

小青龙汤治疗咳喘的有效成分预测*

李思洋,朱芝娴,董 杰**

(南京中医药大学附属医院 南京 210001)

摘 要:目的:通过一种基于通路模式的方法,对传统中药名方小青龙汤治疗咳喘的有效成分进行预测。方法:首先搜索TCM中医药资料库与CTD数据库来整理出小青龙汤已知的各种化合物成分与相关基因,再通过DAVID生物分析系统与KEGG数据库挖掘出与咳喘相关的通路模式。然后利用该通路模式对候选基因打分,筛选出相关性高的基因,并据此推测出治疗咳喘的有效成分。结果:根据上述的计算方式,我们将相关基因根据得分进行排序,得到相关性最好的三种基因分别为MAPK1、MAPK3与RELA。最终预测出的有效成分中效能最高的三种分别是Paeonol(丹皮酚)、Glycyrol(甘草醇)、Geraniin(老鹳草素)。结论:此种基于通路模式的方法可以较好的预测出小青龙汤治疗咳喘的有效成分,加以推广更有利于中药基础药理研究的标准化。

关键词:小青龙汤 数据挖掘 基因打分 有效性指数

doi:10.11842/wst.2018.02.016

中图分类号:R285

文献标识码:A

咳喘是呼吸道系统疾病的一种最基本的症状,可见于多种呼吸系统的疾病中,比如常见的呼吸道感染、支气管哮喘、肺心病、职业性肺病以及肺结核、肺癌等。随着气候环境的恶化以及当今社会生活习惯的改变,呼吸系统疾病的发病率逐年升高,而小青龙汤作为经验方剂常被用于治疗呼吸系统疾病^[1]。随着祖国医学走上世界舞台,越来越多的国内外学者关注于中医药治疗疾病的分子机制,这就需要我们探明各种中药成分的潜在药理学基础。本文中我们使用了一种基于信号通路模式的方法,将小青龙汤中所涉及的化合物成分以及相关基因的效能进行了评分比较,并对该方在治疗咳喘中发挥主要药理作用的成分进行预测。此研究方法可以作为探索和规范中药现代化应用的一种新思路。

小青龙汤最早见于东汉张仲景所著中医经典著作《伤寒论》,主治外感风寒、内停水饮。《伤寒论》第40条曰:“伤寒表不解,心下有水气,干呕,发热而咳,或渴,

或利,或噎,或小便不利,少腹满,或喘者,小青龙汤主之。”第41条曰:“伤寒心下水气,咳而微喘,发热不渴。服汤已渴者,此寒去欲解也,小青龙汤主之。”从仲景原文看,小青龙汤主症应以咳喘为宜,至于咳和喘孰轻孰重,则临床表现不一。在后续多本著作中均有记载,如《宋·太平惠民和剂局方》、明·方贤著《奇效良方》等,甚至日本的“汉方药”中亦有小青龙汤。该方也是目前临床应用比较多的方剂之一^[2]。但小青龙汤及其类方的相关文献报道仍主要集中在中医方证研究、临床治疗和整体动物药效学评价等,少量的药理学研究也仅限于对其中麻黄碱、伪麻黄碱、甲基麻黄碱、芍药苷、肉桂酸等若干含量较高的成分进行定性、定量分析,尚未见对小青龙汤进行系统性的化学成分辨识研究报道^[3]。

中药中通常包含了大量的化学物质,所以如果能够系统性的研究某首方剂或者某味中药,寻找到复方当中的有效单体成分,对于阐明中药治疗疾病的机理意义重大。为此,许多研究人员试图建立一种研究策略来探索各种药物成分如何影响药物的效用^[4]。目前

收稿日期:2017-12-11

修回日期:2018-01-19

* 国家自然科学基金青年基金项目(81700037):IL-35依赖的免疫调节T细胞iT(R)35对抗原特异性记忆T细胞的抑制在过敏性哮喘特异性免疫治疗中的作用和机制研究,负责人:董杰;江苏省中医院课题:IL-35在过敏原特异性免疫治疗中的免疫调节作用及调控机制研究(Y17002),负责人:董杰。

