

使 用 说 明

(如软件未出现错误, 请直接从第二页开始看详细教程)

如果想获取软件更新或软件出现错误, 请到软件更新网址下载最新版。

也可以直接与我联系。E-mail: mwxiao@hnu.edu.cn; QQ: 1306410893

软件更新地址:

<https://pan.baidu.com/s/1o7QRFaM>; 密码: n4n1

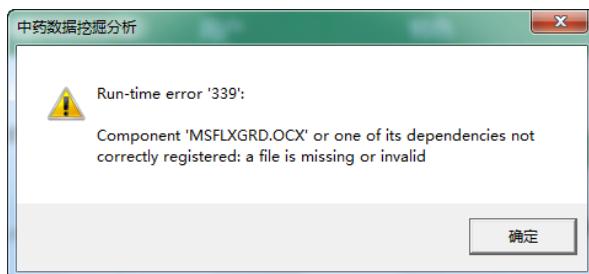
也可以在此网址下载相关软件 gephi, 如果要使用 gephi, 请先安装 java。

软件视频教程:

http://v.youku.com/v_show/id_XMTgyNTQ5NjcyMA==.html; 密码: haha

常见错误处理:

1. 软件出现 MSFLXGRD 未注册 (如下图):



解决方法: 请右键单击->该软件目录下的“注册 MSFLXRD.bat”文件->选择以管理员身份运行, 即可解决该问题。



如果没有“注册 MSFLXRD.bat ”文件, 请到软件更新网址下载, 并放到该软件目录下。

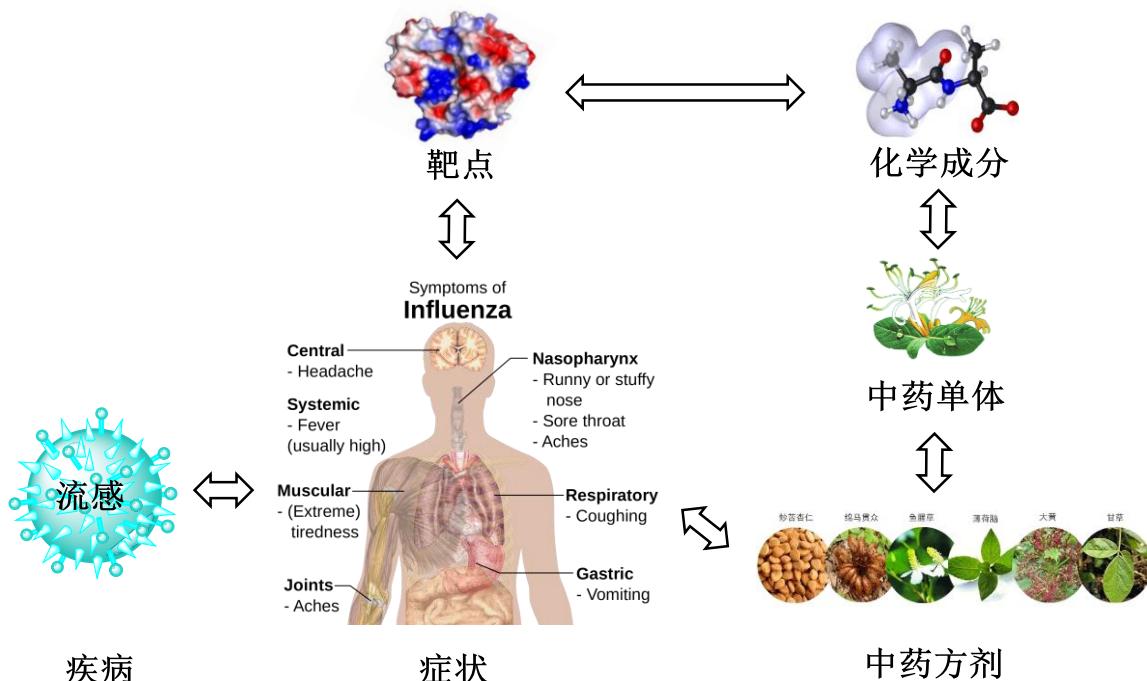
中药数据挖掘小软件说明书

肖梦武¹, 崔满营¹, 张蒙¹, 钟莹²

(¹湖南大学化学与化工学院, ²湖南大学信息科学与工程学院)

