

奶牛 β -酪蛋白基因 5 和 3' 调控区的克隆及序列分析*

刘金龙, 郑月茂, 王玉洁, 张 涌

(西北农林科技大学 生物工程研究所, 陕西 杨凌 712100)

[摘 要] 利用普通及重组 PCR 技术克隆了奶牛 β -酪蛋白基因(CSN 2) 5' 调控成分(2 826 bp)和 3' 调控成分(620 bp)。前者包括 5' 上游调控序列、第一外显子及部分第一内含子; 后者主要包括最后一个不翻译的外显子和 3' 侧翼序列。纯化 PCR 产物通过 pMD 18-T Vector 亚克隆后测序并进行软件分析的结果表明, 2 个克隆片断与奶牛 CSN 2 相应区域同源性分别为 99.0% 和 98.0%, 并且包含有全部的 CSN 2 表达核心调控序列和多个转录、翻译因子的结合位点。

[关键词] 奶牛; CSN 2; 调控区; PCR; 序列分析

[中图分类号] Q 785

[文献标识码] A

[文章编号] 1671-9387(2004)02-0099-05

自 Lovell- Badge 提出用转基因动物乳腺生产重组蛋白的思路以来, 该技术得到了迅猛的发展, 现已成为生物技术领域发展的重要方向。乳腺生物反应器具有产品活性好、产量高、易于纯化和成本低的特点, 利用其大规模生产具有高附加值的蛋白质可获得良好的经济效益。制备乳腺生物反应器的关键前提是获得能够指导靶基因在乳腺中高效、特异性表达的调控成分。 β -酪蛋白(β -casein, CSN 2)是牛乳中含量最高的乳蛋白, 在泌乳激素的刺激下, CSN 2 具有很强的表达活性, 其调控成分能够指导靶基因在乳腺组织中高效表达^[1~3]。Ninomiyama 等^[4]采用 CSN 2 5' 1.7 kb 的调控区指导人生长激素基因(hCG)在转基因大鼠乳腺表达, 得到了 10.9 g/L 的高水平表达。黄赞等^[5]利用山羊 CSN 2 5' 6.7 kb 的调控区指导人凝血因子 IX(hFIX)在转基因小鼠乳腺中也得到最高为 52.9 g/L 的表达量; 张克忠等^[6]研究证实, 牛 CSN 2 5' 端 2.0 kb 的片断即能有效驱动外源基因在乳腺组织中表达。基于此, 并考虑到将要构建表达载体(质粒)的承载能力, 本研究分离克隆了牛 CSN 2 5' 长约 2.8 kb 的调控成分。另外, 由于乳蛋白内源的 3' 非翻译区, 尤其是 polyA 信号对于 mRNA 的加工、转运及稳定性有重要作用, 所以本研究还分离克隆了 CSN 2 3' 长约 620 bp 的下游

调控区, 以保证内源性 CSN 2 表达调控模式的完整性。

1 材料与方法

1.1 材料

1.1.1 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。

1.1.2 质粒和菌株 质粒 pMD 18-T Vector 购自大连宝生物公司, 宿主菌 DH5 α 为本室保存。

1.1.3 酶和试剂 Ex Taq DNA 聚合酶、Pyrobest DNA 聚合酶购自大连宝生物公司, 限制性内切酶和 Marker 购自加拿大 MBI 公司, DNA Gel Extraction Kit 和 Plasmid Miniprep Kit 购自上海生物工程公司。克隆引物由上海生物工程公司合成。

1.2 方法

1.2.1 奶牛基因组 DNA 的提取 奶牛颈静脉采血, 0.5 g/L EDTA 抗凝后按照文献^[7]的方法提取基因组 DNA。

1.2.2 CSN 2 5' 调控区的 PCR 扩增及产物纯化 由于预扩增的 CSN 2 5' 调控区较长, 因此先将之分为两段分别从基因组 DNA 中扩增, 然后采用重组 PCR 方法连接成完整的目的片段。根据 Genbank 中登陆的奶牛 CSN 2 全基因序列(登陆号: X14711),

