牛 β-乳球蛋白基因 5′调控成分的分离、克隆及序列分析*

石玉强, 张 涌, 郑月茂, 刘建忠, 张志平
(西北农林科技大学 生物工程研究所, 陕西 杨陵 712100)

[摘 要] 利用 PCR 方法克隆了牛 β-乳球蛋白基因 5′调控成分(1 449 bp), 其中包括 5′侧翼序列(645 bp), 第一外显子及第一内含子(804 bp)。PCR 产物回收纯化后, 克隆在 pMD 18-T Vector 的 T 位点。序列分析表明, 该序列与牛 β-乳球蛋白 A 型基因、牛 β-乳球蛋白 B 型基因、山羊和绵羊 β-乳球蛋白基因的同源性分别为 98.55%, 99.79%, 90.70% 和 90.09%。计算机分析表明, 此调控成分含有诸多调节因子结合位点或反应元件, 其中包括视黄酸反应元件(RARE)、佛波酯反应元件(TRE)、乳腺特异性因子(MSBF)和核因子 1(NF1)结合位点等, 提示此调控成分可调控外源基因在乳腺细胞中高效表达, 可用于构建乳腺特异性表达载体。

[关键词] β-乳球蛋白基因; 牛; PCR; 序列分析
[中图分类号] Q 823; Q 785 [文献标识码] A [文章编号] 1671-9387(2003)03-0001-04

利用乳腺生物反应器能大规模地生产人类所急需的外源活性蛋白, 该法因具有独特的优点而引起人们的广泛关注。制备乳腺生物反应器最关键的环节是构建高效、特异的表达载体。β-乳球蛋白(β-lactoglobulin, BLG)是反刍动物最主要的乳清蛋白, 大量研究证明, 它的调控成分能指导外源基因在乳腺细胞中特异、高效表达。杨国庆等[1]以牛 BLG 基因 5′调控序列作为启动子与人生长激素基因(hGH)构建表达载体, 在转基因小鼠乳清中 hGH 的表达量为 420 mg/L。陈红星等[2]以牛 BLG 基因 5′调控序列构建表达载体, 获得组织型纤溶酶原激活剂(t-PA)转基因鼠, t-PA 在乳腺中也有较高的表达量。最引人瞩目的是 Wright[3]以绵羊 BLG 基因的 5′调控序列来表达 α1-抗胰蛋白酶(α1-AT), α1-AT 在转基因绵羊乳中表达量高达 35 g/L。鉴于此, 本试验拟以牛 BLG 基因 5′调控成分构建表达载体, 考虑到第一外显子和第一内含子在指导外源基因正确翻译、表达和分泌等方面中的作用, 也克隆了此序列。

1 材料与方法

1.1 材 料

试验动物 美国荷斯坦奶牛由杨凌科元生物工程有限公司提供。

质粒和菌种 宿主菌 DH5α 为本室保存, 质粒 pMD 18-T Vector 购自大连宝生物公司。

试剂和仪器 限制性内切酶、T4 连接酶和 Marker 均购自加拿大 MBI 公司, Taq DNA 聚合酶、10×PCR Buffer 和 dNTP Mixture 购自大连宝生物公司, DNA Gel Extraction Kit 和 Plasmid Miniprep s Kit 购自上海生物工程公司。PCR 仪为 Gene AmpR PCR-System 2400(美国 PE 公司), 凝胶成像分析系统为英国 UVP GDS8000 型, 紫外分光光度计为 Cintra(GBC)型。

1.2 方 法

奶牛基因组 DNA 的提取 奶牛颈静脉采血, 0.5 g/L EDTA 抗凝, RNase 酶、蛋白酶 K 消化后, 利用酚-氯仿法提取基因组 DNA, 真空干燥后溶于 300 μL 灭菌超纯水中, 并用紫外分光光度计测定 DNA 样品在 260, 280 nm 处的 OD 值, 计算 DNA 样品的浓度和纯度, 具体方法见参考文献[4]。

