

DOI: 10.3724/SP.J.1118.2018.17248

吉富尼罗罗非鱼 *Ikaros* 基因 5'调控区的克隆、序列分析及抗无乳链球菌相关 SNP 位点筛选

陈昆平^{1,2,3}, 卢迈新¹, 刘志刚¹, 高风英^{1,2}, 曹建萌¹, 张德锋¹, 李庆勇³, 胡欣欣^{1,2}

1. 中国水产科学研究院 珠江水产研究所, 农业部热带亚热带水产资源利用与养殖重点实验室, 广东 广州 510380;

2. 上海海洋大学 水产与生命学院, 上海 201306;

3. 惠州市渔业研究推广中心, 广东 惠州 516002

摘要: 为获得尼罗罗非鱼(*Oreochromis niloticus*)抗链球菌病相关的单核苷酸多态性(single nucleotide polymorphism, SNP)标记, 本研究通过染色体步移法克隆了尼罗罗非鱼 *Ikaros* 基因 5'调控区序列, 长度为 4178 bp, 并使用生物信息学软件对 *Ikaros* 基因的启动子和转录调控元件进行预测和分析。利用 PCR 产物直接测序法从亲本(P₀)尼罗罗非鱼 *Ikaros* 基因 5'调控区中筛选到 5 个 SNPs, 分别为 SNP1(g.562, G>A)、SNP2(g.217, G>T)、SNP3(g.-53, C>T)、SNP4(g.-220, T>C)和 SNP5(g.-579, T>C)。利用 Snapshot 技术对子一代(F₁)尼罗罗非鱼易感群体和抗病群体进行相应 SNPs 的基因分型, 分析两个群体遗传多样性参数, 结果显示多态信息含量(polymorphism information content, PIC)值在 0.0872~0.3747 之间, 表明 *Ikaros* 基因 5'调控区中所有 SNPs 的多态性均较低。*Ikaros* 基因 5'调控区 SNPs 与抗链球菌病性状关联分析结果表明, SNP2、SNP3、SNP4 和 SNP5 的基因型频率和等位基因频率在易感群体和抗病群体中存在显著差异($P<0.05$)。连锁不平衡分析结果显示 *Ikaros* 基因 5'调控区 SNPs 可形成 1 个单倍块和 5 种单倍型, 其中 GGCTT 单倍型与抗病性显著相关($P<0.05$), GGTCT 和 GTCCC 单倍型与易感性显著相关($P<0.05$)。此外, 还发现 SNP2 和 SNP5 处于完全连锁状态($r^2=1$, LOD=57.25, $D'=1$), 可作为罗非鱼抗链球菌病遗传育种的标签 SNP。综上所述, 在 *Ikaros* 基因 5'调控区筛选到 4 个抗链球菌病相关 SNPs 和 1 个单倍型(GGCTT), 均可作为尼罗罗非鱼分子育种的候选分子标记。

关键词: 尼罗罗非鱼; *Ikaros* 基因 5'调控区; 单核苷酸多态性; 无乳链球菌

中图分类号: S917

文献标志码: A

文章编号: 1005-8737-(2018)02-0237-14

罗非鱼(*Oreochromis*)作为联合国粮农组织(FAO)重点推广的养殖对象, 具有未来动物性蛋白质主要来源之一的美誉。中国的罗非鱼养殖产业发展迅速, 目前罗非鱼养殖产量、产值和出口额均居世界第一位^[1]。但是自 2009 年以来, 罗非鱼链球菌病在罗非鱼养殖中大规模暴发且频繁发生与流行, 给罗非鱼养殖产业造成了巨大的经济损失^[2-3]。培育罗非鱼抗病品系是解决罗非鱼链球

菌病的重要措施之一, 而分子标记辅助育种(molecular marker assisted selection, MAS)是一条快速高效的途径。

作为当前最理想的 DNA 分子标记之一, 单核苷酸多态性(single nucleotide polymorphism, SNP)是指生物体基因组单个核苷酸发生的突变^[4]。SNP 分子标记因具有数量多、分布广、遗传稳定且易于用多种方法进行基因分型等优点, 被广泛应用

收稿日期: 2017-06-29; 修订日期: 2017-09-22.

基金项目: 现代农业产业技术体系专项资金项目(CARS-46); 广东省自然科学基金项目(2016A030313146); 国家自然科学基金项目(31502205).

作者简介: 陈昆平(1991-), 男, 硕士研究生, 研究方向为水产健康养殖及遗传育种. E-mail: colinchenkunping@163.com

通信作者: 卢迈新(1962-), 男, 研究员, 研究方向为水产健康养殖及遗传育种. E-mail: mx-lu@163.com

于关联分析、遗传连锁图谱构建、MAS 和品种鉴定等方面,并表现出良好的应用前景^[5]。目前,在水产动物中已有较多利用免疫相关基因进行抗病关联分析的研究,如已有鲤^[6](*Cyprinus carpio*)、凡纳滨对虾^[7](*Litopenaeus vannamei*)、牙鲆^[8](*Paralichthys olivaceus*)和半滑舌鳎^[9](*Cynoglossus semilaevis*)等的免疫相关基因进行了抗病相关 SNPs 的筛选,为鱼类 MAS 奠定了基础。

Ikaros 基因编码的 *Ikaros* 是淋巴系统发育所必需的转录因子,在 preBCR、Notch、MAPK 等多个信号通路中均不可或缺,通过与 FOXO1、HuD、ETS1 等调控因子的相互作用来调控大量淋巴细胞发育相关下游基因的表达,对于淋巴细胞的正确发育及其免疫功能的发挥具有重要作用^[10-11]。*Ikaros* 基因最早在小鼠^[12](*Mus musculus*)和人类^[13](*Homo sapiens*)中被发现。研究表明该基因的突变会导致恶性肿瘤、系统性红斑狼疮(systemic lupus erythematosus, SLE)等疾病的发生与发展^[14-17]。在鱼类中,虽然 *Ikaros* 基因已在虹鳟^[18](*Onchorynchus mykiss*)、斑马鱼^[19](*Danio rerio*)和半滑舌鳎^[20]等物种进行克隆和组织表达分析,但其多态性与抗病性的关联分析却未见相关报道。

本团队在前期的研究中已成功克隆了尼罗罗非鱼(*Oreochromis niloticus*)*Ikaros* 基因,并通过 qRT-PCR 分析了该基因在健康鱼不同组织和感染无乳链球菌后免疫相关组织的表达情况,根据其表达变化规律,发现 *Ikaros* 基因参与调控尼罗罗非鱼抵御无乳链球菌的免疫应答反应^[21]。因此,本研究以 *Ikaros* 基因作为候选基因,通过染色体步移法克隆尼罗罗非鱼 *Ikaros* 基因的 5'调控区序列,然后利用生物信息学软件对该基因的启动子和转录调控元件进行预测和分析。采用 PCR 产物直接测序法筛查亲本(P₀)尼罗罗非鱼 *Ikaros* 基因 5'调控区中的 SNPs,通过人工感染无乳链球菌(*Streptococcus agalactiae*)的方法构建尼罗罗非鱼子一代(F₁)的易感群体和抗病群体,然后利用 Snapshot 技术对 2 个群体中相应 SNPs 进行基因分型,通过与抗链球菌病性状的关联分析,获得与无乳链球菌抗性相关 SNPs,为罗非鱼的分子育种提供有效的分子标记。

1 材料与方法

1.1 实验材料

1.1.1 实验用鱼 吉富尼罗罗非鱼^[22](简称吉富鱼)(*Oreochromis niloticus*, GIFT)来自中国水产科学研究院珠江水产研究所高要水产种质中心,共 20 个家系,亲本(P₀)经过人工感染无乳链球菌后,存活下来的个体继续繁殖产生子一代(F₁)。F₁ 各家系分别饲养在 10 m² 的水泥池中,当体重长至约为 50 g 时,进行人工感染无乳链球菌实验。

