

Original Paper

Characterization of Global Research Trends and Prospects on Single-Cell Sequencing Technology: Bibliometric Analysis

Quan Wang^{1,2*}, MD, PhD; Ke-Lu Yang^{3*}, MN; Zhen Zhang^{1,2}, PhD; Zhu Wang⁴, MD, PhD; Chen Li^{1,2}, BA; Lun Li⁵, MD, PhD; Jin-Hui Tian⁶, PhD; Ying-Jiang Ye^{1,2}, MD, PhD; Shan Wang^{1,2}, MD, PhD; Ke-Wei Jiang^{1,2*}, MD

¹Department of Gastroenterological Surgery, Peking University People's Hospital, Beijing, China

²Laboratory of Surgical Oncology, Beijing Key Laboratory of Colorectal Cancer Diagnosis and Treatment Research, Peking University People's Hospital, Beijing, China

³Evidence-Based Nursing Center, School of Nursing, Lanzhou University, Lanzhou, China

⁴Department of Gastrointestinal Surgery, Shandong Provincial Hospital Affiliated to Shandong First Medical University, Jinan, China

⁵Department of Breast Surgery, Shanghai Cancer Center, Fudan University, Shanghai, China

⁶Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou, China

* these authors contributed equally

Corresponding Author:

Ke-Wei Jiang, MD
Department of Gastroenterological Surgery
Peking University People's Hospital
No.11 Xizhimen South Street, Xicheng District
Beijing, 100044
China
Phone: 86 010 88326600
Email: jiangkewei@pkuph.edu.cn

Abstract

Background: As single-cell sequencing technology has been gradually introduced, it is essential to characterize global collaboration networks and map development trends over the past 20 years.

Objective: The aim of this paper was to illustrate collaboration in the field of single-cell sequencing methods and explore key topics and future directions.

Methods: Bibliometric analyses were conducted with CiteSpace and VOSviewer software on publications prior to November 2019 from the Web of Science Core Collection about single-cell sequencing methods.

Results: Ultimately, we identified 2489 records, which were published in 495 journals by 14,202 authors from 1970 institutes in 61 countries. There was a noticeable increase in publications in 2014. The United States and high-income countries in Europe contributed to most of the records included. Harvard University, Stanford University, Karolinska Institutes, Peking University, and the University of Washington were the biggest nodes in every cluster of the collaboration network, and SA Teichmann, JC Marioni, A Regev, and FC Tang were the top-producing authors. Keywords co-occurrence analysis suggested applications in immunology as a developing research trend.

Conclusions: We concluded that the global collaboration network was unformed and that high-income countries contributed more to the rapidly growth of publications of single-cell sequencing technology. Furthermore, the application in immunology might be the next research hotspot and developmental direction.

(*J Med Internet Res* 2021;23(8):e25789) doi: [10.2196/25789](https://doi.org/10.2196/25789)

KEYWORDS

single-cell sequencing; bibliometric analysis; cancer; cancer genomics; bioinformatics; cancer subtyping; tumor dissociation; tumor microenvironment; precision medicine; immunology; development trends; hotspots; research topics; Web of Science; CiteSpace; VOSviewer; network

Introduction

A single cell is regarded as the fundamental unit of an organism. Moreover, there are approximately 10^{14} single cells that comprise the complex tissues and organs of *Homo sapiens* [1]. In order to study genomic or transcriptome information, most published studies to date have collected data by analyzing bulk tissue samples and true genomic information representing the average of millions of cells [2]. Accordingly, because of cell-to-cell variability, traditional sequencing methods only provide the average diversity of cells, instead of obtaining information on entire cellular heterogeneity [3]. When facing the complexity of disease, it is difficult for these averaged data sets to represent cell-to-cell variations; therefore, it is difficult for researchers to identify rare cells, including cancer stem cells, that play a key role in cancer progression. To avoid the weakness and limitations of bulk sequencing technologies, the single-cell method offers a novel possibility that focuses on the single-cell level [4], with detailed and comprehensive studies of individual cells rather than traditional analysis of bulk tissue. The first step of sequencing a single cell, however, involves capturing individual cells; therefore, it is of utmost importance to establish various methods for isolating single cells from abundant populations or rare single cells (<1%) from typical populations. Several approaches have been well-established, including mouth pipetting, serial dilution, robotic micromanipulation, flow-assisted cell sorting, microfluidic platforms, and laser-capture microdissection, however, the isolation of rare single cells (<1%) is far more challenging [2].

With increasing developments of high-throughput sequencing technologies over the past couple of decades, an increasing number of commercial platforms have been introduced for experimental and clinical applications, and thousands of genomes from numerous species have been sequenced [3]. Researchers now have a wealth of data to explore genomic or transcriptome information. Single-cell sequencing refers to the sequencing of a single-cell genome or transcriptome, to reveal cell population differences and cellular evolutionary changes [5]. Furthermore, these commercial methods interrogate the single-cell genome, transcriptome, and epigenome. For example, since the introduction of the first single-cell RNA-sequencing technique in 2009 [6], single-cell RNA sequencing has become a powerful and useful approach to study individual cell transcriptomes on a large scale [7], similar to the technology of multiomics sequencing [8]. When it comes to the applications of single-cell sequencing techniques, the important role of single-cell sequencing in various fields, including oncology [9], microbiology [10], neurology [11,12], reproduction [13], and immunology [14] has been highlighted. These rapidly developing methods have and will continue to lead to new discoveries in many scientific fields.

Recently, because single-cell sequencing technology has been gradually introduced for clinical applications, there has been an increase in research papers published online on this topic. There are more than 70 related records in this field, published online each month on PubMed, and these are of a variety of publication types. Some papers offered a brief introduction of single-cell sequencing technology [15], whereas others focused

on investigating its applications [16]. In some reviews [7,17], the current strengths and disadvantages of single-cell sequencing technology were summarized. Published original articles updated or simplified the algorithms of computational methods [18].

Based on this, it is important for us to better understand and learn more from the developmental trends and novel advancements of single-cell sequencing techniques macroscopically; however, rapidly learning about the current landscape of specific biomedical research is still a challenge. Fortunately, the emergence of bibliometric analysis provides an approach to statistically and quantitatively visualize evidence according to the information in published records in a specific field [19,20], which includes data, such as keywords, citation reports, authors, affiliated countries, institutions, and journals, from a large number of published studies. To the best of our knowledge, no specific study has focused on the characterization of research hotspots, global research collaborations, and developmental trends of single-cell sequencing techniques. To characterize current evidence and establish future research directions, it is essential to perform bibliometric analysis to map global collaboration patterns and developmental trends of published single-cell sequencing methods.

The aims of this bibliometric analysis were to map this research landscape through analysis of the information in published records.

Methods

Data Source and Search Strategy

We searched the Web of Science Core Collection (WoSCC) without publication date restrictions on November 12, 2019. We updated the database search on January 29, 2020, to complete data retrieval for the year of 2019; however, only the number of records published annually was updated.

