d'Collection

A THESIS
FOR THE DEGREE OF MASTER OF SCIENCE

GENOMIC IDENTIFICATION AND MOLECULAR CHARACTERIZATION OF TWO INTERFERON REGULATORY FACTORS AND THREE COMPLEMENT COMPONENT GENES FROM ROCK BREAM, *Oplegnathus fasciatus*

S.D.N.K. Bathige

Department of Marine Life Sciences
GRADUATE SCHOOL
JEJU NATIONAL UNIVERSITY

2013.02

Genomic Identification and Molecular Characterization of
Two Interferon Regulatory Factors
and
Three Complement Component Genes
From Rock bream, *Oplegnathus fasciatus*

S.D.N.K. Bathige
(Supervised by Professor Jehee Lee)

A thesis submitted in partial fulfillment of the requirement for the degree of

MASTER OF SCIENCE
2013. 02

This thesis has been examined and approved by

....

Thesis Director, In-Kyu Yeo, Professor of Marine Life Sciences

....

Joon Bum Jeong, Professor of Marine Life Sciences

....

Jehee Lee, Professor of Marine Life Sciences

20.02.2013
Date

Department of Marine Life Sciences
GRADUATE SCHOOL
JEJU NATIONAL UNIVERSITY
REPUBLIC OF KOREA

~ II ~

CONTENTS

한글 요약문

돌돔(*Oplegnathus fasciatus*)은 한국 양식산업에서 주목 받고 있는 주요 양식어종이다.

그러나 해가 거듭될 수록 돌돔에게 여러 질병이 발병되면서 양식생산에 있어 악영향을 주고 있다.

돌돔에서 확인되는 질병 중에는 *Streptococcos iniae*가 원인인 연쇄상구균증(Streptococcosis)과

*Edwardsiella tarda*가 원인인 에드워드증(Edwardsiellosis)과 같은 세균성질병과 이리도

바이러스(Iridovirus)로 인한 바이러스성 질병 등이 있다. 이런 질병은 단독적으로 발병되거나 여러

질병이 복합적으로 발병되어 돌돔 양식에 큰 경제적인 손실을 주고 있는 실정이다.

따라서 양식산업에서 빈번히 발생되는 질병을 예방하고 발병한 질병을 빠르게 대처할 수

있어야만 양식어종을 안정적으로 생산할 수 있다. 이를 위해 감염성 질병에 대한 어류의

방어체계를 이해하는 것이 기본적으로 필요하다. 특히 병원체에 대한 1차 방어체계인 선천면역을

연구해야만 한다. 그리하여 이 논문에서는 감염에 따른 선천면역체계를 분자유전학적인 수준에서

이해하고자 interferon regulatory factor 4 (IRF4) 와 8 (IRF8) 및 3종류의 complement 1q (C1q)

유전자를 동정하였다.

IRF4와 IRF8은 B 림프구를 자극하여 림프구 발달과 항체(immunoglobulin; Ig)의

일부분인 경쇄(light chain)의 전사를 조절하는 분자들의 활성을 촉진시킨다. 이 두 개의 IRF는

매우 밀접하게 연관되어 있으며 면역세포 중 림프 및 골수세포에서 높게 발현된다고 알려져 있다.

이 논문에서는 돌돔에서 발현되는 IRF4 (RbIRF4)와 IRF8 (RbIRF8)의 cDNA 서열과 유전체 서열을

돌돔 cDNA library와 BAC library에서 각각 스크리닝하여 동정하였다. RbIRF4의 cDNA 전체서열은

3442 염기쌍(bp)로 462개의 아미노산서열을 암호화하고 있다. 또한 유전체 서열의 크기는 9262

bp로 8개의 exon과 7개의 intron으로 구성되었다. RbIRF8인 경우 cDNA 전체서열은 422개의

아미노산을 암호화하고 있는 2186 bp로 이루어졌으며 유전체 서열의 크기는 4120 bp로 9개의

exon과 8개의 intron으로 구성되어 있다. RbIRF4와 RbIRF8의 아미노산 서열에서는 공통적으로

tryptophan pentad-repeat가 포함된 DNA-binding domain (DBD)과 IRF-association domain

(IAD)가 확인되었다. RbIRF4와 RbIRF8 유전자의 5′ flanking region에는 면역유전자의 전사를

조절하는 전사조절인자들의 결합부위가 존재하였다. real time PCR를 이용하여 RbIRF4와 RbIRF8

유전자의 발현양상을 확인한 결과 건강한 돌돔에서는 림프골수계 조직에서 가장 많이

발현되었다. 이리도바이러스와 *E. tarda*, *S. iniae*를 인위 감염시킨 돌돔에서는 두신(head

kidney)과 비장(spleen)에서 높은 발현양상을 보였다. 이러한 결과를 통해 돌돔에서 발현되는

RbIRF4와 RbIRF8도 바이러스와 병원균에 대한 보호기능을 담당하고 덧붙여 IRF 분자들이

경골어류에서는 면역조절 분자로서의 기능을 하는 것이라 유추 할 수 있다.

보체인자 1q(C1q)는 보체 C1 복합체를 구성하는 부분인자 중 하나로 면역적 또는

비면역적 ligand를 먼저 인식하고 결합하여 classical complement pathway가 시작되도록 하는

주요 단백질이다. 이 논문에서는 돌돔에서 발현되는 3종류의 C1q 유전자를 최초로 동정하였다.

기존에 밝혀진 C1q A, B, C와 유사한 특징을 보이는 3종류의 유전자를 각각 RbC1qAL, RbC1qBL,

RbC1qCL라고 명명하였다. RbC1qAL는 260개의 아미노산 서열을 암호화한 780 bp의 염기서열로

이루어져 있으며 RbC1qBL인 경우 240개의 아미노산을 구성하는 720 bp, RbC1qCL인 경우 726

bp 서열이 242개의 아미노산을 암호화하고 있다. 3종류의 RbC1q 아미노산 서열에서는 N말단에

신호서열과 collagen-like region(s)을, C말단에는 C1q 분자들이 가지는 공통 domain를 확인할 수

있었다. C1q의 특징은 Gly-X-Y 서열반복으로 3종류의 RbC1q 아미노산 서열에서 확인되었고

RbC1qAL과 RbC1qCL에는 대식세포의 포식작용을 활성화시키는 CLR연계 서열(CLR-associated sequence)인 ^{49}GEKGEP^{54}와 ^{70}GEKGEP^{75}가 각각 위치하였다. RbC1qAL과 RbC1qCL의 유전체서열은 6개의 exon으로 RbC1qBL인 경우 5개의 exon으로 구성되어있었다. 계통수분석을 한 결과 3종류의 RbC1q 분자들은 다른 어류 보체 cluster에 속함을 알 수 있었다. 이들 유전자의 발현양상은 비장과 간에서 가장 많이 발현되었으며 특히 돌돔에 *E. tarda*와 *S. iniae*, 이리도 바이러스를 인위감염시킨 경우 감염 초기에 RbC1q의 발현 양이 증가한 반면 lipopolysaccharide를 인위적으로 주입한 경우 주입 후기에 발현 양이 증가하였다. 이러한 결과를 토대로 돌돔에서는 보체인자인 RbC1q 분자들이 병원체 감염에 대한 면역작용을 담당하며 특히 감염 초기에 결정적인 기능을 담당하는 것으로 생각된다.

SUMMARY

Rock bream *Oplegnathus fasciatus*, is one of the most important fish species in the industry of South Korea. However, the practical production of rock bream is considerably obstructed due to the infectious diseases caused by several microbial pathogens. Streptococcosis caused by *Streptococcus iniae*, and edwardsiellosis caused by *Edwardsiella tarda* are mostly infected bacterial diseases have affected the rock bream aquaculture industries, causing extensive economic losses. In addition, iridovirus, which is emerging viral pathogen, have been isolated and identified as a most important pathogen infecting rock bream in the last decade. Understanding the innate immune response against infectious diseases in rock bream is essential for the sustainable production of this species in the aquaculture industry. Therefore, genes from two types of family were studied to understand the behavior of them after microbial infections. The studied genes were two interferon regulatory factors (IRF4 and IRF8) and three complement components (3 C1qs).

The interferon regulatory factor (IRF) members IRF4 and IRF8 contribute to B-lymphocyte development and can act as regulators of immunoglobulin (Ig) light chain gene transcription. These two IRFs are closely interrelated and are expressed at high levels in the lymphoid and myeloid cells of the immune system. In this study, the complete cDNA and genomic sequences of rock bream IRF4 (RbIRF4) and IRF8 (RbIRF8) were identified by homology screening of a multi-tissue normalized cDNA library and a BAC library, respectively, which had been established using Roche 454 GS-FLX™ technology. The full-length RbIRF4 cDNA is composed of 3442 bp and encodes a polypeptide of 462 amino acids; the genomic DNA is 9262 bp in length, consisting of eight exons and seven introns. The full-length RbIRF8 cDNA is composed of 2186 bp and encodes a 422 amino acid polypeptide; the genomic DNA is 4120 bp in length, consisting of nine exons and

eight introns. The deduced amino acid sequences of RbIRF4 and RbIRF8 include a conserved DNA-binding domain (DBD) encompassing a tryptophan pentad-repeat and an IRF-association domain (IAD). Several putative transcription factor binding sites were also identified in 5′ flanking region of both RbIRF4 and RbIRF8, and include those of immune-related factors. Quantitative real time PCR analysis of healthy rock bream detected the highest expression levels of RbIRF4 and RbIRF8 in lymphomyeloid-rich tissues. In addition, viral (rock bream iridovirus) and bacterial (*Edwardsiella tarda* and *Streptococcus iniae*) infection stimulated RbIRF4 and RbIRF8 expressions in head kidney and spleen. These results suggest not only that RbIRF4 and RbIRF8 may have a protective function against virus and bacteria pathogen invasion in rock bream, but also those IRFs may be immunomodulatory factors of teleost fish.

Complement component 1q (C1q) is a subcomponent of the C1 complex and the key protein that recognizes and binds to a broad range of immune and non-immune ligands to initiate the classical complement pathway. In the present study, we identified and characterized three novel C1q family members from rock bream, *Oplegnathus fasciatus*. The full-length cDNAs of C1q A-like (RbC1qAL), C1q B-like (RbC1qBL), and C1q C-like (RbC1qCL) consist of 780, 720 and 726 bp of nucleotide sequence encoding polypeptides of 260, 240 and 242 amino acids, respectively. All three RbC1qs possess a leading signal peptide and collagen-like region(s) (CLRs) in the N-terminus, and a C1q domain at the C-terminus. The C1q characteristic Gly-X-Y repeats are present in all three RbC1qs, while the CLR-associated sequence that enhances phagocytic activity is present in RbC1qAL ([49]GEKGEP[54]) and RbC1qCL ([70]GEKGEP[75]). Moreover, the coding region is distributed across six exons in RbCqAL and RbC1qCL, but only five exons in RbC1qBL. Phylogenetic analysis revealed that the three RbC1qs tightly cluster with the fish clade. All three RbC1qs are most highly expressed in the spleen and liver, as indicated

by qRT-PCR tissue profiling. In addition, all three are transcriptionally responsive to immune challenge, with liver expression being significantly up-regulated in the early phase of infection with intact, live bacteria *(Edwardsiella tarda* and *Streptococcus iniae)* and virus (rock bream iridovirus) and in the late phase of exposure to purified endotoxin (lipopolysaccharide). These data collectively suggest that the RbC1qs may play roles in immune protection against pathogens, possibly by regulating the early phase response of rock bream against bacterial and viral infections.

List of Figures

List of Tables

CHAPTER 1

Interferon regulatory factors 4 and 8 in rock bream, *Oplegnathus fasciatus:* structural and expressional evidence for their antimicrobial role in teleost

1. Introduction

The interferons (IFNs) are a family of multifunctional cytokines that play crucial roles in antiviral responses, immuno-modulation, cell growth and differentiation, and apoptosis (Pfeffer et al., 1998). Transcriptional expression of the IFNs is governed by a group of transcriptional factors called the interferon regulatory factors (IRFs). To date, ten IRF family members have been identified in vertebrates. All members of this family, IRF1 through IRF10, share significant homology in the ~120 amino acid region of the amino-terminus that forms the DNA binding domain (DBD). This DBD displays a unique helix-turn-helix structure, with the DNA recognition helix defined by a cluster of five highly conserved tryptophan residues. These residues have specific affinity towards GAAA and AANNNGAA sequences in interferon stimulatory response elements (ISRE) that are located in the promoters of a diverse range of immune or immune-related genes. In addition, all of the IRFs, except IRF1 and IRF2, contain an IRF-associated domain (IAD) at the carboxyl-terminus. The IAD mediates formation of homo- or hetero-dimers with other IRFs and other transcription factors, which are required for accurate promoter targeting and regulation of transcription (Eroshkin and Mushegian, 1999; Hiscott, 2007). The capacity of IRFs to form interactions with such a broad array of factors, along with their tissue-specific expression patterns, account for the wide range of immune related activities performed by IRFs (Paun and Pitha, 2007).

IRF4 (also known as Pip, LSIRF, LCSAT and MUM1) and IRF8 (also known as the IFN consensus sequence binding protein) are two distinct members of the IFN family that are highly homologous to one other. These two IRF members are involved in the regulation of both innate and adaptive immunity (Tailor et al., 2006), and have been shown to contribute to dendritic cell development and function (Tamura et al., 2005a). Notably, IRF4 plays roles related to B and T cell development and their regulation of the

immune response. In addition to lymphoid cells, IRF4 expression has been found in macrophages (Marecki et al., 1999; Rosenbauer et al., 1999); however, its function in the myeloid lineage is not well characterized. In contrast, the function of IRF8 in the myeloid lineage has been defined. IRF8 deficiency can lead to a pathophysiological state known as chronic myelogenous leukemia (CML) (Tamura et al., 2003). CML is characterized by an uncontrolled growth of myeloid cells in the bone marrow and myeloid accumulation in the blood. Recently, IRF4 was shown to play a vital role during the interleukin (IL)-21-mediated steps of Th17 development by modulating the expression profile of the three key transcription factors associated with Th17 differentiation, namely Foxp3, RORα and RORγt (Huber et al., 2008). A subsequent study revealed a novel pathway by which the ROCK2 Rho-associated protein kinase phosphorylates IRF4, leading to differentiation of Th17 cells and synthesis of IL-17 (Biswas et al., 2010).

In general, IRF4 and IRF8 form complexes by binding with other transcription factors to stimulate or repress expression of immune-related genes. PU.1 and Spi-B, which belong to the Ets transcription factor family, have been characterized as the most robust partners of IRF4 and 8. The PU.1- and Spi-B-containing dimers bind to the Ets-IRF composite element (EICE; GGAAnnGAAA) that is located in the immunoglobulin (Ig) light chain κ 3′ (Ek3′) and λ gene enhancer regions (Brass et al., 1996; Brass et al., 1999; Eisenbeis et al., 1995). Recently, another IRF-Ets composite sequence (IECS; GAAANN(N)GGAA) was identified and shown to be a target of IRF8 (Tamura et al., 2005b). The presence of EICE and IECS sites in several of the known IRF4 and IRF8 target genes suggested that PU.1 and Spi-B interactions may represent a common regulatory mechanism of the immune system.

Recent molecular studies in ichthyology have identified some IRFs, based on their significant sequence similarity to mammalian orthologous. Thus far, fish IRF4 and IRF8 have been cloned and studied only in rainbow trout (*Oncorhynchus mykiss*) (Holland et al., 2010). In this study, we aimed to determine the genomic structure and function of IRF4 and IRF8 in the teleost species rock bream (*Oplegnathus fasciatus*), which is one of the most valuable fish species of the Korean aquaculture industry. We identified the full-length cDNA and genomic structure of rock bream IRF4 (RbIRF4) and IRF8 (RbIRF8). Apart from investigating their tissue-specific distribution patterns in healthy rock bream, we also evaluated the responsiveness of RbIRF4 and RbIRF8 to immune stimulation by live bacteria and virus, and immune stimulatory agonists lipopolysaccharide (LPS) and polyinosinic:polycytidylic (poly I:C), to examine their roles in bacterial and viral infection.

2. Materials and Methods

2.1. Identification of rock bream IRF4 and IRF8 cDNAs

The rock bream cDNA sequences of IRF4 and IRF8 were identified from a previously established cDNA sequence database (Umasuthan et al., 2011a) by applying the BLAST tool in the NCBI web-based query system (http://www.ncbi.nlm.nih.gov/BLAST). This rock bream sequence database was established by using the Roche 454 Genome Sequencer FLX System (GS-FLX™), a next-generation DNA sequencing technology using the GS-FLX Titanium instrument (Droege and Hill, 2008).

2.2. Identification of rock bream genomic DNA

Genomic sequences of RbIRF4 and RbIRF8 were obtained from a bacterial artificial chromosome (BAC) library of rock bream. We have previously established a rock bream BAC library, designated as the *Oplegnathus fasciatus* random shear BAC library (OfBAC), that contains 92,160 clones arrayed in 240 384-well plates. Using this BAC library, a 1st round PCR was performed on super pools with gene-specific primers to determine which super pool(s) contained the BAC clone(s) of interest. The plate, row and column pools from the positive super pool(s) were then screened in the 2nd round PCR to localize the sequences of RbIRF4 and RbIRF8. Finally, the positive clones were picked and isolated. The corresponding BAC DNA was sequenced to obtain the complete genomic sequences of RbIRF4 and RbIRF8 using the GS-FLX™ system (Macrogen, Korea).

2.3. Molecular characterization of RbIRF4 and RbIRF8

The full-length sequences of RbIRF4 and RbIRF8 were analyzed by the BLAST tool and compared with other known IRF family members. DNAssist (version 2.2) was used to determine the open reading frames (ORFs) and amino acid sequences. The deduced amino acid sequences were analyzed with the Expert Protein Analysis System (http://www.expasy.org). To determine the conserved domains in the RbIRF4 and RbIRF8 predicted proteins, motif scan Pfam hidden Markov models (Local models) were employed (http://hits.isb-sib.ch/cgi-bin/PFSCAN). Pairwise and multiple alignments were made using ClustalW2 (http://www.ebi.ac.uk/Tools/clustalw2/index.html) with default parameters. Exon-intron structures were determined by aligning the cDNA sequences (Section 2.1) to the genomic sequence with the Spidey program (http://www.ncbi.nlm.nih.gov/IEB/Research/_Ostell/Spidey/). Exon-intron structures of

other IRFs were obtained from the Ensembl Genome Browser (http://www.ensembl.org/index.html) for use in further comparison. The transcription initiation site (TIS) was predicted by the Natural Network Promoter Prediction program (http://www.fruitfly.org/seq_tools/promoter.html). The transcription factor binding sites in the 5′-flanking region sequences were predicted by the TFSEARCH ver1.3 (http://www.cbrc.jp/research/db/TFSEARCH.html) and the AliBaba ver2.1 programs (http://www.gene-regulation.com/pub/programs/alibaba2/index.html) with default parameters. The DBD and IAD sequences of RbIRF4 and RbIRF8 were individually submitted to the SWISS-MODEL (http://www.swissmodel.expasy.org) server for protein structure homology analysis. The analogous models were derived using human IRF4 (Protein Data Bank (PDB) entry ID 2dllA) as the structural template, and domain structures were manipulated using the Swiss-PdbViewer program. Phylogenetic trees were constructed by applying the neighbor-joining method using MEGA ver5.05 and employing 1000 bootstrapping tests.

2.4. Experimental fish and immune challenges

Rock bream (~50 g) were obtained from the Ocean and Fisheries Research Institute (Jeju, Republic of Korea) and maintained in 400 L tanks with seawater (salinity 34 ± 0.6%, pH 7.6 ± 0.5) and a recirculating system at 23 ± 1°C. Fish were acclimated for one week prior to sampling.

To evaluate the immune responses of RbIRF4 and RbIRF8, rock bream iridovirus (RBIV), two pathogenic bacterial strains (*Edwardsiella tarda* and *Streptococcus iniae*), and two immunostimulants (LPS and poly I:C) were respectively suspended in phosphate buffered saline (PBS) for use in immune challenge experiments.

For the RBIV challenge, kidney tissue samples were obtained from moribund fish infected with RBIV and homogenized in 20 volumes of PBS. The homogenate was centrifuged (3000 × g for 10 min at 4°C) to separate the virus-containing supernatant, which was filtered through a 0.45 μm membrane and diluted to obtain 10^3 $TCID_{50}$ of RBIV per 100 μl. Then administered to fish via an intraperitoneal (i.p.) injection of 100 μL RBIV in PBS. For the pathogenic bacteria challenge, the *E. tarda* and *S. iniae* strains were obtained from the Department of Aqualife Medicine at Chonnam National University (Republic of Korea). The bacteria were incubated at 25°C for 12 h in brain heart infusion (BHI) broth supplemented with 1% NaCl. The cultures were then diluted to a desired concentration using PBS. Each fish was i.p. injected with 100 μL live *E. tarda* ($5 × 10^3$ $CFU\ μL^{-1}$) or 100 μL live *S. iniae* ($1×10^5$ $CFU\ μL^{-1}$) in PBS. For the immunostimulants challenge, animals were administered an i.p. injection of either 100 μL purified LPS (1.25 $μg\ μL^{-1}$, *E. coli* 055:B5; Sigma-Aldrich, Missouri, USA) or 100 μL poly I:C (1.5 $μg\ μL^{-1}$; Sigma-Aldrich) in PBS. An injection control group was established by injecting an equal volume (100 μL) of PBS alone. A negative control group was established with fish that received no injection. Rock bream head kidney and spleen samples were collected at 3, 6, 12, 24 and 48 h post-injection (p.i.) from all groups. All resected samples were immediately snap-frozen in liquid nitrogen and stored at -70°C until use.

2.5. Total RNA extraction and first strand cDNA synthesis from different tissues

Different tissues (gill, liver, spleen, head kidney, kidney, skin, muscle, heart, brain and intestine) were dissected from healthy fish, snap-frozen in liquid nitrogen, and stored at −80°C until use for total RNA extraction. Blood samples (~1 mL $fish^{-1}$) were collected from the caudal vein of fish and blood cells were harvested by centrifuging at 3000 × g for 10 min at 4°C. Tissue samples were obtained from three biological replicates.

Total RNA was extracted from the collected tissues by using the Tri Reagent™ (Sigma-Aldrich) according to the manufacturer's protocol and diluted to 1 µg µL^{-1} for use as template to synthesize cDNA. The cDNA was synthesized by using PrimeScript™ First Strand cDNA Synthesis kit (TaKaRa, Shiga, Japan), as follows. First, 2.5 µg RNA was incubated with 1 µL of 50 µM oligo(dT)$_{20}$ and 1 µL of 10 mM of dNTPs for 5 min at 65°C. Then, 4 µL of 5x PrimeScript™ buffer, 0.5 µL of RNase inhibitor (40 U µL^{-1}), and 1 µL of PrimeScript™ RTase (200 U µL^{-1}) were added and the mixture was incubated at 42°C for 60 min. Next, the temperature of the reaction mixture was adjusted to 95°C and incubated for 5 min to terminate the reaction. The cDNA was then diluted by 40× and used as template to analyze the differential transcription levels of RbIRF4 and RbIRF8 *in vivo* by using quantitative real time PCR (qRT-PCR).

2.6 qRT-PCR analysis of RbIRF4 and RbIRF8

The TP800 Thermal Cycler Dice™ Real Time System (TaKaRa) was used to measure the mRNA expression levels of RbIRF4 and RbIRF8. The total reaction mixture was 20 µL and contained 4 µL of diluted cDNA (3.125 ng µL^{-1}) from the respective tissue, 10 µL of 2x TaKaRa ExTaq™ SYBR premix, 0.5 µL of each primer (10 pmol µL^{-1}) and 5 µL of deionized H$_2$O. **Table 1** shows the list of primers used in this study. The thermal reaction profile of qRT-PCR was: one cycle of 95°C for 10 s, followed by 45 cycles of 95°C for 5 s, 58°C for 20 s and 72°C for 20 s, and a final single cycle of 95°C for 15 s, 60°C for 30 s and 95°C for 15 s. Rock bream β-actin (GenBank accession no. FJ975145) was amplified with the same qRT-PCR temperature profile for use as the internal reference. The Livak ($2^{-\Delta\Delta CT}$) method (Livak and Schmittgen, 2001) was used to calculate the relative expression levels of RbIRF4 and RbIRF8. All data are presented in terms of relative mRNA expressed as means ± standard deviation (SD). Statistical differences

between the control and treatment groups were determined by Student's *t*-test using the GraphPad software, and differences were considered significant at $p<0.05$.

Table 1. Primers used in this study

Primer name	Application	Sequence (5′-3′)
RbIRF4-F	qRT-PCR amplification	ATGGCTTATACGCTAAGCGCCTCT
RbIRF4-R	qRT-PCR amplification	GTTTGCATGGCTGTTCCTTCTCCA
RbIRF8-F	qRT-PCR amplification	GTGCCAACCAAGAGGGCATCTTTA
RbIRF8-R	qRT-PCR amplification	TTGACCACAGCATCCCTCTCAAGT
Rbβ-actin-F	qRT-PCR internal reference	TCATCACCATCGGCAATGAGAGGT
Rbβ-actin-R	qRT-PCR internal reference	TGATGCTGTTGTAGGTGGTCTCGT

F, forward; R, reverse.

