

一种方便、无 PCR 过程的 DNA shuffling 方法

王强 刘秋云 李刚 李宝健*

(中山大学生命科学学院基因工程教育部重点实验室, 广州 510275. * 联系人, E-mail: lsbrc15@zsu.edu.cn)

摘要 目前大多数的 DNA shuffling 方法都需要 PCR 过程. 由于一些限制内切酶的识别位点与切割位点不重叠而产生不同的黏性末端, 当再连接的时候, 同源片段之间就可以混合起来按照正确顺序组装成新的嵌合体分子. 利用这种性质, 构建了一种新的、无 PCR 过程的、简便的 shuffling 方法.

关键词 DNA shuffling 限制性内切酶 组织型纤溶酶原激活剂

DNA shuffling 是利用同源基因之间的相似序列在体外进行重组的技术. 在 shuffling 过程中, 亲本序列(parental sequence)的多态性被重组, 因而可以探索更大的序列空间和结构空间. Shuffling 的关键步骤是亲本序列的打断和碎片的重组装. 目前的 DNA shuffling 方法通常使用 DNase 消化或者 PCR 交错延伸(staggered extension process)来获得亲本序列的碎片; 重组装则是利用碎片序列间的重叠部分通过无引物 PCR 来完成^[1-9]. 限制性内切酶也可以把亲本序列切成碎片, 连接酶可以把它们重新组装起来, 因此理论上讲它们也可以用来进行 shuffling. 然而, 大多数限制性内切酶的切割位点就是它们的识别位点, 这样的切割一般都产生相同的、且是回文序列的黏性末端. 这样的片段用连接酶连接起来以后, 各个碎片的原有顺序通常遭到破坏, 正确顺序的组装虽然存在, 但是比例很小, 正如一个各个碎片形状都相同的拼图游戏中, 仅仅通过形状很难拼出正确的图案来. 一些限制性内切酶, 如 *Fok*, *Bbv* 等, 其识别位点位于切割位点上游或者下游若干碱基处, 对亲本序列切割后生成的黏端通常是各不相同且非回文序列的黏端, 通过连接酶连接以后可以得到顺序正确的组装. 利用这一性质, 我们构建了一种新的 shuffling 方法, 并且构建了一个微型 shuffling 文库, 其亲本分子分别为人组织型纤溶酶原激活剂(tissue plasminogen activator, *tPA*)基因以及含有 7 个突变位点的 *tPA* 基因, 对这个文库的质量分析表明这种新方法是可行的.

1 材料与方法

() 材料. 突变体 *tPA* 分子(命名为 127PA)由本室构建, 含有 7 个突变位点, mt1~7(图 1). 克隆载体 pBlueScript SK-购自 Stratagen 公司, 转化受体菌 *Escherichia coli* TOP10F' 购自 Invitrogen 公司, 限制

性内切酶以及连接酶购自 NEB 公司, PCR 纯化试剂盒购自 QIAGEN 公司, Taq DNA 聚合酶购自博亚生物工程有限公司. 各突变位点在天然 *tPA* 基因(native *tPA*, 以下记作 *ntPA*)或 127PA 基因上的位置从左到右依次为: mt7, mt5, mt6, mt1, mt3, mt2, mt4. 127PA 的基因型记作 1111111, *ntPA* 的基因型记作 0000000. 如果一个 *tPA* 突变体含有若干突变位点, 如 mt5 和 mt1, 那么它的基因型记作 0101000. *ntPA* 基因(或 127PA)含有 12 个 *Bbv* 位点, 3 个 *Fok* 位点和 1 个 *Apa*L 位点. 所有 *Bbv* 给出的黏端各不相同, 相互之间无互补性, 且都是非回文序列; 所有 *Fok* 给出的黏端情况也是一样. mt7 和 mt5 之间缺乏 *Bbv* 和 *Fok* 位点, 故使用 *Apa*L 来切割这个区域(图 1).

