ISSN 2616-5783 Vol.5, Issue 10: 70-82, DOI: 10.25236/AJHSS.2022.051013

Analysis and Prediction of Topic Research of Transgenic Papers Based on Knowledge Graph

Yongkang Duan

School of Public Administrations, Sichuan University, Chengdu 630065, P.R China

Abstract: *Citespace and other visualization software were used to analyze the knowledge graph of relevant pieces of literature on GM research in the past decade, and to sort out the number trend, core authors, research institutions, number of core journals published, and keyword co-occurrence graph of core research literature in gm research in the past decade. The analysis shows that the research interest in TRANSGENIC has changed in recent ten years. The research interest in the United States, China, and other countries is similar to the contribution of the sun, moon, and stars, while the major institutions led by the University of Chinese Academy of Sciences, China Agricultural University, and Harvard University are more willing to publish their research results in Plos One. Research focuses on transgenic rice, Alzheimer's disease, biochemistry, and molecular biology, in addition, future research will focus on transgenic plants, Alzheimer's disease, and other aspects.*

Keywords: *Knowledge Mapping, Transgenes, Subject Topics, Evolutionary Analysis, Evolutionary Prediction*

1. Introduction

The 14th Five-Year Plan mentions the focus on the modern seed industry and plant disease prevention and control content. on October 17, 2020, the 22nd meeting of the Standing Committee of the 13th National People's Congress adopted the Biosafety Law of the People's Republic of China, emphasizing the need to promote the output and transformation and application of the results of core key technologies and major defense products for biosafety, plus to improve the scientific and technological security capacity of biosafety. On March 12, 2021, the "Outline of the 14th Five-Year Plan and 2035 Visionary Targets for National Economic and Social Development of the People's Republic of China" was officially released, and gene and biotechnology were identified as one of the seven frontier areas of scientific and technological research for the country to strengthen national strategic scientific and technological forces and enhance original and leading scientific and technGingersoftwareological research; at the same time, gene technology was also listed as a future national strategic emerging industry. In terms of the process of issuing GM safety certificates, on December 30, 2019, the Ministry of Agriculture and Rural Affairs publicized the directory of 192 plant varieties to be approved for issuance of agricultural GMO safety certificates; in June this year, the Ministry of Agriculture and Rural Affairs again publicized the directory of 71 plant varieties to be approved for issuance of agricultural GMO safety certificates, and GM development is still ongoing.

Since its introduction, transgenic technology has developed rapidly and has been widely used in many fields in many countries, creating great value and having a huge impact on the world in the past four decades. Since the first transgenic paper was published in Web of Science in October 1980 [1] (Flatt P R, Bailey C J, 1980), to date (October 10, 2021), a cumulative total of 282813 papers or review papers are available in Web of Science. Meanwhile, during 1996-2018, GM technology generated $225 billion in economic benefits, produced 822 million tons of GM crops, saved 231 million hectares of land, reduced pesticide use by 776 million kg and CO2 emissions by 27.1 billion kg, and significantly improved the economic situation of smallholder families, resulting in improved living standards for 16-17 million farmers and their families, a total of more than 65 million people, and more than 65 million people in total [2] (ISAAA Brief 54-2018).

Technological development has simultaneously brought about too many research hotspots and topics. The article analyzes the evolution of subject topics in the field of transgenics in the WoS platform in the past decade through scientific and technical text mining and information visualization techniques and analyzes the change and development of subject topics in transgenics in the time dimension by using words as representations, which implies the alternating evolution process of knowledge from low to high and from old to new [3-4] (Davenport T H, 2003; Ma F, et al, 2013). With the expertise of Library and Information Science (LIS) disciplines, abstract and complex data are transformed into visual and vivid geometric figures, through which the relationships and evolutionary patterns between knowledge points can be easily sorted out. With the help of CiteSpace and VOSviewer software, we can generate visual charts through big data to present the whole picture of scientific information in the form of scientific knowledge mapping, and dig out the hot research and frontier direction in the field of transgenics in a comprehensive, correct and fast way, to provide a reference for future theoretical research and practical exploration.

2. Literature review

2.1 Evolutionary analysis discussion of disciplinary themes

Subject theme evolution analysis and prediction is one of the focuses of Library and Information Science (LIS) subject research. Among them, evolutionary analysis mainly parses the evolutionary path of disciplinary themes, which is an analysis of the events that have occurred and contains co-word analysis [5-6] (Wang X, et al., 2018.), citation analysis, or co-citation network [7-8] (Lucio-Arias, Leydesdorff, 2008; Wu Q, et al, 2014), theme model [9-10] (Qi Y, et al, 2016; Qu Jia-Bin, Ou Shi-Yan, 2018.) and other disciplinary theme evolution analysis studies. Subject theme evolution prediction is the prediction of future trends, such as prediction based on qualitative methods such as evolution diagram, impact diagram, evolution path, scientific knowledge map, and technology research matrix [11-13] (Zhu D, et al, 2018; Gao N, et al, 2020; Li X, et al, 2020.), and prediction based on state indicators, citation indicators, alternative quantitative methods such as metrics for prediction [14-16](Zhu Q, Leng F, 2014; Porter A L, et al,2019; Tong L, et al, 2016.).

In the temporal and spatial dimensions, science has a certain cyberspace structure and shows patterns of continuation, transfer, and convergence over time [17] (Zou X. Sh. et al, 2012.). Scientists characterize disciplinary themes by studying the dynamics of knowledge elements and the structure of the knowledge networks they constitute. Guan Peng et al [18] (Guan Peng et al., 2018.) sorted out the process of disciplinary knowledge network construction through complex network structure analysis theory and summarized the research scheme of "topic mining → network construction → structure evolution → simulation modeling". Kui Ling et al [19] (Kui Ling et al., 2016.) used the network community evolution analysis method to generate a disciplinary topic evolution network and reveal the process of disciplinary topic convergence and mutation. In the research path, scientists study the construction and changes of knowledge networks in two dimensions, time and space, furthermore explore the evolutionary laws of disciplinary themes.

In this paper, we intend to analyze the research in the whole transgenic field with the method of library and information technology and explore the past hot changes and future research centers in the transgenic field by analyzing various attributes such as authors, countries, institutions, and keywords of transgenic papers.

