

综述

细菌必需基因、最小基因组和合成细胞

邱东茹

中国科学院水生生物研究所, 湖北 武汉 430072

邱东茹. 细菌必需基因、最小基因组和合成细胞. 生物工程学报, 2012, 28(5): 540-549.

Qiu DR. Essential genes, minimal genome and synthetic cell of bacteria: a review. Chin J Biotech, 2012, 28(5): 540-549.

摘要: 单细胞原核生物是原始的细胞生命形式, 确定细菌必需基因和最小基因组对理解生命的本质、细胞生命的起源和进化有非常重要的意义。文中简要介绍近年来有关细菌的必需基因、最小基因组和合成细胞的研究方法、理论和进展。还特别介绍人工建立最小细菌基因组的策略以及应用前景。

关键词: 必需基因, 最小基因组, 合成细胞, 虚拟细胞

Essential genes, minimal genome and synthetic cell of bacteria: a review

Dongru Qiu

Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, Hubei, China

Abstract: Single-cell prokaryotes represent a simple and primitive cellular life form. The identification of the essential genes of bacteria and the minimal genome for the free-living cellular life could provide insights into the origin, evolution, and essence of life forms. The principles, methodology, and recent progresses in the identification of essential genes and minimal genome and the creation of synthetic cells are reviewed and particularly the strategies for creating the minimal genome and the potential applications are introduced.

Keywords: essential genes, minimal genome, synthetic cell, virtual cell

Received: November 10, 2011; **Accepted:** January 16, 2012

Supported by: Knowledge Innovation Program of the Chinese Academy of Sciences (No. Y15103-1-401).

Corresponding author: Dongru Qiu. Tel/Fax: +86-27-68780215; E-mail: qiu@ihb.ac.cn

中国科学院知识创新工程项目 (No. Y15103-1-401) 资助。

分子生物学和基因组学为探索生命的起源和进化带来全新的视角和方法。依据核糖体 RNA 序列,现存的生命形式被分为 3 个域 (Domain),即真细菌、古菌和真核生物。这些生物的遗传物质都是核酸,其基因组大小变化很大,从数十万碱基对到几十亿碱基对不等;所含基因数目则为数百乃至数万。原核生物的基因组较小,基因结构和基因调控网络相对简单。目前有 2 000 余个原核生物的基因组序列被测定和注释。许多真核生物包括酵母、秀丽线虫、果蝇、拟南芥、杨树、水稻、斑马鱼、小鼠、猩猩和人类自身的基因组测序也已完成。第 2 代测序技术更大大提高测序的速度和通量并降低了测序成本,如 DNA 双螺旋结构的发现者、诺贝尔奖得主 James Watson 和基因组学家 J. Craig Venter 的个人基因组被测序,个性化医疗很有可能成为现实。这些基因组序列揭开了生命天书的本来面目,但是解读这些天书的意义尚需时日、任重道远。随着技术的进步,以大规模高通量分析为特征的各种组学应运而生,包括基因组学、转录组学、蛋白组学、相互作用组学和代谢组学,将揭示基因组复制、基因转录、翻译和基因调控网络、蛋白质相互作用和物质能量代谢等不同层次的相互关联、错综复杂的生命活动。生物信息学、计算生物学和系统生物学就是为整合和诠释这些海量的数据而产生的,其重要性日益突出。

细胞是生命活动的基本单位,细胞生命的 3 大特征是维持正常代谢平衡、进行繁殖(自我复制)以及进化。早在上世纪上叶 DNA 双螺旋结构揭示之前,著名物理学家薛定谔(Erwin Schrödinger)在《生命是什么?》一书^[1]中提出

了有关生命的本质的问题,试图从物理学的角度来理解生命现象。在当今的后基因组时代,这个有关生命本质的问题依然存在,也更为直观,即维持细胞生命所必需的物质基础和过程是什么?所需的最小基因群(Minimal gene-sets)、或者说最小基因组(Minimal genome)有多大?初步的答案是维持自由生活细胞的必需基因数目大约为 300 个左右,相应的基因组大小约为 300~400 kb。不仅如此,人工建立最小基因组或最小化细菌(Minimal bacteria)的工作已经开始进行尝试,其中最为突出的成果是最近美国科学家成功用人工合成的细菌染色体制造出合成细胞。未来所建成的具有最小基因组工程菌可能用作人工细胞工厂的“底盘”(Chassis)细胞。

