



我国学者在数据驱动的抗菌肽虚拟筛选领域取得进展

日期 2023-02-08 来源: 工程与材料科学部 作者: 丁玉琴 饶静一 【大 中 小】 【打印】 【关闭】

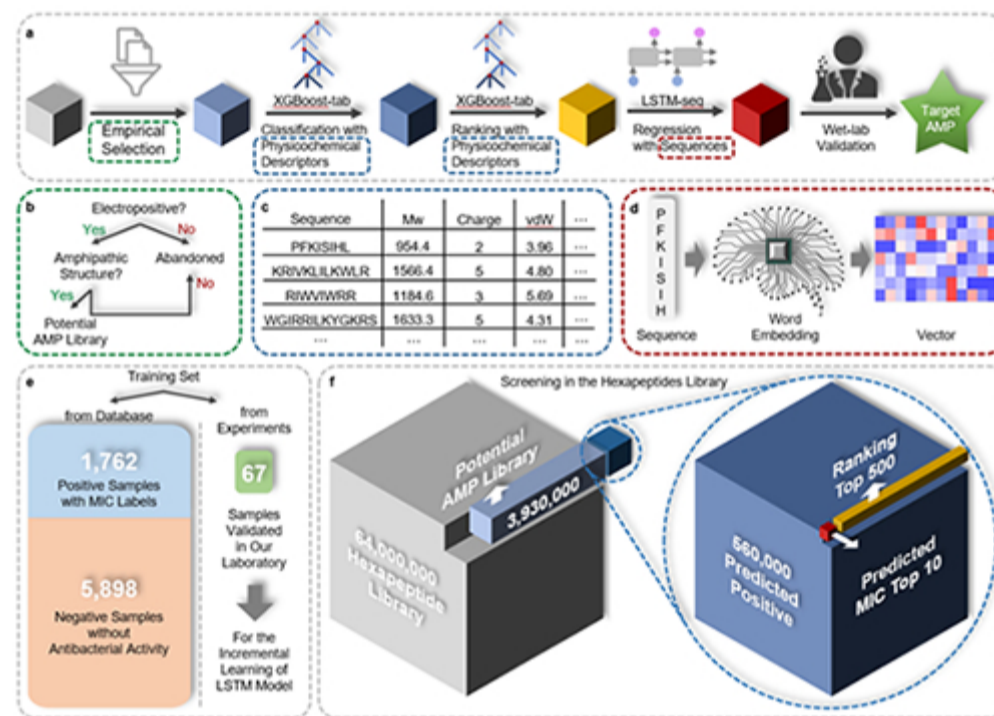


图1 机器学习组合模型识别抗菌肽概述

表1 抗菌六肽对各种细菌的MIC测试 (ug/mL)

Sequence	<i>S. aureus</i> 6538	MRSA 44300	<i>S. aureus</i> 2904	<i>S. haemolyticus</i> 1303	<i>P. aeruginosa</i> 2512	<i>E. coli</i> 1806	<i>A. baumannii</i> 2402
CRR13	8	16	32	8	16	64	32
CRR14	8	32	64	8	64	64	32
CRR17	8	8	16	4	32	32	16
MP196	32	64	128	8	>256	64	32

