



戴俊彪博士, 中国科学院深圳先进技术研究院研究员。1997年于南京大学获得学士学位, 2000年于清华大学获得硕士学位, 2006年于美国爱荷华州立大学(Iowa State University)获得博士学位。之后在美国约翰霍普金斯大学医学院从事博士后研究, 2011年回国入职清华大学生命科学学院。2017年加入中国科学院深圳先进技术研究院。戴俊彪实验室主要从事表观遗传学与合成生物学研究, 开发基因和基因组的合成、组装及转移技术, 通过基因组的设计构建解析基因组功能, 并进行合成生物的改造和优化利用等, 是国际合成酵母基因组计划的主要成员。已在*Cell*、*Nature*、*Science*等刊物上发表学术论文四十多篇。2017年3月与Sc2.0合作团队在*Science*杂志上以封面和专刊的形式发表了五篇染色体合成相关文章, 入选2017年中国科学十大进展、中国高等学校十大科技进展、中国科技进展十大新闻。荣获中组部首批青年“千人计划”、国家杰出青年基金、谈家桢生命科学创新奖、英国皇家学会牛顿高级学者等。

<http://isynbio.siat.ac.cn/dailab/>

合成生物学设计技术

温栾 卢俊南 钟伟 戴俊彪*

(中国科学院深圳先进技术研究院, 合成生物学研究所, 深圳 518055)

摘要 合成生物学以工程化思想为指导, 通过多学科交叉, 设计改造生命系统, 以加深对生命的认识和创造新功能, 为应对人类面临的诸多挑战提供支撑。合成生物学的精髓在于借助精妙的设计实现对生物系统的构建和模拟, 从而更好地了解生命现象。该文主要集中介绍合成生物学研究中的设计技术, 包括生物元件设计、人工基因线路设计和代谢线路设计、人工基因组设计, 归纳总结目前已有的设计技术手段和策略。

关键词 合成生物学; 生物元件设计; 蛋白质设计; 基因线路设计; 代谢途径设计; 基因组设计

Design Technology in Synthetic Biology

WEN Luan, LU Junnan, ZHONG Wei, DAI Junbiao*

(Shenzhen Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China)

Abstract Synthetic biology is an interdisciplinary research field with the goal to create new functional cells and to get deep understanding of life system through ambitious re-designing and engineering biological systems. Synthetic biology could enable us to solve human challenges in future. The key component of Synthetic Biology is to make intelligent design of the biological system, allowing the test of important biological questions. This review

深圳市合成基因组学重点实验室(批准号: ZDSYS201802061806209)、深圳市孔雀团队项目(批准号: KQTD20180413181837372)和广东省合成基因组学重点实验室(批准号: ZDSYS201802061806209)资助的课题

*通讯作者。Tel: 0755-26927780, E-mail: junbiao.dai@siat.ac.cn

This work was supported by Shenzhen Key Laboratory of Synthetic Genomics (Grant No.ZDSYS201802061806209), Shenzhen Peacock Team Project (Grant No.KQTD20180413181837372) and Guangdong Provincial Key Laboratory of Synthetic Genomics (Grant No.2019B030301006)

*Corresponding author. Tel: +86-755-26927780, E-mail: junbiao.dai@siat.ac.cn

网络出版时间: 2019-12-11 10:18:51

URL: <http://kns.cnki.net/kcms/detail/31.2035.Q.20191211.1018.008.html>

summarizes the design technologies including those used in designing new biological parts, how to design genetic circuits and metabolic pathways as well as how to design synthetic genomes in synthetic biology.

Keywords synthetic biology; biological parts design; protein design; circuit design; metabolic pathway design; genome design

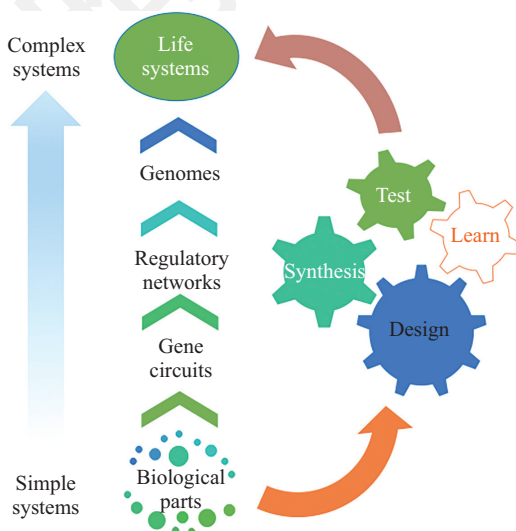
合成生物学(synthetic biology)这个名词最早出现于DNA重组技术火热的上世纪70年代^[1],在本世纪初又被重新提出^[2-3]。合成生物学是一个充满活力的新兴交叉学科,其设计理念类似于工程师建造桥梁的方式,即结合工程化的理念来理性设计改造生命体、创造新功能,用以解决人类食品缺乏、能源紧缺、环境污染、医疗健康等各方面的问題^[4]。合成生物学中的“合成”两字代表了人类对生命进行设计改造的愿望。合成生物学的工程化本质主要体现在两个方面:一是从下到上的方式,通过构建元件、组装线路、构建网络,最终获得功能复杂的生命体;二是从上到下的方式,受限于人类对生命体复杂性的认知能力,可以先通过人工合成储存所有遗传信息的基因组,再用做减法的方式来了解基因组中各个成分的功能^[5]。合成生物学的灵魂在于构思精巧的设计,借助于设计与建造生命来了解生命^[6-7]。在合成生物学研究“设计-合成-检测-学习”循环过程中,“设计”起关键的作用,好的设计既加快进程,又

节省成本。本文着重归纳总结合成生物学领域中常用的设计技术和策略,包括了合成生物学元件标准化与生物元件(BioBricks)、人工基因线路与代谢线路、细胞功能网络以及人工合成基因组的设计策略(图1)。

1 生物元件标准化与设计技术

生物元件是借用工程学的概念来描述具有最基本、最简单生物学功能的基本单元。在生命系统中,构成生命系统的天然生物元件可以是蛋白质、核酸,也可以是有功能的蛋白结构域,或是能形成特异结构的核酸序列。生物元件绝大部分是由DNA编码,因此在合成生物学领域中更加侧重于DNA水平描述生物元件。即在遗传系统中最基本、最简单的生物积块是具有特定功能的氨基酸或核苷酸序列,可以在大规模的设计中与其他元件进一步组合成具有特定生物学功能的生物学装置^[4,8]。

目前,合成生物学中使用的生物元件大部分来



合成生物学在生物元件、基因线路、信号网络和基因组几个层面上,通过“设计-合成-检测-学习”(DSTL)的循环过程来创造新功能生命体。设计是科学与艺术的结合,在科学的基础上,有目的创造新事物。设计在DSTL循环中处于重要位置,是创造新功能生命体的关键,为合成生物学精髓。

The biological engineering processes is hierarchy. The biological parts are used to construct gene circuits, in turn to make regulatory networks and more complex genomes, eventually to create artificial life system. The design-synthesis-test-learn (DSTL) cycle promotes engineering of life system, and is adopt broadly in synthetic biology. The design step plays a pivotal role in DSTL cycle and is the key for biological engineering.

