Prediction of Toll-Like Receptor Gene Family in Asian Sea Bass

Ayşe Cebeci1,*1, Asiye Nur Saltan1, Kumar Sahayaraju2

1 Republic of Turkey Ministry of Agriculture and Forestry Central Fisheries Research Institute Kasustu, Yomra, 61250, TRABZON.
2 Kadalamma-Centre for Ocean Sustainability Research, Education and Training, Thumba, Thiruvananthapuram, Kerala, India-695022.

Article History
Received 16 October 2019
Accepted 25 December 2019
First Online 26 December 2019

Corresponding Author
Tel.: +905318410150
E-mail: ayse.cebeci@tarimorman.gov.tr

Keywords
Genome mining
PRRs
Lates calcarifer
TLR
Innate immunity

Abstract
Toll-like receptors (TLRs) are one of the important components of innate immune system involved in pathogen recognition. Aquaculture species encounter potential diseases caused by pathogens resulting in great economic loss for producers. Asian seabass is a significant aquaculture species, but there is not enough information about their TLR repertoire to comprehend the underlying molecular regulation of pathogen recognition for effective measurements. In this study, members of TLR family in Asian seabass were identified by genome mining to provide basic information for future studies on pathogen recognition receptor-driven signaling in Asian seabass. Then phylogenetic and conserved domain analysis were constructed for predicted TLRs. Analysis resulted with identification of 17 TLR gene that includes mammals and fish specific TLRs and four of the TLR genes (TLR2, TLR5, TLR22, TLR23) had more than one copy. Phylogenetic analysis showed that Asian seabass predicted TLRs made group with their counterparts from other teleost and mammals and also, predicted TLRs exhibit an expected conserved structure. These results mean that TLR family were conserved and even expanded because of selective pressure caused by pathogen in distinct habitat.

Introduction
The role of innate immune system is to provide first line defense against microbial pathogens in invertebrate and vertebrate. In the innate immune system, pathogen recognition receptors (PRRs) recognizes pathogen-associated molecular patterns (PAMPs) which are molecules related to the group of a pathogen such as lipopolysaccharide (LPS), bacterial DNA, viral double strand RNA (dsRNA) or single strand RNA (ssRNA) and then activate the innate immune response. There are 3 major classes of PRRs: Toll-like receptors (TLRs), retinoic acid-inducible gene (RIG)-I-like receptors (RLRs), and nucleotide-oligomerization domain (NOD)-like receptors (NLRs) (Aoki, Hikima, Hwang & Jung, 2013). Unlike cytoplasmic RLRs and NLRs, members of TLR family are type I transmembrane protein localized at the cell surface or in endosomal compartments. TLRs consist of 3 main domains: an intracellular Toll/interleukin1 receptor (TIR) domain that is responsible for signal transduction and most conserved part of TLR, a transmembrane region (TM) and, an extracellular N-terminus with leucine-rich repeat region (LRR) that recognizes PAMPs (Aoki et al., 2013).

TLRs are very important since they recognize various PAMPs of bacteria, viruses, fungi, and protozoa and then activate signaling for stimulation of innate immunity to protect the host against diseases. Upon
recognition of PAMPs by LRR domain of TLRs, TIR domain of that recruits TIR-domain containing adaptor molecules such as Myeloid differentiation primary-response protein 8 (MyD88) and trigger MyD88-dependent pathway or MyD88-independent pathway resulting in production of type I interferons (IFNs), inflammatory cytokines and chemokines (Rauta, Samanta, Dash, Nayak & Das, 2014). Apart from TLR3, all TLRs function through MyD88-dependent pathway (Rauta et al., 2014). TLRs activates directly innate immunity in that way and also they develop adaptive immunity indirectly (Tanekhy, 2014).

TLRs were identified in both vertebrates and invertebrates. In vertebrates, TLRs are categorized into 6 major families: TLR1, TLR3, TLR4, TLR5, TLR7, and TLR11 (Roach et al., 2005). Although TLR gene family are evolutionarily conserved in fish, mammals, and birds, the number of TLR genes changes among various organisms.

