

DOI:10.12113/201911001

共词分析国内外生物信息学领域研究态势

刘亚文,张红燕*,阳灵燕

(湖南农业大学 信息与智能科学技术学院,长沙 410128)

摘要:生物信息学作为自然科学领域中多学科交叉的新兴学科,其发展研究得到了众多学者的关注。为了解生物信息学在国内外的研究态势,以 CNKI 中文数据库和 Web of science 外文数据库中生物信息学领域期刊论文为研究对象,利用 R 语言编程工具,文献计量和共词分析归纳了国内外生物信息学领域的研究现状、热点及趋势。结果表明:国内外生物信息学研究均处于高速发展期,文献量呈逐年增长趋势,研究领域也在不断拓宽;国内外研究热点均聚焦在基因挖掘、蛋白质结构与功能预测、miRNA 分析等,但国内偏向于理论研究,而国际更注重其在疾病治疗、药物设计等方面的应用研究。

关键词:生物信息学;共词分析;R 语言;研究态势

中图分类号:Q-1;G250.2 **文献标志码:**A **文章编号:**1672-5565(2020)03-186-09

Research situation analysis of domestic and international bioinformatics based on co-word analysis

LIU Yawen, ZHANG Hongyan*, YANG Lingyan

(School of Information and Intelligence Science and Technology, Hunan Agricultural University, Changsha 410128, China)

Abstract: Bioinformatics is an emerging interdisciplinary subject in the field of natural science, whose research situation has attracted the attention of many scholars. In order to understand the research situation of bioinformatics at home and abroad, this paper takes the journal articles in the field of bioinformatics in CNKI Database and Web of Science Database as research objects. Based on bibliometrics method and co-word analysis method, the research status, hotspots and trends of bioinformatics at home and abroad are summarized by R programming tool. Results show that the bioinformatics research at home and abroad is both in the stage of high-speed development. The amount of literature is increasing year by year, and the research area is expanding. The research hotspots at home and abroad both focus on gene mining, protein structure and function prediction, and miRNA analysis, etc. However, domestic research tends to be theoretical, while international research pays more attention to the application in disease treatment, drug design and so on.

Keywords: Bioinformatics; Co-word analysis; R language; Research situation

生物信息学(Bioinformatics)是一门由生命科学、计算机科学、数学、物理学和化学等学科交叉而形成的前沿学科,其研究内容涵盖了大规模的基因组、转录组、蛋白质组等多个组学层次的各自及整合性研究,应用领域也覆盖了生物、医学、药学、农学、环境监测等多个领域^[1]。生物信息学发展史可分为3个主要阶段^[2]:(1)前基因组时代(50~80年代)是基本生物信息学理论、方法、模型和软件体系

形成阶段;(2)基因组时代(80年代末~2003年),人类基因组计划的实施进一步推动生物信息学的发展;(3)后基因组时代(2003年-至今)的核心内容之一是充分利用大量生物数据挖掘其生物信息和规律。作为一门新兴学科,生物信息学领域的研究文献逐年增加,这些文献客观记录了该领域的发展概况,开展相关文献分析以了解生物信息学领域的研究现状和研究热点,对把握其未来的发展趋势意义

收稿日期:2019-11-11;修回日期:2019-11-28.

基金项目:湖南省教育厅科学研究重点项目(No.18A105);长沙市工业科技特派员项目(No.201845).

作者简介:刘亚文,女,硕士研究生,研究方向:生物信息学. E-mail:1357818573@qq.com.

*通信作者:张红燕,女,博士,教授,研究方向:生物信息学. E-mail:hongyan_zhang@hunau.edu.cn.

重大。目前,已有不少学者开展了文献分析在生物信息学领域中的应用研究,如宋茂海等基于 CNKI、中华医学会数据库采用共词分析和可视化方法探讨了生物信息学领域的学科分类和热点内容^[3];王俊等基于大数据背景,从迅速发展、主要研究模块、基础背景等方面分析了生物信息学的研究现状^[4];王蕊等基于 Web of science 数据库分析了生物信息学数据库研究的发展趋势^[5]。但利用文献计量分析和关键词共词分析,同时基于 CNKI 中文数据库和 Web of science 外文数据库来对比分析国内外生物信息学领域研究的相关报道却极其鲜见。

当两个或两个以上的关键词同时出现在同一篇文献中时,则称之为存在共现关系^[6]。这种共现关系的紧密程度体现在关键词共现的频次上。关键词共词分析通过捕捉关键词关系,分析它们所代表的学科和主题的结构变化,来跟踪该研究领域的研究热点,揭示其研究主题的演变过程^[7]。本文以国内外生物信息学领域相关论文为研究对象,利用 R 语言编程工具,对文献进行计量分析、关键词共词聚类分析,并以图文展示了国内和国际生物信息领域的研究热点与发展趋势,以期为我国生物信息学领域科研人员的后续研究提供参考。

