

## 赤点石斑鱼抗神经坏死病的基因组选择评估

王卓标, 方铭, 郑乐云, 葛辉, 陈欣欣, 罗辉玉, 王志勇

### Evaluation of genomic selection against neuronecrosis in the red-spotted grouper

WANG Zhuobiao, FANG Ming, ZHENG Leyun, GE Hui, CHEN Xinxin, LUO Huiyu, WANG Zhiyong

在线阅读 View online: <https://doi.org/10.14012/j.cnki.fjsc.2022.03.001>

#### 您可能感兴趣的其他文章

##### Articles you may be interested in

##### [赤点石斑鱼♀与云纹石斑鱼♂杂交F1形态特征和生长性能的研究](#)

Studies on the morphological features and the growth performance of hybrid F1 of *Epinephelus akaara* (♀) × *Epinephelus moara* (♂)  
渔业研究. 2015, 37(2): 148 <https://doi.org/10.14012/j.cnki.fjsc.2015.02.008>

##### [闽南地区斜带石斑鱼\(\*Epinephelus coioides\*\)神经坏死病毒的基因型与分子进化研究](#)

The genotyping and molecular evolution analysis of nervous necrosis virus (NNV) in infected *Epinephelus coioides* in southern Fujian  
渔业研究. 2017, 39(6): 429 <https://doi.org/10.14012/j.cnki.fjsc.2017.06.002>

##### [神经坏死病毒疫苗免疫后3种免疫相关基因在斜带石斑鱼\(\*Epinephelus coioides\*\)组织中的表达研究](#)

Expression patterns of three immune-related genes in tissues of *Epinephelus coioides* after injecting nerve necrosis virus vaccine  
渔业研究. 2019, 41(6): 470 <https://doi.org/10.14012/j.cnki.fjsc.2019.06.002>

##### [斜带石斑鱼\(\*Epinephelus coioides\*\)病毒性神经坏死病毒的纯化分析及检测方法建立](#)

The purification and detection methods research of viral nervous necrosis in orange-spotted grouper *Epinephelus coioides*  
渔业研究. 2019, 41(2): 96 <https://doi.org/10.14012/j.cnki.fjsc.2019.02.002>

##### [海马齿生态浮床对石斑鱼生长及养殖效益的模拟研究](#)

Simulated study on the effect of ecological floating bed of *Sesuvium portulacastrum* L. on the growth and aquaculture benefit of grouper  
渔业研究. 2020, 42(5): 411 <https://doi.org/10.14012/j.cnki.fjsc.2020.05.001>

王卓标,方 铭,郑乐云,等.赤点石斑鱼抗神经坏死病的基因组选择评估[J].渔业研究,2022,44(3):205-212.

# 赤点石斑鱼抗神经坏死病的基因组选择评估

王卓标<sup>1</sup>, 方 铭<sup>1</sup>, 郑乐云<sup>2</sup>, 葛 辉<sup>2</sup>, 陈欣欣<sup>2</sup>, 罗辉玉<sup>2</sup>, 王志勇<sup>1\*</sup>

(1. 集美大学水产学院, 农业农村部东海海水健康与养殖重点实验室, 福建 厦门 361021;  
2. 福建省水产研究所, 福建省海洋生物增殖与高值化利用重点实验室, 福建 厦门 361013)

**摘要:** 赤点石斑鱼 (*Epinephelus akaara*) 是中国东南沿海地区一种重要的海水养殖鱼类, 多年来其苗种生产深受神经坏死病的困扰, 严重制约了其人工养殖业的发展。本研究采集 230 尾赤点石斑鱼神经坏死病毒 (RGNNV) 易感 (染病死亡) 和 230 尾抗性 (最终健康存活) 赤点石斑鱼苗进行基因组重测序, 分析获得 5 412 683 个单核苷酸多态性 (SNPs) 位点的基因型, 并以之对抗病性状进行了遗传评估和基因组选择研究, 获得的估计遗传力均值为 0.566 2, 预测基因组估计育种值 (GEBV) 均值为 0.154 3。随机选择不同数量标记对基因组选择准确性影响进行评估, 结果表明, 采用  $\geq 5\ 000$  个 (5 k) 标记进行赤点石斑鱼抗神经坏死病性状遗传评估就可以获得比较理想的效果。本研究为开展赤点石斑鱼抗神经坏死病基因组选择育种实践提供了有用的理论参考。

