

Mapping new biology research in India and China: An analysis of publications, citations and international collaboration

Subbiah Arunachalam* and Rino S I
M S Swaminathan Research Foundation
Third Cross Street, Taramani Institutional Area
Chennai 600 113, India

*Correspondence: arun@mssrf.res.in

Abstract

New biology research in the two most populous countries of the world, India and China, has been mapped based on publications and citation data. We have identified institutions publishing large number of papers, cities and states contributing the papers, journals used to publish the papers classified by subfield, journal country and impact factor, highly cited papers, and internationally coauthored papers. China's publication output has been growing much faster than India's in the period 1992-2001, and in fact China now publishes a much larger number than India. However, India's 1995 papers have been cited more often, on average, than the Chinese papers. The number of papers from China published in journals of impact factor greater than 3.0 has increased from 1995 to 1998 at about 13% per annum, double the growth rate of such papers from India. The percent share of such papers from India has remained the same at 8.8%, whereas it remained above 11.5% for China. Only a handful of institutions in both India and China have published consistently well in both 1995 and 1998. In both countries a substantial part of the highly cited papers were written in collaboration with foreign institutions, China again having a much higher percent of such papers than India. We have identified the important institutions in India and China that have contributed the better cited papers and the papers from the 1995 data set that have been cited at least 25 times. Contrary to general perception, academic institutions contribute the bulk of India's research papers.

Introduction

This paper is based on the premise that monitoring publications and citations using techniques of comparative bibliometrics can provide some valuable policy-relevant insights, especially if supplemented by more inclusive input-output analysis. In a recent paper Arunachalam had shown that output of scientific research publications was stagnating in India, but increasing rapidly in China, South Korea and Brazil (Arunachalam, 2002). This was picked up by *Nature* (Jayaraman, 2002) and formed the basis of some questions in the Indian parliament.

New biology – life science research done at the molecular level, as distinctive from classical biology - is the area which receives the most funds and which accounts for the lion's share of publications in the advanced countries such as the G7 countries. This term became established when *Nature*, the weekly journal, started new weekly called *Nature New Biology* in the late 1960s. According to NSF's *Science and Engineering Indicators 2002* (National Science Board, 2002), life sciences account for more than half of the publications in the G7 countries as well as in Sweden, The Netherlands, Australia, New Zealand, Switzerland, Israel and Spain. However, life sciences research accounts for less than one-fourth of the scientific research papers from both India and China. China is particularly strong in mathematics and chemistry where it accounts for about 10% of the world's literature (Arunachalam and Gunasekaran, 2002), India publishes more than 5% of the world's agricultural research, but otherwise its contribution to most fields is around 2%.

In this note we compare new biology research in India and China. In a recent note, Padmanabhan (2002) has suggested that the quality of life science research in India has substantially improved and that the frequency of papers published from major institutions in high impact factor journals have significantly increased in recent years. We will see if the data on published literature agree with this perception

Methods

Bibliographic details (author, source, document type, language, title and address of authors) of Indian and Chinese papers were downloaded from *Biochemistry and Biophysics Citation Index (BBCI)* and *Biotechnology Citation Index (BTCI)* in comma delimited format. As *Derwent BTA* has a different format, we used a Visual Foxpro program to convert the data downloaded from text format to a relational database. Also, unlike *BBCI* and *BTCI*, *Derwent BTA* gives only the first author's address and we had no way of capturing papers by Indian and Chinese authors wherever they were not the first authors. The three sets of data were standardized and merged. As many publications are indexed in more than one database, we removed duplicates and retained the unique entries. This involved both writing programs and some manual work. To show the extent of overlap in coverage we drew Venn diagrams. We added information on city, state (province), and subfield (based on ISI deluxe classification of journals), and impact factor (from *JCR* 1997) to each entry. Indian institutions were categorized into academic, government research, etc. International collaboration was studied for all papers indexed in *BBCI* and *BTCI* based on author address field. Data were analysed to reveal the distribution of papers by institution and impact factor, by institution and subfield and by institution and journal.

