

代谢组学研究的文献计量分析

李颖^{1,2*}, 王祁民², 李倩^{1,2}, 李宵², 吴茵², 安静², 董占军^{1,2#} (1. 河北医科大学研究生学院, 石家庄 050017; 2. 河北省人民医院药学部, 石家庄 050051)

中图分类号 Q503 文献标志码 A 文章编号 1001-0408(2017)23-3174-06
DOI 10.6039/j.issn.1001-0408.2017.23.02

摘要 目的:了解国内外代谢组学的研究现状,初步探讨其研究热点及发展规律。方法:采用文献计量方法,借助GoPubMed文献分析工具,以“Metabolomics”为主题词,检索PubMed数据库截至2016年7月31日的所有文献,对相关文献按发文章量、国家和城市、期刊来源、研究主题、核心作者和作者合作网络等信息进行排名并分析。以“Metabolomics”“Metabonomics”为主题词,检索Web of Science数据库截至2016年8月1日的所有文献,利用其被引频次排序功能,分析2007—2016年被引频次较高的文献。结果:共检索得到国内外代谢组学相关文献15 449篇,其中我国研究者发表2 189篇,发文章量总体均呈上升趋势。15 449篇文献分布于全球129个国家和地区,美国、中国、英国、德国等10个国家为核心发文章国家,发文章总量为12 847篇(占83.16%);发文章量最多的国家是美国(4 288篇,占27.76%),其次为中国(2 189篇,占14.17%);发文章量最多的城市是伦敦(467篇),上海和北京分列第2、3位。文献共涉及期刊2 168种,刊均载文章量为7.1篇;发文章量最多的20种期刊均来自于欧美国家(4 377篇,占28.33%)。出现频次最高的前3个主题词分别为代谢组学、代谢组和代谢,相关研究主要集中在代谢过程、代谢通路、生物标志物、蛋白质组和基因组等方面,研究方法主要包括质谱法和磁共振波谱法。主要涉及6个作者群,以Nicholson JK、Holmes E、Lindon JC等研究者为主;核心作者及作者群中未见我国研究者。高被引文献主要涉及代谢与疾病的相关性、代谢组学数据库简介、代谢组学研究方法及与疾病预测有关的生物标志物等内容。结论:代谢组学引起了世界范围内研究者的广泛兴趣,其中代谢通路和生物标志物是该领域的研究热点之一。我国研究者在代谢组学方面的总发文章量较多,但缺乏高质量的期刊载体,且研究者之间的合作较少。我国应鼓励各研究机构间的交流合作,紧紧把握该领域研究的前沿与热点,促进我国代谢组学研究的多元化和深层次化。
关键词 代谢组学; PubMed; Web of Science; 文献计量; 被引频次

Bibliometric Analysis of Metabolomics

LI Ying^{1,2}, WANG Qimin², LI Qian^{1,2}, LI Xiaojiao², WU Yin², AN Jing², DONG Zhanjun^{1,2} (1. Graduate School of Hebei Medical University, Shijiazhuang 050017, China; 2. Dept. of Pharmacy, Hebei General Hospital, Shijiazhuang 050051, China)

ABSTRACT OBJECTIVE: To investigate the present situation of domestic and foreign metabolomics study, and to preliminarily discuss research hotspots and development rules. METHODS: By using bibliometric methods and GoPubMed literature analysis tool, using “Metabolomics” as subject, all literatures were retrieved from PubMed database up to Jul. 31, 2016. Those literatures were ranked and analyzed in respects of publication amount, countries and cities, journal sources, research topics, core authors and authors collaborative networks. Using “Metabolomics” “Metabonomics” as subjects, all literatures were retrieved from Web of

