基于爬虫技术的中药分析平台

# 摘要

网络药理学在预测和辨识中药活性成分群及作用靶点上有重要应用。基于网络药理学的中药研究,需要进行中药成分准备，疾病靶点准备，使用工具进行分子对接等关键步骤。但是现阶段还没有一个集成的，可以从成分获取、靶点获取到分子对接的平台，导致网络药理学研究人员在进行研究时，需要花费大量的时间。为此我们提供了一个基于自动化技术的中药网络药理学在线分析平台，平台集成中药成分的准备、疾病靶点的准备、在线对接三大模块。第一个模块支持中药小分子成分以及成分对应的3D结构文件的获取。第二个模块会根据提供的疾病名称，自动搜索相关数据库完成疾病靶点、靶点蛋白的获取。第三个模块根据前述得到的中药成分、疾病靶点结果进行在线批量分子对接，并提供对接结果。本文提供了一个研究案例，演示了本系统如何获取中药成分的活性成分、疾病的关键靶点以及中药成分和靶点之间的作用关系。

**关键词** 爬虫 中药 网络药理学 预测分析

# 背景

网络药理学是建立在疾病-基因-药物的多层次网络基础上，用于从整体上预测药物靶点、提高药物发现效率的新兴学科。网络药理学技术在中药研究的综合分析中得到了越来越多的应用，并且实践证明，网络药理学技术对中药复杂体系的整体性、系统性分析有独特优势。网络药理学在中药研究中的关键步骤，包括中药成分的准备，中药所治疗疾病靶点的获取，中药成分和靶点之间的相互作用-分子对接。对接得到的结果可以为后续网络分析提供重要的数据支撑。现阶段基于网络药理学的中药研究，需要从公开的数据库中尽可能完整地获取到中药成分信息，靶点的获取也要根据疾病的关键字进行搜寻，最后使用相关的对接工具进行分子对接。

中药成分数据库,如中药系统药理学分析平台（TCMSP）、中药综合数据库（TCMID）等被广大研究者广泛采用，这些数据库可以提供较为全面的中药成分信息，但是需要研究者手动查找并准备。另外，后续分子对接需要使用到化合物成分的结构信息，需要逐条下载，这使得研究者在成分的准备上需要花费大量的时间。疾病靶点的准备需要从NCBI-Gene等网站手动选择并逐条下载，同样过程耗时复杂。分子对接技术按照几何互补、能量互补以及化学环境互补的原则，进行中药成分化合物小分析和靶点蛋白大分子作用识别，并预测2个分子之间的最佳结合模式。现阶段DOCK、AutoDock、Discovery Studio、AutoDock Vina等软件可以进行分子对接，但是这些软件是C-S架构，同时每次对接仅能进行一个成分-一个靶点蛋白。另外，他们的打分函数不能直接反应对接的程度。Systems Dock是一个基于B-S架构的在线分子对接平台，该平台基于AutoDock Vina的，重新设计了打分函数，同时可以根据打分评价分子和靶点结合的好坏。但是该平台限制了每次对接的成分个数。

现在，中药成分准备、靶点准备、分子对接均处于独立的状态。本文基于网络药理学的基本原理，打通三者之间的关系，集成了多个数据库和工具，开发了一个在线的中药分析平台，实现从中药成分准备，疾病靶点准备，分子对接一体化的操作。

本系统设计的中药成分的准备是根据药材的名称，从TCMSP数据库中搜索到该药材的全部成分信息，包括药材名称、Molecule Name、PubChem Cid，首先将这些数据获取到，然后根据PubChem Cid在PubChem数据库中搜索到对应3D结构文件 。这本系统获取中药成分文件的基本操作，从其他数据库中获取成分的操作是类似的。靶点获取的模块就是从NCBI-Gene数据库中输入疾病英文名称，然后限定物种为人类，将对应的所有靶点名称/ID获取到并保存再来，然后将每一个靶点名称一次输入到UniProtKB数据库然后按照分辨率筛选靶点蛋白PDB ID。分子对接，使用System Dock，一次性输入全部的PDB ID，再将成分的3D文件上传到对接系统中，完成对接。系统最后会把中药成分的所有信息综合成起来，最终打分结果表包括中药名称、小分子化合物名称、PubChem Cid、靶点UniprotID、Protein Name、Gene、PDB ID、打分Scores。

