

## 逆向代谢工程的最新研究进展

李桂莹<sup>1,2,3\*</sup>, 张新波<sup>4,5\*</sup>, 王智文<sup>1,2,3</sup>, 石婷<sup>1,2,3</sup>, 陈涛<sup>1,2,3</sup>, 赵学明<sup>1,2,3</sup>

1 天津大学化工学院, 天津 300072

2 系统生物工程教育部重点实验室, 天津 300072

3 天津化学化工协同创新中心, 天津 300072

4 天津城建大学环境与市政工程学院, 天津 300384

5 天津市水质科学与技术重点实验室, 天津 300384

李桂莹, 张新波, 王智文, 等. 逆向代谢工程的最新研究进展. 生物工程学报, 2014, 30(8): 1151-1163.

Li GY, Zhang XB, Wang ZW, et al. Progress in inverse metabolic engineering. Chin J Biotech, 2014, 30(8): 1151-1163.

**摘要:** 随着高通量测序技术的快速发展, 可以快速地通过基因组尺度的序列比较, 来全面、系统地探讨工业生产菌株和实验室优良菌株的遗传基础, 分析基因型-表型之间的关系; 同时结合准确的基因组修饰技术, 完成基于全基因组突变分析的逆向代谢工程。近年来, 基于全基因组突变分析的逆向代谢工程已经成功地应用于微生物优良菌株选育, 成为了研究的前沿和热点。综述了近几年逆向代谢工程发展的研究方法, 突变型-表型相关性分析, 逆向代谢工程的最新进展及其应用, 并讨论其面临的问题及挑战。

**关键词:** 逆向代谢工程, 基因型, 表型, 重构

**Received:** November 25, 2013; **Accepted:** January 14, 2014

**Supported by:** National Basic Research Program of China (973 Program) (Nos. 2011CBA00804, 2012CB725203), National Natural Science Foundation of China (Nos. 21206112, 21176182, 51308373), National High Technology Research and Development Program of China (863 Program) (Nos. 2012AA022103, 2012AA02A702), the Innovation Foundation of Tianjin University (No. 1308).

**Corresponding author:** Zhiwen Wang. Tel/Fax: +86-22-27406582; E-mail: zww@tju.edu.cn

\* These authors contributed equally to this study.

国家重点基础研究发展计划 (973 计划) (Nos. 2011CBA00804, 2012CB725203), 国家自然科学基金 (Nos. 21206112, 21176182, 51308373), 国家高技术研究发展计划 (863 计划) (Nos. 2012AA022103, 2012AA02A702), 天津大学自主创新基金 (No. 1308) 资助。

网络出版时间: 2014-03-25

网络出版地址: <http://www.cnki.net/kcms/doi/10.13345/j.cjb.130597.html>

## Progress in inverse metabolic engineering

Guiying Li<sup>1,2,3\*</sup>, Xinbo Zhang<sup>4,5\*</sup>, Zhiwen Wang<sup>1,2,3</sup>, Ting Shi<sup>1,2,3</sup>, Tao Chen<sup>1,2,3</sup>,  
and Xueming Zhao<sup>1,2,3</sup>

1 School of Chemical Engineering & Technology, Tianjin University, Tianjin 300072, China

2 Key Laboratory of Systems Bioengineering, Ministry of Education, Tianjin 300072, China

3 Collaborative Innovation Center of Chemical Science and Engineering (Tianjin), Tianjin 300072, China

4 School of Environmental and Municipal Engineering, Tianjin Chengjian University, Tianjin 300384, China

5 Tianjin Key Laboratory of Aquatic Science and Technology, Tianjin 300384, China

**Abstract:** In the last few years, high-throughput (or 'next-generation') sequencing technologies have delivered a step change in our ability to sequence genomes, whether human or bacterial. Further comparative genome analysis enables us to reveal detailed knowledge of genetics or physiology of industrial important strains obtained in laboratory, to analyze genotype-phenotype correlations of mutants with improved performance. Based on identified key mutations or mutation combinations, Inverse Metabolic Engineering (IME) can be performed by using accurate genetic modification system. Recently, IME has been successfully used for strain improvement and has become a research hotspot, including improving substrate utilization, engineering the robustness of industrial microbes and enhancing production of bio-based products. Here, we describe recent advances in research methods of IME, with an emphasis on characterization of genotype-phenotype and the latest advances and application of IME. Possible directions and challenges for further development of IME are also discussed.

**Keywords:** inverse metabolic engineering, genotype, phenotype, reconstruction

早在1996年, Baily首次对逆向代谢工程(Inverse metabolic engineering, IME)的概念及其应用进行了详细地阐述<sup>[1]</sup>,认为逆向代谢工程是一种有效的菌种改良策略,主要包括以下步骤:首先,鉴定、构建或设计所需要的目标表型;其次,分析目标表型的遗传基础或环境影响因子;最后,通过直接地遗传修饰或环境条件的改变,将目标表型赋予其他菌株或生物体。经过近20年的发展,逆向代谢工程已在遗传机理分析与胁迫抗性突变株的构建<sup>[2-5]</sup>、底物利用<sup>[6-9]</sup>、生物基产品高产菌株构建<sup>[10-13]</sup>等研究方面取得了巨大进展。

近几年,随着以基因组测序为代表的大规模数据产出和分析技术的发展<sup>[14-16]</sup>,传统的生

物学研究方式正在发生改变,极大加快了微生物遗传学<sup>[17]</sup>、突变轮廓分析(Mutational profiling)和功能基因组学的研究进程<sup>[18]</sup>,促进代谢工程研究模式的转变<sup>[19]</sup>。高通量测序技术能够快速便捷地获取突变菌株基因组和转录组信息,使基于基因组尺度的序列比较来全面、系统地分析工业生产菌株和实验室优良菌株的遗传基础成为可能。通过工业菌种或优良表型突变株比较基因组学,转录组学、蛋白质组学和代谢物组学等系统生物学方法的联合使用,系统地阐明基因型-表型关系,揭示优良表型的遗传基础,结合准确的基因组修饰技术,指导菌种代谢工程改进<sup>[18,20]</sup>,使基于全基因组突变分析的逆向代谢工程成为了研究的前沿和热点,使逆向代

