stimulus	1055 (4158), 25.4%	2.16E-07
	Response to external stimulus	283 (4158), 6.8%	4.54E-07
	Regulation of biological quality	543 (4158), 13.1%	1.74E-06
	Response to stress	528 (4158), 12.7%	5.55E-05
	Defense response	165 (4158), 4.0%	3.10E-04
	Wound healing	29 (4158), 0.7%	5.40E-04
	Immune system process	381 (4158), 9.2%	5.80E-04
	Homeostatic process	292 (4158), 7.0%	5.80E-04
	Hemostasis	50 (4158), 1.2%	2.13E-03
	Cellular amino acid derivative metabolic process	71 (4158), 1.7%	2.59E-03
	Coagulation	26 (4158), 0.6%	2.86E-03
	Negative regulation of biological process	404 (4158), 9.7%	2.98E-03
	Inflammatory response	51 (4158), 1.2%	4.46E-03
	Blood coagulation	22 (4158), 0.5%	6.21E-03
	Regulation of body fluid levels	70 (4158), 1.7%	1.56E-02
	Response to other organism	120 (4158), 2.9%	1.91E-02
	Regulation of blood coagulation	12 (4158), 0.3%	1.94E-02
	Activation of plasma proteins involved in acute inflammatory response	16 (4158), 0.4%	2.38E-02
	Complement activation	16 (4158), 0.4%	2.38E-02
	Protein maturation by peptide bond cleavage	23 (4158), 0.6%	3.28E-02
Component ontology	Extracellular region	282 (4218), 6.7%	3.64E-13
	Extracellular region part	257 (4218), 6.1%	2.07E-12
	Cell surface	57 (4218), 1.4%	9.45E-09
Function ontology	Peptidase activity	221 (4567), 4.8%	8.60E-04
	Endopeptidase activity	128 (4567), 2.8%	1.22E-03
	Peptidase activity, acting on L-amino acid peptides	208 (4567), 4.6%	2.47E-03
	Aldehyde-lyase activity	10 (4567), 0.2%	4.92E-03
	Peptidase inhibitor activity	37 (4567), 0.8%	4.52E-02

Pathway-enrichment analysis of differentially expressed genes

Pathway-enrichment analysis was performed using the KEGG database, and a Q-value <0.05 was considered as significant enrichment. Of the 9145 significantly differentially expressed unigenes, there were 916 in the database, located in 217 pathways. Among the pathways, 32 were significantly enriched, including 4 closely related to coagulation and immunity (the most prominent were complement and coagulation cascades, hematopoietic cell lineage, phagosome, and cytokine-cytokine receptor interaction pathways).

