

DOI: 10.19663/j.issn2095-9869.20200205002

http://www.yykxjz.cn/

崔爱君, 徐永江, 王滨, 姜燕, 柳学周. 黄条鲮生长性状全基因组关联分析. 渔业科学进展, 2021, 42(2): 71-78  
Cui AJ, Xu YJ, Wang B, Jiang Y, Liu XZ. Genome-wide association analysis of growth traits in yellowtail kingfish (*Seriola lalandi*). Progress in Fishery Sciences, 2021, 42(2): 71-78

## 黄条鲮生长性状全基因组关联分析\*

崔爱君<sup>1,2</sup> 徐永江<sup>1,2①</sup> 王滨<sup>1</sup> 姜燕<sup>1</sup> 柳学周<sup>1</sup>

(1. 中国水产科学研究院黄海水产研究所 青岛海洋科学与技术试点国家实验室 海洋渔业科学与食物产出过程功能实验室 青岛 266071; 2. 上海海洋大学水产与生命学院 上海 201306)

**摘要** 利用 2b-RAD 技术对 119 尾黄条鲮(*Seriola lalandi*)个体进行测序, 共获得黄条鲮 SNP 分子标记 26665 个, 对黄条鲮个体的体质量和全长这 2 个重要生长性状进行全基因组关联分析, 筛选与体质量和全长性状相关的 SNP 位点和候选基因。结果显示, 黄条鲮体质量性状中共筛选到 17 个体质量显著关联的 SNP 位点, 找到 17 个可能的候选基因, 全长性状共筛选到 12 个潜在显著关联位点, 找到 12 个可能的候选基因。利用 KEGG 数据库对可能的候选基因进行 Pathway 分析, 得知候选基因主要参与了细胞或组织生长发育相关的代谢通路调控过程, 可能是影响黄条鲮生长性状密切相关的重要候选 SNP 位点和功能基因, 结果可为今后黄条鲮种质资源可持续利用和育种提供遗传信息资料积累。

**关键词** 黄条鲮; 生长性状; 全基因组关联分析; 2b-RAD; 简化基因组

**中图分类号** S917; Q78 **文献标识码** A **文章编号** 2095-9869(2021)02-0071-08

全基因组关联分析 (Genome-wide association study, GWAS) 是应用全基因组范围内的大量分子标记 (一般为 SNP), 将标记基因型结合性状表型进行联合分析, 统计每个标记与目标性状之间的关联性大小 (一般用  $P$  值表示), 鉴定出与目标性状密切相关且具有特定功能和育种潜力的基因位点或分子标记, 主要用于物种经济性状相关 SNP 分子标记以及功能基因的鉴定, 从而达到缩短育种周期和提高育种效率的目的, 目前已在畜禽等脊椎动物育种中广泛应用 (Tavares *et al.*, 2020; Müller *et al.*, 2019; Cui *et al.*, 2016; Zhang *et al.*, 2019)。近年来, 随着基因组高通量测序

技术的发展及测序成本的降低, GWAS 开始应用于水产养殖动物的育种研究, 如在大黄鱼 (*Larimichthys crocea*)、鲢鱼 (*Silurus asotus*)、凡纳滨对虾 (*Litopenaeus vannamei*)、龙胆石斑鱼 (*Epinephelus lanceolatus*)、虾夷扇贝 (*Patinopecten yessoensis*) 等物种的生长性状关联 SNP 位点、候选基因的挖掘和鉴定 (Zhou *et al.*, 2019; Li *et al.*, 2017; Yu *et al.*, 2019; Wu *et al.*, 2019; Ning *et al.*, 2019) 方面应用并取得了一定进展。但是, 与陆生脊椎动物相比, GWAS 在水产动物育种中的应用尚处于起步阶段。

黄条鲮 (*Seriola lalandi*) 属鲈形目 (Perciformes)、

\*山东省支持青岛海洋科学与技术试点国家实验室重大科技专项(2018SDKJ0303-1)、中国水产科学研究院基本科研业务费(2019GH15)、山东省重点研发计划项目(2018GHY115044)、国家重点研发计划项目(2019YFD0900901; 2018YFD0901204)和国家海水鱼产业技术体系(CARS-47)共同资助 [This work was supported by Marine S&T Fund of Shandong Province for Pilot National Laboratory for Marine Science and Technology (Qingdao) (2018SDKJ0303-1), Central Public-Interest Scientific Institution Basal Research Fund, CAFS (2019GH15), Key Research and Development Program of Shandong Province (2018GHY115044), National Key Research and Development Program of China (2019YFD0900901; 2018YFD0901204) and China Agriculture Research System (CARS-47)]. 崔爱君, E-mail: aijun0218@126.com

