Molecular Immunology 106 (2019) 135–142

Contents lists available at ScienceDirect

Molecular Immunology

journal homepage: www.elsevier.com/locate/molimm

# Widespread positive selection on cetacean TLR extracellular domain

Shixia Xu*, Ran Tian, Yurui Lin, Zhenpeng Yu, Zepeng Zhang, Xu Niu, Xiaohong Wang, Guang Yang*

*Jiangsu Key Laboratory for Biodiversity and Biotechnology, College of Life Sciences, Nanjing Normal University, 1 Wenyuan Road, Nanjing, 210023, China*

A R T I C L E   I N F O

*Keywords:*
Cetaceans
Innate immune
Toll-like receptor
Positive selection

A B S T R A C T

Toll like receptors (TLRs), key members of innate immune system, can recognize a wide diversity of pathogens and initiate both innate and adaptive immune responses in vertebrate. Cetaceans must have faced new challenges of pathogens when their terrestrial relatives transitioned from the terrestrial to aquatic environment. Here, we sequenced the extracellular domain (ECD) of 10 TLRs in cetacean lineages because this region involved in the recognition of pathogens. A total of 148 sites ranging between 5–26 codons (0.01%–4.83%) were identified to be robust candidates of positive selection at the ECD of 10 TLRs. In addition, the majority (90.54%) of these positively selected codons were found to have radical amino acid changes, which strengthen the evidence of positive selection. Importantly, more radical amino acid changes in selected sites were enriched in the period of early evolutionary transition from land to semi-aquatic and from semi-aquatic to full-aquatic habitat, which might endow cetaceans with a faster adaptation to new pathogens as they transitioned into novel habitat. Interestingly, similar selective intensity was detected in both viral and non-viral TLRs in cetaceans, which was not in line with previous studies on primates and birds that reported stronger positive selection in non-viral TLRs than in viral TLRs. This result may be explained by the fact that cetaceans might have faced diversity of bacteria and viruses during its transitions from terrestrial to aquatic environment whereas both primates and birds probably being affected by only a restricted number of related viruses due to their homogeneous habitat.

## 1. Introduction

The mammalian immune system is extremely complex, as such loci experience a variety of selection pressures from a wide diversity of coevolving pathogens with which animals are associated (Hedrick, 2004; Piertney and Oliver, 2006). Immunological defenses in mammals comprise two immunological subsystems-innate and adaptive. In the past decades, intensive studies on exploring the selection patterns have been mainly confined to acquired immunity. Among them, genes of the major histocompatibility complex (MHC) have received more attention than other genes since they are considered as a prime example of the effects of balancing selection maintaining the gene polymorphism even in wild populations (Xu et al., 2007, 2008, 2009, 2010). However, the role of selection in shaping diversity of innate immunity remains poorly understood in natural populations (Shen et al., 2012).

The innate immune system is an evolutionarily ancient system and represents the first line of defense against invading pathogens (Iwasaki and Medzhitov, 2010). One important family of innate-immunity genes is the toll-like receptors (TLRs), which can recognize a wide diversity of pathogens and initiate both innate and adaptive immune responses in vertebrate (Kumar and Yu, 2006; Cook et al., 2004; Iwasaki and Medzhitov, 2010). TLRs are a type of pattern recognition receptor (PRR) and recognize molecules that are broadly shared by pathogens but distinguishable from host molecules, collectively referred to as pathogen-associated molecular patterns (PAMPs).

To date, 13 mammalian TLR family members have been identified according to their functions and sequences, of which 10 members are present in the human genome (TLR1–10) and 13 in rodents (TLR1-13). Generally, the TLRs of mammals were classified into two different categories according to their ligand recognition and cellular sublocalization. The first category including TLR1, TLR2, TLR4, TLR5, TLR6, and TLR10 is expressed on cell surface and recognize predominantly bacterial ligands (but also several fungal and parasite ligands), which are termed non-viral TLRs. The second category includes TLR3, TLR7, TLR8 and TLR9 recognize single and double-stranded viral RNA, accordingly, which are termed viral TLRs (Carty and Bowie, 2010; Barton, 2007). These TLRs are expressed mostly within cells into the membranes of the endosomal compartments.

* Corresponding authors.
*E-mail addresses:* xushixia78@163.com (S. Xu), tianrannjnu@163.com (R. Tian), linyurui0214@163.com (Y. Lin), yuzhenpeng2013@163.com (Z. Yu), mtyzzp@163.com (Z. Zhang), niuxu88@163.com (X. Niu), 674003921@qq.com (X. Wang), gyang@njnu.edu.cn (G. Yang).

https://doi.org/10.1016/j.molimm.2018.12.022
Received 21 August 2018; Received in revised form 4 December 2018; Accepted 20 December 2018
Available online 29 December 2018
0161-5890/ 

The TLR molecules are structurally characterized by the presence of an extra-cellular domain (ECD), a signal transmembrance segment and a highly conserved toll/interleukin I-receptor (TIR) domains in their intracellular region (Roach et al., 2005; Werling et al., 2009). The ECD consists a solenoid horseshoe-like structure constituted by at variable number (16–28) of leucine-rich repeats (LRRs) that are responsible for ligand interactions. Specially, the ECD is involved in the Ligand Binding Region (LBR) that is directly responsible for physical interactions with the pathogen-derived structures and as such it evolves faster than the TIR domain identified in the mammals and birds (Mikami et al., 2012; Barreiro and Quintana-Murci, 2010; Wlasiuk and Nachman, 2010; Grueber et al., 2014). Thus, the ECD is likely subject to intensive selection.

Cetaceans (whales, dolphins, and porpoises) are highly specialized mammals that "returned" from the land to the sea approximately 53–56 million years and subsequently adapted to completely aquatic habitats all over the world (Thewissen et al., 2007). Cetaceans must have faced new challenges of pathogens when transitioning from the terrestrial to aquatic environment. Most cetaceans fully rely on the marine ecosystems for their existence whereas several species of cetaceans are found exclusively in riverine systems, such as including the baiji (*Lipotes vexillifer*), the Amazon river dolphin (*Inia geoffrensis*), and La Plata dolphin (*Pontoporia blainvillei*), etc. During the habitat transition, cetaceans have changed their feeding habits from a herbivorous to a carnivorous diet (mainly composed of fishes, large squids, and zooplanktons). Interestingly, some cetaceans are long-distance migrants, such as grey whale (*Eschrichtius robustus*), humpback whale (*Megaptera novaeangliae*) and sperm whale (*Physeter catodon*). During their migrations, these cetaceans are likely to increase parasite infection because parasite faunas may vary not only between geographical areas but also over time. Thus, habitat transition from the land to the sea and even to the riverine systems, adaptive changes in dietary as well as long-distance migration could make contribution in the evolutionary dynamics of cetacean immune system. So its immune system has attracted intense interest for ecologists and immunologists since 1990s. Previous studies showed that cetaceans have considerable gene polymorphism at the MHC genes when compared with terrestrial mammals (Baker et al., 2006; Xu et al., 2007, 2008, 2009, 2010, 2012). These results contradicted with traditional view that selection pressure for MHC polymorphism may not be so pronounced in marine mammals compared to that of the terrestrial mammals because of relatively less pathogenic microorganisms in the aquatic environment compared with the terrestrial ecosystem. However, innate immune studies in cetaceans have just begun recently. Two recent studies investigated the TLR4 gene in cetaceans and found that significant positive selection or unique amino acid residue substitution identified in the import domain of the TLR4 (Shen et al., 2012; Shishido et al., 2010). Therefore, more TLR genes should be investigated in order to comprehensively understand the innate immune evolution in cetaceans.