本文会从系统设计的基本构架，每个部分的实现细节和操作方式以及案例研究等几个部分介绍相关工作。

本文的结构是这样的。在第二部分，我们会介绍本系统的系统设计

# 系统设计与处理方式

本系统基于网络药理学技术的一般流程设计，主要分为三个部分，包括（1）中药成分的获取，包括草药的名称和非草药的数据excel格式上传。2）输入疾病的英文名称 。3)小分子化合物成分和靶点进行对接。下面是系统的结构图，结构图展示了整个数据的处理流程。

|  |
| --- |
| 数据流图 |
| 基于自动化技术的在线分析系统架构 |

为了使用这个基于网络药理学的中药分析预测平台，整个过程需要分成三个主要的步骤，包括（i）如果是由几种草药组成的药剂，需要输入药材的中文名称，同时使用中文的逗号字符隔开，如果还存在非草药成分，需要用户自己从相关文献中查找并上传excel文件，上传的文件命名要以药材的中文或者是英文名字名称，列名必须为molecule,对应的每列是中药成分的英文名称。（ii）准备靶点需要输入靶点的英文名称，例如子宫腺肌症的英文名字adenomyosis（iii）进行对接并且检查对接结果。

## 中药化合物分子获取

现阶段的中药物成分的获取使用开源的数据库，例如TCMSP、TSMID等。但是，中药方剂由多种药材组成，每种药材成分复杂，单一的数据库不能保证数据完整，多个数据库的数据之间也存在异构，需要合并去重，方剂中也可能存在非草药。文献[]中提出，可以根据药物的口服利用度（OB）和药物相似性（DL）来预筛选药物的活性成分，但这会导致某些已经被验证的关键活性成分缺失，所以本系统并未按照这两部分来限制筛选成分。

本系统选择从TCMSP、TCMID数据库中查找小分子化合物，然后将从两个数据库中得到的数据合并。在获取到小分子数据后，再从PubChem中数据库中查找3D结构文件。由于系统需要的关键数据是药物的名称、成分名称、成分的PubChem Cid,成分的3D结构，对接需要小分子化合物的3D文件，参阅相关参考文献中的介绍，这里挑选存在3D结构的化合物。

从TCMSP和TCMID中获取中药化合物小分子的数据设计的总体思路如下所示，首先需要根据名称获取到成分的主页面url，然后解析得到每个成分对应的url，获取靶点后再在PubChem数据库中按照搜索到的PubChem Cid获取到对应的3D结构并下载。

下面是算法的伪代码：

输入：药物的名称

Step 1:在网页首页输入药物名称，查询获取主页面

Step 2:访问主页面遍历网页，获取到该药物对应的所有成分

Step 3:分别获取成分的名称、PubChem Cid

输出：中药对应化合物的PubChem Cid、化合物名称

下面是算法的流程图

|  |
| --- |
| 获取药品  解析成分  获取成分 |
|  |
|  |

Figure 2:中药信息数据获取算法流程

在上述获取到化合物的小分子数据后，还需要获取化合物的3D结构文件，Figure 3是中药化合物分子获取数据流图，展示了从TCMSP、TCMID中获取小分子数据，并最后获取化学物3D结构。

|  |
| --- |
|  |
| 中药化合物分子获取数据流图 |

在最终获取到成分3D结构的同时，需要把药物名称、化合物名称、PubChem CID等信息存储到数据库中，这里3D结构的命名与PubChem Cid保持一致。

对于非草药，系统允许文本(excel)上传，系统会根据上传文件的成分名称信息，从PubChem数据库中查找成分的PubChem Cid,并下载对应的成分的3D文件。然后存储到数据库中。

该部分需要用户在用户框中输入中药的名称，名称确保是中文，并且使用中文逗号分隔，然后提交。如果有非草药成分，再单独提交excel文件，然后提交。Excel文件的名称是中药的名称，列名必须为molecule,对应的每列是中药成分的英文名称。

