Identification and Prediction of Key Technologies in Ginsenosides Based on Genetic Knowledge Persistence Algorithm

Li Qian, Zhang Wenfeng, Yuan Hongmei*

(School of Business Administration, Shenyang Pharmaceutical University, Shenyang 110016, China)

Abstract

Objective To study the key technologies in the field of ginsenosides and to offer a guide for the future development ginsenosides through the main path identification method based on genetic knowledge persistence algorithm (GKPA). **Methods** The global ginsenoside invention authorized patents were used as the data source to construct a ginsenoside patent self-citation network, and to identify high knowledge persistent patents (HKPP) of ginsenoside technology based on the GKPA, and extract its high knowledge persistence main path (HKPMP). Finally, the genetic forward and backward path (GFBP) was used to search the nodes on the main path, and draw the genetic forward and backward main path (GFBMP) of ginsenoside technology. **Results and Conclusion** The algorithm was applied to the field of ginsenosides. The research results show the milestone patents in ginsenosides technology and the main evolution process of three key technologies, which points out the future direction for the technological development of ginsenosides. The results obtained by this algorithm are more interpretable, comprehensive and scientific.

Keywords: ginsenoside; genetic knowledge persistence algorithm (GKPA); high knowledge persistence patent (HKPP); genetic forward and backward path (GFBP); main path analysis

As a national pharmaceutical industry in China, traditional Chinese medicine has always maintained a good development trend. At present, it has become a strategic industry with strong advantages and broad prospects in national economic and social development, playing an important role in the domestic pharmaceutical market. As the main active ingredient of ginseng in traditional Chinese medicine, ginsenoside has a wide range of pharmacological effects and is used in many fields. The main path analysis method of the patent citation network can clarify the path of technological development from the huge and complicated patents, and help researchers clearly understand and grasp the technology evolution, so it is favored by many researchers and widely used.

Aiming at the deficiencies of the existing main path analysis methods, an improved main path analysis method based on the genetic knowledge persistence algorithm (GKPA) is proposed in this paper and it is applied to ginsenosides to accurately identify high-knowledge persistence patent (HKPP) and draw the high persistence main path (HKPMP). Therefore, it can obtain the genetic forward and backward main path (GFBMP), which is used to clarify the technology evolution of ginsenosides, and to guide the direction of future technology development.

* Corresponding author: Yuan Hongmei, professor. Major research area: Drug property right, etc. Tel: 13604027062, E-mail: yuanhm612@163.com.

1 Literature review

1.1 Ginsenosides

Ginseng is one of the most commonly used herbs in Asia and Western countries. Most of the biological activity of ginseng comes from its main component, ginsenoside. So far, about 40 kinds of ginsenoside components have been reported [1]. Current studies focus on using a single ginseng compound (one of the ginsenosides), rather than all ginseng compounds, to determine the mechanism of action of ginsenosides. Ginsenosides have various therapeutic effects, such as anti-cancer, immunomodulatory, anti-inflammatory, anti-allergic, anti-atherosclerotic, anti-hypertensive, and anti-diabetic effects as well as anti-stress activation and effects on the central nervous system [2]. In addition, the extraction and purification of ginsenosides, the technology of qualitative and quantitative analysis of ginsenosides, the quality control of ginseng products and plant materials, the determination of the processing effect of plant materials, and the determination of the metabolism and bioavailability of ginsenosides have all caused extensive attention of scholars [3]. However, there are few studies on ginsenosides at the technical level. Therefore, this paper uses global ginsenoside patents as a data source and associates them with the technological development process to make its technological trajectory more practical, which can help researchers interpret and judge the development trend of cutting-edge technologies.

1.2 Main path identification

To reveal the evolution process of a specific technology, many scholars use the patent citation network to track the evolution trajectory of technology. Traditional network topology analysis can be used to analyze the structure of the network, but it cannot observe the development of technology from the complicated citation network [4]. In 1989, Hummon, et al. [5] proposed the concept of main path first, and they traced the development trajectory of DNA technology through the main path of the citation network. Since then, the main path method based on the patent citation network has gradually become a research hotspot for exploring the technology trajectory, and it has been applied in many fields. To clarify the development trajectory more comprehensively and accurately in a certain field, a variety of improved algorithms or new algorithms have emerged. Based on search path link count (SPLC), search path node pair statistics (SPNP) and node pair projection count (NPPC) proposed by Hummon and Dereian in 1989. Batagelj [6] proposed the search path count (SPC) algorithm in 2003. In 2007, Vespagen [7] proposed the network evolution top path (NETP) algorithm. And in 2009, Choi and Park [8] proposed forward citation node pair (FCNP) statistics algorithm. In addition to the improvement of the link weight calculation method, Liu, et al. [9] proposed forward, backward and critical path search methods. Ma, et al. [10] used the pathfinder algorithm to delete the nodes that interfere with the main path, finding the intermediate terminal to connect the shortest path between the points of the main path, which organically bridge them to the main path. Wang, et al. [11] divided the core patent citation network, combining the main path and the derivative path to identify the technology trajectory.

