

• 专题:双清论坛“组合靶标与原创药物” •

DOI: 10.3724/BNSFC-2025.05.14.0002

基于机器学习理论工具推动的多靶点和单靶点组合药物设计方法*

李睿^{1,2} 隆兰心¹ 朱睿迪¹ 吴承蔚¹ 张健^{1,2,3**}

1. 上海交通大学医学院 药物学与人工智能交叉学系,上海 200025
2. 上海交通大学医学院附属瑞金医院 上海交通大学医学院国家转化医学研究中心 医学基因组学国家重点实验室,上海 200025
3. 宁夏医科大学药学院 多肽与蛋白质药物研究中心 六盘山地区药用资源保护与开发利用教育部重点实验室,银川 750004

[摘要] 单靶点单分子干预作为传统药物研发主流,在复杂疾病中逐渐显现疗效不足与易耐药的局限。研究表明,复杂疾病通常由多通路协同异常驱动,单一靶点难以实现持久控制,组合药物因而成为重要突破方向。在多靶点策略上,临床证据显示BRAF/EGFR双抑制可通过合成致死显著提高缓解率,SHP2/MEK联合阻断可克服KRAS突变耐药,PD-1/CTLA-4双免疫检查点抑制显著延长患者生存期,体现出增强疗效与延缓耐药的优势。在方法学层面,人工智能技术显著加速了组合药物发现。基于多数据库构建的多模态特征体系中,XGBoost模型在大规模样本上取得较高预测性能;DECREASE框架结合约束非负矩阵分解与机器学习,在保持准确性的同时大幅减少实验测试规模,降低研发成本。这表明人工智能(Artificial Intelligence,AI)驱动的组合预测已具备实际应用潜力。此外,单靶点内的协同调控策略亦不断发展。例如过氧化物酶体增殖物激活受体 γ 体系中,正构与变构配体的协同作用可同时提升疗效并降低副作用,为代谢性疾病提供新思路。免疫治疗领域更是组合策略应用前沿,当前大量适应症采用免疫检查点抑制剂联合方案,显示出显著临床价值。未来,组合药物研发需加强多学科融合:构建知识图谱解析协同机制,发展多组学整合平台提升模型泛化能力,并解决数据异质性与可解释性问题。通过AI与实验医学深度结合,有望推动组合药物向精准化与机制驱动的新范式发展。

[关键词] 组合药物;人工智能;变构;双位点;药物设计

创新药物研发体系面临多重挑战,在重大疾病干预领域,单一靶点疗法经常会遇到耐药差、药物效果差等临床问题^[1]。新药研发是一个周期长、投入高、风险大的系统工程(图1)。据统计,一个I类新化学实体(New Chemical Entity,NCE)的研发从靶点发现到最终上市通常需要12~15年的时间,平均研发成本超过20~26亿美元,而其临床转化成功率不足10%。在临床阶段,约有40%~50%的候选药物因疗效不足而被迫终止,另有30%

左右因安全性或药代动力学问题失败。如此高的失败率不仅极大地增加了研发的不确定性,也推高了社会整体的医疗成本。多靶点协同调控的组合药物疗法在一些重大疾病干预领域(如癌症)展现出突破性治疗潜力^[2]。该策略通过药代动力学互补、信号通路网络调控及耐药机制干预等多维度协同作用,在心血管疾病、恶性肿瘤及多重耐药菌感染等复杂疾病治疗中展现出显著的临床获益。组合药物可使肿瘤客观缓解率提升,并降低了

收稿日期:2025-05-14;修回日期:2025-09-22

* 本文根据国家自然科学基金委员会第399期“双清论坛”讨论的内容整理。

** 通信作者,Email:Jian.zhang@sjtu.edu.cn

本文受到国家重点研发计划项目(2023YFF1205103)、宁夏回族自治区重点研究与建设计划项目(2022BEG01002)、浙江大学上海高级研究院星夜科学基金项目(SN-ZJU-SIAS-007)和国家自然科学基金项目(22237005,81925034)的资助。

引用格式: 李睿,隆兰心,朱睿迪,等. 基于机器学习理论工具推动的多靶点和单靶点组合药物设计方法. 中国科学基金,2026,40(2):228-238.

Li R,Long LX,Zhu RD,et al. Current status and prospects of combination drug and its artificial intelligence algorithm development. Bulletin of National Natural Science Foundation of China,2026,40(2):228-238.(in Chinese)

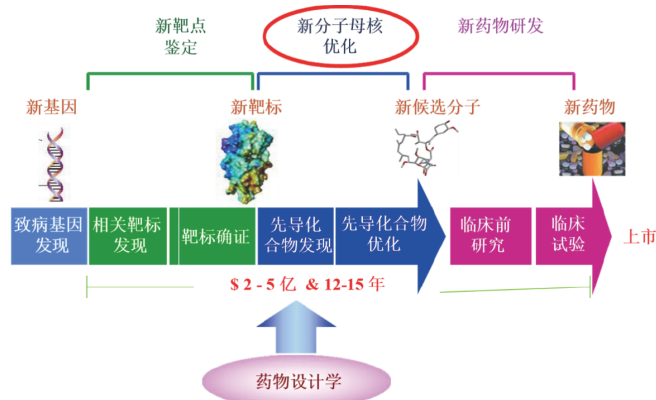


图1 临床前药物研发一般流程

Fig.1 General Process of Preclinical Drug Development

心血管事件复发风险。此外,基于药物相互作用优化的组合方案可降低剂量依赖性毒性,同时减少医疗支出。

一般而言,组合药物可以克服固有性耐药和获得性耐药问题,从而具有更好的干预效果。其中固有性耐药是目前的一类研究热点,即生物体内抑制一条信号通路可能导致第二条通路(或同一通路)出现代偿性增强以维持稳态。此类缓冲机制确保了正常生物系统的稳健性,但在癌症治疗背景下却构成重大挑战。BRAF突变型结直肠癌(Colorectal Cancer, CRC)对BRAF抑制的固有耐药性即一条信号通路的抑制被另一生存信号的激活所代偿。为此,研究者通过RNAi筛选探究了抑制人类基因组中518种激酶中的任意一种是否能在该疾病中与BRAF抑制形成合成致死效应^[3],该筛选及其他研究发现:单独抑制EGFR或BRAF无显著效果^[4],但二者联用则呈现高度致死性。从机制角度分析,这是由于BRAF被抑制时触发了激活EGFR的反馈环路。继概念验证设计试验(采用西妥昔单抗作为EGFR抑制剂、恩考芬尼作为BRAF抑制剂)^[5]及III期试验^[6]成功之后,BRAF与EGFR抑制剂联用方案目前已在美国和欧盟获批用于BRAF突变型结肠癌,这是组合药物通过多靶点协同作用进行癌症干预的代表性临床案例之一。从实验室到临床的案例凸显了耐药机制在发现克服耐药的联合用药方案中的强大潜力。然而,晚期患者最终仍会在BRAF与EGFR抑制剂联用后出现疾病进展^[6]。此外,组合药物也可以克服获得性耐药问题从而展现更好的干预效果。研究人员筛选发现,抑制KRAS突变蛋白下游的MEK激酶会引发MAPK通路信号的快速适应性反馈激活(常通过HER家族或FGFR介导)以限制疗效^[7-9]。尽管在动物模型中,泛HER抑制剂达克替尼与MEK抑制剂联用显示出高效且可接受范围内的毒性效果^[9],但达克替尼联合MEK抑制的I期人体试验却呈现高毒性^[10],提示动物模型对评估联合用药毒性存在局限性。上述对同一信号通路的双重阻断被称为“垂

