

特约评述

DOI: 10.12211/2096-8280.2021-092

进化视角下的定量生物学规律与人工生命合成

赵晓宇^{1,2}, 张浩¹, 李雪飞¹, 胡政¹¹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, 广东 深圳 518055; ²中国科学院大学, 北京 100049)

摘要: 遗传进化是生物系统的一个基本特征, 生物系统的结构与功能都是动态变化的, 生物通过进化更好地适应环境。定量合成生物学作为一门新兴交叉学科, 主要研究如何利用合成系统定量刻画生物学规律, 以及基于理性设计和改造人工生命系统来解答生命科学前沿问题。然而, 目前掌握的知识还不足以满足理性设计和定量可控的要求。尽管人工生命系统产生于实验室, 但其同样受到进化法则的支配, 比如突变、遗传漂变、达尔文自然选择等, 因此需要利用进化规律帮助设计和构建更加稳定的人工生命系统; 反过来, 简单、周期短且可控的定量合成生物系统也可帮助研究生物进化原理提供更佳的生物模型, 两者相互促进, 为更好地探索生命法则提供了理论支撑和技术手段。本文主要综述了目前应用进化原理筛选目标蛋白的连续定向进化方法, 总结了利用进化原理提高合成线路稳定性的策略; 同时, 介绍了用定量合成生物学的手段研究生物进化原理的进展, 并提出了基于进化原理的人工生命设计的研究方向。未来合成生物学与进化生物学的融合发展将为精确控制生命系统提供思路和方法, 增进合成生物学改善人类可持续发展的应用。

关键词: 遗传进化; 定量生物学; 合成生物学; 人工生命系统; 群落进化

中图分类号: Q812 文献标志码: A

An evolutionary perspective on quantitative biological principles and synthetic life design

ZHAO Xiaoyu^{1,2}, ZHANG Hao¹, LI Xuefei¹, HU Zheng¹

¹Shenzhen Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Science, Shenzhen 518055, Guangdong, China; ²University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: Genetic evolution is a defining feature of all biological systems, whereby lives are able to adapt to the rapidly changing environments. Quantitative synthetic biology, an emerging field in life science, mainly studies how to apply synthetic biosystems to reveal quantitative biological rules, which in turn guides the rational design of synthetic life. Although such artificial biosystems are created in the laboratory, they are also governed by the rules of evolution, such as mutation, genetic drift, Darwinian selection, etc. However, the evolutionary potential of a synthetic biosystem

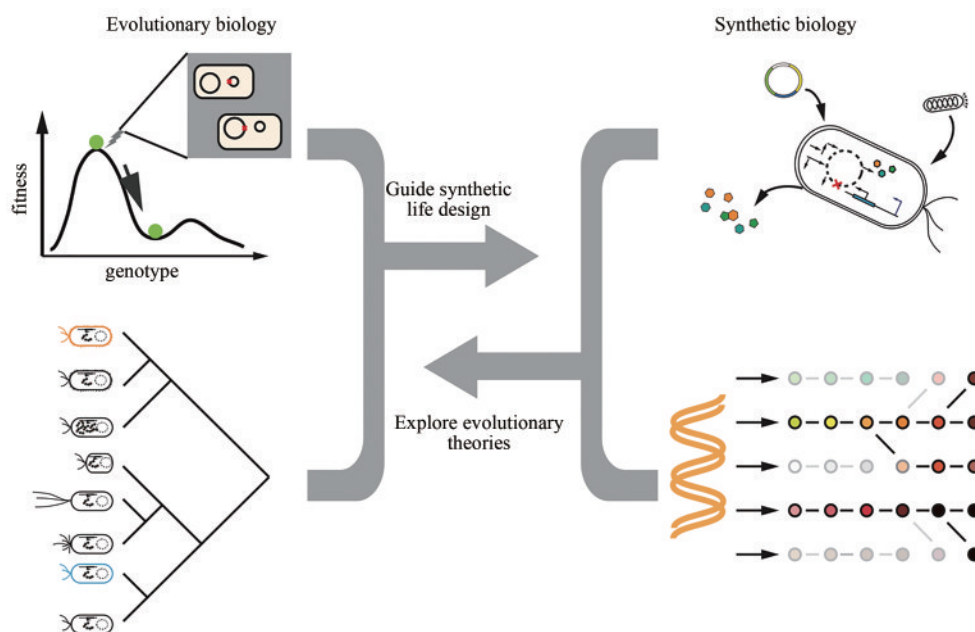
收稿日期: 2021-09-18 修回日期: 2021-11-08

基金项目: 国家自然科学基金青年基金 (32000886); 广东省自然科学基金杰出青年基金 (2021B1515020042); 广东省自然科学基金面上项目 (2114050001502)

引用本文: 赵晓宇, 张浩, 李雪飞, 胡政. 进化视角下的定量生物学规律与人工生命合成[J]. 合成生物学, 2022, 3(1): 6-21

Citation: ZHAO Xiaoyu, ZHANG Hao, LI Xuefei, HU Zheng. An evolutionary perspective on quantitative biological principles and synthetic life design [J]. Synthetic Biology Journal, 2022, 3(1): 6-21

might have been underrated, which often leads to the circuit failures or the emergence of undesigned function. Therefore, overcoming evolutionary perturbations is one of grand challenges in robust circuit design, and harnessing the rules of evolution would be important for more robust design of artificial biosystems. On the other hand, bottom-up synthetic systems provide favorable biological models to test and explore the evolutionary principles that are hard to study based on natural biosystems. In this review, we first summarize the current progress in screening functional proteins through continuous directed-evolution guided by the evolutionary principles. We assess advantages and disadvantages for different architectures in continuous directed-evolution through evolutionary perspective. We also review the evolutionary strategies that are used to improve the stability of synthetic circuits in host cells, including reducing mutation rate, the coupling of host cell fitness with target gene expression, synthetic addiction, compartmentalization, feedback control, orthogonal system, etc. Furthermore, we discuss how to use the synthetic biology approaches to address some basic theories in evolutionary biology, such as the origin of life, the development of multicellularity, genetic constraint and epistasis, and the evolutionary tradeoff of complex life system, etc. We believe that a real-time precise control of the evolutionary process would enhance the quantitative features and predictability in the field of synthetic biology, and greatly advance the applications of robust synthetic biosystems for human sustainable goals.



Keywords: biological evolution; quantitative biology; synthetic biology; artificial life; community evolution

定量合成生物学主要通过结合定量测量与数学模型探索生命的规律，同时基于定量规律利用合成生物技术进行理性改造或创造人工生命^[1]。尽管这样的人工生命系统完全诞生于实验室内，但其也受自然生存法则的支配，如复杂的进化过程经常使得目标功能难以长期稳定地维持^[2]。这主要有以下几个原因：一是构建基因回路带来的

代谢负担经常对底盘细胞产生负向选择压力^[3-4]，使得它们与天然菌株竞争时失败；二是研究人员并没有考虑环境变化对人工生命产生的影响^[5]，比如环境治理工程菌、肠道菌群改造工程菌等，其应用环境往往不是标准的实验室条件，复杂的环境调控和菌群竞争导致很多基因回路无法实现预期功能。因此，从演化的视角来进行合成生命

系统的设计和构建，对更好地实现其功能是非常必要的。

本文主要从两方面来综述进化生物学与定量合成生物学之间的关系：一方面，定向进化、进化动力学和进化预测模型可以帮助合成生物学更高效地筛选出目标人工生命；另一方面，应用合成生物技术和定量方法可以构建简单可控的生物系统，帮助研究进化的基本理论、探索生命起源及了解复杂生命系统中的进化轨迹。两者的紧密结合可为研究各种生物问题提供更实际有效的思路。

1 进化理论帮助人工生命的设计与构建

自从达尔文提出“物竞天择，适者生存”的自然法则，进化思想被人熟知并被广泛应用于生物学各个领域。自然选择作为进化中最核心的观

点，被应用于合成生物学中，对改造后的目标生物进行筛选^[6]。此外进化动力学也被用于探讨合成生命系统的基本规律^[7]；而基于大数据和机器学习的进化预测工具，更是能帮助科学家在计算机中模拟合成与筛选过程，为实验设计提供最佳方案^[8]。

1.1 定向进化帮助改造目标蛋白

定向进化（directed-evolution）是利用大规模突变和条件筛选，期望对目标产物进行定向改造并快速进化，使目标蛋白在短时间内增强现有功能或获得新功能的一种生物技术^[9] [图1(a)]，广泛应用于DNA序列、基因功能和蛋白质结构的优化和筛选。但由于其依赖人为操作的局限性，并不能还原自然进化的大规模与进化深度 [图1(b)]。针对此出现了几种解决方法：一是近年来出现的连续定向进化系统（continuous-