Since TLR was first discovered in Drosophila melanogaster, TLR1-10 has been reported in mammals whilst TLR11,12,13 have been only identified in Muridae species (Takano et al., 2011). Following the first identification of fish TLR gene in goldfish (Carassius auratus) (Stafford, Ellestad, Magor, Belosevic & Magor, 2003), the draft whole genome sequence of Fugu (Takifugu rubripes) and the genome database of zebrafish (Danio rerio) was surveyed for TLR genes (Jault, Pichon & Chluba, 2004; Oshiumi et al., 2003). A comparison of the results of these two surveys showed a set of orthologous genes but also non-mammalian and even teleost unique TLRs. Though both species showed non-mammalian TLRs such as TLR14, 21; TLR4, 19, 20 were only identified in zebrafish while TLR5S and 23 were only found in Fugu. Phylogenetic analysis showed 11 distinct TLR types were identified in Fugu and 14 distinct TLR types were detected in zebrafish, of which three are duplicated (TLR4, 5, and 8). TLR genes has been identified in other fish species, involving Japanese flounder (Paralichthys olivaceus), rainbow trout (Oncorhynchus mykiss), yellow croaker (Pseudosciaena crocea), channel catfish (Ictalurus punctatus), Atlantic salmon (Salmo salar), grass carp (Ctenopharyngodon idella), common carp (Cyprinus carpio), rare minnow (Gobiocypris rarus), orange-spotted grouper (Epinephelus coioides) and gilthead seabream (Sparus aurata) (Aoki et al., 2013). On account of various evolutionary mechanisms such as gene duplication, retrotranscription, and high gene expansion rate and of their distinct habitat, TLR genes display different features with large diversity in aquatic animals (Palti, 2011). Thus far, at least 20 different TLR types (TLR1, 2, 3, 4, 5M, 5S, 7, 8, 9, 14, 18, 19, 20, 21, 22, 23, 24, 25, 26, and 28) have been identified in teleost. Although TLR1, 2, 3, 5, 7, 8, 9 genes were conserved in both mammals and fish, TLR1 subfamily consists of TLR1, 2, 14, 18, 25 with the missed TLR6 and 10 in teleost (Rauta et al., 2014). Because a gene showed equally high identity to mammal TLR1,6,10; this gene is called TLR1 in Fugu genome (Oshiumi et al., 2003). Novel fish specific TLRs, TLR25 and TLR26 were only found in channel catfish lately (Zhang et al. 2013). TLR4 only characterized in the order of Cypriniformes, which includes zebrafish in teleost (Jault, Pichon & Chluba, 2004). Nonetheless, according to Sullivan et al. (2009), zebrafish TLR4s are paralogs rather than orthologs of mammal TLR4 due to tandem rearrangement and differences in LPS signaling. By contrast in mammals, there are two types of TLR5 in teleost: TLR5 membrane form (TLR5M) that is orthologous to mammalian TLR5 and, TLR5 soluble form (TLR5S) that does not contain a TM and TIR domain. TLR5S were found in fugu, Japanese flounder, rainbow trout, and channel catfish genome suggesting TLR5 arose by way of duplication of the LRR domain of TLR5M (Hwang, Asahi, Kondo, Hirono & Aoki, 2010; Oshiumi et al., 2003; Tsujita et al., 2004; Zhang et al. 2013). Similar to the genomic organization of mammalian TLR7 and 8 genes on X chromosome, TLR7 and 8 genes are tandemly located on the same chromosome in zebrafish, fugu, rainbow trout, and Japanese flounder suggesting that they arose through tandem gene duplication in advance of the divergence of teleost and tetrapods (Aoki et al., 2013). TLR11 subfamily consists of fish specific TLRs and murine TLRs (TLR11, 12, 13, 19, 20, 21, 22, 23). It is presumed that TLR11, 12, 13 were lost in teleost whilst fish specific TLRs were lost in mammals following the divergence of mammals from teleost lineage. As is seen from catfish genome with two copy of TLR5M (Zhang et al. 2013), gene duplication has an important role in diversification of TLRs. The known mammalian TLRs result in ancient gene duplication while TLRs could arise from recent gene duplications in teleost. In rainbow trout genome, the presence of two TLR22 genes were shown as a result of a recent gene duplication event (Rebl, Siegl, Köllner, Fischer, & Seyfert, 2007).