Results

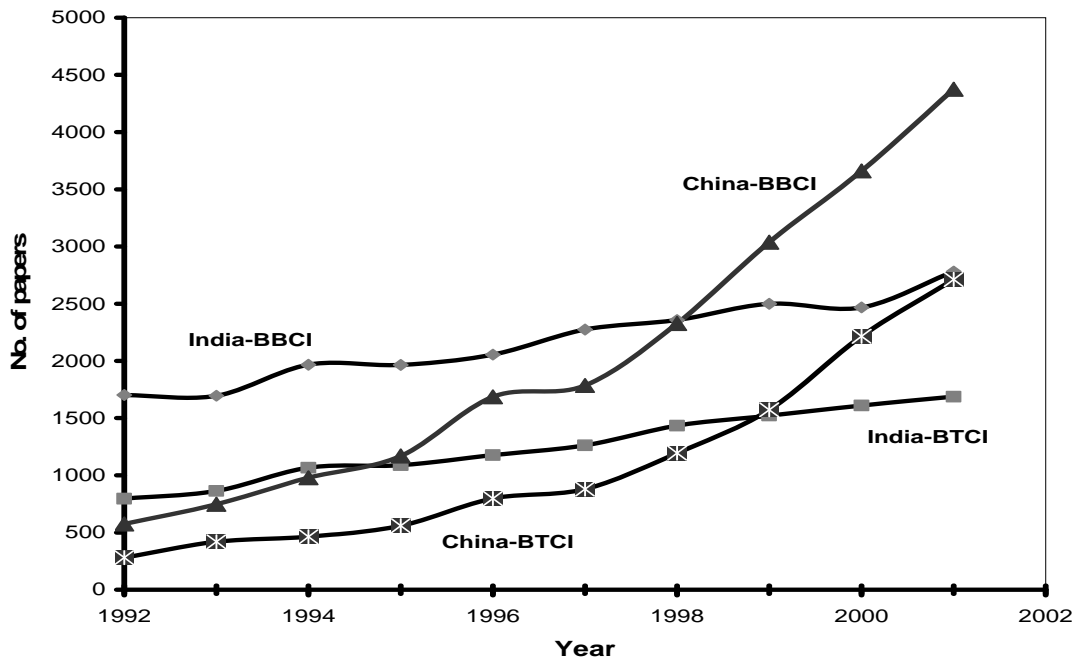
Growth of new biology research in the 1990s - As seen from the three databases, India published 2,362 papers in 1992 (counting entries indexed in more than one database only once), 2,686 papers in 1995 and 3,257 papers in 1998, thus recording a compound annual growth rate of 5.5%. China published 1,448 papers in 1995 and 2,418 papers in 1998, recording a compound annual growth rate of 18.6%. Since 1998, China has overtaken India as seen from papers indexed in *BBCI* and *BTCI*. The numbers of papers published from India and China in each year from 1992 to 2001 and indexed in *BBCI* and *BTCI* are plotted in Fig.1. New biology research output has grown steadily in India throughout the ten years as seen from both *BBCI* and *BTCI*, whereas China's output grew faster than India right from the beginning and accelerated around 1997. While the world's output as seen from the numbers of papers indexed in different disc years of *BBCI* had grown faster than India's output, China's output grew much faster than the world's output.

BBCI indexes a very large proportion of papers from both India and China. In the case of India, the overlap between *BBCI* and *BTCI* was 19% of *BBCI* papers in 1992, 36% in 1995 and 32% in 1998. About 77% percent of Indian papers in 1992 and 1995 and 86% of papers in 1998 were journal articles. Over 81% of Chinese papers in 1995 and over 88% in 1998 were journal articles.

As seen from *BBCI*, India had published 1,702 papers in 1992 (publication year), 1,965 papers in 1995, 2,357 papers in 1998 and 2,782 papers in 2001. In all, there were 102,458 papers in the 1992 annual disc of *BBCI*, 149,766 papers in the 1995 annual disc, 173,353 papers in the 1998 annual disc, and 189,001 papers in the 2001 annual disc. India's compound annual growth rate over the 9 years 1992-2001 was 5.61% as against 7.04% for the world as a whole. China had published 575 papers in 1992, 1,169 papers in 1995, 2,327 papers in 1998, and 4,375 papers in 2001, recording a compound annual growth rate of 25.29%. Between 1998 and

2001, China recorded a compound annual growth rate of 23.42% and India 5.68%, compared to 2.92% for the world as a whole.