Figure 7. Significantly differentially expressed unigenes in infected and control groups.
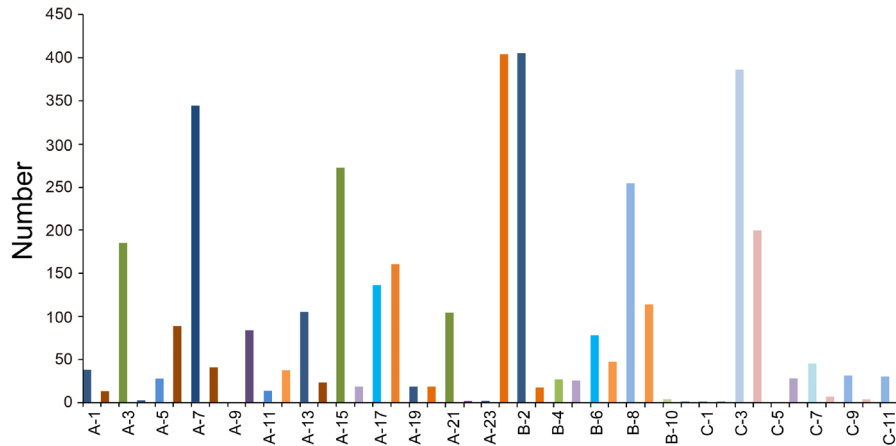
Figure 8. Gene ontology classification of significantly differentially expressed grass carp unigenes. A, biological process; B, cellular component; C, molecular function. A-1, anatomical structure formation; A-2, biological adhesion; A-3, biological regulation; A-4, cell killing; A-5, cellular component biogenesis; A-6, cellular component organization; A-7, cellular process; A-8, death; A-9, developmental process; A-10, establishment of localization; A-11, growth; A-12, immune system process; A-13, localization; A-14, locomotion; A-15, metabolic process; A-16, multi-organism process; A-17, multicellular organismal process; A-18, pigmentation; A-19, reproduction; A-20, reproductive process; A-21, response to stimulus; A-22, rhythmic process; A-23, viral reproduction; B-1, cell; B-2, cell part; B-3, envelope; B-4, extracellular region; B-5, extracellular region part; B-6, macromolecular complex; B-7, membrane-enclosed lumen; B-8, organelle; B-9, organelle part; B-10, synapse; B-11, synapse part; C-1, antioxidant activity; C-2, auxiliary transport protein activity; C-3, binding; C-4, catalytic activity; C-5, electron carrier activity; C-6, enzyme regulator activity; C-7, molecular transducer activity; C-8, structural molecule activity; C-9, transcription regulator activity; C-10, translation regulator activity; C-11, transporter activity.

The complement and coagulation cascade pathway showed the largest difference, and contained 135 differentially expressed unigenes, accounting for 1.47% of 916 located pathways (Table 5). Up-regulated unigenes were: von Willebrand factor (*VWF*), coagulation factor D (*DF*), decay accelerating factor (*DAF*), transforming growth factor act (*TGFp3*), chemokine receptor 4 (*CCR4*), chemokine XC receptor 1 (*XCR1*), nerve growth factor receptor (*NGFR*), unknown factor (*X*), q subcomponent of the first complement component (*C1q*). Down-regulated unigenes in the corresponding pathway were: coagulation factors II (*F2*), V (*F5*), and VIII (*F8*); serpin peptidase inhibitors clade D member 1 (*SERPIND1*), clade C member 1 (*SERPINC1*), and clade A member 5 (*SERPINA5*); carboxypeptidase B2 (*CPB2*); alpha-2-macroglobulin (*A2M*); protein C (*PROC*); urokinase-type plasminogen activator (*u-PA*); urokinase-type plasminogen activator receptor (*u-PAR*); profibrinolysin (*PLG*); kininogen (*KNG*); *C3*; interferon (*IF*); mannan-binding lectin (*MBL*); mannose-binding lectin-associated serine protease 1/2 (*MASP1/2*); *C2*; complement factors IV (*C4*), VII (*C7*), VIII (*C8*), and IX (*C9*); *C5*; a subunit of first protein to react in complement system (*C1S*); serpin peptidase inhibitor clade G member 1 (*SERPING1*); and complement factor V receptor 1 (*C5R1*) (Table 6). Five unigenes (*C2*, *C3*, *C5*, *PLG*, and *F2*), belonging to the complement and coagulation cascade pathway, were chosen for RT-PCR to verify results; all five genes were down-regulated, consistent with results obtained from the KEGG database (Figure 9).

Table 5. KEGG pathway-enrichment analysis of significantly expressed unigenes of grass carp.