## 从多源基因数据库获取靶点

疾病靶点的准备，文献中介绍了从GenBank、GeneCards、Swiss Target Prediction、NCBI-Gene等数据库可以获取到特定疾病对应的靶点。其中，NCBI-Gene数据库为较多的研究者所使用。在NCBI-Gene数据库中，输入疾病的英文名称，然后限定物种为人类，获取相应的靶点基因名称/ID。再从DisGeNET输入疾病的英文名称， 选择diseases，然后下载对应的基因数据，选择Gene-Disease Score打分大于0.2的靶点作为可用靶点。将从NCBI-Gene和DisGeNET获取到的疾病靶点名称合并。

根据基因名称/ID，从UniProtKB数据库中获取到对应的UniprotID和靶蛋白PDB ID。根据成分的名称信息在UniProtKB数据库中进行查找的时候，可能出现对应靶点蛋白不存在的情况，这种数据会被筛掉；对于存在多个靶点蛋白的数据，由于数据量比较大，为了选择较为合理的靶点蛋白，会根据靶点蛋白的分辨率的大小进行排序，选择分辨率较小的蛋白质作为备用的靶点蛋白。系统最后会按照这个方式把靶点的名称、靶点蛋白名称、UniProtID、PDB ID都存储到对应的数据库中。从而完成疾病靶点的准备。下图是疾病靶点数据获取流程图，图中展示了从多源数据库中获取疾病靶点及其分子结构的流程。

|  |
| --- |
|  |
| 疾病靶点数据获取流程图 |

该模块，用户仅需要在输入疾病的名称即可，名称只能输入一个疾病对应的英文名称。

## 在线批量分子对接

分子对接是网络药理学分析的一个关键步骤。传统的对接工具有Dock,

Autodock,Discovery Studio,Autoduck Vina等，用于进行化合物小分子和靶点蛋白之间的对接，并根据对接打分来判断小分子和蛋白质大分子之间的结合程度。上述的对接工具都是基于B-S架构，数据需要在本地进行处理。虽然这些对接软件的打分能反应分子之间的结合程度，但是给出的打分规则都是自己制定，导致打分跨度很大，不能很好地从打分上评价对接的好坏。另外文献中提出了一些在线的工具，例如PharmMapper,可以依据反向药效团匹配方法预测活性成分的作用靶点。从小分子化合物出发，使用反向找靶技术，得到的结果类似。

System Dock是一款在线的分子对接平台，该平台可以根据用户提供的化合物成分，疾病的靶点蛋白进行在线分子对接，同时还可以给出对接的结果，并提供对接后的文件。该平台底层使用的是AutoDock Vina分子对接工具，同时该平台在AutoDock Vina的打分基础上，使用docK-IN打分算法，对原始的打分进行了重新计算，使得打分结果能直接描述结合强度。一般打分值在4.25 以上认为是有一定的结合活性，大于5.0 的说明分子与靶点有较好的结合活性，大于7.0 则说明分子与靶点的结合构型具有强烈的活性。

本系统封装了System Dock，可以直接在中药化合物成分和靶点成分获取的基础上，给出对接结果。原始在线软件每次仅能提供5个化合物成分对接，本系统通过设计，可以一次性批量完成所有成分的对接。另外，原始对接结果仅给出了对接化合物成分的编号和靶点的PDB ID，本系统在该功能的基础上，化合物成分名称，所属中药，靶点的UniprotID，Protein name，Gene name等也会给出。

该部分的操作仅需前两部分完成后，输入邮箱地址和对应的提示信息，然后执行操作。

# 案例研究

基于本系统，我们研究了散结镇痛胶囊对子宫肌腺症的治疗，从而分析中药成分对疾病对应的靶点的作用情况。

首先，使用中药成分模块，根据资料查阅到散结镇痛胶囊的成分为龙血竭、三七、浙贝母、薏苡仁，其中三七、浙贝母、薏苡仁都是草药，这个可以直接使用系统的关键字查找功能来获取化合物成分，龙血竭不是草药，这种药物的成分就需要从文献中查找，然后使用系统的文本上传功能来查找成分。

然后根据治疗的疾病，子宫肌腺症，其对应的英文为adenomyosis,在系统的疾病靶点模块输入这种疾病对应的疾病名称。

前两步完成后，执行对应的对接操作，等待对接完成。下载对接数据后，分析下载数据。下载系统提供的3张数据表，分别是中药成分数据表，靶点信息表和对接结果信息表。根据下载的3个表，可以得到对应的中药成分信息，靶点信息以及对接结果信息。

