

专论与综述

基因体外诱变*

马向东 黄春华 周俊初**

(华中农业大学生命科学技术学院 武汉 430070)

摘 要: 综述了基因体外诱变的一般方法和技术, 并将其分为不依赖于 PCR 体外诱变和依赖于 PCR 体外诱变两大类。着重介绍了基因体外诱变最新突破即 DNA Shuffling 技术。

关键词: 基因, 体外诱变

中图分类号: Q93 **文献标识码:** A **文章编号:** 0253-2654 (2002) 01-0070-04

基因诱变作为改造生物体最直接有效的方法, 已经过了从整体(细胞)诱变到局部(基因)诱变, 从体内的随机诱变到体外的定点诱变。诱变技术发展到今天, 科学家们已经能够控制整个诱变过程, 定点的改造目的基因了, 诱变为生物学基础理论研究和实际应用的重要手段。

早期遗传学家研究突变的材料仅限于一些自然变种和偶然发生的自发突变。随着人们对突变研究的深入, 在这个领域内已取得了很大进展。20 世纪 30 年代发现 X-射线和 40 年代发现有些化学物质能提高突变率, 使研究方法从自发突变一跃而成为诱发突变。这些物理和化学因素的诱变作用, 使人们能部分控制遗传突变的过程, 从而系统的研究基因及其控制的性状。第二次突破性进展是在 50 年代初期 McClintock 发现玉米中可在基因组不同区域转移的控制成分。接着人们在果蝇、细菌和各类真核生物中发现了这种转移成分, 即转座子(transposons)。用转座子通过适当的操作, 理论上可得到任何基因的插入失活突变, 而转座子上带有的特异性序列和抗药性基因极大的方便了被插入基因的克隆。第 3 次突破就是 80 年代初发展起来的基因体外诱变技术。体外诱变技术与体内诱变最大的不同在于可将克隆的基因单独在体外高效突变而不影响遗传背景上任何其他遗传因子, 能大大提高特异性和效率, 成为最重要的生物学研究手段之一。本文将着重介绍体外诱变的传统方法和最新的策略。

根据体外诱变方法的原理不同, 可将其分为两大类: 依赖 PCR 的体外诱变和不依赖 PCR 的体外诱变。

1 不依赖 PCR 的体外诱变

1.1 体外缺失诱变 (Deleting mutagenesis) 采用不同的核酸酶从某一限制性酶切位点开始逐步去掉靶基因一端或两端的部分序列, 形成一系列基因缺失。将这些缺失的基因酶连转化到宿主中检测其表达的不完整的蛋白质, 从而确定截去的这一部分的功能, 如 N-端 DNA 结合域、C-段双体形成的结合部位和中间的 ATP 结合域^[1]。操作过程如下: 先将克隆的基因连在载体多克隆位点上, 接着找出合适的酶切位点酶切后位使基因产生一个可被外切核酸酶Ⅲ消化的 3' 凹端而载体上产生一个不能用外切核酸酶Ⅲ消化的 5' 凹端, 酶切位点如果选在上游就可以得到基因编码蛋白 N-端缺失, 反之得

* 国家自然科学基金资助项目 (No. 39970022)

** 联系作者

收稿日期: 2000-07-10, 修回日期: 2000-09-10

到 C-端缺失。用外切核酸酶 III 部分酶切, 去掉 3' 凹端的部分核苷酸, 再用绿豆核酸酶将两端抹平, 紧接着用连接酶进行连接。最后转化得到一系列去掉不同长度核苷酸的基因的转化子。理论上讲有三分之一的阅读框架正确而能表达截短的蛋白质 (如图 1)。

1.2 酶切定点诱变 使用不同的酶将目的基因上的酶切位点切开再进行一系列的酶切、补平、抹平、酶连等操作, 就能引入特定位点上的定点诱变。如通过使用限制性内切酶 *Hgi*AI 和 *Kpn*I、T4 DNA 聚合酶、连接酶等多种酶, 经过酶切、抹平可将 GAG-CAC 位点突变为 GTGCAC 位点改变密码子编码的氨基酸而不改变阅读框 (Reading frame)^[2] (如图 2)。

