

# Taq DNA 耐热聚合酶在大肠杆菌中的克隆和高表达

杜汉森 季朝能 徐 迈 马匆匆  
郑佐华 毛裕民

(复旦大学遗传学研究所 上海 200433)

从水生栖热菌 (*Thermus aquaticus*) YT-1 中分离得到的 Taq DNA 聚合酶是一种广泛应用于 PCR 的耐热 DNA 聚合酶。由于天然菌株酶产量较低,培养条件要求严格,酶的纯化过程极为繁琐而使产品成本较高,因而促使人们构建适合于大规模生产的基因工程菌株。已有人分别通过不同的途径,使用不同的载体,成功地在*大肠杆菌*菌株中表达了 Taq 耐热 DNA 聚合酶基因<sup>[1,2]</sup>。我们通过与前人不同的途径,把这一基因克隆到*大肠杆菌*的载体质粒 pJLA503 上并使其得以高表达,为降低生产成本和进一步研究该酶的各种特性提供了有利条件。

## 1 材料和方法

### 1.1 材料

*大肠杆菌* TG1 (SupE F' [traD36 proAB<sup>+</sup> lacI<sup>n</sup> lacZ  $\Delta$  M15]), MV1184 (F' [traD36 proAB<sup>+</sup> lacI<sup>n</sup> lacZ  $\Delta$  M154]), CJ236 (dut1 ung1 thi-1relA 1/pCJ105 [cam<sup>r</sup> F']), 质粒 pUC118 和 pUC119、高表达载体 pJLA503、噬菌体 M13mp18 和辅助噬菌体 M13KO7 均系本实验室保存。各种限制酶、T4 DNA Ligase、T4 DNA Pol、T4 DNA Kinase、Klenow 大片段、Mung Bean 核酸酶、Exonuclease III 购自美国 New England Biolabs 或 Promega 公司。T7 DNA 测序试剂盒购自瑞典 Pharmacia 公司,FD DNA 聚合酶由复旦大学复华生物技术公司提供,PCR 扩增引物和碱基置换突变引物由复旦大学遗传所合成,X-Gal 和 IPTG 均为美国 Sigma 公司的产品。

### 1.2 DNA 样品制备

质粒双链 DNA 少量快速制备按文献 [3] 进行。大量制备时在此基础上以等体积 PEG (13%PEG 8000、1.6mol/L NaCl) 沉淀。对于经常使用的质粒 pUC118 和 pUC119 用超离心的方法纯化。M13 噬菌体和质粒单链 DNA 的制备按文献 [3] 进行。DNA 限制性酶切片段纯化采用 DEAE 纤维素膜插片法<sup>[3]</sup>。

### 1.3 DNA 体外操作

限制性酶切和连接:DNA 浓度一般为 0.05 $\mu$ g/ $\mu$ l,限制性酶切时,1 $\mu$ g DNA 加 10u 左右的酶,酶切 1~2h;进行连接反应时,因 DNA 的末端特性不同而适当改变酶量和反应时间<sup>[6]</sup>。

尿嘧啶参入法定点突变见文献[4],体外 PCR 扩增见文献 [3]。

### 1.4 DNA 序列分析

序列分析采用双脱氧方法进行<sup>[5]</sup>。单链和双链样品制备按 Pharmacia 公司的 T7 Sequencing 试剂盒所附说明书进行。电泳系统为 Bio-Red 的 Sequencing Cell,采用模型胶。DNA 序列分析采用 PCGENE 软件包 (Intelli Genetics, Inc. Release 6.60, 1990) 里的程序。

### 1.5 DNA 耐热聚合酶的酶活测定

酶的活力单位定义:在 70 $^{\circ}$ C 中反应 30min,催化掺入 10 nmol/L 的 dNTP 到酸不溶性物质中的酶量为 1 u。

本文于 1994 年 12 月 5 日收到。

酶活测定:取 5 $\mu$ l 待测样品,加反应液 95 $\mu$ l (反应液:67 mmol/L 的 Tris · HCl (pH 8.0)、6.7 mmol/L MgCl<sub>2</sub>、0.01 mol/L  $\beta$ -巯基乙醇、13  $\mu$ g 经 DNase I 处理的小牛胸腺 DNA、0.03 mmol/L dNTP、0.013  $\mu$ mol/L <sup>3</sup>H-dATP),置于 70°C 水浴中反应 30min,加入 0.6ml 7% 预冷 TCA 以终止反应。将反应物转至滤膜上抽干,70°C 真空干燥 20min,取出滤膜放入闪烁杯中,用液体闪烁计数器测 CPM。酶活 (u/ $\mu$ l) 按公式 (CPM-200)  $\div$  12 500  $\div$  5  $\times$  稀释倍数计算。