Figure 1. New Biology research in India and China
(Source: BBCI, BTCI)



As seen from *BTCI*, India had published 796 papers in 1992, 1,089 papers in 1995, 1,434 papers in 1998 and 1,687 papers in 2001. There were in all 56,075 papers in the 1992 annual CD-ROM disc of *BTCI*, 83,507 papers in the 1995 disc, 104,029 papers in the 1998 disc, and 108,248 papers in the 2001 disc. India's compound annual growth rate during 1992-2001 was about 8.7% as against 7.58% for the world as a whole. China had published 281 papers in 1992, 560 papers in 1995, 1,195 papers in 1998 and 2,712 papers in 2001, recording a compound annual growth rate of 28.64%. India's compound growth rate between 1998 and 2001 was 5.56%, and China's 31.41%, compared to 1.33% for the world as a whole.

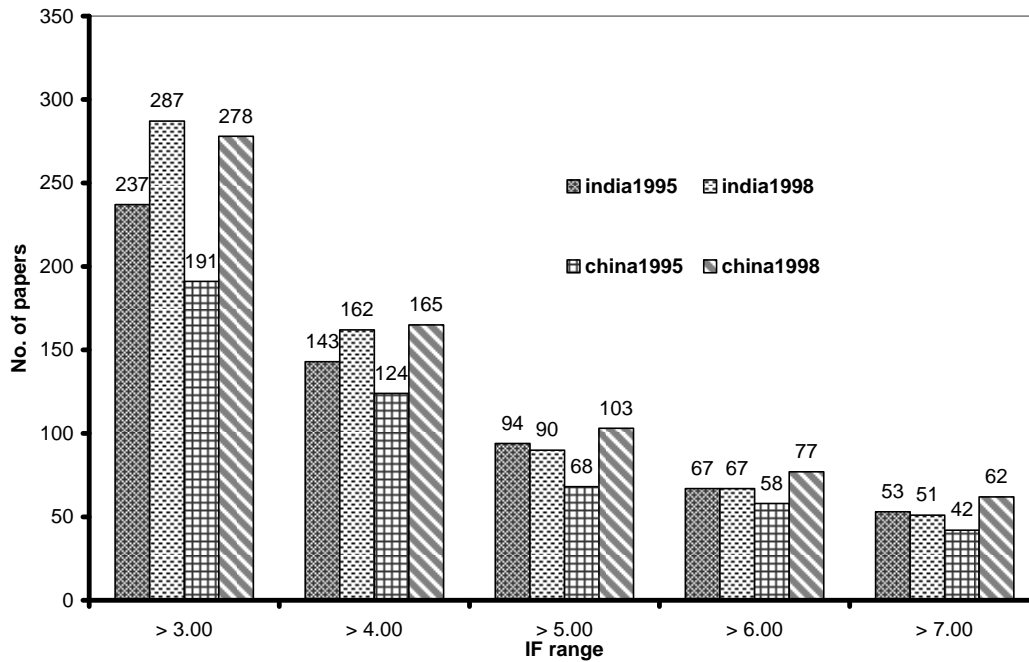
Journals used by Indian and Chinese researchers - Indian scientists used 1294 journals in all to publish these papers – 633 journals in 1992, 727 journals in 1995 and 818 journals in 1998. China used 466 journals in 1995 and 628 journals in 1998 to publish their papers. More than 17 % of the research papers from China were in Chinese language. India published papers in 75 subfields of new biology (*ISI deluxe classification for journals*) in the three years and China published in 59 and 61 subfields in 1995 and 1998 respectively. *Biochemistry & Biophysics* is the area in which India and China publish research the most. *Plant Sciences, Biotechnology & applied microbiology, Microbiology, Pharmacology & toxicology, and Chemistry* are among the prominent subfields where Indian and Chinese scientists publish research. Both India and China publish most number of research articles in US journals. Other than US, India publishes papers in journals originating in 42 countries around the globe and China publishes in journals from 27 countries. Both India and China publish only a small fraction of their papers in high impact journals: 385 Indian papers were published in journals of impact factor higher than 4.0 in the three years (1992, 1995 and 1998) and 289 Chinese papers were published in journals having impact factor greater than 4.0 in 1995 and 1998.

