

Bioinformatics

Microbiology VI

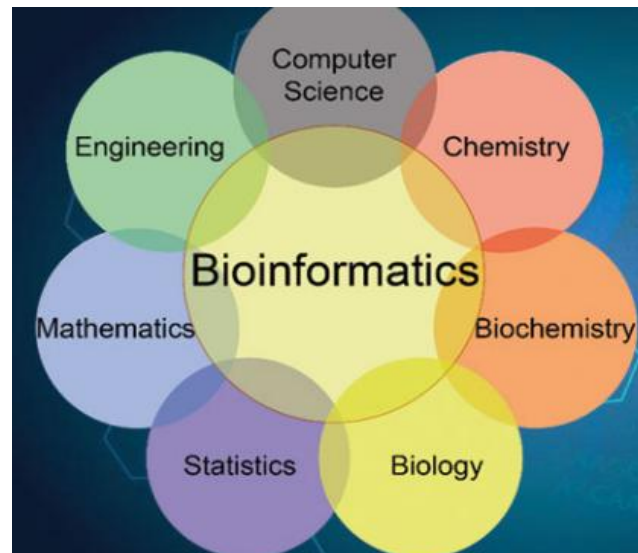
Introduction

- As a large number of prokaryotic and eukaryotic genomes are completely sequenced and more added,
- Access to the **genomic information** and synthesizing it for the discovery of new knowledge have become central themes of modern biological research.
- Mining the genomic information requires the use of sophisticated **computational tools**.

- It is therefore becomes imperative for the new generation of biologists to initiate and familiarize with a field of study that is concerned with the careful **storage, organization and indexing** of information in order to tackle the new challenges in the genomic era.
- Information science has been applied to biology to produce a field is called **bioinformatics**.
- It is concerned with the state of- the-art computational tools available to solve biological research problems.

Bioinformatics:

- The term “**bioinformatics**” was coined by Paulien Hogeweg and Ben Hesper
- To describe “the study of informatic processes in biotic systems”
- It found early use when the **first biological sequence** data began to be shared.
- **Bioinformatics** is an interdisciplinary field that develops methods and software tools for understanding biological data.

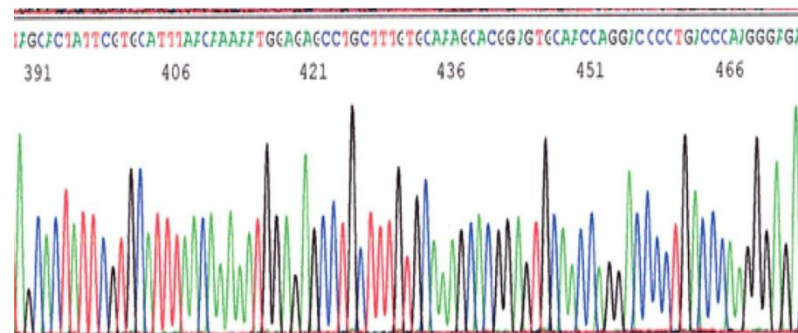


- As an interdisciplinary field of science, bioinformatics combines
 - biology
 - computer science
 - information engineering
 - mathematics and
 - statistics
- to analyze and interpret biological data.
- The key areas of **bioinformatics** include :
 - biological databases
 - sequence alignment
 - gene and promoter prediction
 - molecular phylogenetics
 - structural bioinformatics
 - genomics and
 - proteomics.

