

真鲷(*Pagrus major*)*vasa* 基因 5'端启动子克隆 及表达载体构建*

林帆^{1,2} 刘清华¹ 李军¹ 隋娟^{1,2} 肖志忠¹
徐世宏¹ 马道远¹ 肖永双¹

(1. 中国科学院海洋研究所 青岛 266071; 2. 中国科学院研究生院 北京 100049)

摘要 采用染色体步移(Genome walking)的方法克隆了真鲷 *vasa* 基因 5'侧翼序列, 采用生物信息学方法分析潜在的顺式作用元件, 并与斑马鱼核心启动子进行比对, 在此基础上构建了绿色荧光蛋白表达载体。序列分析结果显示: 克隆得到的真鲷 *vasa* 基因 5'侧翼序列序列长度为 2762bp, 其中包括 TATA-box、CAAT-box、SP-1、GAGA-1、OCT-1、v-Myb、Sox-5、SRY、HNF 等可能对 *vasa* 基因转录调控起重要作用的顺式作用元件。潜在的核心启动子区与斑马鱼核心启动子具有同源性 (61.1%), 推测可能对转录调控具有重要的作用。采用 PCR 方法扩增起始密码 5'侧翼全长片段, 连接入去除 CMV 启动子的绿色荧光蛋白载体, 最终成功构建真鲷 *vasa* 基因 5'侧翼表达载体, 为进一步研究分析启动子效率和绿色荧光蛋白在 PGCs 中的特异表达提供了基础。

关键词 真鲷, *vasa*, 启动子, PGCs

中图分类号 Q344

vasa 基因是 DEAD-box 基因家族的重要成员之一, 最先在果蝇(*Drosophila melanogaster*)中发现。研究者最初的目的是用遗传学方法筛选改变果蝇胚胎前-后极性的母体效应突变体, 结果显示一些基因对极颗粒的形成和生殖细胞的决定起着至关重要的作用, 其中 *vasa* 基因专一定位于极颗粒并在整个生殖细胞的发育过程中特异性表达(Lasko *et al*, 1988)。

在绝大多数物种中, *vasa* 基因仅限在生殖细胞系中特异性表达, 因此作为一种分子标记物广泛应用于研究配子发生和原始生殖细胞(PGCs)起源、迁移、分化等方面(周倩如等, 2007)。早期学者主要采用整体原位杂交技术追踪斑马鱼(*Danio rerio*)(Yoon *et al*, 1997)、青鳉(*Oryzias latipes*)(Shinomiya *et al*, 2000)和金鱼(*Carassius auratus*)(Otani *et al*, 2002)等 PGCs 的起源、迁移及分化。近年来有学者在斑马鱼(Krovel *et al*, 2002)、虹鳉(*Oncorhynchus mykiss*)(Yoshizaki *et al*,

2000)、青鳉(Tanaka *et al*, 2001)、鸡(*Gallus gallus*)(Minematsu *et al*, 2008)和小鼠(*Mus musculus*)(Gallardo *et al*, 2007)等物种中, 利用 *vasa* 基因启动子的功能构建 *vasa* 和绿色荧光蛋白(GFP)的融合蛋白, 以此为标记物, 更直观地追踪 PGCs。在虹鳉中, 利用流式细胞仪分离带绿色荧光蛋白标记的虹鳉 PGCs, 进行了短期冷冻保存并成功移植到鲑鱼中, 产下正常的虹鳉后代(Takeuchi *et al*, 2002), 创建了鱼类种质资源保护的新方法。尽管越来越多的学者利用 *vasa* 启动子构建 *pvasa*-EGFP 载体研究 PGCs, 但目前对于 *vasa* 基因启动子方面的研究报道仅见于鸡(Minematsu *et al*, 2008)、斑马鱼(Bartfai *et al*, 2003)中。

真鲷(*Pagrus major*)隶属鲈形目(Perciformes)、鲷科(Sparidae), 主要分布于印度洋和北太平洋西部, 为黄、渤海名贵经济种类。20 世纪 90 年代起, 随着海水养殖业的迅速发展, 很多经济鱼类出现了野生

* 国家高技术研究发展计划(863 计划)课题, 2006AA10A404 号; 国家自然科学基金项目, 31072212 号, 41076100 号。林帆, E-mail: whhit21@126.com