Table 1. Papers published from India and China in journals of impact factor greater than 3.0

Year	All Papers			Papers in journals of IF > 3.0			
	No. of inst	No. of journals	No of papers	No.of inst	No.of journals	No. of papers	% of papers
India							
1995	454	727	2686	79	77	237	8.8
1998	465	818	3257	106	88	287	8.8
China							
1995	377	467	1448	96	70	191	13.2
1998	555	629	2418	152	113	278	11.5

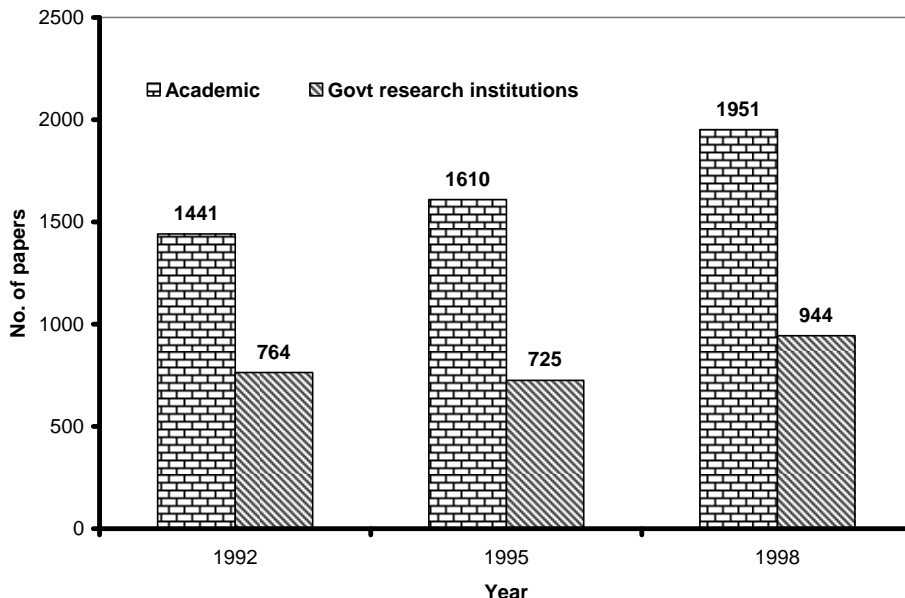
India had used 77 journals of impact factor greater than 3.0 in 1995 to publish 237 papers, and 88 journals of impact factor greater than 3.0 in 1998 to publish 287 papers (Table 1). China had used 70 journals with impact factor greater than 3.0 in 1995 to publish 191 papers and 113 journals with impact factor greater than 3.0 in 1998 to publish 278 papers. While in the case of India the annual increase of papers published in all journals (irrespective of impact factor) and in journals of impact factor greater than 3.0 was roughly equal (about 6.6%) during 1995-1998, in China the overall compound growth rate was 18.64% and the growth rate of papers published in journals with impact factor greater than 3.0 was 13.32%. The distribution of Indian and Chinese research papers by impact factor range of journals is shown in Fig. 2.

Figure 2. Distribution of Indian and Chinese Research papers by impact factor range of journals



Institutions active in new biology research - More than 800 Indian institutions contributed to new biology research. *Indian Institute of Science Bangalore, Banaras Hindu University, Varanasi, and University of Madras, Chennai* had published the largest number of papers from India. In all, new biology research papers in India come from 224 cities, with seven of them (Delhi, Chennai, Bangalore, Hyderabad, Calcutta, Bombay and Lucknow) publishing more than 100 papers in each of the three years 1992, 1995 and 1998. Among the 28 Indian states publishing papers, seven published more than 200 papers each every year. About 60% of Indian research papers come from academic institutions. Contrary to the general perception, academic institutions contribute the bulk of India's papers. The numbers of papers from Indian academic institutions and government laboratories (CSIR, ICMR, ICAR, DBT, DST, DAE and DRDO) in the three years are shown in Fig. 3. While the compound annual growth rate in papers from the academic sector in the six years is 5.07% (from 1,431 papers in 1992 to 1,926 in 1998), the growth rate in the government laboratory sector was merely 3.59% (from 764 in 1992 to 944 in 1998). In particular, there has been a substantial rise in the number of papers from general universities (including Indian Institute of Science and Indian Institutes of Technology) – from 1001 to 1502, a compound growth rate of 6.99%.

