

用牛 S1-酪蛋白基因构建乳腺生物反应器表达载体的研究

管晓斌 李春笑 赖松家* (四川农业大学动物科技学院, 四川雅安 625014)

摘要 牛 S1-酪蛋白基因是制备乳腺生物反应器常用的表达载体, 可以指导外源基因在动物乳腺中高水平表达。对牛 S1-酪蛋白基因的结构、表达载体的各种调控元件及其研究应用进展等方面作一综述。

关键词 牛 S1-酪蛋白基因; 表达载体; 乳腺生物反应器; 转基因

中图分类号 Q789 文献标识码 A 文章编号 0517-6611(2006)23-6140-02

Construction of the Mammary Gland-specific Expression Vector by Bovine S1-casein Gene

GUAN Xiaolin et al (The College of Animal Science and Technology, Sichuan Agricultural University, Yaan, Sichuan 625014)

Abstract Bovine S1-casein gene was usually used as expression vector to construct mammary gland bioreactor. It could make exogenous gene expressed highly in animal galactophore. In this article the structure of bovine S1-casein gene, the various regulative component of expression vector and its development of related research and application were reviewed.

Key words Bovine S1-casein gene; Expression vector; Mammary gland bioreactor; Transgenesis

自从1987年Gordon等将人组织纤溶酶原激活剂(tPA) cDNA与小鼠乳清酸蛋白(WAP)基因启动区构建融合表达载体, 首建乳腺生物反应器小鼠模型以来, 在不到20年的时间里, 人乳铁蛋白(hLF)、人血清白蛋白(hSA)、人促红细胞生成素(hEPO)、人1-抗胰蛋白酶(h1-AT)、人凝血因子(hF)、人尿激酶(hUK)等诸多重要药物蛋白在转基因动物中已经获得有效表达。牛S1-酪蛋白基因是制备乳腺生物反应器常用的调控成分, 可以指导外源基因在动物乳腺中高水平表达。因此深入探讨牛S1-酪蛋白基因在构建乳腺生物反应器表达载体中的作用, 对推动转基因药物的研制和现代动物育种技术的发展意义重大。

1 牛 S1-酪蛋白基因的结构

牛S1-酪蛋白基因全长17 508 bp, 其中19个外显子占1 138 bp, 大小在24~385 bp; 18个内含子占16 370 bp, 长度为90~1967 bp。通过全序列的分析发现, 牛S1-酪蛋白基因结构有以下特点^[1]: 全基因序列中, A和T高达67%, 这是一般的基因序列很少见的。53 bp的第1外显子为5非翻译区。信号肽基因以及成熟蛋白质的前2个氨基酸由63 bp的第2外显子编码。除终止密码子UGA是由第17外显子的最后2个碱基UG和第18外显子的第1个碱基A经剪切后拼接而成外, 其余三联体密码子均不受剪切影响, 因此第3到第16外显子都是三碱基对的倍数。16个编码外显子中有9个均以“GAX”起始, 这与cDNA序列分析结果一致。序列中主要的磷酸化位点是由第10外显子的第1个密码子(GAA)与前一外显子拼接后产生。2段154 bp的区域(+9 101~+9 254和+11 489~+11 642)的碱基顺序几乎完全一致, 同源性高达97.4%, 仅有4个C/T碱基转变。在启动区有2个功能性TATA盒, 分别位于-23~-29和-92~-98处, 序列为“TTTAAAT”。第18外显子和含有polyA信号的第19外显子为3非翻译区。

2 牛 S1-酪蛋白基因表达载体的调控元件

2.1 顺式作用元件 牛S1-酪蛋白基因的2个具有潜在功能的TATA盒分别位于-23~-29和-92~-98的位置上,

其序列是“TTTAAAT”。通过与牛及其他物种的乳蛋白基因序列进行比较, 推断出牛S1-酪蛋白基因的乳腺组织特异性转录因子(MGF)的结合位点可能位于-87~-100的位置, 核苷酸序列为“AATTCTTAGAATT”, 这段序列几乎在所有的乳蛋白基因中都是保守的。另一种乳腺特异性转录因子叫做乳蛋白结合因子(MPBF), 它在牛S1-酪蛋白基因的结合位点可能有2个, 一个位于-333~-346的位置上, 其序列是“TGTCCTAAGAATT”; 另一个位于-11~-24的位置上, 其序列是“ATAGCTTGGAAGC”。在牛S1-酪蛋白基因的-157~-148的位置上有一段叫“milk box”的序列“TCTCCTTGAATTCCTGTTGG”, 是一种类似于CTF/NF-1活性的一般性核转录因子的结合位点。而另一种一般性核转录因子oct-1在牛S1-酪蛋白基因序列上的核心结合位点可能位于-47~-54的位置, 序列是“AATTAGCA”。从以上序列分析的结果来看, 牛S1-酪蛋白基因的重要调控元件均位于启动区-400~-10。利用乳蛋白基因启动区进行乳腺特异性表达的转基因试验表明, 只要乳蛋白基因启动区包括了-500~+500的区域, 一般就能够进行高水平的乳腺特异性表达^[2]。

