Mol Biol Rep (2012) 39:8965–8975
DOI 10.1007/s11033-012-1765-y

Positive selection pressure within teleost toll-like receptors *tlr21* and *tlr22* subfamilies and their response to temperature stress and microbial components in zebrafish

Arvind Y. M. Sundaram · Sonia Consuegra · Viswanath Kiron · Jorge M. O. Fernandes

Received: 16 April 2012 / Accepted: 7 June 2012 / Published online: 24 June 2012

Abstract Toll-like receptors (TLRs) play a crucial role in host defence, since they trigger immune response following recognition of pathogen-associated molecular patterns (PAMPs) in potential infectious agents. TLRs have been found in numerous organisms, including mammals, birds and teleosts. Some TLR members are commonly retained across all species, whilst others were lost, gained or diverged independently during evolution. Our knowledge about the evolution and specific functions of *tlr21*, *tlr22* and *tlr23* in teleosts are still scarce. Phylogenetic analysis of 18 *tlr13*, *tlr21*, *tlr22* and *tlr23* genes from 9 different fish species divided them in two groups. All *tlr21* genes were under the first clade, while the second comprised *tlr22*, *tlr23* and *tlr13* from Atlantic salmon. Evidence of positive selection was detected at three sites within the leucine-rich repeat regions of Tlr22, which may influence PAMP recognition. Immunostimulation experiments revealed that expression of zebrafish *tlr22* is modulated by several unrelated PAMPs. Up to a 3-fold increase in *tlr21* and *tlr22* expression was detected in larvae exposed to immunostimulants such as lipopolysaccharide, peptidoglycan or poly I:C. We found that zebrafish *tlr*s are expressed mainly in immune-related organs, such as spleen and kidney as well as in testis and temperature stress did not have an effect on the expression of *tlr21* and *tlr22* in the early stages of development in zebrafish larvae. Our data indicates that these teleost *tlr*s may play a role in innate host defence. In particular, *tlr22* is evolving under positive selection, which indicates functional diversification and adaptation of the response to different PAMPs.

Keywords Toll-like receptor · *tlr22* · Innate immunity · Positive selection · Teleost fish

Abbreviations

Tlr	Toll-like receptor
PAMPs	Pathogen-associated molecular patterns
LRR	Leucine-rich repeat
LPS	Lipopolysaccharide from *Salmonella enterica typhimurium*
IC	Polyinosinic-polycytidylic acid potassium salt
PG	Peptidoglycan from *Staphylococcus aureus*

Introduction

The innate immune system is the first line of host defence for all vertebrates, particularly during ontogeny when the adaptive immune system is not fully functional. Various factors are involved in innate immune response including antimicrobial peptides, lysosomes and complement system [1–3]. Pattern-recognition receptor (PRR) expressing cells of the innate immune system recognise small molecular motifs conserved within groups of pathogens (pathogen-associated molecular patterns or PAMPs) and activate specific immune pathways. PAMPs are essential for microbial survival and thus form an excellent target for the PRRs [4]. Transmembrane and cytoplasmic signalling

Electronic supplementary material The online version of this article (doi:10.1007/s11033-012-1765-y) contains supplementary material, which is available to authorized users.

A. Y. M. Sundaram · V. Kiron · J. M. O. Fernandes (✉)
Faculty of Biosciences and Aquaculture, University of Nordland, 8049 Bodø, Norway
e-mail: Jorge.fernandes@uin.no

S. Consuegra
Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Ceredigion SY23 3DA, UK

PRRs recognise various PAMPs and trigger the release of inflammatory cytokines and type-I interferons for host defence [2, 5]. Among many known PRRs, Toll-like receptors (TLRs) are recognised as key recognition molecules of the innate immune system [6].

Toll-like receptors are type-I transmembrane proteins with extracellular leucine-rich repeat (LRR) motifs and an intracellular Toll/interleukin-1 receptor (TIR) domain. While the TIR domain initiates intracellular signalling, the LRR specialises in recognition of PAMPs. The extracellular domain consists of 16–28 LRRs with each LRR containing 10–20 amino acids with the conserved motif "LxxLxLxxNxL". TLRs form either homo- or heterodimers with a horseshoe-like structure and this conformation is essential for ligand binding and initiation of downstream signalling pathways [7]. TLRs are expressed in distinct cellular compartments. Many of them are expressed on the cell surface whereas others are found in intracellular vesicles, such as the endosome and endoplasmic reticulum. The intracellular TLRs are transported to the vesicles via a transmembrane protein, which are localised in the endoplasmic reticulum [4]. TLRs expressed on the cell surface recognise mainly microbial membrane components such as lipids, lipoproteins and proteins, while the TLRs expressed in intracellular vesicles specialise in recognising microbial nucleic acids [8].

Toll, the first TLR to be identified, was found to play an important role during embryogenesis in establishing dorsal–ventral axis in *Drosophila melanogaster* [9]. Later it was confirmed that it also provides protection against fungal infections in adult fruit flies [10]. The engagement of TLRs by microbial components triggers a signalling cascade, leading to the induction of genes involved in antimicrobial host defence. Twelve functional TLRs have been identified in mouse, termed TLR1–9 and TLR11–13. TLR1–TLR9 are conserved between mouse and humans, but TLR11–TLR13 are absent in human genome. Also, the TLR10 present in human is not functional in mouse [11]. Additional TLRs absent in mammals are encoded by avian and teleost genomes [12, 13].

Genes from the vertebrate immune system have traditionally been seen as evolving under purifying selection, as TLRs recognise highly conserved pathogens and they should be under strong selection against functional change [14]. However, increasing evidence suggests that positive selection is also involved in the evolution of some innate immune genes, including TLRs. Mammalian TLRs from up to 23 different species, have shown that signatures of positive selection exist in all 10 primate TLRs [14]. In the avian genome, positively selected sites were observed in TLR2 and TLR7 genes [15]. In fish, only TLR9 has been reported to be under positive selection to date [16], which may be related to functional adaptations during teleost evolution. In spite of a large degree of conservation between teleost TLRs and their mammalian orthologues, there are some differences in PAMP recognition and even signalling. For example, immunostimulation experiments have revealed that TLR4 response can be similar to its mammalian counterpart in some fish species but rather different in others [13].

Remarkably distinct features have been discovered in fish Tlr cascades [13], even if most research on TLRs has focused on mammalian systems. *Tlr21*, *tlr22* and *tlr23* are teleost genes that are poorly characterized to date. It is plausible that, similarly to their mammalian counterparts, these teleost Tlrs exhibit specificity in recognising PAMPs. Molecular evolution of teleost *tlr21* and *tlr22* subfamilies was examined in order to address this hypothesis. The zebrafish (*Danio rerio*) is currently used as model for the vertebrate immune system. Its adaptive immune system takes 4–6 weeks to develop and does not exhibit any adaptive immunity markers until 4 days post-hatching [17]. Therefore, embryos, larvae and juveniles up to 6 weeks of age are most useful to study the innate immune system without the complications presented by an adaptive immune response. We exploited this advantage of the zebrafish model to examine *tlr21* and *tlr22* expression during the innate immune response to various microbial components. Moreover, their developmental expression and transcriptional response to thermal stress was also determined.

Materials and methods

Molecular phylogeny and selection pressure

Danio rerio tlr21 and *tlr22* complete cDNA sequences obtained from NCBI were used as probes to search for fish-specific *tlr* coding genes in NCBI and Ensembl databases using blastn. Eighteen *tlr* genes (*tlr13*, *tlr21*, *tlr22* and *tlr23*) were found from 9 different fish species (Supplementary Table 1). cDNA sequences were aligned using MUSCLE (www.ebi.ac.uk/) and Findmodel (www.hiv.lanl.gov/) was used to identify the best nucleotide substitution model. Maximum likelihood phylogenetic analysis was carried out with PhyML [18] and bayesian inference was performed with MrBayes v3.1.2 [19]. In both instances, the GTR nucleotide model was found to best describe the data.