The search was performed using the following keywords and terms: (“single-cell RNA sequencing” OR “scRNA-seq” OR “single-cell RNA-seq” OR “single-cell sequencing” OR “single-cell transcriptomic” OR “single-cell ATAC” OR “single-cell RNA-sequencing” OR “single-cell omics sequencing” OR “scRNA seq”). The search strategy was peer-reviewed and guided by TJH (who has over 10 years of experience as an information specialist). In this study, only publications in English were included; however, there was no restriction on data category.

Inclusion and Exclusion Criteria

Inclusion criteria were published records, including articles, comments, reviews, letters, and brief introductions on single-cell sequencing techniques. We included the records that, not only focused on the single-cell sequencing technology after the paper of single-cell RNA-sequencing published in 2009 [6], but also, generally used the single-cell sorting methods such as Flow sorting in early time. We excluded duplicates, conference abstracts, and manuscripts in a language other than English, for consistent and accurate information collection from multianalysis results based on published records. Two groups of reviewers (WQ/WZ/JKW and YKL/LL) independently screened the titles and abstracts to select the articles after

standard selection training. Full-texts were retrieved when necessary. Disagreements were discussed and solved between the 2 groups.

Social Network Map Software

CiteSpace (version 5.3.R4) was used for social network analysis of developmental dynamics, future trends, hotspots, and key points in scientific literature of a specific topic [21]. Burst detection, to identify keywords and references that appear with an abrupt change in frequency at a certain period, were considered to be hot keywords or references at that point in time. Clustering (or co-citation, ie, both A and B are cited by C) analysis of references was also conducted [21]. We used VOSviewer (version 1.6.9; Leiden University) to visualize the collaborations between authors of a list of publications as well as those between countries, institutions, and high frequency keywords [22]. One classification method uses a metric based on the number of co-authored articles, which allowed authors, institutions, countries or keywords to be clustered, where those belonging to the same group cooperated more with each other [23]. In network maps, unequal-size nodes with different colors represent differences in the number or frequency of published records in clusters among the same research topics [21]. Lines between nodes indicate the strength of collaborations (the

stronger the collaboration, the thicker the line) [24]. We also used overlay visualization, in which the color of the node represents the average year in which each author, institution, country or keyword was used [25].

Data were saved as *Plain Text with Full Record and Cited References* in WoSCC and imported into CiteSpace.

Statistical Analysis

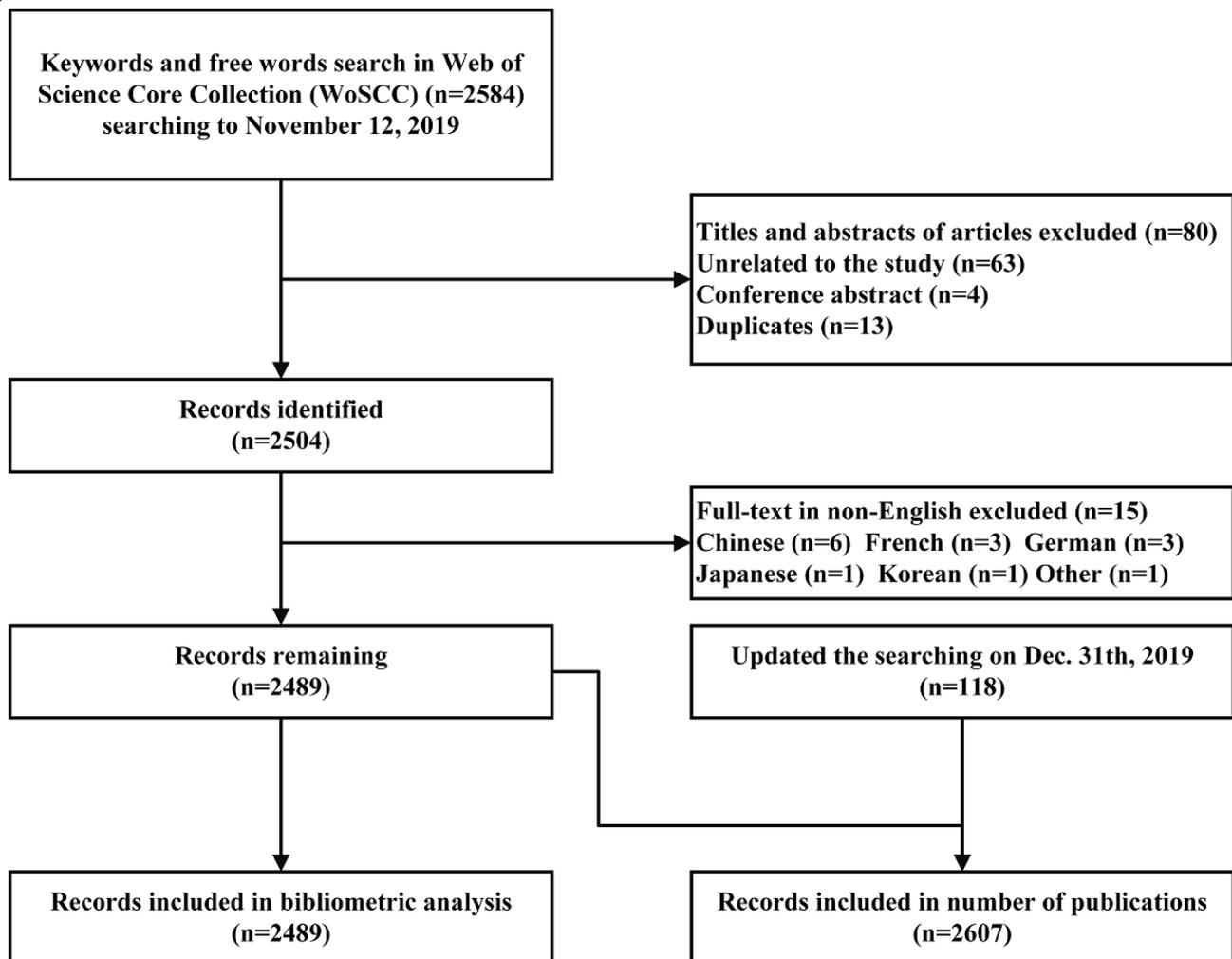
SPSS software (version 22.0; IBM Corp) and Origin (version 2018; OriginLab Corp) were used for data analysis. Continuous variables were presented as mean and standard deviation or median and interquartile range, and categorical variables were expressed using frequencies and percentages.

Results

Search Results

A total of 2584 publications were identified from WoSCC. After screening titles and abstracts, duplicates, unrelated topics, and conference abstracts (n=80) or non-English publications (n=15) were excluded. Finally, 2489 studies were included for bibliometric analysis (Figure 1); 2607 was identified as the number of records published.

Figure 1. Record identification and selection.



Global Publication Trends

Papers published over the past of two decades (Table 1 and Table 2) were primarily classified into 7 types. Most papers were original articles (1858/2607, 71.27%), and there were 388/2607 meeting abstracts (14.84%) and 217/2607 reviews (8.32%). Most publication types were articles, meeting abstracts, and reviews, which can greatly reflect the development trends

and changes in single-cell sequencing technology. The first publication on single-cell sequencing technology was in 2001 [26], which reported single-cell sequencing of dinoflagellate (Class Dinophyceae) nuclear ribosomal genes. In addition, there were only 1 to 3 publication per year before 2010. From 2010 to 2019, the number of publications showed a noticeable upward trend, with increases since the year 2014 (n=59), reaching 987 publications in 2019.

