Chinese Journal of Natural Medicines

Advancing network pharmacology with artificial intelligence: the next paradigm in traditional Chinese medicine

Xin Shao, Yu Chen, Jinlu Zhang, Xuting Zhang, Yizheng Dai, Xin Peng, Xiaohui Fan

Citation: Xin Shao, Yu Chen, Jinlu Zhang, Xuting Zhang, Yizheng Dai, Xin Peng, Xiaohui Fan, Advancing network pharmacology with artificial intelligence: the next paradigm in traditional Chinese medicine, Chinese Journal of Natural Medicines, 2025, 23(11), 1358−1376. doi: 10.1016/S1875−5364(25)60941−1.

View online: https://doi.org/10.1016/S1875−5364(25)60941−1

Chinese Journal of Natural Medicines 23 (2025) 1358-1376

Contents lists available at ScienceDirect

Chinese Journal of Natural Medicines

journal homepage: www.cjnmcpu.com/

Review

Advancing network pharmacology with artificial intelligence: the next paradigm in traditional Chinese medicine

Xin Shao[a,b,c,*], Yu Chen[d], Jinlu Zhang[a,b], Xuting Zhang[a,b], Yizheng Dai[a,b], Xin Peng[b,*], Xiaohui Fan[a,b,c,*]

[a] *Zhejiang Key Laboratory of Chinese Medicine Modernization, Innovation Center of Yangtze River Delta, Zhejiang University, Jiaxing 314100, China*

[b] *The Joint-laboratory of clinical multi-omics research between Zhejiang University and Ningbo Municipal Hospital of TCM, Ningbo Municipal Hospital of TCM, Ningbo 315010, China*

[c] *Pharmaceutical Informatics Institute, College of Pharmaceutical Sciences, Zhejiang University, Hangzhou 310058, China*

[d] *School and Hospital of Stomatology, Wenzhou Medical University, Wenzhou 325035, China*

ARTICLE INFO

Article history:
Received 2 January 2025
Revised 14 February 2025
Accepted 5 April 2025
Available online 20 November 2025

Keywords:
Traditional Chinese medicine
Network pharmacology
Artificial intelligence
Efficacy evaluation
Mechanism elucidation
Network construction
Network analysis

ABSTRACT

Network pharmacology has gained widespread application in drug discovery, particularly in traditional Chinese medicine (TCM) research, which is characterized by its "multi-component, multi-target, and multi-pathway" nature. Through the integration of network biology, TCM network pharmacology enables systematic evaluation of therapeutic efficacy and detailed elucidation of action mechanisms, establishing a novel research paradigm for TCM modernization. The rapid advancement of machine learning, particularly revolutionary deep learning methods, has substantially enhanced artificial intelligence (AI) technology, offering significant potential to advance TCM network pharmacology research. This paper describes the methodology of TCM network pharmacology, encompassing ingredient identification, network construction, network analysis, and experimental validation. Furthermore, it summarizes key strategies for constructing various networks and analyzing constructed networks using AI methods. Finally, it addresses challenges and future directions regarding cell-cell communication (CCC)-based network construction, analysis, and validation, providing valuable insights for TCM network pharmacology.

1. Introduction

Since its initial proposal by Andrew L. Hopkins in 2007, the concept of "network pharmacology" has emerged as a revolutionary approach to drug discovery, particularly in the research of traditional Chinese medicine (TCM) [1-3]. Unlike the conventional "one drug, one target, one disease" paradigm, network pharmacology posits that modulating multiple nodes within a disease-associated biological network is essential for phenotypic regulation [2]. This principle aligns with the holistic philosophy of TCM, which emphasizes a "multi-component, multi-target, and multi-pathway" mode of action [4]. Through the integration of network biology, TCM network pharmacology facilitates systematic assessment of therapeutic efficacy and detailed elucidation of action mechanisms, establishing a novel paradigm for TCM modernization [5]. For instance, Wu et al. developed a network recovery index (NRI) based on constructed diseased and treated biological networks derived from genome-wide transcriptional expression data, providing a quantitative method for evaluating multi-component drug efficacy [6]. Similarly, Liao et al. introduced the NTRA network analysis algorithm, which identifies key nodes by integrating disease network topology and transcriptomic data, revealing the mechanism of Shen Qi Fu Zheng Injection in treating myocardial ischemia-reperfusion injury through pathway enrichment analysis (PEA) [7]. The significant advances in TCM network pharmacology have been facilitated by rapid progress in artificial intelligence (AI)-related disciplines and widespread adoption of high-throughput omics technologies, enhancing the ability to decode TCM's complex therapeutic mechanisms.

AI encompasses the field of science and engineering focused on computational understanding and simulation of intelligent behaviors, specifically achieving human-level performance in real-world tasks. The advancement of machine learning, particularly breakthrough deep learning methods, has accelerated AI technology development and its applications in biological and medical research [8]. For example, Shao et al. developed scDeepSort, a pre-trained cell type annotation method for single-cell ribonucleic acid-sequencing (scRNA-seq) data using a graph neural network (GNN)-based deep learning model [9]. This method achieved superior accuracy in cell type annotation across normal and diseased tissues. Fang et al. introduced KANO, a deep learning method incorporating external fundamental domain knowledge for knowledge graph-enhanced molecular contrastive learning, demonstrating enhanced performance in molecular property prediction [10]. Recently, generative large language models (LLMs) such as DeepSeek [11], ChatGPT [12], and LLaMA [13] have transformed natural language processing (NLP) tasks, leading to specialized medical LLMs, including TCMChat [14], BianQue [15], and HuatuoGPT [16], for healthcare and TCM applications. These AI technologies

* Corresponding author.
E-mail addresses: E-mail: xin_shao@zju.edu.cn (X. Shao); pengx@nit.zju.edu.cn (X. Peng); fanxh@zju.edu.cn (X. Fan)

https://doi.org/10.1016/S1875-5364(25)60941-1

present significant opportunities and challenges for TCM modernization.

This study first outlines four general steps in TCM network pharmacology research: ingredient identification, network construction, network analysis, and experimental validation. Subsequently, it details the core methodological steps of network construction and analysis, summarizing strategies for developing the ingredient-target network (IN), disease/symptom/syndrome-protein/gene network (DN), and background network (BN), along with their analysis. Finally, it addresses challenges and future directions, highlighting the integration of single-cell and spatial omics data for network construction at higher resolutions, anticipated developments in biological cell-cell communication (CCC) network characterization, and the adoption of advanced LLM-related AI methods in TCM network pharmacology research.

2. General steps in TCM network pharmacology research

TCM network pharmacology research encompasses four essential steps: (1) ingredient identification, (2) network construction, (3) network analysis, and (4) experimental validation (Fig. 1A). The fundamental step involves characterizing the chemical constituents of a specific herbal medicine or formula, establishing the molecular foundation for subsequent analyses. Current methodologies primarily utilize liquid chromatography-tandem mass spectrometry (LC-MS/MS) for structural characterization and identification of chemical compounds. For complex multi-herb formulations, DNA barcoding has emerged as an effective complementary approach [17]. This technique enables accurate herb identification through alignment with established marker genes in herbal genomic databases, facilitating the correlation of detected compounds with their botanical sources.

The second step involves constructing three distinct types of networks: IN, DN, and BN, which typically encompass the protein-protein interaction (PPI) network/gene co-expression network, or recently emerged CCC network. These networks are functionally interconnected through protein/gene nodes, which act as essential bridges between the IN and DN. The BN provides biological context for molecular relationships among protein/gene nodes, while the DN establishes molecular signatures within disease, symptom, or syndrome contexts.

Third, network distance metrics and topological analyses are applied to evaluate therapeutic efficacy through integrated assessment of the IN, DN, and BN. This methodology enables the identification of significantly enriched pathways and biological processes based on shared nodes within the integrated network.

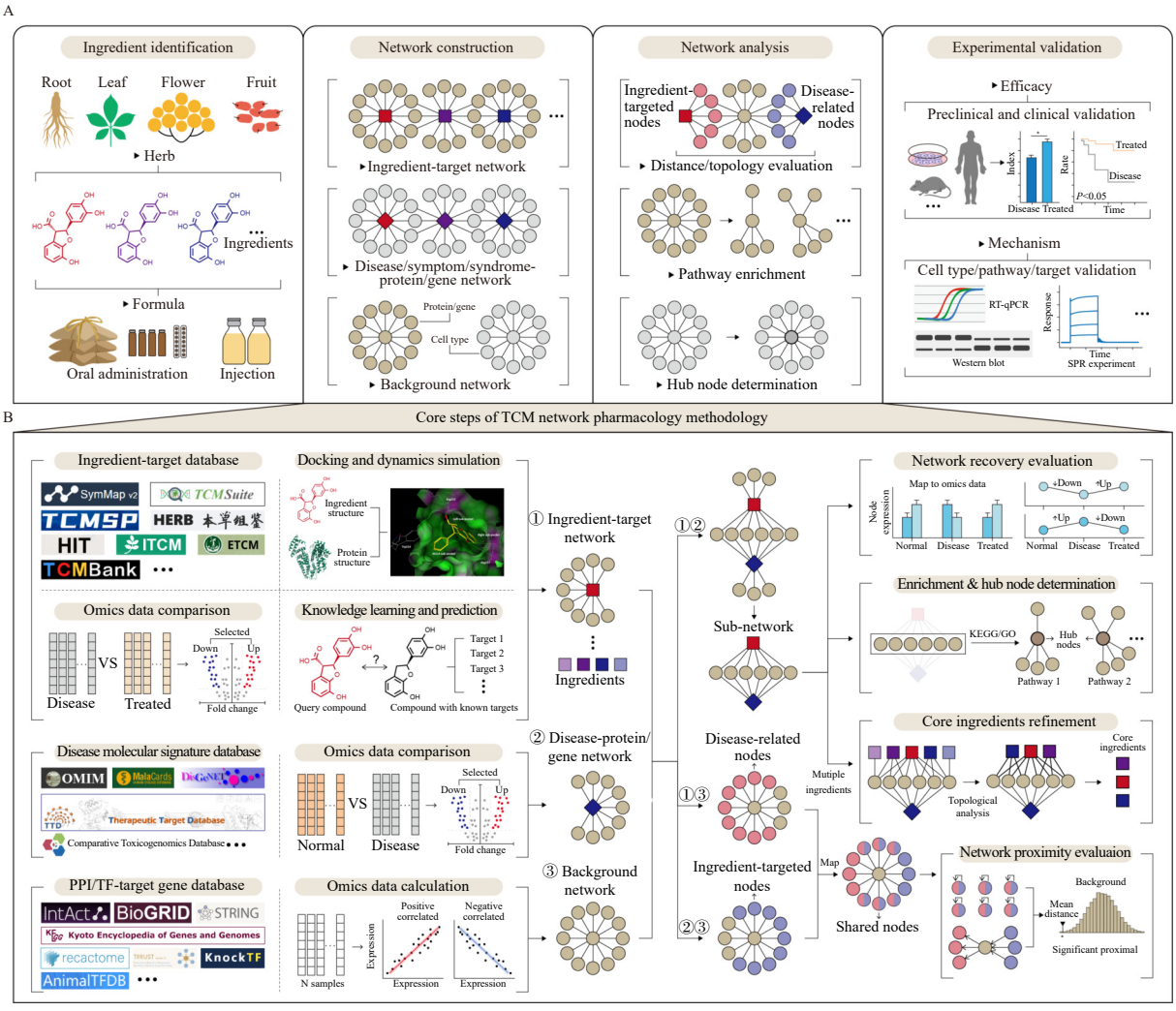

Fig. 1 Overview of TCM network pharmacology research. (A) General steps in TCM network pharmacology research including ingredient identification, network construction, network analysis, and experimental validation. For a given herb or formula, ingredients need to be identified, followed by the construction and analysis of IN, DN, and BN. The preclinical, clinical, and molecular biology experiments are required to further validate the efficacy and mechanisms. (B) Core steps of TCM network pharmacology methodology, i.e., network construction and network analysis. The network construction includes the construction of IN, DN, and BN. The evaluation of efficacy refers to the network recovery and proximity evaluation. The elucidation of mechanisms of action represents enrichment & hub node determination and core ingredients refinement, corresponding to the biological and chemical aspects, respectively.

Furthermore, it facilitates the determination of hub nodes associated with disease pathology and TCM intervention, thereby elucidating the mechanisms underlying TCM's therapeutic effects.

The final and most crucial step involves experimental validation. Based on network analysis results, appropriate preclinical and clinical experiments are necessary to confirm TCM's efficacy in treating the corresponding disease/symptom/syndrome. Regarding mechanisms, cellular and molecular biology experiments must be conducted to verify the identified cell types, pathways, and hub proteins targeted by TCM. Network construction and analysis represent the two core components of TCM network pharmacology methodology. Although these steps vary among current TCM network research, they substantially rely on bioinformatics and computational biology, generating diverse research strategies and numerous relevant AI methods. Therefore, the following discussion focuses on network construction and analysis with AI methods.

3. Network construction with AI methods

Given the distinct characteristics of IN, DN, and BN, the challenges and strategies for constructing these networks differ significantly (Fig. 1B). In IN construction , the primary goal is to identify intrinsic associations between chemical entities and biological molecules, particularly mapping ingredients to their corresponding target proteins or genes. DN construction entails identifying key proteins or genes implicated in the onset or progression of diseases, symptoms, or syndromes, with the objective of establishing robust links between phenotypic traits and molecular mechanisms. BN construction focuses on elucidating molecular interaction networks under unperturbed physiological conditions, providing a foundational baseline. Consequently, considerable research has been devoted to characterizing PPIs and gene regulatory relationships in normal states. This section summarizes the main strategies for constructing IN, DN, and BN, and details the related AI methods in this field (Table 1).

3.1. IN construction

Four primary strategies exist for constructing IN: ingredient-target database-based, docking and dynamics simulation-based, omics data comparison-based, and knowledge learning and prediction-based, as detailed below:

The ingredient-target database-based strategy enables direct acquisition of proteins/genes targeted by ingredients from established TCM database websites. These databases compile direct targets of ingredients, encompassing information about ingredients, targets, their relationships, and associated data regarding formula, herb, disease, and pathway. The construction of ingredient-target databases follows three approaches: 1) manual collection and verification from experimentally validated literature; 2) text mining approaches with manual verification; 3) integration of high-quality databases combined with AI-based prediction. For instance, TCMSP [18] incorporates Chinese herbs manually collected from the Pharmacopoeia of the People's Republic of China. TCMSP obtained herb ingredients through integration with structure databases like PubChem, while target information originated from DrugBank. The drug-target mappings in TCMSP were derived from either the HIT [19] database or predictions using random forest machine learning. Similar databases include SymMap [20], HERB [21], ITCM [22], ETCM [23], TCMBank [24], LTM-TCM [25], and DCABM-TCM [26]. A notable recent addition, TCMSuite [27], features 235 470 biological ingredients and 1 251 548 marker gene sequences. TCMSuite implements DNA-based biological ingredient identification for herb mixtures. These databases provide an alternative method for identifying ingredient targets and constructing IN for subsequent TCM network pharmacology research.

The docking and dynamics simulation-based strategy requires three-dimensional structures of ingredients and proteins for molecular docking or molecular dynamics simulation. Molecular docking represents a frequently employed method for predicting potential targets of specific ingredients. The process involves docking the ingredient (ligand) into the protein's (receptor) binding cavity and exploring possible conformations. Currently, docking processes utilize two types of conformational search algorithms: systematic and stochastic search methods [28]. Methods such as DOCK [29], FRED [30], and Surflex [31] employ systematic search using incremental construction algorithms, progressively building the ligand in the binding site. Alternatively, Gold [32] and AutoDock [33] utilize stochastic search through genetic algorithms to address high computational costs based on evolution and natural selection principles. Molecular docking algorithms then perform quantitative predictions of binding energetics, generating ranking scores based on ligand-receptor complex binding affinity, as seen in eHiTS [34], GLIDE [35], EUDOC [36], FlexX [37], Flog [38], SLIDE [39], Cdocker [40], GlamDock [41], PLANTS [42], MolDock [43], and EADock [44]. Potential targets for specific ingredients can be identified through ranking scores by incorporating target libraries such as PDB and Pocketome. Deep learning models like KarmaDock [45], CarsiDock [46], and transformer-based FeatureDock [47] demonstrate superior performance in efficient, accurate large library ligand docking. These approaches, framing molecular docking as a regression problem, significantly reduce runtime compared to traditional methods. Recent methods conceptualize molecular docking as a generative modeling problem, achieving improved virtual screening performance [48-50]. For example, DynamicBind predicts ligand-specific protein-ligand complex structures using equivariant geometric diffusion networks [49], while SurfDock employs generative diffusion models for reliable binding pose generation [50], demonstrating diffusion models' advantages in generative modeling. Molecular dynamics simulation can verify docking results by assessing the docked ligand-receptor complex stability. Notable molecular dynamics simulation tools include AMBER [51], CHARMM [52], and GROMACS [53].

