

猪血清白蛋白基因5'端调控区的克隆与分析

李 勃, 熊海燕, 陈红星, 邓继先

(中国人民解放军军事医学科学院 生物工程研究所, 北京 100071)

摘 要: 根据已报道的 *hsa* (人血清白蛋白基因) 启动子序列设计了一对引物, 以猪的基因组 DNA 为模板, 采用 PCR 扩增出一条 256 bp DNA 片段, 用 Primer Premier 5.0 和 Promoter Scan II 软件进行分析。结果表明, 此序列为 *psa* 基因 5'-端上游包括启动子在内的调控区, 该区域含有一般启动子典型的 CAA T-box、TA TA-box 等基本顺式作用元件。另外, 还含有 CTF/NF-1、CTF、APF、HNF-1、CP1 等转录因子的结合位点。

关键词: 猪血清白蛋白; 基序; 上游调控区; 启动子

中图分类号: Q 785

文献标识码: A

文章编号: 1001-6600(2009)03-0058-04

人血清白蛋白 (human serum albumin, HSA) 是一种由 586 个氨基酸组成的相对分子质量约为 68 kD 的可溶性单体蛋白^[1], 约占血清蛋白质含量的一半。它在人体内起着载体蛋白的作用, 能与多种金属离子、激素和药物结合; 同时也是维持人体胶体渗透压的主要蛋白^[2], 有增加人体循环血液量和维持血浆渗透压作用。在临床上主要用于治疗大出血、休克、战争烧伤、癌症、红白细胞增多症、肝硬化和白蛋白过少症等^[3]。目前通过转基因方法生产 HSA 为比较科学可靠的途径, 国外曾使用动物乳腺、植物和酵母表达人白蛋白, 但目前还未产业化^[4~6]。2000 年, McCreath 采用体细胞基因打靶和核移植相结合的技术路线, 制备出在羊乳分泌人 $\alpha 1$ -抗胰蛋白酶^[7]。2002 年, 美国密苏里大学赖良学等制备出敲除了猪 β 半乳糖苷转移酶的克隆猪, 使猪的体细胞膜上缺乏半乳糖苷基团, 从而移植到人体不被抗体识别而产生急性排斥反应, 可能成为人类外源性移植器官的潜在供体^[8], 也为猪源转基因产品的安全使用创造了条件。

本研究期望通过基因打靶的方法制备生产 HSA 的转基因猪^[9, 10]: 将猪血清白蛋白基因 *psa* 两翼序列作为同源臂, 加载于人血清白蛋白基因的两端, 构建猪-人同源打靶载体, 通过同源重组使 *hsa* 基因定点整合于猪染色体上原 *psa* 基因座, 置换掉 *psa*, 从而产生只表达 HSA 而不表达 PSA 的转基因猪。故首先应获得完整的猪血清白蛋白基因 (预测其全长在 19 kb 以上)。为此, 我们制备了 DNA 探针, 基于分子杂交原理, 从猪基因组 λ 噬菌体文库中克隆基因。由于 λ 噬菌体文库的平均容量约为 14 kb, 欲从文库中用探针一次克隆一个完整的白蛋白基因很是困难, 所以采取用白蛋白基因 5'端探针和 3'探针分别杂交鉴定, 然后将克隆出的两个片段再体外拼接的策略。本次实验的目标就是制备 *psa* 基因 5'端调控区并考察其作为探针的可行性, 以此为出发点分析 *psa* 基因座的可能调控机制, 为进一步研究 *hsa* 基因在猪体内的调控机制, 为基因打靶载体的构建创造条件。

1 材料和方法

1.1 菌株、载体和引物

E. coli DH10 β 和 DH5 α 感受态细胞由清华大学天为时代公司提供; pMD18-T Vector 为 Promega 公司产品; 引物 D123 和 D61 均由国家生物医学分析中心合成。

1.2 试剂、仪器及软件

QIAprep Spin Miniprep Kit (50) 购自基因有限公司; *Eco*R V 为宝生物 (大连) 有限公司产品; T4 DNA 聚合酶购自华美公司; DNA Standard Marker 购自清华大学天为时代公司, 蛋白酶 K 购自上海生工

收稿日期: 2009-04-20

基金项目: 国家“863”计划资助项目 (2002AA 206621)

