

基于化学信息学的解表与收涩中药成分比较研究

雷 蕾¹, 杨 乐¹, 李小阳¹, 刘 骏², 王 晰¹, 王 忠^{2*}

1. 中国中医科学院中医药信息研究所, 北京 100700

2. 中国中医科学院中医临床基础医学研究所, 北京 100700

摘 要: **目的** 比较解表功效中药化学和收涩功效中药成分的重要理化性质、Murcko 骨架结构特征并计算 Bayesian 距离, 为中药功效和化学成分结构性质的研究提供参考。**方法** 首先将《中药学》记载的解表中药和收涩中药名作为关键词, 在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)、中医药百科全书数据库 (The Encyclopedia of Traditional Chinese Medicine, ETCM)、中药化学数据库 (Traditional Chinese Medicine Database, TCMD) 3 个数据库中进行检索, 得到化学成分, 查重合并后下载成分的 2D 结构文件。然后使用 Discovery Studio 4.5 计算分子的重要性质、统计参数 $|u|$ 和 Bayesian 距离。当 $|u| > 1.96$ 时, 认为解表中药的成分与收涩中药的成分之间存在显著性差异。使用 KNIME 4.3.2 软件 CDK 算法包完成化学成分的加氢、去除糖基等, 使用 RDKit 算法包完成 Murcko 骨架计算。**结果** 解表中药成分与收涩中药成分在芳香环数量、分子环数、相对分子质量、氢受体数量、水溶性、极性表面积、氢给体数量、细胞色素 P450 2D6、血脑屏障、脂水分配系数、血浆蛋白结合预测分数等 11 个理化性质方面都存在明显差异。苯环、黄酮、环己烯、环己烷骨架在解表中药成分中占比大一些, 而羽扇豆烷骨架、联苯环辛烯类木脂素骨架和异喹啉类生物碱骨架在收涩中药成分中占比大一些。从 Bayesian 距离来看, 固精缩尿止带成分-固表止汗成分、固表止汗成分-敛肺涩肠成分的相对距离大于发散风寒成分-发散风热成分。**结论** 中药功效和化学成分性质结构之间存在内在联系。不同功效中药包含的化学成分在理化性质、高频 Murcko 骨架等方面都有明显差别。

关键词: 功效; 化学成分; 理化性质; Murcko 骨架; Bayesian 距离

中图分类号: R284.1

文献标志码: A

文章编号: 0253-2670(2021)06-1822-06

DOI: 10.7501/j.issn.0253-2670.2021.06.025

Comparative study on components of traditional Chinese medicine for exterior-relieving and apocrustic efficacy based on chemical informatics

LEI Lei¹, YANG Le¹, LI Xiao-yang¹, LIU Jun², WANG Xi¹, WANG Zhong²

1. Institute of Information on Traditional Chinese Medicine, China Academy of Chinese Medical Sciences, Beijing 100700, China

2. Institute of Basic Clinical Medicine, China Academy of Chinese Medical Sciences, Beijing 100700, China

Abstract: Objective To provide a reference for the study of the efficacy and structural properties of chemical components of traditional Chinese medicine, the important physical and chemical properties, Murcko skeleton and the Bayesian distance of the chemical components of exterior-relieving herbs and apocrustic herbs were compared. **Methods** First of all, the chemical components were obtained and the plane structure documents after duplicate checking and merging were downloaded by searching the keywords of traditional Chinese medicine for exterior-relieving and apocrustic efficacy in TCMSP, ETCM, TCMD databases. Then, Discovery Studio4.5 was used to calculate the important characters and statistical parameters of molecule $|u|$. When $|u| > 1.96$, it indicated a significant difference between the components of exterior-relieving herb and apocrustic herb. The CDK package of KNIME4.3.2 was adopted to complete the process of hydrogenation and glycosylation removal and the RDKit package was used to calculate the Murcko skeleton. **Results** The components of exterior-relieving herbs and apocrustic herbs were obviously different in 11 physical and chemical properties, including the number of the aromatic rings, the number of molecular rings, molecular weight, the number of hydrogen receptors, water-solubility, polar surface area, the number of hydrogen donors, cytochrome P450 2D6,