2.2 Development of Transgenic Technology

The technology of introducing artificially isolated and modified genes into the genome of an organism to cause heritable modifications to the traits of the organism due to the expression of the introduced genes is called transgenic technology. The terms "genetic engineering", "genetic engineering", and "genetic transformation" are all synonymous with transgenesis. In 1980, Flatt, PR (Flatt, Peter R.) et al. published a paper on genetically modified organisms (GMOs) in Further evidence for a gene-dosage effect in lean heterozygous (ob/+) mice was published in BIOCHEMICAL SOCIETY TRANSACTIONS [1] (Flatt P R, Bailey C J, 1980). In the following four decades, the development of GMO technology has been widely used in various fields around the world, especially to enhance the role of GM agricultural products in people's production and life [20] (De-Ping Wang, 2017.).

Since the world's first commercialization of transgenic maize Bt176 in 1996, transgenic maize has been widely used globally, with a cumulative planting of 7.5 × 108 hm2 in 23 years of commercialization (ISAAA, 2019). In 2018, for example, 14 of the 35 countries and regions where GM crops are grown are planted with GM maize [21] (ISAAA, 2018.). Meanwhile, the International Service for the Application of Agrobiotechnology (ISAAA) measured that in 2013, the area planted with GM crops reached 175.2 million hectares in 27 countries worldwide, which is more than a 100-

fold increase compared to 1.7 million hectares when commercialization began in 1996, with 18 million farmers planting GM crops worldwide [22] (James C, 2013).

2.3 Application of Knowledge Graph

Knowledge Graph (KG) was first developed from the semantic web, and the concept of knowledge graph was introduced by Google in 2012 [23] (Singhal A, 2012.). It is a class of graphs that shows the development process and structural relationships of knowledge by taking knowledge domains as objects. In essence, it is a knowledge base that covers the graph structure, and this storage structure then makes it possible to store efficiently the knowledge graphs based on the association relationships between data and knowledge. Nodes in a graph are used to represent entities or concepts, and edges are used to represent semantic relationships between entities or concepts, and by aggregating various types of data and connected relationships into knowledge in the form of nodes and edges, the knowledge graph can be used to achieve efficient data as well as domain knowledge retrieval through relevant graph matching algorithms [24](Cao S, Zhao B, 2021). Another advantage of knowledge graphs is their inference capability [25] (Chen X, et al, 2020), as they have the dual nature and characteristics of "graph" and "spectrum" and can represent the structural relationships between individual knowledge or It can represent the structural relationships between individual knowledge or knowledge collections (including many potential relationships such as intersection, interaction or evolution), and through these intricate and complex relationships, more new knowledge can be inferred and discovered [26] (Chen, et al., 2015). Thanks to the efficient organization and collation of knowledge in the real world by knowledge mapping, effective communication between human and machine is promoted, which makes it possible to turn abstraction into concretization, to obtain the whole picture of scientific information, and to dig out the development process, hot research and frontier direction in the corresponding knowledge and information field in a comprehensive, correct and fast way [27] (Synnestvedt M B, 2005), thus has been widely used in the field of search engines, finance [28] (Yan Z et al., 2018), education [29] (Zheng Q et al, 2019), agriculture [30] (Shi L, Yang M, 2016), etc.

The field of transgenics has always been the focus of biological research. Using the newer scientometric and information visualization method, knowledge graph, based on Web of Science, the transgenic papers of the last decade are used to visually display the specific situation of transgenic research worldwide in the last decade, aiming to present the research dynamics in the field of transgenics, visually and present the research pulse, hot changes, and development trend of the field, to provide more reliable and objective sources of information for future research scholars and practitioners in the field of transgenics, and promote the further development of the field.

3. Methods and Data

In this study, the analysis and prediction of disciplinary theme evolution based on knowledge graph includes analysis of discipline and journal contribution, analysis of research quality of countries and institutions, analysis of research impact of countries, institutions, and individuals, and analysis of research priorities.

The data of this study are based on the Web of Science core database search in the field of transgenics with TS=(transgenic or trans? genic or trans-genic), the time interval from 2011 to 2020, and the literature types are Article and Review, totaling 78,693 A total of 78,693 journal articles (retrieved on November 6, 2021). During this decade, the number of transgenic papers published each year was around 8,000 and did not fluctuate much.

The indexed English bibliography named “download_xx.txt” was imported into CiteSpace.5.8.R3c for format conversion. The data obtained from Web of Science were analyzed visually using CiteSpace software, with time slices of 2011-2020, “years per slice” of 1 year, and terminology sources selected by default. The Derwent Data Analyzer software can count the documents exported from the core database of Web of Science by year, import the corpus and categorize the papers by journal, country, author, and other attributes. The data can be used to analyze the research trends under these categories.

ISSN 2616-5783 Vol.5, Issue 10: 70-82, DOI: 10.25236/AJHSS.2022.051013

4. Results

4.1 Analysis of the publication of transgenic papers

4.1.1 Subject Analysis

Figure 1 shows the trend of subject themes under the top ten most published topics in transgenics over the last decade. From 2011 to 2017, most of the subject themes were stable with a slight decadence; only after 2017, the trend of publications resumed the upward development again. There was a certain bottleneck in 2017 as the genetic publications under each theme wilted in posting. Transgenic publications in Biochemistry & Molecular Biology and Neurosciences & Neurology have maintained a high position in the last decade, while Transgenic publications in Plant Sciences have shown a more pronounced upward trend in the last decade, while Transgenic publications in other topics have GM publications under other topics have maintained a relatively stable publication trend over the last decade and were, therefore, less affected in 2017. According to the graph, the future publication of GM-related papers is promising, and Plant Sciences, in particular, is likely to be the future center of research on GM papers.