2.2 内含子 从牛S1-酪蛋白基因启动区来看, 它不具有CAAT盒, TATA盒序列“TTTAAAT”为一弱启动区, 但牛S1-酪蛋白基因的表达量却很高, 因此推测除了5端调控机制外, 内含子及3端在表达过程中会起到一定的作用。内含子可充当转录促进子, 在基因组的DNA上内含子核苷酸序列中存在转录调节蛋白的结合位点。此外, 内含子上的剪接信号也能刺激转录, 因为RNA的加工与转录是一个互动的过程, 剪接过程能反馈促进RNA聚合酶的活动^[3]。内含子除能促进转录外还可影响翻译。将成熟的mRNA直接注射到核中, 当它被运送到胞质中时并不能进行翻译, 而在注射核酸结合蛋白的抗体或mRNA的3端非编码区包含1个可剪接的内含子时, 上述抑制被解除^[4]。有很多试验都表明采用基因组基因表达构件比采用cDNA表达构件提高表达量10~100倍。此外, 异源内含子(特别是第1、2内含子)可提高外源基因(特别是以cDNA为表达构件)的表达效率。

2.3 UTR 真核生物mRNA的5端和3端非翻译区(untranslated region, UTR)也是调控基因表达的元件。3-UTR不仅控制mRNA的体内稳定性和降解速度, 而且能够控制基因

作者简介 管晓斌(1982-), 男, 四川阆中人, 硕士研究生, 研究方向: 分子遗传育种与胚胎工程。* 通讯作者, E-mail: 1sj5791@263.net。

收稿日期 2006-08-07

表达的组织特异性;3-UTR 还是糖皮质激素和某些细胞因子如 -干扰素、白介素2 和肿瘤坏死因子等对 mRNA 的调控位点;此外,3-UTR 还可以辨认某些特殊密码子,控制 mRNA 的利用效率。在真核细胞所有已知的 mRNA 中,位于 mRNA 的 5 端的帽子结构都与 40S 核糖体亚基结合。由 40S 核糖体亚基与起始因子所形成的复合物沿着 mRNA 的 5-UTR 迁移,直到碰到起始密码子 AUG。这种搜寻过程可能很慢,甚或被 5-UTR 的二级结构所中断。5-UTR 中富含 GC 的序列能形成一种稳定的二级结构抑制 mRNA 的翻译^[5]。

2.4 MAR 染色体是由许多独立的染色质环组成。染色质环的形成是由染色质上一段特殊的 DNA 序列-核基质结合区(Matrix attachment region, MAR) 锚定到主要由蛋白质成分组成的核基质上,从而把染色体分隔为一个个环状结构。研究表明,作为真核生物特有的一种基因表达调控元件, MAR 与增强子不同,它对邻近基因的表达没有增强作用,只对稳定整合于染色体上的外源基因表现出位点不相关而拷贝数相关的增强作用^[6]。在载体构建时,添加顺式和反式 2 种 MAR 元件比单独添加顺式或反式 MAR 元件效果更好。这说明多个 MAR 元件对外源基因的联合整合作用更能使其高水平表达^[7]。

2.5 LCR 人们在一些基因中发现了一种调控元件,它们具有转录增强活性,能够保障外源基因可以直接被转录而不依赖 DNA 在染色体上随机整合的位置。这种调控元件就是基因座控制区(Locus control region, LCR)。据估计,牛酪蛋白基因座存在 LCR,使得酪蛋白基因可以成为一个完全独立的表达单位。研究表明,含外源 LCR 元件的表达载体比不含 LCR 元件的表达载体更有效且受位置效应和基因沉默的影响更小^[8]。用于乳腺生物反应器表达载体的构建时,LCR 能对同源或异源的启动子起作用,使转基因表达呈位点非依赖性,一般认为它具有增强子和绝缘子的双重功效。