**关键词:** 赤点石斑鱼; 基因组选择; 基因组最佳线性无偏预测 (GBLUP)

**中图分类号:** S943 **文献标识码:** A **文章编号:** 2096-9848(2022)03-0205-08

赤点石斑鱼 (*Epinephelus akaara*) 是一种高经济价值的海洋鱼类, 主要分布于西太平洋地区包括中国东海、南海以及日本和韩国的南部海域<sup>[1]</sup>。20 世纪 60 年代日本学者就研究了赤点石斑鱼的产卵习性和早期生活史<sup>[2]</sup>, 我国在 20 世纪 80 年代取得赤点石斑鱼的人工育苗成功<sup>[3-4]</sup>, 随后在福建、广东等沿海地区开始了网箱养殖。然而多年来赤点石斑鱼的苗种生产受到神经坏死病 (Viral nervous necrosis, VNN) 的严重困扰, 导致育苗的成活率与成功率都很低, VNN 成为制约赤点石斑鱼人工养殖业发展的一个重要因素。

鱼类病毒性神经坏死病是由神经坏死病毒

(Nervous necrosis virus, NNV) 引起的一种全球范围的鱼类流行性传染病。NNV 为海水鱼中较常见、危害严重的传染病之一, 至今已报告的受害鱼类有鳗鲡目 (Anguilliformes)、鲈形目 (Perciformes)、鲽形目 (Pleuronectiformes)、鲉形目 (Tetraodontiformes)、鳕形目 (Gadiformes) 中的 40 余种<sup>[5-6]</sup>, 其中被感染的种类集中在石斑鱼、鲈鱼, 在中国受影响最大的是赤点石斑鱼和斜带石斑鱼。对石斑鱼而言, 神经坏死病的高发期为鱼类的幼鱼阶段, 孵化后的 1~3 周为发病高峰期, 严重时发病率可达 100%<sup>[7]</sup>, 幼鱼成活率低于 10%。神经坏死病的病症通常表现为鱼苗在水中以螺旋状旋转为主的异常游动, 伴随

收稿日期: 2022-04-22

基金项目: 福建省科技计划项目 (2019R1013); 海水鱼产业技术体系项目 (CARS-47-G04)。

作者简介: 王卓标 (1997-), 男, 硕士研究生, 从事水产动物遗传与育种研究. E-mail: 842207943@qq.com

通信作者: 王志勇 (1963-), 男, 教授, 博士生导师, 从事水产动物遗传与育种研究. E-mail: zywang@jmu.edu.cn

着食欲减退, 静止时腹部向上。组织学检测发现细胞空泡化主要集中发生于中枢神经系统细胞以及视网膜<sup>[8]</sup>。受感染的幼鱼绝大多数在短期内死亡, 因此常常导致人工育苗失败, 或成活率极低。近年来, 随着养殖密度的不断提高和受感染鱼类种类增加, 其危害程度愈发严重。

选育抗病品种是解决养殖鱼类病害问题的一个有效途径<sup>[9-10]</sup>。但是传统的选育方法进展慢, 效果较差。2001 年 Meuwissen T H E 等<sup>[11]</sup>提出了基因组选择 (Genomic selection, GS) 的方法, 该方法具有不需构建家系、育种值估计的准确性高、育种效率高, 并可以有效控制近交等多方面优点。如今, 随着高通量 DNA 测序成本不断降低, 已有越来越多的水产动物育种开始使用基因组选择的方法<sup>[12-18]</sup>。目前已经有一些研究者将基因组选择应用于鱼类抗病育种研究, 如 Tsai H Y 等<sup>[19]</sup>报道了对大西洋鲑抗海虱、Liu Y 等<sup>[20]</sup>报道了对牙鲆抗爱德华氏菌的基因组选择研究。Palaiokostas C 等<sup>[21]</sup>对欧洲鲈鱼 (*Dicentrarchus labrax*) 的神经坏死病毒病抗性进行评估, 发现使用基因组选择的方法与比系谱选育预测能力增加 13%。福建省水产研究所石斑鱼研究团队已经完成了赤点石斑鱼全基因组测序组装<sup>[22]</sup>, 华南农业大学 Yang M 等<sup>[23]</sup>通过对 100 尾赤点石斑鱼神经坏死病毒 (Red-spotted grouper nervous necrosis virus, RGNNV) 易感与抗性石斑鱼进行全基因组关联分析, 找到了一些抗病相关的单核苷酸多态性 (Single nucleotide polymorphisms, SNPs) 位点; 但迄今还没有见到对赤点石斑鱼抗神经坏死病基因组选择研究的报道。本研究对 460 尾赤点石斑鱼神经坏死病易感 (染病死亡, 230 尾) 和抗性 (最终健康存活, 230 尾) 鱼苗, 通过基因组重测序获得高密度的 SNPs 集, 进行抗病遗传力评估和基因组选择预测, 以期为后续的抗病育种实践提供必要的理论参考。