Figure 3. Contributions made by Indian academic institutions and government research institutions



In 1995 and 1998, new biology research papers came from 377 and 555 Chinese institutions respectively. The leading institutions were Shanghai Institute of Biochemistry, Beijing Medical University, Institute of Biophysics, Beijing, Nanjing University and Shanghai Medical University. None of them had published more than a hundred papers in a year. Chinese papers in new biology have come from 71 cities in 1995 and 91 cities in 1998. The leading cities are Beijing, Shanghai, Nanjing and Wuhan.

Use of high impact journals by different institutions - Indian Institute of Science, with 25 papers in 1995 and 44 in 1998 and Tata Institute of Fundamental Research, Mumbai, with 15 papers in 1995 and 13 in 1998, are the only Indian institutions to have published more than ten papers in journals of impact factor greater than 3.0 in

Table 2. Research output of prominent institutions

No.	Institution	1995		1998	
		Total	IF >3	Total	IF >3
INDIA					
1	INDIAN-INST-SCI : BANGALORE	114	25	143	44
2	BANARAS-HINDU-UNIV : VARANASI	89	6	128	1
3	UNIVERSITY OF DELHI, NEW DELHI	79	10	78	6
4	UNIV-MADRAS : MADRAS	63	7	80	8
5	CENT-DRUG-RES-INST : LUCKNOW	57	2	57	4
6	BHABHA-ATOM-RES-CTR : BOMBAY	52	4	56	4
7	JAWAHARLAL-NEHRU-UNIV : NEW-DELHI	48	6	71	9
8	NATL-CHEM-LAB : PUNE	45	5	74	2
9	CTR-CELLULAR-&-MOL-BIOL : HYDERABAD	44	9	42	6
10	CENT-FOOD-TECHNOL-RES-INST : MYSORE	42	1	42	0
11	ALIGARH-MUSLIM-UNIV : ALIGARH	40	6	34	1
12	INDIAN-INST-TECHNOL : MADRAS	39	1	36	1
CHINA					
1	ACAD-SINICA, INST BIOPHYS, BEIJING	51	7	82	13
2	SHANGHAI INST BIOCHEM, SHANGHAI	69	4	81	4
3	SHANGHAI-MED-UNIV, SHANGHAI	43	9	78	15
4	BEIJING-UNIV, BEIJING	50	6	67	6
5	NANJING-UNIV, NANJING	44	5	62	4

both 1995 and 1998 (Table 2). University of Delhi had published ten papers in 1995 in journals of impact factor greater than 3.0 and All India Institute of Medical Sciences, Indian Institute of Chemical Biology and Centre for Cellular and Molecular Biology chipped in with nine papers each. All India Institute of Medical Sciences (14), National Institute of Immunology (10) and Indian Institute of Chemical Biology (10) had published ten or more papers in such journals in 1998. Many institutions that have published at least five papers in journals of impact factor greater than 3.0 have published more papers in 1998 than in 1995. These include Indian Institute of Science, All India Institute of Medical Sciences, Tata Institute of Fundamental Research, Indian Institute of Chemical Technology, National Institute of Immunology, Jawaharlal Nehru University, University of Madras, International Centre for Genetic Engineering and Biotechnology, National Institute of Cholera and Enteric Diseases, Institute of Microbial Technology, Bhabha Atomic Research Centre, Christian Medical College, National Chemical Laboratory, University of Hyderabad, and Banaras Hindu University. The National Centre for Biological Science, Bangalore, which was earlier called the TIFR Centre, had published nine papers in 1995 and eight in 1998; of these 14 were published in journals of impact factor greater than 4.5 and two in journals of impact factor between 4.0 and 4.5 (Table 3).

Beijing Medical University is the only Chinese institution to have published more than 15 papers in both 1995 and 1998 in journals of impact factor greater than 3. Shanghai Medical University published 15 papers in 1998 and nine papers in 1995

Table 3. Research papers from NCBS published in 1995 and 1998 along with the number of citations