3. Results

3.1 Molecular characterization of RbIRF4 and RbIRF8

The full-length cDNAs of rock bream IRF4 (**Fig.1**) and IRF8 (**Fig.2**) were found to consist of 3442 bp and 2186 bp, respectively. The RbIRF4 cDNA (GenBank accession no. JQ388475) contains a 221 bp 5′-untranslated region (UTR), a 1835 bp 3′-UTR, and an ORF of 1386 bp that translates into a 462 amino acid putative peptide with a predicted molecular mass of 53.4 kDa. The RbIRF4 3′-UTR contains four mRNA instability motifs (ATTTA) and one more effective instability motif (^{2005}TTATTTATT^{2013}) (Zubiaga et al., 1995). Two polyadenylation signals (^{3357}ATTAAA^{3362} and ^{3399}AATAAA^{3404}) lie upstream of the polyA tail. The RbIRF8 cDNA (GenBank accession no. JQ388476) contains a 153 bp 5′-UTR, a 767 bp 3′-UTR, and an ORF of 1266 bp that translates into a 422 amino acid putative peptide with a predicted molecular mass of 47.7 kDa. Two mRNA instability motifs (ATTTA) and three polyadenylation signals (^{2009}AATAAA^{2014}, ^{2131}TATAAA^{2136}, and ^{2147}AGTAAA^{2152}) lie upstream of the RbIRF8 polyA tail.

In silico analysis revealed that both deduced amino acid sequences of RbIRF4 and RbIRF8 have a DNA-binding domain at the N-terminal region, which contains five tryptophan residues, and an IRF-association domain at the C-terminal region. Rock bream IRF4 and IRF8 displayed a higher degree of identity to other known respective IRFs (**Fig. 3 and 4**). RbIRF4 exhibited the highest overall identity (87.9%) to the homologue from Nile tilapia, with 95.6% identity in the DBD and 90.8% identity in the IAD. RbIRF8 also showed the highest overall identity (84.5%) to homologue from Nile tilapia, with 96.5% identity in the DBD and 85.2% identity in the IAD (**Table 2**). Motif search of the RbIRF4 C-terminal region revealed the presence of a serine-rich domain, which is targeted for virus-induced phosphorylation to facilitate interaction with other IRF members and subsequent activation of virus clearance signaling pathways (Sun et al., 2007). When individually constructed, the domain structures for the DBDs and IADs of RbIRF4 and RbIRF8 clearly depicted the structural homology (**Fig. 5**). The positions of the tryptophan pentad-repeats were conserved in both DBDs, the unique presence of a serine-rich domain in the IAD of RbIRF4 made it distinct from that of the RbIRF8. Phylogenetic analysis showed that the IRF family members diverged into four main clusters, and the RbIRF4 and RbIRF8 clustered within the IRF4 subfamily members (**Fig. 6**). Phylogenetic analysis of the IRF4 subfamily members diverged into two major branches of IRF4 and IRF8, while IRF4 was further sub-clustered into three distinct clades; fish IRF4a, fish IRF4b and IRF4 from other vertebrates (**Fig. 7**). RbIRF4 was clustered with the teleost IRF4a group and showed the closest relationship to the IRF4 of Nile tilapia and IRF4a of medaka, whereas, RbIRF8 showed closest relationship with IRF8 of Nile tilapia.

```
TGCAGAGGAGAAGCTGTGTGGAGCACTAACTAACCCTGGCACCTCTGTCTTTACTAAGAAGTGTTACTGGAAACTGTTTTTAGCTGTTTT      90
AAGCTGAAAATGTGTTAATCAACTGATGAATTTCATCTTGTCAAACTTCTTCTCTCAGGTTCATTTAAATATGAGAATCTCACATTAGTA     180
TAAGTCTAATTAAACTAAATTTTTTGTCTTATTTTGCAGAG**ATG**AACCTGGAGGAGGACAGCGGCCTGTCAGTCAGCTGCGGCAACGGCA     270
                                         M  N  L  E  E  D  S  G  L  S  V  S  C  G  N  G      16
AACTGAGGCAGTGGCTGATTGATCAGATTGACAGCAGGAGATATGCTGGCCTGGTTTGGGAAAACGATGAGAAATCCATCTTTAGGATCC     360
 K  L  R  Q  W  L  I  D  Q  I  D  S  R  R  Y  A  G  L  V  W  E  N  D  E  K  S  I  F  R  I     46
CGTGGAAACATGCAGGGAAACAAGACTATAACAGAGAGGAGGATGCTGCGCTCTTCAAGGCGTGGGCACTGTTTAAGGGAAAATATAAGG     450
 P  W  K  H  A  G  K  Q  D  Y  N  R  E  E  D  A  A  L  F  K  A  W  A  L  F  K  G  K  Y  K     76
AAGGAGTCGACAAGCCAGATCCCCCCACATGGAAAACCAGACTTCGCTGTGCTCTTAATAAAAGCAATGACTTTGATGAGTTAGTGGAGA     540
 E  G  V  D  K  P  D  P  P  T  W  K  T  R  L  R  C  A  L  N  K  S  N  D  F  D  E  L  V  E    106
GAAGCCAACTGGACATCTCAGAGCCCTATAAAGTCTACAGGATCATTCCAGAGGAAGCCAAAAAAGGAATGAAGATGAGCAGTATGGAGG     630
 R  S  Q  L  D  I  S  E  P  Y  K  V  Y  R  I  I  P  E  E  A  K  K  G  M  K  M  S  S  M  E    136
AGACAGCATCACATGTAAATGCCCATGGCTATATTGCTCCATATACATCTCTGCACAACCAGGTACCTGGCTACATGCTTTCTCAGGATA     720
 E  T  A  S  H  V  N  A  H  G  Y  I  A  P  Y  T  S  L  H  N  Q  V  P  G  Y  M  L  S  Q  D    166
GAAGGGACTGGAGGGATTACACACCACCAGAACAGCAACCTCTTCCCCCTCCACATCACCATGGGCCACATGCAGAGGTGCAGTATGGTC     810
 R  R  D  W  R  D  Y  T  P  P  E  Q  Q  P  L  P  P  P  H  H  H  G  P  H  A  E  V  Q  Y  G    196
AGTGCCACTACCCGTCACCGTTCAGCCGGGCTTGGCCTGGGTCACACACAGAAAATGGTTTTCAGCTCTCCTTCCACACCTACTTTTCAG     900
 Q  C  H  Y  P  S  P  F  S  R  A  W  P  G  S  H  T  E  N  G  F  Q  L  S  F  H  T  Y  F  S    226
AGTCCCAACCGCCTGTGTACACAATGAACCACAACAATGCCATAACAGATTTCAGCCTGCATGTGTCCCTATACTACAGAGAATCTTTGG     990
 E  S  Q  P  P  V  Y  T  M  N  H  N  N  A  I  T  D  F  S  L  H  V  S  L  Y  Y  R  E  S  L    256
TGAAGGAGGTTACCACCACCAGCCCGGAAGGCTGTCGGATCACCTCTTCTTCCTCCTCCTCTCCATCCTCCTCCTCCTCTTCTTCCCCAT    1080
 V  K  E  V  T  T  T  S  P  E  G  C  R  I  T  ░S░ ░S░ ░S░ ░S░ ░S░ ░P░ ░S░ ░S░ ░S░ ░S░ ░S░ ░S░ ░P░ P    286
GCCCAGAGGACAAGTTTCACAGCGGGGCAGAGGTCATCCTTTTTCCATTCCCATACCCTGAGTCTCACCGGCAGGGTGCTGAGATGCTTC    1170
 C  P  E  D  K  F  H  S  G  A  E  V  I  L  F  P  F  Y  P  E  S  H  R  Q  G  A  E  M  L    316
CTAATGTACTGGAGAGGGGTGTGCTGCTGTGGATGATGTCTGATGGCTTATACGCTAAGCGCCTCTGCCAGGGACGAGTGTACTGGGAGG    1260
 P  N  V  L  E  R  G  V  L  L  W  M  M  S  D  G  L  Y  A  K  R  L  C  Q  G  R  V  Y  W  E    346
GACCTCTGGCGCCATATATGGATAAACCCAACAAGCTGGAGAAGGAACAGCCATGCAAACTGTTTGACACCCAGCAATTCCTCATTGAGC    1350
 G  P  L  A  P  Y  M  D  K  P  N  K  L  E  K  E  Q  P  C  K  L  F  D  T  Q  Q  F  L  I  E    376
TTCAAGATTTTGCCCATAATGGCCGACATTTACCGAGACTCCAGGTGGTGCTATGTTTTGGAGACGAGTACCCCGACCCACAGCGGCCAA    1440
 L  Q  D  F  A  H  N  G  R  H  L  P  R  L  Q  V  V  L  C  F  G  D  E  Y  P  D  P  Q  R  P    406
GGAAGATGATCACAGCACAGGTGGAGCCAGTGTTTGCCAGAAAACTGGTGTACTATTACCAGCAGAACAATGGCCACTACCTGCGGGGTT    1530
 R  K  M  I  T  A  Q  V  E  P  V  F  A  R  K  L  V  Y  Y  Y  Q  Q  N  N  G  H  Y  L  R  G    436
ATGATCACATCCAGGAACAGAACACATCTCCAGCAATCGACTACCCTTCCCAGAGACCTCTACAGCATATTCAAGAG**TGA**CAAGCATGCA    1620
 Y  D  H  I  Q  E  Q  N  T  S  P  A  I  D  Y  P  S  Q  R  P  L  Q  H  I  Q  E  *              462
CACACACACGCAGGAGGGCATGAATGTGAATGTAAACCGCTGCCCATTTAACGCGCCTTGTGTCTGCAGTGCTGTGTACTGTAAATGGTA    1710
GAAGCTCTCATGGTACTGAGGGGTAGAGATGGGGAAGACCCTGTTGTTTTCACTTGCACACCTATGCGTCGTAGTCTTTGTGTCTGTGAC    1800
CGTGGAGTAATGTTTTGTTTTTTTTCTCTACAAGTATTATTGTGACAGTGTGAAGGATGCGGTTTGAGGGTGGGGGGATGAGGGGAGCGG    1890
AGGGCGACTGAGACAGGAAGTACTAGCACAGGTGCACAGGAAGAATGGCCCCTAAGACAAACACTGATATGTTAAAGGGATAAAACAGAG    1980
TATGAATGAATACTTATTGTATTATTATTATTGCAGTTTATATAATACTTAAGATTTGCTTAGATGAGCAAATGAACAGTCAACAGTAT    2070
CATTGCTTCATCCATTGTAAATCTCAATTTTGGGCTTCATTGCATAAAAATGCTACAACTGGAATGTGAGAGAAACTGAAATTATTATCA    2160
CTGGTCTGGAGGTGTAGTAGAGGAAGCGGTTATCACAAATTGTTTACCTTCGGGCTGCTACATTATTAGTATTGCCGAACGCATTTACTT    2250
ACTTACTCAGCTTTTAGAAATATCTGATTCCTTTTAATTACGTCAAACTGATGAAACAAAAATCATCAAAAGCAGCTTGTGGATGGTAAT    2340
CAAGATAAATGAATTGAGAACATACCTCCATATTATATACATGCCCTTTACTGTCCTGTATTATATTTCTAAGACGGTTATTGGGTTTTG    2430
AATTTTGATAGTAATTTCAAGACATTCAGGGTTTGTTTACATGTGAATCTAATGGTTGTTAGCATTAGAAAGCAAACCTGTGTATTAAGT    2520
TACAAGTGCAGAAGGTTTTCAGACATTTCTTTTTTATATACTTCATTTGTCTTGTGAAATACTGTATAGTTTTAGTTTCTCCGGGTCTGAT    2610
TTGAAAACCCGTTTTTCGATTGAACTTCTTACAACTTATAGACATGAAACTTGCGATGATGACGGTTCGCAAGTCGACATGACTTTCGTC    2700
TCAAGCGCTCACAAGCGAAAATGTTGGGGACCACTGGCCTACAAACATTATCAGAGAGCAGCTTGGGCCTCAAGGTCTGCGTCACCAGA    2790
GAGTGTGGCCTTTATCGGCAGCCTTGAGCTTACTGGACTGAAGGAGGCTGACTGGAACAATCACACTCATCCCGTAACTTGCAAGGAAAC    2880
CGTACCTGTTCCTCTTGACCTGAAACTAGTGGCCTCGCAGATGGGAAGAAAAACTACCTTTGTGGTCTACGATCAAGTTCGGCCAGACAG    2970
GAGCTCTCTGGTTCCCATGCCTCACACGCTGTAGAGGTTCACGCACTGGATGCAACGTCCACACAAAAGCTGTCGTGAGATCCTAACAAC    3060
CTATTTAGGAATGTAAACATATCATAAGACATAAAAAACAACAGTGAGAATATTGGAATTCGTCACGCTATTGAAGTCCCAACGGCTGAT    3150
GCTGCAGTAGAATGGATAAAGGTTTTGTTATGTTTTTTTCTGTGGTCCAGAGTTTCGATCTCTGCATTTACTTCCTTCACACTTAAACCA    3240
CATCACTTTCACAATGGCTTCAAAAGTCTGACCAAAATCACAGGCAGGAATGTTTTGTGTGTCTGTGTGTGTGTGTGTGTGTGTGTCTGT    3330
GTGTGAAACATCTAGTCAGTCCCTCG**ATTAAA**ATTTTACTTCACAGTCAGACATTTGAGCCATTCAA**AATAAA**ATCCTAACAGTGAAAA    3420
CGTTACCAAAAAAAAAAAAAAAA  3442
```

Fig. 1. **Full length cDNA sequence and deduced amino acid sequence of RbIRF4**. The start and stop codons are in boldface. The serine-rich domain in the carboxyl terminal region are shaded in grey color. Two polyadenylation signals (ATTAAA and AATAAA) are underline with boldface.

```
TGGGCATGTGTCGGGGAGAAGTTGAACTCTCTTGATTAAGACTGACAGTGAGCCATCTATCAACATCTCAGGATATCACACACAGGGCCA  90
AGAGCTCAGCTGCCTATCAAACTGTACAAGCGAACGATCCATAAACAGCTTTTCTCCATCAAGATGTCAAACACGGGAGGTCGCAGACTG  180
                                                               M  S  N  T  G  G  R  R  L    9
AAGCAGTGGCTGGTGGAGCAGATCCAGAGCGCGCAGTACTCTGGACTGCAGTGGGAGGATGAAAGTCGCACTTTGTTCCGAATTCCATGG  270
 K  Q  W  L  V  E  Q  I  Q  S  A  Q  Y  S  G  L  Q  W  E  D  E  S  R  T  L  F  R  I  P  W    39
AAACACGCAGGGAAACAGGATTACAACCAAGAGTAGATGCATCCATTTTTAAGGCCTGGGCTGTGTTTAAAGGCAAGTTTAAGGAGGGG  360
 K  H  A  G  K  Q  D  Y  N  Q  E  V  D  A  S  I  F  K  A  W  A  V  F  K  G  K  F  K  E  G    69
GACAAGGCTGAACCTGCAACATGGAAGACCAGACTCCGCTGTGCCCTGAACAAAAGCCCTGACTTTGAGGAGGTGACAGAAAGGTCACAG  450
 D  K  A  E  P  A  T  W  K  T  R  L  R  C  A  L  N  K  S  P  D  F  E  E  V  T  E  R  S  Q    99
TTGGACATCTCTGAGCCCTATAAAGTCTACCGCATTGTACCCGAGGAAGAGCAGAAGCATGGCAAAAGCTCAGTGATGGCCATGGCAGCC  540
 L  D  I  S  E  P  Y  K  V  Y  R  I  V  P  E  E  E  Q  K  H  G  K  S  S  V  M  A  M  A  A   129
ACCACCAGCTCCGGTGATATCACTGACATGGACTGCAGCCCTGCAGACCTAGAGGAGCTCATTAAAGAGGAGGAGGGCTGTAGTATCCAG  630
 T  T  S  S  G  D  I  T  D  M  D  C  S  P  A  D  L  E  E  L  I  K  E  E  E  G  C  S  I  Q   159
TCCAGTCCAGAATATTGGTCCCAAGGCAGCATCAATGCTTTCCCACTGCATCAGGACCCTTTGCCATCGGGAACTCTCAGCTCAGCTCTC  720
 S  S  P  E  Y  W  S  Q  G  S  I  N  A  F  P  L  H  Q  D  P  L  P  S  G  T  L  S  S  A  L   189
TCCCAAATGATGATCAGTTTCTACTATGGAGGAAAGCTGATGCACAACACACTGGTTGCTCATCCTGAAGGCTGCCGGATCTCCCCACAA  810
 S  Q  M  M  I  S  F  Y  Y  G  G  K  L  M  H  N  T  L  V  A  H  P  E  G  C  R  I  S  P  Q   219
CAACACCTGGGTCGTGGCGCCCTATACAGTTCAGACAGCATGCAGTGTGTTAATTTCCCCCCCGCTGAGCTCATCGAGTACGACCGCCAG  900
 Q  H  L  G  R  G  A  L  Y  S  S  D  S  M  Q  C  V  N  F  P  P  A  E  L  I  E  Y  D  R  Q   249
CGCCATGTCACACGCAAGCTCCTGGGCCACCTGGAGAGAGGCGTGCTGGTCCGTGCCAACCAAGAGGGCATCTTTATCAAAAGGCTGTGC  990
 R  H  V  T  R  K  L  L  G  H  L  E  R  G  V  L  V  R  A  N  Q  E  G  I  F  I  K  R  L  C   279
CAGAGCCGTGTCTTCTGGAGCGGGCTGGGAGAAGTGGGCTCACAGTATAGCCCCATGCCTTGTAAACTTGAGAGGGATGCTGTGGTCAAG  1080
 Q  S  R  V  F  W  S  G  L  G  E  V  G  S  Q  Y  S  P  M  P  C  K  L  E  R  D  A  V  V  K   309
ATTTTTGACACAGAAAGGTTTCTTCAAGCTCTACAGCTGTACCAGGAGGGTCAGTTTCCTGCTCCTGATCCGACAGTGACCCTATGTTTT  1170
 I  F  D  T  E  R  F  L  Q  A  L  Q  L  Y  Q  E  G  Q  F  P  A  P  D  P  T  V  T  L  C  F   339
GGAGAGGAGCTCCATGATGTCAGCAATGCTAAGGGCAAACTGATCATTGTGCAGATCACTGTGGTGAACTGCCAGCACCTGCTGGACGCA  1260
 G  E  E  L  H  D  V  S  N  A  K  G  K  L  I  I  V  Q  I  T  V  V  N  C  Q  H  L  L  D  A   369
GTGAACATGCGGCGCACCCAGCCCTTCTGCAACAACCCAAACCTGGACATGTCTGATAATGTGGCCACGGACCAGATGGCCCGCATCTAC  1350
 V  N  M  R  R  T  Q  P  F  C  N  N  P  N  L  D  M  S  D  N  V  A  T  D  Q  M  A  R  I  Y   399
CAGGACTTGTGCAGCTACAGCGGCCCCCAGAGGCCAGCCTGCTACAGGGACAACATGCCCATCACTGCCTGAACTGTGAGCAGTAACACA  1440
 Q  D  L  C  S  Y  S  G  P  Q  R  P  A  C  Y  R  D  N  M  P  I  T  A  *               422
CCTGCCAGGAGCTCTTCAGCGTGTACTAGACTGAATGTACAACATTTTAGGGACATTCCTTATTTTCCTGACTTACATTGACAAGGTGAA  1530
CTCATCACAATTAGTTTTGACATTTATACCAATCACAGTGGAGACATAAAGGAGGAACAAATCTCTTACAGTAAGACTTTTAAGCTGAAA  1620
CTCAATATTAGGATGTCCCTAATATTTTGTATATCAGTAAGTGTTTATATTTGTATAGGTACAGCACAACCTGTGTCTAAACTCTGCTT  1710
TATTATTACAAGAGAGTGTTATCATCCAAGTCTTGCAAAGCCTTAAGAACAGATTCCTTGTGGTGACTTTTTAACAAATCTTTCTTATAT  1800
ATATATATCTTTAAAGACTCTCACTGAAAGCCTCATTGCAGTATGACAGTTTGATATATGTAGGCAATTTAGTTCTCAAAACTGAATTTT  1890
ATTCTGACACGCATCATCTTTTTTGTTGTTGTTTTTGAAGTTGCTTCCTTATGTGACGACCATTTGGATAAACAGGCCAAATAGATATTT  1980
TGTCAGATATTTTTCTATTTTTTACAGTAATAAAAAATCTTTATATGAGTTGATAAATATTGCATGGAATTTTAGATGCAATGAGGCAT  2070
TATATTACAGATAGGTTTTACTATCAGACTGTTTTACTTGACCTTTGTTTCATCATTTTTTATAAACTGCATTTCTAGTAAAAGAAAATT  2160
TTTACTGAACTCCAAAAAAAAAAAAAA  2186
```

Fig. 2. **Full length cDNA sequence and deduced amino acid sequence of RbIRF8**. The start and stop codons are in boldface. Three polyadenylation signals (AATAAA, TATAAA and AGTAAA) are underline with boldface.

Table 2. Putative identity and similarity of the rock bream IRF4 and IRF8 with known orthologues

Gene	Species	Accession number	Overall		DBD		IAD	
			I%	S%	I%	S%	I%	S%
IRF-4	Human	AAH15752	57	68.6	91.2	98.2	57.8	70.8
	Mouse	NP_038702	56.8	68.6	90.4	98.2	56.8	70.3
	Chicken	NP_989630	59.5	71.2	92.1	98.2	60.5	71.9
	Zebra fish	NP_001116182	66.9	77.5	91.2	98.2	74.1	84.3
	Nile Tilapia	XP_003437930	**87.9**	**93.1**	**95.6**	**98.2**	**90.8**	**95.7**
IRF-8	Human	NP_002154	50.1	63.5	85.8	92.9	48.4	63
	Mouse	NP_032346	50.3	62.1	84.3	91.3	48.1	62.3
	Chicken	NP_990747	51.6	64.9	87.6	92.9	49.9	65.2
	Zebra fish	NP_001002622	65.9	77	90.3	95.6	65.8	77.2
	Nile Tilapia	XP_003442371	**84.5**	**90.7**	**96.5**	**99.1**	**85.2**	**92.3**

