Syllabus of

MSc Computational Biology
Specialisation in NGS Data Analytics

(effective from 2021 Admission)

Dept. of Computational Biology & Bioinformatics
UNIVERSITY OF KERALA
2021

GRADUATE ATTRIBUTES (GAs)

The Graduate Attributes (GAs) reflect particular qualities and abilities of an individual learner including knowledge, application of knowledge, professional and life skills, attitudes and human values that are required to be acquired by the graduates of University of Kerala. The graduate attributes include capabilities to strengthen one's professional abilities for widening current knowledge and industry-ready skills, undertaking future studies for global and local application, performing creatively and professionally, in a chosen career and ultimately playing a constructive role as a socially responsible global citizen. The Graduate Attributes define the characteristics of learners and describe a set of competencies that are beyond the study of a particular area and programme.

The GAs of University of Kerala

- Continue life-long learning as an autonomous learner
- Continuously strive for excellence in education
- Apply and nurture critical and creative thinking
- Promote sustainable development practices
- Promote co-operation over competition
- Balance rights with responsibilities
- Understand and respect diversity & difference
- Not be prejudiced by gender, age, caste, religion, or nationality.
- Use education as a tool for emancipation and empowerment of humanity

Note:

1. The following abbreviations are used in this document:

   CB & B : Computational Biology and Bioinformatics
   PSO : Programme Specific Outcome
   CO : Course Outcome
   TL : Taxonomic Level (with reference to Revised Bloom's Taxonomy of educational objectives for cognitive domain)
   Re : Remember
   Un : Understand
   Ap : Apply
   An : Analyse
   Ev : Evaluate
   Cr : Create

2. The syllabus states Module Outcomes (MO) in addition to PSO and CO. Action verbs have been used in MO, rather than CO.
3. The mapping of COs to PSOs is indicated for each module in the module syllabus, rather than in a separate a mapping table.
4. Sample questions to test outcomes have been included after each module.
5. Taxonomy Levels (TL) have been stated for each course outcome in non-laboratory courses. For laboratory courses or ones which involve affective/psychomotor domains, TLs have not been specified as they are not well established and straight forward.

Course structure of MSc Computational Biology Programme
(Specialization in NGS Data Analytics)

MSc. Syllabus of the Programme

Semester	Course Code	Name of the course	Number of Credits
I	**Core Courses (CC)**		
	BIN-CC-511	Introduction to Life Sciences & Bioinformatics	4
	BIN-CC-512	Applied Mathematics	4
	BIN-CC-513	Python Programming	4
	BIN-CC-514	Bioinformatics Lab I	3
	Discipline Specific Electives (DE)		
	BIN-DE-515	Soft Skills (Additional Skill Acquisition Course) (E)	2
	BIN-DE-516	Web programming and Databases (E)	2
II	**Core Courses (CC)**		
	BIN-CC-521	NGS Data Analytics I	4
	BIN-CC-522	Fundamentals of Molecular Biology	4
	BIN-CC-523	Computational Genomics	4
	BIN-CC-524	Bioinformatics Lab II	3
	Discipline Specific Electives (DE)		
	BIN-DE-525	Perl and Bio Perl (E)	2
	BIN-DE-526	Negotiated Studies(E)	2
	BIN-DE-527	Soft Skills (Additional Skill Acquisition Course) (E)	2
III	**Core Courses (CC)**		
	BIN-CC-531	NGS Data Analytics II	4
	BIN-CC-532	Computational Proteomics	4
	BIN-CC-533	Advanced Topics in Computational Biology	4
	BIN-CC-534	Bioinformatics Lab III	3
	Discipline Specific Electives (DE)		
	BIN-DE-535	Programming in R (E)	2
	BIN-DE-536	Seminar (E)	2
	BIN-DE-537	Soft Skills (Additional Skill Acquisition Course) (E)	2
IV	**Core Courses (CC)**		
	BIN-CC-541	NGS Data Analytics III	3
	BIN-CC-542	Project and Viva Voce	12
Generic Courses			
I, II, III	BIN-GC-501	Introductory Bioinformatics	2

Programme Specific Outcomes (PSO) for MSc Computational Biology (Specialization in NGS Data Analytics)

PSO 1: Life Science concepts (especially molecular biology) and skills relevant to CB & B
PSO 2: Mathematical concepts and skills relevant to CB& B
PSO 3: Concepts and skills in processing bio-sequence data
PSO 4: Skill in designing and developing on-line databases
PSO 5: Basic skill in designing and developing simple websites
PSO 6: Basic wet lab skill and exposure to molecular biology experiments
PSO 7: Advanced skill in selected industry-required programming language
PSO 8: Thorough knowledge of basic concepts underpinning Bioinformatics & NGS Data
PSO 9: Moderate skill in using basic computation tools of CB& B
PSO 10: Knowledge in the field of Machine/Deep Learning
PSO 11: Skill in developing simple Deep Learning models
PSO 12: Knowledge about scientific method and skill in research process
PSO 13: Knowledge and skill to process DNA/RNA sequence data
PSO 14: Knowledge and skill in NGS Data Analytics
PSO 15: Basic *in- silico* laboratory skills relevant to CB & B
PSO 16: Skill in designing & implementing medium-scale projects in NGS Data Analytics
PSO17: Awareness of emerging trends and concepts in CB & B
PSO 18: Knowledge and skill to process protein sequence data
PSO19: Concepts and skills in modelling protein structure
PSO20: Knowledge in concept and approaches used in drug discovery
PSO21: Moderate skill in using basic tools of *in-silico* drug discovery
PSO22: Enhanced skills and attitudes for becoming a better learner, thinker, professional and a human being

Semester I

Course Title	Course Code
Introduction to Life Sciences & Bioinformatics	BIN-CC-511
Applied Mathematics	BIN-CC-512
Python Programming	BIN-CC-513
Bioinformatics Lab I	BIN-CC-514
Soft Skills (Additional Skill Acquisition Course) (E)	BIN-DE-515
Web programming and Databases (E)	BIN-DE-516

Semester I	Course Code: BIN-CC-511	Credits:4

INTRODUCTION TO LIFE SCIENCE & BIOINFORMATICS

Programme Specific Outcome relevant to this course:

PSO1: Life Science concepts (especially molecular biology) and skills relevant to CB & BI

Course Outcome

On completion of the course, students should be able to:

CO1: Articulate and exemplify basic knowledge of Life science
CO2: Articulate and exemplify knowledge about macro biomolecules
CO3*:* Articulate and critically evaluate various systematics alternatives
CO4*:* Articulate basic knowledge about cells and subcellular processes
CO5: Articulate basic concepts of DNA, its functions and associated mechanisms
CO6: Articulate basic knowledge about Amino acids, properties and its classification
CO7: Articulate basic structure of proteins and its functions
CO8: Articulate nature and scope of Bioinformatics
CO9: Do basic bio sequence handling
CO10: Demonstrate basic knowledge about bio sequence databases

Semester I	Course Code: BIN-CC-511	Credits:4

INTRODUCTION TO LIFE SCIENCE & BIOINFORMATICS

COURSE CONTENT:

Module I: Characteristics of life, Levels of biological Organization- From atoms to molecules to life, general properties and biological functions of Macromolecules; carbohydrates, Lipids, Proteins and Nucleic acids, intermolecular interactions: covalent bonds, ionic bonds, Non covalent bonds: Hydrogen bonds, electrostatic interactions, Van der Waals interactions, systematics: Binomial nomenclature, Five Kingdom, Six kingdom and Three domain system of classification.

Module II: The cell as basic unit of life, cell theory, structure of Prokaryotic cell and Eukaryotic cell, structure and function of cell membrane, cell organelles- nucleus, endoplasmic reticulum, ribosomes, Golgi complex, mitochondria, peroxisomes, lysosomes, cell division- mitosis, meiosis, cell death: Brief overview about apoptosis and cell necrosis.

Module III: Primary and secondary structure of DNA, Chargaff's Rules, Different forms of DNA, RNA, structural organization of DNA, Gene, chromatin threads, higher order structure of chromosomes.

Module IV: Amino acids, Functional classification of proteins, Representation of amino acids in bio sequence, Structural organization of proteins: Primary, Secondary, Tertiary and Quaternary structure of proteins; Protein folding.

Module V: Bioinformatics: History of Bioinformatics, Definition of Bioinformatics, Bioinformatics versus Computational Biology, Goals of Bioinformatics analysis, Bioinformatics technical tool box, Biological data, File format, conversion of file format, Data retrieval system, Genome browsers.

Module VI: Databases: Bioinformatics databases, Types of databases, Nucleotide sequence databases, Primary nucleotide sequence databases-EMBL, Gene Bank, DDBJ; Secondary nucleotide databases, Protein sequence databases-SwissProt/ TrEMBL, Protein structure databases- Protein Data Bank, SCOP and CATH.

BIN-CC-511 INTRODUCTION TO LIFE SCIENCE & BIOINFORMATICS: Module I

Course Outcomes relevant to this Module:

CO1: Articulate and exemplify basic knowledge of Life science
CO2: Articulate and exemplify knowledge about macro biomolecules
CO3: Articulate and critically evaluate various systematics alternative

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Articulate basic concepts and terminologies of life science	Un
MO2 Articulate characteristic features of life	Un
MO3 Articulate the organization of life, from atoms to Biosphere	Un
MO4 Articulate the Biological functions of macromolecules: Carbohydrates, Proteins, Lipids and Nucleic acids	Un
MO5 Articulate various intermolecular and electrostatic interactions and its relevance in stability and formation of macromolecule	Un
MO6 Explain and critically compare the alternate system of systematics	Ev
MO7 Articulate the binomial nomenclature of organisms	Un
MO8 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: Characteristics of life, Levels of biological Organization- From atoms to molecules to life, general properties and biological functions of Macromolecules; carbohydrates, Lipids, Proteins and Nucleic acids, intermolecular interactions: covalent bonds, ionic bonds, Non covalent bonds: Hydrogen bonds, electrostatic interactions, Van der Waals interactions, systematics: Binomial nomenclature, Five Kingdom, Six kingdom and Three domain system of classification.

BIN-CC-511 INTRODUCTION TO LIFE SCIENCE & BIOINFORMATICS: Module II

Course Outcomes relevant to this Module:

CO4: Articulate basic knowledge about cells and subcellular processes

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Articulate the basic architecture of cells and cell theory	Un
MO2 Differentiate between prokaryotes and eukaryotes	Un
MO3 Articulate structure and function of various sub cellular organelles	Un
MO4 Describe various phases of longitudinal cell division: Mitosis	Un
MO5 Describe various phases of reductive cell division: Meiosis	Un
MO6 Compare and contrast Apoptosis and Necrosis	Ev
MO7 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: The cell as basic unit of life, cell theory, structure of Prokaryotic cell and Eukaryotic cell, structure and function of cell membrane, cell organelles- nucleus, endoplasmic reticulum, ribosomes, Golgi complex, mitochondria, peroxisomes, lysosomes, cell division- mitosis, meiosis, cell death: Brief overview about apoptosis and cell necrosis.

BIN-CC-511 INTRODUCTION TO LIFE SCIENCE & BIOINFORMATICS: Module III

Course Outcomes relevant to this Module:

CO5: Articulate basic concepts of DNA, its functions and associated mechanisms

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Articulate experimental evidence of DNA as genetic material	Un
MO2 Articulate the Primary structure of DNA	Un
MO3 Articulate the secondary structure of DNA	Un
MO4 Articulate the concept of complementary base pairing of DNA	Un
MO5 Articulate polymorphism of DNA	Un
MO6 Articulate and exemplify different types of RNA and its functions	Un
MO7 Articulate the higher order structure, from DNA to chromosomes	Un
MO8 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: Primary and secondary structure of DNA, Chargaff's Rules, Different forms of DNA, RNA, structural organization of DNA, Gene, chromatin threads, higher order structure of chromosomes.

BIN-CC-511 INTRODUCTION TO LIFE SCIENCE & BIOINFORMATICS: Module IV

Course Outcomes relevant to this Module:

CO6: Articulate basic knowledge of Amino acids, classification and its properties
CO7: Articulate basic structure of proteins and its functions

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Articulate basic properties of amino acids	Un
MO2 Articulate different amino acids, its three letter and one letter code	Un
MO3 Classify proteins based on its function	Un
MO4 Articulate the basic architecture of proteins	Un
MO5 Explain the process of protein folding	Un
MO6 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: Amino acids, Functional classification of proteins, Representation of amino acids in bio sequence, Structural organization of proteins: Primary, Secondary, Tertiary and Quaternary structure of proteins; Protein folding.

BIN-CC-511 INTRODUCTION TO LIFE SCIENCE & BIOINFORMATICS: Module V

Course Outcomes relevant to this Module:

CO6: Articulate nature and scope of Bioinformatics
CO7: Do basic bio sequence handling

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Critically discuss definitions of Bioinformatics	Un
MO2 List the different file formats used in Biological databases	Un
MO3 Retrieve genome sequences from genome browsers	Ap
MO4 Use Bioinformatics technical tool box	Ap
MO5 Identify standard bio sequence file formats	Un
MO6 Contrast Computational Biology and Bioinformatics	Un
MO7 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: Bioinformatics: History of Bioinformatics, Definition of Bioinformatics, Bioinformatics versus Computational Biology, Goals of Bioinformatics analysis, Bioinformatics technical tool box, Biological data, File format, conversion of file format, Data retrieval system, Genome browsers.

BIN-CC-511 INTRODUCTION TO LIFE SCIENCE & BIOINFORMATICS: Module VI

Course Outcomes relevant to this Module:

CO8: Articulate basic knowledge about bio sequence databases

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 List different types of databases and tools in Bioinformatics	Un
MO2 Read and interpret standard data files	Ap
MO3 Use specific databases and tools for basic sequence analysis	Ap
MO4 Contrast primary and secondary databases	Un
MO5 Retrieve the protein structure and analyse its features and properties using SCOP and CATH	Ap
MO6 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: Databases: Bioinformatics databases, Types of databases, Nucleotide sequence databases, Primary nucleotide sequence databases-EMBL, Gene Bank, DDBJ; Secondary nucleotide databases, Protein sequence databases-SwissProt/ TrEMBL, Protein structure databases- Protein Data Bank, SCOP and CATH.

BIN-CC-511 INTRODUCTION TO LIFE SCIENCES AND BIOINFORMATICS: ACTIVITIES, LEARNING RESOURCES & ASSESSMENT

SUGGESTED CLASS ROOM ACTIVITIES:

- Seminar Presentation on selected topics
- Debates
- Quiz
- Demonstration of simple experiments

LEARNING RESOURCES:

References:

1. Devasena T. (2012), Cell Biology, Oxford University Press.
2. Karp G. (2014, 2016), Cell and Molecular Biology: Concepts and Experiments, John Wiley & Sons.
3. Kuriyan J., Konforti B., & Wemmer D. (2012), The molecules of life: Physical and chemical principles, Garland Science.
4. Hausman R.E., & Cooper G.M. (2019), The Cell: a molecular approach. ASM, Washington, DC.
5. Lewin B. (2011), Lewin's Genes X (Vol. 10, 11 &12), J. Krebs, S.T. Kilpatrick, & E. S. Goldstein (Eds.). Jones & Bartlett Learning.
6. Lodish H. (2008), Molecular cell biology, Macmillan.
7. Mason K. A., Losos J. B., Singer S. R., & Peter H. Raven, and George B. J. (2011), Biology.
8. Nelson D. L., Lehninger A. L., & Cox M. M. (2008), Lehninger principles of Biochemistry, Macmillan.
9. Rastogi S. C. (2005), Cell biology, New Age International.
10. Snustad D. P., Simmons M. J., Jenkins J. B., & Crow J. F. (2000), Principles of Genetics, John Wiley.
11. Tropp B. E. (2012), Molecular Biology: genes to proteins, Jones & Bartlett Publishers.
12. Watson J. D., & Berry A. (2009), DNA: The secret of life, Random House LLC.
13. Watson J. D., Gilman M., Witkowski J., & Zoller, M. Recombinant DNA.
14. Mount D. W. (2004), Sequence and genome analysis., Bioinformatics: Cold Spring Harbour Laboratory Press: Cold Spring Harbour.
15. Rastogi S. C., Mendiratta N., & Rastogi P. (2013), Bioinformatics: Methods and Applications: (Genomics, Proteomics and Drug Discovery), PHI Learning Pvt. Ltd.
16. Choudhuri S., & Kotewicz M. (2016), Bioinformatics for beginners: Genes, Genomes, Molecular evolution, Databases and Analytical tools, Elsevier.
17. Ghosh Z., & Mallick B. (2014), Bioinformatics: Principles and applications, Oxford University Press.

On-line Resources/MOOCs:

1. Introduction to Biology - The secret of life, Massachusetts Institute of Technology (Edx)
2. Preparation for Introductory Biology: DNA to organisms, University of California (Coursera)
3. Bioinformatics: Introduction and Method, Peking University (Coursera)

ASSESSMENT:

40% Continuous / Formative Assessment (see PG Regulations, Appendix).
60% End-semester/Summative Assessment: 3 hour written Exam (see PG Regulations, Appendix).

Semester I	Course Code: BIN-CC-512	Credits:4

APPLIED MATHEMATICS

Programme Specific Outcome relevant to this course:

PSO2: Mathematical concepts and skills relevant to CB&B

Course Outcome

On completion of the course, students should be able to:

CO 1: Demonstrate the use of Complex numbers in numerical encoding of DNA
CO 2: Demonstrate the use of Scalars and Vectors in database and computation
CO 3: Demonstrate the use matrix algebra to determine the distribution of genotypes
CO 4: Demonstrate the use simultaneous equation in evolutionary tree
CO 5: Demonstrate the use of python for Computational thinking
CO 6: Apply functions in the microbial growth models
CO 7: Apply curve fitting in epidemic models
CO 8: Demonstrate the use of geometric concepts
CO 9: Demonstrate the use trigonometric function/identities
CO10: Apply Fourier theory to derive insight on DNA sequence
CO11: Apply differentiation and integration for biological modelling
CO12: Demonstrate use of python in computing derivatives and integrals
CO13: Formulate problems in the language of sets and perform set operations
CO14: Apply probability theory in prediction problems
CO15: Demonstrate the use of statistical approaches in Bioinformatics
CO16: Demonstrate the modelling of biological data using regression
CO17: Demonstrate the use of statistical tests in data handling
CO18: Demonstrate the use of graph theory as modelling tool

Semester I	Course Code: BIN-CC-512	Credits:4

APPLIED MATHEMATICS

COURSE CONTENT:

Module I: Number systems, Real numbers, Rational numbers and Complex numbers (Use in Numerical encoding of DNA Sequence), Solving equations- first-order equations, Quadratic equations, Simultaneous linear equations(Use in Evolutionary Tree), Linear Algebra: Scalars &Vectors, addition, subtraction, dot, cross & scalar triple products, Matrices, inverse of a matrix, Operations, solution of simultaneous equation by using matrix, Eigen value, Eigen vector (Use in Biological Database/ Population Genetics). Linear Algebra with Python: cmath-complex number processing, Numpy-Matrix, Basic Operations Eigen Value/Vector, Solving Linear Equations with Python

Module II: Functions, Inverse Functions, Exponentials and logs to represent Natural growth and decay, Bacteria/population growth model, Circles and angles, Straights lines, angles, Area and Volume, Pythagoras' Theorem, Basic trigonometric functions: Sine and Cosine, Sinusoidal oscillations, Amplitude, Frequency and phase of sinusoidal oscillation, Damped oscillations, Waves, Fourier theorem (basic Introduction only) Trigonometric functions in python , Fourier theory application in DNA sequence analysis, Mathematical Modelling and Fourier application with Python

Module III: Calculus: Limits, Derivatives & Integrals: Limits, continuity, derivative as rate of change. Graphical treatment of derivative, maxima and minima, Newton's methods, numerical differentiation. Integrals: graphical treatment, Integrating simple expressions, Definite and indefinite integrals, The area under a curve, Numerical integration-Simpson's rule, simple symbolic integration. Introduction to Differential equations, Solving differential equation with Python.

Module IV: Set theory, Permutation, combination & Probability: Set theory, sets, elements, set operation, finite & countable sets, counting, factorial, permutation, combination, binomial coefficients. Basic concepts; sample space & events, laws of probability, conditional probability: Baye's theorem, Random variables: probability distribution, Binomial, Poisson, normal and 't'.