- tion via ZEB2 targets in gastric cancer[J]. *J Cancer Res Clin Oncol*, 2017, 143(3): 385-397.
- [15] Yang H, Wu XL, Wu KH, et al. MicroRNA-497 regulates cisplatin chemosensitivity of cervical cancer by targeting transketolase[J]. *Am J Cancer Res*, 2016, 6(11): 2690-2699.
- [16] Zhou N, Fei D, Zong S, et al. MicroRNA-138 inhibits proliferation, migration and invasion through targeting hTERT in cervical cancer[J]. *Oncol Lett*, 2016, 12(5): 3633-3639.
- [17] Lai XJ, Cheng XY, Hu LD. MicroRNA 421 induces apoptosis of c-33a cervical cancer cells via down-regulation of Bcl-xL[J]. *Genet Mol Res*, 2016, doi:10.4238/gmr15048853.
- [18] Russo V, Bernabò N, Di Giacinto O, et al. H3K9 trimethylation precedes DNA methylation during sheep oogenesis: HDAC1, SUV39H1, G9a, HP1, and Dnmts are involved in these epigenetic events[J]. *J Histochem Cytochem*, 2013, 61(1):75-89.
- [19] Leung DC, Dong KB, Maksakova IA, et al. Lysine methyltransferase G9a is required for de novo DNA methylation and the establishment, but not the maintenance, of proviral silencing[J]. *Proc Natl Acad Sci U S A*, 2011, 108(14):5718-5723.
- [20] 赵丽, 焦运, 杨安宁, 等. MiR-124 及其启动子区 DNA 甲基化在同型半胱氨酸致动脉粥样硬化中的作用[J]. *生理学报*, 2015, 67(2): 207-213.

* 硕士研究生。研究方向:临床药理学。电话:0311-85988998。E-mail:lyyaoda@126.com

通信作者:主任药师,博士。研究方向:医院药学。电话:0311-85988604。E-mail:13313213656@126.com

(收稿日期:2016-12-22 修回日期:2017-06-15)
(编辑:张元媛)

Science database up to Aug. 1, 2016. The literatures with high citation frequency were analyzed during 2007-2016 by using the sorting function of citation frequency. RESULTS: A total of 15 449 domestic and foreign metabolomics literatures were retrieved. The researchers of 2 189 literatures came from China, and the amount of published literatures showed a rising trend. 15 449 literatures were distributed in 129 countries and regions; 10 countries, such as the United States, China, Britain and Germany, were the core countries, and the total number of literatures issued was 12 847 (83.16%). The United States held the largest share of world publications (4 288 literatures, 27.76%), and the following was China in order (2 189 literatures, 14.17%). The city with the largest amount of publications was London (467 literatures), followed by Shanghai, Beijing. A total of 2 168 periodicals were involved, and there were 7.1 published literatures averagely. All of the top 20 periodicals were from the United States and European countries (4 377 literatures, 28.33%). The first 3 subjects with the highest frequency of occurrence were metabolomics, metabolome and metabolism; related researches mainly focused on metabolic processes, metabolic networks and pathways, biological markers, proteome and genome; main research methods included spectrum analysis and magnetic resonance spectroscopy. Main 6 groups of authors were involved, mainly Nicholson JK, Holmes E, Lindon J and other researchers. None of Chinese researchers had been found among the core authors and author groups. The literatures with high citation frequency mainly focused on the association of metabolism with disease, metabolomics database, metabolomics research methods and biological markers related to disease prediction. CONCLUSIONS: Metabolomics has aroused worldwide interest among researchers, and metabolic pathways and biological markers are the focuses in this field. Our researchers have published a large amount of literatures on metabolomics, but there are not high quality periodical carrier or enough cooperation between researchers. It is suggested to enhance the cooperation between various research institutions or grasp the frontier and hotspots of the research in this field so as to push forward the development diversification and depth of metabolomics research in China.

KEYWORDS Metabolomics; PubMed; Web of Science; Bibliometrics; Citation frequency

代谢组学(Metabolomics或Metabonomics)是继基因组学、转录组学、蛋白质组学后系统生物学的另一个重要组成部分,是借助现代分析技术、化学计量学和生物信息学,定量测定生物体受到外界刺激前后细胞中小分子(分子量 $\leq 1\ 000$ Da)的动态变化规律,建立系统代谢图谱,并确定这些变化与生物过程相关性的一门学科^[1]。Metabonomics的概念来源于Nicholson JK小组利用核磁共振(Nuclear magnetic resonance, NMR)技术对体液、细胞和组织中多组分代谢产物的研究,并由Nicholson JK教授于1999年首次提出^[2]。几乎同时,Metabolomics的概念由Fiehn O等^[3]在植物代谢产物研究过程中提出。经过不断发展,代谢组学的4个层次也被学术界广泛认可,即代谢物靶标分析(Metabolite target analysis,对某个或某几个特定组分进行分析)、代谢谱分析(Metabolic profiling analysis,对一系列预先设定的目标代谢物进行定量分析)、代谢组学分析(Metabonomics analysis,对某一生物或细胞的所有代谢物进行定性和定量分析)和代谢指纹分析(Metabolic fingerprinting analysis,不分离鉴定具体单一组分,而是对代谢物整体进行高通量的定性分析)^[4]。与其他组学相比,代谢组学有其独特的优点,包括基因和蛋白的细微变化会在代谢物上得以放大、代谢物的种类远少于基因和蛋白的数量、无需进行全基因组测序、生物体液的代谢物分析可反映生物体的病理生理状态等^[5]。代谢组学的这些优点使其被广泛应用于生命科学的众多领域,成为系统生物学研究领域最活跃的分支之一,已有越来越多的学者加入到了代谢组学的研究中。

