Content available at: https://www.ipinnovative.com/open-access-journals

IP Indian Journal of Library Science and Information Technology

Journal homepage: https://www.ijlsit.org/

Original Research Article

Bibliometric analysis of the structure and evolution of research on epigenetics

R. Santhakumar[1]*, S. Mahalakshmi[2], V. Kalaiappan[3]

[1]*Arulmigu Kabaaleeswarar Arts and Science College, Chennai, Tamil Nadu, India*
[2]*C. Kandaswami Naidu College for Men, Chennai, Tamil Nadu, India*
[3]*Madras University Library, Chepauk Campus, Chennai, Tamil Nadu, India*

ARTICLE INFO

Article history:
Received 12-12-2023
Accepted 23-04-2024
Available online 03-08-2024

Keywords:
Epigenetics
Annual growth rate
Collaboration index
Author productivity
Bibliometrics

ABSTRACT

Aim: Epigenetics, one of the most rapidly intensifying fields of biological research and this field has experienced remarkable research attention in recent years, yet few studies investigated a bibliometric analysis of epigenetics research. In this research, the article analysed the provides a comprehensive overview of literature on epigenetics published between 2012 and 2021.

Methods: The Web of Science Core Collection Database was used to retrieve all related publications. The search was conducted on September 20th, 2023 and select articles or reviews as the document type which resulted in 9,677 publications. The scientometric tools were applied to collect information on publications years of these papers, most prolific authors, famous journals, most productive countries and organizations, language and subjects.

Results: A total of 9,677 publications were retrieved. The publication's output showed a gradual upward trend from 2012 to 2021. The highest number of publications was 1,183 published in 2021. In our study, 117 countries contributed to this area of research in total. The top three countries were the USA, China, and Germany. Proceedings of the National Academy of Sciences of the United States of America was the journal with the highest production. Zhang Y is the most productive author with 65 (0.67%) publications. The 9,677 epigenetics articles were published by 8,145 different authoring organizations. University of Michigan is the institution where the most research related to epigenetics analysis has been developed, with 141 publications.

Discussion: Epigenetics is an emerging area of scientific research. However, most scientific research is conducted in environmental influences and relatively few are on human models. Analysis of the academic communities performing epigenetics research may provide scientific evidence. The method proposed in this research can provide reference for future bibliometric studies.

Conclusion: Advancement and development in a particular area of research are illustrated by the ever-growing body of scientific literature. This accumulation of literature describes various developments and innovations that occur over time and contains potentially valuable information that can be evaluated and classified to explain current emerging trends. In this study, we present a scientometric analysis of the retrieved papers published between 2012 and 2021 in the field of Epigenetics. Since 2012, the number of publications related to this field has grown rapidly. The United States and China are at the fore-front of Epigenetics research and the cooperation between these two countries is relatively close. In this research field, most productive institutes and authors come from United States and China, and the high frequency of Biochemistry and Molecular Biology shows that they are hot subjects.

https://doi.org/10.18231/j.ijlsit.2024.011

1. Introduction

The field of epigenetics has experienced remarkable research attention in recent years and there have been major advances in elucidation of how histone modification can regulate gene expression. Proteins in the sirtuin family are highly conserved from bacteria to humans. These proteins play an important role in epigenetic regulation. Epigenetics involves genetic control by factors other than an individual's DNA sequence.[1] Epigenetics is involved in many normal cellular processes. Consider the fact that our cells all have the same DNA, but our bodies contain many different types of cells: neurons, liver cells, pancreatic cells, inflammatory cells, and others. Epigenetic processes are natural and essential to many organism functions, but if they occur improperly, there can be major adverse health and behavioral effects.[2] While epigenetic changes are required for normal development and health, they can also be responsible for some disease states. The first human disease to be linked to epigenetics was cancer, in 1983.

Many bibliometric analysis methods and tools have been developed to help researchers in different research fields. Statistical analysis of the strength of this relationship between articles can help researchers identify the intellectual base of the discipline, important authors, and other bibliometric information. This article aims to sort out and summarize the research on epigenetics through bibliometric analysis.