Module V: Statistics: Scope of statistical methods, Categorical & Numerical data, frequency distribution, Data Distribution-Uniform, Normal; Measures of central tendency: mean, median, mode, geometric mean, harmonic mean, percentile; Measures of dispersion: range, mean deviation, variance, standard deviation, Z-value, confidence interval, p-value; Types of sampling methods.

Module VI: Applied Statistics: Covariance, Correlation; Linear regression; Statistical testing: Population, Sample, Central limit theorem, Null/alternative hypothesis, F-test, t-test, Chi- square test; ANOVA. Graph terminology: edges, vertices, loop, path, circuit, bridge, Eulers path; Graph representation: Adjacency matrix, incident matrix. Graphs: null, complete, regular, bipartite and complete bipartite, Introduction to SAS, Case study on Clinical trial analysis using python packages.

BIN-CC-512 APPLIED MATHEMATICS: Module I

Course Outcomes relevant to this Module:

CO 1: Demonstrate the use of Complex numbers in numerical encoding of DNA
CO 2: Demonstrate the use of Scalars and Vectors in database and computation
CO 3: Demonstrate the use matrix algebra to determine the distribution of genotypes
CO 4: Demonstrate the use simultaneous equation in evolutionary tree
CO 5: Demonstrate the use of Python for Computational thinking

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Articulate different number systems	Un
MO2 Add, subtract and multiply complex numbers	Ap
MO3 Apply complex number mapping in DNA sequence	Ap
MO4 Apply linear algebra in organizing data and do arithmetic operations	An
MO5 Solve Quadratic and simultaneous equations	Ap
MO6 Add, subtract, multiply (dot, cross & scalar triple) and inverse matrices	Ap
MO7 Compute the genotype of population	Ap
MO8 Compute eigen vector and eigen values of matrices	Ap
MO9 List numpy operators in Python	Un
MO10 Find eigen value for a matrix using Python	Ap
MO11 Write user defined functions to perform translation/modification of given image in Python	Cr
MO12 Perform DNA/Protein Sequence Data handling using Python	Cr
MO13 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **Introductory concepts** : Number systems, Real numbers, Rational numbers and Complex numbers (Use in Numerical encoding of DNA Sequence), Solving equations- first-order equations, Quadratic equations, Simultaneous linear equations(Use in Evolutionary Tree), Linear Algebra: Scalars &Vectors, addition, subtraction, dot, cross & scalar triple products, Matrices, inverse of a matrix, Operations, solution of simultaneous equation by using matrix, Eigen value, Eigen vector (Use in Biological Database/ Population Genetics). Linear Algebra with Python: cmath-complex number processing, Numpy-Matrix, Basic Operations Eigen Value/Vector, Solving Linear Equations with Python

BIN-CC-512 APPLIED MATHEMATICS: Module II

Course Outcomes relevant to this Module:

CO 6: Apply functions in the microbial growth models
CO 7: Apply curve fitting in epidemic models
CO 8: Demonstrate the use of geometric concepts
CO 9: Demonstrate the use trigonometric function/identities
CO10: Apply Fourier theory to derive insight on DNA sequence

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Differentiate independent and dependent variable	Un
MO2 Evaluate domain and range of a given function	Ap
MO3 Contrast one to one and one to many functions	An
MO4 Perform Horizontal line test	Ap
MO5 Do arithmetic operations using logarithmic/exponential functions	Ap
MO6 Plot graph of a given function	Ap
MO7 Distinguish graph functions using horizontal and vertical shift	Ap
MO8 Evaluate logarithmic & exponential function	Ap
MO9 Plot an exponentially growing/decaying function	Ap
MO10 Compute slope of a line	Ap
MO11 Compute angle between two lines	Ap
MO12 Find the area/volume of polygon/circle	Ap
MO13 Apply linear function to model constant rate of change	Ap
MO14 Define different trigonometric functions	Re
MO15 Articulate Pythagoras theorem/Identities	Re
MO16 Describe sinusoidal oscillations	Un
MO17 Solve Trigonometric problems using identities.	Ap
MO18 Articulate Fourier Theorem	Un
MO19 Execute Python programs for simple trigonometric functions	Ap
MO20 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: Functions, Inverse Functions, Exponentials and logs to represent Natural growth and decay, Bacteria/population growth model, Circles and angles, Straights lines, angles, Area and Volume, Pythagoras' Theorem, Basic trigonometric functions: Sine and Cosine, Sinusoidal oscillations, Amplitude, Frequency and phase of sinusoidal oscillation, Damped oscillations, Waves, Fourier theorem (basic Introduction only) Trigonometric functions in python , Fourier theory application in DNA sequence analysis, Mathematical Modelling and Fourier application with Python

BIN-CC-512 APPLIED MATHEMATICS: Module III

Course Outcomes relevant to this Module:

CO11: Apply differentiation and integration for biological modelling

CO12: Demonstrate use of Python in computing derivatives and integrals

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Articulate and exemplify the concept of limit of a function	Un
MO2 Compute the limit of a function	Ap
MO3 Articulate the ideas of continuity and differentiability	Un
MO4 Articulate and exemplify derivative as a model of change	Un
MO5 State the rules of differentiation	Re
MO6 Evaluate the derivatives of simple functions	Ap
MO7 Compute derivative through numeric differentiation (Newton's)	Ap
MO8 Find the maxima or minima of a function by differentiation	Ap
MO9 Compare graph of a function and derivative	Ap
MO10 Articulate integral as a physical process	Un
MO11 Evaluate the critical point of a function as interval	Ap
MO12 Do Simple integration problem	Ap
MO13 Find the area under a curve through integration	Ap
MO14 Compute integral through Simpson's rule	Ap
MO15 Solve simple differential equation	Ap
MO16 Use differentiation and integration in Python	Ap
MO17 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: Calculus: Limits, Derivatives & Integrals: Limits, continuity, derivative as rate of change. Graphical treatment of derivative, maxima and minima, Newton's methods, numerical differentiation. Integrals: graphical treatment, Integrating simple expressions, Definite and indefinite integrals, The area under a curve, Numerical integration-Simpson's rule, simple symbolic integration. Introduction to Differential equations, Solving differential equation with Python.

BIN-CC-512 APPLIED MATHEMATICS: Module IV

Course Outcomes relevant to this Module:

CO13: Formulate problems in the language of sets and perform set operations
CO14: Apply probability theory in prediction problems

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Articulate and exemplify the concept of set and set operations	Un
MO2 Perform set operation: Union, Intersection, Complement, Difference	Ap
MO3 Articulate and exemplify Set identities /Properties	An
MO4 Use Venn diagram to visualize sets and their relationships	Ap
MO5 Apply set theory to solve problem	Ap
MO6 Compute permutations and combinations	Ap
MO7 Solve problems involving permutation and combination	Ap
MO8 Articulate and exemplify basic concepts of probability	Un
MO9 Evaluate probability of an event viz: tossing a con/rolling a die/drawing a card	Ap
MO10 Apply principle of probability to solve problems	Ap
MO11 Compute posterior probabilities using Bayes' theorem	Ap
MO12 Articulate and interpret standard statistical distributions	Un
MO13 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: Set theory, Permutation, combination & Probability: Set theory, sets, elements, set operation, finite & countable sets, counting, factorial, permutation, combination, binomial coefficients. Basic concepts; sample space & events, laws of probability, conditional probability: Baye's theorem, Random variables: probability distribution, Binomial, Poisson, normal and 't'.

BIN-CC-512 APPLIED MATHEMATICS: Module V

Course Outcomes relevant to this Module:

CO15: Demonstrate the use of statistical approaches in Bioinformatics

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Identify/exemplify Categorical/ Numerical/ Discrete/ Continuous data	Un
MO2 Organize data in frequency/grouped frequency distribution table	Ap
MO3 Articulate frequency distribution graph	Un
MO4 Compute AM, GM, HM, SD and variance of given dataset	Ap
MO5 Articulate and Exemplify different measures of central tendency/or dispersion	Un
MO6 Compare the standard deviation of two different dataset	Ap
MO7 Calculate z-score of a given dataset	Ap
MO8 Articulate different sampling methods and suggest one for a given situation	Ev
MO9 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: Statistics: Scope of statistical methods, Categorical & Numerical data, frequency distribution, Data Distribution-Uniform, Normal; Measures of central tendency: mean, median, mode, geometric mean, harmonic mean, percentile; Measures of dispersion: range, mean deviation, variance, standard deviation, Z-value, confidence interval, p-value; Types of sampling methods.

BIN-CC-512 APPLIED MATHEMATICS: Module VI

Course Outcomes relevant to this Module:

CO16: Demonstrate the modelling of biological data using regression
CO17: Demonstrate the use of statistical tests in data handling
CO18: Demonstrate the use of graph theory as modelling tool

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Derive covariance and correlation in the given dataset	Ap
MO2 Articulate the concept of regression	Un
MO3 Solve simple linear regression problems	Ap
MO4 Articulate the different statistical testing methods	Un
MO5 State and Explain central limit theorem	Re
MO6 Articulate and exemplify null/alternative hypothesis	Un
MO7 Solve application involving F/t/Chi Square test/ANOVA	Ap
MO8 Articulate and exemplify the fundamental concepts of graph theory	Un
MO9 Identify edge/vertices/loop/path/circuit/bridge etc for a given graph	Ap
MO10 Do statistical computation on SAS platform	Ap
MO11 Do small scale clinical trial analysis using Python/SAS	Ap
MO12 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: Applied Statistics: Covariance, Correlation; Linear regression; Statistical testing: Population, Sample, Central limit theorem, Null/alternative hypothesis, F-test, t-test, Chi- square test; ANOVA. Graph terminology: edges, vertices, loop, path, circuit, bridge, Eulers path; Graph representation: Adjacency matrix, incident matrix. Graphs: null, complete, regular, bipartite and complete bipartite, Introduction to SAS, Case study on Clinical trial analysis using python packages.

BIN-CC-512 APPLIED MATHEMATICS
ACTIVITIES, LEARNING RESOURCES & ASSESMENT

SUGGESTED CLASS ROOM ACTIVITIES:

- Worksheets /Quizzes
- Maths fun with functions (https://www.wolframalpha.com/examples/mathematics/plotting-and-graphics/)
- White board test for Visualisation of functions
- Learning by doing handout for Python programming
- Activities with rolling a die/tossing coin to understand probability

LEARNING RESOURCES:

References:

1. Olive J. (2000), Maths: A Self-study Guide, Cambridge University Press.
2. M. Harris, G. Taylor & J. Taylor [2007], Maths & Stats for the life and medical sciences, Scion Publishing Ltd.
3. Fred S. [1998], Schaum's outline Theory and Problems of Pre-calculus, Tata McGraw Hill.
4. Lipschutz S. & Schiller J. J. (1998), Schaum's Outline of Theory and Problems of Introduction to probability and statistics, Tata McGraw-Hill.
5. Ramachandran H. Nair A.S. (2012), (Scilab : Free Software to Matlab), S. Chand Limited.
6. Stewart J. (2008), Calculus Early Transcendentals, Bob Pirtle, Belmont, USA.
7. Dawn Griffiths (2008), Head First Statistics, O'Reilly Media Inc.
8. John Stillwell [2018], Elements of Mathematics from Euclid to Godel Princeton, University Press.
9. Mendenhall R.J. Beaver & B. M. Beaver [2017], Introduction to Probability and Statistics, Cengage Learning Products.
10. P. Abbot & H. Neill (2003), Teach Yourself Trigonometry, McGraw Hill.
11. Issac R. (1995), The Pleasures of Probability, Springer, Verlag.
12. Strogatz S. H. (2012), The Joy of X: A Guided Tour of Math, from one to Infinity.

Online Resources:

1. https://www.edx.org/course/ap-introduction-to-statistics
2. https://www.edx.org/course/pre-university-calculus-2
3. https://www.edx.org/course/mitx/mitx-6-041x-introductionprobability-1296.
4. https://www.edx. org/course/tsinghuax/tsinghuax-60240013x-combinatorial-3771
5. https://www.edx.org/course/utaustinx/utaustinx-ut-7-01x-foundations-data-2641
6. http://www.mathtutor.ac.uk/
7. http://www.Scilab .org/
8. https://www.edx.org/course/effective-thinking-through-mathematics-utaustinx-ut-9-01x-0

ASSESSMENT:

40% Continuous / Formative Assessment (see PG Regulations, Appendix).
60% End-semester/Summative Assessment: 3 hour written Exam (see PG Regulations, Appendix).

Semester I	Course Code: BIN-CC-513	Credits:4

PYTHON PROGRAMMING

Programme Specific Outcome relevant to this course:

PSO7: Advanced skill in an industry-required programming language

Course Outcome

On completion of the course, students should be able to:

CO1: All course outcomes CO2 -CO6
CO2: Choose, use and analyse data and control structures for a given task/code
CO3: Know to code using functions, including recursive functions
CO4: Develop Object Oriented and GUI-based event-driven programs
CO5: Use and analyse regular expressions to extract desired patters from data
CO6: Develop code involving data-base connectivity and CGI interface.

Semester I	Course Code: BIN-CC-513	Credits:4

PYTHON PROGRAMMING

COURSE CONTENT:

Module I: Introduction to Python: A quick tour of Python (in Colab and/or IDLE) based on (i) for loops using simple minimal features (ii) exposure to Python Turtle Graphics (focus is on quick feel of language and basic algorithmic thinking, rather than syntax). History of Python language, overview of its features and uniqueness.

Module II: Data Types, Input/Output and Control Structures: Basic data types of Python, Python Operators, basic input/output; basic control structures (if, if-else, elif, continue, break and pass), for loops, while loops, examples based on selective processing (including summing and counting) of natural number sequences for i in range (N).

Module III: Python Collections: Concept of data structures and different types (sequential/non-sequential, mutable/immutable, static/dynamic, linear/non-linear, ordered/unordered). Python Data Structures: Lists, strings, tuples, sets and dictionaries. Basic handling of collections using loops, exposure to methods associated with each class of collections. Algorithms for sorting and merging collections (with focus on lists).

Module IV: Python Files & Functions: Files: Opening and Closing Files, Access modes, File position, file handling with OS commands, Pickles and Shelves, Reading from url: Functions: Functions as a named unit of code with inputs and outputs, arguments and parameters, positional, keyword and default arguments, Anonymous lambda functions, Concept of recursion, analyzing problems as base case and recursive cases, examples of printing members of a list, summing, checking palindromes, printing countdown, factorial etc.

Module V: Classes and Packages: Concept of Object orientation – need for object orientation in relation to software re-use, classes and objects, inheritance, polymorphism; Python classes: definition, inheriting, constructors, overloading, over-riding, class documentation; Exception handling: Types of errors, error handling, Try-except statement; Modules and Packages: Creating modules and packages, import statement, overview of popular packages: math, stat, random.

Module VI: Miscellaneous topics: Basic image processing in Python: Basic concepts of digital images, image file formats, reading and displaying image files, converting images into 2-D lists and vice versa. GUIs in Python: Concept of Event driven programming, introduction to Tkinter; Regular Expressions in Python; Introduction to database connectivity; Introduction to CGI programming in Python

BIN-CC-513 PYTHON PROGRAMMING: Module I

Course Outcomes relevant to this Module:

CO1: All course outcomes CO2 -CO6

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Use IDLE and Google Colab environments	Ap
MO2 Write toy code using for loops and simple assignment and if-else statements	Ap
MO3 Write code for drawing simple graphic patters in Turtle	Ap
MO4 Articulate history and unique features of Python	Un
MO5 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **Introduction to Python:** A quick tour of Python (in Colab and/or IDLE) based on (i) for loops using simple minimal features (ii) exposure to Python Turtle Graphics (focus is on quick feel of language and basic algorithmic thinking, rather than syntax). History of Python language, overview of its features and uniqueness.

BIN-CC-513 PYTHON PROGRAMMING: Module II

Course Outcomes relevant to this Module:

CO2: Choose, use and analyse data and control structures for a given task/code

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 List/exemplify/identify Python data types	Un
MO2 Convert mathematical expressions to Python code	Ap
MO3 Predict effect of code snippets involving multiple assignment statements	Ap
MO4 Using basic input and output in Python programs	Ap
MO5 Write/analyse code snippets which use basic Python control structures	Ap
MO6 Achieve other outcomes dynamically evolved by the teacher during curricular transaction	

Module Syllabus:

Data Types, Input/Output and Control Structures: Basic data types of Python, Python Operators, basic input/output; basic control structures (if, if-else, elif, continue, break and pass), for loops, while loops, examples based on selective processing (including summing and counting) of natural number sequences for i in range (N).

BIN-CC-513 PYTHON PROGRAMMING: Module III

Course Outcomes relevant to this Module:

CO2: Choose, use and analyse data and control structures for a given task/code

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Articulate the concept of data structures and its categories	Un
MO2 Exemplify/identify Python collections	An
MO3 Write/analyse Python code snippets to handle collections	Ap
MO4 Write/analyse Python code for sorting collections	Ap
M05 Achieve other outcomes dynamically evolved by the teacher during curricular transaction	

Module Syllabus: Python Collections: Concept of data structures and different types (sequential/non-sequential, mutable/immutable, static/dynamic, linear/non-linear, ordered/unordered). Python Data Structures: Lists, strings, tuples, sets and dictionaries. Basic handling of collections using loops, exposure to methods associated with each class of collections. Algorithms for sorting and merging collections (with focus on lists).

BIN-CC-513 PYTHON PROGRAMMING: Module IV

Course Outcomes relevant to this Module:

CO3: Know to code using functions, including recursive functions

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Write/analyse code for processing files	Ap
MO2 Write and analyse functions with positional/keyword/default arguments for given tasks	Ap
MO3 Write and analyse recursive functions for given tasks	Ap
MO4 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: Python Files & Functions: Files: Opening and Closing Files, Access modes, File position, file handling with OS commands, Pickles and Shelves, Reading from url: Functions: Functions as a named unit of code with inputs and outputs, arguments and parameters, positional, keyword and default arguments, Anonymous lambda functions, Concept of recursion, analyzing problems as base case and recursive cases, examples of printing members of a list, summing, checking palindromes, printing countdown, factorial etc.

BIN-CC-513 PYTHON PROGRAMMING: Module V

Course Outcomes relevant to this Module:

CO4: Develop object-oriented GUI-based event-driven programs

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Articulate the concept and related terminology of object orientation	Un
MO2 Write and analyze Python classes involving inheritance, overloading and over-riding	Cr
MO3 Use exception facility to handle run-time errors	Un
MO4 Use/list/exemplify popular Python packages	Ap
M05 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: Classes and Packages: Concept of Object orientation – need for object orientation in relation to software re-use, classes and objects, inheritance, polymorphism; Python classes: definition, inheriting, constructors, overloading, over-riding, class documentation; Exception handling: Types of errors, error handling, Try-except statement; Modules and Packages: Creating modules and packages, import statement, overview of popular packages: math, stat, random.

BIN-CC-513 PYTHON PROGRAMMING: Module VI

Course Outcomes relevant to this Module:

CO5: Use and analyse regular expressions to extract desired patters from data
CO6: Develop code involving data-base connectivity and CGI interface.

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Write/analyze code for basic image processing in Python	Ap
MO2 Write/Articulate/analyze GUI-based code for given tasks	Ap
MO3 Write/analyse regular expressions for extracting patterns from data	Ap
MO4 Write/analyse code for connection to and query data bases	Ap
MO5 Write/analyse CGI code in Python to connect to HTML forms	Ap
MO6 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: Miscellaneous topics: Basic image processing in Python: Basic concepts of digital images, image file formats, reading and displaying image files, converting images into 2-D lists and vice versa. GUIs in Python: Concept of Event driven programming, introduction to Tkinter; Regular Expressions in Python; Introduction to database connectivity; Introduction to CGI programming in Python

BIN-CC-513 PYTHON PROGRAMMING:
ACTIVITIES, LEARNING RESOURCES & ASSESSMENT

SUGGESTED CLASS ROOM ACTIVITIES:

1. Program Portfolio: Students shall submit a portfolio with at least 20 Python programs in IDLE environment.
2. Run the programs in Jupyter Notebook and contrast with IDLE.
3. White Board Test/challenge

LEARNING RESOURCES:

References:

1. Mark S. (2018), Programming in Python 3: A complete Introduction to the Python Language, 2nd Edition, Pearson Education.
2. Lutz M. (2013), Learning Python, O'Reilly Media.
3. Tim J. S., Wayne B. (2015), Python Programming for Biology Bioinformatics and Beyond, Cambridge University Press.
4. Downey A. (2012), Think Python: How to Think Like a Computer Scientist, O'Reilly Media.
5. Punch W. F., Enbody R. (2016), The Practice of Computing Using Python, 3rd Edition, Pearson Education.
6. Barry P. (2010), Head First Python, O'Reilly Media.
7. Beazley D. M. (2009), Python Essential Reference, Pearson Education.
8. Dawson M. (2010), Python Programming for the Absolute Beginner, 3rd Edition, Cengage Learning.

