

基于数据挖掘和生物信息学探究糖尿病治疗的用药规律及药理机制研究*

曾珉, 万未希, 黄 铭, 刘思敏, 李 响,
蔡 晶, 邓 锐, 董 颖, 蔡 羽**

(湖北中医药大学药学院 武汉 430065)

摘 要:目的 基于多算法挖掘治疗糖尿病的用药规律,并且基于生物信息学探究关键组方治疗糖尿病的药效药理机制,为糖尿病的辨证防治提供指导。方法 基于古今医案云平台 V2.3.5 通过数据挖掘分析方法研究糖尿病治疗的组方用药规律,同时基于关联规则、聚类分析等算法获得新关键组方,另外基于生物信息学分析关键组方治疗糖尿病的“多成分-多靶点-多通路”的药效药理机制。结果 通过古今医案云平台 V2.3.5 共筛选出 68 首中药处方、214 味中药及 4 首新组方,基于网络药理学分析发现黄芪、甘草、五味子、丹参及当归新组方治疗糖尿病基于 193 种生物活性化合物、30 种关键靶点、4151 种 GO 过程以及 74 条 KEGG 生物通路,并且基于分子对接发现 glycyroside、sitosterol 等关键活性成分。结论 本研究初步阐明了糖尿病治疗的方剂用药规律,并基于生物信息学初步揭示了复方用药的药效药理机制,为糖尿病的中医药临床辨证治疗提供数据算法支持以及生物信息学基础。

关键词:数据挖掘 网络药理学 分子对接 用药规律 药理机制

doi: 10.11842/wst.20210429008 中图分类号: R285 文献标识码: A

World Health Statistics 2020 表明世界人数的 71% 死于非传染性疾病,其中主要的四大疾病为糖尿病 (Diabetes mellitus, DM)、心脑血管疾病、癌症和慢性呼吸系统疾病^[1],同时最新数据显示,2019 年全球约有 4.63 亿 DM 成人患者,而预计到 2045 年 DM 患者会达到 7.002 亿^[2]。DM 定义为以高血糖为特征的代谢性疾病,其具体为胰岛素进行性分泌不足或胰岛素利用障碍,由此可分为 1 型糖尿病和 2 型糖尿病 (Type 2 diabetes mellitus, T2DM)^[3]。DM 的病因尚未阐明,但其临床症状以及病理特征较为明显,即“三多一少”以及多器官衰竭等后期并发症等^[4]。目前 DM 无法治愈,主要的治疗方式是通过降糖药来控制血糖,包含胰岛素类、二甲双胍类和胰岛素增敏剂等,但其往往伴随着

低血糖、胃肠道不良反应等不良反应^[5]。

循证医学表明中西医结合治疗 DM 的疗效明显高于纯西药治疗,单味和复方中药也具降糖作用,同时中药在降低毒副作用以及防治 DM 并发症具有明显优势^[6],中医药在临床治疗 DM 具有治疗效果好、不良反应小和治疗成本低等优势^[7]。

古今医案云平台 V2.3.5 集成云计算、大数据等应用模式及智能信息处理技术,从临床科研出发收集整理 30 余万条古今名医医案,为名医传承与中医药用药规律提供便利的技术工具,其具体包含四气、五味、归经和功效等中药属性分析等分析模块^[8]。网络药理学基于多组学关键分析,其建立在计算机虚拟计算、高通量组学数据分析及网络数据库检索基础上,具体

收稿日期:2021-04-29

修回日期:2022-01-15

* 湖北中医药大学青苗计划 (2020ZZX003): 桑瓜饮通过 JAK2/STAT3 信号通路调节 2 型糖尿病大鼠糖脂代谢的作用机制研究,负责人:蔡羽。

** 通讯作者:蔡羽,博士,讲师,主要研究方向:中药复方物质基础及药理机制。

围绕生物学、生物网络构建分析、冗余性、连接性和多效性等进行药物特性分析,其是中医药现代化的关键分析工具^[9-10]。分子对接基于虚拟计算以及多种算法模拟化学成分和关键靶点的对接,其有利于阐述中药化学物质基础和药效药理机理,同时药物-靶标相互预测可为中医药网络药理学奠定基础^[11]。

本研究采用古今医案云平台 V2.3.5、网络药理学和分子对接技术,同时基于临床研究的药效研究,系统性、特异性探究 DM 治疗的用药规律及药效药理机制,系统性评价中药复方治疗糖尿病的成分、靶点和性味归经等药性特征,另外在辨证论治、已病防变、治病求因和正气存内的中医思想指导下,以期临床中医药治疗 DM 提供基础指导。

1 材料与方法

1.1 数据来源

古今医案云平台 V2.3.5 整理的 40 万医案数据为数据来源(古代医案库、现代医案库、名医医案库、共享医案库)。检索条件:专病大数据,内分泌科,糖尿病;截止时间:2021 年 4 月 25 日;同时相关医案数据由两名人员独立提取以及筛选。

1.2 纳入和排除标准

纳入标准:①西医疾病诊断为糖尿病;②以中药治疗为主,同时中药治疗疗效好,或者中西医联合治疗疗效明显优于纯西药治疗;③中药方剂组成明确;④中西医诊断、中西医症状、临床表现完整以及中医证型、中药组成和剂量记录完备者。

排除标准:①治疗效果不明确或者效果不好;②西医疾病诊断不为糖尿病,排除西医诊断糖尿病并发症;③中药方剂组成不明确,或未使用中药治疗;④针灸等非中药处方临床治疗;⑤重复医案。

1.3 数据标准化

古今医案云平台 V2.3.5 标准化工具,其参照《中医药学主题词表》(2008 版)、《中药大辞典》(第二版)、《中药大辞典》(第二版)、《中医临床诊疗术语国家标准(疾病部分)》等。

1.4 统计分析

采用古今医案云平台 V2.3.5 医案大数据分析的数据挖掘部分,对筛选医案进行四诊统计、病证统计和诊疗统计,主要包括症状、舌质、舌苔、脉象、四诊综合、中医疾病、中医证候、证素、治法、中药和中药属性

等统计分析。

1.5 聚类分析

采用平台对频次前 20 个中药以及症状进行聚类分析,其距离类型采用欧氏距离,聚类方法采用最长距离法。最长距离法为类与类之间的距离为其之间两个最远样本之间的距离,欧氏距离是多维空间的点之间的几何距离。其计算公式 1 如下:

$$\text{dis tan } t_{(a,b)} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (1)$$

1.6 网络分析

利用平台数据挖掘分析模块对临床表现—临床表现以及中药—中药进行关联分析,两者的置信度均为 0.5,支持度为 0.1;同时对临床表现—临床表现以及中药—中药进行复杂网络分析,其连接类型为普通关系,边权重为 5,显示边数为 100;另外对临床表现—临床表现、中药—中药、治法—治法和中医证候—中医证候进行社团分析。

1.7 成分数据库构建

通过聚类分析后得到关联新处方,通过 TCMSP 数据库检索新复方药材成分(黄芪、甘草、南五味子、北五味子、丹参、当归),通过药代动力学(absorption-distribution-metabolism-excretion, ADME)筛选活性化合物,具体为口服生物利用度(Oral bioavailability, OB)≥30%,药物相似性(Drug-likeness, DL)≥0.18,另外通过 YaTCM 数据库以及 PubChem 数据库补充化合物相关性质,从而完成新处方成分数据库的构建^[12-14]。同时成分数据库通过 OBioavail 1.1 系统计算 OB,使用 Tanimoto 系数计算 DL,其公式 2 如下:

$$T_{(a,b)} = \frac{a \cdot b}{\|a\|^2 + \|b\|^2 + a \cdot b} \quad (2)$$

其中 a 代表待测化合物的描述符矢量, b 代表药物化合物的数据库描述符矢量。

1.8 靶点预测

通过 PharmMapper、TargetNet 以及 SwissTargetPrediction 数据库预测活性化合物靶点^[15-17],同时筛选活性靶点,具体参数为 Z-score≥0.8, Maximum Generated Conformations 100, Pharmacophore Mapping Human Protein Targets Only, Number of Reserved Matched Targets 100; TargetNet include models with AUC≥0.7, Fingerprint type ECFP4 fingerprints; Swiss Probability≥0.7。PharmMapper 数据库基于药效团模型,采用 Z-score 模型进行活性靶点

