

合成生物学及其在生物技术中的应用进展*

吕 静¹⁾ 孙洪磊²⁾ 何 皓²⁾ 傅鹏程^{1)**}

¹⁾中国石油大学(北京)化工学院重质油国家重点实验室, 新能源研究中心, 北京 102249;

²⁾中国石油天然气股份有限公司石油化工研究院, 北京 100195)

摘要 合成生物学是一门 21 世纪生物学的新兴学科, 它着眼生物科学与工程科学的结合, 把生物系统当作工程系统“从下往上”进行处理, 由“单元”(unit)到“部件”(device)再到“系统”(system)来设计, 修改和组装细胞构件及生物系统。合成生物学是分子和细胞生物学、进化系统学、生物化学、信息学、数学、计算机和工程等多学科交叉的产物。目前研究应用包括两个主要方面: 一是通过对现有的、天然存在的生物系统进行重新设计和改造, 修改已存在的生物系统, 使该系统增添新的功能。二是通过设计和构建新的生物零件、组件和系统, 创造自然界中尚不存在的人工生命系统。合成生物学作为一门建立在基因组方法之上的学科, 主要强调对创造人工生命形态的计算生物学与实验生物学的协同整合。必须强调的是, 用来构建生命系统新结构、产生新功能所使用的组件单元既可以是基因、核酸等生物组件, 也可以是化学的、机械的和物理的元件。本文跟踪合成生物学研究及应用, 对其在 DNA 水平编程、分子修饰、代谢途径、调控网络和工业生物技术等方面的进展进行综述。

关键词 合成生物学, 系统生物学, 蓝藻, 底盘, 生物燃料

学科分类号 Q6

DOI: 10.3724/SP.J.1206.2011.00583

1 合成生物学概述

1.1 新一代生物学——合成生物学

20 世纪生物学研究一直以“还原论”为指导, 即对生物系统不断分解, 直至细胞中的单个或有限个基因或蛋白质, 然后孤立研究这些基因和 / 或蛋白质的结构和功能, 以此了解生物现象。随着基因组测序和高通量筛选测量为标志的当代分子生物学的迅猛发展, 以系统化设计和工程化构建为理念的合成生物学成为新一代生物学的发展方向。

2000 年 1 月《自然》(*Nature*)同时发表了两篇文章。其一是 Colins 团队研制出由两个抑制基因、两个抑制启动子以及一个作为报告基因的绿色荧光蛋白(GFP)组成的一种双稳态“基因套环开关”, 可对选择的细胞功能进行开关^[1]。其二是 Elowitz 和 Liebler 用转录启动子和抑制基因构建了由连续诱导启动子调控的 3 基因抑制网络, 称为“压缩振荡子”, 同样加上一个 GFP 报告基因。它将交替打开或者关闭 GFP 报告基因, 使细胞能在发光状态和非发光状态之间转换^[2]。随后, 许多合成生物学的基本元件, 例如启动子、核糖体结合位点和转录阻

遏物等, 均用来构建具有特定功能的模块, 将这些模块插入细胞使生物系统具有了新的功能。目前, 合成生物学模块包括了诸如套环开关、串级开关、脉冲发生器、时间延迟电路、振荡器、逻辑门电路等。这些模块和其他模块一起工作时, 可以用来调控基因表达、蛋白质功能、代谢及细胞间的通讯^[3]。以基因工程技术和电子工程的电路设计原则为基础的工作还包括利用启动子和阻遏子等基因元件构建最简单的组件创建可通用组装的, 满足不同的组合要求的最简单的模块库^[4]。应用例子包括逻辑门、门锁(套环开关)^[1, 5]和逆变器^[6]。

可以看出, 合成生物学以信号传导、基因调控以及细胞代谢的元件组装具有我们所希望的细胞功

* 国家重点基础研究发展计划(973)(2011CB200902)资助项目, 中国石油天然气股份有限公司科技研究外协项目《制取生物柴油的工程微藻的筛选与培育》、《浮萍和微藻能源化的资源潜力与过程的中试开发》和《中国航空生物燃料炼制加工技术研究》资助项目。

** 通讯联系人。

Tel: 010-89731283, E-mail: pengcheng@cup.edu.cn

收稿日期: 2011-07-27, 接受日期: 2012-02-13

能的复杂生物系统. 它通过对系统行为的定量预测和细胞改造的工程实施完成合成任务. 必须强调的是, 用来构建生命系统新结构、产生新功能所使用的组件单元既可以是基因、核酸等生物组件, 也可以是化学的、机械的和物理的元件. 从系统设计的角度来看, 合成生物学通过在 DNA 水平编程, 或重新编程来更新和 / 或创造新的生物实体, 如分子修饰、代谢途径、调控网络、生物和生态系统等. 它的发展将大大增强我们以系统论、信息论和控制论为指导对复杂的细胞网络进行微观改造, 进而取得生物系统宏观表型改造的能力.

1.2 基因调控网络和代谢网络的人工构建

基因调控网络的研究是现阶段合成生物学进行得比较多的工作, 它已经从基因调控线路的构建扩展到细胞内信号通路的重连接^[7]、胞内 RNA 逻辑门电路的实现^[8]、具有细胞 - 细胞通讯的多细胞体系^[9]等. 研究表明, 海胆胚胎内中胚层形成的基因调控网络具有五种不同类型的扰动^[10]. 小鼠 B 淋巴细胞和他们的两个胚胎干细胞的祖细胞形成的调控机制也得到了确定^[11]. 在更高的层面上, 一种细菌群体“可调细胞 - 细胞通讯”的工程改造获得了较好的结果^[12-13]. 此外, 对改变配基结合特性的受体和传感器蛋白的计算模型再设计得到了预期的结果^[14-15]. 类似的工作还包括: 合成生物学实现了更改细菌转运和启动子的反应从“全或无”(all-or-none)到可调节逻辑. 通过重组的有丝分裂过程的研究, 合成生物学表征并改变微管和有丝分裂纺锤体蛋白组分的染色体分离、组装、拆卸和运动动力学机制^[16-23].

传统工业生物技术侧重于从环境中筛选具有代谢特定物质能力的天然微生物进行工业应用. 而基于合成生物学的生物技术研发模式则是在细胞中构建由天然或非天然功能元件组成、自然界不存在的合成代谢途径, 形成有生物制造能力的细胞工厂. 合成生物学是从工程系统的角度出发进行设计, 并且通过在细胞内植入适当的生物控制元件达到添加新功能的目的. 这些基本生物控制元件在制造方法和植入上的标准化和模块化, 以及可互换零件进行系统制造, 拥有预测模型和稳健的设计库等, 均可以促使统一生物元件备件库的建立. 标准化使得未来的基因线路就如现在的集成电路一样, 只需用不同的生物元件, 即可设计组装新的基因线路产生所希望的功能. 模块化则使基因线组装变得容易, 并且可以充分利用前人组装出的生物元件, 组合出更

加先进的基因调控网络, 并可应用强大的计算机工程工具(如: CAD 计算机辅助设计)进行综合设计. 进一步的工作包括基因工程技术标准的改进、对遗传材料装配标准化和快速表征; 开发的遗传材料“零件”系列能在构建的基因调控线路和代谢网络中可靠工作, 不产生设计之外的相互作用, 以及生成安全、稳健的宿主细胞.

1.3 非天然氨基酸——合成生物学的元件

合成生物学使用工程的形式去定义和构建由生物单元, 甚至是化工、机械、或计算单元组成的模块. 由此产生的模块可能包含非天然材料. 蛋白质序列中添加非天然氨基酸是合成生物学设计新功能蛋白质的一个策略. 这种策略对天然蛋白质的折叠和功能研究有重要作用. 目前, 已有大约 30 多种非天然氨基酸被人工插入到生物体合成的蛋白质中^[24-29]. 但一种蛋白质中只能插入少数的非天然氨基酸. 随着特别基因组序列合成手段的改进, 将会用 20 个密码子特异性编码 20 种天然氨基酸, 其余的密码子则用来编码非天然氨基酸. 如果成功, 将会实现多种非天然氨基酸的成功插入, 并且对细胞功能没有特别的影响. 加入到现有生物体的非天然氨基酸可以用酰化化学方法合成, 并采用新增加的 tRNA 琥珀突变抑制手段辨识相关蛋白质. tRNA 工程已经用来改造和识别 4 和 5 个碱基的 DNA 密码子^[30-31]. 工业应用方面, 已经有通过半胱氨酸合成途径工程改造的方法发酵生产 O-乙酰-L-丝氨酸和其他非天然 L 型氨基酸作为药品、农药合成的前体的报道^[32].

1.4 底盘与最小基因组

仅凭合成和设计 DNA 获得预期的理想功能是不够的. 这些元件——合成的 DNA 操纵子, 必须能够独立作用, 而不受细胞其他过程的干扰. 解决这个问题方法就是: 合成生物学的另外研究领域——最小基因组细胞. 最小基因组细胞仅含有基本的遗传信息, 在一定的条件下可保持细胞活力. 细胞含有许多生物技术应用之外的元件, 最小基因组细胞提供了一个无干扰的适合于生物工程操作的“底盘”. 目前尚不清楚最小基因组细胞的遗传定义是通用的还是特定的. 因此, 可能需要一系列最小基因组细胞用于生产不同种类的生物产品.

