University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

6-19-2020

Research Productivity in Human Genetics during 2014 – 2018: A Scientometric Approach

Priya Saravanan
Periyar University, Salem, India, pupriya7@gmail.com

Natarajan Radhakrishnan
Periyar University, Salem, India, mkuradha@gmail.com

Abdoulaye Kaba
Al Ain University, kaba_abdoulaye@yahoo.com

Research Productivity in Human Genetics during 2014 – 2018: A Scientometric Approach

Abstract

An attempt is made for bringing out the growth of literature on human genetics. Web of Science database has been used to retrieve the data of five years (2014-2018) by searching a keyword "Human Genetics" in the combined topics, abstract and keywords fields. The scientometric tools and techniques such as document type, year wise distributions, language, and country wise distribution are used for measuring the outputs. Findings of the study indicate that more than 1966 (20.51%) research articles were published in the year 2016. English was the dominant language with 9393 (97.9%) followed by German with 71(0.7%). For the contributing top ten most productive countries, USA topped the list with 46.8%, followed by UK (12.9%), and China (10.1%). The frequency occurrence of words among the publications revealed that the word 'Human' is largely found with 1545 (16.1%) and followed by a word 'genetics' which occurred in 1389 (14.5%) publications. The results of Straight-line Equation and Time Series Analysis predict future trend of growth of research outputs in human genetics to reach 6374 research papers in the year 2034.

Keywords: *human genetics, scientometric, Web of Science*

1. Introduction

Scientometrics is generally concerned with the quantitative features of scientific research. It has been used to mean communication development in science including socio-cultural aspects. Scientometrics is a branch of the library and information science. Assessment of scientific output and exploration on different aspects of research areas can be explored using scientometric studies. Evaluation plays a pivotal role in the decision-making and policymaking process of any academic field. Scientometric methods are recognized as the best way to monitor research trends and scholarly activities. Scientometric study can be carried out on emerging research topics in any discipline[1]. Scientometric methods have been applied to understand the quantitative technique for academic efficiency[2].

Human genetics provides a theoretical structure for considering the biology of the human genus. Fields such as biologic, chemical, medical, and statistical sciences are being applied for solving genetic related problems. Genetics is called DNA molecules which embrace all the genetic information for an organism. It provides cells with the information to allow an organism to develop endure and reproduce. The heritage of genetic substance by offspring influences the emergence and activities of the tissue as well. A few decades ago, human genetics was known as medical field mainly deal with rare syndromes and prenatal diagnostics.

Genetic alterations play an emerging role in human neoplasia. Generally, genetic alterations take place in a particular somatic cell, which divides and develop until it forms cancer. Rarely, when malignant neoplasm occurs as part of a hereditary cancer syndrome, the primary changes are inherited through a germ line current in all cells of the organism. The present study emphasized a devotion on the scientometric investigation of the pattern of publication, time series analysis, and language-wise distribution by using keywords wise analysis for human genetics research.

2. Related literature

Previous studies have been conducted to investigate the research trends of human genetics using scientometrics methods i.e. authorship pattern, activity index, Country wise distribution, keyword wise analysis and collaborative research. The present study made an endeavor to cover human genetics keywords discussions about the diverse techniques of scientometrics by Singh[3] who indicated the growth trends using authorship pattern and collaboration in Biotechnology for sixteen years with a sample of 18918 articles. The data were downloaded from Scopus database during the period 2001-2016. Scientometric tools such as collaboration coefficient (CC), authorship

pattern (AP), activity index (AI), Relative Growth Rate (RGR) and Doubling time of Literature (Dt.) have effectively been applied. It showed that multi-authorship articles are higher (average) and dominant over single authorship pattern. Doubling time for different years has gradually led to an increasing trend. It has been analyzed the characteristics of most productive researchers in Biotechnology through team research rather than the individual. The average activity index of India's contribution was 91.78%. Highest AI year for India is 2016 was 180.3% and the lowest contribution in 2001 was 42.38.

