Journal of Scientific & Industrial Research
Vol. 80, May 2021, pp. 428-437

Technology Forecasting based on Topic Analysis and Social Network Analysis: A Case Study Focusing on Gene Editing Patents

Jia Liu[1], Jiaqi Wei[2]* and Yuqin Liu[3]

[1]State Key Laboratory of Media Convergence and Communication (Communication University of China), Beijing 100 024, People's Republic of China

[2]Department of Business Administration, China University of Political Science and Law, Beijing 100 088, People's Republic of China

[3]Department of Journalism and Publishing, Beijing Institute of Graphic Communication, Beijing 102 600, People's Republic of China

Received 19 September 2020; revised 25 December 2020; accepted 16 February 2021

Technology forecasting research is an indispensable means to master the development trend of technology and provide decision support for scientific research management. For patent documents, it does not provide keyword information, which makes the keyword based technology prediction method have some limitations in revealing research content and hidden topics in specific fields. In order to better reflect the technical information in the patent, this paper combines topic analysis and social network analysis to study the development trends of gene editing technology. First, the patent data of gene editing technology is collected from Derwent Innovations Index. Secondly, text mining software is adopted to draw a network graph of topic words, combined with Inverse Document Frequency (IDF) to construct a weighted adjacency matrix, and Social Network Analysis is used to obtain the degree of centrality of technical topic words. Finally, the technological trends of gene editing technology is explored by identifying the core themes of gene editing, highlighting themes and emerging themes, and some meaningful conclusions are also obtained. Based on the analysis results, this study finds that the development of gene editing technology is limited by factors such as ethics, law and cellular pollution. In addition, future research directions will be more inclined to optimize the safety and efficiency of gene editing technology.

Keywords: Gene editing, Patent analysis, Technology trend, Topic word network

Introduction

Technology forecasting is considered to be important means in tracking the technological trend and future development.[1] Currently, the intensified complexity of emerging technologies and fierce market competition force corporate decision makers to choose the key technologies that may emerge in the future as early as possible. Meanwhile, changes in the social environment also bring new pressure on technology forecasting research, and a large number of researchers participate in the discussion of technology development to help government and enterprise gain a first-mover advantage in the competitive environment.[2]

As one of the most valuable indicators of technology forecasting, patents record clear technical information and R&D activities. By analyzing patent content, researchers can obtain useful information about future technologies to help identify future research directions.[3] Nevertheless, with the massive increase of patent data, the Delphi, Growth curves, patent maps and other traditional methods cannot fully represent the changing characteristics of technology. Therefore, the emergence of text mining techniques makes people start to predict the future technology changes from the patent content, especially the keyword-based semantic analysis method is increasingly utilized as an efficient tool to forecast technology development.[4,5]

Although keyword-based technology forecasting methods are widely used, in the patent literature, due to the lack of keywords provided by patent applicants, this method still performs poorly in revealing the development trend of specific technical topics.[6] In addition, relying on expert knowledge to determine keyword information will also significantly increase research costs. To solve these problems, this paper combines topic analysis and social network analysis to track and predict the development of gene editing technology. This method uses text mining tools to extract topic words and determines the co-occurrence network of topic words, thereby reducing the workload of experts and reducing the impact of expert

---

*Author for Correspondence
E-mail: sjzwjq@163.com

knowledge differences on the results. Under the premise of considering the weight of the topic words, the importance of the topic words in the network is judged through Social Network Analysis, so as to provide more accurate and valuable development information for specific technical fields.

Related Work

Topic analysis and social network analysis have been applied by several authors in the research of technology forecasting.

Much effort has been devoted to adopt the main topics disclosed in the patent documents to forecast the future direction of the technology.[7] In early studies, the method of combining co-word analysis and bibliometric analysis was adopted in topic analysis and was considered a powerful tool to track the rise and fall of topics in different sub-domains.[8] Zhang *et al.* utilized topic clustering and Technology Roadmapping (TRM) to forecast general trends in technology development, which improved the efficiency of R&D management and scientific decision-making for R&D managers and policy makers.[9] Chen *et al.* proposed a topic-based technology forecasting approach to quantitatively estimate the developing trends of the discovered topics.[10]

In terms of social network analysis, Liu *et al.* combined the knowledge of bibliometrics and statistics with social network analysis, summarizing the evolutionary trend of nanotechnology research hotspots.[11] Chen's team used patent analysis and social network analysis in the evolution of dye-sensitized solar cell technology.[12] In addition, in the study of Park and Jun, they proposed a new statistical technique analysis method that combines social network analysis with time series clustering to discover the evolution of 3D printing technology in terms of innovation and sustainability.[13]

However, research on the combination of topic analysis and social network analysis to predict the development trend of technology is still very limited. Therefore, we propose a new method for technical topics forecasting and construct a weighted adjacency matrix for the technical topic words in the patent. By mining the relationship and dynamic relationship between technical subject words from network structure, the application of technology forecasting method in practice is improved.

Materials and Methods

This paper proposes a hybrid method which integrates social network method with patent topic analysis to determine the future trends of gene editing technology. Topic analysis is firstly adopted to extract key topic words of gene editing technology. Then social network analysis is introduced to conduct a quantitative study on the weighted technical topic word co-occurrence matrix, and finally the key technologies in each stage are determined.

Data Collection

The data in this article is based on the Derwent Innovations Index, which covers more than 50 million patents in specific countries and sources, and can obtain English translation patent records in 30 languages from more than 50 patent issuing organizations.[14] It provides high-quality patent documents for this study, ensuring the reliability of research results. Using the basic search function of the Derwent Innovations Index search platform, we searched all relevant records before 2020 with the theme of 'Gene editing', and gained 2582 patent documents. The results were saved in specific text format for subsequent analysis.

