Journal of Cancer Science and Clinical Oncology
Volume 9 | Issue 1
ISSN: 2394-6520

Review Article Open Access

An Omics Study of HBV-Associated Hepatocellular Carcinoma: a Bibliometric Analysis

Qingxiu Li[1], Changzhu Duan[1,*]

[1]Chongqing Medical University, Yuzhong Medical College Road, Chongqing, China

[*]**Corresponding author:** Changzhu Duan Chongqing Medical University, Yuzhong Medical College Road, Chongqing, China Tel: 023-68485804, E-mail: duanchzhu@cqmu.edu.cn

Citation: Qingxiu Li, Changzhu Duan (2022) An Omics Study of HBV-Associated Hepatocellular Carcinoma: a Bibliometric Analysis. J Cancer Sci Clin Oncol 9(1):105

Abstract

Hepatocellular carcinoma (HCC) is a common malignant tumor, and hepatitis B virus (HBV) is the main pathogenic factor. The occurrence of HBV-associated hepatocellular carcinoma (HBV-HCC) is a complex process caused by multi-gene and multi-step interactions. The synergistic effect of multiple cancer-promoting mechanisms accelerates the evolution of the disease from inflammation to tumorigenesis. The maturity of sequencing technology, the advancement of bioinformatics data analysis technology and the development of omics technology have improved the efficiency of research. With the application of single-omics techniques including genomics, transcriptomics, metabolomics, and proteomics in tissue and body fluid samples, and even the newly developed multi-omics analysis on single-cell platforms, HBV-HCC changes can be better analyzed. Bibliometrics is a quantitative statistical analysis tool used to analyze and observe research trends. The aim of this study was to evaluate omics studies of HBV-associated hepatocellular carcinoma and explore the hot spots and frontiers using bibliometric methods from 2001 to 2022.

Keywords: Hepatocellular carcinoma; hepatitis B virus; Omics; Bibliometric

Introduction

Liver cancer is the fifth most frequent cancer in men and the ninth most common cancer in women worldwide[1]. It is also the second most common cause of cancer death. HCC is the most important primary liver cancer, which can be caused by HBV, hepatitis C virus (HCV), alcohol abuse, etc[2,3]. Although hepatitis B virus infection is adequately controlled by vaccines, approximately 350 million people worldwide are chronically infected with hepatitis B virus (CHB), and at least 50% of HCC patients are infected with chronic hepatitis B virus (HBV)[4]. Nonetheless, the role of HBc in the development and progression of liver disease remains unclear. Similar to other types of oncogenic mechanisms, HBV viral-induced hepatocellular carcinogenesis is a multistep process that requires the acquisition of all cellular features responsible for the tumor phenotype, which was originally described by Weinberg in 2000 in "Hallmarks of Cancer" and updated in 2011[5]. HBV core protein (HBc) encoded by HBV genome may play an important role in HBV life cycle. Persistent HBV infection or active replication of the HBV gene can lead to liver injury, fibrosis, cirrhosis and liver cancer. All in all, HBV contributes to hepatocellular carcinoma (HCC) development through direct and indirect mechanisms[6].

On February 15, 2001, the preliminary completion of the Human Genome Project ushered in a new era of life science exploration. Large-scale biological data collection and processing technologies such as DNA sequencing, gene chips and bioinformatics algorithms have rapidly become the core stage of life science, and data have become the basic resources of life science research[7]. "Omics" techniques are primarily used to detect all proteins, transcripts and metabolites to mine available data in biological samples[8]. These high-throughput techniques play a key role in characterizing gene and/or protein expression profiles and their effects on HBV-HCC[9]. A comprehensive understanding of the pathogenesis of HBV-associated carcinoma is crucial for the early diagnosis, therapy and prevention of hepatocellular carcinoma[10]. Recent advances in deep sequencing technologies, including next-generation sequencing, nanopore sequencing, and single-cell sequencing, have helped uncover genetic and epigenetic changes in tumor tissue and chronic liver injury caused by hepatitis virus infection[11-13].

Bibliometrics is a new method to summarize the progress of a research field. Its main purpose is to establish knowledge maps and discover hot spots and even emerging trends in the research field. Mathematical and statistical methods are used to quantitatively analyze numerous articles in a particular research field to reveal many aspects and research trends in this field[14]. In the last decades, an increasing amount of related research articles has been published, and many researchers have used this strategy to evaluate their respective research fields. Liu et al[15] used bibliometrics to outline the map of renal cell carcinoma immunotherapy, and found that the current research hotspots in this field may focus on "immune checkpoint inhibitors", "PD-1" and "mammalian targets of rapamycin". Through bibliometric analysis, Wu [16] found that the current research hotspots of pancreatic cancer mainly focus on the energy metabolism in the hypoxic tumor microenvironment, the regulation of cancer-related fibroblasts in the tumor microenvironment, accurate diagnosis, drug delivery and new therapies. However, up to now, there has been no specific scientometric study on the knowledge graph of HBV-HCC omics research. In summary, this study used bibliometrics method to visually analyze related articles in the field of HBV-HCC omics, so as to understand the current status of HBV-HCC omics research, identify research hotspots in this field, and summarize the future development trend of this field.

Methods

The collection of data

The data analyzed in this paper are from the Sicence Citation Index (SCI) of the core data set of the Web of Science Collection (WoSCC). The articles included in WoSCC are characterized by long time span, large number of articles and high quality. Firstly, we collected "HBV", "hepatocellular carcinoma", "Bioinformatics", "Genome", "Gene Microarry", "NGS", "high-throughput sequencing", "Transcriptome", "Transcriptome sequencing", "Sequence Alignment", "ncRNA", "Proteome", "Network" and "Metabolome" in the title, abstract or keywords of published articles. According to Alex Pollock [17], only the original peer-reviewed paper data can represent the original scientific development. Therefore, the completeness degree of metadata of different document types in WOSCC was considered in data selection, and only "ARTICLE" was retained. The time span of the article is 2001-2021. The data retrieval period

ended on 2 July 2022.Some of the newly published articles have the phenomenon that the Web of Science is not registered. After the search is complete, the search records that meet the search conditions are exported, including the full records and cited documents. The export format is TXT for subsequent analysis. The inventor of CiteSpace, Professor Chen Chaomei[18,19] said there is no need to endlessly optimize search queries to eliminate papers on irrelevant topics. After removing duplications and screening articles, 855 English articles were eventually selected as the research objects of this study.

Software and algorithms for data analysis

In this study, Origin 2019 was used to draw the cumulative line chart, and CiteSpace (6.1.R3), VOSviewer (1.6.18), Scimago Graphica, Carrot2 and HistCite (12.03.17) software were used for visualization analysis, and use R language to process data[20-24]. The analysis settings used for data analysis in CiteSpace software are as follows: Set the period from 2001 to 2022 as a time slice, and each slice is one year. The node type is set to the keyword. In the analysis of collaboration network and co-occurrence network, Top N per slice was selected for the number of objects extracted in each time period, and Top N =50 was selected for the most frequently cited or appeared items for countries and keywords in each slice. It means to extract the number of objects in each time slice. To get a clear view of the network, Pathfinder is used to eliminate redundant, cumbersome or visually poor links. Leave all other settings as default. Sources analysis in VOSviewer software is based on full counting, in which the number of articles published by each journal is set to at least 8. The method of organizational analysis is full counting, in which the number of articles published by each organization is set to at least 22. The Global Total citation score (TGCS) in HistCite software is the number of citations in SCIE, while the local total citation score (TLCS) is the number of citations in the current publication set. In this paper, Scimago Graphica and Carrot2 software are used to maintain their default settings for data analysis. R language is used to process data and construct regression models.

