

文章编号: 1004-0374(2011)09-0931-04

# 从DNA合成技术角度阐述合成生物学

田敬东<sup>1,2</sup>

(1 中国科学院天津工业生物技术研究所, 天津 300308;

2 美国杜克大学生物医学工程系, 基因组科学与政策研究所, 北卡罗来纳州 27708)

**摘要:** 合成生物学是一个拥有巨大潜力的新兴学科, 合成生物学技术的发展将会对未来生物、医药、农业、能源、材料和环保等方面产生巨大的推进作用。基因合成是合成生物学中最基本和使用最多的一种技术手段, 合成生物学的快速发展对基因合成能力提出了空前需求。综述基因合成技术的发展历史、现状和未来趋势, 探讨基因合成技术在合成生物学以及整个生命科学研究中的应用和重要意义。

**关键词:** DNA 合成; 基因合成; 基因组合成; 基因芯片; 合成生物学

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

## Gene synthesis technology and synthetic biology

TIAN Jing-Dong<sup>1,2</sup>

(1 Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin 300308, China;

2 Department of Biomedical Engineering, Institute for Genome Sciences and Technology, Duke University, NC 27708, USA)

**Abstract:** Synthetic biology is an emerging field with great potentials in promoting future developments in biomedical research, agriculture, bioenergy, biomaterial and bioremediation. Gene synthesis is the most basic and most widely used technique in synthetic biology. The rapid development of this field is creating insatiable demand for gene synthesis. This paper will review the history, current status and future trend of gene synthesis and discuss its applications and significance in synthetic biology as well as the life science research as a whole.

**Key words:** DNA synthesis; gene synthesis; genome synthesis; microarray; synthetic biology

合成生物学基本上是通过 DNA 序列的重写来实现其各种目的, 包括设计改造生物大分子、基因网络、代谢途径, 甚至整个染色体或基因组。因此, 基因合成技术对合成生物学便显得尤为重要。

### 1 DNA化学合成

DNA 的化学合成研究从其双螺旋结构被破译后不久便开始进行, 直到 20 世纪 90 年代初才最终发展形成目前 DNA 合成仪上普遍使用的固相亚磷酰胺三酯合成法 (phosphoramidite chemistry)<sup>[1-2]</sup>。该方法以多孔玻璃 (controlled pore glass, CPG) 或聚苯乙烯 (polystyrene, PS) 作为固相载体, 通过一个由活化 (deprotection)、连接 (coupling)、封闭 (capping) 和氧化 (oxidation) 四步化学反应组成的循环反应连

接每一个核苷酸, 由 3' 端向 5' 端延伸合成寡核苷酸 (oligonucleotide), 最后对合成好的寡核苷酸用氨水进行切割和脱保护基处理。

固相亚磷酰胺三酯 DNA 合成法每步的效率一般在 97%~99%。为达到较高的产率, 化学合成的寡核苷酸长度一般在 100 bp 以内, 因特殊需要可到 200~300 bp 以上, 但产率不高, 而且长度越长, 纯度越低, 积累的合成错误越多。以 60-mer 为例, 若每步合成效率为 98%, 最终产物里全长的 60-mer

收稿日期: 2011-05-27

基金项目: 国家重点基础研究发展计划(“973”项目)  
(2011CBA00800)

通信作者: E-mail: jtian@duke.edu

仅有不足 30%，70% 以上为短链或含有其他错误，因此，最好纯化后再使用。常用的纯化全长产物的方法有聚丙烯酰胺凝胶电泳法 (PAGE)、亲和吸附法 (OPC/HAP) 和液相层析法 (HPLC)，纯化后的纯度可达 90% 以上<sup>[3]</sup>。

## 2 基因和基因组合成

### 2.1 基因合成

DNA 化学合成的片断长度有限，因而要合成更长的基因需要将这些短片段拼接起来。常用的体外拼接方法有两种：连接酶拼接法和聚合酶拼接法 (图 1)。

连接酶拼接法 (ligase chain reaction, LCR) 是通过热循环反应，用嗜热性 DNA 连接酶将首尾相连、重叠杂交的寡核苷酸片断连接起来的反应<sup>[4-7]</sup> (图 1A)。聚合酶拼接法 (polymerase cycling assembly, PCA) 则直接利用嗜热性 DNA 聚合酶，通过热循环反应将两头重叠杂交的寡核苷酸片断步步延伸，最终合成全长基因<sup>[8-10]</sup> (图 1B)。两种方法的结果相差不多，聚合酶拼接法允许寡核苷酸片断之间有间隔，因而稍微省料和省时；连接酶拼接法要求使用 5'-端磷酸化的寡核苷酸片断，相邻片断之间不能有间隔。对一些较困难的基因合成可尝试连接酶拼接法。如果杂交条件控制得好，其错误率可能相对较低一些。