收稿日期: 2021-12-04

基金项目: 中国中医科学院基本科研业务费自主选题项目 (ZZ140304)

作者简介: 雷 蕾, 副研究员。E-mail: leilei@mail.cintcm.ac.cn

*通信作者: 王 忠, 研究员。E-mail: zhonw@vip.sina.com

blood-brain barrier, ALogP, and the predictive fraction of plasma protein binding. Benzene ring, flavonoids, cyclohexene, cyclohexane skeleton were accounted for a larger proportion in the components of exterior-relieving herbs, while lupine alkane, dibenzocyclooctadiene lignans, and isoquinoline alkaloid skeleton were accounted for a larger proportion in the components of the apocrustic herbs. According to the Bayesian distance, the relative distance between securing essence and reducing urination components-consolidating exterior and antiperspirant components, and consolidating exterior and antiperspirant components-constraining lung and astringing intestine components was greater than those between wind-cold dispersing components-wind-heat dispersing components. **Conclusion** There was an internal relationship between the efficacy and structural properties of chemical components of traditional Chinese medicine. There are obvious differences in physicochemical properties and high frequency Murcko skeleton in chemical components of TCM with different functions.

Key words: efficacy; chemical components; physical and chemical properties; Murcko skeleton; Bayesian distance

化学信息学是化学领域中一个重要的分支,它利用计算机技术和计算机网络技术,对化学信息进行表示、管理、分析、模拟,揭示化学信息的实质与内在联系。目前,已经有很多报道使用化学信息学方法对不同种类的化学成分进行比较研究,以找到类之间最基本区别和内在联系。Shang 等^[1]使用化学信息学比较了陆地和海洋天然产物的重要理化性质、结构特征和药物相似性。Muigg 等^[2]探索了由物理化学性质定义的 3802 个海洋分子和 29 620 个陆地分子的化学空间。Kong 等^[3-4]分析了陆地和海洋天然产物的 Murcko 框架和 ClogP。Ertl 等^[5]比较了合成药物与天然产物的功能团的差别和联系。

中药功效是中药治疗作用的高度概括与总结,在中医理论指导下,中药的功效是中医辨证用药的基础。而中药发挥作用的物质基础是化学成分,那么中药化学成分的结构和性质与中药功效之间是否存在联系呢?张文晋等^[6]根据“药效团”的科学内涵,提出了适用于中药现代化研究的“药效基团-功效假说”。然而至目前为止,仍然很少有系统探索中药功效对应的化学成分的结构特征和分子性质的报道。中药功效的分类并不完全统一,根据《中药学》教材^[7]的标注,中药功效主要包括解表、泻下、清热等。本研究选择了解表和收涩这 2 类具有相对功效的中药,系统比较解表功效化学成分和收涩功效化学成分的重要理化性质和骨架结构特征,为中药功效和化学成分结构性质的研究提供参考。

1. 资料与方法

1.1 数据来源

在《中药学》^[7]中第 8 章解表药和第 25 章收涩药中检索得到解表功效的中药和收涩功效的中药。将中药名称作为关键词,在 TCMSP (Traditional Chinese Medicine Systems Pharmacology Database and Analysis Platform)^[8]、ETCM (中医药百科全书

数据库)^[9]、TCMD (Traditional Chinese Medicine Database)^[10]3 个数据库中分别进行检索,得到的化学成分名称,查重合并。

在 TCMSP、TCMD 和 Pubchem 中检索化学成分的结构,保存为 sdf 格式文件和 mol2 格式文件。使用 OpenBabel2.4.1 将 mol2 格式文件转换为 sdf 格式文件。