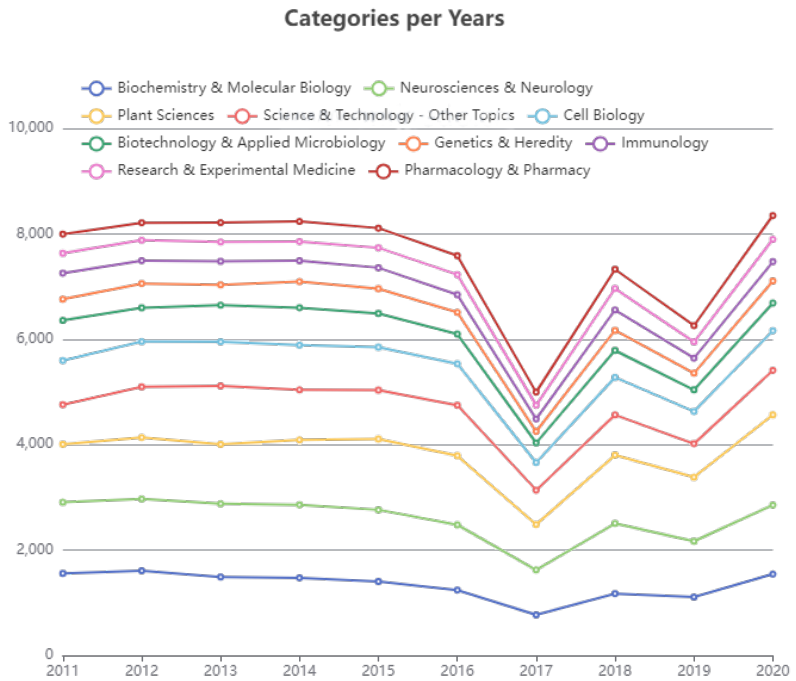

Figure 1: Contribution of disciplinary themes of transgenic papers

Figure 2 shows the analysis of disciplinary theme co-occurrence of papers with the transgenic theme in the last decade. The size of the nodes depends on the number of publications in their disciplinary themes, and the connecting line reflects the centrality of the corresponding discipline, i.e., the degree to which the relevant node is the center, the higher the centrality, the more important the node. It can be seen that BIOCHEMISTRY & MOLECULAR BIOLOGY, NEUROSCIENCES & NEUROLOGY, PLANT SCIENCES, NEUROSCIENCES have more papers published under their subject topics, and the number of papers published exceeds 10,000. At the same time, CELL BIOLOGY, AGRICULTURE, HORTICULTURE, BIOCHEMISTRY & MOLECULAR BIOLOGY, PLANT SCIENCES, BIOTECHNOLOGY & APPLIED MICROBIOLOGY and BIOCHEMICAL RESEARCH METHODS RESEARCH METHODS subject themes are more central and more connected and are the central subject themes in the transgenic papers. Among them, BIOCHEMISTRY & MOLECULAR BIOLOGY, PLANT SCIENCES has more than 10,000 articles published and can be called a key core node. The number of articles published in MICROBIOLOGY and BIOCHEMICAL RESEARCH METHODS is more than 1,000, which can be called secondary core nodes, and the number of articles published in HORTICULTURE is less than 1,000, which can be called general core nodes. Correspondingly, VIROLOGY, ENTOMOLOGY, SCIENCE & TECHNOLOGY - OTHER TOPICS, MULTIDISCIPLINARY SCIENCES, and other disciplinary themes have very low centrality and do not communicate and interact with other disciplinary themes that are the subject of transgenic papers.

ISSN 2616-5783 Vol.5, Issue 10: 70-82, DOI: 10.25236/AJHSS.2022.051013

In between, VIROLOGY, ENTOMOLOGY, and other subject topics have very low publication volume, only hundreds of articles, and are niche independent nodes. In contrast, SCIENCE & TECHNOLOGY - OTHER TOPICS, MULTIDISCIPLINARY SCIENCES, and other subjects have a high publication volume of more than 8,000 articles, which is a mass-independent node.

This shows that the research of papers within the field of transgenics is concentrated in the fields of biochemistry and molecular biology, neuroscience and neurology, plant science, and neuroscience. Among a host of research fields cell biology, agriculture, horticulture, biochemistry, and molecular biology, plant science, biotechnology, and applied microbiology, and biochemistry and biochemistry can intersect with other fields and will lead to more scientific collaborations.

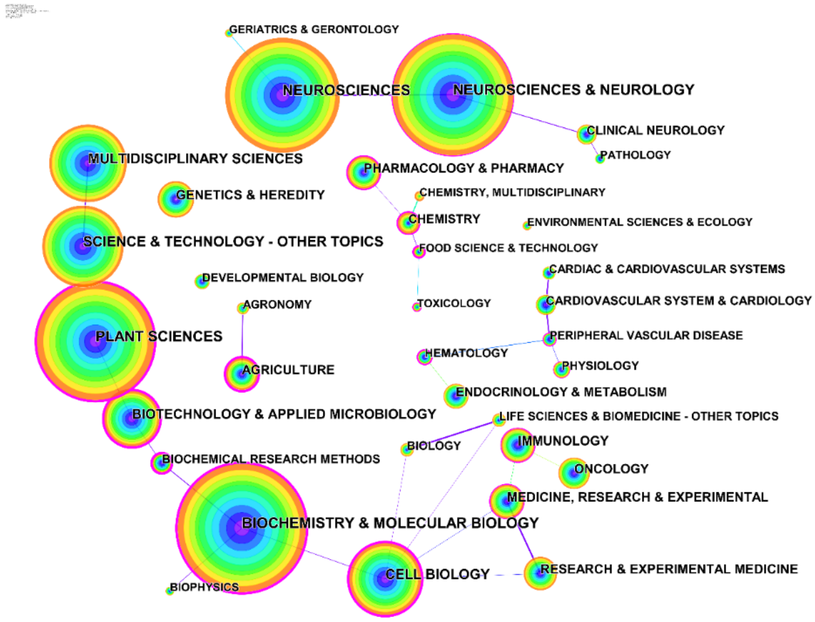

Figure 2: Co-presentation analysis of transgenic paper disciplines

4.1.2 Journal Analysis

Figure 3 illustrates the publication of transgenic papers in journals over the last decade, showing the eight journals with the highest number of published articles. It can be seen that in most cases researchers are significantly more comfortable publishing their transgenic papers in Plos One, which is a good platform for papers and research in the transgenic field. This is followed by Scientific Reports, but still with a significant difference in performance from Plos One. In contrast, the other journals are not performing as well.

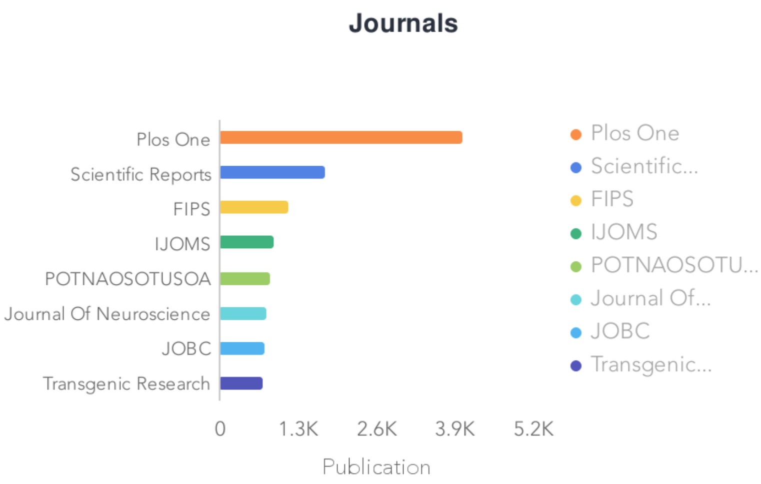

Figure 3: Annual contribution of transgenic papers journals

ISSN 2616-5783 Vol.5, Issue 10: 70-82, DOI: 10.25236/AJHSS.2022.051013

4.2 Research quality analysis of transgenic papers

4.2.1 National level

Figure 4 illustrates the changes in the publication in each country in the field of transgenics in the last decade. As shown in the figure, the publication situation in China shows a linear increase, with the US leading until 2017, but showing a clear decline until 2018 when China overtook the US for the first time, after which it became the country with the most publications in the field of transgenics. In 2011, the contribution of China and Japan was similar, but as time went by, the gap gradually widened, and now the amount of China's publication is five times that of Japan. The contribution of China and the U.S. is outstanding, but the contribution of other countries is low, indicating that China and the U.S. are expanding more widely and deeply in the field of transgenics, showing a polarization phenomenon in the world.

