

多鳞鲢(*Sillago sihama*)高密度遗传连锁图谱构建及生长性状 QTL 定位分析*

田昌绪^{1,2} 朱奕安¹ 钟 键³ 林星桦¹ 叶明慧¹ 黄 洋^{1,2}
张玉蕾¹ 朱春华^{1,2} 李广丽^{1,2}

(1. 广东海洋大学水产学院 广东省名特优鱼类生殖调控与繁育工程技术研究中心 广东省水产动物病害防控与健康养殖重点实验室 广东湛江 524088; 2. 南方海洋科学与工程广东省实验室(湛江) 广东湛江 524088; 3. 湛江海关技术中心 广东湛江 524022)

摘要 为构建多鳞鲢(*Sillago sihama*)遗传连锁图谱并鉴定生长等重要经济性状数量性状位点(QTL), 实验通过基因分型测序(GBS)技术对 163 个多鳞鲢个体(2 个亲本和 161 个全同胞家系 F₁ 代)进行测序及标记分型, 并通过复合区间定位法对该物种的体重、体高、体厚、眼径、体长和背鳍前长 6 个生长性状进行 QTL 定位分析。结果显示, 多鳞鲢首张高密度遗传连锁图谱全长 2 154.803 cM, 标记间平均遗传距离 0.455 cM, 共有 4 735 个 SNP 标记分配到 24 个连锁群。QTL 定位分析结果发现在 6 个生长性状中共检测到 20 个生长显著相关 QTL 位点, 分布在 8 个连锁群上, 单个 QTL 的 LOD 值范围为 3.02~4.23, 可解释的表型变异范围为 0.14%~8.42%。其中, 在连锁群 LG08 聚集了 8 个生长性状显著相关的 QTL。通过对候选 QTL 区间内的基因进行功能注释, 共筛选到了 19 个潜在生长调控相关基因, 包含 *igf1*、*igf2*、*sstr5*、*sstr1a*、*tgfbr2*、*gas1*、*igfals*、*gfg6*、*gfg20*、*bmp7*、*kdm5c*、*ttil* 以及 *rbm10* 等。实验获得的遗传标记及相关候选基因是多鳞鲢生长相关性状标记辅助选择(MAS)的有用基因资源, 为进一步研究鱼类生长调控机制提供了更多的理论依据。

关键词 多鳞鲢; 基因分型测序(GBS); 数量性状位点(QTL); 生长相关性状; 候选基因

中图分类号 Q953; S917.4; S965 **doi:** 10.11693/hyhz20220500118

多鳞鲢(*Sillago sihama* Forsskål)又称沙锥鱼, 隶属鲈形目、鲷科、鲷属, 为热带印度-西太平洋浅海鱼类, 广泛分布于我国沿海地区(Lin *et al.*, 2021)。多鳞鲢肉质细嫩, 味道鲜美, 营养价值高, 具有重要的经济价值。该物种曾在我国近海渔业中占有重要地位, 近年来由于过度捕捞, 其自然个体低龄化、群体趋小化和产量逐年减少等问题突出(黄洋等, 2013), 市场售价已从 2003 年的 20~30 元/kg 持续上升到目前的

100~110 元/kg。本课题组前期突破了多鳞鲢全人工繁育技术, 实现了其全人工规模化育苗, 为下一步多鳞鲢良种选育奠定了基础(黄洋等, 2013)。然而, 多鳞鲢养殖苗种缺乏定向选育, 存在因近亲繁殖导致的生长速度减慢、抗逆性能下降等性状退化风险, 严重限制了其养殖业的健康发展, 亟需开展分子遗传育种研究, 解析重要经济性状相关遗传机制。近年来, 研究人员在多鳞鲢低氧应激(Tian *et al.*, 2020; Pan *et al.*,

* 广东省基础与应用基础研究基金, 2021A1515010733 号, 2019A1515110619 号; 广东海洋大学创新强校工程项目, 2019KTSCX060 号; 2021 年广东省科技创新专项资金, SDZX2021041 号; 广东省南美白对虾现代种业产业园项目, K22221 号; 广东海洋大学科研启动经费资助项目, R19026 号; 广东大学生科技创新培育专项资金资助项目, pdjh2022b0239 号; 2022 年国家级大学生创新创业训练计划项目, 202210566003 号; 海关总署 2021 年科技项目, 2021HK205 号。田昌绪, 博士, E-mail: tiancx@gdou.edu.cn

通信作者: 李广丽, 教授, E-mail: ligl@gdou.edu.cn

收稿日期: 2022-05-04, 收修改稿日期: 2022-07-04

2021)、营养饲料(Huang *et al*, 2020)、群体遗传(Liu *et al*, 2012)、人工繁育(余家旺等, 2022)以及基因组资源(Lin *et al*, 2021)等方面进行了探索研究, 但尚未见分子标记开发及生长性状相关 QTL 定位研究报道。

基于分子标记技术构建遗传连锁图谱是实现重要经济性状 QTL 定位和分子标记辅助育种的有效途径(桂建芳等, 2016)。高通量测序技术简化了遗传标记分型方式, 降低了实验成本, 使得在众多水产养殖物种中通过构建高密度遗传图谱进行性状 QTL 精细定位成为可能(叶华等, 2011)。鱼类生长性状是重要的经济性状, 对生长性状进行遗传改良是最有价值的选育目标之一。借助简化基因组测序技术, 截至 2020 年已有 40 多种鱼类构建了高密度遗传连锁图谱, 并开展了经济性状 QTL 定位研究(You *et al*, 2020), 如鳙(*Hypophthalmichthys nobilis*) (Liu *et al*, 2016)、鲤(*Cyprinus carpio*) (Peng *et al*, 2016)、褐石斑鱼(*Epinephelus bruneus*) (Kessuwan *et al*, 2016)、大口黑鲈(*Micropterus salmoides*) (Dong *et al*, 2019)、锦鲤(*Cyprinus carpio haematopterus*) (Feng *et al*, 2018)以及斑点叉尾鲷(*Ictalurus punctatus*) (Zhang *et al*, 2019), 鉴定出了一批生长性状相关的遗传标记及其候选基因(陈军平等, 2020)。

目前, 多鳞鱮生长相关性状的定位研究尚属空白, 本研究的目的在于利用基因分型测序(genotyping-by-sequencing, GBS)技术构建多鳞鱮高密度连锁图谱, 并进行 6 个生长性状的 QTL 定位分析, 以为多鳞鱮基因组组装、比较基因组学分析及分子标记辅助选择育种提供有力支撑, 并为该物种生长调控机制研究奠定基础。

