

基于 TCMSP 数据库中药成分的新型冠状病毒 3CL 水解酶抑制剂虚拟筛选

刘妍如¹, 唐志书^{1*}, 王 梅², 段金廛^{3*}, 宋忠兴¹, 周 瑞¹, 王 征¹, 李 宏⁴, 李晓红¹, 江大海¹

1. 陕西中医药大学 陕西中药资源产业化省部共建协同创新中心, 秦药特色资源研究与开发国家重点实验室(培育), 陕西省创新药物研究中心, 陕西 咸阳 712083
2. 陕西中医药大学附属医院, 陕西 咸阳 712083
3. 南京中医药大学 江苏省方剂高新技术研究重点实验室, 江苏省中药资源产业化过程协同创新中心, 江苏 南京 210023
4. 陕西中医药大学制药厂, 陕西 咸阳 712046

摘要: 目的 运用分子对接技术模拟预测 TCMSP 数据库中药成分对新型冠状病毒(SARS-CoV-2) 3CL 水解酶的作用。方法 基于 SARS-CoV-2 3CL 水解酶的 PDB 晶型结构(6LU7), 构建虚拟靶标模型, 以 TCMSP 数据库中的 13 143 个化合物为配体筛选对象, 以 6LU7 原配体分子为对照, 采用薛定谔 2018 软件, 以 Glide 程序建立 6LU7 分子对接体系。然后以原配体分子的对接打分结果为阈值, 口服生物利用度(OB) ≥ 30% 或类药性(DL) ≥ 0.18 为标准筛选候选化合物。最后建立“成分-药材-归经-功效”网络, 对候选成分的主要药材归属和作用功效进行总结。结果 按照对接打分、OB、DL 筛选标准, 共筛选得到 60 个化合物。受体-配体相互作用结果显示, 候选化合物主要与靶点蛋白的 GLU166、GLY143、ASP187、CYS145、GLN189 和 LEU141 等发生氢键作用。网络预测结果显示, 所筛选化合物主要归属药材为甘草、桑白皮、满山红、虎杖和车前草等, 肺脏是作用的关键靶器官, 其作用功效以清热解毒、止咳祛痰、泻肺平喘等为主。结论 对 TCMSP 数据库化合物进行虚拟分子对接得到的结果为 SARS-CoV-2 3CL 水解酶抑制剂筛选提供了相关数据。

关键词: 新型冠状病毒; 新型冠状病毒肺炎; 3CL 水解酶; TCMSP 数据库; 分子对接

中图分类号: R285.5 **文献标志码:** A **文章编号:** 0253-2670(2020)07-1694-10

DOI: 10.7501/j.issn.0253-2670.2020.07.003

Potential SARS-CoV-2 3CL protease inhibitors selection from TCMSP platform by molecular docking approach

LIU Yan-ru¹, TANG Zhi-shu¹, WANG Mei², DUAN Jin-ao³, SONG Zhong-xing¹, ZHOU Rui¹, WANG Zheng¹, LI Hong⁴, LI Xiao-hong¹, JIANG Da-hai¹

1. Shaanxi Collaborative Innovation Center of Chinese Medicinal Resources Industrialization, State Key Laboratory of Research and Development of Characteristic Resources of Qin Medicine Research, Shaanxi Institute of Chinese Medicine Industry Research, Shaanxi University of Chinese Medicine, Xianyang 712083, China
2. The Affiliated Hospital of Shaanxi University of Chinese Medicine, Xianyang 712083, China
3. Jiangsu Key Laboratory for Traditional Chinese Medical Formula Research, Collaborative Innovation Center of Chinese Medicinal Resources Industrialization of Jiangsu Province, Nanjing University of Chinese Medicine, Nanjing 210023, China
4. Pharmaceutical Factory of Shaanxi University of Traditional Chinese Medicine, Xianyang 712046, China

Abstract: Objective To identify potential SARS-CoV-2 3CL protease inhibitors from the traditional Chinese medicine systems pharmacology database and analysis platform (TCMSP) by molecular docking approach. **Methods** To alternate extensive compounds experimental screening processes, a Computer-Aided Drug Design (CADD) based molecular docking technology was performed to explore existing drug repurposing possibilities. Molecular docking model with Schrodinger suit 2018 was used to evaluate the binding

收稿日期: 2020-02-18

基金项目: 2020 年度咸阳市重点研发计划“新型冠状病毒肺炎疫情应急防治”科技专项; 陕西省三秦学者创新团队; 陕西省“特支计划”青年拔尖人才项目(陕组通字[2018]33 号)

作者简介: 刘妍如(1985—), 女, 博士, 硕士生导师, 主要从事药物分析研究。Tel: (029)38182207 E-mail: yanzi_2203@aliyun.com

***通信作者** 唐志书(1972—), 男, 教授, 博士生导师, 主要从事中药制剂制备技术与质量分析研究。

Tel: (029)38185060 E-mail: tzs6565@163.com

段金廛(1956—), 男, 教授, 博士生导师, 中国自然资源学会中药及天然药物资源专业委员会主任委员, 国家“973”计划首席科学家。Tel: (025)85811291 E-mail: dja@njucm.edu.cn

abilities between TCMSP 13 143 compounds and SARS-CoV-2 3CL protease receptor-binding domain (PBD ID 6LU7), which involving in mediating viral replication and transcription functions. According to the constructed docking system, potential compounds were screened according to docking score, oral bioavailability (OB), and drug-likeness (DL). At last, a compounds-herb-target organ-function network was constructed. **Results** Compared with 6LU7 original ligand docking score (-7.734), a total of 498 compounds were identified with lower docking score against 6LU7 targets. These compounds were further reduced to 60 high-priority compounds, based on OB (more than 30) and DL (more than 0.18). Meanwhile, these 60 compounds were found to interact with the amino acid residues (GLU166, GLY143, ASP187, CYS145, GLN189, LEU141, etc.) which were critically involved in the 6LU7 domain mainly by hydrogen-bonded interaction. The network exploring results revealed that these potential compounds were mainly attributed to *Glycyrrhizae Radix et Rhizoma*, *Mori Cortex*, *Rhododendron dauricum*, *Polygoni Cuspidati Rhizoma et Radix*, and *Plantaginis Herba*, etc., which associates with acute lung syndromes induced by SARS-CoV-2, with the effect of clearing heat and removing toxin, relieving cough and dispelling phlegm and lung-draining and relieving asthma. **Conclusion** Molecular docking method provides a useful tool for the screening of SARS-CoV-2 3CL protease inhibitors from TCMSP platform.