图1 概括合成生物学创造新功能生命体的简易流程图

Fig.1 The Design-Synthesis-Test-Learn cycle facilitates engineering of biological system in synthetic biology

源于自然界。随着DNA测序能力的快速提高,物种序列信息以指数增长,同时利用生物信息学和系统生物学识别与鉴定以及预测生物元件的能力也显著提升,因而有越来越多的生物元件被发掘出来。对功能保守的生物元件如调控发育的一些关键转录因子,可以通过比对氨基酸序列来识别和预测;对于DNA水平上发挥调控作用的“顺式元件”的识别和预测也有较多的软件和网站可以使用。然而生物信息学预测具有一定局限性,所提供的信息需要生物学实验进行验证。

1.1 生物元件的标准化

合成生物学的工程化理念需要将生物元件标准化。为了便于借助计算机进行生命系统的模拟设计,除了生物元件的实验表征、标准化设计及功能测试,也需要同时搭建生物元件的虚拟数据库^[9-10]。2003年美国科学家建立了标准生物元件注册库(Registry of Standard Biological Parts)^[11],用于收集符合标准化条件的生物元件^[12],包括结构域、蛋白质编码序列、启动子、终止子、核糖体结合位点等非编码序列(表1)。截至2018年,注册的元件已经超过20 000个。然而这只是生物元件实体库,为了能够实现生物系统的模拟搭建和功能预测,提高设计的效率并降低成本,需要对元件进行更为精确的描述和数学建模^[10]。2008年,Goler等^[13]开发了CAD(computer-aided design)工具辅助合成生物系统的设计,但这些工具缺乏可调用的模块化和可重复用数学模型^[10]。实际上

生物元件模型数据库早有建立,如2006年Le Novere等^[14]开发的BioModels数据库,但其存储的模型过于庞大,而且在未经过修饰的情况下无法进一步组合^[10]。2010年,Cooling等^[10]建立了一个标准虚拟元件数据库(SVPs),其收集的标准元件数学模型可以被下载、扩增并可通过组合进行合成生物系统的设计和模拟(表1)。2011年,Galdzicki等^[15]构建了一个生物标准元件知识库(SBPkb),可供研究者查询、检索标准生物元件并用于合成生物学研究及应用,标准生物元件注册库的元件信息经过合成生物学公开语言语义(SBOL-semantic)框架的转换变为可以运算的信息(表1)。当前已经标准化的生物元件仅是非常少的一部分,随着合成生物学发展,越来越多的元件将被发掘,生物元件的设计及其对应数学模型的设计需要并行发展,建立完善的物理库及虚拟数据库才能发挥“标准化”所具备的优势。

1.2 生物元件的设计开发技术

1.2.1 核酸元件的设计技术

核酸的化学和生物合成相对简单,核酸序列的差异可形成如颈(stem)、环(loop)、突出(bulges)、发卡(hairpins)、假结(pseudoknots)、三聚体(triplexes)、四聚体(quadruplexes)等结构,因而可产生大量三维结构不同的分子,可形成具有特异结合功能的核酸适配体(aptamer),或者是具有酶催化活性的ribozymes和DNAzymes^[16]。Tuerk和Gold^[17]开发的指数富集系统配体进化SELEX(Systematic Evolution of Ligand

表1 生物元件标准化数据库网址和生物元件预测相关网址

Table 1 Biobricks databases and biological parts prediction, design related web sites

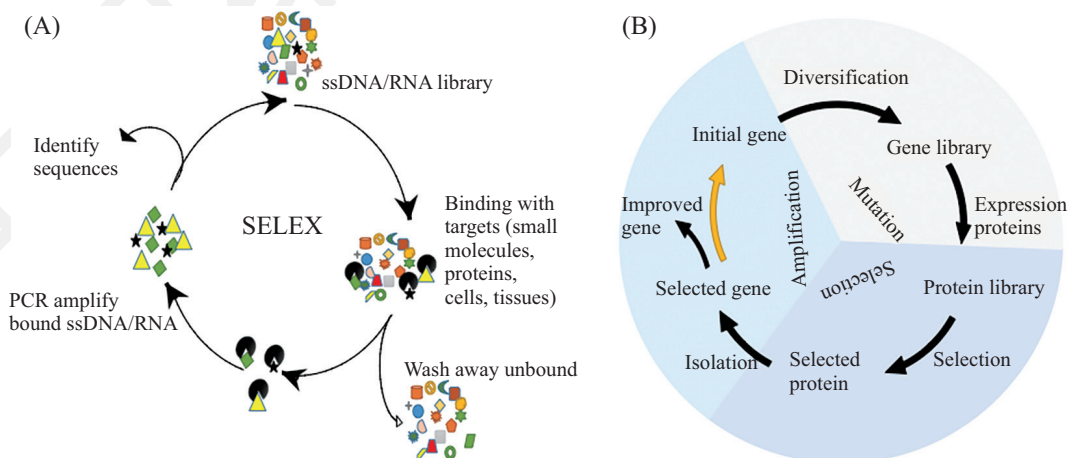
名称	网址
Name	Website
Registry of Standard Biological Parts	http://parts.igem.org/Main_Page
CellML Model Repository	http://models.cellml.org/cellml
BioJADE	http://www.biojade.org
GenoCAD	http://www.genocad.org
BioModels	https://www.ebi.ac.uk/biomodels/
SVPs	https://ico2s.org/servers/virtual_parts.html
SBPkb	https://omictools.com/sbpkb-tool
Pfam	https://pfam.xfam.org/
WWPDB	https://www.wwpdb.org/
BBCU (Bioinformatics & Biological Computing)	https://bip.weizmann.ac.il/toolbox/seq_analysis/promoters.html
CEG	http://atted.jp/help/cis_prediction.shtml
Fold X	http://foldxsuite.org.eu/products#technology_overview
Rosette	https://www.rosettacommons.org/software
Foldit	https://fold.it/portal/info/science

by Exponential Enrichment)方法是核酸功能元件的基本方法。他们利用该方法从一个寡核酸RNA文库中筛选获得噬菌体T4DNA聚合酶结合的核酸序列。同时Ellington和诺贝尔奖得主Szostak^[18]利用类似的方法从RNA文库中筛选小分子(Cibacron Blue和ReactiveBlue 4)特异结合的RNA,他们也首次将这些RNA分子称为RNA核素适配体(Aptamer)。在1992年Ellington和Szostak^[19]利用SELEX在化学合成单链DNA文库中同样筛选到与小分子特异结合的DNA核酸适配体。总的来说,SELEX筛选核酸功能元件首先要建立核酸(DNA或者RNA)随机序列文库通常 $10^{13}\sim 10^{15}$,与靶标(小分子化合物、蛋白质、细胞、活体组织)孵育结合,洗去未结合或结合弱的核酸分子;将结合的核酸分子溶解,通过PCR(DNA)或通过反转录PCR(RNA)来扩增结合靶标的核酸分子;随后可以直接通过克隆测序来获得核酸适配体序列,或者接着把获得扩增的核酸分子重新作为下一轮筛选的文库,重复筛选,经过多轮获得足够强信号时,结束循环,通过克隆测序来获得核酸功能元件序列^[20](图2A)。核酸适配体因为其多样性和结合特异性也被人们称为化学抗体,它们具有在不同环境下稳定、可在体外筛选、可通过化学合成制备修饰等优点。SELEX问世以来,在关键的文库构建、靶标选择、筛选过程、核酸分子扩增各个方面都充分发展,人们开发出众多的SELEX方法。通过该技术筛选到的核酸生物元件也被广泛应用到了基础研究、检测、诊断和治疗中,目前有众多核酸适配体药物在临床