1 鉴定必需基因和最小基因组的方法

有些基因对细胞生命具有至关重要的作用,称为必需基因(Essential gene),其突变通常是致死性的。细菌特定基因的必要性还取决于环境条件。因为寄主细胞内环境条件稳定,营养供应充足,细胞内共(寄)生细菌细胞结构和代谢途径通常极度简化,细胞壁退化乃至消失。相应地,胞内细菌如支原体和蚜虫内共生菌 *Buchnera* 基因组大为缩减,所测序的胞内共生细菌基因组大小的下限纪录不断降低^[2]。目前发现细菌 *Carsonella ruddii* 的基因组最小,仅为 160 kb,基因分布非常致密,有 182 个开放阅读框,90% 的相邻开放阅读框间有所重叠^[3]。更为特别的是缺乏许多必需基因,包括细胞外被膜生成和核酸、脂类代谢基因。作者推测有些必需基因已像

细胞器基因一样转移到寄主的核基因组中。这些细胞内共生细菌的基因组进化证据也印证了线粒体和叶绿体的内共生学说,即真核生物的线粒体和植物的叶绿体分别来源于细胞内寄生的细菌和蓝细菌^[4]。目前可用人工培养基培养的细菌是生殖支原体 *Mycoplasma genitalium*, 基因组仅为 580 kb, 可自由生活细菌的最小基因组应该与此接近。

然而,绝大多数自由生活的细菌往往处在营养缺乏、环境条件多变的生境中。通常意义上的必需基因是指在丰富培养基中细菌生长繁殖所必需的基因。例如,大肠杆菌野生型可在以单糖作为碳源和能源、含无机盐氮源和磷源及微量元素的简单培养基中正常生长,亦即大肠杆菌可利用单糖和无机盐合成自身所需的氨基酸、脂类、多糖类和核酸。但是在丰富培养基上,合成氨基酸的基因对于大肠杆菌而言是非必需基因。总体而言,细菌的必需基因是合成细胞结构成分、信息传递和加工不可或缺的基因。确定必需基因和最小基因组的方法主要有比较基因组学和系统性基因失活法。

1.1 比较基因组学方法

相对而言,原核生物基因组简单,重复序列较少,因此短枪测序法适合于微生物基因组测序。1995年,流感嗜血杆菌 *Haemophilus influenzae* 基因组测序完成,这是第一个细菌全基因组序列。如前所述,目前已有多个细菌基因组序列已知,而且某些重要的病原细菌如炭疽菌、大肠杆菌和工业细菌有多个菌株的基因组序列被测定;有些细菌菌株的基因组序列甚至被不同机构重复测定。我国也完成了钩端螺旋体、表

皮葡萄球菌、野油菜黄单胞菌、腾冲热泉菌等细菌基因组的测序。这些基因组序列为比较基因组学分析提供了很好的机会,可以用来预测必需基因和最小基因组。其基本思路是必需基因应该是在细菌基因组中非常保守的基因,而非必需基因则不会在所有基因组中出现。美国国家生物技术信息中心 (NCBI) 的 Mushegian 和 Koonin^[5]通过对流感嗜血杆菌和生殖支原体基因组的比较分析,发现大约 256 个基因为两者所共有的保守基因,应是细菌的必需基因。稍后 Koonin 等^[6]对 21 个细菌、古细菌和真核生物的基因组进行了比较基因组学分析,认为 150 个基因可以维持基本的转录、翻译和复制系统以及简化的修复系统和一小群分子伴侣,这个最小化细菌的中间代谢简化成糖酵解,拥有原始的跨膜运输体,没有细胞壁。Koonin^[7]进一步发现仅有 60 个蛋白质为所有生物所共有,其中大部分参与蛋白质翻译,推测现存生物的最后共同祖先 (Simple last universal common ancestor, LUCA) 仅有 500~600 个基因。Gil 等^[8]对自由生活型细菌和胞内共生细菌的基因组进行比较分析,总结出含 206 个编码蛋白质的必需基因核心群。