PubMed是由美国国家生物技术信息中心(National Center for Biotechnology Information, NCBI)开发研制的基于网络数据库的检索系统,于1997年6月开通,并免

费向读者提供文献检索服务,现已成为网上检索生物医学文献使用率最高的网站。GoPubMed文献分析工具是以PubMed为数据源,利用基于语义分类工具[基因主体(Gene ontology)和医学主题词(Medical subject headings, MeSH)]对检索结果进行术语提取和匹配,将检索结果转换为分类类目与对应文献之间关系的可视化数据,帮助用户找到最相关的文献,为文献统计和研究热点追踪提供了很大帮助^[6]。

Web of Science是由美国Thomson Scientific公司基于网络开发的大型综合性多学科的核心期刊引文数据库,包括科学引文索引(Science Citation Index, SCI)、社会科学引文索引(Social Science Citation Index, SSCI)、艺术和人文引文索引(Arts & Humanities Citation Index, AHCI),是全球最大、覆盖学科最多的综合性学术信息资源。其中,引文检索是其最具特色的功能之一,研究者可通过该功能获知文献被引用的情况,以获得高质量的文献^[7]。

为深入了解代谢组学研究文献增长及学科发展的规律,本研究利用文献计量学方法,借助PubMed、Web of Science数据库及相关文献分析工具对“代谢组学”相关英文文献进行在线统计和分析,探寻代谢组学的研究轨迹,为研究人员更加深入、系统地开展代谢组学研究提供参考。

1 资料与方法

以“Metabolomics”为主题词(GoPubMed工具自动将“Metabolomics”“Metabonomics”匹配为“Metabolomics”),检索PubMed数据库更新至2016年7月31的所有文献,时间跨度为1998—2016年。借助GoPubMed文献分析工具(www.GoPubMed.com/web/GoPubMed/),将

检索结果通过“Statistics”功能进行统计分析,汇总PubMed数据库中收录代谢组学研究文献的基本情况,包括发文量、发文国家和城市、发文期刊来源、研究主题、核心作者和作者合作网络。浏览GoPubMed左侧的分类导航目录,在“Locations”分类下点击“China”,选择“with China”,以系统自动匹配的形式(“Metabolomics[MeSH] China[Geo]”)检索文献,再通过“Statistics”功能进行分析,了解我国代谢组学文献在PubMed数据库中的收录情况。

以“Metabolomics”“Metabonomics”为主题词,检索Web of Science数据库(www.webofknowledge.com)更新至2016年8月1日的所有文献,时间跨度为2007—2016年。利用Web of Science的被引频次排序功能,汇总并分析近10年的高被引文献。

2 结果与分析

2.1 发文量

以“Metabolomics”检索,共得到相关文献15 449篇,以此作为国内外代谢组学相关文献的统计来源;以“Metabolomics[MeSH] China[Geo]”检索,共得到相关文献2 189篇,以此作为我国研究者代谢组学相关文献的统计来源,详见表1(其中2016年只有截至7月31日的文献)。

表1 1998—2016年PubMed数据库中代谢组学相关文献的年度发文量(篇)

Tab 1 Annual amount of published literatures on metabolomics in PubMed database during 1998-2016(piece)

年份	国内研究者文献	国内外研究者文献
1998年	0	1
1999年	0	2
2000年	0	15
2001年	0	22
2002年	0	63
2003年	0	109
2004年	3	145
2005年	14	251
2006年	14	296
2007年	35	464
2008年	52	600
2009年	98	862
2010年	95	1 011
2011年	159	1 275
2012年	225	1 668
2013年	287	1 883
2014年	345	2 297
2015年	438	2 540
2016年	424	1 945
合计	2 189	15 449

