

人工智能与天然药物 动态监测快报

总第 003 期 2023-11-20



中国科学院昆明植物研究所
云南省植物学会
中国科学院成都文献情报中心

目 录

领域前瞻.....	1
未来 10 年，生物医药需重点关注的 10 大领域.....	1
政策规划.....	7
天津市人民政府关于印发天津市中医药强市行动计划（2023—2025 年）的通知 ...	7
研究前沿.....	8
药物先导化合物优化的人工智能新方法	8
大分子结构测定的挑战与机遇	9
蛋白质结构预测将由机器学习主导	12
药效团引导的深度学习方法用于生物活性分子生成	13
用于药物发现的多种去中心化机器学习技术	14
利用高通量 AI 驱动的平台改变药物发现：Patrimony 平台 5 年经验回顾	15
使用 AlphaMissense 准确预测全蛋白质组错义突变效应	17
利用生成式 AI 预测蛋白质的构象变化	18
从中国传统天然化合物中发现靶向 ADAR2 和 5-HT2CR RNA 结合环的潜在抑制剂	19
近期会议与活动	20

领域前瞻

未来 10 年，生物医药需重点关注的 10 大领域

2023-11-08

近日，西湖大学未来产业研究中心发布了《未来产业发展趋势研究 2023：未来生物医药》。

该报告采用人工智能文本分析技术对全球生物医药技术进行数据挖掘和分析，识别出我国未来生物医药产业发展的重点方向，并有针对性地提出一系列政策建议。基于全球未来生物医药产业布局及我国的重点研究方向，结合当前至 2035 年我国未来生物医药产业发展面临的形势，在充分考虑专家问卷中各位专家对开放性问题（您认为还有哪些值得关注的未来生物医药产业发展方向和潜在技术领域）的回答基础上，我们（西湖大学未来产业研究中心）探索性提出我国未来生物医药产业需重点关注和大力支持的十大领域包括：

1、AI 等信息化技术在生物医药的应用

人工智能等信息化技术在生物医药领域的深度应用，即“人工智能+生物医药”（Artificial Intelligence + Biomedicine），是指企业和研究机构通过将人工智能与生物医药结合以实现生物医药领域的创新突破。

2017 年，《自然》杂志报道了一种的具有专家级皮肤癌诊断能力的 AI 系统，是人工智能赋能生物医药领域的开创性里程碑式事件。

2020 年，DeepMind 公司发布了 AI 算法 AlphaFold 2，其能够基于氨基酸序列准确地预测蛋白质的 3D 结构，其预测的精准程度可与实验技术解析的 3D 结构相媲美。该成果被认为解决了一个生物学 50 年来的重大挑战，引发了科学界的震动，也再次掀起了“人工智能+生物医药”的研究和产业化热潮。

2022 年，Meta 公司基于最新的几何深度学习模型“EquiBind”开发的 ESMFold，以比 AlphaFold 2 快 60 倍速度成功预测 6 亿多种蛋白质的结构。

“人工智能+生物医药”在生物医药的诸多子领域都有所应用（新药开发、酶及蛋白质设计、医学图像分析、疾病预测、疾病预防、智能诊断、精准医疗等），并有望广泛重塑生物医药研究和产业现状。

“人工智能+生物医药”能够实现在生物医药产业自上游到下游的投入使用，并且部分应用场景已经能够为企业带来实际收益。

以新药开发为例：在研究调研阶段，“人工智能+生物医药”可以通过对来自文献等的信息进行自动化的文本分析，抽取和药物研发相关的关键信息，以辅助研发

人员在研究和产业化中的决策；在药物开发阶段，通过对生物医药试验和临床大数据的智能分析以缩短发掘新治疗靶点、新药分子的周期，降低药物研发成本；在临床验证阶段，通过人工智能对临床试验数据的智能化分析，以更好了解新药对于不同患者的治疗效应。

未来，我们将看到更多人工智能应用于生物医药产业链中各环节的案例，通过人工智能的应用来进一步增强生物医药相关单位的产业竞争优势。

2、重组抗体技术

重组抗体（Recombinant antibody）是指利用重组 DNA 等分子生物学技术产生的抗体。重组抗体的最大特点是编码其抗体蛋白质的氨基酸或 DNA 序列是已知的。因此，在制备重组抗体时，人们可以通过重组 DNA 等技术，将编码重组抗体的基因序列插入至表达载体，并将其转入至表达宿主中（如哺乳动物细胞、酵母或细菌），进而表达纯化以获得特定种类的重组抗体。

重组抗体技术持续发展，单链抗体、纳米抗体、双特异性抗体等类型的重组抗体近年来也受到了广泛研究，许多产品也已被批准上市。人工智能等技术的蓬勃发展也使得人们能够更加理性和快速地设计出更高效的重组抗体。另外，重组抗体的制备和生产技术也在持续拓展，其中无细胞表达合成体系值得关注。无细胞表达合成体系由于可以进一步实现无表达宿主式的抗体生产，因而有望在更短的时间产生更广范围的抗体产品。由于配方调节的灵活性，无细胞合成技术亦可适用于那些由人工智能设计产生的、但宿主表达难度高的抗体的制备和生产。

目前，基于重组抗体的药物研发也已成为生物制药的主流之一。未来，重组蛋白类药物将在癌症、传染病、免疫、内分泌代谢和神经系统等疾病防治领域发挥巨大作用。

3、小分子抑制剂技术

小分子药物一直在医学进步中发挥着重要作用，并解决人们未满足的需求，其也是每年新批准药物中占比最大的药物类型（2022 年 FDA 批准的新药中，小分子药物占比超过五成[2]），未来，小分子类药物预计仍然将持续在新药研发中占有较大比重。

小分子抑制剂（Small molecule inhibitor）属于小分子类药物，是指一类能够靶向作用于蛋白，降低蛋白活性或者阻碍生化反应的、分子量小于 1000 道尔顿的有机化合物分子。小分子抑制剂通过直接和靶蛋白结合，以底物竞争、改变蛋白结构，或者阻碍蛋白构象转变等方式降低靶蛋白活性。