另外，最近发现两端重叠杂交的寡核苷酸片断

也可以在酵母细胞内被拼接起来并克隆到同时转化进去的载体中<sup>[11]</sup>。

### 2.2 DNA 大片断和基因组的合成

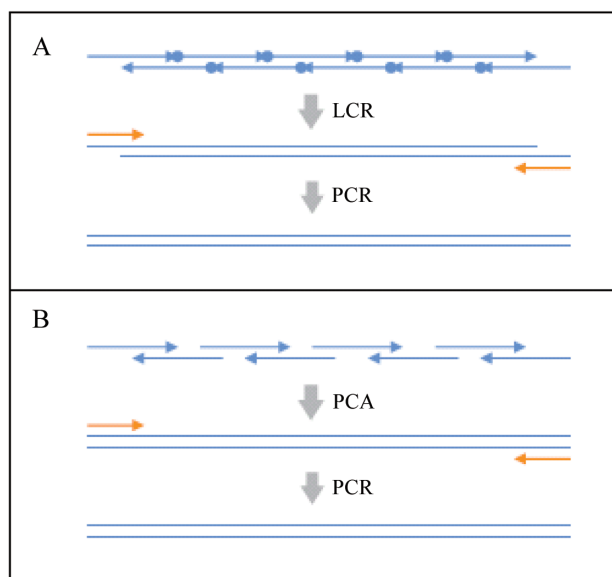
上述两种酶反应，连接酶拼接法和聚合酶拼接法，多用于合成几千个碱基对以下的基因序列。更长的大片断则需由这些中等长度的片断进一步拼接而成，拼接方法包括体外酶反应法和体内连接法。

体外酶反应拼接法有多种 (图 2)。聚合酶拼接法利用一对末端引物和嗜热性聚合酶在温度循环反应中对两端有同源序列的双链片断进行延伸和扩增 (图 2A)<sup>[6,12-15]</sup>，其优点是操作简单，但由于聚合酶功能的限制，合成的大片段长度有限，约在 20 kb 以下。环形聚合酶延伸法 (circular polymerase extension cloning, CPEC) 利用类似的机制 (但不需要扩增引物) 可将末端重叠的多个片断和载体一步连接成完整的环状质粒，然后直接转化入细胞 (图 2B)。该方法已成为一种新型的方便快捷的分子克隆和 DNA 片断组装方法<sup>[16-17]</sup>。体外同源重组法是利用 T5 外切酶将末端有较长同源序列的双链片断的 5'-端修短，使突出的 3'-端相互杂交，然后用嗜热性聚合酶和连接酶将缺口补齐和连接 (图 2C)<sup>[18-19]</sup>。这种方法可以合成长至几十万碱基对的大片段。

体内连接法主要利用酵母细胞的同源重组功能，将同时转化进去的多个同源片断自动修复整合成为一个环状质粒或基因组。该方法已经被用来克隆大片段，组装多酶生化合成路径，以及合成小型细胞基因组<sup>[11,20-22]</sup>。

