

通过原位易错 PCR 一步构建基因突变文库

王未未² 张永昌² 刘明杰² 邵蔚蓝^{1*}

(1. 江苏大学 环境学院 生物质能源研究所 江苏 镇江 212013)

(2. 南京师范大学 生命科学学院 江苏 南京 210046)

摘要: 主要介绍一种通过原位易错 PCR 构建随机突变文库的新技术。本实验室最近发表的一项国际专利中, 利用来源于海栖热孢菌的极耐热性 DNA 连接酶, 在传统 PCR 循环中加入一个连接步骤, 即变性—退火—延伸—连接的四步循环法 PCR, 从而实现环状质粒的 PCR 指数扩增(PPCP)。原位易错 PCR 中所用引物为一段线性双链 DNA, 它含有与模板质粒不同的筛选标记, 产物转化宿主菌后, 模板质粒在筛选平板上被直接剔除。筛选到的阳性突变子可用作模板直接进入突变文库的再次构建, 通过筛选获得二级或多级累加的正突变。利用这种方法构建了一个木聚糖酶基因和一个纤维素酶基因的随机突变文库, 并筛选出具有正向突变的蛋白, 证明以 PPCP 为基础的原位易错 PCR 技术, 为基因定向进化提供了一种快速有效的随机突变文库构建的新方法。

关键词: 定向进化, 随机突变文库, 原位易错 PCR, 环状质粒扩增(PPCP)

One-step construction of mutagenesis libraries via *in situ* error-prone PCR

WANG Wei-Wei² ZHANG Yong-Chang² LIU Ming-Jie² SHAO Wei-Lan^{1*}

(1. Biofuels Institute, School of the Environment, Jiangsu University, Zhenjiang, Jiangsu 212013, China)

(2. College of Life Sciences, Nanjing Normal University, Nanjing, Jiangsu 210046, China)

Abstract: We developed a novel method to construct mutagenesis libraries via *in situ* error-prone PCR. One of our recent PCT patents described a novel method that adding a ligation step mediated by thermostable *Thermotoga maritima* (*Tma*) DNA ligase to form the repeated PCR cycles of “denaturation–annealing–elongation–ligation”, which allows exponential amplification of circular DNA. In this method, circular PCR products are generated by using a long dsDNA primer pair, which is designed to carry a selection marker different from the original selection marker of the template plasmid, the template plasmid carrying the original marker is eliminated when transformed host cells are cultured under the new selection pressure. If the product serves as the template for the next round of amplification, accumulation of positive mutations can be obtained by multiple rounds of *in situ* error-prone PCR. This method has been used to create random mutagenesis libraries of a xylanase gene and a cellulase gene. Screening of these libraries resulted in mutant proteins with desired prop-

基金项目: 国家自然科学基金项目(No. 31170027); 江苏高校优势学科建设工程项目(PAPD)

*通讯作者: Tel: 86-511-88796122; 邮箱: weilanshao@gmail.com

收稿日期: 2013-04-09; 接受日期: 2013-06-03; 优先数字出版日期(www.cnki.net): 2013-10-11

erties, demonstrating the usefulness of *in situ* error-prone PCR for creating random mutagenesis libraries for directed evolution.

Keywords: Directed evolution, Construction of random mutagenesis libraries, *In situ* error-prone PCR, PCR production of circular plasmid (PPCP)

基因定向进化已经成为蛋白质工程中的一项重要技术。体外定向进化不仅为研究酶的底物特异性和确定酶的催化活性位点、改善酶的热稳定性以及提高酶的作用温度等方面提供了强有力的手段,而且有助于了解蛋白结构与功能之间的关系^[1-5]。在分子生物技术的发展过程中人们不断地发展基因体外进化和重组的方法,现在已经建立了多种用于构建随机突变文库的方法^[6]。