由表1可见,国内外代谢组学相关文献的发文量总体呈上升趋势,2014年的年度发文量已超过2 000篇,提示已有越来越多的国内外科学工作者投入到代谢组学的研究中,并获得了大量的研究成果,代谢组学已逐渐成为系统生物学研究领域的热点之一^[8]。我国在这一领域也

紧随国际前沿,自2004年起,PubMed数据库开始收录我国研究者的科研论文,发文量总体呈逐年上升的趋势。

2.2 发文国家和城市

15 449篇文献分布于全球129个国家和地区,其中美国东部、中国东部、欧洲和日本的发文密度较高,详见图1。发文量排名前10位的国家依次为美国、中国、英国、德国、日本、西班牙、意大利、加拿大、荷兰和法国,发文总量为12 847篇,占总发文量的83.16%,为核心发文国家,详见表2。其中,美国的发文量最多(4 288篇),占总发文量的27.76%;其次为中国(2 189篇),占14.17%,提示我国研究者在代谢组学研究领域已获得一定的成果,为代谢组学的发展做出了努力和贡献。



图1 1998—2016年PubMed数据库中代谢组学相关文献的地域分布图

Fig 1 Geographic distribution of published literatures on metabolomics in PubMed database during 1998-2016

表2 1998—2016年PubMed数据库中代谢组学相关文献发文量排名前10位的国家

Tab 2 Top 10 countries in the list of the amount of metabolomics literatures in PubMed database during 1998-2016

排序	国家	发文量,篇(%)
1	美国	4 288(27.76)
2	中国	2 189(14.17)
3	英国	1 366(8.84)
4	德国	1 194(7.73)
5	日本	808(5.23)
6	西班牙	646(4.18)
7	意大利	610(3.95)
8	加拿大	597(3.86)
9	荷兰	586(3.79)
10	法国	563(3.64)
合计		12 847(83.16)

发文量排名前20位的城市多为发达城市,以美国和中国城市最多,均为5个。其中,伦敦以467篇的发文量位居首位。上海和北京分列于第2、3位,发文量分别为413、408篇;南京、武汉和哈尔滨分别位于第6、17和19位,发文量分别为181、145、144篇,这5个城市的发文量占我国总发文量58.98%,详见图2。

2.3 发文期刊来源

代谢组学相关文献发表于2 168种期刊,刊均载文量为7.1篇;发文量最多的20种期刊(见表3)均来自于欧美国家,共发文4 377篇,占总发文量的28.33%,平均

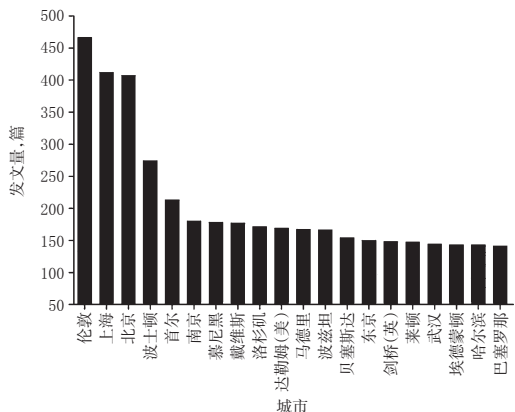


图2 1998—2016年PubMed数据库中代谢组学相关文献发文量排名前20位的城市

Fig 2 Top 20 cities in the list of the amount of metabolomics literatures in PubMed database during 1998-2016

影响因子为4.147。其中,《公共科学图书馆期刊》(*Public Library of Science One, PLoS One*)以764篇的发文量位居首位;《美国国家科学院院刊》(*Proceedings of the*

表3 1998—2016年PubMed数据库中代谢组学相关文献发文量排名前20位的期刊及影响因子
Tab 3 Top 20 periodicals in the list of the amount of metabolomics literatures and impact factor in PubMed database during 1998-2016

