M.Sc. COMPUTATIONAL BIOLOGY

REGULATIONS AND SYLLABI
(Effective from 2011-2012)

Centre for Bioinformatics
SCHOOL OF LIFE SCIENCES
PONDICHERRY UNIVERSITY
PUDUCHERRY
Eligibility for M. Sc. Computational Biology

Students from any of the below listed Bachelor degrees with minimum 55% of marks are eligible.

1. B. Sc. Bioinformatics
   B. Sc. Physics
   B. Sc. Chemistry
   B. Sc. Mathematics

2. B. Sc. Biotechnology
   B. Sc. Biochemistry
   B. Sc. Microbiology
   B. Sc. Plant Biology and Biotechnology / Botany
   B. Sc. Animal Biology and Biotechnology / Zoology

   With Mathematics at +2 level Compulsory

3. B. Tech. Bioinformatics
   B. Tech. Biotechnology
   B. Tech. Industrial Biotechnology
   B. Tech. Pharmaceutical Technology
   B. Tech. Food Technology
   B. Tech. Chemical Engineering

4. B. E. Information Technology
   B. E. Computer Science and Engineering
   B. E. Electrical and Electronics Engineering
   B. E. Electronics and Instrumentation
   B. E. Electronics and Communication Engineering
   B. E. Mechanical Engineering
   B. E. Biomedical Engineering
<table>
<thead>
<tr>
<th>Course Code</th>
<th>Course Title</th>
<th>H/S</th>
<th>Credits</th>
<th>Pg. No.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Semester I</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBIO 403</td>
<td>Probability and Statistics</td>
<td>H</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>CBIO 404</td>
<td>Communication Skills in Science &amp; Technology</td>
<td>H</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>CBIO 406</td>
<td>Programming Language – Introduction to C and PERL</td>
<td>H</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>CBIO 411</td>
<td>Cell &amp; Molecular Biology</td>
<td>H</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>CBIO 412</td>
<td>Biochemistry</td>
<td>H</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>CBIO 413</td>
<td>Analytical Methods in Biotechnology</td>
<td>H</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>CBIO 414</td>
<td>Biology</td>
<td>S</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>CBIO 415</td>
<td>Mathematics</td>
<td>S</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>Lab</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBIO 450</td>
<td>Analytical Techniques in Biotechnology Laboratory</td>
<td>H</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>CBIO 452</td>
<td>Programming Language – Introduction to C and PERL</td>
<td>H</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td><strong>Semester II</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBIO 421</td>
<td>Algorithms in Computational Biology</td>
<td>H</td>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td>CBIO 423</td>
<td>Database Management Systems</td>
<td>H</td>
<td>3</td>
<td>15</td>
</tr>
<tr>
<td>CBIO 425</td>
<td>Structural Biology</td>
<td>H</td>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td>CBIO 426</td>
<td>Biodiversity and IPR</td>
<td>S</td>
<td>2</td>
<td>17</td>
</tr>
<tr>
<td>CBIO 427</td>
<td>Biomedical Informatics</td>
<td>S</td>
<td>2</td>
<td>18</td>
</tr>
<tr>
<td>CBIO 431</td>
<td>Sequence Analysis</td>
<td>H</td>
<td>3</td>
<td>19</td>
</tr>
<tr>
<td>CBIO 432</td>
<td>Molecular Evolution</td>
<td>H</td>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td>Lab</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBIO 454</td>
<td>Database Management Systems</td>
<td>H</td>
<td>1</td>
<td>21</td>
</tr>
<tr>
<td>CBIO 455</td>
<td>Sequence Analysis</td>
<td>H</td>
<td>1</td>
<td>22</td>
</tr>
<tr>
<td><strong>Semester III</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBIO 502</td>
<td>Data Mining and Machine Learning</td>
<td>H</td>
<td>3</td>
<td>23</td>
</tr>
<tr>
<td>CBIO 503</td>
<td>Advanced Programming Language</td>
<td>H</td>
<td>3</td>
<td>24</td>
</tr>
<tr>
<td>CBIO 505</td>
<td>Genomics and Proteomics</td>
<td>H</td>
<td>3</td>
<td>25</td>
</tr>
<tr>
<td>CBIO 506</td>
<td>Systems Biology</td>
<td>S</td>
<td>2</td>
<td>26</td>
</tr>
<tr>
<td>CBIO 507</td>
<td>Immunology &amp; Pharmacology</td>
<td>H</td>
<td>3</td>
<td>27</td>
</tr>
<tr>
<td>CBIO 508</td>
<td>Molecular Modeling and Molecular Dynamics</td>
<td>H</td>
<td>3</td>
<td>28</td>
</tr>
<tr>
<td>Lab</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBIO 550</td>
<td>Molecular Modeling and Molecular Dynamics</td>
<td>H</td>
<td>1</td>
<td>29</td>
</tr>
<tr>
<td>CBIO 551</td>
<td>Advanced Programming Language</td>
<td>H</td>
<td>1</td>
<td>30</td>
</tr>
<tr>
<td><strong>Semester IV</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBIO 521</td>
<td>Project Work</td>
<td>H</td>
<td>12</td>
<td>31</td>
</tr>
</tbody>
</table>
CBIO 403 - Probability and Statistics

