Role of Bioinformatics in Developing Country: Bangladesh

Md. Mirazul Islam
Department of Biochemistry and Molecular Biology, University of Dhaka, Dhaka-1000, Bangladesh
Email ID: mirazul.du.bc@gmail.com

Abstract — This paper is a systematic analysis of bioinformatics base research and education for developing countries. This gives an overview of current bioinformatics education in Bangladesh, from the perspective of the practitioner. Bioinformatics can help pharmaceutical and biotechnological industry of Bangladesh to keep pace with the world. Equal access of biological data can give wide range of research areas in molecular level within limited resource and provide opportunity to healthcare development as well as personalized medicine. Integrative action on bioinformatics will bring a crucial change on biological research and development in Bangladesh. Some potential steps are shown in this paper that have done by Bangladeshi scientists.

Keyword — Bioinformatics, Computational Biology, Medical Informatics, Pharmaceutical Fields, Personalized Medicine.

1. INTRODUCTION

Bioinformatics is a young field of biological and computational collaborative research for optimum use of biological data with the help of computational tools for utmost and quick outcome in life sciences. It is a great opportunity for a developing country like Bangladesh to carry on research works within a limited resource. In spite of being young, the science of bioinformatics exhibits tremendous potential for playing a major role in the future development of science and technology. It is expected that bioinformatics will especially contribute in the future as the leading edge in biomedicine to pharmaceutical companies by appropriately yielding a greater quantity of lead drugs for therapy [1]. Ongoing development of information and communication technology (ICT) can contribute biological data analysis or bioinformatics in Bangladesh. Now it is important to design bioinformatics study policy with the integration of biology and computer science to keep pace with the world.

2. BIOINFORMATICS AND COMPUTATIONAL BIOLOGY

The term bioinformatics arose in the beginning of the 1990s, when large amounts of experimental data were produced by sequencing projects, particularly the Human Genome Project [2]. Computational Biology/Bioinformatics is the application of computer sciences and associated technologies to answer the questions of Molecular Biology, about the mysteries of life and evolution. These two fields are overlapped with each other. Bioinformatics refers to the creation and advancement of algorithms, computational and statistical techniques, and theory to solve formal and practical problems arising from the management and analysis of biological data. Computational biology refers to the hypothesis-driven investigation of specific biological problems using computer, carried out with experimental or simulated data, with the primary goal of discovery and the advancement of biological knowledge. More simply bioinformatics is concerned with the information, while computational biology is concerned with the hypotheses [3]. Developing countries in the Asia-Pacific region are just moving into this new field of information-based life science research. However, the computational infrastructure and biological research in these countries are still at a basic level compared to that in developed countries.

3. BIOINFORMATICS AND MEDICAL INFORMATICS

Bioinformatics and Medical informatics are two research fields that have become mature in the past 20 years. Both of them share the common goal to provide new ICT based solutions to biomedical research and contribute to the treatment and cure of the molecular bases of diseases as well as biotechnology. Bioinformatics provides solutions to scientists doing biological research whereas medical informatics fulfils the demands from clinical personnel, for practitioners and scientists in medical research [4]. Both domains have a common ICT infrastructure (e.g. electronic databases, terminologies, network bandwidths) and scientists in both domains adopt solutions from the other domain if they work in an interdisciplinary environment (e.g. biological research done in a clinical environment) [5]. Medical informatics is a longitudinal electronic records that provide substantial opportunities to organize and manage clinical data in ways that can potentially improve preventive health care, the management of chronic illness, and the financial health of primary care practices. Simultaneous development of bioinformatics and medical informatics in developing country is an urgent need for secondary and tertiary patient care and productive out put on biological research fields to solve the upcoming questions (climate change, agriculture production, salinity, drought, epidemic diseases etc). In all areas of biological and medical research, the role of the computer has dramatically enhanced in the last five to ten-year-period [6].

