

简并 PCR 技术及其在基因克隆中的应用

王洪振^{1,2}, 周晓馥^{1,2}, 宋朝霞¹, 刘文广¹, 曾宪录¹

(1. 东北师范大学遗传与细胞研究所, 长春 130024; 2. 吉林师范大学生物系, 四平 136000)

摘要: 本文简要介绍简并 PCR 技术, 包括什么是简并引物, 如何设计简并引物, 进行简并 PCR 的反应条件, 应用简并 PCR 获得全长基因的方法和简并 PCR 技术的应用范围, 并对简并 PCR 技术的局限性及其新进展进行讨论。在此基础上, 简述基因的克隆策略以及简并 PCR 技术在基因克隆中的应用。简并 PCR 技术是寻找和发现“新”基因或蛋白质家族新成员的一种非常有用的工具。

关键词: 简并 PCR; 简并引物; 基因克隆

中图分类号: Q75

文献标识码: A

文章编号: 0253-9772(2003)02-0201-04

Degenerate PCR and Its Application in Gene Cloning

WANG Hong-Zhen^{1,2}, ZHOU Xiao-Fu^{1,2}, SONG Zhao-Xia¹, LIU Wen-Guang¹, ZENG Xian-Lu¹

(1. Institute of Genetics and Cell Biology, Northeast Normal University, Changchun 130024, China;

2. Department of Biology, Jilin Normal University, Siping 136000, China)

Abstract: Degenerate PCR is introduced in this paper, including what is degenerate PCR, how to design degenerate primers, how to optimize degenerate PCR parameters, how to applying degenerate PCR to obtain full-length gene and which fields can apply degenerate PCR. The limits and recent advances of degenerate PCR are also discussed. Based on this introduction, strategies of gene cloning and applications of degenerate PCR in gene cloning are summarized in brief. Degenerate PCR is a very useful tool for searching and discovering new genes and new members of a protein family.

Key words: degenerate PCR; degenerate primer; gene cloning

在生命科学飞速发展的今天, 新基因的发现及其功能的研究一直是生命科学研究中的热点。尤其随着人类基因组计划即将完成, 以及一些模式生物如酵母、线虫、果蝇、小鼠、拟南芥、水稻和一些原核生物的基因组的全序列已经被测定或即将被测定完成的情况下, 在 GenBank 中积累了大量未知功能的 DNA 序列, 识别和鉴定其中的新基因, 以及对其功能开展研究是进一步深入研究的必然。简并 PCR 技术由于具有能够快速、简便和经济鉴别新基因的优点, 必将在这方面得到越来越广泛的应用。下面就简并 PCR 技术及其在基因克隆中的应用作以简要综述。

1 简并 PCR 技术

简并 PCR 技术是 PCR 技术的发展。PCR 技术也称聚

合酶链反应技术, 其原理是模拟 DNA 复制反应, 通过与特定 DNA 区域两端互补的寡核苷酸引物, 在反应管中选择性地复制合成介于两引物之间的 DNA 片段, 因此, 又被称为 DNA 的体外扩增。PCR 技术已经成为分子生物学中的常规技术。PCR 最初是为扩增已知序列设计的, 简并 PCR 技术却能扩增未知序列。当待研究的基因序列不明, 仅知其一部分蛋白质的氨基酸序列, 可根据氨基酸序列合成一组简并引物, 进行 PCR 扩增, 对基因家族中的未知成员(基因)进行分离。下面简要介绍这一技术。

1.1 简并引物

简并 PCR 与常规 PCR 之间的一个重大差别, 是用简并引物代替具有特定序列的特异的 PCR 引物。简并引物的出现是由于遗传密码具有简并性, 这样可以根据氨基酸的保守

收稿日期: 2002-04-29; 修回日期: 2002-07-15

基金项目: 国家自然科学基金(批准号: 30040031)和吉林省自然科学基金(批准号: 20010553)资助项目

作者简介: 王洪振(1975-), 男, 讲师, 在读博士生, 专业方向: 细胞生物学。E-mail: biostudent@sohu.com, Tel: 0431-5269769