I, identity; S, similarity; DBD, DNA-binding domain; IAD, IRF-association domain. Bold font indicates the highest percentages.

```
                                                  _____
HumanIRF4       MNLEGGGRGGEFGMSAVSCGNGKLRQWLIDQIDSGKYPGLVWENEEKSIFRIPWKHAGKQDYNREEDAAL 70
MouseIRF4       MNLETGSRGSEFGMSAVSCGNGKLRQWLIDQIDSGKYPGLVWENEEKSVFRIPWKHAGKQDYNREEDAAL 70
ChickenIRF4     MNLEPG----ECGMNSVSCGNGKLRQWLIDQIDSGKYPGLVWENEDKSIFRIPWKHAGKQDYNREEDAAL 66
ZebrafishIRF4a  MNLDGD------CIMSVSCGNGKLRQWLIEQIDSGEYSGLVWENDEKTIFRIPWKHAGKQDYNRDEDAAL 64
ZebrafishIRF4b  MCSDEE------RGMTGSSGNGKLRQWLIEQVDTGKYPGLVWENDEKSIFRIPWKHAGKQDYNRDEDAAL 64
RockbreamIRF4   MNLEED------SGLSVSCGNGKLRQWLIDQIDSRRYAGLVWENDEKSIFRIPWKHAGKQDYNREEDAAL 64
                ** :            : *.*********:*:*: .*.******:**::****************:*****

                _____ DBD _____
HumanIRF4       FKAWALFKGKFREGIDKPDPPTWKTRLRCALNKSNDFEELVERSQLDISDPYKVYRIVPEGAKKGAKQLT 140
MouseIRF4       FKAWALFKGKFREGIDKPDPPTWKTRLRCALNKSNDFEELVERSQLDISDPYKVYRIVPEGAKKGAKQLT 140
ChickenIRF4     FKAWALFKGKFREGIDKPDPPTWKTRLRCALNKSNDFEELVERSQLDISDPYKVYRIVPEGAKKGAKQNS 136
ZebrafishIRF4a  FKAWALFKGKYREGLDKPDPPTWKTRLRCALNKSNDFDELVERSQLDISDPYKVYRIVPEGAKRGSKAIS 134
ZebrafishIRF4b  FKAWALFKGKFREGVDKPDPPTWKTRLRCALNKSNDFEEIVERSQLDISDPYKVYRIVPEGSKKGSR--S 132
RockbreamIRF4   FKAWALFKGKYKEGVDKPDPPTWKTRLRCALNKSNDFDELVERSQLDISEPYKVYRIIPEEAKKGMKMSS 134
                **********::**:*************:*******:*:*************************:** :*:* :  :

HumanIRF4       LEDP-QMSMSHPYTMTTPYPSLPAQQVHNYMMPPLDRSWRDYVPDQ-----------PHPEIPYQCP-MT 197
MouseIRF4       LDDT-QMAMGHPYPMTAPYGSLPAQQVHNYMMPPHDRSWRDYAPDQ-----------SHPEIPYQCP-VT 197
ChickenIRF4     MEE--QPLMNHPFPITSPYTSLPS-QVPNYMVP-HERNWREFAPEQ-----------PHPDIPYQCASVP 191
ZebrafishIRF4a  MEENTTHVTPLSYPMHSAYPALQP-QMSGFMLPQERRDWREFGSD-----PPHTQTPHADLPYGQC-PY 196
ZebrafishIRF4b  IEDSQSNSGSPNYPMHPTYAPAPS-QVCNYISP-AERGWREYPTLS--------------DISYSQS--- 183
RockbreamIRF4   MEETASHVNAHGY--IAPYTSLHN-QVPGYMLSQDRRDWRDYTPPEQQQPLPPPHHHGPHAEVQYGQC-HY 200
                 :::          :   ..* .    *: .:: .   *.**:: .               :: *

HumanIRF4       FGPRGHHWQGPACENGCQVTGTFYACAPPESQAPGVPTEPSIRSAEALAFSDCRLHICLYYREILVKELT 267
MouseIRF4       FGPRGHHWQGPSCENGCQVTGTFYACAPPESQAPGIPIEPSIRSAEALALSDCRLHICLYYRDILVKELT 267
ChickenIRF4     FAARGHHWQGPGCENGCQVTGTFYACAPPESQTPGIPIEPSIRSGEALALSDCRLHICLYYREMLVKEVT 261
ZebrafishIRF4a  PPSRSLPWHTAPCDNGYQISGSFYTYSPSESHPV--AMDPSMRSAEAMAISDCRLHVSLFYRESLVKELT 264
ZebrafishIRF4b  --PYTSRWDP-----GYQFSGSFYSCNASDPQPSPFTLDTSMRSAEAMALSDYRLHVMVFYRDALVREVT 246
RockbreamIRF4   PSPFSRAWPGSHTENGFQLS--FHTY-FSESQP------PVVTMNHNNAITDFSLHVSLYYRESLVKEVT 261
                 .    *        * *.:  *::   .:.:.      .    .   *::*  **: ::**: **:*:*

HumanIRF4       TSSPEGCRISHGHT---YDASN------------LDQVLFPYPEDNGQRKNIEKLLSHLERGVVLWMAPD 322
MouseIRF4       TTSPEGCRISHGHT---YDVSN------------LDQVLFPYPDDNGQRKNIEKLLSHLERGLVLWMAPD 322
ChickenIRF4     TSSPEGCRISQGQS---YEVSS------------LEQVIFPYPEDNSQRKNIEKLLSHLERGVILWMAPD 316
ZebrafishIRF4a  TSSPEGCRISSSAS--PGSPSSPSSPSEERLYGGAEPVLFPFPYPQSQRRGAEKLPNVLERGVLLWLSPD 332
ZebrafishIRF4b  VSSPEGCQLGPSREGQAYASPG------------APELVELPHADG---------VPLERGVILWMAPD 294
RockbreamIRF4   TTSPEGCRIT SSSSSSPSSSSSSS PCPEDKFHSGAEVILFPFPYPESHRQGAEMLPNVLERGVLLWMMSD 331
                .:*****::  .              .             :. * :.         ****::**: .*

HumanIRF4       GLYAKRLCQSRIYWDGPLALCNDRPNKLERDQTCKLFDTQQFLSELQAFAHHGRSLPRFQVTLCFGEEFP 392
MouseIRF4       GLYAKRLCQSRIYWDGPLALCSDRPNKLERDQTCKLFDTQQFLSELQVFAHHGRPAPRFQVTLCFGEEFP 392
ChickenIRF4     GLYAKRLCQSRIYWDGPLALCSDRPNKLERDQTCKLFDTQQFLAELQAFAHHGRPLPRYQVALCFGEEFP 386
ZebrafishIRF4a  GLYAKRLCQGRVYWEGPLAPYADKPNKLEKEQTCKLMDTQQFLTELQGFIHHGRPMPRSQVILCFGDEFP 402
ZebrafishIRF4b  GLYARRCCPCRVYWTGAHAPPTDKPNKLEREQNCKLLDTHLFITELQSYTLHARPAPCSQVLLFFEDEST 364
RockbreamIRF4   GLYAKRLCQGRVYWEGPLAPYMDKPNKLEKEQPCKLFDTQQFLIELQDFAHNGRHLPRLQVVLCFGDEYP 401
                ****:* *   *:** *. *   *:*****::* ***:**: *: *** :   :.*  *  ** * * :* .

HumanIRF4       DPQRQRKLITAHVEPLLARQLYYFAQQNSGHFLRGY----DLPEHISNPEDYHR---SIRHSS-IQE 451
MouseIRF4       DPQRQRKLITAHVEPLLARQLYYFAQQNTGHFLRGY----ELPEHVTTP-DYHR---SLRHSS-IQE 450
ChickenIRF4     DPQRQRKLITAHVEPMFARQLYYFAQQNSGHLLRGY----DLPELMTSPEDYHR---SIRHSS-IQE 445
ZebrafishIRF4a  DPQRQSKMITAQVEPMFARQLLYFASQTNGHYLRSY----ELQTPGSLPVEDY-----QRSLQHLTE 460
ZebrafishIRF4b  EGQRPRRTYTVQVEPLFARQLLILTHPGSMNYIRSHELQHLPPEHSLSPTQDYHRVITHHHHNSGPQN 431
RockbreamIRF4   DPQRPRKMITAQVEPVFARKLVYYYQQNNGHYLRGYD---HIQEQNTSPAIDYP---SQRPLQHIQE 462
                : **  :  *.:***::**:*      . : :*.:     *   :      :   :  :
```

Fig. 3. Multiple alignment of RbIRF4 with other known IRF4 amino acid sequences. The DNA-binding domain (DBD) and IRF-association domain (IAD) are highlighted. The conserved tryptophan (W) residues in the DBD are shaded in black. The serine-rich domains in the carboxyl terminal region are boxed. Identical (*) and similar (: or .) residues identified by the ClustalW2 program are indicated.

~ 13 ~

```
                    _____

HumanIRF8       MCDRNGGRRLRQWLIEQIDSSMYPGLIWENEEKSMFRIPWKHAGKQDYNQEVDASIFKAWAVFKGKFKEG 70
MouseIRF8       MCDRNGGRRLRQWLIEQIDSSMYPGLIWENEDKTMFRIPWKHAGKQDYNQEVDASIFKAWAVFKGKFKEG 70
ChickenIRF8     MCDRNGGRRLRQWLIEQIDSEQYPGLIWENEEKTMFRIPWKHAGKQDYNQEVDASIFKAWAVFKGKFKEG 70
ZebrafishIRF8   M--NSGGRRLKQWLIEQINSNIYNGLQWEDEDRTMFRIPWKHAGKQDYNQEVDASIFKAWAIFKGKFKEG 68
RockbreamIRF8   M-SNTGGRRLKQWLVEQIQSAQYSGLQWEDESRTLFRIPWKHAGKQDYNQEVDASIFKAWAVFKGKFKEG 69
                *  ..*****:***:***:*  *  ** **::.:.::*****************************:*******

                      _____DBD_____
HumanIRF8       DKAEPATWKTRLRCALNKSPDFEEVTDRSQLDISEPYKVYRIVPEEEQKCKLG----VATAGCVNEVTEM 136
MouseIRF8       DKAEPATWKTRLRCALNKSPDFEEVTDRSQLDISEPYKVYRIVPEEEQKCKLG----VAPAGCMSEVPEM 136
ChickenIRF8     DKAEPATWKTRLRCALNKSPDFEEVTDRSQLDISEPYKVYRIVPEEEQKCKIG----VGNGSSLTDVGDM 136
ZebrafishIRF8   DKAEPATWKTRLRCALNKSPDFEEVTDRSQLDISEPYKVYRIVPEEEQKLGKG------TVTTVKDTTDM 132
RockbreamIRF8   DKAEPATWKTRLRCALNKSPDFEEVTERSQLDISEPYKVYRIVPEEEQKHGKSSVMAMAATTSSGDITDM 139
                **************************:*:*********************     .          :  :*

HumanIRF8       ECGRSEIDELIKEP-SVDDYMGMIKRSPSPP-EACRSQLLPDWWAQQPSTGVPLVTGYTTYDAHHSAFSQ 204
MouseIRF8       ECGRSEIEELIKEP-SVDEYMGMTKRSPSPP-EACRSQILPDWWVQQPSAGLPLVTGYAAYDTHHSAFSQ 204
ChickenIRF8     DCSPSAIDDLMKEPPCVDEYLGIIKRSPSPPQETCRNPPIPDWWMQQPSPSLPLVNGYTGYEQHHSGYSQ 206
ZebrafishIRF8   DCSP-DLDEIIKES-SNDEYMGILRSSHSPLDERSSMPSVQEWWQQGPLNAAVVHQD--PAGSLNSAFSQ 198
RockbreamIRF8   DCSPADLEELIKE-----------------EEGCSIQSSPEYWSQGSINAFPLHQDPLPSGTLSSALSQ 191
                :*.     :::::**                 *  .     ::* *   .  :  .        *. **

HumanIRF8       MVISFYYGGKLVGQATTTCPEGCRLSLSQPGLPGTKLYGPEGLELVRFPPADAIPSERQRQVTRKLFGHL 274
MouseIRF8       MVISFYYGGKLVGQATTTCLEGCRLSLSQPGLP--KLYGPDGLEPVCFPTADTIPSERQRQVTRKLFGHL 272
ChickenIRF8     MVITFFYSGRLVGHITTSYPEGCRLSLSQPSNHGEKLYTPDSLEHVRFPSAEAIQNDRQKQITKKLFGHL 276
ZebrafishIRF8   MLISFYYGGQMVDNMVTTHPEGCRISPCLPSTANGFLYGSDSLQNIYFPSIDGIKNERQRHVTRKLFSHL 268
RockbreamIRF8   MMISFYYGGKLMHNTLVAHPEGCRISPQQ-HLGRGALYSSDSMQCVNFPPAELIEYDRQRHVTRKLLGHL 260
                *:*:*:*.*:::  :  .:  *****:*           **  .:.:: *  **.  *  :**:::*:**:.**

HumanIRF8       ERGVLLHSSRQGVFVKRLCQGRVFCSGN---AVVCKGRPNKLERDEVVQVFDTSQFFRELQQFYNSQGRL 341
MouseIRF8       ERGVLLHSNRKGVFVKRLCQGRVFCSGN---AVVCKGRPNKLERDEVVQVFDTNQFIRELQQFYATQSRL 339
ChickenIRF8     ERGVLLHSNKQGIFIKRLCQGRVFWSGN---TVVYKDRPSKLDRDEVVKIFDTNLFFRELQQYYNNQGRF 343
ZebrafishIRF8   ERGVLLRANREGIFIKRLCQSRVFWIGQ---DARYN--PCKLERDAVVKIFDTARFLQALQLYQDGHYQA 333
RockbreamIRF8   ERGVLVRANQEGIFIKRLCQSRVFWSGLGEVGSQYSPMPCKLERDAVVKIFDTERFLQALQLYQEGQFPA 330
                *****:::.::*:*:*****.***  *        .  * **:** **::*** *:: ** :   :

HumanIRF8       PDGRVVLCFGEEFPDMAPLRSKLILVQIEQLYVRQLAEEAGKSCGAG----SVMQAPEEPPPDQVFRMFP 407
MouseIRF8       PDSRVVLCFGEEFPDTVPLRSKLILVQVEQLYARQLVEEAGKSCGAG----SLMPALEEPQPDQAFRMFP 405
ChickenIRF8     PDSRVMLCFGEEFPDTVPLRCKLILVQVEQLCVRQVMEEAGKTCSS------PMLPDDVQQEQVYRIFQ 406
ZebrafishIRF8   PEPTVTLCFGEEFNDFSTVKSKLIIVEITAWNCQQLLNAVTARRTQ--CSSGNMEISDNLVSDQMACIYQ 401
RockbreamIRF8   PDPTVTLCFGEELHDVSNAKGKLIIVQITVVNCQHLLDAVNMRRTQPFCNNPNLDMSDNVATDQMARIYQ 400
                 *:  * ******: *    : ***:*::    :::  : .              ::   :*   ::

HumanIRF8       DICA---SHQRSFFRENQQITV 426
MouseIRF8       DICT---SHQRPFFRENQQITV 424
ChickenIRF8     DICG---PHQRPLFRENQQIAV 425
ZebrafishIRF8   DLCSYPVPPRASCFRDNLQIPV 423
RockbreamIRF8   DLCSYSGPQRPACYRDNMPITA 422
                *:*   . : . :*:*  *..
```

Fig. 4. Multiple alignment of RbIRF8 with other known IRF8 amino acid sequences.
The DNA-binding domain (DBD) and IRF-association domain (IAD) are highlighted. The conserved tryptophan (W) residues in the DBD are shaded with black. Identical (*) and similar (: or .) residues identified by the ClustalW2 program are indicated.
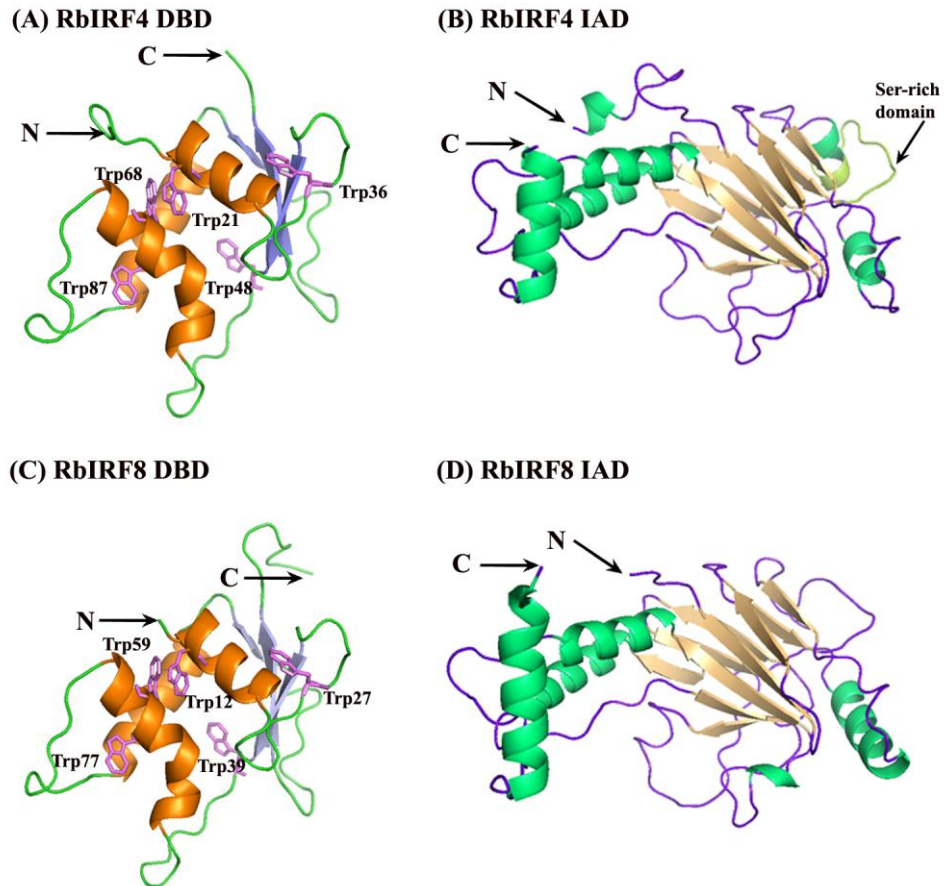
Fig. 5. Comparative 3D-molecular modeling of RbIRF4 (A, B) and RbIRF8 (C, D).
The DNA binding domains (DBD) are presented in **(A, C)** and the IRF-association domains (IAD) in **(B, D)**. The α-helixes and β-sheets are shown in red and yellow, respectively. The conserved tryptophan (W) residues in the DBD are indicated with their corresponding positions. The serine-rich domain of IAD is shown in light green.
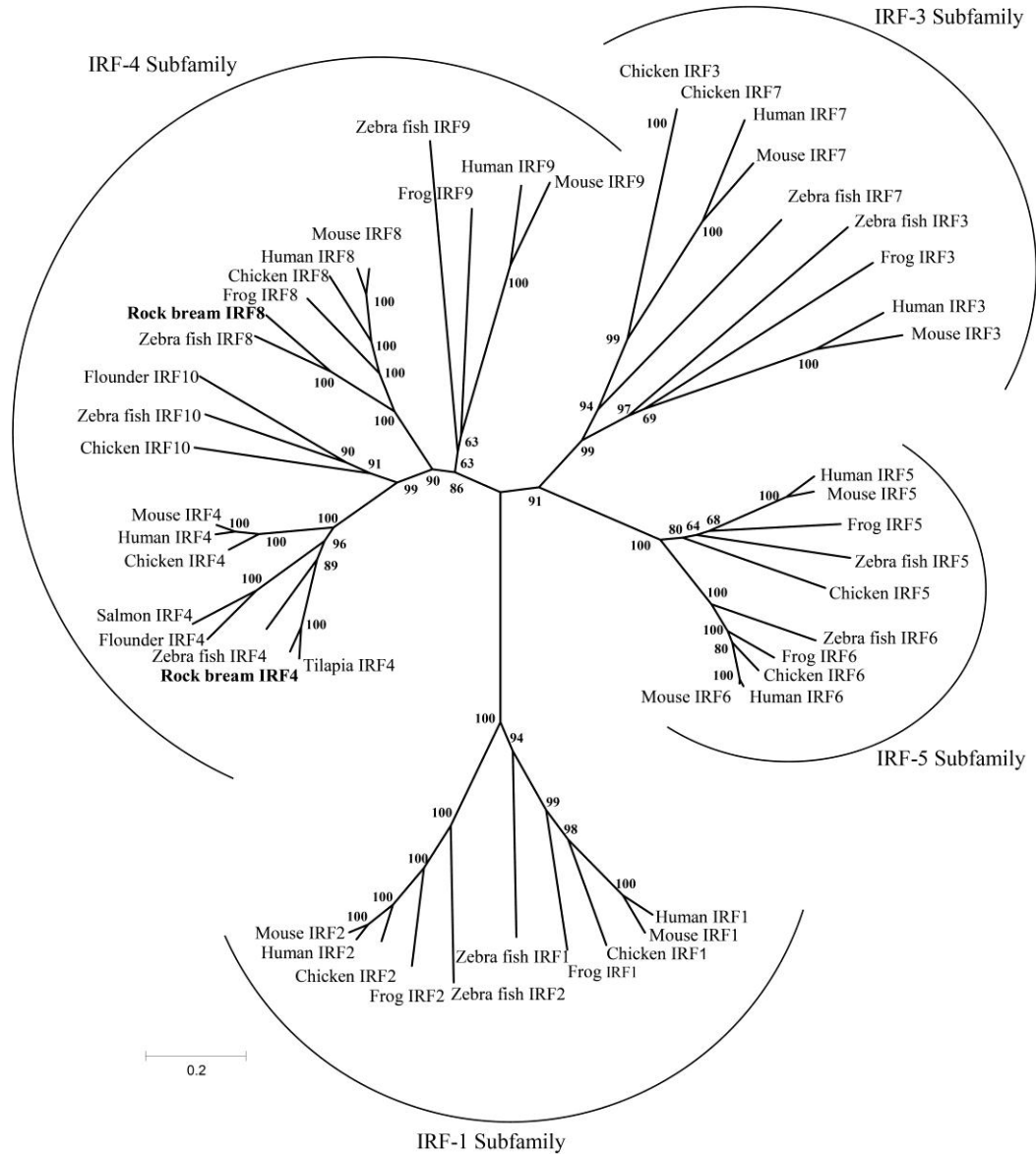
Fig. 6. The neighbor-joining unrooted phylogenetic tree of IRF family members. Bootstrapping was performed 1000 times. GenBank accession numbers of the selected IRF sequences are shown as follows: Human IRF1 (NP_002189), Mouse IRF1 (CAI25046), Chicken IRF1 (NP_990746), Zebrafish IRF1 (NP_991310), Frog IRF1 (NP_001083250), Human IRF2 (NP_002190), Mouse IRF2 (NP_032417), Chicken IRF2 (NP_990527), Zebrafish IRF2 (NP_001008614), Frog IRF2 (NP_001088726), Human IRF3 (CAA91227), Mouse IRF3 (NP_058545), Zebrafish IRF3 (NP_001137376), Frog IRF3 (NP_001079588), Human IRF4 (AAH15752), Mouse IRF4 (NP_038702), Chicken IRF4 (NP_989630), Zebrafish IRF4a (NP_001116182), Salmon IRF4 (NP_001133454), Flounder IRF4 (ADZ96214), Tilapia IRF4 (XP_003437930), **Rock bream IRF4** (JQ388475), Human IRF5 (AAA96056), Mouse IRF5 (NP_036187), Chicken IRF5 (NP_001026758), Zebrafish IRF5 (ABY91289), Frog IRF5 (NP_001088065), Human IRF6 (AEL89176), Mouse IRF6 (AAB36714), Chicken IRF6 (ABB77237), Zebrafish IRF6 (NP_956892), Frog IRF6 (NP_001085345), Human IRF7 (AAI36556), Mouse IRF7 (NP_058546), Chicken IRF7 (NM_205372), Zebrafish IRF7 (NP_956971), Human IRF8 (NP_002154), Mouse IRF8 (NP_032346), Chicken IRF8 (NP_990747), Zebrafish IRF8 (NP_001002622), Frog IRF8 (NP_001087097), **Rock bream IRF8** (JQ388476), Human IRF9 (NP_006075), Mouse IRF9 (AAH12968), Zebrafish IRF9 (NP_991273), Frog IRF9 (NP_001084846), Chicken IRF10 (NP_989889), Zebrafish IRF10 (NP_998044), and Flounder IRF10 (BAI63219).
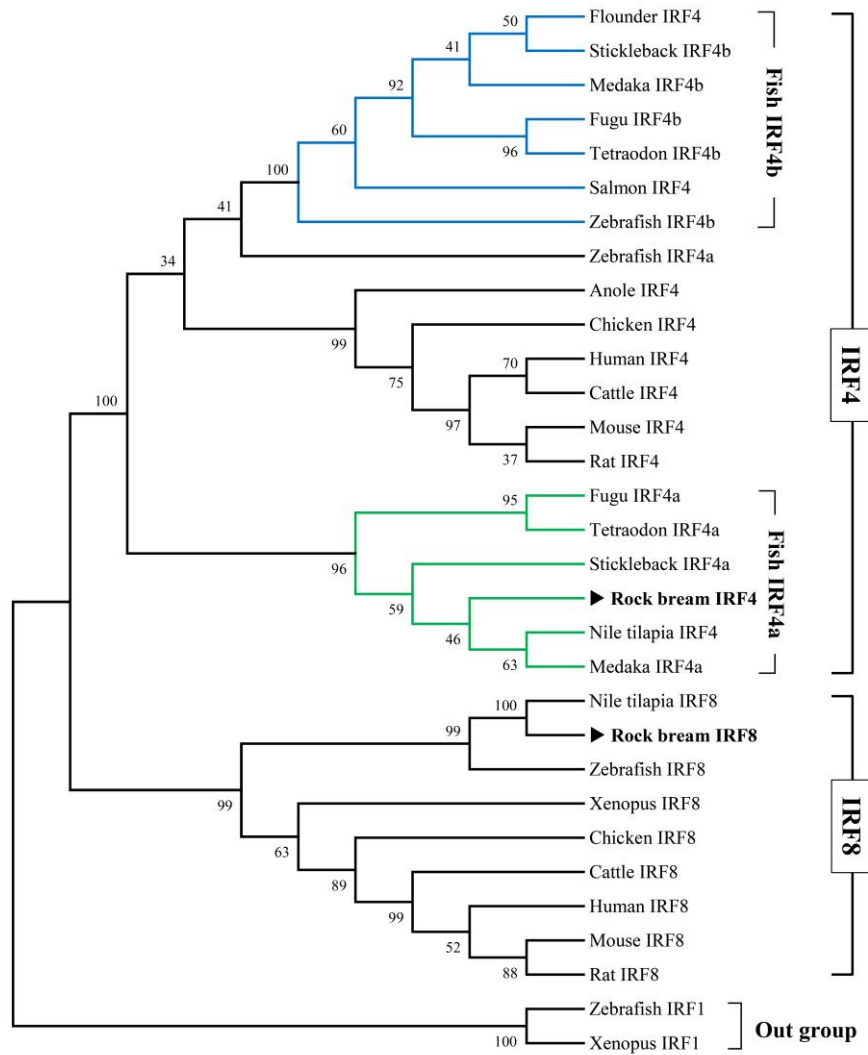
Fig. 7. Rooted phylogenetic tree of Rock bream IRF4 and IRF8 with IRF orthologues. The tree was constructed by the neighbor-joining method with 1000 bootstrapping trials. The numbers at each branch node represent the values given by bootstrapping analysis. GeneBank and Ensembl accession numbers, which are not mentioned in Fig.4, are as follows: Medaka IRF4a (ENSORLG00000017242), Fugu IRF4a (ENSTRUG00000011568), Tetraodon IRF4a (ENSTNIG00000005370), Stickleback IRF4a (ENSGACG00000016461), Medaka IRF4b (ENSORLG00000012712), Fugu IRF4b (ENSTRUG00000002946), Tetraodon IRF4b (ENSTNIG00000008222), Stickleback IRF4b (ENSGACG00000004966), Zebrafish IRF4b (XP_697730), Anole IRF4 (XP_003004441), Cattle IRF4 (NP_001193091), Rat IRF4 (EHB12206), Nile tilapia IRF8 (XP_003442371), Cattle IRF8 (NP_001077238) and Rat IRF8 (NP_001008722). Zebrafish IRF1 (BC115340) and Frog IRF1 (BC075398) were used as an out group.

