BHARATHIDASAN UNIVERSITY

Tiruchirappalli- 620024,
Tamil Nadu, India.

Programme : M.Sc., Biomedical Science

Course Title : Bioinformatics
Course Code : BM35S1BI

Unit-V

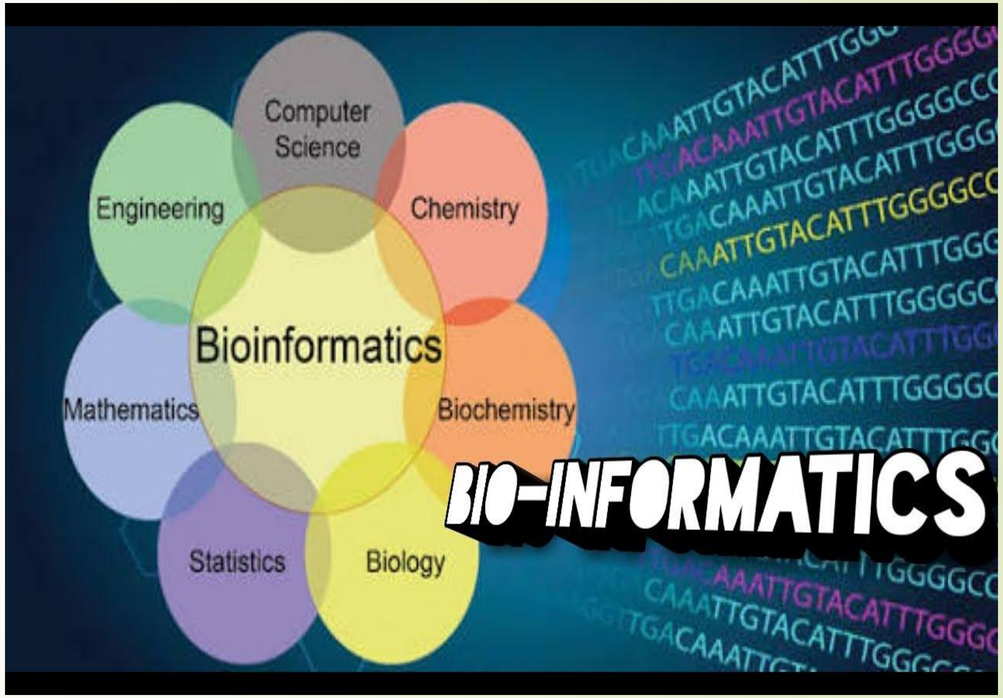
TOPIC: ADVANCEMENT OF BIOINFORMATICS – OVERVIEW,SYSTEM BIOLOGY, E.COIL

Dr. P. JEGANATHAN
Guest Lecturer
Department of Biomedical Science

ADVANCEMENT OF BIOINFORMATICS – OVERVIEW,SYSTEM BIOLOGY, E.COIL

BIOINFORMATICS:

- Bioinformatics is an interdisciplinary area of science composed of biology, mathematics , and computer science. Bioinformatics is the application of information technology to manage biological data that helps in decoding plant genomes.
- ADVANCES OF BIOINFORMATICS:
- Bioinformatics has developed out of need to understand the code of life that is DNA ,the massive DNA sequencing project have evolved and added in growth of science of bioinformatics.
- The ultimate goal of bioinformatics is to uncover the wealth of biological information hidden in the mass of sequence, structure, literature and other biological data.



TECHNOLOGY AND COMPUTING POWER:

Technology is the collection of techniques ,skills, methods, and process applied to make life easier. Technology can be the knowledge of techniques and processes.

- computing power is the ability to undertake or be used for computation.

CREATION OF DATABASES:

- This involves the organizing, storage and management the biological data sets. The databases are accessible to researchers to know the existing information and submit new entries. e.g protein sequence databank for molecular structure.

DEVELOPMENT OF ALGORITHMS AND STASTICS:

- This involves the development of tools and resources to determine the relationship among the members of large data sets e.g; Comparison of protein sequence data with the already existing protein sequence.

ANALYSIS OF DATA AND INTERPRETATION:

The appropriate use of components to analyze data and interpret the results in biologically meaningful manner. This include DNA, RNA and protein sequence, protein structure, gene expression profiles and biochemical pathways.

