The Journal of International Advanced Otology

J Int Adv Otol 2021; 17(6): 541–550 • DOI: 10.5152/iao.2021.21056

Review

A Bibliometric Analysis of Deafness Gene Literature From 2009 to 2018 based on Web of Science

Linhui Li[1,*], Gefei Shi[2,*], Zhigang Yang[1], Zhihao Shen[1]

[1]ENT Department, Affiliated Jiading Central Hospital of Shanghai University of Medicine & Health Sciences, Shanghai, China
[2]Shanghai Key Laboratory of Forensic Medicine, Key Laboratory of Forensic Science, Ministry of Justice, PRC, Shanghai Forensic Service Platform, Academy of Forensic Science, Shanghai, China

ORCID IDs of the authors: L.L. 0000-0002-6881-5949; G.S. 0000-0002-5809-9993; Z.Y. 0000-0002-3671-3542; Z.S. 0000-0001-9137-2923.

Cite this article as: Li L, Shi G, Yang Z, Shen Z. A bibliometric analysis of deafness gene literature from 2009 to 2018 based on web of science. *J Int Adv Otol.* 2021;17(6):541-550.

To find the trend and hot points of deafness gene research in the nearly 10 years, a bibliometric analysis of literature in the fields from Web of Science was performed. Python 3.7, Gephi 0.9.2, and SPSS 22.0 were used for cleaning, restructuring, and analysis. A total of 2828 article documents were collected, which were from 651 Science Citation Index Expanded journals, 93 countries/regions, 18 199 authors, and 3148 organizations. *PLoS One* published more papers than other journals. Among the journals that published more than 30 papers, 10 journals were included in the genetics & heredity class. The top country/region with highest number of papers was United States. However, the top author and top affiliation was Dai Pu (China) and People's Liberate Army General Hospital (China), respectively. The results of co-occurrence network, hierarchical clustering, and multidimensional scale analysis of keywords showed that 2 groups were classified. The 2 groups focused on the gene mutation of deafness and the research objects, methods, factors, and gene locus, respectively. Multidimensional scale analysis told the research character and some research details. The gene mutation and variant of family and population diseases were the hot points.

KEYWORDS: Bibliometric, clinical medicine, deafness, gene, hearing loss, otolaryngology, web of science

INTRODUCTION

According to World Health Organization's report,[1] a person is said to have hearing loss if he or she is not able to hear or has hearing thresholds of 20 dB or better in both ears, and a person is said to be deaf if he or she has a profound hearing loss, which implies very little or no hearing, who often use sign language for communication. It was reported that over 5% of the world's population—or around 466 million people—have disabling hearing loss (hearing loss greater than 40 dB in the better hearing ear in adults and a hearing loss greater than 30 dB in the better hearing ear in children), and 34 million of these are children. In children under 15 years of age, 40% of hearing loss is attributed to unpreventable causes.[1] Can we find the relative gene to protect cue hearing loss and deafness? It is the reason why the deafness gene is researched by otolaryngologists around the world. To explore the research trends and hotspots of deafness gene, we employed bibliometrics and a graph theory to study documents retrieved from the Web of Science (WoS).

Gephi (https://gephi.org), which was developed by Bastian et al.[2] is an open-source computer program that can be used to construct graph-based maps. Gephi can deal with a more than 20 000 nodes network and takes advantage of multi-core processors. It has been claimed that Gephi's ForceAtlas2 has better measured quality, performance, and speed.[3] Thus, using the above mentioned tools, we aimed to determine the trends in deafness gene research studies over the ten years.

MATERIALS AND METHODS

Data Collection

The data were collected from the WoS Core Collection,[4] including the Science Citation Index Expanded (SCIE) and the Social Sciences Citation Index, on September 19, 2019. The keywords were detected with Medical Subject Headings (MeSH) (https://www.ncbi.nlm.nih.gov/mesh). "Deafness," "acquired deafness," "extreme hearing loss," "deaf mutism," "complete hearing loss," "bilateral

*These authors have contributed equally to this work.

Corresponding authors: Zhihao Shen or Zhigang Yang, e-mail: shengzh1010@sina.com; yzgjd@163.com

Received: October 24, 2020 • **Accepted:** April 20, 2021
Available online at www.advancedotology.org

deafness," "prelingual deafness," "cistron," "cistrons," "genetic materials," "genetic material" were the terms researched. The timespan was from 2009 to 2018. The publication type of document was limited by article. The full records were exported to plain text files for analysis with Python 3.7. The pathway for selecting the papers is shown in Figure 1. The data were aggregated as secondary data without personal information. Thus, informed consent was not required.

Analysis

Data cleaning, data reorganization, the descriptive statistics and figures of publishing year, countries/regions, authors, and affiliations were done by Python 3.7. R was used for two-side statistical analysis with $\alpha = 0.05$. The difference among group were accounted by χ^2 test.

Five keywords per article were extracted from the abstract of all documents with jieba package for Python 3.7. According to Zip's second law[5] and cumulative frequency of word frequency (34%),[6-8] the dividing word frequency of high-frequency keywords is determined. KeyWords plus® (ID-keywords) and the papers' WoS categories also were used to analyze the research topic.

The class was determined by modularity[9,10] using Gephi 0.9.2. Modularity is a measure of the relative density of a netwbeork. A community (called a module or modularity class) has high density relative to other nodes in its module but low density relative to those outside.[11] ForceAtlas2 was used to draw the network.

Eigenvector centrality and betweenness centrality were addressed using Gephi 0.9.2. Eigenvector centrality,[12] also called eigencentrality, is a measure of the influence of a node in a network. It assigns relative scores to all nodes in the network based on the concept that connections to high-scoring nodes contribute more to the score of the node in question than equal connections to low-scoring nodes. Betweenness centrality[13,14] is a measure of centrality in a graph based on shortest paths. The betweenness centrality for each vertex is the number of these shortest paths that pass through the vertex. As a result, betweenness centrality of a vertex can express the degree to which nodes stand between each other.