() Shuffling 方法. 首先, 把 *ntPA* 和 127PA 用 *Bam*H 和 *Not* 酶切, 克隆到 SK-载体上. 构建好的 SK-*ntPA*(或 SK-127PA)质粒共含有 7 个 *Fok* 位点, 25 个 *Bbv* 位点以及 3 个 *Apa*L 位点. 所有 *Bbv* 给出的黏端各不相同, 相互之间无互补, 且都是非回文序列; 所有 *Fok* 给出的黏端情况也是一样(图 2). 第二步, SK-*ntPA* 和 SK-127PA 各取 100 ng, 加入 10 U 的 *Apa*L, 37 °C 2 h 后, 用 PCR 纯化试剂盒纯化, 回收产物加入 T4 DNA 连接酶及反应缓冲液, 16 °C 保温 2 h. 转化大肠杆菌, 随机挑选克隆, 从中筛选出 *Apa*L 片段所对应的 2 种重组子(基因型分别为 1000000 和 0111111, 这 2 个质粒分别记作 SK-64PA 和 SK-63PA). 取上述 4 种质粒(SK-*ntPA*, SK-127PA, SK-64PA 和 SK-63PA)各 100 ng 混合, 加入 10 U *Bbv* 酶切 1 h 后, 65 °C 灭活 20 min, 在反应液中直接加 ATP 至终浓度为 1 mmol/L 后, 加入 T4 DNA 连接酶, 16 °C 保温 2 h 后, 取 10 μL 转化大肠杆菌. 转化液中加入 LB, 并且添加氨苄青霉素至终浓度 60 μg/mL, 过夜培养. 取培养物提取的质粒 100 ng, 加 1 U *Fok*, 37 °C 保温 1 h

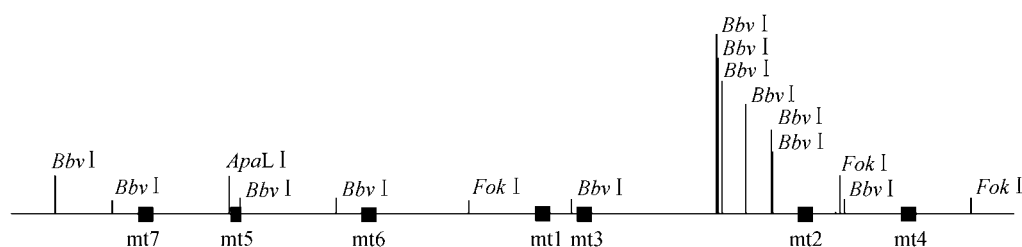


图 1 7 个突变位点以及 3 种内切酶酶切位点在 *ntPA* (或者 *127PA*) 基因上的位置

■ 示突变位点

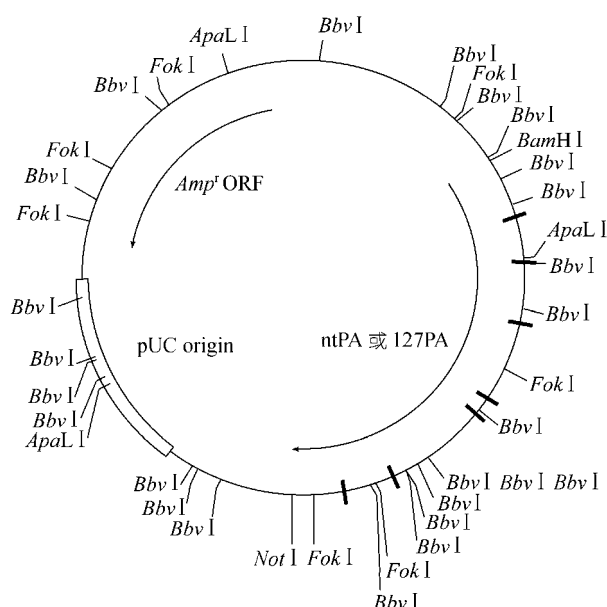


图 2 质粒 SK-*ntPA* 以及 SK-*ntPA* 的图谱

■ 示各突变位点所在位置

后, 65 °C 灭活 20 min. 连接转化后, 取 100 μL 菌液涂板, 剩余转化液 -70 °C 保存. 随机挑取平板上的克隆分别培养提取质粒进行基因型分析.