指导老师: 胡艾希 教授

中药是药物发现的宝库, 从疾病-病症- (靶点) -中药方剂-中药单体-化学成分-生物活性多层次去展示中药, 用更直观、科学的方式去分析挖掘中药中的宝藏——这是我们做这个软件的初衷。



这个中药数据挖掘这个小软件实现了疾病-病症-方剂-单味药-化学成分-生物活性的基本关联, 初步实现了病症-方剂关联, 方剂-单味药分析, 方剂-药对分析, 单味药-化学成分-生物活性(抗流感、抗肿瘤)关联与分析。

由于时间仓促, 能力和专业受限, 软件中有很多不足之处, 还请老师指正。

一、软件用到的数据资源

1. 数据来源

(1) 中国中药方剂库, 是由国家人口与健康科学数据共享平台的中国中医科学院中医药信息研究所数据室提供的。

(2) 中药化学组分数据库, 是从 TCMS 中药数据库收集的。主要参考文献: *J Cheminformatics*. 2014, 6(1):13.

(3) 抗流感天然产物数据库，是由湖南大学化学与化工学院胡艾希课题组搜集文献整理的，目前数据还在完善。未经授权，请勿传播使用。

(4) 天然产物抗癌成分数据库，是从文献中收集的。主要参考文献：Nucl Acids Res. 2013, 41(D1): D1124.

2. 数据处理方法

(1) 数据标准化。对方剂库数据、病症、化学结构、参考文献等进行标准化。

比如：真武汤的方剂“熟附子 3 钱，姜 1 钱，焦白术 2 钱，茯苓 2 钱，归身 2 钱，肉桂 1 钱，炙甘草 8 分，白芍（炒）1 钱 5 分，净枣仁（炒）2 钱。”->标准化为“附子，姜，焦术，茯苓，归身，肉桂，甘草，白芍，枣仁”。

A	B	C
1 方剂名称	方剂药物组成	主治
2 秘传牛黄清心丸	天麻4两，防风2两（去芦），牛胆南星2两半，僵蚕（炒）2两半，全蝎2两小儿惊风，大人中风、中痰、中气，一切风痰。	
3 秘传三仙丹	柏枝、槐角子、生矾。	咳嗽。
4 益气散风汤	人参3钱，黄耆3钱，甘草1钱，半夏1钱，白术5钱，柴胡2钱，茯苓3钱，桔梗虚伤风，头痛发热，身疼腰重，骨节俱痠疼，恶风无汗。	
5 真武汤	熟附子3钱，姜1钱，焦术2钱，茯苓2钱，归身2钱，肉桂1钱，炙甘草8分，产后类中风痉症。	