On-line Resources/MOOCs:

1. Python3 Programming, University of Michigan. (Coursera)
2. Python Programming Essentials, RICE University. (Coursera)
3. Programming for Everybody. (Coursera)
4. Python Tutorial: Learn Python
   http://www.codecademy.com/learn/learn-python-3
5. https://realpython.com/python3-object-oriented-programming
6. https://www.geeksforgeeks.org/tag/python-regex
7. https: www.learnpython.org
8. https://ww.w3schools.com/python

ASSESSMENT:

40% Continuous / Formative Assessment (see PG Regulations, Appendix).
60% End-semester/Summative Assessment: 3 hour written Exam (see PG Regulations, Appendix).

Semester I	Course Code: BIN-CC-514	Credits:3

BIOINFORMATICS LAB 1

Programme Specific Outcome relevant to this course:

PSO 6: Basic wet lab skill and exposure to molecular biology experiments
PSO 15: Basic *in- silico* laboratory skills relevant to CB & B

Course Outcome

On completion of the course, students should be able to:

CO 1: Handle various laboratory instruments and practice safety measures
CO 2: Prepare various reagents based on solute concentration
CO 3: Adjust pH in different solutions
CO 4: Identify bacteria by different staining methods
CO 5: Do Blood typing to understand the basic antigen antibody reaction
CO 6: Demonstrate bacterial growth kinetics
CO 7: Demonstrate understanding of different types of WBCs
CO 8: Detect Blood groups by antigen antibody reaction
CO 9: Analyse different stages of Mitotic cell division
CO10: Retrieve sequence/structure information from Biological databases
CO11: Create small databases, and deploy simple pages in local web browsers
CO12: Process bio sequence using Python

Semester I	Course Code: BIN-CC-514	Credits:3

BIOINFORMATICS LAB 1

COURSE CONTENT:

Module I: Laboratory safety guidelines, equipment handling, and sterilization techniques, Preparation of buffer, acid, base and pH, media preparation: simple media, special media.

Module II: Microscopic examination of bacteria by simple and differential staining, Antibiotic sensitivity test, Microbial growth curve.

Module III: Differential staining of WBC, Blood typing, Mitotic cell division- onion root tip.

Module IV: Basic use of standard biological databases: NCBI, PDB, SWISS PROT, etc.

Module V Database creation and management using PHP-MySQL, Creating Simple web pages with forms; Creating interactive pages with Javasript, setting up PHP-MySQL environment, retrieving data from PHP-MySQL units PHP Pages, case studies.

Module VI: Writing programs using Python features: Sequence analysis – Reading DNA/Protein sequences, sequence length, GC%, handle string search; Programs using string handling function-substring search, count of nucleotides, check RNA/DNA, ORF finding, Transcription, Translation; File handling programs-ReadWrite Fasta; Program with regular expression, Define RE for a set of sequences, search for subsequences/patterns, locations; Programs for processing FASTA files.

BIN-CC-514 BIOINFORMATICS LAB 1: Module I

Course Outcomes relevant to this Module:

CO1: Handle various laboratory instruments and practice safety measures
CO2: Prepare various reagents based on solute concentration
CO3: Adjust pH in different solutions

Module Outcome:

After Completion of this module, the student should be able to:
MO1 Articulate and observe safety guidelines in wet lab experiments
MO2 Safely handle and use basic laboratory instruments
MO3 Use proper sterilization techniques to handle different microorganisms
MO4 Prepare reagents, Adjust pH of the solution
MO5 Prepare and maintain culture media
MO6 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**

Module Syllabus: Laboratory safety guidelines, equipment handling, and sterilization techniques, Preparation of buffer, acid, base and pH, media preparation: simple media, special media.

BIN-CC-514 BIOINFORMATICS LAB 1: Module II

Course Outcomes relevant to this Module:

CO4: Identify bacteria by different staining methods
CO5: Do Blood typing to understand the basic antigen antibody reaction
CO6: Demonstrate bacterial growth kinetics

Module Outcome:

After Completion of this module, the student should be able to:
MO1 Understand the morphological features of bacteria including its size and shape by microscopic examination.
MO2 Do different staining methods for the visualization of bacteria
MO3 Determine the amount of antibodies and its action by disc diffusion methods.
MO4 Measure the growth of bacteria
MO5 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**

Module Syllabus: Microscopic examination of bacteria by simple and differential staining, Antibiotic sensitivity test, Microbial growth curve.

BIN-CC-514 BIOINFORMATICS LAB 1: Module III

Course Outcomes relevant to this Module:

CO7: Demonstrate understanding of different types of WBCs
CO8: Detect Blood groups by antigen antibody reaction
CO9: Analyse different stages of Mitotic cell division

Module Outcome:

After Completion of this module, the student should be able to:
MO1 Identify various WBCs from microscopic images by differential staining
MO2 Do slide agglutination test for detecting Blood groups
MO3 Identify various stages of mitotic cell division from onion root tip
MO4 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**

Module Syllabus: Differential staining of WBC, Blood typing, Mitotic cell division- onion root tip

BIN-CC-514 BIOINFORMATICS LAB 1: Module IV

Course Outcomes relevant to this Module:

CO10: Retrieve sequence/structure information from Biological databases

Module Outcome:

After Completion of this module, the student should be able to:
MO1 Retrieve sequence information from NCBI from given ID
MO2 Access protein sequence and structure information from UniProt, PDB
MO3 Retrieve sequence information of proteins from SWISSPROT
MO4 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**

Module Syllabus: Basic use of standard biological databases: NCBI, PDB, SWISSPROT, etc.

BIN-CC-514 BIOINFORMATICS LAB 1: Module V

Course Outcomes relevant to this Module:

CO11: Create small databases, and deploy simple pages in local web browsers

Module Outcome:

After Completion of this module, the student should be able to:
MO1 Execute simple loops in PHP
MO2 Create simple pages using HTML
MO3 Create interactive page using JavaScript
MO4 Do simple operations in MySQL
MO5 Write commands to create small database and relations
MO6 Connect PHP with MySQL
MO7 Insert and retrieve data from MySQL through PHP forms
MO8 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**

Module Syllabus: Database creation and management using PHP-MySQL, Creating Simple webpages with forms; Creating interactive pages with JavaScript, setting up PHP-MySQL environment, retrieving data from PHP-MySQL units PHP Pages, case studies.

BIN-CC-514 BIOINFORMATICS LAB 1: Module VI

Course Outcomes relevant to this Module:

CO12: Process bio sequence using Python

Module Outcome:

After Completion of this module, the student should be able to:
MO1 Write code snippets to do basic sequence operations
MO2 Manipulate bio sequences using string handling functions
MO3 Write programs with file input and output
MO4 Use regular expressions to retrieve patterns in sequence data
MO5 Write programs to perform conversion of bio sequence
MO6 Select and use functions in Bio-python for processing sequence data
MO7 Write programs to process FASTA file
MO8 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**

Module Syllabus: Writing programs using Python features: Sequence analysis – Reading DNA/Protein sequences, sequence length, GC%, handle string search; Programs using string handling function-substring search, count of nucleotides, check RNA/DNA, ORF finding, Transcription, Translation; File handling programs-Read Write Fasta; Program with regular expression, Define RE for a set of sequences, search for sub sequences/patterns, locations; Programs for processing FASTA files.

Semester I	Course Code: BIN-DE-515	Credits:2

BIN-DE-515 SOFT SKILLS

Programme Specific Outcome relevant to this course:

PSO22: Enhanced skills and attitudes for becoming a better learner, thinker, professional and a human being

---

Course Outcome

On completion of the course, students should be able to:

CO1: Awareness about various aspects of personality
CO2: Skills and motivation to develop personality.
CO3: Healthy attitude towards communicating in English
CO4: Better communication skills in English.
CO5: Better personal management skills
CO6: Better living skills
CO7: Better Creative and Critical Thinking skills

Semester I	Course Code: BIN-DE-515	Credits:4

BIN-DE-515 SOFT SKILLS

COURSE CONTENT:

Module I Understanding and Developing Oneself: Personality: (skills, attitudes, body language, inter-personal skills, multiple intelligence etc.), Values: (trust, self-respect, responsibility, openness, sense of justice, honesty, creativity, forgiveness, corporation, constitutional values), Writing a Personal Development Portfolio.

Module II: Communication Skills in English: Attitude towards English, activities to enhance communication skills: speaking practice, communication games, role plays, blog writing, flash presentations (making strong boards), public speaking and vocabulary enhancement.

Module III: Personal Management Skills: Time management, procrastination management, stress management, managing learning, relationship management, work management, working in teams, leadership, conflict management, social pressure/peer-influence management.

Module IV: Living skills: Green living, legal literacy, managing use of digital equipments and social media, Healthy living, visiting foreign countries or interacting with foreign nationals.

Module V: Smart-Thinking Skills: Creativity: Various views on creativity- Habits of critical thinkers, stimulating creativity, obstructions to creativity, creativity and innovation, creativity and craft, visual thinking through mind mapping, creativity exercises.

Module VI Deep-Thinking skills: Critical thinking- Aims of critical thinking- habits of creative thinkers- logical thinking, logical fallacies, limits to critical thinking. The scientific method and research: observation, question, hypothesis, repeatable experiment and critical debates, research attitude in day-to-day life, characteristics of research attitude, Relationship between research, creativity and critical thinking.

BIN-DE-515 SOFT SKILLS: Module I

Course Outcomes relevant to this Module:

CO1: Awareness about various aspects of personality
CO2: Skills and motivation to develop personality.

Module Outcome:

After Completion of this module, the student should be able to:
MO1 Articulate different aspects of personality
MO2 Assess their own Multiple Intelligent types
MO3 Explain about body language
MO4 List and explain values they profess/plan to profess
MO5 Write a personal development portfolio
MO6 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**

Module Syllabus: Understanding and Developing Oneself: Personality: (skills, attitudes, body language, inter-personal skills, multiple intelligence etc.), Values: (trust, self-respect, responsibility, openness, sense of justice, honesty, creativity, forgiveness, corporation, constitutional values), Writing a Personal Development Portfolio.

BIN-DE-515 SOFT SKILLS: Module II

Course Outcomes relevant to this Module:

CO3: Healthy attitude towards communicating in English
CO4: Better communication skills in English.

Module Outcome:

After Completion of this module, the student should be able to:
MO1 Articulate a healthy attitude towards English as an Indian/Global language
MO2 Communicate more fluently and confidently in English
MO3 Make impressive multimedia presentations
MO4 Deliver a confident public speech
MO5 Enhance vocabulary in English language
MO6 Choose effective ways to continuously improve communication skills
MO7 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**

Module Syllabus: Communication Skills in English: Attitude towards English, activities to enhance communication skills: speaking practice, communication games, role plays, blog writing, flash presentations (making strong boards), public speaking and vocabulary enhancement.

BIN-DE-515 SOFT SKILLS: Module III

Course Outcomes relevant to this Module:

CO5: Better personal management skills

Module Outcome:

After Completion of this module, the student should be able to:
MO1 Articulate time management skills that they use/plan to use
MO2 Manage procrastination effectively, and articulate the same
MO3 Mange stress effectively, and articulate the same
MO4 Better their ability to work in teams, and articulate the same
MO5 Mange social pressure/peer-influence effectively, and articulate the same
MO6 Debate different aspects of personal management skills
MO7 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**

Module Syllabus: Personal Management Skills: Time management, procrastination management, stress management, managing learning, relationship management, work management, working in teams, leadership, conflict management, social pressure/peer-influence management.

BIN-DE-515 SOFT SKILLS: Module IV

Course Outcomes relevant to this Module:

CO6: Better living skills

Module Outcome:

After Completion of this module, the student should be able to:
MO1 Identify and practice a few green living skills
MO2 Demonstrate basic legal literacy
MO3 Identify their own choice of a few healthy living ways
MO4 Articulate physical and mental health issues related to overuse of digital equipments and social media
MO5 Regulate their own use of digital equipment's and social media
MO6 Identify and address issues in inter-cultural communications
MO7 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**

Module Syllabus: Living skills: Green living, legal literacy, managing use of digital equipments and social media, Healthy living, visiting foreign countries or interacting with foreign nationals.

BIN-DE-515 SOFT SKILLS: Module V

Course Outcomes relevant to this Module:

CO7: Better Creative and Critical Thinking skills

Module Outcome:

After Completion of this module, the student should be able to:
MO1 Articulate a few alternate views of creativity
MO2 List common habits of creative thinkers
MO3 List common ways of stimulating creativity
MO4 Differentiate instances of creativity and craft
MO5 Develop mind maps on given topics
MO6 Articulate and exemplify the concept of innovation
MO7 Create multiple creative responses to given situations
MO8 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**

Module Syllabus: Smart-Thinking Skills: Creativity: Various views on creativity- Habits of critical thinkers, stimulating creativity, obstructions to creativity, creativity and innovation, creativity and craft, visual thinking through mind mapping, creativity exercises.

BIN-DE-515 SOFT SKILLS: Module VI

Course Outcomes relevant to this Module:

CO7: Better Creative and Critical Thinking skills

Module Outcome:

After Completion of this module, the student should be able to:
MO1 Debate the aims of Critical Thinking
MO2 List the common habits of Critical thinkers
MO3 Identify common logical fallacies in arguments (from media)
MO4 Articulate limits of critical thinking
MO5 Articulate and exemplify the scientific method
MO6 Develop alternate hypothesis, given common observations
MO7 Acquire skills to enhance their own latent critical thinking skills
MO8 Articulate and enhance their own latent research attitude
MO9 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**

Module Syllabus: Deep-Thinking skills: Critical thinking- Aims of critical thinking- habits of creative thinkers- logical thinking, logical fallacies, limits to critical thinking. The scientific method and research: observation, question, hypothesis, repeatable experiment and critical debates, research attitude in day-to-day life, characteristics of research attitude, Relationship between research, creativity and critical thinking.

BIN-DE-515 SOFT SKILLS:
ACTIVITIES, LEARNING RESOURCES & ASSESSMENT

SUGGESTED CLASS ROOM ACTIVITIES:

- Public speech
- Mock Interviews
- Communication Games
- Personal Development Portfolio Presentation
- Team Presentation on Multiple Intelligences
- Multi-media story board writing, making presentation and presenting it.
- Creativity exercises
- Developing Mind-Maps
- Critical Thinking exercises

LEARNING RESOURCES:

References:

1. Allen R. (2005), Boost Your Creativity: Exercises and Advice for Great Creative Thinking., Anova Books
2. Caroselli M. (2004), Quick Wits: 50 Activities for Developing Critical Thinking Skills, Human Resource Development Press.
3. Mak D. K., Mak A. T., Mak A. B. (2009), Solving Everyday Problems with the Scientific Method: Thinking like a Scientist, World Scientific.
4. Oech R. V. (1983), A Whack on the Side of the Head: How to Unlock Your Mind for Innovation, Warner Books**.**
5. Burn G. (2011), Motivation for dummies. John Wiley & Sons.
6. Klaus P. (2009), The Hard Truth about Soft Skills: Soft Skills for Succeeding in a Check on net, HarperCollins.
7. Reddi B. (2012), Soft Skills and Life Skills: The Dynamics of Success. BSC Publishers and Distributors.
8. Sherfield R. M. (2009), Cornerstone: Developing Soft Skills. Pearson Education.
9. Smith J. (1997), How to be a better time manager. Kogan Page.

On-line Resources/MOOCs:

1. Creative problem solving, from University of Minnesota. (Coursera)
2. Creativity, innovation, and change, from Pennsylvania State University. (Coursera)
3. Enhance Your Career and Employability Skills, University of London. (Coursera)
4. Practical Management for Career Readiness, University of California. (Coursera)

ASSESSMENT:

40% Continuous or Formative Assessment (see PG Regulations, in Appendix).
60% End-semester/Summative Assessment: 3 hour written Exam (see Appendix)

Semester I	Course Code: BIN-DE-516	Credits:2

WEB PROGRAMMING AND DATABASES

Programme Specific Outcome relevant to this course:

PSO 3: Skill in designing and developing on-line databases
PSO 5: Basic skill in designing and developing simple websites

---

Course Outcome

On completion of the course, students should be able to:

CO1: Knowledge in basic HTML and PHP
CO2: Design interactive web pages
CO3: Design databases using ER model.
CO4: Design databases using relational model.
CO5: Do relational algebra operations
CO6: Normalize databases and process them with SQL queries.
CO7: Integrate PHP with MySQL databases and deploy it in web pages
CO8: Skill in popular web scripting languages HTML5 and JavaScript.

Semester I	Course Code: BIN-DE-516	Credits:2

BIN-DE-516 WEB PROGRAMMING AND DATABASES

COURSE CONTENT:

Module I: Introduction to web programming: Introduction to web technology: web application, web server, client and server. Web programming, Different languages in web programming. Introduction to HTML: Basics, elements, attributes, text formatting, images, forms and links. Introduction to PHP, PHP environment setup, PHP syntax, data types, variables, operators, control structures, forms.

Module II: Introduction to databases and ER model: Traditional file system, database approach, Overview of DBMS, Advantages of DBMS, data models, schemas, instances and database state, Components of Database, Data independence, 3 schema architecture, Database users, Classification of database systems. E-R modelling: entities, relationships, attributes, mapping constraints, keys, E-R diagrams.

Module III Introduction to relational model and relational algebra: Basic concepts of Relational model: Domains, attributes, keys, tuples, relations, relational database schemas. integrity constraints. Relational algebra: select, project, rename, union, difference, Cartesian product, join. Overview of relational calculus: TRC and DRC.

Module IV Normalization and Introduction to SQL: Conceptual design of relational database, Normalization theory, Functional dependencies, Lossless decomposition of relations, First, Second, Third and Boyce-Codd normal forms. Introduction to SQL: SQL Statement syntax, SQL data types, domains, attributes, constraints, DDL: CREATE, ALTER, DROP. DML: SELECT, INSERT, DELETE and UPDATE and DCL: GRANT and REVOKE. Views in SQL.

Module V: Introduction to MySQL: Introduction to MySQL, connecting to the server, creating database, creating tables, loading data into a table, retrieving information from tables, integration of PHP with MySQL, Introduction to LAMP and WAMP, examples for small database projects.

Module VI: Familiarization of web programming languages: HTML5: elements, attributes, form elements, form attributes, video, audio. CSS: Advantages and limitations of CSS, external, internal, inline, changing fonts, text attributes. JavaScript: Introduction to JavaScript, Advantages of JavaScript, Client-Side and Server-Side JavaScript. Connecting with HTML, JavaScript syntax, variables, operators, objects, events, and strings.

BIN-DE-516 WEB PROGRAMMING AND DATABASES: Module I

Course Outcomes relevant to this Module:

CO1: Knowledge in basic HTML and PHP
CO2: Design interactive web pages

Module Outcome:

After completion of this module, the student should be able to:	TL
MO1 Describe variety of web applications	Un
MO2 Articulate web-server architecture	Un
MO3 List different web programming languages	Re
MO4 Script web pages using HTML including text, images, forms and links	Ap
MO5 Articulate data types in PHP	Un
MO6 Articulate operators in PHP	Un
MO7 Use PHP code and store values in variables	Ap
MO8 Create forms in PHP using GET and POST method	Cr
MO9 Construct simple control structures in PHP	Ap
MO10 **Achieve other outcome dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **Introduction to web programming:** Introduction to web technology: web application, web server, client and server. Web programming, Different languages in web programming. Introduction to HTML: Basics, elements, attributes, text formatting, images, forms and links. Introduction to PHP, PHP environment setup, PHP syntax, data types, variables, operators, control structures, forms.

