

Indo-US researchers to build next-generation scalable computers for genomics in India*

The Tata Institute of Fundamental Research (TIFR), Mumbai and its Bangalore-based centre, the International Centre for Theoretical Sciences (ICTS) along with the University of Illinois at Urbana-Champaign (UIUC), USA, projected a collaborative research programme to develop a next-generation computer to enable biologists analyse huge genomics sequence data and foster a two-way relationship between computing and genomics.

The meeting organized by the ICTS at the Indian Institute of Science (IISc), Bangalore was directed by R. K. Shyamasundar (TIFR), Spenta Wadia (ICTS-TIFR), Gene Robinson and Ravishankar Iyer (UIUC) and Vijay Chandru (Strand Life Sciences, Bangalore).

Computing large sets of data is a challenge that biologists worldwide face today. The discussion-cum-strategy formulation meeting of the Department of Electronics and Information Technology (DEITY) Government of India was inaugurated by P. Balaram (Director, IISc). It was attended by leading academicians in the field of computer science and biology, industrial researchers, leading educationists from India along with 16 experts from University of Illinois, including their Associate Chancellor Pradeep Khanna.

'The genomic data is doubling every 1.5 years. The current methods are unable to handle large datasets. Therefore, a new technology framework is required', said Gene E. Robinson, an eminent entomologist and director of the Institute for Genomic Biology, UIUC. The current goal in genomics research in general, according to Robinson, is to understand

all functional units of genome, evolution and diversity of life and what knowledge could be derived from the genomic sequences. He felt that there was a need to integrate 'omic' data types to solve big problems. The goal of genomic analysis is to understand genotype-to-phenotype relationship and predict gene function to understand how a gene behaves in a disease.

Data storage is a major challenge related to computational science. Ravishankar Iyer drew attention to the increasing storage cost, which is getting much higher than the cost of genome sequencing itself.

Sam Pitroda, Advisor to the Prime Minister on Public Infrastructures, through a live video conference from California dwelt upon the need to address problems of the poor by developing low-cost health solutions and integrating telecommunication. He viewed supercomputing as the way forward to address societal problems, primarily health and food security.

A proposal has been made to design the next-generation application-specific (genomics) scalable computing test-bed in close liaison with industry. Programmes such as these are expertise-driven. Such a collaboration is extremely important to India as it will have a huge impact on scaling-up biological discoveries, arrive at new computing architectures, develop an affordable sustainable healthcare framework and for building-up a friendly innovative ecosystem for bio-tech and IT industries.

K. VijayRaghavan (Secretary, Department of Biotechnology, New Delhi) said, 'the main challenges are in biodiversity and ecology, biomedical research and agriculture. We have had attempts to sequence native crops, but sequences and functional aspects seem to be the thing. A lot of data in cancer genomics is com-

ing out. Integrating such biology ventures should be important in training and facilitating interaction between computer scientists and biologists.'

J. Satyanarayana (Secretary, DEITY), through a live video conference highlighted the importance of building innovative computing solutions involving both hardware and software in areas of key importance for scientific discoveries and public welfare like healthcare, agriculture, etc. He reiterated treating the discussion meeting as a strategy formulation meeting of DEITY, which shows its commitment. T. Ramasami (Secretary, Department of Science and Technology, New Delhi) in a message to the organizers expressed the need for such joint ventures to push science boundaries and at the same time derive societal benefits.

During the meeting, it was recognized that though several tools were developed by the computer scientists, there was a need to further develop algorithms which would enable the biologists to compute data efficiently. In order to bridge the gap, an interdisciplinary approach is required. It has therefore been proposed to initiate an interdisciplinary graduate and undergraduate training programme in the areas of computing and biology. As a first step, Spenta Wadia announced starting of a winter or summer school during the next four years jointly with the University of Illinois, which will help in bringing students and researchers to work together and exchange knowledge by hosting conferences and mini courses.

The proposal envisages the use of computing systems similar to CompGen at the University of Illinois. Such a system is to be built at ICTS-TIFR for use by researchers from both the countries.

Megha Prakash

*Based on a discussion meeting, 'Challenges in Genomics and Computing' held at Indian Institute of Science, Bangalore during 22-24 July 2013.