4. MAJOR RESEARCH AREAS

Bioinformatics has a wide range of molecular data analysis through algorithms development and statistical data interpretations methodology (Figure-1). Bioinformatics is a dynamic field, so it's research areas exponentially increasing day by day. Computational Drug Design, Sequence Analysis, Genome Annotations, Computational Evolutionary Biology, Measuring Biodiversity, Analysis of Gene Expression, Analysis of Regulation, Analysis of Protein Expression, Analysis of Mutations in Cancer, Prediction of Protein and RNA Structures, TSS Prediction, Comparative Genomics, Modeling of Biol. Systems, High throughput Image Analysis, Protein-Protein Docking, and Data Mining [7]. As a developing country, Bangladesh faces different endemic and epidemic diseases, where bioinformatics can give a cure computational vaccine design, development and mass production.

The applications of systems biology approaches have greatly increased in the past decade largely as a consequence of the human genome project and technological advances in genomics and proteomics. Nowadays, an understanding of molecular systems is the main topic in experimental as well as computational system biology. Systems approaches have been used in the medical and pharmaceutical area for diagnostic purposes and target identification. During this same period, the use of mode of action for risk assessment has been increasing and there is a need for quantitative risk assessments on an ever-growing number of environmental chemicals [8].
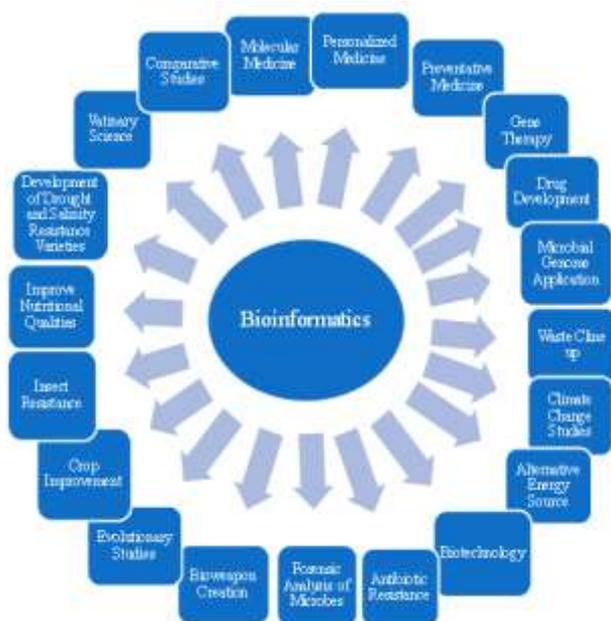
Fig-1: Different Areas of Bioinformatics based research.

5. BIOINFORMATICS BASED RESEARCH IN BANGLADESH

Bangladesh is a young and prospective country on biological research spatially on bioinformatics. Still now no university in Bangladesh has well developed multi-disciplinary bioinformatics course that incorporates many of the public-domain databases, research tools, peer-review journals and research organizations. Some research groups doing well as per need but no collaborative and integrative activity show on bioinformatics research without few exceptions. But some positive and optimistic steps give a strong sign to change the situation very soon. Modern biology and related sciences are increasingly becoming dependent on this new technology. Research and development in this area do not require huge funding to set up a world class laboratory. All it requires is high speed Internet connectivity with a few powerful computers and a team of enthusiastic talented people [1]. The vision of equal and free access to the human genome project data and associated research tools provides an unprecedented research opportunity both for developed and developing countries [9]. Furthermore, the recent movement to deliver scientific publications freely and free HINARI access for developing country has brought an enormous change on biological research all over the world.

Bangladesh needs to form a national research council, under which a group of innovative cross-link scientist will carry out bioinformatics based research. As many Bangladeshi scientists are well placed in various prestigious institutions almost all over the world, could capitalize this opportunity and develop a world class bioinformatics institute in Bangladesh. Bangladesh has a plenty of talented young scientists hunting for opportunities for overseas employment. This institute can train hundreds of students with high tech cutting-edge technologies virtually from anywhere of the world with a minimum cost. Although, conducting world class research and publishing the findings in the top tier journals and conferences are the fundamental driving force of the institute but it is planning to offer the following services to generate funding for long term survival. A world class and well suited bioinformatics course have to develop for graduate and undergraduate students that positively influence their perceptions of how they might contribute to biology. Student-based discoveries are the heart of successful learning. In particular, students should develop thinking skills that improve comprehension, application, analysis, synthesis, and evaluation [10].

