THIRUVALLUVAR UNIVERSITY
MASTER OF SCIENCE

M.Sc. BIOINFORMATICS

(With effect from 2020 – 2021)

The Course of Study and the Scheme of Examination

Sl. No.	*Study Components*		*ins. hrs / week*	*Credit*	*Title of the Paper*	*Maximum Marks*		
	Course Title					*CIA*	*Uni. Exam*	*Total*
SEMESTER I								
1	Core	**Paper-1**	5	4	Basic Bioinformatics	25	75	100
2	Core	**Paper-2**	5	4	Cell and Molecular Biology	25	75	100
3	Core	**Paper-3**	5	4	Programming in C & C++	25	75	100
4	Practical	**Paper-1**	3	0	Biochemistry, Biophysics and Molecular Biology (LAB)	25	75	100
	Practical	**Paper-2**	3	0	Sequence Analysis (LAB)			
	Practical	**Paper-3**	3	0	Programming in C & C++ (LAB)			
Internal Elective for same major students (Choose any one)								
5	**Core Elective**	**Paper-1**	3	3	A. Mathematics & Statistical methods in Bioinformatics B. Biodiversity Informatics C. Genetics & Evolution	25	75	100
External Elective for other major students (Inter/multi disciplinary papers)								
6	**@ Open Elective**	**Paper-1**	3	3	Biological Databases	25	75	100
			30	**18**		**150**	**450**	**600**
SEMESTER II						*CIA*	*Uni. Exam*	*Total*
7	Core	**Paper-4**	5	4	Genomics & Proteomics	25	75	100
8	Core	**Paper-5**	4	4	Relational DBMS & MySQL	25	75	100
9	Core	**Paper-6**	4	4	Structural Biology	25	75	100
10	Practical	**Paper-1**	3	3	Biochemistry, Biophysics and Molecular Biology (LAB)	25	75	100
11	Practical	**Paper-2**	3	3	Sequence Analysis (LAB)	25	75	100
12	Practical	**Paper-3**	3	3	Programming in C & C++ (LAB)	25	75	100
Internal Elective for same major students (Choose any one)								
13	**@ Core Elective**	**Paper-2**	3	3	A. Biophysics & Biochemistry B. Biological Algorithms C. Cheminformatics	25	75	100
External Elective for other major students (Inter/multi disciplinary papers)								
14	**@ Open Elective**	**Paper-2**	3	3	Biological Sequence Analysis	25	75	100
15	**@ *Field Study**		-	2		100	-	100
16	**@ Compulsory Paper**		2	2	**Human Rights**	25	75	100
			30	**22**		**325**	**675**	**1000**

*** Field Study**

There will be field study which is compulsory in the first semester of all PG courses with 2 credits. This field study should be related to the subject concerned with social impact. Field and Topic should be registered by the students in the first semester of their study along with the name of a mentor before the end of the month of August. The report with problem identification and proposed solution should be written in not less than 25 pages in a standard format and it should be submitted at the end of second semester. The period for undergoing the field study is 30 hours beyond the instructional hours of the respective programme. Students shall consult their mentors within campus and experts outside the campus for selecting the field and topic of the field study. The following members may be nominated for confirming the topic and evaluating the field study report.

(i). Head of the respective department
(ii). Mentor
(iii). One faculty from other department

THIRUVALLUVAR UNIVERSITY
M.Sc. BIOINFORMATICS

SYLLABUS
UNDER CBCS
(with effect from 2020-2021)

SEMESTER I
PAPER I

BASIC BIOINFORMATICS

Objective

To introduce classic bioinformatics theory to students by focusing on how computer techniques can be used for the storage, analysis, prediction and simulation of biological sequences (DNA, RNA and Proteins).

UNIT-I

Bioinformatics - Definition - Biological & Specialized Databases - Nucleic acid sequence databases: GenBank, EMBL, DDBJ - Protein sequence databases: SWISS-PROT, TrEMBL, PIR_PSD - Genome Databases at NCBI, EBI, TIGR, SANGER - Virtual Library.

UNIT-II

Bioinformatics servers - NCBI - EBI - GENOMENET - Bibliographic resources and literature databases - PUBMED, MEDLINE, AGRICOLA - Database Searching techniques - ENTREZ - Data Mining - techniques & tools - Data Warehousing - Top Down & Bottom up approaches.

UNIT-III

Sequence patterns & representation - consensus, regular expression, contigs, motifs and blocks - Sequence Analysis - FASTA - BLAST - Scoring matrices - PAM and BLOSUM - Pairwise alignments - Multiple sequence alignments - CLUSTALW and Pileup - dendrograms and its interpretation.

UNIT-IV

Phylogenetic analysis - taxonomy and phylogeny - molecular evolution - Data used in Taxonomy and Phylogeny - Phylogenic trees - Definition and description - types of trees - tree construction - tree analysis - homologous - orthologous - paralogous - Phylip and phylogenetic analysis.

UNIT-V

Application of Bioinformatics - Drug designing - Drug discovery cycle - Role of Bioinformatics in drug design - Target identification - lead discovery - Structure-based drug design - Modeling of target- small molecule interactions.

Text Books

1. Attwood, T.K. and Parrysmith, D.J. 2001. Introduction to Bioinformatics. Pearson Education (Singapore) Pvt. Ltd., New Delhi.
2. Mani, K. and Vijayaraj, N. 2004. Bioinformatics - A practical approach. Aparna Publications, New Delhi.
3. Harshawardhan Bal - Bioinformatics - Primciples and Applications, 1st Edition 2005, TMH, New Delhi.

References

1. Bryan Bergersen, M.D. 2003. Bioinformatics computing. Pearson Education (Singapore) Pvt. Ltd., New Delhi.
2. Rastogi, S.C., Menderatta, M. and Rastogi, P. 2004. Bioinformatics - concepts, skills and applications. CBS Publishers & Distributors, New Delhi.
3. Westhead, D. R., Parish, J. H. and Twyman, R.M. 2003. Bioinformatics. Viva Books Pvt. Ltd., New Delhi.
4. Sahai, S., 1999. Genomics and Proteomics: Functional and computational aspects. Viva Books Pvt. Ltd., New Delhi.
5. Mount, David W. 2001. Bioinformatics sequence and genome analysis. Cold Spring Harbor Laboratory Press, New Delhi.
6. Pennigton, S.R., and Dunn, M.J. 2002. Proteomics. Viva Books Pvt. Ltd., New Delhi.
7. Baxevanis, En Andreas D. and Francis Ouellette, B.F. 2003. Bioinformatics: A practical guide to the analysis of genes and proteins. John Wiley & Sons, New Delhi.

