

Scientometric Analysis of *Nature*

Ming-yueh Tsay

mytsay@nccu.edu.tw

Graduate Institute of Library, Information and Archival Studies, National Chengchi University, Taipei, 116, Taiwan

Introduction

Bibliometric techniques using references made to other documents can be used to map relationships between various channels of scholarly communications. From the studies of citation analysis, one can learn which scholars from which disciplines cite which articles? Which authors are cited more often? Which disciplines cite the journals of other disciplines? The purpose of this paper is to propose a multiple set of indicators for making a bibliometric portrait of a single journal.

Research Methods and Questions

For this study, a scientometric analysis is used to investigate the characteristics of the literature in the *Nature* for the period of 1999-2008. By employing the Excel, 27,581 bibliographic records retrieved from WOS database are processed and analyzed.

The research questions include: what numbers and document type of literature that *Nature* published? What subject area and key word are? And what are the citedness of literature, authors, and the reprint author, etc.?

Growth, Document Type and Reference of the Literature

The number of literature in the *Nature* increased during 1999 and 2000, and reached a peak as 3,021 in year 2000. However, after year 2000, the number of literature gradually decreased to 2,590 in 2003. Then it raised to 2,806 in 2005, and kept steady around 2,700 in the late three years. The average paper per issue (51 issues yearly) published is from 50.8 to

59.2. The growth pattern is typical linear distribution. Figure 1 demonstrates the document types in *Nature* and the main one is Article, followed by Editorial Materials, News Items, Letters, Book Reviews, and others. The Review paper has the average number of 83 references in *Nature*, and ranked first. Article comes next with 26 references. Average number of references in Editorial Material was 6, and less than 3 in other document types.

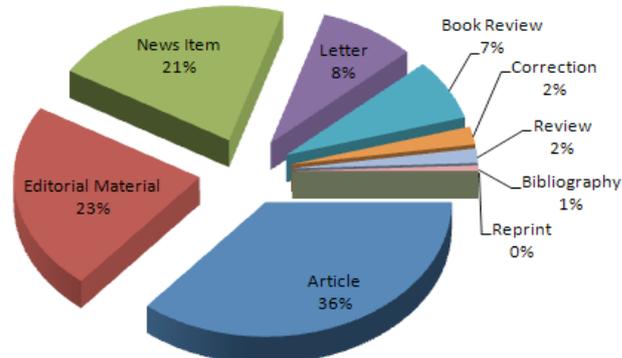


Fig.1. Document type of literature in *Nature*, 1999-2008

Subject Area and Key Words of the Literature

Biochemistry & Molecular Biology is the main subject area in *Nature*, ranking first in this decade. Pharmacology & Pharmacy and Genetics & Heredity ranked among second and third among 1999 and 2003. After 2004, Genetics & Heredity ranked second, and Cell Biology ranked third. In the top ten subject areas, Behavioral Sciences ranked among sixth and seventh, Immunology among fourth and fifth between 2001 and 2003, and Neurosciences & Neurology ranked fourth in 2002. In the late four years, Zoology, Environmental Sciences & Ecology, and Microbiology become more popular,

ranking among fourth and sixth. There are 82,713 key words in *Nature*, and 52 of them appeared more than 100 times. “Expression” was used 775 times, and ranked first. Key words ranking from second to fourth were used more than 500 times, which were “Evolution” (550 times), “Protein” (536 times), and “Gene” (516 times). Key words ranking among fifth and eighth appeared more than 300 times, including “Activation”, “Model”, “Identification”, and “Cells”. Besides, “Mice” ranked ninth and “Dynamics” ranked tenth.

Count of Literature Citedness in Nature

Table 1 shows the number of citedness of 27,581 papers in *Nature* from 1999 to 2008.

