2019自然评论微生物学卷：17期，第12期，第725-741页

**工程微生物组的通用原则和最佳实践**

Christopher E. Lawson 1\*, William R. Harcombe 2, Roland Hatzenpichler 3,4,5,Stephen R. Lindemann 6, Frank E. Löffler 7,8, Michelle A. O’Malley 9,10,Héctor García Martín 10,11,12,13, Brian F. Pfleger 14, Lutgarde Raskin15,Ophelia S. Venturelli14,16,17, David G. Weissbrodt 18, Daniel R. Noguera1,19and Katherine D. McMahon 1,17\*

1美国威斯康星大学麦迪逊分校土木与环境工程系。

2美国明尼苏达大学圣保罗分校生态学、进化与行为学系。

3蒙大拿州立大学化学与生物化学系，美国蒙大拿州波兹曼市。

4美国蒙大拿州立大学生物膜工程中心。

5美国蒙大拿州立大学热生物学研究所。

6美国普渡大学西拉斐特分校食品科学系。

7美国田纳西州诺克斯维尔田纳西大学环境生物技术中心。

8生物科学部，橡树岭国家实验室，橡树岭，田纳西州，美国。

9美国加利福尼亚大学圣巴巴拉分校化学工程系。

10Doe联合生物能源研究所，加利福尼亚州埃默里维尔，美国。

11美国加利福尼亚州伯克利劳伦斯伯克利国家实验室生物系统和工程部。

12Doe Agile生物基金会，加利福尼亚州埃默里维尔，美国。

13巴斯克应用数学中心，毕尔巴鄂，西班牙。

14美国威斯康星大学麦迪逊分校化学与生物工程系。

15密歇根大学土木与环境工程系，美国密歇根州安娜堡。

16美国威斯康星大学麦迪逊分校生物化学系。

17美国威斯康星大学麦迪逊分校细菌学系。

18荷兰代尔夫特科技大学生物技术系。

19Doe大湖区生物能源研究中心，麦迪逊，威斯康星州，美国。

**摘要：**尽管人们对利用地球微生物组的力量有着广泛的科学兴趣，但知识差距阻碍了它们有效地用于应对紧迫的社会和环境挑战。我们认为，围绕设计-建造-测试-学习（DBTL）循环，形成研究和技术开发体系，将推进微生物组工程，并促进控制微生物组功能的基本科学原理的新发现。在这篇综述中，我们提出了微生物组工程中迭代DBTL循环的关键要素，重点介绍了普遍方法，包括自上而下和自下而上的设计过程、合成和自组装的构建方法，以及分析微生物组功能的新兴工具。这些方法可推动微生物组在医药、农业、能源和环境方面的广泛应用。我们还讨论了每种方法的关键挑战和机遇，并将其综合为工程微生物组的最佳实践指南。我们预计，DBTL框架的应用将迅速推进以微生物组为基础的生物技术，旨在改善人类和动物健康、农业和促进生物经济。

**专业词汇**

微生物组学（Microbiome science）：发现和检测控制微生物组组装和功能的基本原理。

微生物组工程（Microbiome engineering）：利用基本的科学原理和定量设计，创造出具有所需功能的微生物组。

共营养（Syntrophy）：一种有义务的互惠过程，由两个或多个有机体之间的代谢物交叉喂养介导，不能由一个有机体单独催化。

宏表型（Metaphenotypes）：由单个微生物基因组（宏基因组）之间的相互作用及其与环境的相互作用而产生的微生物组的新兴功能。

生态工程（Ecological engineering）：设计和操作生物反应器和其他工程系统的过程，以促进能够发挥预期功能的特定微生物群落的发展。

功能团体（Functional guilds）：使用相似资源（例如，电子供体、电子受体或碳源）并占据相似生态位的物种群。

关键种（Keystone species）：

对维持微生物组的功能和微生物相互作用（微生物之间及与环境间）有特别大影响的物种。

通量平衡分析（Flux balance analysis）：一种基于约束的数学建模技术，利用基因组信息重建代谢网络以模拟代谢通量。

集合建模（Ensemble modelling）：通过模拟一组可能性并选择与测量数据一致的，即使用多个模型来解决不确定性问题。

机器学习（Machine learning）：用于建立预测模型的技术，通过样本数据得到模式和推论，而不是明确的或机制的关系。

技术经济评估（Technoeconomic assessment）：通过工艺设计、建模和经济评估相结合的方式，对综合工艺的技术和经济可行性进行评估的工具。

生命周期分析（Life cycle analysis）：用于评估与产品生命周期所有阶段相关的环境影响的工具，如能源和水的消耗，以及空气污染物和温室气体的排放。

自组装微生物组（Self-assembled microbiome）：通过环境操作选择所需功能而建立的微生物组。

合成微生物组（Synthetic microbiome）：通过组合预先确定的纯化或富集培养物以实现所需功能而构建的微生物组。

整合和结合元件（Integrative and conjugative elements）：可移动的遗传元件能够通过位点特异性重组整合到DNA位点，该重组携带编码结合所需机制的基因。

外代谢组学（Exometabolomics）：一种分析技术，通常通过气相/液相色谱-质谱或核磁共振波谱测量环境和/或生物样品中细胞外小分子代谢物。

废气分析（Off-gas analysis）：生物系统产生的气体流速和化学成分（例如二氧化碳、氢气、甲烷）的监测。

结构-功能关系（Structure–function relationships）：微生物组的三维空间组织对其功能的影响。

广义Lotka-Volterra方程（Generalized Lotka–Volterra equations）：基于实验推断的物种相互作用参数，用来表示种群动态的一组常微分方程。

基本生态位（Fundamental niche）：物种能够生存和繁殖的一整套环境条件（即在没有种间竞争的情况下物种的生态位）。

现实生态位（Realized niche）：考虑到种间竞争（竞争、捕食等因素）后，一个物种所使用的环境条件集。

微生物群落似乎有着无限的能力，驱动着地球的生物地球化学循环，占据着每一个环境生态位。工程师和科学家长期以来一直在利用这一能力，例如，通过操纵土壤微生物组提高作物生产力3，通过刺激自然产生的或引入的微生物组来修复受污染的地下水4，或通过建造反应器微生物组来从废水中回收宝贵的资源5。尽管这些成就突出了微生物组的宝贵功能，但微生物世界的绝大多数转化能力尚未被释放和利用。由DNA测序驱动的最新研究揭示了未培养微生物的高度遗传多样性及其在多种生态系统中的关键作用6,7，为潜在的新生物技术应用提供了一个窗口。

认识到这一未被发掘的潜力，资助机构和国际科学界呼吁全球努力推进微生物组研究8,9。这些倡议已经认识到微生物组学需要超越描述性研究，采用一种系统方法，产生机制的、可预测的和可操作的理解，从而使合理的微生物组学工程8成为可能。

然而，由于缺乏对微生物组进行详细功能研究的可操作实验系统，大量的微生物组基因和代谢物功能尚不清楚，微生物之间的许多相互作用（例如，共营养作用）未被表征11，没有足够的工具来精确地测量和模拟微生物在时间和空间上的功能，以及精确操纵微生物结构和功能的方法的有限可用性，因此无法实现这一转变。

将基础科学发现与工程相结合可以克服这些挑战，并开发出支持可持续自然资源管理和人畜健康的创新解决方案。特别是，工程方法可以用来创建实验系统，允许测试概念知识和提取促进微生物组研究的新知识。为了加速科学发现和转化为创新解决方案，我们建议微生物组工程采用迭代设计-构建-测试-学习（DBTL）循环来形成研究和技术开发过程体系。该循环包括开发初始微生物组设计或初步模型系统，以实现规定的工程目标，构建微生物组，根据一组规定的指标测试其功能，以确定设计-构建解决方案是否产生设计目标（即，确定因果关系），学习哪些有用，哪些没用（以及为什么），并将新知识纳入后续DBTL循环的决策过程（图1）。这一方法已成功地应用于制造业12、代谢工程13和创业（“构建、测量、学习”）14，并能迅速提高我们开发急需工具和设计概念的能力，以利用微生物、提供创新解决方案和提高科学知识。

在这篇综述中，我们提出了一种迭代的DBTL方法的关键要素，该方法可用于促进微生物组的合理工程化，以实现有益于社会的功能。我们回顾了在医疗、农业、能源和环境应用中利用微生物组的各种方法，并确定了与实施每个DBTL阶段相关的当前挑战和机遇。最后，我们讨论了DBTL循环如何应用于建立微生物生态系统基本原理的模型系统，并展望了微生物组工程的前景。  
**设计微生物组（**Designing microbiomes**）**

由于分子尺度微生物组工艺的高度复杂性和有限的理解，微生物组设计传统上遵循自上而下的方法。这种方法试图预测生态系统水平的控制如何创造出具有所需功能的微生物组。然而，多组学的最新进展为自下而上设计微生物组提供了机会，通过预测代谢网络如何控制及其相互作用创建具有所需功能的微生物组。这些方法结合起来，提供了为特定工程目标设计微生物组的互补策略，特定工程目标包括从可持续的废水处理到治愈与微生物组相关的人类疾病等。

**自上而下的设计（**top-down design**）**

自上而下的方法不是预先决定哪些物种和详细的代谢途径，而是使用精心选择的环境变量（例如某些底物负载率、平均细胞保留时间和氧化还原条件），通过生态选择强制现有微生物组（自然发生或接种）以执行所需生物过程（或“宏表型”）15（图2）。这里，“上”是指发生所需生物过程的生态系统，“自上而下的设计”是指用于预测生态系统的物理、化学和生物过程（即生态系统过程）的操作如何获得所需功能的方法。生态工程原理16（也称为微生物资源管理17或微生物群落工程18）为预测如何操作生态系统提供了信息。这要求工程师将系统概念化为一个生态系统模型，该模型捕捉系统输入和输出、物理化学条件（pH、温度、氧化还原电位等）、已知的非生物和生物过程以及环境变量，以及它们的操作如何促进或抑制生物过程的优化19，20。随后，利用数学模型对系统中的化学物质和相关微生物进行质量平衡分析，并模拟化学和生化转化率。这些过程模型通过表征具有特定化学计量参数（生长和产物产量）和动力学参数（最大比生长速率，底物摄取率和底物亲和力）的微生物（例如产甲烷菌、发酵罐、硝化菌或光养生物）的关键生理或功能性团体捕获微生物功能21-23。这些模型还可以综合描述作用于化学物质和微生物的三维物理输运过程（扩散、平流和扩散），这在空间结构系统（如生物膜24、25）中尤为重要。

**自下而上的设计（**Bottom-up design**）**

尽管传统的微生物工程自上而下的设计方法为宏观过程提供了框架，并在废水处理21和生物修复4方面取得了广泛的成功，但是它常常忽略了驱动微生物和相关化学转化的复杂的原位代谢网络26，忽略了依赖于群落成员之间复杂相互作用的过程；例如，通过物种间直接电子转移的共营养相互作用27。因此，在设计过程中往往忽略了分子尺度的微生物组过程，限制了通过分子尺度机理分析的系统优化。多组学和自动化技术的最新进展（例如，在宏基因组学和微流体学中）使研究人员能够开发自下而上的方法，并将研究重点放在设计微生物组的代谢网络和微生物相互作用上。这里，“下”是指微生物组中单个物种的代谢网络（从它们的基因组中表达出来的），而“自下而上设计”是指用于预测通过这些相互作用网络的代谢流量如何产生所需输出的方法。一般的设计过程是获取微生物组的单个成员的基因组28（尤其是关键物种29，如果知道30），重建它们的代谢网络31，32，并使用建模33和/或网络分析工具34来指导设计（图2）。现有的基于约束的方法，如通量平衡分析提供了一个合适的框架，用于探索使用定量模型哪些化学转化的组合是可能的，其中个体种群的反应和代谢物可以被划分，并且可以使用最优性原则模拟种群内和种群间的代谢通量35。这些模型还可以模拟随时间和空间变化的稳态通量分布36,37，并且可以集成到基于过程和/或基于个体的模型中38，以预测宏表型、自组织空间模式和其他新兴行为。这种自下而上的工具为工程师提供了一个计算框架，以便系统地评估驱动生物过程和生态相互作用的代谢网络，并为合理设计具有特定特性的微生物组提供了一个平台，例如分布式路径39、40、模块化物种相互作用41，优化生态系统功能和稳定性的群落抵抗力和恢复力42和时空组织43。然而，大多数自下而上的设计实例都是用具有工程依赖性的模式生物（如大肠杆菌和酿酒酵母）构建的简单群落。因此，将这些设计扩展到具有数十到数百种不同物种的非模式生物体的系统，将需要对它们的新陈代谢以及控制它们相互作用和高阶行为的原则有更深入的了解。实施自下而上设计面临重大挑战，包括代谢网络重建不准确和/或不完整，许多基因、蛋白质和代谢物的功能未知，驱动个体和群落水平表型的进化压力缺乏理解，对基因、代谢和生态系统调节体系（例如，群体感应信号-响应系统）的理解有限44。这些限制导致了模型的高度不确定性，因为路径化学计量和酶动力学的关键限制要么不适当，要么缺失，目标函数无法捕捉细胞行为的真正进化驱动力45，最终导致原位宏表型预测不佳。作为自下而上设计的起点，可以从基因组注释和已知的生理信息中重建捕获中心碳和能量代谢的核心代谢模型。这些模型最初的预测能力可能有限，因为它们忽略了调控信息、途径动力学、次级代谢和进化。然而，当这些知识被获取并通过多个测试和学习周期被纳入代谢模型时，系统功能的准确预测（例如，代谢通量和代谢产物交换）可能会出现。作为一种补充方法，数据驱动建模技术，如集成建模和机器学习，可以提供更快速的方法来预测微生物组的代谢过程，或获得微生物组建模所需的约束和参数，而无需对代谢调节进行详细的机制理解46,47。这些模型框架已被用于从蛋白质组和代谢组数据预测途径通量48，通过基于集成模型的通量平衡分析改进代谢物交叉喂养预测49，并获得代谢模型所需的关键催化周转数50。虽然这些方法足够灵活、普遍，可应用到微生物群落，但是他们仍需要大量的单个菌株和互作群落代谢的实验数据。这些信息可以从以前的测试阶段（例如，从高通量的表型筛选和多组学）中利用，以允许数据驱动的设计。

**综合设计（*Integrated design*）**

展望未来，我们认为，成功的微生物组设计需要自上而下和自下而上的方法的合理平衡，特别是当我们处理复杂的微生物组时，如人类微生物组或活性污泥（图2）。混合方法可以包括选择未定义的混合物和已定义的菌群以实现所需的微生物功能，将基于过程的模型与从宏组学信息重建的自下而上的代谢模型合并以模拟生态系统过程、质量平衡和代谢物通量，并使用基因组衍生信息来开发群落选择策略。在设计中获取高阶特性，如功能稳定性和动力学，可能还需要自上而下和自下而上的方法来收敛。特别是，利用代谢框架量化功能退化、生态位互补和网络缓冲机制的新数学建模方法51可使微生物组多样性得到优化，以维持原位所需功能。更全面地描述微生物组代谢的需求将取决于具体的工程目标和生态系统的易处理程度。例如，可能需要更详细地描述厌氧微生物组的代谢，以将生物量转化为特定的商品化学品，而不是甲烷，因为需要更好地控制代谢。在这两种情况下，设计阶段都包括定义工程问题、开发概念和定量模型、确定要操作的关键生物过程以及评估多个候选设计方案。

