

INTEGRAL UNIVERSITY, LUCKNOW

SYLLABUS
&
EVALUATION SCHEME

for

M.TECH. BIOINFORMATICS (Full Time)

(with effect from Session 2020-21)

(Students admitted 2020 onwards)

Integral University
M. TECH. BIOINFORMATICS
(w.e.f. Session 2020-21)
(Students admitted 2020 onwards)

1st Year

1st Semester

| S. No. | Course Category | Subject Code | Subject | Periods and Credits | | | | Evaluation Scheme | | | Subject Total | |
|--------------|-----------------|--------------|---|---------------------|----------|----------|-----------|-------------------|------------|------------|---------------|------------|
| | | | | L | T | P | C | Sessional (CA) | | (ESE) | | |
| | | | | | | | | CT | TA | | | Total |
| 1 | DC | BE-520 | Biological Databases | 3 | 1 | 0 | 4 | 40 | 20 | 60 | 40 | 100 |
| 2 | ESA | MT-503 | Applied Mathematics and Biostatistics | 3 | 1 | 0 | 4 | 40 | 20 | 60 | 40 | 100 |
| 3 | DC | BE-521 | Biochemistry and Molecular Biology | 3 | 1 | 0 | 4 | 40 | 20 | 60 | 40 | 100 |
| 4 | DC | CS-532 | Biocomputing and IPR Related Issues | 3 | 1 | 0 | 4 | 40 | 20 | 60 | 40 | 100 |
| 5 | DC | BE-523 | Biochemistry and Molecular Biology Lab | 0 | 0 | 4 | 2 | 40 | 20 | 60 | 40 | 100 |
| 6 | DC | BE-524 | Bioinformatics and Programming Language Lab | 0 | 0 | 4 | 2 | 40 | 20 | 60 | 40 | 100 |
| Total | | | | 12 | 4 | 8 | 20 | 240 | 120 | 360 | 240 | 600 |

L: Lecture **T:** Tutorial **P:** Practical **C:** Credit **CA:** Continuous Assessment

CT: Class Test **TA:** Teacher's Assessment **ESE:** End Semester Examination

DC: Departmental Core **DE:** Departmental Elective

BIOLOGICAL DATABASES
BE-520

w.e.f Session 2020-21

| Pre-requisite | Co-requisite | L | T | P | C |
|---------------|--------------|---|---|---|---|
| None | None | 3 | 1 | 0 | 4 |

Objective: The objective of this course is getting the introductory idea of the genesis of bioinformatics. The course focuses on briefing the concepts of storage, retrieval and analysis aspects of vast biological data.

| | | |
|-----------------|---|----------|
| UNIT I | Introduction to Bioinformatics | 8 |
| | The Genesis of Bioinformatics, Bioinformatics Versus Other Disciplines, Developments from Linear Information to Multidimensional Structure Organization, Mathematical and Computational Methods: Need of Mathematical Modeling, Fitting Models to Data. | |
| UNIT II | Biological Databases | 8 |
| | Biological databanks: Types, Nucleic Acid and Protein Sequence Databases, Structural Databases, Literature databases, Genome and Organism-specific databases for example OMIM, TAIR, NCBI, TIGR, GOLD. | |
| UNIT III | Data Retrieval | 8 |
| | Building biological databases, Submitting sequences to databases, Database tools for displaying and annotating sequence data, Retrieval of biological data including Entrez, SRS, ARSA and DBGET/Link DB. | |
| UNIT IV | Molecular visualization | 8 |
| | Sequence similarity, identity and homology, definitions of homologues, orthologues, paralogues. Sequence file format: GenBank, Fasta and PIR format, Structure file format: pdb and mmCIF format, Visualization of 3-D structures using PyMOL, RasMol, SPDBV, CHIME, Discovery studio visualizer. | |
| UNIT V | Analysis of Biological data | 8 |
| | Applications of Bioinformatics: Bioinformatics in Life and Environmental Sciences, Structural Biology Interface, Statistical mining of Gene and Protein Databanks. | |

References:

1. Fundamental Concepts of Bioinformatics - Dan E. Krane, Michael L. Raymer, Pearson education.
2. Sequence structure and Database – Des Higgins, Willice Taylor, oxford press
3. Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins, by Andreas D. Baxevanis, B. F. Francis Ouellette, Wiley-Interscience,
4. Sequence and Genome Analysis by David W. Mount - Cold Spring Harbor Laboratory
5. Bioinformatics and Functional Genomics;by Jonathan Pevsner;Wiley-Liss
6. Introduction to Bioinformatics; Arthur M. Lesk; Oxford University Press

APPLIED MATHEMATICS AND BIOSTATISTICS
BE-503

w.e.f Session 2020-21

| Pre-requisite | Co-requisite | L | T | P | C |
|---------------|--------------|---|---|---|---|
| None | None | 3 | 1 | 0 | 4 |

Objective: The objectives of this course are to get acquainted with the advance mathematical and statistical applications for building up the algorithms and needed analysis of complex biological data.

| | | |
|-----------------|---|----------|
| UNIT I | Differential and Integral Calculus | 8 |
| | Differential calculus Function, Limit, Continuity and Differentiability, Differentiation of standard functions, Methods of Differentiation, Maxima and Minima. Integral Calculus Indefinite integration of standard functions, Integration by substitution, by parts, by partial fractions. | |
| UNIT II | Matrices and Determinants | 8 |
| | Definition, Types of Matrix, Properties of Matrices. Addition, subtractions and multiplications, Inverse of Matrix, Determinants:- Properties of Determinants, Eigen Value and Eigen Vectors, solution of simultaneous equations by Matrix Method and Cramer's rule. | |
| UNIT III | Probability Theory | 8 |
| | Sample Space and Events, Axioms of Probability, Conditional Probability, Independent Events, Bayes Theorem. Numerical Description of Data; Discrete and Continuous variable, Mean, Median, Mode, Range Quartiles, Standard Deviation Variance, Coefficient of Variation. Measures of skewness and Kurtosis by moments | |
| UNIT IV | Correlation and Regression Analysis | 8 |
| | Scatter diagram, Karl Pearson's Correlation coefficient, Spearman's Rank Correlation Coefficient. Regression Coefficients and Regression lines. Method of least square, straight line, Parabola, exponential curve. | |
| UNIT V | Discrete and Continuous Distributions | 8 |
| | Binomial, Poisson and Normal distributions, Chi-Square test, Student's t-Test, F-test, Z-test. Analysis of variance: one way ANOVA, Two-way ANOVA. | |

References:

1. Philip Schmidt, Frank Ayres Schaum's Outline of College Mathematics. McGraw Hill, ISBN 13 9780071402279.
2. Topics in Business Mathematics and Statistics; Qazi Shoeb Ahmad, M.V. Ismail and Shadab Ahmad Khan. Laxmi Publications.
3. Wayne W. Daniel, Biostatistics, Wiley, ISBN: 978-0-471-45654-4
4. Bernard Rosner, Fundamentals of Biostatistics Thomson, Brooks/Cole ISBN: 0-534- 418201
5. Alexander Isacev-Introduction to Mathematical Method in Bioinformatics, Springer.
6. Sundar Rao & Richard-An Introduction to Biostatistics PHI.
7. Lipschutz-Theory & Problem of Probability: Schann's Outline' Series, Tata McGraw Hill.

BIOCHEMISTRY AND MOLECULAR BIOLOGY
BE-521

w.e.f Session 2020-21

| Pre-requisite | Co-requisite | L | T | P | C |
|----------------------|---------------------|----------|----------|----------|----------|
| None | None | 3 | 1 | 0 | 4 |

Objective: This course is designed to introduce the basic concepts of Biochemistry of living systems mainly dealing with biomolecules like carbohydrates, proteins, lipids, nucleic acids and their metabolic pathways that governed the entire metabolism. The course is also designed to make the students understand the concepts of DNA replication, Transcription and translation in prokaryotes and eukaryotes.