1 数据与方法

1.1 数据来源

采用的国内文献来源于 CNKI 数据库,以“生物信息学”为主题词进行检索,选择全部期刊文献,共检索到 1996~2018 年共 23 年相关期刊论文 6 893 篇;国际文献来源于 Web of science 数据库,以“bioinformatics”为主题词进行检索,选择核心合集论文,共检索到 2007~2018 年共 12 年相关 SCI 收录论文 32 406 篇。数据均于 2019 年 1 月采集。

1.2 研究方法

1.2.1 数据预处理

数据预处理主要通过利用 R 语言编程环境 (RStudio Version 3.5.1) 编写 R 程序对原始数据进行数据清洗,数据清洗包括删除作者为空的文献记录、删除关键字为空的文献记录等操作,清洗处理后最终得到中文文献 6 717 条记录,外文文献 30 160 条记录。

1.2.2 高频关键词提取

预处理后的数据包含多个字段,CNKI 文献数据中包含的字段有 Title(标题)、Author(作者)、Keyword(关键词)、Summary(摘要)、Year(年份)等,Web of science 文献数据中包含的字段有 AU(作者)、TI(标题)、PY(出版年)、AB(摘要)、DE

(作者添加的关键词)、ID(Web of science 根据文章内容增加的关键词)等。无论是中文文献还是外文文献,本文用于处理分析的主要是年份字段和关键词字段。其中,国际文献的关键词字段有作者关键词“DE”和 WOS 增加的关键词“ID”,考虑到“ID”更具客观性,故选用“ID”关键词字段为后续分析所用。

年份字段主要应用于文献的计量分析和热点变迁的分段研究,关键词字段则主要用于研究热点和研究主题的提取。通常一个关键词字段会包含多个以“,”或“;”分隔开的关键词,我们先采用 R 语言扩展包 tidytext 中的 unnest_tokens 函数来提取关键词;接着将其中的同义关键词合并,如:关键词“mirna”和“microrna”合并为“mirna”;最后,统计关键词词频,并按频次由高到低排序,得到关键词词频表,并从中选取频次较高的前 n 个关键词作为高频关键词。

1.2.3 共词分析

将高频关键词两两配对,构造高频关键词共现矩阵,矩阵中的元素为两两关键词的共现次数。进一步,以共现矩阵为基础,构建关键词共现网络。具体而言,共现网络以 n 个高频关键词为网络节点,当两个关键词之间的共现次数大于等于 k 次时,两个关键词节点间存在连线。

在共词分析中, n 和 k 的设置并无统一标准。高频关键词选取过多或是共现次数设置过小将导致所得研究热点过于宽泛,而高频关键词选取过少或是共现次数设置过大将导致所得研究热点缺乏代表性。综合考虑文献数量和关键词词频情况,最终设置 $n=40$ 、 $k=3$,即选取前 40 个高频关键词用于构建共现矩阵,共现次数大于等于 3 次设置网络连线。

1.2.4 聚类分析

基于关键词共现网络利用 R 中的 cluster_louvain 函数创建可视化聚类图,并对其展开聚类分析。cluster_louvain 函数采用的是 Louvain 聚类算法,即通过计算相邻节点间的模块度增量,实现节点间的动态聚合^[8]。Louvain 算法快速、准确,被公认为性能最好的社区发现算法之一。节点中心度是社会网络分析的量化指标之一,中心度高的节点通常是较为关键的研究热点^[9]。特征向量中心度更是把相邻节点的中心度也考虑在内^[10]。通过分析聚类网络图和特征向量中心度,找出每类的核心关键词及主要特征,进而挖掘出生物信息学领域的研究热点及发展趋势。