上试验,也有获得FDA批准可用于治疗疾病的核酸生物元件^[21]。

随着合成生物学的发展,非天然碱基和修饰核酸的应用也给核酸适配体、ribozymes和DNAzymes的开发带来更广阔的发展和前景^[22],目前也有利用异源核酸(XNA)通过筛选获得具有催化活性新分子的成功案例^[23]。我国科学家在该领域也作出了重要的贡献如湖南大学谭蔚泓教授实验室^[24-25]在开发核酸适配体用于诊断治疗方面持续有重要成果发表。目前还依赖于文库的筛选来获得特异结构的序列。随着计算机运算能力的提升以及机器学习人工智能的发展,将会增加计算机预测特异核酸高级结构的能力和有目的的设计所需功能核酸大分子。

1.2.2 蛋白质生物元件的设计 蛋白质由20种天然氨基酸单元按照序列组成,理论上相同单元长度的蛋白质序列多样性显著高于核酸。蛋白质多肽的化学合成相对于核酸成本高,而且困难,同时无法像DNA那样进行体外扩增。然而自然界生命经过亿万年的进化筛选,各个物种体内都存在大量氨基酸构成的蛋白质生物元件,发挥着众多的生物学功能。在蛋白质生物元件设计方面,一种是理性设计技术,基于已知的生物元件的结构与功能之间的关系来设计,进而在生物元件编码的DNA水平上改变序列,也可以通过密码子扩展技术在特异位点插入非天然氨基酸,实现对蛋白质生物元件的改造;二是基于大规模筛选的蛋白质定向进化技术,即在生物元件编码的DNA水平上引入随机突变,通过设计



A: 指数富集系统配体进化简易流程图; B: 蛋白质定向进化筛选的简易流程图。

A: schematic diagram of systematic evolution of ligand by exponential enrichment (SELEX); B: the workflow of protein direct evolution.

图2 生物元件非理性筛选的重要方法简易流程图

Fig.2 Two key techniques to identify novel functional biological parts in synthetic biologyA

合适的筛选方法进行大规模的筛选,从而获得所需蛋白质生物元件;第三种方式是计算机(辅助)设计,结合理性设计与文库筛选,利用计算机技术的强大计算能力来增强理性设计的成分,随着蛋白质计算机设计能力的提高,目前已经有从头设计全新蛋白质出现。

一方面,理性设计技术在改造已知功能蛋白质生物元件上的使用较多,特别是针对失去活性的功能蛋白质,通常只需把活性位点的关键氨基酸突变即能达到目的,例如为提高CRISPR系统的保真性,cas9一个活性中心位点突变成nickase的设计改造^[26];工程化改造CRISPR系统的cas9使其失去核酸内切酶活性却保留双链DNA结合能力的dcas9的设计改造^[27-28]。另一方面,理性设计通过串联的方式将多个不同功能蛋白质元件或者功能结构域元件连接起来组成一个新的融合蛋白质,获得两者兼有的功能。成功的例子如基因编辑工具:锌指核酸酶(zinc finger nuclease, ZFN)就是工程化生物体内广泛存在具有DNA序列结合特异性锌指蛋白结构域与核酸内切酶Fok I结构域构建新的可识别特异序列的核酸酶^[29];类转录激活因子核酸酶[transcription activator-like (TAL) effector nuclease, TALEN]工程化利用植物病原菌黄单胞菌(*Xanthomonas*)中类转录激活因子(TAL effector, TALE)中DNA结合保守结构域与核酸内切酶Fok I结构域融合形成^[30];利用dcas9蛋白与不同功能结构域串联融合获得新生物元件,实现如DNA碱基编辑、转录调节(激活和抑制)、表观遗传修饰、RNA碱基编辑等。Lim团队^[31]通过将不同功能元件组合在一起构建了一个人工变构调节开关。最近也有课题组^[32]通过结合多个生物元件,效应结构域、RNA发卡结构域和单链RNA结合蛋白组合获得新功能元件,实现对RNA表观遗传的调控从而实现调控基因的功能。我们也利用理性设计的方法构建了一系列具有多个相互作用结构域的支架蛋白称为人工蛋白支架(artificial protein scaffold system, AProSS),并且实现了优化代谢通路的目的^[33]。这些都是理性设计非常典型的例子,对已知功能元件组合也能创造出全新的功能元件。

基于遗传密码子扩展技术的蛋白质生物元件设计,遗传密码子扩展技术是由Schultz实验室^[34]开发的,该技术可在蛋白的特异位点引入非天然氨基

酸,实现有目的改造蛋白质。简单地说,该技术通过具有生物正交性的一对氨酰-tRNA合成酶和识别密码子TAG(常用)的转运RNA(tRNA)系统,该转运RNA只能转运非天然氨基酸而不能转运任何20种天然氨基酸,并且氨酰-tRNA合成酶只能作用于此tRNA和该非天然氨基酸。因此,利用此系统可在不影响20种天然氨基酸编码的同时,将TAG密码子插入到目的基因的特异位点,可在该特异位点引入非天然氨基酸^[34]。在2008年Chin课题组^[35]在一种产甲烷古细菌(*Methanosarcina*)中发现的一对氨酰tRNA合成酶和tRNA^(PyIRS)/tRNA_{CUA}在多个物种如大肠杆菌、酵母及动物中具有很好的正交性,极大推动了基于密码子扩展技术的蛋白质非天然氨基酸引入的发展。随着该技术的发展,目前已有上百种非天然氨基酸被成功引入到蛋白的特异位点,也可以在多细胞动物如线虫和果蝇中进行非天然氨基酸渗入的蛋白质改造,所引入的非天然氨基酸的结构和功能也越来越多样化。我国科学家在该领域也作出了突出的贡献,如北京大学周德敏教授和张礼和教授团队^[36]利用该技术设计更加安全的流感病毒疫苗,中国科学院生物物理所王江运研究员团队^[37]利用该技术设计改造荧光蛋白,模拟天然光合作用光能吸收,将二氧化碳转化为一氧化碳。

蛋白质生物元件设计的另一路径是非理性的筛选技术。通过模拟自然进化选择,开发出蛋白质的定向进化方法来获得所需蛋白质元件。因此,该技术的先驱Arnold在2018年获得了诺贝尔化学奖。蛋白质定向进化的理念最初由Eigen^[38]在1984年发表的一篇理论文章提出,通过多轮进化的方法来筛选突变文库进行蛋白质工程的方案。在1991年Arnold^[39]首先发表利用蛋白质定向进化方法改造酶并获得成功。定向进化设计改进生物元件需要:(1)确定合适的初始生物元件;(2)需要在DNA水平上构建一个足够大的突变文库;(3)需要一套合适的筛选方法能有目的选择所需功能生物元件;(4)能重新多样化DNA序列作为下一轮筛选的文库;(5)还需要确定每轮筛选的合适严谨度(图2B)。蛋白质定向进化经过多年的发展,在技术上得到了多方面的优化改进,在基础研究和酶工程改造中获得广泛的应用^[40]。