~ 17 ~

3.2 Genomic structure analysis of RbIRF4 and RbIRF8

The RbIRF4 genomic sequence (GenBank accession no: JQ388477) was ~9200 bp, and consists of eight exons interrupted by seven introns. The genomic structure of RbIRF4 was compared with mammals (**Fig. 8A**) and fish (**Fig. 8B**) species. The ORF of human and mouse IRF4 contain eight exons and lengths of exons in the RbIRF4 ORF are most similar to those species with exons 2 and 7 being identical. In the case of fish, two types of IRF4 have been reported (IRF4a and IRF4b). In comparison of genomic structure of RbIRF4 with other fish species, no clear-cut differences were observed between IRF4a and IRF4b. Exon numbers varied from 7 to 11 and the one preceding the last exon is identical among fish IRF4s except IRF4a of medaka. In contrast, RbIRF8 genomic sequence (GenBank accession no: JQ388478) was ~4100 bp, and consists of nine exons and eight introns. Several vertebrate IRF8 genomic structures were compared to the genomic structure of RbIRF8. The ORFs of all species compared contained eight exons, with very little variations between exon lengths. Interestingly, the size of each exon in rock bream and stickleback IRF8 were exactly the same in the ORF region. In addition, the first exon of the RbIRF8 belongs to 5′-UTR (**Fig. 8b**).
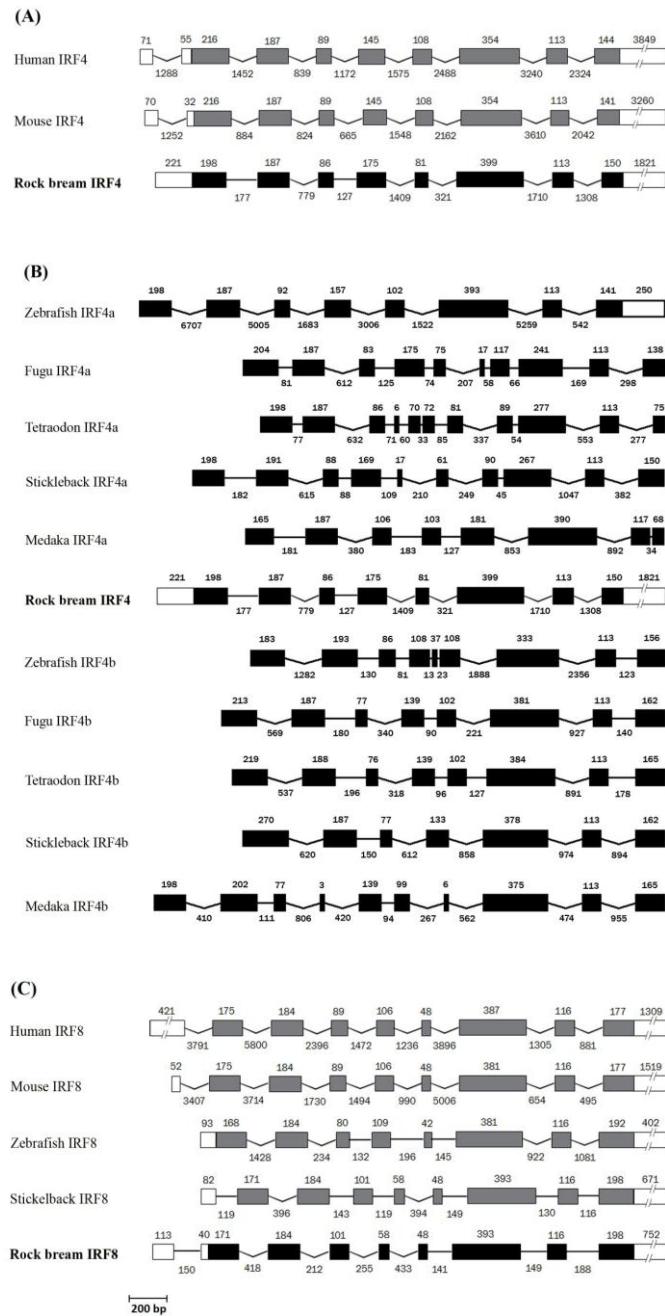
Fig. 8. Comparison of genomic structure of RbIRF4 with mammals (A), RbIRF4 with fish (B) and RbIRF8 (C) with IRF orthologues. Genomic structures were obtained from the Ensembl database, and include: Human IRF4 (ENSG00000137265), Mouse IRF4 (ENSMUSG00000021356), Zebrafish IRF4a (ENSDARG00000006560), Fugu IRF4a (ENSTRUG00000011568), Tetraodon IRF4a (ENSTNIG00000005370), Stickleback IRF4a (ENSGACG00000016461), Medaka IRF4a (ENSORLG00000017242), Zebrafish IRF4b (ENSDARG00000055374), Fugu IRF4b (ENSTRUG00000002946), Tetraodon IRF4b (ENSTNIG00000008222), Stickleback IRF4b (ENSGACG00000004966), Medaka IRF4b (ENSORLG00000012712), Human IRF8 (ENSG00000140968), Mouse IRF8 (ENSMUSG00000041515), Zebrafish IRF8 (ENSDARG00000056407), and Stickleback IRF8 (ENSGACG00000015958).

3.3 Promoter sequence analysis of RbIRF4 and RbIRF8

We obtained 2139 bp and 1836 bp of the 5′-flanking region sequence of RbIRF4 and RbIRF8, respectively, from the BAC clones. The TIS was predicted in each, along with several putative transcription factor binding sites for factors known to be related to immune responses. Pathogen associated molecular pattern (PAMP)-activated transcription factor binding sites were found in both RbIRF4 (**Fig. 9**) and RbIRF8 (**Fig. 10**), and included sites for activator protein-1 (AP-1), CCAAT-enhancer binding protein (C/EBP), C/EBPα, C/EBPβ, cAMP response element-binding protein (CRE-BP), signal transducer and activator of transcription (STAT), nuclear factor kappa-light chain enhancer of activated B cells (NF-κB), GATA-1, CdxA, SRY, acute myeloid leukemia 1a (AML-1a) and the TATA-binding protein, along with ISREs. In addition, a unique IRF1 binding site was found in the RbIRF8 promoter region. Canonical TATA boxes were found in close to the TIS in both RbIRF4 and RbIRF8 promoters.

```
GATCAGAAACCTGGATCATTTCTGAAACCAGGATATGATAGTTACATGTGCAAAACATAACCAGAATACACAAAAAAACCCAATTAAAAC     -1991

CGGCATGTTGACACGCTCATACATCTAAAAGTGTTGAGGAAAAGGGATGCCAATCATGTAGTTTAAAAGTTGAATTTCATGCGTGATACT     -1901

CTTTGTCCATCCCTCCCTTCTTACAGTTTGGCAGGCGTGTCCCGCAGTGTCACCCTGGTGGTAGCTTACATCATGACGGTGACGGGACTG     -1811

GGCTGGCAGGAGGCGCTGGCTGCAGTGAAGGTGGCCCGGCCCTGCGCTGGCCCCAACCTGGGCTTTCAGCGCCAGCTCCAGGAGTTTGAG     -1721

GCTACTCAAGCTGATCAGGTCAGTCCTGGGAACTGTAGTAGGACAGAGTGCAGAACATATCACTGTATTACTGAATGTGGAGAGCACACC     -1631
                         STAT
ATCAGTACTTTGCAATTTGTAAAATTTTTTACATTGTCATTAAATGACTTCCTTCTCCTGCAGTTCAGAGAATGGCTACAGATGGAATAT     -1541
                                                                  STAT
AAGGACAACCCCTTCAGTGACGAGGCTGATATACGTGACTTGCTCGCCAGGGCGTCTAAAGTCAATGGTAAGGAGGTGGGAAAACGGGCA     -1451

TCTACCCCACCTGGAGGTATCTGATCAGATCTTTTGATAAGATACTTTTGCAGCTCAGCCTGGAGGATTTCACCACCGCGTGATGCTACA     -1361
                                                            C/EBPβ
GTATTCTGGACCTATCTGGAAGTCTAATGTGCTCTGGAGTGGACTAATCCAGGTGTTTCTCTCCGCTGGAATGATCCTCTAAATATCATG     -1271
            GATA-1
GACATGAACAGTGAACGAGACAATTAATATGCACAGCACATATTAATTATTTTTTAGCATTGTTCGATGTTTCTGGTGCTTGTTTTTTTT     -1181
                      CdxA             CdxA
TTTTTCTTTTTTTTTTCTAATTTATGAATGTTGTGAACTTTTTGTACTGTGTCACCATGATTTCAAAAAGTGAAATAAAAGGAGTGAAGT     -1091
                   C/EBPα                                            ISRE
CAAGCGCTCTGTGAGGTTGTCGGTCGGTGGGCTGCCTTATTGAAGGTGATAAGACACGTTCTGTGCTGTGTTTACGTCTCTGTTTCCTGG     -1001
CRE-BP
TTGCCTTCCTGTTTGTCTAATGTGCAAGGGAGGTCATGCATGGTGACATGAAAAACACGTTCTTGTATTGCTCCCTCTTCTCAATTATCT     -911
        C/EBPα
TTAGAGACTGTAAAGCATTAAAAAAAAAATAAGATATTGATGAGATATTGATATTAATGAGATATTACATACAGTAGTCTTCTGTAGATAG     -821
                                                       CdxA       TBP
CTGAGCATGTGCATGGATAGCCACTGGAGGGCAGTAACAAAGTAGATTAAAGTACATAGTATTTTGAATTCAATCATGTGTCTTCTTCAG     -731

ACACTGGTTCCTCATGCCACATGAGTCAGGAATCAATAACAAAATTAAGACTTTATAGAAATATTTAGTCAGTAGTGCCATATCACTGGA     -641
                     AP-1                                      AP-1
AATTAGATTGAATTCACAGGCTTTTAAGAAAAACAAAAACATGTGGTGGAAGATTTTGGTGTTGGTAATGTGCACAGGTAATTGTTCCAT     -551
                            SRY       AML-1a
TCCGTCCCGGTGCCTGCTCGGGGTGTTCACCAGAGGGATTCCCCACCGCAGCGCTGTGTGGGCGCCGAGTTTCACAGAAATGAGCGCCTG     -461
                                    NF-κB
ATTGCTGCGCATCTCCTTGTGAACTTCTCAGAGCCCGAGAGAGGCTCCTCCCCGTCACAAATCAAGATGCGATAACGTTTAAGAATAGGA     -371

CGCAGAGAGCTTGGAAACAGATTTCAGCGGGGGAGCTCACATGATCAAAGGTAAGCGGGAGCAGCTCCCTTTTTTTTTTTTTTTGTAAATG     -281

AATGACGGCATTTTCAACATCACAATTAATGAAAACTTTGAATGCATCAGTCAAGCCCAGCAGCTTCAGCCACTCTCCTCTGTGTGGAGT     -191

TTGACAGGTGGAGTCTCAAGATAACATTAACTCAGTGTGATCAGAAGTTTGTCACTCAACTTACAACTTCTAGTAGTTATAAAACTTGAT     -101
                                                                       TATA BOX
CAGTTAAAGTCTGATTAAAGTGGAGCTGATAACATGAAGTTTGAGAAACGTGATAAGGTGCCATTAGCATATATTTCACCGTGCATTGCA     -10
      +1                                 C/EBPβ
GTTGTGTTCAAACTCTTTCTTGCCATGTAATTATGGCTACATGACTGGTTGTGTTTTTATTCATGCATC                          60
```

Fig. 9. Sequence of the 5′-flanking region of RbIRF4. Putative transcription factor binding sites are in bold and underlined, with the corresponding factor indicated directly below. ISRE is highlighted in gray. The putative TATA box is boxed, and the predicted transcription initiation site is indicated by a bent arrow.

```
AGTGGCTTGATATCAATGAATGTAGGTGGATAGCCGGGGAGGAAAAGCTGTGAGTTGGCTCAGTCAGTTCCCAGCAACTCTGCCCCGGGC    -1699
                                                            AP-1
CAGCAAAACTGTCGCTGACTTATATCTTATCAGGCGGAGAGGACAAAAGGAAAAGAAGCACAGTACAACAGCCAAACCCCCTGTAGTTTG    -1609
    ISRE              AP-1
GCCACCAACAACTGATGTCAGATAAGCCATGAGTGGGATTTGTTAGTAAACTAGGCTGCTTCAGTTGGCTACAGAGCTGTGAAGAACATT    -1519
                    GATA-1
AGTTTGTACCGTGGCCATCTAAAGAGGAGTAACATATAGAGTAAGTTTCATAATTGGCAAAATGTTCCCTTAAGTAATACACAGCAGGTT    -1429
                                                            STAT    CRE-BP
AAATTGAGAAGCTACATAAAGTATGGCTGAGTAAAACTGTAACTACACATGAAGGATCTCTGCAGTCCGAATGCTTAAGTCTTCCTGATT    -1339
                            ISRE
TCACTGAAGCTCTGATTTATGTTTTTATTGTGACAACGCTGATTTTTAAAAAAAATTTTTTTAGGTAATTTGTGTTATAAATTCACGGGT    -1249
                CdxA                                          C/EBPα
GCAGTTAACATTTTTAAATTGCGTATTTATGGTGTTTTTTGTTTGTGTTTTATAATAAAATGCAGTGATTGACAAAAAAAAATATGGAAAT    -1159
                        CdxA              TBP
GCTAATAAATCTGTACTATCAACAGCATGGTATGTGTTTGTGTGTAAGTAATGCATGGCTAGTATTGCAGCACATGGCTCACCTTTAGAT    -1069
                                SRY                          C/EBPβ
CAGCTGCTCGTCTCATGAAGCTGCTTCATAACCGCACTGAGTAAACTCCAGAATAACCTTTCTTGAATGTCAGTGTTCGAGATTTGAGAG    -979
                                        AP-1
GTTTCTTGGTTTTACTCAGCAGATCACCTGGTAATGCTGCTTTATACTTTTTACACCACACACTTTGTGTGCAGCTTCATGAAACTAACCA    -889
                                                    AML-1a
CAATGTAGTATGTTTGTGAAAGACACAGATGAGGACAGACCTGTATTGTAGTAGACCATCAATAAGACATTTAATTTCACACTGTATACA    -799
                SRY
AATATATAATAAGACTAAAGCGAAAGAATTGCCTCATAAAATGGAGATCAAACTGTATAAGTGAACTTAAAATAGTATTTAATGACAAAA    -709
                IRF-1
TGATCATTGATTACAGGAGAATTTTAATTTAAAGATACAGTTAGCGGATGATATCGCTCACCAGTTGTGTATTAGGGAAAGATTACTAAT    -619
            NF-κB
ACAGTGCATTTGAATAGCTCTTACTCACAGGTACAGTCAAATAATATTCACATACTGTAAAAAACATTATTCCACTGTAGGCCTGCCTAA    -529
                    AP-1
ATCAACATCACAATTCATATGTTTTCATACTTCAAACATTTACAAATATTTCCCCCATAACCATGATTCCAAATCAGTCAGGTGTAGGCA    -439
                                                            AP-1
TGTTCCCACTCAGCTCATTCCTGGCATGACATCGCAAGAGGTGGTGTTTACACTGTAAATACAAACAGAAAGCATAAAGTCAGACTTACT    -349
                                                            CRE-BP
ATAAATGTAGCTAAATGTAATACTGAAAGGAAATGCAGTTGGAATAAAGAAGAGACTTTCATTTGTTCAGTGTAATCTGTTTAATTATCA    -259
                                        ISRE
AGAAATATCAAGTTAATATAACAAAGATTATAAAAGGCCTATTGAATTGAATGAAATTAATTATTGTTTCAAGGGCTTTTTTTGACAAAT    -169
GATGAACTCTTGATTTTATATGGCATCGGGCACTATACGTGTATCAGCAAAACCATCAGTCTACGGCTGCAACCAACAATTATTTTCATT    -79
                                                                                  +1
ATTGATTAATCTGCAGATTATTTTCTCAATCAATGGATTTATTTGGTCTATAAAATGTCAGAAAATGGTGAAAAATGCCCGTTACATTTT    12
  C/EBPα                                         TATA BOX            C/EBPβ
CCTAAAAACCCAAGCTGATGTCAACGACCAACAGTC                                                        48
```

Fig. 10. Sequence of the 5′-flanking region of RbIRF8. Putative transcription factor binding sites are in bold and underlined, with the corresponding factor indicated directly below. ISRE is highlighted in gray. The putative TATA box is boxed, and the predicted transcription initiation site is indicated by a bent arrow.

~ 22 ~

3.4 Gene expression patterns of RbIRF4 and RbIRF8 in healthy rock bream

To examine the tissue distribution pattern of RbIRF4 and RbIRF8, transcript levels were determined by qRT-PCR in blood, gill, liver, spleen, head kidney, kidney, skin, muscle, heart, brain and intestine of healthy individuals (**Fig. 11**). The expression level for each of the tissues examined was normalized to that of the rock bream β-actin gene. The results were further compared with muscle expression levels to determine the relative tissue-specific profile. RbIRF4 and RbIRF8 were found to be constitutively expressed in all the tissues examined, although the level of expression varied among each. Detailed analysis revealed that RbIRF4 transcription was significantly higher in spleen, blood, head kidney, liver and kidney, was moderately high in heart, skin, intestine and gill, and was poor in muscle. The transcription of RbIRF8 was significantly higher in spleen, liver, kidney, head kidney and heart, was moderately high in blood, gill and intestine, and poor in all other tissues.
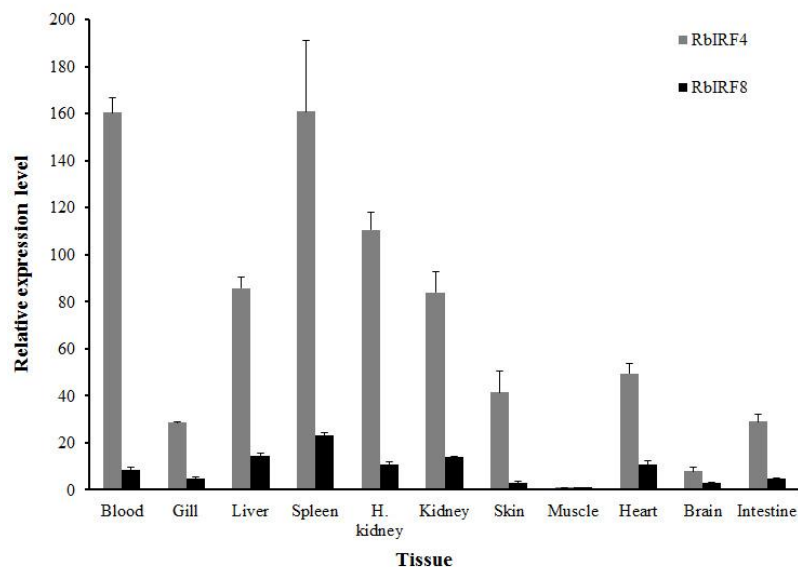

Fig. 11. Tissue-expression analysis of RbIRF4 and RbIRF8 mRNA. The expression was detected by qRT-PCR and normalized to Rbβ-actin. The data is presented as mean ± SD (N=3).

~ 23 ~

3.5 Immune-responsive expression of RbIRF4 and RbIRF8

To evaluate the variations in RbIRF4 and RbIRF8 transcript levels in response to bacterial- and viral-based immune challenges, spleen and head kidney tissues from challenged fish were investigated. The relative mRNA levels of RbIRF4 and RbIRF8 were calculated using β-actin as the internal control and normalized with the respective PBS-injected controls for each time point examined.

The expression profile of RbIRF4 in head kidney after injection of bacterial or viral substances is shown in **Fig. 12A**. The level of IRF4 transcription was significantly up-regulated (2.7-fold; $p<0.05$) at 24 h p.i. of poly I:C, but down-regulated at 6, 12 and 48 h p.i. After injection with RBIV, the highest expression level of IRF4 occurred at 12 h (1.7-fold), and then gradually decreased. After injection with LPS, IRF4 showed no significant up-regulation, but did show significant down-regulation at 6 and 48 h p.i. Stimulation with *E. tarda* and *S. iniae* caused IRF4 to be significantly up-regulated, with the peak levels occurring at 24 h (31.6-fold; $p<0.001$) and 24 h p.i. (3-fold; $p<0.001$), respectively.

The expression profile of RbIRF4 in spleen after injection of bacterial or viral substances is shown in **Fig. 12B**. The level of IRF4 mRNA was not significantly different after challenge with poly I:C, except at 12 h p.i. when it was significantly down-regulated ($p<0.05$). In response to RBIV injection, the IRF4 expression level peaked at 12 h p.i. (2.1-fold). LPS injection resulted in only down-regulation of IRF4. *E. tarda* injection, however, caused the transcript level of IRF4 to be significantly up-regulated at 24 h p.i. (2.4-fold; $p<0.001$), while *S. iniae* injection caused significant up-regulation at 6 h (1.6-fold; $p<0.05$) and 24 h (1.7-fold; $p<0.05$) p.i. Furthermore, both of these bacterial challenges caused significant synchronized down-regulation at 3 h and 48 h p.i. ($p<0.05$).

The expression profile of RbIRF8 in head kidney after injection of bacterial or viral substances is shown in **Fig. 13A**, and revealed a similar pattern to that of RbIRF4 in head kidney for all challenges. The level of IRF8 was significantly up-regulated by poly I:C injection ($p<0.05$), except at 6 and 48 h p.i. After injection with RBIV, IRF8 exhibited down-regulated expression, except at 12 h p.i. LPS challenge up-regulated expression only at 3 h p.i. (3-fold). *E. tarda* injection caused a gradual increase in IRF8 level, which peaked at 24 h p.i. (17.7-fold), while *S. iniae* caused the highest levels of up-regulated expression 6 and 24 h p.i.