1.1.2 菌株 无乳链球菌菌株 WC1535 由本实验室分离保存,为 I a 型强毒株。

1.1.3 实验试剂 常规 PCR 所用的 *Taq* DNA 聚合酶、dNTP mixture 和 Buffer for *Taq* DNA 聚合酶购自上海申能博彩生物科技有限公司; Genome Walker™ Universal Kit 购自美国 Clontech 公司; *Escherichia coli* DH5 α 感受态细胞、pMD19-T 载体和反转录试剂盒购自 TaKaRa 宝生物工程(大连)有限公司; DNA 提取试剂盒和琼脂糖凝胶回收试剂盒购自天根生化科技(北京)有限公司;血平板培养基购自广东环凯微生物科技有限公司;无水乙醇;琼脂糖粉(西班牙, Biowest); 10 $\times$ TAE Buffer; TE 缓冲液。

1.2 实验方法

1.2.1 尼罗罗非鱼抗病群体和易感群体的构建 将无乳链球菌接种于血平板, 32 $^{\circ}$ C 培养 16 h 后用 PBS 缓冲液稀释菌体配置成浓度分别为 10⁵ CFU/mL、10⁶ CFU/mL、10⁷ CFU/mL、10⁸ CFU/mL 和 10⁹ CFU/mL 的菌液。将不同浓度的菌液分别注射于尼罗罗非鱼腹腔中, 每组 40 尾, 每尾 200 μ L, 对照组注射等量 PBS 缓冲液, 获得半致死浓度为 10⁷ CFU/mL。每个家系中均选 90 尾个体大小相近的健康尼罗罗非鱼平均分配到 3 个 0.5 m³ 的充气水箱中, 20 个家系共 1800 尾。每天正常投喂 2 次, 驯养 2 周后以半致死浓度的无乳链球菌菌液进行人工感染实验。前 24 h, 每 4 h 观察 1 次, 24 h 后, 每 2 h 观察 1 次, 观察时间持续一周, 记录鱼的死亡时间及数量, 并取死亡个体的尾鳍于无水乙醇中, -40 $^{\circ}$ C 保存, 用于 DNA 提取。2 周后, 如剩下的鱼稳定不死且鱼体状态良好, 可记录存活个体数量, 并

按上述方法对存活个体进行取样。本实验将人工感染无乳链球菌后 7 天内死亡的尼罗罗非鱼视为易感群体, 而始终存活的尼罗罗非鱼则视为抗病群体。

1.2.2 基因组 DNA 的提取 采用 TIANamp Marine Animals DNA Kit (TIANGEN)试剂盒提取尼罗罗非鱼的尾鳍基因组 DNA, 实验方法参照试剂盒说明书, 琼脂糖凝胶电泳检测其完整性, 并用超微量分光光度计检测其浓度和纯度, -20℃保存备用。

1.2.3 尼罗罗非鱼 *Ikaros* 基因 5'调控区的克隆及序列分析 采用染色体步移法克隆 *Ikaros* 基因的 5'调控区序列, 实验方法参照 Genome Walker™ Universal Kit (Clontech)试剂盒说明书。特异性引物在 *Ikaros* 基因的第一外显子中设计(表 1), 序列参照 *Ikaros* 基因的预测序列(登录号为: NC_022211),

本实验设计的所有引物均由生工生物工程(上海)股份有限公司合成。利用 *Dra* I、*Pvu* II、*Eco*R V 和 *Ssp* I 4 种酶分别对基因组 DNA 进行酶切, 构建相应的 DNA 文库, 并以之为模板, 通过巢式 PCR 克隆 *Ikaros* 基因 5'侧翼序列。琼脂糖凝胶电泳检测到 PCR 产物具有特异性条带后, 将不同片段大小的特异性条带分别胶回收纯化后与载体 pMD19T 连接, 转化 *E.coli* DH5α 感受态细胞, 阳性单克隆送至 Invitrogen 公司测序, 将测序结果放至 NCBI 数据库的 Blast 工具中进行检索, 以片段最长且包含 *Ikaros* 基因第一外显子部分序列的测序结果作为本实验所获得的 *Ikaros* 基因 5'调控区, 然后利用生物信息学软件对 *Ikaros* 基因的启动子和转录调控元件等进行预测和分析。

表 1 尼罗罗非鱼 *Ikaros* 基因引物信息
Tab. 1 The primer of *Ikaros* gene of *Oreochromis niloticus*

引物 primer	序列 (5'–3') sequence (5'–3')	用途 usage
IK-PRO-GSP1	CACTTGAAGGGCTTCTCCCCTGAGTGG	Genome Walking 第一轮扩增
IK-PRO-GSP2	ATGAAGTTCTTCCAACCTCTCCCGGCT	Genome Walking 第二轮扩增
IK-PRO-GSP3	GCTTTGTGCCGTGCCAATAAAAGCTGA	Genome Walking 第三轮扩增
IK-PRO-1F	ACACGGCCTCTCCTGAGATA	筛选 SNP 位点
IK-PRO-1R	CTGTCTGGCATCTGGGCTATTT	
IK-PRO-2F	TGTCTGCAGCCATGTTTCTAGT	筛选 SNP 位点
IK-PRO-2R	GGAATTGCACACCACCGTC	
IK-PRO-3F	AGTGCAGCTTGACGCTTAGA	筛选 SNP 位点
IK-PRO-3R	GAGATCCTGGAGGATAGTGGC	

1.2.4 尼罗罗非鱼 *Ikaros* 基因 5'调控区 SNPs 的筛查 采用 PCR 产物直接测序法筛查尼罗罗非鱼 *Ikaros* 基因 5'调控区的 SNPs。首先根据已获得的 *Ikaros* 基因 5'调控区序列设计 3 对引物(表 1)。然后在 P₀ 尼罗罗非鱼中随机挑选了 19 尾抗病个体和 20 尾易感个体的基因组 DNA 样品, 并以之为模板进行 PCR 扩增。将得到的 PCR 产物送至 Invitrogen 公司测序, 对测序结果进行多重比对分析和观察其峰图结果, 进而获得 *Ikaros* 基因 5'调控区的 SNPs 及其对应的基因型。

1.2.5 尼罗罗非鱼抗链球菌病相关 SNP 分子标记的筛选 对 P₀ 中筛到的 SNPs 与无乳链球菌抗性进行关联分析, 初步获得抗链球菌病相关的 SNP 分子标记。在 F₁ 中, 按各家系存活率和死亡

率的比例挑选了 136 尾抗病个体和 157 尾易感个体的基因组 DNA 样品, 采用 Snapshot 技术对 *Ikaros* 基因 5'调控区的 SNPs 进行基因分型并与无乳链球菌抗性进行关联分析, 最终获得抗链球菌病相关的 SNP 分子标记。基因分型实验由上海捷瑞生物工程有限公司完成, 遗传结构分析采用 Popgen 32 和 PIC 软件进行, 抗病相关性分析采用 SPSS 17.0 统计学软件中的卡方(χ^2)检验进行, 连锁不平衡分析采用 Haploview 4.2 软件进行, $P<0.05$ 为差异显著。

1.2.6 生物信息学分析 利用 DNAMAN 和 NCBI 数据库中 Blast 工具对基因测序结果进行比对分析; 利用 Neural Network Promoter Prediction (http://www.fruitfly.org/seq_tools/promoter.html)和 WWW

Promoter Scan (<https://www.bimas.cit.nih.gov/molbio/proscan/>)软件对启动子和转录调控元件等进行预测; CpG 岛利用 CpG Islands (http://www.bioinformatics.org/sms2/cpg_islands.html)软件分析; 转录因子结合位点(transcription factor binding site, TFBS)利用 Transfac (<http://gene-regulation.com/index2.html>)软件预测; 核苷酸序列的多重比对与分析利用 DNASTar 和 BioEdit 软件进行; 测序结果的峰图观察分析利用 Chromas 软件进行。

2 结果与分析

2.1 尼罗罗非鱼 *Ikaros* 基因 5'调控区序列的克隆

通过染色体步移技术 3 次 PCR 扩增获得了 4 条序列片段, 即大小分别约为 1000 bp、1500 bp、2000 bp 和 4000 bp 的特异性条带(图 1)。利用 NCBI 数据库的 Blast 工具对测序结果进行检索, 发现这 4 个特异性片段之间相互重叠的序列相同, 且均包含 *Ikaros* 基因第一外显子的部分序列 419 bp, 结果表明它们均为 *Ikaros* 基因的 5'侧翼区序列。本研究将以 *Ssp* I 酶切的 DNA 文库为模板 PCR 扩增获得的特异性序列作为 *Ikaros* 基因 5'调控区, 序列长度为 4178 bp(图 2)。

