科学通报 2021 年 第 66 卷 第 3 期: 300 ~ 309
定量工程生物学专辑 评 述
《中国科学》杂志社 SCIENCE CHINA PRESS

自动化合成生物技术与工程化设施平台

唐婷, 付立豪, 郭二鹏, 张振坤, 王子宁, 马辰飞, 张智彧, 张建志, 黄建东*, 司同*

中国科学院深圳先进技术研究院, 深圳合成生物学创新研究院, 中国科学院定量工程生物学重点实验室, 深圳 518055

* 联系人, E-mail: jdhuang@hku.hk; tong.si@siat.ac.cn

2020-04-28 收稿, 2020-07-14 修回, 2020-07-15 接受, 2020-07-16 网络版发表

深圳市科技创新委员会项目(KQTD20150331172101 53)资助

摘要 合成生物学采用工程化设计理念, 对生物体进行有目标的设计、改造, 甚至从头合成具有特定功能的“人造生命”, 用于探索生命活动规律和进行生物技术创新. 由于生命系统高度复杂, 人工设计的合成生命体很难完全按照预期工作, 往往需要长时间的反复调谐. 目前, “试错”过程主要依靠研究者手动完成, 存在通量低、重复性差、迭代慢等局限. 针对这一难题, 自动化合成生物技术通过低成本、多循环地完成海量工程试错性实验, 提高研究通量和效率, 大幅增加实验设计的复杂度和系统性, 从而快速实现特定功能, 揭示人工生命体的设计原理. 近年来, 在合成生物学“设计-构建-测试-学习”的各个研究环节, 自动化技术正在以前所未有的速度加速发展. 与此对应, 在全球范围内已建成或在建多个大型工程化平台, 用于支撑相关研究和应用. 本文旨在对自动合成生物技术的关键要素进行总结, 并对合成生物研究基础设施的发展情况和未来方向进行讨论.

关键词 合成生物学, 自动化, 高通量, 生物铸造厂, 云端实验室, DNA组装

合成生物学“会聚”生命科学、工程学、信息科学等领域, 被认为是继“DNA双螺旋发现”和“人类基因组测序”之后的第三次生物技术革命[1]. 不同于传统生命科学从整体到局部的还原论思想, 合成生物学通过定量、设计、合成的整体论思想, 开展人工基因调控线路、合成基因组、人工细胞工厂等研究, 为认识生命本质提供了新的思维模式. 例如, 国际“基因组编写计划”(the Genome Project-Write)拟通过化学手段合成组装人类染色体所包含的全部DNA[2], 在设计安全的新型治疗细胞系、加速疫苗与药物研发、制造合成移植器官等领域具有潜在的应用价值.

在对生物体进行设计、改造乃至从头合成的过程中, 需要对大量基因元件、线路、系统进行合成与调试, 通过对细胞进行“重编程”(reprogramming)实现特定功能. 由于对合成生命体缺乏理性设计的能力, 需要进行长期、反复的人工实验试错(trial-and-error), 才能逐渐靠近预定目标. 例如, 构建人工酵母用于发酵生产抗疟疾有效成分青蒿素前体青蒿酸的研究, 耗时10年, 投入经费数千万美元[1]. 为了加速合成生物学的研究, 需要在试错过程中引入标准化实验手段, 从而高通量、低成本、多循环地完成“设计-构建-测试”的工程化闭环研发. 基于此发展的自动化合成生物技术, 围绕“如何实现设计合成可预测的生命体”这一关键科学问题, 旨在提升合成生物实验对象、方法、技术的标准化和模块化水平, 实现海量工程试错的自动化闭环运行, 不断发展理性设计合成生命系统能力. 自动化合成生物技术的研究, 不但可以快速积累大批优质基因功能模块, 建立标准化的合成生命工艺流程, 还可以获得高质量的海量实验数据, 从而采用数据驱动的方式开发并优化对合成生命进行系统设计和功能预测的计算模型.

为了促进自动化合成生物技术的发展, 现阶段国

引用格式: 唐婷, 付立豪, 郭二鹏, 等. 自动化合成生物技术与工程化设施平台. 科学通报, 2021, 66: 300–309
Tang T, Fu L H, Guo E P, et al. Automation in synthetic biology using biological foundries (in Chinese). Chin Sci Bull, 2021, 66: 300–309, doi: 10.1360/TB-2020-0498

© 2020《中国科学》杂志社

内外研究机构在世界范围内已搭建或正在建设数十个自动化合成生物铸造厂(biofoundry), 或工程化基础设施[4~6], 包括美国伊利诺伊大学的iBioFAB(Illinois Biological Foundry for Advanced Biomanufacturing)、美国能源部的Agile Biofoundry、英国爱丁堡大学的Edinburgh Genome Foundry(EGF)等. 2019年5月9日, 8个国家的16个隶属于公共研究机构的设施于在日本神户正式成立了“全球合成生物设施联盟”(Global Biofoundry Alliance, GBA)[7]. GBA的成立旨在加强设施间的协作沟通, 共同应对自动化合成生物研究的技术难题, 制定国际统一标准, 将智能制造、智能工厂的理念引入合成生物学. 其中, 中国科学院深圳先进技术研究院牵头建设的深圳合成生物研究重大科技基础设施(以下简称“大设施”)与天津大学合成生物学前沿科学中心是GBA中我国的发起单位(founding member).

本文将首先针对“设计-构建-测试”的不同环节, 简要介绍自动化合成生物技术的进展. 接着, 我们将对GBA中已投入运行的几个合成生物设施进行讨论, 并介绍深圳合成生物“大设施”的规划和进展. 最后将展望自动化合成生物技术及其设施平台的未来发展趋势.

1 自动化合成生物技术——设计

生物设计自动化(bio-design automation, BDA)是将工程化理念引入合成生物学研究的关键步骤和标志. 与合成生物研究过程的不同阶段相对应, 波士顿大学Douglas Densmore团队将现有BDA软件工具按照定义(specification)、DNA序列设计、构建、测试、学习等类别进行详细的归纳与讨论[8]. 按照合成生命体的设计流程, 伊利诺伊大学香槟分校Huimin Zhao团队[6]将BDA工具主要分为零件与系统设计、DNA组装图谱设计等两个类别. 其中, 零件与系统设计将预期实现的功能作为输入, 如催化某个生化反应的酶蛋白, 具有特定蛋白表达强度的核糖体结合位点、具有某种时空性能的基因调控网络、生产某种化合物的代谢通路等, 输出是工程DNA序列. 而DNA组装图谱设计以工程DNA序列为输入, 基于现有DNA组装方法或者基因编辑方法, 输出产生目标DNA序列所需要的零件序列和组装路径.