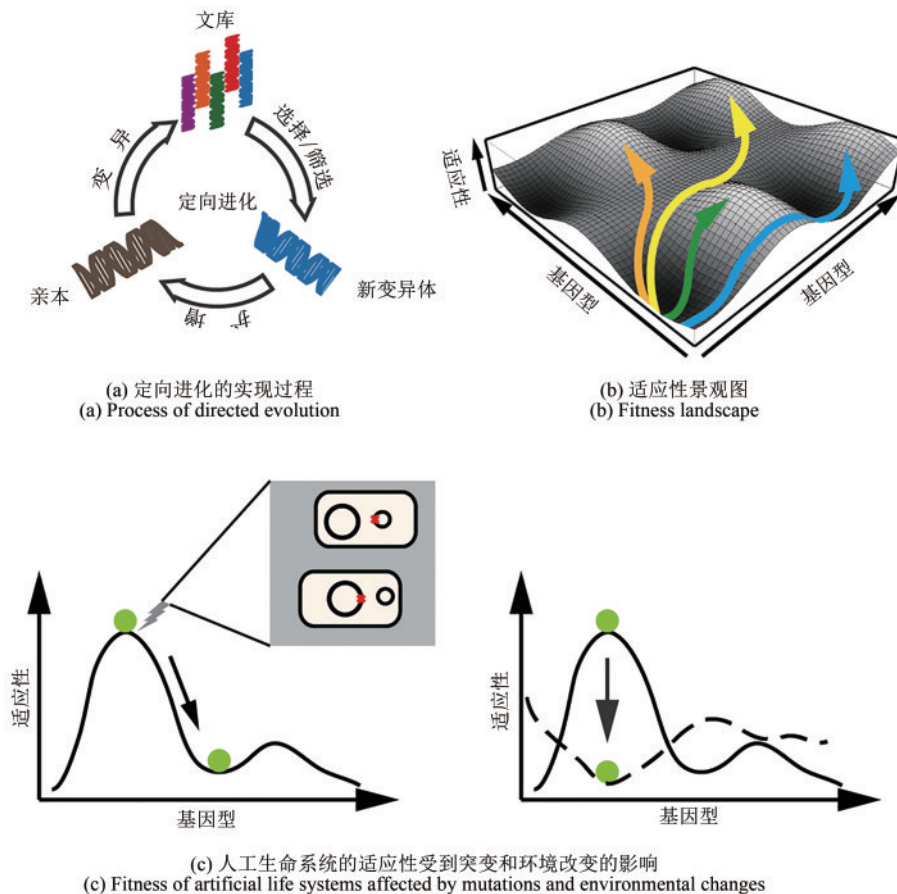


图1 定向进化原理

Fig. 1 Principles for directed-evolution

directed-evolution)^[10] 可利用生物体自身的生命迭代完成自主且连续的进化过程，完美解决了传统进化无法到达的“规模”与“深度”的问题。二是机器学习和定向进化的结合^[11]，可更加理性地设计蛋白突变序列，跳脱出局部筛选的限制，找到适应性更高的进化表型。

连续定向进化方法多样，步骤大致可分为：①构建目的基因的突变文库；②在合适的宿主中表达并被筛选；③扩增后进入下一轮突变筛选^[12-13]。主要区别在于第1步的构建突变文库，其质量对定向进化起着决定性作用，可通过3类框架来实现^[12]：病毒框架、“外部”突变框架和直接突变框架。而第2步的筛选一般是将目的性状与生存必需的基因关联后用抗生素等筛选，或者表达荧光等生物标志物或生物传感器后用流式细胞术等仪器筛选^[14]。下面将根据构建突变框架的不同，介绍几种应用最广泛的连续定向进化系统。

病毒框架 (viral architecture) 是利用病毒搭载目的基因，在非扩增性宿主中进行突变富集。最有名的例子是 David R Liu 实验室开发的噬菌体辅助连续进化技术 (phage-assisted continuous evolution, PACE)^[10, 15] [图2(a)]。该方法是通过突变质粒 MP 来降低大肠杆菌 DNA 聚合酶的校正功能，增加噬菌体携带的目的基因在复制过程中的突变率^[16-17]。目标基因无功能的突变噬菌体同时也丧失了侵染能力，会随着不断流通的环境而被淘汰。该系统进化速度快，无需提供突变文库和人为控

制，可自发朝着目标功能进化；但进化对象有很大的局限性且操作复杂。“外部”超突变框架 (epi-hypermutation architectures) 也可称为核酸编辑技术，是指在独立于 DNA 自身复制的系统之外，利用超突变系统 (如 DNA 编辑酶与 DNA 结合蛋白融合的系统) 对目标基因进行连续突变^[12]。比如 CRISPR-X 系统^[18] [图2(b)]，其原理是 CRISPR 系统的 Cas9 蛋白融合胞苷脱氨酶 (AID)，在 sgRNA 的引导下结合到靶向 DNA，并将胞嘧啶 (C) 脱氨变成尿嘧啶 (U)，进而通过修复转变为胸腺嘧啶 (T)，实现特定碱基对的替换。而类似原理的 CBE 碱基编辑器^[19] 可随着 CRISPR 系统不断导向突变的目标序列实现连续进化。以上两种间接突变框架的优势是宿主细胞的背景突变率没有改变，因此筛选获得的目标突变基因受到宿主自身 DNA 复制系统影响较少；其缺点是脱靶效应。直接超突变框架 (direct hypermutation architectures) 是利用低保真 DNA 聚合酶进行直接突变，推动目标基因在细胞内的快速进化^[12] [图2(c)]。一种常用方法是对错配修复基因进行突变，增加突变速率，即利用增变基因 (mutator) 增加目标基因和宿主基因组的突变速率^[20-21]。然而这种方法由于增加了宿主基因组的突变率，筛选的特异性和稳定性较低。一种改进方式是设计正交的 DNA 复制系统 [图2(c)]，比如 OrthoRep 系统^[22] 是将一对含有自身 DNA 复制元件的质粒 P1 和 P2 转入酵母中，并在 P1 中插入待突变基因和

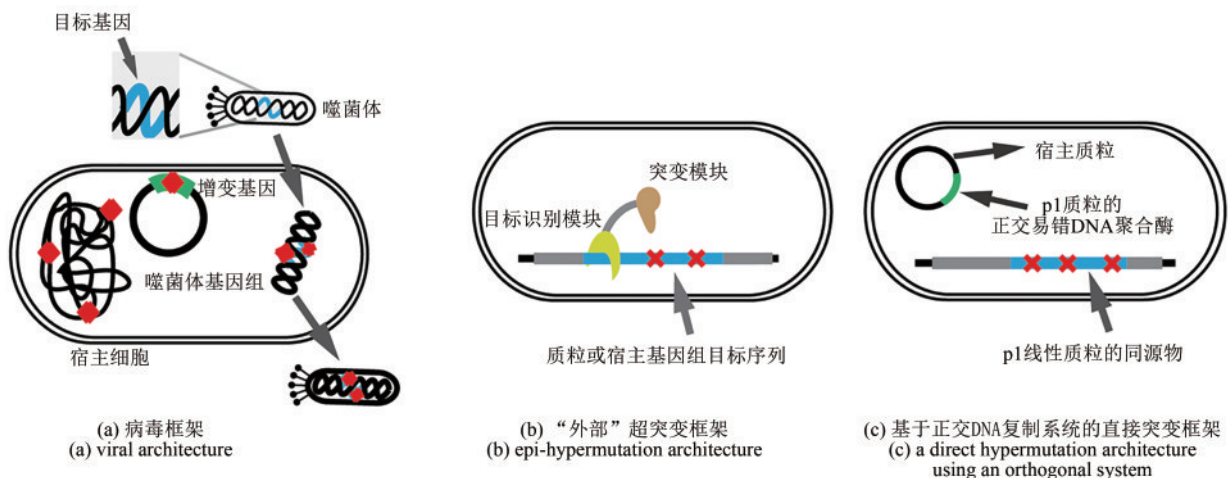


图2 连续定向进化中的突变方法^[12]

Fig. 2 Mutagenesis in continuous directed-evolution^[12]

突变了的DNA聚合酶（可提高目的基因的突变率）。由于该质粒与基因组的复制机制不同，从而实现了体内连续进化的正交化。该系统体内突变率高且可保持超过300代；但其编码基因的表达水平有限，且只能在酵母中实现。常见的连续定向进化方法见表1。

同时机器学习模拟进化过程^[11, 34]也是一种新兴有效的方法。它的原理是利用蛋白定向进化后不同适合度及其对应的测序信息，通过各种算法学习序列的变异特征，选取精确度最高的模型用于模拟筛选突变体并对适合度进行排序，最后构建一个有限范围的突变文库，进行实验筛选验证^[11, 35]。其中构建一个有效的序列函数模型对模拟进化的质量至关重要，常用的算法包括决策树（decision trees, DTs）、高斯过程（Gaussian process, GP）、贝叶斯算法（Bayesian algorithms）、降维算法（dimensionality reduction algorithms）和神经网

络算法（artificial neural networks, ANNs）等。Arnold团队^[34]总结了利用这些机器学习算法进化蛋白质的方法及具体案例。机器学习的优点在于，它可以跳脱出定向进化可能导致的局部最优解，通过有效地学习整个功能景观来探索整体最优突变体，节约了大量试错时间且大大提高蛋白效用。

机器学习可以加速生物元件的理性设计，指导其向特定的性质进化，如提高酶的催化活性^[36]和热稳定性^[37]、预测高亲和力的核酸适配体^[38-40]、设计非生物微蛋白^[41]等。下面具体介绍Arnold团队^[37]如何利用机器学习指导细胞色素P450酶向热稳定性最大化进化的实例。他们用独热编码（one-hot representation）描述蛋白的序列和结构的关系，并建立高斯模型推测出蛋白的适合度景观，然后用一组嵌合细胞色素P450的热稳定性数据进行训练，不断检验后该模型能达到很高的预测能力。

