

DNA 克隆载体的发展和应用

刘建喜 林爱星 丁翔 陈永福

(中国农业大学农业生物技术国家重点实验室, 北京 100094)

The Development and Application of DNA Cloning Vectors

Liu Jianxi Lin Aixin Ding Xiang Chen Yongfu

(National Laboratories of Agrobiotechnology, CAU Beijing 100094)

在分子生物学中, 克隆 DNA 片段是一项必备工作。自从 Cohen 等(1973)⁽¹⁾ 引入质粒(Plasmid)pSC101 作克隆载体以来, 克隆载体的整体结构和效率已有了很大的改善。现在使用的载体除了噬菌体(Bacteriophage)⁽²⁾, 粘粒(Cosmid)⁽³⁾ 外, 还有酵母人工染色体(Yeast Artificial Chromosome YAC)⁽⁴⁾, 细菌人工染色体(Bacterial Artificial Chromosome BAC)⁽⁵⁾, 噬菌体 P1 克隆系统(P1 Clone)⁽⁶⁾, 由 P1 发展而来的 PAC 载体(P1-derived Artificial Chromosome PAC)⁽⁷⁾, 以大肠杆菌小 F 因子为基础构建的载体(Mini F-based Plasmid)⁽⁸⁾, 具有 F 因子和粘粒特征的 Fosmid⁽⁹⁾ 和哺乳动物人工染色体(Mammalian Artificial Chromosome MAC)⁽¹⁰⁾。

1 酵母人工染色体(YAC)

要使人工构建的染色体在细胞中能自主复制, 分离到子细胞中去, 具有一定的遗传稳定性, 该染色体至少应具备复制原点(Origin of Replication), 着丝点(Centromere), 端粒(Telomere)。随着来自酵母菌染色体着丝点, 具复制原点功能的自主复制序列(Autonomous Replicating Sequences)分离成功, Murray 等(1983)⁽⁴⁾ 将这三组元件连接在一个载体上, 构建了第一个人工染色体。进一步研究发现人工染色体的长度是影响复制和分离的重要因素。Burke 等(1987)⁽¹¹⁾ 构建了一个 YAC 载体, 将人的 DNA 大片段插入 YAC, 插入 DNA 的长度至少是噬菌体和粘粒中插入片段的 10 倍(表 1)。

表 1 DNA 克隆载体的比较

载体	结构	宿主细胞	插入片段长度(kb)	载体	结构	宿主细胞	插入片段长度(kb)
Pasmid	环状质粒	大肠杆菌	7~10	Mini F-Based			
Bacteriophage	线状 DNA	大肠杆菌	17~20	Plasmid	环状质粒	大肠杆菌	90~105
Fosmid	环状质粒	大肠杆菌	35~45	BAC	环状质粒	大肠杆菌	< 350
Cosmid	环状质粒	大肠杆菌	35~47	YAC	线状 DNA	酵母菌	100~2 000
P1 Clones	环状质粒	大肠杆菌	70~100	MAC	线状 DNA	哺乳类细胞	?~10 000
PAC	环状质粒	大肠杆菌	100~300				

YAC 质粒 (pYAC) 用 *Bam*HI 切除 HIS3, 用 *Eco*RI 切开得左右两臂。外源 DNA 用 *Eco*RI 部分酶切, 用脉冲电泳(Pulsed Field Gel Electrophoresis PFE)技术分离到含 *Eco*RI 酶切末端的大片段 DNA, 应用连接酶, 连接得到 YAC, 转化到原生质化的酵母菌内。因插入外源 DNA 使赭石突变的抑制基因(SUP4)失活, 故红色的菌落即为所需的 YAC 克隆。YAC 克隆片段从当初的 70kb(千碱基对)很快上升到 400kb。Larin 等(1991)⁽¹²⁾ 制备大片段 DNA 的过程中加入多聚胺(Polyamine), 使 YAC 的平均长度达到 700kb。现已构建好人的大片段 (>1 000kb) YAC 基因组文库。YAC 插入片段长, 只需较少的克隆数便可覆盖整个基因组(Genome), 克服了一般载体容量不足的缺陷, 使构建高等生物基因组的物理图谱(Physical Map)和遗传图谱(Genetic Map)成为可能。YAC 已成为