小分子抑制剂常见于目前临床上使用的各种药物中，包括各类蛋白质、酶、激酶、转录因子、质子泵、离子通道抑制剂等。小分子抑制剂通常能够迅速使其靶标

失活因而在精细调控细胞生命和功能方面具有显著优势，这也使其成为生命科学研究中重要的工具药。

近年来，得益于人工智能、计算化学、分子对接、蛋白质结构解析和预测等技术的发展，人们得以更有效地发掘小分子抑制剂的新靶点，并对小分子抑制剂进行理性药物设计，进而加速小分子抑制剂的新药研发。未来，小分子抑制剂将更广泛地应用于癌症治疗及其他治疗领域，更多类型的小分子抑制剂将问世。

4、高通量测序技术

高通量测序（High-throughput sequencing）是指以高通量、快速、高效、且经济的方式对各种生物序列（例如 DNA、RNA、蛋白质等序列）进行测序。在传统意义上，高通量测序通常是具体代指高通量基因测序，国家发展改革委印发的《“十四五”生物经济发展规划》中所提出的：“要加快发展高通量基因测序技术，推动以单分子测序为标志的新一代测序技术创新，不断提高基因测序效率、降低测序成本。”然而，随着近年来针对蛋白质等非核酸序列进行高通量测序的技术的涌现，高通量测序的含义也随之扩展。

近年来高通量基因测序技术的发展已较为成熟，其测序对象和应用场景已十分多样化，例如：全基因组从头测序、全基因组重测序、全基因组甲基化测序、全外显子组测序、全转录组测序、RNA 测序，等等；许多新型高通量基因测序技术也陆续涌现，长读测序、单分子测序、单细胞测序、空间转录组测序等技术为生物医药研究带来新的可能。

随着人类基因组测序工作的完成，生命科学研究重心或将从基因组学拓展至蛋白质组学。想要深入了解蛋白质组全部的成分和序列信息，进一步认识生命活动和疾病发生的分子机制，其关键即在于要有合适的高通量蛋白质测序技术做支撑。目前，高通量蛋白质测序技术虽然尚不如高通量基因测序技术般强大。然而，非质谱类高通量蛋白质测序、蛋白质高分辨率质谱、单分子蛋白质测序等新型高通量蛋白质测序技术的不断涌现，也让该技术日趋成熟。

近年来，人工智能、生物信息学等软件算法的进步也为高通量测序技术的发展带来新的机遇。通过更智能的信息化技术，人们能够以更高效且自动化的方式对测序大数据进行分析，进而获取有意义的测序结果。

未来，高通量测序技术将被更广泛地应用于生物医药领域，成为精准医疗等现代医学领域的重要基石；高通量蛋白质测序或将成为继高通量基因测序之后的新的产业化增长点。

5、药物偶联物技术

药物偶联物（Drug conjugate）是指一类运用特定的连接子（通常是化学链）将具有靶向定位性的配体和效应分子连接起来而产生的药物，其核心理念是定位配体

发挥靶向投递作用，效应分子发挥治疗作用。总的来说，药物偶联物的构成可以用“定位配体-连接子-效应分子”的公式进行概括，根据定位配体的类型差异，又可将药物偶联物进一步具体细分为：抗体药物偶联物（Antibody-drug conjugate）、多肽药物偶联物（Peptide-drug conjugate）、蛋白药物偶联物（Protein-drug conjugate）、小分子药物偶联物（Small-molecule drug conjugate）、高分子药物偶联物（Polymer-drug conjugate）、放射性核素药物偶联物（Radionuclide-drug conjugate）、病毒样药物偶联物（Virus-like drug conjugate），等等。

以近年来发展较好的、具有代表性的抗体药物偶联物（ADC）为例，通过将抗体用作定位配体，ADC 的构成即可表示为“抗体-连接子-效应分子”。相比传统药物，ADC 具有更好的给药靶向性。2000 年，首个 ADC 被 FDA 批准用于治疗急性髓性白血病，但存在具有致死性毒性等缺点。近年来，ADC 技术日益进步，改进后的 ADC，其不良反应的发生率也显著降低。随着 Brentuximab vedotin（商品名 Adcetris）、Trastuzumab emtansine（商品名 Kadcyla）等新型 ADC 被 FDA 批准用于治疗霍奇金淋巴瘤和 HER2 阳性乳腺癌，ADC 药物再次广泛进入人们的研究视野。目前，ADC 依然存在巨大的发展空间。定向偶联、多价偶联、重组抗体和小分子药物等技术的进步，为 ADC 的药物研发带来新的可能，基于单链抗体、纳米抗体、双特异性抗体等类型抗体的 ADC 药物也持续涌现。

随着药物偶联物技术的持续进步，药物偶联物的定位配体、效应分子和连接子的类型选择也将日趋多样化。未来，我们将看到更多类型的药物偶联物被批准应用于临床，新一代的药物偶联物将持续为患者带来福音。

6、治疗性基因编辑技术

治疗性基因编辑（Therapeutic gene editing）是指一类通过对基因进行靶向编辑（敲除、插入、替换、修饰等）进而获得治疗效果的疗法。

CRISPR-Cas 基因编辑技术具有可编辑范围广、易用、高效、廉价等特点，被广泛应用于生命科学、药物研发等方面的研究。另外，许多新型 CRISPR-Cas 系统也被开发，并应用于基因编辑相关的新兴领域，如 RNA 编辑、单碱基编辑、先导编辑、CRISPR 干扰（CRISPRi）等等。

尽管目前基因编辑在疾病治疗领域的研究数量庞大，但其中大部分研究尚处于临床前的实验室研究阶段，如何进一步优化基因编辑的效率、准确性、可编辑基因序列范围，并降低基因编辑脱靶效应所带来的安全风险或是促进其被广泛应用于治疗领域并实现产业化的关键。

