PONDICHERRY UNIVERSITY
PUDUCHERRY
Ph.D. BIOINFORMATICS

2019-2020 onwards

Eligibility: Master's degree in Bioinformatics / Life Sciences / Computer Science / Physics /Chemistry /Applied Mathematics/ Statistics/ IT & Engineering or any other relevant areas with a minimum of 55% of marks**.**

Pre-PhD Curriculum:

The Pre-PhD Curriculum has three components/papers, each bearing **Four Credits** and is compulsory for all students registering for Ph.D., in Bioinformatics.

Paper – I: **Research Methodology** – The syllabus for Paper will be recommended by the members of the school board of studies and will cover fundamentals of Research Methodology, good laboratory practices, ethical issues, patents and patentability.

Paper –II: **Research Methodology in Bioinformatics**: The syllabus will be recommended by the departmental members of the Board of Studies and will cover the basic and advanced techniques both in wet lab and computational lab which are essential for conducting research in Bioinformatics.

Paper-III: **Guide Paper** – A distinct theory paper specific for each student based on their respective PhD research problem that will address the necessary theoretical background and methodologies as essential for comprehensive assessment of knowledge of the research scholar in their respective research domain.

Course Outcome: After finishing the course the student will be able to design and conduct research in Bioinformatics/Computational Biology using wet lab/dry lab tools. He/ She will have a clear understanding of various research strategies and ethical guidelines in their research.

PAPER-I: RESEARCH METHODOLOGY

COURSE OBJECTIVES: 4 Credits

To understand the fundamentals of designing a scientific problem and experiments to answer the same. To demonstrate the ability to analyze, present and interpret scientific data to draw accurate and appropriate conclusions and identify implications and future directions of the research. To know good laboratory practices, ethical issues in research, intellectual property rights, patents and patentability.

UNIT I: Foundations of Research: 11h

Identification and formulation of scientific problem: Research Question – Defining aims and objectives – hypothesis generation - Concept of theory, empiricism, deductive and inductive theory – validation and interpretation of data - Characteristics of scientific method – Basic and applied research problems -Research Design: Concept and Importance in Research – Features of a good research design – Exploratory and Descriptive Research Designs – concept, types and uses – Reading and interpretation of research papers, Critical analyses of research problems, Patent search - Use of Encyclopedias, Research Guides, Handbooks and Manuals, Academic Databases etc.-translational approach in research.

UNIT II: Experimental Design: 10h

Concept of Independent & Dependent variables– concept of positive and negative controls – Quality Controls-Single and Double Blind Studies-Measurement Issues – Hypothesis – Qualities of a good Hypothesis –Null Hypothesis & Alternative Hypothesis. Hypothesis Testing – Logic & its importance - Qualitative and Quantitative research – Concept of measurement, causality, generalization, replication - Problems in measurement in research – Validity, Reliability and Reproducibility.

UNIT III: Concepts of Statistical Methods: 12h

Population, Sample, Sampling Frame, Sampling Error, Sample Size, Non Response - Characteristics of a good sample - Probability Sample – Simple Random Sample, Systematic Sample, Stratified Random Sample & Multi-stage sampling. Determining size of the sample – Practical considerations in sampling and sample size- Data Preparation – Univariate analysis (frequency tables, bar charts, pie charts, percentages), Bivariate analysis – Cross tabulations and Chi-square test including testing hypothesis of association - p-value, ANOVA (analysis of variance), cluster analysis – SPSS.

UNIT IV: Interpretation of Data and Presentation skills: 10h

Data mining and analysis, preparation and interpretation - Layout of a Research Paper and other communications, Journals in Life Sciences, Open Access Journals, Predatory Journals, Impact factor of Journals, When and where to publish ? - Oral and written presentations- Document preparation, Excel - Power Point Presentation, Scientific editing tools - Hand-outs and Brochures, Paper, Abstract and grant writing skills, Thesis writing - Reference Management and Plagiarism detection softwares.

UNIT V: Good Laboratory Practices, Ethical Issues & IPR: 12h

Good Laboratory Practices – Data management in laboratory – Regulations for recombinant DNA and toxic compounds research - safety and bio- and radio- hazards, disposal of biological and chemical waste, Accuracy of liquid transfer, Preparation of reagents, chemicals and buffers, Handling of sophisticated instruments- Animal

handling and ethics, Maintenance of animals, Various routes of injections and sample collection, CPCSEA guidelines; Institutional ethics and safety committees, Ethical consideration in research on human beings, Regulation of clinical trials and transfer of biological samples - Copyright, Royalty, Intellectual property rights and Patent laws, Reproduction of published material, Ethical issues related to publishing, Plagiarism and Self-Plagiarism, Citation and acknowledgements, Reproducibility and accountability, Conflict of Interest - IPR-related issues, trademarks, copy rights, patents, geographical indicators.

References Books:

1. Gall, M.D., Gall, J.P. Borg, W.R. (2006) Educational research: An introduction, Pearson, London.

2. Willis, J. (2004) Data Analysis and Presentation Skills: An Introduction for the Life and Medical Sciences, Wiley, New Jersey.

3. Green. R.H. (1979) Sampling Design and Statistical Methods for Environmental Biologists. John Wiley & Sons, New Jersey.

4. Ruxton, G.D. and Colegrave, N. (2017) Experimental design for the life sciences, Oxford University Press, Oxford.

5. Snedecor, G.W. and Cochran, W.G. (1989) Statistical methods, Iowa State Press, Iowa.

6. Mitchell, K. and Glover, T. (2001) Introduction to Biostatistics, McGraw-Hill Publishing Co., New York. 7. Padma, N. (2017) An Introduction to Ethical, Safety and Intellectual Property Rights Issues in Biotechnology, Academic Press, Cambridge.