The expression profile of RbIRF8 in spleen after injection of bacterial or viral substances is shown in **Fig. 13B**. After injection with poly I:C, the expression level of IRF8 significantly increased at 6 h p.i. (~2-fold; $p<0.05$). Injection of RBIV led to significantly increased IRF8 transcription at 12 h p.i. (3.1-fold; $p<0.001$). A slight increase in IRF8 transcription was observed at 3 h (1.8-fold) after challenge with LPS. Challenge with *E. tarda* and *S. iniae*, however, led to significant up-regulation at 6 and 12 h p.i., and down-regulation at 3 and 48 h p.i.
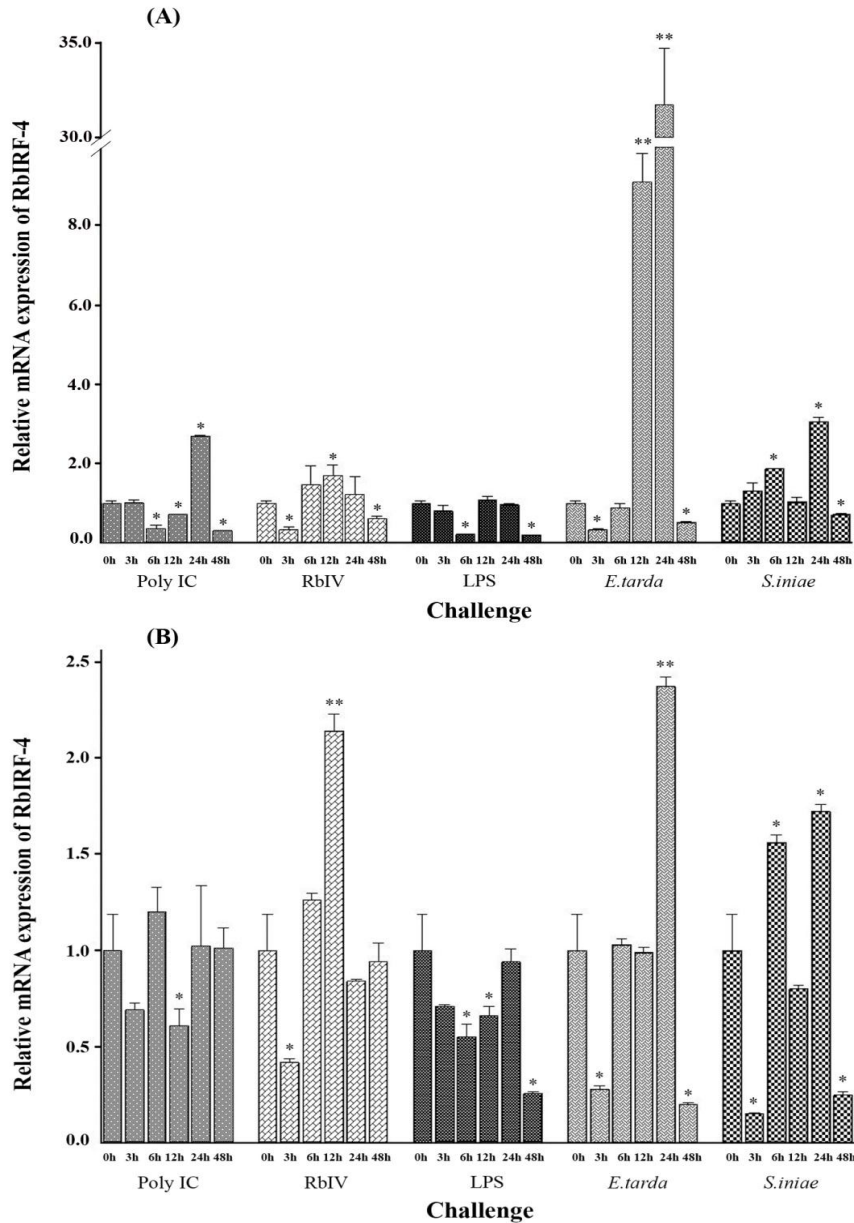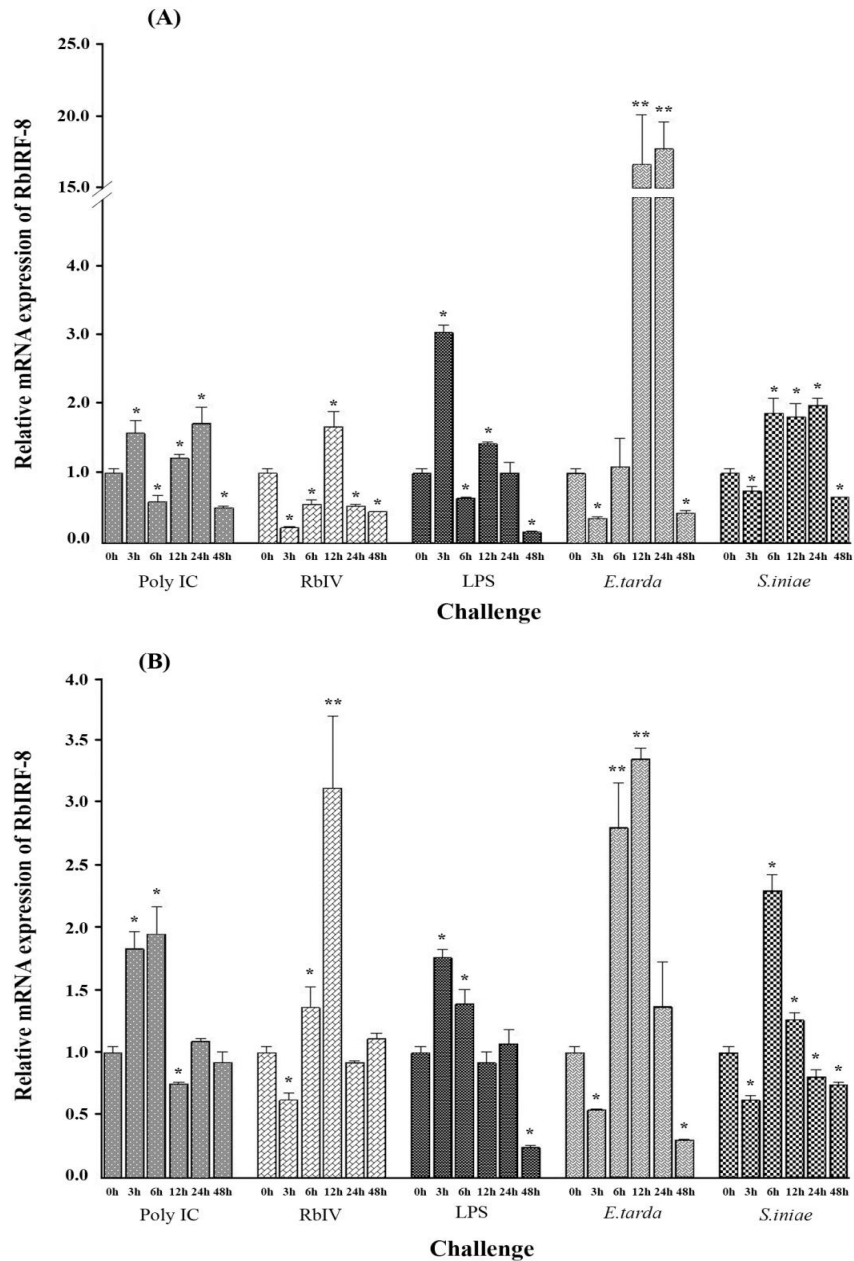
Fig. 12. qRT-PCR analysis of IRF4 transcription in head kidney (A) and spleen (B) in response to poly I:C, rock bream iridovirus (RBIV), LPS, *E. tarda,* and *S. iniae*. The expression was normalized to Rbβ-actin, and is presented as fold-change as compared to the normalized IRF4 in respective controls (PBS-injected). Expression >1 indicates up-regulation, and <1 indicates down-regulation. The data is presented as mean ± SD (N=3). Statistical comparison of the mRNA levels detected at different time points was carried out by the Student's *t*-test. * *p<0.05 vs.* 0 h, ** *p<0.001* vs. 0 h.

Fig. 13. qRT-PCR analysis of IRF8 transcription in head kidney (A) and spleen (B) in response to poly I:C, rock bream iridovirus (RBIV), LPS, *E. tarda*, and *S. iniae*. The expression was normalized to Rbβ-actin, and is presented as fold-change as compared to the normalized IRF4 in respective controls (PBS-injected). Expression >1 indicates up-regulation, and <1 indicates down-regulation. The data is presented as mean ± SD (N=3). Statistical comparison of the mRNA levels detected at different time points was carried out by the Student's *t*-test. * *p*<0.05 *vs*. 0 h, ** *p*<0.001 vs. 0 h.

4.0 Discussion

IRFs are well-characterized transcription mediators that play major roles in immune response, especially against viral infections, in apoptosis, and in cell growth. IRF4 and IRF8 have been extensively studied in mammalian species and are known to modulate anti-viral and anti-bacterial activity; however, information related to their functions in fish is very limited. To obtain molecular level information of IRF4 and IRF8 in teleost fish, we identified and characterized the genome structures and mRNA expression of IRF4 and IRF8 from rock bream.

Both the IRF4 and IRF8 contain two conserved IRF domains, the N-terminus DBD and the C-terminus IAD. The DBD is typical of all IRF family members and mediates binding with the ISRE/IRF-E consensus sequence in the target promoters (Paun and Pitha, 2007). Similar to other vertebrate IRF4 and IRF8, both RbIRF4 and RbIRF8 possess five highly conserved tryptophan residue repeat motifs in the DBD. Two types of association modules for the C-terminal region have been proposed (Takaoka et al., 2008): (1) IAD 1, which was initially found in IRF8 and is conserved in all IRFs, except for IRF1 and IRF2 (Sharf et al., 1997); and (2) IAD 2, which is present only in IRF1 and IRF2. In most cases, these protein complexes enhance the ability of IRF to bind with target DNA sequences, such as ISRE. For example, IRF8 protein complexes with either IRF1 or IRF2 have stronger binding activity to ISRE (Bovolenta et al., 1994; Sharf et al., 1997). IAD 2 of IRF1 and IRF2 is an independent module for this interaction with IRF8. The IRF8/IRF1 complex generally functions as a suppressor of transcription. This may be the case with RbIRF8, but further studies are required. IRF4 and IRF8 are both known to interact with PU.1. As described in the introduction, this interaction allows them to bind with the immunoglobulin light-chain enhancer λB (Brass et al., 1996; Eisenbeis et al., 1995) for activate gene transcription. In addition, the RbIRF4 harbored a serine-rich domain, which

is known to mediate IRF4 functions in viral clearance (Sun et al., 2007). Taken together, these findings support the notion that RbIRF4 and RbIRF8 may function as immune-responsive IRFs in rock bream.

Multiple alignments of fish with higher vertebrate IRF4 and IRF8 sequences revealed the high sequence homology that exists in the DBD and IAD, and suggested that the functions of IRF4 and IRF8 may have also been conserved throughout the vertebrates. In addition, multiple alignments revealed that IRF4a and IRF4b of zebrafish have sequence differences in IAD region although they conserved in DBD region. Phylogenetic analysis of vertebrate IRF family members showed that IRFs diverged into four subfamilies, including the IRF1, IRF3, IRF4 and IRF5 subfamilies, among which the IRF3 and IRF5 subfamilies were shown to be closely related. In the phylogenetic tree, the RbIRF4 and RbIRF8 clustered within the IRF4 subfamily. Phylogenetic analysis of the IRF4 subfamily demonstrated that IRF4 and IRF8 split into two groups, wherein RbIRF4 and RbIRF8 exhibited the closest relationship with other teleost species. Moreover, IRF4 subfamily further split into three sub groups including fish IRF4a, fish IRF4b and IRF4 from other vertebrates. In addition, these results indicated that RbIRF4 belongs to the IRF4a sub family.

Herein, we report the genomic structure analysis of fish IRF4 and IRF8 for the first time in rock bream. The genomic DNA of RbIRF4 was shown to be composed of eight exons, and to have a structure that is similar to other mammals IRF4s. The first exon of RbIRF4 contains a 221 bp length 5′-UTR; in contrast, the first exon of the human and mouse IRF4s are interrupted by an intron. Within the ORF region the size of the second and seventh exons of RbIRF4 were found to be identical to other mammals. However, based on IRF4 genomic structures of fish species, we could not categorize them into two subgroups clearly as found in phylogenetic tree (IRF4a and b), since they have different

제주대학교 중앙도서관
JEJU NATIONAL UNIVERSITY LIBRARY

numbers of exons and there is no relationship between the sizes of those exons. The gene structure of RbIRF8 is composed of nine exons. Within the RbIRF8 ORF region, the sizes of the second and seventh exons were exactly same among all vertebrate homologues. These results suggest that the structural pattern of vertebrate IRF4 and IRF8 are evolutionarily conserved, leading us to speculate that an analogous functional pattern may exist for fish IRF4 and IRF8 as in mammals.

The above hypothesis was bolstered by the fact that our sequence analysis of the 5′-flanking regions of RbIRF4 and RbIRF8 harbor several putative binding sites that are shared with the mammalian homologues. Both of the RbIRFs contained binding sites for AP-1, GATA-1, STAT, CRE-BP, C/EBPα, C/EBPβ, TBP, CdxA, SRY, AML-1a and NF-κB, and ISREs. In addition, an IRF1 transcription factor binding site was found in the predicted promoter region of RbIRF8. All these transcription factors have been previously reported as functionally related to immune responses to bacterial/viral infection, either directly or indirectly, in mammals (Casals-Casas et al., 2009; Collet et al., 2009; Darnell, 1997; Liljeroos et al., 2008; Ramji and Foka, 2002; Skjesol et al., 2010; Taniguchi et al., 2001; Zhang and Gui, 2004). This strongly suggests that the RbIRF4 and RbIRF8 may also play pivotal antimicrobial roles in rock bream fish.

Analysis of qRT-PCR detected RbIRF4 and RbIRF8 transcripts in all tissues tested from healthy rock breams. Both IRFs were strongly transcribed in lymphomyeloid-rich tissues, such as spleen, head kidney, kidney, liver, blood, heart, intestine, and gill. Rainbow trout IRF4 and IRF8 have also been detected in lymphomyeloid-rich tissues (Holland et al., 2010). In that study, however, trout brain and muscle showed higher expression levels of both the IRFs (Holland et al., 2010). The low expression level detected in corresponding rock bream tissues led us to speculate that fish IRF expression may be both tissue- and species-specific.

In this study, we performed viral and bacterial challenges to study the expression responsiveness of RbIRF4 and RbIRF8. We selected head kidney and spleen of experimentally-infected fish to conduct this study, since they represent myeloid and lymphoid tissues and demonstrated significantly higher IRF4/8 expression in healthy rock bream. Poly I:C, a synthetic double-stranded RNA, is a well-established inducer of cytokines and antiviral IRFs, such as IRF1, IRF2, IRF3, and IRF7. Moreover, poly I:C has been identified as a bona fide PAMP of fish Toll-like receptors (TLR)3 and TLR22 (Matsuo et al., 2008). In the current study, poly I:C challenge led to significant up-regulation of RbIRF8 in both head kidney and spleen. In contrast, poly I:C challenge only affected RbIRF4 at 24 h p.i. in head kidney. A previous study using trout splenocytes had reported that poly I:C up-regulated IRF8, but had no effect on IRF4 (Holland et al., 2010). Mammalian studies also demonstrated that poly I:C induced up-regulation of IRF8, along with the other antiviral IRFs (IRF1, IRF5, IRF7 and IRF9), but had no effect on IRF4 (Demoulins et al., 2009).

To date, very few reports on the virus-induced expressions of fish IRF4 and IRF8 have appeared in the literature. In our study, RBIV, a well-known viral pathogen of rock bream, led to significant up-regulation of both RbIRF4 and RbIRF8 transcription in head kidney and spleen. Specifically, RbIRF4 and RbIRF8 exhibited the highest RBIV-induced expression levels at 12 h p.i. in both head kidney and spleen, suggesting that these RbIRFs may have highest antiviral activity in the mid-experimental infection phase. According to the observed expression levels of RbIRF4 and RbIRF8 in head kidney and spleen, it could be further suggested that these RbIRFs may play substantial roles to protect the host from double-stranded DNA viruses. Several previous studies have demonstrated the antiviral activity of IRF4 and IRF8 in vertebrates, but not in fish. For instance, two oncogene viruses, human T cell leukemia virus type I (HTLV-I) and Epstein-Bar virus (EBV) have

been shown to activate the NF-κB pathway and consequently elevate IRF4 expression (Sharma et al., 2002; Sharma et al., 2000; Xu et al., 2008). In addition, IRF8 has been shown to play a vital role in host defense against West Nile Virus (WNV). IRF-8 knock-out mice exhibited increased lethality following intranasal challenge with low dose WNV, as compared to similarly infected wild-type animals (Minten et al., 2011). The results of our study indicated a coordinated expression pattern for these RbIRFs, suggesting an essential antiviral role in rock bream.

LPS, an eminent endotoxin, is the major component of the outer membrane of Gram-negative bacteria. Structurally, it is composed of a core polysaccharide and an O-polysaccharide of variable length, as well as a lipid portion termed "lipid A" that is responsible for activating the innate immune response in mammals (Bishop, 2005; Raetz and Whitfield, 2002). In the current study, we observed that the RbIRF8 mRNA expression level was significantly up-regulated at 3 h p.i. with LPS in both head kidney and spleen; however, the RbIRF4 transcript level was significantly down-regulated after challenged with LPS. Our results were in accordance with a previous finding of down-regulated expression for both IRF4 and IRF8 in trout splenocytes upon stimulation with LPS (Holland et al., 2010). Mammalian studies demonstrated that both IRF4 and IRF8 were up-regulated by LPS in purified B/T cells, macrophages, and dendritic cells (Gautier et al., 2005; Gauzzi et al., 2005). It has been reported that lower vertebrates, such as fish, may be resistant to the toxic effects of LPS (Rodkhum et al., 2006). Furthermore, fish lack the essential co-stimulatory molecules for LPS-induced immune activation via TLR4 (i.e. myeloid differentiation protein 2 and CD14), and some fish lack the TLR4 orthologue (Sepulcre et al., 2009). Thus, it is possible that the recognition mechanism of LPS is different in fish and mammals, providing a putative explanation for the observed down-regulation of RbIRF4 in response to LPS challenge.

Several studies have reported that bacterial infection has emerged as the major threat and caused mortality of rock bream in South Korea (Park, 2009). Therefore, in our study, we challenged the rock bream with *E. tarda*, a Gram-negative bacterium responsible for edwardsiellosis, and *S. iniae*, a Gram-positive bacterium responsible for streptococcosis. These bacterial challenges produced the most robust up-regulation in head kidney and spleen for both RbIRF4 and RbIRF8. Interestingly, *E. tarda* produced the highest up-regulated expression levels of both RbIRFs in head kidney and spleen at 24 h p.i., except for IRF8, which showed the highest expression level in head kidney at 12 h p.i. Similarly, both RbIRFs exhibited the highest expression level at 24 h p.i. with *S. iniae*, except IRF8 in head kidney, which showed the highest expression level at 6 h p.i.

The suppression of transcription at very early-phase, when fish were challenged with live pathogens (**Fig. 10, 11**) may be associated with the immune suppressive capability of the pathogens by evasion mechanism(s) which primarily interfere and block the IFN-signaling cascade (Perdiguero and Esteban, 2009; Ting et al., 1999). Mammalian IRF4 had been shown previously to be up-regulated in mammalian antrum infected with *Helicobactor pylori*, a prevalent Gram-negative bacterium (Hofman et al., 2007). Mammalian IRF8 was also shown to act as a key regulator of host defenses against both Gram-negative and Gram-positive bacteria, such as *Mycobacterium bovis*, *M. tuberculosis*, and *Salmonella typhimurium* (Alter-Koltunoff et al., 2008; Marquis et al., 2011). In addition, macrophages from IRF8-deficient mice were found to remain immature, with altered expression of intrinsic macrophage anti-microbial defenses (Tamura et al., 2008), and to be susceptible to *ex vivo* infection with *M. bovis, S. typhimurium* (Alter-Koltunoff et al., 2008), and *Legionella pneumophila* (Fortier et al., 2009). Hence, our findings, together with these analogous results, suggest that IRF4/8 of rock bream may play essential role(s) as antimicrobial defense agents by activating the IFN-mediated

downstream pathways. However, further analysis required to be performed to investigate the immunological pathways and understand the responses of RbIRFs against microbial infections.

In summary, we have identified and characterized the complete genomic sequences of IRF4 and IRF8 from rock bream, and predicted their associated promoter regions for the first time in fish species. Although these RbIRFs were ubiquitously expressed, the transcripts were rich in lymphomyeloid-rich tissues. Upon challenge with virus and bacteria, both RbIRF genes were significantly up-regulated in head kidney and spleen. Indeed, RbIRF4 and RbIRF8 were highly transcribed in response to the bacterial challenge, suggesting that these genes are highly responsive to bacterial infection. Together, the structural and expression properties of RbIRF4 and RbIRF8 suggest that they may play a role(s) in antiviral and antibacterial defense responses in rock bream, similar to that demonstrated for their mammalian counterparts.

CHAPTER 2

Three complement component 1q genes from rock bream, *Oplegnathus fasciatus***:**
Genome characterization and potential role in immune response against
bacterial and viral infections

1. Introduction

The complement system (CS) is composed of more than 35 proteins in plasma and on cell surfaces, and plays a vital role in clearing pathogenic microorganisms, as well as infected and otherwise damaged cells, from the host system (Gal et al., 2007; Korotaevskiy et al., 2009). The CS can be activated by three major signaling cascades, known as the classical, lectin and alternative pathways, which lead to various but interactive biological processes, such as inflammation, phagocytosis, lysis, and the adaptive immune response (Sunyer and Lambris, 1998). The classical pathway itself is activated upon recognition and binding of antibody-antigen complexes by the complement component 1 (C1) complex, which is composed of the C1q recognition molecule and four serine protease proenzymes, including two units of C1r and two units of C1s (Arlaud et al., 2001).

C1q has been characterized as the primary link between innate immunity driven by classical pathway and acquired immunity mediated by the immunoglobulin (Ig) components of antibody-antigen complexes, IgG and IgM (Kishore and Reid, 2000). Moreover, its activity has been demonstrated as crucial to a broad spectrum of immunological processes, including phagocytic removal of bacteria, retrovirus neutralization, apoptotic cell clearance, immune cell adhesion, and growth modulation of dendritic cells, B cells, and fibroblasts (Kishore et al., 2004b). C1q mediates these effects via strong, targeted binding to a diverse set of host- and pathogen-derived ligands, such as the lipopolysaccharide (LPS) and porin bacterial endotoxins, the envelop proteins of some retroviruses, various phospholipids, β-amyloid fibrils, apoptotic cells, and pentraxins (Kishore et al., 2004a). The unique protein structure of C1q might likely mediate this diverse capacity to recognize and bind such a vast array of ligands.

The teleostean complement system well characterized in common carp and number of complement components were identified including C1r/C1s, C3, C4, C5, Bf/C2, Factor I and MASP (mannose-binding lectin associated serine protease). Some studies have shown that the CS components and functions of mammals are similar to those in rainbow trout (a bony fish) (Nonaka et al., 1981) and the nurse shark (a cartilaginous fish) (Jensen et al., 1981). The C1q subcomponent, in particular, has been identified in nurse shark (Smith, 1998), channel catfish (Dodds and Petry, 1993), zebrafish (C1qA, C1qB and C1qC) (Hu et al., 2010), and mandarin fish (C1qL1 and C1qL2) (Lao et al., 2008). In addition, several studies of fish (Lao et al., 2008) and some other invertebrates (Gestal et al., 2010) have demonstrated that C1q activity is important to immune defense mediated by the classical complement pathway against invading pathogens.

Rock bream (*Oplegnathus fasciatus*) is one of the most economically important fish species of Korean aquaculture. In recent years, significant production losses have been caused by outbreaks of pathogenic diseases and researchers have directed their efforts towards understanding the underlying pathogenic mechanisms and immune responses to create a healthier and sustainable industry environment. In this study, we attempted to characterize the full-length cDNA and genomic sequence of putative C1qs from rock bream. Our data on the temporal expression profile and transcriptional response in liver to infection by common pathogens, including *Edwardsiella tarda* and *Streptococcus iniae* bacteria and the rock bream iridovirus (RBIV), provide novel insights into the C1q-mediated defense mechanisms of rock bream.

2. Materials and Methods

2.1 Rock bream cDNA library construction

The rock bream cDNA sequence database was previously established by our laboratory using the next-generation pyrosequencing-based Roche 454 Genome Sequencer FLX System (GS-FLX™) (Droege and Hill, 2008). In brief, total RNA was extracted from healthy rock bream using the Tri Reagent™ (Sigma, USA) and processed with the FastTrack® 2.0 mRNA isolation kit (Invitrogen, USA). The Creator™ SMART™ cDNA library construction kit (Clontech, USA) and Trimmer cDNA normalization kit (Evrogen, Russia) were used to synthesize and normalize first-strand cDNA, respectively. Subsequently, sequencing was carried out on the GS-FLX™ Titanium instrument (DNA Link, Inc., USA).

2.2 BAC library and genomic sequence screening

The rock bream bacterial artificial chromosome (BAC) library previously generated by our laboratory consists of 92,160 clones arrayed on 240 384-well plates (Wu et al., 1999). PCR-based screening with gene-specific primers for Cq1 identified a putative clone containing the gene of interest, and the BAC clone was sequenced by the Roche GS-FLX™ System.

2.3 Identification of C1q genes from the rock bream database

Three full-length C1q sequences were identified in the rock bream database by applying the Basic Local Alignment Tool (BLAST) in the National Center for Biotechnology Information (NCBI) web-based query system (http://www.ncbi.nlm.nih.gov/BLAST). Two of these sequences originated from the cDNA library and the other from the BAC library. Genomic sequences for all three C1q genes were obtained from the BAC library.

2.4 Sequence analysis of C1q genes and 3D modeling

The open reading frames (ORFs) and amino acid (aa) sequences of the putative C1q genes were analyzed using DNAssist software (version 2.2). The MatGat program was used to evaluate the percentage of identity and similarity of putative aa C1q sequences with known vertebrate sequences published in GenBank. Several programs were used to analyze the deduced aa sequences of the rock bream C1qs, including the expert protein analysis system (http://www.expasy.org), the motif scan Pfam hidden Markov models (Local models) (http://hits.isb-sib.ch/cgi-bin/PFSCAN) for conserved domains, the ClustalW2 (http://www.ebi.ac.uk/Tools/clustalw2/index.html) for pairwise and multiple alignments, the SignalP program (http://www.cbs.dtu.dk/services/SignalP/) for prediction of signal peptides, and the NetNGlyc 1.0 Server (http://www.cbs.dtu.dk/services/NetNGlyc/) for prediction of potential N-glycosylation sites. The Spidey program (http://www.ncbi.nlm.nih.gov/IEB/Research/_Ostell/Spidey/) was used to identify the exon-intron regions in the genomic sequence. Homologous genomic structures were obtained from the Ensembl genome browser (http://www.ensembl.org/index.html). Analysis of the genetic distances between C1q orthologes was carried out by constructing a phylogenetic tree using the neighbor-joining (NJ) method in MEGA software, version 5.05 (http://www.megasoftware.net). The 3D models of the rock bream C1q globular domains were constructed by the Swiss-Model protein modeling server (Arnold et al., 2006) and manipulated with the Swiss-PdbViewer program, version 4.04; the model structures were based on the crystallographic structures of human C1q globular domains (Gaboriaud et al., 2003).

2.5 Animals, immune challenge experiments, and tissue collection

Healthy rock breams, averaging 50 g in body weight, were obtained from the Jeju Special Self-Governing Province Ocean and Fisheries Research Institute (Jeju, Republic of Korea) and acclimatized to laboratory conditions (400 L tanks with filtered sea water: salinity 34±0.6%, pH 7.6±0.5, 23±1°C) for one week. For the live bacteria challenge experiments, each rock bream was intraperitoneally (i.p.) injected with 100 µL of Gram-negative *E. tarda* (5×10^3 CFU/µL) or Gram-positive *S. iniae* (1×10^5 CFU/µL), suspended in phosphate-buffered saline (PBS). These two bacterial strains were obtained from the Department of Aqualife Medicine at Chonnam National University (Republic of Korea) and grown in brain heart infusion broth (BHI). For the live virus challenge experiment, 100 µL of RBIV (10^3 $TCID_{50}$ of RBIV per fish) in PBS was injected (Umasuthan et al., 2011b). For the mitogen stimulation experiment, 100 µL of purified LPS (1.25 µg/µL of *E. coli* 055:B5; Sigma) in PBS was i.p. injected into each fish. A negative injection control group was established by i.p. injection of 100 µL PBS alone. Non-injected fish served as complete negative controls. Random fish were sacrificed from all challenge and control groups for collection of liver samples at post-injection (p.i.) hours 3, 6, 12, 24 and 48 h. Three numbers of fishes were used for each time points and each challenge experiments. Harvested tissues were snap-frozen in liquid nitrogen and stored at -70°C.