Table 1. Number of published records per year.

Year	Records, n
2001	1
2002	0
2003	2
2004	0
2005	1
2006	1
2007	1
2008	2
2009	2
2010	3
2011	10
2012	21
2013	35
2014	59
2015	127
2016	225
2017	406
2018	722
2019	989
Total	2607

Table 2. Percentage of published records by publication type.

Publication type	Records (n=2607), n (%)
Article	1858 (71.27)
Meeting abstract	388 (14.84)
Review	217 (8.32)
Editorial material	97 (3.72)
Correction	29 (1.11)
Letter	13 (0.54)
News item	5 (0.19)

Countries and Districts

In total, 61 countries contributed to publications on single-cell sequencing technology worldwide. The results demonstrated that 7 countries published more than 100 records (Table 3), and the United States (1454/2489, 58.4%), with more than half of all identified records, published the most. The top 2 to 7 were

China (414/2489, 16.6%), the United Kingdom (305/2489, 12.3%), Germany (225/2489, 9.0%), Sweden (134/2489, 5.4%), Japan (108/2489, 4.3%), and Switzerland (104/2489, 4.2%). Although the number of documents published by Australia, Canada, and the Netherlands was less than 100, no significant differences were observed ($P=.10$), and the contribution was almost equal to that of Japan and Switzerland. In addition, the

cooperation among studies in China was a little less than that in the United Kingdom. However, Sweden had the highest average number of citations (52.0 times), followed by Canada

(37.4 times), the United Kingdom (29.6 times), the United States (27.7 times), Australia (24.7 times), and Germany (23.9 times).

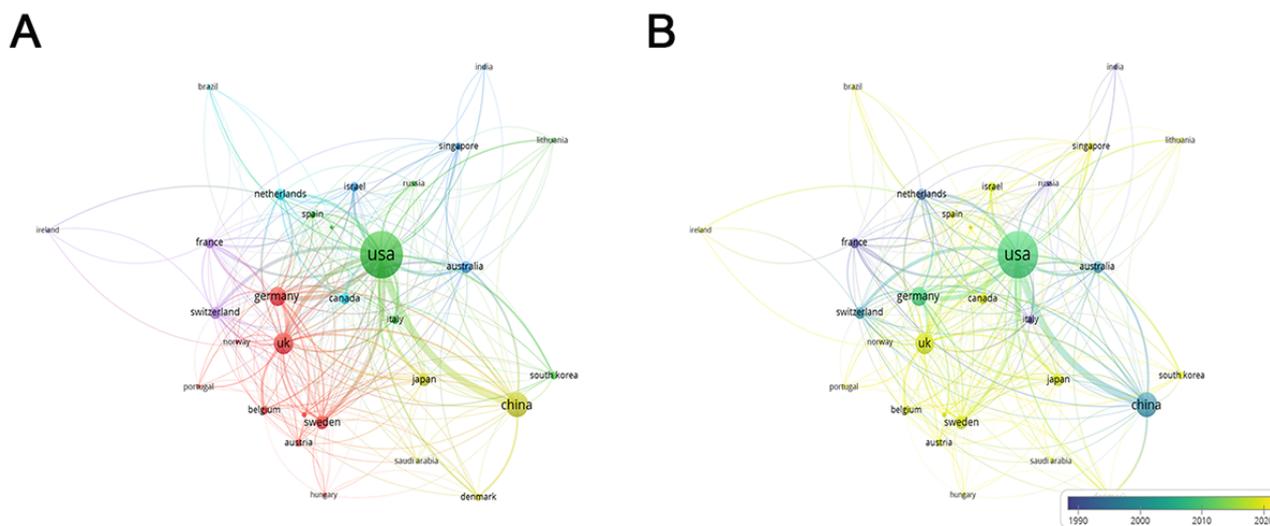
Table 3. Distribution by country.

Rank	Country	Documents, n (%)	Citations, n	Average citations	Total link strength
1	USA	1454 (58.4)	40337	27.7	549
2	China	414 (16.6)	6537	15.8	206
3	UK	305 (12.3)	9026	29.6	214
4	Germany	225 (9.0)	5384	23.9	162
5	Sweden	134 (5.4)	6973	52.0	88
6	Japan	108 (4.3)	1454	13.5	68
7	Switzerland	104 (4.2)	1802	17.3	72
8	Australia	99 (4.0)	2442	24.7	70
9	Canada	85 (3.4)	3177	37.4	53
10	Netherlands	85 (3.4)	1330	15.7	67

For the top 30 most prolific countries, which formed 6 clusters, a network map was created. There were active collaborations among these countries, especially between the United States and China (Figure 2A). From the dynamics and trends, it can be seen that the United States, China, Germany, Switzerland,

France, Australia, and the Netherlands have carried out studies on single-cell sequencing technology dating back to 2009. After 2015, many other investigators worldwide started to pay more attention to this field (Figure 2B).

Figure 2. Distribution of countries and regions. (A) The network map of countries (TOP30). (B) Dynamics and trends of countries/regions over years (TOP30).

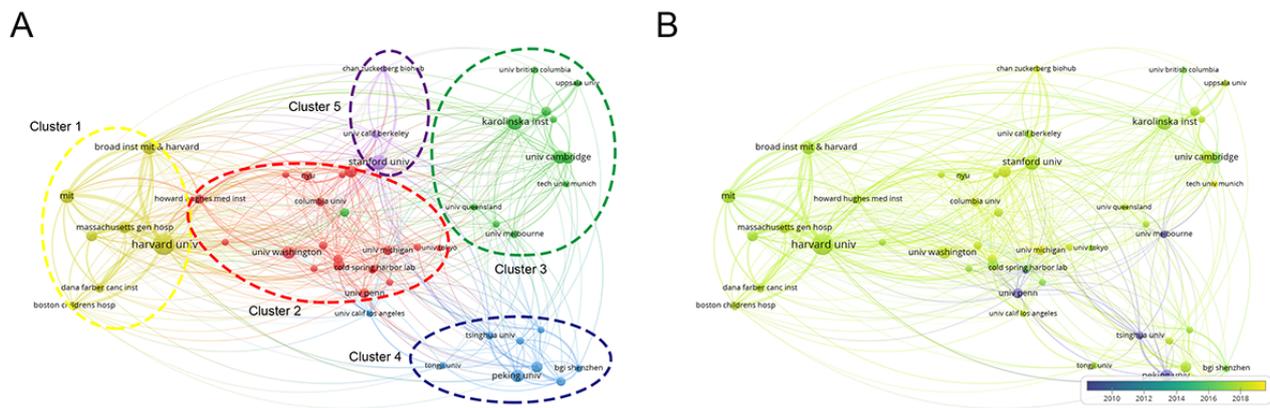


Universities and Institutions

A total of 1970 universities or institutions made contributions to single-cell sequencing technology; extensive cooperation network analysis was carried out among universities or institutions. In a network map and overlay visualization of institutions with the top 50 frequency, 5 clusters were formed (Figure 3A), and Harvard University (n=226, 9.1%, Cluster 1), Stanford University (n=114, 4.6%, Cluster 5), Karolinska Institute (n=111, 4.5%, Cluster 3), Peking University (n=81, 3.3%, Cluster 4), and the University of Washington (n=77, 3.1%, Cluster 2) were the biggest nodes in every cluster,

respectively. Cluster 1 was the biggest cluster, which contained 19 nodes that represented its related different universities or institutions, while Cluster 5 was the smallest, which included 3 nodes. The other clusters included 7 nodes (Cluster 2), 12 nodes (Cluster 3), and 9 nodes (Cluster 4). Furthermore, the results also showed that 4 universities contributed to primary and basic research on these topics earlier in time and started earlier on this field before 2010, which included Peking University, University of Melbourne, University of Penn, and the Tsinghua University. Since then, these studies have gradually gained popularity among other universities and institutes (Figure 3B).