1 材料与方法

1.1 作图家系及 DNA 提取

试验所用多鳞鱮亲本及其 F₁ 代全同胞家系养殖于广东省湛江市广东海洋大学东海岛繁育基地。2018 年 6 月, 挑选体型良好的性成熟雌雄个体为亲本, 一对一繁殖构建全同胞家系。孵化后的幼苗在水泥池(5.8 m×4.8 m×1.8 m)中进行养殖, 在整个养殖过程中, 水温(27±0.5) °C、pH 7.2±0.6 及溶解氧(7.1±0.5) mg/L。2019 年 6 月采集 161 尾 F₁ 代个体及父母本鳍条, 用于基因组 DNA 的提取; 同时, 采集 161 尾个体的体重、体高、体厚、眼径、体长和背鳍前长 6 个生长性状[方法见余家旺等(2022)]。随后, 将样本送至广州基迪奥生物科技有限公司使用 CTAB (十六烷基三甲基

溴化铵, 广州化学试剂厂)法提取基因组 DNA (李进波等, 2014); 并通过 Nanodrop 2000 (Thermo Scientific, 美国)和 1.5%琼脂糖凝胶电泳对提取的 DNA 进行质量检测。将符合试验要求的各样品 DNA 在广州基迪奥生物科技有限公司建库。

1.2 GBS 文库构建及高通量测序

多鳞鱮亲本及 161 个 F₁ 代基因组 DNA 测序采用 GBS 方法(Elshire *et al*, 2011)进行, 由广州基迪奥生物科技有限公司完成。质检合格的基因组 DNA 被两种限制性内切酶 *EcoRI* 和 *NlaIII* (New England Biolabs, Ipswich, MA)消化, 然后进行末端修复, 加 A-尾, 并使用 NEBNext[®] MLtra[™] DNA Library Prep Kit 库(NEB, USA)添加 Illumina 测序接头。选取 300~500 bp 的 DNA 片段进行 PCR 扩增富集, 最后使用 AMPure XP 系统(Beckman Coulter, Brea, CA, USA)对 PCR 产物进行纯化, 使用 Agilent 2100 生物分析仪(Agilent, Santa Clara, CA)对测序文库进行检测, 并使用实时 PCR 进行文库定量。测序在 Novaseq 6000 测序仪上进行, 采用 PE150 测序策略。

1.3 SNP 标记筛选与过滤

使用 FASTP(0.18.0) (Chen *et al*, 2018)对测序后的 raw data 进行 SNP 过滤。筛选标准为: (1) 去除含有未知核苷酸(N)≥10%的 reads; (2) 去除 phred 质量评分≤20 及碱基≥50%的 reads; (3) 删除含接头的 reads。过滤后的 clean reads 用于组装分析。使用 Burrows-Wheeler Aligner (BWA) (Li *et al*, 2009) (0.7.12; 比对参数为-k 32 -M)采用 mem 算法将过滤后的 reads 比对到参考基因组(SRA PRJNA642704); 比对后结果使用软件 picard (<http://sourceforge.net/projects/picard/>) (1.129)进行标记。使用变异检测软件 GATK (Van Der Auwera *et al*, 2013) (3.4~46) (设置参数: -Window 4, -G_filter “QD<2.0 || FS>60.0 || MQ<40.0”)进行群体 SNP 检测, SNP 标记过滤标准为: 去除分型比例低于 30%的位点, 去除杂合比例大于 75%的位点, 按理论比对标记位点的基因型比例进行卡方检验, *p* 值小于 0.001 的位点视为严重偏分离位点并去除, 保留分离类型为母本杂合型 lmxll、父本杂合型 nnxnp 和双亲杂合型 hkxhk 的标记。使用 ANNOVAR 软件(Wang *et al*, 2010)进行功能注释 SNP。

1.4 遗传图谱构建

利用筛选出来的 SNP 标记, 使用 Joinmap4.1 划分连锁群, 构建遗传图谱。对划分好的连锁群采用最大似然法进行排序, 对排序后的结果进行校正, 然后

再使用回归算法进行排序, 使用 Perl SVG 模型获得整合图谱(Van Ooijen, 2011)。整合图谱构建中, 以多鳞鱮基因组为参考, 根据标记的物理位置信息将其归并成对应的连锁群, 去除非染色体上的标记, 在每个遗传位置的标记中取完整度最高(缺失率最低)的标记作为作图标记。

1.5 生长相关性状 QTL 定位及候选基因预测

使用 R/qtl (Broman *et al.*, 2003)的 scanone()函数对 6 个生长性状进行复合区间定位(CIM), 扫描步长为 1 cM, 使用 LOD=3.0 作为阈值筛选 QTL。利用多鳞鱮参考基因组注释信息进行候选基因预测及功能注释。通过 BLASTN 将确定的显著性 SNP 位点比对到多鳞鱮参考基因组, 并根据基因组注释结果筛选

区间的基因, 并进行基因功能注释。

2 结果

2.1 生长相关性状的表型参数

本研究使用 161 尾多鳞鱮 F_1 代全同胞个体进行连锁图谱构建及 QTL 定位分析。课题组前期对该家系的生长性状相关性和途径分析显示, 多鳞鱮体高、体厚、眼径、体长和背鳍前长对多鳞鱮的体质量有显著影响, 初步确定为多鳞鱮的生长性状选育指标(余家旺等, 2022)。形态学分析显示, 6 个主要形态性状的变异系数(CV)值在 11.83%~33.51%, 表明性状变异丰富, 适合于进行 QTL 分析。6 个表型性状的频率分布符合正态分布(图 1), 可用于后续的 QTL 初步定位分析。

