

文章编号: 1004-0374(2011)09-0849-04

应用基因组工程构建新型大肠杆菌细胞工厂

蔡 钒^{1,2}, 方宏清^{2*}, 陈必链¹

(1 福建师范大学生命科学学院, 福州 350108; 2 军事医学科学院生物工程研究所, 北京 100071)

摘要: 基因组工程 (genome engineering) 是指为了实现某一目标对复制系统进行有意地、广泛地遗传修饰。大肠杆菌作为常用细胞工厂之一, 在生物制造领域应用广泛, 已成为合成生物学的主要研究对象。应用基因组工程改造大肠杆菌可以进一步拓展其应用范围。介绍了基因组工程最新技术发展, 如基因组整合、染色体进化、多元自动化基因组工程、可寻迹多元重组工程等, 及其在改造大肠杆菌提高其生产能力、稳定性及环境耐受能力等方面的研究进展。

关键词: 细胞工厂; 大肠杆菌; 基因组工程; 合成生物学

中图分类号: R378.21; Q812; O621.33 **文献标志码:** A

Constructing novel *E. coli* cell factories by genome engineering

CAI Fan^{1,2}, FANG Hong-Qing^{2*}, CHEN Bi-Lian¹

(1 College of Life Sciences, Fujian Normal University, Fuzhou 350108, China;

2 Beijing Institute of Biotechnology, Beijing 100071, China)

Abstract: Genome engineering is the extensive and intentional genetic modification of a replicating system for a specific purpose. *Escherichia coli*, one of the most common cell factories, is widely used in the domains of biological manufacturing and becomes a major subject in synthetic biology. Its application fields could be further expanded by genome engineering. This paper reviews the latest technology development of genome engineering, such as genome integration of heterogenous pathway, chromosomal evolution, multiplex automated genome engineering and trackable multiplex recombineering, and their application in improving the productivity, the stability and the tolerance of *Escherichia coli* as cell factories.

Key words: cell factories; *Escherichia coli*; genome engineering; synthetic biology

微生物细胞工厂 (microbial cell factories) 是以可再生资源为原料, 通过微生物细胞转化制备人类所需的能源、药物、化工原料等等, 或者利用微生物治理日益恶化的环境。大肠杆菌 (*Escherichia coli*) 是研究得最清楚、分析得最透彻的模式微生物, 在生物制造领域应用广泛。大肠杆菌属于紫细菌, 一类生理活性最丰富的细菌, 可以与多种物质代谢及能量代谢途径兼容, 因此成为合成生物学研究领域的主力微生物。然而, 自然界中野生型的微生物由于含有的酶种类有限、代谢速率和产物积累浓度较低, 因此无法有效地满足工业生产的需要^[1], 需要对现有微生物进行改造以适应新的应用需求。基因组工程 (genome engineering) 就是为了实现某一目标

对复制系统进行有意地、广泛地遗传修饰^[2]。本文就基因组工程相关技术的发展现状及在构建新型大肠杆菌细胞工厂中的应用作一介绍。

1 异源途径的基因组整合技术

染色体是细胞内最稳定的载体, 维持合成代谢途径稳定表达的一种有效方法是将代谢途径整合到染色体上。由于存在多复制叉的现象, 整合位点相

收稿日期: 2011-05-25

基金项目: 国家自然科学基金项目(30780049); 国家重点基础研究发展计划(“973”项目)(2011CBA00801)

*通信作者: E-mail: fanghongqing@vip.sina.com

对于复制起始区距离的不同仍然会产生细胞与细胞之间拷贝数的差异。此外，不同整合位点插入可能产生不同的 DNA 构象，这些都有可能影响被整合代谢途径的表达。因此，有时需要考察不同整合位点对被整合途径表达效率的影响。目前，将异源代谢途径整合到基因组上的方法主要有两种：质粒衍生的方法和重组工程。

质粒衍生的方法是应用噬菌体整合系统的介导方法，将合成的 DNA 片段插入到染色体上^[3-4]。这种方法中，一个供体质粒含有一个噬菌体特异性的附着位点 (attP)，当转化到一个表达噬菌体整合酶的宿主细胞内后就可以整合到染色体上的 attB 附着位点。这种方法相当高效，而且当表达适当的噬菌体 xis 切离酶时，还可以很容易地将整合片段从染色体上移除。此外，它对片段的大小并没有限制。但是，这种方法需要构建一个独立的载体，而且插入位点的选择性低，质粒骨架和抗性基因也会同时整合到基因组上。

