

转录组平台技术及其在代谢工程中的应用

史硕博^{1,2}, 陈涛^{1,2}, 赵学明^{1,2}

1 天津大学 系统生物工程教育部重点实验室, 天津 300072

2 爱丁堡大学-天津大学系统生物学与合成生物学联合研究中心, 天津 300072

摘要: 组学技术在系统水平上对细胞代谢进行全面的分析, 极大地促进了代谢工程的发展和應用。全基因组水平的转录分析可以使研究者更加精确地评估细胞表型, 加深对细胞代谢的理解。而且转录组分析也有助于研究者鉴定菌种改良的目标基因, 加速对微生物细胞工厂的合理设计及构建。文中介绍了 3 种主要转录组平台技术的原理, 并总结了转录组学在代谢工程领域中应用的最新进展和未来发展趋势。

关键词: 转录组, 代谢工程, 基因芯片, serial analysis of gene expression (SAGE), massively parallel signature sequencing (MPSS), RNA-seq

Transcriptome platforms and applications to metabolic engineering

Shuobo Shi^{1,2}, Tao Chen^{1,2}, and Xueming Zhao^{1,2}

1 Key Laboratory of Systems Bioengineering, Ministry of Education, Tianjin University, Tianjin 300072, China

2 Edinburg-Tianjin Joint Research Centre for Systems Biology and Synthetic Biology, Tianjin 300072, China

Abstract: Omics technologies have profoundly promoted development and applications of metabolic engineering by analysis of cell metabolism at a system level. Whole genome transcription profiles have provided researchers more rigorous evaluation of cell phenotype and an increased understanding of cellular metabolism. Furthermore, transcriptome analysis can conduce to identification of effective gene targets for strain improvement, and consequently accelerates rational design and construction of microbial cell factories for desired product. In this review, we briefly introduced the principle of three main platforms of transcriptome, and reviewed the recent applications of the transcriptome to metabolic engineering, finally provided conclusions and future prospects.

Keywords: transcriptome, metabolic engineering, microarray, SAGE, MPSS, RNA-seq

近年来, 包括基因组、转录组、蛋白质组、代谢组和代谢通量组的各种组学技术在揭示细胞生理活动规律的研究中起着越来越重要的作用^[1-2]。通过组学技术, 可以系统评价细胞表型, 并根据评价结

果进一步设计、修饰和重构细胞, 提升代谢工程的合理性和有效性^[3-4], 这就是在后基因组时代所强调采取组学技术指导的代谢工程研究方法, 称之为“系统代谢工程”^[5]。

Received: April 30, 2010; **Accepted:** July 22, 2010

Supported by: National Science Foundation of China (Nos. 20806055, 20875068), National Basic Research Program of China (973 Program) (No. 2007CB707802).

Corresponding author: Tao Chen. Tel: +86-22-27406770; E-mail: chentao@tju.edu.cn

国家自然科学基金项目 (Nos. 20806055, 20875068), 国家重点基础研究发展计划 (973 计划) (No. 2007CB707802) 资助。

在各种组学技术中,转录组学是率先发展起来以及应用最广泛的技术^[6]。细胞的功能是从基因的表达开始的,转录组是指某一时间细胞内所有基因转录而来的 RNA 总称。通过分析转录组,可高通量地获得基因表达的 RNA 水平有关信息,可以揭示基因表达与一些生命现象之间的内在联系。据此我们可以高通量表征细胞生理活动规律,确定细胞代谢特性,并进而对细胞进行修饰改造^[7-8]。本文介绍了转录组学主要技术平台及其原理,总结和评述了转录组学在代谢工程中的应用及进展。

1 转录组技术平台

为了高通量并行分析基因的表达情况,各种转录组技术平台得到迅速发展,主要包括:基因芯片技术 (Microarray)、基因表达系列分析技术 (Serial analysis of gene expression, SAGE)、大规模平行测序技术 (Massively parallel signature sequencing, MPSS) 以及最新提出的 RNA 测序技术 (RNA sequencing, RNA-seq), 这些技术的工作流程如图 1 所示。

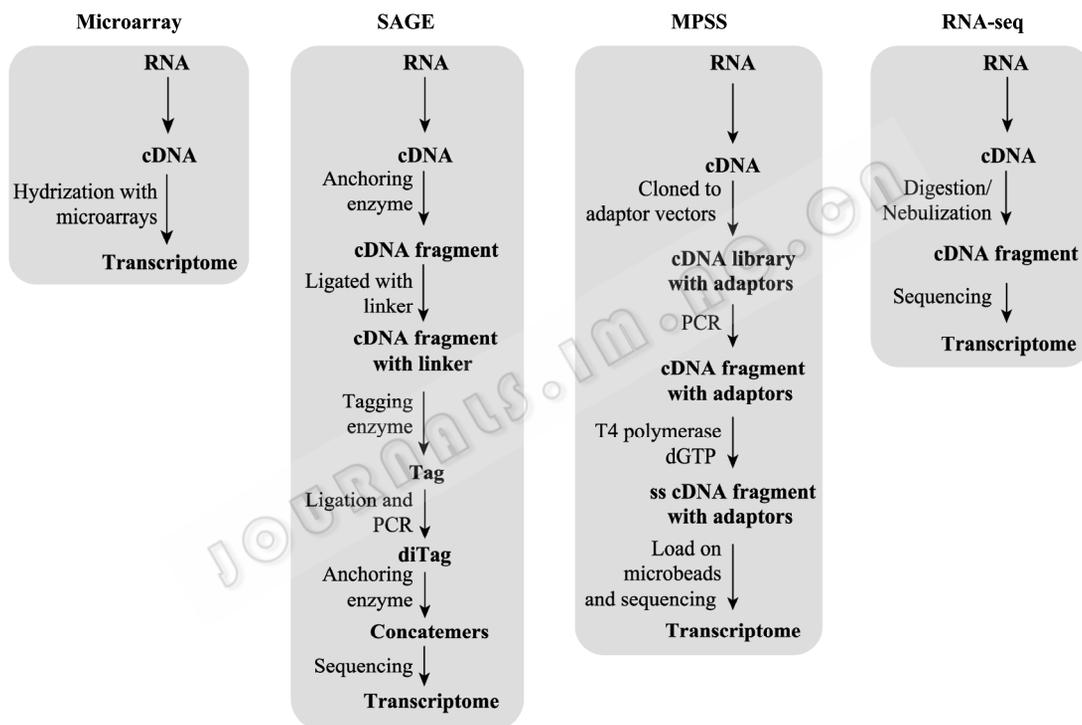


图 1 主要转录组技术平台工作流程示意图

Fig. 1 Outline protocols of the main transcriptome platforms.