谢工程进入了新的发展阶段。

作为逆向代谢工程的第一个环节, 鉴定或构建优良表型菌株方面, 已有很多详细的综述<sup>[21-27]</sup>介绍。本文将重点介绍逆向代谢工程研究方法, 突变型-表型相关性分析, 逆向代谢工程的最新研究进展, 并讨论其面临的问题及挑战。

## 1 逆向代谢工程的原理与方法

目前, 许多重要生物基化学品的工业生产菌株和实验室构建的优良菌株, 是通过多轮理化诱变筛选得到的。由于随机诱变引入突变点的不确定性和随机性, 这些菌株往往具有比较复杂的遗传背景, 使得通过代谢工程进一步合理改进这些突变菌株面临着较大的挑战和困难。例如, 由于随机诱变引入突变点的不确定性, 使突变株与野生株的代谢途径存在很大的差异, 基因组尺度的代谢途径分析与*in silico*代谢网络模型模拟的结果(往往基于野生株的代谢网络)可能并不适用于突变株<sup>[28]</sup>。同时, 大量未知突变的存在也使得对进一步理性修饰的解释和评估变得更为复杂; 另外, 由于诱变的随机性, 这些突变菌株在逐渐积累和目标表型相关有利突变的过程中, 也不可避免积累不利突变以及与目标表型无关的中性突变。通常情况下, 积累的非有利突变远多于有利突变<sup>[29]</sup>。同野生菌株相比, 突变株大都在菌体生长速度、底物利用范围、环境耐受性或遗传稳定性等方面存在缺陷<sup>[30]</sup>。由于和这些不良性状相关的基因型通常非常复杂, 通过代谢工程研究手段消

除这些不良性状面临着巨大困难。由此, 如何全面、系统地确定优良突变株基因型-表型关系, 探索优良表型的遗传调控机理, 消除随机诱变引起不良性状的遗传因素, 是菌种改良面临的关键问题。

基于全基因组突变分析的逆向代谢工程是解决上述问题的有效策略(图 1)。首先, 利用高通量基因组测序技术完成野生型菌株和优良突变株的基因组测序或重测序, 比较基因组学获得突变信息; 或者通过文库构建方法筛选到目的表型菌株, 对特定片段进行测序, 通过序列比对获得全部突变点; 其次, 结合生物化学和生理学知识, 或利用各种“组学”技术(转录组学、蛋白质组学、代谢物组学和代谢通量组学等), 对突变进行相关性分类, 确定潜在重要突变, 同时把突变引入到野生型菌株中验证突变型-表型之间的关系; 最后, 有利突变引入到野生型菌株中, 摒弃不利突变, 迭代循环操作最终构建只含有利突变的重构菌株, 重构菌株可以进一步通过理性代谢工程改造而改进细胞的性能。

重构的菌株遗传背景清晰, 只携带与目标表型相关的有利突变, 保留野生型菌株的生长良好、遗传稳定等优良表型, 可以通过理性代谢工程进一步提高菌种性能。基于全基因组突变分析的逆向代谢工程, 能够从整个代谢网络角度分析突变与表型的关系, 发现有利突变或突变组合; 新发现的优良突变还可以应用到工业菌种的改进, 以及其他相关生物基产品的代谢工程研究。

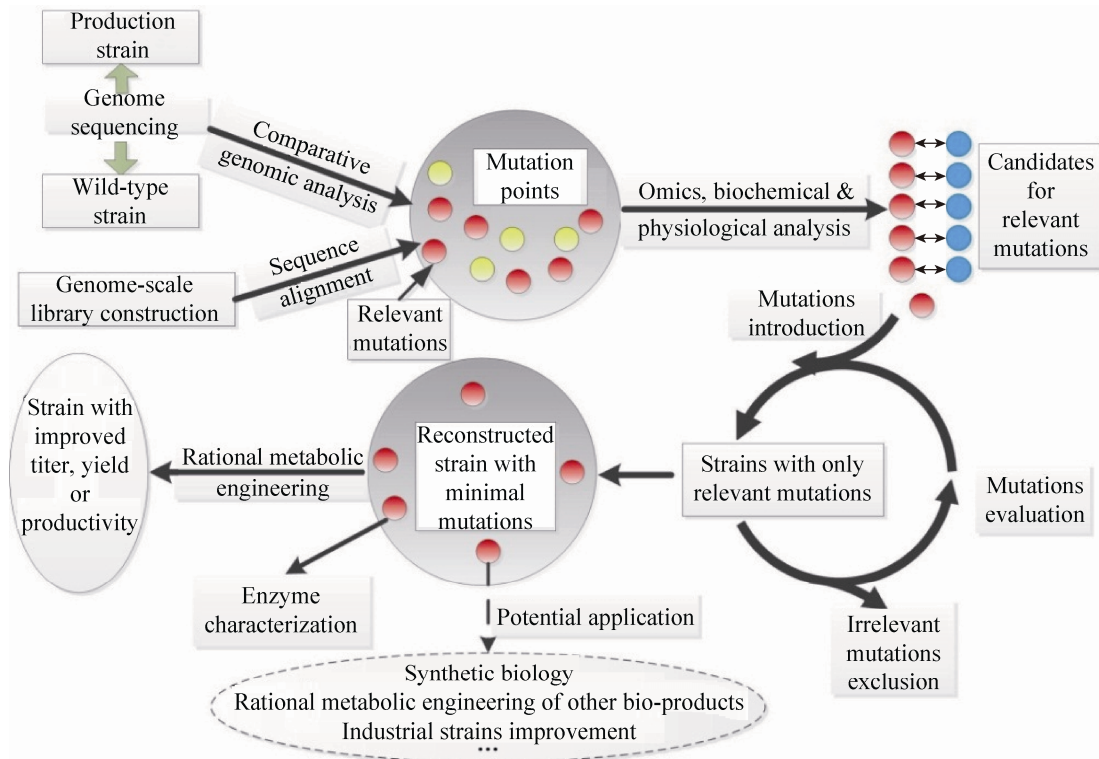


图 1 逆向代谢工程重构菌株流程图

Fig. 1 Workflow of strain reconstruction by inverse metabolic engineering.

