

山羊 β -酪蛋白基因调控区的克隆及 启动子活性的检测

郑艳玲,唐爽,张涌

(西北农林科技大学 生物工程研究所,陕西 杨凌 712100)

[摘要] 为了克隆山羊 β -酪蛋白基因的调控区并检测5'调控区的启动子活性,利用Long PCR技术对西农莎能奶山羊 β -酪蛋白基因(CSN2)5'和3'调控区进行了克隆(5'调控区分两段进行克隆,3'调控区进行直接克隆)、测序,将克隆的5'调控区的两个片段通过共有的单一性酶切位点 *Hind* III融合成一个长为6.7 kb的片段,用其替代表达载体pEGFP-C1上原来的启动子CMV,指导报告基因绿色荧光蛋白(GFP)在山羊乳腺上皮细胞中的瞬时表达。结果表明,克隆的GC 3.3、GC 4.3、GC 6.6 3个片段与GenBank中登录的山羊CSN2基因相应序列相比,同源性均为99%,其包含了TATA box、GC island、CAAT box、SAR、AP1、Oct-1、Sm、Glyco、Pu-1、CAC等复合元件和转录翻译调节因子的作用位点,并且5'调控区可以有效指导报告基因的表达,从而初步验证了所克隆的山羊CSN2基因5'调控区的有效性,表明克隆的调控区可用于乳腺生物反应器的研究。

[关键词] 山羊 β -酪蛋白基因(CSN2);调控区;PCR;GFP基因

[中图分类号] Q785;Q786

[文献标识码] A

[文章编号] 1671-9387(2007)06-0001-05

Cloning of control region of goat CSN2 and the detection of the promoter

ZHENG Yan-ling, TANG Shuang, ZHANG Yong

(Institute of Bioengineering, Northwest A & F University, Yangling, Shaanxi 712100, China)

Abstract: To clone the 5' and 3' control region of goat CSN2 gene, and to detect the promoter's activity of 5' control region. 5' and 3' control region of Xinong Saanen dairy goat CSN2 were partially cloned by Long PCR and sequenced. The 5' control region was divided into two fragments to be cloned. The first fragment of 5' control region was joined to the second one through the *Hind* III restriction site, so a longer fragment which is 6.7 kb in length was constructed. 3' control region was cloned directly. Sequence analysis on NCBI with blast indicated that the three cloned fragments had the homology of 99% respectively with the comparable region of goat CSN2. The three fragments contained many multiple components and factors binding sites, such as TATA box, GC island, CAAT box, SAR, AP1, Oct-1, Sm, Glyco, Pu-1, CAC, and so on. The 5' control region replaced previous promoter CMV cloned in pEGFP-C1 expressional vector and transfected the cultured mammary epithelial cell in vivo. Through observation under UV microscope, it showed that GFP had been successfully expressed in mammary epithelial cell. Then, the cloned control region could be served for the study of mammary gland expression vector construction.

Key words: goat; CSN2; control region; PCR; GFP gene

[收稿日期] 2006-04-28

[基金项目] 国家“863”高技术项目(2004AA213072)