Total Credits: 3
Total: 36 Hrs.

Unit 1
7 Lectures
Numerical descriptive techniques: Measures of central tendency: mean, median, mode, relation between mean, median and mode. Partition values: quartiles, deciles, percentiles; Measures of dispersion: Absolute and Relative Measures, Moments, skewness and kurtosis

Unit 2
7 Lectures
Correlation and Regression: Principles of least squares, scatter diagram, correlation, covariance, correlation coefficient, properties of correlation coefficient, regression, properties of linear regression, rank correlation, multiple correlation

Unit 3
7 Lectures
Probability Theory: Concept of probability: sample space and events, independent events, mutually exclusive events. axioms of probability, conditional probability, additional and multiplication theorem of probability, Baye's theorem, Bernoulli trials, binomial distribution, normal distributions, Poisson distribution

Unit 4
7 Lectures
Sampling Theory: Meaning and objective of sampling, Sampling Error, Types of Sampling, Sampling Distribution, Sampling Distribution of Sample Mean and Sample Proportion, Standard Error

Unit 5
8 Lectures
Test of Hypothesis of Small and Large Samples: Standard Normal distribution, Chi-square distribution, Student’s t distribution, F distribution, Analysis of Variance

Text Books:
2. Schaum’s Outline series - Introduction to Probability and Statistics by Seymour Lipschutz and John Schiller. TATA McGraw-Hill edition. 199,

Reference Books:
CBIO 404 - Communication Skills for Science and Technology

Total Credits: 3

Total: 36 Hrs.

Unit 1
7 Lectures

Unit 2
5 Lectures
**Barriers to Communication:** Definition of Noise, Classification of Barriers

Unit 3
7 Lectures
**Oral/visual Communication:** Active Listening, Speech Structure, The Art of Delivery, Effective Presentation Strategies, Use of audio visual Aids, ICTs, Handling the Audience, Body Language, Conducting Meetings, Interviews, Group Discussion, Negotiation, Small Talk

Unit 4
9 Lectures
**Written Communication:** Letter, Memos and E-mails/ discussion groups, Business Letters, Memos, Reports-Informal and Formal: Characteristics of a Report, Types of Reports, The Importance of Reports, Formats, Prewriting, Structure of Reports, Writing the Report, Revising, Editing and Proofreading.
**Writing Journal Articles:** Word choice and Syntax style, Number use, References, Plagiarism

Unit 5
8 Lectures
**Technical Proposal and Thesis Writing Methodology**

Text Book:
1. Technical Communication, Principles and Practice by Meenakshi Raman, Sangeetha Sharma, Oxford University Press. 2004

Reference Books:
2. Writing for Engineers by Joan van Emden. Palgrave Macmillan. 2005
CBIO 406 - Programming Language - Introduction to C and PERL

Total Credits: 3                       Total: 36 Hrs.

Unit 1                                   7 lectures
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 2                     9 lectures
Programming in C:

Unit 3                                 7 lectures
Procedural Concept:

Unit 4                       6 lectures
Object Oriented Programming:
Programming in C++ : C++ programming – Object Oriented Concept: Encapsulation, Inheritance, Polymorphism – Different forms of Constructor – Destructor – Abstract class – Virtual function

Unit 5                        7 lectures
PERL:
Basic Perl Data Types, References, Matrices, Complex/Nested Data Structures, Scope: my, local, our – Function/Subroutines, System and User Function, File handle and File Tests – stat and lstat Functions – Perl Modules

Text Books:


Reference Books:

2. Beginning PERL for Bioinformatics by James Tisdall. O’Reilly publications. 2001
CBIO 411 - Cell and Molecular Biology

Total Credits: 3
Total: 36 Hrs.

