

李 娟, 王 亚, 曹 霄, 等. 乌鳢生长性状全基因组关联分析[J]. 渔业研究, xxxx, xx(x): 1–10.

Li X, Wang Y, Cao X, et al. Genome-wide association study of growth traits in *Channa argus* [J]. Journal of Fisheries Research, xxxx, xx(x): 1–10.

乌鳢生长性状全基因组关联分析



李 娟¹, 王 亚², 曹 霄², 朱树人¹, 安 丽¹, 朱永安¹,
孟庆磊¹, 董 文¹, 董 俊¹

(1. 山东省淡水渔业研究院, 山东 济南 250013;

2. 微山县渔业发展服务中心, 山东 济宁 277699)

摘要: 【目的】乌鳢 (*Channa argus*) 是中国重要的淡水经济鱼类, 近年来受环境和养殖方式等影响, 其种质资源逐渐退化。本文利用全基因组关联分析 (GWAS) 方法研究乌鳢生长性状的遗传基础, 以期为其种质资源保护和良种选育提供依据。【方法】利用 Illumina Nova (PE150) 平台和 SuperGBS 技术对 405 尾乌鳢进行简化基因组测序; 利用 EMMAX 软件的高效混合模型进行 GWAS 分析, 寻找与乌鳢生长性状显著关联的单核苷酸多态性 (SNP) 位点; 扫描显著性位点上下游 50 kb 的序列, 查找潜在候选基因, 根据 NCBI 数据库和文献检索结果, 注释候选基因的生物学功能, 筛选与目标性状相关的候选功能基因。【结果】测序数据过滤后, 共获得 47536 个 SNP 位点, GWAS 分析筛选到 3 个与体质量性状显著关联的 SNP 位点, 即 2831060、2830864、2830886。扫描显著性位点上下游 50 kb 序列, 注释到 7 个与体质量相关联的候选功能基因; 此外, 还筛选到 6 个全长性状潜在关联的 SNP 位点, 注释到 17 个与全长性状相关联的潜在候选功能基因。在这些功能基因中, 有 7 个基因与全长和体质量性状均存在关联, 分别是 *irak3*、*tuba1a*、*hspa14*、*prl*、*rint1*、*helb* 和 *net1*, 其主要参与乌鳢的生长代谢、发育调控、细胞增殖和免疫调控等生物学过程。【结论】本研究挖掘的 SNP 及功能基因与乌鳢的体质量和全长性状存在显著关联, 可为乌鳢生长性状的机制解析、良种选育和资源保护提供重要参考。

关键词: 乌鳢; 全基因组关联分析 (GWAS); 生长性状; 简化基因组测序

中图分类号: S917.4 **文献标识码:** A **文章编号:** 2096–9848 (xxxx) 00–0001–10

全基因组关联分析 (Genome-wide association study, GWAS) 是针对随机群体设计的关联分析方法, 最初被用于人类疾病相关基因位点的分析, 如今已被广泛应用于育种研究, 成为推动水产养殖业可持续发展的重要手段^[1]。何骞等^[2]使用 GWAS 方法在大黄鱼 (*Larimichthys crocea*) 5 条染色体上定位到 6 个与肌纤维性状显著相关的单核苷酸多态性 (Single nucleotide polymorphism, SNP) 位点; 崔爱君等^[3]利用 GWAS 方法在黄条

鲮 (*Larimichthys crocea*) 中找到 17 个体质量性状和 12 个全长性状相关的 SNP 位点和候选基因; 方家璐等^[4]通过 GWAS 定位到与黄河鲤 (*Cyprinus carpio*) 体质量性状相关的 5 个 SNP 和候选基因。通过 GWAS 筛选出与优良生长性状相关的分子标记, 可以在早期快速筛选出具有优良性状的个体, 加速育种进程, 为水产动物经济性性状的全基因组选择和品种培育提供重要依据。

乌鳢 (*Channa argus*) 属鲈形目 (Perciformes)、

收稿日期: 2025-04-18

修回日期: 2025-05-14

基金项目: 山东省重点研发计划 (2021LZGC029); 国家淡水水产种质资源库 (FGRC: 18537)

第一作者: 李 娟, 女, 副研究员, 研究方向为水产种质资源保护与育种。E-mail: lxla@sina.com

©《渔业研究》编辑部。本文为使用 CC BY-NC-ND 4.0 许可协议的开放获取作品。

© Editorial Office of Journal of Fisheries Research. This is an open access article under the CC BY-NC-ND 4.0 license.

鳢科 (Channidae)、鳢属 (*Channa*), 是中国重要的淡水经济鱼类。乌鳢生长速度快、环境适应性强, 肉质细腻、营养丰富、无肌间刺, 是预制水产品的主要来源之一, 市场前景广阔^[5-6]。然而, 随着养殖规模的扩大和养殖密度的增加, 以及养殖企业长期自养自繁等原因, 乌鳢种质资源逐渐退化, 出现生长变缓、病害频发、成活率低等现象, 严重影响了其商品价值和养殖效益。因此, 开展乌鳢种质资源保护和新品种培育已成为产业可持续发展的重要需求。目前, 乌鳢的良种培育主要以杂交鳢为主^[7], 尚未有采用 GWAS 方法开展乌鳢经济性状选育的报道。为深入分析乌鳢生长性状的遗传基础, 定位相关候选基因, 本研究选择体质量和体长 2 个生长性状为目标性状, 采用 GWAS 方法寻找目标性状关联的 SNP 位点和候选功能基因, 以期对乌鳢的种质资源保护、良种创制和持续利用提供基础资料。

