Status of Human DNA Research in the United States of America: A Scientometric Analysis

P. Murugiah,
(PhD- Part Time Research Scholar, Alagappa University)
CSIR-Central Electrochemical Research Institute, Karaikudi 630 006, Tamil Nadu, India.
murugiahp.cecri@gmail.com

C. Baskaran,
Deputy Librarian,
Department of Library & Information Science,
Alagappa University, Karaikudi 630 006,
Tamil Nadu, India.
E-mail: cbklis@gmail.com

Abstract

Scientometric analysis of human DNA research output of the USA has been studied from 7264 research papers published during 1984-2012 that indexed in Scopus. The study reveals that the highest number of 337 papers or 4.64% published in 1996, followed by 314 (4.32%) papers in 1999. According to the Relative Growth Rate and Doubling Time of the publications value measured from 0.03 in 2010 and 2011 to 0.087 in 1985; from 0.80 in 1985 to 23.10 in 2010 and 2011 respectively found in this study. Among the document types, journal articles were the highest numbers with 7210 papers or 99.26%. From this study, it is observed that the Journal of Biological Chemistry has published with 529 research papers and find top position which is accounted for 7.28% of the total articles.

Keywords

Human DNA; Genetic; Degree of Collaboration; Relative Growth Rate; Doubling Time; Scientometric analysis.

Electronic access

The journal is available at www.jalis.in

Journal of Advances in Library and Information Science
ISSN: 2277-2219 Vol. 3. No.4. 2014. Pp328-334

INTRODUCTION

DNA, or deoxyribonucleic acid, is the hereditary material in humans and almost all other organisms. It is a macromolecule with a double helical structure, built up from four types of nucleotides, each of them containing a sugar, a phosphate group (making up the DNA backbone) and a nucleobase. DNA was first discovered by a Swiss Physician Dr. Friedrich Miescher in 1869. Then its double helical structure was discovered in 1953, by James Watson and Francis Crick. Ever since the discovery of DNA structure, it continues to make great advancement in understanding the human genome and the importance of DNA to life and health. There are only four nucleobases in the DNA structure: Adenine (A), Guanine (G), Cytosine (C) and Thymine (T) (Cardos et al., 2008). The human genome is subdivided into a large nuclear genome with more than 26,000 genes, and a very small circular mitochondrial genome with only 37 genes. The nuclear genome is distributed between 24 linear DNA molecules, one for each of the 24 different types of human chromosome (Organization of the Human Genome). The human genome holds an extraordinary trove of information about human development, physiology, medicine and evolution (Lander E.S et. al., 2001). The modern history of DNA sequencing began in 1977, when Sanger reported his method for determining the order of nucleotides of DNA using Chain – terminating analogs (Sager, F et al., 1977). DNA isn't just about growth. It instructs cells throughout your life — telling them how to respond to the foods you eat, the germs you encounter and the pollutants to which you are exposed. Ultimately, DNA even influences how you age. Genetic tests also are available to help couples learn if they carry genetic mutations for rare diseases and if they are likely to have a child affected by the disease (National Human Genome Research Institute).

REVIEW OF LITERATURE

Baskaran, C. (2013), conducted a study on Research Productivity of Alagappa University during 1999-2011 and found that the multi-authored papers are in higher numbers (750; 96.64%) than single – authored papers (26; 3.35 %). Gupta B.M., et al, (2013), have studied the global publication output of cataract research during 2002-2011 and found that the world publication output of cataract research consisted of 27,053 papers during 2002-2011, which increased from 2,025 papers in 2002 to 3,080 papers in 2011,