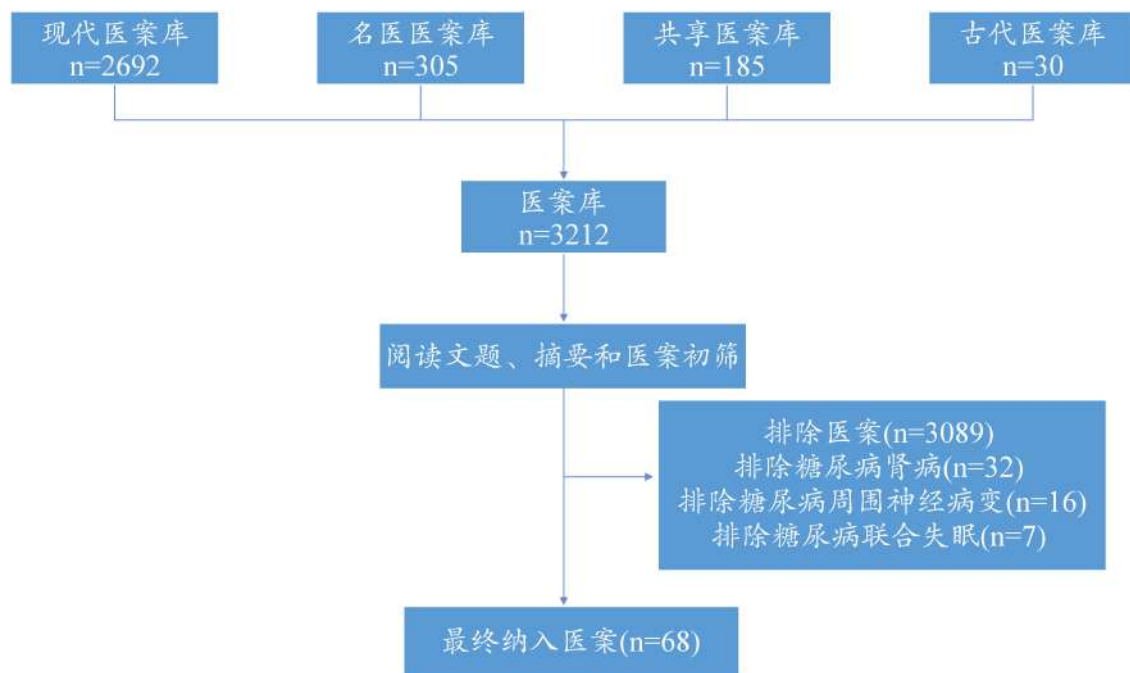


图1 检索流程图

筛选(公式3),同时其包含超过7,000种基于受体的药效团模型。TargrtNet数据库基于定量构效关系(Quantitative structure - activity relationship, QSAR)模型包含623人源关键活性蛋白,其模型见于公式4:

$$Z = \frac{E - \bar{E}_r}{\Delta E_r} \quad (3)$$

其中E是药物的作用评分,E和 ΔE 是随机对应物的相应度量的平均值和标准偏差,Z分数的绝对值越高差异越大。

$$\log\left(\frac{1}{C}\right) = \sum_i \sum_j G_{ij} X_{ij} + \mu \quad (4)$$

其中 X_{ij} 为结构参数,若结构母环中第i个位置有第j类取代基则结构参数取值为1否则为0, μ 为参照分子的活性参数, G_{ij} 为回归系数。

1.9 关键靶点获取

通过 Genecards 以及 DisGetNET 数据库检索“Diabetes Mellitus”获取DM相关疾病基因,同时筛选前数据库前100关键疾病基因,合并去重即得DM相关的疾病基因,同时通过 Uniprot 数据库标准化靶点基因,另外通过利用 Venn 图合并成分靶点以及疾病基因靶点,最终得到关键靶点^[18-20]。

1.10 PPI 网络分析

STRING 数据库是已知和预测的蛋白质-蛋白质相互作用的关键数据库,其涵盖了来自5,090个生物

体的24,584,628个蛋白质数据^[21]。利用STRING分析关键靶点PPI网络,并且通过Cytoscape可视化网络图^[22]。

1.11 GO、KEGG 分析

Metascape 数据库整合了GO, KEGG, DrugBank 和 UniProt 等多个权威数据库,其不仅能完成生物过程和通路富集注释,同时能完成基因相关的蛋白质网络分析和相关的药物分析^[23]。Omicsbean 多组学数据分析系统,可针对多种组合的组学实验设计方案进行数据分析。利用 Metascape 和 Omicsbean 数据库分析关键靶点的GO生物学过程以及KEGG通路机制,同时将关键图表可视化处理。

1.12 分子对接

基于Swissdock 数据库将活性化合物与胰岛素受体(insulin receptor, INSR)(PDB ID: 1gag)进行分子对接,以验证关键活性化合物与能量代谢靶点的结合情况,并且通过 Discovery Studio 3.5 可视化对接结果,并且分析化合物与靶点之间的分子结合力^[24]。

2 结果

2.1 检索筛选结果

基于平台共检索获得医案3212条,依据排除标准剔除,最终纳入符合标准医案68条,涉及中药处方68首、中药214味,平均每首中药处方用药3.14味,总用

表1 数据库及软件汇总表

Database/Software	Website/Version
TCMSP	https://old.tcm-sp-e.com/tcm-sp.php
YaTCM	http://cadd.pharmacy.nankai.edu.cn/yatcm/search
PubChem	https://pubchem.ncbi.nlm.nih.gov/
PharmMapper	http://www.lilab-ecust.cn/pharmmapper/
TargetNet	http://targetnet.scbdd.com/
SwissTargetPrediction	http://swisstargetprediction.ch/
Genecards	https://www.genecards.org/
DisGeNET	https://www.disgenet.org/
STRING	https://www.string-db.org/
RCSB PDB	https://www.rcsb.org/
Metascape	http://metascape.org/gp/index.html#/main/step1
UniProt	https://www.uniprot.org/
Omicsbean	http://www.omicsbean.cn/
Swissdock	www.swissdock.ch
Discovery Studio	V3.5.0
古今医案云平台	V2.3.5
Cytoscape	V3.7.2

表2 基于聚类分析治疗DM的新方剂

No.	新方组合
1	黄芪、甘草、五味子、丹参、当归
2	山药、枸杞子、茯苓、苍术、葛根、白术
3	鸡、鸡内金、泽泻、大黄、知母
4	麦冬、太子参、天花粉、地黄

药频次为756次。检索流程见于图1,同时本研究使用数据库及软件展现于表1。

2.2 统计分析结果

对68条筛选医案进行四诊统计、病证统计和诊疗统计分析(图2),其中四诊统计分析包含舌苔分布(苔白为主),脉象分布(细为主),舌质分布(舌淡为主),以及其汇总分布(口干为主);病证统计分析包含证素-病位分布(阴虚为主)、证素-病性分布(肝为主)、症状分布(口干为主)、治法分布(益气为主)、中医疾病分布(消渴为主)和中医证候分布图(阴虚为主);诊疗统计分析包括功效分布(清热泻火为主)、归经分布(肺为主)、四气分布(平为主)、五味分布(甘为主)和中药分布(黄芪为主),同时将中药归经、四气和五味可视化雷达图。