直通路抑制”。多项临床前研究支持通过MAPK通路双重抑制来克服耐药。同时根据前述合成致死性筛选结果,BRAF或MEK抑制会引发RTKs激活。由于SHP2磷酸酶在几乎所有RTKs下游传递信号至MAPK通路,多个研究团队研究了SHP2抑制剂与MEK抑制剂联用对MAPK信号激活型癌症的疗效,发现该组合在临床前模型中效果显著^[11-13]。针对KRAS突变型癌症中MEK抑制剂内在耐药,MAP2K4(即MKK4)激酶的研究提供了新的联合治疗机遇。总体来讲,针对人类重大疾病的组合药物干预可以克服固有性耐药以及获得性耐药问题,具有很多成熟的临床案例和研究成果。但传统的组合药物研发也是时间、劳动力和成本的综合考验,为确定多种药物在不同浓度范围内的最佳药物组合和药物剂量,传统试错法必须测试多种组合,体系庞大。因此,目前应开发基于AI算法的组合药物研发,加快研究进程。通过组合药物在单一靶点进行组合药物也可以提升药效并克服耐药问题。

综上所述,组合药物可以克服耐药,提升药物干预效果,目前已经有很多成熟的研发案例。传统的组合药物试错筛选法虽能确保组合药物的覆盖度,但其高通量筛选平台需完成指数级实验迭代,需要庞大的时空成本与资源消耗。本综述详细总结了使用AI算法进行组合药物设计所需要的数据库、算法框架和评价指标,以及代表性组合药物研发思路的实际案例,希望可以使用基于AI算法的计算药理学模型,采用智能筛选策略缩短研发周期,并提升组合药物先导化合物的发现效率。通过本文可以为组合药物研发及其AI算法开发寻找更好的组合药物提供帮助,也为组合药物研发攻克肿瘤耐药、神经退行性疾病等复杂病理系统提供新范式。

1 通过AI进行多靶点层面的组合药物研发

单一靶点疗法在实际应用中常面临耐药性与选择性不足等问题,其根本原因在于复杂疾病通常由多条信

号通路协同调控。针对这一特点,基于信号通路的基因上下游关联分析方法可系统性识别具有协同作用的蛋白靶标,并将同一或不同通路中的关键节点进行整合,从而为多靶点联合干预提供理论基础。在此基础上,通过对多个协同靶标实施组合给药,有望实现更优的治疗效果及更持久的耐药控制。因此,多靶点组合策略已成为当前临床与药物研发中的重要方向。在药物开发层面,为实现对上述靶标的有效调控,基于结构的药物设计(Structure-Based Drug Design, SBDD)仍是主流方法。近年来,人工智能技术的引入显著提升了该领域的发展效率。通过对蛋白三维结构的学习,AI模型不仅能够精准识别潜在结合口袋(binding pocket),还可区分正构与变构位点,并辅助设计针对关键靶标的激动剂或拮抗剂,从而加速候选分子的发现与优化过程。相比之下,传统结构解析技术(如X射线晶体学与冷冻电镜)在位点识别中仍存在成本高、通量低等限制,难以满足大规模筛选需求。通过融合物理化学特征分析与机器学习方法,当前计算策略在位点预测的效率与准确性方面均得到显著提升,成为药物设计的重要支撑手段。然而,除单靶点层面的结构识别外,如何在系统层面准确筛选具有协同效应的靶标组合,仍是亟待解决的关键问题。针对这一挑战,当前研究已构建多类用于组合药物预测的数据资源与工具体系,主要可分为两类:一类为药物组合数据库,提供已知组合的标签及多维特征数据;另一类为关联数据库,提供靶标、通路及疾病之间的关联信息与统计特征。上述资源为多靶点组合策略的系统性研究奠定了基础。

药物组合数据库的核心功能在于系统整合已知药物组合信息。本文梳理了当前常用的数据库、网络服务器及软件工具,并总结其对应的算法框架与模型评估表现。这些数据库基于不同标准对药物组合进行分类与量化,可作为机器学习方法开展分类与回归任务的重要基准数据来源。在分类任务中,药物组合通常被标注为“协同、相加或拮抗”,或简化为“有效/无效”;在回归任务中,则通过多种协同效应评分对药物间相互作用进行连续量化描述。DrugComb^[14]、DrugCombDB^[15]、SYNERGxDB^[16]、NCI-ALMANAC^[17]、O'Neil^[18]等研究及AZ-DREAM^[19]主要聚焦抗癌药物组合;而DrugComb、DrugCombDB和SYNERGxDB这3个数据库在药物重叠度与组合实验方面的交集情况除涵盖多种癌症类型外,DrugComb还包含疟疾、COVID-19等其他疾病数据。此外,ASDCD^[20]专门收录抗真菌药物组合信息,而DCDB^[21]和TTD^[22]数据库则广泛整合各类人类疾病的药物组合数据。同时,目前有6个代表性网络服务器(DrugCombDB、SYNERGxDB、NCI-

ALMANAC、DrugR⁺^[23]、ASDCD及SynergyFinder^[24])可以用于该领域,其功能与数据库及软件工具相互补充。用户可通过这些平台检索并可视化特定组织或细胞系中药物组合的协同效应。主流软件工具(包括SynergyFinder、Synergy^[25]、Combenefit^[26]和CompuSyn)基于协同效应模型计算药物组合的协同评分,同时提供效应可视化、分析与量化功能。其他关联数据库则为机器学习模型提供样本特征数据,合理选择特征数据可有效提升模型预测性能、缩短训练时间并降低过拟合风险。这些数据库涵盖以下数据类型:(1)药物化学特征(化学结构、ATC编码、化学—化学相互作用);(2)靶蛋白特征(基因表达、蛋白序列、基因本体术语、通路信息);(3)药理学信息(细胞系药物反应、副作用等)。

研究人员进一步基于机器学习算法和以上数据探究了如何加快组合药物研发进程。机器学习方法展现出多层次技术架构与评估体系,代表性评估指标总结见表1。经典机器学习模型通过特征工程优化与集成策略实现高效预测,其中随机森林(Random Forest, RF)架构在多源数据融合场景表现突出:DCPT平台整合外显子测序、RNA测序及靶点特征,基于HSA模型进行患者特异性协同效应预测,在T-PLL临床数据集($N=3$)中验证了个性化组合方案的有效性^[27];SONAR算法则利用化学遗传互作矩阵(Chemical Genetic Interaction Matrix, CGM)与朴素贝叶斯预测概率构建双层RF分类器,显著提升特征表达能力,达到AUROC 0.91与Bliss评分PCC 0.56的联合指标^[28]。梯度提升框架中,TreeCombo模型基于XGBoost架构处理药物理化特征与基因表达数据,在O'Neil基准库(约22 000样本)中AUROC达0.83,超越同期深度模型DeepSynergy 3.2%,并通过TreeSHAP解析揭示CDKN2A、TP53等关键基因的特征贡献度(权重 > 0.15)。测评方法论层面,针对剂量响应矩阵预测任务,分解机(Factorization Machine, FM)模型在NCI-ALMANAC多模态数据集实现PCC 0.95~0.97的突破性表现^[17],而DECREASE系统通过cNMF-XGBoost混合架构将16×16剂量矩阵的测量点降至10个(成本降低83%)^[29],矩阵重构RMSE < 15%。极端不平衡数据场景中,TLMCS模型采用双层支持向量机(Support Vector

表1 计算效果测评指标总结

Table 1 Summary of Evaluation Indicators for Calculation Effectiveness

评估场景	核心指标	典型阈值范围
二分类任务	AUROC	0.82~0.91(RF模型)
类别不平衡任务	AUPR/F1-score	0.37~0.56(TLMCS)
回归预测任务	皮尔逊相关系数	0.95~0.97(FM模型)
剂量响应预测	均方根误差	矩阵重构误差 < 15%

