3. John, V. O., Holl, G., Allan, R. P., Buehler, S. A., Parker, D. E. and Soden, B. J., Clear sky biases in satellite infrared estimates of upper tropospheric humidity and its trends. *J. Geophys. Res.*, 2011, **116**(D05205); doi:10.1029/2010JD014192.
4. Soden, B. J. and Bretherton, F. P., Upper tropospheric relative humidity from the goes 6.7 µm channel: method and climatology for July 1987. *J. Geophys. Res.*, 1993, **98**(D9), 16669–16688.
5. Soden, B. J. and Fu, R., A satellite analysis of deep convection, upper tropospheric humidity, and greenhouse effect. *J. Climate*, 1995, **8**, 2333–2351.
6. Buehler, S. A., Eriksson, P., von Engeln, A. and Verden, C., Arts, the atmospheric radiative transfer simulator. *Q. J. R. Meteor. Soc.*, 2005, **91**(1), 65–93; doi:10.1029/2004JD005111.
7. Buehler, S. A., John, V. O., Milz, M., Soden, B. J., Jackson, D. L. and Notholt, J., An upper tropospheric humidity data set from operational satellite microwave data. *J. Geophys. Res.*, 2008, **113**(D14110).
8. Xavier, P. K., John, V. O., Buehler, S. A., Ajayamohan, R. S. and Sijikumar, S., Variability of Indian summer monsoon in a new upper tropospheric humidity data set. *Geophys. Res. Lett.*, 2010, **37**(L05705).
9. Gohil, B. S., Gairola, R. M., Mathur, A. K., Varma, A. K., Mahesh, C., Gangwar, R. K. and Pal, P. K., Algorithms for retrieving geophysical parameters from the MADRAS and SAPHIR sensors of the Megha-Tropiques satellite: Indian scenario. *Q. J. R. Meteor. Soc.*, 2012; doi:10.1002/qj.2041.
10. Raju, G., Engineering challenges in the Megha-Tropiques satellite. *Curr. Sci.*, 2013, **104**(12), 1662–1670.
11. Suresh Raju, C., Antony, T., Mathew, N., Uma, K. N. and Moorthy, K. K., MT-MADRAS brightness temperature analysis for terrain characterization and land surface microwave emissivity estimation. *Curr. Sci.*, 2013, **104**(12), 1643–1649.
12. Eymard, L. *et al.*, The SAPHIR humidity sounder, Megha-Tropiques 2nd Scientific Workshop, Paris, France, 2–6 July 2011.
13. Chaboureau, J. P., Chedin, A. and Scott, N. A., Remote sensing of the vertical distribution of water vapor from the TOVS observations: method and validation. *J. Geophys. Res.*, 1998, **103**(D8), 8743–8752.
14. Buehler, S. A. and John, V. O., A simple method to relate microwave radiances to upper tropospheric humidity. *J. Geophys. Res.*, 2005, **110**(D02110); doi:10.1029/2004JD005111.
15. Melsheimer *et al.*, Inter-comparison of general purpose clear sky atmospheric radiative transfer models for the millimeter/submillimeter spectral range. *Radio Sci.*, 2005, **40** (RS1007); doi:10.1029/2004RS003110.
16. Hong, G., Heygster, G. and Kunzi, K., Detection of tropical deep convective clouds from AMSU-B water vapor channels measurements. *J. Geophys. Res.*, 2005, **110**(D05205), 1–15.
17. Burns, B. A., Wu, X. and Diak, G. R., Effects of precipitation and cloud ice on brightness temperature in AMSU moisture channels. *IEEE Trans. Geosci. Remote Sensing*, 1997, **5**(6), 1429–1437.
18. Thapliyal, P. K., Shukla, M. V., Shah, S., Pal, P. K., Joshi, P. C. and Kottayil, A., An algorithm for the estimation of upper tropospheric humidity from Kalpana observations: methodology and validation. *J. Geophys. Res.*, 2011, **116**, 1–16; doi:10.1029/2010JD014291.
19. Meenu, S., Rajeev, K., Parameswaran, K. and Nair, A. K. M., Regional distribution of deep clouds and cloud top altitudes over the Indian subcontinent and the surrounding oceans. *J. Geophys. Res.*, 2010, **115**(D05205); doi:10.1029/2009JD011802.
20. Nair, A. K. M., Rajeev, K., Sijikumar, S. and Meenu, S., Characteristics of a persistent pool of inhibited cloudiness and its genesis over the Bay of Bengal associated with the Asian summer monsoon. *Ann. Geophys.*, 2011, **29**(4), 1247–1252.
21. Zhang, C., Double ITCZ. *J. Geophys. Res.*, 1993, **106**, 11785–11792.
22. Meenu, S., Rajeev, K., Parameswaran, K. and Suresh Raju, C., Characteristics of the double inter-tropical convergence zone over the tropical Indian ocean. *J. Geophys. Res.*, 2007, **112**(D11106); doi:10.1029/2006JD007950.

