

动物启动子的研究策略

冯 政, 梅书棋, 武华玉, 彭先文, 孙 华, 李良华

(湖北省农业科学院 动物胚胎工程及分子育种湖北省重点实验室, 湖北 武汉 430064)

摘要: 动物启动子是调控动物基因表达的重要元件, 开展动物启动子的研究对于阐明基因表达调控机制具有重要意义。简要介绍了动物启动子的一般结构、研究启动子的意义及其克隆和功能分析方法, 并对启动子生物信息学分析的网络资源工具进行了介绍, 最后, 对其发展前景进行了展望。

关键词: 启动子; 克隆; 功能分析

中图分类号: Q789 文献标志码: A 文章编号: 1004-3268(2012)12-0020-04

The Research Strategy of Animal Promoter

FENG Zheng, MEI Shu-qi, WU Hua-yu, PENG Xian-wen, SUN Hua, LI Liang-hua

(Key Lab of Animal Embryo Engineering and Molecular Breeding of Hubei province,
Hubei Academy of Agricultural Sciences, Wuhan 430064, China)

Abstract: Animal promoter is an important element of controlling animal gene expression. Studying the functions of animal promoters has great significance for the researches on the mechanism of gene expression regulation. The structure and methods for both of cloning and functional analysis of animal promoters are reviewed in this paper. Bioinformatics analysis tool available on network for studying animal promoters are also discussed.

Key words: promoter; cloning; function analysis

研究基因互作以及基因的表达与调控机制是基因功能研究的重要内容, 启动子作为调控基因表达的最重要元件之一日益成为研究的热点。启动子是位于基因上游的一段核苷酸序列, 它的主要作用是在基因转录时供 RNA 聚合酶识别和结合, 其长度大多在 200 bp 之内。启动子是基因转录调控的重要顺式作用元件, 它通过与 RNA 聚合酶及转录因子(一般是蛋白质)等反式作用因子相互作用来发挥调控作用。启动子具有 4 个显著特征: ①序列的保守性。启动子序列中通常含有几个结构和位置都相对保守的基序或序列盒, 其结构和位置的变化均会改变启动子的功能。②严格的方向性。作为顺式作用元件启动子是有方向性的, 如果改变它的方向则会失去启动能力。③位置特异性。启动子大多位于其所调控基因的上游, 也有一些启动子位于基因内的 5' 端, 离开这些位置或偏离太远则不能起到应有作用。④启动子具有种属及组织特异性。生物的种、属不同及同一种属生

物的不同组织, 其启动子结构都不一样, 生物亲缘关系越近, 其启动子通用的可能性也越大。从动物启动子的一般结构、研究启动子的意义及其克隆和功能分析方法等方面阐述了动物启动子的研究策略, 以期阐明基因表达调控机制提供帮助。

1 动物启动子的结构模型

动物基因的启动子属于真核生物启动子范畴, 因而其结构特征也与真核生物启动子相同。其一般特征是: 在基因的转录起点上游几十碱基处, 有一段相对保守的基序, 其有 7 对共有碱基序列: 5'-TATA (T) A(A/T)-3', 因其富含 AT 碱基对又称为 TATA 盒 (Hogness 盒)。在昆虫和哺乳动物中, TATA 盒位于 5' 端帽位点上游 20~30 bp 处, 这段序列主要决定了 RNA 聚合酶 II 与核苷酸序列结合的位置, 能影响下游基因的转录水平, 与原核生物启动子的 -10 序列有相似功能^[1]。在基因转录起点上游约 80~220 bp 还存

收稿日期: 2012-08-01

基金项目: 国家自然科学基金项目 (31001000); 湖北省农业科技创新中心创新团队项目 (2007-620-004-003)

作者简介: 冯 政 (1976-), 男, 湖北浠水人, 副研究员, 博士, 主要从事动物分子生物学与育种研究。

E-mail: greatfz@126.com

在一些保守元件,它们能起到提高启动子转录活性的作用,所以通常被称为增强子。这些保守元件可分为 2 类:一类是在动物基因启动子中普遍存在的一般上游元件,如 CCAAT 盒和 GC 盒;另一类是某些启动子所特有的特殊上游元件,它们为启动子或某些基因家族特有,如 8-珠蛋白基因家族特有的-100 盒^[2]。

