Statistics and Application 统计学与应用, 2023, 12(5), 1440-1450
Published Online October 2023 in Hans. https://www.hanspub.org/journal/sa
https://doi.org/10.12677/sa.2023.125148

基于Citation Topic和CiteSpace的长三角地区研究热点的可视化分析

蒋海灵

浙江理工大学图书馆，浙江 杭州

收稿日期：2023年9月26日；录用日期：2023年10月24日；发布日期：2023年10月31日

摘 要

【目的/意义】对长三角地区的研究热点进行可视化分析，助力区域创新研究。【方法/过程】以Web of Science核心合集数据库中收录的近5年(2018~2022年)长三角地区的文献为数据源，首先，借助InCites数据库的Citation Topic功能，对微观主题的发文量、总被引频次、高被引论文和学科规范化的引文影响力等指标进行对比分析；其次，以热门微观主题冠状病毒为例，采用CiteSpace软件对该微观主题文献进行机构合作、关键词共现和关键词聚类的科学知识图谱分析。【结果/结论】超级电容器、氧还原反应、光催化作用和冠状病毒是近5年长三角地区的热门研究领域；新型冠状肺炎疫情下人群的心理健康、新型冠状肺炎疫情传播模拟和预测模型、新型冠状病毒致病机理、单克隆抗体药物的研发、临床特征、药物筛选和疫苗研发等是冠状病毒研究主题的研究热点。【创新/局限】提供了一种结合Citation Topic和CiteSpace工具分析和挖掘区域研究主题与热点的新路径。

关键词

引文主题，CiteSpace，科学知识图谱，研究热点，长三角地区

Visual Analysis of Research Hotspots in Yangtze River Delta Based on Citation Topic and CiteSpace

Hailing Jiang

Library of Zhejiang Sci-Tech University, Hangzhou Zhejiang

Received: Sep. 26th, 2023; accepted: Oct. 24th, 2023; published: Oct. 31st, 2023

Abstract

[Purpose/significance] Visual analysis of research hotspots in the Yangtze River Delta region will

文章引用: 蒋海灵. 基于 Citation Topic 和 CiteSpace 的长三角地区研究热点的可视化分析[J]. 统计学与应用, 2023, 12(5): 1440-1450. DOI: 10.12677/sa.2023.125148

support regional innovation research. [Method/process] Taking journal articles published in Yangtze River Delta retrieved in Web of Science Core Collection in the recent 5 years (2018-2022) as the data source, first, based on the Citation Topics of InCites, a comparative analysis of indicators of micro-topics, such as the number of papers, total cited times, highly cited papers and the category normalized citation impact, etc. is conducted; second, taking the micro-topic of coronavirus as an example, visual analysis of co-institution, keyword co-occurrence and clustering is made by CiteSpace knowledge mapping. [Result/conclusion] Supercapacitors, oxygen reduction reactions, photocatalysis, and coronavirus are the hot research topics in the Yangtze River Delta in the recent 5 year. Effect of COVID-19 on mental health of individuals, spread simulation and prediction models, pathogenesis, monoclonal antibody, clinical characteristics, drug screening and vaccine development of COVID-19, etc. are the hot research topics of the micro-topic of coronavirus. [Innovation/limitation] This paper provides a new path to analyze and mine regional research topics and hotspots by combining Citation Topic and CiteSpace tools.

Keywords

Citation Topic, CiteSpace, Scientific Knowledge Map, Research Hotspots, Yangtze River Delta

Copyright © 2023 by author(s) and Hans Publishers Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/
Open Access

1. 引言

2020 年 12 月 20 日，科技部颁布了《长三角科技创新共同体建设发展规划》，提出联合提升原始创新能力，强化核心技术协同攻关，提高重大创新策源能力，推动长三角地区成为以科技创新驱动高质量发展的强劲动力源。江苏省、浙江省、安徽省和上海市作为长三角地区的三省一市，深入全面地研究当地的科研情况能为政府制定相关科技创新政策提供决策支持，而科技创新政策是合理分配创新资源、提升区域创新能力的重要保障[1]。因此，全面分析和挖掘长三角地区的研究热点显得尤为重要：长三角地区的主要优势领域有哪些？近几年兴新的热门研究领域有哪些？热门研究领域的主要研究人员和机构有哪些？

科睿唯安于 2020 年 12 月初推出了 InCites Benchmarking & Analytics™ Citation Topics (引文主题)，开发并实施了一种基于引用的分类算法，对将近 7000 万篇文献构建文章级别的主题分类体系，可以对研究人员、组织、国家/地区和资助机构的产出进行更精细的分析[2]。目前关于 Citation Topics 在期刊编辑制定选题策划方面的研究已见报道[3]，而在区域研究热点的分析和挖掘方面尚未见报道。

关键词共现网络分析已成为描述学科研究现状并预测未来发展方向的重要手段，在挖掘大规模研究隐藏的深层内涵上具有独特的优势[4]。CiteSpace 是美国德雷赛尔大学计算与信息学院陈超美博士开发的一种重要的科学知识图谱工具，能够绘制作者、机构、国家合作图谱和关键词共现网络图谱。近年来，利用 CiteSpace 对某一特定研究领域的人员、机构、研究热点进行可视化分析的报道比较多[5]，而针对区域研究热点分析的文献非常少。现有的研究通常以某一区域的文献为数据源，采用可视化软件直接对该区域所有的文献进行分析，如陈晓玲等人采用 CiteSpace 软件对东北三省区域 2012~2016 年共 25,752 篇文献进行可视化分析[6]，这种方法受到文献数量的限制，对于发文量较高区域的分析难度较大。为此本文借助 InCites 数据库的 Citation Topic 功能对近 5 年(2018~2022 年)长三角地区的文献进行主题分类，先将庞大的数据源文献分类成许多不同微观主题，解决了因文献数量过大导致可视化软件无法分析的问

