

外显子捕获法分离 FRAXA 位点的新表达序列*

孔建 朱宁宁 尹斌 沈岩¹⁾ 吴冠芸

(中国医学科学院基础医学研究所 医学分子生物学国家重点实验室, 北京 100005)
(中国协和医科大学基础医学院)

摘要 外显子捕获是近年发展起来的一种自基因组 DNA 分离表达序列的有效方法. 我们采用以 pSPL3 为剪接载体的外显子捕获系统, 从覆盖 FRAXA 位点的酵母人工染色体 YAC209G4 中分离到 2 个新外显子序列 A91 和 D12. A91 在人胎肝和骨骼肌文库中丰度较高, 在肝、肾和骨髓文库中也有 PCR 扩增. 而 D12 在所分析的 8 个 cDNA 文库中均未见 PCR 扩增产物. 从骨骼肌文库中克隆到一段包含 A91 的 315 bp 的 cDNA (AM4470). 多种组织 RNA 印迹显示, AM4470 与骨骼肌和心脏 mRNA 杂交, 转录本长度均为 2.8 kb.

关键词 外显子捕获, 表达序列, FRAXA 位点

学科分类号 Q78

从基因组 DNA 中分离编码序列的方法很多. 目前应用最多的主要是外显子捕获 (exon trapping) 和 cDNA 直接筛选. cDNA 筛选需预先了解目的基因表达的组织及发育阶段特异性, 以便选择适宜的 cDNA 文库. 外显子捕获通过真核细胞 RNA 剪接系统对剪接位点的识别, 直接从基因组 DNA 中分离外显子序列. 外显子捕获方法克隆基因只决定于核苷酸序列中剪接位点的存在而无需预知目的基因表达的时空特异性. 这种方法已成功地应用于致病基因, 如 Huntington 基因的定位克隆^[1], 以及转录图谱的构建^[2]等研究.

脆性 X 综合征 (fragile X syndrome, Fra (X)) 是最常见的遗传性智力低下疾病. 1991 年在 X 染色体长臂末端的 FRAXA 脆性位点克隆了 Fra (X) 的致病基因——FMR1^[3]. 一些现象提示, 在 FMR1 基因上游几百 kb 范围内可能还存在未知基因, 其失活也与脆性 X 综合征有关^[4]. 我们采用外显子捕获方法从包含 FMR1 基因及其上游序列的酵母人工染色体 YAC209G4 分离 FRAXA 位点的新外显子, 得到两个潜在的新外显子 A91 和 D12, 并通过 PCR 和 RNA 印迹杂交证实了 A91 在骨骼肌等组织的表达.

1 材料和方法

1.1 重组剪接质粒的构建

以 pSPL3^[5] 为剪接载体的外显子捕获系统工作原理见图 1. 将外源基因组 DNA 片段重组到 pSPL3 载体的 HIV tat 基因内含子中, 转染 COS-7 细胞, 由 SV40 早期启动子 (SV40 early promoter) 引导转录. 当插入片段中包含顺向的完整外显子时, 转录后, 其两端的剪接受位 (SAg) 与供位 (SDg) 会分别与载体内含子两端的剪接供位 (SDv) 与受位 (SAv) 配对并发生剪接. 通过 SAv 与 SDv 两侧特异引物逆转录 PCR 扩增, 可得到目的基因的外显子序列. 该载体多克隆位点两侧各有一个 *Bst* XI 的识别位点 (CCAN₆TGG), 剪接位点边界上各有一 *Bst* XI 的半位点, 用以消除载体自我剪接造成的本底.

取 20 μ g 环状 pSPL3 质粒 DNA (GIBCO-BRL 公司外显子捕获试剂盒), 以 *Bam*HI 线性化、牛小肠碱性磷酸酶 (CIP) 去除 5' 磷酸

* 国家自然科学基金 (39392900) 和国家科委攀登计划“人类疾病基因控制工程的基础研究”项目.

¹⁾ 通讯联系人.