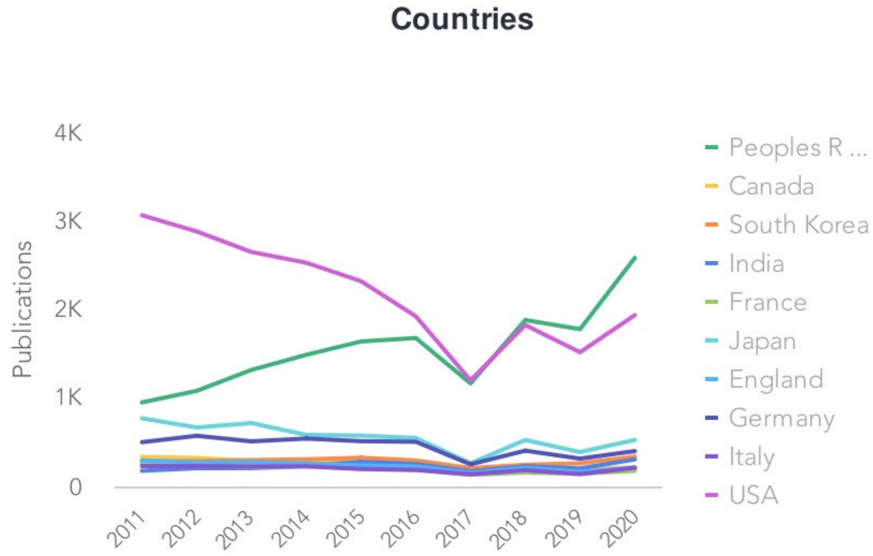

Figure 4: The most productive country or region in the field of genetic modification

4.2.2 Institutional level

Figure 5 shows the number of articles published by the ten most published institutions in the field of transgenics in the last decade. As can be seen, the largest number of articles is published by the UNIVERSITY OF CALIFORNIA SYSTEM, with 3,470 articles, in other words, 4.41% of the corpus is published by this institution. This is followed by the CHINESE ACADEMY OF SCIENCES, with 2,210 articles, accounting for 2.81% of the corpus. Interestingly, four U.S. institutions and one Chinese institution are in the top five in terms of the number of publications, while in the top ten, there are five U.S. institutions, two Chinese institutions, and one each from French, German, and British institutions. This also has some citations for the results in Figure 4.

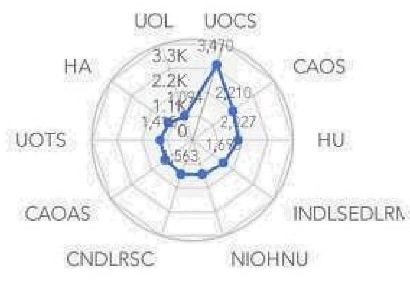

Figure 5: Contribution of the most productive institutions publishing papers in the field of transgenics

ISSN 2616-5783 Vol.5, Issue 10: 70-82, DOI: 10.25236/AJHSS.2022.051013

4.3 Research Impact Analysis of Transgenic Papers

4.3.1 National level

Figure 6 shows the country/organization co-occurrence analysis of papers on the topic of transgenics in the last decade. There are 44 country nodes, and the size of the nodes depends on the number of papers published in their countries/organizations, and the connecting line reflects the centrality of the corresponding country/organization, i.e., the degree of the relevant node as the center, and the higher the centrality, the more important the node is. As can be seen, the United States leads the way with 30,727 papers published, accounting for nearly 40% of the total corpus. China follows with 17,161 papers, accounting for over 20% of the corpus. This is followed by Japan, Germany, the UK, Canada, France, and Korea. From a centrality point of view, the UK is at the center of the collaborative network, followed by Germany and Denmark. The United States, with the highest number of publications, is in sixth place, and China, with the second-highest number of publications, is at the bottom.

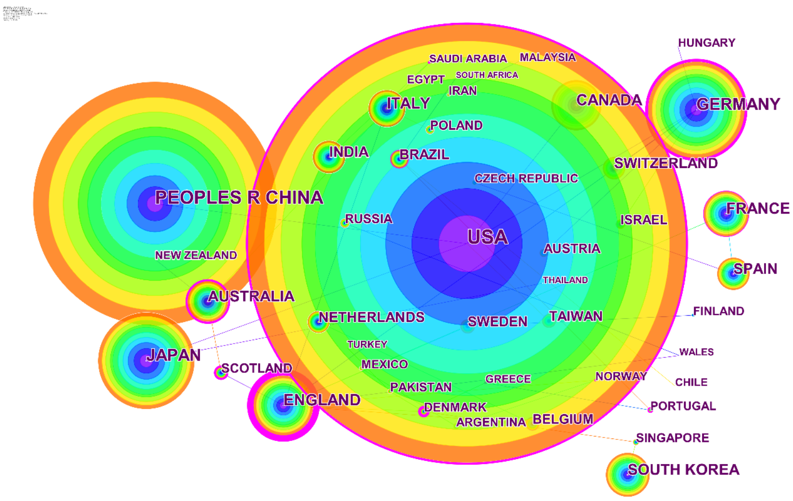

Figure 6: Country co-occurrence analysis of papers published in the field of transgenics