选择一个合适的底盘需要符合以下标准: 具有一些预期的特性, 可以更好地表征和注释这些生物(底盘), 具有适合该底盘的分子生物学改造工具. 在工业水平上, 一个稳定的底盘应该能够高效经济

地生产燃料和化合物. 例如在生物催化方面, 优秀底盘应具备如下特性^[33]: a. 在矿物盐培养基上可以利用廉价碳源生长; b. 可利用五碳和六碳糖, 并能转化利用半纤维素生物质中的糖组分; c. 代谢速率和产率高; d. 发酵过程简单, 降低大规模生产时操作成本和失败风险; e. 底盘耐受性强(高温、低 pH), 可以减少纤维素降解过程中外加纤维素酶和碱的用量; f. 遗传操作简单, 遗传稳定性强; g. 耐受生物质预处理过程产生的抑制剂; h. 耐受高浓度的底物和产物.

创建一个细胞底盘(cellular chassis)用来装配合成生物学家设计的生物零部件, 并集成由此产生的部件以使其生物系统具有新的细胞功能是非常重要的. 细胞底盘必须为细胞生长及组件工作提供各种组分, 应该有各种标准的连接, 而且足够稳定以便能在工业上应用. 大肠杆菌^[34]、丙酮丁醇梭杆菌^[35]、谷氨酸棒杆菌^[36]、酿酒酵母菌^[37]和黑曲霉^[38]可以作为性能优良的微生物细胞底盘.

最小基因组构建策略是改善底盘以生产燃料和化合物的最好方法. 在原核细胞方面目前已开展了最小基因组的研究. 这包括最小支原体基因组^[39]和最小枯草芽孢杆菌基因组(约 271~410 个基因)^[40-42]. Chandonia 等^[43]将大肠杆菌基因组分为必需基因、非必需基因和未知功能基因. 同时进行了相关属必需和非必需的蛋白质比较, 并建立了必需基因数据库 PEC database (<http://shigen.lab.nig.ac.jp/ecoli/pec/>). 在此基础上共删除 16 个大片段, 累积删除了 1.37 Mb. 结果使基因组尺寸减小了 30%, 但突变株生长缓慢, 代时增加了 73%, 而且形态异常, 变长变宽. 类似研究如 Blattnerj^[44]对大肠杆菌 MG1655(4.64 Mb)基因组删除区域的选择是通过比较存在于菌株 K-12, 但在其他 5 株大肠杆菌中不存在的序列后, 获得约 100 个删除位点, 成功构建了不含可移动元件的基因组减小菌株 MDS40-43. 该简适基因组的特点为: a. 基因组减小了 15%; b. 基因组稳定性提高; c. 可表达一些毒性蛋白; d. 电转化效率提高^[45]. 这些研究表明, 通过微生物基因组大量增删基因, 可以获得具有新性能的微生物. 这为建立简适基因组、设计构建简单高效菌株奠定了合成生物学的基础. 除大肠杆菌外, 目前正在酵母、芽孢杆菌、谷氨酸棒杆菌、链霉菌等中建立了大片段基因删除技术, 可用于基因组删减.

支原体是目前发现的最小、最简单的具有自我

繁殖能力的细菌, 其基因组也是原核生物中最小的基因组. Craig Venter 的研究团队使用支原体进行了首例人造细胞的工作. 细胞转化使用的遗传物质是对蕈状支原体(*Mycoplasma mycoides*)的基因组进行系统生物学研究后, 通过化学合成得到的. 即将整个基因组截成 1 078 条 DNA 片段进行合成, 然后把这些 DNA 片段导入酵母细胞中, 在胞内拼接成约 1 Mp 的全长基因组. 这个人造基因组插入到山羊支原体(*Mycoplasma capricolum*)细胞, 产生的人造细胞表现出的特征主要是蕈状支原体的生命特性^[46].

2 合成生物学的生物技术研究应用

合成生物学是一个新型、复杂的生命系统的工程学科. 它强调学科交叉的方法设计或重构生物系统, 并用于制药业、营养保健品、特种化学品及生物燃料等方面产品的生产. 为此可以把细胞作为“可编程机器”, 设计和控制复杂的胞内及胞外活动, 使具有新功能及新结构的生物系统能够实现准确界定的工程或科学目标. 工程师和合成生物学家可以一起设计和建造新的生物组分、网络和途径, 重接线路, 重新编程构建生物. 这些工程化的生物系统在不久的将来可以制造新药物、绿色燃料以及靶向治疗超级耐药菌和癌症等疾病. 这种从基因线路、生物模块到合成途径的从头合成(*de novo*)工程模式已经开始解决关键问题, 并且正在应用实施.

2.1 生物传感器

细胞从 DNA 转录、RNA 翻译和蛋白质翻译后修饰过程中具有无数个信号转导线路来感应各种环境信号, 并作出响应. 这些线路包括精细的感应组分和过滤信号并驱动细胞反应的传感器模块. 参与转录过程的启动子、RNA 聚合酶、转录因子和其他转录组分都可以作为转录生物传感器的工程元件. 大部分合成设计主要在启动子和转录因子. 其中常用的启动子是大肠杆菌的 *lac*, *tet* 和 *ara* 操纵子.

2.1.1 转录水平生物传感器.

通过直接设计改造启动子序列可以调控生物传感器的感应和传导能力. 早期研究策略是构建诱导表达系统^[47-50]. 通过引入、删除或者改变激活子和抑制子位点调控启动子的感应程度. 合成哺乳动物反式激活系统(transactivation systems)是将环境感应的转录因子与哺乳动物反式激活区域融合在一起形成诱导物改变基因表达. 通过该策略已经合成用于

感应抗生素^[51]、群体感应分子^[52]、气体和代谢物^[53-54]以及温度变化^[55]的哺乳动物生物传感器。Fussenegger 等^[56]通过转基因设计哺乳动物线路构建了感应电信号的合成网络。

由于生物元件可能遭遇不可预测的干扰,要做到真正意义上的基因元件模块化是十分困难的,这需要对功能模块仔细解耦(decoupling)。Kobayashi^[57]使用退耦策略构建了一个在可编程模式下对信号做出反应的大肠杆菌全细胞生物反应器。在该设计中,感应模块(由环境反应启动子和相应的转录因子构成)与一个类似中央处理器的基因线路连接,大肠杆菌细胞可以对有害的内源性输入做出反应,特异性地对 DNA 损伤刺激(如:紫外照射和丝裂霉素 C)做出反应。该基因线路采用拨动开关(toggle switch)处理感应信息,当一种信号超过一定的阈值后开关从“关(OFF)”跳跃到“开(ON)”。该生物传感器从表达荧光报告基因到激活天然表型,例如生物膜的形成(表达 *traA* 基因)或细胞的自杀(表达 *ccdB* 基因),均可获得很好的输出信息。

环境因素一般都是很多信号综合在一起。Anderson 设计了一种转录 AND 开关^[58],它使单一基因线路具有多输入环境信号,从而可以获得比较真实的生物感应结果。这种生物传感器可用在很复杂的环境条件下监测特定的微环境(在工业生物反应器)的状态,例如温度、代谢水平或者细胞浓度。

2.1.2 翻译水平的生物传感器。

非编码 RNA 通过剪切 RNA,修饰核糖体 RNA 和催化生物化学反应在转录和翻译水平调控基因表达。非编码 RNA 的调控功能可应用于生物传感器。许多具有调控功能的 RNA 分子是天然的环境传感器,这些分子通过特异的序列区域调控多种模块功能。例如,核糖开关^[59]通过适体域(apptamer domains)和诱导自身 mRNA 5'UTR 构象改变结合特异性小分子配体进而调控基因表达。模块化的适体域作为感应元件已广泛用于基于 RNA 的生物传感器。构建一个完整的基于 RNA 的生物传感器需要在模块化的 RNA 分子骨架上将一个适体域(感应元件)与转录后调控域耦合在一起。

反义 RNA 是通过转录后机制调控基因表达。Bayer 和 Smolke^[60]在酵母中将核糖开关的适体与反义抑制子耦合在单个 RNA 分子上构建了反义-激活、配体-响应的核糖调控子来调控基因表达,适

体与其配体结合,诱导 RNA 传感器构象改变,使其处于开状态(ON switch);或将反义域释放出来抑制基因的翻译,这时传感器处于关状态(OFF switch)。这种生物传感器建立在配体和靶 mRNA 协调作用的基础上,与基因拨动设计类似,表现为配体阈值浓度的双向开关。在实际操作中,通过适体与反义替代耦合成新传感器需要重新筛选相容的二级结构来构建功能开关。未来该平台可以与体内适体快速筛选技术结合,产生一系列 RNA 生物传感器,表征胞内 mRNA 和代谢水平。

此外,可以通过 RNA 自剪切调控翻译,将适体域捕捉的感应信号转导出去。RNA 剪切受核酶催化,核酶具有适体域,可以对代谢做出响应。Yen^[61]利用这种天然结构优势,在报告基因的 mRNA 序列中加入了编码配体感应的核酶。当缺乏同源配体时,报告基因 mRNA 组成型自剪切会导致信号很小或者没有信号。当同源配体抑制核酶活性时, RNA 生物传感器打开。与“反义开关”类似,这些工程改造的 RNA 可以作为响应胞内代谢状态的内源传感器。