Fu, Wang, and Ho[4] conducted a bibliometric study based on the Science Citation Index Expanded database from the Web of Science which was carried out to provide insights into research activities from 1992 to 2011. It emphasizes the performance such as annual output, mainstream journals, Web of Science categories, leading countries, institutions, and research tendencies. The results indicated that the annual output of the related scientific articles was increased steadily. Water Research, Environmental Science & Technology, and Journal American Water Works Association were the majority of three general journals for drinking water research. Keyword obtained from the title, author keyword and Keywords Plus analysis provided ways to determine the recent research trends. Garg, Kumar, Chakraborty[5], demonstrated a study on 2899 articles published by Indian scientists during the period 1991-2008. However, the maximum productivity was in the sub-field of molecular genetics. The study mainly focused on molecular genetics involving to organisms of humans, plants, and animals which published the highest number of papers.

The main aim of this paper is to identify the outputs of genetics research based on the distribution of publications in different sub-specialties. The study compares the research priorities of 16 sub-specialties of genetics from13 countries during two time-spans i.e. 2002-2006 and 2007-2011. It attempts to investigate quantitative growth and expansion of Mass Spectrometry consider in Nuclear Science and Technology in terms of publication productivity as reflected in International Nuclear Information System (INIS) database (1970-2005) by Sagar, Kademani, and Kumar[6]. During 1970-2005, a total of 10913 documents was published in various domains: Chemistry, Materials and Earth Sciences (5286) (48.44%), Physical Sciences (2367) (21.69%), Engineering and Technology (1434) (13.14%), Life sciences and Environmental Sciences (1212) (11.11%) followed by the other aspects of records were Nuclear and Non-Nuclear Energy (492) (4.51%) and Isotopes, Isotope and Radiation Applications (122) (1.12%). There were only three papers published in 1970. The maximum number of records (816) was published in 2004. Molatudi and Pouris[7] conducted a bibliometric method to collect data from the ISI database on the show of microbiology, genetics and molecular biology research over a 20-year during 1980 - 2000.

Recently, Barbosa et.al[8] conducted a study using the Genetics and Molecular Research (GMR) database. GMR is a fully electronic journal available to readers through the Internet and the literature on human genetics. They performed a quantitative analysis regarding genetic polymorphisms on the GMR scientific production during the period 2009 -2013. The organisms most considered in those papers were humans (61.2%). Similarly, Kumaragurupari and Geetha[9] analyzed the genetic research in Glaucoma during 2006-2015 on several parameters such as contribution and citation impact of top most productive countries, growth pattern, leading institutions, authorship pattern, type of communication and output. The data were downloaded from the PubMed database for 10 years (2006-2015) using the keywords "glaucoma" and "Gene" in a combined title, abstract and keyword fields. The results indicated that the total number of published articles was retrieved during 2006 - 2015 was 555. During the ten-year period, they studied the annual output of research articles, which demonstrated steadily increase in quantity. Multiple authors (96.6%) were predominant in genetic research in Glaucoma, and two-thirds of articles were published in Ophthalmology Journals.

The literature review indicated many studies related to human genetics research output which analyzed the scientific research productivity in genetics. Hence, human genetics research output exist in the web of science database using scientometrics approach. However, few studies have been untaken on genetics research using metrics methods.

3. Research Method

The search string "Human Genetics" in the "Topic search" field was selected in the search field and the time span 2014-2018 was used as a restriction for the publication data and downloaded the records on the keyword 'Human Genetics'. It was selected for more setting fields and access to the all Web of Science Core Collection Citation Indexes on Citation Index (SCI), Social Science Citation Index (SSCI), and Arts & Humanities

Citation Index (A&CHI). The Web of Science (WoS) bibliographic database is one of the online databases, which has been subscribed since 2011 by Periyar University Library in Salem, India.

This study analyzed the research output of human genetics in five years during the period 2014 - 2018. The data were downloaded from the Web of Science database and 9590 records were retrieved for the present study. The data were organized, tabulated, analyzed through His cite software and statistical methods. Various scientometric measures like the growth of literature, document type, and year-wise distribution of articles, language, and the country-wise contribution from the keyword analysis have also been appropriately applied. VOS viewer (software) for creating maps based on network records to visualize figures.