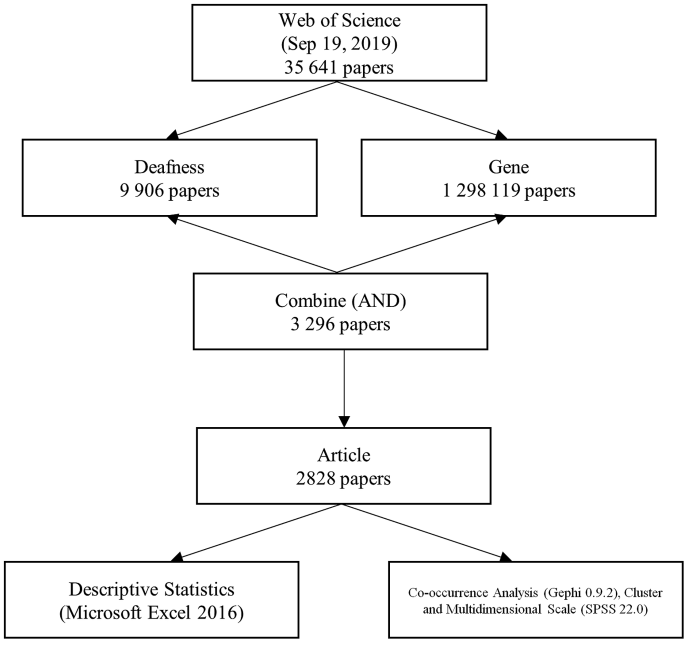

Figure 1. The progress of selecting papers.

Using SPSS 22.0 (IBM SPSS Corp.; Armonk, NY, USA) hierarchical clustering and multidimensional scale analysis were carried out on the co-occurrence matrix of high-frequency keywords, application year high-frequency keyword occurrence matrix, and subject high-frequency keyword occurrence matrix. The hierarchical clustering of high-frequency keyword co-occurrence matrix uses "Euclidean distance" to transform data and uses Ward method to cluster. For multidimensional scale analysis of high-frequency keyword co-occurrence matrix, namely alternating least-squares algorithm (ALSCAL) method is used to calculate and select "Euclidean distance" in interval measurement, select "original" in "level of measurement" in model dialog box, and check "untie tied observations." Whether the model fit is reasonable is judged by stress value <20% and namely, r-squared (RSQ) > 0.6.

RESULTS

General

There are 2828 article documents analyzed in the research. Figure 2 shows that 2015 published more articles (319 articles) than other years, while 2009 published much less than others. On average, there were 282 articles published per year. There was no significant difference among the years (χ^2=13.03, P = .16).

All documents came from 651 journals. On average, 4.34 papers were published per journal (median 1.00 papers; lower and upper quartiles were 1.00 and 3.00, respectively). The journals publishing more than 30 papers (18 journals, 2.76%, Figure 3 and Table 1) published 36.42% (1030) of all documents.

Based on the number of documents (Table 1), *PLoS ONE* published the most articles (168 papers). Among the 18 journals, there were 8 journals from the United States, followed by England (4 journals), Netherlands (2 journals), Denmark (1 journal), Norway (1 journal), Ireland (1 journal), and Japan (1 journal). All journals were included in SCIE. The journals were categorized as follows:

- Genetics & heredity (n = 10)
- Otorhinolaryngology (n = 4)
- Biochemistry & molecular biology (n = 3)

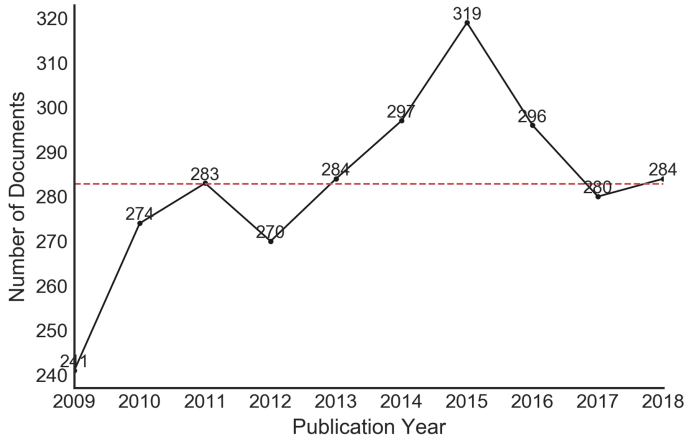

Figure 2. The trend of the article documents from 2009 to 2018.

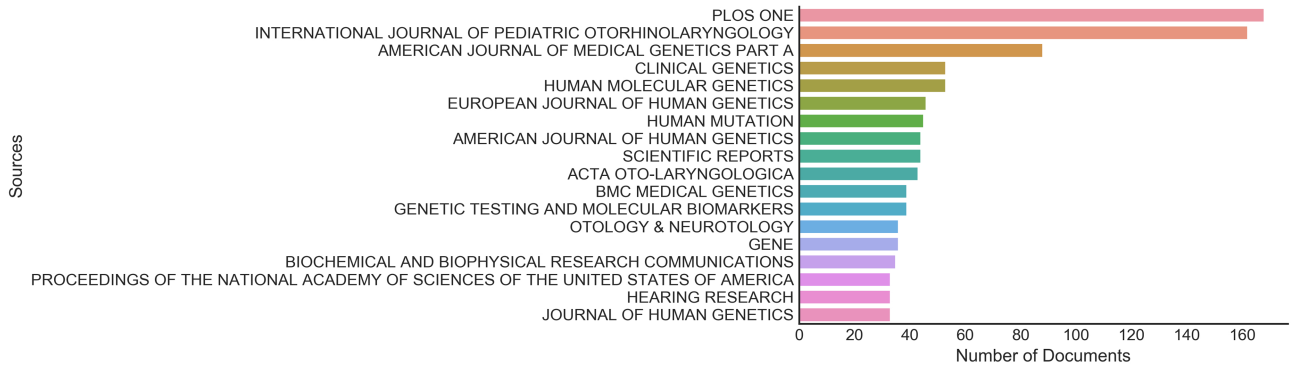

Figure 3. The journals publishing more than 30 articles about deafness gene in 2009 to 2018.

- Multidisciplinary sciences (n = 3)
- Audiology & speech-language pathology (n = 1)
- Biophysics (n = 1)
- Clinical neurology (n = 1)
- Neurosciences (n = 1)
- Pediatrics (n = 1)

However, there were 8, 7, 7, and 3 journals in Q1, Q2, Q3, and Q4, respectively. The average 2020 impact factor (IF) was 4.19 among the 18 journals (Table 1). The articles published on Q2 journals (339 articles, 32.91%) were more than other quartile zones, followed by Q3 (318 articles, 30.87%), Q1 (252 articles, 24.47%), and Q4 (121 articles, 11.75%).