BIN-DE-516 WEB PROGRAMMING AND DATABASES: Module II

Course Outcomes relevant to this Module:

CO3: Design databases using ER model

Module Outcome:

After completion of this module, the student should be able to:	TL
MO1 Contrast traditional file system and database approach	Un
MO2 Explain advantages of database approach	Un
MO3 Explain the concept of DBMS.	Un
MO4 Compare different data models	An
MO5 Draw schema and explain instances and database state	An
MO6 Explain components of database	Un
MO7 Articulate and exemplify data independence	Un
MO8 Articulate 3 schema architecture	Un
MO9 Explain different database users.	Un
MO10 Articulate E-R modelling, entities and attributes	Un
MO11 Identify and suggest keys in relations	An
MO12 Articulate integrity constraints	Un
MO13 Draw ER diagram of a given database	Ap
MO14 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **Introduction to databases and ER model:** Traditional file system, database approach, Overview of DBMS, Advantages of DBMS, data models, schemas, instances and database state, Components of Database, Data independence, 3 schema architecture, Database users, Classification of database systems. E-R modelling: entities, relationships, attributes, mapping constraints, keys, E-R diagrams.

BIN-DE-516 WEB PROGRAMMING AND DATABASES: Module III

Course Outcomes relevant to this Module:

CO4: Design databases using relational model.
CO5: Do relational algebra operations

Module Outcome:

After completion of this module, the student should be able to:	TL
MO1 Explain fundamental concept of relational model	Un
MO2 Exemplify/ identify attributes, keys and tuples	Ap
MO3 Differentiate between different integrity constraints	Un
MO4 Design Relational model	Cr
MO5 Exemplify relational algebra operations	Un
MO6 Apply select, project, rename, union, difference, Cartesian product and join on given cases	Ap
MO7 Explain TRC and DRC	Un
MO8 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **Introduction to relational model and relational algebra:** Basic concepts of Relational model: Domains, attributes, keys, tuples, relations, relational database schemas. integrity constraints. Relational algebra: select, project, rename, union, difference, Cartesian product, join. Overview of relational calculus: TRC and DRC.

BIN-DE-516 WEB PROGRAMMING AND DATABASES: Module IV

Course Outcomes relevant to this Module:

CO6: Normalize databases and process them with SQL queries.

Module Outcome:

After completion of this module, the student should be able to:	TL
MO1 Identify dependencies in relations	Un
MO2 Articulate lossy and lossless decomposition	Un
MO3 Do 1st, 2nd, 3rd and BCNF normalization in a given relation	Ap
MO4 Exemplify data types in SQL	Un
MO5 Write queries to create tables with different constraints	Ap
MO6 Write queries to drop and alter tables	Ap
MO7 Write queries to SELECT, INSERT, DELETE and UPDATE tables	Ap
MO8 Exemplify GRANT and REVOKE	Un
MO9 Exemplify views in SQL	Un
MO10 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **Normalization and Introduction to SQL:** Conceptual design of relational database, Normalization theory, Functional dependencies, Lossless decomposition of relations, First, Second, Third and Boyce-Codd normal forms. Introduction to SQL: SQL Statement syntax, SQL data types, domains, attributes, constraints, DDL: CREATE, ALTER, DROP. DML: SELECT, INSERT, DELETE and UPDATE and DCL: GRANT and REVOKE. Views in SQL.

BIN-DE-516 WEB PROGRAMMING AND DATABASES: Module V

Course Outcomes relevant to this Module:

CO7: Integrate PHP with MySQL databases and deploy it in web pages

Module Outcome:

After completion of this module, the student should be able to:	TL
MO1 Do simple operations in MYSQL	Ap
MO2 Write queries to create small database	Ap
MO3 Articulate MySQL data types	Un
MO4 Write queries to create tables	Ap
MO5 Connect PHP with mySQL	Ap
MO6 Explain the concept of LAMP	Un
MO7 Insert and retrieve data from MySQL through PHP forms	An
MO8 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **Introduction to MySQL**: Introduction to MySQL, connecting to the server, creating database, creating tables, loading data into a table, retrieving information from tables, integration of PHP with MySQL, Introduction to LAMP and WAMP, examples for small database projects.

BIN-DE-516 WEB PROGRAMMING AND DATABASES: Module VI

Course Outcomes relevant to this Module:

CO8: Skill in popular web scripting languages HTML5 and JavaScript.

Module Outcome:

After completion of this module, the student should be able to:	TL
MO1 Code moderately complex web pages in HTML5	Ap
MO2 Embed audio and video in HTML5	Ap
MO3 Articulate external, internal and inline CSS	Un
MO4 Use CSS for style management	Ap
MO5 Differentiate server-side Vs client-side scripting	Un
MO6 Utilize JavaScript in web pages to make them dynamic	Ap
MO7 Do form validation using HTML, CSS and JavaScript	Ap
MO8 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **Familiarization of web programming languages:** HTML5: elements, attributes, form elements, form attributes, video, audio. CSS: Advantages and limitations of CSS, external, internal, inline, changing fonts, text attributes. JavaScript: Introduction to JavaScript, Advantages of JavaScript, Client-Side and Server-Side JavaScript. Connecting with HTML, JavaScript syntax, variables, operators, objects, events, and strings.

BIN-DE-516 WEB PROGRAMMING AND DATABASES:
ACTIVITIES, ASSESSMENT & LEARNING RESOURCES

SUGGESTED CLASS ROOM OR LAB ACTIVITIES:

- Designing and drawing ER diagram of databases, specifying different constraints.
- Designing a relational model of a database specifying required constraints.
- Create registration forms to input user name and password and display appropriate messages using HTML and PHP.
- Create database with 3 tables in MySQL and do insert, update and delete.
- Connect to a database in MySQL using PHP and do insert, update and delete.

LEARNING RESOURCES:

References:

1. Steven H. (2009), PHP: The complete reference, Tata McGraw-Hill.
2. Connolly T. M., & Begg C. (2015). Database systems: A Practical approach to design, implementation, and management. Pearson Education
3. Elmasri R. (2016). Fundamentals of Database Systems. Pearson Education India
4. Leon A., & Leon M. (2010). Database Management Systems. Vikas Publishing House Pvt. Limited
5. Silberschatz A., Korth H. F., & Sudarshan S. (1997). Database system concepts (Vol. 4). New York: McGraw-Hill
6. Robin N. (2014), Learning PHP, MySQL, & JavaScript with jQuery, CSS & HTML5, O'Reilly.
7. Hugh E. W. & David L., (2004). Web Database Applications with PHP & MySQL, O'Reilly.
8. Jon D. (2008). Beginning Web Programming with Html, Xhtml and CSS. Wiley India Pvt. Ltd. New Delhi
9. Terry F. (2015). Web Development and Design Foundations with HTML5, Pearson Education Limited

On-line Resources/MOOCs:

1. HTML online training (tutorialspoint)
2. PHP Introduction (w3schools)
3. PHP - Introduction (tutorialspoint)
4. Building Web Applications in PHP (Coursera)
5. Introduction of DBMS (geeksforgeeks)
6. Normalization (javapoint)
7. Introduction to Structured Query Language (Coursera)
8. javascript Tutorial (w3schools)
9. HTML5 introduction (w3schools)

ASSESSMENT:

40% Continuous or Formative Assessment (see PG Regulations, in Appendix).
60% End-semester/Summative Assessment: 3 hour written Exam (see Appendix)

Semester II

Course Title	Course Code
NGS Data Analytics I	BIN-CC-521
Fundamentals of Molecular Biology	BIN-CC-522
Computational Genomics	BIN-CC-523
Bioinformatics Lab II	BIN-CC-524
Perl and Bio Perl (E)	BIN-DE-525
Negotiated Studies(E)	BIN-DE-526
Soft Skills (Additional Skill Acquisition Course) (E)	BIN-DE-527

Semester II	Course Code: BIN-CC-521	Credits:4

NGS-I

Programme Specific Outcome relevant to this course:

PSO 8: Thorough knowledge of basic concepts underpinning Bioinformatics & NGS Data
PSO 14: Knowledge and skill in NGS Data Analytics

Course Outcome

On completion of the course, students should be able to:

CO1: Understand the different platforms for first and second generation sequencing
CO2: Understand the advantage and disadvantage of different types of first and second generation sequencing
CO3: Understand the different platforms for third generation sequencing.
CO4: Understand the different types of high throughput sequencing.
CO5: Understand the library preparation methods for different types of high throughput sequencing.
CO6: Understand the types of sequencing methods used in different platforms
CO7: Understand the different file formats used in NGS data analysis
CO8: Understand the different sources of publicly available NGS datasets and how to retrieve them
CO9: Assess the quality of an NGS dataset and how to improve the quality
CO10: Understand the basics of NGS data analysis - assembly and mapping
CO11: Able to assess the quality of NGS data assembly
CO12: Understand the different mapping tools and output formats
CO13: Understand the use of BioCompute Object framework
CO14: Understand the ethics involved in sequencing projects

Semester II	Course Code: BIN-CC-521	Credits:4

NGS Data Analytics -I

COURSE CONTENT:

Module I: History & evolution of sequencing: First-generation technologies – Sanger dideoxy sequencing, Maxam-Gilbert sequencing. Technologies used in Human Genome Project. Shotgun sequencing. Advantages and disadvantages of first generation sequencing. Next (second)-generation sequencing – difference between first and next generation sequencing, NGS platforms – Roche 454, ABI SOLiD, Ion torrent, Illumina. Advantages and disadvantages of second generation sequencing.

Module II: Third-generation sequencing: PacBio, Oxford nanopore. Advantages and disadvantages of third generation sequencing. Comparison of output, accuracy and types of errors of first, second and third generation sequencing technologies. Introduction to NGS technologies: DNA-seq, RNA-seq, ChIP-seq, Hi-C, Metagenomics, Single cell sequencing.

Module III: Preparation for sequencing: Different sample preparation methods for different type of NGS (DNASeq, RNASeq, ChIPSeq, Metagenomics, Single cell), Adaptors, Index, Barcode. Library preparation methods - Bridge amplification, Emulsion PCR. Sequencing methods – sequencing by synthesis, ion semiconductor, SMRT, nanopore.

Module IV: NGS Data formats & Pre-processing: Data formats overview – FASTQ, subreads, nanopore data, single cell data. Single-end, Paired-end, Mate-pair. NGS Data sources – NCBI SRA, EBI-ENA, DDBJ-SRA, GEO; Retrieving data from data sources - SRA toolkit; Aspera connect. Sequence quality measures – Phred quality score. Quality check – tool – FASTQC, Pre-processing: Trimmomatic, Fastx-toolkit.

Module V: Introduction to NGS Data Analysis: Assembly (overview) – principles, output file formats – contigs, scaffolds, assembly quality assessment – N50, total length, no. of contigs/scaffolds; Mapping (overview) – Principles, tools – BWA, Bowtie, SAMtools, output file formats – BAM, SAM, mapping alignment assessment – no. of reads mapped, concordant reads; Visualisation tools – IGV.

Module VI: (Flexi-Module) **BioCompute Object (BCO) Project (overview)**: introduction, utility. Bioethics of sequencing – ownership of DNA sequence, storage and security, insurance.

BIN-CC-521- NGS Data Analytics -I: Module I

Course Outcomes relevant to this Module:

CO1: Understand the different platforms for first and second generation sequencing

CO2: Understand the advantage and disadvantage of different types of first and second generation sequencing

Module Outcome relevant to this Module:

After Completion of this module, the student should be able to:	TL
MO1 Articulate on DNA sequencing	Un
MO2 Outline of first generation sequencing technologies: Sanger dideoxy sequencing & Maxam-Gilbert sequencing	Un
MO3 Discuss Human Genome Project and explain the technology used	Un
MO4 Explain different types of second generation sequencing	Un
MO5 Differentiate first and second generation sequencing	Ev
MO6 Explain Next Generation sequencing	Un
MO7 Assess the pros and cons of second generation sequencing	Ev
MO8 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module I: **History & evolution of sequencing**: First-generation technologies – Sanger dideoxy sequencing, Maxam-Gilbert sequencing. Technologies used in Human Genome Project. Shotgun sequencing. Advantages and disadvantages of first generation sequencing. Next (second)-generation sequencing – difference between first and next generation sequencing, NGS platforms – Roche 454, ABI SOLiD, Ion torrent, Illumina. Advantages and disadvantages of second generation sequencing.

BIN-CC-521- NGS Data Analytics -I: Module II

Course Outcomes relevant to this Module:

CO3: Understand the different platforms for third generation sequencing
CO4: Understand the different types of high throughput sequencing

Module Outcome relevant to this Module:

After Completion of this module, the student should be able to:		TL
MO1	Articulate on third generation sequencing and the different platforms available	Un
MO2	Differentiate between PacBio and Oxford Nanopore sequencing platforms	Un
MO3	Compare the output, accuracy & types of errors of first, second & third generation sequencing	Ev
MO4	Briefly give an overview of the various NGS technologies	Un
MO5	Compare and contrast between DNAseq, RNAseq and Chipseq	Ev
MO6	Articulate on Hi-C sequencing	Un
MO7	Describe on metagenomic sequencing and its uses	Un
MO8	Explain how Single cell RNAseq differ from Standard RNAseq	Un
MO9	**Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module II: **Third-generation sequencing**: PacBio, Oxford nanopore. Advantages and disadvantages of third generation sequencing. Comparison of output, accuracy and types of errors of first, second and third generation sequencing technologies. Introduction to NGS technologies: DNA-seq, RNA-seq, ChIP-seq, Hi-C, Metagenomics, Single cell sequencing.

BIN-CC-521- NGS Data Analytics -I: Module III

Course Outcomes relevant to this Module:

CO5: Understand the library preparation methods for different types of high throughput sequencing.
CO6: Understand the types of sequencing methods used in different platforms

Module Outcome relevant to this Module:

After Completion of this module, the student should be able to:	TL
MO1 Articulate sample preparation methods for different types of high throughput sequencing	Un
MO2 Articulate on multiplexing of samples	Un
MO3 Differentiate Bridge amplification and Emulsion PCR	Ev
MO4 Explain different sequencing methods	Un
MO5 Achieve other outcomes dynamically evolved by the teacher during curricular transaction	

Module III: Preparation for sequencing: Different sample preparation methods for different type of NGS (DNASeq, RNASeq, ChIPSeq, Metagenomics, Single cell), Adaptors, Index, Barcode. Library preparation methods - Bridge amplification, Emulsion PCR. Sequencing methods – sequencing by synthesis, ion semiconductor, SMRT, nanopore.

BIN-CC-521- NGS Data Analytics -I: Module IV

Course Outcomes relevant to this Module:

CO7: Understand the different file formats used in NGS data analysis
CO8: Understand the different sources of publicly available NGS datasets and how to retrieve them
CO9: Assess the quality of an NGS dataset and how to improve the quality.

Module Outcome relevant to this Module:

After Completion of this module, the student should be able to:	TL
MO1 Articulate FASTQ file format and basic terms related to NGS output files	Un
MO2 Articulate the different NGS data sources – NCBI-SRA, EI- ENA, DDBJ-SRA and GEO	Un
MO3 Retrieve data from different NGS data sources using different methods	Ap
MO4 Use SRA toolkit and Aspera connect to download NGS data sets	Ap
MO5 Check the quality of NGS data sets using the tool FASTQC	Ap
M06 Execute different pre-processing tools like Trimmomatic, Fastx-toolkit	Ap
MO7 Achieve other outcomes dynamically evolved by the teacher during curricular transaction	

Module IV: **NGS Data formats & Pre-processing**: Data formats overview – FASTQ, subreads, nanopore data, single cell data. Single-end, Paired-end, Mate-pair. NGS Data sources – NCBI SRA, EBI-ENA, DDBJ-SRA, GEO; Retrieving data from data sources - SRA toolkit; Aspera connect. Sequence quality measures – Phred quality score. Quality check – tool – FASTQC, Pre-processing: Trimmomatic, Fastx-toolkit.

BIN-CC-521- NGS Data Analytics -I: Module V

Course Outcomes relevant to this Module:

CO10: Understand the basics of NGS data analysis - assembly and mapping
CO11: Able to assess the quality of NGS data assembly
CO12: Understand the different mapping tools and output formats

Module Outcome relevant to this Module:

After Completion of this module, the student should be able to:	TL
MO1 Understand the basics of NGS data assembly and mapping	Un
MO2 Assess assembly quality using N50, total length, no: of contigs/scaffolds	Ev
MO3 Understand different tools used for NGS data mapping – BWA, Bowtie	Un
MO4 Understand and analyse the different file formats - SAM, BAM	Un
MO5 Assess mapping quality using no: of reads mapped, concordant reads	Ev
M06 Visualize the mapped data using the tool IGV	Ap
MO7 Achieve other outcomes dynamically evolved by the teacher during curricular transaction	

Module V: **Introduction to NGS Data Analysis**: Assembly (overview) – principles, output file formats – contigs, scaffolds, assembly quality assessment – N50, total length, no. of contigs/scaffolds; Mapping (overview) – Principles, tools – BWA, Bowtie, SAMtools, output file formats – BAM, SAM, mapping alignment assessment – no. of reads mapped, concordant reads; Visualisation tools – IGV.

BIN-CC-521- NGS Data Analytics -I: Module VI

Course Outcomes relevant to this Module:

CO13: Understand the use of BioCompute Object framework
CO14: Understand the ethics involved in sequencing projects

Module Outcome relevant to this Module:

After Completion of this module, the student should be able to:	TL
MO1 Articulate BioCompute Object Project framework and its usage	Un
MO2 Articulate the principles of ethics involved in DNA sequencing	Un
MO3 Elucidate the storage and security challenges involved in sequencing	Un
MO4 Achieve other outcomes dynamically evolved by the teacher during curricular transaction	

Module VI: (Flexi-Module) **BioCompute Object (BCO) Project (overview)**: introduction, utility. Bioethics of sequencing – ownership of DNA sequence, storage and security, insurance.

BIN-CC-521- NGS Data Analytics -I

ACTIVITIES, LEARNING RESOURCES & ASSESSMENT

SUGGESTED CLASSROOM ACTIVITIES:

Genome Assembly Excercises using paper cuttings
Genome Mapping Exercises using paper cuttings
Pedigree Analysis – Variant Calling using gene sequences
Seminars
Completion of an online course

LEARNING RESOURCES:

References:

1. Dr. Michal Janitz. (2008), Next Generation Genome Sequencing: Towards Personalized Medicine, Wiley.
2. Stuart M Brown. (2013), Next-Generation DNA Sequencing Informatics, Cold Spring Harbor Laboratory Press.
3. Sara El-Metwally, Osama M. Ouda & Mohamed Helmy. (2014), Next Generation Sequencing Technologies and Challenges in Sequence Assembly, Springer, New York, NY.
4. Gaurav Sablok, Sunil Kumar, Saneyoshi Ueno, Jimmy Kuo & Claudio Varotto. (2015), Advances in the Understanding of Biological Sciences Using Next Generation Sequencing (NGS) Approaches, Springer International Publishing.
5. Xinkun Wang. (2016), Next-generation sequencing data analysis, CRC Press.
6. Urszula Demkow and Rafał Płoski. (2016), Clinical Applications for Next Generation Sequencing, Academic Press.
7. Masoudi-Nejad, Ali, Narimani, Zahra, Hosseinkhan & Nazanin. (2013), Next Generation Sequencing and Sequence Assembly - Methodologies and Algorithms, Springer, New York.
8. Ion Măndoiu & Alexander Zelikovsky. (2016), Computational Methods for Next Generation Sequencing Data Analysis, Wiley.
9. Jerzy K. Kulski (2016), Next Generation Sequencing - Advances, Applications and Challenges, InTech.
10. Emmanuel A. Kornyo. (2017), A Guide to Bioethics, CRC Press.
11. Richard McCombie W, Elaine R. Mardis, James A. Knowles & John D. McPherson. (2019), Next-Generation Sequencing in Medicine, Cold Spring Harbor Laboratory Press.
12. Wing-Kin Sung. (2020), Algorithms for Next-Generation Sequencing, Chapman and Hall/ CRC.

On-line Courses/MOOCs:

1. Introduction to Genomic Technologies, Coursera
2. Analyze Your Genome!, https://www.edx.org/course/analyze-your-genome
3. Finding Hidden Messages in DNA (Bioinformatics I), https://www.classcentral.com/course/hiddenmessages-3289
4. Genome Sequencing (Bioinformatics II), Coursera
5. Algorithms for DNA Sequencing, https://www.coursera.org/learn/dna-sequencing
6. Genomic Data Science Specialization, Coursera
7. Essentials of Genomics and Biomedical Informatics, EdX

ASSESSMENT:

40% Continuous / Formative Assessment (see PG Regulations in Appendix).
60% End-semester/Summative Assessment: 3 hour written Exam.