基因组工程的主要载体和基因组分析的有力工具。YAC 连续克隆群(Contigs: Continuous Sets of Clones)已覆盖了人、猪、牛、绵羊、小鼠、果蝇等动物的基因组。

应用 YAC 载体使研究人员在事先不知道相关基因的功能信息的条件下能直接从染色体的特定位置的上克隆某一基因,即定位克隆(Positional Cloning)。首先从连锁图谱(Linkage Map)分析确定靶位点(Target Locus),接着建立 YAC 克隆,作好精细的物理图谱和基因序列图谱,从中确定靶基因。Vaiman 等(1997)通过流式细胞分离仪分离到特定的染色体^[13],构建特异染色体文库,获取了大量的遗传标记(Genetic Markers)。人 YAC 库和放射杂交(Radiation Hybrid)技术的联和使用有利于 STS(sequence Tagged Sites)和 EST(Expressed Sequence Tags)筛选。大量人的 STS、EST 的获取,通过比较基因组作图(Comparative Genome Mapping)可以带动整个动物基因组工程迅速发展。

由于 YAC 的插入片段长,能获得完整的大而复杂的基因,包括远处的调控序列,因此 YAC 也是基因分析的有力工具。将 YAC 导入哺乳动物细胞的方法有:(1)把含有特定基因的 YAC 通过细胞融合(Cell Fusion)(含目的基因 YAC 的酵母菌原生质体再和哺乳动物细胞融合)。这在小鼠和中国仓鼠中已获成功。(2)通过 PFE 将 YAC 从酵母菌分离出来,然后通过脂质体介导(Lipofection)或显微注射(Microinjection)导入哺乳动物细胞。纯化 YAC 时要非常小心,以防 YAC 被剪切。显微注射大片段 DNA 生产的转基因小鼠传代已获成功^[14]。生产 YAC 转基因动物可为研究基因和基因组的结构和表达调控提供实验模型。

YAC 中克隆片段长,但 YAC 稳定性不好,操作也不方便。(1)YAC 有嵌合现象(Chimaerism)。一个 YAC 中克隆的 DNA 片段可能来自两个或多个不同的染色体上。在基因组文库(Genomic Library)中,嵌合体占克隆总数的 5~50%。由特定的染色体构建的文库,嵌合体占克隆总数的 5~15%。嵌合现象严重妨碍染色体步行(Chromosome Walking)和基因分离(Gene Isolation)。通过荧光原位杂交(Fluorescence *in situ* Hybridization)和细致的图谱分析等方法可将嵌合体检出。(2)YAC 内部有重组现象(Rearrangement)。插入的 DNA 较大,序列发生重排,导致和原来染色体的序列不一致。重组很难检出。通过使用有重组缺陷的菌株,可使重组现象和嵌合现象明显降低。(3)YAC 还有缺失(Deletion)现象,影响 YAC 库的代表性。(4)YAC 结构和酵母天然染色体结构相似,使用常规方法不易将 YAC 和酵母天然染色体分开,要通过较长时间的 PFE 才能分离出 YAC。(5)构建好的 YAC 转化原生质体的酵母菌,转化效率低。Burgess 等(1987)^[15]报道可达 10^{10} 转化子/ μg DNA,但大部分报道远低于这个水平。(6)建好库后,保存方便,但要筛选某个基因时,工作量大。