** 通讯作者:董杰,技师,博士,主要研究方向:中药成分基础药理标准化与现代化探索。

已有多种方法用于分析药物成分和预测药物效用。这些方法大致可分为两种,第一种是将化合物的生物活性与其化学结构和化学特征联系起来,被称作定量结构活性相关性的研究(Quantitative Structure-Activity Relationship analysis)^[5]。但是应用此种方法的缺陷是目前不能获得非常准确的中药药物结构信息。第二种是将复杂化学系统的生物活性与其化学组成相联系,被称为定量组份活性相关性的研究(Quantitative Composition-Activity Relationship analysis)^[6,7]。但是这种方法缺乏精确的计算模型来量化化学成分与其生物活性的关系。

本研究中我们使用一种基于通路模式的分析方法(见图1)^[4],从药物组成中分析活性成分,并预测其基因靶点,在其中寻找疾病发病机制中起着重要作用的基因靶点。再从选定的基因相关的KEGG通路经过数据挖掘进行检索,从而排定各基因在疾病发病过程中的重要性,提示这些活性成分可能在小青龙汤治疗哮喘中发挥主要作用。

1 方法

1.1 筛选小青汤中各味草药所包含的化合物成分

中药的成分复杂,要想系统了解某首方剂的药理基础,首先要收集该药方中各味草药包含的所有化合

物信息。这里我们所选用的小青龙汤方中包含了8味中草药,通过查询中医药资料库@Taiwan(<http://tcm.cmu.edu.tw/zh-tw/index.php>)收集了他们各自目前已知的所有化合物成分(共335个,见附表1)。

1.2 化合物成分的相关基因

为了确定上述这些化合物成分中哪些是对于治疗哮喘有作用的活性成分以及预测其相应的靶基因,我们首先通过搜索CTD数据库(<http://ctdbase.org>)找到可以与这些化合物有相互作用的相关基因。CTD数据库收录了几乎所有FDA批准药物和PubMed收录的小分子化合物及其作用的相关基因和信号通路。由于有些化合物比较罕见,CTD数据库中并没有录入或者没有提供相关信息,我们去掉这些化合物。最终我们在CTD数据库中发现上述335个化合物中有详细分子结构、基因和相关通路的共有50个。这50个化合物能够与被机体细胞识别并产生基因的变化,说明其具有潜在药物活性。进一步分析50个化合物的相关基因有5133个(详见附表2),这些基因涉及的信号通路有11740条。

