

面向智能与数字时代的工程微生物资源库建设： 理念、路径与展望

李小鹏^{1,2,3}, 曹雪强^{1,2,3}, 冯雁⁴, 肖湘⁴, 李雷⁴, 张晓君^{3,4*}

- 1 上海交通大学, 转化医学研究院, 上海
- 2 转化医学国家重大科技基础设施, 上海
- 3 教育部工程微生物资源库, 上海
- 4 上海交通大学 生命科学与技术学院, 微生物代谢全国重点实验室, 上海

李小鹏, 曹雪强, 冯雁, 肖湘, 李雷, 张晓君. 面向智能与数字时代的工程微生物资源库建设: 理念、路径与展望[J]. 微生物学报, 2026, 66(4): 1473-1492.

LI Xiaopeng, CAO Xueqiang, FENG Yan, XIAO Xiang, LI Lei, ZHANG Xiaojun. Building of the Engineered Microorganisms and Genetic Tools Collection Center of China for the intelligent and digital era: concepts, approaches, and prospects[J]. *Acta Microbiologica Sinica*, 2026, 66(4): 1473-1492.

摘要: 工程微生物作为生物制造与新生物经济发展的核心驱动力量, 其战略价值日益凸显。新近建设的教育部工程微生物资源库力求突破传统“菌种保藏中心”的局限, 构建一个以信息技术为支撑, 集资源获取、智能创制、数据分析与开放共享功能于一体, 面向基础研究、技术开发、产业应用需求的综合性资源创新服务平台。本文详细阐述了资源库“三位一体”的建设框架: 以规模化资源保藏为基础, 智能化挖掘分析为核心, 数字化存算共享为纽带。通过开展特种与工业微生物及质粒的规模化获取与标准化保藏工作, 并整合高通量自动化、多组学、单细胞分析、人工智能(artificial intelligence, AI)预测模型及数字孪生等前沿技术, 使资源库实现“获取保藏-分析鉴定-数据共享”的全流程贯通。在借鉴美国典型培养物保藏中心(American Type Culture Collection, ATCC)等国际知名机构标准化与质量控制体系的基础上, 创建“活体-信息-智能”的资源库模式, 为我国合成生物学、代谢工程、生物医药、环境修复等领域的科技创新与产业升级提供强有力的资源保障与技术支撑。

关键词: 工程微生物资源库; 数据共享; 合成生物学; 人工智能; 数字孪生

资助项目: 微生物资源技术中心启动项目(WH510132401/013)

This work was supported by the Microbial Resource Technology Center Launches Project (WH510132401/013).

*Corresponding author. E-mail: xjzhang68@sjtu.edu.cn

Received: 2025-12-08; Accepted: 2026-02-06; Published online: 2026-02-13

Building of the Engineered Microorganisms and Genetic Tools Collection Center of China for the intelligent and digital era: concepts, approaches, and prospects

LI Xiaopeng^{1,2,3}, CAO Xueqiang^{1,2,3}, FENG Yan⁴, XIAO Xiang⁴, LI Lei⁴, ZHANG Xiaojun^{3,4*}

1 Institute of Translational Medicine, Shanghai Jiao Tong University, Shanghai, China

2 National Facility for Translational Medicine, Shanghai, China

3 Engineered Microorganisms and Genetic Tools Collection Center of China, Ministry of Education, Shanghai, China

4 State Key Laboratory of Microbial Metabolism, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, China

Abstract: Engineered microorganisms, as the core driving force of biomanufacturing and the development of the new bio-economy, demonstrate increasingly prominent strategic values. The newly built Engineered Microorganisms and Genetic Tools Collection Center of China (EMTCC) supported by the Ministry of Education aims to break through the limitations of “microbial culture collection centers” and provide a comprehensive resource innovation service platform supported by information technology and integrating resource acquisition, intelligent creation, data analysis, and open sharing functions. This platform intends to meet the needs of basic research, technology development, and industrial applications. This article elaborates on the building framework of EMTCC, which takes large-scale resource preservation as the basis, intelligent mining and analysis as the core, and digital storage and computing sharing as the link. By carrying out large-scale acquisition and standardized preservation of natural and engineered microorganisms and plasmids and integrating cutting-edge technologies such as high-throughput automation, multi-omics, single-cell analysis, artificial intelligence prediction models, and digital twins, the EMTCC achieves a complete process of “acquisition & preservation, analysis & identification, and data sharing”. On the basis of drawing on the standardization and quality control systems of internationally renowned institutions such as American Type Culture Collection (ATCC), a resource repository model integrating living “organisms, information, and intelligence” will be developed. This innovative practice is expected to provide strong resource guarantee and technical support for research innovation and industrial upgrading in the fields of synthetic biology, metabolic engineering, biomedicine, and environmental remediation.

Keywords: Engineered Microorganisms and Genetic Tools Collection Center of China; data sharing; synthetic biology; artificial intelligence; digital twin

微生物资源是国家战略性生物资源的重要组成部分，是支撑微生物学科发展与技术创新的基础条件，与国民食品、健康、环境安全及国家生物安全密切相关，是微生物科学研究、

产业及生物技术发展的核心物质基础^[1-3]。当前，全球生物技术竞争日趋激烈，工程微生物凭借其“活体工厂”特性，其资源库的战略价值相较于传统微生物资源库更具优势。以美国典型培

养物保藏中心(American Type Culture Collection, ATCC)为代表的国际机构, 正从单一保藏功能向集生物标准品、数据服务于一体的综合平台转型。我国在该领域面临核心种质资源储备不足、资源开发深度有限、数据共享机制不畅等突出问题。同时, 随着微生物研究系统性与复杂性的提升, 大规模组学数据与高通量培养、单细胞分析等新兴技术深度融合, 对信息化资源的获取、数据存储与分析、高通量计算建模、可视化及跨区域协同工作提出了更高要求^[4]。在此背景下, 2024年教育部批准建设工程微生物资源库(Engineered Microorganisms and Genetic Tools Collection Center of China, EMTCC), 旨在突破传统菌种保藏中心的局限^[5-6], 通过系统性平台规划与技术整合, 构建面向未来、支撑国家战略需求的新一代工程微生物资源基础设施。

工程微生物是通过基因编辑、合成生物学等技术定向改造以具备特定功能的微生物, 在生物基化学品合成、药物开发、环境修复等领域具有广阔前景^[7]。未来须突破以保藏为主的传统模式, 将人工智能(artificial intelligence, AI)、大数据分析、自动化高通量筛选、数字建模与仿真等新一代信息技术深度融合资源库建设与运营全过程, 形成“智能设计-数字建模-系统验证-应用部署”的闭环, 推动工程微生物向高度可编程、可调控的“细胞工厂”演进, 使其成为生物制造与合成生物学的核心载体。

然而, 当前我国微生物资源库仍面临多重挑战。首先, 现有资源库多以保藏为核心任务, 缺乏从保存、鉴定到功能挖掘与产业转化的全链条开发机制, 导致其在资源开发利用中的推动作用有限^[8-9]。其次, 资源库可持续发展能力不足, 运营资金主要依赖财政拨款, 技术与服务收入有限, 资金来源单一且不稳定, 影响长期运行与服务提升^[10]。再次, 我国微生物资源在空间分布与管理结构上呈现碎片化特征, 大量资源分散于不同单位, 因技术标准与管理规范存在差异, 削弱了整体协同效应^[10]。此外,

资源共享意识与实践水平与 ATCC、德国微生物和细胞培养物收藏中心(Deutsche Sammlung von Mikroorganismen und Zellkulturen, DSMZ)等国际机构相比仍有差距, 开放共享程度不足, 制约了资源的高效流通与广泛应用^[11-12]。最后, 信息管理系统建设滞后, 多数资源库尚未建立统一的数据标准与高效平台, 数字化与智能化水平低, 难以支撑资源的科学管理与智能应用, 削弱了其在数字化时代的竞争力与服务能力。

综上所述, 构建一个集数字化、智能化于一体的国家工程微生物战略资源中心, 已成为推动我国微生物资源高效利用与生物技术自主创新的关键路径。本文将围绕这一目标, 探讨教育部工程微生物资源库建设框架与实施路径。

1 建设理念与框架

在智能与数字技术迅猛发展的背景下, 工程微生物作为生物系统设计与合成生物学的重要载体, 其资源开发与管理正经历从“个体研究”向“系统构建”的深刻转变, 面临前所未有的机遇与挑战。在此背景下, 构建一个面向智能与数字时代的工程微生物资源库, 不仅是实现微生物资源高效利用与安全共享的基础性工程, 更是推动生物制造、绿色化工、医药健康、环境治理等关键领域可持续发展的战略性举措。

工程微生物资源库建设以“实体资源保藏-数字信息管理”双轮驱动、“资源自主创制-产业精准服务”协同发展的新型生态系统为核心理念。通过系统整合资源、推进协同创新、强化信息化管理, 致力于构建一个功能完善、可持续发展、开放共享的现代化微生物资源管理体系, 全面支撑我国微生物科学研究与产业发展的深度融合。教育部工程微生物资源库相较于 ATCC 等国际典型综合性资源库(表 1), 实现了从“基础保藏”向“智能赋能”、从“数据仓库”向“知识引擎”、从“资源提供”向“生态构建”的战略跃迁。其核心贡献体现在: 构建了专业的工程微生物结构化数据库; 首创了人工智能深度赋能的智

表1 国内外典型微生物资源库核心特征对比

Table 1 Comparison of core features among typical microbial resource repositories domestically and internationally

| Name | Resource type | Data integration level | Intelligence level | Service model |
|----------------|---|--|---|---|
| ATCC | Standard strains, cell lines | Basic metadata (taxonomy, source, culture conditions) | Informatized management, online ordering & logistics tracking | Standardized resource sales & professional technical services |
| DSMZ | Systematically classified and functional strains, genomic DNA | Integrated phenotypic, partial genomic, and literature data | Online query tools available, limited intelligent analysis | Resource ordering, custom preservation, identification & training |
| JCM | Extremophiles, strains with industrial/agricultural characteristics | Detailed culture conditions and partial genomic information | High level of informatization, few intelligent prediction tools | Resource supply & international collaboration |
| CGMCC/ CICC | Diverse application-oriented strains, large collection scale | Basic digitization, partially integrated omics data | Transitioning from information management to intelligent services | Combination of public welfare and market-oriented services (preservation, development) |
| EMTCC | Living entities+data+digital twins (strains, plasmids, parts, engineered microbes & their digital counterparts) | Full-spectrum structured integration (from sequence and phenotype to engineering history and performance parameters) | AI-driven (intelligent recommendation, function prediction, experimental simulation, low-code analysis) | Integrated “resource-data-solution” (full-chain customized services & cloud-based innovation ecosystem) |

ATCC: American Type Culture Collection; DSMZ: Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures; JCM: Japan Collection of Microorganisms; CGMCC: China General Microbiological Culture Collection Center; CICC: China Center of Industrial Culture Collection.