Table1. Year-wise distribution of publications

S.No	Publication Year	Records	Percentage	Cumulative Growth	Exponential growth rate
1	2014	1832	19.10	1832	-
2	2015	1907	19.88	3739	0.96
3	2016	1966	20.51	5705	0.97
4	2017	1937	20.19	7642	1.01
5	2018	1948	20.32	9590	0.99
	Total	9590	100		

4. Data analysis and Findings

Findings of the study identified that the highest number of articles i.e. 1966 (20.51%) was published in the year 2016 and the lowest number of articles i.e. 1832 (19.10%) was published in 2014. Overall records of keyword analysis of human genetics during study period (2014–2018) are discussed. Exponential growth rate of publications in human genetics is identified. The highest growth rate of 1.01 was found in 2017 and the lowest growth rate of 0.96 in 2015.

The straight-line equation is applied to arrive at estimates for future growth under the Time Series Analysis.

Straight Line Y c= a+bX:

Since $\sum x = 0$

$$a = \frac{\sum y}{N} = \frac{9590}{5} = 1918$$

$$b = \frac{\sum xy}{\sum x^2} = \frac{2228}{10} = 222.8$$

Estimated literature in 2034 is 20 (when X = 2034 − 2014)

1918 + (222.8 × 20) = 6374

The predicted value of literature output for the year 2034 is **6374.**

On the application of time series analysis and subsequently the results obtained separately for the years 2034. It is found that the future trend of growth of research in human genetics towards International output may lead to an increasing trend in the years to come. The inference from the calculations proved a growth could be seen in India too.

Table 2. Growth of Literature -Time Series Analysis

S.No	Year	No. of Publications	X	X^2	XY
1	2014	1832	-2	4	-3664
2	2015	1907	-1	1	-1907
3	2016	1966	0	0	1966
4	2017	1937	1	1	1937
5	2078	1948	2	2	3896
Total		9590			2228

Different types of documents have are identified from human genetics. It was fact that most of the scholarly statement of scientific research is published in articles i.e. 6900 (71.9%). followed by review with 1972 (20.6%), editorial material with 258 (2.7%) Meeting Abstract with 162(1.7%), Book Review with 99 (1.0%), Article; Proceedings Paper, with 93 (1.0%) and Book Chapter with 38 (0.4%).

Table 3. Document wise distributions

S. No	Document type	Records	%	Cumulative
1	Article	6900	71.9	6900
2	Review	1972	20.6	8872
3	Editorial Material	258	2.7	9130
4	Review; Book Chapter	162	1.7	9292
5	Book Review	99	1.0	9391
6	Meeting Abstract	93	1.0	9484
7	Letter	38	0.4	9522
8	Article; Proceedings Paper	19	0.2	9541
9	Biographical-Item	15	0.1	9556
10	News Item	12	0.1	9568
11	Article; Book Chapter	10	0.1	9578
12	Correction	8	0.1	9586
13	Reprint	4	0.0	9590

	Total	9590	100	

Top ten Languages wise distribution has been explored on human genetics. English language dominate with 9393 (97.9%) and followed by the German with 71(0.7%). The lowest number of contributions was found in Hungarian, Italian, Polish, and Turkish with 6 papers each.

Table 4. Language wise distributions

S.No	Language	Records	Percentage
1	English	9393	97.9
2	German	71	0.7
3	Spanish	42	0.4
4	French	21	0.2
5	Russian	17	0.2
6	Portuguese	8	0.1
7	Hungarian	6	0.1
8	Italian	6	0.1
9	Polish	6	0.1
10	Turkish	6	0.1

For the contributions at the countries' level, USA contributed 4486 (46.8%), UK produced 1242 (12.9%) followed by People R China i.e. 973 (10.1%). Germany contributed 868 (9.0%), France published 595 (6.2%), and Canada contributed 575 (6.0%). Italy published 531 (5.5%) and Australia contributed 502 (5.2%). The lowest contribution of the records was from Spain i.e. 419 (4.4%). The highest number of citations was received by USA i.e. 1115(185).

