

用 RACE 结合 cDNA 文库筛选的方法 获取新的锌指蛋白基因

杜占文,刘立仁,张俊武

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

摘要:大多数有重要功能的蛋白质都含相应的由保守氨基酸顺序组成的功能结构域。本文首先根据蛋白质功能结构域保守氨基酸序列设计简并引物,用 PCR 方法扩增出基因 EST 序列,再利用改进的快速扩增 cDNA 末端(RACE)方法从 cDNA 文库中扩增出基因非同源部位,然后以非同源序列为探针,筛选 cDNA 文库。利用此方法成功地从人骨髓 cDNA 文库中克隆到几个编码锌指蛋白并代表原有 EST 的新的全长 cDNA。这一策略也应适用于筛选编码具有其他序列保守性功能结构域蛋白的基因。

关键词:基因克隆;快速扩增 cDNA 末端(RACE);非同源 DNA 序列;cDNA 文库筛选

中图分类号:Q784

文献标识码:A

文章编号:0253-9772(2002)03-0329-03

Isolation of New Zinc Finger Genes through cDNA Library Screening Combined with RACE

DU Zhan-wen, LIU Li-ren, ZHANG Jun-wu

(National Laboratory of Medical Molecular Biology, Institute of Basic Medical Science, CAMS & PUMC, Beijing 100005, China)

Abstract: Most of the important functionally proteins contain the corresponding function domains that consist of conserved amino acid sequences. The study provided a method to identify novel genes that encode proteins containing important functionally domains with conserved sequences. First, primers were designed according to the sequence of the cDNA library vector and the ESTs that have been obtained by reverse PCR and degenerate primers encoding Zinc finger domain. The cDNA library DNA was used as template for PCR amplification. The amplified fragment that contains nonhomologous sequences of the cDNA was inserted into pGEM-T easy vector. The fragment was recovered and used as a probe for screening the cDNA library. Several cDNAs with full length that encode proteins with Zinc finger domain and represent the original ESTs have been successfully cloned from a human bone marrow cDNA library. This strategy can also be used in screening genes that encode proteins containing differential function domains with conserved sequences.

Key words: gene cloning; rapid amplification of cDNA end (RACE); nonhomologous DNA sequences; cDNA library screening

转录调节蛋白在基因转录调控中具有重要的作用,这类蛋白的一个重要的特征就是它们与 DNA 的特异结合区常具有典型的氨基酸组成和结构,即基序(motif)。其中一类重要的基序就是锌指结构。锌指结构是靠锌原子维系的特

定区域,是蛋白质与 DNA 特异结合的重要结构域,因此多数锌指蛋白与基因的表达调控有关。TFIII-A 是最早发现的 C2H2 型锌指结构蛋白,到目前为止已经发现了许多个具有类似结构的蛋白质。

收稿日期:2001-12-07;修回日期:2002-03-14

基金项目:国家自然科学基金(39893320)和高等学校博士学科点专项科研基金(200001)资助

作者简介:杜占文(1971-),男,博士后,分子生物学专业

通讯作者:张俊武(1946-),男,教授,博士生导师,专业方向:分子生物学。Tel:010-65296423, E-mail: junwu_zhang@hotmail.com

C2H2 型锌指基序大约由 30 个氨基酸残基组成。在这种基序中,一个锌离子被 2 个半胱氨酸残基和 2 个组氨酸残基配位。一般在一个锌指结构域中可以有几个或十几个这样的锌指基序。每个锌指基序都有一些保守的氨基酸序列: HTGEKP 和 CPECGK^[1]。根据这些保守的氨基酸序列设计简并引物,进行 PCR 扩增就可以获得锌指蛋白基因的 EST。利用这一方法,我们已经获得了几十个新的 EST,但是当用这些 EST 直接从 cDNA 文库中筛选 cDNA 克隆时,由于序列的同源性,可能遇到筛选到的阳性克隆并非原 EST 代表的 cDNA 的情况。为解决此问题,我们建立了一种利用快速扩增 cDNA 末端(RACE)结合 cDNA 文库筛选的方法,成功地获得了代表原 EST 的 cDNA 克隆。

1 材料和方法

1.1 试剂

人骨髓 cDNA 文库为 Gibco BRL 公司产品。T-载体试剂盒,PCR DNA 纯化试剂盒、探针标记试剂盒购自 Promega

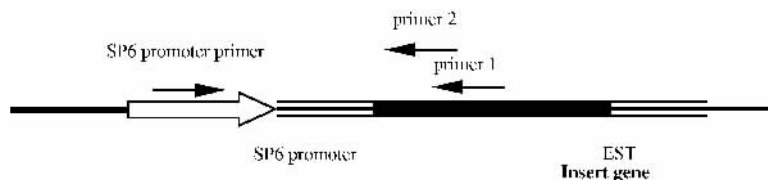


图 1 PCR 扩增 EST 旁侧序列图示
Fig. 1 A scheme of amplifying flanking sequence of EST by PCR