特异性引物设计 参考 Alexander 等[5]发表在 GenBank 中的奶牛 BLG 全基因序列, 利用引物设计软件 Prime Premier 5.0 辅助设计。为方便构建表达载体, 在上、下游引物的 5′端各加上了 1 个 *Eco*RⅠ限制性酶切位点和两个保护碱基, 引物序列如下:

上游引物 5′gcg, aat, tcA, TCC, CAC, GTG,

* [收稿日期] 2002-06-07
[基金项目] 国家"863"高技术资助项目(2001AA213081)
[作者简介] 石玉强(1978-), 男, 山东滕州人, 在读硕士, 主要从事胚胎工程和分子生物学研究。
[通讯作者] 张 涌(1956-), 男, 内蒙古和林格尔人, 教授, 博士生导师, 主要从事胚胎学与胚胎工程研究。

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

CCT, GC 3′

下游引物 5′gag, aat, tcC, TGG, GGA, GGG, ACC, TT 3′

PCR 扩增 取DNA 样品 1 μL (100 ng), Mg^{2+} 4 μL (25 mmol/L), dNTP Mixture 4 μL (各 2.5 mmol/L), 引物各 2 μL (25 μmol/L), 超纯水 33 μL, 10×PCR Buffer 5 μL, Taq 聚合酶 1 μL (5 U/μL)。瞬间离心混匀, 首次 96 ℃变性 5 min 后, 94 ℃变性 30 s, 61 ℃退火 45 s, 72 ℃延伸 90 s, 进行 35 轮循环, 最后 72 ℃延伸 10 min。

琼脂糖电泳 制备 0.7 g/L 的琼脂糖凝胶, 加入溴化乙锭(EB) (20 mg/L)至 1 mg/L。取 PCR 扩增产物 5 μL 与 6×凝胶电泳上样缓冲液 1 μL 混匀后, 点样电泳, 电压 5 V/cm。45 min 后, 在凝胶成像分析系统中观察结果并记录。

克隆 利用DNA Gel Extraction Kit 从琼脂糖凝胶上回收并纯化 PCR 产物后, 克隆在 pMD 18-T Vector 的 T 位点, 并转化感受态大肠杆菌, 蓝白筛选阳性菌落, 具体方法见参考文献[4]。

目的片段的鉴定和序列测定 利用 Plasmid Miniprep s Kit 提取质粒后, *Eco*R Ⅰ酶切鉴定重组质粒, 最后送交大连宝生物公司完成序列测定。

2 结 果

2.1 牛基因组DNA 的完整性

经紫外分光光度计测定, 牛基因组DNA 样品 OD 值为 1.72, 取 2 μL 进行 0.7 g/L 琼脂糖凝胶电泳, 溴化乙锭染色, 紫外灯下观察基因组DNA 无降解。

2.2 PCR 扩增结果

利用提取的基因组DNA 和特异性引物进行扩增, 取 5 μL PCR 扩增产物电泳, 发现与预期的扩增产物大小一致, 结果见图 1。

2.3 重组质粒酶切鉴定结果

利用蓝白筛选法筛选阳性重组菌落, 提取质粒后, *Eco*R Ⅰ酶切(电泳结果见图 2)鉴定, 目的片段已成功地克隆在pMD 18-T Vector 的 T 位点。