1.2 系统性基因插入失活方法

经典遗传学的研究是基于基因突变和表型变化的相关关系的分析来确定基因的功能,在细菌遗传学中营养缺陷型的研究曾经扮演了重要的角色。直到今天,剔除特定基因以分析其功能依然是常用的手段,不过分析和检测的手段更为多样和快捷,如应用基因芯片技术。这个方法也被用来确定必需基因和最小基因组。理论上讲,因为剔除必需基因通常是致死性的,必需基因不

可能剔除, 所以可被剔除的基因是非必需基因。笔者的导师——日本庆应大学先端生命科学研究所以和三菱生命科学研究所以的板谷光泰教授 10 余年前率先利用枯草芽胞杆菌 *Bacillus subtilis* 的自然感受态特性 (Natural competence), 采用基因插入失活的实验方法来估计最小基因组大小, 他随机挑选了 79 个基因位点, 发现仅有 6 个位点不能被抗生素基因标识插入, 因而得不到相关突变株。然后通过统计分析方法估算最小基因组大小约为 562 kb^[9]。此后 J. Craig Venter 实验室的 Hutchinson 等^[10]在生殖支原体和肺炎支原体中利用转座子进行了系统的基因插入失活分析, 即分离大量的转座子插入突变株并测序鉴定转座子插入位点, 能被转座子插入所破坏的基因是非必需基因。生殖支原体染色体为 580 kb, 仅有 482 个编码蛋白质的基因, 是已知能用人工培养基培养的基因组最小的细菌^[11]。他们共定位了 2 200 个转座子插入位点, 鉴定出 130 个生殖支原体基因为非必需基因。据此他们估计必需基因的数目为 265~350 个。后来该实验室进一步分离仅含单个转座子插入位点的细菌单克隆并加以定位, 确认 482 个蛋白质编码基因中的 382 个为实验室培养条件下的必需基因, 还有 5 个必需基因有重复基因, 而且 43 个 RNA 编码基因都没有发现转座子插入。这样他们确认 387 个蛋白质编码基因和 43 个 RNA 基因足以维持细胞的生长繁殖。即便如此, 其中占 28% 的必需基因的功能依然未知。此后利用转座子插入突变分析对其他细菌也进行了必需基因的鉴定工作, 如霍乱弧菌 *Vibrio cholerae*^[12]和流感嗜血杆菌^[13]。在金黄葡萄球菌 *Staphylococcus aureus* 中则用反义 RNA 干扰法鉴

定必需基因^[14]。日本和欧盟的多个研究组长期分工合作对枯草芽胞杆菌进行了系统的研究, 包括测序和功能基因组的研究, 在系统性的定位基因插入失活基础上, 预测芽胞杆菌的必需基因为 271 个, 这些基因大多是细胞生长、分裂、染色体复制、转录、翻译、蛋白质折叠、能量转化和物质代谢所需的持家基因^[15]。

就目前研究最为深入的大肠杆菌而言, 依然有近一半的基因功能未明, 即所谓的 Y 基因。这些 Y 基因可能多为特定生境条件下起作用的基因, 而不是维持最基本的细胞代谢和繁殖所需的基因。不过需要指出的是在各种细菌中所鉴定的必需基因仅有少数功能不明, 这说明我们对原核生物细胞的基本生命过程理解的空白不像以前想象的那么大。

必需基因群 (Essential gene sets) 和最小基因组是相似而有所不同的概念, 最小基因组基因数目应该略大于必需基因数。细菌很多基因在功能上互补, 这些基因可以分别加以剔除而不影响细胞存活, 因此都是非必需基因。然而 2 个功能重叠的基因不可能同时加以剔除, 在甲基因失活后, 非必需的乙基因成为必需基因; 反之亦然。这种基因突变之间关系被称为相互排斥 (Mutually exclusive), 即 2 个基因的失活不可能同时发生。与此相反, 有时某个基因不能被剔除, 该基因失活对细菌是致死性的, 似乎该基因是必需基因。然而在另一个基因被先行剔除的情况下, 该基因可以被剔除, 因此实际上是非必需基因。这种基因突变间关系被称为相互包容 (Mutually inclusive), 即 2 个基因可以被同时剔除。以细菌毒素基因和相应的抗毒素基因以及限制性内切

酶基因和相应的 DNA 修饰酶基因为例, 毒素基因和内切酶基因可以被剔除; 但单独剔除抗毒素基因或修饰酶基因是致死性的。理论上讲, 确定必需基因和最小基因组更可靠方法的是逐一剔除非必需基因, 这需要极大的工作量, 而且存在技术上的困难。