由于其通用性, DNA组装图谱设计工具是目前自动化合成生物设施关注的共性关键技术. 例如美国能源部的Agile BioFoundry研发了DNA自动化组装设计软件j5, 以及跨平台的自动化设备编程语言PR-PR[9], 可以基于Gibson、Golden Gate组装方法高通量设计DNA组装实验, 并编译产生应用于液体工作站、微流控等移液系统的操作指令. 美国伊利诺伊大学的iBioFAB团队研发了DNA组装软件iBioCAD(http://ibiocad.igb.illinois.edu/), 可以基于体外、体内的不同组装方法产生工程DNA的构建设计. 爱丁堡大学的EGF团队, 针对DNA序列设计、序列分析、组装质控、组装指令集产生等不同设计阶段, 开发了一系列DNA组装软件工具(https://cuba.genomefoundry.org/). 值得指出的是, 与传统的DNA组装图谱设计工具如Vector NTI等相比, 自动化合成生物研究对相关软件提出了新的需求, 如组合式平行设计数万个DNA序列、对大片段DNA(如人类染色体等)进行可视化分析与操作、通过机器学习算法根据实验结果反馈优化组装策略等, 而目前尚无软件系统可以完全满足以上需求.

2 自动化合成生物技术——构建与测试

在实验阶段, 合成生物技术主要关注的是: (1) 工程DNA的构建; (2) 工程DNA在目标底盘系统中的功能测试. 其中, 工程DNA的构建流程主要包括基因合成、扩增(PCR、RCA等)、浓度均一化、酶切、体内/外组装、提取等, 其测试流程主要包括核酸浓度分析、片段分析、qPCR、测序分析等. 工程DNA在目标底盘系统中的构建流程包括细胞转化(热击、电转)、菌落涂布、菌落挑取、细胞裂解、荧光分选等, 其测试流程包括转录/蛋白/代谢分析、光学分析(成像、流式、吸收/荧光光谱)、发酵过程放大等. 为了自动化地完成这些实验操作, 需要开发对应的硬件与软件系统.

硬件方面, 首先需要使用标准化的实验容器(如符合SBS标准的96孔、384孔微孔板等), 以及与这些容器配套进行高通量操作的仪器设备, 如具有多通道自动移液机械臂的液体工作站、微孔板离心机、分液器、封膜仪、撕膜仪、自动化震动培养箱/摇床、自动化PCR仪、高通量毛细管电泳仪、自动化菌落涂布挑选仪等(图1). 其次, 需要这些仪器设备具有硬件接口可以与机械手、传送带等自动化转运设备对接, 从而完成样品、试剂、耗材等在不同设备间按照实验流程进行传输.

软件方面, 需要有集成软件系统自动化控制实验操作的仪器设备和转运装置, 目前提供此类解决方案主要为Thermo Fisher、Beckman、Jena、HighRes等国外厂商. 同时, 需要有物料与信息管理系统对操作过程、实验结果以及物料等进行记录与调度. 这方面, 除

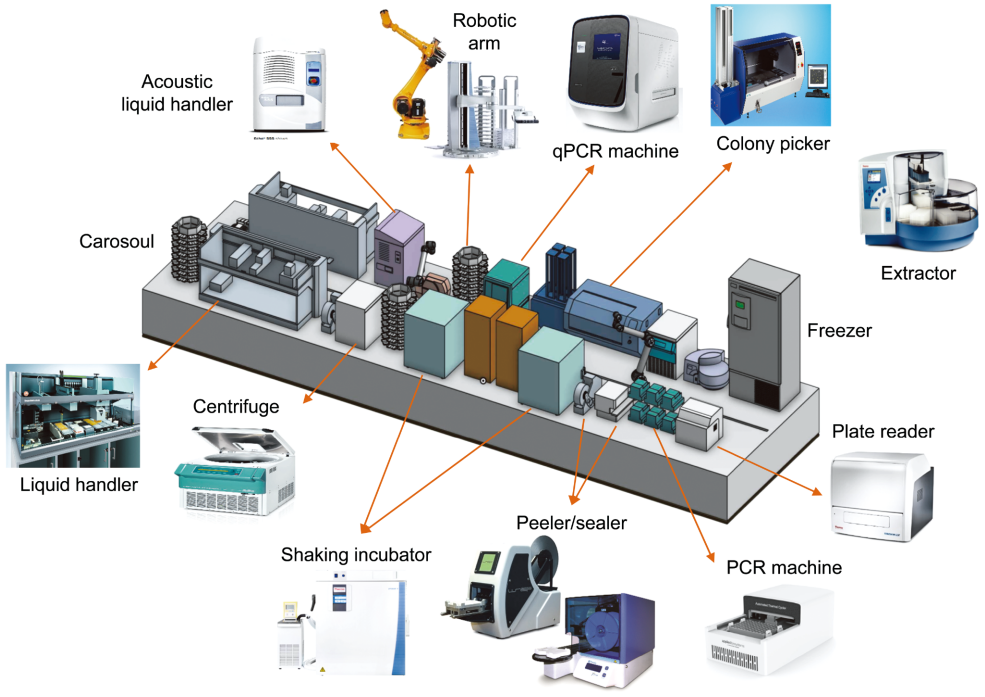


图 1 (网络版彩色)自动化合成生物设施常用仪器设备
Figure 1 (Color online) Automated instrumentation in biofoundries

了商用方案外, 利物浦大学的GeneMill平台研发了基于合成生物任务自动生成实验流程的Leaf LIMS, 华盛顿大学西雅图分校的UW BIOFAB开发了名为Aquarium的实验室操作系统(https://www.aquarium.bio/).

除软硬件系统以外, 由于目前尚无法应用物理方法直接对生物分子进行操作(与此对应, 可以应用机械手直接组装电脑、汽车等), 因此需要依赖分子生物学等手段, 开发与自动化仪器设备操作相匹配的合成生物工艺. 值得指出的是, 目前大多数生物实验过程是面向人工操作进行开发和优化的, 因此自动化工艺的开发往往是难点和关键, 需要合成生物研究者与自动化工程师密切配合, 必要时还需要开发新的仪器设备完成特殊操作. 例如, 琼脂糖凝胶电泳是分子生物研究的标准流程, 用于对不同大小的DNA片段进行分离和提纯. 但由于涉及视觉判断、半固体切胶等复杂操作, 市场上仅有Hamilton公司的Nimbus Select等少数自动化设备可以自动化完成, 且在通量、可靠性等方面仍有很大的优化空间. 下面将对工程DNA构建及其在目标底盘系统测试中的常用合成生物流程的自动化解决方案进行简要介绍.

