

# 人工智能在微生物菌株开发和应用中的研究进展

周小宇, 张宇菲, 袁丁丁, 姚丽华\*, 贺斌\*

江西科技师范大学 生命科学学院, 江西 南昌

周小宇, 张宇菲, 袁丁丁, 姚丽华, 贺斌. 人工智能在微生物菌株开发和应用中的研究进展[J]. 微生物学报, 2026, 66(4): 1554-1568.

ZHOU Xiaoyu, ZHANG Yufei, YUAN Dingding, YAO Lihua, HE Bin. Research progress of artificial intelligence in the development and application of microbial strains[J]. *Acta Microbiologica Sinica*, 2026, 66(4): 1554-1568.

**摘要:** 微生物作为地球上最古老且种类最丰富的生命形式, 因其具有差异化代谢潜能与生物合成能力, 为新型药物、天然活性产物的开发提供了核心资源, 具有重要的开发价值。目前, 人工智能 (artificial intelligence, AI) 与微生物菌株开发的深度融合, 正推动生命科学领域从“经验筛选”向“理性设计”的范式变革。这一变革既源于传统研究方法在解决微生物资源复杂性方面存在局限性, 也得益于 AI 在多组学数据分析、模型预测及实验流程优化方面具有的独特优势。本文系统综述了 AI 在微生物菌株开发与应用中的作用, 涵盖菌种选育、代谢产物开发、疾病诊断与治疗以及外源物质合成 4 个方面。此外, 本文还讨论了 AI 在菌株开发过程中的核心优势及存在的局限性。概言之, AI 通过自动化建模与科学化预测, 不仅加速了微生物菌株开发进程, 还提供了多维度优化策略, 成为推动技术革新的核心驱动力。AI 技术融合有望突破传统产业瓶颈, 推动微生物产业可持续发展。

**关键词:** 人工智能; 菌种选育; 代谢产物开发; 疾病诊疗; 外源物质合成

## Research progress of artificial intelligence in the development and application of microbial strains

ZHOU Xiaoyu, ZHANG Yufei, YUAN Dingding, YAO Lihua\*, HE Bin\*

College of Life Sciences, Jiangxi Science and Technology Normal University, Nanchang, Jiangxi, China

**Abstract:** Microorganisms, as the oldest and most diverse life forms on Earth, possess significant

资助项目: 国家自然科学基金(32260017, 32560212); 国家级大学生创新训练计划(202511318025); 江西省自然科学基金(20242BAB25334)

This work was supported by the National Natural Science Foundation of China (32260017, 32560212), the National College Students' Innovation and Entrepreneurship Training Program (202511318025), and the Jiangxi Provincial Natural Science Foundation (20242BAB25334).

\*Corresponding authors. E-mail: HE Bin, hebin.li@foxmail.com; YAO Lihua, yaolh7905@163.com

Received: 2025-09-05; Accepted: 2026-02-04; Published online: 2026-02-12

development value due to their differentiated metabolic potential and biosynthetic capabilities, serving as core resources for the development of novel drugs and natural active products. Currently, the deep integration of artificial intelligence (AI) with microbial strain development is driving a paradigm shift in life sciences from “empirical screening” to “rational design”. This shift is driven both by the limitations of conventional research methods in addressing the complexity of microbial resources and by the unique advantages of AI in multi-omics data analysis, model prediction, and experimental process optimization. This article systematically reviews the roles of AI in the development and application of microbial strains, covering four aspects: strain breeding, metabolite development, disease diagnosis and treatment, and xenobiotic synthesis. Additionally, it discusses the core advantages and existing limitations of AI in the strain development process. In summary, through automated modeling and scientific prediction, AI not only accelerates the efficiency of microbial strain development but also provides multi-dimensional optimization strategies, serving as a core driver for technological innovation. The integration of AI is expected to break through traditional industrial bottlenecks and promote the sustainable development of the microbial industry.

**Keywords:** artificial intelligence; strain breeding; metabolite development; disease diagnosis and treatment; xenobiotic synthesis

微生物次级代谢产物是天然产物药物研发的重要来源, 在医学、农学、食品科学及生态修复等领域具有重要价值<sup>[1-4]</sup>。青霉素作为首个可规模化生产的抗生素, 不仅验证了微生物次级代谢产物的药用潜力, 更推动了天然产物研究从传统植物提取向微生物资源挖掘的转变; 在政策驱动下, 我国逐步构建了涵盖菌种选育、发酵工程和原料药生产的完整产业链, 产品主要集中在抗生素类药物<sup>[5]</sup>。然而, 病原生物抗药性的持续加剧使得抗微生物药物耐药性(antimicrobial resistance, AMR)成为全球公共卫生领域的重大威胁<sup>[6]</sup>。与此同时, 新发感染病的涌现、节能减排要求以及高产量需求等问题都迫切需要在微生物次级代谢产物的挖掘与应用、工艺体系的创新与优化及生产效率的系统提升等方面实现突破。

微生物既可为人类生存与发展提供关键性支撑, 又可能诱发病原性感染、干扰生产系统稳态, 人类对微生物的理性干预并非以彻底清除为目标, 而是通过定向改造与精准调控, 实现其应用价值的最大化, 具体包括有益菌种选育、病原微生物鉴定、抗菌药物研发、生产环

境优化及个性化治疗方案设计等。上述目标的达成, 亟需相关技术体系提供坚实支撑并实现持续性创新。

传统的微生物菌株开发与应用长期依赖于“试错+优化”的方法, 存在周期长、成本高、盲目性强等局限, 难以深入解析复杂的代谢网络。现代技术如规律成簇的间隔短回文重复序列系统及相关蛋白(clustered regularly interspaced short palindromic repeats, CRISPR-associated, CRISPR-Cas)系统基因编辑技术、高通量筛选(high-throughput screening, HTS)等虽能实现精准改造与快速筛选, 但仍需在海量基因组数据解析、代谢途径理性设计、表型-基因型关联预测等方面寻求更高效的解决方案。

在此背景下, 人工智能(artificial intelligence, AI)作为一项跨学科技术, 通过整合多组学数据, 利用机器学习与深度学习算法, 实现对微生物复杂代谢网络的快速解析、高产菌株的智能预测、合成途径的自动化设计, 以及发酵过程的动态优化等, 为微生物菌株开发与应用开辟了全新路径<sup>[7-9]</sup>。例如, AI辅助的基因组挖掘可高效发现新型生物合成基因簇(biosynthetic gene

cluster, BGC); 机器学习模型可预测蛋白质结构与功能, 指导酶工程改造; 强化学习可优化代谢通路, 提升目标产物产量。AI 有望快速解析微生物复杂的遗传信息和代谢网络, 精准预测菌株性能, 提升微生物菌株开发与应用的效率, 成为推动微生物领域发展的核心驱动力(图 1、图 2)。然而, AI 在微生物领域的应用仍面临数据质量、算法可靠性、生物伦理安全与工程化落地等挑战, 未来需在提升预测可靠性、平衡创新与风险等方面持续完善以推动微生物技术迈向智能化、可持续发展的新阶段。

本文探讨了 AI 在微生物菌株开发和应用中的变革性影响, 系统介绍其在菌株选育、代谢产物开发、疾病诊断与治疗以及外源物质合成等方面的应用现状与优势, 强调了其在微生物开发方面的关键作用。另外, 本文还客观阐明了 AI 应用的局限性以及未来发展方向, 强调了数据驱动、风险评估及伦理安全的重要性, 并提出了可行性建议, 为 AI 赋能优化微生物菌株

开发与应用策略奠定了基础。

## 1 AI 在微生物菌种选育中的应用

微生物菌种选育是微生物工程的基础环节, 指在微生物学、遗传学等理论指导下, 通过人工干预手段对微生物菌株进行筛选、改造和优化以获得具有特定优良性状菌种的过程, 在微生物菌种开发与产业化应用中占据关键性战略地位。AI 技术的引入, 为突破传统局限提供了系统性解决方案, 不仅体现在微生物菌种选育过程中的筛选精度与育种效率的提升, 更在于构建从基因到表型的可解释性预测框架, 推动菌种选育从“经验驱动”向“数据驱动+理性设计”转型。