In the present study, the entire ECD of the 10 TLRs were investigated in representative species of cetaceans because this region responsible for PAMP binding is thought to have greater variation than the TIR domain (Werling et al., 2009; Areal et al., 2011). First, we aimed to document evolutionary histories of these 10 TLRs during cetacean evolution. Particularly, we tested whether some specific lineages that evolved fast radiation or undertook different habitats were subject to strong selective pressure. Second, the selective pattern of non-viral and viral TLRs were compared to reveal whether different patterns of molecular evolution could be detected in the two subclasses. Finally, we evaluated the functional significance of the positively selected sites (PSSs) by determining whether these PSSs were located within or near to the functional domain of predicated 3D structure.

## 2. Materials and methods

### 2.1. TLRs extracellular domains predication, amplification and sequencing

The TLR 1–10 protein sequences of the bottlenose dolphin (*Tursiops truncates*, http://uswest.ensembl.org/Tursiops_truncatus/Info/Index) were first downloaded from Ensemble. The secondary structure of each TLR gene was further predicated by SMART 7.0 (Letunic et al., 2015). Accordingly, the extracellular domains of each TLR were predicated based on the SMART results. Finally, primers for Polymerase Chain Reaction (PCR) were designed for the conserved regions according to an alignment of genomic data of the bottlenose dolphin, cow (*Bos taurus*, http://uswest.ensembl.org/Bos_taurus/Info/Index), sheep (*Ovis aries* http://uswest.ensembl.org/Ovis_aries/Info/Index).

Genomic DNA was extracted from 14 cetacean representative species (2 mysticetes and 12 odonticetes, supplementary Table S1, Supplementary Material online) and a hippopotamus (*Hippopotamus amphibious*, Table S1) using the DNAeasy tissue kit (Qiagen), following the manufacturer's protocol. All these samples used in our study were collected from dead individuals in the wild so that no ethical statement was required. The conditions for the PCR and sequencing were denoted as in previously studies (e.g., Shen et al., 2012). Newly sequenced TLRs by PCRs have been deposited at Genbank under the accessions MH663531-MH663716.

### 2.2. Phylogenetic tree reconstruction of TLRs in cetaceans

We firstly performed multiple sequence alignment of TLRs using PRANK software that makes use of phylogenetic information to distinguish alignment gaps caused by insertions or deletions and produces good alignments for evolutionary inferences (Löytynoja and Goldman, 2008). Then, the Modelgenerator program (Keane et al., 2004) was used to identify the optimal evolutionary model that best fitted our sequence datasets. Finally, maximum-Likelihood (ML) tree and Bayesian inference (BI) tree were reconstructed using RAxML 8.0.26 (Stamatakis, 2014) and MrBayes 3.2.3 (Ronquist et al., 2012) respectively, with hippopotamus and cow as outgroups. The ML tree was reconstructed with 1000 bootstrap replications. For BI tree, the Markov chain Monte Carlo analyses were run for 2,500,000 generations, sampled every 1000 generations, and the first 25% samples were burn-in.

### 2.3. Maximum likelihood (ML) methods for detecting selection

To determine whether TLRs ECD across cetacean phylogeny have undergone positive selection, we estimated the ratio ω of nonsynonymous ($d_N$) / synonymous substitution ($d_S$) using two maximum likelihood (ML) frameworks, the CODEML implemented in PAML version 4.7 (Yang, 2007) and HYPHY package via the Data Monkey Web Server (http:// www.datamonkey.org) (Delport et al., 2010). The well accepted cetacean phylogeny (Xiong et al., 2009; Chen et al., 2011; Gatesy et al., 1999) was used in all analyses.

Site models, which allow the ω ratio to vary among sites but not allow variation among lineages, were first used to identify the probabilities of sites under positive selection in each gene. A pair of site-specific models, i.e. M8a (neutral; beta distribution: $0 < \omega_0 < 1$ and $\omega_1 = 1$) and M8 (positive selection, beta distribution: $0 < \omega_0 < 1$ and $\omega_1 > 1$), implemented in the CODEML (Yang et al., 2007), was tested in the present study. The M8a-M8 comparison is a more refined test of positive selection than the M7 (neutral; beta distribution: $0 < \omega < 1$)-M8 comparison because the former can determine if the $d_N/d_S$ ratio ratio is significantly greater than one (Swanson et al., 2003). Amino acids under selection for model M8 were determined using a Bayes empirical Bayes approach (BEB) by calculating the posterior probability (Yang, 2005). The sites with a posterior probability $\geq 0.9$ were considered as candidates for selection. In addition, PSSs were also identified by another two ML methods, i.e. the fixed-effect likelihood (FEL)