2.1.3 翻译后生物传感器。

信号转导途径中主要的感应元件是受体蛋白。受体蛋白可以在分子相互作用的水平上从头(*de novo*)设计。Looger 等^[62]利用计算机运算的方法设计了与新的靶配体结合的天然受体蛋白。研究人员从大肠杆菌的 5 种周质结合蛋白(PBP)家族入手,用新的非本源靶配体替代野生型配体,用计算的方法研究结合靶位点残基突变与配体的结合情况。通过该方法预测 TNT(三硝基甲苯,一种致癌物和炸药)、L-乳酸(医学上重要的代谢物)和血清素(与精神病学有关的物质)的新受体。预测的受体经过了体外和体内生物传感器的实验验证。

传递感应信号的合成方法需要一种可调并合理的方式。Skerker 等^[63]通过重新构建组氨酸激酶的双组分系统进行信号传递。该构建包括了基于蛋白质生物传感器的模块化方法,以及对上游信号产生的加速、延迟或者超敏反应。

2.1.4 混合方法。

将转录、翻译和翻译后合成线路综合成一种混合方法,这样可以取每种方法的优点,构建一个与天然生物一样,基于细胞层面的生物传感器。Voigt^[64-66]使用混合合成方法构建了基于大肠杆菌的光传感器,合成感应激酶使细胞能识别红光。工程化细胞可以“打印”一种设计图像,相当于生物感

光胶片. 将蓝藻膜结合的光受体与大肠杆菌胞内的组氨酸激酶融合, 诱导受光调节基因的表达. 这种细菌光传感器可以用于图像边缘检测. 将光感应与转录线路连接在一起可以进行细胞间交流(费氏弧菌的群体感应系统), 研究人员使工程化细胞感应光, 周围细胞尽管不能感应光, 但通过群体感应可以产生色素, 进而形成边缘图像.

2.2 合成生物疗法

人类一直饱受各种健康问题困扰, 如抗药性、癌症以及肥胖等等. 而药品的研发面临着多种挑战. 合成生物学在解决上述问题方面将发挥重要作用, 通过构建模块化的生物元件可以了解疾病机理, 寻找药物靶点, 进而创造新的疗法平台.

2.2.1 揭示疾病机理. 合成生物学在病理学研究方面为重构生物系统提供了框架. 该策略可以解释像丙种球蛋白缺乏症(病人不能产生成熟 B 细胞, 因此不能抵抗感染)一类的免疫缺陷疾病. 研究人员在正交环境下系统重构了人 B 细胞抗原受体(BCr)信号途径中的各种组分, 构建了一个合成试验平台^[67]. 这可以鉴定 BCr 信号激活与组装的网络拓扑特征. 为了研究疾病传染的分子基础, 重构致病病毒基因组. 例如合成并重构严重急性呼吸综合征冠状病毒(SARS)^[68]和 1918 西班牙流感病毒^[69]有助于鉴定造成致病力增加的基因突变.

2.2.2 鉴定药物靶位点. 将单个元件构建合成途径和系统可以更好地理解疾病发生的机制, 发现靶向治疗位点. 同时系统的装配合成生物学组件可以鉴定途径中每个组分的功能. Collins 等^[70]通过核糖调控子模块化来调控毒蛋白和生物网络中其他蛋白质的表达. 将核糖调控子 5'UTR 的 mRNA 序列设计成发夹结构, 隔离核糖体结合位点, 防止核糖体结合. 通过调控反式激活 RNA(将颈环解链)可以降低顺式抑制 mRNA 造成的翻译抑制. 用核糖调控子调控 CcdB(抑制 DNA 旋转酶的细菌毒蛋白)的表达, 可以更好地理解细菌细胞死亡的机理^[71]. 合成生物学的这些研究结合喹诺酮类(抑制旋转酶的抗生素)系统生物学的研究发现了所有的细菌抗生素通过刺激氧化伤害而诱导细胞死亡的共同途径^[72]. 这些工作为研究细菌对致死型刺激做出何种反应提供新视角, 为研发更有效的抗细菌类治疗铺平了道路.

2.2.3 新药研发.

使用合成生物学策略开展新药研发可以解决新药研发周期长、成本高的问题, 但需要进行全细胞

筛选检测. Fussenegger 等^[73]建立了筛选具有抗结核分枝杆菌的小分子合成平台. 乙硫异烟胺是一种治疗肺结核的药物, 主要是通过结核分枝杆菌酶 EthA 激活药效. 但由于蛋白 EthR 抑制 *ethA* 基因转录, 因此在治疗时乙硫异烟胺药物总是失效. 为解决该问题, 研究人员设计合成了一个哺乳动物的基因线路, 该线路构建了一个连有 EthR 反式作用子的报告基因来筛选鉴定 EthR 抑制子, 进而消除对乙硫异烟胺的拮抗作用^[73]. 该设计因基于病理机制构建基因工程线路, 并研发出应用于全细胞的新药, 因此该思路可以用于其他疾病的研究.

完全工程化的病毒和组织可用于研究特定病原和病理机制. 研究人员将产生抗药性的细菌装入生物膜, 使用工程化噬菌体来解决细菌产生抗药性的问题. 为了更有效地穿透生物膜, T7 噬菌体经过基因改造可以表达降解生物膜酶 dispersin B (DspB), 降解率达 99.997%^[74]. 非溶源性 M13 噬菌体通过过表达抑制子 *lexA3* 来抑制细菌 SOS DNA 损伤反应^[75]. 工程化的噬菌体提高了三类抗生素的杀死率, 在大肠杆菌感染小鼠实验中提高了杀死抗药性细菌, 降低了细胞因抗生素引起的耐药性.

而在构建工程化病毒时, 可以将其治疗活性与病理信号联系起来, 这样可以避免目前疗法不能区分癌组织和正常组织的弊端. Voigt 等^[76]利用合成生物学方法, 将细菌对癌细胞的入侵与特定环境信号连接起来. 组成型表达来自假结核耶尔森(氏)菌侵袭素(*inv*)基因, 可以诱导大肠杆菌细胞入侵人正常细胞系和癌细胞系. 为了选择性入侵癌细胞系, 研究人员将 *inv* 构建在转录操纵子的下游, 只有在癌细胞的微环境下, 才被特异的环境信号激活表达. 这些工程化细菌可以用于癌症治疗.

2.2.4 药效释放.

合成线路除了应用于上述疗法之外, 还可以用于控制药物释放、基因和代谢疗法. 控制药物释放可以更好地达到治疗疗效, 减轻副作用. 通过振荡器线路合成或者程序化时间延迟线路来调控药物释放. Collins^[77]合成研发了两种不同的基因响应用于上述研究.

基因疗法是一种很有前景的疗法, 主要应用于传统药物治疗无效的情况下, 例如遗传和代谢疾病. 合成线路为基因疗法提供控制方法. 哺乳动物细胞将合成的基因开关与转录抑制蛋白耦连, 使用 RNAi 模式抑制、调控和反向控制基因表达. 该系统^[78]可以抑制目的基因 99%以上的表达, 因此在基

因沉默中具有重要应用前景。

构建非本源途径为治疗代谢紊乱提供了一种基因疗法。Liao 等^[79]在哺乳动物小鼠肝脏细胞中引入乙醛酸分支途径研究脂肪酸代谢, 研究发现脂肪酸氧化增多, 形成了一个分支循环。喂食小鼠高脂肪食物后, 小鼠并没有引起肥胖症, 总脂、血浆甘油三酯和胆固醇水平都很低。该工作为研究代谢网络和失调提供了合成生物学研究的新模式, 为解决肥胖症提供了方法。

2.3 生物燃料, 药品和生物材料

2.3.1 构建生物合成途径。

青蒿素是青蒿产生的倍半萜烯酯的内过氧化物, 是治疗疟疾的特效药。但是该药物供应短缺, 很多病人不能得到治疗。化学合成青蒿素十分困难, 成本昂贵。合成从青蒿素的前体青蒿酸开始进行半合成, 不仅可以降低成本, 而且是个环境友好的产品策略。美国加州伯克利分校的 Keasling 实验室分别使用大肠杆菌和酵母菌为宿主, 插入了合成抗疟疾特效药青蒿素前体青蒿酸的多基因代谢途径, 实现了工业微生物发酵低成本生产青蒿素, 为合成生物学在多基因代谢途径方面的应用起到了示范作用^[80]。此工作通过对青蒿酸等多种酶及代谢途径的组装、精密调控, 对有关代谢途径重新设计, 提高了工业过程中青蒿素产量, 解决了天然或非天然代谢物大量积累对寄主的毒性问题。在大肠杆菌中, 通过对青蒿酸 (ADS) 合成酶密码子优化、共表达 SOE4 操纵子以及引入异源的酵母菌甲羟戊酸途径, 提高了青蒿酸的产量^[81]; 在酵母菌, 引入植物青蒿素的 ADS 合成酶基因, 克隆青蒿类植物转化为青蒿酸的细胞色素 P450 氧化还原酶等。改造后的酵母使青蒿酸的合成能力大大提高^[80]。此外, 紫杉醇合成的关键前体也在重组大肠杆菌中合成出来^[82]。

合成生物学技术已经在生物燃料, 特别是生物制氢以及生物柴油的研究方面得到应用。美国 Virginia Tech 教授 Zhang 等^[83]利用合成生物学原理, 用 13 个已知的酶构成了一条非天然酶催化途径, 可以将淀粉在温和的反应条件下高效低成本地制取氢气。该技术与燃料电池的集成有望解决氢气的生产、储存和运输等方面的难题, 具有生物能源的巨大潜力^[83]。美国加州洛杉矶分校的 Liao 研究室利用大肠杆菌已有的代谢途径, 向其中引入外源酶, 增加两步细菌本身不具备的生化反应, 可使氨基酸代谢中间产物合成醇, 设计高级醇的代谢电