Semester II	Course Code: BIN-CC-522	Credits:4

FUNDAMENTALS OF MOLECULAR BIOLOGY

Programme Specific Outcome relevant to this course:

PSO 1: Life Science concepts (especially molecular biology) and skills relevant to CB & B

Course Outcome

On completion of the course, students should be able to:

CO1: Demonstrate understanding of the concept and structure of genes
CO2: Demonstrate understanding the concept of DNA replication, mutation and repair
CO3: Demonstrate understanding the concept of Central Dogma of Molecular Biology and associated knowledge
CO4: Demonstrate understanding the concept of modification and regulation of gene expression
CO5: Demonstrate understanding the concept of recombinant DNA technology used for Genetic Engineering
CO6: Demonstrate the understanding of concept behind and application of selected latest molecular biology techniques

Semester II	Course Code: BIN-CC-522	Credits:4

Fundamentals of Molecular Biology

COURSE CONTENT:

Module I. Gene structure: introns, exons, splicing; Control Elements- Promoters, Enhancers, Silencers, Insulators; Repeats- tandem, microsatellite, mini satellite, inverted repeats.

Module II. DNA Replication, Mutation & Repair: conservative, semi conservative and dispersive models of DNA replication, Prokaryotic and Eukaryotic DNA replication. Mutations- Gene mutations- silent, missense, nonsense and frameshift, Somatic and germline mutations, spontaneous and induced mutations. DNA repair- common types of DNA repair mechanisms- direct repair, base excision & nucleotide excision repair, mismatch repair.

Module III. Central Dogma of Molecular Biology: Transcription- Prokaryotic and Eukaryotic. Translation- Prokaryotic and Eukaryotic. Genetic code, Codon usage bias, Wobble hypothesis.

Module IV. Modification & Regulation of Gene Expression: Post transcriptional modification, post translational modification, gene expression regulation in prokaryotes - principles of gene regulation, negative and positive regulation, concept of operons, regulatory proteins, activators, repressors, regulation of lac operon, gene expression regulation in eukaryotes.

Module V. Gene Manipulation: Genetic Engineering; recombinant DNA technology, Enzymes used in rDNA technology- Endonuclease, Exonuclease, restriction endonucleases, Ligase, Reverse transcriptase, DNA Polymerase; Foreign DNA, Cloning vectors- plasmids, phages, cosmids, BACs, YACs; cDNA, cDNA construction, cDNA library, Genomic library, Steps involved in rDNA technology.

Module VI. (Flexi-Module) Technologies & Applications of Molecular Biology: Technologies- overview of DNA sequencing, PCR, Gel Electrophoresis, Nucleic acid hybridization. Applications: DNA fingerprinting, RNA interference, gene editing, gene therapy.

BIN-CC-522 Fundamentals of Molecular Biology: Module I

Course Outcomes relevant to this Module:

CO1: Demonstrate understanding of the concept and structure of genes

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Diagrammatically represent and explain the structure of genes	Un
MO2 Differentiate between exons and introns in gene structure	An
MO3 Articulate the process of splicing	Un
MO4 Articulate the nature and function of various control elements	Un
MO5 Articulate and exemplify the different types of repeats	Un
MO6 Identify the type of repeats from a given DNA sequence	Ap
MO7 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **Gene structure:** Gene structure, introns, exons, splicing; Control Elements- Promoters, Enhancers, Silencers, Insulators; Repeats- tandem, microsatellite, mini satellite, inverted repeats.

BIN-CC-522 Fundamentals of Molecular Biology: Module II

Course Outcomes relevant to this Module:

CO2: Demonstrate understanding the concept of DNA replication, mutation and repair

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Explain the conservative, semi conservative and dispersive models of DNA replication	Un
MO2 Compare and contrast DNA replication in prokaryotes and eukaryotes	Ev
MO3 List the functions of helicase, topoisomerase, single strand binding proteins, primase & DNA Polymerase at the replication fork	Re
MO4 Differentiate between the synthesis of the leading and lagging strand	Un
MO5 Articulate the various types of gene mutations – silent, missense, nonsense and frameshift	Un
MO6 Distinguish between the effects of mutations in somatic and germ line cells	An
MO7 Articulate the four types of point mutations- transition, transversion, deletion and inversions	Un
MO8 Identify the types of mutation, given a DNA sequence	Ap
MO9 Articulate the difference between spontaneous and induced mutations	Un
MO10 List DNA repair mechanisms for cells with altered or damaged nucleotides	Re
MO11 Articulate the steps of nucleotide excision repair	Un
MO12 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **DNA Replication, Mutation & Repair:** DNA replication models: conservative, semi conservative and dispersive, Prokaryotic and Eukaryotic DNA replication. Mutations- Gene mutations- silent, missense, nonsense and frameshift, Somatic and germline mutations, spontaneous and induced mutations. DNA repair- common types of DNA repair mechanisms- direct repair, base excision & nucleotide excision repair, mismatch repair.

BIN-CC-522 Fundamentals of Molecular Biology: Module III

Course Outcomes relevant to this Module:

CO3: Demonstrate understanding the concept of Central Dogma of Molecular Biology and associated knowledge

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Articulate the process of transcription in prokaryotes & eukaryotes	Un
MO2 Compare and contrast transcription in prokaryotes & eukaryotes	Ev
MO3 Articulate the process of translation in prokaryotes & eukaryotes	Un
MO4 Compare and contrast the process of translation in prokaryotes & eukaryotes	Ev
MO5 Articulate the salient features of genetic code	Un
MO6 Convert a given DNA sequence into its corresponding amino acids based on the standard genetic code	Ap
MO7 Articulate the principle of codon usage bias	Un
MO8 Convert a given amino acid sequence into its DNA sequence, based on the codon usage table	Ap
MO9 Describe the concept of Wobble hypothesis	Un
MO10 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **Central Dogma of Molecular Biology:** Transcription- Prokaryotic and Eukaryotic. Translation- Prokaryotic and Eukaryotic. Genetic code, Codon usage bias, Wobble hypothesis.

BIN-CC-522 Fundamentals of Molecular Biology: Module IV

Course Outcomes relevant to this Module:

CO4: Demonstrate understanding the concept of modification and regulation of gene expression

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Articulate the post transcriptional modifications that leads to a mature mRNA	Un
MO2 Articulate the different types of post translational modifications	Un
MO3 Explain role of regulatory factors involved in transcriptional regulation	Un
MO4 Differentiate between negative and positive regulation of gene expression	An
MO5 Articulate the concept of operon and the regulatory elements	Un
MO6 Describe the organization of the lac operon and its negative and positive control	Un
MO7 Articulate the control points of gene expression regulation in Eukaryotes	Un
MO8 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: Modification & Regulation of Gene Expression: Post transcriptional modification, post translational modification, gene expression regulation in prokaryotes - principles of gene regulation, negative and positive regulation, concept of operons, regulatory proteins, activators, repressors, regulation of lac operon, gene expression regulation in eukaryotes

BIN-CC-522 Fundamentals of Molecular Biology: Module V

Course Outcomes relevant to this Module:

CO5: Demonstrate understanding the concept of recombinant DNA technology used for Genetic Engineering

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Define and explain genetic engineering and recombinant DNA technology	Un
MO2 List and explain the function of various enzymes used in recombinant DNA technology	Un
MO3 Articulate how sticky end & blunt-end DNA fragments are generated	Un
MO4 Calculate the average sizes and numbers of DNA fragments produced by digesting human genomic DNA with a given restriction enzyme	An
MO5 List and explain various vectors used in recombinant DNA technology	Un
MO6 Outline the procedure for cloning a gene into a vector	Un
MO7 Compare and contrast a genomic library with a cDNA library	Ev
MO8 Explain the use of reverse transcriptase in the construction of a cDNA library	Un
MO9 Illustrate the steps involved in rDNA technology	Un
MO10 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **Gene Manipulation:** Genetic Engineering; recombinant DNA technology, Enzymes used in rDNA technology- Endonuclease, Exonuclease, restriction endonucleases, Ligase, Reverse transcriptase, DNA Polymerase; Foreign DNA, Cloning vectors- Plasmids, Phages, Cosmids, BACs, YACs; cDNA, cDNA construction, cDNA library, Genomic library, Steps involved in rDNA technology.

BIN-CC-522 Fundamentals of Molecular Biology: Module VI

Course Outcomes relevant to this Module:

CO6: Demonstrate the understanding of concept behind and application of selected latest molecular biology techniques

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Articulate on DNA sequencing	Un
MO2 Outline the different methods- Maxam-Gilbert, Chain termination, Whole-genome shotgun, Next-generation of DNA sequencing	Un
MO3 Articulate the roles of DNA polymerase, the template, dideoxyribonucleotides and the primer in a Sanger sequencing reaction	Un
MO4 List and explain the function of each component in a PCR reaction	Un
MO5 Articulate working of PCR and steps of PCR cycle	Un
MO6 Summarize the process by which gel electrophoresis separates DNA fragments	Un
MO7 Articulate the function and significance of gel electrophoresis	Un
MO8 Discuss the basic principle of nucleic acid hybridization	Un
MO9 Describe the process of DNA fingerprinting and explore the various uses for DNA fingerprinting	Un
MO10 Interpret given DNA fingerprints to match child to parents, crime scene evidence to suspects etc.	An
MO11 Define RNA interference and describe its molecular mechanism	Un
MO12 Articulate the mechanisms of CRISPR-Cas9 gene-editing technology	Un
MO13 Assess the pros and cons of the CRISPR-Cas9 gene-editing technology	Ev
MO14 Articulate the process, types, uses and potential issues in gene therapy	Un
MO15 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **(Flexi-Module) Technologies & Applications of Molecular Biology:** Technologies- overview of DNA sequencing, Different methods of sequencing- Maxam-Gilbert sequencing, Chain termination, Whole-genome shotgun sequencing, Next-generation sequencing; PCR, Gel Electrophoresis, Nucleic acid hybridization. Applications: DNA fingerprinting, RNA interference, gene editing, gene therapy.

BIN-CC-522 Fundamentals of Molecular Biology
ACTIVITIES, LEARNING RESOURCES & ASSESSMENT

SUGGESTED CLASSROOM ACTIVITIES:

Quiz
Seminar presentations on selected topics
Poster preparation on selected topics
Simple demonstration activities
Completion of a MOOC course

LEARNING RESOURCES:

References:

1. Berg J. M., Tymoczko J. L., & Stryer L. (2002). Biochemistry. W. H.
2. David R. H. (2010). Genetics and molecular biology. Special Indian edition, Tata McGraw Hill Education private limited.
3. Dubey R. C. (1999). A textbook of biotechnology. S. Chand and Co., New Delhi.
4. Gerald K. (2005). Cell and molecular biology: concepts and experiments. John Wiley and Sons, Hoboken, NJ.
5. Hausman R. E., & Cooper G. M. (2004). The cell: a molecular approach. ASM, Washington, DC.
6. Jogdand S. N. (2009). Gene Biotechnology. Himalaya Publishing house.
7. Lewin B. (2011). Lewin's genes X (Vol. 10). J. Krebs, S. T. Kilpatrick, & E. S. Goldstein (Eds.). Jones & Bartlett Learning.
8. Malacinski G. M. (2005). Essentials of molecular biology. Jones & Bartlett Learning.
9. Pierce B. A. (2010). Genetics: A conceptual approach. Macmillan.
10. Russell P. J. (2004). Cell and Molecular Biology. Cengage Learning India Private Limited.
11. Snustad D. P., Simmons M. J., Jenkins J. B., & Crow J. F. (2000). Principles of genetics. John Wiley.
12. Tropp B. E. (2012). Molecular biology: genes to proteins. Jones & Bartlett Publishers.

On-line Resources/MOOCs:

1. Introduction to Biology - The Secret of Life, https://www.edx.org/course/introduction-to-biology-the-secret-of-life-4
2. Molecular Biology - Part 1: DNA Replication and Repair, https://www.edx.org/course/molecular-biology-part-1-dna-replication-and-repair
3. Molecular Biology - Part 2: Transcription and Transposition, https://www.edx.org/course/molecular-biology-part-2-transcription-and-transposition-0
4. Molecular Biology – Part 3: RNA Processing and Translation, https://www.edx.org/course/molecular-biology-part-3-rna-processing-and-translation-0
5. Introduction to Genomic Technologies, https://www.coursera.org/learn/introduction-genomics
6. Introductory course in Biotechnology, https://www.udemy.com/course/introductory-course-in-biotechnology/

ASSESSMENT:

40% Continuous / Formative Assessment (see PG Regulations in Appendix).
60% End-semester/Summative Assessment: 3 hour written Exam.

Semester II	Course Code: BIN-CC-523	Credits:4

COMPUTATIONAL GENOMICS

Programme Specific Outcome relevant to this course:

PSO 3: Concepts and skills in processing bio-sequence data
PSO 13: Knowledge and skill to process DNA/RNA sequence data

Course Outcome

On completion of the course, students should be able to:

CO1: Articulate the basic structural features of DNA, with respect to sequence data
CO2: Analyse DNA sequence data using basic computational genomic tools
CO3: Articulate various types of sequence alignments and their relevance
CO4: Apply DNA scoring schemes and articulate the same
CO5: Do visual comparison of sequence similarity and their interpretation
CO6: Apply pairwise sequence alignment algorithms using selected tools and articulate it
CO7: Apply multiple sequence alignment algorithms using selected tools and articulate it
CO8: Apply different measures of evaluating the quality of an alignment and articulate it
CO9: Articulate evolution and its molecular trace as a backdrop of modern biology
CO10: Articulate the basic types of phylogenetics and terminology thereof
CO11: Apply selected phylogenetic tree construction techniques and tools and articulate it
CO12: Articulate the concept of transcriptomics and its various techniques
CO13: Use various RNA databases and structure prediction tools to study transcriptome and articulate

Semester II	Course Code: BIN-CC-523	Credits:4

Computational Genomics

COURSE CONTENT:

Module I: Functional elements of DNA & its Analysis: Reading frames +1, +2, +3 and -1, -2, -3, ORFs, Codon usage bias, Basic gene statistics – base counts, word (n-mer) frequencies, ORF finder, Splice site recognition, Gene finding, Transcription factor binding site identification, Sequence profiles, Sequence logos, Sequence chromatograms.

Module II: **Sequence alignment:** Concepts and types – pairwise, multiple, global, local, Need of Scoring schemes/ matrices, Simple DNA scoring schemes, Penalizing gaps – End/ middle gaps, opening/ extension gaps, affine gap penalty, identical vs similar matches, effect of low and high gap penalty, Simple pairwise alignment, EMBOSS Needle, Dot plots for visual interpretation for sequence patterns, characteristic patterns of sequence alignment for standard cases.

Module III: Techniques/ Algorithm for sequence alignment: Dynamic programming – Needleman-Wunsch algorithm, Smith-Waterman algorithm, Hand computing of toy alignments, Modification in algorithms for local alignment, Overview of BLAST algorithm, Interpreting the results, E-value, Bit score, Multiple sequence alignment – Need for MSA, Concept of MSA, Approaches to MSA – Heuristics algorithm for multiple sequence alignment – Progressive Alignment vs Iterative alignment, Alignment criteria – SP Measure, MSA tools – ClustalW – options of word size, matrix, gap open, extension, output format – guide tree, Edit alignment using Jalview, Muscle/MAFFT – output formats – generating tree for alignment.

Module IV: Phylogeny: Basic concepts & terminologies: Molecular Evolution-Micro & Macro, Evolution and Speciation, Taxonomy Vs phylogeny, Traditional Vs Molecular phylogeny-Computational phylogeny, Terminology of phylogenetic tree: Root, Branch, Node, Leaf, Clade, Outgroup, Homology, Orthology, Paralogy, Xenology, Gene phylogeny vs Species phylogeny, Different types of trees- Rooted vs. Unrooted trees, Monophyletic vs. Paraphyletic, Dichotomy vs. Polytomy, Phylogram vs. Cladogram, Model Testing, Molecular clock hypothesis.

Module V: Phylogenetic Tree Construction methods: Clustering based -UPGMA and neighbor joining, Optimality based: Fitch-Margoliash and minimum evolution algorithm; Character based -Maximum Parsimony (MP) and Maximum Likelihood (ML) methods; Bayesian inference, Evaluation of phylogenetic trees-reliability and significance; Bootstrapping; Jackknifing, Tools: MEGA, RAxML, FigTree, ETE3.

Module VI: Transcriptomics: Introduction to Transcriptome and Transcriptomics, Types and functions of coding and non-coding RNAs- mRNAs, rRNA, tRNA, lncRNAs, miRNAs, piRNAs, siRNAs, ceRNAs; Basic introduction to Transcriptomic techniques- EST, SAGE/CAGE, Microarray, RNAseq; RNA databases- RNAcentral, ncRNA, miRBase, Rfam, NONCODE; RNA structure prediction tools- RNAFold, RNA123, RNA STRAND, RNA structure; Applications of Transcriptomics.

BIN-CC-523 Computational Genomics: Module I

Course Outcome relevant to the module:

CO1: Articulate the basic structural features of DNA, with respect to sequence data
CO2: Analyse DNA sequence data using basic computational genomic tools

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Identify and write out the six reading frames of a given DNA sequence	Ap
MO2 Articulate codon usage bias	Un
MO3 Compute base count frequencies of a given DNA sequence	An
MO4 Articulate basic principle for identifying ORFs	Un
MO5 Interpret the result of ORF finder tools	An
MO6 Articulate the principle of TF binding site identification	Un
MO7 Use and interpret the result of GenScan	An
MO8 Interpret seq. profiles for identification of seq. conservation and patterns	An
MO9 To interpret given sequence logos and sequence chromatogram	Un
MO10 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: Functional elements of DNA & its Analysis: Reading frames +1, +2, +3 and -1, -2, -3, ORFs, Codon usage bias, Basic gene statistics – base counts, word (n-mer) frequencies, ORF finder, Splice site recognition, Gene finding, Transcription factor binding site identification, Sequence profiles, Sequence logos, Sequence chromatograms.

BIN-CC-523 Computational Genomics: Module II

Course Outcome relevant to the module:

CO3: Articulate various types of sequence alignments and their relevance
CO4: Apply DNA scoring schemes and articulate the same
CO5: Do visual comparison of sequence similarity and their interpretation

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Define & exemplify sequence alignment methods	Un
MO2 Differentiate different types of alignments from given examples	An
MO3 Explain the need for scoring schemes in alignments	Un
MO4 Explain and justify penalizing various types of gaps	Un
MO5 Explain the need for identical and similar matches	Un
MO6 Align given DNA sequences using EMBOSS Needle tool	Ap
MO7 Compute score of an alignment, given an alignment and a scoring matrix	Ap
MO8 Draw Dot plots manually, given toy sequences	Ap
MO9 Identify and interpret standard patterns, given dot plot images	An
MO10 Use Dot plot software on given sequences	Ap
MO11 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: Sequence alignment: Concepts and types – pairwise, multiple, global, local, Need of Scoring schemes/ matrices, Simple DNA scoring schemes, Penalizing gaps – End/ middle gaps, opening/ extension gaps, affine gap penalty, identical vs similar matches, effect of low and high gap penalty, Simple pairwise alignment, EMBOSS Needle, Dot plots for visual interpretation for sequence patterns, characteristic patterns of sequence alignment for standard cases.

BIN-CC-523 Computational Genomics: Module III

Course Outcome relevant to the module:

CO6: Apply pairwise sequence alignment algorithms using selected tools and articulate it
CO7: Apply multiple sequence alignment algorithms using selected tools and articulate it
CO8: Apply different measures of evaluating the quality of an alignment and articulate it

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Articulate and exemplify sequence alignment techniques/algorithms	Un
MO2 Compute a simple Needleman–Wunsch alignment in given worksheet	An
MO3 Compute a simple Smith-Waterman alignment in given worksheet	An
MO4 Describe the general concept of BLAST algorithm	Un
MO5 Use BLAST tool and evaluate the alignment obtained on the basis of E-value and bit score	An
MO6 Articulate & exemplify the concept of MSA	Un
MO7 Critically compare progressive alignment and Iterative alignment	Un
MO8 Articulate the alignment criteria based on SP score	Un
MO9 Use ClustalW/MAFFT/MUSCLE tool and evaluate the alignment obtained	Ap
MO10 Visualize and analyse the ClustalW output using JalView	An
MO11 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: Techniques/Algorithm for sequence alignment, Dynamic programming-Needleman-Wunsch algorithm, Smith-Waterman algorithm, hand computing of toy alignments, modification in algorithms for local alignment, Overview of BLAST algorithm, Interpreting the results, E-value, Bit score, Multiple sequence alignment: Need for MSA; Concept of MSA; Approaches to MSA-Heuristics algorithm for multiple sequence alignment-Progressive Alignment vs Iterative alignment; Alignment criteria: SP Measure, MSA tools: ClustalW –options of word size, matrix, gap open, extension, output format-guide tree, Edit alignment using JalView, MUSCLE/MAFFT- output formats-generating tree for alignment.