## 2 逆向代谢工程的最新研究进展

### 2.1 基因型-表型相关性分析

#### 2.1.1 基于全基因组测序的基因型-表型相关性分析

构建和鉴定特定表型是逆向代谢工程的基础，全基因组范围的基因型-表型相关性的阐释是逆向代谢工程的关键。在基因型的分析方面，高通量基因组测序手段是最为直接和快捷的方式。2008年，Smith等<sup>[31]</sup>分别利用3种新一代测序技术平台（Roche 454 GS FLX，Illumina Solexa和Applied Biosystems SOLiD）对产乙醇突变菌株树干毕赤氏酵母 *Pichia stipitis* Shi21进行了全基因组测序和突变分析。突变分析结

果表明，3种测序平台均准确地鉴定了14个突变。该研究同时证实，运用第二代测序平台测定10–15倍基因组覆盖倍数的数据，可以低成本、快速、准确地完成基因组尺度的突变点分析。随后Darby和Hall<sup>[18]</sup>在 *Nature Biotechnology* 上撰文高度评价了Smith等工作，提出第二代测序技术给科研工作者提供了前所未有的良好技术平台，使基因组尺度的突变分析、转录组分析、表观组学以及DNA-蛋白相互作用等成为常规研究。

随后，基因组尺度遗传基础或机理阐释得到了快速发展。Brown等<sup>[4]</sup>利用Roche 454 GS FLX和基于芯片的比较基因组测序方法测定了高乙醇耐受性（80 g/L）进化热纤梭菌

*Clostridium thermocellum* 突变株和野生型菌株的全基因组信息。测序发现 230 个共有突变, 并分析突变在基因组上的分布规律。作者发现在乙醇产生途径中, 双功能酶乙醛-CoA/乙醇脱氢酶(*AdhE*)存在两个非同义突变(Pro-704-Leu 和 His-734-Arg), 根据同源建模得到的突变酶 AdhE\*模型, Pro-704-Leu 位于  $\alpha$ -螺旋的表面, 与活性位点和辅因子结合位点较远, 因此推测该突变对辅因子的结合或催化活性影响较小; His-734-Arg 突变与催化活性中心和辅因子结合位点空间距离很近, 而精氨酸替代组氨酸对水溶性分子交互作用和中等极性氢键变化的影响, 对于改变辅因子的特异性是足够的, 因此推测突变酶改变了辅因子的特异性而引起该酶发挥催化功能时结合的辅酶发生改变 (NADH 替换为 NADPH), 辅酶的改变而造成了胞内 NADP<sup>+</sup>/NADPH 比例变化, 影响了其与细胞膜的相互作用, 最终导致高乙醇耐受性。

Yang 等<sup>[2]</sup>利用基因芯片和 Roche 454 GS FLX 焦磷酸测序相结合的方法, 对耐乙酸产乙醇运动发酵单胞菌 *Zymomonas mobilis* 突变株和亲本菌株 *Zymomonas mobilis* ZM4 进行了基因组重测序和比较基因组分析, 探讨影响突变株乙酸耐受性的基因或调控机理。比较基因组分析结果发现, 突变株中 *nhaA* 基因 (编码 Na<sup>+</sup>/H<sup>+</sup> 逆向转运体) 上游序列有 1.5 kb 左右的片段缺失, *ZMO1219* (编码推测性的 K<sup>+</sup>转运蛋白) 和 *ZMO1184* (编码假设性蛋白) 各有一个单核苷酸突变。转录组分析表明, *nhaA* 基因上游大片段的缺失导致基因表达强度提高 16 倍之多。作者根据突变分析的结果, 在野生型菌株中进行逆向代谢工程操作, 敲除了 *nhaA* 基因上游 1 461 bp 片段后, 工程菌株表现出耐乙酸 (NaAc

195 mmol/L, pH 5.0) 的生长特性。

国内也有很多的课题组开展了相关的工作, 中国科学院上海生命科学院的姜卫红和杨晟等<sup>[32]</sup>, 采用相似的策略, 利用 Roche 454 焦磷酸测序对高产丁醇而不产孢子的丙酮丁醇梭菌 *Clostridium acetobutylicum* EA 2018 进行了全基因组测序和转录组分析。比较基因组分析表明, EA 2018 中含有与丁醇高产相关潜在重要突变 (46 个缺失突变和 26 个插入突变)。转录组分析结果表明, *spo0A* 和 *adhE2* 的表达水平显著提高, 而大多数酸的合成途径相关基因表现为较低的表达水平。此外, 作者还发现了一个推断性的转录调控因子 CEA\_G2622 可能与木糖的快速利用相关。高产丁醇突变株的遗传基础阐释以及基因型-表型关系分析, 对于解释突变株不形成孢子、高溶剂产生能力和木糖的快速利用等方面奠定了基础, 也为进一步遗传修饰获得高产菌株提供依据。

全基因组突变分析在微生物产生次级代谢物的基因型-表型相关性分析方面也有重要的应用。红霉素多胞菌是一种革兰氏阳性丝状细菌, 工业上主要用于生产红霉素 A。红霉素多胞菌 *Saccharopolyspora erythraea* E3 是野生型菌株 *S. erythraea* NRRL23338 经过随机诱变后筛选得到的红霉素 A 高产菌株。Li 等<sup>[33]</sup>完成了高产菌株 E3 和野生菌 NRRL23338 的全基因组测序和转录组表达分析。比较基因组结果表明, 突变株 E3 中含有 60 个插入突变、46 个缺失突变和 584 个单核苷酸突变。不同时间点的转录组分析表明, 红霉素 A 合成途径和前体物合成途径的基因表达水平显著上调。而先前鉴定的 *ery* 簇调控因子 BldD 和 *ery* 簇基因的表达没有显著相关性, 推测 *S. erythraea* E3 中红霉素 A 的合成存在