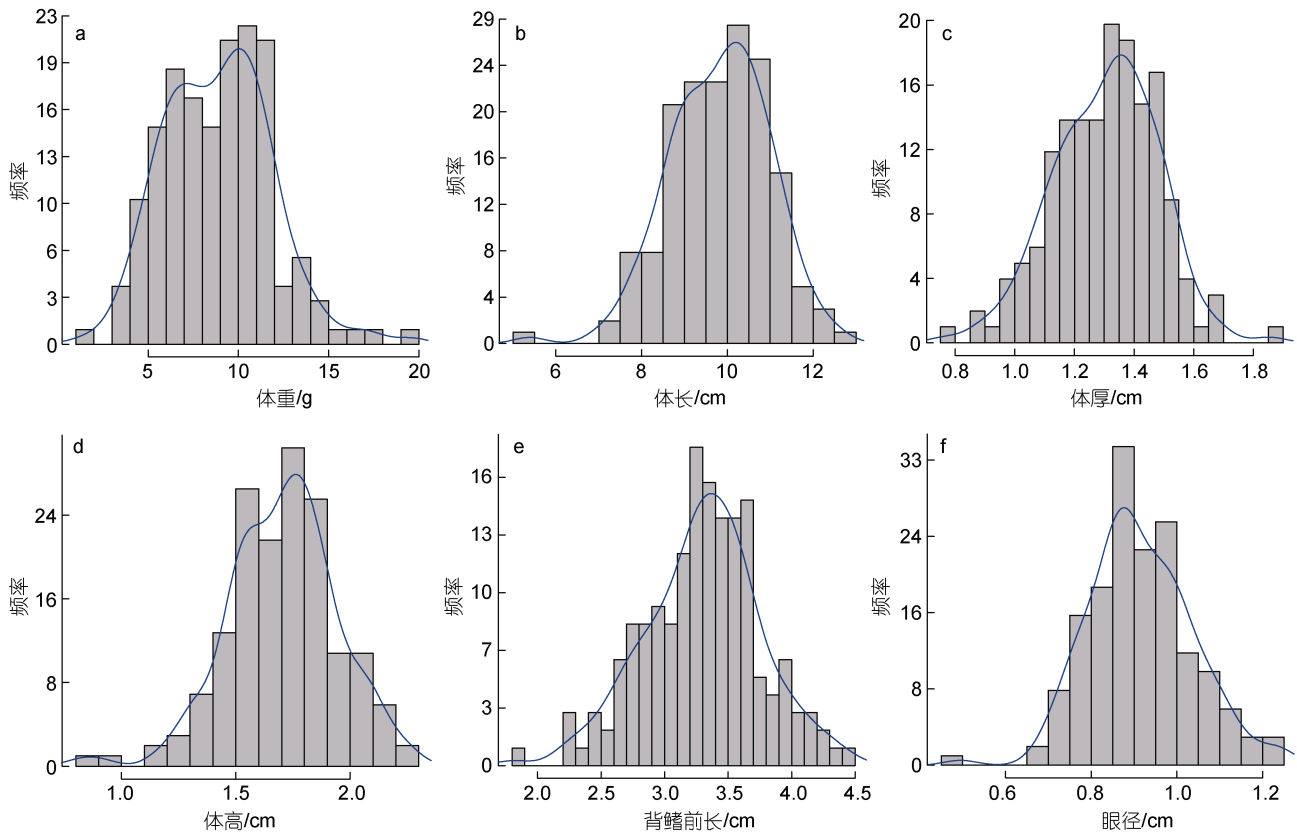


图 1 多鳞鱮 F_1 代全同胞家系体重(a)、体长(b)、体厚(c)、体高(d)、背鳍前长(e)及眼径(f)性状表型数据频率分布($n=161$)
Fig.1 Frequency distribution of phenotypic value of the body weight (a), body length (b), body thickness (c), body height (d), pre-dorsal length (e), and eye diameter (f) of *S. sihama* F_1 full-sib population ($n=161$)

2.2 多鳞鱮作图家系 GBS 测序数据及 SNP 筛选

对多鳞鱮父母本和 161 个 F_1 代进行 GBS 测序。两个亲本共获得 3.40 G bp Raw data (测序深度: $\sim 6.5\times$; 多鳞鱮基因组为 521.63 Mb), 过滤后亲本的平均 Clean data 为 1.60 G bp, 161 个子代样本共获得 880.55 G bp Raw data (平均每个个体 550 M bp), 过滤后平均每

个子代个体的 Clean data 为 534 M bp, 测序数据 $Q_{20} \geq 98\%$, $Q_{30} \geq 94\%$, 样本 GC 分布正常(平均 43%)。过滤后共获得 657.88 M Clean Reads, 平均每个个体 reads 长度为 142 bp。将测序数据与参考基因组进行比对, 比对结果显示, 所有 reads 与参考多鳞鱮基因组(SRA PRJNA642704)的平均比对率为 98%

以上, 用于后续的变异检测及相关分析。

群体变异检测共获得 205 471 个 SNP 位点, 这些位点中转换标记有 120 302 个, 占 58.54%, 颠换标记有 85 169 个, 占 41.46%。根据亲本的基因型确定标记的分离类型(亲本测序深度不低于 4), 保留分离类型为母本杂合型 *lmxl*、父本杂合型 *nnxn* 和双亲杂合型 *hkxhk* 的标记, 共筛选获得 143 886 个多态性 SNP 位点(表 1); 随后, 对标记进一步过滤, 去除分型比例低于 30%、或杂合率大于 75% 的位点、或严重偏分离的位点, 共保留 107 406 个高质量的 SNP 标记用于后续作图分析。

表 1 多鳞鱈作图群体中杂合子 SNP 位点的分离模式
Tab.1 Segregation patterns of heterozygotic SNP locus in the mapping population of *S. sihama*

分离类型	父本基因型	母本基因型	标记数目	百分比/%
hk×hk	hk	hk	22 364	16.00
lm×ll	lm	ll	59 104	41.00
nn×np	nn	np	62 418	43.00
总计			143 886	100.00

2.3 多鳞鱈遗传连锁图谱构建

以筛选获得的 107 406 个 SNP 位点为基础, 根据标记的物理位置信息划分连锁群, 去除非染色体上的标记, 并在每个遗传位置的标记中保留完整度最高的标记, 最终获得 4 735 个 SNP 标记, 被分配到 24 个连锁群, 与多鳞鱈单倍体染色体数目一致(表 2, 图 2)。该遗传连锁图谱全长 2 154.803 cM, 平均遗传距离 0.455 cM。

如表 2 所示, 24 个连锁群长度介于 65.059 cM (LG23)与 114.276 cM (LG08)之间, 平均长度 89.783 cM, 各连锁群标记平均遗传距离介于 0.383 cM (LG04)至 0.554 cM (LG22)之间。其中, LG01 连锁群分布有最多的 SNP 标记(232 个), 其连锁群长度为 112.766 cM, 标记平均距离为 0.486 cM; 而 LG23 上分布有最少的 SNP 标记(131 个), 其长度为 65.059 cM, 平均遗传距离为 0.497 cM。

2.4 生长相关性状 QTL 的定位

基于上述遗传图谱, 利用 R/qtl 软件对多鳞鱈体重、体高、体厚、体长、眼径、和背鳍前长 6 个生长性状进行 QTL 复合区间定位分析, 共定位到 20 个 QTLs (LOD>3.0), 分布在 LG04、LG05、LG07、LG08、LG16、LG17、LG18 以及 LG24 8 个连锁群(图 3)。单个 QTL 的 LOD 峰值范围为 3.02 (qGRT-15) ~ 4.23

(qGRT-13), 可解释的表型变异(PVE)范围为 0.14% (qGRT-7) ~ 8.42% (qGRT-12), 平均为 3.42% (表 3)。连锁群 LG08 上聚集了 8 个生长性状显著相关的 QTL, 其中 79.908~114.276 cM 区间的 6 个显著 QTL 与体重、体高、体长以及背鳍前长 4 个生长性状显著相关。