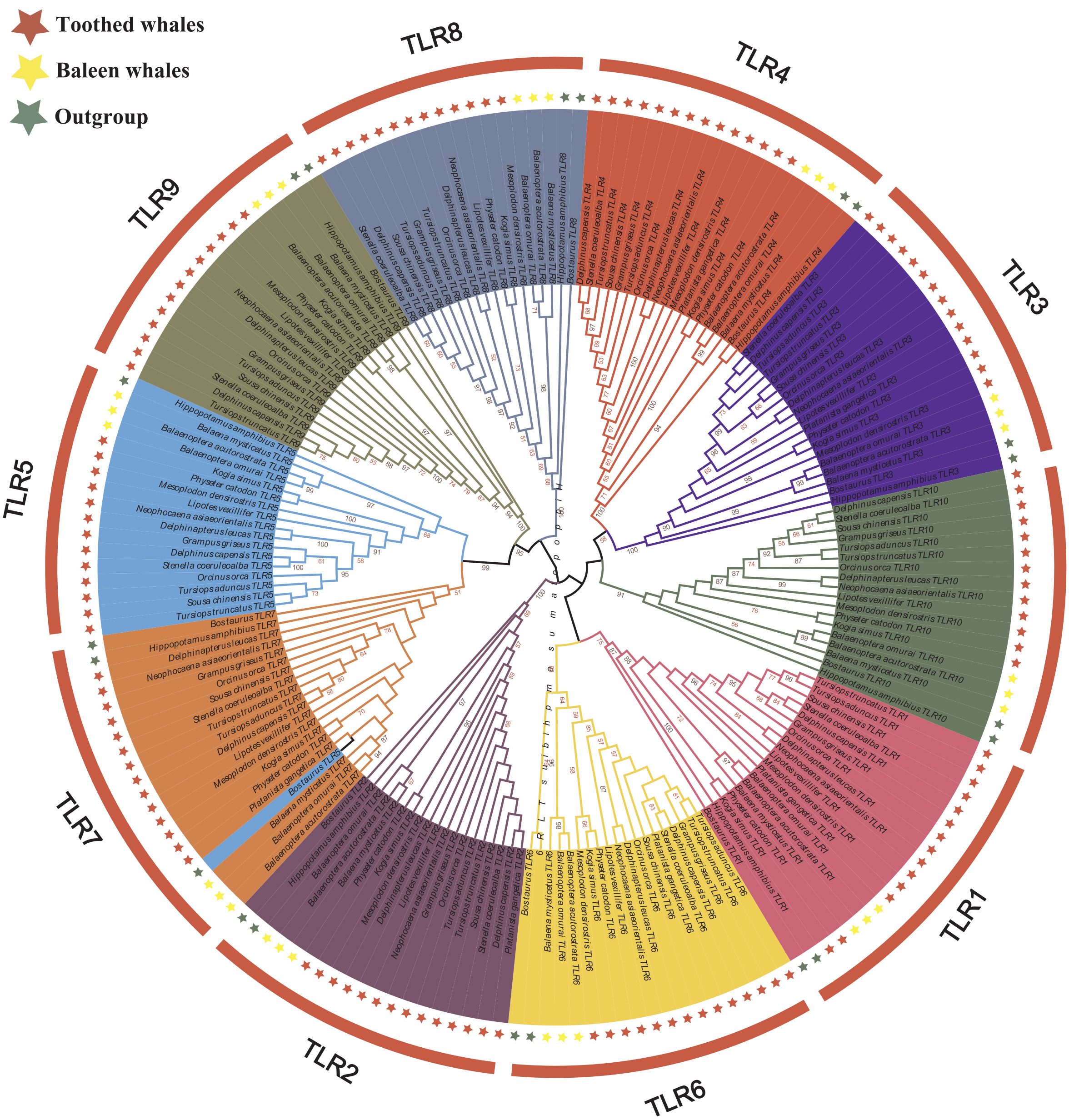


**Fig. 1.** Phylogenetic tree reconstructed using the maximum-likelihood method based on the ECD domains of 10 TLR family members. Numbers above the clades represent bootstrasp values and only higher than 50% bootstrasp values were shown.

and the random effect likelihood (REL) models, implemented in HyPhy package (Pond and Frost, 2005) available on the Datamonkey webserver (http://www.datamonkey.org). The default settings with significance levels of 0.2 for FEL and Bayes factor > 50 for REL were used. FEL and REL have the advantage that they can improve the estimation of the ω by incorporating variation in the rate of synonymous substitution (Pond and Frost, 2005).

Generally, positive selection often operates episodically on a few amino acid sites in a small subset of branches in a phylogenetic tree. Thus, the branch-site model that permits ω to vary among sites in the protein and across branches on the tree (Yang, 2007) was used to detect positive selection affecting a few sites along particular lineages. Modified branch-site model A (test 2) was performed for each TLR in each cetacean lineage. Likelihood ratio test p-values were adjusted for multiple testing with a Bonferroni's correction and threshold of 0.05.

PSSs picked out by more than one ML methods were further to test conservative or radical amino changes using a complementary protein-level approach implemented in TreeSAAP (Woolley et al., 2003) that categorizes amino acids based on 31 physicochemical properties. The BASEML program implemented in PAML 4.7 (Yang, 2007) was used in TreeSAAP to reconstruct ancestral character states at the nodes across the cetacean phylogeny. We then evaluated the effects of amino acid substitution in each branch of cetacean phylogeny. The substitutions that greatly affect a physicochemical property (i.e., in magnitude categories 6-8) can be considered radical changes. All the radical amino acid changes were further mapped onto the cetacean phylogenetic tree.

### 2.4. Three-dimensional (3D) structures of TLRs extracellular domains

To depict the underlying effects of the selection pressures on cetacean TLRs, we mapped these radical amino changes under positive selection identified by more than one ML methods on their 3D structures using PYMOL (http://pymol.sourceforge.net/). An ab-initio 3D model of each of TLR gene was predicated by the I-TASSER (http://zhanglab.ccmb.med.umich.edu/I-TASSER/, Zhang, 2008) because no significant homology modeling was found when using the SWISS-

MODEL (http://swissmodel.expasy.org, Arnold et al., 2006).

### 2.5. Analysis of association between the TLR evolutionary rate (ω) and group size/migratory behavior

To explore relationships between the evolutionary rate (ω) of each TLR family member and potential ecological factors (i.e. group size and migratory behavior), we used the phylogenetic generalized least squares (PGLS) regression, performed in R 3.1.2 using the packages Caper (Orme et al., 2012), to analyze the relationship between log-transformed (root-to-tip ω) and each log-transformed ecological factors. The "root-to-tip" ω is regarded as more suitable for regression because it is more inclusive of the evolutionary history of a locus (Wolf et al., 2009). Root-to-tip ω was firstly estimated by free-ratio model implemented in CODEML in the PAML 4.4 package (Yang, 2007). Then, we calculated the average root-to-tip ω along branches extending from the last common ancestor (LCA) of Cetacea to each extant cetacean species- in our dataset (supplementary Table S2, Supplementary Material online). The group size and migratory behavior were collected from previous published references (supplementary Table S3, Supplementary Material online). For the analysis of migratory behavior, binary variable incorporated into this study were handled as follows: no (0), yes (1).

## 3. Results

The ECD of 10 TLRs was successfully amplified in a hippopotamus and most representative species in cetaceans. TLR4 sequences were from our previous study (Shen et al., 2012). Newly obtained sequences for each TLR genes covered their full CDS range from 85.5% to 100%. The ECD of TLRs in cetaceans ranged from 1614 to 2493 bp (Table S1). Unfortunately, we failed to amplify the target loci of the three genes (i.e. TLR8, TLR9, and TLR10) in the *Platanista gangetica*. TLR5 is likely to be pseudogene in the *L. vexillifer* due to the fact that 1-bp deletion leading to premature stop codon was examined. This indel was further confirmed by re-amplifying this region in different samples in *L. vexillifer* using a different primer sets, suggesting that this deletion might be a biological reality in *L. vexillifer*. No frame-shift mutations or premature stop codons were found in the ECD of TLRs in other cetacean species suggesting that these sequences likely encode functional TLR proteins in cetaceans.

### 3.1. Phylogeny of TLRs in cetaceans

The evolutionary history of cetacean TLR family members was reconstructed using ML and BI methods. The inferred ML and BI trees recovered identical topologies, which showed that cetacean TLR family members clustered into 10 major clusters according to each member with high bootstrap or Bayesian support (Fig. 1, supplementary Fig. S1, Supplementary Material online). But there was one notable exception: *B. taurus*_TLR5 grouped with TLR 7 family in both trees (clustered with *P. gangetica*_TLR7 in the ML tree and clustered with (*Kogia simus*_TLR7 + *P. catodon*_TLR7) in the BI tree). Except the TLR7 and TLR1, the gene trees of eight TLR members were similar to the well-accepted phylogeny with some minor differences within Delphinidae and the position of *Mesoplodon densirostris*, and *P. gangetica* (Figs. 1, 2 and Fig. S1). Actually, the relationships within Delphinidae, especially within the *Sousa-Delphinus-Tursiops-Stenella* complex, as well as the placement of the river dolphin families Platanistidae (*P. gangetica*) remain controversial, despite that many studies have been conducted using a diverse array of systematic markers (Xiong et al., 2009; McGowen et al., 2009; Caballero et al., 2008). In addition, the positions of *M. densirostris* and *Delphinapterus leucas* were significant different to the well-accepted phylogeny at the TLR2, 3, 6 and TLR8. However, in the TLR1, the basal relationships of cetaceans (*K. simus_TLR1* + *P. catodon_TLR1*) disagreed with the well-accepted phylogeny that baleen whales were mixed together other toothed whales. For the TLR7, by contrast, the phylogenetic relationships of most species were significant different to the well-accepted phylogeny (Figs. 1, 2 and Fig. S1).