1.3 PCR 扩增片段的回收与克隆

第二次扩增后的 PCR 产物经低熔点琼脂糖凝胶进行电泳,在长波紫外灯下切下较长的扩增条带置于 1.5ml Eppendorf 管中,加入 5 倍体积的 TE(pH8.0),加热到 60°C 使凝胶融化。然后加入等体积酚抽提,以 4000×g 离心 10min,上层 DNA 溶液再用等体积酚-氯仿以及氯仿各抽提一遍。在 DNA 溶液中加入 0.2 体积的 10mol/L 醋酸铵、2 倍体积预冷的无水乙醇,4°C 放置 15min 后,以 12000×g 离心 20min,所获 DNA 沉淀用 70% 冷乙醇洗涤,以 12000×g 再离心 5min。所得 DNA 片段与 pGEM-T easy 载体经 T4 连接酶连接过夜后转化 XL1-blue 大肠杆菌,利用蓝白斑筛选挑出白色阳性克隆,在 37°C 培养扩增,收获菌体后以碱裂解法提取质粒 DNA。质粒 DNA 用 EcoRI 酶切并经琼脂糖凝胶电泳分离回收插入片段,也即新扩增出的 cDNA 片段。这一片段一般应为 cDNA 中不属于高度同源区的序列,可作为筛选 cDNA 文库的探针。

1.4 cDNA 文库的筛选

按分子克隆所述方法进行。将测定好滴度的细菌铺于含氨苄青霉素的 LB 琼脂平板表面,倒置培养后,进行转膜和

公司。Taq plus DNA 聚合酶购自上海生物工程公司。简并寡核苷酸引物由上海生物工程公司合成。硝酸纤维素膜为 Amersham 公司产品, $[\alpha\text{-}^{32}\text{P}]$ -dCTP 购自亚辉公司。

1.2 巢式 PCR 引物的设计与 PCR 扩增

以 cDNA 文库载体中 SP6 启动子序列设计 5' 端引物(图 1),其排列顺序为: 5' - ATT TAG GTG ACA CTA TAG - 3'。把已获得的 TF-III A 型锌指结构域 EST 的核苷酸序列与 GenBank 数据库核苷酸序列进行比较,找出非同源部位,根据非同源部位的核苷酸序列设计出向 5' 端扩增的巢式 PCR 引物。比如根据我们获得的 EST D26 的核苷酸顺序设计的两条引物为: 引物 1: 5' - CCA AAA GAT TTG AAG CC TC - 3', 引物 2: 5' - CAC AAT TAT ATG GTT TCT CTT - 3'。先以引物 1 和载体 5' 端 SP6 启动子引物进行扩增,用 PCR DNA 回收试剂盒回收 PCR 产物。取少量 PCR 产物为模板,以引物 2 和 SP6 启动子引物进行第二次扩增。这样可获得 cDNA 序列中 EST 一侧的 DNA 序列。

紫外交联。以步骤 1.3 制备的基因 5' 端非同源区片段为模板用无规则引物法合成 $[\alpha\text{-}^{32}\text{P}]$ 标记的探针。进行预杂交、杂交、洗膜和放射自显影。对应 X-光片上的斑点,取下阳性细菌菌落放入 2ml 含 100U 氨苄青霉素的 LB 培养基中,培养 3 小时后重新铺于平板上,进行第二轮筛选,之后再再进行第三轮筛选挑出分隔良好的单个阳性克隆。

2 结 果

2.1 快速克隆未知基因 5' 端

利用引物 1 与人骨髓 cDNA 文库载体的 SP6 promoter 引物一起对 cDNA 文库质粒 DNA 进行扩增,纯化 PCR 产物,取少量作为模板用引物 2 和 SP6 promoter 引物再扩增。PCR 产物经琼脂糖凝胶电泳,可看到 4 条 PCR 扩增条带,分别代表该基因 cDNA 不同的插入片段(图 2)。

2.2 人骨髓 cDNA 文库的筛选

以 $\alpha\text{-}^{32}\text{P}$ 标记的基因 5' 端非同源区片段为探针,筛选人骨髓 cDNA 文库。经过三轮筛选得到单个阳性克隆,用 PCR 方法鉴定插入片段的大小,最大插入片段约 3.9kb(见图 3)。

