国外天然产物化学成分实物库及数据库建设概况

赵丽梅，谭宁华*

（中国科学院 昆明植物研究所 植物化学与西部植物资源持续利用国家重点实验室，云南 昆明 650201）

[摘要]　天然产物是新药发现的重要源泉，天然产物化学成分实物库和数据库的建设对天然产物的研究与开发具有重要意义。目前国外建设的小分子化合物库多为合成或组合合成分子，天然产物实物获取较困难。在信息数据库建设方面由于使用标准不同，信息不统一，开发规范、实用、智能型、综合型的大规模天然产物数据库还存在一定困难。该文就目前国外可以公开查询到的有关天然产物的实物库及数据库建设情况进行了概述和分析，以期对天然产物研究与开发，特别是天然产物化学成分实物库和数据库的建设提供参考。

[关键词]　天然产物；实物库；数据库

Status of libraries and databases for natural products at abroad

ZHAO Li-mei, TAN Ning-hua*

(*State Key Laboratory of Phytochemistry and Plant Resources in West China, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming* 650201, *China*)

[Abstract]　For natural products are one of the important sources for drug discovery, libraries and databases of natural products are significant for the development and research of natural products. At present, most of compound libraries at abroad are synthetic or combinatorial synthetic molecules, resulting to access natural products difficult; for information of natural products are scattered with different standards, it is difficult to construct convenient, comprehensive and large-scale databases for natural products. This paper reviewed the status of current accessing libraries and databases for natural products at abroad and provided some important information for the development of libraries and database for natural products.

[Key words]　natural products; compound library; database

doi: 10.4268/cjcmm20150106

由于生物代谢的多样性和生态环境的复杂性，来源于天然产物的化合物一般都具有丰富的结构多样性，并且大多具有较好的生物活性。在科学技术飞速发展、化学合成药物占主导地位的今天，生物次生代谢的天然产物仍是新药发现的重要源泉，据统计，从 1981 年到 2010 年，批准上市的小分子药物中有 60% 以上的药物直接或间接来源于天然产物[1]。目前研究天然资源中的化学成分，从中筛选生物活性强的、有成药潜力的化合物作为先导化合物，经结构修饰和改造，寻找疗效更高、结构更简单并且便于工业化生产的、安全可控的候选药物，再经临床试验证明其是否能作为新药上市，是创新药物研究的重要途径。随着现代分离、分析、生理、药理、药学、医学等实验技术的不断进步，每年都有大量的化合物被发现，但是到目前为止只有小部分化合物进行了有限的生物活性筛选，所以在活性筛选方面还有很大的潜力。天然产物相关信息数据库和实物库的建立对天然产物的研究及开发具有重要的意义，之前已经对国内中药化学成分库建设现状进行了介绍[2]，本文就目前国外可以公开查询到的有关天然产物的实物库及数据库建设情况进行概述和分析，以期对天然产物研究与开发，特别是天然产物化学成分实物库和数据库的建设提供参考。

1　实物库建设概况

国外很多制药公司和研究机构都建有自己的化合物库，如美国辉瑞、德国拜耳、瑞士诺华、英国葛兰素史克、美国国立癌症研究所等，都在以多种方式大力扩建自己的化学成分库，占领新药研发的源头——分子资源，但多不公开共享，其库存成分多为合成或组合合成分子，分子结构多样性较少，

[收稿日期]　2014-09-24

[基金项目]　国家“重大新药创制”科技重大专项（2011ZX09307-002-02）；云南省自然科学基金项目（2012GA003）

[通信作者]　* 谭宁华，研究员，主要从事植物环肽化学与化学生物学研究，Tel：（0871）65223800，E-mail：nhtan@mail.kib.ac.cn

[作者简介]　赵丽梅，硕士，实习研究员，主要从事中药化学成分库工作，Tel：（0871）65223266，E-mail：zhaolimei@mzil.kib.ac.cn

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

其天然分子多从国外如中国大量收购或合作收集。一方面，由于植物、微生物等天然产物的化学结构独特，一些人工很难合成的化合物在生物体内通过酶的作用就容易形成；另一方面，生物在不断进化的过程中其天然成分大多具有某些生物活性，从中寻找先导化合物比人工合成成功率更高。因此天然产物备受世界各国医药研发者的青睐。目前，美国、欧盟、日本、韩国等一些国家和地区的许多医药研究机构都在加紧进行有关天然植物药的研发工作。不少大型制药公司正尽力把大量的植物物种送入实验室进行大规模筛选，以便从中发现任何可能的生物学功效。如美国国立癌症研究所通过与世界各地的高校或研究所建立合作关系，收集大量的植物、海洋生物、真菌等样品，建立了其天然产物筛选库，据报道，到 2009 年末已收集并制备了 230 000 多个样品[3]。虽然国外目前专门从事天然产物实物库建设的单位不多，但由于在世界各地都有不少从事天然产物的研究和开发的研究单位和公司，且其大多为微生物和海洋天然产物，表 1 列举了一些国外建有天然产物实物库或可提供天然产物的研究单位或公司。

