

## 基因组研究中全长 cDNA 克隆的策略

张庆华 茅 矛 陈 竺

(上海第二医科大学附属瑞金医院 上海血液学研究所  
(卫生部暨上海市人类基因组研究重点实验室, 上海 200025))

**摘要** 全长 cDNA 的克隆是目前人类基因组研究中的一个重要方面,与大规模基因组测序相比较, cDNA 测序克隆用相对较少费用获得更多的功能基因信息,故也是符合我国国情的基因组研究的主要方向。如何更加高效地获得新的全长 cDNA,本文结合作者自己工作中的经验,对从全长文库的建立到全长 cDNA 的测序及克隆方法作了较为详细的介绍。

**关键词** 全长 cDNA 基因克隆 数据库 人类基因组研究

人类基因组计划(HGP)现在已经进入了大规模测序时代,其主要内容为基因组 DNA 测序以及功能性 cDNA 测序两个方面。cDNA 由 mRNA 经逆转录而来,在大规模的研究工作中离不开 cDNA 文库的建立工作。为了得到更多的全长 cDNA 克隆,以便获得新基因, cDNA 文库的质量显得至关重要。在此就 cDNA 文库的构建以及全长 cDNA 克隆的几种较新的方法作简要介绍。

### 1 全长 cDNA 文库的建立

真核生物 mRNA 的结构特点,在 5' 末端有甲基化 G(mG)且 5'-5' 三磷酸键连接的特殊的“帽子”结构,在 3' 端有 polyA 尾巴,经典的 cDNA 文库建立方法是用 oligo(dT)作逆转录引物,也有用随机引物的,然后给所合成的 cDNA 加上适当的连接结构,连接到适当的载体中而成。<sup>[1]</sup>这种建库方法实用、高效、简单方便,但最大的缺点是文库中克隆片段的长度一般较小,一般需若干个克隆才能覆盖一个完整的 cDNA<sup>[2,3]</sup>而对有些较大的基因,甚至无法得到完整的 cDNA,此种情况在 oligo(dT)方法所建立的文库尤甚。

为了更加有效地克隆新的基因,如何克服仅由 mRNA polyA 尾合成 cDNA 技术路线的局限性,以及由逆转录酶作用特点导致的局限性,建立富含全长 cDNA 的文库就显得非常有意义,传统的逆转录酶 AMV 和 MMLV 等都不能逆转录出很长的 cDNA, Superscript<sup>TM</sup> (Gibco BRL) 的推出,使逆转录出更长的 cDNA 成为可能。为了尽量克隆到真正意义上的全长 cDNA, mRNA 5' 的帽子结构就成为人们

首先追逐的对象。

Maruyama 等人<sup>[4,5]</sup>利用 5' 帽的结构特点,用细菌碱性磷酸酶(BAP)去除无 5' 端 mG 保护的部分降解的不完整 mRNA 末端的游离磷酸基团,再用烟草酸性磷酸酶(TAP)去除 5' 的 mG 结构仅保留一个 5'-P,合成一段 r-oligo,用 T4 RNA 连接酶连接到 mRNA 去除了 mG 以及二磷酸基团的 5'-P 上,继而利用修饰的 oligo(dT)作逆转录,并进行 PCR 反应,称为 oligo-capping,然后克隆到载体建库。CLON-TECH 公司推出的试剂盒 CapFinder (现又命名为 SMART<sup>TM</sup>)则是利用帽子结构特征,合成一段 oligo 作为 CapSwitch,在 oligo(dT)引导逆转录后作 PCR 反应的 5' 引物。这两种方法由于利用帽的位置及结构的特殊性,理论上在建库的克隆产物中即含有大量的全长 cDNA 序列,利用极其少量的组织或 mRNA 建库。其缺点是理论上会产生由于 PCR 或其它原因导致的表达谱的偏移(bias)或克隆测序中的冗余性(redundancy)。在我们自己的工作中使用了 CapFinder,克隆的平均长度在 1.2kb 左右,全长克隆为 60%,明显高于普通文库,优化 PCR 扩增条件后尽可能减少了文库的代表性的偏移和冗余性。<sup>[6]</sup>

Kato 等人<sup>[6]</sup>用合成的 DNA-RNA 嵌合连接子代替 Maruyama 等方法中的 PCR 引物,连接到去除 5' mG 的游离 5'-P 上, T4 RNA 连接酶连接、再复制后定向克隆到载体中,得到文库的插入片段平均长度 1.2kb,有 83% 的克隆含有完整的阅读框架。