另外，CRISPR-Cas 也非实现治疗性基因编辑的唯一技术路径，许多其他类型的基因编辑技术也仍值得持续关注，如基于转座子、类转录激活因子效应物核酸酶、锌指核酸酶等的治疗性基因编辑技术。

未来，我们将看到更多治疗性基因编辑相关新技术的诞生，人类将以前所未有的方式治愈一些目前尚无良好疗法的疾病（如基因缺陷型疾病）。

7、细胞治疗技术

细胞治疗（Cell therapy）是指一类将活细胞移植入患者体内以实现治疗效果的疗法。细胞治疗可以根据所使用的治疗用细胞的类型而进一步细分，如基于免疫细胞的细胞免疫疗法、基于干细胞的干细胞疗法等。

2017 年，首个 CAR-T 疗法（由 Kymriah 公司开发）被 FDA 批准以用于治疗急性淋巴细胞白血病。但截至目前，CAR-T 疗法仅在血液瘤上取得较好的临床疗效，其面临的主要挑战是如何使工程免疫细胞能够应用于更广泛癌症类型（特别是实体瘤）的治疗。

干细胞疗法通过利用天然或诱导的干细胞的自我更新能力、多分化潜能等特性，对病变/衰老的细胞/组织进行修复或功能重建，进而达到治疗效应。在临床上，除了可以将天然干细胞（如脐带血干细胞、间充质干细胞等）直接用于干细胞治疗，诱导性多能干细胞、细胞重编程等技术也值得关注。通过诱导性多能干细胞、细胞重编程等技术，人们能够将已分化的细胞在特定条件下逆转恢复到类干细胞状态（或直接转分化为目标细胞类型）以用于治疗。这类技术由于具有逆转细胞命运的潜能，因而也可能为目前市场前景巨大但尚无显著突破的抗衰老领域带来希望。

未来，更多类型的细胞将被开发用于细胞治疗；从患者出发的个性化、定制化细胞治疗有望为更多的不治之症提供新的治疗方案。

8、新型药物递送技术

药物递送系统（Drug delivery system）是指在空间、时间及剂量上全面调控药物在生物体内分布的技术体系，通过增强治疗药物对其目标部位的递送，最大限度地减少目标外积累，从而改善患者的健康。相对于常规口服片剂、胶囊、静脉注射剂、吸入制剂和透皮贴剂等为主的传统药物递送系统，新型药物递送系统是指采用整体具有较高技术壁垒的新型药物递送技术（如基于脂质体、纳米粒、微球、外泌体、工程 AAV 载体、3D 打印药物制剂等的药物递送技术）对各类药物进行递送的系统。新型药物递送系统能够通过调节药物的递送和释放位置，改变药物体内代谢行为，改善药物缓释控释特性、透生理屏障（如血脑屏障）特性等方式，提高药物的疗效，并降低毒副作用。2018 年，FDA 批准脂质纳米粒作为递送载体的首款 RNAi 药物 Onpattro 上市；2021 年批准脂质纳米粒作为递送载体的首款 mRNA 疫苗 Comirnaty 上市。

未来，为满足各种新型药物（如基因药物、mRNA 药物、多肽及蛋白质类药物、细胞药物等）的药物递送需求，我们也需要更多创新的新型药物递送系统。

9、免疫检查点抑制剂

免疫检查点抑制剂（Immune checkpoint inhibitor）是指一类针对肿瘤的免疫治疗药物，其主要机理是通过阻断一类被称为免疫检查点的蛋白质，以恢复免疫系统对肿瘤细胞的杀伤能力，从而起到肿瘤治疗的作用。

2018 年，美国科学家詹姆斯·艾利森（James Allison）和日本科学家本庶佑（Tasuku Honjo）因在免疫检查点抑制剂方面的贡献而获得诺贝尔生理学或医学奖。

目前上市的免疫检查点抑制剂主要是针对免疫检查点 PD-1/PD-L1 或 CTLA-4 的单克隆抗体型药物。2011 年，FDA 批准了首款免疫检查点抑制剂，即靶向 CTLA-4 的 Ipilimumab 单抗，以用于治疗黑色素瘤。之后，陆续有靶向 PD-1/PD-L1 的，适用于黑色素瘤、肺癌、肠癌等肿瘤的免疫检查点抑制剂被批准。

除此之外，目前还有多种针对 LAG-3、TIM-3、TIGIT、VISTA 等免疫检查点靶点的免疫检查点抑制剂在研；另外，双特异性抗体也是免疫检查点抑制剂的一个研发关注点，有多种能够同时靶向 PD-1/PD-L1 和 CTLA-4 的双特异性抗体在研。

未来，随着人们对肿瘤免疫调节机理的更深入研究，有望开发出更多能够适用于不同肿瘤治疗的免疫检查点抑制剂。

10、脑机接口技术

专家问卷反馈中，脑机接口技术是被相对较多提及的，对未来生物医药将有潜在重大影响的技术。脑机接口（Brain-Computer Interface, BCI）是指一种人机连接方式，其可以将脑信号直接与机器（最常见的是计算机、芯片或机器人肢体）建立通信通路。通过脑机接口技术，人们能够通过机器，对大脑中神经系统编码的记忆、决策、情感等信息进行提取和识别；基于机器接收的脑信息，人们又可以进一步对脑活动进行调控。

根据《中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景目标纲要》，脑科学和人工智能为国家战略科技力量，而脑机接口技术（脑机融合技术）是其中的关键技术。

未来，脑机接口技术有望促进脑科学和人工智能领域的研究突破，并在神经精神疾病的诊断、治疗、康复等临床领域取得更广阔的应用。

信息来源：<https://bydrug.pharmcube.com/news/detail/c8170693cadf36428c5aa2d435691029>

政策规划

天津市人民政府关于印发天津市中医药强市行动计划（2023—2025 年）的通知

2023-11-10

2023 年 11 月 3 日，天津市人民政府发布《天津市中医药强市行动计划（2023—2025 年）》。计划主要包括九项内容：

一、完善中医医疗卫生服务体系。包括全面推动建设国家中医医学中心、中西医协同“旗舰”医院、中医特色重点医院、实施天津名医堂工程、打造天津市中医医疗集群。使我市优质中医医疗资源提质、扩容。