2.3 聚类分析结果

本研究采用欧氏距离法,最长距离聚类类型,将前20味药以及前20种症状进行聚类分析,并且最终可视化圆形竖向聚类图(图3)。中药聚类分析结果

最终获取4个核心聚类群:①黄芪、甘草、五味子、丹参、当归;②山药、枸杞子、茯苓、苍术、葛根、白术;③鸡 (*Termitomyces albuminosus* (Berk.)Herm[*Collybia alluminosa* (Berk.)Petch])、鸡内金、泽泻、大黄、知母;④麦冬、太子参、天花粉、地黄,并且4个核心聚类群为治疗DM新处方(表2);症状聚类分析结果最终获取7个核心聚类群:①口干、乏力;②消瘦、苔薄、苔薄白;③气短、尿频、脉沉;④头晕、舌淡、自觉症状;⑤脉细、麻木、疲乏、神倦;⑥苔黄、口苦;⑦苔红、苔白、腰酸。

2.4 网络分析结果

利用平台最终获取网络分析结果,包括关联分析、复杂网络分析和社团分析结果,其中关联分析和复杂网络分析包含临床表现-临床表现关联和中药-中药关联;社团分析结果包含治法-治法、中医证候-中医证候、临床表现-临床表现和中药-中药。通过关联分析可知,临床表现则以口干、乏力等为主,中药以黄芪、山药等为主。通过社团分析可知,临床表现和中药基于社团分类形成聚集。

2.5 成分数据库构建结果

基于TCMSP、YaTCM和PubChem数据库完成新处方药材成分数据库的构建,获取863种化合物,并且通过ADME药代动力学筛选获取193种生物活性化合物。其中生物活性较高的化合物有przewalskin b, formyltanshinone, Stigmasterol, Kanzonol F等。

2.6 关键靶点信息以及PPI网络分析

最终获取新处方治疗DM的关键靶点30种,包含胰岛素受体(INSR)、胰岛素样生长因子受体(insulin-like growth factors, IGF1R)、超氧化物歧化酶(Superoxide dismutase 2, SOD2)等关键靶点,并且关键靶点导入STRING数据库分析其蛋白质互相作用(protein-protein interaction, PPI)网络关系,再利用Cytoscape软件可视化网络图,通过NetworkAnalyzer分析其节点值、关联度等网络节点数据。PPI网络包含29个节点,最短路径为812(100%),聚类系数为0.568。

2.7 GO、KEGG分析结果

通过Metscape、Omicsbean数据库分析关键靶点的基因本体论(Gene Ontology, GO)生物过程以及京都基因和基因组百科全书(Kyoto Encyclopedia of Genes and Genomes, KEGG)通路机制,获取新处方治疗DM的潜在药理机制,结果显示新处方干预DM可能基于

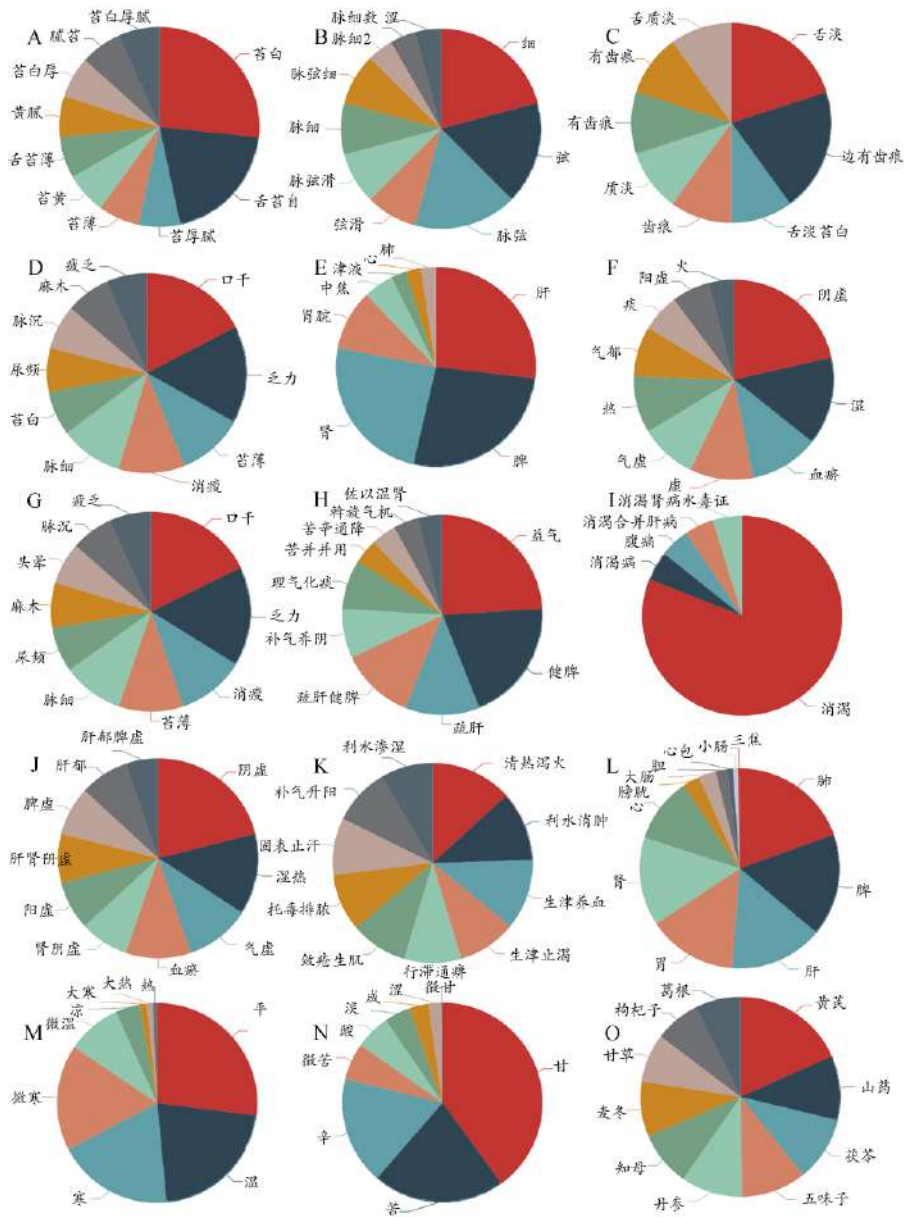


图2 四诊统计、病证统计和诊疗统计分析饼图

注:A舌苔分布图,B脉象分布图,C舌质分布图,D四诊分布图,E证素-病位分布图,F证素-病性分布图,G症状分布图,H治法分布图,I中医疾病分布图,J中医证候分布图,K功效分布图,L归经分布图,M四气分布图,N五味分布图,O中药分布图。

3048种生物过程、189种细胞组分、286种分子功能以及74条生物通路,最终可视化GO(图4)以及KEGG过程。其中GO过程包含生物过程,细胞组分和分子功能,同时GO相关热图展现于图5,KEGG细节系列图展示于图6。GO汇总图关键部分中包含对化学性反应、膜状细胞器、蛋白质结合等。KEGG通路汇总图包含AMPK通路、胰岛素抵抗、MAPK通路、长寿相关通路调节等。GO相关热图解释了关键靶点与DM疾病的相关性,以及包含部分关键基因如肝细胞核因子4

(Hepatocyte nuclear factor 4, HNF4)、肝细胞核因子1 (Hepatocyte nuclear factor 1, HNF1)等,GO与KEGG关联网络图揭示了葡萄糖相关的调控的重要性。由于版面原因部分图表出现于流程图中但无法展现于正文,如生物过程、细胞组分、诊疗统计分析雷达图、复杂网络、关联分析分析结果图、社团分析结果图、PPI网络图、KEGG汇总图、GO与KEGG关联网络图、T-P网络图和分子功能图等,若有需要请联系作者获取相应图表。