Different TLRs would be activated by various ligands and then trigger MyD88-dependent pathway or MyD88-independent pathway. Compared to ligand specificity of TLRs in fish, ligand specificity of TLRs is well-studied in mammals. Dimeric combinations of TLR2-1 and TLR2-6 recognize various lipoproteins and bacterial peptidoglycan in mammals (Palti, 2011). TLR5 detects bacterial flagellin while TLR4 recognizes LPS (Palti, 2011). In mammals, monomeric LPS is transferred to TLR4 via a complex of LPS-binding protein (LBP) and the cluster of differentiation 14 (CD14). Then TLR4 and myeloid differentiation protein 2 (MD2) make up a receptor complex together and they perform as the principal cell-surface LPS-binding component (Kawai & Akira, 2009). Yet, LPS recognition and sensitivity in fish is essentially different from mammals (Iliev et al. 2005) and studies on zebrafish showed that LPS response is not activated by TLR4 (Sepulcre et al., 2009; Sullivan et al., 2009). TLR3 is involved in recognition of dsRNA while TLR7 and 8 is responsible for recognition of ssRNA and TLR9 recognizes viral or bacterial DNA including unmethylated CpG motifs (Palti, 2011). From that
information, TLR1, 2, 4, 5 are assumed as a sensor of bacterial ligands and TLR3, 7, 8, 9 are regarded as a sensor of viral ligands in teleost (Rauta et al., 2014). Nevertheless, evidence of ligand specificity is available for TLR2, 3, 5M, 5S, 22 among 20 TLR types in fish to date (Palti, 2011).

Asian seabass is a catadromous fish species distributing in the Indo-West Pacific region from South Asia to Papua New Guinea and Northern Australia (Frumkin, 2003). It is also significant aquaculture species in southern Asia and Australia due to its high economic value. Various diseases result from different pathogens such as *Streptococcus iniae*, nodavirus, marine vibrios and *C. irritans* (Creager & Buller, 2006; Azad et al., 2005; Krupesha Sharma, Rathore, Verma, Sadhu & Philipose, 2011; Bryant, Lee, Lester & Whittington, 1999) causes significant loss in Asian seabass aquaculture. In spite of great economic losses owing to diseases in Asian seabass, little attention has been paid to understand the PRR-driven signaling mechanism. Until now, nod-like receptor C3 (NLRC3) from NLR family, laboratory of genetics and physiology 2 gene (LGP2) and melanoma differentiation-associated factor 5 (MDA5) from RLR family and only TLR22 from TLR family were characterized in Asian seabass (Paria et al., 2016; Paria et al., 2017; Paria, Makkes, Chaudhari, Purushothaman & Rajendran, 2018a; Paria, Makkes, Chaudhari, Purushothaman & Rajendran, 2018b). Presence and role of other members of TLR family in Asian seabass one of the most important food is unknown.

The aim of this study is to reveal TLR gene family in Asian seabass to provide initial insights into PRR-driven signaling mechanism. In this context, TLR repertoire in Asian seabass were identified, and their primary structures and evolutionary relationship were analyzed. TLRs bind their ligands resulting in improved innate immunity as it was shown in Rainbow trout treated with the ligand of TLR3, Poly I:C (Tafalla, Bøgwald & Dalmo, 2013). Also studies showed that polymorphism on TLR gene region can be responsible for disease resistance (Heng, Su, Huang, Dong & Chen, 2011). Therefore, this study would provide basic data for the development of vaccine adjuvants, the selection of disease resistant breeds and comparative studies.

Materials & Method

Database Mining for TLR Genes and Conserved Domain Analyses

NCBI gene and nucleotide databases were searched for reference sequences to use in blast search. Then reference sequences were retrieved from NCBI and then BLAST search (https://blast.ncbi.nlm.nih.gov/Blast.cgi) (Altschul, Gish, Miller, Myers & Lipman, 1990) was done on Asian seabass assembly (ASM164080v1).