收稿日期: 1997-06-13, 修回日期: 1997-09-10

基团. 1% 低熔点琼脂糖凝胶脉冲场电泳分离 YAC209G4 菌株的染色体 DNA (CHEF Mapper, BioRad; 分离条件: 电压 220 V, 脉冲角度 120° , 脉冲时间变化范围 35.40~44.69 s), 回收 475 kb YAC DNA 片段, β -Agarase I (Biolabs) 消化琼脂糖凝胶. 将样品均分为 7 份, 各加入 0.2 U *Sau*3A I, 于 37°C 分别温育 0 min、5 min、10 min、15 min、20 min、

30 min 和 45 min. 合并各份酶解产物, 抽提、沉淀 DNA, 重新溶解于 TE 缓冲液中. 于 1% 低熔点琼脂糖凝胶电泳回收 2~10 kb DNA 片段. 取 250 ng pSPL3 载体与插入片段重组, 电转化 (Gene Pulser, Bio-Rad 公司) *E. coli* XL1-Blue 受体菌. 转化条件: 电压 2.5 kV、电容 25 μF 、电阻 200 Ω . 碱裂解法提取转化子的混合质粒 DNA^[6].

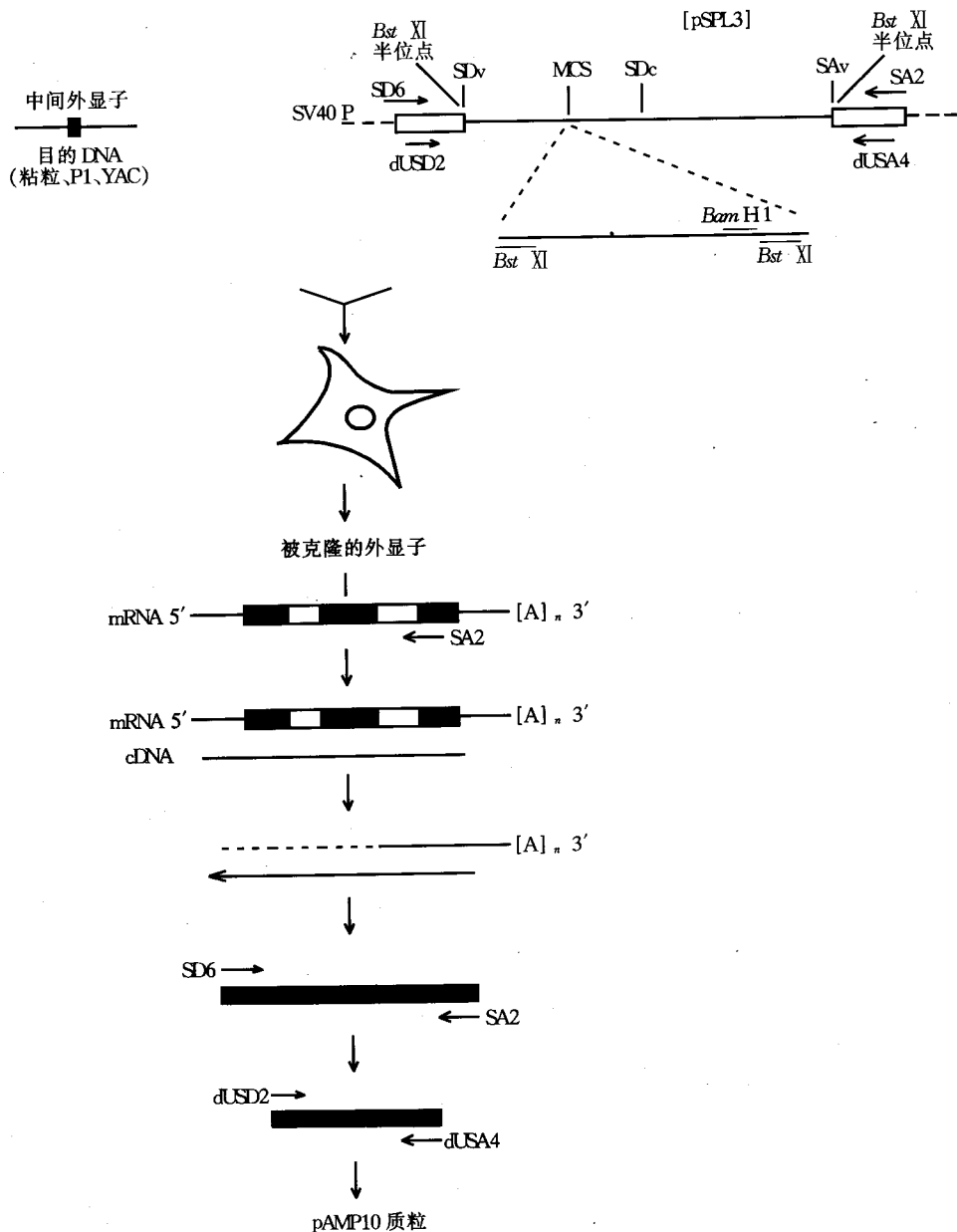


图1 pSPL3 系统工作示意图

P: 启动子, SDv: 载体的剪接供应, SDc: 隐蔽的剪接供应, SA1: 载体的剪接受位, MCS: 多克隆位点.

1.2 转染 COS 7 细胞

取 3 μg 重组质粒, 以 LIPOFECTAMINE (GIBCO-BRL) 介导转染过夜培养、生长至约覆盖瓶壁 80% 的 COS-7 细胞. 37 $^{\circ}\text{C}$, 5% CO_2 , 湿润条件下培养.

1.3 剪接产物的分析

1.3.1 细胞总 RNA 的提取与 RT-PCR: 转染 48 h 后按一步法提取 COS-7 细胞的总 RNA^[7]. 取 3 μg 总 RNA, 在 20 μl 体系中加入 20 pmol SA2 引物 (本工作所用引物均为 GIBCO-BRL 公司外显子捕获试剂盒提供), 200 U SuperScript RT II 逆转录酶 (GIBCO-BRL), 42 $^{\circ}\text{C}$ 温育 30 min 逆转录 cDNA 第一条链. 55 $^{\circ}\text{C}$ 温育 5 min 后加入 1 μl RNase H, 继续温育 10 min.

