Abstract. – OBJECTIVE: Breast cancer (BC) is a prevalent cancer all over the world. We conducted a bibliometric study to analyze global scientific results over the past 10 years, including the hotspots and frontiers of biomarker research in BC.

MATERIALS AND METHODS: From 2011 to 2020 a literature research from the Web of Science Core Collection (WoSCC) was performed. VOSviewer was applied to analyze and visualize the frontiers and hotspots related to biomarker research in BC.

RESULTS: 13,680 papers were retrieved. There was an increasing number of annual publications (Np) related to biomarkers in BC during the past decade. The United States (US) published the greatest number of papers, which had the highest number of citations (Nc) and ranked first in terms of H-index. PLoS One and the University of Texas System were the most productive journals and affiliations, respectively. In 2014, Chetan Bettegowda published a paper with the world’s highest global citation score (GCS). In recent years, keywords such as “expression”, “microRNA”, and “cell” have appeared most frequently. In addition, research related to COVID-19 in this field has become a hot topic in recent years. This bibliometric study found an increasing trend in publications related to biomarkers in British Columbia and the US was found to be an influential producer in this field.

CONCLUSIONS: In the past decade, most research has focused on basic and clinical studies, of which microRNAs (miRNAs) and circulating tumor DNAs (ctDNAs) associated with the inhibition and attenuation of BC have become the focus of recent research.

Key Words: Breast cancer (BC), Biomarkers, Bibliometrics, VOSviewer.

Introduction

Breast cancer (BC) is a very common malignancy among women1,2, accounting for 11.6% of total cancer deaths3,4. Male breast cancer is rare5. It exerts a significant burden on women’s physical and mental health and has attracted extensive social attention. Early diagnosis, effective monitoring, and accurate treatment are the keys to improving prognosis and quality of life6. In the routine diagnosis and treatment of breast diseases, breast palpation by a clinician is still one of the main forms of BC screening. This method is simple, convenient, and economical, but is associated with strong subjectivity and a certain omission rate7. The detection rate of early BC with a diameter of less than 1 cm is low. Although the treatment of BC has been greatly improved in recent years, the prognosis remains poor due to individual patient differences8. Diagnosis of BC via imaging may also be subjective and dependent on the experience and interpretation of the attending radiologist. In addition, although different modality imaging methods can reflect lesion information from different levels, the overall specificity needs to be improved9,10. Therefore, it is necessary to identify sensitive and accurate biomarkers to better diagnose and predict the survival and prognosis of BC patients. With the development of more auxiliary diagnostic techniques, especially the development of molecular biology, rapid progress has been made in the early diagnosis of BC11. The occurrence of BC is believed to be an interaction and multi-stage evolution of genetic and environmental factors. Understanding the progression of precancerous breast lesions and the genetic and epigenetic variations in the evolution of BC, as well as the effects of environmental fac-
tors at the molecular level, is crucial for improving patient outcomes. The identification of meaningful biomarkers may provide an important reference for high-risk BC patients and early diagnosis and treatment. In recent years, the development of molecular biology, genetic theories, and high-throughput omics technology, has enabled the investigation of biomarkers related to the early diagnosis of BC.

During these years, the results of bibliometric analyses were used in orthopedics, gynecology, and other medical fields\textsuperscript{12-14}, providing a guide for further research on disease prevention and treatment\textsuperscript{15,16}. However, there is a paucity of bibliometric studies examining biomarkers in BC. Therefore, this study systematically analyzed the research of biomarkers in BC, so as to assess frontiers and hotspots in this field.

### Materials and Methods

#### Data Collection and Search Methods

The bibliometric study was carried out in Web of Science Core Collection (WoSCC). Because of the rapid updating of the database, publication retrieval was carried out on the same day (January 17, 2021) in order to avoid deviations. Pieces of literature published from 2011 to 2020 were evaluated. The following search terms were used: (TS(Topic) = biomarkers) AND (TS = mammary cancer OR TS = breast cancer). Among multiple document types, only reviews and articles were included. Finally, 13,680 papers were involved in this research. The filtering process is presented in Figure 1.