2 启动子研究的意义

所有生命活动的进程都离不开基因的转录、翻译以及翻译后的修饰加工,这一系列过程是生物体内多种物质参与的有条不紊的动态平衡,从而保证整个生命体的正常运转。只要任何一个环节发生紊乱或存在缺陷,都会打乱生命体的平衡状态,给其造成不良后果。因此,生命科学研究者最核心的工作是要明晰调控基因表达的机制。生命历程的每个环节都存在着严密调控基因精确表达的机制,尤其是转录环节的调控对基因参与的生命活动进程最为关键。启动子发挥它的调控作用主要是在基因的转录环节,因而开展启动子功能的研究对于了解或阐明基因表达的调控机制有着很重要的意义。

3 启动子克隆的方法

目前研究者一般利用启动子探针和 PCR 技术来筛选和克隆启动子序列。

3.1 利用质粒探针筛选启动子

质粒探针筛选启动子的基本原理是将随机 DNA 片段连接到缺失启动子的报告基因载体上,通过检测报告基因的活性来筛选有启动功能的 DNA 片段。其程序是:先选择适当的核酸限制性内切酶,将染色体 DNA 随机打断;再将切割产生的这些随机片段与缺失启动子的质粒报告基因载体重组,使 DNA 片段插入报告基因上游的合适位置;最后把报告基因载体转化到寄主细胞中,构建报告基因载体文库,并检测各个报告基因的表达活性,从而筛选具有启动活性的 DNA 片段^[3]。

Rachael 等^[4]以四环素抗性基因为报告基因,在大肠杆菌中构建了启动子探针质粒 pBRH3B,通过检测四环素的抗性筛选到一些启动子片段。用启动子探针方法克隆筛选启动子的优点是不需要知道基因的序列,免去了引物设计,可随机筛选到大量的启动子片段。其缺点是需要构建报告基因载体、建立质粒文库、转化细胞、检测报告基因活性、筛选阳性质粒,工作量大不说,还费时费力费钱,针对性不强。

3.2 利用 PCR 技术克隆启动子

随着 PCR 技术的飞速发展,这种简便高效的方

法被越来越多地用来从基因组中分离各种已知和未知序列。PCR 方法是根据已知基因序列设计引物克隆启动子,因其简便快捷,逐渐成为克隆启动子的首选方法。贾向志等^[5]通过巢式 PCR 方法,从人的基因组总 DNA 中分离出 CCRK 基因 5'端非翻译区大小为 1 072 bp 的基因片段。分离后经过测序及与收录序列比对,同源率达 98%。山松等^[6]以烟草叶绿体 16S rRNA 基因上游序列为模板设计引物,以烟草 DNA 为模板进行 PCR 扩增,成功扩增到 16S rRNA 基因 5'端的 150 bp 启动子区片段,并将该片段克隆和测序,与基因库进行同源性比较,确定克隆到的片段是烟草叶绿体 16S rRNA 基因启动子。

除了可以根据已知基因序列设计引物进行 PCR 扩增克隆启动子之外,运用反向 PCR 方法还能对已知序列旁的未知 DNA 序列进行扩增,该方法也被称为染色体步移(chromosome walking)^[7]。其一般过程如下:选择适当的限制性核酸内切酶(必须在该序列内部不存在切点)对一段 DNA 进行酶切,然后用 DNA 连接酶将酶切后的含已知序列的片段连接形成环状 DNA 分子;设计与该已知序列两端互补的引物,使引物延伸方向向该序列以外的两端延伸,这样以环状 DNA 分子为模板进行 PCR 扩增,得到的就是已知序列的旁侧 DNA 片段,包括上游和下游核苷酸序列。这种方法的效率与酶切后 DNA 分子的连接环化的效率和准确性直接相关,由于 DNA 在连接环化过程中容易产生多联体,以及以环状 DNA 双链分子为模板的 PCR 扩增效率差,使得该方法成功的概率较低。许多研究者将反向 PCR 方法进行一些细节改进,获得了较满意的效果。如李竹红等^[8]先后设计了 2 对引物进行巢式 PCR 扩增,并在反应体系中加入 5%的去离子甲酰胺,大大提高了 PCR 扩增的特异性,利用这种改进的反向 PCR 方法成功地克隆了 Tc1 转座子的旁侧序列。