6. BIOINFORMATICS IN PHARMACEUTICAL FIELDS

More than a third of the world's population has no access to essential drugs. More than half of these groups of people live in the poorest regions of Africa and Asia. Several factors determine the accessibility of drugs in

developing countries. Hardly any medicines for tropical diseases are being developed, but even existing drugs are often not available to the patients who need them [11]. Among the 49 countries classified as an LDC (of which 32 are WTO members), Bangladesh is the only country with adequate pharmaceutical manufacturing capability and is nearly self-sufficient in pharmaceuticals [12]. Bangladesh's pharmaceutical industry now provides to 96 percent of the country's pharmaceutical needs. It is worth noting that Bangladesh now exports a wide range of pharmaceutical products to 72 countries in Asia, Africa and Europe and in 2006-2007 total exports were US$28.12 million with a growth rate of some 47 percent [13].

The Agreement on Trade Related Aspects of Intellectual Property Rights (TRIPS) which is binding on all members of the World Trade Organization (WTO) aims at establishing strong minimum standards for intellectual property rights (IPRs) including patent protection for pharmaceuticals.

To address the concerns of Doha Declaration, that adopted by the WTO Ministerial Conference, in Doha on 14 November 2001 further extended the transitional period for LDCs to introduce pharmaceutical patent protection, so Bangladesh is required to have a TRIPS compliant patent regime for the pharmaceutical sector from1 January 2016. Developing and Least Developed Countries (LDCs) are apprehensive of strong patent protection on the basis that such patent protection may be harmful to their growing pharmaceutical industries and for the access to affordable medicines to their populations.

What will be the alternative for Bangladesh after 1 January 2016? Bioinformatics would be the alternative solution for pharmaceutical industries. Computational drug design for tropical diseases under research and development (R&D) section of Pharmaceutical industries may be a low cost approach and optimum solution. One study found that a hypothetical pharmacogenetically driven clinical trial of the anticoagulant warfarin could save up to 60% of the cost and reduce possible adverse events [14].

Pharmaceutical companies are achieving increased research efficiency by the introduction of new approaches to the design, synthesis, screening and optimization of drug candidates. Most pharmaceutical development addresses medical problems with a 'one drug fits all' approach. Genetic variation has been shown to influence drug selection, dosing and adverse events[15], and the therapeutic benefits of taking a genetically tailored approach to drug development is now recognized [16]. ICT is an important support function for all of those activities and there are certain functions and operations that cannot be performed without ICT. Long term investment on R&D section is a crucial step for keeping it up the industries. Pharmacogenomics is an involvement of bioinformatics on pharmaceutical fields. This is the important advances in understanding the genetic basis of drug response. Over the past 10 years, new technologies have shifted research from applicant gene pharmacogenetics to phenotype-based pharmacogenomics with a consequent explosion of data [17].TRIPS make bioinformatics as an essential tool for pharmaceutical fields of Bangladesh as well as other developing countries.

7. BIOINFORMATICS IN PERSONALIZED MEDICINE

After 'Human Genome Project' the concept of individual genome based medical treatment has got popular aspects. Personalized medicine is a new challenge for bioinformatics. Now a whole genome or a few dozen exones can be sequenced in <2 weeks with an error rate of ~1 error per 100 kb [18]. Sequencing technologies are becoming affordable and are replacing the microarray-based genotyping methods, which were limited to interrogating regions of known variation [19]. In the coming years, the bioinformatics world will be inundated with individual genomic data. This flood of data introduces significant challenges that the bioinformatics community needs to address. Great opportunities for bioinformatics to integrate with the electronic medical record (EMR) [20]. For example, the BioBank system at Vanderbilt links patient DNA with a deidentified EMRs to provide a rich research database for additional translational research in disease–gene and drug–gene associations [21]. The ultimate challenge for this research is to apply the results for improved patient care. Much of this research has yet to be translated to the clinic. Many physicians are unprepared to incorporate personal genetic testing into their practice and it is unclear how to best apply research results to improve patient care [22]. In fact, personalized medicine is a future generation's expectation in developing countries.