AI 基于微生物基因组学、代谢组学, 结合计算机模拟, 深度挖掘微生物基因组数据及代谢机制, 构建基因与性能之间的关联模型, 可

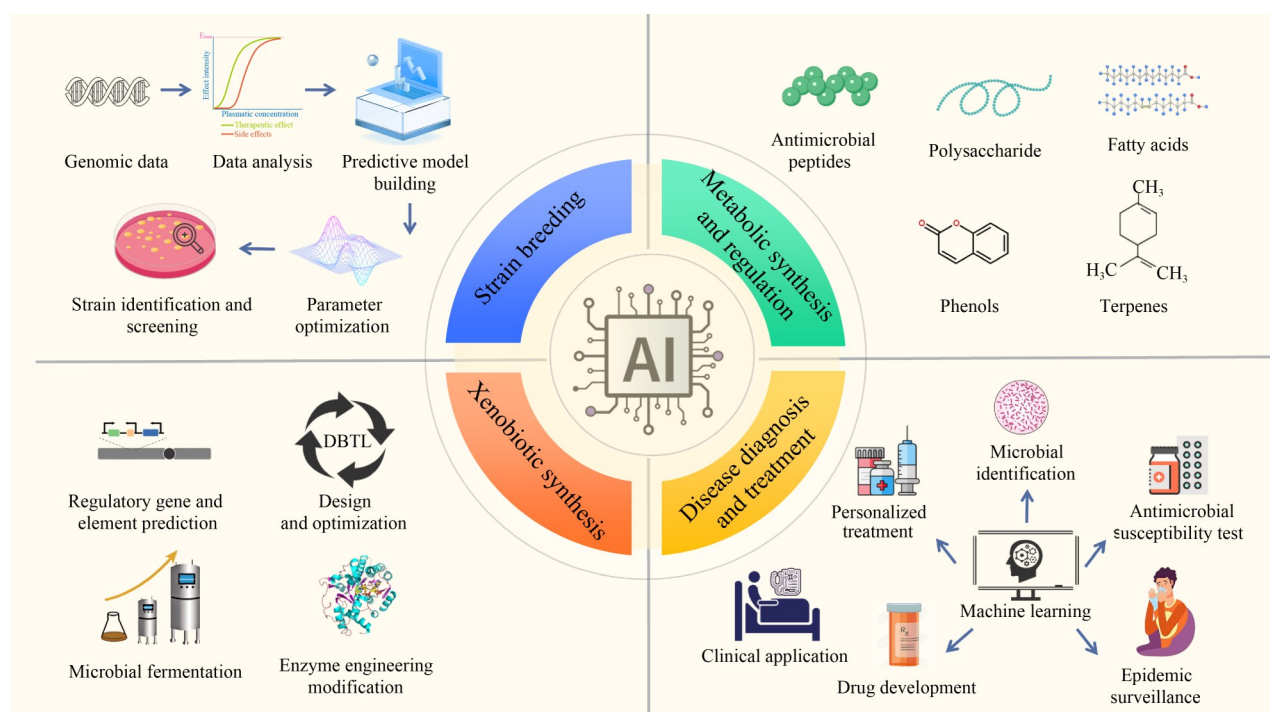


图1 人工智能在微生物菌株开发和应用中的作用

Figure 1 The role of artificial intelligence in the development and application of microbial strains.

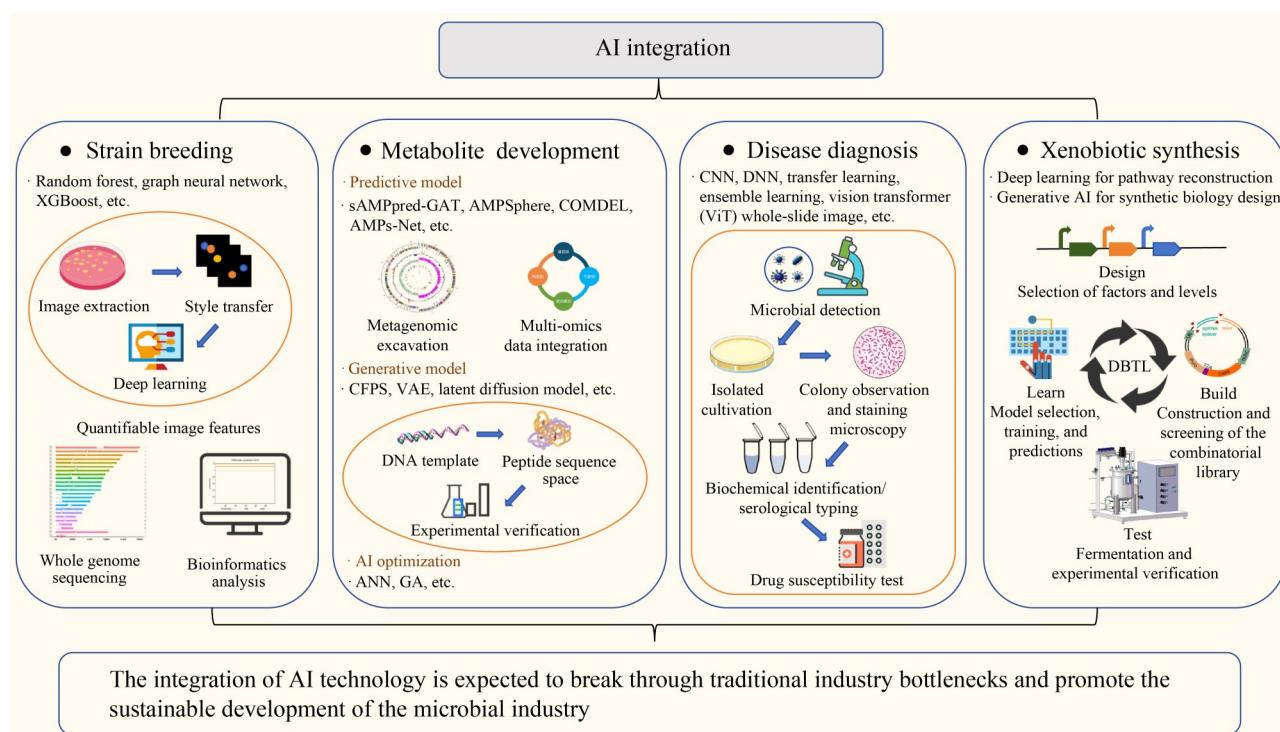


图2 AI与微生物领域融合的系统框架

Figure 2 System framework for the integration of AI and microbiology.

精准预测菌株生产性能。例如，结合全基因组测序、生物信息学分析、AI和机器学习等的精准益生菌基因分型技术，能有效提高数据分析结果的准确性和可靠性，克服数据分析缺乏深入因果关系和机制研究的弊端<sup>[10]</sup>。内蒙古农业大学团队开发出基于机器学习和基因组大数据的益生乳酸菌筛选平台，模型预测准确率高达97.77%，并成功挖掘到与益生菌特性相关的基因和代谢通路，为益生乳酸菌选育提供了高效的筛选方案<sup>[11]</sup>。还有用于优化微生物菌株设计的多智能体强化学习(multi-agent reinforcement learning, MARL)模型能够自主学习微生物菌株代谢调控规律并指导代谢工程策略<sup>[12]</sup>。

AI能将人工观察的定性判断转化为可量化的图像特征。杨帆等<sup>[13]</sup>将分层注意机制(swin transformer)与多阶段级联检测与实例分割(cascade mask R-CNN)模型架构相结合，提出一种样式迁移扩展微生物菌落图像数据集的新方

法，以更好地提取图像特征信息，展现出更强的目标检测性能，其中风格迁移能解决数据稀缺问题，swin transformer的局部注意机制与cascade mask R-CNN逐步优化边界框定位与分类能提升对菌落特征的识别提取能力，在annotated germs for automated recognition (AGAR)数据集上平均精度(mean average precision, mAP)达到61.4%，优于已有的两阶段模型方法，该研究还证明了单阶段模型方法YOLOv8x的mAP更高，在微生物检测任务中更具优势，未来可进一步探索优化cascade mask R-CNN与YOLOv8等模型，扩展更多微生物类别，凸显这类方法在微生物菌落检测与筛选中的潜力。

AI所驱动的方法为微生物菌种选育带来质的飞跃，有效提高了选育质量和效率。中国科学院青岛生物能源与过程研究所单细胞中心与中国科学院南京土壤研究所、青岛星赛生

物科技有限公司等合作开发出的人工智能辅助拉曼激活细胞分选系统(artificial intelligence-assisted Raman-activated cell sorting, AI-RACS)集成了精确的单细胞定位、自动数据收集、光镊捕获和单细胞打印,能在单细胞水平上实现表型和基因型的相关性,自动化、高效地从复杂的土壤环境样本中筛选特定功能的微生物单细胞,细胞功能更接近细胞真实功能,且结果可量化,为微生物资源探测和挖掘提供了有力手段<sup>[14]</sup>。此外,AI可与其他微生物菌株筛选技术结合,通过机器学习模型优化菌株筛选参数实现快速鉴定与筛选,提高筛选效率,例如微流控液滴技术的高通量筛选方法<sup>[15]</sup>。