其他调控机制。红霉素 A 工业化生产菌株与野生型菌株的功能比较基因组学研究,对于从整体上理解 *S. erythraea* 红霉素 A 的合成机制,进一步利用逆向代谢工程修饰提供重要的依据。

### 2.1.2 基于突变文库的基因型-表型相关性分析

相对于基因组测序获得突变型-表型之间的关系,通过文库构建的方法获得优良突变型,进一步对突变型所携带的文库测序,鉴定突变型-表型之间关系,成本更低,而且能更快地发现关键突变型-表型之间的关系。

Gill 课题组在基因组编辑技术和逆向代谢工程方面作出了突出的成绩。早在 2003 年,Gill 就发表了利用基因组学信息进行逆向代谢工程的综述<sup>[34]</sup>。他们开发的可追踪多元重组工程(Trackable multiplex recombineering, TRMR),能同时对基因组上千个位点进行分析和修饰<sup>[35]</sup>,是典型的基于突变文库构建研究基因型-表型相关性的案例。他们通过合成的 DNA 盒子,扰动大肠杆菌的启动子和核糖体结合位点(Ribosome binding site)序列。每个盒子中包含一个灭瘟素抗性基因(用于重组子的筛选)、分子标签、大肠杆菌中不同基因的同源重组序列以及扰动基因表达的序列,其中扰动基因表达的序列可以是强启动子  $P_{LtetO-1}$  和 RBS,以提高下游基因的转录和翻译水平,或者是通过插入序列置换基因原始的 RBS 序列,以降低下游基因翻译的起始效率。Gill 课题组在 1 周时间内,构建了 8 154 条核苷酸扰动片段,对大肠杆菌的 4 077 个基因分别进行基因表达的上调和下调扰动。片段转化、重组后,抗性平板上筛选并保存转化子。利用 Affymetrix Geneflex TAG4 芯片分析重组前的合成片段和重组后提取的混合细

菌基因组 DNA 中每个分子标签的浓度,结果显示重组后有 8 016 个分子标签的信号,即成功构建了 98% 的大肠杆菌基因表达上调或下调的文库。为了进一步验证 TRMR 方法能够快速鉴定突变型-表型之间的关系,他们将构建的文库菌株涂布到碳源分别为水杨苷、D-果糖、丙酮醛和缬氨酸的基本盐平板上(野生型大肠杆菌不能生长)。结果显示,文库菌株在平板上的菌落个数是自发突变的 100 倍以上。作者采用 TAG4 芯片对所有菌落进行了分析,通过芯片数据可以根据适应度,对不同筛选条件下单个突变进行分类。同时,作者对其中的 83 个单菌落的分子标签进行了测序,确定突变的基因型和适应度之间的关系,与芯片分析的结果一致。TRMR 技术最大的优势是在非常短的时间内,利用带标签的合成 DNA,高效重组获得上千个基因弱表达或强表达文库,之后辅助微阵列技术就能对这些文库进行简单快速的追踪,确定与表型相关的基因及其相关系数,使基因功能研究的通量提高了几个数量级。

2013 年,Woodruff 等<sup>[36]</sup>采用多种组学组合运用的方法鉴定大肠杆菌突变文库中基因型-表型的关系。他们首先采用富集文库的多尺度分析方法(multi-Scalar Analysis of Library Enrichments, SCALES),构建了约 100 万个大肠杆菌文库克隆,该方法建立的文库数量几乎可以在基因组尺度上分析大肠杆菌每一个基因对乙醇耐受性的影响。利用基因组芯片和 SCALES 软件分析增强乙醇耐受性的基因及其适应度,并在野生型大肠杆菌中表达得到的 14 个高适应度基因,验证引起乙醇耐受的基因。结果发现了尚未报道的 9 个基因(*lpcA*、*arnB*、*arnC*、*tilS*、

*yaeJ*、*fadE*、*serA*、*zur* 和 *yicE*)能够显著改善乙醇存在时细胞的生长情况,其主要与细胞膜的组成、翻译、丝氨酸的合成和转录调控相关。为了进一步验证乙醇耐受性基因和分析乙醇耐受机制,作者对乙醇耐受性最好的 5 株克隆进行了转录谱分析,利用同位素标记相对与绝对定量技术(Isobaric Tags for Relative and Absolute Quantitation, iTRAQ)对 2 株乙醇耐受性克隆进行了蛋白质组分析。转录谱数据的分层聚类分析表明,表达谱主要分为两个分支,含有过表达基因 *tisS* 和 *lpcA* 的克隆,其全局表达谱与宿主菌的表达谱类似;而含有过表达基因 *yhfUT*、*fadE* 和 *serA* 的克隆为另一簇。其中,由 *serA* 基因引起的乙醇耐受菌株,表达水平变化的基因数量最多。蛋白质组分析表明,LpcA 和 TisS 蛋白是表达上调程度最高的蛋白,同时 *lpcA* 过表达克隆中有 17 个蛋白上调 22 个蛋白下调,而 *tisS* 过表达克隆中有 22 个蛋白上调 17 个蛋白下调。转录谱和蛋白质组分析表明,某些乙醇耐受菌株以相似的方式调控基因的全局表达,其全局表达谱的变化会有一部分的重合,然而不同乙醇耐受菌株,其在转录水平和翻译水平上,重合的程度不同,且不同组学水平之间的相关系数往往较小。多种组学技术的综合运用,能够有力地鉴定基因型-表型的关系,特别是引起全局表达水平变化或多个基因表达水平变化的突变;此外,对于进一步阐释突变引起表型变化的机制具有重要的意义。

随后,Woodruff 等<sup>[37]</sup>采用类似方法构建和筛选具有乙醇耐受性、生长迅速且高产乙醇的文库菌株,鉴定相关表型的遗传基础,并对已鉴定基因进行组合优化,考察不同组合对目标