3 牛 S1-酪蛋白基因表达载体的研究与应用

3.1 国外的研究现状 10 余年来,世界各国的研究者对牛 S1-酪蛋白基因表达载体进行了大量的研究,取得了显著成果。1990 年 Made 等^[9]用牛 S1-酪蛋白基因的 5 端包括 21 kb 的上游序列和第 1、2 外显子翻译起始位点以前的部分,3 端包括编码蛋白羧基端一半的外显子和 PolyA 信号及 2 kb 的下游侧翼序列作为调控序列指导人尿激酶基因在转基因小鼠乳腺中表达,表达量为 1~2 ng/ml。1993 年 Rego 等^[10]用牛 S1-酪蛋白基因 1.6 kb 的调控序列和兔的 -球蛋白基因、SV40 基因的 3-UTR 共同指导人组织纤溶酶原激活剂 cDNA 序列在转基因小鼠和兔乳腺中表达,表达量分别为 50 $\mu\text{g}/\text{ml}$ 和 8~50 ng/L。1994 年 Hatenburg 等^[11]应用 6.2 kb 牛 S1-酪蛋白基因 5 端调控区与人乳铁蛋白基因制备了转基因小鼠,人乳铁蛋白的表达量为 3 ng/ml。而将其中的人乳铁蛋白基因置换为 cDNA 后,乳铁蛋白的表达水平明显下降,在动物乳汁中的表达量仅为 0.1~36 $\mu\text{g}/\text{ml}$ 。2001 年 Zavadskai a 等^[12]用牛 S1-酪蛋白基因调控序列指导内皮细胞抑制素在小鼠乳腺中表达,表达量为 70~300 $\mu\text{g}/\text{ml}$,并且能稳定地遗传给后代。2005 年 Zakhrova 等^[13]用牛 S1-酪蛋白基因调控序列指导人乳铁蛋白在小鼠乳腺中表达,用 Western 印记法在乳

中未检测到人乳铁蛋白,但在乳腺中发现有转录产物 mRNA。

3.2 国内的研究现状 与发达国家相比较,我国对牛 S1-酪蛋白基因表达载体的研究起步较晚,但也取得了可喜的成绩。1997 年张靖溥等^[14]将牛 S1-酪蛋白基因 5 端含转录起始点上游约 8 kb、下游(包括 5-UTR)约 3.5 kb 的基因片段和 3 端含 PolyA 信号的 4.5 kb 基因片段作为调控序列与人乙肝表面抗原(HBsAg)基因的 2.3 kb 片段拼接,构建了融合基因表达载体 106,在转基因山羊乳腺中表达阳性率为 44%,阳性羊的后代的乳汁仍为 HBsAg 阳性,提示构建框架已正确表达且表达性能稳定。1999 年该组研究人员又以小鼠为模型,发现外源基因在转基因小鼠中可以稳定遗传并在后代中表达目标产物^[15]。2001 年谭晓红等^[16]用牛 S1-酪蛋白基因 5 端约 3.6 kb 的片段和 3 端约 1.4 kb 的片段作为调控序列,指导人组织纤溶酶原激活剂基因在小鼠乳腺中表达,表达量为 0.18 $\mu\text{g}/\text{ml}$ 。2001 年张传生等^[17]以小鼠乳清酸蛋白基因 5 端调控区和牛 S1-酪蛋白基因 3 端调控区作为调节元件构建了用于表达人促血小板生成素(hTPO)的乳腺组织特异性表达载体 pWAPTPO,共 9.8 kb。结果 hTPO 在小鼠乳腺中得到表达,表达量在 0.8 $\mu\text{g}/\text{ml}$ 以上。

3.3 在商业生产上的应用 1990 年美国的 GenPharming 公司在世界上首先培育了含有人乳铁蛋白的转基因牛。公司用牛 S1-酪蛋白基因 5、3 端旁侧序列开发了表达载体,构建了人乳铁蛋白 cDNA 表达盒,通过原核显微注射法制备了人乳铁蛋白转基因牛。其生产的人乳铁蛋白于 1997 年进入临床试验,2000 年正式推入市场,这种转基因奶牛年产牛奶 10 t,每年生产的含人乳铁蛋白的奶粉价值达 60 亿美元。这一成就为今后牛 S1-酪蛋白基因乳腺表达载体的更为广泛的应用奠定了坚实基础。