1.2 研究方法

当方差已知且样本量较大时,2 组分子特性总体均值差异的统计学意义用 u 检验或 z 检验来评价。 u 值的计算公式如下。

$$u = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

公式中 $\bar{x}_1$, $\bar{x}_2$ 分别表示解表中药的成分和收涩中药的成分的分子性质的平均值; S_1 和 S_2 分别表示解表中药的成分和收涩中药的成分的一个性质的标准差; n_1 和 n_2 分别表示解表中药的成分和收涩中药的成分的总分子数。因此,当计算出 $|u| > 1.96$ 时,认为解表中药的成分与收涩中药的成分之间存在显著性差异。

本研究使用 Discovery Studio 4.5 计算分子的重要性质和 Bayesian 距离计算,如果没有特别说明,设置均为系统默认参数。由于约 14.8% 的天然产物含有糖单位^[11-12],因此使用 KNIME 4.3.2 软件 CDK 算法包完成化学成分的加氢,去除糖基等,使用 RDKit 算法包完成 Murcko 骨架计算。使用 Excel 完成统计分析。

2 结果

2.1 中药和化学成分统计情况

本研究依据《中药学》^[7]中第 8 章检索得到解表中药 36 个,包括白芷、薄荷、苍耳子等中药。去除动物药和矿物药,植物药中共包含 3031 个化学成

分。依据《中药学》^[7]中第二十五章检索得到收涩中药 29 个,包括:荷叶、鸡冠花、金樱子等。去除了动物药和矿物药,植物药中共包含 1082 个化学成分。解表中药和收涩中药共有 3756 个成分,如图 1 所示,其中解表中药特有化学成分 2674 个,收涩中药特有化学成分 725 个,共享成分 357 个。

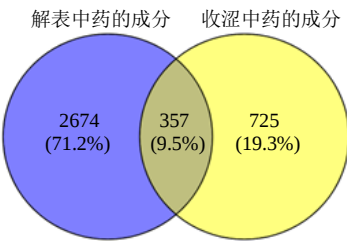


图 1 解表中药的成分和收涩中药的成分韦恩图
Table 1 Venn diagram of chemical components of exterior-relieving herbs and apocrustic herbs

2.2 分子重要理化性质分析

为了更加了解解表药成分与收涩中药成分在分子理化性质方面的差异,本研究对解表中药特有 2674 个化学成分,收涩中药特有 725 个化学成分(图 1)进行了理化性质分析。从表 1 中可以看出,芳香环数量、分子环数、相对分子质量等 11 个重要的分子理化性质的 u 值绝对值都高于 1.96,说明这些性质在解表中药的成分组和收涩中药的成分组中有显著性差异。芳香环数量、分子环数、相对分子质量的 u 检验值分别为-11.53、-11.2、-9.67,说明收涩中药的成分分子量较大,分子环数和芳香环数更多,结构更加刚性。水溶解性和 ALogP 的 u 值分别为 8.68 和-2.78,说明解表中药的成分更易溶于水,而收涩中药的成分更加疏水。细胞色素 P450 2D6 的预测值的 u 值为-4.47,说明收涩中药的成分更多地需要通过肝脏 CYP2D6 酶代谢。血脑屏障的 u 值为 3.93,

表 1 解表药和收涩药成分的分子重要的理化性质
Table 1 Important physicochemical properties of chemical components of exterior-relieving herbs and apocrustic herbs

