

Bioinformatics today and IMTECH's role

Dr G P S Raghava

BE it popular media or scientific forum, Bioinformatics (BI) today is the hottest topic in the world of biotechnology of recent years. With several reports in the media on new ventures and tie-ups in bioinformatics, one finds proponents of "Godfathers" of Bioinformatics arising from unexpected quarters on a regular basis, though they do not profess even a remote link with the science of BI.

On one side we have very limited number of experts in BI, and on the other we find 'n' number of Biotechnology degree /diploma courses of questionable standards offered by colleges, training centers and universities across the country. We have new organizations in BI coming up; we also have record of the ones that got closed down during the last couple of years.

However, there are ample reports indicating the tremendous scope of BI and also the vast job potential it offers. As of now, the number of trained people is much less than the

actual demand. Then how come we find that our students who have completed their degree/diploma/course in BI are not getting jobs? Who is to be blamed if these students feel that they have been cheated in the name of BI?

It's true that there is a gap between demand and availability. The standard/expertise of students is much lower than desired. While the demand is for people who have proved their expertise in terms of contribution (scientific papers in academics), we have fresh graduates coming out of these BI institutes/organizations. Secondly BI is Tough: The students are jumping into this field without realizing its intricacies. A thorough knowledge of both physical and biological sciences is not a simple achievement. Lack of Standards is another major problem with the institutions carrying out the courses without proper infrastructure and expert teaching staff. There is also an element of misconception among the students who feel that a degree or diploma per se can ensure a job in BI field.

Along with the regular courses, the students should also get acquainted with software development, data analysis, database development, prediction methods or web server development.

As a scientist who has been actively involved in the field of biocomputing for the last 12 years, I have seen the up and downs of BI. To my knowledge, biocomputing is core of BI, where India has contributed significantly. The Indian Institute of Science particularly Molecular Biophysics Unit has done lot of work in structure Proteomics for the last 50 years and, I would call Prof. G. N Ramachandran as the godfather of BI in India.

The infrastructure development in BI began with the efforts of Department of Biotechnology (DBT), Government of India which started a project called Distributed Information Centres (DICs). The contribution of these centres in term of research, service and human resource development has been tremendous, with BI centers at Pune University, Pune; Madurai Kamraj University, Madurai; IISc Bangalore; JNU Delhi; DBT Delhi; IMTECH, Chandigarh doing extremely well.

BI at IMTECH: Since its establishment in 1987, Bioinformatics centre at IMTECH has shown consistent growth in different areas of bioinformatics including carrying out peer reviewed research or providing services to overseas clients. The centre has developed databases and softwares both for in house use and for scientists world over. BIC has developed a number of web servers based on peer-reviewed research carried out at the bioinformatics centre (<http://imtech.res.in/raghava/www.html>). BIC staff has published more than 30 research papers in reputed international journals and more than 25 abstract/poster/papers presented in national and international conferences (<http://imtech.res.in/raghava/pub.html>). Our servers are getting more than 5000 hits per day.

Structural Proteomics: We have developed number of programs/web servers for predicting secondary structure of proteins. Recently, our regular sec-

ondary structure (Helix & Sheet) prediction method was classified in category of highly accurate methods in the world (CASP5, CAFASP3 & EVA). As of today, this centre is the only one in Asia which is competing successfully at the international level under CASP, CAFASP and EVA programmes, 'Olympics of proteins structure prediction'. BIC also developed highly accurate methods for predicting tight turns in protein that includes a, b, and g turns. There are more than 10 web servers developed at this centre which are serving the community in field of protein structure prediction or protein modelling.

Genome Wide Search and Genome Annotation: Centre has established the Genome Wide BLAST and FASTA searching facility with integrated sequence analysing tools. It is a world-wide facility for sequence similarity search over large number of databases including genome and proteome. This unique facility allows one to search the potential drug targets. As of today we have genomes and annotated proteomes of all prokaryote and Eukaryote (including human genome) sequenced so far. In addition we have other protein/nucleotide databases. These are powerful web based tools for identifying the drug/vaccine targets. We developed number of genome annotation tools that includes prediction of gene coding and repeat regions in genomes. Best of our knowledge this is the only place in India who have it. There are only a few centres in world which provide such a facility.

Biological Databases: Centre has developed following databases in biology i) MHCBN: A curated database of MHC binders/non-binders & T-Cell epitopes (~23000 entries); ii) Public Domain Software in biology (more than 500 software); iii) Free software for general purpose (more than 1500 entries); and iv) BCIPEP: A curated database of antigenic, immunogenic and non-immunogenic peptides including B cell epitopes (more than 2300 entries). These curated databases have been created from literature and public domain resources and these are recognised by world wide community. The MHCBN and BCIPEP databases are distributed by European Bioinformatics Institute (EBI) Cambridge, only database from

India. Recently, it has been shown by number of groups that BI servers developed at our group are highly accurate.

Mirror sites: Our centre had expertise in field of establishment of mirror sites. We have mirror following websites i) Biological Databases (e.g. genomes, proteomes, PDB, SWISSPROT); ii) Biological software (EBI & Indian Univ.); and iii) General software (GNU, PostgreSQL & SUN Freeware).

Immunoinformatics: We have developed servers such as Propred and Propred1 which allows user to predict the MHC binders for large number of MHC class I and Class II alleles. These servers allow predicting the potential sub-unit vaccine candidates which can work effectively in large number of populations. Our group is considered in category of top groups in Immunoinformatics and most of the sites and papers cite our servers/group. Recently, a research group from UK demonstrate experimentally that our Propred server can identify MHC binders in antigens with high accuracy (>73%).

Human Resource Development: Our centre regularly conducts training and workshops in the area of BI at national level, every year. We have conducted a two weeks training programme for a few personnel of Asia Privacy, South Korea. Recently, We have conducted training for BI trainers for example we trained teachers of DOEACC centers in BI. This shows that our expertise in bioinformatics is being recognised at national and International levels.

BI at IMTECH have powerful infrastructure that include computer hardware (>10 Digital alpha servers; > 5 SGI workstations; SUN server, MAC and plenty of INTEL based PC); Network (Full LAN; Two 24hr Internet Connections; Domain Name reservers; Proxy Servers etc) and Software. We have developed number of stand alone software for BI which are available to public. Almost all web servers are free for public. Some of the web servers have been mirror at University of Arkansas for Medical Sciences (UAMS). So the user can access our sites from <http://bioinformatics.uams.edu/>.

— The author is Scientist & Head, Bioinformatics Centre, Institute of Microbial Technology

Multiply Your Lab Output ...

Automate Your Lab Information With **Caliber LIMS 2.0**

A comprehensive, 21 CFR Part 11 compliant, validatable Laboratory Information Management System that makes Sample Testing easy, fast and authentic.

Highlights

- 3-tier web-based technology
- Easy-to-master workflow
- Add-on modules to cover entire lab functions
- Seamless integration with SAP / ERP
- Instrument Connectivity and Controlled Label Printing
- Fully networked and scalable
- Fast and easy customization, quick ROI
- Different configurations to suit any size lab / budget
- Committed and trustworthy support
- Dr. Reddy's Laboratories (Multiple locations) and Aurobindo Pharma among our valued customers



Email today for a catalogue.

Order for a Tutorial CD, to see for yourself the capability of Caliber LIMS.

Caliber LIMS
Programmed To Deliver



219, 215 Bhanu Enclave, Model Colony, Hyderabad-500038
Tel : +91 (40) 23811322, Fax : +91 (40) 23817972
E-mail: info@caliberindia.com Visit us at: www.caliberindia.com