2.5 候选基因的预测

本研究通过与多鳞鱈参考基因组序列比对, 对 QTL 置信区间进行了基因查找, 共鉴定得到 19 个生长相关候选基因(表 3, 表 4)。相关基因包括组蛋白去甲基化酶(*kdm5c*)、跨膜蛋白 9 (*tmem9*)、Tel2 相互作用蛋白 1 (*ttil*)、RNA 结合基序蛋白 10 (*rbm10*)、胰岛素样生长因子 1 (*igf1*)、胰岛素样生长因子 2 (*igf2*)、转化生长因子 B 受体 2 (*tgfbr2*)、生长抑制特异性蛋白 1 (*gas1*)、类胰岛素生长因子酸不稳定亚基(*igfals*)、生长抑素受体 5 (*sstr5*)、生长抑素 1a (*sst1a*)、成纤维

表 2 多鳞鱈整合遗传图谱基本信息统计
Tab.2 Statistics of the consensus linkage map of *S. sihama*

连锁群	标记数	连锁群长度/cM	平均距离/cM	大于 5 cM 的 gap 数	最大 gap 距离/cM
LG01	232	112.766	0.486	1	5.613
LG02	203	99.516	0.490	0	4.021
LG03	222	86.521	0.390	0	2.764
LG04	218	83.448	0.383	0	1.228
LG05	203	100.428	0.495	1	5.614
LG06	199	94.002	0.472	0	4.337
LG07	203	99.142	0.488	0	3.075
LG08	230	114.276	0.497	0	4.654
LG09	201	83.456	0.415	0	2.456
LG10	168	80.460	0.479	0	4.337
LG11	201	97.999	0.488	0	4.654
LG12	188	82.541	0.439	0	2.461
LG13	227	106.270	0.468	1	9.303
LG14	198	89.941	0.454	0	3.380
LG15	211	81.609	0.387	0	1.536
LG16	189	79.469	0.420	0	2.456
LG17	167	80.754	0.484	0	4.022
LG18	223	92.019	0.413	1	8.264
LG19	165	68.731	0.417	0	1.536
LG20	212	98.225	0.463	0	3.997
LG21	180	73.945	0.411	0	1.843
LG22	174	96.466	0.554	1	9.303
LG23	131	65.059	0.497	0	2.764
LG24	190	87.760	0.462	0	2.153
总计	4735	2 154.803	0.455	5	9.303

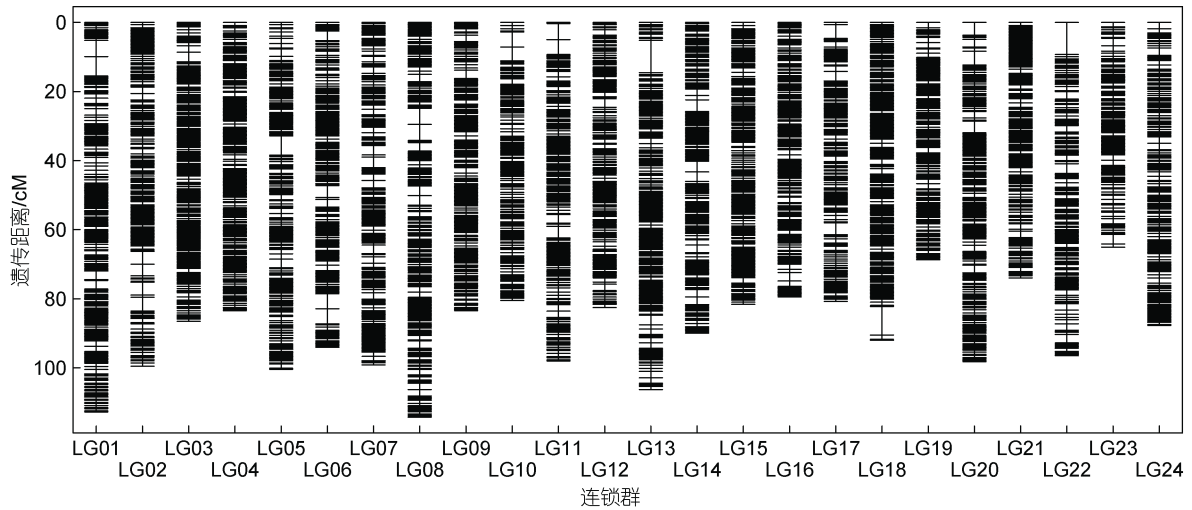


图 2 多鳞鱮高密度遗传连锁图谱

Fig.2 The high genetic linkage map of *S. sihama*

注: 遗传图谱展示了 24 条连锁群(LG01~LG24)的 SNP 标记分布及其遗传距离(cM)。遗传距离由左侧的比例尺表示, 单位为厘摩(cM)。连锁群上的单条线表示 SNP 标记

