基于文献计量学的药用植物基因组学国际研究态势分析

杨俏俏, 黄林芳*

中国医学科学院&北京协和医学院药用植物研究所, 国家中医药管理局中药资源保护重点研究室, 北京 100193

* 联系人, E-mail: lfhuang@implad.ac.cn

收稿日期: 2017-10-16; 接受日期: 2017-12-12; 网络版发表日期: 2018-03-27

国家自然科学基金(批准号: 81274013, 81473315)资助

摘要 自2000年完成人类基因组计划后, 基因组学迎来了高速发展. 基因组学的药用植物研究成为热点. 检索2000~2016年SCIE文献数据库本草基因组学的国际研究论文共计文献3208篇. 利用CiteSpace分析平台和HistCite分析工具进行文献计量分析. 分析结果表明, 中国、印度、美国在本草基因组学研究中占主要地位, 占比分别为29.3%, 19.2%, 10.8%. 中国医学科学院、中国科学院、韩国庆熙大学为主要研究机构, 占比分别为6.1%, 5.1%, 1.5%. 中国学者陈士林在该领域发展发挥重要作用, 其提出的新学科——本草基因组学在该领域的引领优势将持续保持. 未来预测分析表明, 中国将继续为本草基因组学研究主力; 喀麦隆、奥地利会更多参与该项研究; 本草基因组学将更紧密地整合多组学技术, 如代谢组学、转录组学等, 此外, 基因编辑技术将为药用植物基因组学研究中注入新生力量.

关键词 基因组学, 药用植物, 全球变化, 文献计量分析

基因组学(genomics)是20世纪末发展起来的一门学科, 1986年由Tom Roderick提出, 主要对生物体内所有基因进行绘图、核苷酸序列分析以及基因定位和基因功能分析^[1]. 基因组学在药用植物研究中的运用涉及结构基因组、转录组、功能基因组、蛋白质组、代谢组、表观基因组、宏基因组、药用模式生物、基因组辅助分子育种、中药合成生物学、DNA鉴定、生物信息学等多个学科^[2-5]. 狭义的药用植物组学包含药用植物的结构基因组学和功能基因组学等组学概念, 广义的药用植物基因组学可涵盖很多内容, 包括DNA条形码、扩增片段长度多态性(amplified fragment

length polymorphism, AFLP)、分子标记等. 该研究涵盖狭义和广义的药用植物基因组学. 药用植物基因组学研究建立在植物基因组研究的基础上, 把基因组学在药用植物中应用有以下4个方面: (i) 道地药材的生物学本质. 如“沙漠人参”肉苁蓉在新疆与内蒙古不同道地产区 $psbA-trnH$ 存在差异^[6,7]. (ii) 中药分子标记. 如基于结构基因组学开发的DNA条形码鉴定技术已纳入《中国药典》^[8]. (iii) 药用植物基因资源的保护与利用. 目前已建立多个草药基因组数据库, 如转录组数据库(<http://medicinalplantgenomics.msu.edu>)、草药DNA条形码数据库(<http://tcmbarcodes.cn/en>). (iv)

引用格式: 杨俏俏, 黄林芳. 基于文献计量学的药用植物基因组学国际研究态势分析. 中国科学: 生命科学, 2018, 48: 498–508
Yang Q Q, Huang L F. Trend analysis of international studies in herbgenomics (in Chinese). Sci Sin Vitae, 2018, 48: 498–508, doi: 10.1360/N052017-00200

合成生物学研究. 该研究能解决传统的天然提取或人工化学合成难以满足新药研发需求问题, 目前已对丹参、人参等药材开展了有效物质合成调控酶基因的研究^[9~14]. 药用植物基因组学不能脱离动植物基因组学大的学科背景独立存在, 本草基因组学利用动植物基因组学的技术手段解决问题, 如道地性研究、有效成分合成研究等.