* [收稿日期] 2003-08-28

[基金项目] 国家“863”高技术项目(2001AA 213081)

[作者简介] 刘金龙(1979-), 男, 山东潍坊人, 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-), 内蒙古和林格人, 教授, 博士生导师, 主要从事哺乳动物胚胎工程与发育生物学研究。

利用生物软件 Vector NTI Suite 7.0 和 Primer Premier 5.0 辅助设计了 2 对引物, 为了方便将来构建表达载体, 在第一条上游引物 5' 端引入 1 个 *Bcl*I 酶切位点, 引物序列如下:

BCS1 5' TGATCATGAGAAA GGGAAATGTTGAATGGG 3'
BCA1 5' GCTTGCCTTTATGGCTGAGGTATGT 3'
BCS2 5' GCTCCTCCTTCACTTCTTGTCTCTAC 3'
BCA2 5' ATGCCTAATGTTATCTCCCTTTGTCC 3'

首先用 BCS1 和 BCA1 从奶牛基因组 DNA 中扩增出片段 BC1, 同法用 BCS2 和 BCA2 扩增出片段 BC2, 最后将 BC1 和 BC2 用重组 PCR 方法连接成完整的 CSN 2.5 调控区片段 BBC5。PCR 反应体系均按 Ex Taq DNA 和 Pyrobest DNA 聚合酶说明书配置。BC1 和 BC2 循环条件是: 94 °C 变性 4 min 后, 94 °C 变性 30 s, 57 °C 退火 45 s, 72 °C 延伸 90 s, 进行 30 个循环, 最后 72 °C 延伸 7 min。重组 PCR 循环条件为: 94 °C 变性 3 min 后, 94 °C 变性 30 s, 58 °C 退火 45 s, 72 °C 延伸 3 min, 进行 35 个循环, 最后 72 °C 延伸 10 min。

重组 PCR 产物 BBC5 经低熔点胶电泳后观察结果并记录, 然后用 DNA Gel Extraction Kit 回收纯化后用于下一步克隆和测序。

1.2.3 CSN 2.3 调控区的 PCR 扩增及产物纯化 如 1.2.2 所述, 设计 2 条引物从奶牛基因组 DNA 中直接扩增, 并在下游引物 5' 端引入 *Mlu*I 酶切位

点, 引物序列如下:

BC3S 5' AATGATTCCAA GTAA GCCGATG 3'
BC3A 5' ACGCGTTCCAA TTTTAATTTTCCACAGC 3'

PCR 反应体系亦如 1.2.2 所述配置。循环条件为: 94 °C 变性 4 min 后, 94 °C 变性 30 s, 50 °C 退火 45 s, 72 °C 延伸 45 s, 进行 30 个循环, 最后 72 °C 延伸 5 min。

PCR 产物 BBC3 如 1.2.2 电泳, 并回收纯化以用于克隆和测序。

1.2.4 CSN 2.5 和 3 调控区克隆 将回收纯化的 BBC5 和 BBC3 片段用 T-A 克隆法克隆在 pMD 18-T Vector 的 T 位点, 转化 DH5 α 感受态细胞, 蓝白斑筛选阳性菌落, 具体方法见文献[7]。

1.2.5 重组质粒的筛选和序列测定 用 Plasmid Miniprep Kit 提取质粒后, 酶切鉴定重组质粒 pBBC5 和 pBBC3, 最后送交华大上海鼎安生物公司完成序列测定。

