A DOCTORAL DISSERTATION

# Molecular Cloning, Characterization and Expression Analysis of Innate Immune-related Genes in Rock Bream (*Oplegnathus fasciatus*)


**Ilson Whang**

**Department of Life Science**

**Graduate School**
**Jeju National University**


**December, 2010**

# 돌돔에서 선천성 면역관련 유전자들의 특성분석 및 발현 연구

지도교수 김 세 재

황 일 선

이 논문을 이학 박사학위 논문으로 제출함

2010년 12월

황일선의 이학 박사학위 논문을 인준함

심사위원장 _____

위 원 _____

위 원 _____

위 원 _____

위 원 _____

제주대학교 대학원

2010년 12월

# Molecular Cloning, Characterization and Expression Analysis of Innate Immune-related Genes in Rock Bream (*Oplegnathus fasciatus*)

**Ilson Whang**

**(Supervised by Professor Se-Jae Kim)**

**A dissertation submitted in partial fulfillment of the requirement for the degree of Doctor of Philosophy**

**December, 2010**

This dissertation has been examined and approved by

.....

Chairperson of the Committee

.....

.....

.....

.....

.....

Date

**Department of Life Science**

**GRADUATE SCHOOL**

**JEJU NATIONAL UNIVERSITY**

# SUMMARY

The immune system defends the host against infection. Innate immunity is the earliest defense mechanism but lacks the ability to recognize certain pathogens and to provide the specific protective immunity against recurrent infection. Adaptive immunity is based on clonal selection from a repertoire of lymphocytes bearing highly diverse antigen-specific receptors that enable the immune system to recognize any foreign antigen.

In fish immunity, the innate system is of primary importance in combating infections. The strength of innate defense against pathogens is impressive, despite the limited pathogen recognition machinery. This is demonstrated by the very efficient immune defense of fish, which mostly depend on innate parameters for coping with a large variety of pathogens in diverse environmental conditions. The innate immune response of fish is also important in activating an acquired immune response. Although less studied in fish, the available information indicates that the activation of innate recognition components, through the stimulation of phagocytes, production of cytokines and chemokines and activation of the complement system and various cell receptors, stimulates T- and B-cells and antigen presenting cell in fish.

Next-generation sequencing has become a powerful and efficient high throughput technique in whole transcriptome analysis. Rock bream is an economically important cultured marine fish in Korean which frequently affected by diseases such as vibriosis, edwardsiellosis, scuticociellosis, and iridovirus infection. Hence, there is an urgent need to develop and characterize transcriptome of rock bream for providing the extensive genomic information.

To sequence the rock bream whole transcriptome, we employed the Roche 454 pyrosequencing platform on cDNAs constructed using pooled mRNA isolated from

various tissue types (multi-tissue) of un-challenged fish. Nearly, 672,000 reads (average read size: 400 bp) were resulted in a single GS-FLX sequencing which was assembled to ~36,000 contigs. Results showed that pyrosequencing has covered wide range of contig sizes (95-6175 bp) which were assembled into ~1.1 Kb average contig size.

Final aim of present study is to identify and obtain full-length sequences of immune functional genes in rock bream for better understanding of host pathogen interactions. Therefore, this thesis has been focused on molecular characterization, transcriptional analysis and functional aspects of immune-relevant genes such as goose-type lysozyme (g-type lysozyme), myeloid differentiation factor 88 (MyD88), interferon (IFN) regulatory factor-1 (IRF-1), IFN gamma inducible lysosomal thiol reductase (GILT), Cathepsin B and L cystein proteases in rock bream fish.

**1) Characterization and expression analysis of a g-type lysozyme and antimicrobial activity of its recombinant protein**

Lysozyme (muramidase) represents an important defense molecule of the fish innate immune system. Known for its bactericidal properties, lysozyme catalyzes the hydrolysis of β-(1,4)-glycosidic bonds between the *N*-acetyl glucosamine and *N*-acetyl muramic acid in the peptidoglycan layer of bacterial cell walls. In this study, the complete coding sequence of g-type lysozyme from the rock bream (RBgLyz) was identified and characterized as being composed of 669 bp, with a 567 bp open reading frame that encodes 188 amino acids. Protein motif searches indicated that RBgLyz contains the soluble lytic transglycosylase domain involved in balancing cell wall integrity. Furthermore, RBgLyz shares significant identity (81.4%) with Chinese perch *Siniperca chuatsi*. Quantitative real-time PCR (qRT-PCR) analysis results confirmed that RBgLyz transcriptional expression is constitutively expressed in various

tissues from healthy rock breams. RBgLyz expression was analyzed in head kidney following immune challenges with several distinct pathogens, including LPS, poly I:C, *Edwardsiella tarda*, *Streptococcus iniae* and rock bream iridovirus (RBIV). Compared to non-injected control fish, a significant up-regulation of RBgLyz transcripts was observed in response to LPS and *E. tarda* challenge. Poly I:C stimulated a moderate expression of RBgLyz, as did *S. iniae* but to a lesser extent. However, specific time-dependent effects on RBgLyz mRNA expression in response to RBIV infection were not observed. Taken together, the gene expression results indicated a role for g-type lysozyme in innate immune response against LPS, poly I:C, *E. tarda* and *S. iniae* in rock bream. Thus, we generated recombinant RBgLyz in an *Escherichia coli* expression system and characterized its antimicrobial activity. Our results indicated that recombinant RBgLyz had lytic activity against Gram-negative *Vibrio salmonicida*, Gram-positive *Listeria monocytogenes*, *S. iniae* and *Micrococcus lysodeikticus*. In addition, observations by scanning electron microscope (SEM) confirmed that the cell morphology of *M. lysodeikticus* was altered in the presence of recombinant RBgLyz.

**2) Characterization and expression analysis of MyD88**

Myeloid differentiation factor 88 (MyD88) is a universal adaptor protein able to activate nuclear factor-kappa B (NF-κB) through interactions with interleukin-1 receptor (IL-1R) and the Toll-like receptors (TLRs), with the exception of TLR3. The cDNA of rock bream MyD88 was found to be composed of 1626 bp, with an 867 bp open reading frame that encodes 288 amino acids. The deduced amino acid sequence of MyD88 possessed both a conserved death domain at the amino terminus and a typical Toll-IL-1 receptor (TIR) domain at the carboxyl terminus, similar to that found in other fishes, amphibians, avians, mammals and invertebrates. The mRNA expression pattern of MyD88 in healthy and bacterially-challenged rock bream

were examined using qRT-PCR. MyD88 transcripts were found to be strongly expressed in blood, gill, liver, spleen, head kidney and kidney, moderately expressed in skin, brain and intestine, and weakly expressed in muscle. Expression levels of MyD88 in blood, spleen and head kidney were dramatically up-regulated upon exposure to LPS and *E. tarda*, suggesting that MyD88 plays an important role in rock bream defenses against bacterial infection.

**3) Molecular analysis and transcriptional responses of IRF-1 and GILT**

Activation of the interferon (IFN) system and its down-stream IFN-stimulated genes (ISGs) play important roles in innate and adaptive immune responses to pathogens. The cDNA sequence of rock bream IRF-1 (named as RbIRF-1) showed significant evolutionary conservation of its N-terminal 113 amino acids, which encompassed a DNA binding domain (DBD) containing five conserved tryptophan repeats. Meanwhile, the GILT cDNA sequences from rock bream (RbGILT cDNA) had the characteristic GILT signature sequence composed of a functional domain ($^{97}$CQHGEQECLGNMIETC$^{112}$), active site $^{74}$C-XX-C$^{77}$ motif and seven putative disulfide bonds. Therefore, the newly identified rock bream IRF-1 and GILT proteins are similar to those in other fish and mammals as revealed by molecular characterization and phylogenetic analysis. The qRT-PCR analysis revealed that expression of RbIRF-1 and RbGILT transcripts was constitutive in tissues selected from un-challenged rock bream, with the highest levels observed blood and gills. Immune challenge with synthetic poly I:C up-regulated the RbIRF-1 and RbGILT mRNA in blood, gills, spleen and head kidney; however, the magnitude of RbGILT up-regulation was lower than that of RbIRF-1. Only a moderate up-regulation (compared to poly I:C) of RbIRF-1 and RbGILT mRNA was observed in gills and head kidney at the early stage of RBIV challenge, although, in blood, both

transcripts were found to be continuously up-regulated throughout the 48 h observation. The transcriptional up-regulation in response to poly I:C and RBIV strongly suggest that RbIRF-1 and RbGILT are related to IFN signaling and may indicate essential roles in subsequent adaptive immunity in rock bream.

**4) Molecular characterization and expression analysis of Cathepsin B and L cysteine proteases**

Cathepsins are lysosomal cysteine proteases of the papain family that play an important role in intracellular protein degradation. The rock bream cathepsin B (RbCathepsin B) was composed of 330 amino acid residues and having 36 kDa molecular mass. The rock bream cathepsin L (RbCathepsin L) contained 336 amino acid residues encoding for a 38 kDa molecular mass protein. The sequencing analysis results showed that both cathepsin B and L contain the characteristic papain family cysteine protease signature and active sites for the eukaryotic thiol proteases cysteine, asparagine and histidine. In addition, RbCathepsin L contained EF hand $Ca^{2+}$ binding and cathepsin propeptide inhibitor domains. The rock bream cathepsin B and L showed the highest amino acid identity of 90 and 95% to *Lutjanus argentimaculatus* cathepsin B and *Lates calcarifer* cathepsin L, respectively. By phylogenetic analysis, cathepsin B and L exhibited a high degree of evolutionary relationship to respective cathepsin family members of the papain superfamily. The qRT-PCR analysis results confirmed that cathepsin B and L gene expression was constitutive in all examined tissues isolated from un-induced rock bream. Moreover, challenge with LPS and *E. tarda* resulted in significant up-regulation of RbCathepsin B and L mRNA in liver and blood, indicating a role for cathepsin B and L in immune responses against bacteria in rock bream.

# CONTENTS

Page

**LIST OF ABBREVIATIONS** ..... **11**

**PART 1** ..... **14**

**Characterization and expression analysis of a goose-type lysozyme from the rock bream *Oplegnathus fasciatus*, and antimicrobial activity of its recombinant protein**

**1.1 ABSTRACT** ..... **15**

**1.2 LIST OF TABLES** ..... **17**

**1.3 LIST OF FIGURES** ..... **18**

**1.4 INTRODUCTION** ..... **19**

**1.5 MATERIALS AND METHODS** ..... **21**

1.5.1 Experimental fish..... 21

1.5.2 Identification of rock bream g-type lysozyme..... 21

1.5.3 Sequence characterization and phylogenetic analysis of RBgLyz.. 22

1.5.4 Tissue collection, total RNA extraction and cDNA synthesis..... 22

1.5.5 Immune challenge..... 23

1.5.6 RBgLyz mRNA expression analysis by real-time PCR..... 24

1.5.7 Cloning of RBgLyz coding sequences into pMAL-c2X..... 25

1.5.8 Purification of the recombinant RBgLyz protein (rRBgLyz)..... 27

1.5.9 Determination of optimum pH on the activity of rRBgLyz..... 28

1.5.10 Antimicrobial activity of rRBgLyz..... 28

1.5.11 Scanning electron microscope (SEM) observation..... 29

**1.6 RESULTS** ..... **31**

1.6.1 Molecular characterization of RBgLyz..... 31

1.6.2 Phylogenetic analysis of RBgLyz..... 32

1.6.3 Expression of RBgLyz in healthy rock breams..... 32
1.6.4 Expression of RBgLyz in rock breams induced with LPS, poly I:C, *E. tarda, S. iniae* and RBIV..... 33
1.6.5 Expression and purification of recombinant rRBgLyz..... 40
1.6.6 Lysozyme activity..... 40
**1.7 DISCUSSION ..... 46**
**1.8 REFERENCES ..... 52**

**PART 2 ..... 56**
**Characterization and expression analysis of the myeloid differentiation factor 88 (MyD88) in rock bream *Oplegnathus fasciatus***
**2.1 ABSTRACT ..... 57**
**2.2 LIST OF TABLES ..... 58**
**2.3 LIST OF FIGURES..... 59**
**2.4 INTRODUCTION ..... 60**
**2.5 MATERIALS AND METHODS ..... 63**
2.5.1 Identification of rock bream MyD88..... 63
2.5.2 Sequence alignments and phylogenetic analysis of MyD88..... 63
2.5.3 Experimental fish and tissue collection..... 64
2.5.4 Immune challenge..... 64
2.5.5 Total RNA extraction and cDNA synthesis..... 65
2.5.6 MyD88 mRNA expression analysis by qRT-PCR..... 66
**2.6 RESULTS ..... 68**
2.6.1 Molecular characterization and phylogeny of the rock bream MyD88..... 68
2.6.2 Expression of the rock bream MyD88 in untreated individuals.... 68
2.6.3 Expression of the rock bream MyD88 in individuals infected with Gram (-) bacteria..... 69

**2.7 DISCUSSION** ..... **78**

**2.8 REFERENCES** ..... **82**

**PART 3** ..... **87**

**Molecular analysis and transcriptional responses of interferon (IFN) regulatory factor-1 and IFN gamma inducible lysosomal thiole reductase (GILT) from rock bream (*Oplegnathus fasciatus*)**

**3.1 ABSTRACT** ..... **88**

**3.2 LIST OF TABLES** ..... **90**

**3.3 LIST OF FIGURES** ..... **91**

**3.4 INTRODUCTION** ..... **92**

**3.5 MATERIALS AND METHODS** ..... **95**

3.5.1 Fish and tissue isolation ..... 95

3.5.2 Construction of normalized cDNA library ..... 95

3.5.3 GSFLX-454 sequencing and identification of RbIRF-1 and RbGILT ..... 96

3.5.4 Molecular characterization of RbIRF-1 and RbGILT ..... 96

3.5.5 Immune challenge of rock bream ..... 97

3.5.6 cDNA synthesis for transcriptional analysis of RbIRF-1 and RbGILT ..... 97

3.5.7 Transcriptional analysis of RbIRF-1 and RbGILT by qPCR ..... 98

**3.6 RESULTS** ..... **100**

3.6.1 Analysis of pyrosequencing and identification of RbIRF-1 and RbGILT cDNA sequences ..... 100

3.6.2 Sequence characterization of RbIRF-1 ..... 100

3.6.3 Sequence characterization of RbGILT ..... 101

3.6.4 Phylogenetic analysis of RbIRF-1 and RbGILT ..... 102

3.6.5 Transcriptional analysis of RbIRF-1and RbGILT by real time R T-PCR ..... 110
3.6.5.1 Tissue-specific expression ..... 110
3.6.5.2 Transcriptional regulation against polyI:C challenge ..... 111
3.6.5.3 Transcriptional regulation against RbIv challenge ..... 112
**3.7 DISCUSSION ..... 117**
**3.8 REFERENCES ..... 123**

**PART 4 ..... 128**
**Molecular characteriation and expression analysis of Cathepsin B and L cysteine proteases from rock bream (*Oplegnathus fasciatus*)**
**4.1 ABSTRACT ..... 129**
**4.2 LIST OF TABLES ..... 130**
**4.3 LIST OF FIGURES ..... 131**
**4.4 INTRODUCTION ..... 132**
**4.5 MATERIALS AND METHODS ..... 134**
4.5.1 Animal rearing and RNA isolation ..... 134
4.5.2 Multi-tissue cDNA synthesis and normalization ..... 134
4.5.3 Pyrosequencing, sequence processing, assembly and identification of RbCathepsin B and L ..... 135
4.5.4 Molecular characterization of RbCathepsin B and L ..... 135
4.5.5 LPS and bacterial challenge ..... 136
4.5.6 cDNA ..... 136
4.5.7 Expression analysis of RbCathepsin B and L by qPCR ..... 137
**4.6 RESULTS ..... 141**
4.6.1 Identification and characterization of RbCathepsin B and L cDNA sequences ..... 141
4.6.2 Tissue expression profiles of RbCathepsin B and L ..... 143

4.6.3 Transcriptional analysis of RbCathepsin B and L after LPS and *E. tarda* challenge ..... 143
**4.7 DISCUSSION ..... 153**
**4.8 REFERENCES ..... 156**
**SUMMARY IN KOREAN ..... 160**

# LIST OF ABBREVIATIONS

| | |
|---|---|
| aa | amino acid |
| ANOVA | analysis of variance |
| β | beta |
| BHI | brain heart infusion |
| BLAST | Basic Local Alignment Tool |
| bp | base pair(s) |
| cDNA | complementary deoxyribonucleic acid |
| c-type | chicken-type |
| DNA | deoxyribonucleic acid |
| dNTP | deoxynucleotide-triphosphate |
| *E. tarda* | *Edwardsiella tarda* |
| *E. coli* | *Escherichia coli* |
| EDTA | ethylene diamine tetra acetic acid |
| x g | gravity |
| GS-FLX | Genome Sequencer FLX System |
| g-type | goose-type |
| h | hour |
| HEWL | hen egg white lysozyme |
| i.p.(ip) | intraperitoneal |
| IFN | interferon |
| IL | interleukin |
| IL-1R | interleukin-1 receptor |
| IPTG | isopropyl-β-thiogalactopyranoside |
| IRAK | IL-1R-associated kinase |

| | |
|---|---|
| i-type | invertebrate-type |
| Kb | kilo base(s) |
| KDa | kilo dalton |
| KCCM | Korean Culture Center of Microorganisms |
| KCTC | Korean Collection for Type Cultures |
| LB medium | Luria broth medium |
| LPS | lipopolysaccharide |
| LycGL | large yellow croaker g-type lysozyme |
| MBP | maltose binding protein |
| mg | miligram(s) |
| mL | milliliter(s) |
| mRNA | messenger RNA |
| MEGA | Molecular Evolutionary Genetics Analysis |
| MyD88 | Myeloid differentiation factor 88 |
| MW | molecular weight |
| μL | microlitre(s) |
| NAM | N-acetyl muramic acid |
| NCBI | National Center for Biotechnology Information |
| NF-κB | nuclear factor-kappa B |
| NGS | next-generation DNA sequencing |
| OD | optical density |
| ORF | open reading frame |
| PAGE | polyacrylamide gel electrophoresis |
| p.i. | post-infection |
| PAMPs | pathogen-associated molecular patterns |

| | |
|---|---|
| PBS | phosphate buffered saline |
| p*I* | isoelectric point |
| PGN | peptidoglycan |
| poly I:C | polyinosinic:polycytidylic acid |
| PRRs | pattern recognition receptors |
| qRT-PCR | quantitative real-time RT-PCR |
| S.D. | standard deviation |
| SDS | sodium dodecyl sulphate |
| SEM | scanning electron microscopy |
| SLT | soluble lytic transglycosylase |
| SMART | Simple Modular Architecture Research Tool |

# PART 1

**Characterization and expression analysis of a goose-type lysozyme from the rock bream *Oplegnathus fasciatus*, and antimicrobial activity of its recombinant protein**

## 1.1 ABSTRACT

Lysozyme (muramidase) represents an important defense molecule of the fish innate immune system. Known for its bactericidal properties, lysozyme catalyzes the hydrolysis of β-(1,4)-glycosidic bonds between the *N*-acetyl glucosamine and *N*-acetyl muramic acid in the peptidoglycan layer of bacterial cell walls. In this study, the complete coding sequence of g-type lysozyme from the rock bream fish, *Oplegnathus fasciatus*, (RBgLyz) was identified by multi-tissue normalized cDNA pyrosequencing using the Roche 454 GS-FLX$^{TM}$ technology. RBgLyz was characterized as being composed of 669 bp, with a 567 bp open reading frame that encodes 188 amino acids. Protein motif searches indicated that RBgLyz contains the soluble lytic transglycosylase domain involved in balancing cell wall integrity. Furthermore, RBgLyz shares significant identity (81.4%) with Chinese perch *Siniperca chuatsi*. Quantitative real-time RT-PCR analysis results showed that RBgLyz transcripts are constitutively expressed in various tissues from healthy rock breams. RBgLyz expression was analyzed in head kidney following immune challenges with several stimulants or pathogens, including LPS, poly I:C, *Edwardsiella tarda*, *Streptococcus iniae* and rock bream iridovirus (RBIV). Compared to non-injected control fish, a significant up-regulation of RBgLyz transcripts was observed in response to LPS and *E. tarda* challenge. Poly I:C stimulated a moderate expression of RBgLyz, as did *S. iniae* but to a lesser extent. However, specific time-dependent effects on RBgLyz mRNA expression in response to RBIV infection were not observed. Taken together, the gene expression results indicated a role for g-type lysozyme in innate immune response against LPS, poly I:C, *E. tarda* and *S. iniae* in rock bream. Thus, we generated recombinant RBgLyz in an *Escherichia coli* expression system and

characterized its antimicrobial activity. Our results indicated that recombinant RBgLyz had lytic activity against Gram-negative *Vibrio salmonicida*, Gram-positive *Listeria monocytogenes*, *S. iniae* and *Micrococcus lysodeikticus*. In addition, observations by scanning electron microscope (SEM) confirmed that the cell morphology of *M. lysodeikticus* was altered in the presence of recombinant RBgLyz.

**Keywords:** *Oplegnathus fasciatus*, Goose-type lysozyme, innate immunity, gene expression, lytic activity

## 1.2 LIST OF TABLES

Page
Table 1.1 Primer sequences and their use ..... 26

## 1.3 LIST OF FIGURES

Page

Figure 1.1 Nucleotide (GenBank accession no. HM035063) and deduced amino acid sequences of RBgLyz in *Oplegnathus fasciatus* ..... 34

Figure 1.2 ClustalW2 multiple alignment of RBgLyz with g-type lysozymes from various fish species ..... 35

Figure 1.3 Phylogenetic tree based on the amino acid sequences of lysozymes.. 36

Figure 1.4 Relative mRNA expression of RBgLyz in different tissues, including blood, gill, liver, spleen, head kidney (H. kidney), kidney, skin, muscle, brain, and intestine ..... 38

Figure 1.5 Relative RBgLyz mRNA expression in head kidney (H. kidney) in response to challenges with LPS, poly I:C, *E. tarda, S. iniae* or RBIV at 3, 6, 12, 24 and 48 h post-injection ..... 39

Figure 1.6 Expression and purification of recombinant RBgLyz fusion protein..... 42

Figure 1.7 Optimal pH of recombinant RBgLyz as measured by a lysoplate assay ..... 43

Figure 1.8 Antimicrobial activity (UL) of the recombinant RBgLyz, as measured by turbidimetric assay ..... 44

Figure 1.9 SEM images of *M. lysodeikticus* after treatment with the recombinant RBgLyz ..... 45

## 1.4 INTRODUCTION

The immune system is an important physiological mechanism that protects the organism against invading pathogens. The immune system is broadly categorized as either innate or adaptive immunity. Innate immunity plays a critical role in host defense against infectious diseases by discriminating between self and foreign particles, and provides the necessary signals to activate the subsequent adaptive immune response which leads to immunological memory [1]. Fish mostly depend on innate immunity, and as such their immune system is considered less well developed than that of mammals [2].

Lysozyme (muramidase) is an important defense molecule of the fish innate immune system. Its bactericidal properties result from the ability to catalyze the hydrolysis of β-(1,4)-glycosidic bonds between the *N*-acetyl glucosamine (NAG) and *N*-acetyl muramic acid (NAM) in the peptidoglycan layer of bacterial cell walls [3]. Lysozyme is known to act directly on Gram-positive bacteria, causing lysis of their outermost peptidoglycan layer. Gram-negative bacteria, however, are not directly damaged by lysozyme as their outer membrane is significantly coated with lipopolysaccharide (LPS) moieties; instead, the outer membrane of Gram-negative bacteria is disrupted by cationic antimicrobial peptides that act to expose the inner peptidoglycan layer of bacteria, facilitating the activity of lysozyme [4]. Lysozyme is also known to be an opsonin and can activate the complement system and phagocytes [3,5]. In addition to an antibacterial function, some lysozymes have been demonstrated to have antiviral or anti-inflammatory activities [3,6-9], and these may function as multipurpose defense factors.

Based on the differences in structural, catalytic and immunological characteristics, lysozymes are generally classified among six main types: chicken-type (c-type)

lysozyme [10], goose-type (g-type) lysozyme [11], invertebrate (i-type) lysozyme [12], T4 phage (phage-type) lysozyme [13], bacterial lysozyme [14], and plant lysozyme [15]. The g-type lysozyme was initially identified as an antibacterial enzyme located in egg whites from the Embden goose [16], but was later found in the egg whites of several other bird species [17]. Recently, though, the g-type lysozyme has also been identified in vertebrate species, including mammals and fishes [18-28], as well as in invertebrates like scallop [29].

Rock bream *Oplegnathus fasciatus* is one of the most valuable marine fish species in Korean aquaculture. In recent years, the rapid development of the rock bream culture industry has been accompanied by a concomitant increase in prevalence and virulence of infectious diseases afflicting this species, and the economic impact has been significant [30]. At present, information on the genetic and immunological basis of the rock bream fish is limited. Better understanding of the mechanism of the innate immune response in rock bream will facilitate the development of effective approaches for disease prevention.

In the present study, we have identified and characterized the g-type lysozyme from rock bream (denoted as RBgLyz). Phylogenetic analysis was conducted to determine the evolutionary relationships. Gene expression analysis was performed in several tissues from healthy rock bream, and the expression responses were also analyzed following challenge with LPS, polyinosinic:polycytidylic acid (poly I:C), bacteria and virus. To understand the lytic activity of RBgLyz, we purified the recombinant protein using a prokaryotic system, and determined its spectrum of activity against different types of microorganisms. In addition, the morphological changes of bacteria treated with the recombinant RBgLyz were observed by scanning electron microscopy (SEM).

## 1.5 MATERIALS AND METHODS

### 1.5.1 Experimental fish

Healthy rock breams with an average body weight of 30 g were obtained from the Jeju Special Self-Governing Province Ocean and Fisheries Research Institute (Jeju, Republic of Korea). Rock breams were maintained in 400 L tanks with aerated sand-filtered seawater (salinity 34±0.6%, pH 7.6±0.5) at 23±1ºC, and no feeds were supplied. During the experimental challenge, no mortality was observed in experimental tanks.

### 1.5.2 Identification of rock bream g-type lysozyme

We have established a rock bream cDNA sequence database by using the Roche 454 Genome Sequencer FLX System (GS-FLX$^{TM}$), a next-generation DNA sequencing technology using the GS-FLX Titanium instrument [31]. Briefly, total RNA was isolated using Tri Reagent$^{TM}$ (Sigma, USA) from pooled tissues (pituitary gland, brain, gill, blood, liver, spleen, head kidney and kidney) of three healthy rock breams. Then, the mRNA was purified using an mRNA isolation kit (FastTrack® 2.0; Invitrogen, USA). The first strand cDNA synthesis and normalization were carried out with the Creator$^{TM}$ SMART$^{TM}$ cDNA library construction kit (Clontech, USA) and Trimmer cDNA normalization kit (Evrogen, Russia), respectively. Thereafter, the sequencing of rock bream cDNA was performed by Roche GS-FLX$^{TM}$ system (DNA Link, Inc.). From the rock bream cDNA sequence database, we identified a g-type lysozyme gene by means of the Basic Local Alignment Tool (BLAST) program in the National Center for Biotechnology Information (NCBI) web-based query system

(http://www.ncbi.nlm.nih.gov/BLAST).

### 1.5.3 Sequence alignments and phylogenetic analysis of RBgLyz

The open reading frame and amino acid sequence of RBgLyz were obtained by using DNAssist 2.2 (Version 3.0). The sequences of g-type lysozyme from different species were compared by the NCBI BLAST search program (http://www.ncbi.nlm.nih.gov/).

Pairwise and multiple sequence alignment (http://www.ebi.ac.uk/Tools/emboss/align) (http://www.ebi.ac.uk/Tools/clustalw2) were performed using the ClustalW version 2.0 program [32]. Identification of putative secretion features of the amino acid sequences was performed using SignalP3.0 server (http://www.cbs.dtu.dk). The deduced amino acid sequence was analyzed with the Expert Protein Analysis System (http://www.expasy.org/), while protein domain features of RBgLyz were determined by the Simple Modular Architecture Research Tool (SMART) program [33]. The phylogenetic relationship of g-type lysozyme was determined using the Neighbor-Joining method and MEGA version 4 [34].

### 1.5.4 Tissue collection, total RNA extraction and cDNA synthesis

To study the differential expression of RBgLyz *in vivo*, blood (1 mL per fish) was collected from the caudal fin of three healthy rock breams. The sample was immediately centrifuged at 3000×g for 10 min at 4ºC to facilitate blood cell collection, and samples were stored in liquid nitrogen until further use. Various tissues (gill, liver, spleen, head kidney, kidney, skin, muscle, brain and intestine) were excised and immediately snap-frozen in liquid nitrogen and stored at -80ºC until the

total RNA was isolated. All samples were obtained and analyzed from three biological replicates.

Total RNA was extracted from rock bream tissues by means of the Tri Reagent™ (Sigma), as described in the manufacturer's protocol. RNA concentration was determined at 260nm in a UV-spectrometer (Bio-Rad, USA) and diluted to 1 μg/μL concentration before synthesis of cDNA was carried out. A sample of 2.5 μg RNA was used to synthesize cDNA. The respective cDNAs from selected tissues were synthesized by the SuperScript™ III First Strand cDNA Synthesis kit (Invitrogen). Briefly, RNA was incubated with 1 μL of 50 μM oligo$(dT)_{20}$ and 1 μL of 10 mM dNTPs for 5 min at 65ºC. After incubation, 2 μL of 10× cDNA synthesis buffer, 4 μL of 25mM $MgCl_2$, 2 μL of 0.1M dithiothreitol (DTT), 1 μL of RNaseOUT™ (40 U/μL) and 1 μL of SuperScript™ III reverse transcriptase (200 U/μL) were added and the mixture was incubated for 50 min at 50ºC. The reaction was terminated by adjusting the temperature to 85ºC for 5 min. Then, 1 μL of RNase H was added to each cDNA reaction and incubated at 37ºC for 20 min. Finally, the newly synthesized cDNA was diluted 40 times (total 800 μL) before storage at -20ºC until use in further analysis.

### 1.5.5 Immune challenge

Rock breams were immune-challenged using several stimulants or pathogens in time course experiments.

For the mitogen stimulation experiment, each rock bream was administered a single intraperitoneal (i.p.) injection of 100 μL LPS in phosphate buffered saline (PBS) suspension (1.25 μg/μL, *E. coli* 055:B5; Sigma) or 100 μL poly I:C in PBS suspension (1.5 μg/μL Sigma). For the bacterial-challenge experiment, *E. tarda* and *S.*

*iniae* were obtained from the Department of Aqualife Medicine, Chonnam National University, Korea. The bacteria were incubated at 25°C for 12 h using a brain heart infusion (BHI) broth (Eiken Chemical Co., Japan) supplemented with 1% sodium chloride. The cultures were resuspended in sterile PBS, and then diluted to a desired concentration. Each rock bream was i.p. injected with 100 µL live *E. tarda* in PBS ($5\times10^3$ CFU/µL) or 100 µL live S. iniae in PBS (1×105 CFU/µL). For the virus challenge experiment, kidney tissue specimens obtained from the moribund rock bream infected with RBIV were homogenized in 20 volumes of PBS. The tissue homogenate was centrifuged at 3000×g for 10 min at 4°C, and the supernatant of a RBIV sample was filtered through a 0.45 µm membrane. Each rock bream was infected with a single i.p. injection of 100 µL RBIV in PBS. A negative control group was established as non-injected, while a positive control group was injected with an equal volume (100 µL) of PBS. Rock bream head kidney samples were collected at 3, 6, 12, 24 and 48 h post-injection from LPS-, poly I:C-, *E. tarda*-, *S. iniae*- or RBIV-infected rock breams. PBS-injected control samples were isolated at 3 h and 48 h post-injection. Three replicate rock breams were obtained for each time point and the pooled tissues from each group were subjected to total RNA extraction and cDNA synthesis according to the procedure described in section 2.4.

### 1.5.6 RBgLyz mRNA expression analysis by quantitative real-time PCR (qRT-PCR)

The mRNA expression of RBgLyz was analyzed by qRT-PCR. qRT-PCR was carried out using the Real Time System TP800 Thermal Cycler Dice™ (TaKaRa, Japan) in a 20 µL reaction volume containing 4 µL of diluted cDNA (3.125 ng/µL) from each respective tissue, 10 µL of 2× TaKaRa Ex Taq™ SYBR premix, 0.5 µL of each primer (10 pmol/µL) and 5 µL $dH_2O$. The qRT-PCR cycle profile included

one cycle of 95ºC for 10 s, followed by 35 cycles of 95ºC for 5 s, 58ºC for 10 s and 72ºC for 20 s, and a final single cycle of 95ºC for 15 s, 60ºC for 30 s and 95ºC for 15 s. The baseline was set automatically by the Thermal Cycler Dice™ Real-Time System Software (version 2.00). G-type lysozyme expression was determined by the Livak ($2^{-\Delta\Delta CT}$) method [35]. The same qRT-PCR cycle profile was used for the reference gene, rock bream β-actin (GenBank accession no. FJ975146). The primers used in this study are presented in Table 1. All data have been presented in terms of relative mRNA expressed as means ± standard deviation (S.D.). The data were subjected to a one-way analysis of variance (ANOVA), followed by Duncan's Multiple Range test, using the SPSS 11.5 program (USA). P values less than 0.05 were considered statistically significant.

### 1.5.7 Cloning RBgLyz coding sequences into pMAL-c2X

Two primers RBgLyz-F5 and RBgLyz-F6 were designed after checking the restriction enzymes sites of coding sequences of RBgLyz (Table 1.1). Then, RBgLyz coding sequences were cloned into maltose binding protein (MBP)-fused expression vector pMAL-c2X (NEB, USA). Briefly, a 50 µL PCR reaction was carried out with 2.5 U of LA Taq polymerase (TaKaRa), 5 µL of 10× LA PCR buffer II ($Mg^{2+}$ free), 5 µL of 25 mM $MgCl_2$, 8 µL of 2.5 mM dNTP mixture, 25 ng of template (cDNAs from head kidney) and 30 pmol of each primer. The reaction was carried out with an initial incubation at 94ºC for 2 min, 30 cycles of 94ºC for 30 s, 55ºC for 30 s, 72ºC for 60 s, and followed by a final extension at 72ºC for 5 min. The PCR product (approximately 564 bp) was analyzed on a 1% agarose gel stained with ethidium bromide. Then, the PCR product was purified using the Accuprep™ gel

**Table 1.1 Primer sequences and their use**

| Primer name | Purpose | Primer sequences (5′→3′) |
|---|---|---|
| RBgLyz-F1 | qRT-PCR amplification | GGATCCGCAACAAATTTCCTGGCT |
| RBgLyz-R2 | qRT-PCR amplification | TACCACTGAGCTCTGGCAACAACA |
| β-actin-F3 | qRT-PCR reference gene | TCATCACCATCGGCAATGAGAGGT |
| β-actin-R4 | qRT-PCR reference gene | TGATGCTGTTGTAGGTGGTCTCGT |
| RBgLyz-F5 | ORF amplification | $(GA)_3$GAATTCATGGGTTACGGAAACATCATGATGGT(*EcoRI*) |
| RBgLyz-R6 | ORF amplification | $(GA)_3$CTGCAGTTAAAAGCCTCCGTTGGTTTGGTAC(*Pst*I) |
| RBgLyz-F7 | Sequencing | GGTCGTCAGACTGTCGATGAAGC |
| RBgLyz-R8 | Sequencing | GATGTGCTGCAAGGCGATTAAGTTGG |

The number of identical base pair repeats is indicated within parentheses with the appropriate subscripted number.

Restriction enzyme sites of the cloning primers are underlined.

purification kit (Bioneer Co., Korea), after which both insert (PCR product) and vector (pMAL-c2X) were digested with the restriction enzymes *EcoR*I and *Pst*I. Ligation was carried ou tat 16ºC overnight with 50 ng of vector, 100 ng of insert, 1 μL of 10× ligation buffer and 0.5 μL of 1× T4 DNA ligase (TaKaRa). The ligated product was transformed into DH5α cells and the correct recombinant product (as confirmed by sequencing with RBgLyz-F7 and RBgLyz-F8) was transformed into competent *E.coli* BL21(DE3) cells for protein expression.