BLAST search was done for Transcriptome of Asian seabass in order to identify TLR genes. Reference RNA sequence database of Asian seabass on NCBI was searched by blastn using different TLRs (coding region sequences(CDS)) of seven organisms as query. Somewhat similar sequence algorithm was chosen and hits were taken with a cut off e-value< 1.10^{-11}. Then genome of Asian seabass was searched to find not annotated genes, so Refseq genome of Asian seabass was searched by tblastn with same query set. Somewhat similar sequence algorithm was chosen and hits were taken with a cut off e-value<1.10^{-23}.

Finally, conserved domains were predicted by SMART in normal mode and PFAM domain option were chosen (http://smart.embl-heidelberg.de/) (Letunic & Bork, 2018). TMHMM Server, v. 2.0 (http://www.cbs.dtu.dk/services/TMHMM/) (Krogh, Larsson, von Heijne & Sonnhammer, 2001) were additionally used to predict TM region because SMART could not predict all TMs or sometimes show more than one TM region (Jault et al., 2004). When TMHMM Server, v. 2.0 program interpreted that one of the TM regions can be signal peptide, we checked it by SignalP 4.0 Server (http://www.cbs.dtu.dk/services/SignalP-4.0/) (Petersen, Brunak, von Heijne & Nielsen, 2011) with “input sequence do not include TM regions” method using first 350 amino acid sequence for confirmation.

Phylogenetic Analysis

A phylogenetic tree was constructed for predicted TLR genes of Asian seabass to analyze the evolutionary relationships of them with mammalian and other teleost TLRs by phylogeny.fr online version, a la carte mode (http://www.phylogeny.fr/alacarte.cgi) (Dereeper et al., 2008). Multiple protein sequence alignment was conducted by ClustalW (Thompson, Higgins and Gibson, 1994) using protein sequences of predicted Asian seabass TLRs, Fugu TLRs, Mouse TLRs and Brown rat TLR10 and then alignment curation was done with Gblocks for a less stringent selection. Subsequently, a phylogenetic tree was constructed using maximum likelihood method with WAG substitution model and bootstrap sampling was replicated 100 times.

Results

Reference Sequence

Reference sequences of different TLRs from seven different organisms (Zebrafish, Fugu, Channel catfish, Rainbow trout, Chicken, Mouse, Brown rat) were found on the NCBI gene or nucleotide databases (Table 1). The reason is that some genes are specific to fish, mammals, or birds while some genes are present in all vertebrates. All TLR sequence of Fugu, zebra fish, and Rainbow trout were used for BLAST search. Also, all Mouse TLRs sequence and Brown rat TLR 10 sequence were used to
analyze our results in vertebrate context. Furthermore, channel catfish TLR25 and 26 sequences and chicken TLR15 and 16 sequences were used for BLAST search.

**Identification of TLR genes in Asian seabass**

Different TLRs of eight organisms were used as query to search Asian seabass transcriptome and genome using blastn (Table 2) and tblastn (Table 3) to revile TLR gene families in Asian seabass. In this study, 17 TLR genes were identified in Asian seabass genome: TLR1, TLR2-1, TLR2-2, TLR14, TLR3, TLR5M-1, TLR5M-2, TLR5S, TLR7, TLR8, TLR9, TLR21, TLR22-1, TLR22-2, TLR22-3, TLR23-1, TLR23-2. One or more counterpart of mammalian TLRs were found in Asian seabass genome, but it does not include any murine TLRs. In addition, Asian seabass genome has some fish specific TLRs (Figure 1).

**TLR1 Family**

When Asian seabass transcriptome and genomes were searched by using fish TLR1 reference sequence, programs showed a gene (gene ID: 108899238) with low e value (Table 2 and 3). Asian seabass genome and transcriptome were searched using mouse TLR1, 6 and brown rat TLR10 in order to release relationships of this gene with mammalian TLR1, 6 and 10. Results showed the same gene with higher e-value, but there were no significant differences between values for mouse TLR1, 6 and brown rat TLR 10 (Table 2 and 3). Thus, Asian seabass have a TLR gene that includes similar sequence to mammalian TLR 1, 6 and 10.