PAPER II

CELL AND MOLECULAR BIOLOGY

Objectives

To know the terminology and literature of cytogenetics; chromosome structure and function; cytogenetic techniques that can be employed in genetics and improvement of life systems.

UNIT-I

Prokaryotic and eukaryotic cells : Structure and function of extracellular matrix or ECM (cell wall) and membranes - Structure and function of cell organelles (chloroplasts, mitochondria, ER, ribosomes, endosomes, lysosomes, peroxisomes, hydrogenosome). - Nucleus, nucleolus, nuclear pore complex. Chromatin and nucleosome - Cell signalling and cell receptors - Signal transduction.

UNIT-II

Mitosis and meiosis; molecular basis of cell cycle - Phases of Cell Cycle, functional importance of each phase - Numerical and structural variations in chromosomes and their significance. Study of polytene, lampbrush and B-chromosomes - structure, behaviour and significance. Apoptosis- Role of different genes, cell organelles during apoptosis - genetic control of apoptosis.

UNIT-III

Molecular basis of life - DNA as the Genetic Material - Definitions and Chemistry of the Gene - Gene as the unit of mutation and recombination. Genome organizations and mechanism of replication in Prokaryotic and Eukaryotic cells, structure and function of DNA polymerases. Gene as the unit of expression Regulation of gene expression in Bacteria, yeast, mitochondria & chloroplast.

UNIT-IV

Concept of gene Central dogma, updated central dogma, molecular structure of nucleic acids – structure & forms of DNA & RNA. Transcription - components of transcription machinery, RNA polymerases, processing of RNA. Transcription in prokaryotes & eukaryotes, genetic code. Translation – mechanism, post-translational modification.

UNIT-V

Gene regulation in prokaryotes – Operon concept, Lactose, Histidine and Tryptophan operon, Gene regulation in eukaryotes – Transcriptional level, translational level control.

Text Books

1. De Robertis, E.D.P. and De Robertis, E.M.F. 1995. Cell and Molecular Biology. 8th end., B.I. Waverly Pvt. Ltd., New Delhi.
2. Kleinsmith, L.J. & Kish, V.M. 1995. Principles of Cell and Molecular Biology. 2nd edn., McLaughlin, S., Trost, K., Mac Elree, E. (eds)., Harper Collins Publishers, Newyork.

3. Karp, G. (2005) “Cell and Molecular Biology: Concepts and Experiments”; Fourth Edition, Wiley Publishing Co. USA

4. 4. Krieger, M. (2003) “Molecular Cell Biology”; Fifth Edition, W.H. Freeman and Co., New York.

Reference Books

1. Alberts. B., Bray, D., Lewis, J., Raff, M., Roberts, K. and Watson, J. D. 1989. Molecular biology of the Cell (2^{nd} edition). Garland Pub. Inc., New York.
2. Hartl, D. L. and Jones, E. W. 2001. Genetics: Principle and Analysis (4^{th} edition). Jones & Bartlett Pub., USA.
3. Snustad. D. P.and Simmons M. J. 2000. Principles of Genetics (2^{nd} edition). John Wiley and Sons, Inc., USA.

PAPER III

PROGRAMMING IN C & C++

Course Objectives:

- To make the students understand the basic aspects of programming.
- Develop an in-depth understanding of functional, logic, and object-oriented programming paradigms.
- Explain and be able to use fundamental programming constructs such as sequencing, decisions and iteration.
- To enhance problem solving and programming skills in C & C++

Syllabus

Unit-I: Introduction to Programming Languages

Introduction–Programming languages – Problem solving Technique: Algorithm, Flowchart, Compiling, Testing and Debugging, Documentation – Data structures – Array, Stack, Queue, Linked List concepts.

Unit-II: Programming in C

C language Introduction – Tokens – Keywords, Identifier, Variables, Constants, Operators – Expression – Data types –Operator precedence – Statement: Input statement, Output statement, Conditional and Unconditional Control Statement – Looping Statement: while, do–while, for – nested loop – Arrays.

Unit-III: Procedural Concepts in C

C – Procedural Concepts: Structured Programming – Built-in library function – User defined functions – Pointer introduction – Passing pointer in a function – Structure – Union – File handle: Read and Write character from a file.

Unit-IV: Object Oriented Programming and C++

Basic concepts of OOPS– Data hiding–Encapsulation–Inheritance, Polymorphism – Introduction to C++, C vs C++ – data types, variables, constants, operators and statements in C++ – Conditional and looping statements.

Unit-V: Programming and C++

C++ classes - Classes & Objects – Functions in C++ – function prototype-definition–Different forms of Constructor – Destructor – Copy constructor – Inheritance –Single, Multiple and Multi level inheritance – Function & operator overloading –inline functions – Friend and virtual functions – Overloaded functions.

Reference Books:

- B.W.Kernighan and D.M. Ritchie, “The C Programming Language”, 2nd Edition. Prentice Hall of India.
- Byron Gottfried, - “Programming with C” (Schaum's Outline Series) - Tata McGraw Hill Publishing Company – 1998
- E. Balagurusamy - “Programming in C++ ” - Tata McGraw Hill Edition
- Robert Laffore -“Object oriented programming with C++” -Waite series.

SEMESTER I

CORE PRACTICAL I

BIOCHEMISTRY, BIOPHYSICS AND MOLECULAR BIOLOGY

BIOCHEMISTRY

1. Estimation of reducing sugar.
2. Estimation of lipids.
3. Separation of amino acids and lipids using TLC and Paper chromatography.
4. Extraction of secondary metabolites from medicinal plants - Cold percolation method
5. Extraction of secondary metabolites from medicinal plants - Soxhlet method
6. Preliminary Phytochemical analysis of plant extracts
7. Antibiotic bioassay-inhibitory activity (Disc Diffusion)
8. Tissue culture - cell suspension culture
9. Separation of a mixture of proteins (2 or 3) using column chromatography.
10. Estimation of proteins using Bradford and Lowry's methods.
11. Blood analysis, estimation of RBC count, WBC count

BIOPHYSICS

12. Microscopy: Bright field, Phase contrast & Fluorescence microscopy
13. To verify the Lambert Beer's law.
14. Protein crystallization using hanging drop and sitting drop methods.
15. Casting the Gel for SDS-PAGE.
16. Separation of protein and molecular weight determination using SDS-PAGE.
17 . Staining the gel with CBB.