Result

Number of research papers published

In order to analyze the course of HBV-HCC omics research, R language was used to process the published paper data, and these data were used to construct a regression model (f (x) =p0x) n+p1xn−1+p2xn−2+p3xn−3+... + PN), we observed a statistically significant relationship between year and number of publications by fitting the data (R^2=0.8994). Based on the fitted curve, we estimate that in 2022, the number of publications related to HBV-HCC omics research is 82, and the accumulated literature will reach 904. We made an inductive analysis of the number of articles published each year and the accumulated number over the years. As shown in Figure 1, the number of published papers in 2011 doubled compared with that in 2010, and the number of published papers increased significantly for the first time. In recent 10 years (2012-2022), the total number of articles published in this field showed a continuous growth trend, and the cumulative number of articles published from 2017 to 2021 increased by 4.25 times compared with that from 2001 to 2011. In addition, 463 articles (54.15%) were published in the past 5 years, which illustrates that increasingly mature omics technology is popular in the field of HBV-HCC research. Over the past 20 years, the largest increase in the number of articles published in this field was recorded between 2015 and 2016, with 27 articles. Among them, 2016 is the largest increase. And according to statistics, the citation frequency of 855 published articles after removing self-citation was 24756. Each article received an average of 31.21 citations. This illustrates that the research interest in this field is increasing.

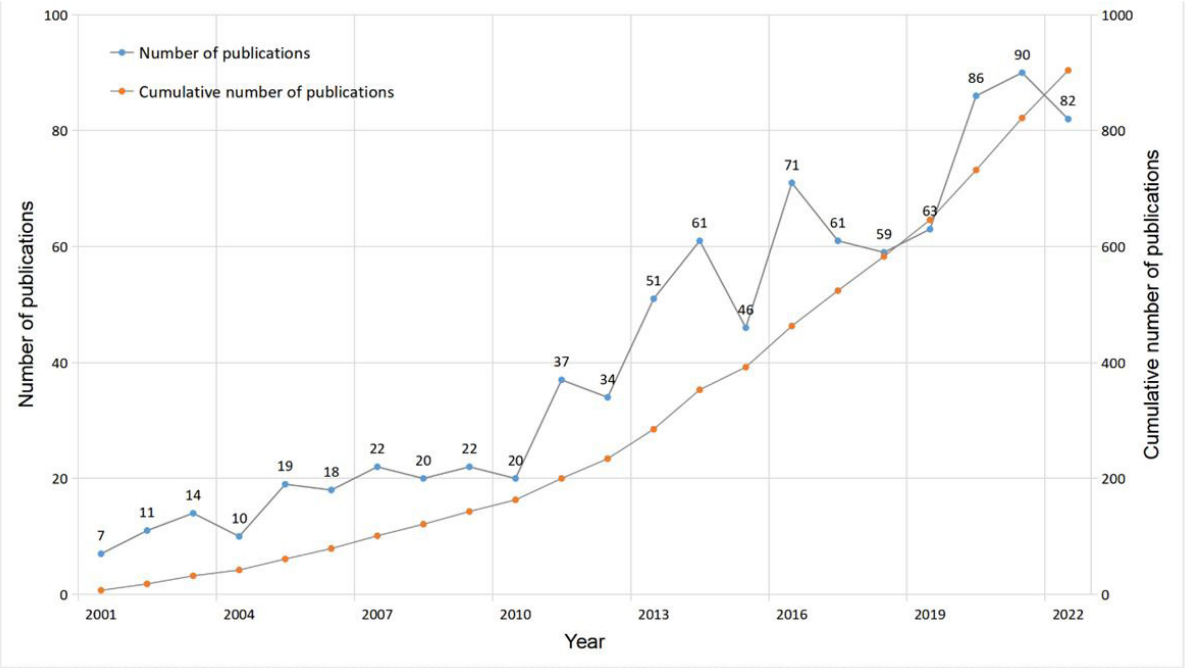

Figure 1: Publications trend of HBV associated hepatocellular carcinoma

Analysis of Country/region and institutional

A total of 58 countries/regions participated in the omics studies of HBV-HCC and published articles. Some of the articles are a collaboration of researchers from different countries. The total number of articles published by the top 10 countries/regions reached 84.09%. At the same time, China is the country with the highest number of papers published in this field, with a cumulative number of 539 papers, accounting for 63.04% of the total number of papers published in this field. This may be due to the high incidence of HBV-HCC in China. Followed by the United States (n = 168,19.6%), Japan (n = 83,9.71%), Korea (n = 56,6.54%) and France (n = 37,4.33%). The number of citations can reflect the recognition of the research peers. As shown in Figure 2, French people have the most citations per article, while German and Italian people have the second and third citation rates respectively.

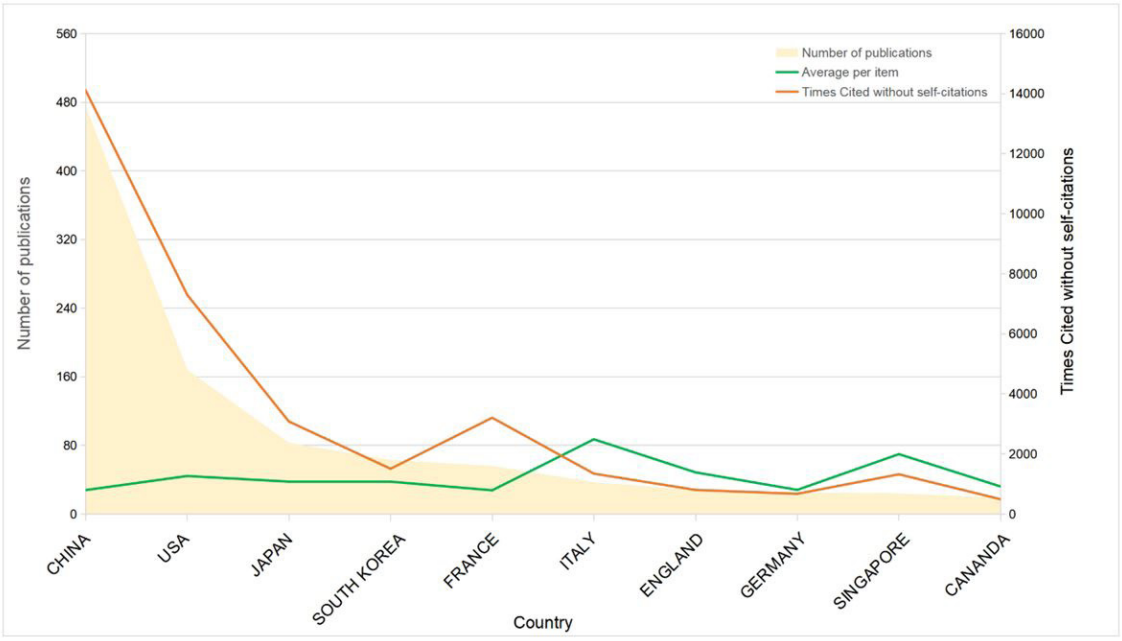

Figure 2: Citations trend of HBV associated hepatocellular carcinoma in top 10 country

In order to further understand the cooperation mode between countries/regions, Scimago Graphica software was used to generate the cooperation relationship network of countries/regions in the field of HBV-HCC omics research as shown in Figure 3. There are 6 clusters in the co-occurrence network, and the strong cooperative relationship occurs in the research group of this cluster. The close relationship between countries is centered on China, and the United States, Australia, Canada, the United Kingdom and Japan have relatively close cooperative ties. It can be seen that the number of articles published by China in this field is much higher than that of other countries in the world, which again proves that China is at the forefront of the research in this subject area[25].