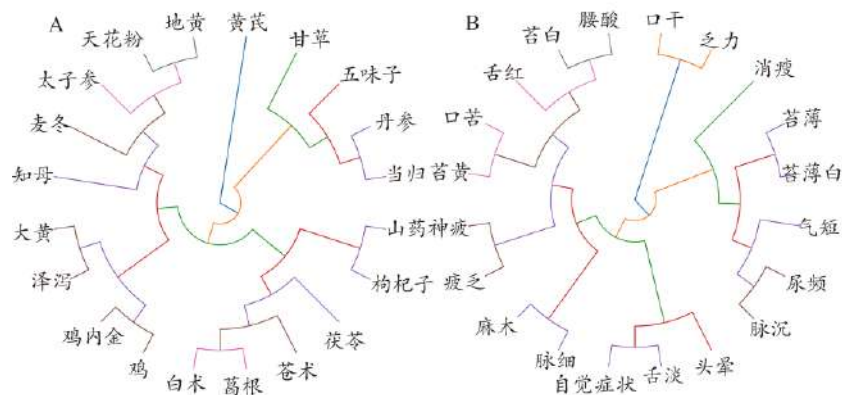


图3 聚类分析圆形竖向图

注:A:药材聚类分析结果图,B:症状聚类分析结果图。

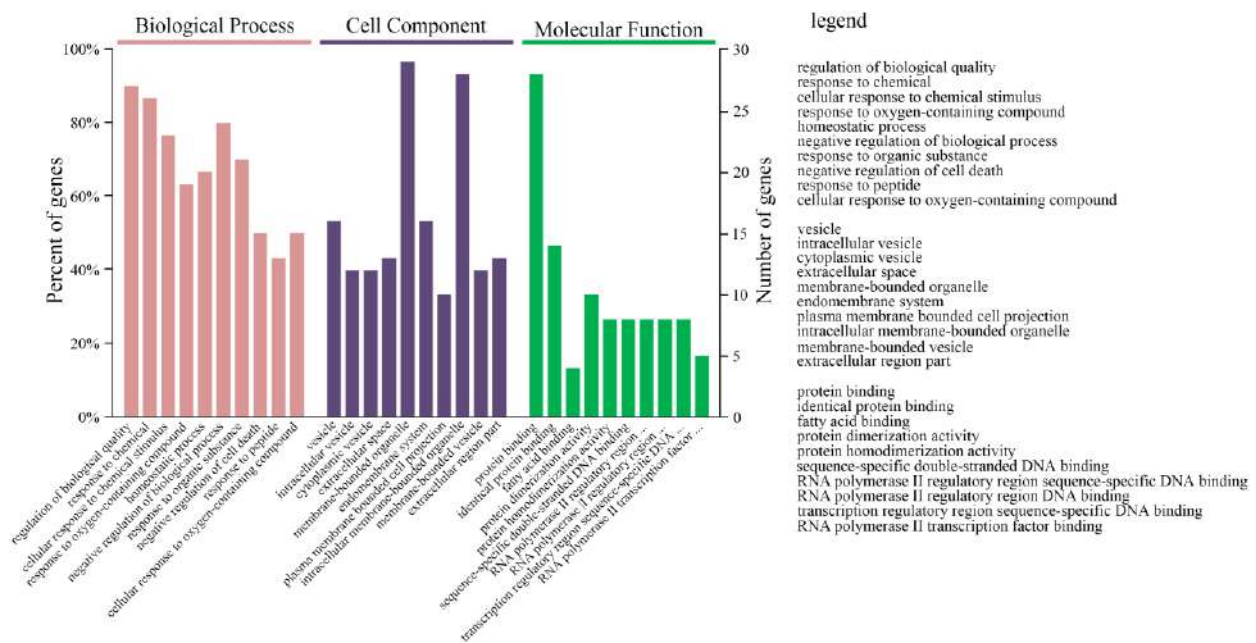


图4 GO汇总图

2.8 分子对接结果

基于SwissDock数据库验证活性化合物与关键蛋白质的结合情况,最终验证了193种生物活性化合物与INSR的分子对接情况(图7、图8),其中6-o-syringyl-8-o-acetyl shanzhiside methyl ester(6-o-丁香酰-8-乙酰山芝思德甲酯),glycyroside(黄甘草苷),sitosterol(谷甾醇)等与INSR蛋白质分子结合情况较好。

2.9 网络图构建

利用Cytoscape揭示新复方干预DM的成分、靶点、通路关系,最终绘制出靶点-通路(Target-pathway,

T-P)网络图,同时T-P网络图较为关键的节点为IGF1R、INSR、胰岛素抵抗(Insulin resistance, IR)通路等。其中较为关键的靶点有TNF、INSR、STAT3等,关键通路有AMPK通路、PI3K-Akt通路、MAPK通路等。同时研究流程图展现于图9。

3 讨论

本研究基于数据挖掘以及生物信息学探究中药复方在DM治疗发挥药效药理作用的多成分、多靶点和多通路的分子机理异同以及复方药味归经的中医联系,同时基于多算法分析DM治疗方剂的用药规律,

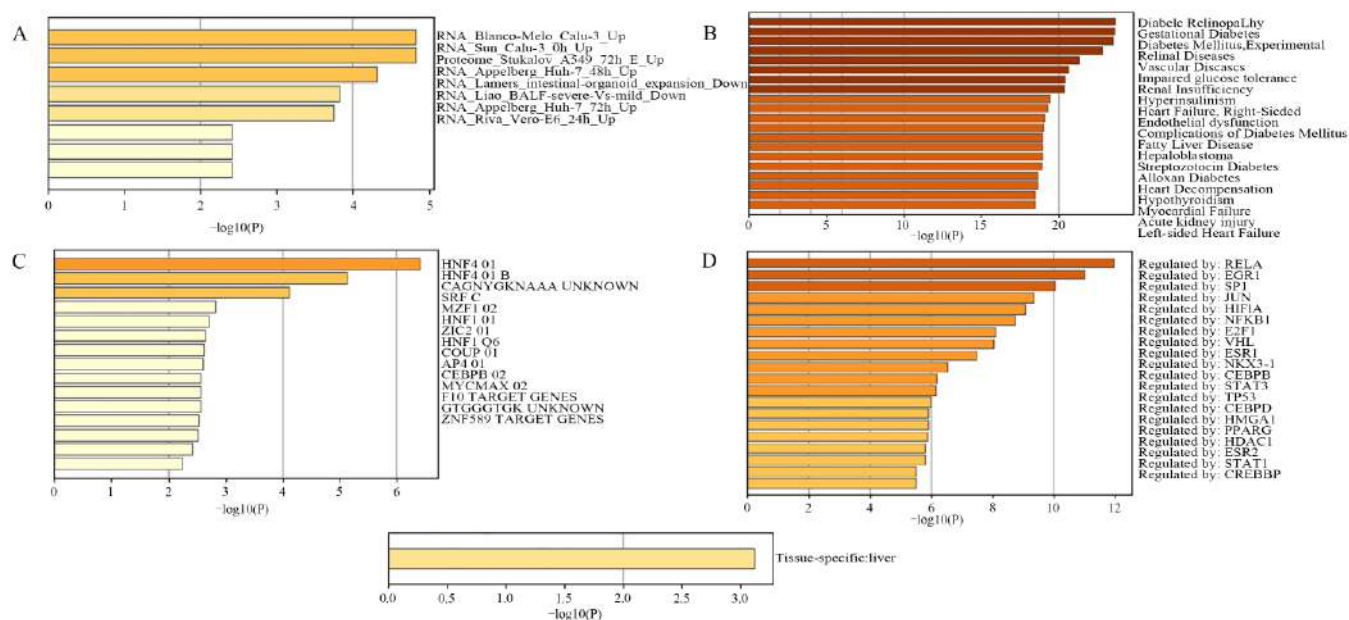


图5 GO相关热图

注:A:GO过程与COVID关联热图,B:GO过程与DisGeNET关联热图,C:GO过程与TRRUST关联热图,D:GO过程与PaGenBase关联热图,E:复方干预DM器官特异性图。

