

**Department of Bioinformatics**  
**School of Earth, Biological and Environmental Sciences**  
**Central University of South Bihar**  
**Syllabus of M.Sc. Bioinformatics**  
**Session-2022-2023 onwards**

**Introduction**

Bioinformatics is a field of interdisciplinary study at the interface of biology and information technology. As such, Bioinformatics brings together molecular biology with areas from statistics, mathematics, and computer science. Bioinformatics now entails the creation and advancement of biological databases, algorithms, computational and statistical techniques, and theory to solve formal and practical problems arising from the management and analysis of biological data. Considering the interdisciplinary nature of Bioinformatics, this master programme shall have a major component from related subjects under one umbrella.

**Objectives**

The core objective of the programme is to provide quality education to the graduates, who want to pursue their career in the areas of Bioinformatics and Computational Biology.

- To serve as a nodal point for Bioinformatics and its applications
- To provide industry interface to the students for research projects
- To build skilled manpower to solve biological problems utilizing computational approaches

**Outcome**

- Students completing the course can efficiently provide bioinformatics solutions for biological research
- Learn problem-solving skills and to develop new algorithms
- Perform biological research utilising computational approaches

**Target Group and Eligibility**

Bachelor's degree in Biological Sciences/ Agricultural Science/ Pharmaceutical Science/ Veterinary Science/ Medical Science/ Mathematics/ Physics/ Chemistry/ Computer Science/ Information Technology or Integrated B.Sc. B.Ed. with 55% marks for General/OBC /EWS candidates and 50% marks for SC/ST/PWD candidates from any recognized University.

**Duration:** Two Years (Four Semester) – Full Time

**Course Structure**  
**M.Sc. Bioinformatics (02 Year)**  
**Total Credits: 80**

<b>Semester - I</b>					
		<b>Credits</b>			<b>Total Credits</b>
<b>DBCC*</b>	<b>Course title</b>	<b>L</b>	<b>T</b>	<b>P</b>	
BIS 8 1 DC 001 04	Linux and Bioinformatics	2	1	1	4
BIS 8 1 DC 002 04	Biochemistry and Molecular Biology	4	0	0	4
BIS 8 1 DC 003 04	Mathematics & Statistics with R	2	1	1	4
BIS 8 1 DC 004 04	Programming with Python	2	1	1	4
<b>Open elective<sup>@</sup></b>	Other department / SWAYAM / NPTEL				4
	<b>Total Credits</b>				<b>20</b>

**\*Discipline Based Core Course – DC**      **@Open elective (Interdisciplinary) Course - OE**  
**OE - Students shall earn 04 Credit from other department / SWAYAM / NPTEL**

<b>Semester – II</b>					
		<b>Credits</b>			<b>Total Credits</b>
<b>DBCC</b>	<b>Course title</b>	<b>L</b>	<b>T</b>	<b>P</b>	
BIS 8 2 DC 005 04	Algorithms in Bioinformatics	2	1	1	4
BIS 8 2 DC 006 04	Biomolecular Modelling and Simulations	3	0	1	4
BIS 8 2 DC 007 04	Introduction to Research Methodology	3	1	0	4
BIS 8 2 DC 008 04	Genomics and Transcriptomics	3	0	1	4
<b>DBCE*</b>	Any one from the list	2	1	1	4
	<b>Total Credits</b>				<b>20</b>
<b>DBCE courses</b>					
BIS 8 2 DE 009 04	Evolution and Molecular Phylogeny	2	1	1	4
BIS 8 2 DE 010 04	Programming with Perl	2	1	1	4
<b>MENC<sup>@</sup></b>					
BIS 8 2 ME 011 00	<b>ICT and Digital Skills</b> / Other department / SWAYAM / NPTEL	Non-credit course equivalent to 2 credits			

**\*Discipline Based Core Elective Course – DE**

**@Mandatory Elective Non-Credit Course - ME**

Students who exit after the successful completion of M.Sc. first year (I and II semesters) shall be awarded 01 year Post-Graduate Diploma in Bioinformatics.

Students wish to go for summer training (non-credit) may join the same at the end of II semester and present work done during summer training within one week of the commencement of semester III.

Experts from industry / alumni placed at various institutions may be called for guest lecture and interaction to get update on requirement in industry.

Educational/Industrial tour or Excursion may be made to visit institutes/industries/laboratories based on availability of funds or otherwise. If excursion took place, student shall submit a report (10 marks) within a week that will be the part of continuous assessment of a specific course decided by the department committee.

Semester -III					
DBCC	Course title	Credits			Total Credits
		L	T	P	
BIS 9 1 DC 012 04	Chemoinformatics and Drug Design	2	1	1	4
BIS 9 1 DC 013 04	DBMS and WEB Technology	2	1	1	4
BIS 9 1 DC 014 04	Applications of Bioinformatics	0	0	4	4
<b>DBCE</b>	Any one from the list	2	1	1	4
Open elective	Other department / SWAYAM				4
	<b>Total Credits</b>				<b>20</b>
<b>DBCE courses</b>					
BIS 9 1 DE 015 04	Systems Biology	2	1	1	4
BIS 9 1 DE 016 04	Exome Sequence Data Analysis	2	1	1	4
<b>MENC</b>					
BIS 9 1 ME 017 00	<b>Research &amp; Publication Ethics</b> / Other department / SWAYAM / NPTEL	Non-credit course equivalent to 2 credits			

Semester - IV					
DBCC	Course title	Credits			Total Credits
		L	T	P	
BIS 9 2 DC 018 20	Dissertation	-	-	20	20
	<b>Total Credits</b>				<b>20</b>

Student would have choice to carry out dissertation internally or externally to be evaluated by a committee of faculties at the end of semester. An elective will run if opted by 33% of the students enrolled in the programme.

#### List of open electives

Open electives	Course title	Credits			Total Credits
		L	T	P	
BIS 8 1 DC 001 04	Linux and Bioinformatics	2	1	1	4
BIS 8 2 DE 009 04	Evolution and Molecular Phylogeny	2	1	1	4
BIS 8 2 DE 010 04	Programming with Perl	2	1	1	4
BIS 9 1 DE 015 04	Systems Biology	2	1	1	4
BIS 9 1 DE 016 04	Exome Sequence Data Analysis	2	1	1	4

#### List of Swayam / NPTEL courses:

- Intellectual property
- Introduction to cell biology
- Mind education
- Environmental studies
- Immunology
- Indian society- social problems and issues
- Ecology and environment
- Academic writing
- Any other relevant course

Course code for SWAYAM/NPTEL courses will be taken as such, if given, or just name of portal will be mentioned otherwise. Students may register one (4 credits) or multiple courses (2 + 2 credits) from SWAYAM / NPTEL to complete the credit requirement.

## Linux and Bioinformatics

<b>Course Details</b>			
<b>Course Title: Linux and Bioinformatics</b>			
<b>Course Code</b>	BIS 8 1 DC 001 04	Linux and Bioinformatics	4
<b>L + T + P</b>	2 + 1 + 1	<b>Course Duration</b>	One Semester
<b>Semester</b>	I	<b>Contact Hours</b>	30 (L) + 15 (T) + 30 (P) Hours
<b>Nature of the Course</b>	Theory cum Practical		
<b>Special Nature/ Category of the Course</b>	Skill Based		
<b>Methods of Content Interaction</b>	Lecture, Tutorials, Practical, Group discussion; self-study, assignment class based as well as take-home type by students		
<b>Assessment and Evaluation</b>	<ul style="list-style-type: none"> <li>• 30% - Continuous Internal Assessment (Formative in nature but also contributing to the final grades)</li> <li>• 70% - End Term External Examination (University Examination)</li> </ul>		

### Course Objectives

Enable students to understand concept of computer architecture and its type. To make them aware of Unix/Linux operating system and various commands and filters to use in bioinformatics in general and online resources related to sequences, structures, pathway, and genome databases in particular to the students.

### Learning Outcomes

- After completing the course, the student will understand Linux operating system and recognize various online bioinformatics resources, able to retrieve relevant information from biological databases and also learn to differentiate between file formats to store biological data.

### Course Contents

#### Unit 1: (50% Weightage)

Fundamentals of computers and software systems: Types of computers, Operating Systems Computer Networks, Remote Login (telnet, ssh, ftp etc.). Introduction to Linux OS, File Management, Directories, File Permissions, Environment Setting, Basic Utilities commands: cp, mv, cd, ls, rm, clear, lp/lpr, cat, pwd, mkdir, man, tar/zip, apt/yum, echo, ps, kill, history, passwd, which, tail/head, more/less, wget. Pipe and Filters: Grep, SED, AWK, Compiling and Running of the Linux Source Code, Make and Makefile. Introduction to shell, writing shell scripts and its execution.

#### Unit 2: (50% Weightage)

General Introduction and history of bioinformatics, GenBank, SwissProt, PDB, Pfam, ZINC, KEGG, Gene expression and genome databases, Interaction databases, Gene ontology, PUBMED, Advanced internet search.

### List of practicals

1. Remote login protocols: telnet, FTP and ssh
2. Basic linux commands: cp, mv, cd, ls, rm, clear, cat, pwd, mkdir, man, tar/zip, apt/yum, echo, ps, kill, history, passwd, which, tail/head, more/less and wget
3. Use of pipes and filters: grep, sed and awk
4. writing shell scripts and execution
5. Retrieval of nucleotide/protein sequences from biological databases NCBI and SwissProt
6. Study of various file formats to store biological data (Fasta, GenBank, PDB)
7. Explore online resources including KEGG, GO Consortium
8. Advanced internet search on biological data
9. Retrieving sequences and annotation with help of genome browsers
10. Downloading sequence data in bulk

### Content Interaction Plan:

<u>Lecture cum Discussion (Each session of 1 Hour)</u>	<u>Unit/Topic/Sub-Topic</u>
1-2	Fundamentals of computers and software systems: Types of computers,
3-4	Operating Systems, Languages, Computer Networks.
5-6	Remote Login (telnet, ssh, ftp etc.). Introduction to Linux OS,
7-10	File Management, Directories, File Permissions, Basic Utilities commands: cp, mv, cd, ls, rm, clear, cat, pwd, mkdir, man, tar/zip, apt/yum, echo, ps, kill, history, passwd, which, tail/head, more/less, wget
11-14	Pipe and Filters: Grep, SED, AWK, Compiling and Running of the Linux Source Code, Make and Makefile, writing shell scripts.
15	Continuous assessment (Class Test/Assignment I)
16-17	General Introduction and history of bioinformatics

18	GenBank
19	SwissProt
20	PDB
21	Pfam (Class Test/Assignment II)
22-23	ZINC, KEGG
24-25	Gene expression and genome databases
26	Interaction databases
27-28	Gene ontology
29	PUBMED
30	Advanced internet search, Continuous assessment (Class Test/Assignment III)
<i>15 Hours</i>	Tutorials
<i>30 Hours</i>	Practical

Suggested Readings:

- Jerry P, Shelley P, O'Reilly T, Loukides M (2002) Unix Power Tools (3rd Edition). O'Reilly Media
- Neil M, Stones R (2007) Beginning Linux Programming (4<sup>th</sup> Edition). Wiley Dreamtech India Pvt Ltd
- Richard Blum C Bresnahan (2020) Linux Command Line and Shell Scripting (4<sup>th</sup> edition), Wiley publication.
- Christopher Negus, Linux Bible (10<sup>th</sup> edition ), Wiley
- Attwood T, Parry-Smith DJ (2001) Introduction to Bioinformatics. Pearson Education
- Mount DW (2004) Bioinformatics: Sequence and Genome Analysis. Cold Spring Harbor Laboratory Press
- Sharma V, Munjal A, Shanker A (2016) A Text Book of Bioinformatics. Rastogi Publications

## Biochemistry and Molecular Biology

<b>Course Details</b>			
<b>Course Title: Biochemistry and Molecular Biology</b>			
<b>Course Code</b>	BIS 8 1 DC 002 04	<b>Biochemistry and Molecular Biology</b>	4
<b>L + T + P</b>	4 + 0 + 0	<b>Course Duration</b>	One Semester
<b>Semester</b>	Odd	<b>Contact Hours</b>	60 (L) Hours
<b>Methods of Content Interaction</b>	Lecture, Tutorials, Group discussion; self-study, seminar, presentations by students, individual and group drills, group and individual field-based assignments followed by workshops and seminar presentations.		
<b>Assessment and Evaluation</b>	30% - Continuous Internal Assessment (Formative in nature but also contributing to the final grades) 70% - End Term External Examination (University Examination)		

### Course Objectives

- To introduce the students to classical and modern concepts in Biochemistry and molecular biology
- To learn their application in biological research.
- To understand the Biochemistry of Biomolecules at the molecular level mechanism and cellular processes.
- To understand the concept of cell and molecular biology in bioinformatics as one of the important key applied areas and its uses in biological research

### Learning Outcomes

After completion of the course, the learners will be able to:

- Understand the concepts of cell and molecular biology and utilize them in bioinformatics applications.
- Basic as well as an advanced application of Cell and Molecular Biology in Biological application

- To Understand the connection of Applied areas of cell and molecular biology and Bioinformatics for human diseases and health

## **Course Contents**

### **UNIT I: Chemistry of Biomolecules and Bioenergetics (25% Weightage)**

Water as Universal solvent, pH, pKa hydrogen bond, Amino acids structure-function, Titration curve of weak acid and bases charges on amino acid at different pH, pKa and pI, Carbohydrate structure and function, Lipids structure and function, Nucleic acids structure and function, Bioenergetics: Laws of Thermodynamics and its applications; Concept of free energy; High energy compounds; ATP as the main source of free energy in biological systems.