4.3.2 Institutional level

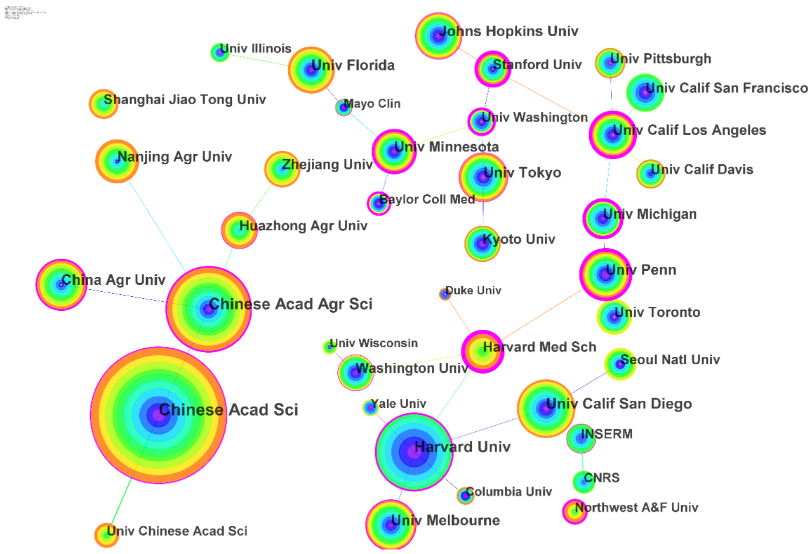

Figure 7: Co-presentation analysis of GM paper publishing institutions

Figure 7 shows the analysis of the co-occurrence of publishing institutions for papers with transgenic themes in the last decade. 44 institutional nodes are shown, and the size of the nodes depends on the number of articles published by their publishing institutions, with the smallest node, Univ N Carolina, publishing 120 articles and the largest node, Chinese Acad Sci, publishing 1930 articles. The linkage between the nodes reflects the centrality of the corresponding publishers, i.e., the degree of centrality of the node in question, the higher the centrality, the more important the node. It

ISSN 2616-5783 Vol.5, Issue 10: 70-82, DOI: 10.25236/AJHSS.2022.051013

can be seen that Chinese Acad Sci, Chinese Acad Agr Sci, and Harvard Univ have significantly more publications than other nodes, with more than 1,000 articles. In terms of centrality, Harvard Med Sch, Univ Calif Los Angeles, Stanford Univ, Univ Michigan, Univ Penn, Univ Washington are more central and located at the center of the collaborative network. Univ N Carolina, NCI, and Ohio State Univ are less central and rarely interact with other institutions.

4.3.3 Individual level

Table 1 shows the ranking of first-author publications of papers with transgenic topics in the last decade, and the top 20 authors are shown in the table. The top 20 authors are shown in the table. Wang Y is the first author with 116 publications, followed by Zhang Y with 102 publications. A total of 485 authors with more than 10 first-author publications were counted in the corpus, but only 2 authors with more than 100 publications and 12 authors with more than 50 publications are all experts in the field of transgenics.

Table 1: Contributions of the most cited authors of transgenic papers from 2011-2020

Rank	author	Number of articles issued	Rank	author	Number of articles issued
1	Wang, Y	116	11	Liu, J	52
2	Zhang, Y	102	12	Wang, X	51
3	Wang, L	91	13	Lee, H J	49
4	Li, Y	90	14	Zhou, Y	49
5	Wang, J	85	15	Kim, J	46
6	Liu, Y	85	16	Zhang, X	46
7	Zhang, L	84	17	Chen, Y	45
8	Zhang, J	76	18	Wang, F	43
9	Li, J	75	19	Wang, H	43
10	Kim, H J	58	20	Liu, X	43

Table 2 shows the top ten most-cited transgenic papers in the last decade, as can be seen in the article The amyloid hypothesis of Alzheimer's disease at 25years by Selkoe, DJ, which has been cited up to 2416 times since its publication in EMBO MOLECULAR MEDICINE in June 2016; followed by Neuroinflammation in Alzheimer's disease by Heneka, MT in LANCET NEUROLOGY in 2015. The next most popular article in the last decade is Neuroinflammation in Alzheimer's disease at 25years, published in EMBO MOLECULAR MEDICINE in June 2016, with 2,416 citations, followed by Heneka, MT's Neuroinflammation in Alzheimer's disease in LANCET NEUROLOGY in 2015, with 2,343 citations. In addition, the United States contributed 40% of the most popular transgenic papers in the last decade, with the remaining six from China, Germany, Pakistan, Australia, Israel, and the United Kingdom, respectively.

Table 2: Contribution of the most cited papers on transgenic papers

Rank	Paper Title	Author	Cited
1	THE AMYLOID HYPOTHESIS OF ALZHEIMER'S DISEASE AT 25YEARS	SELKOE, DJ	2416
2	NEUROINFLAMMATION IN ALZHEIMER'S DISEASE	HENEKA, MT	2343
3	SLEEP DRIVES METABOLITE CLEARANCE FROM THE ADULT BRAIN	XIE, LL	1874
4	SLEEP DRIVES METABOLITE CLEARANCE FROM THE ADULT BRAIN	DELMORE, JE	1839
5	THE PRO-AND ANTI-INFLAMMATORY PROPERTIES OF THE CYTOKINE INTERLEUKIN-6	SCHELLER, J	1755
6	PHYTOREMEDIATION OF HEAVY METALS-CONCEPTS AND APPLICATIONS	ALI, H	1710
7	NEUROVASCULAR PATHWAYS TO NEURODEGENERATION IN ALZHEIMER'S DISEASE AND OTHER DISORDERS	ZLOKOVIC, BV	1494
8	AMYOTROPHIC LATERAL SCLEROSIS	KIERNAN, MC	1475
9	A UNIQUE MICROGLIA TYPE ASSOCIATED WITH RESTRICTING DEVELOPMENT OF ALZHEIMER'S DISEASE	KEREN-SHAUL, H	1335
10	ASCORBATE AND GLUTATHIONE: THE HEART OF THE REDOX HUB	FOYER, CH	1288

Figure 8 shows the author's co-occurrence analysis of papers with transgenes as the topic in the last decade. 54 author nodes are shown, and the size of the author nodes depends on their publication volume, with the largest nodes being Wang Y and Zhang Y, both with 652 publications, and the smallest node, Liu C, with 45 publications. The linkage between nodes reflects the centrality of the

ISSN 2616-5783 Vol.5, Issue 10: 70-82, DOI: 10.25236/AJHSS.2022.051013

corresponding author, i.e., the degree to which the relevant author node is the center; the higher the centrality, the more important the author node is. It can be seen that Wang Y and Zhang Y have significantly more nodes than others, followed by Li Y and Zhang L. This proves that they have contributed more articles than other authors, especially Wang Y, Zhang Y, and Li Y who have more than 500 articles. From the centrality point of view, Zhang Y has not only ranked first in terms of the number of articles published but is also located in the center of the collaborative network, which is a well-deserved leader. They are followed by Xu Y, Li X, Liu Y, and Wang L, who are also located in the sub-center of the collaboration. Compared to Wang Y's centrality performance is significantly inferior.