ACKNOWLEDGEMENTS. We thank Dr Parameswaran and Dr K. Rajiv, Space Physics Laboratory, Vikram Sarabhai Space Centre, Thiruvananthapuram for fruitful discussions and careful reading of the manuscript. We also thank MOSDAC, SAC for supplying the SAPHIR data and humidity products (http://mosdac.gov.in). V.O.J. was supported by the joint DECC/Defra Met Office Hardley Centre Climate Programme (GA01101), the UK JWCRP and EUMETSAT CMSAF. Radiosonde data are from http://weather.uwyo.edu/upperair/sounding.html.

Received 12 September 2014; revised accepted 23 March 2015

Indian and Chinese higher education institutions compared using an end-to-end evaluation

S. Savithri* and Gangan Prathap

CSIR-National Institute for Interdisciplinary Science and Technology, Thiruvananthapuram 695 019, India

The latest (2014) release of the SCImago Institutions Rankings (SIR) allows to compare the research performance of leading higher education institutions in India and China using an end-to-end bibliometric performance analysis procedure. Six carefully chosen primary and secondary bibliometric indicators summarize the chain of activity: input–output–excellence–outcome–productivity. From principal component analysis it is established that the primary indicators are orthogonal and represent size-dependent quantity and size-independent quality/productivity dimensions respectively. Using this insight two-dimensional maps can be used to visualize the results.

Keywords: Bibliometrics indicators, higher education institutions, principal component analysis, research performance.

BOTH in India and China, the higher education institutions (HEIs) taken together are the key contributors to their academic research output. HEIs constitute 19 out of the top 20 research entities in China, and 15 out of the top 20 in India according to the latest SCImago Institutions Rankings (SIR) World Reports (http://www.scimagoir.com/). Overall there are 391 HEIs from China and 156 from India in the SIR list for 2014. Most of these can be considered as significantly research-intensive. So far, the

*For correspondence. (e-mail: sivakumarsavi@gmail.com)

global university rankings of HEIs, e.g. the Academic Ranking of World Universities (http://www.shanghairanking.com/ARWU2014.html), the Leiden rankings, the Taiwan Higher Education Accreditation Evaluation Council university ranking (HEEACT), and the EU Assessment of University-Based Research (AUBR), etc. do not perform an end-to-end evaluation of research performance from size of research manpower participating to excellence and outcome. In this communication, we attempt to do such an evaluation.

The latest SIR report has introduced a new feature that makes such an end-to-end evaluation possible. This is called the scientific talent pool (STP), which gives the number of authors from an institution who have participated in the total publication output of that institution during that particular period of time. This indicator can be taken as a reasonable proxy of the input at the beginning of the chain that performs scientific research activity.