Figure 3. Distribution of universities and institutions. (A) The network map of universities and institutions (TOP50). (B) Dynamics and trends of universities and institutions over time (TOP50).

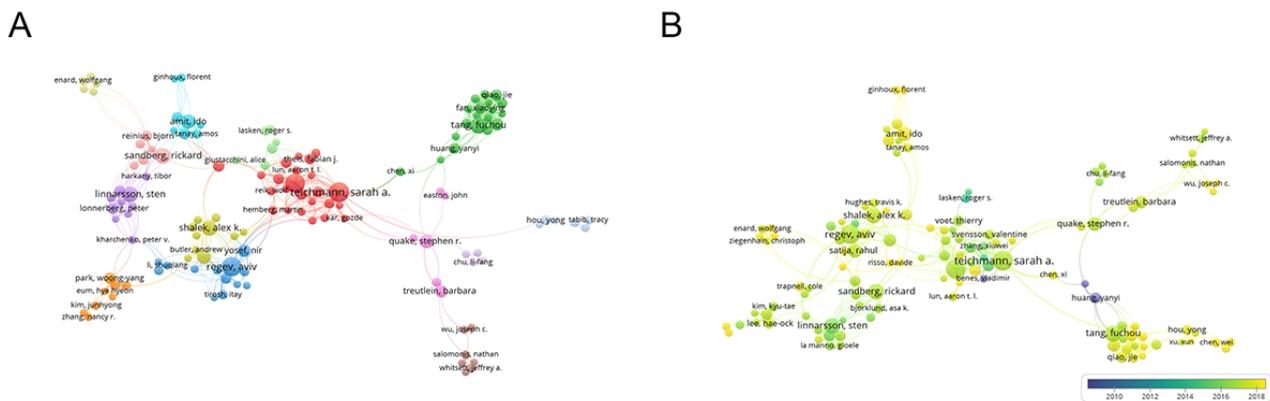


Authors

A total of 14,202 authors contributed to single-cell sequencing technology publications. The network map and overlay visualization of the top 200 cooperatively productive authors formed several clusters (Figure 4). Of these authors, 5 scientific teams contributed to most publications worldwide. The biggest 3 nodes were SA Teichmann (40/14,202, 1.6%) and JC Marioni (38/14,202, 1.5%), who had the biggest cluster and were from

the University of Cambridge. In addition, A Regev (38/14,202, 1.5%) had the second biggest cluster and was from Massachusetts Institute of Technology. The fourth node, FC Tang, represented a group from Peking University (26/14,202, 1.0%), and 2 authors, R Sandberg and S Linnarsson from Sweden contributed to 49 records in total. Moreover, I Amit, from Israel, published 21 records to date (Multimedia Appendix 1).

Figure 4. Distribution of authors. (A) The network map of authors (TOP200). (B) Dynamics and trends of authors over time (TOP200).



Journals

This study showed that a total of 495 journals published articles about single-cell sequencing technology. The results showed that the most productive journal was *Nature Communications* (127/14,202, 5.1%), the most cited journal was *Nature Methods* (55/14,202, 2.2%, 1764 citation times in total), and the journal with the highest average citation was *Nature Biotechnology* (28/14,202, 1.1%, average citation was 56.5 times). The top productive journals, top cited journals, and journals of top average citations are listed (n=20) in Multimedia Appendix 2.

Top 10 Citations of Included Records

Of the 2489 papers, the top 10 publications ranked by citation are listed in Table 4. The first paper was published in the *Journal of Computational Biology* by A Bankevich et al in 2012 [27] and reported a new genome algorithm and its applications to single-cell sequencing, with 5668 citations, which was much higher than that of the second paper (1307 citations) [28]. Of the 10 records, 4 were published in *Science*, and the topics included intratumoral heterogeneity in primary glioblastoma [29], metastatic melanoma [30], and cell types in the mouse cortex and hippocampus revealed by single-cell RNA-sequencing [31]. The latest paper [30], which was published in 2016, has been cited 652 times (from 2016 to 2019).

Table 4. Distribution by top 10-record co-citations.

Rank	Top-cited record	Number of citations	Title	Journal
1	A Bankevich, 2012 [27]	5668	SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing	<i>Journal of Computational Biology</i>
2	N Navin, 2011 [28]	1307	Tumour evolution inferred by single-cell sequencing	<i>Nature</i>
3	AP Patel, 2014 [29]	1168	Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma	<i>Science</i>
4	Y Peng, 2012	1039	IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth	<i>Bioinformatics</i>
5	A Zeisel, 2015 [31]	923	Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq	<i>Science</i>
6	C Trapnell, 2014	773	The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells	<i>Nature Biotechnology</i>
7	S Picelli, 2014	734	Full-length RNA-seq from single cells using Smart-seq2	<i>Nature Protocols</i>
8	I Tirosh, 2016 [30]	652	Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq	<i>Science</i>
9	DA Jaitin, 2014	615	Massively Parallel Single-Cell RNA-Seq for Marker-Free Decomposition of Tissues into Cell Types	<i>Science</i>
10	R Satija, 2015	605	Spatial reconstruction of single-cell gene expression data	<i>Nature Biotechnology</i>

Co-citation References

We used CiteSpace to visualize the co-citation network of references, which were divided into 11 co-citation clusters (Figure 5). The clusters were listed from 2008 to 2016: “transcriptional heterogeneity” (Cluster 10, n=4), “sequencing studies” (Cluster 0, n=53), “high-throughput spatial mapping” (Cluster 13, n=3), “distal lung epithelium” (Cluster 1, n=37), “studying clonal evolution” (Cluster 3, n=33), “early embryo” (Cluster 7, n=9), “assembling single-cell genome” (Cluster 2,

n=34), “high-throughput sequencing” (Cluster 4, n=27), “new genome assembly algorithm” (Cluster 8, n=7), “genomic sequencing” (Cluster 5, n=14), and “future medical applications” (Cluster 6, n=13).

The top 36 references with the strongest citation bursts are presented in Multimedia Appendix 3. The first reference with a citation burst appeared in 2011, and most of the bursts appeared between 2011 and 2016. Only 1 reference [32] appears with a burst in the last 3 years.

Figure 5. The co-citation network of references.