2 创造最小基因组的设想和策略

2.1 自上而下 (Top-down) 剔除基因的方法

即通过逐步剔除现存细菌的非必需基因来缩减基因组大小直至制造出最小基因组。我们知道目前用于基因克隆的大肠杆菌菌株都进行过基因改造和优化, 失活了很多与基因重组相关的基因和限制性内切酶-修饰酶系统基因, 使得基因工程质粒更加稳定。如剔除工程菌株的非必需基因如原噬菌体、转座子、插入序列、毒性因子基因和经水平转移而来的外源基因簇, 可使其在实验室和工业应用中更加有效、稳定、安全。例如, 有些原噬菌体的活化可造成工程菌的溶菌反应, 转座子和插入序列的活动也造成一些关键基因的失活和菌株退化。目前已有一些初步的尝试来降低细菌基因组大小, 如欧盟学者利用同源重组剔除了枯草芽胞杆菌 Marburg 168 株的 332 个非必需基因, 使其基因组缩减了 7.7%^[16], 其中包括一些原噬菌体、高 AT 岛和大型的抗生素合成基因簇。在大肠杆菌中则有更多的尝试^[17-19], 可利用的技术有噬菌体 λ Red 重组酶切除技术, I-SceI 内切酶切除技术和 Cre/loxP 切除技术。在其他一些细菌中, 也有相应的技术可以利用, 原理大同小异, 即利用细菌本身的同源重组系统 (例如枯草芽胞杆菌) 或外源的同源重组系统

(如 λ Red 噬菌体), 首先人为向染色体中插入重复序列 (如 Tn5 转座子), 经染色体内部同源重组可以切除两段重复序列之间的染色体 DNA 序列。威斯康星大学 Fredrick R. Blattner 教授领导的研究组逐步剔除了大肠杆菌 K-12 株 15% 的非必需基因, 发现基因组简化后的细菌不仅生理正常、还具有一些明显的优点, 如电穿孔转化效率提高、质粒更为稳定^[20]。

2.2 自下而上 (Bottom-up) 合成的方法

2002 年美国纽约州立大学石溪分校 Eckard Wimmer 教授实验室成功利用化学方法从头 (*de novo*) 合成了有感染活力的小儿麻痹症病毒 RNA^[21]。2003 年 Venter 和合作者诺贝尔化学奖得主 Hamilton O. Smith 等也合成了长 5 386 碱基对的 FX-174 噬菌体基因组。对 Venter 等而言合成小型的病毒基因组只不过是牛刀小试, 他们试图设计并合成新型细菌基因组, 创造全新的细胞生命。具体方法是首先合成小片段核酸, 再将核酸片段缀合起来形成人工染色体, 然后将这条人工染色体注入本身 DNA 已被破坏的细菌细胞中, 人工染色体可以接管、利用现有的细胞机器进行基本的生命活动并合成新的细胞成分, 以逐步取代原先的成分, 最终变成全新的细胞, 为特定的应用目标服务^[22]。这种设想类似于克隆动物所用的核移植技术和计算机软件启动 (Boot up) 计算机硬件系统运行。虽然这个项目受到生物伦理学家一定程度的批评和质疑^[23], 还是获得美国能源部 (DOE) 的资助和风险投资家的青睐。2007 年他们成功将丝状支原体 *Mycoplasma mycoides* 染色体移植到另一个近缘种山羊支原体 *Mycoplasma capricolum* 细胞中^[24], 他们事先

