**ST ALOYSIUS (**DEEMED To Be **UNIVERSITY)**

# SYLLABUS

FOR THE COURSE

# M.Sc. (BIOINFORMATICS)

Choice Based Credit System (CBCS)

**(With effect from academic year 2024 onwards)**

**ALOYSIUS INSTITUTE OF MGMT & INFORMATION TECHNOLOGY**
**BEERI, MANGALORE – 575 003**
**Phone - 2286890 /6490299**
**www.staloysius.ac.in**

## Master of Science - M.Sc. (Bioinformatics)

## Scheme of Teaching || New CBCS Syllabus from 2024

### I Semester

| Sl No | Subject Code | Subject Title | Type | Credits | Total No of Hours | Lecture Hrs / week | Tutorial Hrs / week | Practical Hrs / week | Self Study Hrs / week |
|---|---|---|---|---|---|---|---|---|---|
| **Hard Core Theory & Lab** | | | | | | | | | |
| 1 | PH 521.1 | Bioinformatics Foundations: From Molecules to Systems | HC | 4 | 45 | 3 | 2 | -- | 2 |
| 2 | PH 522.1 | Algorithms in Computational Biology | HC | 4 | 45 | 3 | 2 | -- | 2 |
| 3 | PH 523.1 | Basic Biochemistry and Immunology | HC | 4 | 45 | 3 | 2 | -- | 2 |
| 4 | PH 524.1 | Cell and Molecular Biology | HC | 4 | 45 | 3 | 2 | -- | 2 |
| 5 | PH 525.1P | Bioinformatics & Computational Biology Lab | HCL | 3 | 90 | -- | (1 x 2)<br>2 | (3 x 2)<br>6 | -- |
| **Soft Core Theory & Lab** | | | | | | | | | |
| 6 | PS 526.1 | Basics of Programming with Python | SC | 4 | 45 | 3 | 2 | -- | 2 |
| 7 | PS 527.1P | Programming with Python Lab | SCL | 3 | 90 | -- | (1 x 2)<br>2 | (3 x 2)<br>6 | -- |
| 8 | PS 528.1 | Research & Soft Skill Development - I | SC | -- | 30 | 2 | 2 | -- | 4 |
| | | **Total** | | **26** | | **17** | **16** | **12** | **14** |

* HC – Hard Core SC – Soft Core HCL – Hard Core Lab SCL – Soft Core Lab OE – Open Elective

## Master of Science - M.Sc. (Bioinformatics)

## Scheme of Teaching || New CBCS Syllabus from 2024

### II Semester

| Sl No | Subject Code | Subject Title | Type | Credits | Total No of Hours | Lecture Hrs / week | Tutorial Hrs / week | Practical Hrs / week | Self Study Hrs / week |
|---|---|---|---|---|---|---|---|---|---|
| **Hard Core Theory & Lab** | | | | | | | | | |
| 1 | PH 521.2 | Structural Proteomics and Biomarkers | HC | 4 | 45 | 3 | 2 | -- | 2 |
| 2 | PH 522.2 | Computational and Translational Immunology | HC | 4 | 45 | 3 | 2 | -- | 2 |
| 3 | PH 523.2P | Structural Proteomics and Computational Immunology Lab | HCL | 3 | 90 | -- | (1 x 2)<br>2 | (3 x 2)<br>6 | -- |
| **Soft Core Theory & Lab** | | | | | | | | | |
| 4 | PS 524.2 | Statistics for Genomics Applications | SC | 3 | 45 | 3 | 2 | -- | 2 |
| 5 | PS 525.2 | Intermediate Python with Bio-Computing | SC | 3 | 45 | 3 | 2 | -- | 2 |
| 6 | PS 526.2P | Programming: Intermediate python and abaiostatistics | SCL | 3 | 90 | -- | (1 x 2)<br>2 | (3 x 2)<br>6 | -- |
| **Open Elective – Offered to other Depts** | | | | | | | | | |
| 7 | PO 527.2 | Bioethics, Biosafety and IPR | OE | 3 * | 45 | 3 | -- | -- | 2 |
| **Add-On** | | | | | | | | | |
| 8 | PS528.2 | Research& Soft Skill Development– II | SC | -- | 30 | -- | 4 | -- | 4 |
| 9 | PS 529.2 | Research Methodology & Ethics | SC | | | | | | |
| | | **Total** | | **20+ 3*** | | **15** | **16** | **12** | **14** |

* HC – Hard Core SC – Soft Core HCL – Hard Core Lab SCL – Soft Core Lab OE – Open Elective

# Master of Science - M.Sc. (Bioinformatics)

## Scheme of Teaching || New CBCS Syllabus from 2024

### III SEMESTER

| Sl No | Subject Code | Subject Title | Type | Credits | Theory Exam duration (hours) | Practical Exam duration (hours) | Max Marks for Internal Assessment | Max. Marks for Term End Exam | Total MArks |
|---|---|---|---|---|---|---|---|---|---|
| **Hard Core Theory & Lab** | | | | | | | | | |
| 1 | PH 521.3 | Computational Approaches in Therapeutics and Biosimilar Development | HC | 4 | 3 | -- | 30 | 70 | 100 |
| 2 | PH 522.3 | Next-Gen Sequencing and Single-Cell Technologies for Immunomics | HC | 4 | 3 | -- | 30 | 70 | 100 |
| 3 | PH523.3 P | Laboratory based on Therapeutics and Single Cell Technologies | HCL | 3 | -- | 3 | 25 | 50 | 75 |
| **Soft Core Theory & Lab** | | | | | | | | | |
| 4 | PS 524.3 | Data Analytics for Biology | SC | 3 | 3 | -- | 30 | 70 | 100 |
| 5 | PS525.3 | Python Applications in Bioinformatics | SC | 3 | 3 | -- | 30 | 70 | 100 |
| 6 | PS 526.3P | Lab based on Python and Machine Learning | SCL | 3 | -- | 3 | 25 | 50 | 75 |
| **Open Elective – Offered to other Depts** | | | | | | | | | |
| 7 | PO 527.3 | Bioethics, Biosafety and IPR | OE | 3 * | 45 | 3 | -- | -- | 2 |
| **Add-On** | | | | | | | | | |
| 8 | PS528.3 | Research& Soft Skill Development– II | SC | -- | 30 | -- | 4 | -- | 4 |
| 9 | PS 529.3 | Research Methodology & Ethics | SC | | | | | | |
| | | **Total** | | **20+ 3*** | | **15** | **16** | **12** | **14** |

# Master of Science - M.Sc. (Bioinformatics)

## Scheme of Teaching || New CBCS Syllabus from 2024

## IV SEMESTER

| **Sl No** | **Subject Code** | **Subject Title** | **Credits** | **Max. Marks for Dissertation / Valuation** | | **External Evaluation** | | **Total** |
|---|---|---|---|---|---|---|---|---|
| | | | | **Dissertation I** | **Dissertation II** | **Dissertation III** | **Max. Marks for Viva-voce (Final)** | |
| 1 | PH 521.4 | Industry Internship/ Project Work/ Dissertation | 16 | SRS & SDD (100) | Mid Term Evaluation (100) | Record Reading (100) | 200 | 500 |
| **Sl No** | **Subject Code** | **Subject Title** | **Credits** | **Internal Evaluation** | | **External Evaluation** | | **Total** |
| 2 | PS 522.4 | Domain Knowledge Project | 04 | Assignments / Paper Presentation (50) | | Documentation / Viva Voce (100) | | 150 |
| | **Total** | | **20** | -- | --- | --- | --- | **650** |

# Master of Science - M.Sc. (Bioinformatics)

## Scheme of Examination || New CBCS Syllabus from 2024

### I Semester

| Sl No | Subject Code | Subject Title | Type | Credits | Theory Exam duration (hours) | Practical Exam duration (hours) | Max Marks for Internal Assessment | Max. Marks for Term End Exam | Total MArks |
|---|---|---|---|---|---|---|---|---|---|
| **Hard Core Theory & Lab** | | | | | | | | | |
| 1 | PH 521.1 | Bioinformatics Foundations: From Molecules to Systems | HC | 4 | 3 | -- | 40 | 60 | 100 |
| 2 | PH 522.1 | Algorithms in Computational Biology | HC | 4 | 3 | -- | 40 | 60 | 100 |
| 3 | PH 523.1 | Basic Biochemistry and Immunology | HC | 4 | 3 | -- | 40 | 60 | 100 |
| 4 | PH 524.1 | Cell and Molecular Biology | HC | 4 | 3 | -- | 40 | 60 | 100 |
| 5 | PH 525.1P | Bioinformatics & Computational Biology Lab | HCL | 3 | -- | 3 | 25 | 50 | 75 |
| **Soft Core Theory & Lab** | | | | | | | | | |
| 6 | PS 526.1 | Basics of Programming with Python | SC | 4 | 3 | -- | 40 | 60 | 100 |
| 7 | PS 527.1P | Programming with Python Lab | SCL | 3 | -- | 3 | 40 | 60 | 75 |
| 8 | --- | Research Project - I | AON | -- | -- | -- | -- | - | -- |
| | | **Total** | | **26** | | | **200** | **450** | **650** |

* HC – Hard Core SC – Soft Core HCL – Hard Core Lab SCL – Soft Core Lab OE – Open Elective AON – Add On

## Master of Science - M.Sc. (Bioinformatics)

## Scheme of Examination || New CBCS Syllabus from 2024

### II Semester

| Sl No | Subject Code | Subject Title | Type | Credits | Theory Exam duration (hours) | Practical Exam duration (hours) | Max Marks for Internal Assessment | Max. Marks for Term End Exam | Total MArks |
|---|---|---|---|---|---|---|---|---|---|
| **Hard Core Theory & Lab** | | | | | | | | | |
| 1 | PH 521.2 | Structural Proteomics and Biomarkers | HC | 4 | 3 | -- | 40 | 60 | 100 |
| 2 | PH 522.2 | Computational and Translational Immunology | HC | 4 | 3 | -- | 40 | 60 | 100 |
| 3 | PH 523.2P | Structural Proteomics and Computational Immunology Lab | HCL | 3 | -- | 3 | 25 | 50 | 75 |
| **Soft Core Theory & Lab** | | | | | | | | | |
| 4 | PS 524.2 | Statistics for Genomics Applications | SC | 3 | 3 | -- | 40 | 60 | 100 |
| 5 | PS 525.2 | Intermediate Python with Bio-Computing | SC | 3 | 3 | -- | 40 | 60 | 100 |
| 6 | PS 526.2P | Programming: Intermediate python and abaiostatistics | SCL | 3 | -- | 3 | 25 | 50 | 75 |
| **Open Elective – Offered to other Depts** | | | | | | | | | |
| 7 | PO 527.2 | Bioethics, Biosafety and IPR | OE | 3 * | 3 | -- | 40 | 60 | 100 |
| **Add - On** | | | | | | | | | |
| 8 | PS 528.2 | Research & Soft Skill Development – II | SC | -- | -- | -- | -- | -- | -- |
| 9 | PS529.2 | Research Methodology and Ethics | SC | | | | | | |
| | | **Total** | | **20+3 *** | | | **200** | **450** | **650** |

* HC – Hard Core SC – Soft Core HCL – Hard Core Lab SCL – Soft Core Lab OE – Open Elective AON – A

