The largest distributed network of bioinformatics centres in the world: Biotechnology Information System Network (DBT-BTISNET)

S. Krishnaswamy[1,*] and T. Madhan Mohan[2]

[1]Institute of Mathematical Sciences, Chennai , India
[2]Apex Bioinformatics Centre, Department of Biotechnology, New Delhi , India

The spread of bioinformatics centres across India is primarily due to the extensive infrastructure and network that was initiated way back in 1980s by S. Ramachandran, the first Secretary of the Department of Biotechnology (DBT), Government of India, and is being supported at 168 locations across the country. Presaging the dawn of bioinformatics globally as an integration of informatics and biology, the seeds were sown by the linkage set up by the National Informatics Centre (NIC) led by N. Seshagiri (Director General of NIC), in a remarkable collaboration with DBT. As with biotechnology, strong teaching initiatives by DBT at the post-graduate level across the country helped in the growth. The recent bioinformatics certification examinations and the consequent fellowships are part of a seminal exercise that tries to provide standards in an area which saw unregulated mushrooming growth in the last decade.

Unlike in the USA and Europe, where bioinformatics was nucleated by sequence analysis, in India the strong crystallography and biophysics structural background from the tradition of G. N. Ramachandran resulted in the initial tilt of bioinformatics in India towards structural perspectives. The Biotechnology Information System Network (BTISNET) captured through development of a plethora of databases by the centres distributed across the country recording the diversity of biological resources in the country. The advent of mega sequencing and the large-scale import of genomics and proteomics technologies resulted in the growth of many bioinformatics groups and companies, several of which have connections to the centres of the BTISNET.

Following the bioinformatics policy document, the recent years have seen the nurturing of the North Eastern Bioinformatics Network and also international collaborative ventures with countries in Asia and Europe. The trend of an admixture of experimental and computational biology approaches that are needed for systems and synthetic biology is becoming common and signals the next phase of expansion and integration of the bioinformatics network.

Keywords: Bioinformatics centres, computational biology, infrastructure and network, national databases.

Introduction

IF one does a *Google* search for the definition of bioinformatics, it gives approximately 25 million hits as of 2016. Several dictionaries, websites and articles give as many definitions of bioinformatics as there are people. Each person, as for the proverbial elephant, gives a different viewpoint. And almost all are valid. A broad definition that encompasses most viewpoints is the use of informatics, computational and experimental methods to generate, access and analyse biological information. Of course, inclusion of the generation of data is a more recent phenomenon due to the advance of technology and instrumentation. The growth of bioinformatics was seeded by the useful information obtained by analysis of protein and DNA sequences and protein structures[1].

Bioinformatics obviously has close links with biotechnology at conceptual, scientific and technological levels. Globally and in India, there has been a growing share in bioinformatics of the biotechnological processes and also markets (Figure 1)[2,3].

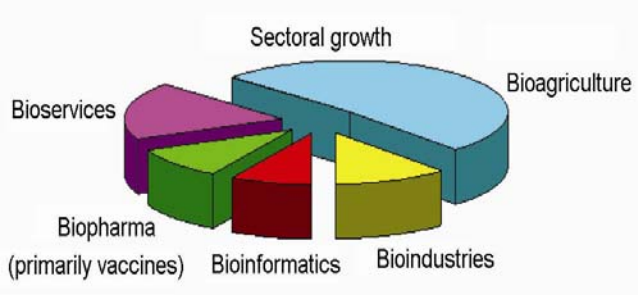

Figure 1. Sectoral contribution to biotechnology.

*For correspondence. (e-mail: mkukrishna@imsc.ernet.in)

Before the 'birth' of bioinformatics

It was only in the mid 1980s that the term 'bioinformatics' was coined. But the reference then was more to the use of informatics to medical literature and biological macromolecular data storage and access. The blossoming of bioinformatics to encompass biological diversity in the molecular and organism level is a more recent phenomenon. However, like in all scientific disciplines, researchers naturally worked at different periods in the history of science in what was classified as bioinformatics after the mid 1980s. Classical examples are Carl Linnaeus and Charles Darwin, who in fact gave the classification and evolutionary perspectives we use routinely in bioinformatics. If they were alive today, they would be dubbed as bioinformaticians.