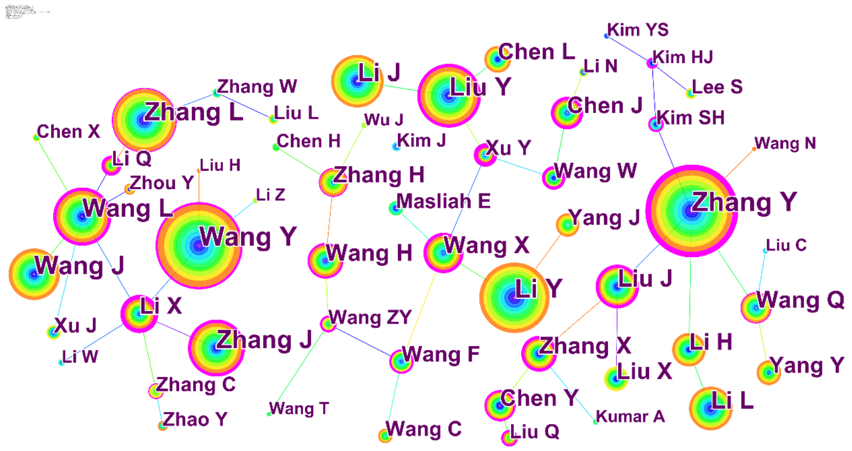

Figure 8: Author co-citation analysis of transgenic papers

Figure 9 shows the author's co-citation analysis of transgenic papers in the last decade. 45 author nodes are shown, and the size of the author node indicates the number of citations, with the largest node being LIVAK KJ cited 1683 times, and the smallest node CHEN J cited 100 times. The linkage between nodes is a reflection of node centrality, which represents the connection with others, the higher the centrality the higher the chance of being co-cited, the closer the connection with other author nodes, and the greater the role of substitution intermediary. It can be seen that the size of the LIVAK KJ node is significantly larger than the others, followed by CLOUGH SJ, MURASHIGE T, ZHANG Y, WANG Y, SELKOE DJ, all of whom have been cited more than 1,000 times. In terms of centrality, the author nodes LI J, WANG J, CLOUGH SJ, MURASHIGE T, HENEKA MT, and SELKOE DJ have more connecting lines, indicating higher centrality, closer connection with other author nodes, and more likely to be co-cited.

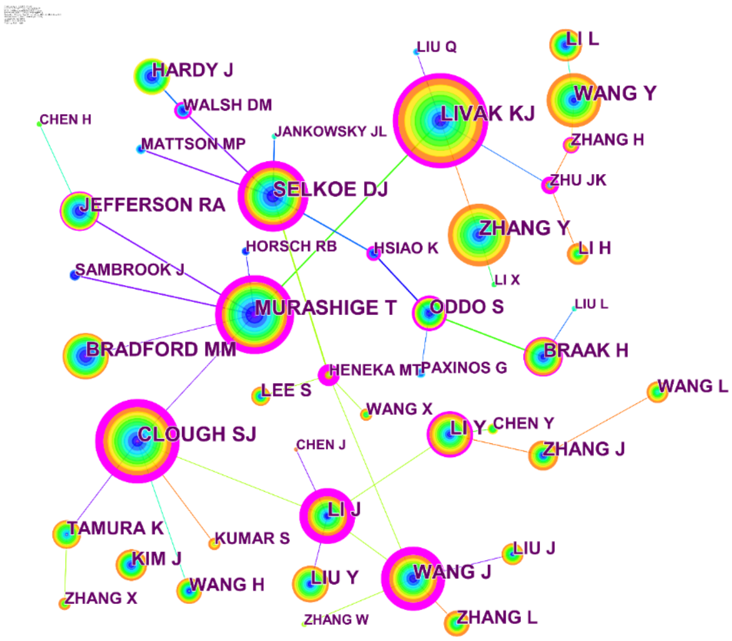

Figure 9: Co-citation analysis of authors of transgenic papers

ISSN 2616-5783 Vol.5, Issue 10: 70-82, DOI: 10.25236/AJHSS.2022.051013

4.4 Analysis of the research focus of transgenic papers

Figure 10 shows the keyword co-occurrence analysis of papers on the topic of transgenics in the last decade, with 79 keyword nodes appearing in the figure. Similarly, the higher the frequency of the keywords, the larger the nodes, with the largest node Alzheimers disease appearing 4366 times and the smallest nodes such as beta-amyloid, cancer, and calcium appearing 37 times. The centrality of a node is determined by the connection between the nodes, which reflects the importance of the node, and the greater the centrality, the more important the node is as an intermediary for other nodes. It can be seen that Alzheimer’s disease, transgenic mice, inflammation, neurodegeneration are among the big nodes with high frequency and are the hot spots of research in the transgenic field. Meanwhile, Alzheimer’s disease, oxidative stress, mitochondria, amyloid-beta, arabidopsis, reactive oxygen species are words with high co-occurrence with other nodes, with many connections and good centrality performance, which are often used as intermediaries for research and are the focus of research in the field of transgenics.

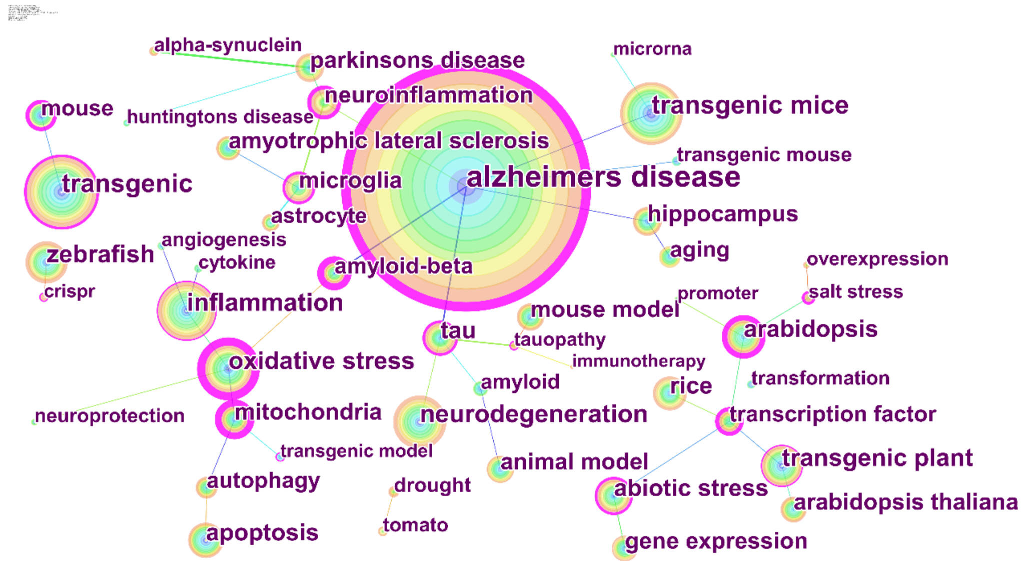

Figure 10: Keyword co-occurrence analysis of GM papers

Figure 11 shows the clustering analysis of the keywords of the papers on transgenic topics in the last decade, i.e., the keywords shown in Figure 12 were clustered into eight categories, namely: rice, neuroinflammation, CRISPR, Alzheimer’s disease, salt stress, tau, They reveal the research areas within the field of transgenics and provide a simple classification of the content of transgenic research.

