

# 基因克隆技术的研究进展\*

钟 军, 李 迅, 官春云

(湖南农业大学 油料作物研究所, 中国湖南 长沙 410128)

**摘要:** 为能快速而准确地克隆目的基因, 综述了一些基因克隆常用技术, 包括差异表达基因分离技术、转座子标签技术、图位克隆技术、同源序列技术、表达序列标签技术的原理、应用及应用潜力, 并对其作了简要的评价。这些技术有利有弊, 应根据不同的实验目的和水平来选择相应的技术。

**关键词:** 基因; 克隆; 差异表达基因分离技术; 转座子标签技术; 图位克隆技术; 同源序列技术; 表达序列标签技术

中图分类号: Q78

文献标识码: A

文章编号: 1007-7847(2002)S1-0148-05

## Advances in Gene Cloning Technique

ZHONG Jun, LI Xun, GUAN Chun-yun

(The Oil Crop Institute of Hunan Agriculture University, Changsha 410128, Hunan, China)

**Abstract:** To clone candidate gene quickly and correctly, advances about gene cloning included map-based cloning, transposon tagging, homology-based candidate gene method, expressed sequence tagging methods and some differentially expressed gene clone method are introduced and appraised. Because of the advantages and disadvantages of those techniques, various technique should be selected according special purpose and level.

**Key words:** gene; clone; differentially expressed gene clone method; transposon tagging; map-based cloning; homology-based candidate gene method; expressed sequence tagging method

(*Life Science Research*, 2002, 6(Suppl): 148~ 152)

克隆基因的途径有两种, 正向遗传学和反向遗传学途径。前者是依据目标基因所表现的功能为基础, 通过鉴定其产物或某种表型突变而进行的; 后者则着眼于基因本身, 通过特定的序列或在基因组中的位置进行。近几十年来, 许多重点实验室致力于植物基因的克隆, 到1992年取得了突破性进展。基因的克隆一般采用下列技术: 差异表达基因分离技术、转座子标签技术、表达序列标签技术、图位克隆技术和同源序列技术等。

### 1 差异表达基因分离技术

#### 1.1 扣除杂交技术

扣除杂交技术的原理是用有特异性表达基因

的目标样提取 mRNA 经逆转录形成 cDNA 探针, 与无特异性表达基因的参照样的过量 mRNA 或 cDNA 杂交, 经两轮充分杂交后, 移去杂交分子和过量的无特异性表达基因的参照样 mRNA 或 cDNA, 将不形成杂交体的有特异性表达基因的目标样 cDNA 纯化富集、扩增, 建立相应 cDNA 文库即为差异表达基因 cDNA 文库。此技术最早是由 Lamar 和 Palmer 于 1984 年提出<sup>[1]</sup>, 他们先用超声波打断雌性小鼠的 DNA, 用 Mbo I 完全消化雄性小鼠 DNA; 将两者一起变性、复性, 再将产物克隆入表达载体的 BamH I 位点中, 只有那些两端有 GATC 序列的基因才能被克隆入载体, 这样就达到了扣除两者共有序列的目的, 并得到雄性小鼠

\* 收稿日期: 2002-06-11; 修回日期: 2002-10-14

作者简介: 钟军(1973), 女, 湖南沅江人, 博士研究生, 从事分子遗传学研究, Tel: + 86-0731-4621882, E-mail: zhhjp@sohu.com

Y染色体的DNA. 该法的缺点是技术要求较高、工作量大,并且有时得出的数据不大可靠. 然而经过多年的改进,扣除杂交技术已发展成为众多克隆技术的基础,很多克隆技术都是从此技术衍生发展而来,如抑制性扣除杂交技术. 它是先将样本mRNA和参照mRNA分别逆转录成cDNA,用4种碱基识别并酶切两种cDNA以产生平端片段;样本mRNA分布接上接头1和接头2,并与过量的经RsaI消化的参照样本杂交;在接头处设计引物,使具有差异表达的片段成为PCR扩增的模板,重复杂交以减少非特异性扩增片段,利用接头上的酶切位点进行克隆和测序.