## Master of Science - M.Sc. (Bioinformatics)

## Scheme of Examination || New CBCS Syllabus from 2024

### III Semester

| Sl No | Subject Code | Subject Title | Type | Credits | Theory Exam duration (hours) | Practical Exam duration (hours) | Max Marks for Internal Assessment | Max. Marks for Term End Exam | Total MArks |
|---|---|---|---|---|---|---|---|---|---|
| **Hard Core Theory & Lab** | | | | | | | | | |
| 1 | PH 521.3 | Computational Approaches in Therapeutics and Biosimilar Development | HC | 4 | 3 | -- | 40 | 60 | 100 |
| 2 | PH 522.3 | Next-Gen Sequencing and Single-Cell Technologies for Immunomics | HC | 4 | 3 | -- | 40 | 60 | 100 |
| 3 | PH523.3 P | Laboratory based on Therapeutics and Single Cell Technologies | HC | 3 | - | -- | 25 | 60 | 75 |
| **Soft Core Theory & Lab** | | | | | | | | | |
| 4 | PS 524.3 | Data Science and Machine Learning for Biology | SC | 4 | 3 | -- | 40 | 60 | 100 |
| 5 | PS525.3 | Python Applications in Bioinformatics | SC | 4 | 3 | 3 | 40 | 60 | 100 |
| 6 | PS 526.3P | Lab based on Python and Machine Learning | SCL | 3 | -- | 3 | 25 | 50 | 75 |
| | | **Total** | | | | | | | |

* HC – Hard Core SC – Soft Core HCL – Hard Core Lab SCL – Soft Core Lab OE – Open Elective AON – A

# Master of Science - M.Sc. (Bioinformatics)

## Scheme of Examination || New CBCS Syllabus from 2024

## IV Semester

| Sl No | Subject Code | Subject Title | Credits | Max. Marks for Dissertation / Valuation | | External Evaluation | | Total |
|---|---|---|---|---|---|---|---|---|
| | | | | Dissertation I | Dissertation II | Dissertation III | Max. Marks for Viva-voce (Final) | |
| 1 | PH 521.4 | Industry Internship/ Project Work/ Dissertation | 16 | SRS & SDD (100) | Mid Term Evaluation (100) | Record Reading (100) | 200 | 500 |
| **Sl No** | **Subject Code** | **Subject Title** | **Credits** | **Internal Evaluation** | | **External Evaluation** | | **Total** |
| 2 | PS 522.4 | Domain Knowledge Project | 04 | Assignments / Paper Presentation (50) | | Documentation / Viva Voce (100) | | 150 |
| | **Total** | | **20** | -- | --- | --- | --- | **650** |

## I Semester

## PH 521.1: BIOINFORMATICS FOUNDATIONS: FROM MOLECULES TO SYSTEMS

**Total No. of Lectures : 45** **Total Marks: 100 [ L – T – P – S ]**

**No. of Lectures / Week: 4** **Credits: 4 [ 3 – 1 - 0 - 2 ]**

**Learning Objectives:** This course provides an introduction to key concepts and methods in bioinformatics. Emphasis will be put on efficient algorithms and techniques used in common applications for the analysis of genetic sequences. Topics covered: comparison and alignment of two or more sequences, indexing and searching of sequence databases, motif discovery, searching with sequence patterns, gene prediction as well as mapping and assembly of data from genome sequencing. Necessary basic knowledge of molecular biology will be communicated throughout.

**Learning Outcomes:** On successful completion of the course students will:

- understand fundamental concepts in bioinformatics
- have an overview of the most important methods and tools that are used
- understand how some of the basic methods for biological sequence analysis works
- appreciate the need for methods to be accurate and efficient to implement some of the algorithms
- be capable of performing simple sequence analyses using existing tools

**UNIT I** **9 hrs**

**Introduction to Bioinformatics**

Overview of bioinformatics and its applications in biological research, History and development of bioinformatics, Basic concepts: sequences, structures, and functions of biological macromolecules (DNA, RNA, proteins), Introduction to databases and resources in bioinformatics ; Nucleotide Sequence Databases: GenBank, EMBL, DDBJ, Genomic Databases; Ensembl, UCSC Genomic Browser, Protein Sequence databases; Uniprot, Structural Databases: Protein Data Bank (PDB), SCOP (Structural Classification of Proteins, Pathway Database; KEGG, Reactome, Gene Expression Databases; GEO

**UNIT II** **9hrs**

**Bioinformatics tools and software:** Biological Databases and Data Retrieval, Types of biological databases: nucleotide, protein, genomic, and structure databases, Data formats (FASTA, GenBank, PDB) and their significance, Searching and retrieving data from biological databases, Introduction to sequence alignment algorithms (pairwise and multiple sequence alignment), Practical exercises on database querying and retrieval.

**UNIT III** **9hrs**

**Sequence Analysis:** Sequence similarity and homology, Basic sequence analysis techniques: sequence alignment, similarity searching, Pairwise sequence alignment algorithms (Needleman-Wunsch, Smith-Waterman), Multiple sequence alignment algorithms (ClustalW, MUSCLE), Phylogenetic analysis and tree-building methods.

**UNIT IV** **9hrs**

**Structural Bioinformatics:** Protein structure prediction methods: homology modeling, ab initio modelling, Protein structure databases (PDB, SCOP, CATH), Protein structure visualization tools, Analysis of protein structure: secondary structure prediction, solvent accessibility prediction, Protein-ligand interaction analysis.

**UNIT V** **9hrs**

**Functional Genomics :** Introduction to gene expression analysis, Microarray technology and data analysis, Next-generation sequencing (NGS) technologies and data analysis pipelines, Gene ontology and functional annotation, Systems biology approaches to understanding biological processes

**Text Book:**

[1] Anna Tramantono, "Introduction to Bioinformatics", 3rd Edition, 2012, CRC Press

[2] Dan E Krane, Raymar, "Fundamental Concepts of Bioinformatics", 1st Ed, 2010, Pearson

[3] Arthur M Lesk, "Introduction to Bioinformatics", 3rd Edition, 2014, Oxford University Press

**Reference Books:**

[1]. Zhumar Ghosh, B Mallick, "Bioinformatics: Principles and Applications", 1st Ed, 2011, Oxford Un Press.

[2]. Andreas D Baxevanis, Alexnder Zelikovisky, "Bioinformatics: Practical Guide to the analysis of Genes and Proteins", 3rd Edition, 2013, Wiley International Ed.

[3]. Orpitha Bosu, S K Thukral, "Bioinformatics: Databases, Tools, Algorithms",2nd Ed, 2012, Oxford Un Press

[4]. S Harisha, "Fundamentals of Bioinformatics", 1st Edition, 2010, I K International Publishers

[5]. Cynthia Gibas, Per Jambeck, "Developing Bioinformatics Computer Skills : An Introduction to Software Tools for Bioinformatics Applications", 1st Edition, 2011, O'Reilly Publishers

[6]. Teresa Attwood, David J, Pary Smith, "Introduction to Bioinformatics", 1st Ed, 2009, Pearson

[7]. Ion Mandoiu, A Zelikovisky, "Bioinformatics Algorithms : Techniques and Applications", 2012, Wiley

[8]. Vittal R Srinivas, "Bioinformatics: A Modern Approach", 1st Edition, 2011, PHI Publishers

[9]. Rastogi S, C Rastogi, "Bioinformatics : Methods and Applications", 2nd Ed, 2009, PHI Publishers

[10]. Prakash S Lohar, "Bioinformatics", 1st Edition, 2009, MJP Publishers, Chennai

## PH 522.1: ALGORITHMS IN COMPUTATIONAL BIOLOGY

**Total No. of Lectures: 45** **Total Marks:100 [ L – T – P – S ]**

**No. of Lectures / Week: 4** **Credits: 4 [ 3 – 1 - 0 - 2 ]**

**Learning Objectives:**

- To provide an introduction to essential concepts of Bioinformatics, Computational Biology, Systems Biology.
- To master abstraction and modeling abilities necessary in precise formulation of computational structures and their interpretations in life sciences.
- To support the computer-based exploratory thinking and experimenting in life sciences.

**Learning Outcomes:** Upon completing the course, the student should:

- Be fluent in using data structures and algorithms to design programs that model basic computational problems in life sciences related to sequence, structure and function of biological entities
- Have substantial knowledge of discrete mathematics issues useful in modeling living systems
- Apply appropriate techniques to deal with complexity of problems and complexity of programs
- Be able to apply this knowledge to independently model fundamental concepts in life sciences such as gene and their product annotation and protein-protein interaction

### UNIT I 9hrs

**Algorithm Design and Analysis Basics:** Introduction to algorithms and their role in bioinformatics, Basics of algorithm analysis: time complexity, space complexity, Greedy algorithms and their applications in bioinformatics, Dynamic programming algorithms and their significance (e.g., Sequence Alignment)

### UNIT II 9hrs

**Sequence Alignment Algorithms:** Needleman-Wunsch algorithm for global sequence alignment, Smith-Waterman algorithm for local sequence alignment, Affine gap penalties in sequence alignment, Optimization techniques in sequence alignment algorithms, Practical implementation of sequence alignment algorithms

**UNIT III** **9hrs**

**Pattern and Motif Search Algorithms**: Introduction to sequence motifs and their significance, Brute-force and heuristic methods for motif searching, Position weight matrices (PWM) and their applications in motif search, Gibbs sampling algorithm for motif discovery, Practical exercises on motif search using software tools, Hidden Markov Models (HMMs)

**UNIT IV** **9hrs**

**Introduction to HMMs and their applications in bioinformatics:** Basics of probability theory and Markov chains, Forward algorithm for HMMs and its significance in sequence analysis, Baum-Welch algorithm for parameter estimation in HMMs, Applications of HMMs in sequence alignment, gene finding, and motif discovery,

**UNIT V** **9hrs**

**Phylogenetic Tree Construction Algorithms**: Distance-based methods: Neighbour-Joining algorithm, Character-based methods: Maximum parsimony, Maximum likelihood, Bootstrap resampling and consensus tree construction, Evaluation of phylogenetic trees, Practical exercises on phylogenetic tree construction using software tools

.

**Text Book:**

[1]. Ka Chun Wang, "Computational Biology & Bioinformatics", 1st Edition, 2011, CRC Press.

[2]. Arabnia, Tran, "Emerging Trends in Computational Biology, Bioinformatics, and Systems Biology, Algorithms and Software Tools", 1st Edition, 2015, Elsevier Publishers

[3]. M Moody, "Structural Biology Using Electrons and X-rays", 3rd Ed, 2012, Elsevier Publishing

**References:**

[1]. Robbe Wunschiers, "Computational Biology", 2nd Ed, 2008, Springer Publications

[2]. John M Hanock, Marketa J, "Bioinformatics and Computational Biology", 1st Ed, 2011, Wiley Blackwell

[3]. Michael S Waterman, "Introduction to Computational Biology", 2nd Ed, 2009, Chapman & Hall Pub

[4]. Roland Eils, Andreas Kriete, “Computational Systems Biology”, 2nd Ed, 2012, Allied Publishers

[5]. Mary Luckey, “Membrane Structural Biology”, 1st Edition, 2008, Cambridge University Press

[6]. Bruce R Donald, “Algorithms in Structural and Molecular Biology”, 1st Ed, 2010, Wiley International

[7]. B. Thiagarajan,Pa Rajalakshmi, “Computational Biology”, 1st Edition, 2009, MJP Publishers

## PH 523. 1: BASIC BIOCHEMISTRY AND IMMUNOLOGY

| | |
|---|---|
| **Total No. of Lectures : 45** | **Total Marks : 100 [ L – T – P – S ]** |
| **No. of Lectures / Week : 4** | **Credits : 4 [ 3 – 1 - 0 - 2 ]** |

**Learning Objectives**: This course provides the knowledge of the body uptake of nutrients and these storing and renewal in different physiological conditions. The student should be able to integrate the received knowledge so that within this field can explain the background of relevant diseases. The study of immunology will enable the student to gain a broad foundation base and build upon that base for understanding the defense mechanisms of the human body.