MOLECULAR BIOLOGY

18. Sterilization techniques and Media preparation
19. Preparation and Maintenance of Microbial Culture
20. Plasmid and Chromosomal DNA Preparation from E. coli
21. Pure microbial culture techniques
22. pH measurements and preparation of buffers
23. Spectrophotometric Analysis of DNA
24. Agarose Gel Electrophoretic Analysis of DNA
25. Restriction digestion of bacterial genomic DNA and plasmid DNA
26. Ligation of DNA fragment with plasmid DNA
27. Microbial genomic DNA isolation

20. Microbial plasmid isolation

21. Plant genomic DNA isolation.

22. Agarose Gel electrophoresis and gel documentation.

23. DNA amplification using Thermocycler.

24.Blotting Techniques - Southern, Northern & Western.

25. Hybridization - Autoradiography - Demonstration

SEMESTER I

CORE PRACTICAL II

SEQEUNCE ANALYSIS

1. Biological databases (sequence, structure and specialized databases)
2. Data retrieval using ENTREZ
3. Database file formats
4. Gene structure and function prediction (Genscan , GeneMark)
5. Sequence similarity searching (NCBI BLAST)
6. Protein sequence analysis (ExPASy proteomics tools)
7. Multiple sequence alignment (Clustal)
8. Molecular phylogeny (PHYLIP)
9. Analysis of protein and nucleic acids sequences,
10. Sequence analysis using EMBOSS or GCG Wisconsin Package
11. Sequence comparison
12. Structure analysis
13. Pattern recognition
14. Proteome analysis using tools
15. Exon finding
16. Genome homology

SEMESTER I

CORE PRACTICAL II

PROGRAMMING IN C & C++

1. Basic arithmetic operations using c program
2. Identification of prime number using c program
3. Sorting biggest of the given numbers using c program
4. Sorting even numbers from a range of numbers using c program
5. Sorting odd numbers from a range of numbers using c program
6. Factorial of a given number using c program
7. Identification of palindrome using c program.
8. Sorting a range of numbers using c program.
9. Sorting given numbers using c program.
10. Factorial using recursion function of the c program.

C++ Programs

1. C++ Program to Display Prime Numbers Between Two Intervals Using Functions
2. C++ Program to Check Prime Number By Creating a Function
3. C++ Program to Check Whether a Number can be Express as Sum of Two Prime Numbers
4. C++ program to Find Sum of Natural Numbers using Recursion
5. C++ program to Calculate Factorial of a Number Using Recursion
6. C++ Program to Convert Binary Number to Decimal and vice-versa
7. C++ Program to Convert Octal Number to Decimal and vice-versa
8. C++ Program to Convert Binary Number to Octal and vice-versa
9. C++ Program to Calculate Average of Numbers Using Arrays
10. C++ Program to Find Largest Element of an Array
11. C++ Program to Calculate Standard Deviation
12. C++ Program to Add Two Matrix Using Multi-dimensional Arrays
13. C++ Program to Multiply Two Matrix Using Multi-dimensional Arrays
14. C++ Program to Find Transpose of a Matrix
15. C++ Program to Multiply two Matrices by Passing Matrix to Function
16. C++ Program to Access Elements of an Array Using Pointer
17. C++ Program to Swap Numbers in Cyclic Order Using Call by Reference
18. C++ Program to Find the Frequency of Characters in a String
19. C++ Program to Remove all Characters in a String Except Alphabets.

20. C++ Program to Find the Length of a String
21. C++ Program to Concatenate Two Strings
22. C++ Program to Store Information of a Student in a Structure
23. C++ Program to Copy Strings

CORE ELECTIVE PAPER - I

A. MATHEMATICAL AND STATISTICAL METHODS IN BIOINFORMATICS

Objective

To introduce the students in mathematical and statistical tools and techniques.

UNIT-I

Role of statistics in Biology and Bioinformatics - Collection and Representation of Experimental data - Measures of Central Tendency and Location: Arithmetic Mean, Median, Mode, Position of averages, Geometric Mean, Harmonic mean and percentile - Measures of Dispersion: Range, Interquartile range, mean deviation, variance and standard deviation.

UNIT-II

Correlation and Regression: Correlation coefficient - Types of correlation - Regression equation - Principles of least squares - Linear regression - Biological significance of correlation and regression - Tests of significance: Basis of statistical inference - Student's 't' test for mean, difference of means and test for correlation and regression coefficients – Chi - square test - Analysis of Variance

UNIT-III

Basic concepts of Probability - Sample space and Events - The use of counting methods in probability - Addition law - Conditional probability - Simple problems involving the estimation of probabilities - Normal Distribution and Binomial and Poisson distributions - Z-score, P-value and E-value - Hidden Markov Models - Neural Networks - Applications of probability concepts in Bioinformatics

UNIT-IV

Matrices: Matrix algebra - Types of matrices - determinant - inverse, rank of matrix - solution of simultaneous equations - rotation matrices and co-ordinate transformation

Vectors: Vector algebra - addition and subtraction of vectors - product of vectors, dot & cross products - scalar triple product - vector calculus - gradient, divergence, curl of a vector & identities - applications.

UNIT-V

Basic differentiation of algebraic and trigonometric functions - Maxima and Minima - Integration of simple functions - Definite and non-definite integrals - Table of integrals - Numerical methods for differentiation and integration - applications to systems biology

Text Books

1. Gupta, S.C. and Kapoor, V.K. 2002.Fundamentals of Mathematical Statistics, 11th Edition, Sultan Chand & Sons, New Delhi.
2. Jordan, D.W. and Smith, P. 2002. Mathematical Techniques, 3rd Edn, Oxford University Press, New Delhi.