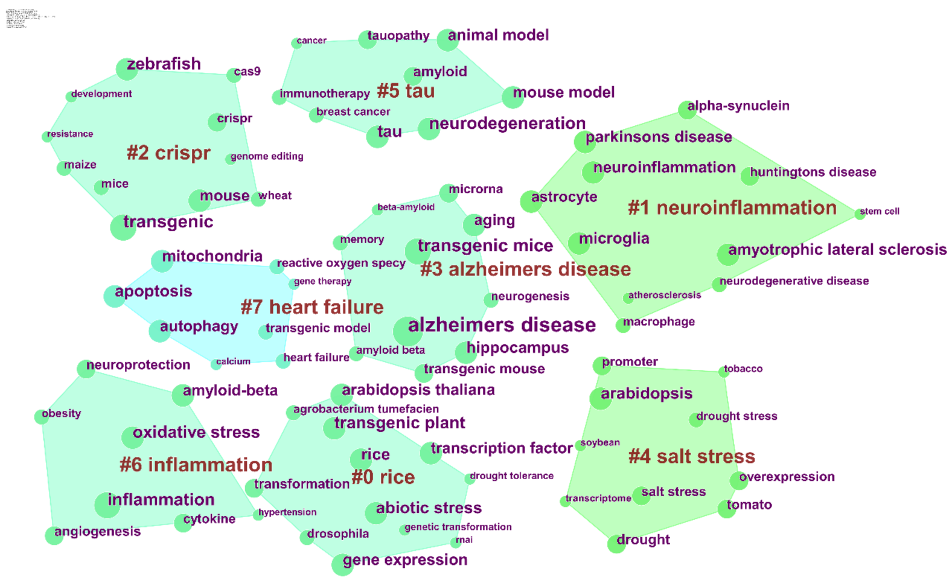

Figure 11: Clustering analysis of keywords of transgenic papers

ISSN 2616-5783 Vol.5, Issue 10: 70-82, DOI: 10.25236/AJHSS.2022.051013

5. Discussion

5.1 Conclusions

(1) From 2011 to 2020, there has been a polarization in the publication of GM papers worldwide, led by the United States, followed by China, and less by other countries, showing a pattern of sun, moon, and stars. China's publication rate has grown rapidly over the past decade, and in 2017, the annual publication volume surpassed that of the United States to become the first, and the future development situation is promising.

(2) From 2011 to 2020, institutions worldwide have placed different levels of importance on transgenic papers, also reflecting the importance of the country to the field of transgenics, led by the University of California system with 3,470 articles published in the decade, in other words, it contributed 4.41% of the core articles on transgenics in Web of Science in the decade. This is followed by the Chinese Academy of Sciences system with 2,210 articles over the decade and Harvard University with 2,027 articles. Of the ten institutions with the most publications, the U.S. institutions hold half of the seats, with two Chinese institutions and one each from France, Germany, and the United Kingdom. At the same time, institutions such as Harvard University, with their high volume of publications and their central position in the network of institutional collaborations, have good prospects for future research. These institutions are the core institutions in the field of transgenics and are the important support and carriers of transgenic research.

(3) From 2011 to 2020, different authors around the world share the same insights on transgenic research. Wang, Y, and Zhang, Y, for example, have published more than 100 articles as first authors over the decade, with an average of more than 10 core transgenic papers published each year. They were followed by 12 authors with more than 50 first-author publications and 485 authors with more than 10 first-author publications. All of these authors are experts in the field and have contributed positively to the advancement of the GM field.

During this period, two of the 78,693 core papers on transgenics have been published with more than 2,000 citations to date, and the top ten cited papers are from around the world, with first-, second-, and third-world authors all contributing to the advancement of the field of transgenics.

(4) From 2011 to 2020, core transgenic papers worldwide prefer to be published in Plos One, followed by Scientific Report; and interestingly in citation its more willing to cite papers in P NATL ACAD SCI USA, NATURE and SCIENCE. These journals are important platforms for research in the field of transgenics, which is constantly advancing.

(5) From 2011 to 2020, transgenic research worldwide shows a multidisciplinary and multi-focused dynamic. With the joint development of multiple disciplinary themes, plant science, biochemistry, and molecular biology have shown strong vitality in recent years, with a surge in publication rate, and may become the new wave in the future; meanwhile, biochemistry and molecular biology, for example, due to their strong disciplinary interactions, are more easily combined with other disciplines for research and may rub off new sparks.

5.2 Research Prospects

Based on the visualization information of the knowledge map, this paper proposes the following outlooks for the future research directions of transgenics.

(1) Focus on the direction of high yield development of transgenic plants. We use transgenic rice, transgenic soybean, and other transgenic crops as research samples to explore how to develop transgenic crops with high yields and thus alleviate the global food shortage problem. The world today is in the midst of unprecedented changes, and only by breaking through and mastering core technologies will we be able to develop in the future without being constrained by others.

(2) Focus on the direction of Alzheimer's disease treatment. Alzheimer's disease has always been one of the problems in the medical field, and various medical drugs are only effective in slowing down the deterioration of Alzheimer's disease, and it is difficult to reverse the disease process. It will be a hot topic in the future to explore what kind of drugs can cure Alzheimer's disease and what kind of protein properties can effectively target the disease from a transgenic perspective.

(3) Focus on the development of transgenic interdisciplinary research. For example, linking biochemistry and molecular biology to other areas such as inflammation, neurodegeneration, oxidative

ISSN 2616-5783 Vol.5, Issue 10: 70-82, DOI: 10.25236/AJHSS.2022.051013

stress, animal experiments, etc., open up new pathways to solving problems in new areas. As the exchange and integration of disciplines continue to advance, more and more interdisciplinary cross-disciplinary research and disciplines will emerge in the future. In terms of clustering, the keywords appearing in the papers, the hot spots of transgenic papers are transgenic rice and various diseases such as Alzheimer's disease, neuroinflammation, heart failure, etc. The key to future transgenic breakthroughs will also appear in these areas. Alzheimer's disease and other keywords often appear at the same time, proving themselves to have strong interdisciplinary research prospects, and collaboration with more disciplines may lead to breakthroughs in the future.