1.1 基因芯片技术 (Microarray)

在转录组研究中应用最早及最广泛的为基因芯片技术,首先从待检测样品中提取 RNA,并利用荧光标记的核苷酸将其反转录成 cDNA,经过标记的核苷酸序列可与基因芯片特定位点上的探针杂交,经检测杂交信号而获取细胞基因表达信息。基因芯片技术已成为一项非常稳定可信的实验技术,目前公布的大量转录组数据主要是利用基因芯片技术产生的。

为了促进基因芯片数据的交流和比较,2001 年

基因芯片表达数据协会 (Microarray gene expression data society) 提出了基因芯片试验的标准方案,即 MIAME 方案 (Minimal information about a microarray experiment)^[9];由美国食品药品监督管理局 (US food and drug administration, FDA) 领导实施的基因芯片质量控制联盟 (MAQC) 发表的报告显示在不同实验室之间基因芯片技术平台内部的数据具有高度的一致性和重现性^[10]。基因芯片是最早开发出来的高通量转录组检测技术,且成本适中,其对较高表达的基因检测比较准确,数据分析软件较多,整个方法

较为成熟, 主要缺点是对低表达基因检测敏感度不够, 前期工作基础要求较高, 而且很难检测出融合基因转录、多顺反子转录等异常转录产物。

1.2 基因表达系列分析技术 (SAGE) 与大规模平行测序技术 (MPSS)

SAGE 技术是一种基于测序技术、开放式的、快速高效的分析细胞基因表达状态的方法, 该技术不需任何基因序列的信息, 能够全局性地检测所有基因的表达水平, 除了具有显示基因差异表达谱的作用外, 还对于那些未知基因特别是那些低拷贝基因的发现起到了巨大的推动作用^[11]。SAGE 技术首先从待检测样品中提取 RNA, 并用生物素荧光标记的核苷酸将其反转录成 cDNA, 随后用一种被称为锚定酶的限制性内切酶 (Anchoring enzyme) 切割双链 cDNA, 将回收得到的 cDNA 片段与不同的接头连接, 再用标签酶酶切处理后得到 SAGE 标签, 连接 SAGE 标签形成标签二聚体并进行 PCR 扩增, 最后锚定酶切除接头序列以形成标签二聚体的多聚体, 对其测序可得转录组。

MPSS 技术是对 SAGE 技术的改进, 但其原理都是基于短标签测序 (Tag-based sequencing) 的方法。MPSS 技术可获得更长的短标签, 因而精度更高; 此外, MPSS 技术特有的微球荧光测序可直接高通量读出序列, 简化了测序过程^[12]。MPSS 技术首先从待检测样品提取 RNA 并反转录为 cDNA, 克隆至带有不同 adaptor 的载体文库中, 随后 PCR 扩增带有不同 adaptor 的 cDNA 片段, 在 T4 DNA 聚合酶和 dGTP 的作用下使其转换为单链文库, 最后通过杂交将其结合在带有 Anti-adaptor 的微载体上进行测序。

SAGE 技术和 MPSS 技术都需要大量的测序工作, 但高通量基因测序仪的迅速发展缓解了这一问题, 有力推动了它们的推广与应用。但是这些技术的难度较大, 而且涉及酶切、PCR 扩增、克隆等可能会产生碱基偏向性的操作步骤, 从而影响转录本的正确识别^[12]。Illumina 公司于 2007 年在 MPSS 技术的基础上推出新一代测序仪 Illumina/Solexa Genome Analyzer, 正迅速得到普及。

1.3 RNA 测序技术 (RNA-seq)

新一代高通量基因组测序仪的迅速发展 (Solexa, 454 GS-FLX, SOLiD, tSMS) 不仅给基因组领域带来革命性的突破, 同时也给转录组检测方法带来重大革新。采用类似 SAGE 技术和 MPSS 技术的理念, 新一代高通量基因组测序仪可以通过测定细胞全部转录产物序列, 通过序列比对得到最后的转录组, 一个新的测定转录组的“RNA 测序”法出现了, 该技术称为 RNA 测序技术 (RNA-seq)^[13-14]。

RNA 测序技术的原理与 SAGE 技术和 MPSS 技术一致, 即对细胞转录产物进行测序, 统计测得的每条序列获得每个特定转录本的表达量, 可提供精确的数字化表达谱检测。该技术首先将细胞中的所有转录产物反转录为 cDNA 文库, 然后将 cDNA 文库中的 DNA 随机剪切为小片段(片段大小根据采取的测序方法的读长而不同), 利用新一代高通量测序仪测序直到获得足够的序列。高通量测序避免了亚克隆过程中引入的偏差, 且其确定的短序列长度显著增加。获得长度显著增加的短序列所包含信息可以提高识别其对应的基因的准确性, 然后计算这些短序列的个数和分析其在整个基因组中的分布, 可以计算出细胞的转录组表达水平^[14]。

RNA-seq 技术可以提供更多信息, 可以得到用基因芯片难以得到的转录可变剪接序列^[14]; 该技术对低表达基因的检测更加准确, 并且可以定量确定转录水平。核苷酸多态性、GC 偏向会影响 SAGE 和 MPSS 测定转录组时的精确性, 而 RNA-seq 技术克服了这些不足。RNA-seq 技术得到的标签长度(由采取的高通量测序技术的读长而定, 可达 400 bp)比 SAGE 和 MPSS 显著增加, 提高了识别转录本的特异性和准确性^[15]。

RNA 测序技术目前发展很快, 在转录组研究中迅速得到广泛应用, 受到高度好评。由于 RNA 测序技术产生海量的数据, 面临着一系列新的信息学方面的难题, 其中包括如何最好地诠释和比对鉴定多个类似的同源基因; 如何确定最佳测序量, 获得高质量的转录图谱等^[16]。

1.4 转录组技术平台比较

目前应用在转录组研究中的几种高通量基因表

达分析技术各有特点 (表 1), 这些技术在应用的某些方面存在重叠和竞争, RNA 测序技术具有的优势较多, 然而就目前来说, 这些技术在更多方面体现的是优势互补、数据互补, 不同方法协同配合使用, 可以互相弥补各自的遗漏部分, 将提供以前的单种技术难以提供的更加全面的转录组分析^[17-19]。例如,

对同样的样品, 分别采用基因芯片和 RNA 测序技术产生的转录组数据其相关性只有 0.45^[20], 分别采用基因芯片和 SAGE 技术产生的转录组数据也存在类似的相关性^[21]。这些数据差异主要是由不同技术的原理特性造成的, 每种技术都会丢失部分信息, 并不是仅仅由于检测灵敏性造成的^[18,22]。

表 1 几种高通量基因表达分析技术的比较

Table 1 Comparison of methods used in global analysis of gene expression of organism

	Advantages	Disadvantages
Hybridization-based (Microarray)	High specificity, low cost and rapid; ease of sample preparation; flexibility in extent to be analyzed; mature informatics and statistics;	Constrained by gene discovery and prediction procedures; difficulties to compare data in different platforms and to obtain absolute quantity of mRNA expressed in a tissue; cross-hybridization and sequence dependent; Poor reproducibility in detect minor differential expressed transcripts;
Tag-based (SAGE and MPSS)	Identification of novel sequences; quantitative, inter-laboratory comparability; sensitive and sequence independent;	Biased sampling; short tag result in sequence ambiguity; require comprehensive reference database; expensive work and complex sample preparation;
RNA-sequencing (RNA-seq)	Identification of novel sequences; quantitative, inter-laboratory comparability; sensitive and sequence independent; avoids the need for cloning; map RNA splice; longer signatures, more accurate annotation.	Expensive work and complex sample preparation; limited bioinformatics tools.