1 材料与方法

1.1 实验材料

实验材料选择微山乌鳢省级原种场同批次繁育的乌鳢, 均一化养殖至 12 月龄, 共选取 13 个家系 405 尾乌鳢用于分析。分别使用卡尺和电子天平测量乌鳢的全长 (精确到 0.01 cm) 和体质量 (精确到 0.01 g)。剪取乌鳢胸鳍, 置于无水乙醇中, 24 h 后换液, 在 -20 °C 冰箱中保存。山东省淡水渔业研究院批准动物实验, 批准编号为 LLSC2025001。

1.2 文库构建与测序

样品使用 TIANGEN 动物基因组抽提试剂盒进行基因组 DNA 提取, 再使用分光光度计检测 DNA 纯度 (OD₂₆₀/OD₂₈₀ 比值), 并通过琼脂糖凝胶电泳分析 DNA 纯度和完整性。DNA 抽提质检合格后, 利用 SuperGBS 技术^[8] 构建测序文库, 采用 PstI-HF/MspI 酶切 DNA, 酶切后的片段两端通过 T4 连接酶加接头和 barcode, 使用磁珠回收系统回收 300~700 bp 的片段, 对回收片段使用高保真酶进行聚合酶链式反应 (Polymerase chain reaction, PCR) 扩增, 使用 Qubit 测定 PCR 产物浓度 (浓度需大于 5 ng/μL), 而后使用 Illumina Nova (PE150) 平台, 将各样本按照一定浓度进行混库上机测序。

1.3 数据分析

1.3.1 表型数据分析

使用 Excel 和 SPSS 软件对测得的体质量和全

长性状进行描述性统计分析, 计算各性状的最大值、最小值、中位数、平均值和标准差等, 并进行独立样本 *t* 检验和相关性分析。

1.3.2 测序数据分析

经 Illumina Nova 高通量测序的原始图像数据序列用碱基识别后转化为原始测序序列 (Raw Reads), 其结果以 FASTQ (简称为 fq) 文件格式存储。使用预处理软件 fastp 对 Raw Reads 进行质量过滤, 去除接头序列、N 碱基 ≥ 5 和平均碱基质量值 < 20 的 Reads, 获取 Clean Reads。利用 BWA^[9] 软件将 Clean Reads 比对到参考基因组上, 根据比对结果统计样品深度信息并计算测序数据对乌鳢参考基因组 (Accession: PRJNA721844) 的覆盖度。使用 Qualimap^[10] 统计各样本的插入片段长度分布。

1.3.3 SNP 注释

基于样本与参考基因组的比对结果, 使用 GATK4^[11] 软件的 Haplotypecaller 模块进行 SNP 检测。QD 是突变质量值 (Quality) 除以覆盖深度 (Depth) 得到的比值, 以 QD ≥ 2.0 进行过滤, 降低 SNP 和 InDel 检测的错误率, 只保留同时满足该条件的突变位点。使用 VCFtools^[12] 进一步过滤, 保留 Reads 支持深度不小于 4 的位点, 去除最小等位基因频率 (Minor allele frequency, MAF) 小于 0.01 的位点, 保留 80% 的个体能够分型的位点。对于有注释文件 (gff 文件) 的基因组, 使用 SnpEff^[13] 软件对得到的 SNP 进行注释, 以确定 SNP 在基因元件的位置、对氨基酸的变化影响等。

1.4 全基因组关联分析

使用 EMMAX 软件的高效混合模型进行 SNP 标记和 GWAS 分析, 以基因型为主要自变量、表型值为因变量, 同时考虑并控制群体结构、亲缘关系等其他因素可能造成的假阳性。GWAS 的主要结果是每个 SNP 位点 (或其他类型的标记) 与表型的相关性程度, 一般用 *P* 值表示, 每个 SNP 能得到 1 个相对应的统计概率 *P* 值。零假设为该 SNP 位点与表型无关, 备择假设为该 SNP 位点与表型有关, 当在零假设下得到的 *P* 值很小时, 表示该 SNP 位点与表型无关的可能性很小, 也就是说该 SNP 位点很可能与表型相关, 这类 SNP 位点优先进行下一步的实验验证。研究使用 2 种阈值筛选强显著性位点, 第一种控制错误发现率 (False discovery rate, FDR) 法是由 Benjamini 和 Hochberg 提出的通过控制错误发现的概率对 *P* 值进行调整的方法^[14]; 第二种 Bonferroni 校正法是多重比较中控

制假阳性最为严格和保守的一种方法^[15]，将单次检验的显著性 P 值 0.05 除以多重比较的次数（在 GWAS 中为 SNP 位点的个数）作为校正后的阈值，将各位点的 P 值与之进行比较，如果某位点的 P 值小于校正阈值，则可判别该位点与性状之间存在显著性关联。根据 GWAS 分析结果绘制 Quantile-Quantile (QQ) plot 和曼哈顿图。

1.5 候选基因鉴定及功能分析

筛选表型 P 值小于 0.05 的位点，利用已发表的乌鳢基因注释信息，扫描显著性位点上下游 50 kb 的序列，查找潜在候选基因。根据 NCBI 数据库和

文献检索结果，注释候选基因的生物学功能，筛选与目标性状相关的候选功能基因。

2 结果

2.1 表型性状描述性统计

本研究所用乌鳢样品的体质量和全长性状的描述性统计结果参见表 1，所用乌鳢全长为 16.61~27.22 cm，平均值为 20.92 cm，变异系数为 7.28；体质量为 22.66~180.66 g，平均值为 77.13 g，变异系数为 24.29。统计分析显示，乌鳢的体质量和全长性状数据呈近正态分布。

表 1 乌鳢表型性状
Tab. 1 Descriptive statistics of *C. argus*

性状 Traits	最大值 Max	最小值 Min	中位数 Median	平均值 Mean	标准差 SD	变异系数/% CV
全长 Total length	27.22	16.61	20.91	20.92	1.52	7.28
体质量 Body mass	180.66	22.66	75.55	77.13	18.73	24.29