剔除了受体菌的修饰-限制系统基因。2008年初 Venter 和 Smith 小组报道成功人工设计和化学合成生殖支原体 *Mycoplasma genitalium* 的染色体^[25]，他们称之为 *Mycoplasma laboratorium* (实验室支原体) JCVI-1.0 (J. Craig Venter 研究所版本 1.0)。但是染色体组装的最后几个步骤未能完全在体外试管中完成，还是借助了细菌和酵母人工染色体技术，因此并非全化学合成，他们也未能将这个人工染色体转植到山羊支原体 *M. capricolum* 受体细胞中。他们有意在人工染色体上缺失天然染色体的几个毒素基因并在添加了供鉴定之用的水印 (Watermarks) 序列，未对基因组做其他重大修改。2009年 Venter 研究组利用酵母克隆了丝状支原体 *M. mycoides* 的染色体并将之移植到山羊支原体受体细胞中^[26]。2010年 Venter 研究所报道人工合成丝状支原体的染色体 (JCVI-syn1.0) 并将之转入聚乙二醇处理过的山羊支原体受体细胞^[27]。在受体细胞分裂时，宿主的天然染色体和转入的人工染色体分配到不同的子细胞。从人工染色体转录、翻译的蛋白质和合成的细胞成分逐渐取代原有的细胞成分，成为全新的丝状支原体细胞，他们称之为合成细胞 (Synthetic cell)。这一途径类似强化的基因克隆，并非像他们最初设计的那样将合成的染色体转入去除本身 DNA 的空细菌细胞中并使其接管细胞机器、合成全新细胞所需的组分。这一成就在世界范围内引起很大的反响，但大多数同行生物学家对声称创造“合成细胞”的说法大多不以为然，因为这个人工染色体绝大部分信息来自天然细胞而且大部分工作的完成也需借助天然细胞。一般认为人造细菌因缺乏抗逆基因和机

制，非常脆弱，不可能独立在自然条件下生存，应该不会带来很大的环境和生态风险。

2.3 人工细胞的计算机辅助设计和虚拟细胞

正如常规的工程和制造项目一样，为制造最优人工细胞，利用计算机模拟来辅助蓝图设计是必要的一步。当今生物学也从描述性向量化发展，生物信息学和系统生物学将遗传发育和细胞物质能量代谢等生命活动信息化、数字化，以整体性、全方位角度来模拟、诠释生命活动的动态性质。笔者的导师——庆应大学富田胜教授领导的研究室与 Venter 实验室合作，根据支原体测序和基因组分析的结果，1997年率先成功利用所开发的电子细胞 (E-Cell) 软件包建立了包括 127 个必需基因的计算机细胞模型^[28]。这个高度简化的虚拟细胞 (Virtual cell) 能进行转录和翻译，利用糖酵解途径分解葡萄糖产生 ATP，从环境中吸收脂肪酸和甘油合成细胞膜所需的磷脂，但它不能繁殖和进化。后来他们又建立了红血球细胞和线粒体等模型。大肠杆菌是已有百年研究历史、研究最为深入的模型生物，时至今日也是最常用的生物工程菌和工具。相关生理生化和代谢物定量数据也最为丰富，代谢网络和代谢物流束 (Metabolite fluxes) 分析与模拟已经有较长的历史。富田实验室计划用 10~20 年时间建立大肠杆菌的全细胞计算机模型。数年前即建立先端生命科学研究所，组织了包括生物学、分析化学和生物信息学等多学科的专家群体，采取“干”、“湿”并举的策略，不但广泛收集文献中的现有数据，并进行大量的分析测试工作，大规模采集包括大肠杆菌的转录组学、蛋白组学、相互作用组学和代谢组学的的数据。其代谢组学分析以毛细管

电泳-质谱联用分析为主,也利用碳同位素标记分析大肠杆菌代谢物质流束,规模居世界之冠。同时也为味精、酿酒、食品和制药等企业及大学和研究开发机构提供代谢组学分析服务。最近他们在美国 *Science* 杂志上报道了大规模、多层次、高通量的分析结果,表明大肠杆菌的代谢物质网络非常健壮 (Robust),个别能量代谢酶基因的剔除 (遗传扰动) 对基因表达、核糖核酸、蛋白质和代谢物质水平的干扰很小,因为代谢流束可能转向进入其他代谢途径。虽然为响应环境条件的变化,大肠杆菌可非常活跃地调节基因的表达 (核糖核酸和蛋白质) 来维持稳定代谢物水平,代谢物质量变化很小^[29]。此外,加拿大阿尔伯塔大学生物分子设计研究院的 Gordon Broderick 和 Michael Ellison 教授所建立的模型成为网络细胞 (Cyber Cell),不同于常用的基于微分方程的代谢网络和流束分析模型,他们致力于高清晰度的原子分子水平细胞活动的模拟。美国加州大学圣迭戈分校的 Bernard Palsson 教授领导的实验室建立了包括大肠杆菌模型在内的一系列硅片上生物 (*In silico* organisms) 和模拟平台 SimPheny。欧洲方面,荷兰阿姆斯特丹自由大学的 Hans Westerhoff 教授也发起了硅片细胞 (Silicon cell) 项目。美国由普渡大学的 Barry Wanner 教授牵头发起了大肠杆菌模型细胞协作组 (*E. coli* model cell consortium, Emc², 源自爱因斯坦质能方程 $E=mc^2$)。2002 年相关学者组成了国际大肠杆菌联盟 (International *E. coli* Alliance, IECA),有北美、日本和欧洲的多位科学家参与合作。1999 年美国康涅狄格大学健康科学中心的 Leslie Loew 博士在国立卫生研究院的

