

中华绒螯蟹卵巢新基因 EJO5 的全长 cDNA 克隆和序列分析

马长艳¹, 周开亚¹, 赵乃刚², 汪朝晖²

(1. 南京师范大学生命科学学院遗传资源研究所 210097, 南京)

(2. 安徽省水产新技术研究所 230088, 合肥)

[摘要] 利用 RACE 技术从中华绒螯蟹卵巢获得了新基因 EJO5(*Eriocheir japonica* ovary gene 5, EJO5) 的全长 cDNA 序列(GenBank 检索号: AY185921)。该 cDNA 序列长度为 1 425 bp, 开放阅读框为 585 bp, 编码 194 个氨基酸。根据氨基酸序列计算的相对分子质量和等电点分别为 22 344 和 9.5。用生物信息学的方法未能得到对该基因功能预测的信息。

[关键词] 中华绒螯蟹, 卵巢, 新基因 EJO5, RACE, cDNA 克隆

[中图分类号] Q953, [文献标识码] A, [文章编号] 1001-4616(2004)01-0071-03

中华绒螯蟹(*Eriocheir japonica sinensis*)又名河蟹、毛蟹,是我国重要的经济蟹类。近年来,随着河蟹养殖业在全国各地的迅速发展,养殖过程中出现的性早熟、种质退化、个体偏小等问题日益严重。为了探讨河蟹性早熟及蜕皮的分子机理和种质资源保护等问题,我们实验室开展了一系列分子水平的研究工作^[1-4]。

作者在用 RACE 方法克隆中华绒螯蟹卵巢发育相关基因的过程中获得了 1 条含有 poly A 尾的非特异的 625 bp 的序列,blast 分析结果表明该序列为 1 条新的 EST(expressed sequence tag, EST),可能代表 1 条新的基因。本文用 5'RACE 的方法从中华绒螯蟹卵巢 RACE cDNA 文库^[2]中获得了此新基因的全长 cDNA 序列,并进一步用生物信息学的方法对该基因的结构和功能进行了初步分析。

1 材料和方法

1.1 材料

卵巢发育Ⅲ期的中华绒螯蟹,由安徽芜湖龙湖水产养殖场提供。

1.2 主要试剂

Taq DNA 聚合酶和 T/A 克隆试剂盒(pGEM-T Easy Vector System)(Promega),切胶纯化试剂盒和质粒提取试剂盒(上海华舜)。

1.3 引物设计和 5'RACE

1.3.1 引物设计

根据中华绒螯蟹卵巢差减 cDNA 文库^[1]中克隆 F386 的序列(EST002, GenBank 检索号: CA591893),在从 RACE cDNA 文库中扩增 EST002 代表的基因全长 cDNA 的过程中,获得了 1 条含有 poly A 尾的非特异的 625 bp 的序列。用 gene runner 软件根据该 625 bp 的序列设计了用于 5'RACE 的基因特异性引物(gene special primer, GSP) C₄-5: 5'-AGT CTT TTA GGG TTT CGG GTG G-3'。

1.3.2 5'RACE

用基因特异性引物 C₄-5 和 5'RACE 接头引物 5'-AAG CAG TGG TAT CAA CGC AGA GT-3'以 RACE cDNA 文库^[2]中的双链 cDNA 为模板(稀释 100 倍)进行 PCR 扩增。PCR 反应总体积为 30 μ L: 10 \times PCR buffer 3.0 μ L, dNTP Mix(2 mmol/L) 2.0 μ L, Mg²⁺(25 mmol/L) 1.8 μ L, Taq DNA 聚合酶 0.2 μ L, 模板 1.5 μ L, 引物(5 μ mol/L) 各 1.0 μ L, 灭菌水 19.5 μ L。PCR 反应程序如下: 95 $^{\circ}$ C 预变性 5 min 随之以 94 $^{\circ}$ C 50 s, 60 $^{\circ}$ C 30 s, 72 $^{\circ}$ C 2 min 进行 35 个循环, 72 $^{\circ}$ C 20 min。

