

Bioinformatics (Computational Biology)

Himanshu Pal

MCA Department

ASM College Thane Mumbai

Abstract

Bioinformatics is the use of computational approach to analyse, manage and store biological data. The research in biotechnology especially that involving sequence data management and drug design occurred at a speedy rate due to development of bioinformatics. A number of tools and software are developed for analysis and interpretation of biological complexity. There are number of applications of bioinformatics viz. sequence analysis and alignment, molecular modelling, docking, annotation and dynamic simulation to accelerate the biotechnological research. It is expected that many future bioinformatics innovations are likely to stimulate analysis of vast biological data. Here, in this communication, we have tried to explain the importance of bioinformatics in various fields of biotechnology viz. genomics, proteomics, transcriptomics, cheminformatics, climate change studies, drug discovery and development, waste cleanup, bioenergy, crop improvement, veterinary sciences, forensic sciences and biodefense.

Keywords

Bioinformatics; Biotechnology; Biological complexity; Molecular modelling; Genomics; Proteomics; Transcriptomics; Cheminformatics; Drug discovery

Introduction

Bioinformatics is the short form for 'Biological Informatics'. This is considered to be an amalgam of biological sciences and computer science and now a days, many scientists prefer to use the term, computational biology. This branch of science became more popular when human genome project came into existence. Bioinformatics merges biology, computer science and information technology to form a single discipline. It covers many areas of biological science especially of modern biology viz. genomics, transcriptomics, proteomics, genetics, and evolution. The ultimate goal of the field is to enable the discovery of new biological insights as well as to create a global perspective from which unifying principles in biology can be recognized.

Bioinformatics is a fascinating subject having input of engineering art as well as of science. Bioinformaticians are mostly engaged in designing new algorithms, software, developing updated databases which all help in solving many biological problems. A number of bioinformatics tools, software and databases are available for better understanding of biological complexity and analyse and store the biological data. Thus, the bioinformatics research is used to avoid time, cost and wet lab practice. Scientists realized the importance of sequence databases in 1950s and that's why, first protein sequence database was created in 1956 just after insulin peptides sequences became available. The human genome sequence data is so huge that if compiled in books, the data would run into 200 volumes of 1000 pages each and reading alone would require 26 years working around the clock. This challenge of handling such a huge data can only be possible because of bioinformatics.

The growth of the biotechnology industry in recent years is unprecedented, and advancements in molecular modelling, disease characterization, pharmaceutical discovery, clinical healthcare, forensics, and agriculture fundamentally impact economic and social issues worldwide. As a result, with people confidence and development of biotechnology, bioinformatics also reached to new heights among all the biological sciences. There exists a number of applications of bioinformatics for accelerating research in the area of biotechnology that include automatic genome sequencing, gene identification, prediction of gene function, prediction of protein structure, phylogeny, drug designing and development, identification of organisms, vaccine designing, understanding the gene and genome complexity, understanding protein structure, functionality and folding. By using bioinformatics in research, many long term projects are turned up so fast like genome mapping of human and other organisms. Similarly, it is expected that bioinformatics innovation in future will also meet the demands of biotechnology. Here, we have tried to explain the role of bioinformatics in various fields of biotechnology like drug designing and development, genomics, proteomics, environment biotechnology and others.

Genomics

The study of genes and their expression is called as Genomics. This field generates a vast amount of data from gene sequences, their interrelation and functions. To manage this vast enormous data, bioinformatics plays a very important role. With the complete genome sequences for an increasing number of organisms, bioinformatics is beginning to provide both conceptual bases and practical methods for detecting systemic functional behaviours of the cell and the organism. Bioinformatics plays a vital role in the areas of structural genomics, functional genomics and nutritional genomics.

Proteomics

The study of protein structure, function, and interactions produced by a particular cell, tissue, or organism is called as proteomics. It deals with techniques of genetics, biochemistry and molecular biology. Advanced techniques in biology led to accumulate enormous data of protein-protein interactions, protein profiles, protein activity pattern and organelles compositions. This vast data can be managed and access easily by using bioinformatics tools, software and databases. Till now, many algorithms in the field of proteomics viz. image analysis of 2D gels, peptide mass fingerprinting and peptide fragmentation fingerprinting have been developed.

Transcriptomics

The study of sets of all messenger RNA molecules in the cell is called as transcriptomics. This can also be called as Expression Profiling where DNA microarray is used to determine the expression level of mRNA in a given cell population. The microarray technique generates vast amount of data, single run generates thousands of data value and one experiment requires hundreds of runs. Analysis of such vast data is done by numerous software packages. In this way, bioinformatics is used for transcriptome analysis where mRNA expression levels can be determined. RNA sequencing (RNAseq) also has been included under transcriptomics. It is carried out using next generation sequencing to determine the presence and quantity of RNA in a sample at a given time. It is used to analyse the continuously changing cellular transcriptome.