一方面,微生物基因组数据庞大,代谢网络涉及数千个反应,传统方法无法解析这种复杂的非线性关系,AI能利用机器学习技术与大数据分析构建预测模型,分析微生物特性与环境因素的关联,具有高维性和复杂性数据处理的优势,从而筛选出不同环境条件下的最适微生物菌株。例如,波兰团队利用XGBoost、SVM、人工神经网络(artificial neural network, ANN)等机器学习模型分析相关微生物菌群与作物耐旱性的关联,结合土壤微生物组数据预测抗旱菌株,最佳模型准确率高达87%,为农业抗逆问题提供了数据驱动的解决方案<sup>[16]</sup>。另一方面,AI的自动化集成优势能满足高效率与高精度需求,如Lin等<sup>[17]</sup>提出一种基于卷积神经网络(convolutional neural network, CNN)和挤压激励模块(squeeze-and-excitation, SE),沿纳米孔测序过程中执行微生物DNA富集与快速自适应采样的新型深度学习模型——NanoDeep,卷积层用于从原始信号中提取局部特征,SE模块对卷积层提取的特征进行加权,自适应地强调重要特征并抑制不重要的特征,从而提升模型性能,相比现有深度学习模型,NanoDeep具有更高的测序效率和准确性,能实现复杂样本目标菌株的快速测序与筛选。各种深度学习方法还能自主学习微生物代谢规律,指导菌种改造,

在微生物菌种选育过程中集成AI预测、CRISPR编辑、高通量筛选的全自动化平台,可实现从基因设计到菌株优化的无人化操作,展现AI强大的数据整合与分析能力以及自动化模拟与运行优势。

尽管AI驱动的方法展现出巨大潜力,但当前研究仍存在局限性。NanoDeep虽在训练集和同类测试集上表现优异,但其数据类型的泛化能力仍有不足,当转移到不同生物环境中时表现出相对较低的准确性,并且该研究并未深入探讨DNA化学修饰是否会影 响模型分类的准确性,未来还需收集和整合更多样化、更大量的微生物物种实测数据进行训练,同时探索并整合多模态特征以明确DNA化学修饰对模型性能的影响。另外,基因组预测、表型识别、单细胞操作等尚未形成标准化技术整合,多数研究停留在计算预测阶段,缺乏大规模实验验证,且实际选育过程中假阳性率也需要予以考虑。AI驱动 的深度学习模型的黑箱特性导致生物学机制解释不足,限制了其在代谢工程中的可迁移性。未来研究应关注跨物种、跨环境的数据集训练,开发可解释的AI模型,并建立“计算预测-自动化实验-反馈优化”的闭环系统,真正实现从数据到产业的转化。

## 2 AI在微生物代谢产物开发中的应用

### 2.1 AI在新型抗菌肽的开发与优化中的应用

抗生素曾是人类抵御微生物感染的关键干预手段,但其滥用现象日益严重,导致大量病原微生物呈现广泛耐药性。抗菌肽(antimicrobial peptide, AMPs)作为传统抗生素的理想替代品,具备广谱抗菌活性且不易产生病菌耐药性。

近年研究表明,利用AI宏基因组挖掘与多组学数据整合扩大了对有前景抗菌肽的搜索范围,肽的功能(如抗菌活性)与其三维空间结构密

切相关, 利用一系列机器学习算法预测的抗菌肽结构信息与活性预测信息可作为识别 AMPs 的有力线索, 有望在短时间内发现更多的新型抗菌肽, 拓展抗菌肽开发空间。

sAMPpred-GAT 是一种基于预测肽结构的 AMPs 预测方法, 可根据预测的肽结构、序列信息和进化信息构建图谱。首先, 算法利用 trRosetta 工具, 仅依据输入的氨基酸序列预测其三级结构, 并提取残基对间的原子距离信息, 进而编码肽的空间折叠信息, 构建肽的图拓扑结构; 图构建完成后, 使用注意力网络(graph attention network, GAT)来学习特征, GAT 的优势在于它能自适应地衡量不同空间相邻残基的重要性从而聚合信息, 有效捕捉关键的局部空间关系, 增强模型对结构特征的判别能力<sup>[18]</sup>。

AMPsphere 是一款基于机器学习方法创建的庞大数据库, 能从环境和宿主相关栖息地的 63 410 个宏基因组和 87 920 个原核基因组中预测抗菌肽, 该数据库结合了一个基于随机森林(random forest, RF)的机器学习模型, RF 能处理高维特征, 适合大规模宏基因组数据, 每个肽序列被转化为一组理化特征(如净电荷、等电点、疏水性、两亲性等), 这些特征与抗菌活性密切相关; 模型会得到每个肽序列的“AMP 概率”, 最终仅保留概率 $\geq 0.5$  的序列作为候选抗菌肽, 预测的抗菌肽需通过实验验证其活性及作用机制, 通过体外活性筛选, 分别测试预测的 100 条合成肽样本对 ESKAPEE 病原体(如耐药鲍曼不动杆菌、大肠埃希氏菌)和人类肠道共生菌的抑制活性, 其中有 79 条肽具有抗菌活性, 63 条靶向特定病原体, 若肽序列被打乱大部分 AMPs 失去活性, 说明肽活性依赖特定序列, 该研究证实 AI 预测模型在抗菌肽挖掘中展现出卓越的预测与识别效能, 但该方法局限于由自身基因编码的肽并且严重依赖公共数据库, 而公共数据库覆盖尚不全面, 需不断更新资源<sup>[19]</sup>。

利用 AI 技术对基因组、转录组以及代谢组等进行多组学数据整合有助于更好地预测微生物

物的代谢产物谱系, 例如 Zhang 等<sup>[20]</sup>基于 AI 深度学习优化改进抗菌肽预测模型 COMDEL 并进行多组学分析, 在 35 种食用益生菌中确定植物乳植杆菌(*Lactiplantibacillus plantarum*)是最有望产生 AMPs 的候选者。预测 AMPs 的理化性质与生物活性在实际应用中至关重要, 深度学习模型 AMPs-Net 能够利用图描述原子之间的关系, 具有高度准确的特异性抗菌和抗病毒活性预测能力<sup>[21]</sup>。

除预测模型外, 生成模型也被广泛应用于抗菌肽的高效生产。例如无细胞蛋白质合成(cell-free protein synthesis, CFPS)方法可直接从 DNA 模板中快速生产 AMPs, 该方法采用变分自编码器(variational autoencoder, VAE)作为核心生成模型, 使用 UniProt 数据库进行蛋白学习, 再用来自 GRAMPA 数据库的约 5 000 个已知 AMPs 序列微调模型, 使潜在空间专注于 AMPs 特性, 最后剔除含非标准氨基酸或化学修饰的序列, 确保生成肽的可合成性; 该方法将设计-测试周期缩短至 24 h, 单次成本低, 最佳模型组合的 AMPs 发现率达 12.7%, 显著高于传统方法, 并且有 6 种从头合成的 AMPs 具有对多重耐药病原体的广谱活性, 在 21 d 连续传代实验下未产生细菌耐药性, 这些优势深刻地揭示了深度学习在开发新型 AMPs 方面发挥的重要作用<sup>[22]</sup>。

传统生成模型倾向于重组训练集中已有的序列模式, 常存在结构单一问题, 并且随着肽链长度增加, 化学空间呈指数级扩张, 传统方法难以有效覆盖, 为此, 山东大学团队提出一种基于扩散模型和分子动力学的从头设计 AMPs 的新计算流程, 开发了一个结合潜在扩散和过滤方法的抗菌肽生产系统, 扩散系统在连续的潜在空间中进行随机采样, 避免局限于训练数据分布, 从根本上解决了序列新颖性问题; 另外, 该系统还采取两阶段训练策略, 第一阶段不施加 AMPs 特异性约束, 促进模型探索完整肽序列空间; 第二阶段引入抗菌活性条件, 平

衡了多样性与功能性；该模型还具有序列设计多样化的优势，生成肽重复率低于 1%，其中 62.5% 的生成肽显示抗菌活性，覆盖革兰氏阳性菌、阴性菌和真菌，首次实现 AI 驱动的抗真菌肽生成<sup>[23]</sup>。在 AMPs 优化进程中机器学习与深度学习算法正展现出显著的应用潜力，借助人工智能工具能够对特定靶向的多肽模板进行优化修改(充分考虑肽长度和氨基酸残基组合)，构建简短而有效的抗菌肽<sup>[24-25]</sup>。

AI 在抗菌肽发现与优化生产中具有巨大潜力，借助深度学习方法和 AI 自动化实验平台不断推进“设计-合成-测试-优化”策略，为开发具有广谱抗菌活性的优质抗菌肽注入新动力。