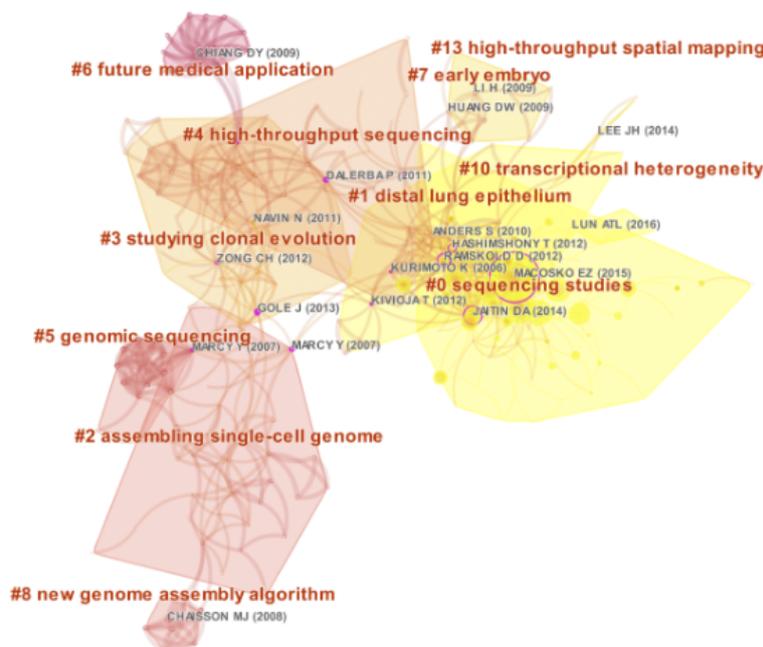


Table 5. Occurrences and total link strength of top 30 keywords.

Rank	Keyword	Occurrences	Total link strength
1	Gene-expression	857	851
2	RNA-seq	563	559
3	Heterogeneity	331	325
4	Stem-cells	235	234
5	Differentiation	231	231
6	Single cell	177	177
7	Reveals	170	170
8	Cancer	155	154
9	Transcriptome	135	135
10	Gene	133	132
11	Mouse	133	132
12	Genome	132	130
13	Identification	125	125
14	Seq	125	122
15	Evolution	115	115
16	Progenitors	93	93
17	Protein	92	92
18	Dynamics	90	90
19	Activation	83	83
20	Transcriptomics	80	78
21	Cells	77	77
22	Receptors	73	73
23	Noise	70	70
24	Mutations	69	69
25	Proliferation	67	67
26	Embryonic Stem-cells	66	66
27	Landscape	65	65
28	Neurons	65	65
29	Models	64	64
30	Diversity	63	63

Discussion

Principal Results

In this study, we summarized research collaborations, new development, and research trends of single-cell sequencing worldwide. In particular, bibliometric analysis offers a new possibility of visualizing current hotspots and trends in this field based on the information of published records [19,20]. We were able to analyze the most productive authors and institutions, by journal and by co-citation references, and we illustrated the different research groups of authors with collaboration networks in the past two decades. In total, 2489 records were identified, published in 495 journals by 14,202 authors from 1970 institutes in 61 countries. The United States had the highest absolute productivity ranking, followed by China and Japan (East Asia),

the United Kingdom, Germany, other European countries (Sweden, Switzerland, the Netherlands), and Australia and Canada. Taken together, these results indicated that single-cell sequencing technology developed rapidly and spread globally in the past 5 years (after 2015), which might be influenced by the amount of funding available for single-cell sequencing projects in each country. Li et al [33] investigated the distribution and development of single-cell sequencing in China, the United Kingdom, and the United States, focusing on government support, and concluded that there was a gap between high-income and low-income countries with respect to the amount of funding and number of projects. When compared with China, the amount of funding and the number of projects for single-cell sequencing in the United States and the United Kingdom had increased dramatically before 2016 (915 projects,

cumulative funding of US \$539 million). On the other hand, the later the starting time of foundation support for studies on single-cell sequencing, the less applications of this technology.

After the first record on single-cell sequencing, on the nuclear ribosomal genes of dinoflagellate (Dinophyceae), was published in 2001 [26], the number of publications was sporadic (1 to 3 annually until 2010). With the advancement of high-throughput sequencing and new computer algorithms, the development trend of publications changed from 2010 to 2019, showing a clear upward trend since 2014. However, as the development of single-cell sequencing technology has lasted for almost 20 years, several studies represent basic milestones of single-cell sequencing [2]. For example, the advancement of next-generation sequencing platforms in 2005 [34] offered a better approach for primary single-cell sequencing for isolated single cells. Subsequently, Tang et al [6] were the first to introduce RNA transcriptome sequencing of a mammalian cell in 2009, and Navin et al [28] published the first genomic sequencing of single human cells in 2011. Furthermore, several representative papers were published in the subsequent 3 years, including not only exome single-cell DNA sequencing [35] and its related single-cell DNA sequencing of sperm cells [36], neurons [37], circulating tumor cells [38], and the microbial tree [39], but also, the first single-cell sequencing of the RNA of immune cells [40] and of epigenomes [41]. Certainly, single-cell sequencing methods have supported advances in other technologies, including multiomics sequencing [8]; methods have led to novel discoveries in many scientific fields, such as tumor research [9], microbiology [10], neurology [11,12], reproduction [13], and immunology [14], which could explain the recent exponential increase in the number of publications.

While the United States published the most records worldwide (1454/2489, 58.42%), Sweden had the highest citation average (52.04 times). Citations differed from the number of documents, as did the cooperation of countries and regions: United States, China, United Kingdom, and Germany showed the strongest cooperation, with higher total link strengths. This was mainly because studies on single-cell sequencing technology were published earlier in these countries, which was indicated by the dynamics and trends of countries and regions over the years (Figure 3B). We believe that publications from Sweden had deeper impacts on the single-cell sequencing technology, although the field seems to be small and young enough that most of the findings can be entirely due to individual pioneering researchers. Furthermore, we also found that Karolinska Institutes, from Sweden, was one of the biggest nodes in its cluster of collaboration network (Figure 4A; Cluster 3), which might provide an indirect explanation. Similarly, the analysis of institutions and authors almost matched the distribution by countries and districts. Of the 1970 universities or institutions, the institutes from United States (Harvard University, Stanford University, University of Washington, and University of Penn), European universities (Karolinska Institute and Cambridge University), and Chinese institutes (Peking University and Tsinghua University), as well as the University of Melbourne in Australia, have taken leading places in single-cell sequencing technology, with steady collaborations in global groups.

Moreover, of the 14,202 authors who contributed to publications, the top 200 authors were situated in the collaboration networks that account for large proportions of research, mainly from the above-mentioned countries and institutes.

Single-cell sequencing avoids the drawbacks of bulk experiments, which only demonstrate average gene expression, and it can be used to study real differences and evolutionary changes in single cells [7]; therefore, it has gained the popularity and interest of the editors of many journals. Of the 495 journals, 20 journals contributed the most. The top journals were *Nature Communications* (most papers with 127 papers), *Nature Methods* (most cited with 1764 citations), and *Nature Biotechnology* (highest average citations with 56.46 citations per paper). Moreover, of included records, all the top 10 most cited papers were published in the top journals. In addition, the research topics of these top-cited records reported or introduced a new genome algorithm and application of single-cell sequencing [27-31], which together have been cited more than 600 times in the past of 5 years. In co-citation analysis, knowledge on “clonal evolution,” “early embryo,” “assembling single-cell genome,” “high-throughput sequencing,” and “new genome assembly algorithm” were the basis of single-cell sequencing research, and most references with citation bursts appeared between 2011 and 2016.