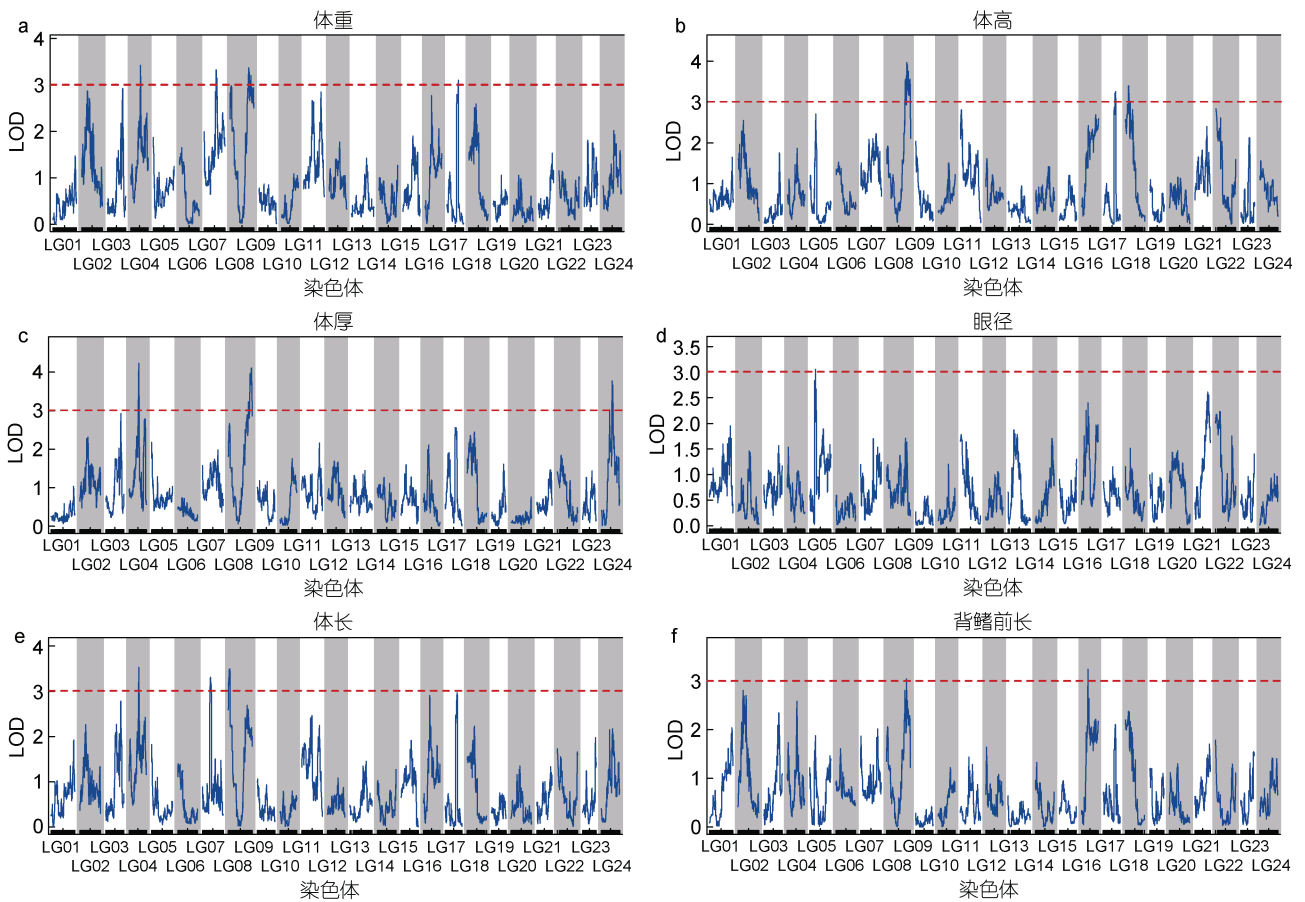


图 3 多鳞鱮体重(a)、体高(b)、体厚(c)、眼径(d)、体长(e)、背鳍前长(f)性状的 QTL 定位以及关联分析

Fig.3 The body weight (a), body height (b), body thickness (c), eye diameter (d), body length (e), and pre-dorsal length (f) mapping and association analysis of *S. sihama* among all linkage maps

注: LOD 曲线, x 轴和 y 轴分别对应所在染色体位置和 LOD 值。红色水平虚线表示 LOD 显著性阈值 3.0

表 3 复合区间定位法定位多鳞鱮生长性状相关 QTL 结果

Tab.3 The results of composite interval mapping for QTLs associated with growth traits of *S. sihama*

性状	QTL	连锁群	LOD	贡献率/%	遗传位置/cM	LOD 峰值遗传位置/cM	LOD 峰值对应标记	候选基因
体重	qGRT-1	LG04	3.42	2.31	43.260~50.009	46.634	LG3_10501634	—
	qGRT-2	LG07	3.33	4.36	52.497~66.925	57.713	LG6_9129684	—
	qGRT-3	LG08	3.37	0.85	79.908~98.625	88.498	LG7_20534125	<i>kdm5c</i>
	qGRT-4	LG08	3.21	1.47	90.031~110.290	100.158	LG7_21679598	—
	qGRT-5	LG17	3.11	3.23	45.763~57.115	55.000	cLG17.loc55	—
体高	qGRT-6	LG08	3.97	1.01	87.269~106.297	95.864	LG7_21084872	<i>tmem9</i>
	qGRT-7	LG08	3.56	0.14	100.463~114.276	111.211	LG7_22598757	<i>rbm10</i>
	qGRT-8	LG17	3.26	4.09	49.444~61.425	57.729	LG16_15413357	—
	qGRT-9	LG18	3.40	1.82	4.295~24.542	14.419	LG17_2181895	—
背鳍前长	qGRT-10	LG08	3.04	5.74	86.041~106.297	96.170	LG7_21093433	<i>ttil</i>
	qGRT-11	LG16	3.23	6.00	25.157~40.498	30.373	LG15_6742111	—
眼径	qGRT-12	LG05	3.06	8.42	21.789~32.219	27.005	LG4_3754362	—
体厚	qGRT-13	LG04	4.23	2.48	43.873~48.167	46.021	LG3_10319846	—
	qGRT-14	LG08	4.11	7.08	100.463~114.276	111.211	LG7_22598757	<i>rbm10</i>
	qGRT-15	LG24	3.02	0.58	32.222~50.025	39.286	LG23_5155271	—
	qGRT-16	LG24	3.77	2.02	50.025~57.388	51.559	LG23_7622430	—
	qGRT-17	LG04	3.52	3.89	43.873~50.008	46.328	LG3_10321212	—
体长	qGRT-18	LG07	3.31	4.27	33.142~43.592	38.369	LG6_4597304	—
	qGRT-19	LG08	3.21	1.74	0~10.738	0.614	LG7_1159083	—
	qGRT-20	LG08	3.49	1.98	0~17.182	6.000	cLG08.loc6	—

表 4 多鳞鱮20个 QTL 区间内已知生长相关基因信息

Tab.4 Growth-related genes in 20 QTL regions of 6 growth-related traits of *S. sihama*

基因 ID	连锁群	基因名缩写	基因全称
EVM0021738	LG08	<i>igf1</i>	insulin-like growth factor-1
EVM0006321	LG05	<i>igf2</i>	insulin-like growth factor-2
EVM0022219	LG17	<i>tgfbr2</i>	TGF-beta receptor type-2-like
EVM0019806	LG17	<i>gas1</i>	growth arrest-specific protein 1-like
EVM0017787	LG16	<i>igfals</i>	insulin-like growth factor-binding protein complex acid labile subunit
EVM0021175	LG16	<i>sstr5</i>	somatostatin receptor type five
EVM0015404	LG24	<i>sst1a</i>	somatostatin-1A-like
EVM0019956	LG05	<i>fgf6</i>	fibroblast growth factor 6
EVM0011026	LG04	<i>fgf20</i>	fibroblast growth factor 20-like
EVM0012898	LG08	<i>mtor</i>	serine/threonine-protein kinase mTOR
EVM0010667	LG24	<i>slc2a4</i>	solute carrier family 2, facilitated glucose transporter member 4-like
EVM0008074	LG08	<i>bmp7</i>	bone morphogenetic protein 7-like
EVM0002813	LG08	<i>ttil</i>	TELO2-interacting protein 1 homolog
EVM0008475	LG08	<i>wnt7a</i>	wingless-type MMTV integration site family, member 7a
EVM0019971	LG08	<i>wnt5a</i>	wingless-type MMTV integration site family, member 5a
EVM0002580	LG24	<i>p53</i>	tumor suppressor protein p53