Identification Technology Development Stage

Considering that the evolution of gene editing technology is a dynamic process, we apply technology life cycle (TLC) to divide different stages of development. TLC is one of the commonly used methods to assess the maturity of technology. Since most technologies need to go through a certain development cycle, TLC is often utilized in conjunction with patent analysis to identify different stages of technologies. Meanwhile, TLC is very effective in investigating patent activities and supporting R&D strategies making.[15] As show in Fig. 1(a), Liu reflected the active degree of technology research through the number of patents, and divided the product technology life cycle into seven research stages according to the turning point of the curve.[16] In this paper, we use Liu's method to divide the life cycle of gene editing technology.

From the data collected in this article, we found that the development of gene editing technology has not yet experienced a complete life cycle. Within 16 years from 1985 to 2000, the number of published patents had been slowly increasing, and the cumulative number of patents that could be retrieved by 2000 is only 30. However, between 2001 and 2002, the number of patents related to gene editing had increased significantly, and more than 90 patents had been published in just two years. After that, the trend of patent application for gene editing had

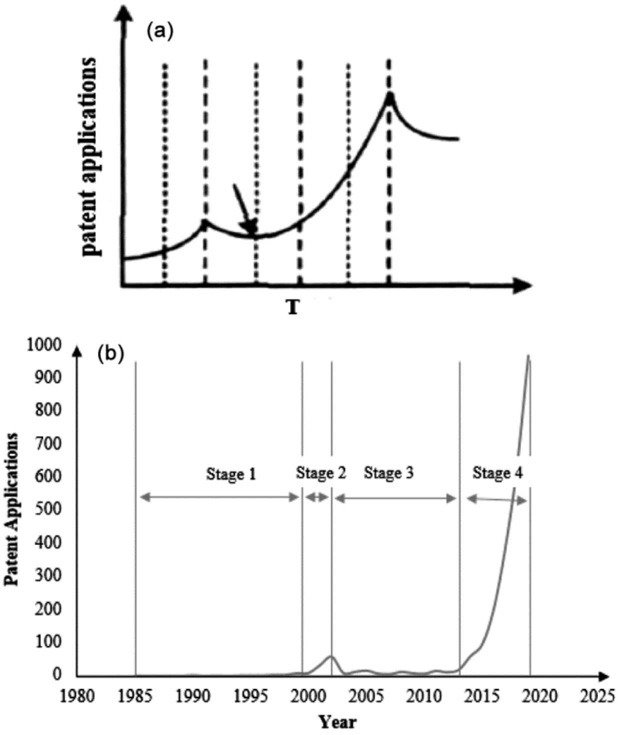

Fig.1 — (a) Technology life cycle diagram; (b) Division of development stages of gene editing patent data

declined, and the number of patents had been at a low level. Until 2014, the number of published gene editing patents began to increase rapidly, and had maintained a strong growth trend. Combining the above analysis, we divided the gene editing technology into four stages according to the development trend, which are 1985–2000, 2001–2002, 2003–2013 and 2014–2019. As shown in Fig. 1.

Extract Topic Words

Topic words reflect the main research direction of the patent documents. Therefore, we take the title and abstract of the patent as the analysis object, and use the text mining tool ITG Insight to identify the list of possible candidate topic words. The ITG Insight tool uses linguistic rules to extract a list of possible terms. Due to the limited ability of a single word to describe technology, the tool only retains terms that contain more than two words, according to: ‘n+n’, ‘n+v’, ‘v+n’, ‘a+n’, ‘d+n’, ‘b+v+n’, ‘n+n+n’, ‘n+m+n’, ‘v+n+b+n’, ‘n+vn+u+n+vn+n’, ‘n+n+vn+c+vn+n’ and other word matching rules to extract the term list. Here ‘n’ is a noun, ‘v’ is a verb, ‘a’ is an adjective, ‘d’ is an adverb, ‘b’ is a distinguishing word, ‘m’ is a numeral, ‘u’ is an auxiliary word, ‘c’ is a conjunction, and ‘vn’ represents either a noun or a verb. After that, ITG Insight judges the previous candidate word list by the value of PC-value, and thinks that words lower than PC-value are unlikely to be technical terms. The calculation formula of PC-value is as follows:

$$PC\text{-}value(a)=\begin{cases}\log_{2^{|a|}} * f(a)+2^{|a|-2} * g(a),\ a \text{ is not nested}\\ \log_{2^{|a|}} * \left(f(a)-\frac{1}{T_a}\sum_{b\in T_a} f(b)\right)+2^{|a|-2} * g(a), \text{Other situations}\end{cases} \quad \ldots (1)$$

where, a and $|a|$ are the extracted candidate words and word length, $f(a)$ is the frequency of a in the corpus, $g(a)$ is the frequency of documents containing a, T_a is the set of nested words in a, $|T_a|$ is the number of nested words, and $f(b)$ represents the frequency of a nested word b in the corpus. After that, the topic word list is handed over to the domain expert for judgment, and the expert deletes the fake topic, and the remaining terms are the required technical terms.

When drawing the co-occurrence network graph of topic words, ITG Insight applies a complex network algorithm Fruchterman-Reingold and Vos Mapping algorithm to visualize the layout of network nodes in the space, and calculate the relationship strength through the relationship matrix of topic word co-occurrence quantity, the topic words that exceed a certain relevance threshold are connected to form a co-occurrence network.