题，再结合 CiteSpace 软件对各微观主题的合作机构和研究热点等进行深入的分析和探究，以揭示长三角地区机构的合作情况以及研究热点，有助于政府全面地了解长三角地区的科研情况，以期为当地政府的创新决策提供支持，同时也为学者探测其他区域的研究热点提供一种全新的、有效的分析方法。

2. 数据来源与研究方法

在数据来源上，通过在 Incites 中设置学科分类体系：Citation Topics，时间范围：2018~2022，国家/地区：Shanghai，China Mainland、Jiangsu，China Mainland、Zhejiang，China Mainland、Anhui，China Mainland，InCites 数据集更新日期为 2023 年 5 月 26 日。在检索结果中选择微观主题级别“Micro”，文献类型“Article”进行精炼，共检索到 894,810 篇文献，以此作为分析数据源。

在研究方法上，本研究借助 InCites 数据库的 Citation Topic 功能对文献进行主题分类，再将热门研究主题文献导出到 Web of Science(WOS)数据库中，并导出为纯文本格式，题录格式为全记录与引用的参考文献格式，作为 CiteSpace 分析的数据源，并结合 CiteSpace6.1.R6 软件进行机构合作、关键词共现和关键词聚类的可视化分析。

Table 1. Micro topics of WOS papers published in the Yangtze River Delta region in recent 5 years
表 1. 近 5 年长三角地区发表的 WOS 论文微观主题

排名	名称	WOS 论文数	被引频次	学科规范化的引文影响力	引文影响力	国际合作百分比/%	期刊规范化的引文影响力	高被引论文
1	Supercapacitor (超级电容器)	10,732	273,836	1.16	25.52	23.64	1.32	397
2	Oxygen Reduction Reaction (氧还原反应)	8736	271,624	1.32	31.09	28.71	1.36	584
3	Photocatalysis (光催化作用)	8944	231,943	1.30	25.93	18.9	1.45	483
4	Coronavirus (冠状病毒)	3445	214,799	1.46	62.35	32.54	3.13	373
5	lncRNA (lncRNA)	7677	138,742	0.96	18.07	6.85	1.58	100
6	Perovskite Solar Cells (钙钛矿太阳能电池)	5349	133,499	1.16	24.96	26.57	1.14	198
7	Metal-Organic Frameworks (金属-有机框架物)	5117	114,779	1.28	22.43	23.39	1.38	211
8	Lithium-Sulfur Batteries (锂硫电池)	4241	111,919	1.23	26.39	26.9	1.16	161
9	Magnetic Nanoparticles (磁性纳米粒子)	2956	88,250	1.64	29.85	17.29	1.34	131
10	Environmental Kuznets Curve (环境库兹涅茨曲线)	3900	86,437	1.29	22.16	34.92	2.08	344
11	MoS_2 (二硫化钼纳米纤维)	4535	81,728	1.17	18.02	30.89	1.11	95
12	Cross-Coupling (交叉偶联)	5030	81,626	1.09	16.23	12.78	1.21	32
13	Aptamer (识体)	4239	80,751	1.17	19.05	13.23	1.16	68
14	PD-1 (PD-1)	4139	76,837	0.84	18.56	16.67	1.35	127
15	Deep Learning (深度学习)	6356	76,099	0.92	11.97	28.82	1.26	152

3. 结果与分析

3.1. 长三角地区微观研究主题分析

本研究筛选的 894,810 篇文献分布在 2482 个微观主题上。按照总被引频次排序，选取 TOP15 微观主题分布如表 1。结合表 1 的各项评价指标可以看出，Supercapacitor (超级电容器)、Oxygen Reduction Reaction (氧还原反应)和 Photocatalysis (光催化作用)这三个研究主题发表的论文数均超过 8000 篇，总被引频次均超过 2 万次，高被引论文的篇数也较高，分别为 397 篇、584 篇和 483 篇，学科规范化的引文影响力高于国际水平，表明这三个微观主题是近 5 年长三角地区最热门的研究主题。值得特别关注的是，冠状病毒主题虽然发文量只有 3445 篇，但总被引频次高达 214,799 次，篇均被引频次 62，位列第一。综合考虑冠状病毒发文量 3445 篇和学科规范化的引文影响力 1.46，表明长三角地区学者在冠状病毒研究主题上发表的论文水平远远高于国际平均水平。此外，Perovskite Solar Cells (钙钛矿太阳能电池)、Metal-Organic Frameworks (金属－有机框架物)、Lithium-Sulfur Batteries (锂硫电池)、Magnetic Nanoparticles (磁性纳米粒子)、Environmental Kuznets Curve (环境库兹涅茨曲线)篇均被引频次都在 20 次以上，高被引论文在 100 篇以上，说明这些也是比较热门的研究主题。lncRNA 和 Deep Learning (深度学习)发文量为 6000~8000 篇，但是学科规范化的引文影响力在 1 以下，表明该微观主题论文的研究水平处于全球平均水平以下，有待进一步提高。

