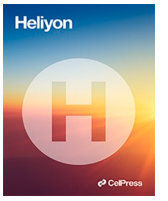
Heliyon 10 (2024) e33219

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

A bibliometric study of the intellectual base and global research hotspots for single-cell sequencing [2009–2022] in breast cancer

Shan Liu [a,1], Xudong Li [a,1], Ying Zhang [c], Yuhan Deng [a], Zehao Li [b], Yunan Zhu [a], Xue Li [a], Yuefeng Shang [a], Guang Yang [a], Xiaolu Zhan [a,***], Yingpu Li [a,**], He Ren [a,*]

[a] *Department of Breast Surgery, Harbin Medical University Cancer Hospital, Harbin, China*
[b] *Jiamusi University School of Clinical Medicine, Jiamusi, China*
[c] *Department of Neurology, Air Force Medical Center, PLA, Beijing, China*

ARTICLE INFO

Keywords:
Single-cell sequencing
Bibliometric analysis
Breast cancer
Immunotherapy
Research hotspots

ABSTRACT

Background: Breast cancer is the most widespread malignant tumor worldwide. Single-cell sequencing technology offers novel insights and methods to understand the onset, progression, and treatment of tumors. Nevertheless, there is currently an absence of a thorough and unbiased report on the comprehensive research status of single-cell sequencing in breast cancer. This study seeks to summarize and quantify the dynamics and trends of research on breast cancer single-cell sequencing by bibliometric analysis.

Methods: Research articles and reviews related to breast cancer single-cell sequencing were selected from the WoSCC database. Visualization of data regarding countries, institutions, authors, references, and keywords was performed by CiteSpace and VOSviewer software.

Results: 583 articles and reviews were analyzed in this study. The quantity of publications related to breast cancer single-cell sequencing has been increasing annually. These studies originate from 302 institutions in 46 countries, with YMAX S WICHA producing the most publications and WANG Y being the most cited author. *Nature Communications* is the most researched journal, while *Nature* holds the highest number of citations. These journals predominantly cover topics in the molecular/biological/immunological fields. Moreover, an analysis of reference and keyword bursts revealed that current research trends in this area are primarily centered on "clonal evolution," "tumor microenvironment," and "immunotherapy."

Conclusion: Breast cancer single-cell sequencing is a rapidly growing area of scientific interest. Future research requires more frequent and in-depth collaborations among countries, institutions, and authors. Furthermore, "clonal evolution," "tumor microenvironment," and "immunotherapy" are likely to become major focal points in upcoming research on breast cancer single-cell sequencing.

* Corresponding author.
** Corresponding author.
*** Corresponding author.
E-mail addresses: xlzhan@hrbmu.edu.cn (X. Zhan), 724224887@qq.com (Y. Li), renhe@hrbmu.edu.cn (H. Ren).
[1] These authors have contributed equally to this work.

https://doi.org/10.1016/j.heliyon.2024.e33219
Received 12 April 2023; Received in revised form 16 June 2024; Accepted 17 June 2024
Available online 18 June 2024

1. Introduction

Breast cancer is the most prevalent malignancy among women globally. In 2020, it overtook lung cancer in the global count of newly diagnosed cases [1]. In 2020, over 685,000 people deaths were attributed to the disease, accounting for 30 % of all new cases of breast cancer [1]. In China, the incidence rate of breast cancer has been climbing at a rate of 3.9 % per year [2,3]. Although the combination of surgery, chemotherapy, radiotherapy, and molecular targeted therapy can effectively improve the mortality rate of breast cancer, tumor heterogeneity, drug resistance, and metastasis still hinder the clinical efficacy of treatment. Therefore, it is necessary to understand the pathogenesis, prevention, and treatment of breast cancer is crucial, necessitating further in-depth research.

With the continuous innovation in sequencing technology in the past few decades, from first-generation to third-generation sequencing, an enormous quantity of genomics data has been released to facilitate basic biology and medical research. scRNA-seq technology involves the sequence analysis and comparison of single-cell DNA and RNA, which can reflect cell heterogeneity and clarify differences in cellular interactions, thus offering unique advantages in exploring cell types associated with tumorigenesis, tumor progression, and metastasis [4,5]. Since 2009, when *Tang* et al. [6] first developed a single-cell RNA transcriptome analysis technique, there has been a continuous improvement in this technology, and scRNA-seq has evolved into a crucial method for the large-scale study of individual cell transcripts. This technology has led to discoveries in many scientific fields [7]. Since the success of precision therapy in breast cancer depends on clarifying the key molecular mechanisms driving tumor growth and metastasis, single-cell sequencing may provide a novel way to uncover the intra-tumor heterogeneity of various types of breast cancer by exploring the tumor microenvironment at high resolution.

Given that it is a rapidly growing field of research, it is vital to investigate the utilization of scRNA-seq techniques in breast cancer from a macroscopic perspective as well as to predict future research trends. Bibliometrics, is a method quantitative analysis of scholarly literature that involve the statistical analysis and visualization of the information in publications to help scholars review the relevant research over time and predict emerging trends [8,9]. To our knowledge, there has not been a bibliometric analysis focusing on scRNA-seq technology in breast cancer research. Therefore, this study was conducted to objectively describe recent advancements and key areas of interest in this domain by utilizing bibliometric tools like VOSviewer and CiteSpace. In addition, we mapped the global collaboration patterns and development trends to serve as a valuable reference for future investigations in this field.

2. Materials and methods

2.1. Data collection

The Web of Science (WoS) serves as world's most extensive and inclusive repository of academic information, encompassing the most foremost influential core academic journals across a wide range of research disciplines, including the natural sciences, engineering, technology, and biomedicine [10]. For this research, data were collected and analyzed by querying the Web of Science Core Collection (WoSCC) database. To avoid omissions due to the database's continuous updates, literature retrieval, and data download were completed within one day (October 1, 2022). The retrieval keywords were as follows: *TS = ("breast cancer" OR "breast*

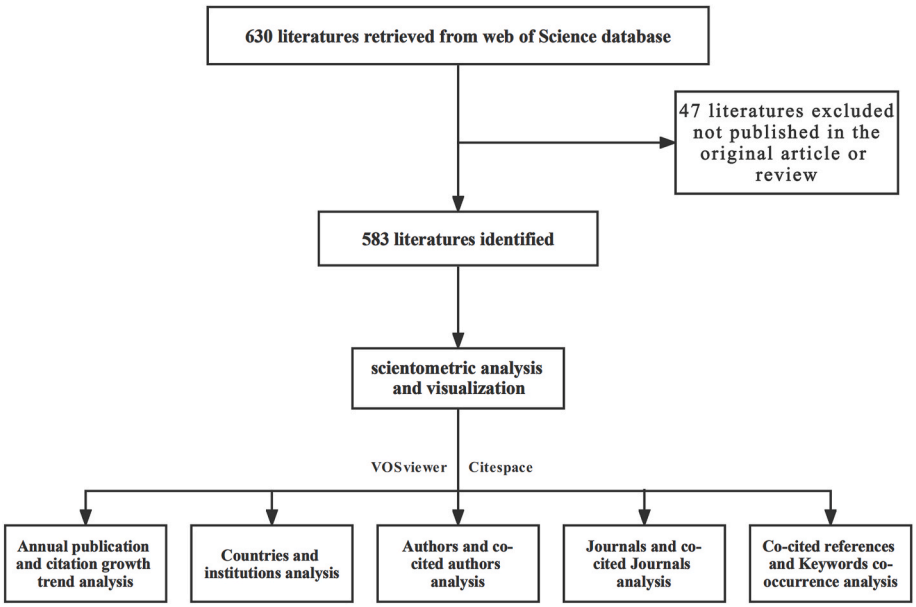


Fig. 1. Flowchart of literature screening and selection.

carcinoma") AND TS = ("single-cell RNA sequencing" OR "scRNA-seq" OR "single-cell RNA-seq" OR "single-cell sequencing" OR "single-cell transcriptomic" OR "single-cell omics sequencing" OR" single-cell omics" OR "single-cell profiling" OR "single-cell immune profiling"), with the publication period from January 1, 2009, to September 30, 2022. The article type was limited to reviews and original articles published in English. 583 articles were identified, with all records and citations being downloaded in TXT format. The process of literature filtering is depicted in Fig. 1. Two researchers (HR and XDL) independently conducted the search.