## 2.2 AI 在其他代谢产物合成与调控中的应用

除抗菌肽外，AI 在其他代谢产物(如多糖、脂肪酸、酚萜类化合物等)合成与调控中也发挥了类似作用。例如，利用深度学习、算法模型和多组学数据分析等可以预测微生物在代谢过程中各种代谢产物的合成潜力，还可为这些产物的代谢途径挖掘以及生产优化提供多组学依据<sup>[26-27]</sup>。AI 通过智能感知与分析微生物的生长条件与代谢调控网络进一步优化培养基组分和发酵工艺参数，其中 ANN 与遗传算法(genetic algorithm, GA)已被广泛应用于菌株发酵过程优化，并且不仅局限于单一模型，研究人员也在试图利用各种混合模型策略不断探究各发酵工艺优化的最优解<sup>[28-30]</sup>。另外，Yu 等<sup>[31]</sup>和 Lv 等<sup>[32]</sup>通过多组学相结合的方法，对蛹虫草中虫草素合成的调控途径进行了深度解析，为改良蛹虫草菌(*Cordyceps militaris*)提供了理论依据；Tang 等<sup>[33]</sup>和 He 等<sup>[34]</sup>基于基因组测序和转录组测序数据对米曲霉(*Aspergillus oryzae*)中脂肪酸去饱和酶基因家族和 HOG 基因家族的特征和表达量变化进行了鉴定和分析，为米曲霉的功能基因组研究提供了素材。基于机器学习指导的环糊精选择策略所构建的酿酒酵母工程菌株，

在发酵过程中努特卡酮、柠檬烯等萜烯化合物的生物合成效率及利用率方面，相较于传统有机溶剂提取法，展现出更高的效率与环境友好性<sup>[35]</sup>。

伴随人工智能技术对微生物次级代谢网络的解析深度持续拓展，有望推动微生物代谢产物系统挖掘、高效生产及产业化应用的协同发展。

## 3 人工智能在微生物疾病诊断、监测和治疗中的应用

### 3.1 AI 在病原微生物疾病诊断上的应用

人工智能驱动的技术正逐步成为人类抵御微生物疾病的强大武器。传统病原菌诊断技术体系包括基因序列鉴定、形态学观察及生化反应测试等核心手段。然而，这些方法普遍存在样本量需求大、耗材试剂成本高昂、依赖专业技术人员操作等局限，且检测周期较长，难以满足快速诊疗的临床需求。为克服这些挑战，AI 在病原微生物的识别与鉴定方面展现出巨大潜力，借助各种机器学习模型能够实现病原微生物的快速检测与精准诊断，不仅优化了传统病原微生物检测和管理流程，还提高了鉴定效率与准确性<sup>[36]</sup>。例如，结合卷积神经网络和递归神经网络的无透镜显微镜延时摄影方法能够快速识别细菌菌落<sup>[37]</sup>；借助神经网络分析反应模式的荧光传感器不仅可以快速、准确鉴定细菌种类及其革兰氏状态，还能减少样本和试剂用量<sup>[38]</sup>。

AI 的应用不仅局限于菌株的识别与鉴定，还涉及对抗微生物药物耐药性的预测与鉴定。例如，可通过机器学习方法从全基因组测序数据中预测微生物抗生素耐药表型。由于结论数据通常与药物、病原体物种、耐药性指标、临床人群等因素直接相关，需要持续更新全基因组序列、耐药表型和治疗结果等数据，并将其纳入 AI 模型训练中，因此基于序列的抗微生物

药物耐药性预测及其应用有待进一步推广<sup>[39]</sup>。与此同时, 人工智能技术已广泛应用于抗菌药物敏感性测试(antibiotic susceptibility test, AST)以鉴定耐药菌。例如, 基于图像拼接技术的拉曼光谱检测分析方法实现了高倍率单细胞水平检测, 能够自动、快速、准确地识别耐药细菌<sup>[40]</sup>。结合液滴微流控技术和深度学习的DropDeepL AST方法可以快速、自动化地测定快速生长的肠杆菌(大肠埃希氏菌和肺炎克雷伯氏菌)的黏菌素敏感性谱<sup>[41]</sup>。

AI技术提供了一种快速、准确、自动化的菌株鉴定与耐药性监测手段, 显著提高了微生物疾病的诊断效率, 不同AI技术路径在关键性

能指标上表现各异(表1), 不仅大幅缩短了检测时间, 还大幅减少了对昂贵试剂和专业人员经验的依赖, 有助于临床医生尽快做出合理的抗菌治疗决策。

### 3.2 AI在病原微生物感染与流行病监测上的应用

利用AI驱动的算法以及AI分析实时数据的能力能够快速监测到微生物病原体, 持续跟踪并分析病原微生物种群变化, 根据反馈的数据对潜在的健康风险迅速作出反应并及时干预, 有助于降低患病风险, 预防疾病传播<sup>[47]</sup>。机器学习模型可以实时分析临床和流行病学数据, 有效预测流行病潜在暴发的趋势和模式, 例如

表1 不同AI技术在病原微生物快速检测中的关键性能指标对比<sup>[42-46]</sup>

Table 1 Comparison of key performance indicators of different AI technologies in rapid pathogen detection<sup>[42-46]</sup>

AI technology	Accuracy (%)	Sensitivity (%)	F1 score (%)	Application	Advantage	References
CNN combined with time-lapse photography (YOLOv4)	98.0	97.0	91.0	<i>Escherichia coli</i> growth phase detection (mAP 98%)	Automated detection and classification of bacteria, efficient and fast testing, high confidence interval	[42]
Deep neural network (DNN) analysis of spectral data	97.8 (surface-enhanced Raman spectroscopy using machine learning, SERS using ML)	-	-	Bacterial detection and classification	The accuracy of identification has greatly improved, and the detection limit has been significantly reduced	[43]
Transfer learning (ResNet)	99.0	98.9	98.9	Longitudinal bacterial fission classification	Precision and recall have significantly improved, and the need for labeled data has decreased by about 70%	[44]
Ensemble learning (XGBoost+matrix-assisted laser desorption-ionization time-of-flight mass spectrometry)	77.0 (serial label prediction)	-	-	Antibiotic resistance serial prediction in <i>Staphylococcus aureus</i>	The accuracy is significantly higher than methods based on a limited number of genetic markers, saving on testing costs	[45]
Vision transformer (ViT) whole-slide image	85.7 (micro-average)	85.7	85.7	Classification of Gram-stained whole-slide images (WSIs)	Simplify the diagnostic workflow, reduce diagnostic delays, and improve the accuracy of early infection management	[46]

基于传染病动力学(infectious disease dynamics, IDD)和动态贝叶斯网络(dynamical Bayesian network, DBN)的模型通过情景推演预测疫情发展状况,其预测结果与实际发展情景基本一致,该方法能够为疫情防控决策提供数据支持<sup>[48]</sup>。另外,采用集成学习方法(如 Random forest, stacking)构建混合模型通常能够有效提高传染病发病率和趋势的预测性能。例如神经网络模型(neural network model, NNM)与易感-暴露-感染-恢复模型(susceptible-exposed-infected-recovered, SEIR)相结合的混合模型,神经网络会纠正流行病学模型在预测过程中的错误,提高预测准确性<sup>[49]</sup>。或者基于堆叠技术的多算法融合学习模型,减少了单一模型的泛化误差,大幅提高了模型的疾病预测性能<sup>[50]</sup>。

病原微生物感染与各种流行病会造成重大的生命损失和持久的社会经济损害,疾病早期预警系统此时显得尤为重要,AI通过实时监控与大数据分析,能够预测疾病的未来流行趋势和传播规模,并提供更早、更精确的警告,帮助公共卫生机构在流行病暴发前及时采取遏制措施,从而保障公众健康和安<sup>[51]</sup>。

随着边缘智能设备与5G技术的普及,AI有望构建覆盖“端-边-云”的分布式监测网络,实现病原体的即时检测与跨区域协同预警。这将促进全球数据共享的同时保障隐私安全,而生成式AI可加速抗微生物药物研发。然而,数据偏见消除、模型泛化能力及监管框架完善仍是关键挑战。AI与合成生物学、纳米传感等技术的融合,或将推动传染病防控进入“智能预测-精准干预”的新纪元。

### 3.3 AI在微生物疾病药物开发与个性化治疗上的应用

除了疾病监测,AI也正不断推进微生物疾病药物开发和个性化治疗进程。各种基于机器学习的方法被用于药物靶点识别以及药物设计中,AI能够整合分析表观遗传学、基因组学、

蛋白质组学、代谢组学和多组学数据,挖掘潜在的疾病通路和药物靶点<sup>[52]</sup>。借助深度学习、免疫信息学、生物信息学的计算系统能够深入分析源自病原体蛋白质编码基因组的疫苗靶点,设计出潜在有效的多表位疫苗<sup>[53]</sup>。借助AI工具能够加快药物开发与设计进程,同时其强大的数据分析能力还能够预测微生物药敏性和安全性。使用AI驱动的语言模型能够加速靶点识别、优化临床试验、促进监管决策和增强药物警戒<sup>[54]</sup>。

