Winter 12-31-2020

BIBLIOMETRIC ANALYSIS OF CHEMISTRY NOBEL LAUREATE GEORGE PEARSON SMITH

Abhay Maurya
Mizoram University, abhaymaurya17@hotmail.com

BIBLIOMETRIC ANALYSIS OF CHEMISTRY NOBEL LARAUTE GEORGE PEARSON SMITH

Abhay Maurya

Department of Library and Information Science, Mizoram University

ABSTRACTS

The scientometric analysis of Nobel Laureate George P. Smith, who had been awarded the Nobel Prize in chemistry in 2018, reveals that the author has published 55 papers in various fields between 1971 and 2019. The subjects included 26 articles in Biomolecular Interactions, 17 papers in Genomics, 10 in Protein, and 2 in Evolutionary Biology. The forms of publications included 45 article papers, 1 book chapter, 2 editorials, 6 reviews, and 1 short survey. 17 of his publications were single-authored, 18 two authored, 9 three and four authored, and 1 five and six authored. The highest no papers were published during the period 1991 - 2000. He has published the highest no papers (6) in BIOTECHNIQUES.

KEYWORDS: Bibliometrics, Bio-bibliometrics; History of science; Scientometric portrait; Scientometrics; Science of science; Sociology of science; Nobel Prize, Nobel Laureate, George Smith, Scopus

1 INTRODUCTION

George Pearson Smith, an American biologist, and Nobel laureate was born on March 10, 1941, in Norwalk, Connecticut, USA. He had earned his A.B. degree from Haverford College in biology, whereafter he worked for one year as a high school teacher and lab technician before earning his Ph.D. degree in bacteriology and immunology from Harvard University. He held a postdoc from the University of Wisconsin with Nobel laureate Oliver Smithies before he had moved to Columbia, Missouri, and joined the University of Missouri as a faculty in 1975. The academic period between 1983–1984 was spent at Duke University with Robert Webster where he began the work that led to him being awarded the Nobel Prize. He is best known for phage display, a technique where a specific protein sequence is artificially inserted into the coat protein gene of a bacteriophage, causing the protein to be expressed on the outside of the bacteriophage. Smith first described the technique in 1985 when he displayed peptides on filamentous phage by fusing the peptide of interest onto gene III of filamentous phage. His profile includes serving the University of Missouri as Curators' Professor in 2000, being elected Fellow of the American Association for the Advancement of Science (AAAS) in 2001, and being conferred the American Society for Microbiology Promega Biotechnology Research

Award in 2007. He won the Nobel Prize in chemistry along with Greg Winter and Frances Arnold in 2018.

2 LITERATURE REVIEW

There are several published scientometric studies on Nobel Laureates and others. **Maurya, A. (2020)** focuses on the scholarly contribution of the Massachusetts Institute of Technology (MIT) to the Nobel Laureates of Chemistry and found that Nobel laureates from MIT are highly effective in their fields. In the course of their study, **Korytkowskia, P. and Kulczycki, E. (2019)** concluded that the publication counting method for national evaluation purposes needs to take into account the current situation in the given country in terms of the excellence of research outcomes, level of internal, external and international collaboration, and publication patterns in the various fields of science. **Kokol, P. (2019)** found that both research literature and funding are witnessing an upward trend in the area of Software Engineering. The results of the study conducted by **Li, K. & Yan, E. (2019)** suggested that the impact of funding size on the keyboard matching ratio is highly moderated by the funding group. **Nazarovets, S. (2018)** focused on highly cited papers of Ukrainian scientists written in collaboration and stated that the number of highly cited documents of Ukrainian scientists forms the top 1%, 5%, of the documents cited and the growth of 10% documents had been steady though these differed significantly in different scientific fields. **Bornmann, Lutz et. al**. suggest that the composition of HCRs was based on one indicator (citation impact) but research excellence cannot generally be determined using one indicator alone but should be measured using multiple dimensions. **Guler A.T. et. al.** stated that the use of scientific workflows in bibliometrics is still in its infancy. The direct support of R inside Taverna workflows is particularly useful for bibliometrics and scientometrics.