最常用的随机突变的方法是易错 PCR^[7]。其原理是在 PCR 过程中通过调整 Mg^{2+} 和 Mn^{2+} 浓度等因素,提高 *Taq* 酶在扩增过程中引入错配碱基的频率,导致目标基因的扩增产物发生适量的随机突变^[8]。早期的通过易错 PCR 构建突变文库的步骤包括:(1) 以目标基因为模板在易错条件下进行 PCR,扩增出含有错配碱基的突变基因;(2) 对易错 PCR 产物和表达载体进行酶切;(3) 用 DNA 连接酶将突变基因连接到表达载体上;(4) 将重组的表达载体导入宿主细胞,获得带有各种不同突变基因的重组细胞群体,即基因突变文库。构建一个多样化的基因突变文库,需要从成千上万个转化子中筛选出所需要的正突变基因;上述方法就显得很繁琐,工作量很大^[9-12]。另外,用 *Taq* 酶进行的易错 PCR 会在 DNA 链的 3'末端带一个多余的碱基,因此在克隆前还需要修补 DNA 链或去除此碱基^[13]。而且通过这种方法构建的文库中常常会产生不带有目的基因片段或是带有多多个目的基因片段的质粒^[14]。因此早期发明的突变文库的构建方法步骤繁琐、效率低下。

为了提高构建基因突变文库的效率, Miyazaki 和 Takenouchi 于 2002 年提出特大引物全质粒扩增法(Mega-WHOP)^[15]; Wei 等 2004 年对这个技术进行了改进和发展^[16]。该扩增方法以随机突变后的目标基因作为双链互补的引物对(即特大引物,

Megaprimers),以含有原始基因的表达质粒为模板(将此模板质粒甲基化),采用高保真 DNA 聚合酶进行 PCR,扩增出质粒的全长片段;再用 *Dpn* I 酶处理,将甲基化了的模板质粒降解,而新扩增出的质粒因为没有被甲基化,所以被完整保留。由于没有 DNA 连接酶存在,利用引物和质粒模板单向延伸产生的线性 DNA 新链不能作为模板进行下一轮的扩增, Mega-WHOP 法不能实现指数扩增;此外,这种方法在转化突变产物前还需要使用类似于 *Dpn* I 这样的酶来进行一个模板消化的过程,步骤繁琐。

为解决以上问题,本文介绍一种新型的通过原位易错 PCR 构建随机突变文库的方法,为广大科技工作者提供一个更为方便快捷的构建随机突变文库的方式。

1 原位易错 PCR (*In situ* error-prone PCR) 技术

原位易错 PCR 技术是本实验室最近发表的一项国际专利中公布的一种最新方法(PCT 专利公开号: US-2012-0004142-A1)^[17]。该发明专利的主要创新点包括:(1) 用一种极耐热性 DNA 连接酶,在传统 PCR 循环中加入一个连接步骤,即“变性—退火—延伸—连接”的四步循环,从而实现环状质粒的 PCR 指数扩增(PPCP);(2) PPCP 和易错 PCR 相结合,在表达载体中对目标基因进行原位易错 PCR,产生带有随机突变基因的环状质粒文库;(3) 在原位易错 PCR 过程中更替筛选标记,产物与模板在平板筛选中自行分离。

通过原位易错 PCR 进行基因定向进化的基本步骤如图 1 所示,将目标基因克隆到表达载体中构建成表达质粒;同时,从序列互补但是带有不同抗药基因的载体质粒中扩增出线性载体 DNA;