In India, G. N. Ramachandran (GNR) and his colleagues in Madras University working in the Department of X-ray Crystallography and Biophysics, showed the power of collecting crystal structure data and subjecting them to computer analysis to provide a brilliant perspective to the delimiting of protein conformations into allowed and non-allowed regions based on steric hindrance. This is now known as the Ramachandran diagram and forms the touchstone to decide acceptance of protein structure data in the Protein Data Base (PDB). In current-day parlance, from a biological dataset (not organized into databases like now), they used informatics to do data-driven research to arrive at biological insights. Though they did not know it, GNR and fellow researchers were bioinformaticians of the first order. According to Vijayan[4]: 'Thanks to G.N. Ramachandran and his students and colleagues, India had a head-start in the area'. The genesis left by GNR in the growth of bioinformatics in India is being seeded and nurtured by macromolecular structural practitioners and different perspectives[5].

The Biotechnology Information System Network (BTISNET)

From the mid 1980s when bioinformatics was systemically growing through the incorporation of the BTISNET by the Department of Biotechnology (DBT), Government of India, to now where there has been a virtual explosion of bioinformatics activity across the country, we have come a long way. The result can be seen in the almost doubling of contribution from India, now with respect to the 1990s, in the realm of bioinformatics and computational biology (Figure 2).

BTISNET is now spread across the country with 168 centres. The network centres are in various levels and include centres of excellence (CoEs), distributed information centres (DICs), distributed information sub centres (Sub DICs) and bioinformatics infrastructure facilities (BIF). The network includes the Supercomputer Facility for Bioinformatics and Interactive Graphics Facilities. Large numbers of R&D projects in bioinformatics are also being supported through this programme. The network supports teaching programmes in M Sc, M Tech and Ph D in bioinformatics and computational biology to generate skilled manpower in bioinformatics. Further, the BTISNET centres conduct short-term training programmes/workshops for the benefit of research community, including experimental biologists. These centres also provide services to the scientific community. These activities are being coordinated by the Apex BTIC which is located in the DBT headquarter. Figure 3 gives a map showing the locations of these centres in the country.

Growth and consolidation of BTIS

According to Seshagiri[6]: 'Dr S. Ramachandran, Secretary of the Department of Biotechnology, who was a renowned Biotechnologist forecast that Bioinformatics will become more and more important in the decades to come and so decided to superpose the National Biotechnology Information System (BTIS) over the National Informatics Centre network (NICNET) to give it a quick start and sustained growth...With my prior involvement as Chairman of SERC Standing Committee on Molecular Biology and early TIFR research experience to computer simulation of protein tertiary structure and enzyme kinetic analysis and then responsibility as Director-General of NIC, Dr S. Ramachandran felt it appropriate to invite me to be the Chairperson for BTIS Standing Committee from its inception in 196–86. As I continued to be Chairman for a decade thereafter, it gave me the opportunity to see

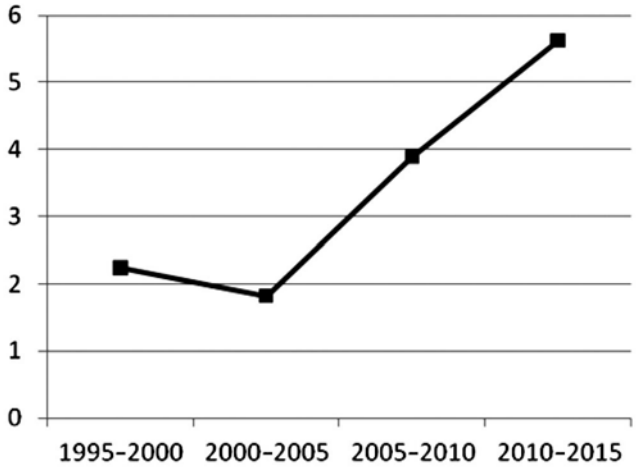

Figure 2. Increase in the contribution of bioinformatics and computational biology in India relative to the global contributions. The ratio of hits for 'bioinformatics or computational biology in India' to the total hits for bioinformatics or computational biology, as seen from *Google Scholar* searches, is shown for the periods 1995–2000, 2000–2005, 2005–2010 and 2010–2015.

at first hand the flowering of the BTIS network during the entire nucleation phase to become one of the largest such domain networks in Asia-Pacific'.