### 3.2. Maximum likelihood methods detection selection

Three ML methods, i.e. M8, REL and FEL, were used to test selection acting on TLR members in cetaceans. A nested models (M8a vs. M8) as implemented in PAML was compared and revealed that M8 including sites with ω > 1 fitted the data significantly better than did a neutral model (M8) at the eight TLRs except for TLR7 (P = 0.671) and TLR10 (P = 0.06) (Table 1). Importantly, M8 model identified a number of PSSs (total 79) ranging from 4 to 23 across the eight TLR loci. By comparison, a series of codons identified to be under positive selection were determined by other two ML methods, i.e. FEL (189) and REL (174) implemented in Hyphy that calculated the ω value by incorporating variation in $d_S$ whereas $d_S$ was fixed across sequences for all the PAML-based analysis. When we combined the three ML methods, a total of 148 codons were commonly yielded by at least two ML methods whereas 45 codons were commonly identified by the three methods (Table 1). In our next analysis, these 148 codons identified by at least two ML methods were considered as robust candidates of positive selection. Thus, these 148 robustly positive selection sites ranged from 5 (TLR 10) to 26 (TLR 4, TLR8) across the 10 TLR loci, with corresponding to 1.03%–4.83% of codons under positive selection. For comparison, we found different evolutionary pressure varied between TLR loci. High percentage of codons subject to positive selection was found at the TLR4 (4.83%), TLR5 (3.18%) and TLR8 (3.18%). In addition, comparable strength of selection examined at the TLR2 (2.16%), TLR 6 (2.22%) and TLR7 (2.29%). However, percentages of PSSs identified in viral-sensing TLRs were comparable with that of bacterial-sensing TLRs in the cetaceans (t-test: $P = 0.889$). We subsequently evaluated whether radical changes occurred in these 148 robust codons under positive selection by using a complementary protein-level approach implemented in TreeSAAP. The result revealed that 89.19% (132 / 148) PSSs had occurred radical changes (supplementary Table S4, Supplementary Material online). More importantly, we mapped these radical changes identified to be under positive selection onto the cetacean phylogeny and found these sites were scattered throughout almost all of the cetacean phylogeny (Fig. 2). The branch-site model was then used to test whether positive selection acting on specific sites in cetacean lineages. We found that TLR5, 7, 8, 9, and 10 showed a significant signature of positive selection along ancestral branch of cetaceans (branch c), the Last Common Ancestor (LCA) of cetaceans and *H. amphibious* (branch b), branch leading to *Sous chinensis*, *M. densirostris*, *Delphinus capensis*, *H. amphibious* and *Balaenoptera omurai*, respectively (supplementary Table S5, Supplementary Material online).

For putatively pseudogenized TLR5, selective pressure was examined by calculating the ω values across cetacean phylogeny. We first assumed that all the branches had the same ω, our analyses showed that TLR5 was evolved under purifying selection according to the comparison between the Model A and Model B ($P < 0.001$) (supplementary Table S6, Supplementary Material online). However, no statistical significance was examined between the Model C and Model D where the lineages with pseudogenes was assumed to have ω2 whereas other lineages have ω1 ($P = 0.331$) (Table S6), suggesting that no significant selective pressure was relaxed in the pseudogenized branch of *L. vexillifer*.

### 3.3. Functional significance of radical amino acid changes under positive selection

To evaluate whether putatively positive selection sites locate in functional domains, those 132 radical amino acid changes found to be under positive selection identified by more than one ML methods were firstly compared to human corresponding TLRs. We then mapped all

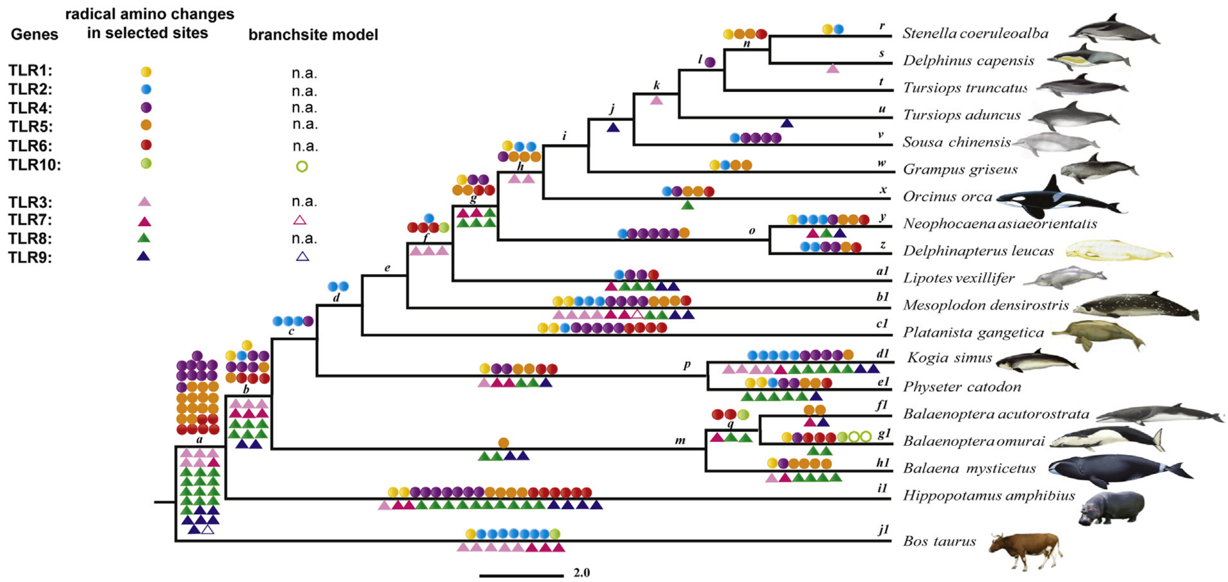


**Fig. 2.** Radical amino acid changes in positively selected sites across the cetacean phylogeny. Positively selected sites identified by site models and branch-site models in the viral-TLRs and non-viral TLRs were marked with triangles and circles, respectively.

these radical amino acid changes onto the 3D of each TLR ECD (Fig. S2). Finally, we evaluated the functional significance of cetacean radical amino acid changes according to functionally annotated sites from human TLRs. The result showed that almost all of these codons were adjacent to or within the key functional domains, such as LRRs or ligand binding regions, which are responsible for ligand interactions and recognition (supplementary and Fig. S2 and Table S7, Supplementary Material online). For example, all these radical changes that had undergone positive selection of TLR2 were located in the known LRR and Ligand binding regions in human. More than half of radical amino acid changes were found to be involved in the LRRs or ligand binding regions at the TLR1 (4/6) and TLR4 (12/24).

### *3.4. Relationship between gene evolution and ecological factors*

Association analysis showed there was no significant association between log ω and log (mean group size) at each TLR ($P > 0.05$) (supplementary Table S8, Supplementary Material online). Similarly, migratory behavior was not related to evolutionary ratios at any TLR loci (Table S8).