witnessing an annual average growth rate of 4.89%. The average citation impact per paper registered by world publications was 6.94 during 2002-11, which decreased from 7.82 during 2002-06 to 5.21 during 2007-11. G. Cantos-Mateos et al. (2012) have studied an overview of stem cell research by Spanish and found the main areas of research. Bibliometric indicators are used for basic nature as well as techniques for the visualization and analysis of networks of scientific information. Gupta, B. M. *et al* (2011) have analysed the research output of typhoid research in India during 2000-2009. They have identified its growth, rank and global publications share, citation impact, share of international collaborative papers, contribution of major collaborative partner countries, and the contribution of various subject fields and patterns of research communication in most productive journals. Pratap. G and Gupta B.M (2011), studied the performance of education and research institutes in India in medical and allied sciences during 1999-2008. The data were collected from the Scopus database and a new composite performance indicator. Garg, K.C, et al. (2010), revealed that most of the papers published by Indian scientists dealt with molecular genetics relating to organisms of humans, plants and animals, and academic institutions published the highest number of papers. However the value of Relative Citation Impact (RCI) for academic institutions was lesser. Molatudi, M and Pouris, A (2006), examined the status of microbiology and molecular biology and genetics research in South Africa. Bibliometric analysis of medical informatics literature has been made by Sundari B.S., et al (2004), and discussed in authorship pattern, collaboration index, degree of collaboration, collaborative coefficient and country wise production. Srivatsava, D (2004) discussed on the concept of collaboration and the methodology in research collaboration in the field of biomedical sciences in India. Almeida-Filho et al. (2003) conducted a bibliometric and content analysis of research on health inequalities produced in Latin American and Caribbean countries. The study found that, recent rapid growth in overall output. Brazil, Chile, and Mexico contributed mostly empirical research, while Ecuador and Argentina produced more conceptual studies. Arunachalam and Gunasekaran (2002a, 2002b) have mapped tuberculosis and diabetes research carried out in India and China based on the data collected from three different databases.

OBJECTIVES

The objectives of the study are:

- To measure the growth of Human DNA research literature
- To analyze pre Relative Growth Rate (RGR) and Doubling Time (DT) of DNA Research Output
- To analyze the degree of collaboration
- To analyze the document type wise distribution
- To observe most collaborating countries with the USA.
- To examine most productivity authors.

DATA METHODS

The bibliographic data on human DNA research carried out in the USA has been collected from the *Scopus* database for the publication year 1984 -2012. The keyword 'human DNA' was used in the Title and Keyword fields to collect publication data pertaining to human DNA. Further, the search results fall under the two subject areas viz. "Life Sciences" and "Health Sciences" were refined. The country 'United States' was used in the affiliation field. Finally, data set of 7,264 publications was retrieved and analyzed using MS-Excel as per objectives of the study. Moreover, the relative growth rates (RGR) and the doubling time (DT) have been studied as suggested by Mahapatra (1985). The degree of collaboration proposed has also been studied as proposed by Subramanyam (1983). We have identified the top productive authors and top institutions in USA in human DNA research. We have also identified the top collaborative countries with the USA. Impact factor values of journals have been assigned from the *Journal Citation Reports* (*JCR*) 2012.

LIMITATIONS

This study is based on the bibliographic data available in the database and restricted to the study period. Some relevant bibliographic data may be missed out.

DATA ANALYSIS

This study discusses the human DNA research carried out in the USA and brought out the following findings.

Relative Growth Rate (RGR)

The overall Relative Growth Rate is the increasing in The unit of time is calculated based on one year. The mean relative growth rate R (1-2) over a specified period of interval can be calculated from the following equation suggested by Mahapatra (1985).

Doubling Time

A direct equivalence exists between the relative growth rate and doubling time. If the number of publications/pages on a subject doubles during a given period, then the difference between the logarithms of the numbers at the beginning and at the end of the period must be the logarithms of the number 2. This difference has a value of 0.693.

the number of publications/pages per unit of time.

DISCUSSION AND ANALYSIS

The distribution of human DNA research papers published in the United States of America is shown in Figure 1. The highest number of papers (337) published in the year 1996 during the study period, which is accounted for 4.64%, followed by 314 papers in 1999 with 4.32% and 293 papers (4.03%) in 1991. On the other hand, 1984, 2011 and 1987 are the years with lower productivity with 160, 196 and 201 papers respectively published by the American scientists on human DNA research. There is a fluctuating trend has been shown in publication productivity during the study period.