最终对68首处方、214进行多算法、多统计分析,同时通过网络药理学分析出新处方治疗DM基于193种生物活性成分、30种关键靶点和74种生物通路,为临床中医药辩证治疗DM提供了指导与思路,为临床中西医结合治疗DM奠定了生物信息学基础。

数据挖掘基于机器学习、模式识别、数据库算法、人工智能和可视化技术等分析大量数据从而搜索隐藏其中的信息,其广泛应用于中医药领域,其是加快中医药现代化的有效技术,吴嘉瑞等基于数据挖掘探究国医大师颜正华牡蛎处方的用药规律^[25],任星等基于数据挖掘、网络药理学探究颈动脉粥样硬化的复方用药规律及药理机制^[26]。网络药理学和分子对接是研究中医药多成分、多靶点和多通路的关键技术,是中医药现代化的新机遇,宋婷等通过网络药理学发现清肺排毒汤治疗COVID-19基于200个相关靶点和30条信号通路的潜在药理机制^[27],居文政等通过网络药理学和分子对接发现达原饮生物活性化合物可能通过与ACE2发挥抗COVID-19活性^[28]。

循证医学表明,中西医联合治疗T2DM的胰岛素抵抗的疗效优于纯西药治疗^[29],同时中西医联合治疗可有效降低T2DM患者餐后两小时血糖和糖化血红蛋白水平^[30],另外中药在缓解DM并发症具有明显的治疗优势,其改善糖尿病周围神经病变、糖尿病肾病疗

效优于纯西药治疗,具体其可降低糖尿病肾病相关病理指标,以及改善不同的神经传导速度^[31]。

数据挖掘结果显示高频中药有黄芪、山药、茯苓、五味子、丹参,同时基于算法分析出新处方包含五味子、丹参、当归、甘草和黄芪。研究发现黄芪皂苷对DM鼠肾脏具有一定的保护作用^[32],山药多糖可改善T2DM鼠氧化应激及糖脂代谢^[33],茯苓多糖可改善而行DM的糖耐受异常以及糖脂代谢紊乱^[34],五味子基于抗氧化因子Nrf2的调节对db/db小鼠氧化应激具有保护作用^[35]。

基于网络药理学发现新处方可能基于193种生物活性化合物,其中Cryptotanshinone(隐丹参酮)可促进T2DM创面愈合,对血管生成、炎症反应和细胞外基质重塑具有调节作用^[36],luteolin(木犀草素)基于MEK/ERK信号通路对DM白内障鼠的晶状体保护作用^[37],formononetin(刺芒柄花素)可基于AKT/FoxO1通路改善DM鼠的心肌损伤^[38],kaempferol(山奈酚)对糖尿病肾病具有肾功能以及组织病理损伤的保护作用^[39]。同时分子对接结果显示6-o-syringyl-8-o-acetylshanzhiside methyl ester、glycoside、sitosterol与AMPK蛋白质分子结合力较高,表现其潜在的抗DM效果,另外研究发现sitosterol基于IRS-1/Akt通路缓解DM大鼠的胰岛素抵抗^[40]。