序号	理化性质	解表药成分			收涩药成分			u	$ u $
		Min	Max	$\bar{x} \pm s$	Min	Max	$\bar{x} \pm s$		
1	芳香环数量	0	6	0.69±0.97	0	10	1.35±1.46	-11.53	11.53
2	分子环数	0	14	2.40±2.31	0	14	3.46±2.24	-11.20	11.20
3	相对分子质量	30.07	1 791.83	311.66±202.3	27.03	1 859.31	397.82±215.57	-9.67	9.67
4	氢受体数量	0	44	4.35±4.75	0	51	6.72±6.69	-8.97	8.97
5	水溶解性	-16.60	3.28	-3.75±2.08	-58.26	3.08	-5.67±5.85	8.68	8.68
6	极性表面积	0	691.4	69.18±78.02	0	871.93	104.52±117.14	-7.68	7.68
7	氢给体数量	0	23	2.02±2.92	0	28	2.99±4.13	-5.98	5.98
8	细胞色素 P450 2 D6	-19.74	3.57	-3.35±3.18	-11.77	34.61	-2.76±3.18	-4.47	4.47
9	血脑屏障	-1.95	1.96	0.22±0.73	-1.98	2.04	0.10±0.74	3.93	3.93
10	ALogP	-6.35	30.57	3.04±2.73	-4.60	16.78	3.36±2.81	-2.78	2.78
11	血浆蛋白结合预测分数	-31.46	15.7	-3.08±6.51	-42.17	40.07	-2.45±7.22	-2.15	2.15

说明解表中药的成分有更高的血脑屏障穿透力。血浆蛋白结合预测分数的 u 值为-2.15,说明收涩中药的成分有更高的血浆蛋白结合能力。

2.3 Murcko 骨架分析

本研究发现,解表中药的成分共筛选 859 种 Murcko 骨架,收涩中药的成分共筛选 365 种 Murcko 骨架,它们共享 68 种 Murcko 骨架,共享 Murcko 骨架占比排列前 5 位的见图 2。可以看出苯环、黄酮、环己烯、环己烷骨架在解表中药的成分中占比大一些,而羽扇豆烷骨架在收涩中药的成分中占比大一些,具体见图 3。

比较了解表中药的成分和收涩中药的成分的 Murcko 骨架,在每类成分中占比前 5 位的骨架列于表 2。可以看出,除了共享的 Murcko 骨架外,每类成分还有一些占比较高的独特的骨架。

2.4 解表中药成分与收涩中药成分的 Bayesian 距离

在《中药材》教材^[7]中,解表中药的功效可以进一步分为发散风寒药和发散风热药 2 种,而收涩中药的功效可以进一步分为固精缩尿止带药、敛肺涩肠药、固表止汗药 3 种。解表中药包含 2674 个成分,其中发散风寒成分 1149 个,发散风热成分 1118 个,还有 407 个成分属于 2 种功效分类,标

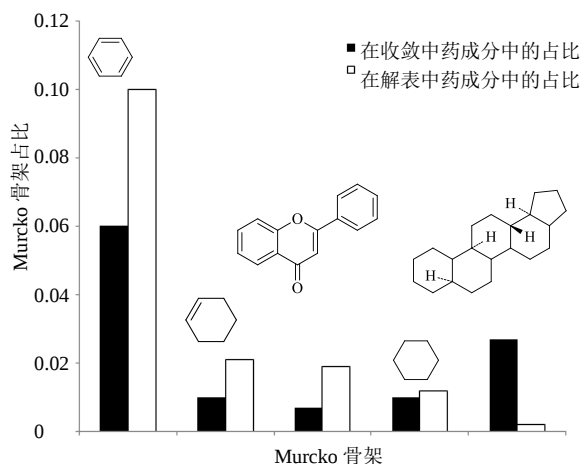


图 2 共享 Murcko 骨架占比情况

Fig. 2 Share Murcko skeleton proportion

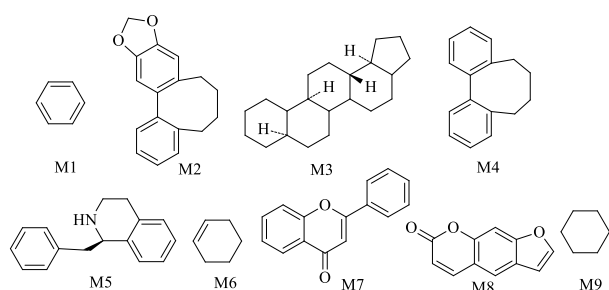


图 3 占比较高的 Murcko 骨架

Fig. 3 Murcko skeletons with high proportion

表 2 收敛中药和解表中药成分 Murcko 骨架占比较高的前 5 个骨架结构

Table 2 Top five Murcko skeletons with high proportion in chemical components of exterior-relieving herbs and apocrustic herbs