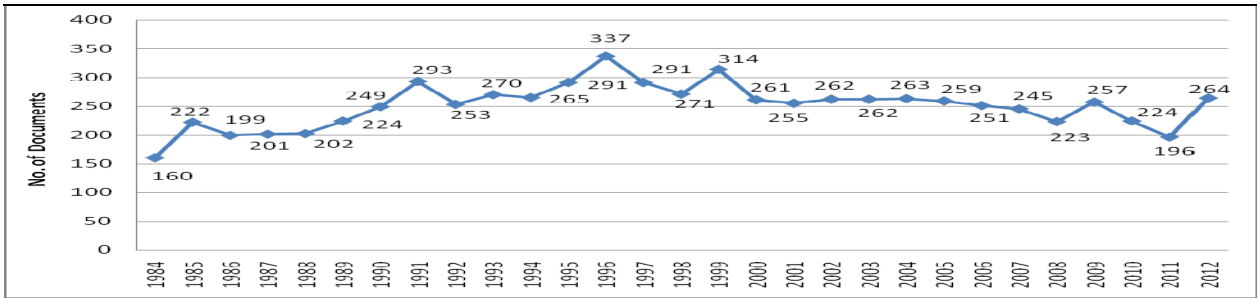


Figure 1 Distribution of human DNA research papers published in USA during 1984-2012

Moreover, there are nine document types were classified as given in the database viz. article, review, conference paper, note, letter, short survey, erratum, editorial and book chapter. The major chunk from article 6743 or 92.83% were classified under journal article. Followed by review with 208 and conference papers with 180 papers were classified. Other six document types were collectively had 133 papers or 1.83%. All document types were considered for this study

.

Table 1: Relative Growth Rate and Doubling Time of publication productivity during 1984-2012

Sl. No	Year	No. of papers	Cumulative No. of papers	W_1	W_2	R (a)	Mean (a) 1-2	Doubling Time	Mean dt (a) 1-2
1	1984	160	160		5.07				
2	1985	222	382	5.07	5.94	0.87		0.80	
3	1986	199	581	5.94	6.36	0.42		1.65	
4	1987	201	782	6.36	6.66	0.3		2.31	
5	1988	202	984	6.66	6.89	0.23		3.01	
6	1989	224	1208	6.89	7.09	0.2		3.47	
7	1990	249	1457	7.09	7.28	0.19		3.65	
8	1991	293	1750	7.28	7.46	0.18		3.85	
9	1992	253	2003	7.46	7.6	0.14		4.95	
10	1993	270	2273	7.6	7.72	0.12		5.77	
11	1994	265	2538	7.72	7.83	0.11		6.30	
12	1995	291	2829	7.83	7.94	0.11		6.30	
13	1996	337	3166	7.94	8.06	0.12		5.77	
14	1997	291	3457	8.06	8.14	0.08		8.66	

15	1998	271	3728	8.14	8.22	0.08	0.22	8.66	4.65
16	1999	314	4042	8.22	8.3	0.08		8.66	
17	2000	261	4303	8.3	8.36	0.06		11.55	
18	2001	255	4558	8.36	8.42	0.06		11.55	
19	2002	262	4820	8.42	8.48	0.06		11.55	
20	2003	262	5082	8.48	8.53	0.05		13.86	
21	2004	263	5345	8.53	8.58	0.05		13.86	
22	2005	259	5604	8.58	8.63	0.05		13.86	
23	2006	251	5855	8.63	8.67	0.04		17.33	
24	2007	245	6100	8.67	8.71	0.04		17.32	
25	2008	223	6323	8.71	8.75	0.04		17.33	
26	2009	257	6580	8.75	8.79	0.04		17.33	
27	2010	224	6804	8.79	8.82	0.03		23.10	
28	2011	196	7000	8.82	8.85	0.03		23.10	
29	2012	264	7264	8.85	8.89	0.04	0.04	17.32	15.55
	Total	7264				3.82	0.13	10.10	10.1

The Relative Growth Rate (RGR) and Doubling Time (DT) of research productivity on human DNA research papers published by US scientists are shown in table 1. It is apparent that the Relative Growth Rate trend has been fluctuating as 0.87 in 1984 and 0.04 in 2012. The mean relative growth rate (mean RGT) 0.22 has been recorded during the period 1984-1998. Similarly, the mean RGT 0.04 has been found during 1999-2012. It is found that the overall mean RGT 0.13 recorded during the study period. There has been an increasing trend on human DNA research papers in terms of doubling time rose from 0.80 in 1984 to 17.32 in 2012. From this study it is clearly found that the mean doubling time (mean DT) for publication as 24.48 years. Moreover, the mean doubling time for the period 1984-1998 is 4.65 years and 15.55 years for the period 1999 to 2012. The doubling time mean value 10.1 has been identified during the study period 1984-2012.