1.3 通路模式挖掘

上一步骤中所发掘出来的50个化合物的所有相关基因可以排出一个候选的靶基因列表。在这个列表中,我们记录下所有的这些基因和他们的出现次数。

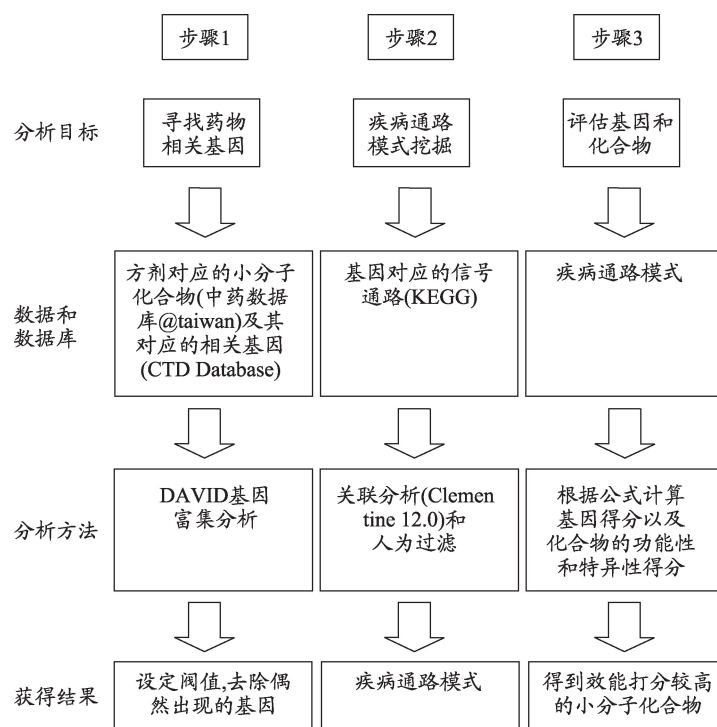


图1 药物活性成分和靶基因预测流程

由于有很多基因是偶然发生,因此可以说其与疾病联系不大,需要设定一个最低阈值来将这些基因排除。因此我们利用 DAVID 生物分析系统 (<https://david.ncifcrf.gov/>) 对哮喘相关基因进行富集^[8,9]。将基因出现频次相同的作为一组,每组基因号通过 ENTREZ_GENE_ID 进行校正后,以“哮喘”为关键词,由系统计算每组基因与疾病的相关性(由 P 值反映)。将基因与疾病相关性最为密切的一组基因的频次设定为阈值,出现频次小于阈值的基因将被舍弃。通过设定阈值筛选之后剩余的基因便可以进行通路信息的收集和研究了。这里我们通过搜索 KEGG 数据库 (<http://www.genome.jp/kegg/>) 将每一个基因对应所有的通路找出来。每一个基因对应一个通路的数据集。在这些数据集中,如果某些通路单独或成组地多次出现,那么这一对或一组信号通路便应该与我们所感兴趣的疾病有着莫大的关联^[4]。同时,我们人为地将与哮喘明显不相关的信号通路提出,保留与肺部疾病,细胞生长和周期以及免疫相关的信号通路。在一定的支持度和置信度设定下,利用 clementine 12.0 统计软件的 Apriori 模型对信号通路进行关联分析。这些在强双向关联规则下的通路被认为是高度相关的,并组成了我们所谓疾病通路模式。

1.4 通过通路模式计算基因得分

通过疾病通路模式便可以计算出一种相关基因的得分。而如果一个基因处于更多上述的强双向关联规则中,那么它理应在疾病的发病机制中发挥更重要的效用,此时它的得分便更高。假设强关联规则中存在两个、三个... k 个信号通路相关,或单个疾病相关信号通路的高频出现。疾病通路模式中单个、两个、三个... k 个信号通路关联规则包含的通路数目分别为 N_1 、 N_2 、 N_3 ... N_k 。

则某基因得分的计算公式为:

$$S_g = \frac{1}{k} \sum_{i=1}^k \frac{h_i}{N_i}$$

其中 h_i 表示在某一项关联规则中涉及的所有通路中,该基因所对应的通路个数。通过这个计算得分的办法,便可以将之前找到的基因根据在哮喘中的重要程度进行量化,而得分高于 0 的这些基因便是后面所谓的候选靶基因,并做成候选靶基因列表。

1.5 化合物的效能计算

为了衡量各种化合物的效用,我们定义了一个效能指标。该指标包括功能性与特异性两个方面。功能

性方面,我们假设某种化合物涵盖了越多的上述得分较高的候选靶基因,那么其功能就越好。使用函数

$fun_c = \frac{\sum_{j=1}^m I_j * S_{gj}}{s}$ 来计算。 m 为化合物对应的基因集中含有的候选靶基因数, s 是所有靶基因得分之和, I_j 表示候选靶基因列表中的第 j 个基因是否是化合物对应的基因集中含有的基因,是则为 1,否则为 0, S_{gj} 是第 j 个基因的基因得分。特异性方面,我们假设化合物对应的基因集中不在上述候选靶基因列表中的基因越少,特异性就越高。因此,可以使用函数 $spe_c = \frac{N_{gt}}{N}$ 来计算特异性。 N 是该化合物相关基因的总数量,而 N_{gt} 是相关基因中包含在上述候选靶基因列表中的数量。最终得到的功效指标定义为 $EI = \frac{1}{2} fun_c + \frac{1}{2} spe_c$ 。

当一个化合物的相关基因与候选靶基因列表完全重合的话,那么其 $fun_c = 1$ 、 $spe_c = 1$, EI 的得分也是最高分 1。 fun_c 和 spe_c 的下降,都会造成 EI 的分值下降。依据 EI 的进行排名,便可整理出方剂中高得分的那些化合物即为所要寻找的有效成分。这些最后得出的有效成分与上面所得到的靶基因一起便可以看作是小青龙汤在治疗哮喘病中潜在的基础药理机制。

2 结果

2.1 通过挖掘通路模式筛选基因

首先通过查找中药数据库@Taiwan,将小青龙汤中单味药所包含的小分子化合物列出(以芍药为例如表 1 所示)。药物成分所包涵的各种化合物相关基因数量非常庞大,因此需要通过挖掘通路模式来进行筛选。我们将 50 个化合物所有的相关基因罗列出来,共 5 133 个。将这些基因根据他们的发生频数进行排列。如同我们在图 2 中所展示出来的,发生频数高的基因数量是相对很少的,而可能因为偶然因素发生一次两次的这些基因却数量庞大。进一步使用 DAVID 生物分析系统对这些基因进行富集分析^[10]。即选择该系统中“GENETIC_ASSOCIATION_DB_DISEASE_CLASS”的分析,以“哮喘”为关键词,从而找出与我们要研究的哮喘以及肺部相关疾病高度相关的基因。该系统所提供的这种分析方法可以计算出这些基因的一个相关性指标 P 值(P 值越低,说明关系越密切)。随着基因发生频数的降低,基因与哮喘的发生关系越来越密切($-\ln P$ 值越高),最后发现在频数为 4 处的这组基因与

