

“小柯”秀

一个会写科学新闻的机器人

《自然 - 生物技术》

深度学习预测

先导编辑效率和产物纯度

瑞士苏黎世大学的 Gerald Schwank 和 Michael Krauthammer 合作, 研究通过深度学习预测先导编辑的效率和编辑产物的纯度。相关成果 1 月 16 日在线发表于《自然 - 生物技术》。

先导编辑是一种通用的基因组编辑工具, 但需要对先导编辑器的引导 RNA(pegRNA) 进行实验优化以实现较高的编辑效率。

研究人员进行了高通量筛选, 以分析 92423 个 pegRNA 对 13349 个人类致病性突变(包括碱基替换、插入和缺失)的主要编辑结果。基于该数据集, 研究人员识别出影响先导编辑的序列背景特征, 并训练了 PRIDICT(先导编辑 pegRNA 预测工具), 这是一个基于注意力的双向递归神经网络。PRIDICT 可靠地预测了所有小规模遗传变化的编辑效率, 对于靶标和非靶标的编辑, Spearman R 分别为 0.85 和 0.78。

研究人员在内源性编辑位点和外部数据集中验证了 PRIDICT, 并显示 PRIDICT 评分高(>70) 和低(<70) 的 pegRNA 在体外不同细胞类型(12 倍)和体内肝细胞(10 倍)中编辑效率显著提高, 突出了 PRIDICT 在基础和转化研究应用中的价值。

相关论文信息:

<https://doi.org/10.1038/s41587-022-01613-7>

《自然 - 遗传学》

增强子与启动子相互作用
绕过 CTCF 介导边界

美国国立卫生研究院的 Pedro P. Rocha 调研组发现, 增强子与启动子的相互作用可以绕过 CTCF 介导的边界, 并有助于表型的稳健性。相关论文 1 月 30 日在线发表于《自然 - 遗传学》。

研究人员表示, 增强子如何激活其远端靶标启动子仍不完全清楚。

研究人员剖析了 CTCF 介导的环化如何促进和限制这种调控相互作用。通过利用小鼠突变体的一个等位基因系列, 研究人员表明 CTCF 既不是 Sox2 基因与远端增强子相互作用所需的, 也不是其表达所需的。在 Sox2 及其远端增强子之间插入 CTCF 基序的各种组合产生了具有不同绝缘程度的边界, 这与转录输出的减少直接相关。然而, 在外胚层和神经组织中, 增强子接触和转录诱导都不能完全消除, 并且插入也不能破坏着床和神经发生。

相比之下, 在携带最强边界的突变体前肠中检测不到 Sox2 的表达, 这些动物在该组织中完全表型地丢失了 Sox2。研究人员认为具有高密度调控活性的增强子簇可以更好地克服物理障碍, 以保持忠实的基因表达和表型一致性。

相关论文信息:

<https://doi.org/10.1038/s41588-022-01295-6>

《自然 - 地球科学》

大规模火山活动触发
白垩纪中期海洋酸化事件

美国西北大学 Jones Matthew M 调研组在研究中取得进展。他们报道了由大规模火山活动引发的白垩纪中期海洋酸化的突然发生。相关研究 1 月 19 日发表于《自然 - 地球科学》。

白垩纪中期的大型火成岩省火山活动引发了大约 9450 万年前全球范围内海洋氧水平下降的事件, 即海洋缺氧事件。有人假设, 火山二氧化碳在地质上的快速脱气改变了海水碳酸盐的化学性质, 影响了海洋生态系统、地球化学循环和沉积。

课题组报告了国际海洋发现计划在澳大利亚西南海域钻深的两个地点, 这些地点以缺乏碳酸盐矿物的地层间隔的形式, 展示了受抑制的远洋碳酸盐沉积的明确证据, 记录了事件期间的海洋酸化。然后, 研究组将大块沉积物中镁同位素的组成与火山活动开始时, 浅滩中约 600 kyr 海相方解石补偿深度的延长联系起来。

相关论文信息:

<https://doi.org/10.1038/s41561-022-01115-w>

更多内容详见科学网小柯机器人频道:
<http://paper.scientenet.cn/Alnews/>

刚成立一年多,
这家中心为何受联合国青睐?

(上接第 1 版)

发展前景如何?