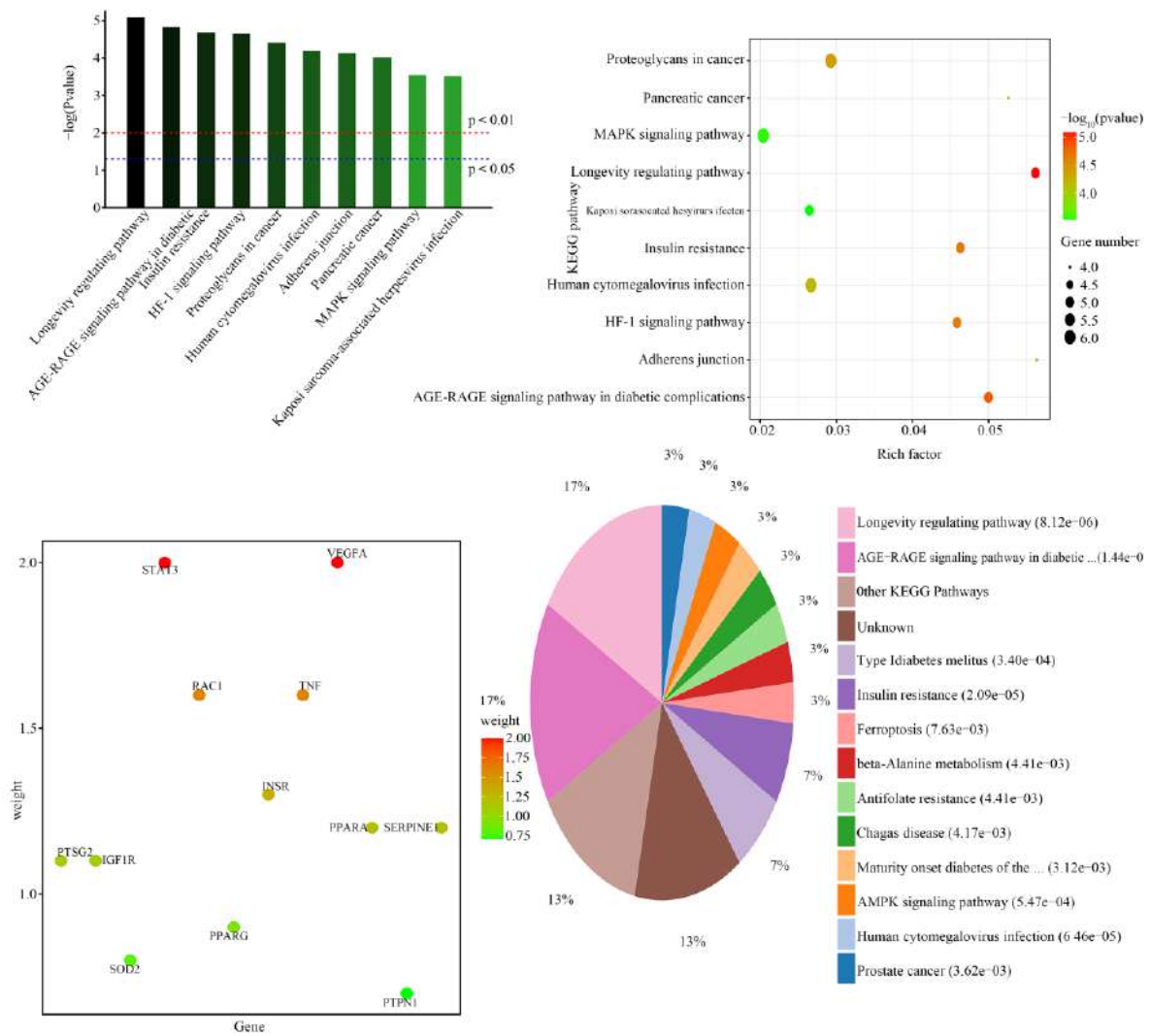


图6 KEGG系列图

注:A:KEGG统计学意义柱状图,B:KEGG汇总气泡图,C:KEGG基因气泡图,D:KEGG统计扇形图。

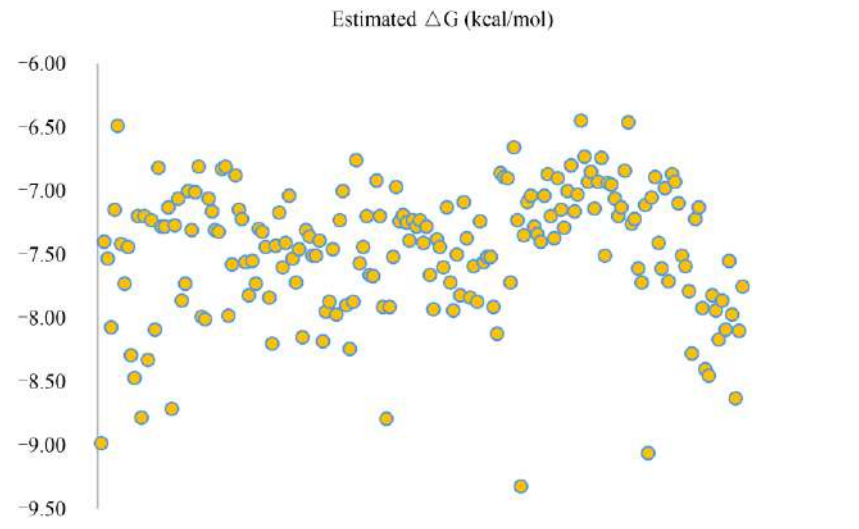


图7 分子对接散点图

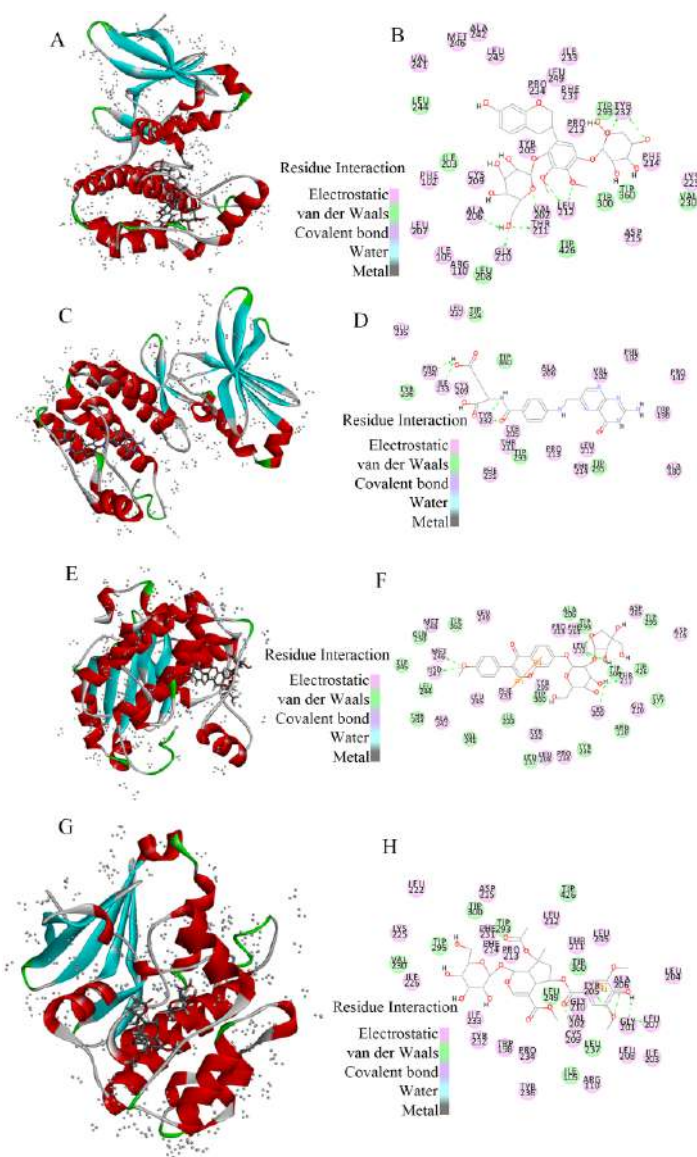


图8 分子对接可视化2D、3D图

注:A, B: MOL000374与INSR,C, D: MOL000433与INSR,E, F: MOL004917与INSR,G, H: MOL007051与INSR。

研究表明,AMPK通路、胰岛素受体抵抗通路、糖尿病并发症中的AGE-RAGE信号通路等与DM以及其并发症紧密相关^[41-43],本研究基于生物信息学发现新处方治疗DM的潜在生物通路机制基于74种生物通路,其机制与AMPK通路、并发症中的AGE-RAGE信号通路、PI3K-Akt通路以及其它炎症氧化应激通路,通过上述关键通路药物有效地干预DM,特别是在同病异治、异病同治和辨证论治的中医思想指导下,在融合西医基础治疗的基础上达到有效控制血糖以及防治并发症症状。

本文利用分子对接验证谷甾醇及胰岛素受体结合情况良好,研究显示 β -谷甾醇(谷甾醇)有潜在的类胰

岛素样作用,有利于降低血糖和缓解胰岛素抵抗。研究显示, β -谷甾醇具有抗高脂血症、抗氧化应激、降低血糖和缓解胰岛素抵抗的作用^[44-46]。同时研究显示 β -谷甾醇与IRS1分子对接良好,同时基于药效实验、PCR和WB实验验证发现 β -谷甾醇通过IRS-1/Akt介导减轻脂肪组织的胰岛素抵抗的T2DM大鼠的胰岛素信号传递,表现出良好的类胰岛素样作用^[47]。同时Ponnulakshmi R等人基于分子对接以及药效药理实验发现, β -谷甾醇通过调节胰岛素信号分子,具有抗高脂和抗高血糖的活性,其使生化状况、氧化应激标志物和抗氧化剂酶水平正常化^[48]。另外有研究表明, β -谷甾醇通过介导氧化应激、脂肪动员、葡萄糖生成、胰