AI驱动的技术贯穿从“找靶点”到“设计药物”,从“效果试验”到“临床落地”的全过程,不仅加速了药物的研发,还为临床医生与患者提供个性化和数据驱动的理解与建议,推动个性化治疗<sup>[55]</sup>。

多模态数据融合技术与边缘计算正在发展,AI有望实现实时、便携化的病原体检测,尤其在资源匮乏地区具有重要应用价值。结合区块链技术,AI可进一步优化全球病原体监测网络,实现早期疾病预警。然而,数据标准化、模型可解释性及伦理问题仍需突破。总体而言,AI将推动传染病防控迈向智能化、个体化时代。

## 4 人工智能在微生物外源物质合成中的应用

微生物合成的外源物质种类丰富,AI能够预测并设计微生物的合成代谢途径。AI算法可通过分析微生物基因组数据、代谢网络等数据预测参与调控特定外源物质合成过程的基因或调控元件等,进而设计出更为高效的代谢途径。例如,组学技术和机器学习平台作为高通量分析工具被用于促进萜烯、类黄酮、生物碱等天然产物生物合成途径的引导优化以提高微生物生产能力<sup>[56]</sup>。Moreno-Paz等<sup>[57]</sup>结合机器学习与设计-构建-测试-学习(design-build-test-learn, DBTL)循环,通过分析酵母突变体的组学数据预测代谢路径限速步骤并设计了启动子文库,该

方法使香豆酸产量提升 68%。

AI 在微生物发酵的监测与优化方面也展现出卓越能力。微生物发酵过程中的温度、pH、溶解氧、底物浓度等多种参数都会对微生物合成外源物质的产量和质量产生影响。利用传感器收集发酵过程中的实时数据, AI 可以通过建立数学模型, 快速分析这些数据之间的复杂关系, 从而精准监测与调控发酵条件。例如利用 AI 化学计量学辅助方法可实现多源传感器的综合监测, 为色氨酸的发酵生产提供一种稳定有效的质量控制策略<sup>[58]</sup>。方柏山教授课题组开发了一个由传感器、预测器、控制器和自动化系统组成的人工智能系统, 用于 1,3-丙二醇的全自动补料分批发酵, 不仅减少了人工依赖, 还提高了原料利用率和产物产量<sup>[59]</sup>。AI 同样在酶工程改造方面发挥作用。酶是微生物合成外源物质过程中的关键催化剂, 通过 AI 技术可以对蛋白质的结构和功能进行精准预测与改造, 提升其催化活性、稳定性等性能。一些基于深度学习、强化学习和迁移学习等技术能够在短时间内对蛋白质序列进行结构预测、结合亲和力优化和酶设计<sup>[60]</sup>。在合成生物学领域, CRISPR-Cas 系统具备核心应用价值, AI 能够指导

CRISPR-Cas 系统核心酶元件的改造, 不仅能准确预测突变效应、实现高通量筛选功能性关键蛋白, 还能基于蛋白质结构-功能关系进行定向设计, 为提升基因编辑效能、降低脱靶风险及扩展应用场景等提供了关键支撑<sup>[61]</sup>。

AI 在微生物生物合成外源物质领域的应用, 正不断突破传统技术的局限。AI 与合成生物学的深度融合, 将加速微生物细胞工厂的智能化升级, 为可持续生物制造开辟新路径。随着 AI 技术的持续进步与完善, 我们有理由相信, 未来微生物生物合成领域将创造出更多令人瞩目的成果, 为人类社会的发展做出更大贡献。

## 5 讨论

AI 凭借强大的数据处理和模型构建能力, 能够快速解析微生物复杂的遗传信息和代谢网络, 精准预测菌株特性并提供优化方案, 极大提升了微生物菌株开发效率和成功率, 成为推动微生物菌株开发技术革新的核心驱动力, 在微生物菌种选育、代谢产物开发、疾病诊断与治疗以及外源物质合成等方面得到广泛应用(表 2)。其核心优势体现在 3 个方面。(1) 多维度数据整合能力, 能将基因组、代谢组等碎片化数

表2 AI在微生物领域应用的主要技术现状<sup>[62-67]</sup>

Table 2 The current major technologies of AI applications in the field of microbiology<sup>[62-67]</sup>

Field	Major AI technologies	Application	Technology maturity	References
Strain breeding	Random forest, graph neural network, XGBoost, etc.	High-producing strains, CRISPR editing for metabolic engineering, etc.	Higher	[62]
Metabolite development	Convolutional neural networks, transformer models, molecular networks (machine learning for redundancy reduction)	BGCs activation in <i>Streptomyces/ Myxobacteria</i> , discovery of natural products from silent clusters, etc.	High	[63]
Disease diagnosis and treatment	Convolutional neural networks for imaging, linear models for phenotypic analysis, multimodal fusion	Pathogen detection, antimicrobial drug resistance prediction, vaccine design, etc.	High	[64-65]
Xenobiotic synthesis	Deep learning for pathway reconstruction, generative AI for synthetic biology design	Alternative proteins, bioactive compounds, bioreactor optimization, etc.	Medium (laboratory scale)	[66-67]

据转化为精准预测模型。例如 BGC-Prophet 模型通过分析微生物基因组，能够准确、快速地大规模检测与定位生物合成基因簇，该模型有助于更深入地了解微生物次生代谢物及其在合成生物学中的应用<sup>[68]</sup>。(2) 实现闭环开发，结合自动化实验平台形成 DBTL 循环。组合途径优化就是基于 DBTL 循环来执行迭代开发产品菌株，基于机器学习的菌株开发可以改善 DBTL 循环以促进将合成菌株从实验室转移到工业<sup>[69]</sup>。(3) 大幅缩短开发周期，AI 驱动数据整合、模型预测和自动化实验，显著提升了菌株设计的效率和精准度，同时，AI 技术还提供一种更加主动、智能的质量管理办法，减少了运营效率低下的问题<sup>[70]</sup>。

AI 通过数据驱动的建模与优化，打破了传统“试错式”开发的局限，同时也面临着诸多挑战，如数据泛化、模型的可解释性不足、方案实施风险以及伦理争议等。首先，复杂代谢性状需要大量地标数据，大多数实验室都难以满足，同时算法的性能完全取决于所用数据的质量，数据的不一致性会影响模型的准确性，并且训练数据的来源对于确定最终应用程序或预期用途至关重要<sup>[71]</sup>。为克服这一点，需不断丰富数据进行参考对比，并加强跨机构数据共享，统一基因测序、发酵参数等标准，解决数据碎片化问题，还可采取迁移学习加强模型对数据处理的训练，提高适应性。

其次，深度学习虽能预测高产菌株，但难以具体阐述背后机制，实验室数据训练的模型在实际生产中常因环境波动失效，AI 提出的设计方案能否正常实施有待衡量，尤其在合成生物学领域，平衡潜在风险和收益变得愈发重要。对此，建议加强 AI 设计方案的风险评估与管理工作，注重生物安全、生物伦理、生态环境和社会风险的管理，国际基因工程机器大赛 (International Genetically Engineered Machine Competition, iGEM) 在该方面就取得了一定的经验<sup>[72]</sup>。无法解释的预测模型会阻碍监管审核、

实验结果的可重复性以及从实验室到工业生物反应器放大过程中故障的排查。在菌株工程中模型的可解释性至关重要，SHapley additive explanations (SHAP) 和 local interpretable model-agnostic explanations (LIME) 等可解释 AI 融合技术通过量化特征揭示如何从多组学数据中优先考虑诸如抗逆性或生长速率等性状，从而增强信任并加速应用，同时减少来自精选数据集的偏差，并确保对未知菌株的稳健泛化。例如通过预测基因编辑结果来加速基于 CRISPR 的微生物设计，发现使酿酒酵母产量提高 300% 的启动子-基因对，或挖掘  $\alpha$ -淀粉酶生产中的关键代谢途径；在精准发酵中可解释的模型将经验试错法转变为机制驱动的设计，使生物学家能够根据数据驱动模型预测结果，并确定可编辑靶点的优先级，使得菌株筛选数量大幅减少，将彻底改变菌株开发以更透明、可重复的结果推动可持续生物制造<sup>[73-76]</sup>。

