

What is Bioinformatics?

Bioinformatics has evolved into a full-fledged multidisciplinary subject that integrates developments in information and computer technology as applied to Biotechnology and Biological Sciences. Bioinformatics uses computer software tools for database creation, data management, data warehousing, data mining and global communication networking.

Bioinformatics is the recording, annotation, storage, analysis, and searching/retrieval of nucleic acid sequence (genes and RNAs), protein sequence and structural information. This includes databases of the sequences and structural information as well methods to access, search, visualize and retrieve the information. Bioinformatics concern the creation and maintenance of databases of biological information whereby researchers can both access existing information and submit new entries. Function genomics, biomolecular structure, proteome analysis, cell metabolism, biodiversity, downstream processing in chemical engineering, drug and vaccine design are some of the areas in which Bioinformatics is an integral component.

Sub-disciplines within bioinformatics

There are three important sub-disciplines within bioinformatics involving computational biology:

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

Activities in bioinformatics

We can split the activities in bioinformatics in two areas (1) the organization and (2) the analysis of biological data

Organization activity in Bioinformatics

- The creation of databases of biological information
- The maintenance of these databases

Analysis activity in Bioinformatics

- Development of methods to predict the structure and/or function of newly discovered proteins and structural RNA sequences.
- Clustering protein sequences into families of related sequences and the development of protein models.
- Aligning similar proteins and generating phylogenetic trees to examine

evolutionary relationships

Aims of Bioinformatics:

The aims of bioinformatics are basically three-fold. They are

- Organization of data in such a way that it allows researchers to access existing information & to submit new entries as they are produced. While data-creation is an essential task, the information stored in these databases is useless unless analyzed. Thus the purpose of bioinformatics extends well beyond mere volume control.
- To develop tools and resources that help in the analysis of data. For example, having sequenced a particular protein, it is with previously characterized sequences. This requires more than just a straightforward database search. As such, programs such as FASTA and PSI-BLAST much consider what constitutes a biologically significant resemblance. Development of such resources extensive knowledge of computational theory, as well as a thorough understanding of biology.
- Use of these tools to analyze the individual systems in detail, and frequently compared them with few that are related.

Three levels of bioinformatics

1. Analysis of a single gene (protein) sequence. For example:
 - Similarity with other known genes
 - Phylogenetic trees; evolutionary relationships
 - Identification of well-defined domains in the sequence
 - Sequence features (physical properties, binding sites, modification sites)
 - Prediction of subcellular localization
 - Prediction of secondary and tertiary structure
2. Analysis of complete genomes. For example:
 - Which gene families are present, which missing?
 - Location of genes on the chromosomes, correlation with function or evolution
 - Expansion/duplication of gene families
 - Presence or absence of biochemical pathways
 - Identification of "missing" enzymes
 - Large-scale events in the evolution of organisms
3. Analysis of genes and genomes with respect to functional data. For example:
 - Expression analysis; microarray data; mRNA conc. measurements
 - Proteomics; protein conc. measurements, covalent modifications
 - Comparison and analysis of biochemical pathways

- Deletion or mutant genotypes vs. phenotypes
- Identification of essential genes, or genes involved in specific processes

Modified From: Suresh Kumar (2005). Bioinformatics web. Retrieved June 11, 2015
from: <http://www.geocities.com/bioinformaticsweb/>