排序	期刊名	影响因子 (2015年)	发文量, 篇	排序	期刊名	影响因子 (2015年)	发文量, 篇
1	《公共科学图书馆期刊》(<i>PLoS One</i>)	3.057	764	11	《生物信息学》(<i>Bioinformatics</i>)	5.766	141
2	《蛋白质组学研究》(<i>Journal of Proteome Research, J Proteome Res</i>)	4.173	534	12	《药学与生物医学分析杂志》(<i>Journal of Pharmaceutical and Biomedical analysis, J Pharm Biomed Anal</i>)	3.169	140
3	《分析化学》(<i>Analytical Chemistry, Anal Chem</i>)	5.886	508	13	《植物生理学》(<i>Plant Physiology, Plant Physiol</i>)	6.258	127
4	《分子生物学方法》(<i>Methods in Molecular Biology, Methods Mol Biol</i>)	-	248	14	《色谱法杂志: B辑—生物医学与生命科学应用》(<i>Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences, J Chromatogr B Analyt Technol Biomed Life Sci</i>)	2.678	125
5	《分子生物系统》(<i>Molecular Biosystems, Mol Biosyst</i>)	2.829	227	15	《生物分析》(<i>Bioanalysis</i>)	2.813	123
6	《科学报告》(<i>Scientific Reports, Sci Rep</i>)	5.228	202	16	《分析化学学报》(<i>Analytica Chimica Acta, Anal Chim Acta</i>)	4.712	123
7	《分析化学和生物分析化学》(<i>Analytical and Bioanalytical Chemistry, Anal Bioanal Chem</i>)	3.125	195	17	《电泳》(<i>Electrophoresis</i>)	2.482	110
8	《农业食品化学杂志》(<i>Journal of Agricultural and Food Chemistry, J Agric Food Chem</i>)	2.912	175	18	《伦敦生物医学中心生物信息学杂志》(<i>BMC Bioinformatics</i>)	2.435	107
9	《色谱A》(<i>Journal of Chromatography A, J Chromatogr A</i>)	3.926	168	19	《美国国家科学院院刊》(<i>Proc Natl Acad Sci U S A</i>)	9.423	104
10	《代谢组学》(<i>Metabolomics</i>)	3.661	161	20	《生物化学》(<i>The Journal of Biological chemistry, J Biol Chem</i>)	4.258	95

注:“-”指该杂志为系列丛书,无影响因子
Note:“-”refers to a series of books, without impact factor

表4 1998—2016年PubMed数据库中代谢组学相关文献中出现频次最高的20个关键词
Tab 4 Top 20 keywords of metabolomics literatures in the list of citation frequency in PubMed database during 1998-2016

排序	关键词	发文量, 篇	排序	关键词	发文量, 篇
1	代谢组学(Metabolomics)	15 449	11	蛋白质组学(Proteomics)	2 909
2	代谢组(Metabolome)	13 439	12	基因(Genes)	2 844
3	代谢(Metabolism)	9 120	13	蛋白质组(Proteome)	2 773
4	代谢过程(Metabolic process)	8 952	14	磁共振波谱(Magnetic resonance spectroscopy)	2 607
5	人类(Humans)	6 014	15	蛋白质(Proteins)	2 588
6	质谱分析(Spectrum analysis)	5 404	16	氨基酸(Amino acids)	2 547
7	动物(Animals)	4 773	17	基因组学(Genomics)	2 489
8	质谱(Mass spectrum)	4 172	18	女/雌性(Female)	2 420
9	生物标志物(Biological marker)	3 854	19	代谢通路(Metabolic networks and pathways)	2 274
10	男/雄性(Male)	3 203	20	基因组(Genome)	2 265

National Academy of Sciences of the United States of America, Proc Natl Acad Sci U S A)的影响因子最高,为9.423,共收录代谢组学相关文献104篇,详见表3。

2.4 研究主题

15 449篇代谢组学相关文献中,出现频次最高的前3个关键词分别为代谢组学(Metabolomics)、代谢组(Metabolome)和代谢(Metabolism)。代谢组学研究涉及人类(Humans)和动物(Animals),包括男/雄性(Male)和女/雌性(Female);其中,研究人类代谢组学的相关文献有6 014篇,高于研究动物模型的4 773篇。代谢组学的研究内容主要集中在代谢过程(Metabolic process)、生物标志物(Biological marker)、蛋白质组(Proteomics)、代谢通路(Metabolic networks and pathways)和基因组(Genome)等方面,以代谢过程(Metabolic process)的相关研究居多,有8 952篇。代谢组学的研究方法主要包括质谱(Mass spectrum)法和磁共振波谱(Magnetic resonance spectroscopy)法,详见表4。