2 细菌人工染色体(BAC), Mini F-based Plasmid, Fosmid

BAC, Mini F-based plasmid, Fosmid 都具有大肠杆菌 F 因子遗传稳定的优点,但构建原理和插入 DNA 片段的能力不一样。Shizuya 等(1992)^[5]以大肠杆菌 F 因子为基础构建 BAC 载体, pBAC108L。BAC 采用 F 因子的复制原点,以氯霉素抗性基因作为选择标记。为进一步方便克隆的筛选,在 pBAC108L 的基础上插入 pGEM3z 的 lacZ 基因片段。当外源 DNA 成功插入 LacZ 时, LacZ 基因插入失活,在含 X-gal、IPTG 的平板上,可通过蓝白斑法筛选白色克隆。BAC 的插入片段小于 350kb,不如 YAC 克隆片段的长,可能的原因有:(1)大肠杆菌本身的基因组只有 4Mb,要复制 1Mb 的 BAC 似乎不可能。(2)电转化法限制大片段 DNA 进入大肠杆菌。BAC 的插入 DNA 片段虽不如 YAC 长,但其有很多优点可弥补 YAC 的不足。(1)BAC 的复制子来源于 F 因子(单拷贝复制),因此, BAC 在宿主菌内只有极少拷贝数,即使传代 100 代以后,仍可稳定遗传,尚未检出缺失、重组、嵌合现象。(2)BAC 以大肠杆菌为宿主,转化效率高。电转化可达 10^9 转化子/ μg DNA。从商业途径获取的感受态细胞,以 pUC 质粒作电转化,转化效率可达到 10^{10} 转化子/ μg DNA,至少比酵母菌的转化效率高 1000 倍。(3)BAC 以环状形式存在于大肠杆菌,转化后分离 BAC 相当方便,使用一般的碱变性法即可做到。使用 PFE 分析 BAC 所花的时间要比 YAC 的少。(4)筛选目的基因方便,可使用菌落原位杂交(Clone *in*

situ Hybridization)方法筛选目的基因。现已构建人、牛、小鼠、水稻、高粱等生物 BAC 基因组文库。绵羊、猪、山羊的 BAC 基因组文库正在构建。Venter 等 (1996)^[16] 提出了一个新的基于 BAC 库的基因组测序策略。该方法相对省时省力,有可能促进人类基因组计划的完成。

Mini F-based Plasmid 是 Hosoda 等(1990)^[8] 使用大肠杆菌 F 因子为基础发展的克隆载体。首先将 PFE 分离获得的大片段目的 DNA 两端连上人工 (cos 右臂,以 F 因子为基础构建的载体两端连上人工 λ cos 的左臂,然后退火连接起来,导入宿主扩增。该系统可克隆较大片段 DNA,并可稳定遗传。Fosmid 是 Kim 等(1992)^[9] 将 pBAC 引入 pUCcos 融合后构建的载体系统。插入片段长度与 Cosmid 接近(见表 1)。现已知道高等生物的基因组 DNA 在大肠杆菌内以多拷贝形式存在时,大部分是不稳定的。用 Fosmid 克隆系统可以构建稳定复杂基因文库,并可用于高等生物基因的最终的特征分析和序列分析。

3 P1 克隆(P1 Clones)和源于 P1 的 PAC(P1-derived Artificial Chromosome)

Sternberg 等(1990)^[6] 发展了噬菌体 P1 克隆系统。P1 克隆可用来分离和扩增最大至 100Kb 的 DNA 片段。该系统利用噬菌体 P1 的包装位点(pac)将 DNA 包装到噬菌体颗粒内(第一次包装),然后噬菌体将 DNA 注入大肠杆菌(第二次包装)。通过噬菌体 P1 的 2 个重组位点 lox P 及在宿主体内表达的 P1 Cre 重组酶使包装的 DNA 环化。P1 克隆系统通过两次包装能获得较高的转化率。在含有 lacI 菌株中,P1 质粒复制子在每个宿主染色体上只有一个拷贝,保证了目的 DNA 的稳定遗传。当加入 IPTG 后,利用 lac 启动子调控 P1 裂解,使目的基因得以扩增。现已建立人、鼠生物等的 P1 克隆文库。Ioannou 等(1994)^[7] 发展了源于 P1 的人工染色体 PAC。PAC 载体以 F 因子和噬菌体 P1 为基础构建,具有 P1 和 F 因子的特点,通过电转化可将 PAC 导入大肠杆菌。插入的 DNA 没有明显的嵌合和缺失现象。PAC 载体不仅有较高的遗传稳定性,还能高效扩增。