(2) 数据可靠性。化学结构-活性数据库是人工录入的，每个化合物都有参考文献可以追踪。数据录入是采用本实验室自己设计的化学数据库软件进行规范化录入和管理的。

二、软件主要功能

1. 中药方剂分析

(1) 方剂搜索。

支持多关键词搜索、在结果中搜索和在结果中排除。比如搜索流感的常见症状“咳嗽发烧”，也可以搜索方剂名如或中药名。

方剂分析——搜索历史：感冒			
数据分析		结果中排除	
方剂名称	方剂药物组成	主治	功效
秘传百解散	人参7钱，甘草（炙）3钱5分，白术6钱，白茯苓6钱，黄耆（炒）5钱，陈皮（去白）5钱，糯米（炒）5钱，升麻（炒）2钱，川芎3钱，白芷3钱，天麻2钱5分，僵蚕（炒）1钱5分，南星（姜制）5分。	小儿感冒风寒，身热咳嗽，欲出疹痒，并痘后欲出麻疹。	
消风宁嗽散	桔梗1钱，枳壳1钱，半夏1钱，陈皮1钱，前胡1钱，干葛1钱，茯苓1钱，苏叶1钱2分，杏仁1钱，桑白皮1钱，甘草4分。	感冒鼻塞，咳嗽。	
润肺汤	知母2钱，紫菀5钱半（洗净），山梔仁5分，甘草1钱半（炒），麻黃5分（滚汤泡，去白沫，晒干），荆芥5分（去梗），马兜铃5分（去筋膜），前胡2钱，（去芦），赤芍1钱，桑白皮2钱半（去红皮，用蜜炒），半夏3钱（汤泡7次，晒），赤茯苓2钱（去皮），杏仁5钱半（去皮尖，研成泥用），黄芩2钱半（去脚土）。	上盛下虚，肺脾湿热，气喘，咳嗽痰盛，心胸气闷，不思饮食，或寒热往来，或感冒风寒，喘嗽气急，五劳七伤，吐血。	
消风散	苍术1钱，麻黃1钱，荆芥1钱，白芷1钱，陈皮1钱，甘草5分。	四时感冒，发热恶寒，头痛身重。	

(2) 方剂单味药分析。

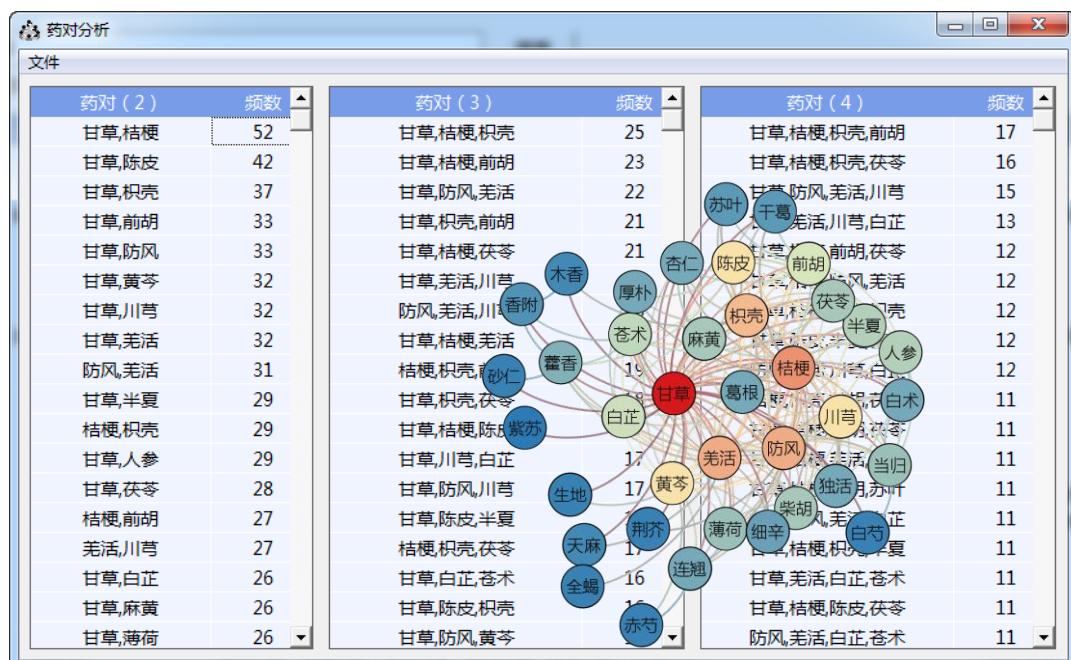
对多个方剂中的中药单体进行频数分析，也支持单味中药分析的功能。而且可以同时分析多个单味药化学成分之间的网络关联。下图是感冒关键词搜索得到的方剂中的单