Codon alignments were generated separately for *tlr21* and *tlr22* genes using the software available at the HIV server (www.hiv.lanl.gov/). The highly variable C-terminal portion of the codon alignments was removed along with any stop codons. Approximately 89 % (*tlr21*) and 84 % (*tlr22*) of the total coding sequence was further used in adaptive selection tests. Tests for positive selection were

performed using the maximum likelihood methods implemented in the CODEML program of PAML [20], as detailed elsewhere [21, 22]. The ratio (ω) of non-synonymous (dN) to synonymous substitution (dS) was calculated for each codon using models M0 (neutral), M1 (nearly neutral), M2 (positive selection), M7 (beta) and M8 (beta&ω). Models were compared against each other using likelihood ratio tests [21, 23]. Bayesian posterior probabilities (p) were calculated for positively selected sites using naive empirical Bayes (NEB) and Bayes empirical Bayes (BEB).

FEL analysis was carried out in Datamonkey server [24] to calculate dN-dS values for each codon, along with the corresponding probability values. First, the nucleotide substitution model that best fitted the subfamily 21 and 22 data was determined using the software available at their server. Based on the Akaike information criterion (AIC), model 012012 (AIC = 27,069.8) and model 012032 (AIC = 50,097.6) were found to best represent the subfamily 21 and 22, respectively, and further used in FEL analysis.

Zebrafish husbandry and sample collection

Wild-type adult *D. rerio* were obtained from the local retailer (Nordland Zoosenter AS, Bodø, Norway). After 4 weeks in quarantine, the fish were transferred to a benchtop recirculation system (Aquatic Habitats, Florida, USA—model: AHT3-3) equipped with a UV lamp for water sterilisation. Adult fish were maintained at 28 °C, 12 h light: 12 h dark cycle with a flow rate of 60 L h^{-1} and 10 % daily water exchange. They were fed twice a day with TetraMin XL flakes (Tetra GmbH, Germany). Six healthy adult fish were humanely killed by over-anaesthesia in a solution of tricaine methanesulfonate (MS222) (200 mg L^{-1}) (Sigma). Samples of gills, kidney, brain, liver, spleen, gut, muscle, skin, heart and gonads were excised and immediately frozen in liquid nitrogen for subsequent RNA extraction.

Breeding tanks (Aquatic Habitats, Florida, USA) containing marbles were placed inside the aquaria holding adult fish in the evening prior to spawning and the eggs were collected the following morning. After cleaning, the embryos were transferred to sterile E3 medium and maintained at 28 °C throughout the experiment. At this temperature, embryos hatched between 48 and 72 h post fertilisation (hpf). All procedures were performed according to the guidelines from the national committee for animal experimentation (Forsøksdyrutvalget, Norway).

Temperature stress experiments

Fertilised eggs were collected after natural spawning and incubated at 3 different temperatures (23.1 ± 0.2, 27.1 ± 0.3 and 31.0 ± 0.2 °C). Four independent pools of 25 embryos were collected at the following developmental stages: fertilised egg, blastula (128 cells), germ-ring, bud, 20 somites, beginning of pharyngula and hatching (protruding mouth) and immediately frozen in liquid nitrogen for subsequent RNA extraction [25].

Stimulation experiment

All immunostimulation experiments were carried out simultaneously, with the corresponding control. A total of 360 larvae were used per experiment, which were evenly divided in 6-well plates. Lipopolysaccharide (LPS) from *Salmonella enterica typhimurium* (Sigma), polyinosinic-polycytidylic acid potassium salt (IC) (Sigma) or peptidoglycan (PG) from *Staphylococcus aureus* (Sigma) were introduced into the E3 medium at 200 μg ml^{-1} and were replaced with fresh E3 medium after one hour. An initial sample (control) was collected before introducing the stimulants. Samples were taken at 1, 2, 6, 12, 24 and 48 h post-exposure (hpe). Twelve larvae were sampled at each time point, frozen in liquid nitrogen and stored at −80 °C for subsequent RNA extraction. Experiments were carried out in triplicate.

RNA extraction and cDNA synthesis

Tissue samples and larvae were homogenised in MagNA Lyser (Roche) using the appropriate silica beads and total RNA was extracted from all samples using mirVana kit (Ambion), as per the manufacturer's protocol. After determining RNA concentration and purity using NanoDrop (Thermo scientific), quality was assessed by electrophoresis on a 1 % (w/v) agarose gel. Ten picograms of luciferase mRNA (Promega) were added to 500 ng of each RNA sample as an external reference and cDNA was synthesised using the Quantitect Reverse transcriptase kit (Qiagen), as previously reported [25].

Quantification of gene expression

Primer design for PCR

Specific primers were designed to quantify expression of the target genes *tlr21* and *tlr22* using semi-quantitative RT-PCR as well as real-time PCR (Supplementary Table 2). *Eef1a* was used as a reference gene in semi-quantitative analysis, while *luciferase* was used as an external control in real-time PCR. *Il1b* and *mxa* are markers for bacterial and viral infection, respectively, and were used to validate the immunostimulation experiment in larvae [26, 27]. Whenever possible, primers were manually designed across intron–exon boundaries to avoid potential genomic DNA contamination [28]. *D. rerio* genomic and cDNA sequences were obtained from NCBI (Supplementary Table 1) and

Spidey [29] was used to find the intron–exon boundaries. All primer pairs were screened for hairpins, homo- and cross-dimers using Netprimer (www.premierbiosoft.com/netprimer/).

Semi-quantitative RT-PCR (RT-PCR)

Tlr21, tlr22 and *eef1a* gene expression in the tissues were determined using RT-PCR. Recombinant Taq DNA polymerase (Invitrogen) was used for RT-PCR with the following thermocycling parameters: 95 °C: 2 min, 35 cycles of (95 °C: 15 s, annealing temperature (Supplementary Table 2): 30 s and 72 °C: 2 min) and 72 °C: 7 min. Amplification was carried out in a programmable thermo cycler (BioRad C1000). Gels were visualised and photographed using Kodak Gel Logic 200 Imaging System (Carestream, USA).

Real-time PCR (qPCR)

Real-time PCR was performed essentially as previously described [30]. After an initial denaturation step of 15 min at 95 °C, the desired PCR product was amplified in 45 cycles (94 °C: 15 s, annealing temperature (Supplementary Table 2): 20 s and 72 °C: 20 s). Fluorescence data were acquired during the last step in each cycle. A dissociation step with a gradient from 65 to 97 °C was performed to check the specificity of the qPCR reaction and the absence of primer dimers. Specificity was further confirmed by Sanger sequencing. Data were acquired using the Light Cycler 480 (Roche, Germany). Ct values were calculated with a fluorescence threshold of 0.5 and the average of two technical replicates was used to calculate relative gene expression. Relative expression of target genes in the temperature stress experiment was evaluated using the $\Delta\Delta C_T$ method, against the luciferase expression levels and one-way ANOVA was used for statistical analysis. For the immunostimulation experiment, REST-MCS v2 was used to calculate the relative expression of target genes. REST-MCS uses pair wise fixed reallocation randomisation test with no distributional assumption about the data to calculate the *p* values. It compares control and treatment sample groups against their mean to indicate the significance of the treatment effect [31]. For every time point in each immunostimulation experiment, data were normalised against luciferase expression levels [25] and expressed as ratios over the initial control. Prior to the reverse transcription step, ten picograms of luciferase mRNA was added to each RNA sample. After reverse transcription, the amount of luciferase cDNA present in each sample was quantified using specific primers as described above. Average Ct value of luciferase expression was provided to REST-MCS in the place of reference gene Ct values and was further used to normalise the expression of target genes.