2 结 果

2.1 CSN 2.5 和 3 调控区的 PCR 扩增结果

经 PCR 扩增, 获得了 2 826 bp 的 5 调控区和 620 bp 的 3 调控区, 与预期的扩增产物大小一致, 电泳结果见图 1。

2.2 重组质粒酶切鉴定结果

琼脂糖凝胶电泳结果见图 2。

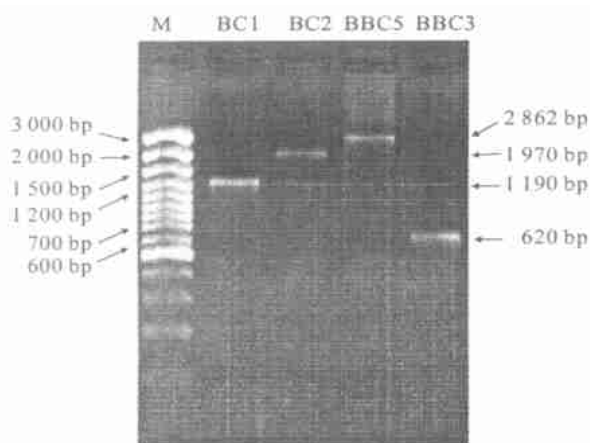


图 1 5 和 3 调控区 PCR 扩增电泳结果

BC1, BC2 分段 PCR 产物; BBC5 5 调控区片段;
BBC3 3 调控区片段; M: DNA marker

Fig 1 PCR products of 5' / 3' control region

BC1, BC2 segmental PCR product; BBC5 5' control region;
BBC3 3' control region; M: DNA marker

蓝白斑筛选阳性菌落并提取质粒后, 由于 BBC5 和 T-Vector 的 MCS (多克隆位点) 区均存在

一个 *Eco*R I 位点, 因此用 *Eco*R I 酶切鉴定 pBBC5; 与此相同, 用 *Kpn*I 酶切鉴定 pBBC3。结果前

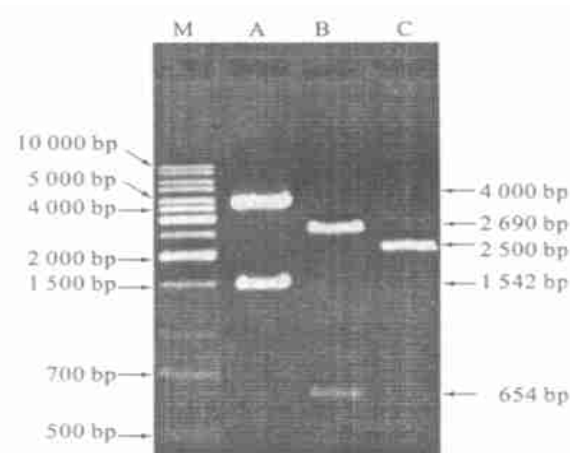


图 2 重组质粒酶切电泳结果

A. pBBC5 *Eco*R I 酶切结果; B. pBBC3 *Kpn*I
酶切结果; C. T-V vector; M: DNA marker

Fig 2 Result of digested recombinant plasmid

A. pBBC5 digested by *Eco*R I; B. pBBC3 digested by *Kpn*I;
C. T-V vector; M: DNA marker

者酶切出了 1 542 bp 和 4 000 bp 的 2 个片断, 后者酶切出了 645 bp 和 2 690 bp 的 2 个片断, 这证明 2 个目的片段已经成功克隆在 T-V ector 的 T 位点, 并且均为正向插入。

2 3 序列分析结果

扩增片断测序后, 经 NCB I 网站数据库 BLA ST 中的 blastn 进行同源性分析, 片断 BBC5 和 BBC3 与奶牛 (Cow) CSN 2 全基因对应区段的同源性分别

为 99 0% 和 98 0%, 证明克隆的片断确为牛 CSN 2 基因的调控区。同时, BLA ST 结果还显示, 克隆片断与山羊 (Goat) 和绵羊 (Sheep) 的 CSN 2 的同源性分别为 90 0%, 89 0%, 94 0% 和 95 0%。利用 DNA Star 软件和启动子区分析软件 AlBaba2 分析 BBC5 和 BBC3 片断, 发现这 2 个调控成分含有多个复合反应元件和转录、翻译调节因子的作用位点, 具体见表 1。

