

# 鲈形目 *TLR1* 和 *TLR9* 基因的选择压力分析

刘 兴<sup>1</sup>, 高 波<sup>2</sup>, 吉红九<sup>2</sup>, 贾超峰<sup>2</sup>, 祝 斐<sup>2</sup>, 孟 乾<sup>2</sup>,  
陈淑吟<sup>2</sup>, 张志勇<sup>2</sup>, 徐士霞<sup>1</sup>

(1. 南京师范大学生命科学学院, 江苏省生物多样性重点实验室, 江苏 南京 210023)

(2. 江苏省海洋水产研究所, 江苏省海水鱼类遗传育种重点实验室, 江苏 南通 226007)

**[摘要]** 鲈形目(Perciformes)是鱼类中物种丰富度最高的一个目,分布在多个水域,在渔业经济发展中占据着重要地位。然而,高密度的养殖模式加剧了病原微生物引起的鱼类病害。Toll 样受体(Toll-like receptors, TLRs)是先天免疫中的重要蛋白质分子,在病毒以及细菌防御中起重要作用。本文以 *TLR1* 和 *TLR9* 为候选基因,共下载了鲈形目 17 个代表物种的同源序列,通过进化分析探讨该目物种先天性免疫的分子进化基础。选择压力分析发现 *TLR1*、*TLR9* 基因中分别有 16 和 8 个正选择位点被两种方法同时检测到,且 86% 的位点位于蛋白质富含亮氨酸重复序列(Leucine rich repeat, LRR)结构域。上述结果提示鲈形目物种面临较高的病原微生物感染的挑战,而这些正选择位点可能导致基因进化出重要的功能改变以抵抗病原体的侵袭。另外,Free-ratio 以及 Branch-site 模型检测发现基因的正选择具有物种特异性,如 *TLR1* 在鲷科的最近共同祖先支以及鲈科最近共同祖先支等受到显著正选择, *TLR9* 也在鲷科的最近共同祖先支受到显著正选择,且这些谱系被报道经常受到病原体的威胁,因此,鲈形目多个谱系受到正选择可能是在长期抵抗病原微生物侵袭中进化的结果。以上研究结果表明,在长期抵御病原微生物过程中,鲈形目物种的 *TLR1*、*TLR9* 基因可能发生了适应性功能改变。

**[关键词]** 鲈形目, *TLR1*, *TLR9*, 适应性进化, 先天性免疫

**[中图分类号]** Q953 **[文献标志码]** A **[文章编号]** 1001-4616(2022)03-0068-11

## Analysis of Selection Pressure on *TLR1* and *TLR9* Genes in Perciformes

Liu Xing<sup>1</sup>, Gao Bo<sup>2</sup>, Ji Hongjiu<sup>2</sup>, Jia Chaofeng<sup>2</sup>, Zhu Fei<sup>2</sup>, Meng Qian<sup>2</sup>,  
Chen Shuyin<sup>2</sup>, Zhang Zhi yong<sup>2</sup>, Xu Shixia<sup>1</sup>

(1. School of Life Sciences, Nanjing Normal University, Jiangsu Provincial Key Laboratory of Biodiversity, Nanjing 210023, China)

(2. Marine Fisheries Research Institute of Jiangsu Province, Jiangsu Key Laboratory for Genetics and Breeding of Marine Fishes, Nantong 226007, China)

**Abstract:** The Perciformes is an order with the highest species richness in fish, distributed in many waters, and plays an essential role in developing the fishery economy. However, the high-density culture mode aggravates the fish diseases caused by pathogenic microorganisms. Toll-like receptors (TLRs) are essential proteins in innate immunity and play an important role in the defense against bacterial or viral infections. In this paper, *TLR1* and *TLR9* were used as candidate genes, and the homologous sequences of 17 representative species of Perciformes were downloaded. The molecular evolutionary basis of innate immunity of this species was discussed by evolutionary analysis. Selective pressure analysis showed that the three methods simultaneously detected 16 and 8 positive selection sites in *TLR1* and *TLR9*, respectively. 86% of sites were located in the Leucine rich repeat (LRR) domain. These results suggest that Perciformes species face the challenge of high pathogenic microbial infection. These positive selection sites may lead to critical functional changes in gene evolution to resist pathogen invasion. In addition, free-ratio and branch-site models found that positive selection of genes was species-specific. For example, *TLR1* is significantly positively selected in the last common ancestor (LCA) branch of Sparidae and Percidae, and *TLR9* is also significantly positively selected in the last common ancestor branch of Sparidae. These lineages are often reported to be threatened by pathogens. Therefore, the positive selection of multiple Perciformes lineages may result from evolution in long-term resistance to pathogenic microorganisms. The above results indicated that

收稿日期: 2022-01-27.

基金项目: 江苏省农业重大新品种创制项目(PZCZ201744)、江苏省种业振兴“揭榜挂帅”项目(JBCS[2021]034)。

通讯作者: 陈淑吟, 研究员, 研究方向: 鱼类分子遗传与育种. E-mail: shuyinchen89@163.com

*TLR1* and *TLR9* of Perciformes species may have changed their adaptive functions during the long-term resistance to pathogenic microorganisms.

**Key words:** Perciformes, *TLR1*, *TLR9*, adaptive evolution, innate immunity

鲈形目(Perciformes)是鱼类中物种丰富度最高的一个目,包括 25 个亚目、160 科、1 539 属、10 033 种<sup>[1]</sup>. 鲈形目物种分布广泛,大多生活在海洋,仅有少数(如鲈科、丽鲷科等)生活在淡水水域<sup>[1]</sup>. 鲈形目多个物种是中国重要的海水养殖鱼类,具有重要食用价值和经济价值. 据报道高密度养殖模式会引起鲈形目物种病害频繁发生,尤其是受到多种细菌、病毒等病原微生物威胁<sup>[2]</sup>. 鱼类病害的爆发会导致其产量和品质的下降,从而造成较为严重的经济损失. 因此,研究鲈形目免疫基因的进化不仅有助于理解宿主和病原体间的作用机制,也为物种遗传育种及疾病防治提供重要的参考依据.

先天性免疫在抵抗外界病原体入侵方面起着重要作用,是动物的主要防御机制,代表着抵御病原体入侵的第一道防线,在维持生物体生长、发育和生存过程中发挥最重要作用<sup>[3-4]</sup>. 鱼类属于变温动物,其淋巴细胞增殖缓慢导致适应性免疫相对滞后且免疫应答能效较低,因此鱼类先天性免疫在抵御病原体感染的过程中扮演更为重要的角色<sup>[5]</sup>. 研究表明,先天性免疫系统主要通过免疫细胞表面的模式识别受体(pattern recognition receptor, PRR)并结合病原体中存在的病原体相关分子模式(pathogen-associated molecular pattern, PAMP),包括微生物的 CpG ODN、脂蛋白、dsRNA 等,从而引发一系列的免疫反应,以消灭入侵的病原体<sup>[6]</sup>. 其中,抗原呈递细胞上表达的 Toll 样受体(Toll-like receptor, TLRs)家族是关键 PRR,在诱导先天性免疫应答以及适应性免疫应答中起着重要作用<sup>[3]</sup>. 根据不同 TLRs 识别不同病原相关分子模式,鱼类 TLR 通常分为两类,第一类包括 *TLR1*、*TLR2*、*TLR4*、*TLR5* 和 *TLR9*,主要参与机体对细菌的识别,被称为非病毒型 TLRs;第二类包括 *TLR3*、*TLR7*、*TLR8* 和 *TLR22*,主要参与病毒的识别,被称为病毒型 TLRs<sup>[7]</sup>.