味药频数分析。具体的单味药分析功能在后面具体讲。



(3) 方剂药对分析。

软件支持二元、三元和四元药对分析，可以从药对分析中看出高频的中药药对。使用方法：数据分析菜单->方剂药对分析。另外，软件还支持导出方剂药对表，用 Gephi 进行网络分析（下图为感冒搜索到方剂的药对结果，使用方法：文件菜单->导出药对列表）。



2. 单味中药分析

(1) 中药类型分析

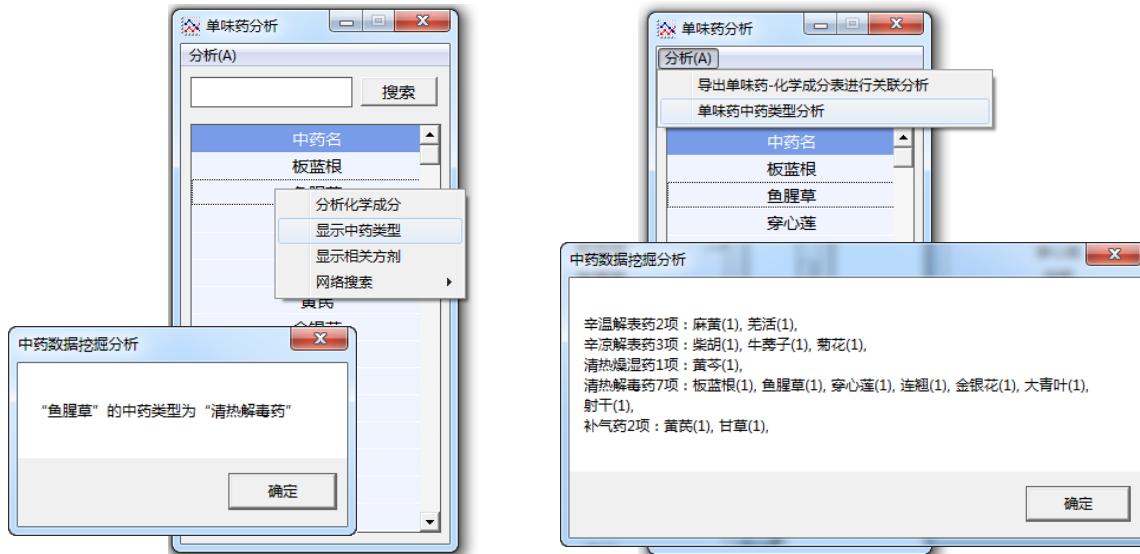
软件支持单味中药的中药类型分析，也支持多个中药单体的中药类型统计。

例子 1：单击右键->显示中药类型->鱼腥草属于清热解毒药。

例子 2：分析文献“中国科学:生命科学, 2015, 45(3):247-255.” 中总结的抗流感药物的中药类型。使用方法：分析菜单->单味药中药类型分析->结果如下：辛温解表药 2 项：麻黄，羌活；辛凉解表药 3 项：柴胡，牛蒡子，菊花；清热燥湿药 1 项：黄芩；清热解毒药 7 项：板蓝根，鱼腥草，穿心莲，连翘，金银花，大青叶，射干；补气药 2 项：黄芪，甘草。

根据中药的功效分为以下 37 个类：辛温解表药，辛凉解表药，清热泻火药，清热燥湿药，清热解毒药，清热凉血药，清虚热药，攻下药，润下药，通水药，祛风湿散寒药，祛风湿清热药，祛风湿筋骨药，化湿药，利水渗湿药，温里药，理气药，消食药，驱虫药，凉血止血药，收敛止血药，化瘀止血药，温经止血药，活血祛瘀药，化痰药，止咳

