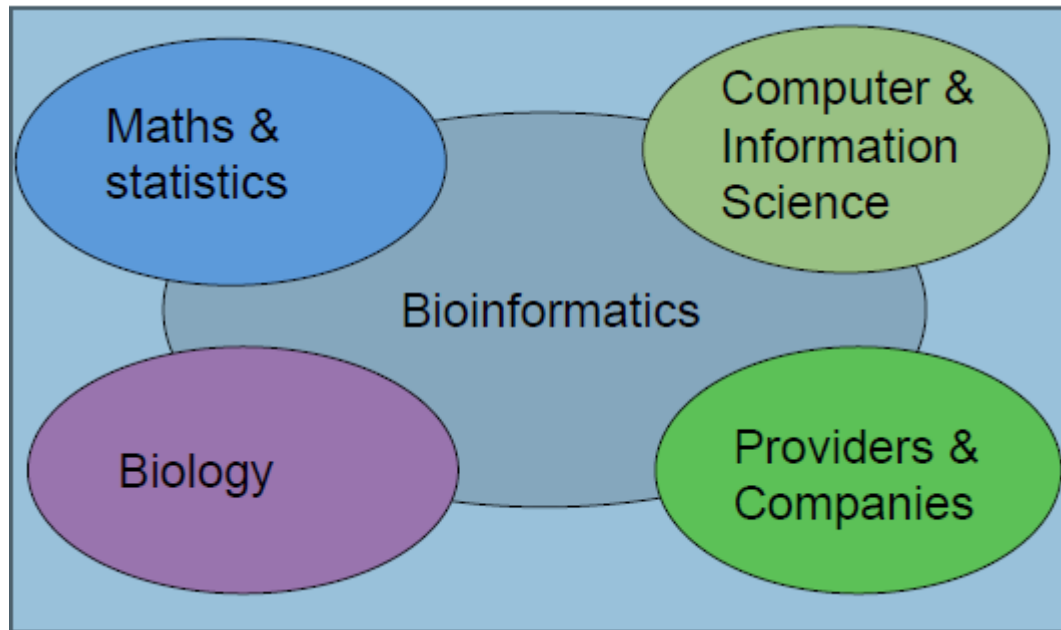


Introduction to Bioinformatics

WHAT IS BIOINFORMATICS?

Bioinformatics is an interdisciplinary research area at the interface between computer science and biological science. A variety of definitions exist in the literature and on the World Wide Web; some are more inclusive than others. Here, we adopt the definition proposed by Luscombe et al. in defining bioinformatics as a union of biology and informatics: *bioinformatics* involves the technology that uses computers for storage, retrieval, manipulation, and distribution of information related to biological macromolecules such as DNA, RNA, and proteins. The emphasis here is on the use of computers because most of the tasks in genomic data analysis are highly repetitive or mathematically complex. The use of computers is absolutely indispensable in mining genomes for information gathering and knowledge building. Bioinformatics differs from a related field known as *computational biology*. Bioinformatics is limited to sequence, structural, and functional analysis of genes and genomes and their corresponding products and is often considered *computational molecular biology*. However, computational biology encompasses all biological areas that involve computation. For example, mathematical modeling of ecosystems, population dynamics, application of the game theory in behavioral studies, and phylogenetic using fossil records all employ computational tools, but do not involve biological macromolecules.

Beside this distinction, it is worth noting that there are other views of how the two terms relate. For example, one version defines *bioinformatics* as the development and application of computational tools in managing *all kinds* of biological data, whereas *computational biology* is more confined to the theoretical development of algorithms used for bioinformatics. The confusion at present over definition may partly reflect the nature of this vibrant and quickly evolving new field.



GOALS

The ultimate goal of bioinformatics is to better understand a living cell and how it functions at the molecular level. By analyzing raw molecular sequence and structural data, bioinformatics research can generate new insights and provide a “global” perspective of the cell. The reason that the functions of a cell can be better understood by analyzing sequence data is ultimately because the flow of genetic information is dictated by the “central dogma” of biology in which DNA is transcribed to RNA, which is translated to proteins. Cellular functions are mainly performed by proteins whose capabilities are ultimately determined by their sequences. Therefore, solving functional problems using sequence and sometimes structural approaches has proved to be a fruitful endeavor.

Bioinformatics consists of two subfields:

1-the development of computational tools and databases

2-the application of these tools and databases in generating biological knowledge to better understand living systems.