2.5 核心作者及作者合作网络

发文量最多的作者是来自英国帝国理工大学的Ni-

cholomon JK 教授,共发表相关文献114篇。对作者合作网络进行可视化展示,详见图3(作者之间通过短虚线、长虚线、实线、粗体实线进行连接,连线越多、线型越粗表明该作者的研究地位越重要)。由图3可见,代谢组学研究主要有6个作者群,其中来自于英国帝国理工大学的Nicholson JK、Holmes E、Lindon J等与其他作者联系最为紧密,为核心研究团队;而其他5个作者群内的联系则较为松散。核心作者及作者群中未见我国研究者。

2.6 高被引文献分析

2007—2016年,代谢组学相关文献被引频次最高的是2011年发表在*Nature*上的“Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease”^[9],内容涉及代谢与疾病的相关性,被引频次为839次;其余高被引文献涉及代谢组学数据库简介、代谢组学的研究方法和与疾病预测有关的生物标志物研究,与表4的内容基本相符,详见表5。

3 讨论

随着人类基因组测序工作的完成,基因功能逐渐成为研究热点,随之出现了一系列“组学”的研究,代谢组学是众多组学中的一种,是随着生命科学的发展而发展起来的。本文利用文献计量学方法对PubMed和Web of Science数据库中代谢组学相关文献进行统计和分析,对该领域的发展趋势和热点进行揭示。

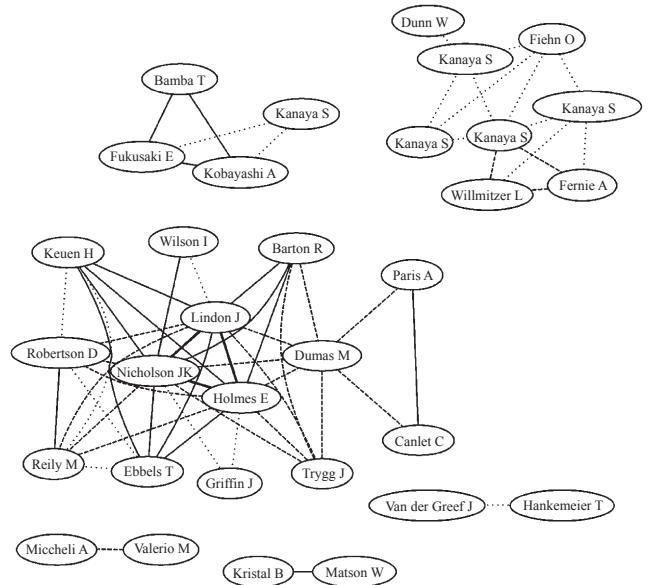
表5 2007—2016年被引频次排名前10位的代谢组学相关文献

Tab 5 Top 10 metabolomics literatures in the list of citation frequency during 2007-2016

题名	作者	来源期刊	发表年份	被引次数
Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease	Wang ZN等 ^[9]	<i>Nature</i>	2011年	839
HMDB: the human metabolome database	Wishart DS等 ^[10]	<i>Nucleic Acids Research</i>	2007年	823
HMDB: a knowledgebase for the human metabolome	Wishart DS等 ^[11]	<i>Nucleic Acids Research</i>	2009年	746
Mass spectrometry-based metabolomics	Dettmer K等 ^[12]	<i>Mass Spectrometry Reviews</i>	2007年	727
Metabolite profiles and the risk of developing diabetes	Wang TJ等 ^[13]	<i>Nature Medicine</i>	2011年	693
Mitochondrial overload and incomplete fatty acid oxidation contribute to skeletal muscle insulin resistance	Koves TR等 ^[14]	<i>Cell Metabolism</i>	2008年	683
Systems biology: metabolomics	Nicholson JK等 ^[15]	<i>Nature</i>	2008年	638
Chemometrics in metabolomics	Trygg J等 ^[16]	<i>Journal of Proteome Research</i>	2007年	611
HMDB 3.0: the human metabolome database in 2013	Wishart DS等 ^[17]	<i>Nucleic Acids Research</i>	2013年	564
Proposed minimum reporting standards for chemical analysis	Sumner LW等 ^[18]	<i>Metabolomics</i>	2007年	528