### **Unit 2: UNIT II: Enzyme Kinetics and Biochemical Pathways (25% Weightage)**

Enzymes, isozymes, and enzyme kinetics, Michaelis Menton Equation, Reciprocal plot/ Line viewer Burk plot Enzyme inhibition, Enzyme regulation, Regulation of Glycolysis, Gluconeogenesis, Regulation of TCA cycle, Glycogen metabolism in animals C16 Fatty Acid oxidation.

### **Unit 3: Transcription, Translational Mechanism and its control (25 % Weightage)**

Replication of DNA in prokaryotes and eukaryotes, Chromatin and its remodelling, Transcription and its machinery: RNA polymerases, formation of initiation complex, elongation and termination of transcription. Regulation of transcription: activators (enhancers) and repressors, Mechanism of translation: Initiation, elongation and termination factors and translational proof-reading in prokaryotes. Regulation of translation, posttranslational modifications of proteins, Protein trafficking and transport, Control of gene expression at transcription and translation level: Regulation of gene expression in prokaryotes and eukaryotes.

### **Unit 4: Cell Communications, receptors and its messengers (25 % Weightage)**

Cell signaling, Hormones and their receptors, cell surface receptor, signaling through G-protein coupled receptors, modulation of G-proteins through small molecules, activators, and antagonism, signal transduction pathways, second messengers, and regulation of signaling pathways.



**Content Interaction Plan:**

<b><u>Lecture cum Discussion</u></b> <b><u>(Each session of 1 Hour)</u></b>	<b><u>Unit/Topic/Sub-Topic</u></b>
(Unit I)1-2	Water as Universal solvent, hydrogen bond etc
3-6	Amino acids structure function, Titration curve of weak acid and bases charges on amino acid at different pH, pKa and pI,
7-9	Carbohydrate structure and function.
10-11	Fatty acids structure and function
12-13	Nucleic acid structure and function, Hydrogen bonding, nucleic acid protein complexes
13-15	Bioenergetics: Laws of Thermodynamics and its Applications; Concept of free energy; High energy compounds; ATP as the main source of free energy in biological systems.
(Unit –II)16-19	Enzymes, isozymes and enzyme kinetics,.Michaelis Menton Equation, Reciprocal plot or Line viewer burk plot
20-21	Enzyme inhibition
	<b>CONTINUOUS ASSESSMENT : 01</b>
22	Enzyme regulation
23-25	Regulation of Glycolysis, Gluconeogenesis,
26	Regulation of TCA cycle
27	Glycogen metabolism
28-30	Fatty Acid oxidation
(Unit –III) 31-32	Replication of DNA in prokaryotes and eukaryotes
32-34	Chromatin and its remodelling
34-35	Transcription and its machinery: RNA polymerases
35-36	Formation of initiation complex
36-37	Elongation and termination of transcription
38-39	Regulation of transcription: activators (enhancers) and repressors
39-40	<b>CONTINUOUS ASSESSMENT : 02</b>
40-41	Mechanism of translation: Initiation, elongation and termination factors
41-42	Translational proof-reading in prokaryotes
42-43	Regulation of translation, posttranslational modifications of proteins
43-44	Protein trafficking and transport
44-45	Control of gene expression at transcription and translation level

45-50	Regulation of gene expression in prokaryotes and eukaryotes
<b>(Unit –IV)</b>	Cell signalling
51-52	Hormones and their receptors
52-53	Cell surface receptor
53-54	Signaling through G-protein coupled receptors
54-55	Modulation of G-proteins through small molecules
55-56	Activators, and antagonism
56-57	Signal transduction pathways
57-58	Second messengers
58-59	Regulation of signaling pathways.
59-60	Signal transduction pathways <b>CONTINUOUS ASSESSMENT : 03</b>

**Suggested Readings:**

1. David LN, Michael MC (2013) Lehninger Principles of Biochemistry Sixth Edition by New York W.H. Freeman
2. Victor R, David B, Kathleen B, Peter K, Anthony W (2018) Harper's Illustrated Biochemistry Thirty-First Edition 31st Edition by McGraw Hill
3. Donald V, Judith GV, Charlotte WP (2016) Fundamentals of Biochemistry: Life at the Molecular Level 5th Edition by John Wiley and Sons Ltd
4. Jeremy MB, Lubert S, et al (2019 ) Biochemistry 9th edition. CBS publishers and distributors.
5. Roger LM, Megan MM (2017) Biochemistry First Edition by W. W. Norton & Company
6. Brown TA (2000) Genomes 2. BIOS Scientific Publishers Ltd, New York
7. Bruce A, Alexander J, Julian L, Martin R, Roberts K, Peter W (2008) Molecular Biology of the Cell (5th Edition). Garland Science
8. Bruce A, Dennis B, Karen H, Alexander JD, Julian, Martin LR, Roberts K, Peter W (2010) Essential Cell Biology (3rd Edition). Garland Science
9. Cooper GM, Hausman RE (2013) The Cell: A Molecular Approach. Sinauer Associates
10. Harris ELV, Angal S (1990) Protein Purification Applications: A Practical Approach. Oxford University Press
11. Michael RL, Stuart EB, Chih-duen CP, Richard CW (1990) Protein Purification: From Molecular Mechanisms to Large-Scale Processes (Acs Symposium Series). An American Chemical Society Publication
12. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual (3 Volume Set; Edition 2<sup>nd</sup>). Cold Spring Harbor Laboratory Press
13. Sambrook JF, Russell DW (2001) Molecular Cloning: A Laboratory Manual (3rd Edition). Cold Spring Harbor Laboratory Press
14. Watson JD (2001) The Double Helix: A Personal Account of the Discovery of the Structure of DNA. Simon & Schuster
15. Watson JD, Baker TA, Bell Stephen P, Gann A, Levine, Richard ML (2008) Molecular Biology of the Gene (6th Edition). Cold Spring Harbor Laboratory Press
16. Wilson K, Walker J (2005) Principles and Techniques of Biochemistry and Molecular Biology. Cambridge University Press

## Mathematics and Statistics with R

Course Details			
<b>Course Title: MATHEMATICS AND STATISTICS WITH R</b>			
<b>Course Code</b>	BIS 8 1 DC 003 04	<b>Credits</b>	4
<b>L + T + P</b>	2 + 1 + 1	<b>Course Duration</b>	One Semester
<b>Semester</b>	I	<b>Contact Hours</b>	30 (L) + 15 (T) + 30 (P) Hours
<b>Nature of the Course</b>	Theory cum Practicum		
<b>Special Nature/ Category of the Course (if applicable)</b>	Skill Based		
<b>Methods of Content Interaction</b>	Lecture, Tutorials, Practical, Group discussion; self-study, assignment class based as well as take-home type by students		
<b>Assessment and Evaluation</b>	<ul style="list-style-type: none"> <li>• 30% - Continuous Internal Assessment (Formative in nature but also contributing to the final grades)</li> <li>• 70% - End Term External Examination (University Examination)</li> </ul>		

### Course Objectives

This course is designed to provide basic understanding of mathematical and statistical concepts required for biological data analysis with R-Statistical software.

### Learning Outcomes

After completion of the course the learners will be able to solve linear equation with help of matrices. Calculate descriptive statistics of given data. Perform probability calculations. Calculate sensitivity and specificity of a diagnostics test. Perform testing of hypothesis in relation to biological problems. Understand correlation and its importance in solving biological problems

### Course Contents

**UNIT I:** **(25% Weightage)**  
 Sequences and series (Ap, GP, HP), Test for convergence, Vector – algebra and its applications, determinants, Gradient, Divergence & Curl, Matrices and its types, their eigenvalues and eigen vectors, transposition and inverse of matrices, Cayley–Hamilton theorem. Graph theory and applications.

**UNIT II:** **(25% Weightage)**

Limits and continuity. Introduction of Differential and Integral calculus; derivatives and partial derivatives, maxima and minima, Differential Equations. Ordinary differential equation (ODE), Linear and nonlinear, first and higher order ODEs. Laplace & Fourier transforms. Numerical Methods: ODE and Euler method, Integration by trapezoidal and Simpson's rule.

**UNIT III: (25% Weightage)**

Measure of central tendency: arithmetic mean, harmonic mean, geometric mean, weighted mean, median and mode of grouped and ungrouped data. Measure of dispersion: standard deviation and variance, Coefficients of dispersion and variance, Measure of skewness and kurtosis. Probability: Classical, frequency and axiomatic approach of calculating probabilities. Conditional probability and Bayes theorem. Random variable and probability distribution: discrete and continuous random variables, probability mass function and probability density function, cumulative distribution function for discrete and continuous random variables. Binomial, normal, Chi-square, t and F distribution.

**UNIT IV: (25% Weightage)**

Sampling theory: Population and sample, random sample, population parameters, Sample statistics, sampling distribution of means, sampling distribution of proportion, sampling distributions of differences and sums, sample variance. Estimation: unbiased, maximum likelihood, Bayesian. Test of hypothesis and significance: Type I and type II error, Power of test, P-values. Curve fitting and regression: method of least squares, standard error of estimate, correlation coefficients, Analysis of variance, nonparametric tests.

**List of practicals**

1. Vector and Matrices exercises with Mathematica
2. Integration and differentiation with Matlab
3. Introduction to R-Statistical software. Vectors and assignments, Vector arithmetics, generating regular sequences, logical vectors, character vector, selecting and modifying subset of a data set. Other objects in R: factors, arrays, matrices, list, data.frame and functions. R-objects their mode and attributes. Reading data from files.
4. Determinants of 2 by 2 single digit matrices. Transpose of an array.
5. Performing matrix multiplication in 'R', solving linear equations with help of 'R'.
6. Examining the distribution of a set of data.
7. Calculating probabilities and p-values.
8. One and two sample T-test.
9. Generating good quality plots in 'R': scattered plot, bar plot, histogram, pi-chart, density plot and boxplots.

**Content Interaction Plan:**

<b><u>Lecture cum Discussion (Each session of 1 Hour)</u></b>	<b><u>Unit/Topic/Sub-Topic</u></b>
1	Sequences and series (Ap, GP, HP), Limits and continuity

2-4	Vector algebra
5-6	Matrices and its types, their eigenvalues and eigen vectors, transposition and inverse of matrices, Cayley–Hamilton theorem.
7	Limits and continuity
8-9	Differential and Integral Calculus
10-11	Differential Equations, Ordinary differential Equations (ODE)
12	Mathematical Transforms numerical methods
13-14	Measure of central tendency: arithmetic mean, harmonic mean, geometric mean, weighted mean, median and mode of grouped and ungrouped data. Continuous assessment (Class Test/Assignment I)
15	Standard deviation and variance, Coefficients of dispersion and variance. Measure of skewness and kurtosis.
16-17	Classical, frequency and axiomatic approach of calculating probabilities. Conditional probability and Bayes theorem.
18-19	Discrete and continuous random variables, p.m.f, p.d.f and c.d.f.
20	Binomial, normal, Chi-square, t and F distribution. Continuous assessment (Class Test/Assignment II)
21-23	Population and sample, random sample, population parameters, Sample statistics. sampling distribution of means, sampling distribution of proportion.
24-25	Estimation: unbiased, maximum likelihood, Bayesian.
26-28	Test of hypothesis and significance: Type I and type II error, Power of test, P-values.
29-30	Curve fitting and regression: method of least squares, standard error of estimate. correlation coefficients, Analysis of variance, nonparametric tests (Class Test/Assignment III)
<i>15 Hours</i>	Tutorials

30 Hours	Practical
<p>Suggested Readings:</p> <ol style="list-style-type: none"> <li>1. Harper, Charlie (1978) Introduction to Mathematical Physics, Prentice-Hall, New Delhi</li> <li>2. Shanti Narayan (1993) A text book of matrices, S. Chand &amp; CO., New Delhi</li> <li>3. Deo N (2000). Graph Theory with Applications to Engineering and Computer Science, Prentice-Hall, New Delhi</li> <li>4. Tenenbaum M, Pollard H (1985) Ordinary differential equation, Dover Publication, New York</li> <li>5. Shanti Narayan, Mittal PK (1942) Differential Calculus. S. Chand &amp; CO., New Delhi.</li> <li>6. William Feller (1971) An introduction to probability theory and its applications. Wiley publication</li> <li>7. Das NG (2017) Statistical methods. Mcgraw Hill Education</li> <li>8. George A (2012) Mathematical methods for physicists. Orlando Academic Press</li> <li>9. Pal and Sarkar (2007) Statistics: Concepts and Applications. PHI Learning Pvt Ltd</li> <li>10. Rosner B (2010) Fundamentals of Biostatistics. Duxbury Press</li> </ol>	