Results and discussion

Phylogeny of teleost *tlr21* and *tlr22* subfamilies

A search for teleost-specific *tlr*s in NCBI and Ensembl databases identified 18 genes from 9 different fish species (Supplementary Table 1) encoding proteins that have a typical TLR structure (Fig. 1a). They have 25–27 LRRs at their N-terminal region, a transmembrane region and a TIR domain at the C-terminal.

The LRR domain is a protein structural motif that forms a horseshoe fold and has been identified in a large number of functionally unrelated proteins. It is involved in the formation of protein–protein interactions and in the case of TLRs, it binds to specific PAMPs thereby recognising pathogens [32]. Comparative sequence analysis of vertebrate TLRs has shown that each of the six TLR families can be characterized by its leucine-rich motifs, their repeat number and the flanking region with the LRR domain (Fig. 1a). All 18 genes analysed in this study encoded for proteins with 25–27 LRRs typical of the TLR11 family, which comprises teleost *tlr13, tlr21*, *tlr22* and *tlr23* as well as murine *Tlr11*, *Tlr12* and *Tlr13* [33].

The LRRs of most TLRs have cysteine clusters flanking either side of the LRR region with two to five cysteine residues and are denoted as LRR C-terminal (LRRCT) and LRR N-terminal (LRRNT) domains. Multiple sequence alignment showed that *tlr22*, *tlr23*, *tlr13* had a $CxCx_{24}Cx_{18}C$ motif at their LRRCT, characteristic of fish *tlr22* genes (Fig. 1b). Similar to other known teleost *tlr21*s, *D. rerio*, *Ictalurus punctatus*, *Takifugu rubripes*, *Oryzias latipes* and *Gasterosteus aculeatus tlr21* encoded a $CxCx_{24}Cx_{15}C$ motif at their LRRCT domain (Fig. 1b). While LRRNT regions are variable among TLRs, LRRCT contains a highly conserved consensus sequence and found to play a crucial role in TLR signalling [34]. The LRRCT forms a compact structure stabilised by disulphide bridges positioning the extracellular domain of the TLR relative to the membrane as seen in the structure of human TLR3 protein [35].

Both maximum likelihood and Bayesian phylogenetic analyses generated an identical phylogenetic tree (Fig. 2), dividing the teleost *tlr*s into two subfamilies. Subfamily 21 contains *tlr21* from *D. rerio*, *I. punctatus*, *T. rubripes*, *O. latipes* and *G. aculeatus*. They encode a 881–989 aa protein that shares 46–61 % identity across taxa. All teleost *tlr22, tlr23* genes along with *Salmo salar tlr13* are grouped under subfamily 22. They encode proteins of 925–973 residues, with 36–93 % identity. As expected, *S. salar*

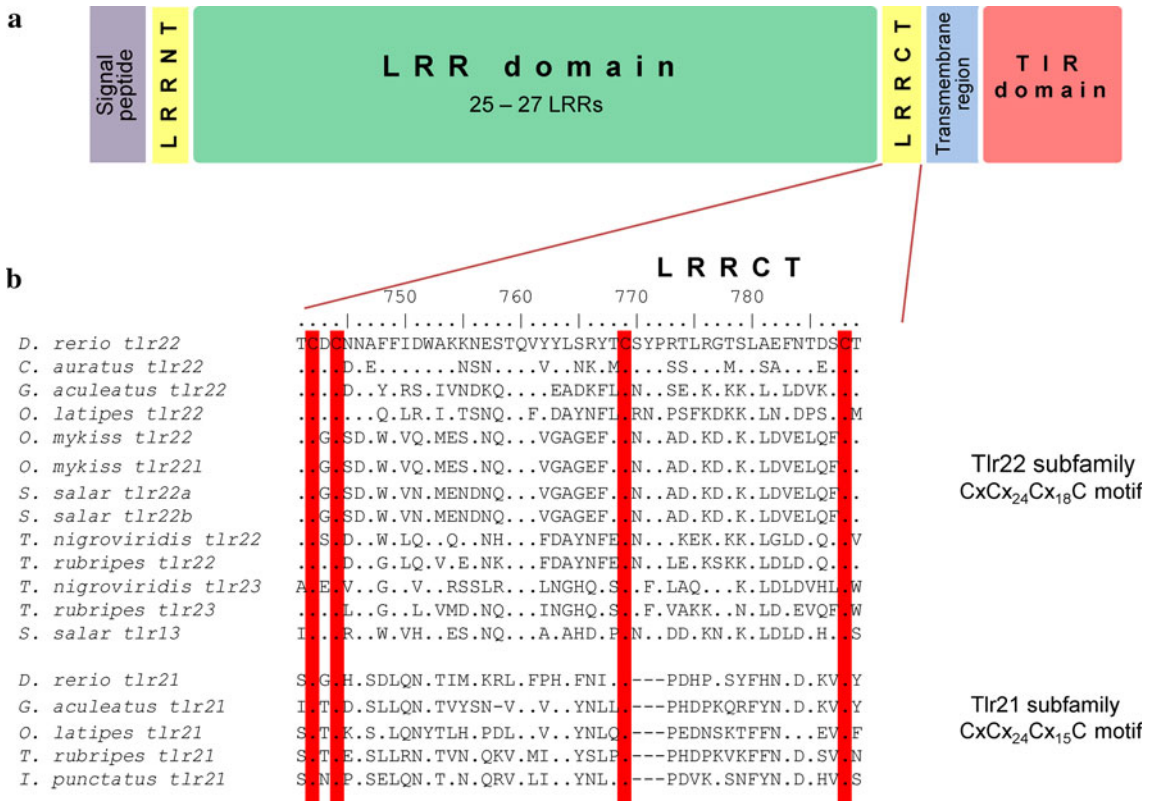

Fig. 1 TLR structure and LRRCT domain. **a** Structure of TLR protein. Signal peptide is indicated in *purple*. LRRNT and LRRCT are highlighted in *yellow*. *Green* represents LRR domain and it contains 25-27 leucine rich repeats in the case of teleost *tlr21*, *tlr22* and *tlr23*. Transmembrane region and TIR domain are represented in *blue* and *red*, respectively. **b** Multiple sequence alignment of teleost *tlr21*, *tlr22* and *tlr23* LRRCT domain. Cysteines in the typical LRRCT motif ($CxCx_{24}Cx_{(15-18)}C$) are highlighted in *red*. (Color figure online)