## 1.6 SDS-PAGE 电泳

采用 Laemmli 的高 pH 不连续缓冲系统,胶浓度为 8%,具体参照文献 [3]。

## 2 结果和讨论

### 2.1 Taq DNA 聚合酶基因在 *E. coli* 中的克隆

2.1.1 YT-1 染色体基因文库的构建:抽提质粒 pUC118 和质粒 pCS4 (克隆有大肠杆菌一段未知序列的 pBR322 质粒),分别用 BamHI 和 PstI 进行双酶切,把二者连接后转化 TG1,从转化子中筛选得到一个大小为 7.7kb 左右的质粒,它是在质粒 pUC118 上克隆了质粒 pCS4 一个约为 4.5kb 的片段的产物,命名为 pFDF1。pFDF1 双酶切后电泳回收 pUC118 载体部分,作为建立 YT-1 基因文库的载体,可基本上清除载体自连。根据已知的酶切图谱<sup>[1]</sup>可以看到在 Taq 耐热 DNA 聚合酶基因的两端有 Bgl II 和 Pst I 酶切位点各一个,在 Taq 聚合酶基因的内部也有一个 Pst I 切点。抽提 YT-1 染色体 DNA,用 Bgl II 完全酶切后再用 Pst I 部分酶切,电泳回收 4~9kb 大小的片段,同时将质粒 pFDF1 用 BamHI 和 Pst I 双酶切,电泳回收小片段,利用 BamHI 和 Bgl II 酶切产生的相同粘性末端,把两种电泳回收片段连接后转化 TG1,构建了包含 6000 个转化子的 YT-1 基因文库。

2.1.2 从文库中筛选含 Taq 耐热 DNA 聚合酶基因的克隆:通过 PCR 方法从文库中筛选阳性克隆具有简便、迅速、灵敏并且不用接触同位素等优点,所以我们选用 PCR 方法从 YT-1 基因文库中筛选含 Taq 耐热 DNA 聚合酶基因的克隆。通过比较 Taq 耐热 DNA 聚合酶基因、T7 DNA 聚合酶基因和 Klenow DNA 聚合酶基因的核苷酸序列,在 DNA 聚合酶基因的保守区域设计了 8 个引物,配成 4 对,准备用于 Taq 耐热 DNA 聚合酶基因以及其他一些 DNA 聚合酶基因的克隆。这里它们被用于 Taq 耐热 DNA 聚合酶基因的克隆。这 8 个引物依次命名为 PTAQ1~8,它们的核苷酸序列分别是:

PTAQ1. 5' ATGAGGGGGATGCTGCCCTCTTTGAGCCCAAGG 3';

PTAQ2. 5' TCACTCCTTGGCGGAGAGCCAGTCCTCC 3';

PTAQ3. 5' GTGGTCTTTGACGCCAAG 3';

PTAQ4. 5' CCTCGT AGCCCTCCTTTTC 3';

PTAQ5. 5' CGGTAGTCGGCCCACTGGTC 3';

PTAQ6. 5' TTCAACATGCCCGTCCAG GG 3';

PTAQ7. 5' GGGGAGTGGACGGAGGAGG 3';

PTAQ8. 5' CTGGTTGAAGCGGGTGTGGAG 3';

预期其中 PTAQ1、5 配对或 3、4 配对能扩增出 Taq 耐热 DNA 聚合酶基因的头部片段,PTAQ2、6 配对能扩增出 Taq 耐热 DNA 聚合酶基因的尾部片段,PTAQ7、8 配对能扩增出 Taq 耐热 DNA 聚合酶基因的中间片段。首先以 YT-1 的染色体 DNA 为模板,用这四对引物进行 PCR 扩增,发现 PTAQ3、4 配对、PTAQ2、6 配对或 PTAQ7、8 配对均能扩增出相应的 DNA 片段,而 PTAQ1、5 配对则不能扩增相应的 DNA 片段。分析 Taq 基因的序列,可以看到在基因的头部,尤其在突变引物的位置 GC 含量很高,有多处连续出现 4、5 个 G 或 C。显然在这些位置上容易形成二级结构,由于退火温度受引物 T<sub>m</sub> 的限制,在一定温度下导致引物不能很好地和模板结合或链不能延伸。David R Engelke 在用 PCR 方法克隆 Taq 基因时所用引物 T<sub>m</sub> 值很高,可在 72°C 退火<sup>[2]</sup>,这从一个侧面也说明了二级结构的存在。这些结果为我们以后用 PCR 方法克隆别的 DNA 聚合酶基因时提供了借鉴和经验。用设计好的引物对 YT-1 文库进行三轮 PCR 筛选,以期获得完整的 Taq 基因:先以文库中的转化子每 100 个为一组,混合