In the omics data comparison-based strategy, researchers commonly analyze differentially expressed proteins (DEPs) or differentially expressed genes (DEGs) from high-throughput proteomics or transcriptomics data to establish potential associations between treated compounds and proteins/genes. This analysis employs statistical models to identify features with significant expression differences between diseased and treated conditions. The current methodological approaches vary based on data distribution assumptions, including the conventional T-test and fold change-based approach, negative binomial-based DESeq2 [54], edgeR [55], and EBSeq [56], log-normal distribution-based Ballgown [57], and non-parametric test-based NOISeq [58]. Specifically, omics data from diseased and treated conditions undergo significance testing using these statistical models to determine significantly differentially expressed genes/proteins through *P* values. By combining the fold change of gene/protein expression between different conditions, significantly high and low expressed proteins/genes (DEPs/DEGs) are filtered and designated as ingredient-associated nodes for IN construction (Fig. 2). Adequate biological replicates are essential to minimize expression variability within conditions and ensure statistical power in detecting differentially expressed features. The comparison can be conducted in pairwise, multi-conditional, or time series formats between conditions to identify significantly expressed proteins/genes for IN construction.

The knowledge-guided prediction strategy requires prior information on drug–target interactions (DTIs) or drug–target binding affinities (DTAs), which serve as essential reference data for model training and validation [59]. This approach involves

Table 1 AI methods for network construction and analysis.

Name	Underlying AI algorithm	Strategy	Network type	Involved core step	URL of web/code
SymMap* [20]	Merge databases, manual collection	Ingredient-target database	IN	Network construction	http://www.symmap.org/
TCMSP [18]	Merge databases, manual collection, random forest	Ingredient-target database	IN	Network construction	https://www.tcmsp-e.com/#/home
HERB* [21]	Merge databases, manual collection, Fisher's test	Ingredient-target database	IN	Network construction	http://47.92.70.12/
HIT [19]	Manual collection, PubTator central api	Ingredient-target database	IN	Network construction	http://hit2.badd-cao.net/
ITCM* [22]	Manual collection, iDrug algorithm	Ingredient-target database	IN	Network construction	http://itcm.biotcm.net/
ETCM [23]	Manual collection, D3CARP similarity search, hypergeometric test	Ingredient-target database	IN	Network construction	http://www.tcmip.cn/ETCM2/front/#
TCMBank* [24]	Merge databases, manual collection, TextRank	Ingredient-target database	IN	Network construction	https://tcmbank.cn
TCMSuite* [27]	Merge databases, manual collection, HMM	Ingredient-target database	IN	Network construction	http://tcm-suite.aimicrobiome.cn/
LTM-TCM [25]	Merge databases, manual collection, BioNLP	Ingredient-target database	IN	Network construction	http://cloud.tasly.com/#/tcm/home
DCABM-TCM* [26]	Merge databases, manual collection	Ingredient-target database	IN	Network construction	http://bionet.ncpsb.org.cn/dcabm-tcm/#/Home
DOCK* [29]	Systematic docking, incremental construction, force-field-based scoring	Docking and dynamics simulation	IN	Network construction	https://dock.compbio.ucsf.edu/
FRED* [30]	Systematic docking, incremental construction	Docking and dynamics simulation	IN	Network construction	https://www.eyesopen.com/oedocking
Surflex* [31]	Systematic docking, incremental construction	Docking and dynamics simulation	IN	Network construction	https://htpsurflexdock.biocomp.uenf.br/
eHiTS* [34]	Systematic docking, incremental construction, and binding site scoring	Docking and dynamics simulation	IN	Network construction	https://bip.weizmann.ac.il/toolbox/structure/ehits.htm
GLIDE* [35]	Systematic docking, incremental construction, empirical scoring	Docking and dynamics simulation	IN	Network construction	https://www.schrodinger.com/
EUDOC [36]	Systematic docking, matching algorithm, database	Docking and dynamics simulation	IN	Network construction	-
FlexX* [37]	Systematic docking, incremental construction	Docking and dynamics simulation	IN	Network construction	http://www.biosolveit.de/FlexX
Flog [38]	Systematic docking, matching algorithm, database	Docking and dynamics simulation	IN	Network construction	-
SLIDE* [39]	Systematic docking, incremental construction	Docking and dynamics simulation	IN	Network construction	https://kuhnlab.natsci.msu.edu/software/slide/
Gold* [32]	Stochastic docking, genetic algorithm, force-field-based scoring	Docking and dynamics simulation	IN	Network construction	https://www.ccdc.cam.ac.uk/solutions/software/gold/
AutoDock* [33]	Stochastic docking, genetic algorithm, force-field-based scoring	Docking and dynamics simulation	IN	Network construction	https://autodock.scripps.edu/
Cdocker* [40]	Stochastic docking, molecular dynamics	Docking and dynamics simulation	IN	Network construction	https://www.computabio.com/cdocker-tutorial.html
GlamDock [41]	Stochastic docking, Monte Carlo algorithm	Docking and dynamics simulation	IN	Network construction	-
PLANTS* [42]	Stochastic docking, ant colony optimization	Docking and dynamics simulation	IN	Network construction	https://github.com/discoverdata/parallel-PLANTS
MolDock* [43]	Stochastic docking, force-field-based scoring	Docking and dynamics simulation	IN	Network construction	http://molexus.io/molegro-virtual-docker/
EADock* [44]	Stochastic docking, hybrid evolutionary algorithm	Docking and dynamics simulation	IN	Network construction	https://github.com/wenyizng/EDock
KarmaDock* [45]	Graph neural network, mixture density network	Docking and dynamics simulation	IN	Network construction	https://github.com/schrojunzhang/KarmaDock
CarsiDock* [46]	Deep learning, geometry optimization procedure	Docking and dynamics simulation	IN	Network construction	https://github.com/carbonsilicon-ai/CarsiDock
FeatureDock* [47]	Local environment learning, transformer	Docking and dynamics simulation	IN	Network construction	https://github.com/xuhuihuang/featuredock
DIFFDOCK* [48]	Diffusion generative model	Docking and dynamics simulation	IN	Network construction	https://github.com/gcorso/DiffDock
DynamicBind* [49]	Equivariant geometric diffusion network	Docking and dynamics simulation	IN	Network construction	https://github.com/luwei0917/DynamicBind
SurfDock* [50]	Surface-informed diffusion generative model	Docking and dynamics simulation	IN	Network construction	https://github.com/CAODH/SurfDock
AMBER* [51]	Molecular dynamics, energy minimization	Docking and dynamics simulation	IN	Network construction	https://ambermd.org/
CHARMM* [52]	Molecular dynamics, energy minimization	Docking and dynamics simulation	IN	Network construction	https://academiccharmm.org/
GROMACS* [53]	Molecular dynamics, energy minimization	Docking and dynamics simulation	IN	Network construction	http://www.gromacs.org
DESeq2* [54]	Negative binomial	Omics data comparison	IN/DN	Network construction	https://github.com/thelovelab/DESeq2
edgeR* [55]	Negative binomial	Omics data comparison	IN/DN	Network construction	https://bioinf.wehi.edu.au/edgeR/
EBSeq* [56]	Negative binomial	Omics data comparison	IN/DN	Network construction	https://github.com/lengning/EBSeq
Ballgown* [57]	Log-normal distribution	Omics data comparison	IN/DN	Network construction	https://github.com/alyssafrazee/ballgown
NOISeq* [58]	Non-parametric test for any distribution	Omics data comparison	IN/DN	Network construction	https://github.com/ConesaLab/NOISeq
SRP [61]	Similarity, nearest neighbors	Knowledge learning and prediction	IN	Network construction	-
LapRLS [62]	Bipartite local model	Knowledge learning and prediction	IN	Network construction	-

Continued

Name	Underlying AI algorithm	Strategy	Network type	Involved core step	URL of web/code
BLM-NII [63]	Bipartite local model	Knowledge learning and prediction	IN	Network construction	-
NBI [64]	Network diffusion, bipartite network topology similarity	Knowledge learning and prediction	IN	Network construction	-
NRWRH [65]	Network diffusion, random walk algorithm	Knowledge learning and prediction	IN	Network construction	-
PSL [66]	Network diffusion, probabilistic rules, and multiple similarity	Knowledge learning and prediction	IN	Network construction	-
DASPfind* [67]	Network diffusion, similarity	Knowledge learning and prediction	IN	Network construction	http://www.cbrc.kaust.edu.sa/daspfind
KBMF2K [68]	Matrix factorization, Bayesian, binary classification	Knowledge learning and prediction	IN	Network construction	http://users.ics.aalto.fi/gonen/kbmf2k
PMF [69]	Matrix factorization, active learning	Knowledge learning and prediction	IN	Network construction	http://www.csb.pitt.edu/Faculty/bahar/files/
CMF [70]	Matrix factorization, dual similarity, self-paced learning	Knowledge learning and prediction	IN	Network construction	-
WGRMF [71]	Matrix factorization, graph regularization	Knowledge learning and prediction	IN	Network construction	-
NRLMF [72]	Matrix factorization, neighborhood regularization	Knowledge learning and prediction	IN	Network construction	-
DNILMF* [73]	Matrix factorization	Knowledge learning and prediction	IN	Network construction	https://github.com/minghao2016/DNILMF
DrugAI* [80]	Feature vector, GNN, CNN, FCL	Knowledge learning and prediction	IN	Network construction	https://github.com/cutezsq9503/DrugAI
iGPCR-drug [74]	Feature vector, fuzzy K-nearest neighbour algorithm	Knowledge learning and prediction	IN	Network construction	http://www.jci-bioinfo.cn/iGPCR-Drug/
EnsemDT [75]	Feature vector, decision tree, kernel ridge regression	Knowledge learning and prediction	IN	Network construction	-
SITAR [76]	Feature vector, logistic regression	Knowledge learning and prediction	IN	Network construction	-
RFDT [77]	Feature vector, rotation forest algorithm	Knowledge learning and prediction	IN	Network construction	-
PDTPS [78]	Feature vector, PCA, relevance vector machine	Knowledge learning and prediction	IN	Network construction	-
ER-Tree [79]	Feature vector, pseudo-substitution matrix representation	Knowledge learning and prediction	IN	Network construction	-
FusionDTA* [81]	Feature vector, BiLSTM, Linear attention	Knowledge learning and prediction	IN	Network construction	https://github.com/yuanweining/FusionDTA
IGT* [82]	Feature vector, graph transformer	Knowledge learning and prediction	IN	Network construction	https://github.com/microsoft/IGT-Intermolecular-Graph-Transformer
MolTrans* [83]	Feature vector, transformer	Knowledge learning and prediction	IN	Network construction	https://github.com/kexinhuang12345/MolTrans
DeepMGT-DTI* [84]	Feature vector, GCN, CNN, FC	Knowledge learning and prediction	IN	Network construction	https://github.com/zhangpl109/DeepMGT-DTI
MHTAN-DTI* [85]	Feature vector, attention, FC	Knowledge learning and prediction	IN	Network construction	https://github.com/ranzhran/MHTAN-DTI
GraphormerDTI* [86]	Feature vector, graph transformer, FC	Knowledge learning and prediction	IN	Network construction	https://github.com/mengmeng34/GraphormerDTI
DrugormerDTI* [87]	Feature vector, graph transformer, residual2vec	Knowledge learning and prediction	IN	Network construction	https://github.com/joannacatj/drugormerDTI
DTITR* [88]	Feature vector, transformer, FC	Knowledge learning and prediction	IN	Network construction	https://github.com/larngroup/DTITR
AttentionMGT-DTA* [89]	Feature vector, graph transformer, cross attention, FC	Knowledge learning and prediction	IN	Network construction	https://github.com/JK-Liu7/AttentionMGT-DTA
MUSE* [90]	Variational expectation-maximization, GNN	Knowledge learning and prediction	IN	Network construction	https://github.com/biomed-AI/MUSE
DTI-LM* [91]	Pretrained language models, graph attention network	Knowledge learning and prediction	IN	Network construction	https://github.com/compbiolabucf/DTI-LM/
DTIAM* [92]	Graph transformer, self-supervised pre-training	Knowledge learning and prediction	IN	Network construction	https://github.com/CSUBioGroup/DTIAM
DMFF-DTA* [93]	Feature vector, graph transformer, GNN	Knowledge learning and prediction	IN	Network construction	https://github.com/hehh77/DMFF-DTA
SP-DTI* [94]	Pretrained language models, graph transformer	Knowledge learning and prediction	IN	Network construction	https://github.com/Steven51516/SP-DTI
OMIM* [95]	Manual collection	Disease molecular signature database	DN	Network construction	https://omim.org/
MalaCards [96]	Merge databases, manual collection, phenotypic disease network	Disease molecular signature database	DN	Network construction	https://www.malacards.org/
DisGeNET [97]	Merge databases, manual collection, text mining	Disease molecular signature database	DN	Network construction	https://www.disgenet.org/
TTD* [98]	Merge databases, manual collection, DESeq2, BLAST, iCn3D	Disease molecular signature database	DN	Network construction	https://db.idrblab.net/ttd/
CTD* [99]	Merge databases, manual collection	Disease molecular signature database	DN	Network construction	https://ctdbase.org/
CHD@ZJU [100]	Merge databases, manual collection, text mining	Disease molecular signature database	DN	Network construction	http://tcm.zju.edu.cn/chd/
MarkerDB* [101]	Merge databases, manual collection	Disease molecular signature database	DN	Network construction	https://markerdb.ca/
DMRdb* [102]	Merge databases, manual collection	Disease molecular signature database	DN	Network construction	http://www.inbirg.com/DMRdb/
Deng et al. [103]	Merge databases, manual collection	Disease molecular signature database	DN	Network construction	https://proteome-phenome-atlas.com/
IntAct* [104]	Manual collection	PPI database	BN	Network construction	https://www.ebi.ac.uk/intact/

Continued

Name	Underlying AI algorithm	Strategy	Network type	Involved core step	URL of web/code
BioGRID* [105]	Manual collection	PPI database	BN	Network construction	https://thebiogrid.org/
STRING* [106]	Merge databases, manual collection, and automatic text mining	PPI database	BN	Network construction	https://string-db.org/
KEGG [107]	Manual collection	PPI database	BN	Network construction	http://www.genome.ad.jp/kegg/
Reactome* [108]	Manual collection	PPI database	BN	Network construction	https://reactome.org/
Hitpredict [109]	Merge databases	PPI database	BN	Network construction	https://www.hitpredict.org/
GeneMANIA [110]	Merge databases	PPI database	BN	Network construction	https://genemania.org/
HumanNet* [111]	Merge databases, Bayesian statistics	PPI database	BN	Network construction	https://staging2.inetbio.org/humannetv3/
HPC-Atlas* [112]	Merge databases, deep forest	PPI database	BN	Network construction	http://www.yulpan.top/HPC-Atlas/
Laman et al.* [113]	Data analysis	PPI database	BN	Network construction	https://www.ebi.ac.uk/biostudies/studies/S-BSST1423
TRRUST* [114]	Manual collection	TF-target gene database	BN	Network construction	https://www.grnpedia.org/trrust/
KnockTF* [115]	Manual collection	TF-target gene database	BN	Network construction	http://www.licpathway.net/KnockTF/index.html
AnimalTFDB* [116]	Merge databases, PfamScan, BLAST	TF-target gene database	BN	Network construction	https://guolab.wchscu.cn/AnimalTFDB4//#/
hTFtarget [117]	Merge databases, manual collection	TF-target gene database	BN	Network construction	http://bioinfo.life.hust.edu.cn/hTFtarge
ChIPBase* [118]	Merge databases, manual collection, liftOver	TF-target gene database	BN	Network construction	http://rna.sysu.edu.cn/chipbase/
TFLink* [119]	Merge databases, manual collection	TF-target gene database	BN	Network construction	https://tflink.net/
TransmiR* [120]	Merge databases, manual collection, TF binding motifs	TF-target gene database	BN	Network construction	http://www.cuilab.cn/transmir
WGCNA* [121]	Biweight midcorrelation or Spearman correlation	Omics data calculation	BN	Network construction	https://github.com/cran/WGCNA
ARACNE [122]	Mutual information, relevance networks method	Omics data calculation	BN	Network construction	-
ARACNe-AP* [123]	Mutual information, adaptive partitioning	Omics data calculation	BN	Network construction	https://github.com/califano-lab/ARACNe-AP
CLR [124]	Mutual information, context likelihood of relatedness	Omics data calculation	BN	Network construction	-
MRNET [125]	Mutual information, max dependency/relevance, min redundancy	Omics data calculation	BN	Network construction	-
parmigene* [126]	Mutual information, k-nearest neighbor distance	Omics data calculation	BN	Network construction	https://github.com/sales-lab/parmigene
SIRENE [127]	Local binary classification, regulatory interaction	Omics data calculation	BN	Network construction	http://cbio.ensmp.fr/sirene
GENIE3* [128]	Tree-based ensemble, regulatory interaction	Omics data calculation	BN	Network construction	https://github.com/vahuynh/GENIE3
SpaTalk* [131]	Graph modeling, non-negative linear regression, random walk	Omics data calculation	BN	Network construction	https://github.com/ZJUFanLab/SpaTalk
COMMOT* [133]	Graph modeling, optimal transport	Omics data calculation	BN	Network construction	https://github.com/zcang/COMMOT
HoloNet* [134]	Graph modeling, attention-based graph learning	Omics data calculation	BN	Network construction	https://github.com/lhc17/HoloNet
CellChat* [135]	Permutation test	Omics data calculation	BN	Network construction	https://github.com/sqjin/CellChat
CellPhoneDB* [136]	Permutation test	Omics data calculation	BN	Network construction	https://github.com/Teichlab/cellphonedb
NicheNet* [137]	Linear regression	Omics data calculation	BN	Network construction	https://github.com/saeyslab/nichenetr
MEBOCOST* [138]	Permutation test	Omics data calculation	BN	Network construction	https://github.com/kaifuchenlab/MEBOCOST
miRTalk* [139]	Permutation test	Omics data calculation	BN	Network construction	https://github.com/multitalk/miRTalk
CITEdb* [140]	Manual collection	CCC database	BN	Network construction	https://citedb.cn/
MACC* [141]	Manual collection	CCC database	BN	Network construction	http://macc.badd-cao.net/
NRI [6]	Regulating score, fold change	Network recovery evaluation	IN + DN	Network analysis	-
Cmap* [142, 143]	Scoring of rank-ordered genes	Network recovery evaluation	IN + DN	Network analysis	https://github.com/cmap/
Ding et al. [144]	CIPHER algorithm	Core ingredients refinement	IN + DN	Network analysis	-
Wang et al. [145]	Text mining, pharmacophore model-based prediction	Core ingredients refinement	IN	Network analysis	-
Cheng et al.* [146]	Shortest path length of nodes	Network proximity evaluation	DN + BN	Network analysis	https://github.com/emreg00/toolbox
Wang et al.* [147]	Community cohesion looseness	Network proximity evaluation	IN + DN + BN	Network analysis	https://github.com/BISLshwang/Community-cohesion-scores
Metascape* [148]	Enrichment, hypergeometric test	Enrichment & hub node determination	IN + DN + BN	Network analysis	https://metascape.org/
GSEA* [149]	Enrichment, fold change-based gene ranking	Enrichment & hub node determination	IN + DN + BN	Network analysis	https://www.gsea-msigdb.org/gsea/index.jsp