基于人类现有的认识水平,完全从头设计有功能的生物元件,还具有非常大的挑战,特别是在氨基酸序列与高级结构方面。目前根据氨基酸序列预测

蛋白质结构常用的计算机软件有多种, 计算机预测蛋白质结构仍然具有很大的挑战, 目前仅限于小蛋白质结构预测, 并且准确率也不高^[41]。然而计算机辅助蛋白质设计在生物元件酶催化设计方面发挥着巨大的作用^[42]。目前在蛋白质生物元件设计方面, 根据空间结构来设计氨基酸序列却较为成功。在这方面的先驱Baker课题组^[43], 他们基于蛋白折叠总是趋于最低自由能和一些实验观察的结果, 如局部氨基酸序列并不特异决定该区域折叠的高级结构, 开发了Rosette蛋白预测程序(表1)。后来他们利用Rosetta程序, 通过人工设计的空间结构来反向预测最低自由能的氨基酸序列, 成功预测了一种完全人工设计的含97个氨基酸的蛋白质TOP7, 通过实验证实了该蛋白结构与初始设计相吻合^[44]。经过十几年在蛋白结构、蛋白质折叠、蛋白质结合以及蛋白质组装上的认识积累, 并在Rosetta能量方程和数据采集上的优化, Baker课题组^[45]成功地从头设计了许多蛋白质和蛋白质相互作用界面, 如从头设计了一个荧光蛋白^[46]、多次跨膜的蛋白质^[47]、全新的功能的IL-2^[48]。在软件方面该课题组^[49]开发了FOLDIT(表1), 一款从头设计蛋白的在线游戏, 给游戏者提供了多肽链, 游戏者需要设计氨基酸序列来形成相应的空间结构。他们还在大肠杆菌中成功从146个游戏者设计的全新蛋白中表达了56个。我国科学家在计算机辅助设计改造蛋白质元件方面也有重要的成果。中国科学院微生物研究所研究团队^[50]利用Rosetta结合高通量MD模拟方法重设计天冬氨酸酶, 获得了一系列具有绝对位置选择性与立体选择性的人工 β -氨基酸合成酶。

我们需要看到在生物元件设计上通常是多种技术结合的, 如理性设计、计算机辅助设计加上非理性的大规模筛选, 用以提高成功率。合成生物学还处在学科发展的初期, 在生物功能元件设计技术方面, 还有待于开发出更简便、更可靠的技术来从头设计新生物元件, 创造新的生物功能。

2 人工基因线路和代谢线路的设计

正如把生命体内最基本功能单元称为生物元件, 那么相应地, 多个不同生物元件组合形成具有一定信息处理能力的通路成为基因线路。生命体自身就具备复杂的信号处理能力, 因此可以把细胞抽象成一个复杂信号处理网络, 把一部分通路单独看作

是基因线路。合成生物学人工基因线路设计就是利用生物元件, 根据类似于电路工程设计的思路, 构建多元件组成信息处理功能的人工线路, 对生命体的运行过程进行干预。在基因线路设计方面也依赖于人们对元件功能的认识, 限于目前的认识水平, 在人工基因线路构建方面还只能设计一些简单信号处理的能力, 更多是改造细胞内已有线路, 实现人们所需的目的(代谢工程)。

早在2000年Collins课题组^[51]首次利用已知的生物元件在大肠杆菌中人工构建一个转录水平的双稳态开关。同时Elowitz和Leibler^[52]利用已知功能生物元件在大肠杆菌中设计了基因表达振荡器, 该线路利用3个元件间的彼此间抑制和解抑制实现了输出信号的规律振荡。借助逻辑电路的概念和思路, 过去20多年, 合成生物学家利用众多已知功能的生物元件, 通过理性设计在原核细胞、真核细胞甚至人体细胞中构建了众多基本逻辑门基因线路^[53]。我国科学家在该领域也作出了重要的贡献, 如北京大学欧阳颀课题组^[54]通过理性设计将两个逻辑“与”门、两个逻辑“或”门和一个记忆模块, 在大肠杆菌中构建成一种具有巴甫洛夫经典条件反射行为的人工基因线路。借助于计算机程序的自动化线路设计和模拟, Voigt课题组^[55]开发了人工基因线路设计的一款计算机程序, 称之为“Cello”(cellular logic)(表2)。该程序借助电子器件设计自动化可以自动将Verilog语言编写的人工线路设计转化为DNA序列, 该软件也考虑了大量生物元件的特性、生物元件组装的经验、元件的生物学限制。研究者利用该软件设计了60个人工基因线路并在大肠杆菌中正确运行了45个基因线路^[55]。设计好的人工基因线路最终需要在底盘细胞中执行, 人工基因线路与底盘细胞的正交性的好坏也决定了人工基因线路的成功与否。最近Baker团队^[56-57]报道的从头合成的全新非天然的分子开关蛋白, 为人工基因线路在不同底盘细胞中的应用提供很好的控制元件。

代谢线路的设计分为两大类。一类是异源性代谢产物合成的设计策略。对于已知的代谢通路在异源底盘细胞中构建, 需要将新的代谢通路与底盘细胞代谢网络与代谢反应数据库相结合, 通过一定的计算方法找出需要在底盘细胞中引入的新生物元件(酶基因), 从而形成新的代谢途径, 实现目标产物的合成。2012年Chatsurachai等^[58]通过逐级算法将来源

表2 人工基因线路与代谢线路设计和人工合成基因组设计工具网址

Table 2 Links for gene circuits, metabolic pathways, and synthetic genomics design tools

名称	网址
Name	Website
Cello	https://www.cellocad.org
ATLAS	https://www.metabolicatlas.org/
KEGG	https://www.genome.jp/kegg/pathway.html
BRENDA	https://www.brenda-enzymes.org/
ENZYME	https://enzyme.expasy.org/
Metacyc	https://metacyc.org/
ViennaRNA	http://rna.tbi.univie.ac.at/
Synthetic Gene Developer	http://visualgenedeveloper.net/VisualGeneDeveloper.html
Gene Designer 2.0	https://www.atum.bio/resources/tools/gene-designer
Codon Optimization OnLine (COOL)	https://omictools.com/cool-tool
BioStudio	http://23.22.137.164/gbrowse2/

于KEGG、BRENDA和ENZYME(表2)三个数据库中的异源代谢物和反应添加到三种代谢模型中以合成新的产物。结果预测出在大肠杆菌和酿酒酵母中都同时引入甘油脱氢酶和1,3-丙二醇氧化还原酶可以实现1,3-丙二醇的合成。通过这种方法,科学家在底盘生物酵母中实现了许多高附加值的产品的全新合成,如青蒿素前体青蒿酸^[59]、阿片类药物^[60]、大麻素^[61]。对于在数据库中没有的生化反应如何扩展到特定底盘细胞的代谢网络中的问题,Hatzimaniakidis研究团队^[62-63]利用化学信息学对KEGG数据库中的酶反应进行了扩展,创建了ATLAS数据库(表2),并且利用该数据库成功发现了多条合成丁酮的全新途径。第二类是内源性代谢产物的高效生产策略。内源性代谢产物的高效生产,是一个从有到优的过程。这个“优”主要体现在目标产物的高得率上。对于那些涉及副产物、多个前体、消耗能量和还原力的复杂代谢途径,人们很难通过简单的观察或手工计算得到其理论得率。2010年Bernhard Ø等^[64]提出通量平衡分析(Flux Balance Analysis, FBA)方法。利用该方法可以准确计算代谢网络中产物的理论得率,并且能获得达到理论得率的最优途径。2015年Lin等^[65]通过在大肠杆菌中利用FBA算法优化3-羟基丁酸酯(P3HB)的合成,发现利用苏氨酸循环途径可以回收甲酸和二氧化碳,进而可提高P3HB的得率。另外,通过表达异源代谢途径来减少碳损失也可实现目标产物得率的提高。2013年Liao团队^[66]在大肠杆菌中引入*Bifidobacterium adolescentis*的戊糖/己糖磷酸转酮酶,构建了非氧化糖酵解途径,实现了代谢