现在,各种经过改进的 PCR 方法越来越多地用于克隆启动子。如黄君健等^[9]利用 Panhandle PCR 方法扩增人端粒酶催化亚基 *hTERT* 基因 5'端上游序列,结果成功克隆出 *hTERT* 基因起始密码子上游长达 2 090 bp 的 DNA 序列。王新国等^[10]采用衔接头 PCR 方法,将胡萝卜基因组 DNA 酶切,再将一特殊的衔接头与将该酶切片段连接,以此连接片段为模板进行 PCR 扩增,成功克隆到一个新的胡萝卜 SII 基因启动子片段。其他比较常见的改进 PCR 方法还有 UFW(universal fast walking,通用快速步移)^[11]、TAIL-PCR(thermal asymmetric interlaced PCR,热不对称交错 PCR)^[12]、panhandle PCR 等^[13],它们都不对 DNA 进行连接环化而直接进行 PCR 扩增,因而扩增

的特异性和成功率都较反向 PCR 方法高。

随着分子生物学技术的发展,以及生物试剂和生物辅助材料的不断进步,一定会出现更多、更好的克隆启动子的方法。

4 启动子功能分析方法

4.1 启动子删除片段分析

克隆到基因上游序列后,一般运用启动子删除片段分析方法来获得核心启动子。其过程是:以克隆到的基因上游序列为模板,设计几条上游引物和一条共用下游引物,上下游引物分别加上在报告基因质粒载体上选择的合适酶切位点和保护碱基,组合扩增多个片段。将扩增到的多个片段回收纯化后克隆提取质粒,再用前述 2 种酶双酶切克隆质粒,纯化回收双酶切后的多个启动子删除片段,分别与进行了同样双酶切并回收纯化的 pGL3 载体连接,用常规方法克隆构建多个启动子删除片段报告基因载体,将载体转化到细胞,检测不同删除片段的报告基因表达活性,从而确定核心启动子区域^[14]。

4.2 点突变分析

得到核心启动子区域后,可以利用转座子插入突变,也可以用重叠延伸 PCR 的方法针对转录因子结合位点中几个最重要的碱基做点突变,或者用某种限制性酶切去启动子中某重要元件来制造启动子的突变,最后把突变后的启动子连接报告基因载体转化细胞,以野生的启动子序列作阳性对照,根据报告基因活性变化来确定启动子的核心序列功能。

4.3 启动子与转录因子的结合验证

启动子是基因上游序列,和转录因子结合的那部分序列一般被称为顺式作用元件,长约 10~20 bp,相应的,转录因子常被称为反式作用因子。在点突变分析证实启动子序列上某顺式作用元件对其功能有影响之后,还要通过实验来验证转录因子和该顺式作用元件在体内外的结合能力。

凝胶迁移试验(electrophoretic mobility shift assay, EMSA)是一种研究 DNA 与蛋白质结合的常用试验技术,其原理是 DNA/蛋白质复合物与单独的 DNA 或蛋白质在聚丙烯酰胺凝胶电泳(PAGE)中迁移率不同,被用来分析启动子与转录因子在体外的相互作用^[15]。启动子是一段 DNA 序列,在适当的条件下与转录因子结合后,结合在一起的 DNA/蛋白复合物在 6% 的非变性聚丙烯酰胺凝胶中的迁移率小于未能结合上转录因子的 DNA,从而可以检测到与启动子结合的转录因子。转录因子既可以用体外表达纯化的蛋白,也可以直接从细胞核抽提蛋白质,从而证实体外与该启动子有特异性结合的转录因子。