**Learning Outcome:** As a result of successfully completing this course, the student will be able to do the following:

- Demonstrate a comprehensive and practical understanding of basic immunological principles
- Differentiate between innate and adaptive immunity.
- Explain the mechanisms and differences between primary and secondary responses and their relevance to immunizations.
- Identify the role of antigen presenting cells, lymphocytes, and phagocytic cells in immune responses.

**UNIT I** **9hrs**

**Proteins & Amino Acids:** Amino acids and their physio-chemical properties, peptide bond and peptides, disulphide cross links, various levels of structural organization of Proteins. Classification- size, shape, degree of association, complexity. Classification of proteins according to biological functions (Enzymes,transport, storage, contractile, structural, defense and regulatory) Ramachandran plot, Alpha-helix, Beta sheet, Helix-coil transitions. Tertiary and Quaternary Structures, Protein folding and domains.

**UNIT II** **9 hrs**

**Basic enzymology**; General properties of enzymes, substrate specificity and active site. Enzyme kinetics, Michales-Menten equation.

Organization of metabolic systems: enzyme chains, multi-enzyme complexes, Substrate Channeling and features, Multifunctional enzymes and Regulatory enzymes, Enzyme Inhibition, Allosteric Enzymes.

**UNIT III** **9hrs**

**Metabolism**: Metabolism: Overview of metabolism, high energy compounds, Oxidation-reduction reactions, the reactions of glycolysis, Control of glycolysis. Glycogen breakdown and synthesis, Control of glycogen metabolism and Gluconeogenesis. Citric acid cycle: Enzymes of the Citric Acid cycle, Regulation of the citric acid cycle. Signal Transduction and Regulation of Metabolism. Computational Resources; KEGG and REACTOME.

**UNIT IV** **9hrs**

**Immune System Components**: Innate and acquired immunity, active and passive immunity, natural and artificial immunity and humoral. Lymphoid system- primary or secondary organ. Complement System- activation, pathways and biological effects. Major Histo-Compatibility complexes- Structure and Function.

**UNIT V** **9hrs**

**Immune Response and Antibody Generation**: T and B Lymphocytes, TCR and BCR, B Cell generation and Differentiation, Antibody Generation: Structure and Functions –Clonal Selection theory-different types of Immunoglobulins, Antibody Diversity and Class Switching. Antigen and antibody reaction/interaction: Precipitation, Haemagglutination, ELISA, Western Blot and Immuno-fluorescence.

**Text Book:**

[1] William E Paul, "Fundamental Immunology", 6th Edition, 2015, Morgan Kaufmann Publishers

[2] David M Nelson and Michael Cox, Lehninger Principles of Biochemistry, Fifth Edition, Freeman and Company

[3] Jeremy M. Berg | John L. Tymoczko | Lubert Stryer Biochemistry, 6th edition

**Reference books:**

[1]. Maurice R.G. O'Gorman, Albert D. Donnenberg, "Handbook of Human Immunology", 2nd Ed, 2014, CRC

[2]. Rafael Fernandez, Vaclav Vetvicka, "Advanced Methods in Cellular Immunology", 2010, CRC

[3]. Peter Wood, "Understanding Immunology", 3rd Edition, 2012, Pearson Asia

[4]. Lauren M. Sompayrac, "How the Immune System Works", 5th Edition, 2015, Wiley International

[5]. George Pinchuk, "Immunology", 3rd Edition, 2010, Tata McGrwaw Hill Publishers

[6]. A A Haleem Khan, "Text Book of Immunology", 2nd Edition, 2010, Taylor & Francis International

[7]. Abdul K Abbas, Andrew Litchman, Shiv Pillai, "Cellular and Molecular Immunology", 6th Edition, 2014, Elsevier Publishers

[8]. Raj Khanna, "Immunology", 1st Edition, 2013, Oxford University Press

[9]. Ashim K Chakravarthy, "Immunology and Immunotechnology", 2nd Ed, 2010, Oxford Un Press

## PH 524.1: CELL AND MOLECULAR BIOLOGY

**Total No. of Lectures : 45** **Total Marks : 100** **[ L – T – P – S ]**
**No. of Lectures / Week : 4** **Credits : 4** **[ 3 – 1 - 0 - 2 ]**

**Learning Objectives:** This course is designed to introduce and study, understanding of the structure, genetics, biochemistry, and physiology of cells. The cell is the basic fundamental unit of life. All the processes of life, including harnessing energy, reproduction, inheritance of characteristics, and responding to the environment, can only be fully appreciated with an understanding of their cellular bases.

**Learning Outcomes:** After the course, the student should - be able to

- Describe the chemical components of the macromolecules of life and their functions
- Describe the structural differences between prokaryotic and eukaryotic cells or between plant and animal cells
- Classify living organisms within taxonomic groups
- Explain the differences and similarities between photosynthesis and cellular respiration
- Predict the outcome of a reaction catalyzed by enzymes under different conditions
- Apply the concepts of replication, transcription and translation

**Unit I** **9 hrs**

**Molecules of life** - Structural organization of prokaryotic and eukaryotic cells- Concept of a composite cell and Molecular composition of cells. Biomembranes- Structural organization- Models of a plasma membrane, Membrane permeability- Transport across cell membranes- Transmembrane signals- Artificial membranes- liposome.

**Unit II** **9 hrs**

**Mitochondrial Structure and Function** – Oxidative Metabolism in the Mitochondrion – The Role of Mitochondria in the formation of ATP – Translocation of Protons and the Establishment of a protonmotive force – The Machinery for ATP formation – Peroxisomes.

**Unit III** **9 hrs**

**Chloroplast structure and function** – An overview of photosynthetic Metabolism – The absorption of light – Photosynthetic units and reaction centers – Photophosphorylation – Carbondioxide fixation and the synthesis of carbohydrates

**Unit IV** **9 hrs**

**Cellular Components** – Cytoskeleton – components of Cytoskeleton, Microtubules, Intermediate filaments – Microfilaments, Cell cycle, Endoplasmic reticulum, Golgi complex, Types of vesicles - transport and their functions, Lysosomes.

**Unit V** **9 hrs**

**DNA and Protein Synthesis** - Structure of DNA - evidence for DNA as genetic material. Gene transfer in microorganisms – conjugation, transformation, transduction - DNA replication, Transcription – mRNA processing, Translation. Protein synthesis – Ribosomes, enzymes, Protein processing.

**Textbook:**

[1]. Gerald Karp, “Cell and Molecular Biology”, 5th Edition, 2013, John Wiley & Sons

[2]. Y Bruce Alberts, Alexander Johnson, Julian Lewis , David Morgan, Martin Raff, Keith Roberts, Peter Walter, “Molecular Biology of the Cell”, 6th Edition, Garland Science, Taylor & Francis Pub

[3]. James D. Watson, Tania A. Baker, Stephen P. Bell, Alexander Gann, Michael Levine, Richard Losick, “Molecular Biology of the Gene”, 7th Edition, Pearson Education International

**References:**

[1]. Nalini Chander, Susan Viselli, “Cell and Molecular Biology”, 3rd Edition, 2009, Pearson

[2]. Philp Sheeler, Donald Bianchi, “Cell and Molecular Biology”, 3rd Edition, 2014, Wiley India Ltd

[3]. P K Gupta, “Cell and Molecular Biology”, 4th Edition, 2008, Rastogi Publishers

[4]. Harvey Lodish, Arnold Berk, Chris A. Kaiser, Monty Krieger, Matthew P. Scott, Anthony Bretscher, Hidde Ploegh, Paul Matsudaira , “Molecular Cell Biology”, 6th Edition, 2008, W H Freeman Publishers

[5]. N Arumugam, “Molecular Biology”, 1st Edition, 2014, Saras Publication

[6]. Bruce Alberts, Dennis Bray , Karen Hopkin, Alexander D Johnson, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts , Peter Walter, “Essential Cell Biology”, 3rd Edition, 2009, Garland Science, T&F

[7]. Geoffrey M. Cooper , Robert E. Hausman, “ The Cell : A Molecular Approach”, 5th Edition, 2009, Sinauer Associates Publishers

[8]. S P Vyas, A Mehtha, “Cell and Molecular Biology”, 1st Edition, 2011, CBS Publishers

[9]. Peter J. Russell , Paul E. Hertz, Cecie Starr, “ Cell and Molecular Biology”, 1st Ed, 2009, Cengage Learning

## PH 525.1P BIOINFORMATICS AND COMPUTATIONAL BIOLOGY LAB

**Total No. of Practicals: 90** **Total Marks : 75 [ L – T – P – S ]**

**No. of Labs / Week: 3** **Credits : 3 [ 0 – 1 - 3 – 0 ]**

**Learning Objectives:**

- Understand the principles and some methods of genomics, gene expression and proteomics
- Understand the concepts of the new, high-throughput and high-noise biology.
- Search large databases, interpret their results.
- Analyze gene expression using data from microarrays or RNA sequencing.
- Analyze metabolomic, proteomics, and protein-protein interaction experiments

**Learning Outcomes:** After this course, the student will be able to:

- Implement computational methods to solve problems involving measurement data.
- Perform data acquisition from raw data.
- Independently search for information and available methods to solve practical problems.
- Present results, methods and conclusions in written and oral reports.

### Bioinformatics Lab

1. Sequence retrieval from nucleic acid and protein databases.
2. Retrieval of information about structure, bioassay, physical and Chemical properties of chemical compounds (such as Drugs and naturally occurring compounds).

3. Gene sequence assembly and contig mapping and identification of Gene.
4. Primer and Promoter design for a given sequences
5. Sequence searches using FASTA and BLAST, and Phylogenetic analysis.
6. Prediction of secondary structure for given protein and RNA sequences.
7. Retrieval of protein structure from PDB and its visualization and modification.
8. Prediction of 3D structure of unknown protein sequence.
9. Prediction of protein-protein interactions.
10. EST clustering and EST mapping, and Genome annotation
11. Microarray data analysis- normalization, clustering.
12. Study of Profiles, Patterns and PSSMs

**Computational Biology Lab**

1. Sequence Analysis Packages – EMBOSS, NCBI ToolKit study
2. Dynamic programming.
3. Analysis of Biological Sequences: Basic Blast and Specialized Blast
4. FASTA
5. Multiple sequence alignment
6. MEME/MAST, eMotif, InterproScan, ProSite, ProDom, Pfam
7. Phylogenetic analysis – PAUP, PHYLIP, MacClade
8. Advanced Visualization Software and 3D representations.
9. Coordinate generations and inter-conversions.
10. Secondary Structure Prediction: Fold Recognition, *ab initio* (Rosetta Server), Homology based comparative protein modeling, Energy minimizations, Validation of models (WHATIF, PROSA, PROCHECK, VERIFY 3D), Protein Structure Alignmemt

**Biological Databases Lab:**

1. Retrieval of sequences from NCBI
2. ORF Finder
3. Retrieval of sequences from EBI
4. Retrieval of structures from PDB
5. Rasmol, Pymol, Jmol
6. Retrieval of sequences from DDBJ
7. KEGG
8. Retrieval of sequences from PIR
9. OMIM

10. Structural databases, SRS

**Sequence Analysis Lab:**

1. Sequence Retrieval - Fetch
2. Importing and Exporting – Reformat
3. Pairwise comparison – Gap, Bestfit, Compare, Fr
4. Multiple comparison - Pileup, Pretty, Profile Make
5. Sequence Searching – BLAST, FASTA
6. Evolution – Distances, Diverge
7. Gene Finding and Pattern Recognition – FindPatt
8. HMMer – HMMer Build, HMMer Align
9. Mapping – Map, Mapsort
10. Protein Analysis – Isoelectric, Helical Wheel
11. Sequence analysis by online tools
12. BLAST, FASTA, CLUSTALW, PHYLIP etc.,

**Note :** Facilitator can modify / add more exercises according to the academic need.

## PS 526.1 BASICS OF PROGRAMMING WITH PYTHON

| | | |
|---|---|---|
| **Total No. of Lectures : 45** | **Total Marks : 100** | **[ L – T – P – S ]** |
| **No. of Lectures / Week : 3** | **Credits : 3** | **[ 3 – 0 - 2 - 2 ]** |

**Learning Objectives:** The objective of this course is to make the students:

To understand the fundamentals of computer and linux environment

Describe the core syntax and semantics of Python programming language.