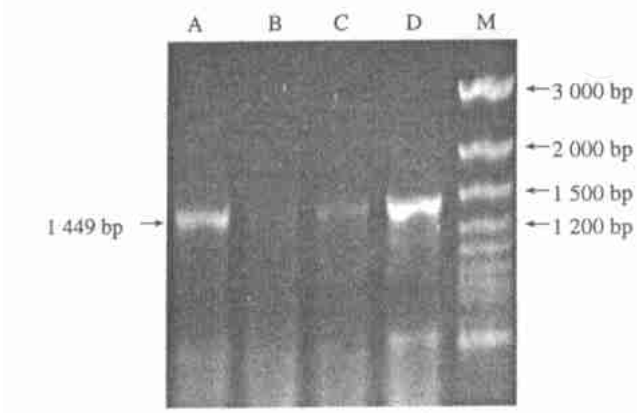

图 1 牛BLG 基因 5′调控成分 PCR 扩增电泳

A, B, C, D. PCR 产物; M. DNA marker

Fig. 1 PCR product of bovine BLG gene 5′regulatory element

A, B, C, D. PCR product; M. DNA marker

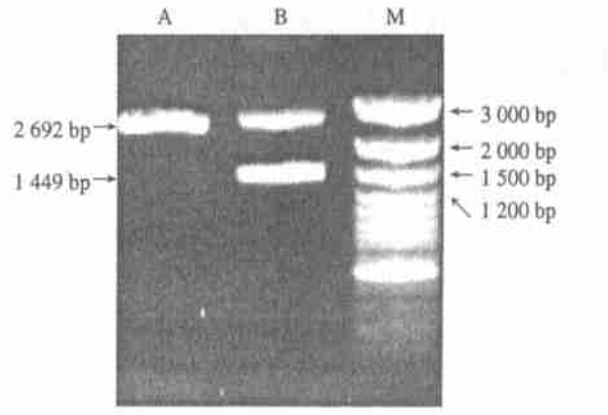

图 2 重组质粒 *Eco*R Ⅰ酶切电泳

A. pMD 18-T vector; B. 重组质粒酶切结果; M. DNA marker

Fig. 2 The recombinant plasmid digested by *Eco*R Ⅰ

A. pMD 18-T vector; B. Recombinant plasmid; M. DNA marker

2.4 序列分析结果

目的片段测序后, 利用BLAST 软件进行同源性分析, 结果见表 1。同时利用DNA Star 软件分析, 发现此调控成分含有多个调节因子作用位点或反应元件, 其中包括视黄酸反应元件(RARE)、佛波酯反应元件(TRE)、乳腺特异性因子(MSBF)和核因子(NF1)结合位点等, 见图 3。

表 1 牛BLG 基因 5 调控成分序列的比较

Table 1 The comparison of bovine BLG gene 5′regulatory element with other mammals

哺乳动物 Mammals	基因 Gene	同源性/% Homology
奶牛 Cow	β乳球蛋白A 型基因 β-lactoglobulin gene variant A	98.55
奶牛 Cow	β乳球蛋白B 型基因 β-lactoglobulin gene variant B	99.79
山羊 Goat	β乳球蛋白基因 β-lactoglobulin gene	90.70
绵羊 Sheep	β乳球蛋白基因 β-lactoglobulin gene	90.09