Murcko 骨架	收敛中药的成分		Murcko 骨架	解表中药的成分	
	数量	占比/%		数量	占比/%
M1	43	0.060	M1	266	0.100
M2	33	0.046	M6	56	0.021
M3	19	0.027	M7	51	0.019
M4	16	0.022	M8	43	0.016
M5	9	0.013	M9	32	0.012

记为不确定。收敛中药包含 725 个成分，其中固精缩尿止带成分 353 个、敛肺涩肠药 350 个、固表止汗 19 个，还有 3 个成分属于 2 种功效分类，标记为不确定。

使用 Discovery Studio4.5 中的 compare libraries 功能为每一个分类的中药成分都建立了贝叶斯分类模型，并根据贝叶斯评分计算贝叶斯距离，用来评

价不同类中药成分之间的差异，Bayesian 距离越大，差异越大。此外，为了更加直观，进一步制作了发散风寒成分-发散风热成分、固精缩尿止带成分-固表止汗成分、固表止汗成分-敛肺涩肠成分的 Bayesian 统计图，见图 4。每个统计图上的横坐标均为实线代表的数据集合的统计堆栈中心的贝叶斯分数，纵坐标则是 2 个数据集落入这个统计堆栈的配体的标准化频率。

3 讨论

在中医理论中，解表药以发散表邪为主要功效，常用来治疗表证。这类药物大多辛散轻扬，主入肺、膀胱经，偏行肌表，能促进肌体发汗，使表邪由汗出而解，从而达到治愈表证，防止疾病传变的目的。即《内经》所谓：“其在皮者，汗而发之。”收敛药则以收敛固涩为主要功效，常用以治疗各种滑脱病证。该类药味多酸涩，性温或平，主入肺、脾、肾、大肠经。具有收敛固涩之功，以敛耗散、固滑脱，即陈藏器所谓“涩可固脱”，李时珍所谓“脱则故而不收，故用酸涩药，以敛其耗散”之意。

本研究从这 2 类功效的中药包含的成分出发，首先比较了解表药成分和收敛药成分情况，发现解表药和收敛药有 357 个化学成分是共有的，解表药独特的成分占比为 88.2% (2674/3031)，收敛药独特的成分占比为 67% (725/1082)，说明在中医理论中虽然发散与收敛固涩为相对的功效，但是有一部分的中药可能具有双向调节作用。

口服给药后的人体肠道吸收 (human intestinal absorption, HIA) 是中药进入人体的主要途径。肠道吸收是吸收的百分比，一种吸收良好的化合物至少 90% 会被人体血液吸收^[13]。在研究药物在细胞内的运输时发现极性表面积和 AlogP 与 HIA 有着明显相关。从表 1 中的分子理化性质可以看出，收敛中药的成分较解表中药的成分的极性表面积和 AlogP 有着显著差别，说明这 2 种成分在 HIA 中有明显差别。这一点从 2 类成分的芳香环数量、分子环数、相对分子质量、水溶性等性质的比较也可得到验证。

Murcko 骨架由 Murcko 等^[14]设计并用于药物分子的形状、结构分析，它可以去除支链，将分子结构中的母核提取出来，是化学信息学中常用的骨架分析方法。经过分析发现，黄酮结构在解表类中药中占比较高。刘娜等^[15]采用 RP-HPLC 法分析其解表的药效组分，结果发现葛根解表的药效组分为黄酮类组分，各组分的比例在一定范围内变化。郭文勇等^[16]