### 1.5.8 Purification of recombinant RBgLyz protein (rRBgLyz)

The recombinant RBgLyz protein was expressed in *E. coli* BL21 (DE3) cells by isopropyl-β-thiogalactopyranoside (IPTG) induction. Briefly, 10 mL volume of *E. coli* BL21 (DE3) starter culture was inoculated in 100 mL of Luria broth (LB) medium with ampicillin (100 μg/mL final concentration) and glucose (2% final concentration) and grown overnight at 37°C with shaking at 200 rpm. Overnight culture was diluted 1:10 into a larger volume of fresh medium and further cultured at 37°C until the optical density (OD) at 600nm reached 0.5. Cells (500 μL aliquots: uninduced cells) were then induced by IPTG (0.5 mM final concentration) for 3 h at 30°C (500 μL aliquots: induced cells). The cells were harvested by centrifugation at 3000×g for 30 min at 4°C, and resuspended with column buffer (Tris-HCl, pH 7.4, 200 mM NaCl, 0.5 M EDTA). The recombinant RBgLyz (rRBgLyz) was purified in the form of fusion protein with MBP by pMAL™protein fusion and purification system (NEB, USA). Briefly, the resuspended *E. coli* BL21 (DE3) cells were placed in an ice-water bath and sonicated 6-8 times in short pulses for 15 s. Then, the sonicated cell suspension was centrifuged at 15000 rpm for 30 min at 4 °C and the resulting supernatant considered as crude rRBgLyz extract. Meanwhile, amylase resin was poured into a 1×5 cm size column and washed with eight volumes of column buffer. In the final purification step, the crude extract was loaded onto the column and

allowed to settle for 20 min, followed by washing with 12 volumes of column buffer. Finally, the rRBgLyz fusion protein was eluted by applying 3 ml of elution buffer (column buffer plus 10 mM maltose) in 500 μL aliquots, and then subsequently analyzed on 12% SDS-PAGE with a protein molecular weight marker (Takara). Protein was visualized by staining with 0.05% Coomassie Blue R-250, followed by the standard de-staining procedure. All of the activity tests performed in this study were conducted using this purified rRBgLyz fused with MBP. The concentration of the purified protein was determined at 280nm in a UV-spectrometer (Bio-Rad).

### 1.5.9 Determination of optimum pH on the activity of the recombinant rRBgLyz

The optimal pH of the recombinant RBgLyz was measured by lysoplate assay, as described by Minagawa *et al*.[36]. Briefly, a gel plate was prepared that contained 1% agarose in 50 mM sodium phosphate buffer with pH's ranging from 4.5 to 8.5 (at intervals of 0.5) and *Micrococcus lysodeikticus*. The concentration of *M. lysodeikticus* was adjusted to 0.4 at $OD_{600}$. Fifty microliters of MBP, standard hen egg white lysozyme (HEWL, sigma) or RBgLyz (500 μg/mL) or a standard hen egg white lysozyme (HEWL) were applied to a well (diameter: 5 mm) in the agarose plate and incubated at 30˚C. After 24 h, the diameter of the clearing zone was measured (mm).

### 1.5.10 Antimicrobial activity of rRBgLyz

Antimicrobial activity of the recombinant lysozyme was assessed with turbidimetric assay, as described by Hultmark [10]. Several bacterial strains and one fungus (yeast) were obtained from the Korean Culture Center of Microorganisms (KCCM) and Korean Collection for Type Cultures (KCTC), including Gram-negative bacteria *Vibrio alginolyticus* (KCCM 40513), *V. salmonicida* (KCCM 41663), *V.*

*parahaemolyticus* (KCCM 11965), *Escherichia coli* (KCCM 70089), Gram-positive bacteria *Listeria monocytogenes* (KCCM 40307), *M. lysodeikticus* (KCTC 1056), and yeast *Pityrosporum ovale* (KCCM 11894). These microbial cells were used as substrates and suspensions of each were prepared in 50 mM sodium phosphate buffer (pH 4.5) adjusted to various concentrations ($OD_{600}$=0.1−0.6). Fifty microliters of MBP (500 μg/mL), HEWL (500 μg/mL) and rRBgLyz (500 μg/mL) were mixed separately with 150 μL of the suspensions, and the initial $OD_{600}$ (designated as $OD_i$) was measured using a spectrophotometer (Bio-Rad). The mixture was incubated in a water bath in 37˚C for 30 min and transferred to an ice bath to stop the reaction immediately. $OD_{600}$ was measured again and designated as $OD_f$. A blank sample was prepared using sodium phosphate buffer (200 μL). Lytic activity (UL) was calculated by using the difference between the initial ($OD_i$) and final ($OD_f$) turbidity, according to the following formula: $UL = (OD_i - OD_f)/OD_f$. Triplicate measurements were performed.

### 1.5.11 Scanning electron microscope (SEM) observation

A suspension of *M. lysodeikticus* ($OD_{600}$=0.2) was prepared in 50 mM sodium phosphate buffer (pH4.5). Then, 300 μL aliquots of the suspension were dispensed into sterile tubes and 500 μL of MBP (500 μg/mL), rRBgLyz (500 μg/mL), or HEWL (500 μg/mL) were added. All tubes were incubated in a water bath for 30 min at 37˚C. The third tube for each was used as control. After the incubation, samples were fixed for 10 h at 4˚C with 2.5% glutaraldehyde containing 0.1 M sodium phosphate buffer (pH 7.3). After being washed with 0.1 M sodium phosphate buffer (pH 7.3), fixed cells were dehydrated by passage through a graded series (30%, 50%, 70%, 90%, and 100%) of ethanol and then followed by substitution processes using the same concentration series of isoamyl acetate. Next, the samples were dried with liquid $CO_2$ at a critical point followed by platinum coating by a sputter coater at 10 mA for 2 min. Finally, morphology differences of *M.*

*lysodeikticus* cells in response to the recombinant RBgLyz treatment were examined by field emission SEM (JEOL Ltd., Japan) and compared to the untreated controls.

## 1.6 RESULTS

### 1.6.1 Molecular characterization of RBgLyz

The nucleotide and deduced amino acid sequences of RBgLyz are shown in Figure 1.1; the 669 bp nucleotide sequence of RBgLyz was deposited in GenBank under accession no. HM035063. The sequence contained an open reading frame (ORF) of 567 bp which encoding a polypeptide of 188 amino acid residues, a 5′ -terminal untranslated region (UTR) of 4 bp, a 3′-terminal UTR of 98bp with a stop codon (TAA), and a putative polyadenylation consensus signal (AATAAA) at the 70 bp followed by the stop codon. The calculated molecular mass was 20 kDa with a theoretical isoelectric point of 6.0. SignalP program was unable to detect any signal peptide sequence. The deduced RBgLyz possessed no cysteine residues that would be expected to generate disulfide bridges in the avian and mammalian g-type lysozyme.

Three amino acid residues (Glu71, Asp84 and Asp95), potentially important for the lysozyme catalytic activity, were found to be completely conserved in RBgLyz (Fig. 2, light gray underlay). Four amino acid residues (Leu91, Leu115, Leu122 and Gly145) that were considered to be involved in binding the substrate were also identified in the corresponding site of RBgLyz (Fig. 1.2, dark gray underlay). In addition, the catalytic bacterial soluble lytic transglycosylase (SLT) domain, located from Ser48 to Ser172, in RBgLyz was found to be well conserved (Fig. 1.2, boxed).

BLAST search indicated that the deduced amino acid sequences of RBgLyz shared significant identity with other reported g-type lysozymes. RBgLyz presented 54.2% identity with common carp *Cyprinus carpio* (BAB91437), 55.3% with grass carp *Ctenopharyngodon idella* (ACF41165), 66.1% with Atlantic cod *Gadus morhua* (EU377606), 73.0% with Japanese flounder *Paralichthys olivaceus* (BAB62406), 75.8% with large yellow croaker *Larimichthys crocea* (ABR66916), 77.3% with brill *Scophthalmus rhombus* (BAF75845), and 81.4% with Chinese perch *Siniperca chuatsi*

(AAU86896) (Fig. 1.2).

### 1.6.2 Phylogenetic analysis of RBgLyz

To classify and analyze the molecular evolution of RBgLyz, a phylogenetic tree of lysozyme amino acid sequences was constructed using 19 lysozymes with 13 g-type, four c-type and two i-type lysozymes by the Neighbor-Joining method. As expected, RBgLyz was clearly grouped in the same clade as other g-type lysozymes. The c- and i-type lysozymes also formed each cluster individually. The relationship revealed in the phylogenetic tree was in agreement with homology comparison with other g-type lysozymes of teleosts; RBgLyz presented the closest distant relationship with sequences from Chinese perch g-type lysozyme. In this way, the closer distant relationship corresponded to brill g-type lysozyme, followed by Japanese flounder, large yellow croaker, Atlantic cod, grass carp and common carp (Fig. 1.3).

### 1.6.3 Expression of RBgLyz in healthy rock breams

Distribution of RBgLyz mRNA transcripts was examined by qRT-PCR in tissues of blood, gill, liver, spleen, head kidney, kidney, skin, muscle, brain and intestine of healthy individuals (Fig. 1.4). The expression level of each tissue examined was normalized to that of β-actin. Relative expression fold differences were calculated based on the expression in muscle to determine the tissue expression profile. RBgLyz was found to be constitutively expressed in all the tissues examined, although its level of expression was variable. The mRNA expression levels in blood, gill and kidney were higher than in the muscle by 112.7-fold, 85.6-fold and 74.1-fold, respectively. A higher mRNA level was also detected in head kidney (H. kidney), liver, intestine and spleen (35.0-fold, 32.8-fold, 32.0-fold and 31.9-fold higher than in muscle, respectively). Lower expression levels (<12-fold of that in muscle) were found in other tissues.

### 1.6.4 Expression of RBgLyz in rock breams induced with LPS, poly I:C, *E. tarda*, *S. iniae* and RBIV

In order to determine whether RBgLyz expression was elevated in rock bream fish after several immune challenges, six groups of fish were i.p. injected with either LPS, poly I:C, *E. tarda*, *S. iniae*, RBIV or PBS, and then the transcriptional response of RBgLyz in head kidney was analyzed at 3, 6, 12, 24 and 48 h post-infection. No significant difference was observed in non-injected control and PBS-injected controls at 3 and 48 h p.i. Therefore, non-injected sample was used to determine the transcriptional regulation in this study.

Diverse expression patterns were induced by different stimuli (Fig. 1.5). Among the five stimulators, LPS and *E. tarda* were the most capable of inducing the expression of RBgLyz. Up-regulation of RBgLyz transcription by LPS treatment was first detected at 3 h p.i., being 2.3-fold higher than the expression level of non-injected control, then enhanced to 7.0-fold at 6 h post-infection, and reaching the highest level (13.0-fold) at 12 h. With *E. tarda* infection, RBgLyz transcription slightly increased at 12 h post-infection (1.3-fold compared to control), then reached its highest level at 24 h (5.9-fold) and decreased (to 0.9-fold) at 48 h. In comparison with LPS and active *E. tarda*, polyI:C stimulated a moderate expression of RBgLyz; the highest transcription levels were detected at 6 h (2.5-fold) and 12 h (2.4-fold). The other stimuli, *S. iniae*, also slightly stimulated the expression levels of RBgLyz at 12 h (1.4-fold) and 24 h (1.3-fold). However, the expression pattern of RBgLyz induced by active RBIV infection was quite different from those of the other stimuli. Specific time-dependent effects on RBgLyz mRNA expression after RBIV infection were not observed.

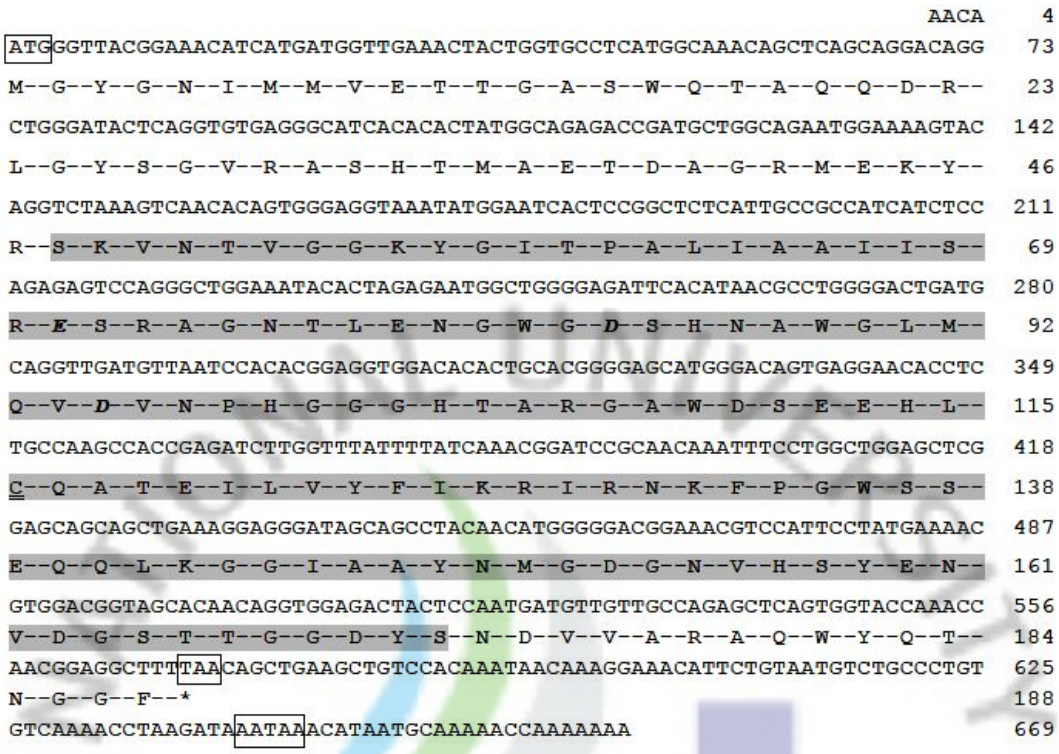



**Figure 1.1 Nucleotide (GenBank accession no. HM035063) and deduced amino acid sequences of RBgLyz in *Oplegnathus fasciatus*.** The nucleotide and amino acids are numbered along the right margin. The boxed letters indicate the start codon (ATG), the stop codon (TAA) and the putative polyadenylation signal (AATAAA). The SLT domain in the C-terminus extends from position 48-172 (gray underlay). The catalytic residues are denoted by bold italicized letters (Glu [71], Asp [84] and Asp [95]).


**Figure 1.2 ClustalW2 multiple alignment of RBgLyz with g-type lysozymes from various fish species.** The amino acids are numbered along the right margin. Identical residues in all sequences are indicated by asterisks (*) under the column. Residues strongly and weakly conserved are indicated by colons (:) and dots (.), respectively. The dashes indicate the gaps introduced to maximize alignment. Light gray and dark gray underlays indicate the proposed catalytic residues and the substrate binding sites of g-type lysozymes, respectively. The SLT domain is boxed.

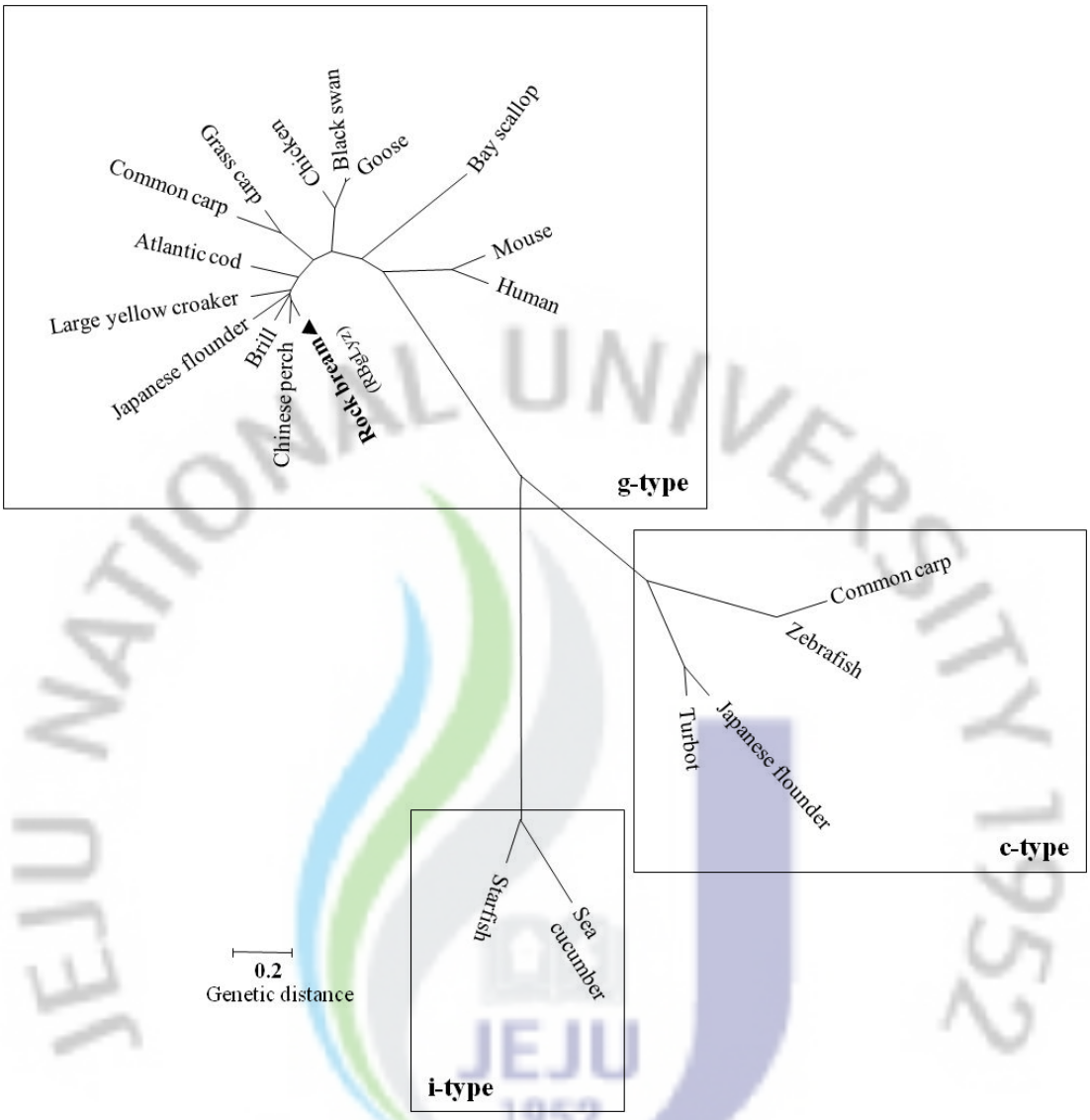



**Figure 1.3 Phylogenetic tree based on the amino acid sequences of lysozymes.** The tree was constructed by the Neighbor-Joining method using MEGA version 4 and based on sequence alignment using ClustalW2. Bootstrapping was performed 1000 times. The scale bar corresponds to 0.2. GenBank accession numbers of the selected g-type lysozyme sequences are shown as follows: rock bream *Oplegnathus fasciatus* (RBgLyz) HM035063; Chinese perch *Siniperca chuatsi* AAU86896; brill *Scophthalmus*

*rhombus* BAF75845; Japanese flounder *Paralichthys olivaceus* BAB62406; large yellow croaker *Larimichthys crocea* ABR66916; Atlantic cod *Gadus morhua* EU377606; common carp *Cyprinus carpio* BAB91437; grass carp *Ctenopharyngodon idella* ACF41165; chicken *Gallus gallus* CAA43319; black swan *Cygnus atratus* P00717; goose *Anser anser anser* P00718; bay scallop *Arogopecten irrradians* AAX09979; mouse *Mus musculus* AAI19397; human *Homo sapiens* AAH29126. The accession numbers of c-type lysozyme are common carp *Cyprinus carpio* AB027305; zebrafish *Dinio rerio* NP_631919; Japanese flounder *Paralichthys olivaceus* AB050469; turbot *Scophthalmus maximus* AJ250732. The i-type lysozymes are sea cucumber *Stichopus japonicas* EF036468 and starfish *Asterias rubens* AY390770.

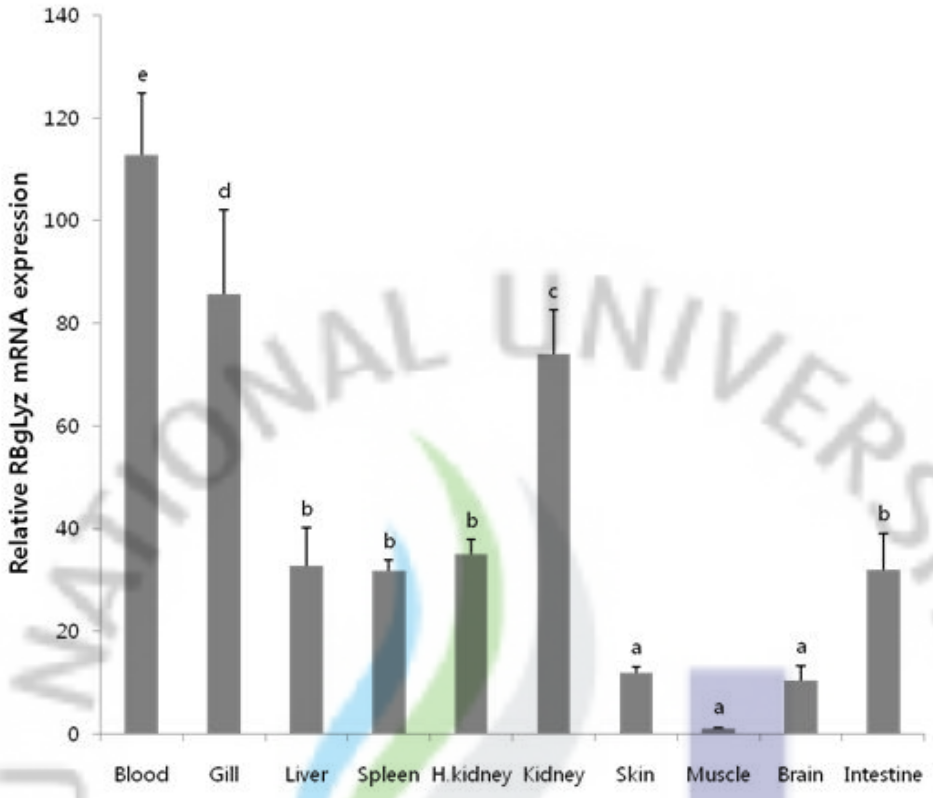



**Figure 1.4 Relative mRNA expression of RBgLyz in different tissues, including blood, gill, liver, spleen, head kidney (H. kidney), kidney, skin, muscle, brain, and intestine.** Expression values were normalized to those of rock bream β-actin. The bars indicate the mean fold change (mean ± S.D., *n*=3) from the calibrator group (muscle). Means with different letters are significantly different ($P<0.05$).

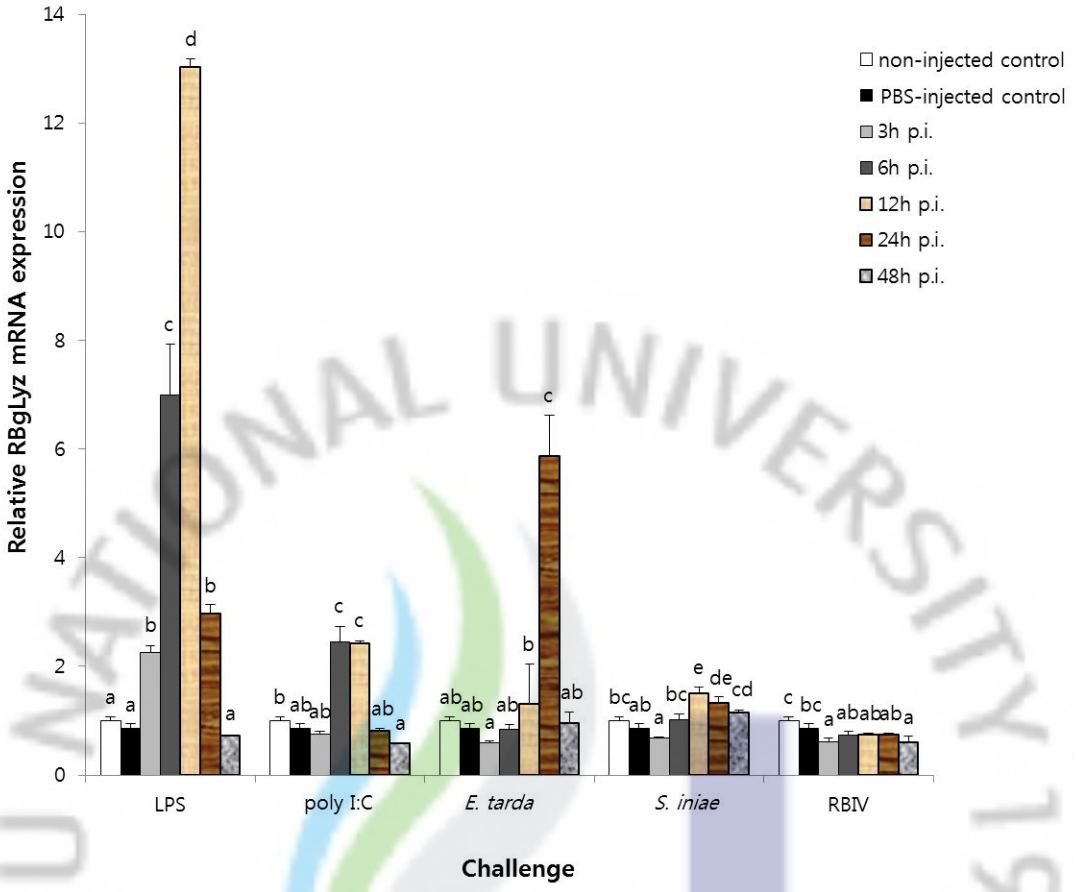



**Figure 1.5 Relative RBgLyz mRNA expression in head kidney (H. kidney) in response to challenges with LPS, poly I:C, *E. tarda, S. iniae* or RBIV at 3, 6, 12, 24 and 48 h post-injection.** Con(-) and con(+) indicate non-injected and PBS-injected control samples, respectively. The bars indicate the mean fold change (mean ± S.D., *n*=3) from non-injected controls. Means with different letters are significantly different ($P<0.05$).

### 1.6.5 Expression and purification of recombinant RBgLyz

The cloned RBgLyz coding sequence was inserted down-stream from *malE* gene of *E. coli*, which encodes the MBP under the strong tac promoter and *malE* translation initiation signal of pMAL-c2X. The re combinant RBgLyz/pMAL-c2X protein was expressed in *E. coli* BL21(DE3) cells by IPTG induction, and a 62.5 kDa fusion protein (RBgLyz 20 kDa + MBP 42.5 kDa) was purified by exploiting MBP's affinity for maltose. Aliquots of different steps during the purification process were analyzed by 12% SDS-PAGE with Coomassie Blue (Fig. 1.6).

We observed a strong fusion RBgLyz band in the induced extract (lane 2) compared to the uninduced crude protein extract (lane 1). The sample containing the fusion protein was characterized by gel analysis as being of high purity (lane 3).

### 1.6.6 Lysozyme activity

The lytic activity of the recombinant RBgLyz was determined by lysoplate assay at different pH's. The results showed that RBgLyz possessed the highest lysozyme activity at relatively low pH value (pH = 4.5), and the activity gradually decreased at pH's above pH 4.5. MBP showed no lytic activity against *M. lysodeikticus*, while HEWL activity at pH 8 was clearly detected (Fig. 1.7).

The antimicrobial activity of the recombinant RBgLyz was evaluated by turbidimetric assay carried out with different types of microorganisms, including Gram-positive, Gram-negative bacteria and yeast. MBP and HEWL was used as control. The results indicated that lytic activity of the recombinant RBgLyz was active against Gram-negative bacteria V. salmonicida, Gram-positive bacteria *L. monocytogenes*, *S. iniae* and *M. lysodeikticus* ($P<0.05$) (Fig. 1.8). On the other hand, HEWL possessed high ($P<0.05$) activity against *M. lysodeikticus*, but low activity against *L. monocytogenes* (Fig. 1.8), *V. alginolyticus*, *V. parahaemolyticus* and *P. ovale* (data not shown). Neither RBgLyz nor HEWL showed any antimicrobial effect

against *E. tarda* or *E.coli* (data not shown). MBP possessed no antimicrobial activity against all microorganisms used in this study.

In addition, we examined the morphology change of *M. lysodeikticus* in response to the recombinant RBgLyz treatment, under different concentrations, as evidenced by SEM imaging. As shown in Figure 1.9, deformed cells with rough, wrinkled surfaces were present in rRBgLyz- or HEWL- treated samples (C, D), as compared to the untreated or MBP control (A, B).

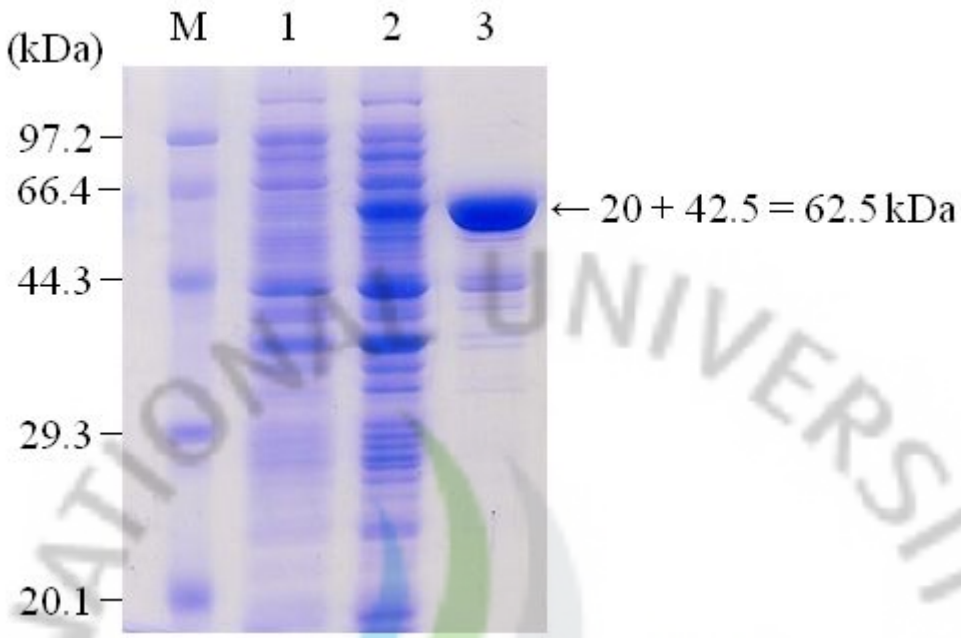



**Figure 1.6 Overexpression and purification of recombinant RBgLyz fusion protein.** Protein samples were separated by 12% SDS-PAGE and stained with Coomassie Blue. Lane M: protein molecular weight marker (Takara). Lane 1: total cellular extract from *E. coli* BL21 (DE3) before induction. Lane 2: total cellular extract after induction with 0.5 mM IPTG. Lane 3: purified recombinant RBgLyz fusion protein. The target protein is indicated by an arrow.

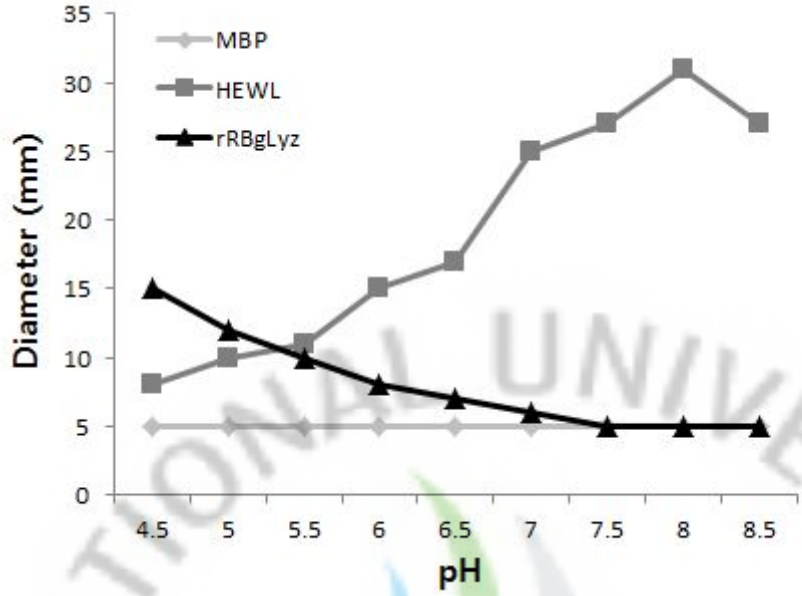



**Figure 1.7 Optimal pH of recombinant RBgLyz as measured by a lysoplate assay.**

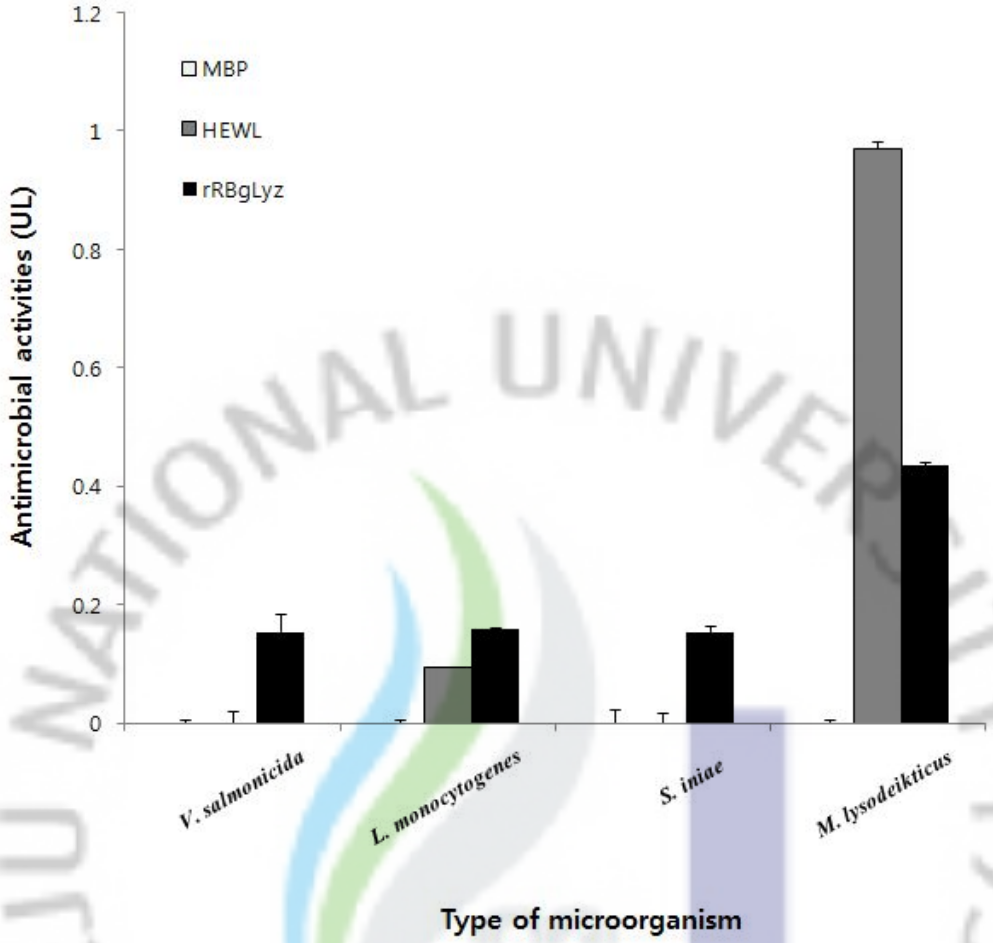



**Figure 1.8 Antimicrobial activity (UL) of the recombinant RBgLyz, as measured by turbidimetric assay.** Lytic activities are expressed as the mean UL (mean ± S.D., *n*=3).

