



iMeta | 华中大宁康等 TCM-Suite 打通中药成分鉴定和网络药理分析，助力中药标准化、数字化和国际化

Received: 1 June 2022 | Revised: 5 July 2022 | Accepted: 20 July 2022
DOI: 10.1002/imt2.47

RESEARCH ARTICLE

iMeta WILEY

TCM-Suite: A comprehensive and holistic platform for Traditional Chinese Medicine component identification and network pharmacology analysis

Pengshuo Yang¹  | Jidong Lang^{2,3,4} | Hongjun Li¹ | Jinxiang Lu¹ |
Hanyang Lin⁵ | Geng Tian^{2,3} | Hong Bai¹ | Jialiang Yang^{2,3,4} | Kang Ning¹ 

TCM-Suite: 一个打通中药成分鉴定和网络药理分析的在线平台

短标题: TCM-Suite: 聚焦中药的成分鉴定和药理学挖掘的在线平台

原文链接: <https://onlinelibrary.wiley.com/doi/10.1002/imt2.47>

第一作者: 杨朋硕

通讯作者: 宁康

主要单位:

华中科技大学生命科学与技术学院 (Key Laboratory of Molecular Biophysics of the Ministry of Education, Hubei Key Laboratory of Bioinformatics and Molecular - Imaging, Department of Bioinformatics and Systems Biology, Center of AI Biology, College of Life Science and Technology, Huazhong University of Science and Technology, Wuhan, Hubei, China)

元码基因科技 (北京) (Geneis Beijing Co., Ltd., Beijing 100102, China)

赛昆思生物科技公司 (Sequenxe Biological Technology Co., Ltd., Xiamen 361000, China)

引文: Pengshuo Yang, Jidong Lang, Hongjun Li, Jinxiang Lu, Hanyang Lin, Geng Tian, Hong Bai, Jialiang Yang, Kang Ning. 2022. TCM-Suite: A comprehensive and holistic platform for Traditional Chinese Medicine component identification and network pharmacology analysis. *iMeta* 1: e47. <https://doi.org/10.1002/imt2.47>

摘要

对于传统中药，基于 DNA 的生物成分鉴定和下游的药理学网络分析受到了越来越多的重视。并且，随着生物信息学工具的进步以及相关数据的积累，该领域已经得到了极大的进步。然而，目前缺少一个平台来无缝联通生物成分鉴定和下游药理学网络分析，阻碍了对传统中药这一药物宝库的深入挖掘。在本研究中，我们开发了由两个子数据库 Holmes-Suite 和 Watson-Suite 组成的 TCM-Suite 平台，分别用于中药生物成分鉴定和网络药理学研究：在 Holmes-Suite 子数据库中，我们收集并处理了六类标记基因序列，总共得到了 1,251,548 条高质量的标记基因序列。在 Watson-Suite 子数据库中，我们收集并整合了来自十几个公共数据库的大量网络药理数据。最重要的是，我们开发了一个数据 workflow，通过机器学习将中药生物成分鉴定和下游网络药理学研究相结合，使用户能够识别中药配方的成分并分析其潜在的药理作用机制。此外，我们定制化地设计了搜索引擎和展示平台，以更好地可视化这些丰富的资源。TCM-Suite 作为一个全面的搜索和分析平台，可用于传统中药的药物发现和再利用。

TCM-Suite 网站：<http://TCM-Suite.AImicrobiome.cn>。

关键词：中药；网络药理学；成分鉴定；整体数据 workflow；



亮点：

- TCM-Suite 平台由两个子数据库 Holmes-Suite 和 Watson-Suite 组成。
- Holmes-Suite 数据库收集了六种标记基因，涵盖了 235,470 种生物成分。
- Watson-Suite 收录了海量数据，可用于挖掘“草药-有效化合物-靶标蛋白质-疾病”的关联性。
- 设计了一条连接 Holmes-Suite 和 Watson-Suite 的数据 workflow，并通过用户友好的平台进行展示。

引言

传统中药 (TCM) 是传统医学的一个分支，经过 3,500 多年的中医实践，传

统中药已经涵盖了各种形式的草药。随着现代医学对于 TCM 的深入解构，发现中药制剂(即方剂)在治疗疾病方面具有广阔的前景，成为药物开发的重要资源。而随着生物信息学工具的进步和相关数据的积累，基于 DNA 测序技术的中药成分(生物成分)鉴定已成为一个重要且紧迫的问题：对于中药生物成分的准确且高效地鉴定，重要的是构建高质量的基于 DNA 的标记序列数据库，以及设计高效且准确的序列搜索引擎。此外，这些中药生物成分鉴定结果可以与靶蛋白和药物作用的分子机制联系起来，以加深对于中药配方的理解和应用。