3 OBJECTIVES

The purpose of this study is to document quantitatively the article publication productivity pattern of George P. Smith. In particular, the study is an effort:

1. To analyze the domain-wise preferred medium of scientific communication
2. To analyze the domain-wise authorship patterns;
3. To find out the channels of communication;
4. To find Authors' Production over Time
5. To find out the citation network

4 METHODOLOGY

The study is based on the list of papers indexed in the Scopus database. The data obtained were analyzed on various parameters: documents authored by George Smith, the number of scientific communications, domain-wise scientific communication, domain-wise authorship pattern, channels of communication, the authorship credits of researchers collaborated with other Nobel Laureates, and citation network. The data has been downloaded from the Scopus database at 10:05 hours (IST) on 24.03.2020 using the keywords "George Smith". The modus operandii included a visit to the Scopus database (https://www.scopus.com/search/form.uri?display=authorLookup) and selecting the author radio button where the surname "Smith" and author's first name "George P" has been entered before pressing 'Search'. Twenty-two results had been obtained and after verification of the affiliation, I had selected one thereby downloading 55 data files in three formats: Text, CSV, and Bibtex. I have used different tools for analysis which included Excel and R programming. For visualization, I have used VOSviewer.

5 ANALYSIS

5.1 Domain-wise preferred mode of scientific communication

Table 1 details the number of scientific communication of George P Smith. An analysis of Table 1 indicates that with 45 articles, forming 81.81% of the total scientific communications, the highest contribution had been in the form of articles followed by 6 reviews forming 10.90%. 2 editorials (3.63%), 1 each from book chapter and short survey (1.81%) make up for the remaining scientific communication. Domain wise, the highest number of contributions is in the field of Biomolecular Interactions 26 (47.27%), followed by Genomics 17 (30.90%), Protein 10 (18.18%), and Evolutionary biology 2 (3.63%).

Table 1: Number of scientific communications

SN	DOMAIN WISE CONTRIBUTION	DOMAIN				TOTAL PAPER	%
		A	B	C	D		
01	**ARTICLE**	19	2	16	8	45	**81.81**
02	**BOOK CHAPTER**	1				1	**1.81**

03	**EDITORIAL**				2	2	**3.63**
05	**REVIEW**	5		1		6	**10.90**
06	**SHORT SURVEY**	1				1	**1.81**
TOTAL		26	2	17	10	55	
%		**47.27**	**3.63**	**30.90**	**18.18**		

A = Biomolecular Interactions, B = Evolutionary Biology, C = Genomics, D = Protein.