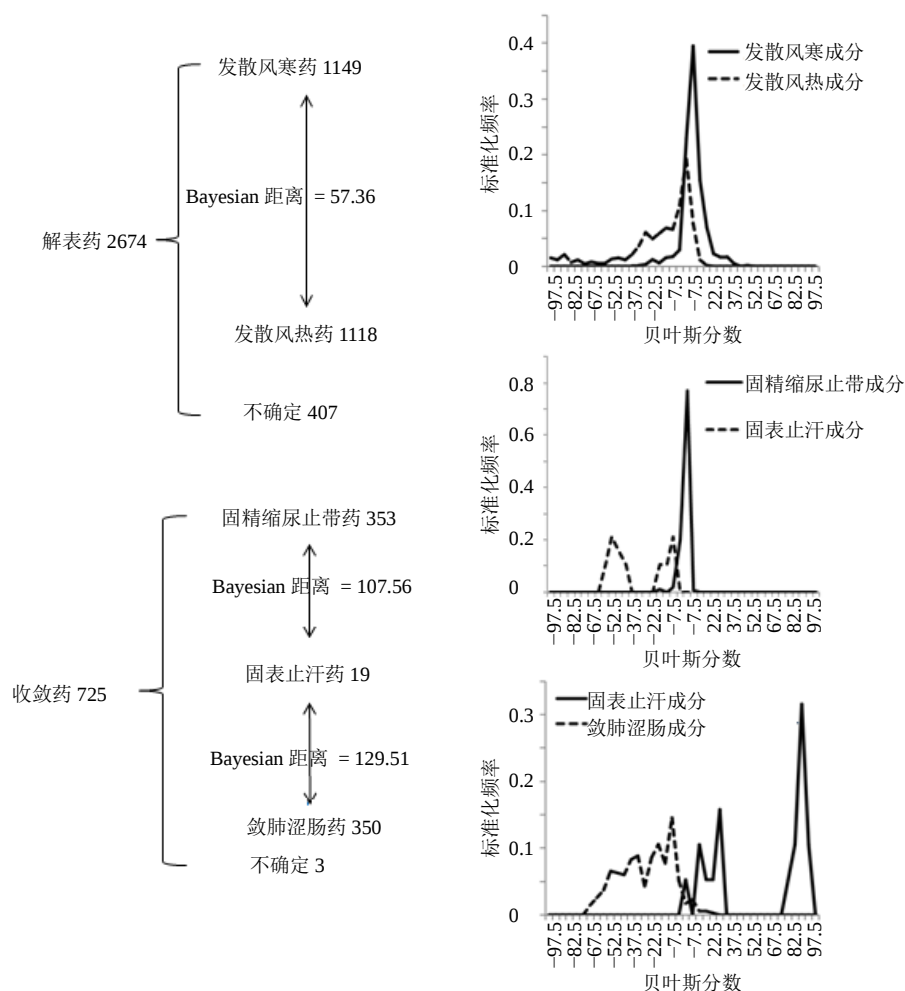


图 4 解表中药成分与收涩中药成分的 Bayesian 距离及 Bayesian 统计图

Fig. 4 Bayesian distance and Bayesian statistics of chemical components of exterior-relieving herbs and apocrustic herbs

研究了解表药淡豆豉“解表除烦”作用机制,发现起作用的主要有效成分是异黄酮和皂苷类成分。

此外,联苯环辛烯类木脂素骨架和异喹啉类生物碱骨架在收涩类中药中占比较高,这 2 类化合物多具有肝保护作用。如余凌虹等^[17]发现联苯环辛烯类木脂素五味子甲素、(±)-五味子乙素、R-(+)-五味子丙素等 10 余种化合物对化学性肝损伤模型均有保护作用。崔文博等^[18]报道了异喹啉类生物碱普罗托品具有抗心律失常、抗疟疾作用、镇痛作用以及对肝损伤的保护作用。

从不同功效中药成分的 Bayesian 距离来看,固精缩尿止带成分-固表止汗成分、固表止汗成分-敛肺涩肠成分的相对距离大于发散风寒成分-发散风热成分。一方面可能因为解表药-收涩药、发散风寒药-发散风热药包含了比较多的相同的成分,另一方面可能因为收涩药下分固精缩尿止带药、固表止汗药和敛肺涩肠药在临床上治疗的疾病差异比较大,