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

```
1    atcccacgtg cctgcacgga gccctagtgc tactca[TRE]gtca tgcccccgca gcaggggtc[CK]a
61   ggtca[reti]ctttc ccatcctggg ggttattatg actgttgtca ttgttgttgc cattttg[Sm]ct
121  accctaactg ggcagcgggt gcttgcagag ccctcgatac tgaccaggtt ccccccctcgg
181  agctcgacct gaaccccatg tcaccc[CAC]tcgc cccagcctgc agagggtggg gtgactgcag
241  agatcccttt acccaaggcc acagtcacat ggtttggagg agatggtgcc caaggcagaa
301  gccaccc[CAC]tcc aggacacacc tgcccccagt gctggctctg acctgtcctt gtctaagagg
361  ctgaccccag aagtgttcct ggcgctggca gccagcctgg acccagagcc tggacaccc[CAC]c
421  ctgcgccccc acttctgggg cg[GC box]taccagga accgtccagg cccagagggg gccttcctgc
481  ttggcctcga atggaagaag gcctcctatt gtcctcgtag aggaa[PU-1]gcaac cccagggccc
541  aaggataggc caggggggat tcggggaacc[MSBF] gcgtggctgg gggcccggcc cgggctggct
601  ggctggccct cctcctgtat aa[TATA box]ggccccga gcccgctgtc tcagccctcc actccctgca
661  gagctcagaa gcgtgacccc agctgcagcc atgaagtgcc tcctgcttgc cct[start]ggccctc
721  acttgtggcg cccaggccct cattgtcacc cagaccatg[Code]a agggcctgga tatccagaag
781  gttcgagggt gcccgggtgg gtggtgagtt gcagggcagg caggggagct gggcctcaga
841  gaccaaggga ggctgtgacg tctgggattc ccatcagtca gctagagccg cctgacaaat
901  cgcccgccag ggcagcttca accaggcgtt tagtgtcttg cattctggag gctggaagcc
961  tgcaatcc[CAAT box]gg gcatcggccc agctggcttc tcctgcggcc actctccggg gagcagacag
1021 ccatcttctc cctgtgtcct ttgcgtgccc tggtttcctc ttcctgtgag gtca[reti]ccaggc
1081 ctgctggatc cacgcccgcc cacacagcct cacgtaacct ttgtcatctc tttaaaggcc
1141 gtgtctccag tcctgtgttg aggttctggg ggtt aatggg ac[glyco]acagttca gcccctaaaa
1201 gagtcactct gcccctcaaa ttttccccac ctccagctat gtctccccaa gatccaaatg
1261 ttgccacgtg tgcgggggct catctgggtc cctctttggg ctcagagtga gtctggggag
1321 agcattcctc agggtgccga gttgggggga gcatctcagg gctgcccagg ccagggtggg
1381 acagagagcc cactgtgggg ctgggggccc cttcccgccc ctggagtgca gctcaaggtc
1441 cctccccag
```

图 3　牛BLG 基因 5 调控成分的DNA 序列及分析

CAAT box. CCAAT 框、核因子-1 结合位点; CAC. GT-1c 蛋白结合位点; CK. 细胞因子 CK-2e 结合位点; Code. 翻译起始位点; GC box. GC 框、SP1 因子结合位点; glyco. 糖基化位点; MSBF. 乳腺特异性因子结合位点; Pu-1. Pu-1 因子结合位点; reti. 视黄酸反应元件(RARE); Sm. 核糖核蛋白 Sm 结合位点; Start. 转率起始位点; TATA box. TATA 框、TBF 因子结合位点; TRE. 佛波酯反应元件、AP1 结合位点

Fig. 3　The sequence and analysis of bovine BLG gene 5′regulatory element

CAAT box. CAAT signal of RNA polymerase Ⅱ. NF-1 factor binding site; CAC. GT-1c protein binding site; CK. cytokine-2e binding site; Code. start code of translation; GC box. GC signal of RNA polymerase Ⅱ. SP-1 factor binding site; glyco. glycosylation site; MSBF. mammary specific binding factor site; Pu-1. Pu-1 factor binding site; reti. retinoic acid response element (RARE); Sm. small nuclear ribonucleoprotein Sm binding site; Start. start code of transcription; TATA box. RNA polymerase binding site. TATA-box binding protein site; TRE. TPA response element. AP-1 factor binding site

3　讨　论

利用乳腺生物反应器制备转基因动物, 是当前生物技术领域的研究热点之一, 但是外源基因在乳腺细胞中特异性表达受诸多因素的影响。①内含子: 目前已经有充分的研究[6]证实, 内含子(特别是第一内含子) 对外源基因正确的剪接、修饰以及提高表达效率有重要的作用。这可能是由于内含子区包含增强子序列以及基质结合区(MAR)、内含子可开放染色体区域, 或通过内含子的剪接可增加mRNA 在细胞质的稳定性, 在本研究中克隆牛BLG 基因第一内含子序列正是基于此。②位置效应: 表达构件在宿主染色体上的整合位点强烈影响外源基因的表达, 若外源基因整合在宿主异染色质区, 它的表达便会被强烈抑制。然而有些表达构件似乎不存在整合效应, 表达水平直接与其在染色体上整合的拷贝数有关。

