Isolation of innate immune response genes, expression analysis, polymorphism identification and development of genetic markers for linkage analysis in common carp (*Cyprinus carpio*)

Pawapol Kongchum

Dissertation submitted to the Faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Fisheries and Wildlife Sciences

Eric M. Hallerman
Yniv Palti
Edward J. Smith
Kathleen A. Alexander

December 17, 2010
Blacksburg, Virginia

Keywords: Common carp, cyprinid herpesvirus-3, immune response genes, single nucleotide polymorphisms, genetic markers

Isolation of innate immune response genes, expression analysis, polymorphism identification and development of genetic markers for linkage analysis in common carp (*Cyprinus carpio*)

Pawapol Kongchum

ABSTRACT

Since the late 1990s, common carp and koi production enterprises around the world have suffered enormous losses due to a viral disease caused by cyprinid herpesvirus-3 (CyHV-3). Genetic variation in resistance to CyHV-3 infection was observed in different common carp strains, indicating that disease resistance can be improved by selective breeding. Marker-assisted selection is a breeding strategy that can accelerate genetic gain; however, this approach requires genetic markers and a genetic linkage map. To develop molecular tools for breeding CyHV-3-resistant aquaculture stock, several candidate genes for antiviral innate immune response from common carp were isolated, and single nucleotide polymorphisms (SNPs) were identified. SNP markers for common carp immune response genes were developed for testing their linkage to disease resistance and for generating a genetic linkage map.

Common carp immune response genes were isolated using degenerate primers developed from conserved peptide regions among other fish species for polymerase chain reaction (PCR) amplification. The amplified products were cloned and sequenced. Gene-specific primers were designed based on the isolated carp gene sequences to amplify gene fragments from genomic DNA of three carp strains and koi. The amplified products were cloned and sequenced to identify SNPs. For the genes that are duplicated, locus-specific primers were used for PCR amplification.

SNPs were identified in several genes, including TLR2, TLR3a, TLR3b, TLR4a, TLR4b, TLR7a, TLR7b, TLR9, TLR21, TLR22, MyD88a, MyD88b, TRAF6a, TRAF6b, type I IFN, IL-1β, IL10a and IL10b. Putative SNPs were genotyped in a SNP discovery panel consisting of different common carp strains and koi to evaluate their allele frequencies and in a full-sib family to validate their segregation patterns using the SNaPshot method. Validated SNPs were used to genotype a mapping family. Twenty-three SNPs (19 exonic and 4 intronic SNPs) were informative in a mapping family. Among these genes, polymorphisms in IL10a suggested a possible association with resistant and susceptible phenotypes of CyHV-3-challenged fish. These SNPs will be analyzed with a set of approximately 300 microsatellites to generate a second-generation genetic map and to identify quantitative trait loci (QTLs) affecting resistance to CyHV-3.

Among the common carp genes that were isolated and sequenced, TLR9 is known for its ability to detect viral DNA and requires adaptor molecules MyD88 and TRAF6 for signal transduction. Therefore TLR9, MyD88 and TRAF6 may be important candidate genes for mediating host antiviral response to CyHV-3. To elucidate possible functions of these genes, full-length cDNAs of common carp TLR9, MyD88 and TRAF6 were isolated and tissue-specific mRNA expression was determined. cDNA sequences of MyD88 and TRAF6 revealed that these genes are duplicated. These findings were the first report of MyD88 and TRAF6 duplications in a vertebrate. Protein domain characterization demonstrated that structural characteristics of these genes are conserved and resemble those of other vertebrates, indicating that common carp TLR9, MyD88 and TRAF6 genes may have identical functions with their mammalian orthologs. The mRNA expression of TLR9, MyD88a and b, and TRAF6a and b varied among tissues. Differential expression of the MyD88 and TRAF6 paralogous transcripts were observed in

muscle tissues, suggesting that one paralog has evolved and attained a non-immune function. This genomic information will facilitate further research to better understand the ligand specificity of TLR9 and the role of TLR9, MyD88 and TRAF6 in the common carp immune response.

Acknowledgements

There are numerous people who have contributed to the successful completion of my doctoral degree at Virginia Tech. First, I would like to thank the co-chairs of my graduate advisory committee, Dr. Eric Hallerman and Dr. Yniv Palti. I was very fortunate to have the opportunity to pursue my doctoral study under their supervision. Dr. Hallerman and Dr. Palti have graciously devoted their time to teach professional skills and greatly contributed to the successful completion during my graduate training at Virginia Tech. I am deeply grateful for their guidance and generosity. I especially thank my committee members, Dr. Edward Smith and Dr. Kathleen Alexander for providing valuable advice, guidance and comments during committee meetings and on my research plan and dissertation manuscript. I would like to thank all staff of the USDA National Center for Cool and Cold Water Aquaculture, West Virginia, especially Renee Fincham, Kristy Shewbridge, Roseanna Long, Brian Smith, Meghan Manor, Dr. Beth Cleveland, Dr. Cecilia Castano, Dr. Sixin Liu, Dr. Guangtu Gao, Dr. Roger Vallejo, and Dr. Gregory Weins for their help and guidance, and the center director, Dr. Caird Rexroad, for giving access to the research facility. I also thank Dr. Lior David of the Hebrew University of Jerusalem and his graduate students, Sigal Lutzky and Eran Sandel, for providing samples for my research. I would like to thank Joanne Davis for spending her time to train me on laboratory technique. I am very grateful to Virginia Tech's Department of Fisheries and Wildlife Sciences, the US-Israel Binational Agricultural Research and Development Fund, and the Ministry of Science and Technology, Royal Thai Government for financial support of my study and dissertation research. Finally, I would like to acknowledge Silpakorn University, Thailand for permission for study leave to pursue my Ph.D. degree while employed as a faculty member of the Faculty of Animal Sciences and Agricultural Technology.

Foreword

This dissertation is composed of an introduction, four potentially publishable manuscripts (chapter 1 – 4) and a summary section. Chapter 1 is a research study on the TLR9 gene of rainbow trout as a preliminary research, since rainbow trout is an aquaculture species with a well-studied genome and similar procedures will be used for subsequent study in common carp. Chapters 2, 3 and 4 are research manuscripts on common carp. The first chapter was submitted to the journal *Animal Genetics* and the second and third chapters were submitted to *Fish & Shellfish Immunology*. Chapter 4 presents preliminary results of gene polymorphisms and disease resistance association. The summary section is a synthesis that summarizes results and conclusions of the research project.

Attribution

There are three potentially publishable research manuscripts in this dissertation (Chapter 1 – 3) and the following individuals serve as co-authors on one or more chapters.

Dr. Eric M. Hallerman (Professor, Department of Fisheries and Wildlife Sciences, Virginia Tech) is the major advisor and committee chair and has provided advice on research works presented in this manuscript, and edited the manuscripts for publication.

Dr. Yniv Palti (Research Geneticist, Agricultural Research Service, United States Department of Agriculture, and Adjunct Professor, Department of Fisheries and Wildlife Sciences, Virginia Tech) is the co-chair and has provided supervision on research works conducted at the National Center for Cool and Cold Water Aquaculture and edited the manuscripts for publication.

Dr. Lior David (Professor, Department of Animal Sciences, Hebrew University of Jerusalem, Israel) and **Dr. Gideon Hulata** (Professor, Institute of Animal Science, Agricultural Research Organization, Israel) have provided considerable insight and editorial review on the research manuscripts.

Dr. Caird E. Rexroad, **III** (Research Leader and Director, National Center for Cool and Cold Water Aquaculture) has helped with linkage analysis and is a co-author on Chapter 1.

Table of Contents

List of Tables

List of Figures

INTRODUCTION

Common carp (*Cyprinus carpio* Linnaeus) is an economically important food fish that has the longest history of domestication (Balon 2004) and is cultured worldwide, with a global production of 2,987,433 metric tons in 2008 (www.fao.org). Japanese ornamental carp, or koi, are members of a subgroup of this species that has been selectively bred for ornamental purposes. Over the past few years, common carp and koi production and trade around the world have suffered enormous losses due to the disease caused by a DNA virus, koi herpesvirus (KHV), which was designated a formal name cyprinid herpesvirus-3 (CyHV-3). The International Committee on Taxonomy of Viruses (ICTV) classified CyHV-3 in the family Alloherpesviridae (Davison et al. 2009). The early outbreaks were reported in Europe and Israel in 1997 – 1998, and later outbreaks were reported worldwide soon afterword (Hedrick et al. 2000; Waltzek and Hedrick 2004; Ilouze et al. 2006). CyHV-3 infection in common carp and koi occurs at temperatures of 17 - 27^{o}C and is highly contagious at 22 - 24^{o}C (Hedrick et al. 2000; Perelberg et al. 2003). CyHV-3 can infect crucian carp and goldfish (El-Matbouli and Soliman, in press), but only cause disease in common carp and koi (both are *Cyprinus carpio*). Fish infected with the virus die between 6 - 22 days post infection, leading to 80 -100% mortality (Ilouze et al. 2006). Currently this disease still remains a considerable risk in carp aquaculture (Hartman et al. 2008). Although CyHV-3 infection is highly contagious, resistance to CyHV-3 infection differed among commercial strains and a wild strain of common carp. Shapira et al. (2005) reported that some carp strains/crossbreds are less susceptible to the disease. This finding suggests that genetic components influence susceptibility to CyHV-3 infection and that genetic improvement for resistance to CyHV-3 infection can be achieved by selective breeding.

Genetic variations due to a single-base mutation or point mutation, referred to as Single Nucleotide Polymorphisms (SNPs), have been identified as a susceptibility factor to infectious diseases in human and livestock. SNPs in immune response genes may result in a change in protein structure or characteristics, leading to altered ability for pathogen recognition or signal transduction. The obvious example is a point mutation in the TLR5 gene in human which changes arginine to a stop codon, making the mutant variant unable to mediate flagellin signaling (Hawn et al. 2003). Such studies in fishes are limited; however, the studies of immune response genes in model fish species demonstrated that the structural characteristics and functional protein domains are conserved and resemble those of mammals' genes. These findings suggest that fish immune response genes may have similar functions as the genes of other vertebrates, and therefore, mutations in common carp immune response genes may be responsible for disease susceptibility and can be used as genetic markers to differentiate susceptible and resistant phenotypes.

The innate immune system includes defense mechanisms that activate immediately or within a few hours following infection. Cells of the innate immune system rely on a set of pattern recognition receptors (PRRs) which have a general ability to detect certain molecular structures present in pathogens (Krieg 2002). The conserved molecular patterns are known as pathogen-associated molecular patterns (PAMPs). Toll-like receptors (TLRs) are PRRs of the innate immune system that play important roles in antiviral immune responses in fish (reviews by Workenhe et al. 2010). Isolation and identification of polymorphisms of genes encoding TLRs and TLR-associated molecules in the signaling pathways will be useful for the development of genetic markers that will enable analysis of genetic linkage to disease resistance and the construction of a genetic map.

The main objective of this dissertation research was to identify genetic variations due to a single-base mutation in candidate antiviral immune response genes, particularly TLRs and genes in TLR signaling pathways, and thereby develop genetic markers for selective breeding to improve CyHV-3 resistance in common carp. SNP markers that were genotyped in a mapping family were used for disease association analysis. Genotype data obtained from this research will be analyzed with the genotype data set for microsatellite markers to generate a second-generation genetic linkage map for common carp. A genetic linkage map is an essential tool that enables the localization of quantitative trait loci (QTLs) affecting resistance to the CyHV-3 virus. SNP markers of candidate genes that are linked to CyHV-3 resistance will be used for selecting resistant individuals as broodstock of the next generation. This process, in which a DNA marker is used for indirect selection of a trait of interest, is known as marker-assisted selection (MAS). Integrating MAS with classical selection by selecting individuals based on marker alleles will increase the rate of genetic gain for resistance to CyHV-3. The use of a genetically improved line for CyHV-3 resistance will reduce the risk and allow profitable common carp production.

LITERATURE CITED

Balon, E. K. (2004). "About the oldest domesticates among fishes." Journal of Fish Biology **65**(s1): 1-27.

Davison, A., R. Eberle, et al. (2009). "The order *Herpesvirales*." Archives of Virology **154**(1): 171-177.

El-Matbouli, M. and H. Soliman "Transmission of cyprinid herpesvirus-3 (CyHV-3) from goldfish to naïve common carp by cohabitation." Research in Veterinary Science, **In Press**.

Hartman KH, Yanong RPE, et al. (2008). "Koi herpesvirus (KHV) disease." Fact Sheet VM-149 Institute of Food and Agricultural Sciences, University of Florida; available at http://edis.ifas.ufl.edu/vm113

Hawn, T., A. Verbon, et al. (2003). "A common dominant TLR5 stop codon polymorphism abolishes flagellin signaling and is associated with susceptibility to Legionnaires' disease." The Journal of Experimental Medicine **198**(10): 1563-1572.

Hedrick, R. P., O. Gilad, et al. (2000). "A herpesvirus associated with mass mortality of juvenile and adult koi, a strain of common carp." Journal of Aquatic Animal Health **12**: 44-57.

Ilouze, M., A. Dishon, et al. (2006). "Characterization of a novel virus causing a lethal disease in carp and koi." Microbiology and Molecular Biology Reviews **70**(1): 147-156.

Krieg, A. M. (2002). "CpG motifs in bacterial DNA and their immune effects." Annual Review of Immunology **20**: 709-760.

Perelberg, A., M. Smirnov, et al. (2003). "Epidemiological description of a new viral disease afflicting cultured *Cyprinus carpio* in Israel." The Israeli Journal of Aquaculture – Bamidgeh **55**(1): 5-12.

Shapira, Y., Y. Magen, et al. (2005). "Differential resistance to koi herpes virus (KHV)/carp interstitial nephritis and gill necrosis virus (CNGV) among common carp (Cyprinus carpio L.) strains and crossbreds." Aquaculture **245**: 1-11.

Waltzek, T. B. and R. P. Hedrick (2004). "Koi Herpesvirus Update 2004." available at: http://www.vetmed.ucdavis.edu/whatsnew/article2.cfm?id=1378.

Workenhe, S. T., M. L. Rise, et al. (2010). "The fight between the teleost fish immune response and aquatic viruses." Molecular Immunology **47**(16): 2525-2536.

CHAPTER 1

Single Nucleotide Polymorphism Identification, Mapping and Tissue Expression of the Rainbow Trout TLR9 Gene

ABSTRACT

Six single nucleotide polymorphisms (SNPs) were identified in a genomic DNA sequence covering 2,489 base pairs of the rainbow trout TLR9 gene (rtTLR9). Five SNPs were synonymous mutations and one was a non-synonymous mutation, 1489C/T (Leu497Phe). SNP markers developed from one non-synonymous and two synonymous SNPs were used to map rtTLR9 by two-point linkage analysis. The rtTLR9 gene was mapped to chromosome 16. Gene-specific amplification of cDNA confirmed the expression of the rtTLR9 mRNA in a variety of adult rainbow trout tissues.

KEYWORDS: Toll-like receptor 9, Single nucleotide polymorphisms, Gene mapping, Gene expression

INTRODUCTION

Toll-like receptors (TLRs) are the important pattern-recognition receptors (PRRs) of the innate immune system that recognize pathogen-associated molecular patterns (PAMPs) and activate the immune response against infection. A number of TLR genes have been identified in rainbow trout, including TLR1, TLR2, TLR3, TLR5, TLR7, TLR8, TLR9, TLR20, TLR22 and TLR23 (Tsujita et al. 2004; Roach et al. 2005; Rodriguez et al. 2005; Palti et al. 2006; Rebl et al. 2007; Ortega-Villaizan et al. 2009; Palti et al. 2010a; Palti et al. 2010b).

Mutation in TLR genes may result in structural alteration of gene products or reduction of gene expression efficiency, which can affect their function or efficiency in recognizing pathogens. Therefore, single nucleotide polymorphisms (SNPs) due to substitution, deletion or insertion of one base pair can have an impact on the ability of individuals to respond to pathogenic agents. SNPs in TLR genes were found to be associated with resistance to pathogen infections in human and livestock (Uenishi and Shinkai 2009).

TLR9 is characterized by its ability to recognize unmethylated CpG motifs in a bacterial genome (Hemmi et al. 2000) and CpG motifs present in a DNA virus (Lund et al. 2003). Therefore, TLR9 has a key role in immune reactions against pathogens and is a candidate gene for disease resistance. Recently, the mRNA and gene sequences of the rainbow trout TLR9 (EU627195 and FJ594277) were deposited in GenBank. However, genetic polymorphism and gene location of rtTLR9 have not been reported to date. In this article, I describe the identification of SNPs, mapping, and tissue expression of the rtTLR9 gene.

MATERIALS AND METHODS

TLR9 gene sequencing and SNP identification

Rainbow trout genomic DNA was screened for TLR9 polymorphisms between two double haploid clonal lines, Swanson and OSU, and between the male and female parents of the National Center for Cool and Cold Water Aquaculture (NCCCWA) reference family mapping panel (Rexroad et al. 2008). Specific PCR primers (Table 1) were designed based on the rainbow trout TLR9 sequence available on GenBank (accession no. EU627195) using the program FastPCR (http://www.biocenter.helsinki.fi/bi/Programs/fastpcr.htm). PCR reactions (11 µl total volume) consisted of 1 µl (~10 ng) genomic DNA, 5 µl nuclease-free water, 1.1 µl 10X PCR buffer, 0.8 µl $MgCl_2$ (25mM), 1 µl deoxyribonucleotide triphosphates (dNTPs) (2mM), 1 µl of each primer (10 µM), and 0.1 µl AmpliTaq Gold Polymerase (Applied Biosystems; 5 units/µl). Amplifications were conducted in PTC-200 thermocyclers (MJ Research) with initial denaturation at 94°C for 10 min; 35 cycles of 94 °C for 10 s, 58 °C for 30 s, and 72 °C extension for 30 s; followed by a final extension of 72 °C for 10 min. PCR products were separated by electrophoresis on a 1% agarose gel stained with ethidium bromide. PCR amplicons were cloned using a TA Cloning kit (Invitrogen, Carlsbad, CA, USA), and the transformations were spread on LB agar plates containing X-gal and 100 µg/ml of ampicillin. The plates were incubated at 37°C overnight, then eight to sixteen clones per sample were picked and grown overnight in LB media, followed by overnight growth in MB media with 50µg/ml of kanamycin. Plasmid DNA was extracted using the QIAprep 96 Turbo kit on a Biorobot 8000 (Qiagen, Valencia, CA, USA) according to the manufacturer's instructions. Sequencing reactions (10 µl total volume) consisted of 3 µl plasmid DNA, 1 µl Big Dye Terminator Ready Reaction Mix (Applied

Biosystems, Foster City, CA, USA), 1.75 µl 5x BD buffer (400 mM Tris, pH 9.0; 10 mM $MgCl_2$), 1 µl of M13 forward or M13 reverse primers (8 µM), and 3.25 µl nuclease-free water. Amplification conditions were 95°C for 3 min; followed by 25 cycles of 95°C for 20 s, 50°C for 20 s, and 60°C for 4 min; and a final extension at 60 °C for 5 min. The amplified products were purified using the CleanSeq magnetic beads protocol according to the manufacturer's instructions (Agencourt Bioscience Corp., Beverly, MA, USA) and sequenced on an ABI 3100 Automated Sequencer (Applied Biosystems, Foster City, CA, USA). To identify single nucleotide polymorphisms (SNPs), the program Sequencher (Gene Codes Corporation, Ann Arbor, MI, USA) was used to assemble gene sequences into contigs at a minimum identity of 85%, with a minimum overlap of 20 bases. In heterozygotes, a sequence variation was determined to be a SNP if the minor allele frequency was at least 30%.

SNP genotyping and mapping of TLR9 gene

The single base extension primers for SNP genotyping (Table 2) were designed for using with the SNaPshot Kit (Applied Biosystems, Foster City, CA, USA). The SNaPshot reactions contained 1 µl of SNaPshot multiplex ready mix, 0.5 µl of 5X BD buffer, 0.2 µl of extension primer (10 mM), 2.3 µl of nuclease-free water, and 6 µl of purified PCR product. Thermocycler conditions for SNaPshot reactions were 96°C for 3 min, followed by 25 cycles of 96°C for 10 s, 50°C for 5 s, and 60°C for 30 s. The reactions were then purified with 1 µl of Shrimp Alkaline Phosphatase (SAP) at 37°C for 60 min and then 75°C for 15 min to remove unincorporated ddNTPs. The purified SNaPshot reactions were run on an ABI 3730 Automated Sequencer and the results were analyzed by GeneMapper v. 3.7 (Applied Biosystems, Foster City, CA, USA).

The SNP markers were mapped using two-point linkage analysis for each parent to identify linkage groups. MULTIMAP and JoinMap linkage analysis software packages were

used to generate male- and female-specific maps as previously described (Palti et al. 2006; Rexroad et al. 2008).

Tissue expression

Tissue expression for the rtTLR9 gene was determined by gene-specific amplification of cDNA prepared from RNAs of various tissues of 10-month-old rainbow trout. The tissues examined included intestine, pyloric ceca, stomach, anterior kidney, trunk kidney, spleen, brachial arch, heart, pituitary, brain, liver, white muscle, red muscle, eye, skin, fin, gill, and abdominal adipose tissues (fat). TLR9 specific primers were: forward 5'-AAGAACCTCCCCAATCTGG-3' and reverse 5'-GATGTTCTCCCAGGGGAAA-3', with a product size of 309 bp. Rainbow trout β-actin was used as the control housekeeping gene, and the primers were forward 5'-GCCGGCCGCGACCTCACAGACTAC-3' and reverse 5-CGGCCGTGGTGGTGAAGCTGTAAC-3. PCR reactions (11 µl total volume) included 1 µl (~ 25 ng) of cDNA sample, 5 µl of nuclease-free water, 1.1 µl 10X PCR buffer, 0.8 µl $MgCl_2$ (25mM), 1 µl dNTPs (2mM), 1 µl of each primer (10 µM), and 0.1 µl AmpliTaq Gold Polymerase (5 units/µl). Amplifications were conducted as follows: initial denaturation at 94^{o}C for 10 min; 35 cycles consisting of 94 ^{o}C for 10 s, 58 ^{o}C for 30 s, and 72 ^{o}C extension for 30 s; followed by a final extension of 72 ^{o}C for 10 min. PCR products were separated on a 2% agarose gel and stained with ethidium bromide for visualization of positive amplifications. To verify TLR9 amplification, PCR products from different tissues were cloned with the TA cloning kit (Invitrogen Carlsbad, CA, USA) and sequenced on an ABI-3100 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA).

RESULTS

Single nucleotide polymorphisms in rtTLR9 gene

Six genomic amplicons covering 2,489 base pairs of the rtTLR9 gene (EU627195; 3,570 bp) were analyzed. Six single nucleotide polymorphisms (SNPs) were detected between Swanson and OSU doubled haploid clonal lines and among the parents of the NCCCWA mapping panel. The positions of these SNPs were 1233C/T, 1489C/T, 1576C/T, 1992C/T, 2358C/T, and 3177C/T (Table 3). Five SNPs were synonymous substitutions and one was a non-synonymous substitution which results in amino acid change (Leu497Phe, Figure 1). Five SNPs from four amplicons were informative for mapping (one parent homozygous and the other parent heterozygous).

SNP genotyping and mapping of rtTLR9 gene

Three SNPs (1233C/T, 1489C/T, and 1992C/T) were useful for genotyping with the SNaPshot kit. Markers 1233C/T and 1992C/T were informative in only one family. All three markers were mapped to chromosome 16. Linkage analyses of families 1, 5 and 6 with the 1489C/T SNP revealed that the rtTLR9 gene is closely linked to microsatellite markers in the 35.7 cM – 41.7 cM interval of chromosome 16 (Table 4).

Tissue expression

Expression of the rtTLR9 gene was detected in various tissues, including intestine, white muscle, pyloric ceca, eye, stomach, skin, trunk kidney, fin, anterior kidney, gill, pituitary, spleen, brain, liver, red muscle and fat tissues (Figure 2). No expression was detected in the brachial arch and heart tissues. The nucleotide sequences of PCR products amplified using cDNA

templates from different tissues were 309 bp and had 100% identity with the rtTLR9 sequence EU627195.

DISCUSSION

In recent years, SNPs have been used widely as genetic markers for mapping genes of interest. Here, I analyzed the TLR9 coding region for SNPs and was able to detect six SNPs in this gene. One of them (1489C/T) resulted in a non-synonymous mutation (Leu497Phe). The properties of the leucine and phenylalanine amino acid side chains are similar, and the mutation occurred between LRR8 and LRR9 in a region of the extracellular receptor that is not currently known to have an important function (Figure 3). However, the replacement occurred in the region of the protein that is involved in binding and recognition of pathogen motifs and hence may affect the binding efficiency of the receptor and its pathogen recognition capabilities. The evidence from recent studies in human indicates a significant relationship between non-synonymous SNPs and disease susceptibility; i.e., mutations in TLR8 (rs3764880; Met1Val) and IFNAR1 (19158C/G) were found to be strongly associated with tuberculosis infection and chronic hepatitis B virus infection, respectively (Davila et al. 2008; Zhou et al. 2009). These findings suggest that non-synonymous mutations, especially in coding regions, are more likely to have a significant impact on the immune response. Most of the SNPs identified in this study are synonymous substitutions and such change may not affect the function of the gene. However, synonymous SNPs are still useful for mapping of the gene, and if they are in strong linkage disequilibrium with allelic variants in the promoter or other regulatory segments of the gene that can alter the level and efficiency of gene expression, they can be associated with disease resistance.

Two-point mapping of TLR9 confirmed that this gene is located on chromosome 16 with a high LOD score. This gene is located on chromosome 8 in zebrafish, chromosome 9 in pufferfish and mouse, and chromosome 3 in human (Table 5).

In the tissue expression testing, I observed strong expression in gill, spleen, brain and eye tissues. However, since regular PCR was used for this purpose, I cannot compare or draw a conclusion on transcription levels among different tissues.

LITERATURE CITED

Davila, S., M. L. Hibberd, et al. (2008). "Genetic association and expression studies indicate a role of toll-like receptor 8 in pulmonary tuberculosis." PLoS Genetics **4**(10): e1000218.

Hemmi, H., O. Takeuchi, et al. (2000). "A Toll-like receptor recognizes bacterial DNA." Nature **408**(6813): 740-745.

Lund, J., A. Sato, et al. (2003). "Toll-like receptor 9–mediated recognition of herpes simplex virus-2 by plasmacytoid dendritic cells." The Journal of Experimental Medicine **198**(3): 513-520.

Ortega-Villaizan, M., V. Chico, et al. (2009). "The rainbow trout TLR9 gene and its role in the immune responses elicited by a plasmid encoding the glycoprotein G of the viral haemorrhagic septicaemia rhabdovirus (VHSV)." Molecular Immunology **46**(8-9): 1710-1717.

Palti, Y., S. A. Gahr, et al. (2010a). "Identification, characterization and genetic mapping of TLR7, TLR8a1 and TLR8a2 genes in rainbow trout (*Oncorhynchus mykiss*)." Developmental & Comparative Immunology **34**(2): 219-233.

Palti, Y., M. F. Rodriguez, et al. (2010b). "Identification, characterization and genetic mapping of TLR1 loci in rainbow trout (*Oncorhynchus mykiss*)." Fish & Shellfish Immunology **28**(5-6): 918-926.

Palti, Y., M. F. Rodriguez, et al. (2006). "Mapping of toll-like receptor genes in rainbow trout." Animal Genetics **37**(6): 597-598.

Rebl, A., E. Siegl, et al. (2007). "Characterization of twin toll-like receptors from rainbow trout (*Oncorhynchus mykiss*): Evolutionary relationship and induced expression by *Aeromonas salmonicida salmonicida*." Developmental & Comparative Immunology **31**(5): 499-510.

Rexroad, C. E., III, Y. Palti, et al. (2008). "A second generation genetic map for rainbow trout (*Oncorhynchus mykiss*)." BMC Genetics **9**: 74.

Roach, J. C., G. Glusman, et al. (2005). "The evolution of vertebrate toll-like receptors." Proceedings of the National Academy of Sciences of the United States of America **102**(27): 9577-9582.

Rodriguez, M., G. Wiens, et al. (2005). "Characterization of toll-like receptor 3 gene in rainbow trout (*Oncorhynchus mykiss*)." Immunogenetics **57**(7): 510-519.