① 通讯作者: 徐永江, 研究员, E-mail: xuyj@ysfri.ac.cn

收稿日期: 2020-02-05, 收修改稿日期: 2020-03-02

鲷科(Carangidae)、鲷属(*Seriola*), 俗称“黄鲷牛”、“黄金鲷”, 是一种全球广泛分布且具有长距离洄游特性的大洋性经济鱼类。黄条鲷体型大, 生长迅速, 肉质鲜嫩, 具有较优的口感鲜度和较高的营养价值, 品质可与三文鱼、金枪鱼媲美, 国内外市场消费需求旺盛, 是鲷属鱼类中经济价值最高的一种。近年来, 全球很多国家纷纷开展黄条鲷的人工繁育和养殖(Sicuro *et al.*, 2016), 目前, 已在中国、日本、澳大利亚、新西兰和智利等国形成一定规模的养殖产业。为支撑养殖产业可持续发展, 初步开展了黄条鲷种质资源的相关研究, 主要包括染色体核型和带型(史宝等, 2017; 刘永山等, 2018)和群体遗传学(Premachandra *et al.*, 2017; Sepulveda *et al.*, 2017; Nguyen *et al.*, 2018a, 2018b)等研究, 相关结果为黄条鲷种质资源的科学利用和育种研究提供了遗传信息参考。我国于 2017 年在黄条鲷人工繁育技术方面取得突破, 培育出批量优质苗种, 开启了我国黄条鲷人工养殖的序幕。目前, 本实验室正在开展黄条鲷的群体遗传特性研究, 利用 2b-RAD 简化基因组测序技术获得了大量 SNP 分子标记, 对亲鱼和 F<sub>1</sub> 代苗种的遗传多样性变动情况进行监测。生长性状在水产养殖业中具有重要的经济价值, 直接和产量相关(Li *et al.*, 2017)。黄条鲷体型较大, 生长速度快, 前期研究发现, 体高和体长是影响黄条鲷体质量的 2 个重要性状, 共同决定程度达 42.88%(李荣等, 2017)。本研究选择全长和体质量这 2 个重要生长相关性状为目标, 利用获得的 SNP 分子标记进行全长和体质量性状的 GWAS 分析, 以期找到影响体质量和全长性状的关键 SNP 分子标记和候选功能基因, 为黄条鲷种质资源评价和持续利用提供遗传信息积累和技术依据。

## 1 材料与方法

### 1.1 实验材料

选择来自我国黄海种群养殖黄条鲷 F<sub>1</sub> 代 1~2 龄成鱼共 119 尾, 用卷尺测量其全长(cm), 精确到 1 mm; 电子天平测量其体质量(g), 精确到 0.1 g。利用高压高温(180℃)灭菌后的手术剪, 分别剪取样品的胸鳍, 以 95%无水乙醇浸润后, 在 4℃冰箱中保存待用。

### 1.2 基因组 DNA 提取与质量检测

使用 QIAGEN 公司生产的动物基因组 DNA 提取试剂盒(DP121221), 参照试剂盒使用说明, 提取鳍条基因组 DNA。用 NanoDrop 2000 分光光度计(Thermo, 美国)测定基因组 DNA 浓度, 通过 1%琼脂糖凝胶电

泳检测 DNA 的完整性, 通过 A<sub>260 nm</sub>/A<sub>280 nm</sub> 的比值来判断 DNA 的质量。将质检合格的 DNA 浓度稀释至 100 ng/μl, 于 -20℃ 条件保存备用。

### 1.3 文库构建与测序

将 ≥200 ng 的各样品基因组 DNA 采用 IIB 型限制性内切酶 *Bsa*XI 进行酶切, 酶切产物分别加入 5 组不同的接头, 使用 T4 DNA Ligase 连接, 然后 PCR 扩增连接产物, 最后根据 5 组接头信息, 将 5 个标签按顺序串联, 连接产物添加 barcode 序列, 混库, 使用 Illumina HiSeq 测序平台对混合好的文库进行 Paired-end 测序。

### 1.4 数据分析

**1.4.1 表型数据分析** 使用 R 语言中 `fivenum()` 函数, 对黄条鲷体质量和全长表型数据进行描述性统计, 计算其平均值、中位数、最小值、最大值、标准差和变异系数。

**1.4.2 测序数据分析与 SNP 分型** Illumina HiSeq 测序平台得到的原始图像数据文件经碱基识别转化为 Raw Reads, 过滤删除含有接头序列的 Reads, 得到 Clean Reads, 过滤删除含有 N 碱基比例大于 8% 的 Reads, 过滤删除低质量 Reads(质量值低于 Q30 的碱基超过 15%); 利用 Pear (Zhang *et al.*, 2014) 软件 (V0.9.6) 将成对的 Clean Reads 拼接, 提取出各样品对应的 Reads, 过滤删除不含酶切识别位点的 Reads 后, 得到各样品的 Enzyme Reads; 利用电子酶切从参考基因组中提取含有酶切识别位点的标签, 作为参考序列, 利用 SOAP 软件将各样品的 Enzyme Reads 比对到参考序列上, 主要参数为 -r0-M4-v2 (-r0 指唯一比对; -M4 指最优比对; -v2 指比对允许 2 个错配), 对比对到相同标签的 reads 聚类, 得到 unique 标签深度, 选择样品深度 >3× 且深度 <500、标签长度为 27 bp 的标签, 利用 SOAP 软件 (V 2.21) (Li *et al.*, 2008) 将测序数据比对到参考序列, 利用最大似然法 (ML) 进行位点的分型 (Fu *et al.*, 2013), 过程中使用的 RAD 分型软件包 (RAD typing), 包含 10 余个软件组分, 覆盖了从数据预处理至最终分型结果输出的全过程。