() 基因型鉴定方法. 预期的重组子将以 *ntPA* 序列为骨架, 上含 *127PA* 的 7 个突变位点的不同组合. 这个组合就是该重组子的基因型. 由于测序比较昂贵, 我们采用自己建立的半引物(hemi-primer)多重 PCR 方法^[10]来鉴定重组子的基因型. PCR 反应可以用来检测突变^[11~13]. 但是如果有 7 个突变位点需要检测则需要做 7 个常规 PCR 反应. 这样繁重的工作可以用 7 个特别设计的半引物作为检测各个突变的正向引物, 与 1 个共同的反向引物在同一个管内同时进行这 7 个 PCR 反应来简化. 这些正向引物除了用于突变检测外, 还作为温度开关来分隔同一个管中发生

的不同 PCR 反应. 它们由 3 部分组成: 取突变位点区域序列作为引物的 3' 末端部分, 其 3' 末端与突变位点序列匹配而与 *ntPA* 上对应序列不匹配; 3' 其余部分与突变位点紧邻序列完全匹配(由这两部分结合的序列的 T_m 值在 45 °C 左右); 引物的 5' 部分与模板不匹配. 整个引物序列的 T_m 值在 65 °C 左右. 把所有 7 种正向引物和 1 个共用的反向引物混合, 加入 *Taq* DNA 聚合酶以及其他 PCR 必需成分, 热启动, 加入待测质粒 50 ng 左右. PCR 程序为: 94 °C, 30 s, 50 °C, 30 s, 72 °C, 1 min, 1~2 个循环; 94 °C, 30 s, 65 °C, 30 s, 72 °C, 1 min, 15 个循环; 72 °C, 10 min. 之后进行凝胶电泳. 分子量标准由上述 7 个半引物为正向引物分别进行的 PCR 的产物混合而成(这些 PCR 反应的模板都是 *127PA*, 均使用上文提及的共用的反向引物). 若凝胶中待测样品泳道上, 某突变位点对应位置出现了 PCR 条带, 则表明该样品含有该突变位点, 否则不含该突变位点. 这个方法的有效性已经在多个突变位点组合已知的 *tPA* 突变体(包括 7 个突变位点都存在的 *127PA*)检测中得到了证实; 用该方法对突变位点组合未知的 *tPA* 突变体进行基因型测定, 所得结果与测序结果完全一致.

2 结果

随机挑取 8 个克隆进行检测, 所获得的重组子其基因型各不相同(图 3); 对这些克隆进行测序也得到了一致的结果; 另外随机挑取的 8 个克隆其基因型也都各异(结果未显示). 这些结果表明这个方法可行的.

3 讨论

构建好的 SK-载体上含有 3 个 *ApaLI* 位点, 由于 *ApaLI* 的切点是回文序列, 酶切后再连接的产物可能出现顺序混乱. 仔细分析载体序列发现, 这 3 个

