



E-ISSN: 2278-4136
P-ISSN: 2349-8234
JPP 2017; 6(4): 1938-1942
Received: 06-05-2017
Accepted: 07-06-2017

Brahm Prakash Dahiya
Assistant Professor, B.L.M. Girls
College, Punjab, India

Bioinformatics impacts on medicine, microbial genome and agriculture

Brahm Prakash Dahiya

Abstract

Bioinformatics is combination of biology and information technology. It is multidisciplinary area of science combination of biological computational, mathematical, statistical techniques and information technology. Using information technology tools, we can analysis biological data sequences. Bioinformatics also helps in biological algorithm development and suitable data analysis tools to collect information and make discoveries. Here we will discuss role of bioinformatics and its various tools for biological data analysing using computational and statistical techniques. In other side many works have done to improve proteomic level in different crop varieties using bioinformatics. We will also discuss various applications of bioinformatics in various fields and impact of bioinformatics in Medicine, Microbial Genome Application and Agriculture.

Keywords: Bioinformatics, biotechnology, genomics, agriculture and DNA

1. Introduction

Bioinformatics is an application of information technology to the processing and analysing the data generated in biological researches and experiments or we can say that bioinformatics is a computerized system for molecular biology and has many practical's use in biology. Bioinformatics objective to enrich biological data and to apply computer based algorithm to analyse biological data. We also can say that it is about analysis, data storage and exchange of huge amounts of biological data. It helps in DNA barcoding; design the patterns of disease outbreaks and new biological products. Bioinformatics has high capacity to analysis data and to influence the attribute of plants, animals and human beings. Many researches have processed to enhance the gene and protein sequences. In proteomics, bioinformatics helps in the study of protein structures and the discovery of sequence sites where protein-protein interactions take place. To help understanding biology at the system level, bioinformatics begins to show promise in unravelling genetic networks. Finally, in the study of metabolomics, bioinformatics is used to study the dynamics in a cell, and thus to simulate the cellular interactions^[1-2].

Bioinformatics today has come in every major discipline in biology. In genomics, Bioinformatics has helped in genome sequencing, and has presented its success in locating the genes, in phylogenetic comparison and in the detection of transcription factor binding sites of the genes. Bioinformatics provides analytical tools for microarray data^[2]. These tools range from image the processing techniques that read out the data, to the visualization tools that provide a first-sight hint to the biologists; from pre-processing techniques that remove the systematic noise in the data to the clustering methods that reveal genes that behave similarly under different experimental conditions. Tools of bioinformatics are providing useful information about the genes existing in the genome of these species. These tools made it promising to predict the function of different genes and factors affecting these genes^[4]. These collections of information are useful to the scientists to produce enhanced species of plants which have drought, herbicide, and pesticide resistance in them. Specific genes can be altered to improve the production of meat and milk. Positive changes can be made in their genome for disease resistant. Bioinformatics deals with the exponential growth in biological data have led to the development of primary and secondary databases of nucleic acid sequences, protein sequences and structures. Some of the well-known databases include Gen-Bank, SWISS-PROT, PDB, PIR, SCOP, CATH, etc. These databases are available as public domain information and hosted on various Internet servers across the world. Basic research and modelling is done using these databases with the help sequence analysis tools such as BLAST, FASTA, CLUSTALW, etc., and the moulded structures are visualized using visualization tools such as Rasmol, Web-Lab, MOLMOL etc.^[5].

