

阿里达摩院研究团队发布细胞世界模型“灵枢细胞” 硅基细胞“数”造生命

■ 本报记者 杨千莹

“AI虚拟细胞”，一个在计算机中生长与消亡的硅基生命单元。

想象一下这样一幅画面：轻点几下鼠标，几秒钟内，计算机上便出现一个跳动的细胞，演示其在给定条件下如何变化。与传统实验室内耗数月甚至数年培养细胞、观察药物反应相比，这一幕颇为科幻的图景已经出现在真实世界。

过去两年，“AI虚拟细胞”概念在生物医学界迅速升温，美国扎克伯格基金会、深度思维、中国科学院、西湖大学等科研机构纷纷入场。

不久前，阿里达摩院研究团队发布了全球首个全转录组尺度的细胞世界模型——“灵枢细胞”(Lingshu-Cell)。这个在业界竞赛中性能领先的生成式模型，能够精准预测细胞对基因扰动和药物的反应，为我们勾勒出从细胞层面构建“数字分身”的可能性。

AI，如何理解并参与设计一个像细胞一样时刻变化的复杂系统？它又将给生物医学带来哪些改变？

让细胞“动起来”

要理解“灵枢细胞”能做什么，首先要了解“AI虚拟细胞”。简单来说，它的目标是在计算机中建立一个细胞的“数字分身”，并由AI模型驱动，预测当细胞遭遇“扰动”时会发生什么。

这个“扰动”，可以是某个基因的敲除、药物分子的刺激，或是细胞因子的作用。通过单细胞测序技术，研究人员能获取细胞遭遇“扰动”前后数以百万计的基因表达数据，并基于此寻找治愈疾病或开发新药的方法。

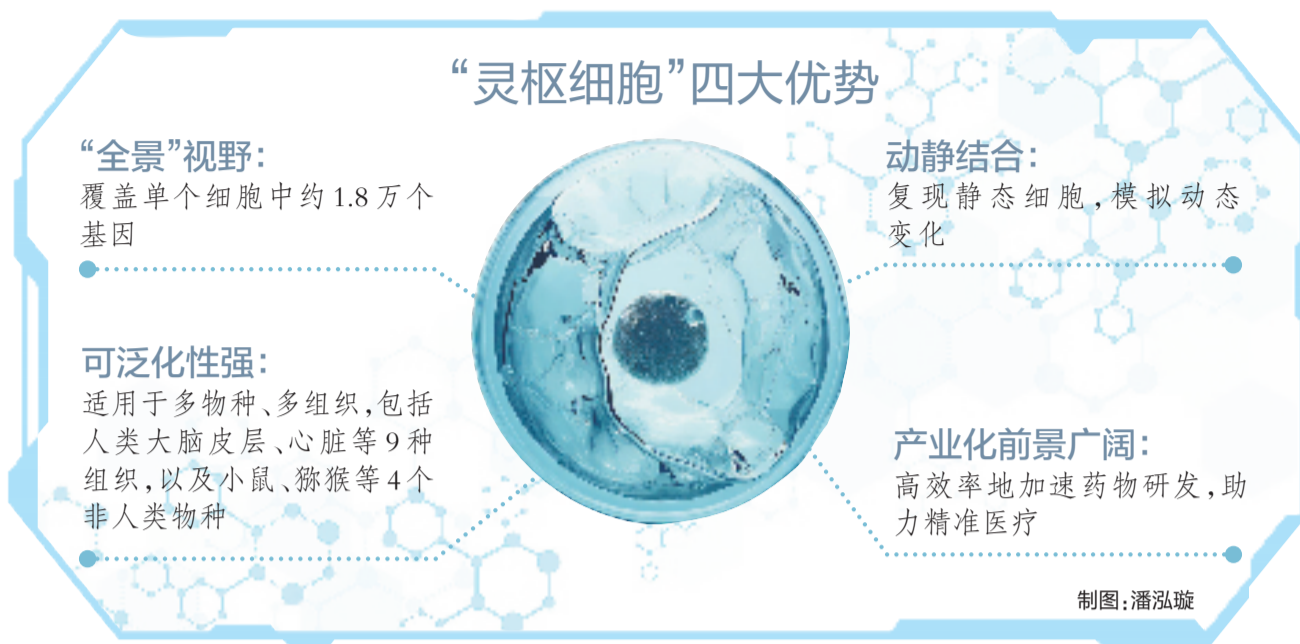
从前的细胞模型大多只能模拟细胞在某一时刻的“定格”样态。然而对于复杂的生命系统而言，这远不足以囊括其变化规律。