Table 1. The Information of Journals Publishing More Than 30 Articles About Deafness Gene Per Journal in 2009 to 2018

Sources	Number of Documents	Country	2020 Impact Factor	2020 Quartile	Categories
PLoS One	168	USA	3.240	Q2	Multidisciplinary sciences—SCIE
International Journal of Pediatric Otorhinolaryngology	162	Ireland	1.675	Q3	Otorhinolaryngology—SCIE
				Q3	Pediatrics—SCIE
American Journal of Medical Genetics Part A	88	USA	2.802	Q3	Genetics & heredity—SCIE
Human Molecular Genetics	53	England	6.150	Q1	Biochemistry & molecular biology—SCIE
				Q1	Genetics & heredity—SCIE
Clinical Genetics	53	Denmark	4.438	Q2	Genetics & heredity—SCIE
European Journal of Human Genetics	46	England	4.246	Q2	Biochemistry & molecular biology—SCIE
				Q2	Genetics & heredity—SCIE
Human Mutation	45	USA	4.878	Q1	Genetics & heredity—SCIE
Scientific Reports	44	England	4.380	Q1	Multidisciplinary sciences—SCIE
American Journal of Human Genetics	44	USA	11.025	Q1	Genetics & heredity—SCIE
Acta Oto-Laryngologica	43	Norway	1.494	Q4	Otorhinolaryngology—SCIE
BMC Medical Genetics	39	England	2.103	Q4	Genetics & heredity—SCIE
Genetic Testing and Molecular Biomarkers	39	USA	1.795	Q4	Genetics & heredity—SCIE
Otology & Neurotology	36	USA	2.311	Q2	Otorhinolaryngology—SCIE
				Q3	Clinical neurology—SCIE
Gene	36	Netherlands	3.688	Q2	Genetics & heredity—SCIE
Biochemical and Biophysical Research Communications	35	USA	3.575	Q3	Biochemistry & molecular biology—SCIE
				Q2	Biophysics—SCIE
Proceedings of The National Academy of Sciences of The United States of America	33	USA	11.205	Q1	Multidisciplinary sciences— SCIE
Journal of Human Genetics	33	Japan	3.172	Q3	Genetics & heredity—SCIE
Hearing Research	33	Netherlands	3.208	Q1	Audiology & speech-language pathology—SCIE
				Q3	Neurosciences—SCIE
				Q1	Otorhinolaryngology—SCIE

SCIE, Science Citation Index Expanded.

There were 26 documents (0.92%) published in the top 10 journals based on the IF in 2020 journal citation reports (Table 2). From Table 2, it was shown that *New England Journal of Medicine* had the top IF (91.245), but only one paper was published in the journal while *Nature* published the most articles (7 papers). There were 8 journals from the United States, followed by England (3 journals). The journals (Table 2) were categorized as follows:

- Medicine, general, & internal (n =3)
- Biochemistry & molecular biology (n = 2)
- Cell biology (n = 2)
- Biotechnology & applied microbiology (n = 1)
- Clinical neurology (n = 1)
- Medicine, research, & experimental (n = 1)
- Multidisciplinary sciences (n = 1)
- Genetics & heredity (n = 1)
- Immunology (n = 1)

Countries/Regions

There were 93 countries/regions (Figure 4) focused on the study of deafness gene. On average, 1 country/region published 30.41 papers in the field. The United States listed in the top with 852 papers, followed by the People's Republic of China (509 papers), Germany (242 papers), Japan (233 papers), England (208 papers), France (208 papers), Italy (204 papers), Netherlands (134 papers), South Korea (111 papers), and Iran (94 papers).

According to the Eigenvector centrality, United States (1.00), England (0.97), Germany (0.93), France (0.92), Netherlands (0.89), Canada (0.88), Italy (0.88), Turkey (0.86), Japan (0.84), and Australia (0.83) listed in the top 10. Meanwhile, United States (609.12), Germany (319.44), England (231.46), France (220.13), Japan (179.41), Turkey (177.86), Canada (164.91), India (142.54), Iran (125.18), and Italy (117.42) were the top 10 countries based on the betweenness centrality.

Authors and Affiliations

From the articles, 18 199 authors were extracted, and 261 authors were found in an article.[15] On average, there were 8.88 authors per paper and 1 author published 0.16 papers. Dai Pu from People's Liberate Army General Hospital (China) had authored 52 papers in the subject from 2009 to 2018 and was listed in the top 1 (Table 3).

When the author's frequency was filtered by 6-52, there were 29 groups separated based on the modularity class of Gephi 0.9.2. There were 9 groups with more than 10 members (Figure 5A). The groups were named by the author's name, whose Eigenvector centrality score listed top in his group. They were Dai Pu group (55 members), Smith Richard J. H. group (43 members), Liu Xuezhong group (27 members), Kim Unkyung group (20 members), Guan Minxin group (19 members), Usami Shinichi group (16 members), Van Camp Guy group (14 members), Griffith Andrew J. group (13 member), Huygen Patrick L. M. group (12 members). Figure 5B detected the important authors through Eigenvector centrality score. The top 10 included Dai Pu (People's Liberate Army General Hospital), Han Dongyi (People's Liberate Army General Hospital), Yuan Yongyi (People's Liberate Army General Hospital), Kang Dongyang (People's Liberate Army General Hospital), Wang Guojian (People's Liberate Army General Hospital), Liu Xuezhong (University of Miami), Huang Shasha (People's Liberate Army General Hospital), Smith Richard J. H. (University of Iowa), Han Bing (People's Liberate Army General Hospital), Yu Fei (People's Liberate Army General Hospital). Figure 5C shows the active authors. Liu Xuezhong (University of Miami) got the top followed by Li Huawei (Fudan University), Dai Pu (People's Liberate Army General Hospital), Lin Xi (Emory University), Smith Richard J. H. (University of Iowa), Tekin Mustafa (University of Miami), Han Bing (People's Liberate Army General Hospital), Han Dongyi (People's Liberate Army General Hospital), King Maryclaire (Univ Washington). Each of those betweenness centrality score was more than 500. Otherwise, it was found that Dai Pu (People's Liberate Army General Hospital) cooperated with Liu Xuezhong (University of Miami), Li Huawei (Fudan University), and Lin Xi (Emory University).

Table 2. Top 10 Journals Based on Impact Factor in 2020 Journal Citation Reports (JCR)

Sources	Number of Documents	Country	JCR		
			2020 IF	2020 Quartile	Categories
New England Journal of Medicine	1	USA	91.245	Q1	Medicine, general, & internal—SCIE
Lancet	1	USA	79.321	Q1	Medicine, general, & internal—SCIE
JAMA - Journal of the American Medical Association	1	USA	56.272	Q1	Medicine, general, & internal—SCIE
Nature	7	England	49.962	Q1	Multidisciplinary sciences—SCIE
Cell	4	USA	41.582	Q1	Biochemistry & molecular biology—SCIE
				Q1	Cell biology—SCIE
Nature Biotechnology	3	USA	54.908	Q1	Biotechnology & applied microbiology—SCIE
Nature Medicine	1	USA	53.440	Q1	Biochemistry & molecular biology—SCIE
				Q1	Cell biology—SCIE
				Q1	Medicine, research, & experimental—SCIE
Lancet Neurology	1	England	44.182	Q1	Clinical neurology—SCIE
Nature Genetics	6	USA	38.33	Q1	Genetics & heredity—SCIE
Immunity	1	USA	31.745	Q1	Immunology—SCIE

SCIE, Science Citation Index Expanded.