支持下建立虚拟细胞 (Virtual cell) 和 NRCAM 项目 (National resource for cell analysis and modelling)。他们主要以真核细胞为模拟对象,如建立神经细胞钙离子转运的模型。目前这些计算机建模工作遇到的最大问题是缺乏统一、可靠的定量数据,最完备的生物代谢途径和代谢物质数据库是日本京都大学化学研究所金久实教授领导建立的京都基因和基因组百科全书 (KEGG),包括 KEGG LIGAND 数据库和 KEGG PATHWAY 数据库^[30-31]。

前几年由德国、英国和匈牙利学者组成的研究组利用大肠杆菌代谢物流束平衡模型来模拟昆虫细胞内共生细菌 *Buchnera aphidicola* 和 *Wigglesworthia glossinidia* 的进化过程。大肠杆菌是这两种细菌的远源祖先。模拟结果表明连续性的基因丢失可导致最小核心代谢网络的形成,并根据两种胞内共生细菌各自的环境条件预测出其基因组成上的差异^[32]。另外,他们认为单个基因失活大大低估了自由生活型细菌所需的最小基因组大小,可能达 45% 之多,即实际上可能的最小基因组应为该方法所估计的 2 倍。这样实际的最小基因组就接近现存可以用培养基培养 (亦即能脱离宿主细胞自由生长) 的细菌支原体的基因组大小。

3 最小基因组的应用前景

必需基因和最小基因组的研究不仅具有理论意义,即探讨生命的本质和生命的起源和进化问题,而且具有重要的实用价值。病原细菌抗药性是愈来愈严重的问题之一,而筛选、开发新型抗生素的努力还远远不够,鉴定细菌必需基因可

以提供新的药物靶标。现存药用抗生素的作用靶标仅集中在 15 个必需基因所编码的细菌正常生长和繁殖所必需的细胞成分,如细胞壁和核糖体^[10]。

人类利用细菌发酵已经有很长的历史,如乳酸菌和醋酸菌。目前经筛选和基因改造的工程菌在工业产品生产、新材料、能源和环境污染净化等多个领域具有更加广泛的应用。最近日本学者将阿维链霉菌 *Streptomyces avermitilis* 8.2 Mb (百万碱基对) 的线状染色体中约 1.4 Mb 的片段切除后可以作为异源基因簇的表达宿主高效合成多种抗生素^[33]。

美国和日本均投入巨资进行微生物基因组和功能基因组学的研究,美国有关研究由能源部支持和管理,日本方面的研究则由经产省新能源开发机构 (NEDO) 支持,日本也特别重视极端微生物如深海微生物的研究。DOE 和世界各国其他机构已支持完成超过 1 000 个细菌基因组的测序,目前主要进行功能基因组学研究。另外,环境基因组学也为微生物的开发利用带来新的机遇。目前可分离培养的微生物不到 1%, 绝大多数的细菌难以人工培养,阻碍了研究开发和应用。以前某些分子生物学技术已被应用于环境微生物学和微生物检测等。现在环境宏基因组学 (Metagenomics) 方兴未艾,可直接分离环境中的 DNA 利用短枪法和第二代测序技术进行大规模测序,以计算机组装、重建还不能分离纯培养的细菌完整的基因组序列,从中可以发现有用的遗传资源,如变形菌视紫红质 (Proteorhodopsin) 基因,并确定特定生境的微生物群落结构。还可以提供可用的遗传资源信息,