因此成分的差异也比较大。

综上所述,本研究通过系统地比较解表药和收涩药这 2 类具有相反功效的中药包含的成分的分子理化性质和 Murcko 骨架特征,从分子的刚性特点、HIA、BBB、疏水性和亲水性等特点阐述了解表药成分和收涩药成分有显著差异,最后通过 Bayesian 距离直观展现出固精缩尿止带成分-固表止汗成分、固表止汗成分-敛肺涩肠成分的差异大于发散风寒成分-发散风热成分,为中药功效和化学成分结构性质的内在规律研究提供参考。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] Shang J, Hu B, Wang J M, et al. Cheminformatic Insight into the differences between terrestrial and marine originated natural products [J]. *J Chem Inf Model*, 2018, 58(6): 1182-1193.
- [2] Muigg P, Rosén J, Bohlin L, et al. *In silico* comparison of

- marine, terrestrial and synthetic compounds using ChemGPS-NP for navigating chemical space [J]. *Phytochem Rev*, 2013, 12(3): 449-457.
- [3] Kong D X, Jiang Y Y, Zhang H Y. Marine natural products as sources of novel scaffolds: Achievement and concern [J]. *Drug Discov Today*, 2010, 15(21/22): 884-886.
- [4] Kong D X, Guo M Y, Xiao Z H, *et al.* Historical variation of structural novelty in a natural product library [J]. *Chem Biodivers*, 2011, 8(11): 1968-1977.
- [5] Ertl P, Schuhmann T. A systematic cheminformatics analysis of functional groups occurring in natural products [J]. *J Nat Prod*, 2019, 82(5): 1258-1263.
- [6] 张文晋, 曹也, 常丽坤, 等. 中药“药效基团-功效假说”及其应用 [J]. *中草药*, 2021, 52(7): 1865-1871.
- [7] 钟赣生. 中药学 [M]. 第 4 版. 北京: 中国中医药出版社, 2016: 148.
- [8] Ru J L, Li P, Wang J A, *et al.* TCMSP: A database of systems pharmacology for drug discovery from herbal medicines [J]. *J Cheminformatics*, 2014, 6(1): 13.
- [9] Xu H Y, Zhang Y Q, Liu Z M, *et al.* ETCM: An encyclopaedia of traditional Chinese medicine [J]. *Nucleic Acids Res*, 2019, 47(D1): D976-D982.
- [10] NeoTrident, Ed. Traditional Chinese Medicine Database [DB]. 2009.
- [11] Koch M A, Schuffenhauer A, Scheck M, *et al.* Charting biologically relevant chemical space: A structural classification of natural products (SCONP) [J]. *Proc Natl Acad Sci USA*, 2005, 102(48): 17272-17277.
- [12] Grabowski K, Baringhaus K H, Schneider G. Scaffold diversity of natural products: Inspiration for combinatorial library design [J]. *Nat Prod Rep*, 2008, 25(5): 892-904.
- [13] 韩卓君, 吴小平, 严明煜, 等. 中药与肠道菌群相互作用机制探讨 [J]. *世界中医药*, 2021, 16(24): 3591-3595.
- [14] Bemis G W, Murcko M A. The properties of known drugs. 1. Molecular frameworks [J]. *J Med Chem*, 1996, 39(15): 2887-2893.
- [15] 刘娜, 张贵君, 金哲雄, 等. 葛根解表药效组分分析 [J]. *现代药物与临床*, 2009, 24(5): 294-296.
- [16] 郭文勇. 中药淡豆豉的质量评价方法及其“解表除烦”作用机制研究 [D]. 上海: 第二军医大学, 2004.
- [17] 余凌虹, 刘耕陶. 五味子联苯环辛烯类木脂素成分的结构与药理活性关系和药物创新 [J]. *化学进展*, 2009, 21(1): 66-76.
- [18] 崔文博, 钱金楸. 普罗托品的研究进展 [J]. *医学综述*, 2012, 18(14): 2260-2262.

[责任编辑 王文倩]