

【实验研究】

基于文本挖掘和生物医学数据库的新型冠状病毒肺炎药物发现

李彦波^{1,2}, 赵鑫², 吕文良^{1,2*}, 武庆娟², 曹正民^{1,2}, 强睿², 张丽丽²

1. 北京中医药大学 研究生院, 北京 100029

2. 中国中医科学院广安门医院 感染疾病科, 北京 100053

摘要: **目的** 基于文本挖掘技术和生物医学数据库对新型冠状病毒肺炎 (COVID-19) 相关文献进行数据挖掘分析, 探究 COVID-19 及其主要症状发热、咳嗽、呼吸障碍相关基因靶点, 筛选潜在有效的化学药和中药。**方法** 使用 GenCLiP 3 网站获取 COVID-19 和其主要症状咳嗽、发热、呼吸障碍共 4 个关键词的共有靶点, 在 METASCAPE 数据库中对其进行基因本体 (GO) 和通路富集分析, 再利用 String 数据库和 Cytoscape 软件构建共有靶点的蛋白质相互作用网络, 筛选获得核心基因, 运用 DGIdb 数据库、SymMap 数据库针对核心基因进行中西医治疗药物预测。**结果** 获得 COVID-19 及其主要症状共有基因靶点 28 个, 其中有 IL2、IL1B、CCL2 等核心基因 16 个, 使用 DGIdb 数据库筛选获得与 16 个关键靶点相互作用的化学药包括沙利度胺、来氟米特、环孢素等 28 种, 中药包括虎杖、黄芪、芦荟等 70 味。**结论** COVID-19 及其主要症状的病理机制可能和 CD4、KNG1、VEGFA 等 28 个共有基因相关, 可能通过介导 TNF、IL-17 等信号通路参与 COVID-19 病理过程。潜在有效药物可能通过作用相关靶点通路起到治疗 COVID-19 的作用。

关键词: 新型冠状病毒肺炎; 文本挖掘; 生物医学数据库; 药物发现; 发热; 咳嗽; 呼吸障碍

中图分类号: R965.1 **文献标志码:** A **文章编号:** 1674-6376 (2022) 01-0037-11

DOI: 10.7501/j.issn.1674-6376.2022.01.005

Drug discovery for coronavirus disease 2019 based on text mining and biomedical databases

LI Yanbo^{1,2}, ZHAO Xin², LÜWenliang^{1,2*}, WU Qingjuan², CAO Zhengmin^{1,2}, QIANG Rui², ZHANG Lili²

1. Beijing University of Chinese Medicine, Beijing 100029, China

2. Guang'Anmen Hospital, China Academy of Chinese Medical Science, Beijing 100053, China

Abstract: Objective Based on text mining technology and biomedical database, data mining and analysis of coronavirus disease 2019 (COVID-19) were carried out, and COVID-19 and its main symptoms related to fever, cough and respiratory disorders were explored. **Methods** The common targets of COVID-19 and its main symptoms cough, fever and respiratory disorder were obtained by GenCLiP 3 website, Gene ontology in metascap database (GO) and pathway enrichment analysis, then STRING database and Cytoscape software were used to construct the protein interaction network of common targets, the core genes were screened and obtained. DGIdb database and Symmap database were used to predict the therapeutic drugs of traditional Chinese and Western medicine for the core genes. **Results** A total of 28 gene targets of COVID-19 and its main symptoms were obtained, including 16 core genes such as IL2, IL1B and CCL2. Through the screening of DGIdb database, 28 chemicals interacting with 16 key targets were obtained, including thalidomide, leflunomide and cyclosporine et al. And 70 kinds of Chinese materia medica including *Polygonum cuspidatum*, *Astragalus membranaceus* and aloe. **Conclusion** The pathological mechanism of COVID-19 and its main symptoms may be related to 28 common genes such as CD4, KNG1 and VEGFA, which may participate in the pathological process of COVID-19 by mediating TNF, IL-17 and other signal pathways. Potentially effective drugs may play a role in the treatment of

收稿日期: 2021-07-03

基金项目: 国家科技重大专项课题资助项目(2018ZX10303-502-003); 国家重点研发计划中医药现代化研究重点专项(2018YFC1705702); 京津冀中医国家临床重点专科肝病协同病房项目(GZY-GCS-2017-60)

第一作者: 李彦波(1997—), 女, 硕士研究生, 研究方向为中西医结合治疗感染性疾病。E-mail: 1418664324@qq.com

*通信作者: 吕文良, 教授, 博士生导师, 研究方向为中西医结合治疗感染性疾病。E-mail: lvwenliang@sohu.com

COVID-19 through action related target pathway.

Key words: coronavirus disease 2019; text mining; biomedical database; drug discovery; fever; cough; respiratory disorders

新型冠状病毒肺炎(COVID-19)是一类感染新型冠状病毒(SARS-CoV-2)引发的急性全球流行性传染病。少部分COVID-19患者无症状,大部分轻、中度患者临床表现多样,但发热、咳嗽、呼吸困难是最常见的症状,也是主要的症状诊断标准^[1-3]。重症患者可能并发严重的急性呼吸窘迫综合征、感染性休克和多器官衰竭,甚至导致死亡。COVID-19的治疗尚无特效药物,主要集中在对症治疗和呼吸支持方面^[4]。然而临床常用的抗病毒药物及激素治疗多缺乏循证依据,且常伴随腹泻、恶心、呕吐、肝功能损害等不良反应^[5]。探究COVID-19及其主要症状产生的病理机制,寻找有效的治疗药物迫在眉睫。

自疫情暴发以来,COVID-19相关的基础和临床研究迅速增加,在PubMed数据库中,2019年12月与COVID-19相关的检索结果仅有17个,而2021年7月检索到与COVID-19相关的已有16万多条相关结果,大量的文献资源为文本挖掘技术(text mining, TM)的实施提供了数据基础。文本挖掘技术是指利用自然语言处理、人工智能、信息检索和数据挖掘方法来发现、检索和提取一个文本语料库中的信息,是大数据时代综合分析和解释文献及高通量数据的一个有效工具。不断成熟并被逐步推广的文本挖掘技术使从大量COVID-19相关文献中自动收集疾病关键词-基因关联成为可能,有助于探索COVID-19的发病机制、治疗用药等。本研究基于文本挖掘技术和生物信息学方法探究COVID-19主要症状的发生机制,并据此筛选治疗COVID-19的药物,为COVID-19的治疗及新药开发提供思路和参考。