根据中药成分信息表可知，这四种中药材对应的成分去重之后得到种103种成分，其中 浙贝母有9种，三七有61种，薏苡仁有21种，龙血竭有20种。下表是该胶囊的成分表

|  |  |  |
| --- | --- | --- |
| **molecule** | **pubchemcid** | **drug** |
| dracaenogenins A | 101389811 | 龙血竭 |
| Cadalin | 10225 | 三七 |
| Daturic acid | 10465 | 三七 |
| oleanolic acid | 10494 | 三七 |
| MBOA | 10772 | 薏苡仁 |
| Heptadekan | 10887 | 三七 |
| myristic acid | 11005 | 三七,薏苡仁 |
| protopanoxadiol | 11213350 | 三七 |
| DFV | 114829 | 三七 |
| Oktadekan | 11635 | 三七 |
| vanillin | 1183 | 薏苡仁 |
| ginsenoside rh2 | 119307 | 三七 |
| tachioside | 11962143 | 龙血竭 |
| 1-Hydroxycumene | 12053 | 三七 |
| 5-METHYLFURFURAL | 12097 | 三七 |
| Ditertbutyl phthalate | 121712 | 三七 |
| Î±-copaene | 12303902 | 三七 |
| isoarborinol | 12305182 | 薏苡仁 |
| Î±-cyperene | 12308843 | 三七 |
| Verticine | 131900 | 浙贝母 |
| trans-feruloylcampesterol | 13786591 | 薏苡仁 |
| 2-ACETYLPYRROLE | 14079 | 三七 |
| ()-beta-Pinene | 14896 | 三七 |
| 2-methoxy-4,4'-dihydroxychalcone | 14977366 | 龙血竭 |
| Peimisine | 161294 | 浙贝母 |
| 2,6-dimethyl-cyclohexanol | 164822 | 三七 |
| WLN | 16627 | 三七 |
| Peiminine | 167691 | 浙贝母 |
| campesterol | 173183 | 薏苡仁 |
| Zhebeiresinol | 192547 | 浙贝母 |
| METHYL ISOPALMITATE | 21205 | 三七 |
| 7-hydroxy-3'-methoxy-4'-butoxyflavone | 21597638 | 龙血竭 |
| Omaine | 220401 | 薏苡仁 |
| sitosterol | 222284 | 三七,浙贝母,薏苡仁 |
| cochinchinenin B | 23634523 | 龙血竭 |
| 7,4'-dihydroxy-3'-methoxyflavan | 25232840 | 龙血竭 |
| Ziebeimine | 3083151 | 浙贝母 |
| butylated hydroxytoluene | 31404 | 三七 |
| isopulegone | 34645 | 三七 |
| 2-ethyl-3-hydroxyhexyl butyrate | 347867 | 薏苡仁 |
| caprylic acid | 379 | 三七,薏苡仁 |
| lauric acid | 3893 | 三七 |
| Diop | 395120 | 三七 |
| NSC 308879 | 4212 | 三七 |
| 7-hydroxy-4'-methoxyflavan | 442277 | 龙血竭 |
| 4'-methoxy-3',7-dihydroxyflavone | 44259961 | 龙血竭 |
| syringaresinol | 443023 | 浙贝母 |
| ergosterol | 444679 | 薏苡仁 |