在 40 μl 体系中加入 8 μl 逆转录产物, 引物 SA2、SD6 各 20 pmol, 1 U Taq DNA 聚合酶 (GIBCO-BRL), 94 $^{\circ}\text{C}$ 1 min、60 $^{\circ}\text{C}$ 1 min、72 $^{\circ}\text{C}$ 5 min, 扩增 6 个循环, 72 $^{\circ}\text{C}$ 延伸 10 min, 然后将温度降至 55 $^{\circ}\text{C}$. 取出 35 μl , 用 40 U *Bst* XI 于 55 $^{\circ}\text{C}$ 温育过夜. 再加入 4 U *Bst* XI, 继续温育 2 h.

在 50 μl 体系中加入 5 μl *Bst* XI 酶解产物, 引物 duSA4、duSD2 各 20 pmol, 1 U Taq DNA 聚合酶, 94 $^{\circ}\text{C}$ 1 min、60 $^{\circ}\text{C}$ 1 min、72 $^{\circ}\text{C}$ 3 min, 扩增 30 个循环, 72 $^{\circ}\text{C}$ 延伸 10 min, 产物保存于 4 $^{\circ}\text{C}$.

1.3.2 RT-PCR 产物的克隆与分析: 在 10 μl 体系中加入 1 μl PCR 产物, 50 ng 非连接酶依赖的 pAMP10 尿嘧啶 DNA 糖基化酶 (UDG) 克隆载体, 1 单位 UDG (GIBCO-BRL), 37 $^{\circ}\text{C}$ 连接 30 min. 取 1 μl 连接产物, 电转化 *E. coli* XL-1-Blue 受体菌. 挑取平板上的单菌落作为模板, 以引物 SA4、SD2 进行 PCR 扩增 (94 $^{\circ}\text{C}$ 1 min、60 $^{\circ}\text{C}$ 1 min、72 $^{\circ}\text{C}$ 1 min, 30 个循环), 筛选阳性克隆.