Blasnt and tblastn search for TLR2 showed 3 possible genes (108901668, 108901669, 108901670) with 5.10^-15 e-values. However, length of CDS of 108901668 is shorter than expected length of CDS based on model organisms. Also, the product of this gene is called as Low-quality protein on NCBI. Therefore, we eliminated 108901668 genes, and 108901669 and 108901670 are present on Asian seabass genome as TLR2 gene.

Asian seabass genome and transcriptome searched for TLR 14 and TLR 18 revealed a gene (108899238) with high e-value (0.058-1.10^-142) (Table 2 and 3). This gene had lower e-value (0-8.10^-19)
when BLAST search was done for TLR 1 sequence. Program showed this gene for the search of TLR 15 and 16 because these genes are in the same subfamily. Consequently, TLR15 and 16 was not found in Asian seabass. Similarly, BLAST search for TLR 25 showed a gene (108901668), which was identified as possible TLR 2 gene and then eliminated above, with high e-value (Table 2 and 3). The result suggests that Asian seabass does not have TLR 25 gene.

**TLR4 Family**

When blast search was done on Asian seabass transcriptome and genome, program showed other TLR genes and unrelated genes with 0,67 - 3.10^{-5} e-value. Blast results showed other TLR genes only due to sequence similarity of a gene family yet TLR4 is not available in Asian seabass genome. In addition, accessory and adapter molecules that are mandatory for TLR4-mediated LPS response in mammals were searched in NCBI gene and nucleotide databases. CD14, MP2, TICAM2 gene and nucleotide sequences are not available for teleost fish and Asian seabass. Also, LBP gene or nucleotide sequence was not found for Asian seabass whilst there was gene sequence of LBP based on predicted sequence for some teleost species.

**TLR5 Family**

BLAST search for TLR5 showed 3 genes (108895805, 108895804 and 108888912) with low e-value (0-4.10^{-10}) and length of their CDS sequences are 2672, 2672 and 1954 respectively. The CDS sequence length of the third gene is similar to the CDS sequence length of the soluble form of TLR5 in Rainbow trout. SMART program was used to predict conserved domains of these genes and results showed that the third gene possess only LRR region whilst first two genes contain LRR region, a transmembrane region, and a TIR domain. Therefore, Asian seabass genome includes a TLR5 soluble form and two TLR5 membrane form.

**TLR3 and 7 Families**

BLAST search showed one gene (108901661) for TLR3 mostly with 0 e-value, so this gene is present in Asian seabass genome (Table 2 and 3). When we have done BLAST search for TLR7, program showed a gene (108901560) with 0-2.10^{-15} e-value and we identified this gene as TLR7. Similarly, BLAST search for TLR8 and 9 showed 108890949 and 108877869 with e-values mostly 0 for TLR8 and 9 respectively (Table 2 and 3). However, TLR9 in mouse showed high e-value with 0, 24

