

中医药生物信息学数据库：从数据整合到知识发现

王豫骞¹, 樊薛津¹, 金博文¹, 尹清晟¹, 张艳军^{1,2,3*}, 庄朋伟^{1,2,3*}

1. 天津中医药大学中药学院, 天津 301617

2. 天津中医药大学第一附属医院, 天津 300193

3. 现代中医药海河实验室, 天津 301617

摘要: 中医药体系庞大而复杂, 数据库为大量信息化数据的积累应用提供有效可靠方法。随着与网络药理学、多组学及深度学习技术的整合, 传统中医药的复杂体系和创新性发展不断取得突破。系统回顾了中医药生物信息学相关数据库的发展历程, 提出了基于功能与应用场景的 4 级分类体系, 涵盖基础资源、应用领域、技术方法及文化传承 4 大维度。通过分析代表性数据库的优劣势发现, 现有数据库在数据标准化、更新频率及临床验证方面仍存在显著不足, 并通过结合人工智能与多组学技术的融合趋势, 展望了中医药生物信息学数据库在精准医疗与国际化发展中的潜力。

关键词: 中医药; 数据库; 网络药理学; 人工智能; 深度学习

中图分类号: TP18; R285 **文献标志码:** A **文章编号:** 0253-2670(2025)17-6450-13

DOI: 10.7501/j.issn.0253-2670.2025.17.032

Bioinformatics databases for traditional Chinese medicine: Harnessing data integration toward knowledge discovery

WANG Yuqian¹, FAN Xuejin¹, JIN Bowen¹, YIN Qingsheng¹, ZHANG Yanjun^{1,2,3}, ZHUANG Pengwei^{1,2,3}

1. School of Chinese Materia Medica, Tianjin University of Traditional Chinese Medicine, Tianjin 301617, China

2. First Teaching Hospital of Tianjin University of Traditional Chinese Medicine, Tianjin 300193, China

3. Haihe Laboratory of Modern Chinese Medicine, Tianjin 301617, China

Abstract: The multifaceted and intricate system of traditional Chinese medicine (TCM) necessitates robust methodologies for managing its extensive knowledge base. Databases have emerged as an effective and reliable method for the systematic aggregation and utilization of large-scale informational datasets. With the integration of network pharmacology, multi-omics, and deep learning technologies, continuous progress has been made in unraveling TCM's intricate mechanisms and fostering innovative development. This study conducts a comprehensive review of the evolutionary trajectory of TCM bioinformatics databases and introduces a functionally stratified classification system encompassing four dimensions: foundational resources, application domains, technical methodologies, and cultural preservation. Comparative assessment of representative databases reveals significant limitations in data standardization, update sustainability, and clinical validation. Finally, by highlighting the convergence of artificial intelligence (AI) and multi-omics technologies, we envision the transformative potential of TCM bioinformatics databases to advance precision medicine and promote global integration.

Key words: traditional Chinese medicine; database; network pharmacology; artificial intelligence; deep-learning

中医药是中华民族数千年来疾病防治实践的经验总结, 其“整体观”与“辨证论治”理论体系与西方还原论医学形成鲜明互补^[1]; 依据中医药理论指导, 通过天然药物及其加工产物发挥对疾病的

预防、治疗、诊断、康复及保健功能^[2]。随着中医药国际化进程的推进, 全球学者对其研究日益深入, 尤其在 1978 年青蒿素的发现以后, 各国科学家开始对中药中多个活性成分及相应靶点进行探索^[3-4]。

收稿日期: 2025-04-29

基金项目: 天津市科技局计划项目 (24ZYJDSY00280); 国家自然科学基金青年项目 (82304909)

作者简介: 王豫骞, 硕士研究生。E-mail: gisellewangyq@163.com

*通信作者: 庄朋伟, 教授, 博士生导师。E-mail: zhuangpengwei@163.com

张艳军, 教授, 博士生导师。E-mail: zyjsunye@163.com

三氧化二砷治疗急性早幼粒细胞白血病、小檗碱治疗 2 型糖尿病、宣肺败毒方等中药方剂治疗新型冠状病毒肺炎等也体现了中医药在治疗慢性病、难治病等疾病的巨大开发潜力^[5-6]。然而，中医药复方多成分作用、证候动态演变等知识体系的复杂性长期制约其现代化进程。21 世纪以来，生物信息学与人工智能（artificial intelligence, AI）技术的快速发展，催生了以数据库为核心的中医药知识工程研究范式。据统计 2023 年已有 1 764 个生物数据库公开在线^[7]。这些数据库不仅为中医药研究提供了海量信息支持，更通过整合中药化学成分、靶点及其生物学效应，助力研究者系统解析中药复杂作用机制。例如，京都基因与基因组百科全书（Kyoto

encyclopedia of genes and genomes, KEGG）、Reactome 等权威数据库为中药化学成分的靶点预测与机制研究奠定了数据基础。通过系统化整合药材、成分、靶点、疾病等多维度数据，中医药生物信息学数据库不仅为机制探索提供支撑，更成为连接传统经验与现代科学的桥梁^[8]。本文旨在系统梳理中医药生物信息学数据库的发展现状，评估其优势与局限，探讨其在基础研究、临床应用中的关键作用（图 1）。通过构建基础资源、应用领域、技术方法及文化传承数据库的分类框架评述技术应用，为研究者、临床医师及相关从业者提供清晰的参考体系，推动中医药生物信息学数据库的标准化建设与优化升级，助力中医药现代化与国际化的协同发展。