Existing research has important theoretical and practical value for the identification of the main path, but there are still limitations. Although Ma Ruimin and Wang Ting considered the importance of nodes, this importance is still from the perspective of a single node that does not take the overall network into consideration. Therefore, the identified main path has problems such as a single path and lack of hub achievements to promote the development of the field, which cannot represent and present the key path of the domain evolution process. This paper proposes a genetic knowledge persistence-based main path search algorithm (GKPM) to generate multiple interconnected main paths covering HKPP, which can reduce the network complexity of the identified master paths without missing any development path in the technology field from time to time and improve the quality and efficiency of qualitative tracking of technology.

2 Research design

2.1 Research ideas

After comprehensively considering the original intention of using the main path and the theory of technological innovation, we can analyze the characteristics of the main path of technological development accurately. First of all, one of the basic purposes of using the main path analysis is to simplify the technological development. At the same time, we must include technically important patents [12], otherwise the identified main path will be misleading. In addition, one of the main driving forces of technological development is the recombination of existing knowledge [13], and a technological field usually consists of several specific sub-knowledge fields, so the main path analysis can show the compositional relationship between different knowledge flows. Therefore, a main path search algorithm based on genetic knowledge persistence is proposed, which extends the main path by GFBP. Knowledge persistence quantifies how much a patent's knowledge is retained in recent inventions based on the patent's topological position in the patent citation network using genetic algorithms. Since HKPP are milestone patents in a specific period, a patent self-citation network is constructed first. Then, the knowledge permanence of patents is calculated based on genetic algorithms to identify HKPP. Next, the SPC algorithm is used to calculate the weight of the high-persistence patent link. Finally, the forward search is performed to reduce the nodes, and the HKPMP is extracted. The reduced high knowledge persistent main path guarantees the maximum information carried, but it cannot guarantee that the analysis results of the main path cover all important paths in this field. Therefore, the GFBP search is finally performed on the nodes on the high knowledge persistent main path, screening high knowledge persistence nodes and nodes with higher genetic value to form a derivative path, repairing the main path, obtaining the GFBMP, and accurately revealing the law of technological evolution.

2.2 Research framework

Firstly, we obtain patent data in a specific field for a certain period of time from the IncoPat patent database as a data source, secondly a patent self-citation citation network is constructed. Then, we calculate the knowledge persistence of patents based on genetic algorithms to identify HKPP, decomposing the high-knowledge persistence patent self-citation network, and extracting the HKPMP. Finally, we search the nodes on the main path for GFBP to obtain the GFBMP. The specific research framework is shown in Fig. 1 .

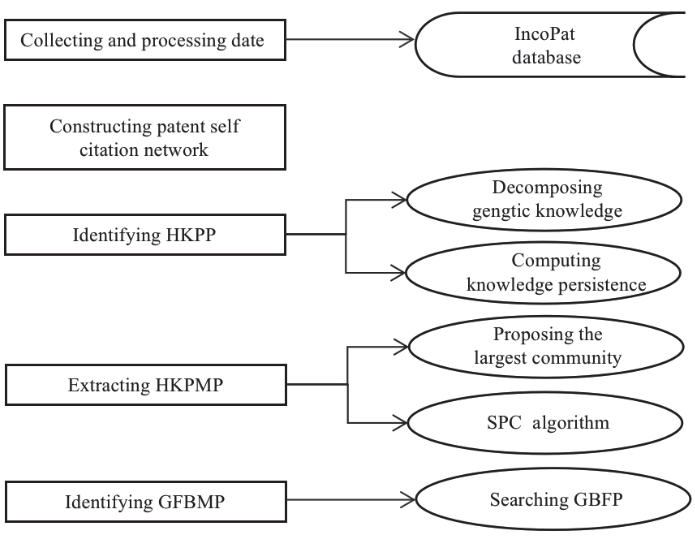
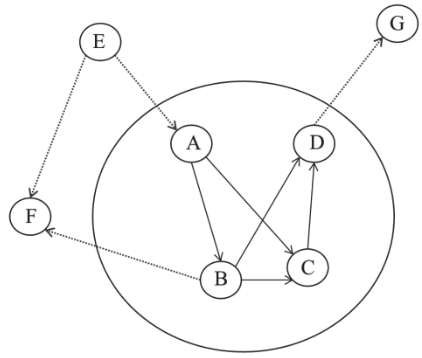


Fig. 1 Research framework

2.3 Research methods

2.3.1 Patent citation network

The patent citation network is generated based on the patent citation relationship, representing as a collection of nodes and edges, and they are shown in Fig. 2. Nodes represent patents with citation relationships, edges represent citation relationships between patents, and edges are from citing patents to cited patents, that is, a technology flows from early patents to late patents. This paper studies the knowledge flow in the technical field. So, only the patent citations in the technical field are considered, and the patent self-citation network is constructed.