随着药用植物基因组学在我国迅速发展, 2010年, 中国学者陈士林等人^[15,16]提出本草基因组计划, 2016年提出新学科“本草基因组学”. 本草基因组学是从基因组水平研究中药及其对人体作用的一门前沿学科, 利用组学技术研究药用植物基原物种的遗传信息及其调控网络, 阐明药用植物防治人类疾病的分子机制^[17,18]. 该新兴的前沿生命科学技术在解决药用植物道地性形成的生物学实质、保存药用植物基因资源等方面有着不可替代的优势^[19,20], 韩国、美国、日本等国也已启动众多药用植物基原物种基因组研究^[21,22]. 从2000年人类全基因组计划完成后, 国际上发表了大量基因组方面的科技文献. 然而, 目前尚未应用文献计量学来分析药用植物基因组学研究的报告. 文献计量学是借助文献的各种特征的数量, 采用数学与统计学的方法来描述、评价和预测科学技术的现状与发展趋势的图书情报学分支学科^[23]. 美国科学信息研究所的科学引文索引扩展版(SCIE)数据库收录了世界各学科领域内最优秀的科技期刊, 其收录引文能及时反映科学前沿的发展动态^[24]. 检索2000~2016年SCIE中关于药用植物基因组学的国际论文, 利用Thomson科技公司开发的HistCite分析工具和CiteSpace分析平台对文献进行数据挖掘, 对药用植物基因组学研究的国际热点与发展态势进行分析及预测.

1 材料与方法

1.1 数据来源

在web of science中选择核心合集进行检索, 限定数据来源为SCIE, 以(“medicinal plant*” OR “medicinal herb*” OR “herbal medicine” OR “traditional medicine”) AND (genome OR genomics OR “structural genomics” OR “functional genomics” OR “function genomics” OR “functional genome” OR “gene expression and function” OR transcriptome OR transcriptomics OR proteome OR

proteomics OR metabolome OR metabonomics OR epigenome OR epigenetics OR epigenomics OR metagenome OR metagenomics OR “synthetic biology”) 和 (gene* OR DNA) AND “medicinal plant*”为检索式, 检索2000~2016年article/proceedings papers/review类型的文章, 得到关于药用植物基因组学研究的论文共4785篇(数据库更新时间: 2017-07-17), 去除重复后得3208篇. 在此基础上补充下载以“genome*” AND “plant*”为主题词的数据, 检索格式同上, 得到检索结果40661条, 按被引频次降序排序下载前5000条数据作为补充分析数据D_{SUP1}, 以“genome*” AND “animal*”为主题词的数据, 检索格式同上, 得到检索结果20684条, 按被引频次降序排序下载前5000条数据作为补充分析数据D_{SUP2}.

1.2 分析工具

采用CiteSpace进行分析, CiteSpace是Citation Space 的简称, 可译为“引文空间”. CiteSpace着眼于分析科学分析中蕴含的潜在信息, 并在科学计量学、数据和信息可视化背景下发展起来的引文可视化分析软件. CiteSpace的理论基础主要为5个方面: (i) 库恩的科学发展模式理论; (ii) Derek John de Solla Price的科学前沿理论; (iii) 社会网络分析的结构洞理论(structure hole); (iv) 科学传播的信息觅食理论(information foraging); (v) 知识单元离散余重组理论^[25]. HistCite=history of cite, 即引文历史, 是一个基于引文的分析软件. 该软件系SCI的发明人Eugene Garfield开发, 可快速绘制分析领域的发展历史, 定位该领域的重要文献及最新的重要文献^[26].