表 1　国外建有天然产物实物库或可提供天然产物的研究单位或公司[4-7]

Table 1　Organizations with natural products library at abroadp[4-7]

名称	天然产物种类及数量	网址
Albany Molecular Research, Inc.	约 300 000 来源于海洋及陆生微生物、植物的天然产物	http://www.amriglobal.com
Ambinter	天然产物和植物提取物	http://www.ambinter.com
Analyticon Discovery	3 900 多个植物天然产物，1 300 多个微生物天然产物	http://www.ac-discovery.com
CB Library	海洋放线菌代谢产物	
Cerylid Biosciences	600 000 多个提取物	http://www.cerylid.com.au
Entocosm Pty Ltd	1 300 多种昆虫和陆生无脊椎动物标本，包括部分提取物	http://www.csiro.au
Gaia Chemical Corporation	天然产物及衍生物	http://www.gaiachem.com
Instituto Biomar, S. A.	15 000 多个海洋微生物提取物	http://www.institutobiomar.com
InterBioScreen Ltd GmbH	41 000 多个天然产物及其衍生物	http://www.ibscreen.com
IRIS Biotech GmbH	1 200 多个天然产物	http://www.iris-biotech.de
Kosan Biosciences	聚酮类天然产物	http://www.kosan.com
Magellan Bioscience Group Inc.	从海洋植物、动物、微生物提取的 10 000 多个天然产物	http://www.magellanbioscience.com
Marine Natural Products	海洋天然产物	
Martek Biosciences	主要是来源于微藻的天然产物	http://www.martekbio.com
Mera Pharmaceuticals, Inc.	水生微生物(如海藻类)天然产物	http://www.merapharma.com
MicroSource Discovery Systems, Inc.	天然产物及其衍生物、植物提取物	http://www.msdiscovery.com
Moscow MedChemLabs, Ltd	200 多个天然产物	http://www.mosmedchemlabs.com
NCI	美国国立癌症研究所小分子化合物库，收集了大量的天然产物	
Sequoia Sciences	植物天然产物	http://www.sequoiasciences.com
Specs	400 多个天然产物及其衍生物	http://www.specs.net
TimTec	13 000 多个天然产物及其衍生物和植物提取物	http://www.timtec.net

2　数据库建设概况

在化学和生物学信息数据库建设方面，国外发展较快，出现了很多信息全面、易于查询、可智能分析的化学、生物、医药类数据库，给科研工作者带来了很大的便利，不仅可以节约成本，还提高了工作效率。随着各国医药科研工作者对天然药物的重视，近些年也建立了一些有关天然产物方面的信息数据库，下面主要从以下几个方面进行介绍。

2.1　综合信息数据库　进入 20 世纪由于合成药物、组合化学、高通量筛选、靶向药物设计等技术的发展，天然药物的研究一度进入低谷，使得有关天然产物的综合信息数据库的发展缓慢。近些年随着化学药的毒副作用不断出现、疾病谱的改变以及医学模式的转变，很多医药科研工作者又把焦点转移到天然产物上，并认识到包含药物生物来源、分布、传统应用、活性成分分子结构与功能、疗效和毒副反应、临床应用、参考文献等信息的综合型数据库的建立对于天然产物的研究与开发具有重要的意义。目前也有一些单位在进行这方面数据库的研究与构建，表 2 列举了国外一些有关天然产物综合信息的数据库，这些数据库的建立在一定程度上促进了天然产物的研究与开发，但是由于信息的收集困难，并且传统医药与现代医药之间存在差别，难以将两者较好的结合起来，在专业术语描述上也缺乏统一的标准，使这一类数据库的发展受到了一定限制，规模不大。

2.2　针对特定疾病的药用植物数据库　天然药物在一些疾病的治疗上显示有比合成药物更好的疗效，因此在新药研发时选择对特定疾病有效且有长期使用经验的传统药用植物进行研发，会提高创新药物成功的概率。目前国外已出现了一些用于治疗癌症、糖尿病、哮喘、关节炎、皮肤病等的药用植物数据库，见表 3。但是由于传统医学在疾病分类、病理描述等方面与现代医学存在差别，怎样将两者联系起来，建立统一标准，是这一类数据库发展面临的一大难题。