基于 DNA 的标记基因序列(即 DNA 条形码)能够有效且准确地鉴定中药生物成分，尤其是在中药这种混合了许多物种的体系中。DNA 条形码的定义是来自物种基因组保守区域的一种可用于划分物种谱系的短序列。经过研究表明，ITS2、matK、trnH-psbA、trnL、rpoC1、ycf1 等典型的标记基因可应用于中医药研究。这些 DNA 条形码使得从中药这一混合体系中准确鉴定生物成分变得可行和方便。目前，高通量测序方法已成为破译基于标记基因的中药分类学特征的最重要和最有效的方法之一。

通常，中药方剂由多种草药组成，其中含有数十、数百甚至数千种有效化合物。因此，中药方剂为药物开发和临床实践提供了丰富的资源，并且为下游药理学研究提供了有力的基础：在鉴定出中药方剂的生物成分后，基于网络药理学，对每种成分的活性成分进行网络药理学分析，来解释活性化合物(或小分子)、人类靶标蛋白质(或基因)和疾病的生物系统之间的复杂关系。此外，通过探究“有效化合物-靶标蛋白质-疾病”这一相互作用网络，可以系统地评估生物成分的潜在药效。

近来，各种中药方剂分析数据库相继出现，有的侧重于生物成分的鉴定，有的侧重于网络药理学的推论。针对中药生物成分鉴定，目前有 ITS2 核糖体 RNA 数据库、TCMBarcode 数据库等数据库。对于网络药理学研究，目前已有许多聚焦于不同种类相关关系的数据库：Super Natural II、TCM@Taiwan、ChEMBL 和 TCMID 数据库，聚焦于识别中药配方-中药成分和中药成分-有效化合物的关系；STITCH 数据库主要解析化合物-蛋白质之间的关联关系，STRING 数据库聚焦于蛋白质-蛋白质相互作用分析，OMIM 和 GAD 关注蛋白质-疾病关联分析。最后，为了全面了解中药的药效作用，toxicity 和 side effect 数据库收录了中药相关的安全性评估数据。

尽管这些中医药相关数据库取得了巨大的成功，但传统中药研究走向现代化和全球化的最大瓶颈仍未克服：中药配方的准确成分应该是什么，治疗疾病的有效化合物是什么，它们的靶标蛋白质是什么？能够治疗什么人类疾病？这些问题在很大程度上仍未被先前描述的任何与 TCM 相关的数据库所解决。目前中医药相关数据库的局限性如下：(一)目前，中药的生物和化学鉴定方法还远远不够完善和准确。(二)虽然已经构建了很多中药网络药理分析数据库，但不同数据库之间缺乏整合导致信息的冗余和错误，导致了网络药理学研究变得不完整且低效。(三)重要的是，对于中药的分析仍然缺乏一个整体的工作流，无法整合包括鉴

定中药配方的具体成分，挖掘其相关的活性化合物和蛋白质靶点，并分析它们与疾病的关联在内的全部功能。

于是，我们开发了 TCM-Suite 数据库，以促进与 TCM 相关的数据管理、数据驱动的假设生成和潜在的药物发现。首先，为了构建一个全面的数据库，我们收集了大量的涵盖不同类型的中药相关的数据。针对中药的生物成分鉴定，我们收集了六种标记基因（ITS2、matK、trnH-psbA、trnL、rpocl 和 ycf1）的基因序列和物种信息，构建了 Holmes-Suite 子数据库，覆盖 235,470 种生物成分的 1,251,548 条标记基因序列，是在我们所知范围内最完整的数据库。对于下游的网络药理学分析，我们检索了大量公共数据库记录，构建了 Watson-Suite 子数据库，包括 6,692 个中药药方、7,322 种草药、704,321 种化合物、19,319 种蛋白质和 15,437 种疾病，以及不同类型数据之间的对应关系。其次，我们开发了一个完整的数据 workflow，将这两个子数据库的数据进行了集成，链接了中药的生物成分识别到下游网络药理学研究的全过程。第三，基于海量条目和完整的数据 workflow，我们设计了一个用户友好的展示和分析平台，供用户探索和可视化这些丰富的资源。我们相信 TCM-Suite 可以作为基于 TCM 的药物发现和再利用的资源平台，并且 TCM-Suite 这一集成数据库和分析平台的问世将有助于全球药物研究和临床实践。