路, 实现了异丁醇、丁醇的大肠杆菌发酵合成^[84]。在此基础上, Keasling 的研究小组又设计和构建了生产生物柴油的大肠杆菌, 开辟了微生物工程化炼制能源的新途径^[85]。

2.3.2 优化代谢途径代谢流。建立生物合成途径后, 所有元件表达需要处于最优化的代谢状态, 并产生高浓度的产物。表达代谢途径元件的标准方法就是与可调的强启动子连接, 例如来源于大肠杆菌操纵子 *tet*, *lac* 和 *ara* 的 P_{tet} , P_{lac} 和 P_{BAD} 启动子。目前, 合成生物学正致力于创建和表征更多可重复使用、调控生物元件的启动子达到调控表达的目的。

2.3.3 创建具有新功能材料。

设计线路可以感应生物反应器环境, 改变代谢阶段提高生物燃料生产效率。自动时间线路可以在一定时期关闭代谢过程。这类生物计时器可以通过模型指导的启动子工程使用基因拨动开关来完成。这些生物计时器可用于工业发酵生产乙醇过程中指导酵母细胞在特定时间絮凝, 简化细胞分离。

合成控制系统可以用于提纯合成产物。这在重组蛋白、生物塑料和其他生物材料生产中尤其重要。因为这些产物在胞内积累容易形成包涵体, 高浓度下毒害细胞。为了将重组蛛丝蛋白单体分泌到胞外, Widmaier 等^[86]构建了一个分泌系统, 可以有效地广泛地将蛋白质分泌到胞外。沙门氏菌 III 型分泌系统 (T3SS) 符合这些要求, 具有调控机制, 可以调控表达蛋白分泌到胞外。为了获得高分泌效率的重组丝蛋白, 研究人员仅需要构建一个调控线路将编码丝蛋白异源基因与内源分泌系统相连。

合成生物学也在细胞群体效应方面开展研究。Hasty 研究小组^[87]已成功合成了多细胞系统, 在细菌群体细胞内同步振荡, 完成细胞间交流和胞内信号处理模块化。Weiss^[9]完成了类似的先前研究, 建立了生物分子信号处理装置, 可以过滤从“发送”细胞传来的交流信号。这些系统有助于组织工程的构建和发展。

3 真核及光合生物合成生物学应用实例

3.1 酵母合成生物学与代谢工程

真核生物系统较原核生物要复杂得多, 对于合成生物学家们来说, 建立真核生物的合成生物学是一个很大的挑战和机遇。在真核生物基因组中存在着很多内含子, 影响基因表达, 同时具有多种不同的细胞器和蛋白在不同的区室定位, 这些加大了组件合成的难度。但是也为合成生物学提供了更多可

用的工具包和组件。酵母是工业生物技术中常用的生产平台。遗传背景清楚, 组学信息丰富, 基因工程操作成熟, 是一个安全可靠的生物系统。合成生物学研究主要集中在原核生物系统, 因此酵母合成生物学研究较滞后。这里就酵母几个关键事例作一综述。

3.1.1 生产异戊二烯类产品。

萜类 (terpenoids, 也称类异戊二烯, isoprenoids) 是一类天然宿主体内含量极低、结构复杂并具有多个手性中心的小分子物质。所有生物细胞的基本过程中都需要萜类次生代谢产物, 功能多样, 如防御、通讯、授粉和种子传播等。在这些次生代谢物中, 许多极具医学应用价值及保健前景, 如紫杉醇、青蒿素、类胡萝卜素和人参皂苷等等。合成生物学家已经将萜类生物合成途径插入大肠杆菌, 在不远的将来实现工业化规模生产^[88]。在酵母中生产异戊二烯可通过提高异戊二烯合成甲羟戊酸途径中起点物质乙酰 -COA 的供应。Shiba 等^[89]过表达酵母乙醛脱氢酶 (ALD6), 导入沙门氏菌乙酰 -COA 合成酶, 提高了碳流向甲羟戊酸途径, 提高了异戊二烯的产量。

在酵母中利用合成生物学和代谢工程策略生产医药品的经典例子就是重构了酵母的甲羟戊酸途径生产青蒿酸, Ro 等^[90]利用工程化酵母生产获得高浓度的青蒿酸。紫杉醇是紫杉次生代谢产生的双萜类物质, 广泛用做抗癌药物。但是紫杉醇类物质产量有限, 因此通过代谢工程生产高浓度的紫杉醇是一个热点研究。Dejong 等^[90]通过表达该途径中的五步反应的关键酶重建了紫杉醇的部分合成途径。但因该途径只有最初反应正确发挥功能, 因此只是积累了中间产物紫杉二烯。

研究人员在医药领域另外感兴趣的是氢化可的松的生产。通过在酵母体内表达来源植物的 1 个酶和来源于哺乳动物的 8 个酶。优化的策略是通过调整基因拷贝数目和启动子强弱来去除副反应, 平衡基因表达。面临最大的挑战就是建立这样复杂的途径会积累很多中间产物和副产品, 造成对宿主细胞的毒害。因此, 合成生物学可以提供必需工具构建调控线路, 减少这些副效应。

3.1.2 生产聚酮类产品。

聚酮是一类天然医药化合物。Mutka 等^[91]在酵母体内引入了聚酮前体——甲基丙二酰 -COA 的合成途径。另外的研究在酵母中异源表达 6- 甲基水杨酸合成酶 (6-MSAS) 基因, 合成 6- 甲基水杨酸

(6-MSA), 参与聚酮合成过程。同时共表达了磷酸泛酰巯基乙胺基转移酶 (PPTase), 用于激活 6-MSAS, 提高了 6-MSA 的浓度。

3.1.3 生产丁醇。Steen^[92]研究小组通过重建丁醇途径合成策略, 利用酵母生产 1- 丁醇, 同时评价了不同来源的酶对于丁醇的影响。酵母是研究真核生物合成生物学的首选, 但仍需要建立更广泛的启动子库、优化密码子的基因库、终止子库、整合位点库和表达载体库。这样可以从工具库中挑选元件, 组装构建新的代谢途径, 为代谢工程提供新的工具, 这样也利于工业生物技术发展。

3.2 蓝藻合成生物学

蓝藻作为最古老的原核生物出现, 在地球的碳平衡和光合放氧方面起着关键作用。蓝藻适应能力非常强, 可忍受高温、冰冻、缺氧、干涸及高盐度、强辐射等极端环境条件。近 10 年来, 蓝藻的基因工程研究迅猛发展, 目前已有 39 种蓝藻完成了基因组测序和注释。其中, 集胞藻 *Synechocystis* sp. PCC 6803 作为第一株光合生物, 在 1996 年完成测序^[93], 因此也成为研究光合作用、进行遗传改造、开展转录组学和蛋白质组学研究的模式藻株。与微藻和植物相比, 蓝藻在基因工程操作方面更容易。以下阐述了蓝藻应用于合成生物学时使用的生物元件以及基因工程操作时使用的工具和步骤。

3.2.1 蓝藻合成生物学简介。合成生物学主要是从工程的视角开展基因工程研究, 使用系统生物学方法进行生物元件的标准化, 细胞内基因和代谢网络的模块化。Agapakis 等^[94]使用光合细菌细长聚球藻 PCC 7942 作为研究进化动力学和两物种合成生物学系统的研究平台, 将蓝藻注射到斑马鱼的胚胎或者哺乳动物的巨噬细胞, 研究发现, 与大肠杆菌相比, 蓝藻对真核宿主细胞毒害性较小。工程转化鼠疫菌的侵袭素和单核细胞增生利斯特菌的李斯特溶素 O 后的细长聚球藻能够侵染哺乳动物细胞, 而且能在巨噬细胞中分裂。由此表明, 将光合细菌工程改造后侵染哺乳动物细胞质, 进而将其用于合成生物学, 该操作和应用是可能的。因此, 工程改造后侵染的光合细菌, 作为合成生物学组件 (devices) 具有很大的潜力。有关蓝藻合成生物学上述这些方面的研究仍很欠缺, 大肠杆菌中广泛应用的启动子在集胞藻 PCC 6803 中的功能表现十分不同。因此仍需开展蓝藻元件特性研究以及新生物元件的开发。

3.2.2 蓝藻底盘及生物元件。使用蓝藻作为生物底

盘, 不仅需要考虑其生长温度、利用的底物(兼养, 固氮)、盐浓度、代时, 还需要考虑基因工程操作时可利用的工具. 到目前为止, 蓝藻可用的标准化生物元件不多, 表征的就更少. 合成生物学研究主要在大肠杆菌中开展, 大肠杆菌中的生物元件(parts)可以应用于蓝藻. 但是这些元件在蓝藻中展示的特征与大肠杆菌中的表现不同, 这也再次表明生物底盘(biological chassis)对于展示元件具体特征的重要性. 在开展蓝藻合成生物学研究和基因工程操作时, 需考虑其以下特异性: 含有多拷贝的染色体(例如集胞藻 PCC 6803 拥有 12 个染色体拷贝), 在基因敲除时需要进行几轮的去除; 胞内存在类囊体膜结构, 含有高达 60% 的天线色素——藻胆蛋白, 蓝藻胞外有厚度不等的黏质鞘, 因此在提取 DNA、RNA 和蛋白质时需要采用特殊方法; 存在明显的昼夜节律, 可以将依赖时间的行为引入合成生物学基因线路研究. 目前蓝藻大部分元件都是内源性的.