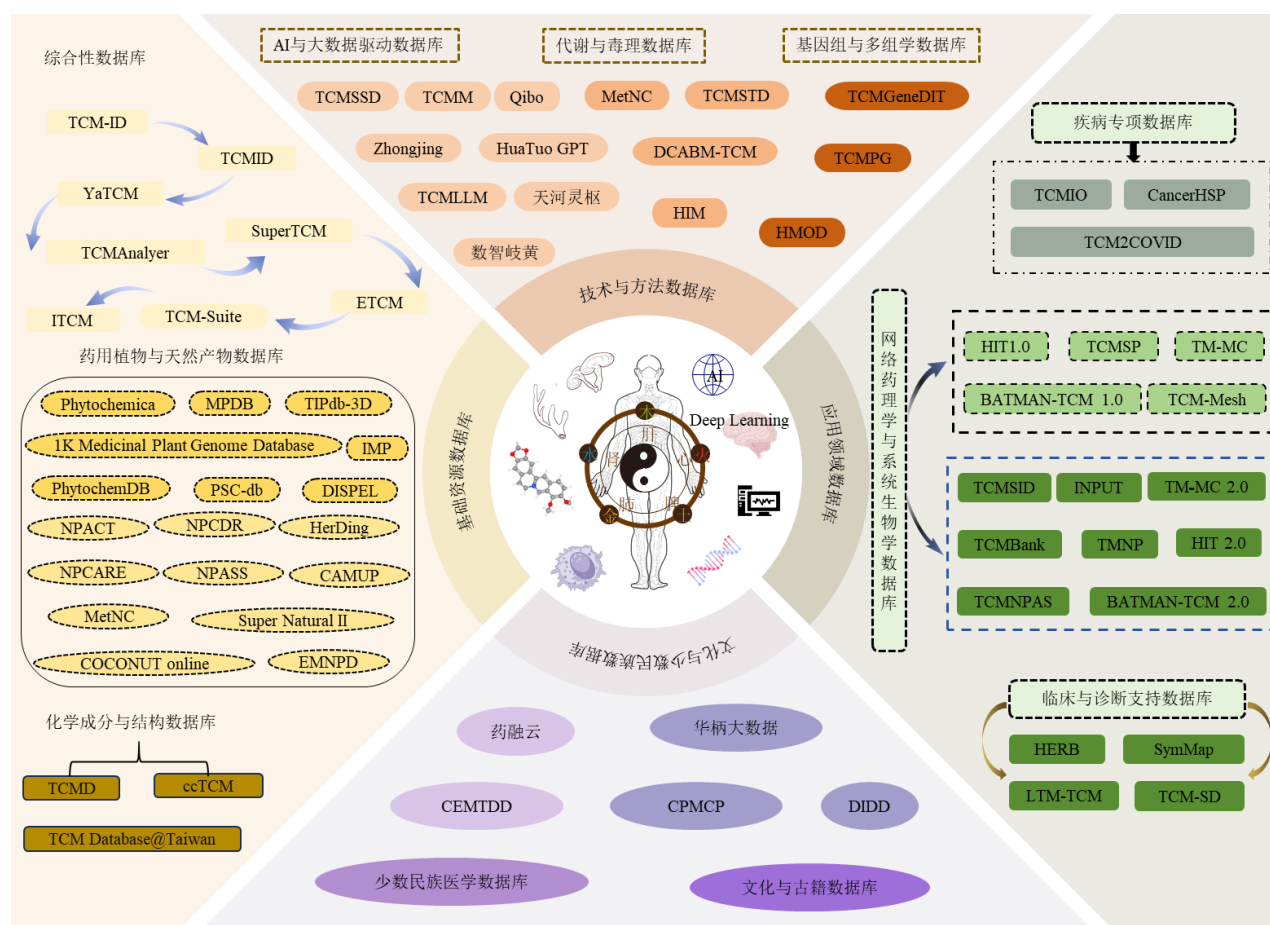


图 1 中医药生物信息学数据库分类

Fig. 1 Classification of traditional Chinese medicine bioinformatics databases

1 基础资源数据库

基础资源数据库是中医药数据体系的基石，主要包含综合性数据库、药用植物与天然产物数据库、化学成分与结构数据库，为药物发现、机制解

析和临床应用提供原始数据支撑。

1.1 综合性数据库

综合性数据库包含药材、方剂、成分、靶点、疾病等多维度信息，为中医药研究提供基础数据支

撑（表 1），在整个发展阶段由静态知识库向动态知识库演进，融合了地理分布、炮制工艺等信息，通过多维数据整合及算法升级等实现从基础研究到精准医疗的跨越（图 2）。

表 1 综合性数据库数据量
Table 1 Data volume of comprehensive databases

数据库	网址	中药数量	处方数量	成分数量	靶点数量	疾病数量	文献
TCM-ID	https://www.bidd.group/TCMID/	1 313	1 588	5 669	—	4 111	9
TCMID	http://www.megabionet.org/tcmid/	8 159	47 000	25 210	17 521	3 791	10
TCMID 2.0	http://www.megabionet.org/tcmid/	8 159	49 000	25 210	17 521	4 633	11
TCMAnalyzer	http://www.rcdd.org.cn/tcmanalyzer	618	1 493	16 437	2 298	1 529	12
YaTCM	http://cadd.pharmacy.nankai.edu.cn/yatcm/home	6 220	1 813	47 696	18 697	—	13
ETCM	http://www.tcmip.cn/ETCM/	3 962	478	7 274	2 266	3 027	14
SuperTCM	https://tcm.charite.de/supertcm/	6 516	—	55 772	543	8 634	15
TCM-Suite	http://TCM-Suite.AImicrobiome.cn	7 322	6 692	704 321	19 319	15 437	16
ITCM	http://itcm.biotcm.net	8 454	25 857	43 430	18 851	11 180	17
ETCM 2.0	http://www.tcmip.cn/ETCM2/front/#/	2 079	58 314	38 296	1 040	8 045	18

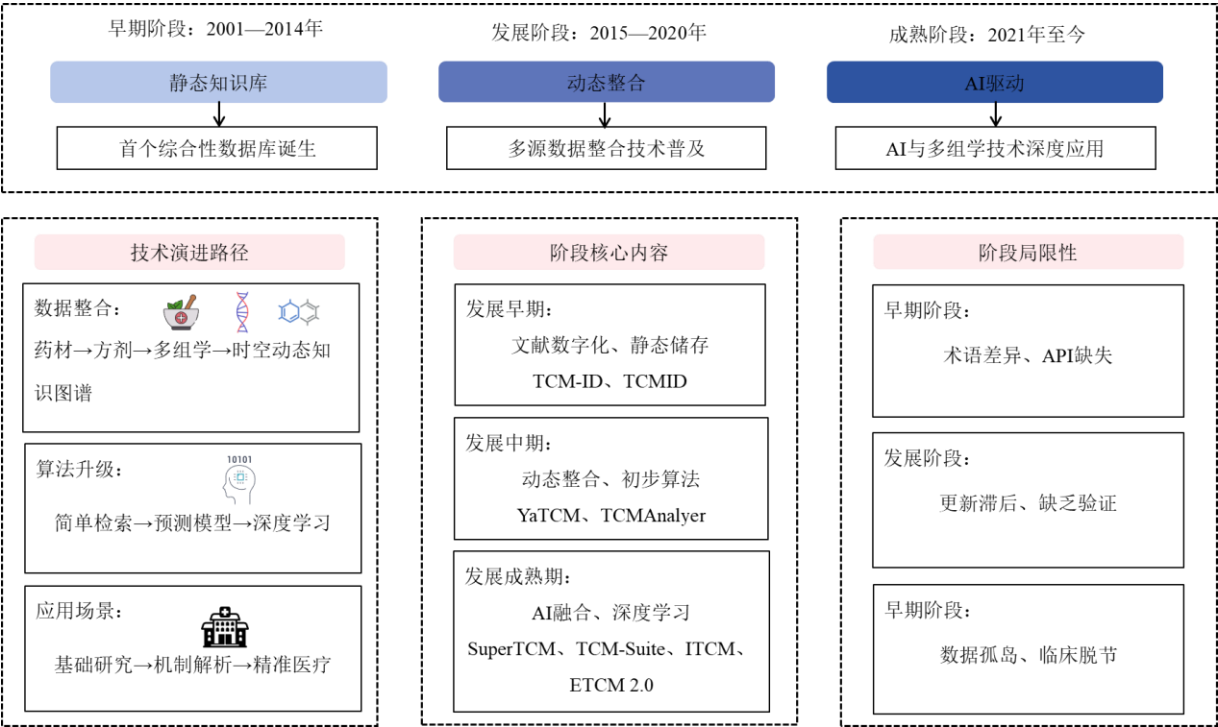


图 2 综合性数据库发展历程

Fig. 2 Development timeline of comprehensive databases

在早期阶段（2001—2014 年），中医药生物信息学数据库多为简单的知识库或文献检索工具，主要基于中医药经典著作和文献。首个综合数据库——中医药信息数据库（traditional Chinese medicine information database, TCM-ID）首创“方剂-药材-成分-靶点”网络模型，其数据主要来源于《中国药典》，具有权威性，且含 3D 化学成分结构，但网络关联简单^[9]。Xue 等^[10]构建的 TCMID1.0 应用虚拟网络，建立“草药-成分-靶点-疾病”之间相互关系网络并进行可视化。

在发展阶段（2015—2020 年），数据维度进一步扩展。TCMID2.0 数据库增加处方成分与质谱信息，以便于合理解释炮制对中药化学成分的影响^[11]。TCMAnalyzer 数据库侧重于分析中药的化学和生物信息学网络，可以通过化合物子结构、相似性等多种方法进行搜索，但其包含的量级相对较少^[12]。YaTCM 数据库整合了古籍、书籍及文本挖掘等多源数据，含独立预测算法，但未对外开放^[13]。ETCM 数据库集成古今方剂术语并进行了标准化，是较为全面的中药综合信息数据库^[14]。

在成熟阶段（2021 年至今），中医药生物信息学数据库融合 AI 技术及深度学习等算法，加速中医药现代化与智能化发展。如 SuperTCM 数据库将生化遗传途径相结合，解决药材异名问题，且包含舌象和脉象等中医药诊断知识^[15]。TCM-Suite 数据库采用多模块设计，支持网络药理学与系统生物学研究可视化，且含有生物成分标记基因序列^[16]。ITCM 数据库支持活性成分转录组分析，支持大规模筛选，预测

ADMET 算法先进，但仅单细胞系数据，表达谱覆盖不足^[17]。ETCM 2.0 数据库较 ETCM 支持标准化术语与定量成分统计，包含证候预测模型、关联性算法等创新性算法，中药网络可视化功能强，但部分网址访问不稳定^[18]。