The SIR reports also give output indicators which can be interpreted as belonging to quantity (size-dependent) and quality (size-independent) dimensions. The ratio of the quantity of output to input is a proxy for productivity, but without taking into account the quality of research. When quality is taken into account, one can compute a size-dependent composite performance indicator which is the measure of the outcome of the research effort. The ratio of the outcome to the input then becomes a measure of the quality-linked productivity of the institution. We thus have an end-to-end performance analysis based on the input–output–excellence–outcome–productivity chain that leads to six variables. Thus for 2014, we have a 391 × 6 matrix of data for China and another 156 × 6 matrix for India.

Hendrix[1] used principal component analysis (PCA) to evaluate institutional-level performance of medical research institutions by classifying and clustering various bibliometric indicators. The variables clustered neatly into three distinct groups: the first cluster refers to size-dependent input and output terms and comprises of the total number of faculty (input), total number of papers (output), and total number of citations (outcome). The second factor comprises of size-independent terms that reflect the impact of a researcher, average number of citations per article, etc. and can be interpreted as a quality or excellence dimension. The third group describes research productivity and impact at the individual level, like the number of papers and number of citations per faculty member.

We follow this approach and use PCA (not reported here) for both China and India to confirm that with six variables, three components are needed and that when the output/input ratio is discarded, so that there are only five variables, two components suffice to account for most of the common variance. These are the size-dependent quantity indicators and the size-independent quality and productivity indicators, which are clearly orthogonal to the former. This insight allows us to represent and visualize the data as two-dimensional maps.

In this communication, we restrict attention only to the 391 and 156 HEIs that are ranked as the top research-intensive tertiary educational institutions from China and India respectively, for 2014. The data are taken from the 2014 release of SIR World Reports. The new release only presents ranks and normalized grades are given. This means that only indirect surrogate indicators can be computed.

Table 1 shows that there are six indicators: three primary indicators and three secondary indicators that are

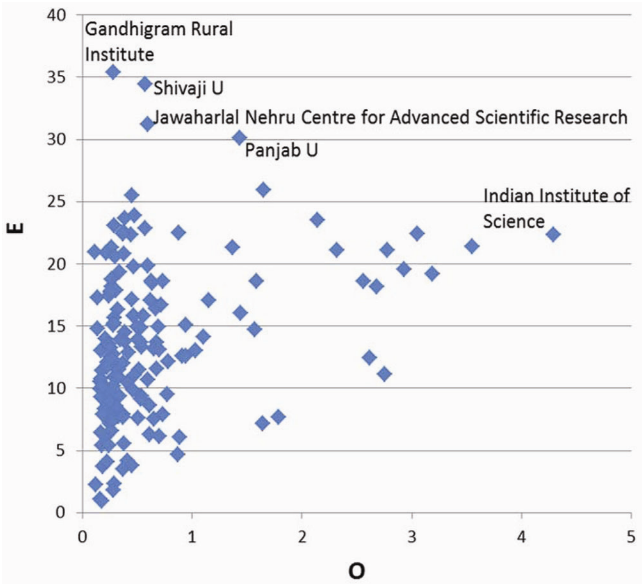

Figure 1. Scatterplot of the 156 higher education institutions (HEIs) from India on the *E* versus *O* parameter space.