DNA合成是合成生物学发展的重要基础, 高通量、高精度地进行DNA合成是自动化合成生物设施需要实现的核心功能之一. 聚合酶链式反应(polymerase chain reaction, PCR)局限于自然界已有DNA片段的获得, 而人工设计DNA序列的获得则需要依赖寡链核苷酸(oligonucleotide, Oligo)的从头合成与拼接[10]. 目前Oligo的自动化合成主要采用的是20世纪80年代Beaucage和Caruthers[11]开发的基于亚磷酰胺的化学合成方法, 但该方法合成的Oligo长度一般不超过200个核苷酸. 20世纪90年代发展起来的DNA芯片合成技术已在通量上有了极大提升, 单次通量可达到12000或94000种寡核苷酸簇, 且单碱基成本大幅降低. 芯片法DNA合成技术较传统柱状合成法在通量与成本上有明显优势, 但存在产出DNA质量不佳及对后续技术工艺要求极高等问题. 而且, 酶促合成法作用条件温和, 对DNA不会造成太大损伤, 有助于提高合成DNA的准确性, 可合成更长的Oligo. 酶促法中研究得最多的是末端脱氧核苷酰转移酶(TdT)介导的酶促合成法. 近期, Palluk等人[12]将dNTP通过可光诱导剪切的连接子与TdT相连形成dNTP-TdT复合物, 可在10~20 s的时间完成DNA链的延伸, 并且可以重复工作, 最终实现较长特定DNA的合成. 该方法为实现商业化的酶促DNA合成法奠定了很好的基础.

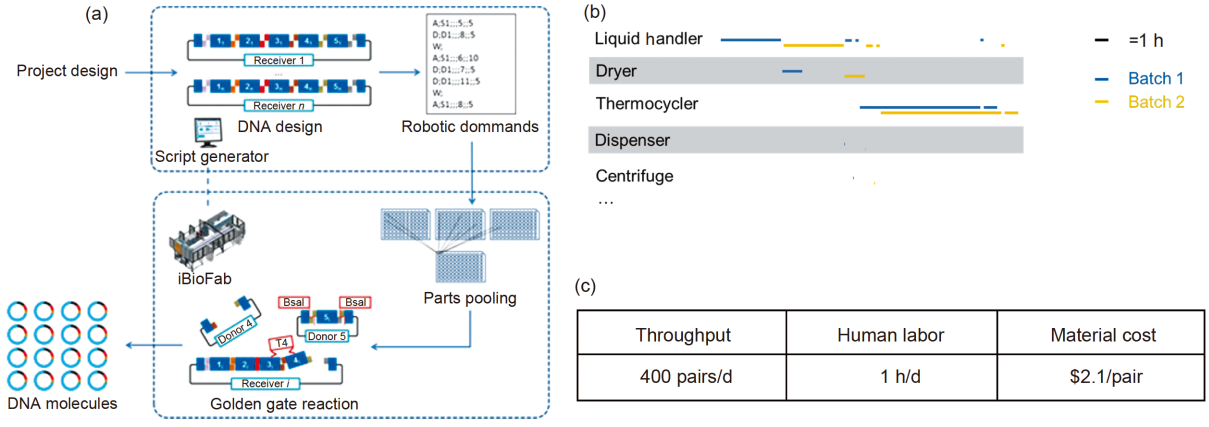
DNA组装是合成生物研究的关键共性技术. 合成

基因组研究是DNA组装技术不断快速发展的主要推动力之一. 2003年美国文特实验室首次通过全基因组合成的方式实现了人工设计合成的X174噬菌体[13], 并且历时5年设计制造了仅具有能够维持生命必需基因且能够实现自我复制的支原体[14]. 由Jef Boeke发起, 美国、中国、英国等多国科学家共同参与的Sc2.0计划, 旨在化学全合成酿酒酵母基因组. 将化学合成的Oligo组装成为Mb乃至Gb级别的大DNA片段的过程中, 存在大量重复性、平行性组装工作, 因此非常有必要引入自动化技术. 例如, Golden Gate组装法是一种依赖于IIs型限制性内切酶的DNA组装方法[15], 可以一次性组装多个片段; 并且可以通过多种限制性内切酶的替换使用, 实现标准化基因元件的多轮逐级组装(如调控和功能元件、单一表达框、多表达框代谢通路或基因线路等). 基于这一组装方法, 伊利诺伊大学Huimin Zhao团队[16]开发了类转录激活因子感受器核酸酶(transcription activator-like effector nucleases, TALEN)蛋白的自动化组装方法(图2). 利用iBioFAB平台, 实现了15个DNA片段的一步法组装; 通过酶反应、大肠杆菌转化和培养、质粒提取、酶切验证等过程的自动化运行, 每天可以构建400对编码TALEN蛋白的DNA序列, 正确组装效率达96%以上, 每对TALEN表达载体成本约2美元(商业售价的0.3%).

核酸提取是DNA构建与分析流程中的常用步骤. 自动化核酸提取的流程相对成熟化和标准化, 可以利用基于磁珠或柱式分离方法[17], 目前市场上已有Tecan、Qiagen、Thermo等公司研发的自动核酸提取设备或模块[18]. 例如, 基于Tecan液体工作站和微孔板离心机的整合系统, 我们开发了大肠杆菌质粒高通量提取方法, 每90 min可以全自动处理近200个样品.

工程DNA构建成功与否必须经过基因型验证, 且需要达到单碱基分辨率, 因为单点突变也可能对人工生命体的功能造成巨大影响. DNA测序是验证工程DNA序列是否与设计符合的终极手段, 但常用的Sanger法测序受到成本及测序长度的限制, 不适用于大规模工程DNA验证. 面对这一挑战, 可以在一代测序分析前, 首先使用分辨率较低但成本更低、通量更高的验证方法, 如结合限制性酶切和高通量毛细管电泳比对酶切指纹图谱[19], 或者利用qPCR验证目标DNA构建中是否包含全部组装片段[20]. 另一方面, 利用高通量测序技术可以进行高通量平行测序[21]. 例如, 可以同时将上千个中等大小的质粒或PCR扩增片段, 经条码序列标记混合建库, 并在测序结果分析后通过算法将混合样品的测序数据分离并与参比目标图谱进行比对, 从而大幅降低单个工程DNA的测序成本[22]. 以上工艺过程在不同的设施平台完成了自动化实现.