Machine, SVM)集成策略,在200:1负样本比例下AUPR保持0.372,但揭示出现有方法在超小临床样本(如T-PLL案例 $N=3$)的统计验证瓶颈^[30]。性能验证表明,基于装袋法的RF模型在AZ-DREAM高维数据集^[19]($> 1\,000$ 特征)分类任务中AUROC达0.88,而SGB增强模型通过特征选择策略明确ATC编码相似性对预测的核心贡献(特征重要性权重占比28.6%),凸显机器学习方法在特征可解释性方面的固有优势。最近,有研究人员提出“MultiComb”模型,使用图卷积网络和注意力机制同时预测药物组合的协同效应(synergy)和敏感性(sensitivity)分数^[31]。另外合成致死(Synthetic Lethality, SL)为肿瘤精准治疗提供了重要的理论基础,其通过靶向两个基因的互作关系来选择性杀伤肿瘤细胞。近年来,基于SL的组合疗法逐渐显示出克服耐药性和增强抗肿瘤疗效的潜力。然而,SL相互作用的数量庞大且机制复杂,仅依赖实验手段进行鉴定效率极低。因此,亟需发展基于AI的预测与机制解析方法,以辅助发现更多潜在的SL相互作用并指导组合药物设计。尽管已有部分深度学习与图神经网络模型用于SL预测,但多数模型偏重于“结果预测”,而对预测背后的生物学机制缺乏解释力。这种“黑箱”问题限制了模型在临床转化中的应用。因此,最新研究报道了一种能够结合生物学知识与数据驱动的可解释性AI框架,成为推动SL研究的重要方向^[32],同时有工作建立基于知识图谱的图神经网络(Graph Neural Network, GNN),能够为每一个SL基因对生成多个可信解释,并编码高阶结构(称为DGIB4SL)^[33]。另外也有工作提出ELISL模型,将蛋白序列嵌入与上下文特定的组学数据结合,用森林集成方法进行SL预测,表现出更强鲁棒性和跨癌种推广能力^[34],还有工作系统性评估近年来12个SL预测方法,包括GNN模型,并指出SLMGAE表现最佳,同时提供方法选择指南^[35]。

2 通过AI进行“正构—变构”的单一靶点组合药物研发

单一靶点具有药效差、容易产生耐药性等问题,目前认为其主要原因是单一靶点+单一分子的给药策略造成靶标突变干扰导致分子失效,例如kras耐药。单一靶点+多分子的给药策略可以在理论上规避上述问题,进一步提高药效和耐药效果。因此“正构—变构”的单一靶点给药策略具有研究价值,依据其设计的分子可以同时在一个靶标的两个位点结合,共同产生协同效果。变构(allosteric)位点的产生是因为生理状态下的蛋白结构是柔性的,能够在体内生理环境中表现出固有的灵活运动属性。这样的性质有助于配体与内源性活性位点之外的其他氨基酸口袋产生相互作用,影响蛋白的结构变化,从

而产生变构调节作用,增强或者抑制蛋白功能^[36,37]。从配体结合蛋白口袋的结构分类来说,内源性配体常常结合在蛋白靶点的内源性口袋(正构)位点,而靶点的内源性正构(orthosteric)口袋具有高度的进化保守性,因此针对这些保守(正构)位点的药物化合物常常面临靶点选择性差或多靶点导致副作用多等问题^[38,39]。变构调节提供了一种高专一性潜力的调节蛋白功能的策略,常见gpcr的正构与变构位点见图2。与直接抑制或激活蛋白的活性位点不同,变构调节涉及调节分子在活性位点之外的位点(变构位点)上的相互作用,这些位点通常在空间上与酶活性或内源性配体结合的位置不同。因此,变构调节剂不会直接与活性位点的内源性底物竞争,在单一靶点上更容易进行双位点组合药物设计的策略^[40]。结合这些变构位点通常也会导致蛋白结构的构象变化,这些变化可以激动或抑制蛋白的活性,从而调节其功能。这种变构调节策略为设计相关疾病蛋白的调节剂分子从而进行潜在专一性治疗干预提供了一种有价值并且副作用较少的途径^[41,42],因此导致在正构位点突变时变构位点依然有可能保持活性。总之,精准识别这些变构位点不仅能够揭示潜在的变构结合口袋(allosteric pocket),更为关键单一靶标的位点活性分子发现提供了理论基础和实际应用路径,从而推动针对单一靶标的组合药物的开发进程。通过融合物理化学特征分析与机器学习技术,显著提升了变构位点的预测效率与准确性。AI算法可以加快变构位点发现,结合药物发现技术,进一步加速探索变构分子化学空间,发现先导化合物,为组合药物分子设计研发提供物质基础。基于AI算法可以开发更有潜力的药物分子,主要工作流程见图3。

为了完成单一靶标的“正构—变构”分子设计,首先需要发现该靶标可以结合的变构分子。具体地说,

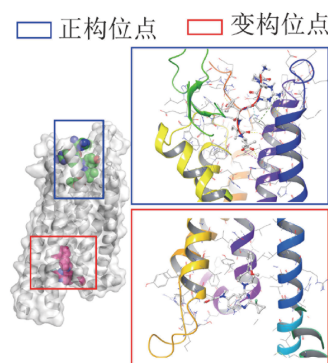


图2 变构效应中的变构位点与正构位点在GPCR蛋白三维空间结构上的分布,PDB编号8JPF

Fig.2 Distribution of Allosteric and Orthosteric Sites in the Three-dimensional Structure of GPCR Protein in the Allosteric Effect, PDB Number 8JPF

TopoAlloSite创新性地构建了基于CATH结构域节点的蛋白质拓扑网络模型,采用支持向量机SVM与子图核匹配算法(Subgraph Kernel Matching, SKM)解析变构位点的空间分布规律^[43]。研究证实,蛋白质拓扑结构是决定变构位点(含隐秘位点)定位的关键因素,其模型在跨蛋白测试中展现出0.71的召回率与0.76的准确率,显著优于传统基于几何特征的方法。该算法成功应用于磷酸泛酰半胱氨酸合成酶(PPC synthase)、亚精胺合成酶(spermidine synthase)及Sirtuin 6等复杂体系的变构位点挖掘,揭示了传统X射线晶体学分析未能检测的3个功能性变构口袋,证实了拓扑特征对功能位点预测的独特价值。然而, PASSerRank提出了基于LambdaMART的集成学习策略,该模型通过梯度提升决策树(Gradient Boosting Decision Tree, GBDT)与LambdaRank函数耦合,实现了变构位点可能性排序预测^[44]。在标准测试集中,其F1分数在ASD数据库达到0.662(MCC=0.645),在CASBench数据库为0.608(MCC=0.589),显示出稳定的跨数据集泛化能力。值得注意的是,该模型在保留测试中优先识别出80%的已知变构位点(Top 5预测精度),表明其在高通量虚拟筛选中具有重要应用潜力。此外,神经网络及其增强方法也在该领域得到了应用。PASSer结合了XGBoost模型和图卷积神经网络(Graph Convolutional Neural Network, GCNN),使用口袋的原子网络表示,并基于FPocket物理