| farnesol | 445070 | 三七 |
| resveratrol | 445154 | 龙血竭 |
| zoomaric acid | 445638 | 三七 |
| 7,4'-dihydroxyflavone | 45116976 | 龙血竭 |
| 4-Hydroxybenzoic acid | 4649 | 龙血竭 |
| EIC | 5280450 | 薏苡仁 |
| Stigmasterol | 5280794 | 三七,薏苡仁 |
| quercetin | 5280804 | 三七 |
| Hexenal | 5281168 | 三七 |
| Gaidic acid | 5282743 | 薏苡仁 |
| 10Z,13Z-nonadecadienoic acid | 5282804 | 三七 |
| panaxydol | 5283280 | 三七 |
| METHYL LINOLEATE | 5284421 | 薏苡仁 |
| cochinchinenin A | 5315988 | 龙血竭 |
| Germacrene D | 5317570 | 三七 |
| 5,7,4'-trihydroxyflavone | 5322064 | 龙血竭 |
| 2'-methoxy-4,4'-dihydroxychalcone | 5355594 | 龙血竭 |
| Butylcyclobutane | 544591 | 三七 |
| NSC692928 | 5469789 | 三七 |
| hexanoic acid | 564 | 三七,薏苡仁 |
| Sitogluside | 5742590 | 三七 |
| PEL | 5858445 | 三七 |
| hexanal | 6184 | 三七,薏苡仁 |
| 4-methyl-7-encholyl alcohol | 637516 | 龙血竭 |
| 2-Coumarate | 637540 | 三七 |
| Glycosides | 637579 | 龙血竭 |
| (-)-alpha-cedrene | 6431015 | 三七 |
| Linolenyl alcohol | 6436081 | 三七 |
| Solatubin | 65727 | 浙贝母 |
| 2,4,4'-trihydroxydihydrochalcone | 68146761 | 龙血竭 |
| OSI-2040 | 6918328 | 浙贝母 |
| beta-elemene | 6918391 | 三七 |
| Octadecyne | 69425 | 三七 |
| 7,4'-dihydroxydihydroflavone | 7 | 龙血竭 |
| 3,4-dichloroaniline | 71413934 | 三七 |
| Hypnon | 71750952 | 三七 |
| DICHLOROANILINE | 7257 | 三七 |
| panaxatriol | 73599 | 三七 |
| Piceol | 7469 | 三七 |
| hydroquinone | 785 | 龙血竭 |
| PTL | 8063 | 三七 |
| Hepanal | 8130 | 三七 |
| nonanoic acid | 8158 | 薏苡仁 |
| methyl palmitate | 8181 | 三七 |
| Cyclooctadiene | 82916 | 三七 |
| Protocatechuic aldehyde | 8438 | 龙血竭 |
| MEHQ | 9015 | 三七 |
| Friedelin | 91472 | 薏苡仁 |
| Picein | 92123 | 三七 |
| panaxytriol | 93484 | 三七 |
| 3691-11-0 | 94275 | 三七 |
| Sitosterol alpha1 | 9548595 | 薏苡仁 |
| 3,5-Dimethoxyacetophenone | 95997 | 三七 |
| palmitic acid | 985 | 三七,薏苡仁 |
| ginsenoside f2 | 9852086 | 三七 |