2.2. Bibliometric analysis

Citation reports and search results were derived from the WoSCC database, which included the annual number of publications, annual citation counts, publication numbers across various countries and institutions, and number of categories. Next, we used Microsoft Excel 2021, Tableau Public, CiteSpace and VOSviewer to conduct the bibliometric analysis and visualization [11,12]. Microsoft Excel 2021 was utilized to illustrate the yearly trends in publication numbers and citations as well as to process the merging of keywords that had different formats but the same or similar meanings. Tableau Public is a visual analytics platform for exploring, creating, and sharing data publicly, and we used it to construct a global map displaying the number of publications.

CiteSpace, created by *Chaomei Chen*, is an open-source Java tool for visualizing and analyzing scientific literature trends [13]. It specializes in identifying key milestones, intellectual turning points, and pivotal moments in a field's evolution [14]. In this study, we applied CiteSpace for country/region and institution co-occurrence analysis, dual mapping of journals, keyword clustering, timeline, and burst detection in references and keywords. The detailed parameter settings: the time span extended from January 2009 to October 2022, with each slice representing one year. Text processing, node type, link strength/range, pruning of sliced networks, and selection criteria adhered to the default settings. In addition, nodes with betweenness centrality values ($\geq$0.1) were used as purple circles to emphasize the importance of the literature (or authors, journals, countries, etc.) in the field [15].

VOSviewer is a tool for building and visualizing bibliometric networks [16]. These networks can encompass journals, researchers, or individual publications, etc. VOSviewer was applied to map the country/institution and author collaboration networks, journal/-literature co-citation networks, and keyword co-occurrence networks.

3. Results

3.1. Annual publication and citation growth trend

The annual publication counts and citation frequencies indicate the research trends within the field. 583 publications were initially retrieved through the WoSCC database (478 articles and 105 reviews). The geographical map of the study countries and regions was used to map the global contribution to the breast cancer single-cell sequencing research field. The USA contributed the most to the field with 272 (46.7 %) articles, followed by China with 167 (28.6 %), England with 37 (6.3 %), Germany with 32 (5.5 %), and Australia with 30 (5.1 %) (Fig. 2A). In addition, the yearly trend in publications and citation frequency for articles on single-cell sequencing in breast cancer from 2009 to 2021 exhibited a continuous upward trend (Fig. 2B).

3.2. Spatial distribution of countries/regions and institutions

This study analyzed 583 papers from 302 institutions across 46 countries/regions. Based on the country collaboration map, the USA and China exhibited the most frequent collaborations, while the frequency of collaborations other nations was relatively lower (Fig. 3A and B). Table 1 presents the top 10 countries/regions and institutions by publication count. Notably, the USA (272) and China (167) produced a significantly more papers than other countries, exceeding four times the number of papers from other nations. The Harvard Medical School published the most papers (23) among all research institutions. Furthermore, 30 % of the top 10 institutions were affiliated with China. Moreover, multiple countries and affiliated institutions, including the USA (0.88), China (0.22), Sweden (0.09), *Harvard Medical School* (0.18), *Stanford University* (0.14), and *Sun Yat-sen University/MD Anderson Cancer Center* (0.13), demonstrated

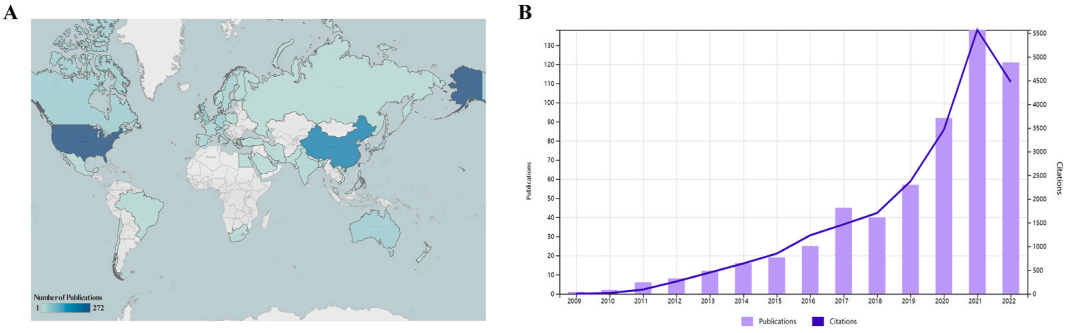


Fig. 2. Distribution of publications and citation trends by country and time. **(A)** Global map of the distribution of publications on single-cell sequencing in breast cancer. **(B)** Annual publication growth trend of research on breast cancer single-cell sequencing.

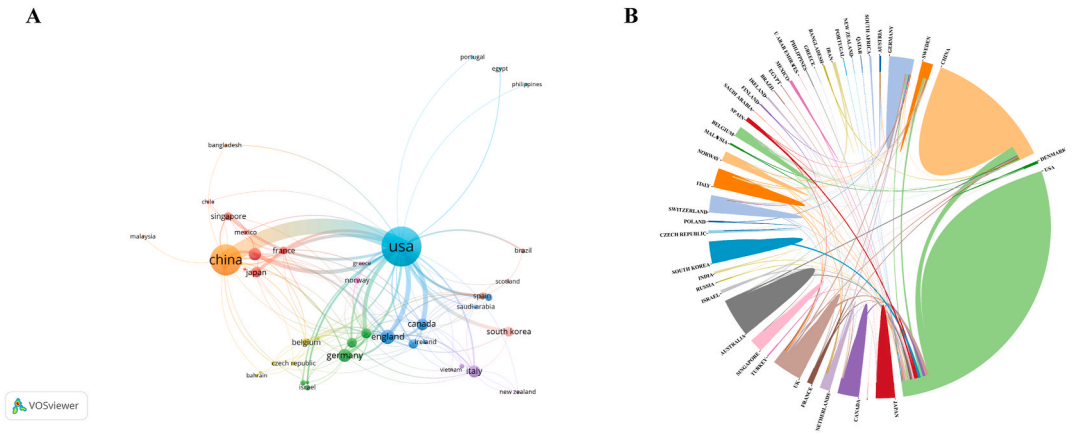


Fig. 3. Contribution of different countries/regions and institutions to publications on this topic. (A) Network map of countries. **(B)** Cooperation analysis between countries/regions.

Table 1
The top 10 high-output countries/regions and institutions.

Rank	Countries/Regions	Number of publications	Centrality	Institutions	Number of publications	Centrality
1	USA	272	0.88	Harvard Med Sch	23	0.18
2	China	167	0.22	Univ Texas MD Anderson Canc Ctr	19	0.13
3	England	37	0.06	Stanford Univ	17	0.14
4	Germany	32	0.08	Dana Farber Canc Inst	16	0.03
5	Australia	30	0	Chinese Acad Sci	15	0.11
6	Italy	28	0.06	Brigham & Womens Hosp	10	0.02
7	Canada	27	0.05	Sun Yat Sen Univ	10	0.13
8	Japan	21	0.01	National Canc Inst	10	0.11
9	South Korea	19	0	Fudan Univ	9	0.09
10	Sweden	19	0.09	Harvard Univ	9	0.12

high centrality in the centrality ranking, as denoted by the purple circles in Fig. S1A and B. This observation highlights the significance and impact of these nations and institutions in the realm of single-cell sequencing of breast cancer.