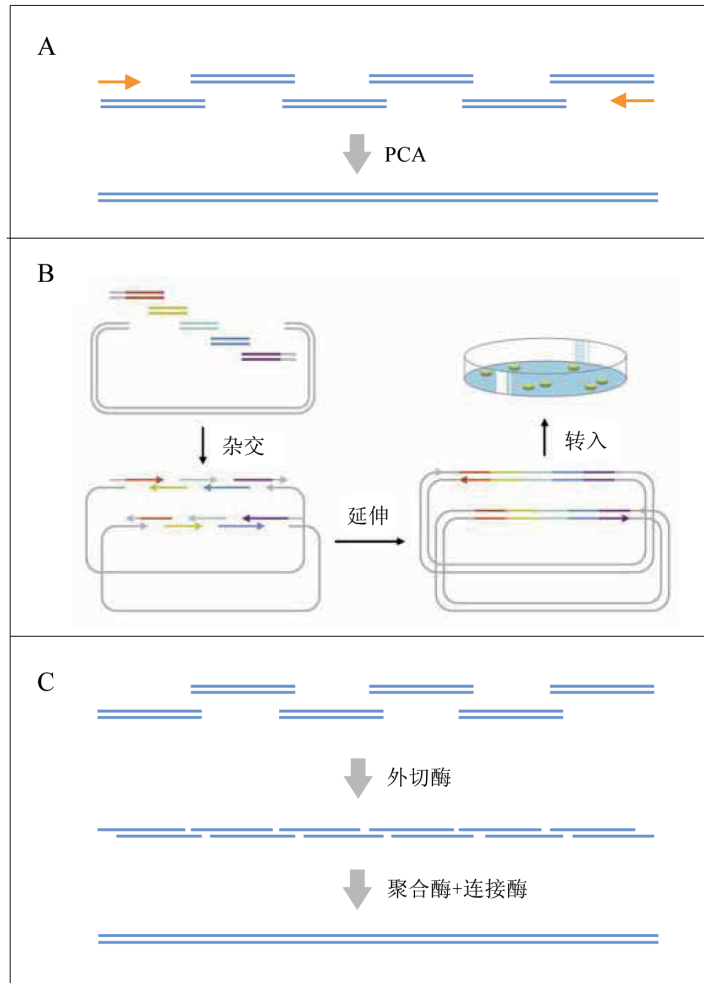
## 3 大规模自动化基因合成

从以上描述不难看出，基因合成包括寡核苷酸化学合成、纯化、连接、扩增，甚至更多后续组装步骤，程序繁琐，成本昂贵。从发展进程上看，基因合成技术远远落后于基因测序技术。为了提高基因合成通量并降低成本，人们开始探索基因合成的微型化和自动化技术<sup>[23]</sup>。2004 年，我们和其他实验室一起研制出用 DNA 芯片合成大量寡核苷酸片断进行基因组组装的方法<sup>[13,24-25]</sup>。由于一块芯片上可合成上万个不同的寡核苷酸序列，而每种序列的量却很少，这给高效基因合成带来一个难题。随后的研究在提高基因组组装效率和选择性扩增寡核苷酸库等方面取得了一定进展<sup>[26-27]</sup>，但仍无法解决微型化和自动化的问题。我们实验室最近研发了一种新的芯片基因合成技术，将寡核苷酸合成、扩增和连接组装等步骤集成到一块塑料芯片上，从而初步解决



A: 连接酶拼接法(LCR)实心圆代表磷酸化的5'端; B: 聚合酶拼接法(PCA)

图1 常用基因合成方法



A: 聚合酶拼接法; B: 环形聚合酶延伸法(CPEC); C: 体外同源重组法

图2 常用大片段和基因组合成方法

了这一难题 (图 3)<sup>[28]</sup>。该技术使用易加工的塑料芯片, 并将整块芯片划分成多个小室, 在每一小室中合成一组寡核苷酸序列, 用恒温扩增的方法将这组寡核苷酸序列扩增并释放到小室中, 最后利用聚合酶拼接法将它们组装成基因。使用该技术可将基因合成

时间从 2~3 周缩短到 2~3 d, 成本从每对碱基 0.5~1 美元降低到 0.5 美分, 同时大大降低了原料用量<sup>[28]</sup>。

#### 4 发展趋势及展望

综上所述, 基因合成技术正在进入一个快速发

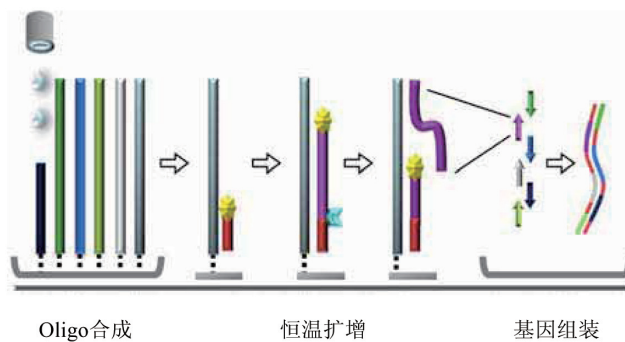


图3 芯片基因合成<sup>[28]</sup>

展的阶段。可以预见在不久的将来,基因合成将会变得更加快捷、廉价。大规模基因和基因组合成将成为现实。基因合成的功用将越来越广泛,有可能取代分子克隆技术而广泛应用于生物医药研究领域。复杂合成基因库的构建会加速蛋白质结构功能的研究和设计,快速优化和完善基因设计。在最近的一次演示实验中,我们实验室通过设计、合成和筛选由同义密码子构成的基因库,快速高效地优化了75个果蝇基因在大肠杆菌中的表达<sup>[28]</sup>,这用目前的基因设计软件是很难做到的。大规模基因和基因组的合成和基因优化技术还极有助于新代谢途径和新基因组的设计和构建,为生物基合成奠定基础。

