

# Post Graduate Diploma In Bioinformatics

## Assam University:: Silchar

### Detailed Course Structure and Syllabus

#### The Programme

The Post Graduate Diploma in Bioinformatics is of one year duration. The Programme has been organized in two semesters, First semester and second semester.

### COURSE STRUCTURE

Sl. No.	Name of the modules	Total marks	Pass Marks
<b>Semester -I</b>			
1.	PGDBI-T1: Fundamentals of Bioinformatics	70	30
2.	PGDBI-T2: Internet, DBMS and Fundamentals of Programming	70	30
3.	PGDBI-T3: Biostatistics and Basic Mathematics	70	30
4.	PGDBI-P1: Practical	30	10
	Sub-Total (A)	240	100
<b>Semester-II</b>			
5.	PGDBI-T4: Programming with C, JAVA and PERL	70	30
6.	PGDBI-T5: Omics and Informatics	70	30
7.	PGDBI-P2: Practical	30	10
8.	PGDBI-Project: Project/Dissertation work	90	30
	Sub-Total (B)	260	100
	<b>TOTAL (A+B)</b>	<b>500</b>	<b>200</b>

## **PGDBI-T1: Fundamentals of Bioinformatics**

### **Unit I Biomolecules**

Introduction to Biomolecules, Structure of atoms, chemical bonds, Chemical properties of water.

Concept of free energy, entropy, enthalpy and basic Biological reactions, Laws of Thermodynamics.

Protein Structure: Primary, Secondary, Super Secondary, Domains, Tertiary, Quaternary. Ramachandran plot.

Enzyme-substrate interactions, Enzyme inhibitions.

### **Unit II Introduction to Bioinformatics and Biological Database**

What is bioinformatics and its relation with molecular biology. Examples of related tools (FASTA, BLAST, BLAT, RASMOL), databases (GENBANK, PubMed, PDB) and software (RASMOL, Ligand Explorer).

Data generation; Generation of large scale molecular biology data (Genome sequencing, Protein sequencing, Gel electrophoresis)

General Introduction of Biological Databases; Nucleic acid databases (NCBI, DDBJ, and EMBL). Protein databases (Primary, Composite, and Secondary). Specialized Genome databases: (SGD, TIGR, and ACeDB). Structure databases (CATH, SCOP, and PDBsum)

File Format (Genbank, DDBJ, FASTA, PDB, SwissProt).

### **Unit III Sequence Alignments and Visualization**

Introduction to Sequences, alignments and Dynamic Programming; Local alignment and Global alignment (algorithm and example),

Pairwise alignment (BLAST and FASTA Algorithm) and multiple sequence alignment (Clustal W algorithm).

Methods for presenting large quantities of biological data: sequence viewers (Artemis, SeqVISTA), 3D structure viewers (Rasmol, SPDBv, Chime, Cn3D, PyMol), Anatomical visualization.

### **Unit IV Computer Aided Drug Designing (CADD)**

Introduction: Principles of drug Development, Bioinformatics in drug development.

Types of Computer Aided Drug Designing

Drug Likeness and ADME/Tox screening, Drug targets

Protein Structure Prediction - Homology Modeling, Fold Recognition, Ab- initio Method, Protein folding.

Molecular Docking, Molecular Dynamics Simulation.

Combinatorial libraries & their design, Quantitative Structure Activity Relationship

### **Suggested Readings:**

- a) David W Mount, Bioinformatics: Sequence And Genome Analysis, Cold Spring Harbor Press.
- b) Attwood, Introduction to Bioinformatics

## **PGDBI-T2: Internet, DBMS and Fundamentals of Programming**

### **Unit-1: Concept of Internet and HTML**

Basic of Computer networks; LAN, MAN, WAN; Concept of Internet; Applications of Internet; What is ISP; Basics of internet connectivity.

World Wide Web; Web Browsing software, Search Engines; Understanding URL; Domain name; IP Address.

HTML: Designing webpages using HTML, HTML Tags, Attributes, Elements etc.

### **Unit-2: DBMS**

Introduction; Characteristics of Database approach; Advantages of using DBMS approach; Data models, schemas and instances; Three-schema architecture and data independence.