1.2 药用植物与天然产物数据库

药用植物与天然产物数据库聚焦植物来源的天然产物及其活性成分，支持药物发现与开发（表 2）。

表 2 药用植物与天然产物数据库数据量

Table 2 Data volume of medicinal plants and natural products databases

数据库	网址	化合物数量	天然产物数量	药物数量	靶点数量	疾病数量	基因数量	药用植物数量	文献
TIPdb-3D	http://cwtung.kmu.edu.tw/tipdb	8 853	—	—	—	—	—	—	19
MPDB	https://www.medicinalplantbd.com/	—	—	—	—	—	—	500	20-21
Phytochemica	home.iitj.ac.in/~bagler/webserver/Phytochemica	8 093	—	—	—	—	—	525	22
PSC-db	http://pscdb.appsbio.utalca.cl	2 853	—	—	—	—	—	—	23
PhytochemDB	https://phytochemdb.com/	—	—	—	—	—	—	963	24
1K Medicinal Plant Genome Database	http://www.herbgenome.com/	—	—	—	—	—	—	—	25
DISPEL	https://compbio.iitr.ac.in/dispele	—	—	—	—	—	100	—	26
IMP	https://www.bic.ac.cn/IMP	—	—	—	—	—	8565 672	—	27
NPACT	http://crdd.osdd.net/raghava/npact/	1 574	—	—	—	—	—	—	28
Super Natural II	http://bioinformatics.charite.de/supernatural	32 6000	—	—	—	—	—	—	29
HerDing	http://combio.gist.ac.kr/herding	25 210	8 159	6 828	17 521	3 721	—	—	30
NPCARE	http://silver.sejong.ac.kr/npcare	6 578	1 952	—	—	—	—	—	31
NPASS	http://bidd.group/NPASS	—	94 413	—	5 863	—	—	—	32-33
CAMUP	http://bidd2.nus.edu.sg/CMAUP/	47 645	5 645	—	—	656	—	—	34
COCONUT online	https://coconut.naturalproducts.net	—	730 441	—	—	—	—	—	35
NPCDR	https://idrlab.org/npcdr/	—	50 000	12 000	—	218	—	—	36
EMNPD	http://emnnpd.idrlab.cn/	—	6 632	—	—	—	—	—	37

药用植物数据库是专门用于收集和整合药用植物信息的数据库，通常包括植物分类、成分、药理作用、临床应用和安全性数据等内容。TIPdb-3D 是中国台湾省开发的药用植物数据库，与 TCM database@Taiwan 相比更侧重于中药天然化合物^[19]。MPDB 是一个关于孟加拉国药用植物的数据库^[20-21]。Phytochemica 数据库涵盖了药用植物和化学物质，尤其涉及喜马拉雅地区的生物资源^[22]。PSC-db 是一个收集药用植物 2D 和 3D 结构信息的数据库^[23]。PhytochemDB 是一个全面的药用植物数据库，其中详细记录了药用植物的化学成分信息^[24]。1K Medicinal Plant Genome Database 是一个药用植物综合性的数据库，将药用植物的基因组和代谢物结合在一起^[25]。DISPEL 是用于治疗人类疾病的药用植

物数据库，包含有关药用植物治疗相关疾病治愈率等相关信息^[26]。IMP 数据库整合了药用植物的基因组和转录组相关信息^[27]。

天然产物数据库整合了有关天然药物（如植物、动物和矿物来源的药物）信息的综合数据库，旨在为研究人员、药剂师和临床医生提供有关天然药物的详细信息，包括其成分、作用机制、临床应用和安全性。NPACT 是一个关于“天然植物抗癌成分-活性-靶点”的数据库，但已不再更新和维护^[28]。Super Natural II 数据库能够预测天然产物的毒性并支持多种方式搜索（更新于 2022 年）^[29]。HerDing 数据库能通过疾病搜索获取治疗草药并关联基因化学物质^[30]。NPCARE 数据库收录与癌症相关的天然产物及其对癌细胞活性^[31]。NPASS

数据库详实记录真菌、细菌及动物源天然产物活性数据并支持跨物种比较^[32-33]。CAMUP 是一个植物活性分子数据库,植物信息丰富,成为植物活性分子研究和药物开发的重要工具^[34]。COCONUT online 数据库中的异构体生物活性精准编码、数据准确且更新及时^[35]。NPCDR 是首个基于天然产物的药物组合信息数据库,药物组合都可得到收集的文献和临床试验的支持,因此比预测数据更可靠^[36]。MetNC 数据库通过预测推荐最有可能在体内发生生物转化的候选药物^[38]。此外,EMNPD 是首个微生物天然产物数据库,包含天然产物的生物活性效力评估,其数据准确性较

高^[37]。

1.3 化学成分与结构数据库

化学成分与结构数据库存储化合物结构、定量信息及代谢路径数据(表3)。2000 年,发布首个中药数据库 TCMD,包含中药基原植物及其化学成分,以及化学成分的二维和三维结构^[39]。TCM Database@Taiwan 是首个虚拟筛选数据库,含化合物 2D、3D 结构,支持中英双语,但相较于 ETCM 等数据库,其在成分-靶点-疾病网络分析等多级数据分析方面的能力有限^[40]。ccTCM 数据库首创定量成分分析,支持 3 级相似度算法,但数据规模有限,未覆盖少数民族药材^[41]。

表 3 化学成分与结构数据库数据量

Table 3 Data volume of chemical composition and structure databases

数据库	网址	中药数量	成分数量	靶点数量	文献
TCMD	—	1 500	—	—	39
TCM Database@Taiwan	http://tcm.cmu.edu.tw/	443	24 033	—	40
ccTCM	http://www.cctcm.org.cn.	273	2 757	40 767	41

2 应用领域数据库

应用领域数据库针对特定应用场景设计的专题数据库,覆盖网络药理学与系统生物学、疾病治疗、临床决策、公共卫生等领域,强调数据与需求的精准对接。

2.1 网络药理学与系统生物学数据库

Li 等^[42]提出以网络的形式治疗中医证候,开创了网络药理学的概念^[43]。作为 AI 与大数据时代催生

的交叉学科,网络药理学因与中医整体观的高度契合,广泛应用于阐释中药复方疗效的综合机制、药效物质基础及组分协同作用,为解析疾病与中药的复杂性提供了全新范式^[44]。网络药理学在发展阶段从网络模型初步关联分析到 AI 驱动多组学整合实现网络药理学与中医药的深层次结合(图 3 和表 4)。

发展早期(2001—2020 年),网络药理学数据库数据来源相对有限,且功能相对简单。HIT 数据

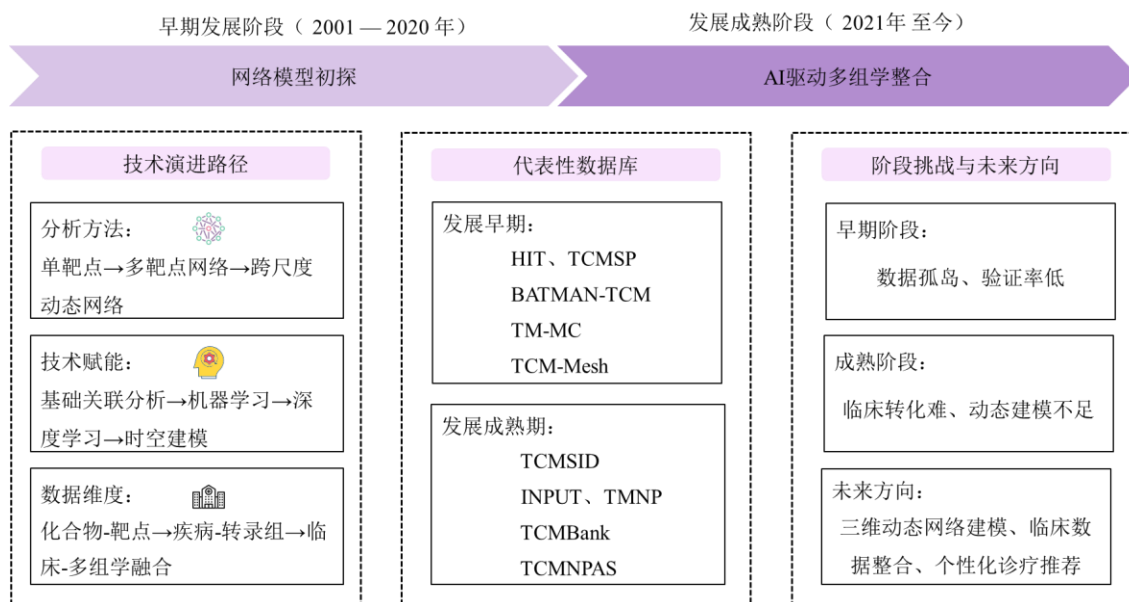


图 3 网络药理学与系统生物学数据库发展历程

Fig. 3 Development timeline of network pharmacology and systems biology databases