Correspondence
Brahm Prakash Dahiya
Assistant Professor, B.L.M. Girls
College, Punjab, India

The integration of bioinformatics impacts on plant science and crop enhancement area. The identification of essential genes based on genomics, expression analysis and functional genomics. In aggregation with the design and construction of transgenic plants this will allow new target genes to be identified that will improve quantitative and qualitative traits in commercially important crops. The design of agrochemicals based on an analysis of the components of signal perception and transduction pathways to select objective, and with chemo-informatics, to identify effective compounds that can be used as herbicides, pesticides, or insecticides. The management of plant genetic resources to preserve genetic diversity in various agricultural species. To search taxonomic data goes far beyond the field of taxonomy and information of all species with requirement of phenotypic and genotypic attributes. The taxonomic provides stability to work of management and conservation bodies. The best management of biological repositories of clones, cell lines, organisms and seeds. Typically, existing repositories are not linked to each other databases. The many commercial databases and repositories are also part of the bioinformatics infrastructure but operate largely outside of the present day cooperative activities. Imagine studying gene structures without the help of a computer. It would take many years to compare the 15,000 genes of Arabidopsis to the genes of a similar plant. And keeping track of the 100,000 genes of a human being would be inconceivable. With computers, the process of comparison is automated. By storing information as it is discovered, computers ease the immense job of genome mapping. But computers can analyse as well as store information. They can be used to construct models that reduce the need for experimentation. In this way, biotechnology has become more efficient. Scientists are able to use fairly reliable computer-assisted predictions of test results on genetic modifications. This complements the time-consuming process involved in growing out every modified plant in the laboratory or greenhouse to test for the desired modification [1, 3, 4].

2. Applications of Bioinformatics

Bioinformatics is wide used in the area of Medicine, Microbial Genome and Agriculture. Following are the application of Bioinformatics in these fields:

2.1 Medicine: In the field of Medicine of bioinformatics is covered following areas:

- **Drug Discovery**

Infectious diseases are now the world's biggest killers of children and young adults. "They account for more than 13 million deaths a year - one in two deaths in developing countries" as stated by the WHO. Most deaths from infectious diseases occur in developing countries. The cause for this has been attributed to the unavailability of efficient drugs and if at all available, the high cost associated with those drugs. Development of cheap and efficient drugs for a disease is one of the major problems faced by mankind. This problem can be solved by rational drug design using Bioinformatics. Pharmacy industry has switched from the trial and error process of drug discovery to a rational, structure based drug design. A successful and reliable drug design process could reduce the time and cost of developing useful pharmacological agents. Computational methods are used for the prediction of 'drug-likeness' which is nothing but the identification and elimination of candidate molecules that are unlikely to survive the later stages of discovery and

development. Drug-likeness could be predicted by genetic algorithm and neural network based approaches. The Idea of using X ray Crystallography in drug discovery emerged more than 30 years ago, when the first 3 dimensional structure of protein was determined. Protein structure can influence drug discovery at every stage in design process. Classically, it is used in lead optimization, a process that uses structure to guide the chemical modification of a lead molecule to give an optimised fit in terms of shape, hydrogen bonds and other non-covalent interactions with the target [3, 16].

- **Personal Medicine**

Personalized medicine is developing practice of medicine that uses particular's genetic profile to advice decisions made in favour to the prevention, diagnosis, and treatment of disease. Information about patient's genetic profile can help doctors to provide proper medication using the proper dose or regimen. It applied for treatment as personalized cancer medicine, Diabetes- related disease and HIV.

Personalized medicine can be defined widely as a model of healthcare that is predictive, personalized, preventive and participatory. Translational bioinformatics is a field that can help address these challenges and is defined by the American Medical Informatics Association as "the development of storage, analytic and interpretive methods to optimize the transformation of increasing voluminous biomedical data into proactive, predictive, preventative and participatory health. Practical application outside of long established considerations like a patient's family history, social circumstances, environment and behaviours are very limited so far and practically no progress has been made in the last decade. Personalized medicine research tries to discover individual solutions based on the susceptibility profile of each individual. It is expected that these areas will enable new approaches to diagnosis, drug development, and individualized therapy [8, 9, 14].