表1 连续定向进化的原理和应用

Table 1 Principles and methods for continuous directed-evolution

方法	框架	原理	进化对象	突变速率	实例	突破性与局限性	方法延伸
PACE ^[10, 15]	病毒框架	将目的蛋白与噬菌体生命周期耦联,利用宿主的致突变质粒和辅助质粒完成突变与扩增,通过自动化连续培养实现选择与连续进化	包含目的基因的噬菌体	2.3×10^{-3} bp/代; 突变文库数量可达到 10^9 个	碱基编辑器 ^[23] 、Cas9蛋白 ^[24] 、肉毒素蛋白酶 ^[25]	优点:应用范围最广且优化速度较快,是其他进化方法的100倍 缺点:只适用于能被噬菌体感染的原核生物;只能突变引起gIII基因表达变化的蛋白;操作和设备复杂	SE-PACE系统 ^[26] 、PANCE系统 ^[27]
CRISPR-X ^[18]	“外部”突变框架	CRISPR系统融合胞苷脱氨酶(AID),将目标基因的C脱氨并修复为T,实现碱基对直接替换	哺乳细胞	1×10^{-3} bp/代	慢性髓系白血病细胞中的耐药突变 ^[28] 和癌症治疗药物硼替佐米的靶点耐药突变 ^[18]	优点:适用于哺乳细胞,实现特定碱基对替换 缺点:只能实现C到T的突变,突变率较低,未实现真正的连续进化	胞嘧啶剪辑编辑器 ^[19] 、腺嘌呤碱基编辑器(ABE) ^[29] 和先导编辑器(Prime Editor, PE) ^[30]
OrthoRep ^[22]	直接突变框架	含有目标基因和突变的DNAP的质粒、含有所有必需基因的质粒负责靶基因的复制与突变,与基因组的复制正交	酵母中的正交突变质粒	1×10^{-5} bp/代; 能保持超过300代的高突变率	进化耐药性疟疾二氢叶酸还原酶 ^[22]	优点:正交系统可保持细胞基因组的稳定,实现连续进化 缺点:局限于酵母;系统设计复杂;外源性基因低表达	OrthoRep基因表达盒 ^[31] 、两个相互正交的DNA复制系统 ^[32] 、扩展OrthoRep的通用性并加速进化效率 ^[33]

(This table describes principles, evolution targets and mutation rate of different continuous directed-evolution methods, and lists their application examples, advantages and disadvantages and subsequent improvements.)

随后他们开发并测试了两种基于贝叶斯决策理论的蛋白质序列设计算法，通过迭代改进高斯过程模型，预测识别优化序列。对这些预测的蛋白变体进行实验验证，发现其中两个酶的耐热性高于之前通过理性设计或定向进化得到的任何酶。他们用机器学习算法建立的定向进化框架，可以高效利用少量数据进行预测，该框架已被用于其他蛋白系统和特性的设计中^[42-43]。

机器学习结合自动化平台已经完全实现了自动进化。如赵惠民团队^[38]设计的基于贝叶斯优化模型的全自动化机器平台“BioAutomata”，可实现合成生物学“设计-构建-测试-学习”模式的完全自动化，效果优于77%的随机筛选。这种“自动进化”已经在优化番茄红素系统中验证其可行性^[38]。当然，机器学习还存在应用上的局限性：除了本身的算法和输入描述存在误差外，它比定向进化要付出额外的测序成本和建模预测的时间；且它的生物学可解释性较差，无法普遍推广^[34]。

1.2 进化动力学和博弈论帮助设计合成生命系统

进化原理不仅适用于个体或基因的进化，也同样适用于生物种群和群落。当生物个体众多时，他们间的关系很大程度上会决定群体的进化方向，比如它们如何在空间上分配生态位、如何分配有限资源、如何应对突变和选择等。这些问题也都是合成生物学在设计微生物群体时需要考虑的重要问题，因此将进化动力学和生态群落的博弈思想应用于合成生物学的设计，可帮助合成群落增强稳定性和可持续性。比如中国科学院深圳先进技术研究院的刘陈立团队^[7]在该领域做了很好的尝试。他们首先构建了细菌迁徙的适应性进化系统，发现迁移快慢不同的细菌分别占据外围和里面的生态位；然后构建了两种迁移速率可控的工程菌，让它们以不同的速率迁移。最后作者得到了菌群的迁徙进化规律，并推导出计算迁徙进化最优策略的定量公式。这对构建稳定的合成群落并维持其多样性提供了重要的理论指导。

进化博弈论主要是在达尔文自然选择的框架下研究个体或群体的生存策略。它可用来指导合成微生物菌群间相互关系的设计，比如利用不同

菌群间的制衡作用帮助维持整个合成群落的功能。Hasty团队^[44]利用微生物群落的竞争关系，设计了3种相互抑制的大肠杆菌，每种细胞内基因线路是通过周期性种群控制来维持稳定。这种“石头-剪刀-布”的相互竞争的生态位大大提高了基因回路的功能稳定性。而Studer团队^[45]利用生态群落间的合作关系，设计了一种多维生物的联合代谢系统，这些微生物具有不同的代谢能力，且占据不同的生态位，可合作将木质纤维素转化为短链脂肪酸。这种微生物的分工合作可大大减少负担并提高效率，为微生物联合体的设计提供了借鉴方案。

进化博弈论也可用于理解人体内环境中出现“入侵种”后的进化策略。肿瘤对正常组织来说可算作“入侵种”，其不断扩增会严重影响机体正常功能，夺取正常组织的生态位。而在肿瘤微环境中，改造的免疫细胞或工程微生物药物对肿瘤微环境可算作“入侵种”，可能会影响肿瘤的进化方向。研究肿瘤如何在肿瘤微环境的选择作用下获取进化优势，产生抗免疫或耐药性对肿瘤防治有重要意义。如Lang等^[46]利用酵母的进化策略研究来模拟肿瘤进化。他们用来自11个实验进化的酵母菌群的116个突变来衡量突变对生长的影响，发现成功的突变只有20%是驱动型的（即有利突变），其余的都是搭便车。其中一个突变组合带来的适合度增加大于全体基因突变的效应的总和。对模式生物进化动力学的研究有助于人们了解并推广到其他系统，如癌症中的基因突变动力学。

1.3 提高合成回路进化稳定性的构建方法

进化在很大程度上能帮助快速筛选到目标产物，但有些时候它也会破坏精心设计的基因线路。合成生物学家常常面对随机突变引起的合成回路功能丢失、基因线路带来的代谢负担等问题^[47] [图1(c)、(d)]。因此如何避免菌群迭代导致的基因线路被破坏，且合理定量地预测进化轨迹，对设计合成生命系统有重要的指导意义。总的来说，提高基因回路进化稳定性有3种策略：①降低丢失突变的发生概率；②抑制丢失突变的适应性；③降低合成回路对宿主细胞产生的负担^[48-50]。下面

将分别对这3种策略进行介绍。

降低背景突变速率可以减小合成回路功能丢失的风险 [图3(a)]。在以质粒作为合成回路载体的细菌底盘细胞中，转座元件 (transposable elements, TEs) 或插入序列 (insertion sequences, ISs) 是引起合成回路功能丢失的主要原因 [47]。因此，从细菌基因组中删除这些元件可以降低突变速率 [51-53]。比如，Chan等 [54] 测试一个毒素介导的细菌合成自杀开关的效果，发现在删除IS元件的大肠杆菌中自杀开关失效的概率降低了3~5个数量级。类似的，在假单胞菌和不动杆菌中删除转座元件也大大降低合成回路的稳定性 [51, 55]。然而由于底盘细胞固有的本底突变无法避免，需要增加调控机制来缓解或消除突变对人工生命系统的影响。

抑制丢失突变体的适应性能显著提高合成回路进化稳定性。比如在代谢工程领域，将目标产

物与底盘细胞的适应性相耦合 [图3(b), growth-coupled production]，突变体因为相对适合度的下降而在群体中被逐渐淘汰 [56]，从而保持群体整体的稳定性。通过适应性实验室进化 (adaptive laboratory evolution) 可以提高目标产物与底盘细胞适应性的耦合性 [57]。比如，Fong等 [58] 通过代谢流分析，对构建的大肠杆菌突变株进行实验室进化，提高了其产乳酸的能力和细胞的生长速度，在增加产量的同时保持了菌株的稳定性。合成成瘾 (synthetic addiction) 也是常用的抑制丢失突变的适应性的方法 [59]。通过将必需基因引入合成回路中，让底盘细胞的活性对必需基因表达产生依赖性 [图3(c)]。突变体的适应性会因为失去该必需基因的高表达而下降。Rugbjerg等 [59] 首次在大肠杆菌中构建了甲羟戊酸的感知器合成回路，能够诱导必需基因的 *glmM* 和 *folP* 表达，从而构建了对甲羟戊酸“成瘾”的大肠杆菌，能维持超过