Pathway	Differentially expressed genes with pathway annotation (% of 916 total)	All genes with pathway annotation (% of 8229 total)	Q-value	Pathway ID
Complement and coagulation cascades	98 (1.07%)	554 (6.07%)	1.24E-34	ko04610
Hematopoietic cell lineage	84 (0.92%)	541 (5.92%)	1.15E-14	ko04640
Phagosome	28 (0.30%)	393 (4.30%)	1.14E-11	ko04145
Cytokine-cytokine receptor interaction	98 (1.07%)	583 (6.37%)	9.25E-10	ko04060
PPAR signaling pathway	34 (0.37%)	400 (4.37%)	9.60E-09	ko03320
Glycolysis / Gluconeogenesis	41 (0.45%)	412 (4.50%)	1.16E-08	ko00010
Systemic lupus erythematosus	51 (0.56%)	234 (2.56%)	1.28E-06	ko05322
Collecting duct acid secretion	49 (0.54%)	168 (1.84%)	7.40E-05	ko04966
Antigen processing and presentation	92 (1.01%)	275 (3.01%)	1.13 E-04	ko04612
Cell adhesion molecules (CAMs)	4 (0.04%)	369 (4.04%)	2.15E-04	ko04514
Glycerolipid metabolism	17 (0.19%)	492 (5.38%)	5.99E-04	ko00561
Steroid biosynthesis	28 (0.31%)	398 (4.35%)	1.46E-03	ko00100
ECM-receptor interaction	96 (1.05%)	187 (2.05%)	1.45E-03	ko04512
Prion diseases	11 (0. 12%)	468 (5.12%)	1.14E-03	ko05020
Arginine and proline metabolism	96 (1.05%)	553 (6.05%)	1.47E-03	ko00330
Maturity onset diabetes of the young	32 (0.35%)	492 (5.38%)	1.47E-03	ko04950
Malaria	4 (0.04%)	278 (3.04%)	2.35E-03	ko05144
Chagas disease	10 (0.11%)	200(2.19%)	2.47 E-03	ko05142
Autoimmune thyroid disease	14 (0.15%)	380 (4.15%)	2.48 E-03	ko05320
Graft-versus-host disease	16 (0.17%)	473 (5.17%)	3.68 E-03	ko05332
Pentose phosphate pathway	13 (0.14%)	379 (4.14%)	4.14 E-03	ko00030

Table 6. Specific differentially expressed unigenes.

Expression	Gene
Up-regulation	*VWF*, *DF*, *DAF*, *TGFp3*, *CCR4*, *XCR1*, *NGFR*, *X*, *DF*, *C1q*
Down-regulation	*F2*, *F5*, *F8*, *SERPIND1*, *CPB2*, *SERPINC1*, *SERPINA5*, *A2M*, *PROC*, *u-PA*, *u-PAR*, *PLG*, *KNG*, *C3*, *IF*, *MBL*, *MASP1/2*, *C2*, *C4*, *C5*, *C7*, *C8*, *C9*, *C1S*, *SERPING1*, *C5R1*
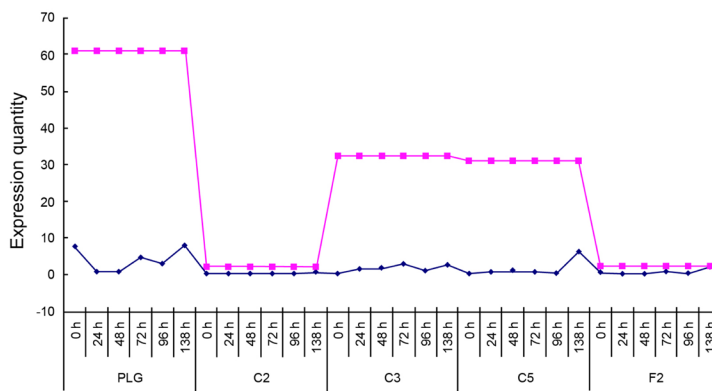
Figure 9. Comparison of expressions of selected unigenes by RT-PCR in infected and control groups.

DISCUSSION

In order to systematically address GCRV in the mechanisms of grass carp hemorrhage during infection, transcriptome analysis of differentially expressed genes in spleens of GCRV-infected groups were compared to uninfected controls, by solexa sequencing. Data obtained from this study are expected to immensely enrich the available genomic data, and help future studies aimed at identifying novel genes and pathways related to grass carp hemorrhagic disease.