二、提升中医药服务能力。强化中医治未病和康复能力建设。健全现代中医医院管理制度。实施基层中医药服务能力提升工程。发挥中医药在治未病中的主导作用、在重大疾病治疗中的协同作用、在疾病康复中的核心作用。

三、提升中医药应急救治能力，建设国家中医疫病防治基地。完善突发公共卫生事件中医药应急救治能力，建立中西医协同救治和中西医联合会诊制度。

四、强化中医药人才培养，支持天津中医药大学建设，培育天津市顶尖学科；实施津沽中医药人才系统培育工程暨中医药杰出人才、领军人才、中青年骨干人才和基层人才的四级人才培养体系。完善人才评价体系。

五、推进中医药传承创新平台建设，推动国家中医针灸临床医学研究中心、组分中药国家重点实验室等科研平台建设，构建产学研医用协同创新模式。

六、促进中医药产业高质量发展。建设现代中医药创新标志区。实施天津市中药制剂工程。促进中成药高质量发展，深化医保供给侧改革，开展中成药带量采购。支持中药工业标准化发展和中药材产业化发展等。

七、促进中医药与健康服务融合发展，推进中医药健康旅游和中医药服务贸易发展，深化国家中医药健康旅游示范基地和国家中医药服务出口基地建设。

八、传承发扬中医药文化。加强中医药文化研究和卫药文化宣传，普及中医药养生保健知识，开展多种形式的中医药文化传播活动。

九、完善医保支持中医药特色发展政策措施。深化中医药医疗服务价格改革。建立目标导向的中医服务价格项目管理机制，优化现有中医价格项目，完善新增中医服务价格项目管理政策。完善符合中医药特点的医保支付政策。

信息来源：https://www.sohu.com/a/733812215_776163

研究前沿

药物先导化合物优化的人工智能新方法

2023-10-23

先导化合物的结构优化（Lead optimization）是药物设计的核心环节，需要通过 DMTA（设计-合成-测试-分析）的反复循环来提高化合物活性、特异性、成药性等性质。长期以来，先导化合物的活性优化高度依赖药物化学家的经验以及大量的人力和资源投入。基于结构的药物设计，尤其是高精度的结合自由能计算，可以通过部分模拟的 DMTA 循环来加速先导化合物活性优化过程。随着分子力场和构象采样算法的改进，自由能微扰（FEP）等相对结合自由能模拟方法的预测值与实验结果的误差可以接近化学精度（1 kcal/mol，约 5-6 倍活性差异以内），但这类方法通常需要复杂的配置与体系搭建过程。此外，消耗计算资源庞大、商业软件价格高昂等问题也限制了这类方法的应用范围。

近期，中国科学院上海药物研究所郑明月课题组提出了一种先导化合物优化的人工智能方法 PBCNet（pairwise binding comparison network）。该方法采用孪生图卷积神经网络架构，通过比较一组相似配体的结合模式差异来预测二者之间的相对结合亲和力，可以较好的兼顾计算速度和精度。此外，研究团队还开发了易于操作的图形界面网络计算服务（<https://pbcnet.alphama.com.cn/index>）。相关研究论文

“Computing Relative Binding Affinity of Ligands Based on a Pairwise Binding Comparison Network” 于 2023 年 10 月 19 日于 Nature Computational Science 在线发表。

信息来源：http://simm.cas.cn/web/xwzx/kydt/202310/t20231023_6904813.html

推荐评论:

上海药物所郑明月课题组提出的 PBCNet 方法，为先导化合物优化领域带来了令人振奋的创新。在药物设计中，先导化合物的结构优化是关键环节，但受限于需要实践经验和大量人力资源的传统方法存在局限性。目前基于结构的药物设计方法也有算力要求高，体系搭建复杂等问题。PBCNet 方法采用了孪生图卷积神经网络架构，通过比较相似配体的结合模式差异来预测它们之间的相对结合亲和力，从而加速了先导化合物的活性优化过程。这一方法兼顾了计算速度和精度的优势，并且研究团队还提供了易于操作的图形界面网络计算服务，为研究人员提供了方便的工具。

大分子结构测定的挑战与机遇

2023-10-27

2023 年 10 月 17 日, Nature reviews molecular cell biology 杂志 (最新影响因子 112.7) 发表文章 Challenges and opportunities in macromolecular structure determination。

在文章中, 德克萨斯大学西南医学中心生物物理系的白晓晨、加利福尼亚大学洛杉矶分校霍华德-休斯医学研究所的 Tamir Gonen、匹兹堡大学结构生物学系的 Angela M. Gronenborn、荷兰癌症研究所生物化学部肿瘤编码研究所的 Anastassis Perrakis、汉堡大学纳米结构与固体物理研究所的 Andrea Thorn、山东大学数学与交叉科学研究中心的杨建益等人讨论了确定大分子结构的主要挑战和机遇, 强调了经验实验和基于人工智能的技术在揭示大分子结构复杂性方面的重要协同作用。

文中对于大分子结构测定的新领域有以下设想:

Angela M. Gronenborn:

大分子结构测定必须超越"分而治之"的方法, 不能再孤立地分析大分子。将来, 我们将研究越来越少的生化重组系统, 更多地研究从天然细胞或组织环境中分离的复合物和组件。

生物现象的复杂性与任何表征的内在不完整性相关联, 因此需要使用实验、计算和理论等多种方法。众所周知, 单个类型的结构数据在范围、准确性和通用性方面都是有限的, 而以整合的方式使用互补信息有助于克服任何缺陷。这种整合将成为常态, 并可通过机器学习加以推进。

白晓晨:

大多数蛋白质的结构和功能都受细胞环境中周围分子的调控, 如细胞骨架、其他蛋白质、膜和小分子代谢物。在研究纯化蛋白质时, 所获得的结构信息可能缺乏这些重要的细胞因素。