## 1 材料与方法

### 1.1 样本的采集与鉴定

实验材料采集于厦门刘五店, 该育苗场在 2019 年赤点石斑鱼育苗生产中遭遇了神经坏死病, 分别在发病时期采集已染病的濒死个体作为易感组 (共 230 尾), 渡过发病期后, 采集同样数量存活的健康个体作为抗病组。对 460 尾鱼苗采集全鱼, 保存于 95% 乙醇中, 用于 DNA 提取。发病鱼苗病症除了根据其病状判断外, 还提取 30 尾病鱼头部组织 DNA 制成 10 个混合样品, 用 RGNNV 特异性检测引物 (正向引物: 5' - CACCCGCTTTGCAATCACAATG - 3', 反向引物: 5' - GTCATCAACGATACGCACTAGG - 3') 进行了 PCR 扩增验证 (结果均为阳性)。

### 1.2 基因分型和质量控制

使用南京诺唯赞生物科技股份有限公司的快速组织基因组 DNA 试剂盒提取鳍条组织基因组 DNA, 进行质检和建库后, 在 Illumina NovaSeq 6000 平台 (Illumina, USA) 进行 WGS 测序。其中 450 尾的目标测序深度为 4 ×, 另随机挑选 10 尾测序深度为 20 ×, 用于提供高质量的 SNPs 变异参考。首先使用 fastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc>) 对测序数据进行质量检测, 后使用 fastp<sup>[24]</sup> 对测序数据进行过滤。最后使用 MultiQC<sup>[25]</sup> 对最终的质控结果进行汇总检查。

通过 BWA-MEM<sup>[26]</sup> 将 clean reads 比对到赤点石斑鱼基因组<sup>[22]</sup>上, 将产生的文件利用 Samtools 进行排序并转化为 bam 文件格式。之后使用 sambamba<sup>[27]</sup> 对 bam 中构建文库时的 PCR 重复进行标记, 对标记重复后的 bam 文件使用 GATK<sup>[28]</sup> 中的 “HaplotypeCaller” 的模块进行单个样本变异的检测, 再通过 “CombineGVCFs” 将单个的变异集整合为群体的 VCF (Variant call format) 文件。使用 “SelectVariants” “Variant-Filtration” 模块进行硬过滤以及双等位基因的提取, 并使用 BCFtools<sup>[29]</sup> 对缺失率大于 20% 的变异进行过滤, 然后对缺失的基因型使用 Beagle<sup>[30]</sup> 进行填充。最后通过 PLINK<sup>[31]</sup> 对 VCF 进行过滤以及格式转换, 过滤标准为: 1) 次等位基因频率 MAF > 0.05; 2) HWE < 1e<sup>-6</sup>。最终获得 460 个样本的高质量 SNPs 数据用于后续分析。

### 1.3 群体结构分析

利用 PLINK 的 PCA 模块进行主成分分析 (Principal component analysis, PCA)。根据 PCA 结果所占比重前 2 个主成分 PC1 ~ PC2 的数据, 通过在线绘图工具 bioinformatics (<http://www.bioinformatics.com.cn>) 绘制 PCA 图。

#### 1.4 遗传参数估计

使用 GCTA<sup>[32]</sup> 和 R 语言的 EMMREML 包 (<https://CRAN.R-project.org/package=EMMREML>) 对遗传参数进行估计以及后续基因组育种值 (Genomic estimated breeding value, GEBV) 的计算。基因组最佳线性无偏预测 (Genomic best linear unbiased prediction, GBLUP) 模型公式为:

$$y = X\beta + Z\mu + e \quad (1)$$

其中  $y$  代表的是性状,  $Z$  是根据 SNPs 效应而设计的矩阵 (基因型 “AA”、“Aa” 和 “aa” 分别对应着 SNPs 基因型中的 “0”、“1” 和 “2”),  $\beta$  是固定效应 (地点效应), 混合线性模型也可以展开表示为:

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + G^{-1}\lambda \end{bmatrix} \begin{bmatrix} \hat{\alpha} \\ \hat{g} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix} \quad (2)$$

其中  $\lambda$  为惩罚因子, 公式为:  $\lambda = \sigma_e^2 / \sigma_a^2 = (1 - h^2) / h^2$ ,  $\sigma_a^2$  为抗病性状的加性遗传方差,  $h^2$  为抗病性状的狭义的遗传力。G 矩阵 (Genomic relationship matrix) 又称为基因组关系矩阵的计算公式由 VanRaden P M 提出<sup>[33]</sup>, 具体公式为:

$$G = \frac{(B - 2P)(B - 2P)'}{2 \sum p_j (1 - p_j)} \quad (3)$$

其中  $B$  代表的是等位基因的  $m \times n$  共享矩阵,  $m$  代表总的标记数量,  $n$  代表样本的总数;  $P$  代表的是每个 SNPs 位点次等位基因的频率 (Minor allele frequency, MAF) 组成的矩阵;  $p_j$  为每个 SNPs 在第  $j$  个位点等位基因 “a” 的频率。本研究的 G 矩阵通过 460 尾的 SNPs 位点的基因型数据, 使用 GCTA<sup>[32]</sup> 的计算得出, 相较于普通的 BLUP 模型, 将系谱推测的个体遗传关系的 A 矩阵改为由全基因组测序得到的 SNPs 标记构建的 G 矩阵, GEBV 估算的准确度更高<sup>[12]</sup>。