Of 4 clusters in visualization maps to explore key topics and future directions, Cluster 1 represented the basis of single-cell sequencing (next generation sequencing and single cell isolation methods researched earlier), while Cluster 2 represented the development of computer methods and algorithms. Cluster 3 showed current research hotspots of single-cell sequencing, such as applications in cancer research, with keywords including, circulating tumor-cells, origin, evolution, landscape, metastasis, resistance, and progression, and Cluster 4 showed that immunology will be the next hotspot of single-cell sequencing technology research, with keywords including dendritic cells, t-cells, macrophages, and diversity. In particular, in immunology, neuronal growth and the activation of regulators have been studied in recent years, rapidly gaining interest from 2015 to 2020.

Limitations

Our study had several limitations. First, we only searched records from the WoSCC and only included records in English, which might result in selection bias. However, the number of included records in our study was large enough to represent the current evidence landscape of single-cell sequencing technology [42]. Second, although the data had been manually standardized, bias might still exist due to the authors with the same name or the keywords of various expressions. Third, by using software, we might have overlooked some information, which may result in errors in data analysis, even if there are some reasonable and unavoidable differences on the same outcomes between CiteSpace and VOSviewer. This was the first study to perform bibliometric analysis on single-cell sequencing technology research, to identify collaboration networks among authors, countries, and institutions and illustrate developmental trends, current hotspots, and future directions in this field.

Recommendations for Future Work

Single-cell sequencing methods have addressed the drawbacks of averaged gene expression in bulk experiments and provide a possible approach to understanding biological diversity and rare cells. Over the past 10 years, the applications have had a broad impact in many fields, which emerged as delineating cell diversity, tracing cell lineages, classifying cell types, and profiling rare cells [2]. We only selected publications related to immunotherapy in the entire year of 2019, to form and label clusters (Multimedia Appendix 5). Eventually, the formed clusters illustrated that current hotspots in immunology involve at least 5 topics. Cancer immunotherapy is an exciting topic, for which single-cell sequencing methods have great potential for investigating the heterogeneity of intratumor immunity and other topics such as immune cell involvement of breast cancer therapy, role of immune cells in nervous system, formation and differentiation of immune cells, and functions of immune cells in the tumor microenvironment. Cancer therapy research to clarifying the functions of T-cells and macrophages, understand the antigenicity in intratumor heterogeneity, and study the role of clonal diversity in transformation, invasion, and evolution of chemoresistance would open new avenues for preventing cancer in the future. Cancer immunotherapy has revolutionized cancer treatment, and a detailed immune cell atlas at single-cell resolution could facilitate a comprehensive understanding the immunity landscape in tumor microenvironments [43] Moreover, future efforts in the development of single-cell sequencing technology will also emerge, which could focus on single-cell

multimodal omics, such as genomic DNA–messenger RNA sequencing [44,45], genome and transcriptome sequencing [46], and single-cell methylome and transcriptome sequencing [47]. Currently, the areas of particular relevance to immuno-oncology based upon single-cell sequencing technology focus on the ability to track individual specific cell clones through paired sequencing of their cell receptor genes and high-dimensional single-cell spatial analysis [48]. High-dimensional single-cell sequencing technologies are likely to generate clinically relevant biomarker signatures in immuno-oncology and may be able to guide clinical decision making.

Conclusions

The number of publications related to single-cell sequencing technology has increased dramatically since 2014. The United States led absolute productivity ranks by contributing approximately 60% of the total publications, followed by China, the United Kingdom, and Germany. Similarly, collaboration networks consisted mainly of institutes and authors from the above-mentioned countries. Moreover, in terms of the top 10 productive journals, the most cited journals, and journals with the highest average citation, there were 20 journals calculated in total. And most of them were considered as top-journals (impact factor > 10, Q1). Single-cell sequencing technology has made a large impact in various fields of biology, with a noticeable increase in publications annually in the last 5 years. We believe that the field of immunology might be a future research hotspot.

Acknowledgments

We thank MogoEdit, a professional copyediting agency, for helping to improve the quality of language in our manuscript. This work was supported by National Key Research and Development Program of China (2017YFC1308805), Industry-University-Research Innovation Fund in Ministry of Education of the People's Republic of China (2018A01013) and National Natural Science Foundation of China (grant 81871962).

Authors' Contributions

QW, KLY, ZZ, SW, and KWJ designed the study. KLY and ZW searched the database using the search strategy from JHT and LL. QW, ZW, KWJ, KLY, ZZ, and YJY screened the titles and abstracts to select articles. KLY, ZW, and YJY conducted analysis and mapping. All authors participated in the interpretation of the results. QW, KLY, SW, and KWJ wrote the manuscript with input from all authors. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Distribution of top 10 productive authors.

[\[DOCX File , 15 KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

Distribution by top journals and its citations.

[\[DOCX File , 17 KB-Multimedia Appendix 2\]](#)

Multimedia Appendix 3

Top 36 references with strongest citation bursts.

[\[PNG File , 438 KB-Multimedia Appendix 3\]](#)

Multimedia Appendix 4

Top 15 keywords with the strongest citation bursts.

[\[PNG File , 72 KB-Multimedia Appendix 4\]](#)

Multimedia Appendix 5

Clusters of immunotherapy in 2019.

[\[PNG File , 677 KB-Multimedia Appendix 5\]](#)