- **Preventive Medicine**

Preventive Medicine is practiced by all physicians to possess their patients healthy. It is also a unique medical subject affiliated by the American Board of Medical Specialties (ABMS). Preventive Medicine attentions on the health of individuals, communities, and defined populations. It is also helpful for the treatment for obesity, blindness. The Epidemiology Division used research methods to understand the patterns and causes of health and disease in the population and to transform this knowledge into programs designed to stop disease. The division has a lengthy history of association in NIH-sponsored multi-site, longitudinal cohort studies, and its faculty oversees many investigator-initiated, NIH-sponsored research projects and trials. Public trust in vaccines is a key to the success of immunization programs worldwide in the era of preventive medicine. Preventive medicine or preventive care contains of measures engaged to prevent diseases in stand of curing and treating their symptoms. Many different techniques such as curative and palliative medicine, and applied public health methods [10].

- **Predictive Medicine**

Predictive medicine is an area of medicine that involves expecting the probability of disease and instituting preventive processes in order to either prevent the disease completely or considerably decrease its influence upon the patient. Techniques and assays include New-born screening, diagnostic testing, Medical bioinformatics, Prenatal testing,

Carrier testing, Preconception testing. New-born screening is a public health program designed to screen infants shortly after birth for a list of conditions that are treatable, but not clinically evident in the new born period [15].

Gene Therapy: Gene therapy is a new method of drug delivery that detail lists of synthetic machinery of the patient's cell to create a therapeutic agent [15]. It contains the efficient introduction of functional gene into the suitable cells of the patient in order to create sufficient amount of protein encoded by transferred gene so as to exactly and permanently correct the disorder. Strategies of Gene Therapy classified into three categories:

- Gene addition
- Removal of harmful gene by antisense nucleotide or ribozymes
- Control of gene expression

2.2 Microbial Genome Applications

In the field of Microbial Genome Applications of bioinformatics is covered following areas [18]:

- **Waste Clean-up:** In bioinformatics bacteria and microbes are helpful in cleaning waste. *Deinococcus radiodurans* Bacterium is point out in the Guinness Book of World Records and this bacterium has the ability to repair damaged DNA and small fragments from chromosomes by isolating damage segments concentrated area. This is because it has additional copies of its genome. Genes from other bacteria have been inserted into *D. radiodurans* for environmental clean-up. It was used to break down organic chemicals, solvents and heavy metals in radioactive waste sites.
- **Climate Change:** All concerned areas of science and technology including bioinformatics plays a main role in the control of climate change to some level. Bioinformatics is also assisting in climate change studies. In the last few years, various genome sequencing and metagenomic projects have produced large amounts of genetic data for cyanobacteria. This wealth of data affords researchers with a new basis for the study of molecular adaptation, ecosystem and evolution of cyanobacteria, as well as for developing biotechnological applications. It also helps the use of multiplex techniques, i.e., expression profiling by high-throughput technologies such as microarrays, RNA-sequence, and proteomics. Microbe *Deinococcus radiodurans* is known for radiation resistance and being used for cleaning up the waste sites that contain toxic chemicals. There are many organisms which use carbon dioxide as their sole carbon source and increasing levels of carbon dioxide emission is one of the major causes of the global climate change. The study of genomes based on microbial organisms, which is possible using bioinformatics, helps in decrease the carbon dioxide content. Climate change is caused by factors that contain oceanic methods variations in solar radiation received by Earth. By studying microorganisms genome scientists can begin to understand these microbes at fundamental level and isolated the genes that give them their unique abilities to survive under extreme conditions [11-12].
- **Nanotechnology and Biotechnology:** The concept of biotechnology includes a wide range of procedures for altering living organisms according to human drives, going to training of animals, farming of plants, and enhancements based on breeding programs that employ

artificial selection techniques. Advance usage also contains genetic, cell and tissue culture engineering technologies. In the field of bioinformatics and biotechnology identified organisms and microorganisms which are very beneficial in dairy industry and food manufacturers. *Lactococcus Lactis* is very important micro-organisms involved in the dairy industry. Nanotechnology is the impacts on an atomic, molecular, and super molecular scale [5].