通讯作者: 曾宪录(1959-), 男, 教授, 博士生导师, 专业方向: 细胞生物学。E-mail: jiegou@nenu.edu.cn, Tel: 0431-5269769

序列反推到 DNA 水平设计引物。由于大多数氨基酸的遗传密码不止一种,所以引物的部分碱基不能确定。可以根据各种不同的遗传密码规律设定 DNA 和氨基酸之间的相互转换,这样设计出来的引物是将可能编码一个给定氨基酸序列的核苷酸组合设计为一组,实际上是多种序列的混合物,序列的大部分是相同的,但在某些位点有所变化,称之为简并引物^[1]。一个给定引物的简并度可通过每一位置可能的核苷酸数相乘得到。例如,“Ala Asn Ile Lys Met”的引物库简并度为 $4 \times 2 \times 3 \times 2 \times 1 = 48$ (编码简并度: Ala=4, Asn=2, Ile=3, Lys=2, Met=1)。简并寡核苷酸链可独立合成然后混合为库,或者在 DNA 合成仪的程序中在一个位置放置多个核苷酸^[2]。

1.2 简并引物的设计

应用简并 PCR 方法成功与否关键在于简并引物的设计^[1~3],应遵循以下几个原则:(1)考虑所选靶扩增区的进化保守性和简并 PCR 的扩增特异性,选择同源性高且简并性低的氨基酸区域,如蛋氨酸和色氨酸均只有一个密码子,最好不选择具有 4 或 6 个密码子的氨基酸的肽段,简并程度最好不超过 516 倍。但也有人用 1024 个简并度的引物得到很好的结果。

(2)根据生物使用密码子的偏性,选择使用率最高的密码子,进一步限定引物的简并度^[4]。

(3)引物的 3' 端不应存在简并,因为单碱基错配会阻碍延伸。对于大多数氨基酸残基来说,意味着引物 3' 末端不要位于密码子的第三位。除 Met 或 Trp 外,末端密码子的最后一个碱基可以省略。

(4)在一些多义位置使用能与多种成分配合的脱氧次黄苷以降低简并度^[5,6]。也可以使用能与嘌呤或嘧啶配对的稀有碱基。这些稀有碱基是 6H、8H-3,4-二氢嘧啶并(4,5-c)(1,2)-恶嗪-7-酮(P)和 2-氨基-6-甲氧基氨基嘌呤(K)。P 可与 A 和 G 配对,K 可与 C 和 T 配对。已经证明含有简并碱基 P 和 K 的寡核苷酸链可用作 PCR 引物^[7,8]。与此类似,普通碱基 1-(2'-脱氧-β-D-核糖咪唑(M)也可在 DNA 引物中用于多义位点的扩增^[9]。

(5)引物的长度可短到对应 4~6 个氨基酸的长度,一般为 15~20 个核苷酸。

遵循的总的原则为:尽量降低引物的简并度,尤其在 3' 末端或近 3' 末端。现在 PCR 引物设计都通过计算机软件进行^[1]。具有引物设计功能的软件有很多,一般使用软件 Primer Primer 5.0 可以满足设计简并引物的要求。

1.3 简并 PCR 的反应条件

由于简并引物碱基的不确定性,反应条件也相应发生一些变化。Wilks 等人首次检测在 PCR 中使用简并寡核苷酸引物的各种参数(引物长度、简并度和退火和延伸温度)变化产生的效果^[10]。现在简并 PCR 的反应条件主要考虑以下几个方面:

1.3.1 热动力学框架很显著地改变简并 PCR 的成功率。

PCR 扩增条件先在非严谨退火温度(35~45℃)下进行 2~5 个循环,随后在较严谨的退火温度下进行 25~40 个循环。缓和的退火条件使得引物的短互补区域与靶序列杂交^[2]。在前几轮 PCR 循环中把退火时间增加到 3~5 分钟,能够产生质量更好的所求产物。若在退火和延伸温度之间安排 4 或 5 分钟的爬升时间也可获得更高的特异性^[2]。有人将简并 PCR 与 Touchdown PCR^[11] 和巢式 PCR^[12] 结合使用也得到较好的结果。

1.3.2 简并 PCR 应当在比较高的引物浓度下进行,即 1~3 μmol/L 而不是 0.2 μmol/L,因为在反应混合物中的大多数寡聚物并不是被用来引发专一的反应,而只是产生高的背景^[3]。

1.3.3 不要使用具有 3'~5' 外切核酸酶的 DNA 聚合酶,以防简并引物被降解,使用 Taq 酶就能满足要求。

1.3.4 模板一般选用 cDNA 较好,因为没有内含子,可以预测 PCR 片段的大小。

1.4 应用简并 PCR 获得全长基因的方法

主要有两种途径。一种是采用简并 PCR 扩增的 cDNA 小片段作为探针筛选 cDNA 文库,从而获得相应的目的基因^[13];另一种方法是可以以此小片段为基础设计特异引物进行 3' 和 5' cDNA 末端快速扩增(Rapid Amplification cDNA Ends, RACE)^[74,15,16],扩增此 cDNA 的其余部分,从而获得全长基因。

1.5 简并 PCR 技术的应用范围主要在两个方面

1.5.1 当分离出一种新蛋白并测定了其中的一段氨基酸序列,要进一步寻找相应的基因时,可以使用简并 PCR 技术;

1.5.2 用简并 PCR 易于找到具有进化上保守结构域的蛋白质家族新成员的基因。由于几乎所有蛋白质都具有与其他蛋白质的相似性,并经常具有共同的进化起源,目前已知存在超过 500 个具有进化上保守结构域的蛋白质家族,以后还会有更多的蛋白质家族被发现。通过比对许多相关蛋白质的氨基酸序列,便能找到保守的蛋白质结构域,这些结构域可被用来设计简并引物,进行简并 PCR。

1.6 简并 PCR 技术的局限性

1.6.1 由于该引物特异性较低,因此在电泳检测 PCR 产物时往往本底较高,且常含有非特异性条带,最终需要对每一个可疑阳性的 PCR 产物进行克隆测序证实。

1.6.2 模板具有很高的 GC 含量,经常会扩增出许多不正确片段。

1.6.3 模板具有很低的 GC 含量且引物设计太短,使引物解链温度很低。

1.6.4 要寻找的基因在所选择的生物中不存在,是“恐龙基因”(指在进化过程中在某些品系中消失的基因)。

1.6.5 简并引物 PCR 本身有一定的偶然性,条件须摸索。

1.7 简并 PCR 技术的新进展

1998 年, Rose 等人在核酸研究杂志(Nucleic Acid Re-

seach)上发表了题为“共有序列—简并杂合寡核苷酸引物用于远缘相关序列扩增”的论文,在简并引物设计上有很大改进^[17,18]。他们采用共有序列—简并杂合寡核苷酸引物(consensus-degenerate hybrid oligonucleotide primer, CODE-HOP)来克服使用高简并性引物使有效引物浓度下降的缺点,以及使用共有序列引物时出现的引物与模板之间会出现错配的问题。这种杂合引物由 5'非简并共有序列夹和很短的 3'简并核心区组成。3'简并核心区很短,降低了引物简并度。5'非简并共有序列夹在不增加引物简并度的情况下起到稳定 3'简并核心区与模板结合的作用,允许反应在较高的退火温度下进行,提高了简并 PCR 反应的特异性。在初始 PCR 循环中这种引物的共有序列夹和靶序列间可能存在错配,但由于距离 3'羟基延伸位点很远,所以引物与模板间错配不破坏聚合酶的延伸反应。他们采用这种引物先后成功地克隆了一种新的猴痘疱疹病毒的 DNA 聚合酶基因、人基因组中多个反转录酶相似基因和在各种植物中存在的 C⁵ DNA 甲基转移酶同源物基因。他们还将这种引物设计策略变成可以在网络上得到的计算机程序,网址为: <http://blocks.fhrc.org/codehop.html>。相信会有越来越多的人使用这种策略进行尝试。