3. Forthofer, L. 1995. Introduction to Biostatistics, Academic Press, New York.
4. Sokal, Robert R. and Rohlf, F.J. 1987. Introduction to Biostatistics (Biology-Statistics Series), W.H. Freeman & Company, New York..
5. Batschelet, E. 1991. Introduction to Mathematics for Life Scientists, 2nd Edn., Springer International Student Edn., Narosa Publishing House, New Delhi.

CORE ELECTIVE PAPER - I

B. BIODIVERSITY INFORMATICS

Objective

To aware of digitized biodiversity data resource available nationally and internationally and to utilize the same effectively to conserve biodiversity.

UNIT-I

Biological diversity of life - Methods for species identification & classification - Information needs in biodiversity assessments and inventorying programmes - Role of information technology in distributing biodiversity information.

UNIT-II

Introduction to Biodiversity Informatics - Assessing, analyzing and documenting biodiversity - Morphological and molecular characterization of biodiversity - Introduction to biodiversity database: endangered animals, endemism and Red data books - Biodiversity registers.

UNIT-III

Designing information systems to support biodiversity conservation - Networks for distributing information - Distributed Databases and Web -Accessible Resources - Species 2000 and Tree of life.

UNIT-IV

Software for identification of Accessing existing biodiversity databases on the World - wide Web-Probabilistic and deterministic identification, Delta, MicroIS, AVIS, ICTV.

UNIT-V

Global biodiversity information system - Overview of the UNEP/GEF biodiversity data management project (BDM) - CBD and bioethics - General agreement on trade and tariffs.

Text Books

1. Kevin J. Gaston and John I. Spicer. Biodiversity - An introduction
2. Agarwal., K.C., Biodiversity

References

1. Global Biodiversity : Status of the Earth's Living Resources. Water Conservation Monitoring Centre (1992), Chapman & Hall, London.
2. Systematics and Conservation Evaluation - Forey, P.L., C.J. Humphries and R.I Vane-Wright (eds) (1994), Clarendon Press, Oxford.
3. Biodivesity: Measurement & Estimation - Hawkswoth, D.I. (Ed.) (1995), Chapman & Hall, London.
4. Alice , 1990. A Bio-Diversity database system. Alice Software Partnership. Cnhos, D.A.L. Canhos, V.P. and Kirsop, B.E (eds) 1994. Linking Mechanisms for biodiversity Information, Tropical Foundation, Campinas, Brazil.
5. Uhlir, P.F,. 1980. The Public international law of Civilian remote sensing: an overview. In: Mink, P.D. (ed), American Enter Prise, The law, and the commercial use of spece, Vol II. National Legal Center for the Public Interest, Washington, Dc.
6. Heywood, V.H., Watson, R.T. 1995. Global Biodiversity Assessment. Published for the United Nations Environment Programme, Cambridge University Press, Cambridge.

Web Resource

www. Biodiv.org

www.wri.org/wri/biodiv/

www.wcmc.org.uk/

CORE ELECTIVE PAPER - I

C. GENETICS & EVOLUTION

Course Objectives

The course is intended, to provide knowledge on the structure and functions of genes and chromosomes, to understand the concepts of mutations, genetic aberrations, and changes due to regular cellular propagations to understand the types of changes and the consequences leading to evolution to understand the process of genetic evolution to be familiar with some of the approaches for assessment of genetic variability and evolution

Unit-1:

Genes and inheritance, laws of inheritance, Mendalian hypothesis and tests, Concept of alleles, dominance, complementation, epistasis, recombination, gene mapping in prokaryotes and eukaryotes, sex-linked inheritance, chromosomes, mitosis and meiosis and evolution.

Unit-2:

Structure of DNA, DNA mutations, types, mechanisms of repair, origin of genetic variability; types of mutagens; allelic variation and gene functions, euploidy, aneuploidy, polyploidy; chromosomal rearrangements - deletion, duplication, inversion, and translocation; linkage, recombination, cross-over and evolution.

Unit-3:

Genetic code; replication and fidelity of DNA polymerase, transcription, alternative splicing, translation; gene-protein relationships; and the transfer of genetic information between generations, Gene transfer in bacteria and viruses, horizontal and vertical transfer of DNA, transposable elements and transposon mutagenesis, jumping genes.

Unit-4:

Structure of genome of eukaryotes, prokaryotes and archaea; evolution of mitochondrial and chloroplast genomes, evolution of nuclear genome and the origin of eukaryotic cells, sequence diversity of chromosomes across taxa, genome duplication, large-scale chromosomal alterations, Genome-wide association studies, Inbreeding, Evolution of sex chromosomes, Evolution of human DNA sequence families.

Unit-5:

Inheritance of complex traits, analysis of quantitative traits and statistics, estimation of heritability, Hardy-Weinberg equilibrium, natural selection and random genetic drift, phylogenetics and phylogenetic estimation, phylogenetic trees, cladograms, clustering methods, similarity and distance tables, Maximum likelihood and maximum parsimony methods, dotplot, dynamic programming, pairwise and multiple sequence alignment, methods of BLAST.

Text Books

1. Peter D. Snustad and Michael J. Simmons. Principles of Genetics. 6th edition, John Wiley & Sons, Inc., 2012
2. Watson JD, Baker TA, Bell, SP, Gann A, Levine M and Losick R, Molecular Biology of the Gene, 5th Edition, CSHL Press, 2004
3. Subramanian C. Genomic Bioinformatics, Dominant Publishers and Distributors, New Delhi, 1st edition 2009.

Reference Items: books, Journal

1. Brooker RJ. Genetics – analysis and principles, 4th edition. McGrawHill publications, 2012

OPEN ELECTIVE

PAPER I

BIOLOGICAL DATABASES

Unit-1

Nature of biological data, Overview of available Bioinformatics resources on the web NCBI/EBI/EXPASY etc, Biological Databases: Nucleic acid sequence databases GenBank/EMBL/DDBJ, Biological Databases: Protein sequence databases - PIR-PSD, SwissProt, UniProtKB

Unit-2

Database search engines Entrez, SRS • Overview/concepts in sequence analysis • Pairwise sequence alignment algorithms o Needleman & Wunsch o Smith & waterman

Unit-3

Scoring matrices for Nucleic acids and proteins MDM BLOSUM CSW • Database Similarity Searches BLAST, FASTA

Unit-4

Multiple sequence alignment PRAS CLUSTALW • Biological databases: Genome & genetic disorders Genome databases: Human, model organisms, microbes & viral OMIM

Unit-5

Biological databases: structural databases PDB NDB CCSD • Derived databases Prosite BLOCKS Pfam/Prodom.