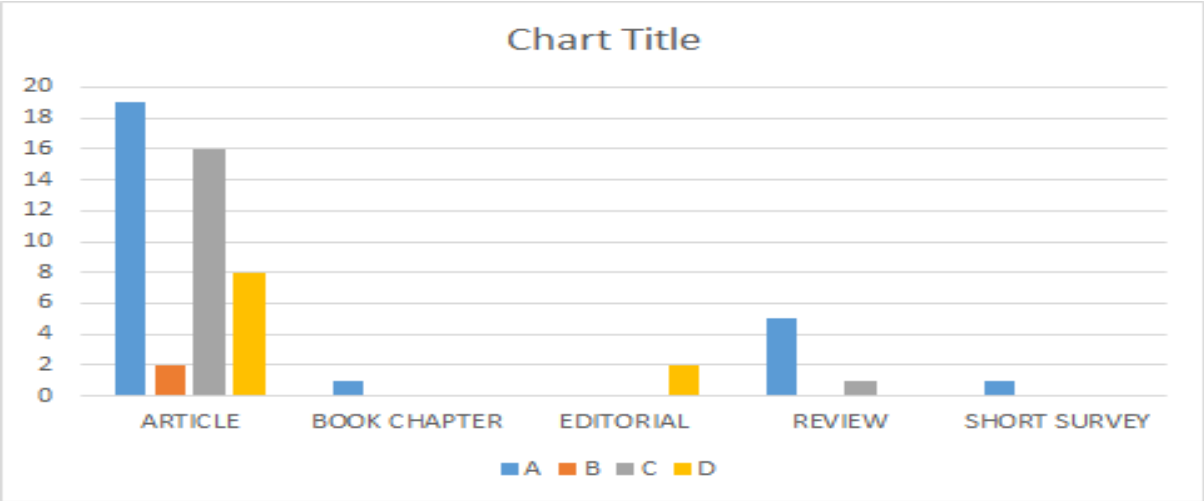

Fig 1: Number of scientific communication

5.2 Domain-wise authorship patterns of George P. Smith

5.2.1 According to collaboration wise

Table 2: Domain wise authorship according to collaboration

SN	AUTHORSHIP	DOMAIN				TOTAL PAPER	%	TOTAL AUTHORSHIP	%
		A	B	C	D				
01	1 Authored	3	1	6	7	17	30.90	17	13.38
02	2 Authored	8	1	7	2	18	32.72	36	28.34
03	3 Authored	5		3	1	9	16.36	27	21.25
04	4 Authored	8		1		9	16.36	36	28.34
05	5 Authored			1		1	1.81	5	3.93
06	6 Authored	1				1	1.81	6	4.72
TOTAL		25	2	18	10	55		127	

A = Biomolecular Interactions, B = Evolutionary Biology, C = Genomics, D = Protein.

An analysis of Table 2 reveals that 18 papers, contributing 32.72% of the total works, are two-authored, followed by 17 single-authored papers (30.90%). There are 9 three and four authored papers forming 16.36% of the total publications. Domain wise, the highest number of papers have been published in the field of biomolecular interaction (25papers, 45.45%), followed by genomics (10 papers, 32.72%), protein (10 papers, 18.18%), and evolutionary biology (2papers, 3.63%).