2 国内生物信息学领域研究态势分析

2.1 国内文献量年变化趋势分析

对 1996~2018 年 CNKI 数据库中生物信息学相

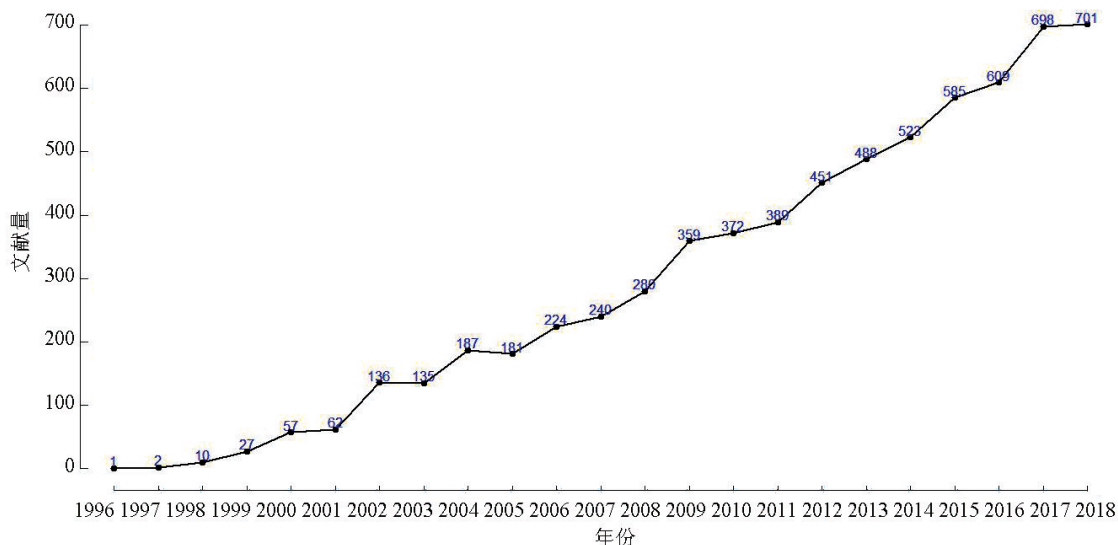


图 1 1996~2018 国内生物信息学文献量变化趋势图

Fig.1 Trend of domestic bioinformatics literature quantity between 1996 and 2018

2.2 国内研究热点分析

关键词是某一研究主题的提炼和对文章核心思想的高度概括^[11]。通过对关键词频次和中心度的统计分析,得到 40 个高频关键词的相关结果(见表 1)。其中,频次排名靠前的关键词有“克隆”、“基因克隆”、“序列分析”、“基因”、“数据库”、“电子克隆”等,各关键词前后的频次差距不大,体现了生物信息学领域研究热点的多方向发展。

进一步共词分析和聚类分析,得到国内生物信息学文献关键词聚类图(见图 2),图中关键词节点的颜色不同表示所属类别不同,40 个关键词被聚成了 4 类。整体分析时所得表 1 中中心度较高的关键词有克隆、基因表达、基因克隆、序列分析等也基本上是各类别中的核心关键词。

依据聚类结果及中心度情况,结合具体文献和专业知

识,可将国内生物信息学领域研究热点归纳如下:
(1)基因克隆及生物信息学分析研究 这一类的高频词有克隆、序列分析、基因克隆、表达分析、race、原核表达、rt-pcr 等。传统的基因克隆方法利用限制性内切酶和连接酶在体外连接的方法构建重组载体,步骤繁琐且受限于酶切位点,随着测序技术的发展,一些简化组装流程的 DNA 克隆及组装新技术竞相发展^[12]。生物信息学分析包括了序列分析、基因分析和蛋白分析等多种方法。利用基因克隆技

术结合生物信息学分析研究以达到对基因功能、结构预测的目的是国内生物信息学领域的研究热点之一。

术结合生物信息学分析研究以达到对基因功能、结构预测的目的是国内生物信息学领域的研究热点之一。

(2)生物信息学数据库与基因挖掘研究分析 这一类中包含的高频关键词有数据库、基因表达、人类基因组计划、数据挖掘、计算生物学、序列比对等。生物信息学的数据库主要存储在数据库中,而数据挖掘则从生物信息学数据库中调用、提取数据^[13]。生物信息学数据库的主要特点有:数据库种类的多样性、数据库数据量增长惊人、逐渐形成综合性生物信息学数据库平台、数据库高度网络化等^[14]。如何利用生物信息学数据库并采用合适的数据挖掘算法进行有效的基因挖掘研究分析是国内生物信息学领域的研究热点之一。

(3)miRNA 研究分析 这一类中包含的高频关键词有生物信息学、mirna、启动子、靶基因、转录因子、基因芯片等。miRNA 也就是 MicroRNA,是基因表达的负调控因子,主要通过 RNA 干扰(RNAi)途径进行调控^[15]。随着生物信息学技术的发展,新 miRNA 的发现及其相关调控机制的研究分析已成为国内生物信息学领域的研究热点之一。

(4)电子克隆分析 这一类的高频词有基因家族、水稻、拟南芥、玉米、电子克隆、表达序列标签。电子克隆技术是近年来发展起来的基因克隆新方

法,其应用是基于数据库、计算机网络和应用软件这三部分生物信息学资源而展开的^[16]。电子克隆技

术是发现新基因和确定基因功能的重要手段,也是国内生物信息学领域的研究热点之一。

表1 1996~2018 生物信息学中文文献高频关键词表

Table 1 High frequency keywords in Chinese literature of bioinformatics between 1996 and 2018

序号	关键词	频次	中心度	序号	关键词	频次	中心度
1	生物信息学	5 083	1.00	21	计算生物学	68	0.24
2	克隆	497	0.57	22	玉米	67	0.25
3	基因克隆	382	0.41	23	序列比对	63	0.16
4	序列分析	235	0.41	24	人类基因组计划	62	0.19
5	基因	161	0.27	25	差异表达基因	56	0.13
6	数据库	161	0.32	26	大豆	56	0.28
7	电子克隆	135	0.26	27	race	51	0.32
8	水稻	134	0.28	28	结构	51	0.12
9	基因表达	131	0.46	29	数据挖掘	51	0.20
10	基因芯片	131	0.21	30	猪	51	0.32
11	蛋白质组学	122	0.21	31	二级结构	50	0.12
12	原核表达	114	0.25	32	基因表达谱	50	0.14
13	转录因子	112	0.33	33	亚细胞定位	50	0.16
14	mirna	107	0.19	34	表达	49	0.16
15	启动子	102	0.31	35	rt-pcr	48	0.21
16	靶基因	95	0.18	36	抗原表位	47	0.10
17	基因组	92	0.32	37	拟南芥	47	0.21
18	表达分析	85	0.41	38	表达序列标签	46	0.13
19	基因组学	74	0.21	39	小麦	44	0.32
20	基因家族	70	0.26	40	cdna	43	0.16