### 1.2 差式筛选技术

分别从有特异性表达基因的目标样和无特异性表达基因的参照样中提取mRNA经逆转录形成cDNA,并构建有特异性表达基因的目标样的cDNA文库,再分别用有特异性表达基因的目标样和无特异性表达基因的参照样的cDNA探针与有特异性表达基因的目标样cDNA文库的菌落或噬菌斑作原位杂交,选出只与有特异性表达基因的目标样杂交的克隆,即为有特异性表达基因的目标样特异表达的克隆. 此技术要经过多轮菌斑杂交,不仅费力费钱,重复性差,而且灵敏度低.

### 1.3 差异显示PCR

差异显示PCR<sup>[2]</sup>是1992年由Liang报道的. 此法的主要原理是由mRNA-PCR共同组成的,利用大多数真核细胞基因的mRNA结尾处有poly(A)结构,在其3'端设计5' poly(T)引物,该引物可与mRNA结合,从而使这部分的基因达到转录. 然后从两种不同类型的细胞中分离mRNA,作反转录,比较PCR产物后即可克隆目的基因. 其具体操作为:用一套锚定引物反转录每一部分材料后,再分别用这一套锚定引物和随机引物从反转录的每一部分材料中扩增cDNA,电泳分离产生扩增片段后再扩增两种不同状态下有差异的片段,并且进行克隆和测序. 如果采用同位素标记的方法,在凝胶电泳或放射自显影照片上就可发现不同长度的扩增产物. 该技术能快速地显示细胞mRNA的组成,可对各样本mRNA的差异同时进行比较和展示,并且所用mRNA量少;可以立即对扩增的cDNA进行测序和克隆及与探针标记、杂交和文库筛选. 它的缺点是假阳性条带较多、对丰度低的基因表达不容易检测,无法定量研究;基因的克隆受mRNA表达时效的影响;其扩增出的条带往往是

在基因的3'端有比较短的一段序列,这部分序列往往是在基因3' UTR区,所提供的信息比较少,为得到cDNA必须进行cDNA文库的筛选.

PCR技术的发展和应用于差异表达基因分离技术注入了新的活力. 如在差式筛选技术中引入了PCR<sup>[3]</sup>,不仅提高了效率,还降低了假阳性. Litsitsy<sup>[4]</sup>等将PCR引入扣除杂交,发展了扣除扩增方案,同时用特异性引物结合技术在PCR中选择性地扩增差异cDNA,形成了代表差示分析法;Straus等<sup>[5]</sup>受扣除杂交的启发,提出了基因组扣除方案,若目的基因存在该基因缺失的突变体,就可以用变性的野生型基因组DNA与缺失基因组DNA进行扣除杂交,纯化出“多余的DNA片段”,然后设计引物进行扩增、富集、制备探针,筛选候选基因的DNA文库. 以上方法在具体操作中也可结合起来使用,如Kang等<sup>[6]</sup>利用扣除杂交后的RNA或DNA样品进行差异显示PCR,发展了交互式差减差异RNA显示技术,成为一种高效、快速克隆差异表达基因的方法.

## 2 转座子标签技术

转座子是McClintock首先在玉米中发现的,以后在金鱼草、飞燕草、甜豌豆、细菌、酵母、拟南芥和一些多细胞绿藻等植物中也发现了转座子的存在<sup>[7]</sup>. 由转座子引起的突变可以转座子DNA为探针,从纯合突变株构建的基因组文库中筛选含该转座子的同源片段,并获得含有部分突变株DNA序列的克隆,进而以该DNA为探针,筛选野生型的基因组文库,最终得到完整的基因(图1). 在植物中利用的转座子有:Ac/Ds、En/Spm和Mutator,其中应用最多的是Ac/Ds双因子系统. Johal等<sup>[8]</sup>利用转座子克隆到玉米抗圆斑病基因HM1,这是第一个克隆的功能产物未知的植物抗病基因;Jones等<sup>[9]</sup>利用Ac/Ds系统克隆到番茄抗叶霉病基因cf-9;Aarts等<sup>[10]</sup>利用En/Spm克隆到拟南芥菜的一个雄性不育基因;Cui等<sup>[11]</sup>利用Mutator克隆到了玉米胞质一个恢复基因. 利用转座子标签技术在克隆植物抗病基因方面也获得了巨大成果(表1).