3.2.3 转录调控.

蓝藻中 RNAP 核心酶的结构和 σ 因子与细菌不同. 蓝藻的 σ 因子属于 σ^{70} 家族^[95], 而细菌的 σ 因子是 σ^{54} 家族. 因蓝藻转录系统核心酶存在的差异, 大肠杆菌中常用的启动子在蓝藻中表现可能不同. Huang 等^[96]研究发现, Plac, Ptet 和 λ PR 这些常用启动子在蓝藻中活性很低, 甚至检测不到.

目前仅有少数几个启动子用于蓝藻, 而且都是蓝藻内源性启动子, 这可能会造成本底噪音很大. 由于蓝藻内源性启动子的结构尚不清楚, 所以对外源启动子研究十分有限. 因此, 需要研究清楚蓝藻内源启动子的结构后设计新的启动子, 利用不同的调控模式开展多种功能的研发, 进而应用于合成生物学.

转录因子: Wu 等^[97]比较 21 个测序的蓝藻基因组, 发现了 1 288 个可能的 transcription factors (TFs), 这为改造代谢途径提供了信息支持. 但是这些内源性 TF 表达的变化可能造成不能预知的影响. 因此使用正交(orthogonal)TFs 会更适合.

终止子: 目前在蓝藻基因工程研究中使用的终止子是噬菌体 T7 终止子和大肠杆菌的 *rrnB* 基因编码的终止子, 或者两者的结合. 但目前在蓝藻中尚未发现内源性的终止子.

3.2.4 翻译调控.

核糖体结合位点: 原核生物的核糖体在 mRNA 的结合位点(RBS)含有关键的 SD 序列

(5'-GGAGG-3'). 对集胞藻 PCC6803 所有基因(对应蛋白 > 100 个氨基酸)的 RBS 序列进行分析, 发现仅有 26% 的基因含有上述 SD 序列. Heidorn 等^[98]设计了 PCC6803 的 RBS 序列(RBS*: TAGT-GGAGGT), 并克隆了该序列和 3 个用于合成生物学中的 BioBrick RBS (BBa_B0030, BBa_B0032 和 BBa_B0034), 并检测了这 4 个序列在集胞藻 PCC6803 和大肠杆菌中的翻译效率. 研究发现, RBS* 在蓝藻中的翻译效率是 BioBrick RBS BBa_B0034 的 5 倍, 而大肠杆菌中 BBa_B0034 比 RBS* 翻译效率高 1/3. 结果再次为我们表明不同生物底盘下同一生物元件特性的巨大差异.

反义 RNA: Holtman 等^[99]用反义 RNA 抑制细长聚球藻(*Synechococcus elongatus*) PCC7942 的蛋白质翻译. 最近, Mitschke 等^[100]用不同的反义 RNA 序列研究集胞藻 PCC6803.

编码蛋白的序列——密码子和基因优化: 蓝藻与大肠杆菌相似, 很少使用稀有密码子. 但是单细胞蓝藻比丝状蓝藻使用稀有密码子频率高. 蓝藻进行基因表达研究时应优化密码子提高基因表达的效率. Lindberg 等^[101]在集胞藻 PCC6803 中异源表达来自植物的异戊二烯合成酶, 将密码子优化为蓝藻常用密码子后该酶表达量提高 10 倍.

3.2.5 翻译后调控.

降解标签 / 蛋白酶. 胞内蛋白质的积累受胞质蛋白酶的调控, 如依赖 ATP 的丝氨酸类型 Clp 蛋白酶. 在大肠杆菌中, 单体 Clp P 蛋白水解中心边缘有 HSP100 伴侣蛋白. 在蓝藻 PCC6803 中已经找到了多个 Clp P 的类似物. Heidorn 等^[102]在集胞藻 PCC6803 中使用 LVA、AAV 和 ASV 三个降解标签 (BioBricks BBa_E0432, BBa_E0434, BBa_E0436), 并验证了标签的功能. 这三个标签已用于大肠杆菌的合成生物学, 同时也可用于蓝藻振荡线路的研究.

蓝藻的基因工程和分子分析技术都已得到很好的研究. 但是蓝藻合成生物学的方法和新标准化生物元件仍需要构建和表征研究. 这包括在蓝藻底盘中应用 *in silico* 的方法, 结合可预测的计算机模型设计具有新功能的人工基因线路. 从长远看, 开发蓝藻的正交系统(orthogonal systems), 例如依赖于 T7 RNAP 的转录和建立在核糖体上的正交翻译, 可以减少与内源代谢的干扰, 做到独立于底盘进行生物元件的使用. 我们的最终目标就是开发得到一个良好的光合细胞平台进行可持续应用, 例如进

行生物修复, 生物燃料的生产以及高附加值产品的生产.

3.2.6 蓝藻合成生物学生产生物燃料.

气体生物燃料: 利用蓝藻代谢工程和基因工程生产气体燃料, 主要两大类, 一个是氢气, 另外是异戊二烯. 蓝藻胞内具有双向催化功能的氢化酶和固氮酶可以产氢. 由于氢化酶和固氮酶对氧气十分敏感, 因此研究人员通过基因工程改造的方法降低氢化酶和固氮酶对氧气的敏感程度, 提高产氢量. 还可以通过调控藻类代谢提高产氢量, 例如改善藻类呼吸代谢, 异源表达己糖吸收酶基因. Melis 等^[103]通过基因操作剪短光合色素天线, 增加光合传递链中质体醌(PQ)池提高藻类产氢能力. 这些研究为合成生物学构建蓝藻光合作用元件奠定了基础.

异戊二烯主要用于橡胶工业(如: 轮胎, 粘合剂)的生产. 生物来源的异戊二烯可以加工成航空燃料, 高分子材料等. 异戊二烯在植物体内主要通过甲羟戊酸途径合成. 蓝藻体内没有编码异戊二烯合成酶基因, 但可以通过糖醇-4-磷酸(2-C-methyl-D-erythritol-4-phosphate, MEP)合成异戊二烯前体(IPP 和 DMAPP). 利用合成生物学的方法, 将异源的异戊二烯合成酶(IspS)基因分别转入蓝藻 *Anabaena* 7120 和 *Synechocystis* PCC 6803^[104]中, 使其产生异戊二烯. 由此看出蓝藻可以作为生物底盘进行绿色生物燃料的生产.

液体燃料乙醇: 目前主要通过粮食和作物废弃物的发酵生产乙醇. 工业上主要利用酿酒酵母(*Saccharomyces cerevisiae*)和运动发酵单胞菌(*Zymomonas mobilis*)生产乙醇. 但用 *Z. mobilis* 生产乙醇存在 3 个缺点: a. 不能降解复杂碳水聚合物, 例如纤维素、半纤维素、淀粉; b. 产生很多副产物, 山梨醇、果聚糖、甘油和乙酸; c. 胞外易形成果聚糖聚合物. 合成生物学的创新方法可以借助“Photanol”方法将光合作用与发酵代谢结合成一体, 在自然转化的蓝藻 *Synechocystis* sp. PCC 6803 中将卡尔文循环的中间产物直接形成生物燃料. Deng 和 Coleman^[105]利用基因工程的方法, 将 *Z. mobilis* 编码丙酮酸脱羧酶(*pdc*)和乙醇脱氢酶(*adh*)基因构建在 pCB4 穿梭载体上, 在温度诱导型内源启动子 *rbclS* 的调控下导入蓝藻 *Synechococcus* sp. PCC 7942 表达生产乙醇. Dexter 和 Fu^[106]利用同源双交换的基因方法, 将 *Z. mobilis* 的 *pdc* 和 *adh* 基因整合到 *Synechocystis* sp. PCC 6803 染色体上, 实

现了在蓝藻中通过光合作用固定 CO₂ 将卡尔文循环的中间产物丙酮酸转化为生物燃料乙醇的过程, 避免了利用抗生素保持质粒复制带来的工业应用的限制, 对启动子进行了改进, 选用光诱导型 *psbA* II 内源强启动子, 这些改变为更好地进行工业应用奠定了基础.

液体燃料丁醇: 丁醇与乙醇相比具有热值高, 性质与汽油相似, 低气压, 疏水等优势. 丙酮丁醇梭菌(*Clostridium acetobutylicum*)通过糖发酵产丁醇, 但由于该菌培养和遗传操作都比较困难, 仅在厌氧条件下生长, 同时还产生大量的副产品如乙醇、丙酮、丁酸等, 所以研究人员寻找新的宿主细胞进行丁醇生产的研发. Liao 研究组^[107]利用代谢工程和基因工程的方法, 将乳酸乳球菌中编码酮酸脱羧酶基因(*kivd*)在 IPTG 诱导型启动子 *P_{trc}* 的调控下转入蓝藻 *Synechococcus elongatus* PCC7942 生产丁醇.

生物柴油: 利用藻类生产生物柴油是生物燃料研究中的热点. 藻类生物柴油与化石柴油性质相似, 具有相同的热值, 而且环境友好, 也不需要改造现有的柴油发动机. 但是目前还需克服几个技术瓶颈: 生长速度慢、很难大规模培养以及油脂合成过程中环境胁迫等问题. Curtiss 等^[108]利用代谢工程和基因工程的方法, 用了 6 次连续复杂基因改造蓝藻 *Synechocystis* sp. PCC 6803 使其产生 C10~C18 脂肪酸, 同将细胞壁的 S 层和肽聚糖层变得容易使脂肪酸直接分泌到胞外. 通过该研究可以看出, 如果将油脂合成的关键基因合成元件形式, 模式化, 会缩短研究周期, 简化复杂的基因改造过程, 加快研究人员在油脂领域的研究速度.