然而, 低温电子显微镜的最新进展与子断层平均 (sub-tomogram averaging) 技术相结合, 开辟了一个令人兴奋的新领域: 直接观察细胞内的大分子机器或大

型膜蛋白。这种方法不需要分离蛋白质，就能确定蛋白质复合物在其原生环境中的结构。

低温相关光电子显微镜（cryo-CLEM）与低温聚焦离子束（cryo-FIB）对细胞样本进行铣削（减薄）的强大组合，可精确定位细胞内的蛋白质复合物。与单粒子分析类似的子断层平均是一种新兴技术，可提高低温电子显微镜重建的分辨率。低温电子显微镜还可用于研究大型蛋白质团簇和液态凝聚体的结构，从而扩展我们对这些复杂细胞结构的了解。

另一个令人兴奋的前沿领域是利用低温电子显微镜对直接从动物组织中纯化的蛋白质复合物进行结构研究。低温电子显微镜所需的样品量比 X 射线晶体学所需的样品量要小，而且可以容许一定程度的杂质，因此分析来自原生来源的蛋白质复合物是可行的。通过将低温电子显微镜分析与质谱分析相结合，我们可以找出对复合物的形成和功能至关重要的不明显成分。

Tamir Gonen:

我相信，在未来 10 年内，我们将对关键膜蛋白有透彻的了解，我相信 MicroED 将在确定其他方法无法确定的结构方面发挥重要作用。

我们还需要更多地了解膜是如何组织的，以及这些蛋白质是如何在膜环境的大背景下工作的。这是人工智能目前正在努力解决的问题。实验性结构测定为我们提供了一个快照，但我们需要收集许多快照才能更好地了解这些蛋白质的功能。有了这些知识，我们就有可能开发出新的药物和治疗方法。

杨建益:

大分子结构测定的新前沿正在从结构的统计分析向结构的生物功能导向转变。

在不久的将来，实验数据（如低温电子显微镜密度图）与计算算法的结合将取得重大进展，将大大加快生物功能导向的结构的确定的预测，并降低成本。希望蛋白质组结构的确定在未来能像基因组测序一样容易，为更好地了解大分子的生物功能铺平道路。

Anastassis Perrakis:

我们已经看到了美好未来的曙光：在细胞内部，我们看到了核糖体的不同构象状态，我们确定了微管内的数十种蛋白质，我们列举了核糖体蛋白质和核糖核酸的翻译后修饰，我们正在拍摄病毒生命周期的快照。

在不久的将来，我们将获得更高的空间和时间分辨率，看到更小、更复杂的大分子机器，我们还可能了解细胞内的无膜区，以及相分离和蛋白质内在无序区在其形成过程中的作用。

Andrea Thorn:

我们的最终目标是了解活细胞中分子的功能和结构。这将需要在原子分辨率方法之间架起桥梁，还需要整合不同分析尺度的见解，并以有意义的计算方式将所有这些信息结合起来。

为此，我们需要克服科学中的可重复性危机，并在未来几年充分利用技术、自动化和机器学习方面的所有进步，这确实是一个激动人心的时刻。

信息来源:

https://mp.weixin.qq.com/s?__biz=MzU2NTA1MzYyOA==&mid=2247549545&idx=1&sn=7865243607309f8f583d6c074e26509f&chksm=fc43d3bacb345aac4411a245f263ed5d94f42699d4d74203ebba6bf99191a2542198e382595b&scene=0&xtrack=1

推荐评论:

经验实验和基于人工智能的技术在揭示大分子结构复杂性方面具有重要协同作用。该研究提出了一些新领域的设想，包括研究细胞或组织环境中的复合物和组件、整合互补信息、利用低温电子显微镜对纯化的蛋白质复合物进行结构研究等。这些设想将推动大分子结构测定领域的发展，帮助理解大分子生物功能。

细胞内生物大分子复合体的结构直接决定复合体的功能，原位解析活细胞中复合物的功能对于深刻理解生命活动至关重要。充分整合多学科技术和机器学习等人工智能手段将辅助加快结构解析。

蛋白质结构预测将由机器学习主导

2023-11-10

本文对蛋白质机构机器学习预测的发展进行介绍。AlphaFold 是蛋白质研究历史上的一项革命性进展，首次提供了凭借序列信息就能够准确预测绝大多数蛋白质的三维氨基酸排列的实际能力。

尽管我们为首次能够从基因序列准确预测基因组规模的蛋白质结构的发明而欢呼，但许多科学家也感到有些不安，因为我们并没有像预期的那样提取出可推广的传统科学解释原则。如同基于开普勒的定律，我们可以准确预测行星的轨道，但开普勒的技术的全部威力只有在牛顿从中提取了力学原理后才被充分释放出来。又如兰登的 X 射线，1895 年发现，于 1901 年获得诺贝尔物理学奖，长期以来在医学上产生了巨大影响，但在通过量子力学理解它们之前，它们一直是未解之谜。随着时间的推移，我们可以希望会出现有关蛋白质实际如何折叠（与最终形状不同）的类似深刻的见解和新原则。

当前版本的 AlphaFold 只是一个开始，并不断改进，包括多亚单位蛋白质复合体。它将吸引开放式创新，并以无法预测的方式发展，触及生物科学的各个角落。作者所知道的每个实验室都在以高度富有想象力的方式使用 AlphaFold。即使在这些探索中不可避免失败，但是也将会在许多方面不断改进。AlphaFold 最长期的影响之一可能是预示 AI 即将成为生物学中被接受的、可靠和有用的发现方法。今天的科学家大多接受了使用模型的培训，但并不真正信任它们，特别是无法明确理解这些模型的时候。我们准备好信任模型并对它们进行验证了吗？

信息来源：

https://mp.weixin.qq.com/s?__biz=MzU2NTA1MzYyOA==&mid=2247550328&idx=1&sn=fb996ebf9df9c69641039531281c5e05&chksm=fc43d4abcb345dbda91ef7ed14e1e5f45940789cfe62130677952f7c9ff850189c87912819c0&scene=0&xtrack=1