Figure 3: The number of published articles and cooperative relationships in the country/region

A total of 1227 institutions published articles in the field of HBV-HCC omics. The top 10 Chinese institutions (Table 1) accounted for two-fifths of the articles published by all institutions. This is mainly because hepatocellular carcinoma (HCC) kills more than 300,000 people in China each year, accounting for about half of all HCC-related deaths worldwide. This means that the research field of hepatocellular carcinoma has received high attention from scientific research institutions in China[1]. At the same time, due to the vigorous development of bioinformatics and the further promotion of high-throughput technology, the study of hepatocellular carcinoma related genome, proteome and metabolome is increasing, which has aroused great attention of domestic scholars. In Figure 4, there are 13 nodes and 4 clusters in the co-occurring network diagram. It is also found that the links between research institutions may be geographically related. For example, the collaboration network between Fudan University, Chinese Academy of Sciences, Nanjing Medical University and the Second Military Medical University is formed. The University of Hong Kong, the Chinese University of Hong Kong and Sun Yat-sen University have formed close ties. As shown in Table 1, the number of articles published by Fudan University and Guangxi University of Traditional Chinese Medicine ranked first and second respectively, with 114 articles, accounting for 13.34% of the total number. Articles from the Chinese Academy of Medical Sciences and Peking Union Medical College and the University of Hong Kong were cited the most, with 58.69 and 58.55 citations respectively.

Institutions	Publication number	Citation per paper
FUDAN UNIVERSITY	68	52.32
GUANGXI MEDICAL UNIVERSITY	46	14.48
NAVAL MEDICAL UNIVERSITY	45	48.89
SUN YAT SEN UNIVERSITY	35	32.26
CHINESE ACADEMY OF SCIENCES	31	29
CHINESE ACADEMY OF MEDICAL SCIENCES PEKING UNION MEDICAL COLLEGE	29	58.69
UNIVERSITY OF HONG KONG	29	58.55
PEKING UNIVERSITY	24	21.58
CHINESE UNIVERSITY OF HONG KONG	23	57.7
HUAZHONG UNIVERSITY OF SCIENCE TECHNOLOGY	22	14.14

Table 1: Top 10 organizations with largest number of publications

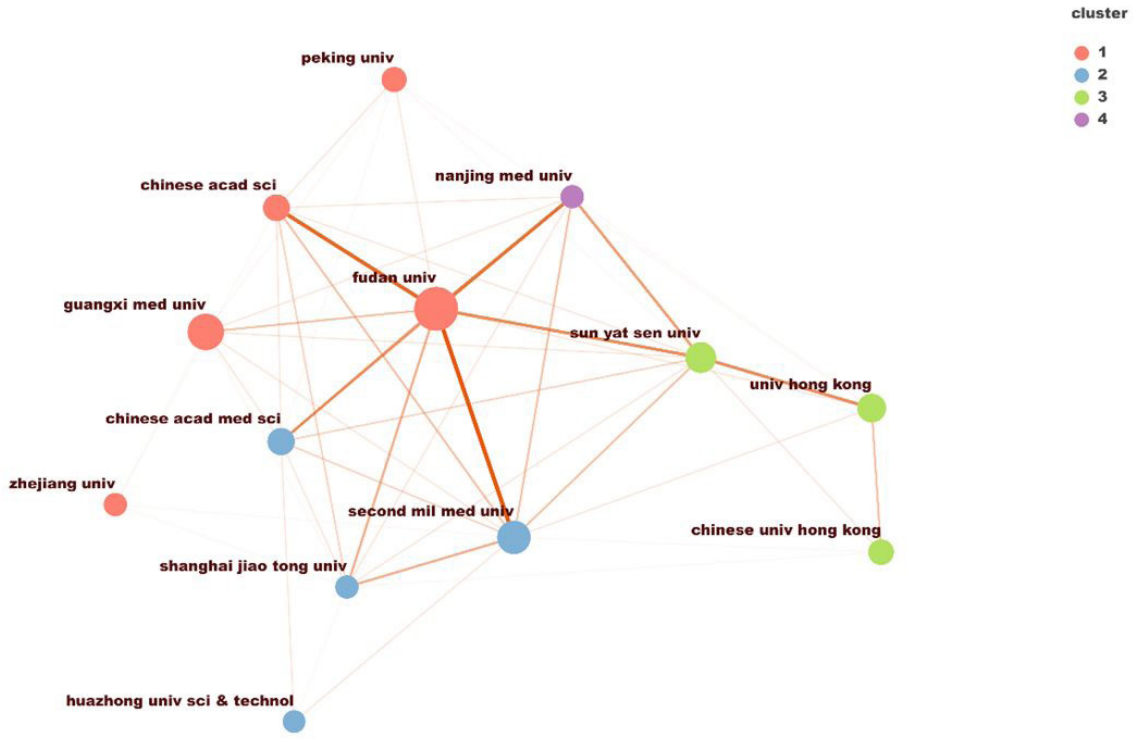

Figure 4: The collaboration network of institution

Research area and journal sources

According to statistics, a total of 308 SCI journals have published articles, and the top 10 journals have published 238 articles, accounting for two seventh of the total number of published articles. Table 2 illustrates the number of citations and the average number of citations per article of the top 10 journals. Using VOSviewer software for analysis, the direct publication articles of these journals are set to greater than 8. Figure 5 illustrates the density of citations of 20 journals. The higher the number of journals near a point in the figure, the higher the weight of neighboring journals, the closer the color of the point is to red. Conversely, the smaller the number of journals near a point, the lower the weight of neighboring journals, and the closer the color of the point is blue. These journals are principally in oncology, cell biology, pathology, pharmacy and cell biology. Examples include PLOS ONE, JOURNAL OF MEDICAL VIROLOGY and HEPATOLOGY.

Journals	Publication number	Average per item
PLOS ONE	51	28.8
JOURNAL OF MEDICAL VIROLOGY	31	25.42
HEPATOLOGY	28	108.71
WORLD JOURNAL OF GASTROENTEROLOGY	27	36.81
SCIENTIFIC REPORTS	25	20.36
ONCOTARGET	22	18.41
JOURNAL OF VIRAL HEPATITIS	17	18.71
JOURNAL OF HEPATOLOGY	14	53.64
ONCOLOGY LETTERS	13	8.46
INTERNATIONAL JOURNAL OF CANCER	10	29

Table 2: The top 10 published journals in the field of bioinformatics research on hepatocellular carcinoma