BIN-CC-523 Computational Genomics: Module IV

Course Outcome relevant to the module:

CO9: Articulate evolution and its molecular trace as a backdrop of modern biology
CO10: Articulate the basic types of phylogenetics and terminology thereof

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Articulate the concept of molecular evolution	Un
MO2 Compare and contrast micro and macro evolution	Ev
MO3 Differentiate between taxonomy and phylogeny	Ev
MO4 Differentiate between traditional and molecular phylogeny	Ev
MO5 Articulate the concept of computational phylogeny	Un
MO6 Describe and exemplify phylogeny and molecular phylogeny	Un
MO7 Analyse a given phylogenetic tree and label root, node, branch, clade and leaf	An
MO8 Differentiate between crown group and stem group	Ev
MO9 Differentiate between homology, orthology, paralogy and xenology	Ev
MO10 Differentiate between phylogram and cladogram	Ev
MO11 Compare and contrast gene phylogeny and species phylogeny	Ev
MO12 Classify a given phylogenetic tree as rooted and unrooted	An
MO13 Identify dichotomy and polytomy in a given tree topology	Un
MO14 Differentiate between monophyletic and polyphyletic groups	Ev
MO15 Articulate the concept of model testing for phylogenetic tree construction	Un
MO16 Articulate the concept of molecular clock hypothesis	Un
MO17 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: Basic concepts & terminologies: Molecular Evolution-Micro & Macro, Evolution & Speciation, Taxonomy Vs phylogeny, Traditional vs. Molecular phylogeny-Computational phylogeny, Terminology of phylogenetic tree: Root, Branch, Node, Leaf, Clade, Outgroup, Crown group vs. stem group, Homology, Orthology, Paralogy, Xenology, Gene phylogeny Vs Species phylogeny, Different types of trees- Rooted vs. Unrooted trees, Monophyletic vs. Paraphyletic, Dichotomy Vs Polytomy, Phylogram Vs Cladogram, Model Testing, Molecular clock hypothesis

BIN-CC-523 Computational Genomics: Module V

Course Outcome relevant to the module:

CO11: Apply selected phylogenetic tree construction techniques and tools and articulate it

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Articulate the different types of phylogenetic tree construction methods	Un
MO2 Differentiate between UPGMA and Neighbor Joining method	Ev
MO3 Construct phylogenetic trees using UPGMA method	Ap
MO4 Construct phylogenetic trees using NJ method, given a distance matrix	Ap
MO4 Differentiate between Fitch-Margoliash and Minimum Evolution method	Ev
MO5 Differentiate between Maximum Parsimony and Maximum Likelihood method	Ev
MO6 Calculate the informative and uninformative sites for a given set of sequence	Ap
MO7 Construct a phylogenetic tree based on the MP method given a set of sequences	Ap
MO8 Articulate the Bayesian method of phylogenetic tree construction	Un
MO9 Articulate the need for tree evaluation and explain Bootstrapping/Jackknifing methods	Un
MO10 Use and interpret the tools- MEGA/RAxML for phylogenetic tree construction	Ap
MO11 Visualize the constructed phylogenetic tree using FigTree/ ETE3	Ap
MO12 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: Phylogenetic Tree Construction methods: Clustering based -UPGMA and Neighbor joining, Optimality based: Fitch-Margoliash and Minimum evolution algorithm; Character based -Maximum Parsimony (MP) and Maximum Likelihood (ML) methods; Bayesian inference, Evaluation of phylogenetic trees-reliability and significance; Bootstrapping; Jackknifing, Tools: MEGA, RAxML, FigTree, ETE3.

BIN-CC-523 Computational Genomics: Module VI

Course Outcome relevant to the module:

CO12: Articulate the concept of transcriptomics and its various techniques
CO13: Use various RNA databases and structure prediction tools to study transcriptome and articulate

Module Outcome:

After Completion of this module, the student should be able to	TL
MO1 Define and articulate transcriptome and transcriptomics	Un
MO2 Articulate various types and functions of coding and non-coding RNAs	Un
MO3 Articulate the various transcriptomic techniques- EST, SAGE/CAGE, microarray and RNAseq	Un
MO4 Critically compare microarray vs RNAseq for gene expression profiling	Ev
MO5 Articulate the features of various RNA databases	Un
MO6 Compare RNAcentral database with ncRNA database	Ev
MO7 Compare various RNA structure prediction tools	Cr
MO8 Articulate the various applications of transcriptomics	Un
MO9 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: Transcriptomics: Introduction to Transcriptome and Transcriptomics, Types and functions of coding and non-coding RNAs- mRNAs, rRNA, tRNA, lncRNAs, miRNAs, piRNAs, siRNAs, ceRNAs; Basic introduction to Transcriptomic techniques- EST, SAGE/CAGE, Microarray, RNAseq; RNA databases- RNAcentral, ncRNA, miRBase, Rfam, NONCODE; RNA structure prediction tools- RNAFold, RNA STRAND, RNAstructure; Applications of Transcriptomics.

BIN-CC-523 Computational Genomics
ACTIVITIES, LEARNING RESOURCES & ASSESSMENT

SUGGESTED CLASSROOM ACTIVITIES:

Quiz
Seminar presentations on selected topics
Worksheet exercises on sequence alignment
Completion of a MOOC course

LEARNING RESOURCES:

References:

1. Claverie J. M., & Notredame C. (2011). Bioinformatics for dummies. John Wiley & Sons.
2. Lesk A. (2013). Introduction to Bioinformatics. Oxford University Press.
3. Mount D. W. (2004). Sequence and genome analysis. Bioinformatics: Cold Spring Harbour Laboratory Press: Cold Spring Harbour, 2.
4. Pevsner J. (2009). Bioinformatics and Functional genomics. John Wiley & Sons
5. Rastogi S. C., Mendiratta N., & Rastogi P. (2013). Bioinformatics: Methods and Applications: (Genomics, Proteomics and Drug Discovery). PHI Learning Pvt. Ltd.
6. Salemi M., Lemey P., & Vandamme A. M. (Eds.). (2009). The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing. Cambridge University Press
7. Semple C., & Steel M. (2003). Phylogenetics. Oxford Lecture Series in Mathematics and its Applications.
8. Xiong J. (2006). Essential bioinformatics. Cambridge University Press
9. Zvelebil M. J., & Baum J. O. (2008). Understanding bioinformatics. Garland Science.

On-line Resources/MOOCs:

1. Bioinformatic methods I, conducted by University of Toronto. https:// www.Coursera.org/ course/bioinfomethods1
2. Bioinformatics algorithms (Part 1), conducted by University of California San Diego. https://www.coursera.org/course/bioinformatics
3. Bioinformatics: introduction and methods conducted by Peking University. https://www.coursera.org/course/pkubioinfo
4. Computational Molecular Evolution, conducted by Technical University of Denmark. https://www.coursera.org/course/molevol.
5. Experimental Genome Science, conducted by University of Pennsylvania. https://www.cour sera.org/course/genomescience

ASSESSMENT:

40% Continuous / Formative Assessment (see PG Regulations in Appendix).
60% End-semester/Summative Assessment: 3 hour written Exam.

Semester II	Course Code: BIN-CC-524	Credits: 3

BIOINFORMATICS LAB II

Programme Specific Outcome relevant to this course:

PSO 6: Basic wet lab skill and exposure to molecular biology experiments
PSO 15: Basic *in- silico* laboratory skills relevant to CB & B

Course Outcome

On completion of the course, students should be able to:

CO1: Hands-on experience in basic Molecular Biology experiments
CO2: Hands-on experience in Computational Genomics software
CO3: Hands-on experience in NGS data download, quality check and preprocessing
CO4: Hands-on experience in *de-novo* whole genome assembly, mapping and its visualization
CO5: Hands-on experience in using Perl to do basic DNA sequence data processing

Semester II	Course Code: BIN-CC-524	Credits: 3

Bioinformatics Lab II

COURSE CONTENT:

Module I: Molecular Biology: Extraction and Quantification of DNA, Electrophoresis: AGE, Polymerase Chain Reaction

Module II: Genomics: ORF Finder, Gene finder, Pairwise sequence alignment, BLAST, EMBOSS, Dot plot analysis, Multiple sequence alignment, Phylogenetic analysis – MEGA

Module III: NGS I: Download NGS datasets from SRA, ENA, DDBJ, GEO – Direct download, using SRA toolkit, using Aspera Connect; NGS data quality check – FastQC; Data Preprocessing – Trimmomatic, Fastx toolkit, Cutadapt

Module IV: NGS I: *De-novo* whole genome assembly of prokaryotes – Velvet, SPAdes, Unicycler; Prokaryotic genome mapping – Bowtie2, BWA, SAM to BAM conversion – Samtools; Genome Visualization – IGV, Tablet

Module V: **Perl:** Data analysis using Perl programming language, Validating DNA/ RNA/ Amino acid sequences, Finding complement & reverse complement of DNA sequence, writing a sequence in Fasta format, computing the nucleotide composition of a DNA sequence. Computing the amino acid composition of a protein sequence

Module VI: **Perl:** Finding the AT Composition of a given DNA sequence, Finding the GC Composition of a given DNA sequence, Finding the ORFs in a given DNA sequence, Transcribe a DNA sequence into RNA, Translate the given DNA sequence into corresponding amino acid sequence, Mapping amino acid sequence with different physiochemical features like hydrophobicity, finding n-mer frequencies in DNA and amino acid sequences

BIN-CC-524 Bioinformatics Lab II: Module I

Course Outcome relevant to the module:

CO1: Hands-on experience in basic Molecular Biology experiments

Module Outcome:

After Completion of this module, the student should be able to:
MO1 Isolate DNA using Phenol chloroform method
MO2 Analyse DNA by Agarose gel electrophoresis
MO3 Quantify DNA using Diphenylamine method
MO4 Perform polymerase chain reaction using basic protocols
MO5 Articulate the principles of the above experiments

Module Syllabus: The laboratory work will consist of the following experiments 1. Isolation of DNA from Bacteria or Plant using phenol chloroform methods 2. Analyse the DNA by Agarose gel electrophoresis 3. Quantify the DNA by Diphenylamine method 4. Amplify the DNA by PCR.

BIN-CC-524 Bioinformatics Lab II: Module II

Course Outcome relevant to the module:

CO2: Hands-on experience in Computational Genomics software

Module Outcome:

After Completion of this module, the student should be able to:
MO1 Locate ORFs in a given sequence using ORF finder
MO2 Locate genes in a given sequence using GENSCAN
MO3 Perform pairwise sequence alignment of DNA sequences using BLAST & EMBOSS
MO4 Select appropriate scoring schemes and perform pairwise sequence alignment with EMBOSS
MO5 Draw Dot plot of two given DNA/ protein sequences
MO6 Differentiate between the nature of the two given dot plots
MO7 Perform multiple sequence alignment of the given sequence using ClustalW
MO8 Derive phylogenetic tree of given sequences using MEGA

Module Syllabus: The laboratory work will consist of 5-10 experiments 1. Finding ORF of a given sequence using ORF finder 2. Predict the location of the gene from given sequences using GENSCAN 3. Perform pairwise sequence alignment of nucleotide/ protein sequences using BLAST and EMBOSS. 4. Draw Dot plot for visual representation of optimal alignment between two sequences 5. Perform multiple sequence alignment using ClustalW 6. Construct phylogenetic tree using MEGA

BIN-CC-524 Bioinformatics Lab II: Module III

Course Outcome relevant to the module:

CO3: Hands-on experience in NGS data download, quality check and preprocessing

Module Outcome:

After Completion of this module, the student should be able to:
MO1 Download NGS datasets from different sequence databases
MO2 Check quality of NGS data
MO3 Remove adapters from NGS data
MO4 Clip overrepresented sequences from NGS data

Module Syllabus: **Module III: NGS I:** Download NGS datasets from SRA, ENA, DDBJ, GEO – Direct download, using SRA toolkit, using Aspera Connect; NGS data quality check – FastQC; Data Preprocessing – Trimmomatic, Fastx toolkit, Cutadapt.

BIN-CC-524 Bioinformatics Lab II: Module IV

Course Outcome relevant to the module:

CO4: Hands-on experience in *de-novo* whole genome assembly, mapping and its visualization

Module Outcome:

After Completion of this module, the student should be able to:
MO1 Run different whole genome assembly software
MO2 Assemble prokaryotic whole genomes
MO3 Execute prokaryotic genome mapping
MO3 Convert SAM files to BAM files
MO4 Visualize mapped reads with respective to reference using IGV and Tablet

Module Syllabus: **NGS I:** *De-novo* whole genome assembly of prokaryotes – Velvet, SPAdes, Unicycler; Prokaryotic genome mapping – Bowtie2, BWA, SAM to BAM conversion – Samtools; Genome Visualization – IGV, Tablet.

BIN-CC-524 BIOINFORMATICS LAB II: Module V

Course Outcome:

CO5: Hands-on experience in using Perl to do basic DNA sequence data processing

Module Outcome:

After Completion of this module, the student should be able to:
MO1 Validate whether a sequence is DNA / RNA/ Protein
MO2 Find the complement and reverse complement of DNA sequence
MO3 Find the presence of the sequence repeat – “CAG” and display the count of the repeat if it is present.
MO4 Compute the nucleotide composition of a DNA sequence
MO5 Compute the amino acid composition of a protein sequence
MO6 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**

Module Syllabus: **Perl:** Data analysis using Perl programming language, Validating DNA/ RNA/ Amino acid sequences, Finding complement & reverse complement of DNA sequence, writing a sequence in Fasta format, computing the nucleotide composition of a DNA sequence. Computing the amino acid composition of a protein sequence.

BIN-CC-524 BIOINFORMATICS LAB II: Module VI

Course Outcomes

CO5: Hands-on experience in using Perl to do basic DNA sequence data processing

Module Outcome:

After Completion of this module, the student should be able to:
MO1 Find the AT composition of a DNA sequence
MO2 Compute the GC composition of a DNA sequence
MO3 Find ORF in a given DNA sequence
MO4 Transcribe a DNA sequence into RNA
MO5 Translate a DNA sequence into corresponding amino acid sequence
MO6 Map amino acid sequence with different physicochemical properties
MO7 Compute n-mer frequencies in DNA and amino acid sequences
MO8 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**

Module Syllabus: **Perl:** Finding the AT Composition of a given DNA sequence, Finding the GC Composition of a given DNA sequence, Finding the ORFs in a given DNA sequence, Transcribe a DNA sequence into RNA, Translate the given DNA sequence into corresponding amino acid sequence, Mapping amino acid sequence with different physicochemical features like hydrophobicity, finding n-mer frequencies in DNA and amino acid sequences.

Semester II	Course Code: BIN-DE-525	Credits: 2

PERL AND BIOPERL

Programme Specific Outcome relevant to this course:

PSO 7: Advanced skill in selected industry-required programming language

---

Course Outcome

On completion of the course, students should be able to:

CO1: Acquire basic Perl programming skills
CO2: Write Perl programs including control structures basic sequence processing
CO3: Use/Predict effect of Regular expression in Perl
CO4: To handle database connectivity to Perl and write CGI scripts in Perl
CO5: To handle data using OOPS concepts in Perl
CO6: To acquire skills in basic BioPerl

Semester II	Course Code: BIN-DE-525	Credits: 2

Perl and BioPerl

COURSE CONTENT:

Module I: Evolution & Environment, Features of Perl, Perl data structures – Scalars, Array and Hashes, Perl Operators – Arithmetic Operators, Relational Operators, Logical Operators, Bitwise Operators, Assignment Operators, Ternary Operator and string operators, Manipulating DNA/RNA/Protein sequence data

Module II: Control Structures – if-else, while, for, foreach, until, Loop control – last, next, built in functions, User defined functions (Subroutine), File handling, Using Perl for ORF Finding, GC content, Translation of DNA to Protein

Module III: Regular Expressions: Features and advantages, Match operation, Substitute operation, Translate operation, Boundary, Alternative matching, Grouping, Character class, Metacharacters, Special character class, Repetition, Anchors, Nucleotide frequency of DNA, Transcribe DNA into RNA, Complement and reverse complement of DNA, Motif finding in DNA string, Regex Applications: Extract Email Headers, Spam filtering

Module IV: DB connection & disconnection, SQL statements with INSERT, SELECT, UPDATE, DELETE. CGI programming– Client Vs Server-Side Scripting-Concept of CGI-Setting CGI option to server-Connecting forms with server scripts-Data transfer between forms and server, GET & POST methods, Environment variables, Simple examples.

Module V: Overview of Principles of Object Orientation, Creating Classes, Instance Methods, Special Methods, Class Variables, Inheritance, Polymorphism, Type Identification, Overview of the Exception Model.

Module VI: (Flexi-Module) BioPerl & References– Basics, Overview of Bio Perl objects, seq objects, Sequence homology-BLAST+ usage, Annotated database sequences, using references

BIN-DE-525 Perl and BioPerl: Module I

Course Outcomes relevant to this Module:

CO1: Acquire basic Perl programming skills

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Compare and contrast Perl scripting with C language	Un
MO2 Articulate the steps in setting the environmental variables in Perl	Un
MO3 List the features of Perl	Re
MO4 List & exemplify various data structures in Perl	Re
MO5 Predict value of Perl script involving operators	Ev
MO6 Use various Perl operators	Ap
MO7 Illustrate data types & conversions in Perl	Ap
MO8 Translate given problem statement into program logic	Ap
MO9 Exemplify string handling operators in Perl	Un
MO10 Uses of Scalar and Array Variables to manipulate DNA/RNA/Protein sequence data	Ap
MO11 Concatenate DNA fragments	Ap
MO12 Predict the output of string processing statements	An
MO13 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **I. Perl Basics:** Evolution & Environment, Features of Perl, Perl data structures – Scalars, Array and Hashes, Perl Operators- Arithmetic Operators, Relational Operators, Logical Operators, Bitwise Operators, Assignment Operators, Ternary Operator and string operators, Manipulate DNA/RNA/Protein sequence data

BIN-DE-525 Perl and BioPerl: Module II

Course Outcomes relevant to this Module:

CO2: Write Perl programs including control structures basic sequence processing

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Use control structures in Perl script	Ap
MO2 Design & Write code snippets using various control flow structures for given requirements	Ap
MO3 Predict output of Perl code snippets including control structures	An
MO4 Design & Write programs with loops incorporating last and next statements	Ap
MO5 Use subroutine to translate DNA to protein	Ap
MO6 “Write” own functions for given tasks and to call them	Cr
MO7 “Write” Perl script for developing ORF finder	Cr
MO8 “Write” Perl script to find GC content	Cr
MO6 “Write” programs to read from a file and write to a file	Ap
MO7 Import Perl package using require/use functions	Ap
MO8 “Write “a Perl script for sending and validating email	Cr
MO9 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: II. Functions: Control Structures – if-else, while, for, foreach, until, Loop control – last, next, built in functions, User defined functions (Subroutine), ORF Finding, GC content, Translation of DNA to Protein, File handling, Perl Packages – installation, Require/Use functions, Foo, Calculator, sending/validating email

BIN-DE-525 Perl and BioPerl: Module III

Course Outcomes relevant to this Module:

CO3: Use/Predict effect of Regular expression in Perl

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Articulate features & syntax of Perl regular expression for pattern matching	Un
MO2 Design & Write regular expression to match simple character class	Ap
MO3 Use Regex to compute the nucleotide frequency	Ap
MO4 Use Regex to transcribe DNA into RNA	Ap
MO5 Use Regex to find complement/reverse complement of DNA	Ap
MO6 Use Regex to find motif in a DNA string	Ap
MO7 List out the special characters in Perl Regex	Re
MO8 Predict the result (string set) of applying regular expression	An
MO9 Write regular expression for greedy match using repetition qualifiers	Cr
MO10 Demonstrate regular expression group method	Un
MO11 Use regex method to split a given string	Ap
MO12 Write Perl script to extract Email Headers/Spam filtering	Cr
MO13 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **Regular expression:** Features and advantages, Match operation, Substitute operation, Translate operation, Boundary, Alternative matching, Grouping, Character class, Metacharacters, Special character class, Repetition, Anchors, Nucleotide frequency of DNA, Transcribe DNA into RNA, Complement and reverse complement of DNA, Motif finding in DNA string, Regex Applications: Extract Email Headers, Spam filtering

BIN-DE-525 Perl and BioPerl: Module IV

Course Outcomes relevant to this Module:

CO4: To handle database connectivity to Perl and write CGI scripts in Perl

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Articulate the concept of CGI	Un
MO2 Use database package in Perl to connect to a database	Ap
MO2 "Write" queries to create tables with different constraints	Cr
MO3 Predict result of the given queries	An
MO4 "Write" a simple CGI program for a form	Cr
MO5 "Write" a CGI program to extract environmental variables	Cr
MO6 "Write" a CGI program to pass information using GET method and POST method	Cr
MO7 "Write" simple CGI scripts for solving given problem using server-side processing	Cr
MO8 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **Database & CGI:** DB connection & disconnection, SQL statements with INSERT, SELECT, UPDATE, DELETE. CGI programming –Client Vs Server-Side Scripting-Concept of CGI-Setting CGI option to server-Connecting forms with server scripts-Data transfer between forms and server, GET & POST methods Environment variables Simple examples.