化学特征进行训练,以增强预测能力^[45]。该方法在多个基准测试中表现出色,在ASD、ASBench和CASBench中的F1得分分布在0.6~0.7之间。此外,其Top 3正确率为0.849,进一步验证了其在准确识别变构位点方面的有效性。其他机器学习方法如RF等也被应用于变构位点识别,使用专门工具进行动态和拓扑特征分析。增强型麻雀搜索算法(Enhanced Sparrow Search Algorithm, ESSA)通过整合分子动力学模拟的构象涨落参数与拓扑网络中心性指标,建立基于统计力学原理的预测标准^[46]。在跨平台验证中(数据集II),其成功识别14个变构位点中的10个(成功率71.4%),显著优于PARS(14.3%)和AlloSitePro(57.1%)等机器学习工具。结果表明,动态运动特征本身可作为独立于机器学习的重要预测因子。PrankWeb提出基于配体可及性表面斑块的三维网格聚类算法^[47]:(1)特征提取,将蛋白质表面离散化为0.5 Å分辨率的3D网格,计算各节点的配体结合势能;(2)聚类策略,采用密度峰值聚类(Density Peak Clustering, DPC)算法生成候选结合斑块;(3)性能验证,在COACH420和HOLO4K测试集中分别获得0.720和0.686的Top-N成功率。可以识别3D网格上的可配体点并将其聚类为表面斑块作为预测结合位点。该模型使用CHEN11数据集进行训练,并使用JOINED数据集进行验证。在测试中,使用了COACH420数据库和HOLO4K数据集,Top-N成功率

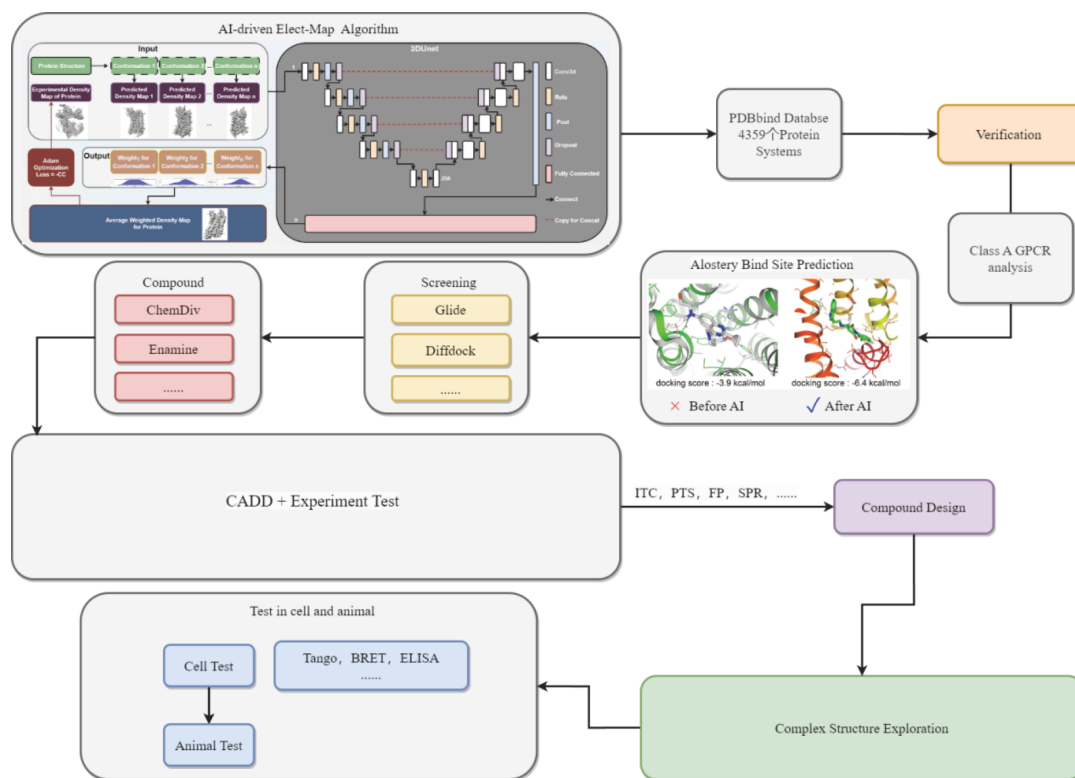


图3 基于AI的变构位点与先导化合物发现算法流程

Fig.3 AI Based Algorithm Flow for Discovering Allosteric Sites and Lead Compounds

分别为0.720和0.686。BiteNet则创新性地应用计算机视觉技术,将蛋白质结构转化为 $128 \times 128 \times 128$ 体素矩阵,通过三维卷积神经网络(3D-CNN)实现变构位点检测^[48]。该模型在相同测试集中的Top-N成功率分别突破0.6与0.7阈值,证实了深度学习在空间特征提取中的优势。该方法的预测能力在COACH420和HOLO4K数据集中进行了评估,其Top-N成功率分别超过0.6和0.7。CorrSite作为CavityPlus平台的一部分,通过结构运动特征和生物物理特性揭示了正构位点和变构位点之间的耦合^[49]。通过引入时间相关运动分析(time-lagged correlation analysis),实现了正变构位点耦合关系的精准解析:在数据集I上,35个变构蛋白中36个已知位点识别出35个(准确率97.2%),覆盖169 917个氨基酸的序列长度范围;在数据集II上,独立测试中预测准确率达90%(20个位点中18个正确识别)。此外,PASSer2.0采用AutoGluon自动化机器学习框架,对FPocket提取的32维物理化学特征进行多模型集成优化:在基准测试中获得0.701的F1分数(较初版提升11.2%);Top 3预测准确率达0.827,表明特征工程的优化对模型性能具有决定性影响^[50]。在增强采样和隐秘口袋方面,PocketMiner通过整合分子动力学(Molecular Dynamics, MD)增强采样、卷积神经网络与几何特征分析,建立了动态隐秘口袋的预测体系^[51]。其训练数据构建具有显著创新性,病毒蛋白数据集包含SARS-CoV-2主要蛋白酶等7种病毒蛋白的长时程MD轨迹(累计模拟时长 > 50 μ s);跨物种验证集包含埃博拉病毒VP35蛋白的构象动态数据及16个配体诱导型隐秘口袋的结构库。模型在PDB来源的独立测试集(39个样本,含24个阳性/4个阴性/11个争议位点)中表现出0.90的总体准确率,其中对临床靶点BCR-ABL1的T315I突变体隐秘口袋的预测与实验验证结果高度一致(RMSD < 1.2 Å)。这些模拟识别了SARS-CoV-2蛋白质组中大多数蛋白质的隐秘口袋、先前模拟的埃博拉病毒VP35蛋白数据,以及一组16个已知配体结合隐秘口袋的蛋白质,这些蛋白质专门为本研究进行了模拟。KeyAlloSite采用直接耦合分析(Direct Coupling Analysis, DCA)算法量化变构残基的进化关联。该方法通过成对比较变构口袋内每个残基与功能位点残基之间的进化耦合分数来推断关键变构残基^[52]。该方法已在典型变构蛋白(如BCR-ABL1、Tar和PDZ3)的先前实验研究中得到验证。AlloReverse通过比较holo/apo状态的副本交换分子动力学(Replica Exchange Molecular Dynamics, REMD)模拟,构建自由能变化矩阵^[53]。该模型的分类测试使用了包含58个蛋白质的测试集进行了基准测试,该测试集包括926个口袋样区域,其中83个被注释为“变构位点”,

该模型的召回率为71.0%。此外,ROC曲线下面积为0.758。这些指标表明了该模型在复杂数据集中有效识别变构位点的能力。总体而言,近年来出现的多类变构位点预测方法在“核心思想”与“适用场景”上各有不同,形成了互补格局。例如,TopoAlloSite基于拓朴网络模型,强调蛋白质结构域的全局联系,其优势在于对隐秘位点具有较强的捕捉能力,但对构象动态变化的描述有限;PASSerRank依托于LambdaMART排序框架,通过集成学习提升跨数据集的稳健性,适合大规模虚拟筛选,但预测结果的可解释性相对较弱;BiteNet借助三维卷积神经网络(3D-CNN)直接从体素化的空间信息中学习复杂特征,能有效识别空间构象关联,但对计算资源依赖较高;而ESSA等方法则突出了分子动力学模拟与拓朴动力学特征的结合,能够捕捉动态构象涨落信息,但在大规模应用中计算成本偏高。与之相对,PrankWeb、CorrSite等工具则更适合快速分析配体可及性和正—变构耦合关系,在功能解释上更具直观性。这些差异说明,不同方法在“预测准确率—可解释性—计算效率”之间存在取舍关系。未来的发展趋势可能在于多策略融合,即在早期使用基于几何或拓朴的轻量级算法进行高通量筛选,再结合深度学习与分子动力学等方法进行精细化解析,从而兼顾预测效率与准确性。这种层次化方法学框架将有望推动变构位点预测向更系统化与实用化的方向发展。