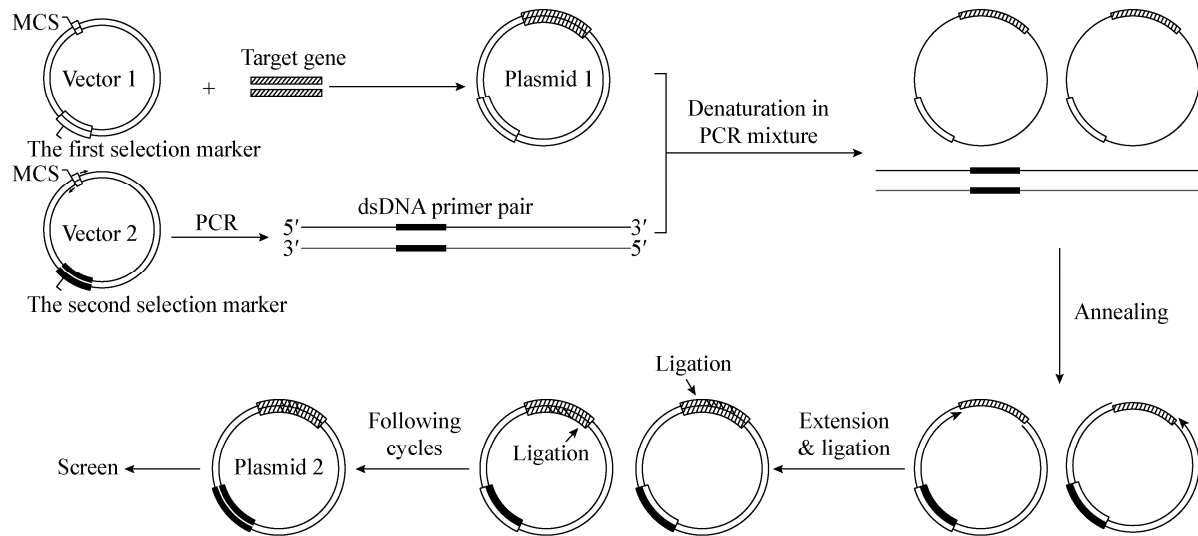


图 1 原位易错 PCR 构建基因突变文库及筛选流程

Figure 1 Schematic of *in situ* error-prone PCR (from PCT/US2010/027364)

再以表达质粒为模板, 线性载体 DNA 作引物, 利用 *Taq* DNA 聚合酶在易错条件下对表达质粒中的目标基因进行延伸; 再通过极耐热性连接酶的连接作用, 扩增出带有突变基因的可直接用于转化的环状质粒, 实现通过易错 PCR 一步构建基因突变文库。

这种方法省略了限制性酶切、酶连等繁琐且低效的操作, 可以用 PCR 产物直接进行转化。另外, 由于引物片段中的抗性基因与 PCR 反应中加入的模板质粒所带的抗性基因不同, 无需对模板质粒进行降解, 可直接将 PCR 产物转化后通过不同的筛选标记选择性地培养含突变基因的转化子, 将基因突变产物与原始模板分离, 完全排除模板质粒对转化的影响。

2 原位易错 PCR 技术的特点

2.1 原位易错 PCR 与 MegaWHOP 的不同之处

原位易错 PCR 和 MegaWHOP 两种方法虽然都要用到易错 PCR 技术, 但是它们在本质上有区别(表 1)。虽然 MegaWHOP 的试剂盒已经在美国市场上销售, 其方法并不符合 PCR 的基本原理; 由于非指数扩增, 并且主要产物不是双链环状的 DNA, MegaWHOP 的效率势必走低。原位易错 PCR

因基因在表达载体中的原位被诱导随机突变而得名, 所用的引物是通过高保真性扩增的同源载体的线性双链 DNA; 相对于 MegaWHOP 的特大引物 (Megaprimer) 而言, 原位易错 PCR 所用的引物可谓“超(级)大引物”, 在国际专利中被称为载体引物 (Vector-primer)。

“超大引物”与“特大引物”之间的另一个区别在于超大引物具有通用性。Mega-WHOP 中所用的特大引物必须针对目标基因及易错条件特别制备, 适宜的片段大小为 0.2–1.0 kb; 原位易错 PCR 中所用的“超大引物”是一个载体系列的公共序列, 不论克隆到载体中的是哪一个基因, 都可以用同一种引物进行原位易错 PCR。

原位易错 PCR 引物的通用性和操作的简易性为我们带来更大的便利是正突变的多轮叠加。如图 2 所示, 当备有两套带有不同筛选标记(如 Amp^r 或 Kan^r) 的超大引物时, 可以轮番运用这两套引物, 在筛选到正突变基因的基础上继续进行随机诱变和筛选, 获得二级或多级累加的正突变。

