

大黄鱼 *Cspg4* 基因的表达调控

罗阿蓉, 李完波, 王志勇

(集美大学水产学院、农业农村部东海海水健康养殖重点实验室, 福建厦门 361021)

[摘要] 从大黄鱼脾脏 eQTL 数据库中, 筛选出调控免疫相关靶基因 *Cspg4* 的顺式作用区 eQTL, 通过 Sanger 测序验证了该 eQTL 区域中最显著的 SNP 位点的基因型。在此基础上, 通过构建候选启动子区不同长度缺失片段的重组质粒测定活性, 确定了 *Cspg4* 基因启动子核心区为 -1202 ~ -762 和 -190 ~ +20, 而携带 Chr17_13952653 突变位点的两种单倍型调控序列的活性均显著低于对照组, 提示该调控序列可能为沉默子。

[关键词] 大黄鱼; *Cspg4*; 表达调控; 沉默子

[中图分类号] S 963

Expression Regulation of *Cspg4* in Large Yellow Croaker

LUO Arong, LI Wanbo, WANG Zhiyong

(Fisheries College & Key Laboratory of Healthy Mariculture for East China Sea of Ministry Agriculture and Rural Affairs, Jimei University, Xiamen 361021, China)

Abstract: In this study, the *Cspg4* gene was cloned based on previous eQTL analysis in *Larimichthys crocea*, which was regulated by cis-eQTLs. The results of Sanger sequencing verified the genotypes of the most significant SNP Chr17_13952653 regulating *Cspg4* gene. Subsequently, the core promoter regions of *Cspg4* were determined to be from -1202 to -762 and from -190 to +20 through recombinant plasmids with different truncated fragments in the candidate promoter region. Further study found that the activity of the two haplotypes containing Chr17_13952653 site was significantly lower than that of the control group, suggesting that the sequence may be a silencer, and weaken the promoter activity in transcription. Our results provided useful information for elucidating the molecular mechanisms of cis-regulation of *Cspg4* gene in *L. crocea*.

Keywords: *Larimichthys crocea*; *Cspg4*; expression regulation; silencer

0 引言

大黄鱼是我国养殖量最大的海水鱼类, 但各种病害频发, 而且病原复杂多样、防控困难, 使养殖业遭受巨大损失。从大黄鱼自身的免疫机制入手, 阐明其免疫抗病机制, 可为制定防治措施提供理论基础^[1-3]。本课题组前期对 81 尾 1.5 龄健康大黄鱼脾脏进行了转录组测序, 结合全基因组测序数据进行表达数量性状位点分析 (expression quantitative trait locus, eQTL), 发掘出诸多调控免疫相关靶基因的顺式作用 eQTL (未发表数据), 这些靶基因有的在免疫系统中充当关键枢纽, 如参与免疫调节

[收稿日期] 2022-06-01

[基金项目] 国家自然科学基金项目 (31872562); 海水鱼类产业技术体系项目 (CARS-47-G04)

[作者简介] 罗阿蓉 (1996—), 女, 硕士生, 从事水产遗传育种研究。通信作者: 李完波 (1982—), 男, 副教授, 硕导, 从事水产遗传与育种研究。E-mail: wanbo@jmu.edu.cn

<http://xuebaobangong.jmu.edu.cn/zkb>

的硫酸软骨素蛋白聚糖 4 (chondroitin sulfate proteoglycan 4, *Cspg4*) 基因。调控 *Cspg4* 基因的 eQTL 中最显著的单核苷酸多态性 (single nucleotide polymorphism, SNP) 位点位于脂质磷酸磷酸酶相关蛋白 1 型 (lipid phosphate phosphatase-related protein type 1, *lppr1*) 的内含子中。

Cspg4 是一种由硫酸软骨素和核心蛋白组成的具有多个亚结构的跨膜蛋白, 在细胞生长、连接、迁移和细胞外基质中起作用。近几年的研究发现, *Cspg4* 基因在人类多种恶性肿瘤细胞及组织中高表达, 如恶性黑色素瘤^[4]、成胶质细胞瘤^[5]、乳腺癌^[6]、头颈肿瘤^[7]、胰腺肿瘤^[8]和难治性混合性白血病^[9]等。它能促进肿瘤细胞的多种生物学行为, 如增殖、运动、侵袭和抗凋亡, 这些行为在肿瘤的发生发展过程中具有重要作用^[10-11]。人类和犬类的 *Cspg4* 氨基酸序列具有 82% 的同源性, Riccardo 等^[12]研究发现人的 *Cspg4* DNA 能使犬口腔恶性黑色素瘤中 *Cspg4* 特异性免疫和生存期延长, 针对 *Cspg4* 的异种疫苗能够克服宿主对“自我”抗原的无反应性, 对治疗犬类恶性黑色素瘤初步有效。另外研究者们利用模式生物也对 *Cspg4* 基因进行了多方面的研究。Lee 等^[13]发现 *Cspg4* 在斑马鱼胚胎发育中通过调控 Wnt/平面细胞极性信号通路调控体轴组织, 而 *Cspg4* 的过表达还会导致睫状视网膜动脉阻塞。受 *Cspg4* 基因座的转录调节, 在没有心功能不全或纤维脂肪症、掌跖角化病和脱发的情况下, 桥粒斑蛋白 (desmoplakin, Dsp) 的缺失可导致小鼠致死性心律失常, 类似人类的心脏皮肤综合症^[14]。鼠体内还存在 *Cspg4* 同源蛋白的神经元胶原抗原 (neuron-glia antigen2, Ng2)。Zhu 等^[15]利用细菌人工染色体修饰技术, 制备了在 Ng2 阳性 (Ng²⁺) 细胞中特异性表达红色荧光蛋白 DsRed 或噬菌体 Cre 重组酶的转基因小鼠, 结果发现 Ng²⁺ 细胞在体内外均可形成原生质星形胶质细胞和少突胶质细胞。Wilms 等^[16]证明脊椎动物特异性氧化还原酶谷氧还蛋白 2 (Grx2c) 的胞浆异构体调节转录因子 Sp-1 的氧化还原状态, 使其与 *Cspg4* 基因的启动子和增强子区域结合, 导致编码蛋白水平的增加, 从而增强了胶质细胞和胶质瘤细胞的迁移能力。