表 4 网络药理学与系统生物学数据库数据量

Table 4 Data volume of network pharmacology and systems biology databases

数据库	网址	中药数量	成分数量	靶点数量	疾病数量	处方数量	文献
HIT	http://lifecenter.sgst.cn/hit/	1 300	586	1 301	—	—	45
TCMSP	http://sm.nwsuaf.edu.cn/lsp/tcmsp.php	499	29 384	3 311	837	—	46
TM-MC1.0	https://tm-mc.kr/	536	—	—	—	14 000	47
BATMAN-TCM	http://bionet.ncpsb.org/batman-tcm/	8 159	25 210	—	—	46 914	48
TCM-Mesh	http://mesh.tcm.microbioinformatics.org/	6 235	38 340	14 298	6 204	—	49
TCMSID	https://tcm.scbdd.com	499	20 015	3 270	—	—	50
INPUT	https://cbcb.cdutcm.edu.cn/INPUT/	490	2 158	10 424	130	—	51
HIT2.0	http://www.badd-cao.net:2345/	1 250	1 237	2 208	—	—	52
TCMBank	https://TCMBank.CN/	9 192	61 966	15 179	32 529	—	53
TM-MC2.0	https://tm-mc.kr	635	34 107	13 992	27 997	5 075	54
TCMNPAS	http://54.223.75.62:3838/	1 630	18 090	—	—	—	55
BATMAN-TCM 2.0	http://bionet.ncpsb.org.cn/batman-tcm/	8 404	39 171	9 927	6 632	54 832	56

库与 Uniprot、PDB 等数据库建立交叉链接, 专注中药活性成分与其靶点相关数据^[45]。TCMSP 数据库含中药化合物及靶点等信息, 其中药动力学参数广泛用于成分筛选, 但其数据最近更新于 2014 年, 网站稳定性差^[46]。TM-MC 1.0 数据库仅包含了中国、日本和韩国的药典信息以及药材的化学成分信息^[47]。BATMAN-TCM 是首个在线的中药分子机制生物信息分析数据库, 广泛用于已知和新发现化合物靶点预测^[48]。TCM-Mesh 数据库的药材信息数据量有所提升, 新增高通量实验数据, 构建“中药-化合物-基因-疾病”网络, 并进行可视化, 以揭示小分子化合物的作用机制^[49]。

发展成熟阶段(2021 年至今)的网络药理学数据库深度整合多维度数据, 功能显著增强。TCMSID 数据库整合中药关键成分、指标成分、图片并进行分类与靶点预测^[50]。INPUT 集成蛋白质相互作用(protein-protein interaction, PPI)网络分析, 支持在线机制解析。例如, 通过整合代谢实验和网络药理学破解逍遥散的抗抑郁机制^[51]。HIT 2.0 数据库的数据量、更新速度、药理学整合及网络分析支持方面大幅提升, 克服了 HIT 1.0 的许多局限性^[52]。TMNP 数据库使用中药的药物转录组数据作为生物活性的标志, 直接评估中药组合的效应, 避免了中

药多成分作用环节。此外, TMNP 能够在分子、细胞、组织、病理过程等不同生物尺度上探索中药组合的效应, 符合中医整体治疗的理念。例如, 在黄芪和血塞通注射液应用中应用 TMNP 验证了其在揭示中药多尺度效应的能力^[57]。TCMBank 是基于自然语言处理的智能挖掘平台, 自动更新 PubChem 最新数据, 是目前中药、疾病等数据量较大的数据库^[53]。TM-MC 2.0 数据库较 TM-MC 扩展了处方、靶点及疾病信息^[54]。TCMNPAS 数据库结合多种算法解析方剂核心化合物与靶点关系并支持在线分子对接^[55]。BATMAN-TCM 2.0 数据库支持分子对接与多层级网络分析, 靶点预测准确率提升^[56]。

2.2 疾病专项数据库

疾病专项数据库是针对特定疾病(如癌症、新型冠状病毒肺炎)的中医药治疗方案数据库(表 5)。TCMIO 是一个免疫肿瘤学数据库, 含抗肿瘤免疫靶点, 采用配体-网络药理解析模块化设计, 在抗肿瘤免疫靶点信息方面较为全面, 但是 TCMIO 中中药仅来源于《中国药典》, 数据较少^[58]。CancerHSP 是一个网络药理学抗癌草药数据库, 可通过算法预测成分靶蛋白^[59]。TCM2COVID 是抗新型冠状病毒肺炎方剂数据库, 支持机制解析, 收录“三药三方”的分子机制与临床证据, 为突发公共卫生事件中的

表 5 疾病专项数据库数据量

Table 5 Data volume of disease-specific databases

数据库	网址	处方数量	中药数量	成分数量	靶点数量	文献
TCMIO	http://tcmio.xielab.net.	1 493	618	16 437	400	58
CancerHSP	http://lsp.nwsuaf.edu.cn/CancerHSP.php	—	2 439	3 575	—	59
TCM2COVID	http://zhangy-lab.cn/tcm2covid/	280	300	—	—	60

中医药方案选择提供快速响应支持^[60]。

2.3 临床与诊断支持数据库

临床与诊断支持数据库通过整合临床数据与诊断标准，助力中医药智能化决策（表 6）。HERB 数据库首含临床试验与 Meta 分析，支持循证医学研究，数据量庞大^[61]。然而，HERB 也存在一些问题，比如中药转录组数据集不足，因此其利用也受

到限制。SymMap 构建“证候-现代症状-靶点”映射网络，是第 1 个中医证候与现代医学症状对应的标准化数据库，但其数据集较小^[62]。LTM-TCM 数据库为中英双语界面，附加人工收集的 41 025 个临床治疗记录以及古籍治疗心血管疾病数据^[63]。TCM-SD 是首个开放且收集临床数据的中医辨证证数据库，但该数据库的处方信息集较小^[64]。

表 6 临床与诊断支持数据库数据量
Table 6 Data volume of clinical and diagnostic support databases

数据库	网址	处方数量	中药数量	成分数量	靶点数量	疾病数量	中医症状数量	现代医学症状数量	证型数量	文献
HERB	http://herb.ac.cn/	—	7 236	49 258	12 933	28 212	—	—	—	61
HERB 2.0	http://47.92.70.12/	6 743	6 892	44 601	15 515	30 170	—	—	—	—
SymMap	http://www.symmap.org/	—	499	19 595	4 302	5 235	1 717	961	—	62
LTM-TCM	http://cloud.tasly.com/#/tcm/home	48 126	9 122	34 967	13 109	—	—	—	—	63
TCM-SD	https://github.com/BororBor/ZY-BERT	—	—	—	—	—	—	—	148	64

3 技术与方法数据库

技术与方法聚焦数据分析技术与方法涵盖 AI、代谢组学等工具与算法，旨在解析中药多成分协同作用的复杂机制。

3.1 AI 与大数据驱动数据库

深度融合为解决人类健康问题、创新药物开发和阐明疾病机制提供了创新思路，从而进一步推进医学研究和治疗创新^[65-66]。AI 算法尤其是机器学习和深度学习技术，擅长识别潜在的模式和趋势，从而促进发现新的中药成分和阐明其作用机制，通过对化学结构、生物活性和靶蛋白之间的关系建模，发现构效关系来解释中医理论^[67]。例如，深度神经网络推

断化合物的生物活性和预测分子，构建算法定期预测蛋白质结构的原子精度^[68-70]。Chen 等^[71]首次将 AI 方法引入生物和药物研究，开创了药物靶点发现的反向对接方法，并应用深度学习方法预测药物性质，并依此创建了知名药物领域的 TTD 药物靶点数据库^[72-73]。与此同时，AI 可以预测中药的疗效和安全性，并优化药物组合以提高治疗效果；建立中医推荐模型，通过 AI 算法实现中医智能化决策。此外，AI 深入分析了中医方剂的配伍规律，通过证素构建了中医诊疗系统，这些深度学习和建模方法使中医数据库不断得到创新^[74-76]，通过利用深度学习、知识图谱等技术优化数据挖掘与预测（表 7）。