Unit 1

Unit 2

Unit 3
Chloroplast structure and function – An overview of photosynthetic Metabolism – The absorption of light – Photosynthetic units and reaction centers – Photophosphorylation – Carbondioxide fixation and the synthesis of carbohydrates

Unit 4

Unit 5

Text Book:

Reference Books:
2. Cell and Molecular Biology by De Robertis and De Robertis. Saunders College, Philadelphia, USA. 2002
Unit 1
7 lectures

Unit 2
8 lectures
Overview of metabolism, high energy compounds, oxidation-reduction reactions, experimental approaches to the study of metabolism, the reactions of glycolysis, fermentation, the anaerobic fate of pyruvate, control of glycolysis, metabolism of hexoses other than glucose. The pentose phosphate pathway, glycogen breakdown and synthesis, control of glycogen metabolism, gluconeogenesis and other carbohydrate biosynthetic pathways.

Unit 3
7 lectures
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, protein degradation, amino acid deamination, the urea cycle, breakdown of amino acids, amino acid biosynthesis, heme biosynthesis and degradation, chemical synthesis of peptides, oligonucleotides and oligosaccharides.

Unit 4
7 lectures
Lipid digestion, adsorption and transport, fatty acid oxidation, ketone bodies, fatty acid biosynthesis, regulation of fatty acid metabolism. Lipid bilayers and membranes. Membrane transport.

Unit 5
7 lectures

Text book:

Reference Books:
CBIO 413 - Analytical Methods in Biotechnology

Total Credits: 3

Unit 1
Microscopy


Unit 2

Unit 3

Unit 4
Biochemical Techniques: Protein and DNA isolation and purification; PCR amplification, Estimation of Carbohydrates, Lipids, Proteins and Nucleic Acids.

Unit 5

Text Books:

1. Instrumental Analysis by Skoog, Holler, Crouch, 2007, Brooks/Cole,

Reference Books:

CBIO 414- Biology

Total Credits: 2                       Total: 24 Hrs.

Unit 1                                  4 Lectures
Diversity in Living World: Diversity of living organisms - Classification of the living organisms (five kingdom classification, major groups and principles of classification within each kingdom). Systematics and binomial System of nomenclature - Salient features of animal (non-chordates up to phylum level and chordates up to class level) and plant (major groups; Angiosperms up to class - linnaeus) classification.

Unit 2                                  7 Lectures
Structural Organization in Plants: Morphology, anatomy and functions of different parts of flowering plants: Root, stem, leaf, inflorescence, flower, fruit and seed.

Unit 3                                  7 Lectures
Structural Organization in Animals: Morphology, anatomy and functions of different systems of an annelid (earthworm), an insect (cockroach) and an amphibian (frog).

Unit 4                                  8 Lectures

Unit 5                                  8 Lectures

Text Books:

Reference Books:
CBIO 415 – Mathematics

Total Credits: 2

Unit I
Matrices and Linear Algebra

Matrices- Properties of Determinants, Minors and Cofactors, Multiplication of Determinants, Adjoint, Reciprocal, Symmetric Determinants, Cramer’s rule, Different types of matrices, Matrix Operations, Transpose of a matrix, Adjoint of a square matrix, Inverse of a matrix, Eigen values and eigen vector

Linear Algebra - Definition of vector space, Subspaces, Linear independence and Bases.

Unit II

Unit III
Trigonometry and Analytical Geometry: Trigonometric ratios, De Moivre’s theorem, The general equation of a Straight line, slope of a line, intercepts of a line, Angle between two lines, Intersection of two lines, The general equation of a Circle.

Unit IV

Unit 5

Text Books:

Reference Books:
1. Basic mathematics by Serge A. Lang, Springer publisher. 1988
2. A First Course in Calculus by Serge A. Lang, Springer publisher. 1986
CBIO 450 – Lab- Analytical Techniques in Biotechnology Laboratory

Total Credits: 1

Analytical Techniques

1. Optical Microscopy – Gram's Staining
3. Fluorescence Microscopy (Demonstration)
4. Protein purification using HPLC (Demonstration)
5. Thin Layer Chromatography - Separation of Chlorophyll / Amino acids
6. Interpretation of NMR, Mass spectra and FTIR data.

