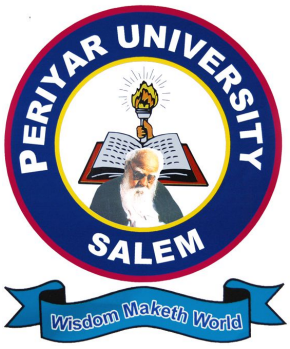

PERIYAR UNIVERSITY
SALEM – 11

M.Sc., (BIOINFORMATICS)

Semester System

(Effective from the academic year 2017-2018)

M.Sc. BIOINFORMATICS

CHOICE BASED CREDIT SYSTEM

(For the students admitted from the year 2017-18 onwards)

Part	Course Code	Course	Hrs.	Credit	Marks		
					CIA	EA	Total
Semester I							
Core I	17PBI01	Introductory Bioinformatics	5	5	25	75	100
Core II	17PBI02	Cell Biology	5	5	25	75	100
Core III	17PBI03	Biomolecules	5	5	25	75	100
Elective	17PBIE01	Mathematics and Biostatistics	5	4	25	75	100
Practical	17PBIP01	Lab Course – I	5	4	40	60	100
Practical	17PBIP02	Lab Course- II	5	4	40	60	100
Semester II							100
Core IV	17PBI04	Programming for Bioinformatics	5	5	25	75	100
Core V	17PBI05	Genomics and Proteomics	5	5	25	75	100
Core VI	17PBI06	Database Management systems	5	5	25	75	100
Elective	17PBIE02	Molecular Biology and Genetic engineering	5	4	25	75	100
Value Edu		Human Rights	2	1	25	75	100
Practical	17PBIP03	Lab Course- III	5	4	40	60	100
Practical	17PBIP04	Lab Course- IV	5	4	40	60	100
Semester III							
Core VII	17PB107	Molecular Modelling and Drug Design	5	5	25	75	100
Core VIII	17PB108	Omics and Systems Biology	5	5	25	75	100
Core IX	17PB109	Big Data Analysis and Next Generation Sequencing	5	5	25	75	100
Elective	17PBIE03	Data Warehousing and Data Mining	5	4	25	75	100
Practical	17PBIP05	Lab Course-V	5	4	40	60	100
Semester IV							
Core X	17PB110	Research Methodology	5	5	25	75	100
	17PB111	Project Work		9			200

PROJECT WORK

Dissertation submission: 120

CIA: 30

Viva-voce: 50

Total: 200

Total Credits: 92

Total Marks: 2100

SEMESTER - I

INTRODUCTORY BIOINFORMATICS

UNIT I

Introduction to Bioinformatics; Computers in Biology to understand biological system; Basic commands of Windows, Unix and Linux operating systems; Concept of open resources in Bioinformatics.

UNIT II

Sequence analysis: Biological background for sequence analysis; Sequence alignment; Global, Local, Pairwise and Multiple sequence analysis; Algorithm for alignments; Database searching; Tools for sequence alignment.

UNIT III

Biological Databases: Database concepts; Introduction to data types and source; Protein sequence and structural databases; Nucleic acid database; Genome databases; Specialized databases; Carbohydrate databases; clinically relevant drug-drug interaction databases; Information retrieval from Biological databases; Entrez system.

UNIT IV

Cheminformatics: Introduction; Chemoinformatics tools; Chemical structure representation (SMILES and SMARTS); Chemical databases: CSD, ACD, WDI, Chembank, PUBCHEM, Chemical structure file formats; Structural isomers; Structure visualization.

UNIT V

Medical and Pharmacy Informatics: Introduction to Pharmacy Informatics, Medical transcription; Role of informatics to enhance the services provided by pharmaceutical care givers. Health Information Systems Architecture, Health data management, Medical coding, Telemedicine and Telehealth, Ethics in medical informatics, Pharmacy systems and automation, Informatics applications in pharmacy, survey and evaluation of on-line resources.

References:

- Mount, D. 2004. "Bioinformatics: Sequence and Genome Analysis"; Cold Spring Harbor Laboratory Press, New York.
- Baxevanis, A.D. and Francis Ouellellette, B.F. 1998. "Bioinformatics – a practical guide to the analysis of Genes and Proteins"; John Wiley & Sons, UK.

- Lesk, A.M. 2002. “Introduction to Bioinformatics”, First edition, Oxford University Press, UK.
- Jean-Michel, C. Notredame, C. 2003. “Bioinformatics for Dummies”; John Wiley & Sons.
- Ole Lund, Nielsen, M., Lundegaard, C. Kesmir, C. and Brnak, S. 2005.“Immunological Bioinformatics”; The MIT press.
- R. Durbin, S. Eddy, A. Krogh and G. Hitchison. 2003. Biological Sequence Analysis, Cambridge University Press, Eighth edition.
- Andrew R. Leach & Valerie J. Gillet. 2007. An Introduction to Chemoinformatics, Springer, Revised Edition.