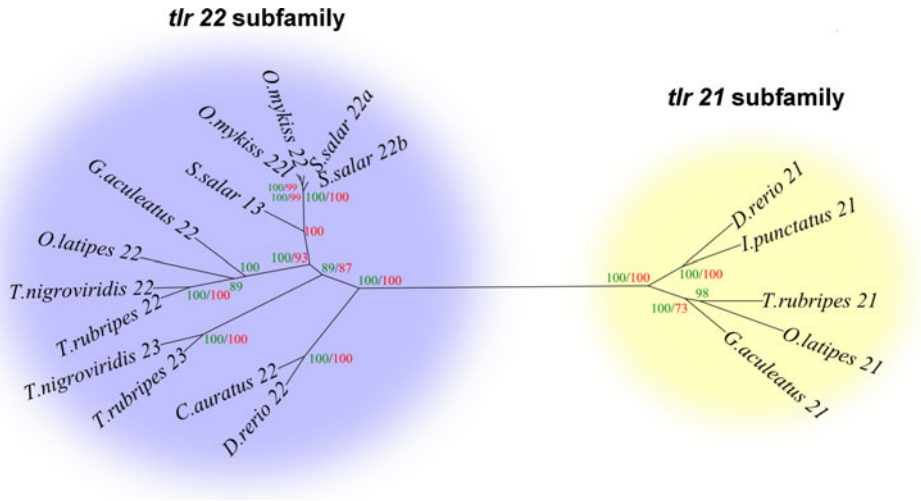

Fig. 2 Phylogenetic tree illustrating the relationship between teleost-specific *tlr*s. The two main clades corresponding to *tlr21* and *tlr22* subfamilies are shaded in *yellow* and *blue*, respectively. Percentages shown in *red* and *green* on the tree nodes indicate Bayesian posterior probabilities and maximum likelihood bootstrap values, respectively. Only values above 85 % are indicated. Accession numbers for the sequences used in the analysis are provided in Supplementary Table 1. (Color figure online)

tlr22a and *tlr22b* were clustered together and corresponded to paralogous genes, which have probably arisen from the salmonid tetraploidisation event. *Oncorhynchus mykiss tlr22* paralogues are most closely related to *S. salar tlr22a* and *tlr22b* and share up to 90 % identity with Tlr13, Tlr22 and Tlr23 orthologues at the protein level. *Tlr23* from *T. rubripes* and *Tetraodon nigroviridis* appear as an independent lineage-specific duplication in the Tetraodontidae family, as previously reported [11]. During the preparation of this manuscript, Palti [36] also raised the issue of renaming *S. salar tlr13* based on *O. mykiss* EST data. Three lines of evidence from our study confirm that the previously termed *S. salar tlr13* [37] belongs in fact to the *tlr22* subfamily: i) it shares a minimum of 41 %

Table 1 Identification of positively selected sites in *tlr22* by maximum likelihood analysis

Model	Parameter estimates	Model comparison	Positively selected sites[a]
M0: neutral	$\omega = 0.29$		None
M1: nearly neutral	$\omega_0 = 0.14$, $\omega_1 = 1.00$,		Not allowed
	$p_0 = 0.59$, $p_1 = 0.41$		
M2: positive selection	$\omega_0 = 0.14$, $\omega_1 = 1.00$, $\boldsymbol{\omega_2 = 9.52}$,	M2 vs. M1	11, **56**, 294 (BEB)
	$p_0 = 0.59$, $p_1 = 0.41$, $p_2 = 0.007$	$2\Delta LnL = 23.34$, df = 2, p = 0	
M7: β	p = 0.55, q = 0.98		Not allowed
M8: $\beta + \omega S > 1$	p = 0.56, q = 1.02, $\boldsymbol{\omega = 7.26}$,	M8 vs. M7	11, **56** (NEB)
	$p_0 = 0.99$, $p_1 = 0.008$	$2\Delta LnL = 22.17$, df = 2, p = 0	
FEL (Datamonkey)	**dN-dS = 5.725**, p = 0.06,		56
	dN-dS = 2.31, p = 0.006		294

[a] Only positively selected sites with Bayesian posterior probabilities above 95 % are indicated and the ones greater than 99 % are highlighted in bold

identity with Tlr22 but only up to 32 % with Tlr21 orthologues, ii) the LRRCT consensus sequence is identical to all other Tlr22 proteins and iii) it is clearly clustered under the subfamily 22 with the a bootstrap value of 100 %. Hence, it would be appropriate to rename *S. salar tlr13* as *tlr22c*.

Molecular evolution

Site-specific analysis in PAML identified positive selection acting on *tlr22* genes (Table 1). Comparing different models implemented in PAML based on likelihood test ratios, revealed the ones that fitted the data better than those which did not (M2 vs. M1, $p = 0.0001$; M8 vs. M7, $p = 0.0001$). All four models, M1, M2, M7 and M8 represent the data better than the neutral model M0 ($p = 0$). The model M2 implemented in PAML that allows for positive selection found three positively selected sites ($\omega = 9.52$) at positions 11 ($p = 0.99$), 56 ($p = 0.994$) and 294 ($p = 0.967$). Model M8, which is more suited for smaller data sets [38] identified positions 11 ($p = 0.975$) and 56 ($p = 0.999$) as positively selected sites ($\omega = 7.26$) (Fig. 3a, b). The FEL algorithm in Datamonkey suggested that codons 56 ($\omega = 5.725$, p value = 0.065) and 294 ($\omega = 2.31$, p value = 0.006) are under positive selection. Neither of the above mentioned methods identified positively selected sites within the teleost *tlr21* subfamily.

Toll-like receptors are generally perceived to undergo purifying selection due to the highly conserved nature of the PAMPs. Early studies did not find any evidence of positive selection in various TLR families such as TLR3, TLR5 and TLR11 using genes from distantly related organisms [11]. Since then, selection pressure in TLR has been demonstrated in closely related species. Positively selected sites were identified in mammalian TLR7 and TLR9 using sliding-window analysis of dN/dS ratios. In the case of TLR7, the codons under positive selection were present in the convex side of the horseshoe structure [39]. Evidence of selection pressure was found in 6 out of 10 primate TLR genes, being particularly strong in TLR4 and TLR1 [14]. Recently, signatures of positive selection have been shown to exist in all 10 mammalian TLRs by analysing genes from up to 23 different species [40]. In the avian genome, positively selected sites were observed in TLR2 and TLR7 genes [15]. Until now, positive selection within teleost *tlr* genes was reported only for the orthologue of mammalian TLR9. Eleven codons in *tlr9* across 8 teleost species were identified to be under positive selection, of which 10 were associated with LRRs [16].

Toll-like receptors are composed of an intracellular TIR domain [41] and an extracellular domain with a horseshoe-like solenoid structure formed by the LRR motifs [32]. The TIR domain is highly conserved across species as well as between different TLRs, since it is involved in cascade signalling [41]. In contrast, the convex side of the extracellular solenoid structure is more divergent and important for ligand binding [34]. Tlr22 codons 11 and 56 were found between LRR1-2 and LRR3-4, respectively, while codon 294 is located within LRR13 immediately after the third leucine of the LRR motif, LxxLxLxxNxL. Codons 11 and 56 were found on the convex side in the coils, whereas codon 294 was found on the concave beta sheet (Fig. 3). Several amino acid substitutions were observed in these positively selected sites. In particular, positively charged amino acids arginine and lysine were present in five out of thirteen Tlr22 proteins at codon 11, which might improve binding to negatively charged PAMPs such as lipopolysaccharides. On the other hand, the tryptophan residue at this position causes spatial hindrance that may affect the ligand specificity of salmon Tlr22a (Fig. 3).