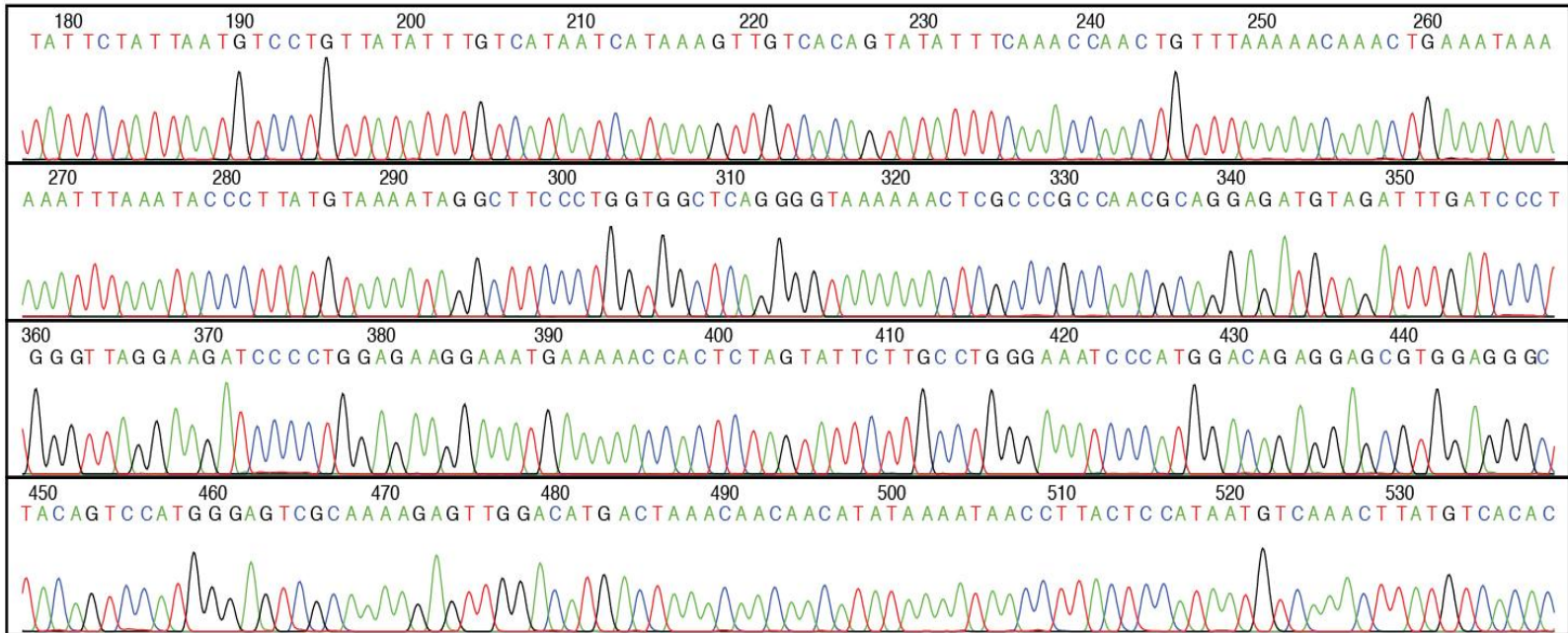
Genomics:

- Is the study of whole genomes of organisms, and incorporates elements from genetics.
- Genomics uses a combination of:
 - recombinant DNA
 - DNA sequencing methods and
 - bioinformatics
- to sequence, assemble, and analyse the structure and function of genomes.
- It is an interdisciplinary field of biology focusing on the structure, function, evolution, mapping, and editing of genomes.

- A genome is an organism's **complete set of DNA**, including all of its genes.
- Genomics aims at:
 - The collective **characterization** and **quantification** of all of an organism's genes
 - Their interrelations and influence on the organism.