Calculate Topic Word Weights

Delphi is a more commonly used method to determine the weight of keywords or topic words, but differences in expert knowledge and experience may affect the final result.[17] In order to reduce the influence of subjective factors when giving weight, we use the Inverse Document Frequency (IDF) method to calculate the weight of each topic word. Term frequency alone may not have the discriminating power to pick up all the relevant documents from other irrelevant documents. Therefore, an IDF factor can increase the term’s discriminating power, and has been widely used in term weighting.[18] The size of IDF is inversely proportional to the common degree of a word. In the previous step, we have selected high-frequency topic words based on word frequency, so we use the IDF method to further determine which of these high-frequency topic words are basic-vocabulary and which are special vocabularies that cause a change in research direction. The IDF value of the topic word i is calculated as follows.

$$IDF_i = \log(\frac{N}{n_i}) \qquad \ldots (2)$$

where, N is the total number of documents in the corpus, and n_i is the number of documents containing the topic word i.

Establish Weighted Adjacency Matrix

A general adjacency matrix can represent a simple geometric relationship between connected and disconnected nodes. After adding the weight between the connections, the structure of the network itself can be used to reflect the tightness between the connections. In step two, we have obtained the co-occurrence network diagram of the topic words in each stage, so we establish the adjacency matrix according to the connection relationship between the topic words in the network. If there is a connection between node i and node j, then $a_{ij} = 1$, otherwise 0, obeying the above rules to generate matrix A.

$$A = \begin{pmatrix} 0 & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & 0 \end{pmatrix} \qquad \ldots (3)$$

After obtaining the matrix A, the social network analysis software UCINET is adopted to calculate the density of the matrix. Network density reflects the closeness between nodes in the network. When the network density of topic words is higher, the connection between topic words becomes closer, which is more conducive to the spread of topics. The density calculation formula is:

$$Density = \frac{2L}{n(n-1)} \qquad \ldots (4)$$

where, L is the number of links actually owned in the network, and n is the number of topic words. When calculating the importance of a node, we first adopt the gravity model to calculate the weight between the connections. The specific formula is as follows.

$$\omega_{ij} = \frac{\omega_i \omega_j}{d_{ij}^2} \qquad \ldots (5)$$

where, ω_{ij} denotes the weight of the connection between node i and node j, ω_i is the weight of node i, ω_j is the weight of node j, and d_{ij} is the distance between node i and node j.

Then multiplying the value of ω_{ij} by the binary value of the intersection of the i-th row and the j-th column in the matrix A to obtain the final weighted adjacency matrix M. On this basis, the mainstream technologies of each stage can be determined according to the degree centrality C_{RD} of the nodes. The form of the matrix M and the calculation formula of C_{RD} are as follows:

$$M = \begin{pmatrix} 0 & a_{12} * \omega_{12} & \cdots & a_{1n} * \omega_{1n} \\ a_{21} * \omega_{21} & 0 & \cdots & \vdots \\ \vdots & \ldots & \ddots & a_{(n-1)n} * \omega_{(n-1)n} \\ a_{n1} * \omega_{n1} & \cdots & a_{n(n-1)} * \omega_{n(n-1)} & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & m_{12} & \cdots & m_{1n} \\ m_{21} & 0 & \cdots & \vdots \\ \vdots & \cdots & \ddots & m_{(n-1)n} \\ m_{n1} & \cdots & m_{n(n-1)} & 0 \end{pmatrix} \qquad \ldots (6)$$

$$C_{RD} = \sum_{j=1}^{n} \frac{m_{ij}}{n-1} \qquad \ldots (7)$$

Results and Discussion

Evolution of Gene Editing Technology from 1985 to 2000

Since the frequency of the topic words or keywords can directly reflect the research focus,[19] we used ITG Insight to analyze the patent titles and abstracts from 1985 to 2000, and determined 40 candidate topic words based on the number of co-occurrences. After selection by experts, 20 technical subject headings were retained, and plotted the network diagram shown in Fig. 2. The node text in the Fig. 2 indicates the name of the topic. The number in parentheses indicates the number of occurrences of the topic. The size of the node is proportional to the number of occurrences of the topic. The line between the nodes indicates that there is a co-occurrence relationship between the two. On the right side of the Fig. 2 is the topic word number.

Through co-occurrence analysis and topic weights calculation, both the adjacency matrix and the weighted adjacency matrix were constructed, and then the network density of the topic words was obtained by the adjacency matrix, and the degree centrality of the nodes was obtained by the weighted adjacency matrix. The specific results are shown in Table 1.

It can be seen from the results that in the budding stage, the network density index is only 0.1263. There are still several isolated nodes in the network, the network density of the topic words is low, and there is no excellent connection between the topics. The topic