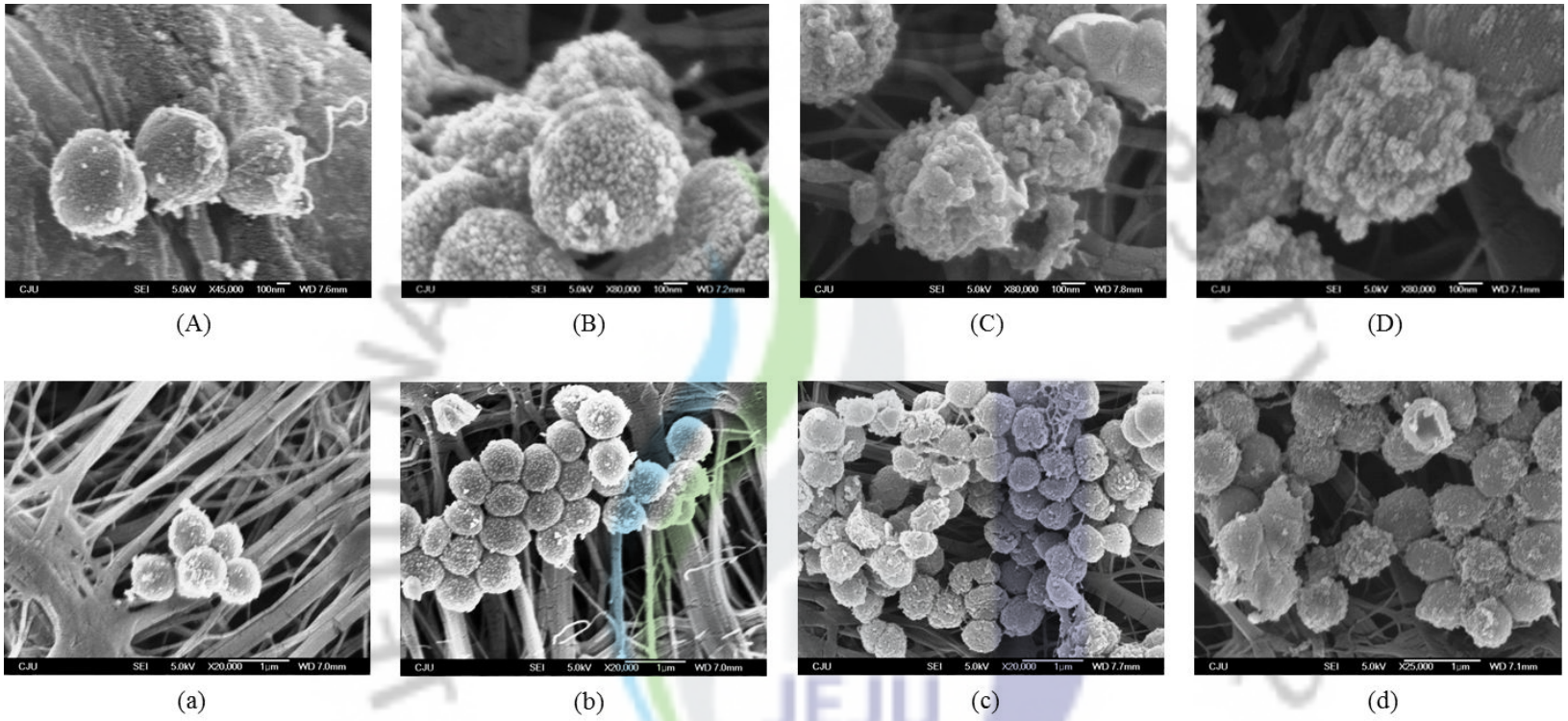



**Figure 1.9** SEM images of M. lysodeikticus after treatment with the recombinant RBgLyz. A, a: untreated control cells (A: magnification ×45,000, a: ×20,000); B, b: cells treated with 500 μL of MBP (500 μg/mL) (B: ×80,000, b: ×20,000); C, c: cells treated with 500 μL of rRBgLyz (500 μg/mL) (C: ×80,000, c: ×20,000). D, d: cells treated with 500 μL of MBP (500 μg/mL) (B: ×80,000, b: ×25,000).

## 1.7 DISCUSSION

The innate immune system in fish has become a focus as researchers attempt to establish a firm understanding of disease resistance in these species. Lysozyme has been characterized as an important protein of the innate immune response, and has been evolutionarily conserved throughout vertebrates, invertebrates, phages, bacteria and plants [2]. Even though lysozyme has been shown to play a significant role in defense against infectious disease in fish, its underlying functional mechanisms remain to be completely understood [2]. In fish, the g-type lysozyme has been identified in Japanese flounder [20], common carp [21], orange-spotted grouper [22], Chinese perch [23], large yellow croaker [24], Atlantic salmon [25], brill [26], Atlantic cod [27], and grass carp [28] and rohu [29].

In the present study, we performed transcriptome sequencing of rock bream using Roche 454 GS-FLX sequencing technology, and we identified a g-type lysozyme that exists in rock bream, designated as RBgLyz. The RBgLyz was characterized as being 567 bp in length, encodes a protein of 188 amino acids, and has a theoretical molecular weight of 20 kDa. Interestingly, this RBgLyz did not include a putative signal peptide, as did the g-type lysozymes found in birds [17], mammals [11] and the recently published two fish species, Atlantic salmon (GenBank accession no. CAM35431) [25] and Atlantic cod (Codg1; GenBank accession no. EU377606) [27]. This fact suggested that the encoded RBgLyz protein(s) were not secreted. However, one particular structural feature was conserved among the g-type lysozymes from birds and mammals, that being four conserved cysteine residues that are involved in two pairs of disulfide bridges (Cys4-Cys60 and Cys18-Cys29) [19]. Kawamura et al. [37] described that the formation of these two disulfide bonds in g-type lysozyme from ostrich was not essential for correct folding into the catalytically-active

conformation, but was crucial for structural stability. The authors also described the formation of the two disulfide bonds of g-type lysozyme as occurring late in the folding process, and in an independent rather than sequential manner. Since Atlantic cod and Atlantic salmon live at temperatures close to the freezing point, the lack of the disulfide bonds could create a more flexible structure necessary for the g-type lysozymes to support their continued function at low temperatures [25,27]. However, the lack of disulfide bridges was a common trait of most fish g-type lysozymes, even for more temperate fish species [37]. In all these cases, the observation that RBgLyz lacked the disulfide bonds supported the presumption that RBgLyz was likely an intracellular protein.

The three amino acid residues (Glu71, Asp84 and Asp95) in RBgLyz were well conserved in other fish [20,22-24,26-28] and avian species [19]. The only exception was that the second catalytic residue, Asp84, in common carp was replaced by Pro86. Kawamura et al. [38] verified that Glu73 (corresponding to Glu71 in RBgLyz) was essential for catalytic activity and structural stability of ostrich g-type lysozyme, based on a site-directed mutational analysis and X-ray structural studies. Likewise, Helland et al. [39] demonstrated similar crucial roles in the structural stability and catalytic activity of Atlantic cod g-type lysozyme. The other two aspartic acid residues of g-type lysozymes in Atlantic cod were also shown to be involved in catalysis, but the third catalytic residue, aspartic acid, appeared to be more critical to the function than the second one [39]. The three conserved residues (Glu71, Asp84 and Asp95) in the RBgLyz presented in this study are theorized to have similar catalytic activity with those of other fish and avian species. SMART program analysis revealed that RBgLyz contained the SLT domain which is known to be involved in the cleavage of most bacterial cell walls by acting upon the β -(1,4)-glycosidic bonds between the NAG and NAM in the peptidoglycan layer.

Originally, the SLT domain was defined as the catalytic region of a bacterial Slt70 enzyme, and was also characterized as a novel class of lysozyme [40,41]. Thunnissen et al. [41] also concluded that the g-type lysozyme, in contrast to the c- and phage-type lysozymes, commonly harbored an SLT domain.

The expression pattern of the g-lysozyme gene has previously been investigated in various organisms. In chicken, a restricted expression pattern of the g-type lysozyme was found. It was only expressed in bone marrow and lung, and not in the oviduct, providing an explanation as to why its absence was observed in chicken egg white [18]. In contrast, g-lysozyme was quite abundant in the egg white of many other birds, including geese, ostriches and swans [19]. Nile et al. [42] identified a second chicken g-type lysozyme sequence (chicken g2) that was expressed in liver, kidney and intestine. In humans, two g-type lysozymes were identified, but neither of these genes was widely nor highly expressed in fetal and adult tissues [19]. Irwin and Gong showed that the mouse also carried two g-type lysozyme genes in its genome. The murine lysozyme g1 was found to be expressed at high level only in the tongue, whereas g2 was expressed in skin [19]. These restricted distribution patterns in birds and mammals contrasted with the broad expression patterns of g-type lysozyme in fish species. The g-type lysozyme has been detected in spleen, kidney, gill, skin, heart, intestine and blood of Japanese flounder, orange spotted grouper, large yellow croaker, Atlantic cod and grass carp [20,22,24,27,28], suggesting that g-type lysozyme was an intracellular protein. This finding was also consistent with the lack of disulfide bridges and a signal peptide. In this study, we also found that the RBgLyz was ubiquitously expressed in all analyzed tissues, including blood, gill, liver, spleen, head kidney, kidney, skin, muscle, brain and intestine. The level of RBgLyz mRNA was highest in blood, gill and kidney, which was consistent with previous suggestions that the g-type lysozyme gene in Japanese flounder and large

yellow croaker is expressed predominantly in tissues of organs exposed to the external environment or in hematopoietic tissues [20, 24].

Expression of g-type lysozyme has been shown to be up-regulated in fish treated with LPS or infected with bacteria. The expression of g-type lysozyme was increased in the intestine, heart and blood of Japanese flounder injected with *E. tarda* [20]. Orange-spotted grouper g-type lysozyme transcripts were up-regulated in the stomach, spleen, kidney, heart, brain and leukocytes after *V. alginolyticus* injection [22]. The mRNA levels of g-type lysozyme in large yellow croaker (called LycGL) were up-regulated in intestine, spleen and head kidney after induction by bacterial vaccine [24]. A clear g-type lysozyme expression response to LPS or *Photobacterium damselae* subsp. *piscicida* inoculation was observed in head kidney of brill [26]. Furthermore, expression level of g-type lysozyme was up-regulated in spleen, liver, head kidney and gill of grass carp challenged with *A. hydrophila* [28].

In the study presented herein, we also detected whether RBgLyz expression was elevated in rock bream fish after several immune challenges with either LPS, poly I:C, *E. tarda*, *S. iniae* or RBIV injection at different time points. As a major component of the Gram-negative bacterial cell wall, the LPS endotoxin elicits a defense response from the innate immune system [43]. The most robust up-regulation we detected for RBgLyz mRNA level in response to LPS exposure occurred in head kidney. *E. tarda*, (responsible for edwardsiellosis, and a Gram-negative bacterium) and *S. iniae* (causing streptococcosis, and a Gram-positive bacterium) have become one of the most problematic infectious outbreaks in rock bream, as well as in many other fresh and marine fish [30]. In rock bream injected with active *E. tarda*, RBgLyz transcripts were found to increase significantly at 24 h post-infection. Only slightly higher mRNA levels were observed after active *S. iniae* inoculation. Taken together, the observed patterns generally agreed with other previously published

information on g-type lysozyme responses to LPS or bacteria injection in other fish species [20,22,24,26,28], thus supporting a defensive role for RBgLyz against bacteria in rock bream.

Poly I:C, a synthetic double-stranded (ds) RNA, has been widely used as an inducer to study the antiviral state triggered by the immune system [44]. Our results showed RBgLyz moderately up-regulated transcripts after poly I:C induction. However, no increase or even a reduction of RBgLyz mRNA levels, was observed after injection of RBIV (ds DNA virus). To our knowledge, the expression of g-type lysozyme detected by qRT-PCR in our present study represents the first report of such in fish treated with poly I:C or infected with virus.

Saurabh and Sahoo [45] have reported that fish are poikilothermic organisms and, therefore, their lysozyme activity may not be as sensitive to temperature. But, pH is known to be an influential factor affecting the enzymatic activity of lysozyme in fish. The recombinant RBgLyz was found to possess the highest lysozyme activity at pH 4.5, with a decrease in activity observed above pH 4.5. This pH activity profile was different from those obtained for other g-type lysozymes using *M. lysodeikticus* as substrate. The Atlantic cod and salmon g-type lysozymes, studied previously by others, showed optimum activity occurred at pH 4.8 and pH 5.2, respectively [25,27], while the g-type lysozymes of Japanese flounder and grass carp have an optimum activity at pH 6.0 and pH 6.5 [22,28].

Other recombinant fish g-type lysozymes have been shown to have antimicrobial activities against fish pathogenic bacteria [20,22,24,28]. In the present study, recombinant RBgLyz also exhibited antimicrobial activity against bacteria *M. lysodeikticus*, and fish pathogenic bacteria, such as *L. monocytogene*s, *S. iniae* and *V. salmonicida*. The SEM observation carried out here provided evidence that recombinant RBgLyz is able to cause considerable morphological damage to treated

*M. lysodeikticus* cells.

In summary, we identified g-type lysozyme sequences from rock bream, for the first time, by applying a pyrosequencing technique. RBgLyz was constitutively expressed in various tissues, but the level of RBgLyz mRNA expression was highest in blood, gill, and kidney, which represent organs exposed to the external environment or hematopoietic tissues. Furthermore, RBgLyz transcription was up-regulated in response to exposure to LPS, poly I:C, *E. tarda*, or *S. iniae*. The recombinant RBgLyz also showed antimicrobial activity against bacteria *M. lysodeikticus*, and a variety of pathogenic bacteria. Additionally, SEM observation results confirmed that *M. lysodeikticus* cells were damaged by recombinant RBgLyz treatment. Collectively, these results indicate a functional role for RBgLyz in the innate immune defense system of rock bream.

## 1.8 REFERENCES


[1] Medzhitov R, Janeway C Jr. Innate immune recognition: mechanisms and pathways. Immunol Rev 2000;173:89−97.

[2] Magnadóttir B. Innate immunity of fish (overview). Fish Shellfish Immunol 2006;20: 137−51.

[3] Jollès P, Jollès J. What's new in lysozyme research? Always a model system, today as yesterday. Mol Cell Biochem 1984;63:165−89.

[4] Hancock RE, Scott MG. The role of antimicrobial peptides in animal defenses. Proc Natl Acad Sci U S A 2000;97:8856−61.

[5] Grinde B. Lysozyme from rainbow trout, *Salmo gairdneri* Richardson, as an antibacterial agent against fish pathogens. J Fish Dis 1989;12:95−104.

[6] Samarayanake YH, Samaranayake LP, Wu PC, So M. The antifungal effect of lactoferrin and lysozyme on *Candida krusei* and *Candida albicans*. APMIS 1997;105:875−83.

[7] Lee-Huang S, Huang PL, Sun Y, Huang PL, Kung HF, Blithe DL, et al. Lysozyme and RNases as anti-HIV components in β-core preparations of human chorionic gonadotropin. Proc Natl Acad Sci US A 1999;96:2678−81.

[8] Ibrahim HR, Thomas U, Pellegrini A. A helix-loop-helix peptide at the upper lip of the active site cleft of lysozyme confers potent antimicrobial activity with membrane permeabilization action. J Biol Chem 2001;276:43767−74.

[9] Zhang X, Wu FX, Sun M, Sang QY, Wang YJ, Wang XK. Study on antimicrobial and antiviral activities of lysozyme from marine strain S-12-86 in vitro. Agric Sci China 2008;7:112−6.

[10] Hultmark D. Insect lysozymes. In: Jollès P, editor. Lysozymes: model enzymes in biochemistry and biology. Basel: Birkhäuser Verlag; 1996. p. 87−102.

[11] Prager EM, Jollès P. Animal lysozymes c and g: an overview. In: Jollès P, editor. Lysozymes: model enzymes in biochemistry and biology. Basel: Birkhäuser Verlag; 1996. p. 9−31.

[12] Jollès J, Jollès P. The lysozyme from *Asterias rubens*. Eur J Biochem 1975;54:19−23.

[13] Fastrez J. Phage lysozymes. In: Jollès P, editor. Lysozymes: model enzymes in biochemistry and biology. Basel: Birkhäuser Verlag; 1996. p. 35−64.

[14] Holtje JV. Bacterial lysozymes. In: Jollès P, editor. Lysozymes: model enzymes in biochemistry and biology. Basel: Birkhäuser Verlag; 1996. p. 65−74.

[15] Beintema JJ, Terwisscha van Scheltinga AC. Plant lysozymes. In: Jollès P, editor. Lysozymes: model enzymes in biochemistry and biology. Basel: Birkhäuser Verlag; 1996. p. 75−86.

[16] Canfield RE, McMurry S. Purification and characterization of a lysozyme from goose egg white. Biochem Biophys Res Commun 1967;26:38−42.

[17] Prager EM, Wilson AC, Arnheim N. Widespread distribution of lysozyme g in egg white of birds. J Biol Chem 1974;249:7295−7.

[18] Nakano T, Graf T. Goose-type lysozyme gene of the chicken: sequence, genomic organization and expression reveals major differences to chicken-type lysozyme gene. Biochim Biophys Acta 1991;1090:273−6.

[19] Irwin DM, Gong ZY. Molecular evolution of vertebrate goose-type lysozyme genes. J Mol Evol 2003;56:234−42.

[20] Hikima J, Minagawa S, Hirono I, Aoki T. Molecular cloning, expression and evolution of the Japanese flounder goose-type lysozyme gene, and the lytic activity of its recombinant protein. Biochim Biophys Acta 2001;30:35−44.

[21] Savan R, Aman A, Sakai M. Molecular cloning of G-type lysozyme cDNA in common carp (*Cyprinus carpio* L.). Fish Shellfish Immunol 2003;15:263−8.

[22] Yin ZX, He JG, Deng WX, Chan SM. Molecular cloning, expression of orange-spotted grouper goose-type lysozyme cDNA, and lytic activity of its recombinant protein. Dis Aquat Org 2003;55:117−23.

[23] Sun BJ, Wang GL, Xie HX, Gao Q, Nie P. Gene structure of goose-type lysozyme in the mandarin fish *Siniperca chuatsi* with analysis on the lytic activity of its recombinant in *Escherichia coli*. Aquaculture 2006;252:106−13.

[24] Zheng W, Tian C, Chen X. Molecular characterization of goose-type lysozyme homologue of large yellow croaker and its involvement in immune response induced by trivalent bacterial vaccine as an acute-phase protein. Immunol lett 2007;113:107−16.

[25] Kyomuhendo P, Myrnes B, Nilsen IW. A cold-active salmon goose-type lysozyme with high heat tolerance. Cell Mol Life Sci 2007;64:2841−7.

[26] Jiménez-Cantizano RM, Infante C, Martin-Antonio B, Ponce M, Hachero I, Navas JI, et al. Molecular characterization, phylogeny, and expression of c-type and g-type lysozymes in brill (*Scophthalmus rhombus*). Fish Shellfish Immunol 2008;25:57−65.

[27] Larsen AN, Solstad T, Svineng G, Seppola M, Jørgensen TØ. Molecular characterization of a goose-type lysozyme gene in Atlantic cod (*Gadus morhua* L.). Fish Shellfish Immunol 2009;26:122−32.

[28] Ye X, Zhang L, Tian Y, Tan A, Bai J, Li S. Identification and expression analysis of the g-type and c-type lysozymes in grass carp Ctenopharyngodonidellus.DevCompImmunol2010;34:501−9.

[29] Mohanty BR, Sahoo PK. Immune responses and expression profiles of some immune-related genes in Indian major carp, *Labeo rohita* to *Edwardsiella tarda* infection. Fish Shellfish Immunol 2010;28:613−21.

[30] Zhao JM, Song LS, Li CH, Zou HB, Ni DJ, Wang W, et al. Molecular cloning of an invertebrate goose-type lysozyme gene from *Chlamys farreri*, and lytic activity of the recombinant protein. Mol Immunol 2007;44:1198−208.

[31] Park SL. Disease control in Korean aquaculture. Fish Pathol 2009;44:19−23.

[32] Droege M, Hill B. The genome sequencer FLX$^{TM}$ system-longer reads, more applications, straight forward bioinformatics and more complete datasets. J Biotechnol 2008;136:3−10.

[33] Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. Bioinformatics 2007;23:2947−8.

[34] Schultz J, Milpetz F, Bork P, Ponting CP. SMART, a simple modular architecture research tool: identification of signaling domains. Proc Natl Acad Sci U S A 1998;95:5857−5864.

[35] Tamura K, Dudley J, Nei M, Kumar S. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 2007;24:1596−99.

[36] Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods 2001;25:402−8.

[37] Minagawa S, Hikima J, Hirono I, Aoki T, Mori H. Expression of Japanese flounder c-type lysozyme cDNA in insect cells. Dev. Comp. Immunol 2001;25:439−45.

[38] Kawamura S, Ohkuma M, Chijiiwa Y, Kohno K, Nakagawa H, Hirakawa H, et al. Role of disulfide bonds in goose-type lysozyme. FEBS J 2008;275:2818−30.

[39] Kawamura S, Ohno, K, Ohkuma M, Chijiiwa Y, Torikata T. Experimental verification of the crucial roles of $Glu^{73}$ in the catalytic activity and structural stability of goose type lysozyme. J Biochem 2006;140:75−85.

[40] Helland R, Larsen RL, Finstad S, Kyomuhendo P, Larsen AN. Crystal structures of g-type lysozyme from Atlantic cod shed new light on substrate binding and the catalytic mechanism. Cell Mol Life Sci 2009;66:2585−98.

[41] Engel H, Kazemier B, Keck W. Murein-metabolizing enzymes from *Escherichia coli*: sequence analysis and overexpression of the slt gene, which encodes the soluble lytic transglycosylase. J Bacteriol 1991;173:6773−82.

[42] Thunnissen AWH, Isaacs NW, Dijkstra BW. The catalytic domain of a bacterial lytic transglycosylase defines a novel class of lysozyme. Protein Struct Funct Genet 1995;22:245−58.

[43] Nile CJ, Townes CL, Michailidis G, Hirst BH, Hall J. Identification of chicken lysozyme g2 and its expression in the intestine. Cell Molo Life Sci 2004;61:2760−6.

[44] Trent MS, Stead CM, Tran AX, Hankins JV. Diversity of endotoxin and its impact on pathogenesis. J Endotoxin Res 2006;12:2005−23.

[45] Zeng W, Liu G, Ao J, Chen X. Expression analysis of immune-relevant genes in the spleen of large yellow croaker (*Pseudosciaena crocea*) stimulated with polyI:C. Fish Shellfish Immunol 2006;21:414−430.

[46] Saurabh A, Sahoo PK. Lysozyme: an important defence molecule of fish innate immune system. Aquac Res 2008;39:223−39.

# PART 2

**Characterization and expression analysis of the myeloid differentiation factor 88 (MyD88) in rock bream *Oplegnathus fasciatus***

## 2.1 ABSTRACT

Myeloid differentiation factor 88 (MyD88) is a universal adaptor protein able to activate nuclear factor-kappa B (NF-κB) through interactions with interleukin-1 receptor (IL-1R) and the Toll-like receptors (TLRs), with the exception of TLR3. Here, we describe the identification of MyD88 from the rock bream fish *Oplegnathus fasciatus* and its characterization based on GS-FLX$^{TM}$ sequencing. The cDNA of rock bream MyD88 was found to be composed of 1626 bp, with an 867 bp open reading frame that encodes 288 amino acids. The deduced amino acid sequence of MyD88 possessed both a conserved death domain at the amino terminus and a typical Toll-IL-1 receptor (TIR) domain at the carboxyl terminus, similar to that found in other fishes, amphibians, avians, mammals and invertebrates. The mRNA expression pattern of MyD88 in healthy and bacterially-challenged rock bream were examined using quantitative real-time polymerase chain reaction (qRT-PCR). MyD88 transcripts were found to be strongly expressed in blood, gill, liver, spleen, head kidney and kidney, moderately expressed in skin, brain and intestine, and weakly expressed in muscle. Expression levels of MyD88 in blood, spleen and head kidney were dramatically up-regulated upon exposure to LPS and the Gram-negative bacteria Edwardsiella tarda, suggesting that MyD88 plays an important role in rock bream defenses against bacterial infection.

**Keywords:** *Oplegnathus fasciatus*; MyD88; innate immunity; LPS; *Edwardsiella tarda* gene expression

## 2.2 LIST OF TABLES

Page

Table 2.1 Primers used for qRT-PCR expression study ..... 67

Table 2.2 ClustalW2 analysis and comparisons of the deduced amino acid sequence of rock bream *O. fasciatus* MyD88 (GenBank accession number HM035064) with MyD88 from other known species ..... 71

## 2.3 LIST OF FIGURES

Page

Figure 2.1 The cDNA and deduced amino acid sequences of rock bream MyD88 (GenBank accession number HM035064)..... 72

Figure 2.2 The architecture of rock bream MyD88 (288 amino acids) ..... 74

Figure 2.3 Phylogenetic analysis of rock bream MyD88 with MyD88 from 16 other species, including large yellow croaker *P. crocea,* pufferfish *T. rubripes,* flounder *P. olivaceus,* tongue sole *C. semilaevis,* sweetfish *P. altivelis altivelis,* rainbow trout *O. mykiss,* zebrafish *D. rerio,* channel catfish *I. punctatus,* frog *X. laevis,* chicken *G. gallus,* mouse *M. musculus,* cattle *B. Taurus,* human *H. sapiens,* chimpanzee *P. paniscus,* scallop *C. farreri*, and fruitfly *D. melanogaster* ..... 75

Figure 2.4 Relative mRNA expression of rock bream MyD88 in different tissues including blood, gill, liver, spleen, head kidney (H. kidney), kidney, skin, muscle, brain, and intestine..... 76

Figure 2.5 Relative rock bream MyD88 mRNA expression in response to challenge with LPS (a) and *E. tarda* (b) at 3, 6, 12, 24 and 48 h post-injection..... **77**

## 2.4 INTRODUCTION

The fish immune system consists of innate and adaptive immune functions, which act to protect against invading pathogens. As in other vertebrates, the innate immune response of fish not only provides a general defense mechanism against the early pathogenic event but also instructs the subsequent specialized adaptive immune response [1-3]. The innate immune system relies on the pattern recognition receptors (PRRs) to recognize 'non-self' molecular patterns expressed on the surface of pathogens [4-6]. Among the particularly well-characterized PRRs are the Toll-like receptor (TLR) family members. As single membrane-spanning non-catalytic receptors, the TLRs recognize highly structurally conserved pathogen-associated molecular patterns (PAMPs) [7]. TLRs together with the interleukin-1 receptors (IL-1R) form a receptor superfamily, known as TLR/IL-1R, all members of which harbor the so-called TIR (Toll-IL-1 receptor) protein domain [7].

To date, a total of 13 different mammalian TLRs have been identified [7]. The various TLRs have been shown to recognize different kinds of PAMPs, including unmethylated bacterial CpG DNA [8], double- [9] or single-stranded viral RNA [10-12], lipopolysaccharide (LPS) [13], peptidoglycan (PGN) [14], other bacterial lipoproteins [15,16] and bacterial flagellin [17]. Importantly, most of the PAMPs recognized by the TLRs play essential roles in the microbes' or viruses' structural integrity, physiologic function or replication ability; thus, the infectious agent is unable to readily mutate the targeted PAMP as a means of escaping detection by host immune surveillance. For example, the major ligand of TLR4 is lipopolysaccharide (LPS), a major component of the outer membrane of Gram-negative bacteria. LPS expression has been shown to be critical for function and ultimate survival of Gram-negative bacteria as mutations that ablate the enzymes required for LPS synthesis are often lethal to the bacteria [13]. Similarly, double-stranded RNA, the ligand for TLR3, is a key replication intermediate for all

RNA viruses, so that evasion of TLR3 recognition is not easily achieved by these viruses [9].

The signaling pathways activated by these TLRs are broadly classified as either dependent on the adaptor molecule myeloid differentiation factor 88 (MyD88) or MyD88-independent. The former uses MyD88 to initiate early activation of transcription factor NF-κB and production of tumor necrosis factor (TNF) and other inflammatory cytokines [18,19]. The latter propagates its signals through a mechanism involving the TIR domain-containing adaptor-inducing interferon (IFN) to instigate IFN-β production [20].

MyD88 was originally isolated as a myeloid differentiation primary response gene in mice that was induced during terminal differentiation of $M1D^{+}$ myeloid precursor cells in response to IL-6 treatment [21]. MyD88 has a bi-partite structure composed of an amino terminal death domain and carboxyl terminal TIR domain. The carboxyl terminal TIR domain interacts with its cognate domains located in the cytoplasmic tails of activated TLRs or IL-1R. The amino terminal death domain mediates the interaction with the corresponding domain in IL-1R-associated kinase (IRAK) in order to recruit further downstream immune molecules [22]. Interestingly, MyD88 knock-out in mice completely abolished all TLR-mediated cytokine production [22]. Collectively, these properties of MyD88 indicate that it is a pivotal component of the innate immune system and plays a critical role in initiating and activating the immune response, especially in the MyD88-dependent TLR/IL-1R signaling pathway [18,22].

MyD88 has been extensively studied in many species, including human [23], mouse [24], frog [25], chicken [26], zebrafish [27], scallop [28] and fruitfly [29]. In fishes, MyD88 has been identified in zebrafish *Danio rerio* [27,30], Japanese flounder *Paralichthys olivaceus* [31], large yellow croaker *Pseudosciaena crocea* [32] and rainbow trout *Oncorhynchus mykiss* [33].

The rock bream is a tropical fish which resides mainly in the coastal areas of the Pacific and Indian Ocean. Rock bream represents an economically important marine species of South Korea, accounting for the largest yield of a single species in the

national net-cage farming sector. In recent years, the rapid development of the rock bream culture industry has been accompanied by a concomitant increase in the prevalence and virulence of infectious diseases afflicting this species, resulting in significant economic losses [34,35]. Unfortunately, little is known about the genetic and immunological mechanisms of the rock bream fish. This lack of knowledge has proven a major obstacle to the establishment of effective measures towards disease control and development of a more genetically robust population of rock breams.

Here, we describe our investigations that led to the discovery and characterization of a MyD88 cDNA from rock bream *O. fasciatus*. To further our understanding of the defense mechanisms in this species, we studied the temporal expression patterns of rock bream MyD88 in response to exposure to LPS and the Gram-negative bacteria *Edwardsiella tarda*.

## 2.5 MATERIALS AND METHODS

### 2.5.1 Identification of rock bream MyD88

We have established a rock bream cDNA sequence database by using the Roche 454 Genome Sequencer FLX System (GS-FLX$^{TM}$), a next-generation DNA sequencing (NGS) technology [36]. Briefly, total RNA was isolated from pooled tissues (pituitary gland, brain, gill, blood, liver, spleen, head kidney and kidney) of three healthy rock bream fishes using the Tri Reagent$^{TM}$ (Sigma, USA), according to the manufacturer's instructions. Then, the polyadenylated messenger RNA was purified by means of an mRNA isolation kit (FastTrack® 2.0; Invitrogen, USA). First strand cDNA synthesis and normalization were carried out with the Creator$^{TM}$ SMART$^{TM}$ cDNA library construction kit (Clontech, USA) and Trimmer cDNA normalization kit (Evrogen, Russia), respectively. Thereafter, the sequencing of rock bream cDNA was performed on a GS-FLX Titanium instrument (DNA Link, Inc.). By applying the Basic Local Alignment Tool (BLAST) algorithm (http://www.ncbi.nlm.nih.gov/BLAST), we identified a MyD88 gene in the rock bream cDNA sequence database.

### 2.5.2 Sequence alignments and phylogenetic analysis of MyD88

The open reading frame and amino acid sequence of rock bream MyD88 were obtained using DNAssist 2.2 (Version 3.0). The sequences of MyD88 from different species were compared by the BLAST search program. Pairwise sequence alignment (http://www.ebi.ac.uk/Tools/emboss/align) and multiple sequence alignment (http://www.ebi.ac.uk/Tools/clustalw2) were performed using the ClustalW2 program. The phylogenetic relationship of MyD88 was determined using the Neighbor-Joining

method and Molecular Evolutionary Genetics Analysis (MEGA) software version 4 [37]. Prediction of protein domains was carried out by the MotifScan scanning algorithm (http://myhits.isb-sub.ch/cgi-bin/motif_scan).

### 2.5.3 Experimental fish and tissue collection

Rock breams with an average body weight of 30 g were obtained from the Jeju Special Self-Governing Province Ocean and Fisheries Research Institute (Jeju, Republic of Korea). The fish were maintained in a controlled environment at 22-24ºC. All individuals were allowed to acclimate for one week prior to experimentation. Whole blood (1 mL per fish) was collected from the caudal fin using a sterilized syringe, and the sample was immediately centrifuged at 3000 × g for 10 min at 4ºC to separate the blood cells from the plasma. The collected cells were stored in liquid nitrogen until use. Meanwhile, the fish was sacrificed and the gill, liver, spleen, head kidney, kidney, skin, muscle, brain and intestine were excised and immediately snap-frozen in liquid nitrogen and stored at -80ºC until total RNA isolation was performed.

### 2.5.4 Immune challenge

*E. tarda* was obtained from the Department of Aqualife Medicine, Chonnam National University, Korea. The bacteria were incubated at 25°C for 12 h using a brain heart infusion (BHI; Eiken Chemical Co., Japan) broth supplemented with 1% sodium chloride. The cultures were briefly spun down and resuspended in sterile phosphate buffered saline (PBS), and diluted to a desired cell number.

For the mitogen stimulation experiment, each rock bream was administered a single intraperitoneal (ip) injection of 100 μL purified LPS in PBS suspension (1.25

μg/μL, *E. coli* 0127:B8; Sigma). For the live bacterial-challenge experiment, each rock bream was infected with a single ip injection of 100 μL *E. tarda* suspended in PBS (5× $10^3$ CFU/μL). A negative control group was established as non-injected, while a positive control group was injected with an equal volume (100 μL) of PBS. Blood, spleen and head kidney were collected at 3, 6, 12, 24 and 48 h post-injection from LPS- and *E. tarda*-injected rock breams. PBS control samples were isolated at 3 h and 48 h post-injection. All samples were obtained and analyzed from three replicates.

### 2.5.5 Total RNA extraction and cDNA synthesis

Total RNA was extracted by Tri Reagent™ (Sigma) from rock bream blood, gill, liver, spleen, head kidney, kidney, skin, muscle, brain and intestine. RNA concentration was determined at 260 nm in a UV-spectrometer (Bio-Rad, USA) and diluted to 1 μg/μL. Two and one-half μg RNA from selected tissues was used to synthesize cDNA with the SuperScriptTM III First-Strand cDNA Synthesis kit (Invitrogen). Briefly, RNA was incubated with 1 μL of 50 μM oligo$(dT)_{20}$ and 1 μL of 10 mM dNTPs for 5 min at 65ºC. After incubation, 2 μL of 10× cDNA synthesis buffer, 4 μL of 25 mM $MgCl_2$, 2 μL of 0.1M dithiothreitol (DTT), 1 μL of RNaseOUT™ (40U/μL) and 1 μL of SuperScript III reverse transcriptase (200 U/μL) were added and the mixture was incubated for 50 min at 50ºC. The reaction was terminated by adjusting the temperature to 85ºC for 5 min. Then, 1 μL of RNase H was added to each cDNA reaction and incubated for 20 min at 37ºC. Finally, the newly synthesized cDNA was diluted 40-fold (total 800 μL) before storage at -20ºC until needed for further analysis.

### 2.5.6 MyD88 mRNA expression analysis by quantitative real-time polymerase chain reaction (qRT-PCR)

qRT-PCR was used to detect the expression levels of MyD88 in blood, gill, liver, spleen, head kidney, kidney, skin, muscle, brain and intestine tissues, and the temporal expression of MyD88 in the blood, spleen and head kidney. Total RNA was extracted at different time points following immune stimulation, and the first-strand cDNA synthesis was carried out as described in section 2.4. qRT-PCR was carried out using a Thermal Cycler Dice$^{TM}$ (Real Time System TP800; TaKaRa, Japan) in a 20 µL reaction volume containing 4 µL of diluted cDNA (3.125 ng/µL) from each tissue, 10 µL of 2× TaKaRa Ex Taq$^{TM}$ SYBR premix, 0.5 µL of each primer (10 pmol/µL; Table 2.1) and 5 µL $dH_2O$. The qRT-PCR cycle profile included one cycle of 95ºC for 10 s, followed by 35 cycles of 95ºC for 5 s, 58ºC for 10 s and 72ºC for 20 s, and a final single cycle of 95ºC for 15 s, 60ºC for 30 s and 95ºC for 15 s. The baseline was set automatically by the Thermal Cycler Dice$^{TM}$ Real-Time System Software (version 2.00). MyD88 expression was determined by the Livak ($2^{-\Delta\Delta CT}$) method [38]. The same qRT-PCR cycle profile was used for the internal control gene, rock bream β-actin (GenBank accession number FJ975146). All data are presented as relative mRNA expressed as means ± standard deviation (S.D.). To determine statistical significance ($P<0.05$), the data were subjected to one-way analysis of variance (ANOVA) followed by Duncan's Multiple Range test, using the SPSS 11.5 program (USA).

**Table 2.1 Primers used for qRT-PCR expression study**

| Gene | Orientation | Primer sequences (5′-3′) |
|---|---|---|
| MyD88 | Forward | AATCCTGACGCAGGTGGAGAGAAA |
| MyD88 | Reverse | AGCTGTCAACCTCTGGAACCTGAA |
| β-actin | Forward | TCATCACCATCGGCAATGAGAGGT |
| β-actin | Reverse | TGATGCTGTTGTAGGTGGTCTCGT |

## 2.6 RESULTS

### 2.6.1 Molecular characterization and phylogeny of rock bream MyD88

The nucleotide and deduced amino acid sequences of MyD88 from rock bream are shown in Figure 2.1. MyD88 nucleotide sequence was deposited in GenBank under accession number HM035064.