| | | |
|-----------------|--|----------|
| UNIT I | Carbohydrates Overview of metabolism, high energy compounds, the reactions of glycolysis, the anaerobic fate of pyruvate, control of glycolysis. The pentose phosphate pathway, glycogen breakdown and synthesis, control of glycogen metabolism, gluconeogenesis. Overview of citric acid cycle. Synthesis of acetyl coenzyme A, enzymes of the citric acid cycle, regulation of the citric acid cycle, reactions related to the citric acid cycle, | 8 |
| UNIT II | Lipids Lipid metabolism: fatty acid oxidation, ketone bodies, fatty acid biosynthesis, regulation of fatty acid metabolism. The mitochondrion, electron transport, oxidative phosphorylation. | 8 |
| UNIT III | Proteins and Nucleic Acids Protein degradation, amino acid deamination, the urea cycle, breakdown of amino acids, amino acid biosynthesis, nucleotide biosynthesis by de novo and salvage pathways, nucleotide degradation. | 8 |
| UNIT IV | DNA replication Roles of DnaA, Helicase, HD protein, Primase, DNA gyrase, Topoisomerase, DNA Polymerase III, DNA Polymerase I, DNA ligase, Fidelity, DNA replication in eukaryotes: Autonomously replicating sequence, Origin recognition complex, DNA polymerases α , δ , ϵ , Nucleases, DNA ligase, Telomeres | 8 |
| UNIT V | Transcription and Translation | 8 |

| | | |
|--|---|--|
| | Transcription in prokaryotes: Outline of process, Promoter, RNA polymerase; Transcription in eukaryotes: Outline of process, Promoters, Enhancers, RNA polymerase I, II, III; Post transcriptional modifications: End modifications (Addition of 5' cap and 3' Poly A tail in mRNA). Translation in prokaryotes and eukaryotes: Outline of process, Types of RNA, Structure of ribosome, Aminoacyl RNA transferase, Genetic codes, Wobble hypothesis. | |
|--|---|--|

References:

1. Voet and Voet, Biochemistry, Wiley : ISBN: 978-0-471-19350-0.
2. Nelson and Cox, Lehninger Principles of Biochemistry, W H Freeman & Co, ISBN: 978- 0- 716-77108-1.
3. Berg, Tymoczko&Stryer, Biochemistry, W.H.Freeman and Co New York. ISBN0716787245.
4. Baltimore- Molecular Biology of the Cell
5. Benjamin Levin – Genes, latest ed.
6. Albert B, Bray Denis et al.: Molecular Biology of The Cell, latest ed.
7. Watson, Hopkin, Roberts et al.: Molecular Biology of the Gene, latest ed.

BIOCOMPUTING AND IPR RELATED ISSUES
CS-532

w.e.f Session 2020-21

| Pre-requisite | Co-requisite | L | T | P | C |
|----------------------|---------------------|----------|----------|----------|----------|
| None | None | 3 | 1 | 0 | 4 |

Objective: The objective of this course is to get acquainted with C, Perl & Python language for the simulation of biological data. It also focuses on the IPR and its related issues for proper scientific conscience and ethical values among learners.

| | | |
|-----------------|--|----------|
| UNIT I | Introduction to programming languages | 8 |
| | Introduction - Flowcharts - Algorithms - Pseudocodes - Programming languages: data types, variables, constants, operators, input output, expressions, control flow constructs (conditional and loop statements) - functions, arrays, structures and unions - Pointers - Data structures - File handling. | |
| UNIT II | Procedural languages | 8 |
| | Programming in C Procedural languages - C language introduction: Variables, Data Types - Arrays (one and two dimensional arrays)- Functions: Types, Parameters, Recursion, Function prototype, Standard C library - Structures. Pointers: Introduction, Pointer with variables, Arrays and Strings, Pointers and structures, Pointers and linked list - Unions - File handling: File I/O, File opening modes - C Preprocessor - Graphical Interfaces: Dialog Boxes, Dynamic Memory Allocation. | |
| UNIT III | PERL & PYTHON | 8 |
| | Perl doc - Data types: scalar data (numbers and strings), lists, arrays, variables, operators, expressions, operators, control flow constructs (conditional and loop statements), miscellaneous control flow built in functions, associative arrays hashes, functions, Basic I/O, file handling. | |
| UNIT IV | Introduction to Web Technology | 8 |
| | Developing web pages with HTML, PHP: Tags, Links, Tables, Forms, Frames - ASP variables and data types, language syntax, controls, structures, functions, strategies and tools for handling input and generating output, error handling. | |
| UNIT V | IPR | 8 |

| | | |
|--|--|--|
| | General Introduction, Patent Claims, the Legal Decision-Making Process, Ownership of Tangible and Intellectual Property. Patentable Subject Matter, Novelty and the Public Domain, Nonobviousness. Disclosure Requirements, Collaborative Research, Competitive Research, Recent Developments in Patent System and Patentability of Biotechnological invention, IPR issues in the Indian Context. Patents in Bioinformatics. | |
|--|--|--|