Cell & Molecular Biology

7. Isolation & Purification of genomic DNA from bacteria
8. Isolation & Purification of plasmid DNA
9. Restriction Digestion, Agarose gel electrophoresis of chromosomal & plasmid DNA
10. Separation of protein on SDS PAGE
11. DNA amplification using PCR
CBIO 452 - Lab - Programming Language - Introduction to C and PERL

Total Credits : 1

**LINUX Operating System:** Overview of Linux Architecture, Installation, Booting and Shutdown Process, System Processes (an overview), User Management - Types of users, Creating Users, Granting Rights, File System management

**C**
1. Working with C tokens
2. Program that illustrate operator precedence
3. Sample program for Switch – case construct
4. Sample program for looping construct
5. Program for creating user defined function
6. Program for passing pointer in a function
7. Program for String Handling (Sequence alignment, Pattern match)
8. Sorting and Binary search
9. Read and write a sequence in a file

**C++**
1. Create a class which shows the various form of constructors
2. Implement any one form of Inheritance
3. Implement static and dynamic Polymorphism

**PERL**
1. Read and Print Matrix
2. Program to find the longest sequence
3. Implementing complex data structure
4. Procedure creation example
5. Reading / Writing Protein / DNA sequences in files.
CBIO 421 - Algorithms in Computational Biology

Total Credits: 3

Unit 1

Unit 2

Unit 3
Combinatorial Pattern Matching- Hash Tables, Repeat Finding, Exact Pattern Matching; Expectation and Maximation (EM) with forward and backward algorithms, discriminative learning; Genetic Algorithm: Basic Concepts, Reproduction, Cross over, Mutation, Fitness Value, Optimization using GAs; Applications in bioinformatics

Unit 4
Hidden Markov Models: Markov processes and Markov Models, Hidden Markov Models, Parameter estimation for HMMs, Optimal model construction, Applications of HMMs
Artificial Neural Networks: Historic evolution – Perceptron, NN Architecture, supervised and unsupervised learning, Back Propagation Algorithm, Training and Testing, Self-organizing Feature Map and Radial Basis Function Network; Overview of Support Vector Machines, Bayesian network

Unit 5
Clustering and Trees: Hierarchical Clustering, k-Means Clustering, Evolutionary Trees, Distance-Based Tree Reconstruction, Reconstructing Trees from Additive Matrices, Character-Based Tree Reconstruction, Small and large Parsimony Problem.

Text Books:

Reference Books:
CBIO 423 - Database Management Systems

Total Credits: 3

Total: 36 Hrs.

Unit 1
Introduction – Database System Versus File Systems, Characteristics of Database, Database Concepts, Schemas & Instances, DBMS architecture and Data Independence, Data Models, Database Languages & Interfaces, View of Data, Database users and Administrators, Database System Structure, Database System Applications

Unit 2
Data models – ER Model: Keys, Constraints, Design Issues, Extended ER features, Reductions of ER Schema to Tables. Relational Model: Structure, Relational Algebra; Hierarchical Model, Network Model, Object Oriented Model

Unit 3
Structured Query Language – Basic Structure, Set Operations, Aggregate Functions, Null Values, Nested Sub queries, Views, Integrity: Domain constraints, Joined Relations, Data-Definition Language

Unit 4
Relational Database and Storage – Pitfalls in Relational Design Database, Functional dependencies, Decomposition Normal Forms – 1NF, 2NF, 3NF & Boyce-Codd NF, Data Storage – Ordered indices, Hashing concepts - Security and Authorization.

Unit 5

Text Books:

Reference Books:
CBIO 425 - Structural Biology

Total Credits: 3                       Total: 36 Hrs.

Unit 1                                      7 Lectures
Basic structural principles, building blocks of proteins, motifs of protein structures, alpha
domain structures, alpha/beta structures, Macromolecular crystallography-concepts

Unit 2                                      7 Lectures
DNA structures, DNA recognition in prokaryotes and eukaryotes, specific transcription factors,
enzyme catalysis and structure. Membrane proteins signal transduction, proteins of the immune
system. Structure of Spherical viruses.

Unit 3                                      7 Lectures
Folding and flexibility, Prediction, engineering and design of protein structures. Methods to
identify secondary structural elements

Unit 4                                      8 Lectures
Determination of protein structures by X-ray and NMR methods. Prediction of secondary
structure- PHD and PSI-PRED methods. Tertiary Structure: homology modeling, fold
recognition and ab-initio approaches. Structures of oligomeric proteins and study of interaction
interfaces.

Unit 5                                      7 Lectures
*In silico* study of biological structures. Structural genomics- concepts and significance.
Structural databases.