References

1. Bianconi E, Piovesan A, Facchin F, Beraudi A, Casadei R, Frabetti F, et al. An estimation of the number of cells in the human body. *Ann Hum Biol* 2013;40(6):463-471. [doi: [10.3109/03014460.2013.807878](#)] [Medline: [23829164](#)]
2. Wang Y, Navin NE. Advances and applications of single-cell sequencing technologies. *Mol Cell* 2015;58(4):598-609 [FREE Full text] [doi: [10.1016/j.molcel.2015.05.005](#)] [Medline: [26000845](#)]
3. Tang X, Huang Y, Lei J, Luo H, Zhu X. The single-cell sequencing: new developments and medical applications. *Cell Biosci* 2019;9:53 [FREE Full text] [doi: [10.1186/s13578-019-0314-y](#)] [Medline: [31391919](#)]
4. Wen L, Tang F. Boosting the power of single-cell analysis. *Nat Biotechnol* 2018;36(5):408-409. [doi: [10.1038/nbt.4131](#)] [Medline: [29734314](#)]
5. Stegle O, Teichmann SA, Marioni JC. Computational and analytical challenges in single-cell transcriptomics. *Nat Rev Genet* 2015;16(3):133-145. [doi: [10.1038/nrg3833](#)] [Medline: [25628217](#)]
6. Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, et al. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods* 2009;6(5):377-382. [doi: [10.1038/nmeth.1315](#)] [Medline: [19349980](#)]
7. Dal Molin A, Di Camillo B. How to design a single-cell RNA-sequencing experiment: pitfalls, challenges and perspectives. *Brief Bioinform* 2019;20(4):1384-1394. [doi: [10.1093/bib/bby007](#)] [Medline: [29394315](#)]
8. Li J, Lin H, Hou R, Shen J, Li X, Xing J, et al. Multi-omics study in monozygotic twins confirm the contribution of de novo mutation to psoriasis. *J Autoimmun* 2020;106:102349 [FREE Full text] [doi: [10.1016/j.jaut.2019.102349](#)] [Medline: [31629629](#)]
9. Zhang Q, He Y, Luo N, Patel SJ, Han Y, Gao R, et al. Landscape and dynamics of single immune cells in hepatocellular carcinoma. *Cell* 2019 Oct 31;179(4):829-845.e20 [FREE Full text] [doi: [10.1016/j.cell.2019.10.003](#)] [Medline: [31675496](#)]
10. Wang Y, Niu Q, Zhang X, Liu L, Wang Y, Chen Y, et al. Exploring the effects of operational mode and microbial interactions on bacterial community assembly in a one-stage partial-nitritation anammox reactor using integrated multiomics. *Microbiome* 2019;7(1):122 [FREE Full text] [doi: [10.1186/s40168-019-0730-6](#)] [Medline: [31462278](#)]
11. Lake BB, Chen S, Sos BC, Fan J, Kaeser GE, Yung YC, et al. Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain. *Nat Biotechnol* 2018;36(1):70-80 [FREE Full text] [doi: [10.1038/nbt.4038](#)] [Medline: [29227469](#)]
12. Rosenberg AB, Roco CM, Muscat RA, Kuchina A, Sample P, Yao Z, et al. Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science* 2018;360(6385):176-182 [FREE Full text] [doi: [10.1126/science.aam8999](#)] [Medline: [29545511](#)]
13. Guo F, Li L, Li J, Wu X, Hu B, Zhu P, et al. Single-cell multi-omics sequencing of mouse early embryos and embryonic stem cells. *Cell Res* 2017;27(8):967-988 [FREE Full text] [doi: [10.1038/cr.2017.82](#)] [Medline: [28621329](#)]
14. Casasent AK, Schalck A, Gao R, Sei E, Long A, Pangburn W, et al. Multiclonal invasion in breast tumors identified by topographic single cell sequencing. *Cell* 2018;172(1-2):205-217 [FREE Full text] [doi: [10.1016/j.cell.2017.12.007](#)] [Medline: [29307488](#)]
15. Suvà ML, Tirosch I. Single-cell RNA sequencing in cancer: lessons learned and emerging challenges. *Mol Cell* 2019;75(1):7-12 [FREE Full text] [doi: [10.1016/j.molcel.2019.05.003](#)] [Medline: [31299208](#)]
16. Zhang X, Liu L. Applications of single cell RNA sequencing to research of stem cells. *World J Stem Cells* 2019;11(10):722-728 [FREE Full text] [doi: [10.4252/wjsc.v11.i10.722](#)] [Medline: [31692946](#)]
17. Ren X, Kang B, Zhang Z. Understanding tumor ecosystems by single-cell sequencing: promises and limitations. *Genome Biol* 2018;19(1):211 [FREE Full text] [doi: [10.1186/s13059-018-1593-z](#)] [Medline: [30509292](#)]
18. Kiselev VY, Andrews TS, Hemberg M. Challenges in unsupervised clustering of single-cell RNA-seq data. *Nat Rev Genet* 2019 May;20(5):273-282. [doi: [10.1038/s41576-018-0088-9](#)] [Medline: [30617341](#)]
19. Agarwal A, Durairajanayagam D, Tatagari S, Esteves SC, Harlev A, Henkel R, et al. Bibliometrics: tracking research impact by selecting the appropriate metrics. *Asian J Androl* 2016;18(2):296-309 [FREE Full text] [doi: [10.4103/1008-682X.171582](#)] [Medline: [26806079](#)]
20. Keathley-Herring H, Van Aken E, Gonzalez-Aleu F, Deschamps F, Letens G, Orlandini PC. Assessing the maturity of a research area: bibliometric review and proposed framework. *Scientometrics* 2016 Aug 13;109(2):927-951. [doi: [10.1007/s11192-016-2096-x](#)]
21. Chen C. CiteSpace II: detecting and visualizing emerging trends and transient patterns in scientific literature. *J Am Soc Inf Sci* 2006 Feb 01;57(3):359-377. [doi: [10.1002/asi.20317](#)]