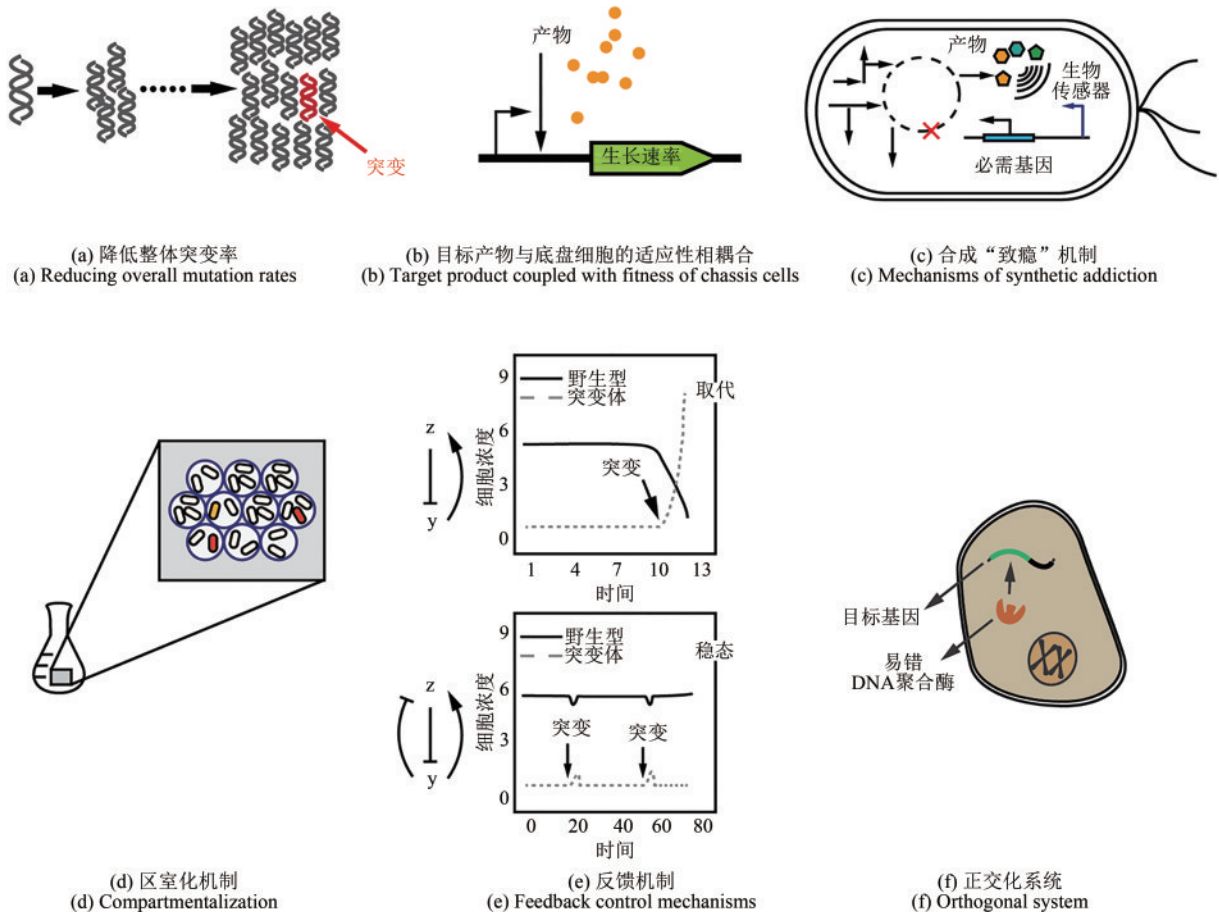


图3 提高合成基因回路进化稳定性的构建方法

Fig. 3 Construction methods for improving the evolutionary stability of synthetic gene circuits

90代的高产量培养。另外一个抑制突变体在群体中扩张的方法是将培养系统区室化 [图3(d)], 比如构建微流控、微孔或微囊等微小反应器 (micro-bioreactor), 这类似于在群体中引入空间结构, 其可以限制优势突变只在区域中扩张而不会取代整个群体^[60]。游凌冲和戴卓君团队^[61]发明了包裹合成大肠杆菌的智能微胶囊, 让蛋白生产在各个区室中分开进行, 因而逃脱合成线路的突变体只局限在各自的微胶囊中, 不会对整体造成污染, 从而大大提高合成线路的稳定性。

外源的基因回路的表达对于底盘细胞而言是一种“负担”^[3-4]。合成回路的功能实现依赖宿主细胞的DNA复制系统、转录和翻译系统以及各种代谢产物前体。外源基因表达将占用有限的DNA、RNA、蛋白质的合成机器以及各种底物, 这种资源上的占用将改变底盘细胞内源的基因表达, 进而改变底盘细胞的生理状态。因此, 合成回路应该尽量减小对底盘细胞的负担, 才能提高其稳定性。比如研究人员利用控制论开发了一种前反馈的基因回路-非一致前馈环 (incoherent feedforward loop, iFFLs) [图3(e)], 无论质粒本身的起始复制量是多少, 都会稳定地表达产物^[62]。结果表明, 即使基因的拷贝数由于基因组突变或者培养基条件改变发生了变化, 这些由iFFLs稳定的启动子控制的目的基因在不同的基因组位点或者质粒上仍能实现接近相同的表达水平。该方法利构建了非协同的拓扑基因回路, 精妙地解决了合成系统的基因表达水平受质粒拷贝数影响的难题, 提高了系统稳定性。近年来, 涌现出了大量基因回路模块化、正交化的设计思路 [图3(f)], 成功减弱或规避合成回路对底盘细胞的“负担”^[63]。

除上述方法外, 模型构建和预测也能帮助提高合成回路的进化稳定性。最近科学家们尝试利用进化的力量来帮助合成生物学设计。Castle等^[64]提出一个由变异 (variation)、功能 (function) 和选择 (selection) 三个维度决定的进化型 (evotype) 概念, 通过对这三者的改造, 使原先设计的基因型 (design type) 最终进化为目标进化型。理想的进化型分为进化稳定型和特异性进化型, 引入突变和选择会使其朝着功能最强的方向连续进化, 最后达到稳定且高效的“进化型”山顶。该理念

的提出重新构想了合成理念, 把进化的变量引入合成系统也极大帮助科学家重新思考生物设计的底层逻辑。另一种利用进化的预测模型是“质粒进化预测工具”^[8]。它是利用大肠杆菌引入质粒失败的案例作为研究对象, 预测两种进化力量“突变”和“选择”是如何使得外源基因失活的。如“进化失败模式 (EFM) 计算器”^[65]给质粒提供相对不稳定性预测 (RIP) 评分; 在体外计算细胞生长速率降低的程度^[66-67]来定量预测基因的表达负担。这些进化预测工具可以明确基因线路的设计限制, 用重构或替代的方法来帮助完成基因设计。

2 人工生命系统帮助探索进化规律

合成生物学是一门人工设计和构建生命系统的工程学科, 它能为研究进化原理提供绝佳的研究对象。使用各种合成生物工具, 可以构建独立且定量的系统, 测试特定生物系统基因型-表型的映射关系, 实现分子和系统层面上的进化理论研究。这不仅可以帮助了解进化的起点、各种生物大分子对构成原始细胞的作用; 还能帮助了解进化的分子机制和驱动因素。此外人工系统也能帮助完善进化的基础理论, 更好地解释生命现象^[68]。而人工生命群落在复杂环境中的相互作用和进化过程也有助于理解复杂生命系统。

2.1 通过合成生物学手段理解生命起源

自然界的进化已持续数十亿年, 生物进化的起点是地球形成早期由非生命物质到生命结构的转化, 即生命起源 [图4(a)]。虽然已无法观察到生命起源, 但合成定量手段可以帮助构建人造生命, 模拟生命最初形态, 对了解生命产生的条件和进化起点提供重要帮助。如化学合成方法可以采用全合成的方法搭建自然界不存在的生物大分子和生命系统, 通过对比自然生物的结构与功能, 可在一定程度上还原自然界生命起源的条件和过程。除此之外, 通过引入或删除生物大分子来探索现有经典分子的必要性, 帮助人类更加了解生命构成和必要功能。

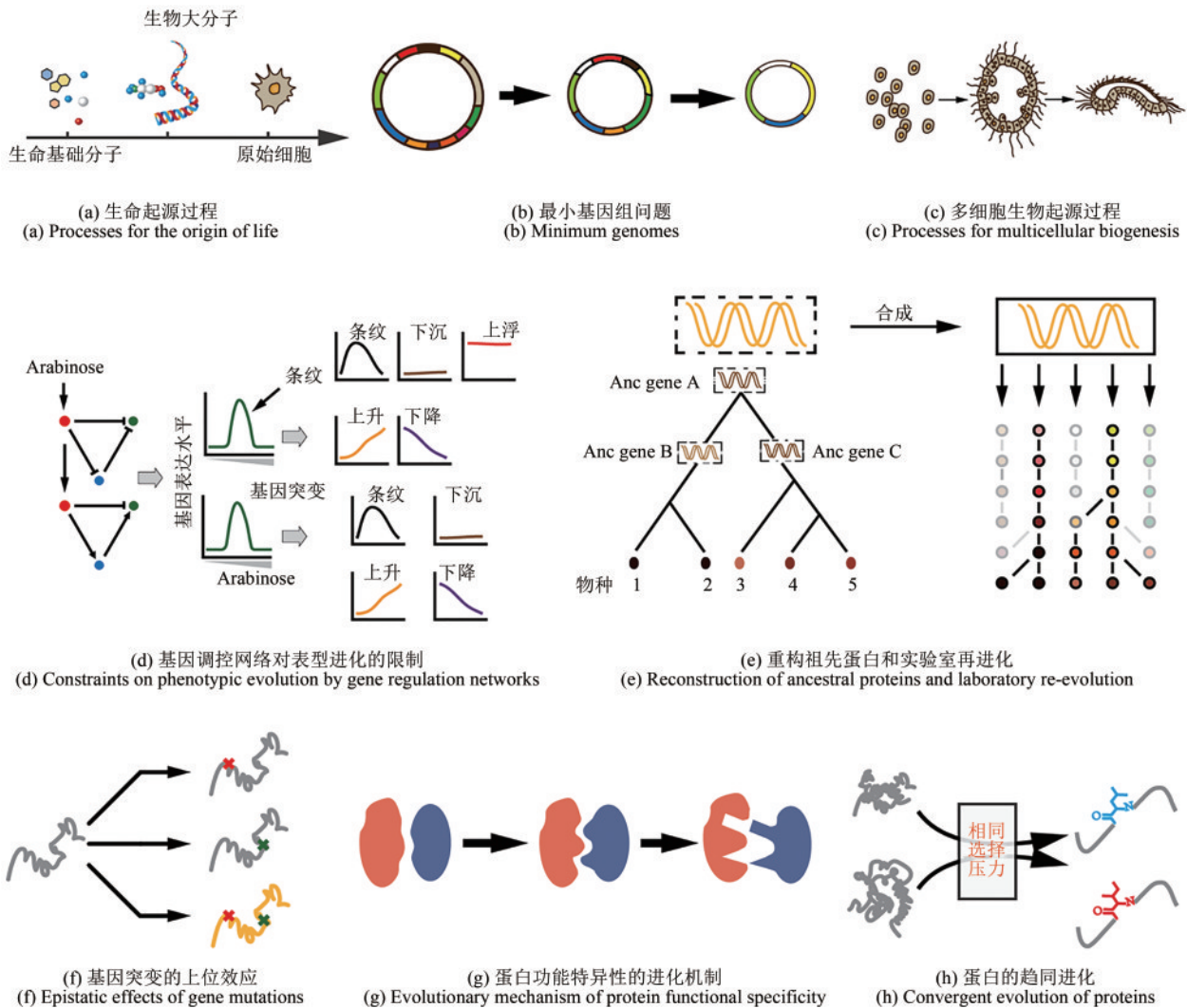


图4 人工合成生物系统帮助探索进化规律

Fig. 4 Synthetic biological systems help exploring laws of evolution

研究进化重点之一是生命起源前分子：核酸和蛋白质是偶然产生的还是自然选择的必然结果；替换或增减这些生物大分子是否会对生命体执行功能造成损害。这些问题都能通过合成基因组来探究。如Chin等^[69]只用了61个密码子重编码了大肠杆菌的全基因组，余下的3个密码子被用于编码非天然氨基酸，成功构建了一株活性和功能良好的大肠杆菌。而美国研究团队通过增加两对新的核酸碱基对S: B和Z: P，成功构建了八碱基遗传生物体Hachimoji^[70]。这些研究都证明了中心法则并不局限于4个DNA碱基和20个氨基酸，帮助我们更好地理解生物大分子的选择和构建原理。