## Programming with Python

Course Details			
<b>Course Title: Programming with Python</b>			
<b>Course Code</b>	BIS 8 1 DC 004 04	<b>Credits</b>	4
<b>L + T + P</b>	2 + 1 + 1	<b>Course Duration</b>	One Semester
<b>Semester</b>	II	<b>Contact Hours</b>	30 (L) +15 (T) + 30 (P) Hours
<b>Nature of the Course</b>	Theory cum Practicum		
<b>Special Nature/ Category of the Course</b>	Skill Based		
<b>Methods of Content Interaction</b>	Lecture, Tutorials, Practical, Group discussion; self-study, assignment class based as well as take-home type by students		
<b>Assessment and Evaluation</b>	<ul style="list-style-type: none"><li>• 30% - Continuous Internal Assessment (Formative in nature but also contributing to the final grades)</li><li>• 70% - End Term External Examination (University Examination)</li></ul>		

### Course Objectives

- Identify/characterize/define a problem
- Design a program to solve the problem
- Enable students with Python Scripting
- Hone students programming capability to write scripts to solve bioinformatics problems

### Learning Outcomes

- After completion of the course the learners will be able to:
- Understand Python programming syntax
- Apply python programming to solve computational biology problems
- Able to write python scripts for various bioinformatics applications

### Course Contents

#### Unit 1:

**(50% Weightage)**

Introduction to python programming, Keywords & Identifier, Statements & Comments, Python I/O, Python Data Types (numbers, string, lists, tuples and dictionary), Python Operators and Operator Precedence, Conditional Statements if-else, for and while loop, break, continue, and pass statement, Python functions: syntax, function call, types of functions, function arguments, anonymous functions.

**Unit 2:****(50% Weightage)**

File handling: open, reading writing and closing of files, Directory handling, User defined exception, Module and packages, import and reloading module, importing module from packages, Object & Class in python: defining an Object and class in python, deleting attributes and objects, namespace and variable scopes, inheritance and method overriding, multiple inheritances, resolution order.

**Lists of Practicals**

- Python Syntax
- Datatype and variables in Python
- Conditional and Logical statements in Python
- Loops in Python
- Function, definition and calling
- File handling and file operation in Python
- Directory handling and its operation in Python
- Python modules and packages and its integration
- OOP in python

**Content Interaction Plan:**

<b><u>Lecture cum Discussion</u></b> <b><u>(Each session of 1 Hour)</u></b>	<b><u>Unit/Topic/Sub-Topic</u></b>
<b>(Unit I)</b> 1-2	Introduction to python programming, Keywords & Identifier, Statements & Comments
2-5	Python Data Types (numbers, string, lists, tuples and dictionary)
5-6	Python Operators and Operator Precedence
7-10	Conditional Statements if-else, for and while loop, break, continue and pass statement
11-15	Python functions: syntax, function call, types of functions, function arguments, anonymous functions
16	<b>CONTINUOUS ASSESSMENT: 01</b>
<b>(Unit –II)</b> 17-18	File Handling: open, reading writing and closing of files
19-20	Directory handling, User defined exception



20-21	Module and packages, import and reloading module, importing module from packages <b>CONTINUOUS ASSESSMENT: 02</b>
22-23	Object & Class in python: defining an Object and class in Python
24-25	Deleting attributes and objects, namespace and variable scopes
26-28	Inheritance and method overriding
29-30	Multiple inheritances, resolution order <b>CONTINUOUS ASSESSMENT: 03</b>
15 Hours	<b>Tutorial</b>
30 hours	Practical
<p><u>Suggested Readings:</u></p> <ul style="list-style-type: none"> <li>• David M, Beazley DM (2009) Python Essential Reference (4th Edition). Pearson Addison-Wesley Professional</li> <li>• Downey AB (2009) Think Python: An Introduction to Software Design. Create Space Independent Publishing Platform</li> <li>• Lutz M (2009) Learning Python (5th Edition). O'Reilly</li> <li>• Lambert KA (2011) The Fundamentals of Python: First Programs, Cengage Learning</li> </ul>	

## Algorithms in Bioinformatics

<b>Course Details</b>			
<b>Course Title: Algorithms in Bioinformatics</b>			
<b>Course Code</b>	BIS 8 2 DC 005 04	Algorithms in Bioinformatics	4
<b>L + T + P</b>	3 + 0 + 1	<b>Course Duration</b>	One Semester
<b>Semester</b>	II	<b>Contact Hours</b>	45 (L) + 30 (P) Hours
<b>Nature of the Course</b>	Theory cum Practical		
<b>Special Nature/ Category of the Course</b>	Skill Based		
<b>Methods of Content Interaction</b>	Lecture, Tutorials, Group discussion; self-study, seminar, presentations by students, individual and group drills, group and individual field based assignments followed by workshops and seminar presentation.		
<b>Assessment and Evaluation</b>	<ul style="list-style-type: none"> <li>• 30% - Continuous Internal Assessment (Formative in nature but also contributing to the final grades)</li> <li>• 70% - End Term External Examination (University Examination)</li> </ul>		

### Course Objectives

- To deals with various algorithms used to solve biological problems
- To understand the algorithms for sequence alignment, promoter and gene prediction.
- To learn machine learning algorithms to build their own computational biology models.
- To Articulate the trade-offs behind algorithms in bioinformatics
- Locate and critically evaluate scientific information

### Learning Outcomes

After completion of the course the learners will be able to:

- Familiar with the use of various computational algorithms
- Algorithms use in solving biological problems.
- Students will be able to design and analyze their algorithms.
- Apply one of the paradigms presented in class to solve real-world problems
- Present scientific content to a small technical audience

## Course Contents

### Unit 1: (25 % Weightage)

A brief introduction to algorithms, history, complexity of an algorithm and pseudocodes, introduction to Dynamic programming, Divide & conquer, Brute force and Greedy algorithms. Representation of algorithms flowchart & pseudocode. Concept of Homologous, orthologous, paralogous and Xenologous sequences, Alignment of pair of sequences, and gap penalties, dot plot method, dynamic programming algorithm as applicable to global (Needleman-Wunsch) and local (Smith - Waterman) sequence alignment. Database searching for similar sequences: Heuristic algorithm FASTA, BLAST and their types.

### Unit 2: (25 % Weightage)

Multiple sequence alignment (MSA), dynamic programming for MSA, heuristic approach to MSA, progressive sequence alignments method, Clustal W/X, Branch and bound and other methods for MSA, Consensus sequence & PSSM, Scoring matrices (PAM and BLOSUM) Whole genome alignments: concepts and applications, Cluster of Orthologous Groups.

### Unit 3: (25 % Weightage)

Promoter prediction in prokaryote and eukaryotes, Prediction of genes and gene structure for prokaryotic and eukaryotic organisms, sequence and structural motifs and motif prediction tools RNA structure prediction algorithms, Prediction of splice sites. Genetic algorithms, Graph Theory with Applications, Adjacency matrix, Network properties.

### Unit 4: (25 % Weightage)

Machine Learning/Deep Learning algorithms, Clustering- K-means, Hierarchical etc., Dimensionality Reduction Algorithms, Logistic Regression, Naive Bayes, Decision Tree; Random Forest, Neural networks, Support vector machine, deep learning.

## List of Practical

1. Finding similar sequences (DNA/proteins) in databases: use of different version of BLAST, installation and use of local BLAST and interpretation of BLAST results.
2. Multiple sequence alignments using web-based and stand-alone version of CLUSTAL. Interpretation of results to identify conserved and variable regions and correlate them with physico-chemical & structural properties.
3. Prediction of ORFs.
4. Gene Prediction in prokaryotes and eukaryotes.
5. Primer designing.
6. Prediction of promoters using Neural Network Promoter Prediction (NNPP).
7. Prediction of alternate splice sites using methods such as Splice Site Prediction by Neural Network.
8. Predicting protein function from 3D structure using ProFunc.

**Content Interaction Plan:**

<b><u>Lecture cum Discussion</u></b> <b><u>(Each session of 1</u></b> <b><u>Hour)</u></b>	<b><u>Unit/Topic/Sub-Topic</u></b>
1-2	A brief introduction to algorithms, history, complexity of an algorithm and pseudocodes, introduction to Dynamic programming, Divide & conquer, Brute force and Greedy algorithms.
3-4	Representation of algorithms flowchart & pseudocode Homologous, orthologous, paralogous and Xenologous sequences
5-6	Alignment of pair of sequences, and gap penalties, Dot plot method
7-8	Dynamic programming algorithm as applicable to global (Needleman-Wunsch) and local (Smith - Waterman) sequence alignment
8-9	Database searching for similar sequences: Heuristic algorithm FASTA, BLAST and their types.
10	<b>CONTINUOUS ASSESSMENT : 01</b>
11	Multiple sequence alignment (MSA)
12-13	Dynamic programming for MSA
13-15	Heuristic approach to MSA, Consensus sequence & PSSM, Scoring matrices
16	Progressive sequence alignments method
17	Clustal W/X
18	Branch and bound and other methods for MSA
19	Whole genome alignments: concepts and applications,
20	Cluster of Orthologous Groups
21	<b>CONTINUOUS ASSESSMENT : 02</b>
22	Promoter prediction in prokaryote and eukaryotes
23	Prediction of genes and gene structure for prokaryotic organisms
24	Prediction of genes and gene structure for eukaryotic organisms
25-26	Sequence and structural motifs and motif prediction tools. RNA structure prediction algorithms
27	Software Information and demo
28	Prediction of splice sites.
29	Application of machine learning Techniques: Genetic algorithms
30	Application of machine learning Techniques: Graph Theory with Applications, Adjacency matrix, Network properties

31	Application of machine learning Techniques Clustering- K-means, Hierarchical
32-34	Application of machine learning Techniques: Logistic Regression, Naive Bayes, Decision Tree; Random Forest
35-36	Application of machine learning Techniques: Neural networks
37-38	Application of machine learning Techniques: Support vector machine.
39-41	Application of machine learning Techniques: deep learning
42-43	Problem Discussion
44	Problem Discussion
45	<b>CONTINUOUS ASSESSMENT : 03</b>
30 Hrs	<b>Practicals</b>

Suggested References:

1. Cormen TH, Leiserson CE, Rivest RL, Stein C (2009) Introduction to Algorithms. The MIT Press
2. Dey DK, Ghosh S, Mallick BK (2010) Bayesian Modelling in Bioinformatics. Chapman and Hall
3. Gusfield D (1997) Algorithms on Strings, Trees, and Sequences. Cambridge University Press
4. Jones NC, Pevzner PA (2004) An Introduction to Bioinformatics Algorithms. The MIT Press
5. Klir GJ, Yuan B (1995) Fuzzy Sets and Fuzzy Logic: Theory and Applications. Prentice Hall
6. Mitchell T (1997) Machine Learning. The McGraw-Hill Company
7. Mitra S, Dutta S, Perkins T, Michailidis G (2008) Introduction to Machine learning and Bioinformatics. Chapman and Hall
8. Mount DW (2001) Bioinformatics: Sequence and Genome Analysis. Cold Spring Harbor Laboratory Press
9. Rao SS (1991) Optimization: Theory and applications. Wiley Eastern Ltd
10. Sharma V, Munjal A, Shanker A (2016) A Text Book of Bioinformatics. Rastogi Publications
11. Waterman M (1995) Introduction to Computational Biology: Maps, sequences and genomes. Chapman & Hall
12. Witten IH, Frank E, Hall MA, Pal CJ (2011) Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann

## Bimolecular Modeling and Simulations

<b>Course Code</b>	BIS 8 2 DC 006 04	<b>Credits</b>	4
<b>L + T + P</b>	3 + 0+ 1	<b>Course Duration</b>	One Semester
<b>Semester</b>	II	<b>Contact Hours</b>	45 (L) + 30 (P) Hours
<b>Course Type</b>	Discipline Based Core		
<b>Nature of the Course</b>	Lecture and Practicals		
<b>Special Nature/ Category of the Course</b> <i>(if applicable)</i>	Skill Based		
<b>Methods of content Interaction</b>	Lecture and Practical, Group discussion; self-study, seminar, presentations by students, individual and group drills, group and individual assignments followed by paper presentation		
<b>Assessment and Evaluation</b>	<ul style="list-style-type: none"> <li>• 30% - Continuous Internal Assessment (Formative in nature but also contributing to the final grades)</li> <li>• 70% - End Term External Examination (University Examination)</li> </ul>		

### Course Objectives:

- ❖ The course will impart essential training in molecular modeling, simulation, structure determination, and prediction to add to research in pharmaceutical and biological sciences. Further, the present course introduces concepts, algorithms, and software in structural biology and protein modeling.