2 结果

2.1 总览分析

SCIE索引的药用植物基因组学研究国际论文共3208篇, 其中article占91.1%, review占8.1%, proceedings paper占0.8%. 99.4%是英文文章, 其他0.6%为日语、中文、德语等文章. 图1可见, 从2000~2016年, 药用植物基因组学研究国际论文发文量增加趋势显著. 文献量在2000~2007上升平缓, 在2008年后上升速度加快, 2016年出现发表文献高峰, 发表文献496篇. 2008~2012年发表的论文被大量引用, 5年间总被引频

图1 SCIE索引论文数和总被引频次的年度分布

次占总被引频次的44.2%。2010年发表的200篇文章,至今共被引用5283次。

发表该领域论文较多的期刊有: *Journal of Ethnopharmacology* (134篇), *PLoS ONE* (129篇), *Planta Medica* (78篇), *Evidence-Based Complementary and Alternative Medicine* (51篇), *Genetics and Molecular Research* (43篇), *Plant Cell Tissue and Organ Culture* (41篇), *Bmc Complementary and Alternative Medicine* (41篇), *Journal of Medicinal Plants Research* (36篇), *African Journal of Biotechnology* (32篇), *Phytotherapy Research* (32篇)等。

2.2 国家分析

2000~2016年, SCIE收录药用植物基因组学研究论文较多的前10个国家包括中国29.3%、印度19.2%、美国10.8%、韩国8.1%、巴西5.0%, 日本4.5%、德国3.9%、伊朗2.5%、马来西亚2.5%、英国2.5% (图2A)。国际合作排名前10的国家为中国、印度、美国、韩国、巴西、日本、德国、马来西亚、伊朗、意大利。中国在文章量与国际合作均排名第一, 主要与新加坡、英国、美国、法国等合作。印度主要与美国、法国、沙特阿拉伯等国合作; 美国主要与中国、印度、埃及、德国、纳米比亚等国合作。对从事

该领域研究的全部108个国家进行突发性检测(burst detection)。突发性检测说明被检测对象的变化情况, 通过突发性检测分析对象的变化可预测未来研究趋势。结果共有10个国家出现突发性, 强度最强的为日本, 强度指数13.6019, 其次为美国, 强度指数8.0232。突发持续时间最长的是日本, 持续时间2000~2009年, 喀麦隆在2016年出现突发性(图2B)。喀麦隆与法国、印度、德国、韩国、巴西、意大利等国之间存在合作, 与中国、美国无合作。

2.3 机构分析

2000~2016年, 药用植物基因组学研究发文前10的机构依次为中国医学科学院发文197篇, 占比6.1%; 中国科学院发文165篇, 占比5.1%; 韩国庆熙大学发文48篇, 占比1.5%; 中国中医科学院发文47篇, 占比1.5%; 韩国首尔国立大学发文46篇, 占比1.4%; 印度机构CSIR (Cent Inst Med & Aromat Plants)发文41篇, 占比1.3%; 德国Johannes Gutenberg Univ Mainz发文32篇, 占比1.0%; 浙江大学发文32篇, 占比1.0%; 沙特阿拉伯King Saud Univ发文30篇, 占比0.9%; 上海交通大学发文30篇, 占比0.9%。

对中国医学科学院进行全球引用得分评估(total global citation score), 结果显示该机构在发文量和全球

图2 发文量及突发性检测

A: 发文量前10国家分布; B: 突发性检测

引用得分(TGCS)评估均为第一, 发文197篇、TGCS:3455分, 其次为中国科学院, 发文165篇, TGCS:1958分。中国科学院在合作网中显示出最广的合作范围, 中国医学科学院次之(图3A)。对2890个在该领域有贡献的机构进行突发性检测, 结果显示, 一共有17个机构出现突发, 其中中国机构9家, 最先出现突发性的机构为韩国机构Korea Res Inst Biosci & Biotechnol, 且其延续时间最长, 2003~2010年。在2016年表现突发的机构有2个, 维也纳大学和中国中医科学院(图3B)。