2 转录组在代谢工程领域的应用

全基因组水平的转录分析可以使研究者更加精确地评估细胞表型与基因表达的关系, 加深对细胞代谢的理解, 而且转录组分析也有助于研究者鉴定菌种改良的目标基因, 加速对微生物细胞工厂的合理设计及构建。目前转录组在代谢工程领域的应用正快速得到普及和发展。

2.1 微生物细胞特性的改造

2.1.1 优化菌种耐受性及减少代谢副产物合成

微生物在发酵生产过程中必然会遭受一些抑制细胞正常生长及产物合成的不利环境因素, 细胞对不利环境的耐受性是一种非常复杂的表型, 通过对不同环境中生长的菌株的转录组比较, 往往可以发现那些和表型密切相关的基因, 从而使研究者可以更好地通过代谢工程来增强菌种对不利环境的耐受性。Hirasawa 等^[23]比较了两株具有不同的乙醇耐受性的 *S. cerevisiae* 转录组差异, 利用聚类分析方法发现色氨酸合成基因的表达水平与乙醇耐受性紧密关系。过量表达色氨酸合成基因可以使乙醇耐受性低的菌株对 5% (V/V) 的乙醇具有耐受性, 而且外源添

加色氨酸并同时增强表达色氨酸透性酶同样可以增加对乙醇的耐受性。Hirasawa 等^[24]还通过分析高盐条件下两株不同渗透压耐受性 *S. cerevisiae* 的转录组, 发现钠离子泵与铜金属硫蛋白基因与菌体渗透压耐性密切相关, 提高上述基因的表达水平可显著提高菌株的渗透压耐性。Veit 等^[25]通过对 *E. coli* 转录组的分析发现 *sdhCDAB*、*sucB*、*sucC*、*acnB*、*lpdA*、*fumC*、*mdh* 等基因以及 *acs-yjcH-actP* 操纵子的表达与乙酸形成负相关, 增强 *sdhCDAB-b0725-sucABCD* 系列基因表达显著降低乙酸的生成, 同时不影响菌株的生长速率。由于 *E. coli* 的生长主要受到其自身分泌的乙酸的抑制, 而采取直接阻断乙酸合成途径的策略会显著降低菌体的生长速率, 上述研究结果无疑是一种更好的替代策略。

2.1.2 扩大底物利用范围

扩大菌株的底物利用范围对于利用生物质资源生产生物基化学品具有重要的意义, 一些微生物对木糖、半乳糖等底物快速利用的表型 (尤其是在厌氧条件下) 也涉及到复杂的基因表达变化。Bro 等^[26]利用转录组分析了具有不同半乳糖摄取速率的

S. cerevisiae 菌株, 鉴定出编码葡糖磷酸变位酶的 *PGM2* 基因为新的靶标, 通过增强该基因的表达使工程菌半乳糖摄取速率提高了 70%, 该研究还表明如果对底物消耗速度不同的多株菌株同时进行转录组分析, 则获得有用信息的效率将大为提高。Guimaraes 等^[27]比较了一株由进化工程获得的可快速利用乳糖的 *S. cerevisiae* 突变株与野生株的转录组差异, 大部分差异表达基因的功能集中在与 RNA 介导的转座、DNA 修复与重组、抗逆性、染色体重建、细胞周期控制、有丝分裂调控, 少部分差异表达基因与糖酵解和乙醇发酵路径相关, 这些差异表达基因能较好地解释突变株的表型, 为进一步的工程菌的设计提供了基础。Bengtsson 等^[28]分析了 4 株具有不同木糖利用能力的 *S. cerevisiae* 与正常菌株的转录组差异, 发现有 13 个基因的表达在 4 个菌株中都发生了变化。在正常菌株中相应过表达或缺失这些差异基因, 发现有 5 个基因可有效提高菌株对木糖的利用能力, 上述研究结果表明转录组是对复杂表型进行高效分析的有力工具, 其在菌株的代谢工程改造以及通过反向代谢工程重构具有目标表型的菌株中具有不可替代的作用。

2.1.3 合成异源蛋白产物

转录组指导的代谢工程加速了构建产外源蛋白微生物细胞工厂。例如, Choi 等^[29]利用基因芯片分析了高细胞浓度下重组 *E. coli* 生产类胰岛素生长因子 I 的融合蛋白 (IGF-I_f) 的转录组, 发现共有 529 个基因的表达水平显著发生改变。结合已有文献经分析后增强表达两个下调基因 *prsA* (编码磷酸核糖焦磷酸合成酶, 核酸和氨基酸合成的关键酶) 和 *glpF* (甘油运输蛋白, 有利于菌体生长), 使 IGF-I_f 的产量由 1.8 g/L 提高到 4.3 g/L。这是一个典型的利用转录组数据表征细胞表型并据此分析进而通过遗传修饰进行菌株改良的例子。Gasser 等^[30]分析了产人类胰蛋白酶 *P. pastoris* 工程菌及野生菌的转录组差异, 共有 524 个基因表达水平发生明显改变。分别扩增其中 13 个与蛋白分泌、应激反应相关的上调表达基因, 发现新确定的 6 个基因 *Bfr2*、*Bmh2*、*Ssa4*、*Sse1*、*Cup5*、*Kin2* 明显提高人类胰蛋白酶的

比生产速率, 并增加了蛋白总产量。由于蛋白质的合成分为设计剪切、折叠、分泌等多个步骤, 传统代谢靶标分析方法在这里受到很大的局限, 包括转录组在内的组学技术可深入对菌体代谢理解, 对增强蛋白质合成能力提供重要的信息。

2.1.4 提高代谢产物产率和产量

基于转录组分析来改良细胞合成目标代谢产物的产率和产量也有大量的应用实例。Lee 等^[5]分析了 *E. coli* 工程菌在过量生产苏氨酸时菌体转录组的变化, 发现编码磷酸烯醇式丙酮酸羧化酶的 *ppc* 基因表达水平没有处于最优状态。作者增强了 *ppc* 基因表达, 同时缺失转录调节子 *iclR* 基因以加强乙醛酸支路通量, 新构建的工程菌苏氨酸产量提高 51.4%, 作者在这篇文章中率先提出了“系统代谢工程”的概念, 强调了组学尺度生理分析对代谢工程的重要意义。Sindelar 等^[31]通过转录组分析确定编码转甲基酶的 *NCgl0855* 基因, 以及编码铵吸收系统的 *amtA-ocd-soxA* 操纵子为影响赖氨酸生产的潜在靶点, 增强 *NCgl0855* 或 *amtA-ocd-soxA* 操纵子的表达, 可使赖氨酸产量提高 40%。Park 等^[32]基于产缬氨酸 *E. coli* 工程菌与野生菌的转录组差异, 分别扩增了通用调节子 *lrp* 基因以及编码缬氨酸运输蛋白的 *ygaZH* 基因, 缬氨酸产量分别增加了 21.6% 和 41.7%; 共扩增上述基因则使缬氨酸产量可增加 113%, 作者指出新确定的基因修饰靶点是常规分析方法难以确定的, 这充分显示出转录组等系统分析方法的优越性。van den Berg 等^[33]利用基因芯片技术对 *P. chrysogenum* DS17690 的青霉素 G 高产工程菌与实验室标准菌株进行转录组比较, 揭示了以前未知的 68 个 β -内酰胺运输蛋白, 同时通过敲除部分差异表达基因确定了这些基因对产青霉素 G 的影响, 最终获得几株产量得到不同程度提高的工程菌株。Harris 等^[34]构建了一株产头孢菌素前体物己二酰化-7-氨基甲酰化头孢烷酸的工程菌 *P. chrysogenum* DS49834, 通过转录组分析发现工程菌中头孢菌素生物合成基因以及引入的外源基因表达水平很高, 导致己二酰化-7-氨基甲酰化头孢烷酸合成过程中消耗了大量的氨甲酰磷酸, 造成氨甲酰磷酸的供给不