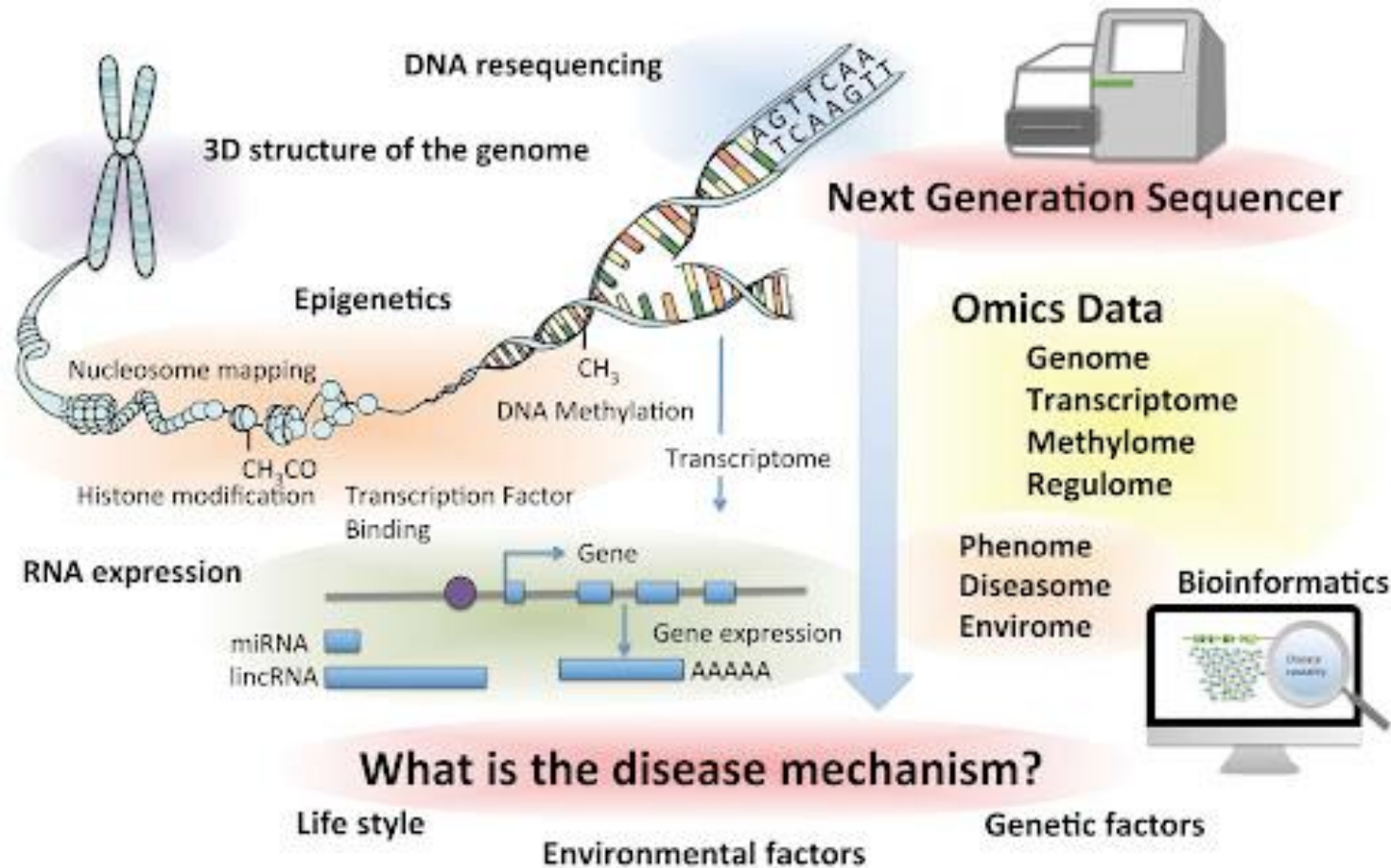


Genome sequencing



DNA sequence data from an automated sequencing machine

- **Genes** may direct the production of **proteins** with the assistance of enzymes and messenger molecules.
- In turn, proteins make up **body structures** such as organs and tissues as well as control **chemical reactions** and carry signals between cells.
- Genomics also involves the **sequencing** and **analysis** of **genomes**
 - Through uses of high throughput DNA sequencing and bioinformatics
 - To assemble and analyze the function and structure of entire genomes.



Proteomics:

- Is the large-scale study of proteins, their **structure** and physiological role or **functions**.
- It involves the **analysis** of the entire protein complement of a cell, tissue, or organism under a specific, defined set of conditions.
- Proteins are vital parts of living organisms, with many functions.
- The **proteome** is the entire set of proteins that is produced or modified by an organism or system.
- Proteomics has enabled the identification of ever increasing numbers of protein.