Text Books:
2. Introduction to Macromolecular Crystallography- A. McPherson., John Wiley

Reference Books:
2. Introduction to Protein Structure, Branden, Carl and Tooze. John Garland, Publication
   Inc. 1991
# CBIO 426 - Biodiversity and IPR

**Total Credits: 2**
**Total: 24 Hrs.**

## Unit 1
**6 Lectures**

## Unit 2
**5 Lectures**
Biodiversity and species extinction: impact of deforestation, hunting, wildlife trade, diseases and climate change on species extinction. Case studies on Amazonian deforestation, amphibian extinction through diseases, REDD, CITES. IUCN Red lists and other legal framework for biodiversity conservation.

## Unit 3
**4 Lectures**
Convention on biological diversity (CBD), Rio Summit, Intergovernmental Committee for the Cartagena Protocol on Biosafety (ICCP). Case study: has the CDB achieved its 2010 target.

## Unit 4
**5 Lectures**
Laws and agreements: IPR- patents, trade secrets, copyrights, trademarks - Plant genetic resources Agreement – GATT (General Agreement on Tariffs and Trade) and TRIPS (Agreement on Trade-Related Aspects of Intellectual Property Rights) Cooperation and implications - Patents of Higher plants, Transgenic organisms, Isolated genes and DNA sequences

## Unit 5
**4 Lectures**
SUI-GENESIS system and its uses - Plant variety protection and UPOV - Terminator and Traitor technologies for seed protection, uses and implications.

### Text Books:
1. Intellectual Property Rights by Deborah E. Bouchoux, Delmar Cenage Learning. 2005

### Reference Books:
3. Intellectual Property Rights on Biotechnology by Singh, KC, BCIL, New Delhi
CBIO 427 - Biomedical Informatics

Total Credits: 2  
Total: 24 Hrs.

Unit 1  
**Introduction:** Biomedical data, Clinical and life sciences standards and databases. Principles and its uses

Unit 2  
**Electronic health records (EMR)** and health Information exchanges—including information retrieval, medical decision making, evaluation and evidence. Patient monitoring systems-ethics in informatics - bayesian networks-learning and decision-data structure in algorithm design and analysis.

Unit 3  
**Networking:** TCP/IP Sockets and DNS clinical database concepts-design of the clinical information systems/Clinical Decision support systems-Sychnornization, concurrency, deadlock, full-text databases, distributed database services and architexture on one of the database any clinical database structure as one example.

Unit 4  
**Methods and Evaluation:** Sampling, appropriate use of controls, data collection including human-testing of statisticl significance, sensitivity and specificity. ROC plots. Methods and issues specific to healthcare.

Unit 5  
**Healthcare informatics:** Understanding and interaction Health organization especially academic health centers, understanding the health care environment, understanding the organization informatics- Interaction between these three units-machine learning approaches to make decision making and discovery. Human factors in clinical systems—use of machine learning to make modeling, datamining, policy design and law. Translation research and its uses and implications Evidence based medicines.

Text Books:


Reference Books:

CBIO 431 - Sequence Analysis

Total Credits: 3

Unit 1
Overview: Biological Literature Information access, storage and retrieval systems- Primary and secondary databases of genomics, transcriptomics, proteomics and metabolomics. Knowledge on freeware and commercial software. Importance of hardware and software creations.

Unit 2
Data- alignment and applications: Collecting and Storing Sequence Data: Sequence assembly; Submission of Sequences; Sequence accuracy; Sequence databases; Sequence formats; Conversion between formats; EST databases; SNP databases; Annotation and Archival.
Sequence alignment and applications: Uses: Choice to be made for alignment; Scoring matrices; Homology and related concepts; Dot Matrix methods; Dynamic programming methods for global and local alignments tools- FASTA, BLAST, statistical and Biological significance.

Unit 3
Nucleic acid sequence analysis: Reading frames; Codon Usage analysis; Translational and transcriptional signals; Splice site identification; Gene prediction methods; RNA fold analysis

Unit 4
Multiple Sequence alignment and applications: Uses; Methods available- Iterative alignment, Progressive alignment – ClustalW, T-Coffee; Profile Methods – Gribskov profile, PSI-BLAST, HMM ; Clustering and Phylogeny; Methods for Phylogeny analysis: Distance and Character based methods; Motif detection ; Protein family databases; Use of Structure based sequence alignment

Unit 5
Protein sequence analysis: Compositional analysis; Hydrophobicity profiles; Amphiphilicity detection; Moment analysis; Transmembrane prediction methods; Secondary structure prediction methods

Text Book:
1) Computational Molecular Biology by P. A. Pevzner, Prentice Hall of India Ltd. 2004

Reference Books:

CBIO 432 - Molecular Evolution

Total Credits: 3                       Total: 36 Hrs.