8. Shomini, P. and Deepa, G. (2013) IPR, Biosafety and Bioethics, Pearson, London.

9. Bouchoux, D.E. (2013) Intellectual Property Rights: The Law of Trademarks, Copyrights, Patents and Trade Secrets, Cenage Learning, Boston, Massachusetts.

Suggested reading:

1. Ministry of Environment, Forest and Climate Change, Govt. of India (2018) Compendium of CPCSEA.

2. ICMR. (2008) Guidelines for Good Clinical Laboratory Practices (GCLP).

3. ICMR (2017) National Ethical Guidelines For Biomedical and Health Research Involving Human Participants. (https://www.icmr.nic.in/guidelines).

COURSE OUTCOME: Demonstrate intellectual independence, knowledge about good research practice and ability to make scientific judgments based on such principles. Have better understanding of the nature of science and values at stake in the practice of science.

PAPER II: RESEARCH METHODOLOGY IN BIOINFORMATICS

4 Credits

Unit-I 8 lectures

Bioinformatics databases:- **Genome Databases**: ICTVdb, GOLD, MBGD; **Genome Browsers**:- Ensembl, VEGA genome browser, NCBI-NCBI map viewer, KEGG, MIPS, UCSC Genome Browser; **Sequence Databases**:- GenBank, EMBL, DDBJ; Swiss-Prot, TrEMBL, UniProt, **Sequence motifs Databases**:- Prosite, ProDom, Pfam, InterPro, **Structure and derived databases** – PDB, NDB, MMDB; SCOP, CATH, FSSP, CSA; KEGG ENZYME database; STRING; Pubchem. **Database search engines** – Entrez, BLAST, FASTA, ScanProsite and eMOTIF, VAST and DALI, EMBOSS. **Literature Databases**:- Open access and open sources, PubMed, PLoS, Biomed Central, etc.; **Bioinformatic Resources**:- NCBI, EBI, ExPASy, RCSB.

Unit -II 7 lectures

Sequence Analyses – Sequence similarity, identity and homology, homologues, orthologues, paralogues and xenologues; PAM and BLOSUM Scoring matrices, Needleman and Wunsch, Smith and Waterman algorithms for pairwise alignments. Multiple sequence alignments (MSA) –Algorithms of BLAST, FASTA, CLUSTALW, PSI-BLAST and PHI-BLAST analyses and interpretations. **Molecular Phylogenetics** :– Introduction, Gene Phylogeny versus Species Phylogeny, Phylogenetic Tree Construction Methods and Programs: Distance-Based Methods, Character-Based Methods, Phylogenetic Tree Evaluation, Phylogenetic Programs_MEGA abnd PHYLIP.

Unit -III 7 lectures

Analytical Techniques: **Chromatography**:- TLC, GC, HPLC, Gel filtration, ion-exchange and affinity chromatography; **Electrophoresis**:- SDS-PAGE, Southern, Northern, Western blotting techniques, DNA sequencing and Electro focusing, **Centrifugation**:- Ultra centrifugation-velocity & density gradient centrifugation in isolation of cells. **Microscopy**:- fixation, staining; Principle and application of light, phase contrast, fluorescence, scanning and transmission microscopy. **Spectroscopy**:- Mass spectroscopy, MALDI, infrared spectroscopy, 1- and 2-dimensional NMR and ESR spectroscopy.

Unit -IV 7 lectures

Macromolecular Simulations and Drug Design: **Biomolecular structure**: Nucleic Acids (DNA and RNA) and Proteins: –Primary, Secondary, super secondary, Teritary and Quaternary structure, biological significance; Structure validation -Ramachandran Map, Chou Fasman – structure prediction; Ramachandran Map, **Molecular Dynamics**: Force Fields: AMBER, CHARMM, Merck Molecular Force Field, etc., PES and characterization, Molecular Dynamics, **Drug Design**: Structure based drug design Docking, De Novo Drug Deisgn.

Unit -V 7 lectures

Systems Biology: Introduction to Systems Biology. Applications in health and diseases. Microarrays and its applications in systems biology. Connectivity maps (CMap) and Library of Integrated Network-based

Cellular Signatures (LINCS) -definition and its uses. **Biological Networks**: Degree distribution, Clustering coefficient, Random networks, Scale-free networks, small-world effect.

Recommended Books:

1. Fundamentals of Molecular Spectroscopy by C.N.Banwell and Colin. 2000
2. Molecular Modeling principles and Applications (2nd Ed.) by Andrew Leach., Prentice Hall, USA. 2001.
3. Introduction to Systems Biology: Design Principles of Biological Circuits by Uri Alon, Chapman & Hall/CRC, 2007.
4. Principles of Protein Structure by G.E.Schulz., Springer 2009.
5. Principles and Techniques of Practical Biochemistry (7th Ed) by Keith Wilson and John Walker, Cambridge University Press. 2010.

Reference Books:

1. Pevzner, P.A. (2004) "Computational Molecular Biology"; Prentice Hall of India Ltd, New Delhi
2. Pevsner, J. (2003) "Bioinformatics and Functional Genomics"; John Wiley and Sons, New Jersey, USA.
3. Lesk, A.M. (2002) "Introduction to Bioinformatics", First edition, Oxford University Press, UK.
4. Sensen, C.W. (2002) "Essentials of Genomics and Bioinformatics"; Wiley-VCH Publishers, USA
5. Mount, D. (2004) "Bioinformatics: Sequence and Genome Analysis"; Cold Spring Harbor Laboratory Press, New York.