80年代中期人们在了解转座子的遗传特性和分子结构的基础上,开始用转座因子作为分子标签来分离由于它的插入而引起突变的基因,它的优点在于:1)分离基因不需要预先知道被标签产物的结构、性质,因此那些背景尚不清楚的如发

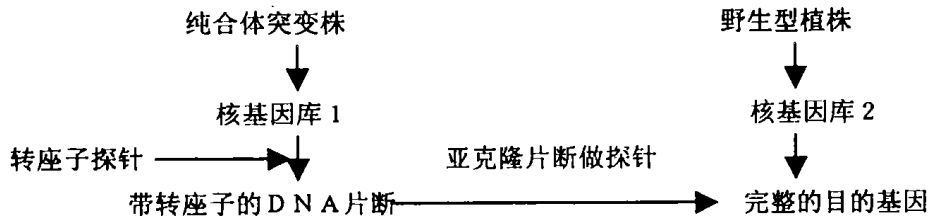


图 1 转座子标签法克隆基因的过程

Fig. 1 The process of cloning gene in transposon tagging

育、抗病及与生理过程有关的基因都有可能用此法来分离; 2) 转座子插入引起的突变会由于转座子的切离而回复; 3) 由于转座子总是沿其邻近的位点转座, 因此只要转座子在染色体上定位后, 就很容易确定与其连锁的突变基因的相对位置; 4) 如果把转座子导入目标植物, 通过自交、杂交、回交等方法便可得到一个较大的含不同插入位点的转座子群体; 5) 把转座子导入异源植物不受转化条件的限制。但由于一些植物存在内源转座子, 像金鱼草、拟南芥、玉米等, 因而利用和植物本身同源的转座子转座就不能真正标签基因的转座, 所以转座子标签技术多用于异源植物; 另外转座子在不同植物中转座的频率和活性相差很大, 并且需要筛选大量的个体来鉴定转座子突变个体, 植物基因组中多拷贝基因的存在也限制了转座子标

签技术的应用。因此, 许多科学家正致力于转座子系统的改进以提高转座子的标签效率和提高突变的频率。在这方面研究较多的是 *Ac/Ds* 系统。首先对 *Ac* 转座子作适当的修饰: 去掉 *Ac* 两端 200 bp 的任一端, *Ac* 转座子就不能发生转座, 成为稳定型; 其次, 把 *Ac* 转座酶基因的启动子去掉, 换成 35S 强启动子, 这样可以大大提高 *Ac* 转座酶的转录速度, 从而提高转座子的切割频率和转录效率; 第三, 构建非自主转座子; 最后, 人们又将 *Ds* 与特异重组系统 *Cre/lox* 成功地结合, 把 *Cre/lox* 重组系统的一个 *lox* 位点放在 *Ds* 内部, 同时将另一个 *lox* 克隆在 T-DNA 区, *Ds* 转座到 T-DNA 附近后, 这两个 *lox* 位点再经 *Cre* 的反式作用而重组, 使 T-DNA 和 *Ds* 之间的 DNA 产生重排, 可使突变率提高, 还可引起位点特异突变。

表 1 已克隆到的植物抗病基因  
Table 1 Resistance gene cloned in plant

| 植物<br>Plant | 基因名称<br>Gene names | 克隆方法<br>Cloning method | 植物<br>Plant | 基因名称<br>Gene names | 克隆方法<br>Cloning method | 植物<br>Plant | 基因名称<br>Gene names | 克隆方法<br>Cloning method |
|-------------|--------------------|------------------------|-------------|--------------------|------------------------|-------------|--------------------|------------------------|
| 拟南芥         | Eds1               | 转座子标签法                 | 亚 麻         | L6                 | 转座子标签法                 | 番 茄         | Pto                | 图位克隆法                  |
|             | RPS2               | 图位克隆法                  |             | M                  | 转座子标签法                 |             | Fen                | 图位克隆法                  |
|             | RPS5               | 图位克隆法                  |             | Rp1-D              | 转座子标签法                 |             | Pf                 | 图位克隆法                  |
|             | RPP1               | 图位克隆法                  | 玉 米         | Hm1                | 转座子标签法                 |             | Cf-2               | 图位克隆法                  |
|             | RPP5               | 图位克隆法                  |             | Hm2                | 图位克隆法                  |             | Cf-4               | 图位克隆法                  |
|             | RPP8               | 图位克隆法                  |             | 水 稻                | Xa21                   |             | 图位克隆法              | Cf-5                   |
|             | RPM1               | 图位克隆法                  | Xa1         |                    | 图位克隆法                  |             | Cf-9               | 转座子标签法                 |
|             | NPP1               | 图位克隆法                  | Pt-b        |                    | 图位克隆法                  |             | l2C                | 图位克隆法                  |
| 烟 草         | N                  | 转座子标签法                 | 小 麦         | Cre3               | 图位克隆法                  |             | 甜 菜                | HsIPre-1               |