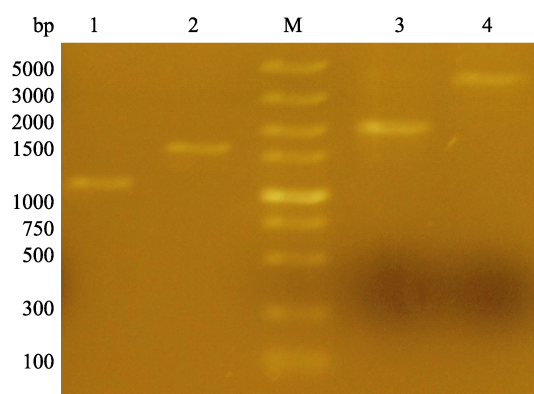


图 1 尼罗罗非鱼 *Ikaros* 基因 5'调控区序列的
第 3 次 PCR 扩增

M 表示 DL5000 DNA Marker, 1、2、3、4 分别代表以 *Dra* I、*Pvu* II、*Eco*R V 和 *Ssp* I 酶切的 DNA 文库为模板的巢式 PCR 产物。

Fig. 1 The third PCR amplification of *Ikaros* gene 5' regulatory region sequence from *Oreochromis niloticus*
M represents DL5000 DNA Marker. 1, 2, 3 and 4 represent nested PCR results from 4 kinds of DNA library which was digested by *Dra* I, *Pvu* II, *Eco*R V and *Ssp* I.

2.2 尼罗罗非鱼 *Ikaros* 基因 5'调控区序列的结构预测和分析

利用 Neural Network Promoter Prediction 和 WWW Promoter Scan 软件对 *Ikaros* 基因 5'调控区序列进行预测和分析, 结果发现该序列具有多个潜在的核心启动子和转录起始位点(transcriptional start site, TSS), 其中以位于起始密码子上游 931 bp 位置的 TSS 及其相应的核心启动子的预测分值最高(score=0.95), 因此将其默认为 RNA 聚合酶与 *Ikaros* 基因结合并开始转录的位置。为方便对 *Ikaros* 基因 5'调控区序列进行分析, 在 DNA 序列中将 TSS 定为 1, 其上游的 DNA 序列用负数表示, 下游的 DNA 序列用正数表示, 结果如图 2 所示。*Ikaros* 基因核心启动子位于-57~48 bp 的序列区间, 而一些重要的启动子元件如 TATA 框、CCAAT 框和八聚体元件等分别位于 TSS 上游-28 bp、-169 bp 和-41 bp 的位置上。利用 Transfac 软件对 *Ikaros* 基因的 5'调控区的各种 TFBS 进行分析, 结果表明在 5'调控区-2200~1200 bp 的序列中存在丰富的 TFBS, 包括 GATA-1、GATA-2、GATA-3、Homeobox、Oct-1、CDP CR3+HD、AP-1、AREB6、C/EBP、TFIID、TGIF、Nkx2-5、Lmo2、Evi-1、Freac-7、Brn-2、YY1、Pax-4、NF-E2、v-ErbA、HNF-4、XFD-1、RREB-1、v-Myb、CHOP-C/EBP alpha、HFH-3、c-Myb、COMP1、Elk-1、c-Ets-1(p54)、SP1、NF-Y、FOXJ2 共 33 种 TFBS(图 2)。此外, 利用 CpG Islands 软件对 *Ikaros* 基因的 5'调控区的 CpG 岛进行分析, 结果发现该序列有 2 个 CpG 岛, 分别分布在-695~-450 bp 的启动子区域和 1048~1351 bp 的第一外显子区域中(图 2)。

2.3 尼罗罗非鱼 *Ikaros* 基因 5'调控区 SNPs 的筛查

本研究利用 PCR 产物直接测序法对 P_0 尼罗罗非鱼 *Ikaros* 基因 5'调控区的 SNPs 进行筛查, 结果发现了该序列有 5 个 SNPs, 分别为 SNP1(g.562, G>A)、SNP2(g.217, G>T)、SNP3(g.-53, C>T)、SNP4(g.-220, T>C)和 SNP5(g.-579, T>C)。结合 *Ikaros* 基因 5'调控区预测结果, 发现这 5 个 SNPs 主要分布在 *Ikaros* 基因的核心启动子和各种调控元件中, 如 SNP1 位于 c-Ets-1(p54)中, SNP2 位于 c-Myb 中, SNP3 位于 v-Myb 中, SNP4 位于 RREB-1 中, SNP5 位于 Barbie box 元件中(图 2)。

图2 尼罗罗非鱼 *Ikaros* 基因 5'调控区序列

Fig. 2 5' regulatory region sequence of *Ikaros* gene from *Oreochromis niloticus*

The transcriptional start site of *Ikaros* gene are denoted by bold italicized letters. Lowercase letters represent core promoter. The regulatory elements and transcription factor binding sites of 5' regulatory region sequence are underlined. The CpG islands are characterized in gray shading. The sites of SNP are characterized in white letters and black background. The initiation codon (ATG) of the *Ikaros* gene is characterized in box.

2.4 F₁ 尼罗罗非鱼抗病选育结果及其抗病群体和易感群体的构建

对 F₁ 尼罗罗非鱼的 20 个家系共 1800 尾鱼进行抗病选育, 感染无乳链球菌后各家系的存活率情况如图 3 所示。其中存活率在 50% 以上的家系有 4 个, 分别为 N5、N6、N14 和 N15 家系, 存活率在 30%~50% 的家系有 7 个, 分别为 N2、N3、N4、N10、N11、N13 和 N16 家系, 而其余家系的

存活率均低于 30%。由此可以看出, 尼罗罗非鱼不同家系在无乳链球菌感染后存活率存在着较大差异, 表明了这些家系存在着非常大的遗传选育潜力。此外, 在抗病选育过程中, 将 7 天内死亡的尼罗罗非鱼视为易感群体, 而始终存活的尼罗罗非鱼则视为抗病群体, 同时分别收集它们的尾鳍材料用于基因组 DNA 的提取, 这为筛选出抗链球菌病相关 SNPs 提供了基础材料。