细胞因子 6 (*gfg6*)、成纤维细胞因子 20 (*gfg20*)、哺乳动物雷帕霉素靶蛋白(*mtor*)、葡萄糖转运蛋白 4 (*slc2a4*)、骨形态发生蛋白 7 (*bmp7*)和肿瘤抑制蛋白 p53 (*p53*)等基因。上述候选基因为研究多鳞鱮生长调控提供了重要的基因资源,其功能有待深入研究。

3 讨论

作图群体的选择是构建遗传连锁图谱的先决条件。不同的作图家系类型各有利弊,主要包括重组近交系(recombined inbred lines, RIL)、单倍体(haploid, HAP)和双单倍体(doubled haploid, DH)、子一代(F_1)、子二代(F_2)和回交(back crossing, BC)等家系群体(Peng *et al.*, 2016)。通常在鱼类中构建 RIL、HAP 和 DH 家系较为困难,而 F_2 和 BC 家系的构建耗费时间长。基于养殖鱼类繁殖周期较长、基因组复杂、遗传多态性高、繁殖率高等特点,目前大部分水产物种采用拟测交策略的 F_1 作图家系构建遗传连锁图谱。除了作图家系的选择对遗传连锁图谱质量产生影响外,作图标记类型的选择也是直接影响遗传图谱质量的原因之一。与其他遗传标记相比,SNP 标记是基因组中最丰富的标记类型,具有高多态性。然而,在具有相对较大作图群体的非模式水产物种中,很难开发足够数量的 SNP (Tsigenopoulos *et al.*, 2014; Zhu *et al.*, 2014)。随着高通量测序技术的发展,各种基于基因组序列的 SNP 基因分型技术(包括 RAD、2b-RAD、GBS 和 SLAF 等)被开发出来,为构建高密度连锁作图提供了快速便捷的手段。GBS-seq 是在第二代测序基础上发展而来的一种基于全基因组酶切位点的基因分型测序技术(Elshire *et al.*, 2011)。GBS-seq 不仅操作简单,性价比高,更为重要的是它一次测序即可获得数以万计的多态性遗传标记,已经广泛用于群体遗传学、基因组学和遗传连锁图谱构建等研究领域(Guo *et al.*, 2021; Weng *et al.*, 2021)。本研究基于拟测交策略,对多鳞鱮 F_1 代作图群体进行基因分型测序,获得了大量分离比为 1:1 的 SNP 位点,成功构建了该物种高密度连锁图谱。多鳞鱮高密度连锁图谱全长 2 154.803 cM,分为 24 个连锁群,包括 4 735 个 SNP 标记,标记间平均遗传距离为 0.455 cM。近几年公布的鱼类遗传图谱中,鳊(*Siniperca chuatsi*)遗传图谱全长 1 694.3 cM,上图标记数 2 328 个,标记间平均遗传距离为 0.7 cM (Guo *et al.*, 2021)。黄姑鱼(*Nibea albiflora*)图谱全长 4 300.20 cM,上图标记数 6 219 个,标记间平均遗传距离为 0.69 cM (Xu *et al.*, 2021)。丝

尾鳊(*Hemibagrus wyckioides*)图谱全长 2 067.35 cM,上图标记数 2 369 个,标记间平均遗传距离为 0.87 cM (Zhou *et al.*, 2021)。与其他鱼类遗传图谱相比,多鳞鱮遗传图谱上图标记数量合适、标记密度高。此外,当遗传连锁图谱构建时的上图标记数量和分布合适时,遗传连锁图谱连锁群的数量应该等于该物种单倍体的染色体数目。多鳞鱮染色体水平基因组分析结果表明,多鳞鱮单倍型染色体数目 $N=24$ (Lin *et al.*, 2021)。在本研究中获得的遗传连锁图谱连锁群数量也为 24 条,与多鳞鱮单倍体染色体数目一致,这表明本研究构建的多鳞鱮遗传连锁图谱质量高。综上,本研究构建了高密度和高分辨率的遗传图谱,为后续 QTL 精细定位提供了重要工具。

QTL 分析有助于发现性状相关的连锁标记和预测候选基因。现有研究表明,鱼类生长性状多受微效多基因的调控。本研究中,在多鳞鱮的全同胞家系中定位到多个 QTL 位点且分布在不同的连锁群上,未能定位到生长性状相关主效位点,这表明多鳞鱮生长相关性状可能受到微效多基因的调控,跟已知的多数鱼类生长性状 QTL 定位结果相似。Zhang 等(2018)在布氏鲟鲂(*Trachinotus blochii*)中成功检测到了 23 个与体长、体高及体重等生长相关性状的 QTL。Zhou 等(2021)通过 QTL 分析显示,在 10 个连锁群上鉴定了丝尾鳊 4 个形态性状(体长、头长、体高和体重)的 17 个 QTL,其可解释的表型变异为 7.4%~13.3%。Zhang 等(2020)对瓦氏黄颡鱼(*Pelteobagrus vachelli*)体重相关的 8 个性状进行 QTL 分析,鉴定到 4 个生长性状显著相关的 QTL,其可解释的表型变异分别为 11.2%~12%。本研究获得的生长相关 QTL 的可解释的表型变异比值相比其他水产物种较低,平均为 3.42%,可能与选取的作图家系类型(F_1 全同胞家系)及其个体数目不足有关,在后续研究中,选用不同的作图群体类型(如 F_2 代群体)或增加作图群体数量或选用不同发育时期的个体可以直接和间接地改善可解释的表型变异,从而提高检测 QTL 的准确性,降低因假阳性而导致的错判。

位于连锁群 L08 上 79.908~114.276 cM 区间的 6 个显著 QTL 与体重、体高、体长以及背鳍前长 4 个生长性状显著相关,定位到的生长相关 QTL 具有成簇分布的特点,这表明连锁群 LG08 的上述区段可能与多鳞鱮的生长发育过程显著相关。鱼类生长相关 QTL 往往呈现成簇分布的特点,如 Wang 等(2018)在黄河鲤(*Cyprinus carpio haematoperus*)中发现,定位

到生长相关 QTL 呈现出成簇分布的特点; Yang 等 (2020)在金钱鱼(*Scatophagus argus*)中研究发现其体长和全长两个生长相关性状的大多数 QTL 均分布在连锁群 LG2 的重叠置信区间内。现有研究表明, 鱼类生长性状多受微效多基因的调控, 生长相关 QTL 在同条染色体上的成簇分布, 区域内的功能基因对生长相关性状的影响可能存在一因多效或不同的功能基因间互作影响。对多鳞鱮生长性状相关 QTL 的深入研究, 可明确各区域内基因位点的相互作用和生长性状间的调控机制, 有助于指导多鳞鱮的选育工作。