Tsujita, T., H. Tsukada, et al. (2004). "Sensing bacterial flagellin by membrane and soluble orthologs of toll-like receptor 5 in rainbow trout (*Onchorhynchus mikiss*)." Journal of Biological Chemistry **279**(47): 48588-48597.

Uenishi, H. and H. Shinkai (2009). "Porcine toll-like receptors: The front line of pathogen monitoring and possible implications for disease resistance." Developmental & Comparative Immunology **33**(3): 353-361.

Zhou, J., D. K. Smith, et al. (2009). "A non-synonymous single nucleotide polymorphism in IFNAR1 affects susceptibility to chronic hepatitis B virus infection." Journal of Viral Hepatitis **16**(1): 45-52.

Table 1 Rainbow trout TLR9-specific primers used for PCR reactions and length of amplicons.

Forward primer (5'-3')	Reverse primer (5'-3')	Length (bp)
TAAGAGCTTTCAGAACCAGAACCG	ACTTAACGCCTCCAGATTGGGGA	544
CATTGCCATACGCACCGGTA	TGGATGATGCCAGGGCTCAT	472
AAGAACCTCCCCAATCTGG	GATGTTCTCCCAGGGGAAA	309
AACTGCTACTATGGTAACCC	GATGAGGCTGATCCAAGTGAGC	504
AAGGTGCTCCGGGTCAGCAA	CAGCCGTAGAGGTGCCTGAGTA	566
TGTTTCCCAAGCTGAAATACCTGC	GCCAAAATGGGGATTTTATTTGAA	348

Table 2 Sequences of primers (5'-3') used for SNP genotyping.

SNP	PCR primer	SNaPshot primer
1233C/T	F-CATTGCCATACGCACCGGTA R-TGGATGATGCCAGGGCTCAT	GAGAAATTCCAGCATGTTCTGGGAAAGGTC
1489C/T	F-TAAGAGCTTTCAGAACCAGAACCG R-ACTTAACGCCTCCAGATTGGGGA	CATGGCACCATTGCAGAGATTATTTTTGAAGGAAA
1992C/T	F-AAGAACCTCCCCAATCTGG R-GATGTTCTCCCAGGGGAAA	GACTACATCACCTTCTTCCAGAACCTGACCAA

Table 3 Single nucleotide polymorphisms (SNPs) in the rainbow trout TLR9 gene.

SNP Position [a]	Allele		Amino acid substitution [a,b]
	1	2	
1233C/T	C	T	No change
1489C/T	C	T	Leu497Phe
1576C/T	C	T	No change
1992C/T	C	T	No change
2358C/T	C	T	No change
3177C/T	C	T	No change

[a] SNP and amino acid positions are given relative to the A base and ATG start codon as position 1 based on the rtTLR9 sequence (GenBank accession no. EU627195).

[b] The rtTLR9 nucleotide sequence was translated to protein sequence by the ExPASy Translate tool (http://www.expasy.ch/tools/dna.html).

Table 4 Genetic distances between the TLR9 gene (1489C/T marker) and microsatellite markers on chromosome 16 of the rainbow trout genetic map.

Marker	Distance from point zero (cM)	LOD score
OMM1264	35.7	22.62
OMM1308	41.1	22.32
OMM1583	41.7	22.02
OMM1325	35.7	20.84
OMM5281	41.7	19.66

Table 5 Comparative chromosome location of the TLR9 gene in rainbow trout, zebrafish, pufferfish, mouse, and human.

Species	Gene location
Rainbow trout (*Oncorhynchus mykiss*)	Chr. 16
Zebrafish (*Danio rerio*) [a]	Chr. 8
Pufferfish (*Tetraodon nigroviridis*) [b]	Chr. 9
Mouse (*Mus musculus*) [a]	Chr. 9
Human (*Homo sapiens*) [a]	Chr. 3

[a] Gene locations are in GenBank (accession numbers NC_007119, NC_000075 and NC_000003 for zebrafish, mouse and human, respectively).

[b] Pufferfish gene location was determined by Blat-search of the Tetraodon Genome Browser (http://www.genoscope.cns.fr/externe/tetranew/entry_ggb.html).

Figure 1 Partial alignment of rtTLR9 amino acid sequences showing amino acid replacement due to nucleotide substitution (1489C/T). TLR9 amino acid sequences were aligned using ClustalW2 (http://www.ebi.ac.uk/Tools/clustalw2/index.html).

```
seq1   NTLAILEDECTQNPTILSFKNNLCNGAMSFDLSQNNILTLNSSLFLGMEKAVCLDLSYNY 540
seq2   NTLAILEDECTQNPTIFSFKNNLCNGAMSFDLSQNNILTLNSSLFLGMEKAVCLDLSYNY 540
       ***************:******************************************
```

Figure 2 Tissue-specific expression of rtTLR9 (a) and β-actin (b) genes in 10-month-old rainbow trout. Lane assignments: 1, intestine; 2, white muscle; 3, pyloric ceca; 4, eye; 5, stomach; 6, skin; 7, trunk kidney; 8, fin; 9, anterior kidney; 10, brachial arch; 11, heart; 12, gill; 13, pituitary; 14, spleen; 15, brain; 16, liver; 17, red muscle; 18, fat; and 19, negative control.

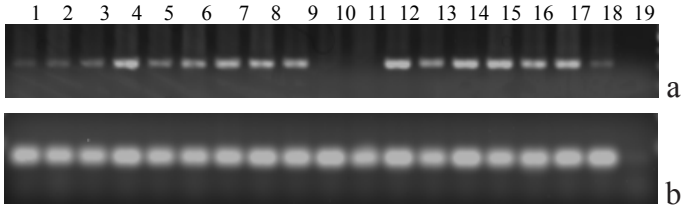
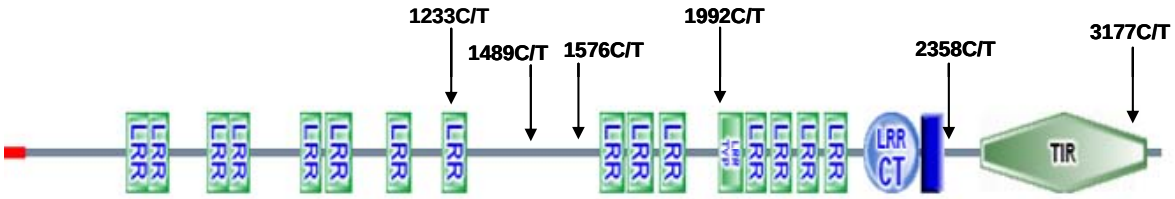


Figure 3 Schematic domain organization of rtTLR9 (EU627195) and SNP positions. The domain organization was predicted using the SMART program analysis tool (http://smart.embl-heidelberg.de/). LRR: leucine-rich repeat; LRR TYP: typical leucine-rich repeat; LRR-CT: C (carboxyl)-terminal leucine-rich repeat; TIR: Toll/IL-1 receptor; (▬) signal peptide; and (▮) transmembrane domain.



CHAPTER 2

SNP Discovery and Development of Genetic Markers for Mapping Innate Immune Response Genes in Common Carp (*Cyprinus carpio*)

ABSTRACT

Single nucleotide polymorphisms (SNPs) in immune response genes have been reported as markers for susceptibility to infectious diseases in human and livestock. A disease caused by cyprinid herpesvirus-3 (CyHV-3) is highly contagious and virulent in common carp (*Cyprinus carpio*). With the aim to develop molecular tools for breeding CyHV-3-resistant carp, I have amplified and sequenced 11 candidate genes for viral disease resistance, including TLR2, TLR3, TLR4ba, TLR7, TLR9, TLR21, TLR22, MyD88, TRAF6, type I IFN and IL-1b. For each gene, I initially cloned and sequenced PCR amplicons from 8 to 12 fish (2-3 fish per strain) from the SNP discovery panel. I then identified and evaluated putative SNPs for their polymorphism in the SNP discovery panel and validated their usefulness for linkage analysis in a full-sib family using the SNaPshot method. My sequencing results and phylogenetic analyses suggested that TLR3, TLR7 and MyD88 genes are duplicated in the common carp genome. I, therefore, developed locus-specific PCR primers and SNP genotyping assays for the duplicated loci. A total of 48 SNP markers were developed from PCR fragments of the 13 loci (7 single-locus and 3 duplicated genes). Thirty-nine markers were polymorphic with estimated minor allele frequencies of more than 0.1. The utility of the SNP markers was evaluated in one full-sib family, which revealed that 20 markers from 9 loci segregated in a disomic and Mendelian pattern and would be useful for linkage analysis.

KEYWORDS: *Cyprinus carpio*, SNPs, Immune response genes, Genetic markers, Disease resistance

INTRODUCTION

The term single nucleotide polymorphism (SNP) describes polymorphisms due to point mutations that create different alleles containing alternative bases at a given nucleotide position within a locus (Liu and Cordes 2004). SNPs in immune response genes can affect the structure of the translated protein, leading to altered ability for pathogen recognition. For example, a mutation in the human Toll-like receptor (TLR) 5 gene (1174C/T) changes arginine to a stop codon which is predicted to prematurely truncate the TLR5 protein in the extracellular domain, rendering the mutant variant unable to mediate flagellin signaling (Hawn et al. 2003). Other studies have shown association between single nucleotide mutations in human TLR signaling pathway genes and susceptibility to infectious diseases or disease progression (Thuong et al. 2007; Leoratti et al. 2008; Wurfel et al. 2008; Prado-Montes de Oca et al. 2009). Although such studies in other species including fishes are limited, there is clear evidence that the immune systems of model fish species such as zebrafish (*Danio rerio*) and pufferfish (*Takifugu rubripes*) resemble those of higher vertebrates (Oshiumi et al. 2003; Meijer et al. 2004). Therefore, SNPs in piscine immune response genes also may influence susceptibility to infection or to result in variation of individual response to pathogenic agents.

The innate immune system, an evolutionarily conserved defense mechanism, is the first line of host defense against microbial infections (Uematsu and Akira 2006). Innate immune response is initiated by Pattern Recognition Receptors (PRRs) which target the conserved molecular patterns generated by microbes (Pathogen-Associated Molecular Patterns; PAMPs) such as lipopolysaccharides (LPS), double-stranded RNA and unmethylated cytocine-guanine (CpG) motifs. TLRs are transmembrane signaling receptors that are expressed in distinct cellular

compartments. In mammals, TLR1, 2, 4, 5, 6, 10 and 11 are expressed on the plasma membrane (cell surface), whereas TLR3, 7, 8 and 9 are expressed in intracellular vesicles such as endosomes and the endoplasmic reticulum (Kumar et al. 2009). They can recognize PAMPs derived from viruses, bacteria, fungi and protozoa. Some of the TLR family members are known for their capability to recognize viral structural components in mammals. TLR9 recognizes unmethylated CpG DNA motifs, TLR7 and TLR8 recognize single-stranded RNA (ssRNA), TLR3 recognizes double-stranded RNA (dsRNA) that can be generated by ssRNA viruses or DNA viruses during replication, and TLR2 and TLR4 can recognize viral-envelope proteins (Akira et al. 2006). The teleost-specific TLR22 recently was shown to recognize long-sized dsRNA, making it another candidate for conferring anti-viral immunity in fishes (Matsuo et al. 2008). TLRs recognize their ligands through interactions with Leucine-Rich Repeats (LRRs) that are found on their extracellular N-terminal domain. TLRs then trigger the activation of intracellular signaling through a cytoplasmic myeloid differentiation primary response protein 88 (MyD88)-dependent pathway or a MyD88-independent pathway. In the MyD88-dependent pathway, MyD88 recruits the interleukin-1 receptor-associated kinases (IRAKs) and TNF receptor-associated factor 6 (TRAF6), which in turn activates downstream genes in TLR signal transduction. Ultimately, the TLR signaling pathways induce production of cytokines including interleukin (IL), tumor necrosis factor (TNF), and type I interferon (IFN) that mediate direct defense responses and alert adaptive immune cells to the presence of a pathogen (Lazarus et al. 2002).

Common carp (*Cyprinus carpio* Linnaeus) is the economically important food fish with the longest history of domestication (Balon 2004) and is cultured worldwide. Japanese ornamental carp, or koi, are members of a subgroup of this species that has been selectively bred

for ornamental purposes. Over the past few years, common carp and koi production enterprises around the world have suffered from enormous losses due to disease outbreaks caused by cyprinid herpesvirus-3 (CyHV-3). This disease remains a considerable risk in carp aquaculture (Hartman et al. 2008). CyHV-3 infection in common carp and koi occurs at temperatures of 17-27 ^{o}C and is highly contagious at 22-24 ^{o}C (Hedrick et al. 2000; Perelberg et al. 2003). Although CyHV-3 is an extremely lethal virus, some carp strains/crossbreds are less susceptible to this disease (Shapira et al. 2005), suggesting that there is a genetic component of susceptibility/resistance to CyHV-3 infection in this species. The long-term research objectives are to investigate the genetic basis of CyHV-3 resistance and develop molecular tools for breeding CyHV-3-resistant common carp. In this study, I cloned and sequenced several TLRs and genes from their signal transduction pathways, and developed SNP markers for genetic mapping and genetic analysis of disease resistance in common carp.

MATERIALS AND METHODS

Samples

The SNP discovery panel was composed of 32 individuals representing two domesticated strains (Našice, n = 10 and Dor-70, n = 6) (Wohlfarth et al. 1975), one strain of Amur wild carp (Sassan, n = 10) that was imported to Israel in 2001 from the Czech Republic (Pokorný et al. 1995; Shapira et al. 2005), and koi carp from the Kibbutz Gan Shmuel Fish Breeding Center in Israel (n = 3) and from a local pet store in Frederick, MD, U.S.A. (n = 3). Twenty-two samples from one full-sib family (2 parents and 20 progeny) of a cross between Dor-70 (♀) x Sassan (♂) were used for validation of SNP markers. Genomic DNA was extracted from blood and fin-clip

tissues using the salting out procedure (Miller et al. 1988; Palti et al. 2002). Tissue sampling was conducted in concordance with the National Center for Cool and Cold Water Aquaculture (NCCCWA) Institutional Animal Care and Use Committee protocol number 043. DNA samples were quantified on a Nanodrop ND-1000 spectrophotometer (Nanodrop Technologies) and the concentrations were adjusted to 20 ng/µl.

Primers and PCR amplification

Specific primers for sequencing and SNP genotyping in TLR2, TLR21, IL-1b and type I IFN genes were designed based on common carp gene sequences retrieved from the GenBank database (accession numbers FJ858800, FJ811455, AJ245635 and AB376667). For TLR3, TLR4, TLR7, TLR9, TLR22, MyD88 and TRAF6 genes for which common carp sequences were not available, degenerate primers were developed from conserved peptide regions among other fish species for gene amplification by polymerase chain reaction (PCR). PCR fragments were subsequently cloned and sequenced to identify carp genes, and gene-specific primers for sequencing and SNP genotyping were designed based on carp sequences. FastPCR software (http://www.biocenter.helsinki.fi/bi/Programs/fastpcr.htm) was used for primer design. The sequences of degenerate and specific primers used are provided in Table 1. The PCR mixture contained 1X PCR buffer, 2 mM $MgCl_2$, 0.2 mM dNTPs, 1 µM of each primer, 0.5 U AmpliTaq Gold (Applied Biosystems) and 20 ng genomic DNA in a final volume of 11 µl. Amplifications were conducted in a PTC-200 thermocycler (MJ Research). The thermal profile consisted of an initial denaturation at 94 ^{o}C for 10 min; 35 cycles of 94 ^{o}C for 10 s, 58 ^{o}C for 30 s, and 72 ^{o}C for 30 s; followed by a final extension of 72 ^{o}C for 10 min. PCR products were visualized on a 2% agarose gel with ethidium bromide.

Cloning, sequencing and SNP identification

Candidate genes that were screened for SNPs included TLR2, 3, 4ba, 7, 9, 21 and 22, and the TLR signaling and response pathways genes MyD88, TRAF6, type I IFN and IL-1b. SNPs were detected by amplifying the candidate genes from carp genomic DNA samples and sequencing the amplification products. For each primer pair, PCR was performed on the SNP discovery panel genomic DNA, and PCR products of 2-3 fish per strain were used for cloning and sequencing. PCR amplicons were cloned into the pCR2.1 vector (TA Cloning Kit, Invitrogen) and transformed into One Shot TOP10 *Escherichia coli* (Invitrogen). The transformations were spread on Luria-Bertani (LB) agar plates containing X-gal and 100 μg/ml of ampicillin and incubated at 37 ^{o}C overnight. A total of 96 insert-positive colonies (8 - 12 transformations per fish from 12 to 8 fish, respectively) were picked and grown overnight in LB broth, followed by overnight growth in Magnificent Broth (MB) with 50 μg/ml of kanamycin. Plasmid DNA was extracted using the QIAprep 96 Turbo Kit on a Biorobot 8000 (Qiagen) according to the manufacturer's instructions. Sequencing reactions (10 μl total volume) were prepared using 3 μl plasmid DNA, 1 μl Big Dye Terminator Ready Reaction Mix (Applied Biosystems), 1.75 μl 5 X Sequencing Buffer (400 mM Tris, pH 9.0; 10 mM $MgCl_2$), 1 μl of M13 forward (50-GTAAAACGACGGCCAG-30) or M13 reverse (50-AACAGCTATGACCATG-30) primers (8 μM), and 3.25 μl nuclease-free water. Cycle conditions were as follows; 95 ^{o}C for 3 min; 25 cycles of 95 ^{o}C for 20 s, 50 ^{o}C for 20 s, 60 ^{o}C for 4 min; 60 ^{o}C for 5 min. The reactions then were purified using the CleanSEQ magnetic beads protocol (Agencourt Bioscience Corporation) and sequenced on an ABI-3100 Automated Sequencer (Applied Biosystems). Sequence assembly software (Sequencher, Gene Codes Corporation) was used to assemble the

DNA sequences into contigs. Single nucleotide polymorphism in DNA sequences was identified as an SNP if the minor allele frequency was at least 25%.

SNP genotyping

The putative SNPs identified by sequencing were genotyped by single nucleotide primer extension assay (SNaPshot) in the SNP discovery panel and in a full-sib family to evaluate their usefulness as genetic markers. The PCR primers and the SNaPshot extension primers used for genotyping are listed in Table 2. PCR reaction mixtures and thermal cycling profiles for preparing templates for SNaPshot reactions were the same as for the pre-sequencing PCR. Following amplification, PCR products were purified using the MinElute 96 UF PCR Purification Kit (Qiagen) to remove primers and unincorporated dNTPs. The SNaPshot reactions were performed using 1 µl of SNaPshot Multiplex Ready Mix (Applied Biosystems), 0.5 µl of 5X Sequencing Buffer, 0.2 µl of extension primer (10 µM), 2.3 µl of nuclease-free water, and 6 µl of purified PCR product. Cycle conditions were 96 ^{o}C for 3 min; and 25 cycles of 96 ^{o}C for 10 s, 50 ^{o}C for 5 s, and 60 ^{o}C for 30 s. The reactions then were purified with 1 µl (1 U) of Shrimp Alkaline Phosphatase (SAP) at 37 ^{o}C for 60 min and 75 ^{o}C for 15 min. SNaPshot products were diluted 1:10 with water, and 1 µl of diluted products were mixed with 9 µl Hi-Di deionized formamide and 0.1 µl LIZ 120 size standard (Applied Biosystems). Samples were denatured by heating at 95 ^{o}C for 5 min, cooled to 4 ^{o}C and run on an ABI 3730 Automated Sequencer. The software GeneMapper version 3.7 (Applied Biosystems) was used to analyze the individual genotypes.

Phylogenetic analyses

Nucleotide sequence haplotypes based on SNPs of the genes described above from each strain of carp and GenBank sequences from other finfishes were used for phylogenetic analyses. Multiple sequence alignments were performed with ClustalW implemented in the MEGA 3.1 software (Kumar et al. 2004). Phylogenetic trees were constructed with MEGA 3.1 by using the neighbor-joining method based on the Jukes and Cantor distance model, with the tree topology reliability tested by 1,000 replicates of bootstrap resampling.

RESULTS AND DISCUSSION

Isolation of common carp immune response genes

I identified the sequences of TLR genes and genes in TLR signaling pathway that have not been previously reported in common carp, including TLR4, TLR7, TLR9, TLR22, MyD88 and TRAF6 genes. I also found that TLR3, TLR7 and MyD88 genes are duplicated in common carp genome. Representative allelic sequences of these gene fragments were deposited in GenBank under accession numbers GU064560, GU321981, GU321982, GU321983, GU321984, GU321985, GU321986, GU321987 and GU459061.

Gene polymorphism and expression

To identify SNPs in these candidate genes and other immune response genes including TLR2, TLR21, IL-1b and type I IFN, I sequenced 26 gene fragments (1 – 4 fragments/gene) in three common carp strains and koi carp. Results revealed that putative SNPs exist at a relatively high frequency. Of 11,446 base pairs sequenced, there was 1 SNP on average in every 160 base

pairs. I identified 70 SNPs in 26 PCR fragments, among which 59 were in coding regions of the genes and 14 were non-synonymous (Table 3).

Common carp has been considered tetraploid because of its chromosome number (2n = 100) and its high DNA content (David et al. 2003). In this study, I observed high frequencies of polymorphic bases in the coding sequences of TLR3 (61 in 763 bp) and TLR7 (67 in 975 bp), suggesting that these genes are duplicated in the carp genome. I then developed locus-specific PCR primers for TLR3 and TLR7 and was able to identify two paralogs for each gene (TLR3a/b and TLR7a/b). The common carp TLR3a and TLR3b fragment sequences had 97% and 87% DNA sequence identity with the GenBank common carp TLR3 gene (accession number DQ885910). The predicted amino acid sequences revealed differences of 23 and 17 residues between TLR3 and TLR7 paralogs, respectively. I performed RT-PCR on cDNA templates prepared from liver, spleen and kidney of Našice carp with TLR3a-, TLR3b-, TLR7a- and TLR7b-specific primers to verify mRNA expression of these duplicated genes and found that TLR3a, TLR3b, TLR7a and TLR7b were expressed in these three tissues (data not shown).

Phylogenetic relationship

To infer the evolutionary relationships of TLR3 paralogs among closely related cyprinids, I conducted phylogenetic analysis based on coding nucleotide sequences. The phylogenetic tree of TLR3 (Figure 1a) showed that the common carp TLR3 paralogs were clustered in two monophyletic branches. The GenBank common carp TLR3 (DQ885910) and TLR3a formed a monophyletic group, whereas carp TLR3b was placed on a separate node and closely related to goldfish and grass carp TLR3, confirming that TLR3 gene is duplicated in the common carp

genome. Similarly, I conducted phylogenetic analysis upon TLR7 (Figure 1b). Again, the topology of the inferred phylogenetic tree suggested that the TLR7 gene also is duplicated.

MyD88 gene is likely duplicated in this species. Amplification by one MyD88 primer pair produced two fragments of 558 bp and 586 bp. Alignment with the zebrafish MyD88 revealed that the two distinct fragments contained introns of 370 bp and 400 bp, respectively. Although the coding sequences from the two MyD88 gene fragments were not long enough for phylogenetic analysis, I was able to design locus-specific PCR primers for each paralog and identified allelic SNPs for each locus which segregated in Mendelian fashion (Table 3). I designated the 558 bp and 586 bp fragments MyD88a and MyD88b, respectively.

A high frequency of duplicated loci is expected in common carp because of the recent genome duplication event that is thought to have occurred in the lineage that led to common carp and goldfish between 11 and 21 MYA (David et al. 2003). My results showed that SNP variation between paralogs was greater than the allelic SNP variation for the three duplicated genes characterized in this study, which is in agreement with previous studies that indicated that the genome duplication was likely caused by hybridization of two genomes (allotetraploidy) (Ohno 1999; David et al. 2003).

In contrast to TLR3 and TLR7, the nucleotide sequences of TLR4ba, 9, 21, 22, TRAF6, IL-1b and type I IFN had fewer polymorphic bases and each fragment was assembled into a single DNA sequence contig. The phylogenetic tree constructed using TLR9 nucleotide sequences (Figure 1c) showed two distinct branches. All carp strains were clustered together on the same branch with their closely related cyprinid, zebrafish. This result suggests that common carp TLR9 paralogous locus may have been lost after the genome duplication event or evolutionarily diverged to the extent that it is not amplified by the gene-specific primer set used

in this study. I also conducted phylogenetic analyses of TLR4ba and TRAF6 sequences (Figure 2 a-b) which showed all common carp sequences clustered closely together. I could not perform phylogenetic analyses for the IL-1β and type I IFN gene fragments, as a large portion of each fragment was composed of intron and the coding sequence was too short (only 200-300 bp) for cross-species comparisons.

Genetic markers

Of a total 70 SNPs detected, I developed 48 SNP markers from exonic and intronic SNPs and genotyped them in the SNP discovery panel. SNPs in coding regions may affect the phenotype as they may cause changes in peptide sequence or gene expression. However, SNPs in noncoding regions of the gene (introns, 5' UTR and 3' UTR) still can be useful genetic markers for linkage disequilibrium analyses. The estimated minor allele frequencies ranged from 0.02 to 0.46, with 39 SNPs having minor allele frequencies larger than 0.1 (Table 3). While the two domesticated strains of European origin (Dor-70 and Našice) were monomorphic for most of the common SNP alleles, the rare alleles mostly were observed in the koi and in some markers also in the Amur wild carp (Sassan strain). This genetic divergence can be attributed to the difference in the geographic origin of the two groups as the Amur carp and koi were originated from Asian common carp (Chistiakov and Voronova 2009).

The usefulness of genetic markers for linkage analysis depends on their polymorphism and segregation pattern in the mapping families. I evaluated the polymorphism and Mendelian inheritance of 44 SNP markers in a full-sib family. Markers were defined as monomorphic if both parents had the same allele; MSV (multisite variants) where I detected heterozygote excess (chi-square test); and PSV (paralogous sequence variants) where both parents and all progeny

were heterozygotes (Fredman et al. 2004; Moen et al. 2008). Twenty-three markers were polymorphic in either one or both of the parents and 21 markers were monomorphic. Of the 23 polymorphic markers, 20 exhibited a disomic and Mendelian segregation, 1 was MSV and 2 were PSVs (Table 3). I was unable to develop useful markers for the type I IFN and IL-1b genes. The SNPs for those two genes that were polymorphic in the panel had either distorted genotypic ratios or fixed allele frequencies in the full-sib family (MSV and PSV markers). I conclude that the polymorphisms observed in the panel for type I IFN and IL-1b likely were caused by PCR amplification of monomorphic duplicated loci or other close members of the same gene family. Additional experiments for cloning and sequencing full-length transcripts are needed to clarify the source of the polymorphisms observed at each of those two gene sequences.

In this study, I evaluated SNP markers for segregation patterns in one full-sib family. Some markers had relatively high minor allele frequencies in the SNP discovery panel, but were not polymorphic in the tested family. Genotyping these SNPs in more families likely would validate their usefulness for genetic analysis. The SNP markers for the common carp innate immune response genes described in this report can be employed for genetic diversity assessment, stock identification, linkage disequilibrium analysis and QTL mapping. Association between MHC class II B gene polymorphism and CyHV-3 resistance in common carp was identified recently (Rakus et al. 2009). My focus on developing genetic markers for common carp innate immune response genes complements that work, as the MHC class II is mostly involved in the adaptive arm of the immune response. In the next phase of my research, I plan to genotype the innate response markers described here in mapping families and to test for genetic linkage with resistance to CyHV-3.

LITERATURE CITED

Akira, S., S. Uematsu, et al. (2006). "Pathogen Recognition and Innate Immunity." Cell **124**(4): 783-801.

Balon, E. K. (2004). "About the oldest domesticates among fishes." Journal of Fish Biology **65**(s1): 1-27.

Chistiakov, D. and N. Voronova (2009). "Genetic evolution and diversity of common carp *Cyprinus carpio* L." Central European Journal of Biology **4**(3): 304-312.

David, L., S. Blum, et al. (2003). "Recent duplication of the common carp (*Cyprinus carpio* L.) genome as revealed by analyses of microsatellite loci." Molecular Biology and Evolution **20**(9): 1425-1434.

David, L., S. Blum, et al. (2003). "Recent duplication of the common carp (*Cyprinus carpio* L.) genome as revealed by analyses of microsatellite loci." Molecular Biology and Evolution **20**(9): 1425 - 1434.

Fredman, D., S. J. White, et al. (2004). "Complex SNP-related sequence variation in segmental genome duplications." Nature Genetics **36**(8): 861-866.