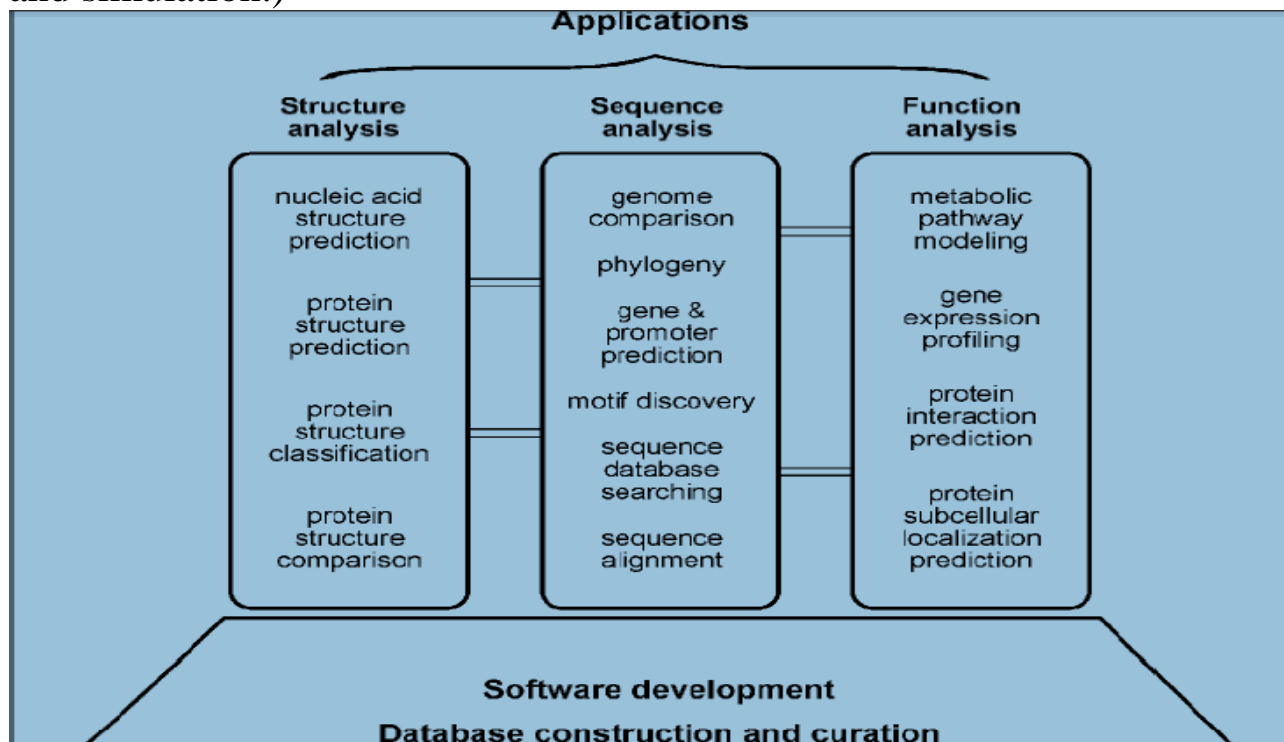
The tool development includes writing software for sequence, structural, and functional analysis, as well as the construction and curating of biological databases.

These tools are used in three areas of genomic and molecular biological research:

1-molecular sequence analysis (the areas of sequence analysis include sequence alignment, sequence database searching , motif and pattern discovery, gene and promoter finding, reconstruction of evolutionary relationships, and genome assembly and comparison.)

2-molecular structural analysis (structural analyses include protein and nucleic acid structure analysis, comparison, classification, and prediction.)

3-molecular functional analysis (the functional analyses include gene expression profiling, protein– protein interaction prediction, protein subcellular localization prediction, metabolic pathway reconstruction, and simulation.)



APPLICATIONS: -

- 1- Knowledge-based drug design.
- 2- Forensic DNA analysis.
- 3- Agricultural biotechnology.
- 4-Computational studies of protein–ligand interactions provide a rational basis for the rapid identification of novel leads for synthetic drugs.
- 5-Knowledge of the three-dimensional structures of proteins allows molecules to be designed that are capable of binding to the receptor site of a target protein with great affinity and specificity.
- 6-Sequence analysis determines those genes which encode regulatory sequences or peptides by using the information of sequencing.
- 7-In genome annotation, genomes are marked to know the regulatory sequences and protein coding. It is a very important part of the human genome project as it determines the regulatory sequences.
- 8-Comparative genomics which determines the genomic structure and function relation between different biological species.