#### 1.5 交叉验证

交叉验证用于测试预测的准确度, 本研究采用 5 折的交叉验证, 将 460 尾个体随机分为 5 组 (参考群体与测试群体比例为 4:1), 使用 GBLUP 对 460 个体的 GEBV 进行计算, 测试群体的

表型设置为空, 通过参考群体的表型对测试群体的表型进行预测, 比较两个值之间的相关系数来衡量模型的预测能力。重复以上步骤 5 次, 使得每一组均有一次机会作为候选群体。交叉验证的重复次数为 100 次。

#### 1.6 不同标记密度对预测能力的影响

为考察不同 SNPs 数量 GEBV 的预测力, 除了使用全部的 SNPs (6 132 865 个 SNPs, 6 132 k) 外, 还设计了 13 个不同个数的标记子集, 分别为 0.5、1、3、5、10、30、50、100、250、500、1 000、2 500 k。为了降低抽样的误差, 使用 GATK SelectVariants 包对每个数量的标记集都进行 50 次的随机抽样后, 对每次抽样的结果都进行 100 次的 5 折交叉验证。

#### 1.7 不同覆盖度标准对预测的影响

根据本研究使用原始的测序深度, 对每个个体在每个 SNPs 上的覆盖度设置了 6 个过滤标准, 即 DP3、DP4、DP5、DP6、DP8 和 DP10, 分别对应于  $3 \times$ 、 $4 \times$ 、 $5 \times$ 、 $6 \times$ 、 $8 \times$  与  $10 \times$  的 reads 覆盖度。使用过滤后的 VCF 文件, 通过 BCFtools 对于每个 SNPs 位点大于 80% 的个体的基因型覆盖度大于过滤标准, 就保留该位点。过滤后使用 BEAGLE 对 VCF 文件进行填充, 后对每个过滤标准产生的 SNPs 标记集进行 100 次 5 折交叉验证。

## 2 结果

### 2.1 标记分型及群体结构分析

本研究对 460 尾赤点石斑鱼苗进行基因组重测序, 共获得 2.67 Tb clean data 数据挖掘 SNPs, 通过哈代-温伯格平衡 (HWE)  $> 10^{-6}$  以及次要等位基因频率 (MAF)  $> 5\%$  经过质控后, 最终用于分析的 SNPs 共有 5 412 683 个 (未进行覆盖度过滤), 平均标记密度约为 200.1 bp/SNPs。对 460 个样品进行主成分分析, 提取前 2 个主成分 (Principal components, PCs) 的结果 (图 1), 可以看出群体存在明显的分层现象, 分成了 2 个小的聚类群, 但易感 (Case) 和抗性 (Control) 个体在两个亚群中均匀分布。

### 2.2 遗传参数估计及基因组选择预测

利用全部 460 尾鱼苗的表型 (230 尾易感、

230尾抗性)和全部SNPs位点的基因型数据计算抗神经坏死病性状的遗传力,使用R包EM-MREML的结果为0.5662, GCTA的AI-REML经过1000迭代的结果为0.5666,两种方法估算结果相近。对460尾赤点石斑鱼进行抗神经坏死病基因组选择的预测力分析,80%个体作参考群,20%个体作验证群,分别进行了100次随机抽样的交叉验证,平均的基因组预测准确度为 $(0.359 \pm 0.019)$ 。

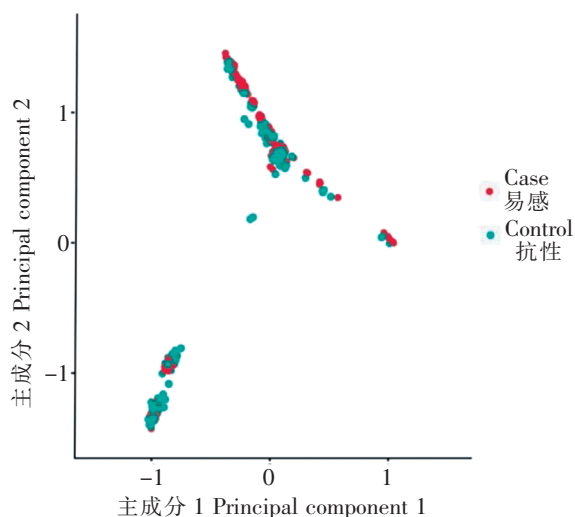


图1 460尾个体的群体遗传结构的主成分分析

Fig.1 Principal component analysis of population genetic structure of 460 individuals

注: Case组为易感组; Control组为抗性组。

Notes: Case group was the susceptible group; Control group was the resistant group.