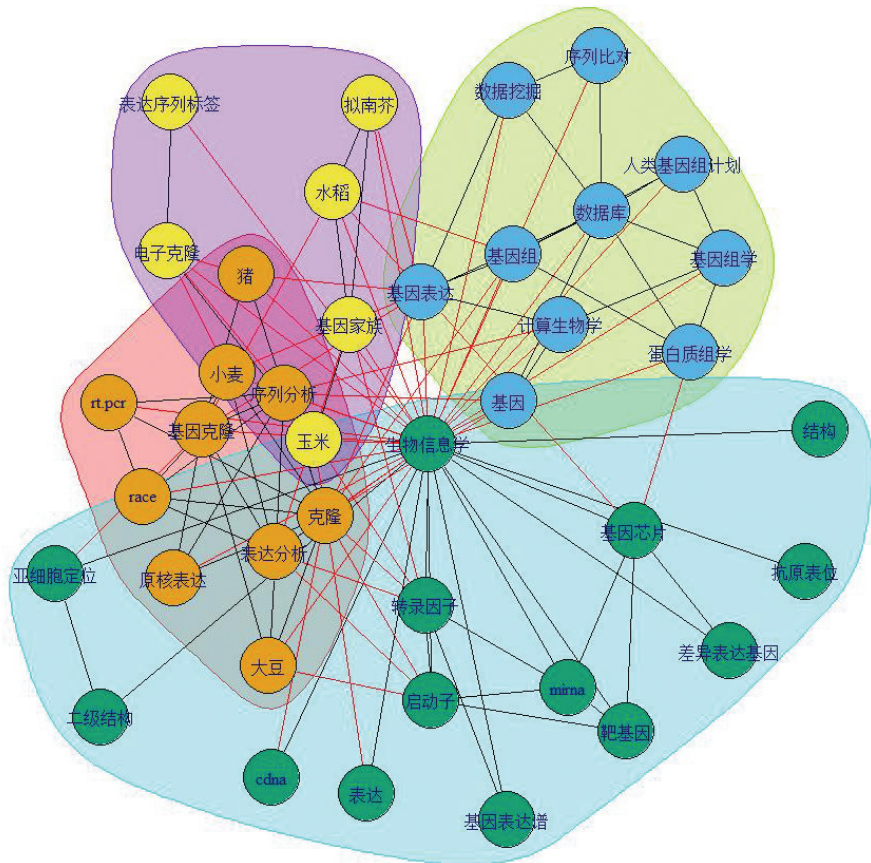


图2 1996~2018 国内生物信息学文献关键词聚类图

Fig.2 Keywords cluster of domestic bioinformatics literature between 1996 and 2018

2.3 国内研究趋势分析

前文对国内生物信息学领域研究热点做了整体分析,此处对国内研究热点随时间变迁作进一步分析。以40个高频关键词为行,年份为列,分年度统计每个关键词出现的频数,构建国内文献年份-关键词统计表(见表2)。

为了便于国内外的对比分析,将1996~2006年的国内文献高频关键词作为整体分析,结果表明国内早期生物信息学的研究热点主要集中在基因数据库研究,其代表高频关键词有“数据库”、“人类基因

组计划”等;对2007~2018年的文献,通过分析各高频关键词的频数相对密集年份等特征可知:2007~2010年更侧重于蛋白质数据库研究以及本体和结构域研究,代表关键词有“蛋白质组学”、“数据挖掘”、“结构”等;2011~2013年主要是电子克隆研究,代表关键词有“电子克隆”、“水稻”、“序列分析”等;2014~2018年该领域研究热点更加多样,主要在miRNA分析、基因和蛋白质结构和功能预测分析等,代表关键词有“亚细胞定位”、“基因芯片”、“miRNA”、“功能”、“计算生物学”等。

表2 国内文献年份-关键词统计表(部分)

Table 2 Year-keyword statistics of domestic literature (partial)

关键词	...	2007年	2008年	2009年	2010年	2011年	...	2018年
生物信息学		184	211	258	273	284		551
克隆		10	19	34	23	36		49
基因克隆		9	11	21	18	21		51
序列分析		10	11	20	19	27		10
基因		6	4	3	6	5		14
数据库		9	6	7	11	3		4
电子克隆		4	10	11	9	11		6
水稻		5	4	9	6	11		12
基因表达		4	3	3	3	9		20
基因芯片		6	4	4	1	9		15
.....								
功能		3	2	2	4	2		5