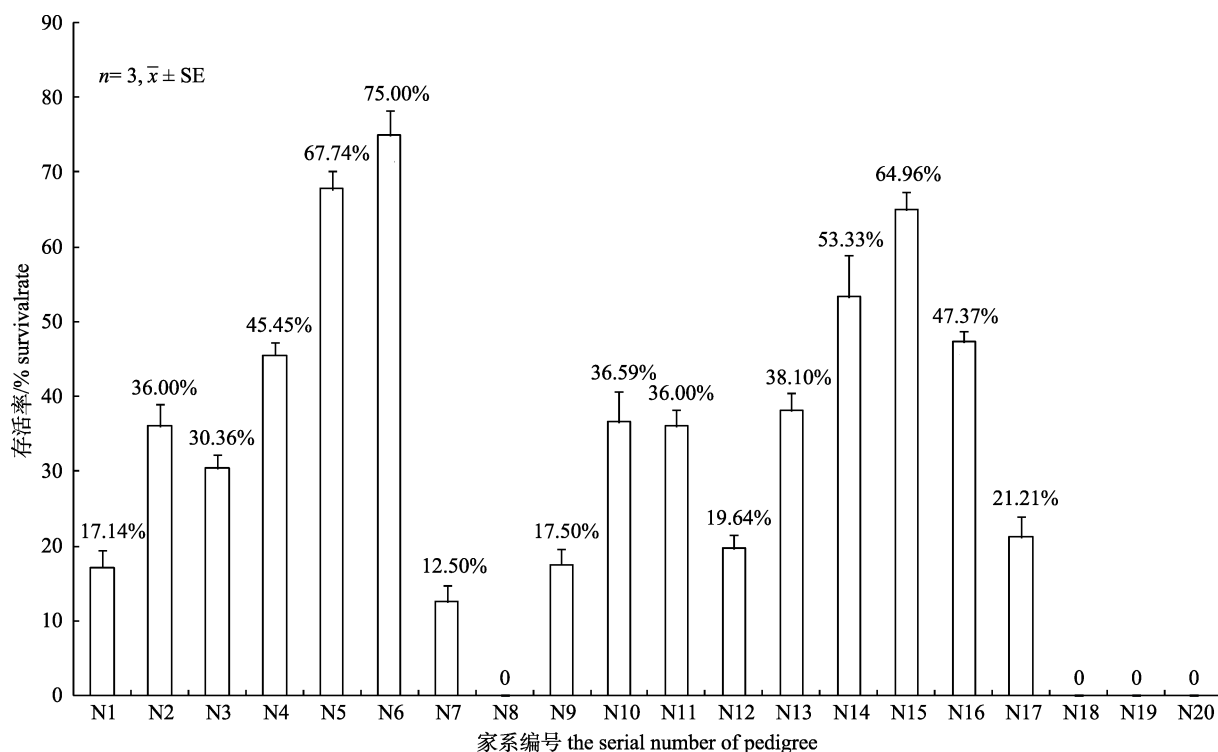


图 3 F₁ 20 个家系尼罗罗非鱼经人工感染无乳链球菌后的存活情况

Fig. 3 The survival rate of *Oreochromis niloticus* within F₁ generation of 20 families after challenged with *Streptococcus agalactiae*

2.5 F₁ 尼罗罗非鱼抗病群体和易感群体的 SNPs 基因分型及其遗传结构分析

利用 Snapshot 技术对 F₁ 尼罗罗非鱼易感群体和抗病群体进行相应 SNPs 的基因分型, 获得有效结果 293 个, 采用 Popgen 32 和 PIC 软件对两个群体遗传多样性参数进行分析, 结果如表 2 所示。在抗病群体中, PIC 值的范围为 0.0872~0.3747, 其中 SNP4 属于中度多态(0.25<PIC<0.5), 其余 SNPs 属于低度多态(PIC<0.25)。在易感群体中, PIC 值的范围为 0.1865~0.3484, 其中 SNP3 和 SNP4 属于中度多态(0.25<PIC<0.5), 其余 SNPs 属于低度多态(PIC<0.25)。χ² 检验表明, 在抗病群体

中, 除 SNP3 不符合 Hardy-Weinberg 平衡外(P<0.05), 其余 SNPs 均处于 Hardy-Weinberg 平衡状态(P>0.05); 在易感群体中, 除 SNP1 和 SNP3 不符合 Hardy-Weinberg 平衡外(P<0.05), 其余 SNPs 均处于 Hardy-Weinberg 平衡状态(P>0.05)。

2.6 尼罗罗非鱼 SNPs 的连锁不平衡分析及与无乳链球菌病抗性的关联分析

采用 SPSS17.0 软件分析 *Ikaros* 基因 5'调控区的 SNPs 在易感群体和抗病群体中的基因型分布频率、等位基因频率以及 χ² 检验其与无乳链球菌病抗性的相关性。结果表明, 在 P₀ 中, 仅 SNP4 的基因型频率与无乳链球菌抗性显著相关(P<

0.05)(表 3), 而在 F_1 中, SNP2、SNP3、SNP4 和 SNP5 的基因型分布频率和等位基因的频率在抗病群体和易感群体之间均存在显著差异性 ($P<0.05$), 表明这 4 个 SNPs 与抗无乳链球菌病性状显著相关(表 4)。利用 Haploview 4.2 软件中的 Four Gamete Rule 计算方法对 *Ikaros* 基因 5'调控区的 SNPs 进行连锁不平衡分析, 结果如图 4 所示。5 个 SNPs 可形成 1 个连锁不平衡单倍块和 5 种单倍型, 其中 GGCTT 单倍型与无乳链球菌抗病性状显著相关($P<0.05$), GGTCT 和 GTCCC 单倍型与无乳链球菌易感性状显著相关($P<0.05$)(表 5)。此外, 本实验还发现 SNP2 和 SNP5 处于完全

连锁状态($r^2=1$, LOD=57.25, $D'=1$), 可作为罗非鱼抗链球菌病遗传育种的标签 SNP。

3 讨论

3.1 尼罗罗非鱼 *Ikaros* 基因 5'调控区结构与功能

Ikaros 基因编码的 *Ikaros* 是免疫相关基因中必不可少的转录调控因子, 对于淋巴细胞的正确发育及其免疫功能的发挥具有重要的作用^[10-11]。位于 5'端的启动子和调控元件是各种转录因子识别与结合的部位, 在转录前起始复合物的装配、基因的特异性表达以及基因的转录效率等方面均发挥重要的调控作用^[23]。本研究成功地克隆了尼

表 2 5 个 SNPs 在 F_1 尼罗罗非鱼抗病群体和易感群体中的遗传信息
Tab. 2 The genetic polymorphism information of 5 SNPs in susceptible groups and resistant groups from *Oreochromis niloticus* F_1

名称 name	抗病群体 resistant group					易感群体 susceptible group				
	H_o	H_e	PIC	N_e	H-W	H_o	H_e	PIC	N_e	H-W
SNP1(g.562, G>A)	0.2113	0.2326	0.1948	0.2318	$P=0.2684, \chi^2=1.2249$	0.1456	0.2074	0.1865	0.2068	$P=0.0001, \chi^2=14.3974$
SNP2(g.217, G>T)	0.1056	0.1129	0.0872	0.1126	$P=0.4270, \chi^2=0.6309$	0.1835	0.1672	0.1531	0.1667	$P=0.2127, \chi^2=1.5531$
SNP3(g.-53, C>T)	0.1761	0.2274	0.2018	0.4980	$P=0.0064, \chi^2=7.4217$	0.2215	0.3698	0.3015	0.3686	$P=0.0000, \chi^2=25.6743$
SNP4(g.-220, T>C)	0.4859	0.4998	0.3747	0.2266	$P=0.7406, \chi^2=0.1096$	0.3924	0.4514	0.3484	0.4499	$P=0.0993, \chi^2=2.7163$
SNP5(g.-579, T>C)	0.1056	0.1129	0.0872	0.1126	$P=0.4270, \chi^2=0.6309$	0.1835	0.1672	0.1531	0.1667	$P=0.2127, \chi^2=1.5531$

注: H_o 代表观测杂合度, H_e 代表期望杂合度, PIC 代表多态信息含量, N_e 代表有效等位基因数, H-W 代表 Hardy-Weinberg 平衡, $P<0.05$ 表示差异性显著。
Note: H_o represents observed heterozygosity; H_e represents expected heterozygosity; PIC represents polymorphism information content; N_e represents effective number of alleles; H-W represents Hardy-Weinberg equilibrium; $P<0.05$ is considered to be statistically significant.

表 3 P_0 尼罗罗非鱼 *Ikaros* 基因在抗病群体和易感群体的 SNPs 统计分析
Tab. 3 The SNPs distributions of *Oreochromis niloticus Ikaros* gene in susceptible groups and resistant groups from P_0

名称 name	基因型(N) genotype	基因型频率 genotype frequency		χ^2 (P)	等位基因 allele	等位基因频率 allele frequency		χ^2 (P)	置信区间 confidence interval
		抗病群体 resistant group	易感群体 susceptible group			抗病群体 resistant group	易感群体 susceptible group		
SNP1 (g.562, G>A)	AA(5)	3(15.8%)	2(10.0%)	0.326 (0.850)	A	10(26.3%)	9(22.5%)	0.154 (0.695)	1.230(0.437, -3.464)
	AG(9)	4(22.1%)	5(25.0%)		G	28(73.7%)	31(77.5%)		
	GG(25)	12(63.1%)	13(65.0%)						
SNP2 (g.217, G>T)	GT(3)	0(0)	3(15.0%)	3.087 (0.079)	T	0(0)	3(7.5%)	2.964 (0.085)	1.081(0.990, -1.181)
	GG(36)	19(100.0%)	17(85.0%)		G	38(100.0%)	37(92.5%)		
SNP3 (g.-53, C>T)	TT(1)	1(5.3%)	0(0)	1.255 (0.534)	T	6(15.8%)	3(7.5%)	1.312 (0.252)	2.313(0.535, -10.002)
	CT(7)	4(21.0%)	3(15.0%)		C	32(84.2%)	37(92.5%)		
	CC(30)	14(73.7%)	17(85.0%)						
SNP4 (g.-220, T>C)	CC(10)	6(31.6%)	4(20.0%)	10.391 (0.006)	T	22(57.9%)	18(45.0%)	1.297 (0.255)	1.681(0.686, -4.117)
	CT(18)	4(20.0%)	14(70.0%)		C	16(42.1%)	22(55.0%)		
	TT (11)	9(50.0%)	2(10.0%)						
SNP5 (g.-579, T>C)	GT(3)	0(0)	3(15.0%)	3.087 (0.079)	C	0(0)	3(7.5%)	2.964 (0.085)	1.081(0.990, -1.181)
	GG(36)	19(100.0%)	17(85.0%)		T	38(100.0%)	37(92.5%)		

注: $P<0.05$ 认为差异性显著。
Note: $P<0.05$ is considered to be statistically significant.