### 2.3 不同SNPs数量的基因组选择预测力

从全部分型位点(5412683个)中随机取不同数量SNPs进行基因组选择分析,每组标记数量随机抽样50次。随机抽取的SNPs标记数量与育种值估计准确度的关系见图2。由图2可知,当标记密度由0.5k升到5k时基因组的预测能力迅速增加,标记密度达到5k以上时,育种值估计准确度随着标记的增加趋于平稳。

### 2.4 不同覆盖度标准对预测的影响

对测序数据进行不同覆盖度的过滤,获得的SNPs数目如表1所示。随着覆盖度过滤标准提高,保留下来的SNPs标记数量急剧减少。用不同覆盖度过滤得到的SNPs分别进行基因组选择预测能力评估,结果如图3所示,当覆盖度标准为DP3(即 $\geq 3\times$ )时,基因组选择预测力最高;覆盖度标准为DP4和DP5(保留的SNPs数量分别为5982和11291)时,基因组选择预测力几乎完全一样,均为0.346。覆盖度标准提高为DP6时,保留的SNPs只有3695个,预测力下降到0.30左右,DP8和DP10过滤剩余的SNPs数分别只有1817和830个,预测力下降为0.26左右。

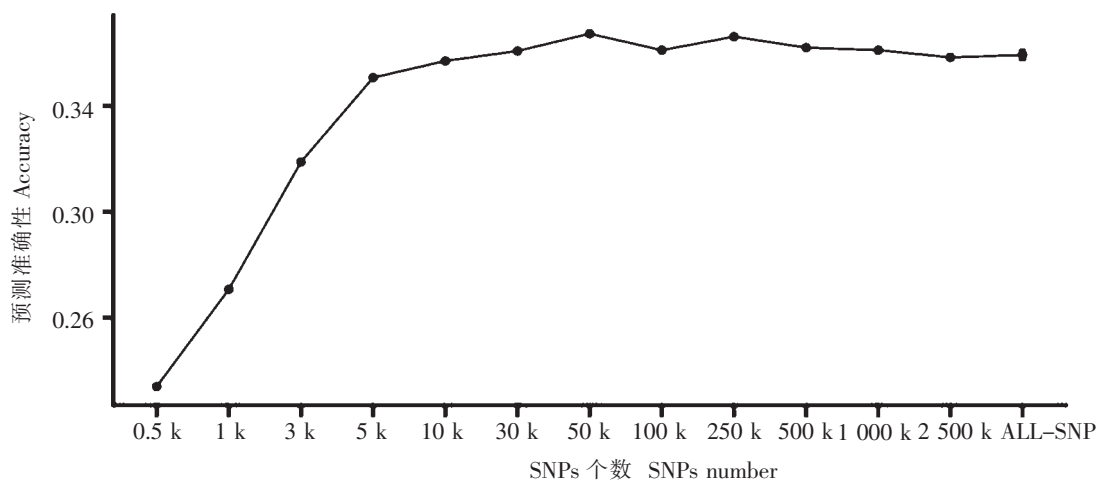


图2 不同标记密度下对赤点石斑鱼抗病性状的预测准确度

Fig.2 Accuracy of prediction model under different marker densities for disease resistance traits of *Epinephelus akaara*

表1 不同过滤标准剩余的 SNPs 数  
Tab. 1 Number of SNPs remaining for different filter standards

过滤标准 Filter standards	填充后 After filling	过滤 MAF 后 After MAF filter
ALL_SNPs	20 795 120	5 412 683
DP3	418 632	94 992
DP4	105 269	11 291
DP5	56 794	5 982
DP6	34 427	3 695
DP8	16 526	1 817
DP10	7 993	830

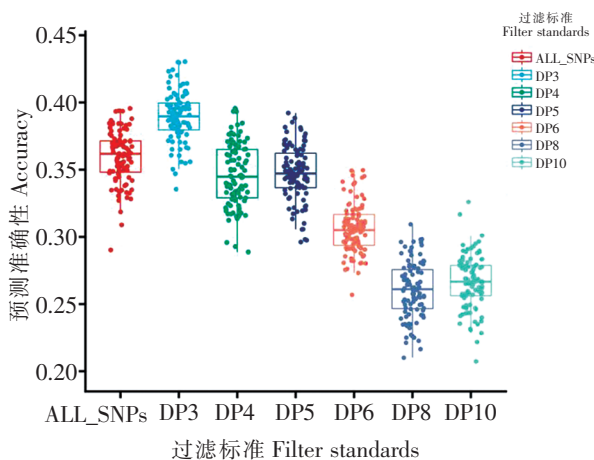


图3 不同覆盖度标准基因组估计的准确性

Fig.3 Accuracy of genome selection estimation for different coverage filter standards