Discover the need for working with the strings and functions.

Illustrate the process of structuring the data using lists, dictionaries, tuples and sets.

Indicate the use of regular expressions and built-in functions to navigate the file system.

Infer the Object-oriented Programming concepts in Python.

**Learning Outcomes:** At the end of this course students will be able

To learn the concept on fundament of computer and linux environment.

To learn how to design and program Python applications, use lists, tuples, and dictionaries in Python programs.

To learn how to identify Python object types, use indexing and slicing to access data in Python programs.

To learn how to build and package Python modules for reusability and to read and write files in Python.

To learn the object oriented programs with Python classes and use class inheritance in Python for reusability.

To learn how to use exception handling in Python applications for error handling.

**Unit I** **9hrs**

**Introduction to computers and programming: Introduction**, Hardware and Software, How Computers Store data, How a program works, Using Python.

Introduction to Python Basics: Entering Expressions into the Interactive Shell, The Integer, Floating-Point, and String Data Types, String Concatenation and Replication, Storing Values in Variables, Your First Program, Dissecting Your Program.

**Unit II** **9hrs**

**Program Flow Control & Functions**: Conditional blocks using if, else and elif, for loops in, for loop using ranges, use of while loops in python, Loop manipulation using pass, continue, break and else. Organizing python codes using functions, organizing python projects into modules, importing own module as well as external modules, Understanding Packages. Inbuilt function, maps, comprehension, anonymous functions, iterators and generators

**Unit III** **9hrs**

**String, List, Dictionary &File handling**:, Introduction to lists and list slicing, Use of Tuple data type. List manipulation, Dictionary manipulation. Programming using string, list and dictionary in-built functions. Reading files in python, writing files in python, Understanding read and write functions, manipulating file pointer using seek, Programming using file operations.

**Unit IV** **9hrs**

**Object Oriented Programming & Regular Expressions**: Concept of class, object and instances, constructor, class attributes and destructors, Inheritance, overlapping and overloading operators, Adding and retrieving dynamic attributes of classes. Pattern matching and searching, Power of pattern searching using regex in python, Pattern finding programs using regular expression.

Exception Handling & Database Interaction: Avoiding code break using exception handling, Safe guarding file operation using exception handling, Handling and helping developer with error code, SQL Database connection using python, Creating and searching tables, Reading, storing and updating information on database.

**Unit V** **9hrs**

**Linux Basics**: Introduction to Linux, File System of the Linux, General usage of Linux kernel & basic commands, Linux users and group, Permissions for file, directory and users, Searching a file & directory.

The File system –The Basics of Files-What's in a File-Directories and File Names-Permissions INodes-The Directory Hierarchy, File Attributes and Permissions-The File Command knowing the File Type-The Chmod Command Changing File Permissions-The Chown Command Changing the Owner of a File-The Chgrp Command Changing the Group of a File, Windows Subsystem for Linux (WSL).

**Text Books:**

1. R. Nageswara Rao, "Core Python Programming", 2nd Edition, 2018, Dreamtech Press.
2. Jason Rees, "Python Programming: A Practical Introduction To Python Programming For Total Beginners Paperback", 2019, Independent Publication.
3. Al Sweigart, "Automate the Boring Stuff with Python", William Pollock, 2015, ISBN: 978-1593275990.
4. Introduction to Unix Shell Programming by M.G.Venkateshmurthy, Pearson.

**Reference Books:**

1. Eric Matthes, "Python Crash Course, 2nd Edition: A Hands-On, Project-Based Introduction to Programming", 2019, No Starch Press.
2. Zed A. Shaw, "Learn Python 3 the Hard Way", 2017, Addison Wesley.
3. John M Zelle, "Python Programming: An Introduction to Computer Science", 2016, Ingram short title; Third edition.

4. David Beazley, Brian K. Jones, “Python Cookbook: Recipes for Mastering Python 3, 2013, Oreilly 3rd Edition.

## PS 527.1P PROGRAMMING with Python LAB

|  |  |
|---|---|
| **Total No. of Practicals : 90** | **Total Marks : 75 [ L – T – P – S ]** |
| **No. of Labs / Week : 3** | **Credits : 3 [ 0 – 1 - 3 – 0 ]** |

**Learning Objectives:**

Understand string manipulation techniques in Python.

Learn various methods to find the length of lists.

Master basic arithmetic operations in Python.

Understand how to manipulate functions and parameters.

Learn to work with files, including reading, writing, and checking file size.

Gain proficiency in tuple operations and summation.

**Learning Outcomes:**

Able to determine if a given string is symmetrical or a palindrome.

Competent in removing specific characters from a string.

Proficient in identifying and printing even-length words in a string.

Competent in interchanging the first and last elements of a list.

Able to swap two elements within a list.

Capable of finding the power of a number using recursion.

Able to create an empty class in Python.

Proficient in implementing a student management system.

Capable of reading content from one file and writing it into another.

Proficient in finding the most repeated word in a text file.

1. Python program to check whether the string is Symmetrical or Palindrome.
2. Reverse words in a given String in Python.
3. Ways to remove i’th character from string in Python.
4. Find length of a string in python.
5. Python program to print even length words in a string.
6. Python program to interchange first and last elements in a list.
7. Python program to swap two elements in a list.
8. Python | Ways to find length of list.
9. Maximum of two numbers in Python.
10. Minimum of two numbers in Python.
11. How to get list of parameters name from a function in Python.
12. How to Print Multiple Arguments in Python.
13. Python program to find the power of a number using recursion.
14. How to create an empty class in Python.
15. Student management system in Python.
16. Read content from one file and write it into another file.
17. Write a dictionary to a file in Python.
18. How to check file size in Python.
19. Find the most repeated word in a text file.
20. Python – Sum of tuple elements.

**Note:** Facilitator can modify/ add more exercises according to the academic need

## PS 528.1 RESEARCH & SOFT SKILL DEVELOPMENT – I

**Total No. of hrs : 30** **Total Marks : --- [ L – T – P – S ]**
**No. of Hrs / Week : 2** **Credits : --- [ 0 – 2 - 0 – 4 ]**

**Objectives:** To give the Masters student experience and training in bioinformatics research through a high quality research project either in industry or academia. To understand and experience:

1. The technical skills required for efficient bioinformatics research;

2. The organisation and implementation of research;
3. Working to deadlines and as part of a team;
4. The skills required to record, report and write up original scientific research.

**Research projects**

The research project will normally be carried out in the workplace by the student and will be supervised by a supervisor. The Research Projects are the part of the Research activities initiated by the Department of Biotechnology, Ministry of Science and Technology, Government of India. For each research project, the Dept can appoint two supervisors i.e One from PG Dept of Bioinformatics and the other from the PG Depts of Life Sciences of the college..

***The responsibilities of the Research supervisor include:***

1. Giving guidance about the nature of research and the standard expected, the planning of the research programme, literature and sources, attendance at taught classes where appropriate and about requisite techniques (including arranging for instruction where necessary);
2. ensuring that, where relevant, the work-based on some needy and quite relevant area in Bioinformatics;
3. maintaining contact through regular interaction in the Lab (Twice a week);
4. being accessible to students at other appropriate times for advice and responding to difficulties raised by the student;
5. giving detailed advice on the necessary completion dates of successive stages of the work so that the thesis may be submitted within the agreed timescale;
6. requesting written work or reports as appropriate and returning written material with constructive criticism and in reasonable time;
7. ensuring that the student is made aware when progress is not satisfactory and facilitating improvement with advice and guidance;
8. establishing, at an early stage, the supervisor's responsibilities in relating to the student's written work, including the nature of the guidance and comments to be offered as the work proceeds and on the draft of the report before it is submitted.
9. encouraging the student to publish the research;
10. providing pastoral support and advising students, where appropriate, of College support services(E.g Softwares/ Lab facility / Library/ Journal);
11. bringing to the attention of the students the health and safety regulations and academic rules, regulations and codes of practice of the College;
12. bringing to the attention of the students, regulations concerning absence due to ill health.

***The responsibilities of the student include:***

1. pursuing the programme with a positive commitment, taking full advantage of the resources and facilities offered by the academic environment and, in particular, contact with the supervisor, other staff and research students;
2. discussing with the supervisor the type of guidance and comment believed to be most helpful, and agreeing a schedule of on-line meetings;
3. successfully completing any training programme arranged within the prescribed time period;
4. taking the initiative in raising problems or difficulties, however elementary they may seem, bearing in mind that prompt discussion and resolution of problems can prevent difficulties and disagreements at a later stage;
5. maintaining the progress of the work in accordance with the stages agreed with the supervisor, including in particular the presentation of written material as required, in sufficient time to allow for comments and discussion before proceeding to the next stage. Where possible, students will be given details of the work programme for the academic year at the beginning of the year;
6. agreeing with the supervisor the amount of time to be devoted to the research and the timing and duration of holiday periods;
7. deciding when to submit the report. The student should take due account of the supervisor's opinion but must recognise that it is only advisory. The student must ensure that appropriate notice of intent to submit the report.
8. checking the completeness and accuracy of the text of the report submitted; failure to check the report carefully may result in the report being failed .

### Choosing your topic

The results of the supervised research project will be presented as a **bound dissertation** for **40 pages** only. The literature project gives experience in abstracting and assessing information you have received during the taught element of the course and reviewing additional information as necessary. It also gives you an opportunity to develop your communication skills through the planning and writing of the report.

### Assessment of the Report

Research dissertations submitted by students in the Faculty of Bioinformatics are examined by internal and external examiners. The external examiner acts in a moderating role in relation to all the dissertations submitted, after the reports have been individually marked by the internal & external examiner separately and independently, with no prior knowledge of the assessment made by each other. The student has to submit two copies of the Project Report.

## SEMESTER II

## PH 521.2: STRUCTURAL PROTEOMICS AND BIOMARKERS

**Total No. of Lectures : 45** **Total Marks : 10 [ L – T – P – S ]**
**No. of Lectures / Week : 4** **Credits : 4 [ 3 – 1 - 0 - 2 ]**

**Learning Objectives:** This paper deals with a rapidly evolving scientific area that introduces students into genomes, proteomes and databases that store various data about genes, proteins, genomes and proteomes. The main objective is to organize the large amount of information about genomics, proteomics and bioinformatics and offer basic knowledge of genome sequencing, major differences between prokaryotic and eukaryotic genomes, basic proteomics and its applications, basics in bioinformatics, comparative and evolutionary genomics and applications.

**Learning Outcome:** Upon successful completion of this subject, students should:

- Master Protein Structures: Understand the principles governing protein architecture, with a focus on tertiary structures.
- Navigate Proteomics: Demonstrate proficiency in proteomic techniques for large-scale protein studies.
- Understand Structure-Function Relationships: Investigate the dynamic links between protein structure and function.
- Integrate Interdisciplinary Knowledge: Generate concepts from biochemistry, structural biology, and computational biology.

Apply Bioinformatics for Biomarker Discovery: Utilize bioinformatics tools to discover potential biomarkers, particularly in medical and diagnostic contexts.

Apply protein function and proteomic knowledge, immersing in real problems and case studies.