途径中碳的完全利用:1个葡萄糖生产3个乙酸。

在内源性或者异源性代谢产物的微生物合成中,有许多与代谢途径相关的酶分布在细胞的不同位置。如果将代谢途径相关的酶集中在细胞的某一特定场所/细胞器,这不仅可以增加特定反应场所中对应酶和代谢物的浓度,还可以方便某些特异性蛋白的功能表达,从而提高生物催化效率,防止有毒代谢物中间体的释放以及解除底物扩散等问题。例如,将代谢反应定向到线粒体、过氧化物酶体、内质网以及液泡中。2016年Yuan等^[67]利用线粒体定向信号(Cox4/Coq3),将法呢基二磷酸酯的整个代谢途径和紫穗槐二烯合成酶同时定向到线粒体中,显著提高了紫穗槐二烯的产量。

对于异源代谢产物或者内源性代谢产物的生产,采用传统的策略可以提高目标产物的得率,如过表达代谢通路中的关键生物元件(酶)或者通过替换改进的生物元件、失活或者关闭副产物途径、解除产物合成抑制等^[68]。但是,这是一种静态调控,不能在胞内或胞外的条件发生改变时对代谢流进行动态控制。采用合成生物学人工基因线路策略,即在特定的时间和细胞密度下关闭或开启基因表达,可以实现代谢流向在细胞生长的内源代谢途径和化学品生产的异源途径之间的切换。比如*HXT1*启动子的强度依赖于葡萄糖浓度:当酵母中胞内葡萄糖丰富,细胞密度低时,*HXT1*基因会高度表达;当细胞密度高,葡萄糖耗尽时,*HXT1*基因会被抑制表达。因此,可以利用*HXT1*启动子动态调控代谢途径中与细胞存活相关的基因,比如*ERG9*。2012年

Scalcinati等^[69]采用*HXT1*启动子动态控制*ERG9*表达, 可以使碳代谢流从内源性甾醇合成方向流向 α -檀香烯, 使 α -檀香烯产量提高了3.4倍。另外, 除了糖感应启动子外, 科学家还开发了基于转录因子的生物感应子。2016年David等^[70]开发了一种基于单酰辅酶A的FapR生物感应子来控制单酰辅酶A还原酶的表达, 强化单酰辅酶A的利用率, 提高了3-羟基丙酸的合成。

3 生物网络的设计(细胞重编程)

在多通路构成的信号网络设计上, 通常由于其复杂程度, 对其设计改造更多是基于非理性的筛选方法。典型的成功例子为诺贝尔奖获得者Yamanaka发现诱导胚胎干细胞重编程技术, 并在这之后的多种细胞的重编程。在真核生物中作为生物元件的转录因子通常具有强大的功能, 其往往能调节一系列下游基因表达, 造成显著的表型变化。早在上世纪80年代人们就发现, 单一转录因子MyoD可以将培养的成纤维细胞转化为肌肉细胞^[71], 因此人们意识到转录因子具有足够大的能力改变细胞的命运。Yamanaka在设计诱导多能干细胞的实验中, 有理性设计成分, 如选择只在胚胎干细胞中表达的24个转录因子作为候选基因, 有好的检测系统, 胚胎干细胞特异基因*Fbx15*启动子下游插入抗性筛选基因。24个候选基因一一尝试并不能得到诱导多能干细胞, 于是Takahashi^[72]尝试将24个候选基因全放到细胞里, 这时才发现有诱导多能干细胞出现, 随之他们开始做减法, 最后确定了4个关键的转录因子, 目前也被称为Yamanaka因子^[72]。从这个过程我们可以看出理性设计和非理性筛选相结合的成功典范。在此之后, 利用类似的思路, 人们在体外和体内发现了众多的改变细胞命运的多因子组合, 如重编程胰腺外分泌腺细胞为内分泌腺beta细胞的Ngn3、Pdx1和Mafa^[73]; 重编程成纤维细胞为心肌细胞的Gata4、Mef2c和Tbx5^[74]; 重编程成纤维细胞为神经元的Ascl1、Brn2和Myt11^[75]; 重编程成纤维细胞为肝细胞的Hnf4a和Foxa^[76]。我国科学家在该领域也作出了重要的贡献, 如北京大学邓宏魁教授和中国科学院广州生物医药与健康研究院裴端卿研究员团队^[77-78]完全利用小分子化合物诱导多能干细胞和中国科学院生物化学与细胞生物学研究所惠利建研究员团队^[76]重编程成纤维细胞为肝细胞。

4 人工合成基因组设计

生命系统的遗传信息绝大部分存储在基因组DNA上, DNA化学合成技术、DNA测序技术、DNA重组技术的进步, 为人工合成基因组奠定了坚实的基础。人类对全基因组合成的探索, 始于1970年的单个基因的人工全合成^[79], 2002年, 第一个与天然病毒几乎无功能差异的人工合成脊髓灰质炎病毒的成功^[80], 拉开了人工合成基因组序幕。2003年, 人工合成第一个噬菌体 Φ 174基因组^[81]。2008年Venter团队^[82]人工合成了582 790 bp的最小原核细胞型微生物生殖道支原体(*Mycoplasma genitalium*)的基因组。该团队在2010年合成出更大的1.08 Mb的覃状支原体(*Mycoplasma mycoides*)基因组DNA, 并将该人工合成基因组在细胞中复活, 成为第一个人造细胞^[83]。在2016年删减并合成了一个只有531 Kb的覃状支原体基因组, 比已知所有能自我存活的物种基因组都小^[84]。2011年开始真核生物酿酒酵母基因组人工合成计划(Sc2.0)的实施, 到目前为止6条酵母染色体相继人工合成^[85-93]。2018年两个课题组^[94-95]研究人员分别通过将酵母的染色体融合, 改变酵母染色体数目并维持细胞存活, 最终分别构成了只有一条或是两条巨大染色体的酵母。2019年Chin课题组^[96]报道合成了只有61个密码子的大肠杆菌基因组。

人工合成基因组, 并非简单的重新合成天然的基因组, 更多是对基因组的设计。鉴于人们目前认识的局限性, 现阶段所设计的人工基因组, 更多还是基于天然基因组的设计改造。比如在基因组中哪些基因是维持细胞功能所必需的? Venter团队^[84]通过不断重复设计-合成-测试流程, 将覃状支原体基因组中编码的901个基因删除了428个, 基因组大小从1.08 Mb缩小到531 Kb。这一基因组比自然界中已知的任何独立复制的细菌基因组都要小。然而有意思的是, 在剩下的473个基因中仍然还有149个基因功能未知。

酵母染色体人工合成计划(Sc2.0)在提出时就对合成的基因组确定了一系列设计的普遍规则, 其核心是首先维持人工合成的酵母与野生型类似的活性。具体主要包括: (1)改所有的终止密码子TAG为TAA; (2)构建包含loxPsym位点的可诱导进化系统SCRaMbLE; (3)删除一些重复序列、大量的intron区域和重新定位所有的tRNA基因^[97]。