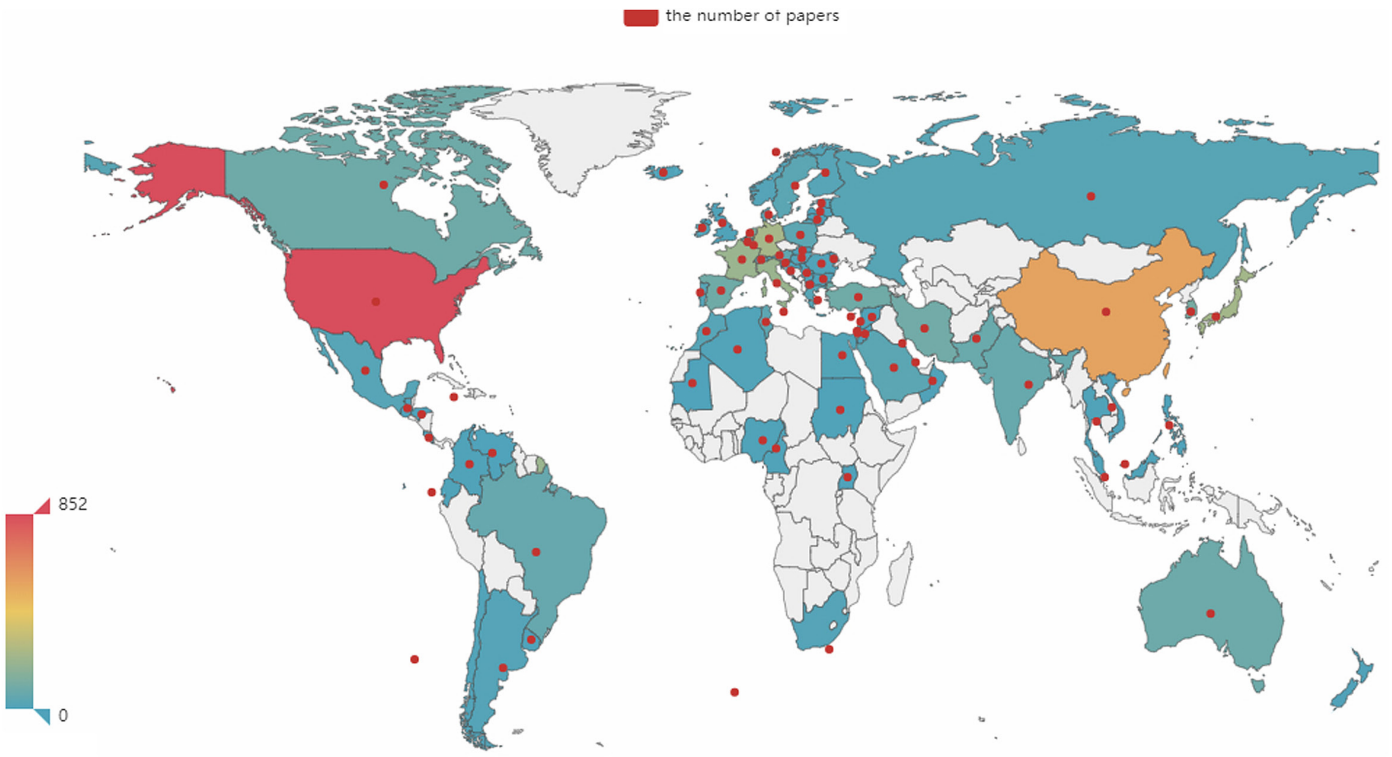

Figure 4. The heatmap of countries/regions based on the number of papers about deafness gene.

In all, 3148 organizations published articles in the deafness gene field. On average, 3.00 organizations published 1 paper, and 1 organization published 0.91 articles. One article[15] with 57 co-organizations listed in the top in the study. The top 10 organizations were People's Liberate Army General Hospital (98 papers, China), University of Miami (68 papers, USA), Radboud University Nijmegen (68 papers, Netherlands), the University of Iowa (62 papers, Italy), Harvard University (51 papers, USA), Shinshu University (46 papers, Japan), Fudan University (46 papers, China), Shanghai Jiao Tong University (44 papers, China), National Institute of Deafness and Other Communication Disorders (44 papers, USA), and Baylor College of Medicine (43, USA).

There were 160 organizations with the frequency between 10 and 98, whose cumulative frequency was 36.93%. From Figure 6A, 19 groups were found based on the modularization.[9,10] Though People's Liberate Army General Hospital (China, 0.15 and 333.9) got the highest frequency, the Eigenvector centrality score and betweenness centrality score were lower than Harvard University (USA, 1.00 and 1166.75) (Figure 6B-C). It was followed by University of Amsterdam (Netherlands, 0.72), University of Groningen (Netherlands, 0.59), Charite (Germany, 0.59), and University of British Columbia (Canada, 0.59) based on Eigenvector centrality score and followed by Radboud University Nijmegen (Netherlands, 1080.72), University of Miami (USA, 1001.92), Baylor College of Medicine (USA, 499.97), and Emory University (USA, 455.93) based on betweenness centrality score.

Keywords and Subjects

From 2828 abstracts, 2858 keywords were extracted by Python. The high-frequency keywords were shown in Table 4, and the total cumulative frequency of these keywords was more than 48%. They were divided into 4 groups based on the modularity class with modularity score 0.315 (reflecting the quality of class), and each group was named with the word who had the highest Eigenvector centrality score in its own group. As a result, the 4 groups were mutation group (including 15 words), cell group (including 6 words), hearing group (including 2 words), and syndrome group (including 2 words). These also could be viewed from the co-occurrence network (Figure 7A). The top 5 keywords based on Eigenvector centrality were mutation, gene, patient, hearing, and loss (Table 4). The top 5 keywords based on betweenness centrality were mutation, gene, family, GJB2, and patient (Table 4, Figure 7B). Mutation and gene listed the top according to all bibliometric indicators.

There were 2786 papers with the content of ID-keywords. 5595 ID-keywords were extracted and analyzed by Gephi 0.9.2. According to Table 5, the citing references of 2786 papers mainly focused on the gene research of deafness because of the inner ear of family diseases or Pendred syndrome.

Hierarchical clustering (Table 4) of 25 high-frequency keywords has shown that 2 groups were classified. Group 1 includes mutation, hearing, loss, gene, patient, family, and *GJB2*, and other keywords were collected in group 2. Group 1 indicated that the documents focused on the gene mutation of deafness. Group 2 has shown the research objects, methods, factors, and gene locus. The phenomena also were discovered by the multidimensional scale (Figure 7C).