4 结 论

通过综述可以发现, 合成生物学作为后基因组时代的新兴学科, 获得了快速发展. 以“合成生物学”为关键词, 通过 Web of Science 搜索的学术论文数、专利数和网络信息可以准确反映上述趋势. 图 1 是合成生物学文献分析.

合成生物学方兴未艾, 其研究目标、对象和使用的技术, 以及潜在的应用范围还没有完全定义. 合成生物学进一步发展需要建立一些可靠的方法, 用于将小的功能线路连接起来形成具有预期功能的有序网络. 为此, 可以通过以下 4 个步骤完成设计周期, 达到与原有生物线路无缝整合对接^[107]: a. 为了减少生物元件间干扰, 以建立更大更复杂

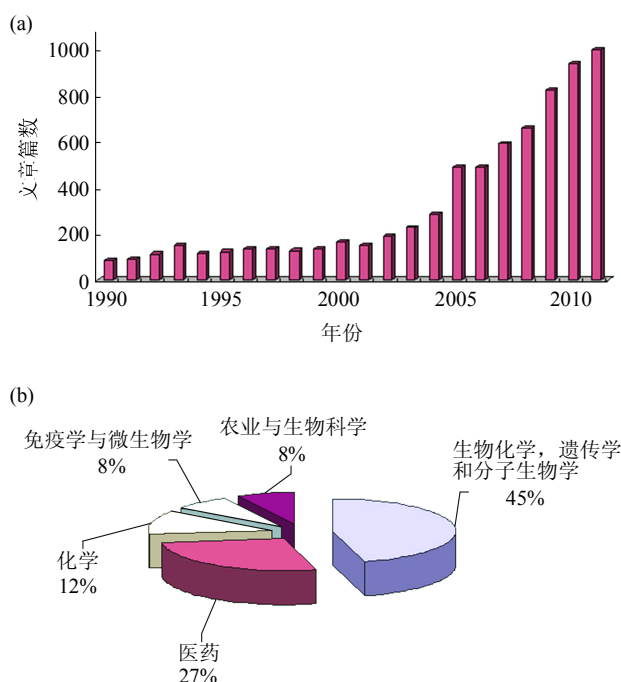


图1 合成生物学论文的年份 (a) 和领域 (b) 分布

Fig.1 Year (a) and area (b) distribution of synthetic biology papers

的生物系统, 需要扩展合成生物学工具包和通用生物元件库. 目前合成生物学组件(devices)主要建立在细菌和古菌调控元件基础之上(elements), 应加大开发真核生物元件. b. 随着网络尺寸和复杂程度的增加, 应强调合成网络的模块化与可调性能. 当完全不同的模块相连时可以简化输入 - 输出的匹配. c. 在体内应开发新探针和高通量方法测定线路动力学, 快速表征元件, 排除网络中的故障(bug). d. 开发细胞测试平台, 加快对网络问题节点的鉴定, 在构建复杂网络过程中减少易发生故障的跃变. 该平台可以是具有最小基因组的工程细胞, 也可以是装备有高等生物特异性体系的低等模式生物.

合成生物学的目标就是产生可以同时完成多项任务的生物. 合成生物学的应用将受益于生物信息学的发展. 许多与人类健康相关的问题出现在哺乳动物系统中, 因此未来更多的努力应投入到加快哺乳动物合成生物学研究, 这对解决医疗问题是十分关键的, 也包括干细胞再生与分化基因网络的合成研究.

合成生物学的一大特点是将工程系统中广泛应用的“系统论”、“控制论”和“信息论”的概念

指导“工程化”改造生物系统. 合成生物学家们认为本领域的研究已经从构建生物合成元件, 并以这些元件构建功能模块的“第一次浪潮”进入细胞作为“底盘”及“可编程”整体, 开发有效组装策略以构建大规模系统的“第二次浪潮”^[107]. 在工程科学领域, 系统工程理论已经成功应用在许多复杂的工程系统当中, 其处理流程图值得合成生物学借鉴. 我们建议合成生物学结合系统生物学开发系统生物技术应当采用的流程如图2所示. 这里系统生物学定义生物系统目标及功能要求, 合成生物学则选择装配的生物元件, 及对组件进行集成和测试. 最后在系统集成与细胞生长环境中测试验收, 完成系统生物技术应用. 为此有必要开发合成生物学辅助工具, 如全基因组预测模型、模块设计库、生物元件库、分子/细胞生物学测试平台、生物生长环境组学测试平台等(图2).

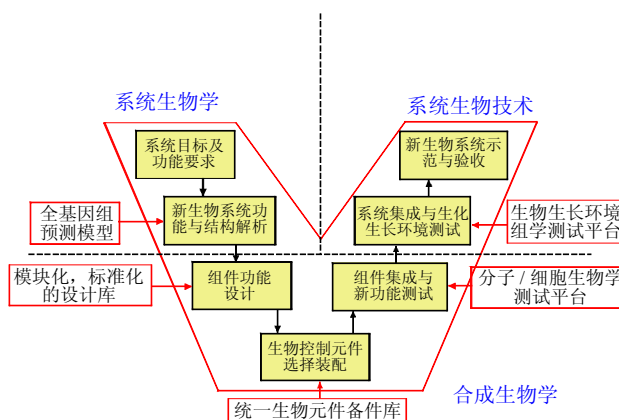


图2 合成生物学与系统生物学结合开发系统生物技术流程图

Fig.2 Flow chart for the combination of synthetic biology and system biology to develop system biotechnology

参 考 文 献

- [1] Gardner T S, Cantor C R, Collins J J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature*, 2000, **403** (6767): 339-342
- [2] Elowitz M B, Leibler S. A synthetic oscillatory network of transcriptional regulators. *Nature*, 2000, **403**(6767): 335-338
- [3] Purnick P E, Weiss R. The second wave of synthetic biology: from modules to systems. *Nature Reviews Molecular Cell Biology*, 2010, **10**(6): 410-422
- [4] Kaern M, Blake W J, Collins J J. The engineering of gene regulatory networks. *Annual Review of Biomedical Engineering*, 2003, **5**: 179-206