“灵枢细胞”的研究动机主要是围绕“把握细胞的动态规律”这一核心目标展开。

阿里达摩院的屏幕上，三块分区清晰可见：扰动前细胞、选定的扰动、扰动后细胞。

“灵枢细胞”研究团队成员张晗现场为记者演示了模型的预测过程：选定扰动的基因或细胞因子，点击“开始模拟”，几秒钟后，屏幕右侧便生成了一个动态的细胞，展示AI模型预测结果。

“这个过程，有点像教AI用基因玩一个填空游戏。”张晗介绍，首先，团队将单个细胞中约1.8万个基因及其表达转换成词元(token)。接着，随机遮盖其中一部分基因，让模型根据学习到的规律，预测基因的表达结果，亦即细胞的状态



分布。在模型充分学习上述静态规律之后，团队叠加上另一个模型，建立起一个更大的词表，从而使得“灵枢细胞”具备动态模拟预测能力。

“这个叠加的模型，就是掩码离散扩散模型。”“灵枢细胞”研究团队负责人、阿里达摩院科学智能负责人荣钰介绍，这也是该模型首次被应用于AI虚拟细胞的研究。

这一叠加并非简单的拼接，而是系统化的改造。“如ChatGPT这样的语言大模型，是基于上下文语境展开预测，而单细胞基因数据的特殊之处，在于它的位置任意性。”荣钰解释，单细胞中的基因表达数据可能以任意顺序出现，具有高度离散的特征，不符合语言大模型“高频词”和“从左到右”的预测逻辑。因此，该模型采用的离散扩散生成方式，更贴合基因表达“天然无序”的内在结构。

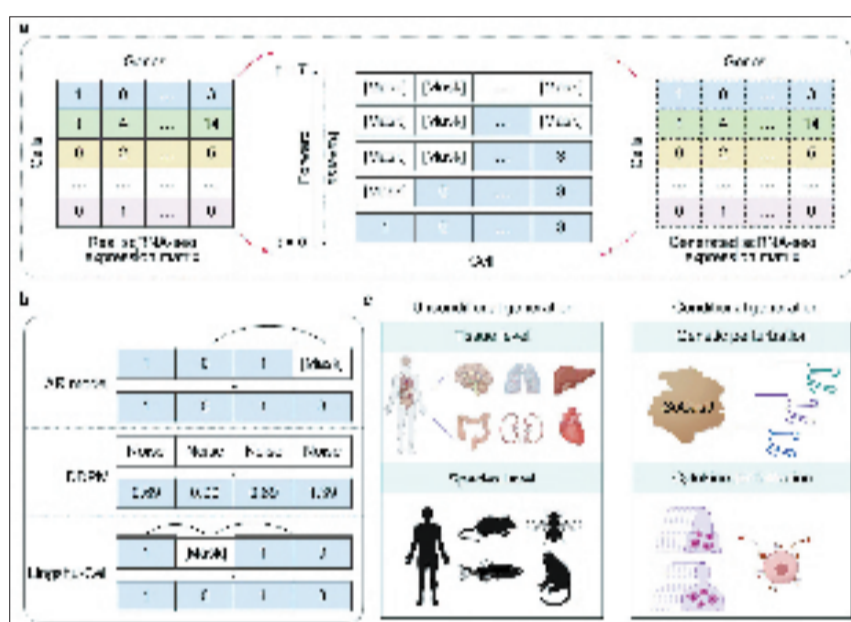
验证人类大脑皮层、心脏等9种组织，以及小鼠、猕猴等4个非人类物种数据后的结果表明，“灵枢细胞”的生成结果已经与真实细胞高度一致。