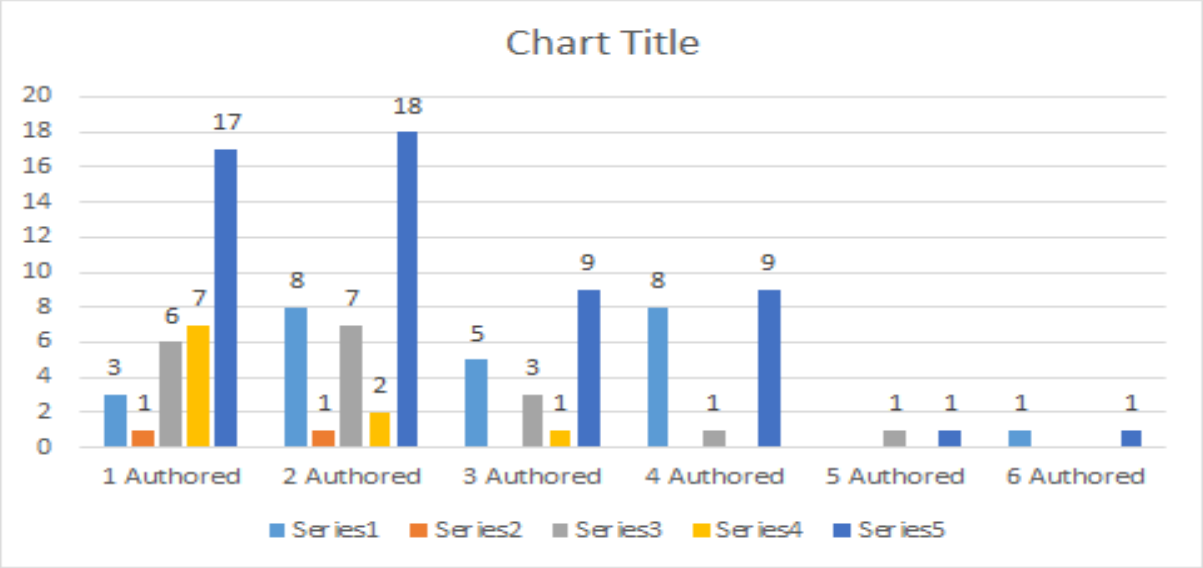

Fig 2: Domain wise authorship

5.2.2 5. Year wise contributions in various Domains

Table 3: Domain wise authorship according to year wise

SN	AUTHORSHIP	DOMAIN				TOTAL PAPER	%
		A	B	C	D		
01	1971 - 1980			10	2	12	21.81
02	1981 - 1990		1	7	2	10	18.18
03	1991 - 2000	11	1		4	16	29.09
04	2001 - 2010	10				10	18.18
05	2011 - 2020	5			2	7	12.72

TOTAL	26	2	17	10	55	

A = Biomolecular Interactions, B = Evolutionary Biology, C = Genomics, D = Protein.

As regards the year-wise authorship pattern which has been incorporated in Table 3, it is observed that 16 papers (29.09%) had been published from 1991 to 2000. This is followed by the publication of 12 papers (21.81%) from 1971 to 1980. The years from 1981 to 1990 and 2001 to 2010 witnessed the publication of 10 papers (18.18%) while 7 papers (12.72%) were published from 2011-2020**.**

5.3 Preferred channels of communication used by George P. Smith

As observed from Fig.3. George Smith published the maximum papers in peer-reviewed journals. Among his published papers include 6 papers in BIOTECHNIQUES, 5 papers in GENE, 4 papers in Science and VIROLOGY, 3 papers in PROTEIN ENGINEERING. More than 50 percent of the publications have been published in journals having good impact factors.