The initiative taken at that time is neatly summarized by Dharmalingam[7]: 'Bioinformatics programme started by the National Biotechnology Board of India is one of the earliest activities of the National Biotechnology Board. The only programme of NBTB, which preceded bioinformatics, was the M Sc Biotechnology programme, initiated at the following five Universities, Jawaharlal Nehru University, Madurai Kamaraj University, Banaras Hindu University, Pune University and University of Baroda in 1985. Ever since the establishment of the Department of Biotechnology in 1986, the bioinformatics programmes were nurtured with adequate funds. Though the initial mandate was only service in the fledgling Biotechnology research, the activities soon enlarged to various research and development aspects as well'.

Six CoEs in bioinformatics, and computational and systems biology have been established as part of BTISNET. These centres are well equipped with infrastructure to support research within the institute as well as neighbouring institutions. The focus of these centres is high-quality research education and services. As part of BTISNET 11 DICs in bioinformatics have been established. Fifty sub-DICs are functional as part of BTISNET at various institutions/universities. These centres were mainly established with the aim to provide service to the research community. They are now also imparting training in bioinformatics through workshops. Many centres have ventured into bioinformatics related R&D activities and have also developed information resources in the form of databases.

Figure 3. Location of BTISNET centres in India.

According to Arora[8]: 'The conceptualisation of the BTIS as a network of some major Distributed Information Centres and a few user centres stood the test of time for the gradual growth of the network...The apex centre set up at the DBT assumed a very important role to explore all possibilities to facilitate this requirement....For the first time, after the return of Dr Madhan Mohan from a training course at EMBL, a demonstration of live access to servers at EMBL was demonstrated using Telnet (Remote Login facility of Internet) via the ERNET connectivity and all the centres were issued guidelines accordingly'.

The connectivity between the centres that is so much taken for granted nowadays was not so in the early days. The BTISNET evolved as the communication technologies grew. As Vijayaditya[9] rightly points out, 'As can be seen the Bioinformatics network evolved along with the usage and technology. At every stage of this evolution a lot of discussion took place regarding the method and modes of connectivity with scientists participating actively. But one has to recognise that the small steps taken at that time, in promoting the use of computer communication technology, has enabled the researchers/scientists to equip themselves with modern computer tools and techniques to process and exchange information among them. It enabled to study the requirements and usage pattern to evolve higher speed network with necessary computing facilities and tools'.

The next level of centres started were the Bioinformatics Infrastructure Facilities (BIFs), many in colleges, especially for biology teaching through bioinformatics (BTBI). So far, 101 educational institutions have been supported under this scheme. The scheme is designed to expose teachers, scientists and students to the use of bioinformatics in solving hard-core biological problems.

As part of the BTIS policy document[2], DBT systematically nurtured the rapid development of bioinformatics centres in the North East under the North Eastern Bioinformatics Network (NEBINET) programme. Twenty-nine Bioinformatics centres were established across eight states. NEBINET has one DIC (at NEHU), two DISCs (at IBSD, and Sikkim State Council of Science and Technology) and 26 BIFs (at various universities, colleges and institutions).

Development of databases

The BTISNET centres have been involved in establishing many databases. There have also been specific initiatives involving these centres and other experts to develop integrated databases on nationally important areas.

Establishment of national rice resource database (NRRD)

Eight centres, viz. NBPGR, NRCPB, CRRI-Cuttack, DRR-Hyderabad, BHU-Varanasi, CSSRI-Karnal, IGKV-Raipur and University of Delhi-South Campus are partners in the project. So far 15,000 accessions (3000 by each centre) have been characterized for 30 descriptors (19 qualitative and 11 quantitative) during the last four years. Excellent variability for characters like leaf blade colour, stigma colour, seed coat colour and hull colour awns and panicle has been observed by various centres. From the data on characterization, a core of 1548 accessions has been developed and is being validated at different centres for certain biotic and abiotic traits. Molecular characterization of the core is in progress. NRRD for accessions has been developed and implemented on the server. It consists of 15,000 records of rice passports, 3000 (accessions) × 5 (centres) × 3 (years) characterization data and circulation data of accession among centres over the years. An NRRD website has also been developed.