**UNIT I** **9hrs Protein Architecture** Primary structure: peptide mapping, peptide, Secondary structure: Alpha, beta and loop structures and methods for analyses. Super-Secondary structure: Alpha-turn alpha, beta-turn- beta (hairpin), beta-sheets, alpha-beta-alpha, topology diagrams, up and down & TIM barrel structures nucleotide binding folds**.**
Protein Structure Classification Schemes and resources in Bioinformatics.

**UNIT II** **9hrs**
**Protein Tertiary Structure** Prediction of substrate binding sites, Tertiary structure: Domains, folding, denaturation and renaturation, Overview of methods to determine 3D structures; X-Ray Crystallography and NMR, Quaternary structure: Modular nature, formation of complexes, Protein-Protein interactions. Computational modelling of protein structures, Structural Assessment of Proteins-Ramachandran Plot and Protein Stability.

**UNIT III** **9hrs**

**Proteomics** Introduction to the concept of proteome, Components and Proteomic analysis, Importance of proteomics in biological functions, Protein arrays, EXPASY platform for Proteomic Analysis and Expression Profiling.

Protein Sequential Patterns and Functional Signatures-Motifs, Patterns, Profiles and Fingerprints. Computational tools and resources for analysis of functional signatures: ScanProsite and InterPro.

## UNIT IV 9hrs

**Bioinformatics assisted Biomarker Discovery** Understanding Flowcytometry and Flowcytometric Data, Cell Populations, Computational resources to analyse Flowcytometry Data, Fluorescent Activated Cell Sorting, Immunophenotyping, CD markers in Cancer Biology, Challenges in Cancer Bio-marker Discovery.

## UNIT V 9hrs

**Structure-Function Relationship** Case Studies; Structural motifs: Helix-turn-Helix, Zn fingers, Leucine zippers, Beta Barrels, Membrane proteins: General characteristics, Trans-membrane segments prediction, Immunoglobulins: IgG Light chain and heavy chain architecture, Enzymes: Domains, Active Sites, Functional Signatures, Multidomain organisation.

## TEXT BOOKS:

1 Branden C. and Tooze J., “Introduction to Protein Structures” 2nd Edition, Garlan Publishing, 1999’

2. Creighton T.E. “Proteins” 2nd Edition. W.H. Freeman, 1993.

3. Pennington, S.R and M.J. Dunn, “Proteomics: Protein Sequence to Function”. VivaBooks, 2002

4. Jeremy M. Berg | John L. Tymoczko | Lubert Stryer Biochemistry, 6th edition

5. David M Nelson and Michael Cox, Lehninger Principles of Biochemistry, Fifth Edition, Freeman and Company

6. Gabriel Waksman Proteomics and Protein-Protein Interactions Biology, Chemistry, Bioinformatics, and Drug Design, Springer 2005.

7. Huzefa Rangwala , George Karypis, Introduction to Protein Structure Prediction- Methods and Algorithms, Springer 2005.

## Reference:

1. Liebler, “Introduction to Proteomics” Humana Press, 2002. 2. David Whitford, “Proteins: Structure and Function” John wiley& Sons ltd, 2005

## PH 522.2: COMPUTATIONAL AND TRANSLATIONAL IMMUNOLOGY

**Total No. of Lectures : 45** **Total Marks : 100[ L – T – P – S ]**
**No. of Lectures / Week : 4** **Credits : 4 [ 3 – 1 - 0 - 2 ]**

**Learning Objectives:** The aim of this course is to prepare students to navigate the dynamic and interdisciplinary field of immunology, applying theoretical knowledge to practical scenarios in antibody discovery, computational analysis, engineering, and in silico vaccine design.

**Learning Outcomes:** Upon completion of this course, students will

- Master Cellular & Molecular Immunology: Understand immune system principles, including cellular functions, signaling, and immune responses.
- Comprehend Antibody Discovery Techniques: Learn methods for isolating and producing antibodies for therapeutic and research purposes.
- Utilize Antibody Computing: Proficiently apply computational tools for antibody design, analysis, and optimization, integrating bioinformatics with immunology.
- Excel in Antibody Engineering: Explore advanced concepts to modify antibody properties for enhanced therapeutic efficacy.
- Acquire Insilico Vaccine Design Skills: Develop competence in designing vaccines using computational methods, optimizing immunogenicity and epitope prediction.

### UNIT I 9hrs

**Cellular & Molecular Immunology** Immune System Components, T and B lymphocytes, T Cell- dependent activation of B cells, B-lymphocyte signal transduction. T cell receptors (TCR), Cytokines. B Cell generation and differentiation: BCR & Antibody diversity, structure and functions.

The Major Histocompatibility Complex (MHC) & Polymorphism, Supertypes. Antigen Processing in the MHC Class I & II Pathways.

### UNIT II 9hrs

**Antibody Discovery** Recognition by B cells, Affinity Maturation, Hybridoma Cell Lines and Production of monoclonal antibodies, Cloning and expression of antibody chains in mammalian-based systems, Antibody Chain Production; RACE, cDNA, PCR and ScFv. Antibody Characterisation, Antibody Neutralisation Study, Recognition of Antigen by B cells. Affinity Maturation, Antibody Characterisation, Neutralizing Antibodies and their therapeutic significance.

### UNIT III 9hrs

**Antibody Computing** Current knowledge of Sequence-Structure-Functional correlations of antibodies,

Databases and resources associated with Immunoglobulins (IG), Immunoglobulins: Structure and functions, CDR analysis and Numbering Schemes,

Generation of antibody repertoires, Antibody Modelling Strategies and Resources, Challenges in Antibody Modelling, Structural Characterisation of CDR-H3; Implications for Antibody Modelling, Compuatational Modelling of TCR and CDR3, Analysis of ScFv.

Computational study of Antibody Interaction; Antibody Docking, Interface Properties.

### UNIT IV 9hrs

**Antibody Engineering** Isotype Characterisation, Generation of Heavy and Light Chains, Computer Assisted Antibody Engineering, CDR grafting, Computational study of affinity maturation, In silico Deimmunisation, Human antibodies by guided selection, Improved antibody fragments for Pharmacokinetics and Pharmacodynamics and Engineering Antigen-Antibody Interactions.

**UNIT V** **9hrs**

**Insilico Vaccine Design** Linear and Conformational Epitopes, Epitope mapping strategies, Computational resources for linear and conformational epitope prediction, Parameters in B cell epitope prediction. Custom Antigen design. Prediction of Cytotoxic T Cell Epitopes. ELISA and Western Blot based characterisation of epitopes.

**Text Book**

1. Gary C. Howard, Delia R. Bethell Basic Methods in Antibody Production and

2. Characterization, CRC Press , 2001
3. Johan Rockberg, Johan Nilvebrant Epitope Mapping Protocols, Humana Press

4. Third Edition
5. Kouhei Tsumoto and Daisuke Kuroda. Computer-Aided Antibody Design

6. Humana Press
7. Roland Kontermann and Stefan Dubel Antibody Engineering Volume 1, Springer

8. Third Edition.
9. Roland Kontermann and Stefan Dubel, Antibody Engineering, Volume 2, Springer,

10. Second Edition.
11. Ole Lund, Darren Flower, “Immunological Bioinformatics”, MIT press, Springer

12. September 2005, 2006.
13. Darren R Flower “Immunoinformatics: Predicting Immunogenicity in Silico” Humana Press 2007

14. Thomas J. Kindt, Richard A. Goldsby, Barbara A borne Kuby Immunology W. H. Freeman & Company 2006

**Reference**

1 Rammensee, “Immunoinformatics- Bioinformatics Strategies for Better Understanding of Immune Function” Wiley, 2003
2 Shyamasree Ghosh, Computational Immunology: Basics CRC Press 2020

## PH 523.2P STRUCTURAL PROTEOMICS AND COMPUTATIONAL IMMUNOLOGY LAB

**Learning Outcome :** Upon completion of this lab course, students will:

- Acquire Structural Proteomics analysis skills:

- Proficiently apply experimental techniques for protein structure determination.

- Develop Understanding of Biomarkers:
- Gain hands-on experience in identifying biomarkers using cutting-edge technologies.
- Analyze complex biomolecular datasets for disease-associated markers.
- Apply Computational Immunology:
- Acquire skills in using bioinformatics tools to analyze immunogenic epitopes.
- Optimize therapeutic protein sequences through computational deimmunization.
- Integrate Structural and Computational Approaches:
- Combine structural proteomics data with computational tools for a comprehensive understanding.
- Collaborate on interdisciplinary projects for biomarker discovery and immunological research.

**STRUCTURAL PROTEOMICS LAB**

Structural Motif Analysis
Protein Structure Classification Schemes
Protein Surface Analysis
Active Site Analysis
Protein Insilico Mutagenesis
Protein Homology Modelling
Protein Model/ Structure Evaluation
Protein Refinement
Computational tools and resources for analysis of functional signatures: ScanProsite and InterPro
EXPASY platform for Proteomic Analysis and Expression Profiling
Flowcytometry Data Visualisation, Immunophenotyping-CD markers Data Analysis
Analysis of the following Structural Motifs/Structures:
Helix-turn-Helix, Zn fingers, Leucine zippers, Beta Barrels, Trans Membrane proteins, Immunoglobulins & Enzymes

**COMPUTATIONAL IMMUNOLOGY LAB**
Introduction to Vaccine design and Reverse vaccinology
Primary protein structure prediction of antigenic protein
Prediction of antigenicity of antigenic proteins
Prediction of allergic nature of antigenic proteins
Prediction of physiochemical properties of antigenic proteins
Prediction of secondary structure of antigenic protein
Prediction of domains and important sites in antigenic protein
Continuous B-cell epitope prediction
Discontinuous B-cell epitope prediction
Prediction of immunogenic regions in antigenic protein
Prediction of glycoprotein antigen epitopes
Cytotoxic T cell epitope prediction
MHC class I and II prediction

T cell epitopes processing prediction
T cell epitopes Immunogenicity prediction
CDR analysis
Generation of antibody repertoires
Automated antigen modelling
Alignment based antigen modelling
Antibody modelling
Antigen-Antibody Docking

## PS 524.2: STATISTICS FOR GENOMICS APPLICATIONS

| | | |
|---|---|---|
| **Total No. of Lectures : 45** | **Total Marks : 100** | **[ L – T – P – S ]** |
| **No. of Lectures / Week : 4** | **Credits : 3** | **[ 3 – 1 - 0 - 2 ]** |

**Learning Objectives:** This course provides an introduction to the use of statistical analysis for Bioinformatics. Provides an understanding of the basic methods and underlying concepts of statistics that are used in Bioinformatics for decision-making and analysis. Among topics explored are descriptive statistics, correlation and regression, sampling, hypothesis testing and design of experiments.

**Learning Outcomes** : After completing this course, students will be able to:

- learn the basic of data classification with appropriate visualization to have a better perspective of the data.
- understand the purposes of measures of central tendency and compare the variability of two or more series.
- apply multiple linear regression analysis, differentiate between simple linear regression analysis and multiple linear regression analysis and predict the model and interpret it.
- recognize the logic behind a hypothesis test and how it relates to the P-value.
- know experimental designs such as Randomized design, completely randomized design and Latin Square designs.

### Unit – I Biostatistics Foundations for Proteomics and Biomarker Research 9 hrs

Basics of Biostatistics: Key concepts: Mean, median, mode, variance, standard deviation. Types of Data: Categorical, continuous, discrete data in proteomics. Statistical Tools for Proteomic and Biomarker Data Analysis. Introduction to Statistical Software: R and Python for data analysis in proteomics. Data Visualization Techniques: Heatmaps, volcano plots, and dimensionality reduction for biomarker and immune expression data. Probability Applications: Basic probability concepts. Probability in pathway enrichment, biomarker discovery, and immune response profiling.

