BIOINFORMATICS – DEFINING LIFE WITH COMPUTERS

Radhika G.
and Shynu M.
Assistant Professors,
College of Veterinary & Animal Sciences,
Pookot

Bioinformatics is the field of science in which biology, computer science and information technology merges to form a single discipline. Immense amount of biological information produced in the recent years, demanded computerized databases to store, organize and index the data, and for specialized tools to view and analyse the data. This led to the rapid development of this field. Infact the first database was created shortly after sequence of insulin was deciphered in 1956. Bioinformatics has at least five major types of activities: data acquisition, database development, data analysis, data integration, and analysis of integrated data. Data gatherers, processors and process product users work hand in hand to analyse and interpret various types of data, including nucleotide and amino acid sequences, protein domains, protein structures etc. Actual process of analyzing and interpreting data is referred to as computational biology. Thus Bioinformatics is conceptualizing biology in terms of molecules and applying "information techniques" to understand and organize the information associated with these molecules. Major applications of Bioinformatics are in the areas of data mining, sequence analysis, phylogenetics, protein structure prediction, genomics, proteomics, drug discovery etc.

Databases:

Primary sequence databases are repositories for raw sequence data and can be accessed freely over the worldwide web. Three such databases are GenBank maintained by NCBI (National Centre for Biotechnology Information), Nucleotide Sequence Laboratory by EMBL (European Molecular Biology Lab), and DNA Databank of Japan (DDBJ)

Protein Database contains sequence data from the translated coding regions from DNA sequence in GenBank, EMBL, DDBJ as well as protein sequence submitted to Protein Information Resource (PIR); SWISSPROT, Protein Data Bank (PDB) etc.Other databases include genome database, structure database, taxonomy database etc.

Database mining and sequence analysis:

Database mining is the process by which testable hyities between genes and between proteins. Algorithms were developed to help researchers rapidly identify similar gene or protein sequences. Such tools were extremely useful for determining whether a newly sequenced piece of DNA was at all similar to sequences already entered in a database.

To determine how multiple sequences align and to view their similarities either locally or over their entire length, multiple alignment programs were developed. Such programs helped scientists compare the sequences of closely related genes or compare the sequence of a particular gene or protein as it appears in several species. Programs widely used for this purpose are FASTA (Fast Approximation) and BLAST. (Basic Local Alignment Search Tool)

Multiple sequence alignments which optimize the alignment of several homologs can be used to search for patterns of highly conserved residue positions. Multiple Sequence Alignment (MSA) is a tool to determine levels of homology and hence relatedness between members of a series of globally related sequences. Programs used are CLUSTAL, PILEUP, MACAW etc.

Evolutionary Biology & Phylogenetics

Phylogenetics is the field of biology dealing with identifying and understanding the relationship between different kinds of life on earth.MSA mentioned above is very important for doing phylogenetic analysis, because the accuracy with which structural and functional attributes, can be inherited between relatives depends on the evolutionary distance between those relatives. With the help of nucleotide and protein sequences, it is possible to find the ancestral ties between different organisms. The more closely related the species are, the more likely they share a recent common ancestor. Thus, similarities and differences between organisms can be used to infer phylogenies (evolutionary relationships). Commonly used program is PHYLIP in which evolutionary relationships are visualized using graphs and trees, which maybe rooted or unrooted with scaled or unscaled branches.

Protein structure prediction and modeling

Knowing the structure of a protein can give an immense amount of information about functioning of a protein. Here bioinformatics has an enormous analytical and predictive potential. Comparative modeling or homology modeling is the most reliable technique for predicting protein structure. Experimentally determined protein structures (templates) are used to predict the three dimensional structure of another protein (target), that has a similar amino acid sequence. Here the structure of a new protein is predicted by comparing its sequences with the sequence of

protein of known structure. If there is strong similarity, it can be assumed that proteins have similar overall structure. It provides a starting point for researchers, wishing to confirm a structure through X- ray crystallography and NMR spectroscopy. Protein modeling is an important tool for scientists to understand about normal and disease-related processes in living organism.

Protein structure databases include PROWL and CDD (Conserved Domain Database). RasMol is popularly used for protein structure visualization. VAST (Vector Alignment Search Tool) is used to identify similar protein three dimensional structures and DART is used to search for proteins with similar domain architecture.

Genomics and Proteomics tools:

Genome of a typical bacterium comes as a single DNA molecule, which if extended would be about 2mm long. The DNA of higher organisms is organized into chromosomes. Genomics starts with the gene and makes inferences about its products. It has provided a vast amount of information which links gene activity with disease. In the areas of structural genomics, functional genomics and nutritional genomics, bioinformatics plays a vital role.

Proteomics begins with functionally modified protein and works back to the gene responsible for its production. Protein in an organism change during growth, disease and death of cells and tissues. Modification of protein, which occurs during and after their synthesis, changes the proteome complement. To discover and monitor the relevance of a protein to a biological process in the cell, it is important to identify where, when and to what extent a protein is expressed. Bioinformatics tools can be widely used for this purpose.

Drug Discovery

Bioinformatics tools help us to structurally modify natural products with medicinal potential into a new drug with desired properties. Assessment of their therapeutic effects can also be done theoretically, *in silico.*It is now possible, through computer algorithm based bioinformatic procedures, to develop patient specific designer drugs, which are free of untoward side effects. These applications and pharmacogenomics are some of the areas in which bioinformatics has become an integral part of Research & Development. The duration of a new drug development which takes 15-20 years by conventional methods, can be considerably reduced with the advent of bioinformatics tools.

The real potential of Bioinformatics lies in its ability to redefine life. This subject mainly focuses on the issues of how to describe and analyse the fundamental life processes in cells and organisms. With the aid of bioinformatics we are on the verge of controlling the coding of all living things. But one should always keep in mind the fact that, it can only suggest certain hypotheses, which should definitely be tested by wet lab experiments. This new field helps life scientists to delve into the complexities of life and apply successfully new techniques to solve the mysteries of mankind.

❏
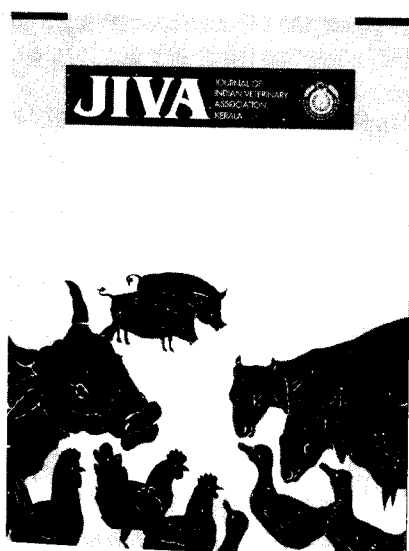
Cover Page of First Issue of
JIVA Realesed on
29 December 1996

INFOMANIA

ANSWERS
1. Drey
2. Cat
3. Golf
4. Tribology
5. Vegans
6. Viticulture
7. Animal Farm
8. Titanic
9. Hyena Disease
10. Acute abdominal pain
11. Pica
12. Fever
13. Selenium and Vitamin E
14. Cuisine
15. Heterakis
16. Taenia saginata
17. Edible testicles
18. Shackling

Complied by : **Dr. Jayesh V.**, Veterinary Surgeon, District Veterinary Centre, Kalpetta, Wayanad