• Alternative Energy

Alternative energy sources can be segmented by Geographies, Applications, Macro Indicators and Companies. Geographies of this market are North America, Europe, Asia-Pacific and Rest of World. Applications of this market are NGS data analysis, storage and management, Metabolomics, Proteomics, Phylogenetic, Transcriptomics and Other Applications. Macro Indicators of this market are Healthcare Expenditure. Companies of this market are Accelrys, Affymetrix, Agilent Technologies, Biomatters Ltd, CLC bio, Ocimum Biosolutions Ltd., IDBS, Illumina, Biomax Informatics AG, Kinexus Bioinformatics Corporation, Integromics, Instem plc, Partek Incorporated and GenoSplice technology [3].

2.3 Agriculture

In the field of Agriculture, applications of bioinformatics are covered following areas:

• Crop Improvement

Different genetics of the plant genomes has shown that the group of their genes has remained more conserved. These results suggest that information obtained from the model crop systems can be used to propose improvements to other food crops. *Arabidopsis thaliana* was the first plant to be sequenced and considered the model species for examining plant genetics and biology. There are many genes which are similar in all plants and the study of genes in a model organism such as *A. thaliana* facilitates understanding of gene expression and function in all plants species. *Oryza sativa* most famous variety of rice. It is the very essential crop for human consumption, providing staple food more than half of the world population [17]. *Oryza sativa* was the cereal a selection of to be sequenced as a priority and has gained the status "model organism". It has the smallest genome all the cereals: 430 million nucleotides and it can serve as a model genome for one of the two main groups of flowering plants, the monocotyledons. Many studies on hybrid vigor, genetic resistance to disease and adaptive responses, scientists have taken advantage of the existence of a multitude of varieties that have adapted to a very wide range of environmental conditions, from dry soil in temperate regions to flooded cultures in tropical regions [1, 6-7].

- **Insect Resistance:** Genes from *Bacillus thuringiensis* that can control a number of serious pests have been effectively transferred to cotton, maize and potatoes. This new ability of the plants to resist insect attack and hence the nutritional quality of the crops production is increased [17]. *Bacillus thuringiensis* is a pathogenic bacteria mostly applied for insect control. It is rod-shaped aerobic bacteria in the genus *Bacillus*. *B. thuringiensis* an insect bacterium, advertised worldwide for resistor of many important plants. They are applied to leaves where the insect larvae feed. The toxin genes have also been

genetically applied into several crop plants to enhance production of crops ^[11].

- **Improve Nutritional Quality:** Agriculture Scientists and research scholars have recently successfully transferred genes into rice to increase levels of Vitamin A, iron and other micronutrients. This work could have profound influence in reducing occurrences of blindness and anaemia affected by deficiencies in Vitamin A and iron respectively. They have successfully inserted a gene from yeast into the tomato, and the result is a plant whose fruit stays longer on the vine and has an extended shelf life time.

Using bioinformatics techniques agricultural Scientists and research scholar will be able to modified breeds of tomatoes that will be more flavourful even after the long journey from the vine to the produce department. Picking tomatoes early means they have chance to develop flavour, colour, and nutrients naturally. By modified the rin gene, agricultural scientists will be able to slow the ripening process, letting the tomato develop on the vine longer - but still keeping it firm enough to ship safely. They hope that their technique may also be applied to other fruits - such as strawberries, bananas, bell peppers, and melons - which suffer from the same shipping and storage complications ^[11, 19].

3. Bio-Informatics Tools

The Bioinformatics tools are the application programs for the saving, modify, retrieving and analysis of Biological data. Factors that must be taken into consideration when designing these tools are ^[1].

- a. The end user (the biologist) may not be a frequent user of Information technology and thus it should be very user friendly.
- b. These software tools are available on internet and play important roles in scientific research community.

The bioinformatics tools may be categorized into following categories:

- **Homology and Similarity Tools**

The term homology means relationship between DNA sequences or bristle patterns on a fly's nose. Homologous sequences are sequences that are linked by divergence from a common ancestor. So the degree of similarity between two sequences can be measured while their homology is a case of being either true or false. This set of tools can be used to identify connections between novel query sequences of unknown structure and function and database sequences whose structure and function have been elucidated ^[1].

