

Course Code	Course Name	Credits
26BT009	GENOMICS AND PROTEOMICS	04

Course Objectives

- To provide a comprehensive understanding of genomics and proteomics
- To develop proficiency in managing and analyzing genome
- To impart knowledge of protein structure, characterization, and proteomics techniques
- To introduce Artificial Intelligence (AI) and Machine Learning (ML) applications
- Gain knowledge of DNA sequencing and proteome analysis techniques.

Learning Outcomes

Upon successful completion of this course it is intended that a student will be able to:

- Demonstrate knowledge of genomics and DNA sequencing techniques
- Retrieve, manage, and interpret genome data
- Perform protein structure and proteome analyses
- Apply AI and ML approaches to genomic and proteomics problems
- Gain knowledge on physical and chemical properties of protein.

Unit 1 - Introduction to Genomics (12 Hrs.)

DNA sequencing methods – manual & automated: Maxam & Gilbert and Sangers method. Pyrosequencing, Genome Sequencing: Shotgun & Hierarchical (clone contig) methods, Computer tools for sequencing projects: Genome sequence assembly software.

Unit 2 – Managing and Distributing Genome Data (12 Hrs.)

Web based servers and softwares for genome analysis: ENSEMBL, VISTA, UCSC Genome Browser, NCBI genome. Selected Model Organisms' Genomes and Databases.

Unit 3 - Protein structure and characterization (14 Hrs.)

Introduction to protein structure, Chemical properties of proteins. Physical interactions that determine the property of proteins. Short-range interactions, electrostatic forces, van der Waal interactions, hydrogen bonds, Hydrophobic interactions. Determination of sizes (Sedimentation analysis, gel filtration, SDS-PAGE); Native PAGE, Determination of covalent structures – Edman degradation.

Unit 4 – Proteome analysis and techniques (10 Hrs.)

Analysis of proteomes. 2D-PAGE. Sample preparation, solubilization, reduction, resolution. Mass spectrometry based methods for protein identification, Metabolomics & global biochemical networks.

Unit 5 - Application of AI in Genomics & Proteomics Analysis (12 Hrs.)

Introduction to AI, ML and deep learning; Algorithms (regression, classification and clustering); AI based methods for sequence alignment, gene prediction and variant detection; Protein structure prediction (AI based approaches); AI in mass spectroscopy data interpretation.

Reference Books:

1. Jonathan Pevsner, Bioinformatics and Functional Genomics, 3rd (2015, Wiley-Blackwell)
2. T.A. Brown, Genomes, 5th edition, Garland Science (Taylor & Francis Group)(2023)
3. S.R. Pennington and M.J. Dunn, Proteomics: From Protein Sequence to Function, BIOS Scientific Publishers (Taylor & Francis), 2001
4. Sushmita Mitra and Tinku Acharya, Machine Learning for Bioinformatics, CRC Press, 2008
5. S.C. Rastogi, Namita Mendiratta, Parag Rastogi, Bioinformatics: Methods and Applications, PHI Learning Pvt. Ltd. (India), 2022
6. S. Rajasekaran and G. A. Vijayalakshmi Pai, Fundamentals of Bioinformatics, Ane Books India, 2005
7. Hooman Zolnoori, Artificial Intelligence in Bioinformatics, Springer Nature, 2022.