Cheminformatics

Cheminformatics (chemical informatics) focuses on storing, indexing, searching, retrieving, and applying information about chemical compounds. It involves organization of chemical data in a logical form to facilitate the retrieval of chemical properties, structures and their relationships. Using bioinformatics, it is possible through computer algorithm to identify and structurally modify a natural product, to design a compound with the desired properties and to assess its therapeutic effects, theoretically. Cheminformatics analysis includes analyses such as similarity searching, clustering, QSAR modelling, virtual screening, etc.

Drug Discovery

Bioinformatics is playing an increasingly important role in nearly all aspects of drug discovery, drug assessment and drug development. This growing importance is not because bioinformatics handles large volumes of data but also in the utility of bioinformatics tools to predict, analyse and help interpretation in clinical and preclinical findings. Traditionally, pharmacology and chemistry-based drug discovery approaches face various difficulties in finding new drugs. The increasing pressure to generate more and more drugs in a short period of time with low risk has resulted in remarkable interest in bioinformatics. In fact, now there is an existence of new separate field known as computer-aided drug design (CADD). Bioinformatics provides a huge support to overcome the cost and time context in various ways. It provides wide range of drug-related databases and software's which can be used for various purposes related to drug designing and development process.

Evolutionary Studies/Phylogenetics

The study of evolutionary relationship among individuals or group of organisms is defined as phylogenetics. Taxonomists find the evolutionary relationship using various anatomical methods that takes too much time. Using Bioinformatics, phylogenetic trees are constructed based on the sequence alignment using various methods. Various algorithmic methods are developed for the construction of phylogenetic tree that are used depending on the various evolutionary lineages.

Crop Improvement

Sustainable agricultural production is an urgent issue in response to global climate change and population increase. Innovations in omics based research improve the plant based research. The integrated 'omics' strategies clarify the molecular system of the plant which are used to improve the plant productivity. Genomics strategy, especially comparative genomics helps in understanding the genes and their functions, and also the biological properties of each species. Bioinformatics databases are also used in designing new techniques and experiments for increased plant production.

Veterinary Science

Food production from livestock can meet demand of human population for food. For better bio-economy, there is a need of efficient animal production and reproduction. This is achieved with better understanding of livestock species. Current and new methods in livestock species using data from experimental or field studies with bioinformatics are helping in understanding the systems genetics of complex traits and provide biologically meaningful and accurate predictions. Finally, almost all of the next generations-omics tools and methods that are used in other fields of biological sciences, can also be used in veterinary sciences.

Forensic Science

Forensic science includes the study regarding identification and relatedness of individuals. It is inherently interdisciplinary with bioinformatics as both are dependent on computer science and statistics. This field is based on the molecular data and many databases are being developed to store the DNA profiles of known offenders. This field is being pushed due to technological and statistical advances in microarray, Bayesian networks, machine learning algorithms, TFT biosensors and others. This provides the effective way of evidence organization and inference.

Biodefense

Biodefense includes measures to restore biosecurity to a group of organisms who are subjected to biological threats or infectious diseases (in context of bio-war or bioterrorism). Today, bioinformatics has a limited impact on forensic and intelligence operations. There is a need of more algorithms in bioinformatics for biodefense so that the developed databases may show interoperability with each other. In order to use next generation genome sequencing for forensic operation, bio-threat awareness, mitigation and medical intelligence, there is a need for development of more computational applications.

Waste Cleanup

Today, the major concern all over the Globe is environmental pollutants. The main concern of the environmentalists is waste generated from the industries. These pollutants progressively deteriorate the environment which in turn affects human health. There are few microorganisms that are considered to remediate the pollutants into the natural biogeochemical cycle. Bioremediation is the recent technology which explores the microbial potentiality for biodegradation. This technology can be further improved by using bioinformatics. Genomic and bioinformatics data provide a wealth of information that would be greatly enhanced by structural characterisation of some protein. Bioinformatics provides data of microbial genomics, proteomics, systems biology, computational biology, and bioinformatics tools for understanding of the mechanisms of biodegradative pathways.

Climate Change Studies

Another Global concern is the Climate change because of loss of sea ice, accelerated sea level rise and longer and more intense heat waves. To solve this issue, bioinformatics may help by way of sequencing microbial genome which can reduce levels of carbon dioxide and other greenhouse gases. This plays an important role in stabilizing the global climate change. Not much work has been done in this area in bioinformatics domain, and more region-specific work must be conducted considering microbes of that region and their capability in CO₂ reduction.

Bioenergy/Biofuels

Biofuels offer great promise in contributing to the growing global demand for alternative sources of renewable energy. Bioinformatics is important in understanding and analysis of biofuel producing pathways. Recent progress in algal genomics, in conjunction with other omics approaches, has accelerated the ability to identify metabolic pathways and genes that are potential targets in the development of genetically engineered micro-algal strains with optimum lipid content.

Acknowledgements

Authors acknowledge the facilities of the Department of Biotechnology, Ministry of Science and Technology, Government of India, New Delhi (DBT) under the Bioinformatics Sub Centre as well as M.Sc. Biotechnology program used in the present work.