Degree of Collaboration

Table 2 presents the year wise distribution of a single author and multi-authored articles and their degree of collaboration. In this study, the degree of collaboration of the years (1984 -2012) is almost same of the mean value as 0.91 (i.e. 91% papers collaborative papers). Out of 7264 papers, 6631 were published by more than one author (accounted for 91.3%) and 633 papers or 8.71% are published by a single author. Overall, there is an increasing trend has been identified on multi authored papers in human DNA research by the US during the period. The highest degree of collaboration 0.99 has been recorded in the year 2003. In the years 1997, 1998 and 2004 the degree of collaboration 0.98 has been recorded. The highest number of multi-authored papers published in 1996 and 1999, with 323 and 304 respectively. On the other hand, there is a fluctuating trend of single authored articles during the period. The largest number of single authored papers 76 was published in 2010, followed by 65 papers in 2012 and 61 papers in 2009.

Table 2: Distribution of the degree of collaboration in human DNA research papers published by US scientists during 1984-2012

Year	No of papers	Single author (No. of papers)	Multiple authors (No. of papers)	Degree of collaboration
1984	160	11	149	0.93
1985	222	15	207	0.93
1986	199	8	191	0.96
1987	201	6	195	0.97
1988	202	16	186	0.92
1989	224	14	210	0.94
1990	249	11	238	0.96
1991	293	12	281	0.96

1992	253	8	245	0.97
1993	270	12	258	0.96
1994	265	8	257	0.97
1995	291	13	278	0.96
1996	337	14	323	0.96
1997	291	6	285	0.98
1998	271	5	266	0.98
1999	314	10	304	0.97
2000	261	8	253	0.97
2001	255	9	246	0.96
2002	262	7	255	0.97
2003	262	2	260	0.99
2004	263	5	258	0.98
2005	259	7	252	0.97
2006	251	56	195	0.78
2007	245	49	196	0.80
2008	223	59	164	0.74
2009	257	61	196	0.76
2010	224	76	148	0.66
2011	196	60	136	0.69
2012	264	65	199	0.75
Total	7264	633	6631	0.91

Table 3: Distribution of human DNA research papers published from USA by journals*

No.	Journal	Journal country	ISSN	EISSN	Impact factor 2012	No. of papers	Percentage to total 7264
1	Journal of Biological Chemistry	US	0021-9258	1083-351X	4.651	529	7.28
2	Cancer Research	US	0008-5472	1538-7445	8.65	356	4.90
3	Proceedings of the National Academy of Sciences of the USA	US	0027-8424	1091-6490	9.737	356	4.90
4	Nucleic Acids Research	UK	0305-1048	1362-4962	8.278	307	4.23
5	Journal of Virology	US	0022-538X	1098-5514	5.076	306	4.21
6	Biochemistry	US	0006-2960		3.377	243	3.35
7	Molecular and Cellular Biology	US	0270-7306	1098-5549	5.372	165	2.27
8	Carcinogenesis	US	0143-3334	1460-2180	5.635	148	2.04
9	Genomics	US	0888-7543		3.01	115	1.58
10	Journal of Clinical Microbiology	US	0095-1137	1098-660X	4.068	87	1.20

* top ten journals are listed.