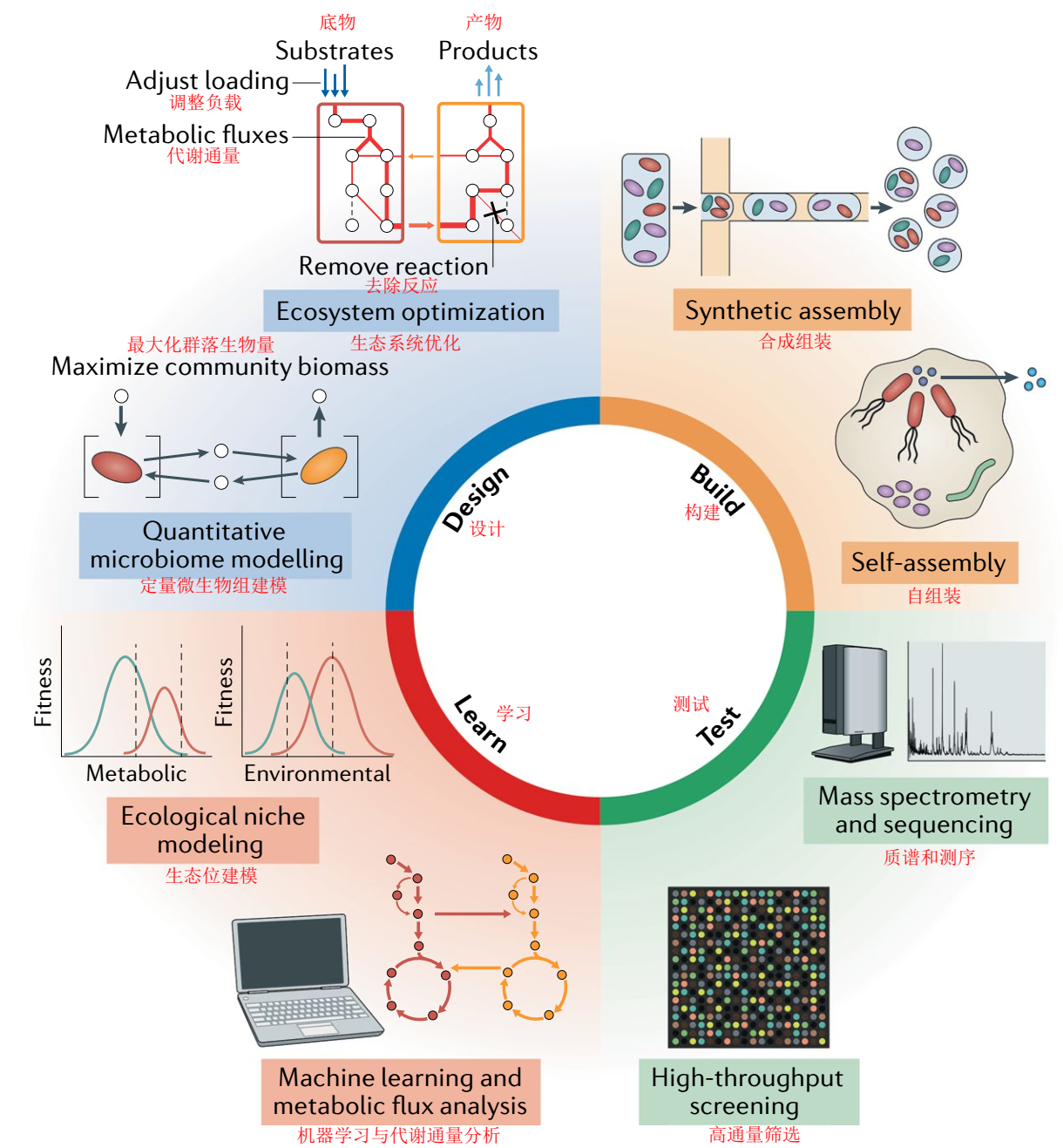


图1 |微生物组工程的设计-构建-测试-学习循环。介绍了设计-构建-测试-学习循环每个阶段的关键方面和方法。该循环从一个确定的工程目标开始，该目标确定了设计并产生了一个执行所需功能的工程化微生物组。

**实用的设计步骤（*Practical design steps*）**

在设计微生物组，特别是复杂微生物组时，有五个关键步骤：定义工程问题、开发概念生态系统模型、创建定量模型、确定要工程化的微生物组过程以及开发和评估候选设计策略。为了推动DBTL循环，必须明确定义问题及建立具有可测量的设计目标。这些目标可以具体说明期望的结果，例如产品的滴定度、速率和产量、污染物去除效率、作物生产力或功能稳定性和鲁棒性的程度。

设计目标应辅以技术经济评估和/或生命周期分析，以确保解决方案在经济上可行，并具有积极的环境和社会影响52,53。概念生态系统模型可用于将问题置于上下文中，这些模型捕捉了系统边界、输入和输出、碳和养分流动的主要途径、导致这些转化的关键物种和种间相互作用以及影响其活动的因素（例如，pH、温度、氧化还原电位和停留时间）19。它们提供了一个概念图，描述了目前对微生物组与生态系统物理、化学和生物组成部分之间相互作用的理解，有助于确定系统理解和数据收集需求方面的重要差距。在这个阶段，所有相关的信息应该从文献中收集，现有的数据（例如，来自人类微生物项目）54和在线数据库（例如，MiDAS（活性污泥微生物数据库）55用于生态系统表征。这包括关键生物的参考基因组和生理信息、以前的多组学数据集、生态系统的物理化学特性（如pH、温度和化学浓度）和过程（如光化学反应和水文地质过程）、场地特征（如养分负荷和动力学，土壤剖面和肠道解剖）以及描述生态系统所需的所有其他信息。缺失的信息，例如未知的生化途径和介导它们的生物体，可以在构建、测试和学习阶段作为目标。这一概念生态系统模型可供科学界用于提出和检验理论，并作为开发定量模拟工具的路线图。建立能够计算和模拟代谢通量、微生物丰度、质量平衡和生态系统物理化学参数的定量建模工具，对于微生物组的系统设计至关重要。可以使用几种方法来创建此类模型，包括机理代谢建模33、基于过程的建模21、数据驱动的建模（例如机器学习）48和基于个体的建模38或其组合。无论采用何种方法，复杂微生物组的模拟都可能需要基于实验有效假设的简化。简化可以包括将模型简化为一组代表重要功能团体并控制主要碳和能量流动的核心或关键物种，或者将物种的代谢网络规模缩小为中心碳和能量代谢。展望未来，重要的是确保模型在构建-测试-学习循环中经过严格的实验验证和迭代，以提高其在微生物组工程中的实用性和广泛应用，并确定建模工作失败的时间，揭示概念理解上的差距，从而进一步促进模型的重新设计和改进。定量微生物组建模（如动态流量平衡分析）有助于确定需要直接操作、添加或移除以实现预期工程目标的核心和外围生化途径。目标可以包括增加丁酸产生和不消化碳水化合物被人类肠道中发酵细菌降解，在淡水生态系统中通过蓝藻防止毒素生物合成或通过生物强化有机卤化物呼吸细菌刺激有毒氯化物的降解。微生物组学建模可以预测环境（如底物负载、pH值和固体保留时间）或基因操作（如基因敲除、途径添加和强制依赖）如何优化微生物组学功能以实现工程目标。如有必要，可设计合成微生物以改善微生物组的功能。这种合成微生物将需要评估它们在相关环境条件下与现有微生物组成员合作和竞争的能力。

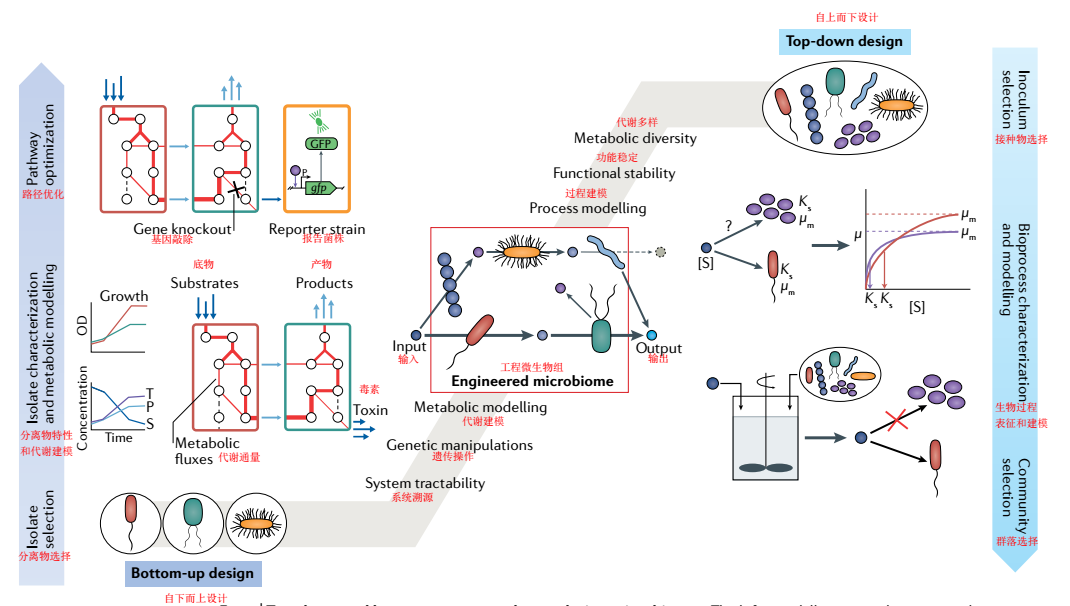


图2 |自上而下和自下而上设计微生物组的方法。左面板说明了从分离物开始的自下而上的设计流程。对单个物种进行生理特性分析，并利用代谢模型为期望的功能（从深蓝化合物中产生浅蓝化合物）设计群落。基因工程和合成生物学策略被用来优化系统功能（识别基因编辑目标，重设远离毒素（紫色）和朝向所需的产品的代谢流路径；毒素报告菌株的设计）。右面板说明了一种自上而下的设计，首先是一种含有环境中未培养微生物的接种物。对未定义的微生物组进行群落特征描述，并利用生物过程建模（包括动力学和微生物生长在内的质量平衡分析）来制定选择策略以实现所需的功能（从深蓝色化合物中产生浅蓝色化合物）。生物过程工程（例如，反应器设计）用于优化系统功能。中间面板显示了自上而下、自下而上的集成设计。选择未定义群落和已定义培养物的组合以实现所需的功能。进行群落特征描述，并将基于过程的模拟与代谢模拟相结合的微生物组建模用于制定选择策略和分析微生物组代谢通量。微生物的形状代表了设计过程中选择的不同分离物或群落。OD，光密度；S，底物。

**构建微生物组（Building microbiomes）**

构建阶段包括物理组装设计的微生物组，或通过自上而下操纵自然群落（即自组装微生物群）或通过自下而上使用自然发生或工程微生物（即合成微生物组）的纯化或富集培养物组装。构建阶段旨在实现设计规范和预测。

**通过自组装构建（*Building by self-assembly*）**

自组装微生物组可能包括那些使用反应器工程（例如，废水处理生物反应器）或生物刺激（例如，添加到土壤、沉积物或地下水含水层）作为开放混合培养物构建的微生物组，其中，构建创造了一个环境，促进土著微生物的生长和期望活性。例如，操纵反应器流体动力学，将生长缓慢的微生物固定在致密颗粒中，使其得以保留和增殖56,57，使用非人类可消化的碳水化合物刺激肠道发酵产生短链脂肪酸58，或添加电子供体，在有毒氯污染的生物修复过程中驱动有机卤化物呼吸细菌的代谢。当功能团体之间的生理和物理化学性质的差异可以通过环境操纵（例如，生长速率59、主要电子供体和受体4、60、底物亲和力、细胞和/或生物膜密度61和氧化还原梯度的差异）来进行组装时，这种方法非常有效。然而，当需要对微生物代谢和相互作用进行更精细水平控制时（例如，控制复杂的竞争性相互作用62、以高产量和纯度生产有价值的生物产品63或以多种生活方式控制物种64），这一点可能会受到限制。

此外，进化工程的新策略已经成为构建自组装微生物组的潜在工具。在多个选择周期和/或制度下控制初始微生物组的暴露会导致微生物组通过适应或进化获得或优化特定功能。例如，连续地转移使植物性状最大化的微生物组已经产生了增加植物生物量65和开花时间66的微生物组。对群落水平选择的响应通常由单个物种的富集或适应所驱动67,68；然而，对群落生物量生产的选择也被证明能增强所定义的两个物种和三个物种共培养中所需的物种相互作用37,69。重新检查选择实验，以了解突变和/或适应改变微生物组表型的时间和方式，可以阐明微生物组适应度优化和信息设计的潜在机制，正如在实验室进化实验中大肠杆菌所示70,71。由于类似的进化方法（例如，适应性实验室进化）也已成功应用于代谢工程菌株的优化72，将已经为单个微生物开发的实验和计算协议扩展到微生物组可以简化设计阶段，减少完成进化实验所需的时间。

**构建合成微生物组（*Building synthetic microbiomes*）**

使用纯化或富集培养物直接构建微生物组也是有希望的，因为降低了复杂性和使用遗传易处理和/或特征良好的微生物。这种自下而上的方法使得越来越多的合成生物学工具可以用于微生物组的构建和优化。早期直接从培养的微生物中构建微生物组的一种方法是生物强化。在这里，定义的实验室菌群被添加回环境中，以提高特定污染物的降解率。一个成功的例子是添加含脱卤球菌纲有机卤化物呼吸细菌的菌群到污染的地下水含水层和沉积物，以加速有毒氯化污染物的降解。这一方法成功的关键在于详细了解关键脱氯菌与其他微生物及地球化学环境的生理学、营养需求和潜在的生态相互作用4。然而，与氯化污染物的成功相比，生物强化方法在溢油事故中基本失败。与有机卤化物呼吸性脱卤球菌成员不同，其填补了一个独特的生态位，没有氯化污染物就无法生长，能够降解油烃的生物体（特别是需氧菌）无处不在，代谢多样，生长不依赖于特定的底物或氧化还原对64。这种代谢的多功能性限制了它们在生物强化中的应用，因为它们不可预测的原位活性。生物强化可能失败的其他原因是，未经确认的互惠相互作用和执行关键功能的微生物缺失（例如，生产多糖表面活性剂以提高碳氢化合物的生物利用度）73，或者在实验室条件下选择的菌群在苛刻条件和/或可变现场条件下不再具有足够的竞争力74–76。这些例子突出表明，需要更好地了解合成菌群的相互作用网络，特别是支持相互作用（次要功能）的功能，以及在复杂生态系统中往往难以预测的原位竞争环境。尽管自下而上构建微生物组和从特定栖息地收集培养微生物的呼声越来越高77,78，但大多数与人类健康、农业和环境应用相关的微生物仍然未培养、特性差、遗传上难以控制且难以维持，使合成微生物组的构建具有挑战性。为了捕获这种非特征化的代谢多样性，需要创新的分离和控制微生物组组装的技术，例如，单细胞分选79与高通量培养（培养组学）80、81相结合，并在多个条件下并行分型82、83。微流体84,85（即，微升液滴的产生和操作）可以促进这种方法。微流控芯片可以通过液滴组合86、消除特定物种87、测序和单个细胞的多组学分型，自动组装和分析纯化或富集培养物中的微生物群落。结合新的基因编辑技术，例如提高了基于同源重组的基因编辑效率的基于CRISPR基因组工具90，91,92，微流控技术还可以自动化合成生物学技术，用于具有新功能的细胞和微生物组的工程93。