REFERENCES

- Bioinformatics: A Practical Guide to the analysis of Genes and Proteins (2nd Ed.) by Baxevanis, A.D. & Ouellettee, B., F. F., New York, John Wiley & Sons, Inc. Publications, 2002.
- Introduction to Bioinformatics by Attwood, T.K. & Parry-Smith, D.J., Delhi, Pearson Education (Singapore) Pte.Ltd., 2001.
- Bioinformatics: Sequence and Genome Analysis by Mount, David, New York, Cold Spring Harbor Laboratory Press, 2004

SEMESTER-II

CORE PAPER I

GENOMICS & PROTEOMICS

Subject description :
This paper deals with genome map, comparative genomics, structural genomics, functional genomics, protein structure prediction and function and various tools for analysis of proteins.

Goals:
To make the students to familiar with genome map, comparative genomics, structural and functional genomics and Proteomics —extensively used in drug discovery, and in learning various tools for analysis of proteins.

Objectives:
To understand the genome architecture and extracting information like gene function, gene regulation, protein evolution and targets for drug designing.

UNIT I
Annotation of the Genome
Various approaches in gene prediction, ORF prediction, Gene prediction in prokaryotes, Gene prediction in eukaryotes, Pattern discrimination, Evaluation of gene prediction method, Prediction of promoter sequences.
Genome analysis
Chromosome rearrangement, Compositional analysis, Clustering of genes, Composite genes.

UNIT II
Functional Genomics
Gene expression analysis by cDNA micro arrays, SAGE, Strategies for generating ESTs and full length inserts; EST clustering and assembly; EST databases (DBEST, UNIGENE); Expression and regulation of entire set of genes, Sporulation Vs Vegetative condition in yeast and *Bacillus.*

UNIT III
Comparative Genomics
Purpose and Methods of comparison
Methods of comparison, Comparison at Nucleotide level, Breakpoints level, Gene cluster level.
Applications of comparative Genomics
Predicting function, Predicting regulatory elements, Analysis of conserved strings.

UNIT-IV
Principles of Protein classification:

Based on Structural features, Phylogenetic relationship, CATH - Classification by Class, Architecture, Topology, Homology, SCOP - Structural Classification Of Protein, FSSP - Fold classification based on structure - structure alignment, MMDB - Molecular Modeling Database, SARF - Spatial arrangement of backbone fragments.

UNIT - V

Proteome analysis

2D Electrophoresis**,** Immobilized pH gradient, Sample preparation, First dimension criteria, second dimension criteria, Stabilization, Detecting protein on gel, Electro blot, Image analysis, Digital imaging, Spot detection and quantification, Gel matching.

REFERENCES

1. Bioinformatics Sequence and Genome Analysis. 2001. David W. Mount. Cold Spring Harbor laboratory Press.
2. Inna Dubchak et al. 2000, Active conservation of noncoding sequences revealed by three way species comparisons. *Genome Research*. **10**, 1304-1306
3. Proteomics. S.R. Pennigton and M.J. Dunn. 2002. Viva Books Private Limited. New Delhi. (for Units III and IV and V.)
4. Introduction to Protein Structure. Carl Branden and John Tooze 1999. Garland Publishing. New York. (for Units I and II)
5. Protein Evolution by Laszlo Patthy 1999. Blackwell Science
6. Shanmughavel, P. 2005. Principles of Bioinformatics, Pointer Publishers, Jaipur, India

SEMESTER-II

CORE PAPER II

RELATIONAL DATABASE MANAGEMENT SYSTEM AND MySQL

Objective

The primary goal of this subject is to provide the knowledge on relational database. It imparts the skill on normalization and database design. It inculcates the knowledge on management of databases.

UNIT-I

Introduction - History of database systems - Applications of database systems - Database systems vs. file systems - View of data: Data abstraction - Instances and Schema - Database system structure - Database architecture - Database administrators and users - Transaction - Homogenous and Heterogeneous data - Advantages and disadvantages.

UNIT-II

Types of data models - Relational model - Relational algebra and calculus - Relational databases - Relational languages - Relational-database design - Object-Relational databases and other hybrid databases; Integrity and security - Constraints - Normalization - Indexing and hashing.

UNIT-III

SQL languages: Data Definition Language (DDL) - Data Manipulation Language (DML) - Transaction Control Language (TCL) - Data Control Language (DCL) - Basics of SQL - MySQL datatypes - MySQL operators - MySQL Functions.

UNIT-IV

Working with databases using MySQL commands - Working with tables using MySQL commands - Working with datas using MySQL commands - Joins - Subqueries - Transactions.

Basics of PL SQL and simple PL SQL programs.

UNIT-V

Managing scientific data: Introduction - Challenges faced in the integration of biological information - Data management and data integration in Bioinformatics - Issues to address while designing a biological information system - SRS: An integration platform for database and analysis tools in bioinformatics - An integration challenges in gene expression data management - discovery link

Text Books

1. Silberschatz, Korth and Sudarshan, Database System concepts, Tata McGraw-Hill, New Delhi, 2006.
2. Vikram vaswani, The complete reference for MySQL, Tata McGraw-Hill, New Delhi, 2004.
3. Lacroix Critchlow, Bioinformatics - Managing scientific data, Elsevier, New Delhi, 2003.

SEMESTER-II

CORE PAPER III

STRUCTURAL BIOLOGY

UNIT 1

Levels of molecular organization, Brief discussions on: Amino acids, Nucleic acids, Adenylates, Carbohydrates, Lipids, Cofactors, Vitamins, and Hormones. Composition and primary structures of proteins, Conformational analysis and forces that determine protein structures, geometries, phi, psi, omega angles, Ramachandran or steric contour diagram, allowed chi angles of side chains in proteins, hydrogen bonding, disulphide bonds, hydrophobic interactions, vanderwaals forces, potential energy calculations, alpha helices, beta sheets, helix to coil transition, general features and thermodynamic aspects of protein folding, folding kinetics, protein-ligand interactions, Scatchard plot, cooperative interactions, allosteric effects, Hill constant; Relationship between the primary, secondary, and tertiary structure of proteins. Structure of IgG, fibrous proteins (structure of collagen, keratin). Quaternary structures - dimers, homo & hetero dimers, trimers, tetramers; Protein folds, structural families and classes, multifunctional domains.