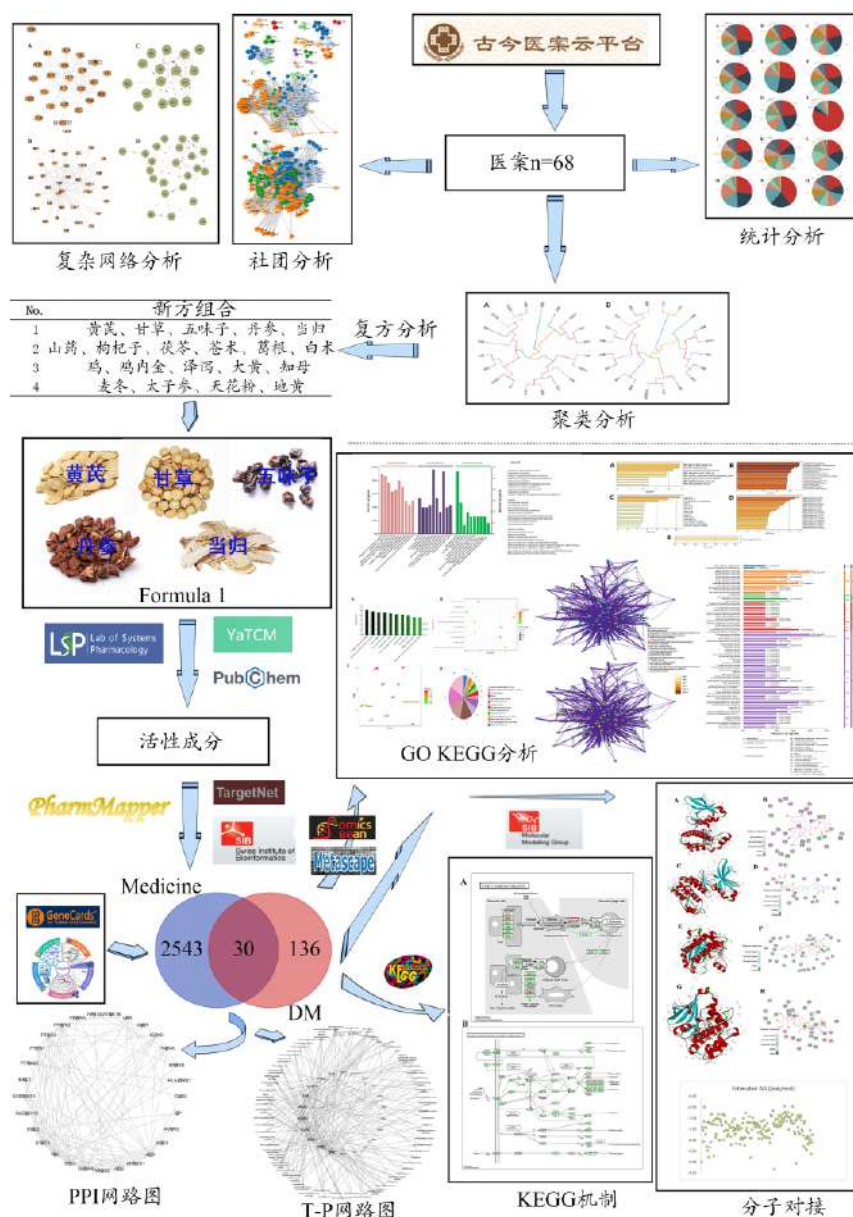


图9 研究逻辑流程图

胰岛素抵抗等参与调控糖脂代谢,从而发挥治疗DM和高脂血症的作用^[49-54]。

本研究基于课题组抗DM中药物质基础、机理研究以及生信分析相关研究,整合多算法数据挖掘和前沿生信分析创新性地分析抗DM中药方剂治疗的用药规律以及药效药理机制,同时系统性基于前沿算法以

及数据库分析中药方剂的物质基础以及药效药理机制,为后续课题组开发抗DM的关键中药方剂奠定了基础,同时本研究缺乏相应实验验证的不足为后续研究提供了新思路,进而指导后续中药方剂治疗DM研究的物质基础、表观药效以及药理机制实验,为抗DM关键中药方剂开发指明了新方向。

参考文献

- 1 World Health Statistics 2020: monitoring health for the SDGs, sustainable development goals. WHO, Geneva (2020). *Revue*
- 2 International Diabetes Federation. IDF Diabetes atlas 9th edition 2019. *Dépidémiologie Et De Santé Publique*, 2016.

- 2019
- 3 Celeste C T, Louis H P. Update on diabetes classification. *Med Clin North Am*, 2015, 99(1):1-16.
 - 4 Warren, Shields. The Pathology of Diabetes Mellitus. *JAMA*, 1952, 149(17):1609.
 - 5 Krentz A J, Bailey C J. Oral antidiabetic agents: current role in type 2 diabetes mellitus. *Drugs*, 2005, 65(3):385-411.
 - 6 陈威妮, 刘志龙, 张弛, 等. 中药治疗2型糖尿病随机对照试验的系统评价. *中国中医药信息杂志*, 2012, 19(1):21-24, 28.
 - 7 Li W L, Zheng H C, Bukuru J, *et al.* Natural medicines used in the traditional Chinese medical system for therapy of diabetes mellitus. *J Ethnopharmacol*, 2003, 92(1):1-21.
 - 8 张文顺, 曲超, 佟奕霖, 等. 基于古今医案云平台的中医药治疗早期糖尿病肾病组方规律研究. *中华中医药学刊*, 2021, 39(2):243-247, 288-290.
 - 9 Wang X, Wang Z Y, Li S, *et al.* TCM network pharmacology: A new trend towards combining computational, experimental and clinical approaches. *Chin J Nat Med*, 2021, 19(1):1-11.
 - 10 Andrew L H. Network pharmacology: the next paradigm in drug discovery. *Nat Chem Biol*, 2008, 4(2):682-690.
 - 11 吴纯伟, 路丽, 梁生旺, 等. 药物靶标预测技术在中药网络药理学中的应用. *中国中药杂志*, 2016, 41(3):377-382.
 - 12 Ru J L, Li P, Wang J N, *et al.* TCMSP: a database of systems pharmacology for drug discovery from herbal medicines. *J Cheminform*, 2014, 6:13.
 - 13 Li B Q, Ma C F, Zhao X Y, *et al.* YaTCM: yet another traditional Chinese medicine database for drug discovery. *Comput Struct Biotechnol J*, 2018, 16:600-610.
 - 14 Kim S, Chen J, Cheng T J, *et al.* PubChem in 2021: new data content and improved web interfaces. *Nucleic Acids Res*, 2021, 49(D1):D1388-D1395.
 - 15 Liu X F, Ouyang S S, Yu B, *et al.* PharmMapper Server: a web server for potential drug target identification via pharmacophore mapping approach. *Nucleic Acids Res*, 2010, 38:W609-W614.
 - 16 Yao Z J, Dong J, Che Y J, *et al.* TargetNet: a web service for predicting potential drug-target interaction profiling via multi-target SAR models. *J Comput Aided Mol Des*, 2016, 30(5):413-424.
 - 17 Daina A, Michielin O, Zoete V, *et al.* SwissTargetPrediction: updated data and new features for efficient prediction of protein targets of small molecules. *Nucleic Acids Res*, 2019, 47(W1):W357-W364.
 - 18 Gil S, Naomi R, Inbar P, *et al.* The genecards suite: from gene data mining to disease genome sequence analyses. *Curr Protoc Bioinformatics*, 2016, 54:1.30.1-1.30.33.
 - 19 Piñero J, Ramírez-Anguita J M, Saüch-Pitarch J, *et al.* The DisGeNET knowledge platform for disease genomics: 2019 update. *Nucleic Acids Res*, 2020, 48(D1):D845-D855.
 - 20 ConsortiumUniProt. UniProt: a hub for protein information. *Nucleic Acids Res*, 2015, 43(Database issue):D204-D212.
 - 21 Damian S, Annika L G, David L, *et al.* STRING v11: protein - protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res*, 2019, 47(D1):D607-D613.
 - 22 Paul S, Andrew M, Owen O, *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*, 2003, 13(11):2498-2504.
 - 23 Zhou Y 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.
 - 24 Grosdidier A, Zoete V, Michielin O, *et al.* SwissDock, a protein-small molecule docking web service based on EADock DSS. *Narnia*, 2011, 39(suppl2):W270-W277.
 - 25 吴嘉瑞, 郭位先, 张晓朦, 等. 基于数据挖掘的国医大师颜正华含牡蛎处方用药规律研究. *中国中药杂志*, 2014, 39(14):2762-2766.
 - 26 任星, 杨静, 张苑桐, 等. 基于数据挖掘和网络药理学的颈动脉粥样硬化用药规律及作用机制研究. *中药新药与临床药理*, 2021, 32(4):518-525.
 - 27 许冬玉, 许玉龙, 王至婉, 等. 基于网络药理学研究清肺排毒汤治疗新型冠状病毒肺炎的作用机制. *中药药理与临床*, 2020, 36(1):26-32.
 - 28 宗阳, 丁美林, 贾可可, 等. 基于网络药理学和分子对接法探寻达原饮治疗新型冠状病毒肺炎(COVID-19)活性化合物的研究. *中草药*, 2020, 51(4):836-844.
 - 29 聂春丽, 李靖, 郑世静, 等. 中西医结合治疗2型糖尿病胰岛素抵抗的Meta分析. *世界中医药*, 2018, 13(3):767-775.
 - 30 代晓颖, 张燕, 易静, 等. 中医药降低早期糖尿病肾病患者蛋白尿水平疗效的Meta分析. *北京中医药大学学报*, 2016, 39(2):144-150.
 - 31 黄勇, 顾静, 李应东. 黄芪桂枝五物汤干预糖尿病周围神经病变临床RCT文献的Meta分析. *甘肃中医学院学报*, 2013, 30(2):65-68.
 - 32 肖峰, 胡雅国, 吴石楠, 等. 黄芪皂苷提取物对糖尿病大鼠肾脏的保护作用. *中国中药杂志*, 2015, 40(10):2014-2018.
 - 33 植飞, 邢琪昌, 汪莹, 等. 佛手山药多糖对2型糖尿病大鼠糖脂代谢及氧化应激的影响. *食品科学*, 2017, 38(5):262-266.
 - 34 黄聪亮, 郑佳俐, 李凤林, 等. 茯苓多糖对II型糖尿病小鼠降糖作用研究. *食品研究与开发*, 2016, 37(4):21-25.
 - 35 董奥, 谭小月, 孔琪, 等. 五味子醇提液对糖尿病肾病小鼠氧化应激的保护作用及机制研究. *中草药*, 2019, 50(24):6038-6044.
 - 36 Song M, Chen L, Zhang L S, *et al.* Cryptotanshinone enhances wound healing in type 2 diabetes with modulatory effects on inflammation, angiogenesis and extracellular matrix remodelling. *Pharm Biol*, 2020, 58(1):845-853.
 - 37 燕洁静, 乔瑛, 王芬芬. 木犀草素对糖尿病性白内障大鼠的晶状体保护作用及对MEK/ERK信号通路的影响研究. *广西医科大学学报*, 2021, 38(3):537-543.
 - 38 程圣杰, 何夕松, 黄健, 等. 刺柄柄素通过AKT/FoxO1信号通路改善糖尿病小鼠心肌损伤. *中药新药与临床药理*, 2020, 31(10):1165-1172.
 - 39 汤利华, 方超, 王浩然, 等. 山奈酚对高糖诱导的糖尿病肾病大鼠肾功能和组织病理损伤的保护作用. *免疫学杂志*, 2018, 34(12):1041-

