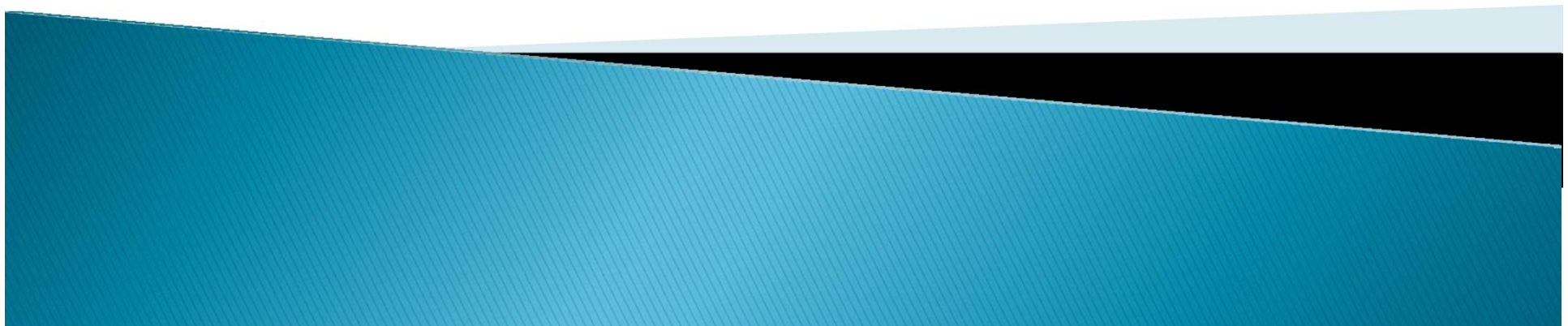


Basics of Bioinformatics



BIOINFORMATICS

UNIT – IV

- **Introduction,**
- **Definition**
- **Objectives of Bioinformatics,**
- **Bioinformatics Databases,**
- **Impact of Bioinformatics in Vaccine Discovery**
- **Applications of Bioinformatics**