能检索、预测推荐与虚拟筛选平台；开发了基于知识图谱的主动知识发现引擎大模型；参与工程微生物领域特色数字化标准的制定；并初步形成了产学研用协同创新的智能生态体系。

这些关键技术突破不仅显著提升了资源库的服务效能与综合价值，更为我国工程微生物资源的深度开发、智能化利用与产业转化提供了关键支撑平台，成为我国在智能与数字时代参与全球生物资源竞争的核心优势所在。如图1所示，其整体架构基于三大核心能力构建。

1.1 资源获取与保藏能力

建立覆盖特种微生物、工业微生物及质粒资源的规模化、标准化采集与保藏体系，是重要的科技基础性工作^[13]，为项目奠定物质基础。工程微生物资源库旨在系统收集与保存包括细

菌、真菌、古菌在内的多类别微生物资源，目标是超越传统的单一微生物分类和功能识别模式，转向“系统整合与多维协同”的理念。该理念强调资源库不仅要涵盖微生物基因组、代谢通路、功能特性等生物学数据，还应集成工程改造策略、合成标准、调控元件及其特征以及数字建模结果等多维信息。通过系统化集成与标准化描述，资源库成为连接微生物“自然属性”与“工程属性”的桥梁，从而实现从“研究材料与研究工具”向“工程平台”的升级。

建立微生物资源的全程数字化追踪机制，构建具有数字身份标识和数字孪生的管理体系，形成“模块化”与“标准化”的建设目标：模块化的设计使资源库中的微生物元件(如启动子、编码序列、终止子、CRISPR系统等)具备可重复

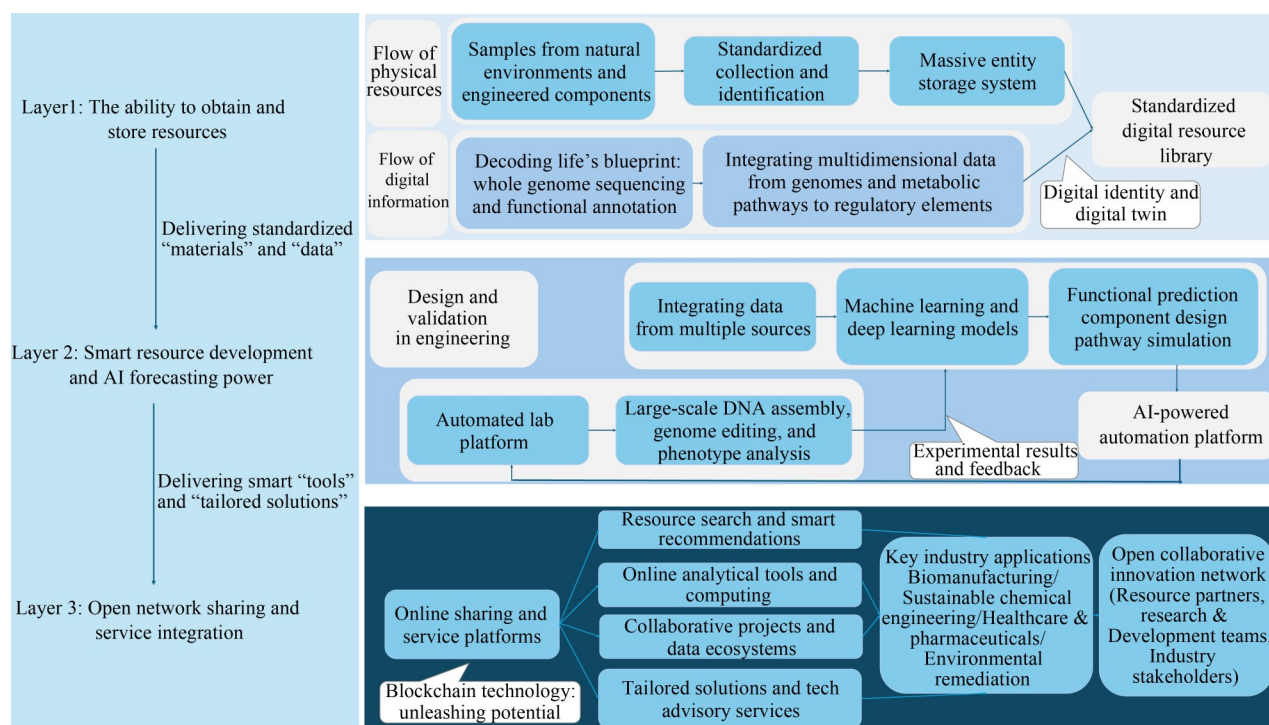


图1 “双轮驱动、协同发展”资源库核心架构关系图

Figure 1 The core framework of the “dual-engine, synergistic growth” resource bank.

使用、可组合利用的特性，为用户提供了可扩展且模块化的设计原则，使微生物能够合理工程化^[14]；标准化的描述体系(如 iGEM Registry 格式、SynBioDB 标准等)，保障资源的互操作性与高效共享，为微生物的理性设计与工程化提供坚实基础。

1.2 智能化资源开发与 AI 预测能力

集成自动化、高通量、多维度的分析模块，实现保藏资源的深度功能解析和高效工程改造。工程微生物资源库的建设秉持“智能驱动”理念，充分调用 AI 辅助设计、数字模拟、功能预测等工具，为合成基因组设计和构建提供了可用性与可预测性指导^[15]。同时，资源库构建开放的数据共享平台，鼓励跨学科、跨领域、跨机构的数据协作，形成“数据-模型-反馈”闭环，推动微生物工程设计的持续进化。

传统微生物功能开发主要依赖实验驱动策略，如基于经验的大规模培养筛选或随机基因

突变等方法，其过程周期长、成本高，且缺乏系统性指导，难以精准定位功能改造的关键靶点。随着 AI 技术的快速发展，微生物资源的评价与设计能力正经历方法学层面的深刻变革。AI 驱动的方法通过整合大规模基因组学^[16]、代谢通路、表型数据及功能注释信息，构建以深度学习和机器学习为核心的预测模型，能够有效识别微生物的关键功能属性^[17]。例如通过基因资源挖掘(如真核表达系统、次级代谢产物、极端酶等)对微生物的环境适应性、代谢产物合成潜力等关键功能进行高精度预测。这不仅加速了功能基因的挖掘，更能为合成生物学中的代谢通路优化与调控网络设计提供理性指导。

我们积极推动 AI 驱动的功能预测与理性设计方法能力建设，助力生物学研究范式由经验驱动向模型驱动转变。随着模型可解释性与精度的不断提升，AI 在跨物种功能预测、工业菌株定向进化等领域展现出广阔前景，显著提升

了资源开发的效率与成功率。

1.3 网络化开放共享与服务能力

随着信息技术的快速发展，资源的数字化与网络化共享已成为推动知识传播与科技创新的重要手段，为全球微生物学相关工作者提供信息服务和交流平台^[18]。工程微生物资源库依托自主开发的线上资源管理系统，构建了一个覆盖资源登记、检索、申请、共享及服务反馈的全流程数字化平台。通过深度融合先进大模型技术，平台实现了资源从静态存储向动态调用的转型，并在资源全生命周期管理中引入智能推荐与个性化服务功能，显著提升了资源配置的效率与精准性。

为确保共享过程的透明与可信，平台深度融合了区块链与智能合约技术。所有资源的访问、使用及数据流转记录均被不可篡改地存证于分布式账本中，建立了可追溯、防篡改的共享生态，有效防范了数据滥用与权益纠纷。

在服务模式上，平台面向科研机构、企业及公众提供多样化接入方式，并支持灵活的定制化服务。用户不仅可申请实体资源，还可按需获取数据分析、模型构建与深度解读等增值服务，从而充分释放资源潜能，形成“资源-数据-服务(解决方案)”一体化的服务闭环。

2 建设与运行的路径

工程微生物资源库的建设与运行，旨在构建一个覆盖资源全生命周期的一站式服务平台。其核心路径包括三大系统的协同建设：资源获取与保藏系统负责实体资源的收集、标准化保藏与实物库管理；资源挖掘与分析评价系统聚焦于资源的功能表征、性能提升与应用潜力评估；与此同时，资源数据存算与共享系统对实体资源衍生的多组学数据及资源本体信息进行整合存储、计算分析与深度挖掘，从而为资源的功能赋能提供数据驱动型指导，驱动微生物种质资源挖掘向通量化、智能化、数字化发展，

极大提高微生物种质资源的开发利用水平^[19]。该系统还构建统一的共享平台，通过集成化信息网站对外提供服务，使用户能够便捷查询资源信息并高效获取实物与数据资源。为进一步提升可用性，平台还将开发系列低代码分析工具，助力用户无需编程即可完成复杂数据分析，快速定位适用于下游应用的工程微生物资源并进行申领。整体上，本项目通过三级资源体系的整合，推动工程微生物资源的系统化管理、深度挖掘与高效共享，形成覆盖“资源-数据-服务”的一体化支撑体系。

2.1 资源获取与保藏系统：规模化与多样性的资源基础平台

近年来，随着组学技术的不断进步与人工智能辅助手段的广泛介入，微生物资源的获取、保藏与功能解析进入了规模化、多样化、系统化的发展阶段。本节重点介绍资源库的建设目标与技术路径，并探讨了如何从数据整合、标准化管理及智能化处理等方面提升资源获取与保藏平台的能力体系。