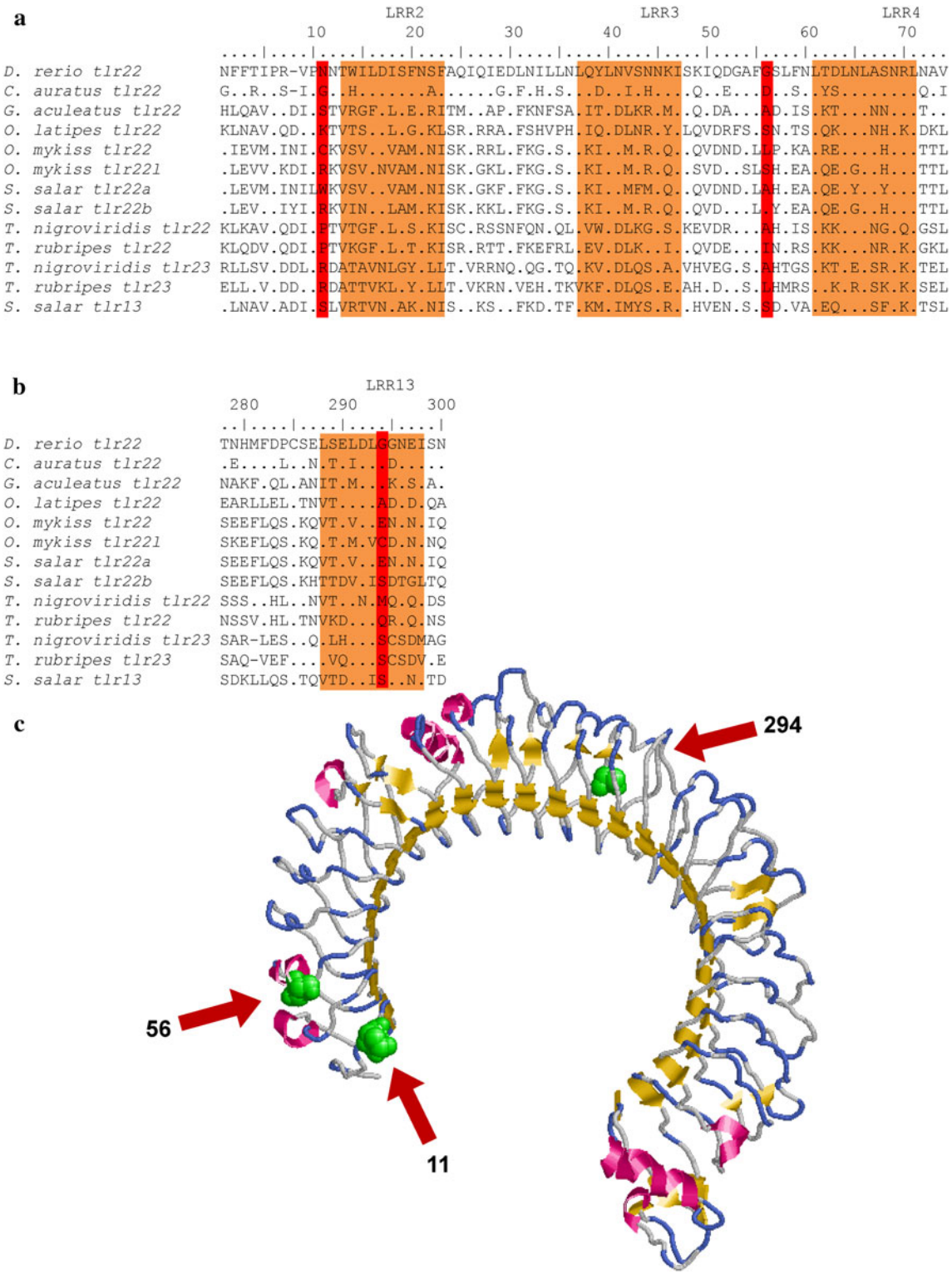

Fig. 3 Multiple sequence alignment of teleost-specific Tlr22 proteins. Amino acid residues identical to *D. rerio* Tlr22 are represented by a *dot* and alignment gaps are indicated by a *dash*. LRR regions are shaded in *orange* and positively selected sites are boxed in *red*. **a** and **b** represents the region containing the first two positively selected positions and the third positively selected site, respectively. **c** Predicted structure of *D. rerio* Tlr22 LRR region with the positively selected sites highlighted in *green* and their amino acid position. (Color figure online)

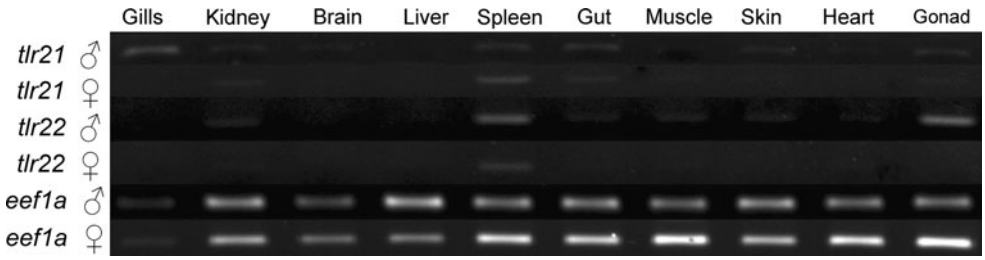

Fig. 4 Tissue distribution of *tlr21* and *tlr22* in adult *Danio rerio*. Representative semi-quantitative PCR showing the expression of *tlr21* and *tlr22* in various tissues of adult zebrafish. *Tlr21* and *tlr22* are mainly expressed in immune-related organs such as spleen and kidney as well as in testis. Expression of *eef1a* was used as endogenous control

Almost all the positively selected sites found in mammalian, avian and teleost *tlr* genes were located in LRR regions, predominantly at their convex side. This suggests that the variation in these sites are involved in pathogen recognition and could be important for host-pathogen co-evolution [13, 40], as for the variation in the peptide binding residues of the MHC molecules involved in the adaptive immune response [42]. Until now Tlr22 proteins were not thought to be under any selection pressure. Our study indicates that Tlr22 proteins are under significant positive selection and the amino acid substitutions may produce striking variations in the structure of the PAMP recognition sites. These changes alter their ability to detect charged as well as structurally varied ligands and may therefore be associated with adaptations to evolving pathogens. This is indeed a plausible hypothesis, since fish are an extremely diverse group and are exposed to numerous different families of pathogens in their aquatic environment.

Tissue distribution

The *D. rerio* genome encodes only for *tlr21* and *tlr22* but the third teleost *tlr* homologue (*tlr23*) is absent. Both *tlr21* and *tlr22* were mainly expressed in immune-related organs, namely spleen and kidney (Fig. 4). A similar pattern of expression was observed in *T. rubripes, O. mykiss*, *Carassius auratus* and *Pseudosciaena crocea* [43–46]. The spleen, which plays a key role in the teleost immune system, had the highest expression of both *D. rerio tlr21* and *tlr22*, similarly to what has been observed for human TLR1, TLR4, TLR6, TLR8 and TLR10 [47] as well as in *O. mykiss* [43]. *D. rerio tlr21* was present in the gills and low transcript levels of *tlr21* and *tlr22* were found in muscle and skin as well. In contrast, *P. crocea* and *I. punctatus* did not express *tlr22* and *tlr21* in muscle, respectively [45, 48]. It is noteworthy that *tlr22* expression was observed in *D. rerio* testis. This is the first line of evidence indicating that this *tlr* might be important in protecting the male reproductive tract. Albeit absent in *D. rerio* ovary, *tlr21* transcripts were found in *I. punctatus* ovary [48], thus reflecting a species-specific variation in *tlr21* expression across tissues.