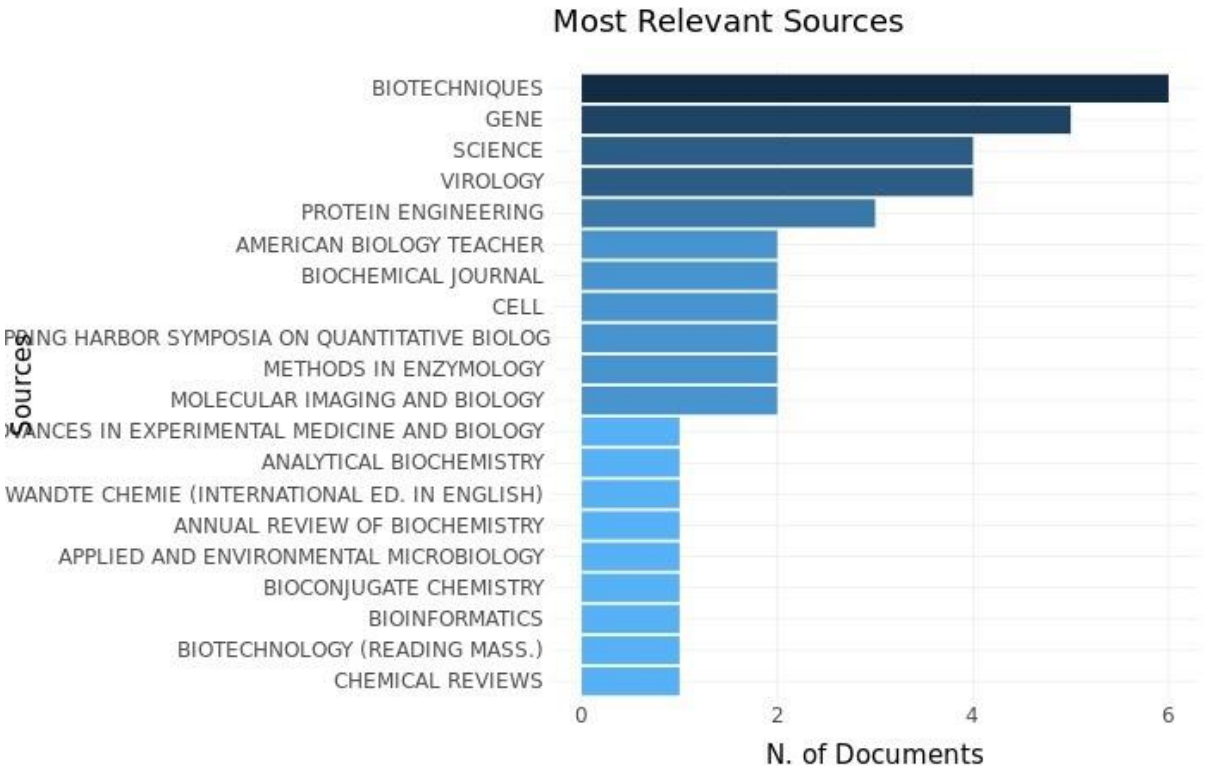

Fig 3: Channels of communication used by George Smith

5.4 To find Authors' Production over Time

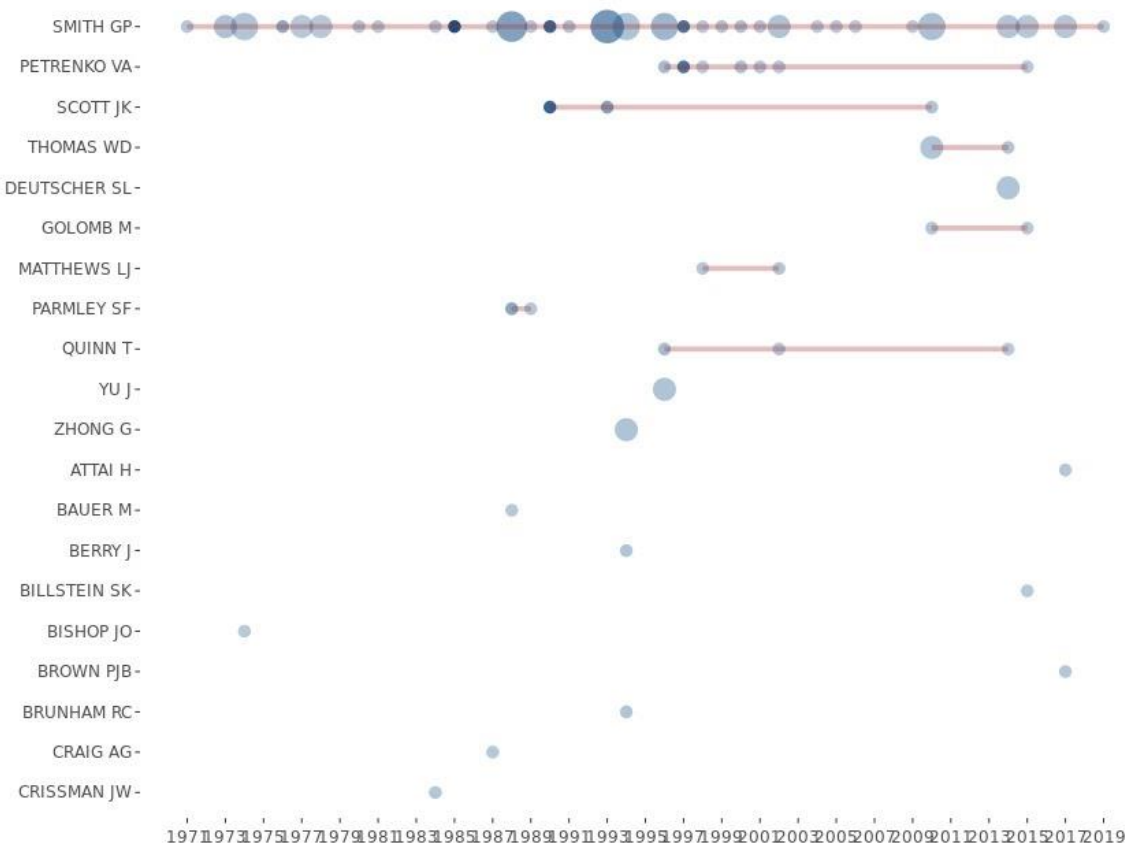

Fig 4: Authors' Production over Time

The year 1971 saw George Smith publish his first paper with a total citation of 1.57 per annum. In the year 1973, two papers were published with a total citation of 4.4 per annum. Three papers were published in 1974 with a per annum citation of 2.45 followed by one article in 1976 having a citation of 18.42 per annum. In 1977, George Smith published two articles with a citation of 0.75 per annum, and so on.