Establishment of national database on tuberculosis

This database is being developed in a multi-centric mode with NJIL&OMID, Agara; NTI, Bengaluru; IOB, Bengaluru; JNU, New Delhi; IISc, Bengaluru and TRC, Chennai as partners. A prototype for centralized website for TBNet India has been developed to facilitate integration and highlight the tuberculosis initiative taken by the country. This website provides links to the participating institutes and the databases/analytical tools developed by them. The tuberculosis reference database provides information related to proteome, which includes 535 PPIs, 117 direct enzyme-substrate associations, 1780 protein subcellular locations and 467 PTMs. A host–pathogen interaction database has also been developed in which over 12,300 gene expression events in human cells upon exposure to *Mycobacterium tuberculosis* have been documented. A prototype has also been developed for the tuberculosis pathway resource, in which modulation of human signalling pathways during tuberculosis pathogenesis and infection has been documented. This database also hosts information on the pathway map for regulatory network of mycolic acid synthesis.

Development of national database on mango

Considering the importance of mango, this project was initiated in a network mode involving 10 research institutes of India, viz. CISH, Lucknow; IIHR, Bengaluru; TNAU, Coimbatore; HARP, Ranchi; BCKVV, West Bengal; FRS (APHU), Sangareddy; RFRS (DR BSKVP), Vengurle; NRCL, Muzaffarpur; AES (NAU), Paria, and ICAR Research Complex, Meghalaya. Accession details of mango (*Mangifera indica*) were collected from 17 field genebanks situated in ecologies of the country. Fifteen new field genebanks have been identified and accession information by the collaborating institution is being collected. For update of genebank information, data have been collected from the National Active Germplasm sites and major germplasm collections of the country. A new module on custodian farmers (conserving noticeable number of genetic resources) has been developed and collaborators have been provided formats for data collection. In the IPR database, 300 records were updated with details from different patent offices of different countries. Scientific literature on mangoes was digitized from 1982 to 1985 for an update of the literature base. Details of 350 phytochemicals found in mangoes were updated. Uses of these phytochemicals, relevant to human health and other aspects were collected. Two new database modules on technology base and reprint repository have been developed. In Genes and Proteins database about 300 ESTs, 310 nucleotides and 75 proteins have been updated. GIS analysis-based thematic maps were generated to develop georeferenced query modules related to variety, suitability, disease, pest and other aspects influencing mango production and trade in the country.

DeLCON consortium

The Electronic Library Consortium (DeLCON) is a significant initiative of DBT, to enhance information resources in its research institutions. It was launched in January 2009 with the ten DBT member-institutions with a large number of prominent on-line journals. It is a topical endeavour for providing access to scholarly electronic resources, including full-text and bibliographic databases in all the life science subject disciplines to the DBT institutional community across the country. It facilitates access to high-quality e-resources to research Institutions to enhance research, teaching and learning.

The access to all major e-resources was given to 10 DBT institutions in the beginning of 2009. It was extended to 17 new more DBT institutions in the second phase of extension in 2010 and further 7 members added in the third phase of extension in the 2011. In 2012, there emerged the enlarged DeLCON Consortium with 33 members. In 2013, total member-institutions of DeLCON Consortium was 34. Besides the DBT institutions, an emphasis has been given to incorporate institutions and universities (both in the State and Central Government sectors) across the States in NE India. DeLCON provides current as well as archival access to more than 1000 core peer-reviewed journals and a bibliographic database (*SCOPUS*) in different disciplines from 22 overseas publishers and aggregators.

The faculty, scientists, research scholars, students and project assistants of institutions covered under DeLCON

are the primary beneficiaries. DBT sponsored the entire expenses for its organizations for providing access to e-journals through DeLCON Consortium (http://www.delcon.gov.in).

International collaboration

Ramachandran and Arora[10] pointed out that 'There is greater thrust all over the world to evolve towards global networking, thus allowing access to information in the open domain by scientists working in different parts of the world. This effort, however, requires necessary infrastructure to be established in the areas of communications and information technology. The Dept of Biotechnology would encourage developments leading to an Asian Network amongst member countries with the help of international organisations such as the FAO, UNESCO, etc.'. The result is much international collaboration, including the recent one with Japan. In this, the bioinformatics projects mainly focus on drug development and delivery systems, senescence and cell proliferation, with the aim to understand cancers and develop novel ways for intervention using Ashwagandha. In order to promote close and effective collaboration, AIST and DBT have set up DAILAB [DBT–AIST International Laboratory for Advanced Biomedicine] at the Biomedical Research Institute of AIST in Tsukuba, Japan.