<table>
<thead>
<tr>
<th>TLR SUBFAMILY</th>
<th>TYPE OF TLR</th>
<th>Zebrasish</th>
<th>Fugu</th>
<th>Rainbow trout</th>
<th>Channel catfish</th>
<th>Chicken</th>
<th>Mouse</th>
<th>Brown rat</th>
</tr>
</thead>
<tbody>
<tr>
<td>TLR1</td>
<td>TLR1</td>
<td>5.10^{53}</td>
<td>0</td>
<td>2.10^{160}</td>
<td>8.10^{39}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLR1</td>
<td>TLR2</td>
<td>2.10^{48}</td>
<td>0</td>
<td>5.10^{129}</td>
<td>2.10^{64}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLR1</td>
<td>TLR6</td>
<td></td>
<td></td>
<td></td>
<td>2.10^{39}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLR1</td>
<td>TLR10</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLR1</td>
<td>TLR14</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLR1</td>
<td>TLR15</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLR1</td>
<td>TLR16</td>
<td></td>
<td></td>
<td></td>
<td>3.10^{17}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLR1</td>
<td>TLR18</td>
<td></td>
<td></td>
<td></td>
<td>6.10^{154}</td>
<td>1.10^{15}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLR1</td>
<td>TLR25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLR3</td>
<td>TLR3</td>
<td>3.10^{-25}</td>
<td>0</td>
<td>0</td>
<td>1.10^{15}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLR4</td>
<td>TLR4</td>
<td>a</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLR5</td>
<td>TLR5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLR5</td>
<td>TLR5M</td>
<td>a</td>
<td>6.10^{34}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLR7</td>
<td>TLR7</td>
<td></td>
<td></td>
<td></td>
<td>2.10^{21}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLR8</td>
<td>TLR8</td>
<td>a</td>
<td>1.10^{48}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLR8</td>
<td>TLR8</td>
<td>b</td>
<td>2.10^{13}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLR7</td>
<td>TLR9</td>
<td>5.10^{61}</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLR11</td>
<td>TLR11</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLR11</td>
<td>TLR12</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLR11</td>
<td>TLR13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLR11</td>
<td>TLR19</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLR11</td>
<td>TLR20</td>
<td>a</td>
<td></td>
<td></td>
<td>5.10^{-22}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLR11</td>
<td>TLR20</td>
<td>b</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLR11</td>
<td>TLR21</td>
<td>2.10^{129}</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLR11</td>
<td>TLR22</td>
<td></td>
<td></td>
<td></td>
<td>4.10^{156}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLR11</td>
<td>TLR23</td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLR11</td>
<td>TLR26</td>
<td></td>
<td></td>
<td></td>
<td>8.10^{14}</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* the ones e-value higher than 1.10^{-10} was removed.
and this is not unusual as these organisms are evolutionarily far from each other.

**TLR11 Family**

When murine TLRs were searched by blastn, it showed unrelated genes and TLR13-like-gene (108880675) in Asian sea bass that is so called in NCBI with high e-value (Table 2). Also, tblastn results showed Asian sea bass TLR13-like-gene with high e-value (Table 3), but when we searched TLR21 gene by blastn and tblastn, they showed this gene with 0 e-value. Program showed these results for murine TLRs due to sequence similarity of the same subfamily. Therefore, Asian sea bass has a TLR21 gene while it does not include any murine TLRs. Whereas TLR19 and 20 is not present in Asian sea bass genome BLAST search showed 3 orthologs of TLR22 and 2 orthologs of TLR23 with 0 e-value.

**Conserved Domain Analysis**

SMART was used to predict conserved domains of TLRs in Asian seabass genome. TLR2-1, TLR2-2, TLR14, TLR3, TLR5M-1, TLR5M-2, TLR7, TLR8, TLR9, TLR21, TLR22-3 showed LRR region, a TM region, and a TIR domain as expected (Figure 1). Nonetheless, TLR5 soluble form possesses only a LRR region. SMART program showed TLR1, TLR22-1 and TLR22-2 have LRR region, 2 TM region (one of them are N side of the protein), and TIR domain. In addition, protein sequence of TLR1, TLR22-1 and TLR22-2 were analyzed by TMHMM version 2.0 and it indicated the possibility of being signal peptide. Then first 350 amino acid sequence of the proteins were checked by SignalP 4.0 Server and it showed the signal peptide. That means one of TM regions is signal peptide and TLR1, TLR22-1 and TLR22-2 have LRR region, a TM region, and TIR domain as expected. SMART showed TLR23-1 and TLR23-2 possess LRR region and TIR domain without transmembrane region. Protein sequence of them were check by TMHMM version 2.0 and it did not predict any TM region. As a result, all the predicted TLRs displayed an expected TLR structure, aside from TLR23-1, TLR23-2, and TLR5S.

**Phylogenetic Analysis**

A phylogenetic tree was conducted to confirm identity of predicted Asian seabass TLRs and analyze their evolutionary relationship in the context of fish and mammalian TLRs. One or more counterparts of mammalian TLRs are present in Asian seabass genome and each of these TLRs made a subgroup with its
Figure 1. Schematic representation of the conserved domain structure of Asian seabass TLRs and their GeneID on NCBI. Dark blue rectangles represent trans-membrane regions while green bands indicate LRRs involved in PAMP recognition (LRR-TYP represents typical LRRs, whereas LRR represents LRR outliers). Light blue rectangles and circles represent L terminal and C terminal LRRs respectively, and thin pink bars segments of low compositional complexity. Motifs with the prefix PFAM represent those that are recognized by the PFAM database.
Figure 2. Phylogenetic tree demonstrating the evolutionary relationship of Asian seabass (As), Takifugu (Ta), and Mouse (Mo) TLRs. The Phylogenetic tree was constructed using the maximum-likelihood method based on the alignment of the amino acid sequences with 100 bootstrap by phylogeny.fr. The numbers on branches shows percentage of bootstrap values of the 100 bootstrap replicates.
Discussion