3.2. 发文量和被引频次

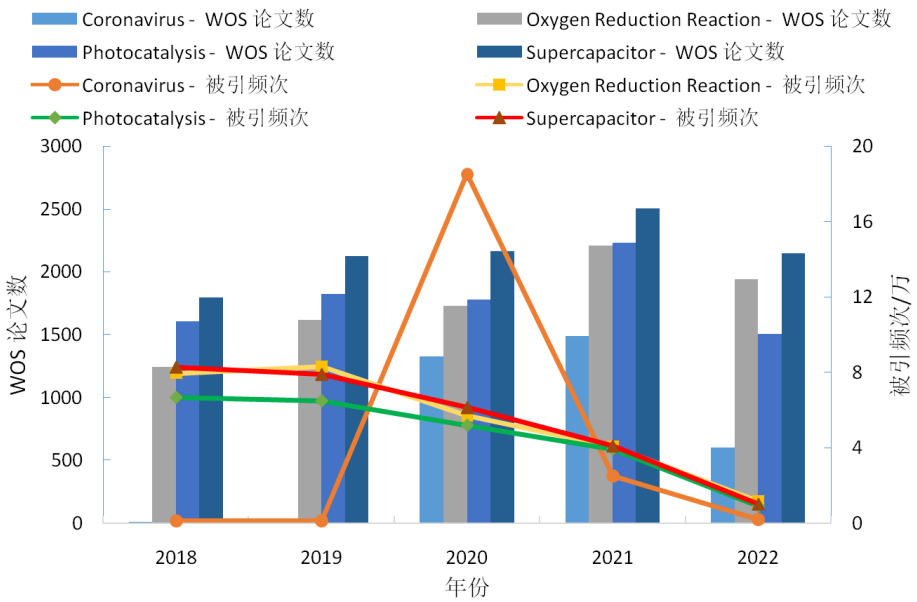

Figure 1. Annual distribution of the number of papers and cited times of hot research topics in the Yangtze River Delta region in recent 5 years

图 1. 近 5 年长三角地区热门研究主题发文量和被引频次年度分布情况

近 5 年来长三角地区被引频次均超过 2 万次以上的热门微观主题有四个，分别是超级电容器、氧还原反应、光催化作用和冠状病毒，其发文量和被引频次年度分布情况如图 1 所示。从发文量来看，超级电容器、氧还原反应和光催化作用三个研究主题发文量在 2018~2021 年间持续稳定增长，到 2022 年略有下降。冠状病毒主题发文量从 2019 年后陡然上升，表明 2019 年底新型冠状肺炎疫情(COVID-19，简称新冠肺炎)爆发后，冠状病毒研究主题迅速得到学者的关注，2020~2021 年发文量稳步上升，直到 2022

年才有所下降。从被引频次来看，超级电容器和氧还原反应研究主题的年度被引频次走势比较一致，光催化作用研究主题的被引频次比前两者稍低。冠状病毒研究主题的被引频次从 2019 年的 1206 次跃升到 2020 年的 185,243 次，在 2020 年超过了前三个热门研究主题，表明冠状病毒是近 5 年来全球新兴的研究前沿，也是近 5 年学者关注最多的研究主题，值得深入分析。

3.3. 冠状病毒微观主题研究结果分析

新冠肺炎疫情从 2019 年底开始肆虐全球，对全球的公共卫生安全带来严重的威胁，同时也是健康中国建设面临的重大挑战，因此进一步分析与之相关的冠状病毒微观主题变得十分有意义。为此，本文将 InCites 中冠状病毒微观主题的文献共 3445 篇导出到 WOS 数据库中，作为进一步分析的数据源，先对该主题的高被引论文进行分析，再用 CiteSpace 软件对该主题进行机构合作、关键词共现和关键词聚类等可视化分析。

Table 2. Top 10 highly cited papers on coronavirus research topics in the Yangtze River Delta region in recent 5 years
表 2. 近 5 年长三角地区冠状病毒研究主题的高被引文献列表 TOP 10

序号	作者	标题	被引次数	发表期刊	发表年份
1	Chen, Nanshan; Zhou, Min; Zhang, Xinxin; Zhang, Li, *et al.*	Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study	7700	LANCET	2020
2	Wu, Fan; Zhao, Su; Zhang, Yong-Zhen, *et al.*	A new coronavirus associated with human respiratory disease in China	5182	NATURE	2020
3	Wu, Chaomin; Chen, Xiaoyan; Song, Yuanlin, *et al.*	Risk Factors Associated With Acute Respiratory Distress Syndrome and Death in Patients With Coronavirus Disease 2019 Pneumonia in Wuhan, China	4631	JAMA INTERNAL MEDICINE	2020
4	Lai, Jianbo; Ma, Simeng; Hu, Shaohua, *et al.*	Factors Associated With Mental Health Outcomes Among Health Care Workers Exposed to Coronavirus Disease 2019	3765	JAMA NETWORK OPEN	2020
5	Wang, Cuiyan; Pan, Riyu; Ho, Roger C. , *et al.*	Immediate Psychological Responses and Associated Factors during the Initial Stage of the 2019 Coronavirus Disease (COVID-19) Epidemic among the General Population in China	3374	INTERNATIONAL JOURNAL OF ENVIRONMENTAL RESEARCH AND PUBLIC HEALTH	2020
6	Yan, Renhong; Zhang, Yuanyuan; Zhou, Qiang, *et al.*	Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2	2971	SCIENCE	2020
7	Jin, Zhenming; Jiang, Hualiang; Rao, Zihe; Yang, Haitao, *et al.*	Structure of M-pro from SARS-CoV-2 and discovery of its inhibitors	2133	NATURE	2020
8	Fang, Yicheng; Zhang, Huangqi; Ji, Wenbin, *et al.*	Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR	1849	RADIOLOGY	2020
9	Xu, Xiao-Wei; Wu, Xiao-Xin; Li, Lan-Juan, *et al.*	Clinical findings in a group of patients infected with the 2019 novel coronavirus (SARS-Cov-2) outside of Wuhan, China: retrospective case series	1799	BMJ-BRITISH MEDICAL JOURNAL	2020
10	Wang, Qihui; Zhang, Yanfang; Yan, Jinghua; Qi, Jianxun, *et al.*	Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2	1759	CELL	2020