通讯联系人: 邓继先 (1956—), 男, 湖北荆州人, 解放军军事医学科学院研究员, 博导。E-mail: hongxingchen@hotmail.com

生物工程公司, 其他常规药品为实验室自备。

Sigma 3K12中速冷冻离心机, Eppendorf Centrifuge 5415高速台式离心机, PCR 仪(伯乐公司产), DYY-III-6B 电泳仪(北京六一仪器厂产品)。

Primer Premier5.0、Promoter Scan II、Biodedit 软件包。

1.3 方法

1.3.1 总DNA 的提取

切取猪肝组织5 g 左右, 除结缔组织, 按照分子生物学常规方法提取该组织总DNA, 然后加30 μL ddH₂O 溶解即可。

1.3.2 5'端上游调控区的引物设计

根据 Genbank 已经发表的 *hsa* 全序列中的上游调控区序列, 基于基因序列的同源性考虑, 运用 Primer Premier5.0软件设计一对同源引物D123/D61。预测上游引物D123应该处于 *psa* DNA promoter region, 而下游的D61即 PHE-1R 应该处于猪人同源的 exon I 上。引物由本所的国家生物医学分析中心合成。其序列分别为D123-promoter: 5'-CAGATC CAGACG GCA AAC ACA CGC-3'; D61/PHE-1R: 5'-CTC GAC GAA ACA CAC CCC TGG-3'。

1.3.3 PCR 扩增5'端上游调控区

25 μL 的 PCR 反应体系中含2 μL 模板DNA (步骤1.3.1的产物猪基因组DNA 稀释100倍作为模板), D123和D61各自0.5 μL, 0.5 μL Taq DNA pol, dNTPs 共计1 μL, 2.5 μL 10×Reaction Buffer 和1.5 μL MgCl₂ solution, 其余的用 ddH₂O 补至25 μL。94℃ 变性30 s, 60℃ 退火30 s, 72℃ 延伸20 s, 共进行30个循环。循环前94℃ 预变性5 min, 循环后72℃ 继续延伸10 min, 然后4℃ 保存。所扩增的DNA 片段经1% 琼脂糖电泳检测, 用电泳凝胶回收试剂盒回收并分析其纯度。

1.3.4 5'上游调控区的分析与检测

将上步回收的目的片段克隆到 TaKaRa 公司的 pMD18-T 载体上, 按照试剂盒操作说明进行。后转化 DH5α感受态细胞, 阳性转化子质粒DNA 的提取、纯化、酶切鉴定参考一般基因工程实验手册即可。

选取有目的基因插入的阳性克隆进行测序(由国家生物医学分析中心完成), 然后利用测序结果与已发表的人 *hsa* 基因启动子区序列做序列比对分析, 探索其作为5'端探针用于杂交的可行性, 以便进一步用于从文库中筛选 *psa* 基因全长。

1.3.5 5'端上游调控区作为探针可行性的分析

运用BLAST 程序对探针与需要筛选的目标基因部分区段的匹配分析。

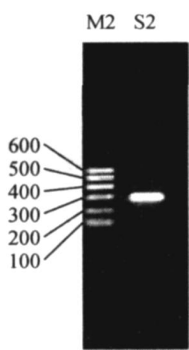


图1 *psa* 基因5'端调控区 PCR 产物的电泳结果
Fig. 1 A garose gel electrophoresis analysis of PCR products
M 2: 100 bp ladder 标准相对分子质量; S2: PCR 产物

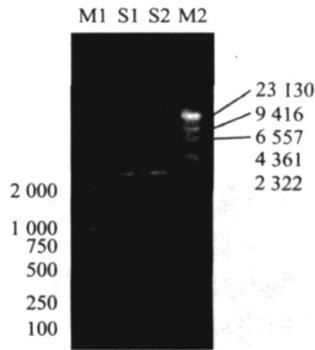


图2 阳性克隆重组子的酶切鉴定结果
Fig. 2 Restriction enzyme analysis of positive recombinant
M 1: DL 2000核酸相对分子质量标准; M 2: DNA/Hind III核酸相对分子质量标准; S1, S2: 阳性克隆酶切产物