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

表 2　国外有关天然产物综合信息的数据库

Table 2　Comprehensive databases on natural products at abroad

数据库名称	国家	内容	网址
Customary MedicinalKnowledgebase [8]	澳大利亚	澳大利亚土著人传统药用植物信息,包括植物分类、植物化学、地理分布、生物活性等信息	http://biolinfo.org/cmkb/index.php
GlobinMed [9]	马来西亚	马来西亚、中国、南非等多个国家的药用植物信息,包括植物名称、传统应用、公开发表的医学研究数据等信息	http://www.globinmed.com
KNApSAcK Family Databases [10]	日本	主要包括 Metabolomics DB system 和 Multifaceted Plant Usage DB 两个部分: Metabolomics DB system 收录了 101 500 多个物种-代谢物关系,包含 20 741 个物种和 59 048 个代谢物; Multifaceted Plant Usage DB 包括药用/食用植物应用、地理分布、生物活性、日本及印度尼西亚复方等信息。	http://kanaya.naist.jp/KNApSAcK_Family/
Medicinal Plants Database of Bangladesh [9]	孟加拉	孟加拉国内外公开发表的有关药用植物的信息,包括植物分类、名称、形态描述、照片、化学成分、应用、地理分布、药理等信息	http://www.mpbd.info
Natural Products Alert [9]	美国	民族药及传统医药、植物药、微生物和海洋药物的文献数据库,数据来源于 700 多种期刊,包含民族医疗、药理、植物化学等信息	http://napralert.org/
TCM-ID [11]	新加坡	1 197 个处方,1 098 种药用植物和 9 852 种中草药成分,包括中草药处方、组成、化学成分、活性成分分子结构与功效、疗效和毒副反应、临床应用、参考文献等信息	http://tcm.cz3.nus.edu.sg/group/tcm-id/tcmid.asp
TradiMed [9]	韩国	3 199 种韩国传统药和中药、11 810 种处方、来源于草药、微生物、海洋资源的 20 012 种化学成分、4 080 种传统疾病、95 种中药炮制方法、502 种中草药的炮制方法,包括草药、处方、化学成分、疾病与症状、炮制等信息	http://www.tradimed.com

表 3　国外针对特定疾病的药用植物数据库

Table 3　Disease-specific databases on natural products at abroad

数据库名称	国家	内容	网址
Arthritis Database [12]	印度	100 多种具有抗关节炎活性的药用植物的名称、地理分布、药用部位、临床试验数据等信息	
Asian Anti-Cancer Materia Database [13]	美国	约 700 种来自亚洲的抗癌药物信息,其中 80% 为植物来源,其他来源于真菌、食物、昆虫、海洋、爬行动物、矿物等	http://asiancancerherb.info/index.html
Asthmaplants [14]	印度	40 属 50 种可治疗哮喘的药用植物的植物分类、药用部位、参考文献、药用用途等信息	http://www.asthmaplants.com
CHMIS-C [15]	美国	临床上用于治疗不同类型癌症的 500 多个草药处方、900 多个成分、203 个癌症相关分子靶标和从中草药中分离得到的 9 366 个小分化合物的相关信息,并建立了参考文献和分子靶标子数据库,可通过多种途径在线查询	http://sw16.im.med.umich.edu/chmis-c/
Database On Anti-diabetic Plants Based On Clinical/Experimental Trials [16]	印度	印度抗糖尿病植物的名称、药用部位、地理分布、活性成分、临床试验数据、植物原料供应商等信息	http://www.biotechpark.org.in/antidia/index.html
DIAB [17]	印度	258 属 287 种具有降血糖作用的植物的名称、参考文献、活性成分等信息,其中 129 种还有抗氧化作用,53 种有抗菌作用	http://www.autogeneralfilters.com/holycross/Home.html
DiaMedBase [18]	印度	309 属 389 种具有抗糖尿病活性的药用植物信息	
InDiaMed [19]	印度	印度用于治疗糖尿病的药用植物的化学、药理、生物、地理分布等信息	http://www.indiamed.info
InPACdb [20]	印度	印度抗癌植物化合物数据库,包括化合物基本信息、三维立体化学结构、癌症类型、分子靶点等信息,并能进行化学信息学分析	http://www.inpacdb.org
Phyto-mellitu [21]	印度	印度传统用于糖尿病治疗的药用植物及其化学成分信息	http://www.bicmlacw.org/bt/
Plant-antivenom [22]	巴西	68 种植物、43 种分泌毒素的动物、53 种可能的抗毒活性、5 900 多篇文献、900 多个毒素蛋白氨基酸序列,包括具有抗蛇毒素的植物、可分泌毒素的动物、毒素蛋白氨基酸系列等信息	http://gbi.fmrp.usp.br/plantantivenom