3.3.1. 高被引论文分析

在 3445 篇文献中高被引论文共 373 篇，按被引频次降序排序，列出第一作者或者通讯作者为长三角地区的研究机构且排名前 10 位的高被引论文见表 2。由表 2 可知，被引次数排名第一的为上海交通大学医学院张欣欣教授团队发表的文章，排名第二和第三的为复旦大学生物医学研究院张永振教授科研团队和复旦大学附属中山医院宋元林教授科研团队发表的文章，前 5 篇文章的被引频次在 3000 次以上，其他几篇文献的被引频次都在 1700 次以上；在国际顶级期刊《科学》(Science)、《自然》(Nature)、《细胞》(Cell)和《柳叶刀》(Lancet)上均有发文，其中 Science 刊文 1 篇、Nature 刊文 2 篇、Cell 刊文 1 篇，Lancet 刊文 1 篇。研究主题涉及新冠病毒的流行病学、临床和生物学特征、新冠肺炎疫情对人们心理健康的影响、新冠病毒细胞表面受体血管紧张素转换酶 2 (ACE2)、新冠病毒关键药物靶点、胸部 CT 在新冠肺炎诊断中的作用等。上述高被引论文的作者张欣欣、张永振、李兰娟、宋元林、周强和杨海涛等也为全球科技战疫作出重要贡献。

3.3.2. 机构合作分析

通过 CiteSpace 的机构合作网络的可视化分析，节点类型选择 institution，其余参数选择默认值，得到长三角地区冠状病毒研究主题的机构合作网络图谱如图 2 所示。图谱显示共有节点 434 个，连线 2099 条，网络密度 0.0223。表 3 为发文量大于 100 的研究机构，排在第一位的是复旦大学，发文量为 543，由其构成的主要合作群体包括上海交通大学、中国科学院上海生命科学研究院、华中科技大学、中国科学院大学，南方科技大学和纽约血液中心等；排名第二的是上海交通大学，发文量为 417，由其构成的主要合作群体包括复旦大学、华中科技大学、武汉大学、香港中文大学、武汉金银潭医院、中山大学和中国医学科学院&北京协和医学院等；排名第三的是浙江大学，发文量为 400，由其构成的主要合作群体包括中山大学、广东省疾病预防控制中心、中国人民大学、温州医科大学和树人大学等。排名第四的是

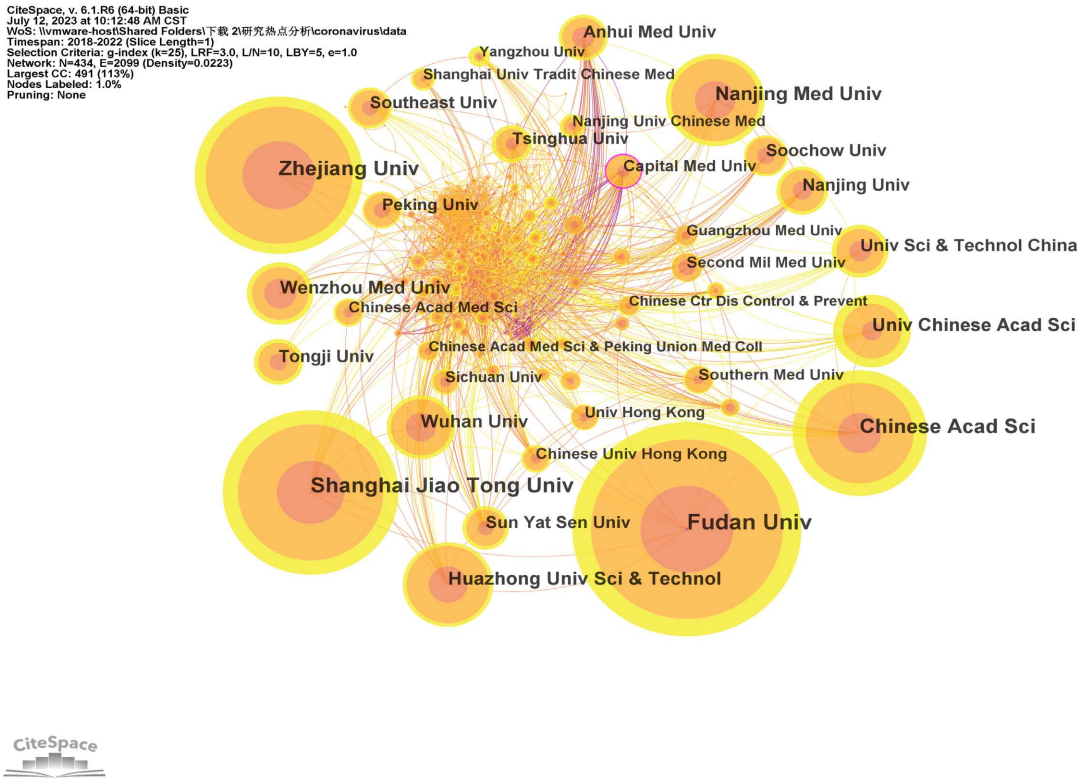

Figure 2. Knowledge map of co-institutional analysis
图 2. 机构合作分析知识图谱

中国科学院，由其构成的主要合作群体包括复旦大学、中国科学院大学、中国科学技术大学、清华大学、上海交通大学、安徽大学、南京医科大学和西湖大学等。

由此看出，长三角地区的核心研究力量主要集中在“双一流”建设高校，如复旦大学、上海交通大学和浙江大学等。由于这些高校有较多附属医院，如复旦大学附属中山医院和华山医院等，上海交通大学附属瑞金医院和仁济医院等，浙江大学附属第一医院和第二医院等，技术实力雄厚，教学、医疗和科研水平较高，科研成果较为丰硕。其次，中国科学院(包括上海生命科学研究院、上海药物研究所和武汉病毒研究所等)、南京医科大学、温州医科大学、中国科学技术大学、南京大学、安徽医科大学和同济大学发文量都在 100 以上，贡献也较大。此外，从图 2 可以看出，长三角地区的机构合作较为密切，尤以同一省份研究机构之间合作紧密，并与中部和东部的机构，如华中科技大学、武汉大学、武汉病毒研究所和中山大学等合作较为紧密，但与西部和东北部地区的机构合作较少。此外，国际合作也较少。因此，长三角地区的研究机构应当进一步打破地域界限，加大跨省之间和国际机构合作力度，共同促进长三角地区的进一步发展。