2.6 Extraction of total RNA and synthesis of first-strand cDNA

Liver, gill, spleen, head kidney, kidney, blood, skin, muscle, heart, brain, and intestine were harvested from healthy fish to evaluate the normal tissue distribution pattern of the rock bream C1qs. All tissues, including those from the immune-challenged and control liver tissues, were processed by the Tri Reagent to extract total RNA. The

~ 40 ~

PrimeScript™ First-Strand cDNA Synthesis Kit (TaKaRa, Japan) was used to synthesize cDNA by reverse transcription (RT). Briefly, 2.5 µg RNA (1 µg/µL) was mixed with 1 µL of 50 µM oligo(dT)$_{20}$ and 1 µL of 10 mM dNTPs, incubated for 5 min at 65°C, and combined with 4 µL of 5 × PrimeScript™ buffer, 0.5 µL of RNase inhibitor (40 U/µL), and 1 µL of PrimeScript™ reverse transcriptase (200 U/µL). The solution was incubated for 60 min at 42°C, after which the reaction was terminated by heating at 95°C for 5 min. The synthesized cDNA was diluted 40 times and stored at −70°C.

2.7 Quantitative analysis of mRNA expression

Quantitative real-time PCR (qPCR) was performed on the TP800 Thermal Cycler Dice™ Real-Time System (TaKaRa) to determine the spatial and temporal levels of C1q expression in healthy and immune-challenged fish. All qPCRs were carried out in triplicate using a 20 µL reaction mixture containing 4 µL of diluted cDNA (3.125 ng/µL) from the respective tissue, 10 µL of 2 × TaKaRa Ex*Taq*™ SYBR premix, 0.5 µL of each gene-specific primer (10 pmol/µL), and 5 µL of dH$_2$O. The thermal cycling conditions included one cycle of 95°C for 10 s, followed by 45 cycles of 95°C for 5 s, 58°C for 20 s and 72°C for 20 s, and a final single cycle of 95°C for 15 s, 60°C for 30 s and 95°C for 15 s. The rock bream β-actin (GenBank accession no. FJ975145) was amplified as an internal reference, and used to normalize the C1q transcripts detected in each tissue. The β-actin and C1q gene-specific primers used in this study are listed in **Table 3**. The $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001) was used to quantify the C1q mRNA expression level detected by qPCR.

To determine the transcriptional response to immune-challenges over time, the normalized C1q expression levels were calculated as fold-changes relative to normalized C1q expression levels detected in PBS-injected controls at the corresponding time points.

All the data are presented as mean (n=3) ± standard deviation of relative mRNA. Differences between the challenge and control groups were analyzed by GraphPad statistical software and a *p*-value of <0.05 was considered significant.

Table 3. Description of primers used in this study

Primer	Purpose	Sequence, 5′-3′
RbC1qAL F	qRT-PCR amplification	TGAGCAACAAGCTGGGATTCTGTG
RbC1qAL R	qRT-PCR amplification	AAGGACTCGAGCCAAACCTTCTGT
RbC1qBL F	qRT-PCR amplification	AGAGGTCGACACTGCCATCAACTT
RbC1qBL R	qRT-PCR amplification	TTATGTGACTGTCGCTGCCCTTCA
RbC1qCL F	qRT-PCR amplification	ATACCCAGACAAAGCCAGCGTCAT
RbC1qCL R	qRT-PCR amplification	AGAGGCGTGGAACGCAAAGTAGTA
Rb β-actin F	qRT-PCR internal reference	TCATCACCATCGGCAATGAGAGGT
Rb β-actin R	qRT-PCR internal reference	TGATGCTGTTGTAGGTGGTCTCGT

3. Results

3.1 cDNA sequence characterization of rock bream C1qs

Three full-length C1qs, C1qA-like (RbC1qAL) and C1qC-like (RbC1qCL) from the rock bream cDNA library and C1qB-like (RbC1qBL) from the BAC library, were isolated and deposited in GenBank under the accession numbers JQ805140, JQ805145, and JQ805141, respectively. Sequence analysis indicated that the RbC1qAL cDNA is 1125 bp in length. The ORF is 780 bp (excluding the termination codon) and encodes a polypeptide of 260 aa with an estimated molecular mass of 27.8 kDa. The 5′-untranslated region (UTR) is 96 bp in length and the 249 bp 3′-UTR contains a polyadenylation signal (^{1056}ATTAAA^{1061}) and an extended polyA tail (**Fig. 14a**). The cDNA of RbC1qBL is 970 bp and contains an ORF of 720 bp (excluding the termination codon) encoding a polypeptide of 240 aa with an estimated molecular mass of 25.3 kDa. Alignment with the known sequence of C1q-like-1 protein from mandarin fish (Lao et al., 2008) revealed that

the 5′-UTR of RbC1qBL is 71 bp in length and the 3′-UTR is 179 bp. The 3′-UTR contains two polyadenylation signals (^{863}ATTAAA^{868} and ^{926}AATAAA^{931}) (**Fig. 14b**). The full-length cDNA of RbC1qCL is 1262 bp and consists of a 726 bp ORF (excluding the termination codon), a 5′-UTR of 146 bp, and a 3′-UTR of 390 bp with two polyadenylation signals (^{1053}ATTAAA^{1058} and ^{1154}ATTAAA^{1159}) and an extended polyA tail. In addition, the ORF of RbC1qCL encodes a polypeptide of 242 aa with an estimated molecular mass of 26.1 kDa (**Fig. 14c**).

The organization of C1q domain-containing (C1qDC) proteins has been characterized, and includes a leading signal peptide (SP), a CLR of variable size, and a globular C1q (gC1q) domain in the C-terminus (Mei and Gui, 2008). An SP was identified in all three rock bream C1qs (RbC1qs), and was 22aa in RbC1qAL, 24 aa in RbC1qBL, and 22 aa in RbC1qCL. Motif analysis of the RbC1qs revealed that this specific sequence, responsible for stimulating phagocytic activity was located in the CLRs of both RbC1qAL (^{49}GEKGEP^{54}) and RbC1qCL (^{70}GEKGEP^{75}), but was absent from RbC1qBL. In addition, the characteristic Gly-X-Y (where X is often proline and Y is often hydroxylysine or hydroxyproline) repeats were clearly recognized in the CLRs of all three RbC1qs, and the number of repeats were 23 in RbC1qAL, 24 in RbC1qBL, and 23 in RbC1qCL. The gC1q domain was present in the C-terminals of all three RbC1qs, and was 142 aa long in RbC1qAL, 135 aa long in RbC1qBL, and 131 aa long in RbC1qCL. Moreover, all three RbC1qs contained a total of five cysteine residues, three of which were located in the gC1q domains of each. Only RbC1qBL contained N-linked glycosylation sites (^{103}NPSN^{106} and ^{149}NGSF^{152}), suggesting a unique glycoprotein function for this RbC1q.
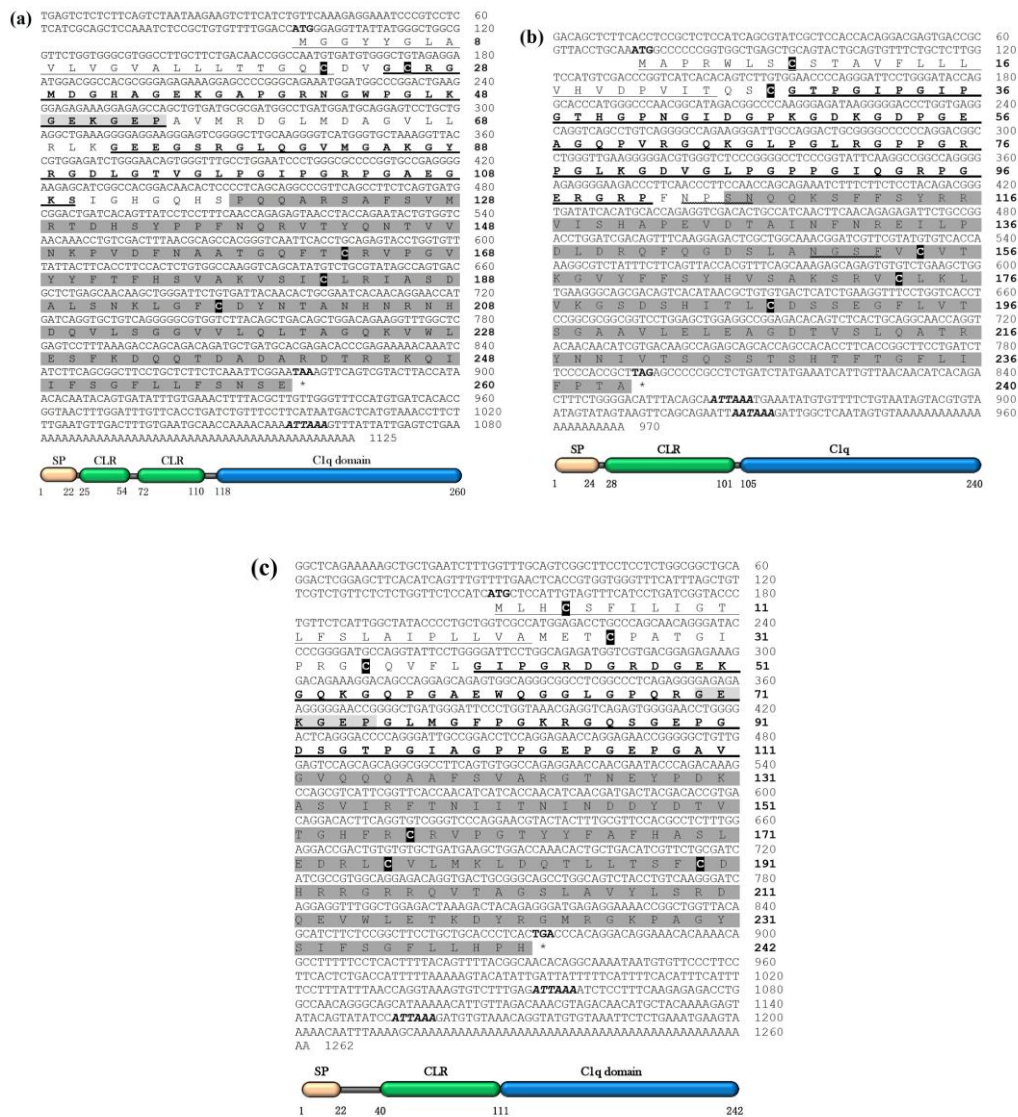
Fig. 14. The complete nucleotide sequence with deduced amino acid sequence of (a) RbC1qAL, (b) RbC1qBL, and (c) RbC1qCL. The start and stop codons are in bold letters. The polyadenylation signals are in bold and italicized. The putative signal peptide (SP) is underlined. The collagen-like region (CLR), comprised of Gly-X-Y repeating triplets, is in bold and underlined. The C1q domain is highlighted in grey. Conserved cysteine residues are highlighted in black. The GEKGEP motif in CLR is highlighted in light grey. The two potential N-linked glycosylation sites in RbClqBL are underlined with waves. The schematic diagram of each C1q is illustrated below the sequence.

3.2 Amino acid sequence comparison and phylogenetic analysis of the RbC1qs

To investigate the evolutionary distance of RbC1qs from C1qs of other species, the pairwise identity and similarity were analyzed. RbC1qAL, RbC1qBL, and RbC1qCL displayed the highest identities with *Siniperca chuatsi* C1qL2 (78.5%), *Siniperca chuatsi* C1qL1 (87.1%), and *Oreochromis niloticus* C1qC (73.1%), respectively (**Table 4**). In contrast, the RbC1qs showed 30-45% identities with human C1qs. Multiple alignment analysis identified several conserved sequence regions among the C1qs of different species (**Fig. 15**), including 19 Gly-X-Y repeats in the CLR and four cysteine residues.

Table 4. Percentage of interspecies amino acid sequence identity and similarity for RbC1qAL, RbC1qBL, and RbC1qCL

	RbC1qAL	S. chuatsi C1qL2	H. sapiens C1qA	D. rerio C1qA	RbC1qBL	S. chuatsi C1qL1	D. labrax C1qB	O. niloticus C1qBL	S. salar C1qB	D. rerio C1qB	H. sapiens C1qB	RbC1qCL	O. niloticus C1qC	D. labrax C1qC	D. rerio C1qC	H. sapiens C1qC
RbC1qAL		78.5	39.8	36.4	31.2	32.7	30.5	29.7	33.1	27.5	33.7	33.8	37.0	44.0	34.0	33.5
S. chuatsi **C1qL2**	82.7		37.2	37.6	30.8	30.8	30.0	28.1	32.7	28.6	34.1	33.6	35.5	43.6	34.5	33.2
H. sapiens **C1qA**	54.2	51.4		35.9	32.7	32.2	33.3	32.4	37.3	30.8	40.3	34.0	34.8	37.4	37.3	37.6
D. rerio **C1qA**	51.2	54.2	51.0		32.9	32.6	32.9	30.7	34.0	29.7	32.7	32.5	35.7	33.7	34.0	34.5
RbC1qBL	46.5	45.0	51.0	47.0		87.1	82.5	57.6	48.2	35.5	38.4	34.0	32.7	31.2	32.5	36.9
S. chuatsi **C1qL1**	46.5	46.6	51.0	48.2	92.9		82.1	60.5	50.2	36.7	36.8	32.4	33.1	31.2	33.2	36.9
D. labrax **C1qB**	47.3	45.0	51.8	46.6	92.9	89.6		60.9	49.8	38.9	37.2	33.1	35.1	31.9	32.4	34.9
O. niloticus **C1qBL**	46.5	45.0	51.4	46.6	71.3	71.7	72.9		47.2	36.1	39.7	32.0	34.7	30.6	33.2	34.5
S. salar **C1qB**	47.7	48.2	53.8	49.4	62.2	61.4	62.2	61.0		41.8	36.8	34.4	36.7	31.6	36.2	34.3
D. rerio **C1qB**	44.6	43.4	48.6	47.0	55.4	54.5	57.0	56.2	61.4		35.0	30.3	32.1	30.3	31.2	33.5
H. sapiens **C1qB**	52.7	54.5	57.7	53.0	56.9	56.1	56.5	55.3	57.3	54.9		37.6	41.1	38.0	38.5	48.4
RbC1qCL	48.1	49.4	53.5	49.4	52.5	50.0	52.5	51.7	51.4	49.2	51.4		73.1	65.6	46.6	44.2
O. niloticus **C1qC**	51.5	54.2	53.1	51.8	53.1	53.1	52.7	53.5	53.0	48.3	53.8	83.1		59.6	50.4	46.0
D. labrax **C1qC**	58.5	55.4	54.2	48.1	47.7	46.5	46.5	48.1	50.4	46.2	51.9	76.5	71.2		45.1	43.6
D. rerio **C1qC**	50.0	49.8	55.9	49.8	48.8	50.0	49.6	48.0	52.2	50.0	54.5	63.1	68.0	58.8		46.6
H. sapiens **C1qC**	53.8	51.8	55.5	51.0	53.5	55.1	53.1	52.7	46.6	50.2	63.2	56.3	59.2	53.8	60.0	

Identity (%) — top axis; Similarity (%) — left axis
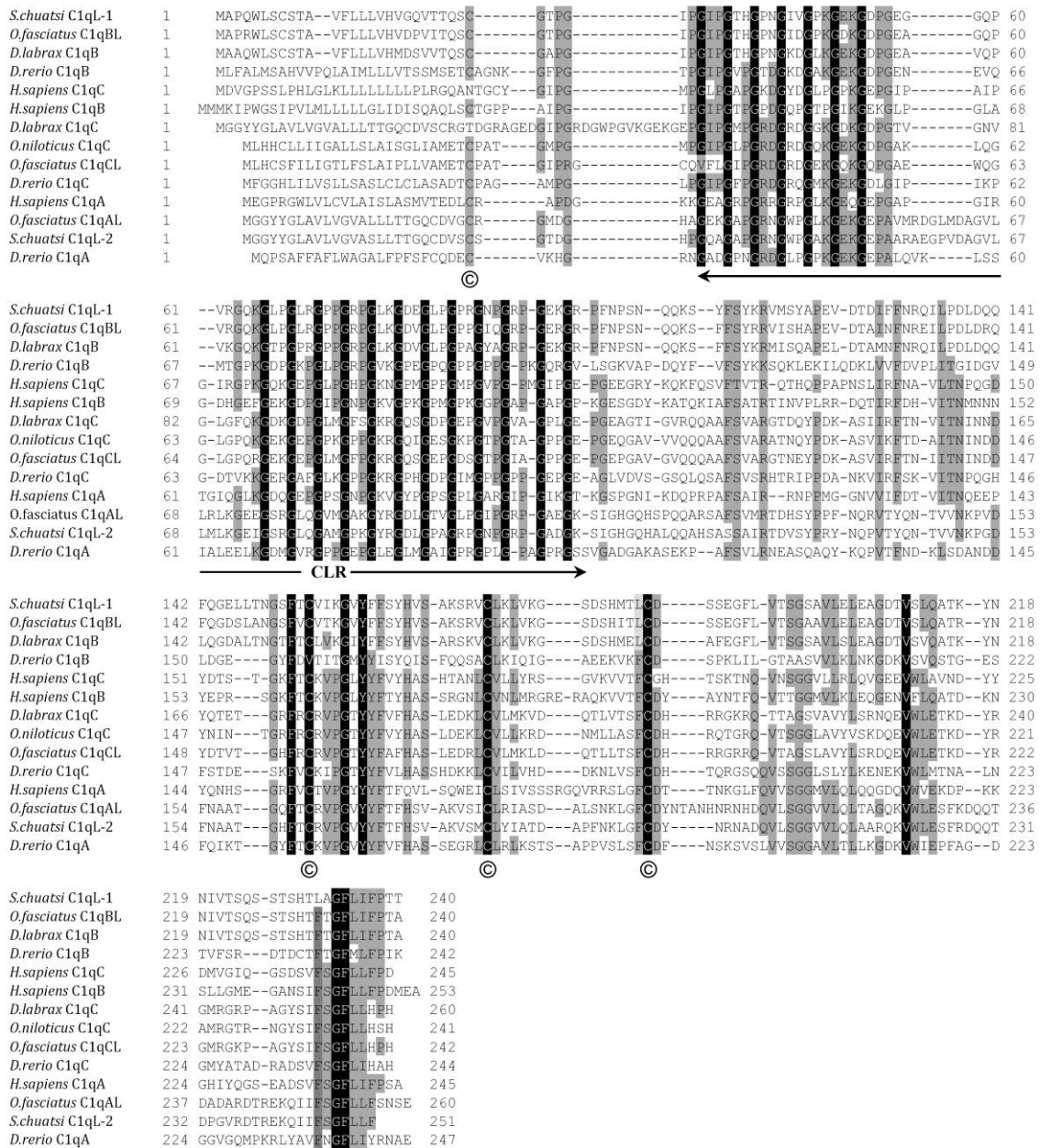
~ 45 ~

Fig. 15. Multiple alignment of deduced amino acid sequences of RbC1qs with putative orthologs. Completely conserved residues are highlighted in black, and highly conserved residues are highlighted in grey. Dashes in the amino acid sequences indicate gaps that were introduced to maximize alignment. Conserved cysteine residues are indicated by ©.

Molecular modeling studies were performed to evaluate the conservation of folding structure in the rock bream gC1q domains by comparison with human gC1qs. Predicted structures of the rock bream gC1qs exhibited a jelly-roll topology with 10 β-strands (**Fig. 16**), and also the position of cysteine residues in the gC1q regions is depicted in the each rock bream gC1q. Only RbC1qCL had a clearly demonstrated intra-chain disulfide bond between last two cysteine molecules, which were closest to one another. In addition, free thiol group was placed in similar position of the each RbC1qs.
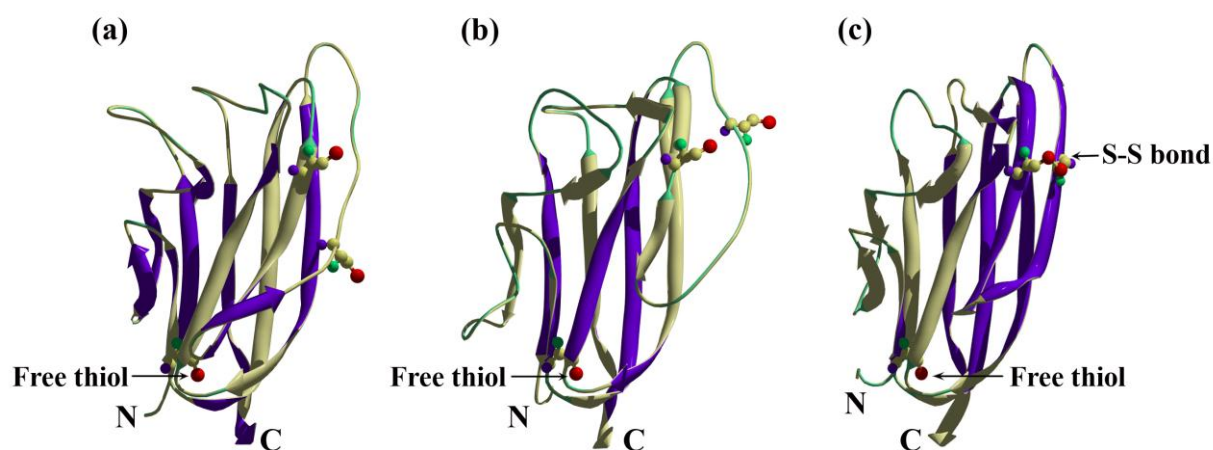
Fig. 16. Modeled 3D structures of the globular domains of (a) RbC1qAL, (b) RbC1qBL, and (c) RbC1qCL. A ribbon diagram is displayed for each RbC1q, with the conserved cysteine residues and sulphur atom highlighted in red. N- and C-terminals are labeled as N and C, respectively.

A phylogenetic tree was constructed for the C1qs based on the three subcomponents from different species in order to ascertain the origin and formation of RbC1qs. As shown in **Fig. 17**, the C1qs fell under three distinct clades. Each clade diverged into two vertebrate subclasses (fish and other vertebrates) and the RbC1qs clustered within the fish subclass.
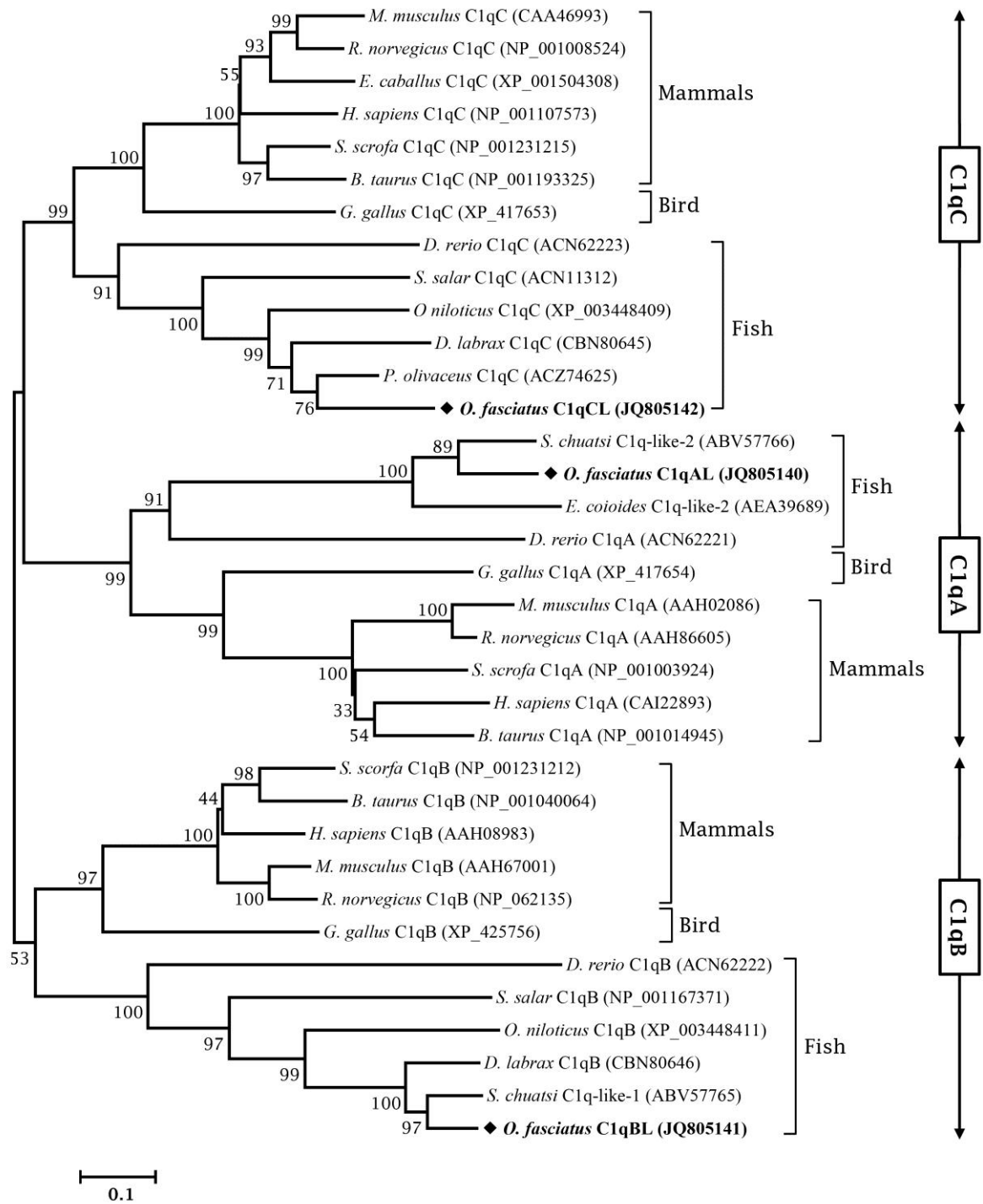
Fig. 17. Phylogenetic tree of RbC1q family constructed by the neighbor-joining method. The numbers at each branch indicate the corresponding percentage of bootstrapping.

