

肠道耐药菌克星如何“万里挑三”

● 本报记者 冯丽妃

人体“宇宙”中居住着数百万种微生物，它们相生相克、相互制衡，维持着人体健康的微生态平衡。

很多常见耐药菌，如金黄色葡萄球菌、肺炎链球菌、大肠杆菌等都是与人体共生的“原住民”，而人体中同样存在着克制它们的微生物及其代谢产物。找到这些耐药菌的“克星”作为药物，是当前生命科学领域的研究前沿。

如何在数以亿计的人体微生物中“找出”这些耐药菌“克星”？面对这一难题，中国科学院微生物研究所研究员陈义华和王军团队合作，利用机器学习的方法，从1万多个肠道微生物组中“海淘”出3个具有显著活性和安全性的抗菌肽“潜力股”。相关研究近日发表于《自然—生物技术》。

“整个研究令人振奋！从人体肠道微生物组挖掘抗菌肽的方法十分新颖。”该刊一位审稿人评论说。

人工智能“海淘”奏效

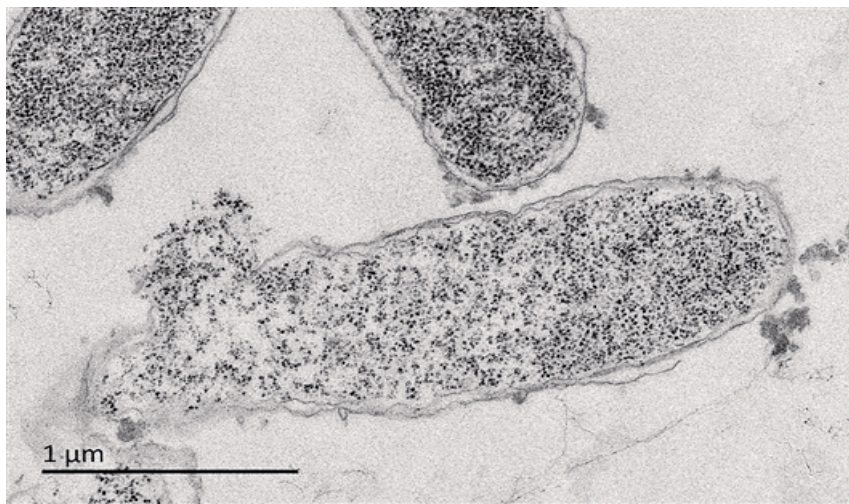
长期以来，微生物产生的活性代谢产物是开发抗菌药物的重要资源。青霉素、头孢菌素、红霉素、万古霉素、达托霉素……临床使用的抗菌药物中，超过一半来源于微生物。

科学家认为，人体肠道中庞大的微生物“原住民”蕴含着与耐药病原菌共生的“克星”，它们的代谢产物能够抑制耐药菌的活性。

抗菌肽就是这样一类具有抗微生物活性的小肽（序列在几个到几十个氨基酸残基之间）。以往研究发现，它们经过简单的后修饰或不经任何修饰就能直接发挥作用，达到抑制病原微生物的效果。能抵御多重耐药菌同时不易产生耐药性的抗菌肽，被认为是可能替代传统抗生素的下一代抗菌剂。

近年来，由于测序技术的发展涌现出海量微生物组数据，为人们寻找活性肽提供了丰富的数据资源。但挖掘方法的落后阻碍了相关研究的步伐。

“人体肠道微生物互作网络错综复杂，存在许多序列短、多样性高、相似性低的功能多肽，怎么区分它们是个难题。”文章共同通讯作者王军向记者



通过人工智能挖掘到的抗菌肽分子（10倍最低抑菌浓度）处理后的大肠杆菌。
论文作者供图

解释，人工智能深度学习方法为这一难题的解决提供了保障。

研究团队采用自然语言处理（NLP）的多种神经网络方法，把氨基酸序列当做“语言”，训练人工智能学习现有抗菌肽的组织方式，从而让其根据“经验”识别氨基酸短序列，区分相似的多肽。通过选取三种NLP神经网络模型来预测鉴定，他们实现了抗菌肽挖掘模型的构建和优化。在测试集中，该模型的精确度达到91.31%。