2.4 关键词突发分析

分析中出现两个属名*Panax* (共63条记录, 占2.0%), *Salvia* (共61条记录, 占1.9%)。对所有7594个进行突发性检测, 得到73个有突发性的(图4)。AFLP在2002~2010年出现突发, 强度指数为5.7926。PCR在2002~2010年出现突发, 强度指数为4.3416。组织培养在2005~2010年出现突发, 强度指数为5.3424。分子标记在2007~2008年出现突发, 强度指数为3.3144。随机扩增多态性DNA标记(random amplified polymorphic DNA, RAPD), 分析在2008~2012年出现突发, 强度指数为16.3731。Phylogeny在2010~2012年出现突发, 强度指数为8.2686。Authentication在2010~2012年出现突发, 强度指数为7.7208。ISSR在2012~2014年出现突发, 强度指数为6.542。Metabolism在2012~2014年出现突发, 强度指数为9.1796。DNA Barcode在2013~2016年出现突发, 强度指数为7.9439。Molecular cloning在

2013~2016年出现突发, 强度指数为6.8679。

使用 D_{SUP1} 数据进行关键词突发性检测, 2016年出现突发的关键词有double strand break, homologous recombination, domestication, Cas9, CRISPRcas system, RNA guided endonuclease, genome modification等(详情见网络版附图1)。使用 D_{SUP2} 数据进行关键词突发性检测, 2016年出现突发的关键词有stem cell, small RNA, epidemiology, genome wide association, CRISPRcas system, Cas9, phylogenomics等(详情见网络版附图2)。

2.5 研究方向动态分析

对研究方向进行突发性检测, 在2016年出现突发的研究方向有3个, 多学科、进化生物学和其他, 其中多学科在2013~2016年出现突发, 强度指数为6.063, 进化生物学在2014~2016年出现突发, 强度指数为5.4873。

2.6 作者分析

对所有12165位学者进行分析, 发文量排名前10学者, 7位为中国学者, 依次为Chen S L (47条记录, 占比1.5%), Wang Y (45条记录, 占比1.4%), Li Y (38条记录, 占比1.2%), 德国学者Efferth T (35条记录, 占比1.1%), Song J Y (29条记录, 占比0.9%), Liu Y (28条记录, 占比0.9%), 印度学者Kumar S (23条记录, 占比0.7%), Sun C (22条记录, 占比0.7%), 韩国学者Kim Y J (21条记录, 占比0.7%) (图6A)。Chen S L亮点工作有DNA条形码ITS2在植物鉴定中应用、转录组学、叶绿体全基因组

图3 国际机构合作和突发性检测
A: 国际机构合作关系网; B: 机构突发性检测

图4 关键词突发性检测

图 5 研究方向突发性检测

测序、灵芝全基因组测序等, Wang Y 主要工作是转录组学功能基因的研究、分子鉴定等, Li Y 主要工作含 DNA 条形码、转录组等, 德国学者 Efferth T 主要工作是本草疗法、分子靶标、分子药理学等, Song J Y 亮点工作含 DNA 条形码 ITS2 在植物鉴定中应用、叶绿体全基因组、灵芝全基因组测序等, Liu Y 主要工作是围绕丹参、冷蒿、地黄、人参等, 印度学者 Kumar S 主要工作有微卫星标记技术、转录组学等, Sun C 主要工作有转录组、生物合成等, 韩国学者 Kim Y J 主要工作是人参转录组、人参皂苷生物合成等。发文量排名前 10 学者中印度学者 Kumar S 在合作网分析中几乎表现为孤立状态, 不与他国研究者合作。图中表现为几乎孤立的作者群含印度学者、韩国学者、德国学者(图 6B)。中国学者在该项研究中占据主要位置, 陈士林在合作网中表现出重要位置。

2.7 文献共被引分析

文献共被引分析可展示具体文献的被引及共被引情况, 从图 7A 可见, 被引文献形成两大中心, 一边以 Grabherr M G 2011 年发表在 *Nature Biotechnology* 题为 Full-length transcriptome assembly from RNA-Seq data without a reference genome 的文章为核心形成, 另

一边以 Tamura K 2011 年发表在 *Molecular Biology and Evolution* 题为 MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods, Chen S L 2010 年发表在 *PLoS ONE* 题为 Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species, Hollingsworth P M 2009 年发表在 *Proceedings of the National Academy of Sciences of the United States of America* 题为 A DNA barcode for land plants 为核心形成。对共被引文献进行关键词提取分析, 关键词叶绿体全基因组表现出强势劲头, 此外, 人参与转录组分析存在较多联系, 说明人参的转录组分析研究较多, 丹参与生物合成、转录组有较多联系(图 7B)。