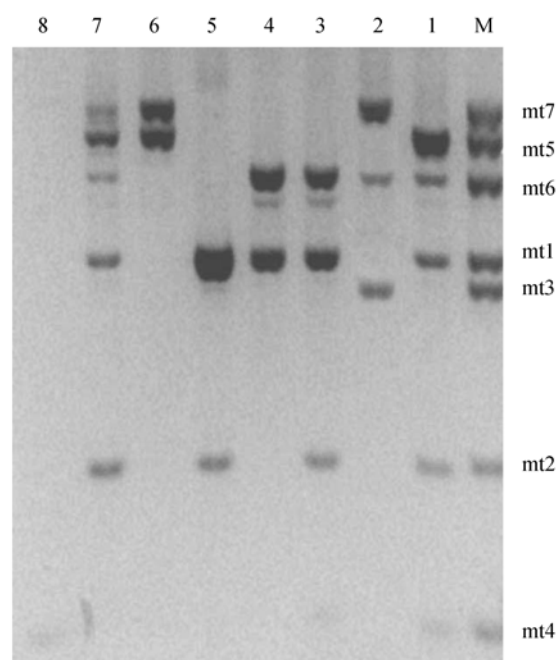


图3 从用本文所述 Shuffling 方法制作的文库中随机挑选的 8 个克隆的基因型检测结果

1~8 示克隆 1~8 的半引物多重 PCR 的产物; M 示分子量标准, 分别与 7 个突变位点相对应

*Apa*L 切点, 1 个位于复制原点区域, 1 个位于氨苄抗性基因的编码区域。这 2 个区域的插入或者缺失会导致质粒无法繁殖, 因此如果能够获得质粒则说明这 2 个区域内 *Apa*L 是以正确顺序组装的。使用 SK-空载体进行的预实验表明使用 *Apa*L 切割-重组装的质粒都是以正确顺序连接的(结果未显示)。第三个 *Apa*L 位点位于插入片段内。这个位置可能出现插入。但连接和转化倾向于较小的质粒, 因此这个区域出现插入的比例应该不高。尽管如此, 为排除干扰, 随机挑选克隆然后从中筛选出 2 种重组子(SK-63PA 和 SK-64PA), 和亲本质粒 SK-ntPA, SK-127PA 混合后进行后续反应。

构建好的质粒共含有 25 个 *Bbv* 位点、7 个 *Fok* 位点, 其中若干片段的长度只有几十个碱基, 回收困难。因此本试验方案采用直接加热灭活内切酶, 然后在酶切反应液中加入连接酶和 ATP。这样即避免了损失小片段而影响 shuffling 效果, 又简化了操作。

致谢 非常感谢谢军博士和王震宇博士对本文提出的宝贵建议。本工作为国家自然科学基金(批准号: 30070420, 30370799)和广州市科委基金(批准号: 2001-Z-006-01)资助项目。

参 考 文 献

- 1 Stemmer W P. DNA shuffling by random fragmentation and reassembly: *In vitro* recombination for molecular evolution. *Proc Natl Acad Sci USA*, 1994, 91(22): 10747~10751
- 2 Stemmer W P. Rapid evolution of a protein *in vitro* by DNA shuffling. *Nature*, 1994, 370(6488): 389~391[DOI]
- 3 Zhao H, Arnold F H. Optimization of DNA shuffling for high fidelity recombination. *Nucleic Acids Res*, 1997, 25(6): 1307~1308[DOI]
- 4 Zhao H, Giver L, Shao Z, et al. Molecular evolution by staggered extension process (StEP) *in vitro* recombination. *Nat Biotechnol*, 1998, 16(3): 258~261
- 5 Shao Z, Zhao H, Giver L, et al. Random-priming *in vitro* recombination: An effective tool for directed evolution. *Nucleic Acids Res*, 1998, 26(2): 681~683[DOI]
- 6 Cramer A, Raillard S A, Bermudez E, et al. DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature*, 1998, 391(6664): 288~291[DOI]
- 7 Kikuchi M, Ohnishi K, Harayama S. Novel family shuffling methods for the *in vitro* evolution of enzymes. *Gene*, 1999, 236(1): 159~167[DOI]
- 8 Kikuchi M, Ohnishi K, Harayama S. An effective family shuffling method using single-stranded DNA. *Gene*, 2000, 243(1-2): 133~137[DOI]
- 9 Ness J E, Kim S, Gottman A, et al. Synthetic shuffling expands functional protein diversity by allowing amino acids to recombine independently. *Nat Biotechnol*, 2002, 20(12): 1251~1255[DOI]
- 10 Wang Q, Liu Q, Li B. Simultaneous detection of 7 mutations with 7 forward primers and 1 common reverse primer in a single PCR step. *J Biochem Biophys Methods*, 2004, 58(2): 153~157
- 11 Huang M M, Arnheim N, Goodman M F. Extension of base mispairs by Taq DNA polymerase: Implications for single nucleotide discrimination in PCR. *Nucleic Acids Res*, 1992, 20(17): 4567~4573
- 12 Kwok S, Kellogg D E, McKinney N, et al. Effects of primer-template mismatches on the polymerase chain reaction: Human immunodeficiency virus type 1 model studies. *Nucleic Acids Res*, 1990, 18(4): 999~1005
- 13 Sarkar G, Cassady J, Bottema C D K, et al. Characterization of polymerase chain reaction amplification of specific alleles. *Anal Biochem*, 1990, 186(1): 64~68

(2003-08-12 收稿, 2004-02-12 收修改稿)