2.2 原位易错 PCR 的技术要点

环状质粒的 PCR 扩增 (PPCP) 技术中所使用的 DNA 连接酶需要有高度的热稳定性。经过研究, 来自极端嗜热菌 *Thermotoga maritima* (*Tma*) 的

表 1 原位易错 PCR 与特大引物全质粒扩增法的比较 Table 1 Comparison of <i>in situ</i> error-prone PCR with MegaWHOP						
方法 Method	引物对 Primer pair	扩增区域 Extension region	PCR 条件 PCR condition	产物 Product	模板去除 Template removing	指数扩增 Exponential amplification
MegaWHOP	Mutated target gene (0.2–1.0 kb)	Vector sequence	High-fidelity	Linear ssDNA and their hybrids (nicked dsDNA)	<i>Dpn</i> I digestion	No
PPCP	Vector sequence	Target gene	High-fidelity or error-prone	Closed-circular dsDNA	Plate selection	Yes

DNA 连接酶在 95 °C 的半衰期超过 30 min，适用于 PCR 过程^[18]。*Tma* DNA 连接酶只对粘性末端有连接作用，对平端则没有连接作用，因此此酶在反应中只将 PCR 产物连成环状，不会对双链的超大引物进行相互连接或自身环化。*Tma* DNA 连接酶连接过程中需要二价离子如 Mg^{2+} 、 Mn^{2+} 和 Ca^{2+} 等的参与；其中添加 Mn^{2+} 既可以增强连接作用，也是易错 PCR 常用的诱导碱基错配的条件。值得注意的是，此酶在反应过程中需辅酶 NAD^{+} ，因此，PPCP 反应体系中一定要添加 NAD^{+} 。

超大引物和普通 PCR 引物的要求一样，要求 3'端的序列具有较强的特异性。因为充当超大引物的双链 DNA 片段变性解链后，正负链各自与互补的模板链退火成为引物，所以这段双链 DNA 的两端序列都应当具有特异性。设计超大引物时应尽量避免在末端使用非编码区的常见序列，如很多基因共有的类似启动子、终止子、转录起始位点至起始密码子之间的引导序列等。超大引物可以通过 PCR

方法合成：以载体质粒为模板，合成一对普通引物，用高保真性且产生平末端的 DNA 聚合酶，如 *Pfu* 和 *Pyrobest* 等来扩增载体上的相应区域，包括选择标记的编码基因及两侧的其余序列。非常重要的是，PCR 扩增的超大引物的 5'端必需经过多聚核苷酸激酶进行磷酸化处理，才能在原位易错 PCR 中实现新链的首尾连接从而产生环状 DNA。

原位易错 PCR 过程中，“变性—退火—延伸—连接”四步循环的温度是个有趣的问题。由于使用的引物非常大：除了筛选标记基因以外，筛选标记的上游和下游各有几百个与模板序列互补的碱基，其退火温度可高于 65 °C；*Taq* 酶的最适聚合温度为 72 °C，但是在 60 °C 也有较高的活性；*Tma* DNA 连接酶的最适连接温度是 60 °C，但是在 70 °C 的活性高达最高活性的 80%^[18]。因此，原位易错 PCR 过程中除变性温度为 94 °C 以外，退火、延伸与连接的温度非常接近，都可以在 65–72 °C 之间选择，甚至可以合并为一步任其连续进行。