收稿日期 2003-06-25。

基金项目 国家自然科学基金重点项目(No. 30130040)。

作者简介 马长艳,女,1970-,博士,南京医科大学讲师,从事动物分子生物学的学习与研究, E-mail: machangyanj@163.com

通讯联系人 周开亚,1932-,南京师范大学生命科学学院教授,博士生导师,从事动物学教学与研究, E-mail: kzyzhouj@jlonline.com

万方数据

1.3.3 5'RACE PCR 产物的纯化和克隆测序

5'RACE PCR 产物切胶纯化后 和 pGEM - T Easy Vector 连接 然后转化 JM109 感受态细胞.挑取单克隆接种于 1 mL Amp^r 液体 LB 培养基进行放大培养.菌落 PCR 鉴定阳性克隆后 随机取 2 个克隆 C₄ - 5(6)和 C₄ - 5(11)用质粒提取试剂盒提取质粒后在 ABI 310 遗传分析仪上进行双向测序.

1.3.4 序列的拼接

将以上 2 个克隆的插入序列分别和 625 bp 的序列进行拼接.

1.4 全长序列的生物信息学分析

用 ORF(open reading frame , ORF)finder 软件(<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>)寻找开放阅读框 blast(<http://www.ncbi.nlm.nih.gov/>) fast3(<http://www.ebi.ac.uk/fasta3>)和 Wu - blast(<http://www.ebi.ac.uk/blast2>)软件进行蛋白质的一致性和相似性分析 ;Compute pI - Mw 软件(http://cn.expasy.org/tools/pi_tool.html)进行蛋白质等电点和相对分子质量的计算 SignalP 软件(<http://www.cbs.dtu.dk/services/SignalP/>)进行信号肽的分析 SMART(Simple Modular Architecture Research Tool)(http://smart.embl-heidelberg.de/smart/show_motifs.pl)软件进行蛋白质结构域的分析.

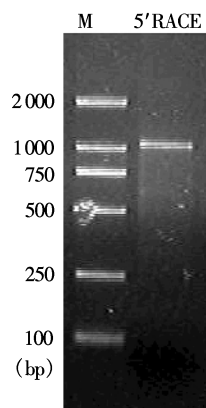
2 结果与讨论

2.1 5' RACE 和序列拼接

5'RACE 扩增得到 1 条约 1 000 bp 的主带(图 1).该主带切胶纯化后 T/A 克隆测序的结果表明,克隆 B₄ - 5(6)的插入片段长 997 bp 和 625 bp 的序列拼接获得成功,拼接处重叠 197 bp.拼接的序列长 1 425 bp,命名为中华绒螯蟹卵巢基因 5(*Eriocheir japonica* ovary gene 5, EJ05).EJ05 基因的 cDNA 序列及推导的氨基酸序列见图 2.1987 年,Kozak 通过对 699 种脊椎动物 mRNA 翻译起始密码 5'及 3'端两侧核苷酸序列的分析,揭示了 - 3 位的 A 和 + 4 位的 G 对于 AUG 被起始识别有最显著的促进作用^[5],即 Kozak 规则.EJ05 基因的 cDNA 序列虽然不符合明显的 Kozak 规则,但:①该基因在 72 ~ 74 位为第一个起始密码子 ATG,且第一个起始密码子上游 - 45 ~ - 43 位为同框终止密码子 TGA;②第一个起始密码子下游有同框终止密码子 TAA 和 polyA 尾.有资料表明,90% ~ 95% 的真核 mRNA 符合第一 ATG 规律^[6].根据以上两点,可判断我们获得的序列是 EJ05 基因的全长 cDNA 序列.