UNIT 2

General characteristics of nucleic acid structures (A, T, G, C, U), forces and stabilizing geometries, glycosidic bond, rotational isomers. Stabilizing ordered forms of DNA (A, B and Z), base pairing types, base stacking, tertiary structure of DNA (Supercoiled DNA), Melting of the DNA double helix (Hyperchromicity), Interaction with small ions and small molecules. Ribose puckering and Tertiary structure of tRNA. Structure and conformational properties of cell membranes, Singer and Nicholson model, integral proteins in membranes, conformational variations during ion transport, Signal transduction and molecular reception.

UNIT 3

Rayleigh scattering, ultra-centrifugation, viscometry. Electron microscopy (SEM-TEM, AFM), luminescence (fluorescence & phosphorescence), Calorimetry, DSC, Mass spectrometry, LCMS, MALDI-TOF, Voltage Clamp and Patch Clamp (measurements of membrane potentials).

UNIT 4

X-ray diffraction: structure determination via single crystal diffraction, fibre diffraction; Neutron diffraction. XAFS. NMR spectroscopy (structure determination). ORD/CD, UV, IR, Laser Raman, ESR/EPR.

UNIT 5

Association of macromolecules, molecular conjugates, supramolecular interactions, protein-protein interactions, protein-nucleic acid interactions, lipid/membrane-protein interactions. Molecular mechanics and dynamics (Newtonian and Monte Carlo simulations), theoretical

principles and its importance towards insilico simulations, results of molecular dynamics calculations and their implications to biological function.

REFERENCE BOOKS

1. Biophysical Chemistry by Cantor R. and Schimmel P.R, W. H. Freeman.

2. Physical Biochemistry by David Freifelder, W H Freeman and Company.

3. Biophysical Principles of Structure & Function by Fred M. Snell & Sidney Shulman.

4. Introduction to Protein Structure by Carl Branden and John Tooze, Garland Publishing.

5. Proteins Structure – A Practical Approach by Creighton, Oxford University Press.

6. Physical Chemistry: Principles and Applications in Biological Sciences by Tinoco and others, Prentice Hall.

CORE PRACTICAL I

BIOCHEMISTRY, BIOPHYSICS AND MOLECULAR BIOLOGY

BIOCHEMISTRY

1. Estimation of reducing sugar.
2. Estimation of lipids.
3. Separation of amino acids and lipids using TLC and Paper chromatography.
4. Separation of a mixture of proteins (2 or 3) using column chromatography.
5. Estimation of proteins using Bradford and Lowry's methods.
6. Extraction of proteins from microbes and plants.
7. Blood analysis, estimation of RBC count, WBC count

BIOPHYSICS

8. Microscopy: Bright field, Phase contrast & Fluorescence microscopy
9. To verify the Lambert Beer's law.
10. Protein crystallization using hanging drop and sitting drop methods.
11. Casting the Gel for SDS-PAGE.
12. Separation of protein and molecular weight determination using SDS-PAGE.
13. Staining the gel with CBB.

MOLECULAR BIOLOGY

14. Histochemical techniques for Plant cells and tissues.
15. Mitosis - Onion root tip squash with heamatoxylin staining.
16. Meiosis - Tradeschantia anther squash with Acetocarmine staining
17. Isolation of Mitochondria
18. Isolation of Chloroplast
19. Microbial genomic DNA isolation
20. Microbial plasmid isolation
21. Plant genomic DNA isolation.
22. Agarose Gel electrophoresis and gel documentation.
23. DNA amplification using Thermocycler.
24. Blotting Techniques - Southern, Northern & Western.
25. Hybridization - Autoradiography - Demonstration

CORE PRACTICAL II

SEQEUNCE ANALYSIS

1. Biological databases (sequence, structure and specialized databases)
2. Data retrieval using ENTREZ
3. Database file formats
4. Gene finding (Genscan)
5. Protein prediction
6. Sequence search
7. Sequence alignment
8. Phylogenetic tree construction
9. Sequence comparison
10. Structure analysis
11. Pattern recognition
12. Proteome analysis using tools
13. Exon finding
14. Genome homology
15. Molecular visualization using Rasmol

CORE PRACTICAL III

PROGRAMMING IN C & C++

I. Character array manipulations

1) Read and Display a character array

2) Reverse print the array (Sting Reverse)

3) Length of the array

4) Copying the contents of one array to another (String Copy)

5) Copy the Uppercase character of one array as Lowercase character to another array

6) Checking whether a string is a palindrome or not

7) Copy the left „n‟ characters of one array to another

8) Copy the last „n‟ characters of one array to another

9) Copy the middle „n‟ characters of one array to another

10) Concatenate two character arrays (String Concatenate)

11) Counting the numbers of Words, Lines and characters in an array

12) Counting the numbers of Uppercase and Lowercase Alphabets, Digits and special characters in an array

13) Check the number of occurrences of a pattern

14) Check the occurrences of a pattern and skip the same.

15) Check the occurrences of a pattern and replace it with a different pattern

II. Pointers and Character Array

16) Pattern Counting

17) Pattern Skipping

18) Pattern Replacing

III. Files and Command Line Arguments

19) Read data from the keyboard and write it in the file (char by char)

20) Read data from the file and display it on the screen (char by char)

21) Display the content of all the files (Cat all the files)

C++ PROGRAMS

1) C++ Program to Add Two Numbers

2) C++ Program to Check Whether Number is Even or Odd

3) C++ Program to Find Largest Number Among Three Numbers

4) C++ Program to Find All Roots of a Quadratic Equation
5) C++ Program to Calculate Sum of Natural Numbers
6) C++ Program to Check Leap Year
7) C++ Program to Find Factorial
8) C++ Program to Generate Multiplication Table
9) C++ Program to Reverse a Number
10) C++ Program to Multiply two Numbers
11) C++ Program to Check Whether a Number is Palindrome or Not
12) C++ Program to Check Whether a Number is Prime or Not
13) C++ Program to Display Prime Numbers Between Two Intervals
14) C++ Program to Display Armstrong Number Between Two Intervals
15) C++ Program to Display Factors of a Number
16) C++ Program to Display Prime Numbers Between Two Intervals Using Functions
17) C++ Program to Check Prime Number By Creating a Function
18) C++ Program to Check Whether a Number can be Express as Sum of Two Prime Numbers
19) C++ program to Find Sum of Natural Numbers using Recursion
20) C++ program to Calculate Factorial of a Number Using Recursion

CORE ELECTIVE - PAPER 2

A) BIOPHYSICS AND BIOCHEMISTRY

Objective

To impart knowledge to the candidates on structural, functional and dynamic aspects of biological components

UNIT-I

Classification, Structure, Properties and Biological role of Carbohydrates. Carbohydrate Biosynthesis, Metabolism - Glycolysis, TCA cycle and ATP bioenergetics.