表 4 F₁ 尼罗罗非鱼 *Ikaros* 基因在抗病群体和易感群体的 SNPs 统计分析
Tab. 4 The SNPs distributions of *Oreochromis niloticus Ikaros* gene in susceptible groups and resistant groups from F₁

名称 name	基因型(N) genotype	基因型频率 genotype frequency		χ^2 (P)	等位基因 allele	等位基因频率 allele frequency		χ^2 (P)	置信区间 confidence interval
		抗病群体 resistant group	易感群体 susceptible group			抗病群体 resistant group	易感群体 susceptible group		
SNP1 (g.562, G>A)	AA(11)	4(2.9%)	7(4.5%)	1.397 (0.497)	A	34(12.5%)	37(11.8%)	0.070 (0.791)	1.069(0.651, -1.758)
	AG(53)	26(19.1%)	23(14.6%)		G	238(87.5%)	277(88.2%)		
	GG(236)	106(77.9%)	127(80.9%)						
SNP2 (g.217, G>T)	TT(1)	1(0.7%)	0(0)	7.698 (0.021)	T	13(4.8%)	29(9.2%)	4.350 (0.037)	2.027(1.032, -3.984)
	GT(40)	11(8.1%)	29(18.5%)		G	259(95.2%)	285(90.8%)		
	GG(252)	124(91.2%)	128(81.5%)						
SNP3 (g.-53, C>T)	TT(27)	6(4.4%)	21(13.4%)	8.671 (0.013)	T	37(13.1%)	77(24.5%)	11.090 (0.001)	0.485(0.315, -0.746)
	CT(60)	25(18.4%)	35(22.3%)		C	245(86.9%)	237(75.5%)		
	CC(206)	105(77.2%)	101(64.3%)						
SNP4 (g.-220, T>C)	CC(110)	37(27.2%)	73(46.5%)	12.093 (0.002)	T	141(51.8%)	207(65.9%)	12.316 (0.000)	0.548(0.393, -0764)
	CT(128)	67(49.3%)	61(38.9%)		C	131(48.2%)	107(34.1%)		
	TT (55)	32(23.5%)	23(14.6%)						
SNP5 (g.-579, T>C)	CC(1)	1(0.7%)	0(0)	7.698 (0.021)	C	13(4.8%)	29(9.2%)	4.350 (0.037)	2.027(1.032, -3.984)
	CT(40)	11(8.1%)	29(18.5%)		T	259(95.2%)	285(90.8%)		
	TT(252)	124(91.2%)	128(81.5%)						

注: $P<0.05$ 认为差异性显著。
Note: $P<0.05$ is considered to be statistically significant.

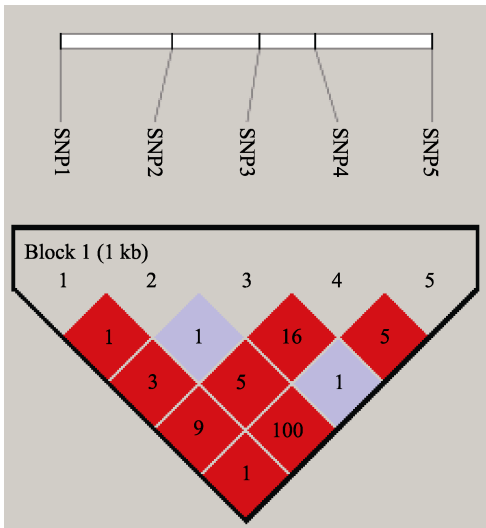


图 4 尼罗罗非鱼 *Ikaros* 基因 SNPs 连锁不平衡分析
标记之间显示配对 LD 值(r^2), 阴影
代表标记之间所处的连锁状态。

Fig. 4 Linkage disequilibrium map showing the pair-wise LD between the SNPs of *Ikaros* gene in *Oreochromis niloticus*
The values inside the squares indicate the LD values (r^2) of the two SNPs. The chain states between the marks were indicated by shadow.

罗罗非鱼 *Ikaros* 基因 5'调控区序列, 并利用生物信息学软件对该序列的结构进行预测和分析。预

测结果显示 *Ikaros* 基因 5'调控区具有多个潜在的核心启动子和 TSS, 但由于 Neural Network Promoter Prediction 和 WWW Promoter Scan 软件是根据 TFBS 的分布密度和 TATA 盒的权重矩阵进行预测的, 因此本研究选择评分最高的预测结果作为 *Ikaros* 基因的启动子和 TSS 具有较高的可信度, 而且所选的启动子还包含 TATA 框、CCAAT 框和起始子等重要调控元件, 是典型的真核生物类型 II 启动子^[24]。

Ikaros 基因 5'调控区-2200~1200 bp 序列中存在丰富的 TFBS, 而且大部分与免疫基因的转录密切相关, 其中一些多次出现且评分较高的 TFBS, 如 GATA-1、Oct-1、AP-1 和 Homeobox 等, 可能对 *Ikaros* 基因的精确表达起着重要的调控作用^[25]。如 GATA-1 是一种非常重要的、且研究十分广泛的造血转录因子, 有研究表明几乎所有的红系特异性基因都存在与 GATA-1 结合的 DNA 基序[T/A(GATA)A/G], 而 GATA-1 通过作用于这些部位来调控相关基因的转录, 对红系和巨核系等造血细胞的分化发育具有重要作用^[26]。Oct-1 是一种组成性转录激活因子, 普遍存在于各种细胞中,

表 5 尼罗罗非鱼 *Ikaros* 基因单倍型与链球菌病抗性的关联分析
Tab. 5 Association of *Ikaros* haplotypes with resistance to *Streptococcus agalactiae* for *Oreochromis niloticus*

单倍块 haploblock	单倍型 haplotype	分布频率/% frequency	易感群体频率/% frequency in susceptible group	抗病群体频率/% frequency in resistant group	χ^2	<i>P</i>
Block 1	GGCTT	40.6	34.1	48.2	11.989	0.0005
	GGCCT	20.6	20.4	21.0	0.029	0.8641
	GGTCT	19.5	20.5	13.6	11.090	0.0009
	AGCCT	12.1	11.8	12.5	0.070	0.7909
	GTCCC	7.2	9.2	4.7	4.350	0.0370

注: $P<0.05$ 表示差异性显著。
Note: $P<0.05$ is considered to be statistically significant.