合成微生物的另一个挑战是在实验室或开放系统（例如人类肠道、土壤和废水处理厂）中保持其功能的稳定性，这些系统易受自然微生物和动态异质环境的入侵。如前所述，有机卤化物呼吸性脱卤球菌成员生物强化成功的主要原因是他们高度专业化的生活方式，使他们能够利用氯化电子受体占据开放的生态位。然而，在开放系统中具有多种生活方式的物种的功能稳定性是不可预测的。很少有研究考察开放系统中合成菌群的功能稳定性，而合理设计稳定的生态相互作用所需的知识是有限的。然而，工程菌已经成功地在哺乳动物肠道中作为诊断传感器部署了200天，维持了强健的功能94,95。这一壮举，再加上脱卤球菌4的生物强化实例，证明只要关键参与者能够与土著微生物竞争，合成菌群就可以与先前建立的群落成员形成稳定的微生物组。自组装微生物组的观察结果表明，构建具有时空组织的群落对于获得稳定和多功能的合成微生物组具有重要意义。高度多样的微生物群落，如人类微生物群或用于废水处理的微生物群，以生物膜、絮体或颗粒的形式自组装，包含多个单一物种的微菌落，通过物种特定的胞外聚合物质（包括多糖、蛋白质和DNA）和其他不明确的大分子（如腐殖质）96，97连接在一起。这些自组织微生物组合创造了不同的微环境和生态位，支持看似不相容的功能（例如，好氧和厌氧过程）98、99和可以补偿干扰的功能多样的种群结构，例如营养物质的变化，物理化学条件或捕食的变化100，101。尽管将这种精细水平和复杂的结构构建成合成微生物组尚处于萌芽阶段，但基于微流控的系统已被用于通过控制空间结构和化学通讯来组装简单的群落，从而提高功能稳定性102。此外，三维生物打印平台可以允许构建空间组织系统，其中种群可以物理分离，同时保持化学互作103，104。如何将这些空间上定义的结构从实验室试验系统扩展到实际应用仍有待解决，尽管从模型系统（如合成多糖颗粒）105，106的测试和学习阶段获得的知识应提供更多的见解。在此之前，现有的基于自上而下组装和/或工程生物膜载体媒介107的方法可用于构建具有更大稳定性和功能性的自组织化合成微生物组。在复杂环境中可靠地执行感知-计算-响应程序的工程宿主中设计合成基因回路也仍然是一个主要的挑战108。因此，研究自然生态系统和工程生态系统中决定微生物组稳定性和适应环境扰动的分子机制，以提取可用于合理工程化稳健功能的设计原则，具有重要意义。鉴于基因工程微生物和微生物组在各种开放环境中的潜在用途，未来诸如生物封闭系统（例如两层基因回路和必要的合成营养缺陷）109等保障措施也需要进一步发展，并需要作为使用基因修饰物种构建合成微生物组的组成部分。

**整合方法（*Integrating approaches*）**

合理微生物组设计的最终目标是开发工具，使工程师能够在一系列理想的操作条件下，直接在原位添加、移除或修改特定的功能和表型。一种新兴的技术，有望实现这种灵活性就是原位宏基因组工程110， 111，这涉及到工程化移动遗传元件输送到土著微生物。例如，利用整合和接合元件工程化的供体菌株，已将携带报告基因和抗生素抗性基因或多基因途径（例如，固氮基因（nif）簇）112的DNA转移到高度异质性和多样性环境中的细菌，例如土壤112和哺乳动物肠道111。与现有的CRISPR-CAS基因编辑技术相结合，这种工具的进一步开发将允许精确地操纵微生物基因组的代谢网络，有效地结合自组装和合成微生物组（方框1；图3）。

**测试微生物组功能（**Testing microbiome function**）**

测试阶段包括测量微生物组的相关表型和特性，以确定设计-构建解决方案的有效性。测量应确定是否达到了设计结果（例如，测量生物产品的滴定度、速率和产量、污染物去除效率或作物生产率），以及设计-构建解决方案是否对观察到的结果负责（确定因果关系）。这通常需要读取生态系统的物理化学性质（如pH、温度和化学浓度），以及关键生态系统过程和微生物组功能（如生物量增长、化学转化、养分同化和代谢通量）的化学计量和动力学。例如，在厌氧消化微生物组中，乙酸降解率和甲烷途径可使用13C标记的乙酸和在线沼气分析进行测试，测定乙酸产甲烷相对于共营养的乙酸氧化耦合氢气产甲烷的流量113。虽然在测试过程中测量的微生物粒级将取决于特定的设计目标和生态系统复杂性，量化分子微生物过程的能力（例如，代谢途径率和路线，酶活性和单个有机体生长速率）超出了批量活性的测量范围，并允许对观察到的微生物组功能的特定机制进行测试。面临的挑战将是开发出高通量、定量、价格合理且易于使用的工具，以便能够在时间和空间以及动态条件下完成微生物组的常规分析。为了实现这一目标，我们设想了一个测试阶段，包括高通量微生物组设计表型筛选-构建解决方案，然后使用多组学和代谢通量分析对有前景的解决方案进行更深入的研究，以获得对潜在机制的更深入了解（图4）。利用液滴微流控技术可以实现对构建的微生物组的高通量表型检测，最近已证明可用于筛选约100000个合成群落114。结合液体处理和微量滴定板先进的传感的全自动微生物反应器平台，或缩小的生物反应器培养也可以使用82，83。结合新兴的测量异质环境中代谢网络活性和代谢过程的方法（方框2），将获得丰富的信息以促进学习。

**方框1 |设计-构建-测试-学习循环，以创建具有所需功能的合成微生物组**

我们提出了一个通用的设计-构建-测试-学习循环，用于创建具有所需功能的合成微生物组，集成了自上而下和自下而上的方法。我们简要描述了循环的两次迭代，并确定了综合采用高通量方法和自动化来提高速度和再现性的机会。

**自上而下的方法**

*设计：识别生物过程。*利用或复制过程的一个例子是将复杂的木质纤维素生物质厌氧转化为有价值的商品化学品。初始设计步骤包括选择可能含有具有所需功能的微生物（例如，酸相厌氧消化污泥、草食性瘤胃微生物组或其他微生物）的不同接种物。包括环境参数（pH、温度、养分等）和预期功能团体（水解菌、发酵菌、产甲烷菌等）的概念生态系统模型用来选择富集变量。

*构建：从多种来源富集微生物组。*来源接种物在不同的环境条件下，利用真实的（如木质纤维素水解液或瘤胃液）和合成介质培养，以选择所需的功能。通过对环境条件和介质组成的调节，达到了预期的功能。对于复杂环境（如土壤），实验室生态系统模型可能是微生物富集的理想平台146。

*测试：评估性能。*利用高通量表型筛选，对真实和合成培养基上富集微生物的性能进行测试。利用微流控或自动微反应器实验可以开发出高通量筛选。更深层的多组学测量（如宏基因组学、宏转录组学和宏蛋白质组学）是从高性能微生物组中收集的。

*学习：确定微生物组成员的关键功能角色。*除了关键功能外，通过代谢重建和多组学分析还确定了所需功能的瓶颈。这种理解有助于完善微生物组功能的概念模型，并创建定量模型。

**自下而上的方法**

*设计：识别新的潜在微生物伙伴。*模拟代谢模型用于筛选相互作用微生物的高效微生物组富集物。组装的宏基因组基因可用于重建关键微生物组成员的代谢模型。自动化的计算工作流（加上人工管理）将加速模型的建立。流量平衡分析用于预测每个微生物对最佳生长和活性的需求，并将单个代谢模型统一到微生物组模型中，以确定新的潜在合作伙伴，从而改进设计目标（例如，更高的滴定度、速率或有价值产品的产量）。

*构建：将关键微生物重组成新的合成菌群。*在分离或富集之后，在不同比例（例如，1:1，1:10）模拟预测的基础上，关键微生物组装成新的合成菌群。微流控设备和/或液体处理机器人可用于高通量分离和重组。

*测试：测试菌群的功能和稳定性。*高通量表型筛选结合多组学测量可用于检测。这一步骤还应包括验证预测的单个分离物或富集物代谢。

*学习：识别控制功能的微生物相互作用。*用代谢通量分析在菌群内生长的微生物与在分离中生长的微生物的代谢，可以确定重要的机制和相互作用。这一认识可用于提出如何通过环境操作和/或原位基因组工程优化微生物组的功能和稳定性。

**微生物组代谢网络活性（*Microbiome metabolic network activity*）**

为了在系统水平上测试微生物组的功能预测，测量微生物组的原位代谢网络结构和活性至关重要。多组学方法（宏基因组学、宏转录组学、宏蛋白质组学和代谢组学）与生物信息学工具相结合，使得能够以基因组为中心分析微生物组内的单个物种（甚至菌株）115，并对序列、蛋白质和代谢物s116–118进行全局测量。这些工具在光谱上测量微生物组分，从功能潜能（例如基因丰度）到表达产物（例如蛋白质和代谢物丰度），并通过其组合活性产生驱动系统功能的微生物组宏表型。目前，用于推断微生物组功能的多组学方法主要集中在将跨时空的基因丰度或基因表达数据与生态系统地球化学数据或过程速率相关联。这包括使用定量PCR分析（例如氨单加氧酶）119、微阵列（例如，GeoChip）120或非靶向高通量方法（宏转录组和/或宏蛋白质组）测量关键功能基因和转录。虽然有助于整个系统表征和发现，这些方法侧重于测量系统的组成部分或“部件列表”，或系统的“零件清单”，由于代谢网络的复杂性、相互作用和调节，通常会限制潜在表型的预测因子。因此，需要新的方法和工具来测量微生物组代谢网络的原位化学计量和流量，以便对设计预测进行直接测试，并为代谢调节提供机制见解。代谢通量分析是测量体内通量最权威的方法。该方法利用代谢网络模型123计算同位素标记实验期间获得的代谢物稳定同位素测量的通量。尽管代谢通量分析已被用于测量共培养中的通量124，但由于代谢产物池不易分配给单个细胞，并且微生物组中可能的反应数量大大超过单个生物体，因此在群落中的通量分析具有挑战性。然而，同位素示踪剂结合外代谢组学和/或尾气分析已被用于驱动重要微生物组功能的确定过程通量，例如厌氧消化116中的共营养的醋酸氧化与甲烷生成。为了规避代谢产物测量的挑战，提出了一种用短肽代替氨基酸进行代谢通量分析的标记模式分析方法。利用高通量的宏蛋白质组学方法，肽可以被分配到微生物组中的单个物种，这为确定微生物群落中的通量（即“宏通量组学”）打开了大门。鉴于通量代表了细胞在所有水平上调节的最终结果126，进一步发展和展示宏通量组学对于推进微生物组工程工作和我们理解微生物组的代谢调节至关重要。这也将需要新的软件包来进行相关的计算分析，类似于现有的13C代谢通量分析软件127。这样的数据还可以让代谢建模者推断，而不是假设，群落水平和单菌水平的目标函数，并确定新的约束条件，允许准确预测和测量驱动微生物组功能的反应速率。

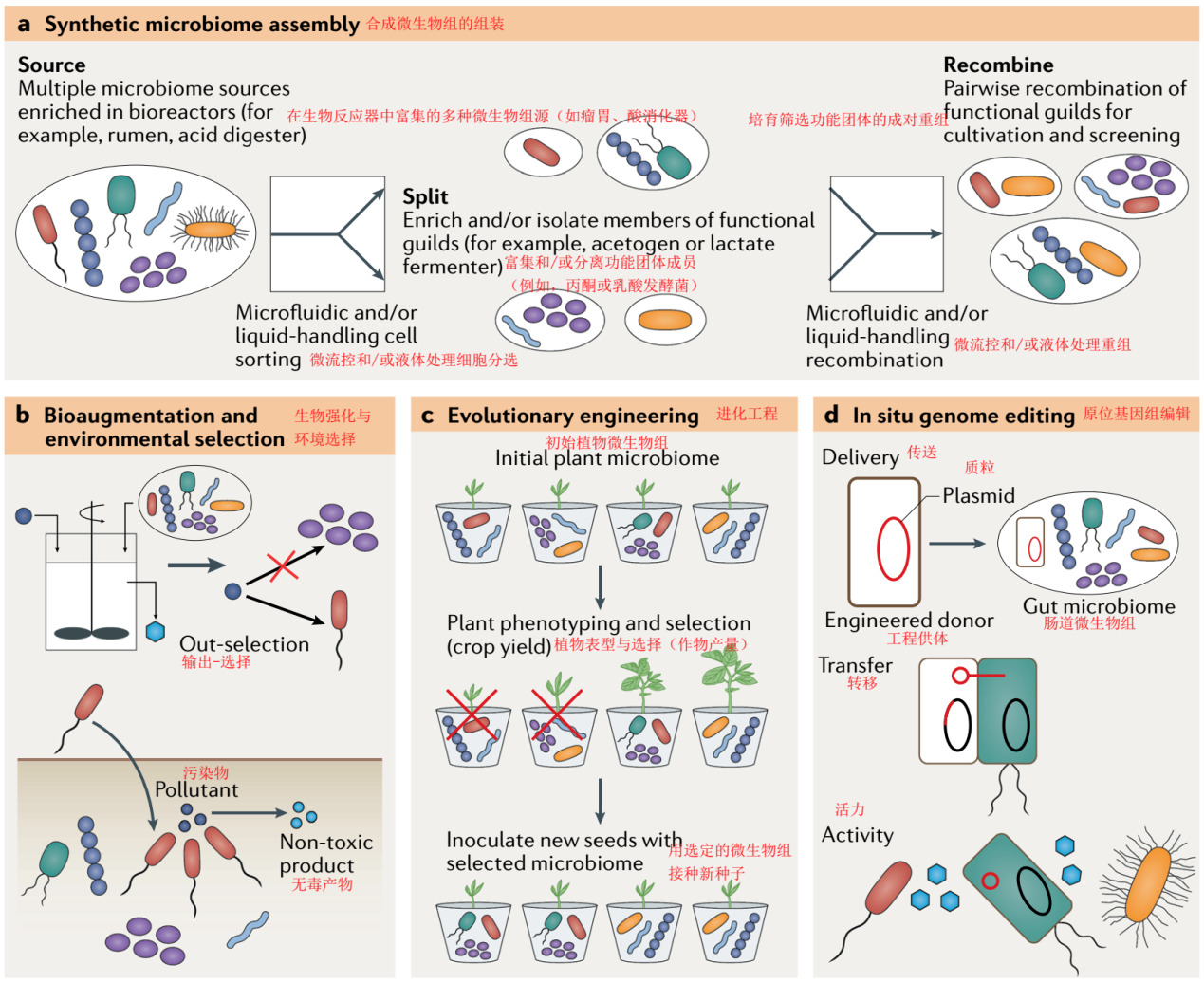


图3 |构建自组装和合成微生物组。a |从多种微生物来源组装合成微生物组的流程。利用自动微流控细胞分选技术，可以将复杂的微生物组分离成关键的功能成员。然后，分离或富集的成员可以使用液体处理机器人重组为合成菌群，用于下游筛选和/或培养。b |微生物组的组装也可以通过生物反应器操作或生物刺激（上）的环境选择来实现，或者通过使用定义的培养物（底部）的生物强化来实现。c |另一种选择是通过微生物组的定向适应和/或进化来获得或优化所需功能的微生物组组装。d |原位微生物组工程可用于为环境中的土著微生物组添加新功能。

*.*

**空间异质环境中的功能测量（*Measuring function in spatially heterogeneous environments*）**

大多数自然微生物组，例如与植物（例如根际）、人类（例如口腔微生物群）和工业过程（例如酸性矿井排水）相关的微生物组，在直接影响微生物组功能的微观物理化学梯度上显示出高度组织化的空间组成。例如，微生物的空间接近性可以控制它们是否通过扩散底物或直接转移相互作用128，而菌落大小的变化可以显著地影响生物膜微生物的表观底物亲和常数和底物竞争129。因此，最大的挑战之一将是创造工具，在所有相关尺度（从微米到千米）上测量和报告微生物组的空间结构和功能。目前，利用荧光原位杂交结合稳定同位素标记130、化学指纹131、质谱图132和/或基于荧光的生物正交非标准氨基酸标记133等方法测量结构-功能关系主要集中在微米到毫米尺度上（方框2）。尽管这些技术已经成功地确定了微生物组中空间分布微生物的底物利用和活性模式，但它们受到通量的限制，只能检测和/或区分有限数量的物种。标记技术（例如稳定同位素标记和生物正交非标准氨基酸标记）与宏蛋白质组学和细胞分选（例如荧光活化细胞分选）133的综合应用可用于高通量和空间分辨率的微生物代谢活性测量。结合描述微环境化学特性的微传感器装置（例如，通过微电极134或工程生物传感器95），可以实时监测微生物组结构、微生物组功能和生态系统理化参数。

**学习微生物组设计原则（*Learning microbiome design principles*）**

在微生物组工程的设计、建造和测试阶段取得进展，为我们提供了一个从以往的失败和成功中吸取教训，并将新知识融入后续循环的独特机会。事实上，DBTL循环的学习阶段对于成功和提高微生物工程效率至关重要。到目前为止，还没有一种通用的策略、技术或方法能够保证成功地将从测试阶段获得的信息转化为新的知识，从而为下一个设计阶段提供信息。因此，我们强调在早期将足够的重视和资源投入到学习阶段的重要性，以避免例如由于在学习步骤中相对缺乏投资而在代谢工程中遇到的困难13。