Table 3. Analysis of co-institution by frequency
表 3. 机构合作频次分布表

频次	中心性	年份	机构
543	0.01	2018	Fudan Univ.
417	0.05	2019	Shanghai Jiao Tong Univ.
400	0.01	2018	Zhejiang Univ.
320	0.04	2018	Chinese Acad. Sci.
236	0.01	2020	Nanjing Med Univ.
158	0.01	2018	Wenzhou Med Univ.
135	0.01	2020	Univ. Sci. & Technol. China
124	0.05	2020	Nanjing Univ.
120	0.08	2018	Anhui Med Univ.
118	0	2020	Tongji Univ.

3.3.3. 关键词共现网络分析

Table 4. Analysis of keyword co-occurrence by frequency
表 4. 关键词共现频次表

序号	频次	中介性	关键词	年份	序号	频次	中介性	关键词	年份
1	461	0	covid 19	2020	11	126	0.03	clinical characteristics	2019
2	443	0.03	coronavirus	2018	12	116	0.04	mental health	2020
3	225	0.01	pneumonia	2020	13	112	0.05	china	2018
4	221	0.04	outbreak	2020	14	111	0.09	virus	2018
5	208	0.18	spike protein	2018	15	109	0.01	impact	2020
6	199	0.1	acute respiratory syndrome	2018	16	105	0.01	ace2	2020
7	189	0.15	sars	2019	17	96	0.06	antibody	2018
8	189	0.03	wuhan	2020	18	94	0.03	transmission	2020
9	149	0.01	infection	2020	19	82	0.09	disease	2018
10	146	0.02	risk	2020	20	67	0.02	epidemic	2020