### 3 讨论

基因组选择的首要工作建立参考群,利用参考群估算各个分子标记位点对性状表型的效应。本文建立了460尾石斑鱼的参考群,主成分分析显示群体聚集为2个不同的亚群,表明群体间存在较分化的亲缘关系。赤点石斑鱼人工育苗中亲本雌雄配比一般达到(20~30):1,使用的雄性亲本数量非常有限,因此同一批、尤其是同一池的受精卵可能来自少数雄性亲本。本研究发现鱼苗亲缘关系分化成2个亚群,推测可能采样的鱼苗来自于2个或几个雄性亲本及为数不多的母本。

候选群与参考群之间的亲缘关系是影响基因组选择效果的关键因素<sup>[34-35]</sup>。在应用基因组选择技术时,候选群亲鱼要尽可能地与参考群中个

体存在亲缘关系,如果是在同一家育苗场持续选育,并在鱼苗繁育中有可追踪的生产记录,则可以较可靠地推断候选亲鱼与参考群之间的亲缘关系,这将更有利于基因组选择技术的应用。另外,如果经费允许,有必要进一步扩大参考群规模,使参考群包含更多家系,增加参考群的多样性,对于扩展所建立的基因组选择技术的适用面、提高基因组选择的预测准确性和选育效率都具有重要意义<sup>[34]</sup>。

此外,基因组选择准确性还受基因组的标记密度的影响<sup>[36]</sup>。本文通过随机抽样研究了不同标记密度对育种值预测准确性的影响,当使用5k个标记时,即可使预测准确性接近利用全基因组SNPs的水平,其后随着标记个数增加,预测准确性变化趋于平缓;当标记个数为50k时,预测准确性与利用全基因组所有SNPs的预测准确性几乎一致,这与Tsai H Y等<sup>[19]</sup>对大西洋鲑抗海虱性状基因组预测的研究结果非常相似,提示如果设计赤点石斑鱼育种芯片,50k的标记密度可满足育种要求。尽管如此,利用全基因组SNPs标记有利于挖掘具有较大效应的分子标记及因果突变,将这些大效应分子标记和因果突变嵌入预测模型,建立基于主效-微效多基因效应相结合的预测模型,能进一步提高育种值预测准确性。利用同一批数据,笔者在基因组上已经发现了两个效应值极强的GWAS信号(结果未列出),计划在信号内部挖掘可靠的分子标记或因果突变,并将其作为协变量加入GBLUP模型中,通过主效标记/因果突变进一步吸收残余误差的方式进一步提高预测准确性。对测序数据进行覆盖度过滤能够提高对挖掘的标记位点基因分型的准确性,从而在一定程度上提高育种值估算的准确性(图3)。但是如表1和图3所示,在每个个体测量量不变的情况下,随着覆盖度过滤标准提高,可保留用于分析的标记数量急剧减少,当 $DP \geq 4$ 时,尽管标记分型的准确度会明显提升,但是由于保留的标记数量减少,育种值预测准确性已低于不进行覆盖度过滤。据Zhang W等<sup>[37]</sup>对大黄鱼研究的结果,采用全基因组重测序,由于可挖掘到的标记数多,即使测序覆盖度低至 $0.5 \times$ ,而且不进行覆盖度过滤,也能获得与 $8 \times$ 覆盖度测序基本一致的基因组选择效果,这将大大降低候选亲本标