UNIT-II

Structure, classification. Properties and Biological role of Lipids. Storage of fatty acids. Lipid, Biosynthesis, Metabolism. Utilization of fatty acids for energy production - β Oxidation**.**

UNIT-III

Introduction to protein structure - Composition and dynamic structural properties, primary and higher level protein organization. Structural components of Nucleic acids. DNA structure, function and sequence. Properties, structure and types of RNA. Nucleic acid Metabolism - Conformational properties of proteins - Ramachandran, Chandrasekaran and Sasisekaran plots, secondary, super secondary, tertiary and quaternary structure of protein.

UNIT-IV

Biophysics - Introduction - Thermodynamics - Laws of thermodynamics - Energy states - Ground and Excited states - Electromagnetic spectrum - Absorption of light by atoms & molecules - Paulis exclusion principle - Coupling of chemical reactions - Endergonic and exergonic coupling - Redox reactions - Hydrogen half cell.

UNIT-V

ATP synthesis - ETC in chloroplast and mitochondria - Oxygen electrode - Classification, Characteristics of enzymes - Biological role - Enzyme Kinetics, Regulation of enzyme activity and Factors affecting enzyme kinetics - Formulation and significance of Michaelis - Menton Equation.

Text Books

1. Berg, J.M., Tymoczko, J. L. and Stryer, L. 2002. Biochemistry. 5th Edn. W.H Freeman and

Company, New York.

2. Devlin, T. M. 2002. Text book of Biochemistry - with clinical correlations. 5th Edn. John Wiley & Sons Inc., New York, USA.

3. Freifelder, D. and Malcinski, G.M. 1993. Essentials of Molecular Biology, 2nd Edn. Jones & Bartletl Publishers Inc., London.

4. Nelson, D.L. and Cox, M.M. 2005. Lehninger's Principles of Biochemistry, 4th Edn. Replika Press Pvt. Ltd., New Delhi.

5. Voet, D. and Voet, J.G. 1990. Biochemistry. John Wiley and Sons Inc., New York.

References

1. Atherly, A.G., Girton, J.R. and McDonald, J.F. 1999. The science of Genetics – Saunders College Publishers, New Delhi.
2. Bickerstaff, G.F.1997. Immobilization of Enzymes and Cells. Humana Press, New Jersey, USA.
3. Bray, A., Raff, L. and Watson, R. 1994. Molecular biology of the Cell. 3rd Edn. Garland Publishing Company, New York.
4. Click, B.R. and Pastumak, J.J. 1998. Molecular Biotechnology - Principles and application of recombinant DNA. American Society of Microbiologists Press, Washington.
5. Cooper, G.M. and Hausman, R. E. 2004. The Cell: A Molecular Approach, 3rd Edn. American Society of Microbiologists Press, Washington.
6. Karp, G. 1996. Cell and Molecular Biology: Concepts and Experiments. John Wiley & Sons, New York.
7. Micklos, D.A., Freyerr, G. A. and Crotty, D. A. 2003. DNA Science, 2nd Edn. Cold Spring Harbor Laboratory Press, New York.
8. Primrose, S. B. 1994. Molecular Biotechnology, 2nd Edn. Blackwell Scientific Publishers, Oxford.

CORE ELECTIVE - PAPER 2

B. BIOLOGICAL ALGORITHMS IN COMPUTATIONAL BIOLOGY

COURSE OBJECTIVES: To make the students understand the application of Artificial Intelligence in Biocomputing.

Unit 1 DNA computing: Motivation, DNA structure, processing and computational operations, steps involved in DNA computation, Filtering models: Adleman's experiment, Lipton's solution, Scope and Applications of DNA computing. Search Algorithms: Hill climbing, Simulated annealing:-introduction, Simulated annealing algorithm,

Unit 2 Combinatorial Pattern Matching: Hash Tables, Repeat Finding, Exact Pattern Matching; Genetic Algorithm: Basic Concepts, Reproduction, Cross over, Mutation, Fitness Value, Optimization using GAs; Applications of GA in bioinformatics.

Unit 3 Hidden Markov Model: Markov processes and Markov Models, Hidden Markov Models. Forward and Backward Algorithms, Most probable state path: Viterbi algorithm, Parameter Estimation for HMMs:-Baum-Welch Algorithm, Applications of profile HMMs for multiple alignment of proteins and for finding genes in the DNA.

Unit 4 Support Vector Machines: Introduction, hyperplane separation (maximum and soft margin hyperplanes), linear classifier, Kernel functions, Large Margin Classification, Optimization problem with SVM, Applications of SVM in bioinformatics. Bayesian network: Bayes Theorem, Inference and learning of Bayesian network, BN and Other Probabilistic Models.

Unit 5 Artificial Neural Network: Historic evolution – Perceptron, characteristics of neural networks terminology, models of neuron Mc Culloch – Pitts model, Perceptron, Adaline model, Basic learning laws, Topology of neural network architecture, single layer ANN, multilayer perceptron, back propagation learning, input - hidden and output layer computation, back propagation algorithm, Applications of ANN.