The human DNA research papers of the USA were published in four document sources as indexed in the database viz. journals, book series, conference proceedings and books. The major papers were published in journals with 7210 (99.26 %). The other three source types, viz. Book series have 45 (0.62%), conference proceedings 5 (0.07%), and books with 4 (0.06%) papers in DNA research by the USA Scientists. The distribution of human DNA research papers published from USA in top journals is given in table 3. About 36% of papers were published in the top ten journals. *Journal of Biological Chemistry* ranked top in the list with 529 papers or 7.28%, followed by *Cancer Research* and *Proceedings of the National Academy of Sciences of the USA* with 356 papers (4.90%) each. *Nucleic Acids Research* has published 307 (4.23%) papers and *Journal of Virology* has published 306 (4.21%) papers. On the other hand, *Journal of Clinical Microbiology* finds tenth place in the list with 87 (1.20%) papers during 1984 to 2012.

Table 4: Distribution of human DNA research papers from USA by institutions*

Sl. No.	Institution	No. of papers	Percentage of 7264
1	National Cancer Institute	322	4.43
2	Harvard Medical School	150	2.06
3	UT Medical Branch at Galveston	141	1.94
4	University of Washington Seattle	140	1.93
5	National Institute of Environmental Health Sciences	139	1.91
6	University of Texas M. D. Anderson Cancer Center	136	1.87
7	The University of North Carolina at Chapel Hill	133	1.83
8	University of California, San Francisco	127	1.75
9	Baylor College of Medicine	116	1.60
10	VA Medical Center	95	1.31

*Only top ten institutions are given in this table

It is found that 258 institutions involved in human DNA research in USA during the period. Distribution of human DNA research papers from USA by institutions are shown in table 4. The top ten institutions have collectively published 1499 papers, which is accounted for 20% of the total publications. National Cancer Institute of NIH, Bethesda occupies the top of the list with 322 papers or 4.43% in DNA research. The other eight institutions have published over 100 papers each during the period. VA Medical Center has published 95 papers with 1.31% in Human DAN research during the period.

Table 5: Countries collaborating with USA in human DAN research

Sl. No.	Country	No. of papers	Percentage to 7264
1	UK	270	3.72
2	Japan	202	2.78
3	Germany	196	2.70
4	France	159	2.19
5	Canada	158	2.18
6	Italy	124	1.71
7	China	95	1.31
8	Spain	74	1.02
9	South Korea	66	0.91
10	The Netherlands	53	0.73

*Only top ten countries are listed

Out of 7264 papers, 2111 have published with international collaboration from 72 countries which is accounted for 29%. Top ten countries collaborating with USA in human DNA research during 1984-2012 are shown in table 5. The US scientists have collaborated in 270 papers (3.72%) with scientists in the United Kingdom, followed by Japan with 202 (2.78%) papers. Apart from the UK and Japan, the other four countries viz. Germany, France, Canada and Italy have collaborated in 196, 159, 158 and 124 papers respectively. China, Spain, South Korea and the Netherlands have collaborated in 95, 74, 66 and 53 papers respectively during the period. The recent National Science Foundation (NSF) report indicates that the percentage of internationally collaborative articles in all of science rose from 16% to 25% between 1997 and 2012 in the USA. [NSF report on S&E Indicators 2014] The report also reveals that the relatively high rates of international collaboration have been recorded for Geosciences, computer sciences, mathematics, physics, and biological sciences in the range of 27%-37% during 1997-2012.

Table 6: The authors contributing to human DNA research in the USA

Sl. No.	Name	No. of papers	Percentage to 7264
1	Prakash, S.	52	0.72
2	Prakash, L.	50	0.69
3	Pegg, A.E.	47	0.65
4	Kunkel, T.A.	44	0.61
5	Mitra, S.	43	0.59
6	Bohr, V.A.	41	0.56
7	Wilson, S.H.	36	0.50
8	Pommier, Y.	35	0.48
9	Copeland, W.C.	35	0.48
10	Brent, T.P.	35	0.48

*Only top ten authors are given

1242 authors have contributed to human DNA research in USA during the period. The top ten prolific authors in human DNA research in the USA are shown in table 6. Prakash, S. and Prakash, L both from the University of Texas Med Branch, Galveston, have published more than 50 papers during the period. Pegg, A.E., Kunkel, T.A., Mitra, S. and Bohr, V.A., have published 47; 44; 43 and 41 papers respectively during the period.