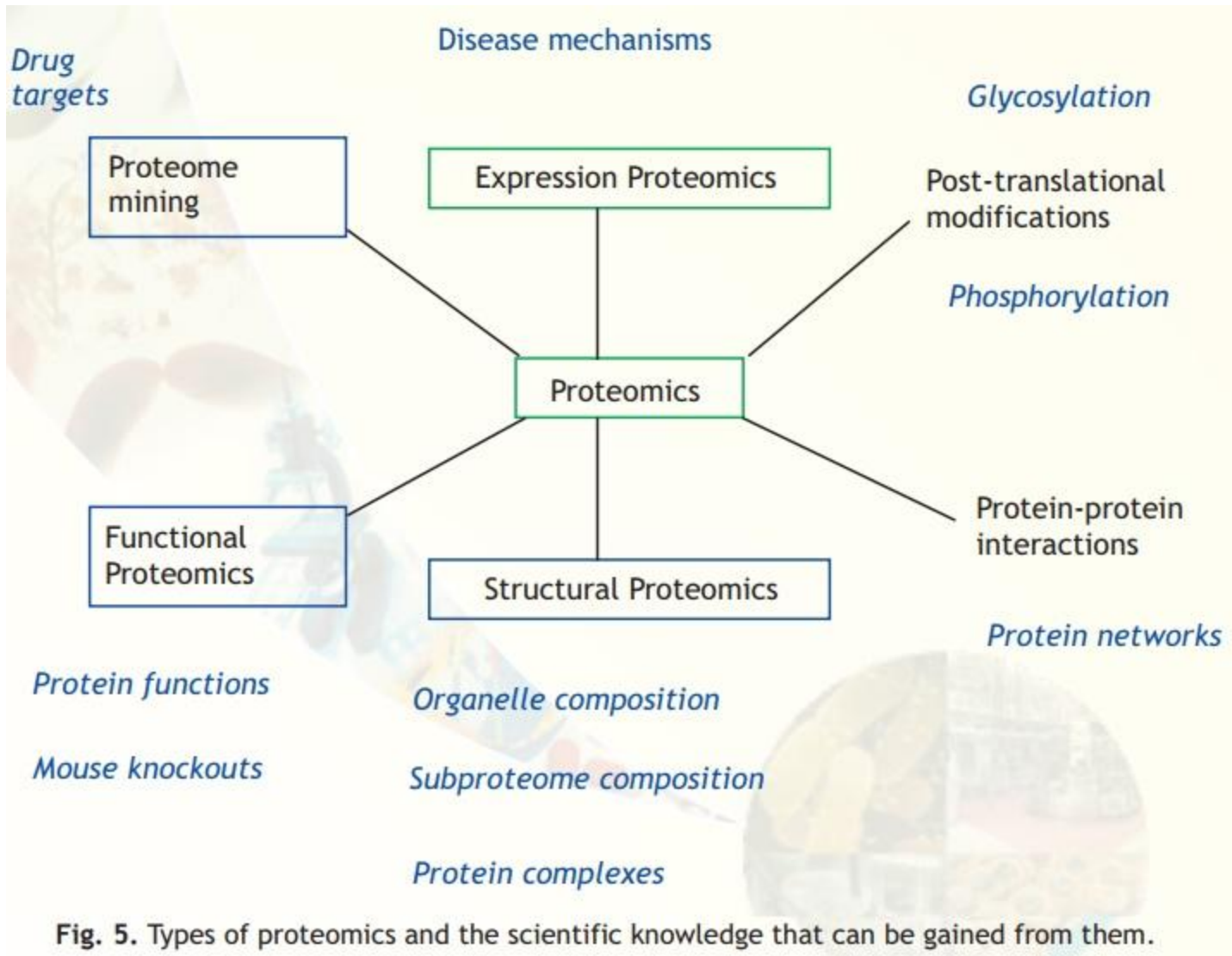
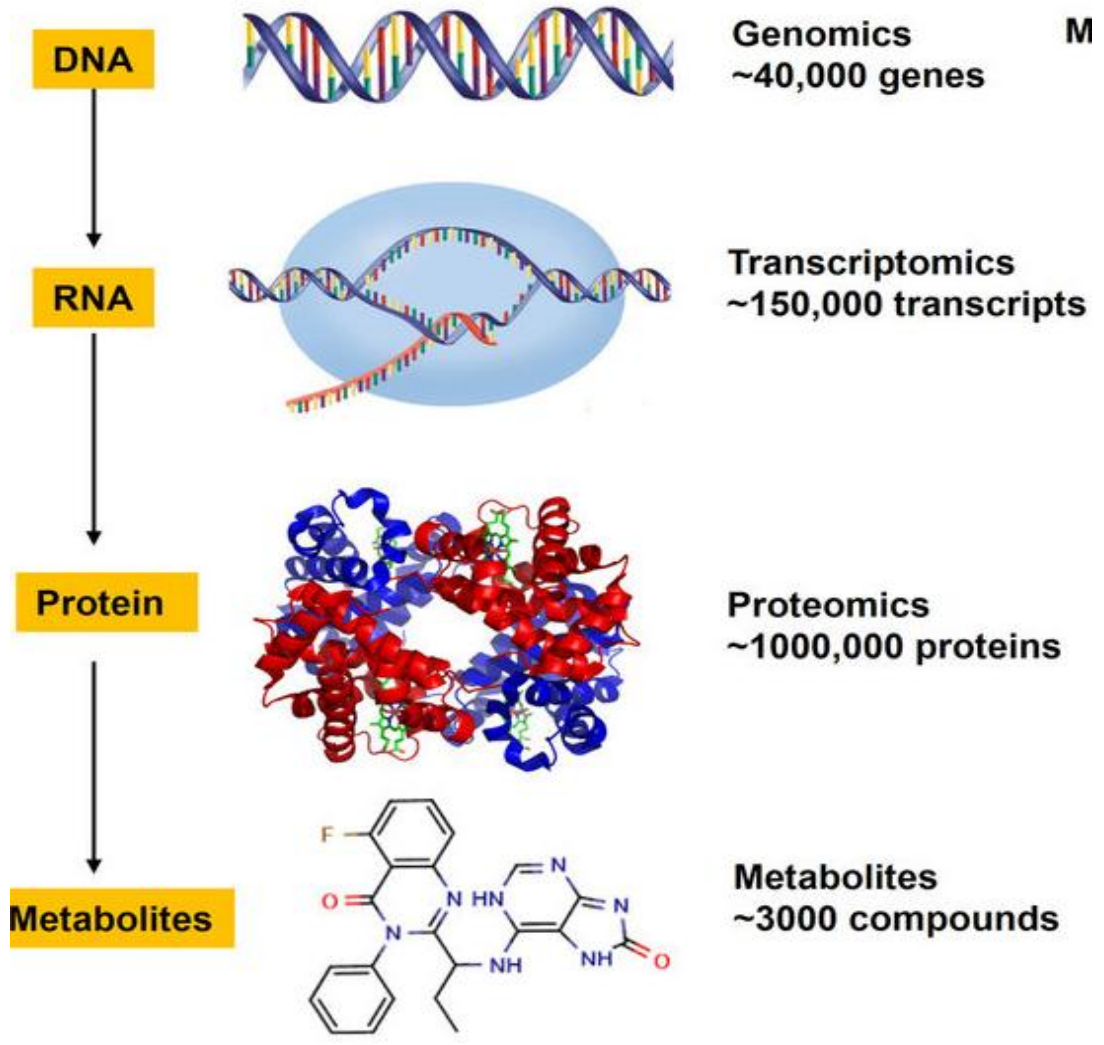