[作者简介] 郑艳玲(1977-),女,内蒙古阿荣旗人,在读硕士,主要从事分子生物学研究。

[通讯作者] 张涌(1956-),男,内蒙古和林格尔人,教授,博士生导师,主要从事哺乳动物胚胎工程与发育生物学研究。

1985 年,Badge 等^[1]首次提出了用转基因动物乳腺生产重组蛋白的思路,1987 年,Gordon 等^[2]以人组织纤溶酶原激活剂(tPA)与小鼠乳清酸蛋白基因启动区构建融合表达载体,研制出了首例乳腺生物反应器小鼠模型。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、特异性表达的乳蛋白调控区。 β -酪蛋白是牛羊乳中含量最高的乳蛋白,山羊 β -酪蛋白在山羊乳中占 50%以上^[3],这说明山羊 β -酪蛋白基因(CSN2)在泌乳激素的刺激下具有很强的表达活性,在乳蛋白的表达过程中处于绝对优势地位,利用山羊 β -酪蛋白基因的调控区作为外源基因的调控序列能够指导外源基因的高效表达。在国外,Ebert 等^[4]用山羊 β -酪蛋白基因 5'端长度约 6.2 kb 和 3'端长度约 7.1 kb 的侧翼序列,构建了人长效 tPA 乳腺表达载体,其在山羊乳中的表达量达到 1~3 mg/mL;在国内,黄赞等^[5]用山羊 β -酪蛋白基因 5'端长度约 6.7 kb 的片段构建了人 F-IX 表达载体,其在转基因小鼠乳腺中的表达量高达 52.9 mg/L。由此可见,国内生产的转基因动物外源基因的表达量与国际水平还有很大差距,这可能是由于国内构建的表达载体都没有采用乳蛋白较长的 3'端调控区。为了构建一个拥有 β -酪蛋白基因较长 5'端和 3'端调控区的乳腺表达载体,本研究通过分段克隆的方法(分 GC 3.3 和 GC 4.3 两段,其长度分别为 3.3,4.3 kb)克隆山羊 β -酪蛋白基因 5'调控区,一次性克隆长度为 6.6 kb 的 3'调控区,并将 GC 3.3 和 GC 4.3 通过共有的单一限制性酶切位点连接成一个长为 6.7 kb 的片段 GC 6.7,以期构建拥有较长 5'端和 3'端调控区的高效乳腺表达载体的研究奠定基础。

1 材料与方法

1.1 材料

1.1.1 试验动物 西农莎能奶山羊,由杨凌科元生物工程有限公司提供。

1.1.2 质粒、菌种和细胞 质粒 pGEM-T Vector 购自美国 Promega 公司,宿主菌 DH5 α 、质粒 pEGFP-C1 及山羊乳腺上皮细胞,由西北农林科技大学生物工程研究所保存。

1.1.3 酶和试剂 LA-Taq DNA 聚合酶和 dNTP 购自大连宝生物工程有限公司;限制性内切酶 Age I 购自 NEB 公司,其他限制性内切酶和 DNA Marker 购自加拿大 MBI 公司;T4 DNA Ligase 购自美国 Promega 公司;DNA Gel Extraction Kit 和 Plasmid

Minipreps Kit 购自安徽优品生物公司;阳离子脂质体购自 Invitrogen 公司。

1.1.4 引物 根据 GenBank 中登陆的山羊 CSN2 基因序列(登陆号:AF409096),利用生物软件 Vector NTI Suite 7.0 和 Primer 5.0 辅助设计了 3 对引物,由上海生物工程公司合成。为了将 GC 3.3 与 GC 4.3 连成一个大片段 GC 6.7,及方便将来构建表达载体,在第 1 对引物的上游引物 GC 3.3 S5'端先后引入 Sal I(“—”所示)和 ApaI I(“=”所示)两个酶切位点;在第 2 对引物的下游引物 GC 4.3A 的 5'端引入一个 Age I(“—”所示)酶切位点,引物序列如下。

第 1 对引物:上游 GC 3.3 S 5'-GTCGAC GTG-CACTCGTGGTAGAGGTCAAGGATGCTGCT-3',

下游 GC 3.3 A 5'-GAA GGGCAAATGA GGGAA G-CAAGAAA-3';

第 2 对引物:上游 GC 4.3 S 5'-ATCTTGCTCC TCA T-TCAGGGTTATCT-3',

下游 GC 4.3 A 5'-ACCGGTTTGTCTGTGTTC-CTTCTTGTTGGTCT-3';

第 3 对引物:上游 GC 6.6 S 5'-TCCTCA TCCTT-GCCTGTCTGGTG-3',

下游 GC 6.6 A 5'-ATAACTCTTCATCTCACCCCATTTCA-3'。

3 对引物分别用于扩增片段 GC 3.3、GC 4.3 和 GC 6.6。由于引物上设计了酶切位点,这些酶切位点在合成引物时加到了引物 5'端,在引物扩增基因过程中可直接扩增到基因端侧。