Continued

Name	Underlying AI algorithm	Strategy	Network type	Involved core step	URL of web/code
EGSEA* [150]	Enrichment, collective gene set scoring	Enrichment & hub node determination	IN + DN + BN	Network analysis	http://www.bioconductor.org/packages/EGSEA/
GIGSEA [151]	Enrichment, weighted linear regression	Enrichment & hub node determination	IN + DN + BN	Network analysis	-
GSVA* [152]	Enrichment, matrix factorization	Enrichment & hub node determination	IN + DN + BN	Network analysis	https://bioconductor.org/packages/GSVA
PLAGE [153]	Enrichment, matrix factorization	Enrichment & hub node determination	IN + DN + BN	Network analysis	http://dulci.biostat.edu/pathways
ReporterScore* [154]	Enrichment, generalized reporter score	Enrichment & hub node determination	IN + DN + BN	Network analysis	https://github.com/Asa12138/ReporterScore
CTpathway* [155]	Enrichment, graph modeling	Enrichment & hub node determination	IN + DN + BN	Network analysis	https://github.com/Bioccjw/CTpathway
CADIA* [156]	Enrichment, graph modeling	Enrichment & hub node determination	IN + DN + BN	Network analysis	https://github.com/pouryany/CADIA
NEA [157]	Enrichment, fast network randomization	Enrichment & hub node determination	IN + DN + BN	Network analysis	http://http/www.meb.ki.se/ ~ yudpaw
EnrichNet* [158]	Enrichment, random walk with restart	Enrichment & hub node determination	IN + DN + BN	Network analysis	http://www.enrichnet.org
LPIA* [159]	Enrichment, network-based edge weighting	Enrichment & hub node determination	IN + DN + BN	Network analysis	https://math.bu.edu/LPIA/
LEGO [160]	Enrichment, network-based gene weighting	Enrichment & hub node determination	IN + DN + BN	Network analysis	http://omics.fudan.edu.cn/static/softwares/Lego.tar.gz.
GANPA* [161]	Enrichment, network-based gene weighting	Enrichment & hub node determination	IN + DN + BN	Network analysis	http://cran.r-project.org/web/packages/GANPA/index.html
MiNEApy* [162]	Enrichment, flux variability analysis	Enrichment & hub node determination	IN + DN + BN	Network analysis	https://github.com/vpandey-om/mineapy
NetworkAnalyzer* [163]	Hub determination, network topology	Enrichment & hub node determination	IN + DN + BN	Network analysis	http://www.cytoscape.org
CentiScaPe* [164]	Hub determination, network topology	Enrichment & hub node determination	IN + DN + BN	Network analysis	http://www.cytoscape.org
cytoHubba* [165]	Hub determination, network topology	Enrichment & hub node determination	IN + DN + BN	Network analysis	http://www.cytoscape.org
MCODE [166]	Hub determination, network topology, local neighborhood density	Enrichment & hub node determination	IN + DN + BN	Network analysis	https://ftp.mshri.on.ca/pub/BIND/Tools/MCODE
ClusterONE* [167]	Hub determination, network topology, greedy growth process	Enrichment & hub node determination	IN + DN + BN	Network analysis	http://www.paccanarolab.org/cluster-one/
IPCA [168]	Hub determination, network topology, sub-graph density	Enrichment & hub node determination	IN + DN + BN	Network analysis	http://netlab.csu.edu.cn/bioinformatics/limin/IPCA
Deepwalk [169]	Hub determination, network topology, random walk	Enrichment & hub node determination	IN + DN + BN	Network analysis	-
SPICi* [170]	Hub determination, network topology, heuristic approach	Enrichment & hub node determination	IN + DN + BN	Network analysis	http://compbio.cs.princeton.edu/spici
Core* [171]	Hub determination, network topology, core-attachment structure	Enrichment & hub node determination	IN + DN + BN	Network analysis	http://alse.cs.hku.hk/complexes/
DPCMNE* [172]	Hub determination, network topology, deepwalk, core-attachment structure	Enrichment & hub node determination	IN + DN + BN	Network analysis	https://github.com/XiangmaoMeng/DPCMNE
PEWCC* [173]	Hub determination, network topology, weighted clustering coefficient	Enrichment & hub node determination	IN + DN + BN	Network analysis	http://faculty.uaeu.ac.ae/nzaki/Research.htm

The asterisk (*) represents databases and AI methods in line with FAIR (Findable, Accessible, Interoperable, Reusable) data principles. -, not available.

learning from known DTI or DTA data to predict unknown interactions using machine learning and deep learning algorithms. The available drug-related data for knowledge learning encompasses: 1) chemical structure features, including molecular properties and descriptors; 2) pharmacological information regarding absorption, distribution, metabolism, excretion, and toxicity; 3) associated diseases and pathways; 4) drug category information, such as ATC codes and chemical taxonomy. Target-related data includes genomic sequences and biological functions, particularly gene ontology (GO) terms. Numerous methods have been developed to predict DTIs or DTA through knowledge learning [60]. Methods like SRP [61] predict by comparing similarities between drugs with known targets and new compounds, while bipartite local model-based methods such as LapRLS [62] and BLM-NII [63] combine predictions from both drug and target perspectives. Network diffusion-based methods employ graph-based techniques for DTI prediction [64-67], while matrix factorization-based methods reconstruct the DTI matrix through latent feature matrices [68-73]. Additionally, some methods utilize machine learning algorithms with fixed-length feature vectors for DTI prediction [74-79]. Recent approaches increasingly incorporate GNN and transformer methodologies to enhance predictions by learning representations from known DTI or DTA data from databases including DrugBank, BIOSNAP, BindingDB, DAVIS, and KIBA [80-94]. The process involves learning separate feature vectors from drugs and targets, which are then integrated into a joint embedding. This embedding is trained with DTI labels and DTA data using attention mechanisms to construct the prediction model.

In identifying relationships between ingredients and their protein targets, the accuracy of DTI predictions, including false positives and false negatives, requires careful consideration when developing the IN construction strategy. While DTI and DTA predictions represent classification and regression tasks, respectively, recent AI-based methods have demonstrated high performance metrics in both areas. Notably, deep-learning, transformer-based models incorporating attention mechanisms have shown superior performance compared to traditional machine learning approaches. The evaluation of DTI prediction methods employs AUROC and area under the precision-recall curve (AUPRC), while DTA prediction methods are assessed using concordance index (CI), mean square error (MSE), and regression toward the mean (r^2 index). Despite achieving high AUROC, AUPRC, CI, and r^2, along with low MSE on benchmarking datasets, these computational methods demonstrate significantly reduced performance in cold-start scenarios compared to warm-start configurations (Table 2). The cold-start problem, encompassing drug, target, and drug-target cold-start scenarios, evaluates model performance with unseen inputs. Given that most TCM formula ingredients represent unseen molecules for models trained on chemical drugs, cold-start evaluation better reflects real-world

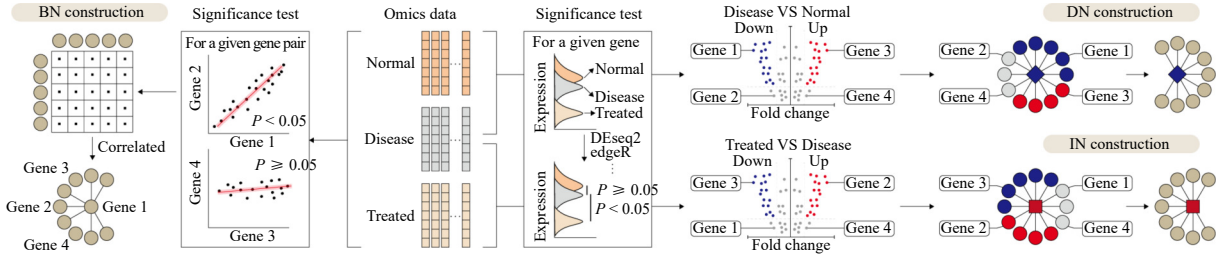

Fig. 2 Schematic pipeline for constructing different types of molecular networks. For BN construction , omics data are subjected to statistical significance testing using established correlation analysis methods to identify strongly correlated gene pairs across samples. Based on the resulting correlation matrix, gene co-expression or gene regulatory networks are generated by linking significantly correlated nodes, while isolated nodes may be excluded to improve network robustness. For DN and IN construction , omics data collected under normal, diseased, and treated conditions are analyzed with statistical models to identify DEGs or DEPs *P* -value thresholds. When combined with fold-change criteria, the filtered DEGs/DEPs are designated as condition-associated nodes and subsequently used to construct the corresponding networks.

TCM network pharmacology applications. Future research should focus on reducing false positive and false negative predictions through enhanced AI methods and increased experimentally validated DTI and DTA data involving TCM ingredients, enabling more accurate IN construction for herbs or formulas with unknown targets.

3.2. DN construction

The construction of DN primarily utilizes disease molecular signature database-based and omics data comparison-based strategies. A straightforward approach involves obtaining disease-related proteins/genes directly from curated high-quality disease molecular signature databases. These databases, including OMIM [95], MalaCards [96], DisGeNET [97], TTD [98], CTD [99], CHD@ZJU [100], and MarkerDB [101], contain experimentally validated proteins/genes compiled through manual collection and verification. These proteins/genes typically either contribute to disease occurrence and development or result from the disease process. Some databases integrate multiple high-quality sources, while others employ AI-based prediction methods to establish disease-protein/gene associations, similar to ingredient-target database construction. For example, DisGeNET combines expertly curated resources with text-mining approaches, specifically the conditional random field one and the BeFree system. The current DisGeNET version presents 429 036 gene-disease associations connecting 17 381 genes to 15 093 diseases, disorders, and clinical or abnormal human phenotypes, providing comprehensive resources for DN construction. Additionally, the recently developed DMRdb evaluates causal relationships between diseases and various molecular and phenotypic factors [102], while a comprehensive plasma proteomics atlas containing 168100 protein-disease associations and 554488 protein-trait associations has been published [103]. These resources represent significant advances in understanding disease mechanisms at the molecular level.

An alternative to the database-based strategy involves identifying potential disease-protein/gene associations through DEPs and DEGs from high-throughput proteomics or transcriptomics data for DN construction. This process begins with statistical analysis of omics data under normal and diseased conditions using established models such as DESeq2 and edgeR to identify significant DEPs/DEGs based on *P* values. The DN is then constructed by selecting disease-associated nodes from DEPs/DEGs, considering expression fold changes between conditions (Fig. 2). Similar to IN construction using omics data comparison, this approach requires adequate biological replicates in pairwise, multi-conditional, or time series formats. The statistical methods for identifying significant expression differences between normal and diseased conditions parallel those used in omics data comparison-based IN construction strategies.

3.3. BN construction

For constructing BN, two primary strategies exist: PPI/transcriptional factor (TF)-target gene database-based and knowledge learning and prediction-based approaches. The conventional method utilizes a human PPI network to elucidate disease and drug patterns, where proteins and physical interactions serve as network nodes and edges, respectively. Several PPI databases have been established, including IntAct [104], BioGRID [105], STRING [106], KEGG [107], Reactome [108], Hitpredict [109], GeneMANIA [110], and HumanNet [111]. STRING, for instance, encompasses undirected known PPIs from curated databases, experimentally determined PPIs, and predicted PPIs represented through various gene relationships and protein homology. The extensively used KEGG contains multiple types of directed PPIs, including processes such as activation, inhibition, repression, binding, and other interactions, manually documented in reference pathway diagrams. The recent HumanNet v3 update expanded its data coverage with enhanced network inference algorithms, encompassing 99.8% of human protein-coding genes. Furthermore, HumanNet v3 offers a functional network extended by co-citation (HumanNet-XC) and a functional gene network (HumanNet-FN) with co-functional links. Recently, researchers developed a human protein complexes atlas (HPC-Atlas) containing 8944 protein complexes involving 16632 proteins through database integration and computational methods [112]. Additionally, the first atlas of tissue-specific protein associations, comprising 116 million protein associations across 11 tissues, demonstrated its utility in prioritizing disease genes and interactions in a tissue-specific context [113]. Beyond the PPI network, the TF-target gene regulatory network represents another form of BN concerning gene expression regulation through TFs. Established databases documenting TF-target gene relationships include TRRUST [114], KnockTF [115], AnimalTFDB [116], hTFtarget [117], ChIPBase [118], TFLink [119], and TransmiR, which contains TF-miRNA regulations [120].

With next-generation RNA-seq technology advancements, omics data calculation presents an alternative strategy for BN construction. Gene co-expression networks, also termed relevance networks, are frequently employed in network pharmacology research as prior knowledge for BN construction. These networks utilize genes as nodes and co-expressed genes as edges between gene pairs. Co-expressed genes are identified through correlation coefficients calculated using methods such as Spearman, Pearson, Kendall Tau, or mutual information from gene expression profiles across multiple samples. Statistical significance testing, including permutation tests, linear regression, Fisher's test, supervised machine learning, or rank-based tests, determines significantly correlated gene pairs. The gene co-expressed network or gene regulatory network is then constructed by filtering and connecting significantly correlated gene nodes and/or removing isolated nodes (Fig. 2). WGCNA [121], for example, enables weighted correlation network construction where edge weights derive from biweight midcorrelation or Spearman correlation. Methods such as ARACNE [122], ARACNe-AP [123], CLR [124], MRNET [125], and parmigene [126] combine mutual information with additional statistical approaches to enhance network accuracy. Incorporat-

Table 2 Summary of representative DTI/DTA prediction methods in recent years.