## 4. Discussion

### *4.1. Adaptive evolution of TLRs during cetacean evolution*

TLRs are evolutionary conserved proteins and are traditionally under strongly functional constraint. However, recent studies have found evidence of positive selection in some vertebrate groups, such as some PSSs identified at TLR3, 4, 5 and 15 in birds (Alcaide and Edwards, 2011), TLR 4, 7 in wild rodents (Fornůsková et al., 2013), and TLR1, 2, 6, 8 in family Suidae (Darfour-Oduro et al., 2015). Particularly, balancing selection was identified at TLR1, 6, and 10 in European populations of human (Ferrer-Admetlla et al., 2008). In the present study, we comprehensively investigated evolutionary patterns of all the 10 TLR family in cetacean species with the goal of providing a general picture of the evolution of cetacean innate immunity. Our results supposed that pathogen mediated positive selection would have occurred at cetacean TLRs with its habitat transition from the land to the completely aquatic environment. Three lines of evidence provided support for this hypothesis. First, a total of 148 robust candidates of sites were identified to be under positive selection at ECD of all TLR loci. Second, the overwhelming majority (89.19%) of these PSSs were found to have radical amino acid changes, which is yet another evidence of positive selection. Finally, radical amino changes under positive selection were adjacent to or within the functional domains of each TLR. All the results suggested widespread positive selection acting on TLR genes in cetacean lineages, which was in line with the result reported in cetacean adaptive immunity, such as balancing selection at the MHC genes. Previous studies revealed that MHC polymorphism in cetaceans were comparable with their terrestrial relatives (Xu et al., 2009, 2010). From the results of TLRs and MHC, we could suggest cetacean might have faced diversity of pathogens during its transitions from terrestrial to aquatic environment. This suggestion was contradiction with traditional view that relatively less pathogenic microorganisms in the aquatic environment compared with the terrestrial ecosystem.

We reconstructed the ancestral node of each TLR ECD across lineages of cetaceans. When we mapped these PSSs on the phylogenetic tree and found these sites were dispersed across almost all of the lineages of cetacean phylogeny from the origin to the subsequent radiation. Specially, lineage-specific evolution was determined across cetacean phylogeny, reflecting rapid host adaptation to varying microbial pressures. For example, the highest concentration of PSSs (55 codons) was detected along the LCA of cetacean + hippopotamus (Fig. 2, branch a), which represented the habitat transition of the terrestrial ancestors of cetaceans from land to semi-aquatic habitat (Gatesy and O'Leary, 2001). Additionally, more significant positive selection (27 condons) was identified along the lineage leading to LCA of cetaceans (Fig. 2, branch b) with habit transition from semi-aquatic to full-aquatic habitat (Gatesy and O'Leary, 2001). The accumulation of PSSs at specific transitions examined in our study might reflect a faster adaptation to new pathogens as cetaceans encountered novel habitat. It was worth to note that the accumulation of PSSs were different in lineages of extant cetaceans. For instance, high accumulations of radical amino acid changes were identified along lineage leading to *M. densirostris* (24 codons), *K. simus* (22 codons), *P. catodon* (15 codons), however, few radical changes in selected sites were found at the lineages of delphinids. The colonization of aquatic habitats by cetaceans implied adaptation not only to a benthic or bathypelagic environment (e.g., *P. catodon*) but also exclusively to offshore (e.g., *T. truncatus*). More importantly, microbial abundances were reported to have different patterns between coastal and pelagic water, sea surface and deeper water, etc. (Fuhrman, 1999). Thus, ecological differences

**Table 1**
Sites under positive selection identified by different ML methods.

| Gene (length) | Lnl M8a | Lnl M8 | −2ln△L | PAML M8[a] (site model) | PAML ma[b] (branch-site model) | FEL[c] | REL[d] | Total no. of sites[e] | % of sites |
|---|---|---|---|---|---|---|---|---|---|
| **TLR1** (1746bp) | −4679.892 | 4670.434 | 18.916 (p < 0.001) | **4**, ***65***, **218**, **290**, ***544***, **578** | – | ***65***, 94, 118, 123, 139, 140, 161, 241, 256, 315, 518, 543, ***544***, 548, 582 | **4**, ***65***, **218**, **290**, ***544***, **578** | 6(2) | 0.01 |
| **TLR2** (1665bp) | −4899.949 | −4886.506 | 26.886 (p < 0.001) | **236**, 265, ***274***, 522, ***532***, ***538***, ***545*** | – | 42, **84**, **140**, 172, **248**, ***274***, 275, **294**, **447**, 520, **521**, 525, **530**, ***532***, ***538***, ***545***, 551 | **84**, **140**, **236**, **248**, ***274***, **294**, **447**, **521**, **530**, ***532***, ***538***, ***545*** | 12(4) | 2.16 |
| **TLR3** (2097bp) | −5496.774 | −5490.570 | 12.408 (p < 0.001) | ***13***, **279**, 384, **566**, 587 | – | **4**, **8**, ***13***, 80, 87, 221, 241, 262, **270**, 286, 325, 339, **357**, **367**, **412**, **435**, 510, **556**, 566, 582, **640**, **676** | **4**, **8**, ***13***, **270**, **279**, **357**, **367**, **412**, **435**, **566**, **640**, **676** | 12(1) | 1.72 |
| **TLR4** (1614bp) | −4610.722 | −4591.049 | 19.673 (p < 0.001) | ***177***, ***179***, ***204***, ***207***, **208**, ***228***, **239**, **250**, ***272***, ***304***, ***324***, **489** | – | **28**, **45**, 84, 150, ***177***, ***179***, ***204***, ***207***, **211**, **221**, ***228***, **247**, ***272***, **280**, 281, **302**, ***304***, **308**, ***324***, **342**, **346**, 404, **408**, 409, **419**, **430**, 447, 456, 482, 525 | 28, 45, 150, ***177***, ***179***, ***204***, ***207***, **208**, 211, 221, ***228***, **239**, 247, **250**, ***272***, 280, 302, ***304***, 308, ***324***, 342, 346, 408, 419, 430, **489** | 26(8) | 4.83 |
| **TLR5** (1890bp) | −5478.917 | −5464.968 | 27.898 (p < 0.001) | ***4***, ***32***, ***117***, ***207***, **239**, ***262***, **274**, **384**, ***429***, ***473***, ***481***, **496**, **511**, ***592***, ***619*** | 94, 119, 268 | ***4***, ***32***, 34, ***117***, **128**, 167, 184, ***207***, **210**, **214**, 242, ***262***, **400**, ***429***, ***473***, ***481***, 540, 589, ***592***, 595, **602**, ***619***,623 | **4**, 14, ***32***, ***117***, **128**, ***207***, **210**, **214**, **239**, ***262***, 272, **274**, **384**, **400**, ***429***, ***473***, 477, ***481***, **496**, **511**, ***592***, **602**, 616, ***619***, 628 | 20(10) | 3.18 |
| **TLR6** (1755bp) | −4400.468 | −4385.040 | 30.856 (p < 0.001) | ***45***, ***90***, ***111***, **335**, ***378***, **577** | – | **12**, 18, ***45***, **73**, ***90***, ***111***, **136**, 215, 222, 240, **251**, **294**, **313**, 365, ***378***, 392,**421** | 12, 41, ***45***, 73, ***90***, ***111***, 136, 251, 294, 313, **335**, 375, ***378***, **421**, **577** | 13(4) | 2.22 |
| **TLR7** (2493bp) | −5864.086 | −5863.996 | 0.18 (p = 0.671) | – | **314** | 36, **82**, 197, **218**, **222**, **310**, **454**, **458**, **475**, **493**, **512**, **525**, **563**, **642**, **648**, **696**, **713**, **725**, **734**, **773** | **82**, 115, 172, **218**, **222**, 276, 280, **310**, **314**, 340, 382, 414, 436, **454**, **458**, **475**, **493**, **512**, **525**, **563**, **642**, 645, **648**, 664, 683, 685, **696**, **713**, **725**, **734**, 748, **773** | 19(0) | 2.29 |
| **TLR8** (2451bp) | −6718.944 | −6716.131 | 5.626 (p = 0.018) | ***38***, 52, ***61***, ***146***, ***156***, **183**, ***186***, ***209***, ***210***, **214**, **228**, ***230***, **263**, **266**, **306**, **380**, ***433***, **436**, **467**, ***476***, **479**, ***558***, ***598*** | 218, 311 | ***38***, ***61***, 95, ***146***, ***156***, ***186***, ***209***, ***210***, ***230***, ***433***, 440, **446**, 465, ***476***, 525, ***558***, ***598***, **642**, **658**, 701, **745**, 778, 798 | ***38***, ***61***, 98, 136, ***146***, ***156***, **183**, ***186***, 187, ***209***, ***210***, **214**, **228**, ***230***, 234, **263**, **266**, **306**, 345, **380**, ***433***, **436**, **446**, **467**, ***476***, 479, ***558***, ***598***, **642**, 650, **658**, 659, 680, 693, **745** | 26(12) | 3.18 |
| **TLR9** (2421bp) | −6469.113 | −6463.251 | 11.724 (p < 0.001) | ***4***, ***6***, **329**, **359** | 38, 115, ***300***, 335,442,645,723, **726**, 775 | ***4***, ***6***, **186**, **296**, ***300***, **329**, **359**, **393**, 444, 466, **726** | **4**, **6**, 186, 296, ***300***, 393 | 9(3) | 1.12 |
| **TLR10** (1725bp) | −4032.733 | −4030.967 | 3.532 (p = 0.060) | – | 130, 244, **245**, ***392*** | **71**, 116, **132**, 134, 146, 235, **245**, **255**, 319, ***392***, 424 | **71**, 99, **132**, **255**, ***392*** | 5(1) | 0.87 |