3 国际生物信息学领域研究态势分析

3.1 国际文献年变化趋势分析

2007~2018年SCI数据库中生物信息学相关文

献量的年份变化趋势(见图3)。从图中可以看出文献量逐年增长,速度较快。总的来说,国际文献的年份统计结果显示了生物信息学领域的研究正处于发展中。

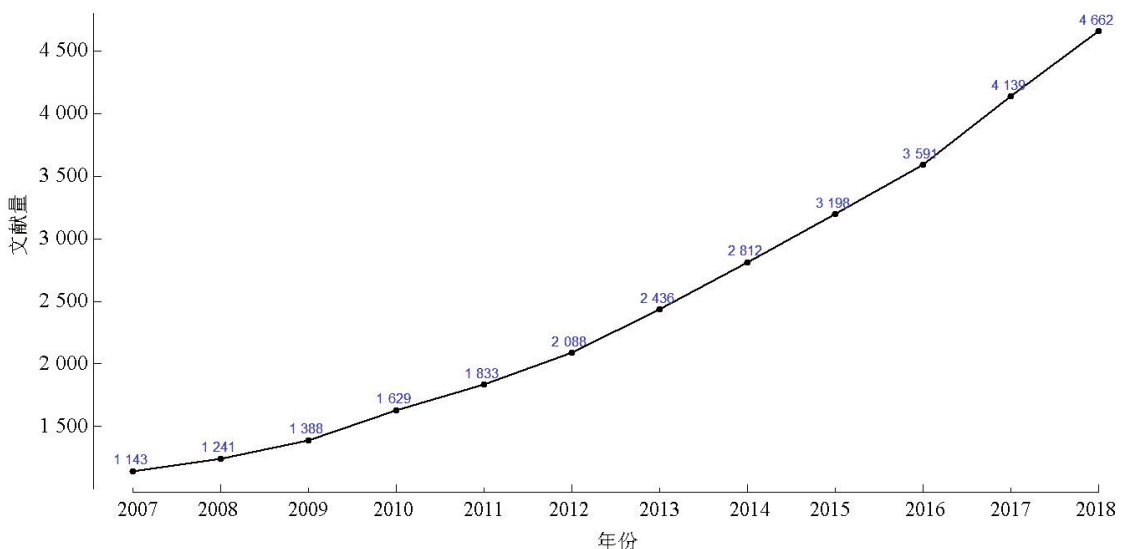


图3 2007~2018国际生物信息学文献量变化趋势图

Fig.3 Trend of international bioinformatics literature quantity between 2007 and 2018

3.2 国际研究热点分析

通过对国际文献关键词频次的统计分析,得到40个高频关键词(见表3)。从表3中不难发现,关键词 expression、identification、protein、gene、gene-

expression 的频次较高。而主题检索词“bioinformatics”并不是频次最高的关键词,这一现象也从侧面说明了选取 WOS 的关键词字段“ID”进行共词分析是合理的。

表3 2007-2018 生物信息学国际文献高频关键词表

Table 3 High frequency keywords in international literature of bioinformatics between 2007 and 2018

序号	关键词	频次	中心度	序号	关键词	频次	中心度
1	expression	4 369	0.97	21	metastasis	767	0.89
2	identification	3 098	1.00	22	classification	720	0.85
3	protein	2 496	0.95	23	mass-spectrometry	720	0.96
4	gene	2 282	1.00	24	micornas	711	0.85
5	gene-expression	1 994	1.00	25	pathway	694	0.92
6	sequence	1 645	0.93	26	tool	689	0.93
7	database	1 541	0.97	27	discovery	661	0.81
8	cancer	1 540	0.98	28	carcinoma	653	0.98
9	cells	1 348	1.00	29	in-vitro	634	0.90
10	prediction	1 305	0.98	30	arabidopsis-thaliana	594	0.91
11	bioinformatics	1 275	0.90	31	differentiation	594	0.74
12	disease	1 109	0.96	32	mutations	594	0.94
13	breast-cancer	1 033	0.93	33	model	591	0.91
14	escherichia-coli	959	0.76	34	progression	578	1.00
15	genome	949	1.00	35	crystal-structure	576	0.84
16	growth	939	0.92	36	dna	542	0.73
17	activation	933	1.00	37	invasion	533	0.88
18	evolution	898	0.95	38	algorithm	532	0.77
19	proliferation	878	0.91	39	in-vivo	530	0.69
20	apoptosis	783	0.97	40	biology	504	

同上,通过共词分析和聚类分析,得到的国际生物信息学文献关键词聚类图,40个关键词被聚成了2类(见图4)。

同上,依据聚类结果及中心度情况,并结合具体文献和专业知识,可将国际生物信息学领域研究热点归纳如下:

(1)miRNA 及与之相关的分子间相互作用研究

这一类包含的高频关键词有 cells(细胞)、proliferation(增殖)、apoptosis(细胞凋亡)、micornas、pathway(神经通路)等。miRNA 及其产物 mRNAs 形成复杂的调控网络,参与细胞的生长、分化、增殖、死亡等一系列重要的生命过程,基因表达谱微阵列的生物信息学分析已被广泛用于识别各种疾病中的重要分子机制和生物标志物^[17]。利用生物信息学分析方法挖掘 miRNA 功能并得到与之相关的分子间相互作用成为了国际生物信息学领域的研究热点之一。

(2)基因、蛋白质等功能预测与应用研究 这一类包含的高频关键词有 identification(识别)、

expression(表达)、protein(蛋白质)、database(数据库)、prediction(预测)、sequence(序列)。众所周知,生物实验的实施大都费用昂贵且费时,而生物信息学的产生就是为了攻破这些传统的生物学难题。例如,蛋白质组学领域的一个重要研究方向就是利用有效的算法对 DNA 结合蛋白进行鉴定与预测^[18]。因此,利用数据库和有效的计算预测方法在基因、蛋白质等功能预测与应用方面的研究是国际生物信息学领域的研究热点之一。

3.3 国际研究趋势分析

分年度统计40个国际文献高频关键词出现的频次,构建国际文献年份-关键词统计表(见表4)。并参照国内文献划分时间段来进行研究趋势分析,即分为2007~2010年、2011~2013年、2014~2018年三个时间段。结果表明,2007~2010年该领域在国际的研究热点侧重于数据库分析以及实验工具、算法的改进;2011~2013年的研究热点集中在蛋白质分析及应用;2014~2018年的研究热点较为偏向于 miRNA 分析及应用。

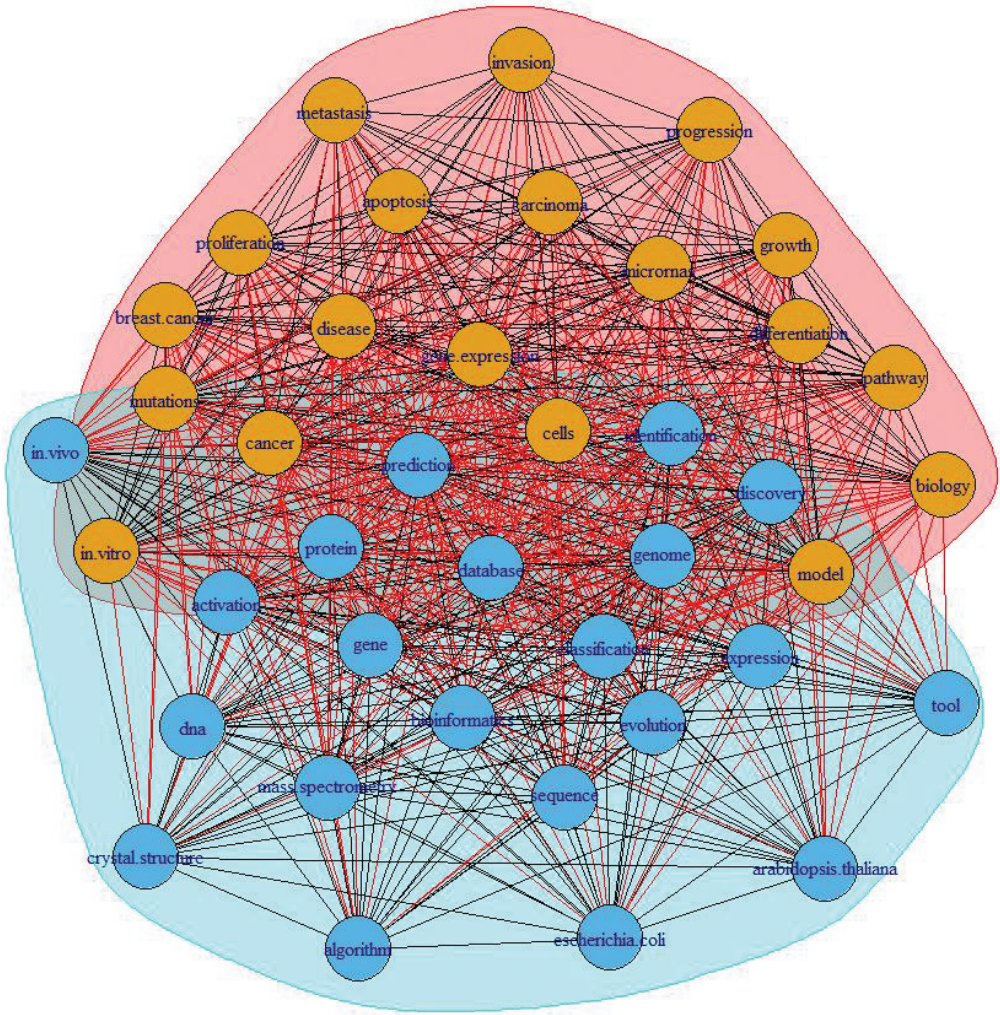


图 4 2007~2018 国际生物信息学文献关键词聚类图

Fig.4 Keywords cluster of international bioinformatics literature between 2007 and 2018