5.5 To find out the citation network of Nobel Laureates

For citation networks, I have analyzed co-authorship, citation, bibliographic coupling, and co-occurrence networks.

5.5.1 Co-authorship network

Fig 5 shows the authorship network between G. P. Smith and other authors. Smith has the highest number of co-authorship collaboration with V. A. Petrenkvo.

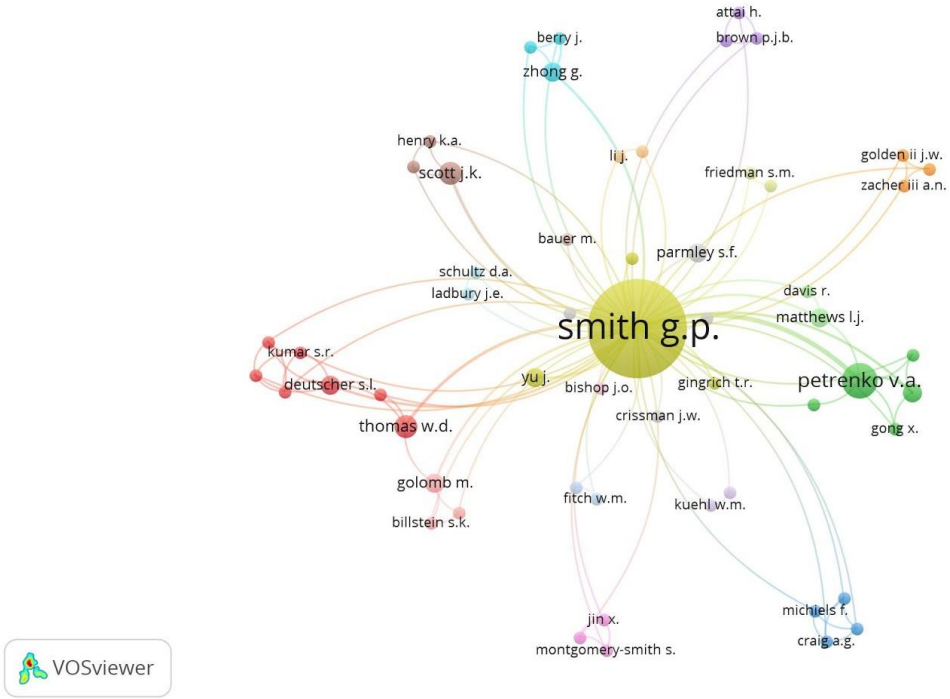

Fig 5: Co-author Network

5.5.2 Citation analysis

The citation analysis of George Smith indicates that "Filamentous fusion phage: Novel expression vectors that display cloned antigens on the virion surface" published in the year 1985 has the highest number of citations which is followed by "Phage display" published during the year 1997 and "Searching for peptide ligands with an epitope library" published in 1990.