### Learning Outcomes

- ❖ The student would be able to understand, model, and analyze the structure of small and macromolecules.
- ❖ Students would be able to predict the structures and grasp the idea of computational structure prediction.
- ❖ Students would develop hands-on experience in molecular modeling and molecular dynamics simulation.

### Course Contents

#### UNIT I: Nomenclature, Protein geometry, and structure determination (25% Weightage)

- Nomenclature & terms used in molecular modeling:
- Protein data file and numbering of side-chain atoms, CORN rule, and torsion angle
- peptide geometry
- Ramachandran plot
- Secondary structure, helix, sheet, and coils
- Reverse turns, beta-hairpin
- Side chain conformation
- Super secondary structure; motifs, folds, and domain
- Protein folding, Affinsen's experiment, Levinthal paradox, protein folding funnel
- Quaternary structure
- Overview of experimental techniques to study macromolecular structures: crystallization method and X-ray crystallography, NMR methods of structure determination, and Cryo-electron microscopy

#### UNIT II: 2D and 3D structure prediction methods (25% Weightage)

- Secondary Structure Prediction Methods: Chou and Fasman, GOR, Neural Network,

- Computational methods of 3D structure prediction:
- Comparative modeling
- Fold recognition method(threading)
- I-TASSER
- Ab-initio structure prediction
- Structure consistency and validation tools: SAVES server, Procheck, and WhatIF MolProbity and ProSA
- Structural comparison & alignment methods DALI

### **UNIT III: Overview of molecular modeling and simulation (25 % Weightage)**

- A brief overview of molecular modeling methods: Integration algorithm: Verlet algorithm, and Velocity Verlet algorithm
- Born Oppenheimer approximation, the potential energy function
- A simple molecular mechanics force field
- The general feature of MM force field, ball-and-spring model, harmonic approximation, bond length, bond angle, torsional angle
- Out-of-plane and cross-terms, electrostatics, and van der Waal term of biomolecules
- Popular force fields - MM, AMBER, CHARMM, OPLS, NAMD, GROMACS
- Energy Minimization: derivative-based method (SD, CG).

### **UNIT IV: Molecular dynamics methods (25%Weightage)**

- Molecular Dynamics:
- Steps in typical MD simulations, minimization, equilibration and
- Data collection and analysis
- Time steps
- Periodic boundary condition
- Analysis of trajectories
- Solvation, water model: implicit and explicit solvation model
- Membrane models
- Constant temp dynamics
- Constant pressure dynamics
- Monte Carlo (MC) methods
- Choice of method: Monte Carlo or Molecular dynamics simulation

### **List of Practicals:**

- Exploring the PDB Database.
- Protein structure visualization: Pymol
- Methods for prediction of the secondary structure of proteins: Prediction of secondary structures of proteins using at least two different approaches with analysis and interpretation of the results. Comparison of the performance of the different methods for various classes of proteins
- Methods for tertiary structure prediction of proteins along with analysis and interpretation of results: Homology modeling: Modeller, SWISS-MODEL
- Fold recognition methods: I-TASSER
- *Ab initio* structure prediction methods
- Use of Steepest Decent and Conjugate Gradient for energy minimization of the biomolecule.
- MD Using GROMACS: MD using implicit solvent, MD using an explicit solvent with PBC/layer of solvent.

- Visualization and analysis of MD trajectory data (VMD)

### **Content Interaction Plan:**

<b><u>Lecture cum Discussion</u></b> <b><u>(Each session of 1 Hour)</u></b>	<b><u>Unit/Topic/Sub-Topic</u></b>
1-3	<b>UNIT I: Nomenclature, Protein geometry, and structure determination</b> Nomenclature & terms used in molecular modeling: Protein data file and numbering of side-chain atoms, CORN rule, torsion angle, stereochemistry, peptide geometry, Ramachandran plot,
4-5	Secondary structure, helix, sheet, and coils, reverse turns, beta-hairpin, side-chain conformation, super secondary structure; motifs, folds, and domain,
5-6	protein folding, Affinsen's experiment, Levinthal paradox, protein folding funnel, quaternary structure,
7-11	Overview of experimental techniques to study macromolecular structures: crystallization method and X-ray crystallography, NMR methods of structure determination, Cryo-electron microscopy
12-13	<b>UNIT II: 2D and 3D structure prediction methods</b> Secondary Structure Prediction Methods: Chou and Fasman, GOR, Neural Network,
14-16	Computational methods of 3D structure prediction: comparative modeling, fold recognition method(threading),
17-22	I-TASSER, ab-initio structure prediction,
23	Structure consistency and validation tools: SAVES server, Procheck, WhatIF, MolProbity, ProSA, Structural comparison & alignment methods DALI
24-25	<b>UNIT III: Overview of molecular modeling and simulation</b> A brief overview of molecular modeling methods: Integration algorithm, Verlet algorithm, Velocity Verlet algorithm, Born Oppenheimer approximation,
26-30	potential energy function, A simple molecular mechanics force field, General feature of MM force field, ball-and-spring model, harmonic approximation, bond length, bond angle, torsional angle, out-of-plane and cross-terms, Non bonded interactions: electrostatics of biomolecules, hydrogen bonding, and van der Waal interactions
31-32	Popular force fields - MM, AMBER, CHARMM, OPLS, NAMD, GROMACS
33-34	Energy Minimization: derivative-based method (SD, CG)
35-39	<b>UNIT IV: Molecular dynamics methods</b> Molecular Dynamics: Steps in typical MD simulations, minimization, equilibration, and data collection
40-41	Time steps, periodic boundary conditions, analysis of trajectories,
42-43	Solvation, water model: implicit and explicit solvation model, membrane models, constant temp dynamics, constant pressure dynamics
44-45	Monte Carlo (MC) methods, choice of method: Monte Carlo or Molecular dynamics simulation



**Essential Readings:**

1. Bourne PE, Weissig H (2003) Structural Bioinformatics (Methods of Biochemical Analysis, V. 44). Wiley-Liss Publisher
2. Leach A (2001) Molecular Modelling: Principles and Applications. Prentice-Hall

**Additional/Advance/Further Readings:(Web Links/Research Papers)**

3. Ramachandran, GN et al. (1963). Stereochemistry of polypeptide chain configurations
4. <https://www.nobelprize.org/uploads/2018/06/popular-chemistryprize2017.pdf>
5. Roman AL (2001) PDBsum: summaries and analyses of PDB structures
6. Orengo, CA et al., (1999) The CATH Database provides insights into protein structure/function relationships
7. <http://www.mdtutorials.com/gmx/lysozyme/index.html>
8. Ambrish R. et al., (2010) I-TASSR: A unified platform for automated protein structure and function prediction
9. Stewart AA (2006) Molecular dynamics: Survey of methods for simulating the activity of proteins
10. Wilfred FVG and Hermen JCB (1990) Computer simulation of molecular dynamics: Methodology, Application, and perspective in chemistry.
11. James UB (1991) A Method to identify protein sequences that fold into a known three-dimensional structure
12. Markush W and Manfred JS (2007) PosSA Web: Interactive web service for the recognition of errors in the 3-D structure of proteins
13. Liisa H and Paivi R (2010) Dali server: conservation mapping in 3D
14. Michel PA (2004) Introduction to molecular dynamics simulation

## Introduction to Research Methodology

<b>Course Details</b>			
<b>Course Title: Introduction to Research Methodology</b>			
<b>Course Code</b>	BIS 8 2 DC 007 04	Introduction to research methodology	4
<b>L + T + P</b>	3 + 1 + 0	<b>Course Duration</b>	One Semester
<b>Semester</b>	II	<b>Contact Hours</b>	30 (L) + 15 (T) Hours
<b>Nature of the Course</b>	Theory		
<b>Special Nature/ Category of the Course</b>	Value based (ethics)		
<b>Methods of Content Interaction</b>	Lecture, Tutorials, self-study, seminar presentations by students.		
<b>Assessment and Evaluation</b>	<ul style="list-style-type: none"> <li>• 30% - Continuous Internal Assessment (Formative in nature but also contributing to the final grades)</li> <li>• 70% - End Term External Examination (University Examination)</li> </ul>		

❖ **Course Objectives:**

- ❖ This course aims to provide an overview on fundamentals of conducting research including scientific terminology, methods, literature review, data collection, analysis and interpretation, preparation of research report and presentation. It is designed to introduce essential aspects related to research such as intellectual property, bioethics and biosafety as well. The course is intended to develop research skills including written and oral communication skills.

❖ **Course Learning Outcomes:**

- ❖ This course will help student develop key skills to carry out research in scientific areas. It will also train them in IPR and bioethics.
- ❖ Additionally, the course is designed to enhance student's communication skill.

❖ **Course Contents**

**UNIT 1: An Overview of research methodology**

**(25% Weightage)**

History of development of science, philosophy of science, Logical Reasoning, introduction of Research Methodology, Defining Research Problem, measurement technique, sampling design, Formulation of hypothesis type and testing, data collection, processing and analysis.

**UNIT 2: Biosafety, Ethics and IPR (25% Weightage)**

Research Ethics & Misconduct, Good Laboratory Practices (GLP), Patents, Copyrights, GI and Trademarks, Patent Treaties and Convention, database of patent, search and retrieval.

**UNIT 3: Scientometrics, Scientific Writing & communication (50% Weightage)**

Types of Scholarly Scientific content, Journal indexing, SCOPUS and WoS, Impact factor & other citation indices and matrices, h-index and i<sub>10</sub>-index, Basics of writing an article in scholarly journals, Writing style and various presentation tools.

Seminar Presentation: Each student would choose a recently published paper or review article from prestigious journals and make a presentation during the semester.

**Content Interaction Plan:**

<b><u>Lecture cum Discussion</u></b> <b><u>(Each session of 1 Hour)</u></b>	<b><u>Unit/Topic/Sub-Topic</u></b>
1-2	History of development of science, philosophy of science
3-4	Research Methodology, Defining Research Problem
5-6	Measurement technique, sampling design, Formulation of hypothesis type and testing, data collection, processing and analysis
7-9	Research Ethics & Misconduct
10	Good Laboratory Practices (GLP)
11-15	Patents, Copyrights, GI and Trademarks, Patent Treaties and Convention
16	Scientometrics
17-20	Journal indexing, SCOPUS and WoS
21-25	Impact factor & other citation indices and matrices, h-index and i <sub>10</sub> -index
26-29	Scientific writing, Writing style and various presentation tools,
30	Problem Discussion
31-45	Presentations

**Suggested Readings:**

- Kothari CR (2019) Research Methodology: Methods and Techniques. New Age International, New Delhi
- Corley RB (2005). A Guide to Methods in The Biomedical Sciences. Springer Germany.
- Matthews RW, Bowen JM, Matthews JR (1996). Successful Scientific Writing: Guide for biological and medical science writing. Cambridge University Press UK.
- Sateesh MK (2013) Bioethics And Biosafety. IK International, New Delhi.

**Additional/Advance/Further Readings:**

- Contents in USPTO, WIPO and IPO site.
- Gupta K, Karihaloo JL, Khetarpal RK (2008). Biosafety regulations of Asia-Pacific countries. In: Biosafety regulations of Asia-Pacific countries. National Bureau of Plant Genetic Resources, New Delhi, India, ISBN 978-92-5-105828-9.
- Richard WS (2000) Intellectual Property: Patents, Trademarks, and Copyrights, 2nd Edition. Delmar Cengage Learning.
- Bird A (1998). Philosophy of Science. Routledge/Talyor & Francis, London

## Genomics and Transcriptomics

<b>Course Details</b>			
<b>Course Title: Genomics and Transcriptomics</b>			
<b>Course Code</b>	BIS 8 2 DC 008 04	<b>Credits</b>	4
<b>L + T + P</b>	3 + 0 + 1	<b>Course Duration</b>	One Semester
<b>Semester</b>	II	<b>Contact Hours</b>	45 (L) + 30 (P) Hours
<b>Nature of the Course</b>	Theory cum Practicum		
<b>Special Nature/ Category of the Course (if applicable)</b>	Skill Based		
<b>Methods of Content Interaction</b>	Lecture, Practical, Group discussion; self-study, assignment class based as well as take-home type by students		
<b>Assessment and Evaluation</b>	<ul style="list-style-type: none"> <li>• 30% - Continuous Internal Assessment (Formative in nature but also contributing to the final grades)</li> <li>• 70% - End Term External Examination (University Examination)</li> </ul>		

### Course Objectives

The course is designed to provide students the knowledge of current tools and techniques used in genomics and transcriptomics.

### Learning Outcomes

Students will be able to understand the nature of genomics and transcriptomic data, how it is generated and tools used in analysis. The course will help students to develop practical knowledge of differential expression analysis, clustering, gene set enrichment analysis, etc.