1 方法

1.1 获取COVID-19及其主要症状相关靶点

通过 GenCLiP 3 网站 (<http://ci.smu.edu.cn/genclip3/analysis.php>) 获得 COVID-19 及主要症状相关的基因靶点。在 GenCLiP 3 数据库中,以“所有人类基因”作为基因数据集,选择“在 MEDLINE 中搜索”,并在期刊过滤器中使用默认值,分别搜索关键词“发热”“咳嗽”“呼吸障碍”和“COVID-19”。得到关联基因列表后,对 4 组关联基因集取交集作为 COVID-19 及其主要症状相关靶点基因集,并进行进一步的处理。

1.2 基因本体和通路富集分析

基因本体分析(gene ontology, GO)是注释基因

产物及其功能特征的通用且有效的方法,其中 GO 分析依据功能的不同可区分为生物过程(biological process, BP)、细胞成分(cellular component, CC)、分子功能(molecular function, MF)。将交集靶点导入 METASCAPE (<https://metascape.org>),物种类型选择为“Homo sapiens”,再选择“Custom Analysis”,使用默认设定值作为筛选条件,进行 GO 和京都基因组百科全书(Kyoto encyclopedia of genes and genomes, KEGG)通路注释及富集分析,并用于基因注解可视化。

1.3 构建蛋白质相互作用网络和聚类功能模块

将关键靶点输入到 STRING 数据库,设定物种为“Homo sapiens”,置信度设置为 0.4 进行筛选,获取目标基因的蛋白质相互作用(PPI)信息的 TSV 格式文件。并将该文件导入 Cytoscape 软件,运用 Cytoscape 软件的分子复合物检测(MCODE)和 STRING 应用程序(APP)对重要的基因模块进行聚类分析,执行 MCODE 和 STRING APP 中的默认参数,得到在 PPI 网络中密集连接的网络模块即为 COVID-19 主要症状相关核心基因。

1.4 药物基因相互作用

将筛选出的核心靶点输入药物基因相互作用数据库(DGIdb, <https://www.dgidb.org>)及 SymMap 数据库(<http://www.symmap.org>)中,筛选可能具有 COVID-19 治疗作用的药物。

1.4.1 潜在化学药筛选 使用 DGIdb 数据库检索“1.3”项得到的核心基因,勾选“FDA Approved”“Immunotherapies”2个筛选值,获得靶向 COVID-19 主要症状相关基因的候选药物,这些药物可能代表了潜在的症状治疗手段。

1.4.2 潜在中药筛选 SymMap 数据库是中国传统医学与现代医学的集成数据库。该数据库提供大量有关中药、中国传统医学症状、现代医学症状、化学成分、靶点和疾病的描述性信息,并且通过直接关联或间接相互作用计算推断这 6 种类型的条目之间的对应关系,并以网络的形式展现。将“1.3”项得到的核心基因依次输入 SymMap 数据库,在 Target 选项下进行搜索,查找每个靶点基因所对应的具有统计意义($P < 0.05$)的中药和功效,导出每次的结果,最后对中药名称及中药功效进行统计,得到出现频次较高(频率 ≥ 11 次)的中药及中药功效,这些

中药可能对 COVID-19 症状起到治疗作用。研究数据分析及其流程框架见图 1。

2 结果

2.1 COVID-19 及其主要症状相关靶点和 GO、KEGG 分析

在探索 COVID-19 的潜在药物的过程中,通过文本挖掘工具搜索分别发现了 1 262、302、83、370 个与发烧、咳嗽、呼吸障碍和 COVID-19 作为主题词相关的基因。对 4 组基因取交集,得到 28 个 4 组共有基因。共有基因的生物过程注释可能揭示了 COVID-19 肺炎的病理过程,使用 Metascape 网站分别对交集基因进行蛋白功能分析,结果见表 1。

对 28 个基因进行 GO 注释和 KEGG 通路富集分析,并将注释结果可视化,收集 $P < 0.01$, 最小计数为 3 且富集因子(富集因子是观察到的计数与偶然预期的计数之间的比率) > 1.5 的术语,并根据其成员相似性将其分组。GO 注释(图 2)包含了探究基因组的 BP、CC、MF 重要富集项。在 BP 注释中,