The MyD88 cDNA was composed of 1626 bp which contains an 867 bp open reading frame encoding a putative protein of 288 amino acid residues. The calculated molecular mass was 33 kDa with a theoretical isoelectric point of 4.9. The start (ATG) and stop (TAA) codons were found at nucleotide positions 123-125 and 987-989 from the 5′ end of cDNA, respectively. Sequence analysis revealed the presence of a 122 bp 5′ untranslated region (5′ UTR) and a 637 bp 3′ UTR containing four RNA instability motifs ($^{1102}$ATTTA$^{1106}$, $^{1277}$ATTTA$^{1281}$, $^{1313}$ATTTA$^{1317}$, and $^{1564}$ATTTA$^{1568}$) with a putative polyadenylation consensus signal ($^{1608}$AATAAA$^{1613}$).

Motif scan analysis indicated a death domain was located at positions 24-103 of the amino terminus. The TIR domain, which is important for TLR signaling events, was identified at positions 151-288 of the carboxyl terminus. Amino acid sequence alignment of MyD88 showed the presence of three highly conserved regions (Box 1, Box 2 and Box 3) within the TIR domain that corresponded to those in other teleost fishes (Fig. 2.2).

The percentages of amino acid identity in MyD88 were calculated using the NCBI BLAST search tool and pairwise ClustalW2 program. The deduced amino acid sequence of rock bream MyD88 shared significant homology (more than 58%) with those from other vertebrates, and had the greatest degree of identity (86.2%) with the

MyD88 of large yellow croaker *P. crocea* (Table 2.2). However, the rock bream MyD88 amino acid sequence had less than 30% identity to MyD88 from invertebrates; there was only 14.4% identity to fruitfly MyD88 and 27.2% identity to that of scallop (Table 2).

Phylogenetic analysis was carried out using the Neighbor-Joining method on a multiple sequence alignment of rock bream MyD88 with a range of vertebrate and invertebrate MyD88 (Fig. 2.3). As expected, rock bream MyD88 grouped in the same clade as other piscine MyD88, with high bootstrap values. The deduced amino acid sequences from amphibians, avians and mammals clustered to a corresponding clade, indicating that rock bream MyD88 arose from a common ancestor of vertebrates.

### 2.6.2 Expression of rock bream MyD88 in untreated individuals

The mRNA transcripts from rock bream MyD88 were detected by qRT-PCR in blood, gill, liver, spleen, head kidney, kidney, skin, muscle, brain and intestine of healthy individuals. Rock bream MyD88 mRNA was found to be constitutively expressed in all the tissues examined, although its level of expression was variable. The mRNA expression levels detected in liver, blood, kidney and spleen were higher than in the muscle by 260.8-fold, 179.0-fold, 110.2-fold and 104.4-fold, respectively. A moderately high mRNA level was also detected in gill and head kidney (93.5-fold and 73.1-fold higher than in muscle, respectively). Lower expression levels (<50-fold of that in muscle) were found in other tissues (Fig. 2.4).

### 2.6.3 Expression of rock bream MyD88 in individuals infected with Gram-negative bacteria

The transcriptional response of MyD88 in the blood, spleen and head kidney was analyzed after rock breams were exposed to LPS or *E. tarda* at 3, 6, 12, 24 and 48 h post-infection. The relative transcriptional level of MyD88 was calculated using rock bream β-actin as the internal control gene. Little change in PBS-injected controls was observed in comparison to non-injected controls (1.1-fold increase in blood and spleen, and 1.2-fold increase in head kidney), and no significant difference was detected among PBS-injected controls at 3 and 48 h post-injection. Therefore, samples from non-injected rock breams were used in subsequent comparisons to determine the fold-change of transcriptional up-regulation.

The temporal expression of MyD88 in the blood, spleen and head kidney of rock bream after LPS stimulation is shown in Figure 2.5a. The expression levels of MyD88 differed in these three tissues. The strongest up-regulation in MyD88 mRNA levels (4.3-fold increase) was detected in the blood. The mRNA level peaked in spleen and head kidney at 3 h post-injection, which was 2.0-fold and 2.1-fold higher, respectively, than that in the non-injected control group, and decreased after 3 h post-injection. The peak of the mRNA level in blood persisted up till 12 h post-injection, which was later than that observed in other tissues.

An obvious increase in the MyD88 mRNA levels was also found in the *E. tarda*-challenged tissues, including the blood, spleen and head kidney of rock bream (Fig. 2.5b). The highest induced level (4.2-fold increase) of MyD88 was observed in spleen at 6 h post-injection. In blood and head kidney, the induction of mRNA occurred at 6 h after *E. tarda* injection and reached maximum expression levels (3.4-fold increase and 2.8-fold increase, respectively) within 24 h post-injection.

**Table 2.2 ClustalW2 analysis and comparisons of the deduced amino acid sequence of rock bream *O. fasciatus* MyD88 (GenBank accession number HM035064) with MyD88 from other known species**

| Species | Accession number | Amino acids | Identity % |
|---|---|---|---|
| *Pseudosciaena crocea* | ACL14361 | 287 | 86.2 |
| *Takifugu rubripes* | NP_001106666 | 288 | 78.1 |
| *Plecoglossus altivelis altivelis* | BAI68385 | 287 | 77.9 |
| *Paralichthys olivaceus* | BAE75959 | 285 | 77.1 |
| *Oncorhynchus mykiss* | NP_001117893 | 282 | 76.5 |
| *Cynoglossus semilaevis* | ACU31062 | 285 | 76.0 |
| *Ictalurus punctatus* | ACD81929 | 279 | 71.4 |
| *Danio rerio* | AAZ16494 | 284 | 70.2 |
| *Mus musculus* | AAH58787 | 296 | 61.3 |
| *Pan paniscus* | BAG55249 | 296 | 60.6 |
| *Gallus gallus* | ABQ17966 | 299 | 60.5 |
| *Xenopus laevis* | NP_001081001 | 283 | 59.8 |
| *Homo sapiens* | AAB49967 | 296 | 59.6 |
| *Bos taurus* | NP_001014404 | 296 | 58.9 |
| *Chlamys farreri* | ABB76627 | 367 | 27.2 |
| *Drosophila melanogaster* | NP_610479 | 537 | 14.4 |

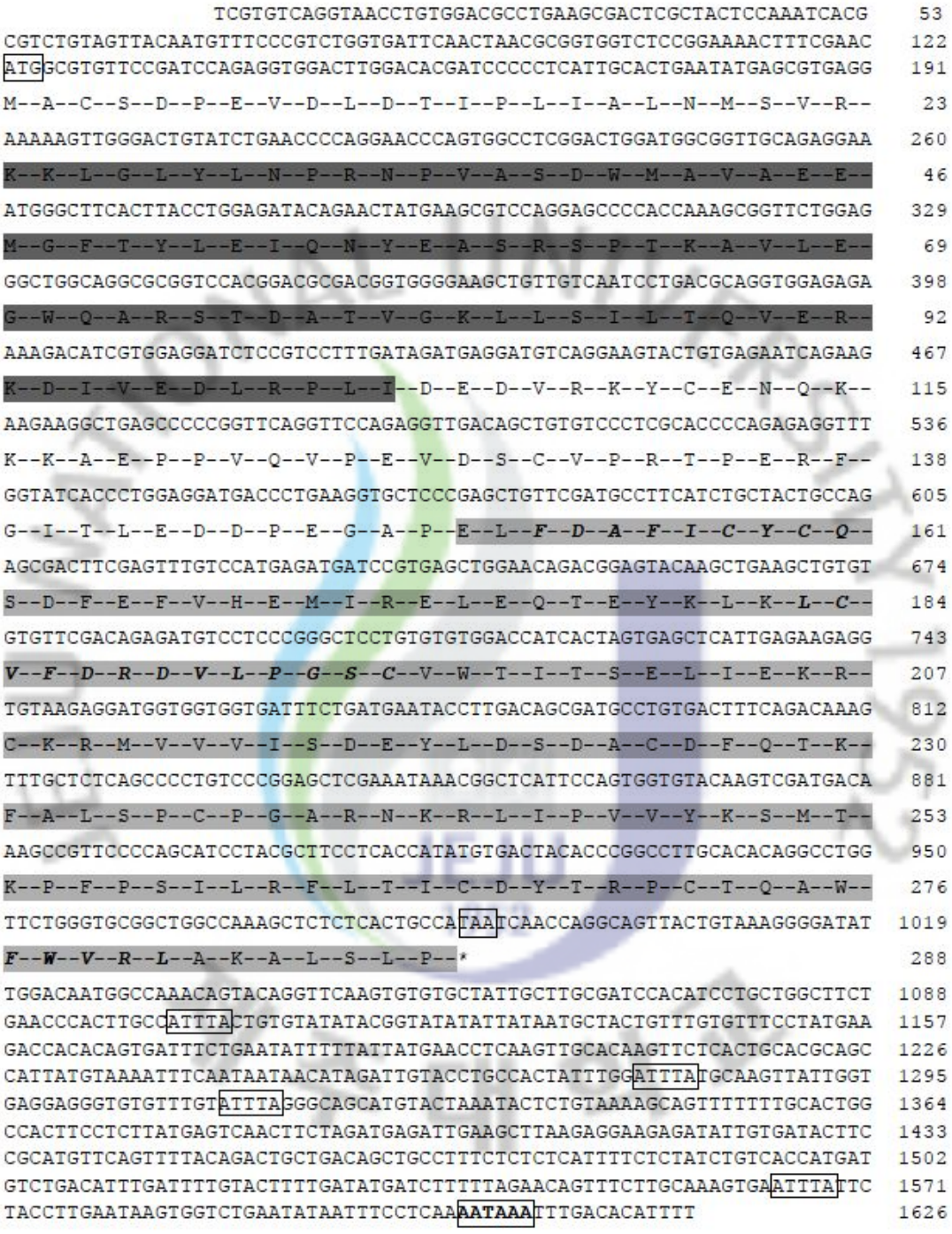



**Figure 2.1 The cDNA and deduced amino acid sequences of rock bream MyD88**

**(GenBank accession number HM035064).** The nucleotides and amino acids are numbered along the right margin. The boxed letters indicate the start codon (ATG), the stop codon (TAA) and the instability motif (ATTTA). The boxed bold letters indicate a putative polyadenylation signal (AATAAA). The death domain regions extend from position 24-103 (dark gray underlay), and the TIR domain extends from position 151-288 (light gray underlay), in which conserved sequences are denoted by bold italicized letters.

| | | Box1 | Box2 | Box3 |
|---|---|---|---|---|
| *O. fasciatus* | 153 | FDAFICYCQ | LCVFDRDVLPGSC | FWVRL |
| *P. crocea* | 152 | FDAFICYCQ | LCVFDRDVLPGSC | FWVRL |
| *T. rubripes* | 153 | FDAFICYCQ | LCVFDRDVLPGSC | FWKRL |
| *P. altivelis_altivelis* | 152 | FDAFICYCQ | LCVFDRDVLPGSC | FWIRL |
| *P. olivaceus* | 150 | FDAFICYCQ | LCVFDRDVLPGSC | FWTRL |
| *I. punctatus* | 144 | FDAFICYCQ | LCVFDRDVLPGTC | FWVRL |
| *O. mykiss* | 147 | FDAFICYCQ | LCVFDRDALPGSC | FWVRL |
| *C. semilaevis* | 150 | FDAFICYCH | LCVFDRDVLPGSC | FWDRL |
| *D. rerio* | 149 | FDAFICYCQ | LCVFDRDVLPGTC | FWTRL |

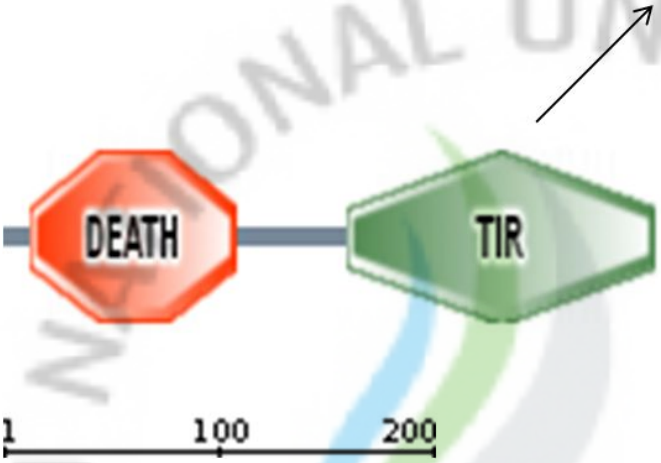



**Figure 2.2 The architecture of rock bream MyD88 (288 amino acids). The TIR domain from rock bream is compared with other fish MyD88 sequences.** The three conserved boxes (Box 1, Box 2 and Box 3) are shown.

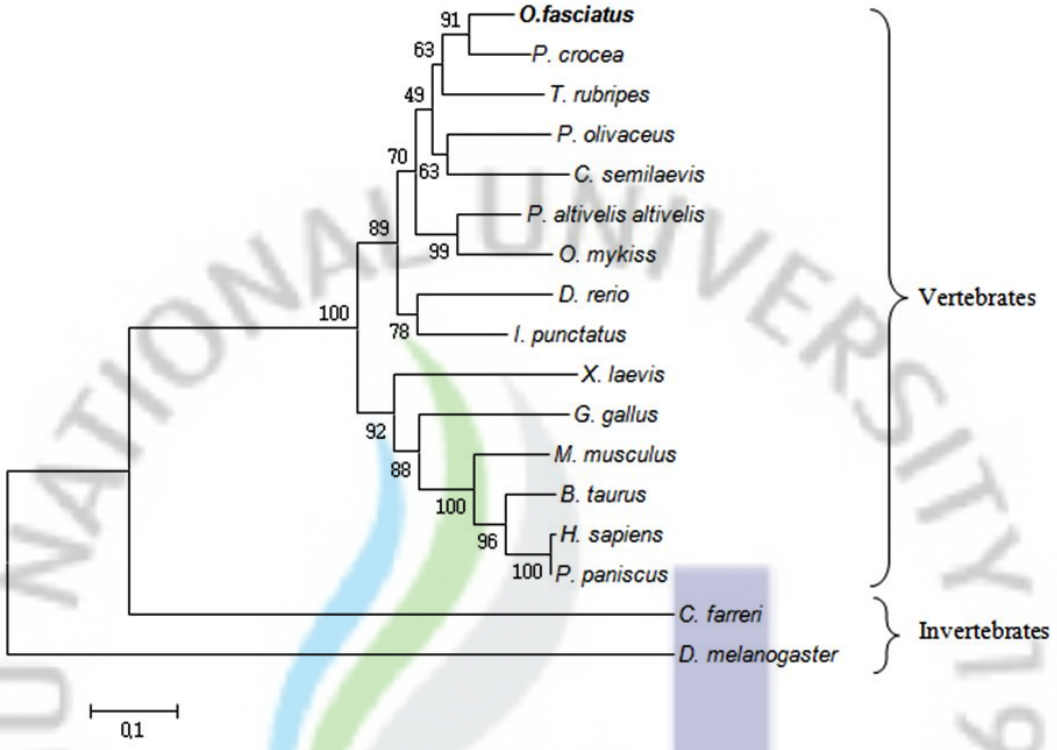



**Figure 2.3 Phylogenetic analysis of rock bream MyD88 with MyD88 from 16 other species, including large yellow croaker *P. crocea*, pufferfish *T. rubripes*, flounder *P. olivaceus*, tongue sole *C. semilaevis*, sweetfish *P. altivelis altivelis*, rainbow trout *O. mykiss*, zebrafish *D. rerio*, channel catfish *I. punctatus*, frog *X. laevis*, chicken *G. gallus*, mouse *M. musculus*, cattle *B. Taurus*, human *H. sapiens*, chimpanzee *P. paniscus*, scallop *C. farreri*, and fruitfly *D. melanogaster*.** The tree is based on an alignment corresponding to full-length amino acid sequences, generated by ClustalW and *MEGA* version 4. The numbers at each of the nodes represent the bootstrap percentages of 1000 replicates. The scale bar corresponds to 0.1.

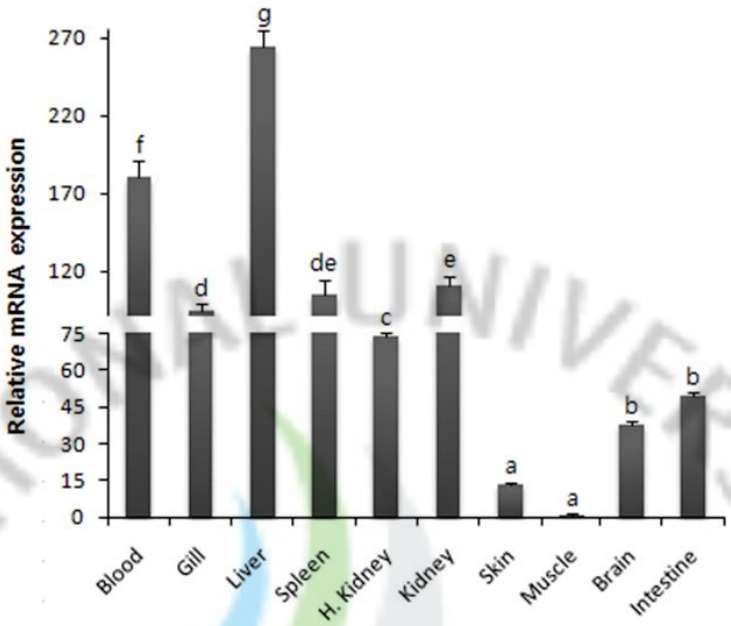



**Figure 2.4 Relative mRNA expression of rock bream MyD88 in different tissues including blood, gill, liver, spleen, head kidney (H. kidney), kidney, skin, muscle, brain, and intestine..** Expression values are normalized to those of rock bream β-actin (internal control). The bars indicate the mean fold change (mean ± S.D., $n$ = 3) from the calibrator group (muscle). Means with different letters are significantly different ($P$<0.05).

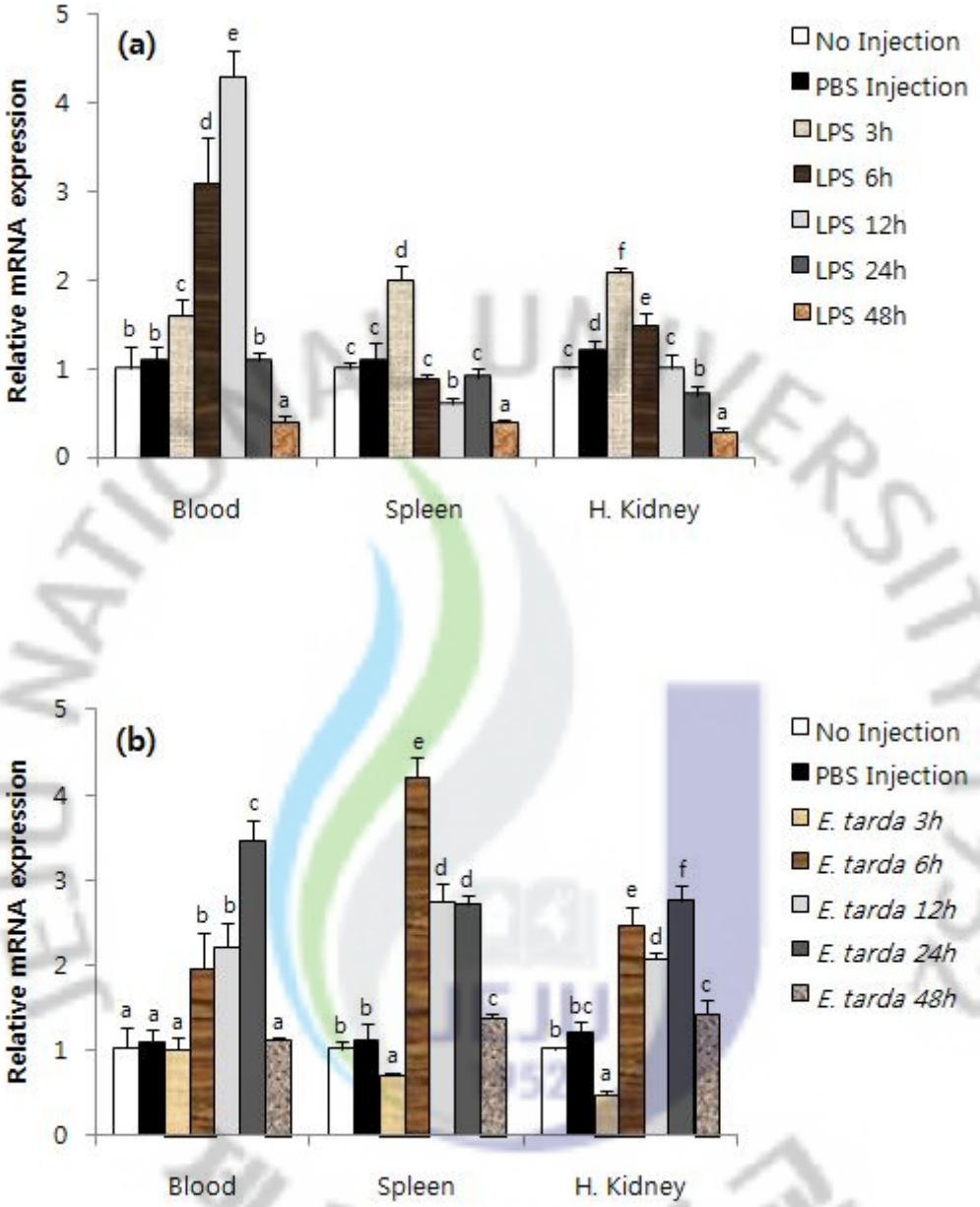



**Figure 2.5 Relative rock bream MyD88 mRNA expression in response to challenge with LPS (a) and *E. tarda* (b) at 3, 6, 12, 24 and 48 h post-injection.** Expression values are normalized to those of rock bream β-actin. The bars indicate the mean fold change (mean ± S.D., $n$ = 3) from non-injected controls. Means with different letters are significantly different ($P$<0.05).

## 2.7 DISCUSSION

Toll-like receptors play an essential role in the detection of invading pathogens within the body and the initiation of the subsequent innate immune response. TLRs are a type of pattern recognition receptor and recognize structurally conserved pathogen-associated molecular patterns that are broadly expressed by pathogens, but are distinguishable from host 'self' molecules [7]. When activated by PAMPs, almost all TLRs (except TLR3) recruit the adaptor protein MyD88 in order to propagate their signal. This adaptor molecule activates certain protein kinases, such as IRAKs, that then act to amplify the signal and ultimately lead to the induction or suppression of downstream genes that orchestrate the immuno-inflammatory response [7].

In this study, we have identified and characterized the TLR/IL-1R adaptor protein MyD88 from rock bream *O. fasciatus*. The MyD88 cDNA was 1626 bp in length and included an open reading frame of 867 bp encoding a putative protein of 288 amino acids, which was consistent with the sequence analysis of the MyD88 from pufferfish *T. rubripes* (GenBank accession number NP_001106666). In its 3′ UTR, the rock bream MyD88 cDNA contained four mRNA instability motifs (ATTTA), which are believed to be responsible for destabilizing mRNA by regulating degradation and suppressing translation mechanisms [39,40]. This instability motif has not only been found in *P. crocea* MyD88 [32] but also in a number of inflammatory mediators in fish, including various cytokines [41,42].

Based on its predicted structural features, the amino acid sequence of the rock bream MyD88 protein was remarkably conserved. First, the death domain was found to be located at positions 24-103 of the amino terminus. To date, more than 300

death domain-containing proteins have been reported [43,44] and, in most of them, the death domains are typically located at the extreme carboxyl terminal. The death domain originally was characterized for its involvement in signaling processes that ultimately led to apoptosis [45]. However, the death domain within the TLR pathway component MyD88 was located at the amino terminus and appears to be used for mediating downstream interactions with the IRAK family [18]. Therefore, the conserved sequences and location of the death domain in rock bream MyD88 suggest that it should have a typical function in TLR signaling transduction. Second, the TIR domain, with three conserved boxes (Box 1, Box 2 and Box 3), was found to comprise the carboxyl terminal half of the protein (amino acid positions 151-288) and exhibited significantly more sequence conservation than did the death domain (data not shown). Box 1, which has been proposed as essential for the interaction of MyD88 with TLRs and IL-1R [46], was found to be highly conserved in rock bream MyD88. This result suggested that MyD88 in rock bream played a similar role to the conserved MyD88 from other species. Third, a 47 amino acid long interdomain was found in MyD88 at positions 104-150. Interestingly, functional relevance of an interdomain was proven after the discovery of a mouse MyD88 splice variant lacking the interdomain. This truncated form of MyD88 (called $MyD88_s$) has been shown to be stimulated by LPS and can inhibit LPS-induced NF-κB activation through inhibition of IRAK activity [47,48]. Thus, $MyD88_s$ may negatively regulate the inflammatory responses triggered by LPS. However, the function of this interdomain in fish species remains to be elucidated.

The mRNA expression of the MyD88 in healthy rock bream was found to be ubiquitous, but at various levels of robustness in different tissues. High expression was observed in immune-relevant tissues, including the liver, blood, kidney, spleen, gill and head kidney. This is similar to the observed expression of MyD88 in other

species, including mammalian [23,24], amphibian [25], avian [26], fishes [27], and invertebrates [28,29], indicating a role of MyD88 in innate immune defense.

The temporal expression patterns of MyD88 mRNA have been previously investigated in response to LPS, the peptidoglycan (PGN) component of Gram-positive bacteria, polyinosinic-polycytidylic acid (poly I:C) and whole cell bacterial stimulation [31,32]. In the study presented herein, an immediate increase in the expression of rock bream MyD88 was observed in blood, spleen and head kidney after LPS stimulation (Fig. 5a). The observed pattern in this study generally agreed with published data on MyD88 response after LPS treatment of peripheral blood leukocytes (PBLs) from *P. olivaceus* [31]. Takano *et al.* [31] suggested that the prompt increase in MyD88 expression indicated that the MyD88-dependent signaling pathway of TLR4 may have been activated. Mammalian TLR4 is known to be involved in the response to LPS from Gram-negative bacteria, and MyD88 has been demonstrated to act as an adaptor molecule in the related signaling pathway [7]. However, the TLR4 gene has only been found in the teleost species zebrafish (Cypriniformes) [30,49]. Sepulcre *et al.* [50] reported that the mechanism of LPS recognition in fish may be distinct from that of mammals; it was suggested that LPS may trigger a TLR4- and MyD88-independent signaling pathway in fish. These authors also showed that the zebrafish TLR4 was able to negatively regulate the MyD88-dependent signaling pathway. On the other hand, Roach and colleagues [51] suggested that *Fugu* TLR23 may participate in LPS recognition. Likewise, Iliev *et al.* [52] speculated that alternative signaling receptors (not TLR4), such as Beta-2 integrins, may play a primary role in the activation of piscine leukocytes by LPS. These reports raise the possibility that different receptors might be involved in the activation of signaling pathway by LPS in fish.

In our study, we observed that infected rock bream were able to mount a

significant response to *E. tarda*, and presented an increased expression of MyD88 in blood, spleen and head kidney (Fig. 5b). Interestingly, in contrast to the rapid response to LPS, in the early stages after live *E. tarda* infection, the expression of MyD88 decreased in the spleen and head kidney; a null induction was observed in blood which was followed by a gradual increase. This study represents the first *in vivo* investigation of MyD88 expression in rock bream exposed to live pathogenic bacteria; the immunologic response pattern observed basically agreed with previous data from fish exposed to formalin-inactivated Gram-negative bacteria *Vibrio parahaemolyticus* [32].

In conclusion, we identified the cDNA of MyD88 from rock bream *O. fasciatus,* and characterized its expression features in healthy and immune-challenged individuals. MyD88 was found to be highly expressed in immune-relevant tissues and significantly up-regulated by immune-challenge. These findings indicated that MyD88 of rock bream can act as a highly inducible molecule after bacterial challenge and might be involved in the first line of defense against bacterial pathogens.

Although further research is needed to understand the dynamic role of MyD88 with respect to other immunoproteins (signaling receptor or costimulatory proteins) in the rock bream innate immune system, we believe that investigations similar to those presented in this paper will provide the critical information necessary to solve the problems caused by diseases currently affecting the culture of this important fish species.

## 2.8 REFERENCES


[1] Magnadóttir B. Innate immunity of fish (overview). Fish Shellfish Immunol 2006;20:137-51.

[2] Fearon DT. Seeking wisdom in innate immunity. Nature 1977;388:323-4.

[3] Medzhitov R, Janeway Jr CA. Decoding the patterns of self and nonself by the innate immune system. Science 2002;296:298-300.

[4] Armant MA, Fenton MJ. Toll-like receptors: a family of pattern-recognition receptors in mammals. Genome Biol 2002;3:3011.1-6.

[5] Beutler B. The Toll-like receptors: analysis by forward genetic methods. Immunogenetics 2005;57:385-92.

[6] Bilak H, Tauszig-Delamasure S, Imler JL. Toll and Toll-like receptors in Drosophila.BiochemSocTrans2003;31:648-51.

[7] Krishnan J, Selvarajoo K, Tsuchiya M, Lee G, Choi S. Toll-like receptor signal transduction. Exp Mol Med 2007;39:421-38.

[8] Hemmi H, Takeuchi O, Kawai T, Kaisho T, Sato S, Sanjo H, et al. A Toll-like receptor recognizes bacterial DNA. Nature 2000;408:740-5.

[9] Alexopoulou L, Holt AC, Medzhitov R, Flavell RA. Recognition of double stranded RNA and activation of NF-kappaB by Toll-like receptor 3. Nature 2001;413:732-8.

[10] Lund JM, Alexopoulou L, Sato A, Karow M, Adams NC, Gale NW, et al. Recognition of single stranded RNA viruses by Toll-like receptor 7. Proc Natl Acad Sci USA 2004;101:5598-603.

[11] Diebold SS, Kaisho T, Hemmi H, Akira S, Reis e Sousa C. Innate antiviral responses by means of TLR7-mediated recognition of single-stranded RNA. Science 2004;303:1529-31.

[12] Heil F, Hemmi H, Hochrein H, Ampenberger F, Kirschning C, Akira S, et al. Species-specific recognition of single-stranded RNA via Toll-like receptor 7 and 8. Science 2004;303:1526-9.

[13] Takeuchi O, Hoshino K, Kawai T, Sanjo H, Takada H, Ogawa T, et al.

Differential roles of TLR2 and TLR4 in recognition of gram-negative and gram-positive bacterial cell wall components. Immunity 1999;11:443-51.

[14] Takada H, Uehara A. Enhancement of TLR-mediated innate immune responses by peptidoglycans through NOD signaling. Curr Pharm Des 2006;12(32):4163-72.

[15] Takeuchi O, Kaufmann A, Grote K, Kawai T, Hoshino K, Morr M, et al. Cutting edge: preferentially the R-stereoisomer of the myoplasma lipopeptide macrophage-activating lipopeptide-2 activates immune cells through a toll-like receptor 2- and MyD88-dependent signaling pathway. J Immunol 2000;164:554-7.

[16] Ozinsky A, Underhill DM, Fontenot JD, Hajjar AM, Smith KD, Wilson CB, et al. The repertoire for pattern recognition of pathogens by the innate immune system is defined by cooperation between toll-like receptors. Proc Natl Acad Sci USA 2000;97:13766-71.

[17] Hayashi F, Smith KD, Ozinsky A, Hawn TR, Yi EC Goodlett DR, Eng JK, et al. The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5. Nature 2001;410:1099-103.

[18] Medzhitov R, Preston-Hurlburt P, Kopp E, Stadlen A, Chen C, Ghosh C, et al. MyD88 is an adaptor protein in the hToll/IL-1 receptor family signaling pathways. Mol Cell 1998;2:253-8.

[19] Janeway Jr CA, Medzhitov R. Innate immune recognition. Annu Rev Immunol 2002;20:197-216.

[20] Kawai T, Takeuchi O, Fujita T, Inoue J, Muhlaradt PF, Sato S, et al. Lipopolysaccharide stimulates the MyD88-independent pathway and results in activation of IFN-regulatory factor 3 and the expression of a subset of lipopolysaccharide-inducible genes. J Immunol 2001;167:5887-94.

[21] Lord KA, Hoffman-Liebermann B, Liebermann DA. Nucleotide sequence and expression of a cDNA encoding MyD88, a novel myeloid differentiation primary response gene induced by IL6. Oncogene 1990;5 (7): 1095–7.

[22] Wesche H, Henzel WJ, Shillinglaw W, Li S, Cao Z. MyD88: an adapter that recruits IRAK to the IL-1 receptor complex. Immunity 1997; 7: 837-47.

[23] Bonnert TP, Garka KE, Parnet P, Sonoda G, Testa JR, Sims JE. The cloning and characterization of human MyD88: a member of an IL-1 receptor related family. FEBS Lett 1997;402: 81-4.

[24] Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD, Collins FS,

et al. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. Proc Natl Acad Sci USA 2002;99:16899-903.

[25] Prothmann C, Armstrong NJ, Rupp RA. The Toll/IL-1 receptor binding protein MyD88 is required for Xenopusaxisformation.Mech Dev 2000;97: 85-92.

[26] Wheaton S, Lambourne MD, Sarson AJ, Brisbin JT, Mayameei A, Sharif S. Molecular cloning and expression analysis of chicken MyD88 and TRIF genes. DNA Seq 2007;18: 480-6.

[27] van der Sar AM, Stockhammer OW, Van der Laan C, Spaink HP, Bitter W, Meijer AH. MyD88 innate immune function in a zebrafish embryo infection model. Infect Immun 2006;74: 2436-41.

[28] Qiu L, Song L, Yu Y, Su W, Ni D, Zhang Q. Identification and characterization of a myeloid differentiation factor 88 (MyD88) cDNA from Zhikong scallop Chlamysfarreri.Fish Shellfish Immunol 2007;23:614-23.

[29] Tauszig-Delamasure S, Bilak H, Capovilla M, Hoffmann JA, Imler JL. DrosophilaMyD88isrequiredfortheresponsetofungalandGram-positivebacterialinfection.Nat Immunol 2001;3: 91-7.

[30] Meijer AH, Krens SFG, Rodriguez IAM, He S, Bitter W, Snaar-Jagalska BE, et al. Expression analysis of the Toll-like receptor and TIR domain adaptor families of zebrafish. Mol Immunol 2004;40: 773-83.

[31] Takano T, Kondo H, Hirono I, Saito-Taki T, Endo M, Aoki T. Identification and characterization of a myeloid differentiation factor 88 (MyD88) cDNA and gene in Japanese flounder Paralichthysolivaceus.Develop Comp Immunol 2006;30: 807-16.

[32] Yao CL, Kong P, Wang ZY, Ji PF, Liu XD, Cai MY, et al. Molecular cloning and expression of MyD88 in large croaker Pseudosciaenacrocea.Fish Shellfish Immunol 2009;26:249-55.

[33] Rebl A, Goldammer T, Fischer U, Köllner B, Seyfert HM. Characterization of two key molecules of teleost innate immunity from rainbow trout (Oncorhynchusmykiss):MyD88andSAA.Vet Immunol Immunopathol 2009;131:122-6.

[34] Zenke K, Kim KH. Functional characterization of the RNase III gene of rock bream iridovirus. Arch Virol 2008;153:1651-56.

[35] Park SL. Disease control in Korean aquaculture. Fish Pathol 2009;44:19-23.

[36] Droege M, Hill B. The genome sequencer FLX$^{TM}$ system-longer reads, more

applications, straight forward bioinformatics and more complete datasets. J Biotechnol 2008;136:3-10.

[37] Tamura K, Dudley J, Nei M, Kumar S. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 2007;24:1596-99.

[38] Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods 2001;25:402-8.

[39] Shaw G, Kamen R. A conserved AU sequence from the 3′ untranslated region of GM-CSF mRNA mediates selective mRNA degradation. Cell 1986;46:659-67.

[40] Han J, Brown T, Beutler B. Endotoxin-responsive sequences control cachectin/tumor necrosis factor biosynthesis at the translational level. J Exp Med 1990;171:465-75.

[41] Lee EY, Park HH, Kim YT, Choi TJ. Cloning and sequence analysis of the interleukin-8 gene from flounder *(Paralichthys olivaceus*). Gene 2001;274:237-43.

[42] Laing KJ, Zou JJ, Wang T, Bols N, Hirono I, Aoki T, et al. Identification and analysis of an interleukin 8-like molecule in rainbow trout *Oncorhynchus mykiss*. Dev Comp Immunol 2002;26:433-44.

[43] Weber CH, Vincenz C. The death domain superfamily: a tale of two interfaces? Trends Biochem Sci 2001;26:475-81.