表 7 AI 与大数据驱动数据库数据量
Table 7 Data volume of AI and big data-driven databases

数据库	网址	处方数量	中药数量	成分数量	靶点数量	疾病数量	中医症状数量	现代医学症状数量	证型数量	文献
TCMSSD	http://tcmssd.ratcm.cn/	133 518	8 259	43 413	17 602	8 073	1 843	6 230	624	77
TCMM	https://www.tcmm.net.cn/	48 043	8 932	69 816	76 449	22 365	1 900	17 079	146	78

目前 AI 与大数据驱动数据库有 TCMSSD 数据库，结合 AI 辨证模型与知识图谱预测证候，准确率较高^[77]。TCMM 数据库是基于知识图谱和图神经网络的处方生成系统，训练集为古籍方剂，同时对在线并发症关联分析，支持动态复方优化^[78]。张仲景大模型采用连续预训练、监督式微调和强化学习训练，用于辅助诊断与辨证论治^[79]。HuaTuo GPT 数据库专注于中医问诊、辨证论治和处方推荐，支持

多轮对话，能结合患者症状进行个性化诊疗建议。“数智岐黄”大模型涵盖中医典籍、方剂、证候等知识，实现中医药知识智能问答和临床辅助诊疗^[80]。Qibo 大模型收集大量真实的医学语料结合知识图谱进行训练，能够实现智能医疗问答及辅助诊断作用^[81]。天河灵枢大模型以“天河天元大模型”作为基础大模型训练进一步融合中医经典名著、针灸临床循证证据库以及中医循证知识图谱等专业数据，

创新性地对人体全部穴位进行了专业的三维建模，成功构建出三维针灸数字人，能够动态展示相关穴位信息和针灸治疗方案。TCMLLM 数据库可以用于中医处方推荐、临床辅助诊疗、中医药知识问答等多个场景^[82]。

3.2 代谢与毒理学数据库

代谢与毒理学数据库主要包含成分代谢途径分析及毒性效应 2 部分（表 8）。系统毒理学数据库 TCMSTD 含 5 类毒性靶点预测，即心脏毒性、肝毒性、肾毒性、神经毒性和生殖毒性，但仅基于结构相似性，未考虑代谢产物毒性^[83]。HIM 是体内代谢路径数据库，交叉链接 PubChem、TCM-ID 等数据库，但动态代谢模拟能力有限^[84]。MetNC 数据库可以进行代谢产物预测，推荐体内生物转化候选分子。DCABM-TCM 是首个系统收集中药及其方剂血液成分的数据库，在整合中药化学成分和生物活性数据方面具有明显优势，并支持网络药理学研究^[85]。此外，DCABM-TCM 数

据库还提供了相应的详细检测条件。

3.3 基因组与多组学数据库

基因组与多组学数据库整合基因组、转录组与代谢组数据，解析中医药分子机制（表 9）。TCMGeneDIT 数据库从大量生物医学文献中挖掘出基因、疾病和中药及其成分之间的作用关系^[86]；将中医药与基因组学和蛋白质组学相结合有助于揭示中药基因调控，但该数据库中的中医药信息组织得不够完善。HMOD 数据库将多组学整合，含代谢通路和转录组关联分析，并且具有很强的时效性^[87]。IGTCM 数据库整理了中药的基因组序列，但总体中药数量仍然较少^[88]。TCMPG 为药用植物基因组数据库，涵盖 255 物种^[89]。TCMPG 2.0 进一步补充和完善了中药植物的信息^[90]。

4 文化与少数民族医学数据库

文化与少数民族医学数据库致力于保存中医药文化遗产与少数民族传统医学知识（表 10），促进多民族医学融合与现代化。

表 8 代谢与毒理学数据库数据量

Table 8 Data volume of metabolism and toxicology databases

数据库	网址	处方数量	中药数量	成分数量	靶点数量	体内代谢产物数量	文献
TCMSTD	https://www.bic.ac.cn/TCMSTD/	22	252	4 361	2 425	—	83
HIM	http://www.bioinformatics.org.cn/	—	673	361	—	1 104	84
DCABM-TCM	http://bionet.ncpsb.org.cn/dcabm-tcm/	192	194	—	—	—	85

表 9 基因组与多组学数据库数据量

Table 9 Data volume of genomics and multi-omics databases

数据库	网址	中药数量	基因数量	疾病数量	转录组数量	代谢物数量	文献
TCMGeneDIT	http://tcm.lifescience.ntu.edu.tw/	2 000	13 167	3 360	—	—	86
HMOD	http://herbalplant.ynau.edu.cn/	23	—	—	172	55	87
IGTCM	http://yeyn.group.96/	83	3610 350	—	—	—	88
TCMPG	http://cbcb.cdutcm.edu.cn/TCMPG/	255	195	—	—	—	89
TCMPG 2.0	http://cbcb.cdutcm.edu.cn/TCMPG2/	414	324	—	—	—	90

表 10 文化与少数民族医学数据库数据量

Table 10 Data volume of cultural and ethnic minority medicine databases

数据库	网址	处方数量	中药数量	成分数量	靶点数量	疾病数量	文献
CEMTDD	http://www.cemtdd.com/index.html	—	621	4 060	2 163	210	91
CPMCP	http://cpmcp.top	2 125	1 557	26 341	20 965	14 086	92
DIDD	—	—	1 338	—	—	—	93

4.1 少数民族医学数据库

少数民族医学数据库通过记录少数民族传统药物与疗法，促进多民族医学融合。CEMTDD 为中国少数民族药物数据库，聚焦新疆维吾尔医药，为“沙疗”等民族疗法的现代化研究提供数据支撑^[91]。

药融云数据资源丰富，不仅包含传统中医药的数据，还对少数民族医药进行了整合。其中，药模块全面，支持多种应用场景。

4.2 文化与古籍数据库

传统方剂与古籍数据库通过整理古籍方剂与

经典文献支持中医药传承研究。CPMCP 是一个中成药与古方数据库,通过中医古籍语义解析映射古症状至现代标准,加速中医药现代化^[92]。DIDD 药食同源数据库主要围绕药食同源物质的成分、功效和安全性,旨在为公众和企业提供实用信息^[93]。华柄数据库数据资源丰富,收集超 20 万首方剂信息,其预测模型基于中医药大数据、方剂用药规律、图计算开发,预测结果包括疾病、症状、功能,可用于自拟方剂的优化、调整、完善等研究。

5 结语与展望

5.1 中医药生物信息学数据库发展的核心成就

中医药生物信息学数据库的发展体现了从“经验驱动”到“数据驱动”的范式转变,其核心成就可概括为以下 3 点。(1) 多维数据整合能力提升:早期数据库如 TCMD、TCM-ID 等仅聚焦药材与成分的静态存储,而现代数据库已实现多组学数据(基因组、代谢组、临床疗效)的深度融合,为中西医结合研究提供了分子层面的证据链。(2) AI 技术的深度赋能:生成式 AI 与自然语言处理大幅提升了数据挖掘效率。TCMBank 通过智能文本挖掘,较传统人工录入效率大幅提高。(3) 应用场景的多元化拓展:数据库功能从基础检索延伸至精准医疗支持。如 TCM2COVID 在新型冠状病毒肺炎疫情期间,为“三药三方”的全球推广提供了实时机制解析与临床证据整合,凸显了中医药生物信息学数据库在公共卫生事件中的应急价值。

5.2 现存挑战的深层次分析

尽管中医药生物信息学数据库发展具有显著的成果,但仍存在一些局限。(1) 数据标准化与互通性问题:同一药材在不同数据库中存在术语差异,导致跨库检索效率降低;数据库之间存在格式壁垒,未提供标准化(application programming interface, API)接口,迫使研究者依赖手工数据清洗,耗费人力物力资源。(2) 动态更新与质量控制机制缺失:部分数据库未及时更新,数据时效性严重滞后。以 TCMSP 数据库为例,其药动学参数有限,未纳入 2014 年后新发现的活性分子。数据库之间的数据质量参差不齐,部分数据库依赖非结构化文本挖掘,错误率高,如化合物结构式错误、计量单位混淆。(3) 临床验证与应用脱节:网络药理学预测结果缺乏实验验证,如 BATMAN-TCM 2.0 的靶点验证率低,且多数验证实验局限于细胞模型,缺乏动物或临床试验支持。临床数据库如 HERB 虽