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

研究证明[7], 乳腺细胞中BLG 基因启动子区存在DNase Ⅰ超敏感位点, 此位点是诸多反式作用因子识别和作用的位点, 与基因的表达息息相关。

本研究通过对牛BLG 基因启动子区的克隆及序列分析, 发现此调控成分存在视黄酸反应元件(RARE)、佛波酯反应元件(TRE)、乳腺特异性因子(MSBF)和核因子(NF1)等诸多调节因子结合位点, 这对调节外源基因在乳腺上皮细胞特异性的表达至关重要。同时, 本研究之所以克隆牛BLG 基因第一外显子, 主要考虑第一外显子中的信号肽序列在指导外源基因特异性表达和分泌中是重要的。但是, 究竟此调控成分的活性如何, 还需要制作转基因动物来验证, 这也是作者目前正在从事的研究。

[参考文献]

[1] 杨国庆, 戴蕴平, 朱宝利, 等. 牛BLG 基因5′调控成分的克隆和制备乳腺生物反应器研究[J]. 中国科学(C 辑), 1996, 26(5): 463-469.

[2] 陈红星, 程 萱, 杨 晓, 等. 牛β乳球蛋白基因调控序列指导组织纤溶酶原激活剂在小鼠乳腺中的表达[J]. 生物工程学报, 2001, 17(2): 135-139.

[3] Wright G. High level expression of active human α1antitrypsin in the milk of transgenic sheep[J]. Biotechnology, 1991, 9: 830-834.

[4] Sambrook J, Fritsch E F, Manatis T, et al. Molecular cloning a laboratory mammal[M]. 2nd ed. New York: Cold Spring Harbor Laboratory Press, 1988.

[5] Alexander L J, Hayes G, Bawden W, et al. Complete nucleotide sequence of the bovine β-lactoglobulin gene[J]. Animal Biotechnology, 1993, 4(1): 1-10.

[6] Hyttinen J M, Korhonen V P, Hiltunen W D, et al. High level expression of bovine beta-lactoglobulin gene in transgenic mice[J]. J Biotechnol, 1998, 61(3): 191-19

[7] Kim S J, Sohn B H, Jeong S, et al. High level expression of human lactoferrin in milk of transgenic mice using genomic lactoferrin sequence [J]. J Biochem, 1999, 126(2): 320-325.

The isolation, cloning and sequence analysis of bovine β-lactoglobulin gene 5′ regulatory elements

SHI Yu-qiang, ZHANG Yong, ZHENG Yue-mao, LIU Jian-zhong, ZHANG Zhi-ping

(*Institue of Bio-Engineering, Northwest Sci-Tech University of Agriculture and Forestry, Yangling, Shannxi* 712100, *China*)

Abstract: For the purpose of constructing mammary gland specific expressional vector, the bovine β-lactoglobulin gene 5′flanking fragment was cloned by PCR amplification. It consisted in part of the gene upstream region about 645 bp, the first exon and the first intron about 804 bp. After PCR product was recovered and purified, the aim fragment was cloned at T site of pMD 18-T Vector. The fragment was sequenced, and it was compared with bovine BLG gene variant A, bovine BLG gene variant B, goat and sheep BLG gene. The results indicated the homology was 98.55%, 99.79%, 90.70% and 90.09% respectively. The nucleotide acid sequence was analyzed by computer, and found many factor or protein binding sites and response elements such as mammary specific binding factor (MSBF), nuclear factor-1 (NF-1), retinoic acid response element (RARE) and TPA response element (TRE). It suggested the cloned regulatory element should direct extra genes to express specifically in mammary gland epithelia cells, and it could be used as constructing mammary gland-specific expressional vector.

Key words: β-lactoglobulin gene; bovine; PCR; sequence analysis

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net