2 简并 PCR 在基因克隆中的应用

2.1 基因的克隆策略

基因克隆,以及基因产物结构与功能的分析,是基因组研究的重要目的。当前,随着理论的发展和技术的进步,基因克隆的策略已有正向遗传学途径和反求遗传学途径两种。正向遗传学途径指的是通过被克隆基因的产物或表型突变去进行;反求遗传学途径指的是依据被克隆基因在染色体上的位置来实现^[19]。由于绝大多数控制重要性状的基因产物及调控机制尚不清楚,走向正向遗传学途径很受限,而反求遗传学途径则显示出较广泛的前景。

采取反求遗传学途径主要有三种方法,分别是图位克隆法、转座子标签法和随机突变体筛选法。随机突变体筛选法由于随机性大且不能控制失活基因的种类和数量,限制其使用。目前主要采用图位克隆法和转座子标签法进行未知功能基因的克隆。例如,至 2000 年 2 月,已有二十多个植物抗病基因被分离克隆出来,其中 60%的抗病基因是利用图位克隆法获得,40%是经采用转座子标签法克隆得到^[20]。对于基因组较大、重复序列较多的一些生物,采用图位克隆法克隆基因不仅投资大而且效率低;而转座子标签法受其种类、活性、数量的制约,在一些生物很难获得转座子突变体,从而不能用该方法克隆目的基因。

在对已经克隆的近 20 个植物抗病基因产物的结构进行比较分析后,人们发现它们存在一些共同的结构域,如核苷酸结合位点(nucleotide binding site, NBS);富亮氨酸重复(leucine rich repeat, LRR);跨膜域(transmembrane domain,

TM),以及蛋白激酶(protein kinase, PK)等。这些结构可能参与蛋白质之间的相互作用,以及细胞信号的识别与传导过程。根据抗病基因的这样一些保守结构域设计简并引物,应用 PCR 技术获得相似序列 DNA 片段,为快速鉴定候选抗性基因提供了新的尝试和途径。这种基于同源序列的候选基因克隆法(homology-based candidate gene method),已引起国内外学者的广泛重视。

2.2 简并 PCR 在基因克隆中的应用

如上所述,基于同源序列的候选基因法主要采用的就是简并 PCR 技术。到 1993 年为止,在小鼠中先后克隆了 42 个基因,其中有 31 个由候选基因法获得,大约占 74%。在植物抗病基因克隆方面,小麦中 Prk10 基因是采用候选基因法克隆的首个基因。已经有许多文献分别报道根据基因同源序列设计引物或简并引物,扩增克隆出与已知基因保守序列同源的序列^[21~24]。国内薛勇彪^[25]、王石平^[26]、李子银^[27]等分别利用 R 基因同源序列获得水稻、小麦的抗病基因同源序列。

其实,简并 PCR 一直是寻找和发现“新”基因或基因家族新成员的一种非常有用的工具。简并 PCR 可用于搜索一个基因家族的新成员^[10]、不同物种的同源基因^[28]或相关病毒^[29~31]。自 1988 年 Lee 采用该方法成功进行尿酸氧化酶基因的克隆以来^[32],这方面的报道也越来越多^[33~40]。随着时间的推移,简并 PCR 技术会越来越受到重视,采用简并 PCR 技术进行基因克隆的尝试和报道也会越来越多。

参 考 文 献(References):