In the innate immune system, a variety of PAMPs are recognized by PRRs that contain TLRs, RLRs, NLRs. Although the members of TLRs were identified in many fish species, there is no information about any members of TLR gene family in Asian seabass, except TLR22 (Paria et al., 2018b). In this study, we identified 17 TLR genes in Asian seabass genome by bioinformatics approach. Their homology was confirmed by phylogenetic tree under 5 subfamilies (TLR1, TLR3, TLR5, TLR7, TLR11) and also most of the predicted TLRs share similar structure with other teleost and mammalian TLRs.

Asian seabass genome contains orthologs of mammalian TLR1, 2, 3, 5, 7, 8, 9 as model organism Fugu (Oshiumi et al., 2003). Phylogenetic tree that showed these genes have a common ancestor suggesting they segregated prior to mammalian divergent from teleost. These genes must have been conserved between mammal and teleost since pathogens cannot change the main structure of their PAMPs such as PGN (Takano et al., 2011).

Asian seabass genome contains TLR1 gene (108899238) whereas there is not TLR6 and 10. Blast search with fish TLR1 showed one possible gene. Also, blast search was done using mouse TLR1, 6 and brown rat TLR10 to understand relationship of them with possible TLR1. Results showed the same gene with higher e-value (Table 2 and 3). That means Asian seabass TLR1 gene share sequence similarity with mouse TLR1, 6 and brown rat TLR10 as Fugu TLR1 showed (Oshiumi et al., 2003). Phylogenetic tree showed mouse TLR1, 6 and brown rat TLR10 are closer to each other than fish TLR1 though mouse TLR1, 6, brown rat TLR10 and fish TLR1 are in the same group. To sum up, fish TLR1 and mammalian TLR1 had a common ancestor and then ancestral TLR1 diverged into TLR1, 6, 10 in mammal lineage.

TLR1 family contains TLR1, 2, 6, and 10 and homo or hetero dimers of them recognize a variety of ligands from bacteria, viruses, and fungi in mammals (Palti, 2011). TLR1 family expanded with other genes in teleost and this study showed TLR1 family expanded with TLR2 duplicate and TLR14 in Asian seabass. TLR2 duplicate and TLR14 might be functional substitution of mammalian TLR6 and TLR10 to make an efficient immune response against different pathogens in water. Phylogenetic tree showed Asian seabass TLR2s and TLR14 are closely related to each other under TLR1 family. This supported the idea that TLR14 diverged from an ancestor of TLR2 before the divergence of agnathans and teleost (Hwang, Kondo, Hirono & Aoki, 2011).

Orthologous of mammalian TLR4 were not found in Asian seabass genome like fugu genome (Oshiumi et al., 2003). This is an expected result since TLR4 were only reported in the order cypriniformes, which includes zebrafish. Yet, studies showed Zebrafish TLR4s are paralog and do not have the same function with mammalian TLR4 (Sullivan et al., 2009; Sepulcre et al., 2009). In addition, accessory and adaptor molecules were not identified in Asian seabass and even in teleost to date. The results suggest that in mammals accessory and adaptor molecules of TLR4 might have evolved for LPS recognition whereas TLR4 might have been lost in teleost lineage. In teleost, alternative receptors or mechanisms could have evolved for LPS recognition.