有朝一日可以通过合成细菌直接组装利用这些遗传资源。

人工建立最小基因组具有应用潜力,工程细菌的基因组越简单,细胞的结构和代谢途径也越简化,用于维持自身生长繁殖所需的资源和能量越少,细菌可更为有效地将资源转化为人类所需的特定产品,减少资源浪费。另一方面,由于工程菌本身的代谢物成分较为简单,目标产物如重组蛋白的分离和纯化更为容易。日本经产省新能源开发机构于 2001 年启动“最小基因组工厂”(Minimum genome factory) 研究计划,旨在构建基因组的工程菌用于工业生产。另外未来人工细菌或工程菌也可能用作疫苗和药物的载体。可以人工合成减毒菌以获得更安全的疫苗,目前美国已有公司在开展这一方面的研发工作。

另外,与缩减基因组的努力相反,日本庆应大学的板谷光泰教授和同事在 NEDO 的支持下,经过数年的努力,成功将蓝藻集胞藻 *Synechocystis* PCC6803 的整个基因组逐渐转入并分成 4 个大片段整合到枯草芽胞杆菌的染色体中,创造了前所未有的杂种细菌^[34]。这个细菌中大部分蓝藻基因沉默,也没有光合活性,因为蓝藻核糖体 RNA 操纵子基因不能成功转入芽胞杆菌,正因为如此,蓝藻染色体才能保持稳定。与此类似,在人工合成生殖支原体染色体工作中所克隆 DNA 序列上的基因产物可能大都是不完整的断片而毒性很小,因为其色氨酸密码子 (UGA) 在宿主中是翻译终止密码子^[25]。后来,板谷小组又成功将水稻的叶绿体基因组和小鼠的线粒体基因组转入芽胞杆菌^[35]。这说明除缩减细菌基因组规模外,目前微生物学家已经有能力

大规模操纵和改造细菌基因组, 可以将分散在适应不同生境和生态位的不同细菌中的基因资源加以整合和集成, 提高资源和能源转化效率, 创造用于分解难降解污染物或合成新型化合物的生物代谢途径和能源转化的生物。此外微生物燃料电池和微藻产油、产氢技术虽然尚不能大规模应用, 但欧洲、美国和日本的相关研究并没有中断。微生物基因组工程、代谢工程和刚刚起步的合成生物学 (Synthetic biology) 也许是解决未来世界能源、资源和环境问题的关键所在。

REFERENCES

- [1] Schrödinger E. What Is Life? The Physical Aspect of the Living Cell. 1st ed. Cambridge: The University Press, 1944: 1–92.
- [2] Pérez-Brocá V, Gil R, Ramos S, et al. A small microbial genome: the end of a long symbiotic relationship? *Science*, 2006, 314(5797): 312–313.
- [3] Nakabachi A, Yamashita A, Toh H, et al. The 160-kilobase genome of the bacterial endosymbiont *Carsonella*. *Science*, 2006, 314(5797): 267.
- [4] Andersson SGE. The bacterial world gets smaller. *Science*, 2006, 314(5797): 259–260.
- [5] Mushegian AR, Koonin EV. A minimal gene set for cellular life derived by comparison of complete bacterial genomes. *Proc Natl Acad Sci USA*, 1996, 93(19): 10268–10273.
- [6] Koonin EV. How many genes can make a cell: the minimal-gene-set concept. *Annu Rev Genomics Hum Genet*, 2000, 1(1): 99–116.
- [7] Koonin EV. Comparative genomics, minimal gene-sets and the last universal common ancestor. *Nature Rev Microbiol*, 2003, 1(2): 127–136.
- [8] Gil R, Silva FJ, Peretó J, et al. Determination of the core of a minimal bacterial gene set. *Microbiol Mol Biol Rev*, 2004, 68(3): 518–537.
- [9] Itaya M. An estimation of minimal genome size required for life. *FEBS Lett*, 1995, 362(3): 257–260.
- [10] Hutchison CA, Peterson SN, Gill SR, et al. Global transposon mutagenesis and a minimal *Mycoplasma* genome. *Science*, 1999, 286(5447): 2165–2169.
- [11] Fraser CM, Gocayne JD, White O, et al. The minimal gene complement of *Mycoplasma genitalium*. *Science*, 1995, 270(5235): 397–404.
- [12] Judson N, Mekalanos JJ. TnAraOut, a transposon-based approach to identify and characterize essential bacterial genes. *Nat Biotechnol*, 2000, 18(7): 740–745.
- [13] Akerley BJ, Rubin EJ, Novick VL, et al. A genome-scale analysis for identification of genes required for growth or survival of *Haemophilus influenzae*. *Proc Natl Acad Sci USA*, 2002, 99(2): 966–971.
- [14] Forsyth RA, Haselbeck RJ, Ohlsen KL, et al. A genome-wide strategy for the identification of essential genes in *Staphylococcus aureus*. *Mol Microbiol*, 2002, 43(6): 1387–1400.
- [15] Kobayashi K, Ehrlich SD, Albertini A, et al. Essential *Bacillus subtilis* genes. *Proc Natl Acad Sci USA*, 2003, 100(8): 4678–4683.
- [16] Westers H, Dorenbos R, Van Dijl JM, et al. Genome engineering reveals large dispensable regions in *Bacillus subtilis*. *Mol Biol Evol*, 2003, 20(12): 2076–2090.
- [17] Yu BJ, Sung BH, Koob MD, et al. Minimization of the *Escherichia coli* genome using a Tn 5-targeted Cre/loxP excision system. *Nat Biotechnol*, 2002, 20(10): 1018–1023.
- [18] Kolisnychenko V, Plunkett G, Herring CD, et al. Engineering a reduced *Escherichia coli* genome. *Genome Res*, 2002, 12(4): 640–647.
- [19] Goryshin IY, Naumann TA, Apodaca J, et al. Chromosomal deletion formation system based on Tn5 double transposition: use for making minimal genomes and essential gene analysis. *Genome Res*, 2003, 13(4): 644–653.
- [20] Pósfai G, Plunkett G III, Fehér T, et al. Emergent properties of reduced-genome *Escherichia coli*. *Science*, 2006, 312(5776): 1044–1046.
- [21] Cello J, Paul AV, Wimmer E. Chemical synthesis