### Course Contents

#### UNIT I: (50% Weightage)

Introduction to genomics, Genetic and physical mapping, Molecular markers: RFLP, RAPD, AFLPs, SSRs, Recombinant DNA technology: Cloning strategies, vectors: YAC and BAC, Methods of DNA sequencing: Sanger's/dideoxy method, Sequence assembly: shotgun and

clone contig approach. Next Generation sequencing: Platforms based on emulsion PCR, bridge amplification and single molecule real time (SMRT) sequencing; Genome assembly algorithms.

**UNIT II:**

**(50 % Weightage)**

Goals of transcriptomics experiment, Basics of RNASeq data generation, read sequence, FASTQ file format and phred quality score, Converting read sequences to expression data, SAM file format. Experimental design for comparing transcriptome, statistical power and number of replicates. Reproducibility of transcriptomic experiments.

Filtering of non-informative genes. Supervised and unsupervised clustering. K-means and hierarchical clustering. Cluster validation. Generating Heatmaps. T-test and ANOVA to identify differentially expressed genes, MA-plot, Volcano plot. Making sense of list of differentially expressed genes, hypergeometric distribution, gene ontology, Gene set enrichment analysis, Gene regulatory networks.

**List of Practicals**

- Transcriptome databases.
- Tools for Genomic Data Mining.
- Transcriptome data visualization and processing.
- RNASeq read sequences quality control.
- Filtering low quality reads and removing adapter contamination.
- FASTA/Q file manipulation using seqkit tool.
- Aligning read sequences to reference genome.
- Biomarker identification: Identification of differentially expressed genes with 'R' and Bioconductor.
- Performing gene ontology and gene set enrichment analysis.

**Content Interaction Plan:**

<b><u>Lecture cum Discussion (Each session of 1 Hour)</u></b>	<b><u>Unit/Topic/Sub-Topic</u></b>
1-3	Introduction to genomics, Genetic and physical mapping
4-5	Molecular markers: RFLP, RAPD, AFLPs, SSRs

6-7	Recombinant DNA technology: Cloning strategies, vectors: YAC and BAC
8-10	Methods of DNA sequencing: Sanger's/dideoxy method, Sequence assembly: shotgun and clone contig approach
11-13	Next Generation sequencing: Platforms based on emulsion PCR, bridge amplification and single molecule real time (SMRT) sequencing;
14-15	Genome assembly algorithms Continuous assessment (Class Test/Assignment I)
16	Understanding Next generation sequencing data
17-18	Experimental design for comparing transcriptome, statistical power and number of replicates, Reproducibility of transcriptomic experiments.
19-20	Basics of RNASeq data generation, read sequence, FASTQ file format and phred quality score, Converting read sequences to expression data, SAM file format.
21-22	Filtering non-informative probes. Supervised and unsupervised clustering, k-means clustering. Continuous assessment (Class Test/Assignment II)
23-24	Hierarchical clustering and Cluster validation, heat maps. Finding genes differentially expressed between two conditions, T-test.
25-26	Finding genes differentially expressed between more than two condition, ANOVA test.
27-30	Gene Ontology, Making sense of list of differentially expressed genes, gene only analysis and hypergeometric distribution. Gene set enrichment analysis and Gene regulatory networks (Class Test/Assignment III)
30 Hours	Practical

- Suggested References:

- Brown TA (2017) Genomes 4. Garland Science
- Primrose SB, Twyman RM (2013) Principles of Gene Manipulation and Genomics. Wiley-Blackwell
- Lesk A (2017) Introduction to Genomics. Oxford
- Bernhard OP (2006) Systems Biology: Properties of Reconstructed Networks. Cambridge University Press
- Melanie Kappelmann-Fenzl (2021) Next Generation Sequencing and Data Analysis. Springer Cham.
- Jonathan P (2015) Bioinformatics and Functional Genomics. Wiley-Blackwell
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Research 43: e47
- Stekel D (2003) Microarray Bioinformatics. Cambridge University Press
- Csaba O, Zsuzsanna O (2017) Molecular Data Analysis Using R. John Wiley & Sons
- Grolemond G (2016) R for Data Science. O'Reilly Media



## Evolution and Molecular Phylogeny

<b>Course Details</b>			
<b>Course Title: Evolution and Molecular Phylogeny</b>			
<b>Course Code</b>	BIS 8 2 DE 009 04	<b>Credits</b>	4
<b>L + T + P</b>	2 + 1 + 1	<b>Course Duration</b>	One Semester
<b>Semester</b>	I	<b>Contact Hours</b>	30 (L) + 15 (T) + 30 (P) Hours
<b>Nature of the Course</b>	Theory cum Practicum		
<b>Special Nature/ Category of the Course</b> <i>(if applicable)</i>	Skill Based		
<b>Methods of Content Interaction</b>	Lecture, Tutorials, Practical, Group discussion; self-study, seminar, presentations and assignment class based as well as take-home type by students,		
<b>Assessment and Evaluation</b>	<ul style="list-style-type: none"> <li>• 30% - Continuous Internal Assessment (Formative in nature but also contributing to the final grades)</li> <li>• 70% - End Term External Examination (University Examination)</li> </ul>		

### Course Objectives

- To understand the concept & fundamental of classical evolution principles and theories.
- To understand Molecular basis of Evolution, theories, Phylogeny and molecular clock.
- To be able to explain the various approaches and algorithm to inferring molecular phylogeny among organisms.
- To be able to explain and evaluate different phylogenetic optimality criteria, and select adequate criterion to solve a given problem.
- To select and apply, for the problem area, existing software on a given biological problem

### Learning Outcomes

The Students will be able to

- Understand the theoretical aspects of classical, molecular evolution and applications.
- Measuring the rate of evolution and concept of molecular clock hypothesis.
- Understand various algorithms and their comparison for deducing phylogenetic tree among species.
- Know how to construct phylogenetic trees, using a variety of methods and to trace characters on phylogenetic trees
- Know how to evaluate the robustness of phylogenetic hypotheses

## **Course Contents**

### **Unit 1:**

**(33 % Weightage)**

Foundations in Molecular Evolution: Evolution, natural selection, convergent evolution, co-evolution. Evolution of genes in populations. Hardy-Weinberg equilibrium. Mutation; Migration; Selection; Genetic drift; measures of genetic diversity. The neutral and nearly-neutral theories of molecular evolution. Evolution of genome and gene families: Lateral gene transfer and transposition. Chromosomal evolution: Genome duplications, Domain shuffling, concerted evolution and molecular drive Calculating evolutionary distances among sequences; correlation and models.

### **Unit 2:**

Molecular clocks. Molecular phylogenetics: Concepts, types of trees, rooting, clades, reconstructing character evolution, consensus trees. Phylogenetic analysis algorithms: Distance-based: UPGMA, Neighbor-Joining, Maximum Parsimony, Reliability of trees: Bootstrap and randomization tests. Phylogenetic trees and their comparison: Definition and description, various types of trees; Consensus tree, Probabilistic models of evolution, Maximum likelihood algorithm.

### **List of Practical:**

- Compute amino acid and nucleotide composition of sequences.
- Compute codon usage bias.
- Estimate transition transversion bias.
- Translating nucleotides sequence into amino acids sequence
- Construction of distance-based tree, UPGMA and NJ using MEGA
- Construction of Character based tree, Maximum parsimony using MEGA

- Construction of Maximum likelihood tree using MEGA
- Analyze user tree by maximum parsimony and maximum likelihood
- Bootstrap analyses.
- Estimating divergence times for sequences.
- Infer ancestral sequences in a parsimony tree.
- Tajima's test of neutrality and relative rate test
- Estimate gamma parameter for site rate.
- Test molecular clock

**Content Interaction Plan:**

<b><u>Lecture cum Discussion (Each session of 1 Hour)</u></b>	<b><u>Unit/Topic/Sub-Topic</u></b>
(Unit I)1-2	Concept of organic Evolution, theories of evolution, natural selection
2-4	Convergent evolution, co-evolution. Evolution of genes in populations.
5-6	Hardy-Weinberg equilibrium.
7-9	Mutation; Migration; Selection; Genetic drift; measures of genetic diversity
10-11	The neutral and nearly-neutral theories of molecular evolution.
(Unit II )12-14	Evolution of genome and gene families: Lateral gene transfer and transposition
15-16	Chromosomal evolution: Genome duplications, Domain shuffling, concerted evolution and molecular drive
<b>CONTINUOUS ASSESSMENT: 01</b>	
17-18	Calculating evolutionary distances among sequences; correlation and models. Molecular clocks hypothesis.
19-20	Concepts, types of trees, rooting, clades, reconstructing character evolution, consensus trees.
21-23	Calculating evolutionary distances among sequences; correlation and models & Molecular clocks.
24-26	Concepts, types of trees, rooting, clades, reconstructing character evolution, consensus trees.
<b>CONTINUOUS ASSESSMENT: 02</b>	

27-29	Phylogenetic analysis algorithms: Distance-based: UPGMA, Neighbor-Joining
30-32	Maximum Parsimony, Maximum likelihood
33-35	Reliability of trees: Bootstrap, randomization tests.
36-39	Probabilistic models and associated algorithms: Probabilistic models of evolution, Maximum likelihood algorithm.
	<b>CONTINUOUS ASSESSMENT: 03</b>
15 Hours	<b>Tutorial</b>
30 Hours	<b>Practical</b>
<p><b><u>Suggested References:</u></b></p> <ol style="list-style-type: none"> <li>1 Dan G, Wen-Hsiung L (2001) Fundamentals of molecular evolution. Willey-VCH</li> <li>2 Bromham L (2016) An Introduction to Molecular Evolution and Phylogenetics. Oxford University Press</li> <li>3 Marco S, Anne-Mieke V (2003) The phylogenetic handbook: a practical approach to DNA and protein phylogeny. Cambridge University Press</li> <li>4 Rastogi VB (2021) Organic Evolution, 14<sup>th</sup> edition, Meditech publisher</li> <li>5 Bernardi G (2005) Structural and evolutionary genomics, Volume 37: natural selection in genome evolution (new comprehensive biochemistry). Elsevier Science, Netherland</li> <li>6 Oren A (2010) Molecular phylogeny of Microorganism. Academic press</li> <li>7 Nei M, Kumar S (2000) Molecular Evolution and Phylogenetics. Oxford University Press</li> </ol>	

## PROGRAMMING WITH PERL

<b>Course Details</b>			
<b>Course Title:</b> Programming with Perl			
<b>Course Code</b>	BIS 8 2 DE 010 04	<b>Credits</b>	4
<b>L + T + P</b>	2 + 1 + 1	<b>Course Duration</b>	One Semester
<b>Semester</b>	I	<b>Contact Hours</b>	30 (L) + 15 (T) + 30 (P) Hours
<b>Nature of the Course</b>	Theory cum Practicum		
<b>Special Nature/ Category of the Course (if applicable)</b>	Skill Based		
<b>Methods of Content Interaction</b>	Lecture, Tutorials, Group discussion; self-study, seminar, presentations by students, individual and group drills, group and individual field-based assignments followed by workshops and seminar presentation.		
<b>Assessment and Evaluation</b>	<ul style="list-style-type: none"> <li>• 30% - Continuous Internal Assessment (Formative in nature but also contributing to the final grades)</li> <li>• 70% - End Term External Examination (University Examination)</li> </ul>		

### **Course Objectives**

- To acquaint the students with basic concepts of Perl scripting.
- Enable students to write Perl scripts for bioinformatics applications.

### **Learning Outcomes**

After completion of the course the learners will be able to:

- On successful completion, students will be equipped with designing and writing Perl scripts that will help in their work and research.

### **Course Contents**

#### **UNIT I: Introduction to Perl, Scalar, Array and List Data (25% Weightage)**

- Introduction to Perl, Use of Perl in Bioinformatics, History of Perl, Availability, Support, Basic Concepts.

- Introduction to Scalar Data, Numbers, Strings, Scalar Variables, Scalar Operators and Functions.
- Introduction to Arrays and List Data, Literal Representation, Variables, Array Operators and Functions, Scalar and List Context.

**UNIT II: Control Structures and Basic I/O File handles (25% Weightage)**

- Control Structures: Statement Blocks, Introduction to Hashes, Hash Variables, Literal Representation of a Hash, Hash Functions, Hash Slices.
- Basic I/O File handles and File Tests, File Manipulation, Directory Manipulation. Formatting Data: Sorting, Transliteration.

**UNIT III: Regular Expressions and Subroutines in Perl (25% Weightage)**

- Regular Expressions: Concepts about Regular Expressions, Simple uses of Regular Expressions, Patterns, Matching Operator, Substitutions, the split and join Functions.
- Subroutines: System and User Functions, The local Operator, Variable-length Parameter Lists, Notes on Lexical Variables.

**UNIT IV: Database Manipulation, CGI Programming and Bioperl (25% Weightage)**

- Database Manipulation: DBM Databases and DBM Hashes, Opening and Closing DBM Hashes, Fixed-Length Random-Access Databases, Variable-Length (Text) Databases, Win32 Database Interfaces.
- CGI Programming: The CGI.pm Module, Your CGI Program in Context, Simplest CGI Program, Passing Parameters via CGI, Perl and the Web.
- Bioperl: Introduction, Installation procedures, Architecture, Uses of Bioperl.