预演“多元宇宙”

“AI虚拟细胞”看似新颖，实际上完成的是一个未竟的任务。

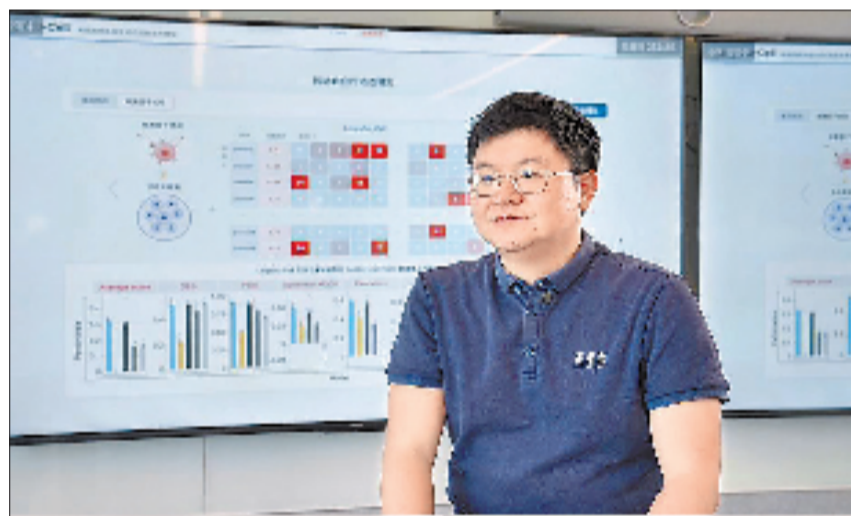
在AI出现以前，用计算机模拟细胞并非新鲜事。早期的虚拟细胞建模实践，可以追溯到20世纪90年代，美国康涅狄格大学教授莱斯利·洛(Leslie M. Loew)团队开发的“虚拟细胞”(Virtual Cell)计算平台。

当时，研究者们尝试用数学公式描述细胞内部的信号通路——某个蛋白质



“灵枢细胞”模型架构。

受访者供图



团队负责人荣钰，背后为“灵枢细胞”模型示意图。

受访者供图

激活了哪条通道，另一个分子又如何反馈抑制。这种思路在逻辑上是清晰的：如果能把细胞里所有的分子关系都写成方程，就能推算细胞的行为。

问题在于，这个“如果”的条件太过苛刻。

“一个哺乳动物细胞中蛋白质分子的平均数量多达数十亿，与现有地球人口规模处于同一量级。”西湖大学医学院及生命科学学院院长特聘教授郭天南说，这样庞大的体系，意味着用数学公式“完整描述”细胞的努力往往会走向还原论，即把复杂系统拆解成少数关键组成部分来理解。也就是说，研究者只挑选最重要、最关键的几个分子来研究，而将大量其他因素暂时忽略。

“这样的局部可以做得非常精细。”郭天南给出了一个形象的比喻，“就像你绘制了一张北京或者波士顿的超高精度地图，这在特定场景下很有价值。但真实细胞里有数十亿个蛋白质分子，模型却往往只保留三四个关键分子。这样一来，就好比拿着这张地图去环球旅行一样，不具备泛化性，离实际的药物治疗与疾病研究非常遥远。”

AI的介入，改变了这个逻辑的起点。作为首个全转录组尺度的细胞世界模型，“灵枢细胞”把研究范围从少数几个基因扩大到细胞中所有RNA，用全景的视野生成预测。

传统方法依赖“人先理解规律，再写出公式”。AI的介入，更像是在海量数据中先观察生命，再从中学习规律。最直观的变化，就是成本降低、效率提升。

“这一能力最直接的应用之一，就是在制药领域。”荣钰说。在传统新药研发中，一个候选药物从实验室走向临床，往往需要经历细胞实验、动物实验、人体临床等多个阶段，仅前期筛选就需要耗费大量时间与成本。“不断调整药物浓度、组合不同分子，再观察细胞是否死亡、有没有毒性、是否出现耐药反应等，每一步都非常耗时。”