**1.4.3 全基因组关联分析** 使用 EMMA eXpedited (EMMAX) 高效混合模型 (Kang *et al.*, 2010), 通过方差分量方法进行 SNP 分子标记和表型性状的全基因组关联分析, 所用模型:

$$y = Xb + Ga + e$$

式中,  $y$  为表型值;  $X$  为固定效应关联矩阵,  $b$  为固定效应向量,  $G$  为通过 SNP 标记计算得到的关

系矩阵,  $a$  为随机加性遗传方差的参数,  $e$  为剩余效应的向量。

每个 SNP 位点能得到 1 个关联值  $P$ 。对 GWAS 给出的  $P$  值划定 2 条显著性水平线, 其中 1 条经 Bonferroni 校正  $P=0.05/N$  来确定全基因组显著性阈值 (Bonferroni, 1936),  $N$  为 SNP 标记的个数, 2 个性状经 Bonferroni 校正后显著关联阈值  $-\lg P=5.726$ ; 另一条使用 R 软件包中的  $p.adjust()$  函数计算得到经 FDR 校正后的阈值, 体质量性状潜在显著关联阈值  $-\lg P=4.091$ , 全长性状潜在显著关联阈值  $-\lg P=4.413$ , 挑选 Scaffold 长度的前 30 使用 R 软件包的 qqman 绘制曼哈顿图, 绘制 QQ 图对关联分析进行评价, 判断关联分析结果是否可靠。

### 1.5 候选基因鉴定及功能分析

将筛选到的具有关联性的 SNP 位点上下游 1 Mb 范围内的碱基序列与 GenBank 数据库中已有的黄条鲮参考基因组 (<https://www.ncbi.nlm.nih.gov/genome/?term=Seriola+lalandi+dorsalis>) 进行序列比对, 使用 SnpEff 软件 (Version 4.3T) (Cingolani *et al*, 2012) 对得到的 SNP 位点进行注释, 以确定 SNP 位点在基因元件的位置、对氨基酸的变化影响等找到距离 SNP 位点最近的基因。将所有关联基因与 KEGG 数据库比对, 进行 Pathway 分析, 并用超几何分布检验的方法计算每个 Pathway 条目中基因富集显著性, 公式如下:

$$p = 1 - \sum_{i=0}^{m-1} \frac{\binom{M}{i} \binom{N-M}{n-i}}{\binom{N}{n}}$$

式中,  $N$  为所有基因中具有 KEGG 注释的基因数目,  $n$  为  $N$  中差异表达基因中具有 KEGG 注释的基因数目,  $M$  为所有基因中注释为某特定 KEGG 的基因数目,  $m$  为注释某特定 KEGG 的差异表达基因的数目。计算的结果会返回一个富集显著性的  $P$  值, 小的  $P$  值表示基因在该 Pathway 中出现富集, 当  $P \leq 0.05$  表示显著富集。

## 2 结果

### 2.1 表型性状描述性统计

根据本研究所用黄条鲮样品的体质量和全长性状的描述性统计结果 (表 1) 可知, 体质量范围为 228~3000 g, 变异系数为 58.12, 全长范围为 27~67 cm, 变异系数为 19.49。统计分析显示, 黄条鲮体质量和全长的表型数据呈近正态分布。

表 1 黄条鲮生长性状表型的统计

Tab.1 Descriptive statistics of growth traits of yellowtail kingfish

性状 Trait	平均值 Mean	中位数 Median	最小值 Min	最大值 Max	标准差 SD	变异系数 CV
体质量 Weight	1237.87	947.85	228.00	3000.00	719.48	58.12
全长 Total length	47.98	46.50	27.00	67.00	9.35	19.49

### 2.2 SNP 分型

对 2b-RAD 简化基因组测序数据按照以下指标进一步过滤。剔除所有样品中低于 80% 个体可以分型的位点; 剔除 MAF 低于 0.05 的位点, 剔除等位基因大于 2 的位点。最终, 测序获得 26665 个 SNP 位点进行 GWAS 分析。

### 2.3 全基因组关联分析

利用 R 软件包分别绘制黄条鲮体质量和全长性状 GWAS 分析的 QQ 图 (图 1 和图 2), 比较观测值 (纵轴) 和期望值 (横轴) 的一致性, 散点分布与阈值实线在前端基本贴合一致, 代表 SNP 假阳性较低, 关联分析结果可靠。黄条鲮体质量与全长性状 GWAS 分析的曼哈顿图分别见图 3 和图 4。根据全基因组关联分析软件 EMMAX 方差分量法的计算和 Bonferroni 校正  $P$  值, 共筛选到 17 个与体质量性状显著相关的 SNP 位点 (表 2)。通过 NCBI 和 Ensembl 数据库比对, 当体质量性状的显著性值  $P=3.49 \times 10^{-10}$  时, 有 1 个 SNP 位点, 在基因组上找到 1 个基因 LOC111655074; 当  $P=1.20 \times 10^{-8}$  时, 鉴定出 15 个显著关联 SNP 位点, 找到 15 个关联基因, 分别是 ptpn14、LOC111651007、LOC111655681、LOC111658683、LOC111659677、

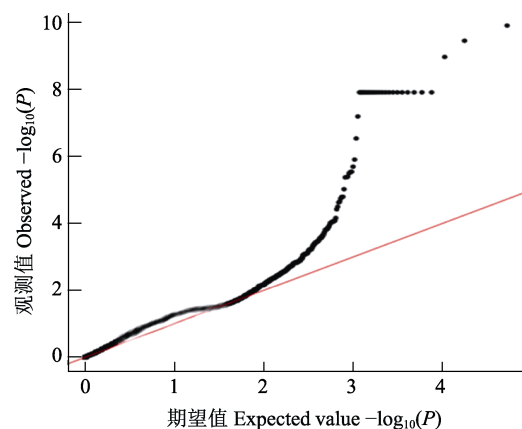


图 1 黄条鲮体质量性状 GWAS 关联分析的 QQ 检验  
Fig.1 QQ-plot of genome-wide association of body weight trait of yellowtail kingfish