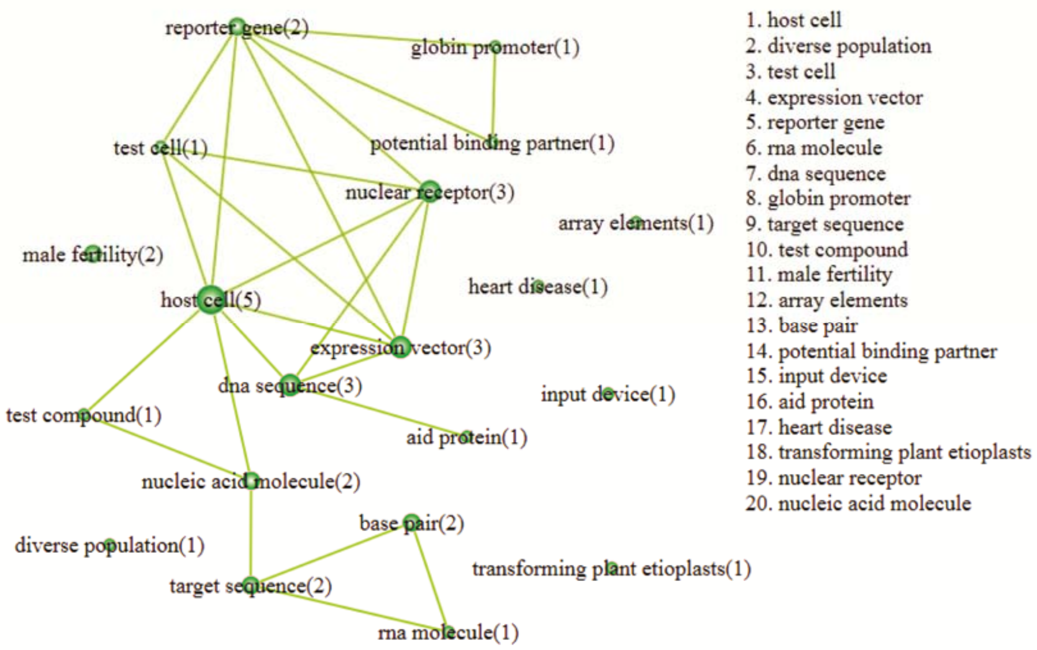

Fig. 2 — Technical topic word network diagram (1985–2000)

Table 1 — Topic weight, degree centrality and network density (1985–2000)

Topic number	Weight ω_i	Degree centrality (C_{RD})	Topic number	Weight ω_i	Degree centrality (C_{RD})
5	2.708	41.026	7	1.003	8.375
3	3.401	26.033	6	2.015	7.670
19	2.303	22.464	13	1.792	7.220
1	1.322	21.731	16	3.401	3.412
8	3.401	20.779	2	3.401	0.000
14	3.401	20.779	11	2.708	0.000
20	2.303	15.514	12	3.401	0.000
4	1.322	14.192	15	3.401	0.000
10	3.401	12.327	17	3.401	0.000
9	2.015	12.310	18	3.401	0.000

Network density: 0.1263

words numbered 5, 3, 19, 1, 8 are ranked relatively high in degree centrality, indicating that the research on gene editing technology in the early stages of development is still in the exploration stage, and most scholars mainly reveal the unknown function of the genome by regulating the nucleic acid and DNA sequence of host cells or test cells. In addition, in isolated nodes, ‘heart disease’ and ‘transforming plant etioplasts’ are also two important topic words, which indicates that scholars at that time had begun to pay attention to the important role that gene editing technology might play in human diseases and plant improvement, and hoped that the implementation of this technology would get the desired gene expression traits.[20–22]

Evolution of Gene Editing Technology from 2001–2002

Similarly, after analyzing the patent data of 2001–2002, the following results were obtained: Fig. 3 shows the network diagram of topic words, Table 2 shows the analysis results.

The network density of topic words in 2001–2002 had increased significantly. The network density index increased from 0.1263 to 0.5263, which indicates that gene editing technology had made remarkable achievements in the development of these two years. At this stage, the connection between the topics was closer, and scholars were more inclined to carry out research work from multiple technical topics, thereby promoting the wider application of gene editing technology. It is not difficult to see from the ranking of the degree centrality of topic words, ‘residue amino acid sequence’, ‘biological activity’ and ‘test compound’ are the three most important topic words, indicating that gene editing technology has been used to generate compounds. Meanwhile, the emergence of ‘immune response’, ‘biological sample’, ‘pathological state’, ‘isolated polypeptide’ also provides the possibility for the application of gene editing technology in drug development, clinical medicine, and mammalian cell modification.[23,24]

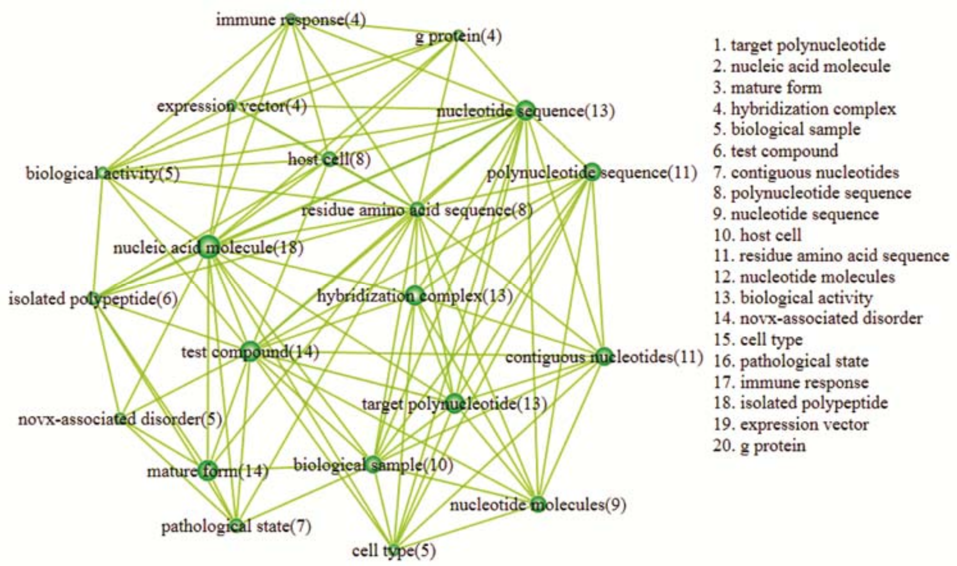

Fig. 3 — Technical topic word network diagram (2001–2002)