TCM-Suite 数据库网站为 <http://TCM-Suite.AImicrobiome.cn>。

网站建设和内容

Holmes-Suite 子数据库的数据收集和处理

对于鉴定中药成分的标记基因（ITS2、matK、trnH-psbA、trnL、rpocl 和 ycf1），原始序列是从 NCBI 核酸数据库中（<https://www.ncbi.nlm.nih.gov/nucleotide/>），使用基因名称作为关键字进行搜索（截止日期 2020 年 12 月，以 GenBank 格式存储）。基于这些数据，我们构建了包含了中药标记基因数据库和中药成分鉴定系统的 Holmes-Suite 子数据库，以满足中药生物成分鉴定的严格要求。对于搜集到的标记基因序列，我们使用了机器学习批量处理低质量的标记基因序列的方法，具体如下：首先，对于下载的标记基因序列，挑选 1,000 条序列代表进行人工检查，确定没有低质量的碱基和越界序列后，得到高质量的标记基因。其次，利用这些高质量的标记基因序列构建隐马尔可夫模型（HMM），以识别这些候选序列的潜在标记基因区域。所有下载的不符合模型的序列都将被过滤掉。第三，手动收集 HMM 模型未涵盖的人工校验过的序列。我们以 Holmes-Suite 数据库中的 ITS2 标记基因序列为例，阐述了详细的流程 workflow，并提供在了我们网站的帮助页面上（<http://TCM-Suite.AImicrobiome.cn/help>）。基于这些高质量的标记基因，我们构建了一个高精度、覆盖率高的基于序列的搜索引擎。

Watson-Suite 的数据收集和处理

Watson-Suite 数据库中的数据由六个数据字段组成：中药配方、草药、草药成分、有效化合物、靶标蛋白质和相关疾病。为保证不同字段之间关联关系的准确性，所有不同字段之间的相关性主要通过从现有数据库中收集和整理得到。根据现有数据库中的数据，相关性主要分为两种：直接相关性包括中药配方-草药、草药-草药成分、草药成分-有效化合物、有效化合物-蛋白质和蛋白质-相关疾病。而其他字段之间的相关性被定义为间接相关性。

对于这些从不同数据库收集的数据，需要进行数据整合和过滤，然后进行人工校正。首先，为了提高数据的兼容性，我们将它们全部转换为标准格式。其次，人工检查每条数据，删除重复记录。详细的流程见我们的网站的帮助页面（<http://TCM-Suite.AImicrobiome.cn/help>）。

Holmes-Suite 和 Watson-Suite 中的数据之间的互连

Holmes 和 Watson 数据库之间的相互联系是通过连接 Holmes-Suite 中的“物种”字段和 Watson-Suite 中的“草药”字段来建立的，因为这两个字段都存储了草药的拉丁名称（图 1，A）。模糊匹配和弹性搜索(<https://www.elastic.co>) 用于匹配这两个字段。对于匹配到的草药，我们整合了网站上两个子数据库的详细信息，用于 Holmes-Suite 中的中药生物成分（图 1，B）或 Watson-Suite 中目标草药的确定（图 1，C），同时提供到相应数据库的超链接。同样，使用 Holmes-Suite 搜索模块也提供了这种互连（图 1，D）。