BIN-DE-525 Perl and BioPerl: Module V

Course Outcomes relevant to this Module:

CO5: To handle data using OOPS concepts in Perl

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Explain & Exemplify the concept of Object Orientation	Un
MO2 Creating and using Objects different methods	Ap
MO2 Articulate principles of object-oriented programming	Un
MO3 Illustrate polymorphism in Perl snippet	An
MO4 Design & Write a simple program using type identification	Cr
MO5 Using exception classes in Perl	Ap
MO6 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **Module V:** Overview of Principles of Object Orientation, Creating Classes, Instance Methods, Special Methods, Class Variables, Inheritance, Polymorphism, Type Identification, Overview of the Exception Model.

BIN-DE-525 Perl and BioPerl: Module VI

Course Outcomes relevant to this Module:

CO6: To acquire skills in basic BioPerl

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Articulate the use of BioPerl	Un
MO2 Manipulate sequence objects using basic sequence functions	Ap
MO3 “Write” a sequence object to a Fasta file	Ap
MO4 Retrieve a sequence object from Swissprot	Ap
MO5 “Translate” a DNA sequence object to a protein sequence object	Ap
MO6 “Blast” a sequence object using BLAST+ and retrieve results	Ap
MO7 Use of Bio::SeqFeature objects /Bio: Annotation	Ap
MO8 Create a variable reference in Perl	Cr
MO9 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **(Flexi-Module) BioPerl & References–** Basics, Overview of Bio Perl objects, seq objects, Sequence Homology, BLAST+ usage, Annotated database sequences, using references

BIN-DE-525 Perl and BioPerl
ACTIVITIES, LEARNING RESOURCES & ASSESSMENT

SUGGESTED CLASSROOM ACTIVITIES:

Program Portfolio: Students shall submit a portfolio with at least 20 Perl programs
White Board Test/challenge

LEARNING RESOURCES:

References:

1. Schwartz R. L., & Phoenix T. (2001). Learning Perl. O'Reilly & Associates, Inc.
2. Tisdall J. (2003). Mastering Perl for bioinformatics. "O'Reilly Media, Inc."
3. Tom C., Jon O., Larry W., Brian F. (2012). Programming Perl, 4th Edition, Unmatched power for text processing and scripting. O'Reilly Media, Inc.

On-line Resources/MOOCs:

1. http://www.bioPerl.org/wiki/BioPerl_Tutorial
2. http://www.ebi.ac.uk/~lehvasla/bioPerl/BioPerlOverview.html
3. http://learn.Perl.org/tutorials/
4. https://www.Perl.org/books/beginning-Perl/
5. https://hackr.io/tutorials/learn-Perl

ASSESSMENT:

40% Continuous / Formative Assessment (see PG Regulations in Appendix).
60% End-semester Evaluation/Summative Assessment: 3 hour written Exam.

Semester II	Course Code: BIN-DE-526	Credits: 2

NEGOTIATED STUDIES (E)

Programme Specific Outcome relevant to this course:

PSO 17: Awareness of emerging trends and concepts in CB & B

Course Outcome

On completion of the course, students should be able to:

To be Negotiated

Semester II	Course Code: BIN-DE-526	Credits: 2

Negotiated Studies (E)

Course Outcome

To be Negotiated

Syllabus

To be Negotiated

Negotiated Studies are included in the MSc Computational Biology & Bioinformatics Programme to make its syllabus dynamic. Students can propose courses for offer depending on their interest. The proposal should be made with involvement of at least 50% of the students in the programme. The proposal shall be duly considered by the Department Council and shall be developed into a course, stating outcomes and sample questions as in existing courses. The Council, in rare cases, may also choose not to offer the course, for reasons to be recorded. The students shall be given an opportunity to feedback on the course. Once developed and approved by the Department Council, the same shall be intimated to the CSSAC for approval, at least 3 weeks before the commencement of the corresponding semester.

Semester II	Course Code: BIN-DE-527	Credits:2

BIN-DE-527 SOFT SKILLS

Programme Specific Outcome relevant to this course:

PSO22: Enhanced skills and attitudes for becoming a better learner, thinker, professional and a human being

---

Course Outcome

On completion of the course, students should be able to:

CO1: Awareness about various aspects of personality
CO2: Skills and motivation to develop personality
CO3: Healthy attitude towards communicating in English
CO4: Better communication skills in English
CO5: Better personal management skills
CO6: Better living skills
CO7: Better Creative and Critical Thinking skills

Course content and other modules same as that of BIN-DE-515

Semester III

Course Title	Course Code
NGS Data Analytics II	BIN-CC-531
Computational Proteomics	BIN-CC-532
Advanced Topics in Computational Biology	BIN-CC-533
Bioinformatics Lab III	BIN-CC-534
Programming in R (E)	BIN-DE-535
Seminar (E)	BIN-DE-536
Soft Skills (Additional Skill Acquisition Course) (E)	BIN-DE-537

Semester III	Course Code: BIN-CC-531	Credits:4

NGS Data Analytics-II

Programme Specific Outcome relevant to this course:

PSO 8: Thorough knowledge of basic concepts underpinning Bioinformatics & NGS Data
PSO 14: Knowledge and skill in NGS Data Analytics

Course Outcome

On completion of the course, students should be able to:

CO1: Understanding genome assembly and its application
CO2: Understanding different types and methods of assembly
CO3: Be familiar with different assembly algorithms
CO4: Hand-on experience with different genome assembly software
CO5: Be familiar with genome assembly assessment methods
CO6: Understand RNAseq technology
CO7: Be familiar with the general structure of plant chloroplast and mitochondrial genome
CO8: Be familiar with different classes of repeats in the genome
CO9: Apply techniques learnt to an independent data set

Semester III	Course Code: BIN-CC-531	Credits:4

NGS Data Analytics-II

COURSE CONTENT:

Module I: Introduction to genome assembly: Introduction to DNA assembly, K-mer, repeats, contig, scaffold, denovo assembly, reference based assembly, applications of DNA assembly (overview) – whole genome assembly, hybrid assembly, transcriptome assembly, metagenome genome study, organelle genome assembly

Module II: Assembly algorithms and assembly assessment: Mapping-based method, OLC-based method, DBG-based method and greedy based-algorithms, scaffolding, genome alignment, genome assembly assessment – N50, L50, coverage, k-mer coverage, genome completeness, no. of contigs/scaffolds. Tools – Velvet, SPAdes, Canu, DBG2OLC

Module III: Transcriptome assembly: RNA-seq overview, workflow, Mapping RNA-seq reads, denovo vs referenced based transcriptome assembly, splice variants, Trinity (*de novo*), functional annotation. ExN50.

Module IV: Organelle genome assembly: Introduction to chloroplast genome and its features, introduction to mitochondrial genome and its features, assembly methods for organelle genome. Applications of organelle genome.

Module V: Repeat annotation: Repeats – types & classification-tandem repeats, satellite DNA microsatellite/SSR, Direct repeats, inverted repeats, palindromic repeats, interspersed repeats, transposable elements, Barbara McClintock, LINE, SINE, DNA transposons, retro-transposons, genetic and evolutionary significance of repeats, application of repeats, repeat databases – Rpbase, Methods of repeat Identification- Ab initio & Homology based methods.

Module VI: (Flexi module- Only for Internal Assessment. Lecturers may expand and/ or interpret the syllabus to update it or suit the particular cohort in any way): Case study of genome/transcriptome/organelle assembly.

BIN-CC-531-NGS Data Analytics II: Module I

Course Outcomes relevant to this Module:

CO1: Understanding genome assembly and its application
CO2: Understanding different types and methods of assembly

Module Outcome relevant to this Module:

After Completion of this module, the student should be able to:	TL
MO1 Explain DNA assembly and its applications	Un
MO2 Define K-mer	Un
MO3 Articulate and exemplify the different types of repeats	Un
MO4 Differentiate between contig and scaffold	Ev
MO5 Articulate and exemplify different types of assembly	Ev
MO6 Differentiate between whole genome and organelle genome assembly	Ev
MO7 Achieve other outcomes dynamically evolved by the teacher during curricular transaction	

Module Syllabus: **Introduction to genome assembly:** Introduction to DNA assembly, K-mer, repeats, contig, scaffold, denovo assembly, reference based assembly, applications of DNA assembly (overview) – whole genome assembly, hybrid assembly, transcriptome assembly, metagenome genome study, organelle genome assembly

BIN-CC-531- NGS Data Analytics II: Module II

Course Outcomes relevant to this Module:

CO3: Be familiar with different assembly algorithms
CO4: Hand-on experience with different genome assembly softwares
CO5: Be familiar with genome assembly assessment methods

Module Outcome relevant to this Module:

After Completion of this module, the student should be able to:	TL
MO1 Compare different genome assembly algoithms	An
MO2 Execute genome assembly using different tools	Ap
MO3 Execute genome assembly completeness analysis	Ap
MO4 Articulate the difference between contig and scaffold, N50 and L50	Un
MO5 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **Assembly algorithms and assembly assessment:** Mapping-based method, OLC-based method, DBG-based method and greedy based-algorithms, scaffolding, genome alignment, genome assembly assessment – N50, L50, coverage, k-mer coverage, genome completeness, no. of contigs/scaffolds. Tools – Velvet, SPAdes, Canu, DBG2OLC

BIN-CC-531- NGS Data Analytics II: Module III

Course Outcomes relevant to this Module:

CO6: Understand RNAseq technology

Module Outcome relevant to this Module:

After Completion of this module, the student should be able to:	TL
MO1 Articulate on RNAseq technology	Un
MO2 Discuss on the workflow of RNaseq data analysis	Un
MO3 Differentiate between de novo and reference based transcriptome assembly	Un
MO4 Describe how splice variants can be analysed from RNAseq data	Un
MO5 Perform a denovo based RNAseq data analysis using Trinity suite	Ap
M06 Discuss the significance of ExN50	Un
MO7 Articulate how functional annotation is performed on RNAseq data and its significance	Un
MO8 Achieve other outcomes dynamically evolved by the teacher during curricular transaction	

Module Syllabus: **Transcriptome assembly:** RNA-seq overview, workflow, Mapping RNA-seq reads, denovo vs referenced based transcriptome assembly, splice variants, Trinity (*de novo*), functional annotation. ExN50.

BIN-CC-531- NGS Data Analytics -II: Module IV

Course Outcomes relevant to this Module:

CO7: Be familiar with the general structure of plant chloroplast and mitochondrial genome

Module Outcome relevant to this Module:

After Completion of this module, the student should be able to:	TL
MO1 Articulate the structural features of plant chloroplast genome	Un
MO2 Articulate the structural features of mitochondrial genome	Un
MO3 Articulate the evolutionary significance of organelle genome	Un
MO4 Achieve other outcomes dynamically evolved by the teacher during curricular transaction	

Module Syllabus: **Organelle genome assembly:** Introduction to chloroplast genome and its features, introduction to mitochondrial genome and its features, assembly methods for organelle genome. Applications of organelle genome.

BIN-CC-531- NGS Data Analytics -II Module V

Course Outcomes relevant to this Module:

CO8: Be familiar with different classes of repeats in the genome

Module Outcome relevant to this Module:

After Completion of this module, the student should be able to:	TL
MO1 Articulate general features of DNA transposons	Un
MO2 Articulate general features of RNA transposons	Un
MO3 Execute repeat identification using different bioinformatics tools	Ap
MO4 Achieve other outcomes dynamically evolved by the teacher during curricular transaction	

Module Syllabus: **Repeat annotation:** Repeats – types & classification-tandem repeats, satellite DNA microsatellite/SSR, Direct repeats, inverted repeats, palindromic repeats, interspersed repeats, transposable elements, Barbara McClintock, LINE, SINE, DNA transposons, retro-transposons, genetic and evolutionary significance of repeats, application of repeats, repeat databases – Rpbase, Methods of repeat Identification- Ab initio & Homology based methods.

BIN-CC-531- NGS Data Analytics -II Module VI

Course Outcomes relevant to this Module:

CO9: Apply techniques learnt to an independent data set

Module Outcome relevant to this Module:

After Completion of this module, the student should be able to:	TL
MO1 Perform a case study analysis on any one of the technologies learnt in the previous modules	Ap
MO2 Achieve other outcomes dynamically evolved by the teacher during curricular transaction	

Module Syllabus: (Flexi module- Only for Internal Assessment. Lecturers may expand and/ or interpret the syllabus to update it or suit the particular cohort in any way): Case study of genome/transcriptome/organelle assembly.

BIN-CC-531- NGS Data Analytics -II

ACTIVITIES, LEARNING RESOURCES & ASSESSMENT

SUGGESTED CLASSROOM ACTIVITIES:

Quiz
Seminar presentations on selected topics
Completion of a MOOC course

LEARNING RESOURCES:

References:

1. Arthur M. Lesk, (2007). Introduction to Genomics. Oxford University Press.
2. Sandy B. Primrose and Richard Twyman (2008). Principles of Genome Analysis and Genomics (Third Edition). Blackwell Publishing.
3. Sara El-Metwally, Osama M. Ouda, Mohamed Helmy (2014). Next Generation Sequencing Technologies and challenges in sequence assembly. Springer-Verlag New York.
4. Ali Masoudi-Nejad, Zahra Narimani, Nazanin Hosseinkhan (2013). Next Generation Sequencing and Sequence Assembly:Methodologies and Algorithms. Springer New York.
5. Adam Voshall (2018). Next-Generation Transcriptome Assembly: Strategies and Performance Analysis. IntechOpen.
6. Michael Chandler, Martin Gellert (2020). Mobile DNA III. Wiley.

On-line Courses/MOOCs:

1. Introduction to Next-Generation sequencing (NGS) https://www.coursera.org/lecture/wgs-bacteria/introduction-to-next-generation-sequencing-ngs-hhpIu

ASSESSMENT:

40% Continuous / Formative Assessment (see PG Regulations in Appendix).
60% End-semester/Summative Assessment: 3 hour written Exam.

Semester III	Course Code: BIN-CC-532	Credits:4

COMPUTATIONAL PROTEOMICS

Programme Specific Outcome relevant to this course:

PSO18: Knowledge and skill to process protein sequence data
PSO19: Concepts and skills in modelling protein structure

Course Outcome

On completion of the course, students should be able to:

CO1: Knowledge of basic concepts and biological background of computational Proteomics
CO2: Working knowledge of primary, secondary, tertiary & quaternary protein structures
CO3: Have knowledge about enzymes
CO4: Demonstrate understanding of the concept behind enzyme activity and enzyme kinetics and structure of genes
CO5: Knowledge of basic laboratory technologies in proteomics
CO6: Familiarity with protein identification and structure determination techniques
CO7: Familiarity with standard protein databases and visualization tools
CO8: Have working knowledge of software tools used to process molecular data files
CO9: Have working knowledge of concepts & techniques related to AA sequence alignment
CO10: Do amino acid sequence alignments and interpret the results
CO11: Demonstrate understanding of multilevel structure of proteins
CO12: Assess and interpret protein structure prediction models

Semester III	Course Code: BIN-CC-532	Credits:4

COMPUTATIONAL PROTEOMICS

COURSE CONTENT:

Module I: Protein Basics: Proteins as work-horse molecules, protein diversity, types of proteins and examples; Proteins as amino acid chains - basic structure of amino acids, classification of amino acids; Peptides, Polypeptides, Backbone Primary & secondary tertiary and quaternary structures. Protein folding, Protein stability, Protein denaturation; Case study of Hemoglobin and NSP3 (SARS-CoV-2).

Module II: Enzymes: Nomenclature, Classification - IUB system, Characteristics of enzymes, lock and key model, enzyme substrate complex. substrate specificity, coenzymes. Rate of enzyme reaction, kinetics, inhibition, effect of pH and temperature, allosteric enzymes. Enzyme Kinetics: Michaelis - Menten Equation, steady state enzyme kinetics. Vmax and Km. Enzyme inhibition - types of inhibitors - competitive, non-competitive and un-competitive, mode of action.

Module III: Brief Overview of laboratory techniques: Protein analysis: SDS Page, 2D Gel Electrophoresis, Liquid chromatography; Protein identification: LC-MS; Protein structure determination: X-ray Crystallography, NMR spectroscopy.

Module IV: Databases: Sequence: UniprotKB, Structure: PDB, Structural Classification- SCOP and CATH; Visualization tools: RasMol, Swiss PDB viewer, Overview of ExPASy Proteomic Tools. Active site prediction using Casp2 Server, protein interaction database: BioGRID, IntAct, protein interaction network: STRING, EIIP. Functional annotation and Protein families; ProtoNet-tracing Protein families, ProtoNet based tools.

Module V: Basic Computational Proteomics: AA Sequence Alignment: Review of basic sequence alignment concepts, Aligning amino acid sequences, scoring matrices (PAM & BLOSUM), scoring schemes, Bit scores & e-values, Dot Plots, Algorithm for Global alignment (Needleman & Wunch) and Local alignment (Smith-Waterman): Hand computing alignment, Online tools: EMBOSS Needle & BLAST; Molecular Phylogeny using AA sequences, Case Study of Covid-19.

Module VI: Protein Structure Prediction: Secondary structural elements –backbone, domains, side chains, native state/conformation Backbone flexibility- Φ and ψ, α and β propensities, Ramachandran plot, Prediction methods : Chou-Fasman Method & GOR Method; Tertiary Structure Prediction: Need for structure prediction, role of hydro-phobicity and chaperones in protein folding, Levinthal's paradox, Denovo Vs Ab-initio approaches to folding prediction, CASP for assessing structure models; Homology Modeling: Swiss Modeller for homology Modelling; Case study of Papain-like protease (NSP3) of Coronavirus.

BIN-CC-532- COMPUTATIONAL PROTEOMICS: Module I

Course Outcomes relevant to this Module:

CO1: Knowledge of basic concepts and biological background of computational Proteomics
CO2: Working knowledge of primary, secondary, tertiary & quaternary protein structures

Module Outcomes relevant to this Module:

After Completion of this module, the student should be able to:	TL
MO1 Draw and explain the basic structure of amino acids and their classification	Un
MO2 Differentiate primary, secondary, tertiary and quaternary structure of protein and articulate each	An
MO3 Articulate various secondary structural features of protein, given a 3Dmodel	Un
MO4 Articulate protein folding and its role in protein stability	Un
MO5 Articulate the concept of denaturation and structural deformity of protein	Un
MO6 Demonstrate familiarity with Haemoglobin and NS3 protease	Un
MO7 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: Protein Basics: Proteins as work-horse molecules, protein diversity, types of proteins and examples; Proteins as amino acid chains - basic structure of amino acids, classification of amino acids; Peptides, Polypeptides, Backbone Primary & secondary tertiary and quaternary structures. Protein folding, Protein stability, Protein denaturation; Case study of Hemoglobin and NSP3 (SARS-CoV-2).