#### Data Collection

The raw data, including year of publication, H-index, number of papers as well as citations, country/region, affiliations, authors, journals, references, and keywords, were extracted from the WoSCC database. The data were then imported into VOSviewer to perform the further analysis.

### Bibliometric Analysis

Bibliometric indicators, including the number of papers (Np) and the number of citations (Nc), were used to represent bibliographic material. In general, the Np is used to measure productivity and the Nc is used to indicate impact, as these are the two main indices for assessing the level of research. More recently, the H-index has increasingly been used to assess a researcher’s academic contributions and predict future scientific achievements\textsuperscript{17,18}. The H-index combines productivity and impact by finding a threshold that connects the Np and the Nc. If a researcher publishes H papers and each paper is cited at least H times, she or he will have an H-index\textsuperscript{19}. Particularly, although originally designed to assess individual academic achievement, the H-index can be extended to describe the publication output of a country or region, an institution, or a journal\textsuperscript{20}. In addition, the impact factor (IF) obtained from Journal Citation Reports (JCR) is regarded as a main indicator to measure the quality and impact of medical journals\textsuperscript{21}. To further explain the variation of annual literature volume, the fitting polynomial model was used to predict annual Np. The variable f(x) represents the number of studies per year and x represents the year of publication. In addition, a bibliometric map was constructed using VOSviewer software to obtain more comprehensive results according to co-occurrence and co-citation\textsuperscript{22}. Co-citation is defined when both items are referenced by a third item. The co-occurrence measure for keywords is the keyword that appears most frequently in the same document\textsuperscript{23}.

![Figure 1. A flowchart showing the literature screening process.](image-url)
Results

A Summary of Papers on Biomarkers in BC

13,680 papers, including 10,554 articles and 3,126 reviews, were retrieved according to our search strategy, indicating that the research in this field has developed rapidly.

Annual Trend in the Number of Publications

Figure 2A presents the Np of each year associated with biomarkers in BC. Overall, during the past decade, Np increased from 798 in 2011 to 1,908 in 2020, with Np peaking in 2020. Since 2011, the annual Np of the US and England has remained stable, while that of China has increased rapidly. Figure 2B shows a fit curve for the trend of each year in the number of print publications. The Np per year was significantly correlated with the year. According to Figure 2B, the correlation coefficient R² reached 0.9638. Taken together, these findings indicated that the study of biomarkers in BC has become a topic of interest and got into a phase of rapid growth.

Performance of Countries/Regions on Global Output

The top 10 highest-producing countries/regions were ranked by Np for all authors (Table I). The US had the most Np (4,767/34.85%), followed by China (3,047/22.27%) and England (948/6.93%). Nc of the US was 180,115, followed by China [66,433] and England [41,724]. Besides, the US processed the highest H-index [173], almost double that of China [99]. Compared with Italy and Germany, Canada and France had slightly lower Np and Nc, but remarkably higher H-index. The regional distribution and national co-occurrence

<table>
<thead>
<tr>
<th>Rank</th>
<th>Country/region</th>
<th>Np</th>
<th>% (out of 13,680 publications)</th>
<th>Nc</th>
<th>H-index</th>
<th>Average per item</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>USA</td>
<td>4,767</td>
<td>34.85</td>
<td>180,115</td>
<td>173</td>
<td>38.88</td>
</tr>
<tr>
<td>2</td>
<td>China</td>
<td>3,047</td>
<td>22.27</td>
<td>66,433</td>
<td>99</td>
<td>22.43</td>
</tr>
<tr>
<td>3</td>
<td>England</td>
<td>948</td>
<td>6.93</td>
<td>41,724</td>
<td>93</td>
<td>44.57</td>
</tr>
<tr>
<td>4</td>
<td>Italy</td>
<td>923</td>
<td>6.75</td>
<td>31,893</td>
<td>71</td>
<td>35.09</td>
</tr>
<tr>
<td>5</td>
<td>Germany</td>
<td>854</td>
<td>6.24</td>
<td>31,873</td>
<td>71</td>
<td>35.07</td>
</tr>
<tr>
<td>6</td>
<td>Canada</td>
<td>707</td>
<td>5.17</td>
<td>26,179</td>
<td>76</td>
<td>37.48</td>
</tr>
<tr>
<td>7</td>
<td>France</td>
<td>637</td>
<td>4.66</td>
<td>25,147</td>
<td>75</td>
<td>39.98</td>
</tr>
<tr>
<td>8</td>
<td>Spain</td>
<td>482</td>
<td>3.52</td>
<td>18,503</td>
<td>60</td>
<td>38.76</td>
</tr>
<tr>
<td>9</td>
<td>Japan</td>
<td>454</td>
<td>3.32</td>
<td>11,842</td>
<td>52</td>
<td>26.33</td>
</tr>
<tr>
<td>10</td>
<td>Australia</td>
<td>422</td>
<td>3.08</td>
<td>22,629</td>
<td>65</td>
<td>54.07</td>
</tr>
</tbody>
</table>
of the number of documents are shown in Figures 3A and 3B, respectively.

