

# SYLLABUS

POST GRADUATE DIPLOMA IN BIOINFORMATICS



ENLIGHTENMENT TO PERFECTION  
2018-19

University of North Bengal  
Raja Rammohunpur, Siliguri  
West Bengal-734013, India

# OUTLINE OF COURSE STRUCTURE

## Post Graduate Diploma in Bioinformatics (PGDBI)

(Session: 2018-19)

### SEMESTER-I

Course Code	Name of the module	Marks	Credit	Hours/Week
<b>Course (Theoretical)</b>				
DBI-101	Fundamentals of Cell Biology, Genetics and Molecular Biology	40	2	4
DBI-102	Introduction to Basic Bioinformatics	40	2	4
DBI-103	Basic Computer Application and Programming	40	2	4
<b>Course (Practical)</b>				
DBI-104	Basic techniques in Cytology and Molecular Biology	40	2	5
DBI-105	Bioinformatics methods and application	40	2	5
DBI-106	Bioinformatics Programming	40	2	5
<b>Continuing Evaluation</b>				
DBI-107	Seminar	15	2	2
DBI-108	Class test	45	2	1
<b>Total Marks &amp; Credits in the semester I</b>		<b>300</b>	<b>16</b>	<b>30</b>

### Semester-II

Course Code	Name of the module	Marks	Credit	Hours/Week
<b>Course (Theoretical)</b>				
DBI-201	Genomics and Proteomics	40	2	4
DBI-202	Structural bioinformatics	40	2	4
DBI-203	Biostatistics and Database Management Systems	40	2	4
<b>Course (Practical)</b>				
DBI-204	Genomics & Proteomics	40	2	5
DBI-205	Structural bioinformatics	40	2	5
DBI-206	Biostatics & Database Creation	40	2	5
<b>Continuing Evaluation</b>				
DBI-207	Educational trip/submission of report	15	2	1
DBI-208	Class test	15	1	1
DBI-209	Dissertation/Project Work	30	1	1
<b>Total Marks &amp; Credits in the semester II</b>		<b>300</b>	<b>16</b>	<b>30</b>
<b>Grand Total</b>		<b>600</b>	<b>32</b>	<b>60</b>

# Detailed Syllabus

## Semester-I

Course Code	Name of the module	Marks
<b>DBI-101</b>	<b>Fundamentals of Cell Biology, Genetics and Molecular Biology</b>	<b>40</b>
<i>Unit-1</i>	<ol style="list-style-type: none"><li>1. Basics of cell biology: Structure and Functions of Cellular Organelles, Cell Cycle, Chromosome Structure, Cellular Metabolic Pathways</li><li>2. Molecular Biology: Structure of Nucleic acids, Structure of a Gene, Central Dogma, Concepts of Genetic code and ORF's, DNA Replication, Chromosomal Aberration, Mutation and its implications; Transcription, translation and post-translational modifications; protein degradation</li><li>3. Genome Sequencing Techniques, Polymerase Chain Reaction, Chromatographic Techniques, DNA Fingerprinting</li></ol>	
<i>Unit-2</i>	<ol style="list-style-type: none"><li>1. Mendel's principles of inheritance and extension of Mendel's law, chromosome theory of heredity; evolving concept of gene; linkage, crossing over and recombination at the molecular level, chromosome mapping; molecular basis of chromosome pairing in mitosis-meiosis; cytological variations-an overview.</li><li>2. Genetic and physical mapping of a genome: Recombination and genetic marker viz. microsatellite / minisatellite markers, RFLP, RAPD, SNPs, FISH &amp; Chromosome painting, clone counting, physical mapping, sequencing and annotation. CpG islands, isochore and gene densities.</li><li>3. Gene amplification and PCR: Basic principles and methodologies of PCR, design of PCR primers, RT-PCR and Real-Time PCR and their utility.</li></ol>	
<b>DBI-102</b>	<b>Introduction to Basic Bioinformatics</b>	<b>40</b>
<i>Unit-1</i>	<ol style="list-style-type: none"><li>1. Bioinformatics – an Overview, Definition and History. Information Networks – Internet in</li><li>2. Bioinformatics, Evolution of Bioinformatics – Scope – Potentials of Bioinformatics, Human Genome Project</li><li>3. Introduction to Biological Databases: NCBI, EMBL, PIR, SWISS-Prot, PubChem</li><li>4. Compound, KEGG-Pathway, ChEMBL, BindingDB. Analysis of Three Dimensional Structures of Proteins, RCSB-PDB. Primary and Secondary database</li><li>5. Various file formats for bio-molecular sequences: genbank, fasta, gcg, msf, nbrf-pir etc.</li><li>6. Basic concepts of sequence similarity, identity and homology, definitions of homologues, orthologues, paralogues.</li><li>7. Scoring matrices: basic concept of a scoring matrix, PAM and BLOSUM series.</li></ol>	
<i>Unit-2</i>	<ol style="list-style-type: none"><li>1. Pairwise and Multiple sequence alignments: basic concepts of sequence alignment, Use of pairwise alignments and Multiple sequence alignment for analysis of Nucleic acid and protein sequences and interpretation of results.</li><li>2. Sequence-based Database Searches: what are sequence-based database searches, BLAST and FASTA algorithms, various versions of basic BLAST and FASTA</li><li>3. Phylogeny: Phylogenetic analysis, Definition and description of phylogenetic trees and various types of trees, Method of construction of Phylogenetic trees [distance based method (UPGMA, NJ), Maximum Parsimony and Maximum Likelihood method]</li><li>4. Codon usage analysis and its implications in modern biology</li></ol>	
<b>DBI-103</b>	<b>Basic Computer Application and Programming</b>	<b>40</b>
<i>Unit-1</i>	<ol style="list-style-type: none"><li>1. Fundamentals of Computer, Basic Applications of Computer; Components of Computer System, Central Processing Unit (CPU), input/output Devices, Computer Memory, Concepts of Hardware and Software; Concept of</li></ol>	