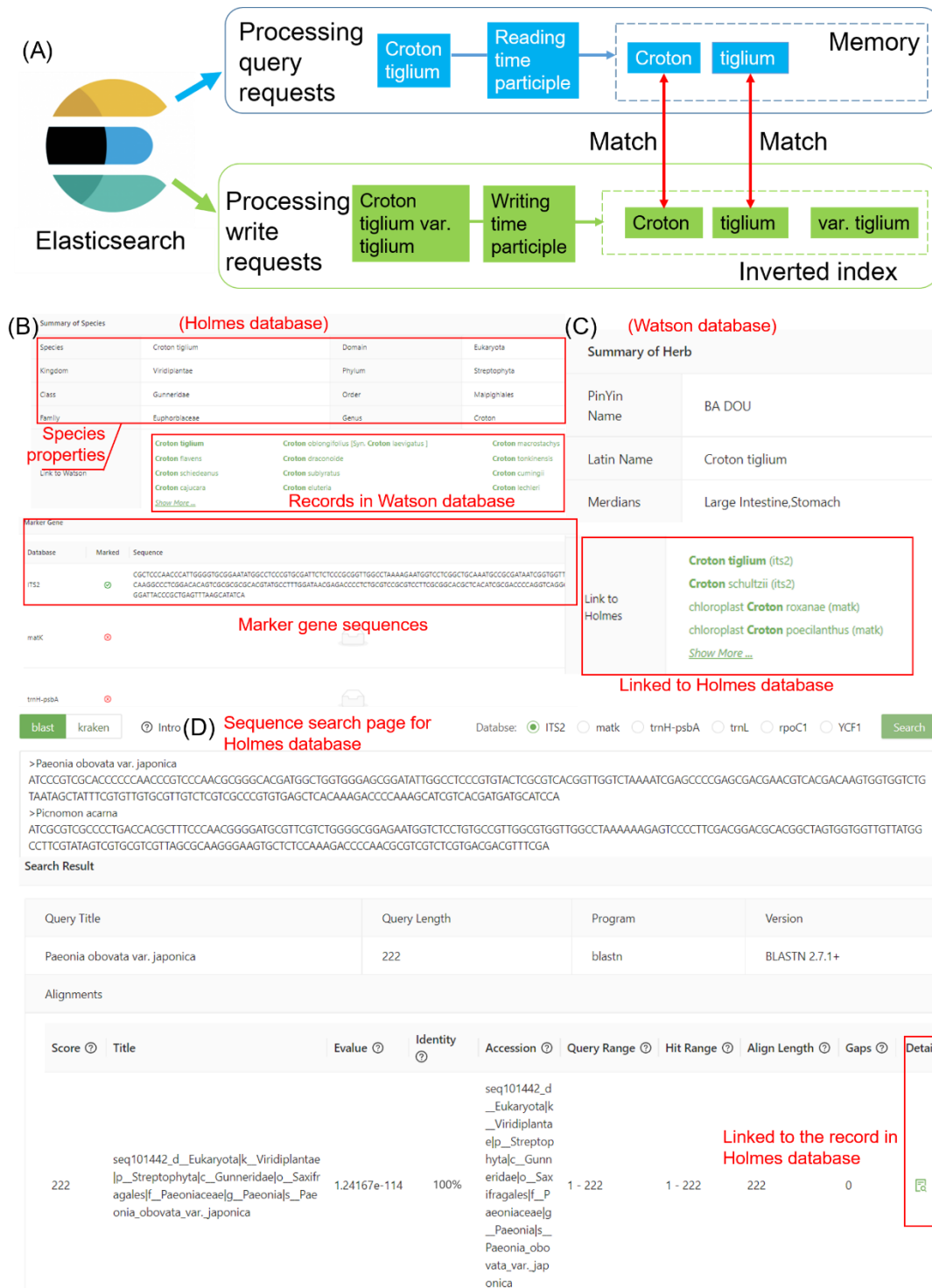


图 1. Holmes-Suite 和 Watson-Suite 两个子数据库的互连。(A) Holmes-Suite 和 Watson-Suite 的模糊匹配。Holmes-Suite 中的“物种”字段通过模糊搜索链接到 Watson-Suite 中的“草本”字段。(B)来自 Holmes-Suite 子数据库的数据信息。每个数据的详细信息都包含物种分类。“Link to Watson”是链接到 Watson-Suite 中相应的网络药理学分析结果。(C)Watson-Suite 数据库的详细草药信息页面。提供了该数据的物种名称和药物相关特性。同样，提供了 Holmes-Suite 中物种的

超链接。(D)对于序列搜索结果, Holmes-Suite 数据库还提供了指向 Holmes-Suite 中记录的超链接。这样用户可以方便地获取该检索结果的分类学信息和相应的网络药理学信息。

TCM-Suite 的网站搭建

TCM-Suite 网站使用了 Python-Flask、Nginx 和 React JavaScript 框架进行开发, 数据存储在后greSQL 数据库中。TCM-Suite 数据库兼容不同操作系统下的大多数主流浏览器(包括 Chrome、Edge、Firefox 和 Safari 测试), 支持 10 人同时搜索数据库和 30 人同时在线浏览。TCM-Suite 致力于为用户提供一个用户友好的 Web 界面进行浏览、搜索、结果可视化和数据下载。TCM-Suite 网址为 <http://TCM-Suite.AImicrobiome.cn>, 无需任何注册。

数据资源与数据库建设

为了建立一个全面的数据库, 我们收集并整合了多种来源的数据, 用于构建 TCM-Suite 数据库: 为了整合用于识别 TCM 生物成分的标记基因序列, 构建了 Holmes-Suite 子数据库(<http://TCM-Suite.AImicrobiome.cn/holmes/browse>)。对于鉴定成分后的下游网络药理学分析, 构建了 Watson-Suite 子数据库(<http://TCM-Suite.AImicrobiome.cn/watson/browse>)。Watson-Suite 包含多达 6,692 种中药配方、65,231 种配方-草药的关联数据、7,322 种草药、54,868 种草药-有效化合物的关联数据、704,321 种有效化合物、4,350,151 种靶标蛋白质-有效化合物的关联数据、19,319 种靶标蛋白质和 160,095 条蛋白质-疾病相关记录、15,437 种疾病和 160,169 条药物副作用记录。

Holmes-Suite 和 Watson-Suite 数据库中数据的数据统计信息如图 2 所示。在 Holmes 数据库中, 六种标记基因的对应物种分布如图 2, A 所示。尽管六个标记基因具有不同的物种分布, 但是仍有 1,131 个物种检测到了所有六种标记基因。基于基因序列的长度分布, 这六种标记基因都有较长的平均长度, 反映了序列的质量高(图 2, B)。同时, 与以前的网络药理学数据库相比, 我们的数据库提高了药理学网络的复杂性: 在 Watson-Suite 数据库中, 不同数据字段的药理学网络特征通过从现有数据库收集和整理的直接相关性(图 2, C)和推导出的间接相关性(图 2, D)来反映。