CELL BIOLOGY

UNIT I

Cell theory- prokaryotic and eukaryotic cells- animal and plant cell. Molecular models of cell wall and cell membrane- structure and function. Ultrastructure and function of cell organelles-Mitochondria, chloroplast, nucleus, endoplasmic reticulum, golgi apparatus and lysosome.

UNIT II

Gene transfer: Vector and vector less modes of gene transfer. Steps in gene cloning, plastid transformation. Cloning vectors- Ti plasmid, Baculovirus, Adenovirus, SV40 & M13.

UNIT III

Chromosomal architecture, classification, its aberration and banding techniques. Replication, transcription, translation, protein sorting and targeting, chaperones.

UNIT IV

Regulation of gene expression in Lac and Trp operon. Epigenetic regulation in eukaryotes, transposons and genetic recombination. Transcription factors, DNA binding, motifs- Zn, HTH, HLH, Leucine zipper.

UNIT V

Cell cycle- mitosis, meiosis. Tools to study cell biology, PCR types and application, Cell adhesion, cell matrix adhesion, collagen and non-collagen components of extra cellular matrix. Communication-second messengers, Cell death and its regulation.

Recommended Books:

- Bruce Alberts , Alexander Johnson , Julian Lewis, Martin Raff , Keith Roberts, Peter Walter, 1994. Molecular Biology of the Cell, Fourth Edition.. Academic Press. New York.
- De Robertis and De Robertis. 2005. 8th Eds. Cell and Molecular Biology. Lippincott Williams & Wilkins
- Lodish, Berk, Baltimore et al. 2000. 6th Eds Molecular Cell Biology. W.H. Freeman & Co.
- Gerald Karp, 2008. Cell and Molecular Biology: Concepts and Experiments, 5th Eds. Wiley
- Geoffrey Cooper, 2000. The Cell: A molecular approach. 2nd Eds. Sinauer Associates Inc.

BIOMOLECULES

UNIT I

Introduction to Biomolecules – Structure and function of DNA, RNA, protein, carbohydrate and lipid. Folding pathways: Principles of protein folding, role of chaperons,

UNIT II

Molecular interactions: protein-protein, protein-DNA, DNA-Drug, Protein-Lipid, Protein-Ligand, Protein-Carbohydrate interaction, Metalloprotein. Pi … Pi interactions, C-H…Pi interactions.

UNIT III

Structure and Stability of Proteins: Peptide bond; Ramachandran Plot; Calculation of conformation of different states of proteins; Properties of α-helix and β-sheets; Secondary structure prediction and determination; connectivities. Tertiary structure: Determination of state of tertiary structure; characteristic balance in rigidity and flexibility; Domain concept (α-,β-,α/β- and α+β-domains) and interacting motifs. Quaternary structure: Geometry, Symmetry and intermolecular interfaces.

UNIT IV

Structure and Stability of Carbohydrates: configuration and Conformation; Step Up and Step down reactions; Conversion of Aldos to Ketos and vice versa. Conformational stability ratings, Anomeric effect, Reverse anomeric effect; Structure and stability of polysaccharides, glycoproteins, glycolipids and proteoglycans. Structure and stability of Nucleic acids (DNA and RNA), topological structure, fine structure of DNA and its organization in genome.

UNIT V

Stereochemistry: Configuration and conformation and stability; Elements of symmetry, Chirality; RS-, EZ-, DL- and dl- system of nomenclature; Stereo-specific and stereo-selective reactions; Determination of relative configuration by asymmetric synthesis. Stereochemistry of proteins and nucleic acids.

Recommended Books:

- Principles of Biochemistry, Lehninger C Rs. Publ. (1982).
- Biochemistry, L. Stryer, W.H. Freeman, San Francisco.
- Schaum's Outline Series of Theory and Problems of Biochemistry, Philip W. Kuchel and G.B. Ralston. Int. Ed., McGraw-Hill Book Co.
- Problem Approaches in Biochemistry. Wood and Hood.
- Biochemistry by Voet and voet
- Biochemistry by Zubay
- Albert cotton, F. 1971. Chemical Application of Group Theory. John Wiley and Sons, Inc. New York. 386 pp.
- Spice, J. E. 1964. Chemical Bonding and Structure. Pergamon Press Ltd., Headington Hill Hall, Oxford. 395 pp.
- Winter, m. j. 1996. Chemical Bonding. Oxford University Press, Inc., New York. 91 pp.
- Ernest Eliel, 1996. Stereochemistry of carbon compounds, Prentice Hall

MATHEMATICS AND BIOSTATISTICS

UNIT I

Trigonometry: Trigonometric Functions, Series Expansion, Inverse, General Values, Graphs, Calculus: Limits, Continuity, Analysis, Differentiation (1D & Partial), Reimann Integration, Definite Integrals, Recursion Theorems, Ordinary and Partial Differential Equation: 1st Order and 2nd Order Ordinary Differential Equations, Self Adjoint Equations, Special Functions, Nature of Partial Differential Equations, Method of Separation of Variables.