最后，AI 合成生物学技术存在各方面的伦理争议，在医学领域，AI 驱动模型构建过程中需要临床数据反馈，存在隐私担忧。另外，AI 设计的微生物可能通过水平基因转移或环境释放引发不可控后果，有一定的生态风险。例如基因组预测模型虽然有助于发现新的生物标志物或治疗靶点，加速药物和疫苗研发、改善作物产量等，但也可以催化高致病性的病原微生物的产生，增强病原体的传播能力或毒力，可能引发大规模流行病甚至全球大流行。对此，AI 模型开发者要作出生物设计责任声明，政府要充分发挥主导作用，建立强制性的、有针对性的、国际协同的监管体系，各国政府应合作制定共同的监管标准，由独立的联邦机构对 AI 模型进行生物安全与生物安保风险评估，并建立数据共享平台<sup>[77]</sup>。AI 合成生物学技术还可能被用于生物武器开发，一方面，AI 可以使复杂的科学任务实现自动化，大幅降低开发生物武器的技术门槛；另一方面，AI 通过定制化设计，可能被用于制造更高效、更稳定的生物战剂或

新型病原体, 增强生物武器的威胁性。对此, 需要完善相关法律和道德法规, 加强伦理审查, 明确各部门和科研人员在研发过程中的技术和伦理责任, 确保生物研究安全, 防止 AI 技术被不法分子滥用。

## 6 总结与展望

随着 AI 技术不断进步与发展, AI 与微生物经济的结合将日益紧密, 进一步加速科技与产业变革的步伐, 这不仅为微生物新资源转化为新质生产力提供了重要契机<sup>[78]</sup>, 也对数据基础与模型能力提出了更高要求, 未来需在持续积累与整合微生物研究数据的基础上, 构建更为复杂、精准的 AI 模型。未来 5-10 年, 新兴 AI 技术将深刻改变微生物学的研究范式与应用前景, 大语言模型和多模态 AI 不仅会加速基础研究突破, 例如微生物代谢网络以及功能注释、微生物互作、抗生素耐药性预测等, 随着技术不断创新以及加速转化, AI 技术更将推动微生物技术从研究工具向设计平台的转变, 如微生物数字孪生平台、AI-自动化实验闭环系统、全球微生物组智能监测网络等。成功的关键在于跨学科深度融合、高质量数据积累以及负责任创新框架。除了能够完成功能预测、筛选、验证和优化, 还应当鼓励发展防御性技术(如推进遗传工程检测、微生物法医学和 DNA 合成筛选技术的研究等)以提高人类对生物技术的可控能力<sup>[79-80]</sup>。正确处理 AI 合成生物学技术问题的关键在于平衡创新与安全。过度监管或抵制可能会阻碍生物技术的进步, 而监管不足则可能带来巨大风险。因此, 必须采取一种更主动、跨学科合作的方法, 加强 AI 研究人员、生物工程师和安全专家之间的深入交流与合作, 制定出灵活、有效且创新的发展策略, 实现菌株开发全流程智能化, 推动微生物产业升级。

### 作者贡献声明

周小宇: 论文撰写和修改、文章构思; 张宇菲:

参与论文讨论、协助论文修改; 袁丁丁: 参与论文讨论; 姚丽华、贺斌: 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] Carlino N, Blanco-Míguez A, Punčochář M, Mengoni C, Pinto F, Tatti A, Manghi P, Armanini F, Avagliano M, Barcenilla C, Breselge S, Cabrera-Rubio R, Calvete-Torre I, Coakley M, Cobo-Díaz JF, de Filippis F, Dey H, Leech J, Klaassens ES, Knobloch S, et al. Unexplored microbial diversity from 2 500 food metagenomes and links with the human microbiome[J]. *Cell*, 2024, 187(20): 5775-5795.e15.
- [2] Zelaya-Molina LX, Chávez-Díaz IF, Ramos-Garza J, Fajardo-Franco ML, Hernández-Rodríguez C. Agriculture assisted by microbial genetic resources: current and future scenarios[J]. *Current Research in Microbial Sciences*, 2024, 6: 100222.
- [3] Waluyo D, Prabandari EE, Pramisanadi A, Hidayati DN, Chrisnayanti E, Puspitasari DJ, Dewi D, Suryani, Kristiningrum, Oktaviani AN, Afrianti KR, Nonaka K, Matsumoto A, Tokiwa T, Adipratiwi N, Ariyani T, Hartuti ED, Putri TZ, Rahmawati Y, Inaoka DK, et al. Exploring natural microbial resources for the discovery of anti-malarial compounds[J]. *Parasitology International*, 2021, 85: 102432.
- [4] Lima N. Preserving and utilizing microbial diversity for innovation and sustainability[J]. *Microbiology*, 2025, 171(3): 001544.
- [5] 于晴, 黄婷婷, 邓子新. 微生物药物产业现状与发展趋势[J]. *中国工程科学*, 2021, 23(5): 69-78.  
Yu Q, Huang TT, Deng ZX. Microbial medicine industry: current status and future trends[J]. *Strategic Study of CAE*, 2021, 23(5): 69-78 (in Chinese).
- [6] GBD 2021 Antimicrobial Resistance Collaborators. Global burden of bacterial antimicrobial resistance 1990-2021: a systematic analysis with forecasts to 2050[J]. *The Lancet*, 2024, 404(10459): 1199-1226.
- [7] Lv XQ, Li Y, Xiu X, Liao C, Xu YM, Liu YF, Li JH, Du GC, Liu L. CRISPR genetic toolkits of classical food microorganisms: current state and future prospects[J]. *Biotechnology Advances*, 2023, 69: 108261.
- [8] Sarnaik A, Liu A, Nielsen D, Varman AM. High-throughput screening for efficient microbial biotechnology[J]. *Current Opinion in Biotechnology*, 2020, 64: 141-150.
- [9] Gong XY, Zhang JL, Gan Q, Teng YX, Hou JX, Lyu YJ, Liu ZL, Wu ZH, Dai RP, Zou YS, Wang XQ, Zhu DJ, Zhu HT, Liu TM, Yan YJ. Advancing microbial production through artificial intelligence-aided biology[J]. *Biotechnology Advances*, 2024, 74: 108399.
- [10] 李佳, 黄京山, 吴红霞, 仇华吉, 孙元. 益生菌的精准筛