2.8 前 10 论文

2000~2016 年关于药用植物基因组学研究论文共有 3208 篇, 对所有论文进行引用频次分析, 以年均本地被引得分 (LCS/t=local citation score per year) 排序, 列举被引次数最高的 10 篇论文(表 1), 年均本地被引得分 (LCS/t) 指下载文献组成的数据库中的文献对这些文献的引用度。基于 LCS 表明被该领域研究者所关注文章。

图 6 发文量及作者间合作
A: 发文量前10作者; B: 作者间合作关系图

3 结论与讨论

纵观全球研究, 论文发表数量涨势明显. 中国、印

度、美国、韩国、巴西、日本、德国、伊朗、马来西亚、英国等国家对药用植物基因组学开展了较多研究. 主要研究机构为中国医学科学院、中国科学院、

图 7 文献共被引分析

A: 文献共被引分析图; B: 文献共被引关键词分析

表 1 2000~2016年年均全球被引得分最多的前10篇SCI论文

论文标题	杂志/年份	年均被引得分 LCS/t
DNA barcoding of medicinal plant material for identification	<i>Current Opinion in Biotechnology</i> /2014	6.33
Identification of medicinal plants in the family <i>Fabaceae</i> using a potential DNA barcode ITS2	<i>Journal of Ethnopharmacology</i> /2010	4.29
454 EST analysis detects genes putatively involved in ginsenoside biosynthesis in <i>Panax ginseng</i>	<i>Plant Cell Reports</i> /2011	4.17
Genome-wide identification and characterization of novel genes involved in terpenoid biosynthesis in <i>Salvia miltiorrhiza</i>	<i>Journal of Experimental Botany</i> /2012	3.80
A renaissance in herbal medicine identification: from morphology to DNA	<i>Biotechnology Advances</i> /2014	3.67
Cloning and characterization of 2-C-methyl-d-erythritol-4-phosphate pathway genes for isoprenoid biosynthesis from Indian ginseng, <i>Withania somnifera</i>	<i>Protoplasma</i> /2013	3.50
Full-length transcriptome sequences and splice variants obtained by a combination of sequencing platforms applied to different root tissues of <i>Salvia miltiorrhiza</i> and tanshinone biosynthesis	<i>Plant Journal</i> /2015	3.50
Authentication of the family <i>Polygonaceae</i> in Chinese pharmacopoeia by DNA barcoding technique	<i>Journal of Ethnopharmacology</i> /2009	3.38
DNA barcoding: an efficient tool to overcome authentication challenges in the herbal market	<i>Plant Biotechnology Journal</i> /2016	3.00
Genome-based approaches to the authentication of medicinal plants	<i>Planta Medica</i> /2008	2.89

韩国庆熙大学、中国中医科学院、韩国首尔国立大学、印度机构CSIR、德国Johannes Gutenberg Univ Mainz、浙江大学、沙特阿拉伯King Saud Univ、上海交通大学. 国家突发性检测表明, 喀麦隆是在2016年出现突发的国家, 可知该国在药用植物基因组学研究上属于新兴力量, 未来有较大可能在该领域继续发力. 机构突发性检测表明, 在2016年共有两个机构出现突发性, 中国中医科学院和维也纳大学, 其中中国中医科学院作为发文量排名前10的机构出现突发性, 实力不容小觑. 未来, 中国将继续担当药用植物基因组学研究主力, 在很长一段时间内, 中国医学科学院会保持在该领域的研究优势, 但仍需加强国际合作.

