



ISSN 2347-2677

IJFBS 2018; 5(2): 03-10

Received: 04-01-2018

Accepted: 05-02-2018

M Younus Wani

Temperate Sericulture Research
Institute, Mirgund, SKUAST-
Kashmir, J&K, India

NA Ganie

Temperate Sericulture Research
Institute, Mirgund, SKUAST-
Kashmir, J&K, India

S Rani

Temperate Sericulture Research
Institute, Mirgund, SKUAST-
Kashmir, J&K, India

S Mehraj

Temperate Sericulture Research
Institute, Mirgund, SKUAST-
Kashmir, J&K, India

MR Mir

Temperate Sericulture Research
Institute, Mirgund, SKUAST-
Kashmir, J&K, India

MF Baqual

Temperate Sericulture Research
Institute, Mirgund, SKUAST-
Kashmir, J&K, India

KA Sahaf

Temperate Sericulture Research
Institute, Mirgund, SKUAST-
Kashmir, J&K, India

FA Malik

Temperate Sericulture Research
Institute, Mirgund, SKUAST-
Kashmir, J&K, India

KA Dar

Temperate Sericulture Research
Institute, Mirgund, SKUAST-
Kashmir, J&K, India

Correspondence

M Younus Wani

Temperate Sericulture Research
Institute, Mirgund, SKUAST-
Kashmir, J&K, India

Advances and applications of Bioinformatics in various fields of life

M Younus Wani, NA Ganie, S Rani, S Mehraj, MR Mir, MF Baqual, KA Sahaf, FA Malik and KA Dar

Abstract

Bioinformatics is an interdisciplinary area of the science composed of biology, mathematics and computer science. Bioinformatics is the application of information technology to manage biological data that helps in decoding plant genomes. The field of bioinformatics emerged as a tool to facilitate biological discoveries more than 10 years ago. With the development of Human Genome Project (HGP), the data of biology increased fabulously and marvelously. The ability to capture, manage, process, analyze and interpret data became more important than ever. Bioinformatics and computers can help scientists to solve it. Here are introduced roles of bioinformatics, meanwhile Web tools and resources of bioinformatics are reviewed and its applications in agriculture and relevance with other disciplines is also highlighted. Application of various bioinformatics tools in biological research enables storage, retrieval, analysis, annotation and visualization of results and promotes better understanding of biological system in fullness. This will help in animal and plant health care based disease diagnosis and treatment.

Keywords: annotation, agriculture, bioinformatics, disease diagnosis, health and web tools

Introduction

Term Bioinformatics was coined by Paulien Hogeweg and Ben Hesper in 1970 as the study of informatic processes in biotic systems. Bioinformatics deals with computational management and analysis of biological information (genes, genomes, proteins, cells, ecological systems, medical information, robots, artificial intelligence etc.

The National Center for Biotechnology Information (NCBI 2001) defined Bioinformatics as the field of science in which biology, computer science, and information technology merge into a single discipline. Fredj Tekaia at the Institute Pasteur defines bioinformatics the mathematical, statistical and computing methods that aim to solve biological problems using DNA and amino acid sequences and related information. Since the sequencing of the first complete microbial genome of *Haemophilus influenzae* in 1995 hundreds of microbial genomes have been sequenced and archived for public research in Gen Bank. The vast amount of data generated by genome sequencing projects is becoming unmanageable. Bioinformatics has silently filled in the role of cost effective data analysis. Bioinformatics analysis has enhanced our understandings about the genome structure and the microorganism restructuring process. Bioinformatics has emerged as an essential field of science that is facilitating biological discoveries since more than a decade. Without the usage of bioinformatics tools it is merely impossible to capture, manage process, analyse and interpret the huge amounts data that is available especially after whole genome sequencing projects. The sequencing of the genomes of plants and animals will have enormous benefits for the agricultural community. Bioinformatics tools can be used to search for the genes within these genomes and to elucidate their functions. This specific genetic knowledge could then be used to produce stronger, drought, disease and insect resistant crops and improve the quality. In agriculture it helps in the insect resistance, improve nutritional quality, rational plant improvement, waste cleanup, climate change studies, and development of drought resistance varieties (Dahiya and Lata, 2017) [4] and in addition to this it also plays an important roles in biotechnology, antibiotic resistance, and forensic analysis of microbes, comparative studies, evolutionary studies and veterinary Sciences