研究表明 Oct-1 能够结合启动子中的八聚体基序 (ATGCAAT), 并在相关基因的转录过程发挥重要调控作用^[27]。Homeobox 蛋白是一类与胚胎发育相关的转录因子, 通过与 DNA 基序特异性结合来激活或抑制一系列相关下游基因的表达^[28]。AP-1 是细胞内的一个转录激活因子, 其参与了多种生理和病理过程的调控, 近期研究表明, 它对细胞的增殖、分化、转化、凋亡以及免疫系统活化具有重要影响^[29]。

此外, 在 *Ikaros* 基因 5'调控区中还发现了 2 个 CpG 岛, 分别位于启动子区域和第一外显子区域中, 这与大多数关于基因 5'端 CpG 岛位置分布的研究相一致^[30]。5'调控区的 CpG 岛甲基化状态对于基因的表达调控具有重要作用, 如 CpG 岛高甲基化会影响基因的转录机制, 主要包括阻碍转录因子结合、募集其他蛋白引起染色质沉默和影响核小体定位等, 进而导致基因表达下降或沉默^[31]。

3.2 尼罗罗非鱼 *Ikaros* 基因 5'调控区 SNP 多态性与抗病的相关性

SNP 是指物种基因组中由单个核苷酸突变所引起的 DNA 序列多态性, 往往与生物体机能的改变和疾病的发生有关^[4]。作为当前最理想的 DNA 分子标记, SNP 已经广泛应用于罗非鱼抗病育种进程中, 目前已有一些研究进行了免疫相关基因的 SNPs 与抗病性状之间的关联分析, 如高风英等^[32]在主要组织相容性复合体(major histocompatibility complex II B, *MHC II B*)基因找到了与链球菌病抗性相关的 SNP 遗传标记。Fu 等^[33]在肥大细胞蛋白酶(mast cell protease 8, *MCP-8*)基因中共找到 5 个 SNPs, 其中有 3 个 SNPs 与无乳

链球菌抗性显著相关。Shen 等^[34]在丝氨酸蛋白酶 *Duodenase-1* 基因中共找到了 7 个 SNPs, 其中有 4 个 SNPs 与无乳链球菌抗性显著相关。曹建萌等^[35]在补体 C9(*OnC9*)基因中共找到了 22 个 SNPs, 抗病关联分析发现有 2 个 SNPs 与无乳链球菌抗性显著相关, 连锁不平衡分析发现有 12 种单倍型, 其中有 2 种单倍型与无乳链球菌易感性显著相关。这些 SNP 资源均可作为重要分子遗传标记用于辅助罗非鱼抗病选育, 为加速罗非鱼的遗传改良进程并培育出罗非鱼优良抗病新品种奠定了基础。

目前认为, 人类 *Ikaros* 基因的突变与 SLE 的易感性和复杂表型相关, 以及与恶性肿瘤(白血病、卵巢癌、肺癌等)等疾病的发生、发展密切相关^[14-17]。Hu 等^[36]采用 GWAS 技术对汉族 SLE 患者进行研究, 发现大部分患者的 *Ikaros* 基因会发生可遗传变异, 而这些遗传变异体与 SLE 易感性的关联程度已经达到全基因组显著水平($p_{\text{comb}}=2.75\times10^{-23}$)。Westra 等^[37]利用 SNPs (rs12718597) 对多种血液指标进行关联分析, 结果发现 SNPs (rs12718597)与平均红细胞容积有关, 并与 31 个基因表达上调和 19 个基因表达下调成反式关联, 表明该突变位点可能与白血病等疾病的发生相关。此外, 解丽然^[38]利用 *Ikaros* 基因的 2 个 SNPs (rs4132601、rs10272724)对自身免疫性甲状腺疾病(autoimmune thyroid disease, AITD)进行关联分析, 发现 SNPs (rs10272724)的 T 碱基和 TT/TC 基因型以及 SNPs (rs4132601)的 T 碱基和 TT 基因型与 AITD 的发病具有显著相关性($P<0.01$)。由此可看出, *Ikaros* 基因作为重要免疫相关基因, 其多态性与众多疾病密切相关。但是关于 *Ikaros* 基因的多

态性与疾病的关联分析在水产动物中却未见相关报道。本团队在前期研究中已发现无乳链球菌的感染可引起尼罗罗非鱼多个免疫相关组织中 *Ikaros* 基因的表达量发生变化, 表明 *Ikaros* 基因参与调控尼罗罗非鱼抵御无乳链球菌的免疫应答反应^[21]。因此本研究选择 *Ikaros* 基因作为抗病性相关 SNPs 筛选的候选基因, 以期在 *Ikaros* 基因中筛选到抗链球菌病相关的 SNP 分子标记。

本研究从 P_0 尼罗罗非鱼的 *Ikaros* 基因 5'调控区中筛选到 5 个 SNPs, 而且其主要分布在核心启动子和各种调控元件中(图 2)。Xie 等^[39]研究表明, 基因 5'端调控区的启动子和调控元件通常含有较多的 SNPs, 而这些 SNPs 往往会改变基因的转录活性和影响基因的正常表达调控。这可能是不同基因型的尼罗罗非鱼抗病力表现出差异的原因之一, 这为本研究在 *Ikaros* 基因 5'调控区中筛选到抗病相关的 SNPs 提供理论依据。在 F_1 抗病群体和易感群体相应 SNPs 的基因分型中, 发现大部分 SNPs 处于低度多态性($PIC < 0.25$), 而且在抗病群体中 SNP3 偏离 Hardy-Weinberg 平衡状态($P < 0.05$), 在易感群体中 SNP1 和 SNP3 偏离 Hardy-Weinberg 平衡状态($P < 0.05$)(表 3)。这可能是因为本研究使用的尼罗罗非鱼材料来自多代选育的抗病家系, 抗病相关 SNPs 的基因型或等位基因在选育过程被定向选择, 导致 SNPs 的多样性发生一定程度的衰退。Liu 等^[40]研究表明, 对于不同基因其受环境等选择因素的影响程度也存在差异, 这可能是 SNPs 完全偏离 Hardy-Weinberg 平衡现象的部分原因。另外, 曹建萌等^[35]研究表明, 在人工构建的抗病群体和易感群体中, 部分 SNPs 出现不符合 Hardy-Weinberg 平衡现象可作为与抗病性状相关的初步判断。

Ikaros 基因 5'调控区 SNPs 与抗病性状进行关联分析中, 在 P_0 中仅 SNP4 与无乳链球菌病抗性显著相关($P < 0.05$)(表 3), 而在 F_1 中 SNP2、SNP3、SNP4 和 SNP5 均与无乳链球菌病抗性显著相关($P < 0.05$)(表 4)。 P_0 和 F_1 的结果出现差异可能是因为 P_0 中使用样本数量较少, 导致部分抗病相关 SNPs 在 P_0 中没有表现出来, 而在 F_1 中使用大样本验证, 其获得的结果应更具有代表性。另外, 通

过连锁不平衡分析, 发现了与无乳链球菌病抗性显著相关($P < 0.05$)的 GGCTT 单倍型和与无乳链球菌易感性状显著相关($P < 0.05$)的 GGTCT 和 GTCCC 单倍型(表 3-5)。本研究结果表明, 在 *Ikaros* 基因 5'调控区筛选到的 4 个抗病 SNPs 和 GGCTT 单倍型均可作为尼罗罗非鱼分子育种的候选分子标记, 这为实现利用分子标记辅助罗非鱼抗病选育奠定了基础。

4 展望

罗非鱼是中国最重要的养殖经济鱼类之一, 尽管科研工作者已经开展了大量的相关研究工作, 但是对罗非鱼的抗病性研究却处于起步阶段, 尤其是在分子水平上阐述其抗病机制。本研究在尼罗罗非鱼 *Ikaros* 基因 5'调控区中筛选到抗病相关的 SNPs 和基因型, 为尼罗罗非鱼的分子育种提供了有效分子标记。但是要想获得更加稳定和可靠的分子标记, 应继续在抗病选育群体的后代或其他家系中作进一步扩大验证。目前, 关于罗非鱼免疫相关基因与抗病性关联分析的研究并不多, 而且能真正应用到罗非鱼抗病选育中的分子标记更是寥寥无几, 这对于罗非鱼的 MAS 是不够的。因此, 应继续加强在免疫相关基因中有效分子标记的筛选, 并构建精密的遗传连锁图谱, 这样才能提高育种的准确性, 为培育出罗非鱼优良抗病新品种奠定基础。

参考文献:

- [1] Xiao J, Gan X, Luo Y J. The progress of tilapia breeding[J]. Journal of Hunan University of Science & Technology (Natural Science Edition), 2014, 29(1): 106-112. [肖俊, 甘西, 罗永巨. 罗非鱼育种研究进展[J]. 湖南科技大学学报(自然科学版), 2014, 29(1): 106-112.]
- [2] Lu M X. Review of research on streptococcosis in tilapia[J]. South China Fisheries Science, 2010, 6(1): 75-79. [卢迈新. 罗非鱼链球菌病研究进展[J]. 南方水产科学, 2010, 6(1): 75-79.]
- [3] Lu M X, Li J, Ye X, et al. Identification and characterizations of *Streptococcus agalactiae* isolated from tilapia cultured in Guangdong and Hainan Provinces[J]. Microbiology China, 2010, 37(5): 766-774. [卢迈新, 黎炯, 叶星, 等. 广东与海南养殖罗非鱼无乳链球菌的分离、鉴定与特性分析