2. Review of Literature

Kringel et al analysed the pubmed database search yielded 3,051 hits on epigenetics and drugs, starting in 1992 and peaking in 2016.[3] Annual citations increased to a plateau in 2000 and show a downward trend since 2008. Approved and investigational drugs in the Drug Bank database included 122 compounds that interacted with 68 unique epigenetic enzymes. Mills and Rahal reviewed of genome-wide association studies (GWAS) from 2005 to 2018 (3,639 studies; 3,508 traits) reveals extraordinary increases in sample sizes, rates of discovery and traits studied.[4] A longitudinal examination shows fluctuating ancestral diversity, still predominantly European Ancestry (88% in 2017) with 72% of discoveries from participants recruited from three countries (US, UK, Iceland). US agencies, primarily NIH, fund 85% and women are less often senior authors.

Carlos Olmeda Gomez analysed the literature on epigenetics published between 2009 and 2017 using qualitative and visualization techniques.[5] CiteSpace V software was used to establish an intellectual overview, based on 13,295 scientific articles and review papers were drawn from Web of science core collection in January 2018. Document co-citations were analysed and a variety of graphics was created. Zhang et al analysed the DNA methylation in CVDs from January 1, 2001, to September 15, 2021, were searched and confirmed from the Web of Science.[6] A total of 2,617 publications were included in 912 academic journals by 15,584 authors from 963 institutions from 85 countries/regions. Among them, the United States of America, China, and England were the top 3 countries contributing to the field of DNA methylation.

Barbosa et al performed a quantitative analysis regarding genetic polymorphisms on the GMR scientific production between 2009 and 2013.[7] We used the keywords polymorphism AND genetics OR molecular marker in order to conduct the literature survey. We found 423 articles related to genetic polymorphism and 87% were original articles. Six countries account for about 89% of publications and China is responsible for 56% of all publications. Shao et al elucidated the status of oncology research from 2001 to 2010.[8] Studies published in 30 representative oncology journals were retrieved from the Web of Science (2001-2010) to compose our dataset. Knowledge domain visualisation, co-citation analysis and social network analysis methods were used. Over the past 10 years, America has led oncology research, while China is the sole developing country to be ranked in the top 10.

Jia et al studied for publications on the Web of Science database in the field of atherosclerosis related to epigenetics was conducted from the earliest mention to 31 December 2020.[9] Data on total and annual publications, citations, impact factors, Hirsch (H)-index, citation times, most prolific authors, and frequently published journals were collected for quantitative and qualitative comparison. A total of 1,848 publications related to epigenetics and atherosclerosis were found. Haotian Jiang analysed the N6-methyladenosine publications were retrieved from Web of Science Core Collection and PubMed from 2000 to 2021, with keywords 'm^6A' and 'cancer', and analyzed in biblioshiny and VOSviewer.[10] A total of 1,013 documents were included, and China and the USA were the top countries with close collaboration.

Dario Kringel analyzed computational functional genomics of drugs with epigenetic interactions.[3] PubMed database search yielded 3,051 hits on epigenetics and drugs, starting in 1992 and peaking in 2016. Annual citations increased to a plateau in 2000 and show a downward trend since 2008. Approved and investigational drugs in the Drug Bank database included 122 compounds that interacted with 68 unique epigenetic enzymes. Chenchen Zhu analysed a total of 867 publications met the inclusion criteria, which spanned the period from 2000 to 2022 indexed in Web of Science database.[11] Over the years, the number of publications and the frequency of citations exhibited a clear upward trend in general, reaching a peak in 2021. The major contributing countries in terms of the

* *Corresponding author*.
E-mail address: santham74@gmail.com (R. Santhakumar).

number of publications were China, the United States, and Japan. Andy Wai Kan Yeung examined the incidence rate of reviews being mislabelled by Scopus, and compared this rate with Web of Science (WoS), PubMed and official websites of publishers. Top 400 cited publications defined by Scopus as 'articles' were examined.[12] Their contents were evaluated to see if any were actually reviews. These publications were cross-checked in WoS, PubMed and publisher websites to identify the assigned document type labels. Out of the 400 Scopus 'articles', 117 were reviews (29.3%).