卫星星座、平台共享、科技服务……

面向未来, 郭华东介绍, 在 SDGSAT-1 基础上, 面向全球可持续发展的空间观测需求, 研究中心正在规划和论证可持续发展系列科学卫星, 构建形成可持续发展卫星星座, 为 2030 年议程和全球发展倡议实施提供定制化、体系化数据支撑服务。

研究中心还启动了“CBAS 全球 SDGs 伙伴关系计划”, 2022 年, 包括联合国人类住区规划署(联合国人居署)、芬兰赫尔辛基大学、泰国亚洲理工学院、北京大学、南京大学在内的 16 家中外机构成为研究中心合作节点, 为科研发、科学数据共享、行业应用与技术研发、国际合作与能力建设四类合作活动提供平台, 助力联合国可持续发展目标的实现。

与此同时, 研究中心还通过大数据平台, 服务于 17 项可持续发展目标落实的分析、评估和监测工作。

“2023 年 3 月底, 落实 2030 年可持续发展的进程就要过半, 当前中国范围内与环境相关的可持续发展目标指标中, 接近一半已经达标, 可以说中国提前 7 年半实现了相关目标。”郭华东说。

他表示, 从全球范围看, 联合国 2030 年可持续发展议程的落实进程依然受限。在全球发展倡议框架下, 研究中心将持续利用科技创新和大数据应用手段, 特别是面向发展中国家紧迫的发展需求, 聚焦消除贫困、零饥饿、气候行动等多项 SDGs 指标, 为落实联合国 2030 年可持续发展议程提供数据支撑和科技服务。

AI 首次成功从零生成原始蛋白质

本报讯 科学家创建了一个能够从头开始生成人造酶的人工智能(AI)系统。在实验室测试中, 尽管人工生成的氨基酸序列与任何已知的天然蛋白质存在显著差异, 但其中一些酶与自然界中发现的酶一样有效。相关研究成果 1 月 26 日发表于《自然 - 生物技术》。

该研究表明, 虽然自然语言处理是为读写语言文本开发的, 但至少可以学习一些生物学的基本原理。Salesforce Research 公司开发了名为 ProGen 的 AI 程序, 使用下一代标记预测将氨基酸序列组装成人造蛋白质。

科学家表示, 这项新技术可能比获得诺贝尔奖的蛋白质设计技术——定向进化更为强大, 它将加速新蛋白质的开发, 为已有 50 年历史的蛋白质领域注入活力。这些新蛋白质几乎可以用于从疾病治疗到降解塑料的任何领域。

“人工设计的性能比受进过程启发的设计表现得更好。”该研究作者之一、美国加州大学旧金山分校药学院生物工程和治疗科学教授 James Fraser 表示, 语言模型正在学习进化的各个方面, 但它不同于正常的进化过程。“我们现在能够针对特定效果调整这些属性的生成, 例如非常耐热或喜欢酸性环境, 不会与其他蛋白

质相互作用的酶。”

为了创建这个模型, 科学家只需将 2.8 亿种不同蛋白质的氨基酸序列输入机器学习模型, 并让它在几周内消化信息。然后, 他们使用来自 5 个溶菌酶家族的 56000 个序列, 以及有关这些蛋白质的上下文信息对模型进行微调。

该模型迅速生成了 100 万个序列。研究团队根据它们与天然蛋白质序列的相似程度, 以及 AI 蛋白质的潜在氨基酸“语法”和“语义”的自然程度, 选择了 100 个序列进行测试。

在第一批由 Tierra 生物科学公司体外筛选的 100 种蛋白质中, 该团队制作了 5 种人工蛋白质用于细胞测试, 并将其活性与鸡蛋清中发现的一种酶(鸡蛋清溶菌酶, HEWL)进行了比较。在人类的眼泪、唾液和牛奶中也发现了类似的溶菌酶, 它们可以抵御细菌和真菌。

其中两种人工酶能够分解细菌的细胞壁, 其活性与 HEWL 相当。但它们的序列只有约 18% 相同, 这两个序列与任何已知蛋白质的相似性分别为 90% 和 70%。

天然蛋白质中的一个突变就能使其停止工作。但在另一轮筛选中, 研究小组发现, 即使只有 31.4% 的序列与任何已知的天然蛋白质相

似, AI 生成的酶仍显示出活性。

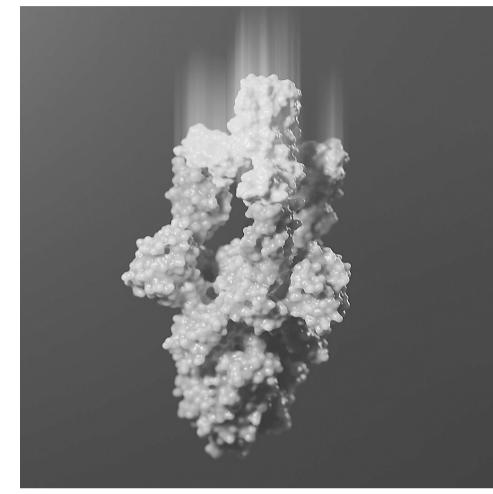
AI 甚至能够通过研究原始序列数据了解酶如何形成。X 射线晶体学测量显示, 人造蛋白质的原子结构看起来和它们应该有的样子一样, 尽管这些序列是前所未有的。