UNIT II

Vector and Matrices: Vector Algebra, Vector Calculus, Basic Computations, Matrices. Measure Theory: Introductory Concepts, Borel Sets, Lebesgue Integration, Complex Variable: Complex Functions, DeMoivre's Theorem, Conformal Map, Complex Integration, Numerical Techniques: Basic Formalism, Methods for Solving Equations, Finding Eigenvalues & Eigenvectors, Solving ODE & PDE, Differentiation and integration.

UNIT III

Data Representation: Types of Numerical data, Tables and Graphs. Measures of central tendency: Arithmetic Mean, Weighted arithmetic mean, Median and Mode - Geometric mean and Harmonic mean. Measures of Dispersion**:** Range, Interquartile range, Average deviation, Standard deviation and Coefficient of variation, Lorenz curve. F ratio prediction and analysis of variance ANOVA.

UNIT IV

Linear Correlation and Regression: Types of Correlation, Methods of studying Correlation: Scatter diagram, Karl Pearson's Coefficient of Correlation, Spearman's Rank Correlation. Regression lines, Equations and Coefficients.

UNIT V

Probability: The concept of probability, Sample space, Independent events, Mutually exclusive events, Addition law of probability, Conditional probability, Bayes theorem.

Probability Distributions: Expected value and Variance, Binomial distribution, Poisson distribution, Normal distribution, Analysis of variance, Chi squared distribution, F- Test, Student's t distribution.

Theory of Sampling: The purpose of sampling, Principles of sampling, Methods of samplings, Techniques of non-probability sampling, Size of Sample, Sampling and Non-Sampling errors.

Recommended Books:

- Isaev, Berlin, A. 2004. "Introduction to mathematical methods in bioinformatics"; Springer.
- Raman, K.V. and Pal Sourav, P. 2005. "Mathematics in chemistry";Vikas publishing house Pvt. ltd., New Delhi.
- Stephenson, F.H. 2003. "Calculations in molecular biology and biotechnology: a guide to mathematics in the laboratory"; Amsterdam, Academic Press.
- Stephenson, G. and Radmore, P. M. 1990. "Advanced mathematical methods for engineering & science students"
- Viergever, M.A. and Todd- Pokropek, Andrew 1988. "Mathematics and computer science in medical imaging"
- Arfken, G. 1970. "Mathematical Methods for Physicists"; Academic Press, NY.
- Roman P., Pergamon. 1975. "Some Modern Mathematics for Physicists and Other Outsiders"; New York, Vol. 2, p. 660.
- Murray, J. D. 1989. "Mathematical Biology"; Springer Verlag.
- Segal, L. 1980. "Mathematical Models in Molecular and Cellular Biology"; Cambridge: Cambridge University Press.
- Balaguruswamy. "Numerical Methods"; Tata Mcgraw Hill.
- Haykin, S. 1999. "Neural Networks: A Comprehensive Foundation"; Prentice Hall.
- Zar, J.H. 1984. "Bio Statistical Methods"; Prentice Hall International Edition, USA.
- Warren, J., Gregory, E. and Grant, R. 2004. "Statistical methods in Bioinformatics"; First edition, Springer-Verlag, Berlin.
- Milton, J.S. 1992. "Statistical methods in the Biological and Health Sciences" Second Edition, McGraw Hill Publishers.
- Rosner, B. 2005. "Fundamentals of Biostatistics"; Duxbury Press.

LAB COURSE – I

LAB IN CELL BIOLOGY

1. Microscopy
2. Stains and staining procedures
3. Instrumental methods for cell biology
4. Microtomy
5. Histochemical techniques
6. Mitosis and Meiosis
7. Cell fractionation method
8. Haemotoxylin and Eosin staining in tissue samples

LAB COURSE – II

LAB IN BIOMOLECULES

1. Estimation of reducing sugar.
2. Estimation of lipids.
3. Separation of amino acids using paper chromatography.
4. Separation of amino acids and lipids using thin layer chromatography (TLC).
5. Separation of two proteins using column chromatography.
6. Estimation of proteins using Bradford and Lowry‟s methods.
7. Estimation of DNA using DPA.
8. Estimation of RNA using Orcinol reagent.
9. Estimation of Tyrosine by Million’s method.
10. Estimation of Lysine

SEMESTER – II

PROGRAMMING FOR BIOINFORMATICS

UNIT I

Programming in C: Introduction, Data types, Operators, Expressions, Control Flow, Structures, Input and Output, Functions, Pointers and References, String Processing, File Handling

UNIT II

Programming in C++: Basic concepts of OOPS-Introduction to C++,C vs C++-data types, variables, constants, operators and statements in c++- Functions in c++- function prototype-definition-inline functions-overloaded functions.