在合成基因组中维持蛋白质基因的编码, 但是

根据密码子偏好性,设计重编了部分DNA序列,在改变DNA序列后要考虑到所转录mRNA的稳定性及形成二级复杂结构性的可能。目前基于密码子偏好的设计算法有许多如*Fop*、*CAI*、*ENc*以及*tAI*等,用于计算RNA二级结构的算法常用的有*ViennaRNA*(表2)等,用于编码区DNA的设计重编有*Synthetic Gene Developer*、*Gene Designer 2.0*和*COOL(Codon Optimization OnLine)*(表2)等^[98]。基于Sc2.0计划的实施,研究者开发了设计人工合成基因组的软件平台——*BioStudio*(表2),研究者可以根据需要,按照一定的原则进行全基因范围的重编设计^[97]。目前这些软件通常只将有限的因素作为权重考虑,因而或多或少都有一定的局限性。随着人们认识水平的加深以及计算机运算能力的提高,将DNA序列各种因素综合考虑的智能化DNA序列设计重编会将合成生物学带到新的高度。在合成基因组中通常为了区分野生型,设计加入了不影响编码的PCR标签或水印、引入或者去除一些限制性内切酶识别位点。Sc2.0中一个重要的替换是将所有编码基因的终止密码子TAG替换成了TAA,从而空出一个密码子为将来利用TAG密码子编码提供了可能。2019年Chin课题组^[96]在人工合成大肠杆菌Syn61的基因组中实现了密码子的压缩,将天然的64种密码子减少到61种,主要是丝氨酸的密码子TCG、TCA分别被替换为同义的AGC、AGT,以及终止密码子TAG全部替换为TAA,总共替换了18 218个位点。遗传密码子的压缩与重编程使非天然氨基酸的体内引入成为了可能,在各个领域有巨大的潜在应用价值。然而在Sc2.0的实施过程中,我们和其他团队都发现有些位点同义替换也会造成显著的表型变化,一方面表明了人们对于基因组的认识有限,另一方面也说明人工合成基因组是一种探索生命系统的强有力工具。在人工合成基因组过程中对于基因组中非编码区序列中的功能区域如启动子区、终止子区域的设计还处在尝试阶段,我们发现,一部分基因是可以由通用元件来替换野生型元件并且维持细胞正常的功能。在Sc2.0的设计中,在每一个非必需基因的3'末端引入一个loxPsym位点,设计了通过Cre重组酶介导的合成染色体重组和修饰进化系统(synthetic chromosome recombination and modification by loxP-mediated evolution, SCRaM-bLE),可以实现基因组的重排,在一定的压力选择下,可以有目的筛选获得所需表型。目前利用该系统,

探索基因组的排列方式都有众多报道^[99-100]。

设计广义上来讲不仅仅是技术,也是一门艺术。我们在此讨论合成生物学设计技术,就是想让读者们了解在合成生物学中设计有理性的技术成分,也有非理性的艺术成分。理性的设计常常带有严谨的逻辑美,而非理性的设计,有时会带来震撼人心灵的艺术美。合成生物学作为一门新兴的交叉学科,由改造到创造生命的转变中需要更多艺术与技术的融合,以应对人类未来面临的挑战,创造美好的未来。

致谢——

感谢课题组方圆博士、郭婷博士、赵孟楷博士、程莉博士在文献收集整理中给予的帮助,向亮博士在稿件修改中提供的宝贵意见。

参考文献 (References)

- 1 Szybalski W, Skalka A. Nobel prizes and restriction enzymes. *Gene* 1978; 4(3): 181-2.
- 2 Rawls RL. 'Synthetic biology' makes its debut. *Chem Engineer News* 2000; 78(17): 49.
- 3 Benner SA. Synthetic biology: act natural. *Nature* 2003; 421(6919): 118.
- 4 Way JC, Collins JJ, Keasling JD, Silver PA. Integrating biological redesign: where synthetic biology came from and where it needs to go. *Cell* 2014; 157(1): 151-61.
- 5 赵国屏. 合成生物学: 开启生命科学“会聚”研究新时代. *中国科学院院刊* 2018; 33(11): 1135-49.
- 6 Agapakis CM. Designing synthetic biology. *ACS Synth Biol* 2014; 3(3): 121-8.
- 7 Keller EF. Knowing as making, making as knowing: the many lives of synthetic biology. *Biol Theory* 2009; 4(4): 333-9.
- 8 Purnick PE, Weiss R. The second wave of synthetic biology: from modules to systems. *Nat Rev Mol Cell Biol* 2009; 10(6): 410-22.
- 9 Rouilly V, Canton B, Nielsen P, Kitney R. Registry of BioBricks models using CellML. *BMC Syst Biol* 2007; 1(s1): p79.
- 10 Cooling MT, Rouilly V, Misirli G, Lawson J, Yu T, Hallinan J, et al. Standard virtual biological parts: a repository of modular modeling components for synthetic biology. *Bioinformatics* 2010; 26(7): 925-31.
- 11 Peccoud J, Blauvelt MF, Cai Y, Cooper KL, Crasta O, DeLalla EC, et al. Targeted development of registries of biological parts. *PLoS One* 2008; 3(7): e2671.
- 12 Knight TF. Engineering novel life. *Mol Syst Biol* 2005; 1(1): 2005.0020.
- 13 Goler JA, Bramlett BW, Peccoud J. Genetic design: rising above the sequence. *Trends Biotechnol* 2008; 26(10): 538-44.
- 14 Le Novere N, Bornstein B, Broicher A, Courtot M, Donizelli M, Dharuri H, et al. BioModels database: a free, centralized database of curated, published, quantitative kinetic models of biochemi-