2.2 SNP 分型

对 SuperGBS 技术简化基因组测序的数据进行进一步过滤，保留 Reads 支持深度不小于 4 的位点，去除 $MAF < 0.01$ 的位点，保留 80% 的个体能够分型的位点。过滤后得到 47536 个 SNP 位点，用于乌鳢表型性状的 GWAS 分析。

2.3 全基因组关联分析

根据 EMMAX 软件运行结果，得到每个 SNP 位点与各性状关联的 P 值，绘制乌鳢全长和体质量的 QQ plot，比较各 SNP 的 P 值观测值（纵轴）与期望值（横轴）的一致性（图 1 和图 2）。如图所示，散点分布与阈值实线在前段基本一致，表明得到的 SNP 假阳性率低，GWAS 分析结果较为准确。

根据 EMMAX 软件结果和 Bonferroni 校正 P 值，得到乌鳢全长和体质量曼哈顿图（图 3 和图 4）。全长和体质量性状未找到 Bonferroni 校正后的阈值线以上显著关联的 SNP 位点，但经 FDR 校正后，找到了 3 个与体质量性状显著关联的 SNP 位点。

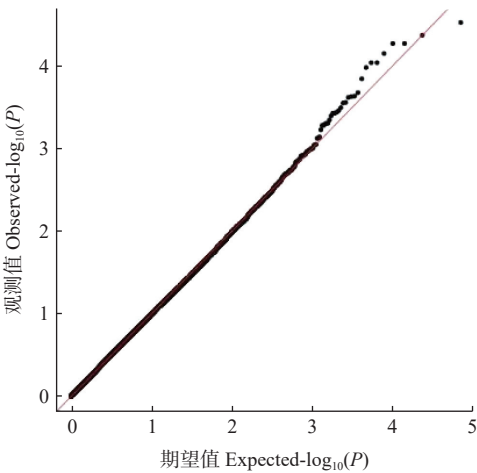


图 1 乌鳢全长性状 QQ plot
Fig. 1 QQ plot of total length trait of *C. argus*

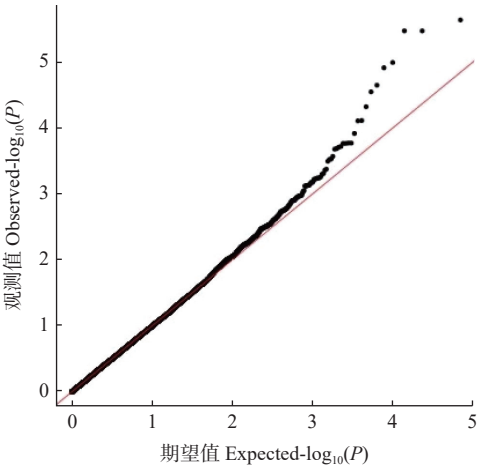


图 2 乌鳢体质量性状 QQ plot
Fig. 2 QQ plot of body mass trait of *C. argus*

2.4 候选基因及功能注释

对获得的体质量显著位点上下游 50 kb 区域进行扫描，比对 NCBI 和 Ensemble 数据库，3 个体质量性状显著关联 SNP 位点共找到 7 个潜在关联基因；对潜在关联基因进行功能注释，共注释到 7 个与乌鳢体质量相关联的候选功能基因（表 2）。

全长性状未找到显著关联 SNP 位点；对全长性状 SNP 位点 P 值从小到大排序，得到 6 个 P 值

小于 1×10^{-4} 的 SNP 位点, 对这 6 个位点上下游 50 kb 区域进行扫描, 寻找潜在候选关联基因并对其进行功能注释, 共注释到 17 个与乌鳢体全长性状相关的潜在功能基因 (表 3)。

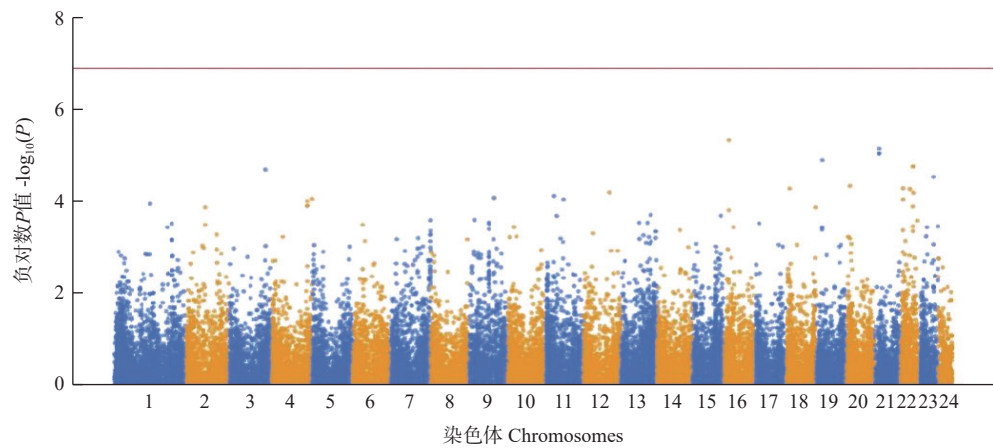


图 3 乌鳢全长性状曼哈顿图

Fig. 3 Manhattan plot of total length trait of *C. argus*

注: 红线是对 GWAS 给出的 P 值划定的显著性水平线 (已经进行 $-\log_{10}$ 转换), 红线经 Bonferroni 矫正后的 $P=0.05$ 。
Notes: The red line represents the significance level line for the P values given by GWAS (which have already been transformed by $-\log_{10}$). The red line indicates the P value of 0.05 after Bonferroni correction.