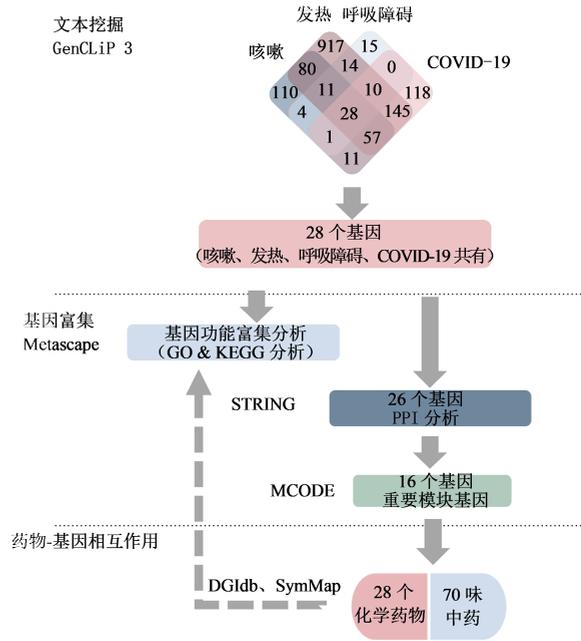


图 1 数据分析及其流程图
Fig. 1 Data analysis and flow chart

表 1 COVID-19 及其主要症状相关靶点和蛋白功能

Table 1 COVID-19 and its major-symptom-related targets and protein functions

序号	关键靶点	蛋白功能
1	CD4	CD 标记; FDA 批准的药物靶点/生物技术药物; 预测细胞内蛋白质
2	NPPB	潜在心血管疾病基因; 预测分泌蛋白
3	IFNA2	预测分泌蛋白
4	MUC1	CD 标记; 预测细胞内蛋白质; 肿瘤相关基因/潜在肿瘤标记物
5	CARD8	预测细胞内蛋白质
6	CRP	预测细胞内蛋白质; 潜在心血管疾病基因; 肿瘤相关基因/潜在肿瘤标记物; 预测分泌蛋白
7	CALCA	转运蛋白/转运蛋白通道和毛孔; 预测分泌蛋白; 潜在肿瘤标记物
8	SERPINA1	预测细胞内蛋白质; 肿瘤相关基因/潜在肿瘤标记物; 潜在心血管疾病基因; 预测分泌蛋白
9	CXCL8	潜在心血管疾病基因; 预测分泌蛋白; 肿瘤相关基因/潜在肿瘤标记物
10	TNF	FDA 批准的药物靶点/(生物技术药物, 小分子药物); 预测细胞内蛋白质; 肿瘤相关基因/(潜在肿瘤标记物, 突变癌症基因); 潜在心血管疾病基因; 预测分泌蛋白
11	MPO	酶/酶蛋白/氧化还原酶; 预测细胞内蛋白质; 肿瘤相关基因/(潜在肿瘤标记物, 突变癌症基因); 潜在心血管疾病基因
12	CSF3	预测分泌蛋白; 肿瘤相关基因/潜在肿瘤标记物
13	VEGFA	RAS 途径相关蛋白; FDA 批准的药物靶点/(生物技术药物, 小分子药物); 预测细胞内蛋白质; 肿瘤相关基因/潜在肿瘤标记物; 潜在心血管疾病基因; 预测分泌蛋白
14	EGR3	转录因子/锌配位 DNA 结合域; 预测细胞内蛋白质
15	GPT	酶/酶蛋白/转移酶; 预测细胞内蛋白质
16	CCL2	FDA 批准的药物靶点/生物技术药物; 潜在心血管疾病基因; 预测分泌蛋白; 肿瘤相关基因/潜在肿瘤标记物
17	TNFSF13	CD 标记; 预测分泌蛋白; 预测细胞内蛋白; 肿瘤相关基因/潜在肿瘤标志物
18	ACE	CD 标记; 酶/(酶蛋白/水解酶, 肽酶/金属肽酶); 预测细胞内蛋白质; FDA 批准的药物靶点/小分子药物; 潜在心血管疾病基因; 预测分泌蛋白
19	LPA	酶/肽酶/丝氨酸型肽酶; 预测细胞内蛋白质; 潜在心血管疾病基因; 预测分泌蛋白质
20	MB	转运蛋白/转运蛋白通道和毛孔; 预测细胞内蛋白质; 潜在心血管疾病基因
21	SOD1	酶/酶蛋白/氧化还原酶; 预测细胞内蛋白质; 肿瘤相关基因/潜在肿瘤标记物
22	IL2	预测分泌蛋白质; 潜在心血管疾病基因; 肿瘤相关基因/潜在肿瘤标记物

续表1

序号	关键靶点	蛋白功能
23	KNG1	预测分泌蛋白质
24	IL1B	FDA 批准的药物靶点/(生物技术药物, 小分子药物); 预测细胞内蛋白质; 转运蛋白/转运蛋白通道和毛孔; 肿瘤相关基因/潜在肿瘤标记物; 预测分泌蛋白质; 潜在心血管疾病基因
25	ALB	预测细胞内蛋白质; 肿瘤相关基因/潜在肿瘤标记物; 潜在心血管疾病基因; 预测分泌蛋白质
26	KANTR	null
27	CSF2	预测分泌蛋白质; 肿瘤相关基因/潜在肿瘤标记物
28	MMP9	酶/(酶蛋白/水解酶, 肽酶/金属肽酶); 肿瘤相关基因/潜在肿瘤标记物; FDA 批准的药物靶点/小分子药物; 潜在心血管疾病基因; 预测分泌蛋白

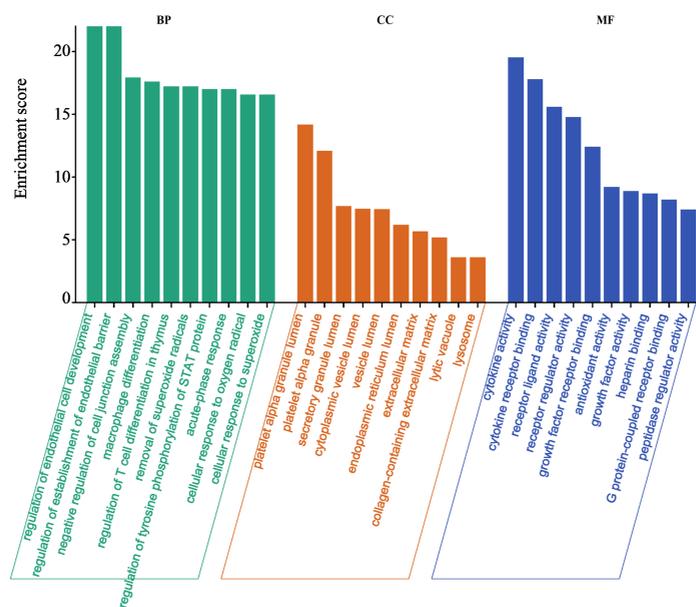


图2 GO 生物功能富集分析结果

Fig. 2 GO function enrichment analysis results

主要参与细胞因子介导的信号通路、白细胞分化、造血调节、细胞黏附调节、基于免疫受体体细胞重组的适应性免疫调节和细胞黏附调节。在CC注释中,显著富集在血小板 α 颗粒管腔、血小板 α 颗粒、分泌颗粒腔、胞质小泡腔、囊腔和细胞外基质。在MF中与细胞因子活性、受体配体活性、细胞因子受体结合、受体调节活性、生长因子受体结合和G蛋白偶联受体结合显著相关。

在KEGG通路分析结果(图3)中,排名前15位的通路分别是细胞因子与细胞因子受体的相互作用、类风湿关节炎、IL-17信号通路、恰加斯病(美国锥虫病)、疟疾、NOD样受体信号通路、造血细胞谱系、糖尿病并发症中的AGE-RAGE信号通路、TNF信号通路、流体剪切应力与动脉粥样硬化、乙型肝炎病毒性肝炎、甲型流行性感冒、阿米巴病、T细胞受体信号通路和Toll样受体信号通路。

2.2 PPI网络分析

将COVID-19及症状共有基因输入STRING数据库中,然后从STRING EXPORT通道以“tsv”格式将数据导出,使用Cytoscape构建网络,得到了具有26个节点,166个边缘的PPI网络,见图4-A。并使用mcode app筛选出16个核心基因,见图4-B。核心基因有趋化因子(C-C基元)配体2(CCL2)、血管紧张素转化酶(ACE)、基质金属蛋白酶9(MMP9)、白细胞介素-2(IL2)、血管内皮生长因子A(VEGFA)、白细胞介素-1(IL1B)、白蛋白(ALB)、C反应性蛋白(CRP)、集落刺激因子2(CSF2)、集落刺激因子3(粒细胞)(CSF3)、重组人白介素-8(CXCL8)、 α 1-抗胰蛋白酶(SERPINA1)、谷氨丙酮酸转氨酶(GPT)、肿瘤坏死因子(TNF)、髓过氧化物酶(MPO)、激肽原1(KNG1)。这些靶点基因在网络中具有重要地位,说明其在COVID-19的发病机