**List of Practicals**

- Basics of Perl scripting: declaration, initialization and printing: Scalar Variable, Array Variable and Hash Variables.
- Manipulation of array variable using Join, Split, Splice, Sort, Merging, Pop, Push, Shift and Unshift function.
- Concatenating DNA fragment.
- Calculating the number of bases in the DNA string.
- Transcribing DNA to RNA.
- Reverse Complement of DNA string.
- Searching motifs.
- Using loops in perl scripting: for, while, do- while, foreach, until, unless etc.
- Counting DNA bases present in file containing DNA string.
- Using Subroutines in perl scripting.

- Finding STOP codons in a DNA string.
- Mutating DNA string at random positions.
- Generating random DNA sequence of specific length.
- Counting nucleotide in DNA string using subroutine.
- Calculating average percentage identity between pairs of random DNA sequence.
- Translation of DNA string using regular expression and hashes.
- Extracting DNA from FASTA files using regular expression.
- Extracting Data from a local RDBMS database server.

**Content Interaction Plan:**

<b><u>Lecture cum Discussion (Each session of 1 Hour)</u></b>	<b><u>Unit/Topic/Sub-Topic</u></b>
<b>1-8</b>	<b>UNIT I: Introduction to Perl, Scalar, Array and List Data</b>
1-2	Introduction to Perl, Use of Perl in Bioinformatics, History of Perl, Availability, Support, Basic Concepts.
3-5	Introduction to Scalar Data, Numbers, Strings, Scalar Variables, Scalar Operators and Functions.
6-8	Introduction to Arrays and List Data, Literal Representation, Variables, Array Operators and Functions, Scalar and List Context.
<b>9-14</b>	<b>UNIT II: Control Structures and Basic I/O File handles</b>
9-11	Control Structures: Statement Blocks, Introduction to Hashes, Hash Variables, Literal Representation of a Hash, Hash Functions, Hash Slices.
12-14	Basic I/O File handles and File Tests, File Manipulation, Directory Manipulation. Formatting Data: Sorting, Transliteration.
<b>15-21</b>	<b>UNIT III: Regular Expressions and Subroutines in Perl</b>
15-17	Regular Expressions: Concepts about Regular Expressions, Simple uses of Regular Expressions, Patterns, Matching Operator, Substitutions, the split and join Functions.
18-21	Subroutines: System and User Functions, The local Operator, Variable-length Parameter Lists, Notes on Lexical Variables.
<b>22-30</b>	<b>UNIT IV: Database Manipulation, CGI Programming and Bioperl</b>

22-24	Database Manipulation: DBM Databases and DBM Hashes, Opening and Closing DBM Hashes, Fixed-Length Random-Access Databases, Variable-Length (Text) Databases, Win32 Database Interfaces.
25-27	CGI Programming: The CGI.pm Module, Your CGI Program in Context, Simplest CGI Program, Passing Parameters via CGI, Perl and the Web.
28-30	Bioperl: Introduction, Installation procedures, Architecture, Uses of Bioperl.
15 Hours	Tutorials
30 Hours	Practicals

**Essential Readings:**

- Bal HP (2003) Perl programming for Bioinformatics. Tata McGraw-Hill
- Curtis JD (2003) Perl programming for biologists. John Wiley & Sons
- Schwartz RL, Phoenix T (2011) Learning Perl. O'Reilly Media, Inc.
- Tisdall JD (2001) Beginning Perl for bioinformatics. O'Reilly Media, Inc.

**Additional/Advance/Further Readings:**

- Moorhouse M, Barry P (2004) Bioinformatics Biocomputing and Perl: an introduction to bioinformatics computing skills and practice. John Wiley & Sons
- Tisdall JD (2003) Mastering Perl for bioinformatics. O'Reilly Media, Inc.
- Guelich S, Gundavaram S, Birznieks G (2000) CGI Programming with Perl. O'Reilly Media, Inc.



## ICT AND DIGITAL SKILLS

<b>Course Code</b>	BIS 8 2 ME 011 00	<b>Credits</b>	02
<b>L + T + P</b>	2 + 0 + 0	<b>Course Duration</b>	One Semester
<b>Semester</b>	II	<b>Contact Hours</b>	20 Hours
<b>Course Type</b>	Non-credit		
<b>Nature of the Course</b>	Theory, Hands-on		
<b>Methods of Content Interaction</b>	Lecture, Hands-on exercise, Short assignments		
<b>Assessment and Evaluation</b>	<ul style="list-style-type: none"><li>● Student would be given assignment/short-project and based on performance, satisfactory/unsatisfactory grade would be provided.</li></ul>		

### ❖ **Course Objectives:**

- ❖ To Develop essential Digital and ICT skills,
- ❖ To learn basic skills required for essential word, numerical and graphical processing of data
- ❖ Learn basic skills related with online activities and cloud services

### ❖ **Course Learning Outcomes:**

After completion of the course the students will be able to:

- ❖ Learn tools required for word and data processing and data visualization
- ❖ Learn how to use cloud services and online tools

### ❖ **Course Contents:**

#### **UNIT I: ICT tools**

**(25% Weightage)**

- Introduction of IT, ICT and IoT
- Numerical processing of data using Excel and other software
- Word processing tools and MS/Libre office essentials
- Graphical Processing & data visualization using Matlab, Sigmaplot, Origin, Tablue & other software
- Essential Unix commands, awk, grep utilities, installation in Unix
- Virtualization- desktop virtualization, application virtualization

## **UNIT II: Online Tools & Cloud services**

**(25% Weightage)**

- Use of Web and multimedia tools, Search Engine and optimization, Network essentials and Firewall
- Webpage design and editing, domain registration and web hosting, Online and software tools - Weebly, Wordpress etc.
- Using Cloud services (Adobe creative cloud, Google, Amazon cloud etc.)

### **Suggested Readings:**

- Microsoft Excel Manual (2015) University of Mary Washington Verginia
- Rui Costa, R (2021) Google Cloud Cookbook: Practical Solutions for Building and Deploying Cloud Services, O'Reilly Media, Inc. USA
- Smith J, Smith C (2021) Adobe Creative cloud, Wiley USA
- Williams A (2020). WordPress for Beginners
- Kusnetzky D (2011). Virtualization, O'Reilly Media, Inc.

## Course Title: Chemoinformatics and Drug Design

<b>Course Code</b>	BIS 9 1 DC 012 04	<b>Credits</b>	4
<b>L + T + P</b>	3 + 0 + 1	<b>Course Duration</b>	One Semester
<b>Semester</b>	III	<b>Contact Hours</b>	45 (L) + 30(P) Hours
<b>Course Type</b>	Discipline Based Core		
<b>Nature of the Course</b>	Theory/Practical		
<b>Special Nature/ Category of the Course (if applicable)</b>	Skill Based and Indian Knowledge System (traditional medicines)		
<b>Methods of Content Interaction</b>	Lecture, Tutorials, Group discussion; self-study, seminar, presentations by students, individual and group drills, group and individual assignments followed by research paper presentation.		
<b>Assessment and Evaluation</b>	<ul style="list-style-type: none"><li>• 30% - Continuous Internal Assessment (Formative in nature but also contributing to the final grades)</li><li>• 70% - End Term External Examination (University Examination)</li></ul>		

### Course Objectives

- ❖ To train students in the area of drug discovery & Chemoinformatics for pharmaceutical research. The course is designed to provide a basic understanding of computational drug design and the applicability of tools for novel drug discovery.

### Learning Outcomes

- ❖ Students would be able to get insights into the modern drug discovery and development process. Students would be able to develop basic skills in drug design and apply them in the pharmaceutical industry.

### Course Contents

#### UNIT I: Chemoinformatics & chemical data formats (25% Weightage)

- Introduction to Chemoinformatics and the role of Chemoinformatics in the pharmaceuticals
- Linear representation of small molecules, Line notation (SMILES, InChI, WLN, ROSDAL, SLN)
- Matrix and Graph representation of small molecules and connection table
- Canonicalization of small molecules; Morgan algorithm,
- Mol/SDF and PDB file formats
- Special notation of small molecules( Makush structure, Fingerprints, and hash notation),
- Molecular surfaces Van der Waal and Connolly surface
- 2D and 3D structures generation through CORINA, conformational analysis, and ensemble conformation,
- File format conversion (open babel),
- Internal coordinates/Z matrix, orthogonal coordinates, and
- Conformation Search Structure-based Search techniques: Exact, Substructure, and similar structure search from Zinc database.

#### UNIT II: Ligand-based Drug Design QSAR and case studies (25% Weightage)

- Ligand-based Drug Design: IC<sub>50</sub>, pIC<sub>50</sub>, K<sub>i</sub>, and LD<sub>50</sub>,

- 1D, 2D, and 3D QSAR descriptors and shape-based descriptors
- History of QSAR Graphs and equations and statistical data for QSAR  
Physiochemical properties: Hydrophobicity, partition coefficient(P), substituent hydrophobicity constant Pi,
- A case study of the binding of the molecule with human serum albumin,
- A case study of general anesthetics
- Electronic effect sigma, Taft steric factors(Es), and molar refractivity MR
- Hansch equation for antimalarial molecule and Crag plot
- 3-DQSAR/Field based QSAR CoMFA and CoMSIA, defining electrostatic and steric fields, A case study of colchicine.
- Pharmacophore features and definition, Pharmacophore modeling, Searching databases using pharmacophores,

**UNIT III: Structure-Based Drug Design: (25 % Weightage)**

- Structure-Based Drug Design:
- Molecular Targets, binding site/ cavity determination, and analysis
- Virtual screening, docking search parameter, and docking scoring functions,
- Quantum polarized docking, rigid and flexible docking
- Preparation of ligand and protein
- Hit/Lead identification and optimization
- The general principle of de novo drug design
- Automated de novo design principle of Ligbuilder, LUDI, and SPROUT
- A case study of *de novo* design
- Database/library of small molecules, ZINC, Maybridge Pubchem, Asinex
- Library generation (focused & diverse)
- Combinatorial & Fragment libraries: Molecular Scaffolds, combinatorial library generation, Scaffold hopping.

**UNIT IV: *In Silico* ADME screening, Preclinical trials, and patent filing (25% Weightage)**

- Approaches & various phases of novel drug discovery,
- Pharmacokinetics: introduction to drug absorption, disposition, metabolism, elimination, toxicity (ADMET),
- *In silico* screening, Lipinski and Weber's rule,
- ADMET and physiochemical properties prediction,
- Prediction of the site of metabolism (SOM),
- the concept of prodrug & soft drug.
- Pharmacodynamics: mechanism of drug action, dose-response relationship, drug potency and efficacy,
- Preclinical trials and clinical trials
- Patenting and regulatory affairs
- Introduction of Ayurvedic and other herbal drugs prevalent in the traditional systems and their databases. Molecular approaches to understand the mechanism of herbal drugs

**List of Practical**

- Practical session for chemical structure representation and storage in particular file formats (SMILES, WLN, sd, and mol)
- Importance of 3D structures and method of generation from 1D & 2D representations
- Substructure/Exact/similar structure-based searching
- Quantitative Structure-Activity Relationship Studies
- Pharmacophore hypothesis and searching

- Virtual screening and Docking studies (Rigid, Flexible & library based)
- De-novo drug designing through Lig-Builder.