表型的影响,发现过表达 *otsA* 基因能显著提高乙醇耐受性和乙醇生产率。Hong 等<sup>[38]</sup>采用基因组文库构建的方法,鉴定了酿酒酵母基因组中引起乙醇耐受性改进的基因修饰位点。在基因组尺度上共鉴定了 4 个基因(*INO1*、*DOG1*、*HAL1* 和 *MSN2*),其中,*INO1* 的过表达不仅能够提高乙醇耐受性,还可以使宿主菌在高浓度糖(10%)和乙醇(5%)的培养基中,比生长速率提高 3 倍。

虽然基于文库构建的方法鉴定基因型-表型关系成本相对较低,能较快地鉴定重要突变,但是突变库的构建相对繁琐,受到高通量筛选方法的限制,且往往不能包含基因组的全突变信息,具有一定局限性。

## 2.2 逆向代谢工程的应用

### 2.2.1 提高底物利用效率

通过逆向代谢工程提高底物的利用效率主要有两种方式:一是通过进化工程,激活潜在的代谢途径,使细胞能够利用新的底物并提高底物的利用效率;或者通过进化工程提高天然底物的利用效率,获得优良突变株,进一步通过基因型-表型关系分析,得到逆向代谢工程的操作靶点,此方面的研究进展主要集中在木糖<sup>[39]</sup>和半乳糖<sup>[40]</sup>等底物利用,以及非天然底物<sup>[41]</sup>或甘油<sup>[42-44]</sup>等重要工业副产物等的利用;其次是在异源微生物或相关模式系统中得到高效利用底物表型,然后通过基因组测序或其他“组学”技术分析其遗传基础,得到关键代谢途径、酶基因,进一步在特定微生物中通过遗传修饰提高底物的利用效率<sup>[9]</sup>。

Palsson 课题组在利用高通量测序技术分析进化菌株的基因型-表型关系,进而指导逆向代谢工程研究方面作了出色工作<sup>[45,41]</sup>。Lee 和

Palsson<sup>[41]</sup>对 *E. coli* K12 MG1655 菌株在以 1,2-丙二醇为唯一碳源的基本盐培养基上进行 700 代适应性进化, 筛选获得的进化菌株不仅能在 1,2-丙二醇为唯一碳源和能源的基本培养基上正常生长, 且比生长速率达到 0.35/h; 进一步通过 Illumina Solexa 全基因组测序和比较基因组分析, 发现进化过程中积累了 6 个突变 (5 个突变位于编码区, 1 个突变位于非编码区), 作者通过基因吞噬 (Gene gorging) 技术和  $\lambda$ -Red 重组系统将突变点引入野生型菌株, 考察每一个突变点对进化表型的贡献度, 最终将全部有利突变引入到野生型菌株中, 构建了一株表型与进化菌株几乎相同的工程菌株。基于基因组测序的突变分析结合有效的基因操作手段, 分析突变型与表型的相关性, 重构只含有利突变的目标表型工程菌株。

将异源微生物的表型赋予特定微生物方面, 通过逆向代谢工程使酿酒酵母代谢木糖是比较成功的案例。酿酒酵母 *Saccharomyces cerevisiae* 不能利用木糖作为碳源, 而树干毕赤酵母 *Pichia stipitis* 天然能够代谢木糖<sup>[46]</sup>, *P. stipitis* 基因组测序阐明了木糖代谢途径<sup>[47]</sup>, 作者将其木糖代谢途径引入至酿酒酵母细胞中, 使其能够利用木糖合成乙醇。然而, 厌氧条件下外源木糖代谢途径辅因子不平衡, 酿酒酵母生长缓慢<sup>[48]</sup>。为进一步理解木糖利用的代谢机制和逆向代谢工程修饰, Wohlbach 等<sup>[9]</sup>最近对两株天然利用木糖的甲虫关联酵母 *Spathaspora passalidarum* 和纤维假丝酵母 *Candida tenuis* 进行基因组测序, 并对现有的 14 株子囊菌基因组进行了比较基因组学分析, 根据真菌的系统发育树进行基因型-表型的比对分析, 同时检测了 5 株具有不同木糖利用能力的

半子囊菌的基因表达情况, 发现了促进木糖利用的潜在基因 (10 个)。在酿酒酵母中分别表达来自 *C. tenuis* 醛酮还原酶 (CtAKR) 和 *S. passalidarum* 未注释蛋白 (SpNA) 能够特异性提高木糖利用效率。好氧或厌氧条件下 CtAKR 的过表达均能够显著提高木糖的利用效率, 其中厌氧条件下, 发酵 72 h 木糖利用率提高了 32%, 同时副产物甘油和乙酸的积累分别降低了 73%和 42%。

### 2.2.2 增强环境胁迫能力

环境胁迫是指不利于细胞生长的外部环境因素, 如抗生素、毒性化学物质、极端温度、pH 或营养限制等。在工业生产过程中, 微生物细胞会面临多种胁迫作用如酸胁迫、高温胁迫、渗透压胁迫等。胁迫诱导的基因调节, 有可能影响细胞的许多重要生理功能, 进而影响生物转化效率, 因此选择生产性能良好、对发酵过程中的主要胁迫因素有较强耐受性的菌株至关重要。付瑞艳和李寅<sup>[25]</sup>对利用代谢工程提高工业微生物胁迫抗性作了系统地论述。工业化生产过程中, 为了追求产品的高得率和高产量, 需要较高的底物和目标产品浓度, 特别是在生产醇类 (如乙醇、丁醇和异丁醇) 以及氨基酸和有机酸 (如琥珀酸、乳酸、乙酸和苹果酸) 等物质, 提高菌株对底物和产物的耐受性尤为重要。强耐受性菌株的发酵过程能够大大简化产品的下游处理过程, 增加了产品浓度, 降低成本<sup>[49]</sup>。

相对于乙醇, 高级醇 (三碳或更高碳链醇) 更适合作为液体燃料, 其具有更高的能量密度, 更低的蒸发压和更低的亲水性。近几年, 1-丙醇、1-丁醇、异丁醇、2-甲基-1-丁醇、3-甲基-1-丁醇等相继通过代谢工程在大肠杆菌中合成出来。值得一提的是, 大肠杆菌异丁醇的发酵产