History of Bioinformatics

1865: Father of Genetics: Gregor Mendel discovers the concept of genetic inheritance

1930: Electrophoresis introduced
1953: Watson and Crick suggest double-helix model for DNA
1955: Bovine Insulin is first protein to be sequenced
1978: the term Bioinformatics first used
1984: FASTP algorithm program published
1990: BLAST program published
1994: PRINTS database published
1995: First bacterial genomes sequenced
2000: EMOTIF database released

Why is bioinformatics important

Bioinformatics has developed out of the need to understand the code of life that is DNA, the massive DNA sequencing projects have evolved and added in the growth of the science of bioinformatics. The ultimate goal of bioinformatics is to uncover the wealth of biological information hidden in the mass of sequence, structure, literature and other biological data.

Components of Bioinformatics

Technology and Computing power

Technology is the collection of techniques, skills, methods, and processes applied to make life easier. Technology can be the knowledge of techniques, processes.

Computing power is the ability to undertake or be used for computation. Specifically the ability of a computer to perform work, often considered in terms of the number of instructions that can be carried out in a given time, or with reference to the amount of random access memory present.

Creation of databases

This involves the organizing, storage and management the biological data sets. The databases are accessible to researchers to know the existing information and submit new entries, e.g. protein sequence data bank for molecular structure. Databases will be of no use until analysed.

Development of algorithms and statistics

This involves the development of tools and resources to determine the relationship among the members of large data sets e.g. comparison of protein sequence data with the already existing protein sequences.

Analysis of data and interpretation

The appropriate use of components to analyse the data and interpret the results in a biologically meaningful manner. This includes DNA, RNA and protein sequences, protein structure, gene expression profiles and biochemical pathways.

Computational Biology

Broadly speaking, computational biology is the application of computer science, statistics, and mathematics to problems in biology. Computational biology spans a wide range of fields within biology, including genomics/genetics, biophysics, cell biology, biochemistry, and evolution. Likewise, it makes use of tools and techniques from many different quantitative fields, including algorithm design, machine learning, Bayesian and frequentist statistics, and statistical physics.

Molecular biology

Molecular biology is the study of biology at a molecular level. The field overlaps with other areas of biology and chemistry, particularly genetics and biochemistry. Molecular biology chiefly concerns itself with understanding the interactions

between the various systems of a cell, including the interrelationship of DNA, RNA and protein synthesis and learning how these interactions are regulated. Researchers in molecular biology use specific techniques native to molecular biology, but increasingly combine these with techniques and ideas from genetics and biochemistry. There is not a hard-line between these disciplines as there once was. Molecular biology is the study of molecular underpinnings of the process of replication, transcription and translation of the genetic material. The central dogma of molecular biology where genetic material is transcribed into RNA and then translated into protein, despite being an oversimplified picture of molecular biology, still provides a good starting point for understanding the field. Much of the work in molecular biology is quantitative, and recently much work has been done at the interface of molecular biology and computer science in bioinformatics and computational biology.

Genomics

Genomics is an interdisciplinary field of science within the field of molecular biology. A genome is a complete set of DNA within a single cell of an organism, and as such, focuses on the structure, function, evolution, and mapping of genomes. Genomics aims at the collective characterization and quantification of genes, which direct the production of proteins with the assistance of enzymes and messenger molecules. Genomics also involves the sequencing and analysis of genomes. Advances in genomics have triggered a revolution in discovery-based research to understand even the most currently complex biological systems such as the brain. In contrast to genetics, which refers to the study of individual genes and their roles in inheritance, genomics uses high throughput DNA sequencing and bioinformatics to assemble, and analyze the function and structure of entire genomes.