在底盘细胞操作方面, DNA转化是自动化合成生物过程中最耗时和最繁琐的步骤之一. 目前, 针对大肠杆菌、酿酒酵母等模式微生物, 伊利诺伊大学iBioFAB、爱丁堡大学EGF、帝国理工大学London Biofoundry等平台已经开发了基于液体工作站的化学转化过程[16,23~25]. 此外, 也有研究针对链霉菌、丝状真菌等操作更为挑战的底盘系统发展高通量DNA转化方法[26,27]. 同时, 植物系统已有报道基于原生质体法和农



Throughput	Human labor	Material cost
400 pairs/d	1 h/d	$2.1/pair

图 2 (网络版彩色)自动化TALEN蛋白表达载体构建[16]. (a) 基于Golden Gate方法的自动化DNA组装流程; (b) 自动化Golden Gate DNA组装流程的甘特图; (c) 通量和成本

Figure 2 (Color online) Fully automated synthesis of TALEN libraries[16]. (a) General workflow for the DNA assembly pipeline based on the Golden-Gate method. (b) Gantt chart for automated Golden Gate DNA assembly workflow. (c) Summary of efficiency, throughput, and cost

杆菌转染法开发了高通量DNA转化技术[28,29]. 电转方面, 利用微流体和纳米流体技术成功开发了包括哺乳动物在内的多种来源细胞的转染方法[30]. 底盘系统的自动化操作方法被应用于加快人工细胞工厂研发. 例如, 我们基于iBioFAB平台开发了酿酒酵母基因组定向进化的全自动方法(图3), 研究通量相当于1位研究人员手工操作的10倍[8]; 经过6周共计3轮的工程菌株自动化构建和筛选, 可以快速获得已报道乙酸耐受性能最高的酵母菌株[24].

工程DNA在底盘细胞内的功能测试可以针对RNA转录、蛋白质翻译、代谢等不同生命过程, 利用光学、色谱、质谱等不同分析手段, 在微孔板、摇瓶、发酵罐等不同尺度开展. 为了与自动合成生物构建过程日益增长的通量相匹配, 需要开发高通量、标准化的样品前处理技术. 例如, 利用液体工作站、微流控等技术, 结合条码序列标记, 可以平行对上千个样品(如单细胞等)进行RNA测序分析[31,32]. 伊利诺伊大学和中国科学院深圳先进院团队利用基质辅助激光解析电离-时间飞行质谱, 可以对微生物菌落的化学成分进行快速定性和定量分析(2 s/样品)[33,34]. 另一方面, 光学手段提供了对人工生命体进行原位、活体、动态观察的能力. 例如, 高内涵成像筛选与高通量合成生物改造的结合, 可以为系统性理解复杂生命过程提供新的研究手段[35]; 而在线光学探针在发酵过程的参数控制和优化方面发挥着越来越重要的作用[36].

基于以上自动化构建与测试技术, 可以产生高通量、标准化的实验数据, 用于完成合成生物"设计-构建-测试-学习"的闭环. 如图4所示, Huimin Zhao团队[37]将iBioFAB自动化平台与机器学习算法相结合, 实现了番茄红素生物合成途径的快速迭代优化. 具体而言, 构建机器学习模型, 以酶蛋白表达的启动子强度为变量、以番茄红素产量为目标函数, 生成实验设计; 由iBioFAB平台自动化完成通路构建、底盘转化与产物分析等步骤, 依据实验结果更新预测模型, 并基于贝叶斯优化算法生成下一轮的实验设计. 已有综述文章对合成生物研究的"学习"环节进行了讨论.

3 自动化合成生物设施平台

围绕自动化合成生物技术, 国内外研究机构正在大力发展相关设施平台, 成立了旨在加强协作沟通、共同制定行业标准的国际组织GBA. 下面对GBA中具有代表性的自动化合成生物设施进行介绍.

伊利诺伊大学的iBioFAB和爱丁堡大学的EGF两个设施均采用全自动化的整合方式, 且都由Thermo公司提供集成方案. 其中, iBioFAB系统由两部分组成: 一部分是支撑"设计-构建-测试"自动化研究的软件设计框架, 另一部分是可自动化进行DNA克隆、组装、异源表达以及产物检测的集成机器人系统. 这个机器人系统经过了3次升级, 第一个版本于2014年搭建完成, 包含1个具有6个自由度的机械臂, 可以在5 m长的轨道上行进, 在平台的20台(套)设备间传送物料, 同时具有1个样品跟踪系统(图5). 第一代iBioFAB实现的合成生物工艺包括基于Golden Gate方法的自动化DNA组装[16]和酿酒酵母全自动基因组定向进化[24]. 第二代iBioFAB增加1个具有4个自由度的机械臂、静置培养箱等用于开发自动化孵育蜜蜂幼虫的工艺. 第三代iBioFAB整合纳升移液工作站、自动毛细管电泳仪、高内涵显微镜等设备, 同时增加了生物安全层流罩, 提

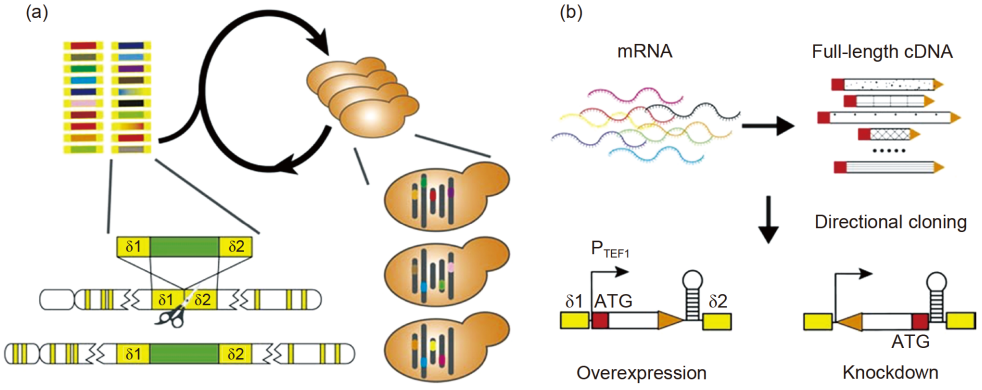


图 3 (网络版彩色)基于iBioFAB的自动化酵母基因组工程[24]. (a) 酵母基因组定向进化方法设计; (b) 流程图
Figure 3 (Color online) Automated yeast genome engineering using iBioFAB[24]. (a) Design of a directed genome evolution workflow in yeast; (b) process flow diagram