表 1 牛 CSN 2 5 和 3 调控成分序列分析

Table 1 The sequence analysis of 5' and 3' regulatory region of bovine CSN 2

结合元件 Elements	(结合) 位置 (Binding) Position	序列特征 Sequence feature (5'→3')	类别 Class		
			5' 远端 5' dist	5' 近端 5' prox	3' 不翻译区 3' UTR
STAT5(MGF)	- 1 574~ - 1 565	TTCTCAGAAAT	✓		
	- 102~ - 91	ATTTCTAGGAAT		✓	
C/EBP	- 1 612~ - 1 604	TGTTGAATG	✓		
	- 1 587~ - 1 579	TTTTGTATT	✓		
1/2GRE	- 1 556~ - 1 549	TTTTGCCTG	✓		
	- 1 600~ - 1 594	AGGACAT	✓		
	- 1 514~ - 1 507	AA GGACAT	✓		
	- 130~ - 124	AGGAAGA		✓	
GC island	- 251~ - 247	GGGCC		✓	
Milk box	- 175~ - 153	TA TTCCCCTTTAAAA TGCTCCCCA		✓	
NF1(CAAT box)	- 89~ - 83	CAAA TCC		✓	
Oct-1	- 56~ - 51	AA TTAGC		✓	
TATA box	- 35~ - 27	TA TATAAA		✓	
Reti	- 11~ - 7	GATCA		✓	
TRE	+ 1 099~ + 1 103	GGTCA		intron 1 (内含子 1)	
	- 717~ - 712	TTCTCA		✓	
	+ 78~ + 83	TACTCA		exon 1 (外显子 1)	
Sm	- 150~ - 144	ATTTTTG		✓	
	+ 8 012~ + 8 018	ATGTTTG			✓
	+ 8 154~ + 8 160	ATGTTTG			✓
Glyco	- 288~ - 282	AA GA GA		✓	
	+ 8 187~ + 8 194	ACTGGAAA			✓
Pu-1	+ 8 025~ + 8 030	GAGGAA			✓
CK	- 738~ - 731	GACCCCTT		✓	
CAC	+ 1 099~ + 1 103	CACCC		intron 1 (内含子 1)	
PolyA	+ 8 477~ + 8 482	AA TAAA			✓
G/T	+ 8 497~ + 8 607	GnTn			✓

注: STAT5(MGF) 表示信号转导子和转录激活子 5(乳腺因子) 结合位点; C/EBP 表示 CCAAT/增强子结合蛋白结合位点; 1/2GRE 表示(半回文) 皮质激素应答元件; GC island 表示 GC 框, SP1 因子结合位点; Milk box 表示“乳盒”; NF1(CAAT box) 表示核因子-1 结合位点(CAAT 框); Oct-1 表示八聚体结合蛋白结合位点; TATA box 表示 TATA 框, TBF 因子结合位点; Reti 表示视黄酸反应元件; TRE 表示佛波酯反应元件; AP1 结合位点; Sm 表示核糖核蛋白 Sm 结合位点; Glyco 表示糖基化位点; Pu-1 表示 Pu-1 因子结合位点; CK 表示细胞因子 CK-2e 结合位点; CAC 表示 GT-1c 蛋白结合位点; PolyA 表示 polyA 加尾信号; G/T. GT 簇。

Note: STAT5 (MSF) stands signal transducers and activators of transcription 5(mammary gland factor); C/EBP stands CCAAT/enhancer binding protein; 1/2GRE stands half-palindromic glucocorticoid response elements; GC island stands GC box GC signal of RNA polymerase; “Milk box” stands mammary specific binding factor site; NF1 (CAAT box) stands NF-1 factor binding site (CAAT signal of RNA polymerase II); Oct-1 stands octamer-binding transcription factors binding site; TATA box stands RNA polymerase binding site; Reti stands retinoic acid response element; TRE stands TPA response element; Sm stands small nuclear ribonucleoprotein Sm binding site; Glyco stands glycosylation site; Pu-1 stands Pu-1 factor binding site; CK stands cytokine-2e binding site; CAC stands GT-1c protein binding site; PolyA stands polyA tail signal; G/T stands GT cluster