CONCLUSION

Research is evaluated by so many criteria and tools. Bibliometric analysis is one among the tools to measure scientific productivity in terms of quantity as

well as quality. This bibliometric study of human DNA research done in the USA reveals that there has been an upward and downward trend shown in research publications during the study period. The mean value of 10.10 Doubling Time (DT) and the value of 3.82 Relative Growth Rate (RGR) are recorded during the period. Overall, there has been an increasing trend identified on multi authored papers in human DNA research by the US during the period. The recent NSF report indicates that the percentage of internationally collaborative articles in all of science rose from 16% to 25% between 1997 and 2012 in the USA. [National Science Board, 2014] The report also reveals that the relatively high rates of international collaboration have been recorded for Geosciences, computer sciences, mathematics, physics, and biological sciences in the range of 27%-37% during 1997-2012. But this study of human DNA research done in the USA, states that 29% papers are internationally collaborating. There is a huge need for international collaboration in this field to combat threatening diseases at the global level. This study will be a useful input for scientists and researchers work on human DNA research. It also needs to look at the global level analysis to compare different countries.

Acknowledgement
The authors wish to thank Mr. Subbiah Gunasekaran, CSIR-Central Electrochemical Research Institute, Karaikudi for his valuable suggestions.

REFERENCES

1. Almeida F, Kawachi I, Filho A.P. and Dachs, J.N.W. 2003. Research on Health Inequalities in Latin America and the Caribbean: Bibliometric Analysis (1971–2000) and Descriptive Content Analysis (1971– 1995), *American Journal of Public Health*, Vol. 93,no.12:2037-2043.

2. Arunachlam , S. and Gunasekaran, S. 2002a. Tuberculosis research in India and China: From bibliometrics to research policy, *Current Science*, Vol.82,no.8: 933-947.

3. Arunachlam ,S. and Gunasekaran, S. 2002b,. Diabetes research in India and China today : From literature - based mapping health – care policy . *Current Science*, Vol.82no.9: 1086-1097.

4. Baskaran, C. 2013. Research Productivity of Alagappa University during 1999-2011: A Bibliometric Study. *DESIDOC Journal of Library & Information Technology*, Vol. 33no.3):236-242.

5. Cantos, G.M., Bvargas, Q. Chinchilla, Z. R. and Zulueta, M.A. 2012. Stem cell research: bibliometric analysis of main research areas through KeyWords Plus. *Aslib Proceedings*, Vol. 64,no.6: 561 – 590.

6. Cardos. G. Rodewal. A, Andrei, and D. Soficaru. 2008. Ancient DNA study on human fossil found Costisa, Romania, dating from de Bronze Age. *Bucarest*, 55–67.

7. Garg, K.C, et al. (2010), Scientometric Profile of genetics and heredity research in India, *Annals of Library and Information Studies*, Vol. 57, PP. 196-206.

8. Gupta, B.M.; Bala, Adarsh; Baidwan, Kiran; Chadhha, Neelima; and Cheema, Harjit Singh. (2011), Mapping of typhoid research in India: A scientometric analysis of publications output in 2000-2009. *Chinese Librarianship: an International Electronic Journal*, 31. URL: http://www.iclc.us/cliej/cl31GBBCC.pdf

9. Gupta, B.M; Bala, Adarsh; and Kshitig, Avinash, "World Cataract Research: A Scientometric Analysis of Publications Output during 2002-11" (2013). *Library Philosophy and Practice (e-journal)*. Paper 895. http://digitalcommons.unl.edu/libphilprac/895

10. Lander E.S et al., *(2001)*, "Initial sequencing and analysis of the human Genome", *Nature*, Vol. 409 (2001)

11. Mahapatra, M. (1985). On the Validity of the Theory of Exponential Growth of Scientific Literature, *Proceedings of the 15th IASLIC Conference*, Bangalore, Pp.61-70.

12. Molatudi,M and Pouris, A (2006), Assessing the Knowledge base for biotechnology in south Africa: A Bibliometric analysis of South Africa Microbiology, Molecular biology and Geneties research, *Scientometrics*, Vol. 68, (1), PP. 97-108.