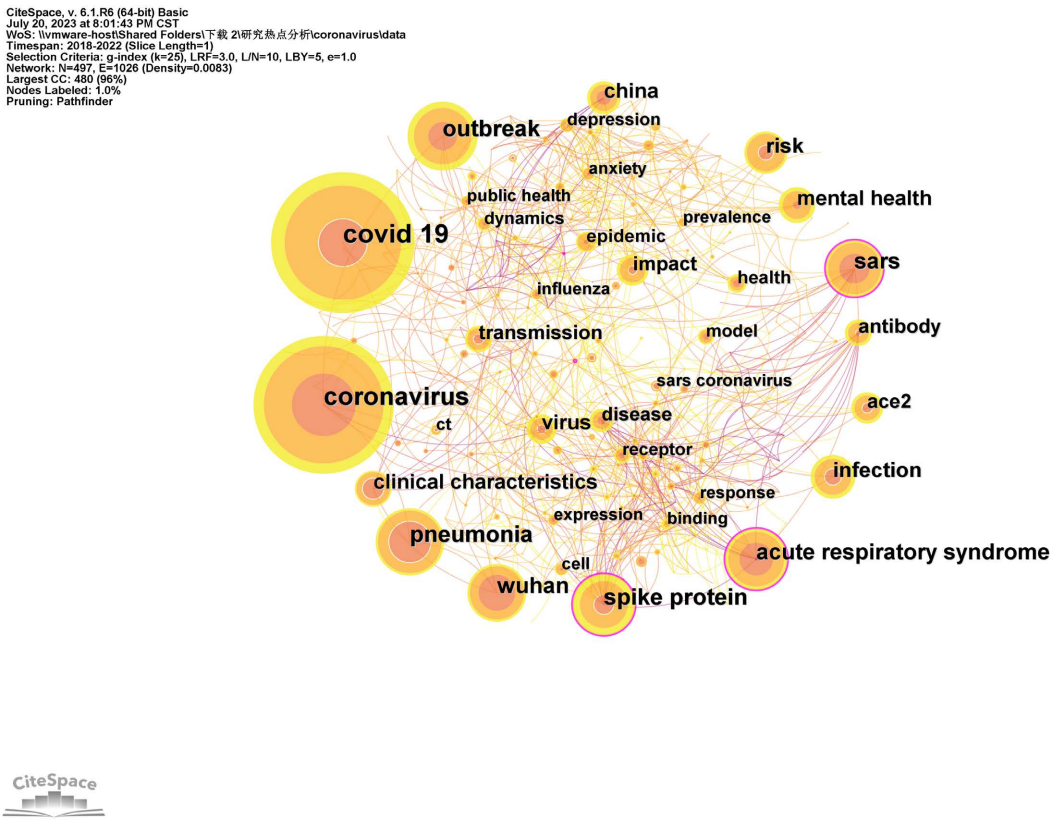

Figure 3. Knowledge map of keyword co-occurrence analysis
图 3. 关键词共现分析知识图谱

关键词是一篇文章的核心与精髓，是文章主题的高度概括及凝炼，频次高的关键词常被用来确定一个研究领域的热点问题[7]。为了挖掘长三角地区冠状病毒主题中的研究热点，本文用 CiteSpace 对 2018~2022 年的文献进行关键词共现分析，节点类型选择 keyword，寻径网络算法选择 pathfinder，网络裁剪策略选择 Pruning sliced networks，其余参数选择默认值。利用 add to the alias list 功能将同义词进行合并，如将 covid 19、coronavirus disease 2019、novel coronavirus pneumonia、coronavirus disease 2019 (covid-19)和 SARS-CoV-2 进行同义词合并，合并为 covid 19，将 risk 和 risk factor 合并为 risk factor 等。最后得到关键词共现分析网络知识图谱，如图 3 所示。将图 3 中出现频次排名前 20 的关键词按照频次高低排序如表 4 所示。根据表 4 可知，节点“covid 19”在图谱中出现的频次最高为 461 次，这表明近 5 年的冠状病毒研究主题是围绕新冠肺炎展开的。此外，共现频次较高的关键词还有 coronavirus (443)，pneumonia (225)，outbreak (221)，spike protein (208)，acute respiratory syndrome (199)，sars (189)，wuhan (189)，infection (149)，risk (146)，clinical characteristics (126)，mental health (116)，china (112)，virus (111)，impact (109)，ace2 (105)，antibody (96)，transmission (94)，disease (82)，epidemic (67)等。其中，pneumonia (肺炎)、outbreak (爆发)、wuhan (武汉)和 China (中国)出现频次较高，表明学者们研究的对象以新冠肺炎疫情爆发初期的中国湖北省武汉市的病例居多；而 spike protein (刺突蛋白)、clinical characteristics (临床特征)、mental health (心理健康)、ace2 (受体血管紧张素转化酶 2)、antibody (抗体)、transmission (传播)则表明研究的主要内容集中在新冠病毒的刺突蛋白的结构、临床特征、细胞表面受体血管紧张素转换酶 2、抗体的研发、传播特性的研究，以及对人们心理健康的影响等。这与高被引论文的研究主题相一致。

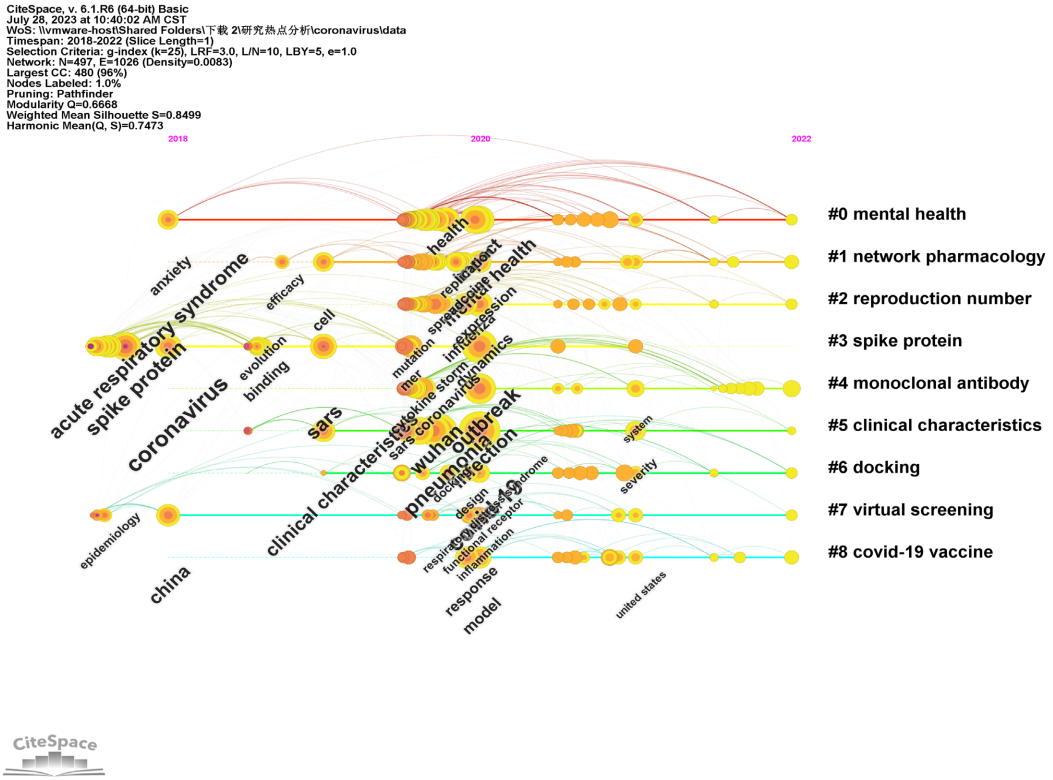

Figure 4. Timeline map of keyword clustering
图 4. 关键词聚类时间线图谱

为深入分析研究热点，在关键词共词分析的基础上，对关键词知识图谱进行聚类分析，生成关键词聚类图谱，并选择时间线图(Timeline View)，以进一步展现长三角地区冠状病毒研究主题中的研究热点。关键词聚类图谱分析结果如图 4 所示，一共生成 16 个聚类，选取前 9 个聚类进行展示和分析。每个聚类的关键词按照出现的时间由远到近在时间轴上由左到右铺展开，聚类大小从大到小垂直排列。我们将该领域的研究热点大致分为以下几个方面：