碱裂解法提取重组质粒. 以 γ - ^{32}P -ATP 标记 SD2 引物, 采用双链 DNA 循环测序法测定阳性克隆的 DNA 序列 (dsDNA 循环测序试剂盒, GIBCO-BRL 公司). 测序反应物以含

7 mol/L 尿素的 6% 聚丙烯酰胺凝胶电泳分离. -70 $^{\circ}\text{C}$ 放射自显影 12~20 h.

根据被捕获的外显子序列设计引物 (表 1 下加线序列), 分别以重组质粒、人基因组 DNA、YAC209G4 DNA、YAC209B6 (与 Xq27.3 染色体位点无关) DNA 为模板进行 PCR, 确认其来源. 分别以人胎脑 (本所生化室袁建刚教授惠赠)、人胎肝 (上海第二医科大学陈竺教授惠赠) 及成人脑、肝、肾、骨髓、骨骼肌和睾丸 (上海复旦大学遗传所余龙教授惠赠) 来源的 8 个 cDNA 文库储存液为模板进行 PCR, 检测被捕获的序列是否在这些组织中表达.

1.4 从人骨骼肌 cDNA 文库中克隆含新外显子序列的 cDNA

以人骨骼肌 cDNA 文库储存液作为模板, 以新外显子 A91 引物与文库载体 $\lambda\text{gt}11$ 引物交叉配对 (即 A91F- $\lambda\text{gt}11\text{F}$ 、A91F- $\lambda\text{gt}11\text{R}$ 、A91R- $\lambda\text{gt}11\text{F}$ 、A91R- $\lambda\text{gt}11\text{R}$ 4 组引物对) 进行 PCR 扩增: 94 $^{\circ}\text{C}$ 50 s、60 $^{\circ}\text{C}$ 50 s、72 $^{\circ}\text{C}$ 90 s, 30 个循环. PCR 产物经 2% 低熔点琼脂糖凝胶电泳分离、纯化, 克隆到 pGEM-T 载体. 菌落 PCR 法筛选阳性克隆. 以 T7 或 SP6 为引物, 采用双链 DNA 循环测序法对阳性克隆进行序列测定.

1.5 RNA 印迹实验

PCR 标记阳性克隆 AM4470 质粒的插入片段 (1.85×10^6 Bq α - ^{32}P -dCTP, 6 个循环), 与人多种组织 RNA 印迹膜 (Human MTN Blot, Cat. 7760-1, 1 of 2, ClonTech) 在 $5 \times \text{SSPE}$, $10 \times \text{Denhardt's}$ 溶液, 100 mg/L 变性鲑精 DNA, 2% SDS 溶液中 42 $^{\circ}\text{C}$ 杂交过夜. $2 \times \text{SSPE}$ -0.1% SDS 室温漂洗 10 min 3 次, $0.3 \times \text{SSPE}$ -0.1% SDS 50 $^{\circ}\text{C}$ 漂洗 30 min. -70 $^{\circ}\text{C}$ 放射自显影.

2 结 果

2.1 FRAXA 位点新外显子的捕获和鉴定

将 2~10 kb 的 YAC DNA/Sau3A I 酶解片段与剪接载体 pSPL3 连接, 转化 XL1-blue

菌, 得到约 8 000 个转化子. 提取总质粒 DNA 转染 COS-7 细胞. 转染 48 h 后提取 COS-7 细胞总 RNA, 通过 RT-PCR 扩增经 COS-7 剪接的转录产物. PCR 产物与 pAMP10 载体重组, 得到约 1×10^4 转化子. 随机挑取转化子用载体引物进行 PCR 扩增. 载体自身

剪接产物经 PCR 扩增得到 153 bp 的 DNA 片段. 插入被捕获的外显子后, 扩增产物长度增加. 挑选扩增片段大于 153 bp 的克隆, 进行 DNA 序列分析. 通过与基因库 (Genbank) 中已知序列的同源比较, 发现了 2 个新的潜在外显子序列: A91 和 D12. 其核苷酸序列见表 1.