足, 由此确定限制目标产物产量提高的代谢靶点为氨甲酰磷酸合成酶。Lum 等^[35]分析泰乐菌素高产菌株的基因组表达情况发现编码酰基辅酶 A 脱氢酶的 *aco* 基因和编码异丁酰辅酶 A 变位酶的 *icmA* 基因的转录水平较野生型菌株有很大提高, 增加这两个基因的拷贝数能为泰乐菌素的生产提供更多的脂肪酸前体。Kang 等^[36]通过转录组分析发现在产阿霉素 *S. peucetius* 工程菌中转录因子 *wblA* 与阿霉素合成负相关。基于此分析, Noh 等^[37]随后通过敲除转录因子 *wblA* 使 *S. peucetius* 工程菌抗生素产量提高 70%, 同时比较了敲除转录因子 *wblA* 的重组菌与其亲株的转录组, 鉴定出受转录因子 *wblA* 调控的靶基因。在这些靶基因中选取 SCO4967 在 *wblA* 敲除菌中进一步增强表达水平可提高 130% 的抗生素产量。Im 等^[38]分析了不同阿维菌素产量 *S. avermitilis* 工程菌株的转录组, 鉴定出与阿维菌素合成正相关的 50 个新基因, 其中 SAV213、SAV3818、SAV4023 等 3 个基因可促进阿维菌素合成。此外 SAV3818 基因的增强表达还可促进放线紫红素的生物合成, 表明 SAV3818 基因为一与抗生素合成相关的重要调控因子^[39]。Koetsier 等^[40]在发酵培养基中添加己二酸, 观察其对 *Penicillium chrysogenum* 转录组影响, 发现 *aclA* 基因为表达提高最显著的基因。缺失 *aclA* 基因使菌体己二酸比消耗速率下降 32%, *aclA* 基因在 *E. coli* 中表达产物可以己二酸为底物, 该基因的确有助于构建头孢菌素高产菌株。本研究组^[41]通过转录组分析从基因表达水平上揭示产核黄素工程菌 *B. subtilis* RH33 的主要生理特点, 发现在产核黄素工程菌中 PurR 蛋白调控基因表达水平全部降低, 而这些基因主要涉及核黄素前体物合成途径。继而通过基因修饰, 增强工程菌胞内磷酸核糖焦磷酸 (PRPP) 浓度, 解除 PurR 蛋白对其调控基因的抑制作用, 增加核黄素前体物的供给, 从而使核黄素产量提高 25%。上述研究工作表明通过转录组分析可以发现与表型有关的重要调节子的表达变化, 并根据该变化对菌种改造作出相应的合理设计。Hirasawa 等^[42]研究了产乳酸 *S. cerevisiae* 工程菌株的转录组, 发现有 388 个基因表达水平明显改变, 从中选取了 289 个基因进行单基因缺失, 同时随机

选取 56 个基因进行单基因缺失。结果表明, 只有基于转录组结果选取的基因可对乳酸产量发生影响, 而随机选取的基因对菌株乳酸产量基本无作用。Wierckx 等^[43]通过转录组分析指出导致 *P. putida* S12TPL3 高产苯酚的重要原因为酪氨酸合成途径基因的上调表达以及色氨酸合成途径基因的下调表达。为了增加胞内酪氨酸浓度, 作者敲除了 *hpd*、*pobA*、*phhA*、*aroP1* 等 4 个基因, 然而苯酚的产量反而下降。作者推测 *P. putida* S12TPL3 合成酪氨酸的代谢通路已经处于很好的优化状态, 酪氨酸供给已非限制因素; 此时限制苯酚生成的为酪氨酸-苯酚裂解酶, 并通过酶活分析以及随后的通量分析验证了这一假设^[44], 这是一个通过转录组分析揭示菌株代谢特征的典型研究实例。

2.2 植物细胞特性的改造

2.2.1 提高植物抗逆性

植物的生长发育及产量与外界环境密切相关, 通过代谢工程手段可以显著提高植物对环境胁迫的抗性, 保护植物免受外界环境的不良影响。Seki 等^[45]对拟南芥转录组分析, 得到了 44 个受干旱诱导基因以及 19 个受冷诱导的基因, 其中有 12 个基因被确定为植物胁迫应答的重要调节因子 CBF/DREB (C-repeat-binding-factor/dehydration-responsive-element-binding) 的靶基因。这些靶基因的确认对深入认识植物产生环境胁迫抗性的机理具有重要意义。Benedict 等^[46]将拟南芥的冷胁迫应答转录因子 *CBF1* 转入杨树, 使其冷耐受性明显增强, 同时发现杨树中受拟南芥 *CBF1* 调控的基因与拟南芥中相应基因具有高度一致性。该研究指出 CBF/DREB1 冷调控机制可增强植株对多种逆境的抵抗力。类似地, Vogel 等^[47]发现在低温环境下拟南芥基因表达水平改变最大的 25 个基因受已知的冷胁迫应答途径的关键转录因子 CBF/DREB 以及新发现的转录因子 *ZAT12* 调控, 通过增加转录因子 *ZAT12* 的表达, 拟南芥可提高其冷耐受性。然而 Fowler 等^[48]通过检测拟南芥转录组发现 CBF/DREB 途径并不是获得冷耐受性的唯一途径, 在表达水平发生改变的 306 个基因中, 只有 12% 为 CBF/DREB 的靶基因。随后 Kim 等^[49]从这些显著上调表达的基因中选择并扩增

了两个编码 DEAD-box RNA 解旋酶的基因 (*AtRH9* 和 *AtRH25*), 只有增强 *AtRH25* 基因表达水平的转基因拟南芥提高了对低温的耐受性。除此之外, Dai 等^[50]发现在低温环境下水稻共有 328 个基因表达水平发生显著变化, 将表达水平增加的基因 *OsMYB3R-2* 引入拟南芥能提高转基因拟南芥对低温、干旱、高盐等的耐受性。何飞等^[51]分析了不同拟南芥野生居群在冷胁迫处理下转录水平的差异, 不同拟南芥野生居群在冷胁迫下皆上调基因 (如 *MAPKs*、*CBFs*、*COR* 等基因) 可能对植物在低温下生存十分重要。李利华等^[52]发现水稻在低磷胁迫后大量基因 (1 207 个) 在转录水平上发生了变化, 这些差异表达基因包括了代谢调节离子转运、信号传导、转录调节、和逆境应答等方面的基因, 为提高水稻低磷耐受性提供了有用的信息。