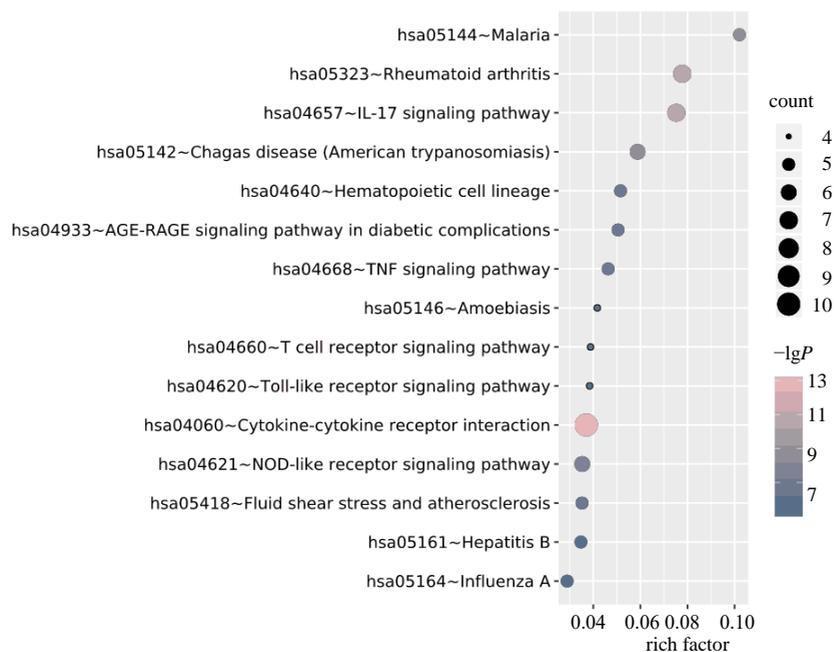
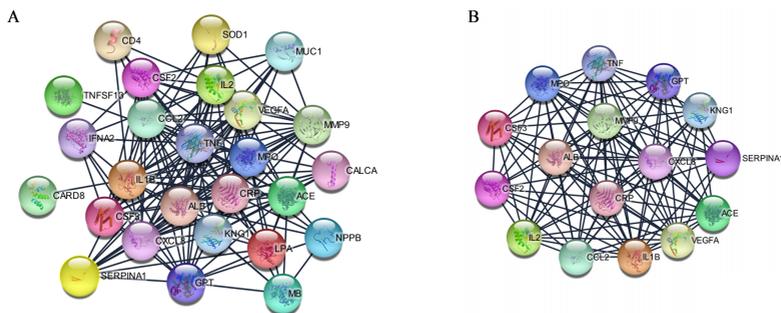


图3 KEGG通路富集分析
Fig. 3 KEGG pathway enrichment analysis



A-PPI网络; B-核心靶点
A-PPI network; B-key targets

图4 PPI网络和核心靶点
Fig. 4 PPI network and key targets

制中也具有重要意义。

2.3 药物基因相互作用

2.3.1 化学药筛选 将筛选所得到的16个核心基因导入DGIdb数据库中,使用预设过滤器默认值,勾选FDA批准、免疫疗法2个筛选项,导出结果中交互评分提示基因与药物交互作用关系证据的可靠程度,筛选出评分 ≥ 0.1 的28种药物(表2),其中大多数药物靶向CSF2、CXCL8、IL1B、VEGFA和TNF,这提示它们能够与这些基因发生相互作用。使用临床试验数据库(ClinicalTrials.gov)检索该药物是否有与COVID-19相关的临床试验并获得临床试验注册号。

2.3.2 中药筛选 将16个核心基因分别输入

SymMap数据库,导出靶点相关中药及功效,并进行统计。共发现具有统计意义的中药556味,功效包括清热解毒、活血祛瘀、化痰和利水渗湿等,统计出排名前10的功效及其频次,结果见表3。并按照中药出现频次进行排序,统计出频次 ≥ 11 的中药共有70味,见表4。

3 讨论

3.1 COVID-19及其主要症状的相关靶点和发生机制

COVID-19作为尚未得到有效控制的一种全球性流行病已经造成了难以计算的经济和生命财产损失。绝大多数轻症患者预后良好,而重症患者则病情进展迅速且具有较高的死亡率。虽然目前针对COVID-19炎症反应具体的发生发展机制仍有待

表 2 化学药筛选结果
Table 2 Prediction results of chemical drugs

药物	基因	PMIDs (部分)	交互评分	COVID-19 相关研究	临床试验注册 号(部分)
Canakinumab(卡那单抗)	IL1B	19169963	10.34	有	NCT04348448
Golimumab(戈利木单抗)	TNF	21079302	4.69	有	NCT05080218
Rilonacept(列洛西普)	IL1B	23319019	3.45	无	
Mechlorethamine(氮芥)	CSF2	10640980	3.11	无	
Adalimumab(阿达木单抗)	CRP、TNF	27096233	2.50	有	NCT04705844
Tacrolimus(他克莫司)	GPT	9825815	1.88	有	NCT05077254
Infliximab(英夫利昔单抗)	TNF、IL1B	16720636	1.59	有	NCT04425538
Etanercept(依那西普)	TNF、CSF2	16720636	1.28	有	NCT05080218
Bevacizumab(贝伐珠单抗)	VEGFA、MMP9、CXCL8	18182667	1.02	有	NCT04305106
Raltitrexed(雷替曲塞)	ALB	15025739	0.87	无	
Procarbazine(甲基苄肼)	CSF2	10640980	0.62	无	
Aldesleukin(阿地白介素)	IL2		0.50	有	NCT04724629
Daclizumab(达克利珠单抗)	IL2		0.50	无	
Propylthiouracil(丙硫氧嘧啶)	MPO、TNF	26509551	0.36	无	
Ustekinumab(乌司奴单抗)	IL1B	28696418	0.34	有	NCT05014555
Rituximab(利妥昔单抗)	IL2	23645042	0.28	有	NCT04877496
Thalidomide(沙利度胺)	TNF	8755512	0.27	有	NCT04273581
Methimazole(甲巯咪唑)	CSF2、TNF、CXCL8	8836317	0.26	无	
Temozolomide(替莫唑胺)	CSF2	12610499	0.23	无	
Glimepiride(格列美脲)	TNF	14686960	0.23	无	
Leflunomide(来氟米特)	CXCL8	10902750	0.19	有	NCT04361214
Oxaliplatin(奥沙利铂)	VEGFA	24090479	0.18	无	
Hydroxychloroquine(羟氯喹)	TNF	9002011	0.17	有	NCT04340544
Capecitabine(卡培他滨)	VEGFA	20125120	0.14	无	
Lenalidomide(来那度胺)	TNF、VEGFA		0.13	有	NCT04361643
Cyclosporine(环孢素)	IL2、TNF	12135706	0.10	有	NCT04412785
Irinotecan(伊立替康)	VEGFA	20125120	0.10	无	
Vinblastine(长春碱)	CSF2	10640980	0.10	无	