3.3. Visual examination of authors and co-cited researchers

Among the 4089 contributors to these publications, we screened a total of 422 authors who had published more than two papers to generate the author collaboration network (Fig. 4A). The size of each node reflects the publication count, whereas the links between nodes denote the collaborative efforts among researchers. The collaborative network is fairly decentralized, forming several distinct clusters, represented by *"Navin, Nicholas," "Aceto, Nicila," "Kendall, jude,"* and *"Perou, Charles M."* Moreover, according to Table 2, *YMAX S WICHA* published the most numbers of papers in this area (n = 5), followed by a group of authors, including *CAROL A*

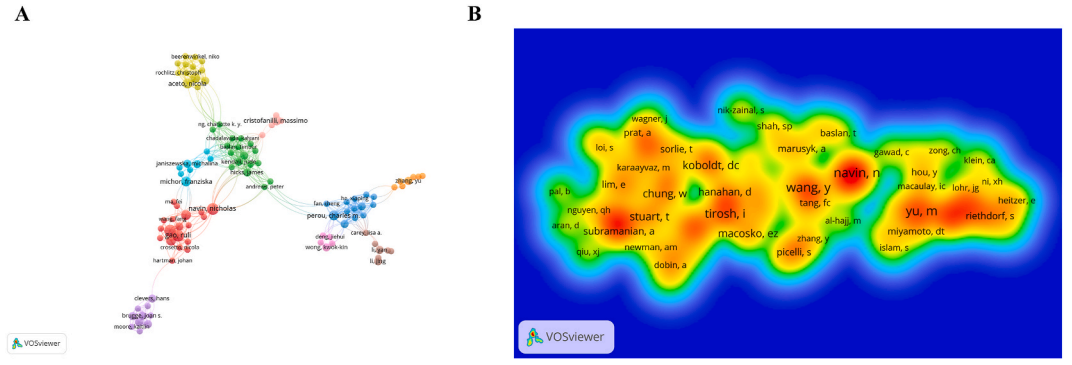


Fig. 4. Contribution of different authors to publications on this topic. (A) Collaboration of the author network map. **(B)** Co-citation analysis density map of authors.

Table 2
The top 10 authors and co-cited authors.

Rank	Author	Country	Count	Co-cited author	Co-citations	Centrality
1	MAX S WICHA	USA	5	WANG Y	81	0.02
2	CAROL A SARTORIUS	USA	4	NAVIN N	76	0.16
3	NICHOLAS NAVIN	USA	4	YU M	71	0.03
4	RULI GAO	USA	4	TIROSH I	68	0.04
5	ALEXANDER SWARBRICK	Australia	4	STUART T	65	0.01
6	FAN YANG	China	4	CHUNG W	64	0.02
7	TAO HUANG	Australia	4	GERLINGER M	62	0.15
8	SUNNY Z WU	Germany	3	KOBOLDT DC	61	0.04
9	BERNHARD POLZER	USA	3	CRISTOFANILLI M	60	0.19
10	JOAN S BRUGGE	Switzerland	3	MACOSKO EZ	60	0.01

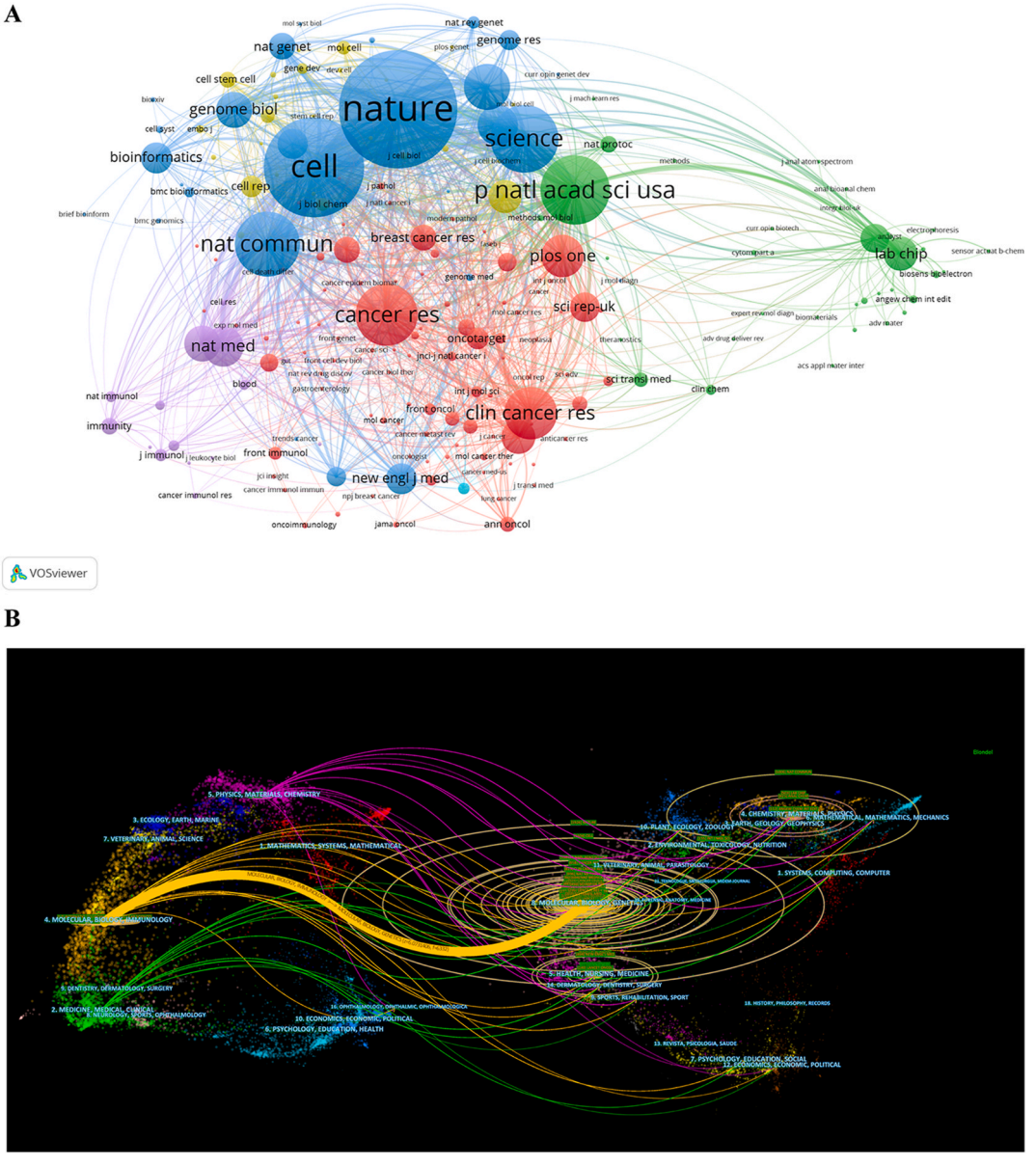


Fig. 5. Contribution of different journals to publications on this topic. (A) Co-citation network of journals based on references. **(B)** The dual-map overlay of journals with publications on single-cell sequencing in breast cancer.