Note: The codons picked out by at least two ML methods were bold, and the codons identified simultaneously by three methods were bold and Italic.
[a] Codons with posterior probabilities of 90% in the BEB analyses.
[b] Codons with posterior probabilities of 70% in the BEB analyses.
[c] Condons with $P < 0.2$.
[d] Codons with Bayes factors > 50.
[e] No. of sites indicate positively selected sites identified by both ML methods and the codons identified simultaneously by three methods in the parentheses.

over long time scales were suggested to be sufficient to drive divergent selection in cetaceans. However, no significant association was examined between the each TLR evolution and some easy-to-obtain ecological factors, such as mean group size and migratory behavior. This result suggested that TLR evolution was a complex process. For this reason, more ecological factors and life history of cetaceans should be collected to further investigation this issue.

### 4.2. Viral TLRs as promising immune receptors in marine system

Previous investigation of primates, bird and rodents TLRs evolution indicated that viral TLRs seemed to evolve under stronger functional constraint than non-viral TLRs (Wlasiuk and Nachman, 2010; Alcaide and Edwards, 2011; Barreiro et al., 2009; Fornůsková et al., 2013; Darfour-Oduro et al., 2015). This result might be contributed to the dual roles of viral TLRs which could recognize both viral nucleic acids and target self components. Thus, viral TLRs must be under strong purifying selection in order to avoid autoimmunity. In addition, viral TLRs under functional constraint were suggested to be due to primates and bird as homogeneous groups probably being faced with the restricted number of viruses (Areal et al., 2011). In our study, however, proportion of sites under selection in viral-sensing TLRs was not significantly different with that of bacterial-sensing TLRs across the cetacean phylogeny (t-test: $P = 0.889$). This indicated that comparable level of positive selection was identified in viral TLRs and non-viral TLR during cetacean evolution. The distinct ecology of marine system might explain the similar level of positive selection at viral and non-viral TLRs in cetaceans. It has been known that viruses are the most abundant pathogens in the oceans, exceeds richness of bacteria and archaea by about 15-fold (Suttle, 2007; Bergh et al., 1989). Thus, viral infections are a major source of mortality, and cause disease in a range of organisms in the ocean (Suttle, 2007). For example, viral pathogens such as influenza A virus and morbillivirus infection have caused heavy losses of marine mammals (Nielsen et al., 2001; Visser et al., 1993). Moreover, marine pathogens have relatively rapid spread than that of terrestrial environment, which accelerated dispersal of pathogens and posed a formidable challenge for mammals in the ocean (McCallum et al., 2003). Therefore, cetacean viral TLRs ECD under strong positive selection was expected to improve their functional integrity in order to protect against infection by such viruses in aquatic environment.

Intriguingly, higher accumulation of PSSs was inferred in the case of viral TLR genes presumably enhanced the biological function to recognize PAMPs from diverse viral sources. For instance, strongest evidence of positive selection was identified at TLR8 in cetaceans (26 positions, 3.18%) with accumulation of 26 radical amino acid changes along 16 lineages, followed by the TLR7 (26 codons, 2.29%) with accumulation of 12 radical amino acid changes along 13 lineages (Table S4). It was noted that TLR7 (9 codons) and TLR8 (20 codons) with a large number of PSSs in the LRR domain (Table S7). Moreover, 20 codons of TLR8 were located in irregular motif, which is more structural flexibility (Matsushima et al., 2007). TLR7 and TLR8, recognizing viral ssRNA, are close phylogenetic relatives and share similar transmembrane regions. In line with this, our finding suggested that TLR7 and TLR8 must have adaptively modified to recognize and bind potential and abundant viral pathogens in aquatic environment.

### 4.3. Differences in evolution of non-viral TLRs

We found divergent signal of positive selection within non-viral TLRs from the consideration of the proportion of sites under selection. TLR4 (4.83%) and TLR5 (3.18%) displayed a higher proportion of PSSs whereas TLR1 (0.01%) and TLR 10 (0.87%) were the receptors with the lower percentage of codons under positive selection (Table 1). This suggested that cetacean TLR genes might be subject to different selective regimes. Although TLRs are evolutionarily highly conserved, diversified TLRs recognize a variety of microbial ligands. For instance, TLR4 and TLR5 detect lipopolysaccharide (LPS) and bacterial flagellin, respectively (Poltorak et al., 1998). Therefore, the specific recognizing of microbial ligands might due to different evolutionary pressures between different TLRs.

The strongest evidence for positive selection was examined at TLR4 in cetaceans. The highest proportion of PSSs (4.83%) was examined at the TLR4 in cetaceans with 26 PPSs identified by at least two ML methods. Of these sites, twenty-four sites (92.3%) were categorized as radical changes at the protein-level. Strong evidence of positive selection examined at TLR4 was concordant with previous observation in primates and rodents (Wlasiuk and Nachman, 2010; Fornůsková et al., 2013). TLR4 is expressed on the cell membrane and response to various pathogens, such as recognizing LPS from Gram-negative bacteria and targeting components of yeast, *Trypanosoma*, and even viruses (Kumar et al., 2009). Crystallographic structures of mouse TLR4 have shown that TLR4 and MD-2 form a heterodimer to facilitate the recognition of LPS (Kim et al., 2007). Interestingly, nine PSSs (136, 330, 338, 341, 363, 371, 393, 395, 399) under radical amino acid changes identified in cetaceans were in or adjoined to the regions that participated in the interaction between TLR4, MD-2 and LPS (Table S7). Nevertheless, these PSSs located in the functional regions of TLR4 remains to be assessed.