References:

1. B.W. Kernighan and D. Ritchie. The C Programming Language, Prentice Hall of India. ISBN: 0131103628.
2. E. Balagurusamy. Programming in ANSI C, Tata McGrawHill Publishing Company Limited.
3. James Tisdall Beginning PERL for Bioinformatics – an introduction to perl biologist. O’Reilly publications. ISBN:978-0-596-00080-6
4. Robert W. Sebesta, Concepts of Programming Languages, Addison-Wesley

BIOCHEMISTRY AND MOLECULAR BIOLOGY LAB
BE-523

w.e.f Session 2020-21

| Pre-requisite | Co-requisite | L | T | P | C |
|---------------|--------------|---|---|---|---|
| None | None | 0 | 0 | 4 | 2 |

Objective: To understand the concept of various techniques required for quantification of carbohydrates, proteins and nucleic acids in known and unknown samples. Basic concept of separation techniques for biological samples like chromatography and electrophoresis are also included here. To learn the techniques used in molecular biology related to DNA isolation and its qualitative and quantitative analysis are also a part of this course.

- 1) Estimation of protein by Folin's-Lowry's method.
- 2) Estimation of glucose by anthrone method.
- 3) Estimation of DNA by DPA method and RNA by Orcinol method.
- 4) Determination of T_m of DNA and RNA.
- 5) Chromatography: Separation of amino acids, and sugars by TLC & paper chromatography.
- 6) Isolation and characterization of plasmid and genomic DNA from Bacteria followed by Agarose Gel Electrophoresis.
- 7) Isolation of plant DNA followed by Agarose Gel Electrophoresis.
- 8) Quantitative Estimation of genomic DNA: Determination of Absorption Spectra of genomic DNA.

References:

1. Keith Wilson John Walker John M. Walker "Principles and Techniques of Practical Biochemistry"
2. Chirikjian "Biotechnology Theory & Techniques"
3. Joseph Sambrook David W. Russell Joe Sambrook "Molecular Cloning: A Laboratory Manual"
4. William M., Ph.D. O'Leary Robert Dony Wu "Practical Handbook of Microbiology"
5. Brown, TA "Gene cloning: An introduction"

BIOINFORMATICS AND PROGRAMMING LANGUAGE LAB
BE-524

w.e.f Session 2020-21

| Pre-requisite | Co-requisite | L | T | P | C |
|---------------|--------------|---|---|---|---|
| None | None | 0 | 0 | 4 | 2 |

Objective: To obtain hands-on-training on the different tools database retrieval and analysis. It involves building the concept of programming through Python language.

1. To perform text based search of
 - a. NCBI Entrez
 - b. EMBL-EBI SRS
2. To retrieve protein sequence or nucleotide sequence of an organism in GenPept format followed by its conversion into FastA format.
3. To retrieve 3-D structure of a protein from PDB and visualize it in PyMol, RasMol or DS Visualizer.
4. To write programs in C language based on basic Biological problems.
5. Python Programming:
 - a. **Python Basics:** Python Literals, Data Types-Variable, Reading & Writing on screen.
 - b. **Conditional Statements:** *if-else, if-elif-else*, Nesting of *if-else*.
 - c. **Decision Making & Control Statements:** *for, while and do-while loops, switch* statement.
 - d. **Python Data Structure:** *List, Tuple, Dictionary & Set*.
 - e. **Function & Module:** Function declaration & Definition, Concepts of Package & module in python.

References:

1. Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins, by Andreas D. Baxevanis, B. F. Francis Ouellette, Wiley-Interscience,
2. E. Balaguruswami, "Programming with ANSI "C", Tata McGraw Hill.
3. Python course in Bioinformatics by Katja Schuerer and Catherine Letondal
4. Timothy A. Budd, —Exploring Python, Mc-Graw Hill Education (India) Private Ltd., 2015.
5. John Hunt and Ian Mackie, —“A Beginners Guide to Python 3 Programming”, Midmarsh Technology Ltd, Springer-2020.