抽提每组的质粒 DNA, 以 PTAQ3、4 作为引物进行 PCR 扩增, 筛选出能扩增预期片段的样品, 再在这一组中以 10 个转化子为一组, 按同样方法扩增, 选出阳性组; 最后对 10 个转化子的每一个进行扩增, 筛选出含有 Taq 基因头部但长度不等的 7 个克隆。由于采用了部分酶切的方法构建文库, 基因的 3' 端有可能缺失, 因此再用 PTAQ2、6 和 7、8 为引物对这 7 个克隆子进行 PCR 扩增, 筛选出能扩增的阳性克隆。PTAQ2、6 能扩增的阳性克隆中 Taq 基因的尾部是完整的, PTAQ7、8 能扩增的阳性克隆中 Taq 基因内部的 Pst I 切点位置未插入其他的 DNA 片段, 因此这些克隆可以认为含有完整的 Taq 基因。选择其中一个, 利用公用引物 1224 和 PTAQ2, 进行了序列测定, 发现我们克隆到的片段的两端序列与文献报道的 Taq DNA 聚合酶基因序列的头部和尾部完全相同, 从而确证了这是一个完整的基因克隆。我们将含有 Taq 基因的质粒命名为 pFDT1。

2.1.3 Taq DNA 聚合酶基因的亚克隆:

pFDT1 在大肠杆菌中虽能表达, 但酶的产率很低。Taq 耐热 DNA 聚合酶基因的表达只是利用了载体 pUC118 上 LaxZ- $\alpha$  基因的启动子。为此我们以高表达质粒 pJLA503 作为载体进行 Taq 耐热 DNA 聚合酶基因的亚克隆。

pJLA503 是一个高表达载体<sup>(7)</sup>, 它含有大肠杆菌的 atpE 翻译起始区域, fd 转录终止子, P<sub>R</sub> 及 P<sub>L</sub> 串联启动子。多克隆位点上有一个 Nde I 酶切位点 (5'CATATG 3')。为

了能用它克隆 Taq DNA 聚合酶基因, 通过尿嘧啶单链定点突变法在基因的 5' 端起始密码处产生了一个 Nde I 酶切位点, 通过 PCR 介导的双链定点突变法在 3' 端产生了 Sph I、Sal I、EcoR I 三种酶切位点。

5' 端的改造: 利用基因的内部和载体上的 Kpn I 酶切位点对 pFDT1 进行酶切, 回收带有 Taq 聚合酶 5' 端基因的小片段, 把它克隆到 M13mp18 上, 转化 CJ236, 从中抽提参入尿嘧啶的单链。另外设计命名为 PTAQ9 的突变引物 (5' GGTAACATGAGGCATATGCTGCCCT 3', 其中黑体是突变碱基, 原来序列为 GGG)。以 PTAQ9 和参入尿嘧啶的单链模板配对进行体外定点突变, 取得带有 5' 端翻译起始密码处产生 Nde I 切点的 Taq DNA 聚合酶基因片段的 M13mp18TAQN, 经序列测定表明 M13mp18TAQN 上的 Taq 耐热 DNA 聚合酶基因片段带有预期突变, 并未发生额外突变。

3' 端的改造: 设计一长 55bp 的突变引物 PTAQ10 (图 3), 该引物 3' 端的 30 个核苷酸顺序与 Taq 耐热 DNA 聚合酶基因尾部的序列相同, 5' 端包含 Sph I、Sal I、EcoR I 三个酶切位点及二个不同读框的翻译终止密码子的核苷酸序列。以 PTAQ10 及 PTAQ6 为引物进行 PCR 扩增得到长 280bp 的片段。该片段经 Xho I 和 Sal I 双酶切后, 克隆到 Phagenid Bluescript M13<sup>+</sup>SK 上进行测