推荐评论：

AlphaFold 的出现标志着蛋白质结构与功能研究的重大突破，通过序列信息准确预测大多数蛋白质的三维结构，是具有革命性意义的进展。然而，该技术的发展也遇到了问题，因为科研工作者尚未完全理解背后的科学原理。尽管如此，AlphaFold 的不断改进和应用将推动生物科学的创新，并在各个领域产生重要影响。它的出现也预示着人工智能将成为生物学中被广泛接受、可靠且有用的工具。

药效团引导的深度学习方法用于生物活性分子生成

2023-10-08

2023 年 10 月 6 日, 中南大学李敏教授团队, 联合中南大学曹东升教授, 在 Nature Communications 上发表文章: A pharmacophore-guided deep learning approach for bioactive molecular generation。

作者提出了一种药效团引导的深度学习方法用于生物活性分子生成

(Pharmacophore-Guided deep learning approach for bioactive Molecular Generation, PGMG)。通过药效团的引导, PGMG 提供了一种灵活的生成生物活性分子的策略。PGMG 使用图神经网络对空间分布的化学特征进行编码, 并使用 Transformer 解码器生成分子。引入隐变量解决药物载体与分子之间的多对多映射, 提高生成分子的多样性。与现有方法相比, PGMG 的灵活性和有效性使其成为加速药物发现过程的有用工具。作者还为 PGMG 开发了一个易于使用的 web 服务器, 用户可以为任何给定的药效团假设生成分子。

信息来源:

https://mp.weixin.qq.com/s?__biz=MzU2NTA1MzYyOA==&mid=2247550300&idx=1&sn=91b286305651229c7dd8bc1a3e581261&chksm=fc43d48fcb345d99a432f4e5399fa4319e1f0064535ff8fe137a3836e3744e404464002fd059&scene=0&xtrack=1

推荐评论:

首先, PGMG 克服了数据稀缺的问题, 通过在统一表述中, 能利用不同类型数据的一种计算方法。第二, 备选药效团包含生物学意义, 因此 PGMG 的生成过程具有比较强大的研究背景和可解释性。此外, PGMG 能针对特定靶点的药效团, 可用于设计多靶点活性分子。PGMG 有潜力用于化学文库的构建, 以, 因为它可以为特定的靶标提供一定数量的候选药物样分子。PGMG 基于药效团假说, 提高虚拟筛选效率, 生成分子具有结构多样性, 可以应用于研究替代药物和克服耐药性等多种药物设计。

用于药物发现的多种去中心化机器学习技术

2023-11-16

2023 年 11 月 5 日, 奥地利维也纳大学 Aljoša Smajić 等人在 Drug Discov Today 发表文章 Privacy-preserving techniques for decentralized and secure machine learning in drug discovery。

数据可用性、数据安全和隐私问题通常会影响机器学习技术的最佳性能。因此, 在药物发现领域, 人们提出了利用私人/敏感数据来完成机器学习模型构建任务的新技术。这些新技术包括安全多方计算、分布式深度学习、同态加密、基于区块链的点对点网络、差分隐私和联邦学习, 以及这些技术的组合。

文章概述了这些被用于药物发现的多种去中心化机器学习技术。这些技术将使各方有可能在不披露内部数据集的情况下共同开发模型。

去中心化机器学习方法对于虚拟筛选和药物设计等应用来说非常有利。研究人员可以利用分布在许多制药公司和研究机构的数据集, 针对治疗靶点训练化合物活性预测模型, 从而提高虚拟筛选的准确性, 加快药物开发过程。

文中还讨论了这些方法在药物发现中的最新应用、策略的适用性、在药物发现早期阶段的意义, 以及未来的潜力、新发展和优缺点。

信息来源:

https://mp.weixin.qq.com/s?__biz=MzU2NTA1MzYyOA==&mid=2247550380&idx=1&sn=823528135cf268a8233b103be02cdbe0&chksm=fc43d47fcb345d691cef352c312e126947661d04bbdceac34148db03be77a754f333b9745d68&scene=0&xtrack=1

推荐评论:

去中心化机器学习技术可用于解决药物发现领域中的数据可用性、数据安全和隐私问题。对于虚拟筛选和药物设计等应用非常有帮助。通过分析多个制药公司和研究机构的数据集, 研究人员可以训练化合物活性预测模型, 提高虚拟筛选的准确性, 加速药物开发过程。

利用高通量 AI 驱动的平台改变药物发现：Patrimony 平台 5 年经验回顾

2023-11-16

2023 年 11 月，来自施维雅 Patrimony 团队的研究人员在 Drug Discovery Today 发表文章 Transforming drug discovery with a high-throughput AI-powered platform: A 5-year experience with Patrimony。文章介绍了施维雅推出的 Patrimony 平台如何利用计算科学和人工智能来整合内部和外部的海量多模态数据，并就该平台在产业化和扩大应用中所面临的主要挑战和关键成功因素分享了施维雅的经验。

Patrimony 的核心是知识图谱（KG），这是一种用于集成和解释大量多模态数据的经典工具。为了表征受干扰的生物系统在具体疾病中的复杂性，Patrimony 被设计为整合各种类型的相互作用，包括蛋白质-蛋白质、基因-分子功能和药物-靶点。Patrimony 集成了内部和外部数据库，包括 Drugbank（药物-靶点），DisGeNET（基因-表型）和反应组（基因-生物过程），以生成将疾病与基因和药物联系起来的综合图（图 1）。