- [5] Atkinson M R, Savageau M A, Myers J T, *et al.* Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in *Escherichia coli*. *Cell*, 2003, **113**(5): 597–607
- [6] Klarreich E. Digital cells: Computer circuits made of genes may soon program bacteria. *Science news*, 2003, **163**(17): 267–269
- [7] Dueber J E, Yeh B J, Chak K, *et al.* Reprogramming control of an allosteric signaling switch through modular recombination. *Science*, 2003, **301**(5641): 1904–1908
- [8] Bayer T S, Smolke C D. Programmable ligand-controlled riboregulators of eukaryotic gene expression. *Nature Biotechnology*, 2005, **23**(3): 337–343
- [9] Basu S, Gerchman Y, Weiss R, *et al.* A synthetic multicellular system for programmed pattern formation. *Nature*, 2005, **434**(7037): 1130–1134
- [10] Davidson E H, Rast J P, Oliveri P, *et al.* A Provisional regulatory gene network for specification of endomesoderm in the sea urchin embryo. *Developmental Biology*, 2002, **246**(1): 162–190
- [11] Medina K L, Pongubala J M, Reddy K L, *et al.* Assembling a gene regulatory network for specification of the B cell fate. *Developmental Cell*, 2004, **7**(4): 607–617
- [12] Bulter T, Lee S G, Wong W W, *et al.* Design of artificial cell-cell communication using gene and metabolic networks. *Proc Natl Acad Sci USA*, 2004, **101**(8): 2299–2304
- [13] Basu S, Mehreja R, Thiberge S, *et al.* Spatiotemporal control of gene expression with pulse-generating networks. *Proc Natl Acad Sci USA*, 2004, **101**(17): 6355–6360
- [14] Looger L L, Dwyer M A, Smith J J, *et al.* Computational design of receptor and sensor proteins with novel functions. *Nature*, 2003, **423**(6936): 185–190
- [15] Marvin J S, Hellinga H W. Manipulation of ligand binding affinity by exploitation of conformational coupling. *Nature Structural Biology*, 2001, **8**(9): 795–798
- [16] Kinoshita K, Arnal I, Desai A, *et al.* Reconstitution of physiological microtubule dynamics using purified components. *Science*, 2001, **294**(5545): 1340–1343
- [17] Howard J, Hyman A A. Dynamics and mechanics of the microtubule plus end. *Nature*, 2003, **422**(6933): 753–758
- [18] Grill S W, Howard J, Schaffer E, *et al.* The distribution of active force generators controls mitotic spindle position. *Science*, 2003, **301**(5632): 518–521
- [19] Cowan C R, Hyman A A. Asymmetric cell division in *C. elegans*: cortical polarity and spindle positioning. *Annu Rev Cell Develop Biol*, 2004, **20**: 427–453
- [20] Tournebise R, Andersen S S, Verde F, *et al.* Distinct roles of PP1 and PP2A-like phosphatases in control of microtubule dynamics during mitosis. *Embo Journal*, 1997, **16**(18): 5537–5549
- [21] Andersen S L, Wittmann T. Toward reconstitution of *in vivo* microtubule dynamics *in vitro*. *BioEssays*, 2002, **24**(4): 305–307
- [22] Heald R, Nogales E. Microtubule dynamics. *Journal Cell Science*, 2002, **115**(1): 3–4
- [23] Pines J, Rieder C L. Re-staging mitosis: a contemporary view of mitotic progression. *Nature Cell Biology*, 2001, **3**(1): E3–E6
- [24] Chin J W, Cropp T A, Anderson J C, *et al.* An expanded eukaryotic genetic code. *Science*, 2003, **301**(5635): 964–967
- [25] Wang L, Schultz P G. Expanding the genetic code. *Chem Commun (Camb)*, 2002, **1**: 1–11
- [26] Mehl R A, Anderson J C, Santoro S W, *et al.* Generation of a bacterium with a 21 amino acid genetic code. *J American Chemical Society*, 2003, **125**(4): 935–939
- [27] Wang L, Zhang Z, Brock A, *et al.* Addition of the keto functional group to the genetic code of *Escherichia coli*. *Proc Natl Acad Sci USA*, 2003, **100**(1): 56–61
- [28] Noren C J, Anthony-Cahill S J, Suich D J, *et al.* *In-vitro* suppression of an amber mutation by a chemically aminoacylated transfer RNA prepared by runoff transcription. *Nucl Acid Res*, 1990, **18**(1): 83–88
- [29] Robertson S A, Noren C J, Anthony-Cahill S J, *et al.* The use of 5' phospho-2-deoxyribocytidylriboadenosine as a facile route to chemical aminoacylation of transfer tRNA. *Nucl Acid Res*, 1989, **17**(23): 9649–9660
- [30] Hohsaka T, Ashizuka Y, Murakami H, *et al.* Five-base codons for incorporation of nonnatural amino acids into proteins. *Nucl Acid Res*, 2001, **29**(17): 3646–3651
- [31] Hohsaka T, Ashizuka Y, Taira H, *et al.* Incorporation of nonnatural amino acids into proteins by using various four-base codons in an *Escherichia coli in vitro* translation system. *Biochemistry*, 2001, **40**(37): 11060–11064
- [32] Maier T H. Semisynthetic production of unnatural L- α -amino acids by metabolic engineering of the cysteine biosynthetic pathway. *Nature Biotechnology*, 2003, **21**(4): 422–427
- [33] Laura R J, Xueli Z, Xuan W, *et al.* Metabolic engineering for production of biorenewable fuels and chemicals: contributions of synthetic biology. *J Biomed Biotech*, 2010: 761042
- [34] Moniruzzaman M, Lai X, Ingram L O. Isolation and molecular characterization of high-performance cellobiose-fermenting spontaneous mutants of ethanologenic *Escherichia coli* KO11 containing the *Klebsiella oxytoca* casAB operon. *Appl Environ Microbiol*, 1997, **63**(12): 4633–4637
- [35] Senger R S, Papoutsakis E T. Genome-scale model for *Clostridium acetobutylicum*—part I: metabolic network resolution and analysis. *Biotech Bioeng*, 2008, **101**(5): 1036–1052
- [36] Kjeldsen K R, Nielsen J. In silico genome-scale reconstruction and validation of the *Corynebacterium glutamicum* metabolic network. *Biotech Bioengin*, 2009, **102**(2): 583–597
- [37] Forster J, Famili I, P. Fu, *et al.* Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. *Genome Research*, 2003, **13**(2): 244–253
- [38] Andersen M R, Nielsen M L, Nielsen J. Metabolic model integration of the bibliome, genome, metabolome and reactome of *Aspergillus niger*. *Molecular Systems Biology*, 2008 (4): 178
- [39] Hutchison C A, Peterson S N, Gill R, *et al.* Global transposon mutagenesis and a minimal mycoplasma genome. *Science*, 1999, **286**(5447): 2165–2169
- [40] Westers H, Dorenbos R, van Dijk J M, *et al.* Genome engineering

- reveals large dispensable regions in *Bacillus subtilis*. *Mol Biol Evol*, 2003, **20**(12): 2076–2090
- [41] Kobayashi K, Ehrlich S D, Albertini A, *et al.* Essential *Bacillus subtilis* genes. *Proc Natl Acad Sci USA*, 2003, **100**, (8): 4678–4683
- [42] Winstead E R. Another minimal genome: microbe needs just 271 genes. *Genome News Network*. 2003. http://www.genomenewsnetwork.org/articles/04_03/minimal.shtml
- [43] Chandonia J M, Konerding D E, Allen D G, *et al.* Computational structural genomics of a complete minimal organism. *Genome Informatics*, 2002, **13**: 390–391
- [44] Ji Y, Zhang B, Van S F, *et al.* Identification of critical staphylococcal genes using conditional phenotypes generated by antisense RNA. *Science*, 2001, **293**(5538): 2266–2269
- [45] Posfai G, Plunkett G, Fehér T, *et al.* Emergent Properties of Reduced-Genome *Escherichia coli*. *Science*, 2006, **312** (5776): 1044–1046
- [46] Gibson, D G, Glass, J I, Lartigue C, *et al.* Creation of a bacterial cell controlled by a chemically synthesized genome. *Science*, 2010, **329**(5987): 52–56
- [47] Brown M, Figge J, Hansen U, *et al.* Lac repressor can regulate expression from a hybrid SV40 early promoter containing a lac operator in animal cells. *Cell*, 1987, **49**(5): 603–612
- [48] Deuschle U, Pepperkok R, Wang F B, *et al.* Regulated expression of foreign genes in mammalian cells under the control of coliphage T3 RNA polymerase and lac repressor. *Proc Natl Acad Sci USA*, 1989, **86** (14): 5400–5404
- [49] Hu M C, Davidson N. The inducible lac operator-repressor system is functional in mammalian cells. *Cell*, 1987, **48**(4): 555–566
- [50] Lutz R, Bujard H. Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucl Acid Res*, 1997, **25** (6): 1203–1210
- [51] Weber W, Fu X C, Daoud-el Baba M, *et al.* Macrolide-based transgene control in mammalian cells and mice. *Nat Biotech*, 2002, **20**(9): 901–907
- [52] Neddermann P, Gargioli C, Muraglia E, *et al.* A novel, inducible, eukaryotic gene expression system based on the quorum-sensing transcription factor TraR. *EMBO Report*, 2003, **4**(2): 159–165
- [53] Weber W, Rimann M, Spielmann M, *et al.* Gas-inducible transgene expression in mammalian cells and mice. *Nat Biotech*, 2004, **22**: 1440–1444
- [54] Mullick, A. Xu Y, Warren R, *et al.* The cumate gene-switch: a system for regulated expression in mammalian cells. *BMC Biotech*, 2006, **6**: 43
- [55] Boersma M, Nieba L, Koller D, *et al.* A temperature-regulated repliconbased DNA expression system. *Nat Biotech*, 2000, **18**(4): 429–432
- [56] Weber W, Luzzi S, Fussenegger M, *et al.* A synthetic mammalian electro-genetic transcription circuit. *Nucl Acid Res*, 2009, **37**(4): e33
- [57] Kobayashi H, Kaern M, Araki M, *et al.* Programmable cells: interfacing natural and engineered gene networks. *Proc Natl Acad Sci USA*, 2004, **101**(22): 8414–8419
- [58] Anderson J C, Voigt C A, Arkin A P. Environmental signal integration by a modular AND gate. *Molecular Systems Biology*, 2007, **3**: 133
- [59] Winkler W C, Breaker R R. Regulation of bacterial gene expression by riboswitches. *Annual Review Microbiology*, 2005, **59**: 487–517
- [60] Bayer T S, Smolke C D. Programmable ligand-controlled riboregulators of eukaryotic gene expression. *Nat Biotech*, 2005, **23**(3): 337–343
- [61] Yen L, Svendsen J, Lee J S, *et al.* Exogenous control of mammalian gene expression through modulation of RNA self-cleavage. *Nature*, 2004, **431**(7007): 471–476
- [62] Looger L L, Dwyer M A, Smith J J, *et al.* Computational design of receptor and sensor proteins with novel functions. *Nature*, 2003, **423**(6936): 185–190
- [63] Skerker J M, Perchuk B S, Siryaporn A, *et al.* Rewiring the specificity of twocomponent signal transduction systems. *Cell*, 2008, **133**(6): 1043–1054
- [64] Levskaya A, Chevalier A A, Voigt C A, *et al.* Synthetic biology: engineering *Escherichia coli* to see light. *Nature*, 2005, **438**(7067): 441–442
- [65] Levskaya A, Weiner O D, Voigt C A, *et al.* Spatiotemporal control of cell signalling using a light-switchable protein interaction. *Nature*, 2009, **461**(7266): 997–1001
- [66] Tabor J J, Salis H M, Voigt C A, *et al.* A synthetic genetic edge detection program. *Cell*, 2009, **137**(7): 1272–1281
- [67] Ferrari S, Lougaris V, Caraffi S, *et al.* Mutations of the Ig β gene cause agammaglobulinemia in man. *J Exper Med*, 2007, **204**(9): 2047–2051
- [68] Becker M M, Graham R L, Donaldson E F, *et al.* Synthetic recombinant bat SARS-like coronavirus is infectious in cultured cells and in mice. *Proc Natl Acad Sci USA*, 2008, **105** (50): 19944–19949
- [69] Tumpey T M, Basler C F, Aguilar P V, *et al.* Characterization of the reconstructed 1918 Spanish influenza pandemic virus. *Science*, 2005, **310**(5745): 77–80
- [70] Dwyer D J, Kohanski M A, Collins J J, *et al.* Gyrase inhibitors induce an oxidative damage cellular death pathway in *Escherichia coli*. *Mol System Biol*, 2007, **3**: 91–99
- [71] Kohanski M A, Dwyer D J, Collins J J, *et al.* A common mechanism of cellular death induced by bactericidal antibiotics. *Cell*, 2007, **130**(5), 797–810
- [72] Kohanski M A, Dwyer D J, Collins J J, *et al.* Mistranslation of membrane proteins and two-component system activation trigger antibiotic-mediated cell death. *Cell*, 2008, **135**: 679–690
- [73] Weber W, Schoenmakers R, Fussenegger M, *et al.* A synthetic mammalian gene circuit reveals antituberculosis compounds. *Proc Natl Acad Sci USA*, 2008, **105**(29): 9994–9998
- [74] Lu T K, Collins J J. Dispersing biofilms with engineered enzymatic bacteriophage. *Proc Natl Acad Sci USA*, 2007, **104** (27): 11197–11202
- [75] Lu T K, Collins J J. Engineered bacteriophage targeting gene