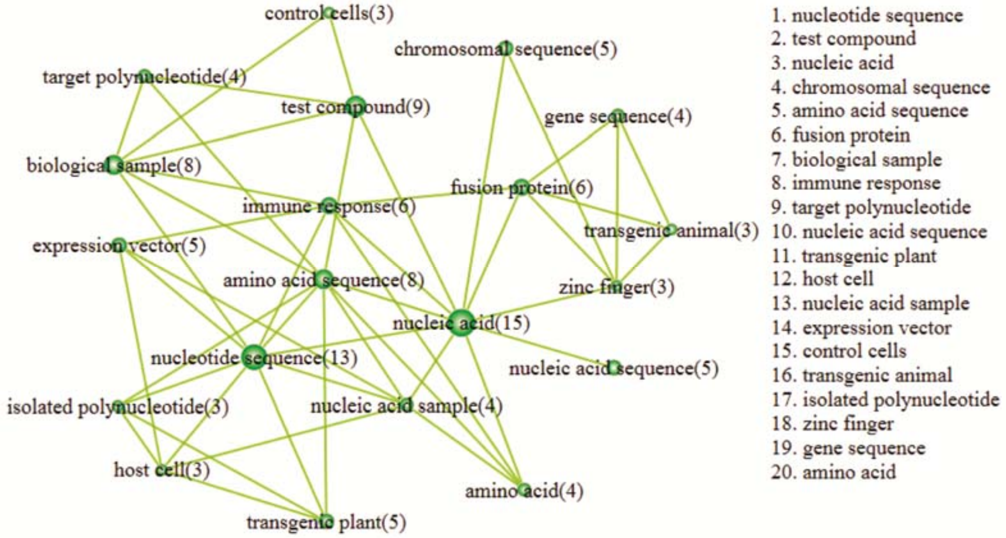

Fig. 4 — Technical topic word network diagram (2003 – 2013)

Table 2 — Topic weight, degree centrality and network density (2001–2002)

Topic number	Weight ω_i	Degree centrality (C_{RD})	Topic number	Weight ω_i	Degree centrality (C_{RD})
11	1.946	46.420	5	1.014	17.011
13	2.208	26.460	10	1.179	15.336
6	1.077	24.579	8	1.215	15.321
12	1.872	22.373	20	1.420	15.027
2	0.927	22.349	7	1.179	14.905
1	1.678	21.934	16	1.566	14.350
4	1.620	21.279	18	1.110	13.696
15	1.872	20.467	3	1.466	13.580
19	2.026	20.215	14	1.620	9.961
17	1.946	19.572	9	0.352	8.549

Network density: 0.5263

Evolution of Gene Editing Technology from 2003 to 2013

Since 2003, the number of patents for gene editing had grown slowly. After extracting the patented topic headings from 2003 to 2013, we obtained the network diagram as shown in Fig. 4. It can be seen from the Fig. 4 that the network density of topic words in this period had decreased significantly, and the co-occurrence frequency of high-frequency topic words was also less than in the previous period. To further analyze the trend of change in this period, we calculated the weight of the topic words, and completed the establishment of the weighted adjacency matrix on this basis, and finally obtained the degree centrality and network density of the topic words. The results are shown in Table 3.

The analysis results show that the research on gene editing technology is more dispersed, and the connection degree of topic words is generally not high. In Table 3, ‘nucleic acid sample’, ‘host cell’, ‘amino acid sequence’ are the top three topic words in importance. In addition to some common topic words, some new topic words have appeared, such as ‘transgenic plant’, ‘zinc finger’, ‘fusion protein’, ‘transgenic animal’, ‘chromosomal sequence’, which indicates that in 2003–2013 gene editing technology had been used in the improvement of actual organisms. To further explain the phenomenon of patent reduction, we did a simple literature analysis of the articles related to gene editing in the WOS database from 2003 to 2013, and found that in addition to common words, ‘mitochondria’, ‘evolution’, and ‘alternative splicing’ are the words that appear more frequently. Afterwards, we reviewed the specific literature and found that mitochondrial replacement and mitochondrial gene editing can not only improve biological traits, but also be an effective means of drug development and treatment of genetic diseases.[25,26] ZFN and TALEN, as the main technologies of gene editing at that time, had been tried in this practice, but due to the lack of technology and the higher requirements for safety, the development of gene editing in this research had stalled, which also explained why the number of patents declines during this period.[27,28]

Table 3 — Topic weight, degree centrality and network density (2003–2013)

Topic number	Weight ω_i	Degree Centrality (C_{RD})	Topic number	Weight ω_i	Degree Centrality (C_{RD})
13	3.401	30.849	17	2.390	20.003
12	2.485	30.738	6	1.954	19.552
5	1.609	30.502	9	3.401	18.768
7	1.954	27.824	16	2.708	16.901
11	3.178	24.098	19	2.390	15.674
1	1.099	22.335	15	3.689	14.418
8	2.485	22.077	3	0.525	10.738
2	1.954	21.846	20	1.050	8.420
14	2.303	21.805	4	2.996	7.255
18	1.897	20.057	10	2.015	1.057

Network density: 0.2526

Evolution of Gene Editing Technology from 2014 to 2019

After analyzing the patent data available in 2014–2019 with ITG Insight, we obtained the topic network diagram shown in Fig. 5 and the analysis results shown in Table 4.