Note: Each circle represents a specific patent and arrows represent knowledge flow (directed links) from cited patents to citing patents. Only the patents in the technical field (A, B, C and D) and the citation relationship (AB, AC, BC, CD) are considered, and the patents and citations outside the technical field are ignored (E, F, G, and H).

Fig. 2 Patent self-citation network

2.3.2 HKPP

Martinelli and Nomaler [14] proposed a measure of knowledge persistence inspired by Mendel's law of inheritance. They found milestone patents with high knowledge persistence in specific fields by quantifying the knowledge persistence of patents. The basic idea of this method is that new knowledge is a process of recombining old knowledge, and new inventions use a certain proportion of previous knowledge. Therefore, inherited knowledge is continuously passed on to later inventions through the same process. In fact, the contribution of each patent depends on the genealogical structure composed of all the paths generated by it. The patent genealogy structure is first identified by understanding the longest patent citation sequence in the patent citation network. Patents without back references are assigned to the first tier, patents without in-domain citation are assigned to the next tier, and patents without forward citations in the domain become the terminal (end tier), as shown in Fig. 3.

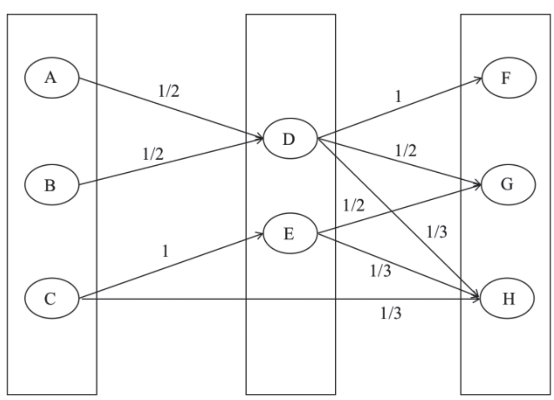


Fig. 3 Patent citation pedigree structure

When the decomposition is completed, it is necessary to quantify the degree of retention of the knowledge of the starting point in the end point, which is knowledge persistence. The knowledge inherited by new patents from old patents is measured by the reciprocal of the number of back-references of new patents. For example, patent D is cited by patents A and B, the number of back-references of A is 2, and the number of back-references of B is 4, then the knowledge inherited by A from D is 0.5, and the knowledge inherited by B from D is 0.25. Therefore, the knowledge persistence of patents in patent citation networks can be expressed as Equation (1).

$$KP_A = \sum_{i=1}^{n} \sum_{j=1}^{m_i} \prod_{k=1}^{l_j - 1} \frac{1}{BWDCit\,(Pijk)} \tag{1}$$

Where KP_A is the knowledge persistence value of patent A (P_A), n represents the number of patents directly or indirectly cited in the last layer. And mi refers to all paths from patent i (P_i) in the last layer to patent A. l_j refers to the number of patents existing in the j-th path from P_i to P_A. P_{ijk} is the k-th patent on the j-th reverse path from P_i to P_A. And $BWDCit$ (P_{ijk}) is the number of back-references of P_{ijk}, which ignores the back-references of patents between the first and t1 tiers. When P_A belongs to t-tier, the back-reference index of Cit (P_{ijk}) measures the number of recently invented patents (end tiers).

Defining HKPP requires global persistence (GP) and layer persistence (LP). GP identifies important patents in the global domain, and LP identifies important patents in each layer. LP is used to cover relatively recent important patents. GP is the knowledge persistence value of the patent divided by the maximum knowledge persistence value in the entire network. LP is the patent's knowledge persistence value divided by the largest knowledge persistence value in the same layer.

2.3.3 HKPMP

The main path first measures the importance of all links in the citation network through the concept of "traversal count", and then it connects the most important links to the main path in turn. We compare the current 6 main traversal counting search algorithms such as NPPC, SPLC, SPNP SPC, NETP and FCNP. Among them, the performance of SPC is slightly better than that of SPLC and SPNP [11]. Therefore, this paper uses SPC to calculate the link weight of high-permanence patents. The basic formula of the algorithm is as follows.

First, we should determine the traversal weight of the edge. In the reference network, N (u, v) is defined as the number of all paths from source node s to node t. To determine N (u, v), two auxiliary quantities are introduced. $N-(v)$ is the path from source node s to node v, and $N+(v)$ is the path from node v to sink t, so source point s to the path π of sink t via edge (u, v) can be expressed as:

$$\pi = \sigma \;\&\; (u, v) \;\&\; \tau.$$

Where δ represents all paths $N-(u)$ from s to u, τ represents all paths $N+(v)$ from v to t, and the weight N (u, v) of the edge (u, v) is calculated as formula 2:

$$N\,(u, v) = N-(u) * N+(v) \tag{2}$$

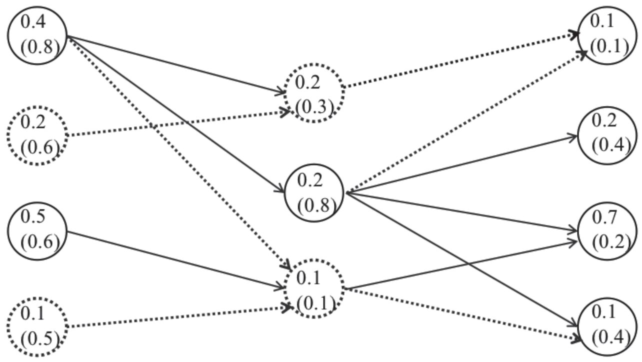
Then we determine the standard traversal weight W (u, v) of the edge. According to Kirchhoff's node law, referring to the total inflow in the network equal to N (s, t), we can obtain a natural way of normalizing the weights.

$$W\,(u,v) = \frac{N\,(u,v)}{N\,(s,t)}\,,\ 0 \le W\,(u,v) \le 1\,.$$

After obtaining the SPC weights of the HKPP, we choose the local forward search algorithm to find out the HKPMP. This local algorithm highlights the importance of a specific time point, and it can track the most important citation links in each possible split point.

2.3.4 GFBP

GFBP can search forward and backward from patents with high knowledge persistence through directly related citing/cited patents of HKPP and other valued patents. To avoid the singleness of the main path in the past, this paper conducts GFBP search on the HKPP on the main path. If there are five high-knowledge persistence patents on the main path, five backward and forward searches are performed. The backward/forward search ends at the start/end point of high knowledge persistence patents, as shown in Fig. 4 below.



Note: ◯HKPP; ◯Non-HKPP; The first row of numbers represents GP index; The second row of numbers represents LP index; ⟶ GFBP search.

Fig. 4 Genetic forward and backward path search

3 Empirical research

In this paper, ginsenosides are selected as the research object, and the patent information is obtained through the IncoPat global patent database. By reading relevant scientific literature and combining expert opinions, the search formula is determined as title abstract = (ginsenoside or triterpene saponin, ginsenoside or ginsenoside Rb1, ginsenoside Rb2 or ginsenoside Rb3, ginsenoside Rc or ginsenoside Rd, ginsenoside Rh2 or ginsenoside Rg3, ginsenoside Re or ginsenoside Rf, ginsenoside Rg1 or ginsenoside Rg2, ginsenoside Rh1 or ginsenoside Ro, ginsenoside Rh3 or ginsenoside Ri, ginsenoside F4 or ginsenoside Rg, ginsenoside Rg5 or Ginsenoside Rh). As of August 30, 2022, 5 374 patents in the field of ginsenoside technology were finally obtained after removing irrelevant patents.

3.1 Patent self-citation citation network

Based on the citation relationship between patent documents, a patent self-citation network in the field of ginsenoside technology was constructed, which had 1 676 nodes and 3 682 links. And with the help of Gephi software visualization, the patent self-citation network in the field of ginsenosides is obtained as Fig. 5.

Fig. 5 Patent self-citation network in the field of ginsenosides

3.2 HKPP

After constructing the ginsenoside patent self-citation network, we calculate the knowledge persistence index of the patent, perform global normalization and layer normalization respectively. Then, we calculate the GP index and the LP index. Usually when GP > (0.3−0.5) or LP > (0.7−0.9), it is most suitable to identify HKPP [15]. So, the cut-off value of HKPP in this paper is GP or LP patents, and a total of 55 HKPPs are identified.

By carefully reading and analyzing the specifications of the 55 HKPPs, we summarize them into the following three technical themes. The first technical topic is the extraction and purification technology of ginsenoside monomers, covering 25 patents, among which the patent US5776460A and patent KR100813997B1 that rank high in the knowledge persistence index, respectively provide two methods to increase the extraction of ginsenoside. The second technical theme is the preparation of ginsenoside derivatives. There are 15 patents under this theme, including patent KR1020030080296A, which has a saponin fraction and a ginsenoside derivative. The third technical subject is pharmacological activity and application of ginsenosides and ginsenoside derivatives, including 23 patents in total, among which the patent with the largest knowledge persistence index is KR1020160086149A. It contains ginseng extract Rb2 as the active ingredient, which has effects of anti-oxidation, inhibiting sebum secretion, and pore shrinkage.

3.3 HKPMP

After identifying 55 HKPPs, the high-knowledge-persistence citation network is decomposed to remove unnecessary interfering main-path nodes and links. Next, with the help of pajek software, the SPC algorithm is used to calculate the link weight. Besides, the forward search is used to extract the HKPMP, see Fig. 6, and the specific node information is shown in Table 1.