- networks as adjuvants for antibiotic therapy. *Proc Natl Acad Sci USA*, 2009, **106**(12): 4629–4634
- [76] Anderson J C, Clarke E J, Voigt, C A, *et al.* Environmentally controlled invasion of cancer cells by engineered bacteria. *J Mol Biol*, 2006, **355**: 619–627
- [77] Friedland A E, Lu T K, Collins J J, *et al.* Synthetic gene networks that count. *Science*, 2009, **324**(5931): 1199–1202
- [78] Dueber J E, Yeh B J, Lim W A, *et al.* Reprogramming control of an allosteric signaling switch through modular recombination. *Science*, 2003, **301**(5641): 1904–1908
- [79] Dean J T, Tran L, Liao J C, *et al.* Resistance to diet-induced obesity in mice with synthetic glyoxylate shunt. *Cell Metabolism*, 2009, **9**(6): 525–536
- [80] Ro D K, Paradise E M, Keasling J D, *et al.* Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature*, 2006, **440**(7086): 940–943
- [81] Martin V J, Pitera D J, Keasling J D, *et al.* Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nat Biotech*, 2003, **21**(7): 796–802
- [82] Ajikumar P K, Xiao W H, Tyo K E, *et al.* Isoprenoid pathway optimization for taxol precursor overproduction in *Escherichia coli*. *Science*, 2010, **330**(6000): 70–74
- [83] Zhang Y H, Evans B R, Mielenz J R, *et al.* High-yield hydrogen production from starch and water by a synthetic enzymatic pathway. *PloS One*, 2007, **2**(5): e456
- [84] Atsumi S, Hanai T, Liao J C. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature*, 2008, **451**(7174): 86–89
- [85] Steen E J, Kang Y S, Keasling J D, *et al.* Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature*, 2010, **463**(7280): 559–562
- [86] Widmaier D M, Tullman-Ercek D, Mirsky E A, *et al.* Engineering the *Salmonella* type III secretion system to export spider silk monomers. *Mol Systems Biol*, 2009, **5**: 309
- [87] Danino T, Mondragon-Palomino O, Hasty J. A synchronized quorum of genetic clocks. *Nature*, 2010, **463**(7279): 326–330
- [88] Martin V H, Pitera D J, Withers S T, *et al.* Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nat Biotech*, 2003, **21**(7): 796–802
- [89] Shiba Y, Paradise E M, Kirby J, *et al.* Engineering of the pyruvate dehydrogenase bypass in *Saccharomyces cerevisiae* for high-level production of isoprenoids. *Metabolic Engineering*, 2007, **9** (2): 160–168
- [90] Dejong J M, Liu Y, Bollon A P, *et al.* Genetic engineering of taxol biosynthetic genes in *Saccharomyces cerevisiae*. *Biotechnology Bioengineer*, 2006, **93**(2): 212–224
- [91] Mutka S C, Bondi S M, Carney J R, *et al.* Metabolic pathway engineering for complex polyketide biosynthesis in *Saccharomyces cerevisiae*. *FEMS Yeast Research*. 2006, **6**(1): 40–47
- [92] Steen E J, Chan R, Prasad N, *et al.* Metabolic engineering of *Saccharomyces cerevisiae* for the production of n-butanol. *Microbial Cell Factories*. 2008, **7**: 36
- [93] Ikeuchi M. Complete genome sequence of a cyanobacterium *Synechocystis* sp. PCC 6803, the oxygenic photosynthetic prokaryote. *Tanpakushitsu Kakusan Koso*, 1996, **41** (16): 2579–2583
- [94] Agapakis C M, Niederholtmeyer H, Noche R R, *et al.* Towards a synthetic chloroplast. *PLoS One*, 2011, **6**(4): e18877
- [95] Imamura S, Asayama M. Sigma factors for cyanobacterial transcription. *Gene Regulation and Systems Biology*, 2009, **3**: 65–87
- [96] Huang H H, Camsund D, Lindblad P, *et al.* Design and characterization of molecular tools for a synthetic biology approach towards developing cyanobacterial biotechnology. *Nucl Acid Res*, 2010, **38**: 2577–2593
- [97] Wu J, Zhao F, Wang S, *et al.* cTFbase: A database for comparative genomics of transcription factors in cyanobacteria. *BMC Genomics*, 2007, **8**: 104
- [98] Heidorn T, Camsund D, Huang H H, *et al.* Synthetic biology in cyanobacteria engineering and analyzing novel functions. *Methods in Enzymology*, 2011, **497**: 539–579
- [99] Holtman, C K, Chen Y, Sandoval P, *et al.* High-throughput functional analysis of the *Synechococcus elongatus* PCC 7942 genome. *DNA Research*, 2005, **12**(2): 103–115
- [100] Mitschke J, Georg J, Scholz I, *et al.* An experimentally anchored map of transcriptional start sites in the model cyanobacterium *Synechocystis* sp. PCC6803. *Pro Natl Aca Sci USA*, 2011, **108**(5): 2124–2129
- [101] Lindberg P, Park S, Melis A, *et al.* Engineering a platform for photosynthetic isoprene production in cyanobacteria, using *Synechocystis* as the model organism. *Metabolic Engineering*, 2010, **12**(1): 70–79
- [102] Tetali S D, Mitra M, Melis A. Development of the lightharvesting chlorophyll antenna in the green alga *Chlamydomonas reinhardtii* is regulated by the novel *Tla1* gene. *Planta*, 2007, **225**(4): 813–829
- [103] Deng M D, Coleman J R. Ethanol synthesis by genetic engineering in cyanobacteria. *Appl Envir Microbiol*, 1999, **65**(2): 523–528
- [104] Dexter J, Fu P C. Metabolic engineering of cyanobacteria for ethanol production. *Energy & Environmental Science*, 2009, **2**: 857–864
- [105] Atsumi S, Higashide W, Liao J C. Direct photosynthetic recycling of carbon dioxide to isobutyraldehyde. *Nat Biotech*, 2009, **27**(12): 1177–1180
- [106] Liu X, Sheng J, Curtiss R 3rd. Fatty acid production in genetically modified cyanobacteria. *Proc Natl Acad Sci USA*, 2011, **108**(17): 6899–6904
- [107] Khalil A S, Collins J J. Synthetic biology: applications come of age. *Nat Rev Gen*, 2010, **11**(5): 367–379

Synthetic Biology: Its Applications in Biotechnology*

LÜ Jing¹⁾, SUN Hong-Lei²⁾, HE Hao²⁾, FU Peng-Cheng^{1)**}

¹⁾State Key Laboratory of Heavy Oil Processing, Unconventional Energy Research Center, Faculty of Chemical Engineering, China University of Petroleum, Beijing 102249, China;

²⁾PetroChina Petrochemical Research Institute, Beijing 100195, China)

Abstract Synthetic biology is an emerging biological discipline in the 21st century. It focuses on the combination of biological sciences and engineering sciences. The engineering formalism was used to design, modify and assemble cellular components into biological systems in the "bottom-up" way that comes up from "unit" to "parts" then to "System". Synthetic biology is a multidisciplinary study of molecular and cellular biology, evolutionary systematics, biochemistry, informatics, mathematics, computer and engineering. Recently, synthetic biology has been redefined as for: (I) the design and construction of new biological parts, devices, and systems that do not already exist in the nature, and (II) the re-design of existing, natural biological systems for novel function or products. As an established post-genome method, synthetic biology emphasizes on the synergistic integration of computational biology and experimental biology for the existing or new life forms. It must be noticed that the elements used in synthetic biology for new cellular structure or function may not only be genes, nucleic acids and other biological components, but also be chemical, mechanical and physical components. This article is a review for the most recently progress in the fundamental research and applications in synthetic biology. In particular, attentions were paid on synthetic biology study on the programming at the DNA level, molecular modification, metabolic pathways, regulatory networks, and industrial biotechnology and other aspects.

Key words biofuels, chassis, cyanobacteria, synthetic biology, systems biology

DOI: 10.3724/SP.J.1206.2011.00583

* This work is supported by National Basic Research Program of China (2011CB200902) and partially funded by PetroChina Petrochemical Research Institute.

**Corresponding author.

Tel: 86-10-89731283, E-mail: pengcheng@cup.edu.cn

Received: July 27, 2011 Accepted: February 13, 2012