Traditional sequencing methods are based on generation of cDNA libraries and screening a large volume of expressed sequence tags (ESTs), to identify differentially expressed genes, a process that is time-consuming and requires high-quality cDNA libraries (Wall et al., 2009). Next-generation DNA sequencing has the potential to accelerate biomedical research by rapid and comprehensive analysis of genomes, transcriptomes and interactomes without the requirement for construction of cDNA libraries and production-scale efforts (Shendure and Ji, 2008). Our results showed that unigenes detected in both infected group and control group participate in biological processes, cell components and molecular function. Chen et al. (2012) also found that annotated unigenes of head kidney in grass carp related to cellular process (25.5%), metabolic process (19.1%) and biological regulation (10.8%). In the current study, KEGG analysis showed related pathways in immunity and disease resistance, including complement and coagulation cascade, phagosome, intestinal IgA, chemokine signaling, and T-cell receptor signaling pathways.

Grass carp hemorrhagic disease can be divided into three categories based on the anatomic location of lesions (muscles, gills and enteritis). Muscle hemorrhagic disease is the most common and is characterized by a lack of obvious hyperemia; dark body color; transparent body under sunlight; and red coloration of the white muscles (Shi et al., 2014). GO classification analysis showed that most differentially expressed unigenes participated in biological processes including anatomical structure formation, growth, and development process. KEGG analysis demonstrated 32 significantly enriched pathways, including 4 closely related to coagulation and immunity, i.e., complement and coagulation cascades, hematopoietic cell lineage, phagosome, and cytokine-cytokine receptor interaction pathways. Of the unigenes involved in these four pathways, *VWF*, *DF*, *DAF*, *TGFp3*, *CCR4*, *XCR1*,

NGFR, *X*, *DF*, *C1q* were up-regulated, while *F2*, *F5*, *F8*, *SERPIND1*, *CPB2*, *SERPINC1*, *SERPINA5*, *A2M*, *PROC*, *u-PA*, *u-PAR*, *PLG*, *KNG*, *C3*, *IF*, *MBL*, *MASP1/2*, *C2*, *C4*, *C5*, *C7*, *C8*, *C9*, *C1S*, *SERPING1*, and *C5R1* were down-regulated. *C2*, *C3*, *C5*, *PLG*, and *F2* belonged to the complement and coagulation cascade pathway and our RT-PCR analysis also confirmed their down-regulation. The complement system is an ancient mechanism to fight against pathogens and is found in a wide range of animals, from higher mammals to lower fishes (Wang et al., 2014). In fish, the complement system appears to be fully equipped with the three C3-activation pathways and the cytolytic pathway, showing many of the typical effector activities, such as target cell killing, opsonization, and anaphylatoxic leukocyte stimulation (Nakao et al., 2011). FI to FXIII play important roles in tissue coagulation, fibrinolysis, and anticoagulation (Campo et al., 2013; Hoppe, 2014). C5 is a key factor linking the complement system and coagulation cascade (Ehrnthaller et al., 2011; Kolev et al., 2014). Chemokines induce leukocyte chemo-attraction and directly regulate T-cell development, priming, and effector functions (Franciszkiewicz et al., 2012). Cytokines play an important role in the fish innate immune system, similar to their mammalian counterparts (Sun et al., 2013; Shin et al., 2014). From our analysis, we propose that the complement and coagulation cascade pathway is a specific defense mechanism activated in response to GCRV infections in grass carps, and that cytokine-cytokine receptor interaction might promote this pathway.

In conclusion, splenic gene expression profiles were compared between GCRV-infected and uninfected green carp and we identified 9145 unigenes and 32 significantly enriched pathways that were differentially expressed between the groups. Results of the transcriptome analysis, combined with RT-PCR validation, identified the complement and coagulation cascade pathways as relevant pathways in defense against GCRV infection. In future studies, we hope to further explore these data in order to better understand precise interactions between different pathways and their connections with grass carp hemorrhagic disease.