整合了 Meta 分析数据,但未与电子病历系统(electronic medical record, EMR)实时对接,难以支持个性化诊疗决策。此外,目前有关中药处方个性化推荐的模型层出不穷,但多依赖症状关键词匹配,未内化“八纲辨证”“六经辨证”等核心中医推理机制,导致推荐处方脱离病机本质,以及忽视体质、地域、节气等个性化因素,如“同病异治”原则,无法生成因人制宜的诊疗方案^[94]。

5.3 未来发展的创新路径

在未来的发展中需突破瓶颈问题,从技术、机制与协作 3 方面协同创新:(1) 在技术层面,构建动态知识图谱与联邦学习架构。联邦学习是一种分布式机器学习框架,其核心思想是在多个数据源共同参与模型训练时,不需要进行原始数据流转,实现“数据可用不可见”的数据应用模式。联邦学习架构作为 AI 技术在传统数据库中应用的重要技术之一,例如 Intel & Penn Medicine 合作使用联邦学习在保护隐私的前提下,训练来自全球多个机构的脑瘤核磁共振成像(magnetic resonance imaging, MRI)分割模型^[95]。未来可以通过动态知识图谱融合地理信息系统(geographic information system, GIS),明确道地药材分布,如可基于 SuperTCM 的生化遗传路径数据,结合气象卫星信息,预测药材有效成分含量的波动。开发中医证候-生理指标动态映射算法,利用可穿戴设备监测脉象、舌象数据。例如,华为光电容积脉搏波(photoplethysmography, PPG)传感器结合 AI 舌诊图像识别,以及检测患者实时生理指标与多组学数据,构建时空动态知识网络^[96]。此外,在保护隐私前提下,通过联邦学习,分布式机器学习可实现跨机构数据共享,如联合临床试验数据与医院 EMR 系统,训练个性化处方推荐模型,避免敏感数据集中存储风险。(2) 在机制层面,建立全生命周期数据管理标准。采用数据标准化协议,可参考 FAIR 原则(可发现、可访问、可互操作、可重用),制定中医药专属数据标准。例如,定义“证候”的临床医学术语标准——SNOMED-CT 编码扩展集,实现 SymMap 数据库与 ICD-11 诊断系统的互操作。此外,在区块链存证,利用智能合约技术记录古籍方剂与民族医药知识的贡献链,确保知识溯源不可篡改。如将《伤寒论》方剂的数字化版本上链,为知识产权保护提供技术支撑。此外,为突破当前中医药数据库与大模型在辨证论治中的局限,亟需构建“动态辨证数理模型-个性诊

疗联邦系统”，基于辨证理论将证候拆解为量化向量，开发中医规则引擎实现“症状→辨证方法→证候→治法”的自动推演，破解逻辑缺失困境；构建患者数字孪生体整合多组学数据，依托联邦学习生成“因人-因时-因地”三因制宜的理法方药链路，如湿毒郁肺证动态优化三仁汤方剂权重，确保辨证与施治一致性^[97]。（3）在协作层面，推动中医药全球化数据建设。成立全球中医药生物信息学数据库，统一数据标准。参考欧盟 Horizon 2020 设计联邦多中心疗效验证框架；并开发跨医学知识图谱以及多语言 Transformer 模型，如将《伤寒论》与印度 *Charaka Samhita* 方剂语义匹配，深度释放传统医学协同潜能。

中医药生物信息学数据库已实现从基础数字化存储向多维度智能分析系统的范式演进，逐步构建了“数据整合-知识挖掘-临床决策”联动的技术架构。AI 等技术的深度应用将有效驱动中医药研究模式革新，突破数据共享壁垒并弥合临床脱节问题，加速从经验医学向精准医学的转型。未来研究应从技术创新、机制完善和协作拓展 3 方面入手，协同推进技术标准化、理论深层次剖析应用及文化互鉴融合，以全面提升中医药在全球健康领域的应用价值。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] Li W F, Jiang J G, Chen J. Chinese medicine and its modernization demands [J]. *Arch Med Res*, 2008, 39(2): 246-251.
- [2] Sorokina M, Steinbeck C. NaPLoS: A natural products likeness scorer-web application and database [J]. *J Cheminform*, 2019, 11(1): 55.
- [3] Tang J L, Liu B Y, Ma K W. Traditional Chinese medicine [J]. *Lancet*, 2008, 372(9654): 1938-1940.
- [4] White N J. Qinghaosu (artemisinin): The price of success [J]. *Science*, 2008, 320(5874): 330-334.
- [5] Wang J G, Wong Y K, Liao F L. What has traditional Chinese medicine delivered for modern medicine? [J]. *Expert Rev Mol Med*, 2018, 20: e4.
- [6] Huang K, Zhang P, Zhang Z H, et al. Traditional Chinese Medicine (TCM) in the treatment of COVID-19 and other viral infections: Efficacies and mechanisms [J]. *Pharmacol Ther*, 2021, 225: 107843.
- [7] Rigden D J, Fernández X M. The 2023 Nucleic Acids Research Database Issue and the online molecular biology database collection [J]. *Nucleic Acids Res*, 2023, 51(D1): D1-D8.
- [8] Wang X, Wang Z Y, Zheng J H, et al. TCM network pharmacology: A new trend towards combining computational, experimental and clinical approaches [J]. *Chin J Nat Med*, 2021, 19(1): 1-11.
- [9] Chen X, Zhou H, Liu Y B, et al. Database of traditional Chinese medicine and its application to studies of mechanism and to prescription validation [J]. *Br J Pharmacol*, 2006, 149(8): 1092-1103.
- [10] Xue R C, Fang Z, Zhang M X, et al. TCMID: Traditional Chinese Medicine integrative database for herb molecular mechanism analysis [J]. *Nucleic Acids Res*, 2013, 41(Database issue): D1089-D1095.
- [11] Huang L, Xie D L, Yu Y R, et al. TCMID 2.0: A comprehensive resource for TCM [J]. *Nucleic Acids Res*, 2018, 46(D1): D1117-D1120.
- [12] Liu Z H, Du J W, Yan X, et al. TCMAnalyzer: A chemo- and bioinformatics web service for analyzing traditional Chinese medicine [J]. *J Chem Inf Model*, 2018, 58(3): 550-555.
- [13] Li B Q, Ma C F, Zhao X Y, et al. YaTCM: Yet another traditional Chinese medicine database for drug discovery [J]. *Comput Struct Biotechnol J*, 2018, 16: 600-610.
- [14] Xu H Y, Zhang Y Q, Liu Z M, et al. ETCM: An encyclopaedia of traditional Chinese medicine [J]. *Nucleic Acids Res*, 2019, 47(D1): D976-D982.
- [15] Chen Q F, Springer L, Gohlke B O, et al. SuperTCM: A biocultural database combining biological pathways and historical linguistic data of Chinese *Materia Medica* for drug development [J]. *Biomed Pharmacother*, 2021, 144: 112315.
- [16] Yang P S, Lang J D, Li H J, et al. TCM-Suite: A comprehensive and holistic platform for Traditional Chinese Medicine component identification and network pharmacology analysis [J]. *Imeta*, 2022, 1(4): e47.
- [17] Tian S S, Zhang J B, Yuan S L, et al. Exploring pharmacological active ingredients of traditional Chinese medicine by pharmacotranscriptomic map in ITCM [J]. *Brief Bioinform*, 2023, 24(2): bbad027.
- [18] Zhang Y Q, Li X, Shi Y L, et al. ETCM v2.0: An update with comprehensive resource and rich annotations for traditional Chinese medicine [J]. *Acta Pharm Sin B*, 2023, 13(6): 2559-2571.
- [19] Tung C W, Lin Y C, Chang H S, et al. IPdb-3D: the three-dimensional structure database of phytochemicals from Taiwan indigenous plants [J]. *Database*, 2014, 2014: bau055.
- [20] Ashraf M A, Khatun A, Sharmin T, et al. MPDB 1.0: A