- [1] <http://210.72.11.60/primer.doc>. 张新宇. 使用 Oligo 6 和 Primer Premier 5.0 等软件设计 PCR 引物.
- [2] Dieffenbach C W, Dveksler G S 著(黄培堂等译). PCR 技术实验指南[M]. 北京:科学出版社,1998,95~100.
- [3] 郑仲承. 寡核苷酸的优化设计[J]. 生命的化学,2001,21(3): 254~256.
- [4] Wada k, Wada Y, Doi H, *et al.* Codon usage tabulated from the GenBank genetic sequence data[J]. Nucleic acids research, 1991,19:1981.
- [5] Knoth K, Roberds S, Poteet C, *et al.* Highly degenerate, inosine-containing primers specifically amplify rare cDNA using the polymerase chain reaction[J]. Nucleic acids research, 1988,16, 10932.
- [6] SHEN Zhi-Yuan, Wells R L, LIU Jing-mei, *et al.* Identification of a cytochrome P450 gene by reverse transcription PCR using degenerate primers containing inosine[J]. Proc Natl Acad Sci, 1993,90:11483~11487.
- [7] Lin P K T, Brown D M. Synthesis of oligodeoxyribonucleotides containing degenerate bases and their use as primers in the polymerase chain reaction[J]. Nucleic acids research, 1992, 20: 5149~5152.