Hartman, K. H., R. P. E. Yanong, et al. (2008). "Koi herpesvirus (KHV) disease, Fact Sheet VM-149." Institute of Food and Agricultural Sciences, University of Florida; available at http://edis.ifas.ufl.edu/vm113.

Hawn, T., A. Verbon, et al. (2003). "A common dominant TLR5 stop codon polymorphism abolishes flagellin signaling and is associated with susceptibility to Legionnaires' disease." The Journal of Experimental Medicine **198**(10): 1563-1572.

Hedrick, R. P., O. Gilad, et al. (2000). "A herpesvirus associated with mass mortality of juvenile and adult koi, a strain of common carp." Journal of Aquatic Animal Health **12**(1): 44-57.

Kumar, H., T. Kawai, et al. (2009). "Toll-like receptors and innate immunity." Biochemical and Biophysical Research Communications **388**(4): 621-625.

Kumar, S., K. Tamura, et al. (2004). "MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment." Briefings in Bioinformatics **5**: 150-163.

Lazarus, R., D. Vercelli, et al. (2002). "Single nucleotide polymorphisms in innate immunity genes: abundant variation and potential role in complex human disease." Immunological Reviews **190**(1): 9-25.

Leoratti, F., L. Farias, et al. (2008). "Variants in the toll-like receptor signaling pathway and clinical outcomes of malaria." The Journal of Infectious Diseases **198**(5): 772-780.

Liu, Z. J. and J. F. Cordes (2004). "DNA marker technologies and their applications in aquaculture genetics." Aquaculture **238**(1-4): 1-37.

Matsuo, A., H. Oshiumi, et al. (2008). "Teleost TLR22 recognizes RNA duplex to induce IFN and protect cells from birnaviruses." The Journal of Immunology **181**(5): 3474-3485.

Meijer, A. H., S. F. G. Krens, et al. (2004). "Expression analysis of the Toll-like receptor and TIR domain adaptor families of zebrafish." Molecular Immunology **40**: 773–783.

Miller, S. A., D. D. Dykes, et al. (1988). "A simple salting out procedure for extracting DNA from human nucleated cells." Nucleic Acids Research **16**(3): 1215.

Moen, T., B. Hayes, et al. (2008). "A linkage map of the Atlantic salmon (*Salmo salar*) based on EST-derived SNP markers." BMC Genomics **9**(1): 223.

Ohno, S. (1999). "Gene duplication and the uniqueness of vertebrate genomes circa 1970-1999." Seminars in Cell and Developmental Biology **10**(5): 517-522.

Oshiumi, H., T. Tsujita, et al. (2003). "Prediction of the prototype of the human Toll-like receptor gene family from the pufferfish, *Fugu rubripes*, genome." Immunogenetics **54**(11): 791-800.

Palti, Y., A. Shirak, et al. (2002). "Detection of genes with deleterious alleles in an inbred line of tilapia (*Oreochromis aureus*)." Aquaculture **206**(3-4): 151-164.

Perelberg, A., M. Smirnov, et al. (2003). "Epidemiological description of a new viral disease afflicting cultured *Cyprinus carpio* in Israel." Israeli Journal of Aquaculture - Bamidgeh **55**: 5-12.

Pokorný, J., M. Flajshans, et al. (1995). "Atlas of common carp populations bred in the Czech Republic.", Victoria Publishing, Praha (in Czech with English summary).

Prado-Montes de Oca, E., J. S. Velarde-Félix, et al. (2009). "SNP 668C (-44) alters a NF-κB1 putative binding site in non-coding strand of human β-defensin 1 (DEFB1) and is associated with lepromatous leprosy." Infection, Genetics and Evolution **9**(4): 617-625.

Rakus, K. L., G. F. Wiegertjes, et al. (2009). "Resistance of common carp (*Cyprinus carpio* L.) to cyprinid herpesvirus-3 is influenced by major histocompatibility (MH) class II B gene polymorphism." Fish & Shellfish Immunology **26**(5): 737-743.

Shapira, Y., Y. Magen, et al. (2005). "Differential resistance to koi herpes virus (KHV)/carp interstitial nephritis and gill necrosis virus (CNGV) among common carp (*Cyprinus carpio* L.) strains and crossbreds." Aquaculture **245**: 1-11.

Thuong, N., T. Hawn, et al. (2007). "A polymorphism in human TLR2 is associated with increased susceptibility to tuberculous meningitis." Genes and Immunity **8**(5): 422-428.

Uematsu, S. and S. Akira (2006). "Toll-like receptors and innate immunity." Journal of Molecular Medicine **84**(9): 712-725.

Wohlfarth, G., R. Moav, et al. (1975). "Genetic variation in seine escapability of the common carp." Aquaculture **5**(4): 375-387.

Wurfel, M., A. Gordon, et al. (2008). "Toll-like receptor 1 polymorphisms affect innate immune responses and outcomes in sepsis." American Journal of Respiratory and Critical Care Medicine **178**: 710-720.

Table 1 A list of degenerate (bold type) and specific primers used for gene amplification

Gene	Forward primer sequence (5'-3')	Reverse primer sequence (5'-3')	Fragment size (bp)
TLR2	CCAGGACATGTGATTGTGACCAGC	GACTAAACCAATGGGACGAGCTGC	530
	AGATGGGCTCACTCATCTGGACG	ACAGTGGGGTAGTTTCCCGTGGA	569
	GCAGTTGGGATTTCTTCTGGTGCA	GCGTTCCTTTGGATAAGCAGTGTC	549
TLR3	**AARATGCARCAYAAYAATTTG** **AARATGCARCAYAAYAACTTG**	**TGRATNCGCCANCCYTGGAA** **TGRATYCTCCANCCYTGGAA**	587
	AATGCCTACTTTAACCGCTCTG	TCGCCTACACCAAGGATCCT	444
TLR4	GGACAAGGACAGGAATGTAC	TCTTTGAAGGTGCTCATAAA	1068
	GAGTACTCATGTTCTGGAAG	TATGTAATAGGGTTTCCAGT TCTTTGAAGGTGCTCATAAA	319 496
	CCTCCATTTATGAGCACCTTCA	**CTATTNCCNGCCATYTTGAG** **CTGTTNCCNGCCATYTTGAG** **GARTTNCCNGCCATYTTAAG** **GARTTNCCNGCCATYTTGAG**	941, 950
	CCACCATACAAAACTAGTCGTA	**ATRAARTGYTKNSWCACCAC** **CCAYTTNARRTANGTGTTCTT** **AGTTCATTCATGACCCAGACTTCA** **GACTGCCCATTATTCCTTCATC**	942 1123 795, 807 911, 923
	GGATGACCAATAAGCCATTTTCCG	GACTGCCCATTATTCCTTCATC	975
TLR7	**AAYTGYCCNAGRTGYTATAAT**	**ATCTTDATRAARTTNGTTCC** **ATCTTDATRAARTTNGTCCC**	458
	GTCACCTTAAATCGCTGAAGGTCC	**GTRAARTTNAGCATGTGAGT**	694

Table 1 (continued)

Gene	Forward primer sequence (5'-3')	Reverse primer sequence (5'-3')	Fragment size (bp)
TLR7	GCACCTTTTCCATGCGTTCC	GCAAAGTTTCTCCTTGAGAAGGCA	532
	ATCTCAATGCCTTCTCAAGGAG	ACCCTCAGCCACAAAGTAATG	747
TLR9	**AAYTGYTAYTAYGGCAACCC**	**RTCRCANCKYTGGCAATTCCA**	249
	GGCAACCCATGTTATCAGTCTTAC	**GGYTCRTARTTRTAAATAAG**	506
	CACTGCATACAATTCCTCAG	**AARTGRAANTCRTTGTTGCT**	867
	TCGCATTGACCTGTACTCTGA	**TCCCANCCRTANARRTGCTTCA**	975
	GATCTTCACTCATGCCAGGA	**GGCCAGCTNAGNACNGATTT**	578
	GGCAACCCATGTTATCAGTCTTAC	AACATTCCGGAGCTCTTCAT	474
	CGACTCCACAAGCTCCAAGAGTTAGATT	CCTTGTCAGAGTACAGGTCAATGC	775
	GACTTTAAATGGCCAGCAGTTCAC	TCAAGGATCTTCAGCTGGTTGTGG	693
	TGAGCTTCGTCTTGTAAGCCTGGA	GAATGCTGGTTGCTGCCAATGTC	443
TLR21	CGTACATCGACATGCGCAGTTGTC	GTGGAGGCACAGGCGAAATGA	440
TLR22	**CARTAYGAYGCNTTCATTTC**	**TGRTGNGTNGGDATGTCCTC** **TGRTGNGTNGGDATATCCTC**	311
	TCCTACAATGCCAAAGATGAGCC	CAGGAACACCAGAATCAGTACATCC	273
MyD88	**GATGCNTTYATHTGCTATTGT**	**AAAYTTNGTYTGRAAATCGCA**	606
	TGACTTCCAGTTTGTGCATGAG	GCATCACTGTCCAGATAGTC	558, 586

Table 1 (continued)

Gene	Forward primer sequence (5'-3')	Reverse primer sequence (5'-3')	Fragment size (bp)
TRAF6	**TTTGTNCACACNATGCAGGG**	**ACATANCCRAANCCYTTGGG**	185
	GGTCGNCTNGTNCGNCAAGA	**AGTGGGCCTCTGGAAKGCCTGC**	471
	AGATCCGGGAGCTGTGCATCC	GCCTCTGGAATGCCTGCAAGTC	441
IL-1β	CTGTGACGCTGAGTGCTGGA	TTCGGGTGGTTGGCATCTGG	540
	TCAACATTCGTGTCGAGTGCAA	CTCAAGTGTGAAGTTTGTGGTTCG	416
Type I IFN	CGCGGGAATTCGATTACAACATTCTCTCG	GCTGTGGTTGTATTGCGCAGCA	284
	TGAAGGTGCCATTTCCAAGGACC	AAGCTCAGATGACTGCCTGC	285
	CCGTCACATATGGAGTCGTACAAG	TCTGAAGGTGGGTTATCACAGCTC	377

Table 2 A list of the SNP marker names, PCR primers and amplicon size, and single-base extension (SBE) primers used for SNP genotyping.

Marker name	PCR primer sequence (5'-3')	Size (bp)	SBE primer sequence (5'-3')
TLR2-SNP59A/G-R	F-CCAGGACATGTGATTGTGACCAGC R-GACTAAACCAATGGGACGAGCTGC	530	GCCAAGCGCATTCGCTGGAACTTTTGGGACCTG
TLR2-SNP126T/G-R	F-CCAGGACATGTGATTGTGACCAGC R-GACTAAACCAATGGGACGAGCTGC	530	GCTCACTGTATGGGCTGAGATCACTCATGT
TLR2-SNP221G/A-F	F-CCAGGACATGTGATTGTGACCAGC R-GACTAAACCAATGGGACGAGCTGC	530	TTATACACAAAGATGCTTTCAAATCTCA
TLR2-SNP484C/T-F	F-CCAGGACATGTGATTGTGACCAGC R-GACTAAACCAATGGGACGAGCTGC	530	AGATGACATTCGTTGGCAGTAATCTGAGGT
TLR2-SNP201T/C-F	F-AGATGGGCTCACTCATCTGGACG R-ACAGTGGGGTAGTTTCCCGTGGA	569	CACCCTGAAACTCTGCTGATAATCAAAGATGT
TLR2-SNP301T/G-R	F-GCAGTTGGGATTTCTTCTGGTGCA R-GCGTTCCTTTGGATAAGCAGTGTC	549	CTGGCCAGCTACAGTTTTCTGGCAT
TLR3a-SNP93G/T-F	F-AAGGATGCCACAAAACTGTCTGTT R-GCAGGAAAGCTGTGAATAACAACATC	493	TAATGCTTTGCGCGGCTTCTTTGAGTTACA
TLR3a-SNP382C/T-F	F-AAGGATGCCACAAAACTGTCTGTT R-GCAGGAAAGCTGTGAATAACAACATC	493	ATGTGTAACACCCCAAATGCCTACTTTAAC
TLR3a-SNP42C/T-F	F-GGGTTCCCTCAAAGTTATATGTGT R-TTATTTTCCTGGATTGTCCCATGT	430	CCTCAAAGTTATATGTGTAACACCCCAAATGCCTA
TLR3a-SNP234T/C-F	F-GGGTTCCCTCAAAGTTATATGTGT R-TTATTTTCCTGGATTGTCCCATGT	430	GTGAACCGTATGCTGGGGTCACTGAAGGA
TLR3b-SNP59T/A-F	F-GGATTCCCTCAAAGTTACATCTGC R-GGATTTTCCTAGACTGTCCCATGT	430	CCCAAATGCCTACTTTAACCGCTCCGTCA
TLR3b-SNP148A/T-F	F-GGATTCCCTCAAAGTTACATCTGC R-GGATTTTCCTAGACTGTCCCATGT	430	GAGTAGCACAGCAGTGTTAATGTTGTTATTT
TLR3b-SNP157C/T-R	F-GGATTCCCTCAAAGTTACATCTGC R-GGATTTTCCTAGACTGTCCCATGT	430	CAGAAGAACTGAATTCTCCATCCCTGGAAGTGTACCA
TLR3b-SNP336A/G-F	F-GGATTCCCTCAAAGTTACATCTGC R-GGATTTTCCTAGACTGTCCCATGT	430	GTGGAACGAAGCTTGCTCCCCTTAGAGGATGA
TLR3b-SNP369A/G-R	F-GGATTCCCTCAAAGTTACATCTGC R-GGATTTTCCTAGACTGTCCCATGT	430	CAATGGTGTCAAGTTGAGAAAAGCCAGGTAT
TLR3b-SNP378C/A-R	F-GGATTCCCTCAAAGTTACATCTGC R-GGATTTTCCTAGACTGTCCCATGT	430	GTTTTCAACAATGGTGTCAAGTTGAGAAAA

Table 2 (continued)

Marker	PCR primer sequence (5'-3')	Size (bp)	SBE primer sequence (5'-3')
TLR4ba-SNP59A/C-F	F-GCCAATTTATATCTTTCGCTGTATGATTAATGCC R-GGAATCAAGTCCATCAAATCCTCCTACATC	385	AATGCCACAAATGTATTAGTAAGGAATGGT
TLR7a-SNP114A/G-R	F-GCACCTTTTCCATGCGTTCC R-CGGTTGGATGTCTTCCTGCT	381	GAGTCAGATTTTGAAACCACTCCGAAGGTATGTG
TLR7a-SNP126G/A-R	F-GCACCTTTTCCATGCGTTCC R-CGGTTGGATGTCTTCCTGCT	381	GACCAAGTTGAGTCAGATTTTGAAACCACTC
TLR7a-SNP348G/A-F	F-GCACCTTTTCCATGCGTTCC R-CGGTTGGATGTCTTCCTGCT	381	GCTGAAGGTCCTAAGAATCAGAGGGTATGT
TLR7b-SNP72G/A-F	F-GCACCTTTTCCATGCGTTCC R-AGGTTGGATGTCTTCCTGTG	381	CTTCAGATACATCCAAACTCCTTTAAGACCCT
TLR9-SNP68T/A-R	F-GGCAACCCATGTTATCAGTCTTAC R-AACATTCCGGAGCTCTTCAT	474	TGGTGGAACAATGGTCACATTGTTGTAACTCAGTGTAAGG
TLR9-SNP190C/T-F	F-CGACTCCACAAGCTCCAAGAGTTAGATT R-CCTTGTCAGAGTACAGGTCAATGC	775	ATCCATAGAGAAGATGACGGCTCTACGAGAG
TLR9-SNP481C/T-F	F-CGACTCCACAAGCTCCAAGAGTTAGATT R-CCTTGTCAGAGTACAGGTCAATGC	775	AGCTCAACTTGCAAACTCAAAAAGTCATCCTA
TLR9-SNP538T/A-R	F-CGACTCCACAAGCTCCAAGAGTTAGATT R-CCTTGTCAGAGTACAGGTCAATGC	775	TTAGAGCAGATCTGCTTTTTGAAATGCCATATAGAAT
TLR9-SNP723G/A-F	F-CGACTCCACAAGCTCCAAGAGTTAGATT R-CCTTGTCAGAGTACAGGTCAATGC	775	GACTTTAAATGGCCAGCAGTTCACACATCTAAGCAA
TLR9-SNP132T/C-F	F-CAGATGCTGAGAATGGACTCT R-TGGCATGAGTGAAGATCTATCG	503	ATGCTTTCATTCTTACCTAATATGCATTTTGAGCT
TLR9-SNP144C/T-R	F-CAGATGCTGAGAATGGACTCT R-TGGCATGAGTGAAGATCTATCG	503	CACAACTAGTCGATTGTGACTTAGGTCCAG
TLR9-SNP372A/T-F	F-CAGATGCTGAGAATGGACTCT R-TGGCATGAGTGAAGATCTATCG	503	GCATGCCAATCCTTTTACTTGTAGTTGTGT
TLR9-SNP460T/C-F	F-CAGATGCTGAGAATGGACTCT R-TGGCATGAGTGAAGATCTATCG	503	ACAGAGGTTCACTGTGGTTATCCAGAGTCA
TLR21-SNP185A/G-F	F-GCAGTTGTCCTCTCAGTTGTGGCTG R-AGAACCCTCGAGGTTCGGCAACAG	404	AGACATAGCACTCTACCTGTTTTCCTCAAC
TLR21-SNP272A/G-F	F-GCAGTTGTCCTCTCAGTTGTGGCTG R-AGAACCCTCGAGGTTCGGCAACAG	404	ACTTTACTGGAAATTTAAGTATGGCTACTATGTGTTTCG
TLR21-SNP365C/T-F	F-GCAGTTGTCCTCTCAGTTGTGGCTG R-AGAACCCTCGAGGTTCGGCAACAG	404	ATGCCTTTGTTTCCTACAACTCAGCAGATGAAGA

Table 2 (continued)

Marker	PCR primer sequence (5'-3')	Size (bp)	SBE primer sequence (5'-3')
MyD88a-SNP142T/G-F	F-TGACTTCCAGTTTGTGCATGAG R-ATGTCCACTATGTAGAATGGCT	281	TCACCAGTGAGCTCATAGAAAAAAGGCAAGT
MyD88a-SNP159T/C-R	F-TGACTTCCAGTTTGTGCATGAG R-ATGTCCACTATGTAGAATGGCT	281	TTAAAGATCAATTTAAATGAGCATTGGATG
MyD88b-SNP83G/A-F	F-TGACTTCCAGTTTGTGCATGAG R-GACTGACACACAATAAATGGCTGA	369	AACCTGAAGCTGTGTGTCTTTGACCGAGAC
MyD88b-SNP282G/A-F	F-TGACTTCCAGTTTGTGCATGAG R-GACTGACACACAATAAATGGCTGA	369	TGAATATCTGTTCTTTATCGTGTACATACT
TRAF6-SNP125C/T-F	F-AGATCCGGGAGCTGTGCATCC R-GCCTCTGGAATGCCTGCAAGTC	441	AGAGCTGGAAGCCCAGCAGTATCAAGGTGTCTA
TRAF6-SNP140G/A-R	F-AGATCCGGGAGCTGTGCATCC R-GCCTCTGGAATGCCTGCAAGTC	441	CCTCTTGGTTGCGCAGGTGAAGTGAGAAGTT
TRAF6-SNP212G/A-F	F-AGATCCGGGAGCTGTGCATCC R-GCCTCTGGAATGCCTGCAAGTC	441	CCATAGTCCTCCACAGCCCACCTTTCTACAC
TRAF6-SNP281T/C-R	F-AGATCCGGGAGCTGTGCATCC R-GCCTCTGGAATGCCTGCAAGTC	441	ACCCTGCATAGTGTGCACGAAAAGTGAGATGTA
TRAF6-SNP380C/T-F	F-AGATCCGGGAGCTGTGCATCC R-GCCTCTGGAATGCCTGCAAGTC	441	CACGATCCGACTGGCAGTGTTGGACCAGGT
IL1β-SNP40A/G-F	F-CTGTGACGCTGAGTGCTGGAGCAATG R-TTCGGGTGGTTGGCATCTGGTT	540	GACGCTGAGTGCTGGAGCAATGCAATACAAAGGTA
IL1β-SNP62G/C-R	F-CTGTGACGCTGAGTGCTGGAGCAATG R-TTCGGGTGGTTGGCATCTGGTT	540	GTCAAAAGGACATTTCGGAAGTCATGCATTAAGATG
IL1β-SNP119T/C-F	F-CTGTGACGCTGAGTGCTGGAGCAATG R-TTCGGGTGGTTGGCATCTGGTT	540	TCCGAAATGTCCTTTTGACGGTTCCCCCTACACATTTTA
IL1β-SNP280C/T-F	F-CTGTGACGCTGAGTGCTGGAGCAATG R-TTCGGGTGGTTGGCATCTGGTT	540	CCTGTCCTGATCTTGAAGGTTTGTGCAGCA
Type I-IFN-SNP27G/A-R	F-AAGATGAACCAAACTCAAATGTGGAC R-CTGATGAACATTTACAAACAAATCATG	188	TCTGCAGAGTTAAAAATATCACAAAAATACA
Type I-IFN-SNP160C/T-F	F-AAGATGAACCAAACTCAAATGTGGAC R-CTGATGAACATTTACAAACAAATCATG	188	GAAATGGTAAGTAGCCTAGACTTTTAAAGC

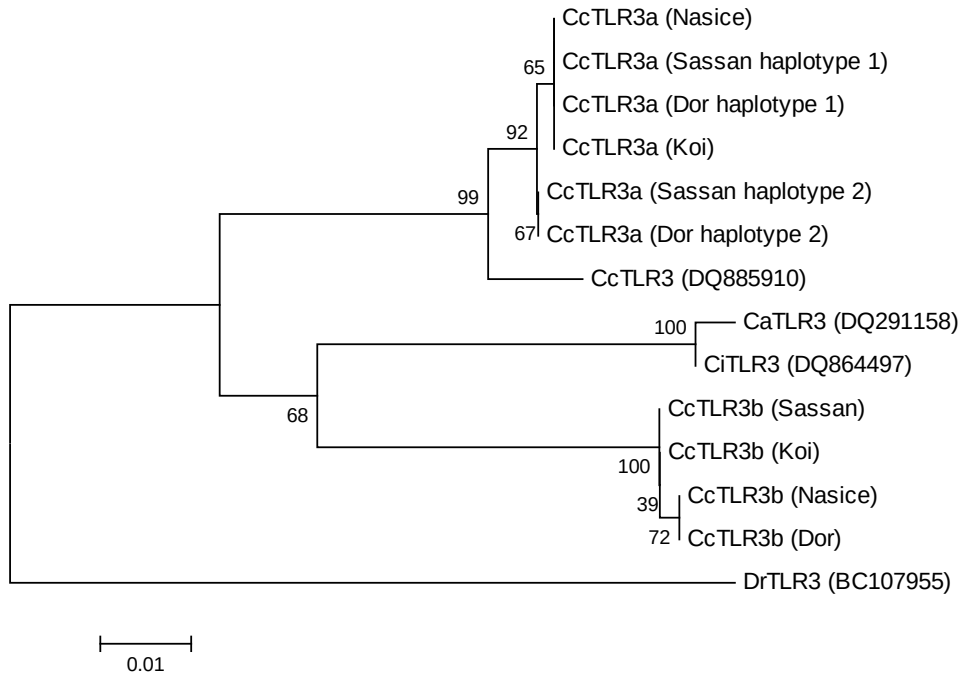
Table 3 Number of single nucleotide polymorphisms (SNP) identified in each gene and polymorphism of SNP markers genotyped.

Gene	Accession number	Base pairs sequenced[1]	SNPs identified[2]	SNPs genotyped[3]	SNP marker validation[4]				
					Normal	MSV	PSV	Homozygous	Total
TLR2	FJ858800	1459 (3)	11 (2)	6 (6)	1	-	-	5	6
TLR3a	DQ885910	763 (2)	4 (2)	4 (1)	-	-	-	4	4
TLR3b	GU321981	763 (2)	6 (2)	6 (6)	2	-	-	4	6
TLR4ba	GU321982	630 (1)	4 (2)	1 (1)	-	-	-	1	1
TLR7a	GU321983	975 (2)	3 (0)	3 (3)	3	-	-	-	3
TLR7b	GU321984	975 (2)	3 (0)	1(1)	1	-	-	-	1
TLR9	GU321985	2077 (4)	14 (2)	9 (7)	3	-	-	3	6
TLR21	FJ811455	440 (1)	5 (2)	3 (3)	3	-	-	-	3
TLR22	GU459061	273 (1)	0	0	-	-	-	-	-
MyD88a	GU321986	558 (1)	2 (0)	2 (1)	2	-	-	-	2
MyD88b	GU321987	586 (1)	3 (1)	2 (1)	2	-	-	-	2
TRAF6	GU064560	441 (1)	5 (0)	5 (4)	3	-	-	1	4
IL-1β	AJ245635	560 (2)	8 (1)	4 (3)	-	1	1	2	4
Type I IFN	AB376667	946 (3)	2 (0)	2 (2)	-	-	1	1	2
Total		11446 (26)	70 (14)	48 (39)	20	1	2	21	44

[1] Number of fragments sequenced is in parentheses.
[2] Number of non-synonymous SNPs identified in coding regions is in parentheses.
[3] Number of SNP markers genotyped in the panel of common carp strains and koi. Number of polymorphic markers with minor allele frequency > 0.1 is in parentheses.
[4] SNP markers evaluated in a family (Dor-70 x Sassan) for polymorphisms and segregation patterns. SNP types are classified according to Fredman et al. (2004) and Moen et al. (2008). Normal = polymorphic and reliably scored single-locus marker. MSV (multiple sequence variant) = likely duplicated with polymorphism at one or both loci. PSV (paralogous sequence variant) = duplicated SNP without homozygotes. Homozygous = all individuals were homozygous.

Figure 1 Phylogenetic trees of common carp TLR3 (a), TLR7 (b) and TLR9 (c) genes constructed by Neighbor-Joining method based on Jukes and Cantor distance model using nucleotide sequences (TLR3, 493bp; TLR7, 975bp; TLR9, 443bp). Sequence haplotypes were obtained from Dor-70, Našice, Sassan and koi carps. GenBank accession numbers for nucleotide sequences obtained from NCBI database are shown in parentheses. The numbers at the branches indicate bootstrap values based on 1,000 replicates. Abbreviations are indicated as follows: Cc, *Cyprinus carpio*; Ca, *Carassius auratus*; Ci, *Ctenopharyngodon idella*; Dr, *Danio rerio*; Om, *Oncorhynchus mykiss*; Ss, *Salmo salar*; Tr, *Takifugu rubripes*.

a.



b.

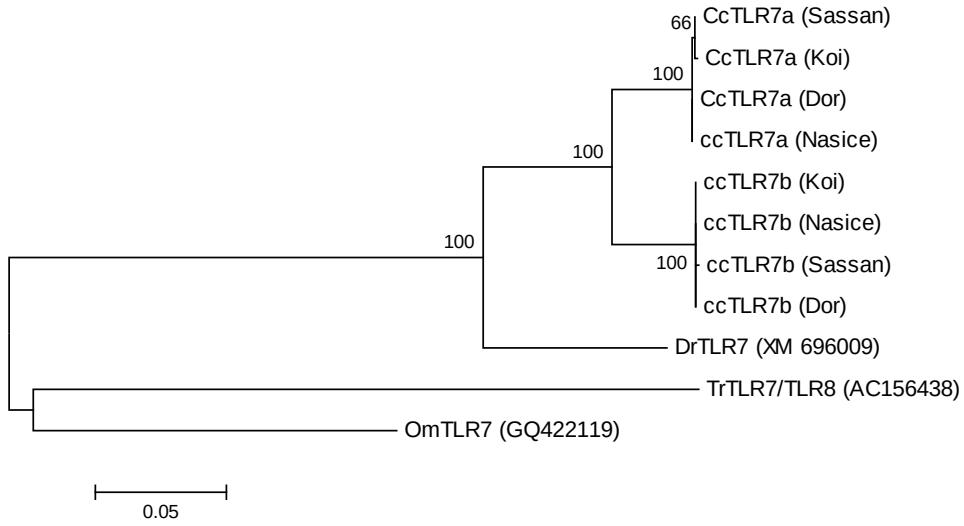


Figure 1 (continued)

c.