**Unit – II Hypothesis Testing and Confidence Intervals in Proteomic and Gene Expression Studies 9 hrs**

Concepts of Hypothesis Testing: Null and alternative hypotheses, p-values, and significance levels. Errors in hypothesis testing: Type I and Type II. Applications in Biomarker Validation: Case studies on proteomic datasets and immune gene expression. Confidence Intervals: Constructing confidence intervals. Applications in quantitative proteomics, biomarker validation, and immune differential expression studies.

**Unit – III Statistical Modeling and Predictive Analysis 9 hrs**

Regression Analysis: Simple and multiple regression in biomarker prediction. Predictive modeling for proteomic trends and immune gene expression data. Logistic Regression: Logistic regression for binary outcomes in biomarker classification. Model Evaluation: Goodness-of-fit, R-squared, and cross-validation techniques. Applications in biomarker-based diagnostics and immune response prediction. Case Studies: Use of regression models in quantitative proteomics and immune differential expression analysis.

**Unit – IV Multivariate Analysis in Proteomics and Immune Data 9hrs**

Multivariate Analysis Techniques: Principal Component Analysis (PCA), t-SNE, and clustering techniques. Applications in proteomic data dimensionality reduction and immune expression profiling. Classification Techniques: K-means clustering, hierarchical clustering. Classification of proteomic data and immune response profiles. Multivariate Statistical Approaches: M-ANOVA and discriminant analysis. Case studies on multivariate analysis of proteomics and immune gene datasets.

**Unit – V Statistical Applications in Proteomics and Biomarker Discovery 9 hrs**

Time-to-Event Analysis:Survival analysis, Kaplan-Meier curves. Applications in biomarker-based survival prediction and immune response studies. Bayesian Statistics in Proteomics and Immune Analysis: Introduction to Bayesian inference. Bayesian models for pathway analysis, biomarker discovery, and immune response profiling. Machine Learning Integration: Overview of machine learning techniques in proteomics. Predictive modeling for biomarker discovery and immune differential gene expression. Advanced Data Integration: Multi-omics data integration for biomarker discovery.

**Text Books:**

[1]. Principles of Proteomics Richard M. Twyman, 2nd Edition, 2014.

[2]. Practical Statistics for Data Scientists: 50+ Essential Concepts Using R and Python Peter Bruce and Andrew Bruce, 2nd Edition, 2020.

[3]. Bayesian Data Analysis Andrew Gelman, John B. Carlin, Hal S. Stern, David B.Dunson, Aki Vehtari, and Donald B. Rubin, 3rd Edition, 2013.

[4]. Handbook of Statistical Systems Biology , James M. Walker ,4th Edition , 2014

**Reference:**

[1]. Statistical Methods in Bioinformatics: An Introduction Warren J. Ewens and Gregory R. Grant, 2nd Edition, 2005.

[2]. Applied Multivariate Statistical Analysis Richard A. Johnson and Dean W. Wichern,6th Edition, 2007.

[3]. Proteome Bioinformatics Helmut Bergeron and Philip Day, 1st Edition, 2011.

[4]. Data Analysis for Omics Sciences: Methods and Applications Florian Huber andAndreas Hildebrandt, 1st Edition, 2015.

[5]. Modern Statistics for Modern Biology Susan Holmes and Wolfgang Huber, 2nd Edition , 2021

## PS 525.2 INTERMEDIATE Python WITH BIO-COMPUTING

| | | | | |
|---|---|---|---|---|
| **Total No. of Lectures** | **: 45** | **Total Marks :** | **100** | **[ L – T – P – S ]** |
| **No. of Lectures / Week :** | **3** | **Credits :** | **3** | **[ 3 – 1 - 0 - 2 ]** |

**Learning Objectives**: The core objectives of the course are :

- Understand the fundamentals of building web applications using Streamlit.

- Develop essential skills in organizing and displaying content within a Streamlit web application.
- Master the manipulation and analysis of numerical data using NumPy arrays and statistical techniques.
- Acquire proficiency in data manipulation and analysis using Pandas, a powerful data analysis library in Python.
- Gain practical experience in bioinformatics tasks and data mining using Python libraries such as Biopython.

**Learning Outcomes** : Upon successful completion of this course, students will be able to:

- By the end of this unit, learners will be able to set up a Python coding environment for Streamlit,
- learners will be proficient in organizing and arranging content using columns and expanders, and dynamically showing or hiding parts of the application based on importance or user interaction
- learners will be able to create and manipulate NumPy arrays, perform slicing, indexing, and fancy indexing, apply basic descriptive statistics, and conduct linear algebra operations.
- Learners will be able to understand data manipulation and analysis using Pandas, a powerful data analysis library.
- learners will be competent in parsing DNA data files, performing sequence analysis tasks, utilizing Python for bioinformatics tasks, analyzing microarray data, and applying text mining techniques for data mining purposes..

**Unit- I** **9 hrs**

**Getting started with Streamlit.**

**Getting to know streamlit- A brief overview:** The importance of turning an idea into a prototype – the changing scenario in computing methods, the importance of a full python web framework, local versus the cloud, python libraries for web applications (Flask and Django), Examples of streamlit's capabilities, the importance of turning an idea into a prototype.

**Setting up the python coding environment:** Technical requirements, setting up the OS, IDE selection, what is a virtual environment.

**Exploring Streamlit's Features and Functionality:** Technical requirements, installing and launching streamlit, streamlit features and widgets, colored texboxes, images, audio and video, inputting text and numbers.

**Unit- II** **9hrs**

**Building a basic web app for Essential streamlit skills.**

**Streamlit Essentials - Configuring the environment, managing packages, and more:** Technical requirements, configuring our environment, installing and importing packages, app skeleton building, building the menu and adding decorations.

**Organizing and Displaying content with columns, expanders:** Technical requirements, organizing and arranging content in a web app, hiding and showing parts depending on importance.

**Unit – III** **9hrs**

**NumPy Arrays:** NumPy array object, Creating a multidimensional array, Selecting NumPy array elements, NumPy numerical types, One-dimensional slicing and indexing, Manipulating array shapes, Creating array views and copies, Fancy indexing, Indexing with a list of locations, Indexing NumPy arrays with Booleans, Broadcasting NumPy arrays.

**Statistics and Linear Algebra:** NumPy and SciPy modules; Basic descriptive statistics with NumPy; Linear algebra with NumPy; eigenvalues and eigenvectors with NumPy; NumPy random numbers; Creating a NumPy-masked array.

**Unit – IV** **9hrs**

**Pandas Primer:** pandas DataFrames; pandas Series; Querying data in pandas; Statistics with pandas DataFrames; Data aggregation with pandas DataFrames; Concatenating and appending DataFrames; Joining DataFrames; Handling missing values.

**Retrieving, Processing and Storing Data:** Writing CSV files with NumPy and pandas, Comparing the NumPy .npy binary format and pickling pandas DataFrames, Storing data with PyTables, Reading and writing pandas DataFrames to HDF5 stores, Reading and writing to Excel with pandas.

**Unit- V** **9hrs**

**Biopython and Bioinformatics**: Parsing DNA data files, Image manipulation, Sequence analysis – Sequence alignment (pair wise and multiple sequence alignment), Dynamic Programming, Detecting tandem repeats and generating Hidden Markov Models, Multiple Sequence Alignment, gapped alignment. pySam, pyGNA, Multipath, networks, Simulation of EST Clustering.

Data mining – Text mining, Simulating Genetic algorithm; Analysis of Microarray data – Spot finding and Measurement. Python for network analysis.

**Text Books:**

[1]. Febrizo Lomano, “Mastering Python”, 1st Edition, 2018, PACKT Publishers

[2]. Ivan Idrics “ Python For Data Analytics”, 1st Edition, 2017, PACKT Publishers

[3]. Magnus Vilhelm Persson, Luiz Felipe Martins, “Mastering Python Data Analysis”, 1st Edition, 2018, PACKT Publishers.

[4]. Jason Kinser, “Python for Bioinformatics”, 2nd Edition, 2014, Jones & Bartlett Publishers.

[5]. Mark Lutz, “Learning Python”, 5th Edition, 2013, O'Reilly Media.

[6]. Alex Martelli, David Ascher, “Python cookbook”, 2nd Edition, 2012, O'Reilly Publishers

[7]. Web App Development Made Simple with Streamlit: A web developer's guide to effortless web app development, deployment, and scalability, By Rosario Moscato, 2024.

**REFERENCE BOOKS:**

[1]. Mark Lutz, “Programming Python”, 4th Edition, 2012, O'Reilly Media.

[2]. Tim Hall and J-P Stacey, “Python 3 for Absolute Beginners”, 2nd Edition, 2013, Wiley Apress.

[3]. Magnus Hetland, “Beginning Python: From Novice to Professional”, 2nd Ed, 2010, Apress

[4]. Shai Vaingast, “Beginning Python Visualization Crafting Visual Transformation Scripts”, 2nd Edition, 2014, Wiley Apress International Edition

[5]. Wes Mc Kinney, “Python for Data Analysis”, 3rd Edition, 2012, O'Reilly Media.

[6]. White, “Hadoop: The Definitive Guide”, 3rd Edition, 2012, O’Reilly Publishers

[7]. Brandon Rhodes and John Goerzen, “Foundations of Python Network Programming: The Comprehensive Guide to Building Network Applications with Python”, 2013, Wiley Apress

[8]. Eric Matthes, “Python Crash Course, 2nd Edition: A Hands-On, Project-Based Introduction to Programming”, No Starch Press; 2 edition

[9]. Zed A. Shaw, “Learn Python 3 the Hard Way”, Addison Wesley, 2017

[10]. John M Zelle, “Python Programming: An Introduction to Computer Science”, Ingram short title

[11]. Jeff Forcier, Paul Bissex, Wesley J Chun “Python Web Development with Django”

## PS 526.2P PROGRAMMING: INTERMEDIATE PYTHON AND BIOSTATISTICS

## A. Python Lab

1. How to create an empty and a full NumPy array?
2. Create a Numpy array filled with all zeros
3. Create a Numpy array filled with all ones
4. Replace NumPy array elements that doesn't satisfy the given condition
5. Get the maximum value from given matrix
6. Make a Pandas Data Frame with two-dimensional list | Python
7. How to iterate over rows in Pandas Dataframe
8. Program to demonstrate on Create a pandas column using for loop
9. Program to demonstrate on Create a Pandas Series from array
10. Pandas | Basic of Time Series Manipulation
11. Understanding the Streamlit basic functions (Title, Header and Subheader, Markdown).
12. Program to demonstrate on Success, Info, Warning, Error, Exception
13. Using write function, we can also display code in coding format. This is not possible using st.text(").
14. Program to demonstrate on Display Images, Radio Button.
15. A checkbox returns a boolean value. When the box is checked, it returns a True value else returns a False value.
16. Program to demonstrate on Selection Box
17. Program to demonstrate on Date and Time
18. Program to demonstrate on Slider
19. Program to demonstrate on Button
20. Program to demonstrate on Multi-Selectbox

## B. Biostatistics Lab

1. Calculate mean, median, mode, variance, and standard deviation for a proteomic dataset.
2. Classify provided datasets into categorical, continuous, and discrete types.
3. Create a heatmap for biomarker expression levels using R or Python.
4. Analyze differential expression data and visualize it as a volcano plot.
5. Compute probabilities for pathway enrichment in proteomic datasets.
6. Perform t-tests on proteomic data to compare case and control groups.
7. Evaluate p-values from a dataset and interpret significance levels.
8. Simulate Type I and Type II errors using a proteomic dataset.
9. Construct confidence intervals for protein abundance values.
10. Analyze immune gene expression data and test hypotheses for differential expression.
11. Perform simple and multiple regression analysis for biomarker discovery.
12. Build a logistic regression model to classify high vs. low biomarker levels.
13. Evaluate regression models using R-squared and cross-validation.
14. Develop a predictive model for immune gene expression trends.
15. Analyze real-world proteomic data for predictive diagnostics.
16. Perform PCA on a high-dimensional proteomic dataset and interpret results.
17. Visualize immune expression data using t-SNE.
18. Apply K-means and hierarchical clustering on proteomic datasets.
19. Conduct multivariate analysis of variance on immune gene datasets.
20. Use discriminant analysis for biomarker classification.