SARTORIUS, NICHOLAS NAVIN, RULI GAO, ALEXANDER SWARBRICK, FAN YANG, and *TAO HUANG* (n = 4). The co-citation density plot reveals that 78 out of 20,919 referenced authors have been cited over 30 times (Fig. 4B), with warmer colors in the plot indicating higher citation counts. Among the top 10 most frequently co-cited authors, *"WANG Y"* had the highest citation count (n = 81), followed by *"NAVIN N"* (n = 76) and *"YU M"* (n = 71), as detailed in Table 2.

3.4. Visual analysis of journal distribution

To identify for the most impactful journals, we used VOSviewer to visualize and identify the most impactful journals in breast cancer scRNA-seq research. The analysis revealed that 583 articles were distributed across 201 academic journals, which were clustered into five groups (Fig. 5A). As shown in Table 3, *Nature Communications* (28 articles, IF: 17.694) published the most articles on breast cancer single-cell sequencing, followed by *Cancers* (21 articles, IF: 6.575) and *Frontiers in Oncology* (19 articles, IF: 5.738). Seven of the top ten journals fall under JCR Q1 with an impact factor above 5. The significance of journals in a research field is measured by how frequently they are cited. Journal co-citation analysis showed that 200 out of 2950 journals were cited over 30 times. As shown in Table 3, *Nature* had the highest co-citation frequency (cited 1928 times, IF: 69.504), followed by *Cell* (cited 1604 times, IF: 66.85) and *PNAS* (cited 1081 times, IF: 12.779). Based on the 2022 Journal Citation Reports (JCR), all but Plos One among the top ten journals were ranked in the Q1 category (Table 3). Furthermore, a dual-map overlay was utilized for visual analysis to compare global academic journal relationships [17]. In Fig. 5B, the left side illustrates the citing journals, while the right side shows the cited journals, with colored paths indicating citation relationships. The orange path demonstrates that papers published in Molecular/Biology/Immunology journals were frequently cited by those in Molecular/Biology/Genetics journals.

3.5. Co-citation references and reference bursts

We analyzed the reference lists of papers included in our study and constructed a co-citation network. Among the 27,217 references cited, 82 papers were cited over 20 times (Fig. 6A). Table 4 details the ten most frequently cited references. The most referenced work was a 2017 paper by *Chung* et al. [18], entitled *"Single-cell RNA-seq enables comprehensive tumor and immune cell profiling in primary breast cancer"* published in *Nature Communications*. The second most highly cited reference was an article, titled *"Integrating single-cell transcriptomic data across different conditions, technologies, and species"* [19]. These two articles discuss the application of scRNA-seq in evaluating the TME of breast cancer, as well as its application in different conditions and species. To highlight the evolving trends and prominent areas in research, we analyzed the temporal analysis of the co-cited references (Fig. 6B). Clusters on the right of each line represent the high-frequency keywords, with purple indicating an earlier appearance and red indicating a more recent appearance. Cluster #0, circulating tumor cells, was the largest, followed by liquid biopsy (Cluster #1), cancer stem cells (Cluster #2), tumor suppressor (Cluster #3), whole genome amplification (Cluster #4), lineage negative/positive cell populations (Cluster #5), single-cell DNA sequencing (Cluster #6), and tumor cell plasticity (Cluster #7). Clusters #0, #1, and #2 are the current research trends and frontiers, with increased activity over the last two years.

Analyzing reference bursts gives an insight into the research hotspots and emerging trends over specific periods [20]. In CiteSpace, burst duration was set to one year, and the 25 most burst co-cited references were identified (Fig. 6C). The first burst of co-citation emerged in 2012 with the paper, entitled *"Tumor evolution inferred by single-cell sequencing"* [21]. The strongest burst paper (strength = 12.61) was a paper, entitled *"Clonal evolution in breast cancer revealed by single nucleus genome sequencing"*, published by *Wang* et al. [22] in *Nature* in 2014 with a burst duration from 2016 to 2018. Notably, four references had bursts before 2022, indicating that the research related to single-cell sequencing in breast cancer is likely to continue its rapid expansion in the future.

Co-occurrence of keywords and keyword bursts.

In scientific articles, keywords encapsulate the essential content and primary theme of the article. Examining keyword co-occurrence can identify emerging research trends and focal points. Among 2636 keywords, 79 were mentioned at least 10 times. As shown in Fig. 7A and Table 5, "breast cancer" appeared most frequently (254 times), followed by "single-cell sequencing" (92 times), "heterogeneity" (86 times), "single-cell analysis" (85 times), and "circulating tumor cells" (81 times). Fig. 7B shows the degree of keyword appearance over time, with different colors representing the order of events. It is worth noting that nodes such as "tumor microenvironment," "immunotherapy," and "prognosis" appear in yellow, indicating that these fields have gained significant

Table 3
The top 10 journals and co-cited journals.

Rank	Journal	Count	JCR	IF (2021)	Co-cited journal	Citation	JCR	IF (2021)
1	Nature Communications	28	Q1	17.694	Nature	1928	Q1	69.504
2	Cancers	21	Q1	6.575	Cell	1604	Q1	66.85
3	Frontiers In Oncology	19	Q2	5.738	Pnas	1081	Q1	12.779
4	Scientific Reports	16	Q2	4.997	Science	1064	Q1	63.832
5	Cancer Research	12	Q1	13.312	Nature Communications	1006	Q1	17.694
6	Frontiers In Cell And Developmental Biology	12	Q1/Q2	6.081	Cancer Research	958	Q1	13.312
7	International Journal Of Molecular Sciences	12	Q1/Q2	6.208	Clinical Cancer Research	813	Q1	13.801
8	Bioinformatics	11	Q1	6.931	Nature Methods	698	Q1	47.99
9	Plos One	11	Q2	3.752	Plos One	644	Q2	3.752
10	Frontiers In Genetics	10	Q1	4.772	Nature Biotechnology	606	Q1	68.164