通讯作者: 李军, 研究员, E-mail: junli@ms.qdio.ac.cn

收稿日期: 2010-12-03, 收修改稿日期: 2011-02-01

资源急剧下降、养殖群体遗传多样性下降等一系列问题(黎中宝等, 2009), 因此对其种质鉴定和资源保护的要求也日益突出(陈松林等, 2007)。尽管精子低温保存技术相对成熟, 但胚胎保存仍未取得令人满意的进展。PGCs 较胚胎更易于保存, 且具有分化为两性配子的潜能, 因此 PGCs 的保存无疑提供了一种新的有潜力的种质资源保存手段(Yoshizaki *et al*, 2003)。本文在真鲷 *vasa* mRNA (GenBank: AB378581)基础上利用 Genome walking(染色体步移)方法, 克隆了 5'端侧翼区域, 然后通过生物信息学手段进行初步序列分析鉴定, 并构建 *pvasa*-EGFP 表达载体, 可为研究 *vasa* 启动子在 PGCs 中特异表达, 进而为标记分离 PGCs 提供基础。

1 材料与方法

1.1 实验材料

真鲷(*Pagrus major*)精子由本实验室保存, 用来提取基因组 DNA。海洋动物组织提取试剂盒、高纯质粒中量小提提取试剂盒购自天根升华科技有限公司, 胶回收试剂盒、PCR 产物纯化试剂盒购自杭州博日科技有限公司, Genome walking kit、PMD19-T 载体、LA *Taq* 酶(含 dNTPs 和 10 × buffer)、DNA marker、DNA ligation kit 购自 TAKARA 公司, 大肠杆菌感受态 JM109 购自北京全式金生物技术有限公司。主要实验仪器为 Life Express Thermal Cycle PCR 热循环仪, 引物合成及测序服务分别由上海生物工程公司和上海桑尼生物科技有限公司提供。pEGFP-N1 表达载体由巴斯德研究所王卓女士惠赠。

1.2 实验方法

1.2.1 真鲷基因组 DNA 的提取 取 15 μl 解冻真鲷精子, 按照试剂盒说明书进行基因组 DNA 提取。

1.2.2 *vasa* 基因第一内含子克隆 由于 *vasa* mRNA 序列 5'UTR 较短(仅为 168bp), 为了获得足够的分析信息, 根据 *vasa* mRNA 序列, 在 5'UTR 和起始密码子之间设计引物扩增第一内含子片段。PCR 产物回收纯化后送往桑尼公司测序。引物序列如下:

Vasa E1F: 5'-GAGTGTCTTGACAAGGCATAGC-3'

Vasa E2R: 5'-TCTTCTTCCCCTCGTCCAT-3'

1.2.3 *vasa* 基因 5'端侧翼序列克隆 根据第一内含子测序结果, 设计 3 条基因特异性上游步移引物分别为:

Vasa UP1-1: 5'-AGTGCTCGGAAAGACAACATA ACC-3'

Vasa UP1-2: 5'-CGTCCCTTCTGCTCTATCAAAA

TGT-3'

Vasa UP1-3: 5'-GTAACTCCGTCGTCTTCAGC TATG-3'

利用设计的 3 条步移引物, 分别与 Genome Walking Kit 试剂盒提供的引物组合进行步移 PCR 扩增, 操作按照说明书进行, 连续进行 3 次步移扩增。1%琼脂糖凝胶电泳, PCR 产物测序, 根据测序结果设计 3 条第二轮步移引物, 同上进行 3 次步移扩增, PCR 产物回收纯化, 连接 PMD19-T 载体, 转化 JM109 大肠杆菌感受态, 于含有 AMP(100 μg/μl) LB 平板, 37 °C 过夜, 挑单菌落, PCR 鉴定, 阳性克隆测序。3 条引物序列如下:

Vasa UP2-1: 5'-CATTTTGTCTTTTCACCATCC AG-3'

Vasa UP2-2: 5'-CACGATGACTTGTTGTTGTGT ATAG-3'

Vasa UP2-3: 5'-TCTGCATCTCTACACCCAAAC ACT-3'

1.2.4 *vasa* 基因 5'端序列分析 利用 DNAMAN 软件拼接两次扩增所得到的 *vasa* 基因 5'端侧翼序列, 利用在线分析工具 NNPP(Neural Network Promoter Prediction)以及 Promoter 2.0 预测潜在转录起始位点; *vasa* 基因 5'端侧翼序列潜在转录因子结合位点预测采用 Alibaba2.1 及 TFSEARCH, CpG 岛预测采用 CpG Island Predicton。预测使用参数均采用在线默认设置, 在线网址如下:

http://www.fruitfly.org/seq_tools/promoter.html

<http://www.cbs.dtu.dk/services/Promoter/>

<http://www.gene-regulation.com/pub/programs/alibaba2/index.html>

<http://www.cbrc.jp/research/db/TFSEARCH.html>

http://www.ualberta.ca/~stothard/javascript/cpg_islands.html

利用 DNASTAR 软件对斑马鱼和真鲷的 5'侧翼进行比对分析。

1.2.5 *pvasa*-EGFP 表达载体构建 在用生物信息学分析 *vasa* 基因 5'侧翼的基础上, 设计引物扩增克隆到的全长 5'端启动子序列并引入 *Hind* 和 *Bam* 位点(下划线部分)。引物序列如下:

Vasa5-F: 5'-TAGCAAGCTTCCAAGCTGTCGCCG CTGAT-3'

Vasa5-R: 5'-GTCAGGATCCTCTTCTTCCCCTC GTCCAT-3'

用 *Ase* 和 *Nhe* 双酶切去掉 pEGFP-N1 的 CMV 启动子序列, 然后用 Klenow Fragment 补平, 自连成 pEGFP-4.1。用 *Hind*、*Bam* 双酶切扩增得到的启动子序列和 pEGFP-4.1 载体连接, 转化 JM109 感受态,

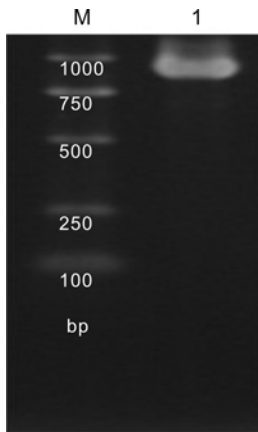


图1 真鲷 *vasa* 第一个内含子扩增电泳图

Fig.1 PCR amplification and electrophoretic analysis of the first intron of the red seabream *vasaM*: DL2000 DNA 分子量标记, 1: *vasa* 第一个内含子扩增产物

5'侧翼经过了两轮染色体步移, 第一轮步移(UP1-1、UP1-2、UP1-3)得到约 1.2kb 的扩增片段(图3上), 根据测序结果设计第二轮步移引物(UP2-1、UP2-2、UP2-3)进行第二轮步移, 得到约 2.2kb 扩增片段(图3下), 连接 PMD19-T 载体、阳性克隆测序, 拼接最终得到 5'端侧翼 2762bp 序列。

2.3 真鲷 *vasa* 基因 5'端生物学信息学分析

利用在线分析工具 NNPP 以及 Promoter 2.0 预测潜在转录起始位点, 并将预测结果(图4, 可能的转录起始位点用大号字体标出)与斑马鱼 *vasa* 核心启动子序列比对(图5), 发现 1130—1180bp 区具有较高的同源序列 (61.1%)。

采用 Alibaba 2.1 及 TFSEARCH 对潜在转录因子结合位点进行预测结果如图6所示, 在克隆所得到的真鲷 *vasa* 基因 5'侧翼中存在 TATA BOX、CAAT BOX、

菌落 PCR 鉴定阳性克隆。在含 Kan 抗性(40 μ g/ μ l) LB 液体培养基中, 37 $^{\circ}$ C, 180r/min 过夜, 少量抽提质粒, 并将产物命名为 *pvasa*-EGFP。

2 结果

2.1 内含子扩增结果

PCR 扩增电泳结果显示扩增片段大小在 900bp 左右(图1), 扩增产物回收后测序, 经过与真鲷 *vasa* mRNA 比对分析后得到 852bp 第一内含子序列(图2)。

2.2 真鲷 *vasa* 基因 5'侧翼染色体步移结果

SP-1、GAGA-1、OCT-1、v-Myb、Sox-5、SRY、HNF 等主要反应元件, 同时发现在 5'非编码区 CG 含量丰富(Obs/Exp>0.65, %GC>50.00)。

2.4 *pvasa*-EGFP 表达载体构建

根据测序拼接结果, 设计引物(*vasa*5-F、*vasa*5-R)扩增 *vasa* 基因起始密码子 5'侧翼(含第一内含子和 5'非编码区)得到预期的 3.8kb 片段(图7)。pEGFP-N1 (4.7kb)载体经用 *Ase* 和 *Nhe* 酶切去除 CMV 启动子之后, Klenow 片段补平自连成 pEGFP-4.1。将 3.8kb 侧翼序列插入 pEGFP-4.1 后, 菌落 PCR 鉴定阳性克隆(图8), 得到 *pvasa*-EGFP 重组载体。

3 讨论

启动子作为基因表达调控的关键序列区域, 是转基因工程研究的热点。因此, 研究启动子功能序列, 对于基因表达调控机制的研究是很有必要的(陈维宇等, 2008; 刘秀霞等, 2009)。目前, 鱼类仅在斑马鱼(Krovel *et al*, 2002)、虹鳟(Yoshizaki *et al*, 2000)、青鳉(Tanaka *et al*, 2001)等物种中有利用 *vasa* 启动子驱动目标蛋白在原始生殖细胞中特异表达的报道。