而在计算机里完成的虚拟实验，大幅节省了初步筛选的时间。在荣钰看来，这有点像电影《奇智博士》中预演“多元宇宙”的过程：“先在虚拟细胞中快速模拟一种药物可能带来的影响，提前筛掉大量无效方案，再把最有潜力的结果拿去做真实实验。”

须臾之间，曾经耗费几个月才显现的无数种可能的“未来”便能展现眼前。

与此同时，真实世界中的生物实验，本身存在难以复现的问题。采访中，荣钰分享了一个学界流传的段子：“今天我左脚先进门，培养细胞不开心了，一天的实验都没成功。”玩笑中，藏着生物实验苛刻的培养条件。

荣钰说，即便是同一种细胞、同一种培养条件，不同实验室、不同时间做出来的结果都可能不一样。细胞的衰老、变

异以及实验环境都会带来影响。

AI介入的意义，正在于从这些庞杂、充满噪声的数据中，寻找更稳定的规律。

走向真实的生命

多位研究者表示，“AI虚拟细胞”仍处于发展的早期阶段。

“‘灵枢细胞’初代模型的泛化能力，目前仅存在于现存的细胞类型范围内。”荣钰说，这意味着，如果出现从未遇见过的细胞类型，模型就难以精准地做出预测。

换言之，现阶段，AI的能力局限于在已知范围内进行精确推演。荣钰透露，“灵枢细胞”的下一步，将加入更多基因组学信息、更复杂的扰动条件，并依据真实场景中的困难情况不断提升泛化能力，朝着更加接近真实生命的方向迭代。同时，类器官水平的研究也在进行中，预计在今年发布成果。

在郭天南看来，“AI虚拟细胞”最终的发展方向，是逐步从细胞扩展到组织、器官，最终走到人体。但实现这一目标，前提仍是先真正理解细胞。“虚拟细胞是承上启下的一环，只有模拟好细胞，才有可能往上构建组织器官，最终构建人体的数字孪生。”

那么，“AI虚拟细胞”是否仍处于实验室的起步阶段？专家们纷纷否定了这个说法。

“‘AI虚拟细胞’天然地带有应用属性。”郭天南说，与很多基础科研不同，“AI虚拟细胞”从诞生开始，就几乎直接面向制药、医疗与生物工业，因此，模型的发展，不仅在高校与研究所以发生，也在产业界出现。

不久前，由诺奖得主联合创立的赛拉治疗公司推出虚拟细胞模型X-Cell，参数规模高达49亿个。此外，全球科技公司也在相继布局相关研究。

去年4月，美国食品药品监督管理局宣布，将使用“新方法”(包括AI计算模型、人类细胞系、类器官等)逐步减少并替代部分传统动物实验。这对整个医药行业都是历史性的时刻。

“在动物模型中有效的药物，在人体中不一定表现出相同的疗效或安全性。”郭天南说，这一在药物研发中沿用了100多年的动物实验传统正因为“AI虚拟细胞”等新技术的崛起而开始松动。

“关于我们关于细胞如何运作的知识，最终或许都将被整合进这些模型中，以实现更全面的预测。”采访中，郭天南描述了一个愿景。这既是生物医学领域的目标，也是它还需翻越的一座座山头。我们期待着，这个在硅基世界里生存的小小“生命”，变得越来越真实。