Unit 1 6 Lectures

Unit 2 10 Lectures
Evolutionary change by mutation, gene flow, genetic drift, natural selection and non-random mating. Role of gene duplication, transitions and transversions- chromosomal deletions and insertions, in evolution. Role of repetitive DNA, transposable elements and junk DNA in evolution. Homology of proteins and DNA in evolution

Unit 3 10 Lectures

Unit 4 6 Lectures
The concept of the Molecular Clock. Calibration. Limitation of molecular clock models. Human molecular clock: deducing evolutionary histories through mitrochondrial DNA and Y chromosome.

Unit 5 4 Lectures
Evolution of the genome: Genomic sequencing and mapping: Genome databases Human Genome Project.

Text Books:

2) Evolution (3rd Edition) by Ridley, M., Blackwell Science. 2004
Exercise in DBMS (MYSQL)

**Data Definition Language (DDL) statements:**
Creating database, Selecting database, Deleting database, Creating table, Modifying Table, Deleting table

**Data Manipulation statements:**
Inserting, updating and deleting records
Retrieving Records
Retrieving specific rows and columns
Use of MySQL operators – Arithmetic operators, Comparison Operators, Logical operators
Math functions, Aggregate functions
String operations
Limiting, Sorting and grouping query results
Handling null values
Renaming or aliasing table and column names
Using subqueries
Using Joins – joining a table to itself, joining multiple tables
Use of Indexes
Security Management
Granting and Revoking rights on tables
CBIO 455 - Lab - Sequence Analysis

Total Credits : 1

1. Accessing Biological databases:
2. Retrieving protein and nucleic acid sequences, structures, EST sequences, SNP data using database browsers and genome browsers.
3. Converting sequences between different formats. Using sequence editors. sequence assembly.
5. Sequence alignment and applications: pairwise alignment-dot matrix comparisons, global and local alignment, Database searching-different pairwise methods. Use of scoring matrices and gap penalties.
7. Protein Sequence analysis:
   Composition, Hydrophobicity and amphilcity. Predictions: transmembrane and secondary.
CBIO 502 - Data Mining and Machine Learning

Unit 1
Introduction 7 lectures
Introduction, Importance of Data Mining, Relational Databases, Data Warehouses, Transactional Databases, Advance Database Systems and Applications, Data Mining Functionalities, Classification of Data Mining Systems, Major issues in Data Mining.

Unit 2
Primitives and System Architectures 7 lectures
Data Mining Primitives, Data Mining Query Language, Designing Graphical User, Interfaces Based on a Data Mining Query Language, Architectures of Data Mining Systems.

Unit 3
Concept Description and Association Rules 7 lectures
Concept Description, Characterization and comparison, Data Generalization and Summarization-Based Characterization, Analytical Characterization, Mining Class Comparisons, Mining Association Rules in Large Databases, Association Rule Mining, Mining Single-Dimensional Boolean Association Rules from Transactional Databases.

Unit 4
Classification and Prediction 7 lectures
Classification and Prediction, Issues: Data preparation for classification and Prediction, Comparing classification Methods, Classification by Decision Tree Induction: Decision Trees and Decision Tress induction.

Unit 5
Clustering Methods 8 lectures
Clustering Analysis, Types data in clustering analysis: Scaled variable, Binary variables, Variables of Mixed Types, Partitioning Methods: K-means and K-Medoids, Model-Based Methods, Data Mining Applications: Data mining for Biomedical and DNA Data Analysis.

Text Books:

Reference Books:
2. Data Mining: Practical machine learning tools Techniques with java implementation by Ian H.Witten, Eibe Frank, 2005.
## CBIO 503 - Advanced Programming Language

**Total Credits: 3**

**Total: 36 Hrs.**

### Unit 1

**Perl & Bioperl:** Uses of Regular expressions: Patterns, Single-character Patterns, Grouping Patterns (Sequence, Multipliers, Parantheses as Memory, Alternation) Anchoring patterns, Precedence, Matching operators, Ignoring case, Different Delimiter, Split and Join functions—Bioperl: Installation, architecture and uses.