Key words: SARS-CoV-2; COVID-19; 3CL protease; TCMSP platform; molecular docking

2019 年 12 月开始, 一种不明原因的急性肺炎在中国武汉爆发, 该病是由新型冠状病毒 (SARS-CoV-2) 感染所致的一种急性肺部传染病, 已从武汉迅速向中国其他城市扩散。由于该次疫情传播的严重性, 世界卫生组织 (WHO) 已于 2020 年 1 月 30 日将此疫情宣布为“国际关注的突发公共卫生事件”, 并将此病命名为新型冠状病毒肺炎 (COVID-19)。截止 2020 年 2 月 17 日, 中国各城市已报告 COVID-19 确诊病例 70 641 例, 死亡 1 772 例, 治愈率 13.74%, 死亡率 2.16%。目前, 中国以外的其他多个国家亦已报告确诊病例, 而且疫情越发严重。SARS-CoV-2 于 2020 年 1 月 24 日被中国疾控中心从临床样本中首次分离, 为正链单链 RNA, 在基因组序列上与 2003 年 SARS 病毒有大约 80% 相似性, 与 Bat SARS-like coronavirus isolate bat-SL-CoVZC45 基因组序列相似度为 88%^[1-3]。目前推测 SARS-CoV-2 可能通过呼吸道入侵肺部, 攻击血管内皮细胞、上皮细胞和包括巨噬细胞在内的多种免疫细胞, 来抑制免疫应答反应, 从而诱发严重的肺炎症状, 很多治疗药物和疫苗也还处于研发或临床试验阶段^[4-10], 因此, 迫切需要研发安全性高、耐药性好的抗病毒药物。

由于 SARS-CoV-2 基因组序列与 SARS 病毒的高度相似性, 其功能蛋白如覆膜蛋白、棘突蛋白、3C 样水解酶 (3CL 水解酶, 也称主蛋白酶, Mpro) 等蛋白从理论上可作为抗 SARS-CoV-2 病毒设计和筛选的主要靶标。其中, 3CL 水解酶作为一类半胱氨酸水解酶, 序列高度保守, 其主要作用是水解冠状病毒复制过程中编码的多聚蛋白 ppla 和 pplab 来产生功能蛋白, 并且人体中尚未发现与其相似的

水解酶^[11-13]。因此, 3CL 水解酶可作为抗 SARS-CoV-2 靶标筛选的首选蛋白。为了在短时间内迅速筛选出潜在的抗病毒化合物, 本研究采用计算机辅助药物设计中的分子对接技术, 在选定虚拟靶点的前提下, 根据几何互补、能量互补和化学环境互补的原则, 通过分析虚拟靶点的疏水“口袋”结构, 搜索受体-配体结合的优势构象, 评价对接分子的结合强度, 从 TCMSP 数据库中筛选出潜在的先导化合物。

1 材料与方法

1.1 靶点选择与处理

2020 年 1 月 26 日上海科技大学饶子和/杨海涛课题组将 SARS-CoV-2 3CL 水解酶的高分辨率晶体结构公布于 PDB 数据库 (PDB ID: 6LU7)。本实验将 6LU7 的复合物数据 (原始配体分子: *N*-[(5-methylisoxazol-3-yl) carbonyl] alanyl-l-valyl-N~1~-((1*R*,2*Z*)-4-(benzyloxy)-4-oxo-1-{[(3*R*)-2-oxopyrrolidin-3-yl] methyl} but-2-enyl)-l-leucinamide) 用于分子对接位点分析。首先用 Schrodinger-2018 软件的 Protein preparation wizard 模块对 6LU7 结构进行修正化学键序, 加氢, 处理二硫键、金属离子, 除去晶体结构中水分子、杂原子等预处理, 然后在 OPLS3 力场条件下对蛋白进行能量优化, 晶体结构重原子 RMSD 收敛标准为 0.05 nm (1 Å=0.1 nm)。

1.2 配体优化

以 6LU7 原配体化合物为参照, 将中药系统药理学数据库与分析平台 (traditional Chinese medicine systems pharmacology database and analysis platform, TCMSP, <http://www.tcmspw.com>) 中的

13 143 个化合物作为 SARS-CoV-2 3CL 水解酶抑制剂筛选的候选分子^[14-17]。采用 Ligprep 模块对各化合物分子进行加氢等优化预处理。在力场 OPLS3 下, 计算 pH 为 7.0±2.0 范围的化合物离子状态, 搜索合理构象作为对接起始构象, 形成可能的互变异构体, 并生成批量处理文件 (*.sd)。

1.3 格点文件生成

采用 Receptor Grid Generation 模块, 盒子大小设为 14 Å, 以原始配体分子为中心生成格点文件。

1.4 分子对接

以 6LU7 作为刚性受体, 各配体分子作为构象可变化(柔性)配体, 进行半柔性对接。采用 Ligand Docking 模块进行虚拟对接, 将格点文件导入 Receptor grid file 中, 对所有化合物进行分子对接, 其他参数为默认值。由于受体-配体对接的打分值 docking score 综合了配体能量、受体能量及两者间的结合能等参数, 分值绝对值越大说明受体亲和力越强, 两者结合越牢固, 因此以打分值小于原配体分子作为初步筛选条件。

化合物成为临床研究候选药物的关键因素在于

其是否具有良好的药理学特征和较低的毒副作用。因此, 以口服生物利用度 (OB) ≥30% 或类药性 (DL) ≥0.18 作为卡值, 对化合物进行进一步筛选。最后选择对接得分绝对值高于原配体分子, 且符合 OB ≥30% 或 DL ≥0.18 标准的化合物作为候选化合物。