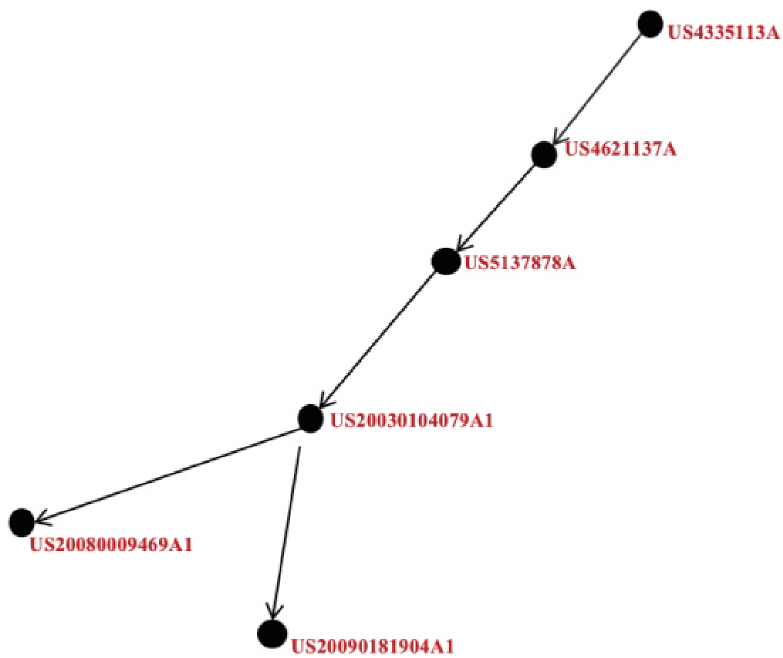


Fig. 6 HKPMP, the main path of high knowledge persistence in the field of ginsenosides

Combined with the patent specification and expert analysis, it can be seen that HKPMP mainly involves the extraction and purification technology and various pharmacological effects of ginsenosides.

Table 1 Node information of high knowledge persistence main path (HPMP) in the field of ginsenosides

Node	Public account	Title	Technology and application
1	US4335113A	Ginsenosides with pharmacological effects	HPLC
2	US4621137A	α-glycosyl ginsenoside	Synthesis of α-glycosyl ginsenosides; Production of foods containing α-glycosyl ginsenoside
3	US5137878A	Rb1 in ginsenoside mixture; Compositions and methods for treating senile dementia	Cacuum chromatography; Treat Alzheimer's type senile dementia
4	US20030104079A1	Skin tissue regeneration promoter containing ginsenoside Rb1	Skin tissue regeneration/reconstruction enhancer or wound healing enhancer
5	US20080009469A1	The drug of ginsenoside Rb1, a drug for the treatment of pain	Treat neuropathic pain
6	US20090181904A1	Methods of ginsenosides to regulate nutrient absorption	Regulates nutrient absorption

Node 1 is patent US4335113A that relates to the method of extracting ginsenosides by using pressurized liquid phase preparative chromatography technology, and the obtained saponins can be used in phlebology, rheumatology and traumatology in human and veterinary treatment. Node 2 is patent US4621137A that contains a new preparation of α-glycosyl ginsenoside, which can be obtained by passing the aqueous solution of ginsenoside and α-glucosyl sugar through α-glucosyltransferase, and it can be used in the production of foods containing α-glycosyl ginsenosides. Node 3 is patent US5137878A that describes a method for extracting pure ginsenoside Rb1 from a mixture of ginsenosides by vacuum chromatography, and further explains the mechanism of action of ginsenosides in treating senile dementia. Node 4 is patent US20030104079A1 that provides formulations containing ginsenosides and their derivatives for external or topical application to mucous membranes or cosmetics, which can be used to promote tissue regeneration in the case of cut wounds, bite wounds and skin or mucous membrane defects. Node 5 is patent US20080009469A1 that describes a drug containing ginsenoside Rb1 as an active ingredient, which can be used to treat and prevent neuropathic pain after peripheral nerve or central nerve injury. Node 6 is patent US20090181904A1 that relates to a method for regulating nutrient absorption of a subject by administering an effective amount of isolated ginsenoside compounds to the subject to regulate the transport of nutrients through the intestinal cells of the subject.

Among them, the inventions of node 1 to 3 describe how to seek the existence of ginsenosides in different fields while optimizing the extraction and purification process of ginsenosides through different approaches. Node 2 to 6 discovered new applications of ginsenosides in food, medicine, Alzheimer's disease, wound healing, pain treatment, and promotion of nutrient absorption, so as to promote the innovative progress of ginseng in drug development.

Therefore, the development of the high knowledge persistence main path (HPMP) in ginsenosides is to expand the new uses of ginsenosides by optimizing the extraction and purification technology, making technological improvements and reducing costs, , thus actively exploring new applications of ginsenosides in different fields.

3.4 GFBP

To prevent the main path analysis results from covering all the important paths, the GFBP is performed on the nodes on the HKPMP, and the nodes with a large knowledge persistence index are selected to repair the main path. Then, the GFBMP is obtained (see Fig. 7).