**Affiliations Performance Analysis**

Table II shows the top 10 affiliations with the largest number of publications associated with biomarkers in BC. The University of Texas ranked first in terms of Np [578], and Harvard University [542] and the University of California [454] were the second and the third. The University of Texas ranked first for Nc [32,414] and H-index [86]. Compared with Unicancer, Institut National de la Sante et de la Recherche Medicale Inserm, and the University of Toronto, the Np of the University of London in England was relatively low, but its H-index was relatively high. Most affiliations were based in the US, followed by France. The co-occurrence of affiliations was shown in Figure 4.
Table II. The top 10 most productive affiliations.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Affiliation</th>
<th>Np</th>
<th>Nc</th>
<th>H-index</th>
<th>Average per item</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>University of Texas System</td>
<td>578</td>
<td>32,414</td>
<td>86</td>
<td>56.49</td>
<td>USA</td>
</tr>
<tr>
<td>2</td>
<td>Harvard University</td>
<td>542</td>
<td>29,696</td>
<td>81</td>
<td>55.23</td>
<td>USA</td>
</tr>
<tr>
<td>3</td>
<td>University of California System</td>
<td>454</td>
<td>20,512</td>
<td>72</td>
<td>45.51</td>
<td>USA</td>
</tr>
<tr>
<td>4</td>
<td>Utmd Anderson Cancer Center</td>
<td>421</td>
<td>26,288</td>
<td>79</td>
<td>62.9</td>
<td>USA</td>
</tr>
<tr>
<td>5</td>
<td>National Institutes of Health NIH USA</td>
<td>346</td>
<td>17,822</td>
<td>63</td>
<td>51.83</td>
<td>USA</td>
</tr>
<tr>
<td>6</td>
<td>Unicancer</td>
<td>329</td>
<td>13,836</td>
<td>55</td>
<td>42.55</td>
<td>France</td>
</tr>
<tr>
<td>7</td>
<td>Institut National de la Sante et de la Recherche Medicale Inserm</td>
<td>318</td>
<td>12,030</td>
<td>55</td>
<td>38.13</td>
<td>France</td>
</tr>
<tr>
<td>8</td>
<td>University of Toronto</td>
<td>306</td>
<td>10,819</td>
<td>52</td>
<td>35.69</td>
<td>Canada</td>
</tr>
<tr>
<td>9</td>
<td>University of London</td>
<td>297</td>
<td>14,180</td>
<td>61</td>
<td>48.05</td>
<td>England</td>
</tr>
<tr>
<td>10</td>
<td>NIH National Cancer Institute NCI</td>
<td>261</td>
<td>14,102</td>
<td>55</td>
<td>54.38</td>
<td>USA</td>
</tr>
</tbody>
</table>

Figure 4. Map of affiliations with publications related to biomarkers in breast cancer. A, The 130 affiliations that occurred more than 45 times were divided into 6 clusters by different colors. Cluster 1: red; cluster 2: green; cluster 3: dark blue cluster 4: yellow; cluster 5: purple; cluster 6: light blue. The size of the nodes represents the frequency of occurrences. B, Visualization of the affiliations according to the average year of publication (APY). Affiliations in yellow started research later than affiliations in blue.
Authors Performance Analysis

The top 10 prolific authors (Table III) published a total of 941 papers, accounting for 6.88% of all publications analyzed. Zhang Y was in the first place in this research field, Li J from Caltech in the US and Wang J from China ranked second and third. Wang J had an obviously high Nc. Besides, great majority of the top 10 authors came from the US or China. The co-occurrence of authors is shown in Figure 5.