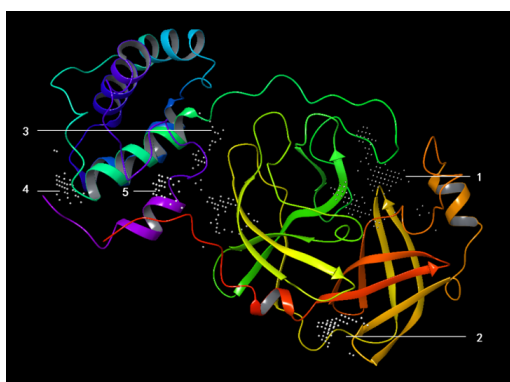
1.5 “化合物-药材-归经-功效”预测网络构建

收集候选化合物对应药材、归经和功效等信息, 将数据导入 Cytoscape 3.7.2 软件, 构建“化合物-药材-归经-功效”预测网络。以网络的节点和边来描述化合物与药材、归经和功效之间的关系, 通过简化网络体系复杂的相互作用, 搜索关键药材, 并对靶器官和作用功效进行预测。

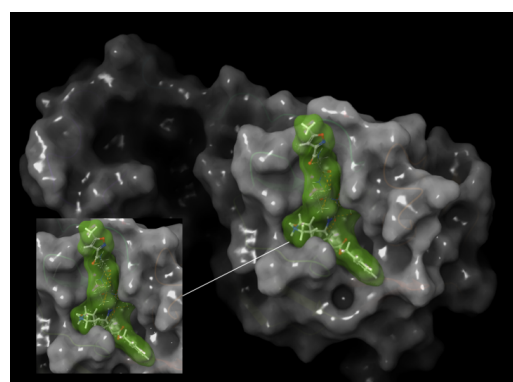
2 结果

2.1 6LU7 结合位点确证

删除 6LU7 原配体分子, 以 Binding Site 搜索受体优势结合位点。计算生成 5 个可能的结合位点, 如图 1 所示, 其中结合位点 1 空腔较大, 且与原配体叠加, 说明该对接口袋与配体结合位点的一致性, 可用于下一步的对接分析。



SARS-CoV-2 3CL 水解酶受体位点预测图 (白色点阵: 1~5 号结合位点)



6LU7-原配体复合物结构

图 1 SARS-CoV-2 3CL 水解酶受体位点及复合物结构 3D 图

Fig. 1 SARS-CoV-2 3CL protease sitemap and crystal structure

2.2 对接体系验证

为了初步确定对接方法的准确度, 将对接前后的 6LU7 原始配体进行构象叠合, 叠加结果显示均方根偏差 (RMSD) 值为 1.281 3 Å < 2 Å。说明所用对接参数能够适用模拟体系, 所选对接体系可初步用于 TCMSP 数据库化合物与 6LU7 的对接预测。

2.3 6LU7 与 TCMSP 数据库化合物对接结果

以 6LU7 作为受体, 原始配体与受体作用参数为依据, 对 TCMSP 的 13 143 个化合物进行分子对接。以打分值 (docking score) 小于原配体分子

(-7.734) 为筛选条件, 以 OB ≥30% 或 DL ≥0.18 作为进一步筛选条件, 得到符合筛选标准的化合物 60 个 (表 1)。分析受体-配体对接结果显示, 大部分化合物和受体结合以发生氢键作用为主。6LU7 原配体分子为肽类化合物, 其 α,β-不饱和酯键与 CYS145 形成共价键, 与 GLU166 形成 3 个氢键, 与 PHE140、THR190、GLN189、HID164 及 GLY143 分别形成一个氢键。相较于原配体分子, 60 个候选化合物与受体的结合位点可按结构分为以下几类 (图 2)。(1) 黄酮及黄酮苷类: 黄酮及二氢黄酮类

表 1 60 个候选化合物对接、筛选结果及关联药材分析

Table 1 Sixty screened compounds from TCMSP platform as potential inhibitors for SARS-CoV-2 3CL protease