2.3　天然产物化学数据库　国外在化合物信息数据库建设方面发展较早,从 1961 年美国化学文摘社建立了第一个化学化工数据库《化学题录》,现称为化学文摘数据库(Chemical Abstract,CA)后,很多国家相继发展和建立了自己的数据库,但是主要是关于合成化合物的,而有关天然产物的信息数据库直到近些年随着科研工作者对天然产物的重视,才慢

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

慢兴起，目前已建立了不少这方面的数据库，表 4 是国外可以公开查到的一些天然产物化学相关数据库。但是目前建立的这些数据库大多只能提供一些比较基础的信息，并多把焦点放在新药发现上。

表 4　国外有关天然产物化学的数据库

Table 4　Phytochemical databases on natural products at abroad

数据库名称	国家	内容	网址
3DMET [23]	日本	天然代谢产物三维结构数据库，包括 KEGG Compound 中收录的大部分天然产物的三维结构和理化常数等信息	http://www.3dmet.dna.affrc.go.jp
Alkamid Database [24]	比利时	*N*-Alkylamides (NAAs) 类化合物的植物来源、名称、结构、分子式、相对分子质量、生物活性、生物合成途径、药用情况等信息	http://alkamid.ugent.be/
Antibase [25]	德国	40 000 多个来自微生物和高等真菌的天然产物的分子式、相对分子质量、CAS 号、来源、理化常数、波谱、药理、毒理、参考文献等信息	http://www.wiley-vch.de/stmdata/antibase.php
Chem-TCM [26]	英国	350 多种中药，9 500 多个化合物的相关植物、化学、对常见靶标的活性预测等信息	http://www.chemtcm.com/
Chemiebase Database [27]		来自 79 种植物的 29 960 多个化合物的相关信息	http://chemiebase.ku.ac.th
Chinese Herbal Constituents Database [28]	英国	从 240 种常用中草药中分离得到的 7 000 多个化合物的名称、化学分类及骨架类型、相关植物的系统名及中文名、CAS 号、药理、毒理、参考文献等信息	
Dictionary of Natural Products Databases [29]	英国	190 000 多个天然产物的名称、分子式、相对分子质量、结构式、CAS 号、立体化学、理化常数、来源、毒性、活性、用途、参考文献等信息	http://dnp.chemnetbase.com
Dr. Duke's Phytochemical and Ethnobotanical Databases [30]	美国	植物化学与民族植物学信息数据库，包括植物、化学成分、活性、民族应用情况等信息	http://www.ars-grin.gov/duke/
Indonesian Herbal Constituents Database [31]	英国	来自印度尼西亚的 33 种不同植物的 1 000 多个化合物的名称、结构、CAS 号、药理、毒理、植物来源、参考文献等信息	
KEGG Compound Database [32]	日本	17 000 多个代谢物及其他小分子化合物的名称、分子式、相对分子质量、结构式、CAS 号、化学反应、代谢途径等信息	http://www.genome.jp/kegg/compound/
Marinlit [33]	新西兰	海洋天然产物文献数据库，数据来源于 1 400 多种期刊和书，已收录了 24 000 多个化合物的名称、结构、分子式、相对分子质量、UV、^{13}C-NMR 和 ^{1}H-NMR 的化学位移等信息	http://pubs.rsc.org/marinlit
NCI [4]	美国	化合物及药物结构和性质检索数据库，可以通过多种途径检索得到化合物的分子式、CAS 号、相对分子质量、抗 HIV 活性、氢键接受体和给予体的数目、可旋转键的数目、logP、有关生物活性的预测值以及化合物结构的 SMILES、sdf 文件等信息	http://cactus.nci.nih.gov/ncidb2.2
RÖMPP Online [4]	德国	德国最大的化学百科全书信息检索系统，RÖMPP Encyclopedia Natural Products 是其中的一部分，收录了 6 000 多个天然产物的信息，包括化学结构、立体化学、绝对构型、名称、分子式、相对分子质量、理化常数，来源、生物合成途径、医疗用途等信息	http://www.roempp.com/prod/
Seaweed Metabolite Database [34]	印度	来自红藻藻属海藻的代谢产物信息，包括来源、提取方法、化学结构、生物活性等信息，可通过结构和文本 2 种方式检索，并且可以和 PubChem 和 ChemSpider 数据库连接	http://www.swmd.co.in/home.html
StreptomeDB [35]	德国	从 1 900 多种链霉菌菌株或突变菌株分离得到的 2 000 多个化合物的来源、结构、分子式、相对分子质量、名称、生物活性、生物合成途径、参考文献等信息	http://www.pharmaceutical-bioinformatics.de/streptomedb/
SuperNatural Ⅱ [36]	德国	352 811 多个天然产物的二维结构、理化常数、供货商等信息	http://bioinf-applied.charite.de/supernatural_new/
The Beilstein Database [4]	德国	数据可追溯到 1771 年，来源于 180 种期刊，其中收录了大量天然产物信息，包括植物来源、化学结构、化学名称、生物活性、药理、毒理、代谢、NMR、MS 和 IR 等信息	