UNIT III

Programming in PERL: Introduction, Basic Operators and Control Structures, Scalars, Lists, Hashes, File Manipulation, Pattern Matching and Regular Expressions, Subroutines, Text and String Processing

UNIT IV

BioPERL Programming: General Bioperl classes, Sequences (Bio::Seq Class, Sequence Manipulation), Features and Location Classes (Extracting CDS), Alignments (AlignIO), Analysis (Blast, Genscan), Databases (Database Classes, Accessing a local Database)

UNIT V

Python Programming: Overview, Data structures, Control Flow, Modules, Basic I/O, Exception Handling, Regular Expressions, File Manipulation, Classes, Standard library

Recommended Books:

- B.W.Kernighan and D.M. Ritchie The C Programming Language, II Edition. Prientice Hall of India.
- Larry Wall, Tom Christiansen & John Orwant Programming Perl –3 ed 2000- O' Reilly.
- Mark Lutz, Programming Python, II Edition., O' Reilly.
- E. Balagurusamy , Programming in C++ - Tata Mc. Graw Hill Edition.
- Byron Gottfried.1998. Programming with C (Schaum's Outline Series) - Tata
- McGrawHill Publishing Company.
- Robert Laffore , Object oriented programming with c++ -Waite series.
- Larry. Wall, Programming Perl - Tom Christiansen, Orielly Publications.

GENOMICS AND PROTEOMICS

UNIT I

Genomes, transcriptomes and proteomes: Viral genome- bacteriophage and eukaryotic viruses. Prokaryotic genome- chromosome, linear or multipartite genome. Eukaryotic genome - yeast and human. Repetitive DNA sequences- tandem repeats of centromeres, mini satellites and micro satellites, interspersed repeats. Mobile genetic elements - Transposition and DNA transposons.

UNIT II

Restriction mapping, FISH, Sequence tagged site. Sequencing genomes-chain termination, chemical degradation, pyro sequencing. Sequence assembly - shot gun, clone contig methods. Human genome project- mapping phase- sequencing.

UNIT III

Genetic mapping-map repositories: Entrez genome map viewer- GDB, NCBI, OMIM, MGI/MGD Linkage map resources. Gene expression analysis by Micro arrays, Serial analysis of gene expression (SAGE). Applications of structural, functional and metagenomics. Comparative genomics and databases: PEDANT, COG, NCBI- gene plot, KEGG, MPG

UNIT IV

3D structure of protein- overview-protein secondary, tertiary, quaternary structure- protein Denaturation-protein folding- reverse turns- Ramachandran plot- Expasy tools. Sequencing of protein-N and C terminal sequencing- proteomics- tools and application of proteomics.

UNIT V

Analytical proteomics-analytical protein and protein separation techniques- 1D SDS-PAGE, isoelectric focusing, 2D SDS-PAGE, image analysis of 2D gels-MALDI-TOF, HPLC-protein digestion techniques. Protein identification and analysis- Mass spectrometry-tandem mass spectrometry-peptide mass finger printing- SALSA algorithm-protein arrays.

Recommended Books:

- Daniel C Liebler, 2000. Introduction to Proteomics- tools for the new biology. Humana Press.
- Pennington, SR., Dunn, MJ. 2001. Proteomics- from protein sequence to function. Viva books Pvt. Ltd.
- David, M. 2004. Bioinformatics sequence and genome analysis. Cold spring harbour laboratory Press, NY.
- Benjamin Lewin. Genes VIII.2003. Benjamin-Cummings Pub Co.
- Sandy B. Primrose , Richard M. Twyman , Robert W. Old, 2002. Principles of Gene Manipulation and genomics. 7th Edition. Blackwell Science.
- Campbell A and Laurie J. Heyer. 2006. Discovering Genomics, Proteomics, and Bioinformatics. 2nd Edition. Pearson Publishers.
- Desmond S. T. Nicholl. 2002. Introduction to Genetic Engineering. 2nd Edition. Cambridge University Press.
- Richard J. Reece. 2003. Analysis of Genes and Genomes. John Wiley High Education.

DATABASE MANAGEMENT SYSTEMS

UNIT I

Introduction to DBMS: Introduction- Purpose of Database system- Concept and Architecture- Data models, Schema and Instances- Database Languages- Database users and administrator- Advantages of using DBMS approach- Introduction to Relational database.

UNIT II

Database Design: Data Modeling using Entity- Relationship (ER) model, Entity types- entity sets, attributes and keys- Relationships- ER Diagram notations- Goals of ER diagram- ER diagram construction- Relational database design by ER- Enforcing Data Integrity Constraints

UNIT III

Structure Query Language: SQL Data definition and Data types- Create table- Specifying constraints- Basic queries- SELECT, INSERT, UPDATE, and DELETE- Complex Queries- Nested Queries- Aggregate Functions- Null Values- Schema Change Statement in SQL.

UNIT IV

Transaction processing and concurrence control: Introduction-Transaction and system concepts- Desirable properties of transaction- Serializability- Overview of concurrency

control- Locking mechanism- Two phase Locking- Time Stamps- Deadlock handling- Detection and resolution.

UNIT V

Trends in database technology: Introduction to distributed databases- Client server technology- Multidimensional and parallel databases- Spatial and multimedia databases- Mobile and Web databases- Data warehouse-Data marts.