重组工程是借助 λ -Red 重组酶，促进染色体与线性 DNA 片段之间的同源重组^[5]。这种方法通过 PCR 扩增获得两侧具有 45 bp 左右同源臂的线性 DNA 片段，在实施上具有很大的灵活性。重组工程是目前改造染色体最有效、易操作的方法。近几年，基于 Red 重组进行精确位点修饰的研究取得一些新的进展。2010 年，Ying 等^[6]优化电转条件以及优化阿拉伯糖诱导的策略，通过抗性筛选成功将长约 4 kb 的 PCR 片段整合到染色体上，并证明基因组整合后能够降低生物噪音。2010 年，Albermann 等^[7]通过将基因整合到非必需的碳源降解途径位点，应用 MacConKey 平板进行筛选，成功地将类胡萝卜素合成基因 *crtE*、*crtB*、*crtI*、*crtY* 和 *crtZ* 逐个整合到大肠杆菌基因组上不同的糖降解途径位点上。然而，对于大片段 DNA 来说这种方法不适合：一方面，通过 PCR 扩增难以获得足够量高保真的产物；另一方面随着片段大小增加，转化及整合效率出现了显著降低，当片段的长度超过 3 kb 时就难以获得成功，尽管也有一些成功的报道。另外，这种方法通常需要一个抗性标记用于筛选。

2010 年，Kuhlman 和 Cox^[8]描述了一种利用着陆垫 (landing pad) 将大片段 DNA 定位插入到染色体上的方法。他们利用这个方法成功地将长达 7 kb 的片段插入到染色体的指定位置上。因着陆区序列很短，需要双链断裂促进整合。因此，这种方法可以连续使用，不需要构建新的着陆区序列。但是，

该方法的一个潜在的缺陷是重复应用标准的着陆序列有可能导致以前插入片段丢失。因此，需要适当的抗生素选择压力来保留已整合到基因组上的片段。在许多应用事例中，需要在不同位点整合不同的序列。因此，仍然需要一种无痕精确修饰方法。

针对现实需求，我们实验室构建了能够表达 I-Sce I 归巢核酸内切酶和 Red $\alpha\beta\gamma$ 重组系统的辅助质粒，将 Red 重组技术和 I-Sce I 特异性切割技术以及 Cm-SacB 正负筛选相结合，建立了一种异源代谢途径无痕无突变基因组定位整合技术。这种方法的重组效率高，可直接利用 PCR 方法筛选重组修饰菌株，避免了现有基因组修饰方法存在的上述问题。利用供体质粒将异源代谢途径引入宿主内，可避免 PCR 扩增产生的突变。应用该方法可将 8 kb 以上的重组途径整合到基因组上，我们已成功将青蒿素前体倍半萜、番茄红素的人工合成代谢途径整合到大肠杆菌基因组上，并获得高效表达，表达水平优于质粒载体。我们建立的染色体整合技术可用于多次连续地将不同的异源基因整合到染色体上的不同位置，对构建超长基因簇稳定表达工程菌具有指导意义，对微生物合成生物学的发展具有一定的促进作用。

2 染色体进化技术

代谢工程以及合成生物学研究中常用的方法是在宿主的细胞内引入一个新的基因或者是直接协调各个基因间的表达活性^[9-10]。在质粒上表达的基因可以获得多拷贝，这样从某些方面来说可以获得一个更高的产量，而且通过质粒介导的方法简单易行，操作简便。因此，质粒介导的途径工程是短时表达重组基因的最好选择，尤其是用来高表达重组蛋白时更是一个理想的选择。但是，对于多酶催化途径合成产物的研究来说，过强的表达基因对长期生产不利。此外，质粒不仅有着结构和分配的不稳定性、需要用选择标记来维持其在宿主细胞内的存在^[11]，还会因为等位基因分离导致非生产质粒替代生产质粒，从而引起质粒表达系统的不稳定^[12]。更为重要的是质粒还会影响细胞代谢途径（尤其是中心代谢途径），除了会增加代谢负荷不利于菌体生长外，还会引起其他的生理变化，如增加氧摄取率和葡萄糖吸收、促进 ATP 合成等^[13-14]。因此，质粒介导的途径工程存在着影响遗传稳定性，进而影响重组途径的缺点，更为甚者还可能通过影响宿主细胞的全局转录网络而影响目标产物的合成。