到目前为止,鱼类中已经鉴定出大约 23 种 TLRs<sup>[8]</sup>,典型的 TLRs 蛋白结构包括亮氨酸重复区(LRRs)、跨膜结构域(TM)、胞内 Toll/IL-1 受体结构域(TIR),其中 LRRs 区通过与相应的配体结合,使得 Toll 样受体被激活,进一步通过 TIR 结构域与接头分子 MyD88(髓样分化因子)结合,诱导产生炎症细胞因子抵御病原菌的入侵,TIR 区相对保守,与分子信号转导有关<sup>[3]</sup>. TLRs 在鱼类免疫中的作用已被广泛研究,病原体感染后 TLRs 相关基因表达量会发生不同的变化来适应特定的生存环境. 大西洋鲷(*Sparus aurata*)被嗜水气单胞菌感染后,*TLR2* 表达发生上调<sup>[9]</sup>;大黄鱼(*Larimichthys crocea*)被嗜水气单胞菌感染后,*TLR1*、*TLR3* 表达上调,*TLR2*、*TLR22* 表达下调<sup>[10]</sup>;海鱼分枝杆菌也被证实可以诱导斑马鱼 *TLR9* 表达上调<sup>[11]</sup>. 值得注意的是,*TLR9* 与其他非病毒型 TLRs 不同,该受体只存在于细胞内部,也是识别细菌和病毒 DNA 中 CpG ODN 的主要受体,通过与 CpG 特定的序列结合最终引起免疫器官的功能性成熟,进而产生一系列免疫效应细胞<sup>[12]</sup>,此外,该受体也被认为能与其配体协同进化,因此是研究鱼类抗病原体免疫适应的重要模型. 在非病毒型 TLRs 中,*TLR1* 似乎也拥有独特的抗病原免疫机制,尽管目前 *TLR1* 在大黄鱼、绿河豚(*Tetraodon nigroviridis*)、斑马鱼等物种中鉴定出来,并且病原感染实验也表明该受体在鱼类病原免疫中发挥着重要作用,但是目前鱼类 *TLR1* 的配体还未明确,相关研究表明 *TLR1* 可能需要与 *TLR2* 形成异源二聚体发挥作用,因此 *TLR1* 仍需要更深入的研究<sup>[13-14]</sup>. 目前国内外对鲈形目 TLRs 家族研究大多局限在免疫学实验中,然而其分子进化基础仍不完全清楚. 本研究以 17 种鲈形目物种的 *TLR1*、*TLR9* 为候选基因,尝试通过分析这两个基因分子进化揭示其先天性免疫机制. 研究结果对于鲈形目鱼类的遗传育种以及疾病防治具有一定的指导意义.

## 1 材料与方法

### 1.1 *TLR1* 和 *TLR9* 序列获得与比对

本研究共选取 17 种鲈形目物种,涵盖攀鲈科、鲮科、狼鲈科、鲈科、鮨科、花鲈科、鳄冰鱼科、龙躄科、南极鱼科、石首鱼科、鲷科,并选取斑马鱼作为外群. 除了真鲷以及黑鲷 *TLR1* 基因外,本研究所使用物种的 *TLR1*、*TLR9* 基因序列信息均来自 NCBI(<http://www.ncbi.nlm.nih.gov/>)数据库(表 1). 进一步,以大西洋鲷 *TLR1* 作为参考序列在真鲷(*Pagrus major*)以及黑鲷(*Acanthopagrus schlegelii*)基因组中进行本地 Blast<sup>[15]</sup>以

获取对应的 *TLR1* 基因,E-value 值设定为  $1e-5$ ,其中真鲷基因组来自 NCBI 数据库,黑鲷基因组为江苏省海洋水产研究所与南京师范大学合作测序数据(未发表).使用 Prank<sup>[16]</sup> 软件对每一个基因基于密码子水平进行多序列比对,然后使用 Gblocks<sup>[17]</sup> 软件对序列非保守区域进行比对和人工校对.

表 1 本文研究所选物种基因序列登录号

Table 1 Accession numbers of species used in this study

物种名	拉丁名	科	TLR1 登录号	TLR9 登录号
真鲷	<i>Pagrus major</i>	鲷科	本地 blast *	EU256334.1
黑鲷	<i>Acanthopagrus schlegelii</i>	鲷科	本地 blast *	EU256333.2
大西洋鲷	<i>Sparus aurata</i>	鲷科	XM_030396315.1	XM_030421964.1
鲈鱼	<i>Micthys miiuy</i>	石首鱼科	KP340899.1	KR709252.1
白梭吻鲈	<i>Sander lucioperca</i>	河鲈科	XM_031319317.1	XM_031316335.1
黄鲈	<i>Perca flavescens</i>	河鲈科	XM_028589377.1	XM_028584076.1
橙胸镖鲈	<i>Etheostoma spectabile</i>	河鲈科	XM_032528600.1	XM_032521582.1
斜带石斑鱼	<i>Epinephelus coioides</i>	鲷科	HM357229.1	GQ358201.1
鞍带石斑鱼	<i>Epinephelus lanceolatus</i>	鲷科	XM_033632980.1	XM_033629987.1
黄尾鲷	<i>Seriola lalandi</i>	鲷科	XM_023396084.1	KR364001.2
卵形鲳鲹	<i>Trachinotus ovatus</i>	鲷科	MG762971.1	KU975048.1
龟壳攀鲈	<i>Anabas testudineus</i>	攀鲈科	XM_026357347.1	XM_026367613.1
眼斑鳢	<i>Anarrhichthys ocellatus</i>	狼鲈科	XM_031858208.1	XM_031844856.1
南乔治亚拟冰鱼	<i>Pseudochaenichthys georgianus</i>	鳕鱼科	XM_034092824.1	XM_034086580.1
尖头裸龙躄	<i>Gymnodraco acuticeps</i>	龙躄科	XM_034230873.1	XM_034207489.1
伯氏肩孔南极鱼	<i>Trematomus bernacchii</i>	南极鱼科	XM_034134081.1	XM_034145045.1
日本真鲈	<i>Lateolabrax japonicus</i>	花鲈科	KY883370.1	MK273053.1
斑马鱼	<i>Danio rerio</i>	鲤科	NM_001130593.1	NM_001130594.1

注: \* 基因序列采用 Blast 获得.

1.2 TLR9 和 TLR1 基因树重建

为探讨 TLRs 在鲈形目物种内部的系统发生关系,通过 Modelgenerator<sup>[18]</sup> 软件选择最佳核苷酸替换模型,使用 RAxML<sup>[19]</sup> 软件构建最大似然系统发生树,斑马鱼作为外群,自展 1 000 次估计节点的支持度(bootstrap value).另外,为了使得树的结果更加可靠,也使用了 Mrbayes<sup>[20]</sup> 软件基于贝叶斯法进行构树,其中后验概率借助马尔科夫链蒙特卡罗方法(markov chain monte carlo,MCMC)运行 100000 代进行估计.