8. APPLICATION OF BIOINFORMATICS IN BIOTECHNOLOGY

Bangladesh is an overpopulated country with high demand for food. Due to scarcity of land, agricultural output is experiencing a downward trend in growth. Biotechnology is a scientific field that needs to be explored to grow the agricultural production. In spite of being agricultural country; Bangladesh regularly faces agricultural problems due to climate change. Bioinformatics would be able to adapt this change by genetic modification of plants. Specific gene identification and modification using bioinformatics tool, helps to express particular characteristics of plants.

Drought, Salinity registrant and more productive rice is urgent need for Bangladesh. Biotechnology is an aspect of biological science providing new strategies to solve emerging problems of agriculture, forensic medicine, human health, industrial production.

9. SOME POSITIVE STEPS IN BANGLADESH

Bangladeshi scientists have successfully decoded the fundamental genome sequence of jute, opening up a new outlook in the development of the golden fibers. A team called the "Swapna Jaatra", or the "Dream Journey", has done this great work by young, energetic and dedicated graduates of Bangladesh. 'Data Soft' gives technical and computational help to do it. This would give a concern about possibility and potentiality of bioinformatics in Bangladesh. The best quality jute is grown in Bangladesh and its genome sequencing is a symbol of national identity. Now pipeline application is bioinformatics knowledge base fiber development and this trick is to keep on dreaming [23].

Another step is to form an organization named Bio-Bio-1 by some passionate individuals in late 2008 to learn bioinformatics for making some senses from the mystery of life. Their aim is to spread out the excitement of R&D to the young and the like-minded through several programs. Bio-Bio-1 arranged a seminar on 'Bioinformatics in Bangladesh' to encourage people and to raising public concern on bioinformatics. They carry out a weekly study session to popularize bioinformatics [24].

Some individual people are also trying to make understanding on bioinformatics and associated tools to carry out biological research in Bangladesh. The fact is to connect different groups and to bring them under same shade of Bioinformatics Research Council.

10. RESOURCE AND FUNDING

Research and development depends on long term investment or funding. Bangladesh has shortage of resource as well as funding. But bioinformatics base research does not need as huge funding as other fields to set up a world class laboratory. All it requires is high speed internet connectivity with a few powerful computers and a collaborative multi-dimensional team of talented people. Cloud-based computing can provide large space to store billions of bioinformatics data and many data freely available from different public websites. Government have to influence bioinformatics as a field of study in its own to maintain a significant shortage of

bioinformatics scientists in academic and industrial research [25]. Public Private Partnership would give a positive feedback on bioinformatics research and publishing the data in high impact factor journals. This is crucial driving force of the industrial outcome to their business requirement. Local and Global investment could promote bioinformatics software development and its usages related to latest tools and technologies (Table 1&2). A well develop bioinformatics course through online resources provide students with exciting entrée into the new fields of genomics, proteomics, and bioinformatics [9].Bangladesh has many reputed pharmaceutical industries and research organizations, that can invest handsome amount on bioinformatics research and development [26] .

11. COORDINATION

Bioinformatics is a collaborative research work where coordination is vital for success. A interlink can be easily established with biological and medical laboratory of home and abroad. This coordination could offer experienced and qualified bioinformaticians to the industry for local and international job markets. Faculty needs to integrate their research and teaching with students. If faculty improves their teaching and research simultaneously, the next generation of faculty will believe that teaching and research are mutually beneficial rather than mutually exclusive.

12. CONCLUSION

The field of biology has a great possibility to contribute to the development of Bangladesh. Twenty First century is the century of biology. Contribution to the development basically depends on biological research. Bioinformatics is a new era of biological research both for developed and developing countries. Researchers have to take some update initiative to spread bioinformatics all over the country. The governmental support and funding is crucial for this type of initiative. Considering the present political and economic condition of Bangladesh, private initiative is vital to pushing the science ahead with the rest of the world. Now is the best time to start the race. Together it's possible to achieve more than imagine.

REFERENCE

[1] Abdur Rahman Sikder. Bioinformatics in Bangladesh: An Optimistic View.Plant Tissue Cult. & Biotech. **18**(1): 85-88, 2008 (June).

[2] Human Genome Project Information. http://www.ornl. gov/sci/techresources/ Human_Genome/home.shtml (11 February 2008, date last accessed).