表 1 新外显子 A91 和 D12 的序列

外显子名称	DNA 序列
A91	AGAGCTTGGCGGTAATCATGGTCATAGCTGTTTCTGTGTGAAATTGTTATCCGCTCACAATT CCACACAACATACGAGCCGGAAGCATAAAGTGTA
D12	GGCTGCAGTTCAGGGGGCAGGCCGACTTGCTAAGCAGAAAATAGCCCATCTCTTCCACAA CATCAC

注: 下划线为进一步研究时所用的引物序列.

为了验证 A91 和 D12 的来源, 分别以人基因组 DNA、YAC209G4 总 DNA、YAC209B6 总 DNA 为模板, 以 A91 和 D12 各自的特异引物进行 PCR 扩增. 两个片段在

YAC209G4 和人基因组 DNA 中均有扩增, 在与 FRAXA 区域无关的 YAC209B6 中无扩增, 说明 A91 和 D12 确实来源于 209G4 的 YAC DNA 片段. 结果如图 2 所示.

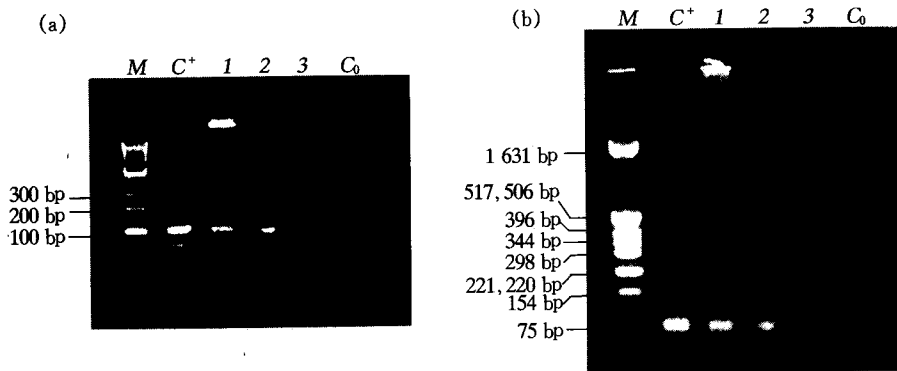


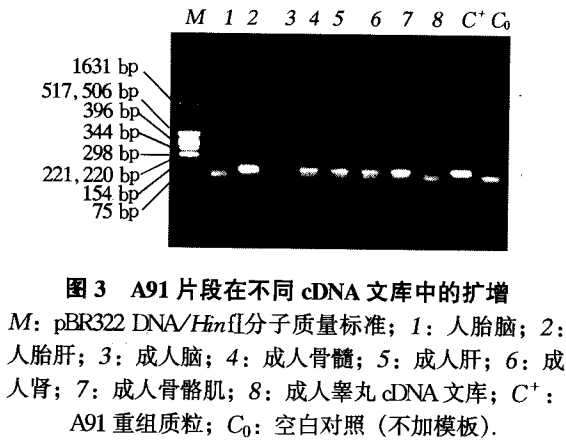
图 2 新外显子 A91 和 D12 来源的鉴定

(a) A91 的扩增结果; (b) D12 的扩增结果; M: 分子质量标准 (100 bp DNA ladder (a), pBR322/*Hinf* I (b)); C+: 阳性对照 (在重组质粒中的扩增); C0: 空白对照 (不加模板); 1: 人基因组 DNA; 2: YAC209G4 DNA; 3: YAC209B6 DNA.

以人胎脑、人胎肝及成人脑、肝、肾、骨髓、骨骼肌和睾丸等 8 个 cDNA 文库储液为模板的 PCR 扩增显示, A91 片段在胎肝和骨骼肌两个文库中扩增较强, 骨髓、肝、肾中也有扩增 (图 3). D12 在这 8 个文库中均无扩增产物 (结果未显示).

2.2 含 A91 序列的 cDNA 片段的分离

选择人骨骼肌文库作为筛选 cDNA 克隆的来源. 将 A91F/R 与 λ gt11F/R 交叉配对, 以人骨骼肌 cDNA 文库储液为模板进行 PCR 扩增, 回收特异扩增带, 克隆并进行序列分析. 经过分析比较, 得到如下两组重叠的含有 A91



全序列的克隆——AM4241 和 AM4470. 二者为来自同一转录本的两个不同片段, 其相互关系及序列见图4.