现有研究结果表明, 鱼类生长与发育受多基因调控, 已有物种的研究成果尚不足以解释鱼类生长调控机理, 有必要在更多的物种中开展鱼类生长调控相关基因的发掘研究。本研究中, 在显著 QTL 区域内鉴定出 22 个生长相关候选基因, 其中包括胰岛素样生长因子 1 (*igf1*)、胰岛素样生长因子 2 (*igf2*)、类胰岛素生长因子酸不稳定亚基(*igfals*)、生长抑制特异性蛋白 1 (*gas1*)、生长抑素 1a (*sst1a*)、生长抑素受体 5 (*sstr5*)、成纤维细胞因子 6 (*fgf6*)、转化生长因子 B 受体 2 (*tgfbr2*)、成纤维细胞因子 20 (*fgf20*)以及骨形态发生蛋白 7 (*bmp7*)等已知参与调控鱼类生长发育的功能基因。此外, 在连锁群 L08 79.908~114.276 cM 区间注释到组蛋白去甲基化酶(*kdm5c*)、Tel2 相互作用蛋白 1 (*ttil*)及 RNA 结合基序蛋白 10 (*rbm10*)等候选基因。KDM5 家族蛋白是激活大量促增殖细胞周期基因所必需的, 在神经发生、DNA 损伤反应和癌症发展等方面发挥重要作用(Shen *et al*, 2021)。*ttil* 基因参与促进哺乳动物组装、稳定和维持 mTORC1 和 mTORC2 复合物的活性, 这些复合物调节细胞生长和存活, 以响应营养和激素信号(Kaizuka *et al*, 2010)。*rbm10* 基因参与调节前体 mRNA 的选择性剪接和 mRNA 稳定, 该基因的过度表达已被证明可抑制肺腺癌的恶性增殖(Li *et al*, 2020)。上述候选基因为研究多鳞鱮生长调控提供了重要的基因资源, 为该物种的为分子遗传育种研究提供有价值的参考信息。

4 结论

本研究基于多鳞鱮高密度遗传连锁图谱进行生长性状的 QTL 初步分析及候选基因的预测, 获得 20 个生长相关性状的显著 QTL 和多个生长调控相关的候选基因。多鳞鱮的 QTL 初步定位及候选基因的初步鉴定将为该物种遗传育种研究提供有用信息和奠定基础, 而在之后的 F₂ 群体中可利用不同养殖环境

的多个家系进行全面而精确的 QTL 分析, 以获取更稳定、精确的生长相关性状 QTL。

参 考 文 献

- 叶华, 王志勇, 2011. 水产动物遗传连锁图谱构建和 QTL 研究现状[J]. 海洋科学, 35(1): 105-110.
- 李进波, 盛婧, 李想, 等, 2014. 五种 DNA 提取方法对鱼加工制品 DNA 提取效果的比较[J]. 生物技术通报(4): 43-49.
- 余家旺, 王耀嵘, 林星桦, 等, 2022. 12 月龄多鳞鱮形态性状对体质量的影响[J]. 广东海洋大学学报, 42(1): 137-143.
- 陈军平, 胡玉洁, 王磊, 等, 2020. 鱼类遗传连锁图谱构建及 QTL 定位的研究进展[J]. 水产科学, 39(4): 620-630.
- 桂建芳, 包振民, 张晓娟, 2016. 水产遗传育种与水产种业发展战略研究[J]. 中国工程科学, 18(3): 8-14.
- 黄洋, 杜涛, 黄海立, 2013. 多鳞鱮*Sillago sihama* Forskál 人工繁殖研究[J]. 广东海洋大学学报, 33(1): 15-21.
- BROMAN K W, WU H, SEN Š, *et al*, 2003. R/qtl: QTL mapping in experimental crosses [J]. Bioinformatics, 19(7): 889-890.
- CHEN S F, ZHOU Y Q, CHEN Y R, *et al*, 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor [J]. Bioinformatics, 34(17): i884-i890.
- DONG C J, JIANG P, ZHANG J F, *et al*, 2019. High-density linkage map and mapping for sex and growth-related traits of largemouth bass (*Micropterus salmoides*) [J]. Frontiers in Genetics, 10: 960.
- ELSHIRE R J, GLAUBITZ J C, SUN Q, *et al*, 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species [J]. PLoS One, 6(5): e19379.
- FENG X, YU X M, FU B D, *et al*, 2018. A high-resolution genetic linkage map and QTL fine mapping for growth-related traits and sex in the Yangtze River common carp (*Cyprinus carpio haematopterus*) [J]. BMC Genomics, 19(1): 230.
- GUO W J, HE S, LIANG X F, *et al*, 2021. A high-density genetic linkage map for Chinese perch (*Siniperca chuatsi*) using 2.3K genotyping-by-sequencing SNPs [J]. Animal Genetics, 52(3): 311-320.
- HUANG Q C, ZHANG S, DU T, *et al*, 2020. Modulation of growth, immunity and antioxidant-related gene expressions in the liver and intestine of juvenile *Sillago sihama* by dietary vitamin C [J]. Aquaculture Nutrition, 26(2): 338-350.
- KAIZUKA T, HARA T, OSHIRO N, *et al*, 2010. Ttil and Tel2 are critical factors in mammalian target of rapamycin complex assembly [J]. Journal of Biological Chemistry, 285(26): 20109-20116.
- KESSUWAN K, KUBOTA S, LIU Q, *et al*, 2016. Detection of growth-related quantitative trait loci and high-resolution genetic linkage maps using simple sequence repeat markers in the kelp grouper (*Epinephelus bruneus*) [J]. Marine Biotechnology, 18(1): 57-84.
- LI H, DURBIN L R, 2009. Fast and accurate short read alignment with Burrows-Wheeler transform [J]. Bioinformatics, 25(14): 1754-1760.
- LI Z Q, XUE Q, XU J X, *et al*, 2020. The role of RBM10 mutations in the development, treatment, and prognosis of