平喘药，安神药，平肝熄风药，开窍药，补气药，补血药，补阴药，补阳药，收涩药，涌吐药，解毒杀虫燥湿止痒药，拔毒化腐生肌药。

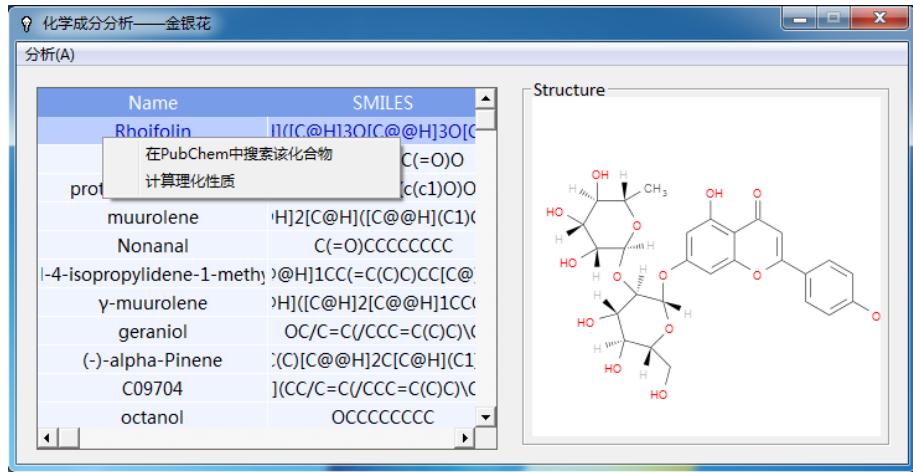


例子 3：涉及感冒的方剂有 215 项，从中分析发现 633 位中药单体，分析出现频数大于 10 的中药的得到的结果如下：

辛温解表药 10 项：防风(48), 羌活(46), 白芷(36), 麻黄(32), 葛根(23), 细辛(19), 紫苏(14), 荆芥(14), 生姜(13), 荆芥穗(12);
辛凉解表药 3 项：薄荷(34), 柴胡(30), 升麻(12);
清热燥湿药 2 项：黄芩(44), 黄连(13);
清热解毒药 1 项：连翘(26);
清热凉血药 1 项：赤芍(15);
攻下药 1 项：大黄(11);
祛风湿散寒药 1 项：独活(18);
化湿药 3 项：苍术(36), 厚朴(19), 砂仁(15);
利水渗湿药 2 项：茯苓(35), 木通(12);
温里药 1 项：干姜(11);
理气药 5 项：陈皮(51), 枳壳(46), 木香(20), 香附(18), 乌药(10);
消食药 1 项：麦芽(11);
活血祛瘀药 1 项：川芎(43);
化痰药 3 项：桔梗(64), 前胡(39), 半夏(36);
补气药 3 项：甘草(126), 人参(38), 白术(28);
补血药 2 项：当归(28), 白芍(13);
其他 14 项：杏仁(30), 苏叶(20), 藿香(20), 干葛(15), 朱砂(15), 天麻(15), 生地(13), 僵蚕(13), 麝香(12), 神曲(12), 紫苏叶(11), 全蝎(11), 白茯苓(10), 雄黄(10)。