Recommended Books:

1. Abraham Silberchatz, Henry F. Korth, S. Sudharshan, Database System Concepts (5th Edition), McGraw Hill, 2002.
2. James Martin, Computer database organization, Prentice Hall of India, 1977.
3. James Martin, Principles of Database management, Prentice Hall of India, 1976.
4. Bipin C. Desai, An introduction to database systems, Galgotia publications pvt. Ltd., New
Delhi, 2003.
5. Peter Rob Carlos Coronel, Database systems, design, implementation & management, Course technology, 2000.
6. Database systems – A practical approach to design, implementation and management, Thomas cannolly and Carolyn begg, Pearson Education, 2002.
7. Vikram Vaswani, The Complete Reference MySQL, Tata McGraw-Hill, New Delhi, 2002.
8. Zoe Lacroix and Ternce Critchlow, Bioinformatics - Managing Scientific Data, Morgan Kaufmann publishers, 2003.

MOLECULAR BIOLOGY AND GENETIC ENGINEERING

UNIT I

Introduction to Molecular Biology and Genetics- DNA as genetic material, properties and organization of DNA and RNA. DNA replication- prokaryotic and eukaryotic, inhibitors of DNA replication. DNA damage, repair and recombination.

UNIT II

Prokaryotic and eukaryotic transcription- RNA polymerase, general and specific transcription factors- regulatory elements and mechanisms of transcription regulation, eukaryotic transcriptional promoters and enhancers. Modification in RNA- transcription, splicing, editing, nuclear export of mRNA, termination, 3' end processing, cap formation, polyadenylation and mRNA stability.

UNIT III

Translation: genetic code and its features and amino acyl t-RNA synthatase, prokaryotic and eukaryotic translation- mechanism of initiation, elongation and termination. Regulation of co-translation and post translational modifications of protein. Synthesis of secretory and membrance proteins – import into nucleus, mitochondria, chloroplast and peroxisomes. Receptor mediated endocytosis, Lactose and Tryptophan operon.

UNIT IV

Biology of plasmids-types and structure of F,RTF, Col factors and Ti. Replication and partitioning. Incompatibility and copy number control. Natural and artificial plasmid transfer and their applications. Transposable genetic elements- oncogenes and tumor suppressor genes. Viral and cellular oncogenes. Tumor suppressor genes from humans-structure, function and mechanism of p53 and RB tumor suppressor proteins.

UNIT V

Molecular mapping of genome: genetic and physical maps- physical mapping and map based cloning- choice of mapping population- simple sequence repeats loci- Southern and FISH for genome analysis- chromosome microdissection and microcloning- molecular markers in genome analysis. RFLP, RAPD and AFLP analysis- application of RFLP in forensic, disease diagnosis, genetic counsellling, taxonomy and biodiversity.

Recommended Books:

- Bruce Alberts , Alexander Johnson , Julian Lewis, Martin Raff , Keith Roberts, Peter Walter. 1994. Molecular Biology of the Cell, Fourth Edition. Academic Press . New York.
- Lodish, Berk, Baltimore et al, 2000. Molecular Cell Biology. 6th Edition. W.H. Freeman & Co.
- Twyman, R.M. 2000. Advanced Molecular Biology. Garland/bios Scientific Publishers.
- Sandy B Primrose. 1991. Molecular Biotechnology. 2nd Edition. Blackwell Scientific Publishers.
- Brown, T.A., 2002. Genomes. 2nd Edition. Wiley-Liss (New York).
- Larry Snyder, Wendy Champness, 2002. Molecular Genetics of Bacteria. 2nd Edition. Amer Society for Microbiology.
- Anthony J.F. Griffiths, Anthony J.F. Griffiths, 2004. An Introduction to Genetic Analysis. 8th Edition.W H Freeman & Co.

LAB COURSE III

PROGRAMMING IN BIOINFORMATICS

MS-Office and HTML.

- Working with MS-Office Packages –One Exercise each in Word, Excel, Power Point and Access.
- Working with HTML Tags and HTML Forms. Creating HTML Pages (At least five different pages to be created using all tags learnt).
- Basic commands in MS-DOS and command line execution in LINUX.

C Programming- simple programs

Write a C program for the following

- To print the size of char, int, float and double data types.
- To convert any given temperature in a) Centigrade scale to Fahrenheit scale & b) Fahrenheit to Centigrade scale.
- To find the area and circumference of a circle with any given radius.
- To find the volume of a cylinder and cone of nay given radius and height.
- To compute the pH of a given solution for a given a) H^+ ion concentration and b) OH^- ion concentration.
- To find the average molecular weight of a double-stranded DNA of length n.
- To calculate the Body Mass Index (BMI) value.

PERL programming

- Program to store a DNA sequence.
- Program to concatenate DNA fragments.
- Program to convert DNA to RNA.
- Program to calculate reverse compliment of DNA sequence
- Program to read protein sequence data from a file
- Program to print the elements of a array
- Program to find motifs in a protein sequence.
- Program to count nucleotides in a sequence.
- Program to find the percentage of hydrophobic amino acids in a sequence.
- Program to find the percentage of G and C in a DNA sequence.
- Program to append ATGC to a DNA sequence using subroutines.
- Program to find if a DNA is stable or not.