- 选与应用[J]. 微生物学报, 2024, 64(11): 4054-4068.
- Li J, Huang JS, Wu HX, Qiu HJ, Sun Y. Precise screening and applications of probiotics[J]. *Acta Microbiologica Sinica*, 2024, 64(11): 4054-4068 (in Chinese).
- [11] Sun Y, Li HC, Zheng L, Li JZ, Hong Y, Liang PF, Kwok LY, Zuo YC, Zhang WY, Zhang HP. iProbiotics: a machine learning platform for rapid identification of probiotic properties from whole-genome primary sequences[J]. *Briefings in Bioinformatics*, 2022, 23: bbab477.
- [12] Sabzevari M, Szedmak S, Penttilä M, Jouhten P, Rousu J. Strain design optimization using reinforcement learning[J]. *PLoS Computational Biology*, 2022, 18(6): e1010177.
- [13] Yang F, Zhong YJ, Yang H, Wan Y, Hu ZH, Peng SS. Microbial colony detection based on deep learning[J]. *Applied Sciences*, 2023, 13(19): 10568.
- [14] Diao ZD, Jing XY, Hou XB, Meng Y, Zhang JP, Wang YS, Ji YT, Ge AL, Wang XX, Liang YT, Xu J, Ma B. Artificial intelligence-assisted automatic Raman-activated cell sorting (AI-RACS) system for mining specific functional microorganisms in the microbiome[J]. *Analytical Chemistry*, 2024, 96(46): 18416-18426.
- [15] Zhang ZD, Guo Q, Wang YT, Huang H. High-throughput screening of microbial strains in large-scale microfluidic droplets[J]. *Frontiers in Bioengineering and Biotechnology*, 2023, 11: 1105277.
- [16] Miller T, Mikiciuk G, Kisiel A, Mikiciuk M, Paliwoda D, Sas-Paszt L, Cembrowska-Lech D, Krzemińska A, Kozioł A, Brysiewicz A. Machine learning approaches for forecasting the best microbial strains to alleviate drought impact in agriculture[J]. *Agriculture*, 2023, 13(8): 1622.
- [17] Lin YS, Zhang YJ, Sun H, Jiang H, Zhao X, Teng XJ, Lin JX, Shu BW, Sun H, Liao YH, Zhou JJ. NanoDeep: a deep learning framework for nanopore adaptive sampling on microbial sequencing[J]. *Briefings in Bioinformatics*, 2023, 25(1): bbad499.
- [18] Yan K, Lv HW, Guo YC, Peng W, Liu B. sAMPpred-GAT: prediction of antimicrobial peptide by graph attention network and predicted peptide structure[J]. *Bioinformatics*, 2023, 39: btac715.
- [19] Santos-Júnior CD, Torres MDT, Duan YQ, Rodríguez del Río Á, Schmidt TSB, Chong H, Fullam A, Kuhn M, Zhu CK, Houseman A, Somborski J, Vines A, Zhao XM, Bork P, Huerta-Cepas J, de la Fuente-Nunez C, Coelho LP. Discovery of antimicrobial peptides in the global microbiome with machine learning[J]. *Cell*, 2024, 187(14): 3761-3778.e16.
- [20] Zhang Y, Liu LH, Xu B, Zhang ZQ, Yang M, He YY, Chen JJ, Zhang Y, Hu YC, Chen XP, Sun ZT, Ge QJ, Wu S, Lei W, Li KZ, Cui H, Yang GZ, Zhao XM, Wang M, Xia JQ, et al. Screening antimicrobial peptides and probiotics using multiple deep learning and directed evolution strategies[J]. *Acta Pharmaceutica Sinica B*, 2024, 14(8): 3476-3492.
- [21] Ruiz Puentes P, Henao MC, Cifuentes J, Muñoz-Camargo C, Reyes LH, Cruz JC, Arbeláez P. Rational discovery of antimicrobial peptides by means of artificial intelligence[J]. *Membranes*, 2022, 12(7): 708.
- [22] Pandi A, Adam D, Zare A, Trinh VT, Schaefer SL, Burt M, Klabunde B, Bobkova E, Kushwaha M, Foroughijabbari Y, Braun P, Spahn C, Preußner C, Pogge von Strandmann E, Bode HB, von Buttlar H, Bertrams W, Jung AL, Abendroth F, Schmeck B, et al. Cell-free biosynthesis combined with deep learning accelerates *de novo*-development of antimicrobial peptides[J]. *Nature Communications*, 2023, 14: 7197.
- [23] Wang YJ, Song MH, Liu FJ, Liang Z, Hong R, Dong YM, Luan HZ, Fu XJ, Yuan WC, Fang WJ, Li G, Lou HX, Chang WQ. Artificial intelligence using a latent diffusion model enables the generation of diverse and potent antimicrobial peptides[J]. *Science Advances*, 2025, 11(6): eadp7171.
- [24] Mishra B, Basu A, Shehadeh F, Felix L, Kollala SS, Chhonker YS, Naik MT, Dellis C, Zhang LY, Ganesan N, Murry DJ, Gu JH, Sherman MB, Ausubel FM, Sotiriadis PP, Mylonakis E. Antimicrobial peptide developed with machine learning sequence optimization targets drug resistant *Staphylococcus aureus* in mice[J]. *Journal of Clinical Investigation*, 2025, 135(12): e185430.
- [25] Deka H, Pawar A, Battula M, Ghfar AA, Assal ME, Chikhale RV. Identification and design of novel potential antimicrobial peptides targeting mycobacterial protein kinase PknB[J]. *The Protein Journal*, 2024, 43(4): 858-868.
- [26] Tian YS, Yang X, Chen NH, Li CY, Yang WL. Data-driven interpretable analysis for polysaccharide yield prediction[J]. *Environmental Science and Ecotechnology*, 2024, 19: 100321.
- [27] Wu SB, Zhou HN, Chen DL, Lu YT, Li YN, Qiao JJ. Multi-omic analysis tools for microbial metabolites prediction[J]. *Briefings in Bioinformatics*, 2024, 25(4): bbac264.
- [28] Xia JY, Long DJ, Chen M, Chen AX. Optimization of fermentation processes in intelligent biomanufacturing: on online monitoring, artificial intelligence, and digital twin technologies[J]. *Chinese Journal of Biotechnology*, 2025, 41(3): 1179-1196.
- [29] Wang YS, Yang G, Sage V, Xu J, Sun GZ, He J, Sun Y. Optimization of dark fermentation for biohydrogen production using a hybrid artificial neural network (ANN) and response surface methodology (RSM) approach[J]. *Environmental Progress & Sustainable Energy*, 2021, 40: e13485.
- [30] Zhang LH, Chao B, Zhang X. Modeling and optimization of microbial lipid fermentation from cellulosic ethanol wastewater by *Rhodotorula glutinis* based on the support vector machine[J]. *Bioresource Technology*, 2020, 301: 122781.
- [31] Yu WB, Zhao ZY, Zeng B, Tu YY, He B. Alanine supplementation enhancing cordycepin production in *Cordyceps militaris* via upregulation of *Cns2* and *Cns3* genes expression levels[J]. *Journal of Food and Drug Analysis*, 2024, 32(4): 589-602.
- [32] Lv GB, Zhu Y, Cheng XJ, Cao Y, Zeng B, Liu XP, He B. Transcriptomic responses of *Cordyceps militaris* to salt treatment during cordycepin production[J]. *Frontiers in Nutrition*, 2021, 8: 793795.

- [33] Tang W, Ouyang CS, Liu LL, Li HR, Zeng CH, Wang J, Fu LJ, Wu QQ, Zeng B, He B. Genome-wide identification of the fatty acid desaturases gene family in four *Aspergillus* species and their expression profile in *Aspergillus oryzae*[J]. *AMB Express*, 2018, 8: 169.
- [34] He B, Tu YY, Hu ZH, Ma L, Dai J, Cheng XJ, Li HR, Liu LL, Zeng B. Genome-wide identification and expression profile analysis of the *HOG* gene family in *Aspergillus oryzae*[J]. *World Journal of Microbiology and Biotechnology*, 2018, 34(2): 35.
- [35] Cai SL, Chen DY, Cai JP, Tan AL, Zhou JT, Zhuo M, Liu MF, Zhu CY, Li S. Machine learning-guided selection of cyclodextrins for enhanced biosynthesis and capture of volatile terpenes[J]. *Journal of Agricultural and Food Chemistry*, 2025, 73(6): 3602-3610.
- [36] Zhang XY, Zhang D, Zhang XF, Zhang X. Artificial intelligence applications in the diagnosis and treatment of bacterial infections[J]. *Frontiers in Microbiology*, 2024, 15: 1449844.
- [37] Paquin P, Durmort C, Paulus C, Vernet T, Marcoux PR, Morales S. Spatio-temporal based deep learning for rapid detection and identification of bacterial colonies through lens-free microscopy time-lapses[J]. *PLoS Digital Health*, 2022, 1(10): e0000122.
- [38] Laliwala A, Svehkarev D, Sadykov MR, Endres J, Bayles KW, Mohs AM. Simpler procedure and improved performance for pathogenic bacteria analysis with a paper-based ratiometric fluorescent sensor array[J]. *Analytical Chemistry*, 2022, 94(5): 2615-2624.
- [39] Hicks AL, Wheeler N, Sánchez-Busó L, Rakeman JL, Harris SR, Grad YH. Evaluation of parameters affecting performance and reliability of machine learning-based antibiotic susceptibility testing from whole genome sequencing data[J]. *PLoS Computational Biology*, 2019, 15(9): e1007349.
- [40] Dou XH, Yang FN, Wang N, Xue Y, Hu HR, Li B. Rapid detection and analysis of Raman spectra of bacteria in multiple fields of view based on image stitching technique[J]. *Frontiers in Bioscience-Landmark*, 2023, 28(10): 249.
- [41] Riti J, Sutra G, Naas T, Volland H, Simon S, Perez-Toralla K. Combining deep learning and droplet microfluidics for rapid and label-free antimicrobial susceptibility testing of colistin[J]. *Biosensors and Bioelectronics*, 2024, 257: 116301.
- [42] Chin SY, Dong J, Hasikin K, Ngui R, Lai KW, Yeoh PSQ, Wu X. Bacterial image analysis using multi-task deep learning approaches for clinical microscopy[J]. *PeerJ Computer Science*, 2024, 10: e2180.
- [43] Hussain M, Zou J, Zhang H, Zhang R, Chen Z, Tang YJ. Recent progress in spectroscopic methods for the detection of foodborne pathogenic bacteria[J]. *Biosensors*, 2022, 12(10): 869.
- [44] Garcia-Perez C, Ito K, Geijo J, Feldbauer R, Schreiber N, zu Castell W. Efficient detection of longitudinal bacteria fission using transfer learning in deep neural networks[J]. *Frontiers in Microbiology*, 2021, 12: 645972.
- [45] Zhang JH, Wang Z, Wang HY, Chung CR, Horng JT, Lu JJ, Lee TY. Rapid antibiotic resistance serial prediction in *Staphylococcus aureus* based on large-scale MALDI-TOF data by applying XGBoost in multi-label learning[J]. *Frontiers in Microbiology*, 2022, 13: 853775.
- [46] McMahon J, Tomita N, Tatishev ES, Workman AA, Costales CR, Banaei N, Martin IW, Hassanpour S. A novel framework for the automated characterization of Gram-stained blood culture slides using a large-scale vision transformer[J]. *Journal of Clinical Microbiology*, 2025, 63(3): e01514-24.
- [47] Shelke YP, Badge AK, Bankar NJ. Applications of artificial intelligence in microbial diagnosis[J]. *Cureus*, 2023: e49366.
- [48] Gao S, Wang HY. Scenario prediction of public health emergencies using infectious disease dynamics model and dynamic Bayes[J]. *Future Generation Computer Systems*, 2022, 127: 334-346.
- [49] Delli Compagni R, Cheng Z, Russo S, Van Boeckel TP. A hybrid Neural Network-SEIR model for forecasting intensive care occupancy in Switzerland during COVID-19 epidemics[J]. *PLoS One*, 2022, 17(3): e0263789.
- [50] Cheng YY, Bai YP, Yang J, Tan XH, Xu T, Cheng R. Analysis and prediction of infectious diseases based on spatial visualization and machine learning[J]. *Scientific Reports*, 2024, 14: 28659.
- [51] Villanueva-Miranda I, Xiao GH, Xie Y. Artificial intelligence in early warning systems for infectious disease surveillance: a systematic review[J]. *Frontiers in Public Health*, 2025, 13: 1609615.
- [52] You YJ, Lai X, Pan Y, Zheng HR, Vera J, Liu SR, Deng SY, Zhang L. Artificial intelligence in cancer target identification and drug discovery[J]. *Signal Transduction and Targeted Therapy*, 2022, 7: 156.
- [53] Rawal K, Sinha R, Abbasi BA, Chaudhary A, Nath SK, Kumari P, Preeti P, Saraf D, Singh S, Mishra K, Gupta P, Mishra A, Sharma T, Gupta S, Singh P, Sood S, Subramani P, Dubey AK, Strych U, Hotez PJ, et al. Identification of vaccine targets in pathogens and design of a vaccine using computational approaches[J]. *Scientific Reports*, 2021, 11: 17626.
- [54] Liu ZC, Roberts RA, Lal-Nag M, Chen X, Huang RL, Tong WD. AI-based language models powering drug discovery and development[J]. *Drug Discovery Today*, 2021, 26(11): 2593-2607.
- [55] Bilgin GB, Bilgin C, Burkett BJ, Orme JJ, Childs DS, Thorpe MP, Halfdanarson TR, Johnson GB, Kendi AT, Sartor O. Theranostics and artificial intelligence: new frontiers in personalized medicine[J]. *Theranostics*, 2024, 14(6): 2367-2378.
- [56] Ramzi AB, Baharum SN, Bunawan H, Scrutton NS. Streamlining natural products biomanufacturing with omics and machine learning driven microbial engineering[J]. *Frontiers in Bioengineering and Biotechnology*, 2020, 8: 608918.
- [57] Moreno-Paz S, van der Hoek R, Eliana E, Zwartjens P, Gosiewska S, Martins dos Santos VAP, Schmitz J, Suarez-Diez M. Machine learning-guided optimization of *p*-coumaric acid production in yeast[J]. *ACS Synthetic Biology*, 2024, 13(4): 1312-1322.
- [58] Ma XB, Tian MY, Huang RQ, Zhang H, Yue JN, Jia HH,