Strong evidence of positive selection was also identified at TLR2 in cetaceans with accumulation 11 radical amino acid changes along 18 lineages (Table S4). It has shown that TLR2 is widely expressed across species and recognizes the greatest number of PAMPs through its unique ability to form heterodimers with TLR1 or TLR6 (Smith et al., 2014). More importantly, four radical amino acid changes evolved under positive selection sites (267, 279, 305, and 325) were adjacent to or within ligand binding regions, which may be as a result of evolutionary arm race between host and pathogens. This was inconsistent with several recent studies conducted on other mammals such as humans and members of Suidae that revealed TLR2 evolved under selective constraint (Darfour-Oduro et al., 2015). However, no sign of positive selection was examined along lineages of mysticetes. We could not exclude the possibility of small samples (3) used in our study, therefore, more samples are needed to add in further work to insight into the evolution of the TLRs.

TLR5 had the second-highest signature of positive selection with 3.18% codons identified under selection. TLR5 is activated by monomeric flagellin, the main component of the bacterial flagella and a potent virulence factor (Gay and Gangloff, 2007). In this case, site 400 fell directly in the flagellin-binding region of extracellular domain (residues 386–407) (Mizel et al., 2003), and consequently, the physiochemical property changes of this site (Table S4) might remodel this protein and contribute to adequate immune response in cetaceans to infection with flagellated bacteria. Moreover, site 262 lay adjacent to a residue 267 that was responsible for specific recognition of flagellin molecules between human and mouse TLR5 (Andersen-Nissen et al., 2007). However, the frame-shift mutation was exclusively identified in L. *vexillifer*, suggesting the TLR5 might have lost its function. Off course, further functional tests are needed to clarify this suggestion.

**Funding**

This work was supported by the National Natural Science Foundation of China (NSFC, grant numbers 3157039 and 31772448 to SX), the National Key Programme of Research and Development, Ministry of Science and Technology (Grant number 2016YFC0503200 to GY and SX), the State Key Program of National Natural Science of China to GY (grant number 31630071), the Priority Academic Program Development of Jiangsu Higher Education Institutions to GY and SX.

**Acknowledgements**

We thank Mr. Xinrong Xu for help with collecting samples for many years. A special thank-you is due to Miss Xiaohui Sun, Xinghua Rong, Xin Huang for their technical supports.

## Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.molimm.2018.12.022.