2.2.2 提高次级代谢产物产率和产量

植物次级代谢物为医药、轻工、化工、食品及农药等工业的发展提供了丰富的资源, 然而其调控是一个十分复杂的系统, 造成其产量极低。通过转录组分析挖掘次级代谢物合成相关基因, 可有力促进相关植物细胞改造。Jennewein 等^[53]分析了受茉莉酮酸甲酯诱导而处于生产紫杉醇阶段的东北红豆杉基因表达水平, 在 10 000 个 cDNA 片段中, 确定了 3 563 个基因由茉莉酸甲酯诱导转录参与生产紫杉醇。这些基因除了包括已知的紫杉烷代谢所需羟化酶外, 还提供了参与紫杉醇合成反应的其他潜在基因, 为改良生产菌株提供了有价值的信息。Teoh 等^[54]在青蒿的转录组文库经序列比对确定一种多功能的 P450 氧化酶 (由 *CYP71AV1* 基因编码), 并将其在酵母中表达, 发现其可以催化 FPP 环化产物紫穗槐-4,11-二烯 (*amorpho-4, 11-diene*) 到青蒿酸的多步氧化反应, 该关键基因的确定推动了高产青蒿素工程菌的构建^[55-56]。Nagel 等^[57]在蛇麻腺体 (*Lupulin glands*) 的含有 10 581 表达序列标签的基因转录文库中鉴定出 4 个编码 S-腺苷甲硫氨酸-甲基转移酶 (*S-adenosyl-L-methionine-dependent O-methyltransferases*, OMTs) 的基因, 为催化黄腐醇生物合成的关键基因, 将加速生产植株的构建。Hirai 等^[58]通过计算数据库中公开发表的 1 388 张基

因芯片数据的皮尔森相关系数发现, 拟南芥中脂肪族芥子油苷生物合成基因与两个未报道的转录因子 *Myb28* (*At5g61420*) 和 *Myb29* (*At5g07690*) 相关性很大。通过基因缺失和基因异位表达表明脂肪族芥子油苷生物合成主要受转录因子 *Myb28* 调控, 增强转录因子 *Myb28* 的表达水平可致使拟南芥合成大量芥子油苷。这些结果表明在转录组指导下可高效确定细胞改造策略。

2.3 动物细胞特性的改造

动物细胞系目前已经被广泛用于蛋白质药物等产品的大量生产上, 利用动物细胞表达蛋白其优势在于有助于蛋白质正确折叠、组装并进行翻译后的修饰, 目标蛋白质可正常行使其功能^[59]。转录组分析在减少细胞代谢负担、控制细胞贴壁性、调控细胞生长活性等方面都有成功的应用。

Khoo 等^[60]发现一株具有单抗生产能力的鼠骨髓瘤细胞 (*GS-NS0 6A1-100*) 与相应野生型 (*NS0 WT*) 生长代谢速率接近, 进而比较了两株细胞系表达谱差异, 推测这些基因表达差异使 *GS-NS0 6A1-100* 细胞可高效利用底物, 补偿了其合成单抗造成的代谢负担, 使两株细胞系生长速率类似。Beer 等^[61]发现小鼠 NIH 3T3 细胞在 32℃ 可最优生产双嗜性白血病毒 4070A, 而当温度提高到 37℃ 时病毒产量降低。他们通过转录组分析发现 37℃ 培养时胆固醇合成与运输相关基因的表达受到抑制, 音猬因子信号通路 (*Sonic hedgehog signalling pathway*) 相关基因的表达降低。这些温度诱导的差异表达基因会影响细胞生长代谢和蛋白合成速率, 是今后改良细胞工厂的靶标基因。Jaluria 等^[62]通过聚类分析算法分析非贴壁依赖性与贴壁依赖性 HeLa 细胞转录组差异, 鉴定出 *siat7e* 和 *lama4* 两个细胞贴壁性相关基因, 降低 *siat7e* 基因的表达或提高 *lama4* 基因的表达皆可改善 HeLa 细胞系的贴壁性。Jaluria 等^[63]在生长迅速且具非贴壁依赖性 HeLa 细胞系的转录组中发现两个上调表达基因 (*cdkl3* 和 *cox15*) 与细胞增殖密切相关, 其中 *cdkl3* 基因编码细胞周期依赖性蛋白激酶, *cox15* 基因编码细胞色素氧化酶亚单位。在 HeLa 细胞、人胚肾 293 细胞 (*Human embryonic kidney-293*, HEK-293)、中国仓鼠卵巢巢

胞 (Chinese hamster ovary, CHO)、犬肾上皮细胞 (Madin-darby canine kidney, MDCK) 中, 引入并过表达这两个基因皆可提高细胞增殖速率同时增加了最高活性细胞密度。Wong 等^[64]通过比较产重组人干扰素 γ 的中国仓鼠卵巢细胞不同时期的转录组差异, 发现 9 个早期细胞凋亡信号基因 (*FasL*、*Fadd*、*Bim*、*Bak*、*Faim*、*Bad*、*Bax*、*Alg-2*、*Requiem*)。 *Fadd*、*Faim*、*Alg-2* 和 *Requiem* 基因在随后的研究中分别在中国仓鼠卵巢细胞增强或降低其表达水平^[65], 这 4 个相关基因表达水平改变后都提高了细胞凋亡抗性及活性细胞密度, 最优的一株重组细胞系重组人干扰素产量提高至原始产量的 2.5 倍。

表 2 多种组学相结合在代谢工程中的应用

Table 2 Examples of combined omics guided metabolic engineering

Omics	Host	Description	References
Transcriptome and proteome	<i>Escherichia coli</i>	Identifying gene targets and understanding phenotypic behavior of wild type strain under high cell density culture	[66]
	<i>Escherichia coli</i>	Identifying the traits leading to L-threonine overproducing	[69]
	Murine NS0 myeloma cell line	Discovery the factors regulating cholesterol biosynthesis, useful for engineering these cells	[71-72]
Transcriptome and metabolome	<i>Aspergillus terreus</i>	Improved understanding of the metabolic pathways and fluxes was conducted to enhance the yields of lovastatin and (+)-geodin	[67]
	Nipponbare	Revealing the limited pathway for producing tryptophan	[70]
Transcriptome and fluxome	<i>Bacillus subtilis</i>	Identifying gene targets and understanding the mechanism leading to high riboflavin productivity	[68]
Transcriptome and genome	<i>Zygaena filipendulae</i>	Identifying gene candidates for biosynthesis of cyanogenic glucosides	[73]

Yoon 等^[66]研究了处于高菌体浓度状态下的 *E. coli* 的转录组与蛋白质组数据, 实验表明在细胞密度增加时表达氨基酸合成的基因受到抑制, 直接导致目标重组蛋白产量降低, 同时发现伴侣基因 (*clpB*、*dnaJK*、*hslSG*、*groEL*、*groES*) 的表达也明显增加, 表明菌体处于营养匮乏状态。Askenazi 等^[67]利用层次聚类 (Hierarchical clustering) 和主成分分析 (Principle component analysis) 等方法分析了具有不同洛伐他汀产量的 *Aspergillus terreus* 菌株转录组、代谢物组数据, 鉴定出影响产物形成的靶标并对其进行成功改造。Askenazi 等研究人员的工作是一个综合运用组学数据提高微生物生产能力的典型案例。Zamboni 等^[68]发现高产核黄素工程菌 *B. subtilis* RB50::pRF69 不同生长阶段产核黄素速率