本文对1998—2016年PubMed数据库中代谢组学相关文献进行计量分析和可视化分析。结果发现,代谢组学相关文献的发量总体呈上升趋势;相关文献来自全球129个国家和地区,但分布不均,目前主要集中在美国东部、我国东部、欧洲和日本。我国的发量较大,主要集中在北京、上海、南京、武汉和哈尔滨等5个城市,但核心作者及作者群中未见我国研究者。因此,我国在加大代谢组学研究力度的同时,还应加强与各地区、各科研机构间的交流与合作,以及同其他地区或国家的跨区、跨国合作。

从发文期刊的来源来看,排名前20位的期刊全部来自于欧美国家,这也在一定程度上表明我国在这一研究领域缺乏高质量的期刊载体,研究基础较为薄弱。高频关键词反映了一个领域的研究热点,代谢组学虽然研



注: ···表示合作1次; - - -表示合作2次; —表示合作3~5次; —表示合作≥6次

Note: ··· refers to 1 collaboration; - - - refers to 2 collaborations; — refers to 3-5 collaborations; — refers to ≥6 collaborations

图3 1998—2016年PubMed数据库中代谢组学的作者合作网络

Fig 3 Author cooperation network of metabolomics in PubMed database during 1998-2016

究比较广泛,但是从高频关键词的统计和分析结果可见,代谢组学的研究热点主要集中在代谢过程和生物标志物等方面。

代谢组学研究对象的数量、大小、官能团、挥发性、极性以及其他物理化学参数差异很大,因此对分析技术的要求也更高。迄今为止,尚无一种能满足上述所有要求的代谢组学分析技术。目前,基于液相色谱-质谱联用的代谢组学技术因其高分离能力和高灵敏度已逐渐成为生命科学研究的重要手段,但由于缺乏有效的通用标准谱图库,使得大量代谢物的结构难以被确定,在一定程度上制约了代谢组学覆盖度的提高和生物标志物的发现,造成化学和生物信息的严重丢失,成为代谢组学发展的主要技术瓶颈^[19]。本研究结果显示,现有代谢组学的研究方法主要包括质谱法和核磁共振波谱法。

通过浏览文献发现,高频关键词“Spectrum analysis”多指液相色谱-质谱联用或者气相色谱-质谱联用技术,可能由于不同作者的表述方式不同,GoPubMed工具未能将其与质谱法(Mass spectrum)合并统计。

文献被引频次可反映一个研究领域的研究方向和热点,由于GoPubMed工具不能统计文献的被引频次,故本文借助Web of Science数据库,对代谢组学领域的高被引文献进行汇总,以更好地了解代谢组学领域的发展趋势。文献被引频次除了受其研究内容的影响,还受发文时间的影响,发文时间越早,被引频次可能越高^[20],因此为排除发表时间的影响,本文选取近10年(2007—2016年)发表的英文文献,按被引频次排序。排名前10位的高被引文献中,仅有3篇发表于2011年以后,且主要来自Nature、Nucleic Acids Research、Mass Spectrometry Reviews、Nature Medicine、Cell Metabolism等期刊,被引次数均超过500次。与其他组学相比,代谢组学更加接近表型,更适用于疾病分型和生物标志物的研究。被引频次排名前10位的文献中有2篇是关于代谢组学与疾病相关性的研究:一篇是Wang ZN等^[9]利用代谢组学方法揭示了磷脂酰胆碱被肠道微生物代谢,并促进心血管疾病的发生;另一篇是Wang TJ等^[13]利用代谢组学方法从多个代谢产物中发现支链氨基酸和芳香族氨基酸可以预测糖尿病的发生。有3篇文献是关于代谢组学数据库的介绍,分别是2007年的“HMDB: the human metabolome database”^[10]、2009年的“HMDB: a knowledgebase for the human metabolome”^[11]和2013年的“HMDB 3.0: the human metabolome database in 2013”^[17],代谢组学数据库的不断扩增也促进了代谢组学的发展。但是从总体来看,代谢组学仍然处于发展阶段,在研究方法、数据处理和生物解释等方面均面临着极大的挑战^[21]。我国科研人员应紧紧把握代谢组学的研究前沿与热点,深入优势领域进行研究,加强区域合作,促进我国代谢组学研究的多元化和深层化。

综上所述,本文利用文献计量学方法对代谢组学相关文献进行汇总分析,初步分析了代谢组学的研究热点及发展规律。由于文献系统的高度复杂性和不稳定性,故不能获得全部的有效信息以揭示代谢组学的宏观发展规律。深入研究代谢组学的发展趋势仍需要多种数学工具和统计学技术的支持,从多角度出发,对代谢组学现有文献进行深入挖掘和研究。

参考文献

[1] 周宁,颜贤忠. 代谢性疾病的代谢组学研究进展[J]. 国际药学研究杂志, 2015, 42(5): 561-567.