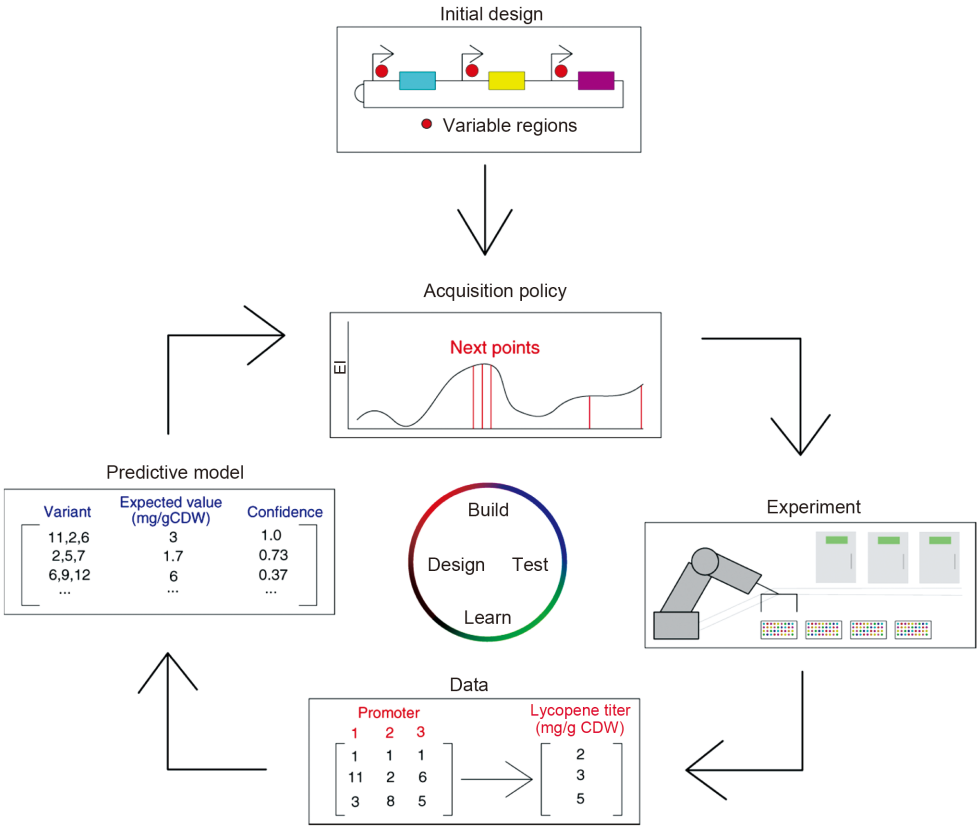


图 4 (网络版彩色)机器学习方法指导下的自动化合成生物研究流程[37]
Figure 4 (Color online) Overall workflow of automated flux balance of biosynthetic pathways guided by machine learning[37]

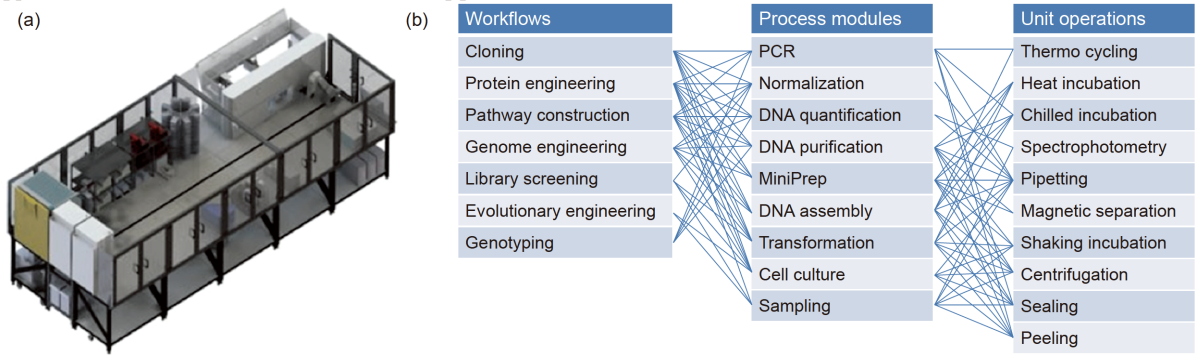


图 5 (网络版彩色)伊利诺伊大学iBioFAB平台[24]. (a) 硬件布局; (b) 工作流程、处理模块和单元操作
Figure 5 (Color online) Illinois Biological Foundry for Advanced Biomanufacturing (iBioFAB)[24]. (a) Physical layout; (b) workflows, process modules, and unit operations

升了对于植物、哺乳动物等底盘系统的兼容能力.

英国爱丁堡大学的EGF设施主要关注大DNA片段的自动化组装. 该平台也是目前学术界已建成最大规模的整合型合成生物平台(图6)[38], 包含3个机械手臂. 围绕大DNA片段与组合DNA文库的设计与构建, EGF已开发了20余种开源的软件工具(图6(c)), 每周可以处理超过2000个DNA组装反应, 通量相当于1位研究人员手工操作的20倍[7]. 目前可以操作的底盘系统包括大肠杆菌、酿酒酵母、蓝细菌等. 该平台还集成了一系列细胞功能测试模块, 进行自动化qPCR分析、高通量微生物发酵及动力学分析等. 其他规划中的整合性设施平台还包括波士顿大学的DAMP Lab等.

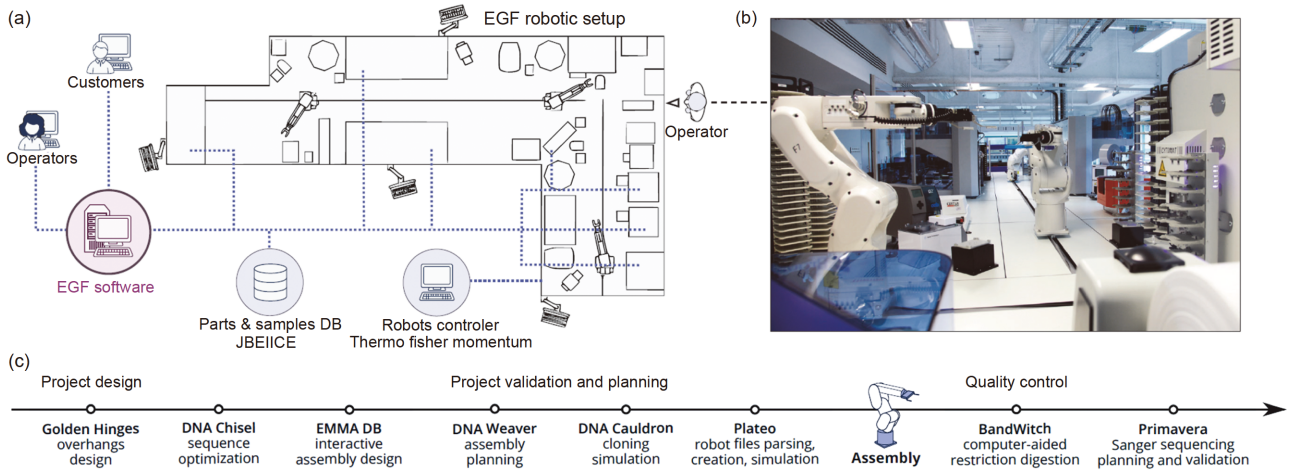


图 6 (网络版彩色)英国爱丁堡大学的EGF设施软硬件系统[39]. (a) EGF平台软件系统将用户、操作者和数据库与自动化平台的不同终端连接起来; (b) EGF平台实景图; (c) 时间线代表不同软件在自动化DNA组装流程中的应用