Fig. 5. Types of proteomics and the scientific knowledge that can be gained from them.

Metabolomics:

- Is the large-scale study of small molecules, commonly known as **metabolites**, within cells, biofluids, tissues or organisms.
- Collectively, these small molecules and their interactions within a biological system are known as the **metabolome**.



Bioinformatics vs Computational Biology

Bioinformatics differs from a related field known as computational biology.

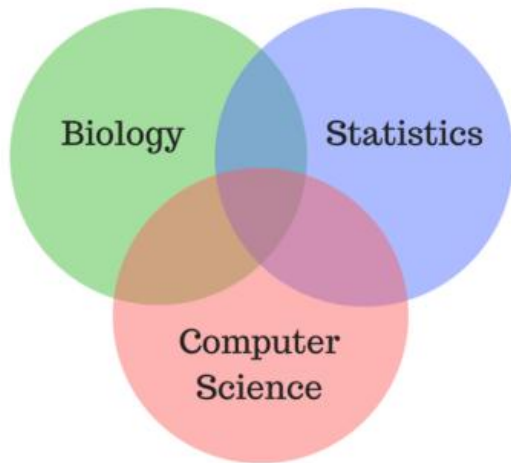
Bioinformatics

- It is limited to sequence, structural, and functional analysis of genes and genomes and their corresponding products and is often considered computational molecular biology.
- It is the development and **application** of computational tools in managing all kinds of biological data.

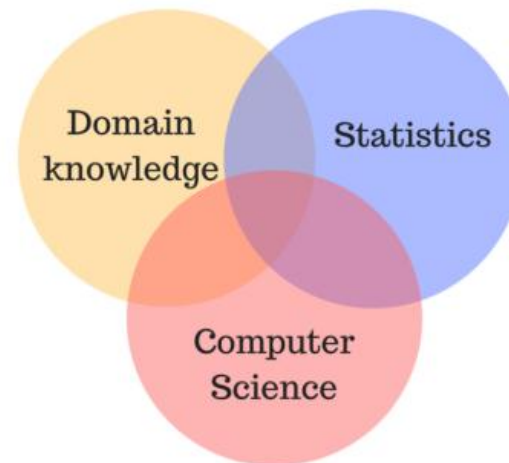
Computational biology

- It encompasses all biological areas that involve computation.
- It is more confined to the **theoretical** development of algorithms used for bioinformatics.