Table 5. Country distributions of citations

S.No	Country	Records	Percentage	Total Citations	ACPP
1	USA	4486	46.8	1115	185
2	UK	1242	12.9	1034	258
3	Peoples R China	973	10.1	784	130
4	Germany	868	9.0	781	130
5	France	595	6.2	648	108
6	Canada	575	6.0	629	209
7	Italy	531	5.5	613	204
8	Australia	502	5.2	560	186
9	Netherlands	493	5.1	506	84
10	Spain	419	4.4	500	83

Figure 1. VOS viewer screenshot of the countries
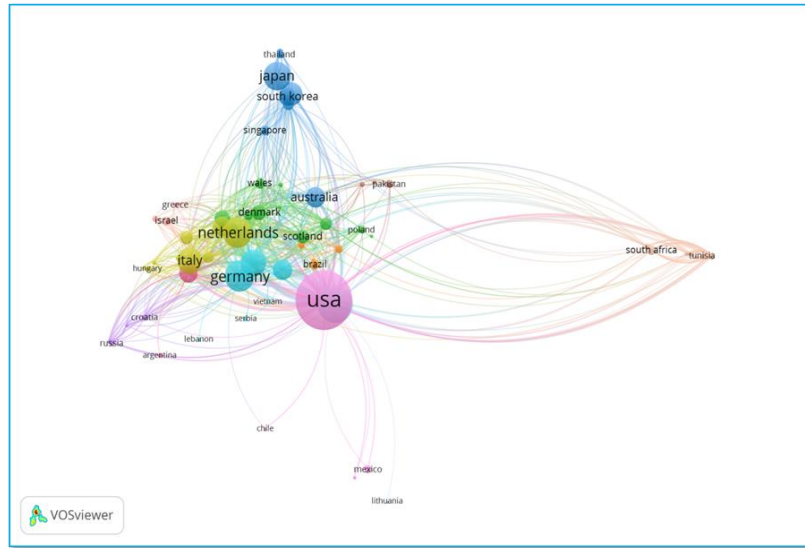
Table 6. Keyword distributions

S.No	Word	Records	Percentage
1	Human	259	21.22
2	Genetics	235	19.26
3	Genetic	167	13.68
4	Disease	105	8.60
5	Gene	95	7.78
6	Population	89	7.29
7	Genome	83	6.80
8	Analysis	78	6.39
9	Association	76	6.22
10	Genes	69	5.69

As present in Figure 2, the keyword 'Human' is found high with 259 (21.22%) frequent occurrence of words. The term 'Genetics' was found in 235 (19.26%) publications, 'Genetic' occurred in 167 (13.68%) publications, and followed by the word 'Disease' with 105 (8.60%) and 'Gene' with 95 (7.78%) 'Population' with 89 (7.29%) and 'Genome' with 83 (6.80%). 'Analysis' contributed in 78 (6.39%) and remaining keywords i.e. 'Association' with 76 (6.22%) and 'Genes' with 69 (5.69%).

Figure 2. VOS viewer screenshot of the keyword
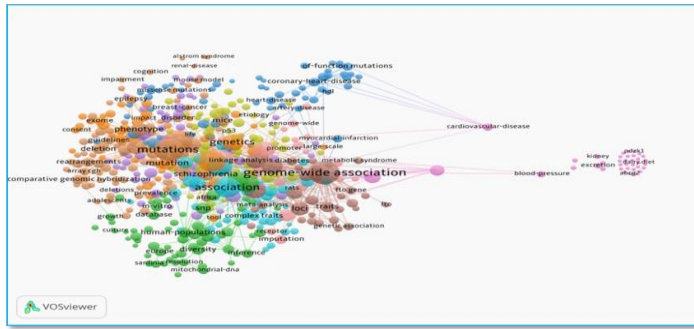
The human genetics research output has brought out in various productive journals. Only top-ranked source journals were selected for the present study. The highest number of the records was published in Plos One i.e. 280 (2.9). Lowest number of records was published by journal of Peripheral Nervous System i.e. 66 (0.7%). IF of 2018 has been taken and presented accordingly.

Table 7. Journal wise distributions

S.No	Journal	Records	Percentage	Country	IF
					2018
1	Plos One	280	2.9	United States	2.766
2	Scientific Reports	145	1.5	United Kingdom	4.011
3	Proceedings of the National Academy of Sciences of the United States of America	126	1.3	United States	9.38
4	Journal of Virology	98	1.0	United States	4.368
5	Forensic Science International Genetics	90	0.9	Netherlands	5.637
6	Plos Genetics	81	0.8	United States	5.540
7	Genetics	74	0.8	United States	4.075
8	BMC Genomics	71	0.7	United Kingdom	3.730
9	Molecular Ecology	71	0.7	United Kingdom	6.131
10	Journal of the Peripheral Nervous System	66	0.7	United Kingdom	2.550

A broader classification on title wise distribution has been identified. Top 15 title wise citations has been listed out for better understanding based on source title during the study period. Genetic Screens in Human Cells Using the CRISPR-Cas9 System article received a large number of citations i.e. 1100 (183.33). Ensemble 2018 article received a lowest number of citations was 343 (171.5). Total global citation and average citation scores were included in the study for easy reference.