本研究就上述 *Cspg4* 基因的表达调控机制进行初步探索。通过 PCR 产物的 Sanger 测序验证该 eQTL 区域中最显著的 SNP 位点的真实性, 随后构建 *Cspg4* 基因候选启动子区不同长度缺失片段的重组质粒, 利用双荧光素酶报告系统检测重组质粒的活性并确定 *Cspg4* 基因的核心转录活性区域, 在此基础上插入此基因的不同单倍型调控序列进行验证, 为揭示该 eQTL 与靶基因之间的表达调控方式打下基础。

1 材料与方法

1.1 实验材料

1.1.1 实验动物

本实验室前期对 81 尾 1.5 龄的健康大黄鱼进行了全基因组重测序 (从鳍条提取 DNA) 和转录组测序 (从脾脏提取 RNA), 所用大黄鱼来自宁德市三都澳养殖渔排。DNA 和 RNA 的提取按照常规方法进行, 动物实验操作遵循集美大学水产学院动物伦理委员会的要求进行。

1.1.2 菌株和细胞

大肠杆菌菌株 DH5 α Competent *E. coli* Strain 购自南京诺唯赞生物科技有限公司; 萤火虫 pGL3-basic 载体菌株、内参海参 pRL-TK 载体菌株为本实验室保存, HEK 293T 细胞购自上海富衡生物科技有限公司, 经传代、冻存保存于本实验室。

1.1.3 引物

本研究所用引物由厦门铂瑞生物科技有限公司和生工生物工程 (上海) 有限公司合成, 其名称及序列如表 1 所示。

1.2 实验方法

1.2.1 SNP 位点验证

前期研究获得了调控 *Cspg4* 基因的顺式 eQTL, 其中最显著的 eQTL 位于大黄鱼基因组 17 号染色

体 13952653 bp 的一个 SNP（Chr17_ 13952653）上。为了检验该 SNP 是否真实存在而非测序错误，搜寻用于 eQTL 分析的 81 个个体的全基因组重测序数据，获取该 SNP 位点的基因型信息，并进行汇总统计。同时，从参考基因组上获取该 SNP 两侧各 250 bp 的序列，使用 NCBI 网站“Primer-BLAST”在线工具分别设计 PCR 引物（chr17_ 13952653-F/R，见表 1）。然后，根据基因组重测序获得该 SNP 位点的基因型，每个位点、每种基因型分别挑选若干个体的鳍条 DNA 为模板，使用高保真 Taq 酶进行 PCR 扩增，扩增产物送至厦门铂瑞生物科技有限公司进行 Sanger 测序，并将测序结果与基因组重测序结果进行比较。PCR 反应程序如下：95 ℃ 3 min；95 ℃ 15 s，55 ℃ 20 s，72 ℃ 30 s，35 个循环；72 ℃ 5 min；4 ℃ 保存。取 2 μL PCR 反应后产物，用 2% 琼脂糖凝胶电泳检测扩增产物是否条带单一。

表 1 本研究所用引物序列及名称
Tab.1 Sequence and name of primers used in this experiment

引物名称 Name	引物序列 Sequence(5′-3′)	用途 Usage
chr17_13952653-F	GTTGAGTGAGGTTTAATTGGAAGTA	SNP 位点验证 Verification of SNP
chr17_13952653-R	ATTTAGGTTTCAGACTGTTGTTTGG	
pGL3-Cspg4-P-R	TCTTACGCGTGCTAGCCCGGGCTCTCGCAGAGCTTCTTCATCC	<i>Cspg4</i> 启动子活性 Activity of <i>Cspg4</i> promote
pGL3-Cspg4-P1-F	TCTTACGCGTGCTAGCCCGGGCCCCACATTAACAAGCCAGTCA	
pGL3-Cspg4-P2-F	TCTTACGCGTGCTAGCCCGGGCAATCTAACATCACAAGTTGAACTATGTT	
pGL3-Cspg4-P3-F	TCTTACGCGTGCTAGCCCGGGGCTGGACGCAGAGACTTCTTG	
pGL3-Cspg4-P4-F	TCTTACGCGTGCTAGCCCGGGATAAATATTGGCATTTACTATCCCAAGG	
pGL3-Cspg4-P5-F	TCTTACGCGTGCTAGCCCGGGGATCTGGGCTCTGGTGAGCA	
pGL3-Cspg4-P6-F	TCTTACGCGTGCTAGCCCGGGTCAACATCTCCACCGGGTTTA	
pGL3-Cspg4-P7-F	TCTTACGCGTGCTAGCCCGGGGAGAGGGCAATAAAATTGAGTTGCC	
pGL3-chr17-1kb-F	AGGTACCGAGCTCTTACGCGTTCTCCTATTCTGGCGCTCCATTTGC	调控序列扩增 Amplification of regulatory sequences
pGL3-chr17-1kb-R	TTTTATTGCCCTCTCCCCGGGGTAATTTGCCTGACATGAAATGGG	
RV primer3	CTAGCAAAATAGGCTGTCCC	pGL3-basic 载体通用引物
GL primer2	CTTTATGTTTTTGGCGTCTTCCA	Universal primers for pGL3-basic vector

注：划线部分为酶切位点，划线前部分为质粒酶切位点前的同源序列；F—上游引物；R—下游引物。
Notes: The underlines and before areas represent enzyme digestion sites, homologous sequence before enzyme digestion sites of the plasmid, respectively; F—Forward primer; R—Reverse primer.