Methods	Type	Condition	AUROC	AUPRC	CI	MSE	r^2	Benchmarking datasets
DrugAI	DTI	Warm start	0.93	0.87	-	-	-	DrugBank
MUSE	DTI	Warm start	0.92	0.92	-	-	-	ConPLex
MolTrans	DTI	Warm start	0.89	0.90	-	-	-	BIOSNAP
		Drug cold-start	0.85	-	-	-	-	BIOSNAP
		Target cold-start	0.77	-	-	-	-	BIOSNAP
		Warm start	0.91	0.40	-	-	-	DAVIS
		Warm start	0.91	0.62	-	-	-	BindingDB
GraphormerDTI	DTI	Warm start	0.87	0.88	-	-	-	DrugBank
		Drug cold-start	0.83	0.84	-	-	-	DrugBank
		Drug-target cold-start	0.69	0.70	-	-	-	DrugBank
		Warm start	0.89	0.79	-	-	-	DAVIS
		Drug cold-start	0.72	0.47	-	-	-	DAVIS
		Drug-target cold-start	0.65	0.42	-	-	-	DAVIS
		Warm start	0.92	0.79	-	-	-	KIBA
		Drug cold-start	0.86	0.66	-	-	-	KIBA
		Drug-target cold-start	0.74	0.46	-	-	-	KIBA
DeepMGT-DTI	DTI	Warm start	0.90	0.77	-	-	-	DrugBank
DTI-LM	DTI	Warm start	0.96	0.86	-	-	-	DrugBank
		Drug cold-start	0.89	0.67	-	-	-	DrugBank
		Target cold-start	0.94	0.82	-	-	-	DrugBank
		Warm start	0.95	0.84	-	-	-	BindingDB
		Drug cold-start	0.90	0.74	-	-	-	BindingDB
		Target cold-start	0.83	0.46	-	-	-	BindingDB
FusionDTA	DTA	Warm start	-	-	0.91	0.21	0.74	DAVIS
		Warm start	-	-	0.91	0.13	0.79	KIBA
		Drug cold-start	-	-	0.75	0.48	-	DAVIS
		Target cold-start	-	-	0.83	0.33	-	DAVIS
		Drug-target cold-start	-	-	0.69	0.72	-	DAVIS
AttentionMGT-DTA	DTA	Warm start	-	-	0.89	0.19	0.70	DAVIS
		Warm start	-	-	0.89	0.14	0.79	KIBA
		Drug cold-start	-	-	0.70	0.73	0.16	DAVIS
		Target cold-start	-	-	0.83	0.42	0.28	DAVIS
		Drug-target cold-start	-	-	0.61	0.61	0.07	DAVIS
DTIAM	DTA	Warm start	-	-	0.91	0.21	0.86	DAVIS
		Drug cold-start	-	-	0.73	0.60	0.50	DAVIS
		Target cold-start	-	-	0.87	0.31	0.79	DAVIS
		Warm start	-	-	0.87	0.16	0.88	KIBA
		Drug cold-start	-	-	0.79	0.33	0.73	KIBA
		Target cold-start	-	-	0.77	0.31	0.75	KIBA
DMFF-DTA	DTA	Warm start	-	-	0.89	0.22	0.70	DAVIS
		Drug cold-start	-	-	0.74	0.55	0.23	DAVIS
		Target cold-start	-	-	0.84	0.33	0.50	DAVIS
		Drug-target cold-start	-	-	0.66	0.76	0.10	DAVIS
		Warm start	-	-	0.89	0.14	0.77	KIBA
		Drug cold-start	-	-	0.75	0.41	0.4	KIBA
		Target cold-start	-	-	0.75	0.41	0.45	KIBA
		Drug-target cold-start	-	-	0.67	0.57	0.24	KIBA

DTI, drug-target interaction; DTA, drug-target affinity; -, not available; AUROC, area under the receiver operating characteristic curve; AUPRC, area under the precision-recall curve; CI, concordance index; MSE, mean square error; r^2, regression toward the mean.

ing known TF genes allows gene co-expression networks to infer gene regulatory networks with directed edges between correlated TF genes and target genes. SIRENE [127] approaches this by decomposing gene regulatory network inference into multiple local binary classification problems, focusing on distinguishing target genes from non-targets for each TF. Similarly, GENIE3 [128] employs tree-based ensemble methods for regulatory network prediction, achieving superior performance in the DREAM challenge. Both undirected gene co-expression networks and directed gene regulatory networks provide alternative approaches for BN construction in TCM network pharmacology studies.

While traditional approaches primarily focus on PPI networks—given proteins' central role in biological functions—recent advances have highlighted the significance of CCC networks in disease pathogenesis and progression [129]. These developments have been facilitated by single-cell omics technologies, enabling systematic inference of CCC at the cell-type level. In multicellular organisms, cells communicate through signaling molecules, forming intricate and dynamic CCC networks. The disruption and restoration of CCC networks correlate with disease onset and recovery, reflecting the TCM principles of bodily discord and reconciliation. For instance, Shao et al. developed a curated ligand-receptor interaction (LRI) database, CellTalkDB [130], utilizing text mining and manual verification, and created SpaTalk [131] with deep learning algorithms and CellTalkDB for LRI-mediated CCC network inference. Subsequently, they employed these methods to elucidate the CCC network dysregulation mechanisms underlying hepatic ischemia-reperfusion injury following liver transplantation [132]. Notably, COMMOT [133] incorporates competition between various ligands and receptors to enhance the inference of spatially proximal CCC using a collective optimal transport approach. Furthermore, HoloNet [134] analyzes target gene expression in each cell by separating baseline expression determined by cell type from expression changes caused by CCC mediated through multiple LRI pairs, achieving functional CCC event inference using an attention-based graph learning method. Similar CCC network inference methods include the widely-used CellChat [135], CellPhoneDB [136], and NicheNet [137] based on single-cell transcriptomics data. Complementing LRI-based CCC, MEBOCOST was developed to study CCC mediated by metabolites from sender cells and sensor proteins on receiver cells [138], while miRTalk enables the inference of CCC mediated by extracellular vesicle (EV)-derived miRNA-target interaction (MITI) [139]. CCC databases present an alternative approach to constructing BN. For example, Shan et al. established CITEdb, a manually curated database of human cell-cell interactions containing 728 pairs of CCC in specific disease or tissue contexts [140]. Similarly, Gao et al. developed MACC, a visual interactive knowledgebase of metabolite-associated cell communications [141]. Specifically, MACC encompasses 244 human endogenous metabolites and 4508 metabolite-related enzymes and transporters, providing a foundation for constructing CCC networks between cells *via* metabolite secretion and adsorption. These databases and AI methods offer new perspectives for network construction and analysis in TCM network pharmacology.

4. Network analysis with AI methods

Network pharmacology offers a systematic framework for evaluating therapeutic efficacy and elucidating mechanisms of action from the perspective of network biology. In TCM research, the primary objectives of network pharmacology are to assess efficacy and clarify mechanisms of action based on the construction of IN, DN, and BN (Fig. 1B). For efficacy evaluation, two principal strategies have been proposed: network recovery and proximity evaluation . The underlying assumption is that therapeutic efficacy correlates with the network distance between perturbed nodes in the IN and DN within the broader biological molecular network, with shorter distances indicating greater efficacy. For mechanism-of-action elucidation, two complementary strategies are employed: enrichment and hub-node determination , which addresses the biological perspective, and core-ingredient refinement , which focuses on the chemical perspective. In this section, we outline the workflows associated with these primary network analysis strategies and present the AI methodologies applied in this domain (Table 1).

4.1. Network recovery evaluation

Network recovery describes the biological network transitioning from a disease state to a normal state after treatment, aligning closely with the disharmony and harmony theory in TCM. Due to the rapid expansion of high-throughput omics data, particularly transcriptomic profiling, across multiple domains of modern molecular biology, network medicine has emerged as a powerful tool for understanding the relationships between biological molecular features and efficacy. Accordingly, Wu et al. proposed a network analysis approach called NRI to quantitatively evaluate the network perturbation recovery from disease to normal state by treatment [6]. The core assumption centers on the recovery of expression of DEGs to the normal level, specifically, the upregulation of lowly expressed genes and downregulation of highly expressed genes in the disease network after treatment. By constructing the disease network of MI, Wu et al. applied NRI to evaluate the efficacy of Shenmai injection (SHENMAI), a widely used TCM composed of red ginseng (RG) and RO in clinically treating MI. Consequently, SHENMAI demonstrated a higher NRI than RG and RO, indicating the synergistic effect of RG and RO on treating MI by reversing the disease network.

Similarly, based on this strategy, Lamb et al. established the first-generation connectivity map (Cmap), comprising genome-wide RNA expression data for 164 distinct bioactive small-molecule perturbagens and corresponding vehicle controls applied to human cell lines for a short duration to connect biological molecular features and efficacy [142, 143]. Specifically, each treatment instance with a rank-ordered signature of approximately 22 000 genes according to the differential expression relative to the control serves as the reference for comparison to the query signature. The fundamental assumption is that upregulated query genes appearing near the top of the reference signature and downregulated query genes near the bottom indicate positive connectivity, while the reverse pattern indicates negative connectivity. When examining a query signature from a chemical perturbation, similar chemicals typically exhibit comparable perturbed signatures by recovering the expression of nodes in the diseased network. Therefore, this approach can evaluate the efficacy of a query ingredient by analyzing the most similar bioactive small molecule in creating the first-generation Cmap.

Notably, a comprehensive transcriptomic perturbation atlas forms the foundation for network recovery evaluation, with its quality substantially influencing the performance of AI methods. Consequently, significant efforts have been directed toward constructing such high-quality atlases. For instance, Tahoe-100M represents a giga-scale single-cell atlas containing 100 million transcriptomic profiles that systematically quantify the effects of 1100 small-molecule perturbations across 50 cancer cell lines [174]. Additionally, researchers proposed the first genome-scale perturbation atlas of morphology phenotypes in human cells, utilizing unbiased, high-dimensional image-based profiling and massively parallel optical pooled CRISPR screening [175]. Specifically, ITCM incorporated a pharmacotranscriptomic map of 496 representative active ingredients using high-throughput experiments [22]. Beyond experimental methods, Zhan et al. developed a genome-scale deep learning model called TranscriptionNet to predict gene expression changes of genetic perturbations [176].

These databases offer potential for systematic evaluation of biological network recovery towards an ingredient, herb, or TCM formula from the perspective of perturbed gene expression profiles.

4.2. Enrichment & hub node determination

A common approach to interpreting the pharmacological mechanism of action for TCM involves the integrated analysis of DN, IN, and/or BN to identify meaningful biological pathways. The conventional strategy involves performing PEA by mapping the gene/protein nodes of IN or the shared nodes between DN and IN to the BN, followed by AI algorithm analysis. In the widely adopted analytical protocol of TCM network pharmacology, researchers conduct over-representation analysis (ORA) using hypergeometric or Fisher's exact tests on shared genes/proteins between DN and IN, constructed from established TCM databases and primary literature. This analysis aims to identify significantly enriched pathways or biological processes related to TCM treatment. Alternatively, DEGs or DEPs derived from omics data can be utilized for PEA, exemplified by DAVID, WebGestalt, and Metascape [148]. However, these methods heavily depend on the number of DEGs or DEPs, which vary with statistical models, resulting in inconsistent outcomes under different parameter selections and thresholds, such as fold change and *P* value.

To address this limitation, certain methods consider changes in all expressed genes to conduct gene set enrichment analysis (GSEA) [149-151] for determining pathway activation or repression. Other approaches like GSVA [152] and PLAGE [153] employ matrix factorization to estimate pathway activity variation across samples. Furthermore, ReporterScore implemented generalized reporter score-based analysis (GRSA) for PEA across multi-group and longitudinal omics data, applicable to microbiome, transcriptome, and metabolome omics data [154]. These methods can identify significantly activated or repressed pathways following TCM treatment. Considering the essential information provided by network nodes' topological properties for identifying significantly enriched pathways, network topology-based methods have emerged that incorporate weighted nodes and/or edges to model the integrated DN, IN, and BN [155-161]. Network topology describes the arrangement or structure of various elements in the IN, DN, or BN, defining node interconnections and communications. For instance, LEGO [160] and GANPA [161] utilize network-based gene weighting, while LPIA [159] employs network-based edge weighting. These methods aim to reflect how incident pathways function to produce similar biological outcomes under specific conditions or perturbations. Notably, CTpathway [155] incorporates pathway crosstalk analysis through graph modeling on a curated global pathway crosstalk map to enhance PEA performance. Regarding metabolic networks, MiNEApy enables minimal metabolic network computation and performs enrichment analysis based on flux variability analysis [162].

Network clustering analysis is commonly employed to identify hub sub-networks or nodes, facilitating the determination of regulatory targets for specific TCM components or herbal formulas. The conventional approach utilizes network topological information to extract dense subgraphs [163-173]. For instance, MCODE [166] employs local neighborhood density to weigh network nodes and segment the network into sub-networks, while Deepwalk [169] identifies similar node clusters through random walk algorithms for node embedding vector learning. These methods enable network node clustering in low-dimensional space. However, this approach is limited by the uncertainty in identifying direct targets within large clustered sub-networks. To enhance the identification of potential regulatory targets, direct topological measures are utilized to rank and screen hub nodes. The widely used software Cytoscape [177] incorporates multiple modules to calculate common metrics, including Degree, Betweenness Centrality, and Closeness Centrality for each node to identify hub genes/proteins as potential TCM active ingredient targets. Specifically, Degree measures the number of node neighbors, Betweenness Centrality calculates the frequency of a node serving as a bridge in shortest paths between other nodes, and Closeness Centrality indicates the proximity of a node to others. Though differently defined, these network topological indexes collectively reflect node importance within the network.

4.3. Core ingredients refinement

After enrichment and hub node determination based on the DN or combined DN and IN with therapeutic effects for ingredients on the disease, potential targets can be identified to refine the core ingredients of specific herbs or formulas. These refined core ingredients subsequently inform quality control standards in TCM production. For instance, Ding et al. implemented a systematic computational approach called CIPHER to develop the Glomerular Cytoskeleton Network (GCNet), a PPI network designed to identify key cytoskeletal components in glomerular diseases [144]. Through GCNet topological metrics such as degree and betweenness centrality for prioritizing functionally critical proteins, the network-based analysis identified 21 consistently regulated cytoskeletal genes across five glomerular diseases. This computational framework demonstrates considerable potential for enhancing herbal medicine efficacy. Through PPI mapping, this methodology can identify crucial molecular targets within disease-associated pathways, enabling the identification of bioactive compounds in herbal formulations that modulate these targets. For example, the method's capacity to highlight critical nodes (e.g., MYL9, PDLIM5) may guide the selection of phytochemicals that restore dysregulated cytoskeletal dynamics in glomerular diseases. The integration of PPI networks with herbal compound databases may expedite the discovery of synergistic multi-target therapies, connecting traditional medicine with modern systems pharmacology.

Traditional methods that focus exclusively on isolated ingredients and their individual pharmacological effects overlook relative concentrations—a crucial factor in TCM's holistic efficacy. Wang et al. addressed this limitation by developing an integrative computational strategy combining network pharmacology and quantitative compositional analysis to systematically identify active ingredients, using Xuesaitong (XST) injection for myocardial infarction (MI) as a case study [145]. Initially, they constructed a cardiovascular disease (CVD)-specific PPI network using databases including STRING and KEGG, followed by topological analysis to identify hub proteins. Subsequently, potential targets of XST's ingredients were predicted through multi-modal data integration, incorporating microarray-based transcriptomics (e.g., GEO datasets), text mining (e.g., PubMed), and pharmacophore modeling (e.g., PharmMapper). These targets were mapped onto the PPI network to evaluate their regulatory roles. To prioritize bioactive compounds, they introduced a composition-weighted index (CWI), which evaluates ingredient efficacy by combining network centrality metrics with HPLC-measured ingredient concentrations. This methodology identified notoginsenoside R1, ginsenosides Rg1, Rb1, Rd, and Re as XST's key active components, subsequently validated in MI rat models. By integrating multi-omics data, network topology, and content-aware scoring, this framework provides a robust, scalable method for TCM bioactive ingredient discovery, advancing beyond reductionist approaches to capture herbal medicine's synergistic complexity. Future enhancements could incorporate machine learning-based target prioritization and dynamic pharmacokinetic modeling to improve predictive accuracy.

To validate the refined core ingredients more thoroughly,

priority should be given to those derived from herbs or TCM formulas with clinical trial verification, thereby strengthening the credibility of computational predictions. Specifically, new TCM drugs with identified core ingredients, evaluated by the Center for Drug Evaluation (CDE) of the National Medical Products Administration (NMPA), provide substantial clinical evidence for direct validation (Table 3). For instance, Lv et al. employed network pharmacology methodology to analyze *Ramulus Mori* (Sangzhi) alkaloid (RMA) in treating type 2 diabetes mellitus (T2DM) [178]. Their analysis identified 1-deoxynojirimycin, fagomine, and *N*-methyl-1-deoxynojirimycin as primary active ingredients of RMA, targeting carbohydrate metabolism and regulating alpha-glucosidase activity to produce anti-diabetic effects. Significantly, the core ingredients of Sangzhi Zongshengwujian Pian, an approved TCM new drug for T2DM treatment with clinical evidence, comprise total alkaloids (Mulberry twig alkaloids) [179], aligning with previous findings. Additionally, a recent study identified several active ingredients, including ginkgolides A, B, C, and K, in modulating ischemic stroke through network pharmacology [180]. These active ingredients target common pathological pathways associated with ischemic stroke, with ginkgolides (A, B, C, and K) emerging as prominent nodes in the network topology analysis combining IN and DN. Correspondingly, the core ingredients of Yinxing Zongneizhi Diwan, a submitted TCM new drug for ischemic stroke treatment with clinical evidence, consist of total terpene lactone, including ginkgolides A, B, and C [181, 182]. In conclusion, core ingredients predicted by TCM network pharmacology methods could inform new TCM drug development and undergo clinical trial validation. Furthermore, identified core ingredients with clinical evidence could enhance AI methods for core ingredient refinement.