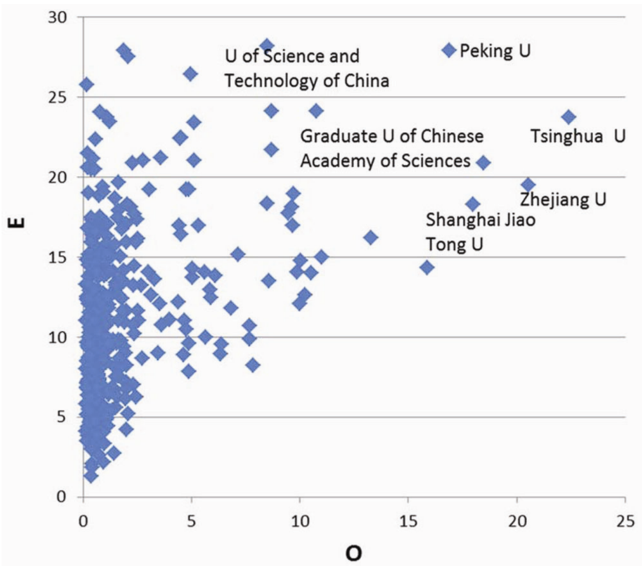

Figure 2. Scatterplot of the 391 HEIs from China on the *E* versus *O* parameter space.

Table 1. The primary indicators or variables and the derived indicators

Indicator or variable	Description	Size dependence	Formula
STP	Scientific talent pool	Dependent	STP
O	Output	Dependent	O
E	Excellence	Independent	E
X	Outcome or performance	Dependent	$X = E^2 \times O$
O/STP	Output productivity	Independent	O/STP
X/STP	Outcome productivity	Independent	X/STP

Table 2. The primary indicators or variables and the derived indicators for principal component analysis (PCA) is shown for the top 20 higher education institutions (HEIs) in India

Institution	STP	O	E	X	O/STP	X/STP
Gandhigram Rural Institute, Gandhigram	0.19	0.28	35.41	351.08	1.47	1847.81
Jawaharlal Nehru Centre for Advanced Scientific Research, Bengaluru	0.45	0.60	31.25	585.94	1.33	1302.08
Shivaji University, Kolhapur	0.57	0.57	34.47	677.26	1.00	1188.18
Panjab University, Chandigarh	1.28	1.43	30.12	1297.32	1.12	1013.53
Tata Institute of Fundamental Research, Mumbai	1.32	1.65	25.94	1110.26	1.25	841.10
Visva-Bharati University, Santiniketan	0.38	0.45	25.46	291.70	1.18	767.62
Indian Institute of Technology, Roorkee	1.85	2.14	23.53	1184.83	1.16	640.45
Indian Institute of Technology, Delhi	2.41	3.05	22.44	1535.84	1.27	637.28
Indian Institute of Science, Bengaluru	3.81	4.29	22.31	2135.29	1.13	560.44
Indian Institute of Technology, Patna	0.09	0.11	20.97	48.37	1.22	537.46
National Institute of Technology Durgapur	0.51	0.48	23.87	273.49	0.94	536.26
Indian Institute of Technology, Kharagpur	3.11	3.55	21.41	1627.28	1.14	523.24
National Institute of Technology, Hamirpur	0.25	0.27	21.35	123.07	1.08	492.29
Doctor Harisingh Gour University, Sagar	0.40	0.38	22.54	193.06	0.95	482.65
Indian Institute of Technology, Guwahati	1.31	1.37	21.29	620.97	1.05	474.02
Indian Institute of Technology, Kanpur	2.19	2.32	21.08	1030.93	1.06	470.74
Tezpur University, Tezpur	0.36	0.38	20.80	164.40	1.06	456.68
National Institute of Technology, Tiruchirappalli	1.01	0.88	22.50	445.50	0.87	441.09
Indian Institute of Technology, Bombay	2.95	2.78	21.12	1240.03	0.94	420.35
University of Jammu, Jammu	0.38	0.29	23.09	154.61	0.76	406.88

derived from these. The key input indicator is the STP, the total number of authors from an institution in the total publication output of that institution during that period of time. It is a measure of the input that participates and contributes to scientific research activity at the institution. The O (or output) indicator in SIR is a measure of the quantity or size of the publication output of an institution and is the total number of documents published in scholarly journals indexed in Scopus. SIR has several indicators which are proxies of the quality of academic research output, but here we restrict attention to E (or excellence rate), which indicates the percentage of an institution's scientific output that is included into the set formed by 10% of the most cited papers in the respective scientific fields, and serves as a measure of the high-quality output of research institutions.