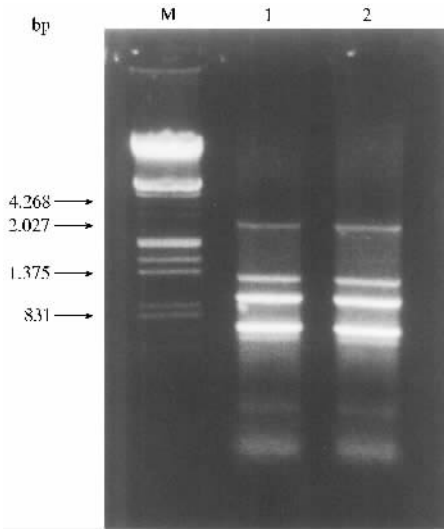


图 2 cDNA 文库用特定引物的 PCR 扩增
Fig. 2 The PCR amplification of the cDNA library using proper primer

M: λ DNA/*Hind* III + *Eco*RI Marker; 1, 2: PCR products.

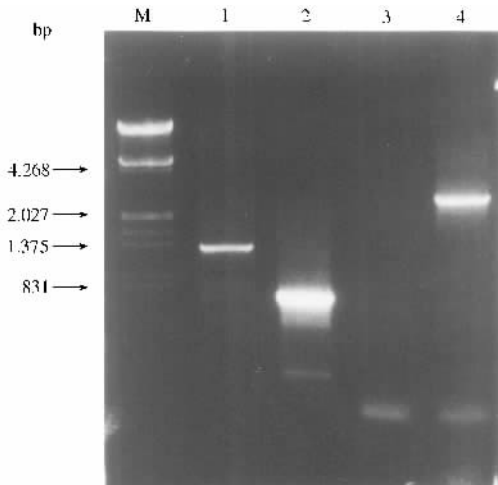


图 3 PCR 鉴别包含 cDNA 插入的阳性质粒
Fig. 3 Identification of the positive plasmids containing insertion cDNA by PCR

M: λ DNA/*Hind* III + *Eco*RI Marker; 1~4: The PCR fragments of the 4 positive plasmids.

为得到全长 cDNA 序列, 取含最大片段的克隆进行测序, 其插入片段的长度为 3888bp, 其中包括 EST 5' 端的部分序列。用 DNAsis 软件分析蛋白表达框, 再通过 GenBank blast 比较该基因与 GneBank 中基因的同源性, 确定了一个新的编码锌指蛋白的全长 cDNA, 在 GenBank 发布, 基因号为 AF246126。顺序比较证明这一 cDNA 准确代表了原 EST, 证明了本文介绍的方法的可行性。

3 讨 论

近年来人们一直致力于获得具有重要功能的新基因。尽管目前已有许多筛选新基因的方法, 如差异显示 PCR (DDRT-PCR)、消减 cDNA 文库等^[2,3], 但它们有着假阳性率高、对低丰度 mRNA 敏感性低及可重复性差等缺点。利用蛋白功能结构域的保守氨基酸序列, 设计出简并核苷酸引物进行 PCR 扩增, 得到 EST 序列, 再用其筛选 cDNA 文库也是一种很好的方法。

蛋白质的功能结构域是蛋白质发挥生物功能的重要区域, 许多蛋白质功能结构域保守氨基酸残基的序列已经清楚。根据这些保守氨基酸残基序列, 利用简并引物扩增的方法可以得到一些编码未知蛋白质基因的表达序列标签 (EST)。但由于这些 EST 具有很高的同源性, 通过 cDNA 文库筛选到代表某一特定 EST 的目的基因也有一定的难度。采用快速扩增 cDNA 末端 (RACE) 的方法, 也是一种获得全长 cDNA 的方法, 但是一次 RACE 只能延伸几百个碱基, 对于较长的基因, 用这种方法很难获得全长 cDNA。本文把 RACE 方法与 cDNA 文库筛选结合起来, 即利用 EST 序列结合 cDNA 文库载体序列设计引物, 通过 PCR 扩增出基因非同源区, 再利用非同源区的 DNA 片段为探针进行 cDNA 文库筛选, 并成功地获得了能够代表原 EST 的全长 cDNA, 为功能基因的筛选提供了一种可行的方法。

参 考 文 献 (References):

- [1] Furukawa T, Yang Y, Nakamoto B, *et al.* Identification of new genes expressed in a human erythroleukemia cell line[J]. Blood Cell, Molecular and Disease, 1996, 31: 11~22.
- [2] Sagerstrom C G, Sun B I, Sive H L, *et al.* Subtractive cloning: past, present, and future[J]. Annu Rev Biochem, 1997, 66: 751~783.
- [3] Peng Liang, Arthur B Pardee. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction [J]. Science, 1992, 257: 967~971.