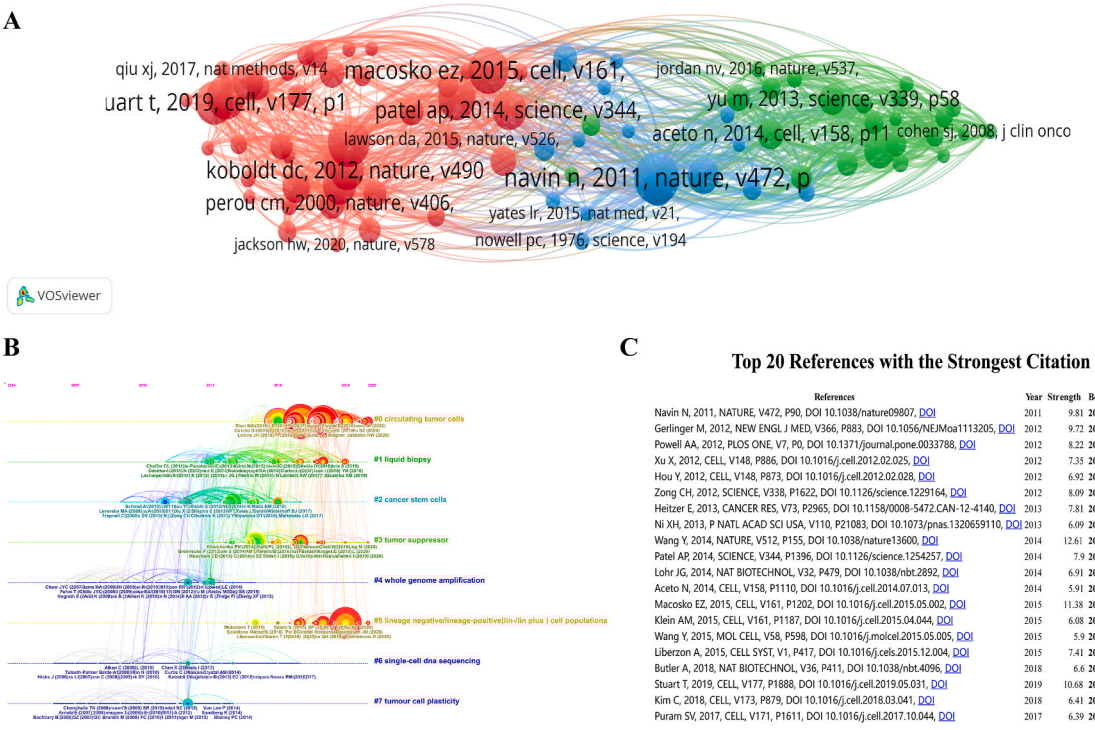


Fig. 6. Co-cited reference network visualization map. (A) Network of co-cited references. **(B)** A timeline view of the co-cited clusters related to single-cell sequencing in breast cancer. **(C)** Top 20 references with the strongest citation bursts.

Table 4
The top 10 co-cited references.

Rank	Title	First author	Publication year	Co-citations	Source
1	Single-cell RNA-seq enables comprehensive tumor and immune cell profiling in primary breast cancer	Chung, W	2017	64	Nature Communications
2	Integrating single-cell transcriptomic data across different conditions, technologies, and species	Butler, A	2018	59	Nature Biotechnology
3	Comprehensive Integration of Single-Cell Data	Stuart, T	2019	58	Cell
4	Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment	Azizi, E	2018	53	Cell
5	Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq	Tirosh, I	2016	48	Science
6	Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets	Macosko, EZ	2015	39	Cell
7	Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer	Puram, SV	2017	35	Cell
8	Chemoresistance Evolution in Triple-Negative Breast Cancer Delineated by Single-Cell Sequencing	Kim, C	2018	32	Cell
9	Unravelling subclonal heterogeneity and aggressive disease states in TNBC through single-cell RNA-seq	Karaayvaz, M	2018	32	Nature Communications
10	Clonal evolution in breast cancer revealed by single nucleus genome sequencing	Wang, Y	2014	31	Nature

popularity recently and are expected to be major topics in the future. In addition, there are distinct 10 clustering patterns in this field, as depicted in Fig. 7C. Using these clustering patterns, we leveraged CiteSpace software to create a timeline view of keyword clusters, which further clarifies the research trends within this domain (Fig. 7D). We found that the research hotspots have shifted from cluster #1 (liquid biopsy), #5 (breast cancer cell), and #6 (bulk RNA-seq) to cluster #0 (differentiation), #3 (triple-negative breast cancer), and #4 (immunotherapy), providing a reference for future research hotspots.

The burst analysis of keywords can assist in identifying keywords that may have academic contributions but have not yet reached the frequency threshold, thus enabling a more comprehensive analysis of the hotspots and emerging areas in single-cell sequencing in breast cancer research. In this research, CiteSpace identified burst keywords in breast cancer scRNA-seq articles from 2010 to 2022. The top 25 strongest burst keywords are shown in Fig. 7E. As illustrated in Fig. 7E, the field of breast cancer single-cell sequencing

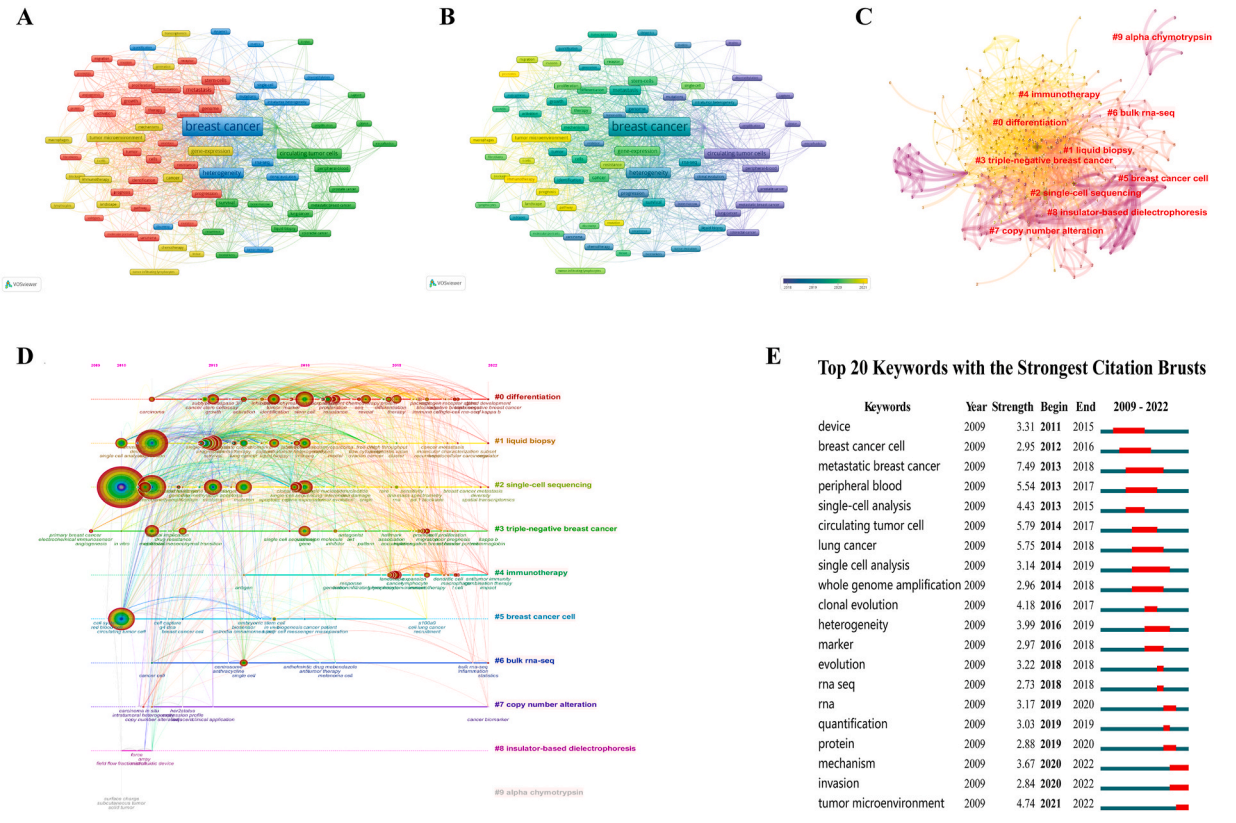


Fig. 7. Visual analysis of keywords in publications. (A) Co-occurrence networks of keywords in research. **(B)** Average year map of keywords. **(C)** Cluster network diagram of keywords on single-cell sequencing in breast cancer. **(D)** The timeline view of the keywords clusters. **(E)** Top 20 keywords with the strongest citation bursts involved in single-cell sequencing of breast cancer.

Table 5
The top 20 keywords.

Rank	Keyword	Occurrences	Total link strength	Rank	Keyword	Occurrences	Total link strength
1	breast cancer	254	935	11	survival	50	206
2	single-cell sequencing	92	407	12	tumor microenvironment	50	225
3	heterogeneity	86	367	13	genome	44	193
4	single-cell analysis	85	352	14	progression	42	198
5	circulating tumor cells	81	379	15	growth	38	148
6	metastasis	65	308	16	rna-seq	36	157
7	gene-expression	61	247	17	resistance	30	119
8	stem-cells	55	202	18	metastatic breast-cancer	27	127
9	cancer	51	193	19	tumor heterogeneity	25	135
10	evolution	51	229	20	immunotherapy	25	101

research exhibits diverse characteristics, with different burst terms appearing in different periods. The highest burst term was “metastatic breast cancer” (n = 7.49), followed by “circulating tumor cell” (n = 5.79) and “lung cell” (n = 5.75). Furthermore, the latest burst words were “mechanism,” “invasion,” and “tumor microenvironment."