LAB COURSE IV

LAB IN DATABASE MANAGEMENT SYSTEMS

1) Design a Database and create required tables. For e.g. Bank, College Database

2) Apply the constraints like Primary Key, Foreign key, NOT NULL to the tables.

3) Write a sql statement for implementing ALTER, UPDATE and DELETE

4) Write the queries to implement the joins

5) Write the query for implementing the following functions:

MAX ( ), MIN ( ), AVG ( ), COUNT ( )

6) Write the query to implement the concept of Integrity constrains

7) Write the query to create the views

8) Perform the queries for triggers

9) Perform the following operation for demonstrating the insertion, updation and deletion using the referential integrity constraints

10) Write the query for creating the users and their role.

SEMESTER-III

MOLECULAR MODELING AND DRUG DESIGN

UNIT I

Quantum mechanics & concepts in molecular modeling: Introduction – Coordinate systems and potential energy surfaces; introduction to quantum mechanics –Schrodinger wave equation – hydrogen molecule – Born-Oppenheimer approximation.

Unit II

Computational quantum mechanics: one electron atom, poly electronic atoms and molecules, Hartree Fock equations; calculating molecular properties using ab initio and semi empirical methods. Molecular mechanics: general features of molecular mechanics force field, bond stretching, angle bending, torsional terms, non-bonded interactions; force field parametrisation and transferability.

Unit III

Energy Minimisation and Computer Simulation: Energy Minimisation - Non-Derivative method, 1st and 2nd order minimisation methods. Computer Simulation Methods. Simple Thermodynamic Properties and Phase Space. Boundaries. Analyzing the Results of a Simulation and Estimating Errors. GROMACS and CNS.

Unit IV

Molecular Dynamics and Monte Carlo simulation: Introduction – Use of single Model time steps, multiple steps; Setting up MD, energy conservation in MD Simulation; Monte Carlo Simulation. BIOSUITE.

Unit V

Drug design: General approach to discovery of new drugs – Structure based drug design, Ligand -based drug design; Lead discovery- Lead optimization– lead modification – Physiochemical principles of drug action – Drug stereo chemistry –Drug action; Molecular Docking, AUTODOCK and HEX; SAR studies; Pharmacophore identification – QSAR, CoMFA.

Recommended Books:

- A.R.Leach. 1996. Molecular Modeling Principles and Application, 2nd edition, Longman Publications.

- D.Baxivanis and Foulette. 2001. Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins, Wiley Indian Edition.
- T K Attwood, D J parry-Smith. 2005. Introduction to Bioinformatics, Pearson Education, 1st Edition, 11th Reprint.
- J.M.Haile. 1997. Molecular Dynamics Simulation Elementary Methods, John Wiley and Sons.
- Satya Prakash Gupta. 2008. QSAR and Molecular Modeling, Springer - Anamaya Publishers.

OMICS AND SYSTEMS BIOLOGY

UNIT I

Introduction and scope of proteomics: Components of a complex mixture and protein sequencing: MALDI TOF MS, QTrap, MS/MS, 2D Gel electrophoresis and protein microarrays. qRT PCR and proteomics. Proteomic approach for clinical studies; Protein biomarker discovery and validation- Body fluid profiles, blood disease profiles, diabetes profiles, infectious diseases.

Unit II

Protein arrays: Basic principles, computational methods for identification of polypeptides, Bioinformatics-based tools for analysis of proteomics data (Tools available at ExPASy proteomics server); databases (such as Inter Pro) Protein-protein interactions; database such as STEINGS and DIP; PPI Modeling in biological systems.

Unit III

Protein complexes and aggregates: Protein interaction networks, Regulatory networks, Structures of regulatory networks, Chou-Fasman and GOR methods for protein structure prediction, Neutral network models.

Unit IV

Glycomics: The challenge and Promise of Glycomics, Identification of carbohydrates, Glycolipids, Glycoproteins, Glycan microarray and Glycan Determinants, Metaglycomes, Glycan recognition molecules, Lipidomics, Fluxomics, Biomics: systems analysis of the biome.

Unit V

Systems biology: Introduction, Integrating Networks. Computer Simulation of the whole Cell. Computer Simulation of the Cell: Human Erythrocyte Model and its Application.

Software for Modeling, E-CELL, V-CELL and GROMOS. Simulation of cellular subsystems, network of metabolites and enzymes, signal transduction networks, Gene 5 regulatory networks, metabolic pathways: databases such as KEGG, EMP, MetaCyc, AraCyc.