Database languages (DDL and DML) and interfaces; Database system environment; Centralized and client-server architectures; Classification of Database Management systems.

ENTITY-RELATIONSHIP MODEL: Entity Types, Entity Sets, Attributes and Keys; Relationship types, Relationship Sets, Refining the ER Design; ER Diagrams

### **Unit-3: Working with MySQL**

MySQL Data Definition and Data Types; queries in MySQL; Database programming issues and techniques.

Informal Design, Guidelines for Relation Schemas; Functional Dependencies; Normal Forms.

Binary Relational Operations: JOIN and DIVISION; Additional Relational Operations.

### **Unit-4: Working with PHP**

Introduction to Php; How php better than other language; How php works with the web server; Php tags. Web architecture: MVC model, API, URL

Basic Development concepts: Php Syntax; Datatype; Operator; Variable; If else statement. Switch case.

Array: Array Structure; Type of array; Single and multi-dimensional array; Creating and deleting file; Reading and writing file; Creating and deleting folder.

### **Suggested Readings:**

- a) Korth, DBMS
- b) Gilmore, Beginning PHP and MySQL 5: From Novice to Professional
- c) Thomas A. Powell, HTML and CSS: The Complete Reference.
- d) Anurag Gupta, Fundamentals of Internet Applications.

## **PGDBI-T3: Biostatistics and Basic Mathematics**

### **Unit - 1: Fundamentals of Biostatistics**

Introduction: Nature of Statistics, Uses of Statistics, Statistics in relation to other disciplines, Abuses of Statistics.

Measures of central tendency, Measures of dispersion.

Types of Data: Concepts of population and sample, quantitative and qualitative data, discrete and continuous data.

### **Unit-2: Correlation, Regression and Sampling**

Correlation analysis – scatter diagram, correlation coefficient and its properties, Correlation ratio, Correlation Index, Spearman's rank correlation.

Concept of Regression and bivariate regression equations.

Principles of least squares and its applications.

Sampling Distributions: Introduction: Concepts of Random Sampling, Statistics and Sampling

### **Unit-3: Hypothesis, Errors and Test of Significance**

Elements of Hypothesis Testing : Null and Alternative hypotheses, Critical Region, Type I and Type II Errors, Level of Significance and Size, p-value.

Distributions of Statistics: Illustrations using different distributions, reproductive properties of the distributions.

Tests of significance: chi-square test, F test and t-test.

### **Unit-4: Basic Mathematics**

Matrix; Determinants and their application;

Complex numbers; Binomial theorem; Coefficient

Bayes' theorem and probability

### **Suggested Readings:**

- a) Zar, J.H. Biostatistical Analysis, Pearson Education.
- b) Chakraborty, A class book of Mathematics, S. Chand & Comp.
- c) Gupta, SP. Statistical Methods, S. Chand & Comp.

## **PGDBI-T4: Programming with C, Java and PERL**

### **Unit-1: Fundamentals of C**

The C Language and its Advantages; The Structure of a C Program; Writing C Programs; Building an Executable Version of a C Program, Data Types, Operands, Operators, and Arithmetic Expressions

Input/Output Management, The getchar() and putchar() Functions and Single-character I/O, Formatted Input and the scanf() Function, The Control-flow Program Statements, Looping Statements, The Data-checking Process

Arrays, Command-line arguments, File Input and Output, Combining Command-line Arguments and File I/O

### **Unit-2: Working with PERL**

Introduction to PERL; variables; Array; Initialization and manipulation.

Arithmetic and logical operators; Conditional statement and Loops; Regular expressions; Function and subroutines

Application of PERL in Bioinformatics; concatenating DNA fragments; DNA to RNA; Reading protein Files; Finding motifs; DNA to protein; Reading sequence form file.

PERL Program for mean, median, mode, variance and coefficient of variance.

PERL script for ANOVA and ANCOVA.

### **Unit-3: OOPs and Java Programming**

OOPs Concept; Class and object; Properties of object; Encapsulation and abstraction; Inheritance; Polymorphism; Abstract class.