- cal and cellular systems. *Nucleic Acids Res* 2006; 34(suppl_1): D689-91.
- 15 Galdzicki M, Rodriguez C, Chandran D, Sauro HM, Gennari JH. Standard biological parts knowledgebase. *PLoS One* 2011; 6(2): e17005.
- 16 Hollenstein M. DNA catalysis: the chemical repertoire of DNAzymes. *Molecules* 2015; 20(11): 20777-804.
- 17 Tuerk C, Gold L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* 1990; 249(4968): 505-10.
- 18 Ellington AD, Szostak JW. *In vitro* selection of RNA molecules that bind specific ligands. *Nature* 1990; 346(6287): 818-22.
- 19 Ellington AD, Szostak JW. Selection *in vitro* of single-stranded DNA molecules that fold into specific ligand-binding structures. *Nature* 1992; 355(6363): 850-2.
- 20 Stoltenburg R, Reinemann C, Strehlitz B. SELEX-a (r) evolutionary method to generate high-affinity nucleic acid ligands. *Biomol Eng* 2007; 24(4): 381-403.
- 21 Zhuo Z, Yu Y, Wang M, Li J, Zhang Z, Liu J, *et al.* Recent advances in SELEX technology and aptamer applications in biomedicine. *Int J Mol Sci* 2017; 18(10): 2142.
- 22 Hamashima K, Soong YT, Matsunaga KI, Kimoto M, Hirao I. DNA sequencing method including unnatural bases for DNA aptamer generation by genetic alphabet expansion. *ACS Syn Bio* 2019; 8(6): 1401-10.
- 23 Taylor AI, Pinheiro VB, Smola MJ, Morgunov AS, Peak-Chew S, Cozens C, *et al.* Catalysts from synthetic genetic polymers. *Nature* 2015; 518(7539): 427-30.
- 24 Liang C, Guo B, Wu H, Shao N, Li D, Liu J, *et al.* Aptamer-functionalized lipid nanoparticles targeting osteoblasts as a novel RNA interference-based bone anabolic strategy. *Nat Med* 2015; 21(3): 288-94.
- 25 Teng IT, Li X, Yadikar HA, Yang Z, Li L, Lyu Y, *et al.* Identification and characterization of DNA aptamers specific for phosphorylation epitopes of tau protein. *J Am Chem Soc* 2018; 140(43): 14314-23.
- 26 Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, *et al.* Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 2013; 154(6): 1380-9.
- 27 Guilinger JP, Thompson DB, Liu DR. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat Biotechnol* 2014; 32(6): 577-82.
- 28 Tsai SQ, Wyvekens N, Khayter C, Foden JA, Thapar V, Reyon D, *et al.* Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat Biotechnol* 2014; 32(6): 569-76.
- 29 Durai S, Mani M, Kandavelou K, Wu J, Porteus MH, Chandrasegaran S. Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells. *Nucleic Acids Res* 2005; 33(18): 5978-90.
- 30 Joung JK, Sander JD. TALENs: a widely applicable technology for targeted genome editing. *Nat Rev Mol Cell Biol* 2013; 14(1): 49-55.
- 31 Dueber JE, Mirsky EA, Lim WA. Engineering synthetic signaling proteins with ultrasensitive input/output control. *Nat Biotechnol* 2007; 25(6): 660-2.
- 32 Rauch S, He E, Srien M, Zhou H, Zhang Z, Dickinson BC. Programmable RNA-guided RNAeffector proteins built from human parts. *Cell* 2019; 178(1): 122-34.
- 33 Li T, Chen X, Cai Y, Dai J. Artificial protein scaffold system (AProSS): an efficient method to optimize exogenous metabolic pathways in *Saccharomyces cerevisiae*. *Metab Eng* 2018; 49: 13-20.
- 34 Liu CC, Schultz PG. Adding new chemistries to the genetic code. *Annu Rev Biochem* 2010; 79: 413-44.
- 35 Chin JW. Expanding and reprogramming the genetic code of cells and animals. *Annu Rev Biochem* 2014; 83: 379-408.
- 36 Si L, Xu H, Zhou X, Zhang Z, Tian Z, Wang Y, *et al.* Generation of influenza A viruses as live but replication-incompetent virus vaccines. *Science* 2016; 354(6316): 1170-3.
- 37 Liu X, Kang F, Hu C, Wang L, Xu Z, Zheng D, *et al.* A genetically encoded photosensitizer protein facilitates the rational design of a miniature photocatalytic CO₂-reducing enzyme. *Nat Chem* 2018; 10(12): 1201.
- 38 Eigen M, Gardiner W. Evolutionary molecular engineering based on RNA replication. *Pure Applied Chemistry* 1984; 56(8): 967-78.
- 39 Chen K, Arnold FH. Enzyme engineering for nonaqueous solvents: random mutagenesis to enhance activity of subtilisin E in polar organic media. *Bio/Technology* 1991; 9(11): 1073.
- 40 Cobb RE, Sun N, Zhao H. Directed evolution as a powerful synthetic biology tool. *Methods* 2013; 60(1): 81-90.
- 41 Schwede T. Protein modeling: what happened to the "protein structure gap"? *Structure* 2013; 21(9): 1531-40.
- 42 Kiss G, Çelebi-Ölçüm N, Moretti R, Baker D, Houk K. Computational enzyme design. *Angew Chem* 2013; 52(22): 5700-25.
- 43 Simons KT, Kooperberg C, Huang E, Baker D. Assembly of protein tertiary structures from fragments with similar local sequences using simulated annealing and Bayesian scoring functions. *J Mol Biol* 1997; 268(1): 209-25.
- 44 Kuhlman B, Dantas G, Ireton GC, Varani G, Stoddard BL, Baker D. Design of a novel globular protein fold with atomic-level accuracy. *Science* 2003; 302(5649): 1364-8.
- 45 Huang PS, Boyken SE, Baker D. The coming of age of de novo protein design. *Nature* 2016; 537(7620): 320-7.
- 46 Dou J, Vorobieva AA, Sheffler W, Doyle LA, Park H, Bick MJ, *et al.* De novo design of a fluorescence-activating beta-barrel. *Nature* 2018; 561(7724): 485-91.
- 47 Lu P, Min D, DiMaio F, Wei KY, Vahey MD, Boyken SE, *et al.* Accurate computational design of multipass transmembrane proteins. *Science* 2018; 359(6379): 1042-6.
- 48 Silva DA, Yu S, Ulge UY, Spangler JB, Jude KM, Labao-Almeida C, *et al.* De novo design of potent and selective mimics of IL-2 and IL-15. *Nature* 2019; 565(7738): 186-91.
- 49 Koepnick B, Flatten J, Husain T, Ford A, Silva DA, Bick MJ, *et al.* De novo protein design by citizen scientists. *Nature* 2019; 570(7761): 390-4.
- 50 Carothers JM, Goler JA, Kapoor Y, Lara L, Keasling JD. Selecting RNA aptamers for synthetic biology: investigating magnesium dependence and predicting binding affinity. *Nucleic Acids Res* 2010; 38(8): 2736-47.
- 51 Gardner TS, Cantor CR, Collins JJ. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* 2000; 403(6767): 339.
- 52 Elowitz MB, Leibler S. A synthetic oscillatory network of transcriptional regulators. *Nature* 2000; 403(6767): 335.
- 53 娄春波, 杜沛, 孟凡康, 季翔宇, 张益豪. 人工基因线路的研究