Computer sciences

The study of the theory, experimentation, and engineering that form the basis for the design and use of computers. It is the scientific and practical approach to computation and its applications and the systematic study of the feasibility, structure, expression, and mechanization of the methodical procedures (or algorithms) that underlie the acquisition, representation, processing, storage, communication of, and access to information. An alternate, more succinct definition of computer science is the study of automating algorithmic processes that scale. A computer scientist specializes in the theory of computation and the design of computational systems. Its fields can be divided into a variety of theoretical and disciplines. Some fields, such as computational complexity theory (which explores the fundamental properties of computational and intractable problems), are highly abstract, while fields such as computer graphics emphasize real-world visual applications. Other fields still focus on challenges in implementing computation. For example, programming language theory considers various approaches to the description of computation, while the study of computer programming itself investigates various aspects of the use of programming language and complex systems. Human-computer interaction considers the challenges in making computers and computations useful, usable, and universally accessible to humans.

Bioinformatics brings together large data bases of biological information and computational techniques of analysis.

What is done in bioinformatics?

The development of new algorithms and statistics with which to assess relationships among members of large data sets.

The analysis and interpretation of various types of data including nucleotide and amino acid sequences, protein domains, and protein structures.

The development and implementation of tools that enable efficient access and management of different types of information.

How do we use Bioinformatics?

Store/retrieve biological information (databases)

Retrieve/compare gene sequences

Predict function of unknown genes/proteins

Search for previously known functions of a gene

Compare data with other researchers

Compile/distribute data for other researchers

Software and tools

Software tools for bioinformatics range from simple command-line tools, to more complex graphical programs and standalone web services available from various bioinformatics companies or public institutions.

The computational biology tool best-known among biologists is probably BLAST, an algorithm for determining the similarity of arbitrary sequences against other sequences, possibly from curated databases of protein or DNA sequences.

BLAST is one of a number of generally available programs for doing sequence alignment. The NCBI provides a popular web-based implementation that searches their databases.

BLAST

Basic Local Alignment Search Tool.

It is an algorithm for comparing biological sequences information, such as amino acid sequence of different proteins or the nucleotides of DNA sequences.

BLAST is used to identify library sequences that resembles the query sequences.

The BLAST program was designed by Eugene Myers, Stephen Altschul, Warren Gish, David J. Lipman and Webb Miller at the NIH and was published in *J. Mol. Biol.* in 1990.

BLAST is a tool for alignment of sequences.

E.g. To identify the unknown gene (query sequences) in the mouse, the scientist will perform a BLAST search of the human genome (library sequences) to see whether the human carrying the similar gene or not.

BLAST was originally developed by NCBI (National Center for Biotechnology Information).

Link: <http://www.ncbi.nlm.nih.gov/BLAST/ed> to map annotations from one organism to another

Kinds of Blast

Nucleotide 6-frame translation-protein (blastx)

This program compares the six-frame conceptual translation products of a nucleotide query sequence (both strands) against a protein sequence database.

Nucleotide 6-frame translation-nucleotide 6-frame translation (tblastx)

This program is the slowest of the BLAST family. The purpose of tblastx is to find very distant relationships between nucleotide sequences.

Protein-nucleotide 6-frame translation (tblastx)

This program compares a protein query against the all six reading frames of a nucleotide sequence database.

Process

BLAST works through use of heuristic algorithm.

Heuristic algorithm, is an algorithm that is able to produce an acceptable solution to a problem in many practical scenarios.

Heuristics are typically used when there is no known method to find an optimal solution, under the given constraint.

Using a heuristic method, BLAST finds homologous sequences, not by comparing either sequences in its entirety, but rather by locating short matches between the two sequences. This process of finding initial words is called seeding.

While attempting to find homology in sequences, sets of common letters, known as words. For example, the sequences contain the following stretch of letters, GLKFA. If a BLASTp was being conducted under default conditions, the word size would be 3 letters.

In this case, using the given stretch of letters, the searched words would be GLK, LKF, and KFA.

The heuristic algorithm of BLAST locates all common words between the sequences of interest (query) and the hit sequences (sequences from database).