总之,靶点的定义已不再局限于单一蛋白整体,不同口袋位点亦可视为独立的“亚靶点”。在此基础上,“正构—变构”策略可视为在同一蛋白上进行双靶点协同干预,从而实现药效增强与耐药缓解的双重目标。通过“正构—变构”组合药物策略可以进一步提升药效,减少耐药发生(图4)。在具体的临床前研究案例中,研究团队深入探究了过氧化物酶体增殖物激活受体 γ (Peroxisome Proliferator-Activated Receptor γ , PPAR γ)的变构激活位点及其在抗糖尿病治疗中的潜力,发现天然产物巴瓦西宁(Bavaclicin, BVC)能够与低剂量的罗格列酮(Rosiglitazone, RSG)协同作用,显著激活PPAR γ ,从而增强胰岛素敏感性,同时有效减轻RSG单药治疗所引起的体重增加和肝毒性等副作用^[54]。案例1通过一系列实验,包括细胞水平的转录报告基因实验、动物模型的代谢分析、核磁共振(Nuclear Magnetic Resonance, NMR)与分子对接技术的结构分析、分子动力学模拟以及基于变构位点的虚拟筛选,研究人员揭示了BVC与RSG共结合PPAR γ 的变构结合位点,这一结合关键位点位于螺旋5的Met329和Ser332附近,能够触发PPAR γ 的变构激活,进而增强RSG的活性。在细胞实验中,BVC展现出剂量依赖性地增强RSG激活PPAR γ 转录活性的能

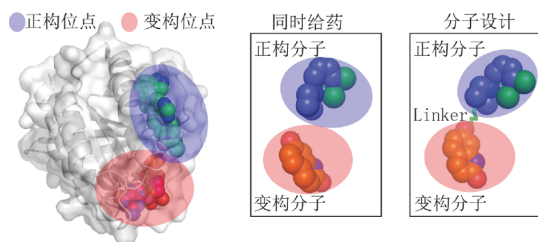


图4 “正构—变构”组合药物与双位点分子设计

Fig.4 “Orthosteric-allosteric” Combination Drug and Dual Site Molecular Design

力。在动物实验中,BVC与低剂量RSG的组合用药显著改善了糖尿病小鼠的葡萄糖耐受性、胰岛素耐受性和血清甘油三酯水平,同时有效抑制了RSG引起的体重增加和肝毒性。分子动力学模拟研究表明,BVC的结合能够稳定PPAR γ 的AF-2表面,使其保持在更活跃的构象,从而增强RSG的活性。基于这一变构热点,研究人员通过虚拟筛选发现了一系列新的PPAR γ 激动剂,这些激动剂能够与RSG协同作用,增强PPAR γ 的活性,为2型糖尿病治疗提供潜在的新药物候选。这些发现不仅证明了BVC与RSG的协同作用在增强胰岛素敏感性和减轻副作用方面的潜力,还揭示了PPAR γ 的一个新的变构激活位点,为开发新型抗糖尿病药物开辟了新的途径。

其次,转移性结肠癌作为结直肠癌致死的主要病理类型,其发生发展与腺瘤性息肉病基因(Adenomatous Polyposis Coli, APC)突变介导的异常信号传导密切相关。病理条件下产生的截短型APC蛋白通过特异性结合Rho鸟苷酸交换因子Asef(APC-Asef Interaction),持续激活Rho GTPase信号通路,进而驱动肿瘤细胞增殖与侵袭转移。针对这一关键致癌互作界面,研究团队前期成功开发了首个靶向APC-Asef蛋白相互作用的抑制剂MAIT-203^[55],但其药效优化因缺乏高灵敏度的动态互作评价体系而遭遇瓶颈。为突破技术障碍,案例2创新性地提出“正构—变构”双位点分子协同调控策略,成功构建了双位点示踪探针(dual-site tracer)^[56]。该探针系统采用模块化设计,配体模块精确锚定APC蛋白的正构结合位点,而荧光模块不仅实现实时信号输出,还可通过特异性结合邻近变构位点产生协同结合效应。经表面等离子体共振(Surface Plasmon Resonance, SPR)验证,该设计使探针结合亲和力提升,检测动态范围显著扩展,从而显著提升了弱效化合物的筛选分辨率。案例2通过“正构—变构”双位点策略开发了组合药物分子,为开发基于肿瘤基因分型的精准治疗策略提供了重要候选分子。

另外,研究人员将该设计理念拓展至G蛋白偶联受体(G Protein-Coupled Receptor, GPCR)靶点开发领域。聚焦于毒蕈碱型乙酰胆碱受体M1亚型(M1 Receptor,

M1R),该受体作为中枢神经系统乙酰胆碱信号传导的核心介质,是治疗阿尔茨海默病、精神分裂症等神经精神疾病的重要药物靶标。然而,由于GPCR家族成员间正构结合域的高度保守性,传统单一位点激动剂普遍存在亚型选择性不足(M1R vs M2R-M5R)与功能活性难以平衡的科学难题。案例3采用结构导向的合理药物设计策略,成功构建具有双位点协同作用的新型分子BAAX^[57]。该分子通过独特的双功能结构域实现精准调控:其正构结合模块通过氢键网络与M1R跨膜区的保守残基形成特异性相互作用,而延伸的变构调节模块则选择性识别M1R特有的胞外环构象口袋。通过等温滴定量热(Isothermal Titration Calorimetry, ITC)与分子动力学模拟分析发现,双位点协同结合诱导M1R发生独特的“双锁扣”构象重排,该构象变化不仅显著增强受体激活效率,更通过空间位阻效应阻断其对M2R-M5R的交叉反应。体外功能实验证实,BAAX对M1R的选择性指数较传统正构激动剂显著提升,且可有效逆转东莨菪碱诱导的认知功能障碍模型小鼠记忆缺失,而未出现传统药物常见的胆碱能副作用。本研究将“正构—变构”双靶向策略成功应用于GPCR受体药物开发,揭示了跨膜受体变构调控的新机制。通过建立“功能活性—亚型选择性”的协同优化范式,为破解GPCR家族药物开发中的选择性困境提供了创新解决方案,也为神经退行性疾病的精准治疗提供了具有转化潜力的候选分子。

3 组合药物策略的代表性临床案例

免疫检查点阻断(Immune Checkpoint Blockade, ICB)疗法堪称癌症治疗领域最具突破性的进展之一。目前,PD1、PDL1及CTLA4抑制剂已获批用于多种癌症治疗,通过激活T细胞杀伤肿瘤细胞。ICB疗法尤为显著的特征在于多数干预后患者可获长期生存获益。例如,约20%接受CTLA4抗体伊匹木单抗治疗的转移性黑色素瘤患者生存期超过5年^[58]。潜在机制可能与T细胞杀伤肿瘤细胞时产生的“伴随效应”相关,该效应甚至可清除邻近抗原阴性细胞^[59],这可解释为何ICB获得性耐药的发生通常晚于靶向治疗。鉴于肿瘤内异质性仍是有效治疗晚期癌症的最大障碍,设计具有伴随效应的疗法至关重要。尽管存在此类效应,ICB获得性耐药仍可能发生,PD1疗法耐药的黑色素瘤中已观察到B2M、JAK1及JAK2基因突变^[60]。此外,ICB内在耐药仍是主要限制因素,亟需探索基于机制设计的联合疗法以惠及更多患者。当前针对ICB联合疗法的探索呈现爆发式增长,美国临床试验注册平台数据显示的已获批的ICB疗法见图5。截至2022年8月,获批的34项ICB适应症中,18项采