All variables in the SIR are normalized on the 0–100 scale. The quantity or size dimension is the number of articles published during a five-year window, normalized on the 0–100 scale, where in 2014, the Centre National de la Recherche Scientifique (CNRS), France was listed as the top-ranking institution in the world with the score of 100. We indicate this normalized quality indicator by O. The second dimension is quality. Again, for each year, these values are normalized so that the highest ranking performer has a score of 100. In 2014, the Broad Institute of MIT and Harvard occupied the top rank with an excellence rate score of 100. We indicate this normalized quality indicator by E. The size-dependent input indicator STP is also normalized in the same manner as above and again in 2014, CNRS, France was listed as the largest institution in the world with the score of 100.

The primary input-side indicator is therefore STP and the primary output-side indicators are O and E. From this we need to generate a single-valued composite outcome indicator. We find that the second-order indicator called the exergy term combining the quantity and quality indicators, $X = E^2 \times O$ is appropriate for this. We have now two possibilities for computing a productivity indicator, an output-based one computed as O/STP and an outcome-based one computed as X/STP. We thus have an end-to-end performance analysis: input–output–excellence–outcome–productivity based on six variables according to the scheme summarized in Table 1.

Table 3. The primary indicators or variables and the derived indicators for PCA shown for the top 20 HEIs in China

Institution	STP	*O*	*E*	*X*	*O*/STP	*X*/STP
University of Science and Technology of China, Hefei	10.98	8.47	28.21	6740.46	0.77	613.89
Peking University, Beijing	24.76	16.87	27.92	13150.61	0.68	531.12
Chongqing University of Arts and Sciences, Chongqing	0.19	0.21	21.47	96.80	1.11	509.48
Tsinghua University, Beijing	24.77	22.41	23.72	12608.73	0.90	509.03
Northeast Normal University, Changchun	2.97	1.87	27.92	1457.71	0.63	490.81
Huazhong Normal University, Wuhan	3.26	2.07	27.52	1567.72	0.63	480.89
Nankai University, Tianjin	7.39	4.97	26.45	3477.02	0.67	470.50
Huzhou Teachers College, Huzhou	0.35	0.39	20.47	163.42	1.11	466.91
Guangdong University of Foreign Studies, Guangdong	0.18	0.19	20.63	80.86	1.06	449.24
Renmin University of China, Beijing	1.55	1.24	23.45	681.88	0.80	439.92
Nanjing University, Nanjing	11.73	8.7	24.1	5053.05	0.74	430.78
Shanghai Normal University, Shanghai	1.47	1.11	23.72	624.53	0.76	424.85
Fudan University, Shanghai	16.59	10.75	24.12	6254.07	0.65	376.98
Bohai University, Jinzhou	0.53	0.44	21.14	196.64	0.83	371.01
South Central University for Nationalities, Wuhan	0.77	0.57	22.35	284.73	0.74	369.78
Graduate University of the Chinese Academy of Sciences, Beijing	23.46	18.49	20.87	8053.45	0.79	343.28
East China University of Science and Technology, Shanghai	6.79	4.52	22.50	2288.25	0.67	337.00
Xiamen University, Xiamen	8.42	5.1	23.43	2799.72	0.61	332.51
Anhui Normal University, Wuhu	1.38	0.79	24.06	457.32	0.57	331.39
East China Normal University, Shanghai	4.87	3.56	21.20	1600.01	0.73	328.54