哮喘关系最为密切,如图3所示。发生频率在4以下的基因与哮喘关系逐渐疏远。所以我们将频数4设为阈值,即将基因频数为3或以下的基因舍弃。通过这种方式对所有的这5 133个基因进行筛选,最终得到了106个候选基因。

2.2 疾病的相关通路模式

将通过筛选得到的106个候选基因在KEGG数据库搜索,其中有着详细通路信息的基因有98个。这些基因的所有相关通路均记录在附表3中。这98个基因所对应的化合物也减少到47个。利用这些基因在KEGG数据库中所查到的通路信息,并设定了对应的支持度0.09与置信度0.9后,我们得到了包括了25个单项目通路、1个双项关联规则与6个多项关联规则的通路模式(见表2)。

2.3 基因得分

通过上面步骤所挖掘得出的通路模式,我们便可以计算所有这98个候选基因的基因得分(见附表4)。表3以CASP3基因为例,展示了根据公式计算基因得分的过程。其中有75个基因的得分在0分以上。将基因得分在前20的基因列在表4中,得分最高的三种基因分别为MAPK1、MAPK3与RELA。这些高得分的基因便被认为与哮喘高度相关,同时也有可能就是小青龙汤在治疗哮喘时所作用的靶基因。

2.4 化合物的效能指数

经过之前的筛选最后剩下来的候选化合物只剩下47个(见附表5)。根据这些化合物相关基因的基因得分,我们可以计算出这些化合物的功能函数值 fun_c 与特异函数值 spe_c ,并最终得出他们的效能指数EI。表5以萝卜硫苷(Glucoraphanin)为例,展示了化合物效能的计算过程。将这些化合物的EI值进行排名后我们在表6中列出了其中的前十名,以及包含了这些化合物成分的草药。其中效能最高的3种化合物为Paeonol(丹皮酚)、Glycyrol(甘草醇)、Geraniin(老鹳草素)。

3 讨论

本文中我们使用了一种基于所谓通路模式的方法对小青龙汤治疗哮喘的有效成分及其发挥药效的靶基因进行了预测。通过使用数据挖掘的方式搜集所有的成分、相关基因、通路等等大量的信息。再使用诸如戴维富集分析等方法设定一定的阈值来对所有的信息进行筛选,将很多偶然发生的基因和无关的信号通路剔除。最终得到需要的化合物成分和靶基因,并作为最

表1 芍药中的化合物单体成分

Compound name	中文名	CASRN
12-Methyl tetradecanoic acid	凝集酯	19207-26-2
4-Phenylbicyclo[2.2.2]octan-1-ol	无	2001-62-9
Albiflorin	芍药内酯苷	39011-90-0
Benzoic acid	苯甲酸	65-85-0
Eugenitin	甲基丁香色原酮	480-12-6
Galuteolin	木犀草苷	5373-11-5
Gamma-Sitosterol	伽马谷甾醇	83-47-6
Oxypeucedanin	氧化前胡素	737-52-0
Paeniflorin	芍药苷	23180-57-6
Paenilactone A	薄荷醇	98751-79-2
Paenonin	芍药素-3,5-二葡萄糖苷	132-37-6
Paenol	丹皮酚	552-41-0
Paenolide	牡丹酚原甙	72520-92-4
Pyrethrin II	除虫菊素	121-29-9
Pyrethrosin	除虫菊精	28272-18-6