1.2 山羊基因组 DNA 的提取

山羊颈静脉采血,0.5 g/L EDTA 抗凝后按照文献[6]的方法提取基因组 DNA。

1.3 山羊 β -酪蛋白基因的 PCR 扩增及产物纯化

1.3.1 PCR 扩增 PCR 反应体系均按 LA-Taq DNA 聚合酶说明书配制,在 0.2 mL PCR 管中依次加入:10 $\times$ PCR Buffer 5 μ L,dNTP Mixture 8 μ L (2.5 mmol/L),上、下游引物各 1 μ L (10 μ mol/L),山羊基因组 DNA 样品 1 μ L (100 ng),LA-Taq DNA 聚合酶(10 U/ μ L)0.5 μ L,加超纯水至 50.0 μ L。PCR 反应条件分别如下:GC 3.3 的反应条件为 94 $^{\circ}$ C 预变性 3 min;94 $^{\circ}$ C 变性 30 s,67 $^{\circ}$ C 退火 30 s,72 $^{\circ}$ C 延伸 220 s,30 个循环;最后 72 $^{\circ}$ C 延伸 10 min。GC 4.3 的反应条件为 94 $^{\circ}$ C 预变性 3 min;94 $^{\circ}$ C 变性 30 s,63 $^{\circ}$ C 退火 30 s,72 $^{\circ}$ C 延伸 300 s,30 个循环;最后 72 $^{\circ}$ C 延伸 10 min。GC 6.6 的反应条件

为 94 °C 预变性 3 min; 94 °C 变性 30 s, 66 °C 退火 30 s, 72 °C 延伸 400 s, 30 个循环; 最后 72 °C 延伸 10 min。

1.3.2 片段 GC 3.3、GC 4.3、GC 6.6 的回收纯化和 T-A 克隆 利用 DNA Gel Extraction Kit 回收纯化各片段, 然后用 T-A 克隆法将 3 个片段分别克隆至 pGEM-T Vector 的 T 位点, T-A 克隆的反应体系按照 pGEM-T Vector 说明书配制, 克隆产物转化 DH5α 感受态细胞, 蓝白斑筛选阳性菌落, 具体方法见文献[6]。GC 3.3、GC 4.3 和 GC 6.6 克隆到 T 载体上得到的阳性菌落分别命名为 pGGC 3.3、pGGC 4.3 和 pGGC 6.6。

1.4 重组质粒的筛选和序列测定

用 Plasmid Minipreps Kit 提取质粒后, 分别用 *Sal* I 和 *Hind* III 双酶切鉴定重组质粒 pGGC 3.3 和 pGGC 4.3, 用 *Not* I 酶切鉴定重组质粒 pGGC 6.6, 最后送交上海英骏生物技术有限公司测序, 测序结果经 NCBI 网站的 blastn 软件进行同源性分析。

1.5 片段 GC 3.3 和 GC 4.3 的连接及重组质粒的筛选、酶切鉴定

由于 pGEM-T Vector 上没有 *Hind* III 酶切位点, 其多克隆位点有一个 *Sal* I 酶切位点, 片段 GC 3.3 和 GC 4.3 均有一个 *Hind* III 单一性酶切位点, 但没有 *Sal* I, 所以可以用 *Sal* I 和 *Hind* III 对 pGGC 3.3 和 pGGC 4.3 进行双酶切。对 pGGC 3.3 酶切后回收长度为 2.5 kb 的小片段 GC 2.5, 对 pGGC 4.3 酶切后回收长度为 7.1 kb 的大片段 GC 7.1, 用 T4 DNA Ligase 系统, 将回收纯化的片段 GC 2.5 和 GC 7.1 进行定向连接, 反应体系按照 T4 DNA Ligase 系统说明书配制。将连接产物转化 DH5α 感受态细胞, 蓝白斑筛选阳性菌落^[6], 用 Plasmid Minipreps Kit 提取质粒, *Apa* I 和 *Age* I 进行双酶切鉴定, 正确的重组质粒命名为 pGGC 6.7。

1.6 GFP 乳腺表达载体的构建

分别用 *Apa* I 和 *Age* I 双酶切载体 pEGFP-C1 和 pGGC 6.7, 并回收目的片段, 然后用 T4 DNA 连接酶连接构建载体, 构建的载体命名为 pGGC 6.7-GFP。对构建的载体进行 *Apa* I 和 *Age* I 双酶切鉴定。

1.7 山羊乳腺上皮细胞的体外培养

从液氮罐中取出冷冻保存的山羊乳腺上皮细胞, 37 °C 快速复苏后加入 5 mL DMEM/F12 培养液, 待细胞长满平皿后, 用 2.5 g/L 胰蛋白酶消化后备用。