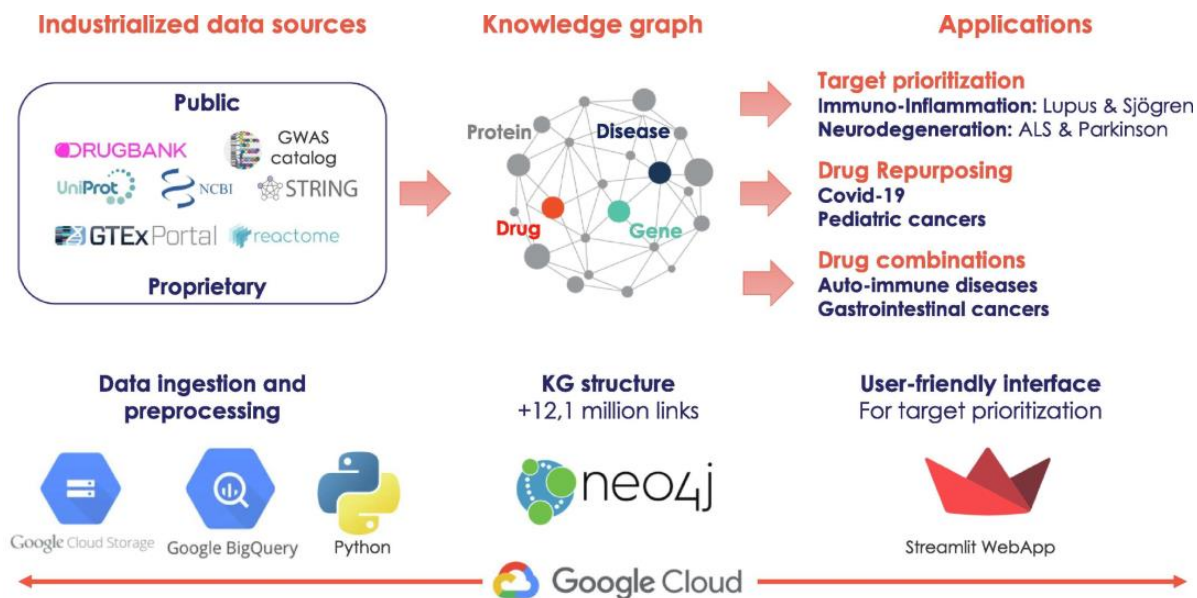


图 1 Patrimony 的主要应用和底层结构概述

作者总结了在实施 Patrimony 并使其产业化的 5 年历程中，几个需要考虑的关键点：首先是在工业化环境中保持不断发展的能力，这需要整合新的数据源，改进

分析方法，开发新的应用，同时确保数据质量和标准化流程；其次是开发或搭建可互操作的工业工具或其他计算平台，将 Patrimony 等平台进一步嵌入基于人工智能的端到端研发价值链支持中，从而提高临床前候选药物的成功概率，并缩短研发时间；第三个要点是，计算平台和人工智能的变革力量在于将计算预测建模与传统的制药研发活动充分结合起来。

信息来源：

https://mp.weixin.qq.com/s?__biz=MzU2NTA1MzYyOA==&mid=2247550161&idx=1&sn=36ea14cae2a787a2171afabdbef45fa7&chksm=fc43d502cb345c14f8a4daf116e05085cfe6d254894cd9cd9482420f613051cf2ca90a0c10ca&scene=0&xtrack=1

推荐评论：

高通量 AI 驱动平台 Patrimony 利用计算科学和人工智能技术，整合了内部和外部的海量多模态数据，并通过知识图谱的构建和应用，实现了对复杂生物（医药）系统的综合解析。它整合蛋白质、基因、分子功能和药物等多种相互作用关系，从而将疾病与基因和药物联系起来。

使用 AlphaMissense 准确预测全蛋白质组错义突变效应

2023-11-08

AlphaFold 使用蛋白质序列作为输入，可以大规模预测高度准确的蛋白质结构。这为理解蛋白质生物学（例如蛋白质变异致病性）提供了一个有价值的起点。

2023 年 9 月 19 日，来自 Google DeepMind 的 Jun Cheng 等人在 Science 上发表了一篇题为 Accurate proteome-wide missense variant effect prediction with AlphaMissense 的文章。

作者对 AlphaFold2 进行了改造，开发了 AlphaMissense。AlphaMissense 利用了 AlphaFold 的两个关键功能：高度准确的蛋白质结构模型和从相关序列中学习进化约束的能力。

AlphaMissense 分两个阶段进行训练。在第一阶段，像 AlphaFold 一样进行训练，通过预测 MSA 中随机位置掩蔽的氨基酸的身份来执行单链结构预测以及蛋白质语言建模。作者对 AlphaFold 进行了一些小的架构修改，并增加了蛋白质语言建模的损失权重，同时仍实现了与 AlphaFold 相当的结构预测性能。在第二阶段，对人类蛋白质进行微调，并为 MSA 第二行中呈现的变异序列定义了额外的变异致病性分类目标。对于训练集，作者为在人类和灵长类动物群体中经常观察到的变异分配良性标记，并为人类和灵长类动物群体中不存在的变异分配致病标记。一旦模型开始在验证集上过度拟合，训练模型就会停止。

AlphaMissense 在临床注释、新发疾病变异和实验分析基准中实现了最先进的错义致病性预测。

信息来源：

https://mp.weixin.qq.com/s?__biz=MzU2NTA1MzYyOA==&mid=2247550116&idx=1&sn=ec22829f235151fc012c0f7b4c4ce02b&chksm=fc43d577cb345c6148390bcd8f7ba447cee15267b9e20e6249b5ae2b8c4db1799a6f1696ca2&scene=0&xtrack=1

推荐评论：

通过对 AlphaFold2 进行训练和改造，AlphaMissense 能够在临床导向、新发疾病变异和实验分析基准中实现最先进的错义致病性预测。这一进展将有助于更深入地了解蛋白质突变或构象改变与疾病之间的关系，为疾病诊断和治疗提供更多信息。

利用生成式 AI 预测蛋白质的构象变化

2023-10-16

2023 年 10 月 10 日,富士通(Fujitsu Limited)与日本理研计算科学中心(RIKEN Center for Computational Science)的高性能计算与 AI 驱动的药物开发平台部门(HPC-and AI-driven Drug Development Platform Division)宣布,双方已于 2023 年 1 月开发出一项人工智能药物发现技术,该技术可利用生成式人工智能,通过电子显微镜图像的三维密度图大范围预测蛋白质的构象变化。