Human Resource Development

As an emerging interdisciplinary area of biotechnology, bioinformatics is progressively attempting to solve biological research problems through systematic application of information technology. Proteomics, genomics, combinatorial chemistry, nanotechnology, spectroscopy and structural and computational biology are increasingly using bioinformatics for data acquisition and analysis. In order to handle biological research problems, highly trained manpower is required to deal with molecular biology and application of software tools. High priority has been accorded by the Department to this area and several innovative educational activities to meet the present requirements, including many long-term and short-term educational programmes to address this gap have been introduced.

As Sharma[11] pointed out, 'Bioinformatics with its speedy growth is also supporting and accelerating biological research...There is of course a dearth of trained professionals in this field. Therefore, there is need and scope for further expansion and strengthening of bioinformatics network in the country'.

Kolaskar[12] writes succinctly that 'The Bioinformatics network of Department of Biotechnology has given high importance to create human resources in Bioinformatics by conducting large number of short term national, international workshops and training programs. Department of Biotechnology has also taken initiative to start post-graduate one year advanced diploma in Bioinformatics and M Sc in Bioinformatics at selected Universities. These courses are very popular and therefore many Universities, institutes and organisations have started Diploma and M Sc Bioinformatics programs without well trained faculty. As a consequence, the quality of bioinformatics students coming out of such institutes was not upto the mark. These students remained unemployed though Bioinformatics and Biotechnology companies are starving for trained human resources. The research in Bioinformatics also suffered. To circumvent these difficulties and to identify well trained students, the Department of Biotechnology has taken a major initiative of conducting BINC examination...A new model for BINC examination was created in which theoretical knowledge as well as practical knowledge in Bioinformatics was tested'.

Bhattacharya[13] sums up the effect of the programs as 'Starting of bioinformatics teaching programs, BINC exams and the computational infrastructure facilities in large number of institutions throughout the country have helped to make the tools of bioinformatics accessible to large number of experimental biologists and generated quality manpower that are needed in large numbers by both academic institutions/research laboratories and industry'.

Current status

The growth of BTISNET in terms of infrastructure, content, teaching and research has made a large impact on the bioinformatics community, within the country and globally. The increase in publications relating to bioinformatics and computational biology has been significant over the years (Figure 4)[14].

Bhan[15] points out that 'Many students and institutions have benefitted from the BTIS network. Bioinformatics

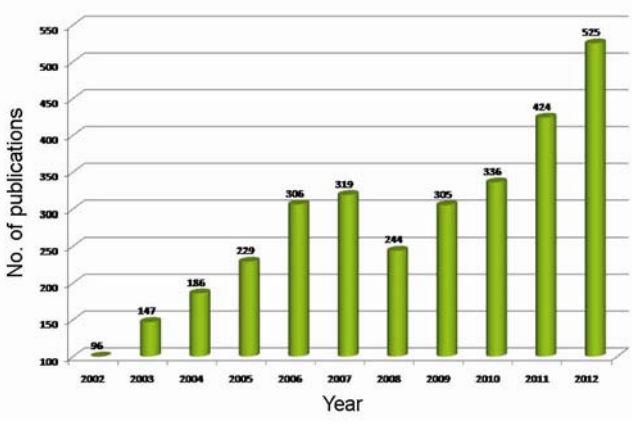

Figure 4. Year-wise publications from the BTISNET centres.

has now become an important component of research and teaching in biotechnology. Now there is a consolidation underway. There is a need for focussing on research and increased interactions between academia and industry'.

According to Madhan Mohan[16], 'India was the first country to conceptualise and establish, during 1986–87, a national distributed Bioinformatics network, which is now the largest in the world. Even as the term "Bioinformatics" was just coined, the Department of Biotechnology took a bold step in initiating the Biotechnology Information System Network (BTISNET). To start with in 1986–87 nine institutions were identified for initiating the distributed information centres and they accepted the challenge then of using computers in Life Sciences'. The organismic growth of BTISNET from nine centres in 1980s to 168 centres spread across India in geographic terms has been remarkable.