由于已报道的真鲷 *vasa* mRNA 中 5'UTR 序列较短仅 168bp, 不足以设计 3 条染色体步移引物, 因此首先克隆了第一个内含子片段, 为后来的染色体步移提供了基础, 结果显示该片段具有明显的真核生物内含子两端序列特征: 5'端为 GT, 3'端为 AG。通过两轮染色体步移最终得到了 2762bp 的侧翼序列。



图2 *vasa* 基因部分序列与 *vasa* mRNA 部分序列比对结果

Fig.2 Sequence alignment of partial *vasa* gene and partial submitted *vasa* mRNA sequence
注: 阴影部分代表 mRNA 序列, 加粗字母代表剪接位点

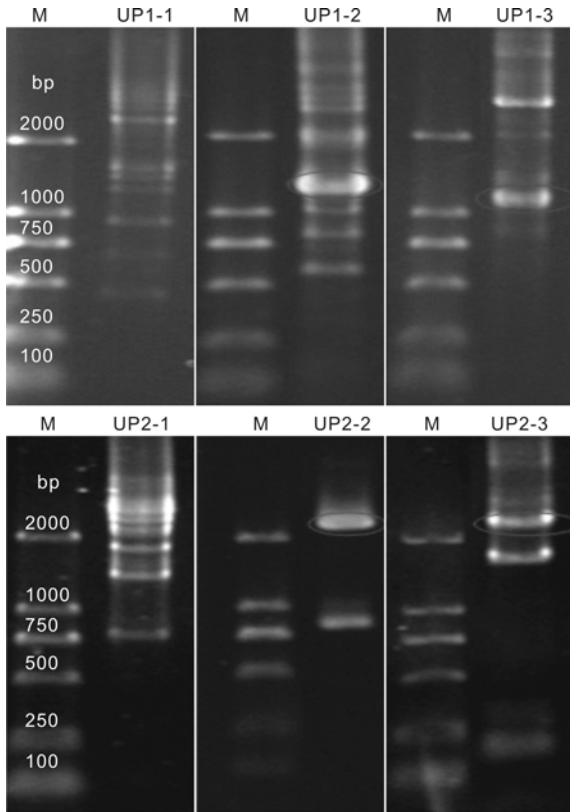


图 3 第一轮和第二轮步移 PCR 产物电泳图

Fig.3 The result of *vasa* 5' flanking genome walking for the first and second times

上图: 第一轮步移; 下图: 第二轮步移。M: DL2000 DNA 分子量标记; UP1-1、UP2-1: 第一次步移 PCR 产物; UP1-2、UP2-2: 第二次步移 PCR 产物; UP1-3、UP2-3: 第三次步移 PCR 产物。圆圈表示目的条带

利用生物信息学软件分析了相应的顺式作用元件, 发现在真鲷 *vasa* 5'侧翼区, 除了包括真核生物比较典型的 TATA-box、CAAT-box 外, 还包括 SP-1、GAGA-1、v-Myb、OCT-1、Sox-5、SRY、HNF 等多个转录因子结合位点。其中 Oct-1 是 POU 同源异型结构域家族的一员, 在细胞中普遍表达, 可调节免疫球蛋白、组蛋白 H2B 以及细胞周期蛋白 D1 等细胞寿命必需相关因子的表达。研究表明, Oct-1 蛋白选择性的结合在小鼠的 Nanog 基因(在 PGCs 和多能干细胞中特异表达)启动子序列。这就意味着 Oct-1 可以通过调节 *vasa* 及 Nanog 基因的表达进而调节生殖细胞增殖和发育(Takeo *et al*, 2008)。SRY

基因, 是 HMG-box 家族的一员, 编码一种转录因子, 它的突变将导致精巢发育不良, 是 Y 染色体上雄性决定基因(Denny *et al*, 1992)。Sox-5 基因是 SRY 相关基因的一员, 具有与 SRY 基因 HMG-box 高度同源的结构, 能与 DNA 特异结合, 研究表明精巢中有 Sox-5 mRNA 表达, 减速第二次分裂的生殖细胞中可以检查到 Sox-5 蛋白(Denny *et al*, 1992)。尽管 SRY、Sox-5 是哺乳动物中的雄性决定基因, 但似乎鱼类性别决定具有某种相似的机制, 即通过某种性别决定基因编码的蛋白直接或间接调节下游与生殖嵴分化相关效应基因的表达。HNF-4 是一种孤独核受体, 在肝脏、肾脏和肠道中表达, 与胆固醇代谢、糖代谢、尿素生物合成、血液凝集、肝脏分化有关(Takeo *et al*, 2008)。在鸡 *vasa* 启动子中也发现类似的转录因子结合位点(Oct-1、HNF-4、SRY 等), 这些转录因子对于真鲷 *vasa* 基因的表达可能起调控作用, 但与其与配子形成的关系仍需要进一步试验证明。