Journals Performance Analysis

PLoS One published the most papers concerning biomarkers in BC (410 publications, IF: 3.24), followed by Oncotarget (358 publications, removed) and Cancer Epidemiology Biomarkers Prevention (297 publications, IF: 4.254). Among these publications, the top 10 journals published 18% of all publications [2,469]. In the 10 journals, except Oncology Letters, BMC Cancer, and Oncotarget, the remaining seven journals have high IF (IF > 3.0). Obviously, Clinical Cancer Research and Cancers have higher IFs, while PLoS One, Clinical Cancer Research, and Cancer Epidemiology have higher H-indexes (Table IV). The co-occurrence of journals is shown in Figure 6.

Analysis of Global Citations (GCs)

Figure 7 shows the annual GCs of the top 10 papers. In 2014, the article written by Bettegowda C ranked first with a GC of 2,039. In this paper, the authors concluded that circulating tumor DNA (ctDNA) were sensitive, specific, and reliable biomarkers. For individuals with central nervous system (CNS) tumors, alternative strategies may be needed to detect cell-free tumor-derived DNA at clinically meaningful levels. In addition, the report by Schwarzenbach et al summarized the potential uses of circulating nucleic acid in cancer, with special attention on the clinical application of acellular nucleic acid as a blood biomarker. Dawson et al proposed that ctDNA was an inherently specific and informative biomarker for BC. Karlsen et al summarized the significance and role of microRNA (miRNA) imbalance in cancer diagnosis, monitoring, and treatment. Gentles et al presented a meta-analysis of pan-carcinomatous resources and expression characteristics from approximately 18,000 tumors, and they identified the FOXM1 as a main predictor of poor prognosis. Using CIBERSORT, a computational method for inferring white blood cell expression in tumor transcriptome, a complex association

Table III. The top 10 authors with the most publications.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Author</th>
<th>Country</th>
<th>Np</th>
<th>Nc</th>
<th>H-index</th>
<th>Average per item</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Zhang Y</td>
<td>China</td>
<td>115</td>
<td>2,450</td>
<td>24</td>
<td>21.44</td>
</tr>
<tr>
<td>2</td>
<td>Li J</td>
<td>USA</td>
<td>114</td>
<td>2,786</td>
<td>29</td>
<td>24.5</td>
</tr>
<tr>
<td>3</td>
<td>Wang J</td>
<td>China</td>
<td>110</td>
<td>2,999</td>
<td>31</td>
<td>27.45</td>
</tr>
<tr>
<td>4</td>
<td>Wang Y</td>
<td>USA</td>
<td>107</td>
<td>2,355</td>
<td>27</td>
<td>22.18</td>
</tr>
<tr>
<td>5</td>
<td>Zhang J</td>
<td>China</td>
<td>92</td>
<td>2,435</td>
<td>28</td>
<td>26.53</td>
</tr>
<tr>
<td>6</td>
<td>Li Y</td>
<td>China</td>
<td>91</td>
<td>1,602</td>
<td>22</td>
<td>17.77</td>
</tr>
<tr>
<td>7</td>
<td>Li L</td>
<td>USA</td>
<td>86</td>
<td>3,097</td>
<td>26</td>
<td>36.05</td>
</tr>
<tr>
<td>8</td>
<td>Liu Y</td>
<td>Belgium</td>
<td>81</td>
<td>2,609</td>
<td>25</td>
<td>32.21</td>
</tr>
<tr>
<td>9</td>
<td>Wang L</td>
<td>USA</td>
<td>80</td>
<td>1,795</td>
<td>21</td>
<td>22.5</td>
</tr>
<tr>
<td>10</td>
<td>Zhang L</td>
<td>USA</td>
<td>65</td>
<td>1,324</td>
<td>20</td>
<td>20.42</td>
</tr>
</tbody>
</table>

