

Unanswered Questions in Bioinformatics

Mohankumar J. Megha*

Department of Bioinformatics, G. N. Khalsa College, Matunga, Mumbai, India

Keywords: Bioinformatics; Genomics; Proteomics; Questions; Challenges; *omics*

Bioinformatics, a discipline that is attached to almost every field of life sciences due to its extravagance application. It is the application of computers to study, analyse and solve the biological problems. Although it has been explained here in a single sentence, bioinformatics has got much more to it. The exponential growth of biological sequences with the help of sequencing technologies, has led to the dependency on computers for data storage. Considering the costly experiments in wet-labs scientists conduct, a need is felt for automated methods that can reliably predict, compare and analyse the data. Bioinformatics has the potential to accelerate the pace of scientific discovery considering the previously discussed need. There exists a huge contribution of this field in the current research of science as it deals with structural biology, molecular modeling, designing and simulation, data mining, machine learning techniques, databases, optimisation programs, statistical analysis, algorithms, etc. Spreading its wings over these many areas, bioinformatics is presently answering many biological problems; but not all.

Questions in Bioinformatics

Today, though we are capable of addressing many challenges, there exists several challenges that still stand unanswered. As of now, we are able to handle the substantial data generated by high-throughput technologies. New technologies are capable of generating long sequencing reads with additionally higher throughput and much lower cost. This will generate huge data, so called 'big data'. Big data symbolises dosens of terabytes (TB) to many petabytes (PB) of data in a single data set. Are we ready to manage the coming 'big data'? Will the current existing tools and methods suffice the need? If not, there is a need to develop high-performance computing techniques using massive resources for large-scale data analytics. Cloud computing technology can be used to address the big data storage and analysis in the bioinformatics field. This accompanies the challenge to cut-off the conflict of sensitivity and speed in heuristic searching techniques for the large data. Genomics field is reproducing many databases, the challenge is to create knowledge-base. The success of genome projects depends on the ease with which we can obtain accurate and timely answers from these knowledge bases.

Few other questions include: How the gene regulatory networks exactly work? Can we model these networks and simulate the biological pathways? Integrating all the biochemical pathways and reactions that undergo in a biological cell, will help in predicting and simulating cell behaviour under different chemical and physical stimuli. It may even facilitate in precise prediction of cancer and other diseases. Need to enhance the accuracy of the existing tools to reach more than 95%, studying protein-protein interactions at molecular level, predicting protein structures more accurately, understanding the mechanism of the molecular recognition, multiple whole genome comparisons, protein function prediction from sequence (*ab initio*), computer simulation of *in-vivo* systems, optimising the algorithms for better performances, ability to predict when and where transcription site will occur in a genome, comparing order of genes with respect to other organisms, clear understanding

of gene evolution with respect to other genomes; are some of the challenges that are in way of bioinformatics.

Another set of challenges include, understanding how speciation occurs at molecular level, knowing how gene network regulates to form organs out of embryo, variation of gene regulatory network in different type of cells (*viz.*, stem cells, cancer cells and normal cells), precise prediction of structure, gene function, protein function prediction models, protein classification models. After human genome sequencing, a big task of understanding the secrets of junk DNA has become very crucial to find its importance in the genome. As very little is known about junk DNA, the new findings will reveal the things which may even alter the current genomics concepts that we have been acquainted with. Mapping human brain using neural network, which stands impossible as of now, is one of the top challenges of bioinformatics. A super-rapid innovations in Support Vector Machines (SVM), machine learning and artificial intelligence techniques will ease the complexity of solving many empirical problems.

Deploying the vast knowledge of bioinformatics to young scientists, under-graduates, graduates, post graduates with appropriate curricula and proper channels is not less than a challenge, nor can be ignored.

Challenges can be many, but looking at the way the bioinformatics is progressing in the current era of *omics*, certainly it will answer these challenges. Over the coming century, challenges from the allied fields *viz.*, genomics, post-genomics, meta-genomics, proteomics, transcriptomics, metabolomics, epigenomics, physiomics, secretomics, kinomics, connectomics, etc will be targetted to understand the insights of life. The field is pushing the current research to the new heights leaving no wonder that it is the fastest growing field of science.

In bioinformatics, *knowledge can be inputted to generate new knowledge*. No matter what the question is, answer lies around 2 digits (0, 1) and four letters (A, T, G, C).

*Corresponding author: Mohankumar J. Megha, M.Sc. Bioinformatics, Department of Bioinformatics, G. N. Khalsa College, Matunga, Mumbai, India; Tel: +91-9987407394; E-mail: meghamohan@in.com

Received January 08, 2013; Accepted February 07, 2013; Published February 10, 2013

Citation: Megha JM (2013) Unanswered Questions in Bioinformatics. J Proteomics Bioinform 6: 030-030. doi:10.4172/jpb.1000257

Copyright: © 2013 Megha JM. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.