BLAST can be used for several purposes

Identifying Species

With the use of BLAST, you can possibly correctly identify a species and/or find homologous species. This can be useful, for example, when you are working with a DNA sequence from an unknown species.

Locating Domains

When working with a protein sequence you can input it into BLAST, to locate known domains within the sequence of interest.

Establishing Phylogeny

Using the results received through BLAST, you can create a phylogenetic tree using the BLAST web-page.

DNA mapping

When working with a known species, and looking to sequence a gene at an unknown location, BLAST can compare the chromosomal position of the sequence of interest, to relevant sequences in the database(s).

Comparison

When working with genes, BLAST can locate common genes in two related species, and can be us.

Proteome and Proteomics

The terms proteome and proteomics were coined by Mark Wilkins and colleagues in the early 1990. It is the complement protein found in a single cell in a particular environment. It is complete collection of proteins encoded by genome of an organism. It is the study of composition, structure, function and interaction of the proteins directing the activities of each living cell.

Applications of proteomics

Oncology

Oncology refers study of Tumor cell, Tumor metastasis, is the

process spread of cancer from one organ to another non-adjacent organ cause death in patients. The major challenge in medicine to describe the molecular and cellular mechanisms underlying tumor metastasis. Analyse the protein expressions correlated to the metastatic process which help to understand the mechanism of metastasis and thus facilitate the development of strategies for the therapeutic interventions and clinical management of cancer. Proteomics is a systematic research, the main aim of this research is to characterize the protein expressions, functions of tumor cells and widely used in biomarker discovery.

Bio-medical applications

The study of interactions between microbial pathogens and their hosts is called “infectomics”. It is very interesting area in proteomics. It deals with the fundamentals of the infections origin and their effect on organs. The main aim of this research is to prevent or cure disease at starting level. Advanced diagnostic issues related to emerging infections, increasing of fastidious bacteria, and generation of patient-tailored phenotypes.

Agricultural applications

Stress is a key limiting factor that impairs the growth and yield of agricultural crops. Stressful conditions often lead to delayed seed germination, reduced plant growth, and decreased crop yield. Proteins associated with the primary function of an organ are specifically accumulated in that organ/tissue or organelle. The need for organ-specific proteomic analyses to identify proteins that are commonly accumulated in organs under a wide range of abiotic stresses (Komatsu and Hossain, 2013) [8]. Jacoby *et al.*, (2013) [7] described the application of the emerging proteomic technology of multiplexed selective-reaction monitoring MS, which has increased accuracy and throughput, for enhancing these approaches and providing a clear method to rank the relative importance of the growing cohort of stress-responsive proteins. The applications of plant proteomics scientific research is still in budding stage. Proteomics is also used to know plant-insect interactions that help identify candidate genes involved in the defensive response of plants to herbivore. Population growth and effect of global climate changes imposing severe limits on the sustainability of agricultural crop production.

Food Microbiology

The use of proteomics in food technology is presented especially for characterization and standardization of raw materials, process development, and detection of batch-to batch variations and quality control of the final product. Further attention is paid to the aspects of food safety, especially regarding biological and microbial safety and the use of genetically modified foods.

MS analysis

The protein mixture is subjected to digestion so as to get peptide mixture, from which a specific peptide is isolated and then identified. Mass spectrometry is an analytical technique that produces spectra of the masses of the atoms or molecules comprising a sample of material. The spectra are used to determine the elemental or isotopic signature of a sample, the masses of particles and of molecules, and to elucidate the chemical structures of molecules, such as peptides and other chemical compounds. Mass spectrometry works by ionizing

chemical compounds to generate charged molecules or molecule fragments and measuring their mass to charge ratios. MALDI-TOF is the most useful technique for protein identification.

Human Genome Project

Genomics is the study of the genome of an organism - its entire genetic material in the form of RNA, DNA, genes and chromosomes. It concentrates on understanding the structure and function of an organism's genetic material from the molecular level upwards, including interactions between genes, interactions between genes and the proteins they produce, and interactions between genes and environmental factors.