需要进一步发展计算方法，使学习阶段形式化，包括机器学习算法48135136、代谢通量分析和基于约束的分析36124125137、生态系统建模方法138和调控网络分析139。总之，这些分析可以从大数据集中分离出微生物组相互作用和功能的主要驱动因素，从而为微生物组的设计提供信息。例如，广义Lotka-Volterra方程可以从作为自下而上设计起点的时间种群动力学数据中推断相互作用物种140或基于约束的分析可以应用于从13C代谢数据中识别关键的代谢物交换反应，从而提高通量模拟的准确度并改进厌氧菌群137。

更广泛地说，我们认为学习阶段的重点是通过不断完善概念知识和提出的理论（例如，从传统的宏观生态学）51，141-144（每个DBTL循环），将数据转化为微生物组工程的一般原理。我们建议利用实验室生态系统模型来推动微生物组工程的研究和学习。模型实验室生态系统是一个实验平台，可以以简化和控制的方式复制复杂环境（自然环境或工程环境）的物理化学条件，并包含模型微生物组（例如，模型根际微生物组（THOR）145可作为学习如何设计、构建和优化工程微生物组的测试依据。这些生态系统减少了复杂性，可以用于实验，并且可以以可重复的方式建立，当人们在自然环境中工作时，这往往是不可能的。

最近，实验室生态系统模型已经开发出来，用于研究植物-土壤微生物组的相互作用146。这些人工构建的生态系统利用三维打印、传感、分析和图像技术，创造了一个复制原生土壤生态系统的实验装置，在该装置中，可以根据变化的变量监测微生物和宿主表型，允许系统地分析影响植物健康的微生物相互作用和代谢物交换146，147。人工生态系统在模式生物和复杂的自然微生物之间提供了一个中间地带，并且可以在专家研究人员之间协作建立，以创建标准化和可复制的设备和协议，传播到更广泛的研究群体。这样的模型系统能够以一种容易处理的方式实验性地开发出具有所需功能的工程微生物组，并允许将结果与自然环境下的结果进行比较。这种模型与自然生态系统之间的交叉研究将是一种有价值和必要的方法，有助于学习与实际系统（而非实验室手工艺品）相关的工程原理和实践，并有助于获取有关将基于实验室的工程策略扩展到全面应用的知识（图5）。例如，以微流体为基础的人体肠道微生物组的体外模型，包含了与不同细菌群共培养的人体细胞，已经产生了生理（包括上皮细胞单层形成、细胞生长和活力、细胞因子水平和代谢谱）和环境（包括氧梯度和层流）可与体内变量148相比较的变量。

将生态系统模型与DBTL循环结合起来，对于理解控制微生物相互作用和功能稳定性的机制可能特别有成效。大量的知识可在特定的微生物上进行共聚集和交换代谢物，如涉及氮循环的细菌、甲烷氧化古菌的联合体和硫酸盐还原菌128149150，以及与氢营养的甲烷菌151152合作的同营养细菌。然而，我们才刚刚开始了解调节群落中微生物行为、相互作用和亲缘关系识别的复杂机制（如群体感应和次级代谢产物153）。虽然研究已经建立了微生物组功能冗余、多样性和稳定性之间的联系154，但尚未建立一个预测或设计功能稳定微生物组的框架。通过使用模型实验室生态系统与现有的微生物生态学和工程设计知识，有可能破译微生态系统的化学语言，并发现其他重要的过程（包括进化、选择、分散限制和中性过程）的机制155，这些机制使得鲁棒性和稳定的微生物组功能。将这一理论转化为工程设计实践需要一个量化的框架，将这些机制与代谢相互作用网络和新的方法联系起来，使代谢模型产生生态特性（框注3）。

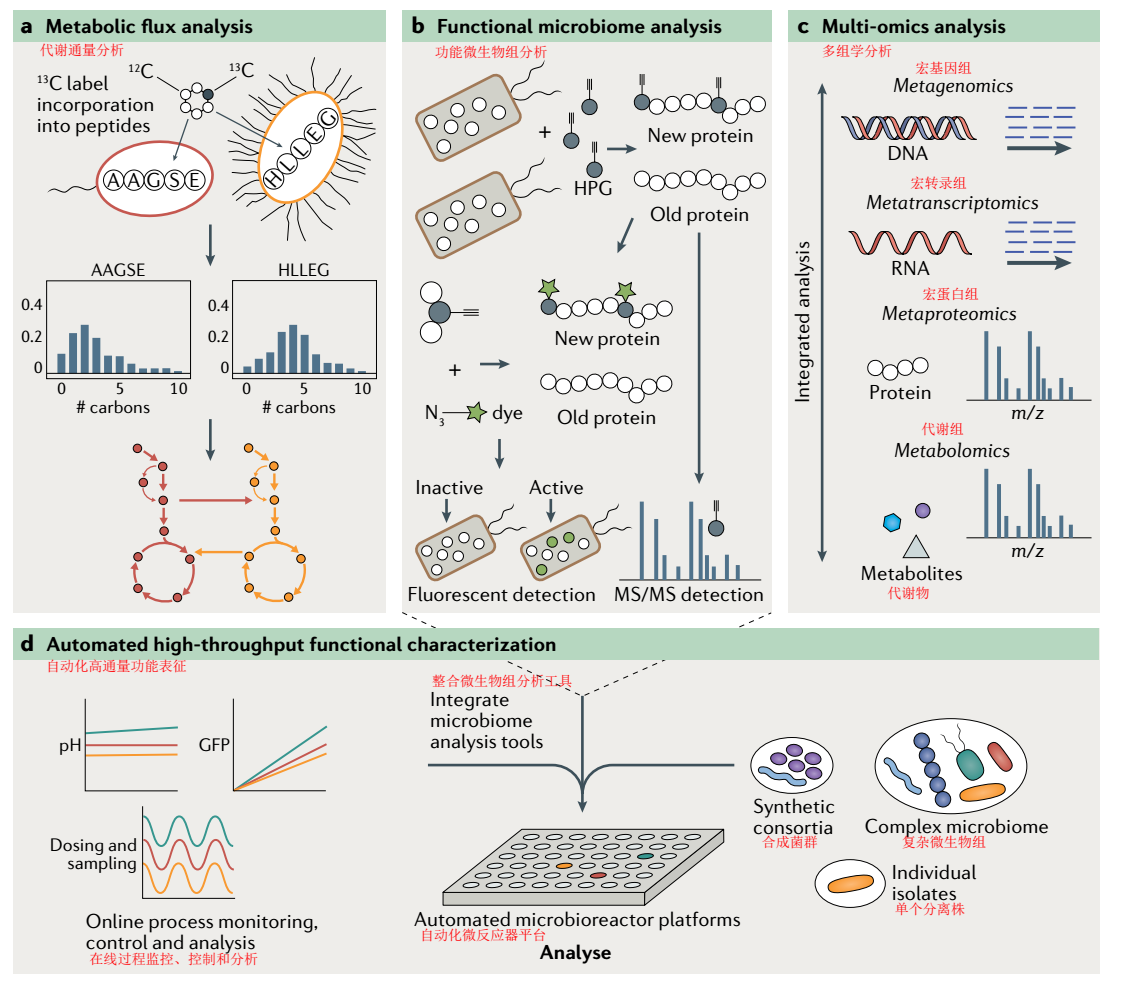


图4 |测试微生物组功能。a |同位素示踪剂结合宏蛋白质组可以通过分析短肽而不是氨基酸（代谢组）的同位素标记模式来测量微生物组的代谢通量。b |生物正交非标准氨基酸标记是一种利用荧光检测或蛋白质组学快速分析原位合成代谢过程（生长）的方法。c |整合宏基因组学、宏转录组学、宏蛋白质组学和代谢组学，可以重建和分析微生物组的代谢网络表达。d |一个自动化的微生物反应器平台允许在不同条件下（例如，随着环境或生理变量的变化）对微生物组过程进行高通量分析。该平台可以集成工具，对微生物组单个成员和复杂群落进行详细的功能分析。过程监测图中的绿色、红色和黄色与微生物反应器孔的位置相对应。HPG，高丙炔甘氨酸；MS，质谱。

**方框2 |测量微生物组功能的工具箱**

**多组学集成**

从宏基因组数据28组装基因组的能力使得能够在不同群落内对单个转录组63和蛋白质组118进行基因组注释分析，并大大提高了多组学数据集的注释能力。向前迈进的一个关键挑战将是整合代谢组学信息163，包括细胞内和细胞外，这些信息不能轻易地分配给微生物组的单个成员，如DNA、RNA和蛋白质。大量未知的或特征性差的基因、酶和代谢物目前限制了多组学信息的注释能力。然而，它确实为进一步的生化研究创造了新的靶点。生物信息工具的进展，例如数据驱动方法（例如，统计或机器学习方法）和基于知识的方法（例如，交互网络或基因组规模的代谢模型）164，165，将是通过连贯的多组学数据集成系统测量微生物功能成功的关键。

**同位素示踪剂**

同位素示踪剂在纯培养物和群落功能分析中有着悠久的历史，并且已经与DNA166、RNA167和蛋白116测量相结合，将个体种群与特定的原位功能联系起来。展望未来，需要更多的努力将同位素示踪技术与多组学（特别是宏蛋白质组学和代谢组学）结合起来，以阐明微生物组内复杂的代谢网络。这些技术的结合还应为测量细胞内和细胞外反应速率（“宏流量组学”）124125铺平道路，124125是在工程化纯培养物中阐明体内表型、途径限制和代谢调节的最有力工具之一。

**质谱成像**

质谱成像（MSI）技术显示了复杂样品中元素及其同位素以及生物分子的分布。MSI非常适合于分析空间结构的微生物组和研究细胞间的相互作用。当与荧光原位杂交结合时，MSI还允许将微生物组结构与功能168169连接起来。用不同的MSI技术可以获得的化学覆盖率、空间分辨率和样品制备取决于使用的电离方法132。尽管纳米二次离子质谱（nanoSIMS）与基质辅助激光解吸电离（MALDI）或解吸电喷雾电离（DESI）相比具有更高的横向分辨率，但其相对化学通用性非常低（元素和同位素与肽、脂类、代谢物和其他分子相比）。因此，nanoSIMS通常被用于研究单个细胞的底物使用，而MALDI–MSI则被用于可视化群体之间的化学相互作用132。尽管MALDI–MSI和DESI-MSI比nanoSIMS170更易接近，并且能够很好地显示微生物体内广泛的化学相互作用，但它们的通量非常低，其横向分辨率和灵敏度目前无法进行单细胞代谢分析132。结合这两种方法的最佳技术是纳米结构引发剂质谱（NIMS）。NIMS是一种无基质解吸/电离技术，它依赖于被困在30nm孔中的引发剂分子来实现吸附在孔表面的小分子的电离。NIMS具有约150nm的横向分辨率，特别适合于肽和代谢物171的分析。到目前为止，NIMS在微生物172，173中的应用还很有限。我们预计，解决这些问题的进展将很快使MSI成为一个有用的、应用更广泛的微生物功能组分析工具174。

**生物正交化学**

代谢标记技术，如生物正交非典型氨基酸标记（BONCAT），提供了额外的方法来测量微生物组原位合成代谢活性。BONCAT基于非标准氨基酸（例如，l-叠氮高丙氨酸，l-蛋氨酸替代物）的体内翻译结合，然后通过叠氮-炔键合化学175对标记的细胞蛋白质进行荧光标记。该技术可与核糖体RNA靶向荧光原位杂交技术相结合，直接将分类学与原位活性175联系起来。BONCAT还与荧光活化细胞分选相结合，从复杂样品中分离出活性细胞，并通过DNA序列133进一步对其进行表征。此外，标记的蛋白质可以通过小珠捕获选择性地富集并进行蛋白质组分析176。这些方法的联合应用可以在不同的理化条件下对未培养微生物合成的蛋白质进行高通量示踪。虽然BOCAT由于细胞氨基酸摄取和代谢紊乱的差异而受到限制，但该技术为在单细胞水平上的原位活性的相对简单、廉价和高通量分析提供了灵活的工具。

**微流体**

能够在单细胞分辨率下对微生物进行高通量分析的设备对于微生物的快速培养和功能分析非常重要。微加工设备，如微流控“芯片上的实验室”技术，可以提供多种应用，包括从复杂的微生物组中分离单个细胞和种群177，建立体外细胞模型，促进合成微生物组的组装和在异质微环境条件下的实验178，以及用于快速监测和检测所需表型的在线诊断。这些应用仍处于开发的早期阶段，仍然存在一些挑战，包括可靠地检测液滴中的微生物、精确控制气体浓度、交叉污染和技术可获得性177179。

**自动化**

为了提高微生物组工程的重现性、生产量、效率和标准化，自动化的进步是必要的。这包括将液体处理机器人、微流控设备、自动化培养系统、在线物理化学测量传感器和软件纳入数据生成和分析工作流程。新出现的例子包括使用液体处理机器人与用于高通量培养的自动化微发酵平台82，或使用微流体技术自动分析数千个探测微生物群落相互作用的液滴实验114，180。这种自动化平台还可以集成一些功能工具（例如，单细胞分析和多组学），从而产生丰富的可重复数据集，可用于机器学习和其他大数据分析。

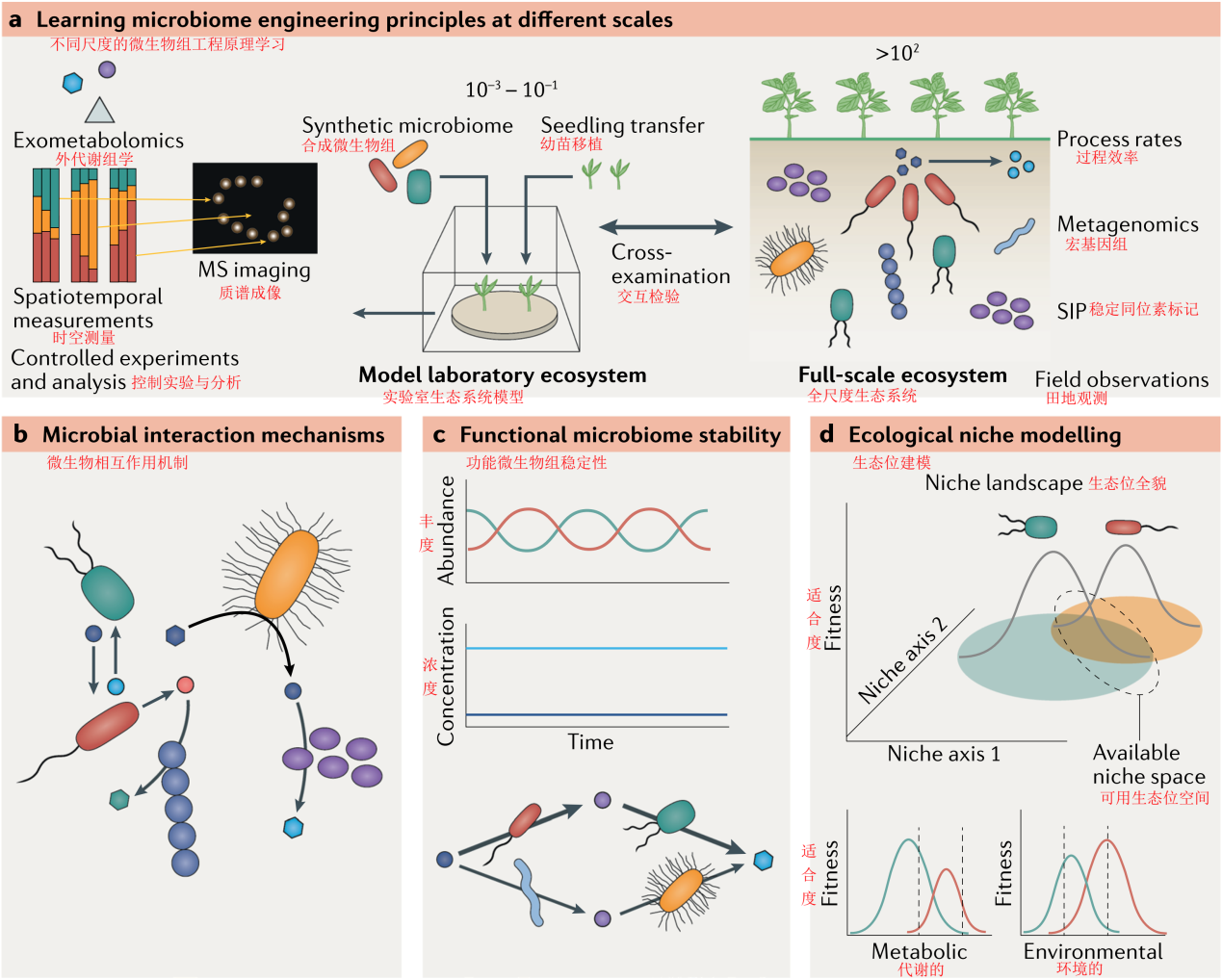


图5 |微生物组工程基本原理的学习。实验室生态系统模型可用于具有简化微生物组和环境特性的控制实验，代表介于纯实验室条件（如试管或烧瓶）和复杂自然环境（如土壤或海洋）之间的系统。需要在实验室规模的模型和自然复杂生态系统之间进行持续的交叉检验，以制定既在实际系统中可靠又在实验室中易于处理的工程原理和实践。这将需要多个利益相关者之间的密切合作，包括研究人员和最终用户（如医院或处理厂），他们在每个规模的具体问题上都有专业知识和经验（a部分）。为使系统化微生物组工程成为可能，需要学习的关键原则包括微生物相互作用机制（b部分）、控制功能稳定性和退化性的机制（c部分）以及定量绘制和模拟复杂生态系统生态位的框架（d部分）。MS，质谱；SIP，稳定同位素标记。