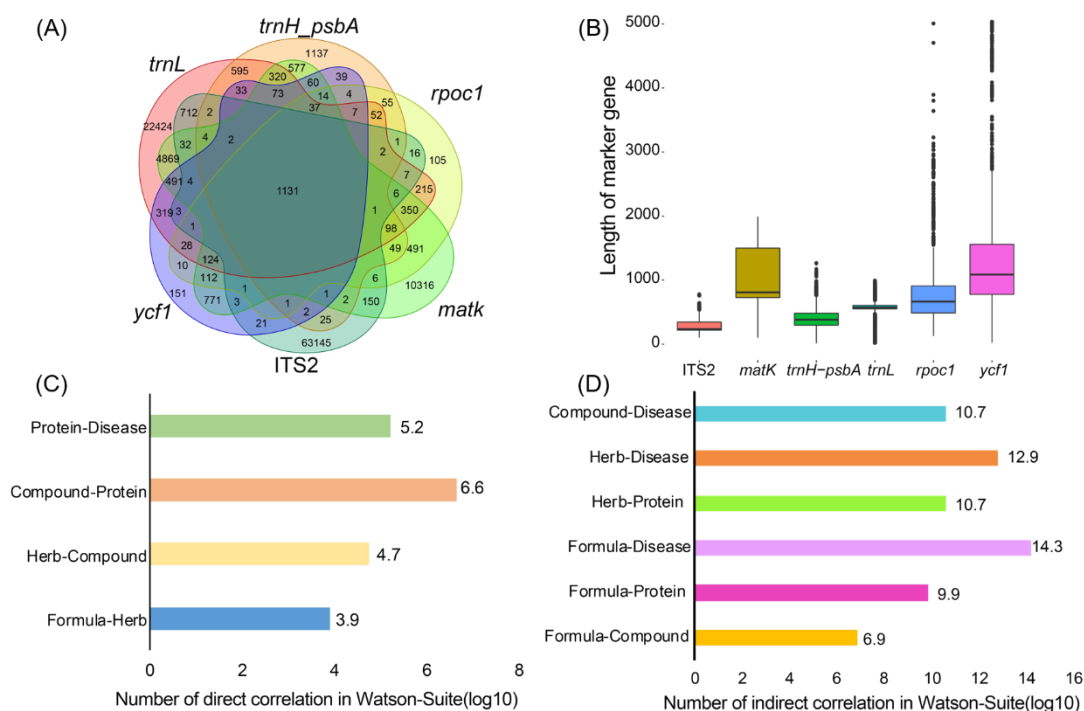


图 2. Holmes-Suite 和 Watson-Suite 子数据库的统计结果。(A)六种标记基因的物种分布。维恩图展示了六种标记序列数据库分组的物种数量。六种标记基因共有 1,131 个常见物种。(B)六种标记基因的序列长度分布。对于每种类型的标记基因，计算长度分布并以不同颜色标记。(C) Watson-Suite 子数据库中六种药理学数据的直接相关性。条形图显示了五种直接关联的直接关联总数。(D) Watson-Suite 数据库中六种药理学网络分类的间接相关性。通过网络药理学分析统计出的间接相关性。条形图显示了五个直接关联的间接关联总数。

实用性和讨论

中药分析的整体管道

为了更深入地挖掘天然产物蕴含的海量信息和探索中药相关的临床应用，我们实现了标记基因数据库(Holmes-Suite)与下游网络药理学分析(Watson-Suite)的互联互通。基于机器学习方法，TCM-Suite 中的 153,016 个物种条目(占有所有 235,470 个物种条目的 65%)可以与 Watson-Suite 中的“草本”字段成功的匹配。对于匹配的草药，我们整合了两个子数据库的详细信息。因此，TCM-Suite 允许用户搜索 TCM 的相关物种序列，TCM-Suite 将使用基于序列的搜索引擎返回 Holmes-Suite 数据库中匹配的条目。此外，由于所有数据相互关联，TCM-Suite 还会返回在 Watson-Suite 里的相关网络药理信息。反之亦然，用户可以在 Watson-Suite 中查询目标数据，并通过网络药理学分析返回 Watson-Suite 数据库中的所有相关条目，同时，TCM-Suite 还会返回 Holmes-Suite 中的相关标记基因信息。该数据 workflow 突出了我们数据库的可解释性，从而拓宽和深化了与中医相关的数据挖掘。