Human Genome Project (HGP) is the name for an international consortium of publicly funded projects to sequence the human genome and map every gene on every chromosome. The consortium includes the US Department of Energy, the US National Institutes of Health and the UK Wellcome Trust, alongside groups in Japan, France, Germany, China, and other countries. HGP Began in 1990 and was successfully completed in April 2003.

The main aims of the Human Genome Project are to Determine the sequence of the three billion or so nucleotides that constitute the human genome.

Identify the 20,000 to 25,000 genes in the human genome.

Develop tools for storing and analyzing this information.

Transfer some of the technologies involved to the private sector, to produce a biotechnology industry that can develop new medical applications.

Examine the ethical, social and legal implications of the information obtained.

The HGP uses the so-called hierarchical shotgun sequencing technique, in which the genome is divided into relatively large sections that are mapped onto the appropriate chromosomes before being sequence. Every single cell in the body contains a complete copy of the approximately 3 billion DNA base pairs.

Life Science Vs Computer Science

Scientific algorithms made it possible to put together a vast amount of data from sequencing machines when the human genome was sequenced. Computer science's computational paradigm has shaped new modes of inquiry in life sciences.

Applications of bioinformatics in various fields

Insect Resistance

Many plants have been made insect resistant by incorporating the desired genes. *Bacillus thuringiensis* is a bacterial species which increases the soil fertility and protects the plants against pests. When the researchers mapped its genome, they used its genes to incorporate into the plant to make it resistant against insects. For example, corn, cotton and potatoes have been made insect resistant so far. By having the genes of bacteria in the plants genome, when insects eat the plants, the bacteria enter in their bloodstream and make them starved, ultimately they die. B.t corn is one species of food plants which have been modified by inserting bacterial genes in it. It is effective against insects by developing resistance against them. The use of B.t genes in the plants genome has made the agriculturists to use the insecticides in very little amount. As a result the productivity and nutritional value of plants will also increase and will be beneficent for human health.

Nutritional Quality

When the changes are made in the genome of the plants, the nutritional value of plants also increases. Golden rice is an important achievement in this endeavour. Here the genes are inserted in the rice genome to increase the Vitamin A level in the crop. Vitamin A is an important component for the eyes and if the Vitamin A deficiency occurs in the body, it may result in blindness. This work has allowed the scientists to reduce the rate of blindness from the world by giving genetically modified rice to the people.

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. Bioinformatics tools are important for understanding of the mechanisms of bio degradative pathways (Sadraeian and Molaei, 2009) ^[10].

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 (Allaby and Woodmark, 2004) ^[11].

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 (Sinha, 2015) ^[11].

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 (Valdivia-Granda, 2010) ^[13].

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 (Bianchi and Lio, 2007) ^[3]. Scientists used their genomic tools to distinguish between the strains of *Bacillus anthracis* that was used in the summer of 2001 terrorist attack in Florida with that of closely related anthrax strains.

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 (Misra *et al.*, 2013) ^[9].

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 (Kadarmideen, 2014).

Plant breeding and crop improvement

Plant genomics helps in understanding the genetic and molecular basis of all biological process which helps in developing new cultivars with improved quality and reduced economic and environmental cost. Now-a-days the Genome program is an important tool for the plant improvement. This genome programme helps in identifying the key genes and their function. This genome project generates data, which includes sequence information, markers etc. These data are then distributed to the multinational research community. The bioinformatics tools helps in the submission of all data through ENTREZ Global Query Cross-Database Search System to the public domain. This helps in retrieving sequence from the NCBI. Genome sequencing of several important plants species has enabled researchers to identify ‘chromosome’ and ‘difference’ factor in sequences. This in turn has been used to identify value traits for crop improvement. For instance, the barley stem rust resistance gene has been identified from rice-barley comparisons and the sugarcane rust resistance gene based on maize-sorghum comparisons.

Development of stress tolerant varieties

Stress tolerant varieties can be developed by identifying the Stress tolerance genes and alleles. Various tools have been developed to study the physiology, expression profiling, comparative genomics. The KEGG database contains all the

metabolic pathways like the pathway for the carbohydrate production. Genes in the ABA production pathway are important for the development of drought resistant varieties. KEGG databases can be important in identifying the pathway for carbohydrate production and ABA production. After the identification of the pathway the genes involved in the same pathway are studied for the development.