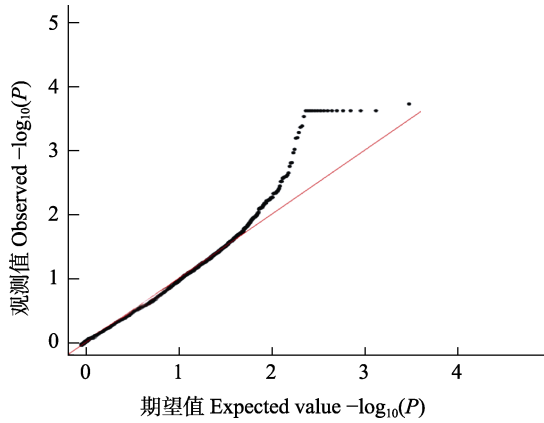


图2 黄条鲷全长性状 GWAS 关联分析的 QQ 检验  
Fig.2 QQ-plot of genome-wide association of total length trait of yellowtail kingfish

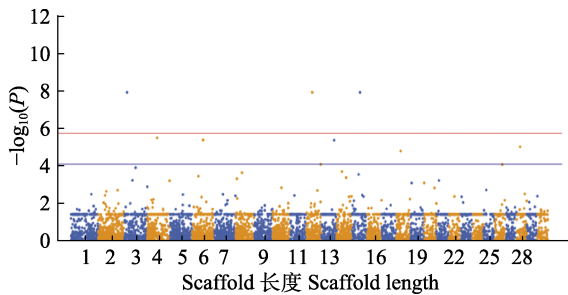


图3 黄条鲷体质量性状 GWAS 分析的曼哈顿图  
Fig.3 Manhattan plot of the genome-wide association analysis for body weight of yellowtail kingfish

红色实线代表全基因显著关联阈值:  $-\log_{10}P=5.726$ ,  
蓝色实线代表潜在显著关联阈值:  $-\log_{10}P=4.091$   
The red solid line indicates the genome wide significant threshold:  $-\log_{10}P=5.726$ . The blue solid line indicates the threshold for the significance of "suggestive association":  $-\log_{10}P=4.091$

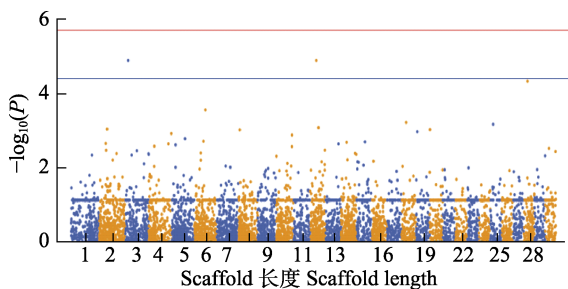


图4 黄条鲷全长性状 GWAS 分析的曼哈顿图  
Fig.4 Manhattan plot of the genome-wide association analysis for total length of yellowtail kingfish

红色实线代表全基因显著关联阈值  $-\log_{10}P=5.726$ ,  
蓝色实线代表潜在显著关联阈值  $-\log_{10}P=4.413$   
The red solid line indicates the genome wide significant threshold:  $-\log_{10}P=5.726$ . The blue solid line indicates the threshold for the significance of "suggestive association":  $-\log_{10}P=4.413$

gpr139、LOC111668306、LOC111670411、LOC111673333、LOC111645011、dag1、LOC111645458、poglut1、LOC111646889 和 LOC111648303; 当  $P=6.29 \times 10^{-8}$  时, 定位了 1 个 SNP 位点, 找到 1 个关联基因 LOC111657860。全长性状未找到经 Bonferroni 校正后的阈值线以上的显著关联位点, 但经 FDR 校正后找到全长性状潜在显著关联 SNP 位点 12 个(表 3), 可能的候选基因 12 个。

## 2.4 候选基因生物信息学分析

本研究共鉴定黄条鲷 17 个体质量性状显著关联 SNP 位点, 12 个全长性状潜在显著关联 SNP 位点。通过 NCBI 和 Ensembl 数据库比对, 分别找到 17 个体质量性状关联基因和 12 个全长性状关联基因。通过与 KEGG 数据库比对, 进行 Pathway 分析, 发现体质量和全长性状共 5 个关联基因参与的 6 条相关代谢通路(表 4), 主要与营养代谢、生长代谢、免疫调控、生殖调控等生物学过程密切相关; 这些与体质量和全长性状关联的 SNP 位点和相关基因参与多个生物学进程, 具有重要的调控作用, 在今后遗传选育中可能具有较大的应用价值。