有了模型，接下来就是放开手脚挖掘数据——研究团队在1万多个微生物组中“海淘”，挖掘并合成了216种潜在的新型抗菌肽。经实验验证，其中有181种新型抗菌肽具有抗菌活性（占83.8%）。进一步实验表明，抗菌肽对多重耐药革兰氏阴性菌具有较强的抑制能力。

但这些抗菌肽是否新颖？对此，王军表示：“它们绝大多数与已知抗菌肽序列同源性低于40%。而且最终选取的抗菌肽作用机制不单一，表明我们的方法没有特定的偏好性。”

层层筛选“万里挑三”

新抗菌肽拿到了，接踵而至的是三个关键问题：它们的抗菌活性如何？作用机理是什么？是否安全可靠？

首先，陈义华团队对181种抗菌肽的活性逐一测试，并在此基础上选取了其中最有效的11种进行了新一轮测试。其测试对象包括粪肠球菌、金黄色葡萄球菌、肺炎克雷伯菌、鲍曼不动杆菌、绿脓杆菌、阴沟肠杆菌和大肠杆菌

等病原细菌。

“我们发现，这些抗菌肽只需要少许剂量就能实现多重耐药菌的显著抑制能力。”论文共同通讯作者陈义华向记者介绍，11种抗菌肽对单个菌的最低抑菌浓度值大部分小于每升50微摩尔（即百万分之五十摩尔）。在针对临床分离得到的鲍曼不动杆菌、肺炎克雷伯菌和大肠杆菌等多重耐药菌活性测试中，有10种抗菌肽显示出强抑制能力。

为了解其作用机理，研究者通过对11种抗菌肽进行测试，确定了其中部分肽的作用位点分别为细胞膜和细胞壁。“如果用电子显微镜观察，可以看到与抗菌肽共同培养的大肠杆菌细胞膜明显破裂。这样细胞内容物外流，耐药微生物就被杀死了。”陈义华解释说。

据介绍，经过长达30天的连续传代培养，研究者并未观察到大肠杆菌产生任何抗性。另外，他们还发现这11种优质抗菌肽中有超过一半的小肽由α-螺旋结构主导，这种小肽或具有更强的抑菌活性。

通过溶血实验（红细胞破裂溶解）和细胞毒性实验，研究人员进一步从11种抗菌肽中筛选出3种溶血性和细胞毒性相对较低的小肽。小鼠模型实验表明，经过这3种小肽治疗后的小鼠比未进行任何治疗的小鼠体重恢复速率明显加快；同时，治疗组的肺部组织病原菌残留量比非治疗组下降了一至两个数量级。验证了这三类多肽的有效性和安全性。

“这项工作非常有意义。它利用机器学习的方式挖掘新颖的抗菌肽，然后通过细致的微生物学测试研究它们的功能，对相关领域有启发作用。”另一位审稿人评论说。

多肽类药物迎爆发期

当前，传统新药开发困难重重，从候选小分子出发，要经过多轮筛选，每一步都存在较高的失败率。一个药物投入几十亿美元最后折戟而归也很常见。

很多人常用“十年十亿美金”来形容一款新药的开发成本。美国塔夫茨大学药物发展研究中心2016年的研究发现，最终获美国食品与药物管理局批准的每种创新药物总平均成本或高达25.58亿美元。

“如果一开始候选分子的范围比较窄，失败的风险和投入的成本会更高。”王军说，“通过深度学习可以更快、更高效地确定更多的候选分子，在此基础上一步步筛选，极大缩减了时间和实验成本。”

他表示，理论上多肽药物可选择的序列是“天文级”的，远超ZINC15数据库（目前最大的有机小分子化合物库之一）收集的1亿余种化学类分子总量。然而，因为开发利用效率低，过去40年里获批的多肽药物数量十分有限。

随着机器学习的深入发展，多肽类药物或将迎来爆发期。生物学家可以通过海量的基因组数据，大量发现乃至主动设计、从头生产自然界不存在的活性肽。

陈义华表示，在此次研究中，微生物在复杂的人体环境中产生抗菌肽的巨大潜能，为从微生物组数据中直接挖掘活性代谢产物提供了证据和范例。

作者表示，目前的研究仅通过动物实验证明部分多肽的有效性和安全性，从而提供了一个成药的候选分子或前体分子。接下来需要进行临床前的成药性等评价，挖掘相关分子在农业和畜牧业方面的潜在应用价值。

相关论文信息：<https://www.nature.com/articles/s41587-022-01226-0>