2.3 多种组织 RNA 印迹实验

以 AM4470 克隆的插入片段作探针, 与入多种组织的 mRNA 进行印迹杂交, 在骨骼肌和心脏组织得到约 2.8 kb 的杂交区带, 提示 A91 所在基因的转录本长度约为 2.8 kb (图5).

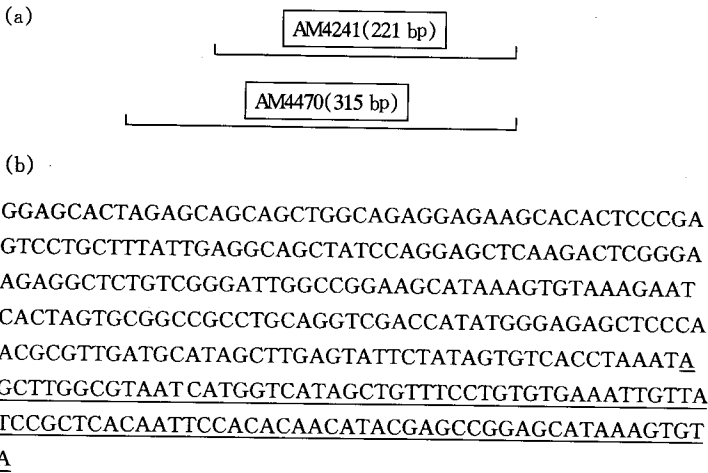


图4 含新外显子 A91 的 cDNA 序列

(a) AM4241 和 AM4470 是两个重叠的 cDNA 片段, 3' 部分为 A91 序列; (b) AM4470 的序列 (划线部分为 A91 序列)。

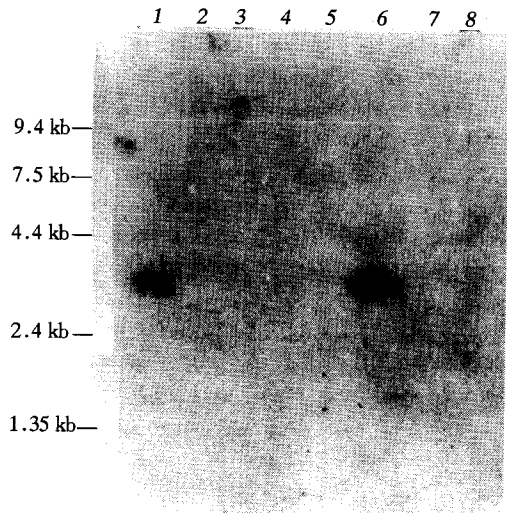


图5 AM4470 克隆与人多种组织 mRNA 的 RNA 印迹实验

1: 心脏; 2: 脑; 3: 胎盘; 4: 肺; 5: 肝; 6: 骨骼肌; 7: 肾; 8: 胰腺。

3 讨 论

采用外显子捕获方法定位克隆基因的最大特点是可以在不了解目的基因表达时空特异性的情况下寻找其表达序列. 外显子捕获的载体系统主要有质粒载体^[5, 8-11]、噬菌体载体^[12]和粘粒载体^[13]三大类. 质粒载体多用于染色体局部位点的研究. 噬菌体载体和粘粒载体更适用于染色体乃至整个基因组范围内的大规模筛选. 我们的研究对象是 X 染色体一个局部区域, 所以使用了质粒载体 pSPL3. 为了提高插入片段的代表性, 我们选择了识别四碱基的 *Sau*3AI (*Bam*H I 的同尾酶) 部分酶解 YAC DNA, 回收 2~10 kb 的片段构建剪接重组体. 取代了以 *Bam*H I 和 *Bgl* II 两个识别六碱基的

同尾酶完全酶解制备插入片段^[5].