Table 8. Title distributions of the citations

S.No	Title	Publication Year	Source	Total Citations	Average Citations
1	Genetic Screens in Human Cells Using the CRISPR-Cas9 System	2014	Science	1100	183.33
2	The amyloid hypothesis of Alzheimer's disease at 25years	2016	EMBO Molecular Medicine	982	245.5
3	Genetics of rheumatoid arthritis contributes to biology and drug discovery	2014	Nature	756	126
4	Human Genetics Shape the Gut Microbiome	2014	Cell	745	124.17
5	The Human Gene Mutation Database: building a comprehensive mutation repository for clinical and molecular genetics, diagnostic testing and personalized genomic medicine	2014	Human Genetics	636	106
6	The family based association test method: strategies for studying general genotype-phenotype associations	2017	European Journal of Human Genetics	627	209
7	Sex-specific, male-line transgenerational responses in humans	2017	European Journal of Human Genetics	610	203.33
8	COSMIC: somatic cancer genetics at high-resolution	2017	Nucleic Acids Research	523	174.33
9	A Proteome-Scale Map of the Human Interactome Network	2014	Cell	489	81.5
10	Xpert (R) MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults	2014	Cochrane Database of Systematic Reviews	486	81
11	The Database of Genomic Variants: a curated collection of structural variation in the human genome	2014	Nucleic Acids Research	413	68.83
12	An Expanded View of Complex Traits: From Polygenic to Omnigenic	2017	Cell	375	125
13	A framework for the interpretation of de novo mutation in human disease	2014	Nature Genetics	349	58.17
14	PGC-l alpha mediates mitochondrial biogenesis and oxidative phosphorylation in cancer cells to promote metastasis	2014	Nature Cell Biology	344	57.33
15	Ensemble 2018	2018	Nucleic Acids Research	343	171.5

The Activity Index between Indian and World output on human genetics research was analyzed. Articles published in India were compared with the publications in human genetics research across the globe, and an activity index is derived. The highest number of the articles published in Activity Index range was 110 (2018). The lowest number of articles published by Activity Index range was i.e. 19 (2014). This study examined the average contribution of Activity Index i.e. 81.4.

Table 9. Activity Index

Publication Year	World Records	Indian Records	Activity Index
2014	1832	30	19
2015	1907	23	78
2016	1966	28	95
2017	1937	31	105
2018	1948	32	110
Total	9590	144	81.4

5. DISCUSSION

A scientometrics study explored the research trends in Human Genetics research during the period 2014 - 2018. The growth of articles by year wise distribution; country-wise distribution; Document type; Activity Index; Time Series Analysis and Journal wise distribution have been identified and presented. Year-wise distribution of publication output during 2014 - 2018 (5 years) has led to publications of 9590. Highest number of publications i.e 1966 (20.51%) was published in 2016 followed by 2018 (1948) (20.32%) Very low number of publications (19.10%) published in 2014. The highest growth rate of 1.01 was found during 2017 with 1937 publications. The frequency occurrence of words among the publications revealed that the word human is highest with 259 (21.22%) and highest number of contributions. A word Genetics occurred in 235 (19.26%) publications. The study mainly considered the publications from the articles using the keywords.

As far as country-wise distribution of the publication is concerned we have found top 10 countries during 2014-2018. It shows that the USA contributed 4486 (46.8) records and got ranked a top among the countries. A maximum number of 1115 citations was received by USA and occupied the first rank and followed by England with 1034. The huge number of average citation per records was 258 from England. Journal wise contribution for top ten records was published from the different countries were scattered across the globe. Among these publishing countries, USA that topped the list with 280 journals followed by England with 145 journals. The highest number of (97.9) research article was published in English language and occupied the first rank and followed by the German language with 71(0.7) which got placed in second position. The maximum number of articles on human genetics research was published in English Language. The Activity Index was high in 2018 (110) whereas the lowest contribution was found in 2014 (19).