1.3 分子进化分析

为了探究 *TLR1* 和 *TLR9* 在鲈形目中的进化模式,本研究使用 PAML4.7(phylogenetic analysis by maximum Likelihood,PAML)软件<sup>[21]</sup> 中 CODEML 程序对非同义替换率(nonsynonymous, $d_N$ )和同义替代率(synonymous, $d_S$ )的比值( $\omega=d_N/d_S$ )进行评估. $\omega$  值为衡量选择压力的重要指标,其中 $\omega<1$ 、 $=1$  和 $>1$  分别意味着受纯化选择(Purify selection)、中性选择(Neutral selection)和正选择(Positive selective).使用目前公认的鲈形目系统发生关系<sup>[22]</sup> 以及 TimeTree 网站(<http://www.timetree.org/>)的系统发育树作为 PAML 分析的输入树(图 1).

运用基于最大似然法的 PAML4.7 软件包<sup>[21]</sup> 和 Datamonkey<sup>[23]</sup> 进行评估鲈形目物种中 *TLR1* 和 *TLR9* 两个基因座中受正选择作用的位点.首先,运用 PAML 软件包中的位点模型(Site model)中的两对模型:M1a(中性模型)vs. M2a(正选择模型)、M8a(中性模型,beta 分布: $0<\omega_0<1$  和  $\omega_1=1$ )vs. M8(正选择模型;beta 分布: $0<\omega_0<1$  和  $\omega_1>1$ )进行检测.运用似然比检验(likelihood ratio test,LRT)评估嵌套模型中的最佳模型,并通过  $2\Delta\ln L$  值与自由度之间的卡方分布(Chi-square test, $\chi^2$ )关系检测上述两个模型的显著性.当模型通过 LRT 检验且  $P$  值显著时,运用贝叶斯经验贝叶斯法<sup>[24]</sup>(bayes empirical bayes,BEB)检测正选择位点,其中后验概率(posterior probabilities,PP)值大于 0.9 的位点作为潜在的正选择位点.另外,基于同义替代的 Datamonkey<sup>[23]</sup> 方法进一步用于正选择位点的检测,运用了三种最大似然法,包括固定效应似然法(fixed effects likelihood,FEL)、随机效应似然法(random effects likelihood,REL)、快速无约束贝叶斯估算法(fast unconstrained bayesian approximation,FUBAR).FEL 显著性水平设置为 0.2,REL 贝叶斯显著水平设置为 50,FUBAR 中后验概率固定阈值 0.8.

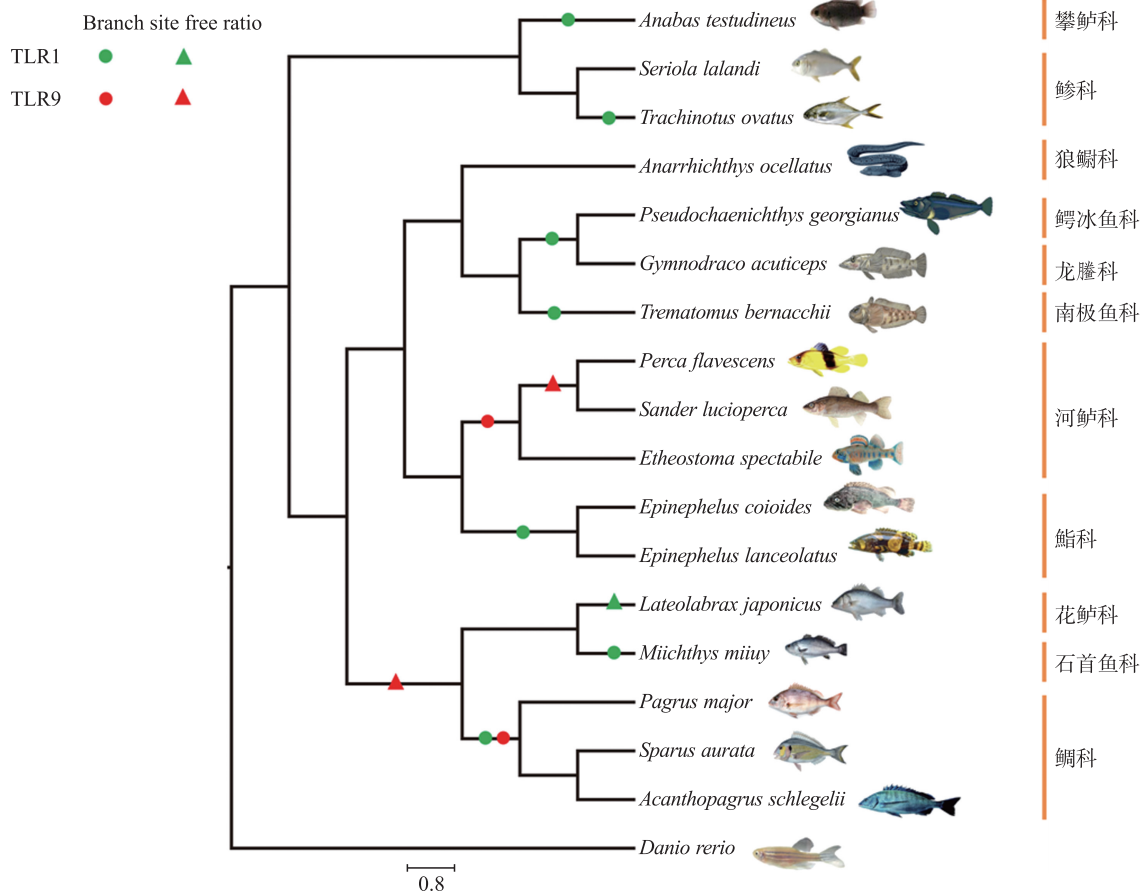


图 1 本研究用于 PAML 分析的输入树. Branch-site、Free ratio 结果分别使用圆圈以及三角符号标记

Fig. 1 The input phylogenetic tree used in PAML analyses in this study. The results of Branch-site and Free ratio models were marked with circles and triangles, respectively

为了进一步检测正选择是否局限于特定的进化谱系,运用 PAML4.7 软件包中的自由比率模型 (Free-ratio model) 和支位点模型 (Branch-site model) 进行检测. 自由比率模型假设每一支系具有独立的  $\omega$  值,该模型与零假设——单一比率模型 (One-ratio model) 即所有支系具有相同的  $\omega$  值进行比较. 支位点模型 (Branch-site model) 是最严格的选择压力检测的方法,不仅能够检测受正选择的支系还能检测受正选择的位点. 该模型需要将进化谱系分为前景支 (foreground branch, 即感兴趣的支系) 与背景支 (background branch, 即其余的支系). 本研究中,将 17 种鲈形目物种分别标记为前景支,零假设 Ma0 (中性模型:  $0 < \omega_0 < 1$ ,  $\omega_1 = 1$  和  $\omega_2 = 1$ ) 和备择假设 Ma (正选择模型:  $0 < \omega_0 < 1$ ,  $\omega_1 > 1$  和  $\omega_2 \geq 1$ ) 进行比较. 运用 LRT 检测嵌套模型的显著性,  $P < 0.05$  时则为模型显著,且所检测的位点  $PP > 0.9$  时,被认为是正选择位点.