- [J]. 微生物学通报, 2010, 37(5): 766-774.]
- [4] Liu Z J, Cordes J F. DNA marker technologies and their applications in aquaculture genetics[J]. Aquaculture, 2007, 238(1-4): 1-37.
- [5] Rasal K D, Chakrapani V, Pandey A K, et al. Status and future perspectives of single nucleotide polymorphisms (SNPs) markers in farmed fishes: Way ahead using next generation sequencing[J]. Gene Reports, 2017, 6: 81-86.
- [6] Kongchum P, Palti Y, Hallerman E M, et al. SNP discovery and development of genetic markers for mapping innate immune response genes in common carp (*Cyprinus carpio*)[J]. Fish & Shellfish Immunology, 2010, 29(2): 356-361.
- [7] Zhao X L, Guo L L, Zhang Y L, et al. SNPs of hemocyanin C-terminal fragment in shrimp *Litopenaeus vannamei*[J]. FEBS Letters, 2012, 586(4): 403-410.
- [8] Fu R X, Zhou N Y, Liu J X, et al. The polymorphism of serine protease I-1 in Japanese flounder (*Paralichthys olivaceus*)[J]. Progress in Fishery Sciences, 2015, 36(5): 55-62. [付瑞雪, 周纳宇, 刘金相, 等. 牙鲆(*Paralichthys olivaceus*)丝氨酸蛋白酶 I-1 基因多态性[J]. 渔业科学进展, 2015, 36(5): 55-62.]
- [9] Xing H F, Gao F T, Zhang Y Z, et al. Molecular cloning, expression and SNP screening of natural resistance-associated macrophage protein (*Nramp*) gene cDNA from half smooth tongue sole (*Cynoglossus semilaevis*)[J]. Progress in Fishery Sciences, 2016, 37(4): 116-127. [邢贺飞, 高峰涛, 张永珍, 等. 半滑舌鲷(*Cynoglossus semilaevis*) *Nramp* 基因克隆与表达分析及 SNP 筛选[J]. 渔业科学进展, 2016, 37(4): 116-127.]
- [10] Fan Y, Lu D. The Ikaros family of zinc-finger proteins[J]. Acta Pharmaceutica Sinica B, 2016, 6(6): 513-521.
- [11] John L B, Ward A C. The *Ikaros* gene family: Transcriptional regulators of hematopoiesis and immunity[J]. Molecular Immunology, 2011, 48(9-10): 1272-1278.
- [12] Georgopoulos K, Moore D D, Derfler B. Ikaros, an early lymphoid-specific transcription factor and a putative mediator for T cell commitment[J]. Science, 1992, 258(5083): 808-812.
- [13] Nietfeld W, Meyerhans A. Cloning and sequencing of *hIk-1*, a cDNA encoding a human homologue of mouse Ikaros/LyF-1[J]. Immunology Letters, 1996, 49(1-2): 139-141.
- [14] Marçais A, Jeannet R, Hernandez L, et al. Genetic inactivation of Ikaros is a rare event in human T-ALL[J]. Leukemia Research, 2010, 34(4): 426-429.
- [15] Dörge P, Meissner B, Zimmermann M, et al. *IKZF1* deletion is an independent predictor of outcome in pediatric acute lymphoblastic leukemia treated according to the ALL-BFM 2000 protocol[J]. Haematologica, 2011, 98(3): 428-432.
- [16] Wen Q, Wu M R, He Y Z, et al. Correlation of *IKZF1* gene with malignant tumors and systemic lupus erythematosus[J]. Journal of Experimental Hematology, 2015, 23(2): 591-595. [温琦, 吴美容, 何颖芝, 等. *IKZF1* 基因与恶性肿瘤及系统性红斑狼疮的相关性研究进展[J]. 中国实验血液学杂志, 2015, 23(2): 591-595.]
- [17] Wen L L, Hu S J, Cui Y, et al. A system review on the association between the critical role of *IKZF1* and the pathogenesis of systemic lupus erythematosus [J]. The Chinese Journal of Dermatovenereology, 2013(6): 629-632. [闻雷雷, 胡素娟, 崔勇. *IKZF1* 基因与 SLE 发病机制相关性的研究进展[J]. 中国皮肤性病学杂志, 2013(6): 629-632.]
- [18] Hansen J D. O1 Characterization of Ikaros in rainbow trout[J]. Developmental & Comparative Immunology, 1997, 21(2): 199.
- [19] Willett C E, Kawasaki H, Amemiya C T, et al. Ikaros expression as a marker for lymphoid progenitors during zebrafish development[J]. Developmental Dynamics, 2001, 222(4): 694-698.
- [20] Li F, Li H, Zhang S. Ikaros expression in tongue sole macrophages: a marker for lipopolysaccharide and lipoteichoic acid induced inflammatory responses[J]. Molecular Biology Reports, 2011, 38(4): 2273-2279.
- [21] Chen K P. Molecular cloning and expression analysis of *Ikaros* gene from *Oreochromis niloticus* and screening of their SNP resistance to *Streptococcus agalactiae*[D]. Shanghai: Shanghai Ocean University, 2017. [陈昆平. 尼罗罗非鱼 *Ikaros* 基因克隆、组织表达及抗无乳链球菌相关 SNP 位点筛选[D]. 上海: 上海海洋大学, 2017.]
- [22] Dong Z J, Liang Z Y, Liu J Q, et al. Genetic diversity analysis in a newly introduced population of GIFT strain Nile tilapia (*Oreochromis niloticus*) [J]. Chinese Journal of Zoology, 2010, 45(5): 129-135. [董在杰, 梁政远, 刘介奇, 等. 我国新引进吉富品系尼罗罗非鱼群体的遗传多样性分析[J]. 动物学杂志, 2010, 45(5): 129-135.]
- [23] Yu F X, Zhu Z W, Chen X Y, et al. Eukaryotes RNA polymerase II core promoters[J]. Chemistry of Life, 2011, 31(6): 822-826. [于福先, 朱志伟, 陈晓宇, 等. 真核生物 RNA 聚合酶 II 核心启动子[J]. 生命的化学, 2011, 31(6): 822-826.]
- [24] Sandelin A, Carninci P, Lenhard B, et al. Mammalian RNA polymerase II core promoters: insights from genome-wide