Bioinformatics



Data Science



Applications of Bioinformatics

- Bioinformatics has not only become essential for basic genomic and molecular biology research,
- But is having a major impact on many areas of biotechnology and biomedical sciences.

The main uses of bioinformatics include:

1. Bioinformatics plays a vital role in the areas of
 - structural genomics
 - functional genomics and
 - nutritional genomics.

2. It covers emerging scientific research and the exploration of proteomes
 - from the overall level of intracellular protein composition (protein profiles),
 - protein structure,
 - protein-protein interaction, and
 - unique activity patterns (e.g. post-translational modifications).

3. It is used for transcriptome analysis where **mRNA** expression levels can be determined.

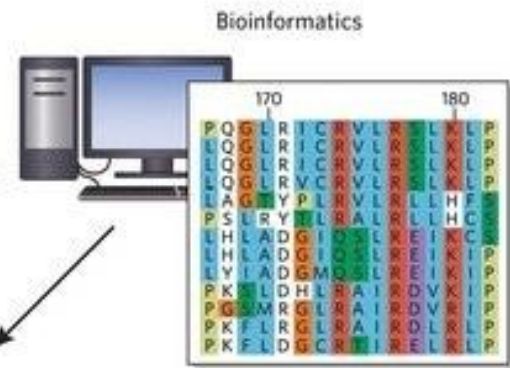
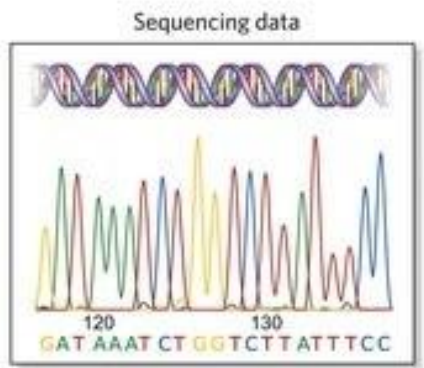
4. It is used to identify and structurally modify a natural product, to design a compound with the desired properties and to assess its **therapeutic effects**, theoretically.

5. **Cheminformatics** analysis includes analyses such as similarity searching, clustering, QSAR modeling, virtual screening, etc.

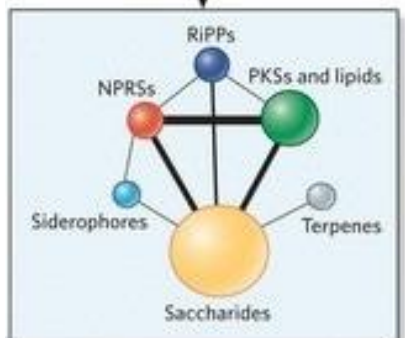
(QSAR – Quantitative structure – activity relationship models. Used to study relationship between chemical structure & biological activity of chemicals)

6. Bioinformatics is playing an increasingly important role in almost all aspects of **drug discovery** and drug development.

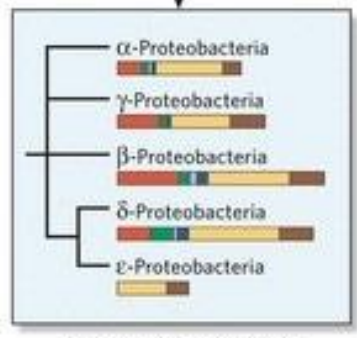
7. Bioinformatics tools are very effective in prediction, analysis and interpretation of **clinical** and preclinical findings.



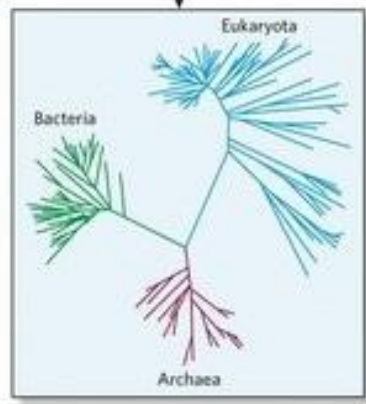
Prediction and analysis of biosynthetic gene clusters



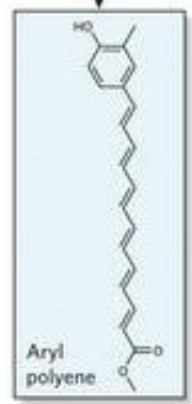
BGC diversity and relationships



Taxonomic distribution



Evolutionary events



Discovery of new molecules