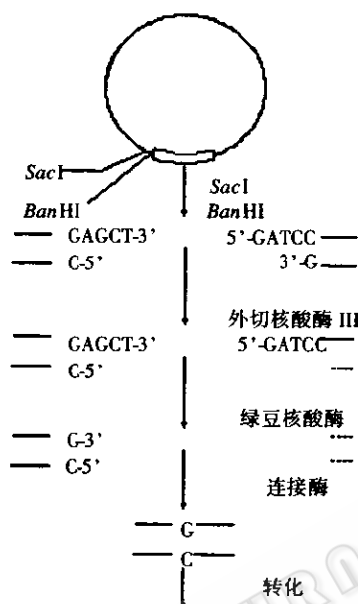


图 1 单向缺失诱变

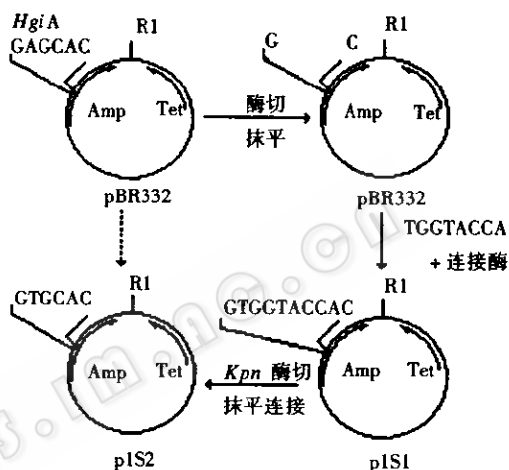


图 2 酶切位点诱变

1.3 定域随机诱变 (Localized random mutagenesis) 因为定域随机诱变是将诱变限制在一段很小的区域内而不影响所改造的基因的其他序列部分, 所以具有很高的位点专一性, 同时筛选时效率很高, 往往只需要很少的转化子就能够筛选到所需要的目的突变。一般过程是: (1) 用限制性酶切下基因中欲突变的目的片段, 对其进行诱变处理, 然后再连回原基因重建成完整的基因^[3]。(2) 使诱变的片段对诱变剂特异性敏感, 使诱变剂作用于目的区域而与其他部分不发生作用。如通过适当的酶操作双链 DNA 使目的区段变成单链 (如缺口、缺失环、置换环等单链结构), 然后用与单链作用的专一性诱变剂 (如亚硫酸氢钠) 诱变, 再用聚合酶和连接酶合成完整的双链, 实现此区域的高效诱变^[4]。

1.4 定点诱变 (Site-specific mutagenesis) 定点诱变有许多方法, 但归纳起来有两大类: 一类是寡核苷酸介导的诱变, 以带有突变的核苷酸为引物, 以单链 DNA 为模板, 用酶进行延伸, 这种异质双链 DNA 转化大肠杆菌, 一部分就是带突变的菌落。第二类是盒式诱变, 此法是使用人工合成的含有突变的 DNA 目的片段在体外取代野生型相应片段造成突变。

(1) 寡核苷酸介导的诱变^[2] (Oligonucleotide-directed mutagenesis): 通过人工合成的少量密码子发生变化的寡核苷酸介导得到诱变的目的基因的一种诱变方式。具体操作过程将目的基因克隆到 M13mp8 上, 设计合成的密码子发生突变的寡核苷酸, 然后以此

寡核苷酸为克隆了目的基因的 M13mp8 上的引物。用大肠杆菌聚合酶 I 大片段延伸和连接酶连接。将得到含突变密码子共价闭合 DNA 电泳, 回收, 纯化, 转染大肠杆菌 JM103, 转移噬菌斑用点杂交获得突变的基因。如为了提高筛选效率可以采取高效选择获得突变^[3], Villafranca 等以二氢叶酸还原酶为材料采用寡核苷酸定点突变的方法改变了二氢叶酸还原酶不同部位的 3 个氨基酸, 由此对此酶的结构与功能进行了研究^[5]。为提高突变效率, 也可以使用甲基化的核苷酸合成突变寡核苷酸进行诱变^[6]。

(2) 盒式诱变 (Cassette mutagenesis): 通过基因定向诱变的过程将限制性内切酶切点引入目的序列的两侧, 然后将两侧的 DNA 序列用相应的限制性内切酶切开。将合成的含突变序列的双链 DNA 诱变盒通过连接酶连接到前面基因的断口上, 使基因的长度恢复到原样, 同时封闭一侧或两侧的酶切位点。由于指定的突变区域 DNA 是合成的, 因此可以得到任何可能的突变, 而又不会产生任何混合的或非目的位点的突变, 因此对于蛋白质功能的研究尤为有利^[7]。