This article also has some shortcomings in the research process, for example, while discussing genetically modified publications, it failed to contact the research results and resumes of the Nobel Prize winners. In future research, breakthroughs in the industry can be linked to publications to explore the relationship between them.

References

[1] Flatt P R, Bailey C J. Further evidence for a gene-dosage effect in lean heterozygous (ob/+) mice.[J]. Biochemical Society transactions,1980,8(5):

[2] ISAAA. Global Status of Commercialized Biotech/GM Crops in 2018

[3] Davenport T H, Prusak L, Wilson H J. Who's Bringing You Hot Ideas and How Are You Responding?[J].Harvard Business Review,2003,81(2):58-64.

[4] Ma Feicheng, Liu Xiang. Evolvement Model for Scientific Knowledge Networks[J]. Systems Engineering-Theory & Practice,2013,33(2):437-443.

[5] Wang X, Cheng Q, Lu W. Analyzing Evolution of Research Topics with NEViewer: A New Method Based on Dynamic Co-word Networks[J].Scientometrics,2014,101(2):1253-1271.

[6] Ji L, Liu C, Huang L, et al. The Evolution of Resources Conservation and Recycling over the Past 30 Years: A Bibliometric Overview[J].Resources Conservation & Recycling,2018,134:34-43.

[7] Lucio-AriasD, Leydesdorff L. Main-path Analysis and Path-dependent Transitions in HistCiteTM-based Historiograms[J].Journal of the Association for Information Science & Technology, 2008,59(12):1948-1962

[8] Wu Q, Zhang C, Hong Q, et al. Topic Evolution Based on LDA and HMM and Its Applicationin Stem Cell Research[J]. Journal of Information Science,2014,40(5):611-620.

[9] Qi Yashuang, Zhu Na, Zhai Yujia. A Comparative Studyon Topic Heats Evolutionin the Field of Information Science Between the Domestic and Foreign Research Based on DTM[J].Library and Information Service,2016,60(16):99-109.

[10] Qu Jia-Bin,Ou Shi-Yan. Analysis of subject matter evolution based on subject matter filtering and subject matter association [J]. Data Analysis and Knowledge Discovery,2018,2(01):64-75.

[11] Zhu Donghua, Wan Dong, Wang Xuefeng, et al. Evolutionary Path Analysis and Forecasting Based on NSFC Funded Topics——A Case Study of Science and Technology Management and Policy[J]. Journal of Beijing Institute of Technology(Social Sciences Edition),2018,20(2):51-57.

[12] Gao Nan, Peng Dingyuan, FuJunying, et al. Research on Technology Fronts Prediction Based on Patent IPC Classification and Text Information: Taking the Field of Artificial Intelligence as an Example[J].Information studies:Theory&Application,2020,43(4):123-129.

[13] Li Xiaosong, Lei Shuai, Liu Tian. Research on General Idea of Front Technologies Forecasting Based on IRD[J]. Information studies: Theory & Application,2020,43(1):56-60.

[14] Zhu Qingsong, Leng Fuhai. Topic Identification of Highly Cited Papers Based on Citation Content Analysis[J]. Journal of Library Science in China,2014,40(1):39-49.

[15] Porter A L, Garner J, Carley S F, et al. Emergence Scoring to Identify Frontier R&D Topics and Key Players[J].Technological Forecasting and Social Change,2019,146:628-643.

[16] Tong,Li, Wen,Xu, Chunjiang, Tao,Limin.Quantitative analysis of research papers on enterprise information resource management in China from 1990 to 2010[J]. Journal of Modern Information,2011,31(07):102-105.

[17] Xiaoshun Zou, Xiaofen Wang, Luohua Deng. Applied mathematics for library intelligence: mathematical methods in knowledge organization, discovery and utilization [M]. Beijing: National Library Press, 2012: 26-27.

[18] Guan Peng, Wang Yifen, Cao Jiajun. A Study on the Construction and Evolutionary Analysis Framework of Disciplinary Knowledge Network Integrating Subjects [J].Information Science, 2018, 36 (9) : 3-8.

[19] Kui Ling, Xu Haiyun, Hu Zhengyin, et al. Multi-model Identification and Prediction of Evolutionary Paths of Academic Subjects[J]. Library and Information Service, 2016, 60 (7) : 71-81.

ISSN 2616-5783 Vol.5, Issue 10: 70-82, DOI: 10.25236/AJHSS.2022.051013

[20] De-Ping Wang. Current Status and Future Strategies for Development of Transgenic Plants in China[J]. Journal of Integrative Plant Biology,2007(09):1281-1283.
[21] ISAAA. Global Status of Commercialized BiotechIGM Crops in 2018: Biotech Crops Continue to Help Meet the Challenges of lncreased Population and Climate Change[R]. ISAAA Brief No.54.ISAAA: Ithaca, NY,2018.
[22] James C. Global Status of Commercialized Biotech/GM Crops:2013[R]. ISAAA Brief No.46.ISAAA: lthaca, NY,2013..
[23] Singhal A. Introducing the knowledge graph: things, not strings. [2012-05-16]. https://www. blog. google/products/search/introducing-knowledgegraph-things-not/
[24] Cao Shujin, Zhao Bang. Construction and application of knowledge graph for innovative content of academic papers [J]. Journal of Modern Information,2021,41(12):28-37.
[25] CHEN X, JIA S, XIANG Y. A review. Knowledge reasoning over knowledge graph[J]. Expert Systems with Applications, 2020,141:112948.
[26] Chen Yue, Chaomei Chen, Zeyuan Liu, et al. Methodological functions of CiteSpace knowledge graphs[J]. Studies in Science of Science, 2015, 33(2): 242-253.
[27] Synnestvedt M B, Chen C, Holmes J H. CiteSpace II: visualization and knowledge discovery in bibliographic databases[J]. AMIA Annu Symp Proc, 2005, 2005: 724-728
[28] Yan Ziming, Du Debin, Liu Chengliang, Gui Qinchang, Yang Wenlong. A visual analysis of knowledge mapping in Western innovation geography research[J].Acta Geographica Sinica, 2018,73(02):362-379.
[29] Zheng Qinghua, Dong Bo, Qian Buyue, Tian Feng, Wei Penfan, Zhang Weizhan, Liu Jun. Current status and development trend of intelligent education research[J]. Journal of Computer Research and Development,2019,56(01):209-224.
[30] Shi LJ, Yang ML. Scientometrics-based research hotspots of agricultural engineering in China [J]. Transactions of the Chinese Society of Agricultural Engineering,2016,32(S2):430-438.