



PERIYAR UNIVERSITY
(NAAC 'A' Grade-State University-NIRF Rank 68)
SALEM-636011
DEPARTMENT OF BIOTECHNOLOGY

Value added course :Protein Sequence analysis
Course code :CBIT-02
Course Mentors : Dr.P.Indra Arulsevi

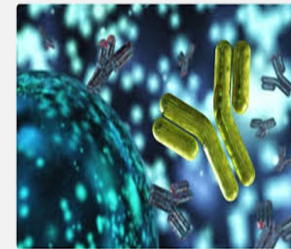
December
2019

ABOUT THE COURSE :

The course will introduce amino acid and protein physical properties as well as the alignment and evolution of protein sequences. Protein structure and methods of structure determination will be presented as well as the use of protein databases and software for visualizing proteins. Methods for secondary and tertiary protein structure prediction will be discussed as well as methods for modeling small/molecule-protein interactions and protein-protein interactions. The course will also cover scoring matrix for comparing protein sequences to judge the quality of the alignment. The pairwise and multiple sequence alignment is used to study identity regions of similarity indicate functional, structural and evolutionary relationships between biological sequences. The purpose of this course is to enlighten the knowledge of students and understanding on technologies for computational aspects of mapping and visualization of protein interaction networks.

COURSE OUTCOME:

- **Database 1:** Ability to understand the basics of sequence alignment and analysis.
- **Database 2:** To gain knowledge about various Biological databases that provide information about protein- gene and protein-protein interaction.
- **Database 3:** Explain about the concept of pairwise and multiple sequence alignment, algorithms and tools for both.
- **Database 4:** To gain knowledge about proteomic tools, classification and Visualization of 3D structure of proteins.



SYLLABUS

Protein sequence database:

Uniprot –KB, SWISS-PROT, TrEMBL, PIR-PSD, Gene expression database : GEO, SAGE, Derived database: Sequence : Interpro, Prosite, Pfam, ProDAM, Gene Ontology, Structure classification database: CATHSCOP, FSSP Protein-Protein interaction database: STRING.

Sequence analysis:

File formats : Various file formats for bio-molecular sequences : GeneBank: FASTA, GSC, MSF, etc.,
Basic concepts : Sequence similarity, identity and homology, definitions of homologues, orthologues, paralogues.
Scoring matrices: Basic concept of a scoring matrix, proteins sequences, principles based on which these matrices are derived.

Sequence Alignments:

Pairwise & Multiple sequence alignments

Protein Properties:

Computation of various parameters using proteomics tools at the ExPASy server and EMBOSS.

Structural biology:

3-D structure visualization and simulation: Visualization of structures using Rasmol or SPDBV or CHIME or VMD.

Classification and comparison of protein 3D structures:

Secondary structure prediction. Tertiary structure prediction.