#### 1.4 TLR1 和 TLR9 蛋白三维结构预测

通过 MEGA<sup>[22]</sup> 软件将 *TLR1*、*TLR9* 完整核苷酸序列翻译为氨基酸序列,利用 SMART 网站 (<http://smart.embl-heidelberg.de/>) 预测蛋白关键结构域. 蛋白质结构域中氨基酸的变化可能引起蛋白质空间构型的改变,进而影响蛋白质功能. 为了进一步为探究正选择位点在蛋白质三维结构上的空间分布,本研究使用 I-TASSER (Iterative Threading ASSEmbly Refinement) 网站对黑鲷 TLR1 和 TLR9 蛋白三维结构进行预测,并通过 EZMOL (<http://www.sbg.bio.ic.ac.uk/ezmol/>) 网站定位正选择位点在蛋白质三维结构上的位置,从而更直观地分析位点对蛋白质生物学功能的影响.

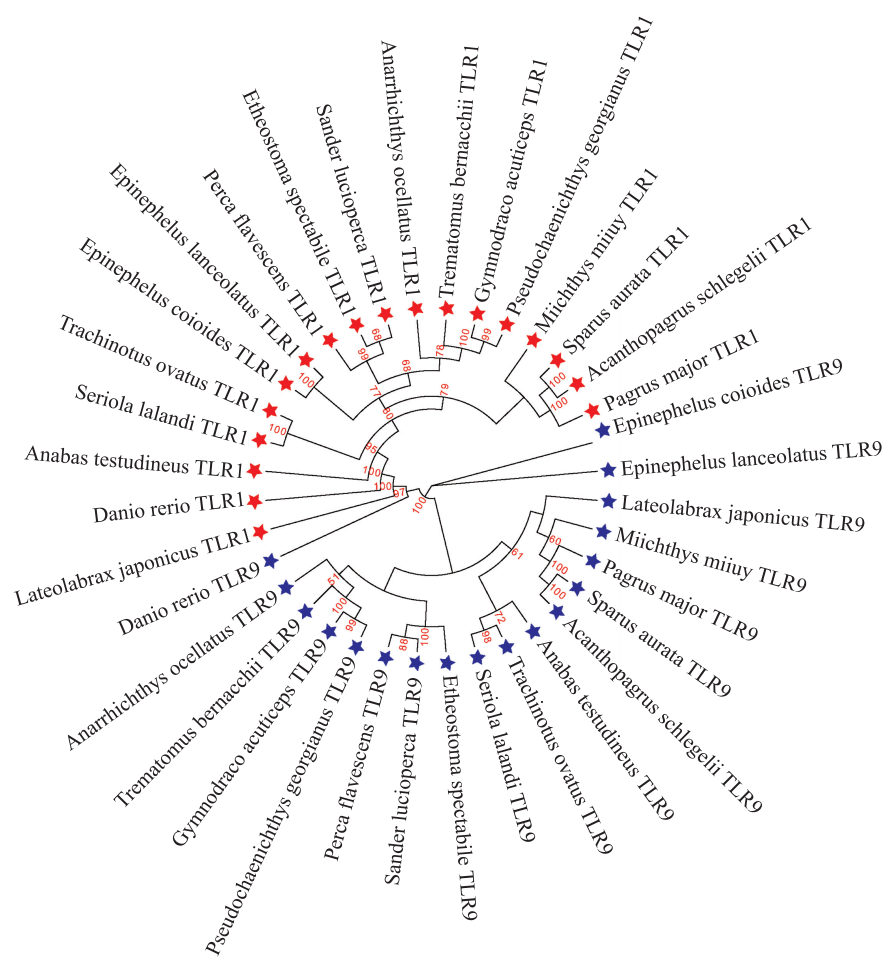
## 2 结果与讨论

### 2.1 系统发育树的构建

为了探讨鲈形目 TLRs 的系统发生关系,首先,基于 Modelgenerator<sup>[19]</sup> 软件分析,发现 GTR+GAMMA+I 模型为核苷酸替代的最佳模型. 其次,使用 RAxML 软件构建基因最大似然树 (图 2). 基因树的拓扑结构



显示鲈形目 TLRs 的基因根据 *TLR1* 和 *TLR9* 两个基因座聚成了两个基因簇。但是,基因树的拓扑结构与物种树存在差异,在 *TLR1* 物种树中,日本真鲈(*Lateolabrax japonicus*)与鲢鱼(*Miichthys miiuy*)聚在一起,而在基因树中,鲢鱼与鲷科的大西洋鲷拓扑关系最近。在 *TLR9* 基因树中,除了两个鲷科物种斜带石斑鱼(*Epinephelus coioides*)以及鞍带石斑鱼(*E. lanceolatus*)没有与黄鲈(*Perca flavescens*)聚在一起,基因树与物种树其他物种亲缘关系基本一致。同时,贝叶斯树也同样支持了最大似然法构建的基因树,两种方法得到的物种基因树在整体拓扑结构上基本一致(图 3)。综上,基因树与物种树的拓扑差异表明在鱼类进化过程中,*TLR1*、*TLR9* 基因已经发生了独特的进化,以适应不同生境下病原微生物的侵袭。



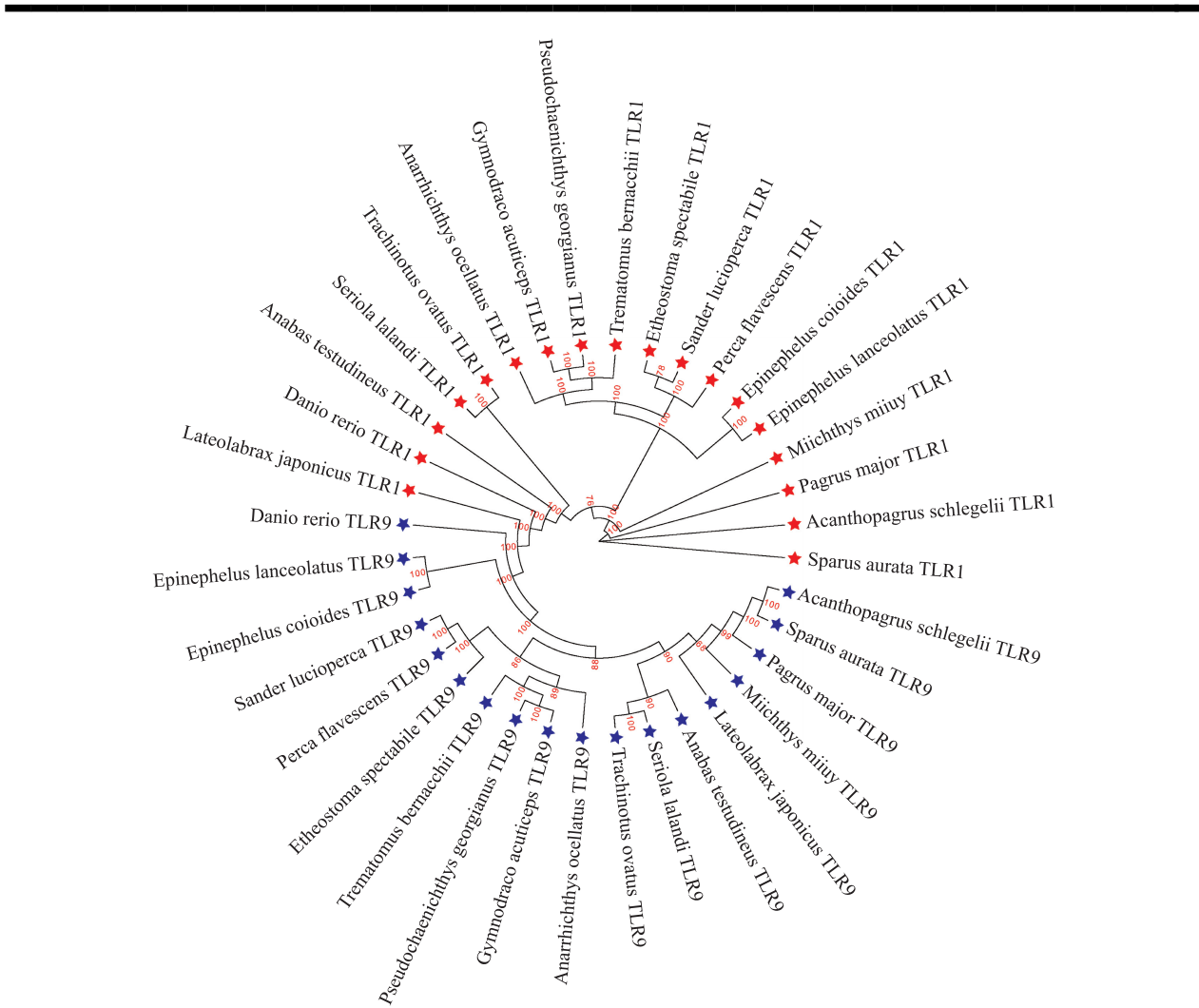
红色和蓝色三角形、五角星分别代表鲈形目 *TLR1* 以及 *TLR9* 支系。节点数字表示支持度大于 50%