3 转录组学与其他组学技术的综合运用

由于细胞内各个层次间的调控互相关联, 单一组学分析具有一定局限性, 另外转录组分析并不能完全准确地反应细胞内基因的翻译情况, 对于带有大量突变的突变株, 基因序列的突变可能对其转录水平没有影响, 但可显著降低翻译的效率, 这时单纯依靠转录组分析就会遗漏重要的差异表达信息。因此转录组与其他组学技术的整合分析显得非常重要, 各种组学技术的综合运用有利于重要信息的补充和整合, 将加深对复杂生物系统的认识、加速对代谢靶标的识别 (表 2)。

不同, 通过研究相应阶段的转录组和代谢通量组数据, 揭示了核黄素生产阶段碳代谢流分布, 即菌株主要采取非氧化磷酸戊糖途径获取核黄素合成前体物, 指出了此时限制核黄素生产的代谢瓶颈。Lee 等^[69]采用转录组和蛋白质组技术综合分析了 *E. coli* W3110 及其过量产苏氨酸突变株 TF5015, 发现有 54 个基因表达水平差异显著, 乙醛酸循环、三羧酸循环以及氨基酸生物合成途径被激活。通过基因测序, 发现苏氨酸合成基因 *thrA* 不仅表达水平提高而且其序列还发生了突变, 此外在表达未改变的异亮氨酸合成基因 *ilvA* 中也发生了突变, 发生的突变有助于解除对苏氨酸合成路径的反馈抑制, 从而使苏氨酸过量合成。将发生突变的苏氨酸合成基因导入 *E. coli* W3110 可提高其苏氨酸产量。Dubouzet 等^[70]

比较了正常日本晴与产色氨酸转基因日本晴的转录组和代谢物组, 发现转基因日本晴中与色氨酸共享莽草酸途径的苯丙氨酸、酪氨酸的浓度未受影响, 其合成途径的转录水平未有明显改变, 说明代谢流未过多进入与色氨酸合成存在竞争的分支代谢途径。此外转基因日本晴中, 转化色氨酸的色氨酸脱羧酶酶活较低, 有利于色氨酸的累积。鼠骨髓瘤细胞 (NS0) 正常生长通常需要外源添加胆固醇, 为了改良培养基而降低成本, Seth 等^[71]比较了需添加胆固醇的鼠骨髓瘤细胞 (NS0) 与不需添加胆固醇的突变鼠骨髓瘤细胞 (NS0 revertant) 转录组与蛋白质组差异, 结果表明表达降低的 *Hsd17b7* 基因导致胆固醇营养缺陷型, 并通过实验验证^[72]。Zagrobelyny 等^[73]通过 Roche 公司 454 的焦磷酸测序法测定了六星燈蛾 *Zygaena filipendulae* 的转录组, 随后结合各转录本序列的长度和拷贝数分析, 以及六星燈蛾与黑腹果蝇、家蚕、产氰蝶 *Heliconius* 的系统发育关系的分析确定了生物活性物质氰苷的潜在合成基因, 该研究表明基于焦磷酸测序技术的转录组分析方法在获取生物体内天然生物活性物质的合成基因等相关信息方面具有非常重要的作用。

4 展望

近些年来代谢工程手段被引入到细胞工厂改良的研究中, 并取得不错的效果, 特别是以转录组学为代表的组学技术的应用增强了代谢工程的指导性可与预见性, 使代谢工程得到重大革新, 有力推动了工业生物技术发展^[74]。目前, 在以转录组学为代表的组学技术指导下的代谢工程正日益成为研究的热点, 国内外关于组学研究的研究中心及公司纷纷成立, 科学工作者获得了空前大量的反映生命活动的数据。与此同时, 如何更好地理解、运用这些生命信息也成为摆在我们面前的一个重大挑战, 目前组学数据分析主要是通过比较组学数据差异完成的, 组学层次模型还不多见。为了更加合理、高效地实现细胞改良, 还需要不同专业背景的科学工作者通力合作, 特别是在信息学、数学的帮助下模拟和设计合成新的生物系统, 建立相应动力学模型和控制模型, 从而提高现有生物基产品的产量、合成

新的生物基产品。

REFERENCES

- [1] Nielsen J, Olsson L. An expanded role for microbial physiology in metabolic engineering and functional genomics: moving towards systems biology. *FEMS Yeast Res*, 2002, **2**(2): 175–181.
- [2] Zhao XM, Wang JY, Chen T, *et al.* Metabolic engineering in the post genomic era: opportunities and challenges. *Chin J Bioprocess Eng*, 2004, **2**: 1–7.
赵学明, 王靖宇, 陈涛, 等. 后基因组时代的代谢工程: 机遇与挑战. *生物加工过程*, 2004, **2**: 1–7.
- [3] Ideker T, Galitski T, Hood L. A new approach to decoding life: systems biology. *Annu Rev Genomics Hum Genet*, 2001, **2**(1): 343–372.
- [4] Lee SY, Lee DY, Kim TY. Systems biotechnology for strain improvement. *Trends Biotechnol*, 2005, **23**(7): 349–358.
- [5] Lee KH, Park JH, Kim TY, *et al.* Systems metabolic engineering of *Escherichia coli* for L-threonine production. *Mol Syst Biol*, 2007, **3**: 149–156.
- [6] Lockhart DJ, Winzeler EA. Genomics, gene expression and DNA arrays. *Nature*, 2000, **405**(6788): 827–836.
- [7] Jewett MC, Oliveira AP, Patil KR, *et al.* The role of high-throughput transcriptome analysis in metabolic engineering. *Biotechnol Bioproc Eng*, 2005, **10**: 385–399.
- [8] Donson J, Fang Y, Espiritu-Santo G, *et al.* Comprehensive gene expression analysis by transcript profiling. *Plant Mol Biol*, 2002, **48**: 75–97.
- [9] Brazma A, Hingamp P, Quackenbush J, *et al.* Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat Genet*, 2001, **29**(4): 365–372.
- [10] Shi L, Reid LH, Jones WD, *et al.* The MicroArray Quality Control (MAQC) project shows inter-and intraplatform reproducibility of gene expression measurements. *Nat Biotechnol*, 2006, **24**(9): 1151–1161.
- [11] Velculescu VE, Zhang L, Vogelstein B, *et al.* Serial analysis of gene expression. *Science*, 1995, **270**: 484–487.
- [12] Brenner S, Johnson M, Bridgham J, *et al.* Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nat Biotechnol*, 2000, **18**: 630–634.
- [13] Mortazavi A, Williams BA, McCue K, *et al.* Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods*, 2008, **5**(7): 621–628.
- [14] Wang Z, Gerstein M, Snyder M. RNA-Seq: a