突发性检测显示, PCR在2005~2010年出现突发, 强度指数为3.5028. AFLP在2007~2010年出现突发, 强度指数为5.2496. RAPD在2008~2012年出现突发, 强度指数为15.5820. Authentication在2010~2012年出现突发, 强度指数为6.0053. ISSR在2012~2014年出现突发, 强度指数为4.2826. Metabolism在2012~2014年出现突发, 强度指数为5.197. DNA Barcode在2013~2016年出现突发, 强度指数为8.1778. Molecular cloning在

2013~2016年出现突发, 强度指数为4.2088. Diversity在2014~2016年出现突发, 强度指数为3.3591. 预测未来药用植物基因组学会更多地整合代谢组学、转录组学等多组学技术^[27], 分子克隆技术等. 2016年出现突发的关键词有Cas9, CRISPRcas system, genome modification, small RNA, genome wide associatioion, phylogenomics等, 预测药用植物基因组学的研究会涉及基因编辑技术^[28]. 领域分析表明, 未来多学科交叉是趋势, 本草基因组学在进化生物学上应用将加强. 通过文献分析, 该领域共有6位学者Tamura K, Kress W J, Hollingsworth P M, Chen S L, Grabherr M G发表了对本领域具有较大影响力的文章. 文献共被引分析显示, 学者陈士林发表文章处于被引中心位置, 在本草基因组学国际研究上占据首要位置(国际研究占1.4%), 说明其在本草基因组学发展中起关键作用, 有重要贡献. 近10年, 陈士林及其团队提出利用基因组学来解决中药面临问题的新学科——本草基因组学, 为利用基因组学解决药用植物面临的品种混乱、退化、品质、道地性等科学问题提供了理论和方法. 未来, 中国在药用植物基因组学研究上将继续保持领军作用.

参考文献

- 1 Bild A H, Chang J T, Johnson W E, et al. A field guide to genomics research. *PLoS Biol*, 2014, 12: e1001744
- 2 Paterson A H, Bowers J E, Burow M D, et al. Comparative genomics of plant chromosomes. *Plant Cell*, 2000, 12: 1523–1540
- 3 Chen S, Xu J, Liu C, et al. Genome sequence of the model medicinal mushroom *Ganoderma lucidum*. *Nat Commun*, 2012, 3: 913
- 4 宋经元, 罗红梅, 李春芳, 等. 丹参药用模式植物研究探讨. *药学报*, 2013, 48: 1099–1106
- 5 Luo H, Zhu Y, Song J, et al. Transcriptional data mining of *Salvia miltiorrhiza* in response to methyl jasmonate to examine the mechanism of bioactive compound biosynthesis and regulation. *Physiol Plantarum*, 2015, 152: 241
- 6 黄林芳, 郑司浩, 武拉斌, 等. 基于化学成分及分子特征中药材肉苁蓉生态型研究. *中国科学: 生命科学*, 2014, 44: 318–328
- 7 徐荣, 陈君, 陈士林, 等. 肉苁蓉种质资源多样性的 AFLP 分析. *中草药*, 2007, 38: 1703–1707
- 8 Chen S, Pang X, Song J, et al. A renaissance in herbal medicine identification: from morphology to DNA. *Biotech Adv*, 2014, 32: 1237–1244
- 9 Wenping H, Yuan Z, Jie S, et al. *De novo* transcriptome sequencing in *Salvia miltiorrhiza* to identify genes involved in the biosynthesis of active ingredients. *Genomics*, 2011, 98: 272–279
- 10 陈士林, 朱孝轩, 李春芳, 等. 中药基因组学与合成生物学. *药学报*, 2012, 47: 1070–1078
- 11 邹丽秋, 匡雪君, 孙超, 等. 天然产物生物合成途径解析策略. *中国中药杂志*, 2016, 41: 4119–4123
- 12 Luo H, Chao S, Sun Y, et al. Analysis of the transcriptome of *Panax notoginseng* root uncovers putative triterpene saponin-biosynthetic genes and genetic markers. *Bmc Genomics*, 2011, 12: 1–15
- 13 Zhang J J, Su H, Zhang L, et al. Comprehensive characterization for ginsenosides biosynthesis in ginseng root by integration analysis of chemical and transcriptome. *Molecules*, 2017, 22: 889
- 14 Li Y, Luo H M, Sun C, et al. EST analysis reveals putative genes involved in glycyrrhizin biosynthesis. *BMC Genomics*, 2010, 11: 268
- 15 陈士林, 宋经元. 本草基因组学. *中国中药杂志*, 2016, 41: 3881–3889
- 16 陈士林, 何柳, 刘明珠, 等. 本草基因组方法学研究. *世界科学技术-中医药现代化*, 2010, 12: 316–324
- 17 Han J, Pang X, Liao B, et al. An authenticity survey of herbal medicines from markets in China using DNA barcoding. *Sci Rep*, 2016, 6: 18723
- 18 Peltonen L, Mckusick V A. Genomics and medicine: dissecting human disease in the postgenomic era. *Science*, 2001, 291: 1224–1229
- 19 陈晓辰, 宋经元, 董林林, 等. 宏基因组学与道地药材研究. *中草药*, 2012, 43: 2315–2320
- 20 Li Q, Li Y, Song J, et al. High-accuracy *de novo* assembly and SNP detection of chloroplast genomes using a SMRT circular consensus sequencing strategy. *New Phytol*, 2014, 204: 1041–1049
- 21 陈士林, 孙永珍, 徐江, 等. 本草基因组计划研究策略. *药学报*, 2010, 45: 807–812
- 22 Mukerji M, Prasher B. Genomics and traditional indian ayurvedic medicine. In: *Genomics and Society*. 2016. 271–292
- 23 Chiu W T, Ho Y S. Bibliometric analysis of tsunami research. *Scientometrics*, 2007, 73: 3–17
- 24 樊怡菁. SCIE 和 Scopus 引文功能的比较分析. *现代情报*, 2006, 26: 80–82
- 25 李杰, 陈超美. Citespace. 北京: 首都经济贸易大学出版社, 2016
- 26 Garfield E., Paris S W, Stock W G. HistCite™: a software tool for informetric analysis of citation linkage. *NFD Information-Wissenschaft Und Praxis*, 2006, 57: 391–400
- 27 Chen S, Yao H, Han J, et al. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLoS ONE*, 2010, 5: e8613
- 28 Sander J D, Joung J K. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol*, 2014, 32: 347–355