1.3.6 5'端上游调控区转录因子潜在结合域分析

利用 Primer Premier5.0以及 Promoter Scan II 软件对测出的5'端上游调控区序列进行Motif 预测, 寻

找其启动子特征区以及潜在的反式作用因子结合调控区。

2 结果与分析

2.1 5 端上游调控区的 PCR 扩增与克隆鉴定

以猪基因组 DNA 为模板进行 PCR 扩增, 扩增产物经琼脂糖凝胶电泳检测呈唯一一条带(图1), 参照 DNA 标准相对分子质量该 DNA 序列大约为 256 bp (包括 *psa* 基因启动子区和第一外显子的部分序列)。

2.2 重组质粒的酶切鉴定

采用凝胶电泳检测质粒酶切产物, 其中 S1 和 S2 为重组质粒 pMD18-T/*psa* 经过 *EcoR* V 酶切的产物, 从上至下分别是 2 692 bp (即空载 pMD18-T 的大小)、260 bp 左右, 结果与预计的完全吻合(图2)。

2.3 5 端上游调控区的测序及其作为探针可行性的分析

根据 5 端调控区测序结果(图3), 运用 BLAST 程序对探针与需要筛选的目标基因部分区段的匹配分析, 结果显示匹配很好, 达 95%, 故作为探针是完全可行的。

表1 Primer Premier5.0对5 端上游调控区序列的预测结果

Tab 1 Prediction result of <i>psa</i> 5'-terminal regulation region			
结构名称	结构	数量	位置(距离翻译起始位点)
CAAT-box	CCAAT	1	63
TATA-box	TATA	1	119
Glyco	AA YNNAC	1	168

2.4 Motif 预测

依据测序结果, 利用 Primer Premier5.0 软件对测出的 5 端上游调控区序列进行 Motif 预测, 结果如表1所示, 其中各基序在 5 端上游调

控区的分布已在图3中标出。

ANTGNTTTA GACCAACAC ATTTTTGGCA AAAAACTAT GAATTTTGTA
ATCAGTTGTG AGCCAATGAA ATACAAAAAT GAGTCTAGTT AATAATCTAC
AATTA TTGGT TAAAGAA GATATTA GTGCT GACTTTCCTC TGTTCGTCCT
ACCTTTTCTC TTCTATCAAC CCCACAA GCC TTTGGCACAA TGAA GTGGGT
GACTTTTATT TCCCTTCTCT TTCTCTTCA G CTCTGCTTA T TCCAGGGGTG
TGTTTCCGTC GAGAAA

图3 5 端上游调控区的测序结果及其内部关键基序的分布

Fig 3 Sequencing of *psa* 5'-terminal regulation region and distribution of key motif

预测结果显示 D123~D61 区 PCR 产物存在着明显的 RNA pol II 启动子核心区域特征, 如图3所示, CAAT-box 和 TATA-box, 确定所克隆的片段就是包括 *psa* 基因启动子区在内的 5 端调控区, 可进一步用它作为探针来筛选 *psa* 基因的全序列。

2.5 扫描预测

利用 Promoter Scan II 软件对测出的 5 端上游调控区序列进行扫描预测, 结果见表2。

预测结果同时显示, 5'-端上游调控区存在着诸多顺式作用元件, 包括转录因子 CTF/NF-1、CTF、APF、HNF-1、CP1 等等结合位点(如表2), 表明这些重要的蛋白因子在猪血清白蛋白基因 *psa* 的表达进程中起到非常重要的作用, 为进一步的基因表达调控研究积累了第一手资料。

3 结 语

本实验以猪基因组 DNA 为模板, 利用 PCR 技术首次分离克隆了猪血清白蛋白基因 *psa* 5 端上游调控区, Primer Premier5.0 和 Promoter Scan II 软件分析结果表明, 该调控区除了具有普通启动子都有的 TATA-box 和 CAAT-box 等顺式作用元件外, 还具其他几种重要的转录因子结合基序, 这些转录因子主要包括 CTF/NF-1、CTF、APF、HNF-1、CP1 等, 提示这些蛋白因子在 *psa* 基因的表达调节中起到非常重要的作用, 为 *psa* 基因表达调节研究提供了基础。

表2 重要调控功能区的分析

Tab 2 Analysis of key functional motif in 5'-terminal regulation region

转录因子	DNA 链	位点	相对分子量/D
Albumin-U S2	+	43	50 000 000
CTF/NF-1	+	61	1 765 000
CTF	+	62	2 993 000
APF	+	86	50 000 000
HNF-1	-	87	10 327 000
CP1	-	111	1 388 000