“病从口入”新解：牙病连全身

■ 本报记者 朱平 通讯员 叶鹏

古语有云：“病从口入”。科学家正为它写下全新注解。

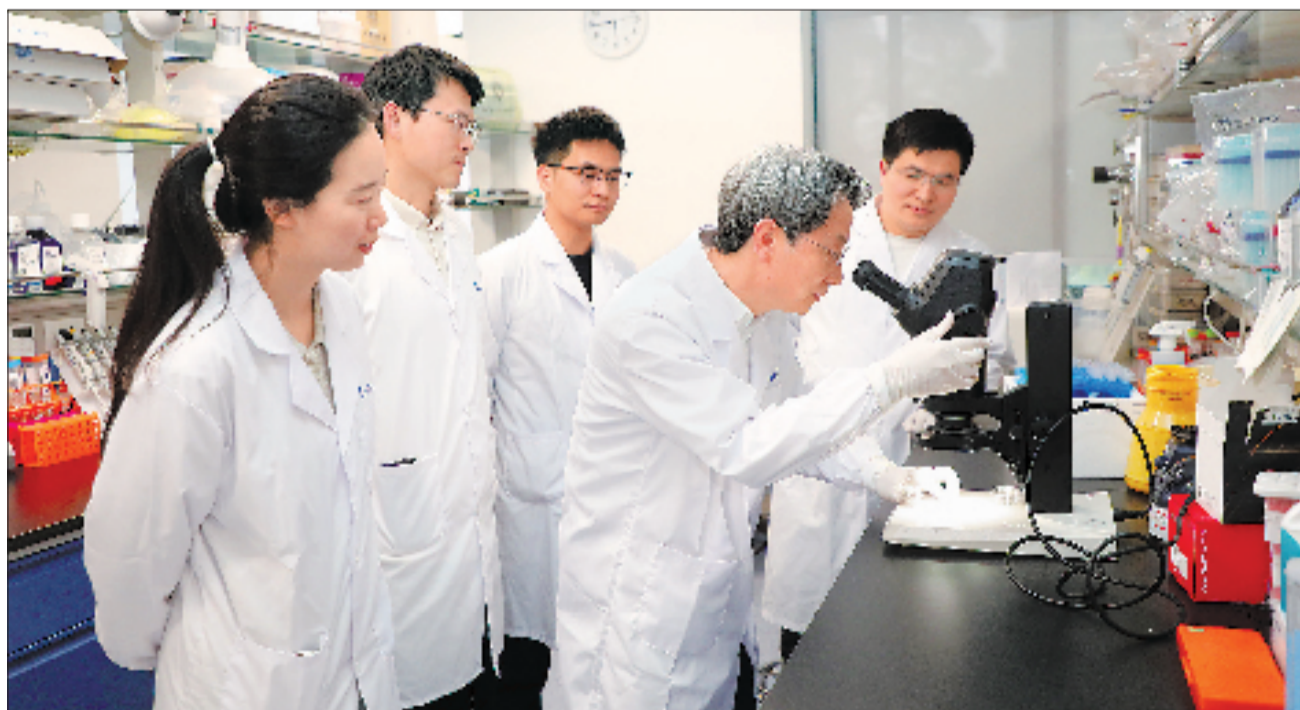
近期，浙江大学医学院附属口腔医院党委书记段胜仲教授团队连续在《先进科学》《循环》等国际知名期刊发表论文，证明牙周炎并非简单的口腔局部感染，而是一个通过重塑全身免疫系统，系统性加重全身慢性病的“隐形炎症开关”。

牙病伤肝又伤心

牙周炎有多普遍？段胜仲给出一个惊人的数字：全世界范围内，牙周炎的患病率比肝炎、关节炎等任何炎症性疾病都要高，近半数的成年人多多少少都有不同程度的牙周问题。然而，大多数人可能将它等同于“上火”，从未想过它会影响远处的器官。

传统观念里，口腔与内脏之间隔着千山万水。即便有研究指出牙周炎患者的心血管疾病风险更高，医学界也长期存在争议：这到底是因果关系，还是仅仅因为两类疾病恰好“伴发”？为了弄清真相，段胜仲团队决定用实验来回答。

在最新的针对2型糖尿病的研究中，团队为一部分肥胖合并糖尿病的实验动物“制造”牙周炎，以此模拟人类的慢性牙周损伤。结果显示：同样是高脂饮食“喂”出肥胖的动物，那些并发牙周炎的个体糖尿病病情显著加重。



段胜仲(右二)和团队成员在进行实验。

受访者供图

进一步分析发现，牙周炎的“攻击靶点”是肝脏。它并未加重脂肪或肌肉组织的胰岛素抵抗，而是特异性地扰乱肝脏的糖代谢功能。

通过一系列免疫追踪，研究团队揭示了一条“口腔—肝脏轴”：牙周炎发生后，肝脏中的巨噬细胞被异常激活，大量释放名为“白介素-18”的炎症因子；这种因子随血液循环，刺激脾脏及其他免疫器官，促使一类名为“B2细