基于质粒介导异源途径存在的诸多缺点, Keith 等^[12]在 2009 年发展了一种能够适用于大肠杆菌的化学诱导染色体进化技术 (chemically inducible chromosomal evolution, CICE), 提高基因组上某个基因或代谢途径的拷贝数。首先, 利用 λ InCh 基因组整合系统^[15]将代谢途径、氯霉素抗性标记和两侧与大肠杆菌基因组非同源的相同 1 kb 外源 DNA 序列整合到染色体的 gal-bio 位置。两侧相同的 1 kb 序列作为后续交叉重组事件中的同源底物。通过逐渐提高氯霉素浓度, 促进依赖 recA 同源重组的染色体进化, 提高重组途径的拷贝数。最高基因拷贝数可以达到 30~40 个。最后将 recA 基因敲除, 获得遗传稳定的工程菌。与质粒载体系统比较, 该方法构建的工程菌的遗传稳定性提高了 10 倍, 在规模化培养时不需要添加抗生素等选择压力, 符合商业生产的要求。在表达合成 PHB 的途径中, 在无抗生素选择压力存在时, 质粒载体系统的生产能力在传 5 代后降至 0; 有抗生素压力时, 在传 40 代后丢失 PHB 合成能力, 实验结果与质粒稳定性模型预测结果相当吻合。而染色体整合系统在传 70 代后仍保持 90% 的生产能力。他们还应用该方法提高番茄红素产量约 60%。

3 多元自动化基因组工程

自然界中微生物基因组的广泛多样性是微生物群体适应各种各样环境的基础。但是, 在实验室里难以获得基因组多样性, 在一个有实际意义的时间尺度内获得一个新表型也不太容易。体外直接进化可以产生有用表型的突变体, 但受限于单个基因操作, 不适于基因网络或基因组的平行进化。2009 年, Wang 等^[16]发展了一种称为多元自动化基因组工程 (multiplex automated genome engineering, MAGE) 技术, 可对细胞进行大尺度编程和进化研究。MAGE 可同时对染色体许多位点进行修饰, 修饰的方法可以是错配、插入或缺失。被研究的主体可以是单细胞, 也可以是一个细胞群体。因此, 可以产生基因组多样性组合。由于这个过程是一个循环的、可放大的, 可以通过自动装置来反复实施。这个装置包括维持细胞生长的培养室和反复转化 DNA 的电转室。他们应用一个合成的复合 DNA 库, 利用这种方法对合成番茄红素的 DXP 途径中 20 个基因的 SD 序列进行协同进化表达, 同时敲除 4 个基因, 每天产生 43 亿个突变子, 3 d 内就获得了番茄红素产量提高了 5 倍的突变株。

这个方法的成功实施是基于三个基础: (1) 基于 Red 单链 DNA 结合蛋白的同源置换技术, 他们通过多个参数优化使同源置换效率最大化, 可使 30% 以上的细胞群体发生遗传修饰; (2) 已清楚哪些基因影响目的产物的合成, 这样就可以有针对性地设计复合 DNA 库; (3) 要有简单快速筛选产物表达水平的方法, 便于从大量的突变库中筛选所需的突变。

4 可寻迹的多元重组技术

微生物基因组具有巨大的组合多样性潜力。尽管大肠杆菌基因组已测序完成很多年, 有 4 000 多个编码基因, 但是其中很多基因的功能还不明确, 它们与微生物性状, 如物质代谢能力、环境抵抗能力等之间的关系知之甚少。2010 年, Warner 等^[17]发展了一种可寻迹的多元重组技术 (trackable multiplex recombineering, TRMR)。TRMR 组合了平行 DNA 合成、重组工程和分子条形码技术, 能够快速修饰所有大肠杆菌基因组, 并可同时定位影响某一性状的遗传修饰位点。应用这种方法, 他们对大肠杆菌基因组上 4 077 个基因的表达进行组合突变, 每一个基因均包括两种突变模式: 表达上调和表达下调。用于突变的 DNA 片段大小约 800 bp, 包含用于重组的同源区、条形码标签、抗性筛选标记、功能区 (启动子/RBS 突变子)。为了构建如此大小 DNA 片段的文库, 应用了寡核苷酸库芯片合成技术、Uracil 切割连接技术、滚环复制技术。Warner 等^[17]应用该技术在一周内定位了几千个影响大肠杆菌在不同环境中生长的基因, 包括纤维素水解物、几种生长抑制剂等。应用这种方法在基因组尺度定位性状基因, 工作效率提高了几个数量级。将其与定向进化等策略结合将会进一步促进生物技术的发展。

5 展望

当前基因组工程正处于一个快速上升的时期。技术的不断发展和创新同时又为我们提供了许多新的设计思路。通过基因组改造也可以构建合成特殊修饰多肽的大肠杆菌宿主, 如乙酰化修饰^[18]、糖基化修饰^[19]等, 建立药物和疫苗生产的新工艺。发展“正交体系”或“遗传隔离”技术, 防止自身的遗传特征转移到环境中野生型的菌株体内^[20-21], 同时构建自身免疫体系防护外源 DNA 的入侵^[22], 可以为构建微生物细胞工厂提供一个更加安全的环