We find that during the year of 2014–2019, scholars paid more attention to the development of gene editing systems, and since the emergence of CRISPR-Cas9 system in 2012, gene editing technology had entered a new era, and gene regulation by RNA guided fusion protein provided a new powerful tool for gene editing technology, which also made gene therapy from ideal to reality.[29-31] In addition, as the research on gene editing had gradually deepened, ‘pharmaceutical composition’ and ‘plant cell’ had also become the

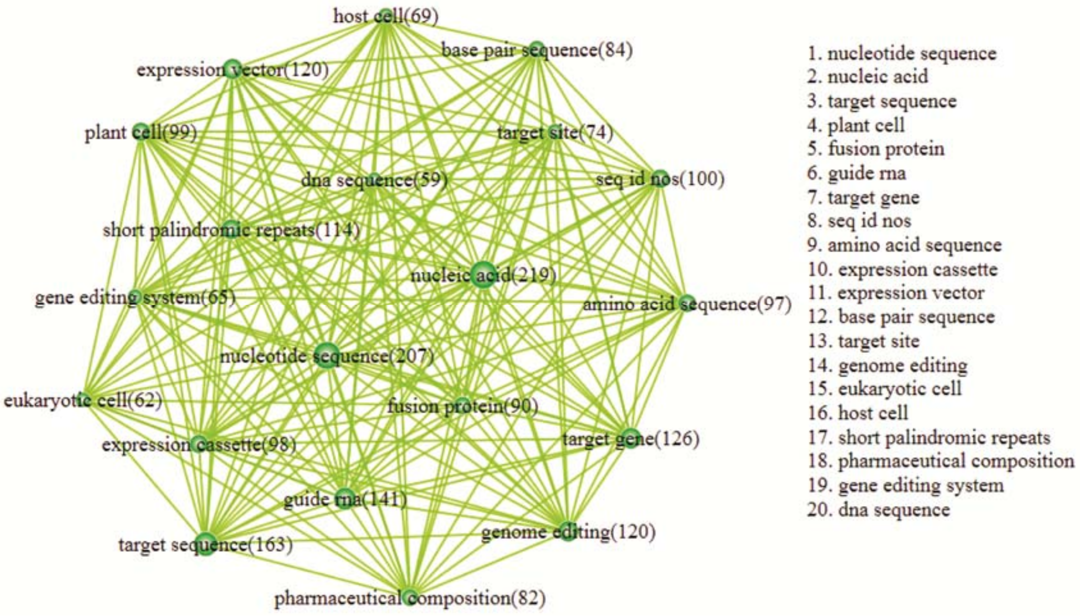

Fig. 5 — Technical topic word network diagram (2014–2019)

focus of research by scholars. Although the above two topic words provided support for the application of gene editing technology in clinical medicine, agriculture and animal husbandry, it also made cell contamination and ethical constraints a key issue that must be considered by relevant personnel.[32,33] After a simple analysis of the relevant papers in the WOS database during this period, we found that since 2013, papers about gene-edited babies and gene enhancement technology have increased year by year, and the controversy over the technology also intensified after Jian-kui HE created the first gene-edited baby in 2015.(34,35) This sounded the alarm for all researchers in the field and made them more cautious when using the technology in the future. Therefore, improving the methods and tools, optimizing the results of gene sequence modification, and getting the desired gene expression level and agronomic traits have become the main direction of gene editing technology in the new era.[36,37]

Table 4 — Topic weight, degree centrality and network density (2014–2019)

Topic number	Weight ω_i	Degree centrality (C_{RD})	Topic number	Weight ω_i	Degree centrality (C_{RD})
19	2.670	91.400	11	1.828	64.114
16	2.450	84.386	3	1.682	59.247
5	2.410	83.141	12	1.911	57.261
10	2.311	79.935	8	1.673	54.826
18	2.570	77.712	9	1.624	53.303
13	2.084	77.712	1	1.211	43.213
4	2.185	70.254	14	1.112	39.814
20	1.931	70.254	6	1.034	37.077
15	2.459	66.951	17	0.970	34.843
7	2.011	65.232	2	0.777	28.085

Network density: 0.9684

The Overall Evolution of Gene Editing Technology

Here, we analyzed the entire development process of gene editing and obtained the evolution graph shown in Fig. 6. The connection between the typical topic words in the Fig. 6 indicates that the topic evolved from the previous topic. On the whole, the evolution of the entire technology is based on the different needs of each period.

In the year of 1985 to 2000, due to the lack of understanding of gene function, all studies were in the exploration stage, in order to provide more useful information for other scholars to understand gene editing technology. So the topic words such as 'reporter gene', 'test cell' and 'nuclear receptor' were more prominent in this period.

With the deepening of research, the application range of gene editing technology was also more extensive, thus promoting its development in the production of pharmaceutical compounds. During this period, topic words such as 'biological activity', 'test compound', 'immune response', and 'isolated polypeptide' also appeared frequently in patents and became the main research direction at that time.[38]

In 2003–2013, scholars have begun to try to use gene editing technology in the research and development of transgenic animals and plants.

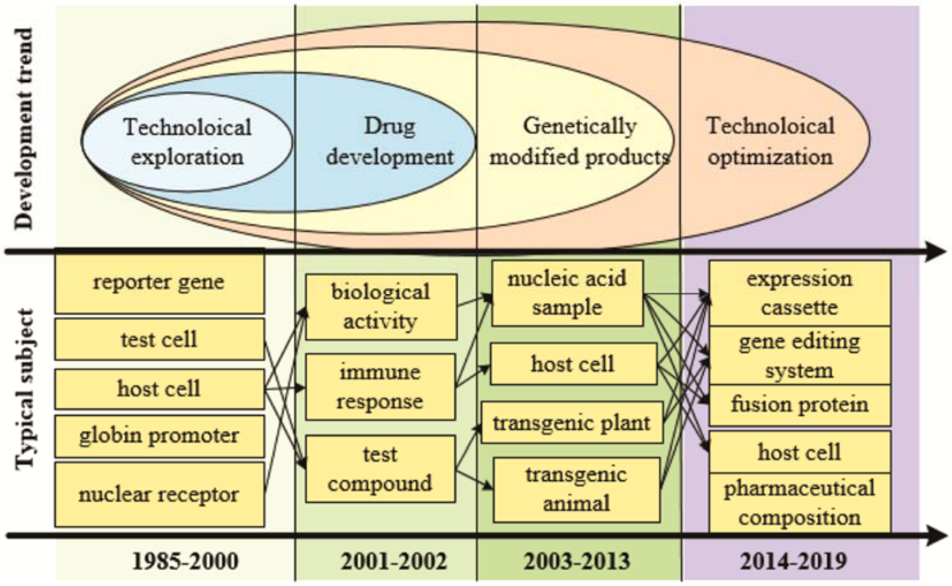

Fig. 6 — The entire process of gene editing evolution

However, when these products may indirectly affect human health, new technical problems also arise, which inhibits the development of gene editing technology, making the technology's progress in these 11 years very slow. At the same time, new technologies were urgently needed to be developed to better support the achievement of this goal.