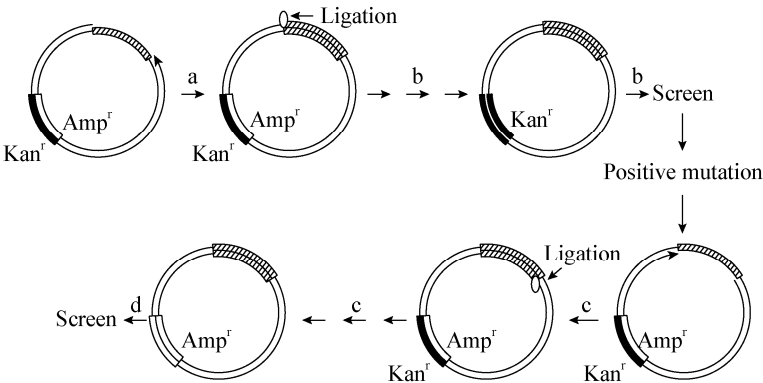


图 2 构建随机突变文库的多级突变筛选示意图
Figure 2 Schematic of directed evolution by multi-round *in situ* error-prone PCR

有效的平板筛选技术或高通量筛选方法是利用原位易错 PCR 技术进行基因定向进化的必要条件。对于许多酶基因实施定向进化, 平板筛选技术的缺乏是主要瓶颈问题, 但是这个问题不在本文讨论之列。

3 原位易错 PCR 应用实例

在中国发明专利(ZL 200910025925.0)文献中介绍的实例包括一个木聚糖酶和一个纤维素酶基因的定向进化过程和试验条件^[19]。实例中采用 pHsh 系列载体构建表达质粒, 其优点包括: (1) 载体质粒小, 转化率高; (2) 载体系列中有不同筛选标记供选择; (3) 该系统质粒为温控型, 可以通过在不同温度下培养来控制表达水平^[20]。试验流程如图 1 所示, 实验者首先将嗜热疏棉状丝孢菌的木聚糖酶基因 *xynA* 和海栖热袍菌纤维素酶基因 *celB* 分别克隆到带有抗药基因 Amp^r 的载体 pHsh-amp (GenBank accession No. FJ571619) 中构建成表达质粒 pHsh-*xynA*^[21] 和 pHsh-*celB*; 再用一对小引物扩增同一系列的带有抗药基因 Kan^r 的载体 pHsh-kan (GenBank accession No. FJ571619), 产生的线性 dsDNA 经过磷酸化即可用作超大引物。

原位易错 PCR 在 20 μL 反应体系中进行, 含有 pHsh-*xynA* 或 pHsh-*celB* 模板 DNA 30 ng 带有 Kan^r 的超大引物 DNA 200 ng, *Taq* DNA 聚合酶 0.5 U, 极耐热性 *Tma* DNA 连接酶 0.1 μg , 以及 0.2 mmol/L dNTPs, 2.0 mmol/L MgCl_2 , 0.5 mmol/L MnCl_2 , 0.5 mmol/L NAD^+ 和 PCR 缓冲液。这次试验中采用的 PCR 循环程序为: 95 $^\circ\text{C}$ 5 min; 94 $^\circ\text{C}$ 1 min, 72 $^\circ\text{C}$ 1.5 min, 60 $^\circ\text{C}$ 2 min, 共 15 个循环; 60 $^\circ\text{C}$ 10 min。琼脂糖凝胶电泳分析结果表明, 加入 *Taq* DNA 聚合酶和 *Tma* DNA ligase 的 PCR 反应中, 环状质粒条带中 DNA 量有明显的提高^[19]。

PCR 产物转化 *E. coli* JM109 感受态细胞 (Promega Corp, Madison, WI, USA), 将 pHsh-*xynA* 的易错 PCR 产物涂布在含有 2% 木聚糖和卡那霉素的 LB 培养基平板上, 将 pHsh-*celB* 的易错 PCR 产物涂布在含有 2% CMC 和卡那霉素的 LB 培养基平板上, 分别放置在 30 $^\circ\text{C}$ 恒温培养箱中培养 10 h 后转到 42 $^\circ\text{C}$ 恒温培养箱中进行诱导表达 5–6 h^[20] (图 3), 菌落周围的透明圈或凹陷圈的大小表明基因突变导致酶活性发生很大的变化。