Conflicts of interest

All authors declare no conflict of interest.

ACKNOWLEDGMENTS

REFERENCES

Audic S and Claverie JM (1997). The significance of digital gene expression profiles. *Genome Res.* 7: 986-995.
Campo G, Pavasini R, Pollina A, Tebaldi M, et al. (2013). Coagulation factors and recurrence of ischemic and bleeding adverse events in patients with acute coronary syndromes. *Thromb. Res.* 132: 151-157. http://dx.doi.org/10.1016/j. thromres.2013.06.007
Chen J, Li C, Huang R, Du F, et al. (2012). Transcriptome analysis of head kidney in grass carp and discovery of immune-related genes. *BMC Vet. Res.* 8: 108. http://dx.doi.org/10.1186/1746-6148-8-108

Ehrnthaller C, Ignatius A, Gebhard F and Huber-Lang M (2011). New insights of an old defense system: structure, function, and clinical relevance of the complement system. *Mol. Med.* 17: 317-329. http://dx.doi.org/10.2119/molmed.2010.00149

Endo A, Tatematsu K, Hanada K, Duermeyer L, et al. (2012). Tissue-specific transcriptome analysis reveals cell wall metabolism, flavonol biosynthesis and defense responses are activated in the endosperm of germinating *Arabidopsis thaliana* seeds. *Plant Cell Physiol.* 53: 16-27. http://dx.doi.org/10.1093/pcp/pcr171

Franciszkiewicz K, Boissonnas A, Boutet M, Combadière C, et al. (2012). Role of chemokines and chemokine receptors in shaping the effector phase of the antitumor immune response. *Cancer Res.* 72: 6325-6332. http://dx.doi.org/10.1158/0008-5472.CAN-12-2027

Hoppe B (2014). Fibrinogen and factor XIII at the intersection of coagulation, fibrinolysis and inflammation. *Thromb. Haemost.* 112: 649-658. http://dx.doi.org/10.1160/TH14-01-0085

Huang T, Su J, Heng J, Dong J, et al. (2010). Identification and expression profiling analysis of grass carp *Ctenopharyngodon idella* LGP2 cDNA. *Fish Shellfish Immunol.* 29: 349-355. http://dx.doi.org/10.1016/j.fsi.2010.04.001

Jia R, Cao LP, Du JL, Liu YJ, et al. (2014). Grass carp reovirus induces apoptosis and oxidative stress in grass carp (Ctenopharyngodon idellus) kidney cell line. *Virus Res.* 185: 77-81. http://dx.doi.org/10.1016/j.virusres.2014.03.021

Kolev M, Le Friec G and Kemper C (2014). Complement--tapping into new sites and effector systems. *Nat. Rev. Immunol.* 14: 811-820. http://dx.doi.org/10.1038/nri3761

Liang HR, Li YG, Zeng WW, Wang YY, et al. (2014). Pathogenicity and tissue distribution of grass carp reovirus after intraperitoneal administration. *Virol. J.* 11: 178. http://dx.doi.org/10.1186/1743-422X-11-178

Liu Q, Xu B, Xiao T, Su J, et al. (2013). Molecular cloning, characterization and expression analysis of coagulation factor VII gene in grass carp (*Ctenopharyngodon idella*). *Fish Shellfish Immunol.* 35: 618-622. http://dx.doi.org/10.1016/j.fsi.2013.05.015

Meyer E, Aglyamova GV, Wang S, Buchanan-Carter J, et al. (2009). Sequencing and de novo analysis of a coral larval transcriptome using 454 GSFlx. *BMC Genomics* 10: 219. http://dx.doi.org/10.1186/1471-2164-10-219

Nakao M, Tsujikura M, Ichiki S, Vo TK, et al. (2011). The complement system in teleost fish: progress of post-homolog-hunting researches. *Dev. Comp. Immunol.* 35: 1296-1308. http://dx.doi.org/10.1016/j.dci.2011.03.003