1. 2. 2 *Cspg4* 启动子活性分析

从大黄鱼参考基因组中获得 *Cspg4* 基因启动子区序列信息，以大黄鱼鳍条 DNA 为模板，扩增 *Cspg4* 候选启动子序列。为了鉴定大黄鱼 *Cspg4* 启动子序列中的核心区域，对其进行不同长度片段的截短。使用 2 × Phanta[®] Max Master Mix，带有 *Sma* I 和 *Hind* III 两个酶切位点的相应引物（pGL3-Cspg4-P-R，pGL3-Cspg4-P1 ~ P7-F，见表 1）分别进行扩增，以获得全长启动子和各个截短的片段。PCR 反应程序如下：95 ℃ 3 min；95 ℃ 15 s，58 ℃ 20 s，72 ℃ 4 min，35 个循环；72 ℃ 5min；10 ℃ 保存。根据各试剂盒说明书进行 PCR 产物纯化和 pGL3-Cspg4-P1 ~ P7 重组质粒的构建及鉴定。构建成功的重组质粒转染至 HEK 293T 细胞并进行双荧光素酶活性测定。

为预测大黄鱼 *Cspg4* 基因的 5′UTR 上游候选启动子序列中潜在核心区域的分布，本研究使用在线

程序神经网络启动子 Neural Network Promoter Prediction 预测 BDGP, 阈值设为 0.8。

1.2.3 Chr17_13952653 位点不同单倍型序列对 *Cspg4* 基因转录的调控作用验证

通过上述 *Cspg4* 启动子活性分析实验确定启动子的核心区域后, 在核心区前插入含 Chr17_13952653 位点的序列, 以此来研究包含该 SNP 位点的序列对靶基因 *Cspg4* 启动子的调控情况。

根据大黄鱼参考基因组设计引物 (pGL3-chr17-1kb-F/R, 见表 1) 扩增包含 Chr17_13952653 位点的序列, 其片段大小为 1157 bp, 2 条引物分别带有 *Mlu* I 与 *Sma* I 酶切位点。从 81 个重测序样品中选取 Chr17_13952653 位点基因型为 AA 和 CC 的个体, 以其鳍条 DNA 为模板, 使用 2 × Phanta[®] Max Master Mix 进行 PCR 扩增, PCR 反应程序如下: 95 ℃ 3 min; 95 ℃ 15 s, 55 ℃ 20 s, 72 ℃ 70 s, 35 个循环; 72 ℃ 5min; 4℃ 保存。根据各试剂盒说明书进行 PCR 产物纯化和 pGL3-*Cspg4*-A/C 重组质粒的构建及鉴定。构建成功的重组质粒转染至 HEK 293T 细胞并进行双荧光素酶活性测定。

1.2.4 *Cspg4* 基因不同单倍型调控序列生物信息学分析

为了计算包含 Chr17_13952653 位点的调控序列内的连锁不平衡, 本研究提取区域内所有的 SNP 及 InDel 并进行可视化展示, 具体步骤如下:

1) 用 GATK V3.8^[17] 的 HaplotypeCaller 模块, 提取含 Chr17_13952653 位点的调控序列 1157 bp 内所有的 SNP 及 InDel, 用 CombineVariants 模块合并 VCF 文件。

2) 通过 PopLDdecay V3.41^[18] 计算该区域内的连锁不平衡, 即每个 SNP 或 InDel 之间的 r^2 值, 统计与 Chr17_13952653 该 SNP 强连锁的 SNP 或 InDel。

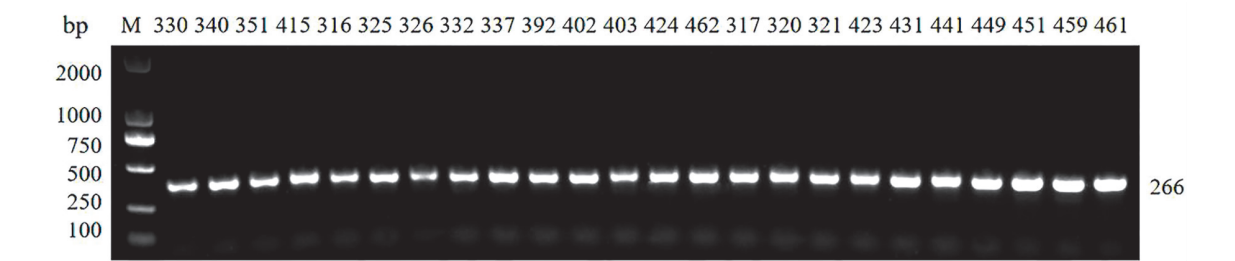
3) 通过 R 4.0.3 中的 TrackViewer 包, 对强连锁的 SNP 或 InDel 在基因组上的位置, 进行可视化展示。

2 结果与分析

2.1 最显著 eQTL 突变位点的验证

2.1.1 最显著 SNP 位点所在短片段序列扩增与测序结果

含 Chr17_13952653 SNP 位点的短片段扩增产物电泳检测结果如图 1 所示, 在 266 bp 左右有单一条带, 符合预期结果。将上述扩增产物进行 Sanger 测序, 测序结果如图 2 所示, 发现各测序样品该 SNP 位点的基因型与基因组重测序获得的结果完全一致, 纯合子测序峰图为单峰, 杂合子测序峰图为双峰。以上结果表明该显著 SNP 位点均真实存在于实验群体中。



注: M 为 2000 bp marker; 330–415 为 AA 基因型的个体; 316–462 为 AC 基因型的个体; 317–461 为 CC 基因型的个体。
Notes: M is a 2000 bp marker; 330–415 are individuals with AA allele; 316–462 are individuals with AC allele; 317–461 are individuals with CC allele.

图 1 大黄鱼含 Chr17_13952653 显著 SNP 位点的短片段扩增产物电泳图
Fig.1 The electrophoresis diagram of the amplified product which containing this significant SNPs Chr17_13952653

2.1.2 最显著 eQTL 位点的基因型统计

从 81 个全基因组重测序数据中, 获取显著影响 *Cspg4* 基因表达的 Chr17_13952653 位点的基因型, 结果显示 Chr17_13952653 位点为 A/C 突变, A 的频率为 0.34, C 的频率为 0.66。其中 AA 基因型个体 4 个, AC 基因型个体 47 个, CC 基因型个体 30 个。