21. Analyze biomarker survival data and plot Kaplan-Meier curves.
22. Apply Bayesian statistics to infer pathways from proteomic datasets.
23. Implement a machine learning model to classify differentially expressed genes.
24. Integrate proteomic and transcriptomic data for biomarker discovery using R.
25. Train and evaluate a predictive model for immune gene response profiling.

## PS 529.2 RESEARCH METHODOLOGY AND ETHICS

**Learning Objectives:** To have clear understanding of the meaning and purpose of Research in academics, research philosophy and strategies of Research.
To acquaint with the knowledge of methodology involved in a scientific Research
To know writing of a good Research Report.
To understand the ethical issues and practices in research with an awareness of rights and obligations of research participants.
Understand the process of Intellectual property Rights and its different forms and implications
To know how to write research papers and publish research papers.

**Learning Outcome:** Research output with philosophical base and greater relevance to the society
Quality research with scientific methodology
Production of good Research Reports
Original Research following ethical guidelines and practices in conducting the research and publication of papers.
More awareness on Intellectual property Rights and Patents.

Pedagogy: Class room lecture, seminars, assignments, case studies, projects, workshops, dissecting research papers.

### Unit 1: Foundation of Research: (6 Hours)

Research – meaning, characteristics, objectives, motivation in research, need and importance of research. Types of Research; Philosophy and Research Philosophy – Ontology, Epistemology, Axiology, positivism, critical realism, interpretivism, post modernism, pragmatism – meaning, relevance and assumptions. Concept of Theory and Theory Building – deduction, induction and abduction. Research Strategies - meaning and types.

### Unit 2: Research Methodology: (8 Hours)

Research Problem – selecting the problem, sources of problem, statement of a problem; Review of Literature – meaning and need for literature review, sources of literature review, reporting the review of literature, identification of research gap; Research Design – meaning, features of good research design, types of research design, significance and preparation of research design; Research Questions; Objectives of the study; Research Methodology: Analytical vs. Empirical Methods, Surveys, Case Studies, Controlled Experiments, Ethnography and Action Research, Quantitative, Qualitative, and Mixed Methods, Choosing research methods, Validity threats; Introduction to Quantitative Research, Study Designs, Controlled Experiments, Elements and Methods, Data Collection – meaning, sources of data, Data Collection Techniques

### Unit 3: Interpretation and Report Writing; (6 Hours)

Analysis of Data – Analysis and Interpretation of Quantitative Data, Descriptive Statistics; Research Report – meaning, features of a good Research Report, elements of Research Report, format of a Research Report, Appendices and References/ Bibliography – styles.

### Unit 4: Research Ethics: (10 Hours)

Ethics – meaning and definition, Ethics Vs moral philosophy, nature of moral judgments and reactions. Rights and obligations of Research Participants. Scientific conduct – ethics with respect to science and research, intellectual honesty and research integrity. Scientific misconduct – falsification, fabrication and plagiarism. Publication ethics – meaning and importance, conflicts of interest, publication misconduct – meaning, problems that lead to unethical behaviors, types of publication misconduct, identification of publication misconduct, complaints and appeal. Redundant publication – duplicate and overlapping publications, salami slicing. Violation of public ethics, authorship and contributor ship. Predatory publishers and journals – software to identify predatory publications – journal finder/journal suggestions tools by JANE, Elsevier journal finder, Springer journal suggestions etc,. Selective reporting and misinterpretation of data. Best practices/standard setting initiatives and guidelines. Self-plagiarism.

### Unit 5: Intellectual Property Rights (IPR) and Publication of Scholarly Papers (10 Hours)

IPR – Concept of IPR, nature and characteristics of IPR, origin and development of IPR, justification and rationale for protecting IPR, IPR and sustainable development, IPR and human rights, IPR issues in physical and biological sciences, Commerce and IPR issues, IPR issues in Social Sciences. Forms of IPR – copyrights, trademarks, patents, industrial designs, trade secrets, geographical indications – meaning, features and application of different forms of IPRs. Filing and Registration process of IPRs.
Publication – Scholarly/research article – meaning and features of scholarly article. Successful scientific writing – process. Reference/ bibliography writing, Plagiarism and how to avoid it. Dissecting research papers. Data base and Research – Data bases
indexing data base, citation data base, Web of science, Scopus etc, Research Metrics
Impact Factor of Journal as per Journal Citation Report, SNIP, SJR, IPP, Cite Score; Metrics – h-index, g-index, i10 index, Altmetric.

### References:

1. Indian National Science Academy (INSA). (2019). Ethics in Science Education, Research & Governance
2. Barbara H Stanley J Joan E Sieber, Gary B Melton. Research Ethics: A Psychological Approach. University of Nebraska Press
3. David I Bainbridge (2012), Intellectual Property Rights. Long man Publication
4. Jayashree Watal. Intellectual Property Rights in the WTO and Developing Countries. Oxford University Press
5. A K Singh. Tests, Measurements and Research Methods in Behavioral Sciences. Bharathi Bhawan (Publishers & Distributors), New Delhi
6. Leedy P D. Practical Research: Planning & Design. Washington: Mc Millan Publishing Co., INC

7. Singh Y K. Fundamentals of Research Methodology and Statistics. New International (P) Ltd., New Delhi.
8. Wallinman N. Your Research Project: A Step by Step Guide for the first time Researcher. Sage Publications, London
9. Kothari C R. Research methodology: Research & Techniques. New Age International Publishers, New Delhi
10. Selecting Empirical Methods for Software Engineering Research, Easterbrook et al
11. Pfleeger, S.L. Experimental design and analysis in software engineering. Ann Software Eng 1, 219–253 (1995). https://doi.org/10.1007/BF02249052
12. Research Design. Qualitative, Quantitative, and Mixed Methods Approaches. By John W. Creswell, Fourth Edition. SAGE Publication, 2014

## III SEMESTER

### PH 521.3 COMPUTATIONAL APPROACHES IN THERAPEUTICS AND BIOSIMILAR DEVELOPMENT

**Total No. of Lectures : 45** **Total Marks : 100[ L – T – P – S ]**
**No. of Lectures / Week : 4** **Credits : 4 [ 3 – 1 - 0 - 2 ]**

**Course Outcomes:**
Students will gain a solid foundation in computational techniques for therapeutic and biosimilar development, including practical experience with relevant tools and software. They will develop the skills to critically assess the quality, efficacy, and safety of therapeutic agents and biosimilars, as well as proficiency in conducting computational experiments and contributing to innovative research. Additionally, students will understand the regulatory and ethical considerations involved in this field, preparing them to navigate the complex landscape of drug discovery and development.

**Course Objectives:**
- **Fundamentals**: Understand basic concepts of computational drug discovery and development.
- **Modeling and Simulation**: Learn molecular modeling and dynamics simulations.
- **Bioinformatics**: Analyze biological data using bioinformatics tools.
- **Biosimilars**: Design and characterize biosimilars computationally.
- **Regulatory and Ethics**: Grasp regulatory requirements and ethical issues.
- **Capstone Project**: Apply knowledge in a practical computational project.

**1: Drug Discovery and Development Cycle 9 hours**

Overview of the drug discovery and development process, Key stages: target identification, lead discovery, preclinical and clinical development, Importance of computational methods in each stage, Integration of computational tools in the discovery cycle, Regulatory pathways for drug approval, Case studies illustrating the complete drug discovery cycle, Challenges and future trends in drug discovery

**Unit 2: Computational Tools and Practices in Drug Design 9 hours**

Introduction to computational drug discovery, Molecular docking: principles, methods, and applications, Quantitative Structure-Activity Relationship (QSAR) models, Virtual screening techniques and their importance, Computer-aided drug design (CADD) tools and software, Molecular dynamics simulations in drug discovery, Practical sessions: hands-on experience with computational tools, Case studies showcasing successful computational drug design

**Unit 3: Biotherapeutics and Importance of Bioinformatics 9 hours**

Introduction to biotherapeutics: types and applications, Role of bioinformatics in biotherapeutics development, Sequence alignment and structural analysis of biotherapeutics, Protein modeling and simulation techniques for biotherapeutics, Stability and immunogenicity predictions using computational methods, Functional genomics in biotherapeutic target identification, Integration of bioinformatics with experimental data in biotherapeutics, Case studies demonstrating bioinformatics in biotherapeutics development

**Unit 4: Peptide Therapeutics 9 hours**

Overview of peptide therapeutics: importance and applications, Computational design and optimization of peptide drugs, Structure-activity relationship in peptide therapeutics, Predictive Modeling for peptide stability and activity, Tools and software for peptide drug design, Case studies of successful peptide therapeutics, Challenges and future directions in peptide drug development.

**Unit 5: Biosimilar Development and Bioinformatics 9 hours**

Introduction to biosimilars: definitions, regulatory perspectives, and market importance, Computational tools in biosimilar development, Sequence and structural comparison of biologics and biosimilars, In silico prediction of biosimilarity, Protein modeling and simulation for biosimilar development, Assessing biosimilar stability and immunogenicity through computational methods, Case studies of biosimilar development using bioinformatics, Regulatory challenges and ethical considerations in biosimilar development

## PH 522.3 Next-Gen Sequencing and Single-Cell Technologies for Immunomics

**Total No. of Lectures : 45 Total Marks : 100[ L – T – P – S ]**

**No. of Lectures / Week : 4 Credits : 4 [ 3 – 1 - 0 - 2 ]**

**Scope and Outcomes**

This course, "Next Generation Sequencing for Immunomics," provides a thorough exploration of NGS technologies and their application in understanding the immune system. It encompasses essential skills in data quality control, preprocessing, and advanced analysis techniques, including genomic and transcriptomic analyses. The course also covers protein structure visualization and molecular docking simulations, focusing on their relevance in immunomics. Students will engage with practical exercises and real-world projects, applying computational tools to study immune repertoires and interactions. By the end of the course, participants will be equipped to handle NGS data, perform comprehensive analyses, and contribute to immunomics research.

**Objectives**

- Proficiency in NGS data handling and analysis.
- Ability to conduct comprehensive genomic and transcriptomic analyses.
- Competence in visualizing protein structures and performing molecular docking simulations.
- Application of bioinformatics methods to practical immunomics research.
- Enhanced understanding of immune system components and their interactions through computational analysis.