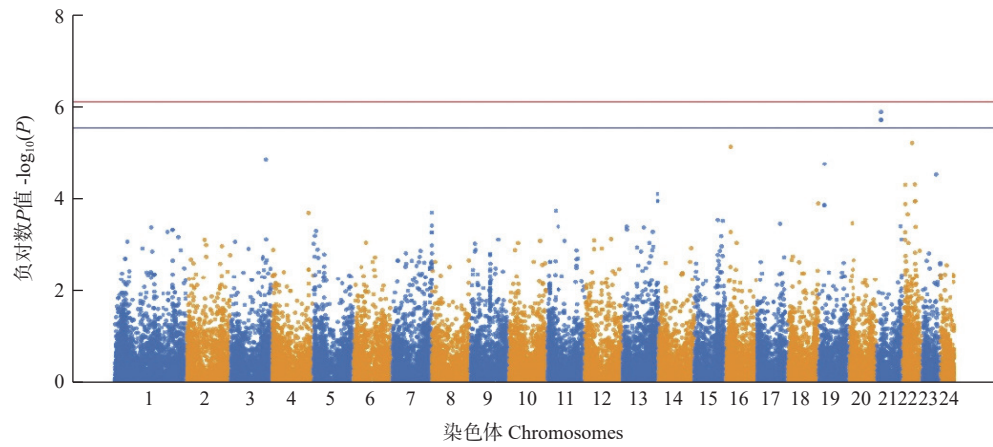


图 4 乌鳢体质量性状曼哈顿图

Fig. 4 Manhattan plot of body mass trait of *C. argus*

注: 红线和蓝线是对 GWAS 给出的 P 值划定 2 条显著性水平线 (已经进行 $-\log_{10}$ 转换), 蓝线是 $FDR=0.05$, 红线是 Bonferroni 矫正后的 $P=0.05$ 。
Notes: The red and blue lines represent two significance level lines for the P values given by GWAS (which have already been transformed by $-\log_{10}$). The blue line corresponds to a false discovery rate (FDR) of 0.05, while the red line indicates the P value of 0.05 after Bonferroni correction.

表 2 乌鳢体质量性状 GWAS 分析结果及候选基因

Tab. 2 GWAS analysis results and candidate genes for body mass trait of *C. argus*

染色体 Chromosome	SNP 位置 SNP location	P 值 P value	基因 ID Gene ID	候选基因 Candidate genes	编码蛋白 Coding protein
21	2831060	2.30E-06	EXN66_Car020554	<i>irak3</i>	Interleukin-1 receptor-associated kinase 3
			EXN66_Car020555		
			EXN66_Car020556		
			EXN66_Car020557	<i>tuba1a</i>	Tubulin alpha-1A chain
			EXN66_Car020558		
			EXN66_Car020559		

续表 2

染色体 Chromosome	SNP 位置 SNP location	P 值 P value	基因 ID Gene ID	候选基因 Candidate genes	编码蛋白 Coding protein
21	2830864	3.37E-06	EXN66_Car020561	<i>hspa14</i>	Heat shock 70 kDa protein 14
			EXN66_Car020562	<i>prl</i>	Prolactin
			EXN66_Car020563	<i>rint1</i>	RAD50-interacting protein 1
			EXN66_Car020560	<i>helb</i>	DNA helicase B
			EXN66_Car020564	<i>net1</i>	Neuroepithelial cell-transforming gene 1 protein
			EXN66_Car020554	<i>irak3</i>	Interleukin-1 receptor-associated kinase 3
			EXN66_Car020555 EXN66_Car020556 EXN66_Car020557 EXN66_Car020558 EXN66_Car020559	<i>tuba1a</i>	Tubulin alpha-1A chain
			EXN66_Car020561	<i>hspa14</i>	Heat shock 70 kDa protein 14
			EXN66_Car020562	<i>prl</i>	Prolactin
			EXN66_Car020563	<i>rint1</i>	RAD50-interacting protein 1
21	2830886	3.37E-06	EXN66_Car020560	<i>helb</i>	DNA helicase B
			EXN66_Car020564	<i>net1</i>	Neuroepithelial cell-transforming gene 1 protein
			EXN66_Car020554	<i>irak3</i>	Interleukin-1 receptor-associated kinase 3
			EXN66_Car020555 EXN66_Car020556 EXN66_Car020557 EXN66_Car020558 EXN66_Car020559	<i>tuba1a</i>	Tubulin alpha-1A chain
			EXN66_Car020561	<i>hspa14</i>	Heat shock 70 kDa protein 14
			EXN66_Car020562	<i>prl</i>	Prolactin
			EXN66_Car020563	<i>rint1</i>	RAD50-interacting protein 1
			EXN66_Car020560	<i>helb</i>	DNA helicase B
			EXN66_Car020564	<i>net1</i>	Neuroepithelial cell-transforming gene 1 protein
			EXN66_Car020554	<i>irak3</i>	Interleukin-1 receptor-associated kinase 3

表 3 乌鳢全长性状 GWAS 分析结果及候选基因

Tab. 3 GWAS analysis results and candidate genes for total length trait of *C. argus*

染色体 Chromosome	SNP 位置 SNP location	P 值 P value	基因 ID Gene ID	候选基因 Candidate genes	编码蛋白 Coding protein
16	3305852	3.01E-05	EXN66_Car016480	<i>foxi1c</i>	Forkhead box protein I1c
			EXN66_Car016481	<i>cfb</i>	Complement factor D
			EXN66_Car016482	<i>dnaaf9</i>	Dynein axonemal assembly factor 9
			EXN66_Car016479	<i>plaat4</i>	Phospholipase A and acyltransferase 4
21	2831060	2.30E-06	EXN66_Car020554	<i>irak3</i>	Interleukin-1 receptor-associated kinase 3
			EXN66_Car020555 EXN66_Car020556 EXN66_Car020557 EXN66_Car020558 EXN66_Car020559	<i>tuba1a</i>	Tubulin alpha-1A chain
			EXN66_Car020561	<i>hspa14</i>	Heat shock 70 kDa protein 14
			EXN66_Car020562	<i>prl</i>	Prolactin
			EXN66_Car020563	<i>rint1</i>	RAD50-interacting protein 1
			EXN66_Car020560	<i>helb</i>	DNA helicase B
			EXN66_Car020564	<i>net1</i>	Neuroepithelial cell-transforming gene 1 protein
			EXN66_Car020554	<i>irak3</i>	Interleukin-1 receptor-associated kinase 3
			EXN66_Car020555 EXN66_Car020556 EXN66_Car020557 EXN66_Car020558 EXN66_Car020559	<i>tuba1a</i>	Tubulin alpha-1A chain
			EXN66_Car020561	<i>hspa14</i>	Heat shock 70 kDa protein 14