Table IV. The top 10 most active journals.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Journal</th>
<th>Np</th>
<th>H-index</th>
<th>Nc</th>
<th>IF</th>
<th>Average per item</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PLoS One</td>
<td>410</td>
<td>51</td>
<td>11,896</td>
<td>3.24</td>
<td>29.13</td>
</tr>
<tr>
<td>2</td>
<td>Oncotarget</td>
<td>358</td>
<td>48</td>
<td>9,427</td>
<td>–</td>
<td>26.48</td>
</tr>
<tr>
<td>3</td>
<td>Cancer Epidemiology Biomarkers Prevention</td>
<td>297</td>
<td>58</td>
<td>11,548</td>
<td>4.254</td>
<td>39.08</td>
</tr>
<tr>
<td>4</td>
<td>Breast Cancer Research and Treatment</td>
<td>279</td>
<td>44</td>
<td>7,693</td>
<td>4.872</td>
<td>27.74</td>
</tr>
<tr>
<td>5</td>
<td>BMC Cancer</td>
<td>231</td>
<td>38</td>
<td>5,592</td>
<td>2.993</td>
<td>24.25</td>
</tr>
<tr>
<td>6</td>
<td>Scientific Reports</td>
<td>224</td>
<td>38</td>
<td>4,819</td>
<td>4.38</td>
<td>21.61</td>
</tr>
<tr>
<td>7</td>
<td>Clinical Cancer Research</td>
<td>181</td>
<td>54</td>
<td>9,223</td>
<td>12.53</td>
<td>51.13</td>
</tr>
<tr>
<td>8</td>
<td>International Journal of Molecular Sciences</td>
<td>169</td>
<td>35</td>
<td>4,105</td>
<td>5.92</td>
<td>24.38</td>
</tr>
<tr>
<td>9</td>
<td>Cancers</td>
<td>164</td>
<td>22</td>
<td>2,479</td>
<td>6.64</td>
<td>15.16</td>
</tr>
<tr>
<td>10</td>
<td>Oncology Letters</td>
<td>156</td>
<td>21</td>
<td>1,673</td>
<td>2.97</td>
<td>10.76</td>
</tr>
</tbody>
</table>
Development in biomarkers of breast cancer: a bibliometric analysis from 2011 to 2020

between 22 different white blood cell subpopulations and cancer survival was identified. This resource and the related analysis tool could help to describe genes and the prognosis of patients with white blood cell subgroups and across cancers and identify potential biomarkers and therapeutic targets. Chandrashekar et al.\textsuperscript{29} introduced UALCAN, allowing users to perform the follow-

Figure 5. Map of the authors with publications related to biomarkers in breast cancer. A, The 101 authors who occurred more than 15 times were divided into 13 clusters by different colors. Cluster 1: red; cluster 2: green; cluster 3: dark blue; cluster 4: yellow; cluster 5: purple; cluster 6: light blue; cluster 7: orange; cluster 8: brown; cluster 9: light purple; cluster 10: pink; cluster 11: light green; cluster 12: grey blue; cluster 13: light yellow. The size of the nodes represents the frequency of occurrences. B, Visualization of the authors according to the average year of publication (APY). Authors in yellow appeared later than authors in blue.
ing operations: (I) analysis of gene expression in tumor and normal samples in different subgroups, based on a single cancer grade, tumor stage, ethnicity, weight or other clinical pathological features; (II) assess gene expression levels and its effect on the survival of patients with clinical pathological features; and (III) identify genes that are upregulated or downregulated.
Cortez et al.\textsuperscript{31} discussed the role of fluid-expressed miRNAs as reliable cancer biomarkers and predictors of treatment response, as well as a potential new criterion for patient selection in clinical trials. In addition, another study\textsuperscript{32} explored the concept that miRNAs could function as hormones. Cohen et al.\textsuperscript{33} described a blood test called CancerSEEK, which was performed on 1,005 patients with non-metastatic cancer. The sensitivity range for detecting different cancer types varies\textsuperscript{34}. Ihara et al.\textsuperscript{35} examined the tumor growth factor (TGF)-β signaling pathway as a potential drug target, the clinical application of TGF-β inhibition, the problems arising from anti-TGF-β therapy, and how these problems can be addressed using personalized administration methods, biomarker monitoring, and simple and/or local administration regimens. Melo et al.\textsuperscript{36} reported that BC-associated exons contain miRNAs associated with RNA-induced silencing complex (RISC) loading complexes (RLC) and demonstrated cell-independent ability to process precursor miRNAs (pre-miRNAs) into mature miRNAs. These papers have played an

\begin{figure}[ht]
\centering
\includegraphics[width=\textwidth]{figure7.png}
\caption{The yearly number of local citations of papers with high global citations (GCs). The size and colors of the circle represent the GCs of the publications.}
\end{figure}
important and arguably groundbreaking role in the study of biomarkers in BC.