- **Protein Function Analysis**

Protein function analysis is Identified and of all functional elements such as coding and non-coding in a genome. The collection of programs allows you to make comparison in your protein sequence to the alternative or secondary protein databases that cover information on motifs, signatures and protein domains. Highly important hits against these different pattern databases ^[1].

- **Structural Analysis**

Using structural analysis we can compare structures with the known structure databases. The function of a protein is directly a importance of its structure instant of its sequence with structural homologs. The importance of a protein's 2D/3D structure is most important study point of view ^[1-2].

- **Sequence Analysis**

Many sequence analysis tools allows you to analysis on your query sequence such as identification of mutations, evolutionary analysis, and compositional biases. The identification of biological properties is search based on specific function of your sequence ^[1, 3].

- **Blast**

The Basic Local Alignment Search Tool (BLAST) for analysis gene and protein sequences against others in public databases. BLASTs are also useful for human, microbial, malaria, and other genomes, and tentative human consensus sequences ^[21].

- **Fasta**

FASTA also a search database tool that compares a nucleotide or peptide sequence with a sequence database based on the rapid sequence algorithm. It was the first and most useful algorithm for database similarity searching ^[22].

- **Emboss**

EMBOSS (The European Molecular Biology Open Software Suite) is a latest and free open source software analysis package specially developed to analyse the molecular data. It is combination around hundred programs for sequence alignment, data searching with sequence patterns, protein motif identification and domain analysis, nucleotide sequence pattern analysis, codon usage analysis for small genomes, and much more ^[23].

- **Clustalw**

ClustalW is multiple sequence alignment programs for DNA or proteins identification. It generates biological suitable or meaningful multiple sequence alignments of different sequences, calculates the best match for the selected sequences ^[24].

- **RasMol**

It is a most powerful research tool. It displays the structure of DNA, proteins and smaller molecules. Protein Explorer, a derivative of RasMol. It is an easier to use program ^[24].

- **JAVA in Bioinformatics:** Java is platform independence. It is plays key role in bioinformatics. Many simulators such as computer-based biological simulation technologies and Bioinformatics Solutions' Pattern Hunter are two examples of Java in bioinformatics ^[1].

- **Perl in Bioinformatics**

Perl is most important software tool to analyse biological data. Bio-Perl project is based on bioinformatics.

- **XML**

XML is also play very important role in bioinformatics. BIOXML project implemented in bioinformatics.

4. Conclusions and future work

Bioinformatics has played very important role in developing country like India in many fields. Bioinformatics is more significance to analyse and identification of biological data using computer based techniques to help understanding

biology at the system level, bioinformatics begins to show promise in unravelling genetic networks. It is more useful in dynamic cell study. It made positively changes in their genome for disease resistant. Bioinformatics is more practical in medicine fields such as drug discovery, personal medicine, preventive medicine, predictive, medicine and Gene therapy. Many application of bioinformatics in fields of microbial genome such as waste clean-up, climate change, nanotechnology and biotechnology and alternative energy. It is more applicable in area of agriculture such as crop improvement, insect resistance improve and nutritional quality. The integration of bioinformatics impacts on plant science and crop enhancement area. Many bioinformatics tools and techniques are applied to analysis biological data and enhancement in genes sequences. Bioinformatics made effective changes in different branches of science and life science. In future work we will try to enhance the vitamin level in rice and protein level in wheat using existing and proposed bioinformatics techniques.