盒式诱变的具体过程^[7]如下: 首先将目的基因克隆到适当的载体上, 接着用定向诱变的方法在准备诱变的目的密码子两侧各引入一个单酶切位点再连到同一载体上; 然后将此载体用新引进的两个酶切位点切开出线型; 最后用人工合成的只有目的密码子发生了变化的双链 DNA 诱变盒和线型载体酶连, 转化筛选所需的突变子。

1.5 脱嘌呤体外诱变 脱嘌呤是自然界中最常见的自发突变, 据估计一个哺乳动物细胞一天脱嘌呤在 10^4 个位点, 其半衰期有几天时间。脱嘌呤是由于连接嘌呤碱基和脱氧核糖的糖苷键断裂造成的, 能由多种因素引起, 如自发水解、特殊的糖苷酶水解、由于化学修饰造成糖苷键的不稳定等。如果 DNA 中存在脱嘌呤位点, 那么在 DNA 聚合酶无模板指导时造成错误掺入, 从而引起突变^[8]。真核生物 DNA 聚合酶是倾向错误的, 很容易通过此脱嘌呤位点不完全随机的发生掺入错误。在酸性条件下脱嘌呤, 如果在不平衡核苷酸体系中用聚合酶处理或使用 Mn^{2+} 则可以使 DNA 聚合酶错配率大为增加。

2 依赖 PCR 的体外诱变

2.1 依赖于 PCR 的定点诱变 早期应用定点诱变的方法, 得到突变子的效率很低^[9]。排除野生型基因的筛选方法既费时间又花精力。随着 PCR 技术的发展, 为定点诱变开辟了一条新途径, 并开发出各种方法^[10], 如采用含有一个和几个突变的寡核苷酸为引物, 对基因的一段或全部进行 PCR 扩增, 这样突变的寡核苷酸就掺入到了基因的相应位置, 通过酶切、与基因其他部分酶连等一系列的操作, 除了引入突变外, 还能重建原基因的结构并恢复原来的大小。将基因连入载体, 导入到宿主菌中, 就能观察突变的表型而研究特定基因特定突变与其产物结构与功能的关系。为了控制 PCR 过程中造成非必需突变, 可以采用保真性高的 DNA 聚合酶, 增加底物浓度, 减少扩增循环的次数。又如一种依赖于反向 PCR 进行定点诱变的策略, 对已经克隆到环状双链质粒上的基因进行诱变时, 首先设计两个方向相背而序列正好相接, 以扩增出整个环状质粒的全部序列的寡核苷酸引物, 其中带有基因特定位点上的突变。对质粒扩增之后, 得到的是完整序列的线性质粒, 基因作两段在质粒 DNA 两侧, 而同时引入了突变。最后只需要进行酶连就可以得到环状的质粒, 用 *Dpn* II 酶切去除模板, 经转化筛选就可以得到所需的突变^[11]。

2.2 倾向错误的 PCR 诱变 (Error-prone PCR) 进行 PCR 扩增反应时最常用的 DNA 聚合酶 Taq DNA 聚合酶是所有已知 DNA 聚合酶中掺入错误碱基几率最高的。每扩增一次其错误率在 $0.1 \times 10^{-4} \sim 2 \times 10^{-4}$ 之间, PCR 反应经 20 ~ 25 次循环, 积累的错误率可

达 10^{-3} /碱基。Taq DNA 聚合酶在一般条件下错误总是有一定的倾向, 即 AT 碱基对转变为 GC 碱基对的倾向^[12]。但这种突变率仍然不够, 特别是对小片段分子, 更难得到突变通过对 $MgCl_2$, dNTP 的浓度比例的改变, 增加聚合酶量, 添加新的 Mn^{2+} 导进一步提高 PCR 反应的突变率。增加 $MgCl_2$ 的浓度能稳定非互补的碱基对; 加入一定量的 $Mn-Cl_2$ 以降低聚合酶对模板的特异性; 调节 dNTP 的浓度促进掺入错误的进行。