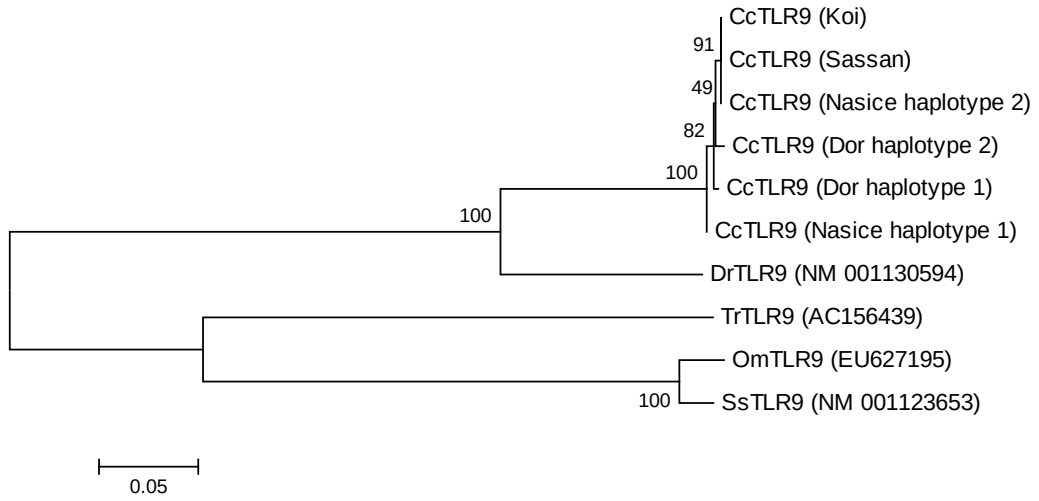
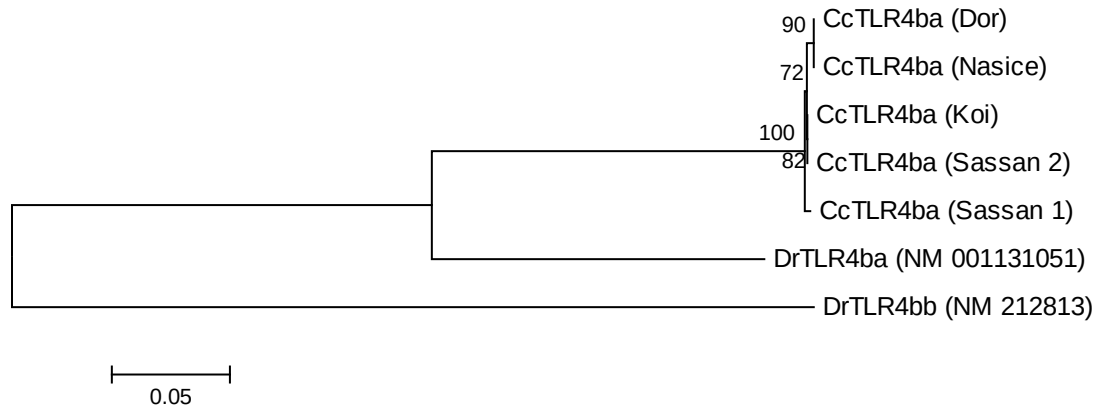
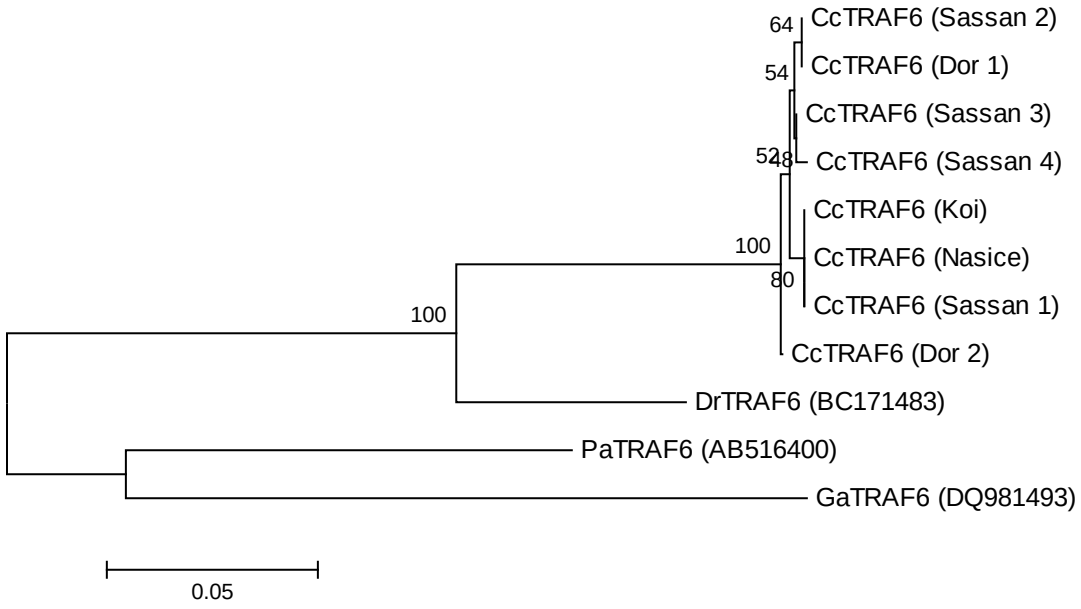


Figure 2 Phylogenetic trees of carp TLR4ba (a) and TRAF6 (b) genes constructed by the Neighbor-Joining method based on the Jukes and Cantor distance model using nucleotide sequences. The fragment lengths used in tree construction were 630 and 441bp, respectively. GenBank accession numbers for gene sequences obtained from NCBI database are shown in parentheses. The numbers at the branches indicate bootstrap values based on 1,000 replicates. Abbreviations are indicated as follows: Cc, *Cyprinus carpio*; Dr, *Danio rerio*; Pa, *Plecoglossus altivelis*; Ga, *Gasterosteus aculeatus*.

a.



b.



CHAPTER 3

Molecular Cloning, Characterization and Expression Analysis of TLR9, MyD88 and TRAF6 Genes in Common Carp (*Cyprinus carpio*)

ABSTRACT

Induction of innate immune pathways is critical for early host defense, but there is limited understanding of how teleost fishes recognize pathogen molecules and activate these pathways. In mammals, cells of the innate immune system detect pathogenic molecular structures using pattern recognition receptors (PRRs). TLR9 functions as a PRR that recognizes CpG motifs in bacterial and viral DNA and requires adaptor molecules MyD88 and TRAF6 for signal transduction. Here, I report full-length cDNA isolation, structural characterization and tissue mRNA expression analysis of the common carp (cc) TLR9, MyD88 and TRAF6 gene orthologs. The ccTLR9 open-reading frame (ORF) is predicted to encode a 1,064-amino acid (aa) protein. MyD88 and TRAF6 genes are duplicated in common carp. This is the first report of TRAF6 duplication in a vertebrate genome, and stronger evidence in support of MyD88 duplication is provided. The ccMyD88a and b ORFs are predicted to encode 288-aa and 284-aa peptides, respectively. They share 91% aa sequence identity between paralogs. The ccTRAF6a and b ORFs are both predicted to encode 543-aa peptides sharing 95% aa sequence identity between paralogs. The ccTLR9 gene is contained in a single large exon. The ccMyD88a and ccMyD88b coding sequences span five exons. The TRAF6b gene spans six exons. PCR amplification to obtain the entire coding sequence of ccTRAF6a gene was not successful. The 2104-bp fragment amplified covers the 3' end of the gene and it contains a partial sequence of one exon and three complete exons. The predicated protein domains of the ccTLR9, ccMyD88 and ccTRAF6 are conserved and resemble orthologs from other vertebrates. Real-time quantitative PCR assays of the ccTLR9, MyD88a and b, and TRAF6a and b gene transcripts in healthy common carp indicated that mRNA expression varied between tissues. Differential

expression of duplicate copies were found for ccMyD88 and ccTRAF6 in white and red muscle tissues, suggesting that paralogs may have evolved and attained a new function. The genomic information described in this paper provides evidence of sequence and structural conservation of immune response genes in common carp.

KEYWORDS: Molecular cloning; TLR9; MyD88; TRAF6; Innate immunity; Gene duplication; Gene expression; *Cyprinus carpio*

INTRODUCTION

Cells of innate immune systems rely on a set of pattern recognition receptors (PRRs) which have a general ability to detect certain molecular structures present in pathogens (Krieg 2002). Toll-like receptors (TLRs) function as PRRs that recognize conserved molecular structures broadly shared by microbes that are known as pathogen-associated molecular patterns (PAMPs) and that trigger the signaling pathways that activate immune cells in response to pathogen infection. Among several TLRs that have been identified, TLR9 was characterized as the receptor that recognized unmethylated CpG dinucleotides in DNA, which are commonly found in bacterial and viral genomes, and that subsequently activated cellular immune responses (Hemmi et al. 2000; Lund et al. 2003; Akira et al. 2006; Barton 2007) Synthetic CpG oligodeoxynucleotides, a potent ligand for TLR9, have been proven to have an immunostimulatory effect on immune cells. Studies in fishes, including common carp (Tassakka and Sakai 2003; Tassakka and Sakai 2004), two flounders (Lee and Kim 2009; Liu et al. 2010), rainbow trout (Carrington and Secombes 2007) and Atlantic salmon (Jørgensen et al. 2001; Jørgensen et al. 2003) demonstrated that CpG stimulation activated antibacterial and antiviral immune responses. Häcker et al. (2000) showed that in mammals, recognition of CpG DNA activated the toll-like receptor/interleukin 1 receptor (TLR/IL-1R) signaling pathways via myeloid differentiation marker 88 (MyD88) and tumor necrosis factor receptor-associated factor 6 (TRAF6).

MyD88 is a cytoplasmic toll/interleukin-1 receptor (TIR) domain-containing adaptor molecule involved in the signaling through the IL-1R and TLR families (Hemmi et al. 2000). This molecule is used by all TLRs, except TLR3, and activates the transcription factor NF-κB and mitogen-activated protein kinases (MAPKs) to induce inflammatory cytokines (Nishiya et al.

2007; Kawai and Akira 2010). MyD88 interacts with TLRs through TIR domain-TIR domain interactions and uses its death domain to interact with the death domain-containing protein IL-1R-associated protein kinase (IRAK). Besides its function as an adaptor molecule in TLR/IL-1R signaling, MyD88 also has an important role in the IFN-γ signaling pathway in mammals, where in macrophages it bridges the cytoplasmic domain of IFN-γ receptor 1 and mixed-lineage kinase 3 (MLK3) in a pathway that, through the activation of p38, regulates mRNA stability of cytokines such as tumor necrosis factors and IFN-γ inducible protein 10 (IP-10) (Sun and Ding 2006).

TRAF6 is a cytoplasmic adaptor protein of the TRAF family that mediates signals induced by the tumor necrosis factor receptor (TNFR) superfamily and the interleukin-1 receptor (IL-1R) (Cao et al. 1996; Arch et al. 1998). It is as an important intermediate of CpG-DNA-induced signal transduction through the TLR/IL-1R – MyD88 pathway which leads to the activation of the IKK complex and JNK kinases (Häcker et al. 2000). A study of TRAF6 function in mice (Kobayashi et al. 2003) suggested that TRAF6 is a critical molecule that regulates the processes required for maturation, activation, and development of dendritic cells. In addition, human TRAF6 was shown to be involved in the stimulation of apoptosis (He et al. 2006).

The ability of TLR9 to detect viral DNA and the important roles of MyD88 and TRAF6 in TLR9 signal transduction make them likely candidates for involvement in the host antiviral response. In teleosts, full-length cDNA sequences encoding TLR9, MyD88 and TRAF6 molecules have been isolated (Meijer et al. 2004; Phelan et al. 2005; Franch et al. 2006; Skjæveland et al. 2008; Ortega-Villaizan et al. 2009; Rebl et al. 2009; Skjæveland et al. 2009; Yao et al. 2009); however, these gene transcripts have not been reported in common carp.

Kongchum et al. (2010) cloned and sequenced genomic fragments of several immune response genes from common carp, including TLR9, TRAF6 and two paralogous MyD88 genes. Here, I describe the isolation of full-length cDNA, genomic organization, phylogenetic relationship and tissue-specific mRNA expression distributions of the common carp TLR9, MyD88 and TRAF6. This information will expand knowledge of the structure and evolution of antiviral immune response genes in teleosts and vertebrates. The gene sequence data will also facilitate on-going research to develop molecular tools for selective breeding of common carp lines that are resistant to cyprinid herpesvirus-3 (CyHV-3), a causative agent of a lethal disease in common carp and koi.

MATERIALS AND METHODS

Experimental fish, tissue collection and RNA extraction

Koi carp (approximately 100 g body weight) purchased from a local pet store were used for gene isolation. Adult common carp (400g – 500g body weight) obtained from the E. W. Shell Fisheries Center, Auburn University, AL were used for tissue expression studies. Different tissues/organs were dissected from anaesthetized fish and immediately frozen in liquid nitrogen and stored at –80 ^{o}C until use. Tissue sampling was conducted in concordance with the National Center for Cool and Cold Water Aquaculture Institutional Animal Care and Use Committee protocol number 043. Total RNA was extracted using TRI Reagent (Sigma-Aldrich) and treated with RQ1 RNase-Free DNase (Promega) following the manufacturer's instructions. The concentration of the total RNA was quantified on a Nanodrop ND-1000 spectrophotometer (Nanodrop Technologies) and its integrity was visualized on an agarose gel.

Full-length cDNA isolation

The full-length cDNA sequences of the common carp TLR9, MyD88a, Myd88b, TRAF6a and TRAF6b were obtained by 5'- and 3'-rapid amplification of cDNA ends (RACE) using the SMART RACE cDNA Amplification Kit (Clontech) according to the manufacturers' instructions. First-strand cDNA was generated from 1μg total RNA extracted from kidney and spleen. The 3'-RACE cDNA was synthesized using 3'-RACE CDS Primer A (5'-AAGCAGTGGTATCAACGCAGAGTAC$(T)_{30}$VN-3'). For 5'-RACE, cDNA was synthesized using 5'-RACE CDS Primer A (5'-$(T)_{25}$VN-3') and SMART II A Oligonucleotide (5'-AAGCAGTGGTATCAACGCAGAGTACGCGGG-3'). The 3'- and 5'-RACE-ready cDNAs were then used as template for cDNA fragment amplification by PCR using gene-specific primers and Universal Primer A Mix (0.4 μM long oligonucleotide, 5'-CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGT-3' and 2 μM short oligonucleotide, 5'-CTAATACGACTCACTATAGGGC-3'). Gene-specific primers used for the amplification of RACE cDNA fragments (Table 1) were designed based on common carp genomic DNA sequences previously deposited in GenBank (accession numbers GU321985, GU321986, GU321987 and GU064560) using FastPCR software (http://www.biocenter.helsinki.fi/bi/Programs/fastpcr.htm). PCR amplifications were performed using a hot-lid thermocycler with the following thermal cycling profiles: 94 °C for 30 s, 68 °C for 30 s, and 72 °C for 3 min for 30 cycles, followed by a final extension of 72 °C for 10 min. Nested PCR was performed with a gene-specific primer and a nested universal primer (5'-AAGCAGTGGTATCAACGCAGAGT-3') to obtain specific PCR product. PCR amplicons were then cloned into a plasmid vector for nucleotide sequencing as previously described (Kongchum et al. 2010). Briefly, PCR amplicons were cloned into the pCR2.1-TOPO vector and

transformed into competent *E. coli* (TOPO TA Cloning Kit, Invitrogen). Transformants were spread on a Luria–Bertani (LB) agar plate containing X-gal and 100 µg/ml ampicillin. Insert-positive colonies were grown in LB broth with 100 µg/ml of ampicillin followed by Magnificent Broth (MB) with 50 µg/ml of kanamycin, and the plasmid DNA was isolated and sequenced.

Characterization of intronic regions

Primer pairs covering entire open reading frames of the common carp TLR9, MyD88 and TRAF6 genes were used to amplify gene fragments from genomic DNA. The primers used are listed in Table 2. PCR was carried out in an 11-µl reaction volume consisting of 10 µl reaction mixture (1X PCR buffer, 2 mM $MgCl_2$, 200 µM dNTPs, 1 µM each of forward and reverse primers, 0.5 U AmpliTaq Gold (Applied Biosystems) DNA polymerase), and 1 µl (~ 20 ng) gDNA template. PCR amplifications were performed as follows: 1 cycle of 94°C for 10 min; then 35 cycles of 94°C for 10 s, 60°C for 30 s, and 72°C for 1 min; followed by 1 cycle of 72°C for 7 minutes. PCR products were cloned and sequenced as described above. The genomic nucleotide sequences were compared to cDNA sequences to determine the exon–intron boundaries of the genes.

Sequence analyses

Sequence assemblies were carried out using Sequencher software (Gene Codes Corporation). Nucleotide sequences were translated to amino acid sequences by the ExPASy Translate tool (http://www.expasy.ch/tools/dna.html). The predicted amino acid sequences were blasted against the NCBI database using the BLASTP program (http://blast.ncbi.nlm.nih.gov/Blast.cgi) to identify sequence identities and similarities. The

identification and annotation of protein domains were performed using the web-based SMART program (http://smart.embl-heidelberg.de/). The presence of a signal peptide was inferred using the web-based SignalP 3.0 (http://www.cbs.dtu.dk/services/SignalP/). Phylogenetic analyses were conducted on amino acid and DNA sequences using MEGA 3.1 software (Kumar, Tamura et al. 2004). Multiple alignments were generated with ClustalW implemented in MEGA 3.1. Phylogenetic trees were constructed by the neighbor-joining (NJ) method using a Poisson correction model with 1,000 bootstrap replicates.

Real-time PCR

Real-time quantitative polymerase chain reaction (Q-PCR) assays were conducted to assess the abundance of TLR9, MyD88a, MyD88b, TRAF6a and TRAF6b mRNA transcripts in a variety of tissues of adult common carp. Tissues examined included kidney, spleen, liver, gill, white muscle, red muscle, and eye. Reverse transcription reactions were performed on 1 μg total RNA in a 20-μl setup using the MMLV RT 1st-Strand cDNA Synthesis Kit (Epicentre Biotechnologies) following the manufacturer's instructions. Briefly, total RNA was mixed with an oligo$(dT)_{21}$ primer in RNase-free water and incubated at 65 ^{o}C for 2 min using a hot-lid thermocycler, and then cooled on ice. The mixture was added with MMLV RT 10X reaction buffer, DTT, dNTPs, RNase inhibitor, and MMLV reverse transcriptase. First-strand cDNA was synthesized by incubating the reaction at 37^{o}C for 60 min, followed by reaction termination by heating to 85^{o}C for 5 min. Reverse transcription negative control (RT–) reactions also were performed and the cDNA samples were tested for genomic DNA contamination by PCR using intron-flanking primers. cDNA samples were kept at –20^{o}C until use. Gene-specific primers used in Q-PCR assays are listed in Table 3. Expression of β-actin was used as an internal control

to correct for RNA loading variation, and the primers were designed based on the common carp β-actin sequence (GenBank accession no. M24113).

Real-time Q-PCR was performed on an ABI Prism 7900HT Sequence Detection System (Applied Biosystems) with a 384-well reaction plate. The reaction (15 µl total volume) contained 4 µl diluted cDNA template equivalent to 4 ng total RNA and 11 µl of reaction mixture consisting of 7.5 µl 2X GoTaq qPCR Master Mix and 0.15 µl CXR Reference Dye (Promega), 0.75 µl each of 10 µM forward and reverse primers, and 1.85 µl nuclease-free water. Reactions were run using a standard cycling program as follows: 50 ^{o}C for 2 min, 95 °C for 2 min, followed by 40 cycles of 95 °C for 15 s and 60 °C for 1 min. Dissociation curve analysis of the PCR amplicons was performed at the end of amplification to verify single product amplification. Representative PCR products were sequenced to confirm the correct gene amplification.

To evaluate PCR amplification efficiency of each primer set used for expression analysis, a standard dilution curve was carried out in triplicate using six serial dilutions of pooled cDNA, with a starting concentration of cDNA corresponding to 20 ng of total RNA. Amplification efficiency (E) was calculated using the equation $E = 10^{[-1/\text{Slope}]}$. The oligonucleotide sequences, product sizes and amplification efficiencies are provided in Table 3.

Five adult common carp were used to examine tissue-specific mRNA expression. For each individual, real-time PCR reactions for each tissue cDNA were prepared in triplicate for target genes and the control gene (β-actin), and run on the same plate. In addition to tissue cDNA, the reactions of a reference sample (pooled cDNA), RT– control and no-template control (nuclease-free water) were included in each gene assay. Expression level of a target transcript

was determined as a relative expression ratio to a reference sample, and expressed in comparison to endogenous control gene according to Pfaffl (2001) as follows;

$$Expression\ ratio = \frac{(E_{target\ gene})^{\Delta Ct(reference-tissue)}}{(E_{control\ gene})^{\Delta Ct(reference-tissue)}}$$

Data analysis

Tissue-specific mRNA expression levels relative to a reference sample were $\log_2$ transformed and subjected to a two-way analysis of variance (ANOVA) with tissues and individuals as the factors. In cases where the results of the *F*-test were significant (p value < 0.05), post-hoc comparisons between group means were performed using the Tukey-Kramer HSD test. Expression levels were considered to be significantly different when $p \leq 0.05$. In addition to tissue distribution analysis for each gene transcript, similar analyses were carried out for MyD88a and MyD88b, and TRAF6a and TRAF6b, with the LSD post-hoc test to compare the difference in expression levels between paralogous transcripts in each tissue. Data were analyzed using JMP 8 software (SAS Institute).

RESULTS

Full-length cDNA cloning

Toll-like receptor 9 (TLR9)

The full-length cDNA of the common carp TLR9 (ccTLR9; GU809229) was obtained from five overlapping RACE cDNA fragments comprising of 4,400 bp. The open reading frame

(ORF) is predicted to encode a 1,064-amino acid protein. The untranslated regions (UTR) were 295 bp and 910 bp for the 5'-UTR and 3'-UTR, respectively. The 3'-UTR has five mRNA instability motifs (ATTTA) and a polyadenylation signal (AATAAA) six bp upstream of the poly-(A) tail. The putative ccTLR9 protein contains a predicted signal peptide sequence encompassing the first 21 amino acid residues of the N-terminal region (Fig. 1). Prediction of protein domains revealed that the putative amino acid sequence consisted of 14 leucine-rich repeat (LRR) domains, a leucine-rich repeat C-terminal (LRR-CT) domain, and a 150-amino acid Toll-interleukin-1 receptor (TIR) domain. Comparative protein domain organization of TLR9 proteins of common carp and other species is presented in Fig. 2. Results of a sequence homology search using BLASTP displayed 80% identity to the zebrafish (*Danio rerio*) ortholog, 51- 56% to other piscine and 38% to human orthologs (Table 4). The amino acid sequence identities and similarities between TIR domains were approximately 10-15% higher than the entire ORF identities and similarities.

Myeloid differentiation primary response gene 88 (MyD88)

The full-length cDNA of common carp MyD88a (ccMyD88a) was 2,389 bp with a putative ORF of 867 bp encoding 288 amino acid residues, 165-bp 5'-UTR and 1,357-bp 3'-UTR. The full-length ccMyD88b cDNA was 1,566 bp, with a putative ORF of 855 bp encoding 284 amino acids, 191-bp 5'-UTR and 520-bp 3'-UTR. In both ccMyD88a and ccMyD88b, the N-terminal death domain is composed of 91 amino acids and the C-terminal TIR domain contains 137 amino acids, with the presence of three highly conserved motifs: box 1, F/Y(DA); box 2, RDXXPG; and box 3 (FW). The two putative peptide sequences shared 91% identity with 23 different amino acid residues between paralogs. Results of a sequence homology search

revealed that the predicted ccMyD88a and ccMyD88b protein sequences were 90% and 88% identical to the zebrafish MyD88, respectively (Table 5). The alignment of amino acid sequences showed that the common carp MyD88 proteins are highly conserved (Fig. 3). Common carp MyD88a and MyD88b cDNA sequences were deposited in GenBank under accession numbers GU809230 and GU809231, respectively.

TNF receptor-associated factor 6 (TRAF6)

The common carp TRAF6 (ccTRAF6) gene sequence was first obtained from genomic DNA amplification using degenerate primers and was used to design RACE primers for cDNA cloning. The assembly of RACE fragments resulted in two contigs. The consensus cDNA sequences were named ccTRAF6a and ccTRAF6b, and deposited in GenBank under accession numbers GU985443 and HM535645, respectively. The full-length ccTRAF6a cDNA consisted of 2,619 bp comprising an ORF of 1,632 bp, a 5'-UTR of 32 bp, and a 3'-UTR of 948 bp. The ORF encoded a protein sequence of 543 amino acid residues. The full-length ccTRAF6b cDNA was 2,845 bp, with a 1,632-bp coding region that encoded a 543-amino acid protein, a 29-bp 5-UTR and a 1183-bp 3'-UTR. The 1,632-bp coding regions of the two cDNA sequences differed by 92 nucleotides. The predicted peptide sequences of ccTRAF6a and ccTRAF6b were 95% identical, with 24 different amino acid residues. Results of sequence homology searches using BLASTP revealed that the carp TRAF6 proteins shared 58 - 88% identity with those of teleosts and mammals (Table 6). Prediction of protein domains revealed the N-terminal RING domain between amino acids 72 and 110, and a C-terminal meprin and TRAF homology (MATH) domain between amino acids 380 and 504. The alignments of ccTRAF6a and ccTRAF6b to the zebrafish and human TRAF6 amino acid sequences are presented in Fig. 4.

Phylogenetic analyses

Phylogenetic analyses were performed on the full-length amino acid sequences of the predicted TLR9, MyD88 and TRAF6 molecules using the neighbor-joining (NJ) method. A phylogenetic tree constructed using 14 full-length amino acid sequences of TLR9 molecules showed two distinct branches for teleosts and mammals (Fig. 5a). The common carp and zebrafish sequences were most closely related, as expected from their relationship in family Cyprinidae, and clustered with other teleost TLR9 sequences with high bootstrap values. Fourteen MyD88 protein sequences from fishes, including two sequences of common carp, and other vertebrates were used for tree construction. The MyD88 tree was composed of two branches (Fig. 5b). The ccMyD88a and b grouped in a branch with other piscine MyD88 protein sequences, with a close relationship to zebrafish and catfish, representing other Ostariophysian fishes. For TRAF6, only a few sequences from teleosts were available. Phylogenetic relationships of TRAF6 amino acid sequences are presented in Fig. 5c. Phylogenetic trees generated based on nucleotide sequences are presented in Fig. 8 a-b. The NJ analysis revealed that ccTRAF6a and b formed a monophyletic group with zebrafish sequences and were clustered in the same branch with the other two teleost sequences. For both duplicated genes (MyD88 and TRAF6), the phylogenetic relationships supports the inference that they were most likely duplicated as part of the recent whole-genome duplication of common carp that took place after common carp and zebrafish ancestors diverged.

Genomic organization

The entire coding regions of the TLR9, MyD88 and TRAF6 genes were successfully amplified from genomic DNA, except for TRAF6a, using primers located outside and inside

coding regions. Comparisons of the genomic and cDNA sequences revealed the exon-intron boundaries in these genes. The common carp TLR9 gene was contained in a single large exon, which translated into a 1,064-amino acid sequence. The MyD88a and MyD88b coding sequences spanned five short exons ranging from 91 to 304 bp and were separated by four introns. The intron sizes in MyD88a are 544, 80, 372 and 725 bp, and in MyD88b are 483, 76, 400 and 1032 bp for introns 1, 2, 3 and 4, respectively. Both sequences translate into 284 amino acids. The MyD88a sequence was four amino acid residues shorter than the translated cDNA sequence due to a C/T mutation at the 3' end. The first exon encodes the entire 91-amino acid N-terminal death domain. The C-terminal TIR domains, which are composed of 137 amino acids, are encoded by exons 3, 4, and 5. The PCR amplification to obtain the entire coding sequence of the TRAF6a gene from genomic DNA was not successful, possibly due to a large insertion in intron 1 or 2. Alignment with the TRAF6a cDNA sequence showed that the 2104-bp fragment covered the 3' end of the gene and contained a partial sequence of one exon (37 bp) and three complete exons (72, 78 and 870 bp), which were separated by three introns (436, 128 and 209 bp). The 870-bp exon at the 3' end encodes the TRAF6 coiled coil region (64 amino acids) and the MATH domain (125 amino acids). The entire length of the TRAF6b gene (exons and introns) was 4,708 bp. It spans six exons (302, 151, 159, 72, 78 and 870 bp) separated by five introns (813, 892, 766, 128 and 477 bp). The predicted peptide sequence is composed of 543 residues. The first two exons encode the RING domain (39 amino acids) and the sixth exon encodes the coiled coil region (64 amino acids) and the MATH domain (125 amino acids). Protein domains and genomic structures of common carp TLR9, MyD88a, MyD88b, TRAF6a and TRAF6b genes are presented in Fig. 6a-c. The TLR9, MyD88a, MyD88b, TRAF6a and TRAF6b gene sequences

were deposited in GenBank under accession numbers GU321985, GU321986, GU321987, GU064560 and HM535646, respectively.