2.4　天然产物波谱数据库　NMR，MS，IR 等波谱信息是目前天然产物化学结构鉴定的主要依据，但是由于这些波谱信息缺乏系统整理，研究人员在进行化学成分结构鉴定时需花费大量时间整理实测谱图和文献中的化合物结构鉴定数据，再进行人工结构解析，所以对这些波谱信息进行系统整理，开发智能波谱信息数据库，对于提高天然产物化学成分研究的效率及准确性具有重要的促进作用。目前国外已有不少单位建立了有机化合物的各种波谱信息数据库，但是有关天然产物化学成分的波谱信息的数据库还不多，表 5 列举了一些有关天然产物的波谱信息的数据库，目前建立的这些数据库大多只能提供波谱信息的查询，智能化程度不高，规模偏小。

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

表 5　国外有关天然产物波谱信息的数据库

Table 5　Spectral databases on natural products at abroad

数据库名称	国家	内容	网址
Biological Magnetic Resonance Data Bank [37]	美国	大量多肽、蛋白质、核酸等的 NMR 信息	http://www.bmrb.wisc.edu/
CH-NMR-NP [38]	日本	28 000 多个公开发表的天然产物的 NMR 信息,可以通过化合物名称、分子式、相对分子质量、化学位移等途径检索	https://www.las.jp/CH-NMR-NP/
Golm Metabolome Database [39]	德国	500 多个植物代谢物的 GC/MS 信息	http://gmd.mpimp-golm.mpg.de/
Madison Metabolomics Consortium Database [40]	美国	20 000 多个代谢产物名称、结构、理化常数等基本信息和 NMR 及 MS 信息	http://mmcd.nmrfam.wisc.edu/
MassBank [41]	日本	有机化合物 MS/MS,EI/MS 信息数据库,可以通过名称、分子式、相对分子质量、谱峰、亚结构等查询	http://www.massbank.jp/
METLIN [42]	美国	大量内源性和外源性代谢物的相对分子质量、分子式、结构式、MS、MS-MS、LC-MS 等信息	http://metlin.scripps.edu/
Natural Products ^{13}C-NMR Database [43]	西班牙	包含 6 000 多个天然产物的 NMR 信息	http://c13.usal.es
NMRShiftDB [44]	德国	有机化合物结构及 NMR 信息数据库,可以通多种途径检索,并可以对化合物碳谱进行预测。目前已有 21 500 多个化合物,其中收录了大量天然产物	http://nmrshiftdb.nmr.uni-koeln.de
Spectral Database for Organic Compounds [45]	日本	有机化合物的 MS、FT-IR、NMR、拉曼光谱、ESR 谱,可以通过名称、分子式、相对分子质量、原子数、CAS 号、红外最大吸收峰值、核磁化学位移等查询	http://sdbs.db.aist.go.jp/sdbs/cgi-bin/cre_index.cgi

2.5　天然产物活性数据库　小分子化合物活性数据库对生物及医药研发具有重要意义,目前已建立的活性数据库,包括 ChemBank,DrugBank,PubChem,ChEBI 等,然而针对天然产物活性的数据库不多,表 6 列举了一些有关天然产物活性的数据库。