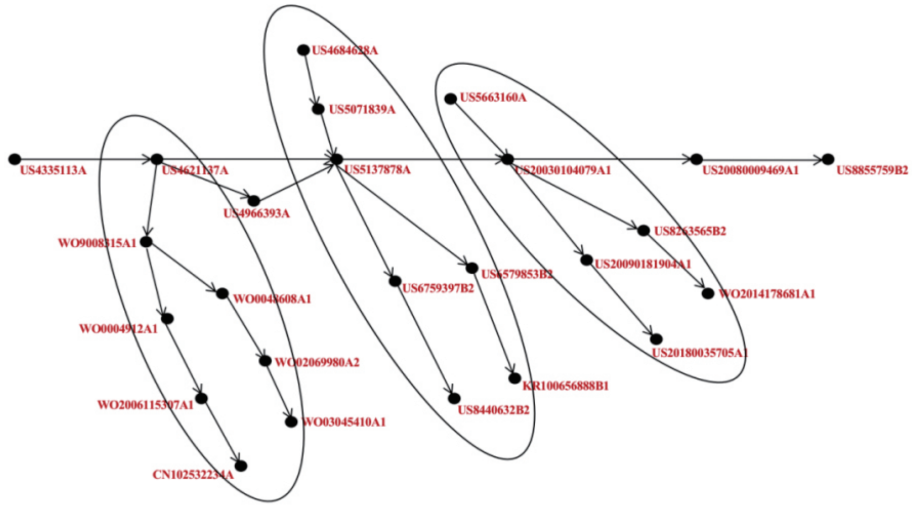


Fig. 7 Genetic forward-backward main pathway (GFBMP) in the field of ginsenosides

The derived path obtained by GFBP search on the patent US4621137A mainly describes the mechanism of ginsenosides that can alleviate Alzheimer's dementia and the evolution process of extraction and purification technology of ginsenosides (Fig. 8).

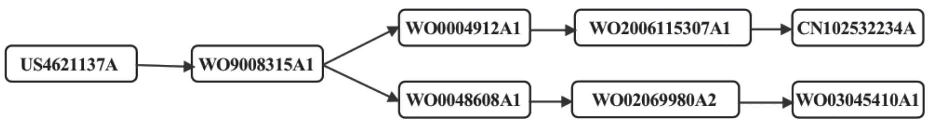


Fig. 8 Genetic forward and backward derivation path I (from left to right)

The derived path is mainly divided into two branches, the patent WO9008315A1 connected to the main path of patent US4621137A in the early stage expounds that ginsenosides Rb1 and Rg1 can alleviate Alzheimer's disease by enhancing the effectiveness of acetylcholine in the cerebral cortex and hippocampus. Patent WO0004912A1 can improve learning and memory by activating the formation of synapses and the signal transduction between synapses by further conducting in-depth research on ginsenoside Rb1 and ginsenoside Rg1. Subsequently, patent WO2006115307A1 uses gradient salting-out method instead of conventional silica gel chromatography to produce ginsenosides on a large scale. Finally, patent CN102532234A further proposes a method to prepare ginsenosides with a purity greater than or equal to 98% by using macroporous absorption resin method and column chromatography. This method can be used to realize the industrial production of ginsenoside Rg1, which has the characteristics of high product purity, high extraction rate, low cost and little environmental pollution. Patent WO0048608A1 explores the positive effect of intravenous administration of biological products containing ginsenoside Rb1 on regeneration and reconstruction of the cerebral vascular network and inhibition of secondary degeneration of nervous tissue after stroke. Patent WO02069980A2 further elaborates that ginsenoside Rb1, as a preferred panaxadiol, has the effect of stimulating angiogenesis. Patent WO03045410A1 uses a nano-delivery system for targeted administration of ginsenosides to improve the therapeutic potential of ginsenoside active substances.

In recent years, some researchers have gradually made substantial progress on the pathogenesis of Alzheimer's disease and drugs, especially the intervention of natural products [16], which further proves the reliability of the derived pathway. Many scholars have confirmed from different angles that ginsenoside Rg1 is the most active component in total ginsenosides by using animal and cell models. Domestic scholars have carried out research on the effect and mechanism of ginsenosides in the treatment of Alzheimer's disease, which aims to provide reference for the development of a new drug for the treatment of AD and other types of memory impairment.

The derived path obtained by GFBP search on patent US5137878A mainly focuses on the multi-disease pharmacological effect research and clinical application of ginsenoside composition (Fig. 9).