此外, DNA 的甲基化能调节基因的表达, 研究报道表明人(*Homo sapiens*)*vasa* 基因 GC 岛的甲基化能导致 *vasa* 基因沉默, 进而导致雄性成熟停滞(Sugimoto *et al*, 2009)。作者通过分析, 在真鲷 *vasa* 基因 5'非编码区域发现 GC 含量丰富区, 可能具有调节 *vasa* 基因表达的作用, 在斑马鱼中也有相似的报告。因此在构建 *pvasa*-EGFP 载体时, 作者将该区域也包含在内。由于虹鳟、青鳉 *vasa* 基因启动子区序列无法获得, 无法比对保守位点, 只比对了真鲷和斑马鱼 *vasa* 的启动子区, 未发现明显的保守序列, 这可能是由于种之间的差异引起的。但在核心启动子区域具有较高的相似性, 说明这段序列可能对转录调控起重要的作用。

起点	终点	分值	序列
1130	1180	0.93	TGGAATTGTATAACTGGGAAACTCAGACGCTTGTGTATTCAATGCTACGT

图 4 真鲷 *vasa* 转录起始位点预测结果

Fig.4 Predicted transcription start site for red seabream *vasa* gene



图 5 真鲷 *vasa* 潜在核心启动子区与斑马鱼 *vasa* 核心启动子区比对结果

Fig.5 Blast result of *vasa* potential core promoters of red seabream and zebrafish