## 3 讨论

黄条鲷作为一种全球分布的大洋性经济鱼类 (Swart *et al*, 2016), 因其高营养和经济价值在国际上诸多国家开始了苗种繁育和养殖 (Whatmore *et al*, 2013)。有学者分析了全球鲷鱼养殖发展现状, 认为种苗早期成活率低、遗传育种研究迟滞以及专用高效饲料缺乏是目前鲷鱼养殖产业发展的主要瓶颈 (Sicuro *et al*, 2016), 加之过度捕捞和环境退化加剧, 近年来, 全球黄条鲷种质资源下降迅速, 有必要加强黄条鲷种质资源与遗传育种相关研究。目前, 有关黄条鲷遗传育种的研究较少, 新西兰已开展了黄条鲷人工选育研究并取得了一定进展 (Symonds *et al*, 2014)。另外, 日本学者 Ohara 等 (2005) 通过微卫星标记建立了黄条鲷和五条鲷 (*Seriola quinqueradiata*) 的遗传连锁图谱, 为其遗传育种提供了重要工具, 但未进行经济性状的关联分析。澳大利亚学者 Nguyen 等 (2018b) 利用 SNP 分子标记构建了黄条鲷的高密度遗传连锁图谱, 并通过全基因组关联分析, 鉴定出 6 个与体质量性状显著关联的位点, 为体质量性状控制的遗传基础解析和下一步育种应用提供了依据。本研究首次利用 2b-RAD 简化基因组测序技术获得了黄条鲷 SNP 位点 26665 个, 并在此基础上进行全长、体质量性状的 GWAS 分析, 筛选到 17 个与体质量性状显著关联的

表 2 黄条鲮体质量性状显著关联基因  
Tab.2 Related genes for body weight traits of Yellowtail kingfish

SNP 位点名称 Ref ID	Scaffold 名称 Chr	位置 Loc	P 值 P value	基因名称 Gene Name	基因描述 Description
NW_019524855.1-9703	NW_019524855.1	6	$3.49 \times 10^{-10}$	LOC1111655074	solute carrier family 23 member 1-like
NW_019523057.1-610962	NW_019523057.1	21	$1.20 \times 10^{-8}$	p1pn14	tyrosine-protein phosphatase non-receptor type 14
NW_019523655.1-214205	NW_019523655.1	16	$1.20 \times 10^{-8}$	LOC1111651007	four and a half LIM domains protein 2-like
NW_019524883.1-915356	NW_019524883.1	22	$1.20 \times 10^{-8}$	LOC1111655681	platelet endothelial aggregation receptor 1-like
NW_019525017.1-364688	NW_019525017.1	3	$1.20 \times 10^{-8}$	LOC1111658683	putative nuclease HARBI1
NW_019525039.1-115887	NW_019525039.1	12	$1.20 \times 10^{-8}$	LOC1111659677	myosin-10-like isoform X2
NW_019525243.1-586198	NW_019525243.1	26	$1.20 \times 10^{-8}$	gpr139	probable G-protein coupled receptor 139
NW_019525288.1-1434838	NW_019525288.1	19	$1.20 \times 10^{-8}$	LOC1111668306	glypican-6-like isoform X2
NW_019525404.1-498088	NW_019525404.1	12	$1.20 \times 10^{-8}$	LOC1111670411	neuromedin-K receptor-like
NW_019525468.1-1066674	NW_019525468.1	14	$1.20 \times 10^{-8}$	LOC1111673333	adhesion G protein-coupled receptor B2-like
NW_019525496.1-2407075	NW_019525496.1	27	$1.20 \times 10^{-8}$	LOC1111645011	zinc finger and BTB domain-containing protein 34-like
NW_019525502.1-828754	NW_019525502.1	2	$1.20 \times 10^{-8}$	dag1	dystroglycan
NW_019525503.1-234341	NW_019525503.1	25	$1.20 \times 10^{-8}$	LOC1111645458	endothelial PAS domain-containing protein 1-like
NW_019525513.1-133817	NW_019525513.1	16	$1.20 \times 10^{-8}$	poglut1	protein O-glucosyltransferase 1
NW_019525527.1-152528	NW_019525527.1	5	$1.20 \times 10^{-8}$	LOC1111646889	transcriptional-regulating factor 1-like
NW_019525660.1-904568	NW_019525660.1	6	$1.20 \times 10^{-8}$	LOC1111648303	neuronal membrane glycoprotein M6-a
NW_019524969.1-162768	NW_019524969.1	2	$6.29 \times 10^{-8}$	LOC1111657860	testis-expressed protein 264-like

表 3 黄条鲮全长性状显著关联基因  
Tab.3 Related genes for total length traits of Yellowtail kingfish

SNP 位点名称 Ref ID	Scaffold 名称 Chr	位置 Loc	P 值 P value	基因名称 Gene Name	基因描述 Description
NW_019523655.1-214205	NW_019523655.1	16	$1.22 \times 10^{-5}$	LOC1111651007	four and a half LIM domains protein 2-like
NW_019524855.1-9703	NW_019524855.1	6	$8.82 \times 10^{-6}$	LOC1111655074	solute carrier family 23 member 1-like
NW_019525017.1-364688	NW_019525017.1	3	$1.22 \times 10^{-5}$	LOC1111658683	putative nuclease HARBI1
NW_019525039.1-115887	NW_019525039.1	12	$1.22 \times 10^{-5}$	LOC1111659677	myosin-10-like isoform X2
NW_019525243.1-586198	NW_019525243.1	26	$1.22 \times 10^{-5}$	gpr139	probable G-protein coupled receptor 139
NW_019525288.1-1434838	NW_019525288.1	19	$1.22 \times 10^{-5}$	LOC1111668306	glypican-6-like isoform X2
NW_019525404.1-498088	NW_019525404.1	12	$1.22 \times 10^{-5}$	LOC1111670411	neuromedin-K receptor-like
NW_019525468.1-1066674	NW_019525468.1	14	$1.22 \times 10^{-5}$	LOC1111673333	adhesion G protein-coupled receptor B2-like
NW_019525502.1-828754	NW_019525502.1	2	$1.22 \times 10^{-5}$	dag1	dystroglycan
NW_019525513.1-133817	NW_019525513.1	16	$1.22 \times 10^{-5}$	poglut1	protein O-glucosyltransferase 1
NW_019525527.1-152528	NW_019525527.1	5	$1.22 \times 10^{-5}$	LOC1111646889	transcriptional-regulating factor 1-like
NW_019525660.1-904568	NW_019525660.1	6	$1.22 \times 10^{-5}$	LOC1111648303	neuronal membrane glycoprotein M6-a