**方框3 |微生物组工程的新兴原理：生态位建模的案例**

生态位建模可用于系统地设计高阶特性，如工程微生物组的功能稳定性和鲁棒性。然而，要开发这样一个框架，就需要理解微生物组中保持多样性的机制，以及多样性如何赋予诸如功能稳定等特性。在这里，我们提出这种理解可应用设计-构建-测试-学习循环来回答关键问题：

**功能退化是否会导致生产力和功能稳定性？**

多样性与大型生物群落的生产力和功能稳定性相关，但多样性在改善微生物组功能和功能稳定性方面的作用仍然是开放的。对于微生物组工程，我们建议通过功能冗余（如前154所述）或更具体地说，功能简并来看待、讨论和定义多样性。这是一组物种在生态系统功能（例如，甲烷氧化、固氮或聚合物水解）中发挥相同作用的程度，但表现出相对于其他生理特征（例如，pH优化或生物膜形成）的简并性。这使得它们能够实现现实生态空间并共存。设计-建造-测试-学习循环期提供了一个极好的机会来理解功能简并的分子基础，并检查如何通过量化基本的和现实的生态位空间来预测出现的群落级特性，如抗干扰能力或对其他物种入侵的敏感性微生物组。我们认为生态位模型可以成为实现这一目标的一个特别有用的框架。

**微生物生态系统如何保持多样性？**

为了建立一个生态位模型框架，了解如何保持多样性将是很重要的。竞争排斥表明，在相同的生态位中144，两个具有相同资源需求的物种不能共存。因此，我们需要了解创造位空间和允许多样性发展和保持的机制。例如，时空变异、休眠、捕食、营养负荷、次生代谢产物产生和抗性、细胞运动和生物膜形成等过程在生态位分化中有什么作用？如何操作这些过程，以达到并保持微生物组中所需的功能简并水平？对这些问题的回答将提供微生物组工程机制，以设计和控制生态位空间以获得所需的微生物组特性。

**生态位模型是如何作为微生物组工程的基础的？**

为了使系统工程具有理想的高阶微生物组特性，我们提出微生物组工程为生态位建模开发了一个框架。该框架的目标是通过整合多组学数据、生理信息、养分有效性和环境参数，量化群落和个体的基本生态位和现实生态位空间，并利用它们制定控制微生物组合作和竞争的策略。为了实现这一目标，需要定义一个物种或团体的基本生态位和现实生态位的新数学表示，以及描述环境变量响应的适应度函数。当纳入微生物组模型时，该框架将允许对高阶特性进行生态预测，并量化合作和竞争的微生物组全貌。此外，这些框架将有助于指导尚未解决的重要微生物组设计问题，例如功能冗余和最小多样性之间的权衡。

**展望**

微生物组工程的真正进展将需要多个DBTL回合来捕获必要的生态学原理，以精确的方式操纵微生物组，并获得可预测的结果（图1）。例如，将在以前的DBTL循环中发现的物种间电子转移直接纳入代谢模型和生物反应器结构（例如，通过添加导电材料）可以优化废物的沼气生产效率27，或者设计工程大肠杆菌来控制先前发现的自我诱导物的水平，可以在生物失调的条件下使肠道微生物组朝着更健康的方向发展156。然而，开发快速周转的新知识和工具将需要下一代基础设施，用于数据收集、数据共享和知识集成。为了加快进度，开发学习阶段所需的预测能力是一个优先事项。实验室生态系统模型结合自动化技术的进步，如液体处理机器人、微流体和数据分析管道157，158，将为以严格和可重复的方式测试多种设计提供一个起点。从这一过程中获取新知识，并将信息整合到随后的DBTL循环中，将加速微生物组工程的发展，创造新颖的生物技术和实践，用于医学、农业、制造业和环境中微生物组的管理。

在这些领域推进微生物工程的例子包括阐明噬菌体与代谢产物交叉喂养在控制瘤胃碳代谢中的作用159，利用未开发的厌氧真菌-细菌联合体提高生物量转化为有价值的生物制品160,161，创造微流控细胞分选技术，从高多样性的样品中自动分选稳定同位素标记的细胞，用于随后的多组学分析或培养162，开发原位宏基因组工程工具，将新功能引入原生环境中的微生物111。为了推进DBTL方法，具有实验（例如，培养、分子遗传学或生物化学）、计算（例如，代谢建模、机器学习或生物信息学）、自动化（例如，机器人学或微流体学）和实践（例如，专业工程师或医生）等专业知识的跨学科研究团队是必不可少的。鉴于我们对微生物生态学的初步理解，微生物组工程的未来之路似乎很漫长；然而，围绕DBTL循环构建研究和技术开发为推进微生物组工程和提供解决紧迫社会和环境问题的创新解决方案提供了一个有希望的途径。

**Common principles and best practices for engineering microbiomes**Christopher E. Lawson 1\*, William R. Harcombe 2, Roland Hatzenpichler 3,4,5, Stephen R. Lindemann 6, Frank E. Löffler 7,8, Michelle A. O’Malley 9,10, Héctor García Martín 10,11,12,13, Brian F. Pfleger 14, Lutgarde Raskin15, Ophelia S. Venturelli14,16,17, David G. Weissbrodt 18, Daniel R. Noguera1,19 and Katherine D. McMahon 1,17\*

1Department of Civil and environmental engineering, university of Wisconsin-Madison, Madison, WI, USA.  
2Department of ecology, evolution and Behavior, university of Minnesota, Saint Paul, MN, USA.  
3Department of Chemistry and Biochemistry, montana State university, Bozeman, MT, USA.  
4Center for Biofilm engineering, montana State university, Bozeman, MT, USA.  
5Thermal Biology Institute, montana State university, Bozeman, MT, USA.  
6Department of Food Sciences, Purdue university, West lafayette, IN, USA.  
7Center for environmental Biotechnology, university of Tennessee-Knoxville, Knoxville, TN, USA.  
8Biosciences Division, oak Ridge National laboratory, oak Ridge, TN, USA.  
9Department of Chemical engineering, university of California, Santa Barbara, Santa Barbra, CA, USA.  
10Doe Joint Bioenergy Institute, emeryville, CA, USA.  
11Biological Systems and engineering Division, lawrence Berkeley National laboratory, Berkeley, CA, USA.  
12Doe Agile BioFoundry, emeryville, CA, USA.  
13Basque Center for Applied mathematics, Bilbao, Spain.  
14Department of Chemical and Biological engineering, university of Wisconsin-madison, madison, WI, USA.  
15Department of Civil and environmental engineering, university of michigan, Ann Arbor, MI, USA.  
16Department of Biochemistry, university of Wisconsin-Madison, Madison, WI, USA.

17Department of Bacteriology, university of Wisconsin-Madison, Madison, WI, USA.  
18Department of Biotechnology, Delft university of Technology, Delft, Netherlands.  
19Doe Great lakes Bioenergy Research Center, Madison, WI, USA.

Abstract | Despite broad scientific interest in harnessing the power of Earth’s microbiomes, knowledge gaps hinder their efficient use for addressing urgent societal and environmental challenges. We argue that structuring research and technology developments around a design–build–test–learn (DBTL) cycle will advance microbiome engineering and spur new discoveries of the basic scientific principles governing microbiome function. In this Review, we present key elements of an iterative DBTL cycle for microbiome engineering, focusing on generalizable approaches, including top-down and bottom-up design processes, synthetic and self-assembled construction methods, and emerging tools to analyse microbiome function. These approaches can be used to harness microbiomes for broad applications related to medicine, agriculture, energy and the environment. We also discuss key challenges and opportunities of each approach and synthesize them into best practice guidelines for engineering microbiomes. We anticipate that adoption of a DBTL framework will rapidly advance microbiome-based biotechnologies aimed at improving human and animal health, agriculture and enabling the bioeconomy.

Microbiome science  
Discovery and testing of fundamental principles governing microbiome assembly and function.

Microbiome engineering

Leveraging fundamental scientific principles and quantitative design to create microbiomes that perform desired functions.

Syntrophy

An obligately mutualistic process that is mediated by metabolite cross-feeding between two or more organisms that cannot be catalysed by one organism alone.

Metaphenotypes

Sets of emergent functions of a microbiome resulting from the interactions between individual microbial genomes (metagenome) and their interaction with the environment.

Ecological engineering

The process of designing and operating bioreactors and other engineered systems to foster the development of specific microbial communities that can perform desired functions.

Functional guilds

groups of organisms that use similar resources (for example, electron donors, electron acceptors or carbon source)and occupy a similar ecological niche.

Keystone species

An organism that has a disproportionately large effect on maintaining the microbiome’s function and microbial interactions (both between microorganisms and with the environment).

Flux balance analysis

A constraint-based mathematical modelling technique for simulating metabolic fluxes through a metabolic network reconstructed from genomic information.

Ensemble modelling

Use of multiple models to address uncertainty by simulating a set of possibilities and selecting those consistent with measured data.

Machine learning

A technique used to build predictive models through patterns and inferences obtained from sample data rather than explicit or mechanistic relationships.

Technoeconomic assessment

A tool used to evaluate the technical and economic viability of an integrated process through a combination of process design, modelling and economic evaluation.

Life cycle analysis

A tool used to evaluate the environmental impacts associated with all stages of a product’s life, such as energy and water consumption, and air pollutant and greenhouse gas emissions.

Self-assembled microbiome

A microbiome built through environmental manipulation that selects for desired functions.

Synthetic microbiome

A microbiome built by combining predefined axenic or enrichment cultures to achieve a desired function. M I C R O B I O M E T R A C TA B

Integrative and conjugative elements

Mobile genetic elements able to integrate into DNA sites via site-specific recombination that carry genes encoding the machinery necessary for conjugation.

Exometabolomics

An analytical technique to quantify extracellular small-molecule metabolites from environmental and/or biological samples typically through gas/liquid chromatography–mass spectrometry or nuclear magnetic resonance spectroscopy.

Off-gas analysis

The monitoring of gas flow rate and chemical composition (for example, carbon dioxide, hydrogen, methane) produced from a biological system.

Structure–function relationships

The influence of the microbiome’s three-dimensional spatial organization on its function.

Generalized Lotka–Volterra equations

A set of ordinary differential equations used to represent population dynamics based on experimentally inferred species interaction parameters.

Fundamental niche

The entire set of environmental conditions in which an organism can survive and reproduce (that is, an organism’s niche in the absence of interspecific competition).

Realized niche

The set of environmental conditions used by a species after consideration of interspecific competition (competition, predation and other factors).

Microbial communities have seemingly limitless capabilities, driving Earth’s biogeochemical cycles and occupying every environmental niche1,2. Engineers and scientists have tapped into this power for a long time; for example, by manipulating soil microbiomes to increase crop productivity3, by stimulating naturally occurring or introduced microbiomes to remediate contaminated groundwater4 or by building reactor microbiomes to recover valuable resources from wastewater5. Although these accomplishments highlight the valuable functions of microbiomes, the vast majority of the microbial world’s transformative capabilities have yet to be unlocked and harnessed. Recent insights driven by DNA sequencing have shed light on the high genetic diversity of not-yetcultured microorganisms and their crucial roles in diverse ecosystems6,7, providing a window on potentially novel biotechnology applications.

In recognition of this unlocked potential, funding agencies and the international science community have called for a global effort to advance microbiome research8,9. These initiatives have recognized the need for microbiome science to move beyond descriptive studies and embrace a systems approach that generates the mechanistic, predictive and actionable understanding that makes possible rational microbiome engineering8.

However, achieving this transition is hindered by the lack of tractable experimental systems that permit the detailed functional investigation of microbiomes, the large pool of microbiome gene and metabolite functions that remain unknown10, the many uncharacterized interactions (for example, syntrophy) between microorganisms11, inadequate tools to accurately measure and simulate microbiome functions across time and space, and the limited availability of approaches to precisely manipulate microbiome structure and function.

Integrating basic scientific discovery with engineering can overcome these challenges and develop innovative solutions that support sustainable natural resources management and human and animal health. In particular, engineering approaches can be used to create experimental systems that permit the testing of conceptual knowledge and extraction of new knowledge that advances microbiome research. To accelerate both scientific discovery and translation into innovative solutions, we propose that microbiome engineering adopt an iterative design–build–test–learn (DBTL) cycle to structure research and the technology development process. This cycle involves developing an initial microbiome design or preliminary model system to achieve a defined engineering goal, building the microbiome, testing its function against a set of specified metrics to determine whether the design–build solution(s) produced the design objective (that is, establish causation), learning what worked and what did not work (and why) and incorporating new knowledge into the decision making process of subsequent DBTL cycles (Fig. 1). This approach has been used successfully in manufacturing12, metabolic engineering13 and entrepreneurship (‘build, measure, learn’)14, and could rapidly advance our ability to develop much needed tools and design concepts for harnessing microbiomes, delivering innovative solutions and advancing scientific knowledge.

In this Review, we present key elements of an iterative DBTL approach that can be implemented to advance the rational engineering of microbiomes for functions that benefit society. We review diverse approaches to harness microbiomes in medical, agricultural, energy and environmental applications, and identify current challenges and opportunities associated with implementing each DBTL phase. Finally, we discuss how the DBTL cycle can be applied to build model systems to establish basic principles of microbial ecosystems and provide an outlook on the frontiers of microbiome engineering.

Designing microbiomes  
Because of the high complexity and limited understanding of molecular-scale microbiome processes, microbiome design has conventionally followed a top-down approach. This approach tries to predict how ecosystem level controls can create a microbiome with desired functions. However, recent advances in multi-omics have provided opportunities to design microbiomes from the bottom up by predicting how the control of metabolic networks and their interactions can create a microbiome with desired functions. Combined, these approaches offer complementary strategies to design microbiomes for specific engineering goals, ranging from sustainable wastewater treatment to curing microbiome-associated human diseases.