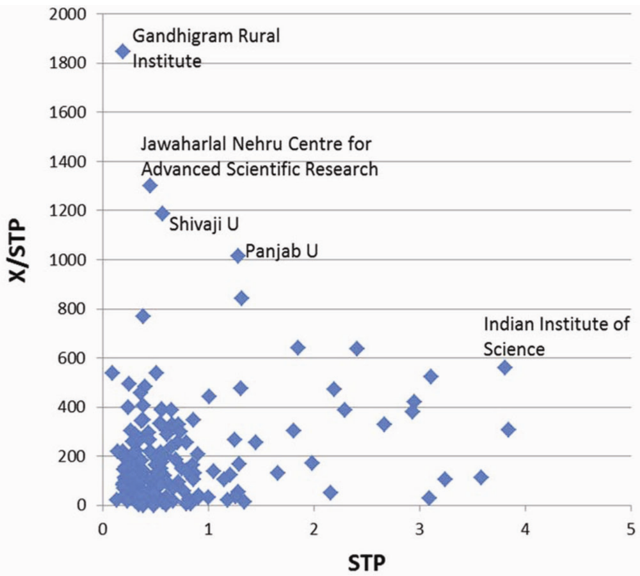

Figure 3. Scatterplot of the 156 HEIs from India on the *X*/STP versus STP parameter space.

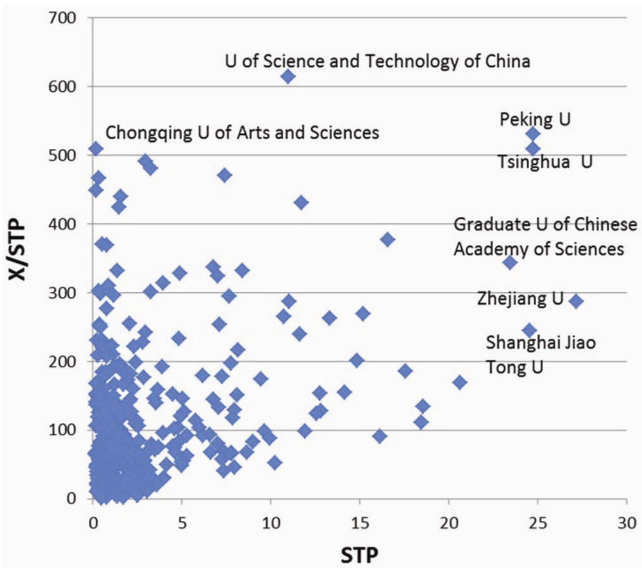

Figure 4. Scatterplot of the 391 HEIs from China on the *X*/STP versus STP parameter space.

Tables 2 and 3 show typically the primary indicators or variables and the derived indicators for PCA for the top 20 HEIs from the list of Indian and Chinese institutions which were studied. We studied various combinations of variables and components (factors) and found that with all six variables, if three components are extracted, the results are difficult to interpret and visualize. When the output-based productivity ratio *O*/STP was discarded, so that there are only five variables, two components suffice to account for most of the common variance and these are the size-dependent quantity indicators and the size-independent quality and productivity indicators which PCA shows to be orthogonal to the former. It was found that the quality term based on excellence *E* is very highly correlated with *X*/STP. This follows from the fact that $X/\text{STP} = O/\text{STP} \times E^2$, and that *O* and STP are highly correlated with each other for the 156 HEIs in India and the 391 HEIs in China. PCA also confirmed (not shown here) that the two primary size-dependent (*O* and STP) indicators are clearly orthogonal to the two size-independent indicators (*E* and *X*/STP). Therefore, the orthogonal quantity–quality parameter space of output *O* and *E* and quantity–productivity parameter space of STP and *X*/STP will display neatly as two-dimensional graphs. Note that the orthogonal pair of O and E corresponds to a base of orthogonal metrics capturing the quality and excellence aspect of research performance and X then becomes a composite performance indicator for this[2]. Similarly, the orthogonal pair of STP and X/STP corresponds to a base of orthogonal metrics capturing the productivity aspect of research performance and X^2/STP then becomes a composite performance indicator for this. Figures 1–4 capture these as scatterplots on the respective pairs of parameter space for both countries. Some interesting results are noticed. The large HEIs in China are nearly five times bigger than their counterparts in India. However, India has a noticeable edge in quality/excellence and productivity. Both in China and India, two small institutions, the Chongqing University of Arts and Sciences (STP = 0.19) and the Gandhigram Rural Institute (STP = 0.19) perform extremely well on the size-independent performance indicators. In terms of size, the Indian Institute of Science, Bengaluru, and the Tsinghua and Peking Universities of China lead in their respective higher education systems.