表 4 体质量性状和全长性状 KEGG 注释结果  
Tab.4 KEGG annotation results of body weight and total length trait

通路 ID KEGG ID	通路名称 Pathway	P 值 P value	基因名称 Gene name
ko00514	Other types of O-glycan biosynthesis	0.000367	poglut1
ko04512	ECM-receptor interaction	0.002148	dag1
ko04380	Osteoclast differentiation	0.004151	LOC111651007
ko04530	Tight junction	0.009224	LOC111659677
ko04020	Calcium signaling pathway	0.01454	LOC111670411
ko04080	Neuroactive ligand-receptor interaction	0.031111	LOC111670411

SNP 位点、12 个与全长性状潜在显著关联的 SNP 位点,为黄条鲮全长与体质量性状遗传基础解析和遗传育种提供了潜在可用的分子标记与候选基因。

GWAS 依赖于连锁不平衡(Linkage Disequilibrium, LD)检测目标种群体的遗传变异与性状之间的关联,然后通过统计基因型和表型的关联性大小筛选出影响显著的遗传变异,定位影响表型性状的重要数量性状位点(QTL)和候选基因,确定其遗传机制(陶林等, 2019; Naha *et al*, 2016)。本研究根据鉴定出与体质量和全长性状关联的 SNP 位点,在每个 SNP 上、下游 1 Mb 序列范围内扫描,共挖掘到 17 个体质量性状显著关联基因、12 个全长性状潜在显著关联基因,但这些基因尚未在黄条鲮研究中报道。已有研究表明,脊椎动物包括鱼类的生长过程是一个由多基因共同参与的级联式网络调控过程,除生长轴相关因子外,其他相关的基因、激素、代谢产物等通过内分泌、旁分泌的通路对生长起到调控作用,成为潜在的育种靶位点(Zhou *et al*, 2019)。Wang 等(2019)通过克隆获得了黄条鲮生长相关关键基因并验证了其在早期生长发育过程中的表达调控作用。但本研究未能关联体质量、全长性状至 GH、IGF 等经典生长调控功能基因。本研究定位到的所有黄条鲮体质量、全长性状关联基因中,通过 KEGG 或文献检索,发现部分基因在其他动物生长性状调控过程中起重要作用。如 adhesion G protein-coupled receptor B2-like、G protein-coupled receptor 139 等均为 G 蛋白偶联受体家族成员,主要参与有丝分裂、肌肉收缩、离子通道调控、基因转录等生理过程(Raise *et al*, 2015)。Myosin-10-like 和 four and a half LIM domains protein 2-like 都是重要的细胞骨架调控蛋白,参与信号转导、细胞黏附和细胞增殖(Matthias *et al*, 2006; Johannessen *et al*, 2006; Pi *et al*, 2008; Woolner *et al*, 2008)。Dag1 通过 ECM-receptor interaction 通路发挥作用,细胞与细胞外基质相互作用导致直接或间接控制细胞粘附、迁移、分化、增殖和凋亡等。LIM 结合结构域基因在鸡(*Gallus gallus*)脚重性状中显著关联并被证明是影响脚重的重要基

因(陈则东等, 2016)。在已报道的其他鱼类生长性状全基因组关联分析研究中,生长性状大多关联到生长轴以外的重要功能基因,这些生长轴以外的基因都通过相关通路对生长起到间接调控作用(Zhou *et al*, 2019; Li *et al*, 2018; Yu *et al*, 2019; Wu *et al*, 2019; Ning *et al*, 2019),提示除经典生长轴功能基因外,其他生长因子也会在鱼类生长性状选育中成为重要的靶向分子标记。

本研究发现了 29 个与黄条鲮体质量、全长性状显著或潜在显著关联的 SNP 位点,并定位到 29 个显著关联基因。这些基因在其他物种中主要参与细胞分化增殖、免疫、组织器官发育等相关的代谢通路调控过程,因此,可能会在今后黄条鲮选育中发挥重要的作用,但其具体的生理功能和遗传育种应用价值尚需研究和验证。全基因组关联分析(GWAS)已广泛应用于脊椎动物经济性状相关 SNP 位点筛选与鉴定,是进行 QTL 精细定位的重要工具,但在水产动物特别是鱼类上的研究报道较少。今后,随着鱼类基因组研究的深入发展和技术进步,GWAS 技术有望在养殖鱼类经济性状解析和遗传育种中得到更加广泛的应用(陈松林等, 2019)。