2.1.1 特种微生物资源库

我国及世界各国在极端环境微生物资源的获取与功能解析方面持续发力。极端环境(如深海、热泉、极地、高盐、高酸或高碱等)蕴藏着大量尚未被开发的微生物资源，其独特的生存机制和代谢能力为新型生物催化剂、降解酶、抗生素及生物修复材料的开发提供了丰富的物质基础。EMTCC 针对极端条件稀有微生物资源，借鉴环境宏基因组指导的定向分离策略^[20]，并采用高通量分离筛选等创新分离培养技术^[21]，大幅提高稀有微生物的获取效率，计划 3-5 年内分离鉴定 1 万-5 万株微生物。重点突破在于确保拟杆菌门(*Bacteroidota*)、假单胞菌门(*Pseudomonadota*)、芽孢杆菌门(*Bacillota*)和放线菌门(*Actinomycetota*)四大门类以外的新菌株占比超 50%，显著提升我国可培养微生物的多样性水平，为新型功能基因和代谢途径的发现奠定

资源基础。

2.1.2 工程微生物资源库

在生物制造、医药开发、能源转化与环境修复等战略性新兴产业中,工程微生物作为合成生物学与代谢工程的核心载体,正发挥着不可替代的关键作用。随着基因编辑技术(如CRISPR-Cas系统)、合成生物学工具链的快速发展,工程微生物的构建效率与功能多样性显著提升,对系统化、标准化、可追溯的工程菌株资源库的需求日益迫切。因此,构建集工程菌株收集、保存、表征与智能分析于一体的工程微生物资源库,已成为推动我国生物工程从“经验驱动”向“理性设计”跃迁的重要基础设施。

工程微生物资源库的构建,是我国生物技术前沿基础研究和产业转型升级的重要依托。微生物资源包括抗生素高产菌株、生物活性物质产生菌株、环境污染物降解菌株、工业发酵菌株、人体共生微生物菌株等。目标是建立覆盖1万-2万株具有明确工程背景的菌株储备体系,涵盖大肠杆菌、酵母、乳酸菌、放线菌等多种模式与非模式宿主。与传统工业菌株资源库不同,EMTCC不仅关注菌株的表型功能(如产物合成能力、环境适应性),更强调对每株工程菌的“工程改造图谱”进行系统记录与数字化保存,包括但不限于:基因编辑策略(如基因敲除、敲入、调控元件替换)、质粒构建信息(载体类型、启动子/终止子组合、抗性标记)、代谢通路重构路径、关键基因元件的序列与功能注释,以及相关的实验条件与表型数据。通过建立标准化的元数据规范,实现工程菌株从“黑箱操作”到“可追溯、可解析、可复用”的转变。在此基础上,资源库进一步构建“菌株工程案例库”,将工程菌株的构建过程、改造逻辑、功能表现与环境条件等多维数据结构化、标准化,形成可用于机器学习与人工智能分析的高质量训练数据集。该案例库不仅记录“是什么”(即最终菌株的性状),更详细记录“如何实现”(即设计路径、改造策略与迭代过程),为揭示生物工程

改造的共性规律、识别关键设计约束与优化节点提供了宝贵的数据基础。此类数据的积累对于揭示生物工程改造的规律性机制、优化菌株设计策略、提升生产效率和稳定性具有重要意义。通过该平台,研究人员可在机器学习辅助下,实现对工程菌构建路径的预测建模与优化推荐,从而从经验驱动的“试错法”转向模型驱动的理性设计^[22]。

2.1.3 基因元件库

目前,国内外在基因元件库的建设方面已取得显著进展,但系统化、标准化的基因元件资源库仍处于发展初期。ATCC作为全球最具影响力的微生物与细胞资源保藏机构之一,在基因元件资源的积累与共享方面具有重要地位。ATCC不仅提供大量标准化的野生型微生物菌株,还逐步拓展其资源范围,涵盖多种基因工程工具和功能元件。然而,ATCC的基因元件库仍以“功能元件”和“质粒载体”为主,缺乏对元件功能参数(如表达强度、动态范围、互作特性)的系统性量化与标准化表征,且多数元件未提供完整的“使用记录”或“工程背景”信息。EMTCC聚焦于工程基因元件的系统性积累与深度挖掘,目标是建设5万个集元件收集、功能注释、性能测试、可追溯性记录于一体的基因元件库。该库应涵盖启动子、RBS、终止子、调控蛋白基因、基因回路等核心元件,并通过高通量实验平台获取其在不同宿主、不同环境下的动态表达数据,形成可供机器学习分析的“基因元件功能数据库”,为工程菌株的理性设计提供底层支撑。该数据库可用于智能推荐、功能预测与合成模块的快速设计,为生物制造与合成生物学提供结构化、可检索的“基因元件”支持^[23]。

2.1.4 微生物资源的长期保藏与质量管理体系

为确保所获取资源的长期稳定性、遗传完整性与功能可靠性,EMTCC建立了系统化的长期保藏与质量管理体系。该体系涵盖多级保藏

策略、全程质量监控、智能环境管理与精准信息追溯，构成了资源库实体管理的核心保障。

(1) 实施分级保藏策略。资源根据使用频率与长期保存需求，分别采用短期冷冻(-20 °C)、中期冷冻(-80 °C)、长期液氮冷冻(-196 °C)及真空冷冻干燥等多级保藏方式，在确保存活率的同时，优化存储成本与存取效率。

(2) 建立保藏过程的质量监控体系。定期对保藏资源进行活力检测、纯度验证与功能稳定性评估，通过生长曲线测定、特定代谢功能复测及基因组稳定性分析，确保资源在长期保藏后仍保持其关键特性。

(3) 实现保藏环境的智能化监控。通过物联网传感器对低温冰箱、液氮罐等设备的温度、湿度、液氮液位进行实时监测与异常报警，并集成至智能化库存管理系统，实现保藏设备的预测性维护与资源库存的动态优化。

(4) 构建全流程可追溯的信息化管理。为每份资源分配唯一标识，并采用条形码或射频识别(radio frequency identification, RFID)技术，将其物理位置、保藏状态、检测记录与数字身份实时关联。

EMTCC 通过构建高度集成的自动化资源保藏系统，实现了全流程的数字化管理。该系统涵盖自动赋码打印、样本追踪、信息化管理、环境监控与安全控制等功能，并规范了入库、出库、移库、核销、盘点、预警、溯源以及数据备份与展示等关键操作流程。系统集成高通量自动化存储模块，配备菌种材料与样本的低温保藏设备及超低温冷冻相关设施，包括仓储式自动化存储系统和可容纳百万级样本的全自动生物样本库，全面支撑微生物资源的高效、安全与可追溯保存与管理。

2.2 资源挖掘与分析评价系统：智能化资源发掘与高效创制平台

随着高通量测序、单细胞分析、人工智能图像识别以及微流控芯片等前沿技术的融合发

展，微生物资源的分析效率、功能解析深度与资源开发的精准度均得到显著提升。构建高效、智能的微生物资源挖掘与分析评价系统，已成为推动微生物资源从“发现”向“利用”转化的关键技术支撑。

传统微生物鉴定与功能评价方法普遍存在通量低、效率差、信息单一等局限性，难以满足现代微生物组学研究对大规模、多维度数据的需求。为突破上述瓶颈，近年来通过整合动态监测、三维形态学成像、图像-质谱联用(image-MS)以及微流控芯片等核心技术，已构建起一套高通量、多维度、高度自动化的微生物分析平台。该平台实现了从菌株分离、表型表征到功能解析的全流程智能化处理。在此基础上，结合经典的地球物理与地球化学调查手段，研发出一系列适用于深海极端环境微生物探索的专用装备、技术流程与分析体系。以海沟沉积物微生物研究为例，通过建立从 DNA 提取到宏基因组测序的半自动化工作流程，已构建起能够支持大样本处理和高测序深度的宏基因组研究框架^[24-25]。在单细胞水平上，平台整合质谱成像与微流控液滴封装技术，实现了对关键代谢酶、转运蛋白及毒力因子等分子标志物的高灵敏度检测，能够在单细胞乃至亚细胞层面解析其功能特征，为功能基因的精准定位与调控机制研究提供重要支撑^[26]。

在此基础上，资源库计划开发基于人工智能图像识别的微生物分选系统(AI-driven microbial sorting system)。该系统依托深度卷积神经网络(convolutional neural network, CNN)与多模态图像分析技术，可实现对单个微生物细胞的自动识别、分类与分选，预计可获取超过100项细胞形态与生理参数，包括细胞形态、纹理特征、运动模式、代谢活性及群体行为等。该技术将广泛应用于新型抗生素开发、极端耐受菌筛选以及污染物降解菌的靶向分离等前沿领域。例如，通过 AI 分选系统对土壤样本中的微生物进行表型筛选与生理特征匹配，可高效

识别具有抗药性或降解特定污染物能力的新型功能菌株^[27]。此外,该系统可与单细胞测序技术无缝衔接,实现“形-序”联用(形态+基因组)的精准筛选策略,为工程菌株的定向设计与高效开发提供高精度起始点。

2.2.1 高通量、智能化资源筛选

传统微生物资源的筛选长期依赖人工操作,存在效率低下、通量有限且易受主观偏差制约等瓶颈。随着生物技术的进步,高通量筛选技术已成为资源挖掘的主流方向,而人工智能等智能化方法的融入,正推动筛选过程向更高效率、更高精度与自适应优化方向演进。

如图 2 所示,本平台通过整合微流控芯片、流式分选与自动化液体处理工作站,构建了涵盖培养与非培养方法的集成化高通量筛选体系。该体系不仅能实现大规模样本的快速并行处理,更关键的是引入了强化学习算法,使其能够根