6. Conclusion

Human genetics comes under the purview of both fundamental and applied sciences. It is a part of genetics which examines the laws of storage, transmission, and realization of information for the development and functions of living organisms. This study has examined the overall research output of Human genetics research during the period 2014 - 2018. The growth of articles by year and volume; language wise distribution; document wise distributions; time series analysis, country wise contributions and journal wise distribution were identified. Year wise distribution of total publications revealed the human genetics research. There were decreasing and increasing trends found during the five years. The study has examined 9590 articles which were published during study period.

The highest number of articles i.e. 1966 (20.51%) was published in 2016 and lowest was 1832 (19.10%) in 2014. Most of the research articles have been published in English language i.e. 9393 (97.9%) in the field of human genetics. It has been identified that human genetics became one of the emerging areas of research in life sciences. Number of funding agencies provides financial support to academics to undertake research on upcoming areas. No such an effort has been made to bring out the research trends in human genetics so far during 2014-2018. Hence it is suggested that the present work will stimulate the readers and academics to concentrate on human genetics for further research with the applications of different tools.

Reference

1. Ginn, L. K. Citation analysis of authored articles in library & information science research, 2001– 2002. *Mississipi Libraries*, 2003, **67,** 106–109.
2. Whitley, E.A. & Galliers, R.D. An alternative perspective on citation classics: Evidence from the first 10 years of the ECIS Conference. *Information & Management*, 2007, **44**(5), 441-455.
3. Singh, M. K. Authorship Pattern and Collaboration Coefficient of India in Biotechnology research during 2001-2016: Based on Scopus database. *Library Philosophy & Practice, 2017* (ejournal 1549).
https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=4442&context=libphilprac (accessed on 15 June 2020).
4. Fu, H.Z., Wang, M.H., & Ho, Y.S. Mapping of drinking water research: a bibliometric analysis of research output during 1992-2011. *Sci Total Environ*, 2013, 443:757-765.
doi:10.1016/j.scitotenv.2012.11.061
5. Garg, K. C., Kumar, S., Dutt, B., Chakraborty, O. Scientometric profile of genetics and heredity research in India. Annals of Library and Information Studies, 2010, 57: 196-206.
http://nopr.niscair.res.in/bitstream/123456789/10233/4/ALIS%2057%283%29%20196-206.pdf (accessed on 1 June 2020).
6. Sagar, Anil, Kademani, B.S., & Vijai Kumar. (2007). Scientometric mapping of mass spectrometry research in nuclear science and technology: a global perspective. *Proceedings of Twelfth ISMAS Symposium cum workshop on mass spectrometry*, (pp. CD-ROM). India: *Indian Society for Mass Spectrometry*, 2007, 39(16).
https://www.researchgate.net/profile/Anil_Sagar4/publication/251949682_Scientometric_Mapping_of_Mass_Spectrometry_Research_in_Nuclear_Science_Technology_A_Global_Perspective/links/564ae9d108ae295f64504d20/Scientometric-Mapping-of-Mass-Spectrometry-Research-in-Nuclear-Science-Technology-A-Global-Perspective.pdf (accessed on 4 June 2020).
7. Molatudi,M. & Pouris, A. Assessing the Knowledge base for biotechnology in south Africa: A bibliometric analysis of South Africa microbiology molecular biology and geneties research. *Scientometrics*, 2006, **68** (1), PP. 97-108.
8. Barbosa, A. M., De Silva, K. S. F., Lagares, M. H., Rodrigues, D. A., Martins, J. V. M., Da Costa, I. R., & Moura, K. K. V. O. Scientometric analysis: five years of genetic polymorphisms. *Genetics and Molecular Research*, 2018, 17(2): 1-7.
https://www.geneticsmr.org/articles/scientometric-analysis-five-years-of-genetic-polymorphisms.pdf (accessed on 10 June 2020).
9. Kumaragurupari, T., & Geetha, V. A Scientometric analysis of genetics research on Glaucoma. *Journal of Advances in Library and Information Science*, 2017, **6**(1), 01-06.
https://jalis.in/pdf/6-1/Guru.pdf (accessed on 15 June 2020).