续表 3

染色体 Chromosome	SNP 位置 SNP location	P 值 P value	基因 ID Gene ID	候选基因 Candidate genes	编码蛋白 Coding protein
			EXN66_Car020561	<i>hspa14</i>	Heat shock 70 kDa protein 14
			EXN66_Car020562	<i>prl</i>	Prolactin
			EXN66_Car020563	<i>rint1</i>	RAD50-interacting protein 1
			EXN66_Car020560	<i>helb</i>	DNA helicase B
			EXN66_Car020564	<i>net1</i>	Neuroepithelial cell-transforming gene 1 protein
21	2830886	3.37E-06	EXN66_Car020554	<i>irak3</i>	Interleukin-1 receptor-associated kinase 3
			EXN66_Car020555	<i>tuba1a</i>	Tubulin alpha-1A chain
			EXN66_Car020556		
			EXN66_Car020557		
			EXN66_Car020558		
			EXN66_Car020559		
			EXN66_Car020561	<i>hspa14</i>	Heat shock 70 kDa protein 14
			EXN66_Car020562	<i>prl</i>	Prolactin
			EXN66_Car020563	<i>rint1</i>	RAD50-interacting protein 1
			EXN66_Car020560	<i>helb</i>	DNA helicase B
			EXN66_Car020564	<i>net1</i>	Neuroepithelial cell-transforming gene 1 protein
19	4093798	7.13E-05	EXN66_Car019069	<i>pfkp</i>	ATP-dependent 6-phosphofructokinase, platelet type
			EXN66_Car019070	<i>klf6</i>	Krueppel-like factor 6
22	8803643	9.22E-05	EXN66_Car021470	<i>trim3</i>	Tripartite motif-containing protein 3
			EXN66_Car021471	<i>fhdc1</i>	FH2 domain-containing protein 1
			EXN66_Car021472	<i>arfip2</i>	Arfapatin-2
			EXN66_Car021468	<i>frem2</i>	FRAS1-related extracellular matrix protein 2

3 讨论

乌鳢的性状关联基因研究已取得一定的进展，主要集中在体色、生长、性别等性状方面。黄苏静等^[16]比较分析了乌鳢 3 种性腺组织的转录组，筛选出大量乌鳢性别分化和精巢、卵巢发育调控相关的候选基因和信号通路；徐成豪^[17]基于构建的乌鳢遗传图谱定位到 3 个全长性状显著相关 SNP 位点，2 个体重性状、2 个体色性状和 1 个性别数量性状位点（Quantitative trait locus，QTL）区间；Luo 等^[18]研究发现 *amh* 基因在乌鳢性腺性别分化中起重要作用；Sun 等^[19]证实 *slc45a2* 基因的无义突变（c.383G>A）是乌鳢体色性状突变的原因；Fan 等^[20]研究发现肝脏和脾脏中的 *pcdhf4*、*nlrc3* 和 *card15-like* 基因可能与白色乌鳢的发育和免疫相关。

研究中定位到 7 个与体质量和全长性状均相关的候选基因。其中，重组 DNA 修复蛋白（*rint1*）基因是染色体结构维护（SMC）蛋白家族成员，参与多种细胞过程，包括 DNA 双链断裂修复、细胞周期检查点激活、端粒维护和减数分裂，在细胞增殖中起重要作用^[21]。DNA 解旋酶 B（*helb*）是一种在高等真核生物中保守的解旋酶，主要在 G 期

核内参与 DNA 复制的起始过程，在 DNA 损伤和复制应激反应中也有重要作用^[22]。神经上皮细胞转化基因 1（*net1*）在细胞内信号传导中起关键作用，主要通过调节 Rho 家族小 G 蛋白的活性来影响包括肌动蛋白细胞骨架的组织、细胞周期进展等多种功能^[23-24]。X-微管蛋白（*tuba1a*）属于微管蛋白家族，是一种重要的细胞骨架蛋白，在细胞分裂、物质运输和细胞形态维持等过程中发挥关键作用^[25]。白介素 1 受体关联激酶 3（*irak3*）是受体相关激酶蛋白家族成员，是 Toll/IL-R 免疫信号转导通路的重要组成部分，主要在单核细胞和巨噬细胞中表达，在副溶血弧菌感染的斑马鱼（*Danio rerio*）幼鱼中参与天然免疫应答^[26-27]。热休克蛋白 14（*hspa14*）是广泛存在的一类小分子热休克蛋白基因，在进化过程中的高度保守，可提高细胞的应激能力，研究发现在中华鲟（*Acipenser sinensis*）、达氏鲟（*Acipenser dabryanus*）和青海湖裸鲤（*Gymnocypris przewalskii*）中均参与机体应激外源性胁迫的生物过程^[28]。催乳素（*prl*）是由动物脑垂体合成与分泌的一种单链多肽类激素，是硬骨鱼类调节渗透压和离子平衡的主要激素，研究发现 *prl* 在草鱼（*Ctenopharyngodon idella*）、罗非鱼（*Oreochromis niloticus*）等鱼类中与生长性状