据实时反馈数据动态优化筛选条件与策略,从而具备“自我进化”的智能调优能力。这一智能化筛选平台不仅显著提升了工程微生物资源的挖掘速度与命中率,也为药物先导化合物发现、环境功能微生物富集、工业菌株性能改良等应用方向提供了强有力的技术支撑。随着技术与数据的不断积累,智能筛选将在微生物资源开发中扮演愈加核心的角色,推动资源挖掘从经验驱动向数据驱动与模型驱动的范式转变。

2.2.2 AI 驱动的功能预测与理性设计

近年来,生成式 AI 方法在分子设计与发现领域展现出显著潜力。与传统依赖于已知分子库的筛选策略不同,生成式 AI 能够学习化学空间的潜在分布规律,直接生成具有目标性质的全新分子结构^[28]。该方法已在抗菌肽设计^[29]以及小分子药物的发现与优化^[30-31]等方面取得初步成果,证明了其在拓展可探索化学边界方面

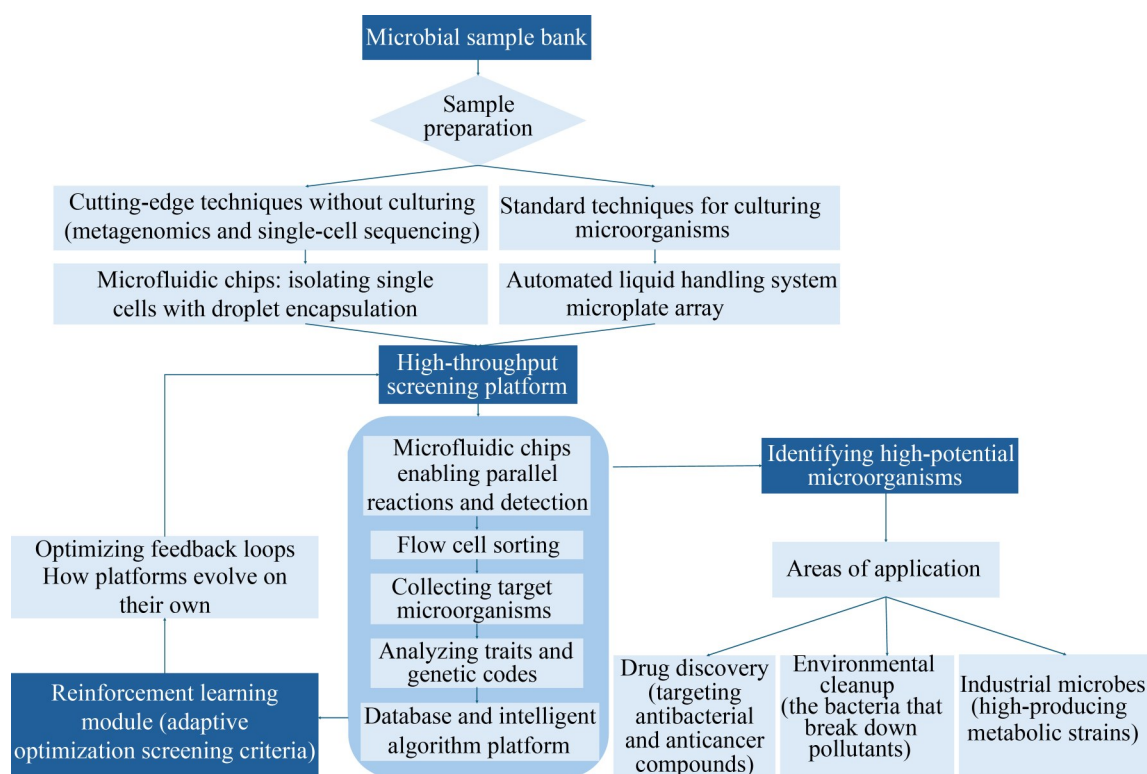


图2 微生物筛选模块技术路线图

Figure 2 Workflow diagram of the microbial screening process.

的应用前景。进一步地, 研究开始融合多种 AI 策略以提升探索效率与针对性。例如, 有工作结合图神经网络(graph neural network, GNN)与蒙特卡罗树搜索(Monte Carlo tree search, MCTS), 系统探索化学空间, 成功发现了在外膜渗透剂存在下对鲍曼不动杆菌具有抗菌活性的新型化合物^[32]。该案例表明, 通过耦合智能生成模型与定向搜索算法, 能够更有效地支撑抗菌药物的研发流程。

基于上述进展, 将生成式 AI 等智能方法应用于工程微生物资源的开发已成为重要方向。其核心在于构建覆盖多组学数据的专用数据库, 并利用图神经网络等模型解析基因型、表型与环境因素之间的复杂关联网络^[33]。通过人工智能、大数据、基因编辑及合成生物学技术的深度融合, 可大幅提升菌种筛选、功能预测与代谢途径优化的效率。更进一步, 可开发面向微生物功能元件与代谢通路的生成式 AI 模型, 实现基因回路和代谢网络的智能化设计与优化, 从而显著加速工程菌株的构建流程。

尽管 AI 模型在功能预测中展现强大潜力, 但其可靠性高度依赖数据质量与生物学约束。为此, 本库构建了三级验证体系: (1) 通过分层交叉验证与时序测试确保模型泛化性; (2) 对高价值预测委托合作实验室进行定向湿实验验证; (3) 开放社区反馈通道持续优化模型。同时, 采用 SHAP、知识图谱约束等技术提升可解释性, 破除“黑箱”困境。除此之外, 对稀有物种/特殊需求的预测仍依赖数据积累, 动态行为模拟需耦合机理模型。为此, EMTCC 引入四级可信度标签(★★★★至★), 引导用户理性决策。未来将通过迁移学习、Lab-automation 实时训练等策略推动模型自主进化。

如图 3 所示, 该智能设计体系可概括为多层架构: 数据层负责整合多源异构生物数据; 认知层通过 AI 模型挖掘数据中的规律与关联; 创造层利用生成模型进行元件与路径的从头设计; 应用层则最终输出可用于实验验证的菌株

设计方案。这一融合数据、认知与创造的框架, 标志着工程微生物研究正迈向可预测、可编程的智能化新阶段。

2.3 资源数据存算与信息平台: 数字化全生命周期管理与数字孪生平台

为支撑 AI 驱动设计及资源深度挖掘, 构建高效、智能的数据基础设施至关重要。资源数据存算与共享系统的核心, 在于建设一个覆盖数据全生命周期的数字化管理与数字孪生平台。该平台以多组学大数据为基石, 构建统一且可扩展的存储与计算资源模块, 形成服务于生物数字孪生模型的高性能数据仓。通过与校级高性能计算中心协同, 平台为海量生物数据提供专用、弹性的算力支撑, 实现了从数据采集、存储、处理到分析的全流程数字化管理。

为促进数据价值的释放, 平台同步研发了低代码、交互友好的一站式生物信息学分析环境, 显著降低了生物学家进行复杂多组学数据整合分析的技术门槛, 加速了从数据到知识的转化。更进一步, 平台深度集成智能 AI 助手, 能够为用户提供分析方案推荐、结果解读与科研假设生成等智能化服务, 从而直接支撑生物数字孪生模型的构建、模拟与迭代推演, 赋能生命科学的数字化探索与创新发现。

如图 4 所示, 该平台的总体架构包含 3 个层次: 数据资源层整合与管理多源异构生物数据; 平台与计算层提供核心的计算引擎、低代码分析工具及 AI 模型服务; 应用服务层则面向用户提供数据共享、智能分析及数字孪生应用等具体功能。这一体系化的建设, 为实现工程微生物资源的深度挖掘与智能化应用奠定了坚实的数据与计算基础。

2.3.1 工程微生物资源库数据标准体系

为确保数据平台与智能分析系统的高效运行与价值释放, 建立统一、规范的数据治理体系至关重要。本平台以“模块化”与“标准化”为核心理念, 构建了涵盖“基础层-工程层-应用层”的工程微生物资源库数据标准体系(表 2), 旨在

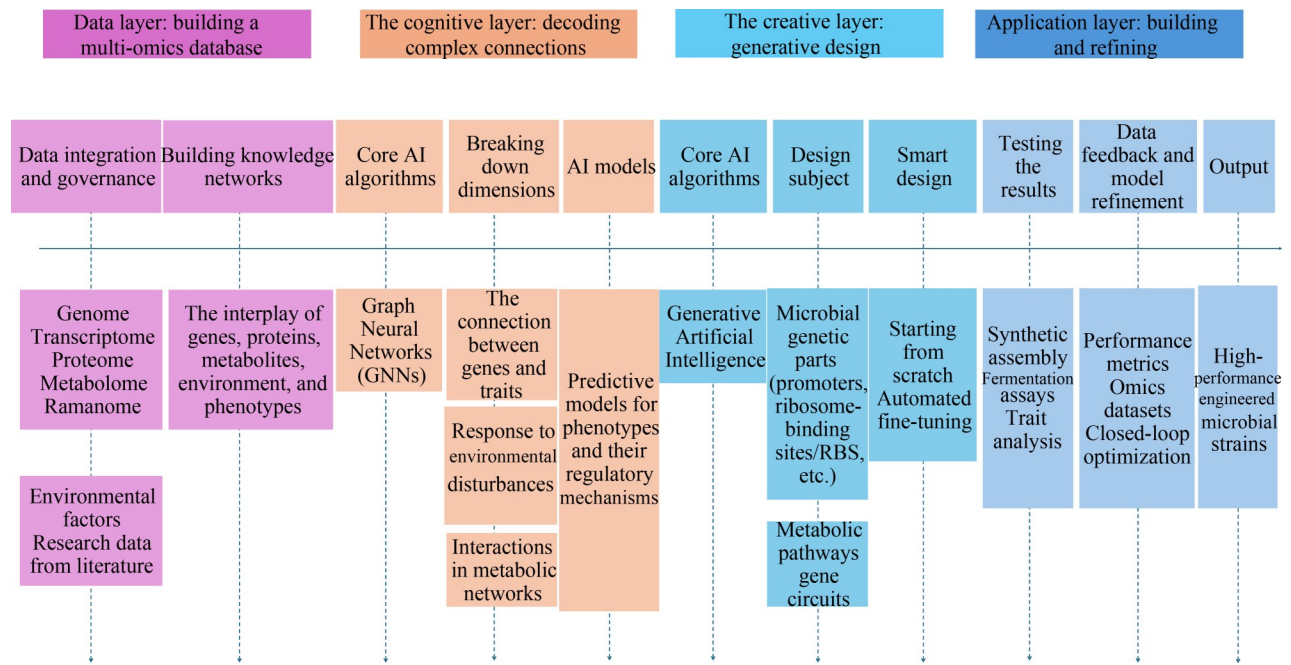


图3 工程微生物智能设计与创制路线

Figure 3 Smart design and engineering pathway for synthetic microbes.