[44] Hofmann K, Tschopp J. The death domain motif found in FAS (APO-1) and TNF receptor is present in proteins involved in apoptosis and axonal guidance. FEBS Lett 1995;371:321-3.

[45] Itoh N, Nagata S. A novel protein domain required for apoptosis. Mutational analysis of human Fas antigen. J Biol Chem 1993;268:10932-7.

[46] Xu Y, Tao X, Shen B, Horn T, Medzhitov R, Manley JL, et al. Structural basis for signal transduction by the Toll/interleukin-1 receptor domains. Nature 2000;408:111-5.

[47] Janssens S, Burns K, Tschopp J, Beyaert R. Regulation of interleukin-1 and lipopolysaccharide-induced NF-κB activation by alternative splicing of MyD88. Curr. Biol 2002;12:467-71.

[48] Burns K, Janssens S, Brissoni B, Olivos N, Beyaert R, Tschopp J. Inhibition of interleukin 1 receptor/toll-like receptor signaling through the alternatively spliced, short form of MyD88 is due to its failure to recruit IRAK-4. J Exp Med

2003;197:263-8.

[49] Jault C, Pichon L, Chluba J. Toll-like receptor gene family and TIR-domain adapters in *Danio rerio*. Mol Immunol 2004;40:759-71.

[50] Sepulcre MP, Alcaraz-Pérez F, López Muñoz A, Roca FJ, Meseguer J, Cayuela ML, et al. Evolution of lipopolysaccaride (LPS) recognition and signaling: Fish TLR4 does not recognize LPS and negatively regulates NF-kB activation. J Immunol 2009;15:1836-45.

[51] Roach JC, Glusman G, Rowen L, Kaur A, Purcell MK, Smith KD et al. The evolution of vertebrate Toll-like receptors. Proc Natl Acad Sci USA 2005;102:9577-82.

[52] Iliev DB, Roach JC, Mackenzie S, Planas JV, Goeta FW. Endotoxin recognition: In fish or not in fish? FEBS Lett 2005;579:6519-28.

# PART 3

**Molecular analysis and transcriptional responses of interferon (IFN) regulatory factor-1 and IFN gamma inducible lysosomal thiole reductase (GILT) from rock bream (*Oplegnathus fasciatus*)**

## 3.1 ABSTRACT

Activation of the interferon (IFN) system and its down-stream IFN-stimulated genes (ISGs) play important roles in innate and adaptive immune responses to pathogens. In this study, we describe the identification and characterization of IFN regulatory transcription factor 1 (IRF-1) and IFN-ɣ inducible lysosomal thiole reducatase (GILT) cDNA sequences from rock bream *Oplegnathus fasciatus*. These two cDNA sequences, named as RbIRF-1 and RbGILT, were discovered by signature pyrosequencing (GS-FLX 454) of a multi-tissue cDNA library of rock bream. RbIRF-1 showed significant evolutionary conservation of its N-terminal 113 amino acids, which encompassed a DNA binding domain (DBD) containing five conserved tryptophan repeats. Meanwhile, the RbGILT cDNA had the characteristic GILT signature sequence composed of a functional domain ($^{97}$CQHGEQECLGNMIETC$^{112}$), active site $^{74}$C-XX-C$^{77}$ motif and seven putative disulfide bonds. Therefore, the newly identified rock bream IRF-1 and GILT proteins are similar to those in other fish and mammals as revealed by molecular characterization and phylogenetic analysis.

Real-time RT-PCR analysis revealed that expression of RbIRF-1 and RbGILT transcripts was constitutive in eleven tissues selected from un-challenged rock bream, with the highest levels observed blood and gills. Immune challenge with synthetic polyinosinic:polycytidylic acid (polyI:C) up-regulated the RbIRF-1 and RbGILT mRNA in blood, gills, spleen and head kidney; however, the magnitude of RbGILT up-regulation was lower than that of RbIRF-1. Only a moderate up-regulation (compared to poly I:C) of RbIRF-1 and RbGILT mRNA was observed in gills and head kidney at the early stage of rock bream iridovirus (RbIV) challenge. In blood, both transcripts were found to be continuously up-regulated throughout the 48 h

observation. The transcriptional up-regulation in response to poly I:C and RbIV, strongly suggest that RbIRF-1 and RbGILT are related to IFN signaling and may indicate essential roles in subsequent adaptive immunity in rock bream.

**Keywords:** Rock bream; *Oplegnathus fasciatus*; IRF-1; GILT; poly I:C; RbIV

## 3.2 LIST OF TABLES

## 3.3 LIST OF FIGURES

Page

Figure 3.1 The nucleotide and deduced amino acid sequences of RbIRF-1 cDNA. The start (ATG) and stop (TAA) codons are underlined ..... 103

Figure 3.2 The nucleotide and deduced amino acid sequences of RbGILT cDNA. The start (ATG) and stop (TGA) codons are underlined ..... 104

Figure 3.3 Multiple sequence alignment of RbIRF-1. Analysis was performed by ClustalW using amino acid sequences of known IRF-1s from different phyla, including *S. chuatsi* (AAV650412), *E. coioides* (ACF95885), *X. tropicalis* (NM_001006694), *G. gallus* (L339766), *R. norvegicus* (M34253) and *H. sapiens* (L05072). The DNA binding domain is marked by a dotted line above the aligned sequences with bold type (RbIRF-1) ..... 105

Figure 3.4 Multiple sequence alignment of RbGILT ..... 107

Figure 3.5 Phylogenetic analysis of RbIRF-1 with other IRF family members .. 108

Figure 3.6 Phylogenetic analysis of RbGILT with known GILT family proteins 109

Figure 3.7 Tissue distribution of RbIRF-1 and RbGILT ..... 114

Figure 3.8 Transcriptional regulation of RbIRF-1 and RbGILT in blood, gills, spleen and head kidney following poly I:C challenge ..... 115

Figure 3.9 Transcriptional regulation of RbIRF-1 and RbGILT in blood, gills, spleen and head kidney following RBIV challenge..... 116

## 3.4 INTRODUCTION

Proper activation of the innate immune system is essential for regulating subsequent host responses associated with adaptive immunity against pathogens [1]. The interferons (IFNs) represent the major group of cytokines involved in cellular antiviral defense, as well as having various other pleiotropic functions, thereby supporting host resistant phenotypes against particular pathogenic agents [2]. In mammals, expression of IFN and certain IFN-stimulated genes (ISGs) is mediated by a family of transcription factors known as the IFN regulatory factors (IRFs) [3]. To date, ten IRF proteins (IRF1-IRF10) have been identified in higher vertebrates, although IRF10 appears to be absent from the genomes of human and mice [4]. New IRF family member, IRF-11 was identified from certain fish species such as *Danio rerio and Fugu rubripes* [5]. The unique characteristic feature of these IRFs is that they possess DNA binding domain (DBD) with tryptophan amino acids cluster [5,6]. IRF-1 was originally identified as a transcriptional activator of the human IFN-β gene. Futhermore the DNA binding domain (DBD) of IRF-1 was shown to specifically bind to *cis*-acting IRF elements present in the promoter sequence of IFN-β [6-7]. Dynamic roles have since been defined for IRF-1 in IFN signaling, as well as in T cell mediated immune response [8], antiviral activity [9], inflammatory response [10], regulation of cell proliferation and differentiation [11], tumor suppression [12] and apoptosis [13]. Also, upon stimulation of IRF-1 by certain viruses, the protein functions as a transcriptional activator of various ISGs which are involved in diverse immune reactions, such as macrophage activation, antigen presentation [14].

Several IFNs and ISGs have been characterized in teleosts, but relatively few IRF-1 genes have been cloned from the species examined to date, including *Salmon salar* [4], *Paralichthys olivaceus* [15], *Fugu rubripes* [16], *Oncorhynchus mykiss* [17], *Scophthalmus maximus*, *Sparus aurata* [18], *Siniperca chuatsi* [19], *Channa*

*argus* [20], *Epinephelus coioides* [21] and *Carassius auratus* [22]. A recombinant plasmid expressing *Paralichthys olivaceus* IRF-1 was able to induce the anti-viral Mx gene transcription upon vaccination [23]. More recently, three IFN-inducible Mx isoforms [24] and the double stranded RNA-dependent protein kinase (PKR) [25] were cloned from rock bream. The study presented herein represents the first successful attempt to identify IRF-1 in rock bream.

The IFN- inducible lysosomal thiol reducatase (GILT) protein is known to play a central role in MHC class II antigen processing and presentation pathways by acting as a reducer of disulfide bonds in exogenous antigens, a process considered as essential for immune responses [26]. It has been demonstrated that GILT also mediates the neutralization of extracellular pathogens and clearance of cell debris resulting from infection [27]. Mammalian GILT is mainly controlled by IFN-ɣ via the JAK-STAT signaling pathway [28]. Only a few numbers of GILT genes have been cloned from fish, including *Danio rerio* [29], *Pseudosciaena crocea* [30], *Epinephelus coioides* [31], catfish (DQ353791) and pufferfish (CR697192). The immune response properties of GILT in fish have been investigated by comparative transcriptional analysis of healthy and immune-challenged individuals, in particular examining the effects of bacteria [30] and LPS [31] exposure. Results of these studies have suggested that induction of GILT could be a result of increased IFN- ɣ produced by stimulated lymphocytes and macrophages. Among the three main classes of IFNs type I (α/β), II (ɣ) and III (λ)), type II is represented by a single member, IFN-ɣ, while the other two types are multigene families [32]. Interestingly, the *C. argus* IRF-1 promoter contains an IFN-ɣ activation site (GAS sequence [21], indicating the potential for interrelated transcriptional regulation. Since IRF-1 is an activator of different sets of genes related to the IFN system, it would be valuable to elucidate the immune related function(s) associated with both IRF-1 and GILT.

Insights into the regulation of IRF-1 and its transcriptional activation of downstream target genes will undoubtedly contribute to our overall understanding of the IFN regulatory mechanisms in fish. To date, however, very little know is known

about such IFN-related molecules in the rock bream, a commercially important marine fish which is frequently infected by rock bream iridovirus (RbIV) [33]. In order to begin tackling the economically devastating effects of RbIV, characterization of IRF-1 and GILT molecules will help to uncover the complexities of the innate and adaptive immune responses in rock bream.

To this end, we sought to identify and molecularly characterize IRF-1 and GILT cDNA sequences from rock bream. We found that the transcriptional regulation of rock bream IRF-1 and GILT are associated with immune responses to poly I:C and RbIV.

## 3.5 MATERIALS AND METHODS

### 3.5.1 Fish and tissue isolation

Healthy rock bream *Oplegnathus fasciatus* fish (mean weight of 30 g) were obtained from the Ocean and Fisheries Research Institute (Jeju Special Self-Governing Province, Republic of Korea). Fish were maintained in a seawater tank (400 L) tank with a salinity level of 34 ‰ at 24±1 ℃. Fish were acclimatized for one week prior to experimentation and maintained on a commercial pellet diet. For tissue sampling, fish were sacrificed and the gills, spleen, heart, head kidney, kidney, liver, intestine, muscle, skin and brain were excised from three healthy (un-challenged) fish. Blood samples (approximately 1 mL/fish) were taken from the caudal vein by using a syringe fitted with a 22 gauge needle containing 100 µL heparin to prevent coagulation. Blood samples were centrifuged 3000 x g at 4 $^{o}$C for 10 min to separate the blood cells from the plasma and serum. All tissues and blood cells were snap-frozen in liquid nitrogen immediately after dissection and stored at -80 $^{o}$C until needed for analysis.

### 3.5.2 Construction of normalized cDNA library

Approximately 150 mg (50 mg × 3) of pooled tissue sample was obtained from each of three animals. Total RNA was isolated from tissues using the TRI Reagent$^{TM}$ (Sigma, USA) according to the manufacturer's protocol. To construct the multi-tissue normalized cDNA library for 454 pyrosequencing, poly(A) mRNA transcripts were isolated by means of the FastTrack® 2.0 kit (Invitrogen, USA), following the

manufacturer's instructions. RNA concentration and purity was determined by measuring the absorbance at 260 and 280 nm in a UV-spectrophotometer (Bio-Rad, USA). Full-length, enriched first strand cDNA was generated from 1.5 mg of poly(A+) RNA using a Creator$^{TM}$ SMART$^{TM}$ cDNA library construction kit (Clontech, USA) and following the manufacturer's instructions. Then, cDNA amplification was carried out with the 50× Advantage 2 polymerase mix (Clontech, USA) at 95°C for 7 sec, 66°C for 30 min and 72°C for 6 min. In order to reduce over-representation of the most commonly expressed transcripts, the resulting double stranded cDNA was normalized using the Trimmer-Direct cDNA normalization kit (Evrogen, Russia).

### 3.5.3 GS-FLX 454 sequencing and identification of RbIRF-1 and RbGILT

To sequence the rock bream whole transcriptome, we employed the Roche 454 pyrosequencing platform to analyze our normalized multi-tissue cDNA library constructed above. Pyrosequencing was carried out using the next-generation GS-FLX Titanium and related reagents (DNA Link, Republic of Korea). A single full plate run was performed using normalized cDNA and the reads obtained were processed and assembled through the ARACHNE assembly program (www.broadinstitute.org). Two unique cDNAs which showed homology to known IRF-1 and GILT were identified from homology searches using the Basic Local Alignment Tool (BLAST) algorithm [33].

### 3.5.4 Molecular characterization of RbIRF-1 and RbGILT

Nucleotide and predicted peptide sequences of RbIRF-1 and RbGILT were

analyzed using DNAsisst and BLAST programs. Characteristic domains or motifs were identified using the PROSITE profile database [34]. Signal sequence and putative cleavage sites of RbIRF-1 and RbGILT were identified using the SignalP 3.0 server. Identity, similarity and gap percentages were calculated using the FASTA program [35]. Pair-wise and multiple sequence alignment were analyzed using the ClustalW program, version 1.8 [36]. Disulfide bond was predicted with DiANNA 1.1 web server (http://clavius.bc.edu/~clotelab/DiANNA). A phylogenetic tree was constructed using the Neighbor-Joining method and plotted with the MEGA version 3.1 program [37].

### 3.5.5 Immune challenge of rock bream

To study the transcriptional regulation of RbIRF-1 and RbGILT, we performed two immune challenge experiments using poly I:C and RbIV. Three groups of 20 fish each were challenged with: i) poly I:C; ii) RbIV; or, iii) phosphate buffered saline (PBS) as control. The first group was intraperitoneally (i.p.) injected with 150 μg/fish of poly I:C (Sigma) in phosphate buffered saline (PBS). The second group was injected with 100 mL of RbIV containing supernatant isolated from iridovirus infected rock bream kidney provided by Dr. Sung-Ju Jung from the Department of Aquatic Life Medicine, Chonnam National University, Republic of Korea. The control group was injected with the same volume (as used in experiments) of PBS alone. Rock bream blood, gill, spleen and head kidney tissues were removed at 3, 6, 12, 24, and 48 h post-challenge. Respective control samples were isolated from PBS injected fish.

### 3.5.6 cDNA synthesis for transcriptional analysis of RbIRF-1 and RbGILT

An equal amount (50 mg) of tissue was mixed from three fish to make a pooled sample for RNA isolation. Total RNA was extracted using TRI reagent according to the manufacture's protocol. Sample of 2.5 mg RNA was used to synthesize cDNA from each tissue using a Superscript III first strand synthesis system for RT-PCR kit (Invitrogen). Briefly, RNA was incubated with 1 mL of 50 mM oligo$(dT)_{20}$ and 1 mL of 10 mM dNTPs for 5 min at 65°C. Afterwards, 2 mL of 10× cDNA synthesis buffer, 4 mL of 25 mM $MgCl_2$, 2 mL of dithiothreitol (DTT; 0.1 M), 1 mL of RNaseOUT™ (40 U/mL) and 1 mL of SuperScript III reverse transcriptase (200 U/mL) were added and the mixture was allowed to incubate for 1 h at 50°C. The reaction was terminated by adjusting the temperature to 85°C for 5 min. Then, 1 mL of RNase H was added to each cDNA reaction and incubated at 37°C for an additional 20 min. Finally, synthesized cDNA was diluted 40-fold (total 800 mL volume) before storing at -20°C.

### 3.5.7 Transcriptional analysis of RbIRF-1 and RbGILT by quantitative real-time RT-PCR (qPCR)

The RbIRF-1 and RbGILT expression in various rock bream tissues and after immune challenge was analyzed by qPCR. Gene-specific primers for RbIRF-1 (Forward primer: 5′-CCAACTTCCGCTGTGCAATGAACT-3′ and Reverse primer: 5′-TTCCTCCGTCTTGACCATTGTGCT-3′) and RbGITL (Forward primer: 5′ -TGAGTTGAGCTGGGACAGTGTCAT-3′ and Reverse primer: 5′ -ACATGCTGCAGACCAGAGTGAAGA-3′) were designed to amplify a specific region of each gene. The rock bream beta actin (Accession No. FJ975145) was selected for

use as a reference gene, and it was amplified using Forward primer: 5′ -TCATCACCATCGGCAATGAGAGGT-3′ and Reverse primer: 5′ -TGATGCTGTTGTAGGTGGTCTCGT-3′ gene-specific primers). Tissue-specific mRNA expression was analyzed in rock bream blood, gills, spleen, heart, head kidney, kidney, liver, intestine, muscle, skin and brain. The RbIRF-1 and RbGILT transcriptional regulation changes in response to poly I:C and RbIV challenge were analyzed in blood, gills, spleen and head kidney. qPCR was carried out in a 20 mL reaction volume containing 4 mL of 1:40 diluted original cDNA, 10 mL of 2× SYBR Green Master Mix, 1.0 mL of each primer (10 pmol/mL), and 4.0 mL of PCR grade water using the Thermal Cycler DiceTM Real Time System (TaKaRa, Japan). The qPCR cycling protocol was as follows: one cycle of 95°C for 3 min, amplification for 35 cycles of 95°C for 20 sec, 58°C for 20 sec, 72°C for 30 sec. The baseline was set automatically by the Thermal Cycler Dice$^{TM}$ Real Time System software (version 2.00). The relative expression of each gene was determined by the Livak ($2^{-\Delta\Delta CT}$) method [38], considering rock bream beta actin as the reference gene. The calculated relative expression level of each gene was compared to respective expression levels of blood to determine the tissue-specific expression level. For analysis of the fold-change in expression after immune challenge, relative expression at each time point was compared to respective PBS-injected controls.

## 3.6 RESULTS

### 3.6.1 Analysis of pyrosequencing and identification of RbIRF-1 and RbGILT cDNA sequences

To attain deep sequencing of the rock bream whole transcriptome, we used the high-throughput next generation GS-FLX 454 sequencing platform. A normalized multi-tissue cDNA library was constructed using RNA isolated from various rock bream organs to obtain a large amount of tissue-specific transcripts in hopes of increasing the overall rock bream transcriptome coverage. Nearly 672,000 reads were generated with an average sequence length of 400 bases. After assembly, we identified ~36,000 contigs and 89,000 singletons. Based on the sequence similarity with known proteins and conserved domains, assembled contig sequences were annotated with gene names representative of the rock bream origin, such as IRF-1 (RbIRF-1) and GILT (RbGILT) cDNAs.

### 3.6.2 Sequence characterization of RbIRF-1

The RbIRF-1 cDNA sequence (GenBank accession no. GQ903769) was 1755 bp in length and composed of a 942 bp single open reading frame (ORF) corresponding to 314 amino acid (aa) residues (Fig. 3.1). The estimated molecular mass and the isoelectic point were 36 kDa and 4.8, respectively. A DBD was identified in the N-terminal region (1-113 aa) which exhibited significant match to the consensus and had an e-value $8.32e^{-53}$. Additionally, a signature of the IRF family, the tryptophan pentad repeat, was found to be present in the 26-59 aa range. A bipartite nuclear localization signal (NLS) profile was predicted (120-134 aa) and had a $2.1e^{+04}$

e-value by motif scan analysis. RbIRF-1 had a 139 bp 5′ untranslated region (UTR) and a 674 bp 3′ UTR with a canonical polyadenylation signal sequence, located at $^{1597}$AATAAA$^{602}$ with respect to ATG start codon. To find the closest species IRF-1 to RbIRF-1, we performed pair-wise sequence analysis using the ClustalW program. RbIRF-1 exhibited the highest amino acid identity (84%) to IRF-1 of *Siniperca chuatsi*, but showed only 39% identity to human IRF-1. Furthermore, RbIRF-1 identity to various fish, amphibian, birds and mammalian IRF-1s ranged from 37-84%, while it showed a lower range to all other IRFs family members (IRF2-IRF9) (data not shown). ClustalW multiple alignment analysis revealed that the N-terminal DBD and the five tryptophan residues of RbIRF-1 were highly conserved among selected sequences (Fig. 3.3). It was further revealed that RbIRF-1 has higher amino acid conservation in the N-terminal region than in the C-terminal: a phenomenon generally considered as a characteristic feature of IRF-1 members [22]. RbIRF-1 was found to contain a basic K and R rich sequence ($^{116}$KSRDKRSKAKETKPRKK$^{132}$) located in a similar region in *C. auratus* IRF-1 and the NLS of mammalian IRF-1 proteins [22].

### 3.6.3 Sequence characterization of RbGILT

We identified 1045 bp long cDNA of RbGILT (GenBank accession no.GQ903766) that was composed of a 765 bp ORF containing a 255 amino acids peptide (Fig 3.2). Signal peptide analysis predicted the first 21 aa residues correspond to a signal peptide expected to result in a mature protein (234 aa) with an estimated molecular mass of 26 kDa and an isoelectric point of 5.2. As expected, the characteristic GILT signature sequence (64-169 aa range with $68e^{-25}$ e-value) comprised of a functional domain ($^{97}$CQHGEQECLGNMIETC$^{112}$) and active site $^{74}$C-XX-C$^{77}$ motif were also

found in RbGILT. Furthermore, the cDNA sequence included a 104 bp 5′ UTR and a 106 bp 3′ UTR. The RbGILT mature peptide contained a high number of cysteine residues (n=15) representing 5% of the entire sequence. Disulfide bond prediction results showed that there were seven putative disulfide bonds in different cysteine positions at $^{5}$C-C$^{56}$, $^{13}$C-C$^{22}$, $^{53}$C-C$^{104}$, $^{97}$C-C$^{112}$, $^{126}$C-C$^{220}$, $^{140}$C-C$^{156}$ and $^{209}$C-C$^{231}$. Therefore, we concluded that RbGILT cDNA encodes a complete coding sequence with all the principal characteristic features of a GILT protein family member.

Pair-wise alignment results showed that RbGILT has the highest amino acid identity (85%) to A. fimbria GILT. Furthermore, it had 77% and 39% amino acid identity to *E. coioides* and human GILT sequences, respectively. Multiple alignment analysis of RbGILT was carried out against known GILT family proteins from several phyla of vertebrates. The conserved amino acids in the N-terminal region were relatively low in comparison to those in the middle and C-terminal regions of RbGILT (Fig. 3.4). However, characteristic CXXC motif elements (CPGC) of RbGILT were found to be completely conserved with all aligned GILT sequences, except that of mouse GILT which contains a CGAC at the same position. Moreover, the GILT signature sequence elements showed higher conservation among aligned sequences.

### 3.6.4 Phylogenetic analysis of RbIRF-1 and RbGILT

A total of 54 full-length cDNAs encoding different IRFs (IRF-1 to IRf-9) from fish, amphibians, avians and mammals were selected from the NCBI database to analyze the phylogenetic relationship of RbIRF-1. A phylogenetic tree was constructed by Neighbor-Joining method (Fig. 3.5). This analysis clearly indicated the existence of nine well-defined clusters or sub-groups of IRFs, as had been described in mammals [10]. Rock bream IRF-1 exhibited the closest evolutional relationship

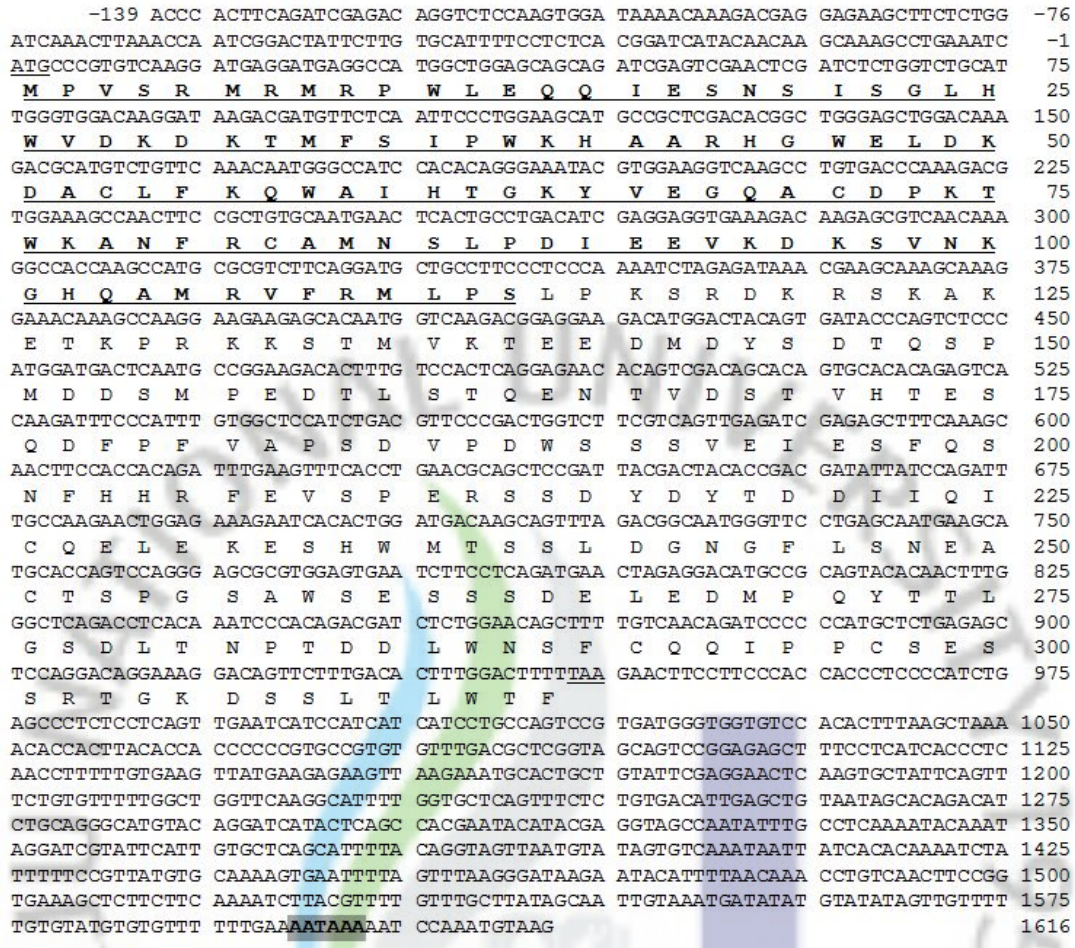



**Figure 3.1 The nucleotide and deduced amino acid sequences of RbIRF-1 cDNA.** The start (ATG) and stop (TAA) codons are underlined. The IFN regulatory factor motif is underlined with bold face type. The predicted polyadenylation sequence is shaded bold type.

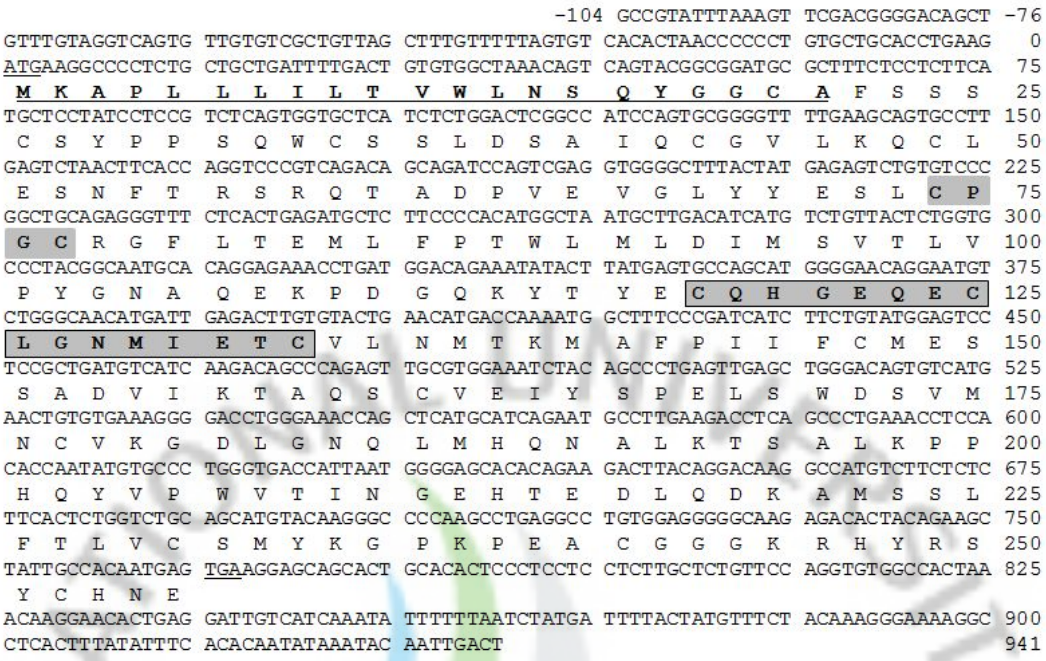



**Figure 3.2 The nucleotide and deduced amino acid sequences of RbGILT cDNA.** The start (ATG) and stop (TGA) codons are underlined. The predicted signal peptide is in bold type and underlined. The predicted C-XX-C motif is shaded bold type. The GILT signature sequence is denoted by a shaded box.

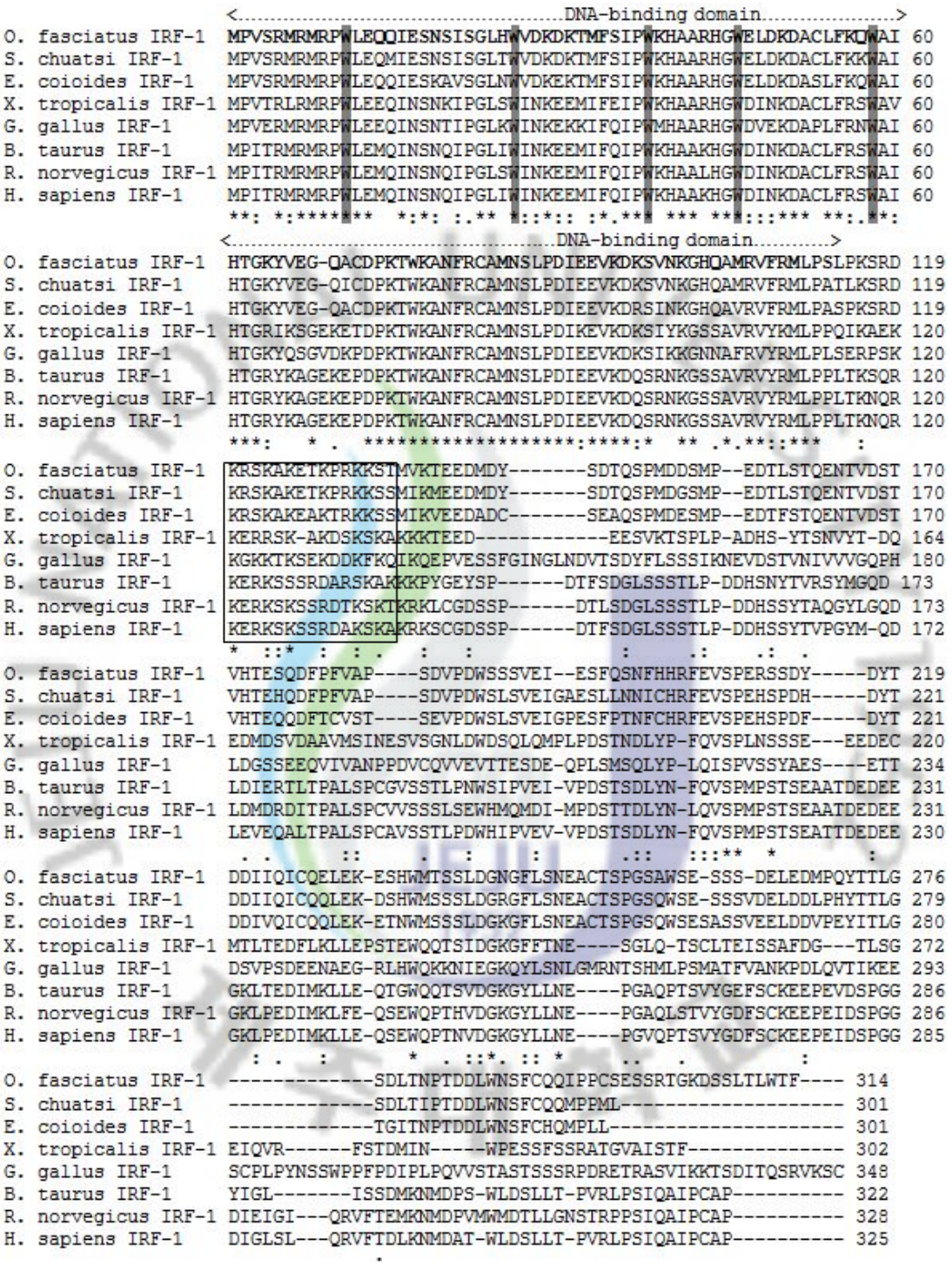



**Figure 3.3 Multiple sequence alignment of RbIRF-1.** Analysis was performed by ClustalW using amino acid sequences of known IRF-1s from different phyla,

including *S. chuatsi* (AAV650412), *E. coioides* (ACF95885), *X. tropicalis* (NM_001006694), *G. gallus* (L339766), *R. norvegicus* (M34253) and *H. sapiens* (L05072). The DNA binding domain is marked by a dotted line above the aligned sequences with bold type (RbIRF-1). The predicted bipartite NLS is boxed. The conserved tryptophan (W) residues are shaded. Identical amino acids are indicated by asterisks (*). Conserved and partially conserved amino acids are indicated by double (:) and single dots (.), respectively.

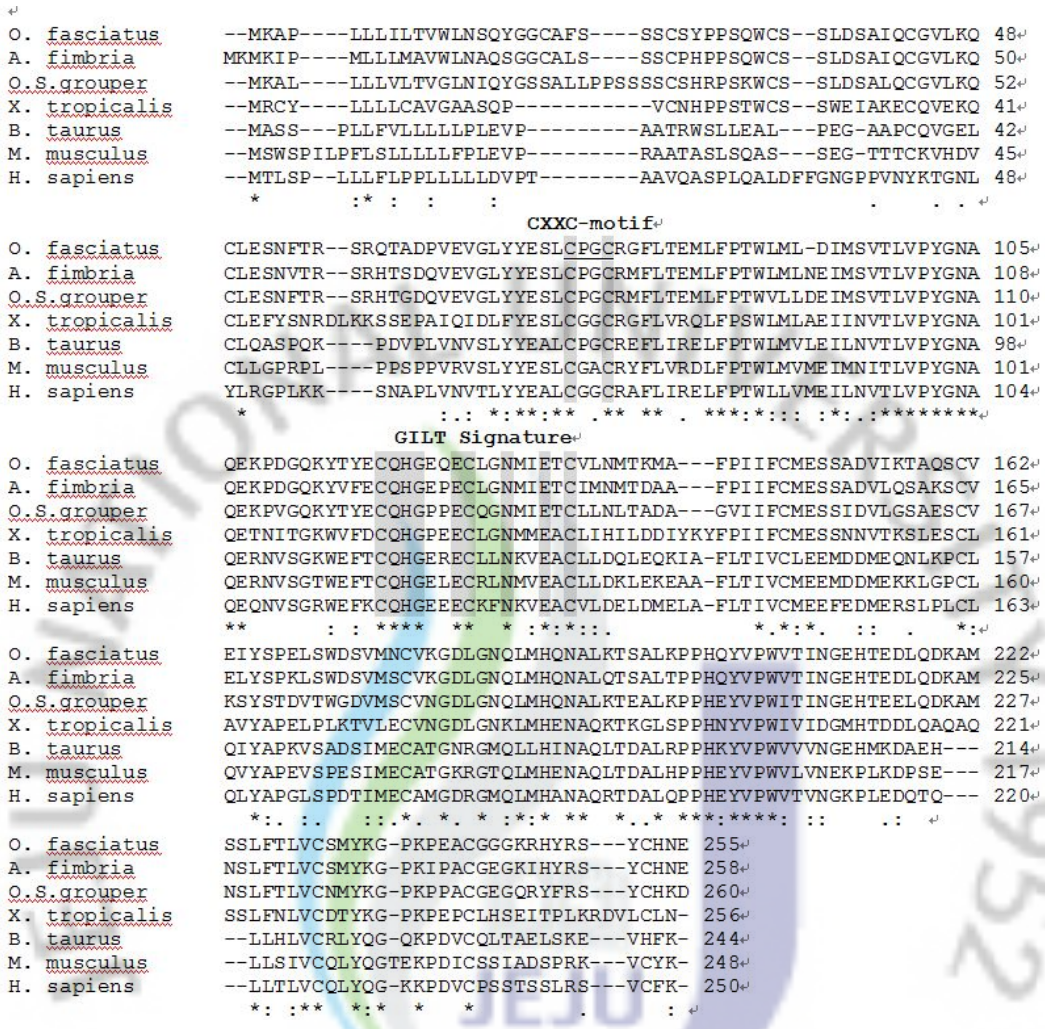



**Figure 3.4 Multiple sequence alignment of RbGILT.** Analysis was performed by ClustalW using amino acid sequences of known GILTs from different phyla, including *A. fimbria* (ACQ58865), *E. coioides* (ABS19625), *X. tropicalis* (NM_001017196), *B. taurus* (NM_001094721), *M. musculus* (NP_075552) and *H. sapiens* (AAH31020). GILT signature sequence and the CXXC motif are indicated above the corresponding aligned sequences, and conserved residues are shaded. Identical amino acids are indicated by asterisks (*). Conserved and partially conserved amino acids are indicated by double (:) and single dots (.), respectively.