**Co-cited References Analysis**

Considering the large number of references, the minimum number of references was set at 110. Of the 430,672 references, 110 were screened for analysis (Figure 8). These papers were divided into different clusters using nodes of different colors. Cluster 1 (red) contains 44 references, and these references mainly focused on studies examining the differences in gene expression patterns from cDNA microarrays to classify BC and to relate tumor characteristics to clinical results.

![Figure 8. Map of co-cited references in studies related to biomarkers in breast cancer. A, A network map of co-cited references. Of the 430,672 references, 130 (classified into 5 clusters) were cited at least 110 times. Cluster 1: red; cluster 2: green; cluster 3: blue; cluster 4: yellow; cluster 5: purple. B, A density visualization for the co-cited references.](image-url)
Cluster 2 (green) publications mainly focused on biomarkers of BC that were identified through modern life science techniques such as genomics and transcriptome analyses. Cluster 3 (blue) mainly focused on the role of miRNAs as novel biomarkers in the diagnosis of BC. The main content of cluster 4 (yellow) publications was the pathogenesis of BC and the survival of patients. Cluster 5 (purple) literature paid attention to the efficacy, safety, and drug resistance of therapeutic agents in the treatment of BC, with the majority being clinical studies. To further investigate the co-citation of references, 26 references in group 3 were analyzed using density visualization. Figure 9 shows that gene expression in BC was the theme of cluster 3 co-citation literature. For instance, references on the roles of genes in the pathogenesis of BC, especially exocrine bodies, have been widely cited in relation to miRNAs. In addition, references on gene mutations and expression profiles were also abundant. Although the keyword "profiles" is referenced relatively few times, in the center of the network, the topic needs to be further investigated.

Analysis of Research Hotspots

In addition to the search terms, VOSviewer was also used to analyze keywords of 13,680 publications (Figure 10). As presented in Figure 10A, publications of cluster 1 majorly focused on the multi-omic study of BC biomarkers. Cluster 2 mainly involved treatments for BC, and most of the reports detailed clinical research. Cluster 3 focused on basic research of BC, most of which involved molecular biology. Cluster 4 pieces of literature were mainly related to the prognostic factors in BC. Cluster 5 mainly involved the pathogenesis of BC. The most frequently occurring keywords were “breast-cancer”, “biomarkers”, “expression”, “survival”, and “prognosis”, suggesting that research related to biomarkers in BC was largely clinical research. According to Figure 5B, VOSviewer divided the colors of all keywords into different types, based on the average year of publication (APY). “Extracellular vesicles” was the latest (cluster 1, APY: 2018.03). The next one was “liquid biopsy” (cluster 1, APY: 2017.89) and “exosomes” (cluster 1, APY: 2017.64). It is noteworthy that exosomes are closely associated with extracellular vesicles. Interestingly, “extracellular vesicles” (cluster 1, APY: 2018.03) and “biomarkers” (cluster 5, APY: 2016.44) were the latest subjects.

Analysis of Research for Biomarkers in BC Related to COVID-19

To screen for COVID-19, BC biomarkers associated with COVID-19 were screened, subsequent analysis was carried out to understand the trends and hotspots of COVID-19-related BC (Figure 11). For COVID-19, as shown in Figure 11, individual patient profiles, inflammation, pandemic, and BC were common. In addition, Figure 11 shows the role of COVID-19-related BC. Anti-inflammatory treatment strategies and the establishment of personalized patient profiles were also revealed.