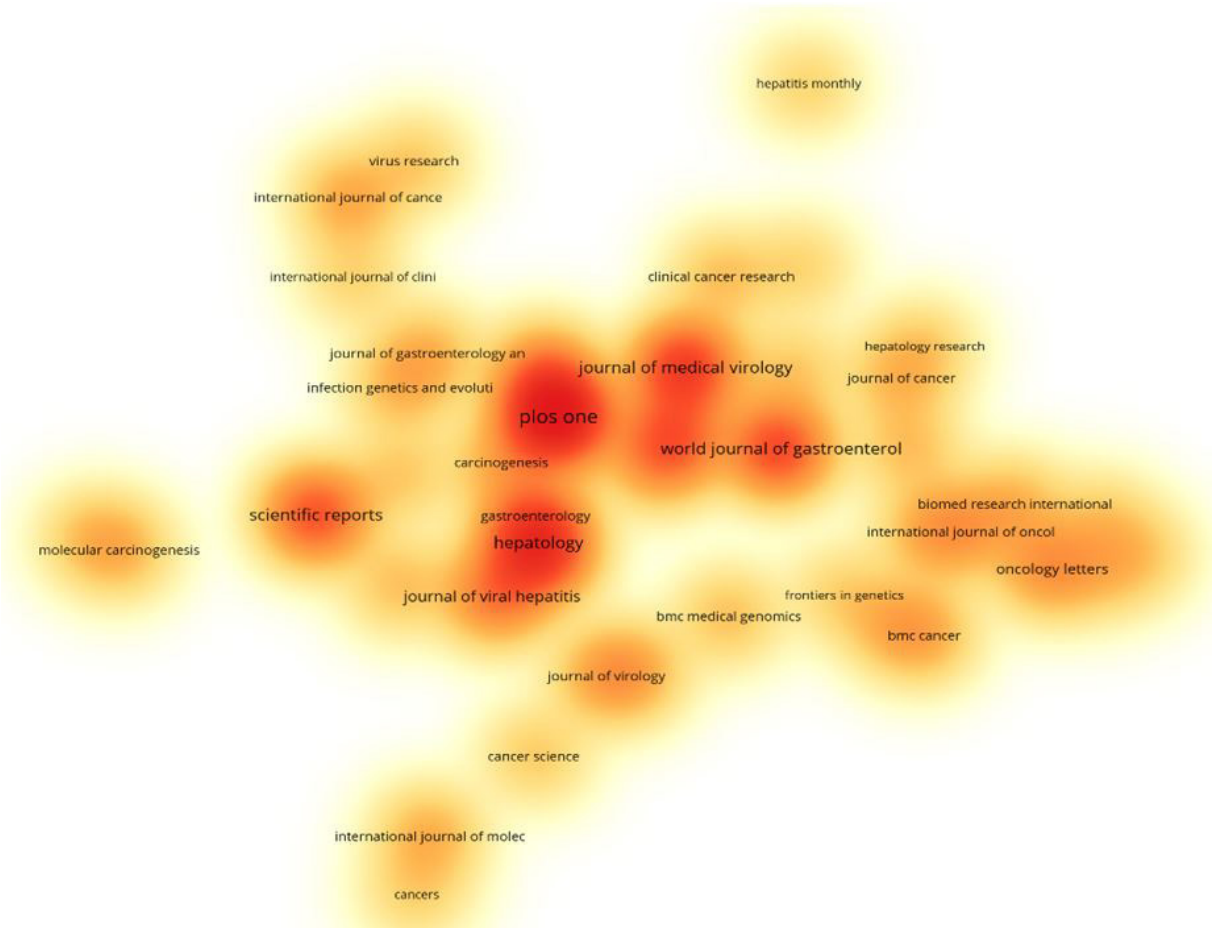

Figure 5: The Density map of journal citations

The double map overlay function of CiteSpace was used to construct the overlay map of the journal. Double graph analysis (also known as superposition graph) refers to revealing the overall scientific contribution. Interplay of journals further divided into different regions on the dual map. Figure 6 depicts the distribution, citation trajectory and the drift of the center of gravity of papers in each discipline. The left side of the figure is the cited part, and the right side is the cited part. The curve is the line of quotation, which can show the whole story. In the left panel, the more papers published in the journal, the longer the vertical axis of the ellipse, and the more authors, the longer the horizontal axis. The label represents the discipline covered by the journal. From left to right, colored lines depict citation paths. There are three different reference paths. The two orange citation paths indicate that research in molecular/biology/genetics journals and health/nursing/medical journals is frequently cited in research in molecular/biology/immunology journals. The green path indicates that research in molecular/biological/genetic journals is frequently cited by research in medical/medical/clinical journals. This different citation path indicates that HBV-HCC omics research is global research that needs the joint efforts of different disciplines.

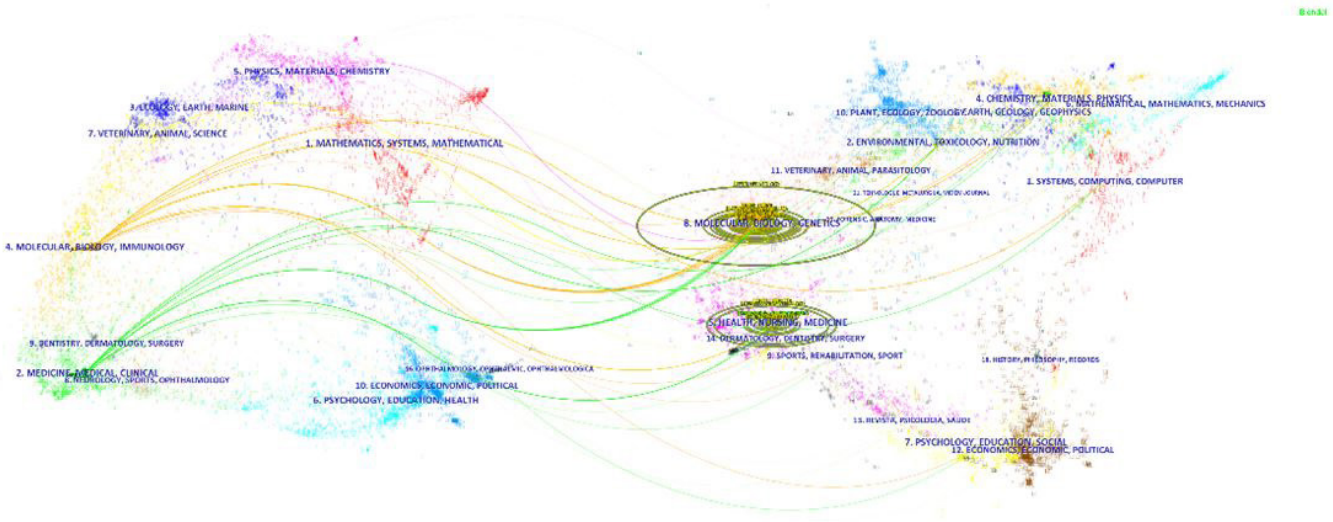

Figure 6: The dual-map overlays of articles

Discussion

Hotspots of Research

Keywords in a paper can be used to summarize the main content of the paper and extract useful information about the paper, such as goals, methods and ideas. The frequency analysis of keywords is the key to the research and development of hot topics related to a particular field. Therefore, co-occurrence analysis of keywords in a research field is beneficial in identifying research hotspots. If one or several keywords appear repeatedly in numerous articles, then the topic represented by these keywords is the research hotspot in this field. For keywords in HBV-HCC omics studies, Crrot2 software was used to display the keyword clustering statistics of all articles. As shown in Figure 7, the relevant hotspots are the research on viruses (including HBV, HCV and HIV) and the significance and credibility of biological information analysis (including CI 95, Treatment significance and P value). This indicates that the focus of omics research on HBV-HCC has always been the study of the mechanism of HBV infection, and the credibility of omics research data is also very important. At the same time, in order to evaluate the high-frequency keywords and the network mappings between them, the running node type is set to "Keyword", and comparable keywords such as "HCC", "HBV" and "omics" are removed. The results showed that the top 5 keywords were "expression", "gene", "risk", "infection" and "mutation", and the frequencies were 201, 126, 112, 92 and 81 times, respectively. HBV is an enveloped double-stranded DNA virus whose classification includes 10 genotypes (A-J). HBV virions contain 3020 to 3320 nucleotides in the partial double-stranded DNA genome (in the form of relaxed circular DNA, rcDNA) in nucleocapsids composed of hepatitis B core antigen (HBcAg) subunits. The nucleocapsid is encapsulated in a host-derived lipid bilayer overlaid with hepatitis B surface antigen (HBsAg). Research discovered by HBV genome integrated into the host genome, induce the host's chromosome instability (one of the typical features of the vast majority of human cancers) or change

the endogenous gene expression and function of normal play, make liver cells have the potential of malignant transformation, closely involved in the process of development of primary liver cancer[26]. For example, HBV integration in CHB-HCC patients in intergenic regions, CpG islands, simple repeats, repeat regions and telomeres, leading to chromosomal instability[27-29]. As confirmation, next generation sequencing (NGS) studies have found that HCC tumors generally have an elevated number of integration events and an increased integration frequency in the coding or promoter region when comparing HBV integration sites between tumors and matched non-tumor tissues [29]. Likewise, Tanaka, Y [30] found that an average of 40 to 60 individual cell alterations were detected in the protein-coding region of the HCC genome. These mutated genes and their copy number changes affect key pathways such as cell cycle control, telomere maintenance, chromatin modification and receptor tyrosine kinases. Among these mutations, there are also genomic alterations that are thought to be directly involved in the activation of important signaling pathways in HCC genesis [9,31,32].