*Top-down design.*

Rather than deciding which organisms and detailed metabolic pathways to use a priori, the top-down approach uses carefully selected environmental variables (such as certain substrate loading rates, mean cell retention times and redox conditions) that force an existing microbiome (naturally occurring or inoculated) through ecological selection to perform the desired biological processes (or ‘metaphenotypes’)15(Fig. 2). Here, ‘top’ refers to the ecosystem in which the desired biological process occurs and ‘top-down design’ denotes the methods used to predict how manipulation of the ecosystem’s physical, chemical and biological processes (that is, ecosystem processes) obtains the desired function. Predicting how to manipulate an ecosystem is informed by principles of ecological engineering16 (also known as microbial resource management17 or microbial community engineering18). This requires engineers to conceptualize the system as an ecosystem model that captures system inputs and outputs, physicochemical conditions (pH, temperature, redox potential and so on), known abiotic and biotic processes, and environmental variables, and how their manipulation may promote or inhibit the biological process(es) being optimized19,20. Subsequently, mathematical modeling is used to perform mass balance analysis of chemicals and relevant microorganisms in the system and simulate chemical and biochemical transformation rates. These process-based models capture microbiome functions by representing key physiological or functional guilds of microorganisms (such as methanogens, fermenters, nitrifiers or phototrophs) with specific stoichiometric parameters (growth and product yields) and kinetic parameters (maximum specific growth rate, substrate uptake rate and substrate affinity)21–23. The models can also integrate equations describing the three-dimensional physical transport processes (diffusion, advection and dispersion) acting on chemicals and microorganisms, which are especially important in spatially structured systems such as biofilms24,25.

*Bottom-up design.*

Although the conventional top down design approach for microbiome engineering offers a framework for macro-scale processes and has been widely successful for wastewater treatment21 and bioremediation4, it often neglects the complex in situ metabolic networks driving microbial and linked chemical transformations26 and ignores processes that depend on intricate interactions between community members; for example, syntrophic interactions through direct interspecies electron transfer27. As a consequence, molecular-scale microbiome processes are often ignored during design, limiting system optimization through molecular-scale mechanistic insight. Recent advances in multi-omics and automation technology (for example, in metagenomics and microfluidics) have enabled researchers to develop bottom-up approaches and focus on engineering the microbiome’s metabolic network and microbial interactions. Here, ‘bottom’ refers to the metabolic networks of individual organisms in the microbiome (expressed from their genomes) and ‘bottom-up design’ denotes the methods used to predict how metabolic flux through these interacting networks generates the desired output. The general design process is to obtain the genomes of individual members of the microbiome28 (especially keystone species29, when known30), reconstruct their metabolic networks31,32 and use modelling33 and/or network analysis tools34 to guide design (Fig. 2). Existing constraint-based methods such as flux balance analysis provide a suitable framework for exploring which combinations of chemical transformations are possible using quantitative models, in which individual populations’ reactions and metabolites can be compartmentalized and metabolic fluxes within and between populations can be simulated using optimality principles35. These models can also simulate steady-state flux distributions over time and space36,37 and can be integrated into process-based and/or individual-based models38 to predict metaphenotypes, self-organizing spatial patterns and other emergent behaviours. Such bottom-up tools provide the engineer with a computational framework to systematically evaluate the metabolic networks driving biological processes and ecological interactions, and a platform for rationally designing microbiomes with specific properties, such as distributed pathways39,40, modular species interactions41, community resistance and resilience42 and spatiotemporal organization43 that optimize ecosystem function and stability. However, most of these bottomup design examples are based on simple communities with model organisms (such as *Escherichia coli* and *Saccharomyces cerevisiae*) that have engineered dependencies. Therefore, extending these designs to systems with non-model organisms of tens to hundreds of different species will require deeper insights into their metabolism and the principles governing their interactions and higher-order behaviour. There are major challenges to implementing bottom up design, including inaccurate and/or incomplete metabolic network reconstructions, unknown functions of many genes, proteins and metabolites, poorly understood evolutionary pressures driving individual and community-level phenotypes and limited understanding of gene, metabolic and ecosystem regulatory schemes (for example, quorum sensing signal–response systems)44. These limitations lead to high model uncertainty because key constraints on pathway stoichiometry and enzyme kinetics are either inappropriate or missing, and objective functions fail to capture the true evolutionary drivers of cell behaviour45, ultimately leading to poor predictions of in situ metaphenotypes. As a starting point for bottom-up design, core metabolic models that capture central carbon and energy metabolism can be reconstructed from genome annotations and known physiological information. The predictive power of these models may be limited initially, as they ignore regulatory information, pathway kinetics, secondary metabolism and evolution. However, when this knowledge is acquired and becomes incorporated into metabolic models through multiple cycles of testing and learning, accurate predictions of system function (for example, metabolic fluxes and metabolite exchange) may emerge. As a complementary approach, data driven modeling techniques such as ensemble modeling and machine learning may offer more rapid methods to predict microbiome metabolic processes or obtain constraints and parameters required for microbiome modeling, without the need for detailed mechanistic understanding of metabolic regulation46,47. Such modeling frameworks have been used to predict pathway fluxes from proteomic and metabolomic data48, to improve metabolite cross-feeding predictions through ensemble modeling-based flux balance analysis49 and to obtain key catalytic turnover numbers needed for metabolic models50. Although these approaches are flexible and generalizable enough to be applied to microbial communities, they require substantial amounts of experimental data on the metabolism of individual strains and interacting communities. This information could be leveraged from prior test phases (for example, from high-throughput phenotypic screens and multi-omics) to allow data-driven design.

*Integrated design.*

Moving forward, we envision that a judiciously balanced blend of top-down and bottom-up approaches will be needed for successful microbiome design, especially when one is working with complex microbiomes, such as human microbiota or activated sludge (Fig. 2). A blended approach could involve selecting both undefined mixtures and defined consortia to achieve desired microbiome functions, merging process-based models with bottom-up metabolic models reconstructed from meta-omic information to simulate ecosystem processes, mass balances and metabolite fluxes, and using genome-derived information to develop community selection strategies. Capturing higher-order properties in design, such as functional stability and dynamics, will likely also require top-down and bottom-up approaches to converge. In particular, new mathematical modeling approaches that quantify mechanisms of functional degeneracy, niche complementarity and network buffering51 using a metabolic framework may allow microbiome diversity to be optimized to sustain desired functions in situ. The need for a more comprehensive representation of microbiome metabolism will depend on the specific engineering objective and the degree of ecosystem tractability. For example, a more detailed representation of anaerobic microbiome metabolism is likely required to convert biomass into a specific commodity chemical instead of methane because finer control over metabolism would be needed. In either case, the design phase encompasses defining the engineering problem, developing conceptual and quantitative models, identifying key biological processes to be manipulated and evaluating multiple candidate design alternatives.

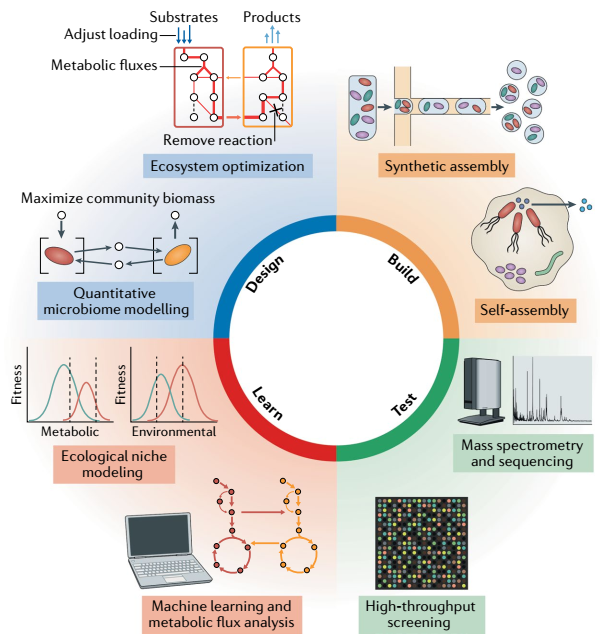


Fig. 1 | The design–build–test–learn cycle for microbiome engineering. The key aspects and approaches of each phase of the design–build–test–learn cycle are presented. The cycle starts with a defined engineering objective that determines the design and produces an engineered microbiome that performs the desired function(s).

M I C R O B I O M E T R A C TA B I L I T Y A N D T R A N S L AT I O N  
*Practical design steps.*

There are five key steps when one is designing microbiomes, in particular complex microbiomes: defining the engineering problem, developing a conceptual ecosystem model, creating an quantitative model, identifying the microbiome process(es) to be engineered and developing and evaluating candidate design strategies. To drive the DBTL cycle, a clear definition of the problem with measurable design objectives must be established. These objectives could specify desired outcomes such as product titre, rates and yields, pollutant removal efficiency, crop productivity, or degree of functional stability and robustness. Design objectives should be complemented by technoeconomic assessments and/or life cycle analysis to ensure that solutions are economically feasible and have positive environmental and societal impacts52,53. Conceptual ecosystem models can be used to contextualize the problem. Such models capture system boundaries, inputs and outputs, major pathways of carbon and nutrient flows, key organisms and interspecies interactions responsible for those transformations and factors influencing their activity (for example, pH, temperature, redox potential and residence times)19. They provide a concept map that describes current understanding of interactions between the microbiome and physical, chemical and biological components of the ecosystem, helping to identify important gaps in system understanding and needs for data collection. At this stage, all relevant information should be collected from the literature, existing data (for example, from the Human Microbiome Project)54 and online databases (for example, MiDAS (microbial database for activated sludge))55 for ecosystem characterization. This includes reference genomes and physiological information for keystone organisms, previous multi-omic datasets, ecosystem physicochemical properties (such as pH, temperature and chemical concentrations) and processes (such as photochemical reactions and hydrogeological processes), site characteristics (such as nutrient loadings and dynamics, soil profiles and gut anatomy) and all other information needed to characterize the ecosystem. Missing information, such as unknown biochemical pathways and organisms that mediate them, can be targeted during the build, test and learn phases. This conceptual ecosystem model can be used by the scientific community for proposing and testing theories and serves as a road map for developing quantitative simulation tools. Construction of quantitative modeling tools that allow the calculation and simulation of metabolic fluxes, microorganism abundances, mass balances and ecosystem physicochemical parameters is critical for the systematic design of microbiomes. Several approaches could be used to create such models, including mechanistic metabolic modelling33, process-based modelling21, data-driven modelling (for example, machine learning)48 and individual-based modelling38 or their combination. Regardless of the approach, the simulation of complex microbiomes will likely require simplification based on experimentally valid assumptions. Simplification could include reducing the model to a set of core or keystone organisms that represent important functional guilds and control major carbon and energy flows, or reducing the metabolic network size of organisms to central carbon and energy metabolism. Moving forward, it will be important to ensure that models undergo rigorous experimental validation and iteration during build–test–learn cycles to increase their utility and widespread use in microbiome engineering and to identify when modelling efforts fail, revealing gaps in conceptual understanding that can further facilitate model redesign and improvement. Quantitative microbiome modelling (such as dynamic flux balance analysis) helps to identify the core and peripheral biochemical pathways that need to be directly manipulated, added or removed to achieve the desired engineering objective. Objectives could include increasing butyrate production and non-digestible carbohydrate degradation by fermenting bacteria in the human gut, preventing toxin biosynthesis by cyanobacteria in freshwater ecosystems or stimulating the degradation of toxic chloroorganics by bioaugmentation with organohalide-respiring bacteria. Microbiome modelling can predict how environmental (such as substrate loading, pH and solids retention time) or genetic manipulation (such as gene knockouts, pathway additions and forced dependencies) could optimize microbiome functions towards the engineering objective. If necessary, synthetic microorganisms could be designed to improve microbiome function. Such synthetic microorganisms will need to be evaluated for their ability to cooperate and compete with existing microbiome members under relevant environmental conditions.

Fig. 2 | Top-down and bottom-up approaches to design microbiomes. The left panel illustrates a bottom-up design workflow starting from isolates. Physiological characterization of individual organisms is performed, and metabolic modelling is used to design consortia for the desired function (produce light blue compound from dark blue compound). Genetic engineering and synthetic biology strategies are used to optimize system function (identifying gene editing targets that reroute metabolic flux away from toxin (purple) and towards the desired product; designing of toxin reporter strain). The right panel illustrates a top-down design starting with an inoculum containing uncultivated microorganisms from the environment. Community characterization of an undefined microbiome is performed, and bioprocess modelling (mass balance analysis including kinetics and microbial growth) is used to develop selection strategies to achieve the desired function (produce light blue compound from dark blue compound). Bioprocess engineering (for example, reactor design) is used to optimize system function. The middle panel shows an integrated top-down, bottom-up design. Combinations of undefined consortia and defined cultures are selected to achieve desired functions. Community characterization is performed, and microbiome modelling that integrates process-based simulation with metabolic modelling is used to develop selection strategies and analyse microbiome metabolic fluxes. The shapes of the microorganisms represent different isolates or communities selected during design. OD, optical density; S, substrate.

Building microbiomes

The build phase consists of physically assembling the designed microbiome by either top-down manipulation of a natural community (that is, a self-assembled microbiome) or bottom-up assembly using axenic or enrichment cultures of naturally occurring or engineered microorganisms (that is, a synthetic microbiome). The build phase aims to bring the design specifications and predictions into reality.

*Building by self-assembly.* Self-assembled microbiomes may include those built as open mixed cultures using reactor engineering (for example, wastewater treatment bioreactor) or biostimulation (for example, additions to soils, sediments or groundwater aquifers), in which construction creates an environment that promotes the growth and desirable activity of resident microorganisms. Examples include manipulating reactor hydrodynamics to immobilize slow-growing microorganisms into compact granules that allow their retention and proliferation56,57, use of non-human-digestible carbohydrates to stimulate fermentative production of short-chain fatty acids in the gut58 or adding electron donors to drive the metabolism of organohalide respiring bacteria during bioremediation of toxic chlorinated contaminants4. This approach is powerful when differences in physiological and physicochemical properties between functional guilds can be exploited for assembly through environmental manipulation (for example, differences in growth rates59, main electron donors and acceptors4,60, substrate affinities, cell and/or biofilm densities61 and redox gradients). However, it can be limited when more fine-scale control over microbial metabolism and interactions is necessary (for example, controlling complex competitive interactions62, producing valuable bioproducts at high yields and purity63 or controlling organisms with versatile lifestyles64). In addition, new strategies for evolutionary engineering have emerged as promising tools to build self assembled microbiomes. Controlled exposure of an initial microbiome to multiple selection cycles and/or regimes results in the microbiome gaining or optimizing specific functions through adaptation or evolution. For example, successively transferring the microbiomes that maximize plant traits has generated microbiomes that increase plant biomass65 and flowering time66. Response to community-level selection will often be driven by enrichment or adaptation of single species67,68; however, selection for production of community biomass has also been shown to enhance desired species interactions in defined two-species and three-species co-cultures37,69. Re-examining selection experiments to understand when and how mutations and/or adaptations altered microbiome phenotypes could elucidate the mechanisms underlying microbiome fitness optimization and inform design, as has been shown for *E. coli* in laboratory evolution experiments70,71. As similar evolutionary approaches (for example, adaptive laboratory evolution) have also been successfully applied to optimize strains for metabolic engineering72, extension of experimental and computational protocols already developed for individual microorganisms to microbiomes could streamline the design phase and reduce the time required to complete evolution experiments.

