

Bioinformatics in the present day

Commentary

Bioinformatics is an interdisciplinary field that applies information technology towards understanding, organising, storing and retrieving the information associated with macromolecule on a large-scale. As an interdisciplinary field of science, bioinformatics combines computer science, statistics, mathematics, engineering and biology. Bioinformatics is at the heart of modern biological research, especially so in whole genome studies and second generation sequencing (2GS) data analysis. Generally bioinformatics include analysing, mapping and storing of protein and DNA sequences, aligning DNA and protein sequences to compare them, and creating and viewing 3-D models of protein structures. As DNA is the basic molecule of life coding for genes which intern code for proteins and proteins determine the biological makeup of all the living organisms. It is errors and variations in the gene that is responsible for the development of a disease or resistance against the same disease. After sequencing the human genome it was necessary to store that data for further research and for scientist all around the world to use that data for their research, bioinformatics tools plays a major role here by allowing the scientist to store the genomic sequence and made it very easy to compare one genome sequence with the other genomic sequences. The complete human genome sequence gained out of Human Genome Project has enabled to improve the biological research and clinical medicine, now scientists are able to find out the cure for several inherited (Huntington's disease and cystic fibrosis) and acquired (heart disease or cancer) diseases due to complete knowledge on human genome.

The bioinformatics software's and tools have helped in drug discovery as they help in understanding protein-protein, protein-ligand and various other interactions. Drug discovery was a very complicated and time consuming process requiring keen observation and thorough knowledge. But now with the developments in bioinformatics, drug discovery have become very easy as various target sights and target bound drugs could be discovered within a short span of time. RNA and DNA are the proteins that store the hereditary information about an organism have a fixed structure, which can be analysed by biologists with the help of bioinformatics databases and tools.

Database

Biological database are just like the library of biological information, collected by scientific experiments, high-throughput experiment technology, published literature and computational analysis. Biological database contain information from various research areas such as, proteomics, metabolomics, genomics, microarray gene expression and phylogenetics. The information present in the biological database includes gene function, structure, localization (both cellular and chromosomal), clinical effects of mutations as well as similarities of biological structures and sequences. A simple database may contain many records of a single file, each of which includes the same set of information. For example, a record associated with a nucleotide sequence database typically contains information such as the scientific name of the source organism from which it was isolated; contact name; the input sequence with a description of the type of molecule; and, often, literature citations associated with the sequence.

Volume 3 Issue 1 - 2016

Syed Rizwan Hassan Razvi, Shailima Rampogu

Department of Biochemistry, Celesta Labs, India

Correspondence: Shailima Rampogu, Celesta Labs, Department of Biochemistry, Yousufguda, Hyderabad, Telangana, India, Tel 9666440706, Email shailima.rampogu@gmail.com

Received: January 25, 2016 | **Published:** January 29, 2016

Biological databases can be broadly classified into structure and sequence databases. Nucleic acid and protein sequences are stored in sequence databases and structure database only store structure of proteins. A few popular databases are GenBank from NCBI (National Centre for Biotechnology Information), SwissProt from the Swiss Institute of Bioinformatics and PIR from the Protein Information Resource.

Genbank

(Genetic Sequence Databank) is one of the fastest growing repositories of known genetic sequences. In addition to sequence data, GenBank files contain information like accession numbers and gene names, phylogenetic classification and references to published literature. It has a flat file structure that is an ASCII text file, readable by both computers and humans.

Swissprot

This is a protein sequence database that provides a high level of integration with other databases and also has a very low level of redundancy (means less identical sequences are present in the database).

Tools

Bioinformatics tools are software programs which are designed to extract the meaningful information from the mass of molecular biology/biological databases and also to carry out sequence or structural analysis. These software tools are made available over the internet given it a global distribution of the scientific research community. These tools retrieve data from genomic sequence databases and are also used to visualise, analyse and retrieve information from proteomic databases. Bioinformatics tools can be broadly classified as homology and similarity tools, protein functional analysis tools, sequence analysis tools and miscellaneous tools.

Homology and similarity tools

This set of tools can be used to identify similarities between the sequences of unknown structural and functional sequences whose structure and function have already been known. Homologous sequences are sequences that are related by divergence from a common ancestor.

Protein function analysis

This programs allow us to compare our protein sequence to the secondary (or derived) protein. Highly significant hits against these

different pattern databases allow us to approximate the biochemical function of a query protein.

Structural analysis

This set of tools helps us to compare the structures with the known structure databases. The function of a protein mostly depend on its structure than its sequence. The determination of a protein's 2D/3D structure is crucial in the study of its function.

Sequence analysis

These tools helps us to carry out further, more detailed analysis on a query sequence including evolutionary analysis, identification of hydrophathy regions, mutations, compositional biases and CpG islands. The identification of these and other biological properties are all clues that aid the search to elucidate the specific function of a selected sequence.

Systems biology

It is an emerging engineering approach applied to biological scientific research, systems biology is a biology-based interdisciplinary field of study that focuses on complex interactions within biological systems, using a holistic approach (holism instead

of the more traditional reductionism) to biological research. Systems biology is the study of systems of biological components, such as cells, organisms, molecules or entire species. To study them, we use systematic measurement technologies, quantitative measurements of the behaviour of groups of interacting components such as genomics, bioinformatics and proteomics, and mathematical and computational models to describe and predict dynamical behaviour. Systems problems are emerging as central to all areas of biology and medicine.

Summary

In summary it can be stated that bioinformatics has been proven to be one of the most sophisticated and highly appreciated fields in the present day. Its helps the researchers and scientists to analyse and understand certain complicated pathways in a very simpler manner making the study more interesting. Not only that, it also allows one to understand the drug interactions and their toxicological studies.

Acknowledgements

None.

Conflict of interest

The author declares no conflict of interest.