## 参 考 文 献

- Bonferroni CE. Teoria statistical delle classi e calcolo delle probabilita. Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze, 1936, 8: 3-62
- Chen SL, Xu WT, Liu Y. Fish genomic research: Decade review and prospect. *Journal of Fisheries of China*, 2019, 43(1): 1-14 [陈松林, 徐文腾, 刘洋. 鱼类基因组研究十年回顾与展望. *水产学报*, 2019, 43(1): 1-14]
- Chen ZD, Wang WH. Genome-wide association study on feet weight in chicken (*Gallus gallus*). *Journal of Agricultural Biotechnology*, 2016, 24(10): 1569-1577 [陈则东, 王文浩. 鸡脚重性状的全基因组关联分析. *农业生物技术学报*, 2016, 24(10): 1569-1577]
- Cingolani P, Platts A, Wang LL, *et al*. A program for annotating and predicting the effects of single nucleotide polymorphisms,

- SnEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly*, 2012, 6(2): 80–92
- Cui ZH, Luo JD, Qi CY, *et al.* Genome-wide association study (GWAS) reveals the genetic architecture of four husk traits in maize. *BMC Genomics*, 2016, 17(1): 946
- Fu X, Dou J, Mao J, *et al.* RAD typing: An integrated package for accurate *de novo* codominant and dominant RAD genotyping in mapping populations. *PLoS One*, 2013, 8(11): e79960
- Johannessen M, Moller S, Hansen T, *et al.* The multifunctional roles of the four-and-a-half-LIM only protein FHL2. *Cellular and Molecular Life Sciences*, 2006, 63(3): 268–284
- Kang HM, Sul JH, Service SK, *et al.* Variance component model to account for sample structure in genome-wide association studies. *Nature Genetics*, 2010, 42(4): 348–354
- Li N, Zhou T, Geng X, *et al.* Identification of novel genes significantly affecting growth in catfish through GWAS analysis. *Molecular Genetics and Genomics*, 2017, 293(3): 1–13
- Li R, Li Y, Kristiansen K, *et al.* SOAP: Short oligonucleotide alignment program. *Bioinformatics*, 2008, 24(5): 713–714
- Li R, Xu YJ, Liu XZ, *et al.* Morphometric analysis and internal anatomy of yellowtail kingfish (*Seriola aureovittata*). *Progress in Fishery Sciences*, 2017, 38(1): 142–149 [李荣, 徐永江, 柳学周, 等. 黄条鲮(*Seriola aureovittata*)形态度量与内部结构特征. *渔业科学进展*, 2017, 38(1): 142–149]
- Liu YS, Liu XZ, Shi B, *et al.* Analysis of the banding patterns of *Seriola aureovittata*. *Journal of Fisheries of China*, 2018, 42(9): 138–147 [刘永山, 柳学周, 史宝, 等. 黄条鲮染色体多种显带的形态特征分析. *水产学报*, 2018, 42(9): 138–147]
- Matthias C, Edwige T, Bernadette B, *et al.* FHL2 interacts with both ADAM-17 and the cytoskeleton and regulates ADAM-17 localization and activity. *Journal of Cellular Physiology*, 2006, 208: 363–372
- Müller BSF, de Almeida Filho JE, Lima BM, *et al.* Independent and joint-GWAS for growth traits in *Eucalyptus* by assembling genome-wide data for 3373 individuals across four breeding populations. *New Phytologist*, 2019, 221(2): 818–833
- Naha BC, Prasad A, Sailo L, *et al.* Concept of genome wide association studies and its progress in livestock. *International Journal of Science and Nature*, 2016, 7(1): 39–42
- Nguyen NH, Premachandra HKA, Kilian A, *et al.* Genomic prediction using DArT-Seq technology for yellowtail kingfish *Seriola lalandi*. *BMC Genomics*, 2018a, 19(1): 107
- Nguyen NH, Rastas PMA, Premachandra HKA, *et al.* First high-density linkage map and single nucleotide polymorphisms significantly associated with traits of economic importance in yellowtail kingfish *Seriola lalandi*. *Frontiers in Genetics*, 2018b, 9: 127
- Ning XH, Li X, Wang J, *et al.* Genome-wide association study reveals E2F3 as a candidate gene for scallop growth. *Aquaculture*, 2019, 73(4): 734216
- Ohara E, Nishimura T, Nagakura Y, *et al.* Genetic linkage maps of two yellowtails (*Seriola quinqueradiata* and *Seriola lalandi*). *Aquaculture*, 2005, 244: 41–48
- Pi X, Ren R, Kelley R, *et al.* Sequential roles for myosin-X in BMP6 dependent filopodial extension, migration, and activation of BMP receptors. *Journal of Cell Biology*, 2008, 179(7): 1569–1582
- Premachandra HKA, De la Cruz FL, Takeuchi Y, *et al.* Genomic DNA variation confirmed *Seriola lalandi* comprises three different populations in the Pacific, but with recent divergence. *Scientific Reports*, 2017, 7(1): 9386
- Raise A, Stefanie W, Ralf J, *et al.* Hunting for the function of orphan GPCRs-beyond the search for the endogenous ligand. *British Journal of Pharmacology*, 2015, 172(13): 3218–3228
- Sepulveda FA, Gonzalez M. Spatio-temporal patterns of genetic variations in populations of yellowtail kingfish *Seriola lalandi* from the southeastern Pacific Ocean and potential implications for its fishery management. *Journal of Fish Biology*, 2017, 90(1): 249–264
- Shi B, Liu YS, Liu XZ, *et al.* Study on the karyotype of yellowtail kingfish (*Seriola aureovittata*). *Progress in Fishery Sciences*, 2017, 38(1): 139–144 [史宝, 刘永山, 柳学周, 等. 黄条鲮(*Seriola aureovittata*)染色体核型分析. *渔业科学进展*, 2017, 38(1): 136–144]
- Sicuro B, Luzzana U. The state of *Seriola* spp. other than yellowtail (*S. quinqueradiata*) farming in the world. *Reviews in Fisheries Science and Aquaculture*, 2016, 24(4): 314–325
- Swart BL, Merwe BVD, Kerwath SE, *et al.* Phylogeography of the pelagic fish *Seriola lalandi* at different scales: Confirmation of inter-ocean population structure and evaluation of southern African genetic diversity. *South African Journal of Marine Science*, 2016, 38(4): 513–524
- Symonds JE, Walker SP, Pether S, *et al.* Developing yellowtail kingfish (*Seriola lalandi*) and hāpuku (*Polyprion oxygeneios*) for New Zealand aquaculture. *New Zealand Journal of Marine and Freshwater Research*, 2014, 48(3): 371–384
- Tao L, He XY, Di R, *et al.* Research progress on genome-wide association study for growth-related traits in livestock and poultry. *Chinese Journal of Animal Science*, 2019, 55(11): 34–41 [陶林, 贺小云, 荻冉, 等. 畜禽生长发育相关性状的全基因组关联分析研究进展. *中国畜牧杂志*, 2019, 55(11): 34–41]
- Tavares V, Pinto R, Assis J, *et al.* Venous thromboembolism GWAS reported genetic makeup and the hallmarks of cancer: Linkage to ovarian tumour behavior. *Biochimica et Biophysica Acta - Reviews on Cancer*, 2020, 1873(1): 188331
- Wang B, Xu Y, Liu X, *et al.* Molecular characterization and expression profiles of insulin-like growth factors in yellowtail kingfish (*Seriola lalandi*) during embryonic development. *Fish Physiology and Biochemistry*, 2019, 45(1): 375–390
- Whatmore P, Nguyen NH, Miller A, *et al.* Genetic parameters for