Ederly 等<sup>[8]</sup>报道一种称为 CAPture 的方法,原理与前面几个方法完全不同,且能够非常有效地建

立全长的 cDNA 文库。利用“帽结合蛋白”专门结合 mRNA 的帽结构,含有完整帽子结构的 mRNA 被富集,逆转录后用 RNaseA 作用于 mRNA/cDNA 杂合子,未被 CAP 结合蛋白结合的“部分降解”的不完整 mRNA,有一些虽结合了“帽子结合蛋白”但 cDNA 并不完整的 mRNA 都被降解,仅完整 mRNA 又有全长 cDNA 第一链结合保护的那些 mRNA 被保护,进一步纯化富集后建库,获得的克隆全长比例很高,但每次需极其大量的 mRNA (100 $\mu$ g),另外,被帽结合蛋白结合的 mRNA 5' 端约 10-50bp 被丢失。

Carninci 等<sup>[9]</sup>最近报道 Cap trapper 的方法,其原理是用过磷酸盐标记 5' 的末端甲基化的 mG,然后用免疫磁珠法分离出并富集逆转录后全长的 cDNA 进行建库。标记的方法有其特殊性,即根据 5' 末端帽结构上的 mG 核糖基团上有 2,3 邻位两个 -OH 基团,可被生物素酰肼作用并标记上生物素,而进一步被富集分离,作几次连接及扩增即可克隆入载体。结果可以得到相当长度的 cDNA,95% 为全长,得率和产量较前增加 10 - 50 倍,并且由于不作 PCR 反应,可减少理论上的偏移和冗余性。但该方法中 mRNA 模板在化学试剂及酶反应体系暴露时间较长,造成效率的降低,目前已有一些改进的方法<sup>[10]</sup>。另外用 tRNA 封闭吸附用的柱子,也可能造成一些非特异的污染。

目前,也有些研究人员根据逆转录后的 cDNA 的大小进行分级,即通过凝胶电泳,只回收分子量较大的片段,这样,虽然建库的 cDNA 总量减少了,但可大大提高所建文库中的较大插入片段的比例,有利于对较大基因进行克隆。

以上方法 mRNA 的逆转录都是在 37 ~ 42 °C 下进行的,此温度不能解开 mRNA 可能形成的二级结构,使逆转录酶不能通过二级结构区段,而得不到全长的 cDNA 第一链,这也是制约大基因全长 cDNA 文库的一个“瓶颈”。Carninci 等<sup>[11]</sup>最近报道,一种二糖海藻糖 (trehalose) 可以提高某些酶如 MMLV 逆转录酶和 DNase、Mlu 等的热稳定性(热稳定效应),可提高耐热温度 5-10 °C 对某些酶还可在适当高温下增加酶的活性(热激活效应)。这可以减少逆转录过程中的二级结构,增加全长 cDNA (尤其对较大基因)第一链的合成比例。这无疑给大基因全长 cDNA 文库的建立带来新的希望。

## 2 全长 cDNA 的克隆

cDNA 文库建立后,无论是在研究基因表达的

EST 大规模测序工作中,对其克隆进行测序并与数据库进行比较后,确定的非已知基因克隆,还是用特定探针筛选出的克隆,皆可以进行新基因的全长克隆。如果文库中的 cDNA 已经是全长,进行测序工作即可。全长测序方法有多种,亚克隆、部分缺失法、引物延伸等都是行之有效的方法,对大规模工作而言,后两种更加可行。

对一个可靠的 cDNA 序列,判断它是否为全长顺序,可以考虑如下几个方面:1. 直接从序列上评价:5' 端:(1)有同源全长基因的比较,通过与其它生物已有的对应基因末端进行比较来判断,(2)无同源基因的新基因,判断编码框架是否完整,a. 在开放阅读框架的第一个 ATG 上游有同框架的终止密码,b. 无终止密码的则考虑有保守的 Kozak 序列<sup>[12]</sup>,判断是否自转录起始点,有资料表明,在 5' 帽结构后一般都有一段富含嘧啶的区域,另外如果 cDNA 5' 序列与基因组序列中经 S1 酶切保护的部分相同,则可以确定得到的 cDNA 5' 是全长的。3' 端:(1)有同源全长基因的比较,(2)编码框架的下游有终止密码,(3)有一个以上的 polyA 加尾信号,(4)无明显加尾信号的则也有 polyA 尾。2. 用实验方法来证实:(1)Northern Blot 证实大小一致,(2)用引物延伸法确定 5' 和 3' 的长度。

对一段不完整(非全长)的 cDNA 序列,克隆全长的方法可以从以下几方面进行:1. 传统的标记探针文库筛选,需选择合适的文库,考虑合适的组织来源、建库方法,费时较长,工作量较大,适合于少量基因的研究;

2. 快速 cDNA 末端扩增法 (RACE)<sup>[13]</sup>,合成一段基因特异的引物 (GSP),结合使用末端的引物作 PCR 扩增,可自己选用试剂进行扩增,也有已构建成连接有末端引物的 cDNA,如 CLONTECH 的试剂盒 Marathon-Ready<sup>TM</sup>等;