Recommended Books:

- Hiroaki Kitano (editor). 2001. "*Foundations of Systems Biology"*; MIT Press: ISBN 0-262-11266-3.
- Fall, CP., Marland, E. Wagner, J and Tyson, JJ (Editors). 2002. "Computational Cell Biology"; Springer Verlag: ISBN 0-387-95369-8
- Bock, G and Goode, JA (eds). 2002. "*In Silico Simulation of Biological Processes*"; Novartis Foundation Symposium 247. John Wiley & Sons: ISBN 0-470-84480-9
- Klipp, E., Herwig, R. Kowald, A.Wierling, C and Lehrach, H. 2005. "*Systems Biology in*
  Practice"; Wiley-VCH: ISBN 3-527-31078-9
- Alberghina, L. and Westerhoff, H(Editors). 2005. "*Systems Biology: Definitions and Perspectives*, Topics in Current Genetics 13"; Springer Verlag, ISBN 978-3540229681
- Szallasi, Z., Stelling, J. and V.Periwal (eds). 2006. "System Modeling in Cellular Biology: From Concepts to Nuts and Bolts (Hardcover)"; MIT Press:, ISBN 0-262-19548-8
- James M. Bower, Hamid Bolouri, 2000. Computational Modeling of Genetic and Biochemical Networks, MIT Press.
- Julio Collado-Vides (Editor), Ralf Hofestadt (Editor) 2002. Gene Regulation and Metabolism: Postgenomic Computational Approaches, MIT Press.
- Sandy B. Primrose & Richard M. Twyman, 2004. GENOMICS: Applications in Human Biology, Blackwell Publishing Ltd.,
- C.W. Sensen, 2002, Essentials of Genomics & Bioinformatics, Wiley-VCH
  Malcolm Campbell & Laurie J. Heyer Laurie J. Heyer, 2006. Discovering Genomics, Proteomics and bioinformatics
- Rodney F. Boyer 2000. Modern Experimental Biochemistry, 3rd Edition, Publisher: Prentice Hall.

BIG DATA ANALYSIS AND NEXT GENERATION SEQUENCING

UNIT I

R-Statistical package: Essentials of R-package and libraries, mathematical operations, string operations, Data structures: vectors, data framers, lists, matrices, control loop: if else, while for loops. File Input/Outpur operations. R plots and the graphics library. Overview of statistical packages and Bioconductor libraries in R. Data representation: qualitative and quantitative data types, Tabulation and visual display of data, plotting line plot, scatter plot, frequency histograms, pie-chart, heat map and 3D plots.

UNIT II

Epigenomics: Different platforms of NGS, Overview of metagenomics principles, methylation of DNA and genetics, histone modifications, ChIP-chip ChIP-seq- techniques. Impact of transcriptome on biology, volume of data produced and imported repositories.

UNIT III

Microarray data analysis: Gene expression analysis, statistical method, relative merits of various platforms, Mapping algorithm such as Burro-wheeler, Measuring gene, IncRNA, siRNA, for RNA-seq NGS data. Sequence assembly concepts and challenges in assembling short reads. Hamiltonian cycle and de Brjin.Gene prediction and annotation.

UNIT IV

Identification of genetic variants from genome sequence: SNPs, SNVs, translocation, copy number variation. Concept behind genome-aide association studies. Introduction to various applications. Concepts and algorithms to measure transcriptional regulation; methylation and alternative splicing; small RNA analysis, validation of whole-genome database.

UNIT-V

Data analysis interpretation: gene expression analysis, differential expression analysis, Allele-specific expression, prioritizing genetic variants, non-synonym variants (SIFT, Polyphen), Synonyms variants, regulatory variants, statistical methods on rare variants, statistical considerations, Hidden Markov model annotating histone markers.

Recommended Books:

- R Development Core Team 2011. R: A Language and Environment for Statistical Computing. Vienna, Austria.
- Wang J. 2012. Next Generation Microarray Bioinformatics: Methods and Protocols. Hamana publishers. NY.
- Anna Maria Paganoni. 2014. Advances in Complex Data Modeling and Computational Methods in Statistics. Springer.
- Norman Maltoff-UC Davis. 2009. The Art of R Programming
- Hadley Wickham.2013. Advanced R Programming
- Venables, WN., Smith, DM., 2008. Network Theory: An Introduction to R.

DATA WAREHOUSING AND DATA MINING

Unit-I

Overview and Concepts: Need for data warehousing, Basic elements of data warehousing, Planning and Requirements: Project planning and management, Collecting the requirements. Architecture And Infrastructure: Architectural components, Infrastructure and metadata.

Unit-II

Data Design And Data Representation: Principles of dimensional modeling, Dimensional modeling advanced topics, data extraction, transformation and loading, data quality. Information Access and Delivery: Matching information to classes of users, OLAP in data warehouse, Data warehousing and the web. Implementation and Maintenance: Physical design process, data warehouse deployment, growth and maintenance.

Unit-III

Introduction: Basics of data mining, related concepts, Data mining techniques. Data Mining Algorithms: Classification, Clustering, Association rules. Knowledge Discovery: KDD Process. Web Mining: Web Content Mining, Web Structure Mining, Web Usage mining.