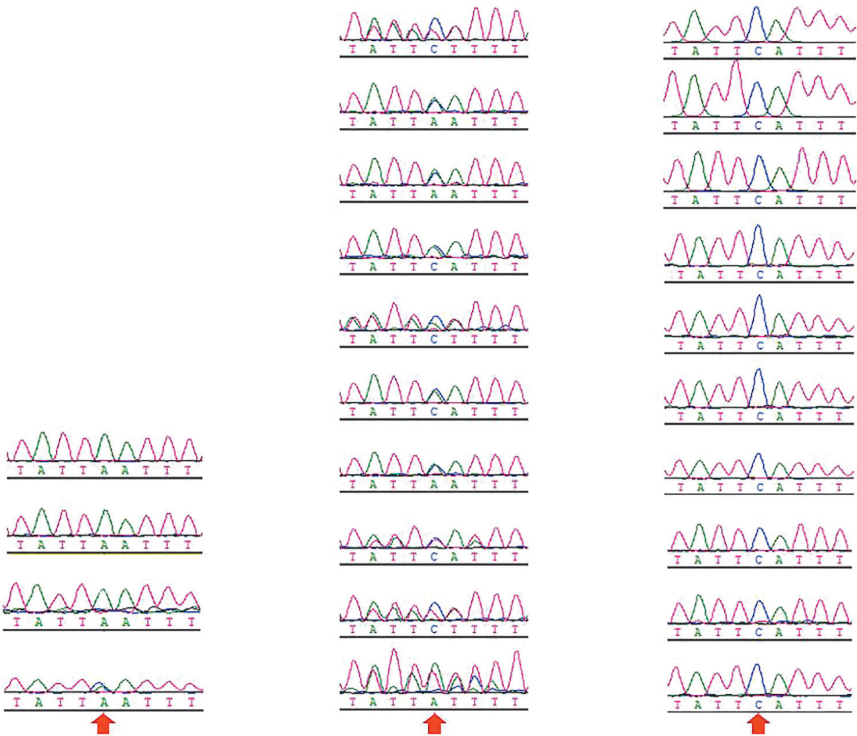


图 2 大黄鱼 Chr17_13952653 显著 SNP 位点的基因型测序峰图

Fig.2 Sequencing peak map of allelic genotypes of Chr17_13952653 in large yellow croaker

2.2 Cspg4 启动子活性分析

2.2.1 Cspg4 候选启动子序列分析及潜在核心区域预测

本研究成功扩增到 *Cspg4* 基因 2076 bp 的启动子序列，另包含 20 bp 的外显子 1 的部分编码序列，具体序列信息见图 3。使用在线程序 BDGP 启动子预测大黄鱼 *Cspg4* 候选启动子序列中潜在的核心区域，结果表明有 1 个潜在的核心区域存在于该基因的候选启动子序列中，位置及得分见表 2，序列具体位置如图 3 所示。

表 2 大黄鱼 *Cspg4* 候选启动子序列中预测的核心区域的分布

Tab.2 The distribution of predicted core regions within the candidate promoter sequence of *Cspg4* in large yellow croaker

开始 Start	结束 End	得分 Score (0-1)	启动序列 Promoter sequence(5'3')
-586 bp	-536 bp	0.93	ggcagctggactaaataacccacaccttctcaacatctccaccgggttta

2.2.2 Cspg4 候选启动子区活性分析

双荧光素酶报告系统分析结果显示，*Cspg4* 基因候选启动子区不同长度缺失片段的相对荧光活性由强到弱的顺序为：P4、P3、P7、P2、P1、P6、P5（见图 4），其中 pGL3-*Cspg4*-P4 重组质粒的启动子活性最高。pGL3-*Cspg4*-P1 ~ P3 重组质粒相较于 pGL3-*Cspg4*-P4 的启动子活性逐渐降低，表明在 -2076 ~ -1202 区域可能存在负调控元件（negative regulatory element，NRE）；在 -1202 ~ -762 和 -190 ~ +20 区域可能为该启动子的核心区域，存在基本调控元件；而 pGL3-*Cspg4*-P5 ~ P7 重组质粒的启动子活性逐渐增强，表明在 -762 ~ -190 区域也可能存在负调控元件。

ccccacattaacaagccagtc aatacagattaatggagtagcctttgctgggagaaggc -1959
ctttggtcaccaaacgaaaaagctgcggtgaaacctgtttcaaattaagctaatatg -1900
attctgactatgctgcaagtcaaaacagtacaatatgaataccttacaataatgtaaac -1841
aacaacagtttttttaacacgt caatctaacatcacaagttgaacta tgttatatttct -1782
tgcaggggtcttttgaatgcattagatttcacaggtgtttctgatacagtgggcaccg -1723
ctgtatgtaccctgcgagcactgaaactgcttaattgtgtggaaaccagaaatcaacc -1664
ttctacaattacattgtgaatcagtaaaaactttccatacacctttatggacttctcat -1605
ctgtctcctttttctttgtcgtagtgtctttctacggcgatggctacatccacttgccg -1546
acagtggaggcgctcaatcagacattgctccatgtacgatttgaacctccagtcaggt -1487
ggggctgctattttggca gctggacgcagagacttcttgttgttggagctgatctctg -1428
ggcacctgcaggtaggctaacgcacatattaagataacagcacatgtgtaaaaagcaac -1369
tgaaagggaatatactcgcagaatgcaaacacacaagacaaatgtatcattttaaagata -1315
agacaaatggtatgacatgaagtttgttttgtgcttttggctgagggatggaatgct -1310
tcaattgtagtaatacgaatacaaatgagtc aaagttaaacttgtccagaggctctcat -1251
aaaacaaatatttgagaaaaaatttgttccttggggccattttagctc ataaatattggc -1192
atttactatccc aaggagagccagagagaagactgtacacactcagcttgtcacagag -1133
agaaaaaaaaattacccctggcactggcttcatttcagaacaagttcttcatgacc -1074
aagttaagagagagagcgcaatagttaatgaaaacttcaagtacatgtcaagggaagc -1015
tgcttatgaaagtgtctgttgctttatgtttgctgaattatccttggatctacagtt -956
tacactcaactcattgattctttaattgtgcctactcaagatctcatcaccatgttat -897
gtctgtttttatgtgttttggcatagaagcaaatattagaaacactaatttcatgtgtc -838
ttggttatcacctttaaccctttttctcattctttaacctctc atctctccgatat -779
ccctcaggtgctgtg gatctgggctctggtgagcattcac tacgctcagaaaaaggca -720
tccatcttagtgacctggcatggcactcaatggagctgacctatgactaccataatgtc -661
accatgaccgtaggaccgaaactctcataccagactccgfatgccaggaccagatctgg -602
agctcagtggttgaggatgggctttttgttggcag ggcagctggactaaataaccacac -543
cttctcaacatctccaccgggttagagg ttgtgtagatgaagttgtattcaatgagca -484
taacttgtgtgccagcctaaggccttattctggatacaagagcatccatgaggtatctc -425
tgggctgtagtcaacagttttctgcaacagaagaagattctgtcagtttttcagctct -366
aaagcctttatttctctcccaccgtgggaggtgccacaggagggggtgtttgaatgtga -307
attacatccctctgcaaaagaagaagatggcatgacctgtacagctctggcaaccaag -248
gaggggtcgttctgttagagatcagagagggtcgtctggtggcaacagtaggaaatgga -189
gagggcaataaaattgagttgcg ttctttaacatatgtccacagtaacctcatcctggt -130
acccatacagcttcacctgctgccccagagtgtccagttgaagggtgggcgaggagttg -71
gtcaaggccaacctaaagtctggagctccaggtcatccagctcaaagggcctctcttct -12
gggagggctgg **ATGAAGAAGCTCTGCGAGAG** +20