另一个进化起点问题是原始细胞如何存活并实现功能。这可通过合成生物学的构建最小基因组、合成人工细胞等研究得到启示 [图4(b)]。最小基因组很可能是最初的细胞具备的基因组，可以帮助了解生命的基本需求和进化起点。目前Pelletier等^[71]已成功创造出世界上第1种可以正常生长和分裂的人工合成细胞——JCVI-syn3A，它由原来Venter团队^[72]构造的只含有473个基因的JCVI-syn3.0进化而来，可正常生长和分裂，这有助于了解生命的本质。

合成生物学手段还可以帮助了解多细胞生物的起源。如同生命的起源一样，从单细胞生物发展为多细胞生物也是生物进化的里程碑事件 [图4(c)]。

多细胞生物相比于单细胞生物学在一些条件下更具生存优势, 比如对抗捕食者和细胞间分工合作 (division of labor) 等^[73], 因此很可能是自然选择的结果。实际上, 多细胞生物的起源在生命树上至少独立发生了25次^[74]。关于多细胞生物的最初起源存在两种理论: 聚集性多细胞起源 (aggregative multicellularity) 和克隆性多细胞起源 (clonal multicellularity)。尽管目前现存的多细胞生物基本上是克隆性多细胞体, 实际上很多微生物可以通过聚集产生细胞的分化, 比如黏细菌^[75]。通过细胞的聚集和产生分化一般认为是单细胞微生物抵抗逆境的一种生存策略。但是由于聚集体内细胞遗传背景的多样性, “背叛者” (cheater) 也容易产生^[76]。盘基网柄菌 (*Dictyostelium discoideum*) 进化出了一种“亲缘识别”的分子机制: *lag* 基因家族类似于哺乳动物中的MHC基因, 存在高度的遗传多态性, 对遗传差异比较大的细胞会产生排斥作用^[76]。克隆性多细胞生物体因为可以降低细胞间遗传多样性, 可以有效防止“背叛者”的发生, 因而可以进化出更丰富的细胞表型和分工, 可能更具有达尔文选择优势。Ratcliff等^[77-78]设计了一个有利于单细胞生物聚集的培养条件, 利用实验进化的手段对酿酒酵母进行连续传代和选择, 经过60次传代后发现酵母都进化出雪花状 (snowflake-like) 表型, 进一步遗传学实验表明受到选择的酵母细胞由于ACE2突变而导致不完全裂解的细胞分裂, 细胞间的黏附能力大大增强。同时, 这些聚集性的酵母都是单克隆的。实验室进化揭示了多细胞生物形成的一种可能过程, 即环境选择了细胞分裂受损而产生的克隆性多细胞生命体。多细胞生物的一个典型特征是细胞形成空间上的表型分化和合作关系, 通过合成生物学方法, 可以构建回路实现单细胞群体的表型分化或自发形成特定的空间特征, 即图灵图案 (Turing pattern)^[79-80], 尽管单细胞生物的图灵图案不代表多细胞生物的产生, 却能在一定程度上反映多细胞生物起源后细胞分化的可能产生机制^[81]。

2.2 利用合成生物学和定量模型手段解析分子与系统进化机制

分子进化是对系统、个体在分子层面的突变、

选择等进化机制提出合理解释。近年来, 科学家使用合成定量手段可以构建相应的基因回路, 并建立数学模型描述关键变量的影响, 研究生物进化的分子机制, 从而归纳推广成进化的基本原理用于生命系统的解释^[82]。本节主要关注以下几个分子进化问题: 不同的基因调控网络如何影响进化方向、祖先蛋白构建及实验室重进化、群落间的相互作用如何影响进化动力学、进化系统是否有最优解及实现方式。

关于基因调控网络, 用合成基因线路可以来理解进化约束 (evolutionary constraint) 和上位效应 (epistasis)。如Schaeerli等^[83]构建了两种表型一致但调控回路不同的大肠杆菌, 细胞可以随阿拉伯糖 (arabinose) 浓度的变化表达不同水平的GFP, 表达图谱形成不同的空间条纹分布, 然后引入随机突变探究突变如何影响新表型产生的能力。结果两种线路在突变后各能产生5种和4种基因表达表型 [图4(d)]。实验证明了具有相同表型但不同基因拓扑调控机制的系统在突变后可能产生不一样的表型结果, 并可能限制生物体进化多样性或适应性的能力。Lagator等^[84]在大肠杆菌中建立了一个简单的合成调控系统, 由两个反式调控元件和一个顺式调控元件组成, 三者相互作用且共同调控了荧光蛋白的表达。然后他们给这3个调控元件引入随机突变, 发现当只有1个元件产生突变时没有新的基因表达表型产生; 当在2种元件中同时引入突变时, 调控系统产生了新的基因表达表型。作者将这些新表型归因于调控线路之间的上位效应, 这可能会促进随后的适应性进化。

大多时候我们无法直接观察分子进化的发生过程, 但是通过生物信息学算法我们可以构建同源蛋白的祖先序列, 在实验室合成祖先蛋白并进行功能表征和再进化 [图4(e)], 系统探究分子进化的路径、约束性、偶然性、上位效应和蛋白特异性功能的形成过程 [图4(e)~(h)], 为理解分子进化的机制提供直接的证据^[85]。比如通过构建祖先蛋白和实验室重进化, 可以鉴定基因突变对蛋白或调控网络功能的上位效应 [图4(f)]^[85-87]。Starr等^[88]构建了单鞭毛生物 (Amorphea) 的热休克蛋白Hsp90在10亿年内的261个祖先序列, 并在酿酒酵母中合成和测试了这些祖先序列的功能,

发现超过半数的祖先 Hsp90 蛋白使酵母的适应性下降, 由此可推断现存物种 Hsp90 蛋白中的一些额外突变与祖先序列产生上位效应, 抵消了祖先序列变异的有害性, 因此证明了上位效应的普遍性和在物种分子进化中的重要作用。利用祖先蛋白重构方法揭示了蛋白特异性功能的产生机制。比如, Clifton 等^[89]发现谷氨酰胺特异性结合蛋白的祖先蛋白还可以结合精氨酸, 揭示了从祖先多功能蛋白到特异功能蛋白的进化过程。

构建合成微生物群落可以探究群落间的相互作用及其对进化动力学规律。如 Fraebel 等^[90]利用细菌迁移速度的进化模型, 解释了细菌的生长速率和运动速率之间的权衡会限制种群的趋化迁移速度的进化; 游凌冲团队^[91]则对该现象做了进一步解释: 他们利用合成基因回路在菌落中形成自组织的核环结构, 发现在菌落扩张形成的环的宽度与细菌大小之间存在着完美的固定比例, 并揭示了一个集群的空间感应机制。同时, 合成生命系统可以被用来观察生物网络如何影响进化方向和导致进化权衡 (evolutionary tradeoff) 的。如 Poelwijk 等^[92]在大肠杆菌内构建了一个包含两种基因操纵子的调控系统, 它们的表达量受两种相反的诱导因子的调控。在含蔗糖的培养基中, 操纵子被诱导后会降低细菌生长速率; 在含抗生素的培养基中, 操纵子被诱导后会提高生长速率。随后他们引入随机突变并通过反复改变环境来分析两个对立环境下适应性的进化权衡, 发现系统产生了全新的调控蛋白来克服约束, 并产生了最优化表型。这证明了一个调控系统在面临不同环境时可进化为最优调控反应。另一个例子则探索了个体的生物网络最优性: Çağatay 等^[93]设计合成了一个与自然系统完全相反的负反馈回路来调节枯草芽孢杆菌的感受态系统, 发现合成系统的噪声动力学分布与天然系统不同, 这导致了合成系统在感受态发生的时机控制不够精准。这说明不同的拓补网络可能根据基因结构的噪声特点进行选择, 一些情况下天然系统更具有进化优势。

2.3 合成生物学帮助了解复杂环境中的共进化

进化的基础理论研究最终是为了解释复杂生

命系统中的自然现象。而合成生物学可以通过“设计-构建-测试-学习”的模式构建简单的合成生命系统, 将复杂的生命现象抽离并简化。通过调控基因线路模拟突变、修改环境参数模拟外界压力, 人工模型可用来观察复杂生命系统将如何进化; 建立定量的生物学模型予以解释, 可揭示进化的普遍规律并推广到复杂生命系统。