Discussion

BC is a very common malignancy among women, which has a high rate of incidence\(^3\). Although the treatment of BC has improved significantly in recent years, the prognosis is not ideal due to individual patient differences\(^4\). Therefore, it is necessary to identify sensitive and accurate biomarkers to better predict the survival and prognosis of BC patients. In this research, we used VOSviewer to analyze biomarkers related to BC. The WoSCC database was accessed to investigate the research trends and hotspots. 13,680 publications were retrieved and analyzed. Based on the polynomial fitting curve, the annual number of publications was on the rise. Among all countries, the US was first in terms of Np. Six US affiliations and five US authors were among the top ten affiliations and authors of biomarker studies in British Columbia, meaning that the US has some of the best institutions and professional researchers in the world, which partly explains the rapid growth of this field in the US during the past 10 years. Although China has a high H-index, the Np and Nc average per item was relatively low. This suggested that Chinese scholars and affiliated institutions should increase the quality of publications. There is a contradiction between the quality and quantity of papers in Australia. 6 journals had high IF, suggesting that publishing a study of biomarkers in BC in high-IF journals was not difficult. In terms of GCs, the top 10 papers were all published in journals with high IF, suggesting that the journals were publishing more potential breakthroughs in this field, and more attention should be given to these journals.

This study facilitated a better understanding of the trends and hotspots in this field. At the same time, our research improved the understanding of this field, with GCs as the indicator. However, this study had some limitations. First, it failed to ana-
Figure 9. The network mapping of keywords in cluster 3 of the co-cited references network map. A, A network map of the keywords. B, Visualization of the keywords according to the average year of publication (APY). Keywords in yellow appeared later than those in blue. C, Density visualization for keywords.
Figure 10. Map of keywords related to biomarkers in breast cancer. A, The 131 keywords that occurred more than 162 times were divided into 5 clusters by different colors. Cluster 1: red; cluster 2: green; cluster 3: blue; cluster 4: yellow; cluster 5: purple. The size of the nodes represents the frequency of occurrences. B, Visualization of keywords according to the average year of publication (APY). Keywords in yellow appeared later than keywords in blue.
proteins have been used to study the relationship between miRNAs and tumors to elucidate its mechanisms. The high sensitivity makes miRNAs an ideal novel biomarker for diagnosis, treatment, and postoperative monitoring. However, our understanding of miRNAs is still in its infancy, and the precise roles of miRNAs in BC remain to be fully elucidated. With advancements in experimental methods, the miRNA regulatory network will gradually become clearer, and it may be possible to decode the mechanisms of other tumors and improve the survival rate of patients.

Second, gene promoter methylation is very common in BC. Over 100 gene promoters have been reported to be hypermethylated in BC tissues, and this modification may lead to the inactivation of tumor suppressor genes. However, the exact roles of these gene promoters in BC development and progression are not fully understood. Further research is needed to elucidate the mechanisms underlying gene promoter methylation in BC.

The full text of the article, VOSviewer might miss some important information. Second, due to the exclusion of high-quality articles published in 2022, this study, to some extent, may not be entirely up to date.