图 2 最大似然法构建鲈形目 *TLR1*、*TLR9* 系统发育树

Fig. 2 Reconstruction of the *TLR1* and *TLR9* phylogenetic tree of Perciformes by the maximum likelihood method

2.2 选择压力分析

为了检测 *TLR1* 和 *TLR9* 两个基因座在鲈形目物种中是否存在正选择位点,首先运用 PAML4.7 软件中位点模型(M1a vs. M2a、M8a vs. M8)进行检测。LRT 结果显示正选择模型 M2a 和 M8 均显著优于中性模型 M1a 和 M8a。在 *TLR1* 基因座位中,正选择模型 M2a、M8 中  $\omega$  值分别为 2.55 和 2.05,提示 *TLR1* 受正选择,并且 M2a、M8 通过 BEB 方法分别检测到 3 个、47 个正选择位点的后验概率大于 0.9,在 *TLR9* 检测结果中,M2a、M8 模型分别检测到 3、30 个正选择位点。当后验概率提高到 0.95 时,在 *TLR1* 检测结果中,M2a、M8 分别筛选到 1 个、20 个正选择位点,而在 *TLR9* 中,M2a 模型未检测到后验概率大于 0.95 的位点,M8 模型则检测到 10 个正选择位点。另外,采用 DataMonkey 网站中的三种最大似然法(FEL、REL 以及 FUBAR)进一步对两个基因座位进行选择压力检测。结果表明,运用三种方法在 *TLR1* 基因座位(FEL:28, REL:8,FUBAR:13)以及 *TLR9* 基因座位(FEL:24,REL:6,FUBAR:3)共检测 53 个正选择位点。综合以上五种最大似然法,*TLR1* 与 *TLR9* 中被两种以上方法同时检测到的正选择位点分别为 16 和 8 个,提示这些位点为强烈的正选择位点,同时被三种以上方法检测到的正选择位点分别为 7 个和 3 个(表 2)。



红色和蓝色五角星三角形、五角星分别代表鲈形目 *TLR1* 以及 *TLR9* 支系. 节点数字表示后验概率

图 3 贝叶斯法构建鲈形目 *TLR1*、*TLR9* 系统发育树

Fig. 3 Reconstruction of the *TLR1* and *TLR9* phylogenetic tree of Perciformes by Bayesian method

表 2 PAML 和 Datamonkey 检测到的正选择位点

Table 2 Sites under positive selection detected by PAML and Datamonkey

Gene	PAML		Datamonkey			Total no. of sites
	M1a	M8	FEL	FUBAR	REL	
<i>TLR1</i>	236	5, 48, 58, 61, 156, 177, 195, 207, 223, 226, 236, 251, 264, 271, 279, 289, 305, 364, 568, 578	46, 48, 111, 120, 125, 128, 164, 177, 195, 205, 214, 236, 239, 251, 331, 364, 365, 409, 418, 421, 515, 553, 564, 592, 604, 635, 730, 736	111, 120, 125, 195, 205, 214, 236, 251, 364, 418, 421, 564, 592	48, 120, 177, 195, 236, 239, 251, 253	16(7)
<i>TLR9</i>	NA	41, 122, 242, 269, 348, 432, 495, 545, 692, 716	14, 29, 41, 81, 210, 242, 327, 374, 391, 406, 428, 453, 470, 487, 514, 613, 620, 692, 730, 733, 753, 811, 819, 858	14, 242, 327	14, 41, 242, 428, 730, 819	8(3)

注:被两种以上方法检测到的位点标粗,被三种以上方法检测到的位点标粗并且斜体.

运用单一比率模型(One-ratio model)假设所有支系具有相同的进化速率,结果显示鲈形目物种 *TLR1* 和 *TLR9* 的  $\omega$  值分别为 0.38 与 0.36,显著小于 1,提示这两个基因整体受到强烈的选择约束. 为了进一步探究正选择是否局限于鲈形目特异的支系,运用自由比率模型(Free-ratio model)进行检测. LRT 结果显示嵌套模型(One ratio vs. Free ratio)差异显著( $P<0.01$ ),说明自由比率模型显著优于单一比率模型. Free-ratio 模型发现 *TLR1* 在日本真鲈显著受正选择;而 *TLR9* 基因座位中则检测到多个正选择支系,例如黄鲈与白梭吻鲈(*Sander lucioperca*)最近共同祖先支,日本真鲈、鲈鱼、真鲷、大西洋鲷以及黑鲷的最近共同祖先

支(图 2、表 3)。Free ratio 模型检测到一些支系中存在正选择信号( $\omega>1$ ),据报道这些支系长期受到病原微生物的侵袭,比如鲷科支系已被报道经常受到虹彩病毒侵袭,引起组织细胞肿大坏死<sup>[25]</sup>。

运用最严格的支位点模型(Branch-site model)检测选择压力,结果提示在 *TLR1* 基因座位中,正选择分别发生在龟壳攀鲈(*Anabas testudineus*)、鲢鱼、卵形鲳鲹(*Trachinotus ovatus*)、伯氏肩孔南极鱼(*Trematomus bernacchii*)以及南乔治亚拟冰鱼(*Pseudochaenichthys georgianus*)、尖头裸龙𩚰(*Gymnodraco acuticeps*)最近共同祖先支,斜带石斑鱼和鞍带石斑鱼最近共同祖先支,真鲷、黑鲷及大西洋鲷最近共同祖先支;在 *TLR9* 中,黄鲈、白梭吻鲈以及橙胸镖鲈(*Etheostoma spectabile*)最近共同祖先支,真鲷、黑鲷及大西洋鲷最近共同祖先支也分别检测到正选择位点(表 4)。

表 3 基因位点模型和分支模型检测结果  
Table 3 The result for site model and branch model(PAML) tests of genes