The performance of leading HEIs in India and China has been compared using an end-to-end bibliometric performance analysis. Six carefully chosen primary and secondary bibliometric indicators summarize the chain of activity: From PCA it is established that the primary indicators are orthogonal and represent size-dependent quantity and a size-independent quality/productivity dimensions respectively. Two-dimensional maps can be used to visualize the results. Although the key Chinese institutions are considerably larger, the Indian counterparts have the edge in productivity and in maintaining excellence.

1. Hendrix, D., An analysis of bibliometric indicators, National Institutes of Health funding, and faculty size at Association of American Medical Colleges medical schools, 1997–2007. *J. Med. Libr. Assoc.*, 2008, **96**(4), 324–334.
2. Franceschet, M., A cluster analysis of scholar and journal bibliometric indicators. *J. Am. Soc. Inf. Sci. Technol.*, 2009, **60**(10), 1950–1964.

Received 31 October 2014; revised accepted 11 March 2015

Distribution pattern of bacteria in the two geographic poles and Southern Ocean from the reported 16S rDNA sequences

Pratibha Gupta[1], Harsh Kumar Agrawal[1] and Rajib Bandopadhyay[1,2,*]

[1]Department of Bio-Engineering, Birla Institute of Technology, Mesra, Ranchi 835 215, India
[2]UGC-Centre of Advanced Study, Department of Botany, The University of Burdwan, Golapbag, Bardhaman 713 104, India

16S rDNA bacterial sequences (913) from the Arctic Ocean, Southern Ocean and Antarctic Iceland were studied to understand the bacterial distribution pattern. Through phylogenetic study, it was observed that some bacteria were common in both the Arctic Ocean and Antarctic Iceland. *γ-Proteobacteria* occupied 77.7% of the total bacterial population in the Antarctic Iceland, whereas in the Southern Ocean it was 72.5% and in the Arctic Ocean it was 50.9%. GC (Guanine + Cytosine) content of the bacteria in the Arctic Ocean and Antarctic Iceland region was 54.4% and 53.8% respectively. Bacterial diversity was calculated using Shannon–Weiner index and was found to be highest in the Antarctic Iceland (1.6926).

Keywords: Bacterial phylogeny, bioinformatic tools, geographic poles and oceans, microbial diversity.

ABOUT 75% of the Earth's surface is covered by oceans. There are major five oceans in the world, namely Pacific Ocean, Atlantic Ocean, Indian Ocean, Southern Ocean (earlier known as Antarctic Ocean) and Arctic Ocean. The Antarctic and Arctic regions are different from each other. The Arctic Ocean is surrounded by continents and there is only 10% of freshwater inflow, whereas there is no inflow of freshwater in the Southern Ocean[1]. The Southern Ocean surrounding the Antarctic continent is driven by a current system known as Antarctic Circumpolar Current (ACC), which is the strongest current system in the world oceans[2]. At the beginning of the 20th century, there was a concept 'everything is everywhere' implying that prokaryotic population genetics can never be broken by physical isolation, but by adaptation alone[3]. One of the major questions regarding biodiversity in the two geographic poles is whether marine bacterial species are the same at both the poles or not. According to endimicity theory some microbes require special environment (hot springs, cryosphere and hyperhalophilic habitats) that are not commonly found throughout the globe[4] and thus they have restricted geographical range[5]. Studies till date have shown that members of the same genera occur

*For correspondence. (e-mail: rajib_bandopadhyay@yahoo.com)