Figure 6 (Color online) Hardware and software platform of Edinburgh Genome Foundry (EGF)[39]. (a) EGF software interfaces customers, operators, databases, and a robotic platform; (b) inset photo showing the robotic setup from the operator's viewpoint; (c) timeline representing how software are utilized during a DNA assembly project

与iBioFAB、EGF的全整合型方案不同, 帝国理工大学搭建的London Biofoundry采用模块化设计, 由Analytik Jena提供集成方案. 该设施平台包含若干个侧重于不同合成生物工艺的功能岛, 如以纳升移液工作站为核心的酶反应功能岛, 及以常规液体工作站为核心的细胞转化与核酸提取功能岛等, 主要关注细菌、酵母、无细胞底盘系统. 神户大学Biofoundry也采用模块化设计, 针对>100 kb的大DNA片段组装、细胞转化、代谢组样品前处理等流程分别建有自动化功能岛, 主要关注细菌、酵母和真菌等底盘系统. 加拿大康考迪亚大学在建中Concordia Genome Foundry包含两个功能岛, 分别关注DNA自动化组装和iPSC细胞自动化操作. 其他规划和在建中的采用模块化设计的平台新加坡国立大学的Singapore BioFoundry、深圳合成生物"大设施"等.

其他非盈利合成生物设施主要由可以进行高通量操作的单台仪器设备组成, 如美国能源部的Agile BioFoundry、澳大利亚皇后大学的AusFAB、澳大利亚麦考瑞大学的Australian Genome Foundry、英国Earlham研究所的DNA Foundry、天津大学Biofoundry等.

除公共研究机构外, 许多企业也搭建了自动化合成生物研究平台, 如美国Amyris、Ginkgo Bioworks、Zymergen、Transcriptic公司等. 例如, 位于美国波士顿的Ginkgo Bioworks公司开发了多个自动化平台(图7). 其中Biowork2平台占地1700 cm^2, 采取模块化功能岛设计, 包含31个机械臂. 其整合方案采用Transcript公司的云实验室软件架构, 主要开展人工细胞工厂的自动化研发. 截至2018年初, Ginkgo Bioworks公司每月执行的自动

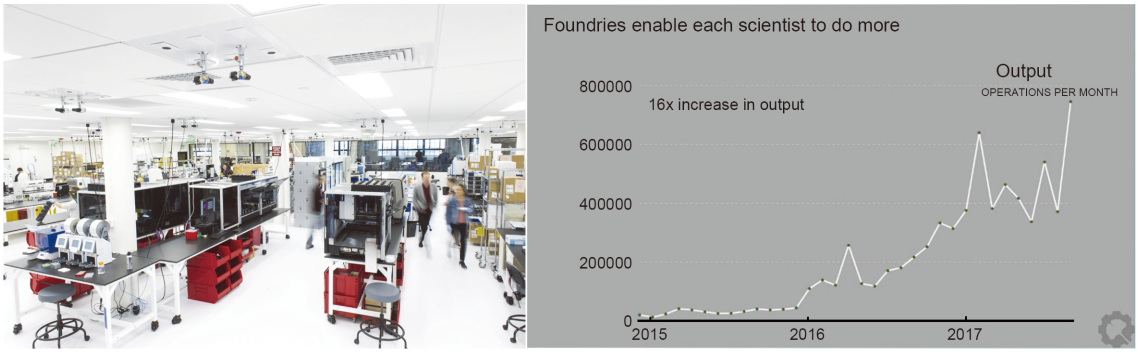


图 7 (网络版彩色)Gingko Bioworks公司自动化设施和操作通量[39]. 图片来源https://ginkgobioworks.com

Figure 7 (Color online) A commercial biofoundry at Ginkgo Bioworks and its operation throughput[39]. Image from https://ginkgobioworks.com

化操作步骤数超过一百万次. 与公共研究机构的合成生物设施相比, 公司研发有较为固定的实验流程, 可以针对特定工艺采取效率最优的功能岛设计; 而研究机构的平台通常需要基于学科发展, 不断开发新的合成生物工艺, 或进行更为频繁的设备升级, 因此对软硬件系统整合方案的柔性、重编程性等提出了更高的要求.

在建的深圳合成生物"大设施"的一期工程建设主要由三大平台构成, 分别是组成物理世界的"合成测试"平台、提供个性化检测的"用户检测"平台, 以及构成数字世界的偏重软件层面的"设计学习"平台. 其中, "合成测试"平台将采用模块化的设计理念, 针对合成生物研究中的主要流程和工艺, 设计和开发集成型功能岛, 用于Oligo组装基因片段、DNA组装、细胞转化、菌落挑取、核酸提取、细胞在线分析等. 模块化设计在提高实验效率的同时, 可以显著降低由于单台设备故障等原因造成的停车风险, 也为将来设备、功能岛和系统的升级提供更好的扩展空间. "合成测试"平台主要关注的研究对象为大片段DNA、噬菌体、细菌、酵母等. 此外, "用户检测"平台由蛋白质与代谢产物分析系统、高级光学检测系统、底盘细胞放大培养系统组成, 为"合成测试"平台构建的人工生命体提供高通量功能测试的同时, 也将满足用户的个性化检测及转化应用需求. 最后, "设计学习"平台的构建基于工业4.0理念和技术发展趋势, 从构建合成生物大设施的运营场景开始, 结合工业互联网技术, 搭建以基础平台层为主体的软件基础设施平台, 实现经营决策层、核心应用层、边缘计算层与物理实验室无缝互联的4层结构的云端实验室平台. 其中的合成生物应用设计模块部分则利用生物信息、数理模型及人工智能等手段, 针对特定科学需求, 基于生物合成大数据, 设计新反应、新酶、新途径、新菌株, 建立一站式的设计技术体系、软件工具及生物信息数据库. 设计学习平台将实现合成生物实验室的柔性自动化, 完善"设计-构建-测试-学习"的研究闭环, 实现物理和数字领域的交汇融合. 大设施的"云端实验室"平台将传统的合成生物研究模式转变为实验服务模式, 帮助科研人员从长时、艰苦、易出错的实验工作中解脱出来, 专注于实验设计、规律总结等更具创造性的思维活动.