1 AGTGGAGTAGCATAGGCCAACTGTCGCCGTGATCAGCCCCAGAACACTGTTTGAAGCTAGAAAGGTGGCAGGTCCG
 81 CCACATATAAA^{TATA}CAAAGTAAACA^{TATA}TTGTGTGTCTTGGAGTCA^{TTTGTTTA}TCAGTCATGAAAACAAGGAG
 TATA SRY TATA SRY
 161 AGTTTGTTTAGGCAGAAAAAATCAGTCAAAGATCTTCT^{CGTCTGATTACA}TTTCTTTCCA^{AAACT}CATAGTGCAC
 OCT-1 SRY
 241 CTTTAA^{TATA}AAAGGAAC^{GTTGT}TTTTTAAAAAAAAAAAAAAAAACA^{TAAACAATT}TTGGATCAATCTGACA
 TATA SRY Sox-5
 321 TCATCGGAATGGTCTGGTCTGTCTTTACAGGAAGTCTGTACAATGAAAAACATGTTATGTTCACTGGCATATAACATT
 401 CCATTTCTCAAGTATAAGGCTGTCAATGGAATTGAAATTAGGTTACAGCATTACTACTGGATCACAATGAAGCCATTTA
 481 ACCACTGGGTGAAAGCAAAAAATGTGTCA^{ATAATCTTTGT}TTTATGAAATAAAAAGACATTTTATGTCACTGTTC
 HNF-3B
 561 AATTACATGCCATGTAATGTGTACAGCACTGTTCAAGTAAAGTGGGACACCATTTTATTCTTCATTATGCTC^{AAAGAT}
 641 ^{AAACAGCT}TGCACCCAACTGTAGTCTAAAGCCAAAGTCTCGGGTTATAGACTCTAAGAGACCCACAGGATCACTCG
 OCT-1
 721 CAATTTAGCCGACAGTACTGTCTGAAGAAGCTGATGTCCACTCATTTTAAAGCTAAGCTGAAGGTAGTGGTATAATG
 801 TGAACCACAAAACCTCAGGCATCTATATCATGCTGGTCTGTCGGGAAGGGATAAATAAT^{CTCCAGAGT}^{CAGCAA}
 SPI
 881 ^{ATTATGG}GAGGGAAAACCTGGCATGCCATTTTCAAGCGTTCCTTGACCTCTCAACTCTAGAATCACAACCTCTCCC
 HNF-3B
 961 TTTATTGGTACACCACTTTATGATAATAATCCATGCAGTTTTTTTTGGCAAAAACCATGCAATTTCTTCCATGCAGT
 1041 GTAATGTATTATTTCCAC^{CAAT}TCTATATGGTGTATTGAGTAATCTGCATACTGGAGTCATTAATGTCTTGGAA
 CAAT
 1121 TTGCATAAATGGAATT^{TATA}CTGGGAAAACCTCAGACGCTGTGTATTC^AATGCTACGTTTGAAGCGGTGAGCCTTTTA
 TATA
 1201 AA^{AAAGATGCATATTC}CAATTTAATCTTACT^{TAAAATGTAATAAGG}TCTAAAATGGTTTATAGAAGCAGTTTATGT
 OCT-1 OCT-1
 1281 AAGTCTACTAGAGCATCTACCAAAAAAGTGGCACATTACCTGAATTCACCCTATATTTTATTTATATAGTCTTTTCATG
 1361 ACACTCAAAGATACTTTACATAGTTTCATTA AAAACATGTGA^{AAACAAC}ATCATCAGAGATCAAAATCAGTACTGAAAA
 SRY GATA-1
 1441 GGTGAGTTTAGAGGGGTAGCTGTGGAAGGCTAATGCCACAGGCTGGTGCTTATTCCTGAGTGGTGGCAGCAGAT
 1521 TGGCATCAGAGGAGCGGAGGCTGCCAGGAGTAATGTTGTGGAGGAGCAGTTTCATGGGTTAGGAAGGAGTCTAGTTATGG
 1601 AGGCTTTGTGAGGTAGGAGAGGGACTTTGAATTGGATCTGTTGTGGAGCAGGAGCCCGTGAAGTTCTGGTCATGGGA
 1681 GTGAGAGTGGGTCGCACTCAAGCAGCAGAGTCTCGATCGTCATTTAAGACGATGATGTACCATCAATAACGCAGTTGC
 1761 AGTAGTCAGGCTGGATGTGATGAATTAATGGATCAGTTCACGACGAGTATCTTTAGCTTCACTGATCTCAAGT
 1841 TCTTGTGAGTCTCCTGCACAATACAGTGCCTCCAGCCATGAGAGATCTACATC^{TACCTACACTAC}CTACAGTTG
 SPI
 1921 AAGTGAGGACTCACTTCAACTAAATTGAAGCAAAAGCAAAATCAGTCACAAAACATATGTACCAACATACATGTTACC
 2001 AATGTGGTTGAACAGGTTGGTTTTGGTTTTGTTAAGAAAACGAAGAAAAAGAAGCCACACACAAGTGAAGTGTGGGTG
 2081 TAGAGATGCAGAGACATGCAATAAAGCAGAGTGGAGTGGGACCGACCTGCTCACTGTCAGGTCCTTTAGCTATACAAA
 2161 CAACAAGTCATCGTGCAC^{GGTGATTGT}TGTTTACATGTATTAGACTAGCTTA^{CATGATGGT}TGGATGGTGAAG
 GATA-1 SRY GATA-1
 2241 ACAAATGTCACCTTCAGACAACATTCAGAGAACGTGTTCTGAGGGTCTTTAGAAACCTTGAAGAACCTTAACGAACG
 2321 TTTCTAAAGGTTCTTAAAGGTTAAAT^{ACCGTTAGA}CAGCCTTCAGGGAACATTTTCAGTTTCGAAGCTGTCACA
 V-Myb
 2401 TTTGTCTGGTCAAATGTGTTAATTTAGCCTACCTGGCCCTCTGTTCTTCACTTGTCTTCTAATTGATACACTTGAGTTT
 2481 TACAGTTAAATTAAGTGGTAAAGAATATGAGTAACTGAACATTGAGGCTGCCTATAGGTGGAAGGGTGAAGTGGTGGT
 2561 TGAGTTTGCTTGAATTTGAATCTGGGGTCGGTATATCATGTTGTTGGTACTGTTGCTGTTTTTCTGTCAATTTTCCA
 2641 AAGTAGTTCTTTAAATGGAATACTGGACCTTGAATCTTTGAGCCAGACAAGGCAATAGGAAAAGCGACCCCTTCC
 2721 TCCATCATCTGAATCTGAATTCATCTCAATTCACCTTTAGTTGATTTTCAAAAAAGAAGTCCAGCTGGGCGGAAATC
 2801 AGCGGAACCTGGCAACCTAGCGAGCTGCTCCGTGTGAGAGTGCAGAGTGCACACGGCTCTGAGTGTCTTGACAAGGCA
 2881 TAGCTGAAGACGACGGAGTTAACCCTGAATTCACCAGAGAA