In this study, we found two TLR5M and one TLR5S based on length of CDS and conserved domain analysis and also catfish genome includes two TLR5M and one TLR5S (Zhang et al., 2013). After identification of TLR5S in many fish species to date, it was suggested that TLR5S arose through duplication of the LRR domain of TLR5M (Aoki et al., 2013). Phylogenetic tree supported this idea for Asian seabass TLR5S because Asian seabass TLR5S made a subgroup together and then TLR5S separated from TLR5M in Asian seabass lineage. TLR 7, 8 and 9 had been identified for Asian seabass but phylogenetic tree showed an unexpected result for TLR7 subfamily. The predicted Asian seabass TLR7 and TLR9 are under the same subclade with their counterparts from Fugu and Mouse. However, Mouse TLR8 made a subgroup with TLR7s with 48% bootstrap value while the predicted Asian seabass TLR8 and Fugu TLR8 made another subgroup. This could be result of close relationship of TLR8 with TLR7 due to TLR8 arose via gene duplication of TLR7 and mouse are evolutionarily far from teleost (Du, Poltorak, Wei & Beutler 2000). On the other hand, low bootstrap value might indicate there could be a computational error.

Evolutionary mechanisms such as gene duplication causes expansion of TLRs and fish specific TLRs in teleost. Asian seabass genome has more than one copy of 4 genes: TLR2, TLR5M, TLR22, TLR23. Even if most duplicates convert into pseudogenes, the fate of duplicates can be subfunctionalization or neofunctionalization (Mazet & M. Shimeld, 2002). Twelve TLR22 and two TLR23 arose via lineage specific duplication in Atlantic cod and underwent neofunctionalization (Sundaram, Kiron, Dopazo & Fernandes, 2012). As a result of subfunctionalization or neofunctionalization, duplicate genes might recognize different PAMPs or their expression pattern could be different in Asian seabass as duplicate genes of Atlantic cod and this idea can be proven by functional studies.

In Asian seabass genome, there are 3 copies of TLR22 and blast search showed one of them
**Figure 3.** Asian seabass TLR22-1, TLR22-2, TLR22-3 DNA sequences were aligned by Clustal Omega. The identical DNA sequences are indicated in (*).
(108885101) were equal to Asian seabass TLR22 characterized before. Rebêl et al. (2007) showed that recent gene duplication result in twin TLR22 in rainbow trout and this event might occur in Asian seabass genome. Alignment of Asian seabass TLR22s by Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) (Sievers et al., 2011) showed a deletion on TLR22-3 (Figure 3). Although TLR22s are under the same clade on phylogenetic tree, TLR22-1 and TLR22-2 made a subgroup together (Figure 1). TLR22-3 may result from a recent gene duplication of TLR22-1 and TLR22-2 via unequal crossing over. Conserved domains of TLR22-1, TLR22-2, TLR22-3 was predicted by SMART and it showed that TLR22-1 and TLR22-2 has 17 LRR while TLR22-3 has 14 LRR. Variations in repeat numbers, sequence, and length of LRRs are considered to be associated with recognition of specific PAMPs by individual TLRs (Aoki et al., 2013). Therefore, TLR22-3 might have recognized different PAMPs than TLR22-1 and TLR22-2. Further studies are needed to prove this suggestion.

Murine specific TLRs were not found in Asian seabass genome like other fish. Phylogenetic tree showed murine specific TLRs made a group with fish specific TLRs (Figure 1). These consequences support the idea that mammals lost fish specific TLRs whilst teleost lost murine specific TLRs after divergence (Palti, 2011).

In conclusion, identification of all members of the TLR family has been done for the first time through genome mining in Asian seabass. In this study, we identified 17 TLR genes that contain mammalian and fish specific TLRs and come across that TLR repertoire is widened as consequence of gene duplication. Phylogenetic analyses displayed that members of TLRs are under 5 major subfamilies, except TLR4. All TLR members in Asian seabass presented LRR region, a TM region, and a TIR domain apart from TLR23-1, TLR23-2, and TLR5S. In this study, validation analyses have not been done but, in future we are planning to make further analyses on this topic broadly. This study provided first insights into TLR driving signaling mechanism in Asian seabass and next step will be leading us to validation and ligand studies of these TLRs. Also, challenge studies with different microbes can be done to compare expression pattern. The basic information obtained from this study can be used in future studies and could support development of more effective vaccine, therapeutics, and the selection of disease-resistant breed stocks.

References