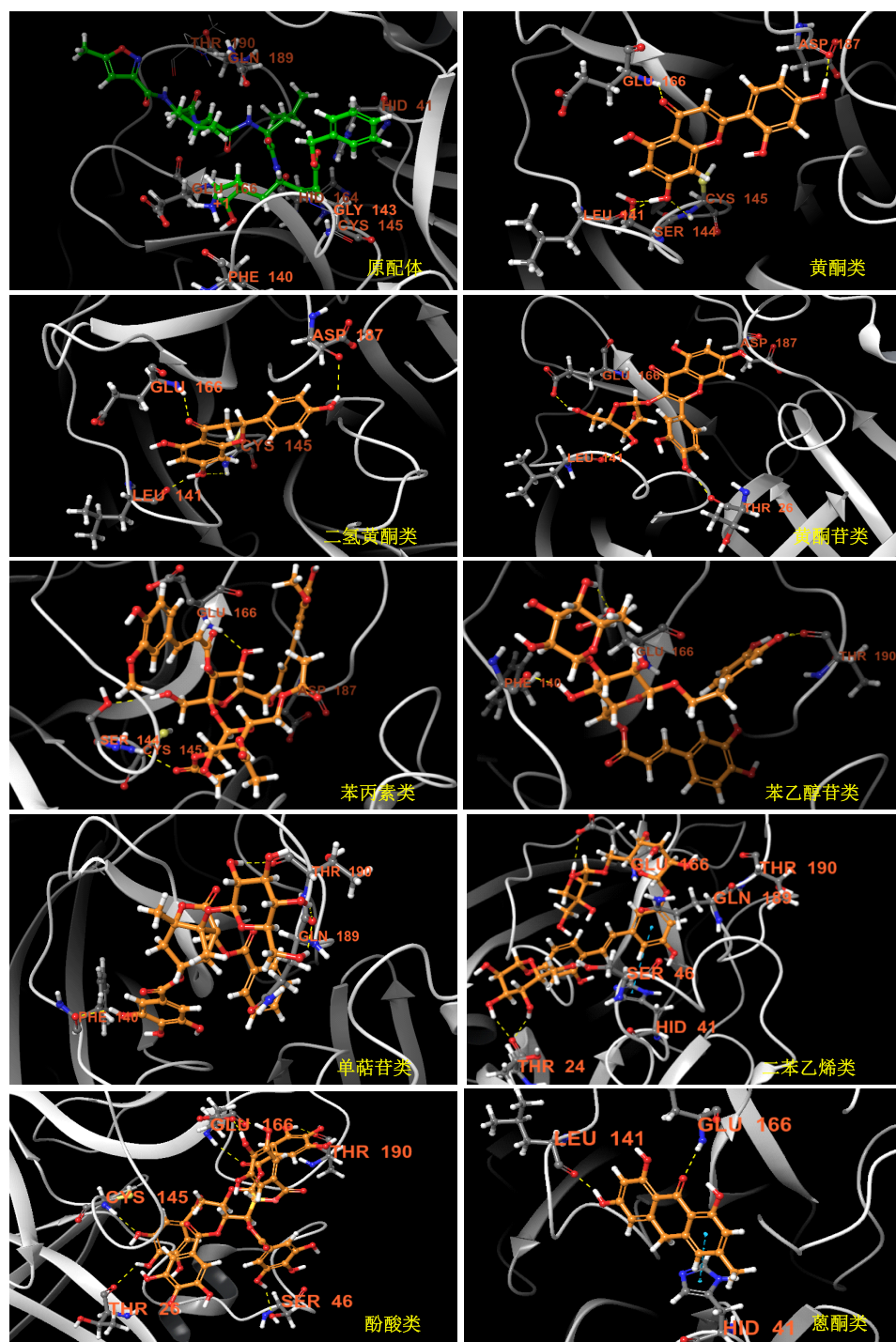
编号	MOL ID	化学成分	类型	相对分子质量	Docking score	OB/%	DL	关联药材	作用位点
1	MOL005860	3-O-没食子酰前矢车菊素 B-3 (3-O-galloylprocyanidin B-3)	黄酮醇	730.67	-10.524	30.06	0.33	地榆	GLU166、HID164、 HID163、THR26
2	MOL012598	山柰酚-3-O- α -L-(2'',4''-二阿魏酸奎尼酸)-鼠李糖苷 [kaempferol-3-O- α -L-(2'',4-di- <i>E</i> -feruloyl)-rhamnoside]	酚酸	784.77	-9.464	3.03	0.31	枇杷叶	GLU166、CYS145、 LEU144
3	MOL012796	山柰酚-3-O-(6'-没食子酰)- β -D-葡萄糖苷 [kaempferol-3-O-(6'-galloyl)- β -D-glucoside]	黄酮醇苷	600.52	-9.263	3.00	0.66	桑白皮	GLU166、ASP187、 SER144、ASN142
4	MOL001519	异千日红苷 II (isogomphrenin II)	β -花青苷	696.67	-9.243	10.49	0.41	千日红	GLU166、GLY143、 SER46
5	MOL012795	山柰酚-3-O-(2',6'-二没食子酰)- β -D-葡萄糖苷 [kaempferol-3-O-(2',6'-digalloyl)- β -D-glucoside]	黄酮醇苷	752.63	-9.028	3.01	0.37	桑白皮	GLU166、HID163、 PHE140、THR26
6	MOL009093	1,2,3,6-四没食子酰葡萄糖 (1,2,3,6-tetragalloylglucose)	酚酸	788.62	-8.814	3.01	0.33	核桃仁	HID164、PRO168、 THR26、CYS145、 ASN142
7	MOL007810	车前草苷 D ([2 <i>R</i> ,3 <i>R</i> ,4 <i>S</i> ,5 <i>R</i> ,6 <i>R</i>)-6-[2-(3,4-dihydroxyphenyl)ethoxy]-3,5-dihydroxy-4-[(2 <i>S</i> ,3 <i>R</i> ,4 <i>S</i> ,5 <i>S</i> ,6 <i>R</i>)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxyoxan-2-yl)methyl(<i>E</i>)-3-(3,4-dihydroxyphenyl)prop-2-enoate)	苯乙醇苷	640.65	-8.767	3.39	0.57	车前草	GLU166、ASN142、 THR190、HID164
8	MOL007931	异鼠李素 3-O-洋槐苷 (iso-rhamnetin-3-O-robinobioside)	黄酮醇苷	624.60	-8.716	5.09	0.65	金钱草	THR190、CYS145、 SER144、GLY143、 THR45
9	MOL002762	山柰酚-3-O- β -D-芸香苷 (kaempferol-3-O-beta-D-rutinoside)	黄酮醇苷	610.57	-8.695	5.51	0.65	红花	GLU166、ASP187、 LEU141
10	MOL007060	紫草酸 B (lithospermic acid B)	酚酸	718.66	-8.682	3.01	0.41	丹参、西洋参、紫草	GLU166、CYS145、 ASN142、THR26、 THR25
11	MOL010961	阿福豆苷 (kaempferol 3-O-rhamnoside)	黄酮醇苷	432.41	-8.581	5.31	0.70	矮地茶、银杏叶	GLU166、ASP187、 THR190
12	MOL012783	槲皮素-3-O-(2',6'-二没食子酰)- β -D-葡萄糖 [quercetin-3-O-(2',6'-digalloyl)- β -D-glucoside]	黄酮醇苷	768.63	-8.485	3.01	0.34	桑白皮	GLU166、THR26
13	MOL008905	槲皮素-3-巢菜糖苷 (β -vicianosyl-3-quercetin)	黄酮醇苷	596.54	-8.481	7.22	0.71	王不留行	HID41、GLU166、 THR190、THR26、 GLY143
14	MOL000367	槲寄生新苷 III (viscumneoside III)	黄酮醇苷	596.59	-8.463	10.88	0.73	槲寄生	GLU166、ARG188、 GLY143、CYS145
15	MOL006114	山柰酚-3-O- α -L-鼠李糖(1 $\rightarrow$ 2)- β -D-葡萄糖苷 [kaempferol-3-O- α -L-rhamnosyl(1 $\rightarrow$ 2)- β -D-glucoside]	黄酮醇苷	446.44	-8.458	27.11	0.74	蒲黄	HID163、GLU166、 ARG188、SER144、 CYS145
16	MOL010261	水仙苷 (isorhamnetin-3-O- β -rutinoside)	黄酮醇苷	624.60	-8.458	5.09	0.65	沙棘	GLU166、THR26、 TYR54、ASN142、 LEU141
17	MOL000738	桑黄酮 C (kuwanon C)	黄酮	422.51	-8.41	1.22	0.58	桑白皮、桑叶、桑枝、 无患子皮	ASP187、LEU141
18	MOL006683	连翘酯苷 B (forsythoside B)	苯乙醇苷	756.78	-8.369	3.30	0.42	密蒙花	GLU166、GLY143、 PRO168