2020 年, Salesforce Research 基于研究人员最初开发用于生成英语文本的一种自然语言编程开发了 ProGen。他们从之前的工作了解到, 人工智能系统可以自学语法和单词的含义, 以及其他使写作井井有条的基本规则。

“当你用大量数据训练基于序列的模型时, 它们在学习结构和规则方面确实非常强大, 能够了解哪些词可以同时出现, 以及组合性。”该研究通讯作者之一、Salesforce Research 人工智能研究主管 Nikhil Naik 说。

对于蛋白质, 设计选择几乎是无限的。溶菌酶和蛋白质一样小, 最多约有 300 个氨基酸。但是有 20 种可能的氨基酸, 这样就有 20³⁰ 种可能的组合。鉴于无限的可能性, 该模型能够如此轻松地产生酶是非常了不起的。

该研究第一作者、Profluent Bio 创始人 Ali Madani 说:“从头开始生成功能性蛋白质的能力表明, 我们正在进入蛋白质设计的新时代。对



图片来源:IAN C. HAYDON

于蛋白质工程师来说, 这是可用的多功能新工具, 我们期待看到它的治疗应用。”(辛雨)

相关论文信息:

<https://doi.org/10.1038/s41587-022-01618-2>

■ 科学此刻 ■

食脑虫
有药了

2021 年夏, 当一名 54 岁男子疾病发作被送往美国北加利福尼亚州的一所医院时, 没人知道他的脑袋里有“虫子”肆虐。

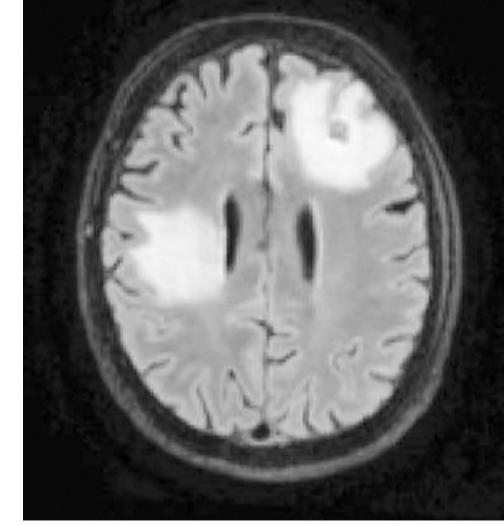
当时的核磁共振成像(MRI)显示, 该男子大脑左侧有个神秘肿块, 于是他被转移到加利福尼亚大学旧金山分校(UCSF)医学中心接受进一步治疗。然而, 脑活检组织检查和其他测试显示, 该男子脑部肿块并非肿瘤, 而是罕见的由巴氏阿米巴原虫引发的中枢神经系统感染。

巴氏阿米巴原虫是几种登上过新闻头条的“食脑虫”之一。这种阿米巴原虫 1986 年于圣地亚哥动物园的一只死亡待产母山魈体内被发现。该病原体可导致 90% 以上的感染者死亡, 是不折不扣的大脑“杀手”。

幸运的是, 上述患者顽强地活了下来, 并在药物治疗后基本康复。

之所以能够找到救命药, 源于 UCSF 医疗团队的努力。当他们在寻找挽救患者的一切方法时, 发现了一篇几年前 UCSF 生化学家 Joseph DeRisi 团队发表的论文。该研究指出, 一种最初在欧洲开发的用于抑制尿路感染的药物——硝羟喹啉, 在实验室中对治疗巴氏阿米巴原虫感染有效。于是, UCSF 医疗团队匆忙从国外找到这种药, 并首次将其用于阿米巴患者治疗。相关研究与治疗结果近日发表于《新发传染病》。

未参与该研究的专家表示, 这名男子的康复, 是长期以来被认为是“死刑”的阿米巴原虫脑部感染治疗的一个突破。



MRI 显示, 一名脑部感染巴氏阿米巴原虫的男子在接受实验药物治疗前(左)后炎症(白色部分)存在显著差异。

图片来源:N. SPOTTISWOODE ET AL., EMERG. INFECT. DIS.