3 讨 论

本次研究分离克隆了奶牛 CSN 2 长为 2 826 bp

的 5 调控区, 包括 1.7 kb 的上游调控序列以及 5 不翻译区中的第一外显子和部分第一内含子。考虑到目的片段的长度和直接从基因组 DNA 中分离扩增

基因的难度, 研究中采用了重组 PCR 的方法, 先分别进行 2 个短片断的扩增, 而后利用重叠区进行拼接重组。实验结果证实此法有效可靠, 在使用较长引物以保证特异性的情况下, 重组 PCR 一次性获得成功, 而且产物纯, 杂带极少。序列分析表明, 试验扩增的 5 调控区包含有许多重要的转录因子结合位点, 尤其是 STAT (M GF)、C/EBP、1/2GRE、milk box 以及 TATA box, 这些 β 酪蛋白基因优势表达的核心调控序列的存在, 保证了所克隆 5 调控区能够驱动外源基因特异性在乳腺中的高效表达^[8,9]。另外, 研究还表明, β 酪蛋白基因第一内含子中含有 SBS (silence-binding site) 和 PR (progesterone receptor) 2 种 L 2 型核元件, 其对于 CSN 2 基因表达的作用是负性的^[10]。因此, 本研究中只扩增了第一内含子的部分上游区域, 对于 2 种负性核元件可能存在的下游区域没有扩增。实际运用中, 也可将外源基因直接克隆在编码信号肽的第一外显子后面。另外, Persuy 等^[11]用山羊 CSN 2 基因长 3 0 kb 的 5 调控区指导

CSN 2 和 κ CSN 基因在细胞水平上高效表达, 也支持了本研究的合理性。

本研究分离克隆的 3 端调控区包含 CSN 2 基因最后一个内含子的小部分、最后一个不翻译的外显子及 150 bp 的侧翼区。调控区主要包含了 polyA 加尾信号及控制转录终止的 G/T 簇, 这些序列结构对 mRNA 的转录、稳定性及介导其翻译调控起到重要的作用。以之构建乳腺特异性表达载体的 3 端调控区, 可与 5 调控区通过分子内的“长程”互作 (long-range interaction) 在靶基因转录及翻译中发挥作用^[12]。

用本研究分离克隆的奶牛 CSN 2 5 和 3 端启动子构建靶基因的乳腺特异性表达载体, 仅仅相当于以靶基因的 cDNA 或基因组代替 CSN 2 的结构基因部分, 基本符合内源性 CSN 2 表达调控模式。如果表达载体在靶细胞中整合的位点合适, 便能够指导靶基因在乳腺上皮细胞或乳腺组织中进行有效表达, 作者目前正在从事这方面的研究。

[参考文献]