No.	Author	Source	IF-1997	Times Cited	Int. Coll.
1	Vijayraghavan-K	CURR BIOL, 1998, 8 , R327-R327	6.667		No
2	Roy-S Vijayraghavan-K	J CELL BIOL, 1998, 141 , 1135-1145	12.005	6	No
3	Bhuyan-AK Udgaonkar-JB	PROTEINS STRUCT FUNCT GENET, 1998, 30 , 295-308	4.161	9	No
4	Anant-S Roy-S Vijayraghavan-K	DEVELOPMENT, 1998, 125 , 1361-1369	9.781	9	No
5	Bhuyan-AK Udgaonkar-JB	BIOCHEMISTRY, 1998, 37 , 9147-9155	4.572	7	No
6	Mayor-S Sabharanjak-S Maxfield-FR	EMBO J, 1998, 17, 4626-4638	12.643	35	Yes
7	Varma-R Mayor-S	NATURE, 1998, 394 , 798-801	27.368	138	No
8	Bhuyan-AK Udgaonkar-JB	PROTEINS STRUCT FUNCT GENET, 1998, 32 , 241-247	4.161	1	No
Subtotal				205	
9	Agashe-VR Udgaonkar-JB	BIOCHEMISTRY, 1995, 34 , 3286-3299	4.572	57	No
10	Daniel-B Mukherjee-G Seshadri-L Vallikad-E Krishna-S	J GEN VIROL, 1995, 76 , 2589-2593	2.863	22	No
11	Khurana-R Hate-AT Nath-U Udgaonkar-JB	PROTEIN SCI, 1995, 4 , 1133-1144	4.6	15	No
12	Lee-JC Vijayraghavan-K Celniker-SE Tanouye-MA	PROC NAT ACAD SCI USA, 1995, 92 , 10344-10348	9.04	17	Yes
13	Udgaonkar-JB Baldwin-RL	BIOCHEMISTRY, 1995, 34 , 4088-4096	4.572	22	Yes
14	Vijayraghavan-K	BIOESSAYS, 1995, 17 , 195-198	7.053	2	No
15	Agashe-VR Shastry-MCR Udgaonkar-JB	NATURE, 1995, 377 , 754-757	27.368	50	No
16	Nath-U Udgaonkar-JB	BIOCHEMISTRY, 1995, 34 , 1702-1713	4.572	22	No
17	Raghu-P Hasan-G	DEV BIOL, 1995, 171 , 564-577	5.289	8	No
Subtotal				215	

in journals of impact factor greater than 3. During these three years, China's annual percentage increase of papers published in journals of impact factor greater than 3.0 is 13.32%, which is more than twice that of India's (6.58%).

Citations - Of the 2686 papers from India published in 1995, only 1935 were cited at least once up to 2000 and in all they received 8209 citations as seen from *Science Citation Index* (CD edition). Thirty Indian papers were cited more than 25 times, of which six were cited more than 50 times (Table 4). Of the 1448 Chinese papers published in 1995, only 578 were cited at least once up to 2000. In all these were cited 3151 times. Of these, 17 papers were cited more than 25 times and four were cited more than 50 times.

International collaboration - About 13% of Indian papers have resulted from international collaboration as compared to 32% of Chinese papers. India collaborated

Table 4. List of highly cited papers

No.	Articles	Institution	No. of citations	Int. Coll
INDIA				
1	Prasad R, CURR GENET,1995, 27 ,320	JAWAHARLAL NEHRU UNIV : NEW DELHI	93	Yes
2	Caldwell M, AMBIO,1995, 24 ,166	MADURAI KAMARAJ UNIV :MADURAI	68	Yes
3	Celliar M, P NATL ACAD SCI USA,1995, 92 , 89	TATA INST FUNDAMENTAL RES : BOMBAY	65	Yes
4	Agashe VR, BIOCHEMISTRY,1995, 34 , 3286	TATA INST FUNDAMENTAL RES : BANGALORE	57	No
5	Mountain JL, AM J HUM GENET, 56 , 979	INDIAN INST SCI : BANGALORE	55	Yes
6	Agashe VR, NATURE,1995, 377 ,754	NATL CTR BIOL SCI : BANGALORE 560012, KARNATAKA	50	No
7	Barrett T, STRUCTURE,1995, 3 , 951	NATL CHEM LAB, DIV BIOCHEM SCI, POONA	49	Yes
8	Ruby AJ, CANCER LETT,1995, 94 , 79	UNIV KERALA, TRIVANDRUM, KERALA	48	No
9	Mitra AB, J NATL CANCER I, 1995, 87 , 742	INDIAN COUNCIL MED RES :NEW DELHI, NEW DELHI	45	Yes
10	Raman B, FEBS LETT, 1995, 365 , 133	CENTER FOR CELLULAR & MOLECULAR BIOLOGY, HYDERABAD	44	No
Other 1925 articles			7635	
Total			8209	
CHINA				
1	Moises HW, NAT GENET,1995, 11 , 321	W CHINA UNIV MED SCI, CHENGDU	147	Yes
2	Burnashev N, J PHYSIOL LONDON,1995, 485 , 4	HUAZHONG UNIV SCI & TECHNOL, WUHAN	81	Yes
3	Vansoolingen D, J CLIN MICROBIOL,1995, 33	BEIJING TB & CHEST TUMOR RES INST, BEIJING	59	Yes
4	Song JL, EUR J BIOCHEM,1995, 23 , 312	INST BIOPHYS, BEIJING	55	No
5	Zhang ZG, PLANT J,1995, 8 , 139	HUAZHONG AGR UNIV, WUHAN	47	Yes
6	Ci YX, ANAL CHEM,1995, 67 , 1785	BEIJING UNIV, BEIJING	45	No
7	Sun Y, ONCOGENE,1995, 10 , 785	HUNAN MED UNIV, CHANGSHA	36	Yes
8	Zhang RW, J BONE MINER RES, 1995, 10 , 415	BEIJING MED UNIV, BEIJING	33	Yes
9	Jiang CJ, GENETICS,1995, 140 , 1111	JIANGSU AGR COLL, JIANGSU	30	Yes
10	Vaca CE, CARCINOGENESIS, 1995, 16 , 1847	INST OCCUPAT MED, BEIJING	30	Yes
Other 568 articles			2588	
Total			3151	