- lung adenocarcinoma [J]. *Cell Cycle*, 19(21): 2918-2926.
- LIN X H, HUANG Y, JIANG D N, *et al*, 2021. Chromosomal-level genome assembly of silver sillago (*Sillago sihama*) [J]. *Genome Biology and Evolution*, 13(2): evaa272.
- LIU H Y, FU B D, PANG M X, *et al*, 2016. QTL fine mapping and identification of candidate genes for growth-related traits in bighead carp (*Hypophthalmichthys nobilis*) [J]. *Aquaculture*, 465: 134-143.
- LIU D, GUO Y S, WANG Z D, *et al*, 2012. Phylogenetics inferred from mitogenome and control region of Silver Sillago, *Sillago sihama* [J]. *Mitochondrial DNA*, 23(4): 255-263.
- PAN Y Y, LIN X H, CHEN F Y, *et al*, 2021. Genome-wide identification and expression profiling of glutathione S-transferase family under hypoxia stress in silver sillago (*Sillago sihama*) [J]. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, 40: 100920.
- PENG W Z, XU J, ZHANG Y, *et al*, 2016. An ultra-high density linkage map and QTL mapping for sex and growth-related traits of common carp (*Cyprinus carpio*) [J]. *Scientific Reports*, 6(1): 26693.
- SHEN H F, ZHANG W J, HUANG Y, *et al*, 2021. The dual function of KDM5C in both gene transcriptional activation and repression promotes breast cancer cell growth and tumorigenesis [J]. *Advanced Science*, 8(9): 2004635.
- TSIGENOPOULOS C S, LOURO B, CHATZIPLIS D, *et al*, 2014. Second generation genetic linkage map for the gilthead sea bream *Sparus aurata* L [J]. *Marine Genomics*, 18: 77-82.
- TIAN C X, LIN X H, SAETAN W, *et al*, 2020. Transcriptome analysis of liver provides insight into metabolic and translation changes under hypoxia and reoxygenation stress in silver sillago (*Sillago sihama*) [J]. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, 36: 100715.
- VAN DER AUWERA G A, CARNEIRO M O, HARTL C, *et al*, 2013. From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline [J]. *Current Protocols in Bioinformatics*, 43(1110): 11.10.1-11.10.33.
- VAN OOIJEN J W, 2011. Multipoint maximum likelihood mapping in a full-sib family of an outbreeding species [J]. *Genetics Research*, 93(5): 343-349.
- WANG X H, FU B D, YU X M, *et al*, 2018. Fine mapping of growth-related quantitative trait loci in Yellow River carp (*Cyprinus carpio haematooperus*) [J]. *Aquaculture*, 484: 277-285.
- WANG K, LI M Y, HAKONARSON H, 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data [J]. *Nucleic Acids Research*, 38(16): e164.
- WENG Z Y, YANG Y, WANG X, *et al*, 2021. Parentage analysis in giant grouper (*Epinephelus lanceolatus*) using microsatellite and SNP markers from genotyping-by-sequencing data [J]. *Genes*, 12(7): 1042.
- XU D D, ZHANG W C, CHEN R Y, *et al*, 2021. Chromosome-scale assembly and high-density genetic map of the yellow drum, *Nibea albiflora* [J]. *Scientific Data*, 8(1): 268.
- YANG W, WANG Y R, JIANG D N, *et al*, 2020. ddRADseq-assisted construction of a high-density SNP genetic map and QTL fine mapping for growth-related traits in the spotted scat (*Scatophagus argus*) [J]. *BMC Genomics*, 21(1): 278.
- YOU X X, SHAN X X, SHI Q, 2020. Research advances in the genomics and applications for molecular breeding of aquaculture animals [J]. *Aquaculture*, 526: 735357.
- ZHANG G S, LI J, ZHANG J J, *et al*, 2020. A high-density SNP-based genetic map and several economic traits-related loci in *Pelteobagrus vachelli* [J]. *BMC Genomics*, 21(1): 700.
- ZHANG S Y, ZHANG X H, CHEN X H, *et al*, 2019. Construction of a high-density linkage map and QTL fine mapping for growth- and sex-related traits in channel catfish (*Ictalurus punctatus*) [J]. *Frontiers in Genetics*, 10: 251.
- ZHANG G Q, ZHANG X H, YE H Z, *et al*, 2018. Construction of high-density genetic linkage maps and QTL mapping in the golden pompano [J]. *Aquaculture*, 482: 90-95.
- ZHOU Y L, WANG Z W, GUO X F, *et al*, 2021. Construction of a high-density genetic linkage map and fine mapping of QTLs for growth and sex-related traits in red-tail catfish (*Hemibagrus wyckiioides*) [J]. *Aquaculture*, 531: 735892.
- ZHU C, TONG J, YU X, *et al*, 2014. A second-generation genetic linkage map for bighead carp (*Aristichthys nobilis*) based on microsatellite markers [J]. *Animal Genetics*, 45(5): 699-708.

CONSTRUCTION OF A HIGH-DENSITY GENETIC LINKAGE MAP AND QTL DETECTION OF GROWTH TRAITS OF SILVER SILLAGO (*SILLAGO SIHAMA*)

TIAN Chang-Xu^{1,2}, ZHU Yi-An¹, ZHONG Jian³, LIN Xing-Hua¹, YE Ming-Hui¹, HUANG Yang^{1,2},
ZHANG Yu-Lei^{1,2}, ZHU Chun-Hua^{1,2}, LI Guang-Li^{1,2}

(1. Guangdong Provincial Key Laboratory of Aquatic Animal Disease Control and Healthy Culture, Guangdong Research Center on Reproductive Control and Breeding Technology of Indigenous Valuable Fish Species, Fisheries College, Guangdong Ocean University, Zhanjiang 524088, China; 2. Southern Marine Science and Engineering Guangdong Laboratory (Zhanjiang), Zhanjiang 524088, China; 3. Zhanjiang Customs District Technology Center, Zhanjiang 524022, China)

Abstract Construction of high-density genetic map and quantitative trait loci (QTL) mapping are powerful tools for identifying genetic markers and candidate genes that may be responsible for such polygenic trait as growth. The first SNP-based high-density genetic linkage map was constructed by sequencing 163 silver sillago (*Sillago sihama*) individuals (2 parents and 161 F₁ offspring) according to a genotyping-by-sequencing (GBS) method. The consensus map spanned 2 154.803 cM, on average marker interval of 0.455 cM. In total, 4 735 SNPs were assigned to 24 linkage groups (LGs). Then, the QTL of 6 growth related traits was mapped via composite interval mapping (CIM), including body weight, body length, body thickness, body height, pre-dorsal length, and eye diameter. Twenty significant QTLs were identified on 8 LGs and explained 0.14%~8.42% of the phenotypic variance. The logarithm of odds (LOD) value ranged from 3.02 to 4.23. Specially, 8 QTLs were distributed on one linkage group (LG08), and the regions showed overlapping on LG08. Through the functional annotation of the genes in the candidate QTL interval, 19 potential growth-related genes were screened, including *igf1*, *igf2*, *sstr5*, *sst1a*, *tgfbr2*, *gas1*, *igfals*, *gfg6*, *gfg20*, *bmp7*, *kdm5c*, *tti1*, and *rbm10*. These genetic markers and candidate genes are useful genomic resources for marker-assisted selection (MAS) in silver sillago and the QTLs are useful tools for growth mechanism analysis of this fish.

Key words *Sillago sihama*; genotyping-by-sequencing (GBS); quantitative trait loci (QTL); growth-related traits; candidate genes