作为 2022 年 5 月启动的联合研究项目的一部分,富士通和理研开发了一种生成式人工智能技术,该技术可从电子显微镜拍摄的大量投影图像中准确估算目标蛋白质的各种构象形式及其可能的比例,还开发了一种技术,可根据估算比例预测目标蛋白质的构象变化。在这两项技术的基础上,双方开发出了能够大范围预测蛋白质结构变化的人工智能药物发现技术,旨在大幅降低药物发现的开发时间和成本。

该技术能够根据实验数据准确获取蛋白质的构象和变化,所需的时间是传统程序的十倍以上,从而实现了与细菌和病毒等靶蛋白结合的药物设计过程的创新。

今后,富士通和理研将把新开发的生成式人工智能技术作为实现新一代 IT 药物发现技术的核心技术之一,该技术可以分析目标蛋白质和抗体之间的复杂关系,并高精度、高速度地预测分子的全局结构变化。

信息来源:

https://mp.weixin.qq.com/s?__biz=MzU2NTA1MzYyOA==&mid=2247549978&idx=2&sn=6a8ea698e340d2c1829d4fafbc018208&chksm=fc43d5c9cb345cdf0ebfa6d0a366d4af1676a950c48878ba19c2d1f214e03a0fb09fd7497331&scene=0&xtrack=1

推荐评论:

这项技术利用生成式人工智能,能够大范围预测蛋白质的构象变化,这是人工智能在生物医学领域应用的一个重要工具。它不仅能够帮助科研工作者更深入地理解蛋白质的结构和功能,而且还能够为药物设计和开发提供更多的可能性。

从中国传统天然化合物中发现靶向 ADAR2 和 5-HT₂CR RNA 结合环的潜在抑制剂

2023-09-12

作用于 RNA2 的腺苷脱氨酶 (ADAR2) 是一种参与 RNA 编辑过程的重要酶,特别是在 RNA 分子中腺苷转化为肌苷的过程中。ADAR2 活性的失调与各种疾病有关,包括神经系统疾病 (包括精神分裂症)、炎症性疾病、病毒感染和癌症。因此,用小分子靶向 ADAR2 为调节 RNA 编辑和潜在治疗相关病理提供了一种很有前途的治疗策略。但有效抑制 ADAR2 反应的化合物有限,因此,本研究采用计算方法从中药库中筛选天然化合物。本次猎取的化合物对 ADAR2 的结合亲和力 (<-9.5 kcal/mol) 比已知的抑制剂 8-azanebularine (-6.8 kcal/mol) 更强。还观察到亲和力最强的化合物对 5-HT₂CR 具有高结合亲和力,结合能范围为 -7.8 至 -12.9 kcal/mol。

进一步对顶部 ADAR2 配体复合物进行分子动力学模拟和分子力学泊松-玻尔兹曼表面积 (MM/PBSA) 计算显示,包括 ZINC000014637370、ZINC000085593577、ZINC000042890265、ZINC000039183320 和 ZINC000101100339 的五种潜在命中化合物具有 -174.911 、 -137.369 、 -117.236 、 -67.023 和 -64.913 kJ/mol 的同人 ADAR2 蛋白有利结合自由能。推测残基 Lys350、Cys377、Glu396、Cys451、Arg455、Ser486、Gln488 和 Arg510 在配体识别和结合中至关重要。这一发现将为 ADAR2 和小分子之间的分子相互作用提供有价值的见解,有助于设计未来具有潜在治疗应用的 ADAR2 抑制剂。

信息来源: Broni E, et al. INTERNATIONAL JOURNAL OF
MOLECULAR SCIENCES, 2023, 24 (16)

推荐评论:

中药及其天然产物库在药物发现中具有重要价值。通过计算方法筛选中药库,研究人员能够发现一些具有潜在治疗效果的新化合物,这为药物研发提供了新的可能性。根据筛选结果,设计并合成类似物,并在实验中验证他们的生物学功能,是虚拟筛选和实验验证相结合的研究思路。

近期会议与活动

时间	题目	地点	相关链接
25-26 Apr 2024	20th International Congress on Advances In Natural Medicines, Nutraceuticals & Neurocognition	London, United Kingdom	https://www.clocate.com/international-congress-on-advances-in-natural-medicines-nutraceuticals-and-neurocognition/55207/
24-26 Oct 2024	The Global Summit on Agriculture and Plant Science (GSAPS 2024)	Bern, Switzerland	https://www.clocate.com/the-global-summit-on-agriculture-and-plant-science-gsaps/101308/
12-14 Dec 2023	Middle East Organic & Natural Product Expo 2023	Dubai	https://www.clocate.com/middle-east-organic-and-natural-product-expo/18595/
22-24 May 2024	Drug Design & Medicinal Chemistry Congress	Basel, Switzerland	https://www.clocate.com/drug-design-and-medicinal-chemistry-congress/36332/
23-24 May 2024	11th International Conference on Natural, Traditional & Alternative Medicine	Zurich, Switzerland	https://www.clocate.com/international-conference-and-expo-on-natural-traditional-and-alternative-medicine/72721/
2024 年 4 月 12-13 日	2024 第二届合成生物学产业嘉年华暨展览会	南京	https://www.instrument.com.cn/news/20231114/691866.shtml

特别鸣谢（快报信息评论专家）

姓名	单位（按姓氏拼音排序）
丁 骁	中国科学院昆明植物研究所
刘将新	中国科学院昆明植物研究所

人工智能与天然药物动态监测快报

编写委员会

主 编：庄会富 陆颖

执行主编：田雅娟

编 委：张于 颜秉超 胡坤 卿立燕 杨帅

编 辑：杜宁 田雅娟 黄蓉 史继强

工作联系人：杜 宁 电话：0871-65223176 邮箱：duning@mail.kib.ac.cn

田雅娟 电话：028-81258707 邮箱：tyj@clas.ac.cn