由于前体 mRNA 的剪接不仅与剪接位点自身的序列保守性有关, 还受到相邻序列的影响^[14]. 此外, 插入质粒载体的基因组 DNA 片段长度有限, 某些天然前体 mRNA 的序列并不包含在重组体的转录产物中, 使得外显子捕获假阳性或假阴性的现象均有可能发生. 因此对于捕获到的“外显子”应做进一步的检验.

本工作捕获到的两个新外显子 A91 和 D12 在 8 个不同组织或发育阶段来源的 cDNA 文库中丰度不同. A91 在骨骼肌、胎肝和成人骨髓、肝、肾等组织中都可检出, 在骨骼肌和胎肝文库中 PCR 扩增信号较强. 而 D12 在 8 个文库中均未能用 PCR 方法检出. 原因可能是 D12 不在这 8 个文库所代表的组织和发育阶段表达, 或是这些 cDNA 文库经过了数轮扩增, 代表性发生偏差. 另外, 不排除 D12 并非体内天然剪接产物的可能. 因为 D12 的 5' 部分存在嘌呤富含区, 有可能作为外显子识别序列 (exon recognition sequence, ERS)^[15] 增加了在其上游发生剪接的机会, 从而产生假阳性.

本工作采用了从 cDNA 文库中 PCR 扩增含 A91 序列的 cDNA 片段的技术路线^[16]. 由于所用人骨骼肌 cDNA 文库的载体 λ gt11 为非定向克隆载体, 其外源 cDNA 片段可以任意方向插入, 故在做 A91 特异引物与 λ gt11 引物的配对扩增时考虑了 4 种可能. 以 A91 的顺、反向引物分别与 λ gt11 的顺、反向引物配对进行扩增, 得到 315 bp 的 cDNA 片段 (克隆 AM4470). RNA 印迹进一步证实该基因在骨骼肌和心脏中有表达, 转录本长度约 2.8 kb.

参 考 文 献

- 1 The Huntington's Disease Collaborative Research Group. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell*, 1993, **72** (6): 971~ 983
- 2 Koyama K, Sudo K, Nakamura Y. Isolation of 115 human chromosome 8-specific expressed sequence tags by exon amplification. *Genomics*, 1995, **26** (2): 245~ 253
- 3 Verkerk A J M H, Pieretti M, Sutcliffe J S *et al.* Identification of a gene (FMR1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell*, 1991, **65** (5): 905~ 914
- 4 Hansen R S, Canfield T K, Lamb M M *et al.* Association of fragile X syndrome with delayed replication of the FMR1 gene. *Cell*, 1993, **73** (7): 1403~ 1409
- 5 Church D M, Stotler C J, Rutter J L *et al.* Isolation of a gene from complex sources of mammalian genome DNA using exon amplification. *Nature Genet*, 1994, **6** (1): 98 ~ 105
- 6 Sambrook J, Fritsch E F, Maniatis T. *Molecular Cloning*. 2nd ed. New York: Cold Spring Harbor Laboratory Press, 1989. 19~ 22
- 7 Chomczynski P, Sacchi N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem*, 1987, **162** (1): 156~ 159
- 8 Duyk G M, Kim S, Myers R M *et al.* A genetic screen to identify candidate transcribed sequences in cloned mammalian genomic DNA. *Proc Natl Acad Sci USA*, 1990, **87** (22): 8995~ 8999
- 9 Auch D, Reth M. Exon trap cloning: using PCR to rapidly detect and clone exons from genomic DNA fragment. *Nucleic Acids Res*, 1990, **18** (22): 6743~ 6744
- 10 Buckler A J, Chang D D, Graw S L *et al.* Exon amplification: A strategy to isolate mammalian genes based on RNA splicing. *Proc Natl Acad Sci USA*, 1991, **88** (9): 4005~ 4009
- 11 Burn T C, Connors T D, Klinger K W *et al.* Increased exon-trapping efficiency through modifications to the pSPL3 splicing vector. *Gene*, 1995, **161** (2): 183~ 187
- 12 Nehls M, Pfeifer D, Boehm T. Exon amplification from complete libraries of genomic DNA using a novel phage vector with automatic plasmid excision facility: application to the mouse neurofibromatosis 1 locus. *Oncogene*, 1994, **9** (8): 2169~ 2175
- 13 Datson N A, Vosse E V, Dauwerse H G *et al.* Scanning for genes in large genomic regions: cosmid-based exon trapping of multiple exons in a single product. *Nucleic Acids Res*, 1996, **24** (6): 1105~ 1111
- 14 Watakabe A, Tanaka K, Shimura Y. The role of exon sequences in splice site selection. *Genes Dev*, 1993, **7** (3): 407~ 418
- 15 Horowitz D S, Krainer A R. Mechanisms for selecting 5' splice sites in mammalian pre-mRNA splicing. *Trends Genet*, 1994, **10** (3): 100~ 106
- 16 Chiang P W, Trbusek M, Osemlak-Hanzlik M *et al.* Rapid PCR-based method to directionally pull out longer cDNA fragments from cDNA libraries. *BioTech*, 1995, **18** (1): 36~ 40