<u>Lecture cum Discussion</u> (Each session of 1 Hour)	<u>Unit/Topic/Sub-Topic</u>
1-2	<b>UNIT I: Chemoinformatics &amp; chemical data formats</b> Introduction to Chemoinformatics, Role of Chemoinformatics in the pharmaceuticals, Linear representation of small molecules, Line notation (SMILES, InChI, WLN, ROSDAL, SLN)
3-7	Matrix and Graph representation of small molecules, Connection table, Canonicalization of small molecules, Morgan algorithm, Mol/SDF, and PDB file formats
8	Special notation of small molecules( Makush structure, Fingerprints, and hash notation),
9-10	Molecular surfaces Van der Waal and Connolly surface 2D and 3D structures generation through CORINA, conformational analysis, Ensemble conformation, File format conversion (open babel), internal coordinates/Z matrix, orthogonal coordinates, and Conformation Search Structure-based Search techniques: Exact, Substructure, and similar structure search from Zinc database
11-12	<b>Ligand-based Drug Design, QSAR, and Case Studies</b> Ligand-based Drug Design: IC <sub>50</sub> , pIC <sub>50</sub> , Ki, LD <sub>50</sub> , 1D, 2D, and 3D QSAR descriptors, shape-based descriptors, History of QSAR, Graphs and equations, statistical data for QSAR
13-15	physiochemical properties: Hydrophobicity, partition coefficient(P), substituent hydrophobicity constant Pi, a case study of the binding of the molecule with human serum albumin, a case study of general anesthetics, electronic effect sigma, Taft steric factors(E <sub>s</sub> ), Molar refractivity MR, Hansch equation for antimalarial molecule, Crag plot,
16-17	3-DQSAR/Field based QSAR CoMFA and CoMSIA, defining electrostatic and steric fields, A case study of colchicine.
18-19	Pharmacophore features and definition, Pharmacophore modeling, Searching databases using pharmacophores,
20	<b>UNIT III: Structure-Based Drug Design</b> Molecular Targets, binding site/ cavity determination and analysis (CastP and Cavity program),
21-23	virtual screening, docking search parameter and scoring functions, rigid and flexible docking, quantum polarized docking
24-25	preparation of ligand and protein, Hit/Lead identification, and optimization
26-30	the general principle of <i>de novo</i> drug design, automated <i>de novo</i> design principle of Ligbuilder, LUDI, and SPROUT, case study,
31-32	database/library of small molecules, ZINC, Maybridge Pubchem, Asinex, Library generation (focused & diverse). Combinatorial & Fragment libraries: Molecular Scaffolds, Scaffold hopping.
33-34	Combinatorial library generation

35-36	<b>UNIT IV: <i>In Silico</i> ADME screening, Preclinical trials, and patent filing</b> Approaches & various phases of novel drug discovery, pharmacokinetics: introduction to drug absorption, disposition, metabolism, elimination, toxicity (ADMET), <i>In silico</i> screening, Weber's rule,
38-40	ADMET and physiochemical properties prediction, prediction of the site of metabolism (SOM), the concept of prodrug & soft drug.
41-42	Pharmacodynamics: mechanism of drug action, dose-response relationship, drug potency, and efficacy
43-45	Preclinical trials, clinical trials, patenting, and regulatory affairs Introduction of Ayurvedic and other herbal drugs prevalent in the traditional systems and their databases. Molecular approaches to understand the mechanism of herbal medicines.
<b>Suggested References:</b>	
<ol style="list-style-type: none"> <li>1. Johann G, Thomas E (2004) Chemoinformatics: A textbook.: Willey-VCH</li> <li>2. Graham L. Patrick, (2018) An introduction to medicinal chemistry</li> <li>3. Holtje HD, Sippl D, Rognan, Folkers G (2008) Molecular modelling basic Principles and applications (3rd Edition). Willey-VCH</li> </ol>	
<b>Additional references:</b>	
<ol style="list-style-type: none"> <li>1. Leach AR, Gillet VJ (2003) An introduction to Chemoinformatics. Kluwer Academic.</li> <li>2. Osman F Güner (2000) Pharmacophore Perception, Development, and Use in Drug Design. International University Line</li> <li>3. Thomas L, Raimund M, Hugo K, Hendrik T (2002) Bioinformatics: From genomes to drugs Volume 1: Basic technologies. Willey-VCH</li> <li>4. Young DC (2009) Computational drug design: a guide for computational and medicinal chemists: John Wiley &amp; Sons</li> <li>5. Barry AB, Brian S, Guillermo M, Jürgen B (2006) Chemoinformatics: Theory, practice, &amp; products. Springer</li> </ol>	

## DBMS AND WEB TECHNOLOGY

Course Details			
<b>Course Title:</b> DBMS and Web Technology			
<b>Course Code</b>	BIS 9 1 DC 013 04	<b>Credits</b>	4
<b>L + T + P</b>	2 + 1 + 1	<b>Course Duration</b>	One Semester
<b>Semester</b>	III	<b>Contact Hours</b>	30 (L) + 15 (T) + 30 (P) Hours
<b>Nature of the Course</b>	Theory cum Practicum		
<b>Special Nature/ Category of the Course (if applicable)</b>	Skill Based		
<b>Methods of Content Interaction</b>	Lecture, Tutorials, Group discussion; self-study, seminar, presentations by students, individual and group drills, group and individual field based assignments followed by workshops and seminar presentation.		
<b>Assessment and Evaluation</b>	<ul style="list-style-type: none"><li>• 30% - Continuous Internal Assessment (Formative in nature but also contributing to the final grades)</li><li>• 70% - End Term External Examination (University Examination)</li></ul>		

### Course Objectives

- This course is intended to provide students with an in-depth understanding of the database management systems and to equip them with design and implementation of database applications.

### Learning Outcomes

After completion of the course the learners will be able to:

- Understand basic database concepts of the relational data model. They will be able to construct database queries using SQL and designing web applications using PHP and HTML.

### Course Contents

#### **UNIT I: Introduction to databases, Data models and E-R model (25%**

#### **Weightage)**

- Introduction to databases, Database designing, data acquisition, Data Abstraction.
- Data Models: Relational Data Model, Network Data Model, Hierarchical Data Model.
- E-R Model: Entity and entity sets, Relations and relationship sets, E-R diagrams, Reducing E-R Diagrams to tables.

**UNIT II: Data normalization, Indexing & Hashing and Data mining (25%**

**Weightage)**

- Data normalization, Basic concepts of Indexing and Hashing. Textual Databases, Introduction to Distributed Database Processing, Data warehousing and Data mining.
- Organizing biological species information using various database techniques. Need of meta-data standards.

**UNIT III: MySQL Architecture and SQL statements (25%**

**Weightage)**

- MySQL Architecture. Objects - Tables, Views, Indexes, Sequences; Synonyms, Snapshots, Clusters. Database - Table space, Data files, Blocks, Extents, Segments; SQL DBA - Export-Import/SQL Monitor Backup & Recovery (Archiving), Physical Storage & Logical Storage.
- Select Statements, Data Definition Statements, Data Manipulation Statements, Data Control Statements.

**UNIT IV: Introduction to Web technology, HTML and PHP (25%**

**Weightage)**

- Introduction to Web Technology, HTML, Basic Tags, CSS, Table and Forms.
- Introduction to PHP, Syntax, Operators, Variables, Constants, Control structure, Language construct and functions.
- Arrays-Enumerated Arrays, Associative array, array iteration, Multi-dimensional array.
- Array function, Date and Time functions, Web Features- Sessions, Forms, GET and POST data, database programming.

**List of Practicals**

- Based on “Data Definition Language”. A set of SQL commands used to create table, modify table structure, drop table, rename table.
- Based on “Data Manipulation Language” A set of SQL commands used to change the data within the database. It consists of inserting of records in the tables, updation of all or specific set of records in tables, viewing the attributes of table’s column.
- Based on “Data Query Language” It allows extracting the data out of the database. Selecting the data from table using Arithmetic and logical operators. Range searching and pattern matching. Function, group function, scalar function.
- Based on defining Constraints. Types: I/O constraints like Primary Key, Foreign key,



Null and Unique constraints. Business constraints like check constraints. Levels: Table level constraints, column level constraints, creating and deletion of constraints using the Alter Table clause.

- Based on using joins. Joining multiple tables, joining a table to itself.
- Based on using Indexes, Sequences.
- Based on Security Management using SQL, Granting rights on user objects such as Tables, Views, and Sequences. Revoking rights on user objects such as Tables, Views, and Sequences.

**Content Interaction Plan:**

<b><u>Lecture cum Discussion (Each session of 1 Hour)</u></b>	<b><u>Unit/Topic/Sub-Topic</u></b>
<b>1-6</b>	<b>UNIT I: Introduction to databases, Data models and E-R model</b>
1-2	Introduction to databases, Database designing, data acquisition, Data Abstraction.
3-4	Data Models: Relational Data Model, Network Data Model, Hierarchical Data Model.
5-6	E-R Model: Entity and entity sets, Relations and relationship sets, E-R diagrams, Reducing E-R Diagrams to tables.
<b>7-13</b>	<b>UNIT II: Data normalization, Indexing &amp; Hashing and Data mining</b>
7-11	Data normalization, Basic concepts of Indexing and Hashing. Textual Databases. Introduction to Distributed Database Processing. Data warehousing and Data mining.
12-13	Organizing biological species information using various database techniques. Need of meta-data standards.
<b>14-20</b>	<b>UNIT III: MySQL Architecture and SQL statements</b>
14-16	MySQL Architecture. Objects - Tables, Views, Indexes, Sequences; Synonyms, Snapshots, Clusters. Database - Table space, Data files, Blocks, Extents, Segments; SQL DBA - Export-Import/SQL Monitor Backup & Recovery (Archiving), Physical Storage & Logical Storage.
17-20	Select Statements, Data Definition Statements, Data Manipulation Statements, Data Control Statements.

<b>21-30</b>	<b>UNIT IV: Introduction to Web technology, HTML and PHP</b>
21-23	Introduction to Web Technology, HTML, Basic Tags, CSS, Table and Forms.
24-26	Introduction to PHP, Syntax, Operators, Variables, Constants, Control structure, Language construct and functions.
27-28	Arrays-Enumerated Arrays, Associative array, array iteration, Multi-dimensional array.
29-30	Array function, Date and Time functions, Web Features- Sessions, Forms, GET and POST data, database programming.
<i>15 Hours</i>	<i>Tutorials</i>
<i>15 Hours</i>	<i>Practicals</i>

**Essential Readings:**

- Date CJ (1999) An Introduction to Database Systems (6th edition). Addison Wesley
- Silberschatz A, Korth HF, Sudarshan S (2002) Database system concepts (4th Edition). McGraw-Hill Education.
- Guleria P (2018) Php Beginner'S Practical Guide. BPB Publications.
- Tahaghoghi Saied MM, Williams HE (2006) Learning MySQL. O'Reilly Media, Inc.
- Matthews M (2015) Php And Mysql Web Development: A Beginner's Guide. McGraw Hill .

**Additional/Advance/Further Readings:**

- Ivan B (2003) SQL, PL/SQL The Programming Language of Oracle (2nd Edition). BPB Publications.
- Nixon R (2006) Learning PHP, MySQL, JavaScript, CSS & HTML5 3ed: A Step-by-Step Guide to Creating Dynamic Websites. O'Reilly Media, Inc.
- Vaswani V (2017) MySQL(TM): The Complete Reference. McGraw Hill Education.

## Applications of Bioinformatics

Course Details			
<b>Course Title: Applications of Bioinformatics</b>			
<b>Course Code</b>	BIS 9 1 DC 014 04	<b>Credits</b>	4
<b>L + T + P</b>	0 + 0 + 4	<b>Course Duration</b>	One Semester
<b>Semester</b>	III	<b>Contact Hours</b>	--
<b>Nature of the Course</b>	Practical		
<b>Special Nature/ Category of the Course (if applicable)</b>	Skill Based, Value added		
<b>Methods of Content Interaction</b>	<ul style="list-style-type: none"> <li>• Discussion, software execution, report preparation and data analysis skills</li> </ul>		
<b>Assessment and Evaluation</b>	<ul style="list-style-type: none"> <li>• 30% - Continuous Internal Assessment by supervisor</li> <li>• 70% - End Term External Examination (Presentation and viva to be evaluated by a committee of faculties)</li> </ul>		

### Course Objectives

- To impart biological problem-solving skills using computational approaches.

### Learning Outcomes

- The student learn to identify biological problems and find computational solutions for the same.

## Systems Biology (Elective)

Course Details			
<b>Course Title: Systems Biology</b>			
<b>Course Code</b>	BIS 9 1 DE 015 04	<b>Credits</b>	4
<b>L + T + P</b>	2 + 1 + 1	<b>Course Duration</b>	One Semester
<b>Semester</b>	III	<b>Contact Hours</b>	30 (L) +15 (T)+ 30 (P)
<b>Nature of the Course</b>	Theory cum Practical		
<b>Special Nature/ Category of the Course</b>	Skill Based		
<b>Methods of Content Interaction</b>	Lecture, Tutorials, Group discussion; self-study, seminar, presentations by students, individual and group drills, group and individual field based assignments followed by paper presentation.		
<b>Assessment and Evaluation</b>	<ul style="list-style-type: none"> <li>• 30% - Continuous Internal Assessment (Formative in nature but also contributing to the final grades)</li> <li>• 70% - End Term External Examination (University Examination)</li> </ul>		

### Course Objectives

This course will provide an introduction to systems biology by focusing on the behaviours expected from interactions between only a few genes, taking examples from microbes to mammals. Cells are dynamic systems, and we will build intuition about the types of responses expected from different gene circuits by running, adapting, and analysing computer simulations. Throughout, the course will use such simulations and analysis as research tools to understand biology. After an introduction to motifs and modules, we will focus on the role of feedback in genetic networks and how feedback can sometimes create permanent switches, in, for example, stem cells, or at other times can generate oscillations such as circadian rhythms in neurons. We will show how these behaviors can be undermined when numbers of molecules become low, an effect that cells may exploit or regulate away. Finally, we discuss experimental techniques that allow direct comparison between simulations and real biological systems. Finally overall aims is to understand, design principles of biological circuits i.e. transcriptional, signal transduction, neuronal and developmental networks in unified way.

### Learning Outcomes

- Students will gain an appreciation of how interactions between genes can explain some of the behaviour we see in cells.