As the stress value (6.24%) and RSQ (0.99) were both reasonable, the multidimensional scale got the high model fit. So, the result was reliable. The keywords were classified with X-axis zero scale in Figure 7D. Group 1 was on the left of X-axis zero scale, and group 2 on the other side. The keywords in the group 1 were all farther from Y-axis zero scale than that in group 2. The keywords above Y-axis zero

Table 3. Authors Published More Than 20 Articles About Deafness Gene From 2009 to 018

Author	Number of Documents	Eigenvector Centrality	Betweenness Centrality
Dai Pu (People's Liberat Army Gen Hosp)	52	1.00	1303.29
Smith Richard J. H. (Univ Iowa)	44	0.72	930.20
Usami Shinichi (Shinshu Univ)	42	0.22	139.15
Nishio Shinya (Shinshu Univ)	38	0.14	86.94
Kim Unkyung (Kyungpook Natl Univ)	33	0.13	213.36
Tekin Mustafa (Univ Miami)	32	0.36	687.22
Yuan Yongyi (People's Liberat Army Gen Hosp)	32	0.85	135.08
Han Dongyi (People's Liberat Army Gen Hosp)	31	0.95	636.89
Choi Byung Yoon (Seoul Natl Univ)	30	0.05	1.93
Liu Xuezhong (Univ Miami)	30	0.76	2092.58
Huang Shasha (People's Liberat Army Gen Hosp)	27	0.74	43.06
Kremer Hannie (Radboud Univ Nijmegen)	25	0.22	246.40
Wang Guojian (People's Liberat Army Gen Hosp)	25	0.77	93.25
Friedman Thomas B. (Natl Inst Deafness & Other Commun Disorders)	25	0.52	294.98
Avraham Karen B. (Tel Aviv Univ)	25	0.04	106.00
Kang Dongyang (People's Liberat Army Gen Hosp)	24	0.84	77.05
Van Camp Guy (Univ Antwerp)	23	0.11	35.61
Najmabadi Hossein (Univ Social Welf & Rehabil Sci)	23	0.42	90.00
Yan Denise (Univ Miami)	23	0.44	470.23
Kim Ah Reum (Seoul Natl Univ)	23	0.05	1.93
Lee Kyuyup (Kyungpook Natl Univ)	23	0.08	1.03
Shearer A. Eliot (Univ Iowa)	22	0.54	208.93
Wu Hao (Shanghai Jiao Tong Univ)	22	0.02	0.00
Wang Qiuju (People's Liberat Army Gen Hosp)	22	0.44	373.57
Yang Tao (Shanghai Jiao Tong Univ)	21	0.02	0.00
Huygen Patrick L. M. (Radboud Univ Nijmegen)	20	0.27	246.93

scale indicated the research character, namely which told that the disease's causes were the mutation or variant of relative gene. The keywords below Y-axis zero scale gave the research details, namely which told that the research focused on the patients about hearing, syndrome or family disease, hair cell, *GJB2* or (and) *SLC26A4* gene loci.

The keywords from Dai Pu and Smith Richard J. H.'s articles were analyzed, respectively. Figure 8A shows the co-occurrence network of keywords from Dai Pu's articles. The top 10 keywords based on the Eigenvector centrality score were mutation (1.0 score, 33 times), hearing (0.73 score, 23 times), patient (0.67 score, 16 times), family

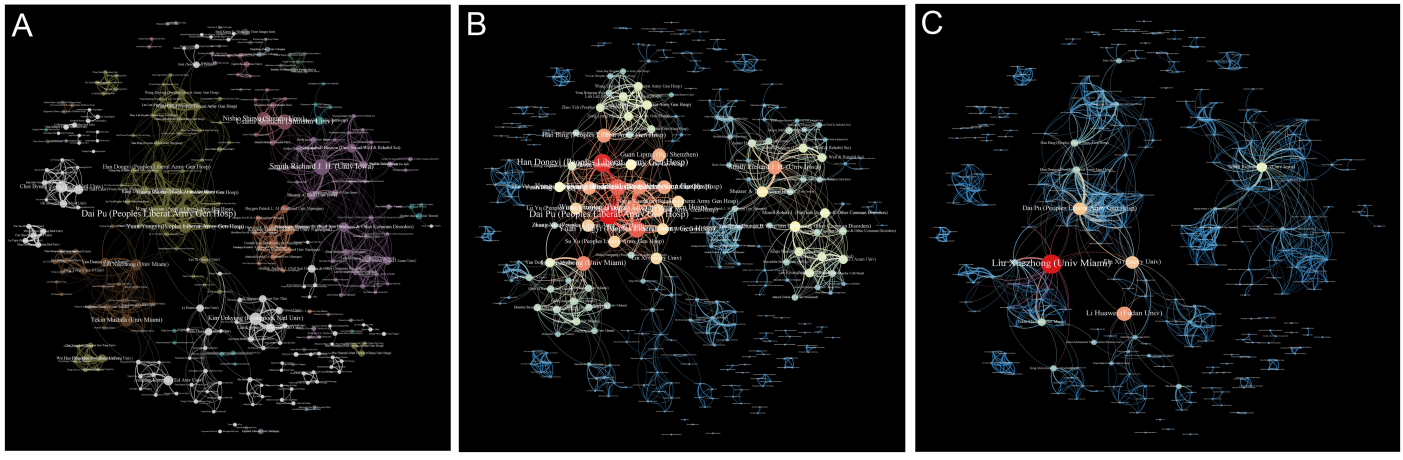

Figure 5. The co-occurrence network based on authors. (A) Nodes size and color based on the frequency and modularity class, respectively; (B) nodes size and color based on the Eigenvector centrality; (C) nodes size and color based on the betweenness centrality (the lines between the nodes represent the number of co-occurrence, the nodes were filtered by 6-52 frequency, and the edges were filtered by 2-36 frequency).