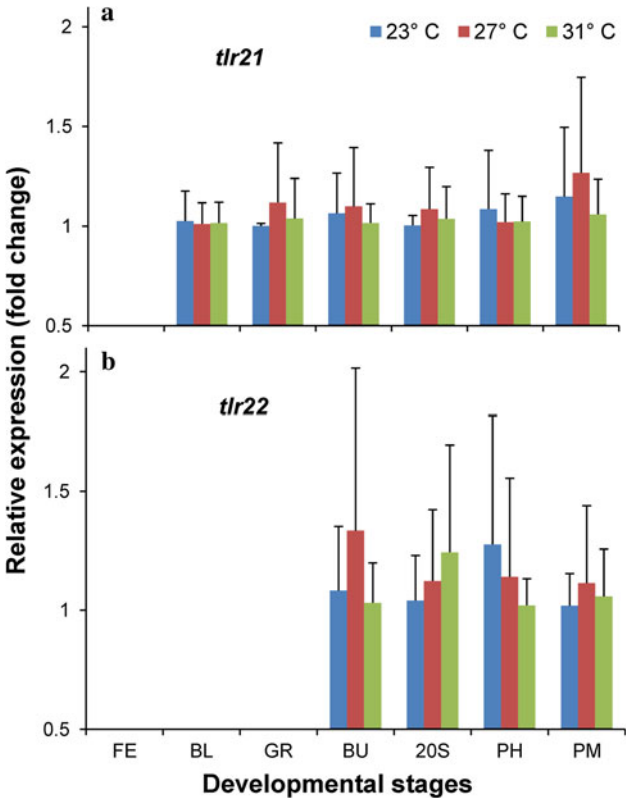

Fig. 5 Temperature stress of *tlr21* and *tlr22* in adult *Danio rerio*. Relative expression of *tlr21* (**a**) and *tlr22* (**b**) in zebrafish embryos incubated at 23 °C (*blue bars*), 27 °C (*red bars*) or 31 °C (*green bars*). Expression was quantified by qPCR and normalised against luciferase mRNA levels as external control. Separate analyses were performed for each gene and developmental stage. Abbreviations are *FE* fertilised egg, *BL* blastula, *GR* germ-ring, *BU* bud, *20S* 20 somites, *PH* start of pharyngula, *PM* protruding mouth. (Color figure online)

Effect of temperature stress

Developmental expression patterns of *tlr21* and *tlr22* subfamily genes had not been examined in any fish species to date. In the present study, expression of *D. rerio tlr21* and *tlr22* and the effect of temperature on their transcript levels

at 23, 27 and 31 °C were determined using qPCR. *Tlr21* was detected at all stages, while *tlr22* was not found to be expressed until bud stage (Fig. 5). Both genes were not detected in fertilised eggs, indicating that they are not maternally transferred. Taking into account biological variation, it is plausible to conclude that temperature did not have any significant effect on the expression of *tlr21* and *tlr22* during the early embryonic development of *D. rerio*.

Differential expression of *tlr21* / *tlr22* upon immunostimulation

To ascertain the specificity and response of *tlr21* and *tlr22* to different PAMPs in *D. rerio*, an immunostimulation experiment was carried out by exposing 72 hpf *D. rerio* larvae to microbial components that mimic Gram-positive and Gram-negative bacteria as well as virus. Pilot experiments were performed to determine the lethal dose of microbial components by exposing 72 hpf *D. rerio* larvae to LPS, PG or IC up to 250 μg ml^{-1}. *Escherichia coli* 0111:B4 LPS at 150 μg ml^{-1} and *Pseudomonas aeruginosa* LPS at 50 μg ml^{-1} have been previously reported as the lethal doses for two day old *D. rerio* larvae [26]. Amongst other factors, immune system development, the process of dechorionation and the LPS virulence influence the lethal dose of microbial components. For example, dechorionated embryos are more sensitive to LPS, tolerating concentrations up to only 100 μg ml^{-1} [49]. In our pilot experiments, 250 μg ml^{-1} of LPS was found to be lethal, whilst in both IC and PG trials more than 75 % of the larvae survived at the same concentration. Therefore larvae were exposed to 200 μg ml^{-1} of any one of the immunostimulants for a period of one hour.

Il1b, a key cytokine involved in the inflammatory response is known to be up-regulated in *D. rerio* larvae upon stimulation with bacterial components [26]. Exposure to IC did not induce *il1b* expression in larvae, whereas a two- to three-fold increase in *il1b* transcript level was seen in LPS and PG stimulated larvae, albeit not statistically significant (Fig. 6). *Mxa*, which is used as a viral infection marker in vertebrates [27], was measured in IC experiment samples and a two-fold change in expression was observed at one hpe (data not shown). Both *il1b* and *mxa* expression changes confirm the immunostimulatory effect of LPS, PG and IC.

Tlr21 expression increased between two- to three-fold immediately after one hour exposure to LPS, PG or IC and returned to basal levels at 2 hpe. This is the first study to document the induction of *tlr21* in fish larvae upon immunostimulation. Chicken *Tlr21* is considered as a functional homologue to mammalian TLR9 and it recognises CpG oligonucleotides [50]. Based on the high

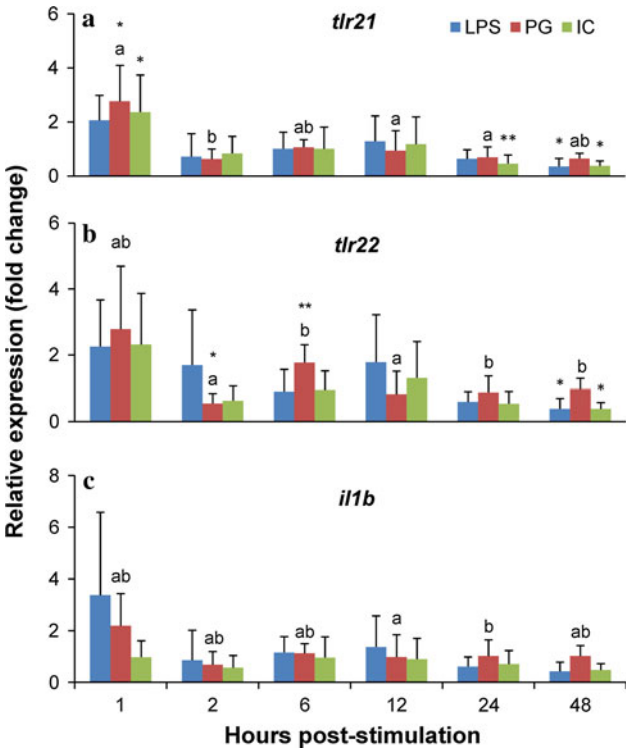

Fig. 6 Relative expression of *tlr21*, *tlr22* and *il1b* in *Danio rerio* larvae exposed to various immunostimulants *D. rerio* larvae were exposed to lipopolysaccharide (LPS) from *Salmonella enterica typhimurium*, polyinosinic-polycytidylic acid potassium salt (IC) or peptidoglycan (PG) from *Staphylococcus aureus* for one hour at 200 μg ml^{-1} concentration. The graph illustrates the relative expression of *tlr21*, *tlr22* and *il1b* calculated as a ratio over the control (sample before exposure) and normalised against *luciferase* mRNA levels as external control. Expression changes were determined by qPCR at 1, 2, 6, 12, 24 and 48 h post-stimulation. **, * represents the significant expression ratios (experiment/control) with $p < 0.05$ and $p < 0.1$, respectively. **a**, **b**, **c** show significant differences within each immunostimulant experiment across the time scale ($p < 0.05$)

similarity shared by chicken and teleost *tlr21*, it is plausible that the latter one also recognises CpG oligonucleotides but also responds to various microbial components as shown from the immunostimulation experiment. *Tlr22* showed a similar pattern to *tlr21* expression with the highest fold change of 2.8 in PG stimulation immediately after exposure. Following an initial up-regulation in *tlr22* expression upon exposure to PG, transcript levels returned to basal levels at 2 hpe before increasing over twofold at 6 hpe. After this time point, *tlr22* expression was similar in immunostimulated and control samples. This correlates with a previous report of *tlr22* response to PG in Japanese flounder peripheral blood leucocytes [51]. Following LPS stimulation, *tlr22* expression was nearly twice as that of control until 12 hpe (except 6 hpe) before returning to normal levels. Response of teleost *tlr22* to LPS has been confirmed in *C. auratus* leucocytes and macrophages [44]. *T. rubripes tlr22* is thought to be a functional substitute for

human cell-surface TLR3, as it responds to long dsRNAs on the cell surface and may therefore promote antiviral protection in teleosts [52], whereas, *D. rerio tlr22* responded significantly towards PG exposure compared to IC stimulation experiment.