Unit-IV

Advanced Topics: Spatial mining, temporal mining. Visualization : Data generalization and summarization-based characterization, Analytical characterization: analysis of attribute relevance, Mining class comparisons: Discriminating between different classes, Mining descriptive statistical measures in large databases Data Mining Primitives, Languages, and System Architectures: Data mining primitives, Query language, Designing GUI based on a data mining query language.

Unit-V

DBMS- Introduction, overview and types. Relational and transactional Database. Relational database-Introduction to relational DB, Data Definition-Manipulationcontrol-Objects, Views, sequences and Synonyms. Data Abstraction; Data Models; Instances & Schemes; E-R Model - Entity and entity sets; Relations and relationship sets; E-R diagrams; Reducing E-R Diagrams to tables. Network Data Model: Basic concepts; Hierarchical Data Model: Basic Concepts; Multimedia Databases - Basic Concepts and Applications; Indexing and Hashing; Text Databases; Introduction to Distributed Database Processing, Data Security. ORACLE and SQL- introduction and functions in DBMS.

Recommended Books:

- Dunham, M.H."Data Mining Introductory and Advanced Topics", Pearson Education.
- Kamber, H., Kaufmann, M. "Data Mining Concepts and Techniques".
- Adriaans, P., Zantinge , D."Data Mining", Pearson Education Asia.
- Berry, M and Linoff, G. "Mastering Data Mining", John Wiley.
- Inmon, W.H. "Building the Data Warehouses", Wiley Dreamtech.
- Kimball, R. "The Data Warehouse Toolkit", John Wiley.
- Silberschatz, A., Korth, H.F. and Sudarshan, S. 2002. "Database system Concepts"; Fourth Edition, McGraw Hill Publishers.
- Date, C.J. 2000. "An introduction to Database systems"; Seventh Edition, Addison Wesley Publishers.
- Xiaohua Hu and Yi Pan 2007. Knowledge Discovery In Bioinformatics, John Wiley & Sons, Inc
- Yan-Qing Zhang 2009. Machine Learning Approach In Bioinformatics, John Wiley & Sons, Inc
- Feldman, R and Sanger, J. 2007. "The Text Mining Handbook: Advanced approaches in analyzing unstructured data"; Cambridge University Press.
- Pierre Baldi & Søren Brunak 2001. Bioinformatics: The Machine Learning Approach, The MIT Press, London.

LAB-V

LAB IN DRUG DISCOVERY AND GENOMICS

- Protein structure prediction
- Primer design and analysis
- DNA fragment contig assembly (GCG, SeqMerge)
- Restriction mapping and analysis
- Protein domain and motif prediction
- Use of web resources cheminformatics studies
- Visualization of micro array informatics
- Homology modelling Validation
- Analysis of protein-protein interaction using protein interaction databases
- ADME/T property prediction
- Molecular docking: Rigid, flexible and QM/MM
- 2D and 3D QSAR along with CoMFA and CoMSIA

SEMESTER- IV

RESEARCH METHODOLOGY

UNIT I

Research: Definition, types, approaches, significance, research methods, criteria of good research, Literature collection and citation, research design: Basic principles of experimental design. Result analysis and interpretation, Report writing and manuscript preparation.

UNIT II

Bio-instrumentation: Microscopy-Light, Phase Contrast, Fluorescence, Confocal, Scanning and Transmission Electron Microscopy, Atomic Force Microscope, Scanning Tunneling Microscope - principles and applications.

UNIT III

Spectroscopy and spectrometry: Fluorescence, UV-Visible spectrophotometer, NMR and ESR Spectroscopy. Mass Spectrometry, GC-MS. Blotting Techniques-Principles and techniques of Southern, Northern and Western blotting techniques and hybridization.

UNIT IV

Algorithms: Protein and Nucleic acid sequence Algorithms: Sequence Databases, Use of the algorithms BLAST, Multiple sequence alignments and Clustering algorithms. Phylogeny: Evolutionary trees; Biological networks: Pathway analysis. Protein structure analysis: Protein structure databases; Protein structure comparison; Fold recognition; 3D-1D Profiles; Threading and Comparative structure modeling.

UNIT V

Computational applications: Introduction to internet – firebox, flock, mozilla, netscape., Clustal W, Wingene, oligo primer analysis, SPDBV, Swiss-prot, PDB, restriction analysis, primer design, data mining methods for sequence analysis, web-based tools for sequence searches, motif analysis and presentation.

Recommended Books:

1. Andrews D. Baxevanis and BF. Francis Ouellette, 2001. Bioinformatics-A Practical guide to the Analysis of Genes and Proteins, A John Wiley & Sons Publishers
2. David W.Mount, 2003. Bioinformatics-Sequence and Genome Analysis, CBS

Publishers.

3. Ian Korf, Mark Xandell & Joseph Bedell, 2003. BLAST. O'REILLY Publisher.
4. Kothari C.R, 2004. Research Methodology - Methods and Techniques, New Age International (P) Ltd.
5. Gupta S.P, 2008. Statistical Methods. Sultan and Sons Company, New Delhi.