1.8 pGGC 6.7-GFP 转化山羊乳腺细胞

将密度为 $1 \times 10^5 \text{ mL}^{-1}$ 的山羊乳腺上皮细胞接种于 24 孔板中, 培养 24 h 后, 待细胞约长满培养皿底部面积的 80 % 后, 用脂质体包裹的 pGGC 6.7-GFP 质粒转染细胞, 48 h 后在荧光显微镜下观察报告基因的表达情况。

2 结果与分析

2.1 片段 GC 3.3、GC 4.3 和 GC 6.6 的 PCR 扩增结果

经 PCR 扩增, 获得了长度为 3.3 kb 的 GC 3.3、长度为 4.3 kb 的 GC 4.3 和长度为 6.6 kb 的 GC 6.6 3 个片段(图 1), 与预期的扩增产物大小一致。

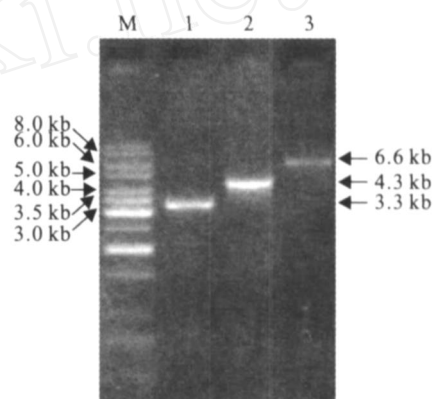


图 1 GC 3.3、GC 4.3 和 GC 6.6 的 PCR 扩增结果

M. DNA Marker; 1. GC 3.3 扩增产物;

2. GC 4.3 扩增产物; 3. GC 6.6 扩增产物

Fig. 1 PCR products of GC 3.3, GC 4.3 and GC 6.6

M. DNA Marker; 1. PCR product of GC 3.3;

2. PCR product of GC 4.3; 3. PCR product of GC 6.6

2.2 重组质粒的酶切鉴定结果

重组质粒 pGGC 3.3 进行双酶切后得到 3.0, 2.5 和 0.8 kb 3 条片段(图 2), 说明 GC 3.3 反连于 pGEM-T Vector 上; 重组质粒 pGGC 4.3 进行双酶切后得到 7.1 和 0.2 kb 两条片段(图 2), 说明 GC 4.3 正连于 pGEM-T Vector 上; 用 *Not* I 对重组质粒 pGGC 6.6 进行单酶切后, 切出 6.6 和 3.0 kb 两条片段(图 2), 以上结果表明重组质粒的酶切结果均与预期结果一致。用 *Apa* I 和 *Age* I 对重组质粒 pGGC 6.7 进行双酶切, 用 *Hind* III 对其进行单酶切, 双酶切切出 6.7 和 3.0 kb 两条片段, 单酶切切出一条长度为 9.7 kb 的片段(图 3)。

2.3 重组质粒的序列分析结果

序列分析结果表明, 片段 GC 3.3、GC 4.3 和 GC 6.6 与 GenBank 中山羊 CSN2 基因对应区段的

同源性均为 99 %。表明克隆的目的片段为山羊 CSN2 基因的序列,所克隆的 5' 调控区包括 4.2 kb 的 5' 侧翼区、第 1~8 个外显子、第 1~7 个内含子以及部分第 8 内含子;3' 包括部分第 2 外显子、部分第 8 内含子、第 2~8 外显子以及第 3~7 内含子。应

用 DNA Star 软件和启动子区分析软件 Ali Baba2 分析 GC 3.3、GC 4.3 和 GC 6.6,结果表明,3 个片段包含了 TA TA box、GC island、CAAT box、SAR、AP1、Oct-1、Sm、Glyco、Pur-1、CAC 等复合元件和转录翻译调节因子的作用位点。

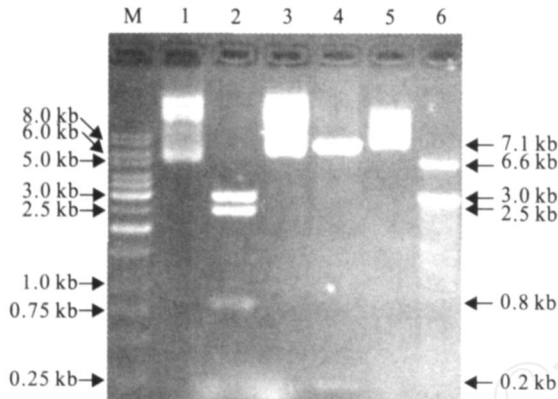


图 2 质粒 pGGC 3.3、pGGC 4.3 和 pGGC 6.6 的酶切鉴定结果

M. DNA Marker;1. 质粒 pGGC 3.3;2. 质粒 pGGC 3.3 的 *Hind* III 和 *Sal* I 酶切结果;3. 质粒 pGGC 4.3;4. 质粒 pGGC 4.3 的 *Hind* III 和 *Sal* I 酶切结果;5. 质粒 pGGC 6.6;6. 质粒 pGGC 6.6 的 *Not* I 酶切结果