Teleost *tlr*s have been shown to respond to different PAMPs, suggesting that these *tlr*s play a crucial role in the innate defence against pathogens. Increase in expression of *tlr22* was found in LPS stimulated macrophages as well as in LPS, *Aeromonas salmonicida* or *Mycobacterium cheloni* stimulated leucocytes in *C. auratus* [44]. In vivo pathogen infection with *M. marinum* up-regulated expression of *tlr22* in adult *D. rerio* after an extended incubation period [53]. Continuous exposure of *O. mykiss* PBL, spleen and kidney to inactivated *A. salmonicida*, induced up to eightfold increase in expression of *tlr22* and *tlr22l* after 24 h [43]. *D. rerio tlr21* and *tlr22* have showed similar expression patterns upon challenge with either immunostimulants. Response of *tlr22* to immunostimulants in *D. rerio* larvae and other fish species along with the evidence of positive selection within the LRR region of Tlr22 confirms the fact that TLRs are capable of recognising structurally unrelated PAMPs. Further research has to be carried out to investigate the specific PAMPs being recognised by these teleost Tlrs, in order to ascertain their role in host defence.

Conclusions

To date, seventeen different *tlr* genes have been identified in fish, some of which are absent in mammals. Ohnologous genes have also been recorded in teleost genomes. We have shown that all *tlr13, tlr21, tlr22* and *tlr23* genes can be grouped under two clades corresponding to *tlr21* and *tlr22* subfamilies. Expression of *tlr21* and *tlr22* genes in immune-related organs and gonads highlights the potential involvement of these genes in host defence. Immunostimulation experiments confirmed that they were activated by various microbial components and may therefore be involved in pathogen recognition. Positive selection has been shaping the extracellular LRR domain of *tlr22* genes in teleosts, which is likely involved in adaptation to new or evolving pathogens and their PAMPs.

Acknowledgments Mr. Sundaram's PhD is funded by a scholarship from the University of Nordland (Norway). We are grateful to Dr. Alessia Giannetto, Dr. Kazue Nagasawa, Ms. Catarina Campos and Mr. Spyros Kollias for their help in laboratory work and fish husbandry.

References

1. Smith VJ, Fernandes JMO (2009) Antimicrobial peptides of the innate immune system. In: Zaccone G, Perriere C, Mathis A, Kapoor BG (eds) Fish defences: pathogens, parasites and predators. Science Publishers, USA, pp 241–276. doi:10.1201/b10188-9
2. Janeway CA Jr, Medzhitov R (2002) Innate immune recognition. Annu Rev Immunol 20:197–216. doi:10.1146/annurev.immunol.20.083001.084359
3. Medzhitov R, Janeway CA Jr (1998) Innate immune recognition and control of adaptive immune responses. Semin Immunol 10(5):351–353. doi:10.1006/smim.1998.0136
4. Kumar H, Kawai T, Akira S (2009) Toll-like receptors and innate immunity. Biochem Biophys Res Commun 388(4):621–625. doi:10.1016/j.bbrc.2009.08.062
5. Medzhitov R (2007) Recognition of microorganisms and activation of the immune response. Nature 449(7164):819–826. doi:10.1038/nature06246
6. Akira S, Uematsu S, Takeuchi O (2006) Pathogen recognition and innate immunity. Cell 124(4):783–801. doi:10.1016/j.cell.2006.02.015
7. Jin MS, Lee JO (2008) Structures of the toll-like receptor family and its ligand complexes. Immunity 29(2):182–191. doi:10.1016/j.immuni.2008.07.007
8. Kawai T, Akira S (2010) The role of pattern-recognition receptors in innate immunity: update on toll-like receptors. Nat Immunol 11(5):373–384. doi:10.1038/ni.1863
9. Hashimoto C, Hudson KL, Anderson KV (1988) The *Toll* gene of drosophila, required for dorsal-ventral embryonic polarity, appears to encode a transmembrane protein. Cell 52(2):269–279. doi:10.1016/0092-8674(88)90516-8
10. Lemaitre B, Nicolas E, Michaut L, Reichhart JM, Hoffmann JA (1996) The dorsoventral regulatory gene cassette spatzle/Toll/cactus controls the potent antifungal response in *Drosophila* adults. Cell 86(6):973–983. doi:10.1016/S0092-8674(00)80172-5
11. Roach JC, Glusman G, Rowen L, Kaur A, Purcell MK, Smith KD, Hood LE, Aderem A (2005) The evolution of vertebrate toll-like receptors. Proc Natl Acad Sci USA 102(27):9577–9582. doi:10.1073/pnas.0502272102
12. Brownlie R, Allan B (2011) Avian toll-like receptors. Cell Tissue Res 343(1):121–130. doi:10.1007/s00441-010-1026-0
13. Rebl A, Goldammer T, Seyfert HM (2010) Toll-like receptor signaling in bony fish. Vet Immunol Immunopathol 134(3–4):139–150. doi:10.1016/j.vetimm.2009.09.021
14. Wlasiuk G, Nachman MW (2010) Adaptation and constraint at toll-like receptors in primates. Mol Biol Evol 27(9):2172–2186. doi:10.1093/molbev/msq104
15. Cormican P, Lloyd AT, Downing T, Connell SJ, Bradley D, O'Farrelly C (2009) The avian Toll-Like receptor pathway-subtle differences amidst general conformity. Dev Comp Immunol 33(9):967–973. doi:10.1016/j.dci.2009.04.001
16. Chen JS, Wang TY, Tzeng TD, Wang CY, Wang D (2008) Evidence for positive selection in the TLR9 gene of teleosts. Fish Shellfish Immunol 24(2):234–242. doi:10.1016/j.fsi.2007.11.005
17. Novoa B, Figueras A (2012) Zebrafish: model for the study of inflammation and the innate immune response to infectious diseases. Adv Exp Med Biol 946:253–275. doi:10.1007/978-1-4614-0106-3_15
18. Guindon S, Delsuc F, Dufayard JF, Gascuel O (2009) Estimating maximum likelihood phylogenies with PhyML. Methods Mol Biol 537:113–137. doi:10.1007/978-1-59745-251-9_6
19. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19(12):1572–1574. doi:10.1093/bioinformatics/btg180
20. Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol 24(8):1586–1591. doi:10.1093/molbev/msm088
21. Fernandes JM, Ruangsri J, Kiron V (2010) Atlantic cod piscidin and its diversification through positive selection. PLoS One 5(3):e9501. doi:10.1371/journal.pone.0009501
22. Consuegra S, Megens HJ, Leon K, Stet RJM, Jordan WC (2005) Patterns of variability at the major histocompatibility class II