References

1. Ujjawal Kumar Singh Kushwaha, Indra Deo, Jai Prakash Jaiswal, Birendra Prasad. Role of Bioinformatics in Crop Improvement, Global Journal of Science Frontier Research: D Agriculture and Veterinary. 2017; 17:1, Version 1.0 Year.
2. Batley Jacqueline, David Edwards. The application of genomics and bioinformatics to accelerate crop improvement in a changing climate. Current Opinion in Plant Biology, 2016; 30:78-81.
3. Siddharthan N, Raja Prabu M, Sivasankari B. Bioinformatics in Drug Discovery a Review, International Journal of Research in Arts and Science. 2016; 2:2.
4. Ohyanagi H *et al.* Oryza Genome: genome diversity database of wild Oryzaspecies, Plant Cell Physiol. 2016; 57:1.
5. Aken BL *et al.* The Ensembl gene annotation system, Database (Oxford), 2016.
6. Hao L *et al.* Information commons for rice (IC4R), Nucleic Acids Res. 2016; 44(D1):D1172-D1180.
7. Brozynska M, Furtado A, Henry RJ. Genomics of crop wild relatives: expanding the gene pool for crop improvement. Plant Biotech. 2015.
8. Li Zhang, Huixiao Hong. Genomic Discoveries and Personalized Medicine in Neurological Diseases, pharmaceuticals, ISSN 1999-4923, 2015; 7:542-553.
9. Genome Sequencing to Uncover Disease. Available online: <http://hudsonalpha.org/wp-content/uploads/2014/04/Genome-Sequencing-to-Uncover-Disease.pdf> (accessed on 2 December 2015).
10. Hemlata, Archana Tiwari. Applications of Bioinformatics tools to combat the Antibiotic Resistance, International Conference on Soft Computing Techniques and Implementations- (ICSCTI) Department of ECE, FET, MRIU, Faridabad, India, 2015.
11. Abberton M, Batley J, Bentley A, Bryant J, Cai H, Cockram J *et al.* Global agricultural intensification during climate change: a role for genomics. Plant Biotechnology, 2015.
12. Barros VR, Field CB. Climate Change Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change; Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 688, 2014.
13. Hong H, Hong Q, Liu J, Tong W, Shi L. Estimating relative noise to signal in DNA microarray data. Int. J. Bioinforma. Res. Appl. 2013; 9:433-448.
14. Hong H, Zhang W, Shen J, Su Z, Ning B, Han T *et al.* Critical role of bioinformatics in translating huge amounts of next-generation sequencing data into personalized medicine. Sci. China Life Sci. 2013; 56:110-118.
15. Hack C, Kendall G. Bioinformatics: current practice and future challenges for life science education. BIOCHEM. MOLECULAR BIOL. EDUC. 2013; 33:82-85.
16. Iskar M, Zeller G, Zhao XM, Van Noort V, Bork P, Drug discovery in the age of systems biology: the rise of computational approaches for data integration, Curr Opin Biotechnol, 2012; 23:609-616.
17. Gu H *et al.* PRIN: a predicted rice interactome network, BMC Bioinf., 2011; 12:161.
18. Li J, Jiang Y, Fan R. Recognition of Biological Signal Mixed Based on Wavelet Analysis. In: Y. Jiang, *et al.* (eds.). Proc. of UK-China Sports Engineering Workshop. Liverpool: World Academic Union (World Academic Press), 2007, 1-8.
19. Mohammadi SA, Prasanna. Analysis of Genetic Diversity in Crop Plants Salient Statistical Tools and Considerations, Crop Science. 2003; 43:1235-1248.
20. Baxevanis A, Ouellette F. Bioinformatics: A practical guide to the analysis of genes and proteins. John Willey & Sons, Inc. N.Y., USA. 2001, 518.
21. Document created by Wilson Leung, Washington University. Available online:<http://blast.ncbi.nlm.nih.gov/Blast.cgi>.
22. https://en.wikipedia.org/wiki/FASTA_format.
23. Rice P1, Longden I, Bleasby A. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet. 2000; 16(6):276-267.
24. NCBI Insights. Available online: <https://ncbiinsights.ncbi.nlm.nih.gov/2017/01/25/visualize-and-interpret-alignment-data-with-the-multiple-sequence-alignment-viewer/>.