显著相关。研究中筛选出的这些候选基因在生长代谢、发育调控、细胞增殖、免疫调控、生理调控等生物学功能中起重要作用^[29-30]，与乌鳢的生长性状可能存在密切的关联，是乌鳢体质量和全长性状的重要候选功能基因，但在乌鳢生长发育过程中的作用有待进一步的验证。

徐成豪^[17]在采用 2b-RAD (2b-restriction site-associated DNA) 测序技术对 130 尾 1 月龄杂交乌鳢 (野生型乌鳢♀×突变型乌鳢♂) 全长性状 QTL 定位研究中，发现了 3 个显著相关的编码蛋白基因 *znf208*、*znf91*、*mxs*，且均位于 17 号连锁群上。本研究定位到的 7 个生长性状候选基因均位于乌鳢 21 号染色体上，在鲤、红罗非鱼 (*Oreochromis* spp.) 和大黄鱼的 GWAS 分析中也有类似现象^[31-33]，即目标性状 SNP 和关联基因表现出成簇分布的特点，说明某些特定性状的关联位点往往集中于部分染色体，甚至集中于染色体的某个特定区域。

本研究发现，体质量性状显著关联的 SNP 位点及 7 个关联基因均出现在全长性状潜在关联的 SNP 位点及候选基因中，且 *P* 值仅高于位点 SNP 3305852，表现出与全长性状较高的关联性。目前，大多数 GWAS 分析是根据单性状进行目标性状和遗传位点之间的关联分析，但当对多个表型性状同时进行分析时，会存在 1 个位点或基因同时影响多个性状的情况，即一因多效性^[34]。鱼类多性状的 GWAS 研究处于起步阶段，蒋仪等^[35]使用表型正交化混合模型对大菱鲆 (*Scophthalmus maximus*) 体质量、体长和尾柄长性状进行了多性状的 GWAS 关联分析，提高了检索效率并检测到更多数量性状核苷酸 (Quantitative trait nucleotides, QTNs)。Porter 等^[36]认为多性状 GWAS 分析与单性状 GWAS 分析相比，具有更高的遗传变异检测效率。本研究中有 7 个潜在关联基因同时与乌鳢体质量和全长性状表现出较高的关联性，说明在乌鳢的生长性状中也存在一因多效现象。后续研究可以考虑引入多性状联合分析方法，进一步提高检测效率和定位的准确性。

参考文献 (References):

- [1] Korte A, Farlow A. The advantages and limitations of trait analysis with GWAS: a review[J]. *Plant Methods*, 2013, 9(1): 29.
- [2] 何骞, 吴怡迪, 周志雄, 等. 大黄鱼肌纤维性状的全

基因组关联分析 [J]. 水产学报, 2022, 46 (8): 1313 – 1323.

He Q, Wu Y D, Zhou Z X, *et al.* Genome-wide association study of muscle fiber related traits in large yellow croaker (*Larimichthys crocea*) [J]. *Journal of Fisheries of China*, 2022, 46(8): 1313 – 1323.

- [3] 崔爱君, 徐永江, 王滨, 等. 黄条鲮生长性状全基因组关联分析 [J]. 渔业科学进展, 2021, 42 (2): 71 – 78.

Cui A J, Xu Y J, Wang B, *et al.* Genome-wide association analysis of growth traits in yellowtail kingfish (*Seriola lalandi*) [J]. *Progress in Fishery Sciences*, 2021, 42(2): 71 – 78.

- [4] 方家璐, 海佳薇, 周林燕, 等. 基于黄河鲤鱼体质量性状的全基因组选择模型评估 [J]. 大连海洋大学学报, 2024, 39 (3): 437 – 444.

Fang J L, Hai J W, Zhou L Y, *et al.* Evaluation of genome-wide selection model based on body weight trait of Yellow River carp (*Cyprinus carpio*) [J]. *Journal of Dalian Ocean University*, 2024, 39(3): 437 – 444.

- [5] 李娟, 陈红菊, 安丽, 等. 乌鳢种质及其养、繁殖技术研究进展 [J]. 浙江海洋大学学报 (自然科学版), 2024, 43 (2): 175 – 182.

Li X, Chen H J, An L, *et al.* Research progress on germplasm resources, breeding and reproduction of *Channa argus* [J]. *Journal of Zhejiang Ocean University (Natural Science)*, 2024, 43(2): 175 – 182.

- [6] 王桂芹, 高开进, 陈秀梅. 乌鳢营养研究与饲料产业发展的现状和趋势 [J]. 饲料工业, 2021, 42 (8): 1 – 9.

Wang G Q, Gao K J, Chen X M. Current situation and nutrition and development trend of feed industry of snakehead (*Channa argus*) [J]. *Feed Industry*, 2021, 42(8): 1 – 9.

- [7] 张新铨, 陈昆慈, 朱新平. 乌鳢、斑鳢及杂交种养殖研究现状 [J]. 广东农业科学, 2011, 38 (22): 132 – 134.

Zhang X C, Chen K C, Zhu X P. Current status of research on the cultivation of black snakehead, spotted snakehead, and hybrid species [J]. *Guangdong Agricultural Sciences*, 2011, 38(22): 132 – 134.