## References


Alcaide, M., Edwards, S.V., 2011. Molecular evolution of the Toll-Like receptor multigene family in birds. Mol. Biol. Evol. 28 (5), 1703–1715.
Andersen-Nissen, E., Smith, K.D., Bonneau, R., Strong, R.K., Aderem, A., 2007. A conserved surface on Toll-like receptor 5 recognizes bacterial flagellin. Exp. Med. 204 (2), 393–403.
Areal, H., Abrantes, J., Esteves, P., 2011. Signatures of positive selection in Toll-like receptor (TLR) genes in mammals. BMC Evol. Biol. 11 (1), 368.
Arnold, K., Bordoli, L., Kopp, J., Schwede, T., 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22 (2), 195–201.
Baker, C.S., Vant, M.D., Dalebout, M.L., Lento, G.M., O'Brien, S.J., Yuhki, N., 2006. Diversity and duplication of DQB and DRB-like genes of the MHC in baleen whales (suborder: Mysticeti). Immunogenetics 58 (4), 283–296.
Barreiro, L., Quintana-Murci, L., 2010. From evolutionary genetics to human immunology: how selection shapes host defence genes. Nat. Rev. Genet. 11 (1), 17–30.
Barreiro, L.B., Ben-Ali, M., Quach, H., Laval, G., Patin, E., Pickrell, J.K., Bouchier, C., Tichit, M., Neyrolles, O., Gicquel, B., 2009. Evolutionary dynamics of human Toll-like receptors and their different contributions to host defense. PLoS Genet. 5 (7), e1000562.
Barton, G.M., 2007. Viral recognition by Toll-like receptors. Semin. Immunol. 19 (1), 33–40.
Bergh, O., Børsheim, K.Y., Bratbak, G., Heldal, M., 1989. High abundance of viruses found in aquatic environments. Nature 340 (6233), 467–468.
Caballero, S., Jackson, J., Mignucci-Giannoni, A.A., Barrios-Garrido, H., Beltrán-Pedreros, S., Montiel-Villalobos, M.G., Robertson, K.M., Baker, C.S., 2008. Molecular systematics of South American dolphins Sotalia: sister taxa determination and phylogenetic relationships, with insights into a multi-locus phylogeny of the Delphinidae. Mol. Phylogenet. Evol. 46 (1), 252–268.
Carty, M., Bowie, A.G., 2010. Recent insights into the role of Toll-like receptors in viral infection. Clin. Exp. Immunol. 161 (3), 397–406.
Chen, Z.S., Zhou, K., Yang, G., 2011. Whale phylogeny and rapid radiation events revealed using novel retroposed elements and their flanking sequences. BMC Evol. Biol. 11 (1), 314.
Cook, D.N., Pisetsky, D.S., Schwartz, D.A., 2004. Toll-like receptors in the pathogenesis of human disease. Nat. Immunol. 5 (10), 975.
Darfour-Oduro, K.A., Megens, H.J., Roca, A.L., Groenen, M.A.M., Schook, L.B., 2015. Adaptive evolution of toll-like receptors (TLRs) in the family Suidae. PLoS One 10 (4), e0124069.
Delport, W., Poon, A.F.Y., Frost, Frost S.D.W., Kosakovsky Pond, S.L., 2010. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. Bioinformatics 26 (19), 2455–2457.
Ferrer-Admetlla, A., Bosch, E., Sikora, M., Marquès-Bonet, T., Ramírez-Soriano, A., Muntasell, A., 2008. Balancing selection is the main force shaping the evolution of innate immunity genes. J. Immunol. 181 (2), 1315–1322.
Fornůsková, A., Vinkler, M., Pagès, M., Galan, M., Jousselin, E., Cerqueira, F., Cosson, J.F., 2013. Contrasted evolutionary histories of two Toll-like receptors (Tlr4 and Tlr7) in wild rodents (MURINAE). BMC Evol. Biol. 13 (1), 194.
Fuhrman, J.A., 1999. Marine viruses and their biogeochemical and ecological effects. Nature 399 (6736), 541.
Gatesy, J., O'Leary, M.A., 2001. Deciphering whale origins with molecules and fossils. Trends Ecol. Evol. 16 (10), 562–570.
Gatesy, J., O'Grady, P., Baker, R.H., 1999. Corroboration among data sets in simultaneous analysis: hidden support for phylogenetic relationships among higher level artiodactyl taxa. Cladistics 15 (3), 271–313.
Gay, N.J., Gangloff, M., 2007. Structure and function of Toll receptors and their ligands. Annu. Rev. Biochem. 76, 141–165.
Grueber, C.E., Wallis, G.P., Jamieson, I.G., 2014. Episodic positive selection in the evolution of avian Toll-like receptor innate immunity genes. PLoS One 9 (3), e89632.
Hedrick, S.M., 2004. The acquired immune system: a vantage from beneath. Immunity 21 (5), 607–615.
Iwasaki, A., Medzhitov, R., 2010. Regulation of adaptive immunity by the innate immune system. Science 327 (5963), 291–295.
Keane, T.M., Naughton, T.J., McInerney, J.O., 2004. ModelGenerator: Amino Acid and Nucleotide Substitution Model Selection. Natl Univ of Ireland, Maynooth, Ireland.
Kim, H.M., Park, B.S., Kim, J.I., Kim, S.E., Lee, J., Oh, S.C., Enkhbayar, P., Matsushima, N., Lee, H., Ook, J.Y., Lee, J.O., 2007. Crystal structure of the TLR4-MD-2 complex with bound endotoxin antagonist Eritoran. Cell 130 (5), 906–917.
Kumar, A., Yu, F.S.X., 2006. Toll-like receptors and corneal innate immunity. Curr. Mol. Med. 6 (3), 327–337.
Kumar, H., Kawai, T., Akira, S., 2009. Pathogen recognition in the innate immune response. Biochem. J. 420 (1), 1–16.
Letunic, I., Doerks, T., Bork, P., 2015. SMART: recent updates, new developments and status in 2015. Nucleic Acids Res. 43 (Database issue), 257–260.
Löytynoja, A., Goldman, N., 2008. Phylogeny-aware gap placement prevents errors in sequence alignment and evolutionary analysis. Science 320.
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 (1), 124.
McCallum, H., Harvell, D., Dobson, A., 2003. Rates of spread of marine pathogens. Ecol. Lett. 6 (12), 1062–1067.
McGowen, M.R., Spaulding, M., Gatesy, J., 2009. Divergence date estimation and a comprehensive molecular tree of extant cetaceans. Mol. Phylogenet. Evol. 53 (3), 891–906.
Mikami, T., Miyashita, H., Takatsuka, S., Kuroki, Y., Matsushima, N., 2012. Molecular evolution of vertebrate Toll-like receptors: evolutionary rate difference between their leucine-rich repeats and their TIR domains. Gene 503 (2), 235–243.
Mizel, S.B., West, A.P., Hantgan, R.R., 2003. Identification of a sequence in human toll-like receptor 5 required for the binding of Gram-negative flagellin. J. Biol. Chem. 278 (26), 23624–23629.
Nielsen, O., Clavijo, A., Boughen, J.A., 2001. Serologic evidence of influenza A infection in marine mammals of arctic Canada. J. Wildlife Dis. 37 (4), 820–825.
Orme, C.D.L., Freckleton, R.P., Thomas, G.H., Petzoldt, T., Fritz, S.A., Isaac, N., Pearse, W., 2012. Caper: comparative analyses of phylogenetics and evolution in R. Methods. Ecol Evol. 3, 145–151.
Piertney, S.B., Oliver, M.K., 2006. The evolutionary ecology of the major histocompatibility complex. Heredity 96 (1), 7.
Poltorak, A., He, X., Smirnova, I., Liu, M.Y., Van Huffel, C., Du, X., Freudenberg, M., 1998. Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. Science 282 (5396), 2085–2088.
Pond, S.L.K., Frost, S.D.W., 2005. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. Bioinformatics 21 (10), 2531–2533.
Roach, J.C., Glusman, G., Rowen, L., Kaur, A., Purcell, M.K., Smith, K.D., Hood, L.E., Aderem, A., 2005. The evolution of vertebrate Toll-like receptors. Proc. Natl. Acad. Sci. U. S. A. 102 (27), 9577–9582.
Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D.L., Darling, A., Höhna, S., et al., 2012. rBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61 (3), 539–542.
Shen, T., Xu, S., Wang, X., Yu, W., Zhou, K., Yang, G., 2012. Adaptive evolution and functional constraint at TLR4 during the secondary aquatic adaptation and diversification of cetaceans. BMC Evol. Biol. 12 (1), 39.
Shishido, R., Ohishi, K., Suzuki, R., Takishita, K., Ohtsu, D., Okutsu, K., Murayama, T., 2010. Cetacean Toll-like receptor 4 and myeloid differentiation factor 2, and possible cetacean-specific responses against Gram-negative bacteria. Comp. Immunol. Microb. 33 (6), e89–e98.
Smith, S.A., Haig, D., Emes, R.D., 2014. Novel ovine polymorphisms and adaptive evolution in mammalian TLR2 suggest existence of multiple pathogen binding regions. Gene 540 (2), 217–225.
Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30 (9), 1312–1313.
Suttle, C.A., 2007. Marine viruses–major players in the global ecosystem. Nat. Rev. Microbiol. 5 (10), 801.
Swanson, W.J., Nielsen, R., Yang, Q., 2003. Pervasive adaptive evolution in mammalian fertilization proteins. Mol. Biol. Evol. 20 (1), 18–20.
Thewissen, J.G., Cooper, L.N., Clementz, M.T., Bajpai, S., Tiwari, B.N., 2007. Whales originated from aquatic artiodactyls in the Eocene epoch of India. Nature 450 (7173), 1190.
Visser, I.K.G., Van Bressem, M.F., Barrett, T., Osterhaus, A.D.M.E., 1993. Morbillivirus infections in aquatic mammals. Vet. Res. 24 (2), 169.
Werling, D., Jann, O.C., Offord, V., Glass, E.J., Coffey, T.J., 2009. Variation matters: TLR structure and species-specific pathogen recognition. Trends Immunol. 30 (3), 0–130.
Wlasiuk, G., Nachman, M.W., 2010. Adaptation and constraint at Toll-like receptors in primates. Mol. Biol. Evol. 27 (9), 2172–2186.
Wolf, J.B., Künstner, A., Nam, K., Jakobsson, M., Ellegren, H., 2009. Nonlinear dynamics of nonsynonymous ($d_N$) and synonymous ($d_S$) substitution rates affects inference of selection. Genome Biol. Evol. 1, 308–319.
Woolley, S., Johnson, J., Smith, M.J., Crandall, K.A., McClellan, D.A., 2003. TreeSAAP: selection on amino acid properties using phylogenetic trees. Bioinformatics 19 (5), 671–672.
Xiong, Y., Brandley, M.C., Xu, S.M., Zhou, K.Y., Yang, G., 2009. Seven new dolphin mitochondrial genomes and a time-calibrated phylogeny of whales. BMC Evol. Biol. 9 (1) 20-0.
Xu, S.X., Sun, P., Zhou, K.Y., Yang, G., 2007. Sequence variability at three MHC loci of finless porpoises (*Neophocaena phocaenoides*). Immunogenetics 59 (7), 581–592.
Xu, S.X., Chen, B.Y., Zhou, K.Y., Yang, G., 2008. High similarity at three MHC loci between the baiji and finless porpoise: trans-species or convergent evolution? Mol. Phylogenet. Evol. 47 (1), 36–44.
Xu, S.X., Ren, W.H., Li, S.Z., Wei, F.W., Zhou, K.Y., Yang, G., 2009. Sequence polymorphism and evolution of three cetacean MHC genes. J. Mol. Evol. 69 (3), 260–275.
Xu, S.X., Ren, W.H., Zhou, X.M., Zhou, K.Y., Yang, G., 2010. Sequence polymorphism and geographical variation at a positively selected MHC-DRB gene in the finless porpoise (*Neophocaena phocaenoides*): implication for recent differentiation of the Yangtze Finless porpoise? J. Mol. Evol. 71 (1), 6–22.
Xu, S.X., Ju, J.F., Zhou, X.M., Wang, L., Zhou, K.Y., Yang, G., 2012. Considerable MHC diversity suggests that the functional extinction of baiji is not related to population genetic collapse. PLoS One 7 (1), e30423.
Yang, Z., 2005. Bayes empirical Bayes inference of amino acid sites under positive selection. Mol. Biol. Evol. 22 (4), 1107–1118.
Yang, Z.H., 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24 (8), 1586–1591.
Zhang, Y., 2008. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics 9 (1) 40-0.