[2] Nicholson JK, Lindon JC, Holmes E. Metabonomics: understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data[J]. *Xenobiotica*, 1999, 29(11): 1181-1189.

[3] Fiehn O, Kopka J, Dormann P, et al. Metabolite profiling

for plant functional genomics[J]. *Nat Biotechnol*, 2000, 18(11): 1157-1161.

- [4] 赵春霞,许国旺. 基于液相色谱-质谱技术的代谢组学分析方法新进展[J]. *分析科学学报*, 2014, 30(5): 761-766.
- [5] 许国旺,路鑫,杨胜利. 代谢组学研究进展[J]. *中国医学科学院学报*, 2007, 29(6): 701-711.
- [6] 吕霖. PubMed及其衍生数据库在医药领域的应用[J]. *中国发明与专利*, 2014(8): 107-113.
- [7] 齐青. Web of Science的检索和应用[J]. *图书馆工作与研究*, 2013(2): 110-112.
- [8] 黄寅,许风国,张伟,等. 药物代谢组学研究进展[J]. *中国药科大学学报*, 2013, 44(2): 105-112.
- [9] Wang ZN, Klipfell E, Bennett BJ, et al. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease[J]. *Nature*, 2011, 472(7341): 57-82.
- [10] Wishart DS, Tzur D, Knox C, et al. HMDB: the human metabolome database[J]. *Nucleic Acids Res*, 2007, 35(Database issue): D521-D526.
- [11] Wishart DS, Knox C, Guo AC, et al. HMDB: a knowledgebase for the human metabolome[J]. *Nucleic Acids Res*, 2009, 37(Database issue): D603-D610.
- [12] Dettmer K, Aronov PA, Hammock BD. Mass spectrometry-based metabolomics[J]. *Mass Spectrom Rev*, 2007, 26(1): 51-78.
- [13] Wang TJ, Larson MG, Vasan RS, et al. Metabolite profiles and the risk of developing diabetes[J]. *Nat Med*, 2011, 17(4): 448-453.
- [14] Koves TR, Ussher JR, Noland RC, et al. Mitochondrial overload and incomplete fatty acid oxidation contribute to skeletal muscle insulin resistance[J]. *Cell Metab*, 2008, 7(1): 45-56.
- [15] Nicholson JK, Lindon JC. Systems biology: metabonomics[J]. *Nature*, 2008, 455(7216): 1054-1056.
- [16] Trygg J, Holmes E, Lundstedt T. Chemometrics in metabolomics[J]. *J Proteome Res*, 2007, 6(2): 469-479.
- [17] Wishart DS, Jewison T, Guo AC, et al. HMDB 3.0: the human metabolome database in 2013[J]. *Nucleic Acids Res*, 2013, 41(Database issue): D801-D807.
- [18] Sumner LW, Amberg A, Barrett D, et al. Proposed minimum reporting standards for chemical analysis[J]. *Metabolomics*, 2007, 3(3): 211-221.
- [19] 孔宏伟,戴伟东,许国旺. 基于液相色谱-质谱联用的代谢组学研究中代谢物的结构鉴定进展[J]. *色谱*, 2014, 32(10): 1052-1057.
- [20] 杨谨成,贾贝,陈伟,等. 化疗药物引起血糖变化的文献计量分析[J]. *中国药房*, 2016, 27(11): 1561-1565.
- [21] Xia J, Mandal R, Sinelnikov IV, et al. MetaboAnalyst 2.0: a comprehensive server for metabolomic data analysis[J/CD]. *Nucleic Acids Res*, 2012, 40 (Web Server issue): W127-W133.

(收稿日期:2016-08-12 修回日期:2017-02-13)

(编辑:张元媛)