Le Guan et al conducted a comprehensive search of the literature from January 1981 to December 2021 using the Web of Science core database.[13] The medical subject term 'visceral pain' was searched. A total of 5,047 articles were included in the analysis. The number of articles on visceral pain has continued to grow steadily over the past 40 years. The United States (1,716 articles), University of California (159 articles), and Neurogastroenterology and Motility (276 articles) were the country, institution, and journal with the most publications, respectively. Zou and Sun studied on 3131 publications of Web of Science to identify the current research status and research trends in this field.[14] The results show that since 2010, the number of publications has been growing rapidly. Cooperative network analysis shows that the United States, the University of Toronto and Roger S Mcintyre are the most influential countries, research institutes and scholars, respectively. Insulin resistance, obesity, and metabolic syndrome are hot topics in this field.[14,15]

3. Objectives for the Study

The following objectives have been framed for the study:

1. To study the annual growth rate of publications
2. To study the most prolific authors
3. To study the highly productive countries
4. To study the highly productive institutes
5. To study the most preferred source titles
6. To study the high productive subject areas

4. Materials and Methods

The research sample included 9,677 articles related to epigenetics obtained from the Web of science database from January 2012 to December 2021. Bibliometric exploration of web of science listed publications on the topic of epigenetics. In the present study, only articles and reviews were included. The information for the documents that meet the requirements contained year of publication, source title, author, affiliation, subjects which were exported into excel format. The date of the retrieval was 20th August 2022.

5. Data Analysis and Interpretations

5.1. Year-wise distribution of publications

A total of 9,677 epigenetics publications were published during 2012-2021. The highest number of publications was 1,183 (12.22%) published in 2021. The least no. of publication is found to be of 2012, because it was the first period of the study. The average number of publications published per year was 967.7. It reveals that research productivity does not remain constant each year. It increases or decreases year after year. The research productivity of epigenetics fluctuates year after year.

There was variation in annual growth during the study period and suddenly decreased from 12.30 in 2014 to -1.31 in 2015. Where as in suddenly increased up to 4.88 in the year 2016, it was decreased to -3.92 in 2017 and it was increased to 23.35 in the year 2018. Likewise, there was a variation after year as indicated in Figure 1 in the AGR for the publications. The significant reason for variations is that there is no constant growth of publications every year in the area of study.

Table 1: Annual growth rate

Year	Publications	Percentage	AGR
2012	668	6.90	-
2013	813	8.40	21.71
2014	913	9.43	12.30
2015	901	9.31	-1.31
2016	945	9.76	4.88
2017	908	9.38	-3.92
2018	1120	11.57	23.35
2019	1109	11.46	-0.98
2020	1117	11.54	0.72
2021	1183	12.22	6.60

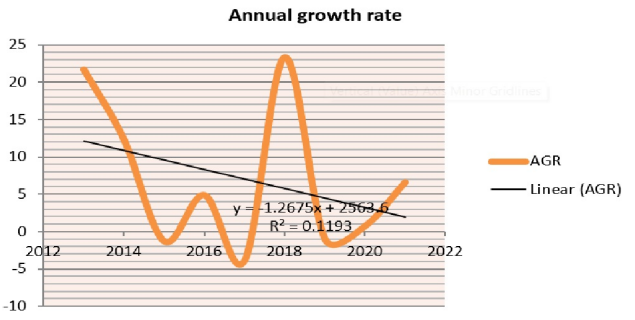

Figure 1: Relative growth rate for research output

The study reveals that the maximum number of publications has been published in English language with 9,593 publications (99.13%), followed by German language with 36 publications (0.37), French language ranks third position with 22 publications (0.23%). And the remaining languages such as Spanish, Russian, Japanese and other

Table 2: Language-wise distributions

S. No.	Language	Publications	Percentage
1	English	9593	99.13
2	German	36	0.37
3	French	22	0.23
4	Spanish	14	0.14
5	Russian	7	0.07
6	Japanese	2	0.02
7	Turkish	2	0.02
8	Hungarian	1	0.01
Total		**9677**	**100.00**

languages are constituted in negligible percentage. The English language superiority was found in every year in total productivity of the discipline.