3. 利用生物信息学,使用公共数据库中的资料进行 EST 查询(电子数据库查询 electro database screening),具体可分下面两种情况:(1)芯片延伸 (silico extension) 或电子步移 (electro walking) 对一段获得的 cDNA 序列,查询数据库得到一些与该片段部分重叠的 EST 数据,进行拼接后可以部分延伸,继续进行下一轮的查询、拼接和延伸,就有可能得到全长的 cDNA 顺序信息,这在我们的日常工作中显得切实可行<sup>[6]</sup>;(2)从头芯片克隆 (de novo silico cloning) 对某些与人类或其它动物(尤其是灵长类和其它哺乳动物)的某个基因(或基因家族中的保守

区段)有高度同源性的基因,可将同源基因或保守的区段进行电子筛库,得到人类的 EST 信息,进一步进行拼接、延伸,而有可能得到全长的 cDNA 顺序信息,这方面的研究方法也已有文献报道<sup>[14,15]</sup>,值得一提的是, Capone 等报道的 TSA-1/ SCA-2 与我们前期工作中克隆的一个基因 RIG E 是同一转录单位,且基因顺序完全相同。我们目前的工作中也有如此的例子。用数据库克隆到的 cDNA 顺序,尚需进一步的 Northern blot 和/或 RT-PCR 加以证实。

### 参考文献

- [ 1 ] Gubler U and Hoffman BJ. Gene ,1983;25:263 - 269.
- [ 2 ] Mao M, Yu M, Tong J-H, et al. Proc Natl Acad Sci ,1996;93: 5910-5914.
- [ 3 ] Yu M, Tong J-H, Mao M, et al. Proc Natl Acad Sci ,1997;94: 7406 - 7411.
- [ 4 ] Maruyama K and Sugano S. Gene ,1994;138:171 - 174.
- [ 5 ] Suzuki Y, Yoshitomo-Nakagawa K, Maruyama K, et al. Gene , 1997;200(1 - 2):149 - 156.
- [ 6 ] Mao M, Fu G, Wu J - S, et al. 1998(submitted) .
- [ 7 ] Kato S, Sekine s, Oh S-W, et al. Gene ,1994;150:243 - 250.
- [ 8 ] Ederly I, Chu LL, Sonenberg N, et al. Mol Cell Biol ,1995;15: 3363 - 3371.
- [ 9 ] Carninci P, Kvam C, Kitamura A, et al. Genomics ,1996;37:327 - 336.
- [ 10 ] Carninci P, Westover A, Aishiyama Y, et al. DNA Res. 1997;4 (1):61 - 66.
- [ 11 ] Carninci P, Nishiyama Y, Westover A, et al. Proc Natl Acad Sci USA ,1998;95(2):520 - 524.
- [ 12 ] Meyerson M, et al. Cell ,1997;90:785 - 795.
- [ 13 ] Schaefer BC. Anal Biochem ,1995;227(2):255 - 273.
- [ 14 ] Capone MC, Grman DM, Ching EP, et al. J Immunol ,1996;157 (3):969 - 973.
- [ 15 ] Rossi DL, Vicari AP, Franz-Bacon K, et al. J Immunol ,1997; 158(3):1033 - 1036.

## Progress of the full-length cDNA cloning strategy in human genome research

Zhang Qinghua Mao Mao Chen Zhu

(Shanghai Institute of Hematology, Shanghai Second Medical University, Key Laboratory of Human Genome Research, Ministry of Public Health and Shanghai, Shanghai 200025, China)

**Abstract** Full-length cDNA cloning is one important aspect of the Human Genome Project (HGP). And it is the main aim of HGP in China, comparing with the large scale genome sequencing, it costs less and may get more functional genes' information relatively. Here, we introduce some new strategies of full-length cDNA library construction, and full-length cDNA cloning with laboratory bench-work and bioinformatics usage.

**Key words** full-length cDNA gene cloning database human genome project

(接第 15 页)

## Study on the Cloning of Animals by Nuclear Transplantation

Fan Biqin

(Jiangsu Academy of Agricultural Sciences, Nanjing 210014 P. R. China)

**Abstract** In this review, it briefly introduces the research history of cloning animals by nuclear transplantation in the world and in China. Methods and procedure are described in detail, including oocyte enucleation (cytoplasmic recipient), preparation of nuclear donor, nuclear transfer, fusion of the donor nucleus with the recipient oocyte, activation of the recipient oocyte, reconstructed embryos culture in vivo and in vitro, and embryo transfer to recipient females for producing offspring. Prospects for the research and applications of cloning animals by nuclear transfer are discussed. Recent results indicate that repeated cloning to produce large number of genetically identical animals for commercial use will be a reality in the near future.

**Key words** Mammal Cloning animals Nuclear transplantation Asexual reproduction