记分型费用。因此,可以预期,随着高通量 DNA 测序价格进一步降低,基因组重测序将越来越多被用于基因组选择,也将成为石斑鱼全基因组选择育种的主要工具。

#### 参考文献:

- [ 1 ] Zorriehzahra M J. Viral Nervous Necrosis Disease [ G ] // Ennaji M. Emerging and Reemerging Viral Pathogens. Newyork: Academic Press, 2020: 673 - 703.
- [ 2 ] 正雄鵜川, 正毅樋口, 敏水戸. キジハタの産卵習性と初期生活史 [ J ]. 鱼类学杂志, 1966, 13 ( 4 - 6 ): 156 - 161.
- [ 3 ] 陈瑞雯, 王祖衍. 赤点石斑鱼 (*Epinephelus akaara*) 在我国首次人工育苗成功 [ J ]. 海洋通报, 1986 ( 4 ): 63.
- [ 4 ] 戴庆年, 张其永, 蔡友义, 等. 福建沿岸海域赤点石斑鱼年龄和生长的研究 [ J ]. 海洋与湖沼, 1988 ( 3 ): 215 - 224, 302.
- [ 5 ] 黄文, 陈偿, 赵哲, 等. 海水养殖鱼类病毒性神经坏死病防控技术研究进展 [ J ]. 渔业研究, 2016, 38 ( 5 ): 419 - 426.
- [ 6 ] 葛辉, 吴丽云, 周宸, 等. 斜带石斑鱼 (*Epinephelus coioides*) 病毒性神经坏死病毒的纯化分析及检测方法建立 [ J ]. 渔业研究, 2019, 41 ( 2 ): 96 - 105.
- [ 7 ] 朱志煌, 林克冰, 周宸, 等. 闽南地区斜带石斑鱼 (*Epinephelus coioides*) 神经坏死病毒的基因型与分子进化研究 [ J ]. 渔业研究, 2017, 39 ( 6 ): 429 - 436.
- [ 8 ] 陈信忠. 石斑鱼病毒性神经坏死病研究 [ D ]. 厦门: 厦门大学, 2006.
- [ 9 ] Fuji K, Hasegawa O, Honda K, et al. Marker - assisted breeding of a lymphocystis disease - resistant Japanese flounder (*Paralichthys olivaceus*) [ J ]. Aquaculture, 2007, 272 ( 1 ): 291 - 295.
- [ 10 ] 陈松林. 牙鲆“鲆优 2 号” [ J ]. 中国水产, 2018 ( 4 ): 65 - 72.
- [ 11 ] Meuwissen T H E, Hayes B J, Goddard M E. Prediction of total genetic value using genome - wide dense marker maps [ J ]. Genetics, 2001, 157 ( 4 ): 1819 - 1829.
- [ 12 ] 董林松. 大黄鱼和黄姑鱼基因组选择研究 [ D ]. 厦门: 集美大学, 2018.
- [ 13 ] Dong L, Xiao S, Chen J, et al. Genomic selection using extreme phenotypes and pre - selection of SNPs in large yellow croaker (*Larimichthys crocea*) [ J ]. Marine Biotechnology, 2016, 18 ( 5 ): 575 - 583.
- [ 14 ] Wang Q, Yu Y, Li F, et al. Predictive ability of genomic selection models for breeding value estimation on growth traits of Pacific white shrimp *Litopenaeus vannamei* [ J ]. Chinese Journal of Oceanology and Limnology, 2017, 35 ( 5 ): 1221 - 1229.
- [ 15 ] Palaiokostas C, Kocour M, Prchal M, et al. Accuracy of genomic evaluations of juvenile growth rate in common carp (*Cyprinus carpio*) using genotyping by sequencing [ J ]. Frontiers in Genetics, 2018, 9: 82.
- [ 16 ] Wang Y, Sun G, Zeng Q, et al. Predicting growth traits with genomic selection methods in zhikong scallop (*Chlamys farreri*) [ J ]. Marine Biotechnology, 2018, 20 ( 6 ): 769 - 779.
- [ 17 ] Liu G, Dong L, Gu L, et al. Evaluation of genomic selection for seven economic traits in yellow drum (*Nibea albiflora*) [ J ]. Marine Biotechnology, 2019, 21 ( 6 ): 806 - 812.
- [ 18 ] Shan X, Xu T, Ma Z, et al. Genome - wide association improves genomic selection for ammonia tolerance in the orange - spotted grouper (*Epinephelus coioides*) [ J ]. Aquaculture, 2021, 533: 736214.
- [ 19 ] Tsai H Y, Hamilton A, Tinch A E, et al. Genomic prediction of host resistance to sea lice in farmed *Atlantic salmon* populations [ J ]. Genetics Selection Evolution, 2016, 48 ( 1 ): 47.
- [ 20 ] Liu Y, Lu S, Liu F, et al. Genomic selection using BayesC $\pi$  and GBLUP for resistance against *Edwardsiella tarda* in Japanese flounder (*Paralichthys olivaceus*) [ J ]. Marine Biotechnology, 2018, 20 ( 5 ): 559 - 565.
- [ 21 ] Palaiokostas C, Cariou S, Bestin A, et al. Genome - wide association and genomic prediction of resistance to viral nervous necrosis in European sea bass (*Dicentrarchus labrax*) using RAD sequencing [ J ]. Genetics Selection Evolution, 2018, 50 ( 1 ): 30.
- [ 22 ] Ge H, Lin K, Shen M, et al. De novo assembly of a chromosome - level reference genome of red - spotted grouper (*Epinephelus akaara*) using nanopore sequencing and Hi - C [ J ]. Molecular Ecology Resources, 2019, 19 ( 6 ): 1461 - 1469.
- [ 23 ] Yang M, Wang Q, Chen J, et al. Identification of candidate SNPs and genes associated with anti - RGNNV using GWAS in the red - spotted grouper, *Epinephelus*