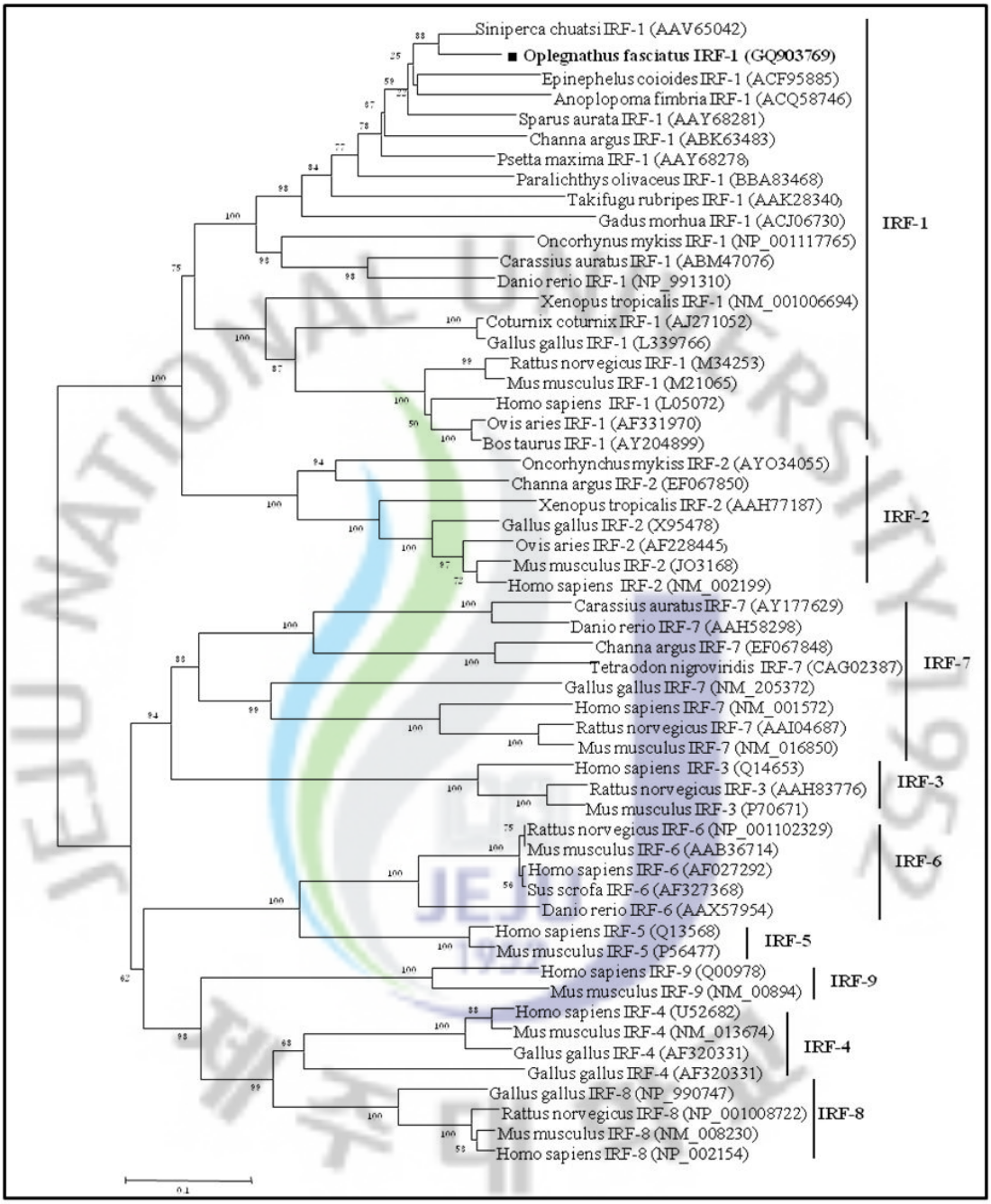



**Figure 3.5 Phylogenetic analysis of RbIRF-1 with other IRF family members.** The tree was constructed based on amino acid sequences. The number at each node indicates the percentage of bootstrapping after 1000 replications.

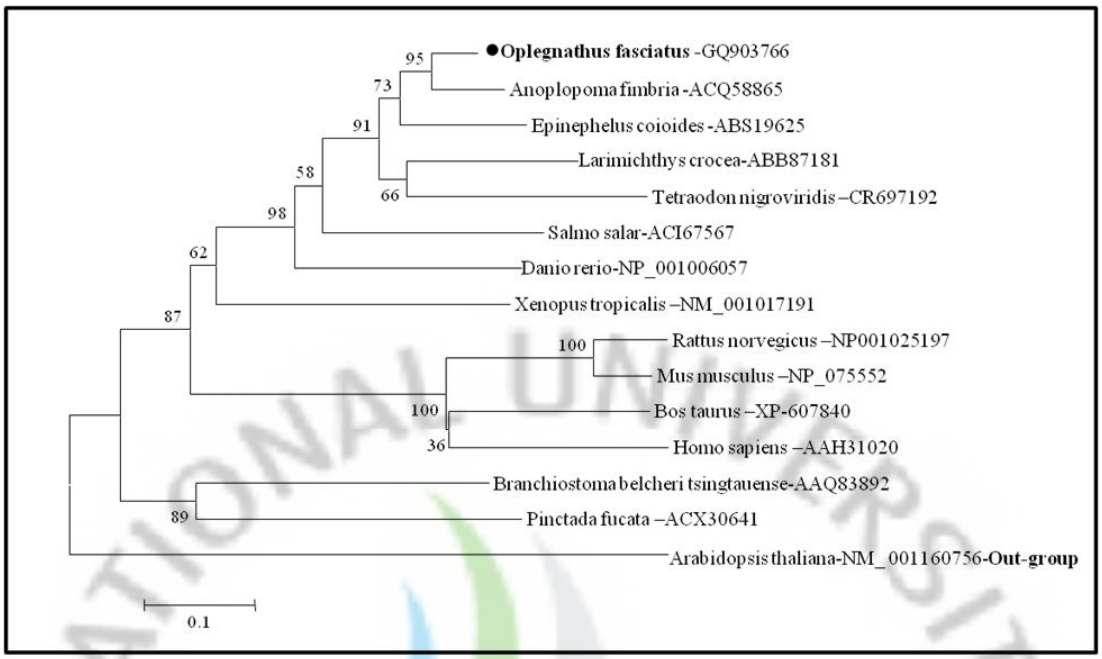



**Figure 3.6 Phylogenetic analysis of RbGILT with known GILT family proteins.** The tree was constructed based on amino acid sequences. The number at each node indicates the percentage of bootstrapping after 1000 replications. *Arabidopsis thaliana* GILT was selected as an out-group member.

to *Siniperca chuatsi* IRF-1, with 88 bootstrap values. Moreover, mammalian species, such as human and mouse, contain IRF1-9 sub family members, but in fish that diversification is limited to a few IRFs or IRF-1 and -2 in a particular species. For example, *Salmon salar* (IRF-1, -2, -3, -7 A and B) [4], *O. mykiss* (IRF-1 and -2) [17] and *C. argus* (IRF-1, -2 and -7) [20] have been identified, but not all other members. Furthermore, we noticed that all members of IRF-1 and IRF-2 grouped under one clade, while all other IRFs formed the second major clade showing the distinct sub-families under IRF 3-9. The newly identified RbIRF-1 from this study is the first type of IRF from rock bream; other IRF sub-families are expected to yield yet unrecognized forms whose identification and characterization will contribute to our understanding of the overall function of IRFs in this fish species.

Likewise, the evolutional relationship of RbGILT was explored by constructing a phylogenetic tree using representative amino acid sequences from invertebrates (mollusk) and vertebrates (fish, amphibians and mammals), and plant *A. thaliana* as an out-group. As shown in Figure 3.6, two distinct classes were observed as vertebrate and invertebrates GILT members. The main vertebrate GILT members were further sub-clustered among the fish, amphibian and mammalian GILTs, where RbGILT was shown to be more evolutionally related to *A. fimbria* GILT, with 95 bootstrap values. This finding suggests that RbGILT may have evolved through a common ancestral gene route.

### 3.6.5 Transcriptional analysis of RbIRF-1and RbGILT by real time RT-PCR

#### 3.6.5.1 Tissue-specific expression

Real-time RT-PCR analysis indicated constitutive expression of RbIRF-1 and RbGILT transcripts occurred in all selected (eleven) tissues of un-challenged rock

bream (Fig. 3.7). Compared to the expression level seen in gill, the highest expression of RbIRF-1 was observed in blood (4.1-fold); RbGILT was highest in gills. The lowest expression level of either gene was observed in muscle. RbIRF-1 expression was higher than RbGILT in heart, liver, intestine, blood and skin. However, almost similar expression levels of RbIRF-1 and RbGILT were identified in spleen, head kidney, muscle, skin and brain.

#### 3.6.5.2 Transcriptional regulation against poly I:C challenge

In order to investigate the rock bream antiviral immune responses through transcriptional regulation analysis of RbIRF-1 and RbGILT, we challenged fish with poly I:C, a synthetic double-stranded RNA which induces the IFN associated immune responses in host cells similar to many viruses [39]. Transcriptional levels of RbIRF-1 and RbGILT were analyzed in blood, gills, spleen and head kidney by real-time RT-PCR with actin as the internal housekeeping gene. As shown in Figure 8A, strong up-regulation of RbIRF-1 was detected in blood, specifically 6.3- and 6.5-fold higher levels were observed at 6 and 12 h post-poly I:C challenge. This induction gradually decreased, nearly to basal levels (1.6-fold) by 48 h post-challenge. In contrast, RbGILT up-regulation in blood was not strong as compared to RbIRF-1. RbGILT expression was up-regulated to its highest level (1.3-fold) at 6 h, and gradually decreased to the basal level by 48 h. In gills, RbIRF-1 and RbGILT transcripts reached their highest levels (2.1 and 1.6 fold, respectively) at an early stage (3 h post-poly I:C challenge (Fig. 3.8B). Both transcripts also showed a gradual decrease in expression and eventually presented expression even below the basal level. In poly I:C challenged fish spleen, RbIRF-1 was up-regulated early (3 h), similar to the gills which showed the highest level at 3.4-fold and then decreased to basal level by 48 h (Fig. 3.8C). However, RbGILT

expression was not increased significantly in spleen and was maintained near basal expression for all the time points examined, with the exception of 12 h. RbIRF-1 and RbGILT transcripts were induced in head kidney at 3 h post-challenge, with the highest peaks being reached at 3.6- and 1.2-fold, respectively (Fig. 3.8D). The induced level of RbIRF-1 was maintained until 12 h post-challenge, while the RbGILT level decreased at 6 h and continued decreasing thereafter. When the transcriptional responses of RbIRF-1 and RbGILT were compared among four rock bream tissue types, we identified RbIRF-1 as having greater expression levels in blood and spleen tissues, all higher than those of RbGILT at all time points. In contrast, a higher RbIRF-1 level was detected in the early stage of poly I:C challenged rock bream gill and head kidney. At the later stage of poly I:C challenge (>24 h), RbGILT expression in gills and head kidney decreased to levels even lower than the basal expression, despite the fact that it had been initially higher than RbIRF-1.

#### 3.6.5.3 Transcriptional regulation against RbIV challenge

To verify whether the transcriptional level of RbIRF-1 and RbGILT was induced in rock bream after RbIV challenge, two groups of fish were i.p. injected with virus and PBS, respectively. Then, transcriptional responses of RbIRF-1 and RbGILT were determined in a manner similar to the polyI:C challenge described above. As shown in Figure 3.9A, both RbIRF-1 and RbGILT expression levels were up-regulated in blood; the highest levels were 1.8-fold (at 48 h) and 2.3-fold (12 h), respectively. We also clearly noticed greater expression (elevated) of RbGILT, as opposed to RbIRF-1, was induced by the RbIv challenge, except at the 24 h time point. This result was completely different from the expression profile of poly I:C-challenged blood, where higher elevated expression occurred with RbIRF-1 instead of RbGILT.

In RbIV injected gills, both RbIRF-1 and RbGILT were up-regulated mainly at 3 and 6 h and then decreased thereafter to levels lower than the basal expression (Fig. 3.9B). Interestingly, there was no significant change in the transcriptional levels in spleen for both genes after RbIV challenge; generally, the expression level was maintained near to the basal expression level, at around 0.8- to 1.1-fold for RbIRF-1 and RbGILT, respectively (Fig. 3.9C). Head kidney displayed early activation (3 h) of RbIRF-1 (1.5-fold) and RbGILT (2.6-fold) transcription; however, the up-regulated status was not maintained beyond 6 h post-challenge (Fig. 3.9D). Both RbIRF-1 and RbGILT expression levels were lower than the PBS injected control at 12-48 h post-challenge of RbIV. The expression profile of RbIRF-1 in response to RbIV was quite different from that to poly I:C, since there was no strong up-regulation observed for the RbIV challenge in any of the four tissues examined. Moreover, differences in the fold change of expression for RbIRF-1 and RbGILT (at each time point) were lower in the RbIV challenge than in the poly I:C challenge.

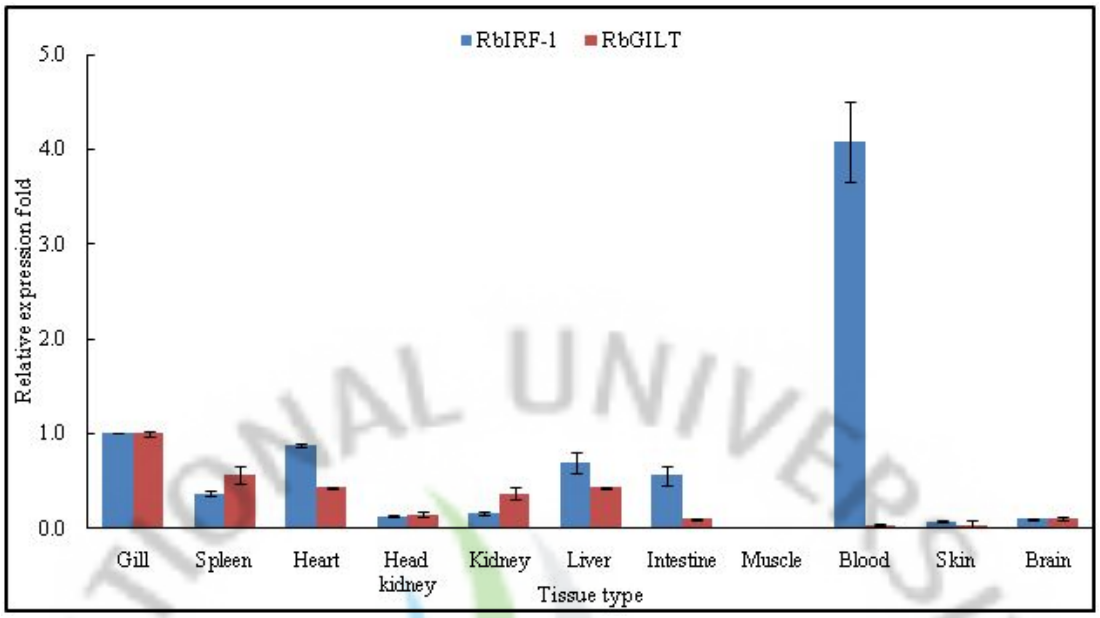



**Figure 3.7 Tissue distribution of RbIRF-1 and RbGILT.** Analysis of mRNA was carried out by qPCR. The relative fold-change in expression was calculated by the $2^{-CT}$ method using beta actin as a reference gene. The relative fold-change in expression of each tissue was compared to the expression in blood in order to determine the levels of tissue-specific expression. Data are presented as the mean of the relative expression ± SD calculated for three replicate real-time reactions using pooled tissues of three individual rock bream at each time point.

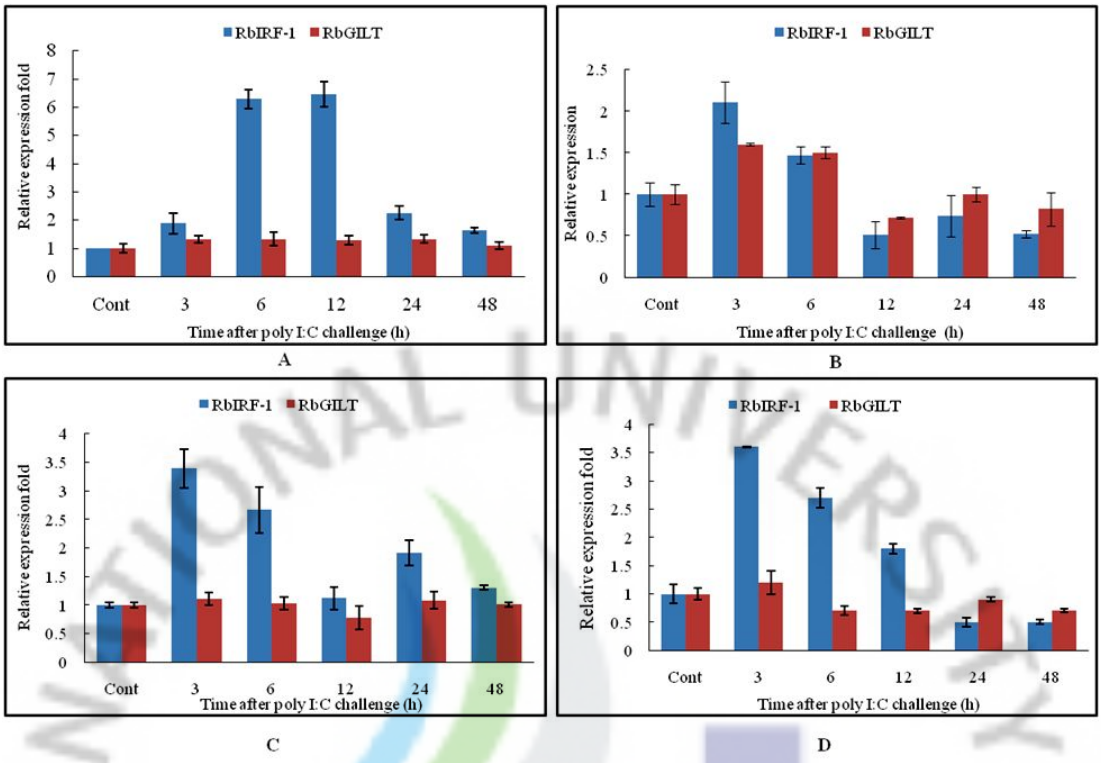



**Figure 3.8 Transcriptional regulation of RbIRF-1 and RbGILT in blood, gills, spleen and head kidney following poly I:C challenge.** Analysis of transcriptional level was carried out by qPCR. The expression fold was calculated by the $2^{-CT}$ method using rock bream beta actin as a reference gene. The relative level of expression at each time point was compared to that of the PBS-injected control (Cont). Data are presented as the mean relative expression ± SD calculated for three replicate real-time reactions using pooled tissues of three individual rock bream at each time point. (A) blood, (B) gills, (C) spleen and (D) head kidney.

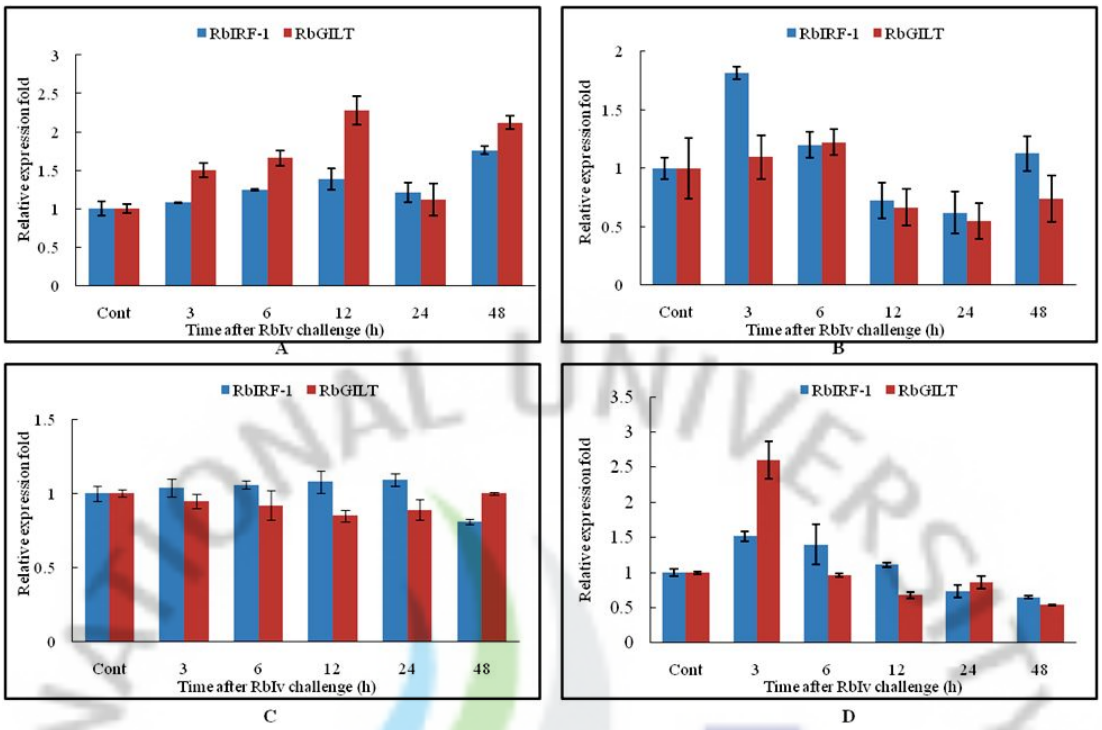



**Figure 3.9 Transcriptional regulation of RbIRF-1 and RbGILT in blood, gills, spleen and head kidney following RBIV challenge.** Analysis of transcriptional level was carried out by qPCR. The expression fold was calculated by the $2^{-CT}$ method using rock bream beta actin as a reference gene. The relative level of expression at each time point was compared to that of the PBS-injected control (Cont). Data are presented as the mean relative expression ± SD calculated for three replicate real-time reactions using pooled tissues of three individual rock bream at each time point. (A) blood, (B) gills, (C) spleen and (D) head kidney.

## 3.7 DISCUSSION

Next generation sequencing technologies, such as Roche 454 pyrosequencing, can be effectively used for gene discovery [40]. These methods have not yet been widely applied to marine fish, as compared to their use in mammalian genome studies. Here, we describe the application of whole transcriptome sequencing analysis to the rock bream by which we identified IRF-1 and GILT as two IFN regulatory molecules. Moreover, we were able to illustrate that immune response reactions of rock bream IRF-1 and GILT are likely associated with host defense mechanisms against virus infection. Ozato *et al*. [11] previously made structural and functional comparisons of IRF proteins and determined that IRF-1 is composed of 325 amino acids and harbors an N-terminal DBD. Mammalian IRFs share this characteristic homology feature with the N-terminal 115 amino acids (first) encompass a DBD [41]. The DBD itself contains five tryptophan (W) repeats, spaced by 10-18 aa which are essential for the tertiary structure. The promoter regions of many ISGs contain a short DNA sequence motif, termed the interferon consensus sequence (ICS), which facilitates binding of regulatory proteins such as IRF-1 [42]. The newly identified RbIRF-1 contained 314 amino acids and had a highly conserved N-terminal DBD (113 aa) with five tryptophan residues. Moreover, RbIRF-1 was found to be most similar to *Siniperca chuatsi* IRF-1 (84% identity) and grouped within the fish IRF-1 clade in phylogenetic analysis. Very recently, Bergan *et al.* [4] (2010) described that Atlantic salmon IRF-1 has an amino acid identity range of 40-84% with other fish IRF-1s, while it showed a lower range with other IRF family members. Similarly, RbIRF-1 displayed over 50% amino acid identity to other fish IRF-1s and 30-39% to other vertebrate IRFs. Further, it has been reported that the C-terminal of IRF-1 has no significant similarity among fish, chicken or mammals [20]. Likewise, we identified few

conserved amino acid residues in the C- terminus of RbIRF-1 that corresponded to other fish, amphibian, avian or mammalian IRF-1 proteins. Also, members of the IRF family harbor a functional NLS which supports localization to the nucleus [9]. Shi *et al*. [23] identified two NLS signals in *C. auratus* IRF-1 (CaIRF-1), one of which has structural and functional similarity to that of mammalian IRF-1. The second NLS of CaIRF-1 contained basic amino acid sequence, including $K_{75,}$ $K_{78,}$ $R_{82,}$ $K_{95,}$ $K_{97}$ and $K_{101}$. Similarly, we identified basic amino acids at $K_{116}$, $R_{118}$, $K_{120}$, $R_{121}$, $K_{123}$, $K_{125}$, $K_{128}$, $R_{129}$, $K_{131}$ and $K_{132}$ (amino acids 116-132) from RbIRF-1, which may be associated with the nuclear localization function. The main difference is that the second NLS of CaIRF-1 is positioned in the DBD, where as it is outside of the DBD in RbIRF-1. Therefore, all the sequence data summarized in the present study suggest that RbIRF-1 is a true member of the mammalian IRF-1s and is more similar to fish IRF-1.

In this study, we also described the molecular characterization of GILT cDNA sequence from rock bream, which is associated with IFN-ɣ regulated immune response. RbGILT mature protein showed all characteristic motifs of the GILT family members, including a C-XX-C active site motif ($^{74}$CPGC$^{77}$) and GILT signature sequence ($^{118}$CQHGEQECLGNMIETC$^{120}$); it also exhibited a relatively high percentage (15%) of cysteine residues (n=15). The RbGILT active site C-XX-C motif is highly conserved among fish, amphibian and mammalian GILT proteins, and it exactly matched in terms of amino acid composition (CPGC) with all aligned fish GILT sequences. The consensus GILT signature sequence CQHGX2ECX2NX4C was found to be $^{118}$CQHG(SE)EC(LG)N(IIAS)C$^{120}$ in RbGILT, with a high level of conserved amino acid residues to other known GILT proteins. RbGILT contains a putative N-linked glycosylation site at $^{53}$NFTR$^{56}$ thus, it could be derived by mannose-6-phosphate (M6P), which is considered as an essential requirement for

transportation of GILT to the lysosomal system [26]. Collins et al. [43] proposed that disulfide reduction in lysosomes was largely catalyzed by a high concentration of cysteine, which is transported into the lumen by amino acid transporters. Therefore, we suggest that the high amounts of cysteine residues associated with the seven disulfide bonds in RbGILT may be related to the function of reducing disulfide bonds during antigen processing and presentation. In phylogenetic analysis, RbGILT was evolutionally connected with other members of fish GILT genes, indicating a common ancestral origination during evolution. Based on the molecular characterization of IRF-1 and GILT in rock bream, we suggest that ancient origin and persistence of these genes have occurred throughout evolution in order to maintain efficient immune defense functions in fish.

Analysis of the differential expression pattern of RbIRF-1 and RbGILT mRNA in various rock bream tissues revealed that the two genes were constitutively expressed in eleven organs but to different extents in each tissue. The RbIRF-1 mRNA expression observed in various rock bream tissues was generally in agreement with ubiquitously and constitutively expressed fish IRF-1 reported in *Salmon salar* [4], *P. olivaceus* [15], *F. rubripes* [16], *O. mykiss* [17], *S. maximus* and *S. aurata* [18]. More similarly, *Salmon salar* IRF-1 was highly expressed in gills (compared to the expression level in brain) and to a much lower extent in muscle [4]. Rock bream IRF-1 also showed the highest expression in blood, followed by gills and having the lowest level in muscle. Since IRFs are known to be involved in cell growth and differentiation, ubiquitous expression of IRF-1 would be associated with alternative functions other than merely the maintenance of innate immunity. Blood circulation is important to maintain homeostasis between different organs and it flows through the gill filaments to facilitate collection of available oxygen from the passing water. Therefore, the higher level of IRF-1 in blood may be associated with the necessity

of readily available IRF-1 to circulate via blood for activation of different IFNs or their regulatory genes in other tissues. However, tissue distribution of IRF-1 has shown wide variation even among fish; constitutive expression patterns of IRF-1 differ in *P. olivaceus* [15], *F. rubripes* [16], *Oncorhynchus mykiss* [17], *S. maximus* and *S. aurata* [18]. This variation may be a result of the particular growth stage, physiological status and other culturing conditions of fish examined in each study.

On the other hand, constitutive expression of GILT has been detected in APCs, indicating its role in the MHC class II antigen processing pathway [44-45]. In fish, constitutive expression (basal level) of GILT has been identified in *P. crocea* liver, gills, brain and heart [30], and in *E. coioides* liver, gills, brain, heart, kidney, muscle and spleen [31], which are similar to the expression levels of RbGILT observed in this study. The outcome from tissue specific expression revealed that all these fish organs express IFN regulatory factor (RbIRF-1), as well as IFN inducible protein (RbGILT), indicating that the IFN system may function even without stimulation and act to maintain the first line of defense against pathogens.

It was reported that IRF-1 can be induced by IFN, polyI:C and viruses in the mammalian host defense response to regulate IFNs and interferon inducible genes [46]. Furthermore, IRF-1 has shown interactions with various immune response molecules such as TNF-α [47]. In fish, polyI:C treatment stimulated the IRF-1 expression in muscle, spleen and kidney of *S. maximus*, *S. aurata* [18], and in gill, head kidney, liver, spleen and gonad cells of *O. mykiss* [17]. In rock bream, IRF-1 mRNA expression was induced in blood, gills, spleen and head kidney by polyI:C treatment, and this result was expected as those organs are recognized as immune functional organs in fish [48]. In contrast to RbIRF-1, RbGILT up-regulation against polyI:C challenge was not strong in any of the analyzed tissues. This situation could be due to the inherent nature of polyI:C, since it is a stronger inducer of type I IFN

(αβ) than of type II IFN (ɣ) [49] and the transcription of GILT is mainly activated by IFN (ɣ). In fact, from the present results we are unable to explain whether RbGILT induction occurs via direct transcriptional activation by polyI:C or indirectly through induced IRF-1 molecules. Also, it is important to note that polyI:C induced the early IRF-1 expression in rock bream tissues, suggesting that this activity could lead to activating IFN and/or IFN regulatory genes.

At present, iridovirus has become one of the major pathogens causing serious epizootics to cultured rock bream in Korea [32]. Kim et al. [50] produced polyclonal antibody against recombinant protein (ORF049L) of RbIV, but otherwise only few attempts have been made towards developing vaccines against RbIV infection. In the present study, RbIRF-1 expression was induced by RbIV challenge in blood, gills and head kidney, but not in spleen. In general, fish IRF-1 has shown up-regulation against several viruses, including the turbot kidney against VHSV [18] and in *C. auratus* blastulae embryonic cells against grass carp hemorrhagic virus [22]. In fact, certain virus-like VHSV was not able to induce IRF-1 in fish [34]. Jia et al. [21] characterized the C. argus IRF-1 (CgIRF-1) promoter which contains putative NF-kB, ISRE like and GAS sequences. The presence of GAS in IRF-1 indicates that it has regulatory role with IFN (ɣ) and downstream genes like GITL. Caipang et al. [23] have shown that transfected Japanese flounder IRF-1 (JFIRF-1) is able to induce the antiviral state against hirame rhabdovirus and VHSV and suggested that JFIRF-1 may regulate the production of type II IFN (ɣ). The present study also may support this idea since RbIv challenge induced the expression of both RbIRF-1 and RbGILT in blood, gills and head kidney. Also, greater transcriptional activation of RbIRF-1 and RbGILT occurred at the early stage of RbIV challenge and then declined; a similar situation has been observed in Atlantic salmon IRF-1 in response to infectious salmon anemia virus (ISAV). Such an early action of IRF-1 in the IFN system was

suggested by Tanaka et al. [6], which suggests that it is essential and critical to activate ISGs in a timely manner to result in efficient host resistance against pathogens.

In conclusion, we have found that rock bream IFN regulator transcription factor (RbIRF-1) and downstream target gene of IFN- (RbGILT) are constitutively expressed in all tissues examined. In addition, poly I:C treatment was able to up-regulate rock bream IRF-1 and GILT transcripts in blood, gills, spleen and head kidney, mainly at the early stages. RbIv challenge induced the RbIRF-1 and RbGILT to a lesser extent (in blood, gills and head kidney) than did the poly I:C. On the basis of the observed ubiquitous and constitutive expression, as well as the transcriptional up-regulation against poly:IC and RbIv suggest that RbIRF-1 and RbGILT are related to (innate) immune responses against viral infection, and that this activation may be essential for subsequent adaptive immunity in rock bream. IFN-ɣ has not been identified from rock bream until now, and the findings from the current study will help to advance our understanding of the roles and molecular mechanisms of IRF-1 and IFN-ɣ in regulation of GILT gene transcription.

## 3.8 REFERENCES


[1] Koyama S, Ishii KJ, Coban C, Akira S. Innate immune response to viral infection. Cytokine 2008;43:336-41.

[2] Takaoka A, Yanai H. Interferon signalling network in innate defense. Cell Microbiol 2006;8:907-22.

[3] Stark GR, Kerr IM, Williams BR, Silverman RH, Schreiber RD. How cells respond to interferons. Annu Rev Biochem 1998;67:227-64.

[4] Bergan V, Kileng O, Sun B, Robertsen B. Regulation and function of interferon regulatory factors of Atlantic salmon. Mol Immunol 2010;47:2005-14.

[5] Huang B, Qi ZT, Xu Z, Nie P. Global characterization of interferon regulatory factor (IRF) genes in vertebrates: Glimpse of the diversification in evolution. BMC Immunol 2010;11-22.

[6] Escalante CR, Yie J, Thanos D, Aggarwal AK. Structure of IRF-1 with bound DNA reveals determinants of interferon regulation. Nature 1998;391:103-6.

[7] Fujita T, Sakakibara J, Sudo Y, Miyamota M, Kimura Y, Taniguchi T. Evidence for a nuclear factor(s), IRF-1, mediating induction and silencing properties to human IFN-beta gene regulatory elements. EMBO J 1988;7:3397-405.

[8] Tanaka N, Ishihara M, Lamphier MS, Nozawa H, Matsuyma T, Mak TW, Aizawa S, Tokino T, Oren M, Taniguchi T. Cooperation of the tumor suppressor IRF-1 and p53 in response to DNA damage. Nature 1996;382:816-8.

[9] Kim EJ, Park CH, Park JS, Um SJ. Functional dissection of the transactivation domain of interferon regulatory factor-1. BioChem BioPhys Res Commun 2003;304:253-9.

[10] Nguyen H, Hiscott J, Pitha PM. The growing family of interferon regulatory factors. Cytokine Growth Factor Rev 1997;8:293-312.

[11] Ozato K, Tailor P, Kubota T. The interferon regulatory factor family in host defense: Mechanism of action. J Biol Chem 2007;282:20065-9.

[12] Romeo G, Fiorucci G, Chiantore MV. Percario ZA, Vannucchi S, Affabris E. IRF-1 as a negative regulator of cell proliferation. J Interferon Cytokine Res 2002;22:39-47.

[13] Tanaka N, Ishihara M, Kitagawa M, Harada H, Kimura T, Matsuyama T, Lamphier MS, Aizawa S, Mak TW, Taniguchi T. Cellular commitment to oncogene-induced transformation or apoptosis is dependent on the transcription factor IRF-1. Cell 1994;77:829-39.

[14] Kroger A, Koster M, Schroeder K, Hauser H, Mueller PP. Activities of IRF-1. J Interferon Cytokine Res 2002;22:5-14.