4. Discussion

4.1. Principal information

Breast cancer is the most prevalent malignancy globally, with increasing incidence. Recently, “precision therapy” has garnered increasing attention [1]. The utilization of scRNA-seq and other advanced technologies in breast cancer research can reveal cell heterogeneity and functional characteristics, thus providing new perspectives on the occurrence, progression, and treatment of breast cancer [23,24]. Bibliometrics is a research method based on statistics that aims to evaluate the academic research level, hotspots, and trends within a field by analyzing metrics such as the volume, quality, and citations of literature [9]. In the context of single-cell

sequencing research on breast cancer, the application of bibliometrics can help researchers identify research hotspots and trends, evaluate the research level and academic contributions, and guide future research directions and decision-making. This study conducted a bibliometric analysis of single-cell sequencing and breast cancer publications from 2009 to 2022 using VOSviewer and CiteSpace. By of September 30, 2022, 583 publications were identified from 201 journals, authored by 4089 authors from 46 countries/regions and 302 institutions. The annual publication count indicates research intensity and trends in single-cell sequencing in breast cancer. From 2009 to 2021, both the annual publication trend and citation frequency on single-cell sequencing in breast cancer exhibited exponential growth, highlighting the rapid rise of this technology as a focal point among breast cancer researchers and institutions worldwide.

In the field of single-cell sequencing for breast cancer, the United States exhibits the highest centrality in the international publication network, evidenced by a centrality degree of 0.88, underscoring its leadership in global research. This preeminence is attributed to its advanced professional equipment, a large cohort of high-level researchers, and substantial funding support. Following closely is China, with a centrality degree of 0.22, reflecting its rapid scientific advancement, strong governmental support for scientific research, and increasingly robust connections with the international scientific community. At the institutional level, *Harvard Medical School* (0.18), *Stanford University* (0.14), and *Sun Yat-sen University/MD Anderson Cancer Center* (0.13) exhibit higher centrality degrees. These institutions distinguish themselves through their capabilities in securing funding, state-of-the-art research infrastructure, and outstanding academic reputation. *Harvard Medical School* leverages extensive collaborations with numerous affiliated hospitals to accumulate a vast array of clinical samples, providing rich data resources for single-cell sequencing studies. *Stanford University* is celebrated for its innovative culture and applications in bioinformatics and computational biology, driving forward the development of scRNA-seq technologies. Simultaneously, *Sun Yat-sen University* and *MD Anderson Cancer Center* are recognized as pivotal nodes in research due to their exceptional performance in disease-specific research and clinical applications. The exceptional performance of these institutions is not solely derived from their geographic and financial advantages but also their excellence in scientific research, technological innovation, and international collaborations, collectively propelling the advancement of scRNA-seq technology in breast cancer research. Moving forward, it is evident that enhancing global collaborations is essential. By harnessing the capabilities of top institutions, the international reach and influence of research can be expanded, thereby promoting a more cohesive global research network.

To determine the most productive authors in the realm of scRNA-seq for breast cancer, we ranked authors according to their total number of publications in this field. This approach provided a thorough understanding of the most prolific authors in this research area. *YMAX S WICHA* (n = 5) emerged as the most productive author in this field. The author developed a high-throughput sequencing method, Hydro-Seq, to detect and study circulating tumor cells (CTCs) based on scRNA-seq, and detailed the characteristics of CTCs in breast cancer patients, shedding light on the potential metastatic process of breast cancer [25]. Additionally, Professor Wicha, founding director of the University of Michigan Rogel Cancer Center, led his laboratory in identifying breast cancer stem cells in human solid tumors for the first time. Subsequently, his team developed widely used models and markers, which have been translated into clinical trials targeting cancer stem cells, pioneering new avenues for breast cancer treatment. Interestingly, the centrality of the top ten most productive authors was relatively low ($\leq$0.01), suggesting a lack of significant academic communication among researchers. Therefore, we highly recommend that scholars from the USA, China, and other countries strengthen academic collaboration in the future to break academic barriers and advance research on scRNA-seq in breast cancer. Among the co-cited authors, *"WANG Y"*, from *MD Anderson Cancer Center*, garnered the highest citation frequency (n = 81) for the paper *"Clonal evolution in breast cancer revealed by single nucleus genome sequencing,"* which systematically describes the clonal evolution process of breast cancer and the diagnostic, therapeutic, and evolutionary significance of chemoresistance through the development of a whole-genome and exome scRNA-seq method known as nucu-seq [22]. Additionally, as this study was published in Nature, a leading global scientific journal, it has garnered significant interest from the global scientific community. The research is also associated with a leading cancer research institution, and its findings have significant potential for clinical application, aiding in the advancement of personalized medicine and having a broad impact in both the scientific and clinical practice communities. These factors together have made *WANG Y* an extremely influential figure in cancer research, particularly regarding breast carcinoma management.

Among the 201 journals identified, *"Nature Communications"* (n = 28) published the most articles in the area of scRNA-seq in breast cancer, ranking fifth in total citations (n = 1006). *"Cancers"* and *"Frontiers in Oncology"* journals also had a significant impact on the dissemination of research on single-cell sequencing in breast cancer, ranking second and third in article count, respectively. Regarding the analysis of co-citation, *"Nature"* had the most co-citations (n = 1928), showing its significant influence in scRNA-seq breast cancer research. Although "*Cancer Research*" ranked sixth in total citations, it published 10 of the 12 most cited articles, highlighting its importance for researchers in this field. Notably, the above journals all focus on the fields of molecular biology, immunology, cell biology, cell and tissue engineering, aligning with the results of the dual-map analysis. Frequently co-cited references appeared together in the third most cited paper, forming a co-citation relationship [26]. The 10 most cited articles were all published in top journals. The research on "clonal evolution," "tumor microenvironment," "chemotherapy resistance," and "clonal heterogeneity and aggressiveness" was cited more than 450 times [22,27–29]. Regarding burst detection analysis of the literature, there are currently four articles in a burst state that are worth paying attention to, all involving scRNA-seq technology in breast cancer [18,27].