续表 1

编号	MOL ID	化学成分	类型	相对分子质量	Docking score	OB/%	DL	关联药材	作用位点
19	MOL000159	异麦角甾昔 ([[(2R,3R,4S,5R,6R)-6-[2-(3,4-dihydroxyphenyl) ethoxy]-3,5-dihydroxy-4-[(2R,3S,4S,5S,6R)-3,4,5-trihydroxy-6-methyl-tetrahydropyran-2-yl]oxy-tetrahydropyran-2-yl]methyl(E)-3-(3,4-dihydroxyphenyl)prop-2-enoate])	苯乙醇昔	624.65	-8.340	3.05	0.60	车前子	GLU166, THR190, PHE140
20	MOL004328	柚皮素 (naringenin)	二氢黄酮	272.27	-8.321	59.29	0.21	化橘红、菊花、橘红、麻黄、麻黄根、青果	GLU166, LEU141, ASP187, CYS145
21	MOL011869	荊芥素 A (schizotenuin A)	酚酸	716.64	-8.292	3.01	0.39	荊芥	GLU166, THR190, LEU167, PRO168
22	MOL012432	杜鹃素 (farrerol)	二氢黄酮	300.33	-8.280	42.65	0.26	满山红	GLU166, SER144, CYS145, ASP187
23	MOL012444	去甲杜鹃素 (poriol)	二氢黄酮	286.30	-8.269	38.45	0.24	满山红	GLU166, LEU141, ASP187
24	MOL007381	牡丹皮昔 H (mudanpioside H)	单萜昔	618.64	-8.215	6.77	0.61	牡丹皮	GLU166, ASP187, SER144, THR190, PHE140
25	MOL004022	甲基 3,3,6-三-O-没食子酰基-β-D-吡喃葡萄糖 (methyl 3,3,6-tri-O-galloyl-β-D-glucopyranoside)	酚酸	650.54	-8.202	3.01	0.53	吴茱萸	GLU166, GLY143, THR25
26	MOL009773	7-羟基-2-(4-羟苯基)-3,5-二甲氧基色原酮 [7-hydroxy-2-(4-hydroxyphenyl)-3,5-dimethoxy-chromone]	黄酮	314.31	-8.200	22.76	0.29	黄药子	GLU166, ASP187, CYS145, LEU141
27	MOL009092	五没食子酰葡萄糖 (pentagalloylglucose)	酚酸	940.72	-8.182	3.01	0.21	白芍、赤芍、诃子、核桃仁、牡丹皮、余甘子	GLU166, THR26, CYS145, THR190, SER46
28	MOL007979	蕃昔 (avicularin)	黄酮醇昔	434.38	-8.173	2.06	0.70	蒺藜、满山红、三白草、沙苑子、桑白皮	GLU166, HID163, THR26, LEU141
29	MOL002236	3,5-二-O-没食子酰基-4-O-二没食子酰基奎宁酸 (3,5-di-O-galloyl-4-O-digalloylquinic acid)	酚酸	800.63	-8.134	3.01	0.31	大黄	GLU166, GLY143, THR26
30	MOL006115	山柰酚-3-O-α-L-鼠李糖基(1→2)-β-D-葡萄糖苷 (kaempferol-3-O-α-L-rhamnosyl(1→2)-β-D-glucoside)	黄酮醇	284.28	-8.103	62.87	0.24	蒲黄	GLU166, SER144, ASP187
31	MOL001792	甘草素 (DFV)	二氢黄酮	256.27	-8.100	32.76	0.18	板蓝根、甘草、葛花	GLU166, ASP187, CYS145, LEU141
32	MOL007987	洛罗兰糖昔 (loroglossin)	二氢黄酮	742.80	-8.081	3.49	0.35	山慈菇	GLU166, GLY143, SER46
33	MOL004938	夏佛托昔 (schaftoside)	黄酮	596.54	-8.074	7.88	0.75	甘草	HID41, GLU166, LEU141
34	MOL006504	(-)-儿茶素没食子酸酯 [(-)-catechingallate]	酚酸	442.40	-8.065	53.57	0.75	白藜、金荞麦、苦丁茶	GLU166, ASN142, THR26, ASP187
35	MOL008430	5,7-二羟基-2-(4-羟苯基)-3-甲基-色原酮 [5,7-dihydroxy-2-(4-hydroxyphenyl)-3-methyl-chromone]	黄酮	284.28	-8.057	45.05	0.24	儿茶	GLU166, SER144, ASP187
36	MOL000008	芹菜素 (apigenin)	黄酮	270.25	-8.051	23.06	0.21	败酱草、板蓝根、半边莲、半枝莲、薄荷、侧柏叶	HID163, ASP187, SER144
37	MOL006614	苦参醇 A (kushenol A)	二氢黄酮	408.53	-8.018	1.03	0.55	苦参	GLU166, HID164, LEU141