生命系统所处的外部环境对其有着重要影响, 有时可直接决定整体的进化方向。其中一个有趣的问题是快速变化的环境如何对生物体产生影响, 生物又是如何通过快速进化来适应环境。合成和定量手段可以构造生命系统并改变环境来用于该问题的研究。如 Gerland 团队^[94]通过改变营养物供应模拟环境快速变化, 构建了一种从稳态生长到动态变化的细菌模型, 他们用一个自上而下的调节稳态模型的方程来定量理解生物适应环境变化的机制。而 Ma 和 Isaacs^[95]则利用噬菌体做了一个有趣的进化实验, 他们改变了噬菌体宿主的遗传密码然后用天然的噬菌体侵染, 结果发现噬菌体会快速调整并改变自身遗传信息来适应宿主, 以实现与宿主的兼容。在极端环境中, 生物进化出可存活的表型后, 是否还应保持多样性也是生物面临的一个难题。Bódi 等^[96]在酿酒酵母中设计了两种合成线路, 来控制不同程度的基因表达异质性和耐药性。他们发现, 具有较高异质性的合成线路不仅对抗真菌药物有较高的初始耐药性, 而且允许酵母细胞在几轮筛选后进化出较高的耐药性。而高异质性合成线路的有益突变取决于基因的高表达随机性, 即在基因表达低随机性的背景下, 它们的适应性将大大降低。这也就回答了稳定性和适应性间的矛盾: 生物在面对极端环境时, 可以通过高表达随机性来获取有利突变, 从而更好地适应进化。

除了环境对生命系统的进化有重要影响外, 群落间和种群内部的关系也在很大程度上决定生态位的划分以及进化方向。合成微生物群落可以构建简单的人工生态系统, 通过调控细菌群落的比例、代谢结构和多样性等, 可在复杂度低且可控性高的系统中重复观察群落的空间分布和资源分配, 建立共进化的定量模型, 为解释复杂生态系统的进化机制搭建桥梁。刘陈立课题组^[7]利用

合成生物改造的菌株在不同培养环境中反复“演绎”细菌竞争迁移的过程，得到了种群竞争迁移的空间分布规律。他们提出了群落共进化策略：生长繁殖和迁移速率不同的群落，能自发在不同空间定植并稳定共存。

进一步推广到人体，这是一个环境和生物间相互作用都极其复杂的生态系统，研究其内部的进化对了解和治疗疾病有重要意义。合成手段可以帮助构建体外模型来模拟人体内环境，可控的系统也便于展开清晰定量的研究；另一种方式是将工程改造的可控细胞或细菌输入体内，观察其如何与环境作用并进化。在肿瘤中，肿瘤细胞和微环境中的组分（如免疫细胞、瘤内菌等）都会发生适应性进化。2020年发表在 *Science* 上迄今最全面的肿瘤微生物组研究表明，细菌在不同肿瘤中广泛存在并形成独特群落，其代谢能力和优势生态位会随着宿主环境改变而发生适应性进化^[97]。想要进一步研究瘤内菌群和肿瘤的进化关系，工程菌和肿瘤类器官的组合似乎是个不错的选择：目前已有多种 CRISPR 技术用于类器官编辑中^[98]，在体外模拟肿瘤异质性，探究特定亚群肿瘤在面对生存压力（如药物治疗）时如何产生适应性进化；而利用标记的工程菌，可以代替自然菌群在体内或体外肿瘤模型中进行模拟进化，探究其在肿瘤发病机制中的作用及耐药性进化进程^[99]。

3 总结与展望

综上所述，定量合成生物学和进化系统生物学的交叉结合已被用于各个领域的研究，并展现出广阔的科研潜力和应用前景^[100-101]。一方面，进化论的原理和规律可以应用于指导合成生物学的设计和定量生物模型的建立^[102]，诸如定向进化等进化法则可以在不了解具体作用机制的前提下筛选出符合要求的合成生物；而博弈论和进化动力学等进化规律可以帮助理性设计合成生命系统，如目前一些进化预测模型可实现虚拟进化，大大减少人为“设计-构建-测试-学习”得到目标工程生物的时间。另一方面，通过合成手段和定量规律能为进化规律的探索提供简单可控的平台，如人工合成的基因组和细胞可被用于探索生命起源

的过程和必要条件、合成微生物群落间的互动可用来观察进化的分子机制和动力学规律、体外模拟复杂生态系统（如肿瘤）的人工模型有助于窥探并厘清生物体内各组分的相互作用和进化轨迹，为疾病的治疗提供指导。未来，合成系统和进化思维有望实现更紧密的结合，两者的相辅相成将对生命的理解和创造（即“造物致知”和“造物致用”）有巨大的帮助。

除此之外，进化理论作为生物学最基础的原理之一，还可与各个新兴学科交叉融合，提高合成生物学可预测性和理性设计原理的研究。相比基于黑箱模型的机器学习方法，生物进化思想可在动态建模中提供更佳方案，它充分考虑了生物在分子、个体和系统层面上的持续动态变化^[103]。例如分子进化思想可帮助预测复杂遗传线路的稳定性^[104]；进化动力学和博弈论可帮助预测和设计人工合成细菌群落的分布比例^[105-108]。进化思想还能应用到工程、环境和材料生物学的合成设计中^[109]，目前已有许多工程改造的生物用于工农业生产或环境处理中，然而暴露在自然环境中的人造生命是否能保持功能稳定？自然进化又会对其产生怎样的影响？这些还都没有明确且普遍适用的说明。如何结合进化原理来理性审慎地设计工程生物，如利用人为可控的基因和染色体装置使其在不同环境下自我优化，也是应用合成技术在工程和环境科学的一个重要研究方向。

过去数十年，进化生物学、合成生物学和定量生物学经历了从诞生、发展到逐渐成熟的过程。三者交叉融合的研究日益增多，但仍未有一个明确清晰的理论阐明三者间的关系。其实无论是合成生物学领域的人工合成生命和群落，还是基于数学理论的定量生物学模型，都应该充分考虑进化这种动态因素，才能更好地模拟生命必须遵守的自然规律，这样真实的模型应用在科学研究中才是有意义、有效果的。因此我们提出用生物进化的视角探索定量规律和人工生命的展望，为实现合成生物学的理性设计发挥进化的力量。

参 考 文 献

- [1] 刘陈立, 汤超, 汤雷翰, 等. 定量至简, 工程至繁: 定量工程生物学[J]. 科学通报, 2021, 66(3): 261-263.

- LIU C L, TANG C, TANG L H, et al. Quantifying to simplicity, engineering to complexity: quantitative engineering biology[J]. Chinese Science Bulletin, 2021, 66(3): 261-263.
- [2] SLEIGHT S C, SAURO H M. Visualization of evolutionary stability dynamics and competitive fitness of *Escherichia coli* engineered with randomized multigene circuits[J]. ACS Synthetic Biology, 2013, 2(9): 519-528.
- [3] KURLAND C G, DONG H. Bacterial growth inhibition by overproduction of protein[J]. Molecular Microbiology, 1996, 21(1): 1-4.
- [4] CARBONELL-BALLESTERO M, GARCIA-RAMALLO E, MONTAÑEZ R, et al. Dealing with the genetic load in bacterial synthetic biology circuits: convergences with the Ohm's law[J]. Nucleic Acids Research, 2015, 44(1): 496-507.
- [5] QIAN Y L, HUANG H H, JIMÉNEZ J I, et al. Resource competition shapes the response of genetic circuits[J]. ACS Synthetic Biology, 2017, 6(7): 1263-1272.
- [6] WANG Y J, XUE P, CAO M F, et al. Directed evolution: methodologies and applications[J]. Chemical Reviews, 2021, 121(20): 12384-12444.
- [7] LIU W R, CREMER J, LI D J, et al. An evolutionarily stable strategy to colonize spatially extended habitats[J]. Nature, 2019, 575(7784): 664-668.
- [8] ELLIS T. Predicting how evolution will beat us[J]. Microbial Biotechnology, 2019, 12(1): 41-43.
- [9] MILLS D R, PETERSON R L, SPIEGELMAN S. An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule[J]. Proceedings of the National Academy of Sciences of the United States of America, 1967, 58(1): 217-224.
- [10] ESVELT K M, CARLSON J C, LIU D R. A system for the continuous directed evolution of biomolecules[J]. Nature, 2011, 472(7344): 499-503.
- [11] 蒋迎迎, 曲戈, 孙周通. 机器学习助力酶定向进化[J]. 生物学杂志, 2020, 37(4): 1-11.
- JIANG Y Y, QU G, SUN Z T. Machine learning-assisted enzyme directed evolution[J]. Journal of Biology, 2020, 37(4): 1-11.
- [12] RIX G, LIU C C. Systems for *in vivo* hypermutation: a quest for scale and depth in directed evolution[J]. Current Opinion in Chemical Biology, 2021, 64: 20-26.
- [13] 陈珏, 黄佳敏, 燕天鹤, 等. 随机突变文库构建与筛选研究进展[J]. 生物工程学报, 2021, 37(1): 163-177.
- CHEN J, HUANG J M, YAN T H, et al. Progress in the construction and screening of random mutation library[J]. Chinese Journal of Biotechnology, 2021, 37(1): 163-177.
- [14] 郭园, 赵仲麟. 微生物系统定向进化与合成生物学应用研究进展[J]. 生物技术通报, 2017, 33(1): 76-82.
- GUO Yuan, ZHAO Zhonglin. Advances on applications of synthetic biology and directed evolution in microbial systems[J]. Biotechnology Bulletin, 2017, 33(1): 76-82.
- [15] CARLSON J C, BADRAN A H, GUGGIANA-NILO D A, et al. Negative selection and stringency modulation in phage-assisted continuous evolution[J]. Nature Chemical Biology, 2014, 10(3): 216-222.
- [16] 翟昊天, 祁庆生, 侯进. 体内连续进化技术的研究进展[J]. 生物工程学报, 2021, 37(2): 486-499.
- ZHAI H T, QI Q S, HOU J. Recent advances of continuous *in vivo* evolution[J]. Chinese Journal of Biotechnology, 2021, 37(2): 486-499.
- [17] FIJALKOWSKA I J, SCHAAPER R M. Mutants in the Exo I motif of *Escherichia coli* dnaQ: defective proofreading and inviability due to error catastrophe[J]. Proceedings of the National Academy of Sciences of the United States of America, 1996, 93(7): 2856-2861.
- [18] HESS G T, FRÉSARD L, HAN K, et al. Directed evolution using dCas9-targeted somatic hypermutation in mammalian cells[J]. Nature Methods, 2016, 13(12): 1036-1042.
- [19] KOMOR A C, KIM Y B, PACKER M S, et al. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage[J]. Nature, 2016, 533(7603): 420-424.
- [20] COX E C, HORNER D L. Structure and coding properties of a dominant *Escherichia coli* mutator gene, *mutD*[J]. Proceedings of the National Academy of Sciences of the United States of America, 1983, 80(8): 2295-2299.
- [21] SHIMODA C, ITADANI A, SUGINO A, et al. Isolation of thermotolerant mutants by using proofreading-deficient DNA polymerase delta as an effective mutator in *Saccharomyces cerevisiae*[J]. Genes & Genetic Systems, 2006, 81(6): 391-397.
- [22] RAVIKUMAR A, ARZUMANYAN G A, OBADI M K A, et al. Scalable, continuous evolution of genes at mutation rates above genomic error thresholds[J]. Cell, 2018, 175(7): 1946-1957.e13.
- [23] THURONYI B W, KOBLAN L W, LEVY J M, et al. Continuous evolution of base editors with expanded target compatibility and improved activity[J]. Nature Biotechnology, 2019, 37(9): 1070-1079.
- [24] HU J H, MILLER S M, GEURTS M H, et al. Evolved Cas9 variants with broad PAM compatibility and high DNA specificity[J]. Nature, 2018, 556(7699): 57-63.
- [25] BLUM T R, LIU H, PACKER M S, et al. Phage-assisted evolution of botulinum neurotoxin proteases with reprogrammed specificity[J]. Science, 2021, 371(6531): 803-810.
- [26] WANG T N, BADRAN A H, HUANG T P, et al. Continuous directed evolution of proteins with improved soluble expression[J]. Nature Chemical Biology, 2018, 14(10): 972-980.
- [27] ROTH T B, WOOLSTON B M, STEPHANOPOULOS G, et al. Phage-assisted evolution of bacillus methanolicus methanol dehydrogenase 2[J]. ACS Synthetic Biology, 2019, 8(4): 796-806.