境。另外, 删除基因组上一些影响遗传稳定性的序列, 能够为构建微生物细胞工厂提供一个更加稳定的环境^[23]。而那些占用细胞资源的非必需组分的删除则能够为微生物细胞工厂节约大量的资源, 从而更有利于提高目的产物的产量。基因组工程能够让我们在工程上合成生态系统、多细胞发育系统等, 构建一些新型微生物细胞工厂高效生产化工原料、生物燃料以及医药健康产品, 必将为社会经济的可持续发展发挥巨大作用。

[参 考 文 献]

- [1] 马延和. 生物炼制细胞工厂: 生物制造的技术核心. 生物工程学报, 2010, 26(10): 1321-6
- [2] Carr PA, Church GM. Genome engineering. Nat Biotechnol, 2009, 27(12): 1151-62
- [3] Basu S, Gerchman Y, Collins CH, et al. A synthetic multicellular system for programmed pattern formation. Nature, 2005, 434(7037): 1130-4
- [4] Chiang CJ, Chen PT, Chao YP. Replicon-free and markerless methods for genomic insertion of DNAs in phage attachment sites and controlled expression of chromosomal genes in *Escherichia coli*. Biotechnol Bioeng, 2008, 101(5): 985-95
- [5] Sharan SK, Thomason LC, Kuznetsov SG, et al. Recombineering: A homologous recombination-based method of genetic engineering. Nat Protoc, 2009, 4(2): 206-23
- [6] Ying B, Ito Y, Shimizu Y, et al. Refined method for the genomic integration of complex synthetic circuits. J Biosci Bioeng, 2010, 110(5): 529-36
- [7] Albermann C, Trachtmann N, Sprenger GA. A simple and reliable method to conduct and monitor expression cassette integration into the *Escherichia coli* chromosome. Biotechnol J, 2010, 5(1): 32-8
- [8] Kuhlman TE, Cox EC. Site-specific chromosomal integration of large synthetic constructs. Nucleic Acids Res, 2010, 38(6): e92
- [9] Jones KL, Kim SW, Keasling JD. Low-copy plasmids can perform as well as or better than high-copy plasmids for metabolic engineering of bacteria. Metab Eng, 2000, 2(4): 328-38
- [10] Pflieger BF, Pitera DJ, Smolke CD, et al. Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes. Nat Biotechnol, 2006, 24(8): 1027-32
- [11] Friehs K. Plasmid copy number and plasmid stability. Adv Biochem Eng Biotechnol, 2004, 86: 47-82
- [12] Keith EJT, Kumaran PA, Gregory S. Stabilized gene duplication enables long-term selection-free heterologous pathway expression. Nat Biotechnol, 2009, 27(8): 760-5
- [13] Wang Z, Xiang L, Shao J, et al. Effects of the presence of ColE1 plasmid DNA in *Escherichia coli* on the host cell metabolism. Microb Cell Fact, 2006, 5(1): 34
- [14] Ow DSW, Lee DY, Yap MG, et al. Identification of cellular objective for elucidating the physiological state of plasmid-bearing *Escherichia coli* using genome-scale *in silico* analysis. Biotechnol Prog, 2009, 25(1): 61-7
- [15] Boyd D, Weiss DS, Chen JC, et al. Single-copy gene expression systems making gene cloning physiologically relevant: λ InCh, a simple *Escherichia coli* plasmid-chromosome shuttle system. J Bacteriol, 2000, 182(3): 842-7
- [16] Wang HH, Isaacs FJ, Carr PA, et al. Programming cells by multiplex genome engineering and accelerated evolution. Nature, 2009, 460(7257): 894-8
- [17] Warner JR, Reeder PJ, Karimpour-Fard A, et al. Rapid profiling of a microbial genome using mixtures of barcoded oligonucleotides. Nat Biotechnol, 2010, 28(8): 856-63
- [18] Ren YT, Yao XQ, Fang HQ, et al. Production of N^α-acetylated thymosin α 1 in *Escherichia coli*. Microb Cell Fact, 2011, 10: 26
- [19] Ihssen J, Kowarik M, Dilettoso S, et al. Production of glycoprotein vaccines in *Escherichia coli*. Microb Cell Fact, 2010, 9: 61
- [20] An W, Chin JW. Synthesis of orthogonal transcription-translation networks. Proc Natl Acad Sci USA, 2009, 106(21): 8477-82
- [21] Church GM. Safeguarding biology. Seed, 2009, 20: 84-6
- [22] Garneau JE, Dupuis M-È, Villion M, et al. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature, 2010, 468: 67-71
- [23] Umenhoffer K, Feher T, Posfai G, et al. Reduced evolvability of *Escherichia coli* MDS42, an IS-less cellular chassis for molecular and synthetic biology applications. Microb Cell Fact, 2010, 9(38): 1-38