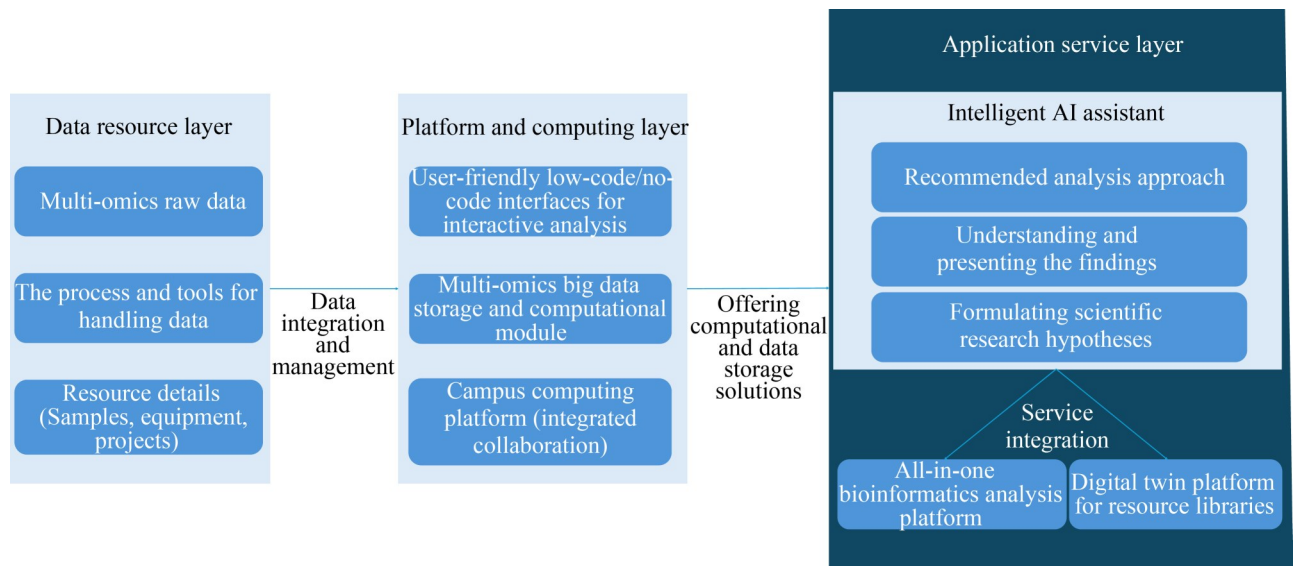


图4 数字平台逻辑架构图

Figure 4 The digital platform of logical architecture diagram.

从源头保障数据的规范性、可互操作性，并为后续的深度挖掘与智能化应用奠定基础。

基础层严格遵循国际通用的 MIxS 核心包，确保资源基本生物学属性的规范描述以及与

GenBank 等主流数据库的基础兼容。在此基础上，扩展定义了工程化状态、主要应用领域及知识产权状态等必填元数据，初步体现工程微生物资源的特色。

表2 具体数据标准与元数据规范

Table 2 Data standard system of the engineering microbial resource repository

| Level | Standard | Content | Relationship | Innovation |
|--------------------------------|--|--|--|---|
| Fundamental level | MiXS core package based | Outline the key biological characteristics of the resource, such as its classification, origin, growth conditions, and biosafety level, to maintain consistency and accuracy in the core dataset | Follow the MiXS guidelines set by the Genomic Standards Consortium (GSC) to maintain seamless compatibility with global databases like NCBI and European Bioinformatics Institute (EBI) | Establishing the groundwork for globally comparable and interoperable data |
| | Extended basic metadata | Building on the MiXS framework, the following required extension fields are defined: engineered status (indicating whether the sample is genetically modified), primary application domain (such as medical or industrial use), and intellectual property status (e.g., patented or open-source) | The MiXS framework allows for expansion while keeping core data parsing intact | Early-stage demonstration of the core characteristics of engineered microbial resources |
| Technical implementation level | Genetic engineering specification (GESpec) | Documenting genetic modifications systematically: track editing methods, target locations, added or deleted components, vector details, selection markers, and stability data | The component specifications align with SBOL standards, while the workflow descriptions follow the MISO framework | For the first time, researchers have systematically and comprehensively recorded the entire genetic modification history of engineered microbial strains, enabling traceability and advanced data-driven analysis |
| | Standardized parts registry description (SPRD) | Enhanced component specifications: requires using SBOL Visual standard diagrams; adds details like quantitative performance metrics and compatibility for interactions; includes links to functional validation data | The data model supports both SBOL v2 and v3 standards, while its performance metrics build upon and expand the SynBioDB evaluation framework | Offering clear, measurable, and visually intuitive descriptions of components to support data-driven design and virtual assembly |
| Practical application level | Performance characterization metadata (PCMeta) | Track key industrial performance metrics: product output/yield, how efficiently raw materials are used, resilience under harsh conditions, reaction rates, critical fermentation parameters, and real-world application test results | Taking cues from the minimum information about a marker gene sequence/minimum information about a metagenome sequence (MIMARKS/MIMS) approach to documenting environmental factors, we developed the first standardized metadata framework specifically designed to assess how well engineered microorganisms perform in real-world applications | The key to unlocking the potential of resources lies in converting raw experimental data into actionable, high-value insights that are both comparable and predictable. This shifts resources from mere “storage” to real-world “utilization” |

工程层是本标准体系的核心创新。针对工程微生物遗传背景复杂、功能元件多样化的特点,重点制定了2项关键规范:(1)《工程微生物遗传操作描述规范》实现了对菌株遗传改造历史的系统化、结构化记录,填补了该领域的标准空白,并与合成生物学开放语言(Synthetic Biology Open Language, SBOL)、合成生物体最低信息标准(Minimum Information about a Synthetic Organism, MISO)等国际规范兼容映射;(2)《标准化功能元件库描述规范》在兼容iGEM Registry与SBOL数据模型的基础上,强制采用SBOL Visual图示标准,并创新性引入量化性能参数元数据,显著提升了元件的可设计性与虚拟组装的可行性。

应用层首创《工程微生物性能表征元数据规范》,聚焦于资源产业化应用,系统化记录了包括产物产量、底物利用效率、环境耐受性及动力学参数在内的关键性能数据,实现了实验室数据向可比较、可预测的高价值信息的转化,为资源的精准筛选与产业对接提供直接依据。

在交互技术实现上(表3),本体系通过提供RESTful API(支持JSON-LD格式)、采用OAuth 2.0认证机制,并将核心字段与生物科学数据分析与管理本体(ontology of bioscientific data analysis and data management, EDAM)、生物医学调查本体(ontology for biomedical investigations, OBI)、生物

兴趣化学实体(chemical entities of biological interest, ChEBI)等国际生物本体进行语义映射,确保了与国内外主流科学数据平台间高效、语义化的互操作。

该体系的设计原则是主动兼容MIxS、SBOL等国际基石标准,同时在遗传操作深度描述(genetic engineering specification, GESpec)、功能元件量化定义(standardized parts registry description, SPRD)与应用性能标准化(performance characterization metadata, PCMeta)等关键环节实现特色化引领。这一机器可读、语义明确的标准体系,不仅为构建资源知识图谱、实现智能检索与AI预测提供了数据基石,也为打通产学研数据流、构建协同创新生态提供了核心保障。目前,相关特色规范正积极推动,以期贡献于我国工程微生物资源数字化标准的进步。

2.3.2 资源信息全流程数字化

为实现资源从入库到应用的全生命周期可追溯与智能化管理,建立资源信息的全流程数字化体系至关重要。该体系的核心是为每份入库的工程微生物资源分配具有唯一性的数字身份标识(unique digital identifier),并以此为基础,完整集成其来源信息、遗传操作历史、生理生化特性、多组学实验数据及实体存储位置等全维度属性。通过将二维码或射频识别(RFID)等物

表3 数据接口与交互

Table 3 Data interface and interaction

| Standard | Content | Relationship | Innovation |
|---------------------------------------|---|---|---|
| Data exchange interface specification | The system offers RESTful API-based data access, featuring support for JSON-LD formatting and OAuth 2.0 security protocols, while allowing flexible queries across all data tiers | By leveraging commonly adopted web services and data exchange standards, the system ensures seamless integration with third-party platforms | Enabling open data sharing and smooth interoperability across systems |
| Ontology mapping | Aligning core metadata fields with widely-used biological ontologies like EDAM, OBI, and ChEBI | Boosting data's machine-readability and logical inference capabilities using semantic mapping | Boosting data's machine-readability and unlocking its knowledge discovery potential in the Semantic Web |

联感知技术与该数字身份标识绑定, 可实现实体资源在自动化低温存储系统中的精确定位、快速检索与状态监控。这一数字化链路能够与自动化存取设备无缝对接, 从而在确保资源长期稳定保藏的同时, 大幅提升资源库运营中存取、盘点与分发环节的效率和准确性, 为后续的高通量筛选、数据挖掘与智能服务奠定了坚实的物联数据基础。

2.3.3 构建资源库数字孪生体

在实现资源全流程数字化的基础上, 本平台的最终目标是构建一个与实体资源库深度映射、虚实交互的数字孪生体。这一孪生系统不仅是实体库在虚拟空间中的动态镜像, 能够实时可视化展示资源库存、设备运行状态及资源流转轨迹, 更核心的创新在于其实现了以下智能化功能:

(1) 虚拟筛选与智能推荐: 用户可基于目标功能需求(如特定产物合成、底物利用或环境耐受性)进行查询。平台整合前述数据标准与 AI 模型, 对数字孪生体中的资源进行多维度比对与智能分析, 自动筛选并推荐最优候选菌株或元件, 极大缩短实验试错周期。

(2) 远程协同与模拟验证: 平台支持科研人员在线设计实验方案, 并通过接口远程调度自动化实验设备, 形成“云端实验室”操作范式。尤为关键的是, 数字孪生系统可在物理实验前, 基于积累的多组学数据与机理模型进行过程仿真与结果预测, 从而优化实验参数与方案设计, 提升研发效率与成功率。

(3) 全生命周期追溯与透明化管理: 依托于唯一的数字身份标识与物联网数据流, 该孪生体实现了对资源从采集、入库、表征、改造到应用的全链条数字化监控与可视化呈现。这不仅显著提升了资源管理的精细度与透明度, 也为资源的质量控制、使用权溯源与科学价值评估提供了完整的数据基础。

数字孪生平台通过深度融合物联网、人工智能与仿真技术, 将传统的静态资源库升级为

一个可预测、可交互、可优化的智能生态系统, 从而真正实现了工程微生物资源的数字化赋能与智能化运维。

2.3.4 建立安全高效的专业数据库

为确保数字孪生平台中数据流的可信、安全与合规, 本平台构建了集存证、保护与价值挖掘于一体的数据治理体系。具体而言, 基于区块链技术建立了数据全流程存证机制, 对数据的产生、流转与访问进行不可篡改的记录, 从而在技术层面保障数据的真实性、完整性与可追溯性。同时, 平台构建了多层次的数据安全管理体系, 通过权限分级、访问审计与数据脱敏等手段, 在严格保护资源知识产权与个人隐私的前提下, 实现数据的受控、合规共享。

为解决数据“可用不可见”的共享难题, 平台积极探索并引入联邦学习等隐私计算技术。该技术使得多方能够在原始数据无需离开本地存储的情况下, 共同参与模型训练与联合分析。这不仅有效避免了数据汇聚带来的泄露风险, 更在保障数据主权和安全的同时, 充分释放了分散数据的潜在价值, 为跨机构、跨领域的协同创新提供了可行的技术路径。

通过融合区块链存证、动态安全管控与隐私计算, 本平台旨在建立一个既安全可靠、又开放协同的专业数据库环境, 为工程微生物资源的可信流通与深度智能化应用奠定坚实的信任基础。

3 共享与可持续发展

3.1 创新 AI 赋能的个性化服务与共享机制

在构建安全可信的数据基础设施之上, 本平台进一步创新性地打造了 AI 赋能的个性化服务与智慧共享机制。如图 5 所示, 该机制以资源的标准化、信息化管理为基础, 依托分级共享策略与用户需求智能分析, 实现服务模式的深度变革。AI 系统能够动态解析用户的具体研

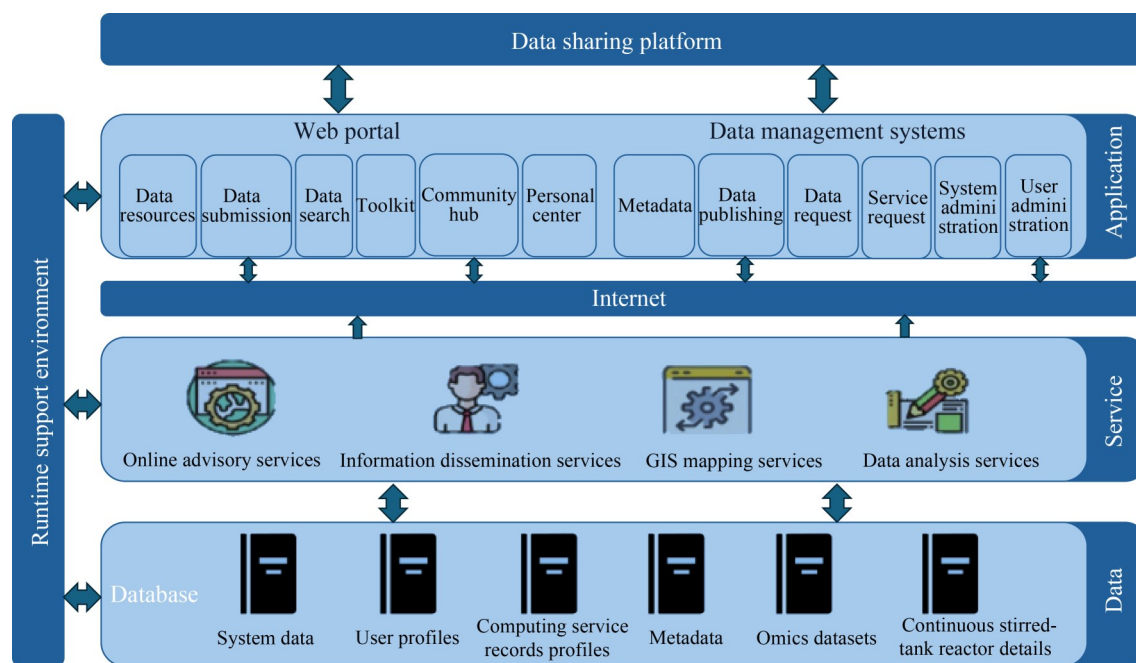


图5 微生物数据共享机制

Figure 5 Microbial data sharing mechanism.

发目标与应用场景，主动推荐适配的菌株资源、数据工具及分析流程，从而提供高度定制化的知识服务与解决方案。这一机制不仅为我国及相关友好国家的科研与产业机构提供了坚实的资源保障与技术服务，更通过精准匹配资源所有者、使用者与平台方的价值诉求，构建了良性互动的协同创新生态，显著提升了资源流转效率与整体创新活力。

3.1.1 提供“资源-数据-服务”一体化服务

为将平台的数据与智能能力有效转化为产业生产力，本平台致力于推动服务模式从传统资源供应向“资源-数据-服务”三位一体的综合服务体系转型，以精准应对中小企业研发资源有限、技术门槛高的核心挑战。

具体而言，平台面向企业用户提供“菜单式”标准服务与“定制化”深度服务。企业不仅可获得严格质控的工程微生物菌株，还能获得由平台 AI 系统生成的多维度功能评价报告，该报告整合了菌株的生长特性、代谢功能、工艺适配性与安全评估等关键数据，为产业化决策提

供科学依据。针对特定生产场景(如农业、工业生物制造、环境修复等)，平台可提供定制化技术解决方案及后续菌株优化支持，扮演企业可靠的“外部微生物研发部”，显著降低其研发成本与风险。此外，平台将部署 AI 咨询机器人，提供 7×24 h 在线服务，能够即时响应咨询并根据企业需求进行初步的菌株筛选与方案推演，大幅缩短技术对接路径。如图 6 所示，通过资源、数据与解决方案的深度耦合，本平台构建了覆盖“需求洞察-智能匹配-方案交付-持续优化”全链条的产业服务生态，致力于成为驱动微生物技术成果高效转化与应用的关键支撑平台。

3.1.2 打造智慧共享平台与创新生态

为推动工程微生物资源的汇聚、整合与高效利用，教育部工程微生物资源库构建了系统的资源整合与智慧共享机制，打造了开放协同的创新生态，建立了高效、规范的资源汇交整合机制。平台通过统一的线上数据录入系统简化资源提交流程，并严格遵循国际通用数据标准确保信息规范性。为激励广泛参与，设立了

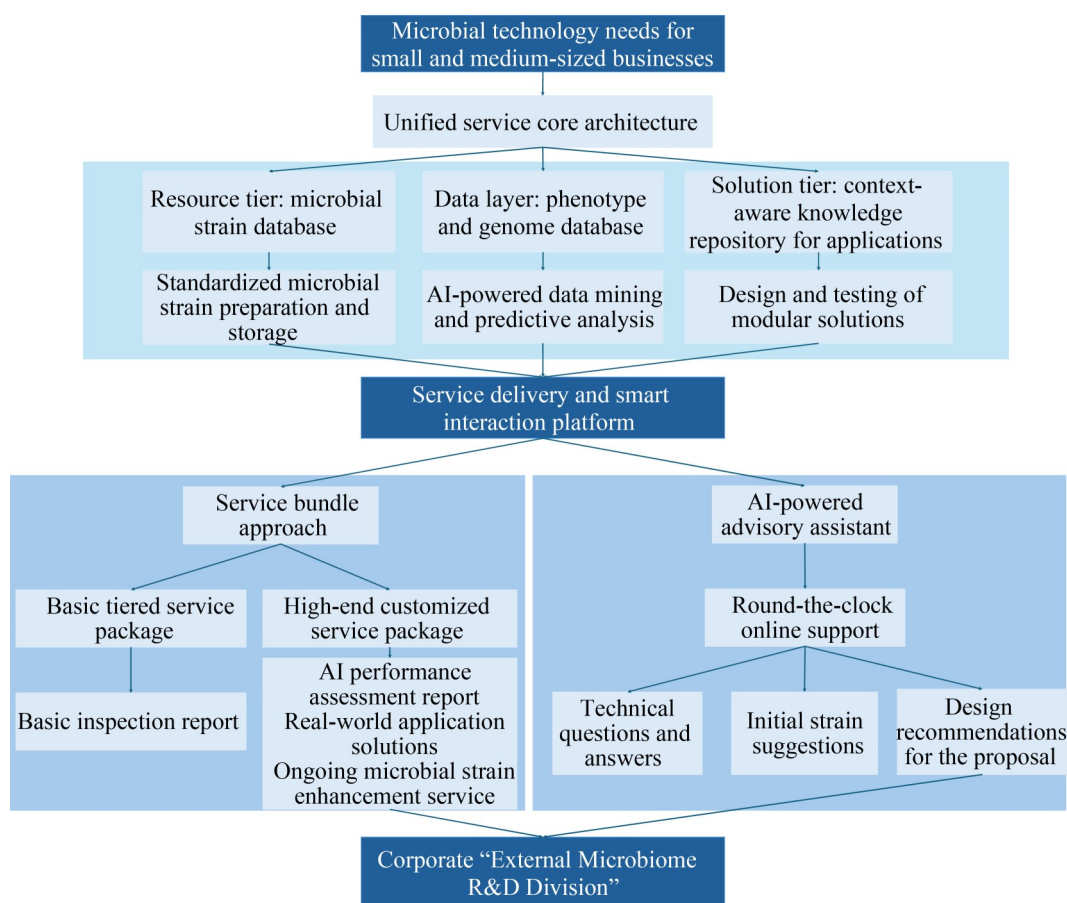


图6 面向企业的一体化服务机制

Figure 6 Integrated service mechanism for enterprises.