Java: tokens and keywords; Data type; Declaring objects; garbage collection; Function overriding; Recursion; Try & Catch; Package & Interface

Java Thread Model; Thread priorities; JDBC, its type and application;

### **Unit-4: Beginners in Linux**

Introduction to Linux, Linux Packages, Installing and Uninstalling programmes, Basic commands, Introduction to Shell, Configuring Gromacs and Modeller using Linux. Common errors and Troubleshooting.

#### **Suggested Readings:**

- a) Balaguruswamy, Java Programming.
- b) Yashwant Kanetkar, Let us C.
- c) James D Tisdall, Beginning PERL for Bioinformatics

## **PGDBI-T5: Omics and Informatics**

### **Unit- I Genomics**

Human Genome Project, cloning vectors, concept of maps, shotgun libraries, DNA polymorphism. Recombinant DNA technology, restriction enzymes, resource for restriction enzyme (REBASE).

Polymerase chain reaction, primer selection for PCR, BLASTn, application of BioEdit. Genome information and special features, coding sequences (CDS), untranslated regions (UTR's), cDNA library, expressed sequence tags (EST). Codon-bias detection, detecting functional sites in the DNA. Detection of functional sites, gene expression.

### **Unit-II DNA microarray and Protein Microarray**

Microarray: Concept of microarrays; spotted arrays, oligonucleotide arrays, designing the experiment Tools for microarray analysis; soft-finder, xCluster, MADAM, SAGE, Computational analysis of Microarray data, Applications of microarray technology.

Protein Microarray in protein expression, profiling and diagnostics, drug target discovery.

### **Unit- III Proteomics and Phylogenetic analysis**

Proteomics: Protein sequence information, composition and properties, physicochemical properties based on sequence, sequence comparison, Tools and techniques in proteomics; 2-D gel electrophoresis, PAGE, isoelectric focusing, affinity chromatography, HPLC, Mass spectroscopy for protein analysis, MALDI-TOF, Electrospray ionization (ESI), Tandem mass spectroscopy (MS/MS) analysis; peptide fingerprinting (PMF).

Phylogenetic analysis: Evolution, elements of phylogeny, methods of phylogenetic analysis, Phylogenetic tree of life, comparison of genetic sequence of organisms, phylogenetic analysis tools

### **Unit- IV Applications of Bioinformatics**

Bioinformatics in Environment, biotechnology, molecular biology, neurobiology, agriculture, drug designing, biomedical genome medicines, medical microbiology.

Chemoinformatics: Introduction to cheminformatics, Use and Prospects of cheminformatics. Prodrugs and soft drugs, Drug solubility, Natural resources of lead compounds, Pharmacokinetics & drug metabolism. Biological testing and bioassays, Preclinical testing and clinical trial.

### **Suggested Readings:**

- a) A. Malcolm Campbell, Laurie J. Heyer. Discovering Genomics, Proteomics and Bioinformatics.
- b) R. Guha and A. Bender, Computational Approaches in Cheminformatics and Bioinformatics

### **PGDBI-P1: Practical**

Retrieval of DNA/RNA/protein sequences; Biological sequence editing and analysis; PCR primer designing; Plasmid vector designing; Retrieving, visualization and analysis of macromolecular structure.

Sequence alignment: database searches (BLAST, FASTA etc.), PSA and MSA; Draw and analysis of chemical structures; Submission of nucleotide sequence using BankIT; Gene finding and analyses of Genomes;

Webpage design using HTML; Database preparation in SQL; SQL commands; operating PHP, WAMP, LAMP etc;

### **PGDBI-P2: Practical**

Basic C programming (use of operator and loops); Use of Array and structures.

Basic Java Programming; Creating and operating objects and classes

Basic PERL programming; Designing PERL program for analysis of nucleotide and protein sequences;

### **PGDBI-Project: Project/Dissertation**

The students will be assigned with a practical project in the first semester which they will have to complete by the end of second semester under the supervision of a teacher concerned to the following themes.

***1. Computer Aided Drug Designing***

***2. Proteomic analyses***

***3. Database Management System***

***4. Developing Bioinformatics tools***

The project report should be submitted to the centre and should comply with the guidelines given in the student corner of the webpage.