4.4. Network proximity evaluation

Network proximity denotes the network distance between IN and DN within the BN context. This analysis suggests that shorter distances between targeted nodes and disease nodes indicate greater efficacy. Prof. Barabasi's group initially proposed this concept for network-based prediction of drug combinations [146]. By quantifying the network-based relationship between drug targets and disease proteins in the human protein–protein interactome, complementary exposure representing separated drugs that individually overlap with the disease module demonstrates statistically significant efficacy for combination therapies. Notably, overlapping exposure, where drug–target modules overlap with each other and the disease module, shows no significant therapeutic advantage over monotherapy but exhibits significant adverse effects. This combination approach aligns with TCM compatibility principles, as both seek to enhance efficacy while reducing adverse effects through action on different disease targets. Subsequently, Prof. Barabasi's group applied this principle to evaluate TCM treatment efficacy by examining the topological relationship between disease symptoms and TCM herb targets on the human protein interactome, revealing that network proximity of herb targets to symptom modules predicts therapeutic effectiveness [183].

Beyond node-based distance for network proximity evaluation, Wang et al. developed a method utilizing edge-based distance to assess proximity between normal and disease networks. Specifically, the distance represents the perpendicular distance between actual pairwise gene expression levels of a diseased case sample and the simple linear regression modeled with normal samples [147]. This approach compares normal and disease gene co-expression networks based on network edges between identical pairwise gene nodes. By incorporating treated sample data, this method can evaluate TCM efficacy through edge-based network proximity assessment. While this approach does not depend on omics data, it requires information similar to network recovery evaluation methods, where BN, DN, and IN correspond to omics data in normal, disease, and treated conditions, respectively.

Recent studies increasingly utilize this strategy for TCM efficacy evaluation. For example, Zong et al. implemented a modular-based network proximity approach to compare three formulae: Shirebi tablets (SRB), Yuxuebi capsules (YXB), and Wangbi-fukang granules (WBFK) [184]. Their analysis identified optimal clinical indications based on network proximity Z-scores. SRB demonstrated the strongest association with gouty arthritis, while YXB showed connections to atherosclerosis, myocardial ischemia, rheumatic fever, osteoarthritis, and gouty arthritis. WBFK exhibited efficacy for myocardial ischemia, systemic lupus erythematosus, rheumatic fever, and gouty arthritis. These findings establish precise clinical positioning for these formulae. Similarly, Zhang et al. employed network-based proximity analysis to evaluate network distance between *Schisandra chinensis* active components and diabetic skin wounds [185]. The study revealed that Gomisin A targets displayed the greatest network proximity to disease-associated nodes within the PPI network. This computational finding was corroborated by *in vivo* experiments, which confirmed the therapeutic efficacy and clarified the underlying mechanisms of Gomisin A —an active constituent of *Schisandra chinensis*—in facilitating diabetic wound healing.

Both network recovery evaluation and proximity evaluation assess efficacy based on the network distance between DN and IN, with their fundamental assumption being that greater shared nodes or closer network distance indicates enhanced efficacy. This strategy demonstrates utility in refining core ingredients, as evidenced by the aforementioned cases. Notably, recent research has shown that incorporating non-coding RNAs (ncRNAs) into the PPI network expands both gene quantity and interactions, substantially enhancing the network's capacity to identify disease modules [186]. This finding highlights the essential role of ncRNAs in strengthening network pharmacology's predictive capabilities. Recent technological advances, including snRandom-seq and STRS, have achieved enhanced coverage of ncRNAs, particularly miRNAs, enabling concurrent quantification of coding and ncRNAs at single-cell or spatial resolution [187, 188]. Future research should prioritize the development of sophisticated AI methods for constructing and analyzing heterogeneous biological networks using coding RNA and ncRNA co-sequencing data.

5. Challenges and future directions

TCM network pharmacology, an emerging interdisciplinary field, has achieved broad recognition in TCM research. This acceptance stems from its alignment with both the holistic perspective and the classical discord and reconciliation theories of TCM, which are increasingly applied to herbal medicine studies worldwide. The integration of TCM network pharmacology with AI and omics technologies marks a significant advancement in TCM research. This approach enables intelligent evaluation of therapeutic efficacy and precise elucidation of action mechanisms, establishing a promising direction for TCM modernization. Despite substantial AI-driven progress in TCM network pharmacology, pressing challenges and future directions concern the advancement of TCM network pharmacology through rapidly evolving AI technologies, such as LLMs and single-cell and spatial omics technologies [189-191], particularly in ingredient identification, network construction, and network analysis (Fig. 3).

Ingredient identification constitutes the fundamental step in TCM network pharmacology research. The current approach combines nuclear magnetic resonance (NMR) spectroscopy with LC-MS/MS data to characterize chemical components qualitatively and quantitatively in herbal medicines or formulas. Traditional manual identification and annotation of chemical mo-

Table 3 Submitted new TCM drugs with identified core ingredients in the past decade (2016-2025) from the CDE of NMPA.

Name	Core ingredients	Related disease	Related herb	Accession
Zhishi Zonghuangtong Pian*	Total flavonoid	Dyspepsia	Immature fruit of Seville orange	CXZS2200008
Guangjinqiancao Zonghuangtong Jiaonang*	Total flavonoid	Ureteral calculi	Snowbellleaf Tickclover	CXZS2101003
Yinyanghuosu Ruanjiaonang*	Icaritin	Hepatocellular carcinoma	Epimedium	CXZS2101001
Huangshukuihua Zonghuangtong Kouqiang Tiepian*	Total flavonoid	Mouth ulcer	Setose Abelmoschus	CXZS1700004
Sangzhi Zongshengwujian Pian*	Total alkaloid	Type 2 diabetes	Mulberry Leaf	CXZS1700008
Yinxing Zongneizhi Diwan	Total terpene lactone	Ischemic stroke	Ginkgo	CXZS2500018
SOC201 Jiaonang	-	-	-	CXZL2500017
Meilianfuxinye	-	-	-	CXZL2500012
Baixintiquwu Keli	-	Acute upper respiratory tract infection	Chinese Bastardtoadflax	CXZL2400097
Anemoside B4 Shuan	Anemoside B4	Ulcerative colitis	Chinese Pulsatilla	CXZL2400092
LBZ-18 Koufuru	-	-	-	CXZL2400071
Kudingcha Zaogan Pian	Total saponin	Hyperlipidemia	Plumleaf Cratoxylum	CXZL2400068
ZY13 Pian	-	Macular degeneration	-	CXZL2400042
CRA Pian	-	Chronic heart failure	-	CXZL2400030
JKN2303 Xirurongye	-	-	-	CXZL2400028
Zhusheyong Qiangjihonghuahuangsesu A	Hydroxysafflor yellow A	Acute ischemic stroke	Safflower	CXZS2300027
Yangchangchuntengye KoufuYe	-	Bronchitis	English Ivy	CXZL2300058
Sheganzhike Jiaonang	-	-	-	CXZL2300046
SZ1108 Pian	-	Diabetic nephropathy	-	CXZL2300043
Maqianzijian Ningjiaotiegao	Strychnine	Osteoarthritis	Nux Vomica	CXZL2300020
Z018B Jiaonang	-	Nonalcoholic steatohepatitis	-	CXZL2300019
Renshen Zongcigan Pian	-	Macular edema in diabetes	Ginseng	CXZL2300006
Lizhihe Jiaonang	-	-	Litchi Seed	CXZL2200039
Gancan Zongchaertong Jiaonang	Total chalcone	-	Root of Ural Licorice	CXZL2200021
Xiaozhongshengji Koufuye	-	Radiotherapy-induced oral mucositis	-	CXZL2200012
Kaituopu Pian	Catalpol	-	-	CXZL2200007
Yanzao Duotang Jiaonang	Fucoidan	-	Brown seaweed	CXZL2200002
Buguzhi Zonggan Jiaonang	Total psoralenoside	Osteoporosis	Malaytea Scurfpea	CXZL2101037
Zhusheyong Danfensuan A	Salvianolic acid A	Myocardial infarction	Root of Ligulilobe sage	CXZL2101034
Yinxing Neizhi Diwan	Ginkgo diterpene lactones	Cerebral infarction	Ginkgo	CXZL2000014
Congrongruntong Koufuye	Total sugar alcohol	Constipation	Desertliving Cistanche	CXZL1900010
Renshen Zongcigan Kouqiang Bengjiepian	-	Stableanginapectoris	Ginseng	CXZL1800001
Maqianzi Zongjian Nangpao Ningjiao	Total strychnine	Osteoarthritis	Nux Vomica	CXZL1700016
Zhimuzaogan BII Jiaonang	Timosaponin BII	Vascular dementia	Rhizome of Common Amarrhe	CXZL1700057
Zexie Jiangzhi Jiaonang	Alisol	Hyperlipidemia	Oriental Waterplantain	CXZL1600022

The asterisk (*) represents the approved new TCM drugs. -, not available.

lecules, relying solely on experimental methods, present considerable challenges: they are time-intensive, lack scalability, and face reproducibility issues. Recent developments have addressed these limitations by integrating AI technologies with experimental data for automated ingredient identification. This advancement has been facilitated by increasing the availability of large-scale experimental datasets for known chemical compounds. A significant example is MassKG, developed by Zhu et al., which combines knowledge-based fragmentation strategies with deep learning-based molecular generation models [192]. This method provides novel perspectives for ingredient identification in TCM network pharmacology. The emergence of LLMs offers additional opportunities for advancement. Their exceptional in-context learning capabilities could significantly enhance ingredient identification by effectively utilizing chemical structures' contextual information. The integration of AI methodologies promises to revolutionize the efficiency and accuracy of ingredient identification in TCM network pharmacology research.

In TCM network pharmacology, accurate construction and modeling of biological networks are essential for understanding the complex mechanisms of herbal compounds, multi-target interactions, and holistic therapeutic effects. While conventional approaches—including PPI networks and heterogeneous molecular networks incorporating ncRNAs and proteins—have shown significant progress over recent decades, these models frequently overlook cellular heterogeneity. This limitation is particularly significant given that PPI networks can vary across cell types, disease states, and individuals, as demonstrated by previous research. Advancing TCM network pharmacology requires developing more precise biological network models that accurately reflect the molecular characteristics of specific cell types, diseases, and individuals.

Recent advances in single-cell and spatial omics technologies have enabled detailed characterization of cellular and molecular events in diseased individuals [189, 190]. Mounting evidence demonstrates the presence of distinctly different cell types and complex CCC networks between pairwise cell types through various signal transmissions within tissue microenvironments. CCC, also termed cell-cell interaction, represents a fundamental characteristic of multicellular organisms, serving a vital role in biological processes [193]. The dysregulation and restoration of CCC networks correlate with disease onset and recovery, reflecting TCM concepts of bodily discord and reconciliation. Given CCC networks' capacity to capture precise molecular features in individual- and disease-specific contexts, the construction and analysis of CCC networks using single-cell and spatial omics data present both opportunities and challenges for TCM network pharmacology.

In CCC network construction, a primary challenge involves accurately inferring CCC between paired cell types. CCC operates through three primary modes: LRI-based, metabolite-based, and EV-based mechanisms. While current computational methods primarily emphasize LRIs, AI-driven approaches for metabolite- and EV-based CCC inference remain limited. Addressing this limitation, Cui et al. developed MEBOCOST, a tool for studying metabolite-based CCC through sender-cell metabolites and receiver-cell sensor proteins. Their research demonstrated that macrophage Epsins significantly influence lipid metabolism in atherosclerosis through CCC [194]. Additionally, Shao et al. introduced miRTalk, an innovative method for inferring CCC mediated by EV-derived MITI, revealing detailed MITI-mediated CCC mechanisms in various disease contexts [139]. Despite these developments, subsequent research must utilize advanced AI technologies to enhance the precision of metabolite- and EV-based CCC inference. Moreover, the integration of multi-modal CCC networks remains crucial for understanding bodily disharmony and reconciliation mechanisms from an integrated TCM and Western medicine perspective.

In CCC network analysis, the primary challenge involves accurately identifying key CCC network targets, including essential CCC sub-networks, cell types, and signaling molecules such as secreted ligands, metabolites, and EVs from sender cells, along with their corresponding targets in receiver cells. Recent advances in single-cell and spatial omics technologies facilitate high-resolution classification of numerous cells while accurately quantifying molecular profiles at single-cell resolution. These developments provide unprecedented technical support for investigating TCM targets at the cellular and molecular levels. For example, Li et al. developed scRank, an AI-driven method that determines drug-responsive cell types by modeling in silico perturbations using untreated single-cell RNA-seq data through target-perturbed gene regulatory networks [195]. This method demonstrated how tanshinone IIA, a bioactive compound from Danshen, achieves anti-MI effects by targeting specific pro-inflammatory macrophage subpopulations. As single-cell and spatial multi-omics technologies progress, the integration of multi-modal omics data for CCC network target inference increasingly relies on AI methods such as transfer learning [196]. These methods have demonstrated significant success in integrating multi-view biological networks, providing robust tools for systematic and precise exploration of CCC mechanisms in TCM network pharmacology.

In experimental validation of CCC network targets, the primary challenge centers on accurately identifying target cell types and corresponding molecules for specific ingredients, herbs, or TCM formulas. Organoid technology has recently emerged as a transformative approach for understanding CCC mechanisms in specific contexts [197]. This system enables researchers to examine how autocrine, paracrine, juxtacrine, and EV-based signaling regulate tissue homeostasis, disease progression, and therapeutic responses. Notably, organoids maintain native human genetic and epigenetic features, enabling more accurate disease modeling and real-time manipulation (e.g., genetic perturbations, compound treatments) to validate CCC mechanisms, as demonstrated in recent studies [198-200]. However, TCM syndromes (e.g., "Kidney Yang Deficiency" or "Liver Qi Stagnation") represent system-level imbalances that prove particularly challenging to model using current single-tissue organoid systems, while key syndrome indicators like pulse quality, tongue coating, or subjective symptoms remain unmeasurable *in vitro*. Future developments require breakthroughs in integrated multi-tissue organoid systems that mimic whole-body "visceral mani-

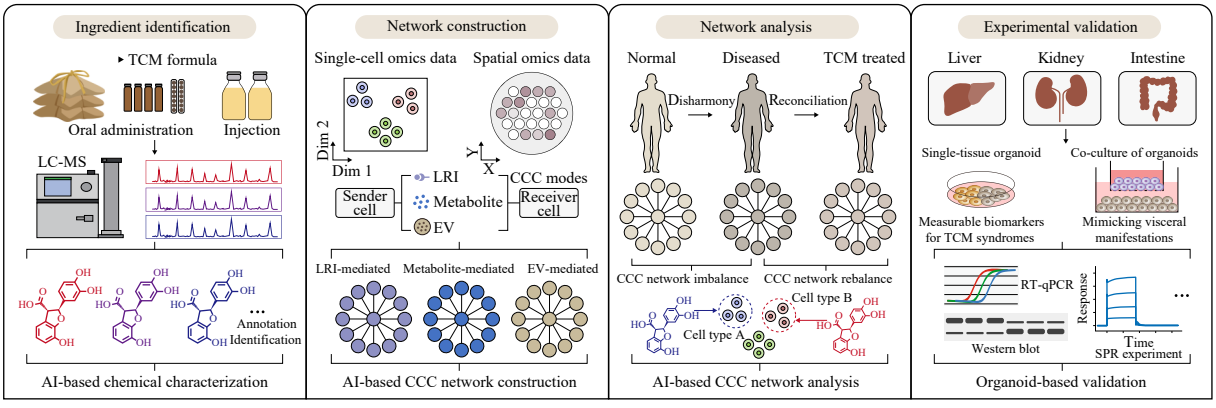

Fig. 3 Challenges and future directions. For ingredient identification, the incorporation of AI methodologies such as LLMs promises to transform the efficiency and accuracy of ingredient intelligent annotation and identification. For the network construction, the AI-based approaches for metabolite- and EV-based CCC inference are urgent except for the LRI-mediated CCC inference. For the network analysis, the key challenge is the accurate identification of the key CCC network targets, including key CCC sub-networks, cell types, and signaling molecules such as secreted ligands, metabolites, and EVs derived from sender cells, as well as their corresponding targets in receiver cells. For the experimental validation, future direction lies in the breakthrough of integrated multi-tissue organoid systems mimicking whole-body "visceral manifestations" and measurable biomarkers for TCM syndromes.

festations" and measurable biomarkers for TCM syndromes.