5.2. *Most productive authors*

Table 3: Most productive authors

S. No.	Author	Publications	Percentage
1	Zhang Y	65	0.67%
2	Wang Y	62	0.64%
3	Xu Y	60	0.62%
4	Li Y	59	0.61%
5	Liu Y	56	0.58%
6	Wang J	48	0.50%
7	Chen Y	45	0.46%
8	Li J	42	0.43%
9	Wang L	42	0.43%
10	Li X	40	0.41%

The list of top 10 prolific authors in terms of productivity count is listed in Table 3. The list is ranked in the order of decreasing productivity. It is found that, Zhang Y is the most productive author with 65 (0.67%) publications followed by Wang Y with 62 (0.64%) publications, Xu Y with 60 (0.62%) publications, Li Y with 59 (0.61%) publications, Liu Y with 56 (0.58%) publications, Wang J with 48 (0.50%) publications, and Chen Y with 45 (0.46%) publications. And a total of 46,566 authors are contributed entire research output of the study period.

5.3. *Author's productivity*

In a study of author productivity, a large number of authors were classified according to the number of articles they had published during a certain period. To calculate Author productivity a formula has been applied. The formula is mathematically represented as below:

Average Author per Paper = No. of Authors/No. of Papers

Productivity per Author = No. of Papers/No. of Authors

It is one of the productivity of the total authors involved in contributing the research productivity. But

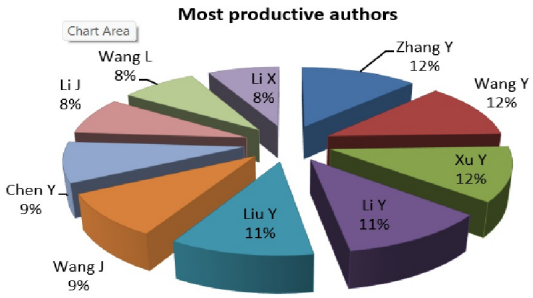

Figure 2: Most productive authors

Table 4: Author productivity

Year	No. of Authors	No. of Articles	Average Author Per Paper	Productivity per Author
2012	2940	668	4.40	0.23
2013	3822	813	4.70	0.21
2014	4454	913	4.88	0.20
2015	3883	901	4.31	0.23
2016	4056	945	4.29	0.23
2017	4322	908	4.76	0.21
2018	5365	1120	4.79	0.21
2019	5995	1109	5.40	0.18
2020	5511	1117	4.93	0.20
2021	6218	1183	5.26	0.19
Total	**46566**	**9677**	**4.77 (Average)**	**0.21 (Average)**

the contribution of individual authors can't measured in these methods of measurements. Table 4 depicts the data pertaining to author productivity and average author per paper. It is revealed from the table 4 that the average number of authors per articles is 4.77 for 9,677 articles published between the periods 2012 to 2021.

The average productivity per author for the period 2012 to 2021 is 0.21. It is also clear from above table that for the years 2012, 2015, 2016 and for the years 2013, 2017, 2018 years are equal average productivity per author is recorded i.e., 0.23 and 0.21 respectively.

5.3.1. *Degree of collaboration*

The degree of collaboration is defined as the ratio of the number of collaborative research papers to the total number of research papers in the discipline during a certain period of time. The formula suggested by Subramanyam (1983) is used. It is expressed as:

$C = N_m/N_m + Ns$

Where, C is the degree of collaboration in a discipline. N_m is the number of multiauthored research papers in the discipline published during a year. N_s is the number of single authored papers in the discipline published during the same year.