2.2 生物信息学分析

ORF finder 软件分析结果表明,EJ05 基因在 72 ~ 656 bp 之间有一个长度为 585 bp 的开放阅读框,编码 194 个氨基酸.用 Compute pI - Mw 软件根据氨基酸序列计算的等电点和相对分子质量分别为 9.50 和 22 344. blastp, fast3 和 Wu - blast 软件在 Swissprot 数据库中皆未搜索到有意义的匹配序列,由此判断 EJ05 基因为新基因,该基因的 GenBank 检索号为 AY185921. SignalP 软件分析结果表明该基因编码的蛋白无信号肽. SMART 软件进行蛋白质结构域的分析结果表明未发现 EJ05 基因编码蛋白与已知蛋白相似的结构域.通常克隆到 1 条新基因时,首要的工作应是进行生物信息学分析,因为生物信息学分析的结果尤其是得到蛋白的结构功能域的信息后,将对研究思路的制定提供重要的指导^[7].本研究中对 EJ05 基因的生物信息学分析并未得到关于该基因功能预测的有价值的信息,这就给该基因的功能研究带来了一定的难度,但可尝试通过基因敲除(gene knockout)反义 RNA(antisense RNA)或 RNA 干扰(RNA interference , RNAi)等技术封闭该基因的表达,观察个体表型以确定其功能.



(M :DNA 标记)

图 1 EJ05 基因的 5'RACE 产物

1 gagtgatcatccagaagaggttagagtgaggttactcgcgtgacatcgacaacagagtccc
61 aaaagtcgtgcatgttcggcgacaccacggcggcgtacgtcgagggtagaccatgcacgt
M F G D T T V A Y V E G R P C T 16
121 gccccgagaagtgtatgtatcttgggtgtgccaggaaggacgatcgtggcgaagactaca
C P E K C M Y L G V P G R D D R G E D Y 36
181 ggacgagatcaacgacactgatatgtctgtttggacgagatactatcgtaacccccctc
R T R S T T L I C L F G R R Y Y R T P P 56
241 gttctcgcgacccccgacactatcataccctatgacttgcaccaacggctcctgggatcac
R S R D P D T I I P Y D L H Q R L L G S 76
301 ggagaacttcacgcacccctgttgtggccacatggtgaaatctacgggtacagtcaactt
R R T S S H P C C G H M V K S T G T V T 96
361 cagcacttggaacgtattcgtctacgactggcagcgttctgcaattacacccctcgtct
S R H W N A I R Y D W H G F C N Y T L V 116
421 cgtcgcagccacggacagggcaccctccgcgcagctcttcgcagagatgaagcgcgtgg
S L D A T D R A P S A A V F A E M K R V 136
481 gtccctcgcgccacgtcttcgacagtgtcaccttcacacgacccctcacaccataatca
G P C A T V F D S V T F H N D P H T I I 156
541 acttgaccgcgcagccaacaacaaaagtgcagcaacagatcggtatcaacggagagggcc
N L T R D A N N K S Q R T D R Y Q R R G 176
601 atgaccgttccccgcaccggcgcccaacttggtcaattcctctgcggcagacataacgtg
H D R S P H R R P L G Q F L C R Q T * 194
661 ctagtgttctggcagaaggactgtctcactcactgagccggggtcatatcacgctccgg
721 tattgccccctgcacatggatatctggatcccccgccacctcacccggcgtgtggacggt
781 atctgcggtcacttcaactacaacaccacggatgacttctccaccgccacgagggcagatc
841 ctccccctggagccattccccgtcaacttccctgacatgcattggcgggctcagaccag
901 cgcacgcgcccttgcaatgacccaatgatcaccacctccatctgtaacgggtaccacgggt
961 gacaagtgcctccttccacccgaaaccctaaaagacttccaacgcgcctgcatggacttc
1021 ctactcagctgttccccagcagccaggttgcgactactaccttccccctgcgtg
1081 gccgacctctgcttgggtgtcccaacgtgatgtgaacttcacggctccttccccgccttc
1141 gtcacgcaggtgtcctcaaaatgtacaggaacatcacagtgcagcattacgaagaccgcagg
1201 gtcacccagctattaatggacccaagtatgtcttcgctatggacgacaccactaccaaca
1261 tcccgctgcgtacggacggcacaccaccaccagacccccgtgcagaggacagcgctc
1321 acaccacccctatcgtcgtttaatacatacacttgtaactctcaaaacatctgtacact
1381 aacaaattgagaagctaaaaaaaaaaaaaaaaaaaaaaaaaaaaa

图2 EJ05 基因的核苷酸及推导的氨基酸序列(下划线表示终止密码子)

[参考文献]