Isolation of Novel Transcribed Sequences from FRAXA Site via Exon Trapping. KONG Jian, ZHU Ning-ning, YIN Bin, SHEN Yan, WU Guanyun (*National Laboratory of Medical Molecular Biology, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100005, China*).

Abstract Exon Trapping is a powerful method developed recently for isolating transcribed sequences from genomic DNA. Two novel exons, A91 and D12, were isolated from YAC209G4 spanning FRAXA site using the

exon trapping system with pSPL3 as splicing vector. A91 was amplified strongly in fetal liver and skeletal muscle cDNA library and mildly in liver, kidney and bone marrow library, respectively. While D12 was failed to be detected from the 8 libraries. A fragment of 315 bp containing A91 has been cloned (AM4470) from the skeletal muscle cDNA library. A single transcript of 2.8 kb was detected from mRNA samples of human heart and skeletal muscle on multiple tissues Northern blot using AM4470 probe.

Key words exon trapping, transcribed sequences, FRAXA site

一种改进的构建杂交扣留 cDNA 文库的方法*

赵大中¹⁾

(北京大学生命科学学院, 北京 100871)

陈民 种康 谭克辉

(中国科学院植物研究所发育室, 北京 100093)

摘要 介绍了一种改进的利用减法杂交构建 λ 噬菌体 cDNA 文库的方法, 该方法利用处理的 sscDNA 减去两种对照的 mRNA 富集目的 sscDNA 并用磁珠法分离杂交单体, 结果表明其增加了杂交单体的产量、简化了实验步骤, 而且更有利于富集所需基因, 是构建 cDNA 文库的一个有效改进.

关键词 cDNA 文库, 磁珠, 减法杂交

学科分类号 Q946.2

减法杂交与差异筛选是在未知基因产物的情况下克隆低丰度发育调控基因的经典方法^[1,2]. 生物在其发育过程中, 不同组织或处于不同发育阶段的同一组织, 其基因具有特异性的表达, 其 mRNA 也就不同. 利用基因相减的方法 (即减法杂交) 构建一个 cDNA 文库, 然后再进行差异筛选克隆出特异表达的基因, 这就是减法杂交与差异筛选的原理^[3]. 在实行减法杂交时, 一般利用羟基磷灰石柱层析的方法分离杂交单体, 这种技术存在着操作复杂、周期长、核酸损失量大等缺点. 本文介绍了一种新的利用磁珠分离杂交单体的技术, 同时在核酸杂交方式上也作了改进.

1 材料与方法

1.1 材料

冬小麦京冬 1 号 (*Triticum aestivum* L. cv Jingdong No. 1) 种子均匀播于垫有滤纸的培养皿中, 吸胀露白后: 在 25℃ 培养箱中暗培养 48 h, 即为未春化处理材料 (nonvernalization, NV); 在 0~4℃ 冰箱中暗培养 21 d 为春化处理材料 (vernalization 21 d, V₂₁); 将春化处理的材料转移到无菌的石英砂基上于

* 国家自然科学基金资助项目 (39330010).

¹⁾ 通讯联系人.

收稿日期: 1997-06-18, 修回日期: 1998-01-05