表 6　国外有关天然产物活性的数据库

Table 6　Bioactive databases on natural products at abroad

数据库名称	国家	内容	网址
Bioactive Plant Compounds Database [23]	英国	已知靶标的 2 597 个化合物的化学结构、名称、CAS 号、化学分类、骨架类型、立体结构、药理、毒理、靶标、IC_{50} 等信息	http://www.onco-informatics.com/
CamMedNP [46]	喀麦隆	用于虚拟筛选的 2 500 多个天然产物及其衍生物的三维结构信息	
ChEBI [47]	英国	大量活性小分子化合物,其中有一部分为天然产物,所使用的术语和符号与 IUPAC 和 NC-IUBMB 2 个国际标准相一致,通过 ChEMBL 数据库链接可以获得生物活性信息,包括与靶标的结合、功能、ADMET 等信息	http://www.ebi.ac.uk/chebi
ChemBank [48]	美国	一个开放的小分子活性数据库,包含化合物信息、实验结果、实验的原始数据、计算分子描述符、用户信息等,其中有部分化合物是天然产物,其最大的特点是收录了大量高通量筛选和小分子微阵列实验结果	http://chembank.broad.harvard.edu
EuroFIR-BASIS [49]	英国	多种类型的活性化合物及 330 个主要食用植物的食用部位、活性成分、生物功能等信息	http://ebasis.eurofir.org
Human Metabolome Database [50]	加拿大	40 000 多个内源性代谢物的信息,包括名称、结构、理化常数、NMR、MS、体内浓度、疾病相关性、代谢途径、酶、基因序列、SNP 和突变、参考文献等,除了从文献、书籍、数据库等收集的信息,还收集了大量的实验数据	http://www.hmdb.ca/
MetaCYC [51]	美国	有关代谢途径及与其相关化合物、酶、基因的相关信息	http://metacyc.org/
NPACT [52]	印度	来源于植物的具有抗癌活性的天然产物信息数据库,收录了 1 500 多个化合物的结构、理化常数、癌症类型、细胞系、抑制值(IC(50)/ED(50)/EC(50)/GI(50))、分子靶标、供货商和药物类似物等信息	http://crdd.osdd.net/raghava/npact/
PubChem [53]	美国	小分子生物活性公共数据库,主要包含物质信息、化合物的结构信息和生物活性信息,检索可得到化合物基本的结构、理化常数、该化合物作为药物的剂型及商品信息、药理、毒理、生物活性、参考文献等信息	http://pubchem.ncbi.nlm.nih.gov
Therapeutic Target Database [54]	新加坡	专业靶点数据库,收录了大量药物靶点和天然产物来源的药物信息,包括药物构效关系模型、靶点确证信息、多靶点药物数据,药物组合、天然来源的化合物等,可通过自定义、靶点相似性、药物相似性等多种方式查询	http://bidd.nus.edu.sg/group/ttd/ttd.asp
ZINC [55]	美国	35 000 000 多个可进行虚拟筛选的化合物信息及供应商信息,库中的分子结构均可免费下载,支持 SMILES,mol2,3D SDF 等多种文件格式,化合物的相关信息可以以表格形式下载,其中一部分是天然产物	http://zinc.docking.org/

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

3 问题与展望

国外在小分子化合物库建设方面发展较早，建立了一些规模较大、分子结构多样的小分子化合物库，但是这些化合物库的分子多为合成或组合合成化合物，随着医药科研工作者对天然产物的关注，国外不少企业和科研单位尝试建设自己的天然产物库，但是由于国外对天然产物的重视不够，其实物获取比较困难，所以目前国外天然产物化学成分库实物库的建设还不多。而在信息数据库建设方面，近年来发展较快，目前已建立了各种各样的天然产物相关数据库，但是也存在一些问题，主要表现为以下几个方面：多为专题型数据库，综合型数据较少；多为文献型数据库，不便于进行数据挖掘和统计分析；全文数据库较少，获取全文困难；缺乏统一标准，数据库之间难以实现整合与共享。所以通过数据优化、行业共同标准制定，建立规范、实用、智能型、综合型的大规模天然产物数据库，研究开发天然产物数据库的整合与扩展技术，实现不同数据库间资源的共享，并建立相应的高储量、大规模的天然产物实物库，对天然产物的研究开发具有重要的意义。

参考文献

[1] Newman D J, Cragg G M. Natural products as sources of new drugs over the 30 years from 1981 to 2010 [J]. J Nat Prod, 2012, 75 (3): 311.

[2] 赵丽梅，谭宁华. 国内中药化学成分库建设的现状分析与思考 [J]. 中国中药杂志, 2012, 37 (20): 3012.

[3] McCloud T G. High throughput extraction of plant, marine and fungal specimens for preservation of biologically active molecules [J]. Molecules, 2010, 15 (7): 4526.

[4] Fullbeck M, Michalsky E, Dunkel M, et al. Natural products: Sources and databases [J]. Nat Prod Rep, 2006, 23 (3): 347.

[5] Takagi M, Shin-ya K. Construction of a natural product library containing secondary metabolites produced by *Actinomycetes* [J]. J Antibiot, 2012, 65 (9): 443.

[6] Bugni T S, Richards B, Bhoite L, et al. Marine natural product libraries for high-throughput screening and rapid drug discovery [J]. J Nat Prod, 2008, 71 (6): 1095.

[7] Kato N, Takahashi S, Nogawa T, et al. Construction of a microbial natural product library for chemical biology studies [J]. Curr Opin Chem Biol, 2012, 16: 101.