- studies[J]. Nature Reviews Genetics, 2007, 8(6): 424-436.
- [25] Songdej N, Rao A K. Hematopoietic transcription factor mutations and inherited platelet dysfunction[J]. F1000Prime Reports, 2015, 7: 66.
- [26] Yu H C, Wu J, Zhai P F, et al. Construction of *GATA-1* knockout K562 cell line by CRISPR/Cas9 technology[J]. Chinese Journal of Cell Biology, 2016, 38(5): 593-602. [于海川, 吴娇, 翟鹏飞, 等. 基于 CRISPR/Cas9 技术的 *GATA-1* 基因敲除 K562 细胞系的构建[J]. 中国细胞生物学学报, 2016, 38(5): 593-602.]
- [27] Voleti B, Hammond Jr D J, Thirumalai A, et al. Oct-1 acts as a transcriptional repressor on the C-reactive protein promoter[J]. Molecular Immunology, 2012, 52(3-4): 242-248.
- [28] Rodrigues M F, Esteves C M, Xavier F C, et al. Methylation status of *homeobox* genes in common human cancers[J]. Genomics, 2016, 108(5-6): 185-193.
- [29] Zhou Q. Research progress of AP-1 in transcriptional regulation[J]. Journal of Medical Molecular Biology, 2001, 23(2): 65-67. [周钦. AP-1 在基因转录调控中的研究进展[J]. 医学分子生物学杂志, 2001, 23(2): 65-67.]
- [30] Tsellou E, Troungos C, Moschovi M, et al. Hypermethylation of CpG islands in the promoter region of the *p15INK4B* gene in childhood acute leukaemia[J]. European Journal of Cancer, 2005, 41(4): 584-589.
- [31] Takeshima H, Yamashita S, Shimazu T, et al. Effects of genome architecture and epigenetic factors on susceptibility of promoter CpG islands to aberrant DNA methylation induction[J]. Genomics, 2011, 98(3): 182-188.
- [32] Gao F Y, Pang J C, Lu M X, et al. Major histocompatibility complex *MHC II B* gene and its association with resistance/susceptibility to *Streptococcus agalactiae* in Nile tilapia (*Oreochromis niloticus*)[J]. Chinese Agricultural Science Bulletin, 2014, 30(2): 76-83. [高风英, 庞纪彩, 卢迈新, 等. 尼罗罗非鱼 *MHC II B* 基因多态性及其与链球菌病抗性的关系[J]. 中国农学通报, 2014, 30(2): 76-83.]
- [33] Fu G H, Wan Z Y, Xia J H, et al. The *MCP-8* gene and its possible association with resistance to *Streptococcus agalactiae* in tilapia[J]. Fish & Shellfish Immunology, 2014, 40(1): 331-336.
- [34] Shen Y, Fu G H, Liu F, et al. Characterization of the *duodenase-1* gene and its associations with resistance to *Streptococcus agalactiae* in hybrid tilapia (*Oreochromis spp.*)[J]. Fish & Shellfish Immunology, 2015, 45(2): 717-724.
- [35] Cao J M, Hu X X, Lu M X, et al. SNPs of *C9* gene and their association with resistance to the infection of *streptococcus agalactiae* in *Oreochromis niloticus*[J]. Journal of Agricultural Biotechnology, 2017, 25(3): 354-365. [曹建萌, 胡欣欣, 卢迈新, 等. 尼罗罗非鱼补体 *C9* 基因单核苷酸多态性及其与抗无乳链球菌感染的关联分析[J]. 农业生物技术学报, 2017, 25(3): 354-365.]
- [36] Hu S J, Wen L L, Hu X, et al. *IKZF1*: a critical role in the pathogenesis of systemic lupus erythematosus?[J]. Modern Rheumatology, 2013, 23(2): 205-209.
- [37] Westra H J, Peters M J, Esko T, et al. Systematic identification of trans eQTLs as putative drivers of known disease associations[J]. Nature Genetics, 2013, 45(10): 1238-1243.
- [38] Xie L R. Association of polymorphisms in *IKZF1* and *AFF3* with the risk of type 1 diabetes mellitus and autoimmune thyroid disease in a Chinese population[D]. Changchun: Jilin University, 2016. [解丽然. *IKZF1* 和 *AFF3* 基因多态性与中国 1 型糖尿病及自身免疫性甲状腺疾病的关联[D]. 长春: 吉林大学, 2016.]
- [39] Xie P, Li Q J, Liu M L, et al. Effect of a novel mutation in 5'-regulatory region of *aldose reductase* gene on its expression[J]. Acta Biochimica et Biophysica Sinica, 2003, 35(10): 904-908.
- [40] Liu J W, Yu Y, Li F H, et al. A new anti-lipopolysaccharide factor (*ALF*) gene with its SNP polymorphisms related to WSSV-resistance of *Litopenaeus vannamei*[J]. Fish & Shellfish Immunology, 2014, 39(1): 24-33.

Molecular cloning analysis of the 5' regulatory region of *Ikaros* gene from *Oreochromis niloticus* and screening of its SNP markers for *Streptococcus agalactiae* resistance

CHEN Kunping^{1,2,3}, LU Maixin¹, LIU Zhigang¹, GAO Fengying^{1,2}, CAO Jianmeng¹, ZHANG Defeng¹, LI Qingyong³, HU Xinxin^{1,2}

1. Pearl River Fisheries Research Institute, Chinese Academy of Fishery Sciences, Guangzhou 510380, China;

2. College of Fisheries and Life Science, Shanghai Ocean University, Shanghai 201306, China;

3. Fisheries Research and Extension Center of Huizhou, Huizhou 516002, China

Abstract: Tilapia (*Oreochromis*) is one of the most important fishery species in the world and has been introduced and cultured widely because of its fast growth rate, strong reproductive capacity, good adaptability and omnivorous feeding habit. However, *Streptococcus* disease has become the biggest threat for tilapia breeding. Recently, the epidemic and outbreak of tilapia *Streptococcus* disease caused substantial economic losses to aquaculture industry. Breeding resistant variety in tilapia was an important way to solve the problem of *Streptococcus* disease. The molecular marker assisted selection (MAS) has become an efficient breeding method for selecting and breeding tilapia, and will accelerate genetic improvement and increase selection intensity for disease resistance. The single nucleotide polymorphism (SNP) markers are used in many genetic and breeding studies because they are abundant in genomes, and can be genotyped easily. *Ikaros* is a kind of transcription factor with zinc finger structure that is essential to the development of lymphocyte and to the maintenance of normal immune function. Therefore, to obtain large amount of effective SNP molecular genetic markers and to perform MAS for disease resistance in tilapia, it is essential to study immune related candidate *Ikaros* gene in *Oreochromis niloticus* and to examine whether the SNPs in the gene are associated with disease resistance. In this study, the 5' regulatory region sequence of *Ikaros*, length of 4178 bp, were obtained through Genome Walking method from *O. niloticus*. Bioinformatics software was used to analyze the 5' regulatory region sequence of *Ikaros* gene. The predicted transcriptional start site (TSS) was in the initiation codon (ATG) upstream of 931 bp, and the core promoter regions was located at -57 bp to 48 bp when the TSS was specified as 1. The predicted promoter regions of *Ikaros* gene included basic start of substructure components: TATA box, CCAAT box and octamer. The analysis of transcription factor binding sites (TFBS) showed that abundant of TFBS were located at -2200 bp to 1200 bp in 5' regulatory region of *Ikaros* gene, such as GATA-1, Homeobox, CDP CR3+HD and AP-1. The analysis of CpG islands showed that two CpG islands were in 5' regulatory region sequence of *Ikaros* gene, one of which was located in promoter regions and the other of which was located in the first exon. Five SNPs in the 5' regulatory region of *Ikaros* gene were detected by direct sequencing method from the parents (P_0), which are named SNP1 (g.562, G>A), SNP2(g.217, G>T), SNP3(g.-53, C>T), SNP4(g.-220, T>C) and SNP5(g.-579, T>C). The five SNPs were sited in various regulatory elements in promoter regions which could have major implications for exact expression of *Ikaros* gene. Based on Snapshot method analysis, the frequencies of alleles and genotypes were calculated in susceptible groups and resistant groups of *O. niloticus* from the first filial generation (F_1). The polymorphisms and genetic parameter of the SNPs in resistant groups and susceptible groups were calculated by software Popgen 32 and PIC. The result showed that the polymorphism information content (PIC) of the 5' regulatory region of *Ikaros* SNPs was 0.0872~0.3747, suggesting that all the SNPs had moderate of intermediate polymorphism. The correlation between SNPs and resistance to *S. agalactiae* was analyzed by software SPSS 17.0, the results showed that four of them (SNP2, SNP3, SNP4 and SNP5) were significantly associated with the resistance to *S. agalactiae*.

($P < 0.05$). Moreover, All of the SNPs in the 5' regulatory region of *Ikaros* could formed one haplotype block and five haplotypes from the prediction of linkage disequilibrium analysis used software Haploview 4.2. The haplotypes (GGCTT) were significantly associated with the resistance to *S. agalactiae* ($P < 0.05$), and two of the haplotypes (GGTCT and GTCCC) were significantly associated with *S. agalactiae* susceptible ($P < 0.05$). Furthermore, the SNP2 and SNP5 were completely linked with each other ($r^2=1$, $LOD=57.25$, $D'=1$), which could be selected as tag SNP for research of genetic breeding in *O. niloticus*. The results suggest that the four SNPs (SNP2, SNP3, SNP4 and SNP5) and the haplotypes (GGCTT) in the 5' regulatory region of *Ikaros* could be potential genetic markers for future molecular selection of *O. niloticus* resistant to *S. agalactiae*.

Key words: *Oreochromis niloticus*; 5' regulatory region of *Ikaros* gene; single nucleotide polymorphism; *Streptococcus agalactiae*

Corresponding author: LU Maixin. E-mail: mx-lu@163.com