

---

# 上海药物所发展出基于知识图谱的计算模型

作者：writer 来源：中国科学院

本文原地址：<https://www.iikx.com/news/progress/29511.html>

本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

## 上海药物所发展出基于知识图谱的计算模型。

9月19日，中国科学院上海药物研究所郑明月课题组在《细胞-基因组学》（*Cell Genomics*）上，发表了题为 *Identifying compound-protein interactions with knowledge graph embedding of perturbation transcriptomics*

的学术论文。该研究提出了基于知识图谱的计算模型PertKGE。这一模型以数据驱动的形式在微扰转录组数据中解耦出化合物-蛋白质相互作用。

化合物和蛋白质是药物发现中两个基本实体，而关于二者相互作用的建模是药物发现的基础。目前，尚无通用的计算方法来预测和解释所有的化合物-蛋白质相互作用，但科研人员可以利用不同的生物数据并从多种研究视角构建化合物-蛋白质相互作用图谱。

在药物发现领域，微扰转录组学连接了药物发现中最重要的实体（化合物）和组学数据，提供了化合物微扰受试体后产生的直观微扰结果，为解耦化合物-蛋白质相互作用（CPI）提供了新视角。然而，由于生物数据的噪声、细胞稳态及转录组表达的动态变化等原因，CPI通常不会直接反应在差异表达基因上，这为微扰转录组数据的分析带来了挑战。

该研究提出了基于知识图谱的PertKGE模型。PertKGE模型受启发于微扰组学分析中的因果推断模型以及知识图谱在药物发现中的应用，从因果视角来构建新型的生物知识图谱，包含原因部分（化合物与受试体中的特定靶标发生CPI）、过程部分（生物系统中多层级的调控事件对CPI信号进行放大）、结果部分（微扰结果以差异表达基因的形式被观测）。在此基础上，研究结合知识图谱嵌入技术来建模因果链接，以达到从差异表达基因中解耦CPI的目的。广泛的计算实验表明，PertKGE在化合物冷启动和靶标冷启动场景中均取得了优异的CPI预测性能，同时，PertKGE通过因果策略引入的多级调控事件缓解了CPI数据集中代表性偏差对模型学习的影响，从而提高了预测的可靠性。

进一步，该研究将PertKGE整合到基于表型和基于靶标的药物发现流程，发现了TNKS抑制剂K-756的新靶标ENPP1，为独特的抗肿瘤免疫效应提供了分子层面的机制解释。研究以10.8%的高命中率为治疗结直肠癌和胰腺癌的新型靶标ALDH1B1筛选到5个具有全新骨架的苗头化合物。

该团队将CPI预测任务对应到药物发现的真实场景中，从计算模拟和实验验证两方面证明了PertKGE在药物发现流程中的作用，为CPI图谱的构建和推理提供了全新策略。

---

研究工作得到国家自然科学基金、国家重点研发计划、中国科学院基础与交叉前沿科研先导专项、中国科学院青年创新促进会会员项目、上海药物所与上海中医药大学中医药创新团队联合研究项目、中国科协青年托举人才工程及上海市科技重大专项的资助，并获得国家蛋白质科学研究（上海）设施规模化蛋白质制备系统的技术支持。

[论文链接](#)

PertKGE工作流

研究团队单位：上海药物研究所

更多 科学进展 请访问 <https://www.iikx.com/news/progress/>

本文版权归原作者所有，请勿用于商业用途，[爱科学iikx.com](#)转发