量已经超过 20 g/L<sup>[11]</sup>。虽然异丁醇的生产在细胞生长停止后仍然继续进行,但是 8 g/L 的异丁醇对细胞生长就有抑制作用。所以,提高菌体对异丁醇的耐受性是提高异丁醇产量的关键环节。

Liao 课题组的工作<sup>[3]</sup>具有很好的代表性,首先他们通过进化工程的方法得到了异丁醇耐受菌株大肠杆菌 SA481,利用 Solexa 测序平台对出发菌株 JCL260 和突变株 SA481 进行全基因组测序,并进行了比较基因组分析,获取进化菌株的突变信息。与 JCL260 相比,SA481 含有 1 个 SNP、25 个插入突变和一段含有 62 个基因的大片段缺失。为了识别异丁醇耐受性关键突变点,作者对 SA481 中的突变进行了单回复突变和不同组合的回复突变,绝大部分单回复突变会降低菌株的异丁醇耐受度,但是没有任何单回复突变可以完全破坏菌株的耐受度,而 *acrA*、*yhbJ* 和大片段缺失的回复可以显著降低异丁醇耐受度。组合回复结果发现,其中的 5 个突变组合回复后获得的突变株 TW313,其生长速度与出发菌株 JCL260 基本相同。作者在出发菌株的基础上,进行了菌株重构,构建了高异丁醇耐受性的突变株 TW306,8 g/L 的异丁醇耐受性检测显示,出发菌株 JCL260 完全停止了生长,SA481 和 TW306 则可以继续生长,同时生产异丁醇,这表明提高异丁醇耐受度可以改善异丁醇胁迫压力下菌体生长和异丁醇生产能力,成功地通过逆向代谢工程方法在大肠杆菌中重构高异丁醇耐受的复杂表型。

### 2.2.3 构建生物基产品高产菌株

逆向代谢工程构建高产菌株方面,Stephanopoulos 课题组构建的产 L-酪氨酸大肠杆

菌重构菌株是非常成功的案例<sup>[12]</sup>。他们首先通过全局转录机器工程 (global Transcription Machinery Engineering, gTME),构建了  $10^5$ – $10^6$  的大肠杆菌文库,从中筛选到了 3 株 L-酪氨酸增产菌株 (*rpoA14*、*rpoA27*和*rpoD3*),其 L-酪氨酸的产量比宿主菌提高了 91%–113%。进一步分析了不同突变质粒与遗传背景之间的关联性,并通过转录组分析了不同基因的表达情况,探讨突变型的遗传基础和 L-酪氨酸的高产机制。通过突变株的全基因组测序和比较基因组学分析,鉴定出 3 株高产菌株中引起 L-酪氨酸高产的突变基因。3 个突变菌株中,每个菌株仅发生了一个单碱基突变(*hisH* L82R、*purF* V5G和*purF* 上游 17 bp 处 T 突变为 C),最后他们将突变整合到宿主菌中,得到遗传背景清晰的工程菌 *rpoA14<sup>R</sup>*。摇瓶发酵检测,L-酪氨酸产量高达 902 mg/L,得率为 0.18 g 酪氨酸/g 葡萄糖,是工业化生产菌株的 1.5 倍,达到理论得率的 85%。2 L 发酵罐中,L-酪氨酸的生产率为 2.1 g/(L·h),36 h 产量达到 13.8 g/L,生长速率为 0.405/h。逆向代谢工程不仅构建了高产菌株,降低了工业化生产成本,而且进一步阐释了 L-酪氨酸合成的遗传机制。

近来,基于全基因组突变分析的逆向代谢工程在三萜烯 $\beta$ -香树精<sup>[50]</sup>、重组蛋白<sup>[13]</sup>和 N 端糖基化蛋白<sup>[51]</sup>高产菌株的构建方面均有成功的报道。我们研究组也正在开展基于高产核黄素突变菌株全基因组突变分析的逆向代谢工程研究,利用 Roche 454 GS FLX 技术对两株核黄素高产菌株进行了全基因组测序和突变分析,两株菌均发现了超过 500 个的单核苷酸突变,通过无痕重组技术已经对超过 30 个突变基因及其组

合进行了表型分析,在野生型菌株中初步构建了重构菌株。

### 3 展望

从近几年的研究可以看出,基于基因组测序的突变分析在连接基因型-表型关系,阐释优良表型的遗传调控机理,改进工业化生产菌株性能方面显示了巨大的应用潜力。由中国科学院微生物研究所牵头的“十二五”863计划“工业微生物基因组及分子改造”,拟从工业微生物的应用基因组科学出发,建立我国重要工业生产菌种的基因组数据库,利用微生物分子改造技术,针对氨基酸、维生素、抗生素、发酵食品等产品,研发新一代工业菌株(<http://www.cncbd.org.cn/>)。

基因组尺度的突变分析,结合有效的基因操作手段,可以分析突变型与表型的相关性,重构只含有利突变的目标表型工程菌株。而且,全基因组突变分析和重要突变点的深入研究,不仅能够阐释突变引起表型变化的机理,还可为解释某些复杂表型的调控机制提供依据,为相关生物基产品的代谢工程提供修饰靶点。基因组尺度的突变分析具有全面、系统的特点,利用全基因组突变分析,阐释生物基产品高产表型的遗传基础或代谢调控机制是重要的发展趋势,逆向代谢工程必将为新一代工业化菌株的改良作出重大贡献。

基于全基因组突变分析的逆向代谢工程面临的挑战之一,是如何从大量的突变信息中“沙里淘金”,得到与表型正相关的突变及其突变组合。特别对于经过诱变、筛选得到的目标表型菌株,可能会含有数百甚至上千个突变。各种“组学”和生物信息学技术是对突变点进行分类,