- 1046.
- 40 Shyamaladevi B, Madhan K, Ponnulakshmi R, *et al.* Beta-sitosterol attenuates insulin resistance in adipose tissue via IRS-1/Akt mediated insulin signaling in high fat diet and sucrose induced type-2 diabetic rats. *Eur J Pharmacol*, 2020, 873:173004.
 - 41 Winder W W, Hardie D G. AMP-activated protein kinase, a metabolic master switch: possible roles in Type 2 diabetes. *Am J Physiol*, 1999, 277(1):E1-E10.
 - 42 李斌, 范源, 李鑫. 基于PI3K/Akt信号通路的中药治疗2型糖尿病胰岛素抵抗研究进展. *中成药*, 2017, 39(1):151-154.
 - 43 Ravichandran R, Yan F S, Ann M S, *et al.* Receptor for AGE (RAGE): signaling mechanisms in the pathogenesis of diabetes and its complications. *Ann N Y Acad Sci*, 2011, 1243:88-102.
 - 44 Babu S, Jayaraman S. An update on β -sitosterol: A potential herbal nutraceutical for diabetic management. *Biomed Pharmacother*, 2020, 131:110702.
 - 45 Rahimifard M, Manayi A, Baeri M, *et al.* Investigation of β -sitosterol and prangol extracted from achillea tenifolia along with whole root extract on isolated rat pancreatic islets. *Iran J Pharm Res*, 2018, 17(1): 317-325.
 - 46 Ramalingam S, Packirisamy M, Karuppiah M, *et al.* Effect of β -sitosterol on glucose homeostasis by sensitization of insulin resistance via enhanced protein expression of PPR γ and glucose transporter 4 in high fat diet and streptozotocin-induced diabetic rats. *Cytotechnology*, 2020, 72(3):357-366.
 - 47 Babu S, Krishnan M, Rajagopal P, *et al.* Beta-sitosterol attenuates insulin resistance in adipose tissue via IRS-1/Akt mediated insulin signaling in high fat diet and sucrose induced type-2 diabetic rats. *Eur J Pharmacol*, 2020, 873:173004.
 - 48 Ponnulakshmi R, Shyamaladevi B, Vijayalakshmi P, *et al.* In silico and in vivo analysis to identify the antidiabetic activity of beta sitosterol in adipose tissue of high fat diet and sucrose induced type-2 diabetic experimental rats. *Toxicol Mech Methods*, 2019, 29(4):276-290.
 - 49 Gupta R, Sharma A K, Dobhal M P, *et al.* Antidiabetic and antioxidant potential of β -sitosterol in streptozotocin-induced experimental hyperglycemia. *J Diabetes*, 2011, 3(1):29-37.
 - 50 Chai J W, Lim S L, Kanthimathi M S, *et al.* Gene regulation in β -sitosterol-mediated stimulation of adipogenesis, glucose uptake, and lipid mobilization in rat primary adipocytes. *Genes Nutr*, 2011, 6(2): 181-188.
 - 51 Xu F, Wu H, Wang X, *et al.* RP-HPLC characterization of lupenone and β -sitosterol in rhizoma musae and evaluation of the anti-diabetic activity of lupenone in diabetic Sprague-Dawley rats. *Molecules*, 2014, 9(9):14114-14127.
 - 52 Jayaraman S, Devarajan N, Rajagopal P, *et al.* β -sitosterol circumvents obesity induced inflammation and insulin resistance by down-regulating IKK β /NF- κ B and JNK signaling pathway in adipocytes of type 2 diabetic rats. *Molecules*, 2021, 26(7):2101.
 - 53 Pandey J, Dev K, Chattopadhyay S, *et al.* β -sitosterol-d-glucopyranoside mimics estrogenic properties and stimulates glucose utilization in skeletal muscle cells. *Molecules*, 2021, 26(11):3129.
 - 54 Ramu R, Shirahatti P S, Nayakavadi S, *et al.* The effect of a plant extract enriched in stigmasterol and β -sitosterol on glycaemic status and glucose metabolism in alloxan-induced diabetic rats. *Food Funct*, 2016, 7(9):3999-4011.

Analysis of Medication Rules and Pharmacological Mechanisms of Diabetes Mellitus Treatment Based on Data Mining and Bioinformatics

Zeng Min, Wan Weixi, Huang Ming, Liu Simin, Li Xiang, Cai Jing, Deng Rui, Dong Ying, Cai Yu
(College of Pharmacy, Hubei University of Chinese Medicine, Wuhan 430065, China)

Abstract: Objective To explore the pharmacological mechanism of essential formulas for the treatment of diabetes, and to guide the prevention and treatment of diabetes, based on multi-algorithm mining the medication rule of treating diabetes and based on bioinformatics. Methods The ancient and modern medical case cloud platform V2.3.5 was used for data mining about the prescription and medication rules of diabetes treatment, and it studied the new essential prescriptions based on algorithms such as association rules and cluster analysis. In addition, bioinformatics was used to analyze the pharmacological mechanism of the critical prescription in the treatment of diabetes. Results A total of 47 prescriptions were screened in the ancient and modern medical case cloud platform V2.3.5, 68 Chinese medicine prescriptions, 214 Chinese medicines and four new formulas were screened Based on network pharmacological analysis, it was found that the new prescriptions for the treatment of diabetes were based on 193 biologically active compounds, 30

key targets, 4151 GO processes and 74 KEGG biological pathways, glycyroside, sitosterol and other vital active ingredients were discovered based on molecular docking. Conclusion This study preliminarily clarifies the rules of prescriptions and medications for diabetes treatment and reveal the pharmacological mechanism of compound medications based on bioinformatics, and it provides data algorithm support and bioinformatics basis for the clinical dialectical treatment of diabetes in traditional Chinese medicine.

Keywords: Data mining, Network pharmacology, Molecular docking, Medication rule, Pharmacological mechanism

(责任编辑:周阿剑、郭思宇, 责任译审:周阿剑, 审稿人:王瑀、张志华)