*Building synthetic microbiomes.*

Direct construction of microbiomes using axenic or enrichment cultures is also promising because of reduced complexity and the use of microorganisms that are genetically tractable and/or well characterized. This bottom-up approach makes the growing suite of synthetic biology tools accessible for microbiome construction and optimization. An early approach for building microbiomes directly from cultured microorganisms was bioaugmentation. Here, defined laboratory consortia are added back to the environment to enhance the degradation rates of specific contaminants. A successful example was the addition of consortia containing organohalide-respiring bacteria of the class Dehalococcoidia to contaminated groundwater aquifers and sediments to speed up the degradation of toxic chlorinated pollutants. Crucial for the success of this approach was detailed knowledge of the physiology, nutritional requirements and potential ecological interactions of the keystone dechlorinators with other microorganisms and the geochemical environment4. However, in contrast to the success for chlorinated contaminants, bioaugmentation approaches have largely failed for oil spills. Unlike organohalide-respiring Dehalococcoidia members, which fill a unique ecological niche and cannot grow without the chlorinated contaminants, organisms capable of degrading oil hydrocarbons (especially aerobic bacteria) are ubiquitous, metabolically versatile and do not depend on a specific substrate or redox couple for growth64. This metabolic versatility limits their utility for bioaugmentation given their unpredictable in situ activity. Other reasons why bioaugmentation can fail are that unrecognized mutualistic interactions and microorganisms performing critical functions are missing (for example, production of polysaccharide surfactants to increase hydrocarbon bioavailability)73 or that consortia selected under laboratory conditions are no longer competitive enough under harsh and/or variable field conditions74–76. These examples highlight the need to better understand the interaction networks of synthetic consortia, especially the roles of supporting interactions (secondary functions), and the competitive landscape in situ, which are often difficult to predict in complex ecosystems. Despite the appeal of building microbiomes from the bottom up and the growing collection of cultured microorganisms from specific habitats77,78, most microorganisms relevant for human health, agriculture and environmental applications remain uncultured, poorly characterized, genetically intractable and difficult to maintain, making the construction of synthetic microbiomes challenging. To capture this uncharacterized metabolic diversity, innovative isolation and controlled microbiome assembly techniques are needed, such as single-cell sorting79 coupled with high-throughput culturing (culturomics)80,81 and phenotyping82,83 across multiple conditions in parallel. Microfluidics84,85 (that is, creation and manipulation of microlitre droplets) can facilitate this approach. Microfluidic chips can allow automated assembly and analysis of microbial communities from axenic or enrichment cultures through droplet combination86, elimination of specific species87, sequencing, and multi-omics phenotyping of individual cells88,89. Combined with new gene editing techniques, such as CRISPR-based genomic tools90 that increase the efficiency of homologous recombination-based gene editing91,92, microfluidics could also automate synthetic biology techniques for the engineering of cells and microbiomes with novel capabilities93.

Another challenge with synthetic microbiomes is maintaining their functional stability in the laboratory or in open systems (for example, human gut, soil and wastewater treatment plants), which are susceptible to invasion by naturally occurring microorganisms and dynamic heterogeneous environments. As mentioned earlier, the major reason for the success of bioaugmentation with organohalide-respiring Dehalococcoidia members is their highly specialized lifestyle that enables them to occupy an open ecological niche using chlorinated electron acceptors. However, the functional stability of organisms with versatile lifestyles in open systems is much less predictable. Few studies have examined the functional stability of synthetic consortia in open systems, and the knowledge required to rationally engineer stable ecological interactions is limited. However, engineered bacteria have been successfully deployed as diagnostic sensors in the mammalian gut for up to 200 days maintaining robust function94,95. This feat, together with the bioaugmentation example of Dehalococcoidia4, demonstrates that synthetic consortia can form stable microbiomes with previously established community members, provided key players can compete with resident microorganisms. Observations from self-assembled microbiomes suggest that building communities with spatiotemporal organization will be important for achieving stable and multifunctional synthetic microbiomes. Highly diverse microbial communities, such as human microbiota or those used for wastewater treatment, self-assemble as biofilms, flocs or granules comprising multiple singlespecies microcolonies attached together via speciesspecific extracellular polymeric substances (including polysaccharides, proteins and DNA) and other poorly defined macromolecules (such as humics)96,97. These self-organizing microbial assemblages create diverse microenvironments and ecological niches that support the combination of seemingly incompatible functions (for example, both aerobic and anaerobic processes)98,99 and functionally diverse population structures that can compensate for disturbances, such as changes in nutrients, changes in physicochemical conditions or predation100,101. Although building such fine-scale and sophisticated architectures into synthetic microbiomes is nascent, microfluidic-based systems have been used to assemble simple communities with improved functional stability by controlling spatial structure and chemical communication102. Additionally, three-dimensional bioprinting platforms could allow the construction of spatially organized systems, in which populations can be physically separated while remaining chemically interactive103,104. How to scale these spatially defined structures from experimental laboratory systems to real-world applications remains to be resolved, although knowledge gained from test and learn phases with model systems (such as synthetic polysaccharide particles)105,106 should provide more insights. Until then, existing approaches based on top-down assembly and/or engineered biofilm carrier media107 could be used to build self-organized synthetic microbiomes with greater stability and functionality. Designing synthetic genetic circuits in engineered hosts that can robustly perform sense–compute–respond programmes in complex environments also remains a major challenge108. Therefore, it will be important to examine the molecular mechanisms that determine microbiome stability and adaptation to environmental perturbation in natural and engineered ecosystems to extract design principles that can be used for rationally engineering robust functions. Given the potential utility of genetically engineered microorganisms and microbiomes in diverse open environments, safeguards such as biocontainment systems (such as two-layered gene circuits and essential synthetic auxotrophies)109 will also require further development and will be needed as integral components of constructed synthetic microbiomes that use genetically modified organisms in the future.

*Integrating approaches.*

The ultimate goal for rational microbiome design is to develop tools that enable engineers to directly add, remove or modify specific functions and phenotypes in situ over a range of desirable operational conditions. One emerging technique with promise to achieve such flexibility is in situ metagenomic engineering110,111, which involves delivery of engineered mobile genetic elements to resident microorganisms. For example, donor strains engineered with integrative and conjugative elements have transferred DNA carrying a reporter and antibiotic resistance genes or multigene pathways (for example, nitrogen fixation gene (*nif*) cluster)112 to bacteria in highly heterogeneous and diverse environments, such as soil112 and the mammalian gut111. Further development of such tools in combination with existing CRISPR–Cas gene editing techniques would allow the precise manipulation of the microbiome’s metabolic network in situ, effectively combining self-assembled and synthetic microbiomes (Box 1; Fig. 3).

Testing microbiome function

The test phase involves measuring microbiome associated phenotypes and properties to determine the efficacy of the design–build solution. The measurements should determine whether the design outcomes were achieved (for example, measuring the titre, rate and yield of a bioproduct, pollutant removal efficiency or crop productivity) and whether the design–build solution was responsible for the observed outcome (establishing cause and effect). This typically requires readouts of ecosystem physicochemical properties (such as pH, temperature and chemical concentrations), as well as the stoichiometry and kinetics of key ecosystem processes and microbiome functions (such as biomass growth, chemical transformations, nutrient assimilation and metabolic fluxes). For example, acetate degradation rates and pathways to methane in an anaerobic digester microbiome could be tested using 13C-labelled acetate and online biogas analysis that measures the flux through acetoclastic methanogenesis versus syntrophic acetate oxidation coupled to hydrogenotrophic methanogenesis113. While the level of microbiome granularity measured during testing will depend on the specific design objectives and ecosystem complexity, the ability to quantify molecular microbial processes (for example, metabolic pathway rates and routes, enzyme activities and individual organism growth rates) goes beyond bulk activity measurements and allows testing of the specific mechanisms responsible for the observed microbiome functions. The challenge will be to develop tools that have highthroughput and are quantitative, affordable and easy to use, such that routine analyses of the microbiome over time and space and under dynamic conditions can be accomplished. Towards this goal, we envision a test phase comprising high-throughput phenotypic screening of microbiome design–build solutions, followed by deeper investigation of promising solutions using multi-omic and metabolic flux analyses to obtain greater insights into underlying mechanisms (Fig. 4). High-throughput phenotypic testing of constructed microbiomes could be achieved using droplet microfluidics, as recently demonstrated for screening ~100,000 synthetic communities114. Fully automated microbioreactor platforms that combine liquid handling and advanced sensing with microtitre plates or scaled-down bioreactor cultivation could also be used82,83. Combined with emerging methods to measure metabolic network activity and metabolic processes in heterogeneous environments (Box 2), rich information will be obtained to facilitate learning.

Box 1 | A design–build–test–learn cycle to create synthetic microbiomes with desired functions

We present a generalized design–build–test–learn cycle for creating synthetic microbiomes with desired functions, integrating both top-down and bottom-up approaches. We briefly describe two iterations of the cycle and identify opportunities for incorporating high-throughput approaches and automation to increase speed and reproducibility.

Top-down approach

*Design: identify biological process(es).* An example of a process to harness or replicate is anaerobic conversion of complex lignocellulosic biomass into valuable commodity chemicals. The initial design step includes selection of different innocula that may contain microorganisms with desired functions (for example, acid phase anaerobic digester sludge, herbivore rumen microbiome or others). Conceptual ecosystem models that include environmental parameters (pH, temperature, nutrients and so on) and expected functional guilds (hydrolytic bacteria, fermenting bacteria, methanogens and so on) are used to select enrichment variables.

*Build: enrich microbiomes from multiple sources.* Source innocula are cultivated under different environmental conditions to select for the desired function using real (for example, lignocellulosic hydrolysate or rumen fluid) and synthetic media. modulation of environmental conditions and medium composition is done to improve the desired function. For complex environments (such as soil), model laboratory ecosystems could be ideal platforms for microbiome enrichment146.

*Test: evaluate performance.* The performance of enriched microbiomes is tested on real and synthetic media using high-throughput phenotypic screens. High-throughput screens could be developed using microfluidic or automated microbioreactor experiments. Deeper multi-omic measurements (such as metagenomics, metatranscriptomics and metaproteomics) are collected from high-performing microbiomes.

*Learn: identify key functional roles of microbiome members.* Besides key functions, bottlenecks for the desired function are identified using metabolic reconstruction and multi-omic analysis. This understanding helps to refine conceptual models of microbiome function and create quantitative models.

Bottom-up approach

*Design: identify new potential microbial partners.* In silico metabolic modelling is used to screen high-performing microbiome enrichments for interacting microorganisms. metagenome-assembled genomes can be used to reconstruct metabolic models of key microbiome members. Automated computational workflows (together with manual curation) will accelerate model building. Flux balance analysis is used to predict each microorganism’s requirements for optimal growth and activity, and unify individual metabolic models into a microbiome model to identify new potential partners that improve the design objective (for example, higher titres, rates or yields of valuable product).

*Build: recombine key microorganisms into new synthetic consortia.* Following their isolation or enrichment, key microorganisms are assembled into new synthetic consortia on the basis of in silico predictions at various ratios (for example, 1:1, 1:10). microfluidic devices and/or liquid-handling robots could be used for high-throughput isolation and recombination.

*Test: test function and stability of consortia.* High-throughput phenotypic screening coupled with multi-omic measurements can be used for testing. This step should also include validation of predicted metabolisms of individual isolates or enrichments.

*Learn: identify microbial interactions that control function.* Analysing the metabolism of microorganisms growing in consortia versus in isolation using metabolic flux analysis can identify important mechanisms and interactions. This understanding can be used to propose how microbiome function and stability could be optimized by environmental manipulation and/or in situ genome engineering.

*Microbiome metabolic network activity.*

To test predictions of microbiome function at a systems level, measurement of the microbiome’s in situ metabolic network structure and activity is critical. Multi-omic approaches (metagenomics, metatranscriptomics, metaproteomics and metabolomics) combined with bioinformatic tools have allowed the genome-centric analysis of individual species (or even strains)115 within microbiomes and global measurement of sequences, proteins and metabolites116–118. These tools measure the microbiome’s components on a spectrum from functional potential (for example, gene abundance) to expressed products (for example, protein and metabolite abundance) and through their combined activity produce microbiome metaphenotypes that drive system function. Currently, multi-omic approaches used to infer microbiome function have focused on correlating gene abundances or gene expression data across time and space with ecosystem geochemical data or process rates. This has included measurements of key functional genes and transcripts using quantitative PCR assays (for example, ammonia monooxygenase)119, microarrays (for example, GeoChip)120 or untargeted high-throughput approaches (metatranscriptome and/or metaproteome). Although useful for overall system characterization and discovery, these approaches focus on measuring the components or ‘parts list’ of the system, which are often limited predictors of emergent phenotypes due to metabolic network complexity, interactions and regulation121,122. Therefore, new approaches and tools are needed to measure the in situ stoichiometry and fluxes of microbiome metabolic networks to permit the direct testing of design predictions and offer mechanistic insights into metabolic regulation. Metabolic flux analysis is the most authoritative method for measuring in vivo fluxes. This method calculates fluxes from metabolite stable isotope measurements obtained during isotopic labelling experiments using metabolic network modelling123. Although metabolic flux analysis has been used to measure fluxes in co-cultures124, flux analysis in communities is challenging because metabolite pools cannot be easily assigned to individual cells, and the number of possible reactions in a microbiome greatly exceed those of an individual organism. Nonetheless, isotopic tracers combined with exometabolomics and/or off-gas analysis have been used to determine process fluxes driving important microbiome functions, such as syntrophic acetate oxidation and methanogenesis during anaerobic digestion116. To circumvent the challenges with metabolite measurements, a method analysing labelling patterns from short peptides instead of amino acids for metabolic flux analysis was proposed125. Peptides can be assigned to individual species in a microbiome using high-throughput metaproteomic approaches, which opens the door to determining fluxes in microbial communities (that is, to ‘metafluxomics’). Given that fluxes represent the final outcome of cellular regulation across all levels126, further development and demonstration of metafluxomics will be essential for advancing microbiome engineering efforts and our understanding of metabolic regulation in microbiomes. This will also require new software packages for associated computational analyses, similar to existing 13C metabolic flux analysis software127. Such data may also allow metabolic modelers to infer, rather than assume, community-level and individual-level objective functions and to identify new constraints, allowing the accurate prediction and measurement of reaction rates driving microbiome function.

Fig. 3 | Building self-assembled and synthetic microbiomes. a | A protocol for assembling synthetic microbiomes from multiple microbiome sources. Complex microbiomes can be taken apart into key functional members using automated microfluidic cell sorting techniques. Isolated or enriched members can then be recombined into synthetic consortia using liquid-handling robots for downstream screening and/or cultivation. b | Microbiome assembly can also be achieved through environmental selection via bioreactor manipulation or biostimulation (top) or by bioaugmentation with defined cultures (bottom). c | Another option is microbiome assembly through directed adaptation and/or evolution of the microbiome to acquire or optimize a desired function. d | In situ microbiome engineering can be used to add new functions to microbiomes residing in the environment.

*Measuring function in spatially heterogeneous environments.*

Most natural microbiomes, such as those associated with plants (for example, rhizosphere), humans (for example, oral microbiome) and industrial processes (for example, acid mine drainage), display highly organized spatial organization across micro-scale physicochemical gradients that directly influences microbiome function. For example, the spatial proximity of microorganisms can control whether they interact through diffusible substrates or direct transfer128, whereas variations in colony size can dramatically influence apparent substrate affinity constants and substrate competition between biofilm microorganisms129. Therefore, one of the biggest challenges will be to create tools that measure and report on microbiome spatial structure and function across all relevant scales (from micrometres to kilometres). Current methods to measure structure–function relationships have focused on the micrometre to millimetre scale using approaches such as fluorescence in situ hybridization combined with stable isotope labelling130, chemical fingerprinting131, mass spectrometry imaging132 and/or fluorescence-based bio-orthogonal non-canonical amino acid tagging133 (Box 2). Although these techniques have successfully identified the substrate use and activity patterns of spatially distributed microorganisms in microbiomes, they are limited by throughput and can examine and/or differentiate only a limited number of organisms. The integrated application of labelling techniques (for example, stable isotope labelling and bio-orthogonal non-canonical amino acid tagging) with metaproteomics and cell sorting (for example, fluorescence-activated cell sorting)133 could be used to measure the metabolic activity of microorganisms with high-throughput and with spatial resolution. Combined with microsensor devices that profile microenvironmental chemical properties (for example, through microelectrodes134 or engineered biosensors95), microbiome structure, microbiome function and ecosystem physicochemical parameters could be monitored in real time.

*Learning microbiome design principles.*

Progressing through the design, build and test phases of microbiome engineering presents a unique opportunity to learn from previous failures and successes, and to incorporate new knowledge into subsequent cycles. Indeed, the learn phase of the DBTL cycle is critical for success and for increasing microbiome engineering efficiency. To date there are no general strategies, techniques or approaches that guarantee success in translating information obtained from the test phase into new knowledge that informs the next design phase. Therefore, we stress the importance of devoting enough emphasis and resources to the learn phase early on so as to avoid, for example, the difficulties encountered in metabolic engineering due to a relative lack of investment in the learn step13.