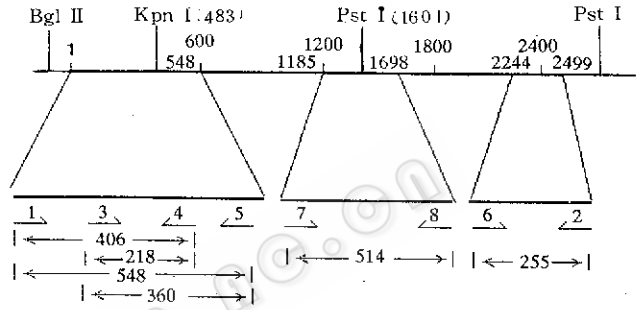


图 1 PCR 引物位置及酶切图谱示意图

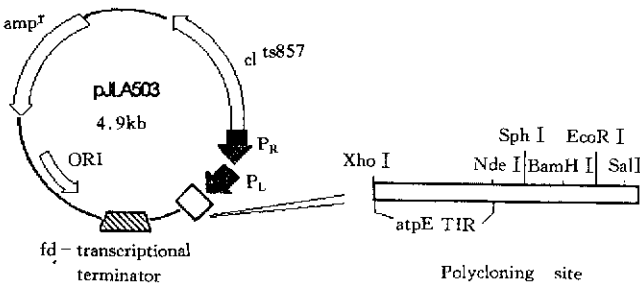


图 2 pJLA503 的示意图<sup>(7)</sup>

序, 结果表明在 Taq 耐热 DNA 聚合酶基因的 3' 端已发生了所需要的突变。将带有 3' 端突变的重组质粒用 Hind III 和 Xho I 双酶切, 割胶回收长 180bp 的片段, 同时将 pFDT1 用 EcoR I 完全酶切, Xho I 部分酶切后割胶回收长 2.4kb 的片段, 质粒 pUC118 经 EcoR I 和 Hind III 双酶切, 割胶回收大片段, 三个片段连接后转化 TG1, 筛选得到 Taq 耐热 DNA 聚合酶基因尾部发生了突变的质粒 pFDT2。

3' CCCTCCTGACCGAAAGGCGGTTCTCACTTAAGATCGATCAGCTGCGTACGCGT 5'

EcoRI

Sal I

Sph I

Stop

Coden Stop

Coden

图3 PTAQ10的序列

对基因的5'端和3'端进行改造以后,又把它们与基因的中间片段拼接成完整的Taq聚合酶基因并克隆到高表达载体pJLA503上。由于抽提噬菌体M13mp18的双链DNA效率不高,且经常混有单链DNA,因此,我们先把基因克隆到高拷贝的pUC118上。于是在pUC118的多克隆位点上插入了一段含Nde I识别位点的序列,构建成中间载体质粒pFDV118N。M13mpTAQN和pFDV118N分别用Kpn I和Nde I双酶切后连接,将已发生突变的Taq耐热DNA聚合酶基因5'端约450bp大小的片段直接克隆到pFDV118N上,构成pFDV118NK。取pFDV118NK用Nde I和Acc I双酶切,割胶回收长120bp的片段。pFDT2用Acc I和Sal I双酶切,割胶回收长约2.5kb的片段。pJLA503则用Nde I和Sal I双酶切,割胶回收长4.9kb的片段。连接上述三个片段,转化TG1,构建含完整的Taq耐热DNA聚合酶基因的高表达质粒pFDT3。经多种酶切鉴定证明其结构正确(图4)。



图4 pFDT3的酶切鉴定

1.  $\lambda$ Bst I marker. 2. pJLA503 cut by Nde I ;
3. pFDT3 cut Nde I ; 4. pJLA503 cut by Acc I ;
5. pFDT3 cut by Acc I ; 6. pJLA503 cut by Xho I ;
7. pFDT3 cut by Xho I

1 2 3 4 5 6 7 8

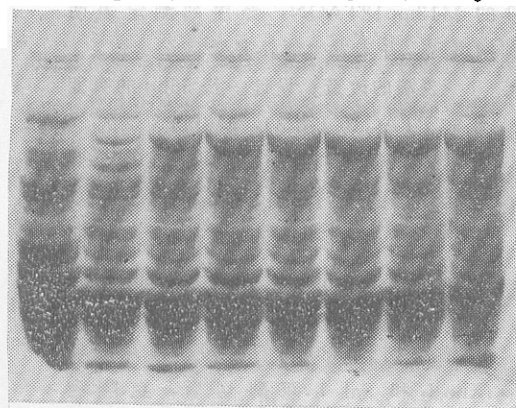


图5 不同诱导时间下粗酶液的SDS-PAGE电泳

1. pJLA503 40°C 4h. 2. pFDT3 42°C 2h, 4h, 6h,
- 8h, 10h, 12h, 14h

## 2.2 Taq耐热DNA聚合酶在大肠杆菌中的高表达

从TG1(pFDT3)平板上挑取菌落,接种于3ml 2YT培养基中,30°C培养过夜作为种子液。取种子液按2%接种量转接至50ml 2YT培养基中,30°C培养2h后加入经预热的等体积2YT培养液,在42°C进行热诱导。诱导不同时间后取出菌液,5000r/min离心5min收集菌体并称其湿重,每克湿菌中加入8ml磷酸缓冲液悬浮菌体,在冰浴中超声波破菌,4°C 15000r/min离心15min后测定上清液的酶活和蛋白质含量,不同诱导时间的酶活和蛋白质含量测定结果见表1。