Gene expression

The abundance of target gene transcripts was determined in different tissues of five adult fish by real-time PCR. The tissue mRNA distributions are shown in Fig. 7 a-c. TLR9 was highly expressed in kidney (~2.5 fold) and moderately in gills (1.4 fold) relative to the reference sample (pooled tissue cDNA). Its expression was low to moderate in all other tissues, with the lowest expression observed in white muscle. MyD88a and MyD88b mRNA transcripts were detected in all tissues analyzed, with highest levels in liver and gills (1.93 and 2.51 fold, respectively). MyD88a was expressed slightly less than MyD88b in the spleen (0.5 fold) and slightly more in gill, muscle and eye tissues. TRAF6a and TRAF6b mRNA transcripts were abundant in liver and gills (1.7 and 1.9 fold, respectively). TRAF6b expression was 0.9-, 0.73- and 0.79-fold higher than TRAF6a in white muscle, red muscle and eye tissues, respectively. Comparison between the two paralogs in each tissue showed that MyD88a and TRAF6b expression was significantly higher ($p < 0.05$) than the paralogous transcripts in white and red muscles (Fig. 7 b-c).

DISCUSSION

In this study, I cloned and sequenced cDNAs encoding TLR9 and the key TLR9-signaling adaptor molecules MyD88 and TRAF6 from common carp. I found two paralogous loci encoding each MyD88 and TRAF6 protein. To my knowledge, this is the first report of

compelling evidence for MyD88 and TRAF6 gene duplications in vertebrates. I also sequenced genomic DNA to determine exon-intron boundaries in the ccTLR9, ccMyD88a, ccMyD88b, ccTRAF6a and ccTRAF6b genes. To explore the evolutionary relationships and elucidate the functions of these genes, I conducted phylogenetic analysis, characterized predicted functional domains of amino acid sequences, and examined mRNA expression patterns in tissues of healthy adult common carp.

TLR9 has been identified in mammals and several fish species. Sequence homology and phylogenetic analysis clustered the fish TLR9 sequences on one branch with ccTLR9 showing closest relationship to the zebrafish TLR9, another member of the family Cyprinidae. The predicted ccTLR9 TIR domain had high sequence similarity to the TIR domains of other known TLR9 molecules. The two CXXC motifs important for CpG binding (Lee et al. 2001; Bell et al. 2003) and a conserved motif of Asp and Tyr amino acids within an LRR domain important for interacting with CpG-DNA (Rutz et al. 2004), which is present in TLR9 of zebrafish (Asp^{546} and Tyr^{548}) and human (Asp^{534} and Tyr^{536}), also was found in ccTLR9 (Asp^{552} and Tyr^{554}) (Fig. 1). Highly conserved motifs in the TIR domain important for signaling and receptor localization also were observed in ccTLR9. The conserved TLR9 functional domains present in the ccTLR9 suggest that it has a function similar to other vertebrates' TLR9. By sequencing genomic DNA with overlapping primer pairs that covered the entire coding region, I found that the ccTLR9 gene was contained in a single large exon. This genomic organization is similar to that of zebrafish (Ensembl database, chromosome 8: 55,380,305-55,383,596), but different from those of other known teleosts. The fugu TLR9 gene (AC156439) has one intron located within the TIR domain. The rainbow trout TLR9 gene (FJ594277) (Ortega-Villaizan et al. 2009) is composed of four exons with three introns located before the start codon, between the

transmembrane and TIR domain, and within the TIR domain. The Japanese flounder TLR9 gene (AB234024) (Takano et al. 2007) has two introns located in the signal peptide region and within the TIR domain. Diversity of TLR gene organization among teleosts and between teleosts and mammals also was reported for TLR1, 2, 3 and 5 (Oshiumi et al. 2003; Rodriguez et al. 2005; Palti et al. 2010), but it is yet unclear whether there are functional differences arising from differences in the genomic organization of these genes.

MyD88 mediates signal transduction from all activated TLRs except TLR3 (Ohnishi et al. 2009). The MyD88 molecular structure consists of an N-terminal death domain, a C-terminal Toll-IL-1 (TIR) domain and a short connecting linker. In this study, I identified two common carp Myd88 cDNA sequences, ccMyD88a and ccMyD88b, in which the complete ORFs encoded 284- and 288-amino acid proteins, respectively. The two amino acid sequences shared high identities with those of fishes and mammals, with highest homology to zebrafish. I sequenced genomic DNA and found that each ccMyD88a and ccMyD88b genes span five exons. The two loci had the same exon sizes, but the intron sizes were different. Phylogenetic analysis showed that the common carp MyD88 sequences were on the same branch with other teleosts' MyD88 and were most closely related to those of zebrafish and channel catfish. Prediction of the protein domains revealed that ccMyD88a and ccMyD88b possessed a typical N-terminal death domain and a C-terminal Toll-IL-1 (TIR) domain with the three highly conserved sequence motifs called box 1, box 2 and box 3, which are (F/Y)DA, RDXXPG and FW, respectively (Slack et al. 2000). The death domain functions as an important mediator necessary for regulation of apoptosis (Weber and Vincenz 2001). The TIR domain functions as an adaptor between members of the TLR/IL-1R superfamily and downstream signaling proteins (Janssens and Beyaert 2002). The

presence of highly conserved peptide domains shared with zebrafish and human MyD88 sequences suggests that ccMyD88a and ccMyD88b are functional adaptor proteins.

TRAF6 is an important adaptor molecule that mediates signals induced by TNFR and IL-1R. Studies in mammals (Häcker et al. 2000; Poole et al. 2006) and in zebrafish (Phelan et al. 2005) demonstrated that TRAF6 is involved in the regulation of the NF-κB pathway. Here I cloned and sequenced two paralogous transcripts that share high sequence identities with the zebrafish TRAF6 ortholog. The ORFs of the two sequences differed by 92 nucleotides, and the predicted peptides differed by 24 amino acid residues. I also analyzed the genomic structure of the two loci and found that the intron sizes were different (Fig. 6c.). Although I was not able to obtain the entire coding sequence of ccTRAF6a from genomic DNA, the polymorphisms in the nucleotide and predicted amino acid sequences together with intron size variations suggest that the TRAF6 gene is duplicated in common carp. Predicted amino acid sequences indicated that both ccTRAF6a and ccTRAF6b ORFs encode functional proteins. Characterization of protein domains revealed that the predicted ccTRAF6a and ccTRAF6b proteins possess a RING domain and a MATH domain, which are also conserved in zebrafish and human TRAF6. Phylogenetic analysis using amino acid sequences further demonstrated that ccTRAF6a and b were closely related to other TRAF6 teleost orthologs.

Common carp is known to be a tetraploid species (David et al. 2003; Chistiakov and Voronova 2009), and a number of duplicated genes have been reported (Futami et al. 2001; Hermesz et al. 2001; Ferencz 2008; Stolte et al. 2008; Hermesz and Ferencz 2009). I observed a large number of single nucleotide polymorphisms (SNPs) between the two paralogs of MyD88 and TRAF6 and also found allelic SNPs in exons and introns within each paralog. I then developed locus-specific SNP markers and genotyped them in an F_2 mapping family (n =50)

using methods previously described (Kongchum et al. 2010) to further elucidate whether the two genes are duplicated. SNPs in MyD88a, TRAF6a and TRAF6b were polymorphic and showed Mendelian segregation, however, SNPs in MyD88b were not polymorphic in the family genotyped. I used chi-square to test for independence of segregation of TRAF6a and TRAF6b SNPs, and I found that they segregated independently (*P*-value = 0.162). The independent segregation of TRAF6a and b SNPs and distinct nucleotide and predicted amino acid sequences of the two MyD88 and TRAF6 paralogs presented in the phylogenetic trees (Figures 5b and c and 8) support the interpretation that MyD88 and TRAF6 are duplicated in the common carp genome.

To assess basal mRNA expression levels, I performed expression assays in a variety of tissues. Common carp TLR9 was predominantly expressed in kidney and gill, but was detected at low levels in other tissues examined (Fig. 7a). A study in Japanese flounder (Takano et al. 2007) showed that TLR9 was highly expressed in epithelial and lymphoid organs, such as the gills, intestines, kidney, spleen and stomach. In my study, TLR9 was expressed at a relatively low level in spleen. This difference in tissue expression pattern between the two studies may be due to the fact that cell populations that express TLR9 differ between species, such as the immune cells of the myeloid lineage express TLR9 in mice, but not in humans (Klinman 2004). Similar differences between species' expression profiles among fishes were previously observed for TLR1, 3, 7 and 8 (Rodriguez et al. 2005; Palti et al. 2010; Palti et al. 2010). However, comparison of TLR gene expression profiles should be interpreted with caution since such discrepancies may result not only from species variation, but also from differences in individuals' immunological status, developmental stage (Renshaw, Rockwell et al. 2002), and genetic background.

Transcripts of the ccMyD88a and b genes were observed in all tissues examined, with predominant expression in liver and gills, but I also identified differential tissue expression profiles between the two transcripts. ccMyD88a expression levels were higher in eye, red muscle and white muscle, whereas ccMyD88b mRNA was expressed at a higher level in spleen (Fig. 7b). Similarly, to MyD88, TRAF6 mRNAs were predominantly expressed in liver and gills. The two paralogous transcripts showed similar expression patterns in kidney, spleen, liver and gill, but had substantially different expression levels in white muscle, red muscle and eye. The significantly higher expression of ccMyD88a and ccTRAF6b in muscle may indicate that these two isoforms have attained a non-immune signal transduction function in common carp, as often occurs when two duplicated genes remain active (Walsh 1995; Ohno 1999; Van de Peer 2005; Coulibaly et al. 2006; Palti et al. 2010). Differential expression of duplicated loci previously has been described in other common carp genes. Constitutive expression of interferon gamma (IFNγ), an important cytokine for antiviral immunity and tumor control in mammals, differed between the IFN-*γ*-1 and IFN-*γ*-2 isotypes (Stolte et al. 2008). IFN-*γ*-1 was expressed 10-fold lower in kidney, spleen, blood and gut, but was expressed 5-fold higher in muscle. Two common carp c-myc genes (CAM1 and CAM2) also were found to be differentially expressed, and their phylogenetic analysis suggested that CAM1 has evolved faster than CAM2 after gene duplication, which may suggest that the CAM1 gene has obtained a new function (Futami et al. 2005).

In summary, I cloned and sequenced the immune response genes TLR9, MyD88a and b and TRAF6a and b in common carp. I then characterized their predicted protein domains and determined their mRNA expression to demonstrate that these genes are functional. The data I present in this paper provide evidence that the structural and functional characteristics of these

immune response genes in common carp are conserved and resemble those of other vertebrate orthologs. The genomic information obtained in this study will be useful for identifying gene polymorphisms that may be associated with differential immune responses and susceptibility to bacterial and viral diseases in common carp. Further research is needed to better understand the ligand specificity and the role of TLR9 in the common carp immune response. The development of assays to measure ligand binding and its effect on downstream signaling are critical for further elucidation of ccTLR9 functions.

LITERATURE CITED

Akira, S., S. Uematsu, et al. (2006). "Pathogen recognition and innate immunity." Cell **124**(4): 783-801.

Arch, R. H., R. W. Gedrich, et al. (1998). "Tumor necrosis factor receptor-associated factors (TRAFs)—a family of adapter proteins that regulates life and death." Genes & Development **12**(18): 2821-2830.

Barton, G. M. (2007). "Viral recognition by Toll-like receptors." Seminars in Immunology **19**(1): 33-40.

Bell, J. K., G. E. D. Mullen, et al. (2003). "Leucine-rich repeats and pathogen recognition in Toll-like receptors." Trends in Immunology **24**(10): 528-533.

Cao, Z., J. Xiong, et al. (1996). "TRAF6 is a signal transducer for interleukin-1." Nature **383**: 443–446.

Carrington, A. C. and C. J. Secombes (2007). "CpG oligodeoxynucleotides up-regulate antibacterial systems and induce protection against bacterial challenge in rainbow trout (*Oncorhynchus mykiss*)." Fish & Shellfish Immunology **23**(4): 781-792.

Chistiakov, D. and N. Voronova (2009). "Genetic evolution and diversity of common carp *Cyprinus carpio* L." Central European Journal of Biology **4**(3): 304-312.

Coulibaly, I., S. A. Gahr, et al. (2006). "Genomic structure and expression of uncoupling protein 2 genes in rainbow trout (Oncorhynchus mykiss)." BMC Genomics **7**(203): 203.

David, L., S. Blum, et al. (2003). "Recent duplication of the common carp (*Cyprinus carpio* L.) genome as revealed by analyses of microsatellite loci." Molecular Biology and Evolution **20**(9): 1425 - 1434.

Ferencz, Á. and E. Hermesz (2008). "Identification and characterization of two mtf-1 genes in common carp." Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology **148**(3): 238-243.

Franch, R., B. Cardazzo, et al. (2006). "Full-length sequence and expression analysis of Toll-like receptor 9 in the gilthead seabream (*Sparus aurata* L.)." Gene **378**: 42-51.

Futami, K., T. Komiya, et al. (2001). "Differential expression of max and two types of c-myc genes in a tetraploid fish, the common carp (*Cyprinus carpio*)." Gene **269**(1-2): 113-119.

Futami, K., H. Zhang, et al. (2005). "Functional divergence of duplicated c-myc genes in a tetraploid fish, the common carp (*Cyprinus carpio*)." Gene **363**: 61-66.

Häcker, H., R. M. Vabulas, et al. (2000). "Immune cell activation by bacterial CpG-DNA through myeloid differentiation marker 88 and tumor necrosis factor receptor–associated factor (Traf) 6." The Journal of Experimental Medicine **192**(4): 595-600.

He, L., X. Wu, et al. (2006). "TRAF6 regulates cell fate decisions by inducing caspase 8-dependent apoptosis and the activation of NF-κB." Journal of Biological Chemistry **281**(16): 11235-11249.

Hemmi, H., O. Takeuchi, et al. (2000). "A Toll-like receptor recognizes bacterial DNA." Nature **408**(6813): 740-745.

Hermesz, E., M. Ábrahám, et al. (2001). "Identification of two hsp90 genes in carp." Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology **129**(4): 397-407.

Hermesz, E. and Á. Ferencz (2009). "Identification of two phospholipid hydroperoxide glutathione peroxidase (gpx4) genes in common carp." Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology **150**(1): 101-106.

Janssens, S. and R. Beyaert (2002). "A universal role for MyD88 in TLR/IL-1R-mediated signaling." Trends in Biochemical Sciences **27**(9): 474-482.

Jørgensen, J. B., A. Johansen, et al. (2001). "CpG oligodeoxynucleotides and plasmid DNA stimulate Atlantic salmon (*Salmo salar* L.) leucocytes to produce supernatants with antiviral activity." Developmental & Comparative Immunology **25**(4): 313-321.

Jørgensen, J. B., L.-H. Johansen, et al. (2003). "CpG DNA induces protective antiviral immune responses in Atlantic salmon (*Salmo salar* L.)." Journal of Virology **77**(21): 11471-11479.

Kawai, T. and S. Akira (2010). "The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors." Nature Immunology **11**(5): 373-384.

Klinman, D. M. (2004). "Immunotherapeutic uses of CpG oligodeoxynucleotides." Nature Reviews Immunology **4**(4): 249-259.

Kobayashi, T., P. T. Walsh, et al. (2003). "TRAF6 is a critical factor for dendritic cell maturation and development." Immunity **19**(3): 353-363.

Kongchum, P., Y. Palti, et al. (2010). "SNP discovery and development of genetic markers for mapping innate immune response genes in common carp (*Cyprinus carpio*)." Fish & Shellfish Immunology **29**(2): 356-361.

Krieg, A. M. (2002). "CpG motifs in bacterial DNA and their immune effects." Annual Review of Immunology **20**: 709-760.

Kumar, S., K. Tamura, et al. (2004). "MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment." Briefings in Bioinformatics **5**: 150-163.

Lee, E. H. and K. H. Kim (2009). "CpG-ODN increases resistance of olive flounder (*Paralichthys olivaceus*) against *Philasterides dicentrarchi* (Ciliophora: Scuticociliatia) infection." Fish & Shellfish Immunology **26**(1): 29-32.

Lee, J.-H., K. S. Voo, et al. (2001). "Identification and characterization of the DNA binding domain of CpG-binding protein." Journal of Biological Chemistry **276**(48): 44669-44676.

Liu, C.-S., Y. Sun, et al. (2010). "Identification and analysis of the immune effects of CpG motifs that protect Japanese flounder (*Paralichthys olivaceus*) against bacterial infection." Fish & Shellfish Immunology **29**(2): 279-285.

Lund, J., A. Sato, et al. (2003). "Toll-like receptor 9–mediated recognition of herpes simplex virus-2 by plasmacytoid dendritic cells." The Journal of Experimental Medicine **198**(3): 513-520.

Meijer, A. H., S. F. Gabby Krens, et al. (2004). "Expression analysis of the Toll-like receptor and TIR domain adaptor families of zebrafish." Molecular Immunology **40**(11): 773-783.

Nishiya, T., E. Kajita, et al. (2007). "Distinct roles of TIR and non-TIR regions in the subcellular localization and signaling properties of MyD88." FEBS Letters **581**(17): 3223-3229.

Ohnishi, H., H. Tochio, et al. (2009). "Structural basis for the multiple interactions of the MyD88 TIR domain in TLR4 signaling." Proceedings of the National Academy of Sciences U.S.A. **106**(25): 10260-10265.

Ohno, S. (1999). "Gene duplication and the uniqueness of vertebrate genomes circa 1970-1999." Seminars in Cell and Developmental Biology **10**(5): 517-522.

Ortega-Villaizan, M., V. Chico, et al. (2009). "The rainbow trout TLR9 gene and its role in the immune responses elicited by a plasmid encoding the glycoprotein G of the viral

haemorrhagic septicaemia rhabdovirus (VHSV)." Molecular Immunology **46**(8-9): 1710-1717.

Oshiumi, H., T. Tsujita, et al. (2003). "Prediction of the prototype of the human Toll-like receptor gene family from the pufferfish, *Fugu rubripes*, genome." Immunogenetics **54**(11): 791-800.

Palti, Y., S. A. Gahr, et al. (2010). "Identification, characterization and genetic mapping of TLR7, TLR8a1 and TLR8a2 genes in rainbow trout (*Oncorhynchus mykiss*)." Developmental & Comparative Immunology **34**(2): 219-233.

Palti, Y., M. F. Rodriguez, et al. (2010). "Identification, characterization and genetic mapping of TLR1 loci in rainbow trout (*Oncorhynchus mykiss*)." Fish & Shellfish Immunology **28**(5-6): 918-926.

Pfaffl, M. W. (2001). "A new mathematical model for relative quantification in real-time RT-PCR." Nucleic Acid Research **29**(9): 2002-2007.

Phelan, P. E., M. T. Mellon, et al. (2005). "Functional characterization of full-length TLR3, IRAK-4, and TRAF6 in zebrafish (*Danio rerio*)." Molecular Immunology **42**(9): 1057-1071.

Poole, E., C. A. King, et al. (2006). "The UL144 gene product of human cytomegalovirus activates NFκB via a TRAF6-dependent mechanism." EMBO Journal **25**(18): 4390-4399.

Rebl, A., T. Goldammer, et al. (2009). "Characterization of two key molecules of teleost innate immunity from rainbow trout (*Oncorhynchus mykiss*): MyD88 and SAA." Veterinary Immunology and Immunopathology **131**(1-2): 122-126.

Renshaw, M., J. Rockwell, et al. (2002). "Cutting Edge: Impaired Toll-Like Receptor Expression and Function in Aging." The Journal Immunology **169**(9): 4697-4701.

Rodriguez, M., G. Wiens, et al. (2005). "Characterization of Toll-like receptor 3 gene in rainbow trout (*Oncorhynchus mykiss*)." Immunogenetics **57**(7): 510-519.

Rutz, M., J. Metzger, et al. (2004). "Toll-like receptor 9 binds single-stranded CpG-DNA in a sequence- and pH-dependent manner." European Journal of Immunology **34**(9): 2541-2550.

Skjæveland, I., D. B. Iliev, et al. (2009). "Identification and characterization of TLR8 and MyD88 homologs in Atlantic salmon (*Salmo salar*)." Developmental & Comparative Immunology **33**(9): 1011-1017.

Skjæveland, I., D. B. Iliev, et al. (2008). "A TLR9 homolog that is up-regulated by IFN-γ in Atlantic salmon (*Salmo salar*)." Developmental & Comparative Immunology **32**(6): 603-607.

Slack, J. L., K. Schooley, et al. (2000). "Identification of two major sites in the type I interleukin-1 receptor cytoplasmic region responsible for coupling to pro-inflammatory signaling pathways." Journal of Biological Chemistry **275**(7): 4670-4678.

Stolte, E. H., H. F. J. Savelkoul, et al. (2008). "Differential expression of two interferon-γ genes in common carp (*Cyprinus carpio* L.)." Developmental & Comparative Immunology **32**(12): 1467-1481.

Sun, D. and A. Ding (2006). "MyD88-mediated stabilization of interferon-γ-induced cytokine and chemokine mRNA." Nature Immunology **7**(4): 375-381.

Takano, T., H. Kondo, et al. (2007). "Molecular cloning and characterization of Toll-like receptor 9 in Japanese flounder, *Paralichthys olivaceus*." Molecular Immunology **44**(8): 1845-1853.

Tassakka, A. C. M. A. R. and M. Sakai (2003). "The in vitro effect of CpG-ODNs on the innate immune response of common carp, *Cyprinus carpio* L." Aquaculture **220**(1-4): 27-36.

Tassakka, A. C. M. A. R. and M. Sakai (2004). "Expression of immune-related genes in the common carp (*Cyprinus carpio* L.) after stimulation by CpG oligodeoxynucleotides." Aquaculture **242**(1-4): 1-12.

Van de Peer, Y. (2005). "Evolutionary genetics: When duplicated genes don't stick to the rules." Heredity **96**(3): 204-205.

Walsh, J. B. (1995). "How often do duplicated genes evolve new functions?" Genetics **139**(1): 421-428.

Weber, C. H. and C. Vincenz (2001). "The death domain superfamily: a tale of two interfaces?" Trends in Biochemical Sciences **26**(8): 475-481.

Yao, C.-L., P. Kong, et al. (2009). "Molecular cloning and expression of MyD88 in large yellow croaker, *Pseudosciaena crocea*." Fish & Shellfish Immunology **26**(2): 249-255.

Table 1 Primers used for cDNA cloning.

Primer name	Sequence (5'-3')	Purpose
TLR9-5RACE-1	GGCAGAGCTGCTTGAGAACTGAGATGTTGC	5' RACE PCR
TLR9-5RACE-2	GGTGCCGTTCTGGATGGCGTAGGCT	5' RACE PCR
TLR9-RACE-N	AGCTTGTGGAGTCGTATGAAGA	Nested PCR
TLR9-3RACE	TCGCATTGACCTGTACTCTGACAAGGCC	3' RACE PCR
TLR9-3RACE-N	CAGATGCTGAGAATGGACTCT	Nested PCR
TLR9-3RACE-N	TGGCAGCAACCAGCATTCCCCTCCTGAAGC	Nested PCR
MyD88-5RACE	GCCTTGGCCAGCCGGGTCCAGAACCAG	5' RACE PCR
MyD88-5RACE-N	CACATGCATCACTGTCCAGATAGTCATCAG	Nested PCR
MyD88a-5RACE-N	GGAACCTGGACAGGCTTCCTCAGCTGTC	Nested PCR
MyD88b-5RACE-N	CACTGGAACCTGGACAGGCTTCCGCTGT	Nested PCR
MyD88-3RACE	GTCTTTGACCGAGACGTCCTTCCTGGCAC	3' RACE PCR
TRAF6-5RACE	ACCCTGCATAGTGTGCACGAAAAGTGAGATGTA	5' RACE PCR
TRAF6-3RACE	GCAGTATCAAGGTGTCTACGTATGGCGTTTGGA	3' RACE PCR

Table 2 Primers used for gene fragment cloning.

Primer name	Sequence (5'-3')	Amplicon size (bp)
TLR9-F1	TCGTTCCATTACACCACTGAGCCCA	
TLR9-R1	CTCAGGGAGATCTTGAAACACACG	726
TLR9-F2	GGCAACCCATGTTATCAGTCTTAC	
TLR9-R2	AACATTCCGGAGCTCTTCAT	474
TLR9-F3	CGACTCCACAAGCTCCAAGAGTTAG	
TLR9-R3	CCTTGTCAGAGTACAGGTCAATGC	775
TLR9-F4	GACTTTAAATGGCCAGCAGTTCAC	
TLR9-R4	TCAAGGATCTTCAGCTGGTTGTGG	693
TLR9-F5	TGAGCTTCGTCTTGTAAGCCTGGA	
TLR9-R5	GAATGCTGGTTGCTGCCAATGTC	443
TLR9-F6	TCCAGAGTCATTGGCTGGTGTCAA	
TLR9-R6	TACTCCCCTTGCAGGTGGTGTGTA	924
MyD88a-5eF	GTCTTTCAGCTTATCGGTAGCGTCA	
MyD88a-5eR	TCACTGGTGATGGTCCACACACAT	1336
MyD88b-5eF	TCCTTCAGTTTATCGCTAGCGTCG	
MyD88b-5eR	TGGCGATGGTCCACACGCAC	1269
MyD88-F1	TGACTTCCAGTTTGTGCATGAG	
MyD88-R1	GCATCACTGTCCAGATAGTC	558, 586
MyD88-F2	GACTATCTGGACAGTGATGCATGTG	
MyD88-R2	TGAGTCTGAGAAACTCTCCCAAGC	1922, 1372
TRAF6-F1	ACCGAGGCAAGATGGCTTGCAGTG	
TRAF6-R1.2	ACTGCATCCTACATTAGGACATCGG	1250
TRAF6-F2	TGTCCAGTTGACAATGAGGTGCTG	
TRAF6-R2.3	GCATACACAAAGCTCCCTGCACAGG	1176
TRAF6-3F	TGCCTGTGCAGGGAGCTTTGTGTATGC	
TRAF6-R1	GCATCTGTGTAAATTCCTGCATGTGCT	1009, 1607
TRAF6-R4	GAGCATTGTAGGAACTCAAAGTGAAGGT	1845
TRAF6b-i-F1	ACTCGTGTTCAGCAAGGATGCA	
TRAF6-R2	TGGGAGTTTGCAGATGCAGTCGTA	873
TRAF6-F3	TACGCCGAAAGCTGAGCTCACT	
TRAF6-R6	CACCCTTTAGCTTGGTTGACAGG	843, 842

Table 3 Primers used for gene expression analysis.

Gene	Forward primer sequence (5'- 3')	Reverse primer sequence (5'- 3')	Size (bp)	Efficiency (*E*)
TLR9	TCCAGAGTCATTGGCTGGTGTCAA	GAATGCTGGTTGCTGCCAATGTC	119	1.83
MyD88a	GATCAAACAGTTGGAACAGACCG	CATGCATCACTGTCCAGATAGTC	165	1.93
MyD88b	ACCATCGCCAGTGAGCTCATAGAG	GAGAATGCTGGGAAAAGGCTTCTC	189	1.92
TRAF6a	TTCCAGAGGCCCACTGTCCT	GGCTAGCATCAAATCGTGGTGTC	145	1.99
TRAF6b	TTCCAGAGGCCCACTGTGCAA	GCTAGCGTCGAATCGTGGTGTG	144	1.99
β-actin	GATCATTGCCCCACCTGAGCGTA	TGGACGATGGATGGTCCAGACTC	129	2.00

Table 4 Comparison of TLR9 amino acid sequences of common carp with those of other teleosts and human.