Table 5: Degree of collaboration

Year	Single Author		Multi Authors		Total	Degree of Collaboration (DC)	Collaboration Index (CI)
	No. of Output	Percentage	No. of Output	Percentage			
2012	73	7.08	595	6.88	668	0.86	4.40
2013	69	6.69	744	8.61	813	0.91	4.70
2014	89	8.63	824	9.53	913	0.90	4.88
2015	107	10.38	794	9.18	901	0.88	4.31
2016	123	11.93	822	9.51	945	0.87	4.29
2017	116	11.25	792	9.16	908	0.87	4.76
2018	104	10.09	1016	11.75	1120	0.91	4.79
2019	91	8.83	1018	11.77	1109	0.92	5.40
2020	132	12.80	985	11.39	1117	0.88	4.93
2021	127	12.32	1056	12.21	1183	0.89	5.26
Total	**1031**	**100.00**	**8646**	**100.00**	**9677**	**0.89**	**4.77 (Average)**

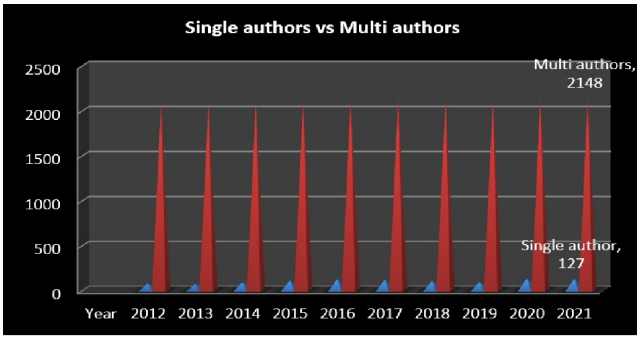

Figure 3: Single author vs. multi-authored articles

The authorship pattern was analysed to determine the percentage of single and multiple authors. Table 5 presents the single and multiple authors productivity pattern on yearly basis. There were 8,646 (89.35%) multi authored and only 1,031 (10.65%) single authored publications. The productivity patterns on the epigenetics publications are much contributed by the multiple authors than the single author since 2012 to 2021.

The degree of collaboration is determined by using this formula based on this study, the result of degree of collaboration C = 0.89. i.e, 89 percents of collaborative author's articles is published in this study. The degree of collaboration in producing research output on epigenetics research has shown a fluctuating trend during the study period.

Collaboration Index means number of authors per joint papers. Analysis in the Table 5 shows the variation in the Collaboration Index. It varies from 4.29 in 2016 and highest collaboration notices in 2019 i.e. 5.40. The average collaboration index is 4.77. It implies the research team falls between 4 and 5 authorship patterns in field of epigenetics.

Table 6: Highly productive countries

S. No.	Country	Publications (%)	S. No.	Country	Publications (%)
1	USA	3941 (40.72%)	9	France	438 (4.53%)
2	China	1526 (15.77%)	10	Australia	373 (3.85%)
3	Germany	929 (9.60%)	11	Netherlands	372 (3.84%)
4	England	814 (8.41%)	12	Sweden	277 (2.86%)
5	Japan	593 (6.13%)	13	India	266 (2.75%)
6	Canada	589 (6.09%)	14	Brazil	221 (2.28%)
7	Italy	521 (5.38%)	15	Switzerland	219 (2.26%)
8	Spain	448 (4.63%)	16	South Korea	204 (2.11%)

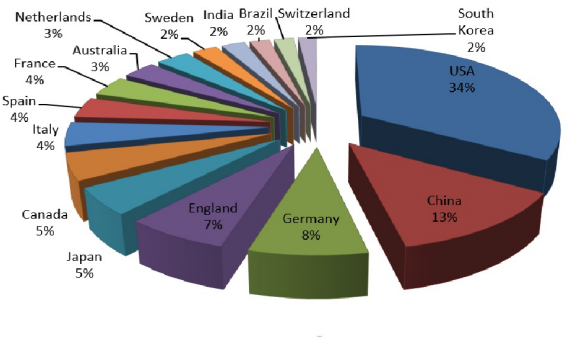

Figure 4: Highly productive countries

5.4. Highly productive countries

Analysis of worldwide research activity revealed that publications on epigenetics originated from 117 countries. 3,941 (40.72%) articles published in the United States ranked first place, which was 25.02% higher than those

in China, whose publication number was 1,526 (15.77%), thereby ranking second. India should ranks 11th in the list, which is a developing country with 266 publications. The countries are ordered according to the total number of publications.