- [28] MA Y Q, ZHANG J Y, YIN W J, et al. Targeted AID-mediated mutagenesis (TAM) enables efficient genomic diversification in mammalian cells[J]. *Nature Methods*, 2016, 13(12): 1029-1035.
- [29] GAUDELLI N M, KOMOR A C, REES H A, et al. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage[J]. *Nature*, 2017, 551(7681): 464-471.
- [30] ANZALONE A V, RANDOLPH P B, DAVIS J R, et al. Search-and-replace genome editing without double-strand breaks or donor DNA[J]. *Nature*, 2019, 576(7785): 149-157.
- [31] ZHONG Z W, RAVIKUMAR A, LIU C C. Tunable expression systems for orthogonal DNA replication[J]. *ACS Synthetic Biology*, 2018, 7(12): 2930-2934.
- [32] ARZUMANYAN G A, GABRIEL K N, RAVIKUMAR A, et al. Mutually orthogonal DNA replication systems *in vivo*[J]. *ACS Synthetic Biology*, 2018, 7(7): 1722-1729.
- [33] JAVANPOUR A A, LIU C C. Genetic compatibility and extensibility of orthogonal replication[J]. *ACS Synthetic Biology*, 2019, 8(6): 1249-1256.
- [34] YANG K K, WU Z, ARNOLD F H. Machine-learning-guided directed evolution for protein engineering[J]. *Nature Methods*, 2019, 16(8): 687-694.
- [35] WU Z, KAN S B J, LEWIS R D, et al. Machine learning-assisted directed protein evolution with combinatorial libraries[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2019, 116(18): 8852-8858.
- [36] MUSDAL Y, GOVINDARAJAN S, MANNERVIK B. Exploring sequence-function space of a poplar glutathione transferase using designed information-rich gene variants[J]. *Protein Engineering, Design and Selection*, 2017, 30(8): 543-549.
- [37] ROMERO P A, KRAUSE A, ARNOLD F H. Navigating the protein fitness landscape with Gaussian processes[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2013, 110(3): E193-E201.
- [38] HAMEDIRAD M, CHAO R, WEISBERG S, et al. Towards a fully automated algorithm driven platform for biosystems design[J]. *Nature Communications*, 2019, 10: 5150.
- [39] BASHIR A, YANG Q, WANG J P, et al. Machine learning guided aptamer refinement and discovery[J]. *Nature Communications*, 2021, 12: 2366.
- [40] Knight C G, Platt M, Rowe W, et al. Array-based evolution of DNA aptamers allows modelling of an explicit sequence-fitness landscape[J]. *Nucleic Acids Res*, 2009, 37(1): e6.
- [41] SCHISSEL C K, MOHAPATRA S, WOLFE J M, et al. Deep learning to design nuclear-targeting abiotic miniproteins[J]. *Nature Chemistry*, 2021, 13(10): 992-1000.
- [42] BEDBROOK C N, YANG K K, RICE A J, et al. Machine learning to design integral membrane channelrhodopsins for efficient eukaryotic expression and plasma membrane localization[J]. *PLoS Computational Biology*, 2017, 13(10): e1005786.
- [43] BEDBROOK C N, YANG K K, ROBINSON J E, et al. Machine learning-guided channelrhodopsin engineering enables minimally invasive optogenetics[J]. *Nature Methods*, 2019, 16(11): 1176-1184.
- [44] LIAO M J, DIN M O, TSIMRING L, et al. Rock-paper-scissors: engineered population dynamics increase genetic stability[J]. *Science*, 2019, 365(6457): 1045-1049.
- [45] SHAHAB R L, BRETHAUER S, DAVEY M P, et al. A heterogeneous microbial consortium producing short-chain fatty acids from lignocellulose[J]. *Science*, 2020, 369(6507): eabb1214.
- [46] BUSKIRK S W, PEACE R E, LANG G I. Hitchhiking and epistasis give rise to cohort dynamics in adapting populations[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2017, 114(31): 8330-8335.
- [47] RUGBJERG P, MYLING-PETERSEN N, PORSE A, et al. Diverse genetic error modes constrain large-scale bio-based production[J]. *Nature Communications*, 2018, 9: 787.
- [48] RACKHAM O, CHIN J W. A network of orthogonal ribosome-mRNA pairs[J]. *Nature Chemical Biology*, 2005, 1(3): 159-166.
- [49] DEL VECCHIO D. Modularity, context-dependence, and insulation in engineered biological circuits[J]. *Trends in Biotechnology*, 2015, 33(2): 111-119.
- [50] ISAACS F J, DWYER D J, DING C M, et al. Engineered riboregulators enable post-transcriptional control of gene expression[J]. *Nature Biotechnology*, 2004, 22(7): 841-847.
- [51] UMENHOFFER K, FEHÉR T, BALIKÓ G, et al. Reduced evolvability of *Escherichia coli* MDS42, an IS-less cellular chassis for molecular and synthetic biology applications[J]. *Microbial Cell Factories*, 2010, 9: 38.
- [52] SUÁREZ G A, RENDA B A, DASGUPTA A, et al. Reduced mutation rate and increased transformability of transposon-free acinetobacter baylyi ADP1-ISx[J]. *Applied and Environmental Microbiology*, 2017, 83(17): e01025-e01017.
- [53] GENG P, LEONARD S P, MISHLER D M, et al. Synthetic genome defenses against selfish DNA elements stabilize engineered bacteria against evolutionary failure[J]. *ACS Synthetic Biology*, 2019, 8(3): 521-531.
- [54] CHAN C T Y, LEE J W, CAMERON D E, et al. 'Deadman' and 'Passcode' microbial kill switches for bacterial containment[J]. *Nature Chemical Biology*, 2016, 12(2): 82-86.
- [55] LIANG P X, ZHANG Y T, XU B, et al. Deletion of genomic islands in the *Pseudomonas putida* KT2440 genome can create an optimal chassis for synthetic biology applications[J]. *Microbial Cell Factories*, 2020, 19(1): 70.
- [56] VON KAMP A, KLAMT S. Growth-coupled overproduction is feasible for almost all metabolites in five major production organisms[J]. *Nature Communications*, 2017, 8: 15956.
- [57] GODARA A, KAO K C. Adaptive laboratory evolution for growth coupled microbial production[J]. *World Journal of Mi-*