注:黄色区域为 *Cspg4* 启动子各个缺失片段重组质粒的上游引物同源序列;下划线区域为预测的潜在核心区域;以 ATG 第一个碱基为正向位置+1。

Notes: The yellow area is the homologous sequence of forward primers of recombinant plasmids with each missing fragment of *Cspg4* promoter; the underlined area is the potential core area for prediction;the first base at ATG is +1.

图3 大黄鱼 *Cspg4* 启动子的核苷酸序列

Fig.3 The nucleotide sequence of large yellow croaker *Cspg4* promoter

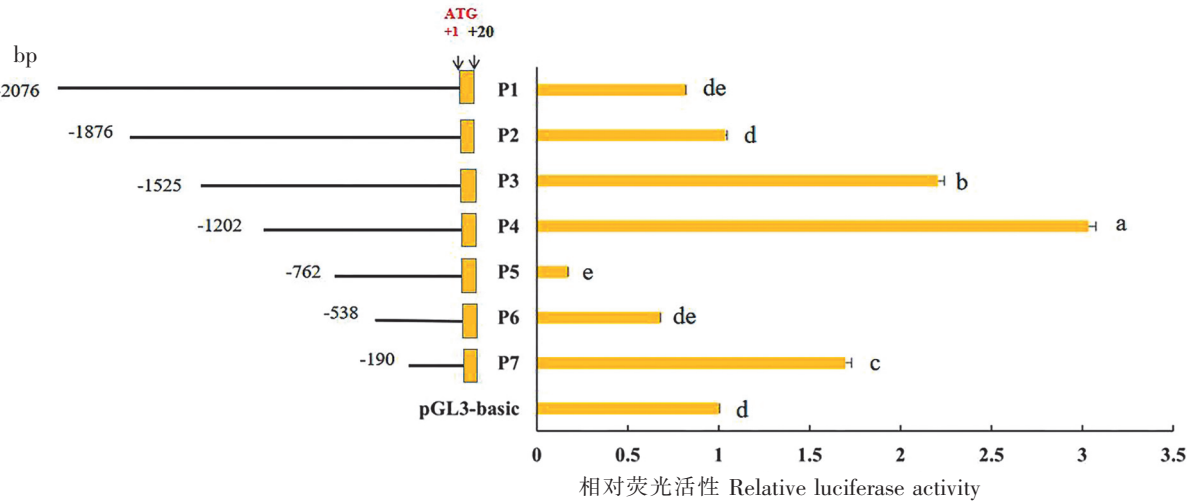
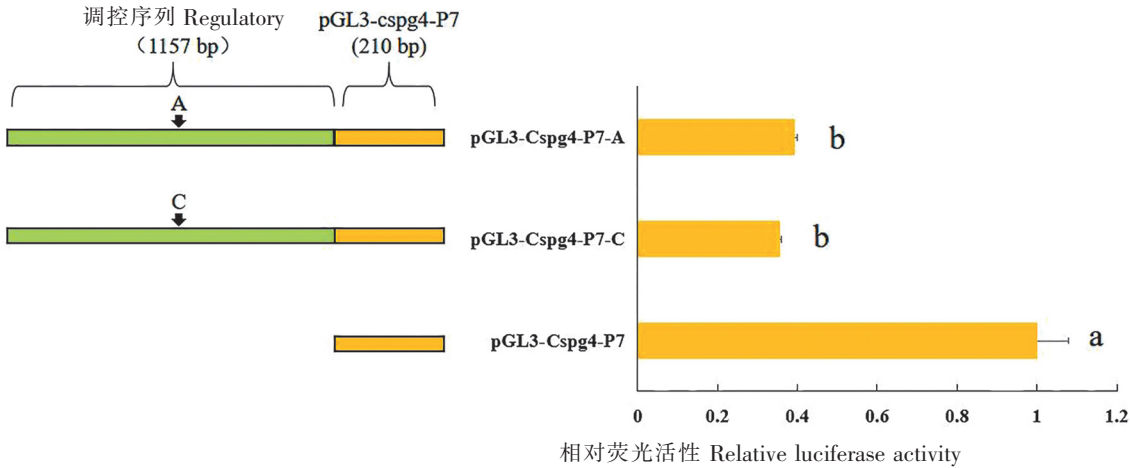


图 4 大黄鱼 *Cspg4* 基因候选启动子不同长度缺失片段重组质粒活性分析
Fig.4 Activity of different length missing fragments in large yellow croaker *Cspg4* promoter

2.3 Chr17_ 13952653 位点不同单倍型序列对 *Cspg4* 基因转录的调控作用

据上述大黄鱼 *Cspg4* 候选启动子区活性分析，推测可知在 -1202 ~ -762 和 -190 ~ +20 区域可能为该启动子的核心区域。本研究选择 pGL3-*Cspg4*-P7（-190 ~ +20）作为 *Cspg4* 基因的核心区域，在此核心区域前插入含 Chr17_ 13952653 位点共 1157 bp 的调控序列，以此来研究该序列对靶基因 *Cspg4* 启动子的表达调控情况。

双荧光素酶报告系统分析结果显示，pGL3-*Cspg4*-P7-A 和 pGL3-*Cspg4*-P7-C 重组质粒的相对荧光活性均显著低于对照组 pGL3-*Cspg4*-P7，而 pGL3-*Cspg4*-P7-A 和 pGL3-*Cspg4*-P7-C 重组质粒之间相对荧光活性无显著差别，pGL3-*Cspg4*-P7-C 重组质粒活性略低（如图 5， $P < 0.05$ ），表明含 Chr17_ 13952653 位点的 1157 bp 序列中可能含有减弱启动子活性的调控序列，即含有沉默子。



注：黄色方块代表 *Cspg4* 基因启动子核心区域，绿色方块代表插入的调控序列，A 和 C 分别代表不同单倍型个体。
Notes: The yellow squares represent the *Cspg4* promoter core region, the green squares represent inserted regulatory sequences, individuals with different alleles are represented by A and C, respectively.