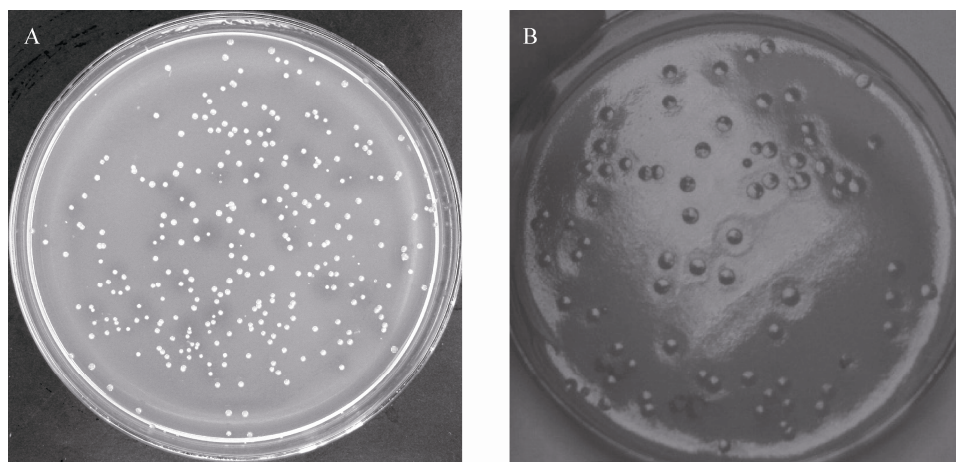


图 3 实施例中易错 PCR 产物转化到宿主大肠杆菌中, 突变子在筛选平板上生长

Figure 3 Directed evolution of xylanase and cellulase as examples of *in situ* error-prone PCR

注: A: 转化子在含 2% 木聚糖的 LB 平板上生长, 阳性克隆会在菌落周围产生透明圈。B: 转化子在含有 2% CMC 的 LB 平板上生长, 阳性克隆会在菌落周围产生凹陷圈。

Note: A: Functional screening of *E. coli* transformants for xylanase activity using xylan-overlay assay. Positive clones were identified by clear zones surrounding xylanase-expressing colonies. B: Functional screening of *E. coli* transformants for cellulase activity by CMC-overlay assay. Positive clones were identified by depressed area surrounding cellulase-expressing colonies.

4 总结

自 1993 年第一次成功运用易错 PCR 以来^[22], 在短短的十几年中, 各种进化方法层出不穷。文中介绍的原位易错 PCR 方法适用于任何可用于克隆的表达载体, 最好使用能够在大肠杆菌中表达的载体。已知的适用于此方法的选择标记也有很多, 比如文中所使用的抗性基因。同样, 需要突变的目的基因也可以是任何编码序列。可以是一个蛋白的全长编码序列, 也可以是一段 RNA 序列或一段调控序列等。

此方法中连接反应和 PCR 反应同步进行, 建议参考文中提供的反应温度设定热循环程序。 MnCl_2 浓度与易错 PCR 的突变频率相关, 推荐浓度范围为 0.05–0.50 mmol/L, 建议对小于 1 kb 的目标基因采用较高 Mn^{2+} 浓度, 对大于 1 kb 的基因适当降低其浓度。极耐热性 *Tma* DNA 连接酶和 NAD^+ 按照文中实例用量使用, 其他酶和试剂浓度按常规 PCR 条件使用。此方法的反应条件需要根据具体情况, 比如突变基因的大小、类型或研究者需要的突变率进行微调, 并没有具体的、通用的标准, 这可能需要使用者根据实际情况对反应条件进行优化。突变子的筛选方法也因目标基因性质而异, 与本技术没有直接关系。