图6 真鲷 *vasa* 基因 5'侧翼序列及对部分潜在转录因子结合位点、GC 含量丰富区预测结果

Fig.6 The 5' flanking region of red seabream *vasa* gene and prediction of potential transcription factor binding site and GC-rich region
 注: 小方框或下划线代表转录因子结合位点, 大号字母代表潜在转录起始位点, 阴影表示 GC 含量丰富区

综上所述, 本实验成功克隆了真鲷 *vasa* 基因 5'侧翼 2762bp, 根据生物信息学分析, 发现了可能对基因转录调控起重要作用的潜在顺式作用元件, 并构建了 *pvasa*-EGFP 表达载体, 为后续研究 *vasa* 启动子启动绿色荧光蛋白基因在原始生殖细胞中特异表达奠定了基础, 此启动子序列的效率, 以及作者预测的调控元件的真实性, 需要通过片段缺失、序列突变分析等方法进一步试验证明。

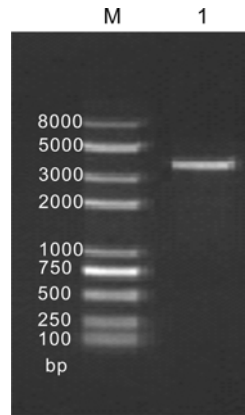


图 7 *vasa* 基因 5'端侧翼扩增结果

Fig.7 Amplification result of 5' flanking region of *vasa* gene
M: Trans2K DNA 分子量标记物, 1: *vasa* 5'端扩增产物

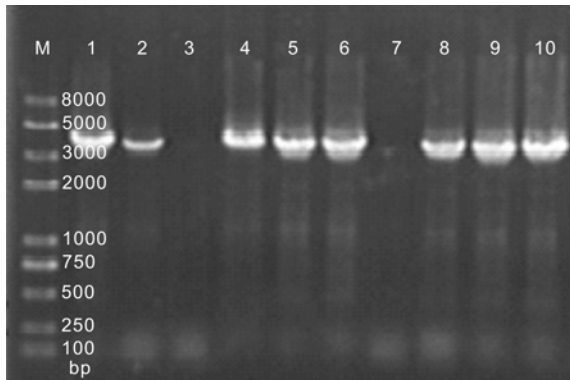


图 8 菌落 PCR 鉴定阳性克隆

Fig.8 Identification of *pvasa*-EGFP by colony PCR
M: Trans2K DNA 分子量标记物, 1—10 表示菌落 PCR 产物

参 考 文 献

- 刘秀霞, 梁旭方, 王琳等, 2009. 鳊鱼(*Siniperca chuatsi*) - 肌动蛋白基因 cDNA 全序列与 5'侧翼区的克隆与分析. 海洋与湖沼, 40(1): 102—108
- 陈松林, 田永胜, 李军等, 2007. 鱼类精子和胚胎冷冻保存理论与技术. 北京: 中国农业大学出版社, 366—367
- 陈维宇, 李卫东, 柯才焕, 2008. 杂色鮑肌动蛋白的克隆与序列分析. 厦门大学学报, 47(5): 624—629
- 周倩如, 邵明瑜, 张志峰, 2007. *vasa* 基因编码蛋白的结构特征和应用展望. 海洋湖沼通报, 4: 129—134
- 黎中宝, 方秀, 陈锦等, 2009. 大黄花鱼(*Pseudosciaena crocea*)养殖群体遗传多样性的降低. 海洋与湖沼, 40(4): 446—450
- Bartfai R, Orban L, 2003. The *vasa* locus in zebrafish: Multiple RGG box from intragenic duplications. DNA and Cell Biology, 22(1): 47—54
- Denny P, Swift S, Connor F *et al*, 1992. An SRY-related gene expressed during spermatogenesis in the mouse encodes a sequence-specific DNA-binding protein. The EMBO Journal, 11(10): 3705—3712
- Gallardo T, Shirley L, John G *et al*, 2007. Generation of a Germ Cell-Specific Mouse Transgenic Cre Line, *Vasa-Cre*. Genesis, 45: 413—417
- Krovel A, Olsen L, 2002. Expression of a *vasa*: EGFP transgene in primordial germ cells of the zebrafish. Mechanisms of Development, 116: 141—151
- Lasko P F, Ashburner M, 1988. The product of the *Drosophila* gene *vasa* is very similar to eukaryotic initiation factor-4A. Nature, 335: 661—617
- Minematsu T, Harumi T, Naito M, 2008. Germ cell-specific expression of GFP Gene Induced by Chicken *vasa* Homologue (*Cvh*) Promoter in Early Chicken Embryos. Molecular Reproduction and Development, 75(10): 1515—1522
- Otani S, Maegawa S, Inoue K *et al*, 2002. The Germ Cell Lineage Identified by *vasa*-mRNA during the Embryogenesis in Goldfish. Zoological Science, 19: 519—526
- Shinomiya A, Tanaka M, Kobayashi T *et al*, 2000. The *vasa*-like gene, *olvas*, identifies the migration path of primordial germ cells during embryonic body formation stage in the medaka, *Oryzias latipes*. Develop Growth Differ, 42: 317—329
- Sugimoto K, Koh E, Narimoto K *et al*, 2009. Tissue-specific differentially methylated regions (TDMRs) of the human *vasa* gene are potentially associated with maturation arrest phenotype in the testis. Journal of Human Genetics, 54(8): 450—456
- Takeuchi Y, Yoshizaki G, Kobayashi T *et al*, 2002. Mass Isolation of Primordial Germ Cells from Transgenic Rainbow Trout Carrying the Green Fluorescent Protein Gene Driven by the *vasa* Gene Promoter. Biology of Reproduction, 67: 1087—1092
- Tanaka M, Kinoshita M, Kobayashi D *et al*, 2001. Establishment of medaka (*Oryzias latipes*) transgenic lines with the expression of green fluorescent protein fluorescence exclusively in germ cells: a useful model to monitor germ cells in a live vertebrate. Proceedings of the National Academy of Sciences (PNAS), 98(5): 2544—2549
- Yoon C, Kawakami K, Hopkins N, 1997. Zebrafish *vasa* homologue RNA is localized to the cleavage planes of 2- and 4-cell-stage embryos and is expressed in the primordial germ cells. Development, 124: 3157—3166
- Yoshizaki G, Takeuchi Y, Sakatani S *et al*, 2000. Germ cell-specific expression of green fluorescent protein in transgenic rainbow trout under control of the rainbow trout *vasa*-like gene promoter. Int J Dev Biol, 44: 323—326
- Yoshizaki G, Takeuchi Y, Kobayashi T *et al*, 2003. Primordial germ cell: a novel tool for fish bioengineering. Fish Physiology and Biochemistry, 28: 453—457