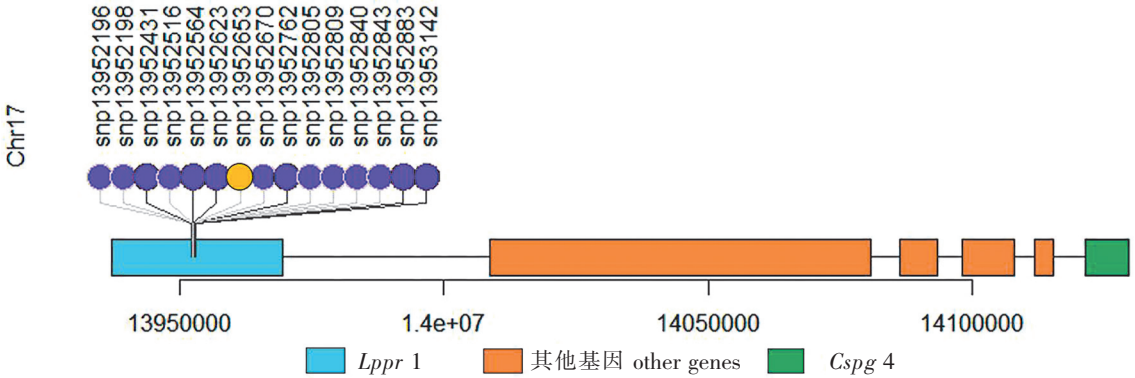
图 5 大黄鱼 *Cspg4* 不同单倍型调控序列重组质粒活性分析
Fig.5 Luciferase analysis of different regulatory haplotypes of *Cspg4* in large yellow croaker

连锁不平衡分析发现，包含 Chr17_ 13952653 位点的 1157 bp 序列中有 14 个 $r^2 = 1$ 的与 Chr17_ 13952653 位点强连锁的 SNP 位点（见表 3、图 6）。这提示该序列对 *Cspg4* 基因转录的弱化作用，可能也涉及诸多强连锁碱基的协同作用。

表 3 *Cspg4* 不同单倍型调控序列区域内连锁不平衡分析

Tab.3 Results of linkage disequilibrium in the region of different allele sequences of *Cspg4*

Chr	位点 1 Site 1	位点 2 Site 2	r^2	Dist
Chr17	13952096	13952653	0.0085	557
Chr17	13952136	13952653	0.0063	517
Chr17	13952187	13952653	0.4286	466
Chr17	13952191	13952653	0.1000	462
Chr17	13952196	13952653	1.0000	457
Chr17	13952198	13952653	1.0000	455
Chr17	13952237	13952653	0.1385	416
Chr17	13952259	13952653	0.0118	394
Chr17	13952263	13952653	0.0118	390
Chr17	13952272	13952653	0.2800	381
Chr17	13952324	13952653	0.0085	329
Chr17	13952377	13952653	0.0140	276
Chr17	13952431	13952653	1.0000	222
Chr17	13952516	13952653	1.0000	137
Chr17	13952539	13952653	0.3333	114
Chr17	13952555	13952653	0.0050	98
Chr17	13952563	13952653	0.0003	90
Chr17	13952564	13952653	1.0000	89
Chr17	13952587	13952653	0.0044	66
Chr17	13952589	13952653	0.3333	64
Chr17	13952602	13952653	0.0808	51
Chr17	13952604	13952653	0.0375	49
Chr17	13952623	13952653	1.0000	30
Chr17	13952653	13952670	1.0000	17
Chr17	13952653	13952706	0.3846	53
Chr17	13952653	13952762	1.0000	109
Chr17	13952653	13952805	1.0000	152
Chr17	13952653	13952809	1.0000	156
Chr17	13952653	13952840	1.0000	187
Chr17	13952653	13952843	1.0000	190
Chr17	13952653	13952883	1.0000	230
Chr17	13952653	13953076	0.0053	423
Chr17	13952653	13953102	0.4000	449
Chr17	13952653	13953142	1.0000	489
Chr17	13952653	13953196	0.3571	543



注:黄色圆圈代表 Chr17_13952653 位点;紫色圆圈代表与其强连锁的 14 个 SNP 位点;蓝色方块代表 *Lppr1* 基因;橙色方块代表序列与 *Cspg4* 之间的其他基因;绿色方块代表 *Cspg4* 基因。

Notes: The yellow circle represents the Chr17_13952653 locus; the purple circles represent the 16 SNP loci strongly linked to it; the blue square represents the *Lppr1* gene; the orange square represents the other genes between the sequence and *Cspg4*; the green square represents the *Cspg4* gene.

图 6 大黄鱼中与 Chr17_13952653 强连锁的 14 个 SNP 相对位置

Fig.6 Relative positions of 14 SNP strongly linked to Chr17_13952653 locus in large yellow croaker

3 讨论

表达 QTL 分析被认为是定位基因表达调控位点的一种有效方法。为了解大黄鱼脾脏 eQTL 中挖掘到的显著 SNP 位点对生物信息学分析预测到的靶基因是否存在调控作用,本研究首先通过 Sanger 测序验证了 81 尾实验鱼中该 SNP 位点的基因型,进而通过构建候选启动子与调控序列片段的重组质粒测定调控序列的活性。