The findable, accessible, interoperable, and reusable (FAIR) principles and data standardization are fundamental to advancing modern research, particularly in data-driven fields such as bioinformatics, network pharmacology, and systems biology. The FAIR principles provide a conceptual framework for optimizing data usability, and data standardization ensures practical implementation by establishing consistent formats, protocols, and metadata schemas. While some databases and AI methods involving TCM network pharmacology are not currently available, most align with FAIR data principles (Table 1 in our revised manuscript). Regarding data standardization for databases and AI methods related to TCM network pharmacology, its implementation faces significant challenges. One of the primary challenges is the diversity of data formats, structures, and processing protocols across different databases, technologies, AI methods, etc. For instance, the bulk RNA-seq data is a dense matrix with limited samples. ScRNA-seq data are typically represented as sparse matrices and often require imputation of gene expression values, resulting in diverse and sometimes incompatible standardization protocols [201]. The lack of universally accepted standards has further encouraged the use of proprietary or domain-specific solutions, which hampers data interoperability. Under such circumstances, researchers frequently prioritize convenience over compliance, leading to inconsistent practices —an issue particularly problematic for AI-based analyses. This is partly because certain AI methods require customized or task-specific standardization to achieve optimal predictive performance in TCM network pharmacology. Beyond adherence to the FAIR data principles, future work should emphasize the development of robust computational and experimental standardization frameworks, capitalizing on the rapid progress and reproducibility of AI methodologies to ensure greater consistency, interoperability, and reliability in data use.

In summary, the integration of AI-driven computational methods with high-resolution and high-throughput single-cell or spatial omics technologies establishes a transformative research paradigm for advancing ingredient identification, network construction, and network analysis in TCM network pharmacology. The field requires the development of TCM-specific AI methodologies and validation approaches that incorporate traditional medical knowledge and theoretical frameworks. These specialized approaches will enhance the systematic evaluation of therapeutic efficacy and elucidate the mechanisms of action and active ingredients, also known as functional components of TCM. The convergence of advanced computational and experimental technologies with TCM principles creates a bridge between modern systems biology and traditional medicine, providing significant opportunities for advancing TCM network pharmacology research and modernization.

Funding

This work was supported by the "Pioneer" and "Leading Goose" R&D Program of Zhejiang (No. 2024C03106, X.F.), the National Natural Science Foundation of China (No. 82474160, X.S.), the Joint Funds of the Zhejiang Provincial Natural Science Foundation of China (No. LBZ24H270001, X.P.), the Major Joint Projects Supported by the National Administration of TCM and Zhejiang Province (No. GZY-ZI-KJ-23037, X.P.), and the Ningbo Top Medical and Health Research Program (No. 2022030309, X.P.).

Declaration of competing interest

These authors have no conflict of interest to declare.

References

1 Hopkins AL. Network pharmacology. *Nat Biotechnol*. 2007;25(10):1110-1111. https://doi.org/10.1038/nbt1007-1110.
2 Hopkins AL. Network pharmacology: the next paradigm in drug discovery. *Nat Chem Biol*. 2008;4(11):682-690. https://doi.org/10.1038/nchembio.118.
3 Li S, Zhang B. Traditional Chinese medicine network pharmacology: theory, methodology and application. *Chin J Nat Med*. 2013;11(2):110-120. https://doi.org/10.1016/S1875-5364(13)60037-0.
4 Li X, Liu Z, Liao J, et al. Network pharmacology approaches for research of traditional Chinese medicines. *Chin J Nat Med*. 2023;21(5):323-332. https://doi.org/10.1016/S1875-5364(23)60429-7.
5 Zhang P, Zhang D, Zhou W, et al. Network pharmacology: towards the artificial intelligence-based precision traditional Chinese medicine. *Brief Bioinform*. 2023;25(1):bbad518. https://doi.org/10.1093/bib/bbad518.
6 Wu L, Wang Y, Li Z, et al. Identifying roles of "Jun-Chen-Zuo-Shi"component herbs of QiShenYiQi Formula in treating acute myocardial ischemia by network pharmacology. *Chin Med*. 2014;9:24. https://doi.org/10.1186/1749-8546-9-24.
7 Liao J, Hao C, Huang W, et al. Network pharmacology study reveals energy metabolism and apoptosis pathways-mediated cardioprotective effects of Shenqi Fuzheng. *J Ethnopharmacol*. 2018;227:155-165. https://doi.org/10.1016/j.jep.2018.08.029.
8 Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. *Nat Biomed Eng*. 2018;2(10):719-731. https://doi.org/10.1038/s41551-018-0305-z.
9 Shao X, Yang H, Zhuang X, et al. scDeepSort: a pre-trained cell-type annotation method for single-cell transcriptomics using deep learning with a weighted graph neural network. *Nucleic Acids Res*. 2021;49(21):e122. https://doi.org/10.1093/nar/gkab775.
10 Fang Y, Zhang Q, Zhang N, et al. Knowledge graph-enhanced molecular contrastive learning with functional prompt. *Nat Mach Intell*. 2023;5(5):542-553. https://doi.org/10.1038/s42256-023-00654-0.
11 Guo D, Yang D, Zhang H, et al. Deepseek-r1: incentivizing reasoning capability in llms via reinforcement learning. *arXiv*. 2025;2501:12948. https://doi.org/10.48550/arXiv.2501.12948.
12 Brown T, Mann B, Ryder N, et al. Language models are few-shot learners. *Adv Neural Inf Process Syst*. 2020;33:1877-1901. https://doi.org/10.48550/arXiv.2005.14165.
13 Touvron H, Lavril T, Izacard G, et al. Llama: open and efficient foundation language models. *arXiv*. 2023;2302:13971. https://doi.org/10.48550/arXiv.2302.13971.
14 Dai Y, Shao X, Zhang J, et al. TCMChat: a generative large language model for traditional Chinese medicine. *Pharmacol Res*. 2024;210:107530. https://doi.org/10.1016/j.phrs.2024.107530.
15 Chen Y, Wang Z, Xing X, et al. Bianque: balancing the questioning and suggestion ability of health llms with multi-turn health conversations polished by chatgpt. *arXiv*. 2023;2310:15896. https://doi.org/10.48550/arXiv.2310.15896.
16 Zhang H, Chen J, Jiang F, et al. Huatuogpt, towards taming language model to be a doctor. *arXiv*. 2023;2305:15075. https://doi.org/10.48550/arXiv.2305.15075.
17 Chen S, Yin X, Han J, et al. DNA barcoding in herbal medicine: retrospective and prospective. *J Pharm Anal*. 2023;13(5):431-441. https://doi.org/10.1016/j.jpha.2023.03.008.
18 Ru J, Li P, Wang J, et al. TCMSP: a database of systems pharmacology for drug discovery from herbal medicines. *J Cheminform*. 2014;6:13. https://doi.org/10.1186/1758-2946-6-13.
19 Yan D, Zheng G, Wang C, et al. HIT 2.0: an enhanced platform for herbal ingredients' targets . *Nucleic Acids Res*. 2022;50(D1):D1238-D1243. https://doi.org/10.1093/nar/gkab1011.
20 Wu Y, Zhang F, Yang K, et al. SymMap: an integrative database of traditional Chinese medicine enhanced by symptom mapping. *Nucleic Acids Res*. 2019;47(D1):D1110-D1117. https://doi.org/10.1093/nar/gky1021.
21 Fang S, Dong L, Liu L, et al. HERB: a high-throughput experiment- and reference-guided database of traditional Chinese medicine. *Nucleic Acids Res*. 2021;49(D1):D1197-D1206. https://doi.org/10.1093/nar/gkaa1063.
22 Tian S, Zhang J, Yuan S, et al. Exploring pharmacological active ingredients of traditional Chinese medicine by pharmacotranscriptomic map in ITCM. *Brief Bioinform*. 2023;24(2):bbad027. https://doi.org/10.1093/bib/bbad027.
23 Zhang Y, Li X, Shi Y, et al. ETCM v2.0: an update with comprehensive resource and rich annotations for traditional Chinese medicine . *Acta Pharm Sin B*. 2023;13(6):2559-2571. https://doi.org/10.1016/j.apsb.2023.03.012.
24 Lv Q, Chen G, He H, et al. TCMBank-the largest TCM database provides deep learning-based Chinese-Western medicine exclusion prediction. *Signal Transduct Target Ther*. 2023;8(1):127. https://doi.org/10.1038/s41392-023-01339-1.
25 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. *Pharmacol Res*. 2022;178:106185. https://doi.org/10.1016/j.phrs.2022.106185.
26 Liu X, Liu J, Fu B, et al. DCABM-TCM: a database of constituents absorbed into the blood and metabolites of traditional Chinese medicine. *J Chem Inf Model*. 2023;63(15):4948-4959. https://doi.org/10.1021/acs.jcim.3c00365.
27 Yang P, Lang J, Li H, et al. TCM-Suite: a comprehensive and holistic platform for traditional Chinese medicine component identification and network pharmacology analysis. *iMeta*. 2022;1(4):e47. https://doi.org/10.1002/imt2.47.
28 Pinzi L, Rastelli G. Molecular docking: shifting paradigms in drug discovery. *Int J Mol Sci*. 2019;20(18):4331. https://doi.org/10.3390/ijms20184331.
29 Allen WJ, Balius TE, Mukherjee S, et al. DOCK 6: impact of new features and current docking performance. *J Comput Chem*. 2015;36(15):1132-1156. https://doi.org/10.1002/jcc.23905.
30 McGann M. FRED and HYBRID docking performance on standardized

datasets. *J Comput Aided Mol Des*. 2012;26(8):897-906. https://doi.org/10.1007/s10822-012-9584-8.
31 Spitzer R, Jain AN. Surflex-Dock: docking benchmarks and real-world application. *J Comput Aided Mol Des*. 2012;26(6):687-699. https://doi.org/10.1007/s10822-011-9533-y.
32 Verdonk ML, Cole JC, Hartshorn MJ, et al. Improved protein-ligand docking using GOLD. *Proteins*. 2003;52(4):609-623. https://doi.org/10.1002/prot.10465.
33 Morris GM, Huey R, Lindstrom W, et al. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J Comput Chem*. 2009;30(16):2785-2791. https://doi.org/10.1002/jcc.21256.
34 Zsoldos Z, Reid D, Simon A, et al. eHiTS: an innovative approach to the docking and scoring function problems. *Curr Protein Pept Sci*. 2006;7(5):421-435. https://doi.org/10.2174/138920306778559412.
35 Friesner RA, Banks JL, Murphy RB, et al. Glide: a new approach for rapid, accurate docking and scoring.1. Method and assessment of docking accuracy. *J Med Chem*. 2004;47(7):1739-1749. https://doi.org/10.1021/jm0306430.
36 Pang YP, Perola E, Xu K, et al. EUDOC: a computer program for identification of drug interaction sites in macromolecules and drug leads from chemical databases. *J Comput Chem*. 2001;22(15):1750-1771. https://doi.org/10.1002/jcc.1129.
37 Rarey M, Kramer B, Lengauer T, et al. A fast flexible docking method using an incremental construction algorithm. *J Mol Biol*. 1996;261(3):470-489. https://doi.org/10.1006/jmbi.1996.0477.
38 Miller MD, Kearsley SK, Underwood DJ, et al. FLOG: a system to select 'quasi-flexible' ligands complementary to a receptor of known three-dimensional structure. *J Comput Aided Mol Des*. 1994;8(2):153-174. https://doi.org/10.1007/BF00119865.
39 Zavodszky MI, Rohatgi A, Van Voorst JR, et al. Scoring ligand similarity in structure-based virtual screening. *J Mol Recognit*. 2009;22(4):280-292. https://doi.org/10.1002/jmr.942.
40 Wu G, Robertson DH, Brooks CL, et al. Detailed analysis of grid-based molecular docking: a case study of CDOCKER-A CHARMm-based MD docking algorithm. *J Comput Chem*. 2003;24(13):1549-1562. https://doi.org/10.1002/jcc.10306.
41 Tietze S, Apostolakis J. GlamDock: development and validation of a new docking tool on several thousand protein-ligand complexes. *J Chem Inf Model*. 2007;47(4):1657-1672. https://doi.org/10.1021/ci7001236.
42 Korb O, Stutzle T, Exner TE. Empirical scoring functions for advanced protein-ligand docking with PLANTS. *J Chem Inf Model*. 2009;49(1):84-96. https://doi.org/10.1021/ci800298z.
43 Thomsen R, Christensen MH. MolDock: a new technique for high-accuracy molecular docking. *J Med Chem*. 2006;49(11):3315-3321. https://doi.org/10.1021/jm051197e.
44 Grosdidier A, Zoete V, Michielin O. EADock: docking of small molecules into protein active sites with a multiobjective evolutionary optimization. *Proteins*. 2007;67(4):1010-1025. https://doi.org/10.1002/prot.21367.
45 Zhang X, Zhang O, Shen C, et al. Efficient and accurate large library ligand docking with KarmaDock. *Nat Comput Sci*. 2023;3(9):789-804. https://doi.org/10.1038/s43588-023-00511-5.
46 Cai H, Shen C, Jian T, et al. CarsiDock: a deep learning paradigm for accurate protein-ligand docking and screening based on large-scale pre-training. *Chem Sci*. 2024;15(4):1449-1471. https://doi.org/10.1039/d3sc05552c.
47 Xue M, Liu B, Cao S, et al. FeatureDock for protein-ligand docking guided by physicochemical feature-based local environment learning using transformer. *npj Drug Discov*. 2025;2(1):4. https://doi.org/10.1038/s44386-025-00005-6.
48 Corso G, Stärk H, Jing B, et al. Diffdock: diffusion steps, twists, and turns for molecular docking. In: International Conference on Learning Representations. 2023;2210.01776. https://doi.org/10.48550/arXiv.2210.01776.
49 Lu W, Zhang J, Huang W, et al. DynamicBind: predicting ligand-specific protein-ligand complex structure with a deep equivariant generative model. *Nat Commun*. 2024;15(1):1071. https://doi.org/10.1038/s41467-024-45461-2.
50 Cao D, Chen M, Zhang R, et al. SurfDock is a surface-informed diffusion generative model for reliable and accurate protein-ligand complex prediction. *Nat Methods*. 2025;22(2):310-322. https://doi.org/10.1038/s41592-024-02516-y.
51 Case DA, Cheatham TE, Darden T, et al. The Amber biomolecular simulation programs. *J Comput Chem*. 2005;26(16):1668-1688. https://doi.org/10.1002/jcc.20290.
52 Brooks BR, Brooks CL, Mackerell AD, et al. CHARMM: the biomolecular simulation program. *J Comput Chem*. 2009;30(10):1545-1614. https://doi.org/10.1002/jcc.21287.
53 Van Der Spoel D, Lindahl E, Hess B, et al. GROMACS: fast, flexible, and free. *J Comput Chem*. 2005;26(16):1701-1718. https://doi.org/10.1002/jcc.20291.
54 Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550. https://doi.org/10.1186/s13059-014-0550-8.
55 Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26(1):139-140. https://doi.org/10.1093/bioinformatics/btp616.
56 Leng N, Dawson JA, Thomson JA, et al. EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments. *Bioinformatics*. 2013;29(8):1035-1043. https://doi.org/10.1093/bioinformatics/btt087.
57 Frazee AC, Pertea G, Jaffe AE, et al. Ballgown bridges the gap between transcriptome assembly and expression analysis. *Nat Biotechnol*. 2015;33(3):243-246. https://doi.org/10.1038/nbt.3172.
58 Tarazona S, García F, Ferrer A, et al. NOIseq: a RNA-seq differential expression method robust for sequencing depth biases. *EMBnet J*. 2011;17(B):18-19. https://doi.org/10.14806/ej.17.B.265.
59 Yu JL, Dai QQ, Li GB. Deep learning in target prediction and drug repositioning: recent advances and challenges. *Drug Discov Today*. 2022;27(7):1796-1814. https://doi.org/10.1016/j.drudis.2021.10.010.
60 Zhang H, Liu X, Cheng W, et al. Prediction of drug-target binding affinity based on deep learning models. *Comput Biol Med*. 2024;174:108435. https://doi.org/10.1016/j.compbiomed.2024.108435.
61 Shi JY, Yiu SM. SRP: a concise non-parametric similarity-rank-based model for predicting drug-target interactions. In: International Conference on Bioinformatics and Biomedicine (BIBM). IEEE. 2015;1636-1641. https://doi.org/10.1109/BIBM.2015.7359921.
62 Xia Z, Wu LY, Zhou X, et al. Semi-supervised drug-protein interaction prediction from heterogeneous biological spaces. *BMC Syst Biol*. 2010;4 Suppl 2(Suppl 2):S6. https://doi.org/10.1186/1752-0509-4-S2-S6.
63 Mei JP, Kwoh CK, Yang P, et al. Drug-target interaction prediction by learning from local information and neighbors. *Bioinformatics*. 2013;29(2):238-245. https://doi.org/10.1093/bioinformatics/bts670.
64 Cheng F, Liu C, Jiang J, et al. Prediction of drug-target interactions and drug repositioning *via* network-based inference. *PLoS Comput Biol*. 2012;8(5):e1002503. https://doi.org/10.1371/journal.pcbi.1002503.
65 Chen X, Liu MX, Yan GY. Drug-target interaction prediction by random walk on the heterogeneous network. *Mol Biosyst*. 2012;8(7):1970-1978. https://doi.org/10.1039/c2mb00002d.
66 Fakhraei S, Huang B, Raschid L, et al. Network-based drug-target interaction prediction with probabilistic soft logic. *IEEE/ACM Trans Comput Biol Bioinform*. 2014;11(5):775-787. https://doi.org/10.1109/TCBB.2014.2325031.
67 Ba-Alawi W, Soufan O, Essack M, et al. DASPfind: new efficient method to predict drug-target interactions. *J Cheminform*. 2016;8:15. https://doi.org/10.1186/s13321-016-0128-4.
68 Gonen M. Predicting drug-target interactions from chemical and genomic kernels using Bayesian matrix factorization. *Bioinformatics*. 2012;28(18):2304-2310. https://doi.org/10.1093/bioinformatics/bts360.
69 Cobanoglu MC, Liu C, Hu F, et al. Predicting drug-target interactions using probabilistic matrix factorization. *J Chem Inf Model*. 2013;53(12):3399-3409. https://doi.org/10.1021/ci400219z.
70 Ling C, Zeng T, Dang Q, et al. Predicting drug-target interactions using matrix factorization with self-paced learning and dual similarity information. *Technol Health Care*. 2024;32(S1):49-64. https://doi.org/10.3233/THC-248005.
71 Ezzat A, Zhao P, Wu M, et al. Drug-target interaction prediction with graph regularized matrix factorization. *IEEE/ACM Trans Comput Biol Bioinform*. 2017;14(3):646-656. https://doi.org/10.1109/TCBB.2016.2530062.
72 Liu Y, Wu M, Miao C, et al. Neighborhood regularized logistic matrix factorization for drug-target interaction prediction. *PLoS Comput Biol*. 2016;12(2):e1004760. https://doi.org/10.1371/journal.pcbi.1004760.
73 Hao M, Bryant SH, Wang Y. Predicting drug-target interactions by dual-network integrated logistic matrix factorization. *Sci Rep*. 2017;7:40376. https://doi.org/10.1038/srep40376.
74 Xiao X, Min JL, Wang P, et al. iGPCR-drug: a web server for predicting interaction between GPCRs and drugs in cellular networking. *PLoS One*. 2013;8(8):e72234. https://doi.org/10.1371/journal.pone.0072234.
75 Ezzat A, Wu M, Li XL, et al. Drug-target interaction prediction using ensemble learning and dimensionality reduction. *Methods*. 2017;129:81-88. https://doi.org/10.1016/j.ymeth.2017.05.016.
76 Perlman L, Gottlieb A, Atias N, et al. Combining drug and gene similarity measures for drug-target elucidation. *J Comput Biol*. 2011;18(2):133-145. https://doi.org/10.1089/cmb.2010.0213.
77 Wang L, You ZH, Chen X, et al. RFDT: a rotation forest-based predictor for predicting drug-target interactions using drug structure and protein sequence information. *Curr Protein Pept Sci*. 2018;19(5):445-454. https://doi.org/10.2174/1389203718666161114111656.
78 Meng FR, You ZH, Chen X, et al. Prediction of drug-target interaction networks from the integration of protein sequences and drug chemical structures. *Molecules*. 2017;22(7):1119. https://doi.org/10.3390/molecules22071119.
79 Huang YA, You ZH, Chen X. A systematic prediction of drug-target interactions using molecular fingerprints and protein sequences. *Curr Protein Pept Sci*. 2018;19(5):468-478. https://doi.org/10.2174/1389203718666161122103057.
80 Zhang S, Yang K, Liu Z, et al. DrugAI: a multi-view deep learning model for predicting drug-target activating/inhibiting mechanisms. *Brief Bioinform*. 2023;24(1):bbac526. https://doi.org/10.1093/bib/bbac526.
81 Yuan W, Chen G, Chen CY. FusionDTA: attention-based feature polymerizer and knowledge distillation for drug-target binding affinity prediction. *Brief Bioinform*. 2022;23(1):bbab506. https://doi.org/10.1093/bib/bbab506.
82 Liu S, Wang Y, Deng Y, et al. Improved drug-target interaction prediction with intermolecular graph transformer. *Brief Bioinform*. 2022;23(5):bbac162. https://doi.org/10.1093/bib/bbac162.
83 Huang K, Xiao C, Glass LM, et al. MolTrans: molecular interaction transformer for drug-target interaction prediction. *Bioinformatics*. 2021;37(6):830-836. https://doi.org/10.1093/bioinformatics/btaa880.
84 Zhang P, Wei Z, Che C, et al. DeepMGT-DTI: transformer network incorporating multilayer graph information for drug-target interaction prediction. *Comput Biol Med*. 2022;142:105214. https://doi.org/10.1016/j.compbiomed.2022.105214.
85 Zhang R, Wang Z, Wang X, et al. MHTAN-DTI: metapath-based hierarchical transformer and attention network for drug-target interaction prediction. *Brief Bioinform*. 2023;24(2):bbad079. https://doi.org/10.1093/bib/bbad079.
86 Gao M, Zhang D, Chen Y, et al. GraphormerDTI: a graph transformer-based