Fig. 9 Genetic forward and backward derivation path II

This derivation path is originally disclosed by the patent US4684628A, which is a new pharmaceutical composition for treating and preventing cardiovascular diseases and improving immune function, and it provides a method for producing these components and compositions. The composition contains some ingredients such as ginsenoside, ophiopogon saponin, sesquiaromatic hydrocarbon and jasmine. Patent US5071839A provides a safe pharmaceutical composition (SDI) II for reducing the side effects of antiviral drugs and improving immune function and its preparation method. The pharmaceutical composition is composed of seven polysaccharides and ginsenosides. The patent US5137878A, which has the highest knowledge persistence, describes the composition and method for treating senile dementia. The composition contains ginsenosides Rb1 and Rg1. Subsequent derivation path patent US6579853B2 provides a pharmaceutical composition containing ginsenoside Rb1 or it can inhibit apoptosis or apoptosis-like death, which can be used as a brain cell or nerve cell protective agent. The composition of patent KR100656888B1 containing ginsenoside has neuroprotective effect against neurotoxicity of brain striatum. Patent US6759397B2 provides the use of ginsenoside Rh2 in a method for inhibiting the proliferation of prostate cancer cells or breast cancer cells. And the patent US8440632B2 describes the application of protopanaxadiol-type ginsenoside composition for reducing the concentration of acetaldehyde in the human body. Besides, it can prevent or improve the symptoms of elevated acetaldehyde concentration, or reduce the risk of diseases or disorders caused by ethanol intake.

Therefore, ginsenosides and their derivatives and compositions have different pharmacological effects on various diseases. This point has been further confirmed in current research [17], including that ginsenosides have broad-spectrum pharmacological activities such as anti-oxidation, anti-tumor, anti-myocardial ischemia, anti-aging, and memory improvement. In addition, they also can be used to treat neurological and metabolic diseases.

The derived path obtained by GFBP on patent US5137878A mainly focuses on the use of ginsenosides in skin tissue regeneration and in promoting human nutrition absorption (Fig. 10).

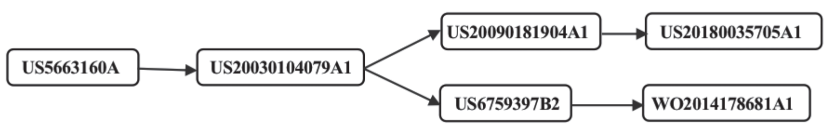


Fig. 10 Genetic forward and backward derivation path III

As an external preparation, ginsenosides in this route have a certain effect on promoting hair growth, regulating skin atopic dermatitis and sebum. Patent US5663160A is the first to propose the application of a cosmetic or dermatological composition containing ginsenosides to human skin and hair. Patent US20030104079A1, as a high-persistence node patent, further describes the promoting effect of ginsenoside Rb and its derivatives on skin tissue regeneration and healing. Patent US20090181904A1 proposes a method for regulating nutrient absorption using isolated ginsenoside compounds. Patent US20180035705A1 describes a composition that provides sustained energy for up to 8 hours with caffeine of less than 25 mg. The composition comprises at least one astragaloside compound, one ginsenoside compound, one stilbene compound and some caffeine. Patent US8263565B2 discloses a nanoemulsion containing ginsenoside metabolites as an active ingredient, a preparation method thereof, and an anti-aging skin care composition containing the nanoemulsion. Patent WO2014178681A1 relates to a composition containing ginsenosides that are excellent antioxidant, anti-inflammatory, moisturizing, improved skin acne and atopic dermatitis. Besides, they are anti-dandruff, which can promote hair growth and prevent hair graying.

In recent years, the related research on the treatment of skin wound repair by traditional Chinese medicine monomers has gradually increased. Different scholars have confirmed the antioxidant effect of human saponin Rg1, anti-skin photoaging [18], and good inhibitory effect on skin hypertrophic scars. Ginsenoside Rg3 has obvious effects on inhibiting the repair of skin wounds caused by hypertrophic scars and diabetes [19]. The construction of a targeted nano-drug delivery system increases drug delivery efficiency and therapeutic effect, which provides a theoretical basis for the subsequent development and application of related ginseng products [20].

Combining the analysis results of 3.2 and 3.3, the HKPMP of ginsenosides and the patents with high knowledge persistence reveal the macroscopic development direction of ginsenosides. In addition, the GFBP search is a supplement to the analysis results of the main path of high knowledge persistence. The description of the derived path clarifies the development of ginsenoside technology and reveals the frontier of technology development.

4 Conclusion and outlook

As a new round of technological revolution and industrial transformation is coming, it is a rare opportunity for China to catch up with this opportunity. Under the global technological innovation trend, to accurately follow the future development not only depend on more inventions, but also a global perspective. In this paper, based on the authorized invention patents in the field of ginsenosides retrieved through IncoPat before August 30, 2022, a novel, effective, comprehensive and scientific main path identification method is proposed. High knowledge in ginsenosides is identified by this method. The development context and the forefront of research represented by the HKPP, the HPMP and the GFBMP are more in line with the current hotspots of this field. It is confirmed that the proposed method has a certain robustness.