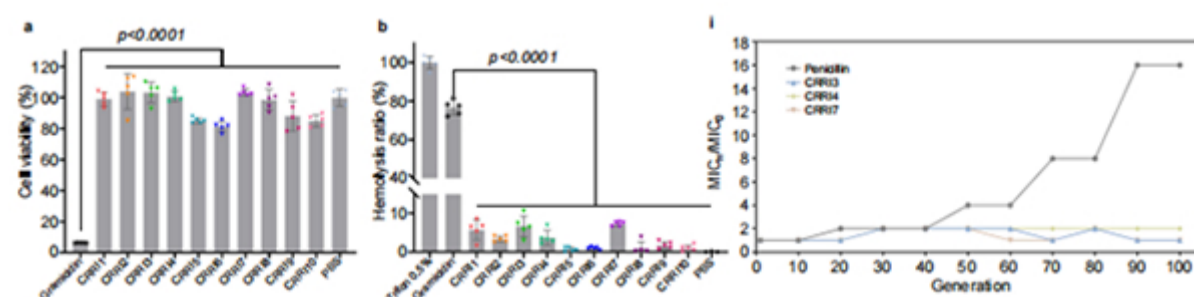


图2 抗菌六肽体外毒性、细菌耐药性测试

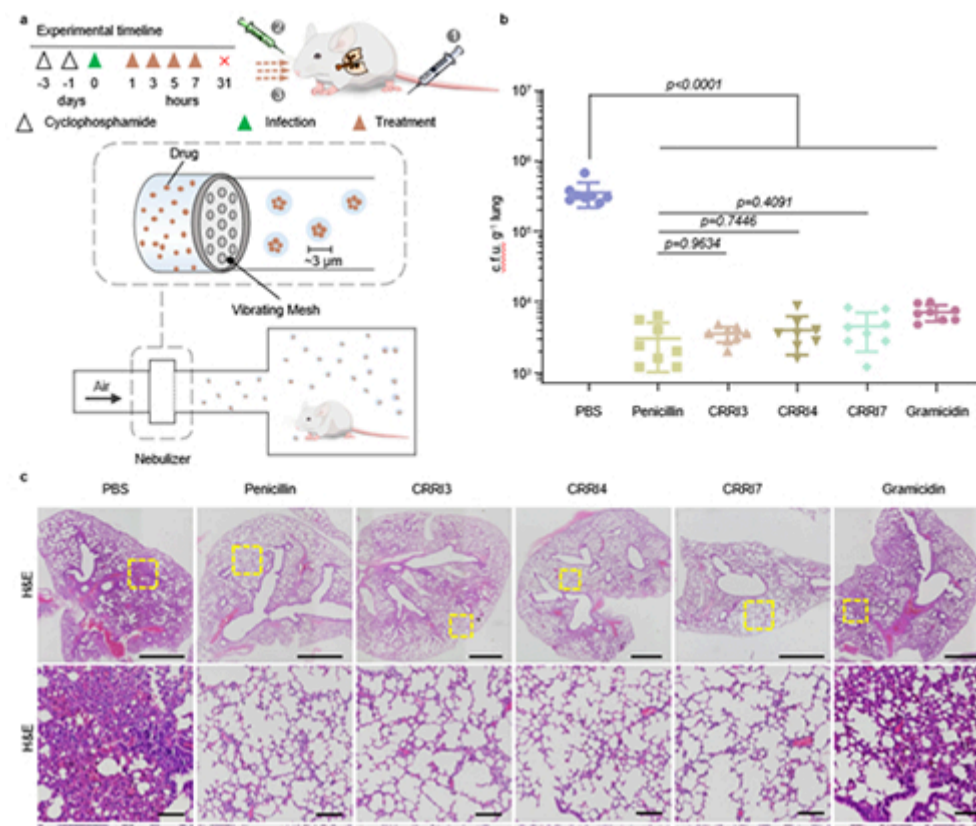


图3 抗菌六肽体内治疗效果

在国家自然科学基金项目（批准号：51933009）资助下，浙江大学高分子科学与工程学系计剑教授课题组在数据驱动的抗菌肽虚拟筛选领域取得进展，相关成果以“机器学习组合模型全肽库挖掘高效抗菌肽（Identification of potent antimicrobial peptides via a machine-learning pipeline that mines the entire space of peptide sequences）”为题，于2023年1月12日在线发表于《自然·生物医学工程》（Nature Biomedical Engineering）杂志。论文链接：<https://www.nature.com/articles/s41551-022-00991-2>。

具有序列多样性的多肽分子构成了一个近乎无限的潜在药物储藏库。其中，抗菌多肽是应对耐药细菌感染的潜在手段。然而针对抗菌多肽，当前最先进的实验技术也仅实现了在80万序列文库中的筛选。如何构建全新的研究模式，实现对抗菌肽全文库高效筛选，是该领域面临的重要挑战。

针对抗菌肽全文库高通量筛选的难题，研究者采用创新的人工智能框架，结合经验判断、分类、排序和回归任务组成的全新机器学习组合模型（图1），在极短时间内实现了对六肽全库6400万序列的识别。抗菌实验证实，从中筛选出的前10条肽均表现出了较高的抗菌活性，并且抗菌活性最强的3条抗菌肽均超越了现有文献报道的最强抗菌六肽。在未改动任何模型参数的前提下，该模型框架成功在极短时间内完成了对含有12.8亿序列的七肽库、256亿序列的八肽库及5120亿序列的九肽库的抗菌肽筛选，成功率高达98.2%（54/55）。作者对其中3条代表性的抗菌六肽进行了实验研究，其对多种临床分离耐药菌表现出了高杀菌活性（表1）、低脱靶毒性、不易诱导耐药性（图2），在小鼠细菌性肺炎模型中表现出了良好的治疗效果（图3）。

该项工作建立的组合式机器学习模型，为分步实现生物功能肽的全库优选提供了崭新的模式，有望进一步拓展到包括抗菌肽、组织再生肽和靶向肽的高效筛选中。为高通量实验和数据挖掘双轮驱动的生物医用材料研究模式提供了有效工具。

机构概况: 概况 职能 领导介绍 机构设置 规章体系 专家咨询 评审程序 资助格局 监督工作

政策法规: 国家科学技术相关法律 国家自然科学基金条例 国家自然科学基金规章制度 国家自然科学基金发展规划

项目指南: 项目指南

申请资助: 申请受理 项目检索与查询 下载中心 代码查询 常见问题解答 科学基金资助体系

共享传播: 年度报告 中国科学基金 大数据知识管理服务平台 优秀成果选编

国际合作: 通知公告 管理办法 协议介绍 进程简表

信息公开: 信息公开制度 信息公开管理办法 信息公开指南 信息公开工作年度报告 信息公开目录 依申请公开

[相关链接](#)

政府

新闻

科普



中华人民共和国
中央人民政府网站

版权所有：国家自然科学基金委员会 京ICP备
05002826号

地址：北京市海淀区双清路83号 邮编：100085

京公网安备 11040202500068号



事业单位



政府网站
找错