注:通过查询CTD数据库得到每种单体化合物的CASRN号作为其唯一标识。

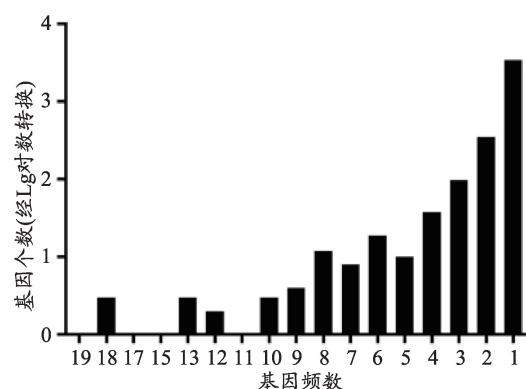


图2 50个化合物对应的相关基因出现的频数

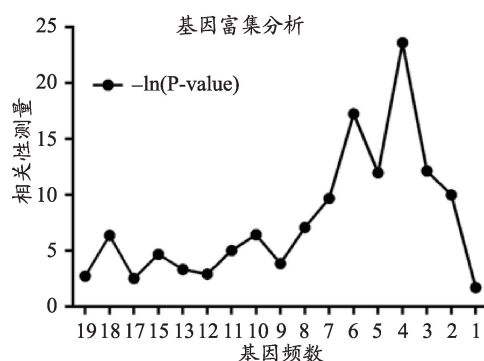


图3 通过DAVID生物分析系统进行基因的富集分析

注:选择哮喘及肺部相关疾病与候选基因谱中各基因进行相关性分析。使用一个计算出的P值来表达该相关性。其中P值越小则相关性越强。图中纵坐标使用 $-\ln(P\text{-value})$,即 $-\ln(P\text{-value})$ 越大则相关性越强。

表2 疾病相关的通路模式

KEGG 通路	发生次数	通路描述
hsa04210	26	Apoptosis
hsa04668	21	TNF signaling pathway
hsa04657	19	IL-17 signaling pathway
hsa01100	18	Metabolic pathways
hsa04151	18	PI3K-Akt signaling pathway
hsa04010	17	MAPK signaling pathway
hsa04218	17	Cellular senescence
hsa04621	17	NOD-like receptor signaling pathway
hsa04659	17	Th17 cell differentiation
hsa04115	16	p53 signaling pathway
hsa04620	16	Toll-like receptor signaling pathway
hsa04060	15	Cytokine-cytokine receptor interaction
hsa04064	15	NF-kappa B signaling pathway
hsa04217	15	Necroptosis
hsa04510	12	Focal adhesion
hsa04658	12	Th1 and Th2 cell differentiation
hsa04660	12	T cell receptor signaling pathway
hsa04062	11	Chemokine signaling pathway
hsa04215	11	Apoptosis - multiple species
hsa04630	11	Jak-STAT signaling pathway
hsa00140	9	Steroid hormone biosynthesis
hsa04110	9	Cell cycle
hsa04650	9	Natural killer cell mediated cytotoxicity
hsa04662	9	B cell receptor signaling pathway
hsa04920	9	Adipocytokine signaling pathway
hsa04210/hsa04215	11	Apoptosis/Apoptosis - multiple species
hsa04659/hsa04658/hsa04660	9	Th17 cell differentiation/Th1 and Th2 cell differentiation/T cell receptor signaling pathway
hsa04657/hsa04658/hsa04660	9	IL-17 signaling pathway/Th1 and Th2 cell differentiation/T cell receptor signaling pathway
hsa04659/hsa04658/hsa04657	9	Th17 cell differentiation/Th1 and Th2 cell differentiation/IL-17 signaling pathway
hsa04210/hsa04660/hsa04620	9	Apoptosis/T cell receptor signaling pathway/Toll-like receptor signaling pathway
hsa04620/hsa04659/hsa04621	9	Toll-like receptor signaling pathway/Th17 cell differentiation/NOD-like receptor signaling pathway
hsa04657/hsa04658/hsa04659/hsa04660	9	IL-17 signaling pathway/Th1 and Th2 cell differentiation/Th17 cell differentiation/T cell receptor signaling pathway

注:在使用支持度0.09与置信度0.9后得到的通路模式,包含了上述25个单项目关联规则、1个双项目关联规则、5个三项目关联规则与一个四项目关联规则。

后的预测结果。

在我们最后得到的结果中,有一些在以前的相关研究中已经报道过。比如我们计算的出EI分值较高,在所有化合物排名第九的山柰酚(Kaempferol),就被发现可以通过调节NF-κB信号通路可以减轻流感造成的呼吸系统损伤,同时有效降低了小鼠肺损伤模型血液中肿瘤坏死因子(TNF)、白细胞介素6(IL-6)的含量,二者也正是我们所预测小青龙汤发挥药效靶基因(即基因得分排名前二十)^[11]。同样的结论在一些慢性阻塞性肺疾病和慢性支气管炎的研究中也有报道^[12,13]。当然这些呼吸系统的疾病均包含在中医的咳证与喘证中^[14]。而没有文献报道的这些化合物与靶基因也可能成为以后对于小青龙汤治疗咳喘的基础药理机制研究的方向。