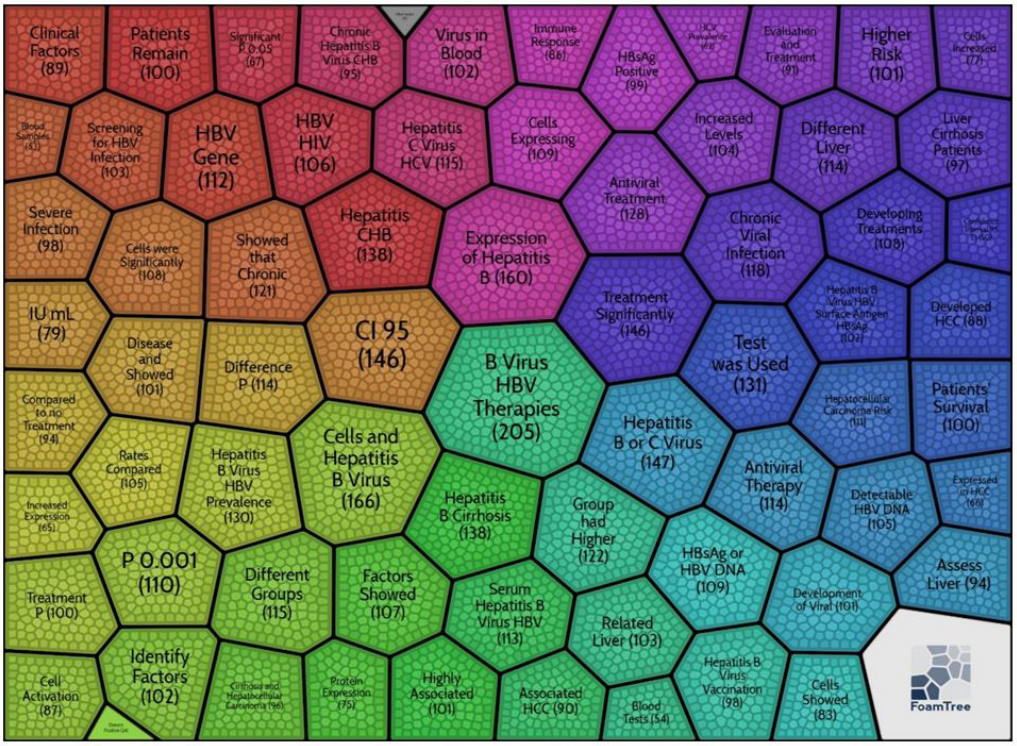

Figure 7: The keyword co-occurrence analysis of carrot2

In order to clarify the research hotspots and focus of HBV-HCC omics, we clustered co-occurrence keywords to better present the research topic. The effect of clustering is determined by the number of lines between keywords, the Modularity Q (Q) of map information and the Sihouette S (S) of map profile. The more lines between keywords, the larger Q value means the better clustering effect. When Q > 0.3, the modularity of map information is significant, and when the S value is greater than 0.5, it means that the clustering result is credible. For keyword clustering with modularity in figure 8, Q is 0.3723 and the average contour is 0.6944, which illustrates that the clustering results are suitable for analysis. The co-occurrence network graph has a total of 223 nodes and 1022 relevant connections, and the network density is 0.0413, indicating that although there are not many links between keywords, they are relatively close. The order of the numbers in the figure represents the importance of each topic in the field. As can be seen from the figure 8 the top 3 hot topics in the research field of HBV-HCC omics in the past 20 years were #0 mutation, #1identification and #2 Long non-coding RNA.

The genome contains four genes (P, preC/C, S and X) that encode for five main proteins: polymerase (gene P), HBcAg (gene C), hepatitis B envelope antigen (HBeAg) (product of preC), HBsAg (gene S), and a replication cofactor X (gene X)[33,34].The current study indicated that preS/S region mutations are related to vaccine failure, immune escape, occult HBV infection and the occurrence of HCC[35,36]. Whereas P region mutations may lead to drug resistance to NA antivirals[37]. PreC/C region mutations are associated with HBeAg negativity, immune escape, and persistent hepatitis[25,30,38]. Additionally, X region mutations play an important role in the development of HCCt[39-42]. As mentioned above, HBV genomic DNA can be integrated into the host genome to induce embodied genomic mutations, which is one of the leading hotspots of HBV-HCC omics research.

More and more research has uncovered that the integration of hepatitis B virus in to host genome is closely involved in liver tumorigenesis. Integration of these hepatitis B viruses can induce host gene copy number changes (CNV), chromosomal changes, or host gene expression changes[43,44]. Some studies have reported that hepatitis B virus is more likely to integrate into the hot target genes TERT, FAR2 and MLL4 in the host tumor genome[30,45,46]. Meanwhile, other studies have reported that hepatitis B virus tends to integrate into the genomes of adjacent non-tumor tissues, targeting genes FN1 and SMAD5[47-49]. It can be seen that there is a great controversy among several studies on HBV integration, and the identification of HBV integration sites into the host genome is likewise one of the current hotspots. The process of hepatocellular carcinoma is multifaceted. Although it has not been sufficiently elucidated, some studies have demonstrated that numerous long non-coding RNAs (lncRNAs) are contributors to the development of HCC. These host-derived lncRNAs, which are typically targeted due to viral infection have been shown to modulate gene expression as signaling factors at the epigenetic, transcriptional, post-transcriptional and even post-translational levels[50]. These lncRNAs predominantly promote HBV replication and oncogene expression or down-regulate tumor suppressors. Few lncRNAs are known to inhibit tumorigenesis, and these are generally downregulated in HCC[51-53]. Hence, the mechanism of lncRNA dysregulation in HBV-related HCC promoting tumorigenesis and cancer progression is also a research hotspot.

HistCite software computed 30 important literature and citation relationships in this field, which also affirmed that these hot words are the focus of omics research field of HBV-HCC. The relevant diagrams and tables are shown in the attachment.The top eight cluster clusters are displayed in Figure 8. It should be noted that different hot topics contain multiple keywords. In the process of HBV-HCC omics research, hot keywords are constantly changing over time. Therefore, it is necessary to further explore the change of hot topics in HBV-HCC omics research over time, and give different research frontiers and research trends in this field based on this.