alpha locus in Atlantic salmon contrast with those at the class I locus. Immunogenetics 57(1):16–24. doi:10.1007/s00251-004-0765-z
23. Sarropoulou E, Fernandes JMO, Mitter K, Magoulas A, Mulero V, Sepulcre MP, Figueras A, Novoa B, Kotoulas G (2010) Evolution of a multifunctional gene: the warm temperature acclimation protein Wap65 in the European seabass *Dicentrarchus labrax*. Mol Phylogenet Evol 55(2):640–649. doi:10.1016/j.ympev.2009.10.001
24. Delport W, Poon AFY, Frost SDW, Kosakovsky Pond SL (2010) Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. Bioinformatics 26(19):2455–2457. doi:10.1093/bioinformatics/btq429
25. Campos C, Valente LMP, Fernandes JMO (2012) Molecular evolution of zebrafish *dnmt3* genes and thermal plasticity of their expression during embryonic development. Gene 500(1):93–100. doi:10.1016/j.gene.2012.03.041
26. Novoa B, Bowman TV, Zon L, Figueras A (2009) LPS response and tolerance in the zebrafish (*Danio rerio*). Fish Shellfish Immunol 26(2):326–331. doi:10.1016/j.fsi.2008.12.004
27. Johansen A, Collet B, Sandaker E, Secombes CJ, Jørgensen JB (2004) Quantification of Atlantic salmon type-I interferon using an Mx1 promoter reporter gene assay. Fish Shellfish Immunol 16(2):173–184. doi:10.1016/s1050-4648(03)00060-3
28. Campos C, Valente LMP, Borges P, Bizuayehu T, Fernandes JMO (2010) Dietary lipid levels have a remarkable impact on the expression of growth-related genes in Senegalese sole (*Solea senegalensis Kaup*). J Exp Biol 213(2):200–209. doi:10.1242/jeb.033126
29. Wheelan SJ, Church DM, Ostell JM (2001) Spidey: a tool for mRNA-to-genomic alignments. Genome Res 11(11):1952–1957. doi:10.1101/gr.195301
30. Fernandes JM, Mommens M, Hagen O, Babiak I, Solberg C (2008) Selection of suitable reference genes for real-time PCR studies of Atlantic halibut development. Comp Biochem Physiol B Biochem Mol Biol 150(1):23–32. doi:10.1016/j.cbpb.2008.01.003
31. Pfaffl MW, Horgan GW, Dempfle L (2002) Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. Nucleic Acids Res 30(9):e36. doi:10.1093/nar/30.9.e36
32. Kobe B, Kajava AV (2001) The leucine-rich repeat as a protein recognition motif. Curr Opin Struct Biol 11(6):725–732. doi:10.1016/s0959-440x(01)00266-4
33. Matsushima N, Tanaka T, Enkhbayar P, Mikami T, Taga M, Yamada K, Kuroki Y (2007) Comparative sequence analysis of leucine-rich repeats (LRRs) within vertebrate toll-like receptors. BMC Genomics 8:124–143. doi:10.1186/1471-2164-8-124
34. Bell JK, Mullen GED, Leifer CA, Mazzoni A, Davies DR, Segal DM (2003) Leucine-rich repeats and pathogen recognition in toll-like receptors. Trends Immunol 24(10):528–533. doi:10.1016/s1471-4906(03)00242-4
35. Bell JK, Botos I, Hall PR, Askins J, Shiloach J, Segal DM, Davies DR (2005) The molecular structure of the Toll-like receptor 3 ligand-binding domain. Proc Natl Acad Sci USA 102(31):10976–10980. doi:10.1073/pnas.0505077102
36. Palti Y (2011) Toll-like receptors in bony fish: from genomics to function. Dev Comp Immunol 35(12):1263–1272. doi:10.1016/j.dci.2011.03.006
37. Leong J, Jantzen S, von Schalburg K, Cooper G, Messmer A, Liao N, Munro S, Moore R, Holt R, Jones S, Davidson W, Koop B (2010) Salmo salar and Esox lucius full-length cDNA sequences reveal changes in evolutionary pressures on a post-tetraploidization genome. BMC Genomics 11(1):279–295. doi:10.1186/1471-2164-11-279
38. Yang Z, Wong WS, Nielsen R (2005) Bayes empirical Bayes inference of amino acid sites under positive selection. Mol Biol Evol 22(4):1107–1118. doi:10.1093/molbev/msi097
39. Park S, Park D, Jung Y-J, Chung E, Choi S (2010) Positive selection signatures in the TLR7 family. Genes Genomics 32(2):143–150. doi:10.1007/s13258-009-0837-4
40. Areal H, Abrantes J, Esteves P (2011) Signatures of positive selection in Toll-like receptor (TLR) genes in mammals. BMC Evol Bio 11(1):368. doi:10.1186/1471-2148-11-368
41. Werling D, Jann OC, Offord V, Glass EJ, Coffey TJ (2009) Variation matters: TLR structure and species-specific pathogen recognition. Trends Immunol 30(3):124–130. doi:10.1016/j.it.2008.12.001
42. John T (2011) The MHC, disease and selection. Immunol Lett 137(1–2):1–8. doi:10.1016/j.imlet.2011.01.002
43. Rebl A, Siegl E, Köllner B, Fischer U, Seyfert H-M (2007) Characterization of twin toll-like receptors from rainbow trout (*Oncorhynchus mykiss*): evolutionary relationship and induced expression by *Aeromonas salmonicida salmonicida*. Dev Comp Immunol 31(5):499–510. doi:10.1016/j.dci.2006.08.007
44. Stafford JL, Ellestad KK, Magor KE, Belosevic M, Magor BG (2003) A toll-like receptor (TLR) gene that is up-regulated in activated goldfish macrophages. Dev Comp Immunol 27(8):685–698. doi:10.1016/s0145-305x(03)00041-7
45. Xiao X, Qin Q, Chen X (2011) Molecular characterization of a toll-like receptor 22 homologue in large yellow croaker (*Pseudosciaena crocea*) and promoter activity analysis of its 5′-flanking sequence. Fish Shellfish Immunol 30(1):224–233. doi:10.1016/j.fsi.2010.10.014
46. Oshiumi H, Tsujita T, Shida K, Matsumoto M, Ikeo K, Seya T (2003) Prediction of the prototype of the human Toll-like receptor gene family from the pufferfish, *Fugu rubripes*, genome. Immunogenetics 54(11):791–800. doi:10.1007/s00251-002-0519-8
47. Nishimura M, Naito S (2005) Tissue-specific mRNA expression profiles of human toll-like receptors and related genes. Biol Pharm Bull 28(5):886–892. doi:10.1248/bpb.28.886
48. Baoprasertkul P, Xu P, Peatman E, Kucuktas H, Liu Z (2007) Divergent toll-like receptors in catfish (*Ictalurus punctatus*): TLR5S, TLR20, TLR21. Fish Shellfish Immunol 23(6):1218–1230. doi:10.1016/j.fsi.2007.06.002
49. Watzke J, Schirmer K, Scholz S (2007) Bacterial lipopolysaccharides induce genes involved in the innate immune response in embryos of the zebrafish (*Danio rerio*). Fish Shellfish Immunol 23(4):901–905. doi:10.1016/j.fsi.2007.03.004
50. Brownlie R, Zhu J, Allan B, Mutwiri GK, Babiuk LA, Potter A, Griebel P (2009) Chicken TLR21 acts as a functional homologue to mammalian TLR9 in the recognition of CpG oligodeoxynucleotides. Mol Immunol 46(15):3163–3170. doi:10.1016/j.molimm.2009.06.002
51. Hirono I, Takami M, Miyata M, Miyazaki T, Han H-J, Takano T, Endo M, Aoki T (2004) Characterization of gene structure and expression of two toll-like receptors from Japanese flounder, *Paralichthys olivaceus*. Immunogenetics 56(1):38–46. doi:10.1007/s00251-004-0657-2
52. Matsuo A, Oshiumi H, Tsujita T, Mitani H, Kasai H, Yoshimizu M, Matsumoto M, Seya T (2008) Teleost TLR22 recognizes RNA duplex to induce IFN and protect cells from birnaviruses. J Immunol 181(5):3474–3485
53. Meijer AH, Gabby Krens SF, Medina Rodriguez IA, He S, Bitter W, Ewa Snaar-Jagalska B, Spaink HP (2004) Expression analysis of the toll-like receptor and TIR domain adaptor families of zebrafish. Mol Immunol 40(11):773–783. doi:10.1016/j.molimm.2003.10.003