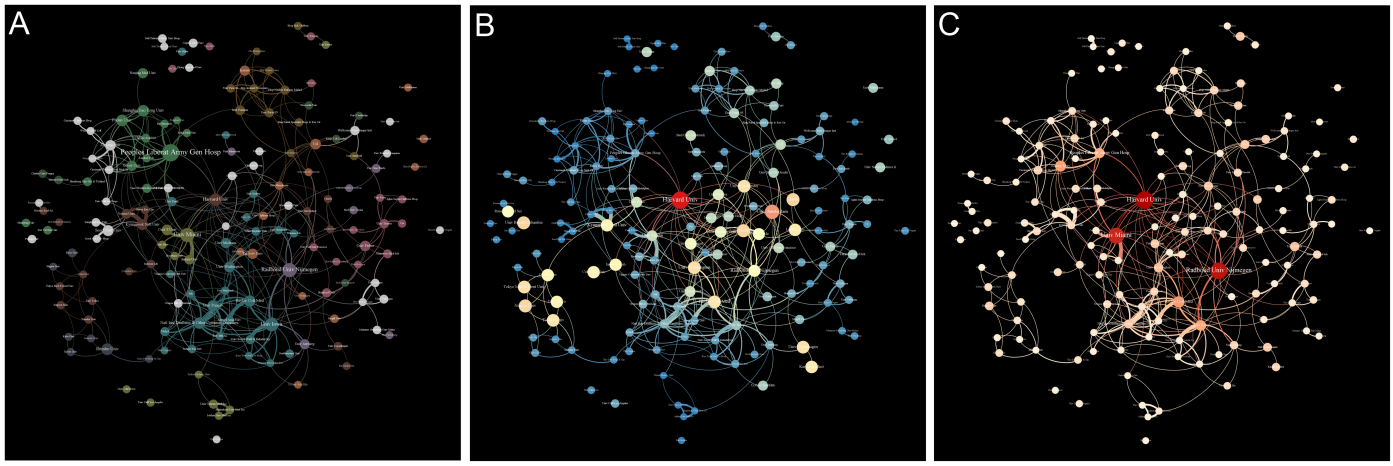

Figure 6. The co-occurrence network based on organizations. (A) Nodes size and color based on the frequency and modularity class, respectively; (B) nodes size and color based on the Eigenvector centrality; (C) nodes size and color based on the betweenness centrality (the lines between the nodes represent the number of co-occurrence, the lines were filtered by 2-20 weights, and the nodes were filtered by 10-98 frequency).

(0.61 score, 16 times), gene (0.55 score, 10 times), loss (0.50 score, 12 times), *GJB2* (0.43 score, 8 times), *SLC26A4* (0.40 score, 7 times), deafness (0.32 score, 5 times), and *TMC1* (0.28 score, 3 times). While the top 10 keywords (Figure 8B) from Smith Richard J. H.'s articles based on the Eigenvector centrality score were hearing (1.0 score, 23 times), loss (0.93 score, 20 times), gene (0.72 score, 11 times), mutation (0.68 score, 14 times), genetic (0.60 score, 10 times), variant (0.56 score, 8 times), family (0.53 score, 8 times), deafness (0.46 score, 4 times), population (0.46 score, 4 times), and cell (0.42 score, 7 times).

Table 4. The High-Frequency Keywords According to the Cumulative Frequency

Rank	Keywords	Number	Frequency (%)	Cumulative Frequency (%)	Eigenvector Centrality	Betweenness Centrality	Modularity Class
1	mutation	1070	7.60	7.60	1.00	1014484.37	1
2	gene	573	4.07	11.67	0.79	509593.45	1
3	hearing	539	3.83	15.50	0.69	319243.50	8
4	patient	430	3.06	18.56	0.70	430559.01	1
5	loss	394	2.80	21.36	0.58	196437.89	8
6	family	285	2.03	23.39	0.52	152337.04	1
7	cell	251	1.78	25.17	0.49	402848.22	10
8	syndrome	226	1.61	26.78	0.45	231216.90	2
9	GJB2	212	1.51	28.28	0.36	44176.32	1
10	deafness	189	1.34	29.63	0.48	131672.42	1
11	mouse	170	1.21	30.83	0.40	271344.38	10
12	variant	157	1.12	31.95	0.43	86525.98	1
13	mitochondrial	118	0.84	32.79	0.32	76501.96	1
14	genetic	106	0.75	33.54	0.33	55188.02	1
15	protein	104	0.74	34.28	0.32	144753.02	10
16	disease	93	1.66	35.94	0.33	83559.87	1
17	hair	93	1.66	37.61	0.25	63865.95	10
18	sequence	84	1.50	39.11	0.29	36715.19	1
19	study	84	1.50	40.61	0.31	60210.17	1
20	cause	74	1.32	41.93	0.32	46276.60	1
21	expression	72	1.29	43.22	0.26	85807.97	10
22	slc26a4	70	1.25	44.47	0.23	10089.12	1
23	identify	67	1.20	45.67	0.27	33023.91	1
24	case	67	1.20	46.87	0.28	42042.18	2
25	ear	64	1.14	48.01	0.23	62691.10	10