表 4 国际文献年份-关键词统计表(部分)

Table 4 Year-keyword statistics of international literature (partial)

Keywords	2007 年	2008 年	2009 年	2010 年	2011 年	2018 年
expression	91	108	112	160	180		1018
identification	149	154	143	205	190		414
protein	87	96	109	139	159		354
gene	79	78	112	121	139		381
gene-expression	66	84	84	137	131		267
cancer	51	47	41	51	70		323
cells	29	29	39	59	59		309
sequence	122	105	102	106	112		165
disease	19	27	28	42	61		230
proliferation	6	5	4	11	14		312
.....
biology	21	31	40	34	44		52

4 讨论与总结

以国内外生物信息学领域期刊论文为研究对象,从年文献量变化趋势、研究热点、分阶段研究趋势三个方面进行分析,得出结论如下:

1) 国内外年文献量变化趋势分析 国内外生物信息学研究均处于高速发展期,文献量均呈逐年增长趋势,不过国内文献的增长速度小于国际文献的增长速度。

2) 研究热点分析 国内的研究热点有基因克隆及生物信息学分析研究、生物信息学数据库与基因挖掘研究分析、miRNA 研究分析、电子克隆分析等;国际的研究热点有基因、蛋白质等功能预测与应用研究和 miRNA 及与之相关的分子间相互作用研究等。国内外研究热点均聚焦在基因挖掘、蛋白质结构与功能预测、miRNA 分析等,但国内的生物信息学领域研究更侧重于理论研究,而国际生物信息学领域研究则更加注重其在疾病(癌症)治疗、药物设计等方面的实际应用。

3) 分阶段研究趋势分析 国内生物信息学领域研究热点在不同阶段内略有波动,呈现了较为明显的热点变迁状况;而国际生物信息学领域研究热点在不同阶段内的变化很小,高频关键词的频次几乎都是逐年增加,而且关键词之间的联系较强,是较为均衡的发展。整体来看,国内外生物信息学的研究范围均在不断拓宽。

参考文献(References)

[1] 严文颖. 生物信息专业 R 语言课程教学的探索和研究[J]. 现代信息科技, 2018, 2(11): 108-109+112.
YAN Wenying. Exploration and research of R language teaching for bioinformatics[J]. Modern Information Technology, 2018, 2(11): 108-109+112.

[2] 胡昇. 基于生物信息技术筛选慢性乙型肝炎血液相关基因的研究[D]. 广州: 南方医科大学, 2013.
HU Sheng. Study on screening chronic hepatitis B related blood genes by bioinformatics[D]. Guangzhou: Southern Medical University, 2013.

[3] 宋茂海, 李东方. 基于共词分析的国内生物信息学热点领域研究[J]. 生物信息学, 2014, 12(1): 46-52. DOI: 10.3969/j.issn.1672-5565.2014.01.08.
SONG Maohai, LI Dongfang. Hot spots analysis of China's bioinformatics based on co-word analysis method[J]. Chinese Journal of Bioinformatics, 2014, 12(1): 46-52. DOI: 10.3969/j.issn.1672-5565.2014.01.08.

[4] 王俊, 郭丽, 吴建盛, 等. 大数据背景下的生物信息学研

究现状[J]. 南京邮电大学学报(自然科学版), 2017, 37(4): 62-67. DOI: 10.14132/j.cnki.1673-5439.2017.04.011.
WANG Jun, GUO Li, WU Jiansheng, et al. Status of bioinformatics research in big data[J]. Journal of Nanjing University of Posts and Telecommunications (Natural Science Edition), 2017, 37(4): 62-67. DOI: 10.14132/j.cnki.1673-5439.2017.04.011.

[5] 王蕊, 胡德华. 生物信息学数据库研究文献引文与热点分析[J]. 生物信息学, 2014, 12(4): 305-312. DOI: 10.3969/j.issn.1672-5565.2014.04.12.
WANG Rui, HU Dehua. Analysis of citation features in bioinformatics database literature studies[J]. Chinese Journal of Bioinformatics, 2014, 12(4): 305-312. DOI: 10.3969/j.issn.1672-5565.2014.04.12.

[6] 袁润, 李莹, 王琦, 等. 用 R 语言分析关键词集共现网络研究[J]. 现代情报, 2018, 38(7): 88-94. DOI: 10.3969/j.issn.1008-0821.2018.07.013.
YUAN Run, LI Ying, WANG Qi, et al. Research on co-occurrence network of keywords set by R[J]. Journal of Modern Information, 2018, 38(7): 88-94. DOI: 10.3969/j.issn.1008-0821.2018.07.013.

[7] RADHAKRISHNAN S, ERBIS S, ISAACS J A, et al. Correction: Novel keyword co-occurrence network-based methods to foster systematic reviews of scientific literature[J]. PLoS One, 2017, 12(9): e0185771. DOI: 10.1371/journal.pone.0172778.