It is not only the infrastructure that is important in any activity, but the human component is what gives it life. In that aspect, BTISNET has been extremely fortunate as Arora[8] correctly puts it: 'BTIS is not only a network of centres comprising infrastructure alone, it has evolved as an entity of a glorious example of human networking with an unique spirit of collaboration and progressing on the shoulders of each other's achievements. The credit for this goes not to a single individual but to all those working in these centres who whole heartedly provided useful services to the scientific community around them. The leadership provided by DBT through all its Secretaries was also a very substantial reason for the recognition and significant progress of the BTIS over the years'.

The importance of research has been increasingly recognized. As Vijayan[4] points out: 'Synergy between experimental data and computational studies based on them are fully realised only in an ambience involving the generation of such data...Happily, India now has a culture of producing relevant experimental data, especially macromolecular structural data. Thus, good synergy has now developed between computational and experimental efforts. All these facts put together have led to a remarkable resurgence of bioinformatics and computational biology in the country. The role of BTIS in this resurgence has been very important'.

Ramachandran[17] provides the final perspective when saying that 'Biotechnology and Bioinformatics have nowadays become integral partners in research, teaching and industrial production. There has been a tremendous growth in both disciplines and in the last 25 years the Dept of Biotechnology, Govt of India, has set up a network of centres that comprise the Biotechnology Information System (BTIS)...The growth of BTIS has facilitated the thrust on research and teaching that have been kept as essential components to this grid'.

1. Krishnaswamy, S., Bioinformatics. *J. Biol. Educ*., 1990, **7**(4), 327–328.
2. http://www.btisnet.nic.in/publication.asp
3. Datta, S., Bioinformatics Policy of India (1986 to 2013) – a report, 2014; https://www.kcl.ac.uk/sspp/departments/politicaleconomy/research/biopolitics/publications/workingpapers/Working-Paper-43-(2014).pdf
4. Vijayan, M., Twenty-five years of association with BTIS. *Biobytes* (*Spec. Issue*), 2011, 22–23.
5. Srinivasan, N., Computational biology and bioinformatics: a tinge of Indian spice. *Bioinformation*, 2006, **1**(3), 105–109.
6. Seshagiri, N., Biotechnology Information System – the nucleation phase. *Biobytes* (*Spec. Issue*), 2011, 5–6.
7. Dharmalingam, K., 25 years of bioinformatics in India. *Biobytes* (*Spec. Issue*), 2011, 15–16.
8. Arora, J. R., BTIS – a journey down the memory lane: from inception to growth. *Biobytes* (*Spec. Issue*), 2011, 9–12.
9. Vijayaditya, N., BTIS. *Biobytes* (*Spec. Issue*), 2011, 13–14.
10. Ramachandran, S. and Arora, J. R., Bioinformatics – an emerging area in biotechnology. *Biobytes*, 1992, **1**(1), 3–4.
11. Sharma, Manju, Bioinformatics initiatives. *Biobytes* (*Spec. Issue*), 2011, 7–8.
12. Kolaskar, A. S., Bioinformatics network of India – reminiscences. *Biobytes* (*Spec. Issue*), 2011, 17–21.
13. Bhattacharya, A., Biotechnology Information System (BTIS). *Biobytes* (*Spec. Issue*), 2011, 24–25.
14. http://mbu.iisc.ernet.in/~mvlab/DBTDocument_Bioinformatics_in_India.pdf
15. Bhan, M. K., Foreword. *Biobytes* (*Spec. Issue*), 2011, 3.
16. Madhan Mohan, T., 25 years of India's Biotechnology Information System Network (BTISNET). *Biobytes* (*Spec. Issue*), 2011, 26–30.
17. Ramachandran, S., Biotechnolgy Information System (BTIS). *Biobytes* (*Spec. Issue*), 2011, 4.

ACKNOWLEDGEMENTS. We thank the Department of Biotechnology, Government of India for support through BTISNET, and the students, staff, faculty and scientists who have helped the bioinformatics network to grow.

doi: 10.18520/cs/v110/i4/556-561