### [参 考 文 献]

- [1] Caruthers MH, Beaucage SL, Becker C, et al. Deoxyoligonucleotide synthesis via the phosphoramidite method. *Gene Amplif Anal*, 1983, 3: 1-26
- [2] Caruthers MH. Gene synthesis machines: DNA chemistry and its uses. *Science*, 1985, 230: 281-5
- [3] Ellington A, Pollard JD, Jr. Introduction to the synthesis and purification of oligonucleotides[M]// Beaucage SL. *Current protocols in nucleic acid chemistry*. Appendix 3, Appendix 3C. USA: John Wiley & Sons, 2001
- [4] Wiedmann M, Wilson WJ, Czajka J, et al. Ligase chain reaction (LCR)--overview and applications. *PCR Methods Appl*, 1994, 3: S51-64
- [5] Au LC, Yang FY, Yang WJ, et al. Gene synthesis by a LCR-based approach: high-level production of leptin-L54 using synthetic gene in *Escherichia coli*. *Biochem Biophys Res Commun*, 1998, 248: 200-3
- [6] Smith HO, Hutchison CA 3rd, Pfannkoch C, et al. Generating a synthetic genome by whole genome assembly: phiX174 bacteriophage from synthetic oligonucleotides. *Proc Natl Acad Sci USA*, 2003, 100: 15440-5
- [7] Bang D, Church GM. Gene synthesis by circular assembly amplification. *Nat Methods*, 2008, 5: 37-9
- [8] Dillon PJ, Rosen CA. A rapid method for the construction of synthetic genes using the polymerase chain reaction. *Biotechniques*, 1990, 9: 298, 300
- [9] Stemmer WP, Cramer A, Ha KD, et al. Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. *Gene*, 1995, 164: 49-53
- [10] Prodromou C, Pearl LH. Recursive PCR: a novel technique for total gene synthesis. *Protein Eng*, 1992, 5: 827-9
- [11] Gibson DG. Synthesis of DNA fragments in yeast by one-step assembly of overlapping oligonucleotides. *Nucleic Acids Res*, 2009, 37: 6984-90
- [12] Cello J, Paul AV, Wimmer E. Chemical synthesis of poliovirus cDNA: generation of infectious virus in the absence of natural template. *Science*, 2002, 297: 1016-8
- [13] Tian J, Gong H, Sheng N, et al. Accurate multiplex gene synthesis from programmable DNA microchips. *Nature*, 2004, 432: 1050-4
- [14] Kodumal SJ, Patel KG, Reid R, et al. Total synthesis of long DNA sequences: synthesis of a contiguous 32-kb polyketide synthase gene cluster. *Proc Natl Acad Sci USA*, 2004, 101: 15573-8
- [15] Shevchuk NA, Bryksin AV, Nusinovich YA, et al. Construction of long DNA molecules using long PCR-based fusion of several fragments simultaneously. *Nucleic Acids Res*, 2004, 32: e19
- [16] Quan J, Tian J. Circular polymerase extension cloning for high-throughput cloning of complex and combinatorial DNA libraries. *Nat Protocols*, 2011, 6: 242-51
- [17] Quan J, Tian J. Circular polymerase extension cloning of complex gene libraries and pathways. *PLoS ONE*, 2009, 4: e6441
- [18] Gibson DG, Benders GA, Andrews-Pfannkoch C, et al. Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science*, 2008, 319: 1215-20
- [19] Gibson DG, Young L, Chuang RY, et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods*, 2009, 6: 343-5
- [20] Ma H, Kunes S, Schatz PJ, et al. Plasmid construction by homologous recombination in yeast. *Gene*, 1987, 58: 201-16
- [21] Gibson DG, Benders GA, Axelrod KC, et al. One-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic *Mycoplasma genitalium* genome. *Proc Natl Acad Sci USA*, 2008, 105: 20404-9
- [22] Shao Z, Zhao H. DNA assembler, an *in vivo* genetic method for rapid construction of biochemical pathways. *Nucleic Acids Res*, 2009, 37: e16
- [23] Tian J, Ma K, Saaem I. Advancing high-throughput gene synthesis technology. *Mol BioSyst*, 2009, 5: 714-22
- [24] Zhou X, Cai S, Hong A, et al. Microfluidic PicoArray synthesis of oligodeoxynucleotides and simultaneous assembling of multiple DNA sequences. *Nucleic Acids Res*, 2004, 32: 5409-17
- [25] Richmond KE, Li MH, Rodesch MJ, et al. Amplification and assembly of chip-eluted DNA (AACED): a method for high-throughput gene synthesis. *Nucleic Acids Res*, 2004, 32: 5011-8
- [26] Borovkov AY, Loskutov AV, Robida MD, et al. High-quality gene assembly directly from unpurified mixtures of microarray-synthesized oligonucleotides. *Nucleic Acids Res*, 2010, 38: e180
- [27] Kosuri S, Eroshenko N, Leproust EM, et al. Scalable gene synthesis by selective amplification of DNA pools from high-fidelity microchips. *Nat Biotechnol*, 2010, 28: 1295-9
- [28] Quan J, Saaem I, Tang N, et al. Parallel on-chip gene synthesis and application to optimization of protein expression. *Nat Biotechnol*, 2011, 29: 449-52