Table 7: Highly productive institutes

S. No.	Institutions	Country	Publications	Percentage
1	University of Michigan	USA	141	(1.46%)
2	Johns Hopkins University	USA	136	(1.40%)
3	Harvard Medical School	USA	132	(1.36%)
4	University Pennsylvania	USA	132	(1.36%)
5	Harvard University	USA	130	(1.34%)
6	McGill University	Canada	122	(1.26%)
7	Karolinska Institute	Sweden	114	(1.18%)
8	Chinese Academy of Science	China	112	(1.16%)
9	Emory University	USA	111	(1.15%)
10	University of Cambridge	England	108	(1.12%)

The 10 most productive institutions are shown in Table 7, taking into account the number of publications since 2012. The 9,677 epigenetics articles were published by 8,145 different authoring organizations. These institutions are related to the number of publications per author and the frequency per affiliation; Thus, University of Michigan is the institution where the most research related to epigenetics analysis has been developed, with 141 publications, followed by Johns Hopkins University with 136 publications. Interestingly, most of the universities in this table are from USA, it is no surprise that US organizations occupy the top five spots.

5.5. Most preferred source titles

A total of 1,945 journals published publications related to deep learning in epigenetics research. The top 10 journals are presented in Table 8. The highest count belonged to the Proceedings of the National Academy of Sciences of the United States of America (n=263), followed by International Journal of Molecular Sciences (n=168). Among these journals, Genes Development had the highest impact factor (12.89).

Table 8: Source title of publications

S. No.	Source Title	Publications	Percentage	Impact Factor
1	Proceedings of the National Academy of Sciences of the United States of America	263	2.72	12.78
2	International Journal of Molecular Sciences	168	1.74	6.208
3	Journal of Biological Chemistry	145	1.50	5.486
4	Scientific Reports	133	1.37	4.996
5	Biochemical and Biophysical Research Communications	77	0.80	3.575
6	FASEB Journal	76	0.78	5.834
7	BMC Cancer	64	0.66	4.4
8	Genes Development	63	0.65	12.89
9	Frontiers in Plant Science	61	0.63	6.627
10	Biology of Reproduction	58	0.60	4.285

Table 9: High productivity subject areas

S. No.	Subject	Articles	Percentage
1	Biochemistry Molecular Biology	1345	13.90
2	Research Experimental Medicine	1017	10.51
3	Behavioral Sciences	990	10.23
4	Neurosciences Neurology	935	9.66
5	Oncology	817	8.44
6	Genetics Heredity	682	7.05
7	Biophysics	663	6.85
8	Chemistry	629	6.50
9	Science Technology	603	6.23
10	Pharmacology Pharmacy	557	5.76

5.6. High productivity subject areas

The scientific literature on epigenetics research is spread over 87 different subjects. The top 10 subjects with their frequencies are shown in Table 9. It is found that Biochemistry Molecular Biology has highest number of articles with 1,345 (13.90%) followed by Research Experimental Medicine contributing 1,017 (10.51%) articles. Behavioral Sciences occupy the third position with 990 (10.23%) articles. The fourth highest articles belonged to the subject Neurosciences Neurology with 935 (9.66%), Oncology with 817 (8.44%) and Genetics Heredity with 682 (7.05%) articles respectively.

6. Conclusions

In the present analysis, research and development of epigenetic scientometric information queried from web of science database.