UCSF 医疗团队表示, 该药物在美国尚未被批准常规使用, 但实验室测试显示它对其他致病性阿米巴原虫同样有效。

巴氏阿米巴原虫生活在土壤、灰尘和水中, 能通过肺部或皮肤切口进入人体。过去几十年中, 美国仅确认了约 100 例巴氏阿米巴原虫感染人类病例。

“绝大多数医生, 即便是传染病医生可能也从未见过相关病例。”UCSF 医疗团队领导者、传染病医师 Natasha Spottiswoode 说。因此, 两年前, 即使在上述患者 MRI 结果出来后, 他们也没有往阿米巴原虫感染上想, 直到活检结果疑似发现阿米巴原虫、相关聚合酶链式反应检测呈阳性后, 他们才确定遇到了这种罕见的感染。

最初, 医疗团队采用抗寄生虫、抗菌和抗真菌药物联合疗法, 患者一天就要服用 47 粒药物, 此外还要接受静脉注射治疗。这种疗法毒副

作用很大, 患者接受上述治疗后血糖水平和白细胞数目都降至危险水平, 并且出现了肾衰竭, 与此同时, 病变仍在不断扩大。直到 Spottiswoode 等人发现硝羟喹啉有望治疗该疾病时, 患者才有了一线生机。

Spottiswoode 联系到了 DeRisi, 但对方表示他们还没有机会使用硝羟喹啉在实际感染者身上进行检验, 且该药物没有获得美国食品药品监督管理局的批准。最终, 在多方协调下, 美国亚虹医药公司为 UCSF 医疗中心无偿提供了数百粒药片, 仅用硝羟喹啉治疗 1 周后, 患者的病变即缩小。随访 MRI 也显示病情正持续改善, 患者现已出院在家中接受后续治疗。近期, 第二名患者也开始服用该药物, 初步检测结果同样令人鼓舞。

相关论文信息:

<https://doi.org/10.3201/eid2901.221531>

科学家创造出新型冰

本报讯 研究人员发现了一种无固定形状的新型冰——它没有整齐有序的晶体结构, 这有助于揭开液态水的奥秘。相关研究 2 月 2 日发表于《科学》。

人们已经知道两种类型的无定形冰——高密度和低密度的。研究人员认为没有办法制造出中等密度的无定形冰(MDA)。但是, 当英国伦敦大学学院的 Christoph Salzmann 等研究人员将普通冰(具有六边形晶体结构)放入一个装有冷却至 -200°C 钢球轴承的玻璃杯时, 它们碰撞产生的剪切力制造出了 MDA。

“这是一个周五下午进行的实验。我们只是把冰打成小块, 本以为什么都不会发生。”Salzmann

说, “但令人惊讶的是, 确实发生了一些事情。”

实验中产生的细白色粉末密度正好介于其两种已知形式的无定形冰之间, 且几乎与液态水的密度完全相同。研究员表示, 这可能是水的另一种存在状态——玻璃相, 即一种在极低温度下也能表现为液态的物质。在短时间内, 玻璃相可能看起来是固体, 但在长时间内它会像黏性液体一样流动。

液态水看似平凡, 而一旦冷却到极低温度, 它就能表现出特殊的性能。基于低密度和高密度无定形冰之间的差距, 研究人员曾提出, 过冷水可能同时以两种不同的液相存在, 其中一种浮在另一种之上, 但 MDA 的存在使这一想法

受到质疑。

Salzmann 说:“它表面和普通冰相似, 但密度和液态水一样, 所以最大的问题是, 它到底是什么? 如果能弄清楚 MDA 是什么, 那我们将更好地了解液态水。”

MDA 也可能是外太阳系冰卫星的重要组成部分。这些星球由于其主行星的引力而经历了强烈的剪切力, 这可能为 MDA 的形成创造了合适的条件。研究员还发现, 这种冰升温和释放大量热量, 这可能会对这些星球的地质活动产生巨大影响。

相关论文信息:

<https://doi.org/10.1126/science.abg2105>

■ 自然要览

(选自 Nature 杂志, 2023 年 2 月 2 日出版)

超新星前身星系统识别非常重要

科学家展示了最近发现的高质量 X 射线双星 CPD-29 2176 所具有的演化历史并表明, 中子星成分是在极度剥离超新星时形成的。此双星的轨道元素在周期和离心率上都与已知轨道周期和离心率的 14 个 Bex 射线双星中的 1 个相似。

对极度剥离超新星前身星系统的识别是必要的, 因为它们的演化路径导致了双中子星系统的形成。双中子星如被观测到有电磁和引力能, 能产生千新星 GW170817 的系统, 则会产生大量的重元素。

相关论文信息:

<https://doi.org/10.1038/s41586-022-05618-9>新研究为实现偶极超流体
和分子超固体铺平道路

近日, 研究人员展示了一种适用范围广泛的新型极性分子。所谓的场联共振发生在微波修饰的分子散射中, 因为分子间势中存在稳定的宏观四聚体状态。

研究者确定了超冷基态钠