	Carp	**Zebrafish**	**Trout**	**Flounder**	**Fugu**	**Human**
ORF (bp)	3195	3174	3225	3198	3138	3099
Amino acid residues [1]	1064	1057	1074	1065	1045	1032
Percent identity (similarity) [2]						
Entire ORF	-	80 (88)	56 (71)	51 (66)	51 (66)	38 (55)
TIR domain	-	90 (96)	72 (83)	72 (83)	69 (82)	52 (68)
Number of LRR domain [3]	14	16	16	14	16	19

[1] GenBank accession numbers: common carp GU809229; zebrafish NP_001124066; rainbow trout NP_001123463; Japanese flounder BAE80691; fugu AAW69377; and human AAZ95519

[2] Sequence identity and similarity identified by the NCBI blastp program

[3] Predicted LRR domain based on results from the SMART program

Table 5 Percent identity and similarity of common carp MyD88a and MyD88b amino acid sequences to those of other fishes and human MyD88 amino acid sequences.

Species	MyD88a		MyD88b		Accession no.
	Identity	Similarity	Identity	Similarity	
Danio rerio	90	96	88	96	AAZ16494
Ictalurus punctatus	76	84	73	82	ACD81929
Plecoglossus altivelis	73	84	71	84	BAI68385
Salmo salar	71	85	69	83	NP_001130017
Oncorhynchus mykiss	69	82	68	81	NP_001117893
Takifugu rubripes	69	80	68	79	NP_001106666
Paralichthys olivaceus	69	81	68	80	BAE75959
Homo sapiens	62	74	62	74	AAP36509

Table 6 Percent identity and similarity of common carp TRAF6a and TRAF6b amino acid sequences to those of other fishes and human TRAF6 amino acid sequences.

Species	TRAF6a		TRAF6b		Accession no.
	Identity	Similarity	Identity	Similarity	
Danio rerio	88	94	88	93	NP_001038217
Plecoglossus altivelis	67	77	66	77	BAI68387
Gasterosteus aculeatus	64	75	63	75	ABJ15863
Homo sapiens	56	70	57	70	NP_004611

Figure 1 Alignment of common carp (cc), zebrafish (dr) and human (hs) TLR9 amino acid sequences performed using CLUSTALW v2.0. Alignment gaps are marked by dashes. Asterisks (*) and colons (:) below the alignments indicate identity and similarity, respectively, among all species examined. The two CXXC motifs important for CpG binding and gene regulation are shown in bold font and boxed. The two essential amino acids for interacting with CpG-DNA are indicated by closed triangles (part of LRR8). In the common carp sequence, the predicted signal peptide is in italic and underlined, the leucine-rich repeats (LRR) are shaded in grey, the transmembrane domain is underlined and the Toll/IL-1 receptor (TIR) domain is in bold font. Highly conserved motifs in TIR domains that are important for signaling and receptor localization are double underlined. Domain names are labeled above the alignments. CT: C (carboxyl)-terminal. GenBank accession numbers for the aligned sequences are: common carp GU809229; zebrafish NP_001124066; and human AAZ95519.

```
                          Signal Peptide
ccTLR9          -------MFGHMLYLALILNQFNHFATLHPEFYPCEIHTTKDGDINVDCQRRHLANVPKF 53
drTLR9          -------MFGPMVSLILLLNQFQLFAASHPQFYPCESHSTKDGHINVDCQHRRLSKVPRF 53
hsTLR9          MGFCRSALHPLSLLVQAIMLAMTLALGTLPAFLPCELQPHG----LVNCNWLFLKSVPHF 56
                       :.   : :  ::  :       * * *** :.       *:*:   * .**:*
                           LRR1
ccTLR9          TSLS----VISLNLNENHIHHIKGATFSGLANLKHLSLMWNCIPDRFKELRWPSCSLKID 109
drTLR9          TSPS----VISLNLNNNHIHRIKGDAFSGLPNLKYLSLMWNCISDRLKEARWPLCSVNID 109
hsTLR9          SMAAPRGNVTSLSLSSNRIHHLHDSDFAHLPSLRHLNLKWNCPPVGLSPMHFP-CHMTIE 115
                :  :    * **.*..*:**::::.  *: *..*::*.* *** .  :.  ::* * :.*:
                              LRR2
ccTLR9          PNAFSGLKNLTSLQLAGNSLKTIPPLPKQLEILGLEFNHIFQIVKP--LGTPLLKQLLLN 167
drTLR9          PDAFVGLKNLTSLQLAGNSLKMIPPLPKQLEVLGLEFNNIFNIVKP--LGTPQLKQLLLS 167
hsTLR9          PSTFLAVPTLEELNLSYNNIMTVPALPKSLISLSLSHTNILMLDSASLAGLHALRFLFMD 175
                *.:* .: .* .*:*: *.:  :*.***.*  *.*...:*: : ..   *   *: *::.
                                                    LRR3          LRR4
ccTLR9          KNCFYANPCYQSYFIDPRVFQDLPELLNLTLSYNNVTIVPPYLPLSLESLDLGENKITHI 227
drTLR9          KNCFYANPCHQPYFINSSVFQDLPELLNLTLSYNNLTAIPSYLPGSLESLDLRENTIDHI 227
hsTLR9          GNCYYKNPCRQALEVAPGALLGLGNLTHLSLKYNNLTVVPRNLPSSLEYLLLSYNRIVKL 235
                 **:* *** *.  : . .: .* :* :*:*.***:* :*  ** *** * *  * * ::
                LRR4                                                    LRR5
ccTLR9          NKESFANLKNLRHLNLGWNCQRCDHASEPCFPCPNNQSLNLHQDAFLDQRDSLISLSLRG 287
drTLR9          NKESFANLRNLRHLNLGWNCQRCDHASDPCFPCPNNQSLDLHQDAFLDQRDSLVSLHLQG 287
hsTLR9          APEDLANLTALRVLDVGGNCRRCDHAPNPCMECPR-HFPQLHPDTFS-HLSRLEGLVLKD 293
                  *.:***  ** *::* **:***** .:**: **. :  :** *:*  : . * .* *:.
                     LRR5                  LRR6
ccTLR9          NSLHTIPQHLFIRLHKLQELDLSDNFLAYAIQNGTFYEELRNVVSLSLLYNYEPLTSFSE 347
drTLR9          NSLRTLPRHLFINLHKLQELDLSSNFLAFTIQNGTFYEELQNVVILNLLYNYEPLKTFPE 347
hsTLR9          SSLSWLNASWFRGLGNLQVLDLSENFLYKCITKTKAFQGLTQLRKLNLSFNYQKRVSFAH 353
                .**  :    *  * :** ****.***   * : . :: * ::  *.* :**:   :*..
                                  LRR7
ccTLR9          LILSPSIEKMTALRELHLSGLFFRVLSNHSLAPLVKLPRFEFLELRMNFICSVSMDAISQ 407
drTLR9          LNLSPYIEKMASLRELYLSGFFFKKLSNRSIAPLVKLPRLEVLDLRMNFICDISIDGLSQ 407
hsTLR9          LSLAPSFGSLVALKELDMHGIFFRSLDETTLRPLARLPMLQTLRLQMNFINQAQLGIFRA 413
                * *:* : .:.:*:** : *:**: *.: :: **.:** :: * *:**** . .:. :

ccTLR9          LRTLRWVGLSQNMIAFSSCFSTCTSEAIPNNYRTLEKRDNGQLNLQTQKVILPTSELNTM 467
drTLR9          LRTLRRVDLSQNMLAFSSCFSTCTSEAEH---QIPERYGNEQFNLQMQELPILNAET--- 461
hsTLR9          FPGLRYVDLSDNRISGASELTATMGEADG----------GEKVWLQPGDLAPAPVDT--- 460
                :  ** *.**:* :: :* :::  .**           . :. **  .:     :
```

Figure 1 (continued)

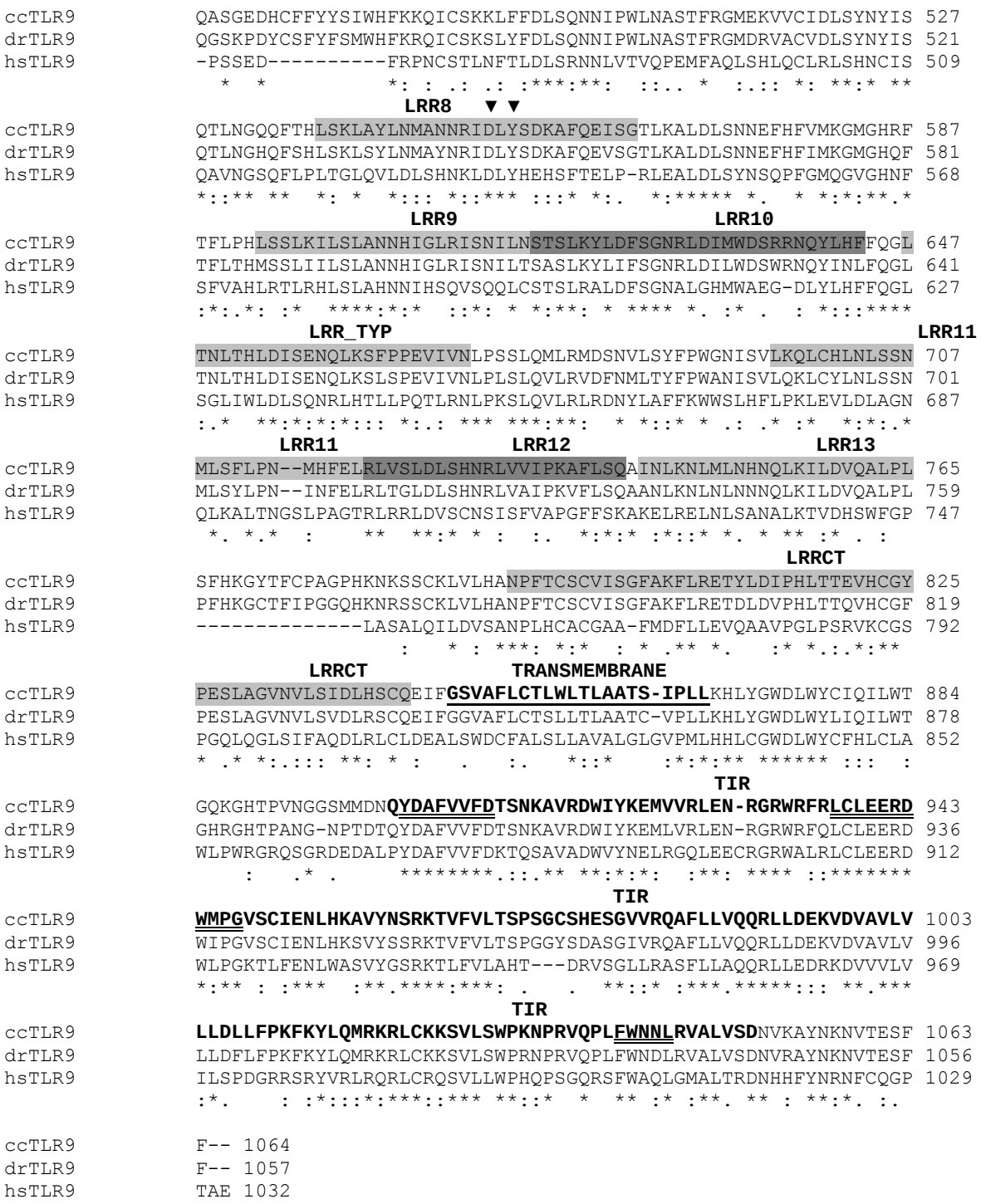


Figure 2 Schematic comparative domain organization for the common carp (cc) TLR9 protein with those of zebrafish (dr), rainbow trout (om) and human (hs) TLR9 molecules (accession numbers GU809229, NP_001124066, NP_001123463, AAZ95519). The domain organization was predicted using the SMART program (http://smart.emblheidelberg.de/). LRR: leucine-rich repeat; LRR TYP: typical leucine-rich repeat; LRR-CT: C (carboxyl)-terminal leucine-rich repeat; TIR: Toll/IL-1 receptor; (▬) signal peptide; (▮) transmembrane domain; (▬) low compositional complexity segment.

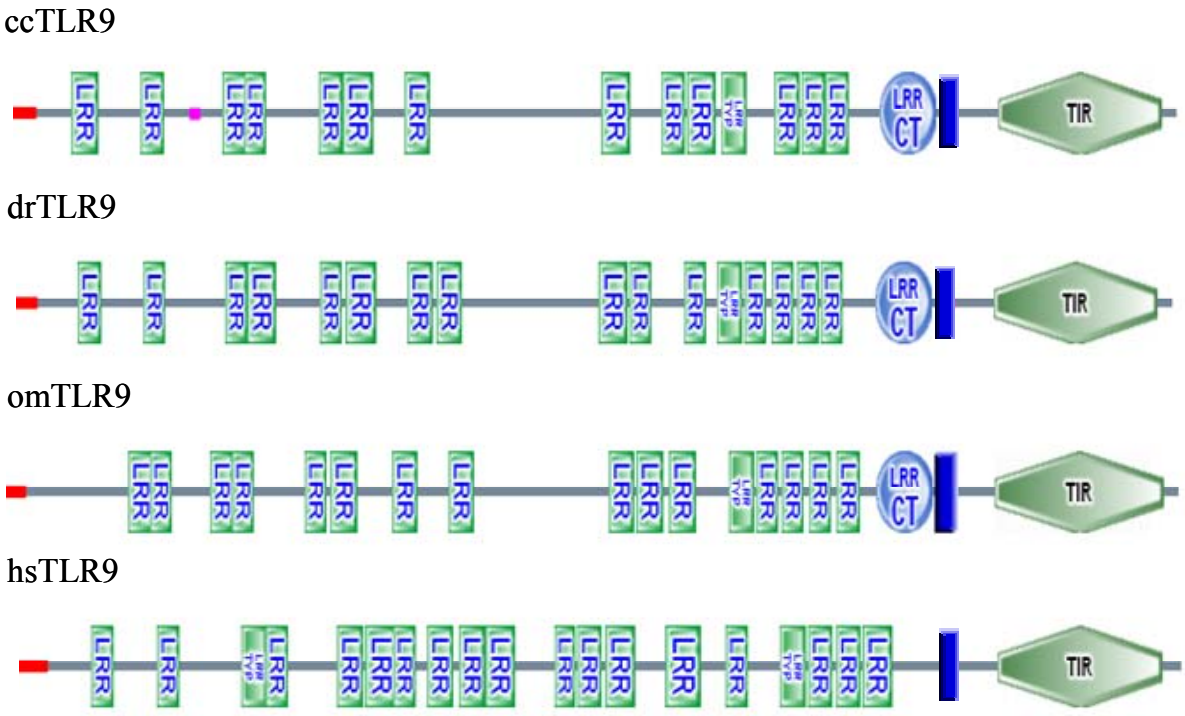


Figure 3 Alignment of common carp (cc), zebrafish (dr) and human (hs) MyD88 amino acid sequences performed using CLUSTALW v2.0. Alignment gaps are marked by dashes. Asterisks (*) and colons (:) below the alignments indicate identity and similarity, respectively, among all species examined. In the common carp sequences, the death domains are marked in grey and TIR domains are underlined. The three conserved short sequence motifs between TIR domains, which are (F/Y)DA, RDXXPG and FW, are in bold font. GenBank accession numbers: common carp GU809230, GU809231; zebrafish AAQ90476; and human Q99836.

```
ccMyD88a        MAS--------KSSIDYEAIPVTALNCSFRKKLGLYLNPTNAVAADWRTVAEMMDFTYLE 52
ccMyD88b        MAS--------KSSIDYEAIPITALNCSFRKKLGLYLNPTNAVAADWRTVAEMMDFTYLE 52
drMyD88         MAS--------KLSLDHEAIPVTALNCSFRKKLGLFLNPTNTVAADWRTVAELMDFTYLE 52
hsMyD88         MAAGGPGAGSAAPVSSTSSLPLAALNMRVRRRLSLFLNVRTQVAADWTALAEEMDFEYLE 60
                **:            . .::*::***  .*::*.*:**  . ***** ::** *** ***

ccMyD88a        IKNFEKREYPFERVLTEWETRPEATVANLLSFLEKAERKDVISDLKDIIDDDCRKYLERQ 112
ccMyD88b        IKNFENREYPFEKVLKEWETRPEATVENLLSILKKAERKDVISDLKDMIDDDCRKYLERQ 112
drMyD88         IKNFEKRDCPFEKVLTDWETRPDATVANLLSILEKAERKDVISELKEILDDDCRKYMERQ 112
hsMyD88         IRQLETQADPTGRLLDAWQGRPGASVGRLLELLTKLGRDDVLLELGPSIEEDCQKYILKQ 120
                *:::*.:  *  ::*  *: ** *:* .**.:* *  *.**: :*   :::**:**: :*

ccMyD88a        LR----KPVQVPVVDSCGPRTQEREGITLCDDPQGLTPETFDAFICYCQNDFQFVHEMIK 168
ccMyD88b        QR----KPVQVPVVDSCGPRTQEREGITLYDDPQGLTPETFDAFICYCKSDFQFVHEMIK 168
drMyD88         QR----KPLQVPVVDSCGPRTQEREGITLYDDPQGLTPETFDAFICYCQSDIQFVHEMIK 168
hsMyD88         QQEEAEKPLQVAAVDSSVPRTAELAGITTLDDPLGHMPERFDAFICYCPSDIQFVQEMIR 180
                 :    **:**..***. *** *  ***  *** *  ** ******** .*:***:***:

ccMyD88a        QLEQTEYNLKLCVFDRDVLPGTCVWTITSELIEKRCKRMVVVISDDYLDSDACDFQTKFA 228
ccMyD88b        ELEQTKYNLKLCVFDRDVLPGTCVWTIASELIEKRCKRMVVVISDDYLDSDACDFQTKFA 228
drMyD88         QLEHTEYNLKLCVFDRDVLPGTCVWTIASELIEKRCKRMVVVISDDYLDSDACDFQTKFA 228
hsMyD88         QLEQTNYRLKLCVSDRDVLPGTCVWSIASELIEKRCRRMVVVVSDDYLQSKECDFQTKFA 240
                :**:*:*.***** ***********:*:********:*****:*****:*. ********

ccMyD88a        LSLCPGARSKRLIPVVYKTMKRPFPSILRFLTICDYTRPCTQAWFWTRLAKALALPRSFQ 288
ccMyD88b        LSLCPGTHSRRLIPVVYKTMEKPFPSILRFLTICDYTRPSTQVWFWTRLAKALSLP---- 284
drMyD88         LSLCPGARTKRLIPVVYKSMKRPFPSILRFLTICDYSKPCTQVWFWTRLAKALSLP---- 284
hsMyD88         LSLSPGAHQKRLIPIKYKAMKKEFPSILRFITVCDYTNPCTKSWFWTRLAKALSLP---- 296
                ***.**:: :****: **:*:: *******:*:***:.*.*: **********:**
```

Figure 4 Alignment of TRAF6 proteins from common carp (cc), zebrafish (dr) and human (hs) performed with CLUSTALW v2.0. Alignment gaps are marked by dashes, and asterisks (*) and colons (:) indicate identity and similarity, respectively, for all species examined. In the common carp sequences, the RING (really interesting new gene) domains are marked in grey and the MATH (meprin and TRAF homology) domains are in bold. GenBank accession numbers: common carp GU985443 and HM535645; zebrafish NP_001038217; and human AAB38751.

```
ccTRAF6a        ---MACSDMEKLSLDVDPYCDSDLSGCAAAMEKERESYLSPTENPSTISVSASIPPDQQG 57
ccTRAF6b        ---MACSDMEKLSLDVDPYCDSDLSGCAAAMEKERESYLSPTENTSTISVSASIPPDQQG 57
drTRAF6         ---MACNDVDKSSFD-DVCCDSGHSSCAAAMEKERESFLSPTENPSTISVSSSMPPDQQG 56
hsTRAF6         MSLLNCENSCGSSQSESDCCVAMASSCSAVTKDD-----SVGGTASTGNLSSSFMEEIQG 55
                   : *.:    * . .  * :  *.*:*. :.:     *   ..** .:*:*:  : **
                                          RING Domain
ccTRAF6a        YDVEFDPPLESKYECPICLMGLRSAVQTPCGHRFCNSCIKKSIRDTGQKCPVDNEVLLEE 117
ccTRAF6b        YDVEFDPPLESKYECPICLMGLRSAVQTPCGHRFCNSCIKKSIRDTGQKCPVDNEVLLEE 117
drTRAF6         YDVEFDPPLESKYECPICLMGLRSAVQTPCGHRFCDSCIRKSIRDTGQKCPVDNEVLLEE 116
hsTRAF6         YDVEFDPPLESKYECPICLMALREAVQTPCGHRFCKACIIKSIRDAGHKCPVDNEILLEN 115
                ********************.**.***********.:** *****:*:*******:***:

ccTRAF6a        QLFPDNFAKREILSLTVRCPNVGCSDKMELRQLEKHLSHCKFATVPCPQCLESVRKSHLD 177
ccTRAF6b        QLFPDNFAKREILSLTVRCPNVGCSDKMELRQLEKHLSQCKFATVQCPQCLESVRKSHLD 177
drTRAF6         QLFPDNFAKREILSLTVKCSNFGCSEKMELRQLEKHLSQCRFATAPCPQCQESVPISHLD 176
hsTRAF6         QLFPDNFAKREILSLMVKCPNEGCLHKMELRHLEDHQAHCEFALMDCPQCQRPFQKFHIN 175
                *************** *:*.* ** .*****:**.* ::*.**   **** ...   *::

ccTRAF6a        EHKSQQCLQRLMTCPACAGSFVYANKQIHEQICPFANTVCEYCEMELIRDQLALHCDTDC 237
ccTRAF6b        EHKSEQCLQRLMTCPACAGSFVYANKQIHEQICPFANTVCEYCEMELIRDQLALHCDTDC 237
drTRAF6         EHKSQHCLQRIMTCPDCAGSFVYAVKQNHEQFCPFANTVCEYCEMELIRDQLALHCDTDC 236
hsTRAF6         IHILKDCPRRQVSCDNCAASMAFEDKEIHDQNCPLANVICEYCNTILIREQMPNHYDLDC 235
                 *  :.* :* ::*  **.*:.:  *: *:* **:**.:****:  ***:*:. * * **

ccTRAF6a        LKAPVACTFSTFGCREKMPRNELAQHMQEFTQMHMRYMAEFLRSQSLSSCPVPSIAAHSS 297
ccTRAF6b        LKAPVACTFSTFGCHVKMPRNELAQHMQEFTQMHMRCMAEFLRSQTLSSCPVPSVAAHSA 297
drTRAF6         LKAPVACTFSTFGCREKMTRNELAQHMQEFTQMHMRYMAEFLRSQTLNNCTMPSAAAHLS 296
hsTRAF6         PTAPIPCTFSTFGCHEKMQRNHLARHLQENTQSHMRMLAQAVHSLSV----IPDSGYIS- 290
                 .**:.********: ** **.**:*:** ** *** :*: ::* ::    :*. .

ccTRAF6a        SDERGATARAPESCQCKPELMNLRETVLELEGRLVRQDQQIRELCIHNETQKNQVTELRR 357
ccTRAF6b        SDERGASARAQDSCQCRQELMNLRETVLELEGRLVRQDQQIRELCIHSETQKNQVTELRR 357
drTRAF6         SDDRGASARSPDSCQCKQELLNLRETVLELEGRLVRQDQQIRELCIHNDTQKNQVTELRR 356
hsTRAF6         ------------------EVRNFQETIHQLEGRLVRQDHQIRELTAKMETQSMYVSELKR 332
                                 *: *::**: :*********:*****  : :**.  *:**:*

ccTRAF6a        KLSSLEEATRELEAQQYQGVYV[WRLENFSLHLRNQEAGQPIVLHSPPFYTGRPGYKLCLR] 417
ccTRAF6b        KLSSLEDATRELEAQQYQGVYV[WRLENFSLHLRNQEAGQPIVLHSPPFYTGRPGYKLCLR] 417
drTRAF6         KLVSLEESTRELEAQQYQGIYVWRVENFSHHLRNQEAGQPIVLHSPPFYTGRPGYKLCLR 416
hsTRAF6         TIRTLEDKVAEIEAQQCNGIYIWKIGNFGMHLKCQEEEKPVVIHSPGFYTGKPGYKLCMR 392
                .: :**: . *:**** :*:*:*:: **. **: **  :*:*:*** ****:******:*
                                       MATH Domain
ccTRAF6a        [LHLQTPNAPRCSNYISLFVHTMQGEFDSQLSWPLQGTIRLAVLDQVEG---EHHVEVMET] 474
ccTRAF6b        [LHLQTPSAPRCSNYISLFVHTMQGEFDSQLSWPLQGTIRLAVLDQVEG---QHHVEVMET] 474
drTRAF6         LHLQTPSAPRCSNFISLFVHTMQGEFDSQLSWPLQGTIRLAVLDQVEG---QHHIEVMET 473
hsTRAF6         LHLQLPTAQRCANYISLFVHTMQGEYDSHLPWPFQGTIRLTILDQSEAPVRQNHEEIMDA 452
                **** *.* **:*:***********:**:*.**:******::*** *.   ::* *:*::

ccTRAF6a        [KPDLQAFQRPTVLRNPKGFGYVTFLHLQVL]QQRGFVKDDVLLVRCEVTPRFDA-SLRREG 533
ccTRAF6b        [KPDLQAFQRPTVQRNPKGFGYVTFLHLQAL]RQRGFVKEDVLLVRCEVTPRFDA-SLRREG 533
drTRAF6         KPDLQAFQRPTVMRNPKGFGYVTFLHLQALRQRGFVKEDVLLVRCEVTPRFDA-SLRREG 532
hsTRAF6         KPELLAFQRPTIPRNPKGFGYVTFMHLEALRQRTFIKDDTLLVRCEVSTRFDMGSLRREG 512
                **:* ******: ***********:**:.*:** *:*:*.*******:.***  ******

ccTRAF6a        VQPRGSEPSL 543
ccTRAF6b        VQPRGSEPSL 543
drTRAF6         VQPRGPEPSI 542
hsTRAF6         FQPRSTDAGV 522
                .***..:..:
```

Figure 5 Phylogenetic trees showing relationships of TLR9 (a), MyD88 (b) and TRAF6 (c) deduced amino acid sequences among common carp and other teleosts and mammals. The trees were constructed by the neighbor-joining method based on the Poisson correction model with 1,000 bootstrap replicates. The numbers at the branches indicate bootstrap values. The bar (0.1) indicates the genetic distance. GenBank accession numbers: TLR9 – common carp GU809229, zebrafish NP_001124066, Atlantic salmon NP_001117125, rainbow trout NP_001123463, grouper ACV04893, yellowback seabream ABY79218, gilthead seabream AAW81697, large yellow croaker ACF60624, fugu AAW69377, Japanese flounder BAE80691, human AAZ95519, pig ACZ82294, mouse BAB19260, and cattle ABN71650; MyD88 – common carp GU809230 and GU809231, zebrafish AAZ16494, channel catfish ACD81929, ayu BAI68385, Atlantic salmon NP_001130017, rainbow trout NP_001117893, fugu NP_001106666, Japanese flounder BAE75959, human AAP36509, pig ABW74617, mouse NP_034981, cattle AAY16578, and chicken ABJ98316; TRAF6 – common carp GU985443 and HM535645, zebrafish NP_001038217, ayu BAI68387, three-spined stickleback ABJ15863, human NP_004611, pig NP_001098756, mouse NP_033450, and cattle NP_001029833.

a.