EMSA 试验能够证实启动子与转录因子在体外的特异性结合作用,但不能证明它们在体内是否同样有这种作用,因为体外与体内的环境有很大的差别。因此,还必须研究在模拟体内生理条件下转录因子与启动子的特异性结合能力。染色质免疫沉淀技术(chromatin immunoprecipitation, CHIP)就是研究体内环境下蛋白质与 DNA 相互作用的一项生物学试验技术^[16]。其原理是:在生理状态下让启动子序列与细胞内的蛋白质交联,用试剂停止交联反应后分离和裂解细胞核,再通过超声波破碎染色质,使染色质随机打断为一定长度的染色质片段,用转录因子的特异性抗体与破碎的染色质进行免疫共沉淀反应,最后将沉淀的 DNA 与蛋白质解除偶联,洗脱并纯化目的 DNA 片段进行克隆测序。该技术在细胞生理状态下进行,利用抗原抗体反应的特异性来检测结合在蛋白质上的 DNA,能较真实地检测体内环境中 DNA 与目的蛋白质的特异性结合。

转录因子和启动子结合后不一定就会影响该启动子所调控基因的表达。很多时候,很多转录因子一起形成复合体才有调控作用,于是要在过表达和干扰该转录因子的条件下,运用 RT-PCR 技术来检测目的基因的表达变化,这样才能充分说明转录因子和启动子的结合确实能够调控目的基因的表达。过表达可以用过表达载体或增强型启动子等来加大转录因子的表达变化,而干扰表达可以用转录因子的拮抗物或阻断减弱其表达的手段达到降低转录因子表达量的目的。

5 启动子的生物信息学分析

启动子的生物信息学分析主要包括启动子序列与结构的预测和转录结合位点及其结合因子的预测,早在 1997 年, Fickett 等^[17]就对前人在启动子预测方面的工作做了详细的介绍。随着不同物种基因组计划的相继实施和完成,生物数据库中收录的基因组序列的信息量越来越丰富。生物信息学研究者通过分析大量启动子和转录因子序列,总结和研制出各种算法和软件来预测潜在的启动子。到目前为止,互联网上关于启动子预测和分析的数据库及软件资源非常多,其中最大的真核基因启动子数据库是欧洲研究人员建立的 EPD 数据库(<http://epd.vital-it.ch/>),该数据库对启动子进行了详细的注释,其中所有转录起始位点都经过严格的试验鉴定。

生物信息学技术以及计算机软硬件和网络的飞速发展,使得启动子的预测分析可以在全世界范围内共享,利用大型服务器作为运算终端,以互联网作为服务通道,大量的启动子预测与分析软件 and 数据库被开发出来,并公布到网上,为科研人员从生物信

息学角度研究启动子提供服务。笔者搜集了主要的真核生物启动子分析网络资源,具体见表 1。

表 1 启动子分析网络资源

资源名称	网址	主要功能
FirstEF	http://rulai.cshl.org/tools/FirstEF/	人的启动子预测
BDGP	http://www.fruitfly.org/seq_tools/promoter.html	果蝇的启动子预测
BIMAS	http://www-bimas.cit.nih.gov/molbio/proscan/	预测真核生物的启动子
CONSITE	http://asp.ii.uib.no:8090/cgi-bin/CONSITE/consite	预测转录因子结合位点
Promoter 2.0	http://www.cbs.dtu.dk/services/Promoter/	预测脊椎动物转录起始位点
PromoSer	http://biowulf.bu.edu/zlab/PromoSer/	人、鼠的启动子预测
TRES	http://users.ugent.be/~efouquae/artikels/augustus%202004/artikels%20promoter%20juli%202004/TRES%20A%20Tool%20for%20Comparative%20Promoter%20Sequence%20Analysis.htm	转录调节因子分析
gene-regulation	http://www.gene-regulation.com/pub/databases.html	真核生物转录因子结合位点分析
TESS	http://www.cbil.upenn.edu/cgi-bin/tess/tess	DNA 序列的转录因子结合位点分析
TRRD	http://www.mgs.bionet.nsc.ru/mgs/gnw/trrd/	真核生物转录调节区域分析
TRANSFAC	http://www.biobase-international.com/product/transcription-factor-binding-sites	真核生物转录因子结合位点分析
TFSEARCH	http://www.cbrc.jp/research/db/TFSEARCH.html	DNA 序列的转录因子结合位点分析
MatInspector	http://www.genomatix.de/solutions/dl_matinspector.html	转录因子结合位点分析