Integral University
M. TECH. BIOINFORMATICS
(w.e.f. Session 2020-21)
(Students admitted 2020 onwards)

1st Year

2nd Semester

| S. No. | Course Category | Subject Code | Subject | Periods and Credits | | | | Evaluation Scheme | | | Subject Total | |
|--------------|-----------------|--------------|--------------------------------------|---------------------|----------|----------|-----------|-------------------|------------|------------|---------------|------------|
| | | | | L | T | P | C | Sessional (CA) | | (ESE) | | |
| | | | | | | | | CT | TA | | | Total |
| 1 | DC | BE-525 | Biomolecular Modeling and Simulation | 3 | 1 | 0 | 4 | 40 | 20 | 60 | 40 | 100 |
| 2 | DE | BE-526 | Algorithms in Molecular Biology | 3 | 1 | 0 | 4 | 40 | 20 | 60 | 40 | 100 |
| | | BE-527 | Metabolomics | | | | | | | | | |
| 3 | DC | BE-528 | Sequence Analysis and Phylogenetics | 3 | 1 | 0 | 4 | 40 | 20 | 60 | 40 | 100 |
| 4 | ESA | CS-513 | Database Management System | 3 | 1 | 0 | 4 | 40 | 20 | 60 | 40 | 100 |
| 5 | DC | BE-529 | Modeling and Phylogenetics Lab | 0 | 0 | 8 | 4 | 40 | 20 | 60 | 40 | 100 |
| Total | | | | 12 | 4 | 8 | 20 | 200 | 100 | 300 | 200 | 500 |

L: Lecture **T:** Tutorial **P:** Practical **C:** Credit **CA:** Continuous Assessment
CT: Class Test **TA:** Teacher's Assessment **ESE:** End Semester Examination
DC: Departmental Core **DE:** Departmental Elective

BIOMOLECULAR MODELING AND SIMULATION

BE-525

| Pre-requisite | Co-requisite | L | T | P | C |
|---------------|--------------|---|---|---|---|
| None | None | 3 | 1 | 0 | 4 |

Objective: The objective of the course is learning and understanding the entire concepts of in silico drug design through molecular modelling and simulation. The application of the course focuses on current trends and recent developments in the modelling of biological macromolecules.

| | | |
|-----------------|---|----------|
| UNIT I | Protein Modeling | 8 |
| | Basic concepts, Generations of secondary structure prediction methods, Chou-Fasman, and GOR algorithms, Knowledge- and Ab-initio-based methods of tertiary structure prediction, 3D model evaluations and visualization. | |
| UNIT II | Molecular Interactomics | 8 |
| | Basic principle and concepts, Rigids and flexible molecular docking, Molecular scoring functions, Analysis of molecular interactions, Common software used in molecular docking viz., AutoDock Tools, GOLD, FlexX, ICM and Dock. Implications of molecular docking concepts in Biomedical and Life Science R&D. | |
| UNIT III | Concepts of Molecular Modeling | 8 |
| | Molecular mechanics, Simulations of Free Energy changes, Force fields, Common force fields and their uses. Energy minimization, Programs and Web Servers useful in energy minimization. Implication of energy minimization in protein modeling. | |
| UNIT IV | Molecular Dynamics Simulation | 8 |
| | Basic concepts, Molecular forces stabilizing 3D structure of biological macromolecules, Simulation methods for conformational analysis, Design constraints, Recent trends, Applications, Pros, and Cons. | |
| UNIT V | Nucleic acid Structure prediction | |
| | Basic concepts, Current algorithms for RNA secondary structure prediction, Methods viz., Ab-initio, and Comparative approach, Performance evaluation of RNA prediction programs. Applications in context to Proteogenomics. | |

References:

1. Andrew R. Leach. Molecular Modelling Principles and applications. Prentice Hall.
2. Fenniri, H. "Combinatorial Chemistry – A practical approach", Oxford University Press, UK.
3. Lednicer, D. "Strategies for Organic Drug Discovery Synthesis and Design"; Wiley International Publishers.
4. Gordon, E.M. and Kerwin, J.F "Combinatorial chemistry and molecular diversity in drug discovery". Wiley-Liss Publishers.