Table 1 Number of citedness of literature, by document types, in *Nature*

Highly Cited Paper

There are 34 papers cited over 1,500 times in *Nature*, 24 of which were Articles, and 10 of which were Reviews. The most highly cited paper, cited 6,429 times, during the study period is a review paper entitled “Initial sequencing and analysis of the human genome,” authored by Lander, ES and other 243 authors from England, USA, Japan, France, German, Hong Kong (Peoples R China), Ireland, and Israel published in 2001. This paper “report the results of an international collaboration to produce and make freely available a draft sequence of the human genome “ and “present an initial analysis of the data, describing some of the insights that can be gleaned from the sequence” The correspondent authors are ES Lander, RH Waterston, JE Sulston and FS Collin. John Edward Suston was a Nobel laureate on physiology or medicine in 2002. The most cited Article papers entitled “Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells,” authored by Elbashir, SM and other 6 authors, whose country of institution was German, was cited 3,712 times.

In the countries of authors' institution of top 24 highly cited Article papers, USA accounted for 58%, with a total of 14 papers. As for Japan, German, England, and French, they just had 3-4 papers. About the number of authors, the top 24 highly cited Articles were all written by more than 2 authors, and at most 45 persons. On the other hands, these articles were mainly published during 1999 and 2002. In fact, the top 3 highly cited papers were published in 2001, 2000, and 2001 respectively.

Author Number of Literature in Nature

Papers written by a single author are mainly Biographical-Item, Book Review, Correction, Editorial Material, Letter, News Item, and Reprint. The number of authors of Articles is mostly 2-4 persons (47%), following by 6-10 authors (31%).

Cited Times Document Type	3,000↑	2,501~ 3,000	2,001~ 2,500	1,501~ 2,000	1,001~ 1,500	501~ 1,000	1~500	0
Article	2	1	8	13	41	234	9,258	368
Review	1	1	3	5	18	33	401	15
Biographical-Item							43	135
Book Review							272	1,748
Correction							254	339
Editorial Material						1	3,711	2,535
Letter							989	1,341
News Item							2,090	3,718
Reprint								3

There are 33 articles by more than 100 authors. As for Review, the number of authors is mainly two persons (38%). Besides, there are 159 papers of Review written by single author (33%), and 9 by more than 100 authors.

Institution of Authors and Country of Institution

USA ranked first, accounting for 31%. England ranked second (10%) and German ranked third (5%). Ranking from fourth to tenth, accounting for 1% to 3%, were France, Japan, Canada, Switzerland, Netherlands, Australia, and Italy. With a

total of 36,410 institutions, University of Harvard was the most productive institution, accounting for 2%. UC Berkeley ranked second, and University of Oxford ranked third. Ranking from fourth to tenth were University of Cambridge, Stanford University, Massachusetts Institute of Technology, California Institute of Technology, UC San Diego, University of Washington, and University College London. On the other hand, most institutions were located in USA. Out of the 59 institutions, 44 were in USA, accounting for 75%, and 6 in England, accounting for 10%. In addition, Japan still ranked first in Asia. That was University of Tokyo.

University, Taipei, Taiwan, is very much appreciated.

Author and Reprint Author in Nature

The author's role is based on the assumption that the first author or the corresponding author is considered as the one mainly responsible for the intellectual output of the research. However, whether or not the first author is the corresponding author requires further study. Out of the 27, 581 papers in *Nature*, there were 18,643 papers with reprint author. The reprint author was chiefly the first author with 74%. Among them, 49% are Articles and 83% are Reviews. Editorial Material and Letter each accounts for 99%.

Conclusions

By employing a scientometric analysis using a set of multiple indicators, a bibliometric portrait of *Nature* has been obtained. The results may help authors, editors and publishers to know the bibliometric characteristics of *Nature* in more depth. This may help a potential author to prepare an appropriate manuscript for successful consideration by the editor of this much well-known journal.

Acknowledgments

Data collected by Chi-chieh Chiang, Graduate Institute of Library, Information and Archival Studies, National Chengchi