mostly with *US, followed by Germany, Japan, UK and France* and 65 other nations in new biology research. Half of the highly-cited Indian papers are written in collaboration with authors from abroad. China is collaborating mostly with US, Japan, Germany and UK. The number of countries China collaborated with rose from 43 in 1995 to 48 in 1998.

Conclusions

Unlike in advanced countries, life science research in India and China accounts for less than a quarter of their S&T research output. China has overtaken India in the number of new biology research papers in 1999, much later than overtaking India in S&T as a whole, mathematics and chemistry. Also, China's publication output is growing at four times the rate at which India's output is growing. Contrary to widely prevailing notions, academic institutions are contributing the bulk of India's papers in new biology research (about 60%), although they account for only about 40% of the 1995 papers cited at least 25 times up to the end of 2000. Only a few institutions have published papers in high impact journals in both 1995 and 1998. About a dozen Indian institutions have published a larger number of papers in journals of impact factor greater than 3.0 in 1998 than in 1995, although overall the share of such papers from India as a whole has remained the same at 8.8%.

It is not just in research that China is doing better than India. India joined the Internet in 1988, six years before China, but by 1999 China had already overtaken India. What is more, China is maintaining and even extending the lead (Press et al., 2002). According to a recent report, China is second only to USA in the number of Internet users (UNCTAD, 2002). Both India and China are assiduously seeking foreign investment, and here again China, with \$47 billion, attracted more than thirteen times foreign direct investment as India (\$3.4 billion) in 2001 (Narasimhan 2002). In 2002, one year after China joined WTO, China is expecting to attract more than \$50 billion in foreign direct investment.

Acknowledgment – Mr S Selvam and Mr S Gunasekaran assisted in data processing. The Department of Biotechnology, Government of India, provided financial assistance.

References

- Arunachalam, S. (2002). Is science in India on the decline?. *Current Science*, 83, 107-108.
- Arunachalam, S. & Gunasekaran, S. (2002). Tuberculosis research in India and China: From bibliometrics to research policy. *Current Science*, **82**, 933-947.
- Jayaraman, K. S. (2002). India's scientists agonize over fall in publication rate. *Nature*, 419, 100.
- Narasimhan, C. R. L. (2002). *The Hindu*, 12 October. Quoting UNCTAD's World Investment Report.
- National Science Board (2002). *Science and Engineering Indicators 2002*. National Science Foundation, Arlington, VA, USA.
- Padmanabhan, G. (2002). Has Indian science slowed down?. *Current Science*, 83, 1055.
- Press, L., Foster, W., Wolcott, P. & McHenry, W. (2002). *The Internet in India and China*. *First Monday*, 7 October.
[http://firstmonday.org/issues/issue7_10/press/index.html].
- UNCTAD. (2002). *E-commerce and development report 2002*. Geneva.
[http://r0.unctad.org/ecommerce/ecommerce_en/edr02_en.htm].