- economically important traits in yellowtail kingfish *Seriola lalandi*. *Aquaculture*, 2013, 400(25): 77–84
- Woolner S, O'Brien LL, Wiese C, et al. Myosin-10 and actin filaments are essential for mitotic spindle function. *Journal of Cell Biology*, 2008, 182(1): 77–88
- Wu LN, Yang Y, Li, BJ, et al. First genome-wide association analysis for growth traits in the largest coral reef-dwelling bony fishes, the giant grouper (*Epinephelus lanceolatus*). *Marine Biotechnology*, 2019, 21(5): 707–717
- Yu Y, Wang QC, Zhang Q, Genome scan for genomic regions and genes associated with growth trait in pacific white shrimp *Litopenaeus vannamei*. *Marine Biotechnology*, 2019, 21(3): 374–383
- Zhang J, Kobert K, Flouri T, et al. PEAR: A fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics*, 2014, 30(5): 614–620
- Zhang YF, Zhang JJ, Gong HF, et al. Genetic correlation of fatty acid composition with growth, carcass, fat deposition and meat quality traits based on GWAS data in six pig populations. *Meat Science*, 2019, 150: 47–55
- Zhou, Z. Han, K. Wu, Y, et al. Genome-wide association study of growth and body-shape-related traits in large yellow croaker (*Larimichthys crocea*) using ddRAD sequencing. *Marine Biotechnology*, 2019, 21(5): 655–670

(编辑 冯小花)

## Genome-Wide Association Analysis of Growth Traits in Yellowtail Kingfish (*Seriola lalandi*)

CUI Aijun<sup>1,2</sup>, XU Yongjiang<sup>1,2①</sup>, WANG Bin<sup>1</sup>, JIANG Yan<sup>1</sup>, LIU Xuezhou<sup>1</sup>

(1. Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Laboratory for Marine Fisheries Science and Food Production Processes, Pilot National Laboratory for Marine Science and Technology (Qingdao), Qingdao 266071;  
2. College of Fisheries and Life Science, Shanghai Ocean University, Shanghai 201306)

**Abstract** The genetic resources available for the commercially important pelagic yellowtail kingfish (*Seriola lalandi*) are relative sparse. 2b-RAD simplified genome sequencing technology was applied to screen single nucleotide polymorphisms (SNPs) in yellowtail kingfish, and a total of 26,665 SNPs were obtained. A genome-wide association study was carried out to detect body weight- and total length-associated SNPs in 119 individuals from the yellowtail kingfish population in the Yellow Sea. The results showed that 17 SNPs associated with body weight and with potential genome-wide significance were found. Genes in the candidate regions with 1 Mb windows were screened, and 17 candidate genes were obtained. A total of 12 SNPs associated with total length and with potential genome-wide significance were identified, and 12 candidate genes were found. For these candidate genes, KEGG pathway analysis showed that they are mainly involved in the metabolic regulation pathway of growth and development in other vertebrates, which may be important candidate SNP loci and functional genes closely related to the growth traits of yellowtail kingfish. The present results could provide genetic information for the sustainable utilization of germplasm resources and genetic breeding of yellowtail kingfish in the future.

**Key words** *Seriola lalandi*; Growth trait; Genome-wide association study (GWAS); 2b-RAD; Simplified genome

① Corresponding author: XU Yongjiang, E-mail: xuyj@ysfri.ac.cn