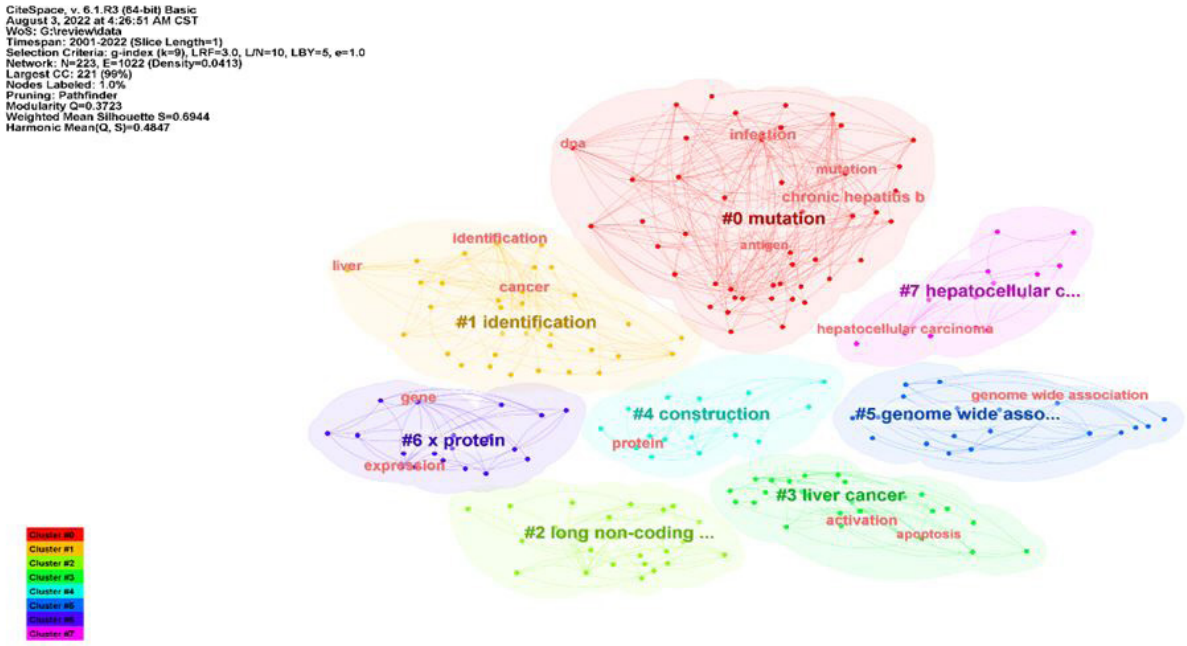

Figure 8: Co-occurrence cluster analysis of keywords

Trend of Research

Timeline time zone graph is a knowledge evolution view displayed from time gradient in CiteSpace software, which can analyze the research context and hotspot changes of a certain field under time series. The publication date is displayed at the top of the view, the most recent article is placed to the right, the nodes of the cluster share a horizontal line and the connections between nodes represent references. The more nodes in the cluster, the more important the research content is. The timeline time zone graph of HBV-HCC omics illustrates the popular phrases in the research process and the content of the research frontier at different stages. As shown in Fig. 9, the keywords "activation", "expression" and "mutation" became the research hotspot earlier in the relevant articles, which implied that the early study of HBV-HCC omics was closely related to this field. As can be seen from the above, HBV

genome integration will change the expression of genes near the integration site, causing damage to the host genome. Through large-scale genome sequencing analysis of HCC patients, TERT, TP53, and CTNNB1 / AXIN1 have been identified as core oncofactors because these genes regulate several pathways, including cell cycle (p53, p16), apoptosis (BCL2), cell proliferation and differentiation (b-catenin, P16). C-myc, APC, E-cadherin), metastasis (MMP4, MMP9, topoisomerase, Rb, Cyclin D1, osteopontin), angiogenesis (VEGFR-2, angiopoietin-2) and other growth factor signaling components (IGF-II, TGF, EGFR, HGF/c-met, PTEN, k-ras) , will lead to a large number of mutations in the somatic genome of the host, and then some relevant mutations suitable for the survival of HBV-infected cells will be screened to promote the formation of hepatocellular carcinoma[54,55]. Proteogenomic analysis identified five markedly mutated genes in tumor tissues of 159 HBV-associated HCC patients, including TP53 (58%), CTNNB1 (19%), Axin1 (18%), Keap1 (7%) and RB1 (6%)[56]. Among them, TP53, Axin1 and RB1 are typical tumor suppressor genes, while CTNNB1 is an oncogene. KEAP1 is an important regulator of cellular resistance to oxidative stress[54]. A disturbance of the KEAP1-NRF2 pathway is closely related to the induction of drug resistance in malignant tumors[57]. These studies suggest that hepatocellular carcinoma is not caused by mutations in specific driver genes, but is affected by multiple oncogenic pathways. "DNA integration", "pathway", "sequencing" and "suppressor gene" have also become hot words in this field.

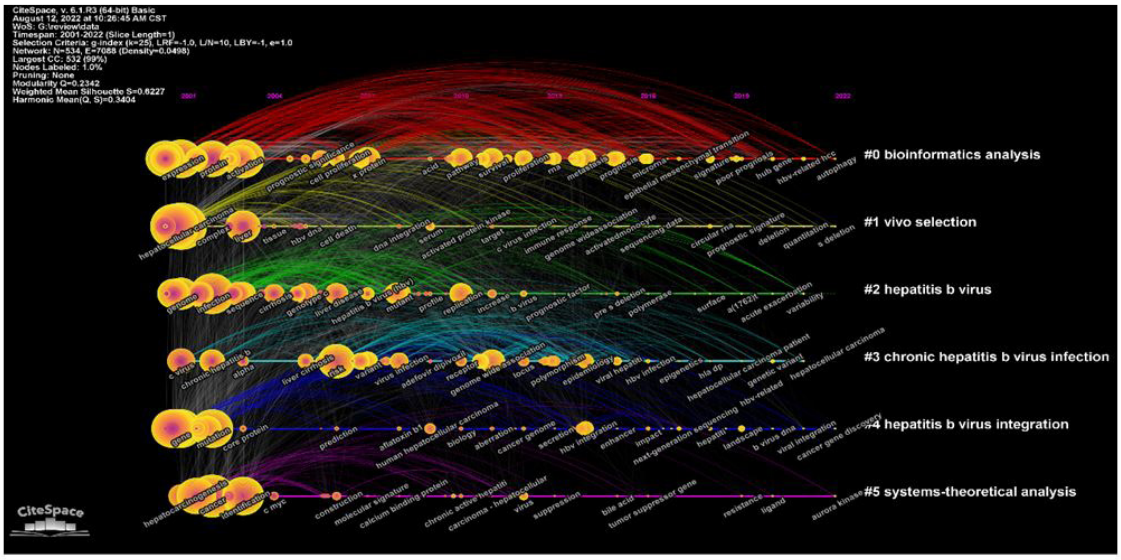

Figure 9: Timeline view of popular keywords

To speculate the development trend and direction of a particular field, we can summarize the emergent terms or keywords that are frequently cited in a short period of time in this field, which can be used as an evaluation index of the research frontier and trend in this field. In figure 10 the blue line represents the time interval, the red line represents the period of the burst keyword, and the endpoints of the red line represent the beginning and end of the time interval for each burst. The top 3 keywords of burst intensity were "surface antigen", "whole genome" and "carcinogenesis", with the intensity of 9.16,7.9 and 7.64, respectively. Over the past two decades, an increasing number of published articles on the study of HBV-HCC omics have demonstrated their wide application in many areas of life sciences. Through the combination of genomics (global gene analysis), transcriptomics (global gene expression analysis), proteomics (global protein expression analysis) and other comprehensive measurements, a variety of physiological and/or pathological events are systematically explained[58-61]. Metabolome is the study of the profiles of metabolites (e.g., amino acids, glycolipid, and hormones) that can be detected under certain conditions[62-64]. There may be changes in tumor metabolic pathways in HCC patients, and the resulting changes in nutrient availability are crucial to overcome nutritional starvation and changes in environmental conditions[65,66]. HBV is related to liver metabolism in the viral life cycle. This proof that cellular metabolism is involved in the prognosis of HBV-associated tumors has led to interest in targeting cancer therapy during metabolic processes[65,67]. In the study of metabolic pathways, omics evidence similarly indicates that the immune pattern between HBV and host is closely related to disease progression in patients infected with the virus and can modulate metabolic alterations in HCC cells through HBV proteins. As for HBV-associated hepatocellular carcinoma, although a large amount of individual omics data has been generated and some valuable information has been provided, incorporating data across multiple omics or utilizing systems biology approaches has rarely been

applied to models or clinics[68]. Consequently, utilizing multi-omics and systems biology analysis methods, discovering biomarkers, identifying new therapeutic targets, developing personalized therapy plans and ultimately preventing these diseases will be the future development direction.