The keyword analysis results demonstrated that tumor mechanisms, tumor markers, DNA methylation, apoptosis, and proliferation have been the research hotspots of tumor exosomes in recent years. This may be due to the following reasons. First, BC-related miRNA has differential expression in tissues, peripheral blood plasma, and even the urine of cancer patients. Immunohistochemistry, qRT-PCR, gene knockout, exosome marker proteins have been used to study the relationship between miRNAs and tumors to elucidate its mechanisms. The high sensitivity makes miRNAs an ideal novel biomarker for diagnosis, treatment, and postoperative monitoring. However, our understanding of miRNAs is still in its infancy, and the precise roles of miRNAs in BC remain to be fully elucidated. With advancements in experimental methods, the miRNA regulatory network will gradually become clearer, and it may be possible to decode the mechanisms of other tumors and improve the survival rate of patients. Second, gene promoter methylation is very common in BC. Over 100 gene promoters have been reported to be hypermethylated in BC tissues, and this modification may lead to the inactivation of tumor suppressor genes. However, the exact roles of these gene promoters in BC development and progression are not fully understood. Further research is needed to elucidate the mechanisms underlying gene promoter methylation in BC.
to have hypermethylation\textsuperscript{38}. This includes cell cycle regulatory genes, such as cyclin D2 (\textit{CCND2}) and cyclin-dependent kinase inhibitor (\textit{CDKN2A})\textsuperscript{39,40}; DNA repair genes, such as breast cancer susceptibility (\textit{BRCA}) genes 1 and 2\textsuperscript{41,42} and glutathione S-transferase P1 (\textit{GSTPI})\textsuperscript{43}; tissue invasion and metastasis genes, such as Ras-related region family 1A (\textit{RASSF1A})\textsuperscript{44}; cell transcription genes, such as source frame gene A (\textit{HOXA1}, \textit{HOXA5}, \textit{HOXA9}, \textit{HOXA10}, etc.)\textsuperscript{45,46}; cell adhesion genes, such as cadherin 1 gene (\textit{CDH1})\textsuperscript{47}; and cell signal transduction genes mediated by excin, such as estrogen receptor (\textit{ER})\textsuperscript{α}\textsuperscript{48}. The hypomethylation of the proto-oncogene promoter may also lead to cancer, such as in trefoil factor 1 (\textit{TFF1})\textsuperscript{49}. This indicates that gene methylation plays an important role in the growth and cancer metastasis of BC. Most BC patients are already advanced or metastatic at the time of diagnosis\textsuperscript{50}. Therefore, early screening markers for cancer are of great significance. The auxiliary examination methods for BC include physical examination, ultrasonography, targeted examination, and magnetic resonance imaging (MRI). Ultrasound and targeted examination are prone to false positives, high radiation exposure, pain, anxiety, and negative psychology. False positives will lead to an overdiagnosis of BC, and MRIs are expensive\textsuperscript{51,52}. Therefore, it is important to identify novel economic cancer screening prognostic biomarkers with high sensitivity and specificity. DNA methylation modification is involved in the early process of cancer, and there are different patterns of DNA methylation with the progression of cancer. In recent years, a great number of works of literature investigated the roles of circulating or cell-free DNA (cfDNA) methylation status in the diagnosis and prognosis of various tumors, including BC\textsuperscript{53-55}. CfDNA detection has the advantages of convenient acquisition, non-invasive, and reproducibility, and has broad prospects in the detection and diagnosis. In triple-negative BC, there is no significant difference in the methylation level of the \textit{BRCA1} promoter in cfDNA and tissue samples. Therefore, cfDNA, as a biomarker, may have broad applications in cancer detection. DNA methylation plays a role in the early screening of BCs and is also closely related to poor prognosis of patients. Researchers have found that the methylation of the paired homeobox transcription factor 2 gene (\textit{PITX2}) increased the risk of poor prognosis in BC patients\textsuperscript{56,57}. \textit{RARRA} and \textit{HIN-1} methylation frequencies are associated with distant metastasis of BC, and methylation was observed at higher levels in metastatic BC tissues compared to primary BC tissues. Furthermore, researchers\textsuperscript{58,59} showed that methylation levels were closely associated with survival time in BC patients.

**Conclusions**

This bibliometric analysis demonstrated that research on biomarkers in BC is developing rapidly. The US is the major producer and has lots of outstanding achievements. The role of p53 in tumor cell circulation has become a research hotspot. The identification of biomarkers for BC has been affected by the COVID-19 disease.

**Conflict of Interest**

The Authors declare that they have no conflict of interests.

**Acknowledgments**

We thank the Hainan Province Clinical Medical Center for providing support.

**Funding**

Hainan Provincial Natural Science Foundation of China (No. 822MS171), Hainan Affiliated Hospital of Hainan Medical University National Natural Science Foundation of China Cultivation 530 Project (2021QNXM04) and Innovative Research Projects for Graduate Students in Hainan Province (Qhys2021-346).

**Data Availability**

All data are available upon request by contact with the corresponding author.

**Authors’ Contributions**

Hai Ying Liu, Yan Chen, and Ya Ping Yu wrote the original draft. Yang Yu collected, processed data, and created visualizations.

**Ethics Approval**

Not applicable.

**Informed Consent**

Not applicable.
References


Development in biomarkers of breast cancer: a bibliometric analysis from 2011 to 2020