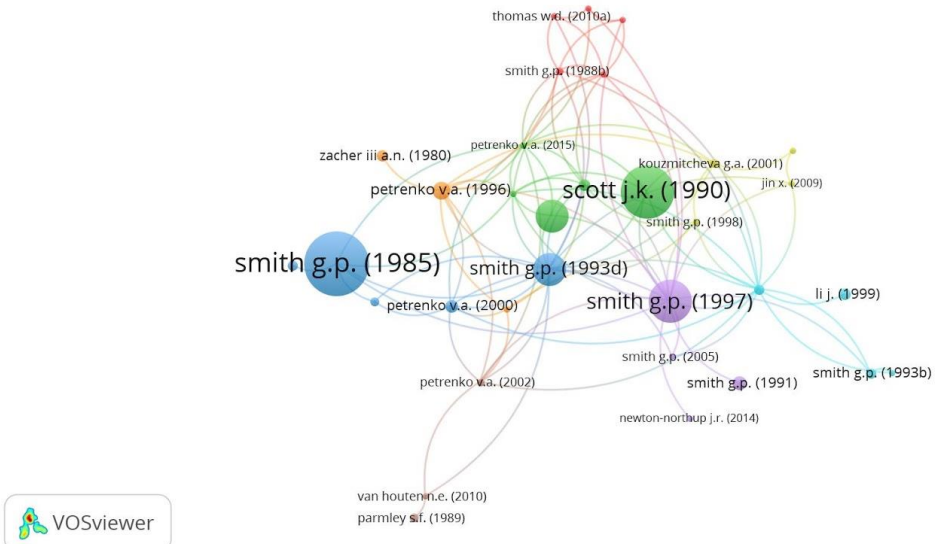

Fig 6: Citation Network

5.5.3 Bibliography Coupling

Bibliographic coupling is done when a reference is used by two papers äs a unity of coupling between those two papers. Fig 7, shows the bibliography coupling of George P Smith.

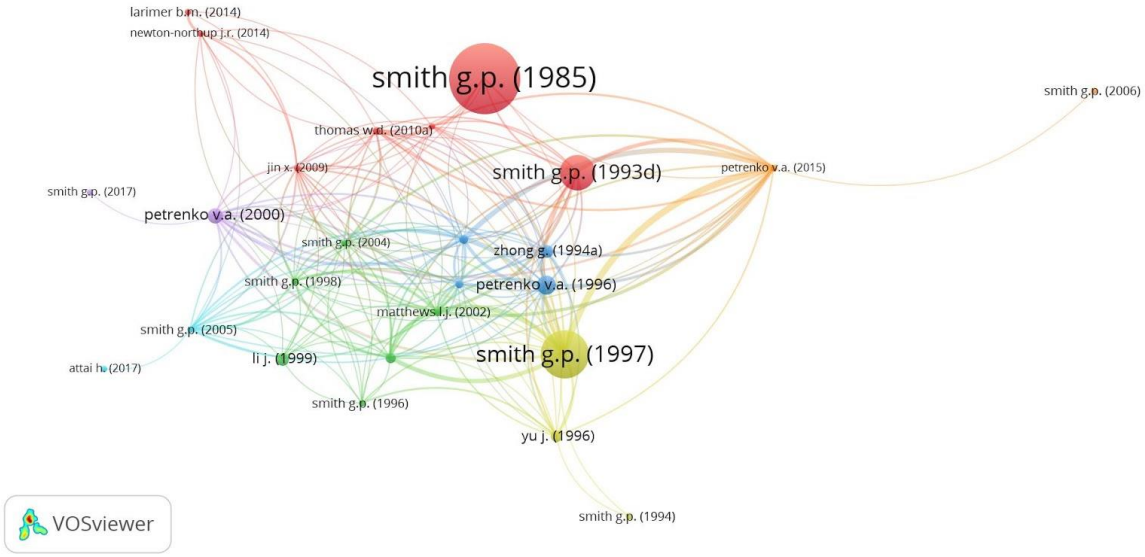

Fig 7: Bibliography Coupling

5.5.4 Co-occurrence

Fig 8 shows the co-occurrence networks of G. P. Smith.