[8] Gaikwad J, Khanna V, Vemulpad S, et al. CMKb: A web-based prototype for integrating Australian aboriginal customary medicinal plant knowledge [J]. BMC Bioinformatics, 2008, 9 (Suppl 12): S25.

[9] Ningthoujam S S, Talukdar A D, Potsangbam K S, et al. Challenges in developing medicinal plant databases for sharing ethnopharmacological knowledge [J]. J Ethnopharmacol, 2012, 141(1): 9.

[10] Afendi F M, Okada T, Yamazaki M, et al. KNApSAcK family databases: integrated metabolite-plant species databases for multifaceted plant research [J]. Plant Cell Physiol, 2012, 53(2): 1.

[11] Chen X, Zhou H, Liu Y B, et al. Database of traditional Chinese medicine and its application to studies of mechanism and to prescription validation [J]. Br J Pharmacol, 2006, 149 (8): 1092.

[12] Asif S M, Asad A, Poonam V, et al. Arthritis database: a composite web interface for anti-arthritic plants [J]. J Med Plants Res, 2011, 5 (12): 2457.

[13] Asian Anti-Cancer Materia Database: about this project [DB/OL]. 2014-08-10. http://asiancancerherb.info/index.html.

[14] Kasirajan B, Maruthamuthu R, Gopalakrishnan V, et al. A database for medicinal plants used in treatment of asthma [J]. Bioinformation, 2007, 2 (3): 105.

[15] Fang X, Shao L, Zhang H, et al. CHMIS-C: A comprehensive herbal medicine information system for cancer [J]. J Med Chem, 2005, 48 (5): 1481.

[16] Singh S, Gupta S K, Sabir G, et al. A database for anti-diabetic plants with clinical/experimental trials [J]. Bioinformation, 2009, 4 (6): 263.

[17] Arulrayan N, Rangasamy S, James E, et al. A database for medicinal plants used in treatment of diabetes and its secondary complications [J]. Bioinformation, 2007, 2 (1): 22.

[18] Babu P A, Suneetha G, Boddepalli R, et al. A database of 389 medicinal plants for diabetes [J]. Bioinformation, 2006, 1 (4): 130.

[19] Tota K, Rayabarapu N, Moosa S, et al. InDiaMed: a comprehensive database of Indian medicinal plants for diabetes [J]. Bioinformation, 2013, 9 (7): 378.

[20] Vetrivel U, Subramanian N, Pilla K. InPACdb—Indian plant anticancer compounds database [J]. Bioinformation, 2009, 4(2): 71.

[21] Middha S K, Mittal Y, Usha T, et al. Phyto-mellitus: a phytochemical database for diabetes [J]. Bioinformation, 2009, 4 (2): 78.

[22] Amui S F, Puga R D, Soares A M, et al. Plant-antivenom: database of anti-venom medicinal plants [J]. Electron J Biotechn, 2011, 14 (1): 1.

[23] Maeda M H, Kondo K. Three-dimensional structure database of natural metabolites (3DMET): A novel database of curated 3D structures [J]. J Chem Inf Model, 2013, 53 (3): 527.

[24] Boonen J, Bronselaer A, Nielandt J, et al. Alkamid database: Chemistry, occurrence and functionality of plant N-alkylamides [J]. J Ethnopharmacol, 2012, 142 (3): 563.

[25] Antibase 2012: the Natural Compound Identifier [DB/OL]. 2014-08-10. http://www.wiley-vch.de/stmdata/antibase.php.

[26] CHEM-TCM: chemical database of traditional Chinese medicine [DB/OL]. 2014-08-10. http://www.chemtcm.com/.

[27] Sawatdichaikul O, Hannongbua S, Sangma C, et al. In silico screening of epidermal growth factor receptor (EGFR) in the tyrosine kinase domain through a medicinal plant compound database [J]. J Mol Model, 2012, 18 (3): 1241.

[28] Ehrman T M, Barlow D J, Hylands P J. Phytochemical databases of Chinese herbal constituents and bioactive plant compounds with known target specificities [J]. J Chem Inf Model, 2007, 47 (2): 254.

[29] Dictionary of Natural Products: introduction to the Dictionary of

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

natural products online [DB/OL]. 2014-08-10. http://dnp.chemnetbase.com/intro/index.jsp.

[30] Dr. Duke's Phytochemical and Ethnobotanical Databases [DB/OL]. 2014-08-10. http://www.ars-grin.gov/duke/.

[31] Naeem S, Hylands P, Barlow D. Construction of an Indonesian herbal constituents database and its use in random forest modelling in a search for inhibitors of aldose reductase [J]. Bioorg Med Chem, 2012, 20 (3): 1251.