由表可见,不经42°C热诱导的TG1(pFDT3)的上清液中酶活很低,随诱导时间的增加,表达量稳定上升,大约12h比活达到高峰。同样实验条件下TG1(pJLA503)不出现酶活。取不同样品的上清液走

SDS-PAGE 凝胶电泳, 染色后

可看到明显的酶带, 且酶蛋白的表达随着诱导时间增加而增高 (图5)。

工程菌 TG1 (pFDT3) 的未纯化酶的比活可达 4472u/mg 蛋白, 水生栖热菌 YT-1 的未纯化酶的比活只有 2.13u/mg 蛋白<sup>[8](10)</sup>, David R. Engelke 构建的工程菌 pTTQ18 的未纯化酶的比活为 827u/mg 蛋白<sup>[2,9]</sup>, 与他们的工作相比, 该菌株具有巨大的生产潜力。

表1 不同时间诱导后的 DNA 聚合酶活性

	Induced time/h	Protein/mg	Enzyme activity/u
TG1 (pFDT3)	0	13.53	$8.56 \times 10^2$
TG1 (pFDT3)	2	14.69	$8.36 \times 10^3$
TG1 (pFDT3)	4	20.82	$2.62 \times 10^4$
TG1 (pFDT3)	6	29.46	$4.55 \times 10^4$
TG1 (pFDT3)	8	21.38	$6.45 \times 10^4$
TG1 (pFDT3)	10	23.62	$7.83 \times 10^4$
TG1 (pFDT3)	12	21.52	$9.63 \times 10^4$
TG1 (pFDT3)	14	21.75	$7.82 \times 10^4$
TG1 (pJLA503)	14	22.47	0

致 谢 盛祖嘉先生多次提出建设性意见和修改此文, 特此致谢。

### 参 考 文 献

- [1] Frances C. Lawyer, Susanne Stoffel, Saik R L *et al.* J Bio Chem, 1989, **264**: 6427~6437.
- [2] David R Engelk, Alexandra Krikos, Mary E Bruck *et al.* Anal Biochem, 1990, **191**: 396~400.
- [3] Sambrook J, Fritsch E F, Maniatis. 《Molecular Cloning》, Cold Spring Harbor Laboratory Press, 1989.
- [4] Kunkel T A. Proc Natl Acad Sci USA, 1985, **82**: 488~492.
- [5] Sanger F *et al.* Proc Natl Acad Sci USA, 1977, **74**: 6008~6013.
- [6] 蔡武城, 袁厚积主编, 《生物物质常用化学分析法》, 上海: 复旦大学出版社, 1983.
- [7] Birgit Schauder, Helmut Blocker, Ronald Frank *et al.* Gene, 1987, **52**: 279~283.
- [8] Chien A, Edgar D B, Trela J M. J Bacteriol, 1976, **127**: 1550~1557.
- [9] Fred G. Pluthero Nucleic Acids Research, 1993, **21** (20): 4850~4851.
- [10] Patrick Argos, Michael G R, Ulrich M G *et al.* Biochemistry, 1979, **18**: 5698~5703.

## Cloning and Expression of Taq DNA Polymerase Gene in *Escherichia coli*

Du Hansen Ji Chaoneng Xu Mai Ma Congcong  
Zhen Zuohua Mao Yumin

(Institute of Genetics, Fudan University, Shanghai 200433)

**Abstract** Using phagemid pUC118, a Genomic library of *Thermus aquaticus* strain YT-1 constructed, and according to the reported sequence of the Taq thermostable DNA polymerase gene, eight primers were designed and synthesized from which four pair were used for PCR reaction. Through three rounds of screening, a clone which contain the whole Taq gene was obtained. In addition, two mutagenic primer was synthesized according to the nucleotide sequence of the 5'-end and 3'-end fo the gene and Nde I and Sal I cleavage sites were created by means of oligonucleotide-mediated mutagenesis. The Taq gene was then subcloned in-frame into the high expression vector pJLA503 making use of these cleavage sites. With the transcription and translation system of pJLA503, the expression of Taq DNA polymerase in *E. coli* can reach 4472 u/mg protein.

**Key words** Genomic library, PCR amplification, site-directed mutagenesis, enzyme activity