22. van Eck NJ, Waltman L. Citation-based clustering of publications using CitNetExplorer and VOSviewer. *Scientometrics* 2017;111(2):1053-1070 [[FREE Full text](#)] [doi: [10.1007/s11192-017-2300-7](https://doi.org/10.1007/s11192-017-2300-7)] [Medline: [28490825](#)]
23. Devos P, Ménard J. Trends in worldwide research in hypertension over the period 1999-2018: a bibliometric study. *Hypertension* 2020;76(5):1649-1655 [[FREE Full text](#)] [doi: [10.1161/HYPERTENSIONAHA.120.15711](https://doi.org/10.1161/HYPERTENSIONAHA.120.15711)] [Medline: [32862706](#)]
24. van Eck NJ, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics* 2010 Aug;84(2):523-538 [[FREE Full text](#)] [doi: [10.1007/s11192-009-0146-3](https://doi.org/10.1007/s11192-009-0146-3)] [Medline: [20585380](#)]
25. van Eck NJ, Waltman L. Text mining and visualization using VOSviewer. *Arxiv* 2011 Sep 09 [[FREE Full text](#)]
26. Ruiz Sebastián C, O’Ryan C. Single-cell sequencing of dinoflagellate (Dinophyceae) nuclear ribosomal genes. *Mol Ecol Notes* 2001;1(4):329-331. [doi: [10.1046/j.1471-8278.2001.00084.x](https://doi.org/10.1046/j.1471-8278.2001.00084.x)]
27. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012;19(5):455-477 [[FREE Full text](#)] [doi: [10.1089/cmb.2012.0021](https://doi.org/10.1089/cmb.2012.0021)] [Medline: [22506599](#)]
28. Navin N, Kendall J, Troge J, Andrews P, Rodgers L, McIndoo J, et al. Tumour evolution inferred by single-cell sequencing. *Nature* 2011;472(7341):90-94 [[FREE Full text](#)] [doi: [10.1038/nature09807](https://doi.org/10.1038/nature09807)] [Medline: [21399628](#)]
29. Patel AP, Tirosh I, Trombetta JJ, Shalek AK, Gillespie SM, Wakimoto H, et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* 2014;344(6190):1396-1401 [[FREE Full text](#)] [doi: [10.1126/science.1254257](https://doi.org/10.1126/science.1254257)] [Medline: [24925914](#)]
30. Tirosh I, Izar B, Prakadan SM, Wadsworth MH, Treacy D, Trombetta JJ, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* 2016;352(6282):189-196 [[FREE Full text](#)] [doi: [10.1126/science.aad0501](https://doi.org/10.1126/science.aad0501)] [Medline: [27124452](#)]
31. Zeisel A, Muñoz-Manchado AB, Codeluppi S, Lönnerberg P, La Manno G, Juréus A, et al. Brain structure, cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* 2015;347(6226):1138-1142 [[FREE Full text](#)] [doi: [10.1126/science.aaa1934](https://doi.org/10.1126/science.aaa1934)] [Medline: [25700174](#)]
32. Kolodziejczyk AA, Kim JK, Svensson V, Marioni JC, Teichmann SA. The technology and biology of single-cell RNA sequencing. *Mol Cell* 2015;58(4):610-620 [[FREE Full text](#)] [doi: [10.1016/j.molcel.2015.04.005](https://doi.org/10.1016/j.molcel.2015.04.005)] [Medline: [26000846](#)]
33. Li A, Tang X. Foundation support for studies on single-cell sequencingits application in USA, Britain and China. *Chin J Med Lib Info Sci* 2017;26(7):44-51. [doi: [10.3969/j.issn.1671-3982.2017.09.008](https://doi.org/10.3969/j.issn.1671-3982.2017.09.008)]
34. Metzker ML. Emerging technologies in DNA sequencing. *Genome Res* 2005;15(12):1767-1776 [[FREE Full text](#)] [doi: [10.1101/gr.3770505](https://doi.org/10.1101/gr.3770505)] [Medline: [16339375](#)]
35. Xu X, Hou Y, Yin X, Bao L, Tang A, Song L, et al. Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. *Cell* 2012;148(5):886-895 [[FREE Full text](#)] [doi: [10.1016/j.cell.2012.02.025](https://doi.org/10.1016/j.cell.2012.02.025)] [Medline: [22385958](#)]
36. Wang J, Fan HC, Behr B, Quake SR. Genome-wide single-cell analysis of recombination activity and de novo mutation rates in human sperm. *Cell* 2012;150(2):402-412 [[FREE Full text](#)] [doi: [10.1016/j.cell.2012.06.030](https://doi.org/10.1016/j.cell.2012.06.030)] [Medline: [22817899](#)]
37. Evrony GD, Cai X, Lee E, Hills LB, Elhosary PC, Lehmann HS, et al. Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. *Cell* 2012;151(3):483-496 [[FREE Full text](#)] [doi: [10.1016/j.cell.2012.09.035](https://doi.org/10.1016/j.cell.2012.09.035)] [Medline: [23101622](#)]
38. Ramsköld D, Luo S, Wang Y, Li R, Deng Q, Faridani OR, et al. Full-length mRNA-seq from single-cell levels of RNA and individual circulating tumor cells. *Nat Biotechnol* 2012;30(8):777-782 [[FREE Full text](#)] [doi: [10.1038/nbt.2282](https://doi.org/10.1038/nbt.2282)] [Medline: [22820318](#)]
39. Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, Cheng J, et al. Insights into the phylogeny and coding potential of microbial dark matter. *Nature* 2013;499(7459):431-437. [doi: [10.1038/nature12352](https://doi.org/10.1038/nature12352)] [Medline: [23851394](#)]
40. Shalek AK, Satija R, Adiconis X, Gertner RS, Gaublomme JT, Raychowdhury R, et al. Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. *Nature* 2013;498(7453):236-240 [[FREE Full text](#)] [doi: [10.1038/nature12172](https://doi.org/10.1038/nature12172)] [Medline: [23685454](#)]
41. Nagano T, Lubling Y, Stevens TJ, Schoenfelder S, Yaffe E, Dean W, et al. Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* 2013;502(7469):59-64 [[FREE Full text](#)] [doi: [10.1038/nature12593](https://doi.org/10.1038/nature12593)] [Medline: [24067610](#)]
42. Peng C, He M, Cutrona SL, Kiefe CI, Liu F, Wang Z. Theme trends and knowledge structure on mobile health apps: bibliometric analysis. *JMIR Mhealth Uhealth* 2020;8(7):e18212 [[FREE Full text](#)] [doi: [10.2196/18212](https://doi.org/10.2196/18212)] [Medline: [32716312](#)]
43. Zheng Y, Chen Z, Han Y, Han L, Zou X, Zhou B, et al. Immune suppressive landscape in the human esophageal squamous cell carcinoma microenvironment. *Nat Commun* 2020;11(1):6268 [[FREE Full text](#)] [doi: [10.1038/s41467-020-20019-0](https://doi.org/10.1038/s41467-020-20019-0)] [Medline: [33293583](#)]
44. Zhu C, Preissl S, Ren B. Single-cell multimodal omics: the power of many. *Nat Methods* 2020;17(1):11-14. [doi: [10.1038/s41592-019-0691-5](https://doi.org/10.1038/s41592-019-0691-5)] [Medline: [31907462](#)]
45. Dey SS, Kester L, Spanjaard B, Bienko M, van Oudenaarden A. Integrated genome and transcriptome sequencing of the same cell. *Nat Biotechnol* 2015;33(3):285-289 [[FREE Full text](#)] [doi: [10.1038/nbt.3129](https://doi.org/10.1038/nbt.3129)] [Medline: [25599178](#)]
46. Macaulay IC, Haerty W, Kumar P, Li YI, Hu TX, Teng MJ, et al. *Nat Methods* 2015;12(6):519-522. [doi: [10.1038/nmeth.3370](https://doi.org/10.1038/nmeth.3370)] [Medline: [25915121](#)]

47. Angermueller C, Clark SJ, Lee HJ, Macaulay IC, Teng MJ, Hu TX, et al. Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nat Methods* 2016;13(3):229-232 [FREE Full text] [doi: [10.1038/nmeth.3728](https://doi.org/10.1038/nmeth.3728)] [Medline: [26752769](https://pubmed.ncbi.nlm.nih.gov/26752769/)]
48. Gohil SH, Iorgulescu JB, Braun DA, Keskin DB, Livak KJ. Applying high-dimensional single-cell technologies to the analysis of cancer immunotherapy. *Nat Rev Clin Oncol* 2021 Apr;18(4):244-256. [doi: [10.1038/s41571-020-00449-x](https://doi.org/10.1038/s41571-020-00449-x)] [Medline: [33277626](https://pubmed.ncbi.nlm.nih.gov/33277626/)]

Abbreviations

WoSCC: Web of Science Core Collection

Edited by R Kukafka; submitted 16.11.20; peer-reviewed by X Wang, L Ge; comments to author 09.12.20; revised version received 22.12.20; accepted 13.05.21; published 10.08.21

Please cite as:

Wang Q, Yang KL, Zhang Z, Wang Z, Li C, Li L, Tian JH, Ye YJ, Wang S, Jiang KW

Characterization of Global Research Trends and Prospects on Single-Cell Sequencing Technology: Bibliometric Analysis

J Med Internet Res 2021;23(8):e25789

URL: <https://www.jmir.org/2021/8/e25789>

doi: [10.2196/25789](https://doi.org/10.2196/25789)

PMID: [34014832](https://pubmed.ncbi.nlm.nih.gov/34014832/)

©Quan Wang, Ke-Lu Yang, Zhen Zhang, Zhu Wang, Chen Li, Lun Li, Jin-Hui Tian, Ying-Jiang Ye, Shan Wang, Ke-Wei Jiang. Originally published in the *Journal of Medical Internet Research* (<https://www.jmir.org>), 10.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the *Journal of Medical Internet Research*, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.