[15] Yabu T, Hirose H, Hirono I, Katagiri T, Aoki T, Yamamoto E. Molecular cloning of a novel interferon regulatory factor in Japanese flounder, *Paralichthys olivaceus*. Mol Mar Biol Biotechnol 1998.7:138-44.

[16] Richardson MP, Tay BH, Goh BY, Venkatesh B, Brenner S. Molecular cloning and genomic structure of a gene encoding interferon regulatory factor in the pufferfish (*Fugu rubripes*). Mar Biotechnol 2001;3:145-51.

[17] Collet B, Guido C, Hovens J, Mazzoni D, Hirono I, Aoki T, Secombes C. Cloning and expression analysis of rainbow trout *Oncorhynchus mykiss* interferon regulatory factor 1 and 2 (IRF-1 and IRF-2). Dev Comp Immunol 2003;27:111-26.

[18] Ordas MC, Abollo E, Costa MM, Figueras A, Novoa B. Molecular cloning and expression analysis of interferon regulatory factor-1 (IRF-1) of turbot and sea bream. 2006. Mol Immunol 2006;43:882-90.

[19] Sun BJ, Chang MX, Song Y, Yao WJ, Nie P. Gene structure and transcription of IRF-1 and IRF-7 in the mandarin fish *Siniperca chuatsi*. Vet Immunol Immunopathol 2007;116:26-36.

[20] Jia W, Guo Q. Gene structures and promoter characterization of interferon regulatory factor 1 (IRF-1), IRF-2 and IRF-7 from snakehead *Channa argus* Mol Immunol 2008;45:2419-28.

[21] Shi Y, Zhu XP, Yin JK, Zhang QY, Gui JF. Identification and characterization of interferon regulatory factor-1 from orange-spotted grouper (*Epinephelus coioides*). Mol Biol Rep 2010;37:1483-93.

[22] Shi Y, Zhang YB, Zhao Z, Jiang J, Zhang QY, Gui JF. Molecular characterization and subcellular localization of Carassius auratus interferon regulatory factor-1. Dev Comp Immunol 2008;32:134-46.

[23] Caipang CM, Hirono I, Aoki T. Induction of antiviral state in fish cells by Japanese flounder *Paralichthys olivaceus*, interferon regulatory factor-1. Fish Shellfish Immunol 2005;19:79-91.

[24] Zenke K, Kim KH. Molecular cloning and expression analysis of three Mx isoforms of rock bream *Oplegnathus fasciatus*. Fish Shellfish Immunol 2009;26:599-605.

[25] Zenke K, Nam YK, Kim KH. Molecular cloning and expression analysis of double strand RNA dependant protein kinase (PKR) of rock bream *Oplegnathus fasciatus*. Fish Shellfish Immunol 2010;133:290-5.

[26] Arunachalam B, Phan UT, Geuze HJ, Cresswell P. Enzymatic reduction of disulfide bonds in lysosomes: Characterization of interferon inducible lysosomal thiol reductase (GILT). Proc Natl Acad Sci USA 2000;97:745-50.

[27] Lackman RL, Cresswell P. Exposure of the promonocytic cell line THP-1 to Escherichia coli induces IFN-gamma-inducible lysosomal thiol reductase expression by inflammatory cytokines. J Immunol 2006;177:4833-40.

[28] O'Donnell PW, Haque A, Klemsz MJ, Kaplan MH, Blum JS. Cutting edge: Induction of the antigen-processing enzyme IFN-gamma inducible lysosomal thiol reductase in melanoma cells is STAT1-dependant but CIITA-independent. J Immunol 173,731-5.

[29] Phan UT, Maric M, Dick TP, Cresswell P. Multiple species express thiol oxidoreductase related to GILT. Immunogenetics 2001;53:342-56.

[30] Zheng W, Chen X. Cloning and expression analysis of interferon gamma-inducible lysosomal thiol reductase gene in large yellow croaker (*Pseudosciaena crocer*). Mol Immunol 2006;43:2135-41.

[31] Dan WB, Ren F, Zhang C, Zhang SQ. Molecular cloning and expression analysis of interferon-gamma inducible lysosomal thiol reductase gene in orange-spotted grouper, *Epinephelus coioides*. Fish Shell fish Immunol 2007;23:1315-23.

[32] Robertsen B. The interferon system of teleost fish. Fish Shellfish Immunol 2006;20:172-91.

[33] Jung SJ, Oh MJ. Iridovirus-like infection associated with high mortalities of striped beakperch, *Oplegnathus fasciatus* (Terminck et Schlegel), in southern coastal area of the Korean peninsula. J Fish Dis 2000;23:223-6.

[33] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic alignment search tool. J Mol Biol 1990;215:403-10.

[34] Bairoch A, Bucher P, Hofmann K. The PROSITE data base, its status in 1997. Nucleic Acids Res 1997;25:217-21.

[35] Pearson WR. Rapid and sensitive sequence comparison with FASTP and FASTA. Methods Enzymol 1990;183:63-98.

[36] Thompson JD, Higgins DG, Gibson TJ. CLUSTALW: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap Penalties and weight matrix choice. Nucleic Acids Res 1994;22:4673-80.

[37] Kumar S, Tamura K, Nei M. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform 2004;5:150-63.

[38] Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods 2001;25:402-8.

[39] Alexopoulou L, Holt AC, Medzhitov R, Flavell RA. Recognition of double-stranded RNA and activation of NF-kappaB by Toll-like receptor 3. Nature 2001;413:732-8.

[40] Cheung F, Haas BJ, Goldberg SMD, May GD, Xiao Y, Town CD. Sequencing *Medicago truncatula* expressed sequenced tags using 454 Life Sciences Technology. BMC Genomics 2006;7:272.

[42] Weisz A, Marx P, Sharf R, Appekka E, Driggers PH, Ozato K, Levi BZ. Human interferon consensus sequence binding protein is a negative regulator of enhancer elements common to interferon inducible genes. J Biol Chem 1992;35:25589-96.

[43] Collins DS, Unanue ER, Harding CV. Reduction of disulfide bonds within lysosome is a key step in antigen processing. J Immunol 1991;147:4054-9.

[44] Luster AD, Weinshank RL, Feinman R, Ravetch JV. Molecular and biochemical characterization of a novel gamma interferon inducible protein. J Biol Chem 1998;263:12036-43.

[46] Taniguchi T, Ogasawara K, Takaoka A, Tanaka N. IRF family of transcription factors as regulators of host defense. Annu Rev Immunol 2001;19:23-55.

[47] Davidson GA, Lin SH, Secombes CJ, Ellis AE. Detection of specific and constitutive antibody secreting cells in the gills, head kidney, and peripheral blood leucocytes of dab (*Limanda limanda*). Vet Immunol Immunpathol 1997;58:363-74.

[49] Kim YI, Ha YM. Ahn SJ, Nam YK, Kim KH, Kim SK. Production and characterization of polyclonal antibody against recombinant ORF049L of rock bream (*Oplegnathus fasciatus*) iridovirus. Process Biochem 2007;42:134-40.

[50] Tanaka N, Kawakami T, Taniguchi T. Recognition DNA sequence of interferon regulatory factor 1 (IRF-1) and IRF-2, regulators of cell growth and the interferon system. Mol Cell Biol 1993;13:4531-8.

# PART 4

**Molecular characterization and expression analysis of Cathepsin B and L cysteine proteases from rock bream (*Oplegnathus fasciatus*)**

## 4.1 ABSTRACT

Cathepsins are lysosomal cysteine proteases of the papain family that play an important role in intracellular protein degradation. The full-length sequences of cathepsin B (RbCathepsin B) and L (RbCathepsin L) were identified after transcriptome sequencing of rock bream *Oplegnathus fasciatus* mixed tissue cDNA. Cathepsin B was composed of 330 amino acid residues and having 36 kDa molecular mass. RbCathepsin L contained 336 amino acid residues encoding for a 38 kDa molecular mass protein. The sequencing analysis results showed that both cathepsin B and L contain the characteristic papain family cysteine protease signature and active sites for the eukaryotic thiol proteases cysteine, asparagine and histidine. In addition, RbCathepsin L contained EF hand $Ca^{2+}$ binding and cathepsin propeptide inhibitor domains. The rock bream cathepsin B and L showed the highest amino acid identity of 90 and 95% to *Lutjanus argentimaculatus* cathepsin B and *Lates calcarifer* cathepsin L, respectively. By phylogenetic analysis, cathepsin B and L exhibited a high degree of evolutionary relationship to respective cathepsin family members of the papain superfamily.

Quantitative real time RT-PCR analysis results confirmed that cathepsin B and L gene expression was constitutive in all examined tissues isolated from un-induced rock bream. Moreover, challenge with LPS and *Edwardsiella tarda* resulted in significant up-regulation of RbCathepsin B and L mRNA in liver and blood, indicating a role for cathepsin B and L in immune responses against bacteria in rock bream.

**Keywords:** Cathepsin; lysosomal cysteine proteases; papain superfamily; rock bream; *Oplegnathus fasciatus*

## 4.2 LIST OF TABLES

Page
Table 4.1 Description of primers used in the study ..... 140

## 4.3 LIST OF FIGURES

Page

Figure 4.1 The nucleotide and deduced amino acid sequences of rock bream cathepsin B (A) and L (B) cDNAs ..... 145

Figure 4.2 Multiple sequence alignment of the RbCathepsin B (A) and L (B). 147

Figure 4.3 Phylogenetic relationship of cathepsin family members was constructed based on amino acid sequences ..... 149

Figure 4.4 Tissue distribution of cathepsin B and cathepsin L of rock bream .. 150

Figure 4.5 Transcriptional regulation of cathepsin B and cathepsin L of rock bream in liver (A) and blood (B) following LPS challenge ..... 151

Figure 4.6 Transcriptional regulation of cathepsin B and cathepsin L of rock bream in liver (A) and blood (B) following *E. tarda* challenge ..... 152

## 4.4 INTRODUCTION

The cathepsin protein family is composed of lysosomal proteolytic enzymes that play an important role in maintaining homeostasis in organisms. In particular, these proteases are involved in intracellular protein degradation/turn over [1], antigen processing [2], hormone maturation [3] and host immune responses [4]. The cathepsin cysteine proteases belong to the C1 family, which is also known as the papain family [5]. Different cathepsin isoforms have been identified and classified based on sequence homology, specific or conserved amino acid motifs (cysteine, serine, aspartate), and specific tissue distribution pattern. Two major cathepsins groups have been classified according to their tissue distribution. The first group is composed of ubiquitously expressed members, including cathepsin B, C, F, H, L, O and Z, while the second group consists of cathepsin J, K, S and W which have restricted expression in certain tissues [6]. Moreover, the cathepsins can be divided into three major sub-groups based on their sequence homology and specific amino acid motifs; these include the cathepsin B-like, cathepsin L-like and cathepsin F-like genes [7].

Cathepsin B has been characterized for its function to degrade proteins in the lysosomal system that entered from outside the cell. In addition, cathepsin B function has also been implicated in many pathological and physiological processes, including inflammation, infection clearance, apoptosis and cancer [8,9]. On the other hand, several studies have demonstrated the ability of human cathepsin L to cleave the various substrates of extracellular, serum, cytoplasmic and nuclear proteins. For example, cathepsin L can act on fibronectin, collagen, and elastin to modulate the extracellular matrices [10,11]. Tryselius et al. [12] described the expression of cathepsin L-like cysteine protease 1 in *Drosophila melanogaster* hemocyte cell line (mbn-2), which supposed to be involved in phagocytosis reaction. Another recent

study revealed that mouse cathepsin L can be suppressed the TH1 immune response which drastically increases susceptibility to bacterial infections [4]. Therefore, cathepsin exhibits a multifunctional role in various organisms. However, cathepsins B and L were cloned only from few fish species such as *Danio rerio* (XM692675/BC066490), *Cyprinus carpio* (AB215097/AB128161), *Paralichthys olivaceus* (AY686604), *Fundulus heterolitus* (AY217741/AY212286) and *Oncorhynchus mykiss* (AF358667/AF358668). Therefore, further studies are required to elucidate the specific immune responses and other associated functions of cathepisn B and L in low-order vertebrates such as fish. To date, no cathepsin isoform has been cloned or identified from the rock bream.

Immune responses in fish can be experimentally stimulated by either injection of bacterial LPS or bacteria [13]. It was reported that low concentration of LPS could be beneficial to fish for enhancing non-specific immune responses such as phagocytic activity and proliferation of B and T lymphocytes [14]. Rock bream is a commercially important fish in marine aquaculture. Edwardsiellosis is one of the most problematic bacterial diseases of rock bream as well as in many other fresh and marine fish caused by gram negative bacteria *Edwardsiella tarda* [15]. *E. tarda* infection has been shown to induce several immune reactions in fish [16-18]. Therefore, fish challenge with LPS or *E. tarda* has great potential to understand immune regulatory effects or responses of cathepsins as well as host pathogen interactions. Additionally, results of LPS challenge could be used to compare the effect of *E. tarda* on cathepsin responses in rock bream.

In the present study, we have identified and characterized the cathepsin B and L cDNAs from rock bream using GS-FLX-based transcriptome sequencing technique. Transcriptional analysis was investigated to test immune responses of cathepsin B and L after experimentally challenge of rock bream using LPS and *E. tarda*.

## 4.5 MATERIALS AND METHODS

### 4.5.1 Animal rearing and RNA isolation

Healthy rock bream Oplegnathus fasciatus fish (mean weight 30g) were obtained from the Ocean and Fisheries Research Institute (Jeju Special Self-Governing Province, Republic of Korea). Fish were acclimatized for one week in a 400 L stock tank at 22-24ºC in our laboratory. For RNA isolation, rock bream brain, gills, liver, kidney, head kidney, spleen, skin, intestine and muscle were dissected from three un-challenged fish. Blood samples (approximately 1-2 mL/fish) were taken from the caudal fin using a 22 G syringe; samples were immediately centrifuged 3000 x g at 4ºC for 10 min to separate blood cells. All tissues and blood cells were directly frozen in liquid nitrogen before storage at -80ºC. Total RNA was isolated from all selected tissue samples separately using the TRI Reagent™ (Sigma, USA) according to the manufacturer's protocol. Poly(A) mRNA was then isolated using FastTrack® 2.0 kit according to the manufacturer's instructions (Invitrogen, USA). The RNA concentration and purity was determined by measuring the absorbance at 260 and 280 nm in a UV-spectrophotometer (BioRad, USA).

### 4.5.2 Multi-tissue cDNA synthesis and normalization

Purified mRNA from various tissues was diluted up to 1 mg/uL and pooled equally before synthesis of multi-tissue cDNA. Full-length, enriched first strand cDNA was generated from 1.5 mg of poly(A+) RNA using a Creator™ SMART™ cDNAlibrary construction kit (Clontech, USA) following the manufacturer's instructions. Then, cDNA amplification was carried out with the 50× Advantage 2 polymerase mix (Clontech, USA) for 95°C for 7 sec, 66°C for 30 min and 72°C for 6 min. In order to reduce over-representation of the most commonly expressed

transcripts, the resulting double stranded cDNA was normalized using the TRIMMER-DIRECT cDNA normalization kit (Evrogen, Russia).

### 4.5.3 Pyrosequencing, sequence processing, assembly and identification of RbCathepsin B and L

The rock bream multi-tissue normalized cDNA was used for parallel pyrosequencing in an automated sequencing platform using GS-FLX titanium system (454, Roche, USA). The creation of rock bream multi-tissue cDNA GS-FLX shotgun library and GS-FLX pyrosequencing (using the GS-FLX Titanium reagents) were carried out using a next-generation sequencing platform (DNA Link, Republic of Korea). A single full plate run was performed using normalized cDNA and the reads obtained were processed and assembled through the assembly program (ARACHNE). Two unique cDNAs which showed homology to known cathepsin B and L were identified from homology searches using the Basic Local Alignment Tool (BLAST) algorithm [19].

### 4.5.4 Molecular characterization of RbCathepsin B and L

Nucleotide and predicted peptide sequences of RbCathepsin B and Lwere analyzed using DNAsisst and BLAST programs. Characteristic domains or motifs were identified using the PROSITE profile database [20]. Signal sequence and putative cleavage site of RbCathepsin B and L were identified using the SignalP 3.0 server. Prediction of the pro-region cleavage sites and active sites were based on the alignment of cathepsin sequences with the vertebrate orthologues. Identity, similarity and gap percentages were calculated using the FASTA program [21]. Pair-wise and multiple sequence alignment were analyzed using the ClustalW program, version 1.8 [22]. The phylogenetic tree was constructed using the Neighbor-Joining method and plotted with MEGA version 3.1 program [23].

### 4.5.5 LPS and bacterial challenge

To study the immune response of RbCathepsin B and L, we performed two immune challenge experiments using bacteria *E.tarda* and purified LPS. The *E. tarda* strain was kindly provided by Dr. Sung-Ju Jung from the Department of Aquatic Life Medicine, Chonnam National University, Republic of Korea. *E. tarda* was cultured in BHIS broth at 30°C for 16 h. Then, the bacterial culture was centrifuged at 7000 x g at 4°C for 5 min. The supernatant fluid was discarded and the bacteria lpellets were resuspended in 1x phosphate-buffered saline (PBS). Fish were intraperitonealy (i.p.) injected with 100 mL of *E. tarda* (5 x $10^6$CFU/ml). In the LPS challenge experiment, fish were injected with 125 mg (50 g fish) of LPS purified from *E.coli* 055:B5 (Sigma). The control group was injected with the same volume (as used in experiments) of PBS. Rock bream liver and blood were removed at 3, 6, 12, 24, and 48 h post-challenge. Respective samples were isolated from PBS injected fish. The tissue samples were immediately snap-frozen in liquid nitrogen and stored at -80°C until use for RNA isolation.

### 4.5.6 cDNA synthesis

An equal amount (50 mg) of tissue was mixed from three fish to make a pooled sample for RNA isolation. The total RNA was extracted from the pooled tissue (150 mg) and blood using TRI Reagent according to the manufacturer's protocol. The total RNA was stored at -80°C until further use. The RNA concentration and purity was determined by measuring the absorbance at 260 and 280 nm in a UV-spectrophotometer (BioRad). Purified RNA was diluted upto 1mg/uL concentration before synthesis of cDNA. A sample of 2.5 mg RNA was used to synthesize cDNA from each tissue using a Superscript III first-strand synthesis system for RT-PCR kit (Invitrogen). Briefly, RNA was incubated with 1 mL of 50 mM oligo$(dT)_{20}$ and 1 mL of 10 mM dNTPs for 5 min at 65°C. After incubation, 2 mL of 10x cDNA

synthesis buffer, 2 mL of 25 mM $MgCl_2$, 2 mL of dithiothreitol (DTT, 0.1 M), 1 mL of RNaseOUT$^{TM}$(40U/mL) and 1 mL of SuperScript III reverse transcriptase (200 U/mL) were added and incubated for 1 h at 50°C. The reaction was terminated by adjusting the temperature to 85°C for 5 min. Then, 1 mL of RNase H was added to each cDNA reaction and incubated at 37°C for an additional 20 min. Finally, synthesized cDNA was diluted 10-fold (total 200 mL) before storing at -20°C.

### 4.5.7 Expression analysis of RbCathepsin B and L by quantitative real-time RT-PCR

The RbCathepsin B and L expression in various tissues in un-challenged and immune challenged fish were analyzed by quantitative real-time RT-PCR. Gene-specific primers were designed to amplify corresponding regions of RbCathepsin B, L and beta actin (Table 1). The rock bream beta actin (Accession No. FJ975145) was selected as a house keeping gene in this study. Tissue-specific mRNA expression was analyzed in blood cells, gills, liver, spleen, head kidney, kidney, muscle and intestine. The RbCathepsin B and L mRNA response was determined in liver and blood cells after LPS and *E. tarda* challenge. The quantitative real-time RT-PCR was carried out in a 20 mL reaction volume containing 4 mL of 1:10 diluted original cDNA, 10 mL of 2× SYBR Green Master Mix, 1.0 mL of each primer (10 pmol/mL), and 4.0 mL of PCR grade water using Thermal Cycler Dice$^{TM}$ Real Time System (TaKaRa, Japan). The quantitative real-time RT-PCR cycling protocol was as follows: one cycle of 95°C for 3 min, amplification for 35 cycles of 95°C for 20 sec, 58°C for 20 sec, 72°C for 30 sec. The baseline was set automatically by the Thermal Cycler DiceTM Real Time System software (version 2.00). The relative expression of each gene was determined by the Livak ($2^{-\Delta\Delta CT}$) method [24], considering the rock bream beta actin as a reference gene. The calculated relative expression level of each gene was compared with respective expression levels of blood cells for the tissue-specific expression. For analysis of the fold-change after LPS and E. tarda injection, relative expression at each time point of challenged fish

was compared to respective PBS-injected and un-injured controls, respectively.

**Table 4.1 Description of primers used in the study**

| Name | Target | Primer sequence (5'- 3') |
|---|---|---|
| RbCathepsin B-F | Real time PCR amplification | GTTATTGGCTGCCAGCTTGTCGTT |
| RbCathepsin B-R | Real time PCR amplification | AAGTTGTGACCAGCCTTCCAGGTA |
| RbCathepsin L-F | Real time PCR amplification | TGGTGTGGGAGAAGAACCTGAAGA |
| RbCathepsin L-R | Real time PCR amplification | CGATTTGCGCTTGTAGCCGTTCAT |
| Rbbeta actin-F | Reference gene Real time PCR | TCATCACCATCGGCAATGAGAGGT |
| Rbbeta actin-R | Reference gene Real time PCR | TGATGCTGTTGTAGGTGGTCTCGT |

Rb rock bream; F, forward; R, reverse

## 4.6 RESULTS

### 4.6.1 Identification and characterization of RbCathepsin B and L cDNA sequences

The pyrosequencing of rock bream normalized cDNA library using Roche GS-FLX technology generated a total of 915,242 reads of raw nucleotide sequences. BLAST analysis of our rock bream transcriptome data led to identify two cathepsin sequences, namely rock bream cathepsin B (RbCathepsin B) and cathepsin L (RbCathepsin L). We then cloned the complete coding sequence of RbCathepsin B and L to confirm the identified sequence using two sets of gene specific primers and cDNA generated from liver tissue. The sequence results showed that there was no difference between respective pyrosequencing sequence and amplified coding sequence. Hence, RbCathepsin B and L were deposited in NCBI under accession numbers of GQ903767 and HM060314, respectively.

The cDNA of RbCathepsin B contained 1753 bp full-length and a 990 bp ORF that translated into a putative peptide of 330 amino acid residues (Fig. 4.1a). It was consisted of 96 bp length 5' untranslated region (UTR) and 667 bp 3' UTR, that harbored three RNA instability motifs (ATTTA). The RbCathepsin B polypeptide had a putative molecular mass of 36 kDa with a 5.3 isoelectric point (p*I*). SignalP prediction results revealed that there was a putative signal peptide in the N-terminal of RbCathepsin B. Several characteristic signature domains or active sites, such as the papain family cysteine protease (79-328 aa), propeptide C1 peptidase family (25-65 aa), and eukaryotic thiole proteases cysteine (101-112 aa), asparagine (292-311 aa) and histidine (275-285 aa), were identified within the amino acid sequence of RbCathepsin B. Moreover, a single ASN-glycosylation site (37-40 aa) and 16 cysteine residues were identified in the mature cathepsin B peptide.

The full-length (1450 bp) of RbCathepsin L encompassed a 1008 bp ORF that encodes a 336 amino acid putative protein (Fig. 1b). Predicted molecular mass and

p*I* of RbCathepsin L was 38 kDa and 5.5, respectively. It had 57 bp and 385 bp lengths for the 5' UTR and 3' UTR, respectively. Amino acid sequence analysis revealed that RbCathepsin L contained several characteristic features of the cathepsin family, including peptidase C1 papain family cysteine protease signature (115-335 aa), EF hand $Ca^{2+}$ binding domain (80-97 aa) and cathepsin propeptide inhibitor domain (28-97), thiol protease active sites of cysteine (133-144), histidine (227-287) and asparagin (398-317). Moreover, RbCathepsin L had a predicted signal peptide that showed the cleavage site located between the 16-17 amino acid position.

To distinguish the RbCathepsin B and L from other cathepsin family members, we analyzed the pair-wise identity using ClustalW program. Results revealed that RbCathepsin B was 90, 70 and 54% identical to fish *Lutjanus argentimaculatus*, human and *Brancheostoma belcher*i (invertebrate) cathepsin B amino acid sequences, respectively. On the other hand, it was shown that RbCathepsin L has 95, 75, 60% identities to Asian seabass (*Lates calcarifer*), human and *Artemia franciscana* (invertebrate) cathepsin L sequences, respectively. To compare with known cathepsins B and L, we aligned complete amino acid sequences of RbCathepisn B and L using ClustalW multiple analysis. It was shown that all selected cathepsin B and L from fish, amphibian, birds and mammalian species contain an N-terminal signal peptide. Also, we identified that previously reported mature light chain, heavy chain of cathepsin B sequences were aligned with RbCathepsin B showing higher amino acid matching (Fig. 4.2A). Multiple alignment of RbCathepsin L with six cathepsin L sequences from other organisms demonstrated that the ERF/WNIN-like motif was present in the pro region of the rock bream cathepsin L, which was more or less identical to the respective motif of other species.

We were interested in determining the phylogenetic relationship of RbCathepsin B and L by the Neighbor-Joining method using human cathepsin A as an out-group (Fig. 4.3). Phylogenetic tree construction was accomplished using 39 different cathepsin sequences (cathepsin A, C, D, H, L, S, W, V and Z) selected from fish, amphibian, avian and mammalian species. We were able to distinguish that

RbCathepsin B and L positions in two separate clades. RbCathepsin B was closely related to *Lutajanus argentimaculatus* cathepsin B, while RbCathepsin L was more related to cathepsin L of *Hippoglossus hippoglossus*. Fish and mammalian cathepsin sub-clusters were identified in both cathepsin B and L clades. Selected cathepsin sequences of other organisms, such as cathepsin A, C, D, H, S, W, V and Z, were grouped in relevant positions with other family members of their respective group.

### 4.6.2 Tissue expression profiles of RbCathepsin B and L

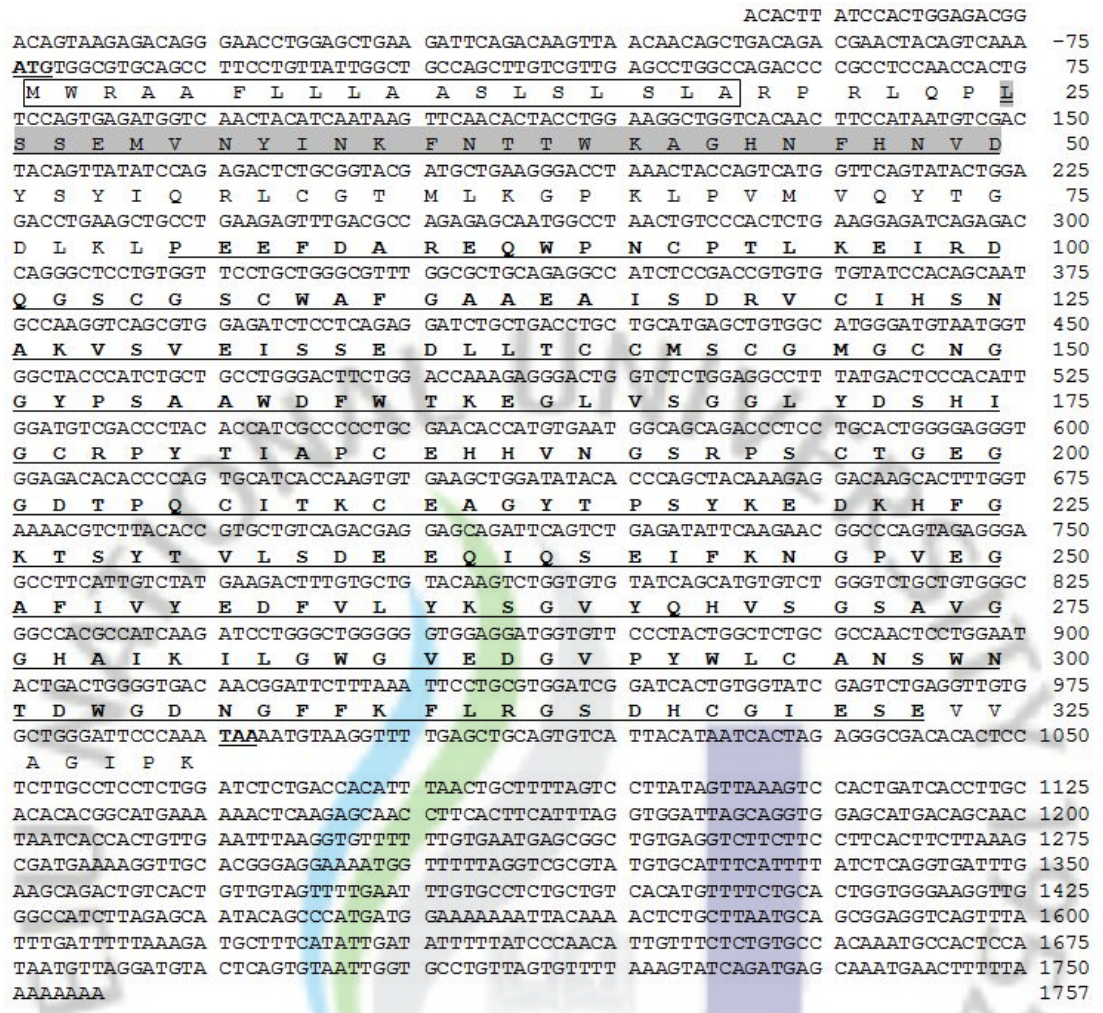
To determine the tissue-specific mRNA expression profiles of RbCathepsin B and L, quantitative real-time RT-PCR was carried out using various tissues of un-challenged rock bream. The expression level of selected tissues was normalized to that of beta actin as a house keeping gene. Relative expression fold was calculated based on the expression in blood cells to determine the tissue expression profile (Fig. 4.4). Expression results revealed that both RbCathepsin B and L transcripts were ubiquitously expressed in all analyzed tissues, namely blood, gill, liver, spleen, head kidney, kidney, muscle and intestine.

### 4.6.3 Expression analysis of RbCathepsin B and L after LPS and *E. tarda* challenge

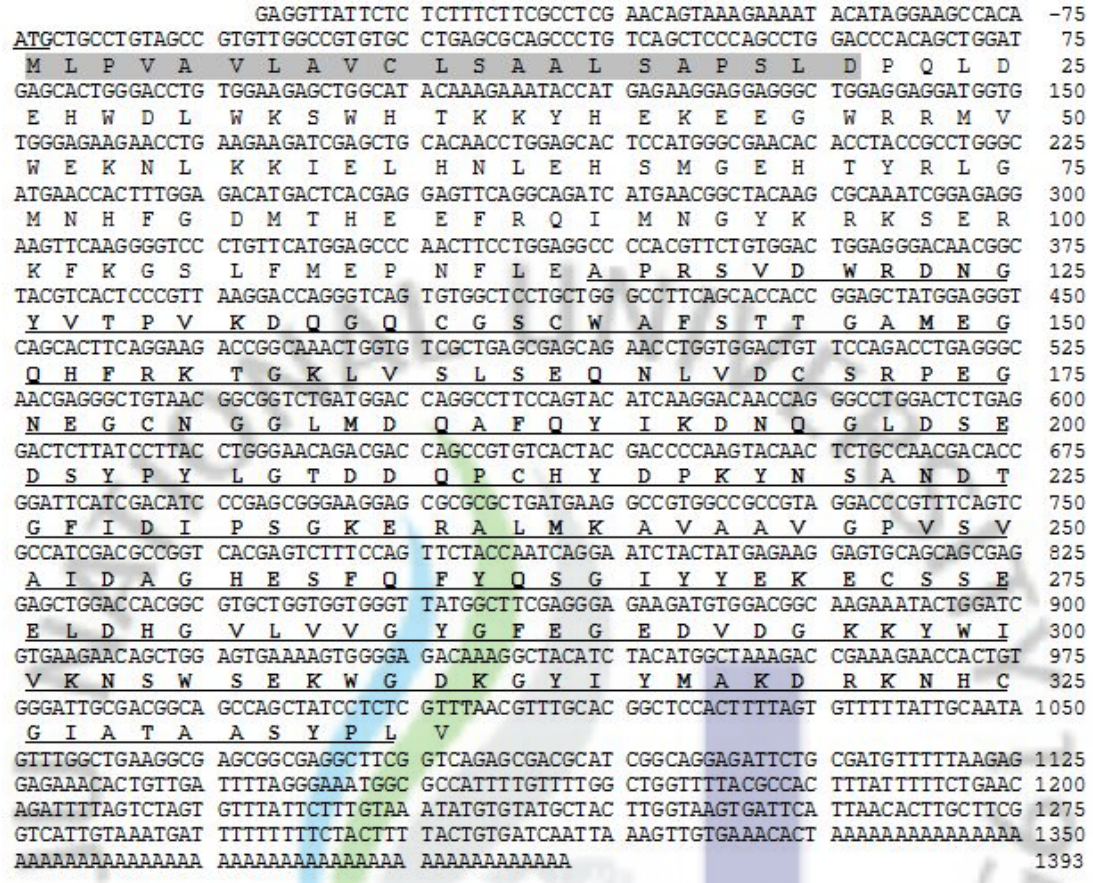
To examine the in vivo transcriptional responses of RbCathepsin B and L, fish were challenged with LPS and bacteria (*E. tarda*) separately. Expression profiles were determined in liver and blood using qPCR. Rock bream Cathepsin B and L mRNA expression analysis after LPS challenge is shown in Figure 4.5. Following the LPS challenge in liver, RbCathepsin B mRNA level was decreased up to the 24 h time point and then increased by the 48 h time point (Fig. 4.5A). In contrast, RbCathepsin L was increased at 3 h and then decreased during the following 3-48 h post-LPS challenge, as compared to the control. Present results showed that RbCathepsin B and L expression profiles were different in liver tissue after LPS

challenge. Additionally, we analyzed Cathepsin B and L transcriptional activation in rock bream blood cells. Expression profiles of Cathepsin B and L in blood cells is shown in Figure 5B. Both Cathepsin B and L transcription was activated at 3 h post-challenge with LPS. However, they were decreased at 6 and 12 h and then increased to the highest induced states of 6.2-fold (cathepsin B) and 38.1-fold (cathepsin L) at 24 h post-LPS challenge. Notably, both cathepsin B and L expression in blood cells was reduced at 48 h (compared to levels at 24 h) but was higher than the respective control. Both RbCathepsin B and L induction was higher in blood cells than liver against LPS challenge.

To examine the expression of RbCathepsin B and L during bacterial infection, we analyzed mRNA expression after challenge with *E. tarda*. In liver, RbCathepsin B expression remained un-changed at 3 and 6 h, compared to control (Fig. 4.6A). Then, the level decreased below the control level during the 12-48 h post- *E. tarda* challenge. However, induced RbCathepsin L was noticed only at 6 and 24 h in liver after *E. tarda* challenge. Interestingly, transcriptional up-regulation of RbCathepsin B was gradually increased in blood up to 48 h after *E. tarda* challenge (Fig. 4.6B). Meanwhile, RbCathepsin L level was gradually up-regulated up to the 24 h and then decreased at 48 h in blood cells but remained higher than the respective level in the control sample. The highest expression level of RbCathepsin B (3.7-fold) and L (2.0-fold) was detected at 48 h and 24 h, respectively after *E. tarda* challenge. Transcriptional activation of RbCathepsin B and L by LPS and *E. tarda* revealed that stronger up-regulation occurred in rock bream blood cells than in liver tissue.