Fig. 2 Digested result of pGGC 3.3、pGGC 4.3 and pGGC 6.6

M. DNA Marker;1. Plasmid pGGC 3.3;2. pGGC 3.3 digested by *Hind* III and *Sal* I;3. Plasmid pGGC 4.3;4. pGGC 4.3 digested by *Hind* III and *Sal* I;5. Plasmid pGGC 6.6;6. pGGC 6.6 digested by *Not* I

2.4 GFP 基因乳腺表达载体的构建

用 *Apa*I 和 *Age*I 双酶切 pGGC 6.7-GFP 乳腺表达载体,结果见图 4。图 4 表明,克隆的 CSN2 基因 5' 调控序列已连接在 GFP 上游。

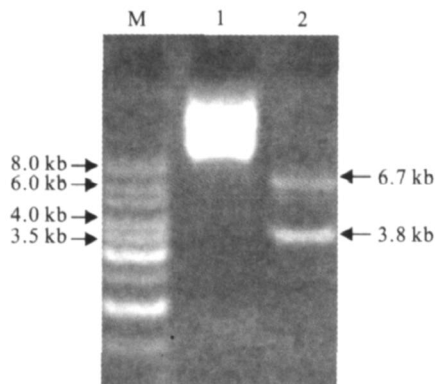


图 4 质粒 pGGC 6.7-GFP 的酶切鉴定结果
M. DNA Marker;1. 质粒 pGGC 6.7-GFP;2. *Apa*I 和 *Age*I 酶切结果

Fig. 4 Identified result of pGGC 6.7-GFP vector
M. DNA Marker;1. Plasmid of pGGC 6.7-GFP;
2. Digest result by *Apa*I and *Age*I

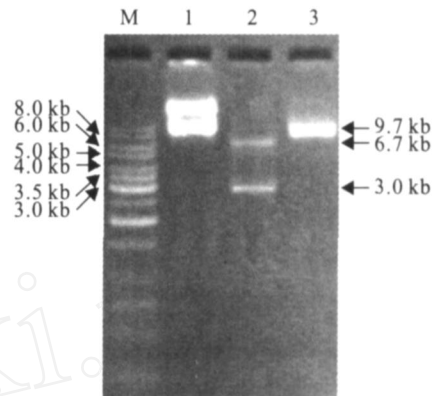


图 3 质粒 pGGC 6.7 的酶切鉴定结果

M. DNA Marker;1. 质粒 pGGC 6.7;2. *Apa*I 和 *Age*I 酶切结果;3. *Hind* III 酶切结果

Fig. 3 Identified result of pGGC 6.7 vector
M. DNA Marker;1. Plasmid of pGGC 6.7;2. Digested result by *Apa*I and *Age*I;3. Digested result by *Hind* III

2.5 pGGC 6.7-GFP 转化乳腺上皮细胞的试验结果

pGGC 6.7-GFP 转化乳腺上皮细胞后荧光显微镜观察结果见图 5。由图 5 可知,转化细胞的细胞质中有报告基因 GFP 表达。

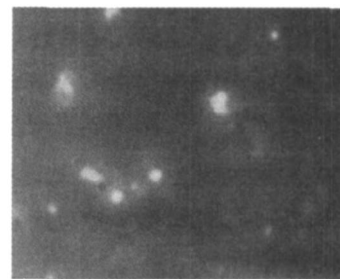


图 5 荧光显微镜检测山羊乳腺上皮细胞中 GFP 的表达

Fig. 5 GFP expression in goat mammary epithelial cell through observation under UV microscope