approach for drug-target interaction prediction. *Comput Biol Med*. 2024;173:108339. https://doi.org/10.1016/j.compbiomed.2024.108339.
87 Hu J, Yu W, Pang C, et al. DrugormerDTI: drug graphormer for drug-target interaction prediction. *Comput Biol Med*. 2023;161:106946. https://doi.org/10.1016/j.compbiomed.2023.106946.
88 Monteiro NRC, Oliveira JL, Arrais JP. DTITR: end-to-end drug-target binding affinity prediction with transformers. *Comput Biol Med*. 2022;147:105772. https://doi.org/10.1016/j.compbiomed.2022.105772.
89 Wu H, Liu J, Jiang T, et al. AttentionMGT-DTA: a multi-modal drug-target affinity prediction using graph transformer and attention mechanism. *Neural Netw*. 2024;169:623-636. https://doi.org/10.1016/j.neunet.2023.11.018.
90 Rao J, Xie J, Yuan Q, et al. A variational expectation-maximization framework for balanced multi-scale learning of protein and drug interactions. *Nat Commun*. 2024;15(1):4476. https://doi.org/10.1038/s41467-024-48801-4.
91 Ahmed KT, Ansari MI, Zhang W. DTI-LM: language model powered drug-target interaction prediction. *Bioinformatics*. 2024;40(9):btae533. https://doi.org/10.1093/bioinformatics/btae533.
92 Lu Z, Song G, Zhu H, et al. DTIAM: a unified framework for predicting drug-target interactions, binding affinities and drug mechanisms. *Nat Commun*. 2025;16(1):2548. https://doi.org/10.1038/s41467-025-57828-0.
93 He H, Chen G, Tang Z, et al. Dual modality feature fused neural network integrating binding site information for drug target affinity prediction. *npj Digit Med*. 2025;8(1):67. https://doi.org/10.1038/s41746-025-01464-x.
94 Liu S, Liu Y, Xu H, et al. SP-DTI: subpocket-informed transformer for drug-target interaction prediction. *Bioinformatics*. 2025;41(3):btaf011. https://doi.org/10.1093/bioinformatics/btaf011.
95 Amberger JS, Bocchini CA, Schiettecatte F, et al. OMIM.org: online mendelian inheritance in man (OMIM(R)), an online catalog of human genes and genetic disorders. *Nucleic Acids Res*. 2015;43(D1):D789-D798. https://doi.org/10.1093/nar/gku1205.
96 Rappaport N, Twik M, Plaschkes I, et al. MalaCards: an amalgamated human disease compendium with diverse clinical and genetic annotation and structured search. *Nucleic Acids Res*. 2017;45(D1):D877-D887. https://doi.org/10.1093/nar/gkw1012.
97 Pinero J, Bravo A, Queralt-Rosinach N, et al. DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants. *Nucleic Acids Res*. 2017;45(D1):D833-D839. https://doi.org/10.1093/nar/gkw943.
98 Zhou Y, Zhang Y, Zhao D, et al. TTD: therapeutic target database describing target druggability information. *Nucleic Acids Res*. 2024;52(D1):D1465-D1477. https://doi.org/10.1093/nar/gkad751.
99 Davis AP, Wiegers TC, Johnson RJ, et al. Comparative toxicogenomics database (CTD): update 2023. *Nucleic Acids Res*. 2023;51(D1):D1257-D1262. https://doi.org/10.1093/nar/gkac833.
100 Wu L, Li X, Yang J, et al. CHD@ZJU: a knowledgebase providing network-based research platform on coronary heart disease. *Database*. 2013;2013:bat047. https://doi.org/10.1093/database/bat047.
101 Wishart DS, Bartok B, Oler E, et al. MarkerDB: an online database of molecular biomarkers. *Nucleic Acids Res*. 2021;49(D1):D1259-D1267. https://doi.org/10.1093/nar/gkaa1067.
102 Zheng X, Tian Z, Che X, et al. DMRdb: a disease-centric Mendelian randomization database for systematically assessing causal relationships of diseases with genes, proteins, CpG sites, metabolites and other diseases. *Nucleic Acids Res*. 2025;53(D1):D1363-D1371. https://doi.org/10.1093/nar/gkae853.
103 Deng YT, You J, He Y, et al. Atlas of the plasma proteome in health and disease in 53,026 adults. *Cell*. 2025;188(1):253-271.E7. https://doi.org/10.1016/j.cell.2024.10.045.
104 Del Toro N, Shrivastava A, Ragueneau E, et al. The IntAct database: efficient access to fine-grained molecular interaction data. *Nucleic Acids Res*. 2022;50(D1):D648-D653. https://doi.org/10.1093/nar/gkab1006.
105 Oughtred R, Stark C, Breitkreutz BJ, et al. The BioGRID interaction database: 2019 update. *Nucleic Acids Res*. 2019;47(D1):D529-D541. https://doi.org/10.1093/nar/gky1079.
106 Szklarczyk D, Franceschini A, Wyder S, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res*. 2015;43(Database issue):D447-D452. https://doi.org/10.1093/nar/gku1003.
107 Kanehisa M, Goto S, Kawashima S, et al. The KEGG resource for deciphering the genome. *Nucleic Acids Res*. 2004;32(D1):D277-D2780. https://doi.org/10.1093/nar/gkh063.
108 Rothfels K, Milacic M, Matthews L, et al. Using the reactome database. *Curr Protoc*. 2023;3(4):e722. https://doi.org/10.1002/cpz1.722.
109 Patil A, Nakai K, Nakamura H. HitPredict: a database of quality assessed protein-protein interactions in nine species. *Nucleic Acids Res*. 2011;39(D1):D744-D749. https://doi.org/10.1093/nar/gkq897.
110 Warde-Farley D, Donaldson SL, Comes O, et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res*. 2010;38(Web Server issue):W214-W220. https://doi.org/10.1093/nar/gkq537.
111 Kim CY, Baek S, Cha J, et al. HumanNet v3: an improved database of human gene networks for disease research. *Nucleic Acids Res*. 2022;50(D1):D632-D639. https://doi.org/10.1093/nar/gkab1048.
112 Pan Y, Li R, Li W, et al. HPC-Atlas: computationally constructing a comprehensive atlas of human protein complexes. *Genomics Proteomics Bioinformatics*. 2023;21(5):976-990. https://doi.org/10.1016/j.gpb.2023.05.001.
113 Laman Trip DS, van Oostrum M, Memon D, et al. A tissue-specific atlas of protein-protein associations enables prioritization of candidate disease genes. *Nat Biotechnol*. 2025. https://doi.org/10.1038/s41587-025-02659-z.
114 Han H, Cho JW, Lee S, et al. TRRUST v2: an expanded reference database of human and mouse transcriptional regulatory interactions. *Nucleic Acids Res*. 2018;46(D1):D380-D386. https://doi.org/10.1093/nar/gkx1013.
115 Feng C, Song C, Song S, et al. KnockTF 2.0: a comprehensive gene expression profile database with knockdown/knockout of transcription (co-)factors in multiple species. *Nucleic Acids Res*. 2024;52(D1):D183-D193. https://doi.org/10.1093/nar/gkad1016.
116 Shen WK, Chen SY, Gan ZQ, et al. AnimalTFDB 4. 0: a comprehensive animal transcription factor database updated with variation and expression annotations. *Nucleic Acids Res*. 2023;51(D1):D39-D45. https://doi.org/10.1093/nar/gkac907.
117 Zhang Q, Liu W, Zhang HM, et al. hTFtarget: a comprehensive database for regulations of human transcription factors and their targets. *Genom Proteom Bioinf*. 2020;18(2):120-128. https://doi.org/10.1016/j.gpb.2019.09.006.
118 Zhou KR, Liu S, Sun WJ, et al. ChIPBase v2.0: decoding transcriptional regulatory networks of non-coding RNAs and protein-coding genes from ChIP-seq data. *Nucleic Acids Res*. 2017;45(D1):D43-D50. https://doi.org/10.1093/nar/gkw965.
119 Liska O, Bohar B, Hidas A, et al. TFLink: an integrated gateway to access transcription factor-target gene interactions for multiple species. *Database (Oxford)*. 2022;2022:baac083. https://doi.org/10.1093/database/baac083.
120 Tong Z, Cui Q, Wang J, et al. TransmiR v2.0: an updated transcription factor-microRNA regulation database. *Nucleic Acids Res*. 2019;47(D1):D253-D258. https://doi.org/10.1093/nar/gky1023.
121 Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9:559. https://doi.org/10.1186/1471-2105-9-559.
122 Margolin AA, Nemenman I, Basso K, et al. ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinformatics*. 2006;7 Suppl 1(Suppl 1):S7. https://doi.org/10.1186/1471-2105-7-S1-S7.
123 Lachmann A, Giorgi FM, Lopez G, et al. ARACNe-AP: gene network reverse engineering through adaptive partitioning inference of mutual information. *Bioinformatics*. 2016;32(14):2233-2235. https://doi.org/10.1093/bioinformatics/btw216.
124 Faith JJ, Hayete B, Thaden JT, et al. Large-scale mapping and validation of Escherichia coli transcriptional regulation from a compendium of expression profiles. *PLoS Biol*. 2007;5(1):e8. https://doi.org/10.1371/journal.pbio.0050008.
125 Meyer PE, Kontos K, Lafitte F, et al. Information-theoretic inference of large transcriptional regulatory networks. *EURASIP J Bioinform Syst Biol*. 2007;2007(1):79879. https://doi.org/10.1155/2007/79879.
126 Sales G, Romualdi C. parmigene--a parallel R package for mutual information estimation and gene network reconstruction. *Bioinformatic*. 2011;27(13):1876-1877. https://doi.org/10.1093/bioinformatics/btr274.
127 Mordelet F, Vert JP. SIRENE: supervised inference of regulatory networks. *Bioinformatics*. 2008;24(16):i76-i82. https://doi.org/10.1093/bioinformatics/btn273.
128 Huynh-Thu VA, Irrthum A, Wehenkel L, et al. Inferring regulatory networks from expression data using tree-based methods. *PLoS One*. 2010;5(9):e12776. https://doi.org/10.1371/journal.pone.0012776.
129 Shao X, Lu X, Liao J, et al. New avenues for systematically inferring cell-cell communication: through single-cell transcriptomics data. *Protein Cell*. 2020;11(12):866-880. https://doi.org/10.1007/s13238-020-00727-5.
130 Shao X, Liao J, Li C, et al. CellTalkDB: a manually curated database of ligand-receptor interactions in humans and mice. *Brief Bioinform*. 2021;22(4):bbaa269. https://doi.org/10.1093/bib/bbaa269.
131 Shao X, Li C, Yang H, et al. Knowledge-graph-based cell-cell communication inference for spatially resolved transcriptomic data with SpaTalk. *Nat Commun*. 2022;13(1):4429. https://doi.org/10.1038/s41467-022-32111-8.
132 Shao X, Wang Z, Wang K, et al. A single-cell landscape of human liver transplantation reveals a pathogenic immune niche associated with early allograft dysfunction. *Engineering*. 2024;36:193-208. https://doi.org/10.1016/j.eng.2023.12.004.
133 Cang Z, Zhao Y, Almet AA, et al. Screening cell-cell communication in spatial transcriptomics *via* collective optimal transport. *Nat Methods*. 2023;20(2):218-228. https://doi.org/10.1038/s41592-022-01728-4.
134 Li H, Ma T, Hao M, et al. Decoding functional cell-cell communication events by multi-view graph learning on spatial transcriptomics. *Brief Bioinform*. 2023;24(6):bbad359. https://doi.org/10.1093/bib/bbad359.
135 Jin S, Guerrero-Juarez CF, Zhang L, et al. Inference and analysis of cell-cell communication using CellChat. *Nat Commun*. 2021;12(1):1088. https://doi.org/10.1038/s41467-021-21246-9.
136 Efremova M, Vento-Tormo M, Teichmann SA, et al. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat Protoc*. 2020;15(4):1484-1506. https://doi.org/10.1038/s41596-020-0292-x.
137 Browaeys R, Saelens W, Saeys Y. NicheNet: modeling intercellular communication by linking ligands to target genes. *Nat Methods*. 2020;17(2):159-162. https://doi.org/10.1038/s41592-019-0667-5.
138 Zheng R, Zhang Y, Tsuji T, et al. MEBOCOST maps metabolite-mediated intercellular communications using single-cell RNA-seq. *Nucleic Acids Res*. 2025;53(12):gkaf569. https://doi.org/10.1093/nar/gkaf569.
139 Shao X, Yu L, Li C, et al. Extracellular vesicle-derived miRNA-mediated cell-cell communication inference for single-cell transcriptomic data with miRTalk. *Genome Biol*. 2025;26(1):95. https://doi.org/10.1186/s13059-025-03566-x.
140 Shan N, Lu Y, Guo H, et al. CITEdb: a manually curated database of cell-cell interactions in human. *Bioinformatics*. 2022;38(22):5144-5148. https://doi.org/10.1093/bioinformatics/btac654.
141 Gao J, Mo S, Wang J, et al. MACC: a visual interactive knowledgebase of

