

PCR 引物设计

THE DESIGN OF PRIMER FOR PCR

王艳秋 张培军

(中国科学院海洋研究所实验海洋生物学开放研究实验室 青岛 266071)

多聚酶链式反应(Polymerase Chain Reaction, PCR)又称无细胞分子克隆系统,是一种体外扩增特异 DNA 片断的技术,1985年由美国 Cetus 公司和加利福尼亚大学联合创建,是 90 年代分子生物学领域的一项革命性突破,很快地在分子克隆、遗传病的基因诊断、法医学、考古学等方面得到了广泛应用。

PCR 技术实际上就是在模板 DNA(Template DNA)引物(Primer)和 4 种脱氧核糖核苷酸(dATP, dGTP, dCTP, dTTP)存在的条件下依赖于 DNA 聚合酶(Taqse)的酶促合成反应,其效率和特异性取决于两个方面:一是引物与模板的特异结合;二是多聚酶对引物的有效延伸。引物与模板的特异结合取决于引物的浓度和引物的设计。引物的设计对于 PCR 极为重要,引物设计得不好,就不会得到扩增片断。

正象根据多年来的实践经验,PCR 已经有了标准的反应体系一样,引物的设计也有一些书中可查到的原则,加上笔者在实验过程中对引物设计的一些体会,在此一并总结出来,与大家探讨。

表 1 近 DNA 片断末端的几个限制性内切酶的切割情况^①

酶	切割序列	反应时间(h)	切割效率(%)
BamH I	CGG↓GATCC	2	>90
ECOR I	CGG↓AATTC	2	>90
Hind III	CGA↓AGCTT	2	X
Kpn I	CGGGTAC↓C	2	90
XBA I	GCT↓CTAGA	2	>90

注:切割序列的前两个碱基为保护碱基

对于不同的使用目的,引物设计的要求是不同的,但对所有的引物设计,都有以下几条基本要求^[1,2,3]:1. 引物长度:15~30nt 为宜(是指与模板配对的引物长度,不包括 5' 末端添加的序列)。2. 引物碱基尽可能随机分布。所谓随机分布,就是避免出现嘌呤、嘧啶堆积现象。

1995 年第 5 期

另外,G+C 含量约在 45~55%左右。3. 引物内部不应自身形成二级结构。4. 两个引物之间不能形成互补结构。5. 引物 3' 末端与模板配对。

若有条件则可借助于计算机进行设计。不管有无条件,最初的设计按以下顺序抓住几点,可事半功倍。设计好之后再借助计算机检查一下更好。

1. 引物的碱基序列不应与非扩增区域有同源性。设计之前用计算机检索分析最令人放心。但一般而言,18nt 的序列在人类基因组中(3×10^9 bp)只会出现一次^[2],引物长度 20nt 左右,可保证其特异性。所以首先根据需扩增的目的片断的两端已知序列写出一定长度的引物。

2. 检查引物的 3' 末端。引物的 3' 末端,尤其最后一个碱基应尽可能严格地与模板互补配对。因为 3' 端配对情况影响酶的延伸,决定了合成的特异性和效率。应注意以下几个方面:1) 引物的模板即目的片断两端序列应选择外显子(Extron)部分,因为外显子比内含子(Intron)保守性强。2) 引物 3' 端最后一个碱基应选在密码子的第一个或第二个核苷酸上。由于碱基摆动,大多数密码子具有简并性,编码同一个氨基酸的密码子第三个碱基可以有差异^[4]。3) 若模板不很清楚,引物 3' 端最后一个碱基最好选择 T, G 或 C, 不选 A。国外有资料表明,当末位为 T 时,即使在错配的情况下也能引发链的合成;而末位为 A 时,错配时引发大大降低;G, C 居于中间。由此可以推知,模板很清楚时,选 A 可提高特异性。

3. 检查引物内部及两个引物之间是否有互补配对的情况。这种情况一般是有的,需要判断的是它们是否影响模板与引物的结合。这需要估计模板与引物结合的

① 百万碱基作图,华美手册。

收稿日期:1995 年 4 月 10 日

自由能及引物配对结合的自由能。A=T 由两个氢键连接, G=C 由 3 个氢键连接。一般来说, 出现 3 个碱基的序列互补便有些危险, 尤其在引物 3' 末端, 若 3' 末端出现 3 个碱基的序列互补或碱基为 GC 的两个碱基的序列互补, 只好将这个引物放弃, 重新设计。

4. 若以上几条没什么问题, 可计算一下 T_m , 适当作些调整; T_m 应达到一定范围, 且两个引物的 T_m 差异不应超过 4℃。然后根据 T_m , 确定退火温度。

T_m 是指 DNA 的溶解温度 (Melting temperature) 即 DNA 的双螺旋结构失去一半时的温度。下面介绍计算 T_m 并确定退火温度的两种方法。

1) 定量计算: 碱基长度为 10~40nt 时^①:

$$T_m = [\Delta H / \{\Delta S + S^\circ - R \ln^{(6/4)}\}] - 273.15 + 16.6 \lg [K^+]$$

ΔH , ΔS 分别指最近邻的焓变和熵变。 S° 指初焓。 C 指碱基序列的浓度。 $[K^+]$ 泛指单种离子浓度。此公式适用于引物 T_m 的计算。碱基长度 > 40nt 时:

$T_m = 81.5 + 16.6 \lg [K^+] - 675/I + 0.41(GC\%)$
 I 指 DNA 序列中碱基的个数。 $GC\%$ 指序列中 G+C 碱基所占的百分数。此公式适用于计算 PCR 扩增产物的 T_m 。所需退火温度的计算公式是:

$$T_{anneal} = 0.3T_m(\text{引物}) + 0.7T_m(\text{PCR 产物}) - 14.9$$

2) 经验公式:

$$T_m = 2(A + T) + 4(G + C)$$

(A+T)、(G+C) 分别指碱基的个数。

退火温度应低于 T_m 。退火温度越高, 特异性越高, 但杂交率越低。一般 PCR 采用的条件是退火温度 55℃, 变性温度 94℃, T_m 在 58~70℃ 之间比较合适。但应根据

引物的 T_m 来选择合适的退火温度。

实验中应用经验公式计算方便可靠。

两个引物的 T_m 应尽可能接近。 T_m 相差过大时, 可延长 T_m 低的引物。

5. 引物的随机分布这一要求在一定范围内是有缓冲余地的, 不必太过拘泥于这一点。

6. 使用目的不同, 引物的设计要求不同。若用于克隆, 引物 5' 端最好加上限制性内切酶的酶切位点, 而且最好是双酶切位点, 有利于定向克隆。酶切位点前应加 2~3 个保护碱基, 具体情况具体分析, 一般加 CG。同样, 加入的酶切位点不能与引物内部形成互补结构。

PCR 技术方便迅速有效, 已经成为分子生物学研究领域的常规手段。抓住 PCR 的各个环节, 才会得到满意的结果。引物的设计是第一步, 也是关键性的一步, 对此, 每个做 PCR 的人都应引起足够的重视。大家的经验的不断积累, 将会使 PCR 技术更加完善。

参考文献

- [1] 卢圣栋等, 1993. 现代分子生物学实验技术. 高等教育出版社, 412.
- [2] 倪灿荣, 1992. 免疫组织化学实验新技术及应用. 北京科学技术出版社, 287.
- [3] J. Sambrook *et al.* 1992. 分子克隆实验指南. 科学出版社, 680.
- [4] David Freifelder, 1989. 分子生物学导论. 复旦大学出版社, 109.

^① Hans Willy Muller and James M. Duclos, 1993. *Amplifications*, Perkin almer 10-11.