ALGORITHMS IN MOLECULAR BIOLOGY
BE-526

w.e.f Session 2020-21

| Pre-requisite | Co-requisite | L | T | P | C |
|---------------|--------------|---|---|---|---|
| None | None | 3 | 1 | 0 | 4 |

Objective: The objective of the course is to get a detailed overview of various Universal sets of algorithms available for solving complex prediction problems. It mainly introduces with Genetic algorithm, Artificial Neural Network, Hidden markov model with Fuzzy logic as a part of search, prediction and optimization strategies.

| | | |
|-----------------|--|----------|
| UNIT I | Genetic Algorithms | 8 |
| | Genetic Algorithms: Basic concepts and applications; Encodings and optimization problems, Search Spaces as Hypercubes, Hyperplane sampling, GA Operators and Schemata; The Schema Theorem, Concept of Reduced Surrogates. | |
| UNIT II | An Executable Model of the genetic algorithm | 8 |
| | A generalized form based on Equation Generators, Generating String Losses for 1-point crossover, Generating String Gains for 1-point crossover. Other Models of Evolutionary Computation: Genitor, CHC; Hill-climbers or Hyperplane Samplers. | |
| UNIT III | Artificial Neural Networks | 8 |
| | A Simple Neuron, Firing rule, Network layers, Architectures of Artificial Neural Network: Feed-Forward networks, Feed-Back networks, Perceptrons, Pattern recognition problems, Back Propagation Algorithm, Applications of Neural Networks. | |
| UNIT IV | Hidden Markov Models | 8 |
| | Applications of HMM: finding CpG islands in long DNA sequence, Solution of the decoding problem: Viterbi Algorithm and its complexity, Posterior decoding through forward and backward algorithms, Profile HMM: ungapped profile alignment. Derivation of profile HMM for multiple sequence alignments; Alignment of the sequences to a Profile HMM. | |
| UNIT V | Fuzzy Logic | 8 |
| | An Introduction of Fuzzy Logic, Fuzzy Vs Non-Fuzzy, K-Mean Clustering, Fuzzy C-Mean Clustering, Basic tipping problems. | |

References:

1. Mount DW 'Bioinformatics: Sequence and genome analysis', CBS publications.
2. Lesk 'Introduction to Bioinformatics', Oxford publications.
3. Page G 'Algorithms in Bioinformatics'.

METABOLOMICS
BE-527

w.e.f Session 2020-21

| Pre-requisite | Co-requisite | L | T | P | C |
|---------------|--------------|---|---|---|---|
| None | None | 3 | 1 | 0 | 4 |

Objective: The objective of the course is learning and understanding the metabolic process of various Biomolecules along with the knowledge of their respective databases. It also underlines the basic Engineering and Reconstructional strategies for crucial metabolic pathways.

| | | |
|-----------------|--|----------|
| UNIT I | Metabolism | 8 |
| | Carbohydrate metabolism, Lipid metabolism, Transport metabolism; Significance of anaplerotic reactions and amphibolic pathways; Feedback control of metabolic pathways. | |
| UNIT II | Enzymes | 8 |
| | Classification of enzymes; Enzymes, Compounds and Reactions databases: LIGAND-Biochemical compounds and reactions, ENZYME–Enzymes, BRENDA-Comprehensive Enzyme Information System. | |
| UNIT III | Metabolic Pathway | 8 |
| | Classification of metabolic pathways (with respect to enzymes); Metabolic Pathway databases: KEGG, EMP, Malaria parasite metabolic pathways, EcoCyc, Boehringer Mannheim-Biochemical Pathways. | |
| UNIT IV | Engineering of metabolic pathways | 8 |
| | Mathematical representation of metabolic pathways, Generation and dynamic representation of metabolic pathways, Deriving common principles from the metabolic pathways knowledge: Deriving sets of enzymes specific for various reactions. | |
| UNIT V | Reconstruction of metabolic pathway | 8 |
| | Predicting regulatory elements; Identifying targets; Full genome annotation through knowledge of metabolic pathways. | |

References:

1. Bioinformatics A practical Guide to the Analysis of Genes and Proteins. Ed. Andreas D. Baxevanis and B. F. Francis Ouellette. John Wiley & Sons, Inc., Publications (For Micro array).
2. Charles M. Perou, 2000. Molecular portraits of human breast tumors. *Nature*, 406: 747-752.
3. Christopher H. Schilling et al. 1998. The underlying pathway structure of biochemical reaction networks. *PNAS*. 95:4193-8.
4. Christopher H. Schilling et al. 1999. Towards metabolic phenomics: Analysis of Genomics Data Using Flux Balances, *Biotechnology. Prog.* 15: 288-295.
5. Chu et al. 1998. The Transcriptional program of sporulation in budding yeast. *Science*. 282:699-705.