After that, the development of the gene editing system and the advent of the CRISPR-Cas9 system broke the research bottleneck, brought the gene editing technology into a new era, and gradually shifted from laboratory applications to clinical medicine. When operating on protein-level genes, the reliability and efficiency of the expression results should be ensured, which also made ‘pharmaceutical composition’, ‘fusion protein’, and ‘expression cassette’ the main research topics at that time. Meanwhile, the research focus was more inclined to improve the technology to make up for the negative impact on cell contamination and moral constraints.

So far, we have discovered that the emergence of new technologies does affect the development direction of gene editing technology. But considering its application prospects in the medical field, the first problem to be solved is to reduce the negative effects of gene editing, improve efficiency and reliability, which is in line with the current research focus. Therefore, we can boldly predict that in the next few years, relevant research will still focus on optimizing gene editing technology to make it better benefit mankind and society.

Conclusions

Some key findings and contributions are listed as follows:

1. From the perspective of the trend of gene editing patent applications, the technology is in a rapid development stage in the past few years. In addition to continuing to expand the scope of its application, researchers should also pay attention to the factors that cause the decline in the number of patent applications. Especially in the case of imperfect technology supervision and ethics, avoiding technology abuse and biological risk is an important prerequisite to ensure the healthy development of gene editing.
2. At present, technology optimization is the research focus of gene editing, and its development direction is mainly affected by the different needs of different stages, and the emergence of new technology will accelerate the transformation of this evolutionary process. Therefore, scholars in the field should not only be alert to the emergence of new technologies, but also comply with the actual needs of society, continue to take technology optimization as one of the objectives of future research, reduce potential risks and promote the further development of gene editing technology.

This article also has its limitations. On one hand, this paper has shortcomings in fully considering the impact of policies on the development of gene editing technology, and a more comprehensive study should be conducted based on the consideration of the above factors in the future. On the other hand, the division of the research stage needs to be judged by the development trend, and the differences in personal views will also make the division result subjective, and more objective methods for stage division are required in future research.

Acknowledgement

The authors would like to acknowledge the support from “the Fundamental Research Funds for the Central Universities” (CUC210B001)

References

1 Li X, Xie Q, Jiang J, Zhou Y & Huang L, Identifying and monitoring the development trends of emerging technologies using patent analysis and Twitter data mining: The case of perovskite solar cell technology, *Technol Forecast Soc Change*, **146** (2019) 687–705.

2 Huang Y, Zhu D, Qian Y, Zhang Y, Porter AL, Liu Y & Guo Y, A hybrid method to trace technology evolution pathways: a case study of 3D printing, *Scientometrics*, **111** (2017) 185–204.

3 Durmuşoğlu A, A pre-assessment of past research on the topic of environmental-friendly electronics, *J Clean Prod*, **129** (2016) 305–314.

4 Yoon B, Lee S & Lee G, Development and application of a keyword-based knowledge map for effective R&D planning, *Scientometrics*, **85** (2010) 803–820.

5 Joung J & Kim K, Monitoring emerging technologies for technology planning using technical keyword based analysis from patent data, *Technol Forecast Soc Change*, **114** (2017) 281–292.

6 Wang B, Liu S, Ding K, Liu Z & Xu J, Identifying technological topics and institution-topic distribution probability for patent competitive intelligence analysis: a case study in LTE technology, *Scientometrics*, **101** (2014) 685–704.

7 Daiha K G, Angeli R, Oliveira S D & Almeida R V, Are Lipases Still Important Biocatalysts? A Study of Scientific Publications and Patents for Technological Forecasting, *PLoS One*, **10** (2015) e0131624.

8 Neff M W & Corley E A, 35 years and 160,000 articles: A bibliometric exploration of the evolution of ecology. *Scientometrics*, **80** (2009) 657–682.

9 Zhang Y, Zhang G, Chen H, Porter A L, Zhu D & Lu J, Topic analysis and forecasting for science, technology and innovation: Methodology with a case study focusing on big data research, *Technol Forecast Soc Change*, **105** (2016) 179–191.

10 Chen H, Zhang G, Zhu D & Lu J, Topic-based technological forecasting based on patent data: A case study of Australian patents from 2000 to 2014, *Technol Forecast Soc Change*, **119** (2017) 39–52.

11 Fan W, Liu Y, Kay L & Cheng Y, Two poles in global nano research: structure and evolution of the global nano collaborative innovation network, *Proc Int Conf Manag Eng & Technol* (IEEE, New York) 2014, 140–150.

12 Chen J K C, Pham V K, Lin F C & Chen Y R, Studying the patent of technology development on dye sensitized solar cell, *Proc Int Conf Technol Manag in the IT-Driven Services* (IEEE, New York) 2013, 985–996.

13 Park S & Jun S, Statistical Technology Analysis for Competitive Sustainability of Three Dimensional Printing. *Sustainability*, **9(7)** (2017) 1–16.

