19th International Conference on Genome Informatics (GIW-2008)
Marriott Surfers Paradise, Gold Coast, 1-3 December 2008

FORUM ON THE DEVELOPMENT OF GENOME-SCALE BIOINFORMATICS IN CHINA AND INDIA

2 December 2008

Participants:
Professor Ashok Kolaskar (India)
Professor Ming Li (Canada)
Professor Yi-xue Li (China)
Professor Shoba Ranganathan (Australia), chair

Genome-scale biology is generating unprecedented quantities of data relevant to most fundamental and applied areas of the life sciences including agriculture, fisheries and forestry, ecological sciences and biodiversity, human health and disease, and molecular and cellular biotechnology. Existing and new technologies for automated DNA sequencing, gene expression analysis, proteomics, systems analysis, imaging and high-content screening are widely available, adding locally produced data sources to the wealth of information available on-line from international data facilities such as EBI or NCBI. The quantity and complexity of these data requires researchers, institutions and countries to apply information technologies of corresponding scale and sophistication. *Bioinformatics* encompasses the management and analysis of these data, and development and provision of software tools, and immediately proximate areas of mathematics, statistics and algorithmics.

In Asia, the economically developing countries China and India have been deeply involved in bioinformatics for several decades. With the emergence of these countries as economic superpowers, this is an opportune moment to discuss where these countries are currently poised; what challenges they face; what strategies they have developed or are developing in research, development, infrastructure, education and training for bioinformatics; what the prognosis might be for their future development in this important area; what opportunities are available for international coordination and collaboration; and not least, what broader lessons can be learnt from their development and success in genome informatics.

"Rising bioinformatic stars"

CHINA

The article "Bioinformatics in China: a personal perspective" by Liping Wei and Jun Yu (published in *PLoS Computational Biology* 4(4):e1000020, April 2008) is absolutely essential reading in this context. The authors present statistics on the growth of publications, and information on online databases, webtools and software developed and maintained in China and on bioinformatics training programs offered by institutions in China.

Specific issues discussed in this workshop included:

The recruitment of overseas Chinese back to China
Diversity and flexibility of arrangements, *e.g.* appointment as guest professors
Wide range of salaries, start-up packages across institutions
The possibility of "reverse culture shock"

R&D investment plans
Multiple national Ministries and bodies are involved
Ministry of Science & Technology: 2001-2010 US$3.8B in biological & biomedical Research
Phase I (2001-05): main fields are bioengineering, gene manipulation, bioinformation technologies, biomedical technologies
Phase II (2006-10): main fields are industrial technologies, gene manipulation, biomedical technologies, bioinformation & biocomputing technologies(*)
(*) National "863" program in bioinformatics, 2006-2010
Increasing focus on neuro-informatics

In addition, there is ongoing investment by other national ministries (Health, Education) and by local governments.

New national programs (November 2008):
New drugs (US$ 1B) – note that China doesn't have a large pharmaceutical industry, although international pharmaceutical companies have recently located R&D facilities in China and graduates are finding jobs there.
Infectious disease (US$ 800M)
3%-5% of these funds set aside for bioinformatics
Funds must be spent in next two years

Heavy investment nationally in data production (genome sequencing, metagenomics, microbial genomics, functional genomics, proteomics, metabolomics).

Job prospects for bioinformatics graduates in China are good. Some graduates go overseas, *e.g.* for further training. Many others are finding jobs within China.

INDIA

Much information is available at the following two websites:
http://www.btisnet.gov.in (Biotechnology Information System Network)
http://bioinfo.ernet.in (Bioinformatics Centre at University of Pune)

Bioinformatics in India started in 1981
Bioinformatics task force – 1983
Recommends a distributed network approach: BioGrid
Triggered internet connectivity for scientific research across India
More recently, development of a three-tiered system:
Centres of Excellence (6 nationally)
Distributed information centres (10) and sub-centres (51)
Infrastructure facilities (76)

Main components of national strategy (for further detail, see below):
Advanced research
Development of human resources
Academia-industry interface
International linkages

Advanced Research

About 1050 bioinformatics research papers, plus >3000 that use bioinformatics

Development of human resources

Goal is to train >2000 bioinformaticians per year at all levels

Central government support for 7 undergraduate, 28 postgraduate, 12 certificate / diploma, 9 PhD and 9 other training programs; many other programs funded by private institutions. Quality, especially of the latter, is variable. Hence BINC, Bioinformatics National Certification Examination. The BINC is open to anyone, including self-taught candidates, but is (intentionally) very difficult and the pass rate is "very, very low". The BINC is implemented in three steps: short-answer questions, long-answer questions, and a practical; >50% correct answers is needed to proceed to each successive step. Website: http://bioinfo.ernet.in/BINC

Discussions are underway with Malaysia about adopting the BINC there.

BioGrid as a "virtual classroom"

Market is (so far) imperfect in matching graduates to jobs. There is difficulty in finding qualified teachers. Some graduates have trouble finding jobs, although these are mostly from the programs of questionable quality. There was much discussion in the workshop about how to match graduates to jobs lies: Professor Kolaskar advised that the principal approach must be through improvement in quality.

Academia-industry interface

Since 1982, central government support of the bioinformation industry has been about US$ 25M in total, of which >$5M is being spent this year.

There is a very active ICT industry in India, including large and small companies in bioinformatics (*e.g.* Molecular Connections, already mentioned in the Innovation and Commercialisation forum).

Note that there is a large drug industry in India. This has attracted bioinformatics and software companies from overseas, *e.g.* Accelrys, to invest in India.

International linkages

The main partners in bioinformatics at this point are Japan, Australia, and others in the Asia-Pacific region; and Europe, particularly Italy.

IBIN – Indian Bioresource Information Network

In three regional languages

Pictorial, common-language focus to facilitate broad public impact

Development of databases in highly relevant: rice, tuberculosis

The Indian government and institutions have noted China's success in attracting overseas Chinese back to China, and are considering how to emulate this vis-à-vis overseas Indians, whether for permanent relocation or on a part-time basis. One possibility being discussed is that overseas Indians might be made eligible to apply for government research grants.