用化疗联合方案,10项为双免疫检查点阻断策略(即2种ICB疗法联用)。需要强调的是,在三阴性乳腺癌模型中,各类化疗药物与抗PD1治疗的协同程度呈现明显差异,这种差异不知是否与化疗诱导衰老能力的强弱相关,这些发现提示未来需建立更精细的化疗—ICB配伍选择标准,以突破当前经验性联合治疗模式的局限性。

除肿瘤领域外,组合药物策略在其他疾病领域同样

具有广泛应用。在心血管疾病中,不同机制的降压药(如ACEI/ARB、钙通道阻滞剂、利尿剂)的联合用药已成为国际指南推荐的一线策略,可在改善降压疗效的同时降低药物相关不良反应。在神经退行性疾病方面,尽管目前尚无根本性治愈手段,但已有研究探索联合用药以多维度干预疾病进程。例如,胆碱酯酶抑制剂与NMDA受体拮抗剂联合治疗阿尔茨海默病,以期同时改善认知功

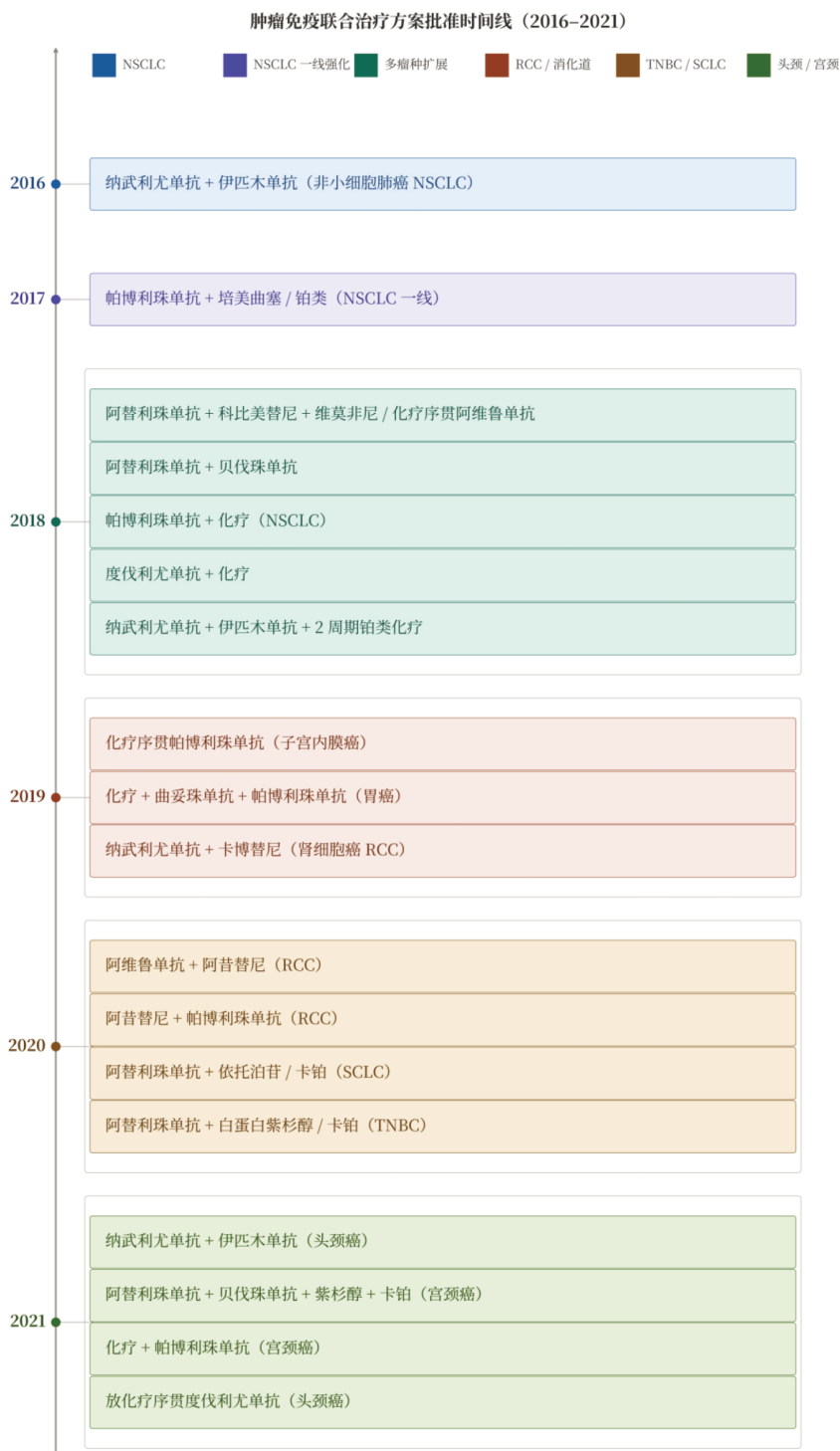


图5 已经批准的不同病症组合用药策略

Fig.5 Approved Combination Medication Strategies for Different Diseases

能与延缓神经元损伤。这些跨领域探索表明,组合治疗作为应对复杂疾病多通路、多靶点病理机制的重要手段,正在逐步从经验性用药走向机制驱动和精准化设计。

4 结论与展望

尽管AI在药物性质预测和分子优化方面取得了快速进展,但现有研究仍存在若干不足。第一,数据层面的异质性与标准化缺失导致模型在跨数据集和跨化学空间时性能波动,制约了预测结果的可重复性与临床转化价值。第二,模型层面的可解释性瓶颈依旧突出,部分“毒性亚结构”的标注与化学直觉相悖,难以为药物化学家提供可靠的机制启示。第三,泛化性不足的问题在面临结构新颖或复杂的化合物(如大环分子、共价抑制剂)时尤为明显,使得当前方法难以覆盖多样化的药物化学空间。面向未来,AI在组合药物研究中展现出重要潜力。通过引入实验—计算闭环体系,科研人员可在预测与高通量验证之间建立反馈机制,提升模型的动态适应能力;多模态数据融合则有助于同时整合分子结构、药效学、药代动力学与临床真实世界数据,从而实现对药物间协同与拮抗作用的系统性建模;基于强化学习或元学习的动态优化策略有望实现对活性、类药性与毒性等多目标性质的同步优化,推动更为安全高效的组合药物方案设计。因此,未来的研究应重点关注以下几个方面:(1)建立统一的数据标准与开放共享平台,以缓解数据碎片化带来的模型不稳定性;(2)发展基于因果推理的可解释性方法,提升预测结果的可信度与实用性;(3)加强跨领域迁移与跨模态融合,以增强模型对复杂分子类型与多药联合情境的泛化能力。只有在上述方向持续突破,AI才能真正成为推动药物研发范式转型的核心动力。