Parchman TL, Geist KS, Grahnen JA, Benkman CW, et al. (2010). Transcriptome sequencing in an ecologically important tree species: assembly, annotation, and marker discovery. *BMC Genomics* 11: 180. http://dx.doi.org/10.1186/1471-2164-11-180

Shendure J and Ji H (2008). Next-generation DNA sequencing. *Nat. Biotechnol.* 26: 1135-1145. http://dx.doi.org/10.1038/nbt1486

Shi M, Huang R, Du F, Pei Y, et al. (2014). RNA-seq profiles from grass carp tissues after reovirus (GCRV) infection based on singular and modular enrichment analyses. *Mol. Immunol.* 61: 44-53. http://dx.doi.org/10.1016/j.molimm.2014.05.004

Shin GW, White SL, Dahms HU, Jeong HD, et al. (2014). Disease resistance and immune-relevant gene expression in golden mandarin fish, *Siniperca scherzeri Steindachner*, infected with infectious spleen and kidney necrosis virus-like agent. *J. Fish Dis.* 37: 1041-1054. http://dx.doi.org/10.1111/jfd.12182

Su J, Heng J, Huang T, Peng L, et al. (2012). Identification, mRNA expression and genomic structure of TLR22 and its association with GCRV susceptibility/resistance in grass carp (*Ctenopharyngodon idella*). *Dev. Comp. Immunol.* 36: 450-462. http://dx.doi.org/10.1016/j.dci.2011.08.015

Sun C, Liu Y, Hu Y, Fan Q, et al. (2013). Gig1 and Gig2 homologs (CiGig1 and CiGig2) from grass carp (*Ctenopharyngodon idella*) display good antiviral activities in an IFN-independent pathway. *Dev. Comp. Immunol.* 41: 477-483. http://dx.doi.org/10.1016/j.dci.2013.07.007

Wall PK, Leebens-Mack J, Chanderbali AS, Barakat A, et al. (2009). Comparison of next generation sequencing technologies for transcriptome characterization. *BMC Genomics* 10: 347. http://dx.doi.org/10.1186/1471-2164-10-347

Wang Q, Zeng W, Liu C, Zhang C, et al. (2012). Complete genome sequence of a reovirus isolated from grass carp, indicating different genotypes of GCRV in China. *J. Virol.* 86: 12466. http://dx.doi.org/10.1128/JVI.02333-12

Wang YD, Huang SJ, Chou HN, Liao WL, et al. (2014). Transcriptome analysis of the effect of *Vibrio alginolyticus* infection on the innate immunity-related complement pathway in *Epinephelus coioides*. *BMC Genomics* 15: 1102. http://dx.doi.org/10.1186/1471-2164-15-1102

Wen Y, Shao JZ, Pan XX and Xiang LX (2005). Molecular cloning, characterization and expression analysis of QM gene from grass carp (Ctenopharyngodon idellus) homologous to Wilms' tumor suppressor. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 141: 356-365. http://dx.doi.org/10.1016/j.cbpc.2005.04.007

Ye X, Tian YY, Deng GC, Chi YY, et al. (2012). Complete genomic sequence of a reovirus isolated from grass carp in China. *Virus Res.* 163: 275-283. http://dx.doi.org/10.1016/j.virusres.2011.10.014

Zeng W, Wang Y, Liang H, Liu C, et al. (2014). A one-step duplex rRT-PCR assay for the simultaneous detection of grass carp reovirus genotypes I and II. *J. Virol. Methods* 210C: 32-35. http://dx.doi.org/10.1016/j.jviromet.2014.08.024

Zhang QL, Yan Y, Shen JY, Hao GJ, et al. (2013). Development of a reverse transcription loop-mediated isothermal amplification assay for rapid detection of grass carp reovirus. *J. Virol. Methods* 187: 384-389. http://dx.doi.org/10.1016/j.jviromet.2012.11.005