14 Norhasyima R S & Mahlia T M I, Advances in CO_2 utilization technology: A patent landscape review, *J CO_2 Util*, **26** (2018) 323–335.

15 Gao L, Porter A L, Wang J, Fang S, Zhang X, Ma T, Wang W & Huang L, Technology life cycle analysis method based on patent documents, *Technol Forecast Soc Change*, **80** (2013) 398–407.

16 Liu Y, Zhu D & Lv L, Technology maturity of product forecasting based on text mining, *Comput Integr Manuf Syst*, **14** (2008) 506–510, 542.

17 Wu C C, Constructing a weighted keyword-based patent network approach to identify technological trends and evolution in a field of green energy: a case of biofuels, *Qual Quantity*, **50** (2014) 213–235.

18 Lan M, Tan CL, Su J & Lu Y, Supervised and traditional term weighting methods for automatic text categorization, *IEEE Trans Pattern Anal Mach Intell*, **31** (2008) 721–735.

19 Liang T & Liu Y, Research landscape of business intelligence and big data analytics: A bibliometrics study, *Expert Syst Appl*, **111** (2018) 2–10.

20 Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y & Honjo T, Class Switch Recombination and Hypermutation Require Activation-Induced Cytidine Deaminase (AID), a Potential RNA Editing Enzyme, *J immunol*, **201** (2000) 2530–2540.

21 Castresana J, Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis, *Mol Biol Evol*, **17** (2000) 540–552.

22 Muramatsu M, Sankaranand V S, Anant S, Sugai M, Kinoshita K O, Davidson N & Honjo T, Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B Cell, *J Biol Chem*, **274** (1999) 18470–18746.

23 Aphasizhev R, Aphasizheva I, Nelson R E, Gao G, Simpson A M, Kang X, Falick A M, Sbicego S & Simpson L, Isolation of a U-insertion/deletion editing complex from Leishmania tarentolae mitochondria, *EMBO J*, **22** (2003) 913–924.

24 Gaddis N C, Sheehy A M, Ahmad K M, Swanson C M, Bishop K N, Beer B E, Marx P A, Gao F, Bibollet-Ruche F, Hahn B H & Malim M H, Further investigation of simian immunodeficiency virus Vif function in human cells, *J Virol*, **78** (2004) 12041–12046.

25 Gómez-Tatay L, Hernández-Andreu J M & Aznar J, Mitochondrial modification techniques and ethical issues, J Clin Med, **6** (2017) 1–16.

26 Cestari I, Kalidas S, Monnerat S, Anupama A, Phillips M A & Stuart K, A multiple aminoacyl-tRNA synthetase complex that enhances tRNA-aminoacylation in African trypanosomes, *Mol Cell Biol*, **33** (2013) 4872–4888.

27 Cornu T I, Thibodeau-Beganny S, Guhl E, Alwin S, Eichtinger M, Joung J K & Cathomen T, DNA-binding specificity is a major determinant of the activity and toxicity of zinc-finger nucleases, *Mol Ther*, **16** (2008) 352–358.

28 Mashimo T, Kaneko T, Sakuma T, Kobayashi J, Kunihiro Y, Voigt B, Yamamoto T & Serikawa T, Efficient gene targeting by TAL effector nucleases coinjected with exonucleases in zygotes, *Sci Rep*, **3** (2013) 1253.

29 Maeder M L, Linder S J, Cascio V M, Fu Y, Ho Q H & Joung J K, CRISPR RNA–guided activation of endogenous human genes, *Nat Methods*, **10** (2013) 977–979.

30 Tsai S Q, Wyvekens N, Khayter C, Foden J A, Thapar V, Reyon D, Goodwin M J, Aryee M J & Joung J K, Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing, *Nat Biotechnol*, **32** (2014) 569–576.

31 Hilton I B, D'ippolito A M, Vockley C M, Thakore P I, Crawford G E, Reddy T E & Gersbach C A, Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers, *Nat Biotechnol*, **33** (2015) 510–517.

32 Wang T, Wei J J, Sabatini D M & Lander E S, Genetic screens in human cells using the CRISPR-Cas9 system, *Science*, **343** (214) 80–84.

33 Ishino T, Hashimoto M, Amagasa M, Saito N, Dochi O, Kirisawa R & Kitamura H, Establishment of protocol for preparation of gene-edited bovine ear-derived fibroblasts for somatic cell nuclear transplantation, *Biomed Res-Tokyo*, **39** (2018) 95–104.

34 Li J, Walker S, Nie J & Zhang X, Experiments that led to the first gene-edited babies: the ethical failings and the urgent need for better governance, *J Zhejiang Univ-Sci B*, **20** (2019) 32–38.

35 Al-Balas Q A E, Dajani R & Al-Delaimy W K, The ethics of gene editing from an islamic perspective: A focus on the recent gene editing of the chinese twins, *Sci Eng Ethics*, **26** (2020) 1851–1860.

36 Ran F A, Cong L, Yan W X, Scott D A, Gootenberg J S, Kriz A J, Zetsche B, Shalem O, Wu X, Makarova K S, Koonin E V, Sharp P A & Zhang F, In vivo genome editing using *Staphylococcus aureus* Cas9, *Nature*, **520** (2015) 186–191.

37 Wang Z, Wang S, Li D, Zhang Q, Li L, Zhong C, Liu Y & Huang H, Optimized paired-sg RNA/Cas9 cloning and expression cassette triggers high-efficiency multiplex genome editing in kiwifruit, *Plant Biotechnol J*, **16** (2018) 1424–1433.

38 Perez-Pinera P, Ousterout D G& Gersbach CA, Advances in targeted genome editing,*Curr Opin Chem Biol*, **16** (2012) 268–277.