Introduction

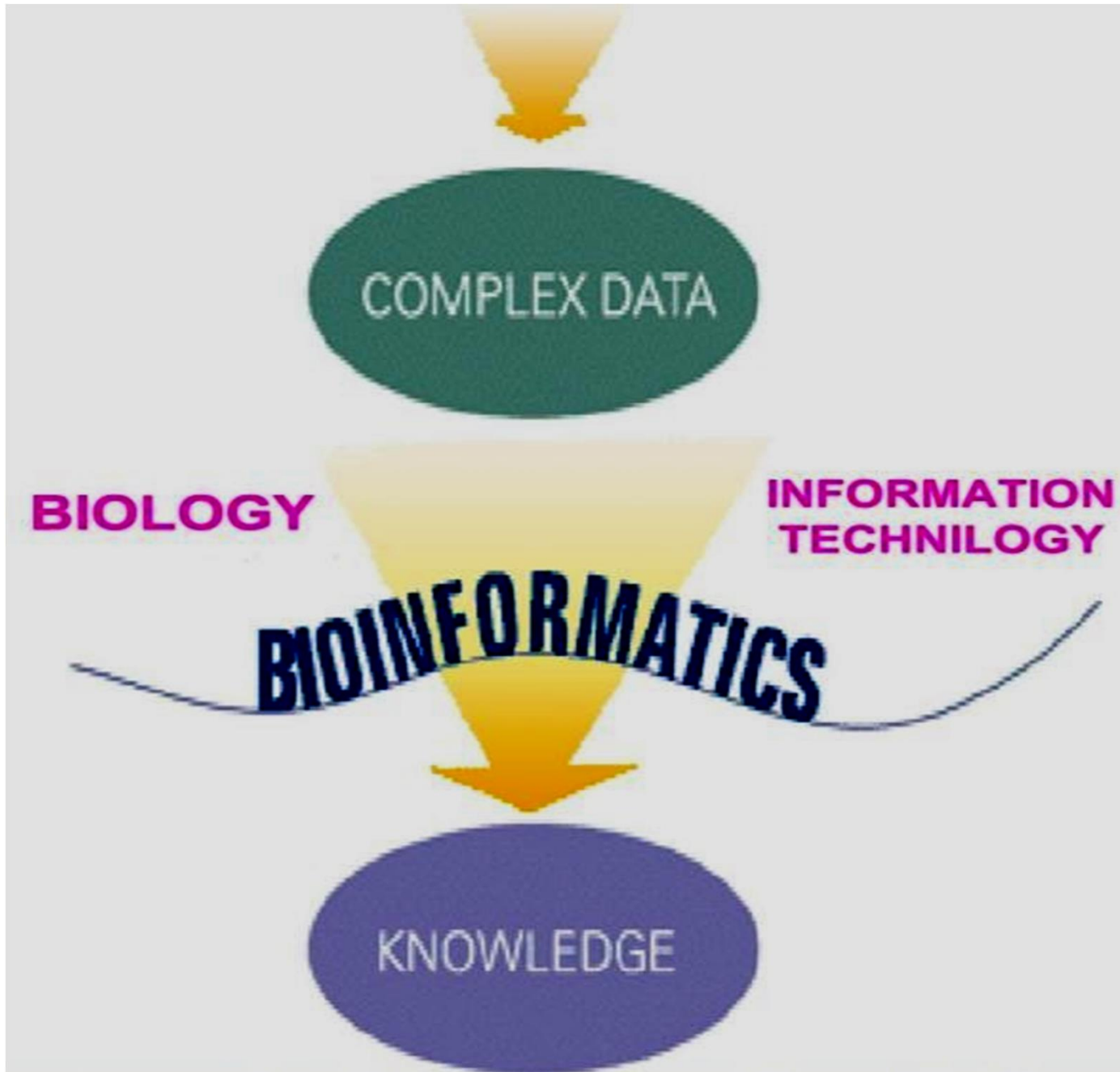
Chronology of Events-I

Sr.	Year	Event
1	1872	Gregor Mendel, —"Father of Genetics through cross fertilization experiment of colored flowers discovered that biological information was passed on from generation to generation.
2	1972	Paul Berg made the first recombinant DNA molecule using ligase
3	1972	Stanley Cohen, Annie Chang and Herbert Boyer produced the first recombinant DNA organism.
4	1973	Joseph Sambrook led a team that refined DNA electrophoresis using agarose gel.
5	1973	Herbert Boyer and Stanely Cohen invented DNA cloning.
6	1977	A method for sequencing DNA was discovered and the first genetic engineering company, Genetech was founded.
7	1981	579 human genes had been mapped and mapping by <i>in situ</i> hybridization had become a standard method. Marvin Carruthers and Leory Hood invented a method for automated DNA sequencing.
8	1988	The Human Genome Organization (HUGO) was founded. This is an international organization of scientists involved in Human Genome Project.
9	1989	The first complete genome map was published of the bacteria <i>Haemophilus influenza.</i>