- [8] Hill F, Loakes D, Brown D M. Polymerase recognition of synthetic oligodeoxyribonucleotides incorporating degenerate pyrimidine and purine bases[J]. Proc Natl Acad Sci, 1998, 95: 4258~4263.
- [9] Nichols R, Andrews P C, Zhang P, *et al.* A universal nucleotide for use at ambiguous sites in DNA primers[J]. Nature 369: 492~493.
- [10] Wilks F F, Kurban R R, Hovens C M, *et al.* The application of polymerase chain reaction to cloning members of the protein tyrosine kinase family[J]. Gene, 1989, 85, 67~74.
- [11] 王佑春, 庄辉, R Ling, 等. 简并引物在扩增 GBV-C/HGV NS3 高度变异区的作用[J]. 中华微生物学和免疫学杂志, 1998, 18(5): 417~419.
- [12] 丁兆军, 王台, 种康. 水稻减数分裂相关基因 OsDMC1 的克隆[J]. 科学通报, 2001, 46(10): 838~843.
- [13] 杜占文, 刘立仁, 张俊武. 用 RACE 结合 cDNA 文库筛选的方法获取新的锌指蛋白基因[J]. 遗传, 2002, 24(3): 329~331.
- [14] 梁国栋主编. 最新分子生物学实验技术[M]. 北京: 科学出版社, 2001, 32~37.
- [15] 金晓琳, 胡福泉. 利用 3'-RACE 法扩增到钙通道基因的 3' 一端片段[J]. 细胞与分子免疫学杂志, 1999, 15(3): 210~212.
- [16] 李冰, Demetris S. 岸蟹 (*Carcinus maenas*) 金属硫蛋白 cDNA 及其基因的克隆[J]. 生物化学与生物物理学报, 2000, 32(6): 640~644.
- [17] Rose T M, Schultz E R, Henikoff J G, *et al.* Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly-related sequences [J]. Nucleic Acids Research, 1998, 26(7): 1628~1635.
- [18] Richmond T. Designing degenerate PCR primers. CODEHOP: consensus-degenerate hybrid oligonucleotide primers[J]. Genome Biology, 2000, 1(1): 240.
- [19] 王泽立, 戴景瑞, 王斌. 植物基因的图位克隆[J]. 生物技术通报, 2000, 4: 21~27.
- [20] 张增艳, 孔凡晶, 辛志勇. 植物抗病基因克隆的进展[J]. 生物技术通报, 2000, 2: 8~11.
- [21] Yu Y G, Buss G R, Maroof M A S. Isolation of a superfamily of candidate disease-resistance genes in soybean based on a conserved nucleotide-binding site[J]. Proc Natl Acad Sci, 1996, 93: 11751~11755.
- [22] Leister D, Ballvora A, Salamini F, *et al.* A PCR-based approach for isolating pathogen resistance genes from potato with potential for wide application in plants[J]. Nature Genet, 1996, 14(4): 421~429.
- [23] Leister D, kurth J, Laurie D A, *et al.* Rapid reorganization of resistance gene homologues in cereal genomes[J]. Proc Natl Acad Sci, 1998, 95: 370~375.
- [24] 李梅章, 褚嘉祐, 杨昭庆, 等. 一个 NEK 基因家族新成员的克隆和鉴定[J]. 遗传, 2001, 23(2): 97~102.
- [25] 薛勇彪, 唐定中, 张燕生, 等. 水稻基因组中 R 类抗病基因同源序列的分离[J]. 科学通报, 1998, 43: 277~281.
- [26] 王石平, 刘克德, 王江, 等. 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: 42~50.
- [27] 李子银, 陈受宜. 水稻抗病基因同源序列的克隆、定位及其表达[J]. 科学通报, 1999, 44: 727~723.
- [28] Kopin A S, Wheeler M B, Leiter A B. Secretin: Structure of the precursor and tissue distribution of the mRNA[J]. Proc Natl Acad Sci, 1990, 87: 2299~2303.
- [29] Mack D, Sninsky J J. Sensitive method for the identification of uncharacterized viruses related to known virus groups: Hepadnavirus model system[J]. Proc Natl Acad Sci, 1988, 85: 6977~6981.
- [30] Manos M M, Ting Y, Wright D K, Lewis A J, Broker T R, Wolinsky S M. Use of polymerase chain reaction amplification for the detection of genital human papillomaviruses [J]. Cancer Cells, 1989, 7: 209~214.
- [31] Shih A, Misra R, Rush M G. Detection of multiple, novel reverse transcriptase coding sequences in human nucleic acids: Relation to primate retroviruses[J]. J Virol, 1989, 63: 64~75.
- [32] Lee C C, Wu X, Gibbs R A, *et al.* Generation of cDNA probes directed by amino acid sequence: Cloning of urate oxidase[J]. Science, 1988, 239: 1288~1291.
- [33] 陈枝楠. 简并 PCR 技术研究洋水仙 PVY 组病毒初报[J]. 中国病毒学, 1995, 10(4): 346~350.
- [34] 张振臣, 李大伟, 陈健夫, 等. 甘薯羽状斑驳病毒外壳蛋白基因在大肠杆菌中的表达及特异抗血清的制备[J]. 农业生物技术学报, 2000, 8(2): 177~179.
- [35] 张宪省, 李全梓, 李兴国, 等. 风信子 HAG 基因的克隆与表达[J]. 中国科学(C 辑), 2000, 30(4): 376~381.
- [36] 王晓晨, BAUW W A D G, 徐庆, 等. Willy DILLEN 天麻中一种抗真菌蛋白基因的克隆[J]. 植物学报, 1999, 41(10): 1041~1045.
- [37] 刘宝, 戎均康, 董英山, 等. 普通小麦 7B 染色体的显微分离和低拷贝专化 DNA 序列的克隆[J]. 科学通报, 1999, 44(4): 389~393.
- [38] 王毅, 单祥年, 张悦, 等. 黑鹿 Y 染色体的鉴别和 Sry 基因的克隆及定位[J]. 遗传, 2001, 23(2): 114~118.
- [39] 张悦, 单祥年, 鲁晓萱, 等. 染色体显微切割与 DOP-PCR 结合对赤鹿 Sry 基因克隆、测序及初步定位[J]. 遗传学报, 2001, 28(4): 322~326.
- [40] 陈静, 倪坚, 陈江野. PCR 法用于白色念珠菌 MAPK 新基因的筛选[J]. 生物化学与生物物理学报, 2000, 32(3): 305~301.