**Unit 1: Foundations of Immunomics and Sequencing Technologies 9 hours**

Overview of immunomics: importance and applications in research and medicine, Basic concepts of the immune system relevant to immunomic, Evolution and principles of sequencing technologies, Comparison of NGS with traditional methods, Overview of NGS platforms (Illumina, PacBio, Oxford Nanopore) and single-cell technologies, Strengths and limitations of each platform, Workflow of a typical NGS and single-cell sequencing experiment

**Unit 2: Experimental Design and Sample Preparation 9 hours**

Defining research questions and objectives for NGS studies in immunomics, Considerations for choosing NGS methods, Design of controls and replicates, Types of samples used in immunomics (e.g., blood, tissue), Best practices for sample collection in immunomics, Sample storage and preservation techniques, Overview of library preparation steps for NGS in immunomics, DNA/RNA extraction methods, Fragmentation, adapter ligation, and amplification

**Unit 3: Immune Repertoire Sequencing and Data Generation 9 hours**

Principles of immune repertoire sequencing, Applications of immune repertoire sequencing in research and clinical diagnostics, Case studies showcasing immune repertoire sequencing, Loading samples for immune repertoire analysis, Running the sequencing reaction and generating raw data, Assessing data quality: base calling accuracy, read length, coverage, Common quality control metrics and tools for immune repertoire sequencing, Troubleshooting low-quality data

**Unit 4: Bioinformatics and Data Analysis for Immune Repertoires 9 hours**

Overview of bioinformatics pipelines for NGS data in immunomics, Introduction to key software and tools for immune repertoire analysis (e.g., FASTQC, MiXCR, IgBLAST), Data formats and conversion tools used in immune repertoire sequencing, Read alignment to reference genomes for immune repertoire sequencing, De novo assembly of immune repertoires, Tools and algorithms for alignment and assembly in immune repertoire analysis, Detecting variants (SNPs, indels) in immune repertoires, Annotation of immune repertoire variants using databases (e.g., VDJdb, IMGT), Functional interpretation of immune repertoire variants 43

**Unit 5: Advanced Applications and Integrative Analysis in Immunomics 9 hours**
Single-cell sequencing technologies in immunomics, Insights into immune cell heterogeneity through single-cell sequencing, Applications of single-cell sequencing in immunology and oncology, Integrating NGS data with other omics data (transcriptomics, proteomics), Future trends and innovations in NGS and immunomics, Ethical considerations and challenges in NGS research, Case studies and practical applications in immune repertoire sequencing, Designing an NGS-based immunomics study, Project presentations and discussions

## PH523.3 P Laboratory based on Therapeutics and Single Cell Technologies

I Computations in Therapeutics:

II Drug Databases

III Molecular Docking

IV Virtual Screening

V Protein Structures

VI Protein Stability

VII Peptide Therapeutics

VIII Peptide Structures

IX Comparing Biologics and Biosimilars

X Predicting Biosimilarity

XI NGS Analysis:

XII Data Quality Control and Preprocessing

XIII Genomic Analysis: Variant Calling

XIV Transcriptomic Analysis: Differential Expression

XV Single-cell RNA-Seq Analysis

XVI Immune Repertoire Sequencing: Data Analysis

XVII Annotation of Immune Repertoire Variants

XVIII Integrative Multi-Omics Analysis

XIX Project: Comprehensive Immunomics Study

## PS 524.3 DATA ANALYTICS FOR BIOLOGY

**Total No. of Lectures: 45** **Total Marks: 100 [ L – T – P – S ]**
**No. of Lectures / Week: 4** **Credits : 3 [ 3- 1 – 0 – 2 ]**

**Course outcome:**
This course provides a comprehensive introduction to data science and machine learning techniques, with a specific focus on their applications in biology. It covers fundamental concepts, methodologies, and tools used in analysing and interpreting biological data. Students will gain hands-on experience with real-world biological datasets, learning to apply data science and machine learning methods to solve complex biological problems.

**Course Objectives:**

- Understand the basic concepts of data science and machine learning, Learn to pre-process and analyse biological data • Gain proficiency in using Pandas and other libraries for data manipulation, cleaning, and analysis.
- • To understand supervised and unsupervised learning and their applications.
- • To apply machine learning techniques to biological datasets for predictive modelling and classification.
- • Develop skills to interpret and visualize complex biological data.
- • Explore case studies in genomics, proteomics, and personalized medicine.
- • Gain hands-on experience through practical exercises and projects, reinforcing theoretical knowledge with real-world applications

**Course outcomes**

- • Understand data science principles,
- • Be proficient in Python programming and basic data manipulation techniques.
- • Apply exploratory data analysis and visualization methods to derive insights from data.
- • Interpret and visualize complex biological data
- • To implement a neural network for an application of your choice using an available tool
- • To use a tool to implement typical clustering algorithms for different types of applications.

**Unit 1: Introduction to Data Science and Machine Learning** 9 hours
Overview of data science and its importance in biology, Basic concepts in machine learning: supervised and unsupervised learning, Introduction to Python, Data exploration and preprocessing: handling missing data, normalization, and scaling, Overview of biological data types: genomic, proteomic, transcriptomic, and clinical data 45

**Unit 2: Supervised Learning Techniques** 9 hours
Regression techniques: linear regression, regularization (Lasso, Ridge), Classification techniques: logistic regression, decision trees, random forests, support vector machines, Evaluation metrics: accuracy, precision, recall, F1 score, ROC-AUC,
Case study: Predicting disease risk based on genomic data,

**Unit 3: Unsupervised Learning and Clustering** 9 hours
Clustering techniques: K-means, hierarchical clustering, DBSCAN, Dimensionality reduction: Principal Component Analysis (PCA), t-SNE, Analyzing and visualizing clusters in biological data.
Case study: Identifying cancer subtypes from gene expression data,

**Unit 4: Advanced Machine Learning Techniques** 9 hours
Ensemble learning: boosting, bagging, and stacking, Introduction to deep learning: neural networks, convolutional neural networks (CNNs), recurrent neural networks (RNNs), Overview of deep learning frameworks (TensorFlow, Keras, PyTorch), Case study: Image analysis in biomedical imaging

**Unit 5: Applications and Ethical Considerations** 9 hours
Applications of machine learning in personalized medicine and drug discovery, Integrating multi-omics data: genomic, transcriptomic, proteomic, and metabolomic data analysis, Ethical considerations in data science and machine learning: data privacy, bias, and fairness, Future trends and challenges in data science and machine learning in biology

References:

1. Data Science from Scratch: First Principles with Python" by Joel Grus

2. DATA SCIENCE: Beginner's Guide to Learn the Realms Of Data Science by William Vance
3. Ethem Alpaydin, “Introduction to Machine Learning”, 1st 2004, MIT press
4. Kevin P. Murphy, “ Machine Learning: A Probabilistic Perspective”, 1st 2012, MIT press
5. Identifying Ethical Considerations for Machine Learning Healthcare Applications Danton S. Char, Michael D. Abràmoff & Chris Feudtner

## PS525.3 PYTHON APPLICATION IN BIOINFORMATICS

**Total No. of Lectures : 45** **Total Marks : 100[ L – T – P – S ]**
**No. of Lectures / Week : 4** **Credits : 3 [ 3 – 1 - 0 - 2 ]**

**Scope of the Course**

This course, "Applied Python for Bioinformatics," provides a comprehensive introduction to using Python in the field of bioinformatics. It covers essential programming concepts, advanced data analysis techniques, and specialized bioinformatics applications. Students will gain hands-on experience with widely-used Python libraries such as Biopython, Pandas, Matplotlib, and SciPy. The course also explores the integration of machine learning and artificial intelligence in analyzing biological data, including sequence analysis, data visualization, and structural bioinformatics. By the end of the course, participants will be equipped with the knowledge and practical skills needed to address real-world challenges in bioinformatics research and applications.

**Course Outcomes & Objectives**

- **Fundamental Understanding:** Develop a solid foundation in Python programming and its application in bioinformatics.
- **Data Analysis Proficiency:** Gain proficiency in using Pandas and other libraries for data manipulation, cleaning, and analysis.
- **Bioinformatics Tools:** Learn to utilize Biopython for sequence analysis, data retrieval, and annotation.
- **Visualization Skills:** Master the use of Matplotlib and Seaborn for creating insightful data visualizations.
- **Advanced Applications:** Understand the methodologies for analyzing NGS data, including quality control, variant calling, and differential expression.
- **Structural Bioinformatics:** Acquire knowledge in protein structure visualization, molecular docking, and simulation techniques.
- **Machine Learning and AI:** Apply machine learning techniques to biological datasets for predictive modeling and classification.
- **Practical Experience:** Gain hands-on experience through practical exercises and projects, reinforcing theoretical knowledge with real-world applications.

**Unit 1: Introduction to Python for Bioinformatics** 9 hours

Python programming basics: data types, variables, control structures, Data structures: lists, tuples, sets, dictionaries. String manipulation and regular expressions, File handling: reading and writing files, working with CSV and TSV files, Parsing and analyzing biological data files (e.g., FASTA, FASTQ), Introduction to Jupyter Notebooks for interactive coding and visualization

**Unit 2: Biopython Fundamentals** 9 hours

Overview of Biopython and its modules, Working with Seq and SeqRecord objects, Sequence alignment and comparison, Translating DNA to protein sequences, Accessing biological 47 databases (e.g., NCBI, UniProt), Parsing sequence data and annotations, Fetching and analyzing gene and protein sequences, Annotating sequences with functional information

**Unit 3: Data Analysis and Visualization** 9 hours

Data manipulation with Pandas: DataFrames, Series, data cleaning, Basic plotting with Matplotlib and advanced visualization with Seaborn, Descriptive statistics and hypothesis testing, Introduction to SciPy for statistical analysis, Analyzing and visualizing biological datasets (e.g., gene expression data, sequence motifs), Identifying patterns and correlations in biological data

**Unit 4: Advanced Python Applications in Bioinformatics** 9 hours

*Scientific Computing (NumPy and SciPy)***:** Array manipulation, mathematical operations, scientific computing functions, and linear algebra. *Pandas:* Advanced data manipulation, aggregation, merging, and reshaping techniques. *Data Visualization: Matplotlib and Seaborn:* Creating and customizing advanced plots, interpreting graphical data. *Machine Learning: Scikit-learn:* Overview of machine learning models, implementing classification, clustering, and regression. *Biological Computations: Biopython:* Advanced sequence analysis, accessing and utilizing biological databases. *Structural Bioinformatics: PyMOL and RDKit:* Molecular visualization and analysis, computational chemistry, and drug discovery.

**Unit 5: Machine Learning and AI in Bioinformatics** 9 hours

Basics of machine learning and key concepts, Overview of supervised and unsupervised learning, Feature extraction and selection, Classification, clustering, and regression techniques, Introduction to deep learning frameworks (e.g., TensorFlow, PyTorch), Neural networks for biological sequence analysis, Building predictive models for biological data (e.g., disease prediction, protein function), Implementing neural networks for sequence and structural data analysis

## PS 526.3P: Lab based on Python and Machine Learning

### I. Python Lab:

1. Parsing and Analyzing DNA Sequences:
Read a FASTA file using Biopython.
Calculate nucleotide frequencies and GC content.
2. Translating DNA to Protein
Use Biopython to translate DNA to protein.
Identify the protein sequence and check for known functions.
3. Sequence Alignment
Use Biopython to align two sequences.
Compare sequences and calculate similarity.
4. Data Analysis with Pandas 48

Load and clean a dataset (e.g., gene expression data).
Summarize and visualize the data.
5. Visualizing Data
Plot data using Matplotlib and Seaborn.
Create a histogram, scatter plot, and heatmap.

**II. Machine Learning:**

1. Data Preprocessing and Cleaning
Handling missing values, normalization, outlier detection
2. Exploratory Data Analysis (EDA)
Visualization (histograms, box plots, pair plots)
3. Linear Regression for Biomarker Identification
Linear regression model, R-squared, residual plots
4. Classification with Decision Trees and Random Forests
Decision trees, random forests, confusion matrix
5. Clustering Analysis of Single-Cell RNA-seq Data
K-means, hierarchical clustering, silhouette score
6. Dimensionality Reduction with PCA and t-SNE
PCA, t-SNE, visualization
7. Building a Neural Network for Protein Function Prediction
Neural network (Keras), accuracy, precision
8. Image Classification in Biomedical Imaging
Convolutional neural networks (CNNs)
9. Integration of Multi-Omics Data for Disease Prediction
Multi-omics integration, logistic regression, random forests
10. Capstone Project: Comprehensive Analysis of a Biological Dataset
    End-to-end project, data preprocessing, model building, evaluation