Through the analysis of high-knowledge persistence patents in ginsenosides, it can be seen that the research on ginsenosides is comprehensive and multi-angle. The research on the extraction and purification process of ginsenosides is beneficial to the separation of effective components of ginseng that provides technical support for basic scientific research. Continuous and in-depth research is carried out on the pharmacological effects of various ginsenosides, including anti-tumor, anti-oxidation, anti-fatigue, and anti-aging. Such patents have laid a scientific foundation and provided a basis for in-depth understanding of the functions of ginseng. In addition, the preparation of ginseng derivatives also plays an active role in safeguarding people’s health and promoting economic development.

Analysis of the HPMP and GFBP search of ginsenosides shows the technological development of traditional Chinese medicine industry. Besides, the basic substances of ginseng and various ginsenosides for disease prevention and treatment have received much attention. According to the main path, ginsenosides have been extensively studied from the initial extraction, separation and purification to pharmacological effects and new applications in food, medicine, and health care products. With the development of science and technology, the level of domestic innovation has been improved gradually. As the exchanges in technological field are gradually deepening, the key technologies are constantly absorbed, which will promote the innovative progress of ginseng in the development of medicine and medical care.

References

[1] Ji HK, Yi YS, Kim MY, et al. Role of ginsenosides, the main active components of Panax ginseng, in inflammatory responses and diseases [J]. Journal of Ginseng Research, 2017, 41 (4): 435-443.

[2] Christensen LP. Ginsenosides: Chemistry, biosynthesis, analysis, and potential health effects [J]. Advances in Food and Nutrition Research, 2008, 55: 1-99.

[3] Kim YJ, Zhang D, Yang DC. Biosynthesis and biotechnological production of ginsenosides [J]. Biotechnology Advances, 2015, 33 (6): 717-735.

[4] Liu Xiang, Jiang Tingting, Ma Feicheng. Collective dynamics in knowledge networks: Emerging trends analysis [J]. Journal of Informometrics, 2013, 7 (2): 425-438.

[5] Hummon NP, Dereian P. Connectivity in a citation network: The development of DNA theory [J]. Social Networks, 1989, 11 (1): 39-63.

[6] Batagelj V. Efficient algorithms for citation network analysis [EB/OL]. (2021-01-15)[2023-02-01]. https://arxiv.org/abs/cs.DL/0309023.

[7] Verspagen B. Mapping technological trajectories as patent citation networks: A study on the history of fuel cell research [J]. Advances in Complex Systems, 2007, 10 (1): 93-115.

[8] Choi C, Park Y. Monitoring the organic structure of technology based on the patent development paths [J]. Technological Forecasting and Social Change, 2009, 76 (6): 754-768.

[9] Liu JS, Lu LY. An integrated approach for main path analysis: The development of the Hirsch index as an example [J]. Journal of the American Society for Information Science and Technology, 2012, 63 (3): 528-542.

[10] Ma Ruimin, Yang Yuhua. New exploration of domain main path discovery based on node importance [J]. Journal of Intelligence, 2018, 37 (3): 71-78+93.

[11] Wang Ting, Huang Yufeng, Liu Xinrui, et al. Technology trajectory identification based on main path-derivative path − taking the traditional Chinese medicine industry as an example [J]. Information Theory and Practice, 2020, 43 (8): 128-133.

[12] Fontana R, Nuvolari A, Verspagen B. Mapping technological trajectories as patent citation networks. An application to data communication standards [J]. Economics of Innovation and New Technology, 2009, 18 (4): 311-336.

[13] Youn H, Strumsky D, Bettencourt LM, et al. Invention as a combinatorial process: Evidence from US patents [J]. Journal of the Royal Society Interface, 2015, 12 (106): 20150272.

[14] Martinelli A, Nomaler Ö. Measuring knowledge persistence: A genetic approach to patent citation networks [J]. Journal of Evolutionary Economics, 2014, 24 (3): 623-652.

[15] Huang Ying, Zhu Donghua, Qian Yue, et al. A hybrid method to trace technology evolution pathways: A case study of 3D printing [J]. Scientometrics, 2017, 111 (1): 185-204.

[16] Ahmed T, Raza SH, Maryam A, et al. Ginsenoside Rb1 as a neuroprotective agent: A review [J]. Brain Research Bulletin, 2016, 125 (1): 30-43.

[17] Attele AS, Wu JA, Yuan CS. Ginseng pharmacology: Multiple constituents and multiple actions [J]. Biochemical Pharmacology, 2000, 58 (11): 1685-1693.

[18] Liu Mengna, Li Zhengyong, Shao Yalan, et al. New progress in the anti-aging effect of ginsenosides on skin [J]. Journal of Nanchang University (Medical Science), 2018, 58 (1): 91-95.

[19] Yang Min, Ding Chuanbo, Wu Yaru, et al. Research progress of ginsenosides on skin wound repair [J]. Ginseng Research, 2022, 34 (4): 37-39.

[20] Ye Chao. Preparation of diol-type ginsenoside W/O nanoemulsion and its anti-skin aging research [D]. Yanbian University, 2019.