- Computing, Data and Information
2. Basics of Operating System; Popular Operating Systems (Windows, Linux, DOS);
  3. 3. Data structure and its relevance to biological science

- Unit-2*
1. Communication using the Internet: Basic of Computer networks; LAN, MAN, WAN;
  2. Concept of Internet; WWW and Web Browsers; Search Engines; Understanding URL
  3. 2. Design & Structure of biological databases

- Unit-3*
1. Introduction to PERL as scripting language; variables; Array; Initialization and manipulation
  2. Arithmetic and logical operators; Conditional statement and Loops; Regular Expressions; Function and subroutines
  3. Application of PERL in Bioinformatics; concatenating DNA fragments; DNA to RNA; Reading protein Files; Finding motifs; ORFs; DNA to protein

- Unit-4*
1. Use of R-Programming for statistical analysis, data structure, regex, functions, loops and if statements

**DBI-104** **Basic techniques in Cytology and Molecular Biology (Practical)** **40**

- ✓ Orcein and Feulgen staining of metaphase plates; preparation of karyotype and idiogram.
- ✓ Study of Meiotic chromosome complements.
- ✓ Restriction digestion and electrophoresis
- ✓ Isolation of genomic and plasmid DNA.
- ✓ PCR and RAPD analysis.
- ✓ PCR primer designing; Plasmid vector designing;
- ✓ Probability and chi-square test for genetic analyses for goodness of fit.

**DBI-105** **Bioinformatics methods and application (Practical)** **40**

- ✓ Retrieval of DNA/RNA/protein sequences from databases; Biological sequence editing and analysis; Data mining tools
- ✓ Exploring the integrated database system at NCBI server and querying the PUBMED and GenBank databases using the ENTREZ search engine
- ✓ Sequence alignment: database searches (BLAST, FASTA etc.), PSA and MSA
- ✓ Phylogenetic tree construction, Bootstrapping
- ✓ Gene Identification and Function Prediction
- ✓ Comparative Genome analysis, alignment and visualization
- ✓ Codon Usage analysis bias prediction

**DBI-106** **Bioinformatics Programming (Practical)** **40**

- ✓ Basic Unix/Linux commands and Scripting
- ✓ Basic PERL programming; Designing PERL program for analysis of nucleotide and protein sequences; Uses of Scalar and Array Variables to manipulate DNA/RNA/Protein sequence data
- ✓ Reading a sequence data from a file and writing the results to a file; Concatenation DNA fragments, Transcribing DNA into RNA; Calculating the Reverse complement of a DNA strand; Uses of subroutine to read FASTA files
- ✓ Uses of Bioperl modules for sequence manipulation, accessing local database
- ✓ R scripting, Use of R for graph plotting, Heatmap construction, Gene expression analysis using R.