### Unit 2

**Object Oriented Language II:**

**Java Basics** - Importance and features of java, Modifiers, Access Controls, Data types, Expressions, Declarations, Statements & Control Structures, Program Structures, String handling, Packages, Interfaces, Working with java util Package, Garbage Collection

### Unit 3

**Exception Handling, I/O & JDBC** – Exception Handling: built in exception, creating your own exceptions, Input Stream & Output Stream: Streams, Byte and Character stream, Predefined streams, Reading and Writing from Console and Files, Buffered Reader & Writer, Serialization, Database: JDBC Basics

### Unit 4

**Multithreading and Communication** – Java Thread Model: Priorities, Synchronization, Messaging, Life Cycle of Thread, Thread class, Runnable interface, Interthread Communication, Suspending, Resuming and Stopping threads, Multithreading, Synchronization, Scheduling and Priority of Threads.

### Unit 5

**HTML:** Introduction – Formatting tags for creating a web page

**AWT & Event Handling in java** – Creating user interface with AWT - Applets, Applet Life Cycle, Simple Graphics, Fonts and Colors, Events, Listeners, Components, Containers, Working with Layouts, Image Processing, Delegation Event Model, Event Classes, Event Listener Interfaces, Adapter and Inner Classes

### Text Books:

1. Advanced Perl Programming by Sriram Srinivasan, O-Reilly, 1997

### Reference Books:

CBIO 505 - Genomics and Proteomics

Total Credits: 3                            Total: 36 Hrs.

Unit 1

Genomics and Metagenomics: Genome databases of Plants, animals and pathogens; Gene networks: basic concepts on identification of disease genes, drought stress response genes, insect resistant genes and nutrition enhancing genes.

Unit 2

Epigenetics: DNA microarray: database and basic tools, Gene Expression Omnibus (GEO), ArrayExpress, SAGE databases: understanding of microarray data, normalizing microarray data, detecting differential gene expression, correlation of gene expression data to biological process and computational analysis tools (especially clustering approaches)

Unit 3

Comparative genomics: Basic concepts and applications, whole genome alignments: understanding the significance; Artemis, BLAST2, MegaBlast algorithms, PipMaker, AVID, Vista, MUMmer, applications of suffix tree in comparative genomics, synteny and gene order comparisons Comparative genomics databases: COG, VOG

Unit 4

Functional genomics: Application of sequence based and structure-based approaches to assignment of gene functions – e.g. sequence comparison, structure analysis (especially active sites, binding sites) and comparison, pattern identification, etc. Use of various derived databases in function assignment, use of SNPs for identification of genetic traits. Gene/Protein function prediction using Machine learning tools viz. Neural network, SVM etc

Unit 5

Proteomics: Protein arrays: basic principles, bioinformatics-based tools for analysis of proteomics data (Tools available at ExPASy Proteomics server); databases (such as InterPro) and analysis tools. Protein-protein interactions: databases such as DIP, PPI server and tools for analysis of protein-protein interactions

Text Books:

Reference Books:
Total Credits: 2                       Total: 24 Hrs.

Unit 1                        5 lectures

Unit 2                        5 lectures

Networks and Motifs: Gene Networks: basic concepts, computational models. Lambda receptor and lac operon as an example. – all types of networks and its uses.

Unit 3                        5 lectures

Robustness and optimality in Biology: model and integral feedback-signaling/bifunctional enzymes. Perfect robustness- Role and its measurement. Linking models and measurement, concepts, calibration and identification, data Vs metadata

Unit 4                        4 lectures

Unit 5                        5 lectures
Synthetic Biology: Introduction, definition and Basics, Synthetic Oligonucleotide/DNA-based, RNA-based, Peptide-based and polyketide Technologies and Applications, Technologies and Applications of Directed Evolution and Microbial Engineering, Potential Hazards of Synthetic Biology

Text Books:
2. Synthetic Biology, A New Paradigm for Biological Discovery, a report by Beachhead Consulting, 2006

Reference Books:
3. Systems Biology and Synthetic Biology by Pengcheng Fu, Sven Panke, Wiley InterScience. 2009
CBIO 507 - Immunology & Pharmacology

Total Credits: 3
Total: 36 Hrs.

Unit 1
Introduction and Antibodies: Innate and acquired immunity, active and passive immunity, natural and artificial immunity and humoral. Lymphoid system- primary or secondary organ. Cells- Lymphocytes, mononuclear, phagocytes, antigen presenting, polymorphs, mast cells, cluster designation (CD) and antigen specific receptors – Principles and its uses.

Unit 2
Antibody generation: structure and function –clonal selection theory-different types of immunoglobulins, effectors, receptors and antibody diversity. complement system-activation, pathways and biological effects. Major Histochemical molecules/peptide complexes-Structure and Function and production of MHC Locus in Mice and Human. t-lymphocytes and cytokine network, receptors, production from TH1 and TH2 CD4+ T- cells.