Progress has been made in developing cereal varieties that have a greater tolerance for soil alkalinity, free aluminium and iron toxicities. These varieties will allow agriculture to succeed in poorer soil areas, thus adding more land to the global production base. Research is also in progress to produce crop varieties capable of tolerating reduced water conditions. The publication of the completed *Arabidopsis thaliana* genome sequence (Arabidopsis Genome Initiative, 2000) [2] and draft sequence for rice genome (Goff *et al.*, 2002) [5], the plant research and industry has step over the threshold of the genomics era.

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 (Hack and Kendall, 2013) [6]. 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 are gene addition, removal of harmful gene by antisense nucleotide or ribozymes and control of gene expression.

Alternative energy

With the help of bioinformatics scientists are studying the genome of the microbe *Chlorobium tepidum* which has an unusual capacity for generating energy from light.

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 (Zhang and Hong, 2015) [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 (Tiwari, 2015) [12].

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 (Hack and Kendall, 2013) [6].

Molecular medicine

Every disease has a genetic component. There are 3000-4000 hereditary disease including Cystic Fibrosis and Huntingtons disease) or a result of the body's response to an environmental stress which causes alterations in the genome (e.g. cancers, heart disease, diabetes). Search for the genes directly associated with different diseases and begin to understand the molecular basis of these diseases more clearly. This new knowledge of the molecular mechanisms of disease will enable better treatments, cures and even preventative tests to be developed.

Biotechnology

The archaeon *Archaeoglobus fulgidus* and the bacterium *Thermotoga maritima* have potential for practical applications in industry and government-funded environmental remediation. These microorganisms thrive in water temperatures above the boiling point and therefore may provide the DOE, the Department of Defence, and private companies with heat-stable enzymes suitable for use in industrial processes. Other industrially useful microbes include, *Corynebacterium glutamicum* which is of high industrial interest as a research object because it is used by the chemical industry for the biotechnological production of the amino acid lysine. The substance is employed as a source of protein in animal nutrition. Lysine is one of the essential amino acids in animal nutrition. Biotechnologically produced lysine is added to feed concentrates as a source of protein, and is an alternative to soybeans or meat and bonemeal. *Xanthomonas campestris* is grown commercially to produce the exopolysac charidexanthan gum, which is used as

a viscosifying and stabilizing agent in many industries. *Lactococcus lactis* is one of the most important microorganisms involved in the dairy industry, it is a non-pathogenic rod-shaped bacterium that is critical for manufacturing dairy products like buttermilk, yogurt and cheese. This bacterium, *Lactococcus lactis* is also used to prepare pickled vegetables, beer, wine, some breads and sausages and other fermented foods. Researchers anticipate that understanding the physiology and genetic make-up of this bacterium will prove invaluable for food manufacturers as well as the pharmaceutical industry, which is exploring the capacity of *L. lactis* to serve as a vehicle for delivering drugs.

Drug development

At present all drugs on the market target only about 500 proteins. With an improved understanding of disease mechanisms and using computational tools to identify and validate new drug targets, more specific medicines that act on the cause, not merely the symptoms of the disease can be developed. These highly specific drugs promise to have fewer side effects than many of today's medicines.

Antibiotic resistance

Scientists have been examining the genome of *Enterococcus faecalis* a leading cause of bacterial infection among hospital patients. They have discovered a virulence region made up of a number of antibiotic-resistant genes that may contribute to the bacterium's transformation from a harmless gut bacteria to a menacing invader. The discovery of the region, known as a pathogenicity island, could provide useful markers for detecting pathogenic strains and help to establish controls to prevent the spread of infection inwards.

The reality of bioweapon creation

Scientists have recently built the virus poliomyelitis using entirely artificial means. They did this using genomic data available on the Internet and materials from a mail-order chemical supply. The research was financed by the US Department of Defence as part of a bio warfare response program to prove to the world the reality of bioweapons. The researchers also hope their work will discourage officials from ever relaxing programs of immunization.