GENOMICS:

- Genomics is the interdisciplinary field of science within the field of molecular biology. Genomics aims at collective characterization and Quantification of genes, which direct the protection of protein with the assistance of enzymes and messenger molecules.
- Genomics also involve in the sequencing and analysis of genomes.advances in genomics have triggered a revolution in discovery based research to understand even the most currently complexes biological systems such as brain. In contrast to genetics, which refers to study of individual genes and their role in inheritance, genomics uses high thorough put DNA sequencing and bioinformatics to assemble ,and analyze the function and structure of entire genomes.

Genomics

Whole genome analysis, SNP analysis, exome analysis, CGH analysis, epigenome analysis

Transcriptomics

Gene expression analysis,
RNA-Seq analysis, transcription factor binding site analysis

Proteomics

Protein expression analysis, phosphorylation analysis, post-translational modification analysis, polyspecimen analysis

Metabolomics

In vivo lipid profile analysis & identification, metabolic pathway analysis, rapid carbohydrate trace analysis

Data Management

AI, machine learning, quantum computing, big database construction, semantic web full-text search, data mining, graph mining, HPCI, cloud services

Chemoinformatics

Lead compounds, and virtual chemical compound screening, side-effects and toxicity prediction

Structural Informatics

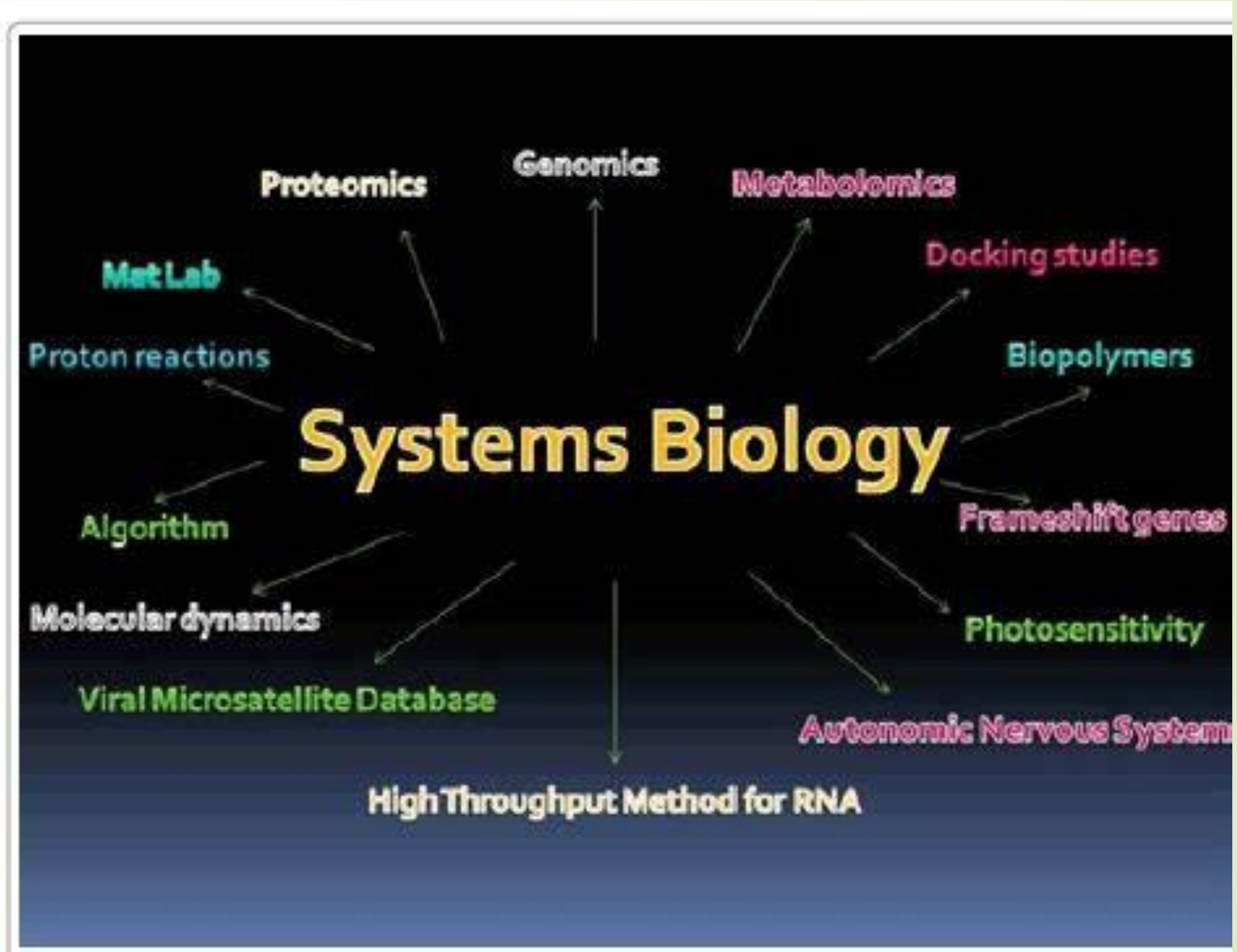
Amino acid sequence/protein structure analysis,
molecular dynamics (MD) simulation, protein tertiary structure prediction, docking simulation