根据得到的靶点成分信息，得到该疾病对应的靶点有56个，如下所示。

|  |  |  |  |
| --- | --- | --- | --- |
| **UniProtID** | **Protein** | **Gene** | **PDBID** |
| P08253 | 72 kDa type IV collagenase | MMP2 | 3AYU |
| P15692 | Vascular endothelial growth factor A | VEGFA | 1MKK |
| P03372 | Estrogen receptor | ESR1 | 3CBP |
| P35222 | Catenin beta-1 | CTNNB1 | 3FQN |
| P06401 | Progesterone receptor | PGR | 1SQN |
| P14780 | Matrix metalloproteinase-9 | MMP9 | 6ESM |
| P10451 | Osteopontin | SPP1 | 3CXD |
| P01138 | Beta-nerve growth factor | NGF | 1WWW |
| Q13651 | Interleukin-10 receptor subunit alpha | IL10RA | 1Y6K |
| P35354 | Prostaglandin G/H synthase 2 | PTGS2 | 5F19 |
| P21964 | Catechol O-methyltransferase | COMT | 3BWY |
| Q92731 | Estrogen receptor beta | ESR2 | 3OLL |
| P05106 | Integrin beta-3 | ITGB3 | 1MIZ |
| P03956 | Interstitial collagenase | MMP1 | 1HFC |
| P16035 | Metalloproteinase inhibitor 2 | TIMP2 | 1BR9 |
| P04629 | High affinity nerve growth factor receptor | NTRK1 | 5JFW |
| P08138 | Tumor necrosis factor receptor superfamily member 16 | NGFR | 3EWV |
| P08887 | Interleukin-6 receptor subunit alpha | IL6R | 1N26 |
| Q16665 | Hypoxia-inducible factor 1-alpha | HIF1A | 4H6J |
| P19838 | Nuclear factor NF-kappa-B p105 subunit | NFKB1 | 1SVC |
| P42771 | Cyclin-dependent kinase inhibitor 2A | CDKN2A | 1BI7 |
| P29474 | Nitric oxide synthase, endothelial | NOS3 | 4D1P |
| O00206 | Toll-like receptor 4 | TLR4 | 2Z62 |
| P60484 | Phosphatidylinositol 3,4,5-trisphosphate 3-phosphatase and dual-specificity protein phosphatase PTEN | PTEN | 4O1V |
| P10145 | Interleukin-8 | CXCL8 | 4XDX |
| P37231 | Peroxisome proliferator-activated receptor gamma | PPARG | 3U9Q |
| P09488 | Glutathione S-transferase Mu 1 | GSTM1 | 1XW6 |
| P42345 | Serine/threonine-protein kinase mTOR | MTOR | 4DRI |
| P24385 | G1/S-specific cyclin-D1 | CCND1 | 2W96 |
| P38936 | Cyclin-dependent kinase inhibitor 1 | CDKN1A | 5E0U |
| P13500 | C-C motif chemokine 2 | CCL2 | 1DOK |
| O15392 | Baculoviral IAP repeat-containing protein 5 | BIRC5 | 2QFA |
| P46531 | Neurogenic locus notch homolog protein 1 | NOTCH1 | 5L0R |
| O95256 | Interleukin-18 receptor accessory protein | IL18RAP | 3WO4 |
| P17693 | HLA class I histocompatibility antigen, alpha chain G | HLA-G | 3KYO |
| P06400 | Retinoblastoma-associated protein | RB1 | 2R7G |
| Q15109 | Advanced glycosylation end product-specific receptor | AGER | 5D7F |
| P04798 | Cytochrome P450 1A1 | CYP1A1 | 4I8V |
| P09038 | Fibroblast growth factor 2 | FGF2 | 4OEE |
| P08254 | Stromelysin-1 | MMP3 | 1HY7 |
| P01033 | Metalloproteinase inhibitor 1 | TIMP1 | 3V96 |
| P05093 | Steroid 17-alpha-hydroxylase/17,20 lyase | CYP17A1 | 5IRQ |
| P09237 | Matrilysin | MMP7 | 2Y6D |
| O95863 | Zinc finger protein SNAI1 | SNAI1 | 4QLI |
| Q99497 | Protein/nucleic acid deglycase DJ-1 | PARK7 | 2RK3 |
| Q9Y6K1 | DNA (cytosine-5)-methyltransferase 3A | DNMT3A | 4QBS |
| P13591 | Neural cell adhesion molecule 1 | NCAM1 | 2HAZ |
| Q13464 | Rho-associated protein kinase 1 | ROCK1 | 3V8S |
| Q9UBC3 | DNA (cytosine-5)-methyltransferase 3B | DNMT3B | 5NRR |
| Q969J5 | Interleukin-22 receptor subunit alpha-2 | IL22RA2 | 3G9V |
| O00622 | Protein CYR61 | CYR61 | 4D0Z |
| P05230 | Fibroblast growth factor 1 | FGF1 | 1RG8 |
| P42702 | Leukemia inhibitory factor receptor | LIFR | 3E0G |
| Q9NZH6 | Interleukin-37 | IL37 | 5HN1 |
| O96013 | Serine/threonine-protein kinase PAK 4 | PAK4 | 2J0I |
| P49757 | Protein numb homolog | NUMB | 5NJJ |

根据所有的打分结果，一般Score 值在4.25 以上认为是有一定的结合活性，大于5.0 的说明分子与靶点有较好的结合活性，大于7.0 则说明分子与靶点的结合构型具有强烈的活性。