数据库功能和网络界面

TCM-Suite 数据库构建了一个易于使用的界面,供用户浏览和搜索数据,挖掘相关信息。TCM-Suite 的首页和数据浏览页面如图 3 所示:在 TCM-Suite 的首页(图 3, A),提供了 TCM-Suite 的主要功能的所有链接:右上角,链接到所有功能模块。此外,我们为用户提供了一个界面,可以直接搜索所有数据库的条目。搜索界面显示了基本的数据库统计,还描述了 TCM-Suite 的创建概念(图 3, B),提供了我们数据库背后的逻辑原理。在 Holmes-Suite 子数据库中,浏览页面(图 3, C)列出了所有标记基因条目。标有“详细信息”的链接提供了指向每个数据条目的分类信息的超链接以及相应网络药理学信息的 Watson-Suite 的链接。Watson-Suite 的浏览页面(图 3, D)显示了选择网络药理学类别后,数据库中的所有条目。对于 Watson-Suite 数据库中的每个条目,还列出了从其他分类中获得的相关数据(图 4, F)。为了全面直观地显示不同条目分类之间的复杂相关性,我们使用 Cytoscape Web (<http://cytoscapeweb.cytoscape.org/>) 设计了一个可视化界面来显示药理学网络中的关联关系。如果详细信息包含“草本”类别(即,在 Holmes-Suite 可匹配到标记序列),则会在“详细信息”页面中提供指向 Holmes-Suite 中相应记录的链接(图 3, E)。

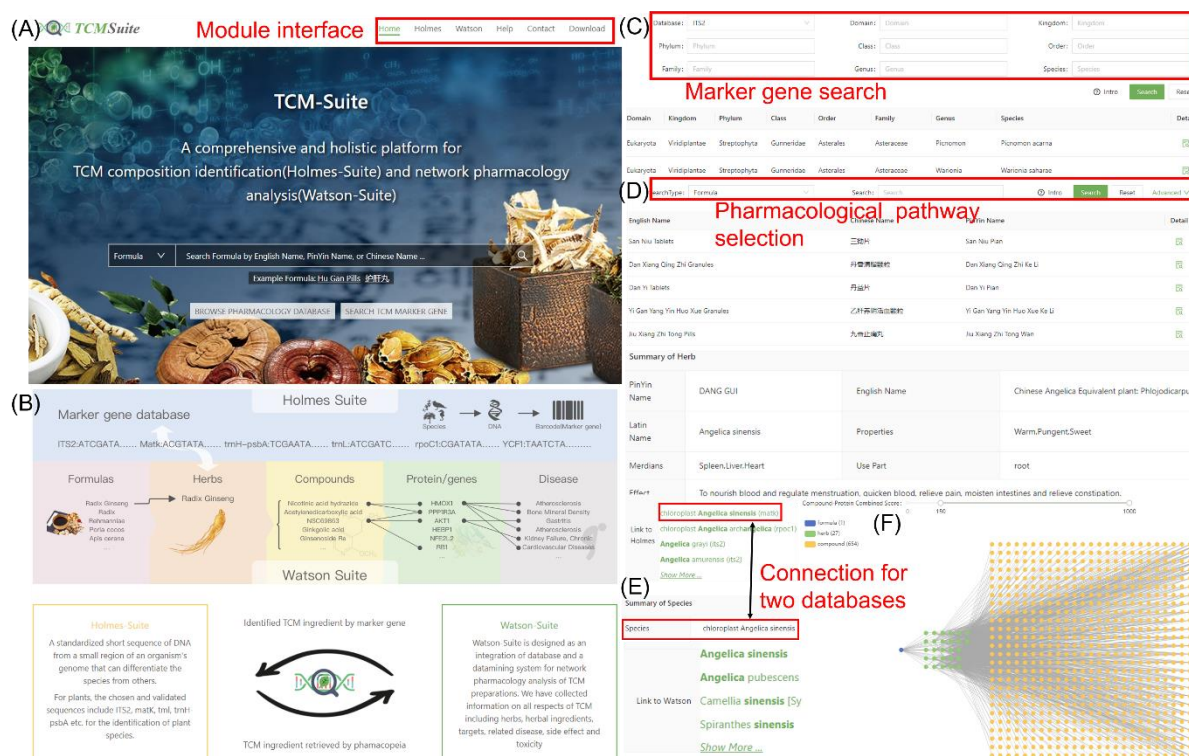


图 3. TCM-Suite 数据库界面截图。(A) TCM-Suite 主页。在右上角,提供了所有模块的超链接。(B) Holmes-Suite 的标记基因 ITS2 浏览页面。显示了详细的分类学分类,以及链接到维基百科的详细描述。(C) Holmes-Suite 的首页。Holmes-Suite 中标记基因条目的浏览页面由分类信息和用于挖掘下游网络药理学信息的 Watson-Suite 链接两部分组成。(D) Watson-Suite 草药的详细信息页面。这种草药的所有记录都列在页面顶部,然后是 Holmes-Suite 数据库中标记基