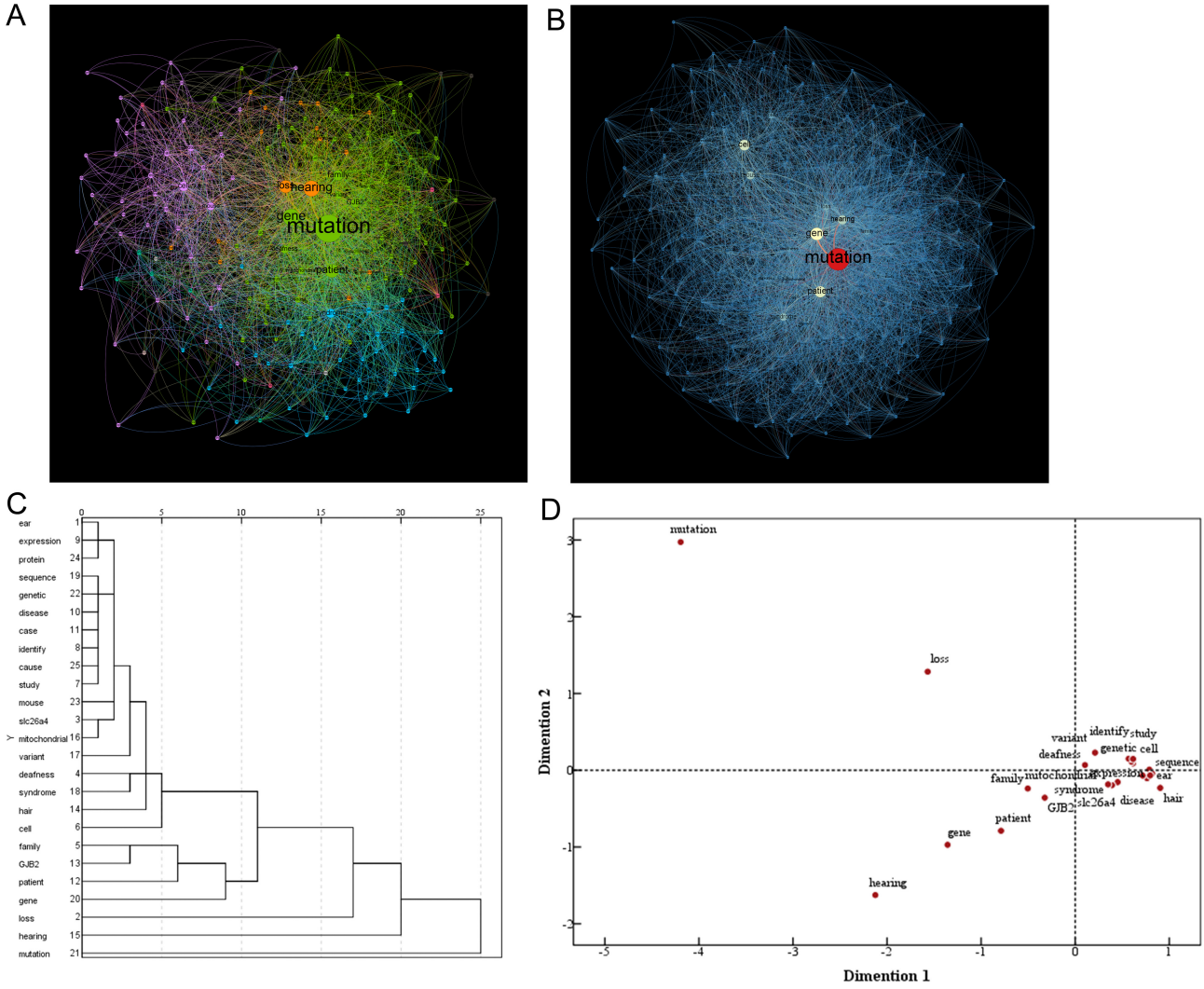

Figure 7. The keywords analysis. (A) Nodes size and color based on the frequency and modularity class, respectively; (B) nodes size and color based on the betweenness centrality; (C) hierarchical clustering; (D) multidimensional scale ((A) and (B), the lines between the nodes represent the number of co-occurrence; the nodes were filtered by 10-1070 frequency; (C) and (D) analysis of 25 high-frequency keywords)).

A total of 2828 papers were from 90 WoS categories (Table 6), in which genetics & heredity, otorhinolaryngology, and biochemistry & molecular biology were listed in the top 3. The top 3 categories occupied more than 40%.

DISCUSSION

The results indicated that the research of deafness gene was stable in the nearly 10 years though the article number of every year was up and down. The deafness gene research articles were published on 651 journals. *PLoS ONE* published the most articles (168 articles) among the journals, followed by *International Journal of Pediatric Otorhinolaryngology* (162 articles). It meant that the publication of research results was relatively scattered. Among the 18 journals publishing more than 30 articles, 8 journals were in Q1 zone and 591 articles were published on Q1 and Q2 journals. These indicated that the articles had high academic level. By the way, there was 1 article published in *New England Journal of Medicine,*[16] *Lancet,*[17] *and JAMA – Journal of the American Medicine Association,*[18] *respectively.* However, it should be noted that *PLoS ONE* was not in otorhinolaryngology class, *International Journal of Pediatric Otorhinolaryngology* was located in Q3, and 10 journals were from genetics & heredity class. So, the higher-quality professional journal in otorhinolaryngology should be promoted in the future. The results of papers' WoS categories study were common with these. But it was indicated that otorhinolaryngology and biochemistry & molecular biology also were concerned by researchers.

Eight of 18 journals were from United States. This was in common with the results of where authors came from. The co-occurrence network also showed that United States was more notable and active than other countries/regions. The number of articles from People's Republic of China was located in the second in the list; however, the Eigenvector Centrality score and Betweenness Centrality score were not both in the top 10. It indicated that there is less cooperation between People's Republic of China and other countries/regions in the field of deafness gene research.

Dai Pu from People's Liberate Army General Hospital (China) published the most articles among the authors. According to modularity class operated by Gephi 0.9.2, Dai Pu group was listed in the first with

Table 5. Top 30 ID-Keywords on Frequency

ID-keywords	Frequency	ID-keywords	Frequency
deafness	876	mice	129
gene	597	children	124
mutations	513	families	120
hearing-loss	330	sensorineural hearing-loss	114
impairment	325	family	112
expression	276	variants	111
inner-ear	258	Pendred-syndrome	107
identification	225	hair-cells	101
sensorineural deafness	216	recessive deafness	98
protein	212	phenotype	94
prevalence	193	DNA	87
spectrum	166	mouse	86
mutation	159	enlarged vestibular aqueduct	86
disease	146	connexin 26 gene	82
population	145	frequency	81

55 members. They cooperated with Liu Xuezhong group and Kim Unkyung group. Smith Richard J. H. group, Huygen Patrick L. M. group, and Usami Shinichi group had cooperation relationship in the study. But Dai Pu group and its partners did not take any touch with Smith Richard J. H. group and its partners. According to the results, authors in the Dai Pu group got more important location in the study field about deafness gene. Liu Xuezhong and Smith Richard J. H. were followed. Meanwhile, Liu Xuezhong was more active than other authors. Except Dai Pu and Smith Richard J. H., Li Huawei (Fudan University) and Lin Xi (Emory University) also played an important role among the cooperation relationships.

Like the authors' analysis, People's Liberate Army General Hospital (98 papers, China) published more articles than other organizations. But its position and activation were not notable according to the study. The reason would be that cooperating organization with People's Liberate Army General Hospital was less. Harvard (United States) though published only 51 papers cooperated with more organizations than others. It is suggested to get more and more opportunities to operate with other organizations at home and abroad for People's Liberate Army General Hospital (China). And in my opinion, it is important to build an international research system or group.

The results of keywords and ID-keywords analysis with modularity class in Gephi 0.9.2 have shown that the score was very low. It meant that the keywords were not classed. It may be related to the overconcentration of research topics. Of course, from the other side, it indicated that the articles studied in the paper focused on our research topic and that the keywords extracting function of jieba package for Python 3.7 were reliable. However, the presentation of hierarchical clustering and multidimensional scale could divide the 25 high-frequency keywords into 2 groups. Especially, the results of multidimensional scale further discovered the research character and research details. And the results also were proved by the analysis of keywords from the top 2 authors' articles. Though Dai Pu and Smith Richard J. H. all focused on the gene mutation in deafness diseases, there were differences between their research points. Dai Pu mainly studied the mutation of gene *GJB2*, *SLC26A4*, and *TMC1* in some family hearing diseases, while Smith Richard J. H. studied the mutation and variant of gene and the changes of cell in some family and population diseases.