SEQUENCE ANALYSIS AND PHYLOGENETICS
BE-528

w.e.f Session 2020-21

| Pre-requisite | Co-requisite | L | T | P | C |
|----------------------|---------------------|----------|----------|----------|----------|
| None | None | 3 | 1 | 0 | 4 |

Objective: The objective of the course focuses on the various strategies available for Pairwise and Multiple sequence alignment studies. The course also outlines the detailed overview of phylogenetic studies as an important application of multiple sequence alignment approach.

| | | |
|-----------------|---|----------|
| UNIT I | Sequence alignment Pairwise Sequence Alignment: Hamming Distance, Dot-plot Matrix, Dynamic programming strategies and their complexities: Global alignment (Needleman and Wunsch algorithm), Global alignment in linear space: Divide and Conquer algorithm, Local alignment (Smith and Waterman algorithm), End space free alignment, Gap penalties. | 8 |
| UNIT II | Database similarity searching Heuristic approaches to Pairwise Sequence Alignment: FASTA and BLAST, PSI- BLAST, PHI- BLAST, Amino acid substitution matrices and their applications: Construction of PAM matrix, BLOSUM matrix. | 8 |
| UNIT III | Multiple sequence alignment Introduction and significance of MSA, Algorithms for MSA: Center Star algorithm, Multiple alignments with Consensus; Alignment of string to a profile; Pairwise and Iterative Multiple Alignment strategies, Common MSA tools: Clustal W, Pile Up. | 8 |
| UNIT IV | Phylogenetics Introduction to Phylogenetics and phylogenetic trees; Reconstruction of Phylogenetic trees: Distance based methods: Ultrametric property and Additive property, UPGMA and Neighbor joining methods, Jukes Cantor model, BootStrapping. | 8 |
| UNIT V | Character based methods Occam's Razor; Parsimony: Small parsimony (Fitch's Algorithm), Weighted parsimony (Sankoff's Algorithm) and Large parsimony problems. Probabilistic approaches like Maximum likelihood method, Overview of MEGA software. | 8 |

References:

1. A.D. Baxevanis et al., Current Protocols in Bioinformatics, Wiley Publishers.
2. David W. Mount Bioinformatics, Cold Spring Harbor Laboratory Press, ISBN 0-87969-608-7
3. Computational Molecular Biology by P. A. Pevzner, Prentice Hall of India Ltd, ISBN 81-203-2550-8.
4. D.E. Krane and M.L. Raymer Fundamental concepts of Bioinformatics, Pearson Education ISBN 81-297-0044-1
5. N. Gautham, Bioinformatics Narosa publications. ISBN-13: 9781842653005.

**DATABASE MANAGEMENT SYSTEM
CS-513**

w.e.f Session 2020-21

| Pre-requisite | Co-requisite | L | T | P | C |
|----------------------|---------------------|----------|----------|----------|----------|
| None | None | 3 | 1 | 0 | 4 |

Objective: The course curriculum helps to understand the designing of the database models, understanding of the relationships between different records and databases. Its major objective is the storage, manipulation and working with the transactions. It helps in the connectivity of the large databases with the various webpages. It also introduces the concept of several technologies which helps in resolving the integrity, atomicity problems. It helps in the transformation of one type to different types. It removes multiple duplicacy, allows access to multiple users and helps in data protection.

| | | |
|-----------------|--|----------|
| UNIT I | DBMS | 8 |
| | Introduction - concepts and overview - Types DBMS- Relational and transactional Database, Database planning and Design concepts:General Database Planning and Design – Documentor forms - preparation and architecture. Entity-Relational ship Model- entities, Attributes, keys, tablesdesign, relationships, roles and dependencies. Advanced E-R model. - concepts. | |
| UNIT II | Relational DB | 8 |
| | Introduction to relational DB and transactions. SQL-statements-Data Definition-Manipulation-control-Objects, - Views, sequences and Synonyms. Working with code and forms- Frontend development-query sublanguage-modifying relations in SQL. | |
| UNIT III | Internals of RDBMS | 8 |
| | Physical data structures, query optimization. Join algorithmstatisca and cost base optimization. Transaction processing. Concurrency control and recovery management. Transaction model properities, state serializability, lock base protocols, two phase locking. | |
| UNIT IV | Database technologies | 8 |
| | JDBC, ODBC standard and CORBA - extended entity relationship model, object data model UML diagram. File organizations and data structures. Distributed database environment and its overview. Different databases and internet. Use of XML. | |
| UNIT V | Datawarehouse | 8 |