- Students will gain an understanding of the different behaviours expected in dynamical systems and how to biochemically code for some of these behaviours.
- Students will develop skills in simulation and learn how to use computers as tools to help decide between different hypotheses

## Course Contents

### UNIT I: The Assessment and Evaluation Process (50% Weightage)

Gene Networks basic concepts, Transcription Networks: Basic concepts, Auto-regulation a network motif AR and NAR , The Feed-Forward Loop Network Motif, Coherent and Incoherent FFL, Temporal programme and the global structure of transcription network, Network motifs in developmental, signal transduction and neuronal networks, Genetic Switches –I and II.

### UNIT II: Tools of Assessment and Evaluation (50% Weightage)

FIFO, DOR, Global, Development, memory and irreversibility- signaling networks and neuron circuits-robustness, Robustness of protein circuits: The example of bacterial Chemotaxis, Robust patterning in development, Kinetic proofreading, Optimal Gene circuit design, Demand rules for gene regulation, Epilogue: simplicity in biology, Cell Designer and SBML. Pathway Mapping through KEGG, VCell/ Virtual Cell.

### List of practicals

- Study of models in systems Biology, Defining model, Reaction Network, reaction kinetics, Kinetic law, parameter, and simulation
- Simulation and analysis of transcription network, signal transduction network, developmental network and metabolic network in cell designer/ cytoscape. Interpretations of simulation results biological context.
- Map Kinase and M Phase of cell Division Pathways simulation and interpretation.

### Content Interaction Plan:

<u>Lecture cum Discussion (Each session of 1 Hour)</u>	<u>Unit/Topic/Sub-Topic</u>

Unit-1 1-5	Gene Networks basic concepts, Transcription Networks: Basic concepts, Auto-regulation a network motif AR and NAR , The Feed-Forward Loop Network Motif.
6-10	Temporal programme and the global structure of transcription network,
11-13	Coherent and In coherent FFL, Network motifs in developmental, signal transduction and neuronal networks.
14-15	Genetic Switches –I and II
Unit-2 16-17	FIFO, DOR, Global, Development
17-18	Memory and irreversibility of developmental network
19-20	signaling networks and neuron circuits
21-23	,Robustness of protein circuits: The example of bacterial Chemotaxis,
23-25	robustnessRobust patterning in development,
25-27	Kinetic proofreading,
27-28	Optimal Gene circuit design,
29	Demand rules for gene regulation, Epilogue: simplicity in biology,
30	Pathway Mapping through KEGG, VCell/ Virtual Cell.
15 Hours	Tutorials
<p><u>Suggested References:</u></p> <ul style="list-style-type: none"> <li>• An introduction to systems biology, u alon (chapman &amp; hall, 2006)</li> <li>• First steps in computational systems biology: a practical session in metabolic modeling and simulation*<i>biochemistry and molecular biology education</i> vol. 37, no. 3, pp. 178–181, 2009 <a href="https://iubmb.onlinelibrary.wiley.com/doi/pdf/10.1002/bmb.20281">https://iubmb.onlinelibrary.wiley.com/doi/pdf/10.1002/bmb.20281</a></li> <li>• First steps in computational systems biology: a practical session in metabolic modeling and simulation. <a href="https://www.ncbi.nlm.nih.gov/pubmed/21567728">https://www.ncbi.nlm.nih.gov/pubmed/21567728</a></li> </ul>	

## Exome Sequencing Data Analysis (Elective)

<b>Course Details</b>			
<b>Course Title: Exome Sequencing Data Analysis</b>			
<b>Course Code</b>	BIS 9 1 DE 016 04	<b>Credits</b>	4
<b>L + T + P</b>	2 + 1 + 1	<b>Course Duration</b>	One Semester
<b>Semester</b>	III	<b>Contact Hours</b>	30 (L) + 15 (T) + 30 (P) Hours
<b>Nature of the Course</b>	Theory cum Practicum		
<b>Special Nature/ Category of the Course (if applicable)</b>	Skill Based		
<b>Methods of Content Interaction</b>	Lecture, Tutorials, Practical, Group discussion; self-study, assignment class based as well as take-home type by students		
<b>Assessment and Evaluation</b>	<ul style="list-style-type: none"> <li>• 30% - Continuous Internal Assessment (Formative in nature but also contributing to the final grades)</li> <li>• 70% - End Term External Examination (University Examination)</li> </ul>		

### Course Objectives

- To acquaint the students with the basic concepts and practices adopted for whole exome sequencing data analysis.
- To help the students understand, how WES helps to bridge cutting edge science and the delivery of genomic medicine in clinical practice.
- To orient the students with tools and techniques of WES data analysis.
- To develop skills and competencies to analyze and interpret WES data.
- To make the students understand how genomic medicine has impacted on a patients diagnosis, care and or treatment.

### Learning Outcomes

After completion of the course the learners will be able to:

- Describe the wider issues affected by whole exome sequencing and Bioinformatics.

- Define the challenges and opportunities facing bioinformaticians in WES data analysis.
- Describe the steps involve in WES data analysis workflow.
- Explain how WES data is taken from the computational level to the human level
- Describe the importance of accuracy and precision in genomic data analysis.
- Evaluate quality of read sequences generated by NGS experiments.
- Align read sequences to reference genome.
- Perform variant calling given aligned read sequences.
- Triaging the variants.

## **Course Contents**

### **UNIT I: Whole Exome sequencing data acquisition (25% Weightage)**

- Exome sequencing techniques: Library preparation, Exome enrichment methods, cluster generation, sequencing and base calling.
- File format for sequencing data: FASTQ, BAM, SAM, VCF.

### **UNIT II: Exome data processing (25 % Weightage)**

- Exome data analysis pipeline
- Quality control for raw read data: Phred quality score, FASTQC tool
- Tools for sequence read alignment: HISAT2 and Bowtie.
- Refining the alignment

### **UNIT III: Variant detection (10 % Weightage)**

- Variant calling algorithms, variant annotation
- Removing common variants.
- Databases that capture the frequency of different variations in the population.

### **UNIT IV: Exome sequencing and ethical issues (10% Weightage)**

- What is ethics and governance? Who owns the data?
- Difference between privacy and confidentiality.
- Data sharing and the law



**List of Practicals****(30% Weightage)**

- Examining file format to store NGS data
- Performing quality control checks on raw sequencing reads
- Aligning reads to reference genome
- Variant calling
- Filtering common variants
- Generating reports for clinicians

**Content Interaction Plan:**

<b><u>Lecture cum Discussion (Each session of 1 Hour)</u></b>	<b><u>Unit/Topic/Sub-Topic</u></b>
1-2	Exome sequencing over view
3-6	Library preparation, Exome enrichment methods
7-9	cluster generation, sequencing and base calling Continuous assessment (Class Test/Assignment I)
10-14	File format for sequencing data: FASTQ, BAM, SAM, VCF
15-16	Exome data analysis pipeline
17-18	Quality control for raw read data: Phred quality score, FASTQC tool
19-20	Tools for sequence read alignment: HISAT2 and Bowtie Continuous assessment (Class Test/Assignment II)
21	Refining the alignment
22-24	Variant calling algorithms, variant annotation
25-26	Removing common variants, Databases that capture the frequency of different variations in the population
27	What is ethics and governance? Who owns the data?

28-30	Difference between privacy and confidentiality, Data sharing and the law (Class Test/Assignment III)
15 Hours	Tutorials
30 Hours	Practical
<ul style="list-style-type: none"> <li>• <u>Suggested References:</u></li> <li>• Genomes 2 by Brown TA. BIOS Scientific Publishers Ltd, New York.</li> <li>• Molecular Biology of the Cell, Bruce A, Alexander J. Garland Science</li> <li>• Essential Cell Biology, Bruce A, Dennis B, Karen H, Alexander JD, Julian. Garland Science</li> <li>• Bioinformatics and Computational Biology Solutions Using R and Bioconductor by Gentleman R. Springer.</li> <li>• ClinVar: public archive of relationships among sequence variation and human phenotype. Landrum MJ, et al, Nucleic Acids Res. 2014 Jan 1;42(1):D980-5. PubMed PMID: 24234437.</li> <li>• Next-generation Sequencing in the Diagnosis of Metabolic Disease Marked by Pediatric Cataract. R. L. Gillespie et al Ophthalmology. 2016 Jan;123(1):217-20.</li> <li>• Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology Genetics in Medicine 17;5 405-424 doi:10.1038/gim.2015.30.</li> <li>• Next Generation Microarray Bioinformatics: Methods &amp; Protocols, by Wang J. Humana Press</li> <li>• Next Generation Sequencing Data Analysis, by Wang X. CRC Press</li> </ul>	

## RESEARCH & PUBLICATION ETHICS

<b>Course Code</b>	BIS 9 1 ME 017 00	<b>Credits</b>	2
<b>L + T + P</b>	2 + 0 + 0	<b>Course Duration</b>	One Semester
<b>Semester</b>	I/II	<b>Contact Hours</b>	30 Hours
<b>Course Type</b>	Core		
<b>Nature of the Course</b>	Theory		
<b>Methods of Content Interaction</b>	<i>Lecture, Tutorials, primary data collection &amp; analysis, Hands-on Exercise</i>		
<b>Assessment and Evaluation</b>	Student would be given assignment/short-project and based on performance, satisfactory/unsatisfactory grade would be provided.		

❖ **Course Objectives:**

- ❖ This course is designed to provide an overview on ethics, misconduct in scientific research particularly focusing on scientific research and publications. Course also aim to introduce various aspects of scientometrics

**Course Learning Outcomes:**

After completion of the course the students will be able to:

- Develop skills in scientometrics
- Develop comprehensive understanding on ethical principles and scientific misconduct in scientific research and communication

❖ **Course Contents:**

**UNIT I: Research and Publication Ethics: THEORY**

**(50 % Weightage)**

- **RPE01: Philosophy and Ethics:**
  - Introduction to Philosophy: definition, nature and scope, concept, branches
  - Ethics: Definition, moral philosophy, nature of moral judgments and reactions.
- **RPE02: Scientific Conduct**
  - Ethics with respect to science and research
  - Intellectual honesty and research integrity
  - Scientific misconducts: Falsification, Fabrication and Plagiarism (FFP)
  - Redundant publications: duplicate and overlapping publications, salami slicing
  - Selective reporting and misrepresentation of data
- **RPE03: Publication Ethics**
  - Publication ethics: definition, introduction and importance
  - Best practices/standards setting initiatives and guidelines: COPE, WAME etc.
  - Conflicts of interest
  - Publication misconduct: Definition, concept, problems that lead to unethical behavior and vice versa, types

- Violation of publication ethics, authorship and contributorship
- Identification of publication misconduct, complaints and appeals
- Predatory publishers and journals

## UNIT II: Research and Publication Ethics: PRACTICE

(50 % Weightage)

- **RPE04: Open Access Publishing**
  - Open access publications and initiatives
  - SHERPA/RoMEO online resource to check publisher copyright & self-archiving policies
  - Software tool to identify predatory publications developed by SPPU: UGC-CARE list of journals
  - Journal finder/journal suggestion tools viz. JANE, Elsevier Journal Finder, Springer Journal Suggester, etc.
- **RPE05: Publication Misconduct**
  - A. Group discussions**
    - Subject specific ethical issues, FFP, authorship
    - Conflicts of interest
    - Complaints and appeals: examples and fraud from India and abroad
  - B. Software tools**
    - Use of plagiarism software like Turnitin, Urkund and other software tools (reference management software like EndNote, Mendeley, Zotero etc.)
- **RPE 06: Databases and research metrics**
  - A. Databases**
    - Indexing databases
    - Citation databases: Web of Science, Scopus etc.
  - B. Research Metrics**
    - Impact factor of journal as per Journal Citation Report, SNIP, SJR, IPP, Cite Score
    - Metrics: h-index, g-index, i-10 index, altmetrics

### Suggested Readings:

- MacIntyre, Alasdair (1967) A Short History of Ethics. London
- P.Chaddah, (2018) Ethics in Competitive Research: Do not get Scooped; do not get Plagiarized, ISBN :978-9387480865
- National Academy of Sciences, National Academy of Engineering and Institute of Medicine. (2009). On Being a Scientist: A Guide to responsible conduct in Research: Third Edition, National Academies Press.
- Resnik, D.B. (2011) What is ethics in research & why is it important. National institute of Environmental Health Science, 1-10 Retrieved from <https://www.niehs.nih.gov/research/resources/bioethics/whatis/index.cfm>
- Beall, J: (2012) Predatory publishers are corrupting open access. Nature, 489(7415), 179-179. <https://doi.org/10.1038/489179a>
- Indian National Science Academy (INSA), Ethics in Science Education, Research and Governance (2019), ISBN:978-81-939482-1-7. [http://www.insaindia.res.in/pdf/Ethics\\_Book.pdf](http://www.insaindia.res.in/pdf/Ethics_Book.pdf)