metabolite-associated cell communications. *Nucleic Acids Res*. 2024;52(D1): D633-D639. https://doi.org/10.1093/nar/gkad914.
142 Subramanian A, Narayan R, Corsello SM, et al. A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. *Cell*. 2017;171(6):1437-1452.e17. https://doi.org/10.1016/j.cell.2017.10.049.
143 Lamb J, Crawford ED, Peck D, et al. The connectivity map: using gene-expression signatures to connect small molecules, genes, and disease. *Science*. 2006;313(5795):1929-1935. https://doi.org/10.1126/science.1132939.
144 Ding F, Tan A, Ju W, et al. The prediction of key cytoskeleton components involved in glomerular diseases based on a protein-protein interaction network. *PLoS One*. 2016;11(5):e0156024. https://doi.org/10.1371/journal.pone.0156024.
145 Wang L, Li Z, Shao Q, et al. Dissecting active ingredients of Chinese medicine by content-weighted ingredient-target network. *Mol Biosyst*. 2014;10(7):1905-1911. https://doi.org/10.1039/c3mb70581a.
146 Cheng F, Kovacs IA, Barabasi AL. Network-based prediction of drug combinations. *Nat Commun*. 2019;10(1):1197. https://doi.org/10.1038/s41467-019-09186-x.
147 Wang S, Lee D. Community cohesion looseness in gene networks reveals individualized drug targets and resistance. *Brief Bioinform*. 2024;25(3): bbae175. https://doi.org/10.1093/bib/bbae175.
148 Zhou Y, Zhou B, Pache L, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun*. 2019; 10(1):1523. https://doi.org/10.1038/s41467-019-09234-6.
149 Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43):15545-15550. https://doi.org/10.1073/pnas.0506580102.
150 Alhamdoosh M, Ng M, Wilson NJ, et al. Combining multiple tools outperforms individual methods in gene set enrichment analyses. *Bioinformatics*. 2017;33(3):414-424. https://doi.org/10.1093/bioinformatics/btw623.
151 Zhu S, Qian T, Hoshida Y, et al. GIGSEA: genotype imputed gene set enrichment analysis using GWAS summary level data. *Bioinformatics*. 2019;35(1):160-163. https://doi.org/10.1093/bioinformatics/bty529.
152 Hanzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics*. 2013;14:7. https://doi.org/10.1186/1471-2105-14-7.
153 Tomfohr J, Lu J, Kepler TB. Pathway level analysis of gene expression using singular value decomposition. *BMC Bioinformatics*. 2005;6:225. https://doi.org/10.1186/1471-2105-6-225.
154 Peng C, Chen Q, Tan S, et al. Generalized reporter score-based enrichment analysis for omics data. *Brief Bioinform*. 2024;25(3):bbae116. https://doi.org/10.1093/bib/bbae116.
155 Liu H, Yuan M, Mitra R, et al. CTpathway: a crosstalk-based pathway enrichment analysis method for cancer research. *Genome Med*. 2022;14(1): 118. https://doi.org/10.1186/s13073-022-01119-6.
156 Yeganeh PN, Mostafavi MT. Causal disturbance analysis: a novel graph centrality based method for pathway enrichment analysis. *IEEE/ACM Trans Comput Biol Bioinform*. 2020;17(5):1613-1624. https://doi.org/10.1109/TCBB.2019.2907246.
157 Alexeyenko A, Lee W, Pernemalm M, et al. Network enrichment analysis: extension of gene-set enrichment analysis to gene networks. *BMC Bioinformatics*. 2012;13:226. https://doi.org/10.1186/1471-2105-13-226.
158 Glaab E, Baudot A, Krasnogor N, et al. EnrichNet: network-based gene set enrichment analysis. *Bioinformatics*. 2012;28(18):i451-i457. https://doi.org/10.1093/bioinformatics/bts389.
159 Pham L, Christadore L, Schaus S, et al. Network-based prediction for sources of transcriptional dysregulation using latent pathway identification analysis. *Proc Natl Acad Sci U S A*. 2011;108(32):13347-13352. https://doi.org/10.1073/pnas.1100891108.
160 Dong X, Hao Y, Wang X, et al. LEGO: a novel method for gene set over-representation analysis by incorporating network-based gene weights. *Sci Rep*. 2016;6:18871. https://doi.org/10.1038/srep18871.
161 Fang Z, Tian W, Ji H. A network-based gene-weighting approach for pathway analysis. *Cell Res*. 2012;22(3):565-580. https://doi.org/10.1038/cr.2011.149.
162 Pandey V. MiNEApy: enhancing enrichment network analysis in metabolic networks. *Bioinformatics*. 2025;41(3): btaf077. https://doi.org/10.1093/bioinformatics/btaf077.
163 Assenov Y, Ramirez F, Schelhorn SE, et al. Computing topological parameters of biological networks. *Bioinformatics*. 2008;24(2):282-284. https://doi.org/10.1093/bioinformatics/btm554.
164 Scardoni G, Petterlini M, Laudanna C. Analyzing biological network parameters with CentiScaPe. *Bioinformatics*. 2009;25(21):2857-2859. https://doi.org/10.1093/bioinformatics/btp517.
165 Chin CH, Chen SH, Wu HH, et al. cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol*. 2014;8(Suppl 4):S11. https://doi.org/10.1186/1752-0509-8-S4-S11.
166 Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics*. 2003;4:2. https://doi.org/10.1186/1471-2105-4-2.
167 Nepusz T, Yu H, Paccanaro A. Detecting overlapping protein complexes in protein-protein interaction networks. *Nat Methods*. 2012;9(5):471-472. https://doi.org/10.1038/nmeth.1938.
168 Li M, Chen JE, Wang JX, et al. Modifying the DPClus algorithm for identifying protein complexes based on new topological structures. *BMC Bioinformatics*. 2008;9:398. https://doi.org/10.1186/1471-2105-9-398.
169 Perozzi B, Al-Rfou R, Skiena S. Deepwalk: online learning of social representations. In Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. 2014;701-710. https://doi.org/10.1145/2623330.2623732.
170 Jiang P, Singh M. SPICi: a fast clustering algorithm for large biological networks. *Bioinformatics*. 2010;26(8):1105-1111. https://doi.org/10.1093/bioinformatics/btq078.
171 Leung HC, Xiang Q, Yiu SM, et al. Predicting protein complexes from PPI data: a core-attachment approach. *J Comput Biol*. 2009;16(2):133-144. https://doi.org/10.1089/cmb.2008.01TT.
172 Meng X, Xiang J, Zheng R, et al. DPCMNE: detecting protein complexes from protein-protein interaction networks *via* multi-level network embedding. *IEEE/ACM Trans Comput Biol Bioinform*. 2022;19(3):1592-1602. https://doi.org/10.1109/TCBB.2021.3050102.
173 Zaki N, Efimov D, Berengueres J. Protein complex detection using interaction reliability assessment and weighted clustering coefficient. *BMC Bioinformatics*. 2013;14:163. https://doi.org/10.1186/1471-2105-14-163.
174 Zhang J, Ubas AA, de Borja R, et al. Tahoe-100M: a giga-scale single-cell perturbation atlas for context-dependent gene function and cellular modeling. *bioRxiv*. 2025:2025.02.20.639398. https://doi.org/10.1101/2025.02.20.639398.
175 Ramezani M, Weisbart E, Bauman J, et al. A genome-wide atlas of human cell morphology. *Nat Methods*. 2025;22(3):621-633. https://doi.org/10.1038/s41592-024-02537-7.
176 Zhan L, Wang Y, Wang A, et al. A genome-scale deep learning model to predict gene expression changes of genetic perturbations from multiplex biological networks. *Brief Bioinform*. 2024;25(5):bbae433. https://doi.org/10.1093/bib/bbae433.
177 Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498-2504. https://doi.org/10.1101/gr.1239303.
178 Lv Q, Lin J, Wu X, et al. Novel active compounds and the anti-diabetic mechanism of mulberry leaves. *Front Pharmacol*. 2022;13:986931. https://doi.org/10.3389/fphar.2022.986931.
179 Chen G, Yang R, Yang X, et al. Mulberry twig alkaloids for type 2 diabetes mellitus: a systematic review and meta-analysis. *Front Pharmacol*. 2025;16: 1475080. https://doi.org/10.3389/fphar.2025.1475080.
180 Gomez-Verjan JC, Zepeda-Arzate EA, Santiago-de-la-Cruz JA, et al. Unraveling the neuroprotective effect of natural bioactive compounds involved in the modulation of ischemic stroke by network pharmacology. *Pharmaceuticals (Basel)*. 2023;16(10):1376. https://doi.org/10.3390/ph16101376.
181 Zhang Q, Wang A, Xu Q, et al. Efficacy and safety of ginkgo diterpene lactone meglumine in acute ischemic stroke: a randomized clinical trial. *JAMA Netw Open*. 2023;6(8):e2328828. https://doi.org/10.1001/jamanetworkopen.2023.28828.
182 Wang L, Liang X, Chen Y, et al. Ginkgo diterpene lactone meglumine for functional recovery in patients with acute ischemic stroke: a systematic review and Meta-analysis. *J Ethnopharmacol*. 2025;342:119416. https://doi.org/10.1016/j.jep.2025.119416.
183 Gan X, Shu Z, Wang X, et al. Network medicine framework reveals generic herb-symptom effectiveness of traditional Chinese medicine. *Sci Adv*. 2023; 9(43):eadh0215. https://doi.org/10.1126/sciadv.adh0215.
184 Zong W, Tian S, Niu Q, et al. Comparable clinical advantages identification of three formulae on rheumatic disease using a modular-based network proximity approach. *J Ethnopharmacol*. 2025;337(Pt 1):118764. https://doi.org/10.1016/j.jep.2024.118764.
185 Zhang Z, Yang W, Chen J, et al. Efficacy and mechanism of *Schisandra chinensis* active component Gomisin A on diabetic skin wound healing: network pharmacology and *in vivo* experimental validation. *J Ethnopharmacol*. 2025;337(Pt 1):118828. https://doi.org/10.1016/j.jep.2024.118828.
186 Morselli Gysi D, Barabasi AL. Noncoding RNAs improve the predictive power of network medicine. *Proc Natl Acad Sci U S A*. 2023;120(45): e2301342120. https://doi.org/10.1073/pnas.2301342120.
187 Xu Z, Zhang T, Chen H, et al. High-throughput single nucleus total RNA sequencing of formalin-fixed paraffin-embedded tissues by snRandom-seq. *Nat Commun*. 2023;14(1):2734. https://doi.org/10.1038/s41467-023-38409-5.
188 McKellar DW, Mantri M, Hinchman MM, et al. Spatial mapping of the total transcriptome by in situ polyadenylation. *Nat Biotechnol*. 2023;41(4):513-520. https://doi.org/10.1038/s41587-022-01517-6.
189 Sun F, Li H, Sun D, et al. Single-cell omics: experimental workflow, data analyses and applications. *Sci China Life Sci*. 2025;68(1):5-102. https://doi.org/10.1007/s11427-023-2561-0.
190 Liao J, Lu X, Shao X, et al. Uncovering an organ's molecular architecture at single-cell resolution by spatially resolved transcriptomics. *Trends Biotechnol*. 2021;39(1):43-58. https://doi.org/10.1016/j.tibtech.2020.05.006.
191 Zhang J, Fang Y, Shao X, et al. The future of molecular studies through the lens of large language models. *J Chem Inf Model*. 2024;64(3):563-566. https://doi.org/10.1021/acs.jcim.3c01977.
192 Zhu B, Li Z, Jin Z, et al. Knowledge-based *in silico* fragmentation and annotation of mass spectra for natural products with MassKG. *Comput Struct Biotechnol J*. 2024;23:3327-3341. https://doi.org/10.1016/j.csbj.2024.09.001.
193 Armingol E, Baghdassarian HM, Lewis NE. The diversification of methods for studying cell-cell interactions and communication. *Nat Rev Genet*. 2024;25(6):381-400. https://doi.org/10.1038/s41576-023-00685-8.
194 Cui K, Gao X, Wang B, et al. Epsin nanotherapy regulates cholesterol transport to fortify atheroma regression. *Circ Res*. 2023;132(1):e22-e42. https://doi.org/10.1161/CIRCRESAHA.122.321723.
195 Li C, Shao X, Zhang S, et al. scRank infers drug-responsive cell types from untreated scRNA-seq data using a target-perturbed gene regulatory network. *Cell Rep Med*. 2024;5(6):101568. https://doi.org/10.1016/j.xcrm.

2024.101568.
196 Qian J, Shao X, Bao H, et al. Identification and characterization of cell niches in tissue from spatial omics data at single-cell resolution. *Nat Commun*. 2025;16(1):1693. https://doi.org/10.1038/s41467-025-57029-9.
197 Verstegen MMA, Coppes RP, Beghin A, et al. Clinical applications of human organoids. *Nat Med*. 2025;31(2):409-421. https://doi.org/10.1038/s41591-024-03489-3.
198 Kumar N, Prakash PG, Wentland C, et al. Decoding spatiotemporal transcriptional dynamics and epithelial fibroblast crosstalk during gastroesophageal junction development through single cell analysis. *Nat Commun*. 2024;15(1):3064. https://doi.org/10.1038/s41467-024-47173-z.
199 Onesto MM, Kim JI, Pasca SP. Assembloid models of cell-cell interaction to study tissue and disease biology. *Cell Stem Cell*. 2024;31(11):1563-1573. https://doi.org/10.1016/j.stem.2024.09.017.
200 Li L, Mou X, Xie H, et al. *In vitro* tests to evaluate embryotoxicity and irritation of Chinese herbal medicine (Pentaherbs formulation) for atopic dermatitis. *J Ethnopharmacol*. 2023;305:116149. https://doi.org/10.1016/j.jep.2023.116149.
201 Kang Y, Zhang H, Guan J. scINRB: single-cell gene expression imputation with network regularization and bulk RNA-seq data. *Brief Bioinform*. 2024;25(3):bbae148. https://doi.org/10.1093/bib/bbae148.

The cover design illustrates the advancements in the network pharmacology of traditional Chinese medicine (TCM) driven by artificial intelligence. The herbal plant symbolizes the chemical characterization of the material basis in TCM. The glowing and translucent human figure, along with several luminous spheres, represents the complex and dynamic cell-cell communication (CCC) network in the healthy and diseased human body, highlighting its crucial for understanding the physiological, pathological, and pharmacological mechanisms. Furthermore, the background visible binary codes and the connection between the herbal plant and human body emphasize the development trend and future direction of AI-driven TCM network pharmacology, including the ingredient identification, CCC network construction, and CCC network analysis using AI methods. In summary, the cover employs symbolic visuals to align with the article′s emphasis on the AI technology and CCC network as the next paradigm-shifting force in TCM network pharmacology.