- medicinal plant database of Bangladesh [J]. *Bioinformatics*, 2014, 10(6): 384-386.
- [21] Hussain N, Chanda R, Abir R A, *et al.* MPDB 2.0: A large scale and integrated medicinal plant database of Bangladesh [J]. *BMC Res Notes*, 2021, 14(1): 301.
- [22] Pathania S, Ramakrishnan S M, Bagler G. Phytochemica: A platform to explore phytochemicals of medicinal plants [J]. *Database*, 2015, 2015: bav075.
- [23] Valdés-Jiménez A, Peña-Varas C, Borrego-Muñoz P, *et al.* PSC-db: A structured and searchable 3D-database for plant secondary compounds [J]. *Molecules*, 2021, 26(4): 1124.
- [24] Mahmud S, Paul G K, Biswas S, *et al.* Phytochemdb: A platform for virtual screening and computer-aided drug designing [J]. *Database*, 2022, 2022: baac002.
- [25] Su X J, Yang L L, Wang D L, *et al.* 1 K medicinal plant genome database: An integrated database combining genomes and metabolites of medicinal plants [J]. *Hortic Res*, 2022, 9: uhac075.
- [26] Singh K, Maurya H, Singh P, *et al.* DISPEL: Database for ascertaining the best medicinal plants to cure human diseases [J]. *Database*, 2023, 2023: baad073.
- [27] Chen T, Yang M, Cui G H, *et al.* IMP: Bridging the gap for medicinal plant genomics [J]. *Nucleic Acids Res*, 2024, 52(D1): D1347-D1354.
- [28] Mangal M, Sagar P, Singh H, *et al.* NPACT: Naturally occurring plant-based anti-cancer compound-activity-target database [J]. *Nucleic Acids Res*, 2013, 41(Database issue): D1124-D1129.
- [29] Banerjee P, Erehman J, Gohlke B O, *et al.* Super Natural II: A database of natural products [J]. *Nucleic Acids Res*, 2015, 43(Database issue): D935-D939.
- [30] Choi W, Choi C H, Kim Y R, *et al.* HerDing: Herb recommendation system to treat diseases using genes and chemicals [J]. *Database*, 2016, 2016: baw011.
- [31] Choi H, Cho S Y, Pak H J, *et al.* NPCARE: Database of natural products and fractional extracts for cancer regulation [J]. *J Cheminform*, 2017, 9: 2.
- [32] Zeng X, Zhang P, He W D, *et al.* NPASS: Natural product activity and species source database for natural product research, discovery and tool development [J]. *Nucleic Acids Res*, 2018, 46(D1): D1217-D1222.
- [33] Zhao H, Yang Y, Wang S Q, *et al.* NPASS database update 2023: Quantitative natural product activity and species source database for biomedical research [J]. *Nucleic Acids Res*, 2023, 51(D1): D621-D628.
- [34] Zeng X, Zhang P, Wang Y L, *et al.* CMAUP: A database of collective molecular activities of useful plants [J]. *Nucleic Acids Res*, 2019, 47(D1): D1118-D1127.
- [35] Sorokina M, Merseburger P, Rajan K, *et al.* COCONUT online: Collection of open natural products database [J]. *J Cheminform*, 2021, 13(1): 2.
- [36] Sun X N, Zhang Y T, Zhou Y, *et al.* NPCDR: Natural product-based drug combination and its disease-specific molecular regulation [J]. *Nucleic Acids Res*, 2022, 50(D1): D1324-D1333.
- [37] Xu H Q, Xiao H, Bu J H, *et al.* EMNPD: A comprehensive endophytic microorganism natural products database for prompt the discovery of new bioactive substances [J]. *J Cheminform*, 2023, 15(1): 115.
- [38] Chen Z K, Yan D Y, Zhang M, *et al.* MetNC: Predicting metabolites *in vivo* for natural compounds [J]. *Front Chem*, 2022, 10: 881975.
- [39] He M, Yan X, Zhou J, *et al.* Traditional Chinese medicine database and application on the Web [J]. *J Chem Inf Comput Sci*, 2001, 41(2): 273-277.
- [40] Chen C Y. TCM Database@Taiwan: The world's largest traditional Chinese medicine database for drug screening in silico [J]. *PLoS One*, 2011, 6(1): e15939.
- [41] Yang D Q, Zhu Z, Yao Q, *et al.* ccTCM: A quantitative component and compound platform for promoting the research of traditional Chinese medicine [J]. *Comput Struct Biotechnol J*, 2023, 21: 5807-5817.
- [42] Li S, Zhang Z Q, Wu L J, *et al.* Understanding ZHENG in traditional Chinese medicine in the context of neuro-endocrine-immune network [J]. *IET Syst Biol*, 2007, 1(1): 51-60.
- [43] Hopkins A L. Network pharmacology [J]. *Nat Biotechnol*, 2007, 25(10): 1110-1111.
- [44] Zhang P, Zhang D F, Zhou W A, *et al.* Network pharmacology: Towards the artificial intelligence-based precision traditional Chinese medicine [J]. *Brief Bioinform*, 2023, 25(1): bbad518.
- [45] Ye H, Ye L, Kang H, *et al.* HIT: Linking herbal active ingredients to targets [J]. *Nucleic Acids Res*, 2011, 39(Database issue): D1055-D1059.
- [46] Ru J L, Li P, Wang J N, *et al.* TCMSP: A database of systems pharmacology for drug discovery from herbal medicines [J]. *J Cheminform*, 2014, 6: 13.
- [47] Kim S K, Nam S, Jang H, *et al.* TM-MC: A database of medicinal materials and chemical compounds in Northeast Asian traditional medicine [J]. *BMC Complement Altern Med*, 2015, 15: 218.
- [48] Liu Z Y, Guo F F, Wang Y, *et al.* BATMAN-TCM: A bioinformatics analysis tool for molecular mechanism of traditional Chinese medicine [J]. *Sci Rep*, 2016, 6: 21146.
- [49] Zhang R Z, Yu S J, Bai H, *et al.* TCM-Mesh: The database