(2) 化学成分分析

使用方法 1：单味药窗口->选择中药->右键菜单->化学成分分析。

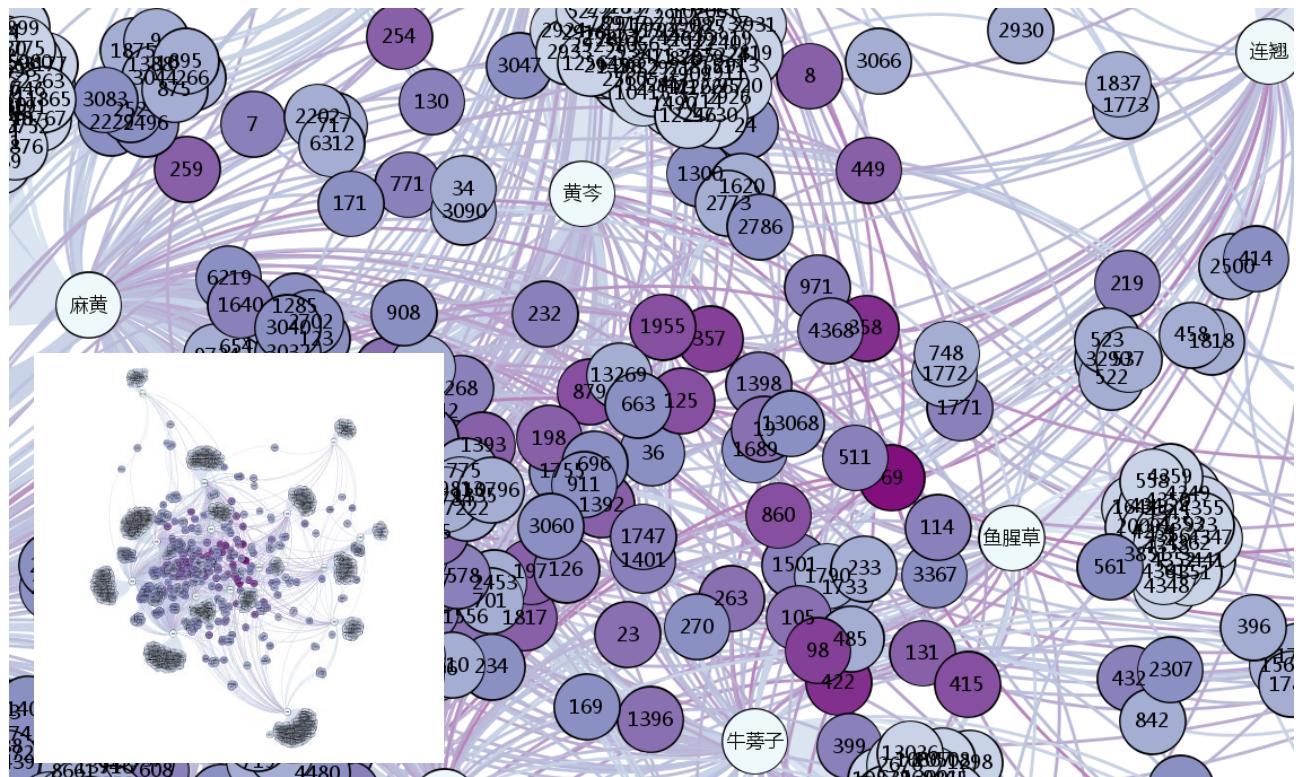


使用方法 2:

主程序菜单-文件->导入单味中药列表->分析->导出单味药-化学成分表->用 gephi 打开分析多个单味药-化学成分的网络关系图。

此处我们研究一下抗流感中药单体的化学成分关联。

文献“邱峰, 张晗, 刘叔文, 等. 抗流感病毒中药及抗埃博拉病毒活性化合物的发现策略[J]. 中国科学: 生命科学, 2015, 45(3):247-255.” 中总结了抗流感病毒的单味中药: 板蓝根、鱼腥草、穿心莲、连翘、黄芩、黄芪、金银花、柴胡、牛蒡子、甘草、麻黄、菊花、大青叶、羌活、射干。本软件对这些中药进行化学成分网络分析 (下图), 可以看出它们的相似性、共有的化学成分 (深色) 及其他关联信息。