MOLECULAR CLONING AND CONSTRUCTION OF THE 5' FLANKING REGION EXPRESSION VECTOR OF THE *vasa* GENE OF RED SEABREAM (*PAGRUS MAJOR*)

LIN Fan^{1,2}, LIU Qing-Hua¹, LI Jun¹, SUI Juan^{1,2}, XIAO Zhi-Zhong¹,
XU Shi-Hong¹, MA Dao-Yuan¹, XIAO Yong-Shuang¹

(1. Institute of Oceanology, Chinese Academy of Science, Qingdao, 266071;

2. Graduate School, Chinese Academy of Sciences, Beijing, 100049)

Abstract Red seabream (*Pagrus major*) is a commercially important marine species in China. With the rapid development of marine aquaculture industry, resource conservation and identification of red seabream is increasingly important. Although sperm cryopreservation of the red seabream is relatively successful, little progress has made on embryo cryopreservation. Primordial germ cell (PGC) is the progenitor of the germ cell lineage, giving rise to either egg or sperm, with the potential to create complete individual organisms after fertilization. Cryopreservation of PGCs provides a new way for resource conservation. *Vasa*, a member of DEAD-box gene family, is strictly expressed in germ cell lineage. It is widely used as a molecular marker to study origin, migration and differentiation of PGCs. The objective of this study is to investigate potential cis-acting elements that involve in the transcription regulation of *vasa* gene in red seabream and set up a basis for further research on promoter efficiency and isolation of PGCs with GFP special expression. As the *vasa* 5'UTR reported (GenBank: AB378581) is not long enough to design three nested primers, we first amplified the first intron to get enough information. Next, 5' flanking region of red seabream *vasa* was amplified from genomic DNA by Genome Walking. Potential transcript factor binding sites were analyzed by bioinformatics method, the potential core promoter was blasted with the core promoter of zebrafish and finally *pvasa*-EGFP expression vector was constructed. Sequence analysis revealed that the 5' flanking region amplified in this study was 2762bp, and the region contained several potential transcription factor binding sites that may have important function on the gene transcription, such as TATA-box, CAAT-box, SP-1, GAGA-1, Oct-1, v-Myb, Sox-5, SRY and HNF. The potential core promoter of red seabream was similar to that of zebrafish, suggesting that it may play important role in the transcription regulation. Based on the above information, full length of the *vasa* 5' flanking region above the start codon was amplified and linked to EGFP report gene for the further research.

Key words *Pagrus major*, *vasa*, Promoter, PGCs