Gingrich 等 (1994)^[17] 以 PAC 作载体,建立了人的第二条染色体的克隆文库。P1 克隆系统和 PAC 系统,在基因分离和序列分析当中,可作为 YAC 连续克隆群的重要补充。

4 哺乳动物人工染色体(MAC)

MAC 和 YAC 一样必须具备三个基本组分,即复制原点、着丝点和两个端粒^[10]。在哺乳动物中,端粒是唯一分离,并得到可靠分析的部分^[18]。在哺乳动物基因组中,复制原点有多个,大约 50~300kb 便有一个。对着丝点的原位杂交分析显示,在染色体的主缢痕区域中心有不同的卫星 DNA。在人染色体的主缢痕区有高拷贝的(-卫星 DNA(Alpha-satellite DNA)。大量实验说明(-卫星 DNA 在着丝点中起着重要的作用,将(-卫星 DNA 引入细胞,整合到染色体上后将会诱发或干扰染色体的分离^[19]。

总而言之,对哺乳动物染色体的着丝点和复制原点的研究尚不透彻。进展缓慢的原因有:(1)染色体的着丝点远大于端粒。端粒 DNA 只有几个 kb 长,而着丝点则有数百个 kb 长。分离该区域必须使用大片段克隆载体 YAC,但(-卫星及类(-卫星 DNA(Alphoid-DNA)在酵母菌中是不稳定的^[20]。(2)通过物理图谱方法确定完整的 MAC 所需的最小片段和必须序列最终还需作体内功能分析。

构建 MAC 有两种策略。第一种策略和构建 YAC 相似,即在体外将各个组份相连。运用这种策略目前尚有两方面的制约因素:(1)端粒已分离,复制原点的广泛存在,从结构成分上而言,只有着丝点无法获取了。(2)即使将各种成分连接起来,插入大片段 DNA 后,要将 MAC 完整地导入哺乳动物细胞也并非易事。第二种策略是将已分离到的端粒,在体内将之连接在染色体着丝点的附近,构成小染色体。Farr 等(1995)^[21] 以第二种策略构建了一个 MAC。将含有端粒和 HPRT 基因的质粒 pHpTS1 电转化进入人仓鼠杂合体细胞株 HyTM1/36。HyTM1/36 细胞中的 Xp22.23 的 MIC2 位点有 HISD 基因(细菌组氨醇脱氢酶基因 Bacterial Dehydrogenase Gene)。X 染色体的 q 端有 Hyg(Hygmycin)基因、端粒,近 q 端有(-卫星 DNA。质粒 pHpTS1 进入细胞后随机

整合到 X 染色体上。首先在含 Hyg 和 HAT 的培养基上将能在 Hyg 和 HAT 培养基中生长的克隆筛选出来。接着加入组氨酸(Histidinol)和 5-溴尿嘧啶, 具有 HISD 基因的克隆大量增殖, 光照处理。掺入 5-BrU 的克隆被杀死。由于 p 端和 pHpT51 发生交换, 不具备 HISD 基因, 不能利用组氨酸, 未掺入 5-BrU, 光照后不会致死, 富集起来。建立细胞系, 供进一步研究使用。