4.2. Hotspots and trends in scRNA-seq in breast cancer

Amid an ever-growing explosion of information, keywords summarize the main research topics and academic ideas, facilitating a more effective understanding of the evolving trends within the research domain [30]. In this study, we assessed the key areas and emerging trends in breast cancer studies utilizing scRNA-seq by conducting keyword co-occurrence analysis, constructing keyword

timelines, and performing keyword burst analysis [20,31,32]. According to the results, the research directions include heterogeneity, circulating tumor cells, tumor microenvironment, and immunotherapy, each of which has its distinct themes [33]. In oncology research, scRNA-seq technology has revealed the molecular mechanisms of tumor heterogeneity and evolution, thus providing new ideas and approaches for precise cancer treatment [34]. Recently, the keywords used were "tumor microenvironment," "immunotherapy," and "prognosis." We found that the research focus has shifted to clinical applications, particularly in the field of immunology [35,36], where immunogenic cell death-based cancer vaccines and immune checkpoint inhibitors are being widely studied. This suggests that pinpointing treatment targets for breast cancer is crucial to scRNA-seq research.

scRNA-seq technology is revolutionizing immunotherapy research in breast cancer, particularly by enhancing our understanding of the tumor microenvironment. This technology allows researchers to identify and classify various cell types within the tumor microenvironment and uncover the heterogeneity among immune cells like T cells and macrophages. This is crucial for discovering new biomarkers and therapeutic targets. For instance, detailed gene expression analysis can pinpoint which patients might benefit from therapies targeting immune checkpoints such as PD-L1, furthering the development of personalized treatment plans. Additionally, single-cell sequencing can dynamically monitor the tumor microenvironment before and after treatment, tracking changes in the functional states of immune cells and the escape mechanisms of tumor cells. This monitoring is vital for understanding why some patients develop resistance to treatments and provides a foundation for adjusting therapeutic strategies. By elucidating the complex interactions among immune cells, this technology also aids in developing new combination therapy strategies. These might integrate immune checkpoint inhibitors with other treatments like chemotherapy, radiotherapy, or targeted drugs, thereby enhancing treatment efficacy. Therefore, it is necessary to continue investigating this issue. In addition, keywords identified some "latest" research topics, including differentiation, whole genome amplification, and circulating tumor cell. Although these topics may not represent the entire field, they still attempt to answer important research questions and ensure future research.

4.3. Recommendations for future work

Through the burst analysis of keywords, we were surprised to find that the keywords "tumor microenvironment" (TME) and "invasion" have shown prominent bursts over past three years, which are also the keywords that we are particularly interested in. Recent research increasingly emphasizes the TME's crucial role in breast cancer progression and treatment effectiveness [37]. The term "invasion" typically describes how breast tumor cells disrupt the original and surrounding normal tissue structures, eventually detaching from the primary tumor and dispersing to varying degrees in the surrounding tissue [38,39]. Moreover, advancements in single-cell sequencing have provided deeper insights into the TME, revealing its immense cellular heterogeneity and identifying novel biomarkers and therapeutic targets. From the summary of high-frequency keywords in breast cancer single-cell sequencing research (Table 5) and results of keyword burst detection (Fig. 7), it is evident that immunotherapy and the TME have been persistent research focal points, with notable activity peaks from 2010 to 2022. This suggests that future research will likely delve deeper into the mechanisms by which the TME influences tumor progression and response to therapy, potentially leading to the development of more effective and personalized immunotherapeutic strategies.

The integration of AI and machine learning with traditional research methods promises to revolutionize our approach to understanding the TME. AI algorithms can handle and interpret vast amounts of data from single-cell sequencing to identify complex patterns and interactions within the TME that may not be discernible through traditional methods. This capability allows for the prediction of tumor behavior and response to treatments more accurately. Furthermore, machine learning models can be trained to predict the efficacy of specific immunotherapeutic agents in individual patients, thereby enhancing the personalization of treatment regimens. This sophisticated technology not only speeds up research and development but also improves the precision of therapeutic interventions, making significant strides toward truly personalized medicine in breast cancer treatment.

4.4. Significance of scRNA-seq in breast cancer research

scRNA-seq is critical in breast cancer research, driving advancements in personalized and precision medicine. This technology enables the precise analysis of individual cells within a tumor, uncovering both genomic and transcriptomic diversity. Such insights are crucial for scientists to grasp the tumor's intricate and evolving nature. By pinpointing the molecular traits of distinct cell subpopulations, scRNA-seq helps both in uncovering new biomarkers and in formulating customized treatment plans. It is particularly effective in targeting subpopulations that respond to specific drugs. Additionally, Intrinsic drug-resistant cells within tumors often survive treatment and quickly form highly resistant tumors. To effectively treat heterogeneous tumors, it is necessary to eliminate resistant cells through targeted therapy before conventional treatment. Treatment strategies include inducing a response by phenotypic conversion or specifically eliminating resistant cells. Targeting the root is the best way to describe the impact of tumor heterogeneity, and scRNA-seq does just that. scRNA-seq has changed the way researchers explore the impact of tumor heterogeneity in patient samples and has provided significant insights into the astonishing cellular and genetic diversity within tumors and their microenvironment. The technology plays a vital role in analyzing treatment resistance and refining therapeutic approaches, thereby boosting treatment effectiveness and ensuring optimal patient outcomes and quality of life. This research is the first systematic bibliometric analysis of single-cell sequencing in breast cancer over the past decade, outlining its knowledge framework and progression. The study's limitations involve relying exclusively on WoSCC data, possibly overlooking significant articles from other databases, and potential inconsistencies due to continual database updates.

5. Conclusion

Our findings underscore the essential role of scRNA-seq in breast cancer research. The number of publications on scRNA-seq in breast cancer has increased dramatically since 2009. The USA leads this field, with China following closely. Nonetheless, there remains a need for enhanced collaboration and communication among countries, regions, and authors in breast cancer scRNA-seq research. Additionally, an increasing number of papers have been published in top international journals, highlighting the significant impact of scRNA-seq in this area. Notably, the investigation of immune-related therapies using scRNA-seq in breast cancer is emerging as a key research focus and trend.

Funding

None.

Data availability statement

The data can be acquired by Web of Science.

Declarations

None.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

CRediT authorship contribution statement

Shan Liu: Software, Methodology, Investigation. **Xudong Li:** Visualization, Methodology, Formal analysis. **Ying Zhang:** Data curation. **Yuhan Deng:** Methodology, Investigation. **Zehao Li:** Formal analysis, Conceptualization. **Yunan Zhu:** Formal analysis. **Xue Li:** Methodology. **Yuefeng Shang:** Methodology, Formal analysis. **Guang Yang:** Software, Methodology. **Xiaolu Zhan:** Writing – original draft, Resources, Conceptualization. **Yingpu Li:** Writing – review & editing, Conceptualization. **He Ren:** Writing – review & editing, Writing – original draft, Software, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Not applicable.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e33219.

References

[1] H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, F. Bray, Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, Ca - Cancer J. Clin. 71 (2021) 209–249.
[2] W. Chen, R. Zheng, P.D. Baade, S. Zhang, H. Zeng, F. Bray, A. Jemal, X.Q. Yu, J. He, Cancer statistics in China, 2015, Ca - Cancer J. Clin. 66 (2016) 115–132.
[3] S. Lei, R. Zheng, S. Zhang, R. Chen, S. Wang, K. Sun, H. Zeng, W. Wei, J. He, Breast cancer incidence and mortality in women in China: temporal trends and projections to 2030, Cancer Biol Med 18 (2021) 900–909.
[4] J. Liang, W. Cai, Z. Sun, Single-cell sequencing technologies: current and future, J Genet Genomics 41 (2014) 513–528.
[5] O. Stegle, S.A. Teichmann, J.C. Marioni, Computational and analytical challenges in single-cell transcriptomics, Nat. Rev. Genet. 16 (2015) 133–145.