Trend analysis of international studies in herbgenomics

YANG QiaoQiao & HUANG LinFang

Key Research Laboratory of Traditional Chinese Medicine Resources Protection, State Administration of Traditional Chinese Medicine (National Administration of Traditional Chinese Medicine), Institute of Medicinal Plants, Beijing Academy of Medical Sciences, Chinese Academy of Medical Sciences, Beijing 100193, China

Since the completion of the Human Genome Project in 2000, the science of genomics rapidly developed. The application of genomics in medicinal plant research has become popular. In this study, we retrieved international research essays on medicinal plant genomics from 2000 to 2016 from the SCIE bibliographic database and performed data mining using the analysis platform CiteSpace and the analysis tool HistCite. The analysis revealed that studies conducted in China, India and the United States accounted for 29.3%, 19.2% and 10.8% of the total studies, respectively. The Chinese Academy of Medical Sciences, the Chinese Academy of Sciences, and Kyung Hee University as the main research institutions, accounted for 6.1%, 5.1% and 1.5% of the total studies, respectively. Chinese scholar Shilin Chen has played an important role in this research area, putting forward a new discipline, called “herbgenomics”. Future forecasting analysis revealed that China will continue to be the leading country in medicinal plant genomics, Cameroon and Vienna would engage in this research, and genomics of medicinal plants will be more closely integrated with multi-omics technologies, such as metabolomics, transcriptomics, etc. In addition, gene editing technology may provide new approaches into genomic studies of medicinal plants.

genomics, medicinal plant, global change, bibliometric analysis

doi: [10.1360/N052017-00200](https://doi.org/10.1360/N052017-00200)