[1] 马长艳,周开亚,郭豫杰,等. 中华绒螯蟹卵巢差减 cDNA 文库的构建[J]. 动物学研究, 2003, 24(1):53—56.
[2] 马长艳,周开亚,郭豫杰,等. 中华绒螯蟹卵巢 RACE cDNA 文库的构建[J]. 动物学杂志, 2003, 38(2):13—15.
[3] 宋霞,周开亚,马长艳. 中华绒螯蟹蜕皮抑制激素 1(Eri - MIH 1)基因的 cDNA 片段克隆和 Northern 印迹分析[J]. 中国水产科学, 2003, 20(5):353—358.
[4] 孙红英,周开亚,杨光,等. 中国大陆绒螯蟹线粒体 16S rDNA 序列变异与分子鉴定标记[J]. 自然科学进展, 2002, 12(5):485—490.
[5] Kozak M. The scanning model for translation : an update[J]. J Cell Biol , 1989 , 108 : 229.
[6] 沈翔珩,方福德. 真核基因表达调控[M]. 北京: 高等教育出版社, 1999. 63—64.
[7] 张成岗,贺福初. 生物信息学方法与实践[M]. 北京: 科学出版社, 2002. 110—111.

(下转第77页)

(group C), lactobacillus rich in selenium group(group Se), CCl₄ loading group(group CCl₄) and CCl₄ plus lactobacillus rich in selenium group(group CCl₄ - Se). Liver injury was induced by abdominal injection of CCl₄ on every other day over six weeks. The blood samples were collected at the 2nd ,4th and 6th weeks , respectively , and the enzyme activities of glutathione superoxidase(GSH - Px) , glutathione S - transferase(GST) , superoxide dismutase(SOD) , the content of malondialdehyde (MDA) in RBC haemolytic liquid were determined. Effect of lactobacillus rich in selenium on lipid peroxidation damage of red blood cell(RBC) in mice with CCl₄ - loading liver injury were studied. The results showed that during the whole experiment period , there were no obvious differences among the groups for SOD activity ; GST activity in group Se tended to increase with the highest values , and rose or did significantly as compared with other three groups at the 6th week , GST in group CCl₄ was low or significantly low to groups C and CCl₄ - Se , there were similar values between groups C and CCl₄ - Se ; GSH - Px activity in groups C , Se and CCl₄ - Se had no significant differences between each other , but group CCl₄ tended to decrease with time of CCl₄ - loading , and was low to other three groups after the 4th week. MDA content had no obvious differences between groups C and CCl₄ - Se , the entire level of MDA in group Se was lower or significantly lower than that of other three groups , group CCl₄ had opposite changes to group Se. It was suggested that the violent changes of lipid peroxidation of RBC in mice with CCl₄ - induced liver injury were confirmed ; beneficial action of lactobacillus rich in selenium was mediated through the protection of some antioxidant enzyme activity in RBC.

Key words :Lactobacillus rich in selenium , Liver injury , Red blood cell , Lipid peroxidation

[责任编辑 孙德泉]

(上接第73页)

Molecular Cloning and Sequence Analysis of the Full-length cDNA of a Novel Gene EJO5 From the Ovary of *Eriocheir japonica sinensis*

Ma Changyan¹ , Zhou Kaiya¹ , Zhao Naigang² , Wang Zhaohui²

(1. Institute of Genetic Resources School of Life Sciences ,Nanjing Normal University , 210097 , Nanjing , PRC)

(2. Anhui Fisheries New Technology Research Institute , 230088 , Hefei , PRC)

Abstract :A full-length cDNA sequence of a novel gene EJO5 (*Eriocheir japonica* ovary gene 5 , EJO5)(GenBank Accession Number : AY185921) from the ovary of *Eriocheir japonica sinensis* was obtained by RACE. The full-length cDNA sequence of EJO5 is 1 425 bp in size and has an open reading frame of 585 bp encoding a polypeptide of 194 amino acids. The pI/Mw deduced from the amino acids sequence is 9.5 and 22 344 respectively. However , we were unable to predict the function of EJO5 gene by means of bioinformatics.

Key words :*Eriocheir japonica sinensis* , ovary , novel gene EJO5 , RACE , cDNA cloning

[责任编辑 孙德泉]