Systems Biology

Network analysis, pathway analysis, metabolic flux analysis, biosimulation, bioprocess design

SYSTEM BIOLOGY:

- A Field that seeks to study of relationship and interaction between various part of biological system (Metabolic pathway, organelles, cells and organism) and to integrate this information to understand how biological function.
- System biology is relatively new biological study field that focuses on systematic study of complex interaction in biological system.



OTHER DEFINITIONS:

- System biology is a comprehensive quantitative analysis of manner in which all components of biological system interact functionally over time(Alan Adrem,Cell,vol121,611-613,2005.Institute of systems biology, Seattle).
- System biology can be described as “Integrative Biology” with the ultimate goal of being able to predict de novo biological outcome given the list of components involved(Edison T.Liu, cell,vol121,505-506,2005.Genome Institute of Singapore).
- “System biology” aims at a quantitative understanding of biological systems to an extent that one is able to predict systematic features (peer Bork and luis serrano ,cell,vol121,507-509,2005.EMBL Germany).

DATABASE RESOURCES FOR SYSTEM BIOLOGY:

MICROARRAY GENE EXPRESSION DATABASE:

- Array Express
- GEO-Gene express omnibus
- CIBEX-Centre for information biology gene expression database.

PROTEIN –PROTEIN INTERACTION DATABASE:

- BIND plus-Biomolecular object network databases
- DIP- Databases of interacting proteins.
- MINT-Molecular interaction of databases.
- STRING-Search tool for the retrieval of interacting genes/proteins.

PATHWAY DATABASE:

- KEGG (Kyoto encylopedia of genes and genomes)
- REACTOME
- METACYC
- Ecocyc

KEGG:

- It is a bioinformatics resource for linking genomes to life and environment.

ECOCYC:

Encyclopedia of Escherichia coli K-12 Genes and Metabolism .

- It is a scientific database for an bacterium Escherichia coli K-12 MG1655.

BIO MODEL DATABASE:

- It is a data resource that allows biologists to store , search, and retrieve published mathematical model of biological interest.

BIGG DATABASE:

- It is a metabolic reconstruction of human metabolism designed for system biology stimulation and metabolic flux balance modelling.
- It is a comprehensive literature-based genome-scale metabolic reconstruction that accounts for function of 1,496 ORFs,2,004 proteins, 2,766 metabolites, and 3,311 metabolic and transport reaction.
- It was assembled from build 35 of human genome. Need user name and password.

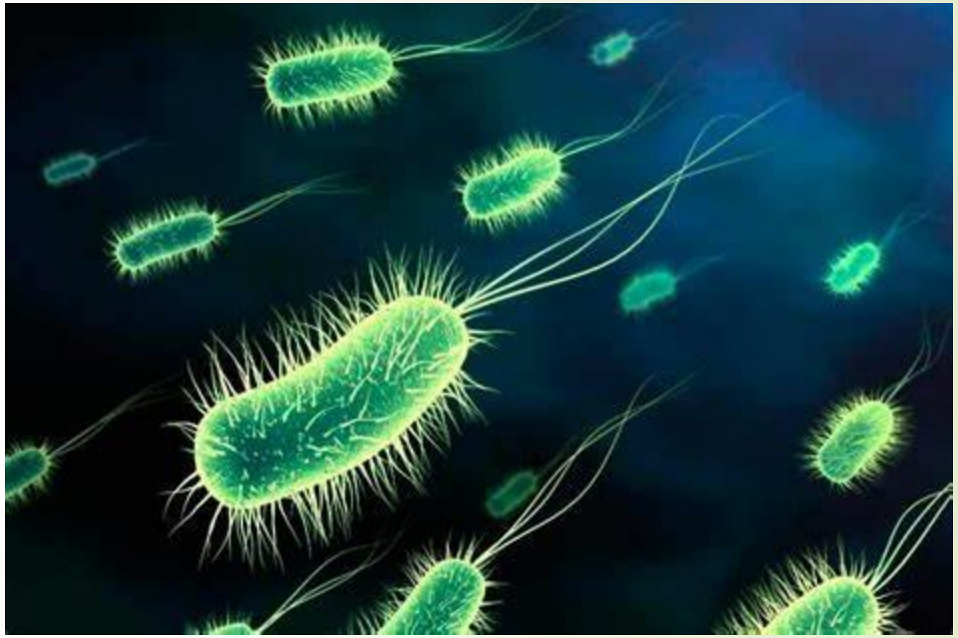
APPLICATION OF SYSTEM BIOLOGY:

- System biology in crop improvement.
- System biology in drug discovery.
- Network biology
- Topological analysis of biological network.

E.COIL:

ECO CYC DATABASE:

- ECO CYC word comes from E.coli(ECO) and Encyclopedia (cyc)
- The encyclopedia of Escherichia coli K-12 genes and metabolism (ECOCYC) is a database that combines the information about the genome and the intermediary metabolism of E.coli.
- It describes the known genes of E.coli , the enzymes encoded by these genes , the reaction catalyzed by each enzyme and the organizations of these reaction into metabolic pathway.
- It is freely accessible ,comprehensive databases that collect and summarizes experimental data for Escherichia coli K-12 at Eco cyc.org.



- The eco cyc system consist of knowledge base (KB) that describes gene and intermediary metabolism of E.coli, and a graphical user interface for accessing that knowledge.
- Ecocyc is a joint work between our group at SRI International , and at a group at Marine Biological laboratory (MBL) led by Monica Riley.
- BY integrate the genomic data with comprehensive Knowledge of the metabolic functions of gene products, we greatly increase the value of the genomic data and the types of analyses that can be applied to it.
- The development of Ecocyc is funded by NIH grants GM77678 and GM 71962 from the NIH National institute of general medical sciences.

GENOME:

- ECOCYC contains the complete genome sequence of E.COLI and describes the nucleotide position and function of all known protein-coding and RNA coding E.COLI genes .genome related classes that are populated with in ECOCYC include genes, pseudo-genes, promoters,DNA Binding site and rep elements.

PROTEOME:

- ECOCYC describes all known monomers and multimeric protein complexes of e.coli. Ecocyc contains extensive annotation of the features of e.coli proteins such as phosphorylation sites , metal ion binding sites , and enzyme active sites assigned by ecocyc curators and imported from uniport. Relevant classes within Ecocyc include polypeptide and protein complexes.

RNA ome:

- Ecocyc describes all known RNAs and protein-RNA complexes of E.coli. Relevant classes with in Ecocyc include RNAs, rRNAs, and regulatory –RNAs. Ecocyc does not explicitly represent messenger RNAs.

REGULATION:

- Ecocyc contains the most complete description of the regulatory network of any organism. It covers E.COLI operons, promoters, transcription factor binding sites,attenuators, and small –RNA regulators ,as well as substrate level regulation of E.COLI enzymes. Each molecular regulatory interaction is described as an instance of class regulation , whose subclasses describe different types of regulation.

METABOLISM:

Ecocyc describes all known metabolic and signal –transduction pathways of E.coli. It describes each metabolic enzyme of E.coli including its cofactors ,activators ,inhibitors, and subunit structure.

MEMBRANE TRANSPORTERS:

- Ecocyc annotates E.COLI Transport protein and the associated transport reaction that they mediate.

GROWTH OBSERVATIONS:

- Ecocyc integrates data on the growth of E.coli under many different growth conditions.

DATABASE LINKS:

- Ecocyc is linked to other biological database containing protein and nucleic acid sequences data, bibliographic data, protein structures , and description of different E.coli stains.

REFERENCES

- HTTPS://WWW.GENOME.JP/TOOLS/MOTIF/
- HTTPS://WWW.GOOGLE.COM/URL?SA=T&SOURCE=WEB&RCT=J&URL=HTTPS://PROSITE.EXPASY.ORG/&VED=2AHUKEWIK7OG4H-Z9AHWXS2WGHWPWCAWQFNOECAUQAQ&USG=AOVVAW2YSYV7B25SZFM7D4PADBFC


thank you