表 3 高频中药功效

Table 3 Efficacy of high frequency traditional Chinese medicine

药物功效	频次	占比/%
清热解毒药	606	14.38
活血祛瘀药	293	6.95
化痰药	255	6.05
利水渗湿药	255	6.05
理气药	250	5.93
祛风湿散寒药	187	4.44
辛温解表药	178	4.22
止咳平喘药	173	4.11
补气药	157	3.73
辛凉解表药	154	3.65

进一步研究,但多项临床研究已证实细胞因子风暴在 COVID-19 病情的恶化中起到了非常重要的作用。高细胞因子血症导致炎症物质的大量聚集和渗出,破坏组织器官,引发多器官衰竭和急性呼吸窘迫可能是导致 COVID-19 重型、危重型患者死亡的重要原因^[6]。本研究发现的 28 个与 COVID-19 及其主要症状相关的靶点中与细胞因子相关的有 11 个基因(IFNA2、IL1B、IL2、CXCL8、CCL2、TNFSF13、TNF、MPO、CSF2、CSF3、VEGFA), 占总基因数的 39.3%, 且 GO 注释结果多与细胞因子、免疫功能相关, KEGG 分析结果也多涉及细胞因子、炎症的调控领域。

表4 高频中药(频次≥11)
Table 4 High frequency Chinese materia medica (frequency≥11)

中药	频次	占比/%	中药	频次	占比/%
大枣	15	0.36	篇蓄	11	0.26
虎杖	14	0.34	草果	11	0.26
黄芪	14	0.34	柴胡	11	0.26
芦荟	14	0.34	沉香	11	0.26
马齿苋	14	0.34	丹参	11	0.26
牛蒡子	14	0.34	丁香	11	0.26
桑椹	14	0.34	凤尾草	11	0.26
沙棘	14	0.34	广藿香	11	0.26
乌梅	14	0.34	桂枝	11	0.26
吴茱萸	14	0.34	荷叶	11	0.26
高良姜	13	0.31	红花	11	0.26
金樱子	13	0.31	花生衣	11	0.26
麻黄	13	0.31	鸡屎藤	11	0.26
砂仁	13	0.31	积雪草	11	0.26
玉米须	13	0.31	景天三七	11	0.26
白果	12	0.29	苦丁茶	11	0.26
楮实子	12	0.29	苦杏仁	11	0.26
覆盆子	12	0.29	老鹳草	11	0.26
枸骨叶	12	0.29	麦芽	11	0.26
枸杞子	12	0.29	玫瑰花	11	0.26
贯叶金丝桃	12	0.29	木瓜	11	0.26
黑豆	12	0.29	木蝴蝶	11	0.26
金沸草	12	0.29	牛膝	11	0.26
金荞麦	12	0.29	人参	11	0.26
菊苣	12	0.29	山茱萸	11	0.26
辣椒	12	0.29	水飞蓟	11	0.26
枇杷叶	12	0.29	松子仁	11	0.26
蒲黄	12	0.29	娑罗子	11	0.26
蕤仁	12	0.29	天冬	11	0.26
桑叶	12	0.29	土茯苓	11	0.26
沙苑子	12	0.29	王不留行	11	0.26
芫荽	12	0.29	乌药	11	0.26
艾叶	11	0.26	银杏叶	11	0.26
白鲜皮	11	0.26	余甘子	11	0.26
半夏	11	0.26	紫苏子	11	0.26

3.2 COVID-19及其主要症状涉及的关键靶点

通过PPI网络分析,从28个相关靶点中筛选出16个核心基因,包括GPT、CRP、IL1B、ACE、MPO、KNG1、ALB、CXCL8、CCL2、TNF、IL2、CSF2、KNG1等。上述基因大多已被证实是COVID-19患者的差异表达基因,参与COVID-19的发病过程,且与重症死亡相关。研究表明,与轻症患者相比,重

症COVID-19患者外周血中IL-2^[7]、TNF-α^[8]、CRP^[9]水平显著升高。巨噬细胞中的CCL2活化也参与COVID-19的发病^[10-11]。Shao等^[12]发现血清ALB上升是重症COVID-19的独立危险因素,应密切监测。研究表明在疾病早期调控趋化因子CXCL8的表达有助于抑制病毒感染诱发的炎症反应,增强机体免疫力,降低炎症因子风暴发生率^[13]。调节上述核心基因的表达水平可能在疾病早期起到抑制细胞因子活化、干预疾病进程的作用。

上述核心基因还能通过影响细胞因子的释放、血管通透性、氧化应激反应等多靶点、多途径加重肺泡组织损伤。研究发现COVID-19患者中的外周血清中VEGFA及其内源性抑制剂sFlt-1(VEGF-A受体Flt-1的截短循环形式)显著增加^[14],而VEGFA表达的上调能够提高微小血管的渗透性,破坏肺泡-毛细血管膜,最终引起肺水肿^[15],与COVID-19严重程度相关。另外,COVID-19患者体内ACE2会因SARS-CoV-2感染而消耗,导致des-Arg(9)-缓激肽的水平上调,从而加重肺损伤和炎症。Zuo等^[16]发现COVID-19患者MPO水平升高,Ueland等^[17]推测COVID-19患者的ICU重症死亡率与血浆MPO水平相关,且MPO可以通过多途径激活细胞因子并加重组织氧化损伤。