- and analytical system for network pharmacology analysis for TCM preparations [J]. *Sci Rep*, 2017, 7(1): 2821.
- [50] Zhang L X, Dong J, Wei H, *et al*. TCMSID: A simplified integrated database for drug discovery from traditional Chinese medicine [J]. *J Cheminform*, 2022, 14(1): 89.
- [51] Li X H, Tang Q, Meng F B, *et al*. INPUT: An intelligent network pharmacology platform unique for traditional Chinese medicine [J]. *Comput Struct Biotechnol J*, 2022, 20: 1345-1351.
- [52] Yan D Y, Zheng G H, Wang C C, *et al*. HIT 2.0: An enhanced platform for Herbal Ingredients' Targets [J]. *Nucleic Acids Res*, 2022, 50(D1): D1238-D1243.
- [53] Lv Q J, Chen G X, He H H, *et al*. TCMBank: Bridges between the largest herbal medicines, chemical ingredients, target proteins, and associated diseases with intelligence text mining [J]. *Chem Sci*, 2023, 14(39): 10684-10701.
- [54] Kim S K, Lee M K, Jang H, *et al*. TM-MC 2.0: An enhanced chemical database of medicinal materials in Northeast Asian traditional medicine [J]. *BMC Complement Med Ther*, 2024, 24(1): 40.
- [55] Liu Y S, Li X, Chen C, *et al*. TCMNPAS: A comprehensive analysis platform integrating network formulaology and network pharmacology for exploring traditional Chinese medicine [J]. *Chin Med*, 2024, 19(1): 50.
- [56] Kong X R, Liu C, Zhang Z Z, *et al*. BATMAN-TCM 2.0: An enhanced integrative database for known and predicted interactions between traditional Chinese medicine ingredients and target proteins [J]. *Nucleic Acids Res*, 2024, 52(D1): D1110-D1120.
- [57] Li P, Zhang H R, Zhang W X, *et al*. TMNP: A transcriptome-based multi-scale network pharmacology platform for herbal medicine [J]. *Brief Bioinform*, 2022, 23(1): bbab542.
- [58] Liu Z H, Cai C P, Du J W, *et al*. TCMIO: A comprehensive database of traditional Chinese medicine on immuno-oncology [J]. *Front Pharmacol*, 2020, 11: 439.
- [59] Tao W Y, Li B H, Gao S, *et al*. CancerHSP: Anticancer herbs database of systems pharmacology [J]. *Sci Rep*, 2015, 5: 11481.
- [60] Ren L P, Xu Y, Ning L, *et al*. TCM2COVID: A resource of anti-COVID-19 traditional Chinese medicine with effects and mechanisms [J]. *Imeta*, 2022, 1(4): e42.
- [61] Fang S S, Dong L, Liu L, *et al*. HERB: A high-throughput experiment- and reference-guided database of traditional Chinese medicine [J]. *Nucleic Acids Res*, 2021, 49(D1): D1197-D1206.
- [62] Wu Y, Zhang F L, Yang K, *et al*. SymMap: An integrative database of traditional Chinese medicine enhanced by symptom mapping [J]. *Nucleic Acids Res*, 2019, 47(D1): D1110-D1117.
- [63] Li X, Ren J, Zhang W, *et al*. LTM-TCM: A comprehensive database for the linking of Traditional Chinese Medicine with modern medicine at molecular and phenotypic levels [J]. *Pharmacol Res*, 2022, 178: 106185.
- [64] Ren M C, Huang H Y, Zhou Y X, *et al*. TCM-SD: A benchmark for probing syndrome differentiation via natural language processing [A] // *Chinese Computational Linguistics* [M]. Cham: Springer International Publishing, 2022: 247-263.
- [65] Ngiam K Y, Khor I W. Big data and machine learning algorithms for health-care delivery [J]. *Lancet Oncol*, 2019, 20(5): e262-e273.
- [66] Muldowney M W, Duncan K R, Elsayed S S, *et al*. Artificial intelligence for natural product drug discovery [J]. *Nat Rev Drug Discov*, 2023, 22(11): 895-916.
- [67] Xu L, Zhang Y X, Zhang P, *et al*. Integrated metabolomics and network pharmacology strategy-driven active traditional Chinese medicine ingredients discovery for the alleviation of cisplatin nephrotoxicity [J]. *Chem Res Toxicol*, 2019, 32(12): 2411-2421.
- [68] Berman H M, Battistuz T, Bhat T N, *et al*. The protein data bank [J]. *Acta Crystallogr D Biol Crystallogr*, 2002, 58(pt 6 no 1): 899-907.
- [69] Stokes J M, Yang K, Swanson K, *et al*. A deep learning approach to antibiotic discovery [J]. *Cell*, 2020, 180(4): 688-702.
- [70] Jumper J, Evans R, Pritzel A, *et al*. Highly accurate protein structure prediction with AlphaFold [J]. *Nature*, 2021, 596(7873): 583-589.
- [71] Chen X, Ung C Y, Chen Y Z. Can an in silico drug-target search method be used to probe potential mechanisms of medicinal plant ingredients? [J]. *Nat Prod Rep*, 2003, 20(4): 432-444.
- [72] Shen W X, Zeng X, Zhu F, *et al*. Out-of-the-box deep learning prediction of pharmaceutical properties by broadly learned knowledge-based molecular representations [J]. *Nat Mach Intell*, 2021, 3: 334-343.
- [73] Zhou Y, Zhang Y T, Zhao D H, *et al*. TTD: Therapeutic Target Database describing target druggability information [J]. *Nucleic Acids Res*, 2024, 52(D1): D1465-D1477.
- [74] Lam C S, Peng L W, Yang L S, *et al*. Examining patterns of traditional Chinese medicine use in pediatric oncology: A systematic review, meta-analysis and data-mining study [J]. *J Integr Med*, 2022, 20(5): 402-415.
- [75] Yang Y, Rao Y L, Yu M H, *et al*. Multi-layer information

- fusion based on graph convolutional network for knowledge-driven herb recommendation [J]. *Neural Netw*, 2022, 146: 1-10.
- [76] Wang J, Duan L, Li H Z, *et al*. Construction of an artificial intelligence traditional Chinese medicine diagnosis and treatment model based on syndrome elements and small-sample data [J]. *Engineering*, 2022, 8: 29-32.
- [77] Huang L, Wang Q, Duan Q C, *et al*. TCMSSD: A comprehensive database focused on syndrome standardization [J]. *Phytomedicine*, 2024, 128: 155486.
- [78] Ren Z X, Ren Y M, Li Z T, *et al*. TCMM: A unified database for traditional Chinese medicine modernization and therapeutic innovations [J]. *Comput Struct Biotechnol J*, 2024, 23: 1619-1630.
- [79] Yang S H, Zhao H J, Zhu S B, *et al*. Zhongjing: Enhancing the Chinese medical capabilities of large language model through expert feedback and real-world multi-turn dialogue [J]. *Proc AAAI Conf Artif Intell*, 2024, 38(17): 19368-19376.
- [80] WANG H, LIU C, XI N, *et al*. HuaTuo: Tuning LLaMA Model with Chinese Medical Knowledge [J/OL] 2023, arXiv:2304.06975[https://ui.adsabs.harvard.edu/abs/2023arXiv230406975W. 10.48550/arXiv.2304.06975]
- [81] Jia Y Z, Ji X Y, Wang X, *et al*. Qibo: A large language model for traditional Chinese medicine [J]. *Expert Syst Appl*, 2025, 284: 127672.
- [82] Tian H Y, Yang K, Dong X, *et al*. TCMLLM-PR: Evaluation of large language models for prescription recommendation in traditional Chinese medicine [J]. *Digit Chin Med*, 2024, 7(4): 343-355.
- [83] Song L L, Qian W X, Yin H Q, *et al*. TCMSTD 1.0: A systematic analysis of the traditional Chinese medicine system toxicology database [J]. *Sci China Life Sci*, 2023, 66(9): 2189-2192.
- [84] Kang H, Tang K L, Liu Q, *et al*. HIM-herbal ingredients *in-vivo* metabolism database [J]. *J Cheminform*, 2013, 5(1): 28.
- [85] Liu X Y, Liu J Y, Fu B Z, *et al*. DCABM-TCM: A database of constituents absorbed into the blood and metabolites of traditional Chinese medicine [J]. *J Chem Inf Model*, 2023, 63(15): 4948-4959.
- [86] Fang Y C, Huang H C, Chen H H, *et al*. TCMGeneDIT: A database for associated traditional Chinese medicine, gene and disease information using text mining [J]. *BMC Complement Altern Med*, 2008, 8: 58.
- [87] Wang X, Zhang J J, He S M, *et al*. HMOD: An omics database for herbal medicine plants [J]. *Mol Plant*, 2018, 11(5): 757-759.
- [88] Ye Y N, Liang D F, Yi J H, *et al*. IGTCM: An integrative genome database of traditional Chinese medicine plants [J]. *Plant Genome*, 2023, 16(2): e20317.
- [89] Meng F B, Tang Q, Chu T Z, *et al*. TCMPG: An integrative database for traditional Chinese medicine plant genomes [J]. *Hortic Res*, 2022, 9: uhac060.
- [90] Meng F B, Chu T Z, Hu L J, *et al*. TCMPG 2.0: An enhanced database of traditional Chinese medicine plant genomes [J]. *Med Plant Biol*, 2024, 3(1).
- [91] Huang J, Zheng Y X, Wu W X, *et al*. CEMTDD: The database for elucidating the relationships among herbs, compounds, targets and related diseases for Chinese ethnic minority traditional drugs [J]. *Oncotarget*, 2015, 6(19): 17675-17684.
- [92] Sun C, Huang J P, Tang R, *et al*. CPMCP: A database of Chinese patent medicine and compound prescription [J]. *Database*, 2022, 2022: baac073.
- [93] Hong Y F, Xu H Q, Liu Y H, *et al*. DDID: A comprehensive resource for visualization and analysis of diet-drug interactions [J]. *Brief Bioinform*, 2024, 25(3): bbac212.
- [94] Yue W J, Ji W D, Wang X Y, *et al*. SDPR: Prescription recommendation with syndrome differentiation in traditional Chinese medicine [J]. *IEEE J Biomed Health Inform*, 2025, 29(5): 3736-3749.
- [95] Pati S, Baid U, Edwards B, *et al*. Federated learning enables big data for rare cancer boundary detection [J]. *Nat Commun*, 2022, 13(1): 7346.
- [96] 孙孟姊, 张斌, 孙美晨, 等. 持续健康监测可穿戴设备技术融合和临床应用与发展挑战 [J]. *中国预防医学杂志*, 2025, 26(5): 535-541.
- [97] 许晓娜. 基于数据挖掘和网络药理学分析中医药辨治新冠肺炎的用药规律和机制 [D]. 郑州: 河南中医药大学, 2023.

[责任编辑 潘明佳]