6.1. Summary Points

1. In total, 9,677 documents were included in this study from 2000 to 2021, with 7,840 articles and 1,837 reviews.
2. Review articles were the lowest publication rate compared to original research.
3. The publications showed fluctuated during the study period.
4. USA and the China were the top two ranked countries of corresponding authors. Meanwhile, USA institutions occupy the top five spots.
5. Multi authored documents contained an average of 8.9 authors, with only 11% of the documents being single authored. The collaboration index (CI) averaged 4.77.
6. The United States has been leading the rank for years, but China's publications have been gradually growing.
7. USA and University of Michigan were the most influential country and institution, respectively.
8. 9,677 articles related to epigenetics were published in 1945 academic journals. Proceedings of the National Academy of Sciences of the United States of America published the most papers (n = 263; 2.72%).

Overall, the present study uses multiple complementary bibliometric methods to generate a panoramic view of the recent developments in epigenetics research.

7. Source of Funding

None.

8. Conflict of Interest

None.

References

1. Hamilton JP. Epigenetics: Principles and practice. *Dig Dis*. 2011;29(2):130–5.
2. Moosavi A, Ardekani AM, A. Role of epigenetics in biology and human diseases. *Iran Biomed J*. 2016;20(5):246–58.
3. Kringel D, Malkusch S, Lötsch J. Drugs and epigenetic molecular functions. A pharmacological data scientometric analysis. *Int J Mol Sci*. 2021;22(14):7250.
4. Mills M, Rahal C. A scientometric review of genome-wide association studies. *Commun Biology*. 2019;2(1).
5. Olmeda-Gomez C, Roma-Mateo C, Ovalle-Perandones MA. Overview of trends in global epigenetic research (2009–2017). *Scientometrics*. 2009;119(18).
6. Zhang Y, Jia Z, Zhou Q, Zhang Y, Li D, Qi Y, et al. A bibliometric analysis of DNA methylation in cardiovascular diseases from 2001 to 2021. *Med (Baltimore)*. 2022;101(33):e30029.
7. Barbosa AM, De Silva K, Laqares MH, Rodriques DA, Martins JV, Moura KK, et al. Scientometrics analysis: Five years of genetic polymorphisms. *Genetics Molecular research*. 2018;17(2).
8. Shao H, Yu Q, Bo X, Duan Z. Analysis of oncology research from 2001 to 2010: A scientometric perspective. *Oncol Rep*. 2013;29(4):1441–52.
9. Jia L, Cheng A, Alam N, Qian Y, Ma Z, Ren H, et al. Global trends in atherosclerosis research in the epigenetics field: Bibliometric and visualization studies. *Int J Environ Res Public Health*. 2021;18(24):13154.
10. Jiang H, Li Y, Liu Y, Dong W. Bibliometric analysis of N6-methyladenosine in cancer: Landscapes, highlights and trending topics. *Future Oncol*. 2022;18(17):2141–53.
11. Zhu C, Zhong W, Gong C, Chen B, Guo J. Global research trends on epigenetics and neuropathic pain: A bibliometric analysis. *Front Mol Neurosci*. 2023;16:1145393.
12. Yeung AW. Comparison between Scopus, Web of Science, PubMed and Publishers for mislabelled review papers. *Curr Sci*. 2019;116(11):1909–14.
13. Guan L, Liu Y, Wu B, Chen A, Tao W, Lin C. Research hotspots and trends in visceral pain research: A global comprehensive bibliometric analysis. *Front Mol Neurosci*. 2022;15:1022463.
14. Zou X, Sun Y. Bibliometrics analysis of the research status and trends of the association between depression and insulin from 2010 to 2020. *Front Psychiatry*. 2021;12:683474.
15. Ho YS, Ranasinghe P. A bibliometric analysis of highly cited insulin resistance publications in Science Citation Index Expanded. *Obes Med*. 2022;31:100399.

Author biography

R. Santhakumar, Librarian https://orcid.org/0000-0001-5374-9910

S. Mahalakshmi, Librarian

V. Kalaiappan, University Librarian Incharge

Cite this article: Santhakumar R, Mahalakshmi S, Kalaiappan V. Bibliometric analysis of the structure and evolution of research on epigenetics. *IP Indian J Libr Sci Inf Technol* 2024;9(1):65-71.