基因 Gene	模型 Model	-Ln 值 -LnL	LRT 值 2ΔlnL	$\omega$ (dN/dS)	正选择位点/支系(pp>0.95) Positively selected sites/branches	P value
<i>TLR1</i>	M1a	1 3897.28		1		
	M2a	1 3877.94	38.68	2.55	236(0.987)	<0.01
	M8a	13 896.79		1	5(0.953);48(0.980);58(0.984);61(0.959);	
	M8	13 878.99	35.60	2.05	156(0.982);177(0.968);195(0.951);207(0.970);	<0.01
					223(0.964);226(0.980);236(0.996);251(0.984);	
					264(0.990);271(0.990);279(0.995);289(0.970);	
<i>TLR9</i>	one	19 219.59		0.38	305(0.982);364(0.957);568(0.961);578(0.984)	
	Free-ratio	19 090.37	258.44		<i>L. japonicus</i>	<0.01
	M1a	18 979.44		1		<0.01
	M2a	18 965.50	27.88	2.01	NA	
	M8a	18 971.51		1	41(0.964);122(0.955);242(0.986);269(0.957);	
	M8	18 951.29	40.44	1.59	348(0.953);432(0.964);495(0.955);545(0.956);	<0.01
	one	22 721.93		0.36	692(0.979);716(0.967)	
	Free-ratio	22 613.34	217.18		LCA of <i>S. lucioperca</i> ;LCA of Sparidae and Lateolabracidae	<0.01

注:LCA(Last common ancestors),最近共同祖先支;PP(Posterior probability),后验概率。

表 4 *TLR1* 和 *TLR9* 基因支位点模型检测结果  
Table 4 The result for Branch-site model test of genes

基因 gene	支位点模型 Branch-site model	-Ln 值 -LnL	P value	$\omega$ (dN/dS)	正选择位点(PP) Positively selected sites
<i>TLR1</i>	<i>A. testudineus</i>				
	Ma0	18 697.14		$\omega_0=0.088\ 38,\omega_1=1.000\ 00,\omega_2=13.239\ 90$	
	Ma	18 692.26	<0.01	$\omega_0=0.087\ 78,\omega_1=1.000\ 00,\omega_2=1.000\ 00$	449(0.994)
	<i>M. miiuy</i>				
	Ma0	18 694.60		$\omega_0=0.084\ 17,\omega_1=1.000\ 00,\omega_2=1.000\ 00$	
	Ma	18 692.36	0.034	$\omega_0=0.084\ 63,\omega_1=1.000\ 00,\omega_2=6.083\ 43$	583(0.993);766(0.967)
	<i>T. ovatus</i>				
	Ma0	18 694.55		$\omega_0=0.084\ 36,\omega_1=1.000\ 00,\omega_2=1.000\ 00$	
	Ma	18 692.52	0.044	$\omega_0=0.084\ 97,\omega_1=1.000\ 00,\omega_2=6.592\ 38$	3(0.972);452(0.985)
	LCA of <i>P. georgianus</i>				
	Ma0	18 695.17		$\omega_0=0.086\ 80,\omega_1=1.000\ 00,\omega_2=85.873\ 21$	
	Ma	18 692.38	0.018	$\omega_0=0.086\ 67,\omega_1=1.000\ 00,\omega_2=1.000\ 00$	332(0.995)
	LCA of <i>E. coioides</i>				
	Ma0	18 697.26		$\omega_0=0.087\ 33,\omega_1=1.000\ 00,\omega_2=1.000\ 00$	
	Ma	18 692.51	<0.01	$\omega_0=0.087\ 78,\omega_1=1.000\ 00,\omega_2=33.552\ 93$	773(0.945)
	LCA of <i>P. major</i>				
	Ma0	18 696.08		$\omega_0=0.086\ 24,\omega_1=1.000\ 00,\omega_2=1.000\ 00$	
	Ma	18 692.15	<0.01	$\omega_0=0.087\ 84,\omega_1=1.000\ 00,\omega_2=37.503\ 84$	155(0.985)
<i>T. bernacchii</i>	Ma0	18 694.79		$\omega_0=0.085\ 96,\omega_1=1.000\ 00,\omega_2=37.014\ 13$	
	Ma	18 692.19	0.022	$\omega_0=0.085\ 80,\omega_1=1.000\ 00,\omega_2=1.000\ 00$	151(0.947);356(0.944);523(0.945)

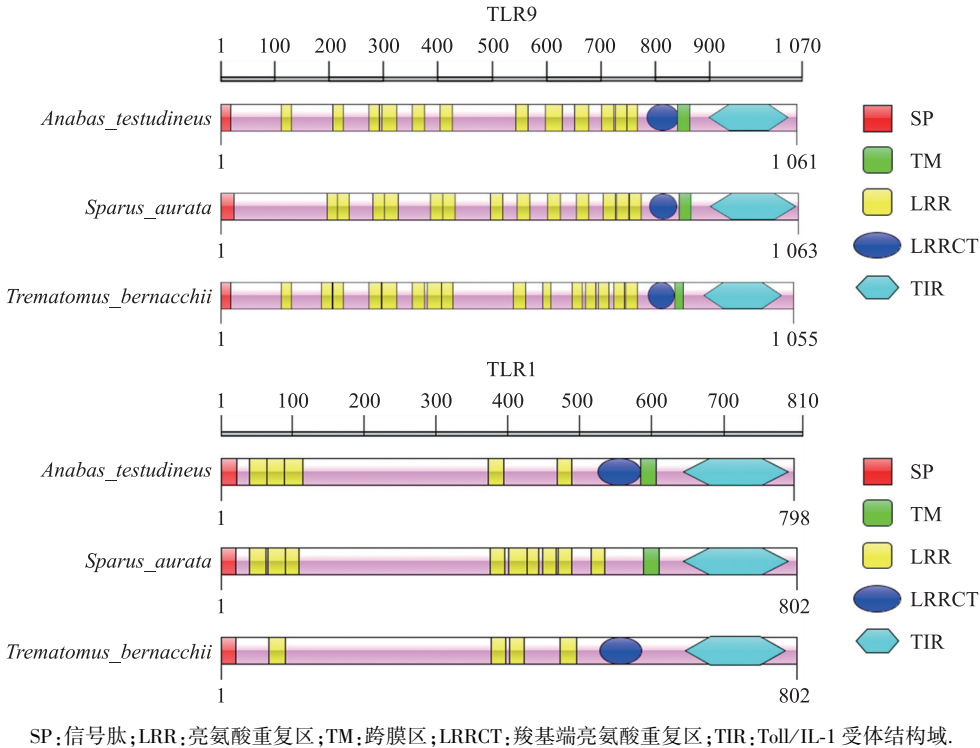
续表 4 Table 4 continued

基因 gene	支位点模型 Branch-site model	-Ln 值 -LnL	<i>P</i> value	$\omega$ ( dN/dS)	正选择位点 (PP) Positively selected sites
<i>TLR9</i>	LCA of <i>E. spectabile</i>				
	Ma0	22 010.87		$\omega_0 = 0.085\ 81, \omega_1 = 1.000\ 00, \omega_2 = 1.000\ 00$	
	Ma	22 006.72	<0.01	$\omega_0 = 0.085\ 78, \omega_1 = 1.000\ 00, \omega_2 = 52.601\ 61$	790( 0.983 )
	LCA of <i>P. major</i>				
	Ma0	22 011.39		$\omega_0 = 0.086\ 49, \omega_1 = 1.000\ 00, \omega_2 = 1.000\ 00$	
	Ma	22 004.18	<0.01	$\omega_0 = 0.086\ 62, \omega_1 = 1.000\ 00, \omega_2 = 75.680\ 34$	124( 0.966 )