Text Books:

1. An introduction to bioinformatics algorithms by Neil C. Jones, Pavel Pevzner. MIT Press.2004
2. Biological sequence analysis: Probabilistic models of proteins and nucleic acids by Richard Durbin, Eddy, Anders Krogh, 1998
3. Algorithms for Molecular Biology by Ron Shamir Lecture, Fall Semester, 2001
4. Neural Networks: A Systematic Introduction by Raul Rojas. Springer. 1996
5. Artificial Intelligence and Games by Georgios N. Yannakakis and Julian Togelius, Springer 2018

Reference Books:

1. Bioinformatics: the machine learning approach by Pierre Baldi, SørenBrunak. MIT Press.2001.

2. Bioinformatics: Sequence and Genome Analysis: by David Mount, University of Arizona,Tucson. 2005
3. Fundamentals of natural computing : Basic concepts, Algorithms and Applications, Chapman & Hall / CRC, Taylor & Francis group, 2006
COURSE OUTCOME: Students are trained in the application of Artificial Intelligence in Biocomputing

CORE ELECTIVE - PAPER 2
C. CHEMINFORMATICS

Course Objectives:

- To make the students understand the basics of cheminformatics and their application.
- To aware the various chemical information sources.
- To analyze the pharmacokinetic properties of small molecules using ADMET calculation.
- To understand the steps in pro drug design.
- To utilize the bioinformatics tools and software in different aspects.

Syllabus

Unit-I: Basic Mathematics and Statistics

Graph theory and molecular numerology; Logic, sets and functions; Algorithms, integers and matrices; Mathematical reasoning, induction and recursion; Counting; graphs, trees and sets, basic probability and statistics; Markov processes.

Unit-II: Foundations of Chemistry and Biology

Basic Stereochemistry, Group Theory, Amino acids and Proteins and Properties; pKa, pH and ionization of acids and bases; Protein structure - Primary structure, Secondary structure - helix & sheet; Tertiary structure; Quaternary structure; covalent and non-covalent forces that maintain structures. Physical properties of proteins - charge, size, hydrophobic, protein binding – structural aspects; antibodies; transport; nucleotide binding; catalytic enzymes; basic concepts of combinatorial chemistry. Introduction to drug action, pro drug design and applications.

Unit-III: Chemical information sources

History of scientific information communication-chemical literature-chemical information chemical information search-chemical information sources-chemical name and formula searching-analytical chemistry-chemical history-biography-directories and industry sources.

Unit-IV: Bioinformatics

Introduction; Experimental sources of biological data; Publicly available databases; Gene expression monitoring; Genomics and Proteomics; Metabolomics; Visualisation of sequence data; Visualization of structures using Rasmol or SPDB Viewer or CHIME; Genetic basis of disease; Personalised medicine and gene-based diagnostics; Legal, ethical and commercial ramifications of bioinformatics.

Unit-V: Pharmaceutical applications of molecular modeling

Introduction to drugs, structure-based drug design. QSAR and 3D-QSAR Methods. Pharmacophore Design, Ligand-Based Design and De Novo Drug Design Virtual screening/docking of ligands. Protein structure. Drug action enzymes. Drug action receptors. Drug design target interaction. Prediction of Binding Modes, Protein–ligand binding free

energies, Fragment-Based Drug Design; Absorption, Distribution, Metabolism, Excretion & Toxicology (ADMET) prediction; Calculation of Physico-Chemical Properties, Biological and Physico-Chemical Predictive Model Building.

References:

- "Mathematical Methods for Physicists" Arfken, Academic Press 1985
- Schaum's Outline of Probability and Statistics, Murray R Spiegel, John J. Schiller, R. Alu Srinivasan
- Stereochemistry, by David G. Morris, Eddie Abel
- Introduction to Protein Structure: Second Edition ,Carl Branden , John Tooze
- Combinatorial Chemistry and Molecular Diversity in Drug Discovery, Eric M. Gordon , James F. Kerwin
- Computer-Aided Drug Design: Methods and Applications, T.J. Perun C.L. Propst
- Chemical Information Sources (Mcgraw-Hill Series in Advanced Chemistry) ,Gary Wiggins
- Introduction to Bioinformatics, Teresa K. Attwood, David Parry-Smith
- Molecular Modeling: Basic Principles and Applications, 3rd Edition,Hans-Dieter Höltje, Wolfgang Sippl, Didier Rognan, GerdFolkers

OPEN ELECTIVE PAPER 2

BIOLOGICAL SEQUENCE ANALYSIS

Subject description :
This paper describes how to acquire information from biological databases, use of computational approaches to analyze this information, and interpret the results as a guide to experiments in biology.

Goals: The goal of this course is to introduce the main principles of bioinformatics. The coverage will include concepts like sequence alignments, phylogenetic trees, and structure prediction.

Objectives: Understand Genomic data acquisition and analysis, comparative and predictive analysis of DNA and protein sequence, Phylogenetic inference etc

UNIT-I
Introduction to bioinformatics, Classification of biological databases, Biological data formats, Application of bioinformatics in various fields. Introduction to single letter code of aminoacids, symbols used in nucleotides, data retrieval- Entrez and SRS.

UNIT-II
Introduction to Sequence alignment. Substitution matrices, Scoring matrices – PAM and BLOSUM. Local and Global alignment concepts, Dot plot. Dynamic programming methodology: Needleman and Wunsch algorithm. Smith–Waterman algorithm. Statistics of alignment score. Multiple sequence alignment. Progressive alignment. Database search for similar sequences using FASTA and BLAST Programs.

UNIT-III
Evolutionary analysis: distances, Cladistic and Phenetic methods. Clustering Methods. Rooted and Unrooted tree representation. Bootstrapping strategies, Use of Clustal and PHYLIP.

UNIT-IV
Gene finding methods. Gene prediction: Analysis and prediction of regulatory regions. Fragment assembly. Genome sequence assembly, Restriction Mapping, Repeat Sequence finder.

UNIT-V
Concepts of secondary structure prediction of RNA and Protein. Probabilistic models: Markov chain, Hidden Markov Models-other applications.

REFERENCES

1. **Bioinformatics – Concepts, Skills, Applications".** S.C. Rastogi, Namita Mendiratta, Parag Rastogi.
2. **Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins.** Andréa's D. Baxevanis, B.F. Francis Ouellette.
3. **Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids.** Richard Durbin et al.
4. **Computer Methods for Macromolecular Sequence Analysis**. Doolittle R.F. (Ed.) (Methods in Enzymology, VOl. 266).
5. Shanmughavel, P. 2005. **Principles of Bioinformatics,** Pointer Publishers, Jaipur, India.
6. **DNA and Protein Sequence Analysis. A Practical approach.** Bishop M.J. Rawlings C.J. (Eds.).
7. **Introduction to Bioinformatics.** Teresa. K. Atwood and David J. Parry-Smith.

****