4 结语

自动化合成生物技术围绕"如何实现设计合成可预测的生命体"的关键科学问题, 针对传统生命科学人工实验操繁琐、耗时、易错、难以规模化等问题, 通过提升合成生物实验对象、方法、技术的标准化和模块化水平, 实现 "设计-构建-测试-学习"闭环的自动化运行[1,4]. 目前, 基于伊利诺伊大学、爱丁堡大学、帝国理工学院等小型合成生物设施, 研究者对合成生物自动化工程试错的可行性进行了概念验证; 但仍然存在底盘细胞单一、大片段DNA的制造成本高、复杂通路设计能力低、高通量测试手段少等局限, 许多研发需求仍未能满足. 未来自动化合成生物研究及工程化设施建设需要突破的关键技术包括如下内容.

(1) 自动化技术. 开发与高通量合成生物操作兼容的仪器设备、自动化物料转运机器人等硬件系统; 开发"云端实验室软件"架构, 实现设备互联互通、智能调度、动态监控、信息整合等.

(2) 智能设计技术. 开发针对合成生物系统及其实验验证方案的计算机辅助设计方法, 包括合成生物数据库知识库、数据驱动的合成生物设计预测算法和模型、实验方案规划(如实验原料、基因合成及组装策略、设备操作参数等)及指令集生成软件等; 应特别针对自动化设施多循环、海量试错的特点, 结合人工智能方法实现算法和模型的动态优化, 提高合成生物研究效率.

(3) 合成生物支撑技术. 开发与自动化流程适配的合成生物构建与测试技术, 如大片段DNA合成与组装、非模式底盘转化与培养、高通量多维度光学与化学分析等, 从而扩展工程化设施研究对象的范围和规模.

自动化合成生物技术尚处在发展初期, 需要合成生物、自动化、分析化学、信息技术等多领域的研究人员和工程师共同协作, 进行设施平台自身的多轮工程化迭代优化, 为合成生物领域的基础与应用研究的快速发展提供支撑, 并有望带来革命性影响.

参考文献

1 Zhao G P. Synthetic biology: Unsealing the convergence era of life science research (in Chinese). Bull Chin Acad Sci, 2018, 33: 1135–1149 [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: 1135–1149]

2 Boeke J D, Church G, Hessel A, et al. The Genome Project-Write. Science, 2016, 353: 126–127

3 Paddon C J, Keasling J D. Semi-synthetic artemisinin: A model for the use of synthetic biology in pharmaceutical development. Nat Rev Microbiol,

2014, 12: 355–367
4 Cui J M, Zhang B Z, Ma Y F, et al. Engineering platforms for Synthetic Biology research (in Chinese). Bull Chin Acad Sci, 2018, 33: 1249–1257 [崔金明, 张炳照, 马迎飞, 等. 合成生物学研究的工程化平台. 中国科学院院刊, 2018, 33: 1249–1257]
5 Chao R, Mishra S, Si T, et al. Engineering biological systems using automated biofoundries. Metab Eng, 2017, 42: 98–108
6 Chao R, Yuan Y B, Zhao H M. Building biological foundries for next-generation synthetic biology (in Chinese). Sci China Life Sci, 2015, 58: 658–665 [晁然, 原永波, 赵惠民. 构建合成生物学制造厂. 中国科学: 生命科学, 2015, 45: 967–984]
7 Hillson N, Caddick M, Cai Y, et al. Building a global alliance of biofoundries. Nat Commun, 2019, 10: 2040
8 Appleton E, Madsen C, Roehner N, et al. Design automation in synthetic biology. Cold Spring Harb Perspect Biol, 2017, 9: a023978
9 Hillson N J. j5 DNA assembly design automation. Methods Mol Biol, 2014, 1116: 245–269
10 Kosuri S, Church G M. Large-scale *de novo* DNA synthesis: Technologies and applications. Nat Methods, 2014, 11: 499–507
11 Beaucage S L, Caruthers M H. Deoxynucleoside phosphoramidites—A new class of key intermediates for deoxypolynucleotide synthesis. Tetrahedron Lett, 1981, 22: 1859–1862
12 Palluk S, Arlow D H, de Rond T, et al. *De novo* DNA synthesis using polymerase-nucleotide conjugates. Nat Biotechnol, 2018, 36: 645–650
13 Smith H O, Hutchison Clyde A I, Pfannkoch C, et al. Generating a synthetic genome by whole genome assembly: X174 bacteriophage from synthetic oligonucleotides. Proc Natl Acad Sci USA, 2003, 100: 15440–15445
14 Gibson D G, Glass J I, Lartigue C, et al. Creation of a bacterial cell controlled by a chemically synthesized genome. Science, 2010, 329: 52–56
15 Engler C, Kandzia R, Marillonnet S. A one pot, one step, precision cloning method with high throughput capability. PLoS One, 2008, 3: e3647
16 Chao R, Liang J, Tasan I, et al. Fully automated one-step synthesis of single-transcript TALEN pairs using a biological foundry. ACS Synth Biol, 2017, 6: 678–685
17 Nagy M, Otremba P, Krüger C, et al. Optimization and validation of a fully automated silica-coated magnetic beads purification technology in forensics. Forensic Sci Int, 2005, 152: 13–22
18 Frégeau C J, Lett C M, Fourney R M. Validation of a DNA IQ^{TM}-based extraction method for TECAN robotic liquid handling workstations for processing casework. Forensic Sci Int-Genet, 2010, 4: 292–304
19 Dharmadi Y, Patel K, Shapland E, et al. High-throughput, cost-effective verification of structural DNA assembly. Nucleic Acids Res, 2014, 42: e22
20 Mitchell L A, Phillips N A, Lafont A, et al. qPCRTag analysis—A high throughput, real time PCR assay for Sc2.0 genotyping. J Vis Exp, 2015, 2015: e52941
21 Metzker M L. Sequencing technologies—The next generation. Nat Rev Genet, 2010, 11: 31–46
22 Shapland E B, Holmes V, Reeves C D, et al. Low-cost, high-throughput sequencing of DNA assemblies using a highly multiplexed nextera process. ACS Synth Biol, 2015, 4: 860–866
23 Hughes S R, Patrick F, Dowd R E H, et al. Cost-effective high-throughput fully automated construction of a multiplex library of mutagenized open reading frames for an insecticidal peptide using a plasmid-based functional proteomic robotic workcell with improved vacuum system. J Lab Autom, 2007, 12: 202–212
24 Si T, Chao R, Min Y, et al. Automated multiplex genome-scale engineering in yeast. Nat Commun, 2017, 8: 15187
25 Rajakumar P D, Gowers G O F, Suckling L, et al. Rapid prototyping platform for saccharomyces cerevisiae using computer-aided genetic design enabled by parallel software and workcell platform development. SLAS Technol, 2019, 24: 291–297
26 Chen L, Wang Y, Guo H, et al. High-throughput screening for Streptomyces antibiotic biosynthesis activators. Appl Environ Microbiol, 2012, 78: 4526–4528
27 Kuivanen J, Korja V, Holmström S, et al. Development of microtiter plate scale CRISPR/Cas9 transformation method for *Aspergillus niger* based on in vitro assembled ribonucleoprotein complexes. Fungal Biol Biotechnol, 2019, 6: 3
28 Lenaghan S C, Neal Stewart C Jr. An automated protoplast transformation system. Methods Mol Biol, 2019, 1917: 355–363
29 Li J F, Park E, von Arnim A G, et al. The FAST technique: A simplified *Agrobacterium*-based transformation method for transient gene expression analysis in seedlings of *Arabidopsis* and other plant species. Plant Methods, 2009, 5: 6
30 Wang S, Lee L J. Micro-/nanofluidics based cell electroporation. Biomicrofluidics, 2013, 7: 11301
31 Liu L, Dalal C K, Heineike B M, et al. High throughput gene expression profiling of yeast colonies with microgel-culture Drop-seq. Lab Chip, 2019, 19: 1838–1849
32 Kumar R, Ichihashi Y, Kimura S, et al. A high-throughput method for illumina RNA-Seq library preparation. Front Plant Sci, 2012, 3: 202
33 Si T, Li B, Comi T J, et al. Profiling of microbial colonies for high-throughput engineering of multistep enzymatic reactions via optically guided matrix-assisted laser desorption/ionization mass spectrometry. J Am Chem Soc, 2017, 139: 12466–12473
34 Xue P, Si T, Mishra S, et al. A mass spectrometry-based high-throughput screening method for engineering fatty acid synthases with improved production of medium-chain fatty acids. Biotechnol Bioeng, 2020, 117: 2131–2138
35 Boutros M, Heigwer F, Laufer C. Microscopy-based high-content screening. Cell, 2015, 163: 1314–1325
36 Esmonde-White K A, Cuellar M, Uerpmann C, et al. Raman spectroscopy as a process analytical technology for pharmaceutical manufacturing and bioprocessing. Anal Bioanal Chem, 2017, 409: 637–649
37 HamediRad M, Chao R, Weisberg S, et al. Towards a fully automated algorithm driven platform for biosystems design. Nat Commun, 2019, 10: 5150
38 Hu R Y, Zhang S Y, Meng H L, et al. Machine learning for synthetic biology: Methods and applications (in Chinese). Chin Sci Bull, 2021, 66: 284–299 [胡如云, 张嵩亚, 蒙海林, 等. 面向合成生物学的机器学习方法及应用. 科学通报, 2021, 66: 284–299]
39 Zulkower V, Bleda A, Luo I, et al. Software projects of the Edinburgh Genome Foundry. In: International Workshop on Bio-Design Automation, 2018

