

Ph.D. BIOINFORMATICS

REGULATIONS AND SYLLABI (Effective from 2009-2010)



**Centre for Bioinformatics
SCHOOL OF LIFE SCIENCES
PONDICHERRY UNIVERSITY
PUDUCHERRY**

**PONDICHERRY UNIVERSITY
PUDUCHERRY**

Ph.D. BIOINFORMATICS

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

**Details of the papers and Scheme of Examinations effective from the academic year
2009-2010 onwards**

Course Code	Name of the Papers	Duration of Exam	Maximum Marks
BINF 601	Research Methodology in Bioinformatics	3 hrs.	100
BINF 602	Advanced Bioinformatics	3 hrs.	100
BINF 603	Guide Oriented Paper	3 hrs.	100

BINF – 601: RESEARCH METHODOLOGY IN BIOINFORMATICS

UNIT-I

8 hours

Introduction; Definition; Objectives of research; Types of research; Research approaches; Significance of research; Research methods; Research and scientific method; Importance of knowing how research is done; Research process; Layout of the Research project; Criteria of good research; Problem encountered by researchers in India; Need for research design: Features of good design; Important concepts relating to design; Different research designs; Basic principles of Experimental design.

UNIT-II

6 hours

Interpretation and Report writing Meaning of Interpretations; Techniques of Interpretation, Precautions in Interpretations, Significance of Report writing; Different steps in Report writing; Types of reports and Oral presentation.

UNIT-III

7 hours

Hypothesis and Statistical evaluation Basic concepts of Hypothesis; Procedure for Hypothesis testing; Probability: Markov models and Hidden Markov models; Probability distribution. Binomial, Poisson, Normal distribution, and Multiple testing methods. ANOVA; Test of significance: T-test, F-test.

UNIT-IV

7 hours

Elements of C Programming Features of C, Variables, Constants, Keywords, Data types, Operators, Statements, Loops-Simple programs using Loops, Arrays-Integer arrays, Character arrays, Simple programs using arrays. Introduction to Functions - Simple programs using functions, Introduction to Pointers, Structures - String manipulations using Pointers and Arrays. Files: Defining & Opening a file, Closing a file, Input/Output operations on files.

UNIT – V

8 hours

Introduction to Bioinformatics; Biological problems needs computational methods; sequencing file formats – NBRF/PIR, FASTA, GDE; Files for multiple sequence format, ALN format; Files for structural data – PDB format and NMR files; Annotated sequence databases – primary databases (GenBank-NCBI, the nucleotide sequence database-EMBL, DDBJ; Subsidiary data storage (ESTs, dbESTs, GSSs), organisms specific databases (EcoGene, SGD, MatDB, TAIR, FlyBase, OMIM, etc.); Protein sequence and structure databases (PDB, SWISS-PROT and TrEMBL); List of Gateways (NCBI, GOLD, MIPS, TIGR, UniGene)

Recommended Books:

1. Mathews. "Successful scientific writing: A step-by-step guide for Biomedical Scientists", Second edition, Cambridge University Press, 2001.
2. Warren, J., Gregory, E. and Grant, R. (2004) "Statistical methods in Bioinformatics"; First Edition Springer-verlag, Berlin
3. Habert Schildt (2007) "The Complete reference – C", Fourth Edition, The McGraw – Hill Companies, New York.

BINF – 602: ADVANCED BIOINFORMATICS

UNIT I

7 hours

Data : Access, Retrieval and Submission Data retrieval tools – Entrez, DBGET and SRS; Submission of new and revised data. Sequence Similarity Searches – Sequence alignment – global, local and free-space; its significance, Alignment scores and gap penalties; Measurement of sequence similarity; Similarity and homology.

UNIT II

7 hours

Sequence Analysis – Methods of sequence alignment: Scoring matrices – Percentage accepted Mutation (PAM). Block Substitution Matrices (BLOSUM). Dynamic programming algorithms; Needleman-Wunch and Smith Waterman; Pairwise Sequence Alignment – Programmes (Dot matrix, Dot plot, Dynamic programming) ; Database searching (BLAST and FASTA). Multiple Sequence alignment (MSA) – significance; softwares (PIMA, Clustal, Pileup, ClustalW, Meme, MACAW);

UNIT III

7 hours

Comparative Genome Analysis – Relevance; orthologs and paralogs; Comparative genomics of prokaryotes and Eukaryotes, Vertical and horizontal gene transfer. Applications of comparative genomics in reconstruction of metabolic pathways.

UNIT IV

7 hours

Phylogenetics: Phylogenetic analysis, Phylogenetic representations – graphs, trees and cladograms; Steps in phylogenetic analysis; Methods of phylogenetic analysis – similarity and distance tables, distance matrix method; Method of calculation of distance matrix (UPGMA, WPGMA); The Neighbour Joining Method; The Fitch/Margoliash method; Character-based Methods – maximum parsimony, maximum likelihood: Limitations of phylogenetic algorithms; Phylogenetic softwares – PAUP, PHYLIP, MacClade.

UNIT-V

8 hours

Systems Biology - Objectives of Systems Biology, Strategies relating to In silico Modeling of biological processes, Metabolic Networks, Signal Transduction Pathways, Gene Expression Patterns. E-cell and V-cell Simulations and Applications.

RECOMMENDED TEXTS:

1. Mount, D. (2004) “Bioinformatics: Sequence and Genome Analysis”; Cold Spring Harbor Laboratory Press, New York.
2. Baxevanis, A.D. and Francis Ouellette, B.F. (2004) “Bioinformatics – a Practical guide to the analysis of Genes and Proteins”; Third Edition, John Wiley & Sons, UK.

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