[3] Barnes MR and Gray IC eds**.** (2003). Bioinformatics for Geneticists, 1st edition. Wiley, 2003. ISBN 0-470-84394-2

[4] Luscombe NM, Greenbaum D, Gerstein M: What is bioinformatics? A proposed definition and overview of the field**.** *Methods Inf Med* 2001, **40(4):**346-58

[5] Friend SH: How DNA microarrays and expression profiling will affect clinical practice**.** *BMJ* 1999, **319:**1306-1307

[6] Baldi P and Brunak S (2001). Bioinformatics: The Machine Learning Approach, 2nd edition. MIT Press, 2001. ISBN 0-262-02506-X

[7] Aluru Srinivas ed**.** (2006). Handbook of Computational Molecular Biology. Chapman & Hall/Crc, 2006. ISBN 1584884061 (Chapman & Hall/Crc Computer and Information Science Series).

[8] Stephen W. Edwards1 and R. Julian Preston. Systems Biology and Mode of Action Based Risk Assessment. TOXICOLOGICAL SCIENCES 106(2), 312–318 (2008).

[9] A. Malcolm Campbell. Public Access for Teaching Genomics, Proteomics, and Bioinformatics**.** Cell Biology Education. Vol. 2, 98–111, Summer 2003

[10] Allen, D., and Tanner, K. (2002). Approaches to cell biology teaching: Questions about questions. *Cell Biol. Educ. 1*, 63–67

[11] Sigrid Sterckx. Patents and Access to Drugs in Developing Countries: An Ethical Analsis. *Developing World Bioethics* ISSN 1471-8731 (print); 1471-8847 (online). *Volume 4 Number 1 2004*

[12] Mohammad Abu Yusuf and Qamrul Alam, 'WTO TRIPS Agreement-Current state of Pharmaceutical Industry and Policy Options for Bangladesh' (2008) 1(1) *Journal of International Business Research* 12, 12-15

[13] Directorate of Drug Administration, Bangladesh, (10 June 2010). http://www.ddabd.org/exporting_country.htm

[14] Ohashi,W. and Tanaka,H. (2010) Benefits of pharmacogenomics in drug developmentearlier launch of drugs and less adverse events. *J. Med. Syst.*, **34**, 701–707

[15] Giacomini,K.M. *et al.* (2007) The pharmacogenetics research network: from SNP discovery to clinical drug response. *Clin. Pharmacol. Ther.*, **81**, 328–345

[16] Foot,E. *et al.* (2010) Pharmacogenetics–pivotal to the future of the biopharmaceutical industry. *Drug Discov. Today*, **15**, 325–327

[17] Caroline F Thorn *et al*. Pharmacogenomics and bioinformatics: PharmGKB Pharmacogenomics. 2010 April; 11(4): 501–505. doi:10.2217/pgs.10.15

[18] Drmanac,R. *et al.* (2010) Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science*, **327**, 78–81

[19] Ng,S.B. *et al.* (2010) Exome sequencing identifies the cause of a mendelian disorder. *Nat. Genet.*, **42**, 30–35

[20] Busis,N.A. (2010) How can I choose the best electronic health record system for my practice? *Neurology*, **75**, S60–S64

[21] Denny,J.C. *et al.* (2010) PheWAS: demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations. *Bioinformatics*, **26**, 1205–1210

[22] McGuire,A.L. and Burke,W. (2008) An unwelcome side effect of direct-to-consumer personal genome testing: raiding the medical commons. *JAMA*, **300**, 2669–2671

[23] Hasina Khan. Prospect and application of genome sequencing with special reference to jute; Harun K. M. Yusuf Memorial Lecture, University of Dhaka. 4 December 2010

[24] http://bio-bio-1.wikispaces.com. 2 May 2012

[25] Ina Koch and Georg Fuellen. A review of bioinformatics education in Germany: *Briefing in Bioinformatics*. VOL 9. NO 3.March 1, 2008

[26] Mohammad M. Azam, Kristy Richardson. Pharmaceutical Patent Protection and Trips Challenges for Bangladesh: An Appraisal of Bangladesh's Patent Office and Department of Drug Administration. *Bond Law Review*: Vol. 22: Iss. 2, Article 1