胞”的免疫细胞大量增殖。这些增生的B2细胞反过来又会刺激肝脏巨噬细胞释放更多的白介素-18，形成恶性循环。

最终，肝脏被“淹没”在炎症中，糖代谢紊乱加剧，胰岛素信号通路严重受损，进而显著加剧糖尿病的严重程度。

在更早前的另一项追踪中，段胜仲团队还发现，被细菌劫持的颈部淋巴结

中的免疫细胞直接充当了细菌，袭击心脏的“内应”。

研究显示，合并牙周炎的心肌梗死动物心脏泵血功能显著更差，梗死面积更大。最关键的是，他们利用基因测序技术及原位杂交技术在梗死的心脏组织中直接检测到了来自口腔的细菌DNA和完整细菌。而没有心梗的动物心脏中并不存在这些细菌。这说明，心肌梗死造成的局部损伤，为口腔细菌的“定

居”敞开了大门。

破译致病密码

究竟是哪些细菌有能力完成这场“远程袭击”？研究团队将从重度牙周炎患者口腔中采集的全部菌群，移植到同时患有心肌梗死和牙周炎的实验动物体内，然后通过高通量测序，找到出现在梗死心脏中，并与病情加重显著相关的细菌。

最终，五种攻击心脏的“细菌刺客”浮出水面：牙龈卟啉单胞菌、具核梭杆菌、中间普雷沃菌、生痰月形单胞菌和普通拟杆菌。这五种细菌有一个共同特征：它们都携带破坏细胞屏障的工具，为自己打开从口腔进入血液的通道。

抵达心脏，只是开始。这些致病菌还会劫持人体的免疫系统。通过巧妙的“细胞追踪技术”，研究团队标记了实验动物颈部淋巴结中的免疫细胞。他们惊讶地发现，在牙周炎和这五种致病菌存在的情况下，颈部淋巴结中的大量B2细胞被“动员”起来涌入血液，最终聚集到梗死的心脏组织中。这些B2细胞并没有产生更多的抗体，反而被细菌“策反”，大量分泌促炎因子，加剧了心脏的炎症和心肌梗死。

研究团队破解了这套“远程调度系统”的导航密码：细菌一方面驱使免疫系统中重要的B细胞离开淋巴结进入血液；另一方面在梗死心脏中释放“导航信号”，引导这些B细胞精准抵达梗死部位

发起攻击。

为验证这一切并非动物实验中独有的现象，段胜仲团队将实验室的发现“带回”临床，与心血管学联合，发现在心肌梗死患者中，合并牙周炎的人心脏病功能更差，血液中有毒B细胞比例更高。而他们口腔中的牙龈卟啉单胞菌、具核梭杆菌、中间普雷沃菌、生痰月形单胞菌这四种致病菌的丰度，与病情的严重程度呈显著正相关。

这些发现正推动“共病同治”医学理念的形成与发展。

过去，牙周炎归口腔科，糖尿病归内分泌科，心肌梗死归心内科，患者在问诊中容易陷入“头痛医头，脚痛医脚”的误区。临床医疗也常“分科而治”，没有探究它们之间可能存在的共同病理基础。

段胜仲认为，现有的研究表明，针对糖尿病或心血管疾病患者，积极治疗牙周炎可能就是一种有效的辅助治疗手段。早前，欧洲心脏病学会与世界心脏联盟也已达成专家共识，建议所有心血管病患者都应去看牙周科。

目前，全球还没有相关指南。段胜仲团队已经将目光投向具体干预方法的探索。他们研发了一种可以局部治疗牙周炎的不含抗生素的水凝胶。实验证明，牙周炎合并高血压的动物血压得到了有效控制，合并心肌梗死的动物症状得到有效缓解。此外，他们还发现健康人唾液中的黏蛋白和铁离子可以促进特定肠道益生菌的富集，能够帮助降低血压。这为未来的益生菌疗法打开了想象空间。