注:LCA(Last common ancestors),最近共同祖先支;PP(Posterior probability),后验概率

2.3 TLR1、TLR9 结构预测以及正选择位点标注在三维结构中

本研究选取了亲缘关系较远的大西洋鲷(鲷科)、伯氏肩孔南极鱼(南极鱼科)、龟壳攀鲈(攀鲈科)三个物种为代表进行氨基酸序列分析. 利用 SMART 网站分别预测 TLR1、TLR9 蛋白结构域并通过 IBS<sup>[26]</sup> 软件对结构域进行可视化(图 4). 龟壳攀鲈、大西洋鲷、伯氏肩孔南极鱼三个物种中,TLR1 编码序列长度相似,分别为 798、802、802 个氨基酸,其中大西洋鲷具有最多的亮氨酸重复区(13 个),但值得注意的是,此物种未检测到羧基端亮氨酸重复区;伯氏肩孔南极鱼亮氨酸重复区数量最少(4 个)并且未检测到跨膜区. TLR9 在三个物种中相对更为保守,都含有典型的 TLRs 蛋白结构域(LRR、TM、TIR),其中亮氨酸重复区数量分别为 12、13、15 个. 为了进一步探究正选择位点对蛋白质功能的影响,将这些位点标注在蛋白三维结构中,首先利用 I-TASSER 网站预测了黑鲷蛋白三维结构模型并通过 EZMOL 在线网站将同时被两种最大似然法鉴定到的正选择位点(TLR1:16 和 TLR9:8)标注到三维模型中(图 5),发现 58%的正选择位点位于蛋白质重要功能域,且 86%强烈的正选择位点位于病原体识别区域亮氨酸重复区(LRRs).



SP:信号肽;LRR:亮氨酸重复区;TM:跨膜区;LRRCT:羧基端亮氨酸重复区;TIR:Toll/IL-1 受体结构域.

图 4 不同物种 TLR1、TLR9 蛋白结构比较

Fig. 4 Comparison of the TLR1、TLR9 domain structures among various species

3 结论

TLRs 是进化保守的蛋白质,传统上受到强烈的功能约束. 然而,最近的研究已经在一些脊椎动物群体中发现正选择的证据,如鸟类中 *TLR3*、4、5 和 15 基因、啮齿动物的 *TLR4* 基因等<sup>[27-29]</sup>. 相对于恒温哺乳类、鸟类物种,生活在水体中的鱼类因没有自我调节体温的机制,当面临病原微生物入侵时,低温诱导的淋巴



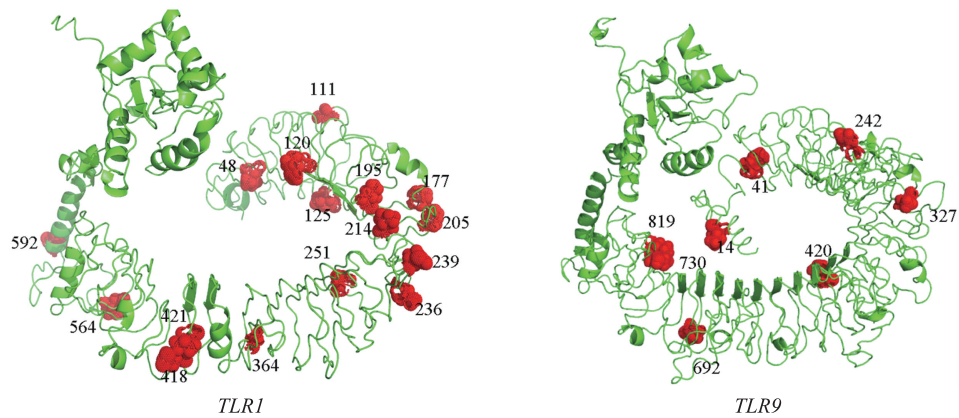


图 5 基因正选择位点在蛋白三维结构中的标注

Fig. 5 Annotation of the positive selection sites of genes in the three-dimensional structure of the protein

细胞的增值减少将导致适应性免疫相对滞后且免疫应答能效较低,因此对栖息于病原体较多的水环境鱼类而言,先天性免疫在抵御病原体侵染的过程中扮演更为重要的角色<sup>[28]</sup>. 作为 Toll 样受体基因家族中的重要成员,*TLR1*、*TLR9* 分别可以识别三酰基脂肽、细菌与病毒非甲基化的 CpG 寡核苷酸序列,尽管在鱼类免疫识别应答方面的功能已被广泛的研究,但是基因的分子进化基础尚不清楚. 目前,已经报道了鲈形目物种容易受到不同的病原微生物的感染,包括真鲷虹彩病毒(RSIV)、鲈鱼虹彩病毒(WVIV)、淋巴囊肿病毒(LCDV)、红细胞坏死病毒(VENV)、传染性胰腺坏死病毒(IPNV)、流行性造血器官坏死病毒(EHNV)、神经坏死症病毒(NNV)等<sup>[25,30]</sup>. 本研究通过分析鲈形目物种先天免疫相关基因 *TLR1*、*TLR9* 的进化模式,发现物种普遍受到显著的正选择. 在位点模型中,*TLR1* 与 *TLR9* 中被两种以上方法检测到的正选择位点分别为 18、6 个,且 86%位点处于 LRR 区. 已报道 LRR 结构域在识别病原体组分中发挥核心作用,且 LRR 和 LRR\_CT 区域还参与病原体刺激后的信号转导过程<sup>[31]</sup>. 该区检测到更多正选择位点可能有利于物种在进化过程中识别不同的配体进而有效的抵抗病原微生物. 另外,相比于 LRR 区,TIR 区具有更高的保守性,并且在信号转导过程中发挥重要作用. 鲈形目物种在 TIR 结构域未检测到正选择位点,提示该区域可能受到强烈的功能约束.

特别重要的是,也在鲈形目一些支系中检测到特有的正选择信号,并且检测到的正选择位点与位点模型所检测到的位点不同,这些支系可能经历了更多的选择来应对不同病原体的感染. 两个基因在鲷科的最近共同祖先支均检测到正选择信号,之前研究也证明鲷科物种经常感染淋巴囊肿病、病毒性表皮坏死病等<sup>[25]</sup>. 石斑鱼属是环境适应能力较强的暖水性鱼类,尤其是斜带石斑鱼,生活温度可以在 11~41℃ 之间. 目前,病毒性神经坏死症是石斑鱼流行病之一<sup>[32]</sup>,研究也在两个石斑鱼属物种(斜带石斑鱼、鞍带石斑鱼)的祖先支检测到 *TLR1* 基因正选择的信号,可能与其独特的环境适应能力有关. 另外,河鲈科物种被列为病毒性出血性败血症的易感物种<sup>[33]</sup>. 作为全球有鳍鱼类致病性最强的病毒性疾病之一,宿主致死率可高达 90%,在河鲈科物种的祖先支中也检测到 *TLR9* 基因存在正选择信号. 这些物种检测到显著的正选择信号,表明随着时间的推移,物种抗病毒免疫反应可能已经发生了进化(图 1,表 4),提示鲈形目物种 *TLR1*、*TLR9* 的分子进化可能是由环境中的病原微生物介导的选择压力所驱动的,从而具有较强的抵御病原微生物侵袭的能力.