RNA 聚合酶及转录因子通常结合作用在基因的启动子区,欲了解基因表达调控模式必须研究启动子的分子结构和功能^[19]。本研究成功扩增到大黄鱼 *Cspg4* 候选启动子 2096 bp 序列并预测该序列中潜在的核心区域,结果表明有 1 个潜在的核心区域存在于该基因的候选启动子序列中。随后构建 *Cspg4* 基因候选启动子区不同长度缺失片段的重组质粒。分析结果显示: pGL3-*Cspg4*-P4 重组质粒的启动子活性最高,而 pGL3-*Cspg4*-P1 ~ P3 重组质粒相较于 pGL3-*Cspg4*-P4 的启动子活性逐渐降低。表明:在 -2076 ~ -1202 区域可能存在负调控元件;在 -1202 ~ -762 和 -190 ~ +20 区域可能为该启动子的核心区域,存在基本调控元件。而 pGL3-*Cspg4*-P5 ~ P7 重组质粒的启动子活性逐渐增强,说明在 -762 ~ -190 区域可能存在负调控元件。

在转录激活过程中发生的一系列级联反应最终会作用到核心启动子上,直接或间接影响核心启动子上发生的基础转录作用。核心启动子是所有能转录调控 RNA 聚合酶的转录因子的终极作用靶标^[20]。因此,为了研究含 Chr17_13952653 位点的序列对 *Cspg4* 的调控作用,据上述启动子活性分析结果,将存在 *Cspg4* 基因的核心区域的 pGL3-*Cspg4*-P7 (-190 ~ +20) 重组质粒作为对照组,构建含 Chr17_13952653 位点的不同单倍型调控序列的重组质粒。双荧光素酶报告系统结果显示, pGL3-*Cspg4*-P7-A 和 pGL3-*Cspg4*-P7-C 重组质粒的相对荧光活性均显著低于对照组 pGL3-*Cspg4*-P7,而 pGL3-*Cspg4*-P7-A 和 pGL3-*Cspg4*-P7-C 重组质粒之间相对荧光活性无显著差别。pGL3-*Cspg4*-P7-C 重组质粒活性略低,表明含 Chr17_13952653 位点的 1157 bp 序列可能为减弱启动子活性的调控序列,但该位点的 A/C 变异可能不直接影响该调控序列的功能。连锁不平衡分析发现,该序列中有 14 个与 Chr17_13952653 位点强连锁的 SNP 位点 (见图 6),且 r^2 值均等于 1 (见表 3)。

生物的基因组存在大量调控基因表达的序列,包括启动子、增强子、沉默子和绝缘子等。其中的沉默子是一类主动行使抑制启动子活性的调控元件,这些序列能招募转录因子到启动子上,进而执行

特定的功能。已有研究表明许多不同类型的沉默子可影响基因调控的多个方面,如转录因子活性、内含子剪接、染色质结构、3'上游非翻译信号识别及转录起始复合物的装配等,最终下调基因表达^[21]。目前已报道的研究显示:沉默子无特定结构或保守序列,它们可能是一段包含有特异性阻遏蛋白结合位点的短核酸序列;也可能作为一个包含多个调控元件的复杂区域存在,如酵母 HMR-E (hidden mating region-E) 沉默子^[22]。大部分沉默子具有非启动子特异性的特点。Zhao 等^[23]报道在多巴胺转运蛋白基因 (human dopamine transporter gene, *hDAT*) 内含子中发现一个 121 bp 的沉默子,它不仅可在体外下调 *hDAT* 启动子活性,还可以使异源性 SV40 (simian virus 40) 启动子活性下降 80%。典型沉默子对基因表达的阻遏作用表现出位置和方向的非依赖性,主要通过干扰转录起始复合物的装配这一主动抑制过程调控转录,但也有些沉默子对此表现出不同程度的依赖^[24]。B 淋巴细胞瘤-2 基因 (*bcl-2*) 是细胞凋亡的抑制因子。Young 等^[25]在此基因的 5'UTR 中鉴定出一个约 1.3 kb 的负调控元件,将其插入 P1 启动子下游约 2.4 kb 处,启动子的活性未受影响,表明该 NRE 以特定位置的方式降低 *bcl-2* P1 启动子的表达。一些沉默子元件具有远距离作用顺式连接启动子的特性,本研究中含 Chr17_13952653 位点的调控序列相对荧光活性显著低于对照组,该结果表明在大黄鱼脾脏基因中也可能存在类似沉默子抑制转录过程的顺式作用元件。目前研究者们对大黄鱼的顺式调控元件及其功能的研究,主要体现在启动子与转录因子结合层面。如 Sun 等^[26]发现大黄鱼 Toll 样受体家族中 *tlr5m* 和 *tlr5s* 的基因组结构和启动子结构可能存在差异,其中 *oct1* 和 *nf-κb* 结合位点可能是 *tlr5s* 启动子的顺式调控元件, *nf-κb/p65* 通过与 *nf-κb* 结合位点结合,在 *tlr5s* 启动子激活中发挥重要作用。Zhang 等^[27]发现大黄鱼中存在对哺乳动物巨噬细胞的发育至关重要的转录因子 *pu.1* 的两种转录本 *pu.1a* 和 *pu.1b*,它们都可以通过结合其启动子的不同 E26 转化特异性基序,上调组织蛋白酶 (Cathepsin S) 的表达,并且它们的三个结构域都是启动 Ctss 表达的必要条件。Xu 等^[28]发现磷酸化的转录因子 *stat1* 通过结合其启动子区域,显著上调了干扰素调节因子 1 (*irf1*) 的启动子活性。但是本研究在启动子基础上进一步探究 eQTL 分析获得的不同单倍型调控序列对靶基因的作用,为表型与变异之间建立起连接。因此,利用启动子、增强子、沉默子等元件构建人工表达载体,精确调控目的基因的表达,深入研究有关基因表达调控机制可作为未来研究的基础。

综上所述,通过对大黄鱼 17 号染色体 13952653 bp 前后长度为 1157 bp 的调控序列对靶基因 *Cspg4* 启动子的表达调控研究,为 *Cspg4* 基因在大黄鱼脾脏中的转录调控及功能研究奠定了基础。但迄今为止,对该基因参与免疫系统的调节且在大黄鱼脾脏组织表达的分子机制还知之甚少,其具体表达调控方式与作用机制有待进一步深入研究。

[参考文献]