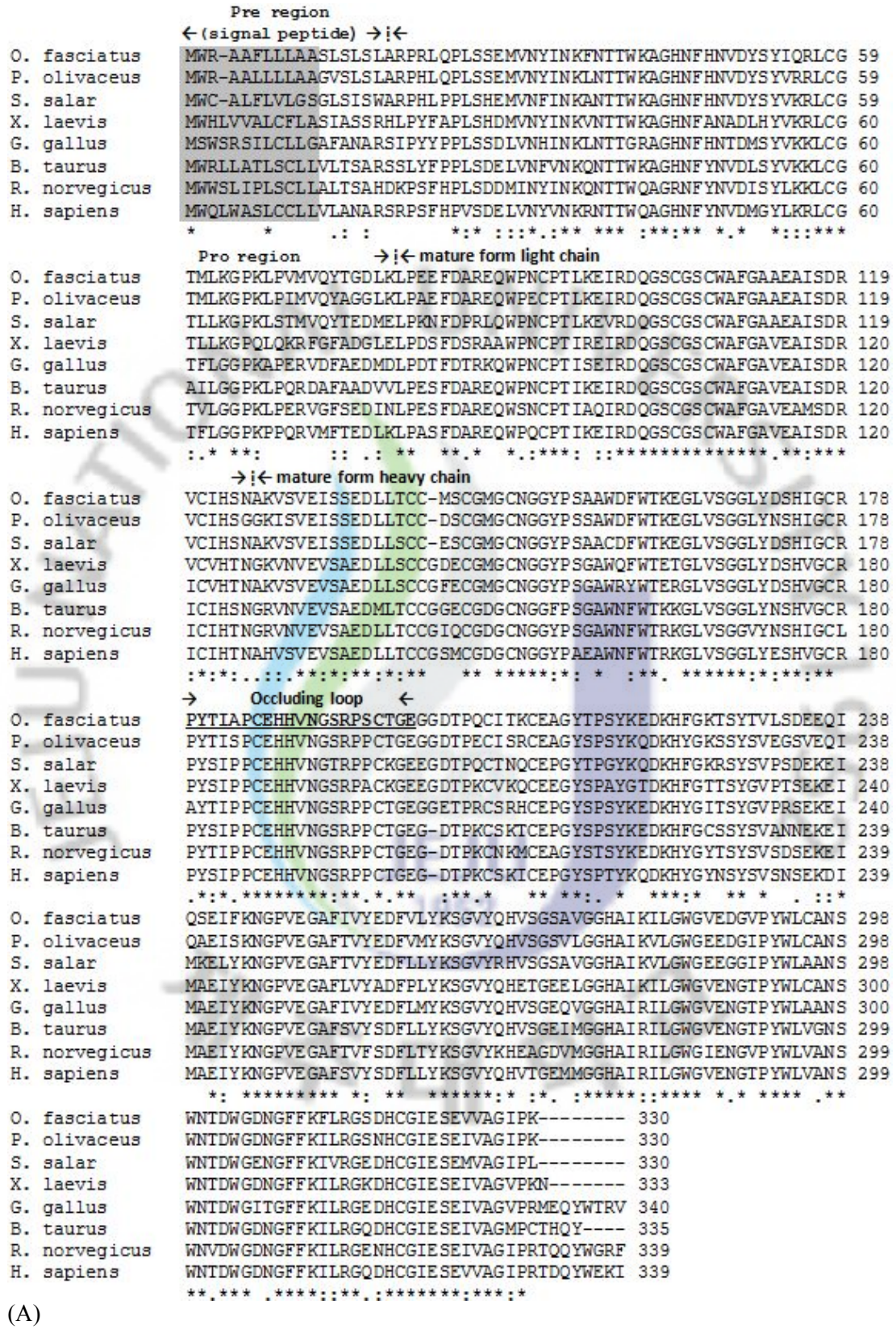



(A)



(B)

**Figure 4.1 The nucleotide and deduced amino acid sequences of rock bream cathepsin B (A) and L (B) cDNAs.** The start (ATG) and stop (TAA) codons are in bold and underlined. The predicted signal peptides are in boxes. The predicted papain family cysteine protease signatures are bold underlined. The propeptide C1 peptidase family motif domain is shaded and underlined.




(A)


(B)

**Figure 4.2 Multiple sequence alignment of the RbCathepsin B (A) and L (B).** Analysis was performed by ClustalW (1.81) using representatives of cathepsin B selected from different phyla. Identical amino acids are indicated by asterisks.

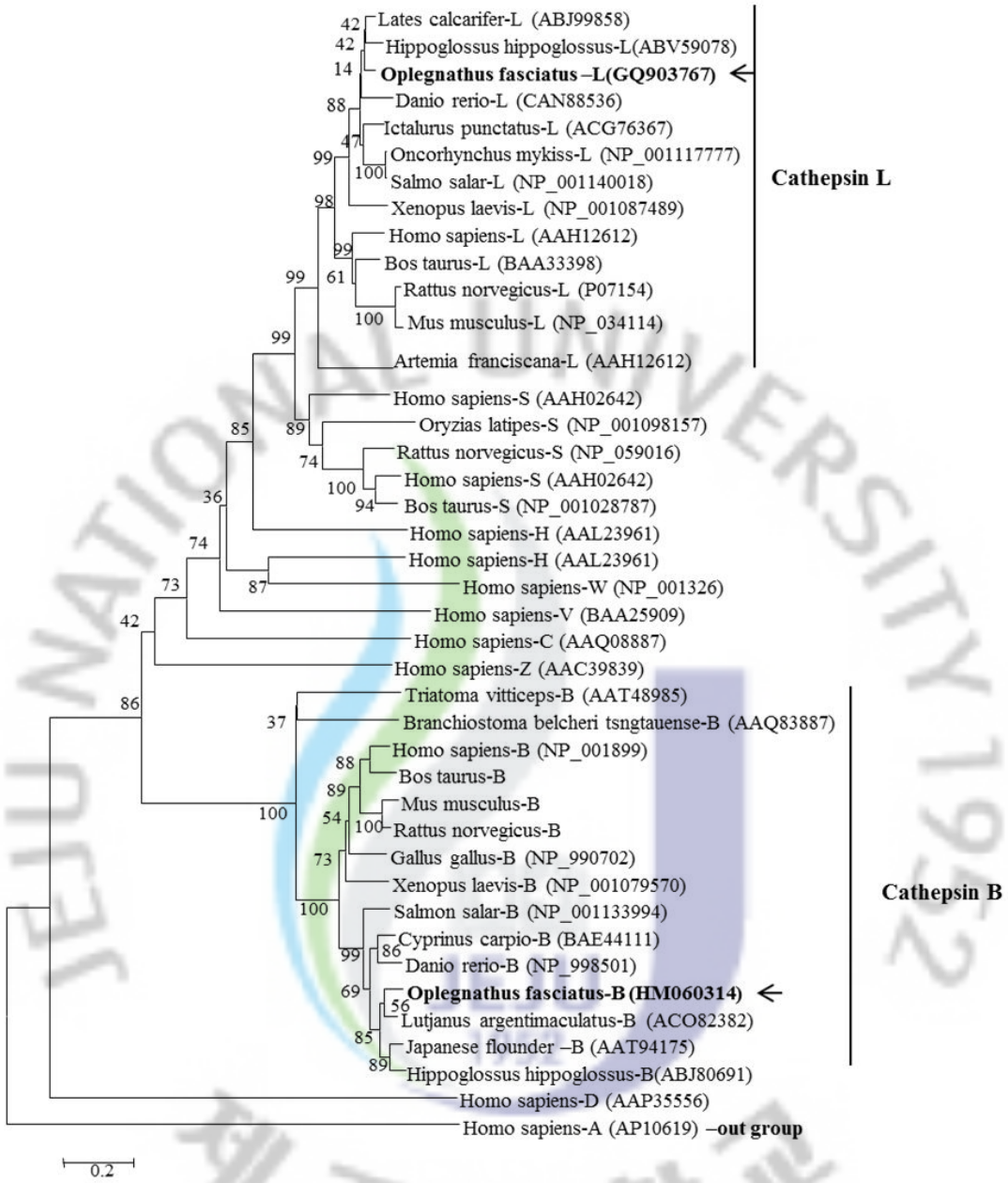



**Figure 4.3 Phylogenetic relationship of cathepsin family members was constructed based on amino acid sequences.** The number at each node indicates the percentage of bootstrapping after 1000 replications. The GenBank accession numbers of selected cathepsins are indicated within brackets. Human cathepsin A was selected as an out-group member.

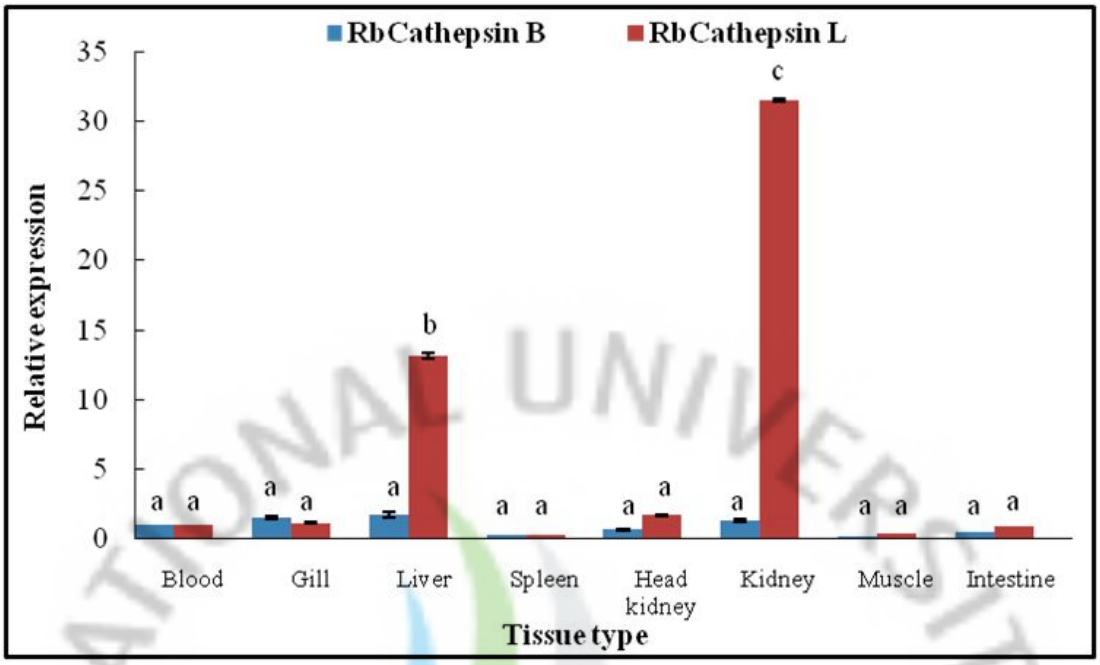



**Figure 4.4 Tissue distribution of cathepsin B and cathepsin L of rock bream.** Analysis of mRNA was carried out by qPCR. The relative fold-change in expression was calculated by the $2^{-CT}$ method using beta actin as a reference gene. The relative fold-change in expression of each tissue was compared to the expression in blood for determining the levels of tissue-specific expression. Data are presented as the mean of the relative expression ± SD for three replicate real-time reactions from pooled tissue of three individual rock bream at each time point. Differences were considered statistically significant at $P<0.05$.

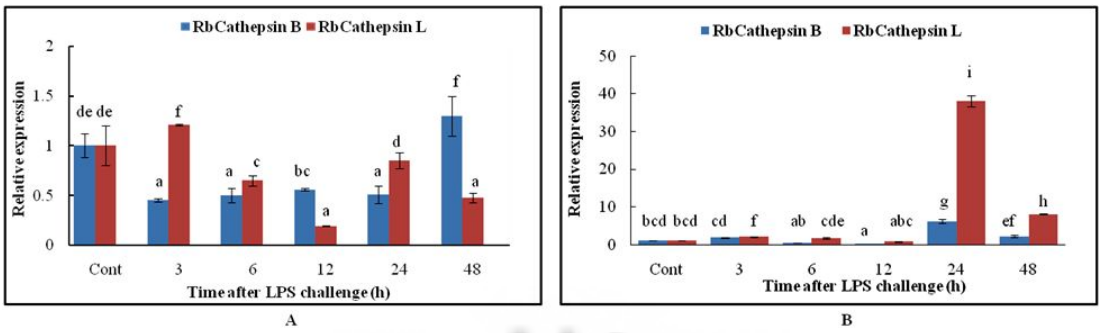



**Figure 4.5 Transcriptional regulation of cathepsin B and cathepsin L of rock bream in liver (A) and blood (B) following LPS challenge.** Transcriptional analysis was carried out by quantitative real-time RT-PCR. The relative level of expression at each time point was compared to that of PBS-injected control (Cont).

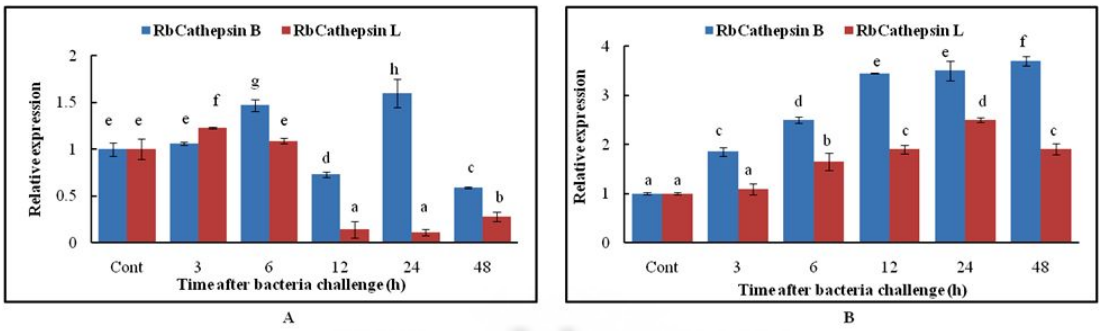



**Figure 4.6 Transcriptional regulation of cathepsin B and cathepsin L of rock bream in liver (A) and blood (B) following *E. tarda* challenge.** Transcriptional analysis was carried out by quantitative real-time RT-PCR. The relative level of expression at each time point was compared to that of PBS-injected control (Cont).

## 4.7 DISCUSSION

In order to obtain the transcriptome sequences of rock bream, we constructed a multi-tissue normalized cDNA and employed the GS-FLX sequencing technology in this study. We identified two cDNAs of cathepsin isoforms and named as rock bream cathepsin B (RbCathepsinB) and cathepsin L (RbCathepsinL). Jaillon et al., [25] reported that the teleost fish underwent whole genome duplication about 350 million years ago. This event could explain the presence of multiple cathepsin isoforms, like cathepsin B and L, in rock bream. It was reported that cathepsin B, C, H, L and S could be diverged at an early stage in evolution, while some of other cathepsins, such as F, K, W and X, could be due to recent gene duplication events [26,27]. Therefore, present study would be a sound foundation to compare cathepsin sequences and their immune responses at molecular level. The rock bream cathepsin B and L genes were characterized as proteins which encodes 330, 336 amino acids, respectively. The amino acid length of cathepsin B/L was more similar to those from *P. oliveceus* (330/338 aa) and *S. salar* (330/338 aa). Lecaille et al., [7] reported that all papain-like cysteine proteases contain conserved active sites of cysteine, histidine and asparagine residues which play crucial roles in the formation and stabilization of the catalytic site of the activated enzyme. The identified RbCathepsin B and L genes encode structural features typical of the vertebrate cathepsin family, including papain family cysteine protease domain, and eukaryotic thiol proteases cysteine, asparagine, histidine active sites. Additionally, RbCathepsin L gene encodes for EF hand $Ca^{2+}$ binding, cathepsin propeptide inhibitor domains and ERF/WNIN-like motif (E-$X_3$-R-$X_3$-F-$X_2$-N-$X_3$-I-X-N). It has been suggested that the ERF/WNIN-like motif is the principal characteristic feature of cathepsin L [28]. Likewise, the deduced RbCathepsin L presents the ERF/WNIN-like motif which could be used to identify it as a member of the cathepsin L family. Moreover, RbCathepsin B and L appear to show higher amino acid identity values and conserved amino acids with known

cathepsin B and L. Also, the phylogenetic analysis revealed that RbCathepsin B and L are positioned with other respective sequences of fish cathepsins. Based on the molecular characterization, pairwise-multiple alignment and phylogenetic results, we could confirm that the newly identified cathepsin B and L are members of the cathepsin family.

Expression of different cathepsin isoforms in fish tissues have been confirmed, but such information remains unknown in rock bream. Ubiquitously expressed cathepsin B was identified in 11 tissues of un-challenged *Paralichthys olivaceus* [29]. Similarly, we identified cathepsin B and L in all tested tissues of rock bream. Constitutive expression of RbCathepsin B and L suggested that both genes may be involved actively in protease function.

Cathepsin expression is known to be induced by several stimulators, such as LPS, virus, poly I:C [29], bacteria Vibrio harveyi [30], tumor promoters [31] and growth factors [32].As a part of the Gram-negative bacteria cell wall, the LPS endotoxin induced the expression of large number of genes related to inflammation, cytokine activity, antigen presentation and binding [33]. Nair et al. [34] reported the induction of cathepsin B and L expression level in sea urchin coelomocytes in response to LPS challenge. In Japanese flounder, cathepsin B has shown moderately up-regulated transcripts in flounder embryonic cells [29]. We observed higher induction of RbCathepsin L in blood than liver after LPS challenge; however, transcriptional induction has not occurred continuously throughout the 48 h post LPS-challenge. A similar situation was explained previously for Theromyzon tessulatum cathepsin L. It was suggested that the protease could be regulated at translational or post-translational stage and therefore was not detectable by northern-blot or RT-PCR experiments [35]. It was reported that oral immunization with *E. trada* ghost can be increased the serum and mucas antibody titers in olive flounder to support the immune activation [18].The present study also showed the higher induction of RbCathepsin B and L in blood than liver against *E. tarda* challenge. Gradual induction of both RbCathepsin B and L was identified in blood

during first 24 h post challenge of E. tarda which could be due to increase cell density of *E. tarda* with the time. Up- and down-regulation of RbCathepsin B and L especially in liver might be due to varied pathogenicity levels during the infection. Darawiroj et al., [36] has described the possibility of transcriptional variation of immune and stress protein against same pathogen or LPS and ConA like lymphocytes mitogens in different fish or tissues. Therefore, transcriptional variation of RbCathepsin B and L in due to LPS and *E. tarda* could be further determined by conducting tissue specific and dose dependant immune challenge.

In conclusion, we identified cathepsin B and L cDNA sequences from rock bream for the first time by applying a pyrosequencing technique. Both cathepsins are constitutively expressed in various tissues, suggesting they may be involved in a multifunctional role in healthy fish. Moreover, cathepsin B and L genes were found to be potentially inducible when exposure to LPS and *E. tarda* occurred, indicating a potential role in the rock bream immune defense system. At present, we are developing a bacterial expression system using *E. coli* to overproduce recombinant RbCathepsin B and L that will facilitate further study of the functional role for each of these enzymes.

## 4.8 REFERENCES


[1] Bond JS, Buter PE. Intracellular proteases. Annu Rev Biochem 1987;56:333-64.

[2] Hsing LC, Rudensky AY. The lysosomal cysteine proteases in MHC class II antigen presentation. Immunol Rev 2005;207:229-41.

[3] Yasothornsrikul S, Greenbaum D, Medzihradszky, KF, Toneff T, Bundey R, Miller R, et al. Cathepsin L in secretory vesicles functions as a prohormone-processing enzyme for production of the enkephalin peptide neurotransmitter. Proc Natl Acad Sci USA 2003;100:9590-5.

[4] Dixit AK, Dixi PJ, Sharma RL. Immunodiagnostic/protective role of Cathepsin L cysteine proteases secreted by *Fasciola species*. Vet Parasitol 2008;154:177-84.

[5] Chwieralski CE, Welte T, Buhling F. Cathepsin-regulated apoptosis. Apoptosis 2006;11:143-9.

[6] Rawlings ND, Barrett AJ. Families of serine peptidases. Methods Enzymol 1994;244:19-61.

[7] Lacaille F, Kaleta J, Bromme D. Human and parasite papain-like cysteine proteases: their role in physiology and pathology and recent developments in inhibitor design. Chem Rev 2002;102:4459-88.

[8] Dong Z, Katar M, Linebaugh BE, Sloane BF, Berk RS. Expression of cathepsin B, D and L in mouse corneas infected with *Pseudomonas aeruginosa*. Eur J Biochem 2001;268:6408-16.

[9] Gocheva V, Zeng W, Ke D, Klimstra D, Reinheckel T, Peters C, et al. Distinct roles for cysteine cathepsin genes in multistage tumorigenenesis. Genes Dev 2006;20:543-56.

[10] Gal S, Gottesman MM. The major excreted protein of transformed fibroblast is an activable acid protease. J Biol Chem 1986;262:1760-5.

[11] Mason RW, Johnson DA, Barrett AJ, Chapman HA. Elastinolytic activity of human cathepsin L. Biochem J 1986;233:925-7.

[12] Tryselius Y, Hultmark D, Cysteine protease 1 (CP1), a cathepsin L-like enzyme

expressed in the *Drosophila melanogaster* hemocytes cell line mbn-2. Insect Mol Biol 1997;6:173-81.

[13] Watzke J, Schirmer K, Scholz S. bacterial lipopolysaccharides induce genes involved in the innate immune response in embryos of the zebrafish (*Daniorerio*). Fish Shellfish Immunol 2007;23:901-5.

[14]] Nya EJ, Austin B. Use of bacterial lipopolysaccharide (LPS) as an immunostimulant for the control of *Aeromonas hydrophila* infections in rainbow trout *Oncorhynchus mykiss* (Walbaum). J App Microbiol 2010;108:686-94.

[15] Park SI. Disease control in Korean Aquaculture. Fish Pathol 2009;44:19-23.

[16] Castro N, Toranzo AE, Nunez S, Magarinos B. Development of an effective *Edwardsiella tarda* vaccine for cultured turbot (*Scophthalmus maximus*). Fish Shellfish Immunol 2008;25:208-12.

[17] Kwon SR, Nam YK, Kim SK, Kim KH. Protection of tilapia (*Oreochromic mosambicus*) from edwardsiellosis by vaccination with *Edwardsiella tarda* ghosts. Fish Shellfish Immunol 2006;20:621-6.

[18] Kwon SR, Lee EH, Nam YK, Kim SK, Kim KH. Efficacy of oral immunization with *Edwardsiella tarda* ghosts against edwardsiellosis in olive flounder (*Paralichthys olivaceus*). Aquaculture 2007;269:84-8.

[19] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic alignment search tool. J Mol Biol 1990;215:403-10.

[20] Bairoch A, Bucher P, Hofmann K. The PROSITE data base, its status in 1997. Nucleic Acids Res 1997;25:217-21.

[21] Pearson WR. Rapid and sensitive sequence comparison with FASTP and FASTA. Methods Enzymol 1990;183:63-98.

[22] Kumar S, Tamura K, Nei M. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform 2004;5:150-63.

[23] Thompson JD, Higgins DG, Gibson TJ. CLUSTALW: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acid Res 1994;22:4673-80.

[24] Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real time quantitative PCR and the $2^{CT}$ method. Methods 2001;25:402-8.

[25] Jaillon O. Aury JM, Brunet F, Petit JL et al., Genome duplication in the teleost fish. Tetraodon nigroviridis reveals the early vertebrate proto-karyotype. Nature 2004;431:946-57.

[26] Berti PJ, Storer AC. Alignment/Phylogeny of the papain superfamily of cysteine proteases. J Mol Biol 1995;246:273-83.

[27] Wex Th, Levy B, Wex H, Bromme D. Human cathepsins F and W: A new sub group of cathepsins. Biochem Biophys Res Commun 1999;259:401-07.

[28] Karrer KM, Peiffer SL, Ditomas ME. Two distinct gene subfamilies within the family of cystene protease genes. Proc Natl Acad USA 1993;90:3063-7.

[29] Zhang FT, Zhang YB, Chen YD, Zhu R, Dong CW, Li YY, Zhang QY, Gui JF. Expressional induction of Paralichthys olivaceus cathepsin B gene in response to virus, poly I:C and lipopolysaccharide. Fish Shellfish Immunol 2008;25:542-9.

[30] Jia A, Zhang XH. Molecular cloning, characterization and expression analysis of cathepsin D from turbot *Scophthalmus maximus*. Fish Shellfish Immunol 2009;26:606-13.

[31] Gottesman MM, Sobel ME. Tumor promoters and kirsten sarcoma virus increase synthesis of a secreted glycoprotein by regulating levels of translatabale mRNA. Cell 1980;19:449-55.

[32] Scher CD, Dick RL, Whipple AP, Locatell KL. Identification of a BALB/c-3T3 cell protein modulated by platelet-derived growth factor. Mol Cell Biol 1983;3:70-81.

[33] Trent MS, Stead CM, Tran AX, Hankins JV. Diversity of endotoxin and its impact on pathogenesis. J Endotoxin Res 2006;12:2005-23.

[34] Nair SV, Del Valle H, Gross PS, Terwilliger DP, Smith LC. Microarray analysis of coelomocyte gene expression in response to LPS in the sea urchin. Identification of unexpected immune diversity in an invertebrate. Physiol Genomics 2005;22:33-47.

[35] Lefebvre C, Vandenbulcke F, Bocquet B, Tasiemski A, Desmons A, Verstraete M, Salzet M, Cocquerelle C. Cathepsin L and B gene expression discriminates

immune coelomic cells in the leach *Theromyzon tessulatum*. Dev Comp Immunol 2008;32:795-807.

[36] Darawiroj D, Kondo H, Hirono I, Aoki T. Immune-related gene expression profiling of yellowtail (*Seriola quinqueradiata*) kidney cells stimulated with ConA and LPS using microarray analysis. Fish Shellfish Immunol 2008;24:260-6.

# 초 록

면역체계는 외부감염으로부터 자신을 방어하는 것이다. 선천성 면역이란 어떤 특정한 병원체에 대항하는 것이 아니며, 일반적으로 재감염에 대한 방어능력이 없으며 비특이적으로 작용하는 1차 방어 작용이다. 반면 후천성 면역은 외부항원에 대해 특이적으로 인식하고 그에 맞는 특이적인 항체를 다양하게 만들어내는 림프구에 의해 일어난다.

하등척추동물인 어류는 다른 고등척추동물보다 선천성 면역체계가 1차적으로 매우 중요한 면역체계이다. 비록 특이적인 병원체을 인식하는 능력이 부족하지만 병원체에 대항하는 능력은 아주 강하다. 이는 어류가 처해 있는 다양한 환경에서 특정병원체에 상관없이 면역인자들이 이미 생체 내에 존재하여 감염발생 시 언제든지 바로 작용할 준비가 되어 있음을 말한다. 또한 어류의 선천성 면역반응은 후천성 면역반응의 활성화를 위해서도 중요하다. 어류의 선천성과 후천성 면역반응의 상호 관계가 비록 포유동물보다는 잘 연구되지 않았지만, 어류의 선천성 면역인자(식세포 자극, 사이토카인과 케모카인의 생성, 보체시스템 활성 등)가 T세포와 B 세포, 그리고 항원제시 세포들을 활성화한다고 알려져 있다.

차세대 염기서열분석 기술은 전사체 분석을 수행함에 있어 동시에 방대한 양의 유전정보를 매우 효율적으로 분석할 수 있게 되었다. 돌돔은 한국의 양식산업에 있어서 경제적으로 중요한 가치를 갖는 해산어류지만, Vibriosis, Edwardsiellosis, Scuticociellosis 그리고 iridovirus 감염 등과 같은 질병으로 인한 피해가 빈번하다. 따라서 돌돔의 세균성/바이러스성 질병 원인체에 대한 면역학적/전사체학적 연구가 절실히 필요하다. 그러므로 돌돔의 광범위한 유전정보를 제공하기 위한 돌돔의 전사체학 분석이 선행되어야 한다.

건강한 돌돔의 다양한 조직으로부터 분리된 mRNA pool을 이용하여 cDNA를 합성하였고, 이를 Roche 454 pyrosequencing platform을 이용하여 대량 염기서열 분석을 수행하였다. 단일 GS-FLX sequencing 반응을 통하여 672,000 reads (평균적인 read의 길이는 400 bp) 를 분석하였고, 이를 통하여 약 36,000개의 contig를 얻을 수 있었다. Pyrosequencing 결과는 다양한 contig 길이 (95-6175 bp)를 나타내었고, contig의 평균 길이는 약 1.1 kb를 나타내었다.

현재 수행된 연구의 최종적인 목표는 숙주와 병원체간의 상호작용에 대한 더 나은 이해를 위하여, 돌돔에서 면역기능에 관여하는 유전자들의 전체 서열을 분석하는 것이다. 이러한 목표 하에서, 이 논문은 돌돔의 goose-type lysozume (g-type lysozyme), myeloid differentiation factor 88 (MyD88), interferon (IFN) regulatory factor-1 (IRF), IFN gamma inducible lysosomal thiol reductase (GILT), Cathepsin B와 L cysteine protease같은 면역 관련 유전자들의 분자적 특성, 전사체적, 기능적 특성을 분석하는데 중점을 두고 있다.

**1) G-type lysozyme의 특성분석과 발현 그리고 재조합 단백질의 항균활성**

Lysozyme (muramidase)은 어류 선천성 면역시스템의 중요한 방어 분자이다. 박테리아 살균능력이 있다고 알려진, lysozyme은 박테리아 세포벽의 peptidoglycan 층에 있는 N-acetyl glucosamine과 N-acetyl muramic acid 사이의 β-(1,4)-glycosidic 결합들의 가수분해를 촉진시킨다. 이 연구에서는 돌돔으로부터 전체 길이 669 bp중, 188개의 아미노산을 암호화 하는 567 bp의 open reading frame으로 이루어진 g-type lysozyme의 전체 coding 서열(RBgLyz)의 특성분석을 수행하였다. 단백질 motif 검색결과 RBgLyz는 세포벽의 균형적인 보존에 관여하는 soluble lytic transglycosylase domain을 포함하고 있었다. RBgLyz는 Chinese perch (*Siniperca chuatsi*)와 매우 높은 유사도(identity, 81.4%)를 나타내었다. 건강한

돌돔에서의 RBgLyz 전사 발현은 다양한 조직에서 일정하게 이뤄지고 있음을 quantitative real-time RT-PCR (qRT-PCR)분석을 통하여 확인할 수 있었다. LPS, poly I:C, *Edwardsiella tarda, Streptococcus iniae*, rock bream iridovirus (RBIV)등의 다양한 병원체를 이용한 면역 자극실험을 통하여 head kidney에서의 RBgLyz 발현을 분석하였다. LPS와 *E. tarda*를 주사한 실험구의 어류들은 면역자극을 주지 않은 대조구 어류와 비교했을 때, 면역자극물질에 반응하여 RBgLyz transcripts의 확연한 발현 증가가 관찰되었다. Poly I:C를 처리하였을 때도 *S. iniae* 보다는 변화의 폭이 작았지만 조금 증가하였다. 하지만, RBIV 을 감염시켰을 때 감염시간에 따른 RBgLyz mRNA의 변화는 관찰되지 않았다. 이러한 결과들을 종합해보면, 돌돔의 면역 자극 실험을 통한 유전자 발현 결과는 LPS, poly I:C, *E. tarda, S. iniae*에 대응하기 위한 선천면역 반응에서 g-type lysozyme의 역할을 보여주었다. 추가적으로, 대장균 발현 시스템을 이용하여 재조합 RBgLyz를 발현시키고, 발현된 RBgLyz재조합 단백질의 항균활성을 분석하였다. 그 결과 재조합 RBgLyz단백질은 그람-음성균인 *Vibrio salmonicida*와 그람-양성균인 *Listeria monocytogenes*, *S. iniae*, *Micrococcus lysodeikticus*에 대하여 lytic activity을 갖고 있음을 나타내었다. 또한 주사전자현미경에 의한 관찰을 통하여 재조합 RBgLyz단백질에 의한 *M. lysodeikticus*의 세포형태 변화를 확인할 수 있었다.

**2) MyD88의 특성과 발현 분석**

Myeloid differentiation factor 88 (MyD88)은 Toll-like receptor3 (TLR3)를 제외한 Toll-like receptors (TLRs)들과 interleukin-1 receptor (IL-1R)의 상호작용을 통하여 nuclear factor-kappa B (NF-kB)를 활성화시키는 universal adaptor protein이다. 돌돔 MyD88 cDNA는 288개의 아미노산을 암호화하는 867 bp의 open reading frame을 포함하는 1626 bp의 염기서열로 구성되어 있었다. MyD88의 아미노산 서열은 다른 어류, 양서류, 조류와 포유류 및 무척추 동물에 이르기까지 대부분의 생물들과 유사한 아미노말단의 conserved death domain과 카르복실말단에 존재하는

typical Toll-Il-1 receptor (TIR) domain을 모두 포함하고 있었다. 건강한 돌돔과 박테리아 감염 돌돔에서의 MyD88의 mRNA 발현 양상을 qRT-PCR을 통하여 분석하였다. Myd88 transcripts는 혈액과, 아가미, 간, 비장, 전췌장(head kidney) 그리고 췌장에서는 매우 강하게 발현되는 한편, 피부조직, 뇌, 내장에서는 중간정도의, 그리고 근육에서는 약하게 발현됨을 알 수 있었다. 혈액과, 비장, 그리고 전췌장에서의 MyD88의 발현수준은 LPS와 *E. tarda*에 노출 되었을 때, 극적으로 증가하였다. 이는 MyD88이 박테리아의 감염에 대응하기 위한 방어반응에 있어서 중요한 역할을 한다는 것을 제시한다.

### 3) IRF-1과 GILT의 전사반응과 특성분석

인터페론(IFN) system과 그 하위 대사의 경로에 존재하는 IFN-stimulated genes (ISG)의 활성은 병원체에 대한 선천성면역과 후천성면역 반응에 중요한 역할을 수행한다. 돌돔 IRF-1 cDNA sequence (RbIRF-1)는 보존된 5개의 tryptophan 반복부위를 포함하는 DNA binding domain (DBD)으로 이루어진 아미노말단의 113개 아미노산이 진화적으로 매우 잘 보존되어 있다. 한편, 돌돔 GILT cDNA (RbGILT cDNA) 서열은 functional domain ($^{97}$CQHGEQECLGNMIETC$^{112}$), active site $^{74}$C-XX-C$^{77}$motif와 7개의 disulfide 결합으로 구성된 GILT signature 서열을 갖고 있다. 이 연구에서, 분리된 돌돔 IRF-1과 GILT 단백질의 분자적 특성분석과 계통분류학적 분석의 결과를 통하여 다른 어류나 포유류의 IRF-1과 GILT 단백질들과 유사함을 확인할 수 있다. qRT-PCR 분석은 RbIRF-1과 RbGILT transcript의 발현은 건강한 돌돔의 실험 조직에서 지속적인 발현을 나타내었으며, 특히 혈액과 아가미에서 가장 높은 수준의 발현이 관찰되었다. Poly I:C를 처리한 돌돔의 혈액, 아가미, 비장, 그리고 전췌장에서의 RbIRF-1과 RbGILT mRNA의 수준이 증가하였다. 하지만, RbGILT mRNA 발현 수준의 증가 정도는 RbIRF-1에 비하여 낮은 결과를 보였다. RBIV를 처리한 돌돔 치어를 48시간 동안 관찰한 결과 혈액에서는 RbIRF-1과 RbGILT 모두의 transcripts가 지속적으로 증가함이

관찰되었지만, 아가미와 전췌장에서는 poly I:C를 처리한 돌돔에 비하여 RbIRF-1과 RbGILT mRNA 수준 증가의 폭이 작았다. Poly I:C와 RBIV에 대한 반응에 의한 전사량 증가는 RbIRF-1과 RbGILT가 IFN 신호전달과 연관되어 있으며, 연속적으로 후천성 면역을 활성화 하는 데에도 중요한 역할을 수행하고 있음을 제시하고 있다.

**4) Cathepsin B와 L cysteine protease의 분자적 특성과 발현 분석**

Cathepsin은 papain의 lysosomal cysteine protease로 세포 내 단백질 분해에 중요한 역할을 수행한다. 돌돔 cathepsin B (RbCahepsin B)는 330 개의 아미노산으로 이루어져있으며, 36 kDa의 분자량을 갖고 있다. 돌돔 cathepsin L (Rbcathepsin L)은 336 개의 아미노산으로 이루어져있으며, 38 kDa의 분자량을 갖고 있다. 서열분석결과는 cathepsin B와 L 모두 papain family cysteine protease signature와 eukaryotic thiol protease cysteine, asparagine, histidine의 active site를 포함하고 있다. RbCathepsin L은 EF hand $Ca^{2+}$ binding과 cathepsin propeptide inhibitor domain을 포함하고 있다. 돌돔의 cathepsin B와 L은 *Lutjanus argentimaculatus* cathepsin B 와 *Lates calcarifer* cathepsin L과 각각 90%와 95%의 유사도(identity)를 나타내었다. 계통분류학적 분석에서 cathepsin B와 L은 각각의 cathepsin family member들과의 높은 진화적 연관성을 나타내었다. qRT-PCR분석을 통하여 자극을 주지 않은 돌돔에서의 cathepsin B와 L 유전자 발현이 모든 실험 조직에서 지속적인 발현을 나타냄을 확인하였다. LPS와 *E. tarda*을 처리한 돌돔의 간과 혈액에서의 RbCathepsin B와 L mRNA는 확연한 발현증가를 나타내었고, 이는 돌돔이 박테리아의 감염에 대한 면역반응에 cathepsin B와 L이 관여함을 암시한다.