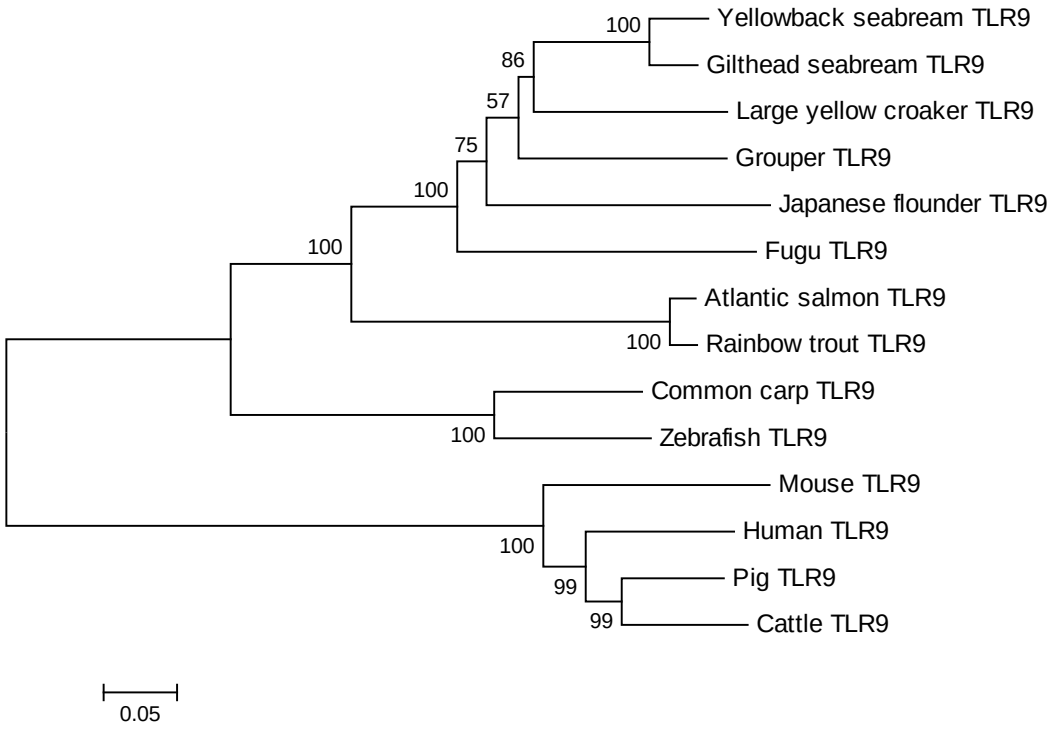
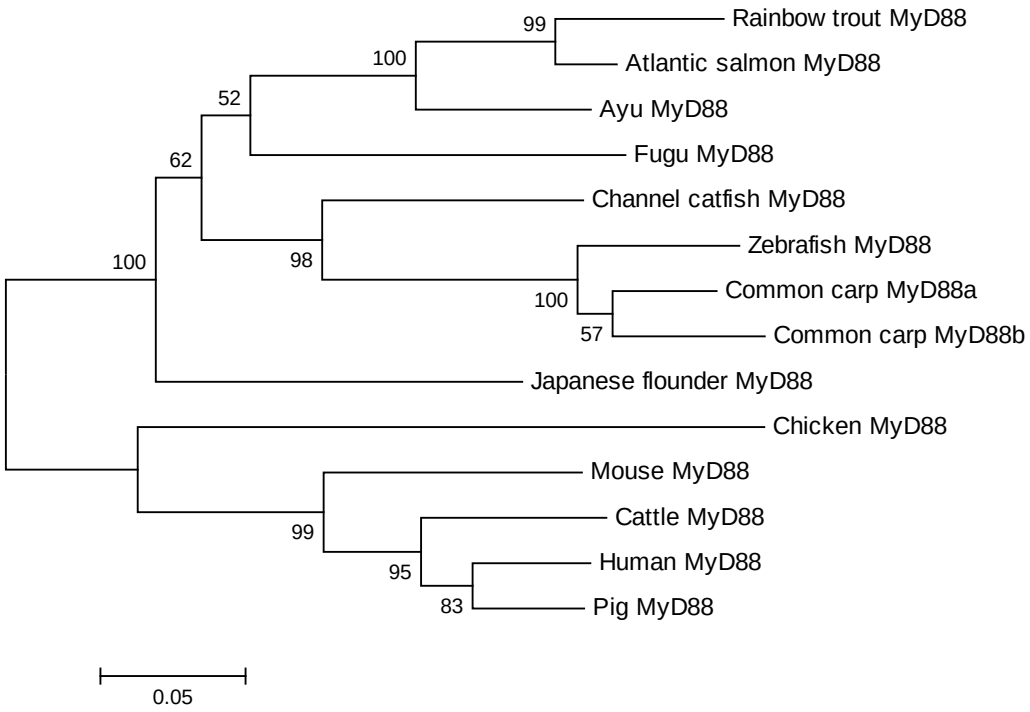


Figure 5 (continued)

b.



5c.

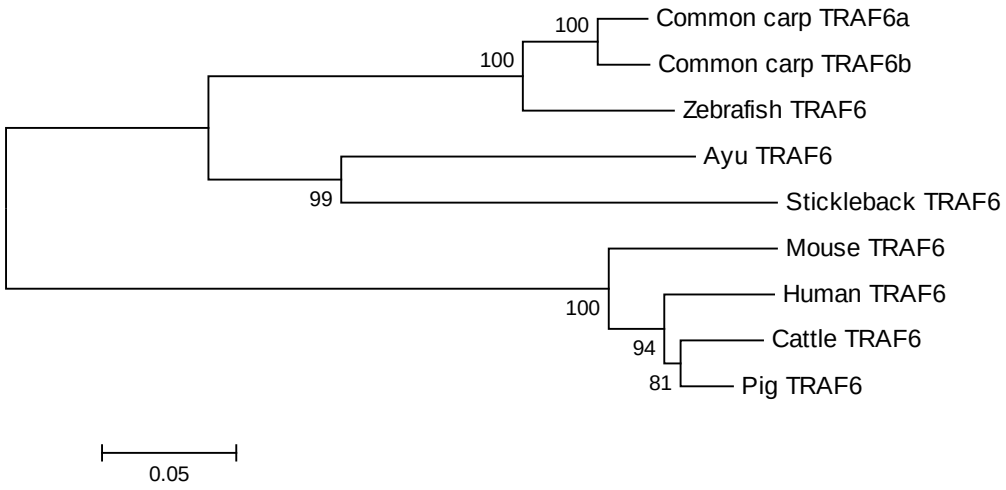
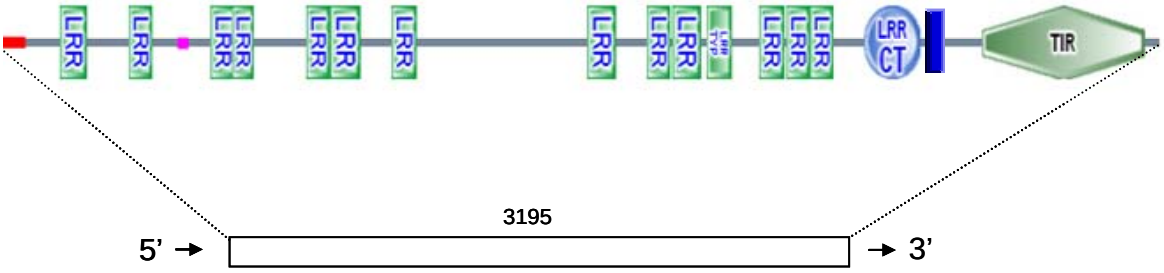
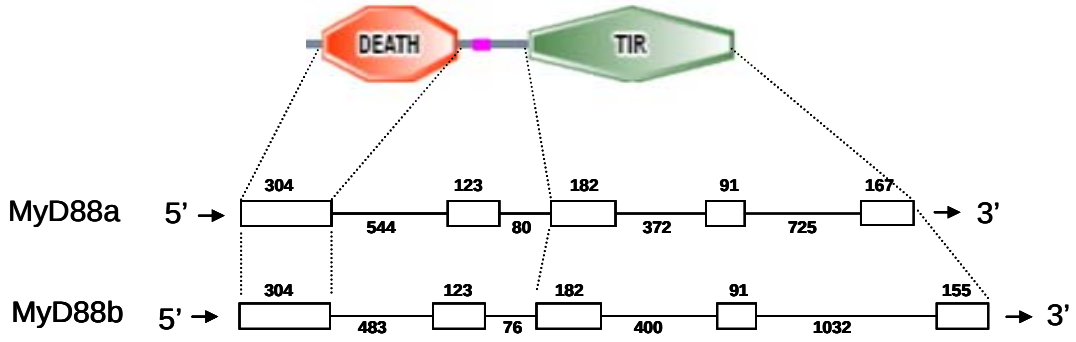


Figure 6 Schematic representations of protein domains and genomic structures of common carp TLR9 (a), MyD88 (b) and TRAF6 (c). Protein domains were predicted using SMART. Exons are represented by boxes, with the exon sizes (bp) shown above the boxes. An exon with partial genomic sequence is represented by a dashed box (panel c). The lines between exons represent introns, with intron sizes (bp) indicated below the lines.

a.



b.



c.

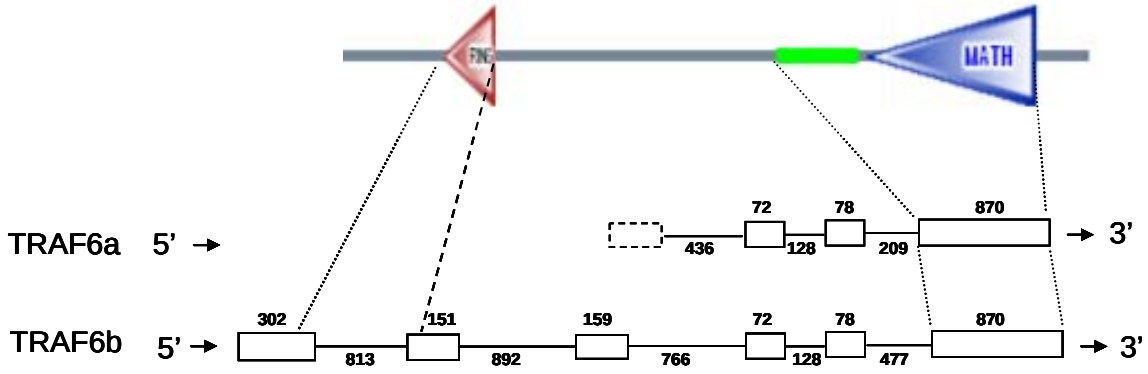
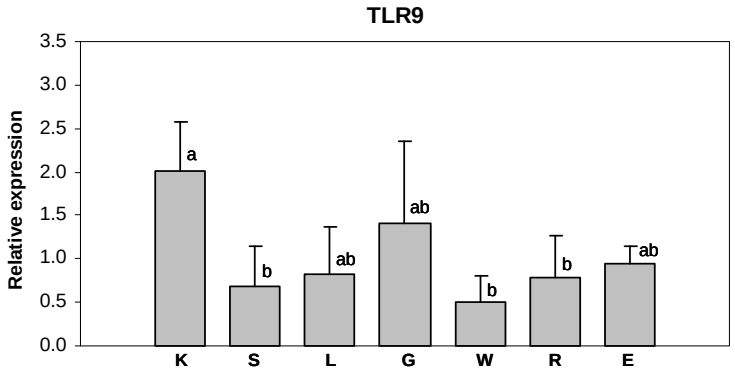


Figure 7 Tissue distribution of TLR9 (a), MyD88 (b) and TRAF6 (c) mRNA transcripts in adult common carp analyzed by real-time PCR. Bar graphs represent the mean expression levels (n = 5) of target transcript in kidney (K), spleen (S), liver (L), gill (G), white muscle (W), red muscle (R), and eye (E) relative to a reference sample (pooled cDNA), and normalized to β-actin gene transcript. In panels b-c, grey bars represent MyD88a and TRAF6a and white bars represent MyD88b and TRAF6b. In each target transcript, the bars with the same letter are not significantly different as evaluated using the Tukey-Kramer HSD test. An asterisk (*) indicates significant difference by LSD test ($p < 0.05$) between paralogous transcripts in the same tissue. The threshold cycles (Ct values) of the reference sample were 32.7, 27.2, 28.3, 28.3 and 27.0 for TLR9, MyD88a, MyD88b, TRAF6a and TRAF6b, respectively.

a.



b.

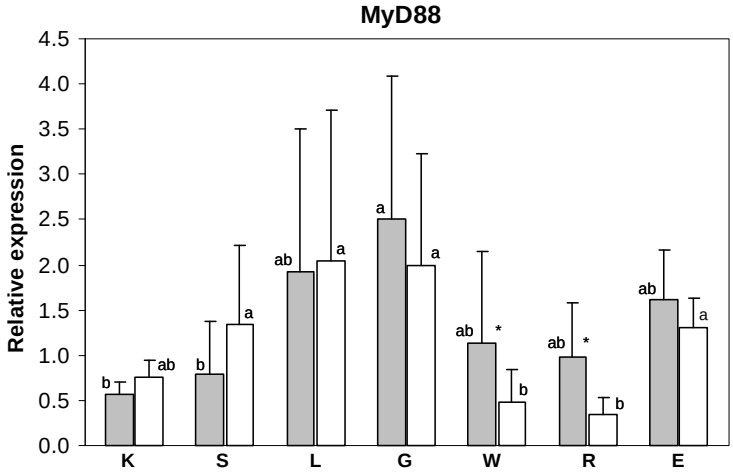


Figure 7 (continued)

c.

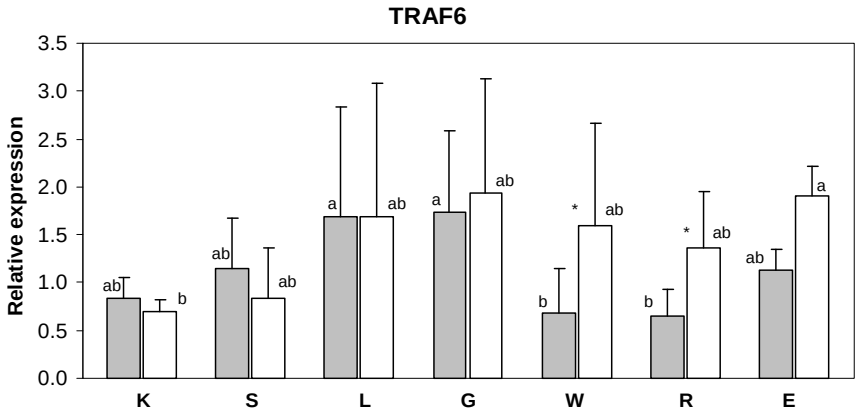
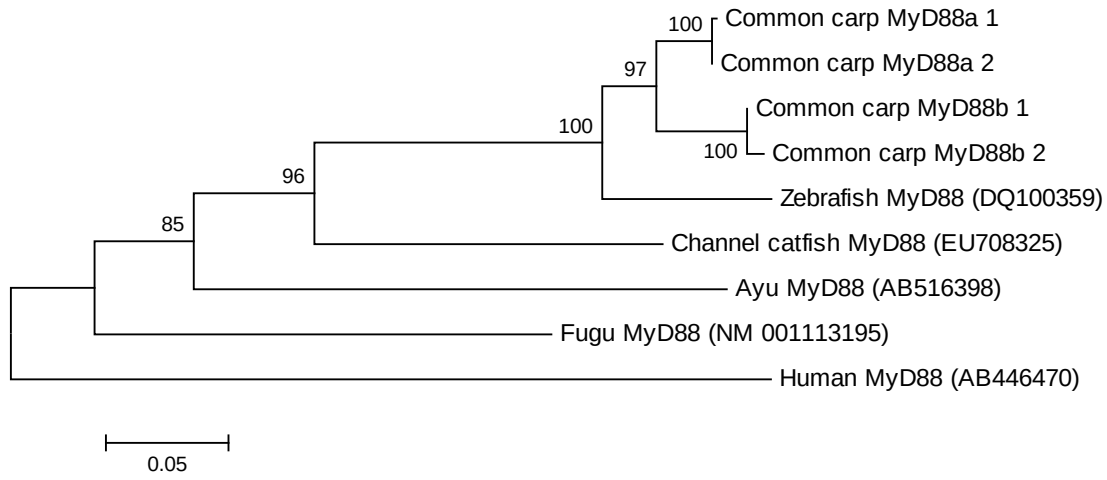
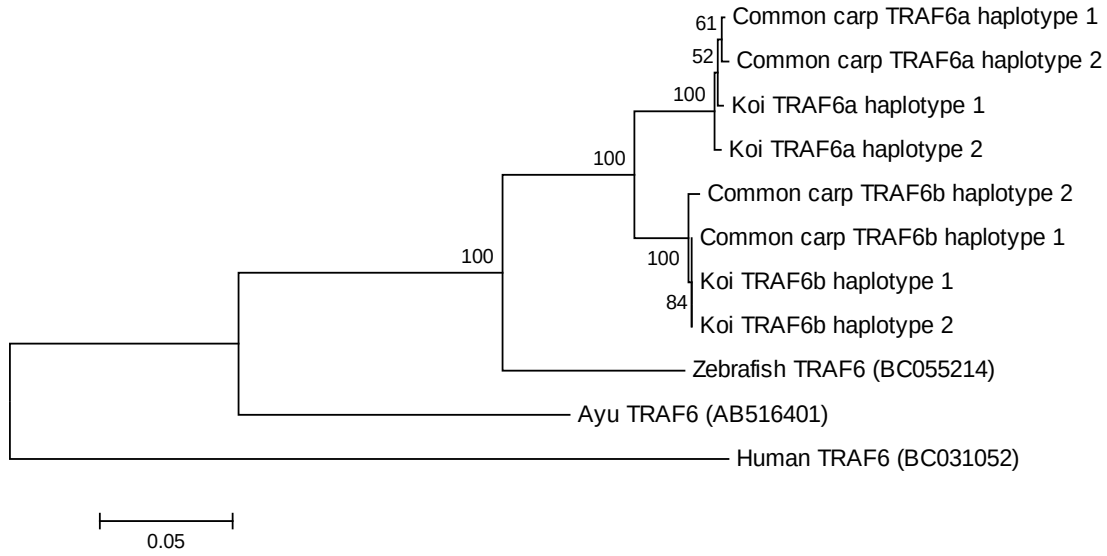


Figure 8 Phylogenetic trees of MyD88 (a) and TRAF6 (b) genes constructed by the neighbor-joining method using nucleotide sequences. The common carp MyD88a and MyD88b sequences of the first three exons (609 bp) were used in the phylogenetic analysis. The two paralogous sequences differed by 34 nucleotides with 1 and 3 putative SNPs within MyD88a and b, respectively. The common carp TRAF6a and TRAF6b gene fragments (566 bp) differed by 27 nucleotides with 6 and 3 putative SNPs within TRAF6a and b, respectively. The numbers at the branches indicate bootstrap values based on 1,000 replicates. Accession numbers of nucleotide sequences obtained from NCBI database are given in parentheses.

a.



b.



CHAPTER 4

Association between Single Nucleotide Polymorphisms in Innate Immune Response Genes and Resistance to Cyprinid Herpesvirus-3 Infection in Common Carp (*Cyprinus carpio*)

ABSTRACT

Gene polymorphisms and disease association analysis are essential for identifying putative candidate genes affecting disease susceptibility or resistance. In this chapter, results of a preliminary association analysis of SNPs in common carp immune response genes with CyHV-3 resistance are reported. Twenty-three SNPs from 14 genes were genotyped in the progeny of a mapping family which was subjected to a CyHV-3 challenge. The challenge was carried out by cohabitation of naïve fish with CyHV-3-infected fish. Fish dying within 15 days after virus exposure were scored as susceptible and the survivors were scored as resistant. Fisher exact probability and chi-square tests were used to assess the association of SNP genotypes or alleles with CyHV-3 resistance. For the IL10a SNPs, individuals carrying the CGA haplotype and the CGA homozygous genotype were significantly associated with CyHV-3 resistance (p = 0.001 and 0.009, respectively). One of the IL10a SNPs was a non-synonymous substitution (Gln71Glu); however, the functional role of this amino acid substitution is currently unknown. It should be noted that this analysis was based on a small sample size (n = 28 – 47); therefore, the result should be interpreted with caution, and a study with a larger number of samples is necessary to confirm this finding. No significant association of SNP genotype or allele with the resistance to CyHV-3 was identified in the other 13 genes examined.

KEYWORDS: Association analysis, disease resistance, single nucleotide polymorphisms, innate immune response genes, cyprinid herpesvirus-3, common carp, *Cyprinus carpio*

INTRODUCTION

The causative agent of massive mortality of farmed common carp and koi (*Cyprinus carpio*) occurring in Europe and Israel during the late 1990s has been identified as a DNA virus termed koi herpesvirus (KHV) or cyprinid herpesvirus-3 (CyHV-3) (Hedrick et al. 2000). CyHV-3 is specific to common carp and koi in which infection could lead to 80 – 100% mortality, but does not cause disease in other cyprinid or non-cyprinid species (Ilouze et al. 2006). Despite its extreme virulence, Shapira et al. (2005) found differential resistance to the CyHV-3 virus among commercial and wild common carp strains and crosses, implying that genetic variation influences susceptibility to the CyHV-3 infection in this species.

Genetic variations due to single-base mutations in immune-related genes have been identified as a risk factor for susceptibility to infectious diseases in human (Leoratti et al. 2008; Cardoso et al. 2010; Motsinger-Reif et al. 2010; Pontillo et al. 2010; Pulido et al. 2010) and in livestock (Ye et al. 2006; Kataria et al. 2010; Koets et al. 2010; Ruiz-Larrañaga et al. 2010). Such genetic risk in teleosts has not been documented to date. However, a number of studies on fish immune-related genes reported conserved structural and functional characteristics between fish and mammalian genes, suggesting that their functions in the respective immune systems may be conserved across species. Therefore, genetic alterations in immune response genes which are known to be involved in stimulatory or regulatory functions in the mammalian immune system may influence susceptibility to disease in fish, including a viral disease caused by CyHV-3 in common carp.

In previous chapters, I described the isolation, characterization, and detection of polymorphisms at several common carp immune response genes, particularly toll-like receptor

(TLR) and TLR-associated genes, and the development of SNP markers for genotyping allelic variations of the genes isolated. TLRs are important pathogen recognition receptors that activate the innate immune response against infection, and several TLRs are known for their ability to recognize viral pathogen-associated molecular patterns (PAMPs) (reviewed by Kumar et al. 2009). Cytoplasmic adaptor molecule myeloid differentiation primary response protein 88 (MyD88) and tumor necrosis factor receptor-associated factor 6 (TRAF6) are essential signal transducers in the TLR signaling pathways (Häcker et al. 2000; Kobayashi et al. 2003; Nishiya et al. 2007). Cytokines are soluble proteins important in regulating inflammatory response. Because these molecules are presumed to play a role in antiviral immunity, I investigated whether genetic variants in immune and inflammatory response genes are associated with resistance to CyHV-3 infection in common carp. I present in this chapter a preliminary association analysis of polymorphisms in common carp immune response genes with CyHV-3 resistance in a mapping family using SNP genotype data and survival data from a challenge experiment conducted by a collaborative research group at the Hebrew University of Jerusalem's Department of Animal Sciences. The analysis of SNPs and disease association would assess the usefulness of the candidate antiviral gene markers for marker-assisted breeding to improve genetic resistance to the CyHV-3 in common carp. In addition, understanding the relationship between variations in immune response genes and disease susceptibility will provide a fundamental insight into host-pathogen relationships and the functions of genes in the fish immune system.

MATERIALS AND METHODS

Experimental fish and challenge experiment

Fish for the challenge experiment (a mapping family) were produced by crossing a koi female to a common carp male produced from a cross between a domesticated commercial strain (Yugoslavian) and a wild-caught strain (Sassan). Progeny were grown to an average weight of 15 grams in a virus-free environment. All progeny were tagged with a passive integrated transponder (PIT) tag for individual identification. Two days prior to the CyHV-3 challenge, fish were treated with formaldehyde and malachite green to remove ectoparasites, followed by Doxycyclin antibiotic treatment to reduce the risk of mortality due to factors other than the virus. The viral challenge was carried out by cohabitation; 64 experimental fish (naïve progeny) and CyHV-3-infected sick fish were stocked in the same tank at a permissive temperature of 20°C. Mortality of experimental fish was monitored daily, and dead fish were removed from the tank. The presence of the virus in dead fish was verified in blood DNA by PCR using virus-specific primers. The challenge experiment was finished at day 24 post-exposure.

PCR amplification and SNP genotyping

Genomic DNA was extracted from fin clips using the standard phenol/chloroform method. Common carp gene fragments were amplified by polymerase chain reaction (PCR) using the gene-specific primers listed in Table 1. Each PCR reaction was carried out using 1X PCR buffer, 2 mM $MgCl_2$, 200 µM dNTPs, 1 µM each primer, 0.5 U AmpliTaq Gold DNA polymerase and 20 ng genomic DNA in a total volume of 10 µl. Cycle conditions were: initial denaturation at 94°C for 10 min; 35 cycles of 94°C for 10 s, annealing (58 °C, 60 °C or 62°C) for

30 s, and 72°C for 1 min; followed by final extension at 72°C for 10 minutes. PCR product was purified using MinElute 96 UF PCR Purification Kit (Qiagen). Single-base extension primers developed for genotyping SNPs in common carp immune response genes (Table 1) were used to genotype progeny of a mapping family. Genotyping of SNPs was performed using the SNaPshot Multiplex Kit (Applied Biosystems). SNaPshot reactions (10-µl volume) were prepared using 1 µl SNaPshot Multiplex Ready Mix, 0.5 µl 5X Sequencing Buffer, 0.2 µl extension primer (10 µM), 2.3 µl nuclease-free water, and 6 µl purified PCR product. Cycle conditions were: 96 ^{o}C for 3 min; and 25 cycles of 96 ^{o}C for 10 s, 50 ^{o}C for 5 s, and 60 ^{o}C for 30 s. The reactions then were purified with 1 U of shrimp alkaline phosphatase (SAP) at 37 ^{o}C for 60 min and 75 ^{o}C for 15 min. The reactions were diluted 1:10 with nuclease-free water, and 1 µl of diluted products were mixed with 9 µl Hi-Di deionized formamide and 0.1 µl LIZ 120 size standard (Applied Biosystems). Samples were denatured at 95 ^{o}C for 5 min, then cooled to 4 ^{o}C, and run on an ABI 3730 Automated Sequencer. The software GeneMapper version 3.7 (Applied Biosystems) was used to analyze the individual genotypes.

Data Analysis

Experimental fish that were subjected to viral challenge were scored "susceptible" if they died within 15 days and "resistant" if they survived after 15 days following exposure to CyHV-3. Numbers of resistant and susceptible fish were counted for each genotype class. The Fisher exact test and chi-square test were used to assess association between SNP genotype and CyHV-3 resistance, and between allele/haplotype and CyHV-3 resistance based on the binary survival score.

RESULTS AND DISCUSSION

Polymorphisms and segregation of SNP alleles in a mapping family

Twenty-three of 65 SNPs in common carp immune response genes that were polymorphic in a SNP discovery panel and in a full-sib family in a previous study were also informative in the mapping family. The SNPs for TLR3a, TLR7b and MyD88b genes were not polymorphic in this family. Genotype frequencies at 23 polymorphic SNPs are summarized in Table 2. Of the 23 polymorphic SNPs, 19 were in coding regions and 4 were in introns of genes. All SNPs in coding regions, except for a SNP in TLR22b and IL10a, were synonymous. TLR22b and IL10a non-synonymous SNPs resulted in isoleucine→methionine and glutamine→glutamic acid substitutions, respectively. Chi-square tests for goodness-of-fit to a 1:1 ratio where one parent was heterozygous, and a 1:2:1 ratio where both parents were heterozygous at a 0.05 level of significance revealed that the segregation of SNP alleles in TLR2, TLR3b, TLR4a, TLR4b, TLR9, TLR22, TRAF6a, TRAF6b, IL1β, IL10a and IL10b genes followed Mendelian segregation patterns, whereas the segregation of SNPs in TLR7a and TLR21b genes did not fit expected ratios ($p < 0.05$). The distorted ratios may be due to the small numbers of progeny genotyped.

Association between SNPs and resistance to CyHV-3 infection

The first mortality following virus exposure of progeny from the mapping family was observed on day 11 and mortality peak was on day 13 (Figure 1). Fish that died during days 11 – 15 were scored as susceptible, and fish that died after day 17 (the second blood sampling) were scored as resistant. Of the total 64 fish, 44 were susceptible and 20 were resistant.

Of the 50 progeny genotyped for 23 SNPs, there were 28 individuals with survival data. The number of susceptible and resistant individuals genotyped with all 23 SNPs is shown in Table 3. Association of immune response gene polymorphisms and survival of experimental fish following viral challenge was assessed using the Fisher exact probability test. Associations of SNPs with susceptibility or resistance to CyHV3 infection were not significant (p-value > 0.1), except for three SNPs in IL10a (p-value = 0.003).