参考文献

- [1] Chirumamilla RR, Muralidhar R, Marchant R, et al. Improving the quality of industrially important enzymes by directed evolution[J]. *Molecular and Cellular Biochemistry*, 2001, 224(1/2): 159-168.
- [2] Cherry JR, Fidantsef AL. Directed evolution of industrial enzymes: an update[J]. *Current Opinion in Biotechnology*, 2003, 14(4): 438-443.
- [3] Eijssink VG, Gaseidnes S, Borchert TV, et al. Directed evolution of enzyme stability[J]. *Biomolecular Engineering*, 2005, 22(1/3): 21-30.
- [4] Yuan L, Kurek I, English J, et al. Laboratory-directed protein evolution[J]. *Microbiology and Molecular Biology Reviews*, 2005, 69(3): 373-392.
- [5] Bhat MK. Cellulases and related enzymes in biotechnology[J]. *Biotechnology Advances*, 2000, 18(5): 355-383.
- [6] Otten LG, Quax WJ. Directed evolution: selecting today's biocatalysts[J]. *Biomolecular Engineering*, 2005, 22(1/3): 1-9.
- [7] Cadwell RC, Joyce GF. Randomization of genes by PCR

- mutagenesis[J]. *PCR Methods and Applications*, 1992, 2(1): 28-33.
- [8] Kammann M, Laufs J, Schell J, et al. Rapid insertional mutagenesis of DNA by polymerase chain reaction (PCR) [J]. *Nucleic Acids Research*, 1989, 17(13): 5404.
- [9] Miyazaki K, Takenouchi M. Creating random mutagenesis libraries using megaprimer PCR of whole plasmid[J]. *Biotechniques*, 2002, 33(5): 1033-1038.
- [10] Nakaniwa T, Tada T, Takao M, et al. An *in vitro* evaluation of a thermostable pectate lyase by using error-prone PCR [J]. *Journal of Molecular Catalysis B-enzymatic*, 2004, 27(2/3): 127-131.
- [11] Zhang JH, Dawes G, Stemmer WP. Directed evolution of a fucosidase from a galactosidase by DNA shuffling and screening[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 1997, 94(9): 4504-4509.
- [12] Kim MS, Lei XG. Enhancing thermostability of *Escherichia coli* phytase AppA2 by error-prone PCR[J]. *Applied Microbiology and Biotechnology*, 2008, 79(1): 69-75.
- [13] Hu G. DNA polymerase-catalyzed addition of nontemplated extra nucleotides to the 3' end of a DNA fragment[J]. *DNA and Cell Biology*, 1993, 12(8): 763-770.
- [14] Shen B. PCR approaches to DNA mutagenesis and recombination: an overview[J]. *Methods in Molecular Biology*, 2002, 192: 167-174.
- [15] Miyazaki K, Takenouchi M. Creating random mutagenesis libraries using megaprimer PCR of whole plasmid[J]. *Biotechniques*, 2002, 33(5): 1033-1038.
- [16] Wei D, Li M, Zhang X, et al. An improvement of the site-directed mutagenesis method by combination of megaprimer, one-side PCR and *Dpn* I treatment[J]. *Analytical Biochemistry*, 2004, 331(2): 401-403.
- [17] Shao W, Le Y, Pei J. Method for construction mutagenesis libraries *in situ*[P]. US, 2012/0004142 A1. 2012-1-5.
- [18] Le Y, Peng J, Pei J, et al. Properties of an NAD^+ -dependent DNA ligase from the hyperthermophile *Thermotoga maritima* and its application in PCR amplification of long DNA fragments[J]. *Enzyme and Microbial Technology*, 2010, 46(2): 113-117.
- [19] 邵蔚蓝, 乐易林, 裴建军. 原位构建基因突变文库的方法及试剂盒[P]. 中国发明专利, ZL 200910025925.0, 2011-7-20.
- [20] Wu H, Pei J, Jiang Y, et al. pHsh vectors, a novel expression system of *Escherichia coli* for the large-scale production of recombinant enzymes[J]. *Biotechnology Letters*, 2010, 32(6): 795-801.
- [21] Yin E, Le Y, Pei J, et al. High-level expression of the xylanase from *Thermomyces lanuginosus* in *Escherichia coli*[J]. *World Journal of Microbiology and Biotechnology*, 2008, 24(2): 275-280.
- [22] Chen K, Arnold FH. Tuning the activity of an enzyme for unusual environments: Sequential random mutagenesis of subtilisin E for catalysis in dimethylformamide[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 1993, 90(12): 5618-5622.