| | | |
|--|---|--|
| | Descriptive Data Summarization, Data Cleaning, Data Integration and Transformation, Data Reduction, Data Discretization and Concept Hierarchy Generation, Data Warehouse and OLAP technology - A Multidimensional Data Model - From Data Warehousing to Data Mining - Efficient Methods for Data Cube Computation - Further Development of Data Cube and OLAP Technology. | |
|--|---|--|

References:

1. Abraham Silberschatz, Henry F. Korth and S. Sudhasan, Database system concepts. McGraw Hill Publications.
2. Elmasri Ramez and Novathe Shamkant, "Fundamentals of Database systems", Benjamin cummings Publishing Company. ISBN-10: 0321369572.
3. P.Ramakrishnan Rao: Database Management system, McGraw Hill Publications. 9780071230575
4. Jim Gray and A.Reuter "Transaction processing : Concepts and Techniques" Morgan Kaufmann Press. ISBN-10: 1558601902
5. V.K .Jain. Database Management system. Dreamtech Press ISBN 8177222279
6. Date C.J. "Introduction to database management" Wesley.
7. Ullman, JD "Principles of Database systems" Galgottia publication
8. James Martin Principles of Database Management systems" PHI.

MODELING AND PHYLOGENETICS LAB
BE-529

w.e.f Session 2020-21

| Pre-requisite | Co-requisite | L | T | P | C |
|----------------------|---------------------|----------|----------|----------|----------|
| None | None | 0 | 0 | 8 | 4 |

Objective: To obtain hands-on-training on the different tools for modelling and phylogenetics studies.

| List of Experiments | |
|----------------------------|---|
| | <ol style="list-style-type: none">1. To draw the 2D and 3D structures of chemical compounds in the ChemSketch software.2. To build the dipeptides, tripeptides and oligonucleotides molecules in the Argus Lab.3. To carry out the Energy minimization studies of the biological molecules using SPDB viewer by applying the GROMOS force field.4. To virtually mutate the small parts of protein molecules and carry out its Molecular Dynamics studies with the help of GROMACS.5. To perform the secondary structure prediction studies of the given protein using GOR and nnPredict tools available at ExpASy sever.6. To perform the homology based comparative protein modeling of the given protein using MODELER.7. To validate the Homology based predicted model using following tools:<ol style="list-style-type: none">a) WHATIFb) PROSAc) PROCHECKd) VERIFY 3D8. To carry out multiple sequence alignment of the given nucleotide sequences using Clustal W and T-Coffee tools and draw out the important inferences.9. To perform the phylogentic analysis of the given protein sequences of a protein family using MEGA/PHYLIP package and draw the important inferences. |

References:

1. Bioinformatics: A Practical Approach by K. Mani and N. Vijayaraj, Aparna Publications, Coimbatore
- A.D. Baxevanis et al., Current Protocols in Bioinformatics, Wiley Publishers.
2. Andrew Leach; Molecular Modelling: Principles and Applications (2nd Edition), Prentice Hall, 2001, ISBN 13: 9780582382107.
3. Barry A Bunin, Brian Siesel, Guillermo Morales, Jurgen Bajorath; Chemoinformatics: Theory, Practice, & Products, Springer Science & Business Media, 2006.
4. Wolff, M E Ed.; Burger's Medicinal Chemistry and Drug Discovery, John Wiley and Sons, 2010, New York.
5. H. Fenniri; Combinatorial Chemistry–A practical Approach, Oxford University Press, 2000, UK.

6. D. Frenkel, B. Smit; *Understanding Molecular Simulation: From Algorithms to Applications*, Elsevier 2012.
7. Stephen Misener, Stephen A. Krawetz. *Bioinformatics Methods and Protocols*, Humana Press, 1999, ISBN 978-0-89603-732-8.