另外,两个基因受正选择的比例也明显不同,在 *TLR1* 中检测到更多的正选择信号,并且系统发育树也表明 *TLR1* 基因树与物种树相比,发生了更多的拓扑结构变化,表明鱼类在适应病原微生物的过程中,*TLR1*、*TLR9* 基因受到不同的选择机制且 *TLR1* 进化更为明显. 本研究也发现 58%的正选择位点位于蛋白质重要功能域,并且大多数位于病原体识别区域亮氨酸重复区(LRRs),该区域的进化有利于识别不同病原体 PAMP,从而在抗病原免疫过程中发挥重要作用(图 4). 以上研究结果从进化的角度初步揭示了鲈形目物种先天性免疫基因 *TLR1*、*TLR9* 与病原体的相关性,为物种的疾病防治与预防提供理论依据.

#### [参考文献]

- [1] 陈大刚,张美昭. 中国海洋鱼类[M]. 山东:中国海洋大学出版社,2015.

- [2] LAFFERTY K D, HOFMANN E E. Marine disease impacts, diagnosis, forecasting, management and policy[J]. Philosophical transactions of the royal society b: biological sciences, 2016, 371(1689): 20150200.
- [3] AKIRA S, UEMATSU S, TAKEUCHI O. Pathogen recognition and innate immunity[J]. Cell, 2006, 124(4): 783–801.
- [4] IWASAKI A, MEDZHITOV R. Regulation of adaptive immunity by the innate immune system. [J]. Science, 2010, 327(5963): 291–295.
- [5] 王卫卫. 牙鲆免疫系统的结构及其部分免疫细胞的功能研究[D]. 青岛: 中国海洋大学, 2009.
- [6] MACLEOD C, BRYANT C E. Visualising pattern recognition receptor signalling[J]. Biochemical society transactions, 2017, 45(5): 1077–1085.
- [7] PIETRETTI D, WIEGERTJES G F. Ligand specificities of Toll-like receptors in fish: indications from infection studies[J]. Developmental & comparative immunology, 2014, 43(2): 205–222.
- [8] WANG Y, BI X, CHU Q, et al. Discovery of toll-like receptor 13 exists in the teleost fish: Miiuy croaker (Perciformes, Sciaenidae)[J]. Developmental & comparative immunology, 2016, 61: 25–33.
- [9] REYES-BECERRIL M, LÓPEZ-MEDINA T, ASCENCIO-VALLE F, et al. Immune response of gilthead seabream (Sparus aurata) following experimental infection with *Aeromonas hydrophila*[J]. Fish & shellfish immunology, 2011, 31(4): 564–570.
- [10] MU Y, DING F, CUI P, et al. Transcriptome and expression profiling analysis revealed changes of multiple signaling pathways involved in immunity in the large yellow croaker during *Aeromonas hydrophila* infection[J]. BMC genomics, 2010, 11(1): 1–14.
- [11] MEIJER A H, KRENS S G, RODRIGUEZ I A M, et al. Expression analysis of the Toll-like receptor and TIR domain adaptor families of zebrafish[J]. Molecular immunology, 2004, 40(11): 773–783.
- [12] KLASCHIK S, GURSEL I, KLINMAN D M. CpG-mediated changes in gene expression in murine spleen cells identified by microarray analysis[J]. Molecular immunology, 2007, 44(6): 1095–1104.
- [13] WU X Y, XIANG L X, HUANG L, et al. Characterization, expression and evolution analysis of Toll-like receptor 1 gene in pufferfish (*Tetraodon nigroviridis*) [J]. International journal of immunogenetics, 2008, 35(3): 215–225.
- [14] QUINIOU S M, BOUDINOT P, BENGTEÁN E. Comprehensive survey and genomic characterization of Toll-like receptors (TLRs) in channel catfish, *Ictalurus punctatus*: identification of novel fish TLRs[J]. Immunogenetics, 2013, 65(7): 511–530.
- [15] CAMACHO C, COULOURIS G, AVAGYAN V, et al. BLAST+: architecture and applications[J]. BMC bioinformatics, 2009, 10(1): 1–9.
- [16] LÖYTYNOJA A, GOLDMAN N. Phylogeny-aware gap placement prevents errors in sequence alignment and evolutionary analysis[J]. Science, 2008, 320(5883): 1632–1635.
- [17] CASTRESANA J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis[J]. Molecular biology and evolution, 2000, 17(4): 540–552.
- [18] KEANE T M, CREEVEY C J, PENTONY M M, et al. Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified[J]. BMC evolutionary biology, 2006, 6(1): 1–17.
- [19] STAMATAKIS A. RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies[J]. Bioinformatics, 2014, 30(9): 1312–1313.
- [20] HUELSENBECK J P, RONQUIST F. MRBAYES: Bayesian inference of phylogenetic trees[J]. Bioinformatics, 2001, 17(8): 754–755.
- [21] YANG Z. PAML 4: phylogenetic analysis by maximum likelihood[J]. Molecular biology and evolution, 2007, 24(8): 1586–1591.
- [22] SLOSS B L, BILLINGTON N, BURR B M. A molecular phylogeny of the Percidae (Teleostei, Perciformes) based on mitochondrial DNA sequence[J]. Molecular phylogenetics and evolution, 2004, 32(2): 545–562.
- [23] POND S L K, FROST S D. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments[J]. Bioinformatics, 2005, 21(10): 2531–2533.
- [24] YANG Z, WONG W S, NIELSEN R. Bayes empirical Bayes inference of amino acid sites under positive selection[J]. Molecular biology and evolution, 2005, 22(4): 1107–1118.
- [25] 吴金炉, 曾志南. 鱼类病毒病与鱼类病毒疫苗[J]. 海洋科学, 1999(4): 6.
- [26] LIU W, XIE Y, MA J, et al. IBS: an illustrator for the presentation and visualization of biological sequences[J]. Bioinformatics, 2015(20): 3359.

- [27] ALCAIDE M, EDWARDS S V. Molecular evolution of the toll-like receptor multigene family in birds[J]. *Molecular biology and evolution*, 2011, 28(5): 1703–1715.
- [28] FORNÛSKOVÁ A, VINKLER M, PAGÈS M, et al. Contrasted evolutionary histories of two Toll-like receptors (Tlr4 and Tlr7) in wild rodents (MURINAE)[J]. *BMC evolutionary biology*, 2013, 13(1): 1–17.
- [29] GOUNDASHEVA D, ANDONOVA M, IVANOV V. Changes in some parameters of the immune response in rats after cold stress[J]. *Journal of veterinary medicine, series b*, 1994, 41(1–10): 670–674.
- [30] 冯东岳, 温周瑞. 鱼类病毒病的研究进展及其展望[J]. *养殖与饲料*, 2011(6): 4.
- [31] WU M H, LI G Y. Structure and function of the leucine—rich repeat as a protein recognition motif[J]. *International journal of pathology and clinical medicine*, 2006, 26(1): 35–38.
- [32] LI L, HE J, MORI K I, et al. Mass mortalities associated with viral nervous necrosis in hatchery-reared groupers in the People's Republic of China[J]. *Fish pathology*, 2001, 36(3): 186–188.
- [33] SKALL H F, OLESEN N J, MELLERGAARD S. Viral haemorrhagic septicaemia virus in marine fish and its implications for fish farming—A review[J]. *Journal of fish diseases*, 2005, 28(9): 509–529.

[责任编辑:黄 敏]