[8] 林定, 徐颖, 黄国新, 等. 基于 Louvain 算法的图数据三维树形可视化[J]. 计算机工程与应用, 2018, 54(7): 96-101. DOI: 10.3778/j.issn.1002-8331.1702-0346.
LIN Ding, XU Ying, HUANG Guoxin, et al. Visualizing graph data in 3D tree-style based on Louvain algorithm.[J]. Computer Engineering and Applications, 2018, 54(7): 96-101. DOI: 10.3778/j.issn.1002-8331.1702-0346.

[9] 郑俊巍. 我国管理创新研究结构与发展的文献计量与网络图景[J]. 科研管理, 2018, 39(9): 102-112. DOI: 10.19571/j.cnki.1000-2995.2018.09.012.
ZHENG Junwei. Bibliometrics and network mapping of China's management innovation research structure and development[J]. Science Research Management, 2018, 39(9): 102-112. DOI: 10.19571/j.cnki.1000-2995.2018.09.012.

[10] 许海云, 方曙, 付鑫金. 基于特征向量中心度加权的期刊影响因子研究[J]. 情报理论与实践, 2011, 34(11): 108-112. DOI: 10.16353/j.cnki.1000-7490.2011.11.024.
XU Haiyun, FANG Shu, FU Xinjin. Research on the influencing factors of journals weighted based on eigenvector centrality[J]. Information Studies: Theory & Application, 2011, 34(11): 108-112. DOI: 10.16353/j.cnki.1000-7490.2011.11.024.

[11] 赵蓉英, 余波. 中外近五年智库研究演进脉络与热点主题对比分析[J]. 情报科学, 2018, 36(11): 3-9+16. DOI:

- 10.13833/j.issn.1007-7634.2018.11.001.
- ZHAO Rongying, YU Bo. Comparative research on the evolution venues and hot topics of think tanks in recent five years in china and abroad[J]. Information Science, 2018, 36(11):3-9+16. DOI:10.13833/j.issn.1007-7634.2018.11.001.
- [12] 盛晓菁, 齐晓雪, 徐蕾, 等. 基因克隆及组装技术的研究进展[J/OL]. 中国生物工程杂志:1-11[2019-09-18]. <http://kns.cnki.net/kcms/detail/11.4816.Q.20190807.1030.002.html>.
- SHENG Xiaojing, QI Xiaoxue, XU Lei, et al. The research progress of gene cloning and assembly[J/OL]. China Biotechnology:1-11[2019-09-18]. <http://kns.cnki.net/kcms/detail/11.4816.Q.20190807.1030.002.html>.
- [13] 吴金华, 张艳秋, 唐毅. 数据挖掘在生物信息学中的应用——文献计量学视角[J]. 生物信息学, 2016, 14(4):249-253. DOI:10.3969/j.issn.1672-5565.2016.04.09.
- WU Jinhua, ZHANG Yanqiu, TANG Yi. Application of data mining in bioinformatics—bibliometrics perspective[J]. Chinese Journal of Bioinformatics, 2016, 14(4):249-253. DOI:10.3969/j.issn.1672-5565.2016.04.09.
- [14] 张桂荣. 生物信息数据库资源及其应用[J]. 河北科技师范学院学报(社会科学版), 2010, 9(1):121-124.
- ZHANG Guirong. Bio-information database resources and their application[J]. Journal of Hebei Normal University of Science & Technology(Social Science), 2010, 9(1):121-124.
- [15] 赵孟良, 田闵玉, 李莉, 等. MicroRNA 参与调控植物抗逆的现状 & 展望[J/OL]. 基因组学与应用生物学:1-7[2019-07-19]. <http://kns.cnki.net/kcms/detail/45.1369.Q.20190701.1202.002.html>.
- ZHAO Mengliang, TIAN Minyu, LI Li, et al. Present situation and prospect of microRNA involved regulation of plant stress resistance[J/OL]. Genomics and Applied Biology:1-7[2019-07-19]. <http://kns.cnki.net/kcms/detail/45.1369.Q.20190701.1202.002.html>.
- [16] 王冬冬, 朱延明, 李勇, 等. 电子克隆技术及其在植物基因工程中的应用[J]. 东北农业大学学报, 2006, 6(3):403-408.
- WANG Dongdong, ZHU Yanming, LI Yong, et al. Application of in silico cloning technique in plant gene engineering[J]. Journal of Northeast Agricultural University, 2006, 6(3):403-408.
- [17] QU Yinmeng, SUN Xin, YAN Xiuli, et al. Identification of microRNAs and messenger RNAs involved in human umbilical cord mesenchymal stem cell treatment of ischemic cerebral infarction using integrated bioinformatics analysis[J]. Neural Regeneration Research, 2019, 14(9):1610-1616. DOI:10.4103/1673-5374.255998.
- [18] XU Ruifeng, ZHOU Jiyun, WANG Hongpeng, et al. Identifying DNA-binding proteins by combining support vector machine and PSSM distance transformation[J]. BMC Systems Biology, 2015, 9(S1). DOI:10.1186/1752-0509-9-S1-S10.

[责任编辑:吴永英]