Fig. 4 | Testing microbiome function. a | Isotopic tracers combined with the metaproteome could be used to measure microbiome metabolic flux by analysing isotopic labelling patterns of short peptides rather than amino acids (metabolome). b | Bio-orthogonal non-canonical amino acid tagging is a method for rapid profiling of anabolic processes (growth) in situ using either fluorescent detection or metaproteomics. c | Metagenomics, metatranscriptomics, metaproteomics and metabolomics can be integrated to reconstruct and analyse metabolic network expression in microbiomes. d | An automated microbioreactor platform allows high-throughput analysis of microbiome processes across diverse conditions (for example, with changing environmental or physiological variables). The platform can integrate tools for detailed functional analysis of individual microbiome members to complex communities. Green, red and yellow colours in process monitoring graphs correspond to microbioreactor well locations. HPG, homopropargylglycine; MS, mass spectrometry.

M I C R O B I O M E T R A C TA B I L I T Y A N D T R A N S L AT I O N  
R E V I E W S

Box 2 | A toolbox for measuring microbiome function

Multi-omics integration

The ability to assemble genomes from metagenomic data28 has enabled the genome-resolved analysis of individual transcriptomes63 and proteomes118 within diverse communities and greatly increased the interpretive power of multi-omic datasets. A key challenge moving forward will be the integration of metabolomic information163, both intracellular and extracellular, which cannot be readily assigned to individual members of the microbiome such as DNA, RNA and proteins can be. The large amount of unknown or poorly characterized genes, enzymes and metabolites currently limits the interpretive power of multi-omic information. It does, however, create novel targets for further biochemical studies. Advances in bioinformatic tools, such as data-driven approaches (for example, statistical or machine learning methods) and knowledge-based approaches (for example, interaction networks or genome-scale metabolic modelling)164,165, will be key to the success of systematic measurements of microbiome function through coherent multi-omics data integration.

Isotopic tracers

Isotopic tracers have a long history in functional analysis in both pure cultures and communities, and have been combined with DNA166, RNA167 and protein116 measurements to link individual populations to specific in situ functions. Moving forward, more efforts to incorporate isotopic tracers with multi-omics (especially metaproteomics and metabolomics) are needed to illuminate the complex metabolic networks within microbiomes. The combination of these techniques should also pave the way for measurement of intracellular and extracellular reaction rates (‘metafluxomics’)124,125, which has been one of the most powerful tools for elucidating in vivo phenotypes, pathway constraints and metabolic regulation in pure cultures used for engineering purposes.

Mass spectrometry imaging

mass spectrometry imaging (MSI) techniques visualize the distribution of elements and their isotopes as well as biomolecules within complex samples. MSI is well suited for the analysis of spatially structured microbiomes and for the investigation of cellular interactions. When combined with fluorescence in situ hybridization, MSI also allows the linking of microbiome structure with function168,169. The chemical coverage, spatial resolution and sample preparation that can be obtained with different MSI techniques depend on the ionization method used132. Although nanoscale secondary ion mass spectrometry (nanoSIMS) has superior lateral resolution compared with matrix-assisted laser desorption/ionization (MALDI) or desorption electrospray ionization (DESI; 50 nm, 3–50 mm and 100 mm, respectively), its relative chemical versatility is very low (elements and isotopes versus peptides, lipids, metabolites and other molecules). Therefore, nanoSIMS has generally been applied to study substrate use of single cells, whereas mAlDI–MSI has been used to visualize chemical interactions between populations132. Although mAlDI– MSI and DeSI–MSI are more accessible than nanoSImS170 and could be well positioned to visualize the broad range of chemical interactions within microbiomes, they have very low throughput and their lateral resolution and sensitivity currently prohibit single-cell metabolic profiling132. A technique that combines the best of these two methods is nanostructureinitiator mass spectrometry (NImS). NImS is a matrix-free desorption/ ionization technique that depends on initiator molecules trapped in 30-nm pores to achieve the ionization of small molecules adsorbed to the pore surface. NImS offers a lateral resolution of ~150 nm and is particularly well suited for the analyses of peptides and metabolites171. So far, NImS has seen only limited application in microbiology172,173. We expect advances that resolve these issues will soon make MSI a useful and more widely applied tool for functional analysis of microbiomes174.

Bio-orthogonal chemistry

metabolic labelling techniques, such as bio-orthogonal non-canonical amino acid tagging (BoNCAT), offer additional approaches to measure microbiome anabolic activity in situ. BONCAT is based on the in vivo translational incorporation of a non-canonical amino acid (for example, l-azidohomoalanine, an l-methionine surrogate) followed by fluorescent labelling of tagged cellular proteins by azide–alkyne click chemistry175. The technique can be used together with ribosomal RNA-targeted fluorescence in situ hybridization to directly link taxonomy with in situ activity175. BoNCAT has also been combined with fluorescence-activated cell sorting to separate active cells from complex samples and further characterize them by DNA sequencing133. In addition, tagged proteins can be selectively enriched through bead capture and subjected to proteomic analysis176. The combined application of these methods could allow the high-throughput tracking of newly synthesized proteins from uncultivated microorganisms under different physicochemical conditions. Although BoNCAT can be limited due to differences in cellular amino acid uptake and metabolic perturbation, the technique offers a flexible tool for the comparatively simple, inexpensive and high-throughput analysis of in situ activity on a single-cell level.

Microfluidics

Devices that allow the high-throughput analyses of microorganisms at single-cell resolution will be important for the rapid cultivation and functional analysis of microbiomes. microfabricated devices such as microfluidic ‘lab-on-chip’ technology could offer multiple applications, including isolation of individual cells and populations from complex microbiomes177, the creation of in vitro cell-based models that facilitate assembly of synthetic microbiomes and experimentation under heterogenous microenvironmental conditions178, and online diagnostics for rapid monitoring and detection of desired phenotypes. These applications are still in early stages of development, and several challenges remain, including reliable detection of microorganisms in droplets, precise control of gas concentrations, cross-contamination and technology accessibility177,179.

Automation

To increase the reproducibility, throughput, efficiency, and standardization of microbiome engineering, advances in automation will be necessary. This includes incorporating liquid-handling robots, microfluidic devices, automated cultivation systems, online physicochemical measurement sensors and software into data generation and analysis workflows. emerging examples include the use of liquid-handling robots coupled to automated microfermentation platforms for high-throughput cultivation82, or microfluidics to automate the analysis of thousands of droplet experiments that probe microbial community interactions114,180. Such automated platforms could also integrate several functional tools (for example, single-cell analyses and multi-omics), resulting in rich reproducible datasets that could be leveraged for machine learning and other big data analytics.

Further development of computational methods to formalize the learn phase will be needed, including machine learning algorithms48,135,136, metabolic flux analysis and constraint-based analysis36,124,125,137, ecosystem modelling approaches138 and regulatory network analysis139. Together, these analyses could isolate the principal drivers of microbiome interactions and function from large datasets to inform microbiome design. For example, generalized Lotka–Volterra equations could infer interacting species from temporal population dynamics data that become the starting point for bottom-up design140 or constraint-based analysis could be applied to identify key metabolite exchange reactions from 13C metabolomic data that increase flux simulation accuracy and improve design of anaerobic consortia137. More broadly, we envision the learn phase to focus on translating data into generalizable principles for microbiome engineering through the continuous refinement of conceptual knowledge and proposed theory (for example, from traditional macroecology)51,141–144 with each DBTL cycle. We propose that model laboratory ecosystems should be used to drive microbiome engineering inquiry and learning. Model laboratory ecosystems are experimental platforms that can replicate the physicochemical conditions of a complex environment (natural or engineered) in a simplified and controlled manner and contain model microbial communities (for example, the model rhizosphere microbiome THOR)145 that can be used as testing grounds for learning how to design, construct and optimize engineered microbiomes. These ecosystems have reduced complexity, are accessible for experimentation and can be established in a reproducible manner, which is often not possible when one is working in natural environments. Recently, model laboratory ecosystems have been developed to study plant–soil microbiome interactions146. These fabricated ecosystems use three-dimensional printing, sensing and analytical and imagining technologies to create an experimental device that replicates the native soil ecosystem, in which microorganism and host phenotypes can be monitored in response to changing variables, allowing the systematic dissection of microbial interactions and metabolite exchanges influencing plant health146,147. Fabricated ecosystems offer a middle ground between model organisms and complex natural microbiomes, and can be established collaboratively between expert investigators to create standardized and reproducible devices and protocols for dissemination to the broader research community. Such model systems offer the ability to experimentally develop engineered microbiomes with desired functions in a tractable manner, and permit results to be compared with results from natural settings. This cross-examination between model and natural ecosystems will be a valuable and necessary approach for learning engineering principles and practices that are relevant to real-world systems (not laboratory artefacts), and for acquiring knowledge on scaling up laboratory-based engineering strategies to full-scale applications (Fig. 5). For example, microfluidicbased in vitro models of the human gut microbiome that contain co-cultures of human cells with different bacterial consortia are already producing physiological (including epithelial cell monolayer formation, cell growth and viability, cytokine levels and metabolomic profiles) and environmental (including oxygen gradients and laminar flow) variables that are comparable to in vivo variables148. The combination of model ecosystems with the DBTL cycle may be particularly fruitful for understanding the mechanisms governing microbial interactions and functional stability. Substantial knowledge is available on specific microorganisms that undergo co-aggregation and exchange metabolites, such as bacteria involved in nitrogen cycling2, consortia of methane-oxidizing archaea and sulfate-reducing bacteria128,149,150, and syntrophic bacteria partnered with hydrogenotrophic methanogens151,152. However, we are only beginning to understand the complex mechanisms (such as quorum sensing and secondary metabolites) involved in regulating the behaviour, interactions and kin discrimination of microorganisms in communities153. Although studies have established links between microbiome functional redundancy, diversity and stability154, a framework to predict or engineer functionally stable microbiomes has not been attained. Through the use of model laboratory ecosystems together with existing knowledge of microbial ecology and engineering design, it may be possible to decipher the chemical language of microbiomes and discover mechanisms of other important processes (including evolution, selection, dispersal limitation and neutral processes)155 that enable robust and stable microbiome function. Translating this theory into engineering design practice will require a quantitative framework that links these mechanisms to metabolic interaction networks and new approaches that allow ecological properties to emerge from metabolic models (Box 3).

Fig. 5 | Learning fundamental principles for microbiome engineering. Model laboratory ecosystems can be used for controlled experiments with simplified microbiomes and environmental properties, representing a system in between pure laboratory conditions (such as test tubes or flasks) and complex natural environments (such as soil or the ocean). Continuous cross-examination between laboratory-scale models and natural complex ecosystems will be needed for development of engineering principles and practices that are robust in real systems while also tractable in the laboratory. This will require close collaboration between multiple stakeholders, including researchers and end users (such as hospitals or treatment plants) that have expertise and experience with issues specific to each scale (part a). Key principles that need to be learned to make possible systematic microbiome engineering include microbial interaction mechanisms (part b), mechanisms governing functional stability and degeneracy (part c), and frameworks for quantitatively mapping and simulating ecological niches in complex ecosystems (part d). MS, mass spectrometry; SIP, stable isotope labelling.

Box 3 | Emerging principles for microbiome engineering: a case for niche modelling

ecological niche modelling could be used to systematically design higher-order properties such as functional stability and robustness into engineered microbiomes. However, to develop such a framework, mechanistic understanding of how diversity is maintained within microbiomes and how it imparts properties such as functional stability is needed. Here, we propose that this understanding could come from applying the design–build–test–learn cycle to answer key questions:

Does functional degeneracy lead to productivity and functional stability?

Diversity has been correlated with productivity and functional stability in communities of macroorganisms143,181, yet the role that diversity has in improving microbiome function and functional stability remains open. For microbiome engineering, we propose that diversity be viewed, discussed and defined through the lens of functional redundancy (as described previously154), or more specifically, functional degeneracy. This is the degree to which a set of organisms perform an identical role in ecosystem functionality (for example, methane oxidation, nitrogen fixation or polymer hydrolysis) but exhibit degeneracy with respect to other physiological traits (for example, pH optima or biofilm formation), which enables them to achieve realized niche space and coexistence51. The design–build–test–learn cycle offers an excellent opportunity to understand the molecular basis of functional degeneracy and to examine how emergent community-level properties, such as resilience to perturbation or susceptibility to invasion by another species, are predictable from quantifying the fundamental and realized niche space in microbiomes. We propose that ecological niche modelling could be a particularly useful framework to achieve this goal.

How is diversity maintained in microbial ecosystems?

To create a framework for ecological niche modelling, it will be important to understand how diversity is maintained. Competitive exclusion suggests that two species with identical resource requirements cannot coexist in the same ecological niche144. Therefore, we need to understand the mechanisms that create niche space and allow diversity to develop and be maintained. For example, what role do the processes of spatiotemporal variability, dormancy, predation, nutrient loading, secondary metabolite production and resistance, cell motility and biofilm formation have in niche differentiation? And how can these processes be manipulated to achieve and maintain a desired level of functional degeneracy in a microbiome? Answers to these questions will offer microbiome engineering mechanisms to design and control ecological niche space for desired microbiome properties.

How does ecological niche modelling underlie microbiome engineering?

To allow the systematic engineering of desirable higher-order microbiome properties, we propose that microbiome engineering develops a framework for ecological niche modelling. The goal of this framework would be to quantify community and individual fundamental niche and realized niche space by integrating multi-omic data, physiological information, nutrient availability and environmental parameters, and use them to develop strategies for controlling cooperation and competition in microbiomes. To achieve this goal, new mathematical representations of the fundamental niche and the realized niche of an organism or guild will need to be defined, together with fitness functions that describe responses to environmental variables. When incorporated into microbiome modelling, this framework will allow the ecological forecasting of higher-order properties, as well as quantification of cooperative and competitive microbiome landscapes. moreover, such frameworks will help guide important unresolved microbiome design questions, such as the trade-off between functional redundancy and minimal diversity.

Outlook

True advancement in microbiome engineering will need multiple DBTL rounds to capture the necessary ecological principles to manipulate microbiomes in a precise manner with predictable outcomes (Fig. 1). For example, incorporating direct interspecies electron transfer discovered during previous DBTL cycles into metabolic models and bioreactor construction (for example, by adding conductive materials) could optimize the efficiency of biogas production from waste27, or designing engineered *E. coli* to control levels of previously discovered autoinducers could tailor gut microbiota under conditions of dysbiosis towards a healthier state156. However, developing new knowledge and tools with fast turnaround will require next-generation infrastructure for data collection, data sharing and knowledge integration. To accelerate progress, developing the predictive capabilities needed for the learn phase is a priority. Model laboratory ecosystems combined with advances in automation, such as liquid-handling robots, microfluidics and data analysis pipelines157,158, will offer a starting point for the testing of multiple designs in a rigorous and reproducible manner. Capturing new knowledge from this process and integrating information into subsequent DBTL cycles will accelerate microbiome engineering developments, creating innovative biotechnologies and practices for the management of microbiomes across medicine, agriculture, manufacturing and environmental stewardship. Examples that show particular promise for advancing microbiome engineering across these fields include illuminating the roles that phages and metabolite cross-feeding have in controlling ruminal carbon turnover159, harnessing untapped anaerobic fungal–bacterial consortia to increase biomass conversion to valuable bioproducts160,161, creating microfluidic cell sorting techniques to automatically sort stable isotope-labelled cells from high-diversity samples for subsequent multi-omic analysis or cultivation162 and developing in situ metagenomic engineering tools to introduce new functions into microbiomes in their native environment111. To move the DBTL approach forward, interdisciplinary research teams with expertise in experimentation (for example, in culturing, molecular genetics or biochemistry), computation (for example, metabolic modelling, machine learning or bioinformatics), automation (for example, robotics or microfluidics), and practice (for example, professional engineers or medical doctors) are essential. The road ahead for microbiome engineering seems long, given our nascent understanding of microbial ecology; however, structuring research and technology developments around the DBTL cycle offers a promising approach for advancing microbiome engineering and providing innovative solutions for addressing pressing societal and environmental problems.

M I C R O B I O M E T R A C TA B I L I T Y A N D T R A N S L AT I O N  
R E V I E W S