To investigate whether increased sample size could reveal significant association of polymorphisms at putative candidate genes and resistance to CyHV-3, an additional 19 progeny with phenotype data were genotyped for one TLR3b, one TLR4b, three TLR9, one MyD88a, and three IL10a SNPs. With an increased number of progeny genotyped, chi-square test results revealed no significant association of gene polymorphisms to CyHV-3 resistance, except for IL10a. At the IL10a gene, however, a highly significant association was found (p-value < 0.01) in which the progeny carrying the homozygous CGA/CGA genotype were associated with CyHV-3 resistance, and the progeny having the homozygous GCG/GCG genotype were associated with susceptibility to CyHV-3 (Table 4). Haplotype analysis showed that the CGA haplotype was associated with resistance and the GCG haplotype conferred susceptibility to CyHV-3 infection ($\chi 2 = 10.17$, $p = 0.001$). Sequence analysis of SNPs in IL10a revealed that the 219C/G mutation resulted in a non-synonymous substitution (71Gln/Glu) in the IL10 protein. The presence of the Glu allele was associated with CyHV-3 resistance. Although the functional role of this amino acid substitution is unknown, results of the association analysis suggested that allelic variations in the IL10a gene might be involved in resistance or susceptibility to CyHV-3 infection.

Recently the genome of CyHV-3 was completely sequenced (Aoki et al. 2007). It was found that the CyHV-3 genome contains genes similar to host genes that have a crucial role in regulating immune response, including IL-10 and TNFR. IL-10 is considered an anti-inflammatory cytokine and has a significant role in counteracting inflammatory action in human immune cells (Benítez et al. 2007). Kotler and Ilouze (unpublished data) compared the infection of wild type, attenuated, and IL-10-knockout CyHV-3 in common carp, and found that IL-10-knockout virus was less effective in infection and/or propagation compared to wild type and attenuated strains. This finding supports the importance of the IL-10 gene in CyHV-3 infection. It is possible that the IL-10-like molecule released by the virus diminishes the inflammatory cytokines released by host cells, which benefits the virus in spreading in the host. I hypothesize that the SNPs in the common carp IL-10a gene may be associated with lower level of IL-10 production. Increased IL-10 release may reduce inflammatory response against infection, which increases the likelihood of viral infection or distribution, and ultimately results in early mortality of fish infected with the virus. An alternative hypothesis would be that IL-10a is not effective in activating the IL-10 receptor (IL-10R) signal to suppress inflammatory response, but rather its main role is competitive binding as a negative regulator of IL-10R signaling. In that case, the non-synonymous mutation which occurred in a receptor-binding region likely affects the IL-10a receptor-binding efficiency so that the common carp IL-10a with the Glu allele has stronger affinity to the IL-10R than the IL-10a with the Gln allele and therefore also competes better with the IL10-like molecule from the virus in binding to the IL-10R.

In conclusion, I have demonstrated the importance of genetic variations in common carp immune response genes that were isolated and characterized in previous studies. Results from

this study suggest that mutations in common carp immune response genes may affect susceptibility to infectious diseases as previously observed in human and livestock.

It should be noted that the sample size (n = 47) was not large enough to provide strong evidence of significant association between polymorphisms in the IL-10a gene and CyHV-3 resistance. On the other hand, for those genes for which no association was detected, or for the genes with a suggestive association (TLR4b; *p*-value = 0.11 and MyD88a; *p*-value = 0.15), increasing sample size would increase the statistical power of the test to exclude their association with disease resistance. Therefore, a study with a larger number of samples per family and from several families is recommended to test for association of polymorphisms in putative candidate genes, i.e., IL10a, TLR4b and MyD88a, and other immune response genes that detect DNA viruses, e.g., TLR9 and TLR21 with resistance to CyHV-3 infection in common carp. For the genes that revealed significant association with resistance to the virus, it would be interesting to elucidate whether the mutations alter gene expression or structure/function of proteins that might cause fish immune response to be less effective following infection. Recently, it was found that the CyHV-3 could also infect and replicate in other cyprinids such as goldfish, but did not develop clinical signs of disease (El-Matbouli and Soliman, in press). It would also be interesting to screen and compare polymorphisms of the genes that confer resistance to infection in CyHV-3-resistant cyprinids. The finding will provide insight into the gene functions in fish antiviral immunity.

The purpose of gene polymorphisms and disease resistance/susceptibility association analysis is to employ the genetic variants as candidate markers for marker-assisted selection (MAS). Development of genetic markers as a prerequisite for detection of quantitative trait loci (QTL) and the application of MAS in aquaculture species have been previously described by

Poompuang and Hallerman (1997) and Korol et al. (2007). SNP variants that showed strong association with disease resistance may be used directly as genetic markers for selecting resistant broodstock. Candidate genes with suggestive associations may indicate that they are located close to the causative loci and may be useful for fine-mapping of QTLs influencing resistance to CyHV-3 infection.

LITERATURE CITED

Aoki, T., I. Hirono, et al. (2007). "Genome sequences of three koi herpesvirus isolates representing the expanding distribution of an emerging disease threatening koi and common carp worldwide." Journal of Virology **81**(10): 5058-5065.

Benítez, S., C. Bancells, et al. (2007). "Pro-inflammatory action of LDL(-) on mononuclear cells is counteracted by increased IL10 production." Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids **1771**(5): 613-622.

Cardoso, C., A. Pereira, et al. (2010). "IFNG +874 T>A single nucleotide polymorphism is associated with leprosy among Brazilians." Human Genetics **128**(5): 481-490.

Häcker, H., R. M. Vabulas, et al. (2000). "Immune cell activation by bacterial CpG-DNA through myeloid differentiation marker 88 and tumor necrosis factor receptor–associated factor (Traf) 6." The Journal of Experimental Medicine **192**(4): 595-600.

El-Matbouli, M. and H. Soliman "Transmission of cyprinid herpesvirus-3 (CyHV-3) from goldfish to naïve common carp by cohabitation." Research in Veterinary Science, **In Press**.

Hedrick, R. P., O. Gilad, et al. (2000). "A herpesvirus associated with mass mortality of juvenile and adult koi, a strain of common carp." Journal of Aquatic Animal Health **12**(1): 44-57.Ilouze, M., A. Dishon, et al. (2006). "Characterization of a novel virus causing a lethal disease in carp and koi." Microbiology and Molecular Biology Reviews **70**(1): 147-156.

Kataria, R., R. Tait, et al. (2010). "Association of toll-like receptor four single nucleotide polymorphisms with incidence of infectious bovine keratoconjunctivitis (IBK) in cattle." Immunogenetics: 1-5.

Kobayashi, T., P. T. Walsh, et al. (2003). "TRAF6 is a critical factor for dendritic cell maturation and development." Immunity **19**(3): 353-363.

Koets, A., W. Santema, et al. (2010). "Susceptibility to paratuberculosis infection in cattle is associated with single nucleotide polymorphisms in toll-like receptor 2 which modulate immune responses against *Mycobacterium avium* subspecies *paratuberculosis*." Preventive Veterinary Medicine **93**(4): 305-315.

Korol, A., A. Shirak, et al. (2007). "Detection and analysis of quantitative trait loci (QTL) for economic traits in aquatic species." In Aquaculture Genome Technologies, Z. Liu (ed.), Blackwell Publishing Ltd.

Kumar, H., T. Kawai, et al. (2009). "Toll-like receptors and innate immunity." Biochemical and Biophysical Research Communications **388**(4): 621-625.

Leoratti, F., L. Farias, et al. (2008). "Variants in the toll-like receptor signaling pathway and clinical outcomes of malaria." The Journal of Infectious Diseases **198**(5): 772-780.

Motsinger-Reif, A., P. Antas, et al. (2010). "Polymorphisms in IL-1beta, vitamin D receptor Fok1, and toll-like receptor 2 are associated with extrapulmonary tuberculosis." BMC Medical Genetics **11**(1): 37.

Nishiya, T., E. Kajita, et al. (2007). "Distinct roles of TIR and non-TIR regions in the subcellular localization and signaling properties of MyD88." FEBS Letters **581**(17): 3223-3229.

Pontillo, A., L. Brandão, et al. (2010). "A 3'UTR SNP in NLRP3 gene is associated with susceptibility to HIV-1 infection." Journal of Acquired Immune Deficiency Syndromes **54**(3): 236-240.

Poompuang, S. and E. M. Hallerman (1997). "Toward detection of quantitative trait loci and marker-assisted selection in fish." Reviews in Fisheries Science **5**: 253-277.

Pulido, I., M. Leal, et al. (2010). "The TLR4 ASP299GLY polymorphism is a risk factor for active tuberculosis in Caucasian HIV-infected patients." **8**(3): 253-258.

Ruiz-Larrañaga, O., J. M. Garrido, et al. (2010). "Genetic association between bovine NOD2 polymorphisms and infection by *Mycobacterium avium* subsp. *paratuberculosis* in Holstein-Friesian cattle." Animal Genetics: **41**(6):652-655.

Shapira, Y., Y. Magen, et al. (2005). "Differential resistance to koi herpes virus (KHV)/carp interstitial nephritis and gill necrosis virus (CNGV) among common carp (*Cyprinus carpio* L.) strains and crossbreds." Aquaculture **245**(1-4): 1-11.

Ye, X., S. Avendano, et al. (2006). "Association of twelve immune-related genes with performance of three broiler lines in two different hygiene environments." Poultry Science **85**(9): 1555-1569.

Table 1 PCR and single-base extension primer sequences used for SNP genotyping at given genes of common carp. AT = Annealing temperature.

Primer name	Sequence (5' - 3')	Size (bp)	AT (^{o}C)
TLR2-F1	CCAGGACATGTGATTGTGACCAGC	530	60
TLR2-R1	GACTAAACCAATGGGACGAGCTGC		
TLR2-S59A/G-R	GCCAAGCGCATTCGCTGGAACTTTTGGGACCTG		
TLR2-F2	AGATGGGCTCACTCATCTGGACG	569	60
TLR2-R2	ACAGTGGGGTAGTTTCCCGTGGA		
TLR2-S201T/C-F	CACCCTGAAACTCTGCTGATAATCAAAGATGT		
TLR3b-F	GGATTCCCTCAAAGTTACATCTGC	430	58
TLR3b-R	GGATTTTCCTAGACTGTCCCATGT		
TLR3b-S336A/G-F	GTGGAACGAAGCTTGCTCCCCTTAGAGGATGA		
TLR4a-F	TGCCATGTGGTAGAGTTACATCCA	567	58
TLR4a-R	GACCCAGACTTCATCATAGCTGGA		
TLR4a-S291G/A-R	CTTGGTGAAAACGTTTTACAGAAAATATTTTCAGGTTG		
TLR4b-F	GGTTCTTGATATCTCACACTGTGGCA	586	58
TLR4b-R	GACCCAGACTTCATCATAGCTGGA		
TLR4b-S88CT-R	GCCAGAAAATCCAAAATCATCAACTTGTTTCG		
TLR7a-F	GCACCTTTTCCATGCGTTCC	381	58
TLR7a-R	CGGTTGGATGTCTTCCTGCT		
TLR7a-S348G/A-F	GCTGAAGGTCCTAAGAATCAGAGGGTATGT		
TLR9-F1	CGACTCCACAAGCTCCAAGAGTTAGATT	775	58
TLR9-R1	CCTTGTCAGAGTACAGGTCAATGC		
TLR9-S190T/C-F	ATCCATAGAGAAGATGACGGCTCTACGAGAG		
TLR9-S723A/G-F	GACTTTAAATGGCCAGCAGTTCACACATCTAAGCAA		
TLR9-F2	CAGATGCTGAGAATGGACTCT	503	58
TLR9-R2	TGGCATGAGTGAAGATCTATCG		
TLR9-S372T/A-F	GCATGCCAATCCTTTTACTTGTAGTTGTGT		
TLR21-F	GCAGTTGTCCTCTCAGTTGTGGCTG	404	58
TLR21-R	AGAACCCTCGAGGTTCGGCAACAG		
TLR21-S185A/G-F	AGACATAGCACTCTACCTGTTTTCCTCAAC		
TLR21-S272A/G-F	ACTTTACTGGAAATTTAAGTATGGCTACTATGTGTTTCG		
TLR21-S365C/T-F	ATGCCTTTGTTTCCTACAACTCAGCAGATGAAGA		
TLR22-F	ACCATCTGCCTGATCACGCACAA	333	58
TLR22-R	ACAATCGGGCTACACTGATCT		
TLR22-S36T/C-F	TCTGCCTGATCACGCACAACTACCTGAAGAG		
TLR22-S282C/G-R	TAAATATGATCTTCATCTTTCCCAGAAAG		
MyD88a-F	TGACTTCCAGTTTGTGCATGAG	281	58
MyD88a-R	ATGTCCACTATGTAGAATGGCT		
MyD88a-S142T/G-F	TCACCAGTGAGCTCATAGAAAAAAGGCAAGT		
TRAF6a-F	AGATCCGGGAGCTGTGCATCC	557	60
TRAF6a-R	TGTCACCTCGCAGCGCACCAAC		
TRAF6a-S140G/A-R	CCTCTTGGTTGCGCAGGTGAAGTGAGAAGTT		
TRAF6a-S380C/T-F	CACGATCCGACTGGCAGTGTTGGACCAGGT		
TRAF6b-F	ACTCGTGTTCAGCAAGGATGCA	414	60
TRAF6b-R	GCATCTGTGTAAATTCCTGCATGTGCT		
TRAF6b-S272TC-F	TAAATGCAGCCTTGTAGAATATAAGAGACT		

Table 1 (continued)

Primer name	Sequence (5' - 3')	Size (bp)	AT (^{o}C)
IL1β-F	CTGTGACGCTGAGTGCTGGAGCAATG	540	62
IL1β-R	TTCGGGTGGTTGGCATCTGGTT		
IL1β-S119T/C-F	TGCATGACTTCCGAAATGTCCTTTTGACGGTTCCCCCTAC ACATTTTA		
IL10a-F1	CGACTGCAAGACTGACTGTTGCT	463	60
IL10a-R1	GTGCAGATGCAAGATGTTAAGCAG		
IL10a-S52G/C-F	CTCATTTGTGGAGGGCTTTCCAGTGAGACT		
IL10a-S290C/G-F	GACATGGAACCATTACTGGACGAAAACGTG		
IL10a-F2	TGCACATTGCTACAGCGGTGG	401	60
IL10a-R2	CTAGCCTAAACTGGTTTGCTGC		
IL10a-S374G/A-F	AACCAGCCACTTACATGCCCAACCAGGCT		
IL10b-F	ACTAGTGCCTGTTGTGCAATGGTA	318	62
IL10b-R	GAGGATATCAAGCTCTCCCATGGCT		
IL10-2-S92T/G-R	TATGCTGGCGATCTCAAGGGGATTTTGGCA		

Table 2 SNPs in common carp immune response genes, genotype frequencies, and results of chi-square tests for goodness-of-fit of SNP allele segregation in a mapping family.

Gene	SNP Position [a]	Amino acid substitution	N	Genotype frequencies	χ^2
TLR2	129 A/G	No change	48	0.50 A/A 0.50 A/G	0
(FJ858800)	702 T/C	No change	50	0.50 T/T 0.50 T/C	
TLR3b	669 A/G	No change	69	0.48 A/A 0.52 A/G	0.13
(GU321981)					
TLR4a	1781 G/A	No change	49	0.55 G/G 0.45 G/A	0.51
(GU321982)					
TLR4b	1577 T/C	No change	69	0.58 T/T 0.42 T/C	1.75
(HQ229652)					
TLR7a	349 G/A	No change	50	0.36 G/G 0.64 G/A	3.92*
(GU321983)					
TLR9	1382 T/C	No change	69	0.39 T/T 0.61 T/C	3.26
(GU809229)	1915 A/G	No change	69	0.39 A/A 0.61 A/G	
	2692 T/A	No change	69	0.39 T/A 0.61 A/A	
TLR21b	200 A/G	No change	50	0.34 A/G 0.66 G/G	5.12*
(HQ323757)	287 A/G	No change	50	0.34 A/G 0.66 A/A	
	380 C/T	No change	50	0.34 C/T 0.66 C/C	
TLR22b	210 T/C	No change	50	0.60 T/T 0.40 T/C	2.0
(HQ293118)	456 C/G	Ile → Met	50	0.60 C/C 0.40 C/G	2.0
MyD88a	1358 T/G	Intron	68	0.57 T/G 0.43 T/T	1.47
(GU321986)					
TRAF6a	1347 G/A	No change	50	0.40 G/G 0.60 G/A	2.0
(GU064560)	1587 C/T	No change	49	0.41 C/T 0.59 C/C	1.65
TRAF6b	3756 T/C	Intron	46	0.39 C/T 0.61 TT	2.17
(HM535646)					
IL1β	1667 T/C	Intron	50	0.58 T/T 0.42 T/C	1.28
(AJ245635)					
IL10a	82 G/C	No change	67	0.23 G/G 0.59 G/C 0.19 C/C	0.16
(HQ323755)	320 C/G	Gln → Glu	66	0.23 C/C 0.57 C/G 0.19 G/G	0.27
	862 G/A	Intron	68	0.22 G/G 0.59 G/A 0.18 A/A	0.09
IL10b	1189 T/G	No change	46	0.50 T/T 0.50 T/G	0
(HQ323756)					

[a] SNP positions are assigned relative to the base position in the GenBank nucleotide sequence. Accession numbers are given in parentheses.

* Significant at $p < 0.05$ level.

Table 3 Results of association analysis of SNP genotypes for fourteen genes and resistance to CyHV-3 infection assessed by the Fisher exact probability test.

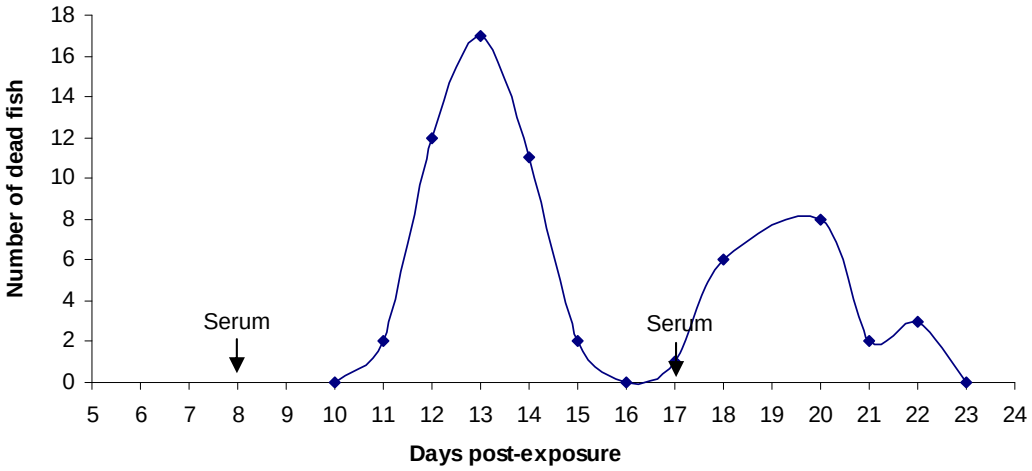
SNPs	Genotype	Susceptible N	Resistant N	*P*-value
TLR2-129A/G	AA	6	6	0.213
	AG	10	5	
TLR2-702T/C	TT	6	6	0.188
	TC	11	5	
TLR3b-669A/G	GA	9	3	0.132
	AA	8	8	
TLR4a-1781G/A	GG	7	5	0.304
	GA	9	6	
TLR4b-1577T/C	TT	9	9	0.102
	TC	8	2	
TLR7a-349G/A	GA	12	7	0.296
	GG	5	4	
TLR9-1382T/C	TT	4	4	0.253
	TC	13	7	
TLR9-1915A/G	AA	4	4	0.253
	AG	13	7	
TLR9-2692T/A	AA	13	7	0.253
	AT	4	4	
TLR21b-200G/A	GG	11	8	0.296
	GA	6	3	
TLR21b-287A/G	AA	11	8	0.296
	AG	6	3	
TLR21b-380C/T	CC	11	8	0.296
	CT	6	3	
TLR22b-210T/C	TC	8	5	0.300
	TT	9	6	
TLR22b-456C/G	CG	8	5	0.300
	CC	9	6	
MyD88a-1358T/G	TT	7	2	0.155
	TG	10	9	
TRAF6a-1347G/A	AG	10	7	0.299
	GG	7	4	
TRAF6a-1587C/T	TC	7	4	0.313
	CC	10	6	
TRAF6b-3756T/C	TC	4	4	0.288
	TT	11	7	
IL1β-1667 T/C	TC	7	4	0.299
	TT	10	7	
IL10a-82G/C	GG	5	0	0.003
	GC	11	6	
	CC	1	5	
IL10a-320C/G	CC	5	0	0.003
	CG	11	6	
	GG	1	5	
IL10a-862G/A	GG	5	0	0.003
	GA	11	6	
	AA	1	5	
IL10b-1189 T/G	TT	9	3	0.168
	TG	8	7	

Table 4 Results of association between IL10a genotype/haplotype and resistance to CyHV-3 infection assessed using the chi-square test (N = 47).

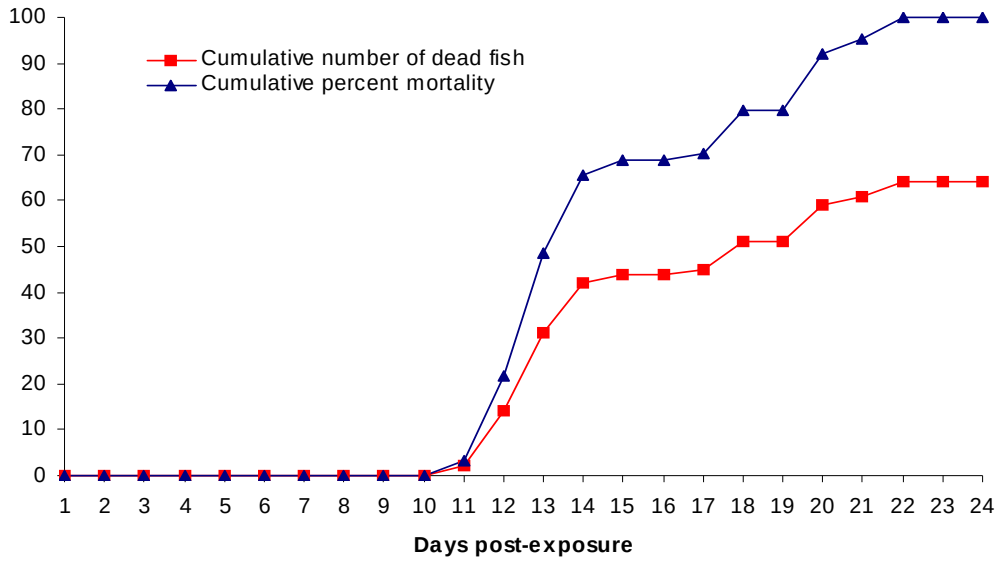
	Susceptible N (Freq)	Resistant N (Freq)	χ^2 (*P-value*)
Genotype			
GCG/GCG	11 (0.35)	1 (0.06)	9.52 (0.009)
GCG/CGA	15 (0.48)	6 (0.38)	
CGA/CGA	5 (0.16)	9 (0.56)	
CGA-dominant			
GCG/GCG	11 (0.35)	1 (0.06)	4.74 (0.029)
[GCG/CGA + CGA/CGA]	20 (0.65)	15 (0.94)	
CGA-recessive			
[GCG/GCG + GCG/CGA]	26 (0.84)	7 (0.44)	8.12 (0.004)
CGA/CGA	5 (0.16)	9 (0.56)	
Haplotype			
GCG	37 (0.60)	8 (0.25)	10.17 (0.001)
CGA	25 (0.40)	24 (0.75)	

Figure 1 Mortality curves showing actual number (a) and cumulative number and percent (b) of dead fish following CyHV-3 exposure. The mortality peak was observed on day 13. Fish dying after day 17 were possibly due to the blood sampling.

a.



b.



SUMMARY AND CONCLUSIONS

1. The common carp genome is less well studied than those of other aquaculture species such as rainbow trout, Atlantic salmon, channel catfish and tilapia. There is limited sequence information available for immune response genes in this species. However, common carp are evolutionarily closely related to a model fish species, zebrafish (*Danio rerio*), whose genome has been completely sequenced. In addition, the molecular structures of immune response genes are conserved across species, making it possible to develop degenerate primers from conserved peptide regions for amplifying common carp genes. I successfully used degenerate primers to isolate several immune response genes from common carp (*Cyprinus carpio*). These immune response gene sequences, including TLR3b, TLR4a, TLR4b, TLR7a, TLR7b, TLR9, TLR22a, TLR22b, MyD88a, MyD88b, TRAF6a and TRAF6b genes, were deposited in the GenBank database under accession numbers GU321981, GU321982, HQ229652, GU321983, GU321984, GU321985, GU459061, HQ293118, GU321986, GU321987, GU064560 and HM535646, respectively.

2. Common carp is known to be a tetrapolid species (2n = 100) as a result of a whole-genome duplication event that was estimated to have been occurred 11 – 21 MYA, and a number of duplicated genes have been identified in this species. I found that several immune response genes, including TLR3, TLR7, TLR21, TLR22, MyD88, TRAF6 and IL10, are duplicated in the common carp genome. Evidence to support the case that these genes are duplicated include polymorphisms in nucleotide and predicted amino acid sequences, phylogenetic analyses and SNP allele segregation patterns.

3. I developed gene-specific and locus-specific primers for gene sequencing and SNP genotyping of 14 common carp immune response genes. Of approximately 100 SNP markers genotyped, 65 were polymorphic and reliably scored, and were used for genotyping a mapping family. Twenty-three SNPs were informative in a mapping family. These genotype data will be used to generate a genetic linkage map for common carp. I was not successful in developing SNP markers for TNF-α and type I IFN genes. The segregation patterns of SNPs genotyped were either multiple sequence variants (MSV) or paralogous sequence variants (PSV), which likely were due to non-specific PCR amplifications of duplicate loci or different isoforms of TNF and IFN genes.

4. Among the TLR genes that have been isolated, TLR9 is known for its ability to recognize viral DNA and requires MyD88 and TRAF6 for signal transduction. Therefore, these genes may be important candidate genes for CyHV-3 resistance. I used the Rapid Amplification of cDNA Ends (RACE) technique for cDNA cloning and examined mRNA expression in a variety of tissues. I successfully isolated the full-length cDNA sequences of TLR9 and two paralogous sequences of MyD88 (MyD88a and b) and TRAF6 (TRAF6a and b). These cDNA sequences were deposited in GenBank under accession numbers GU809229, GU809230, GU809231, GU985443 and HM535645, respectively. Gene expression assays by real-time PCR showed that mRNA expression of TLR9, MyD88a, MyD88b, TRAF6a and TRAF6b varied among tissues and that differential expression was observed between transcripts of the duplicated loci. Interestingly, one form was expressed more highly in non-immune tissue (muscles). This finding suggests that both paralogs are functional, but one locus may have evolved and attained a non-immune function.

5. To determine exon-intron boundaries, I sequenced genomic DNA covering the entire coding regions of TLR9, MyD88 and TRAF6 genes. The TLR9 gene was contained in a single large exon. The coding sequences of both MyD88a and MyD88b, and TRAF6a and TRAF6b genes are

conserved, but the intron sizes differed between the two paralogs. SNPs within each form were found in both exons and introns. The results of genomic DNA cloning confirm that MyD88 and TRAF6 are duplicated in the common carp genome. This is the first report of MyD88 and TRAF6 duplication in a vertebrate.

6. Sequence analysis and characterization of protein domains in the common carp TLR9, MyD88 and TRAF6 genes revealed that the characteristics of functional protein domains are conserved and similar to those of other vertebrates, suggesting that the function of these genes in the common carp immune system may resemble those of other vertebrate immune response genes.

7. To study how polymorphisms in innate immune response genes affect the resistance to CyHV-3 infection, I analyzed the association between polymorphisms at 23 SNPs of 14 immune response genes and CyHV-3 resistance among progeny of a mapping family that were subjected to viral challenge. IL10a genes revealed significant association, and TLR4b and MyD88a showed suggestive associations. No significant or suggestive association of polymorphisms and resistance to CyHV-3 infection was identified at other genes. However, the analysis was based on a small sample size (N = 28 – 47) and experimental fish were from a single full-sib family; therefore, data from a larger-scale experiment are needed to justify the significance of gene polymorphisms and CyHV-3 resistance association.