3.3 COVID-19及其主要症状涉及的通路

针对COVID-19及其主要症状的关键基因进行KEGG信号通路富集分析发现,通路主要与氧化应激、免疫调节、肿瘤和炎症反应等相关。其中有7个核心基因参与的典型代表IL-17信号通路,参与了中性粒细胞的浸润及肺部炎症反应,且与SARS-CoV-2病毒进入人体细胞的机制相关,且通过干预IL-17A通路能缓解炎症反应从而保护肺泡细胞。ACE2能够通过下调STAT3通路限制IL-17信号通路的传导,从而缓解中性粒细胞浸润和肺部炎症^[18]。TNF是炎症反应的关键调节因子,其受体TNFR1和TNFR2能激活复杂的信号通路致使血管内皮细胞产生一系列的炎症反应,包括白细胞的黏附、血管的渗漏、血栓形成等^[19];TNF信号通路中的一些促炎因子已被证实参与COVID-19患者的免疫反应过程^[20]。Toll信号通路参与人体的免疫调节,可通过信号转导通路诱导许多反应快速的信号通路,通过产生共刺激分子、趋化因子等参与机体防御反应^[21]。另外,AGE-RAGE信号通路同样属于与靶点相关性较高的15条通路之一,AGE-RAGE信号传导途径与炎症密切相关,该

通路可激活 MAPK 和 NF- κ B 途径,干预免疫和氧化应激反应^[22]。进一步表明 COVID-19 的病理机制与炎症反应、免疫调节、氧化应激等关系密切,以上信号通路以及相关的其他靶点可能做为缓解细胞因子风暴、减轻肺泡组织炎症以及抗病毒治疗的潜在治疗靶点,值得进一步研究。

3.4 潜在药物筛选

3.4.1 化学药筛选 本研究针对 COVID-19 相关靶点基因进行潜在中西医药物筛选,并统计出交互评分排名靠前的化学药物 28 种,包括抗肿瘤药物(15 种)、免疫调节剂(9 种)、抗甲状腺药物(2 种)、抗生素(1 种)、抗疟药(1 种)。其中羟氯喹是一种具有抗炎特性的抗疟药物,具有广谱的抗病毒活性,因其对 SARS-CoV-2 表现出较强的体外抗病毒作用而成为首批进入临床研究的药物之一^[23]。羟氯喹可抑制巨噬细胞系中 TNF 的产生,进而影响 IL-1 β 水平起到减少免疫因子的释放的作用,还可通过抗炎、抑制自噬缓解急性肺损伤^[24]。多项临床研究已证实了羟氯喹对于 COVID-19 的疗效,包括缓解临床症状、提高病毒清除率以及减少重症率^[25]等。但其有效性与安全性仍存在一定争议,1 项纳入 8 721 例受试者的 Meta 分析结果表明羟氯喹对于 COVID-19 患者死亡率并无影响^[26]。

本研究结果中共有 15 种化学药物有已注册的完结或者正在进行的 COVID-19 相关研究,可能通过影响细胞因子的产生、活化,调节免疫功能以及抑制病毒复制等途径影响 COVID-19 的病理过程。例如沙利度胺能通过下调 NF- κ B 的活性而抑制炎症细胞释放的细胞因子和趋化因子,达到强效的抗炎、免疫调节作用^[27-28]。有临床案例报道,沙利度胺联合低剂量糖皮质激素和抗病毒治疗能有效改善 COVID-19 患者的缺氧状态^[29]。来氟米特的主要活性成分特立氟胺与 nCoV-RdRp 酶在分子对接结果中表现出强抑制性亲和力^[30],而 RdRp 酶在冠状病毒基因组复制和转录中起关键作用^[31],nCoV-RdRp 酶的性能和活性的有效抑制将显著阻碍 SARS-CoV-2 的复制,且有小样本临床试验结果表明来氟米特可能会提高难治性 COVID-19 患者的病毒清除率^[32],提示来氟米特是一种潜在的抗 COVID-19 药物。

3.4.2 潜在中药功效及药物筛选 分析现有的 COVID-19 中医诊疗方案发现,各地区的 COVID-19 中医诊疗方案主要围绕 COVID-19 “寒湿疫毒”的病邪特点,遵循由表入里、由浅入深的发病过程进行

辨治^[33]。COVID-19 临床观察期多为表证,当以解表为先;临床治疗期初期以湿邪郁肺为主,枢机不利,治以化湿解毒、宣肺透邪之法。中期多为湿热壅肺,治以宣肺通腑、清热解毒;在恢复期主要为气阴两伤,以补气养阴为主。本研究利用核心基因筛选所得排名前 10 的中药功效为清热解毒(14.38%)、活血祛瘀(6.95%)、化痰(6.05%)、利水渗湿(6.05%)、理气药(5.93%)、祛风湿散寒药(4.44%)、辛温解表药(4.22%)、止咳平喘(4.11%)、补气药(3.73%)和辛凉解表(3.65%),治疗方向切合 COVID-19 临床观察期、治疗初期、中期的病机特点。

筛选所得高频中药中,解表类代表药麻黄在中药治疗中 COVID-19 已得到广泛应用,使用频次高^[34],其作为君药的“麻黄类方”如麻黄汤、麻黄加术汤、麻杏石甘汤、小青龙汤等具有宣肺透邪、散寒除湿之功,直击 COVID-19 之“寒湿”核心病机;祛湿类代表药物半夏、藿香、虎杖、玉米须等也在被临床应用证实疗效可靠^[35-37],藿香与半夏配伍既可行气燥湿又能理肺醒脾,是 COVID-19 诊疗方案中常见的核心药对^[35];补气类代表药物黄芪归脾、肺经,为补气要药,普遍应用于 COVID-19 的中医药预防方案中^[38],可通过扶助正气起到未病先防、既病防变的作用。另有乌梅^[39]、砂仁^[40]、栀子^[41]等也已被应用于 COVID-19 临床防治中,而马齿苋、土茯苓、金荞麦等具有清热解毒作用的中药虽尚未有 COVID-19 相关的临床报道,但既往研究发现清热解毒类中药在抗病毒、抗炎、调节机体免疫方面具有优势,已被列为中医药治疗 COVID-19 的潜在药物^[42]。本研究筛选所得其他尚未有 COVID-19 相关临床报道的潜在中药也可能通过作用相关靶点及通路影响炎症反应、免疫调节以及氧化应激反应等延缓疾病进程,并缓解相关症状,可能成为治疗 COVID-19 的有效药物。