- 进展和未来挑战. 中国科学院院刊 2018; 33(11): 1158-65.
- 54 Zhang H, Lin M, Shi H, Ji W, Huang L, Zhang X, *et al.* Programming a Pavlovian-like conditioning circuit in *Escherichia coli*. *Nat Commun* 2014; 5: 3102.
- 55 Nielsen AA, Der BS, Shin J, Vaidyanathan P, Paralanov V, Strychalski EA, *et al.* Genetic circuit design automation. *Science* 2016; 352(6281): aac7341.
- 56 Ng AH, Nguyen TH, Gómez-Schiavon M, Dods G, Langan RA, Boyken SE, *et al.* Modular and tunable biological feedback control using a de novo protein switch. *Nature* 2019; 572(7768): 265-9.
- 57 Langan RA, Boyken SE, Ng AH, Samson JA, Dods G, Westbrook AM, *et al.* De novo design of bioactive protein switches. *Nature* 2019; 572(7768): 205-10.
- 58 Chaturachai S, Furusawa C, Shimizu H. An in silico platform for the design of heterologous pathways in nonnative metabolite production. *BMC Bioinformatics* 2012; 13(1): 93.
- 59 Paddon CJ, Westfall PJ, Pitera DJ, Benjamin K, Fisher K, McPhee D, *et al.* High-level semi-synthetic production of the potent anti-malarial artemisinin. *Nature* 2013; 496(7746): 528.
- 60 Galanie S, Thodey K, Trenchard IJ, Interrante MF, Smolke C. Complete biosynthesis of opioids in yeast. *Science* 2015; 349(6252): 1095-100.
- 61 Luo X, Reiter MA, d'Espaux L, Wong J, Denby CM, Lechner A, *et al.* Complete biosynthesis of cannabinoids and their unnatural analogues in yeast. *Nature* 2019; 567(7746): 123-6.
- 62 Hadadi N, Hafner J, Shajkofci A, Zisaki A, Hatzimanikatis V. ATLAS of biochemistry: a repository of all possible biochemical reactions for synthetic biology and metabolic engineering studies. *ACS Syn Bio* 2016; 5(10): 1155-66.
- 63 Tokic M, Hadadi N, Ataman M, Neves D, Ebert BE, Blank LM, *et al.* Discovery and evaluation of biosynthetic pathways for the production of five methyl ethyl ketone precursors. *ACS Syn Bio* 2018; 7(8): 1858-73.
- 64 Orth JD, Thiele I, Palsson BØ. What is flux balance analysis? *Nat Biotech* 2010; 28(3): 245.
- 65 Lin Z, Zhang Y, Yuan Q, Liu Q, Li Y, Wang Z, *et al.* Metabolic engineering of *Escherichia coli* for poly(3-hydroxybutyrate) production viathreonine bypass. *Microbial Cell Factories* 2015; 14(1): 185.
- 66 Bogorad IW, Lin TS, Liao JC. Synthetic non-oxidative glycolysis enables complete carbon conservation. *Nature* 2013; 502(7473): 693.
- 67 Yuan J, Ching CB. Mitochondrial acetyl-CoA utilization pathway for terpenoid productions. *Metab Eng* 2016; 38: 303-9.
- 68 Lian J, Mishra S, Zhao H. Recent advances in metabolic engineering of *Saccharomyces cerevisiae*: new tools and their applications. *Metab Eng* 2018; 50: 85-108.
- 69 Scalcinati G, Knuf C, Partow S, Chen Y, Maury J, Schalk M, *et al.* Dynamic control of gene expression in *Saccharomyces cerevisiae* engineered for the production of plant sesquiterpene α -santalene in a fed-batch mode. *Metab Eng* 2012; 14(2): 91-103.
- 70 David F, Nielsen J, Siewers V. Flux control at the malonyl-CoA node through hierarchical dynamic pathway regulation in *Saccharomyces cerevisiae*. *ACS Syn Biol* 2016; 5(3): 224-33.
- 71 Davis RL, Weintraub H, Lassar AB. Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* 1987; 51(6): 987-1000.
- 72 Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 2006; 126(4): 663-76.
- 73 Zhou Q, Brown J, Kanarek A, Rajagopal J, Melton DA. In vivo reprogramming of adult pancreatic exocrine cells to β -cells. *Nature* 2008; 455(7213): 627.
- 74 Ieda M, Fu J-D, Delgado-Olguin P, Vedantham V, Hayashi Y, Bruneau BG, *et al.* Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. *Cell* 2010; 142(3): 375-86.
- 75 Wapinski OL, Vierbuchen T, Qu K, Lee QY, Chanda S, Fuentes DR, *et al.* Hierarchical mechanisms for direct reprogramming of fibroblasts to neurons. *Cell* 2013; 155(3): 621-35.
- 76 Huang P, He Z, Ji S, Sun H, Xiang D, Liu C, *et al.* Induction of functional hepatocyte-like cells from mouse fibroblasts by defined factors. *Nature* 2011; 475(7356): 386.
- 77 Hou P, Li Y, Zhang X, Liu C, Guan J, Li H, *et al.* Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. *Science* 2013; 341(6146): 651-4.
- 78 Cao S, Yu S, Li D, Ye J, Yang X, Li C, *et al.* Chromatin accessibility dynamics during chemical induction of pluripotency. *Cell Stem Cell* 2018; 22(4): 529-42, e5.
- 79 Agarwal KL, Buchi H, Caruthers MH, Gupta N, Khorana HG, Kleppe K, *et al.* Total synthesis of the gene for an alanine transfer ribonucleic acid from yeast. *Nature* 1970; 227(5253): 27-34.
- 80 Cello J, Paul AV, Wimmer E. Chemical synthesis of poliovirus cDNA: generation of infectious virus in the absence of natural template. *Science* 2002; 297(5583): 1016-8.
- 81 Smith HO, Hutchison CA, Pfannkoch C, Venter JC. Generating a synthetic genome by whole genome assembly: ϕ X174 bacteriophage from synthetic oligonucleotides. *Proc Natl Acad Sci USA* 2003; 100(26): 15440-5.
- 82 Gibson DG, Benders GA, Andrews-Pfannkoch C, Denisova EA, Baden-Tillson H, Zaveri J, *et al.* Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science* 2008; 319(5867): 1215-20.
- 83 Gibson DG, Glass JI, Lartigue C, Noskov VN, Chuang RY, Algire MA, *et al.* Creation of a bacterial cell controlled by a chemically synthesized genome. *Science* 2010; 329(5987): 52-6.
- 84 Hutchison CA, 3rd, Chuang RY, Noskov VN, Assad-Garcia N, Deerinck TJ, Ellisman MH, *et al.* Design and synthesis of a minimal bacterial genome. *Science* 2016; 351(6280): aad6253.
- 85 Dymond JS, Richardson SM, Coombes CE, Babatz T, Muller H, Annaluru N, *et al.* Synthetic chromosome arms function in yeast and generate phenotypic diversity by design. *Nature* 2011; 477(7365): 471.
- 86 Annaluru N, Muller H, Mitchell LA, Ramalingam S, Stracquadanio G, Richardson SM, *et al.* Total synthesis of a functional designer eukaryotic chromosome. *Science* 2014; 344(6179): 55-8.
- 87 Mercy G, Mozziconacci J, Scolari VF, Yang K, Zhao G, Thierry A, *et al.* 3D organization of synthetic and scrambled chromosomes. *Science* 2017; 355(6329): eaaf4597.
- 88 Mitchell LA, Wang A, Stracquadanio G, Kuang Z, Wang X, Yang K, *et al.* Synthesis, debugging, and effects of synthetic chromosome consolidation: synVI and beyond. *Science* 2017; 355(6329): eaaf4831.

- 89 Richardson SM, Mitchell LA, Stracquadanio G, Yang K, Dymond JS, DiCarlo JE, *et al.* Design of a synthetic yeast genome. *Science* 2017; 355(6329): 1040-4.
- 90 Shen Y, Wang Y, Chen T, Gao F, Gong J, Abramczyk D, *et al.* Deep functional analysis of synII, a 770-kilobase synthetic yeast chromosome. *Science* 2017; 355(6329): eaaf4791.
- 91 Wu Y, Li BZ, Zhao M, Mitchell LA, Xie ZX, Lin QH, *et al.* Bug mapping and fitness testing of chemically synthesized chromosome X. *Science* 2017; 355(6329): eaaf4706.
- 92 Xie ZX, Li BZ, Mitchell LA, Wu Y, Qi X, Jin Z, *et al.* "Perfect" designer chromosome V and behavior of a ring derivative. *Science* 2017; 355(6329): eaaf4704.
- 93 Zhang W, Zhao G, Luo Z, Lin Y, Wang L, Guo Y, *et al.* Engineering the ribosomal DNA in a megabase synthetic chromosome. *Science* 2017; 355(6329): eaaf3981.
- 94 Shao Y, Lu N, Wu Z, Cai C, Wang S, Zhang LL, *et al.* Creating a functional single-chromosome yeast. *Nature* 2018; 560(7718): 331-5.
- 95 Luo J, Sun X, Cormack BP, Boeke JD. Karyotype engineering by chromosome fusion leads to reproductive isolation in yeast. *Nature* 2018; 560(7718): 392-6.
- 96 Fredens J, Wang K, de la Torre D, Funke LFH, Robertson WE, Christova Y, *et al.* Total synthesis of *Escherichia coli* with a re-coded genome. *Nature* 2019; 569(7757): 514-8.
- 97 Richardson SM, Mitchell LA, Stracquadanio G, Yang K, Dymond JS, DiCarlo JE, *et al.* Design of a synthetic yeast genome. *Science* 2017; 355(6329): 1040-4.
- 98 罗周卿, 戴俊彪. 合成基因组学: 设计与合成的艺术. *生物工科学报*(Luo Zhouqing, Dai Junbiao. *Synthetic genomics: the art of design and synthesis. Chinese Journal of Biotechnology*) 2017; 33(3): 331-42.
- 99 Luo Z, Wang L, Wang Y, Zhang W, Guo Y, Shen Y, *et al.* Identifying and characterizing SCRaMbLEd synthetic yeast using ReSCuES. *Nat Commun* 2018; 9(1): 1930.
- 100 Liu W, Luo Z, Wang Y, Pham NT, Tuck L, Pérez-Pi I, *et al.* Rapid pathway prototyping and engineering using *in vitro* and *in vivo* synthetic genome SCRaMbLE-in methods. *Nat Commun* 2018; 9(1): 1936.