以“保藏号”与“资源优先使用权”为核心的激励机制，吸引高校、科研院所及企业主动汇交优质资源。所有入库资源均须经过严格的复核鉴定与质量控制流程，从源头保障了资源库的权威性与可靠性。

除此之外，打造了功能集成的线上智慧共享社区。该社区融合了多项智能化技术：用户可通过智能搜索引擎精准定位所需菌株、数据或技术方案；引入基于区块链的智能合约，自动化执行资源使用授权、费用结算与收益分配，保障共享过程的透明、可信与公平；通过举办线上研讨会、开源数据集及算法模型，构建了活跃的交流与协作生态。此外，社区部署了 AI 驱动的“创新匹配”系统，能够自动识别并推荐

潜在的合作机会与技术互补方，促进协同创新。

最后，构建了网络化、可信任的服务能力。在信息化时代背景下，平台依托上述技术架构，实现了资源的跨地域、跨机构安全高效共享。区块链与智能合约技术不仅解决了传统共享中的信任与权益分配难题，还支持按需付费、自动授权等灵活服务模式，为用户提供了兼具安全性、可信度与个性化的服务体验。这一智慧共享平台，正逐步成为驱动工程微生物领域开放创新与合作的核心枢纽。

3.2 构建多方协同的“资源-技术-产业”融合体

深化与全国重点实验室、国家技术创新中心及行业领军企业的战略合作，建立国家微生物

物资源共享平台及科学研究联盟,整合现有菌种资源库,推动菌种采集、保存和功能挖掘的系统化、标准化发展^[33]。打造“资源库-实验室-产业界”三位一体的协同创新机制。可通过联合申报国家重大科技项目、协同攻克产业瓶颈技术等途径,加快实验室工程菌株的中试转化与产业化进程,构建“产业需求驱动资源开发、优质资源反哺产业升级”的正向循环机制。依托资源库的数字孪生平台和 AI 预测模型,可在虚拟环境中预先完成工艺优化和放大模拟,显著降低中试阶段的试错成本与失败风险。

3.3 可持续运营模式与保障机制

针对生物资源库普遍面临的“可持续发展能力不足”的瓶颈,工程微生物资源库探索并构建了一套以公益为基石、多元投入为支撑、创新机制为驱动的可持续运营与保障体系,旨在确保其公共平台职能的同时,实现长期稳定发展。

3.3.1 多元化的资金筹措与市场化服务机制

我们旨在建立多层次、互补性的资金流入渠道。首先,积极争取并稳定获得国家及地方层面的科技基础条件平台专项经费,用于保障基础设施、核心团队及基础公益服务的可持续运行;其次,对平台自主研发形成的专利技术、专用菌株、软件工具等具有明确市场价值的成果,以技术许可形式进行市场化转化;再次,对于定制化菌株构建、深度组学分析、中试工艺开发等高附加值技术服务,实施基于全成本核算的透明收费;最后,通过与企业共建联合实验室或孵化项目,吸引社会资本投入,共享未来收益,深化产学研融合。

3.3.2 制度化的知识产权保护与利益分配框架

为保障协同创新的活力与公平,平台建立了前瞻性的知识产权管理体系。预先明确了基于平台资源产生成果的权属界定与利益分配原则,并创新性引入区块链技术对研发全过程进行可信存证,为成果确权与贡献量化提供客观

依据,有效降低合作各方的风险与顾虑。此外,平台利用 AI 技术对全球专利与文献进行动态监控,为用户提供知识产权预警与布局建议。

通过融合多元化的财务支撑与制度化的创新保障,该体系不仅为资源库的日常运营与长期发展奠定了坚实的经济基础,更通过构建清晰、可信的利益分享与保护机制,从根本上激发了参与各方的创新动力,构成了资源库实现良性循环与可持续发展的核心引擎。

3.4 生物安全与伦理治理体系

为确保微生物资源的合规、安全与负责任使用,工程微生物资源库构建了系统化的生物安全与伦理治理体系,严格遵循“预防为主、分级管控、全程追溯”原则,形成了覆盖技术监管、伦理审查、公众信任及风险防控与应急预案的四级治理框架。

在技术监管层面,平台建立了基于全链条数据流的安全管控机制。通过“物理菌株-RFID 标签-生物信息 ID-实验日志-访问记录-区块链存证”的闭环链路,实现资源的动态追踪与实体受控。同时,系统实施数据动态脱敏、访问权限分级与安全风险实时监测,全面满足我国《基因编辑生物安全评价指南》等法规的强制要求。在伦理审查层面,设立由多学科专家组成的伦理委员会,并引入一定比例的独立外部委员。该委员会对入库资源与技术方案的伦理评估,对基因驱动等争议性技术实施明确禁止与一票否决,从源头防控伦理风险。在公众信任层面,平台通过实验室数据可视化系统与定期发布开源伦理报告,主动推进科研过程的透明化,积极接受社会监督,并探索公众参与机制,以构建负责任创新的社会基础。在应急防控层面,体系建立了完备的风险预案与响应机制。依据风险评估对工程菌株实施分级管控,并制定详细的泄漏与意外释放应急预案,涵盖立即隔离、源头灭活、环境监测与事故追溯等环节,通过定期演练确保应急处置能力,筑牢生物安全的最后一道防线。

该体系旨在实现从预防性规范、过程性审查、社会性共建到应急性保障的全周期治理,为资源库的可持续发展提供坚实的合规与安全基础。

4 总结与展望

教育部工程微生物资源库的建设不仅是我国生物制造战略布局的重要组成部分,也是推动科技创新与产业转型升级的关键举措。随着生命科学、合成生物学和人工智能等技术的迅猛发展,微生物资源的研究与应用已迈入精准化、智能化的新阶段。在此背景下,建立一个集资源收集、功能预测、设计合成、产业转化于一体的工程微生物资源库,具有重要的战略意义和现实价值。

首先,工程微生物资源库肩负着资源保藏与功能挖掘的双重使命。通过高通量基因组测序、多组学分析^[34]、人工智能建模和数字孪生技术,资源库不仅能够全面表征和分类微生物资源,还具备功能预测与合成设计的能力。这不仅提高了微生物资源的利用效率,还为工业发酵、药物开发、环境治理等下游应用提供了强有力的支持。例如通过量身定制微生物“军团”,让农田废弃物变废为宝,实现温室气体减排与土壤改良的双赢^[35]。通过合成生物学手段对需钠弧菌(*Vibrio natriegens*)进行改造:独创“Vmax 迭代自然转化法”,成功将总长达 43 kb 的 5 组基因武器库植入菌体;经测试,这款超级工程菌能同时降解联苯、苯酚、萘等 5 种典型污染物,从简单的单环化合物到复杂的多环结构都不在话下,在氯碱厂和炼油厂的真实废水样本中展现出强大修复潜力^[36]。

其次,数字化与智能化的深度融合将推动该资源库向“下一代生命科学基础设施”方向发展。通过构建标准化数据库^[37]、开放共享平台和智能分析系统,资源库能够实现全球分散生物资源的整合与协同研究。此外,AI 算法的不断优化将增强资源库的智能筛选、功能推演

和合成路线设计能力,从而加速从发现到应用的全流程闭环。在产业赋能方面,工程微生物资源库将积极对接绿色生物制造与新兴产业的发展需求,围绕生物基材料、生物燃料、生物制药和农业微生物等应用方向,开展定向资源开发与技术转化。通过持续的技术创新与制度完善,工程微生物资源库必将成为智能时代生物工程领域的核心基础设施和创新引擎,为可持续发展注入强大的科技动能。

致谢

本文撰写过程中参考了上海交通大学工程微生物资源库项目申报材料。谨向该项目书的撰写团队成员致以诚挚的谢意,特别是许杰、王景、周水荣、于晴等。同时,感谢张显明、孙洁林、李冬凉、朱顺英等众多老师在资源库项目规划与实施过程中给予的指导与支持。

作者贡献声明

李小鹏:论文构思、文献调查、论文写作;曹雪强:文献调查;冯雁:论文指导;肖湘:资料支持;李雷:论文指导;张晓君:论文构思、论文修改。

作者利益冲突公开声明

作者声明不存在任何可能会影响本文所报告工作的已知经济利益或个人关系。

参考文献

- [1] 姜瑞波,阮志勇. 推进国家微生物资源平台建设,支撑微生物学科发展:“第七届全国微生物资源学术暨国际微生物系统与分类学研讨会”专刊序言[J]. 微生物学报, 2016, 43(5): 891-892.
Jiang RB, Ruan ZY. Microbial resources are the basis for the development of Microbiology: the 7th National Conference of Microbial Resources & the International Symposium on Microbial Systematics and Taxonomy[J]. *Microbiology China*, 2016, 43(5): 891-892 (in Chinese).
- [2] 吴林寰,刘柳,孙清岚,刘斌,于建荣,徐萍,陈方,陈云伟,李祯祺,施慧琳,周宇光,马俊才. 中国微生物资源研究现状及未来发展态势分析[J]. 微生物学报, 2018, 58(12): 2123-2133.
Wu LH, Liu L, Sun QL, Liu B, Yu JR, Xu P, Chen F, Chen YW, Li ZQ, Shi HL, Zhou YG, Ma JC. The status