4 问题与展望

天然乳蛋白的表达量都高达毫克级,而用牛 S1-酪蛋白基因的调控序列来指导外源基因的表达结果往往偏低,并且表达先定性和遗传稳定性往往不高。主要原因在于人们还没有完全掌握载体构建的规律,对牛 S1-酪蛋白基因表达载体的各种调控元件的基本功能、作用机制等都有待进一步的研究。另外,目前的研究大都在小鼠、兔等小动物上进行,且大都还处于实验室研究阶段。在小鼠上的研究结果未必与在牛、羊等大家畜上的研究结果一致,表达结果的不确定性阻碍着牛 S1-酪蛋白基因表达载体的广泛应用。但是目前国外已经在商业生产应用方面走出了开创性的一步,并且取得了不错的经济效益。1997 年美国红十字会和遗传学会预测,至 2010 年,所有基因工程药物中利用动物乳腺生物反应器生产的份额将高达 95%。尽管牛 S1-酪蛋白基因表达载体的研究道路还会很曲折,但由于各国研究者卓有成效的工作,相信在不久的将来牛 S1-酪蛋白基因表达载体在乳腺生物反应器技术中占有重要的地位,并极大地推动乳腺生物反应器的商业化进程。

参考文献

- [1] KOCZAND, HOBOMG, SEYFERT H M Genomic organization of the bovine alpha-S1 casein gene[J]. Nucleic Acids Res, 1991, 19(20): 5591-5596.
- [2] 李宁, 吴常信, 陈永福. 牛 S1 酪蛋白基因启动区的克隆和序列分析[J]. 遗传, 1997, 19(1): 4-8.

(上接第6141页)

- [3] MANLEY J L. Nuclear coupling : RNA processing reaches back to transcription [J] . *Nat Struct Biol* ,2002 ,9(11) :790 - 791 .
- [4] LE HR H,NOTT A,MOORE MJ. How introns influence and enhance eukaryotic gene expression[J] . *Trends Biochem Sci* ,2003 ,28(4) :215 - 220 .
- [5] TANG W, TSENG H. A GC-rich sequence within the 5' untranslated region of human basonudin mRNA inhibits its translation[J] . *Cere* , 1999 ,237(1) :35 - 44 .
- [6] JENWEINT, FORRESTER WC. Extension of chromatin accessibility by nuclear matrix attachment regions[J] . *Nature* ,1997 ,385(6 613) :269 - 272 .
- [7] GROD PA, ZAHN ZABAL M, MERMOD N. Use of the chicken lysozyme 5' matrix attachment region to generate high producer CHO cell lines[J] . *Botechnol Bioeng* ,2005 , 91(1) :1 - 11 .
- [8] WANG H, SHAYAKHMETOV DM, LEEGE T, et al . A capsid-modified helper-dependent adenovirus vector containing the beta-globin locus control region displays a nonrandom integration pattern and allows stable , erythroid-specific gene expression[J] . *J Virol* ,2005 ,79(17) :10 999- 11 013 .
- [9] MEADE H, GATES L, LACY E, et al . Bovine alpha S1-casein gene sequences direct high level expression of active human urokinase in mouse milk[J] . *Botechnology* , 1990,8(5) :443 - 446 .
- [10] RIEGO E, LIMONTA J, AGUILAR A, et al . Production of transgenic mice and

rabbits that carry and express the human tissue plasminogen activator cDNA under the control of a bovine alpha S1 casein promoter[J] . *Theriogenology* , 1993,39(5) :1 173 - 1 185 .

- [11] PLATENBURG GJ, KOOTWIJK EP, KOOMAN PM, et al . Expression of human lactoferrin in milk of transgenic mice[J] . *Transgenic Res* ,1994 ,3(2) :99 - 108 .
- [12] ZAVADSKAYA ES, ZAKHAROVA ES, KADUIN SG, et al . Production of recombinant endostatin in milk from transgenic mice[J] . *Genetika* ,2001 ,37(9) :1 207 - 1 212 .
- [13] ZAKHAROVA ES, PRVZHKOVA MV, KIBARDIN AV, et al . Transcription and mRNA splicing of the human lactoferrin gene controlled by the regulatory region of the bovine alpha S1 casein gene in the mammary gland of transgenic mice and in mouse embryonic stem cells[J] . *Genetika* ,2005 ,41(3) :299 - 306 .
- [14] 张靖溥, 劳为德, 成勇, 等. 牛-S1-酪蛋白-乙肝病毒表面抗原融合基因在转基因羊中的表达[J] . *生物工程学报*,1997,13(2) :154- 159 .
- [15] 张靖溥, 劳为德, 张旭晨, 等. 外源基因在转基因小鼠中遗传和表达的稳定性的研究[J] . *遗传学报*,1999,26(2) :135 - 141 .
- [16] 谭晓红, 程萱, 周江, 等. 牛-S1酪蛋白基因序列指导hPA突变体小基因在小鼠乳腺中的表达[J] . *遗传学报*,2001,28(5) :405- 410 .
- [17] 张传生, 高虹, 王玉阁, 等. 人促血小板生成素在转基因小鼠乳腺中定位表达的研究[J] . *Developmental and Reproductive Biology* ,2001 ,10(2) :9 - 14 .