- akaara[J]. *Fish & Shellfish Immunology*, 2021, 112: 31–37.
- [24] Chen S, Zhou Y, Chen Y, et al. Fastp: an ultra-fast all-in-one FASTQ preprocessor[J]. *Bioinformatics*, 2018, 34(17): i884–i890.
- [25] Ewels P, Magnusson M, Lundin S, et al. MultiQC: summarize analysis results for multiple tools and samples in a single report[J]. *Bioinformatics*, 2016, 32(19): 3047–3048.
- [26] Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA – MEM[J]. arXiv:1303.3997[q-bio], 2013: 1–3.
- [27] Tarasov A, Vilella A J, Cuppen E, et al. Sambamba: fast processing of NGS alignment formats[J]. *Bioinformatics*, 2015, 31(12): 2032–2034.
- [28] 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.
- [29] Danecek P, Bonfield J K, Liddle J, et al. Twelve years of SAMtools and BCFtools[J]. *GigaScience*, 2021, 10(2): giab008.
- [30] Browning S R, Browning B L. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering[J]. *American Journal of Human Genetics*, 2007, 81(5): 1084–1097.
- [31] Purcell S, Neale B, Todd-Brown K, et al. PLINK: A tool set for whole-genome association and population-based linkage analyses[J]. *American Journal of Human Genetics*, 2007, 81(3): 559–575.
- [32] Yang J, Lee S H, Goddard M E, et al. GCTA: A tool for genome-wide complex trait analysis[J]. *American Journal of Human Genetics*, 2011, 88(1): 76–82.
- [33] VanRaden P M. Efficient methods to compute genomic predictions[J]. *Journal of Dairy Science*, 2008, 91(11): 4414–4423.
- [34] Pszczola M, Strabel T, Mulder H A, et al. Reliability of direct genomic values for animals with different relationships within and to the reference population[J]. *Journal of Dairy Science*, 2012, 95(1): 389–400.
- [35] Clark S A, Hickey J M, Daetwyler H D, et al. The importance of information on relatives for the prediction of genomic breeding values and the implications for the makeup of reference data sets in livestock breeding schemes[J]. *Genetics Selection Evolution*, 2012, 44(1): 4.
- [36] Wang Q, Yu Y, Yuan J, et al. Effects of marker density and population structure on the genomic prediction accuracy for growth trait in Pacific white shrimp *Litopenaeus vannamei*[J]. *BMC Genetics*, 2017, 18(1): 45.
- [37] Zhang W, Li W, Liu G, et al. Evaluation for the effect of low-coverage sequencing on genomic selection in large yellow croaker [J]. *Aquaculture*, 2021, 534: 736323.



## Evaluation of genomic selection against neuronecrosis in the red – spotted grouper

WANG Zhuobiao<sup>1</sup>, FANG Ming<sup>1</sup>, ZHENG Leyun<sup>2</sup>, GE Hui<sup>2</sup>,  
CHEN Xinxin<sup>2</sup>, LUO Huiyu<sup>2</sup>, WANG Zhiyong<sup>1\*</sup>

- (1. Key Laboratory of Healthy Mariculture for the East China Sea, Ministry of Agriculture and Rural Affairs, Fisheries College, Jimei University, Xiamen 361021, China;
2. Key Laboratory of Cultivation and High – value Utilization of Marine Organisms in Fujian Province, Fujian Fisheries Research Institute, Xiamen 361013, China)

**Abstract:** Red – spotted grouper (*Epinephelus akaara*) is an important maricultured fish in the southeast coastal area of China. It is highly infected by red – spotted grouper nerve necrosis virus (RGNNV) during the artificial cultivation, which seriously restricted the development of the artificial breeding industry. In this study, 230 ind RGNNV susceptible (disease – infected death) and 230 ind resistant (eventually survived fish) red – spotted grouper fry were collected for genome resequencing, and 5 412 683 single nucleotide polymorphisms (SNPs) were discovered, which were used for genetic evaluation of the RGNNV resistance of the red – spotted grouper. The estimated heritability of RGNNV was 0.566 2, and the accuracy of the genomic estimated breeding value (GEBV) with whole genome SNPs markers was 0.154 3. The evaluation of the effect of random selection of different numbers of markers on the genomic prediction accuracy showed that using  $\geq 5\ 000$  (5 k) SNPs markers could obtain more desirable results. This study provided a useful theoretical reference for application of genomic selection for breeding of red – spotted grouper against RGNNV.

**Key words:** red – spotted grouper; genomic selection; genomic best liner unbiased prediction (GBLUP)