3 讨论

国内对山羊 β -酪蛋白基因调控区的克隆一般都

是采用直接克隆的方法,本研究由于未找到能够直接克隆 5'调控区的引物,同时考虑到应用 Long PCR 方法突变率高的因素,所以采用分段克隆的方法,分两段(GC 3.3、GC 4.3)克隆出了长度为 6.7 kb 的 5'调控区。

研究表明,乳蛋白调控区越长,被驱动基因的表达量就越高^[7],这可能是由于内含子及远端的增强子等成分对基因的表达水平有很大影响。在对相关研究^[4-5]的分析中可以看出,所构建的乳腺表达载体在拥有较长的 5'调控区的同时,拥有一个较长的 3'调控区也很重要,这可能与石统东等^[8]研究中提到的 3'调控区与 5'调控区在基因转录后翻译过程中,起着“长程”的相互作用(long range interactions)有关。目前,国内还没有克隆并采用较长 3'调控区构建载体的报道,因此,本试验还克隆了长为 6.6 kb 的 3'端调控区 GC 6.6。登陆 NCBI 网站进行 blastn 分析发现,GC 3.3、GC 4.3、GC 6.6 与山羊 CSN2 基因对应区段的同源性均为 99%,并且应用生物软件分析发现,3 个片段包含了 TATA box、GC island、CAAT box、SAR、AP1、Oct-1、Sm、Glyco、Pur-1、CAC 等复合元件和转录翻译调节因子的作用位点。结果表明,扩增出的片段正确,所以本试验成功地扩增出山羊 β 酪蛋白基因的 5'和 3'调控区。

细胞转染后的瞬时表达,常用来评价和预测体外构建的外源基因表达结构的合理性。目前,使用的标记基因,如 lacZ、CAT、luc 等均是利用酶促催化反应,需要加入外源底物和辅助因子,不仅操做繁琐而且价格昂贵^[9]。研究表明,GFP 基因可在哺乳动物细胞中表达并产生绿色荧光,而且对细胞无毒性作用^[10]。以 GFP 为报告基因构建的乳腺表达载体,可以直接通过荧光显微镜观察,是一个较好的评价方法,具有快速、简便和直观的优点。本试验通过检测报告基因 GFP 在体外培养山羊乳腺上皮细胞中的表达结果,证明克隆的 5'调控区可以指导 GFP 基因的表达,具有启动子的活性。同时利用生物软件分析的结果表明,克隆的 5'与 3'调控区可以作为

外源基因的调控序列,用于乳腺表达载体的构建。因此,可以用本试验克隆的 GC 6.7 作为 5'调控区,以 GC 6.6 作为 3'调控区,在两者之间设计合适的酶切位点,将目的基因插入其中,构建含目的基因的乳腺特异性表达载体。本研究为构建一个拥有较长 5'端和 3'端调控区的高效乳腺表达载体的研究奠定了基础。

[参考文献]

- [1] Badge L, Bygrave A E, Bradley A, et al. Transformation of embryonic stem cells with the human type II collagen gene and its expression in chimeric mice[J]. Cold Spring Harb Symp Quant Biol, 1985, 50:707-711.
- [2] Gordon K, Lee E, Vitale J, et al. Production of human tissue plasminogen activator in transgenic mouse milk[J]. Biotechnology, 1987, 5:1183-1187.
- [3] Provot C, Persuy M A, Merrier J C. Complete sequence of the ovine beta-casein-encoding gene and interspecies comparison[J]. Gene, 1995, 154(2):259-263.
- [4] Ebert K M, Ditullio P, Barry C A, et al. Induction of human tissue plasminogen activator in the mammary gland of transgenic goats[J]. Bio Technology, 1994, 12:699-702.
- [5] 黄赞, 颜景斌, 黄缨, 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[J]. 遗传学报, 2002, 29(3):206-221.
- [6] 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.
- [7] Brophy B, Smolenski G, Wheeler T, et al. Cloned transgenic cattle produce milk with higher levels of β -casein and κ -casein[J]. Nature Biotechnology, 2003, 21:157-162.
- [8] 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.
- [9] 李寿东, 齐义鹏, 胡建红, 等. 用绿色荧光蛋白基因作为筛选标记的新型克隆载体的构建[J]. 生物工程学报, 1997, 13(3):323-325.
- [10] Sheen J, Hwang S, Niway, et al. Green-fluorescent protein as a new vital marker in plant cell[J]. Plant, 1995, 8(5):777-784.