[6] F. Tang, C. Barbacioru, Y. Wang, E. Nordman, C. Lee, N. Xu, X. Wang, J. Bodeau, B.B. Tuch, A. Siddiqui, et al., mRNA-Seq whole-transcriptome analysis of a single cell, Nat. Methods 6 (2009) 377–382.
[7] E. Hedlund, Q. Deng, Single-cell RNA sequencing: technical advancements and biological applications, Mol. Aspect. Med. 59 (2018) 36–46.
[8] H.F. Moed, New developments in the use of citation analysis in research evaluation, Arch. Immunol. Ther. Exp. 57 (2009) 13–18.
[9] A. Ninkov, J.R. Frank, L.A. Maggio, Bibliometrics: methods for studying academic publishing, Perspect Med Educ 11 (2022) 173–176.
[10] P.F. Lyu, J.T. Li, T. Deng, G.X. Lin, P.M. Fan, X.C. Cao, Research trends and hotspots of breast cancer management during the COVID-19 pandemic: a bibliometric analysis, Front. Oncol. 12 (2022) 918349.
[11] L. Miao, J. Zhang, Z. Zhang, S. Wang, F. Tang, M. Teng, Y. Li, A bibliometric and knowledge-map analysis of CAR-T cells from 2009 to 2021, Front. Immunol. 13 (2022) 840956.
[12] K. Cheng, Q. Guo, Z. Shen, W. Yang, Y. Zhou, Z. Sun, X. Yao, H. Wu, Frontiers of ferroptosis research: an analysis from the top 100 most influential articles in the field, Front. Oncol. 12 (2022) 948389.
[13] C. Chen, Searching for intellectual turning points: progressive knowledge domain visualization, Proc. Natl. Acad. Sci. U.S.A. 101 (Suppl 1) (2004) 5303–5310.
[14] C.M. Chen, F. Ibekwe-SanJuan, J.H. Hou, The structure and dynamics of cocitation clusters: a multiple-perspective cocitation analysis, J. Am. Soc. Inf. Sci. Technol. 61 (2010) 1386–1409.
[15] B. Shao, Y.F. Qin, S.H. Ren, Q.F. Peng, H. Qin, Z.B. Wang, H.D. Wang, G.M. Li, Y.L. Zhu, C.L. Sun, et al., Structural and temporal dynamics of mesenchymal stem cells in liver diseases from 2001 to 2021: a bibliometric analysis, Front. Immunol. 13 (2022) 859972.
[16] N.J. van Eck, L. Waltman, Software survey: VOSviewer, a computer program for bibliometric mapping, Scientometrics 84 (2010) 523–538.
[17] C. Chen, L. Leydesdorff, Patterns of connections and movements in dual-map overlays: a new method of publication portfolio analysis, Journal of the Association for Information Science and Technology 65 (2014) 334–351.
[18] W. Chung, H.H. Eum, H.O. Lee, K.M. Lee, H.B. Lee, K.T. Kim, H.S. Ryu, S. Kim, J.E. Lee, Y.H. Park, et al., Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer, Nat. Commun. 8 (2017) 15081.
[19] A. Butler, P. Hoffman, P. Smibert, E. Papalexi, R. Satija, Integrating single-cell transcriptomic data across different conditions, technologies, and species, Nat. Biotechnol. 36 (2018) 411–420.
[20] C. Chen, Z. Hu, S. Liu, H. Tseng, Emerging trends in regenerative medicine: a scientometric analysis in CiteSpace, Expet Opin. Biol. Ther. 12 (2012) 593–608.
[21] N. Navin, J. Kendall, J. Troge, P. Andrews, L. Rodgers, J. McIndoo, K. Cook, A. Stepansky, D. Levy, D. Esposito, et al., Tumour evolution inferred by single-cell sequencing, Nature 472 (2011) 90–94.
[22] Y. Wang, J. Waters, M.L. Leung, A. Unruh, W. Roh, X. Shi, K. Chen, P. Scheet, S. Vattathil, H. Liang, et al., Clonal evolution in breast cancer revealed by single nucleus genome sequencing, Nature 512 (2014) 155–160.
[23] Y. Zhang, H. Chen, H. Mo, X. Hu, R. Gao, Y. Zhao, B. Liu, L. Niu, X. Sun, X. Yu, et al., Single-cell analyses reveal key immune cell subsets associated with response to PD-L1 blockade in triple-negative breast cancer, Cancer Cell 39 (2021) 1578–1593.e1578.
[24] P. Savas, B. Virassamy, C. Ye, A. Salim, C.P. Mintoff, F. Caramia, R. Salgado, D.J. Byrne, Z.L. Teo, S. Dushyanthen, et al., Single-cell profiling of breast cancer T cells reveals a tissue-resident memory subset associated with improved prognosis, Nat. Med. 24 (2018) 986–993.
[25] Y.H. Cheng, Y.C. Chen, E. Lin, R. Brien, S. Jung, Y.T. Chen, W. Lee, Z. Hao, S. Sahoo, H. Min Kang, et al., Hydro-Seq enables contamination-free high-throughput single-cell RNA-sequencing for circulating tumor cells, Nat. Commun. 10 (2019) 2163.
[26] D. Ma, B. Yang, B. Guan, L. Song, Q. Liu, Y. Fan, L. Zhao, T. Wang, Z. Zhang, Z. Gao, et al., A bibliometric analysis of pyroptosis from 2001 to 2021, Front. Immunol. 12 (2021) 731933.
[27] E. Azizi, A.J. Carr, G. Plitas, A.E. Cornish, C. Konopacki, S. Prabhakaran, J. Nainys, K. Wu, V. Kiseliovas, M. Setty, et al., Single-cell map of diverse immune phenotypes in the breast tumor microenvironment, Cell 174 (2018) 1293–1308.e1236.
[28] C. Kim, R. Gao, E. Sei, R. Brandt, J. Hartman, T. Hatschek, N. Crosetto, T. Foukakis, N.E. Navin, Chemoresistance evolution in triple-negative breast cancer delineated by single-cell sequencing, Cell 173 (2018) 879–893.e813.
[29] M. Karaayvaz, S. Cristea, S.M. Gillespie, A.P. Patel, R. Mylvaganam, C.C. Luo, M.C. Specht, B.E. Bernstein, F. Michor, L.W. Ellisen, Unravelling subclonal heterogeneity and aggressive disease states in TNBC through single-cell RNA-seq, Nat. Commun. 9 (2018) 3588.
[30] C. Chen, Science mapping: a systematic review of the literature, Journal of Data and Information Science 2 (2017) 1–40.
[31] X. Yuan, C. Chang, X. Chen, K. Li, Emerging trends and focus of human gastrointestinal microbiome research from 2010-2021: a visualized study, J. Transl. Med. 19 (2021) 327.
[32] Q. Li, W. Yang, J. Li, Z. Shan, Emerging trends and hot spots in autoimmune thyroiditis research from 2000 to 2022: a bibliometric analysis, Front. Immunol. 13 (2022) 953465.
[33] W.L. Jin, M.Z. Jin, Y.Y. Tu, Organoids: a platform ready for glioblastoma precision medicine? Trends Cancer 6 (2020) 265–267.
[34] J. Liu, S. Qu, T. Zhang, Y. Gao, H. Shi, K. Song, W. Chen, W. Yin, Applications of single-cell omics in tumor immunology, Front. Immunol. 12 (2021) 697412.
[35] V. Semiglazov, A. Tseluiko, A. Kudaybergenova, A. Artemyeva, P. Krivorotko, R. Donskih, Immunology and immunotherapy in breast cancer, Cancer Biol Med 19 (2022) 609–618.
[36] M.Z. Jin, X.P. Wang, Immunogenic cell death-based cancer vaccines, Front. Immunol. 12 (2021) 697964.
[37] M.Z. Jin, W.L. Jin, The updated landscape of tumor microenvironment and drug repurposing, Signal Transduct. Targeted Ther. 5 (2020) 166.
[38] D.R. Sherwood, Cell invasion through basement membranes: an anchor of understanding, Trends Cell Biol. 16 (2006) 250–256.
[39] Z. Yuan, Y. Li, S. Zhang, X. Wang, H. Dou, X. Yu, Z. Zhang, S. Yang, M. Xiao, Extracellular matrix remodeling in tumor progression and immune escape: from mechanisms to treatments, Mol. Cancer 22 (2023) 48.