3.3 Genomic analysis of the RbC1qs

To determine the genomic structures of RbC1qs, BAC library were screened and sequenced by next generation sequencing strategy. The lengths of the genomic sequences of RbC1qAL, RbC1qBL and RbC1qCL were 5198bp, 7560bp and 5385bp respectively and submitted to the GenBank (accession numbers JQ805142, JQ805143 and JQ805144). The ORFs of RbC1qAL, RbC1qBL, and RbC1qCL were distributed among six, five, and six exons, respectively, with the sixth exons being located in the 5′-UTRs (**Fig. 18**). All the exon-intron boundaries of each were confirmed by the presence of splice acceptor (AG) and donor (GT) sites. Comparison of the RbC1qs genomic structures with those of other species demonstrated that two fish species, medaka (*Oryzias latipes*) and tetraodon (*Tetraodon nigroviridis*), exhibited a similar structural pattern even though tetraodon having one more extra exon in C1qA. The size of the first, second and fourth exons of RbC1qAL were comparable with those in the ORF of medaka C1qA; whereas, all of the exons of RbC1qBL were identical to those in the ORF of tetraodon C1qB. For RbC1qCL, the last two exons in the ORF were of similar sizes to those in both medaka and tetraodon ORFs. The ORFs of all mammalian C1qs are limited to two exons. Notably, zebrafish C1qs exhibit similar exon patterns to mammalian C1q counterparts.
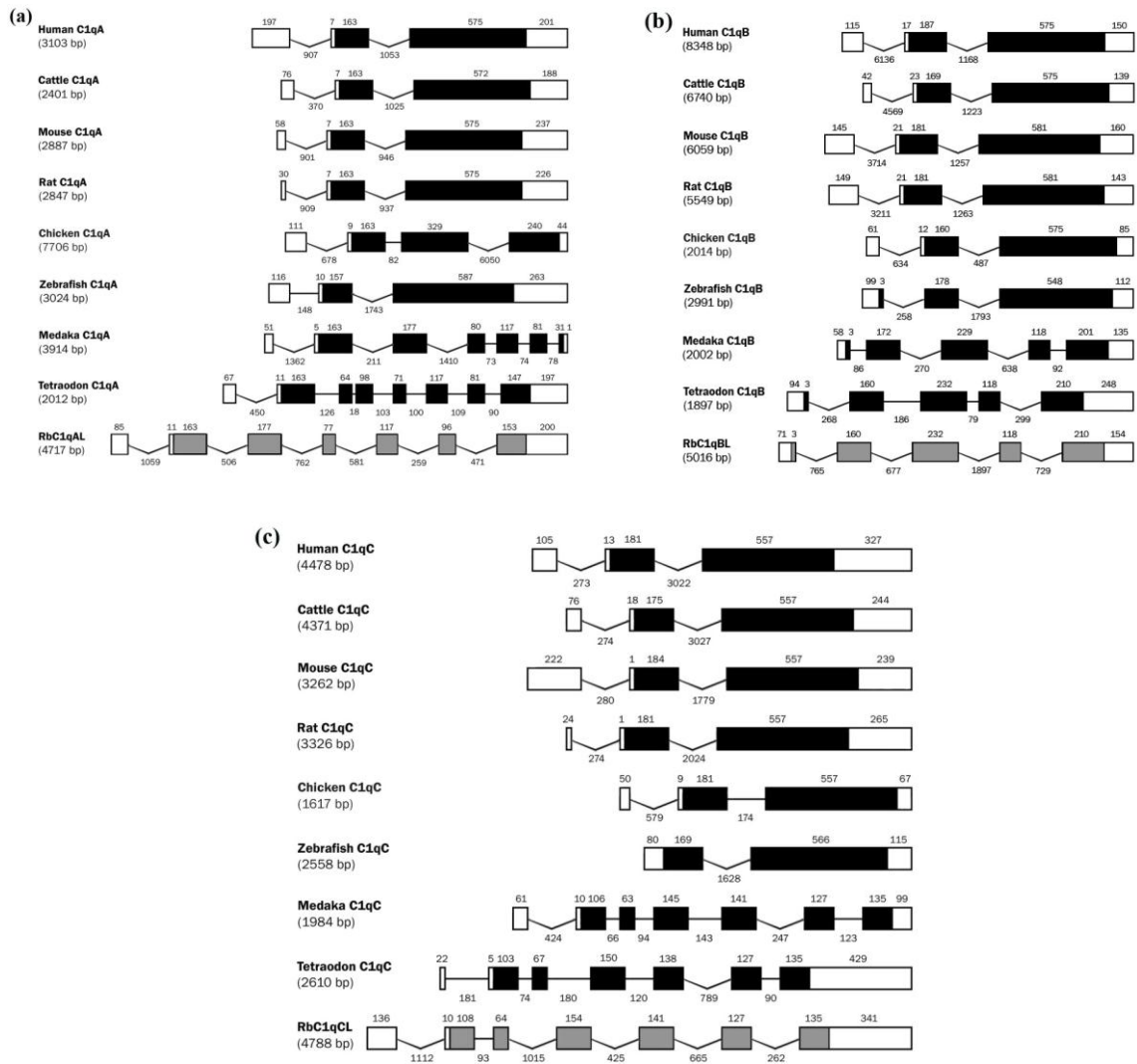
Fig. 18. Comparison of genomic structure and organization of (a) RbC1qAL, (b) RbC1qBL, and (c) RbC1qCL among vertebrate species. The genomic structures were obtained from the Ensembl database and included: C1qA of human (ENSG00000173372), cattle (ENSBTAG00000007153), mouse (ENSMUSG00000036887), rat (ENSRNOG00000012807), chicken (ENSGALG00000021569), zebrafish (ENSDARG00000044613), medaka (ENSORLG00000019405), and tetraodon (ENSTNIG00000006561); C1qB of human (ENSG00000173369), cattle (ENSBTAG00000011196), mouse (ENSMUSG00000036905), rat (ENSRNOG00000012749), chicken (ENSGALG00000004771), zebrafish (ENSDARG00000044612), medaka (ENSORLG00000017261), and tetraodon (ENSTNIG00000000825); C1qC of human (ENSG00000159189), cattle (ENSBTAG00000011193), mouse (ENSMUSG00000036896), rat (ENSRNOG00000012804), chicken (ENSGALG00000023605), zebrafish (ENSDARG00000095627), medaka (ENSORLG00000017257), and tetraodon (ENSTNIG00000006560).

~ 50 ~

3.4 Tissue distribution analysis of the RbC1qs

To determine the tissue-specific expression level of RbC1qs, qPCR was carried out using various tissues from healthy fish. Transcripts of all three RbC1qs were constitutively expressed in all tissues analyzed, and each C1q member displayed a similar expression pattern (**Fig. 19**). The highest expression level was detected in spleen, followed by liver; the lowest expression levels were detected in muscle, skin, brain, intestine, and blood.
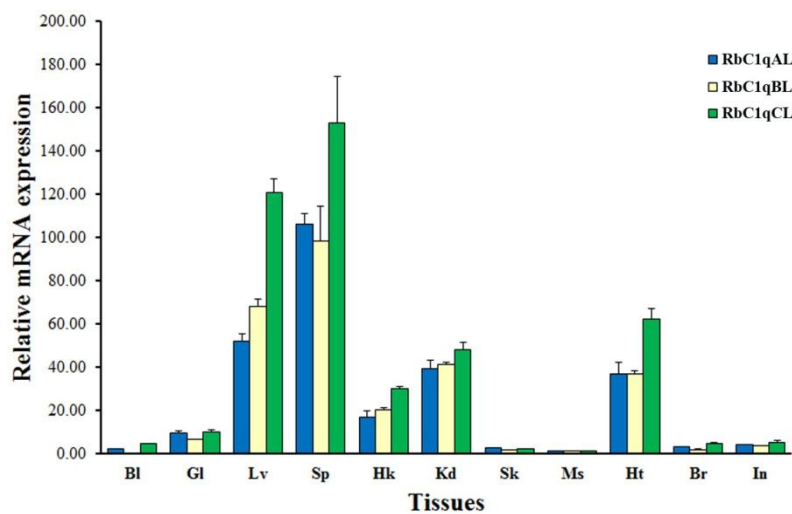
Fig. 19. Tissue-specific transcriptional profile of RbC1q in healthy rock bream. The RbC1q mRNA detected in each tissue by qPCR was normalized to RbC1q mRNA detected in muscle. Data are presented as mean (n=3) ± SD. Bl, blood; Gl, gill; Lv, liver; Sp, spleen; Hk, head kidney; Kd, kidney; Sk, skin; Ms, muscle; Ht, heart; Br, brain; In, intestine.

3.5 Temporal expression analysis of RbC1qs after immune challenges

The transcriptional response of RbC1qs to bacterial and viral challenges were assessed in rock bream liver in order to determine the potential defense activity of these proteins against pathogen infection. Two live bacterial strains (*E. tarda* and *S. iniae*), a viral strain (RBIV), and a bacterial endotoxin (LPS) that are commonly encountered by rock bream in the natural environment were used as immune stimulants. All three RbC1qs

showed similar expression patterns in response to each of the immune stimulants. At 3 h after the *E. tarda* challenge, levels of RbC1qAL, RbC1qBL, and RbC1qCL transcripts were significantly elevated ($p<0.05$ vs. PBS control). This time point was also when the highest *E. tarda*-induced levels were detected (8.8-, 11.2- and 14.9-fold, respectively), indicating an early phase response. After 24 h p.i., the expression levels began to significantly decrease (**Fig. 20a**). In contrast, the *S. iniae* challenge led to up-regulation of RbC1qCL transcription at 12 h p.i., but of RbC1qAL and RbC1qBL transcription at 24 h p.i. (**Fig.20b**); compared to the *E. tarda* transcriptional response, the *S. iniae* response was mild. Collectively, these data showed that the RbC1qs actively respond to Gram-negative bacterium at the early phase of infection, as compared to Gram-positive species. When rock bream were infected with RBIV, the expression level of all three RbC1qs were significantly up-regulated in the liver within 3 h p.i. ($p<0.05$). However, the levels varied between 7- and 20-fold compared to PBS controls (**Fig. 20c**).

LPS, a major component of the Gram-negative bacterial cell wall, can induce the expression of number of immune-related genes in the host system. Induction of RbC1qs was examined after challenge with LPS. The highest expressions of RbC1qAL (1.3-fold), RbC1qBL (2-fold), and RbC1qCL (1.4-fold) were observed at 48 h after injection with LPS. Although significant down-regulation occurred between 6 h and 24 h p.i (**Fig. 20d**), the trends were similar among all three of the RbC1qs.
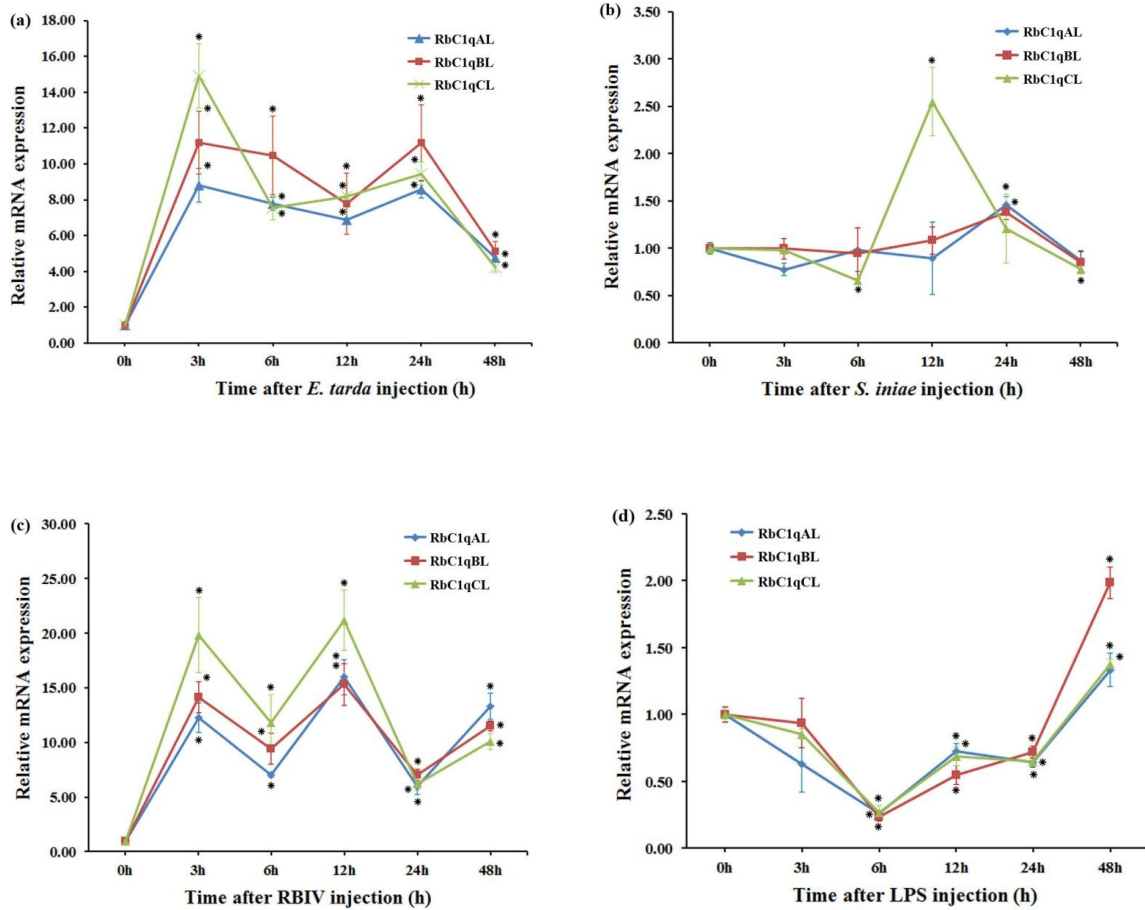
Fig. 20. Temporal expression profile of RbC1q in liver after challenge with (a) *E. tarda*, (b) *S. iniae*, (c) rock bream iridovirus, and (d) LPS. The RbC1q mRNA detected by qPCR was normalized to β-actin expression. Data are expressed as mean fold-induction (n=3) relative to the PBS control ± SD. *t-test, $p<0.05$ vs. unchallenged control at 0 h.

4. Discussion

C1q, the first subcomponent of the C1 complex initiate the classical complement pathway in mammal complement system. This can recognize various pathogenic microbial including viruses, bacteria and infected cells and subsequently activate the proteolysis of protein complexes in complement system to eliminate them.

Mammalian C1q is a hexamer, composed of 18 polypeptides that represent six C1qA, C1qB, and C1qC chains. Each C1q polypeptide chain contains a collagen-like region (CLR) and a C-terminal globular region known as the C1q domain (Sellar et al., 1991). Four conserved cysteine residues are present in all C1q polypeptide chains identified to date. One of these cysteines forms an inter-chain disulfide bond to generate A-B and C-C dimers, while two others form an intra-chain disulfide bond. The fourth cysteine residue acts as a free thiol group, which may interact with ligands. Several studies have suggested that each C1q chain can bind independently to its preferred ligands (Kishore and Reid, 1999), thereby creating an exponential potential for different binding partners.

We have identified and characterized 3 C1q domain containing genes from rock bream. The characteristic features of SP, CLR and C1q domain are well identified in all 3C1q proteins. The CLR is composed of repeating triplets of Gly-X-Y (where X is often proline and Y is often hydroxylysine or hydroxyproline) (Martin-Antonio et al., 2009), wherein the number of Gly-X-Y repeats can range from 14 to 100 (Innamorati et al., 2006). Gly-X-Y repeats were clearly recognized in the CLRs of all three RbC1qs, and the numbers of repeats were 23 in RbC1qAL, 24 in RbC1qBL, and 23 in RbC1qCL. This finding is similar to the number of Gly-X-Y repeats present in C1qA, C1qB, and C1qC of zebrafish and C1qL1 and C1qL2 of mandarin fish, which range from 20 to 25 (Hu et al., 2010; Lao et al., 2008). Experimental studies of human C1q have demonstrated that the

CLR can interact with neutrophils and trigger the activation of NADPH, which will enhance the phagocytic activity via the production of microbicidal oxygen radicals (Ruiz et al., 1995). Later studies of the human mannan-binding lectin, which activates the CS via the lectin pathway, identified a specific sequence (GEKGEP) in the CLR that was responsible for stimulating phagocytic activity. In addition, several studies have reported that complement components are involved in phagocytosis of apoptotic cells (Gershov et al., 2000; Mevorach et al., 1998). Nauta et al. (Nauta et al., 2002) demonstrated that the classical complement pathway could be initiated by binding of C1q directly to apoptotic cells through the globular head region. Therefore, we speculate that RbC1qAL and RbC1qCL may be involved in stimulating phagocytic activity in rock bream to eliminate pathogenic materials and/or apoptotic cells.

To study the evolutional mechanism o RbC1qs, we have compared the each amino acid sequences with other known vertebrate counterparts. The high identity was recognized with fish species, i.e. RbC1qAL, RbC1qBL, and RbC1qCL presented the highest identities with *S. chuatsi* C1qL2, *S. chuatsi* C1qL1, and *O. niloticus* C1qC respectively. Whereas, the RbC1qs depicted very low identities with human C1qs. In the case of molecular modeling of each RbC1qs displayed similar pattern of topology as reported for the gC1qs of *H. sapiens*, *D. rerio*, *M. edulis*, and *B. cereus* (Carland and Gerwick, 2010; Hu et al., 2010). And also, each RbC1qs depicted a free thiol group as shown in **Fig. 16**. This has been previously demonstrated by *in vitro* study that the free thiol group in the gC1q region can interact with several immune complexes through IgG (Martin et al., 1990). However, teleosts do not possess IgG, and the presence of a free thiol group in all three of the RbC1qs may indicate the emergence of potential IgG-binding site in teleostean C1qs. Accordingly, we can speculate that the rock bream gC1qs have become evolutionarily altered to support an adaptive immune response. In addition, phenylalanine

(F), glycine (G), and valine (V) were found to be completely conserved in the gC1q regions examined (**Fig. 15**), suggesting that these residues may be essential to maintain the globular shape of the C1q.

According to the phylogenetic study, we were able to recognize that fish group in each clade has clustered into three separate groups. Interestingly, RbC1qAL, RbC1qBL, and RbC1qCL clustered into the distinct fish groups of C1qA, C1qB, and C1qC, respectively. In addition, each zebrafish C1qs has evolved in separate sister clade in fish group. This information may indicate how the RbC1qs were conserved throughout evolution. Although the phylogenetic tree depicted a common ancestral origin for the three RbC1qs, RbC1qA and RbC1qC were categorized proximal to one another, while RbC1qB appeared to have evolved independently (Hu et al., 2010).

The genomic organization of C1qs revealed that fish C1qs have large number of exons (5-7) compare to other vertebrates (3-4) except in zebrafish. According to the genomic structural analysis of RbC1qs, it appears that the fish and other vertebrate C1qs may have evolved via different evolutionary pathways, even though they originated from a common ancestor. The results from the phylogenetic analysis also support this premise.

The tissue specific expression of RbC1qs was investigated in different tissues and found that highest expression in spleen and liver. Interestingly, all the RbC1qs exhibited similar expressional pattern. In previous studies, C1q expression in zebrafish showed a very similar tissue distribution pattern (Hu et al., 2010), while mandarin fish exhibited the lowest expression level in liver (Lao et al., 2008). In humans, the main sources of C1q synthesis are myeloid cells, macrophages, monocytes, and dendritic cells (Rabs et al., 1986). Investigation of C1q expression in rat liver revealed that the immune-related Kupffer cells produced considerably higher amounts of C1q than other types of liver cells, such as the sinusoidal cells (Armbrust et al., 1997). Similarly, in mice, the interdigitating

제주대학교 중앙도서관
JEJU NATIONAL UNIVERSITY LIBRARY

cells of the spleen, which are associated with high levels of monocyte-derived dendritic cells, exhibited high C1q expression (Lu et al., 2008). According to these evidences we can substantiate that the spleen and liver might be the main organs of RbC1qs synthesis.

Temporal changes in RbC1qs expressions have been analyzed using qPCR in bacterial and viral infected liver tissues. Remarkably, all RbC1qs exhibited similar expressional patterns. Strong early RbC1qs expressions were observed after Gram negative *E. tarda* infection, while recognizing mild expressions against Gram positive *S. iniae*. Previous studies of *Mytilus galloprovincialis* (Gestal et al., 2010), *Chlamys farreri* (Zhang et al., 2008), and *Argopecten irradians* (Kong et al., 2010) also demonstrated that the immune responsive expression of C1q occurred in the early phase (1-6 h) of infection with Gram-negative and Gram-positive bacteria. Recently, a study of *S. iniae* infection in *D. rerio* revealed that C1q expression becomes significantly elevated at the early stage (within 24 h p.i.) of infection (Carland et al., 2012). The expressions of RbC1qs demonstrated strong early responses against RbIV infection compare to the PBS control. A previous study of C1q expression in mandarin fish showed no significant response to challenge with the infectious spleen and kidney necrosis virus in liver, but remarkable up-regulation in spleen (Lao et al., 2008), where endogenous expression is highest. Moreover, a study of C1q expression in monkey found up-regulation in the early asymptomatic stage of SIV (simian immunodeficiency virus) infection (Depboylu et al., 2005). Although having contradicting observations, we can suggest that the RbC1qs possibly involve into the immunological responses beside bacterial and viral infections. In the case of LPS challenge, RbC1qs demonstrated a mild up regulation at 48h p.i. These results further indicate that the transcription of RbC1qs varies upon exposure to purified LPS and intact *E. tarda* bacteria, although the LPS is a component of the Gram-negative bacterial membrane. Moreover, amphioxus C1q has been reported as significantly induced by LPS

at 48 h p.i. (p<0.05) (Yu et al., 2008). Experimental studies in mice have revealed that LPS can stimulate the maturation of dendritic cells, which will in turn reduce the expression of C1q (Armbrust et al., 1997; Castellano et al., 2004). In addition, C1q binding to purified LPS has been shown to be inefficient, while its binding to two other bacterial outer membrane proteins, presumably porins, was shown to be efficient and capable of initiating the classical complement pathway (Alberti et al., 1993). Such inefficient binding may explain the observed down-regulation of RbC1qs in the early phase of challenge with LPS. However, further studies need to be carried out to determine the C1q-related mechanisms of LPS recognition in fish complement system.

C1q is known to activate the classical pathway by binding with its ligands in antibody complexes, which represents a major link between the innate and adaptive arms of immunity (Kishore and Reid, 1999). Nevertheless, in some lower vertebrates, such as lampreys, and invertebrates, C1q has been shown to be dispensable for activation of the classical complement pathway. Due to the lack of antibodies, such as Ig, related to the adaptive immune response, it is possible that C1q may act as a lectin to stimulate the innate immune response (Matsushita et al., 2004). Previous studies have reported that the classical pathway originated in fish. Moreover, some studies have shown that fish are the oldest species with antibody-dependent (IgM) activation of the complement pathway (Boshra et al., 2006; Nonaka and Smith, 2000). Recent studies in zebrafish demonstrated involvement of C1q in the classical complement pathway (Hu et al., 2010); however, antibody-independent binding of C1q with a wide range of pathogens, including Gram-negative and -positive bacteria, viruses and parasites, has been reported in mammals. Human C1q has been shown to directly bind various outer membrane proteins and porins of Gram-negative bacteria, such as *Legionella pneumophila*, *Escherichia coli, Salmonella typhimurium, Salmonella Minnesota*, and *Klebsilla pneumonia* (Alberti et al., 1993; Mintz

et al., 1995). Yet another study revealed that the C1q binding to OmpK36 porin of *Klebsiella pneumonia* was mediated by the globular region and showed that the binding site of OmpK36 was identical to that of IgG (Kojouharova et al., 2003). Several mammalian studies of virus infections have demonstrated C1q-mediated neutralization of the viral pathogens, including human immunodeficiency virus (HIV), human T lymphotrophic virus-1 (HTLV-1), murine leukemia virus (MuLV), herpes simplex virus (HSV), and cytomegalovirus (CMV) (Thielens et al., 2002). Further studies of HIV, HTLV-1, and MuLV indicated that the C1q molecule directly interacts with the viruses to trigger the classical complement pathway (Mehlhop and Diamond, 2006). In the current study, expressional kinetic analysis of all three RbC1qs demonstrated a clear early phase up-regulation in response to experimental injection of *E. tarda* and RBIV. Our findings suggest that RbC1q likely interact with live Gram-negative bacteria and virus through an antibody-independent mechanism to mount early-phase immune responses. However, further studies are required to achieve a better understanding of the rock bream defensive mechanism against microbial pathogens and involvement of RbC1qs in the classical complement pathway.