- of poliovirus cDNA: generation of infectious virus in the absence of natural template. *Science*, 2002, 297(5583): 1016–1018.
- [22] Zimmer C. Tinker, tailor: can Venter stitch together a genome from scratch? *Science*, 2003, 299(5609): 1006–1007.
- [23] Cho MK, Magnus D, Caplan AL, et al. Ethical considerations in synthesizing a minimal genome. *Science*, 1999, 286(5447): 2087–2090.
- [24] Lartigue C, Glass JI, Alperovich N, et al. Genome transplantation in bacteria: changing one species to another. *Science*, 2007, 317(5838): 632–638.
- [25] Gibson DG, Benders GA, Andrews-Pfannkoch C, et al. Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science*, 2008, 319(5867): 1215–1220.
- [26] Lartigue C, Vashee S, Algire MA, et al. Creating bacterial strains from genomes that have been cloned and engineered in yeast. *Science*, 2009, 325(5948): 1693–1696.
- [27] Gibson DG, Glass JI, Lartigue C, et al. Creation of a bacterial cell controlled by a chemically synthesized genome. *Science*, 2010, 329(5987): 52–56.
- [28] Tomita M, Hashimoto K, Takahashi K, et al. E-CELL: software environment for whole-cell simulation. *Bioinformatics*, 1999, 15(1): 72–84.
- [29] Ishii N, Nakahigashi K, Baba T, et al. Multiple high-throughput analyses monitor the response of *E. coli* to perturbations. *Science*, 2007, 316(5824): 593–597.
- [30] Kanehisa M. A database for post-genome analysis. *Trends Genet*, 1997, 13(9): 375–376.
- [31] Kanehisa M, Goto S, Hattori M, et al. From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Res*, 2006, 34(Suppl 1): D354–D357.
- [32] Pál C, Papp B, Lercher MJ, et al. Chance and necessity in the evolution of minimal metabolic networks. *Nature*, 2006, 440(7084): 667–670.
- [33] Komatsu M, Uchiyama T, Ōmura S, et al. Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism. *Proc Natl Acad Sci USA*, 2010, 107(6): 2646–2651.
- [34] Itaya M, Tsuge K, Koizumi M, et al. Combining two genomes in one cell: stable cloning of the *Synechocystis* PCC6803 genome in the *Bacillus subtilis* 168 genome. *Proc. Natl Acad Sci USA*, 2005, 102(44): 15971–15976.
- [35] Itaya M, Fujita K, Kuroki A, et al. Bottom-up genome assembly using the *Bacillus subtilis* genome vector. *Nat Methods*, 2007, 5(1): 41–43.