### 3 图位克隆技术

图位克隆技术是近几年随着各种植物的分子标记图谱的相继建立而发展起来的一种新的基因克隆技术, 是根据功能基因组中都有相对稳定的基因座, 在利用分子标记技术对目的基因进行精

确定位的基础上, 用与目的基因紧密连锁的分子标记筛选 DNA 文库(包括 YAC、BAC、TAC、PAC), 从而构建目的基因区域的物理图谱, 再利用此物理图谱通过染色体步查逐步逼近目的基因或通过染色体登陆的方法最终找到包含该目的基因的克隆, 并通过遗传转化试验验证目的基因的功能(图

2). 随着各种植物的高密度遗传图谱和物理图谱的建立, 图位克隆技术在植物的基因克隆中有着广阔的应用前景. 图位克隆首先被用于模式植物拟南芥上, Giraudat 等成功地克隆了与拟南芥种子形成和发芽有关的脱落酸信号传导基因  $ABI3^{[12]}$ ; Arondel 等则克隆到拟南芥脂肪酸脱饱和基因酶基因  $FAD3^{[13]}$ ; 而番茄的  $Pto$  基因是第一个用图位克隆法克隆到的抗病基因<sup>[14]</sup>, 图位克隆技术的发展加速了抗病基因的克隆(表 1).

然而, 此技术的缺点是: 如果要构建完整的基因组文库, 建立饱和的分子标记连锁图和完善的遗传转化体系, 对基因组较大、标记数量不多而重复序列较多的生物采用此法投资大且效率低, 因而图位克隆技术目前仅局限于拟南芥、水稻、番茄等图谱饱和的模式植物上. 随着比较基因组学的

兴起, 可以利用同科异种间染色体的共线性, 通过比较基因组分析和染色体登陆法, 将控制该物种有关性状的基因或分子标记直接定位在它的分子标记连锁图谱中, 如利用小麦与水稻基因组间的同线性, 将小麦控制开花时间的基因(URN) 和控制谷粒硬度的基因(HA) 定位在第 5 组染色体上, 并利用水稻图谱信息筛选水稻小麦同线区的 BAC 文库或 YAC 文库, 用于测序进而搜索候选基因<sup>[15]</sup>; 另外也可以利用同科异种间染色体的共线性以模式植物为种间图位克隆中介来克隆其他植物基因. 可以预测, 在不久的将来, 在比较基因组学中统一遗传系统理论的指导下, 在具有高多态性水平的以 PCR 为基础的分子标记日渐成熟的背景下, 在作物的研究中应用图位克隆技术必将有更大的成果.