- revolutionary tool for transcriptomics. *Nat Rev Genet*, 2009, **10**: 57–63.
- [15] Siddiqui AS, Delaney AD, Schnerch A, *et al.* Sequence biases in large scale gene expression profiling data. *Nucleic Acids Res*, 2006, **34**(12): e83.
- [16] van Vliet A. Next generation sequencing of microbial transcriptomes: challenges and opportunities. *FEMS Microbiol Lett*, 2010, **302**: 1–7.
- [17] Oudes AJ, Roach JC, Walashek LS, *et al.* Application of affymetrix array and massively parallel signature sequencing for identification of genes involved in prostate cancer progression. *BMC Cancer*, 2005, **5**: 86.
- [18] Liu F, Janssen TK, Trimarchi J, *et al.* Comparison of hybridization-based and sequencing-based gene expression technologies on biological replicates. *BMC Genomics*, 2007, **8**: 153–166.
- [19] Hanriot L, Keime C, Gay N, *et al.* A combination of LongSAGE with Solexa sequencing is well suited to explore the depth and the complexity of transcriptome. *BMC Genomics*, 2008, **9**: 418–426.
- [20] Weber APM, Weber KL, Carr K, *et al.* Sampling the *Arabidopsis* transcriptome with massively parallel pyrosequencing. *Plant Physiol*, 2007, **144**(1): 32–42.
- [21] van Ruissen F, Ruijter JM, Schaaf GJ, *et al.* Evaluation of the similarity of gene expression data estimated with SAGE and Affymetrix GeneChips. *BMC Genomics*, 2005, **6**: 91–106.
- [22] Sultan M, Schulz M, Richard H, *et al.* A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science*, 2008, **321**(5891): 956–959.
- [23] Hirasawa T, Yoshikawa K, Nakakura Y, *et al.* Identification of target genes conferring ethanol stress tolerance to *Saccharomyces cerevisiae* based on DNA microarray data analysis. *J Biotechnol*, 2007, **131**: 34–44.
- [24] Hirasawa T, Nakakura Y, Yoshikawa K, *et al.* Comparative analysis of transcriptional responses to saline stress in the laboratory and brewing strains of *Saccharomyces cerevisiae* with DNA microarray. *Appl Microbiol Biotechnol*, 2006, **70**: 346–357.
- [25] Veit A, Polen T, Wendisch VF. Global gene expression analysis of glucose overflow metabolism in *Escherichia coli* and reduction of aerobic acetate formation. *Appl Microbiol Biotechnol*, 2007, **74**(2): 406–421.
- [26] Bro C, Knudsen S, Regenbreg B, *et al.* Improvement of galactose uptake in *Saccharomyces cerevisiae* through overexpression of phosphoglucosmutase: Example of transcript analysis as a tool in inverse metabolic engineering. *Appl Environ Microbiol*, 2005, **71**(11): 6465–6472.
- [27] Guimaraes P, Le Berre V, Sokol S, *et al.* Comparative transcriptome analysis between original and evolved recombinant lactose-consuming *Saccharomyces cerevisiae* strains. *Biotechnol J*, 2008, **3**: 1591–1597.
- [28] Bengtsson O, Jeppsson M, Sonderegger M, *et al.* Identification of common traits in improved xylose-growing *Saccharomyces cerevisiae* for inverse metabolic engineering. *Yeast*, 2008, **25**(11): 835–847.
- [29] Choi JH, Lee SJ, Lee SJ, *et al.* Enhanced production of insulin-like growth factor I fusion protein in *Escherichia coli* by coexpression of the down-regulated genes identified by transcriptome profiling. *Appl Environ Microbiol*, 2003, **69**: 4737–4742.
- [30] Gasser B, Sauer M, Maurer M, *et al.* Transcriptomics-based identification of novel factors enhancing heterologous protein secretion in Yeasts. *Appl Environ Microbiol*, 2007, **73**: 6499–6507.
- [31] Sindelar G, Wendisch VF. Improving lysine production by *Corynebacterium glutamicum* through DNA microarray-based identification of novel target genes. *Appl Microbiol Biotechnol*, 2007, **76**: 677–689.
- [32] Park JH, Lee KH, Kim TY, *et al.* Metabolic engineering of *Escherichia coli* for the production of L-valine based on transcriptome analysis and in silico gene knockout simulation. *Proc Natl Acad Sci USA*, 2007, **104**: 7797–7802.
- [33] van den Berg MA, Albang R, Albermann K, *et al.* Genome sequencing and analysis of the filamentous fungus *Penicillium chrysogenum*. *Nat Biotechnol*, 2008, **26**: 1161–1168.
- [34] Harris DM, Westerlaken I, Schipper D, *et al.* Engineering of *Penicillium chrysogenum* for fermentative production of a novel carbamoylated cephem antibiotic precursor. *Metab Eng*, 2009, **11**: 125–137.
- [35] Lum A, Huang J, Hutchinson C, *et al.* Reverse engineering of industrial pharmaceutical-producing actinomycete strains using DNA microarrays. *Metab Eng*, 2004, **6**(3): 186–196.
- [36] Kang SH, Huang J, Lee HN, *et al.* Interspecies DNA microarray analysis identifies WblA as a Pleiotropic Down-Regulator of antibiotic biosynthesis in *Streptomyces*. *J Bacteriol*, 2007, **189**(11): 4315–4319.
- [37] Noh J, Kim S, Lee H, *et al.* Isolation and genetic manipulation of the antibiotic down-regulatory gene, wblA ortholog for doxorubicin-producing *Streptomyces* strain improvement. *Appl Microbiol Biotechnol*, 2010, **86**: 1–9.