The number of articles on HBV-HCC omics is increasing every year. The data we collected consisted of 855 articles and 22,795 references. The study has several limitations. Manuscript co-citation analysis is only feasible for papers in the WOSCC database. CiteSpace search results from other databases (including PubMed, Ovid, Scopus, and Google Scholar) were not feasible.

In summary, it can be seen from the keywords and literature analysis that hepatitis B virus infection can promote hepatocellular carcinogenesis through direct or indirect mechanisms. On the one hand, HBV can expand the instability of host cell genome by integrating or inducing host gene mutations, causing epigenetic remodeling of host DNA and abnormal expression of oncogenes and tumor suppressor genes. It can also induce malignant transformation of hepatocytes by activating various cancer-related signaling pathways, regulating cell metabolism and other mechanisms. On the other hand, the liver microenvironment is modified by chronic inflammation induced by HBV infection, interactions between the virus and innate immune cells, and adaptive immune cells, which help the virus evade immune surveillance and promote the evolution of the disease from inflammation to tumor formation. Further integration of data from multiple omics studies and the use of systems biological methods to study the mechanism of HBV infection-induced hepatocellular carcinoma can provide reliable new methods and new ideas for the prevention, diagnosis and treatment of HBV-HCC. This paper reviews the research progress of HBV-HCC, introduces the main mechanism of HBV-HCC, quantitatively summarizes the omics research hotspots and trends of HBV-HCC using bibliometric method, and speculates the possible future research directions in this field.

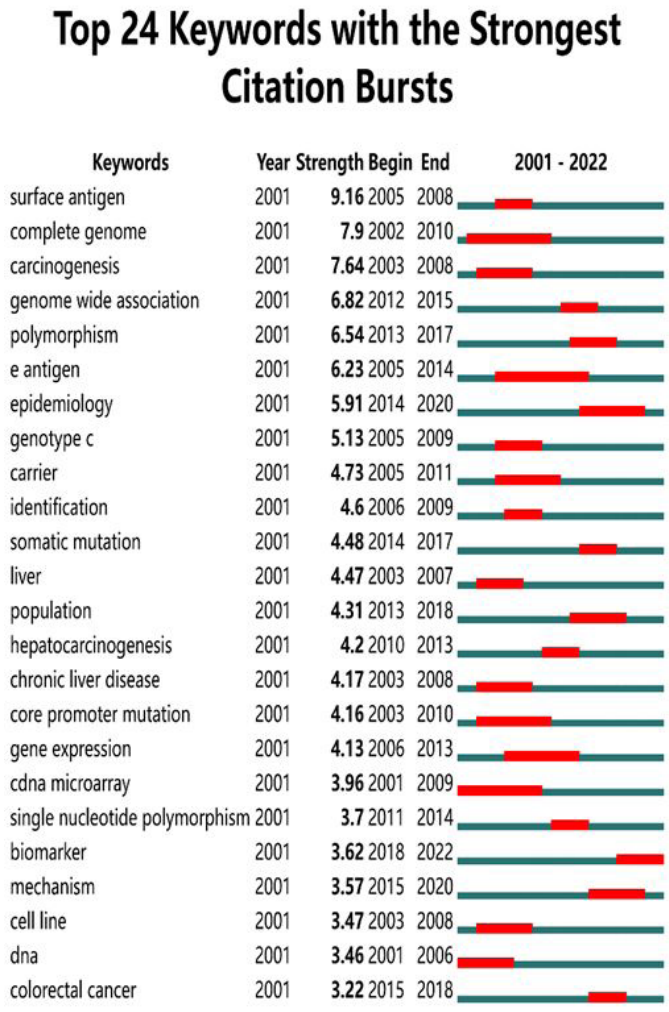

Figure 10: Keywords with the strongest citation bursts in 2001-2022

Compliance with ethical standards

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. Liu M et al (2022) Gender differences in demographic and clinical characteristics in patients with HBV-related liver diseases in China. PeerJ 10, e13828.

2. Sung H et al (2021) Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin 71, 209-49.

3. Yang JD et al (2019) A global view of hepatocellular carcinoma: trends, risk, prevention and management. Nat Rev Gastroenterol Hepatol 16, 589-604

4. Kim YJ et al (2013) A genome-wide association study identified new variants associated with the risk of chronic hepatitis B. Hum Mol Genet 22, 4233-8

5. Hanahan D, Weinberg RA (2000) The hallmarks of cancer. Cell 100, 57-70

6. Levrero M, Zucman-Rossi J (2016) Mechanisms of HBV-induced hepatocellular carcinoma. J Hepatol 64, S84-101

7. Masys DR (1989) New Directions in Bioinformatics. J Res Natl Inst Stand Technol 94, 59-63

8. Wooley JC, Varma MN (1994) Computational biology opportunity and challenges for the future. Basic Life Sci 63, 1-9

9. Kan, Z. et al (2013) Whole-genome sequencing identifies recurrent mutations in hepatocellular carcinoma. Genome Res 23, 1422-33

10. Cui D et al (2021) Advances in Multi-Omics Applications in HBV-Associated Hepatocellular Carcinoma. Front Med (Lausanne) 8, 754709

11. Ho DW et al (2021) Single-cell RNA sequencing shows the immunosuppressive landscape and tumor heterogeneity of HBV-associated hepatocellular carcinoma. Nat Commun 12, 3684

12. Jia J et al (2017) Next-generation sequencing revealed divergence in deletions of the preS region in the HBV genome between different HBV-related liver diseases. J Gen Virol 98, 2748-58

13. Kiyose H et al (2022) Comprehensive analysis of full-length transcripts reveals novel splicing abnormalities and oncogenic transcripts in liver cancer. PLoS Genet 18, e1010342

14. Klein WC, Bloom M (2005) Bibliometrics: The best available information? Soc Work Health Care 41, 117-21

15. Liu K et al (2022) Knowledge mapping and research hotspots of immunotherapy in renal cell carcinoma: A text-mining study from 2002 to 2021. Front Immunol 13, 969217

16. Wu K et al (2022) Emerging Trends and Research Foci in Tumor Microenvironment of Pancreatic Cancer: A Bibliometric and Visualized Study. Front Oncol 12, 810774

17. Pollock A. et al (2018) Stakeholder involvement in systematic reviews: a scoping review. Syst Rev 7, 208

18. Synnestvedt MB, Chen C, Holmes JH (2005) CiteSpace II: visualization and knowledge discovery in bibliographic databases. AMIA Annu Symp Proc 2005, 724-8

19. Synnestvedt MB, Chen C, Holmes JH (2005) Visual exploration of landmarks and trends in the medical informatics literature. AMIA Annu Symp Proc 2005, 1129

20. Eck N, Waltman L (2010) Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics 84, 523-38

21. Chen C (2006) CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. Journal of the American Society for Information Science and Technology 57, 359-77

22. Falagas ME, Kouranos VD (2008) Arencibia-Jorge, R. & Karageorgopoulos, D.E. Comparisonof SCImago journal rank indicator with journal impact factor. Faseb j 22, 2623-8

23. Paynter RA et al (2021) AHRQ Methods for Effective Health Care. in A Prospective Comparison of Evidence Synthesis Search Strategies Developed With and Without Text-Mining Tools (Agency for Healthcare Research and Quality (US), Rockville (MD)

24. Gao H, Huang FY, Wang ZP (2018) Research Trends of Macrophage Polarization: A Bibliometric Analysis. Chin Med J (Engl) 131, 2968-75

25. Yuan J et al (2007) Hepatitis B virus (HBV) genotypes/subgenotypes in China: mutations in core promoter and precore/core and their clinical implications. J Clin Virol 39, 87-93

26. Jiang ZS et al (2012) The effects of hepatitis B virus integration into the genomes of hepatocellular carcinoma patients. Genome Research 22, 593-601