Table-1: Online Database

CONFLICT OF INTEREST
None Declared

Resource	Description	Data	Web Address
GenBank	Searchable database of protein and DNA sequences, based in the U.S	135440924 sequences	http://www.ncbi.nlm.nih.gov/Genbank/index.html
Protein Data Bank	Respository of all protein structures that is searchable by protein names and key words	82347 structures	http://www.rcsb.org/pdb/
DIP: Database	Searchable database that shows protein–protein	24569 portions	http://dip.doe-mbi.ucla.edu

of Interacting Proteins	interactions based on experimental and computational information		
Osprey	Nice graphical representation of physical and genetic biological interactions; requires special plug-in/ can function as a stand-alone program	50,000 interactions	http://biodata.mshri.on.ca/osprey/servlet/Index
ArrayExpress	Searchable British public repository for microarray-based gene expression data	30620 expressions	www.ebi.ac.uk/arrayexpress/
The SNP Consortium	A public/private database containing about 1.8 million SNPs (single nucleotide polymorphisms)	600000 SNPs	http://snp.cshl.org/
PubMed Central	U.S. National Library of Medicine's digital archive of life sciences journal	2.4 millions articles	http://www.pubmedcentral.nih.gov
HighWire Press	Largest repository of free online journals with varying periods of delay before freely available	2131844 articles	http://highwire.stanford.edu/
MIPS	Searchable database of the molecular structure and functional network of *Saccharomyces cerevisiae*		http://mips.gsf.de/proj/yeast/CYGD/db/index.html
Alternative Splicing ASAP	Searchable database with graphical presentations of alternative splicing		http://www.bioinformatics.ucla.edu/ASAP/
MHC Haplotype Project	A work in progress that will provide the complete genomic sequences of at least eight different HLA-homozygous typing haplotypes		http://www.sanger.ac.uk/HGP/Chr6/MHC
Ensembl	European repository for many genome projects. Very user-friendly and good graphic displays		http://www.ensembl.org/
GeneOntology	A consortium that is unifying terminology used to describe genes and their cellular roles in all species		http://www.geneontology.org/
GEO: Gene Expression Omnibus	Gene expression and hybridization array data for searching and visualizing results repository with methods		http://www.ncbi.nlm.nih.gov/geo

Table-2: Some tools for bioinformatics

Bioinformatics Research Area	Tool (Application)	Reference.
Sequence alignment	BLAST	http://blast.ncbi.nlm.nih.gov/Blast.cgi
	CS-BLAST	ftp://toolkit.lmb.uni-muenchen.de/csblast/
	HMMER	http://hmmer.janelia.org/
	FASTA	www.ebi.ac.uk/fasta33
Multiple sequence alignment	MSAProbs	http://msaprobs.sourceforge.net/
	DNA Alignment	http://www.fluxus-engineering.com/align.htm
	MultAlin	http://multalin.toulouse.inra.fr/multalin/multalin.html
	DiAlign	http://bibiserv.techfak.uni-bielefeld.de/dialign/
Gene Finding	GenScan	genes.mit.edu/GENSCAN.html
	GenomeScan	http://genes.mit.edu/genomescan.html
	GeneMark	http://exon.biology.gatech.edu/
Protein Domain Analysis	Pfam	http://pfam.sanger.ac.uk/
	BLOCKS	http://blocks.fhcrc.org/
	ProDom	http://prodom.prabi.fr/prodom/current/html/home.php
Pattern Identification	Gibbs Sampler	http://bayesweb.wadsworth.org/gibbs/gibbs.html
	AlignACE	http://atlas.med.harvard.edu/
	MEME	http://meme.sdsc.edu/
Genomic Analysis	SLAM	http://bio.math.berkeley.edu/slam/
	Multiz	http://www.bx.psu.edu/miller_lab/
Motif finding	MEME/MAST	http://meme.sdsc.edu
	eMOTIF	http://motif.stanford.edu

Authors Description:

Md. Mirazul Islam, Student of Biochemistry and Molecular Biology, University of Dhaka, Bangladesh.
Research Interest: Bioinformatics, System Biology, Computational Biology, Molecular Biology.