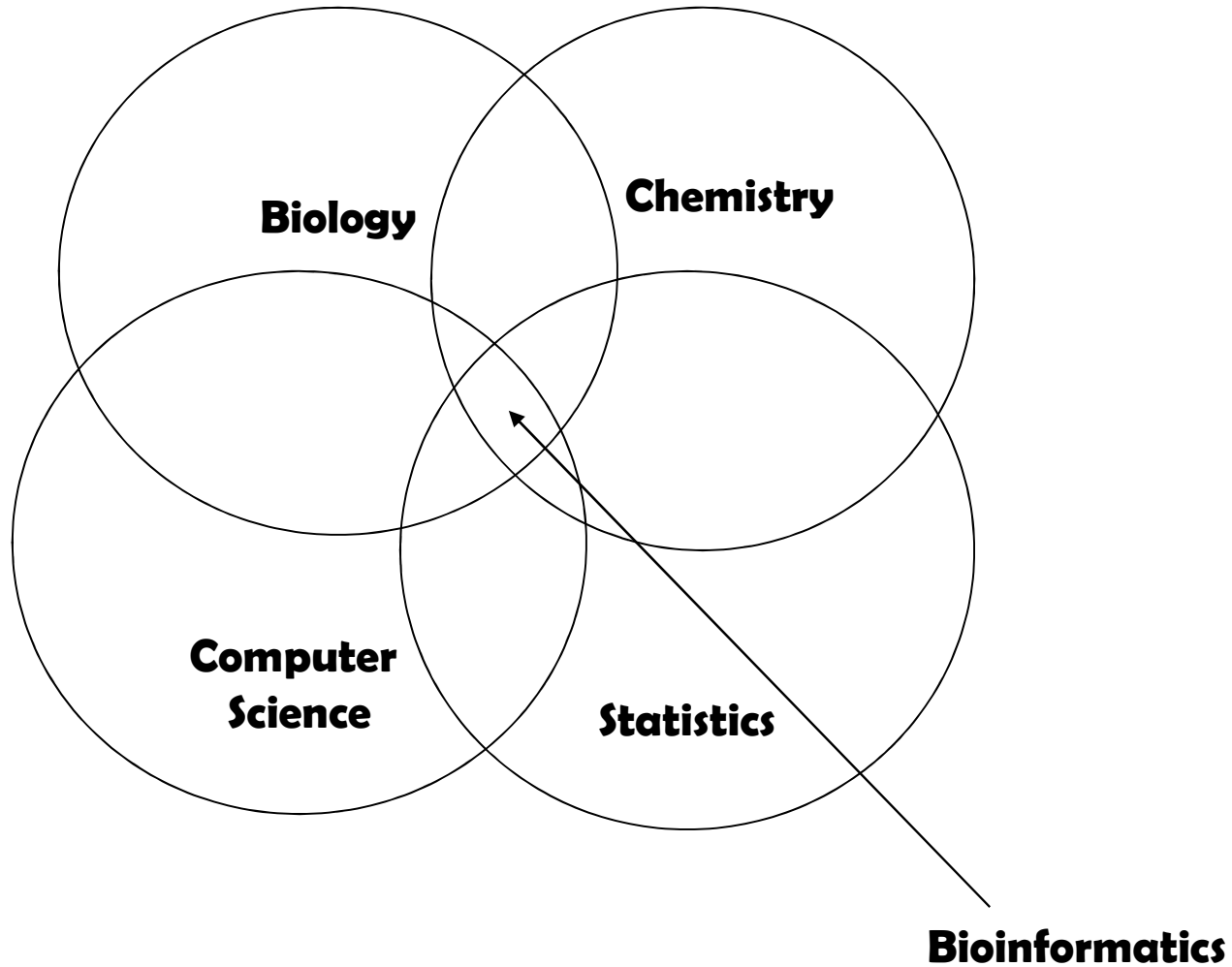
What is Bioinformatics?

- **The marriage between computer science and molecular biology.**
- The algorithm and techniques of computer science are being used to solve the problems faced by molecular biologists
- **'Information technology applied to the management and analysis of biological data'**
- It broadly involve the **computational tools and methods used to manage, analyse and manipulate** volumes and volumes of biological data
- **Storage and Analysis** are two of the important functions – bioinformaticians build tools for each.

Introduction



What is Bioinformatics?



Definitions

The term **Bioinformatics** was coined first time **by Paulien Hogeweg and Ben Hesper in 1970** as "the study of informatic processes in biotic systems".

Various Definitions

1. Bioinformatics is the field of science in **which biology, computer science, and information technology merge to form a single discipline for the collection, organization, analysis, manipulation, presentation, and sharing of biological data** to solve biological problems on the molecular level.
2. Bioinformatics is the **collection of mathematical, statistical and computing methods that aim to solve biological problems** using DNA and amino acid sequences and related information.

The field of science in which **biology, computer science** and **information technology** merge into a single discipline is Bioinformatics.

Biologists

collect molecular data:
DNA & Protein sequences,
gene expression, etc.



Bioinformaticians

Study biological questions by
analyzing molecular data with
the computer tools



Computer scientists

(+Mathematicians, Statisticians, etc.)
Develop tools, softwares, algorithms
to store and analyze the data.

Objective of bioinformatics

- To study **normal biological processes**.
- Store/retrieve biological information (databases)
- Retrieve/compare gene sequences
- Predict function of unknown genes/proteins
- Analyse data and interpret results at molecular level.
- **To analyze the cause of disease** due to malfunctioning of biological process.
- To **design various approaches to improve** biological processes.
- To improve drug discovery techniques **to help in developing new target drugs for fatal diseases**.
- Apply this knowledge in various fields
- Enables study and research on **development of preventive medicines for life threatening diseases like cancer, HIV, Corona virus, etc**

Components of bioinformatics

1. Creation of databases:

- Involves organising, storage and management of biological data sets.
- The databases are accessible to researcher to know the existing information and submit new entries.
- eg. Protein sequence databank for molecular structure.