- and future aspects of the researches of microbial resources in China[J]. *Acta Microbiologica Sinica*, 2018, 58(12): 2123-2133 (in Chinese).
- [3] 汪建峰, 蒙海林, 熊智强, 王勇. 基于合成生物学的微生物制造研究进展[J]. *生物产业技术*, 2013(4): 12-23.
- [4] Kyrpides NC, Eloe-Fadrosh EA, Ivanova NN. Microbiome data science: understanding our microbial planet[J]. *Trends in Microbiology*, 2016, 24(6): 425-427.
- [5] 张甜. 微生物菌种保藏方法及标准菌种管理[J]. *中国城乡企业卫生*, 2014, 29(1): 139-141.
- [6] 国家质量监督检验检疫总局. SN/T 2632—2010 微生物菌种常规保藏技术规程[S]. 北京: 中国标准出版社, 2010.
- Inspection and Quarantine of the People's Republic of China. SN/T 2632—2010 Technical codes for routine microbial culture collection[S]. Beijing: Standards Press of China, 2010.
- [7] 周志华, 赵宗保, 牛国清. 新功能生物元器件的筛选、设计与功能鉴定[J]. *生物产业技术*, 2015(6): 49-58.
- [8] 全拓, 彭楠, 彭东海, 孙明, 郑金水. 我国农业微生物资源库建设与运行的几点思考[J]. *微生物学报*, 2025, 65(4): 1684-1694.
- Quan T, Peng N, Peng DH, Sun M, Zheng JS. Suggestions on the construction and operation of agricultural microbial resource repositories in China[J]. *Acta Microbiologica Sinica*, 2025, 65(4): 1684-1694 (in Chinese).
- [9] 郭静利, 尼鲁帕尔·迪力夏提, 王大庆. 我国农业微生物产业发展的对策建议[J]. *中国农业科技导报*, 2023, 25(11): 1-7.
- Guo JL, Dilxat Nilufar, Wang DQ. Development strategy on agricultural microbial industry in China[J]. *Journal of Agricultural Science and Technology*, 2023, 25(11): 1-7 (in Chinese).
- [10] 周杨, 邓名荣, 杜娟, 宋仲骞, 吴清平, 朱红惠. 我国农业微生物产业发展研究[J]. *中国工程科学*, 2022, 24(5): 197-206.
- Zhou Y, Deng MR, Du J, Song ZJ, Wu QP, Zhu HH. Development of agricultural microbial industry in China[J]. *Strategic Study of CAE*, 2022, 24(5): 197-206 (in Chinese).
- [11] American Type Culture Collection (ATCC). Who we are [EB/OL]. [2025-03-02]. <https://www.atcc.org/about-us/who-we-are>.
- [12] German Collection of Microorganisms and Cell Cultures GmbH (DSMZ). Bioresources[EB/OL]. [2025-03-02]. <https://www.dsmz.de>.
- [13] 宁国赞. 农业微生物资源在国家经济及生态环境建设中的重要作用[J]. *中国农业科技导报*, 2002, 4(3): 23-26.
- Ning GZ. The important role of agro-microbiological resources in national economy and ecological environment improvement[J]. *Review of China Agricultural Science and Technology*, 2002, 4(3): 23-26 (in Chinese).
- [14] Jeong KJ. Advancing microbial engineering through synthetic biology[J]. *Journal of Microbiology*, 2025, 63(3): e2503100.
- [15] James JS, Dai JB, Chew WL, Cai YZ. The design and engineering of synthetic genomes[J]. *Nature Reviews Genetics*, 2025, 26(5): 298-319.
- [16] Overmann J, Abt B, Sikorski J. Present and future of culturing bacteria[J]. *Annual Review of Microbiology*, 2017, 71: 711-730.
- [17] Ürel H, Benassou S, Marti H, Reska T, Sauerborn E, Pinheiro Alves De Souza Y, Perlas A, Rayo E, Biggel M, Kesselheim S, Borel N, Martin EJ, Venegas CB, Schloter M, Schröder K, Mittelstrass J, Prospero S, Ferguson JM, Urban L. Nanopore- and AI-empowered microbial viability inference[J]. *GigaScience*, 2025, 14: giaf100.
- [18] 范国梅, 孙清岚, 张荐轅, 吴林寰, 马俊才. 国家微生物科学数据中心: 国家微生物资源大数据体系建设[J]. *生命科学*, 2023, 35(12): 1599-1607.
- Fan GM, Sun QL, Zhang JY, Wu LH, Ma JC. National Microbiology Science Data Center: construction of national big data system for microbial resources[J]. *Chinese Bulletin of Life Sciences*, 2023, 35(12): 1599-1607 (in Chinese).
- [19] 谷医林, 张晓霞, 魏海雷, 阮志勇, 顾金刚, 李世贵, 罗会颖, 姚斌. 我国农业微生物种质资源保护利用发展思考[J]. *中国工程科学*, 2025, 27(5): 180-189.
- Gu YL, Zhang XX, Wei HL, Ruan ZY, Gu JG, Li SG, Luo HY, Yao B. Protection and utilization of agricultural microbial germplasm resources in China[J]. *Strategic Study of Chinese Academy of Engineering*, 2025, 27(5): 180-189 (in Chinese).
- [20] Yang KW, Zhang L, Ruiz-Valencia A, Song X, Vogel TM, Zhang XJ. Heterogeneity in the composition and catabolism of indigenous microbiomes in subsurface soils cocontaminated with BTEX and chlorinated aliphatic hydrocarbons[J]. *Environmental Science & Technology*, 2025, 59(9): 4540-4550.
- [21] Lewis WH, Tahon G, Geesink P, Sousa DZ, Ettema TJG. Innovations to culturing the uncultured microbial majority[J]. *Nature Reviews Microbiology*, 2021, 19(4): 225-240.
- [22] Saha R, Chauhan A, Rastogi Verma S. Machine learning: an advancement in biochemical engineering[J]. *Biotechnology Letters*, 2024, 46(4): 497-519.
- [23] Slusarczyk AL, Lin A, Weiss R. Foundations for the design and implementation of synthetic genetic circuits[J]. *Nature Reviews Genetics*, 2012, 13(6): 406-420.
- [24] Marchesi JR, Ravel J. The vocabulary of microbiome research: a proposal[J]. *Microbiome*, 2015, 3: 31.
- [25] Roy G, Prifti E, Belda E, Zucker JD. Deep learning methods in metagenomics: a review[J]. *Microbial Genomics*, 2024, 10(4): 001231.
- [26] Xiao X, Wang J, Ding K. MEER: extraordinary flourishing ecosystem in the deepest ocean[J]. *Cell*, 2025, 188(5): 1175-1177.

- [27] Wan FP, Torres MDT, Peng J, de la Fuente-Nunez C. Deep-learning-enabled antibiotic discovery through molecular de-extinction[J]. *Nature Biomedical Engineering*, 2024, 8(7): 854-871.
- [28] Krishnan A, Anahtar MN, Valeri JA, Jin WG, Donghia NM, Sieben L, Luttens A, Zhang Y, Modaresi SM, Hennes A, Fromer J, Bandyopadhyay P, Chen JC, Rehman D, Desai R, Edwards P, Lach RS, Aschtgen MS, Gaborieau M, Gaetani M, et al. A generative deep learning approach to *de novo* antibiotic design[J]. *Cell*, 2025, 188(21): 5962-5979.e22.
- [29] Gómez-Bombarelli R, Wei JN, Duvenaud D, Hernández-Lobato JM, Sánchez-Lengeling B, Sheberla D, Aguilera-Iparraguirre J, Hirzel TD, Adams RP, Aspuru-Guzik A. Automatic chemical design using a data-driven continuous representation of molecules[J]. *ACS Central Science*, 2018, 4(2): 268-276.
- [30] Zhavoronkov A, Ivanenkov YA, Aliper A, Veselov MS, Aladinskiy VA, Aladinskaya AV, Terentiev VA, Polykovskiy DA, Kuznetsov MD, Asadulaev A, Volkov Y, Zholus A, Shayakhmetov RR, Zhebrak A, Minaeva LI, Zagribelnyy BA, Lee LH, Soll R, Madge D, Xing L, et al. Deep learning enables rapid identification of potent DDR1 kinase inhibitors[J]. *Nature Biotechnology*, 2019, 37(9): 1038-1040.
- [31] Swanson K, Liu G, Catacutan DB, Arnold A, Zou J, Stokes JM. Generative AI for designing and validating easily synthesizable and structurally novel antibiotics[J]. *Nature Machine Intelligence*, 2024, 6(3): 338-353.
- [32] De Souza Neto LR, Moreira-Filho JT, Neves B, Maidana RLBR, Guimaraes ACR, Furnham N, Andrade CH, Silva FP Jr. *In silico* strategies to support fragment-to-lead optimization in drug discovery[J]. *Frontiers in Chemistry*, 2020, 8: 93.
- [33] 王薇. 陈卫代表: 构建微生物资源共享机制 推动多学科交叉融合[N]. *中国食品报*, 2025-03-12(3).
- [34] Routhier E, Mozziconacci J. Genomics enters the deep learning era[J]. *PeerJ*, 2022, 10: e13613.
- [35] Wang XH, Xiang BY, Li J, Zhang MH, Frostegard A, Bakken L, Zhang XJ. Using adaptive and aggressive N₂O-reducing bacteria to augment digestate fertilizer for mitigating N₂O emissions from agricultural soils[J]. *Science of the Total Environment*, 2023, 903: 166284.
- [36] Guo LH, Ouyang XY, Wang WW, Qiu XY, Zhao YL, Xu P, Tang HZ. Fine-tuning an aromatic ring-hydroxylating oxygenase to degrade high molecular weight polycyclic aromatic hydrocarbon[J]. *Journal of Biological Chemistry*, 2024, 300(6): 107343.
- [37] 全国科技平台标准化技术委员会(SAC/TC 486). 微生物资源机构数据管理及发布规范: GB/T 46408-2025[S]. 北京: 中国标准出版社, 2025. National Technical Committee 486 on Science and Technology Infrastructure of Standardization Administration of China. Specification for Data Management and Publication of Microbiological Resource Centre: GB/T 46408-2025[S]. Beijing: Standards Press of China, 2025 (in Chinese).