[32] KEGG Compound: metabolome informatics resource integrating genomics and chemistry [DB/OL]. 2014-08-10. http://www.genome.jp/kegg/compound/.

[33] MarinLit: a database of the marine natural products literature [DB/OL]. 2014-08-10. http://pubs.rsc.org/marinlit/introduction.

[34] Davis G D J, Vasanthi A H R. Seaweed metabolite database (SWMD): A database of natural compounds from marine algae [J]. Bioinformation, 2011, 5 (8): 361.

[35] Lucas X, Senger C, Erxleben A, et al. StreptomeDB: A resource for natural compounds isolated from *Streptomyces* species [J]. Nucleic Acids Res, 2013, 41 (Database issue): D1130.

[36] Dunkel M, Fullbeck M, Neumann S, et al. SuperNatural: A searchable database of available natural compounds [J]. Nucleic Acids Res, 2006, 34 (Database issue): D678.

[37] Ulrich E, L. Akutsu H, Doreleijers J F, et al. BioMagResBank [J]. Nucleic Acids Res, 2008, 36 (Database issue): D402.

[38] ^{13}C and ^{1}H NMR databse of organic natural products [DB/OL]. 2014-08-10. https://www.las.jp/CH-NMR-NP/English/English_help.html.

[39] Hummel J, Selbig J, Walther D, et al. Metabolomics: a powerful tool in systems biology. 18 [M]. Berlin: Springer Berlin Heidelberg, 2007: 75.

[40] Cui Q, Lewis I A, Hegeman A D, et al. Metabolite identification via the Madison metabolomics consortium database [J]. Nat Biotechnol, 2008, 26 (2): 162.

[41] Horai H, Arita M, Kanaya S, et al. MassBank: a public repository for sharing mass spectral data for life sciences [J]. J Mass Spectrom, 2010, 45 (7): 703.

[42] Smith C A, O'Maille G, Want E J, et al. METLIN: a metabolite mass spectral database [J]. Ther Drug Monit, 2005, 27 (6): 747.

[43] Lopez-Perez J L, Theron R, del Olmo E, et al. NAPROC-13: a database for the dereplication of natural product mixtures in bioassay-guided protocols [J]. Bioinformatics, 2007, 23 (23): 3256.

[44] Steinbeck C, Krause S, Kuhn S. NMRShiftDB-constructing a free chemical information system with open-source components [J]. J Chem Inf Comput Sci, 2003, 43 (6): 1733.

[45] Spectral Database for Organic Compounds SDBS [DB/OL]. 2014-08-10. http://sdbs.db.aist.go.jp/sdbs/cgi-bin/cre_index.cgi.

[46] Ntie-Kang F, Mbah J A, Mbaze L M, et al. CamMedNP: building the Cameroonian 3D structural natural products database for virtual screening [J]. BMC Complement Altern Med, 2013, 13 (1): 88.

[47] Hastings J, de Matos P, Dekker A, et al. The ChEBI reference database and ontology for biologically relevant chemistry: enhancements for 2013 [J]. Nucleic Acids Res, 2013, 4 (Database issue): D456.

[48] Kathleen P S, Gregory A G, Mary P H, et al. ChemBank: a small-molecule screening and cheminformatics resource database [J]. Nucleic Acids Res, 2008, 36 (Issue suppl 1): D351.

[49] Gry J, Black L, Eriksen F D, et al. EuroFIR-BASIS-A combined composition and biological activity database for bioactive compounds in plant-based foods [J]. Trends Food Sci Tech, 2007, 18 (8): 434.

[50] Wishart D S, Tzur D, Knox C, et al. HMDB: The human metabolome database [J]. Nucleic Acids Res, 2007, 35 (Database issue): D521.

[51] Caspi R, Altman T, Dreher K, et al. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases [J]. Nucleic Acids Res, 2008, 36 (Database issue): D623.

[52] Mangal M, Sagar P, Singh H, et al. NPACT: Naturally occurring plant-based anti-cancer compound-activity-target database [J]. Nucleic Acids Res, 2013, 41 (Database issue): D1124.

[53] Li Q L, Chen T J, Wang Y L, et al. PubChem as a public resource for drug discovery [J]. Drug Discovery Today, 2010, 15 (23-24): 1052.

[54] Zhu F, Shi Z, Qin C, et al. Therapeutic target database update 2012: A resource for facilitating target-oriented drug discovery [J]. Nucleic Acids Res, 2012, 40 (Database issue): D1128.

[55] Irwin J J, Shoichet B K. ZINC-A free database of commercially available compounds for virtual screening [J]. J Chem Inf Model, 2005, 45 (1): 177.

[责任编辑 丁广治]

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net