2. Development of algorithm and statistics:

- Involves development of tools and resources to determine relationship among the members of large data sets.
- Eg. Comparison of protein sequence data with already existing protein sequences.

3. analysis of data and interpretation

- **Appropriate use of component 1 and 2 to analyze the data and interpret the results in a biologically meaningful manner.**
- **This includes DNA, RNA, protein sequence, protein structure etc.**

Broad coverage of Bioinformatics

Functional Genomics- Identification of genes and their respective functions

Structural Genomics – Predictions related to functions of proteins

Comparative genomics – understanding the genomics of different species of organism.

DNA microarray– designed to measure the level of gene expression in different tissues, various stages of development and in different diseases

Medical informatics- involves the management of biomedical data with special reference to bio molecule.

Biological databases

Collection of biological data on computer which can be manipulated **to appear in varying arrangements and subsets is regarded as database.**

Biological information can be stored in different databases as:

1. Nucleotide sequence databases
2. Protein sequence databases
3. Molecular structure of data bases
4. Other data bases

1. Nucleotide sequence database:

- Nucleotide sequence data submitted by the scientists and genome sequencing groups is at the data bases namely **genbank, EMBL (European Molecular Biology Laboratory) and DDVJ (DNA data bank of Japan).**
- There is a good coordination between **these three data bases as they are synchronised on daily basis.**
- Besides the primary nucleotide data bases there are some other data bases **also to provide information on genes, genomes and ongoing research projects.**

2. Protein sequence databases:

- Protein sequence databases are **usually prepared from the existing literature and in consultation with the experts.**
- In fact these data bases represent **the translated DNA data bases.**

3. Molecular structure of data bases:

- The three-dimensional structures of macromolecules are determined by **x-ray crystallography and nuclear Magnetic resonance.**
- PDB and SCOP (Structural classification of protein) are the primary data bases 3D structures of biological molecules.

4. Other databases:

KEGG (Kyoto Encyclopedia of Genes and Genomes database) is an important one that provides information on the current knowledge of molecular biology and cell biology with special reference to information on metabolic pathways interacting molecules and genes.

Thank

you



Basic Concepts of Bioinformatics

- Its main priority has been implicated in data storage and genome sequence analysis.
- The unprecedented growth of information technology and extraordinary growth in molecular biology and recombinant DNA technologies and their interrelated studies culminated into cutting edge technology like bioinformatics.
- Thus, bioinformatics is also termed as bio-computing or computational biology.
- Recently, genomics occupies central role in bioinformatics i.e., understanding the basic life process.

It refers to its approach towards balancing biology, mathematics and computer programming in order to ensure availability of tools to compute and analyze biological data in a best possible manner.

It provides deeper understanding of molecular biology software programmes, programme coding and decoding, genetic data analysis and study of DNA structure and composition.

Bioinformatics involve integration of computers, software tools and databases that are used to address various biological queries. The two major activities involved in bioinformatics are study of genomics and proteomics.

Genomics- analysis of genomes

Proteomics- analysis of proteins.

Impact of Bioinformatics in Vaccine Discovery

- Bioinformatics help in discovery of vaccines in a more effective way and in shorter span of time. This is because it combines Biology with Pharmacology. Bioinformatics reduces the time and cost required to develop hypertensive drugs with fewer Side Effects. Science of genomics play a vital role in improving human health globally.
- It has been found that if a genome sequence of pathogen is available a vaccine can be easily created to destroy that sequence and hence occurrence of disease can be prevented. Genomic data are processed by a variety of software programs that help identify individual chains and the outcomes.
- Designing an ideal vaccine largely depends on targeted pathogens and their interactions with existing drugs.
- Study of genome sequences of various pathogens along with rapid advancements in biotechnology allows us to collect large amount of useful information about hosts and pathogens. That play an important role in discovery of vaccines.
- Thus we can say that Bioinformatics has the potential to improve process of vaccine development.

Helpful vaccine Informatics support.

- Analysis to help find good molecular targets for vaccines.
- Private data basis to help collect and organise lab and clinical data.
- Public data bases and web works to help display and analyse the results of clinical trials.
- Integrating in additional data from samples in clinical trials into database.