参 考 文 献

- [1] Goodman LS, Wintrobe MM, Dameshek W, et al. Nitrogen mustard therapy. Use of methyl-bis (beta-chloroethyl) amine hydrochloride and tris (beta-chloroethyl) amine hydrochloride for Hodgkin's disease, lymphosarcoma, leukemia and certain allied and miscellaneous disorders. *Journal of the American Medical Association*, 1946, 132: 126—132.
- [2] Shafer RW, Vavrek DA. Highly active antiretroviral therapy (HAART) for the treatment of infection with human immunodeficiency virus type 1. *Biomedicine & Pharmacotherapy*, 1999, 53: 73—86.
- [3] Prahallad A, Sun C, Huang SD, et al. Unresponsiveness of colon cancer to BRAF(V600E) inhibition through feedback activation of EGFR. *Nature*, 2012, 483(7387): 100—103.
- [4] Corcoran RB, Ebi H, Turke AB, et al. EGFR-mediated re-activation of MAPK signaling contributes to insensitivity of BRAF mutant colorectal cancers to RAF inhibition with vemurafenib. *Cancer Discovery*, 2012, 2(3): 227—235.
- [5] van Geel R, Tabernero J, Elez E, et al. A phase I study of the BRAF inhibitor encorafenib combined with cetuximab ± alpelisib. *Cancer Discovery*, 2017, 7: 610—619.
- [6] Tabernero J, Grothey A, Van Cutsem E, et al. Encorafenib plus cetuximab as a new standard of care for previously Treated-BRAFV600E-mutant metastatic colorectal cancer: Updated survival results and subgroup analyses from the BEACON study. *Journal of Clinical Oncology*, 2021, 39(4): 273—284.
- [7] Manchado E, Weissmueller S, Morris JP, et al. A combinatorial strategy for treating KRAS-mutant lung cancer. *Nature*, 2016, 534(7609): 647—651.
- [8] Kitai H, Ebi H, Tomida S, et al. Epithelial-to-mesenchymal transition defines feedback activation of receptor tyrosine kinase signaling induced by MEK inhibition in KRAS-mutant lung cancer. *Cancer Discovery*, 2016, 6(7): 754—769.
- [9] Sun C, Hobor S, Bertotti A, et al. Intrinsic resistance to MEK inhibition in KRAS mutant lung and colon cancer through transcriptional induction of ERBB3. *Cell Reports*, 2014, 7(1): 86—93.
- [10] van Geel RMJM, van Brummelen EMJ, Eskens FALM, et al. Phase 1 study of the pan-HER inhibitor dacomitinib plus the MEK1/2 inhibitor PD-0325901 in patients with KRAS-mutation-positive colorectal, non-small-cell lung and pancreatic cancer. *British Journal of Cancer*, 2020, 122(8): 1166—1174.
- [11] Mainardi S, Mulero-Sánchez A, Prahallad A, et al. SHP2 is required for growth of KRAS-mutant non-small-cell lung cancer *in vivo*. *Nature Medicine*, 2018, 24(7): 961—967.
- [12] Wong GS, Zhou J, Liu JB, et al. Targeting wild-type KRAS-amplified gastroesophageal cancer through combined MEK and SHP2 inhibition. *Nature Medicine*, 2018, 24(7): 968—977.
- [13] Ruess DA, Heynen GJ, Ciecieski KJ, et al. Mutant KRAS-driven cancers depend on PTPN11/SHP2 phosphatase. *Nature Medicine*, 2018, 24(7): 954—960.
- [14] Zheng SY, Aldahdooh J, Shadbahr T, et al. DrugComb update: A more comprehensive drug sensitivity data repository and analysis portal. *Nucleic Acids Research*, 2021, 49(W1): W174—W184.
- [15] Liu H, Zhang WH, Zou B, et al. DrugCombDB: A comprehensive database of drug combinations toward the discovery of combinatorial therapy. *Nucleic Acids Research*, 2020, 48(D1): D871—D881.
- [16] Seo H, Tkachuk D, Ho C, et al. SYNERGxDB: An integrative pharmacogenomic portal to identify synergistic drug combinations for precision oncology. *Nucleic Acids Research*, 2020, 48(W1): W494—W501.
- [17] Holbeck SL, Camalier R, Crowell JA, et al. The national cancer institute ALMANAC: A comprehensive screening resource for the detection of anticancer drug pairs with enhanced therapeutic activity. *Cancer Research*, 2017, 77(13): 3564—3576.
- [18] O'Neil J, Benita Y, Feldman I, et al. An unbiased oncology compound screen to identify novel combination strategies. *Molecular Cancer Therapeutics*, 2016, 15(6): 1155—1162.
- [19] Menden MP, Wang D, Mason MJ, et al. Community assessment to advance computational prediction of cancer drug combinations in a pharmacogenomic screen. *Nature Communications*, 2019, 10: 2674.
- [20] Chen X, Ren B, Chen M, et al. NLLSS: Predicting synergistic drug combinations based on semi-supervised learning. *PLoS Computational*