CONCLUSION

The results of deafness gene research were smoothly published in a wide range of journals in the nearly 10 years. It was a pity that otorhinolaryngology field lacked the higher-quality professional journal in otorhinolaryngology. United States published more articles than other countries/regions, while Dai Pu and his organization—People's Liberate Army General Hospital (China)—published more documents than other authors and organizations. However, People's Liberate Army General Hospital (China) and People's Republic of China both need more cooperation with other countries and institutions in the world. Harvard University plays an important and active role in the field. Hearing loss and deafness research mainly focused on the gene

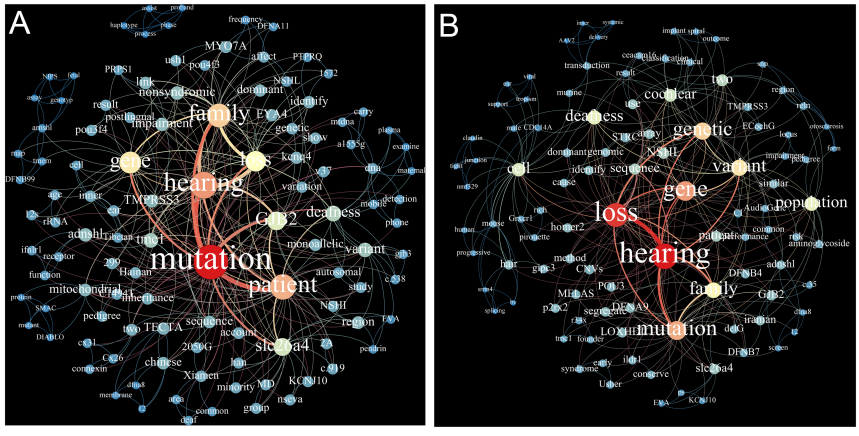

Figure 8. The co-occurrence network based on keywords of Dai Pu and Smith Richard J. H. (A) The keywords from Dai Pu's articles; (B) the keywords from Smith Richard J. H. articles (nodes size and color based on the Eigenvector centrality, the lines between the nodes represent the number of co-occurrence; the nodes were filtered by 10-1070 frequency).

Table 6. WoS Categories

WoS Categories	Frequency	Rate (%)	Cumulative Frequency	Cumulative Rate (%)
Genetics & heredity	862	21.02	862	21.02
Otorhinolaryngology	496	12.10	1358	33.12
Biochemistry & molecular biology	337	8.22	1695	41.34
Pediatrics	283	6.90	1978	48.24
Neurosciences	270	6.59	2248	54.83
Multidisciplinary sciences	265	6.46	2513	61.29
Medicine, research, & experimental	199	4.85	2712	66.15
Cell biology	145	3.54	2857	69.68
Clinical neurology	126	3.07	2983	72.76
Endocrinology & metabolism	126	3.07	3109	75.83
Other	991	24.17	4100	100.00

mutation and variant of family and population diseases. But there were subtle differences between different research groups.

Peer Review: Externally peer-reviewed.

Author Contributions: Design and Review – L.L., G.S.; Analysis and/or Interpretation – L.L., G.S.; Literature Research – L.L., Z.Y.; Supervision – Z.S., Z.Y.

Conflict of Interest: The authors have no conflict of interest to declare.

Financial Disclosure: The study was funded by the project of Science and Technology Commission Shanghai Municipality (19441900500); Funding projects of basic research business fees of central scientific research institutes (GY2021G-12); Jiading District Health Committee Fund(2020-KY-07, 2020-KY-03); Fund (Natural General Items) of Shanghai University of Medicine & Health Sciences (No. SSF-21-16-002); the Shanghai Key Laboratory of Forensic Medicine (21DZ2270800); the Shanghai Forensic Service Platform (19DZ2292700).

REFERENCES

1. World Health Organization (WHO). Deafness and hearing loss. 2019. Available at: https://www.who.int/en/news-room/fact-sheets/detail/deafness-and-hearing-loss.
2. Bastian M, Heymann S, Jacomy M, Jacomy M. Gephi: an open source software for exploring and manipulating networks. *Proceedings of the International AAAI Conference on Weblogs and Social Media*. 2009;3(1):361-362.
3. Jacomy M, Venturini T, Heymann S, Bastian M. ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PLoS One*. 2014;9(6):e98679. [CrossRef]
4. Clarivate Analytics. *Web of Science*. Available at: http://apps.webofknowledge.com/UA_GeneralSearch_input.do?product=UA&search_mode=GeneralSearch&SID=4CGQGJrHdaoHKMuhRsh&preferencesSaved=.
5. Donohue JC. *Understanding Scientific Literature: A Bibliographic Approach*. Cambridge: The MIT Press; 1973.
6. Ding H, Cao P. Research hotspots and fields of foreign media economics in 2013–based on bibliometrics. *J Comm*. 2015;22(4):61-82.
7. Zhang H, Cui L. Study of bioinformatics through co-word analysis. *J China Soc Sci Tech Info*. 2003;22(5):613-617.
8. An XR. The research on the threshold of high-frequency words based on the normal distribution in word frequency analysis. *J Intell*. 2014;10:129-136.
9. Lambiotte R, Barahona M, Delenne JC. Dynamics and multiscale modular structure in networks. Available at: http://www.monmeetings.org/meeting8/lambiotte.pdf.
10. Blondel VD, Guillaume J, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. *J Stat Mech Theor Exp*. 2008;2008(10):10008. [CrossRef]
11. Ladd J, Otis J, Warren CN, Weingart S. Exploring and analyzing network data with python. Available at: https://programminghistorian.org/en/lessons/exploring-and-analyzing-network-data-with-python.
12. Wikipedia. Eigenvector centrality. Available at: http://en.m.wikipedia.org/wiki/Eigenvector_centrality.
13. Freeman LC. A set of measures of centrality based on betweenness. *Sociometry*. 1977;40(1):35-41. [CrossRef]
14. Wikipedia. Betweenness centrality. Available at: https://en.wikipedia.org/wiki/Betweenness_centrality#CITEREFFreeman1977.
15. FANTOM Consortium and the RIKEN PMI and CLST (DGT), Forrest AR, Kawaji H, et al. A promoter-level mammalian expression atlas. *Nature*. 2014;507(7493):462-470. [CrossRef]
16. Bockenhauer D, Feather S, Stanescu HC, et al. Epilepsy, ataxia, sensorineural deafness, tubulopathy, and KCNJ10 mutations. *N Engl J Med*. 2009;360(19):1960-1970. [CrossRef]
17. Schapira AH. Mitochondrial diseases. *Lancet*. 2012;379(9828):1825-1834. [CrossRef]
18. Taylor RW, Pyle A, Griffin H, et al. Use of whole-exome sequencing to determine the genetic basis of multiple mitochondrial respiratory chain complex deficiencies. *JAMA*. 2014;312(1):68-77. [CrossRef]