MAC 将在基因治疗(Gene Therapy), 染色体和基因结构分析当中发挥巨大的作用。使用 MAC 有很多优点: (1) MAC 在宿主细胞内可以自主复制, 不会因为以插入的方式整合到宿主染色体导致插入突变而影响宿主染色体的功能。(2) MAC 可以转移大片的 DNA, 使之有能力容纳大片的基因及其跨越上兆碱基对的调控成分。(3) 有的遗传病是多基因所致, 称为连续多基因缺失综合症(Contiguous Gene Deletion Syndrome), MAC 有能力容纳整簇的基因, 因而 MAC 能同时纠正多个位点不同基因缺失所致的遗传病。

参 考 文 献

- 1 Cohen S N, *et al.* Construction of biologically functional bacterial *in vitro*. Proc. Natl. Acad. Sci. USA, 1973, 70: 3240~3244
- 2 Murry N E, Murry K. Manipulation of restriction targets in λ phage to form receptor chromosomes for DNA fragments. Nature, 1974, 251: 476~481
- 3 Royal A A, *et al.* The ovalbumin gene region: common features in the organization of three genes expressed in chicken oviduct under hormonal control. Nature, 1979, 279: 125~136
- 4 Murry A W, Szostak J W. Construction of artificial chromosome in yeast. Nature, 1983, 305: 189~193
- 5 Shizuya H, *et al.* Cloning and stable maintenance of 300 kilobase-pair fragments of human DNA in Escherichia coli using an F-factor-based vector. Proc. Natl. Acad. Sci. USA, 1992, 89: 8794~8797
- 6 Sternberg N. Bacteriophage P1 cloning system for the isolation, amplification, and recovery of DNA fragments as large as 100 kilobase pairs. Proc. Natl. Acad. Sci. USA, 1990, 87: 103~107
- 7 Ioannou P A, *et al.* A new bacteriophage P1-derived vector for the propagation of large human DNA fragments Nature Genet., 1994, 6: 84~89
- 8 Hosoda F, *et al.* An F factor based cloning system for large DNA fragments. Nucleic Acids Res., 1990, 18: 3863~3869
- 9 Kim U J, *et al.* Stable propagation of cosmid sized human DNA inserts in an F factor based vector. Nucleic Acids Res., 1992, 20: 1083~1085
- 10 Farr C, *et al.* Functional reintroduction of human telomeres into mammalian cells. Proc. Natl. Acad. Sci. USA, 1991, 88: 7006~7010
- 11 Burke D T, *et al.* Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vector. Science, 1987, 236: 806~812
- 12 Larin Z, *et al.* Yeast artificial chromosome libraries containing large inserts from mouse and human DNA. Proc. Natl. Acad. Sci. USA, 1991, 88: 4123~4127
- 13 Vaiman D, *et al.* Mass production of genetic markers from a limited number of sorted chromosomes. Mammalian Genome, 1997, 8: 153~156
- 14 Schedl A, *et al.* Transgenic mice generated by pronuclear injection of a Yeast artificial chromosome. Nucleic Acids Res., 1992, 20: 3073~3077
- 15 Burgess P M, Kimberly J. Transformation of yeast spheroplasts without cell fusion. Anal. Biochem., 1987, 163: 391~397
- 16 Venter J, *et al.* A new strategy for genome sequencing. Nature, 1996, 381: 363~366
- 17 Gingrich J C, *et al.* Construction and characterization of human 2-specific cosmid, fosmid, and PAC clone libraries. Genomics, 1996, 32: 65~74
- 18 Brown W A. Molecular cloning of human telomeres in yeast. Nature, 1989, 338: 774~776
- 19 Haaf T, *et al.* Integration of human (-satellite DNA into simian chromosome: centromere protein binding and disruption of normal chromosome segregation. Cell, 1992, 70: 681~698
- 20 Neil D L, *et al.* Structural instability of human tandemly repeated DNA sequences cloned in yeast artificial chromosome. Nucleic Acids Res., 1990, 18: 1412~1418
- 21 Farr C J, *et al.* Generation of human X-derived minichromosome using telomere-associated chromosome fragmentation. The EMBO, 1995, 14: 5444~5454