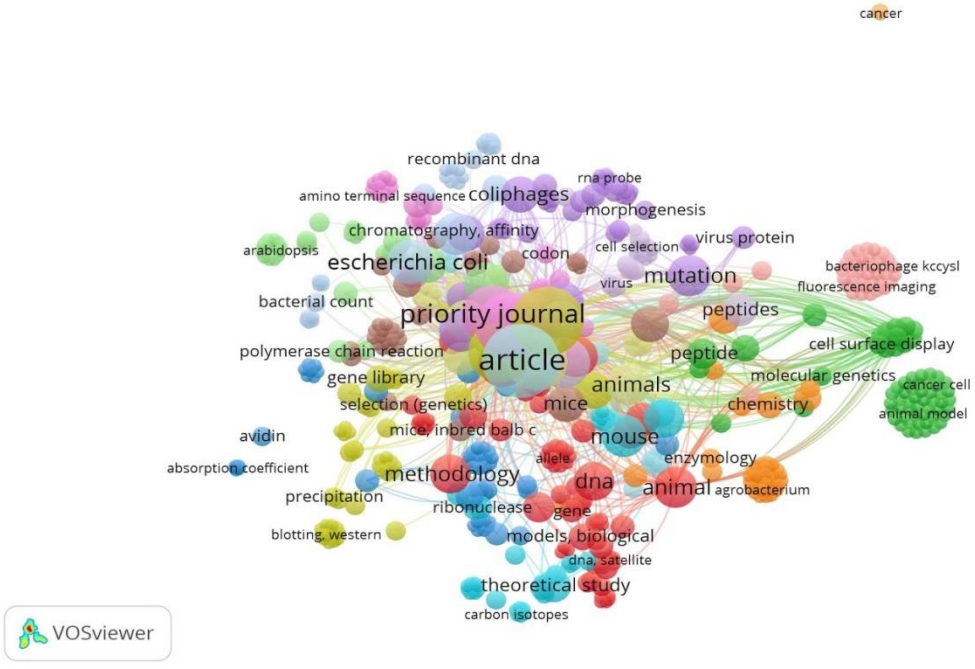

Fig 8: Co-occurrence

6 ANOTHER OUTCOME

1.	Average citations per documents	175.7
2.	Documents per Author	0.982
3.	Authors per Document	1.02
4.	Co-Authors per Documents	2.31
5.	Collaboration Index	1.45
6.	Annual Scientific Production	31.617
7.	Article fractionalized	31.617
8.	H index	29
9.	G index	55
10.	M index	0.580
11.	Total citation	9664

7 CONCLUSIONS

Publication productivity of Nobel laureate George Pearson Smith was found to be consistent throughout the period of his scientific career under study. His papers were published in 32 journals. He received innumerable awards and honors including the Nobel prize in 2018 at the age of 77 years. This pattern suggests that honours and awards are directly related to the number of collaborators resulting in an increase in publication productivity.

8 REFERENCE

Maurya, A. (2020). Scientometric Study of Nobel Laureate of Chemistry Department, Massachusetts Institute of Technology (MIT). Library Philosophy & Practice.

Kokol, P. (2019). Funded and non-funded research literature in software engineering in relation to country determinants. COLLNET Journal of Scientometrics and Information Management, 13(1), 103–109. https://doi.org/10.1080/09737766.2018.1560637

Li, K., & Yan, E. (2019). Are NIH- funded publications fulfilling proposed research? An examination of concept- matchedness between NIH research grants and their supported publication. Journal of Informetrics, 13, 226–237.

Korytkowski, P., & Kulczycki, E. (2019). Publication counting methods for a national research evaluation exercise. Journal of Informetrics, 13(3), 804–816. https://doi.org/10.1016/j.joi.2019.07.001

Nazarovets, S. (2018). Highly cited papers of Ukrainian scientists written in collaboration: A bibliometric analysis (2011-2015). COLLNET Journal of Scientometrics and Information Management, 12(1), 35–47. https://doi.org/10.1080/09737766.2017.1332613

Bornmann, L., Bauer, J., & Schlagberger, E. M. (2018). Highly Cited Researchers 2014 and 2015: An investigation of some of the world's most influential scientific minds on the institutional and country level. COLLNET Journal of Scientometrics and Information Management, 12(1), 15–33. https://doi.org/10.1080/09737766.2017.1332610

Guler, A. T., Waaijer, C. J. F., & Palmblad, M. (2016). Scientific workflows for bibliometrics. Scientometrics, 107(2), 385–398. https://doi.org/10.1007/s11192-016-1885-6

Of, M., & Communication, M. (1994). MODEL OF MANIFESTED COMMUNICATION THROUGH PUBLICATIONS. Scientometrics, 31(3), 223–239.