发现关键突变点,特别是引起全局调控变化的突变的重要工具。此外,随着基因组尺度代谢网络模型的快速发展,结合不同的代谢网络和菌株优化算法,如 MOMA、FBA、OptKnock 和 OptGene 等<sup>[52]</sup>,可以预测某些表型修饰的靶目标基因,将突变信息与预测的靶目标基因组合分析,有助于快速发现有利突变或突变组合。一旦确定了关键(突变)基因之后,下一步的任务是如何调控和优化不同的基因组合和表达水平。笔者认为逆向代谢工程面临的另一挑战,是如何将有利突变信息进行合理地组合,进而得到表型优良的重构菌株。因为代谢网络复杂调控关系,某些复杂表型如生长和产品的高产,往往需要多基因不同表达水平的组合,以及基因“异位显性”等因素的存在,逆向代谢重构过程中需要考虑途径内突变基因的组合,以及不同代谢或调控途径内突变的组合。因此,快速构建不同基因表达水平组合的多重自动基因组编辑技术是未来发展的重要趋势,也是代谢工程和逆向代谢工程的强有力工具。

### REFERENCES

- [1] Bailey JE, Sburlati A, Hatzimanikatis V, et al. Inverse metabolic engineering: a strategy for directed genetic engineering of useful phenotypes. *Biotechnol Bioeng*, 1996, 52(1): 109–121.
- [2] Yang S, Land ML, Klingeman DM, et al. Paradigm for industrial strain improvement identifies sodium acetate tolerance loci in *Zymomonas mobilis* and *Saccharomyces cerevisiae*. *Proc Natl Acad Sci USA*, 2010, 107(23): 10395–10400.
- [3] Atsumi S, Wu TY, Liao JC, et al. Evolution, genomic analysis, and reconstruction of isobutanol tolerance in *Escherichia coli*. *Mol Syst Biol*, 2010, 6: 449–460.

- [4] Brown SD, Guss AM, Karpinets TV, et al. Mutant alcohol dehydrogenase leads to improved ethanol tolerance in *Clostridium thermocellum*. Proc Natl Acad Sci USA, 2011, 108(33): 13752–13757.
- [5] Minty JJ, Lesnefsky AA, Lin F, et al. Evolution combined with genomic study elucidates genetic bases of isobutanol tolerance in *Escherichia coli*. Microb Cell Fact, 2011, 10: 18.
- [6] Le Crom S, Schackwitz W, Pennacchio L, et al. Tracking the roots of cellulase hyperproduction by the fungus *Trichoderma reesei* using massively parallel DNA sequencing. Proc Natl Acad Sci USA, 2009, 106(38): 16151–16156.
- [7] Conrad TM, Joyce AR, Palsson BO, et al. Whole-genome resequencing of *Escherichia coli* K-12 MG1655 undergoing short-term laboratory evolution in lactate minimal media reveals flexible selection of adaptive mutations. Genome Biol, 2009, 10(10): R118.
- [8] Otero JM, Vongsangnak W, Nielsen J, et al. Whole genome sequencing of *Saccharomyces cerevisiae*: from genotype to phenotype for improved metabolic engineering applications. BMC Genomics, 2010, 11: 723–740.
- [9] Wohlbach DJ, Kuo A, Sato TK, et al. Comparative genomics of xylose-fermenting fungi for enhanced biofuel production. Proc Natl Acad Sci USA, 2011, 108(32): 13212–13217.
- [10] Hemme CL, Fields MW, He Q, et al. Correlation of genomic and physiological traits of thermoanaerobacter species with biofuel yields. Appl Environ Microbiol, 2011, 77(22): 7998–8008.
- [11] Smith KM, Liao JC. An evolutionary strategy for isobutanol production strain development in *Escherichia coli*. Metab Eng, 2011, 13(6): 674–681.
- [12] Santos CN, Xiao W, Stephanopoulos G. Rational, combinatorial, and genomic approaches for engineering L-tyrosine production in *Escherichia coli*. Proc Natl Acad Sci USA, 2012, 34: 13538–13543.
- [13] Ghosh C, Gupta R, Mukherjee KJ. An inverse metabolic engineering approach for the design of an improved host platform for over-expression of recombinant proteins in *Escherichia coli*. Microb Cell Fact, 2012, 11: 93.
- [14] Shendure J, Ji H. Next-generation DNA sequencing. Nat Biotechnol, 2008, 26(10): 1135–1145.
- [15] Metzker ML. Sequencing technologies-the next generation. Nat Rev Genet, 2010, 11(1): 31–46.
- [16] Zhou XG, Ren LF, Li YT, et al. Next-generation sequencing technology: technology review and prospect. Sci China Life Sci, 2010(1): 23–37 (in Chinese).  
周晓光, 任鲁风, 李运涛, 等. 下一代测序技术: 技术回顾与展望. 中国科学: C 辑, 2010(1): 23–37.
- [17] MacLean D, Jones JD, Studholme DJ. Application of 'next-generation' sequencing technologies to microbial genetics. Nat Rev Microbiol, 2009, 7(4): 287–296.
- [18] Darby AC, Hall N. Fast forward genetics. Nat Biotechnol, 2008, 26(11): 1248–1249.
- [19] Yadav VG, De Mey M, Lim CG, et al. The future of metabolic engineering and synthetic biology: towards a systematic practice. Metab Eng, 2012, 14(3): 233–241.
- [20] Oud B, van Maris AJ, Daran JM, et al. Genome-wide analytical approaches for reverse metabolic engineering of industrially relevant phenotypes in yeast. FEMS Yeast Res, 2012, 12(2): 183–196.
- [21] Zhang YX, Perry K, Vinci VA, et al. Genome shuffling leads to rapid phenotypic improvement in bacteria. Nature, 2002, 415(6872): 644–646.
- [22] Stephanopoulos G, Alper H, Moxley J. Exploiting biological complexity for strain improvement through systems biology. Nat Biotechnol, 2004, 22(10): 1261–1267.
- [23] Patnaik R. Engineering complex phenotypes in industrial strains. Biotechnol Progr, 2008, 24(1): 38–47.
- [24] Gong J, Zheng H, Wu Z, et al. Genome shuffling:

- progress and applications for phenotype improvement. *Biotechnol Adv*, 2009, 27(6): 996–1005.
- [25] Fu RY, Li Y. Improving industrial microbial stress resistance by metabolic engineering: a review. *Chin J Biotech*, 2010, 26(9): 1209–1217 (in Chinese).  
付瑞艳, 李寅. 利用代谢工程技术提高工业微生物对胁迫的抗性. *生物工程学报*, 2010, 26(9): 1209–1217.
- [26] Zhu L, Zhu Y, Zhang Y, et al. Engineering the robustness of industrial microbes through synthetic biology. *Trends Microbiol*, 2012, 20(2): 94–101.
- [27] Lo TM, Teo WS, Ling H, et al. Microbial engineering strategies to improve cell viability for biochemical production. *Biotechnol Adv*, 2013, 31(6): 903–914.
- [28] Conrad TM, Lewis NE, Palsson BO. Microbial laboratory evolution in the era of genome-scale science. *Mol Syst Biol*, 2011, 7: 509–570.
- [29] Warner JR, Patnaik R, Gill RT. Genomics enabled approaches in strain engineering. *Curr Opin Microbiol*, 2009, 12(3): 223–230.
- [30] Demain AL. Microbial biotechnology. *Trends Biotechnol*, 2000, 18(1): 26–31.
- [31] Smith DR, Quinlan AR, Peckham HE, et al. Rapid whole-genome mutational profiling using next-generation sequencing technologies. *Genome Res*, 2008, 18(10): 1638–1642.
- [32] Hu S, Zheng H, Gu Y, et al. Comparative genomic and transcriptomic analysis revealed genetic characteristics related to solvent formation and xylose utilization in *Clostridium acetobutylicum* EA 2018. *BMC Genomics*, 2011, 12: 93.
- [33] Li YY, Chang X, Yu WB, et al. Systems perspectives on erythromycin biosynthesis by comparative genomic and transcriptomic analyses of *S. erythraea* E3 and NRRL23338 strains. *BMC Genomics*, 2013, 14(1): 523–536.
- [34] Gill RT. Enabling inverse metabolic engineering through genomics. *Curr Opin Biotechnol*, 2003, 14(5): 484–490.
- [35] 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): 856–862.
- [36] Woodruff LBA, Pandhal J, Ow SY, et al. Genome-scale identification and characterization of ethanol tolerance genes in *Escherichia coli*. *Metab Eng*, 2013, 15(1): 124–133.
- [37] Woodruff LBA, Boyle NR, Gill RT. Engineering improved ethanol production in *Escherichia coli* with a genome-wide approach. *Metab Eng*, 2013, 17(1): 1–11.
- [38] Hong ME, Lee KS, Yu BJ, et al. Identification of gene targets eliciting improved alcohol tolerance in *Saccharomyces cerevisiae* through inverse metabolic engineering. *J Biotechnol*, 2010, 149(1/2): 52–59.
- [39] Kim SR, Skerker JM, Kang W, et al. Rational and evolutionary engineering approaches uncover a small set of genetic changes efficient for rapid xylose fermentation in *Saccharomyces cerevisiae*. *PLoS ONE*, 2013, 8(2): e57048.
- [40] Hong KK, Vongsangnak W, Vemuri GN, et al. Unravelling evolutionary strategies of yeast for improving galactose utilization through integrated systems level analysis. *Proc Natl Acad Sci USA*, 2011, 108(29): 12179–12184.
- [41] Lee DH, Palsson BO. Adaptive evolution of *Escherichia coli* K-12 MG1655 during growth on a nonnative carbon source, L-1,2-propanediol. *Appl Environ Microbiol*, 2010, 76(13): 4158–4168.
- [42] Roca C, Albuquerque G, Reis M. Evolutionary engineering of *Actinobacillus succinogenes* for improved succinic acid production on glycerol. *J Biotechnol*, 2010, 150: 373–373.
- [43] Applebee MK, Joyce AR, Conrad TM, et al. Functional and metabolic effects of adaptive glycerol kinase (GLPK) mutants in *Escherichia coli*. *J Biol Chem*, 2011, 286(26): 23150–23159.
- [44] Merico A, Ragni E, Galafassi S, et al. Generation of an evolved *Saccharomyces cerevisiae* strain with a high freeze tolerance and an improved ability to grow on glycerol. *J Ind Microbiol Biotechnol*,

- 2011, 38(8): 1037–1344.
- [45] Herring CD, Raghunathan A, Honisch C, et al. Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. *Nat Genet*, 2006, 38(12): 1406–1412.
- [46] Chung BKS, Lakshmanan M, Klement M, et al. Metabolic reconstruction and flux analysis of industrial *Pichia* yeasts. *Appl Microbiol Biotechnol*, 2013, 97(5): 1865–1873.
- [47] Jeffries TW, Grigoriev IV, Grimwood J, et al. Genome sequence of the lignocellulose-bioconverting and xylose-fermenting yeast *Pichia stipitis*. *Nat Biotechnol*, 2007, 25(3): 319–326.
- [48] Van Vleet JH, Jeffries TW. Yeast metabolic engineering for hemicellulosic ethanol production. *Curr Opin Biotechnol*, 2009, 20(3): 300–306.
- [49] Portnoy VA, Bezdán D, Zengler K. Adaptive laboratory evolution-harnessing the power of biology for metabolic engineering. *Curr Opin Biotechnol*, 2011, 22(4): 590–594.
- [50] Madsen KM, Udatha GD, Semba S, et al. Linking genotype and phenotype of *Saccharomyces cerevisiae* strains reveals metabolic engineering targets and leads to triterpene hyper-producers. *PLoS ONE*, 2011, 6(3): e14763.
- [51] Pandhal J, Woodruff LBA, Jaffe S, et al. Inverse metabolic engineering to improve *Escherichia coli* as an N-glycosylation host. *Biotechnol Bioeng*, 2013, 110(9): 2482–2493.
- [52] Zomorodi AR, Suthers PF, Ranganathan S, et al. Mathematical optimization applications in metabolic networks. *Metab Eng*, 2012, 14(6): 672–686.

(本文责编 陈宏宇)