本研究中也获得了一些既与小青龙汤中的小分子化合物相关联,又与咳喘发生相关的基因,小青龙汤对

表3 CASP3(ID:836)基因得分

关联规则	通路编号	k	Ni	hi	得分 1/k*(hi/Ni)
包含通路	has01524 has04010 has04115 has04210 has04215 has04650 has04657 has04668 has04726 has04932 has04933 has05010 has05012 has05014 has05016 has05120 has05133 has05134 has05145 has05146 has05152 has05161 has05165 has05167 has05168 has05200 has05203 has05205 has05206 has05210 hsa05416				
单项规则	has04010 has04115 has04210 has04215 has04650 has04657 hsa04668	1	25	7	0.28
双项关联规则	has04210 hsa04215	2	2	2	0.5
三项关联规则	hsa04657	3	6	1	0.055556
四项关联规则	hsa04657	4	4	1	0.0625
总计					0.898056

表4 排序前二十的相关基因

基因号	基因名称	基因得分
5594	MAPK1	1.565278
5595	MAPK3	1.565278
5970	RELA	1.315278
4790	NFKB1	1.315278
1147	CHUK	1.275278
4792	NFKBIA	1.195278
3725	JUN	1.155278
841	CASP8	1.049167
207	AKT1	1.035556
2353	FOS	1.019722
7124	TNF	0.959167
836	CASP3	0.898056
596	BCL2	0.875556
329	BIRC2	0.835556
331	XIAP	0.795556
355	FAS	0.78
7157	TP53	0.74
3569	IL6	0.707222
3553	IL1B	0.707222
637	BID	0.7

表5 萝卜硫苷(Glucoraphanin)效能(EI)计算

	基因编号	基因集包含情况	基因得分
化合物 对应基因	2877	否	0
	1728	否	0
	1543	是	0.08
	1544	是	0.08
	1545	是	0.04
	2052	否	0
	4780	否	0
功能性(fun)			(0.08+0.08+0.04)/ 67.064
特异性(spe)		3/7	
化合物效能(EI)		0.5*fun+0.5*spe=0.2173	

注:所有靶基因得分总和为 67.064

哮喘的治疗作用可能依赖于这些信号分子和相关信号通路,如丝裂原活化蛋白激酶(MAPK)信号通路。实际上,已经有研究发现,一些中药方剂主要是通过抑制 MAPK 信号,改善小鼠哮喘症状^[15]。

当然这种研究方法也有很多缺陷,比如所建立的通路模式缺乏生物学价值。这种通路模式其实是通过数据挖掘的方式,从大量的信息中统计出来的,它更多

表6 效能指数EI前十名的化合物

CASRN	化合物(Compound name)	中文名	对应的单味中药	效能指数(EI)
552-41-0	Paeonol	丹皮酚	芍药	0.568778719
23013-84-5	Glycyrol	甘草醇	甘草	0.543656474
60976-49-0	Geraniin	老鹳草素	干姜	0.53390708
520-36-5	Apigenin	芹菜素	麻黄	0.525707581
124-18-5	Decane	癸烷	干姜	0.518137732
1391-73-7	Gingerol	姜辣素	干姜	0.515390343
57-13-6	Urea	尿素	甘草	0.510747237
93-91-4	1-Phenyl-1,3-butanedion	1-苯基-1,3-二丁酮	五味子	0.50759225
520-18-3	Kaempferol	山柰酚	麻黄	0.503520159
117038-80-9	Licopyranocoumarin	甘草吡喃香豆素	甘草	0.50119289

的是符合统计学意义,而在生物学上还不能找到合适的解释。所以寻找其生物学意义也可作为后面研究的一个重点和方向。并且由于所研究化合物的详细成分组成和基因资料并不完全,研究过程中排除的很多化合物在治疗疾病的过程中可能也起到了很重要的作用,但以现在的资料储备水平我们也不得而知了。另外在本研究的基础上,还需进一步进行动物实验,探索

本研究中预测为治疗效能较高的化合物单独或者联合使用,阐明其是否能够治疗或改善哮喘症状。

总的来说,这项研究通过数据挖掘这一利用当今网络大数据的方法,成功预测了小青龙汤治疗哮喘病的很多有效成分及其作用的靶基因。未来,如果可以使用这种方法探索更多的传统中药名方,那对于中药治疗基础药理研究的标准化有着重要意义。