- [1] 张丹枫,安树伟,周素明,等. 大黄鱼 (*Pseudosciaena crocea*) 内脏白点病的组织病理和超微病理分析[J]. 渔业科学进展,2017, 38(4): 11-16.
- [2] 胡娇,张飞,徐晓津,等. 大黄鱼 (*Pseudosciaena crocea*) 内脏白点病病原分离鉴定及致病性研究[J]. 海洋与湖沼, 2014, 45(02): 409-417.
- [3] ZHANG J T, ZHOU S M, AN S W, et al. Visceral granulomas in farmed large yellow croaker, *Larimichthys crocea* (Richardson), caused by a bacterial pathogen, *Pseudomonas plecoglossicida* [J]. Journal of Fish Diseases, 2014, 37(2): 113-121.
- [4] STALLCUP W B. The NG2 antigen, a putative lineage marker; immunofluorescent localization in primary cultures of rat brain [J]. Developmental Biology, 1981, 83(1): 154-165.
- [5] PLUSCHKE G, VANEK M, EVANS A, et al. Molecular cloning of a human melanoma-associated chondroitin sulfate proteoglycan [J]. Proceedings of the National Academy of Sciences of the United States of America, 1996, 93(18): 9710-9715.
- [6] GIACOMINI P, VEGLIA F, FEI P C, et al. Level of a membrane-bound high-molecular-weight melanoma-associated antigen and a cytoplasmic melanoma-associated antigen in surgically removed tissues and in sera from patients with melanoma [J].

- Cancer Resaerch,1984,44(3):1281-1287.
- [7] RETTIG W J, REAL F X, SPENGLER B A, et al. Human melanoma proteoglycan: expression in hybrids controlled by intrinsic and extrinsic signals[J]. Science,1986, 231(4743): 1281-1284.
- [8] MAJUMDAR M, VUORI K, STALLCUP W B. Engagement of the NG2 proteoglycan triggers cell spreading via rac and p130cas[J]. Cellular Signalling,2003, 15(1): 79-84.
- [9] ASHER R A, MORGENSTERN D A, PROPERZI F, et al. Two separate metalloproteinase activities are responsible for the shedding and processing of the NG2 proteoglycan *in vitro*[J]. Molecular and Cellular Neuroscience,2005, 29(1): 82-96.
- [10] CHEKENYA M, KRAKSTAD C, SVENDSEN A, et al. The progenitor cell marker NG2/MPG promotes chemoresistance by activation of integrin-dependent PI3K/Akt signalling[J]. Oncogene,2008, 27(39): 5182-5194.
- [11] JOO N, WATANABE T, CHEN C, et al. NG2, a novel proapoptotic receptor, opposes integrin $\alpha 4$ to mediate anoikis through PKC α -dependent suppression of FAK phosphorylation[J]. Cell death and differentiation,2008, 15(5): 899-907.
- [12] RICCARDO F, IUSSICH S, MANISCALCO L, et al. CSPG4-specific immunity and survival prolongation in dogs with oral malignant melanoma immunized with human CSPG4 DNA[J]. Clinical Cancer Research,2014, 20(14): 3753-3762.
- [13] LEE Y H, KAWAKAMI K, H FU W C, et al. Chondroitin sulfate proteoglycan 4 regulates zebrafish body axis organization via Wnt/planar cell polarity pathway[J]. Plos One,2020, 15(4): e0230943.
- [14] KARMOUCH J, ZHOU Q Q, MIYAKE C Y, et al. Distinct cellular basis for early cardiac arrhythmias, the cardinal manifestation of arrhythmogenic cardiomyopathy, and the skin phenotype of cardiocutaneous syndromes[J]. Circulation Research, 2017, 121(12): 1346-1359.
- [15] ZHU X, BERGLES D E, NISHIYAMA A. NG2 cells generate both oligodendrocytes and gray matter astrocytes[J]. Development,2008, 135(1): 145-157.
- [16] WILMS C, LEPKA K, HÄBERLEIN F, et al. Glutaredoxin 2 promotes SP-1-dependent *Cspg4* transcription and migration of wound healing NG2 glia and glioma cells: enzymatic taoism[J]. Redox Biology,2022,49: 102221.
- [17] Van der AUWERA G A, CARNEIRO M O, HARTL C, et al. From FastQ data to high confidence variant calls: the genome analysis toolkit best practices pipeline[J]. Current Protocols in Bioinformatics,2013, 43(1110): 11.10.1-11.10.33.
- [18] ZHANG C, DONG S S, XU J Y, et al. PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files[J]. Bioinformatics,2019, 35(10): 1786-1788.
- [19] 曾晓玲, 赵昶灵, 文国松, 等. 启动子结构、功能预测和验证方法的研究进展[J]. 分子植物育种,2018, 16(12): 3915-3925.
- [20] SMALE S T, KADONAGA J T. The RNA polymerase II core promoter[J]. Annual Review of Biochemistry,2003, 72(1): 449-479.
- [21] OGBOURNE S, ANTALIS T M. Transcriptional control and the role of silencers in transcriptional regulation in eukaryotes[J]. Biochemical Journal,1998, 331(Pt 1): 1-14.
- [22] CHESNOKOV I N. Multiple functions of the origin recognition complex[J]. International Review of Cytology,2007,256:69-109.
- [23] ZHAO Y, ZHOU Y, XIONG N, et al. Identification of an intronic cis-acting element in the human dopamine transporter gene[J]. Molecular Biology Reports,2012, 39(5): 5393-5399.
- [24] 陈妍, 王金发. 负调控元件: 沉默子[J]. 遗传,1998,20(1):45-48.
- [25] YOUNG R L, KORSMEYER S J. A negative regulatory element in the bcl-2 5'-untranslated region inhibits expression from an upstream promoter[J]. Molecular and Cellular Biologg,1993,13(6):3686-3697.
- [26] SUN Q X, HUANG X N, PAN Y, et al. Promoter analysis of TLR5M and TLR5S revealed NF- κ B might be a conserved cis-element in TLR5S promoter of large yellow croaker[J]. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology,2022,257: 110675.
- [27] ZHANG X Y, ZHUO X, CHENG J, et al. PU.1 regulates cathepsin s expression in large yellow croaker (*Larimichthys crocea*) macrophages[J]. Frontiers in Immunology,2022,12: 819029.
- [28] XU D, LI Q, ZHOU Y, et al. Functional analysis and regulation mechanism of interferon gamma in macrophages of large yellow croaker (*Larimichthys crocea*) [J]. International Journal of Biological Macromolecules,2022,194: 153-162.

(责任编辑 朱雪莲 英文审校 黄力行)