因序列的链接。显示了为 Watson-Suite 数据库中所选草药生成的药理学网络。当光标指向相应区域时显示信息。最后，在网络下提供了详细信息和相应的链接，并按条目类别组织。

基于 TCM-Suite 数据库的关于 COVID-19 的网络药理学调查

2019 年冠状病毒 (COVID-19) 的爆发引起了国际关注。大量的临床研究已证实了几种中药方剂对 COVID-19 的治疗有效，并被纳入 COVID-19 的诊治指南。作为整合中医药资源的强大平台，TCM-Suite 平台提供了大量的潜在资源来探究 COVID-19 相关药理学网络，以发现潜在的基因靶点和治疗药物。

基于前人研究，我们筛选出了五个公认有效的候选基因作为治疗 COVID-19 的靶点 (ACE2、SLC6A20、SIT1、CCR9 和 CXCR6)。对于这些候选基因，我们使用 TCM-Suite 数据库构建了“草药-靶标蛋白质-相关疾病”的药理学网络。对于五个查询基因，我们基于 TCM-Suite 数据库鉴定了有关的有效化合物 (分别为 12、23、1、2 和 14) 和相关草药 (分别为 58、128、2、8 和 54)。在这五个基因中，我们数据库中的大多数记录为与基因 SLC6A20 和 ACE2 相关，与之前的研究相符。对这些潜在靶标基因的进一步研究表明，152 种草药的大部分作用是“清热解毒”和“通肺解毒”。

从另一个角度，我们从 COVID-19 的临床症状为切入点构建了药理学网络。基于之前的研究中确定的 COVID-19 临床症状，肺炎和发烧是 COVID-19 感染的重要症状。于是我们检索了 TCM-Suite 中与这些症状相关的药理学网络，获得了 COVID-19 的目标候选基因簇和潜在的治疗药物。在统计了这些网络中的共同靶标基因和中药成分之后，共得到 15 个药物靶点、52 种有效化合物和 86 种中药成分。其中许多靶点在之前的研究中被证明为可用于治疗 COVID-19，例如 ACE2、SLC6A20、SIT1、CCR9 和 CXCR6。这 86 种中药中有许多在临床上被确定为治疗 COVID-19 的有效成分。例如，中药成分金银花、黄芩、连翘是双黄连的有效成分；连翘、麻黄、忍冬、板蓝根、薄荷、厚皮草是莲花清瘟胶囊的主要成分。综上所述，TCM-Suite 平台在挖掘治疗疾病的中医药方面显示出巨大的潜力。

结论

现有的中药生物成分鉴定和网络药理学数据库在搜索准确性、数据库完整性和搜索速度方面存在局限性。于是，我们开发了 TCM-Suite 平台，旨在整合现有的数据库资源，以促进将中药生物成分鉴定 (Holmes-Suite) 和下游网络药理分析 (Watson-Suite) 无缝集成到一个整体分析平台中。针对中药生物成分鉴定，我们开发了鉴别精度高、数据涵盖全面、鉴定效率高的 Holmes-Suite 子数据库，该系统人性化、交互性强，并通过 HMM 和模糊搜索等机器学习方法进行数据匹配和序列搜索。Holmes-Suite 子数据库收集了六种标记基因 (ITS2、matK、trnH-psbA、trnL、rpoC1 和 ycf1) 的数据，涵盖了 235,470 种生物成分和 1,251,548 条标记基因序列。对于下游中药相关的网络药理学分析，我们构建的 Watson-Suite 子数据

库可以更深入地挖掘“中药-有效化合物-靶标蛋白质-相关疾病”的相关性。Watson-Suite 子数据库包含 6,692 种中药配方、7,322 种草药、704,321 种有效化合物、19,319 种靶标蛋白质和 15,437 种疾病。通过如此全面的数据，Watson-Suite 可以通过整合现有数据来建立关于网络药理学的全面视图。

更重要的是，TCM-Suite 实现了两个子数据库 Holmes-Suite 和 Watson-Suite 之间的无缝互联，这一互联使得 TCM-Suite 在基于中药的药物发现方面具有巨大潜力。在 Holmes-Suite 的数据库测试中，从 NCBI 中随机抽取 1000 条 ITS2 的原始序列作为测试数据集查询 TCM-Suite 数据库，发现其中 552 条（55.2%）可以在 Watson-Suite 中检索到相关网络药理学数据。这些结果表明 Holmes-Suite 和 Watson-Suite 数据库之间存在很强的关联性。

此外，TCM-Suite 的生物成分鉴定和网络药理学搜索引擎使 TCM-Suite 能够高效运行。与其他数据库相比，Holmes-Suite 的查询速度更快，并且查询所需的时间与查询次数之间是线性相关的。这两个结果都反映了 Holmes-Suite 子数据库的搜索引擎具有很高的效率。同时，Watson-Suite 的查询速度也很快：查询 100 种药物时，数据库的反应时间仅为 0.6 秒。此外，凭借用户友好的服务器界面，TCM-Suite 可以应用于药理学发现。总而言之，TCM-Suite 是基于 TCM 的药物发现和再利用的丰富资源，可以为基础研究和临床研究做出巨大贡献。