根据数据，得到打分在7分以上的有566个，打分在5-7分之间的2566个，打分在4.25到5分之间的755个。从而得到了该种药物的成分、靶点打分信息。

下面是打分前50的靶点-成分数据，

|  |  |  |  |
| --- | --- | --- | --- |
| **靶点** | PDB ID | **成分** | docking scores |
| NOTCH1 | 5L0R | Solatubin | 8.59 |
| NOTCH1 | 5L0R | campesterol | 8.58 |
| NOTCH1 | 5L0R | ergosterol | 8.56 |
| NOTCH1 | 5L0R | sitosterol | 8.55 |
| NOTCH1 | 5L0R | Ziebeimine | 8.54 |
| NOTCH1 | 5L0R | Friedelin | 8.51 |
| NOTCH1 | 5L0R | Sitosterol alpha1 | 8.51 |
| PGR | 1SQN | dracaenogenins A | 8.50 |
| NOTCH1 | 5L0R | panaxatriol | 8.50 |
| CYP1A1 | 4I8V | Friedelin | 8.49 |
| PPARG | 3U9Q | Friedelin | 8.49 |
| NOTCH1 | 5L0R | dracaenogenins A | 8.48 |
| NOTCH1 | 5L0R | Verticine | 8.48 |
| PGR | 1SQN | Friedelin | 8.47 |
| ESR2 | 3OLL | ergosterol | 8.47 |
| ESR2 | 3OLL | Friedelin | 8.47 |
| NOTCH1 | 5L0R | Stigmasterol | 8.46 |
| PTGS2 | 5F19 | isoarborinol | 8.451 |
| ESR2 | 3OLL | panaxatriol | 8.451 |
| NOTCH1 | 5L0R | Peiminine | 8.448 |
| CYR61 | 4D0Z | isoarborinol | 8.446 |
| PGR | 1SQN | panaxatriol | 8.445 |
| NOTCH1 | 5L0R | Peimisine | 8.444 |
| ESR2 | 3OLL | dracaenogenins A | 8.441 |
| ROCK1 | 3V8S | Friedelin | 8.44 |
| ESR2 | 3OLL | Stigmasterol | 8.439 |
| NOTCH1 | 5L0R | isoarborinol | 8.438 |
| PPARG | 3U9Q | isoarborinol | 8.437 |
| CYP1A1 | 4I8V | isoarborinol | 8.437 |
| CYP1A1 | 4I8V | panaxatriol | 8.434 |
| ESR2 | 3OLL | sitosterol | 8.421 |
| PGR | 1SQN | ergosterol | 8.42 |
| CYP1A1 | 4I8V | Ziebeimine | 8.416 |
| CYP1A1 | 4I8V | Stigmasterol | 8.416 |
| CYR61 | 4D0Z | Friedelin | 8.412 |
| ESR2 | 3OLL | campesterol | 8.411 |
| CYR61 | 4D0Z | panaxatriol | 8.409 |
| CYR61 | 4D0Z | Sitosterol alpha1 | 8.409 |
| NOTCH1 | 5L0R | protopanoxadiol | 8.408 |
| CYP1A1 | 4I8V | protopanoxadiol | 8.406 |
| ESR2 | 3OLL | isoarborinol | 8.405 |
| PPARG | 3U9Q | panaxatriol | 8.405 |
| CYP1A1 | 4I8V | Solatubin | 8.405 |
| NOS3 | 4D1P | isoarborinol | 8.401 |
| PGR | 1SQN | isoarborinol | 8.394 |
| COMT | 3BWY | Friedelin | 8.39 |
| CYP1A1 | 4I8V | sitosterol | 8.39 |
| PTGS2 | 5F19 | Sitogluside | 8.389 |
| ESR2 | 3OLL | trans-feruloylcampesterol | 8.388 |
| NTRK1 | 5JFW | isoarborinol | 8.388 |

# 讨论

本文基于网络药理学的基本原理，开发了一套分析的系统。

# 展望

本文基于网络药理学的原理，针对现阶段中药分析中存在的问题，设计开发了一个在线的中药分析平台，可以完成从中药成分的获取，靶点成分的准备到在线分子对接整个流程的作业。最后展示了使用本系统的一个案例，解释了本系统的使用方式，并展示了相关的效果(Xiaofeng, Sisheng et al. 2010)。

参考文献

Xiaofeng, L., et al. (2010). "PharmMapper server: a web server for potential drug target identification using pharmacophore mapping approach." **38**(Web Server issue): W609.