- Biology, 2016, 12: e1004975.
- [21] Liu YB, Wei Q, Yu GS, et al. DCDB 2.0: A major update of the drug combination database. Database, 2014, 2014: bau124.
- [22] Li YH, Yu CY, Li XX, et al. Therapeutic target database update 2018: Enriched resource for facilitating bench-to-clinic research of targeted therapeutics. Nucleic Acids Research, 2018, 46(D1): D1121—D1127.
- [23] Masoudi-Sobhanzadeh Y, Omididi Y, Amanlou M, et al. DrugR+: A comprehensive relational database for drug repurposing, combination therapy, and replacement therapy. Computers in Biology and Medicine, 2019, 109: 254—262.
- [24] He LY, Kuleskiy E, Saarela J, et al. Methods for high-throughput drug combination screening and synergy scoring. Methods in Molecular Biology, 2018, 1711: 351—398.
- [25] Zheng SY, Wang WY, Aldahdooh J, et al. SynergyFinder plus: Toward better interpretation and annotation of drug combination screening datasets. Genomics, Proteomics & Bioinformatics, 2022, 20(3): 587—596.
- [26] Di Veroli GY, Fornari C, Wang D, et al. CombeneFit: An interactive platform for the analysis and visualization of drug combinations. Bioinformatics, 2016, 32(18): 2866—2868.
- [27] He LY, Tang J, Andersson EI, et al. Patient-customized drug combination prediction and testing for T-cell prolymphocytic leukemia patients. Cancer Research, 2018, 78(9): 2407—2418.
- [28] Wildenhain J, Spitzer M, Dolma S, et al. Prediction of synergism from chemical-genetic interactions by machine learning. Cell Systems, 2015, 1(6): 383—395.
- [29] Ianevski A, Giri AK, Gautam P, et al. Prediction of drug combination effects with a minimal set of experiments. Nature Machine Intelligence, 2019, 1(12): 568—577.
- [30] Shi JY, Li JX, Mao KT, et al. Predicting drug synergism using deep learning. Computer Methods and Programs in Biomedicine, 2019, 168: 1—10.
- [31] Monem S, Hassanien AE, Abdel-Hamid AH. A multi-task graph deep learning model to predict drugs combination of synergy and sensitivity scores. BMC Bioinformatics, 2024, 25: 327.
- [32] Wang J, Wen Y, Zhang Y, et al. Deep learning-based prediction of synergistic drug combinations. Journal of Advanced Research, 2024, 65: 329—343.
- [33] Chen X, Cai R, Huang Z, et al. Network-based deep learning for drug combination prediction. Briefings in Bioinformatics, 2025, 26: bbaf524.
- [34] Tepeli YI, Seale C, Goncalves JP. Interpretable machine learning for drug synergy prediction. Bioinformatics, 2024, 40: btad764.
- [35] Feng YM, Long YH, Wang H, et al. Benchmarking machine learning methods for synthetic lethality prediction in cancer. Nature Communications, 2024, 15: 9058.
- [36] Weng C, Faure AJ, Escobedo A, et al. Systematic identification of genetic interactions across human cell lines. Nature, 2024, 626: 643—652.
- [37] Wu N, Barahona M, Yaliraki SN. Multiscale network models of protein structure and dynamics. Current Opinion in Structural Biology, 2024, 84: 102737.
- [38] Lu SY, Li S, Zhang J. Harnessing allostery: A novel approach to drug discovery. Medicinal Research Reviews, 2014, 34(6): 1242—1285.
- [39] Gabizon R, Friedler A. Allosteric modulation of protein oligomerization: An emerging approach to drug design. Frontiers in Chemistry, 2014, 2: 9.
- [40] May LT, Leach K, Sexton PM, Christopoulos A. Allosteric modulation of G protein-coupled receptors. Annual Review of Pharmacology and Toxicology, 2007, 47(1): 1—51.
- [41] Nussinov R, Tsai CJ. Allostery in disease and in drug discovery. Cell, 2013, 153(2): 293—305.
- [42] Ferraro M, D'Annessa I, Moroni E, et al. Allosteric modulators of protein-protein interactions. Journal of Medicinal Chemistry, 2019, 62: 60—87.
- [43] Xie J, Pan GX, Li YB, et al. How protein topology controls allosteric regulations. The Journal of Chemical Physics, 2023, 158(10): 105102.
- [44] Tian H, Xiao SA, Jiang X, et al. PASSerRank: Prediction of allosteric sites with learning to rank. Journal of Computational Chemistry, 2023, 44(28): 2223—2229.
- [45] Tian H, Xiao SA, Jiang X, et al. PASSer: Fast and accurate prediction of protein allosteric sites. Nucleic Acids Research, 2023, 51(W1): W427—W431.
- [46] Kaynak BT, Bahar I, Doruker P. Essential site scanning analysis: A new approach for detecting sites that modulate the dispersion of protein global motions. Computational and Structural Biotechnology Journal, 2020, 18: 1577—1586.
- [47] Jakubec D, Skoda P, Krivak R, et al. PrankWeb 3: Accelerated ligand-binding site predictions for experimental and modelled protein structures. Nucleic Acids Research, 2022, 50(W1): W593—W597.
- [48] Kozlovskii I, Popov P. Prediction of allosteric sites using normal mode analysis. Communications Biology, 2020, 3: 618.
- [49] Xie J, Wang S, Xu Y, et al. Allosteric site prediction via ensemble learning. Journal of Chemical Information and Modeling, 2022, 62: 187—195.
- [50] Xiao S, Tian H, Tao P. Advances in computational allosteric drug discovery. Frontiers in Molecular Biosciences, 2022, 9: 879251.
- [51] Meller, A., Ward, M., Borowsky, J. et al. Predicting locations of cryptic pockets from single protein structures using the PocketMiner graph neural network. Nat Commun 2023, 14: 1177.
- [52] Xie J, Zhang W, Zhu X, et al. Global analysis of protein allosteric regulation. eLife, 2023, 12: e83700.
- [53] Zha J, Li Q, Liu X, et al. Allosteric Database 3.0 (ASD 2023): Towards the integrating landscapes of allosteric knowledgebase. Nucleic Acids Research, 2023, 51(W1): W33—W38.
- [54] Feng L, Lu S, Zheng Z, et al. Computational identification of druggable allosteric sites. Science Bulletin, 2021, 66: 1559—1570.
- [55] Jiang H, Deng R, Yang X, et al. Targeting dynamic protein allostery. Nature Chemical Biology, 2017, 13: 994—1001.
- [56] Zhong J, Guo Y, Lu S, et al. Allosteric regulation mechanisms revealed by structural dynamics. Nature Communications, 2022, 13: 4961.
- [57] Zhao N, Wu W, Wang Y, et al. Chemical modulation of protein allostery. Chem, 2023, 9: 901—923.
- [58] Schadendorf D, Hodi FS, Robert C, et al. Pooled analysis of long-term survival data from phase II and phase III trials of ipilimumab in unresectable or metastatic melanoma. Journal of Clinical Oncology, 2015, 33(17): 1889—1894.
- [59] Hoekstra ME, Bornes L, Dijkgraaf FE, et al. Long-distance modulation of bystander tumor cells by CD8⁺ T cell-secreted IFN γ . Nature Cancer, 2020, 1(3): 291—301.
- [60] Zaretsky JM, Garcia-Diaz A, Shin DS, et al. Mutations associated with acquired resistance to PD-1 blockade in melanoma. The New England Journal of Medicine, 2016, 375(9): 819—829.

Current Status and Prospects of Combination Drug and Its Artificial Intelligence Algorithm Development

Rui Li^{1, 2} Lanxin Long¹ Ruidi Zhu¹ Chengwei Wu¹ Jian Zhang^{1, 2, 3*}

1. Department of Pharmaceutical and Artificial-Intelligence Sciences, Institute of Medical Artificial Intelligence, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China

2. State Key Laboratory of Medical Genomics, National Research Center for Translational Medicine at Shanghai, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China

3. Key Laboratory of Protection, Development and Utilization of Medicinal Resources in Liupanshan Area, Ministry of Education, Peptides & Protein Drug Research Center, School of Pharmacy, Ningxia Medical University, Yinchuan 750004, China

Abstract Traditional single-target, single-molecule therapies have shown clear limitations in the treatment of complex diseases, often yielding insufficient efficacy and rapid resistance. Increasing evidence indicates that such diseases are driven by the coordinated dysregulation of multiple signaling pathways, making combination strategies an essential approach. Clinically, rationally designed combinations have demonstrated significant benefits: Dual BRAF/EGFR inhibition enhances objective response rates through synthetic lethality, SHP2/MEK co-inhibition overcomes KRAS mutation—driven resistance, and PD1/CTLA4 blockade markedly prolongs survival in metastatic melanoma. Advances in artificial intelligence (AI) are accelerating combination drug discovery. Leveraging large-scale databases such as DrugComb, predictive models like XGBoost achieve high accuracy (AUROC 0.83 across 22 000 samples), while systems such as DECREASE reduce experimental dose—response testing by over 80% without compromising precision, substantially lowering time and cost. Beyond multi-target applications, innovative strategies are also emerging at the single-target level. For example, orthosteric—allosteric dual-site modulation of PPAR γ synergistically enhances efficacy while mitigating side effects, offering new directions for metabolic disease interventions. Immunotherapy has become a leading clinical frontier, with nearly half of approved immune checkpoint inhibitor indications involving combination regimens—either with chemotherapy, targeted therapies, or dual checkpoint blockade—underscoring the value of immune co-regulation. Looking forward, the development of combination therapies requires multidisciplinary integration. Knowledge graph—based approaches can uncover novel synergistic mechanisms, while multi-omics—driven computational platforms combining structural biology, pharmacology, and clinical data will enhance model generalizability. Addressing challenges such as data heterogeneity and limited interpretability through standardized frameworks will be critical for translation from high-throughput prediction to mechanistic insight. The convergence of AI and experimental medicine promises to shorten development cycles, reduce clinical attrition, and drive the evolution of combination drug discovery toward a precision- and mechanism-based paradigm.

Keywords combination drugs; artificial intelligence; allosteric; dual site; drug design

张 健 教授,上海交通大学医学人工智能研究院副院长,六盘山区域药用资源保护与开发利用教育部重点实验室主任,兼任宁夏医科大学学术副校长、国务院学位委员会学科评议组成员,英国皇家化学会会士,BT-IT国家重点研发计划项目及国家创新药物重大专项项目首席科学家。主要从事人工智能驱动的精准靶标识别和First-in-class原创药物先导发现研究。主持国家自然科学基金青年科学基金项目(A类)等项目。

李 睿 博士后,助理研究员。2024年获得中国药科大学与中国科学院上海药物研究所联合培养的药物化学博士学位。以第一作者或通讯作者在*Journal of Medicinal Chemistry*、*Journal of Chemical Information and Modeling*(封面)等期刊发表论文5篇;在*Nature Communications*、*Cell System*等期刊合作发表论文12篇。主要研究方向为AI药物设计、蛋白动态构象采样算法开发以及先导化合物活性机制分析与结构改造。

(责任编辑 王 磊 张 强)

* Corresponding Author, Email: Jian.zhang@sjtu.edu.cn