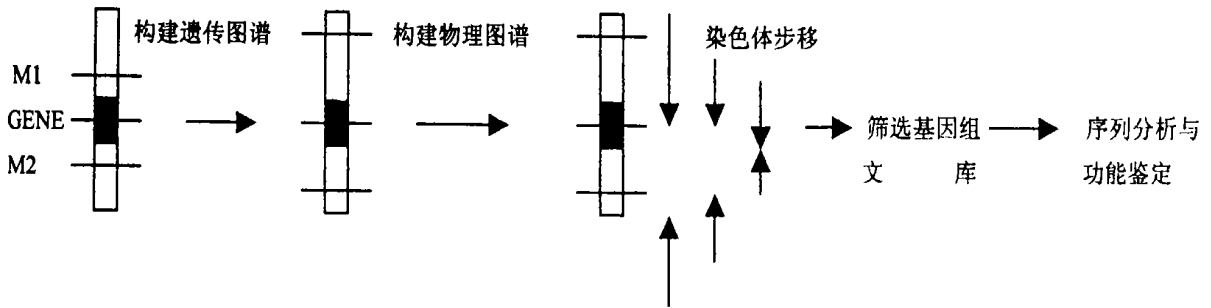


图 2 图位克隆技术克隆基因的过程

Fig. 2 The process of cloning gene in map-based cloning technique

#### 4 同源序列技术

这是一种克隆抗性基因的新策略. 根据基因家族保守氨基酸序列设计兼并引物, 并用兼并引物对含有目的基因的 DNA 文库进行扩增, 再对扩增产物进行扩增、克隆和鉴定, 如通过对已克隆的抗性基因蛋白序列的分析, 发现这些抗病基因所编码的蛋白都存在同源序列, 如富含亮氨酸重复区(LRR)、核苷酸结合位点(NBS)、蛋白激酶(PK)等, 所以人们设想用同源序列来克隆新的抗病基因. 王石平<sup>[16]</sup>等根据这些保守序列设计合成了特异性和兼并性引物, 对水稻基因组 DNA 进行扩增, 得到了 100 个大小不同的克隆, 定位了 26 个克隆, 其中的 10 个克隆与 8 个已定位的水稻抗病基因在分子标记连锁图上的位置对应或毗邻, 表明这些克隆的抗病基因在染色体位置上有较好的对应关系, 证明了同源序列分离和克隆基因的可能性. 黄烈健<sup>[7]</sup>等根据已克隆的抗病基因同源序列设计引物, 对玉米 DNA 进行扩增、克隆并对部

分克隆进行测序, 测序结果与 Genbank 进行了序列同源的比较, 两者同源性达 85% 以上. 此法的长处在于: 若同源序列与抗病基因紧密连锁或带有抗性基因的同源序列, 则可用其作探针来筛选基因组文库, 获得候选基因, 但植物基因组中有许多与 LRR 或 NBS 同源的区域, 其功能并不都是与抗病有关. 所以, 该技术应与图位克隆技术结合使用. 同源序列技术存在的问题也是值得重视和思考的: 1) 由于密码子的兼并性和不同同源序列间同源程度的差异, 兼并引物的特异性要设计得当; 2) 由于某些同源序列并不专属于某一家族, 因而扩增产物不一定是某一基因家族的成员; 3) 基因家族成员成簇存在, 克隆的基因片段是否为目的基因尚需进一步判断. 因此对 PCR 扩增产物和克隆产物, 有必要进行基因与性状共分离分析、插入失活或遗传转化等功能鉴定, 以便最终筛选到目的基因.

## 5 表达序列标签技术(EST)

EST 是完整基因中能特异性标记基因的一部分序列,通常包含了基因足够的结构信息区,从而与其它基因相区分,大规模 EST 克隆和资料库的建立,为利用生物信息学克隆基因提供了条件.其基本原理是从组织特异性或细胞特异性的 DNA 文库中随机挑选克隆,并进行 5' 和 3' 端部分测序,通过对基因库的检索,可以检测所测序以及翻译的多肽氨基酸序列与基因库中已知的是否有同源性,最后对发现的新基因进行突变检测和表达分析.这样不仅可检测到许多已知的基因,更可以发现许多未知的基因.

该技术建立在大量已有的生物信息资源基础上的,同时,结合了目前的新技术,为大规模克隆基因提供了捷径.目前,国际上大部分的基因克隆是从分析同源 EST 开始的,如田芳<sup>[18]</sup>探讨了如何利用 EST 发现并克隆肿瘤基因;Okano<sup>[19]</sup>利用此技术克隆到一簇新的哺乳动物的 DNA-5 胞苷甲基转移酶基因;Dymock 等通过 EST 数据库分析克隆到细胞分裂素的 1 个跨膜突变受体基因 GCR1<sup>[20]</sup>.另外,可利用对体细胞杂交体 DNA 进行扩增来达到 EST 在染色体上的定位.美国密执安大学利用此法将 36 个 EST 定位于人的第 6 号染色体.