(3) 抗流感化学成分分析

使用方法：化学成分分析窗口-分析菜单->可能的活性成分分析->抗流感活性。

该窗口还支持化学成分-活性的文献追踪，理化性质计算、PubChem 搜索和由化学活性成分找被包含的中药。下图显示了金银花中的具有抗流感活性的 12 个化学组分。

Structure					
中药名	成分中文名	成分英文名	抗流感活性	参考文献	
Chaenomeles	3,4-二羟基苯甲	dihydroxyben:	IC50(H1N1)=8.24μM;	Grienke U, Schmidtke M, von Grafenstein S, et al. Influenza	
杜鹃花	熊果酸	Ursolic acid	熊果酸药物血清可降低轮状病毒的毒力，使其感	王婧,宋怀燕.熊果酸药理作用研究进展[J].中华医学研究杂志,	
金銀花	绿原酸	Chlorogenic 颜	搜索含该活性成分的中药 打开参考文献 在PubChem中搜索该化合物 计算理化性质		葛司他
草豆蔻	金合欢酸	ns,trans-farnes	活性		沈霞,罗施,李娜.金银花抗流感病毒H5N1的分子机制研究[J].陕
	紫云英苷	astragalin	活性		Grienke U, Schmidtke M, von Grafenstein S, et al. Influenza
	金丝桃苷	Hyperoside	活性		Grienke U, Schmidtke M, von Grafenstein S, et al. Influenza
	芦丁	Rutinum	活性		于焱.中药单体化合物中神经氨酸酶抑制剂的筛选及其抗流感病毒
天景红	堪非酮	kaempferol	活性		于焱.中药单体化合物中神经氨酸酶抑制剂的筛选及其抗流感病
细皱蔷薇	木犀草素	Luteolin	活性		Hu Ge,Yi-Fei Wang,Jun Xu .Anti-influenza agents from
	大波斯菊甙	cosmosin	活性		Ritter,A.,杜冠华,秦海林,等.细皱蔷薇中主要化学成分及有效组分
	芹菜素	apigenin	活性		Grienke U, Schmidtke M, von Grafenstein S, et al. Influenza
			活性		Hu Ge,Yi-Fei Wang,Jun Xu .Anti-influenza agents from

(4) 抗癌化学成分分析

使用方法：化学成分分析窗口-分析菜单->可能的活性成分分析->抗流感活性。同样支持文献追踪，理化性质计算、PubChem 搜索等。

Structure	SMILES	NPACT_ID
	Oc1cc(ccc1O)C(=O)O	881
	OCl([C@H]1CC=C(CC1)O)C(C)	240
	Oc1cc(O)ccc1c2c2c3c(O)C(=C3)C(C)Ccc2c2c1C=C(C(C)O)O	778
	Oc1ccc(c1O)C@12CC[C@@H]3([C@@H]([C@H]2[C@@H]([C]C(=O)[C@]12CC[C@@H]3(Cl=C(C([C@H]4[C@]13CC(=O)c2cc(Cl)O)C)C)C)C)C	676
	Oc1ccc(c1O)C@12CC[C@@H]3(Cl=C(C([C@H]4[C@]13CC(=O)c2cc(Cl)O)C)C)C	716
	搜索含该活性成分的中药	
	打开参考文献	
	在PubChem中搜索该化合物	
	计算理化性质	
	@[H]2[C@]1(C)CC[C@@H](C2(C)C(=O)O)C([C@H]2CC(CC1)C)C	333
	831	831
	Oc1cc2cc([C@H](CC(=O)c2cc1O)O)cccc1	868
	O[C@H]1CC[C@H]2(C(=CC[C@@H]3[C@@H]2CC[C@H]3CC[C@@H]2[C@H]([C]CC[C@H](C(C)C)CC)C)C	329
	Oc1(O)Cc2cc(c1O)C(=O)c2	941
	c1oc2cc(O)cc(c2c(=O)c1O[C@H]1O[C@H](CO[C@H]2O[C@H]([C]C[C@H]([C]([C@H]([C]([C@H]2O)O)O)O)C[C@H]([C@H]([C@H])	930

(5) 由化学成分找包含该成分的中药

使用方法：右键菜单->搜索包含该活性成分的中药。

例子：分析中药单体“柴胡”中包含 4 个抗流感化学成分，其中异甘草素也出现在其他的 11 个中药单体中。

The figure shows a software interface for analyzing the chemical components of Catechins. On the left, the chemical structure of Catechins is displayed, consisting of two hydroxylated flavonol aglycones linked by a central carbon atom. The structure is shown in a 2D projection with red atoms representing oxygen.