续表 1

编号	MOL ID	化学成分	类型	相对分子质量	Docking score	OB/%	DL	关联药材	作用位点
38	MOL012703	桑黄酮 Q (kuwanon Q)	查耳酮	662.78	-7.981	3.01	0.57	桑白皮	GLU166, THR26
39	MOL007009	4-O-没食子酰白芍苷 (4-O-galloylalbiflorin)	单萜	632.62	-7.973	3.09	0.61	赤芍	HID41, PHE140, GLU189, THR190
40	MOL007245	3-甲基山奈酚 (3-methylkempferol)	黄酮	300.28	-7.966	60.16	0.26	枸骨叶、木棉花、梔子	GLU166, CYS145, LEU141
41	MOL002328	好望角芦荟内酯 (feralolide)	苯丙素	344.34	-7.936	1.37	0.39	防己、芦荟	GLY143, SER144
42	MOL000142	木通苯乙醇苷 B (calceolarioside B)	苯乙醇苷类	478.49	-7.931	1.58	0.70	车前草、车前子	GLU166, HID164, HID163, THR26
43	MOL000155	异类叶升麻苷 (acteoside isomer)	酚酸	624.65	-7.923	3.05	0.61	车前子	GLU166, THR190, PHE140
44	MOL010616	褐藻多酚 (eckol)	酚酸	372.30	-7.910	87.06	0.63	昆布	ASN142, THR190
45	MOL004567	双辣薄荷基厚朴酚 (dipiperitylmagnolol)	黄酮苷	538.88	-7.889	6.60	0.86	独活	GLU166, ARG188, THR190, SER144, HID163
46	MOL012714	桑辛素 A (moracin A)	酚酸	286.30	-7.885	64.39	0.23	桑白皮	LEU141
47	MOL013122	苯丙素糖苷 C (smiglaside C)	苯丙素苷	820.82	-7.875	4.21	0.31	土茯苓	GLU166, SER144, CYS145, ASN142
48	MOL003353	大黄素蒽酮 (emodinanthrone)	蒽酮	256.27	-7.862	24.72	0.21	百合、大黄、虎杖、连翘	HID41, GLU166, LEU141
49	MOL000422	山奈酚 (kaempferol)	黄酮醇	286.25	-7.857	41.88	0.24	矮地茶、八角莲、白果、败酱草、半边莲、蒿蓄、侧柏叶	SER144, GLN189, ASP187
50	MOL005014	甘草苷 A (licorice glycoside A)	糖苷	726.74	-7.843	5.95	0.35	甘草	GLU166, HID41, LEU141
51	MOL004816	(2R)-1-[2,4-二羟基-5-(3-甲基-2-丁烯)苯基]-2-羟基-3-[4-羟基-3-(3-甲基-2-丁烯)苯基]-1-丙酮 [(2R)-1-[2,4-dihydroxy-5-(3-methylbut-2-enyl)phenyl]-2-hydroxy-3-[4-hydroxy-3-(3-methylbut-2-enyl)phenyl]propan-1-one]	酚酸	410.55	-7.826	1.06	0.48	甘草	GLU166, HID41, GLY143, SER144
52	MOL012742	氧化白藜芦醇-4-O-β-D-吡喃葡萄糖基-3'-O-β-D-吡喃葡萄糖基-(1→6)-O-β-D-葡萄糖基 [oxyresveratrol-4-O-β-D-glucopyranosyl-3'-O-β-D-glucopyranosyl-(1→6)-O-β-D-glucopyranoside]	二苯乙烯苷	730.74	-7.822	3.24	0.36	桑白皮	GLU166, THR26, THR190, PHE140
53	MOL000737	桑色素 (morin)	黄酮醇	302.25	-7.814	46.23	0.27	桑椹、桑枝、夏枯草、香薷	SER144, GLN189, ASP187
54	MOL011457	3-甲氧基蜀葵苷元 (3-methoxyherbacetin)	黄酮	316.28	-7.785	23.50	0.30	西洋参	GLU166, ASP187, CYS145, SER144
55	MOL012741	氧化白藜芦醇-4-O-β-D-吡喃葡萄糖基-3'-O-β-D-芹糖基-(1→6)-O-β-D-葡萄糖基 [oxyresveratrol-4-O-β-D-glucopyranosyl-3'-O-β-D-apiofuranosyl-(1→6)-O-β-D-glucopyranoside]	二苯乙烯苷	700.71	-7.784	3.97	0.42	桑白皮	GLU166, SER46, THR24, HID41
56	MOL013305	3,4',7-三羟基黄酮 (garbanzol)	二氢黄酮醇	272.27	-7.772	83.67	0.21	葛花	GLU166, GLN189, SER144, ASP187
57	MOL002933	5,7,4'-三羟基-8-甲氧基黄酮 (5,7,4'-trihydroxy-8-methoxyflavone)	黄酮	300.28	-7.772	36.56	0.27	黄芩、马鞭草	GLU166, LEU141
58	MOL006630	桑叶活性黄酮 (norartocarpetin)	黄酮	286.25	-7.764	54.93	0.24	苦参、桑叶	SER144, ASP187
59	MOL000445	8-异戊二烯基怀特酮 (8-prenylwightone)	异黄酮	406.51	-7.761	23.22	0.54	甘草、海桐皮	HID164, GLY143, THR24
60	MOL012942	鞣质 (tannin)	酚酸	636.51	-7.736	3.01	0.54	大枣	GLU166, ASN142, GLY143, THR24



绿色球棍模型为原配体分子，橙色球棍模型为候选化合物分子；蓝色-氮原子；红色-氧原子

The green ball-and-stick models represent original 6LU7 ligand; orange ones represent potential inhibitors; blue atom means nitrogen atom; red atom means oxygen atom

图 2 6LU7 与不同类型化合物相互作用

Fig. 2 Crystal structures of potential compounds complexed with receptor

化合物的 4 位羰基氧和 GLU166 形成氢键，黄酮类化合物 7 位和 4'位羟基分别和 LEU141、SER144、CYS145、ASP187 形成氢键；二氢黄酮类化合物 7

位和 4'位羟基分别和 LEU141、CYS145、ASP187 形成氢键；黄酮醇类化合物的 7 位和 4'位羟基分别和 LEU141、ASP187 形成氢键；黄酮苷类化合物的

吡喃糖羟基主要和 GLU166 形成氢键, 苷元的 7 位和 4' 位羟基分别和 THR26、LEU141、ASP187 形成氢键。(2) 苯丙素及糖苷类: 主要和 CYS145、GLY143、SER144、ASN142 形成氢键, 苯丙素糖苷类化合物的吡喃糖羟基主要和 GLU166 形成氢键。(3) 苯乙醇苷类: 主要通过化合物苯环的酚羟基与 THR190 结合形成氢键, 吡喃糖环羟基和 GLU166、PHE140 形成氢键。(4) 单萜及单萜糖苷类: 主要和 THR190、PHE140 形成氢键。(5) 二苯乙烯类: 主要通过苯环的酚羟基与 THR190 形成氢键, 和 HID41 形成 π - π 堆积, 吡喃糖羟基和 GLU166 形成氢键。(6) 酚酸类: 主要为苯环的酚羟基和 GLU166、THR190、THR26 和 PRO168 形成氢键。(7) 蒽酮类: 主要是 9 位羰基氧和 GLU166 形成氢键, 6 位羟基和 LEU144 形成氢键, 苯环与 HID41 形成 π - π 堆积。