EST 具有广泛的用途.利用对独立 EST 的拼接可获得新基因的全长 cDNA 序列;利用 EST 标签克隆的斑点杂交可鉴定涉及组织或器官发育过程不同阶段各个基因表达的不同,与其它差别杂交方法相比,它能进行初步筛选和鉴定,因而节省了时间;由于事先获得了某个基因的 EST,所以可加速对该基因的克隆以及序列测定;通过构建 cDNA 文库,可得到有关植物的 EST,从而可了解植物生长发育过程中不同的基因表达,为植物育种改良打下基础.

### 参考文献(References):

- [1] LAMAR H R. Y-encoded, species-specific DNA in the mice: evidence that the Y chromosome exists in two polymorphic forms in hybrid strains[J]. Cell, 1984, 37: 171-177.
- [2] LIANG P, PARADEE A B. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction[J]. Science, 1992, 257: 967-970.
- [3] SANG J, THOMPSON N J. An efficient procedure for obtaining long cDNA clones from phage library screening[J]. Biotechnology, 1994, 17: 447-451.
- [4] LISITSYN N, WIGLER M. Cloning the difference between two complex genomes[J]. Science, 1993, 259: 946-951.
- [5] STRAUS D, AUSUBEL F M. Genomic subtraction for cloning DNA corresponding to deletion mutations[J]. Proc Natl Acad Sci USA, 1990, 87: 1889-1893.
- [6] KANG D C, FRANCE R L, SU Z Z, et al. Reciprocal subtraction differential RNA display: an efficient rapid procedure for isolation differential expressed gene sequences[J]. Proc Natl Acad Sci USA, 1998, 95(23): 13788-13793.
- [7] McClintock. Cytogenetic studies of maize neurospora[J]. Carnegie Inst Wash York, 1974, 46: 146-152.
- [8] JOHAL G S, BRIDGS S P. Reductase activity encoded by the Hm1 disease resistance gene in maize[J]. Science, 1992, 258: 985-987.
- [9] JONES D A, HAMMOND K E. Isolation of the tomato cf 9 for resistance to *cladoporium fulvum* by transposon tagging[J]. Science, 1994, 266: 789-793.
- [10] AARTS M G, DIRKS W G, STIEKEMA W T. Transposon tagging of male sterility gene in *Arabidopsis*[J]. Nature, 1993, 363: 715-717.
- [11] CUI X Q, WISE R. The rf2 nuclear restore gene of male sterile T-cytoplasm maize[J]. Science, 1996, 272: 1334-1336.
- [12] GIRAUDAT J, HAUGE B M, VALONC C. Isolation of the *Arabidopsis* ABI3 gene by positional cloning[J]. Plant Cell, 1992, 4: 1251-1261.
- [13] ARONDEL V, LEMIEUX B, HWANG I, et al. Map-based cloning of a gene controlling omega-3 fatty acid desaturation in *Arabidopsis*[J]. Science, 1992, 258: 1353-1355.
- [14] MARTIN G B, BROMMON S H, CHUMWONGSE J, et al. Map-based cloning of a protein kinase gene conferring disease in tomato[J]. Science, 1993, 262: 1432-1436.
- [15] SARMA R N, FISH L, GILL B S, et al. Physical characterization of the homologous groups chromosome of wheat in terms of rice linkage blocks, and physical mapping of some important genes[J]. Genome, 2000, 43: 191-198.
- [16] 王石平. 用同源序列的染色体定位寻找水稻抗病基因 DNA 片段[J]. 植物学报, 1998, 40(1): 42-45.
- [17] 黄烈健. 玉米抗病相关候选基因 DNA 片段克隆[J]. 中国农业科学, 2000, 23(增刊): 141-146.
- [18] 田芳. 运用生物信息学快速获得与肿瘤基因同源的 EST 及其新基因克隆的策略[J]. 生命科学, 2000, 12(2): 72-75.
- [19] OKANO M, XIE S P, LI E. Cloning and characterization of family of novel mammalian DNA (cytosine-5) methyltransferases[J]. Nature Genetics, 1998, 19(3): 219-220.
- [20] DYMOCK C N. A higher plant seven transmembrane receptor that influences sensitivity to cytokines[J]. Current Biology, 1998, 8(16): 315-324.