真核基因表达调控体系中最关键的是转录水平的调控,它主要包括RNA聚合酶、顺式作用原件和反式作用因子三方面的相互作用。顺式作用原件必须结合特定的反式作用因子才能促使RNA聚合酶启动转录。目前在真核细胞中已经发现多种反式作用因子,它们都直接或间接地与顺式作用原件相互作用,调控着真核基因组中各个基因在特定组织活细胞中,按照一定的发育阶段有序表达^[11]。

若要通过基因打靶技术^[12]将猪染色体上的 *psa* 基因被 *hsa* 基因置换掉,首先必须对 *psa* 基因座的表达调控背景了如指掌。而真核基因表达调控元件主要分布于基因上游,故本研究对于 *psa* 基因表达调控的深入研究具有极其重要的意义。同时,利用克隆的5'端上游调控区作为探针也可以进一步通过噬菌斑原位杂交筛选 *psa* 的全序列。

参 考 文 献:

- [1] YAN Xia, NAGAKIM A, RIKURA J, et al Albumin-producing hepatocytes derived from cryopreserved F344 rats bone marrow cells transplanted in the livers of congenic Nagase's analbuminemic rats[J]. Surg Today, 2005, 35(11): 955-960.
- [2] GHIRARDELLO S, MOSCA F. Albumin versus saline: an open question[J]. Journal of Perinatology, 2008, 28(8): 586-587.
- [3] ASHMAN N, STEVEN M. H. Albumin stimulates cell growth, L-arginine transport, and metabolism to polyamines in human proximal tubular cells[J]. Kidney Int, 2005, 67: 1878-1889.
- [4] RAFALK, ANDRZEJ L, JACEK P. Influence of intralipid on free propofol fraction assayed in human serum albumin solutions and human plasma[J]. Acta Pharmacol Sin, 2006, 27(12): 1637-1641.
- [5] WANG Jun, GARY S, MANSFIELD C, et al Trans-splicing into highly abundant albumin transcripts for production of therapeutic proteins *in vivo*[J]. Mol Ther, 2008, 17: 343-351.
- [6] NIEMAN H, KUES W. A. Transgenic livestock: promises and promises[J]. Animal Reproductive Science, 2000, 60: 277.
- [7] KOLBA F, ANSELL R, MOWHR J, et al Insertion of a foreign gene into the beta-casein locus by Cre-mediated site-specific recombination[J]. Gene, 1999, 227: 21-31.
- [8] LAI Liang-xue, DONNA K S, KWANG W, et al Production of (alpha)-1, 3-galactosyltransferase knockout pigs by nuclear transfer cloning[J]. Science, 2002, 295(5557): 1089-1092.
- [9] PIEDRAHITA J A. Targeted modification of the domestic animal genome[J]. Theriogenology, 2000, 53: 105-116.
- [10] ZHANG H B, WU C C. BAC as tools for genome sequencing[J]. Plant Physiology Biochem, 2001, 39: 195-209.
- [11] MICHAEL C, STEPHEN T S. Transcriptional regulation in eukaryotes: concepts, strategies and techniques[M]. New York: Cold Spring Harbor Laboratory Press, 2000: 61-62.
- [12] GRETECHEN V. A knockout award medicine[J]. Science, 2007, 318: 178-179.

Cloning and Motif Analysis of *psa* 5'-Terminal Regulation Region from *Porcine*

LI Bo, XIONG Hai-yan, CHEN Hong-xing, DENG Ji-xian

(Institute of Bioengineering, Academy of Military Medical Sciences, Beijing 100071, China)

Abstract: A 256bp base pair of *psa* promoter fragment was amplified from *Porcine* genomic DNA by PCR method with the primers designed according to the *hsa* hpromoter of *Homo sapiens*. Primer Premier 5.0 and Promoter Scan II software analysis sequence show this fragment was 5'-terminal upstream regulation sequence of *psa* containing typical CAAT-box and TATA-box consensus and cis-acting elements. And many transcription factor binding sites such as CTF/NF-1, CTF, APF, HNF-1, CP1 etc were found in this sequence.

Key words: porcine serum albumin; motif; upstream regulation region; promoter

(责任编辑 马殷华)