In the center, a table lists traditional Chinese medicine (TCM) names and their corresponding chemical components:

中药名	成分中文名	成分英文名	抗流感活性
甘草	异甘草素	isoleucoplatannin	IC50 values of 9.0 μg/mL
黄芩	黄芩苷	baicalin	搜索会话活性成分的中药
天景红	堪非醇	kannaphenol	打开参考文献
	金丝桃苷	geniposide	在PubChem中搜索该化合物
	芦丁	quercetin	计算理化性质

A context menu is open over the entry for Baicalin (黄芩), listing options such as "Search active ingredients in the session", "Open reference文献", "Search the compound in PubChem", and "Calculate physicochemical properties".

On the right, a window titled "Single味药分析" (Analysis of Single TCM) shows a list of TCM names:

- 板蓝根
- 柴胡
- 垂盆草
- 甘草
- 葛花
- 黄精
- 苦杏仁
- 肉豆蔻
- 薤白
- 淫羊藿
- 灯盏细辛

Below this list is a section titled "参考文献" (References) which includes a link to "Jun Xu_Anti-Influenza" and other related research papers.

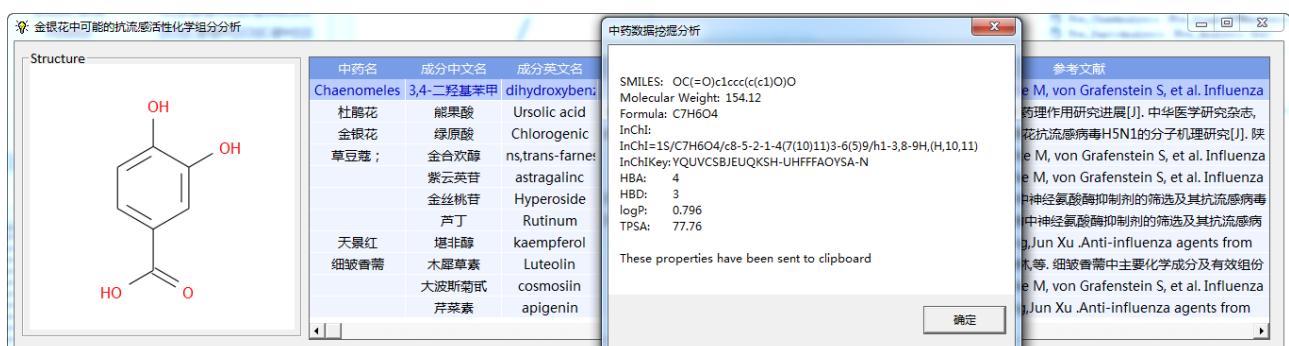
(6) 追踪活性化合物的参考文献

使用方法：右键->打开参考文献。可以打开相应中药化学组分活性数据的参考文献网址。

Cell Lines	Cancer	IC ₅₀	EC ₅₀	ED ₅₀	GI ₅₀	Reference
KC-7	Breast	-	-	-	181±20 μM	[16448184]
UL-1	Breast	15.0±0.3 μM	-	-	-	[16731741]
DX-171	Colon	-	-	9.2±0.3 μM	-	[16735446]
K-562	Hematopoietic and Lymphoid tissue	-	-	15±1.3 μM	-	[16735446]
A549	Lung	10.3±0.4 μM	-	-	-	[167387747]
A549	Lung	-	-	14±2.2 μM	-	[16735446]
PC-3	Prostate	-	-	15±0.5 μM	-	[16735446]
B16(F10)	Skin	-	-	14±0.2 μM	-	[16735446]
B16-F1	Skin	16.1±0.5 μM	-	-	-	[167387747]
SK-MEL-2	Skin	-	-	7.2±0.6 μM	-	[16735446]

(7) 计算化学成分的理化性质

可以计算分子基本理化性质，特别是类药性参数 logP、HBA、HBD 和分子量。



(8) 网络搜索中药信息

软件暂时设置了以下 6 个网站的搜索：“中医百科、中医药搜索、百度、百度学术、TCMSP、TM-MC”。使用方法：右键->网络搜索->中医百科



由于时间和能力的不足，本软件还存在很多不足之处，还请您批评指正！

我的 E-mail: mwxiao@hnu.edu.cn