- [8] Wang X H, Liu H Y, Pang M X, *et al.* Construction of a high-density genetic linkage map and mapping of quantitative trait loci for growth-related traits in silver carp

- (*Hypophthalmichthys molitrix*)[J]. *Scientific Reports*, 2019, 9(1): 17506.
- [9] Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform[J]. *Bioinformatics*, 2009, 25(14): 1754 – 1760.
- [10] Okonechnikov K, Conesa A, García-Alcalde F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data[J]. *Bioinformatics*, 2016, 32(2): 292 – 294.
- [11] McKenna A, Hanna M, Banks E, *et al.* The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data[J]. *Genome Research*, 2010, 20(9): 1297 – 1303.
- [12] Danecek P, Auton A, Abecasis G, *et al.* The variant call format and VCFtools[J]. *Bioinformatics*, 2011, 27(15): 2156 – 2158.
- [13] Cingolani P, Platts A, Wang L L, *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff[J]. *Fly*, 2012, 6(2): 80 – 92.
- [14] Yoav B, Yosef H. Controlling the false discovery rate: A practical and powerful approach to multiple testing[J]. *Journal of the Royal Statistical Society: Series B (Methodological)*, 1995, 57(1): 289 – 300.
- [15] Hommel G. A stagewise rejective multiple test procedure based on a modified Bonferroni test[J]. *Biometrika*, 1988, 75(2): 383 – 386.
- [16] 黄苏静, 吴玉霞, 夏威夷, 等. 乌鳢三种性腺组织的转录组比较分析 [J]. *农业生物技术学报*, 2024, 32 (2): 379 – 395.
- Huang S J, Wu Y X, Xia W W, *et al.* Comparative transcriptome analysis of three types of gonads in northern snakehead (*Channa argus*)[J]. *Journal of Agricultural Biotechnology*, 2024, 32(2): 379 – 395.
- [17] 徐成豪. 乌鳢遗传连锁图谱构建及生长、体色和性别性状 QTL 定位 [D]. 上海: 上海海洋大学, 2023.
- Xu C H. Construction of genetic linkage map and QTL mapping of growth, body color and sex traits in *Channa argus*[D]. Shanghai: Shanghai Ocean University, 2023.
- [18] Luo Q, Ou M, Zhao J, *et al.* Expression profile and estrogenic regulation of Amh during gonadal sex differentiation in northern snakehead (*Channa argus*)[J]. *Genes & Genomics*, 2020, 42(7): 827 – 835.
- [19] Sun D L, Qi X, Wen H S, *et al.* The genetic basis and potential molecular mechanism of yellow-albino northern snakehead (*Channa argus*)[J]. *Open Biology*, 2023, 13(2): 220235.
- [20] Fan W, He Y, Su J, *et al.* Effects of leucism on organ development and molecular mechanisms in northern snakehead (*Channa argus*) beyond pigmentation alterations [J]. *Scientific Reports*, 2023, 13(1): 19689.
- [21] 赵金龙, 许少全, 周润, 等. 细粒棘球绦虫 DNA 损伤修复蛋白 Rad50 的生物信息学分析 [J]. *中国病原生物学杂志*, 2024, 19 (11): 1305 – 1311.
- Zhao J L, Xu S Q, Zhou R, *et al.* Preliminary validation and bioinformatics analysis of protein Rad50 in *Echinococcus granulosus*[J]. *Journal of Pathogen Biology*, 2024, 19(11): 1305 – 1311.
- [22] 马晓慧. 中缅边境地区 3 个黄牛品种全基因组遗传多样性研究 [D]. 咸阳: 西北农林科技大学, 2022.
- Ma X H. Whole-genome genetic diversity of three cattle breeds at the China-Burma border area[D]. Xianyang: Northwest A & F University, 2022.
- [23] 田雯雯. 氮氮对菲律宾蛤仔 miRNA 表达谱的影响及代表性 miRNAs 的响应特征研究 [D]. 烟台: 烟台大学, 2024.
- Tian W W. Study on the toxic effects of ammonia nitrogen on miRNA expression profiles in *Ruditapes philippinarum* and the responsive characteristic of typical miRNAs[D]. Yantai: Yantai University, 2024.
- [24] 顾婷婷, 陈莉. 肿瘤相关基因 NET-1 的研究进展 [J]. *现代肿瘤医学*, 2016, 24 (12): 1999 – 2002.
- Gu T T, Chen L. Progress of NET-1 gene in tumors[J]. *Modern Oncology*, 2016, 24(12): 1999 – 2002.
- [25] 吕梁, 张子平, 万海付, 等. 拟穴青蟹不同发育时期胚胎基因表达的内参基因的筛选 [J]. *中国水产科学*, 2019, 26 (3): 457 – 464.
- Lü L, Zhang Z P, Wan H F, *et al.* Identification of internal reference genes for quantification of gene expression in different embryonic developmental stages of the mud crab *Scylla paramamosain*[J]. *Journal of Fishery Sciences of China*, 2019, 26(3): 457 – 464.
- [26] 李尚阳, 周燕, 柴方超, 等. 香鱼白细胞介素-1 受体相关激酶 3 (IRAK-3) 的序列特征及对鳃弧菌感染的免疫相关性分析 [J]. *农业生物技术学报*, 2020, 28 (10): 1711 – 1721.
- Li S Y, Zhou Y, Chai F C, *et al.* Sequence characteristics of interleukin-1 receptor-related kinase 3 (IRAK-3) in ayu (*Plecoglossus altivelis*) and its immune correlation