- Wang Y, Li YD, Gai C, Li L, Zang HC. AI-chemometric assisted real-time monitoring of tryptophan fermentation process using a sensor fusion strategy[J]. *Food Chemistry*, 2025, 489: 145037.
- [59] Huang JC, Li CD, Zhao HD, Yu M, Zhang AH, Fang BS. Artificial intelligence system for enhanced automated 1,3-propanediol green biosynthesis[J]. *Green Chemistry*, 2023, 25(22): 9175-9186.
- [60] Son A, Park J, Kim W, Yoon Y, Lee S, Park Y, Kim H. Revolutionizing molecular design for innovative therapeutic applications through artificial intelligence[J]. *Molecules*, 2024, 29(19): 4626.
- [61] Mao YF, Chu GY, Liang QL, Liu Y, Yang Y, Liao XP, Wang M. Artificial intelligence-assisted design, mining, and modification of CRISPR-Cas systems[J]. *Chinese Journal of Biotechnology*, 2025, 41(3): 949-967.
- [62] Watcharawipas A, Rungphan W, Khamwachirapithak P, Laothanachareon T. Integrating yeast biodiversity and machine learning for predictive metabolic engineering[J]. *FEMS Yeast Research*, 2025, 25: foaf072.
- [63] Sahayasheela VJ, Lankadasari MB, Dan VM, Dastager SG, Pandian GN, Sugiyama H. Artificial intelligence in microbial natural product drug discovery: current and emerging role[J]. *Natural Product Reports*, 2022, 39(12): 2215-2230.
- [64] Tian T, Zhang X, Zhang F, Huang XH, Li ML, Quan ZW, Wang WY, Lei JW, Wang YT, Liu Y, Wang JH. Harnessing AI for advancing pathogenic microbiology: a bibliometric and topic modeling approach[J]. *Frontiers in Microbiology*, 2024, 15: 1510139.
- [65] Suster CJE, Pham D, Kok J, Sintchenko V. Emerging applications of artificial intelligence in pathogen genomics[J]. *Frontiers in Bacteriology*, 2024, 3: 1326958.
- [66] Merzbacher C, Oyarzún DA. Applications of artificial intelligence and machine learning in dynamic pathway engineering[J]. *Biochemical Society Transactions*, 2023, 51(5): 1871-1879.
- [67] Goshisht MK. Machine learning and deep learning in synthetic biology: key architectures, applications, and challenges[J]. *ACS Omega*, 2024, 9(9): 9921-9945.
- [68] Lai QL, Yao S, Zha YG, Zhang HH, Zhang HB, Ye Y, Zhang YH, Bai H, Ning K. Deciphering the biosynthetic potential of microbial genomes using a BGC language processing neural network model[J]. *Nucleic Acids Research*, 2025, 53(7): gkaf305.
- [69] Liao XP, Ma HW, Tang YJ. Artificial intelligence: a solution to involution of design-build-test-learn cycle[J]. *Current Opinion in Biotechnology*, 2022, 75: 102712.
- [70] Yu WB, Ouyang ZW, Zhang YF, Lu Y, Wei CH, Tu YY, He B. Research progress on the artificial intelligence applications in food safety and quality management[J]. *Trends in Food Science & Technology*, 2025, 156: 104855.
- [71] DeYoung B, Morales M, Giglio S. Microbiology 2.0: a “behind the scenes” consideration for artificial intelligence applications for interpretive culture plate reading in routine diagnostic laboratories[J]. *Frontiers in Microbiology*, 2022, 13: 976068.
- [72] Millett P, Alexanian T. Implementing adaptive risk management for synthetic biology: lessons from iGEM’s safety and security programme[J]. *Engineering Biology*, 2021, 5(3): 64-71.
- [73] Sufi F. Generative AI in microbial evolution and resistance: toward robust, explainable, and equitable predictions[J]. *Frontiers in Microbiology*, 2025, 16: 1705320.
- [74] Priyadharshini D, Muthuvel I, Saraswathy S, Kavitha PS, Jegadeeswari V. Precision to plate: AI-driven innovations in fermentation and hyper-personalized diets[J]. *Frontiers in Nutrition*, 2025, 12: 1659511.
- [75] Wang Y, Wang YH, Xu F. Artificial intelligence-driven fermentation optimization for  $\alpha$ -amylase hyperproduction enabled by Raman monitoring and metabolic network analysis[J]. *Bioresource Technology*, 2026, 439: 133287.
- [76] Xiong LS, Yu DQ, Li HF. Artificial intelligence-based strain screening and optimization strategies: technical framework, current status, and future prospects[J]. *Advances in Resources Research*, 2025, 5(3): 1571-1588.
- [77] Bloomfield D, Pannu J, Zhu AW, Ng MY, Lewis A, Bendavid E, Asch SM, Hernandez-Boussard T, Cicero A, Inglesby T. AI and biosecurity: the need for governance[J]. *Science*, 2024, 385(6711): 831-833.
- [78] 张晓玲, 杨桥. 藻际微生物学: 全面助推可持续发展的新兴交叉学科[J]. *微生物学报*, 2025, 65(5): 1831-1848.
- Zhang XL, Yang Q. Phycosphere Microbiology: an emerging interdisciplinary comprehensively promoting sustainable development[J]. *Acta Microbiologica Sinica*, 2025, 65(5): 1831-1848 (in Chinese).
- [79] Ou YK, Guo SJ. Safety risks and ethical governance of biomedical applications of synthetic biology[J]. *Frontiers in Bioengineering and Biotechnology*, 2023, 11: 1292029.
- [80] Wheeler NE. Responsible AI in biotechnology: balancing discovery, innovation and biosecurity risks[J]. *Frontiers in Bioengineering and Biotechnology*, 2025, 13: 1537471.