Microbial genome applications

Microorganisms are ubiquitous, that is they are found everywhere. They have been found surviving and thriving in extremes of heat, cold, radiation, salt, acidity and pressure. They are present in the environment, our bodies, the air, food and water. Traditionally, use has been made of a variety of microbial properties in the baking, brewing and food industries. The arrival of the complete genome sequences and their potential to provide a greater insight into the microbial world and its capacities could have broad and far reaching implications for environment, health, energy and industrial applications. For these reasons, in 1994, the US Department of Energy (DOE) initiated the MGP (Microbial Genome Project) to sequence genomes of bacteria useful in energy production, environmental cleanup, industrial processing and toxic waste reduction. By studying the genetic material of these organisms, scientists can begin to understand these microbes at a very fundamental level and isolate the genes that give them their unique abilities to survive under extreme conditions.

Conclusions

Bioinformatics joins mathematics, statistics, and computer science and information technology to solve complex biological problems. The tools of bioinformatics are helpful in every field of life. In agriculture the genome sequencing of the plants and animals provides number of benefits. Serioinformatics has energized research and fostered a surge of new ideas for mulberry and silkworm improvement. The ultimate goal of bioinformatics is to integrate large-scale data for understanding the molecular mechanism involved in various developmental processes. This understanding can help in producing the Climate Smart agricultural crops. Application of various bioinformatics tools in biological research enables storage, retrieval, analysis, annotation and visualization of results and promotes better understanding of biological system in fullness. This will help in animal and plant health care based disease diagnosis.

Conflict of interest

There is no conflict of interest among the authors.

Acknowledgement

Authors are highly thankful to TSRI, Mirgund-SKUAST-K for the internet facilities.

Reference

- Allaby RG, Woodwark M. Phylogenetics in the bioinformatics culture of understanding. *Comp Funct Genomics*. 2004; 5:128-46.
- Arabidopsis Genome Initiative. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*. 2000; 408(6814):796-815.
- Bianchi L, Lio P. Forensic DNA and bioinformatics. *Brief Bioinform*. 2007; 8(2):117-28.
- Dahiya BL, Lata M. Bioinformatics impacts on medicine, microbial genome and agriculture. *Journal of Pharmacognosy and Phytochemistry*. 2017; 6(4):1938-1942.
- Goff SA, Ricke D, Lan Presting TH, Wang G, Dunn RM. A Draft Sequence of the Rice Genome (*Oryza sativa* L. ssp. *japonica*). *Science*. 2002; 296(5565):92-100.
- Hack C, Kendall G. Bioinformatics: current practice and future challenges for life science education. *Biochem. Molecular Biol. Educ*. 2013; 33:82-85.
- Jacoby RP, Millar AH, Taylor NL. Application of selected reaction monitoring mass spectrometry to field-grown crop plants to allow dissection of the molecular mechanisms of abiotic stress tolerance. *Front. Plant Sci*. 2013; 4:20.
- Komatsu S, Hossain Z. Organ-specific proteome analysis for identification of abiotic stress response mechanism in crop. *Front. Plant Sci*. 2013; 4:71.
- Misra N, Panda PK, Parida BK. Agrigenomics for microalgal biofuel production: An overview of various bioinformatics resources and recent studies to link OMICS to bioenergy and bio economy. *OMICS*. 2013; 17:537-49.
- Sadraeian M, Molaei Z. Bioinformatics Analyses of *Deinococcus radiodurans* in order to waste clean-up. In environmental and computer science. Second International Conference, 2009, 254.
- Sinha S. Role of bioinformatics in climate change

- studies. *J Science*. 2015; 1:1-9.
12. Tiwari HA. 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.
 13. Valdivia-Granda WA. Bioinformatics for biodefense: Challenges and opportunities. *Biosecur Bioterror*. 2010; 8:69-77.
 14. Zhang L, Hong H. Genomic discoveries and personalized medicine in neurological diseases, pharmaceuticals. 2015; 7:542-553.