- [38] Im J, Kim M, Kim E. Comparative transcriptome analysis for avermectin overproduction via *Streptomyces avermitilis* microarray system. *J Microbiol Biotechnol*, 2007, **17**(3): 534.
- [39] Duong C, Lee H, Choi S, *et al.* Functional expression of SAV3818, a putative TetR-family transcriptional regulatory gene from *Streptomyces avermitilis*, stimulates antibiotic production in *Streptomyces* species. *J Microbiol Biotechnol*, 2009, **19**(2): 136–139.
- [40] Koetsier MJ, Gombert AK, Fekken S, *et al.* The *Penicillium chrysogenum acla* gene encodes a broad-substrate-specificity acyl-coenzyme A ligase involved in activation of adipic acid, a side-chain precursor for cephem antibiotics. *Fungal Genet Biol*, 2010, **47**(1): 33–42.
- [41] Shi SB, Chen T, Zhang ZG, *et al.* Transcriptome analysis guided metabolic engineering of *Bacillus subtilis* for riboflavin production. *Metab Eng*, 2009, **11**: 243–252.
- [42] Hirasawa T, Ookubo A, Yoshikawa K, *et al.* Investigating the effectiveness of DNA microarray analysis for identifying the genes involved in l-lactate production by *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol*, 2009, **84**(6): 1149–1159.
- [43] Wierckx NJP, Ballerstedt H, de Bont JAM, *et al.* Transcriptome analysis of a Phenol-producing *Pseudomonas putida* S12 construct: genetic and physiological basis for improved production. *J Bacteriol*, 2008, **190**: 2822–2830.
- [44] Wierckx N, Ruijsenaars HJ, de Winde JH, *et al.* Metabolic flux analysis of a phenol producing mutant of *Pseudomonas putida* S12: verification and complementation of hypotheses derived from transcriptomics. *J Biotechnol*, 2009, **143**(2): 124–129.
- [45] Seki M, Narusaka M, Abe H, *et al.* Monitoring the expression pattern of 1300 *Arabidopsis* genes under drought and cold stresses by using a full-length cDNA microarray. *Plant Cell*, 2001, **13**(1): 61–72.
- [46] Benedict C, Skinner J, Meng R, *et al.* The CBF1-dependent low temperature signalling pathway, regulon and increase in freeze tolerance are conserved in *Populus* spp. *Plant Cell Environ*, 2006, **29**(7): 1259.
- [47] Vogel J, Zarka D, Van Buskirk H, *et al.* Roles of the CBF2 and ZAT12 transcription factors in configuring the low temperature transcriptome of *Arabidopsis*. *Plant J*, 2005, **41**(2): 195–211.
- [48] Fowler S, Thomashow MF. *Arabidopsis* transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF cold response pathway. *Plant Cell*, 2002, **14**(8): 1675–1690.
- [49] Kim JS, Kim KA, Oh TR, *et al.* Functional characterization of DEAD-Box RNA helicases in *Arabidopsis thaliana* under abiotic stress conditions. *Plant Cell Physiol*, 2008, **49**(10): 1563–1571.
- [50] Dai X, Xu Y, Ma Q, *et al.* Overexpression of an R1R2R3 MYB gene, *OsMYB3R-2*, increases tolerance to freezing, drought, and salt stress in transgenic *Arabidopsis*. *Plant Physiol*, 2007, **143**(4): 1739–1751.
- [51] He F, Kang JQ, Zhou X, *et al.* Monitoring expression profiles of *Arabidopsis* genes during cold acclimation. *Chin Sci Bull*, 2008, **53**(18): 2206–2215.
何飞, 康菊清, 周鑫, 等. 中国野生拟南芥居群冷胁迫下的表达谱变异. 科学通报, 2008, **53**(18): 2206–2215.
- [52] Li LH, Qiu XH, Li XH, *et al.* Expression profiles of low nitrogen stress in rice. *Chin Sci Bull*, 2010, **3**: 12–19.
李利华, 邱旭华, 李香花, 等. 水稻低磷胁迫基因表达谱分析. 科学通报, 2010, **3**: 12–19.
- [53] Jennewein S, Wildung M, Chau M, *et al.* Random sequencing of an induced *Taxus* cell cDNA library for identification of clones involved in Taxol biosynthesis. *Proc Natl Acad Sci USA*, 2004, **101**(24): 9149–9154.
- [54] Teoh K, Polichuk D, Reed D, *et al.* *Artemisia annua* L.(Asteraceae) trichome-specific cDNAs reveal CYP71AV1, a cytochrome P450 with a key role in the biosynthesis of the antimalarial sesquiterpene lactone artemisinin. *FEBS Lett*, 2006, **580**(5): 1411–1416.
- [55] Arsenaault P, Wobbe K, Weathers P. Recent advances in artemisinin production through heterologous expression. *Curr Med Chem*, 2008, **15**(27): 2886–2896.
- [56] Kirby J, Keasling J. Metabolic engineering of microorganisms for isoprenoid production. *Nat Prod Rep*, 2008, **25**(4): 656–661.
- [57] Nagel J, Culley L, Lu Y, *et al.* EST analysis of hop glandular trichomes identifies an O-methyltransferase that catalyzes the biosynthesis of xanthohumol. *Plant Cell*, 2008, **20**(1): 186–200.
- [58] Hirai MY, Sugiyama K, Sawada Y, *et al.* Omics-based identification of *Arabidopsis* Myb transcription factors regulating aliphatic glucosinolate biosynthesis. *Proc Natl Acad Sci USA*, 2007, **104**(15): 6478–6483.
- [59] Wurm FM. Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat Biotechnol*, 2004, **22**(11): 1393–1398.
- [60] Khoo S, Falciani F, Al-Rubeai M. A genome-wide transcriptional analysis of producer and non-producer NS0 myeloma cell lines. *Biotechnol Appl Biochem*, 2007, **47**: 85–95.

- [61] Beer C, Buhr P, Hahn H, *et al.* Gene expression analysis of murine cells producing amphotropic mouse leukaemia virus at a cultivation temperature of 32 and 37°C. *J Gen Virol*, 2003, **84**(7): 1677–1686.
- [62] Jaluria P, Betenbaugh M, Konstantopoulos K, *et al.* Application of microarrays to identify and characterize genes involved in attachment dependence in HeLa cells. *Metab Eng*, 2007a, **9**(3): 241–251.
- [63] Jaluria P, Betenbaugh M, Konstantopoulos K, *et al.* Enhancement of cell proliferation in mammalian cells by gene insertion of a cyclin-dependent kinase homolog. *BMC Biotechnol*, 2007b, **7**: 71–81.
- [64] Wong DCF, Wong KTK, Lee YY, *et al.* Transcriptional profiling of apoptotic pathways in batch and fed-batch CHO cell cultures. *Biotechnol Bioeng*, 2006a, **94**(2): 373–382.
- [65] Wong DCF, Wong KTK, Nissom PM, *et al.* Targeting early apoptotic genes in batch and fed-batch CHO cell cultures. *Biotechnol Bioeng*, 2006b, **95**(3): 350–361.
- [66] Yoon SH, Han MJ, Lee SY, *et al.* Combined transcriptome and proteome analysis of *Escherichia coli* during high cell density culture. *Biotechnol Bioeng*, 2003, **81**: 753–767.
- [67] Askenazi M, Driggers EM, Holtzman DA, *et al.* Integrating transcriptional and metabolite profiles to direct the engineering of lovastatin-producing fungal strains. *Nat Biotechnol*, 2003, **21**(2): 150–156.
- [68] Zamboni N, Fischer E, Muffler A, *et al.* Transient expression and flux changes during a shift from high to low riboflavin production in continuous cultures of *Bacillus subtilis*. *Biotechnol Bioeng*, 2005, **89**(2): 219–232.
- [69] Lee JH, Lee DE, Lee BU, *et al.* Global analyses of transcriptomes and proteomes of a parent strain and an L-threonine-overproducing mutant strain. *J Bacteriol*, 2003, **185**(18): 5442–5451.
- [70] Dubouzet J, Ishihara A, Matsuda F, *et al.* Integrated metabolomic and transcriptomic analyses of high-tryptophan rice expressing a mutant anthranilate synthase alpha subunit. *J Exp Bot*, 2007, **58**(12): 3309–3321.
- [71] Seth G, Philp RJ, Denoya CD, *et al.* Large-scale gene expression analysis of cholesterol dependence in NSO cells. *Biotechnol Bioeng*, 2005, **90**(5): 552–567.
- [72] Seth G, Ozturk M, Hu WS. Reverting cholesterol auxotrophy of NSO cells by altering epigenetic gene silencing. *Biotechnol Bioeng*, 2005, **93**(4): 820–827.
- [73] Zagrobelny M, Scheibye-Alsing K, Jensen NB, *et al.* 454 pyrosequencing based transcriptome analysis of *Zygaena filipendulae* with focus on genes involved in biosynthesis of cyanogenic glucosides. *BMC Genomics*, 2009, **10**: 574–588.
- [74] Li Y. Metabolic engineering: an evolving technology for strain improvement. *Chin J Biotech*, 2009, **25**(9): 1281–1284.
- 李寅. 代谢工程: 一项不断发展的菌株改造技术. *生物工程学报*, 2009, **25**(9): 1281–1284.