本研究利用文本挖掘技术和生物医学数据库对 COVID-19 主要症状相关的分子靶点和病理机制进行了探究,并筛选了多种潜在中西医治疗药物。研究发现 IL2、IL1B、CCL2、CXCL8、TNF- α 、VEGFA 等核心基因可能通过介导细胞因子与细胞因子受体的相互作用、类风湿关节炎、IL-17 信号通路、恰加斯病、疟疾、NOD 样受体信号通路、糖尿病并发症中的 AGE-RAGE 信号通路、TNF 信号通路、T 细胞受体信号通路和 Toll 样受体信号通路等信号通路参与 COVID-19 病理过程。经查阅相关文献,本研究部分筛选结果已被 COVID-19 相关的临床基础研究证实,在 COVID-19 防治方案中得到应用。

希望通过本研究为COVID-19的分子机制研究、临床治疗方案提供新的思路和方向,帮助扩充COVID-19治疗的分子、靶点、通路和药物选择。但本研究基于文献及数据库资料进行计算和筛选,在一定程度上具有局限性,筛选内容的有效性和正确性有待后续基础研究及临床试验进一步验证。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] Zu Z Y, Jiang M D, Xu P P, et al. Coronavirus disease 2019 (COVID-19): A perspective from China [J]. *Radiology*, 2020, 296(2): E15-E25.
- [2] Pascarella G, Strumia A, Piliago C, et al. COVID-19 diagnosis and management: A comprehensive review [J]. *J Intern Med*, 2020, 288(2): 192-206.
- [3] Huang C L, Wang Y M, Li X W, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China [J]. *Lancet*, 2020, 395(10223): 497-506.
- [4] Guo Y R, Cao Q D, Hong Z S, et al. The origin, transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak-an update on the status [J]. *Mil Med Res*, 2020, 7(1): 11.
- [5] Abd El-Aziz T M, Stockand J D. Recent progress and challenges in drug development against COVID-19 coronavirus (SARS-CoV-2)-an update on the status [J]. *Infect Genet Evol*, 2020, 83: 104327.
- [6] Luo X H, Zhu Y, Mao J, et al. T cell immunobiology and cytokine storm of COVID-19 [J]. *Scand J Immunol*, 2021, 93(3): e12989.
- [7] Basheer M, Saad E, Hagai R, et al. Clinical predictors of mortality and critical illness in patients with COVID-19 pneumonia [J]. *Metabolites*, 2021, 11(10): 679.
- [8] Chen G, Wu D, Guo W, et al. Clinical and immunological features of severe and moderate coronavirus disease 2019 [J]. *J Clin Invest*, 2020, 130(5): 2620-2629.
- [9] Paces J, Strizova Z, Smrz D, et al. COVID-19 and the immune system [J]. *Physiol Res*, 2020, 69(3): 379-388.
- [10] Yilla M, Harcourt B H, Hickman C J, et al. SARS-coronavirus replication in human peripheral monocytes/macrophages [J]. *Virus Res*, 2005, 107(1): 93-101.
- [11] Ziegler T, Matikainen S, Rönkkö E, et al. Severe acute respiratory syndrome coronavirus fails to activate cytokine-mediated innate immune responses in cultured human monocyte-derived dendritic cells [J]. *J Virol*, 2005, 79(21): 13800-13805.
- [12] Shao S, Zhao Z L, Wang F, et al. Risk factors associated with disease aggravation among 126 hospitalized patients with COVID-19 in different places in China: A retrospective observational study [J]. *Medicine (Baltimore)*, 2020, 99(45): e22971.
- [13] Peter A E, Sandeep B V, Rao B G, et al. Calming the storm: Natural immunosuppressants as adjuvants to target the cytokine storm in COVID-19 [J]. *Front Pharmacol*, 2020, 11: 583777.
- [14] Rovas A, Osiaevi I, Buscher K, et al. Microvascular dysfunction in COVID-19: The MYSTIC study [J]. *Angiogenesis*, 2021, 24(1): 145-157.
- [15] Ma Y N, Su K N, Pfau D, et al. Cardiomyocyte d-dopachrome tautomerase protects against heart failure [J]. *JCI Insight*, 2019, 4(17): 128900.
- [16] Zuo Y, Yalavarthi S, Shi H, et al. Neutrophil extracellular traps (NETs) as markers of disease severity in COVID-19 [J]. *medRxiv*, 2020: 2020.04.09.20059626.
- [17] Ueland T, Heggelund L, Lind A, et al. Elevated plasma sTIM-3 levels in patients with severe COVID-19 [J]. *J Allergy Clin Immunol*, 2021, 147(1): 92-98.
- [18] 梁宇鹏, 徐嘉敏, 刘培中, 等. 基于数据挖掘探究新型冠状病毒肺炎中医诊疗方案用药特点 [J]. *陕西中医*, 2020, 41(5): 567-570.
Liang Y P, Xu J M, Liu P Z, et al. Exploring the characteristics of New Coronavirus pneumonia diagnosis and treatment based on Data Mining [J]. *Shaanxi J Tradit Chin Med*, 2020, 41(5): 567-570.
- [19] Bradley J R. TNF-mediated inflammatory disease [J]. *J Pathol*, 2008, 214(2): 149-160.
- [20] Robinson P C, Liew D F L, Liew J W, et al. The potential for repurposing anti-TNF as a therapy for the treatment of COVID-19 [J]. *Med (NY)*, 2020, 1(1): 90-102.
- [21] Majewska M, Szczepanik M. The role of Toll-like receptors (TLR) in innate and adaptive immune responses and their function in immune response regulation [J]. *Postepy Hig Med Dosw (Online)*, 2006, 60: 52-63.
- [22] Medeiros M C, Frasnelli S C, Bastos A d e S, et al. Modulation of cell proliferation, survival and gene expression by RAGE and TLR signaling in cells of the innate and adaptive immune response: Role of p38 MAPK and NF-KB [J]. *J Appl Oral Sci*, 2014, 22(3): 185-193.
- [23] Yao X T, Ye F, Zhang M, et al. *In vitro* antiviral activity and projection of optimized dosing design of hydroxychloroquine for the treatment of severe acute