6 展望

启动子是基因转录环节调控的重要顺式作用元件,人们已经利用启动子来改变微生物和植物基因的表达方式,从而得到自己需要的生物工程产品。然而,动物基因的启动子大多是组成型的,较难对其进行干扰从而干预动物基因的表达。目前,启动子的克隆和分离技术手段越来越丰富,启动子的功能分析研究也由于生物信息学的发展取得了一定进展。但由于启动子复杂多样,研究者对生物的转录调控机制认识还不完善,启动子的研究和应用存在很多困难。随着各种生命科学技术的发展和生物信息资源的丰富,动物启动子的序列信息和功能会逐渐明确,启动子的研究和应用也将会步入更快更好的发展轨道。

参考文献:

[1] 朱玉贤,李毅. 现代分子生物学[M]. 3 版. 北京:高等教育出版社,2002.

[2] 吴乃虎. 基因工程原理[M]. 2 版. 北京:科学出版社,2006.

[3] Mervyn J B, Stanley N C. Gene expression in *Streptomyces*; Construction and application of promoter-probe plasmid vectors in *Streptomyces lividans* [J]. Molecular and General Genetics, 1982, 187(2): 265-277.

[4] Rachael L N, Robert W. Eukaryotic DNA fragments which act as promoters for a plasmid gene[J]. Nature, 1979, 277: 324-325.

[5] 贾向志, 林李家宓, 何明亮, 等. 人细胞周期相关激酶启动子的克隆和初步分析[J]. 生物技术通讯, 2004, 15(3): 222-225.

[6] 山松, 张中林, 陈曦, 等. 烟草叶绿体 16S 启动子的克隆

改造及转基因植株的获得[J]. 作物学报, 1999, 25(5): 536-540.

[7] Tony Triglia, Inverse PCR (IPCR) for obtaining promoter sequence [J]. Methods in Molecular Biology, 2000, 130: 79-84.

[8] 李竹红, 刘德培, 梁植权. 改进的反向 PCR 技术克隆转移基因的旁侧序列[J]. 生物化学与生物物理进展, 1999, 26(6): 600-602.

[9] 黄君健, 李杰之, 林坚, 等. 人端粒酶催化亚基 hTERT 基因启动子的克隆[J]. 生物技术通讯, 1999, 10(3): 167-170.

[10] 王新国, 肖成祖, 张国华, 等. 用衔接头 PCR 克隆新的胡萝卜 II 型转化酶基因启动子[J]. 中国生物化学和分子生物学报, 2001, 17(1): 61-65.

[11] Myrick K V, Gelbart W M. Universal fast walking for direct and versatile determination of flanking sequence [J]. Gene, 2002: 125-131.

[12] Liu Y G, Robert F. Thermal asymmetric interlace PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking[J]. Genomics, 1995, 25(3): 674-681.

[13] Jones D H, Winistorfer S C. Sequence specific generation of a DNA panhandle permits PCR amplification of unknown flanking DNA [J]. Nucleic Acids Res, 1992, 20(3): 595-600.

[14] 冯政. 猪骨骼肌发育相关新基因家族 BTG/TOB 的克隆和功能研究[D]. 武汉: 华中农业大学, 2007.

[15] Lawrence D K. Electrophoretic mobility shift assay [J]. Methods in Enzymology, 1995, 254: 619-632.

[16] Kuo M H, Allis C D. *In vivo* cross-linking and immunoprecipitation for studying dynamic protein: DNA associations in a chromatin environment[J]. Methods, 1999, 19(3): 425-433.

[17] Fickett J W, Hatzigeorgiou A G. Eukaryotic promoter recognition[J]. Genom Res, 1997, 7: 861-878.