2.4 “成分-药材-归经-功效”网络构建

收集上述 60 个化合物关联的药材、归经和功效信息, 将其关联数据和节点数据导入 Cytoscape

3.7.2 软件, 构建“成分-药材-归经-功效”关联网络(图 3)。候选化合物、归属药材、归经和功效分别以不同颜色节点表示, 其相互关系以边表示, 每个节点的重要性以度(degree)值表示, 度值越大, 节点越大, 说明该节点在网络中的重要性越大。通过网络分析, 60 个候选化合物分别归属于 67 种药材, 每一个成分至少与一种药材相关, 与其中的 9 个药材有较强的关联作用, 这些药材包括甘草、桑白皮、虎杖、满山红、车前子、车前草、西洋参、败酱草和核桃仁。对“药材-归经-功效”网络特征分析显示, 这些药材主要归为肺经, 其作用功效与清热解毒、止咳祛痰、泻肺平喘、清热利湿、补气养阴重点相关, 表明这些药材在治疗 COVID-19“湿热毒瘀”相关症状方面具有潜在优势。

3 讨论

冠状病毒的变异速度是造成 COVID-19 治愈困难的主要原因, 疫苗、抗病毒药物的研发速度远远跟不上病毒的变异速度。目前, 传统的抗免疫缺陷疾病药物、抗癌药物已投放至临床, 参与抗

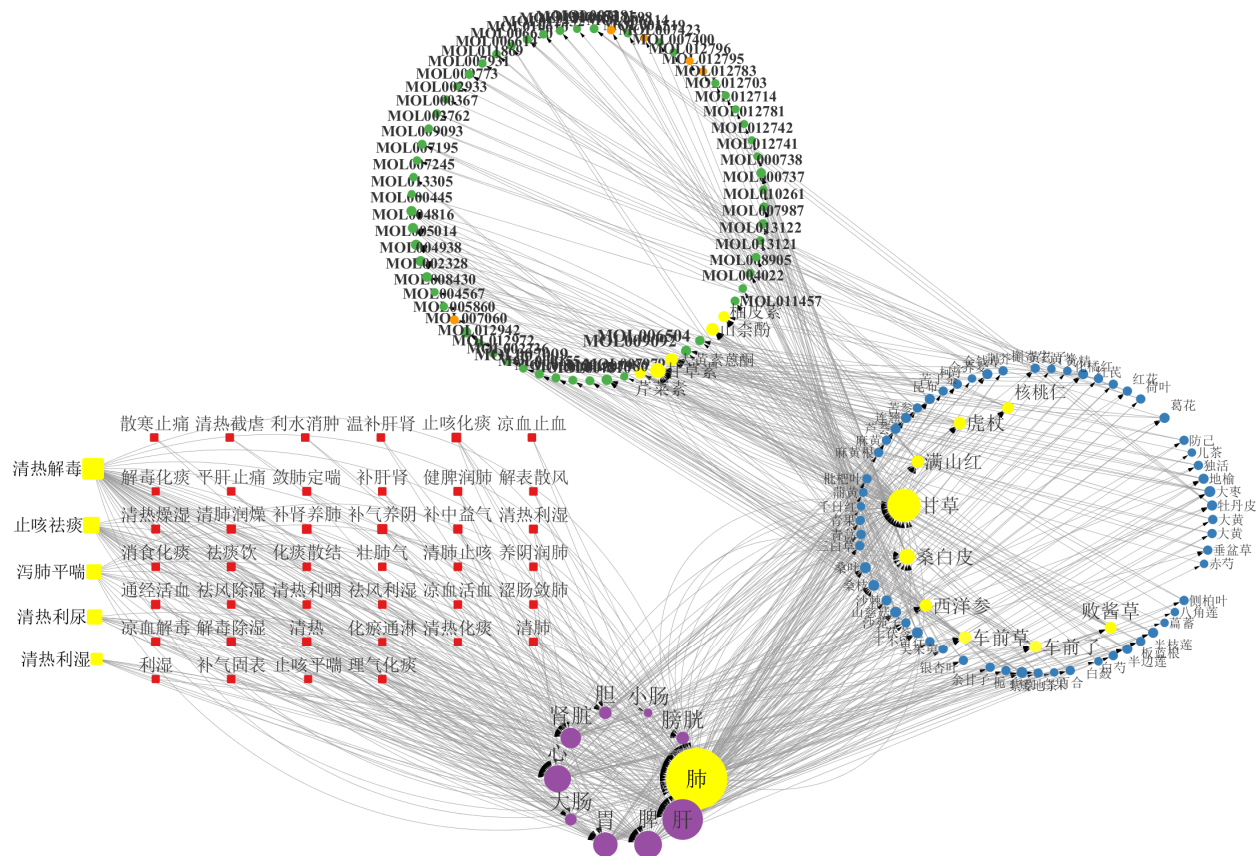


图 3 候选化合物“成分-药材-归经-功效”关联网络

Fig. 3 “Compound-herb-target organ-function” network for potential compounds

COVID-19 的治疗, 而安全有效的抗病毒疫苗仍在研制之中, 因此迫切需要快速、高效地研发安全、经济的药物。

按照传统中医药理论可将 COVID-19 归类于“疫疔”“温病”“伤寒”范畴。《温病条辨》《伤寒论》《黄帝内经》等中医典籍中也记载了大量对这类疾病的诊疗方法和处方。国家卫生健康委员会印发的《国家新型冠状病毒肺炎诊疗方案(试行第七版)》已明确将 COVID-19 归为中医疫病范畴, 病因为感受疫戾之气, 病位在肺, 基本病机特点为“湿、热、毒、瘀”^[18-19]。因此, 从我国丰富的中药宝库中寻找治病方法, 找到具有治愈潜力的药物是可行之途。