27. Yan H et al (2015) Characterization of the genotype and integration patterns of hepatitis B virus in early- and late-onset hepatocellular carcinoma. Hepatology 61, 1821-31

28. Tamori A et al (2005) Alteration of gene expression in human hepatocellular carcinoma with integrated hepatitis B virus DNA. Clin Cancer Res 11, 5821-6

29. Tu T, Budzinska MA, Shackel NA, Urban S (2017) HBV DNA Integration: Molecular Mechanisms and Clinical Implications. Viruses 9

30. Tanaka Y et al (2006) Specific mutations in enhancer II/core promoter of hepatitis B virus subgenotypes C1/C2 increase the risk of hepatocellular carcinoma. J Hepatol 45, 646-53

31. Llovet JM, Montal R, Sia D, Finn RS (2018) Molecular therapies and precision medicine for hepatocellular carcinoma. Nat Rev Clin Oncol 15, 599-616

32. Kao JH, Chen PJ, Lai MY, Chen DS (2003) Basal core promoter mutations of hepatitis B virus increase the risk of hepatocellular carcinoma in hepatitis B carriers. Gastroenterology 124, 327-34

33. Nguyen DH, Ludgate L, Hu J (2008) Hepatitis B virus-cell interactions and pathogenesis. J Cell Physiol 216, 289-94

34. Bianca C, Sidhartha E, Tiribelli C, El-Khobar KE, Sukowati CHC (2022) Role of hepatitis B virus in development of hepatocellular carcinoma: Focus on covalently closed circular DNA. World J Hepatol 14, 866-84

35. Kim H, Kim BJ (2015) Association of preS/S Mutations with Occult Hepatitis B Virus (HBV) Infection in South Korea: Transmission Potential of Distinct Occult HBV Variants. Int J Mol Sci 16, 13595-609

36. Pollicino T et al (2004) Hepatitis B virus maintains its pro-oncogenic properties in the case of occult HBV infection. Gastroenterology 126, 102-10

37. Liu W et al (2012) Coexistence of hepatitis B surface antigen and anti-HBs in Chinese chronic hepatitis B virus patients relating to genotype C and mutations in the S and P gene reverse transcriptase region. Arch Virol 157, 627-34

38. Kim H, Lee SA, Do SY, Kim BJ (2016) Precore/core region mutations of hepatitis B virus related to clinical severity. World J Gastroenterol 22, 4287-96

39. Akrami H, Monjezi M, Ilbeigi S, Amiri F, Fattahi M (2022) The Association Between Hepatitis B Virus Mutations and the Risk of Liver Disease and Hepatocellular Carcinoma. Current molecular medicine 22, 514-23

40. Kim H, Lee SA, Kim BJ (2016) X region mutations of hepatitis B virus related to clinical severity. World J Gastroenterol 22, 5467-78

41. Ma NF et al (2008) COOH-terminal truncated HBV X protein plays key role in hepatocarcinogenesis. Clin Cancer Res 14, 5061-8

42. Sze KM, Chu GK, Lee JM, Ng IO (2013) C-terminal truncated hepatitis B virus x protein is associated with metastasis and enhances invasiveness by C-Jun/matrix metalloproteinase protein 10 activation in hepatocellular carcinoma. Hepatology 57, 131-9

43. Ferber MJ et al (2003) Integrations of the hepatitis B virus (HBV) and human papillomavirus (HPV) into the human telomerase reverse transcriptase (hTERT) gene in liver and cervical cancers. Oncogene 22, 3813-20

44. Gozuacik D et al (2001) Identification of human cancer-related genes by naturally occurring Hepatitis B Virus DNA tagging. Oncogene 20, 6233-40

45. Saigo K et al (2008) Integration of hepatitis B virus DNA into the myeloid/lymphoid or mixed-lineage leukemia (MLL4) gene and rearrangements of MLL4 in human hepatocellular carcinoma. Hum Mutat 29, 703-8

46. Pezzuto F, Buonaguro L, Buonaguro FM, Tornesello ML (2017) Frequency and geographic distribution of TERT promoter mutations in primary hepatocellular carcinoma. Infect Agent Cancer 12, 27

47. Zhang H et al (2010) Genome-wide association study identifies 1p36.22 as a new susceptibility locus for hepatocellular carcinoma in chronic hepatitis B virus carriers. Nat Genet 42, 755-8

48. Lee WY, Bachtiar M, Choo, CCS, Lee CG (2019) Comprehensive review of Hepatitis B Virus-associated hepatocellular carcinoma research through text mining and big data analytics. Biol Rev Camb Philos Soc 94, 353-67

49. Bi J, Ge S (2014) Potential roles of BMP9 in liver fibrosis. Int J Mol Sci 15, 20656-67

50. Yang F et al (2011) Long noncoding RNA high expression in hepatocellular carcinoma facilitates tumor growth through enhancer of zeste homolog 2 in humans. Hepatology 54, 1679-89

51. Alfano V, Zeisel MB, Levrero M, Guerrieri F (2021) The lncRNAs in HBV-Related HCCs: Targeting Chromatin Dynamics and Beyond. Cancers (Basel) 13

52. Qiu L et al (2017) Long Non-Coding RNAs in Hepatitis B Virus-Related Hepatocellular Carcinoma: Regulation, Functions, and Underlying Mechanisms. Int J Mol Sci 18

53. Damas ND, Fossat N, Scheel TKH (2019) Functional Interplay between RNA Viruses and Non-Coding RNA in Mammals. Noncoding RNA 5

54. Villanueva A, Newell P, Chiang DY, Friedman SL, Llovet JM (2007) Genomics and signaling pathways in hepatocellular carcinoma. Semin Liver Dis 27, 55-76

55. Jordi B, Loreto B, Margarita S, Llovet JM (2004) Focus on hepatocellular carcinoma. Cancer Cell 5, 215-9

56. Gao Q et al (2019) Integrated Proteogenomic Characterization of HBV-Related Hepatocellular Carcinoma. Cell 179, 1240

57. Liu Y et al (2020) TRIM25 promotes the cell survival and growth of hepatocellular carcinoma through targeting Keap1-Nrf2 pathway. Nat Commun 11, 348

58. Schulze K, Nault JC, Villanueva A (2016) Genetic profiling of hepatocellular carcinoma using next-generation sequencing. Journal of Hepatology, 1031-42

59. Boyault S et al (2007) Transcriptome classification of HCC is related to gene alterations and to new therapeutic targets. Hepatology 45, 42-52

60. Huang Q et al (2011) RNA-Seq analyses generate comprehensive transcriptomic landscape and reveal complex transcript patterns in hepatocellular carcinoma. PLoS One 6, e26168

61. Li W et al (2013) HIVID: an efficient method to detect HBV integration using low coverage sequencing. Genomics 102, 338-44

62. Fornari F et al (2015) Circulating microRNAs, miR-939, miR-595, miR-519d and miR-494, Identify Cirrhotic Patients with HCC. PLoS One 10, e0141448

63. Li LM, Liu ZX, Cheng QY (2019) Exosome plays an important role in the development of hepatocellular carcinoma. Pathol Res Pract 215, 152468

64. Dunn WB, Bailey NJ, Johnson HE (2005) Measuring the metabolome: current analytical technologies. Analyst 130, 606-25

65. Ferrarini A et al (2019) Metabolomic Analysis of Liver Tissues for Characterization of Hepatocellular Carcinoma. Journal of proteome research

66. Hessein M et al (2005) Hit-and-run mechanism of HBV-mediated progression to hepatocellular carcinoma. Tumori 91, 241-7

67. Vander Heiden MG (2011) Targeting cancer metabolism: a therapeutic window opens. Nat Rev Drug Discov 10, 671-84

68. Zhao P et al (2020) Expression of S100 proteins is associated with HBV intrauterine transmission. Arch Gynecol Obstet 302, 1389-99