2.3 通过 DNA shuffling 技术进行基因体外诱变 借助于 PCR 扩增法在体外进行的 DNA Shuffling 技术能模拟生物的基因在数百万年间发生的分子进化过程, 并能在短期的实验循环中定向筛选出特定基因编码的酶蛋白活性提高成千上万倍的功能性突变基因。1994 年, 美国的 Stemmer 采用单个基因的 DNA Shuffling 和回交技术, 在大肠杆菌中使编码 β -内酰胺酶的 TEM-I 基因的抗生素抑制活性 (MIC) 从 $0.02\mu g/mL$ 提高到 $640\mu g/mL$, 即提高了 32, 000 倍。次后, 他们又相继选取了绿荧光蛋白基因 (gfp) (1996) 和 β -半乳糖苷酶基因 (lac) (1997) 用 DNA Shuffling 技术模拟并加速了分子进化过程。筛选获得了酶活性大幅度提高的突变株^[13]。1996 年, Moore 等亦采用该技术对降解污水中有机污染物的对硝基苯基酯酶基因进行了定向模拟分子进化研究。Cramer 等 (1998) 采用 4 个不同来源的先锋霉素基因混合进行异源基因组的 DNA Shuffling, 使单一循环的 Mic 活性提高了 270 ~ 540 倍。而同时只用单一基因进行的实验仅提高了 8 倍^[14]。DNA shuffling 体外基因突变一般过程如下^[15]: (1) 获得目的基因: 常用的有两种方法。为了增加突变率采用 PCR 扩增目的基因回收; 为了增强保真性, 可以采取抽提含有目的基因的质粒, 通过酶切回收。可以根据不同的需要任取其中一种方法。(2) 基因酶切回收: 采用 DNAase I 酶切基因 DNA 得到 10-50 bps 或 100-300 bps 的片段, 酶切过程中用 Mn^{2+} 能提高保真性 3 倍左右。(3) 无引物 PCR: 让基因片段互为引物进行扩增, 扩增时基因片段将由小到大。为确保得到高保真性的片段, 应在较高的温度下复性。(4) 有引物 PCR: 将无引物 PCR 产物作适当稀释, 作为模板进行 PCR, 于无引物 PCR 不同的是, 增加引物进行 PCR 扩增出目的基因。(5) 目的突变基因的获得: 回收目的片段, 克隆到载体上, 转化选择所需的突变的了的目的基因。如筛选抗性基因、高酶活基因等。

诱变的新方法不断涌现, 为生物学各基础学科提供了强有力的工具和极为丰富的研究材料, 尤其是最近出现的 DNA Shuffling 技术更是强大有效。体外诱变已经使我们进入了改造生物的自由王国, 我们相信它将为人类征服自然作出更大贡献。

参 考 文 献

- [1] Turner P C, Melennan A G, Bates A D, *et al.* Molecular Biology, 北京: 科学出版社, 1999, 158 ~ 162.
- [2] Irving S S, Betty G H, Rene A. Proc. Natl. Acad. Sci. USA, Biochemistry, 1982, 79: 7157 ~ 7160.
- [3] Solnick D. Nature, 1981, 291: 508.
- [4] Hayatsu. prog H. Nucleic Acids Res, Mol. Biol, 1976, 16: 75.
- [5] Jesus E V, Elizabeth E H, Donald H V. *et al.* Science, 1983, 222: 782 ~ 788.
- [6] Mark A V, Michael P W, Carolyn J H. *et al.* Gene, 1988, 65: 129 ~ 133.
- [7] James A W, Mark V, David B P. Gene, 1985, 34: 315 ~ 323.
- [8] Lawrence A L. Cell, 1985, 40: 483 ~ 484.
- [9] Weiner M W, Felts K A, Sisim T G. *et al.* Gene, 1993, 126: 35 ~ 41.
- [10] Olfert L, Hans-Peter G, Ulrich H. Gene, 1990, 96: 125 ~ 128.
- [11] Hemsley A, Arnheim N, Toney M D, *et al.* Nucleic Acids Res, 1989, 17: 6545 ~ 6551.
- [12] Keohavong P, Thilly W G. Proc. Natl. Acad. Sci. USA, 1989, 86: 9253 ~ 9257.
- [13] Zhang J H, Dawes G, Stemmer W P C. Proc. Natl. Acad. Sci. USA, 1997, 94: 4504 ~ 4509.
- [14] Cramer A. Nature, 1998, 191 (15): 288 ~ 291.
- [15] Zhao H M, Arnold F H. Nucleic Acids Res, 1997, 25: 1307 ~ 1308. F