参考文献

- 1 张立山, 戴雁彦. 小青龙汤应用之我见. 中华中医药杂志, 2008, 23(09): 847-848.
- 2 朱晓宏. 小青龙汤现代药理研究及临床应用概况. 国医论坛, 2007, 22(6): 52-53.
- 3 范晓辉, 肖舜, 艾妮, 等. 基于网络方剂学的小青龙汤类方功效物质组研究. 中国中药杂志, 2015, 40(13): 2634-2638.
- 4 Dai W, Chen J, Lu P, *et al.* Pattern-based prediction of active drug components and gene targets from H1N1 influenza's treatment with maxingshigan-yinqiaosan formula. *Mol Biosyst*, 2013, 9(3): 375-85.
- 5 Burbidge R, Trotter M, Buxton B, *et al.* Drug design by machine learning: support vector machines for pharmaceutical data analysis. *Comput Chem*, 2001, 26(1): 5-14.
- 6 Wang Y, Wang X, Cheng Y. A computational approach to botanical drug design by modeling quantitative composition-activity relationship. *Chem Biol Drug Des*, 2006, 68(3): 166-172.
- 7 Cheng Y, Wang Y, Wang X. A causal relationship discovery-based approach to identifying active components of herbal medicine. *Comput Biol Chem*, 2006, 30, 148-154.
- 8 Huang da W, Sherman B T, Lempicki R A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res*, 2009, 37(1): 1-13.
- 9 Huang da W, Sherman B T, Lempicki R A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*, 2009, 4(1): 44-57.
- 10 Chen L, Tai J, Zhang L, *et al.* Global risk transformative prioritization for prostate cancer candidate genes in molecular networks. *Mol Biosyst*, 2011, 7(9): 2547-2553.
- 11 李妍, 王春富, 张瑞华, 等. 山柰酚通过下调NF- κ B信号通路减轻猪源甲型H9N2流感病毒所致小鼠急性肺损伤. 中国病理生理杂志, 2017, 33(2): 315-321.
- 12 刘禹翔, 王峰, 曲敬来, 等. 小青龙汤对慢性阻塞性肺疾病急性发作期患者细胞因子及肺功能的影响. 新中医, 2013, 45(7): 24-26.
- 13 孟学峰, 范晔, 薛连峰. 小青龙汤加味改善慢性支气管炎急性发作期患者血清IL-6和TNF- α 的研究. 中医研究, 2009, 22(6): 27-29.
- 14 王有奎. 呼吸病的中医诊治与调理. 北京: 人民军医出版社, 2007.
- 15 罗永峰, 吴壮, 徐军. 天龙咳喘灵改善慢性哮喘小鼠气道重塑的机制. 辽宁中医杂志, 2011, 38(2): 357-359.

Bioinformatics Prediction of Active Components in Xiao-Qing-Long Decoction against Symptoms of Cough and Dyspnea

Li Siyang, Zhu Zhixian, Dong Jie

(The Affiliated Hospital of Nanjing University of Chinese Medicine, Nanjing 210001, China)

Abstract: This study was aimed to identify active components from traditional Chinese medicine (TCM) formula *Xiao-Qing-Long* (XQL) decoction against cough and dyspnea through a pathway pattern-based method. All compounds contained in candidate herbs and associated genes of XQL decoction were collected from TCM Database and CTD database. And then, the pathway pattern of cough and dyspnea was revealed using the DAVID Functional Annotation Bioinformatics Microarray Analysis and KEGG database. The extracted pathway pattern was used to calculate gene scores, which was employed to reflect the importance of genes to the disease. The effectiveness of compounds was

calculated through matching associated genes of compounds to the disease associated gene target list. The results showed that based on analysis mentioned above, the top three ranked genes closely related to the disease were MAPK1, MAPK3 and RELA. The compounds with top three efficacy indices were paeonol, glycyrol and geraniin. It was concluded that the pathway pattern-based method can better predict active components from XQL decoction, which may be effective as a small molecule drug against the symptom of cough and dyspnea. The data mining method would be beneficial for standardization and modernization of TCM.

Keywords: *Xiao-Qing-Long* decoction, data mining, gene scores, efficacy indices

(责任编辑:张 静,责任译审:王 晶)