In conclusion, we have isolated and characterized the full-length cDNA and genomic sequences of three complement component 1q genes, the subcomponents of the C1 complex in the classical complement pathway of rock bream. Structural and phylogenetic evidence suggest that these newly-identified molecules are three distinct members of the C1q family. Expression analyses revealed that while the RbC1qs are constitutively expressed in various tissues, they are most strongly expressed in myeloid-rich tissues. Furthermore, the responsive expression pattern of RbC1qs upon bacterial and viral challenge suggests that these proteins may play a major role in immune defense of rock bream.

References

Alberti, S., Marques, G., Camprubi, S., Merino, S., Tomas, J.M., Vivanco, F., Benedi, V.J., 1993. C1q binding and activation of the complement classical pathway by Klebsiella pneumoniae outer membrane proteins. Infection and immunity 61, 852-860.

Alter-Koltunoff, M., Goren, S., Nousbeck, J., Feng, C.G., Sher, A., Ozato, K., Azriel, A., Levi, B.Z., 2008. Innate immunity to intraphagosomal pathogens is mediated by interferon regulatory factor 8 (IRF-8) that stimulates the expression of macrophage-specific Nramp1 through antagonizing repression by c-Myc. The Journal of biological chemistry 283, 2724-2733.

Arlaud, G.J., Gaboriaud, C., Thielens, N.M., Rossi, V., Bersch, B., Hernandez, J.F., Fontecilla-Camps, J.C., 2001. Structural biology of C1: dissection of a complex molecular machinery. Immunological reviews 180, 136-145.

Armbrust, T., Nordmann, B., Kreissig, M., Ramadori, G., 1997. C1Q synthesis by tissue mononuclear phagocytes from normal and from damaged rat liver: up-regulation by dexamethasone, down-regulation by interferon gamma, and lipopolysaccharide. Hepatology 26, 98-106.

Arnold, K., Bordoli, L., Kopp, J., Schwede, T., 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22, 195-201.

Bishop, R.E., 2005. Fundamentals of endotoxin structure and function. Contributions to microbiology 12, 1-27.

Biswas, P.S., Gupta, S., Chang, E., Song, L., Stirzaker, R.A., Liao, J.K., Bhagat, G., Pernis, A.B., 2010. Phosphorylation of IRF4 by ROCK2 regulates IL-17 and IL-21 production and the development of autoimmunity in mice. The Journal of clinical investigation 120, 3280-3295.

Boshra, H., Li, J., Sunyer, J.O., 2006. Recent advances on the complement system of teleost fish. Fish & shellfish immunology 20, 239-262.

Bovolenta, C., Driggers, P.H., Marks, M.S., Medin, J.A., Politis, A.D., Vogel, S.N., Levy, D.E., Sakaguchi, K., Appella, E., Coligan, J.E., et al., 1994. Molecular interactions between interferon consensus sequence binding protein and members of the interferon regulatory factor family. Proceedings of the National Academy of Sciences of the United States of America 91, 5046-5050.

Brass, A.L., Kehrli, E., Eisenbeis, C.F., Storb, U., Singh, H., 1996. Pip, a lymphoid-restricted IRF, contains a regulatory domain that is important for autoinhibition and ternary complex formation with the Ets factor PU.1. Genes & development 10, 2335-2347.

Brass, A.L., Zhu, A.Q., Singh, H., 1999. Assembly requirements of PU.1-Pip (IRF-4) activator complexes: inhibiting function in vivo using fused dimers. The EMBO journal 18, 977-991.

Carland, T.M., Gerwick, L., 2010. The C1q domain containing proteins: Where do they come from and what do they do? Developmental and comparative immunology 34, 785-790.

Carland, T.M., Locke, J.B., Nizet, V., Gerwick, L., 2012. Differential expression and intrachromosomal evolution of the sghC1q genes in zebrafish (Danio rerio). Developmental and comparative immunology 36, 31-38.

Casals-Casas, C., Alvarez, E., Serra, M., de la Torre, C., Farrera, C., Sanchez-Tillo, E., Caelles, C., Lloberas, J., Celada, A., 2009. CREB and AP-1 activation regulates MKP-1 induction by LPS or M-CSF and their kinetics correlate with macrophage activation versus proliferation. European journal of immunology 39, 1902-1913.

Castellano, G., Woltman, A.M., Nauta, A.J., Roos, A., Trouw, L.A., Seelen, M.A., Schena, F.P., Daha, M.R., van Kooten, C., 2004. Maturation of dendritic cells abrogates C1q production in vivo and in vitro. Blood 103, 3813-3820.

Collet, B., Ganne, G., Bird, S., Collins, C.M., 2009. Isolation and expression profile of a gene encoding for the Signal Transducer and Activator of Transcription STAT2 in Atlantic salmon (Salmo salar). Developmental and comparative immunology 33, 821-829.

Darnell, J.E., Jr., 1997. STATs and gene regulation. Science 277, 1630-1635.

Demoulins, T., Baron, M.L., Kettaf, N., Abdallah, A., Sharif-Askari, E., Sekaly, R.P., 2009. Poly (I:C) induced immune response in lymphoid tissues involves three sequential waves of type I IFN expression. Virology 386, 225-236.

Depboylu, C., Schafer, M.K., Schwaeble, W.J., Reinhart, T.A., Maeda, H., Mitsuya, H., Damadzic, R., Rausch, D.M., Eiden, L.E., Weihe, E., 2005. Increase of C1q biosynthesis in brain microglia and macrophages during lentivirus infection in the rhesus macaque is sensitive to antiretroviral treatment with 6-chloro-2',3'-dideoxyguanosine. Neurobiology of disease 20, 12-26.

Dodds, A.W., Petry, F., 1993. The phylogeny and evolution of the first component of complement, C1. Behring Institute Mitteilungen, 87-102.

Droege, M., Hill, B., 2008. The Genome Sequencer FLX System--longer reads, more applications, straight forward bioinformatics and more complete data sets. Journal of biotechnology 136, 3-10.

Eisenbeis, C.F., Singh, H., Storb, U., 1995. Pip, a novel IRF family member, is a lymphoid-specific, PU.1-dependent transcriptional activator. Genes & development 9, 1377-1387.

Eroshkin, A., Mushegian, A., 1999. Conserved transactivation domain shared by interferon regulatory factors and Smad morphogens. Journal of molecular medicine 77, 403-405.

Fortier, A., Doiron, K., Saleh, M., Grinstein, S., Gros, P., 2009. Restriction of Legionella pneumophila replication in macrophages requires concerted action of the transcriptional regulators Irf1 and Irf8 and nod-like receptors Naip5 and Nlrc4. Infection and immunity 77, 4794-4805.

Gaboriaud, C., Juanhuix, J., Gruez, A., Lacroix, M., Darnault, C., Pignol, D., Verger, D., Fontecilla-Camps, J.C., Arlaud, G.J., 2003. The crystal structure of the globular head of complement protein C1q provides a basis for its versatile recognition properties. The Journal of biological chemistry 278, 46974-46982.

Gal, P., Barna, L., Kocsis, A., Zavodszky, P., 2007. Serine proteases of the classical and lectin pathways: similarities and differences. Immunobiology 212, 267-277.

Gautier, G., Humbert, M., Deauvieau, F., Scuiller, M., Hiscott, J., Bates, E.E., Trinchieri, G., Caux, C., Garrone, P., 2005. A type I interferon autocrine-paracrine loop is involved in Toll-like receptor-induced interleukin-12p70 secretion by dendritic cells. The Journal of experimental medicine 201, 1435-1446.

Gauzzi, M.C., Purificato, C., Conti, L., Adorini, L., Belardelli, F., Gessani, S., 2005. IRF-4 expression in the human myeloid lineage: up-regulation during dendritic cell differentiation and inhibition by 1alpha,25-dihydroxyvitamin D3. Journal of leukocyte biology 77, 944-947.

Gershov, D., Kim, S., Brot, N., Elkon, K.B., 2000. C-Reactive protein binds to apoptotic cells, protects the cells from assembly of the terminal complement components, and sustains an antiinflammatory innate immune response: implications for systemic autoimmunity. The Journal of experimental medicine 192, 1353-1364.

Gestal, C., Pallavicini, A., Venier, P., Novoa, B., Figueras, A., 2010. MgC1q, a novel C1q-domain-containing protein involved in the immune response of Mytilus galloprovincialis. Developmental and comparative immunology 34, 926-934.

Hiscott, J., 2007. Triggering the innate antiviral response through IRF-3 activation. The Journal of biological chemistry 282, 15325-15329.

Hofman, V.J., Moreilhon, C., Brest, P.D., Lassalle, S., Le Brigand, K., Sicard, D., Raymond, J., Lamarque, D., Hebuterne, X.A., Mari, B., Barbry, P.J., Hofman, P.M., 2007. Gene expression profiling in human gastric mucosa infected with Helicobacter pylori. Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc 20, 974-989.

Holland, J.W., Karim, A., Wang, T., Alnabulsi, A., Scott, J., Collet, B., Mughal, M.S., Secombes, C.J., Bird, S., 2010. Molecular cloning and characterization of interferon regulatory factors 4 and 8 (IRF-4 and IRF-8) in rainbow trout, Oncorhynchus mykiss. Fish Shellfish Immunol 29, 157-166.

Hu, Y.L., Pan, X.M., Xiang, L.X., Shao, J.Z., 2010. Characterization of C1q in teleosts: insight into the molecular and functional evolution of C1q family and classical pathway. The Journal of biological chemistry 285, 28777-28786.

Huber, M., Brustle, A., Reinhard, K., Guralnik, A., Walter, G., Mahiny, A., von Low, E., Lohoff, M., 2008. IRF4 is essential for IL-21-mediated induction, amplification, and stabilization of the Th17 phenotype. Proceedings of the National Academy of Sciences of the United States of America 105, 20846-20851.

Innamorati, G., Bianchi, E., Whang, M.I., 2006. An intracellular role for the C1q-globular domain. Cellular signalling 18, 761-770.

Jensen, J.A., Festa, E., Smith, D.S., Cayer, M., 1981. The complement system of the nurse shark: hemolytic and comparative characteristics. Science 214, 566-569.

Kishore, U., Gaboriaud, C., Waters, P., Shrive, A.K., Greenhough, T.J., Reid, K.B., Sim, R.B., Arlaud, G.J., 2004a. C1q and tumor necrosis factor superfamily: modularity and versatility. Trends in immunology 25, 551-561.

Kishore, U., Ghai, R., Greenhough, T.J., Shrive, A.K., Bonifati, D.M., Gadjeva, M.G., Waters, P., Kojouharova, M.S., Chakraborty, T., Agrawal, A., 2004b. Structural and functional anatomy of the globular domain of complement protein C1q. Immunology letters 95, 113-128.

Kishore, U., Reid, K.B., 1999. Modular organization of proteins containing C1q-like globular domain. Immunopharmacology 42, 15-21.

Kishore, U., Reid, K.B., 2000. C1q: structure, function, and receptors. Immunopharmacology 49, 159-170.

Kojouharova, M.S., Tsacheva, I.G., Tchorbadjieva, M.I., Reid, K.B., Kishore, U., 2003. Localization of ligand-binding sites on human C1q globular head region using recombinant globular head fragments and single-chain antibodies. Biochimica et biophysica acta 1652, 64-74.

Kong, P., Zhang, H., Wang, L., Zhou, Z., Yang, J., Zhang, Y., Qiu, L., Wang, L., Song, L., 2010. AiC1qDC-1, a novel gC1q-domain-containing protein from bay scallop Argopecten irradians with fungi agglutinating activity. Developmental and comparative immunology 34, 837-846.

Korotaevskiy, A.A., Hanin, L.G., Khanin, M.A., 2009. Non-linear dynamics of the complement system activation. Mathematical biosciences 222, 127-143.

Lao, H.H., Sun, Y.N., Yin, Z.X., Wang, J., Chen, C., Weng, S.P., He, W., Guo, C.J., Huang, X.D., Yu, X.Q., He, J.G., 2008. Molecular cloning of two C1q-like cDNAs in mandarin fish Siniperca chuatsi. Veterinary immunology and immunopathology 125, 37-46.

Liljeroos, M., Vuolteenaho, R., Rounioja, S., Henriques-Normark, B., Hallman, M., Ojaniemi, M., 2008. Bacterial ligand of TLR2 signals Stat activation via induction of IRF1/2 and interferon-alpha production. Cellular signalling 20, 1873-1881.

Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25, 402-408.

Lu, J.H., Teh, B.K., Wang, L., Wang, Y.N., Tan, Y.S., Lai, M.C., Reid, K.B., 2008. The classical and regulatory functions of C1q in immunity and autoimmunity. Cellular & molecular immunology 5, 9-21.

Marecki, S., Atchison, M.L., Fenton, M.J., 1999. Differential expression and distinct functions of IFN regulatory factor 4 and IFN consensus sequence binding protein in macrophages. Journal of immunology 163, 2713-2722.

Marquis, J.F., Kapoustina, O., Langlais, D., Ruddy, R., Dufour, C.R., Kim, B.H., MacMicking, J.D., Giguere, V., Gros, P., 2011. Interferon regulatory factor 8 regulates pathways for antigen presentation in myeloid cells and during tuberculosis. PLoS genetics 7, e1002097.

Martin-Antonio, B., Jimenez-Cantizano, R.M., Salas-Leiton, E., Infante, C., Manchado, M., 2009. Genomic characterization and gene expression analysis of four hepcidin genes in the redbanded seabream (Pagrus auriga). Fish & shellfish immunology 26, 483-491.

Martin, H., Kaul, M., Loos, M., 1990. Disulfide bridge formation between C1q and IgG in vitro. European journal of immunology 20, 1641-1645.

Matsuo, A., Oshiumi, H., Tsujita, T., Mitani, H., Kasai, H., Yoshimizu, M., Matsumoto, M., Seya, T., 2008. Teleost TLR22 recognizes RNA duplex to induce IFN and protect cells from birnaviruses. Journal of immunology 181, 3474-3485.

Matsushita, M., Matsushita, A., Endo, Y., Nakata, M., Kojima, N., Mizuochi, T., Fujita, T., 2004. Origin of the classical complement pathway: Lamprey orthologue of mammalian C1q acts as a lectin. Proceedings of the National Academy of Sciences of the United States of America 101, 10127-10131.

Mehlhop, E., Diamond, M.S., 2006. Protective immune responses against West Nile virus are primed by distinct complement activation pathways. The Journal of experimental medicine 203, 1371-1381.

Mei, J., Gui, J., 2008. Bioinformatic identification of genes encoding C1q-domain-containing proteins in zebrafish. Journal of genetics and genomics = Yi chuan xue bao 35, 17-24.

Mevorach, D., Mascarenhas, J.O., Gershov, D., Elkon, K.B., 1998. Complement-dependent clearance of apoptotic cells by human macrophages. The Journal of experimental medicine 188, 2313-2320.

Minten, C., Terry, R., King, N., Campbell, L., 2011. Increased lethality but decreased encephalitis in IRF8-deficient mice following West Nile Virus infection. The Journal of Immunology 186, 154.

Mintz, C.S., Arnold, P.I., Johnson, W., Schultz, D.R., 1995. Antibody-independent binding of complement component C1q by Legionella pneumophila. Infection and immunity 63, 4939-4943.

Nauta, A.J., Trouw, L.A., Daha, M.R., Tijsma, O., Nieuwland, R., Schwaeble, W.J., Gingras, A.R., Mantovani, A., Hack, E.C., Roos, A., 2002. Direct binding of C1q to apoptotic cells and cell blebs induces complement activation. European journal of immunology 32, 1726-1736.

Nonaka, M., Smith, S.L., 2000. Complement system of bony and cartilaginous fish. Fish & shellfish immunology 10, 215-228.

Nonaka, M., Yamaguchi, N., Natsuume-Sakai, S., Takahashi, M., 1981. The complement system of rainbow trout (Salmo gairdneri). I. Identification of the serum lytic system homologous to mammalian complement. Journal of immunology 126, 1489-1494.

Park, S.L., 2009. Disease Control in Korean Aquaculture. Fish Pathology 44, 19-23.

Paun, A., Pitha, P.M., 2007. The IRF family, revisited. Biochimie 89, 744-753.

Perdiguero, B., Esteban, M., 2009. The interferon system and vaccinia virus evasion mechanisms. Journal of interferon & cytokine research : the official journal of the International Society for Interferon and Cytokine Research 29, 581-598.

Pfeffer, L.M., Dinarello, C.A., Herberman, R.B., Williams, B.R., Borden, E.C., Bordens, R., Walter, M.R., Nagabhushan, T.L., Trotta, P.P., Pestka, S., 1998. Biological properties of recombinant alpha-interferons: 40th anniversary of the discovery of interferons. Cancer research 58, 2489-2499.

Rabs, U., Martin, H., Hitschold, T., Golan, M.D., Heinz, H.P., Loos, M., 1986. Isolation and characterization of macrophage-derived C1q and its similarities to serum C1q. European journal of immunology 16, 1183-1186.

Raetz, C.R., Whitfield, C., 2002. Lipopolysaccharide endotoxins. Annual review of biochemistry 71, 635-700.

Ramji, D.P., Foka, P., 2002. CCAAT/enhancer-binding proteins: structure, function and regulation. The Biochemical journal 365, 561-575.

Rodkhum, C., Hirono, I., Stork, M., Di Lorenzo, M., Crosa, J.H., Aoki, T., 2006. Putative virulence-related genes in Vibrio anguillarum identified by random genome sequencing. Journal of fish diseases 29, 157-166.

Rosenbauer, F., Waring, J.F., Foerster, J., Wietstruk, M., Philipp, D., Horak, I., 1999. Interferon consensus sequence binding protein and interferon regulatory factor-

4/Pip form a complex that represses the expression of the interferon-stimulated gene-15 in macrophages. Blood 94, 4274-4281.

Ruiz, S., Henschen-Edman, A.H., Tenner, A.J., 1995. Localization of the site on the complement component C1q required for the stimulation of neutrophil superoxide production. The Journal of biological chemistry 270, 30627-30634.

Sellar, G.C., Blake, D.J., Reid, K.B., 1991. Characterization and organization of the genes encoding the A-, B- and C-chains of human complement subcomponent C1q. The complete derived amino acid sequence of human C1q. The Biochemical journal 274 (Pt 2), 481-490.

Sepulcre, M.P., Alcaraz-Perez, F., Lopez-Munoz, A., Roca, F.J., Meseguer, J., Cayuela, M.L., Mulero, V., 2009. Evolution of lipopolysaccharide (LPS) recognition and signaling: fish TLR4 does not recognize LPS and negatively regulates NF-kappaB activation. Journal of immunology 182, 1836-1845.

Sharf, R., Meraro, D., Azriel, A., Thornton, A.M., Ozato, K., Petricoin, E.F., Larner, A.C., Schaper, F., Hauser, H., Levi, B.Z., 1997. Phosphorylation events modulate the ability of interferon consensus sequence binding protein to interact with interferon regulatory factors and to bind DNA. The Journal of biological chemistry 272, 9785-9792.

Sharma, S., Grandvaux, N., Mamane, Y., Genin, P., Azimi, N., Waldmann, T., Hiscott, J., 2002. Regulation of IFN regulatory factor 4 expression in human T cell leukemia virus-I-transformed T cells. Journal of immunology 169, 3120-3130.

Sharma, S., Mamane, Y., Grandvaux, N., Bartlett, J., Petropoulos, L., Lin, R., Hiscott, J., 2000. Activation and regulation of interferon regulatory factor 4 in HTLV type 1-infected T lymphocytes. AIDS research and human retroviruses 16, 1613-1622.

Skjesol, A., Hansen, T., Shi, C.Y., Thim, H.L., Jorgensen, J.B., 2010. Structural and functional studies of STAT1 from Atlantic salmon (Salmo salar). BMC immunology 11, 17.

Smith, S.L., 1998. Shark complement: an assessment. Immunological reviews 166, 67-78.

Sun, B.J., Chang, M.X., Song, Y., Yao, W.J., Nie, P., 2007. Gene structure and transcription of IRF-1 and IRF-7 in the mandarin fish Siniperca chuatsi. Veterinary immunology and immunopathology 116, 26-36.

Sunyer, J.O., Lambris, J.D., 1998. Evolution and diversity of the complement system of poikilothermic vertebrates. Immunological reviews 166, 39-57.

Tailor, P., Tamura, T., Ozato, K., 2006. IRF family proteins and type I interferon induction in dendritic cells. Cell research 16, 134-140.

Takaoka, A., Tamura, T., Taniguchi, T., 2008. Interferon regulatory factor family of transcription factors and regulation of oncogenesis. Cancer science 99, 467-478.

Tamura, T., Kong, H.J., Tunyaplin, C., Tsujimura, H., Calame, K., Ozato, K., 2003. ICSBP/IRF-8 inhibits mitogenic activity of p210 Bcr/Abl in differentiating myeloid progenitor cells. Blood 102, 4547-4554.

Tamura, T., Tailor, P., Yamaoka, K., Kong, H.J., Tsujimura, H., O'Shea, J.J., Singh, H., Ozato, K., 2005a. IFN regulatory factor-4 and -8 govern dendritic cell subset development and their functional diversity. J Immunol 174, 2573-2581.

Tamura, T., Thotakura, P., Tanaka, T.S., Ko, M.S., Ozato, K., 2005b. Identification of target genes and a unique cis element regulated by IRF-8 in developing macrophages. Blood 106, 1938-1947.

Tamura, T., Yanai, H., Savitsky, D., Taniguchi, T., 2008. The IRF family transcription factors in immunity and oncogenesis. Annual review of immunology 26, 535-584.

Taniguchi, T., Ogasawara, K., Takaoka, A., Tanaka, N., 2001. IRF family of transcription factors as regulators of host defense. Annual review of immunology 19, 623-655.

Thielens, N.M., Tacnet-Delorme, P., Arlaud, G.J., 2002. Interaction of C1q and mannan-binding lectin with viruses. Immunobiology 205, 563-574.

Ting, L.M., Kim, A.C., Cattamanchi, A., Ernst, J.D., 1999. Mycobacterium tuberculosis inhibits IFN-gamma transcriptional responses without inhibiting activation of STAT1. J Immunol 163, 3898-3906.

Umasuthan, N., Whang, I., Kim, J.O., Oh, M.J., Jung, S.J., Choi, C.Y., Yeo, S.Y., Lee, J.H., Noh, J.K., Lee, J., 2011a. Rock bream (Oplegnathus fasciatus) serpin, protease nexin-1: Transcriptional analysis and characterization of its antiprotease and anticoagulant activities. Dev Comp Immunol.

Umasuthan, N., Whang, I., Lee, Y., Lee, S., Kim, Y., Kim, H., Jung, S.J., Oh, M.J., Choi, C.Y., Yeo, S.Y., Lee, S.J., Lee, J., 2011b. Heparin cofactor II (RbHCII) from rock bream (Oplegnathus fasciatus): molecular characterization, cloning and expression analysis. Fish & shellfish immunology 30, 194-208.

Wu, C., Asakawa, S., Shimizu, N., Kawasaki, S., Yasukochi, Y., 1999. Construction and characterization of bacterial artificial chromosome libraries from the silkworm, Bombyx mori. Molecular & general genetics : MGG 261, 698-706.

Xu, D., Zhao, L., Del Valle, L., Miklossy, J., Zhang, L., 2008. Interferon regulatory factor 4 is involved in Epstein-Barr virus-mediated transformation of human B lymphocytes. Journal of virology 82, 6251-6258.

Yu, Y., Huang, H., Wang, Y., Yu, Y., Yuan, S., Huang, S., Pan, M., Feng, K., Xu, A., 2008. A novel C1q family member of amphioxus was revealed to have a partial function of vertebrate C1q molecule. Journal of immunology 181, 7024-7032.

Zhang, H., Song, L., Li, C., Zhao, J., Wang, H., Qiu, L., Ni, D., Zhang, Y., 2008. A novel C1q-domain-containing protein from Zhikong scallop Chlamys farreri with lipopolysaccharide binding activity. Fish & shellfish immunology 25, 281-289.

Zhang, Y., Gui, J., 2004. Molecular characterization and IFN signal pathway analysis of Carassius auratus CaSTAT1 identified from the cultured cells in response to virus infection. Developmental and comparative immunology 28, 211-227.

Zubiaga, A.M., Belasco, J.G., Greenberg, M.E., 1995. The nonamer UUAUUUAUU is the key AU-rich sequence motif that mediates mRNA degradation. Molecular and cellular biology 15, 2219-2230.

Acknowledgements