- respiratory syndrome coronavirus 2 (SARS-CoV-2) [J]. Clin Infect Dis, 2020, 71(15): 732-739.
- [24] Uzunova K, Filipova E, Pavlova V, et al. Insights into antiviral mechanisms of remdesivir, lopinavir/ritonavir and chloroquine/hydroxychloroquine affecting the new SARS-CoV-2 [J]. Biomed Pharmacother, 2020, 131: 110668.
- [25] Das S, Ramachandran A K, Birangal S R, et al. The controversial therapeutic journey of chloroquine and hydroxychloroquine in the battle against SARS-CoV-2: A comprehensive review [J]. Med Drug Discov, 2021, 10: 100085.
- [26] Crichton M L, Goeminne P C, Tuand K, et al. The impact of therapeutics on mortality in hospitalised patients with COVID-19: Systematic review and Meta-analyses informing the European Respiratory Society living guideline [J]. Eur Respir Rev, 2021, 30(162): 210171.
- [27] Khalil A, Kamar A, Nemer G. Thalidomide-revisited: Are COVID-19 patients going to be the latest victims of yet another theoretical drug-repurposing? [J]. Front Immunol, 2020, 11: 1248.
- [28] Zhu H Y, Shi X L, Ju D W, et al. Anti-inflammatory effect of thalidomide on H1N1 influenza virus-induced pulmonary injury in mice [J]. Inflammation, 2014, 37(6): 2091-2098.
- [29] Chen C S, Qi F, Shi K Q, et al. Thalidomide combined with low-dose short-term glucocorticoid in the treatment of critical coronavirus disease 2019 [J]. Clin Transl Med, 2020, 10(2): e35.
- [30] Rabie A M. Teriflunomide: A possible effective drug for the comprehensive treatment of COVID-19 [J]. Curr Res Pharmacol Drug Discov, 2021, 2: 100055.
- [31] Rabie A M. Discovery of (E)-N-(4-cyanobenzylidene)-6-fluoro-3-hydroxypyrazine-2-carboxamide (cyanorona-20): The first potent and specific anti-COVID-19 drug [J]. Chem Pap, 2021, 75(9): 4669-4685.
- [32] Wang Q, Guo H P, Li Y, et al. Efficacy and safety of leflunomide for refractory COVID-19: A pilot study [J]. Front Pharmacol, 2021, 12: 581833.
- [33] 杨珺涵, 孔德昭, 段佳芥, 等. 新型冠状病毒肺炎中医现代诊疗方案证治分析 [J]. 辽宁中医药大学学报, 2021, 23(5): 197-202.
- Yang J H, Kong D Z, Duan J Q, et al. Analysis of syndromes and treatments of *Corona* virus disease 2019 in traditional Chinese medicine [J]. J Liaoning Univ Tradit Chin Med, 2021, 23(5): 197-202.
- [34] 杨灿, 吕晓东, 庞立健, 等. 中药复方治疗新型冠状病毒肺炎用药规律分析 [J]. 海南医学院学报, 2020, 26(13): 961-966.
- Yang C, Lv X D, Pang L J, et al. Analysis of novel coronavirus pneumonia treatment with Chinese herbal compound [J]. J Hainan Med Univ, 2020, 26(13): 961-966.
- [35] 滕龙飞, 陈佳丽, 王 镓, 等. 基于网络药理学及分子对接探讨半夏-藜香抗新型冠状病毒肺炎作用机制 [J/OL]. 世界中医药. (2021-01-27). [2021-07-31]. <https://kns.cnki.net/kcms/detail/11.5529.r.20210126.1330.004.html>.
- Teng L F, Chen J L, Wang J, et al. To explore the mechanism of pinella-ageratum against covid-19 based on network pharmacology and molecular docking [J]. World Chin Med (2021-01-27). [2021-07-31]. <https://kns.cnki.net/kcms/detail/11.5529.r.20210126.1330.004.html>.
- [36] 杨 磊, 刘佳琦, 边 原, 等. 中药活性成分大黄素用于治疗新型冠状病毒肺炎的可能性 [J]. 今日药学, 2020, 30(4): 229-234.
- Yang L, Liu J Q, Bian Y, et al. The possibility of using emodin from Chinese herbal medicine as treatment for COVID-19 [J]. Pharm Today, 2020, 30(4): 229-234.
- [37] 江国庆, 方 鸣, 李 海, 等. 基于历代新安医家抗疫经验指导新型冠状病毒肺炎防治实录 [J]. 中医药临床杂志, 2020, 32(5): 811-821.
- Jiang G Q, Fang M, Li H, et al. Guidance on the prevention and treatment of COVID-2019 based on the past experience of Xin'an physicians [J]. Clin J Tradit Chin Med, 2020, 32(5): 811-821.
- [38] 徐 旭, 张 莹, 李 新, 等. 各地区中医药预防新型冠状病毒肺炎(COVID-19)方案分析 [J]. 中草药, 2020, 51(4): 866-872.
- Xu X, Zhang Y, Li X, et al. Analysis on prevention plan of coronavirus disease 2019 (COVID-19) by traditional Chinese medicine in various regions [J]. Chin Tradit Herb Drugs, 2020, 51(4): 866-872.
- [39] 朱 坤, 申毅锋, 甘文吉, 等. 新型冠状病毒肺炎中医治疗方用药规律分析 [J]. 辽宁中医杂志, 2021, 48(6): 67-72.
- Zhu K, Shen Y F, Gan W J, et al. Analysis on drug administration of *Corona* virus disease 2019 in traditional Chinese medicine [J]. Liaoning J Tradit Chin Med, 2021, 48(6): 67-72.
- [40] 张鹏葛, 王一瑞, 舒占钧, 等. 基于数据挖掘分析治疗新型冠状病毒肺炎不同阶段中药处方用药规律 [J]. 中国医药, 2020, 15(8): 1164-1169.

- Zhang P G, Wang Y R, Shu Z J, et al. Analyze the rule of prescription of traditional Chinese medicine in *Corona* virus disease-19 of different stage: Based on data mining [J]. *China Med*, 2020, 15(8): 1164-1169.
- [41] 杨 炆, 刘 强. 新冠疫情下"火郁发之"在心系疾病诊治中的运用体会 [J]. *浙江中医药大学学报*, 2020, 44(10): 973-976, 985.
- Yang Y, Liu Q. Treatment of heart diseases based on the theory of "releasing stagnated fire" under the COVID-2019 epidemic situation [J]. *J Zhejiang Chin Med Univ*, 2020, 44(10): 973-976, 985.
- [42] 杨安辉, 刘宇灵, 林龙飞, 等. 清热解毒类中药抗新型冠状病毒肺炎研究进展 [J]. *中华中医药学刊*, 2021, 39(1): 181-186.
- Yang A H, Liu Y L, Lin L F, et al. Research progress of heat-clearing and detoxifying traditional Chinese medicine in treatment of *Corona* virus disease 2019 [J]. *Chin Arch Tradit Chin Med*, 2021, 39(1): 181-186.

[责任编辑 刘东博]