中医药研究，尤其是分子生物学相关的研究越来越火爆，积累了大量的数据。为了反映生物成分识别和网络药理学的最新进展，TCM-Suite 计划每两年更新一次。此外，我们计划应用更先进的人工智能技术，用于更准确的中药成分鉴定模型的构建，以及更好的数据挖掘技术，例如用于网络药理学网络的挖掘。我们相信，在这个集成数据库和分析平台上的持续努力将为全球药物研究和临床实践做出巨大贡献。

第一作者简介



杨朋硕，华中科技大学创新研究院博士生。2017 年本科毕业于华中科技大学，之后在生命科学与技术学院宁康教授的指导下攻读博士学位。以第一作者身份发表 SCI 论文 6 篇 (Genome Biology, 2019; PNAS, 2021 等)，并参与五项国家自然科学基金委项目(其中一项为第二完成人)，七项发明专利。研究方向为生物大数据挖掘，以及微生物组学研究中的方法开发。重点关注环境和健康领域的生物大数据处理。



郎继东毕业于天津大学数学与应用数学系，深耕生物信息学领域十余年，且在科学研究及方法应用上拥有丰富的项目经验，主要研究方向包括生物信息学、基因组学、宏基因组学、单细胞组学等。在 PNAS、Nature Protocols、Cell Reports、NAR Genomics and Bioinformatics 等国际知名期刊发表 SCI 论文近 40 篇，其中第一作者或通讯作者论文 12 篇；获得国家发明专利授权 12 项，其中第一发明人 8 项；拥有 9 项软件著作权；部分研究成果在领域知名会议（如 AACR、ASCO、CSCO）上以海报和摘要的形式展出。担任 Frontiers in Bioengineering and Biotechnology 等多个 SCI 杂志 Guest Associate Editor 以及审稿人。

通讯作者简介



宁康，华中科技大学生命科学与技术学院教授，博士生导师，生物信息与系统生物学系系主任。2008 年博士毕业于新加坡国立大学计算机学院生物信息专业。在生物信息学和微生物组学领域从事科研工作 10 余年，作为通讯作者在 *PNAS*、*Gut*、*Genome Biology*、*Gut Microbes*、*Briefings in Bioinformatics*、*Bioinformatics* 等生物学、医学和生物信息学顶级学术期刊等高水平学术期刊发表学术论文 60 余篇，文章总引用超过 3000 次，H 指数 30 (Google Scholar)。目前主持国家自然科学基金项目、科技部重大研究计划课题等。担任中国生物信息学学会（筹）-基因组信息学分会副主任，中国生物工程学会-计算生物学与生物信息学专业委员会委员，中国计算机协会-生物信息学专业委员会委员等。担任 *Genomics*、*Proteomics*、*Bioinformatics*、*Microbiology Spectrum*、*Scientific Reports* 等国际期刊编委。详细情况请参见：<http://www.microbioinformatics.org/>。



杨家亮，博士，现任元码基因科技（北京）股份有限公司副总经理、首席信息官、中国海洋大学特聘教授。他的研究领域包括肿瘤基因检测、癌症早筛、人工智能和大数据挖掘在肿瘤和衰老研究中的应用等。迄今为止，发表学术论文共计 140 余篇，其中部分文章发表于顶级科学期刊《*Science*》、《*Cell Metabolism*》、《*Nature Communications*》、《*Molecular Biology and Evolution*》等，谷歌学术引用 7000 余次。主导研发原发灶不明转移癌组织溯源、医学图像分析等产品，申请 10 余项发明专利和多项软件专著。



白虹，博士，华中科技大学生命科学与技术学院实验教学中心副主任，高级工程师。多年来主要从事中药活性成分研究和质量评价，以及基于元基因组学的中药物种成分分析。为首承担国家自然科学基金项目 4 项，为主参与省部级课题多项，发表学术论文 60 余篇，其中以第 1 作者或通讯作者在 *Plant Biotechnology Journal* , *Frontiers in Pharmacology* , *Marine Biotechnology*, *Tetrahedron*, *Steroids* 等高水平期刊发表 SCI 论文 20 余篇。

引文： Pengshuo Yang, Jidong Lang, Hongjun Li, Jinxiang Lu, Hanyang Lin, Geng Tian, Hong Bai, Jialiang Yang, Kang Ning. 2022. TCM-Suite: A comprehensive and holistic platform for Traditional Chinese Medicine component identification and network pharmacology analysis. *iMeta* 1: e47. <https://doi.org/10.1002/imt2.47>