从复杂的中医药系统中筛选有效的治疗药物, 解析活性成分的疗效, 评价治病的机制是一项长期、困难的工程。而从已有化合物数据库、已上市药物或临床试验药物中进行高通量地药物筛选显然是在急性爆发疫情局面下的优势策略。基于“锁-钥匙”理论的计算机药物辅助设计中的分子对接技术, 可以快速、高效完成早期分子虚拟筛选、药物分子设计、潜在作用靶点测试、相互作用机制等工作, 是药物发现过程中非常重要的工具之一。因此, 本研究采用了分子对接技术, 结合药动力学和药物设计学参数, 对 TCMSP 数据库中上万个分子化合物进行了靶点对接测试, 初步找到可能对 SARS-CoV-2 3CL 水解酶产生抑制作用、且具有良好 OB 及 DL 的小分子化合物 60 个。在本研究所测试的 TCMSP 数据库中药成分化合物数据的基础上, 仍需进行分子动力学模拟测评、基因组序列靶点筛选与抗病毒的测试等研究工作。

本研究通过挖掘候选化合物归属药材、归经及作用功效, 构建“成分-药材-归经-功效”网络, 对其进行可视化分析, 结果显示, 柚皮素、山柰酚、甘草素、芹菜素等化合物归属的甘草、桑白皮、虎杖、满山红等药材, 具有清热解毒、清热利湿、止咳平喘的功效, 治疗机制符合 COVID-19 因“湿热毒瘀”导致的肺失宣发肃降、肺气上逆、气机紊乱等症状。这些结果期望对未来抗 SARS-CoV-2 药物研究提供相关数据。

参考文献

[1] Yang H T, Yang M J, Ding Y, *et al.* The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor [J]. *Proc Natl Acad Sci*

USA, 2003, 100(23): 13190-13195.

- [2] Wu F, Zhao S, Yu B, *et al.* Complete genome characterisation of a novel coronavirus associated with severe human respiratory disease in Wuhan, China [J/OL]. *BioRxiv*, [2020-01-24]. <https://doi.org/10.1101/2020.01.24.919183>.
- [3] Zhao S, Lin Q Y, Ran J J, *et al.* Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak [J]. *Int J Infect Dis*, 2020, 92(1): 214-217.
- [4] Guo Q, Li M, Wang C H, *et al.* Host and infectivity prediction of Wuhan 2019 novel coronavirus using deep learning algorithm [J/OL]. *BioRxiv*, [2020-01-21]. <https://doi.org/10.1101/2020.01.21.914044>.
- [5] Chen T M, Rui J, Wang Q P, *et al.* A mathematical model for simulating the transmission of Wuhan novel coronavirus [J]. *BioRxiv*, [2020-01-19]. <https://doi.org/10.1101/2020.01.19.911669>.
- [6] Li Y, Zhang J Y, Wang N. Therapeutic drugs targeting 2019-nCoV main protease by high-throughput screening [J/OL]. *BioRxiv*, [2020-01-28]. <https://doi.org/10.1101/2020.01.28.922922>.
- [7] Letko M, Marzi A, Munster V. Functional assessment of cell entry and receptor usage for lineage B β -coronaviruses, including 2019-nCoV [J/OL]. *Nat Microbiol*, [2020-02-24]. <https://doi.org/10.1038/s41564-020-0688-y>.
- [8] Tian X L, Li C, Huang A L, *et al.* Potent binding of 2019 novel coronavirus spike protein by a SARS coronavirus-specific human monoclonal antibody [J]. *Merg Microb Infect*, 2020, 9(1): 382-385.
- [9] Liu X, Wang X J. Potential inhibitors for 2019-nCoV coronavirus M protease from clinically approved medicines [J/OL]. *BioRxiv*, [2020-01-29]. <https://doi.org/10.1101/2020.01.29.924100>.
- [10] Xu Z J, Peng C, Shi Y L, *et al.* Nelfinavir was predicted to be a potential inhibitor of 2019-nCoV main protease by an integrative approach combining homology modelling, molecular docking and binding free energy calculation [J/OL]. *BioRxiv*, [2020-01-27]. <https://doi.org/10.1101/2020.01.27.921627>.
- [11] Paraskevis D, Kostaki E, Magiorkinis G N, *et al.* Full-genome evolutionary analysis of the novel coronavirus (2019-nCoV) rejects the hypothesis of emergence as a result of a recent recombination event [J]. *Infect Genetics Evol*, 2020, 79(1): 104212-104215.
- [12] Dong N, Yang X M, Ye L W. Genomic and protein

- structure modelling analysis depicts the origin and infectivity of 2019-nCoV, a new coronavirus which caused a pneumonia outbreak in Wuhan, China [J/OL]. *BioRxiv*, [2020-01-20]. <https://doi.org/10.1101/2020.01.20.913368>.
- [13] Zhou P, Yang X L, Wang X G, *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin [J]. *Nature*, 2020, 579(7798): 270-273.
- [14] Wang M L, Cao R Y, Zhang L K, *et al.* Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) *in vitro* [J]. *Cell Res*, 2020, 30(3): 269-271.
- [15] Ru J L, Li P, Wang J N, *et al.* TCMSP: A database of systems pharmacology for drug discovery from herbal medicines [J]. *J Cheminform*, 2014, 6(1): 13-18.
- [16] 谭金芝. SARS 冠状病毒 3CL 蛋白酶性质的动力学模拟和 EGFR、hCyPJ 的药物分子设计 [D]. 上海: 中国科学院上海药物研究所, 2005.
- [17] Thomsen R, Christensen M H. MolDock: A New technique for high-accuracy molecular docking [J]. *J Med Chem*, 2006, 49(11): 3315-3321.
- [18] 贺福元, 邓凯文, 潘雪, 等. 基于中医药超分子“气析”理论探讨抗新型冠状病毒药物的研究策略 [J]. *中草药*, 2020, 51(3): 557-562.
- [19] 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 (2019-nCoV) 肺炎活性化合物的研究 [J]. *中草药*, 2020, 51(4): 836-844.