- [1] Cerdan M G, Yang J I, Zino E, et al Accurate spatial and temporal transgene expression driven by a 3.8-kilobase promoter of the bovine beta-casein gene in the lactating mouse mammary gland[J]. Mol Reprod Dev, 1998, 49(3): 236- 245
- [2] Oh K B, Choi Y H, Kang Y K, et al A hybrid bovine beta-casein/bGH gene directs transgene expression to the lung and mammary gland of transgenic mice[J]. Transgenic Res, 1999, 8(4): 307- 311
- [3] Gutierrez A, Meade H M, Dittullo P, et al Expression of a bovine kappa-CN cDNA in the mammary gland of transgenic mice utilizing a genomic milk protein gene as an expression cassette[J]. Transgenic Res, 1996, 5(4): 271- 279
- [4] Ninomiya T, Hirabayashi M, Sagara J, et al Functions of milk protein in gene 5'-flanking region on human growth hormone gene[J]. Mol Reprod Dev, 1994, 37: 276- 283
- [5] 黄 赞, 颜景斌, 黄 纓, 等 山羊 β 酪蛋白基因启动子指导的转基因小鼠乳汁高效表达人凝血因子 IX [J]. 遗传学报, 2002, 29(3): 206- 221
- [6] 张克忠, 陈 立, 卢大儒, 等 牛 β 酪蛋白基因控制的人凝血 IX 因子基因在小鼠乳腺组织中的表达调控[J]. 中国科学(C 辑), 1998, 28(5): 463- 469
- [7] Sambrook J, Fritsch E F, Maniatis T, et al Molecular cloning a laboratory manual[M]. 2nd ed New York: Cold Spring Harbor Laboratory Press, 1998
- [8] Jeffrey M R, Shannon L W, Darryl H. Regulation of milk protein gene expression[J]. Annu Rev Nutr, 1999, 19: 407- 436
- [9] 郭继彤, 张肇英 外源基因在转基因动物乳腺中的特异性表达研究[J]. 内蒙古大学学报(自然科学版), 1998, 29(6): 853- 859
- [10] 王 强, 陈美钰, 任兆瑞, 等 山羊 β 酪蛋白基因克隆及序列分析[J]. 生物化学与分子生物学报, 2003, 19(1): 17- 23
- [11] Persuy M A, Printz C, Medrano J F, et al A single nucleotide deletion resulting in a premature stop codon is associated with marked reduction of transcripts from a goat β casein null allele[J]. Animal Genetics, 1999, 30: 444- 451
- [12] 石统东, 吴玉章, 朱锡华 真核 mRNA 3' 非翻译区在基因表达中的作用[J]. 生物化学与生物物理进展, 1998, 25(3): 195- 197

Molecular cloning and sequence analysis of bovine beta-casein gene 5' , 3' control region

L IU J in-long, ZHENG Yue-mao, WANG Yu-jie, ZHANG Yong

(Institute of Bio-Engineering, Northw est Sci-Tech University of Agriculture and Forestry, Yangling, Shaanxi 712100, China)

Abstract: 5' and 3' control region of bovine CSN 2 were partially cloned by normal and recombinant PCR. The former consisted in all of the 5' upstream sequence, the first exon and part of first intron, while the later included the last exon and 3' flanking fragment. The purified PCR products were sub-cloned to pMD 18-T Vector, then sequenced correspondently and analysed by bio-soft. The results indicated that the two cloned fragments have the homology of 99.0% and 98.0% respectively with the comparable region of bovine CSN 2. Moreover, all the vital regulatory elements and many factors binding sites were found in the two fragments. The above suggested that the expressing vector, constructed by the two cloned fragments and work under the control model of endogenous CSN 2, could direct the target gene to express specifically and efficiently in mammary gland.

Key words: cow; CSN 2; control region; PCR; sequence analysis

(上接第 98 页)

A molecular biology method to differentiate C-strain from other strains of CSFV

HU Jian-he¹, CHEN Yong-yao¹, Wei Gang-cai¹,

Wang San-hu¹, ZHANG Yan-ming², XIE Qing-ge³

(1 Animal Science Department of Henan Vocational-technical Teacher's College, Xinxiang, Henan 453003, China;

2 College of Animal Science and Technology, Northw est Sci-Tech University of Agriculture and Forestry, Yangling, Shaanxi 712100, China;

3 Lanzhou Veterinary Research Institute, Lanzhou, Gansu 730046, China)

Abstract: A difference marker sequence CTTTTTCTTTT was selected from C-strain and other strains of CSFV, and specific PCR primer pair and nested-PCR primer pair were designed at the basis of previously published sequences of CSFV. Total RNA was extracted from spleens or lymph nodes of suspicious ill pigs. Target fragments were amplified utilizing RT-PCR and nested-PCR techniques. After purification and recovery, the target fragments were cloned into pMD18-T vectors and further sequenced. According to sequence analysis, sequence containing the marker sequence belongs to C-strain and others belongs to other strains of CSFV.

Key words: classical swine fever; RT-PCR; nested-PCR; marker sequence; differentiation diagnosis