- crobiology & Biotechnology, 2020, 36(11): 175.
- [58] FONG S S, BURGARD A P, HERRING C D, et al. In silico design and adaptive evolution of *Escherichia coli* for production of lactic acid[J]. Biotechnology and Bioengineering, 2005, 91(5): 643-648.
- [59] RUGBJERG P, SARUP-LYTZEN K, NAGY M, et al. Synthetic addiction extends the productive life time of engineered *Escherichia coli* populations[J]. Proceedings of the National Academy of Sciences of the United States of America, 2018, 115(10): 2347-2352.
- [60] SON H I, WEISS A, YOU L C. Design patterns for engineering genetic stability[J]. Current Opinion in Biomedical Engineering, 2021, 19: 100297.
- [61] DAI Z J, LEE A J, ROBERTS S, et al. Versatile biomanufacturing through stimulus-responsive cell-material feedback[J]. Nature Chemical Biology, 2019, 15(10): 1017-1024.
- [62] SEGALL-SHAPIRO T H, SONTAG E D, VOIGT C A. Engineered promoters enable constant gene expression at any copy number in bacteria[J]. Nature Biotechnology, 2018, 36(4): 352-358.
- [63] LIU Q J, SCHUMACHER J, WAN X Y, et al. Orthogonality and burdens of heterologous AND gate gene circuits in *E. coli*[J]. ACS Synthetic Biology, 2018, 7(2): 553-564.
- [64] CASTLE S D, GRIERSON C S, GOROCHOWSKI T E. Towards an engineering theory of evolution[J]. Nature Communications, 2021, 12: 3326.
- [65] JACK B R, LEONARD S P, MISHLER D M, et al. Predicting the genetic stability of engineered DNA sequences with the EFM calculator[J]. ACS Synthetic Biology, 2015, 4(8): 939-943.
- [66] CERONI F, ALGAR R, STAN G B, et al. Quantifying cellular capacity identifies gene expression designs with reduced burden[J]. Nature Methods, 2015, 12(5): 415-418.
- [67] BORKOWSKI O, BRICIO C, MURGIANO M, et al. Cell-free prediction of protein expression costs for growing cells[J]. Nature Communications, 2018, 9: 1457.
- [68] BAIER F, SCHAERLI Y. Addressing evolutionary questions with synthetic biology[M]// CROMBACH A. Evolutionary Systems Biology-Advances, Questions, and Opportunities. 2nd ed. Villeurbanne, France: Springer Press, 2021:135-157.
- [69] FREDENS J, WANG K H, DE LA TORRE D, et al. Total synthesis of *Escherichia coli* with a recoded genome[J]. Nature, 2019, 569(7757): 514-518.
- [70] HOSHIKA S, LEAL N A, KIM M J, et al. Hachimoji DNA and RNA: a genetic system with eight building blocks[J]. Science, 2019, 363(6429): 884-887.
- [71] PELLETIER J F, SUN L J, WISE K S, et al. Genetic requirements for cell division in a genomically minimal cell[J]. Cell, 2021, 184(9): 2430-2440.e16.
- [72] HUTCHISON C A 3rd, CHUANG R Y, NOSKOV V N, et al. Design and synthesis of a minimal bacterial genome[J]. Science, 2016, 351(6280): aad6253.
- [73] GROSBERG R K, STRATHMANN R R. The evolution of multicellularity: a minor major transition? [J]. Annual Review of Ecology, Evolution, and Systematics, 2007, 38: 621-654.
- [74] NIKLAS K J. The evolutionary-developmental origins of multicellularity[J]. American Journal of Botany, 2014, 101(1): 6-25.
- [75] KAISER D, ROBINSON M, KROOS L. Myxobacteria, polarity, and multicellular morphogenesis[J]. Cold Spring Harbor Perspectives in Biology, 2010, 2(8): a000380.
- [76] BENABENTOS R, HIROSE S, SUCGANG R, et al. Polymorphic members of the lag gene family mediate kin discrimination in dictyostelium[J]. Current Biology, 2009, 19(7): 567-572.
- [77] RATCLIFF W C, DENISON R F, BORRELLO M, et al. Experimental evolution of multicellularity[J]. Proceedings of the National Academy of Sciences of the United States of America, 2012, 109(5): 1595-1600.
- [78] RATCLIFF W C, FANKHAUSER J D, ROGERS D W, et al. Origins of multicellular evolvability in snowflake yeast[J]. Nature Communications, 2015, 6: 6102.
- [79] LANDGE A N, JORDAN B M, DIEGO X, et al. Pattern formation mechanisms of self-organizing reaction-diffusion systems[J]. Developmental Biology, 2020, 460(1): 2-11.
- [80] LIU C L, FU X F, LIU L Z, et al. Sequential establishment of stripe patterns in an expanding cell population[J]. Science, 2011, 334(6053): 238-241.
- [81] SOLÉ R, OLLÉ-VILA A, VIDIELLA B, et al. The road to synthetic multicellularity[J]. Current Opinion in Systems Biology, 2018, 7: 60-67.
- [82] 肖敏凤, 张炳照, 刘陈立. 合成生物学在生命起源、进化、结构和功能相互关系研究中的作用[J]. 中国科学: 生命科学, 2015, 45(10): 915-927.
- XIAO M F, ZHANG B Z, LIU C L. Synthetic biology in studying the origin of life, evolution, and structure-function relation[J]. Scientia Sinica (Vitae), 2015, 45(10): 915-927.
- [83] SCHAERLI Y, JIMÉNEZ A, DUARTE J M, et al. Synthetic circuits reveal how mechanisms of gene regulatory networks constrain evolution[J]. Molecular Systems Biology, 2018, 14(9): e8102.
- [84] LAGATOR M, SARIKAS S, ACAR H, et al. Regulatory network structure determines patterns of intermolecular epistasis[J]. eLife, 2017, 6: e28921.
- [85] HARMS M J, THORNTON J W. Evolutionary biochemistry: revealing the historical and physical causes of protein properties[J]. Nature Reviews Genetics, 2013, 14(8): 559-571.
- [86] STARR T N, THORNTON J W. Epistasis in protein evolution[J]. Protein Science, 2016, 25(7): 1204-1218.
- [87] STARR T N, PICTON L K, THORNTON J W. Alternative evo-

- lutionary histories in the sequence space of an ancient protein[J]. *Nature*, 2017, 549(7672): 409-413.
- [88] STARR T N, FLYNN J M, MISHRA P, et al. Pervasive contingency and entrenchment in a billion years of Hsp90 evolution[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2018, 115(17): 4453-4458.
- [89] CLIFTON B E, JACKSON C J. Ancestral protein reconstruction yields insights into adaptive evolution of binding specificity in solute-binding proteins[J]. *Cell Chemical Biology*, 2016, 23(2): 236-245.
- [90] FRAEBEL D T, MICKALIDE H, SCHNITKEY D, et al. Environment determines evolutionary trajectory in a constrained phenotypic space[J]. *eLife*, 2017, 6: e24669.
- [91] CAO Y, RYSER M D, PAYNE S, et al. Collective space-sensing coordinates pattern scaling in engineered bacteria[J]. *Cell*, 2016, 165(3): 620-630.
- [92] POELWIJK F J, DE VOS M G J, TANS S J. Tradeoffs and optimality in the evolution of gene regulation[J]. *Cell*, 2011, 146(3): 462-470.
- [93] ÇAĞATAY T, TURCOTTE M, ELOWITZ M B, et al. Architecture-dependent noise discriminates functionally analogous differentiation circuits[J]. *Cell*, 2009, 139(3): 512-522.
- [94] ERICKSON D W, SCHINK S J, PATSALO V, et al. A global resource allocation strategy governs growth transition kinetics of *Escherichia coli*[J]. *Nature*, 2017, 551(7678): 119-123.
- [95] MA N J, ISAACS F J. Genomic recoding broadly obstructs the propagation of horizontally transferred genetic elements[J]. *Cell Systems*, 2016, 3(2): 199-207.
- [96] BÓDI Z, FARKAS Z, NEVOZHAY D, et al. Phenotypic heterogeneity promotes adaptive evolution[J]. *PLoS Biology*, 2017, 15(5): e2000644.
- [97] NEJMAN D, LIVYATAN I, FUKS G, et al. The human tumor microbiome is composed of tumor type-specific intracellular bacteria[J]. *Science*, 2020, 368(6494): 973-980.
- [98] HOFER M, LUTOLF M P. Engineering organoids[J]. *Nature Reviews Materials*, 2021, 6(5): 402-420.
- [99] SEPICH-POORE G D, ZITVOGEL L, STRAUSSMAN R, et al. The microbiome and human cancer[J]. *Science*, 2021, 371(6536): eabc4552.
- [100] PEISAJOVICH S G. Evolutionary synthetic biology[J]. *ACS Synthetic Biology*, 2012, 1(6): 199-210.
- [101] TAYLOR T. Evolutionary innovations and where to find them: Routes to open-ended evolution in natural and artificial systems[J]. *Artificial Life*, 2019, 25(2): 207-224.
- [102] ROTHSCCHILD L J. A powerful toolkit for synthetic biology: Over 3.8 billion years of evolution[J]. *BioEssays*, 2010, 32(4): 304-313.
- [103] LOPATKIN A J, COLLINS J J. Predictive biology: modelling, understanding and harnessing microbial complexity[J]. *Nature Reviews Microbiology*, 2020, 18(9): 507-520.
- [104] LOPATKIN A J, SYSOEVA T A, YOU L C. Dissecting the effects of antibiotics on horizontal gene transfer: Analysis suggests a critical role of selection dynamics[J]. *BioEssays*, 2016, 38(12): 1283-1292.
- [105] ENKE T N, DATTA M S, SCHWARTZMAN J, et al. Modular assembly of polysaccharide-degrading marine microbial communities[J]. *Current Biology*, 2019, 29(9): 1528-1535.
- [106] GOLDFORD J E, LU N X, BAJIĆ D, et al. Emergent simplicity in microbial community assembly[J]. *Science*, 2018, 361(6401): 469-474.
- [107] SANCHEZ A, GORE J. Feedback between population and evolutionary dynamics determines the fate of social microbial populations[J]. *PLoS Biology*, 2013, 11(4): e1001547.
- [108] HEKSTRA D R, LEIBLER S. Contingency and statistical laws in replicate microbial closed ecosystems[J]. *Cell*, 2012, 149(5): 1164-1173.
- [109] GALLUP O, MING H, ELLIS T. Ten future challenges for synthetic biology[J]. *Engineering Biology*, 2021, 5(3): 51-59.



通讯作者: 胡政(1987—), 男, 研究员, 博士生导师。研究方向为基因组学、进化生物学和定量生物学。

E-mail: zheng.hu@siat.ac.cn



第一作者: 赵晓宇(1999—), 女, 硕士研究生。主要研究方向为细菌与肿瘤。

E-mail: xy.zhao1@siat.ac.cn