聚类#0 mental health (心理健康)是最大的聚类，涉及新冠肺炎疫情爆发对普通人群和医护人员心理健康的影响，如 Lai 等主要通过健康问卷、广义焦虑症量表、失眠严重程度指数和抑郁、焦虑等症状程度进行多变量 logistic 回归分析分析，确定了影响医护人员心理健康的相关因素[8]；Wang 等通过使用滚雪球抽样方法对疫情初期普通人群在线调查，并采用事件影响量表修订版评估心理影响，采用抑郁、焦虑和压力量表评估心理健康状况[9]。聚类#2 reproduction number (再生数)主要涉及新冠肺炎疫情传播和预测模型的研究。基本再生数 R_0 和实时再生数 R_t 是评估疫情传染强度的重要指标。R_0 表示一个感染者在感染期内平均传染的人数，而实时再生数 R_t 是指 t 时刻开始一个感染者平均传染的人数。因此，通过流行病动力学时序模型推算 R_t 动态变化，有助于掌握疫情走势，分析疫情防控政策的效果[11]，如 Zhao 等通过时间序列方法发现新冠肺炎疫情早期 R_0 估值在 2.24~3.58 之间[11]；Yeoh 等研究了非药物干预措施对 7 个世界卫生组织辖区的实时再生数的影响[12]。而疫情预测模型的研究则主要通过传染病动力学模型(SI、SIS、SIR、SEIR 等)结合冠状病毒的传播特性，建立传染病数学模型[13]，来预测疫情发展趋势。聚类#3 spike protein (刺突蛋白)主要涉及新冠病毒致病机理的研究，包括病毒刺突蛋白的结构和功能的研究[14] [15]。由于新冠病毒是利用其表面刺突蛋白受体结合域(RBD)结合宿主细胞上的 ACE2 侵入人体细胞。因此，新冠病毒刺突蛋白的结构与功能的研究对进一步揭示病毒致病机理、新冠靶向药物和疫苗的研发

具有极为重大的意义。聚类#4 monoclonal antibody (单克隆抗体)主要涉及单克隆抗体药物的研发，如 Chi 等发现了首个靶向刺突蛋白 N 端结构域的高效中和单克隆抗体[16]；Zhang 等报道了托珠单抗作为一种针对细胞因子白细胞介素-6 (IL-6)受体的单克隆抗体，可有效阻断 IL-6 信号转导通路，有望成为重症新冠肺炎患者的有效治疗药物[17]。聚类#5 clinical characteristics (临床特征)的主要内容涉及临床症状、影像学特征、实验室检查、预后转归等方面[18] [19]，主要研究发现有：常见症状有发热和咳嗽；其他症状如肌肉疼痛、精神错乱、头痛、喉咙痛、腹泻等较少见；入院时胸部计算机断层扫描以磨玻璃影最常见，部分非重症患者和少部分重症患者未发现影像学异常。聚类#1 network pharmacology (网络药理学)、#6 docking (对接)和#7 virtual screening，涉及新冠肺炎药物筛选，主要研究基于网络药理学和分子对接虚拟筛选抗新冠肺炎病毒药物的活性成分、潜在作用靶点及作用机制[20] [21]。聚类#8 covid-19 vaccine (新冠病毒疫苗)，涉及新冠病毒疫苗的研发，如 Zhu 等对重组腺病毒 5 型载体 COVID-19 疫苗进行了单中心、开放标签、非随机、剂量递增的 1 期试验[22]；Yang 等研究发现了 ACE2 的编号为 319-545 的一段氨基酸序列通过重组蛋白方式制备新冠病毒疫苗[23]。

4. 结论

本文以 WOS 核心合集数据库中收录的 2018~2022 年长三角地区的文献为研究对象，先借助 InCites 数据库的 Citation Topic 功能，对热门研究主题进行分析，再结合 CiteSpace 软件对冠状病毒热门研究主题文献进行机构合作、关键词共现和关键词聚类的科学知识图谱分析，可以得出以下结论：

按照总被引频次排序的 Top15 微观研究主题中，被引频次均超过 2 万次以上的微观主题有超级电容器、氧还原反应、光催化作用和冠状病毒，是近 5 年长三角地区的热门研究主题。从发文量来看，超级电容器、氧还原反应和光催化作用三个研究主题发文量在 2018~2021 年间持续稳定增长。冠状病毒主题发文量从 2019 年新冠肺炎疫情爆发后陡然上升，其被引频次从 2019 年的 1206 次跃升到 2020 年的 185,243 次，表明是近几年兴新研究前沿。

对冠状病毒研究主题的高被引论文分析表明，长三角地区的学者在国际顶级期刊上均有发文，表明长三角地区学者的研究水平较高，得到国际同行的认可。高被引论文的作者张欣欣、张永振和李兰娟等也为全球科技战疫作出了重要贡献。对机构合作分析发现长三角地区的研究力量主要集中在以复旦大学、上海交通大学和浙江大学等为核心的“双一流”建设高校，同省合作以及与中部和东部的研究机构合作较为紧密，但国际合作力度不够。

通过关键词共现分析和关键词聚类，得到长三角地区冠状病毒研究主题的研究热点。涉及的主要聚类包括：新冠肺炎疫情爆发对普通人群和医护人员心理健康的影响、新冠肺炎疫情传播和预测模型的研究、新冠病毒致病机理的研究、单克隆抗体药物的研发、新冠病毒的流行病学、临床和生物学特征、新冠病毒药物筛选和新冠病毒疫苗的研发等。

因此，政府在制定科技政策时需要进一步加大对本区域热门研究领域的投入，巩固已有研究成果；同时，针对薄弱的研究领域，培育和扶持一批新的产业；建立对外交流与合作机制，促进科技人才国际交流合作，加强与国际科研机构的合作力度，从而推动长三角地区的科技创新和高质量发展。

基金项目

浙江省教育厅项目“基于科学知识图谱的学科前沿识别研究”(21176120-F)。

参考文献

[1] 贾永飞, 王振宇, 张欣, 等. 科技创新政策对区域创新能力的影响机理研究——基于 fsQCA 方法的组态分析[J].

山东财经大学学报, 2023, 35(2): 87-96+121.

[2] Introducing Citation Topics in InCites. https://clarivate.com/blog/introducing-citation-topics/

[3] 王燕, 姚蔚, 杜敏, 等. 利用 InCites 新功能 Citation Topics 助力学术期刊编辑制定选题策划方案——以园艺学科研究领域为例[J]. 中国科技期刊研究, 2021, 32(6): 777-785.