**DBI-107** **Continuing Evaluation** **15**

Seminar from a bunch of topics to be offered by the teachers

**DBI-108** **Continuing Evaluation** **45**

Class tests

## Semester –II

Course Code	Name of the module	Marks
<b>DBI-201</b>	<b>Genomics and Proteomics</b>	<b>40</b>
<i>Unit-1</i>	<ol style="list-style-type: none"><li>2. Nucleic acids and their structure; synthesis, modification and repair of DNA; repetitive and unique DNA sequences; split genes, overlapping genes and pseudo-genes. Plasmids, IS elements; transposons and retro-elements</li><li>3. Gene Identification: Genome information and special features, coding sequences (CDS), untranslated regions (UTR's), cDNA library, expressed sequence tags (EST).</li><li>4. Gene expression and DNA Microarray: Introduction, Basic steps for gene expression, concept of microarrays; gene expression analysis, Public Microarray data sources</li><li>5. DNA sequencing methods: manual and automated methods. Chain termination method; Basecalling and sequence accuracy. Quality assessment, NGS data assembly, annotation, Polymorphisms, Single Nucleotide Polymorphisms (SNPs)</li></ol>	
<i>Unit-2</i>	<ol style="list-style-type: none"><li>1. Proteomics: Definition–Transcriptomics; Proteomics, metabolomics. Techniques of proteomics – 2D PAGE, Mass spectrophotometer–(MALDI-TOF) MS, Protein Micro array in protein expression, profiling</li></ol>	
<b>DBI-202</b>	<b>Structural Bioinformatics</b>	<b>40</b>
<i>Unit-1</i>	<ol style="list-style-type: none"><li>1. Proteins: Amino acid components and structural features; Primary, secondary, tertiary, quaternary and supra-molecular structures; Non-covalent interactions in relation to structural conformation; Ramachandran plots; Amino acid sequencing and purification strategies</li><li>2. Methods for determining structure of biomolecules (X-ray, CD, NMR, IR, UV Visible, Fluorescence, SEM etc ) and their limitations.</li></ol>	
<i>Unit-2</i>	<ol style="list-style-type: none"><li>1. Prediction of protein structure from sequences, Homology modeling, functional sites, Protein folding problem, Secondary structure analysis and prediction, motifs, profiles, patterns and fingerprints search, protein structural databases (PDB), Purpose of 3-D structure comparison and concepts, RMSD, Z-score.</li></ol>	
<i>Unit-3</i>	<ol style="list-style-type: none"><li>1. Molecular Interaction Fields and Docking, Concept of Active Site of Enzymes, Protein-Protein and Protein-Ligand Docking; Basics of Molecular simulation, Computer Aided Drug Designing</li></ol>	
<b>DBI-203</b>	<b>Biostatistics and Database Management Systems (Theoretical)</b>	<b>40</b>
<i>Unit-1</i>	<ol style="list-style-type: none"><li>1. Measure of central tendency and dispersion; Concept of Correlation and Regression; . Types of Data: Concepts of population and sample, quantitative and qualitative data, cross-sectional and time-series data</li><li>2. Statistics and Sampling, Student's t2 test, X2 test, F-test</li><li>3. Concept of Set and matrices</li></ol>	
<i>Unit-2</i>	<ol style="list-style-type: none"><li>1. Introduction; Characteristics of Database approach; Advantages of using DBMS approach; Data models, schemas and instances; Database System Architecture.</li></ol>	
<i>Unit-3</i>	<ol style="list-style-type: none"><li>1. SQL Data Definition and Data Types; Specifying basic constraints in SQL; Basic queries in SQL; Insert, Delete and Update statements in SQL.</li><li>2. Data Storage and Querying: Storage and File Structure, Indexing and Hashing, Query Processing, Query Optimization.</li></ol>	
<b>DBI-204</b>	<b>Genomics &amp; Proteomics (Practical)</b> ✓ Exploring the Gene expression databases like GEO	<b>40</b>

- ✓ Analysis of a sample microarray data.
- ✓ Assembling and editing of Genomic data; Genome alignment and analysis tools- BWA (BurrowsWheeler Aligner), SAMtools, GATK (The Genome Analysis Toolkit), IGV (Integrative Genomics Viewer)
- ✓ NGS Data analysis; Data quality assessment; De-novo Genome assembly
- ✓ Genome Annotation
- ✓ Use of PAUP for Phylogenetic analysis based on RAPD Data.

<b>DBI-205</b>	<b>Structural bioinformatics (Practical)</b>	<b>40</b>
	<ul style="list-style-type: none"> <li>✓ Separation of amino acid mixture by chromatography</li> <li>✓ SDS-PAGE analysis of proteins</li> <li>✓ Protein Structure Determination, Homology modelling</li> <li>✓ Evaluation and Validation of protein models</li> <li>✓ Molecular visualization tools, 3D Representation</li> <li>✓ Structural analysis, Domain and motif identification, Ligplot interactions</li> <li>✓ Molecular Docking</li> <li>✓ Basic MD simulation with Gromacs.</li> </ul>	
<b>DBI-206</b>	<b>Biostatistics &amp; Database Creation (Practical)</b>	<b>40</b>
	<ul style="list-style-type: none"> <li>✓ Use of spreadsheet for data handling and graph plotting</li> <li>✓ Formulating and interpreting the result of a statistical test, measures of central tendency, measures of dispersion, measures of skewness</li> <li>✓ Creating database, Selecting database, Deleting database, Creating table, Modifying Table, Deleting table</li> <li>✓ SQL; Basic SQL commands; Inserting, updating and deleting records, retrieving records; Using Joins – joining a table to itself, joining multiple tables</li> <li>✓ Search engine Page design using HTML and PhP and connecting them to database</li> <li>✓ Submitting sequence data to NCBI database.</li> </ul>	
<b>DBI-207</b>	<b>Continuing Evaluation</b>	<b>15</b>
	Educational Trip along with submission of a final tour report	
<b>DBI-208</b>	<b>Continuing Evaluation</b>	<b>15</b>
	Class Tests	
<b>DBI-209</b>	<b>Continuing Evaluation</b>	<b>30</b>
	<p><u>Project/Dissertation Work:</u></p> <p>The students will be assigned with a practical project at the end of the first semester which they will have to complete by the end of second semester to the following themes (which is not an exhaustive list):</p> <ul style="list-style-type: none"> <li>➤ Phylogenetic Analyses</li> <li>➤ Construction of databases</li> <li>➤ Comparative genome analysis</li> <li>➤ Developing Bioinformatics tools/software</li> </ul>	