Unit 3
Antigen and antibody reaction/interaction: Haemagglutination, direct and indirect immunofluorescence, hybridoma technology for mass production.
Vaccine design and development: Reverse vaccinology and, immunoinformatics, databases in immunology, prediction methods-B-cell and T-cell resources to study antibodies. DNA, Plant and protein based recombinant antigens as vaccines.

Unit 4
Introduction to Pharmacology: Introduction to the principles of pharmacokinetics and pharmacodynamics. ADMET (Drug Absorption, Metabolism, distribution & toxicity of drugs) and bioavailability of drugs. General pathways of drug metabolism (Phase I and phase II metabolism), toxicity due to drug-drug interactions with one specific example.
Receptor theory and mechanism of drug action (Pharmacodynamics): How drugs work, receptor occupancy theory and characterization of drug action, including dose-response relationships, agonists and antagonists.
Autonomic nervous system: Outline of autonomic nervous system, Receptor systems, second messengers and location/specificity of action of alpha and beta receptor systems in the autonomic nervous system. Mechanism of action of antihypertensive drug – beta adrenergic antagonist with one example.

Unit 5
Chemotherapy: Antibiotics - antibacterial – antiviral and anticancer - drug types and mechanism of action with one example each.

Text Books:
2. Rang and Dale’s Pharamcology ed Churchill Livingstone, 2007,

Reference books:
1. Text book of Immunology by Riott, 2006
# CBIO 508 - Molecular Modeling and Molecular Dynamics

**Total Credits: 3**

**Total: 36 Hrs.**

## Unit 1

### Molecular Mechanics


### Unit 2


## Unit 3

**Molecular Dynamics Simulation:** Introduction, Radial distribution functions, Pair Correlation function, Newtonian dynamics, Integrators- Leapfrog and Verlet algorithm, Potential truncation and shifted-force potentials, Implicit and explicit Solvation models, Periodic boundary conditions, Temperature and pressure control in molecular dynamics simulations

## Unit 4

**Drug design:** Drug discovery process. Target identification and validation, lead optimization and validation. Methods and Tools in Computer-aided molecular Design, Analog Based drug design: Pharmacophores (3D database searching, conformation searches, deriving and using 3D Pharmacophore, constrained systematic search, Genetic Algorithm, clique detection techniques, maximum likelihood method) and QSAR. Structure based drug design: Docking, De Novo Drug Design (Fragment Placements, Connection Methods, Sequential Grow), Virtual screening.

## Unit 5

**Structure Activity Relationship:** Introduction to QSAR, QSPR, Various Descriptors used in QSARs: Electronics; Topology; Quantum Chemical based Descriptors. Regression Analysis, The Significance and Validity of QSAR Regression Equations, Partial Least Squares (PLS) Analysis, Multi Linear Regression Analysis. Use of Genetic Algorithms, Neural Networks and Principle Components Analysis in the QSAR equations.

## Text Books:


## Reference:

1. Molecular Modelling for Beginners, (2nd Edition) by Alan Hinchliffe, John Wiley & Sons Ltd. 2008
3. Computational medicinal chemistry for drug discovery edited by Patrick Bultinek, Marcel Dekker Inc. 2004
CBIO 550 - LAB: Molecular Modeling and Molecular Dynamics

Total Credits : 1

Exercises

1. Molecular Visualization Softwares: Pymol and Rasmol
2. Geometry Optimization
3. Tutorial on Molecular Dynamics: Gromacs
4. Binding Site Identification
5. Structure based Drug Design:- Molecular Docking
6. Ligand based Drug Design:- QSAR
CBIO 551 - Lab: Advanced Programming Language

Total Credits : 1

1. Reading/Writing Protein/DNA sequences in files.
2. Mutation and randomization in Bioperl.
3. DNA manipulation: Transcription DNA to RNA, Reverse complementing.
4. Passing Data to Subroutines
5. Local and Global alignment of sequences
7. Simple Animation and Threads.
8. Creating simple JAVA graphical user interface
CBIO 521 - Project

Total Credits: 12

The course is designed to result in the satisfactory completion and defense of the Masters dissertation.

This process includes

a) the conceptualization of the independent research that will comprise the dissertation,
b) the preparation of and satisfactory defense of the dissertation proposal,
c) the collection, analysis, and interpretation of data,
d) presentation of findings in the dissertation format, and
e) oral defense of the dissertation.

Dissertation activity must be completed within prescribed time frame for the semester.