[4] 田沛霖, 赵蓉英, 常茹茹, 等. 信息计量学研究的知识结构与发展态势[J]. 情报科学, 2022, 40(7): 186-193.

[5] 刘迪, 张会来. 我国高校网络舆情的研究热点和前沿探析[J]. 情报科学, 2022, 40(11): 176-185.

[6] 陈晓玲, 刘东亮. 基于科学知识图谱的东北三省区域研究热点分析[J]. 情报学报, 2018, 37(12): 1224-1231.

[7] 张立, 尹萍. 我国图情领域近十年国际发文研究热点和前沿剖析[J]. 情报科学, 2020, 38(5): 80-87.

[8] Lai, J.B., Ma, S.M., Wang, Y., *et al.* (2020) Factors Associated With Mental Health Outcomes among Health Care Workers Exposed to Coronavirus Disease 2019. *Jama Network Open*, **3**, e203976. https://doi.org/10.1001/jamanetworkopen.2020.3976

[9] Wang, C.Y., Pan, R.Y., Wan, X.Y., *et al.* (2020) Immediate Psychological Responses and Associated Factors during the Initial Stage of the 2019 Coronavirus Disease (COVID-19) Epidemic among the General Population in China. *International Journal of Environmental Research and Public Health*, **17**, 1729-1753. https://doi.org/10.3390/ijerph17051729

[10] 裴韬, 王席, 宋辞, 等. COVID-19 疫情时空分析与建模研究进展[J]. 地球信息科学学报, 2021, 23(2): 188-210.

[11] Zhao, S., Lin, Q.Y., Ran, J.J., *et al.* (2020) Preliminary Estimation of the Basic Reproduction Number of Novel Coronavirus (2019-nCoV) in China, from 2019 to 2020: A Data-Driven Analysis in the Early Phase of the Outbreak. *International Journal of Infectious Diseases*, **92**, 214-217. https://doi.org/10.1016/j.ijid.2020.01.050

[12] Yeoh, E.K., Chong, K.C., Chiew, C.J., *et al.* (2021) Assessing the Impact of Non-Pharmaceutical Interventions on the Transmissibility and Severity of COVID-19 during the First Five Months in the Western Pacific Region. *One Health*, **12**, Article ID: 100213. https://doi.org/10.1016/j.onehlt.2021.100213

[13] Wang, H.W., Wang, Z.Z., Dong, Y.Q., *et al.* (2020) Phase-Adjusted Estimation of the Number of Coronavirus Disease 2019 cases in Wuhan, China. *Cell Discovery*, **6**, Article No. 10. https://doi.org/10.1038/s41421-020-0148-0

[14] Yao, H.P., Song, Y.T., Chen, Y., *et al.* (2020) Molecular Architecture of the SARS-CoV-2 Virus. *Cell*, **183**, 730-738. https://doi.org/10.1016/j.cell.2020.09.018

[15] Wang, Q.H., Zhang, Y.F., Wu, L.L., *et al.* (2020) Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell*, **181**, 894-904. https://doi.org/10.1016/j.cell.2020.03.045

[16] Chi, X.Y., Yan, R.H., Zhang, J., *et al.* (2020) A Neutralizing Human Antibody Binds to the N-Terminal Domain of the Spike Protein of SARS-CoV-2. *Science*, **369**, 650-655. https://doi.org/10.1126/science.abc6952

[17] Zhang, C., Wu, Z., Li, J.W., *et al.* (2020) Cytokine Release Syndrome in Severe COVID-19: Interleukin-6 Receptor Antagonist Tocilizumab May Be the Key to Reduce Mortality. *International Journal of Antimicrobial Agents*, **55**, Article ID: 105954. https://doi.org/10.1016/j.ijantimicag.2020.105954

[18] Guan, W., Ni, Z., Hu, Y., *et al.* (2020) Clinical Characteristics of Coronavirus Disease 2019 in China. *New England Journal of Medicine*, **382**, 1708-1720. https://doi.org/10.1056/NEJMoa2002032

[19] Chen, N.S., Zhou, M., Dong, X., *et al.* (2020) Epidemiological and Clinical Characteristics of 99 Cases of 2019 Novel Coronavirus Pneumonia in Wuhan, China: A Descriptive Study. *Lancet*, **395**, 507-513. https://doi.org/10.1016/S0140-6736(20)30211-7

[20] Zhang, D.H., Wu, K.L., Zhang, X., *et al.* (2020) *In silico* Screening of Chinese Herbal Medicines with the Potential to Directly Inhibit 2019 Novel Coronavirus. *Journal of Integrative Medicine-Jim*, **18**, 152-158. https://doi.org/10.1016/j.joim.2020.02.005

[21] Jin, Z.M., Du, X.Y., Xu, Y.C., *et al.* (2020) Structure of M-Pro from SARS-CoV-2 and Discovery of Its Inhibitors. *Nature*, **582**, 289-293. https://doi.org/10.1038/s41586-020-2223-y

[22] Zhu, F.C., Li, Y.H., Guan, X.H., *et al.* (2020) Safety, Tolerability, and Immunogenicity of a Recombinant Adenovirus Type-5 Vectored COVID-19 Vaccine: A Dose-Escalation, Open-Label, Non-Randomised, First-in-Human Trial. *Lancet*, **395**, 1845-1854. https://doi.org/10.1016/S0140-6736(20)31208-3

[23] Yang, J., Wang, W., Chen, Z., *et al.* (2020) A Vaccine Targeting the RBD of the S Protein of SARS-CoV-2 Induces Protective Immunity. *Nature*, **586**, 572-577. https://doi.org/10.1038/s41586-020-2599-8