- analysis against *Vibrio anguillarum* infection[J]. Journal of Agricultural Biotechnology, 2020, 28(10): 1711 – 1721.
- [27] 季策. 转录水平分析 Notch1a 参与副溶血弧菌感染斑马鱼幼鱼引起的天然免疫应答 [D]. 上海: 上海海洋大学, 2017.
- Ji C. Transcriptome analysis of Notch1a in innate immune response in zebrafish (*Danio rerio*) larvae challenged by *Vibrio parahaemolyticus* [D]. Shanghai: Shanghai Ocean University, 2017.
- [28] 柳芳, 李长忠, 祁洪芳, 等. Cu²⁺胁迫后青海湖裸鲤 *DNAJC2* 及其相关基因的表达分析 [J]. 基因组学与应用生物学, 2023, 42 (9): 962 – 972.
- Liu F, Li C Z, Qi H F, *et al.* Expression analysis of *DNAJC2* and related genes in *Gymnocypris przewalskii* under Cu²⁺ stress[J]. Genomics and Applied Biology, 2023, 42(9): 962 – 972.
- [29] 傅建军, 张猛, 沈玉帮, 等. 草鱼 *PRL* 基因多态性与幼鱼生长性状和肌肉成分的关联分析 [J]. 中国水产科学, 2016, 23 (3): 491 – 499.
- Fu J J, Zhang M, Shen Y B, *et al.* Polymorphisms of the *PRL* gene and their associations with growth traits and flesh components in juvenile grass carp, *Ctenopharyngodon idella* [J]. Journal of Fishery Sciences of China, 2016, 23(3): 491 – 499.
- [30] 成果, 温彬, 高建忠, 等. 七彩神仙鱼催乳素基因的克隆、定位及表达分析 [J]. 广东农业科学, 2024, 51 (4): 35 – 44.
- Cheng G, Wen B, Gao J Z, *et al.* Cloning, localization and expression of prolactin gene in discus fish (*Symphysodon aequifasciatus*) [J]. Guangdong Agricultural Sciences, 2024, 51(4): 35 – 44.
- [31] 吴碧银, 许建, 曹顶臣, 等. 鲤低氧适应性性状的全基因组关联分析 [J]. 渔业科学进展, 2022, 43 (2): 98 – 106.
- Wu B Y, Xu J, Cao D C, *et al.* Genome-wide association analysis of hypoxia adaptation traits in common carp (*Cyprinus carpio*) [J]. Progress in Fishery Sciences, 2022, 43(2): 98 – 106.
- [32] 徐鸿飞, 陈诏, 黄彩林, 等. 基于简化基因组测序的红罗非鱼低温体色变异全基因组关联分析 [J]. 西南农业学报, 2021, 34 (3): 667 – 672.
- Xu H F, Chen Z, Huang C L, *et al.* Genome-wide association analysis of low temperature body color variation in red tilapia based on RAD sequence [J]. Southwest China Journal of Agricultural Sciences, 2021, 34(3): 667 – 672.
- [33] 陈小明, 李佳凯, 王志勇, 等. 基于简化基因组测序的大黄鱼耐高温性状全基因组关联分析 [J]. 水生生物学报, 2017, 41 (4): 735 – 740.
- Chen X M, Li J K, Wang Z Y, *et al.* Genome-wide association study of thermal tolerance in large yellow croaker *Larimichthys crocea* based on SLAF-seq technology [J]. Acta Hydrobiologica Sinica, 2017, 41(4): 735 – 740.
- [34] 高进, 杨润清. 大菱鲆体重和体尺性状联合 GWAS 分析 [J]. 渔业科学进展, 2021, 42 (2): 63 – 70.
- Gao J, Yang R Q. Joint genome-wide association study of body mass and morphological traits in turbot (*Scophthalmus maximus*) [J]. Progress in Fishery Sciences, 2021, 42(2): 63 – 70.
- [35] 蒋仪, 宋禹昕, 高进, 等. 大菱鲆生长性状表型正交化混合模型关联分析 [J]. 中国水产科学, 2021, 28 (7): 845 – 851.
- Jiang Y, Song Y X, Gao J, *et al.* Multivariate mixed model association analysis for growth trait orthogonalization in turbot (*Scophthalmus maximus*) [J]. Journal of Fishery Sciences of China, 2021, 28(7): 845 – 851.
- [36] Porter H F, O'Reilly P F. Multivariate simulation framework reveals performance of multi-trait GWAS methods [J]. Scientific Reports, 2017, 7: 38837.

Genome-wide association study of growth traits in *Channa argus*

LI Xian¹, WANG Ya², CAO Xiao², ZHU Shuren¹, AN Li¹, ZHU Yong'an¹,
MENG Qinglei¹, DONG Wen¹, DONG Jun¹

(1. Shandong Freshwater Fisheries Research Institute, Jinan 250013, China;

2. Fisheries Development Service Center of Weishan County, Jining 277699, China)

Abstract: [Objective] *Channa argus* is an important freshwater economic fish in China. In recent years, due to environmental impacts and breeding methods, the genetic resources of *C. argus* have gradually deteriorated. This study aims to investigate the genetic basis of growth traits in *C. argus* using genome-wide association studies (GWAS) to provide a basis for the conservation of genetic resources and the selection of superior breeds in *C. argus*. [Methods] A total of 405 individuals of *C. argus* were subjected to reduced-representation genome sequencing using the Illumina Nova (PE150) platform and SuperGBS technology. The GWAS analysis was performed using the efficient mixed model of the EMMAX software to identify SNPs significantly associated with growth traits. Sequences within 50 kb upstream and downstream of the significant loci were scanned to identify potential candidate genes. The biological functions of the candidate genes were annotated based on the NCBI database and literature search results, and candidate functional genes related to the target traits were screened. [Results] After filtering the sequencing data, a total of 47 536 SNP loci were obtained. GWAS analysis identified three SNPs significantly associated with body mass traits: 2831060, 2830864, and 2830886. By scanning sequences within 50 kb upstream and downstream of the significant loci, seven candidate functional genes related to body mass in *C. argus* were annotated. In addition, six potential SNPs associated with total length traits were identified, and 17 potential candidate functional genes related to total length traits were annotated. Among these functional genes, seven were associated with both total length and body mass traits, namely *irak3*, *tuba1a*, *hspa14*, *prl*, *rint1*, *helb* and *net1*. These functional genes are mainly involved in biological processes such as growth metabolism, developmental regulation, cell proliferation, and immune regulation in *C. argus*. [Conclusion] The identified SNPs and functional genes are significantly associated with body mass and total length traits in *C. argus*, providing important references for the mechanism analysis of growth traits, superior breed selection, and resource conservation in *C. argus*.

Key words: *Channa argus*; genome-wide association study (GWAS); growth traits; SuperGBS