Summary for “自动化合成生物技术与工程化设施平台”

Automation in synthetic biology using biological foundries

Ting Tang, Lihao Fu, Erpeng Guo, Zhenkun Zhang, Zining Wang, Chenfei Ma, Zhiyu Zhang, Jianzhi Zhang, Jiandong Huang* & Tong Si*

Key Laboratory of Quantitative Engineering Biology, Shenzhen Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China
* Corresponding authors, E-mail: jdhuang@hku.hk; tong.si@siat.ac.cn

Synthetic biology applies engineering principles for the deliberate design, engineering, and *de novo* creation of artificial biological systems with certain functions. Due to the complexity of living systems and lack of rational design principles, iterative trial-and-error experiments are often necessary, but the dependence on human researchers limits the throughput, efficiency, and consistency of such endeavor. To overcome these limitations, biofoundries are developed as an integrated infrastructure for accelerating the “design-build-test-learn” cycles in synthetic biology research and biotechnology applications. Computer-aided design and robotic automation are applied in the physical manufacturing and prototyping of engineered DNA and genetically reprogrammed organisms. Currently, many biofoundries are being created around the world, and a Global Biofoundry Alliance has been established to promote coordination and collaboration. This paper aims to summarize the recent advances in synthetic biology automation and introduce design and construction of current and future biofoundries.

We start with key technologies that promote automation in synthetic biology, including computer-aided design, high-throughput instrumentation, robotic integration, automation-compatible workflows for DNA assembly and chassis engineering, and analytical approaches. For bio-design automation, we introduce existing software tools such as j5 from Agile Biofoundry, iBioCAD from Illinois Biological Foundry for Advanced Biomanufacturing (iBioFAB), and CUBA from Edinburgh Genome Foundry (EGF). Moreover, software and hardware for creating automated build and test workflows are summarized, followed by recent advances in DNA synthesis, DNA assembly, nucleic acid extraction and analysis, chassis manipulation, and high-throughput testing. Notable examples are discussed, including automated DNA assembly using the Golden-Gate method, multiplex automated yeast genome engineering, and artificial intelligence (AI)-guided optimization of biosynthetic pathways.

Then, representative biofoundries and their software systems, robotic platforms, and available workflows are discussed. Biofoundries are categorized by various formats of integration, including full integration (e.g., iBioFAB and EGF), modular integration (e.g., London Biofoundry, Singapore Biofoundry, and Concordia Genome Foundry), and manual integration. Commercial biofoundries are also compared and contrasted with their academic counterparts, using the ones at Gingko Bioworks Inc. as an example. Furthermore, the design of Shenzhen Biofoundry is discussed in detail, whereby a centralized cloud lab is envisioned to serve the domestic and international synthetic biology communities. Shenzhen Biofoundry will consist of a build platform for engineered DNA, phage, bacteria and yeast, a test platform for optics, chromatography and mass spectrometry, and fermentation scale-up, and a design/learn/cloud platform to coordinate and integrate the whole facility.

We conclude with the future challenges and promises of biofoundries and synthetic biology automation. We propose key research directions, including new automation technologies, data-driven and intelligent design, and automation-compatible workflows in synthetic biology. Systems approach and interdisciplinary collaboration are necessary for biofoundries development, which requires synergistic integration of synthetic biology, analytical chemistry, robotics, instrumentation, informatics, and smart manufacturing. Just like foundries initiated the magnificent prosperity of the semiconductor industry, biofoundries will unleash the potential of synthetic biology to revolutionize our society.

synthetic biology, automation, high throughput, biofoundry, cloud lab, DNA assembly

doi: 10.1360/TB-2020-0498