BIN-CC-532- COMPUTATIONAL PROTEOMICS: Module II

Course Outcomes relevant to this Module:

CO3: Have knowledge about enzymes
CO4: Demonstrate understanding of the concept behind enzyme activity and enzyme kinetics and structure of genes

Module Outcome relevant to this Module:

After Completion of this module, the student should be able to:	TL
MO1 Articulate on enzymes and its classification based on IUB system	Un
MO2 Demonstrate understanding of lock and key model of enzyme substrate binding	Un
MO3 Describe substrate specificity and coenzymes	Un
MO4 Describe various factors affecting enzyme activity	Un
MO5 Articulate the concepts of enzyme kinetics	Un
MO6 Articulate Michaelis - Menten Equation	Un
MO7 Interpret different types of enzyme inhibition	An
MO8 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **Enzymes:** nomenclature, Classification - IUB system, Characteristics of enzymes, lock and key model, enzyme substrate complex. substrate specificity, coenzymes. Rate of enzyme reaction, kinetics, inhibition, effect of pH and temperature, allosteric enzymes. Enzyme Kinetics: Michaelis - Menten Equation, steady state enzyme kinetics. Vmax and Km. Enzyme inhibition - types of inhibitors - competitive, non-competitive and un-competitive, mode of action.

BIN-CC-532- COMPUTATIONAL PROTEOMICS: Module III

Course Outcomes relevant to this Module:

CO5: Knowledge of basic laboratory technologies in proteomics
CO6: Familiarity with protein identification and structure determination techniques

Module Outcome relevant to this Module:

After Completion of this module, the student should be able to:	TL
MO1 Articulate on various protein analysis techniques	Un
MO2 Briefly discuss the protein analysis by peptide gel electrophoresis and Liquid chromatography techniques	Un
MO3 Articulate on the Protein identification methods	Un
MO4 Describe LC-MS based protein identification	Un
MO5 Demonstrate different methods for protein structure determination	Un
MO6 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: Brief Overview of laboratory techniques: Protein analysis: SDS Page, 2D Gel Electrophoresis, Liquid chromatography; Protein identification: LC-MS; Protein structure determination: X-ray Crystallography, NMR spectroscopy.

BIN-CC-532- COMPUTATIONAL PROTEOMICS: Module IV

Course Outcomes relevant to this Module:

CO7: Familiarity with standard protein databases and visualization tools
CO8: Have working knowledge of software tools used to process molecular data files

Module Outcome relevant to this Module:

After Completion of this module, the student should be able to:	TL
MO1 Demonstrate familiarity with protein databases	Un
MO2 Use protein visualization software	Ap
MO3 Predict active site using Casp2 server & articulate the same	Ap
MO4 Demonstrate features of protein interaction databases: BioGrid, IntAct	Un
MO5 Demonstrate features of protein interaction network	Un
MO6 Demonstrate protein functional classification and annotation using ProtoNet	Ap
MO7 Achieve other outcomes dynamically evolved by the teacher during curricular transaction	

Module Syllabus **Databases:** Sequence: UniprotKB, Structure: PDB, Structural Classification-SCOP and CATH; Visualization tools: RasMol, Swiss PDB viewer, Overview of ExPASy Proteomic Tools. Active site prediction using Casp2 Server, protein interaction database: BioGRID, IntAct, protein interaction network: STRING, EIIP. Functional annotation and Protein families; ProtoNet-tracing Protein families, ProtoNet based tools.

BIN-CC-532- COMPUTATIONAL PROTEOMICS: Module V

Course Outcomes relevant to this Module:

CO9: Have working knowledge of concepts & techniques related to AA sequence alignment
CO10: Do amino acid sequence alignments and interpret the results

Module Outcome relevant to this Module:

After Completion of this module, the student should be able to:	TL
MO1 Articulate basic sequence alignment methods	An
MO2 Interpret given sequence alignment	Ap
MO3 Demonstrate knowledge of deriving PAM scoring matrices	Un
MO4 Compare PAM and BLOSUM matrices	Un
MO5 Hand compute alignment scores, given scoring schemes and alignments	Ap
MO6 Hand compute local and global alignments	Ap
MO7 Emboss Needle & BLAST to do alignment & interpret the results	Ap
MO8 Develop phylogenetic tree of AA sequences using PHYLIP & articulate the same	Ap
MO9 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **Basic Computational Proteomics:** AA Sequence Alignment: Review of basic sequence alignment concepts, Aligning amino acid sequences, scoring matrices (PAM & BLOSUM), scoring schemes, Bit scores & e-values, Dot Plots, Algorithm for Global alignment (Needleman & Wunch) and Local alignment (Smith-Waterman): Hand computing alignment, Online tools: EMBOSS Needle & BLAST; Molecular Phylogeny using AA sequences, Case Study of Covid-19.

BIN-CC-532- COMPUTATIONAL PROTEOMICS: Module VI

Course Outcomes relevant to this Module:

CO11: Demonstrate understanding of multilevel structure of proteins
CO12: Assess and interpret protein structure prediction models

Module Outcome relevant to this Module:

After Completion of this module, the student should be able to:	TL
MO1 Identify secondary structure elements, given an image of protein model	Ap
MO2 Use and interpret Ramachandran Plot for structure validation	An
MO3 Predict secondary structure using Chou-Fausman/GOR Methods and interpret the results	Ap
MO4 Articulate tertiary structure and associated phenomenon & concepts	An
MO5 Articulate need and approaches to folding prediction	Ap
MO6 Assess structure models using CASP & articulate the same	An
MO7 Model protein structure, using Homology Modelling & articulate the same	Ap
MO8 Achieve other outcomes dynamically evolved by the teacher during curricular transaction	

Module Syllabus: **Protein Structure Prediction:** Secondary structural elements –backbone, domains, side chains, native state/conformation Backbone flexibility- Φ and ψ, α and β propensities, Ramachandran plot, Prediction methods : Chou-Fasman Method & GOR Method; Tertiary Structure Prediction: Need for structure prediction, role of hydro-phobicity and chaperones in protein folding, Levinthal's paradox, Denovo Vs Ab-initio approaches to folding prediction, CASP for assessing structure models; Homology Modeling: Swiss Modeller for homology Modelling; Case study of Papain-like protease (NSP3) of Coronavirus.

BIN-CC-532- COMPUTATIONAL PROTEOMICS

ACTIVITIES, LEARNING RESOURCES & ASSESSMENT

SUGGESTED CLASSROOM ACTIVITIES:

Quiz
Seminar presentations on selected topics
Completion of a MOOC course
Case studies on homology modeling with different softwares

LEARNING RESOURCES:

References:

1. Kessel A., Ben-Tal N, (2018). Introduction to Proteins: Structure, Function, and Motion. CRC Press.
2. Eidhammer I., Jonassen, I., & Taylor, W. R. (2004). Protein Bioinformatics: An algorithmic approach to sequence and structure analysis (pp. 3-23). J. Wiley & Sons
3. Higgins, D., & Taylor, W. (2000). Bioinformatics: sequence, structure, and databanks: a practical approach. Oxford University Press, Inc.
4. Krane D. E. (2003). Fundamental concepts of Bioinformatics. Pearson Education India.
5. Lesk A. (2013). Introduction to bioinformatics. Oxford University Press.
6. Mount, D. W. (2004). Sequence and genome analysis. Bioinformatics: Cold Spring Harbour Laboratory Press: Cold Spring Harbour, 2.
7. Orengo C., Jones, D. T., & Thornton, J. M. (2003). Bioinformatics: Genes, proteins and computers. Garland Science.
8. Palzkill T. (2002). Kluwer. Proteomics. Springer.
9. Pevsner J. (2009). Bioinformatics and functional genomics. John Wiley & Sons

On-line Courses/MOOCs:

1. Introduction to Proteomics: Swayam: https://www.classcentral.com/course/swayam-introduction-to-proteomics-7910

ASSESSMENT:

40% Continuous / Formative Assessment (see PG Regulations in Appendix).
60% End-semester/Summative Assessment: 3 hour written Exam.

Semester III	Course Code: BIN-CC-533	Credits:4

Advanced Topics in Computational Biology

Programme Specific Outcome relevant to this course:

PSO 10: Knowledge in the field of Machine/Deep Learning
PSO 11: Skill in developing simple Deep Learning models
PSO20: Knowledge in concept and approaches used in drug design
PSO21: Moderate skill in using basic tools of in-silico drug design

Course Outcome

On completion of the course, students should be able to:

CO1: Demonstrate understanding on the historic evolution of AI and be familiar with the basic concepts, nature and scope of the field of AI
CO2: Demonstrate understanding of architectures and working of ANN/MLP
CO3: Develop a small scale ANN and critically evaluate its performance
CO4: Understand and apply basic image processing including filtering
CO5: Articulate the architecture and working of a CNN
CO6: Develop a small scale CNN and critically evaluate its performance
CO7: Have knowledge on popular applications of deep learning
CO8: Have molecular view of drug
CO9: Have working knowledge of software tools used to process molecular data files
CO10: Have Knowledge about vaccines
CO11: Have molecular view of diseases, drug and drug-target molecules
CO12: Have working knowledge of software tools used to process molecular data files
CO13: Have knowledge about Immune response
CO14: Contrast traditional and *in-silico* drug design
CO15: Working knowledge of *in-silico* drug design pipeline
CO16: Working knowledge of Auto dock software

Semester III	Course Code: BIN-CC-533	Credits:4

Advanced Topics in Computational Biology

COURSE CONTENT:

Module I. Introduction to Artificial Intelligence (AI) and Artificial Neural Networks (ANNs): AI- Historic evolution of AI; Overview of definitions and branches of AI; Machine learning as a subset of AI; Learning Types: supervised/unsupervised/semi-supervised learning, transfer learning; Concepts of Deep learning, Major ML problem categories: Classification, Regression and Clustering. Artificial Neural Networks (ANNs**)-** Learning Process: Feature extraction, labelling, training, testing; Generalization: under-fitting and over-fitting; Concept of learning as determination of connection weights (Only concept, based on scalar example y=Wx); Architecture of MLPs: hidden, input and output layers; Activation function; Error/loss function; Concept of back-propagation algorithm. Training/Validating/Testing; Gradient Descent; Stochastic Gradient Descent; Trainable parameters vs hyper-parameter; Performance measures for ML models: accuracy, sensitivity/recall and specificity.

Module II. Basics of Image Processing and Convolution Neural Networks (CNNs): Basics of Image Processing- Basics of Image sampling and quantization; Basic concepts of simple image processing; Process of convolution; Concept of image filtering; Filtering as a means of feature extraction; Depth; Stride; Zero Padding; Use of ReLu on feature Map; Max & average Pooling. Convolution Neural Networks (CNNs): Architecture of convolution layer; Filter/Kernel; Feature Map; Fully connected/classification layer; Learning of Weights and Filter coefficients; Repetition of convolution/classification layers to achieve Deep Learning; Hyper Parameters in CNN, Developing a small-scale CNN for digit recognition.

Module III. Miscallaneous topics (Only overview) Popular applications of CNN., LeNet, AlexNet, VGGNet, R-CNN; **RNN**, LSTM, Deep Generative Models; Generative adversarial networks (GAN); Generative Pre-trained Transformer 3 (GPT-3); Driverless cars, Face Detection Cameras, Sentiment Analysis, Auto correction, Machine Translation, Assisted writing and coding, Generative AI.

Module IV: Drugs and vaccines: Drug Molecules: Molecular and chemical properties of drugs, Lipinski's rule of five, Ligands, small molecules, Leads, Phytochemicals (example: Curcumin), peptide drugs (Oxytocin). Mechanism of drug action: Agonist, antagonist, Functional groups of drugs, Small Molecule Databases: PubChem,DrugBank. Representation of Drugs: Smile notation, IUPAC name, Chemical formula, molecular descriptors, 2D representation, Formats: SDF, MOL, MOL2. QSAR. Vaccines, Drug Vs Vaccines

Module V: Drug Targets and immune response: Molecular Concept of Disease, Drug-Targets; Types of Drug targets: Proteins/Enzymes, Receptor Proteins, Nucleic Acids & other drug targets, Active sites and pharmacophores, allosteric sites, Overview of SARS-CoV-2 Drug Targets, Intermolecular binding force of drugs with targets. Protein interaction database: BioGRID, IntAct, protein interaction network: STRING, Overview of Immune response: antigens & antibodies.

Module VI: Computer Aided Drug Design: Traditional Drug Discovery: Blind search/Serendipitous discovery -limitations in high throughput screening; In-silico Drug discovery Pipeline: Disease Pathways in KEGG, Target identification & validation, Active site identification, pharmacophore, removal of water molecules, Lead/Ligand identification, Binding energy prediction by docking (Auto dock), Energy Minimisation, concept of Molecular Dynamics, concept of ADME; Overview of in-vitro & in vivo clinical trials, and approval IP issues related to drugs IdMOC for drug testing.

BIN-CC-533-Advanced Topics in Computational Biology: Module I

Course Outcomes relevant to this Module:

CO1: Demonstrate understanding on the historic evolution of AI and be familiar with the basic concepts, nature and scope of the field of AI

CO2: Demonstrate understanding of architectures and working of ANN/MLP

CO3: Develop a small scale ANN and critically evaluate its performance

Module Outcome:

After Completion of this module, the student should be able to:		TL
MO1	Narrate historic evolution of AI	Un
MO2	Exemplify and articulate nature and scope of different branches of AI	Un
MO3	Contrast AI, Machine learning & Deep learning	Un
MO4	State and exemplify major ML problem categories	Un
MO5	Articulate the concept of learning as determination of weights	Un
MO6	Articulate the typical architecture of an MLP	Un
MO7	Explain the concept of backpropagation algorithm	Un
MO8	Compute forward pass, given a simple MLP	Ap
MO9	Define and differentiate training, testing and validation	An
MO10	Define, exemplify and contrast learning parameters Vs hyper parameters	An
MO11	**Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **Introduction to Artificial Intelligence (AI) and Artificial Neural Networks (ANNs): AI-** Historic evolution of AI; Overview of definitions and branches of AI; Machine learning as a subset of AI; Learning Types: supervised/unsupervised/semi-supervised learning, transfer learning; Concepts of Deep learning, Major ML problem categories: Classification, Regression and Clustering. Artificial Neural Networks (ANNs**)-** Learning Process: Feature extraction, labelling, training, testing; Generalization: under-fitting and over-fitting; Concept of learning as determination of connection weights (Only concept, based on scalar example y=Wx); Architecture of MLPs: hidden, input and output layers; Activation function; Error/loss function; Concept of back-propagation algorithm. Training/Validating/Testing; Gradient Descent; Stochastic Gradient Descent; Trainable parameters vs hyper-parameter; Performance measures for ML models: accuracy, sensitivity/recall and specificity

BIN-CC-533-Advanced Topics in Computational Biology: Module II

Course Outcomes relevant to this Module:

CO4: Understand and apply basic image processing including filtering
CO5: Articulate the architecture and working of a CNN
CO6: Develop a small scale CNN and critically evaluate its performance

Module Outcome:

After Completion of this module, the student should be able to:		TL
MO1	Articulate the basics of image processing	Un
MO2	Compute filtered image, given filter coefficients	Ap
MO3	Articulate use of filters as feature extractors	Un
MO4	Articulate the structure and working of the convolution layer	Un
MO5	Articulate how deep learning is achieved in CNN	Un
MO6	Articulate Learning as determination of weights &filter coefficients & tuning hyper-parameters	Un
MO7	**Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: Basics of Image Processing and Convolution Neural Networks (CNNs): Basics of Image Processing- Basics of Image sampling and quantization; Basic concepts of simple image processing; Process of convolution; Concept of image filtering; Filtering as a means of feature extraction; Depth; Stride; Zero Padding; Use of ReLu on feature Map; Max & average Pooling. Convolution Neural Networks (CNNs): Architecture of convolution layer; Filter/Kernel; Feature Map; Fully connected/classification layer; Learning of Weights and Filter coefficients; Repetition of convolution/classification layers to achieve Deep Learning; Hyper Parameters in CNN, Developing a small scale CNN for digit recognition.

BIN-CC-533-Advanced Topics in Computational Biology: Module III

Course Outcomes relevant to this Module:

CO7: Have knowledge on popular applications of deep learning

Module Outcome:

After Completion of this module, the student should be able to:		TL
MO1	Exemplify popular applications of CNN in various fields	Un
MO2	Briefly describe architecture of LeNet/AlexNet/VGGNet	Un
MO3	Articulate models other than CNN such as RNN, GAN, GPT-3 etc	Un
MO4	**Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **Miscallaneous topics: (Only overview)** Popular applications of CNN., LeNet, AlexNet, VGGNet, R-CNN; **RNN**, LSTM, Deep Generative Models; Generative adversarial networks (GAN); Generative Pre-trained Transformer 3 (GPT-3); Driverless cars, Face Detection Cameras, Sentiment Analysis, Auto correction, Machine Translation, Assisted writing and coding, Generative AI.

BIN-CC-533-Advanced Topics in Computational Biology: Module IV

Course Outcomes relevant to this Module:

CO8: Have molecular view of drug
CO9: Have working knowledge of software tools used to process molecular data files
CO10: Have Knowledge about vaccines

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Articulate molecular and chemical properties of Drugs	Ap
MO2 Demonstrate understanding of molecular/chemical concepts related to drugs	Un
MO3 Describe and exemplify mechanism of drug action	Un
MO4 Demonstrate familiarity with small molecule databases	Un
MO5 Interpret various structural representation of drugs	Ap
MO6 Articulate the concepts of QSAR	Ap
MO7 Articulate different types of vaccines	Ap
MO8 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **Drugs and vaccines:** Drug Molecules: Molecular and chemical properties of drugs, Lipinski's rule of five, Ligands, small molecules, Leads, Phytochemicals (example: Curcumin), peptide drugs (Oxytocin). Mechanism of drug action: Agonist, antagonist, Functional groups of drugs, Small Molecule Databases: PubChem, DrugBank. Representation of Drugs: Smile notation, IUPAC name, Chemical formula, molecular descriptors, 2D representation, Formats: SDF, MOL, MOL2. QSAR. Vaccines, Drug Vs Vaccines

BIN-CC-533-Advanced Topics in Computational Biology: Module V

Course Outcomes relevant to this Module:

CO11: Have molecular view of diseases, drug and drug-target molecules
CO12: Have working knowledge of software tools used to process molecular data files
CO13: Have knowledge about Immune response

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Articulate molecular concept of diseases & drugs	An
MO2 List different types of drug targets	Un
MO3 List/Describe intermolecular binding forces of drugs with targets	Ap
MO4 Describe important SARS-CoV-2 drug targets	Un
MO5 Demonstrate features of protein interaction databases: BioGrid, IntAct	Un
MO6 Demonstrate features of protein interaction network	Un
MO7 Articulate immune response, antigens & antibodies	Ap
MO8 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: Drug Targets and immune response: Molecular Concept of Disease, Drug-Targets; Types of Drug targets: Proteins/Enzymes, Receptor Proteins, Nucleic Acids & other drug targets, Active sites and pharmacophores, allosteric sites, Overview of SARS-CoV-2 Drug Targets, Intermolecular binding force of drugs with targets. Protein interaction database: BioGRID, IntAct, protein interaction network: STRING, Overview of Immune response: antigens & antibodies.

BIN-CC-533-Advanced Topics in Computational Biology: Module VI

Module Course Outcomes relevant to this Module:

CO14: Contrast traditional and *in-silico* drug design
CO15: Working knowledge of *in-silico* drug design pipeline
CO16: Working knowledge of Auto dock software

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Compare traditional and i*n-silico* drug discovery	Un
MO2 List phases of *in-silico* drug discovery pipeline	Un
MO3 Articulate processing of drug and drug-target molecules for docking	An
MO4 Define basic terms related to drug-discovery pipeline	Un
MO5 Articulate in-vitro and in-vivo clinical trials	An
M06 Articulate concept of molecular dynamics	Un
MO7 Articulate need and process of ADME prediction.	An
MO8 Articulate IP issues relate to drug development & use.	An
MO9 Achieve other outcomes dynamically evolved by the teacher during curricular transaction	

Module Syllabus: **Computer Aided Drug Design:** Traditional Drug Discovery: Blind search/Serendipitous discovery -limitations in high throughput screening; In-silico Drug discovery Pipeline: Disease Pathways in KEGG, Target identification & validation, Active site identification, pharmacophore, removal of water molecules, Lead/Ligand identification, Binding energy prediction by docking (Auto dock), Energy Minimisation, concept of Molecular Dynamics, concept of ADME; Overview of in-vitro & in vivo clinical trials, and approval IP issues related to drugs IdMOC for drug testing.

BIN-CC-533-Advanced Topics in Computational Biology

ACTIVITIES, LEARNING RESOURCES & ASSESSMENT

SUGGESTED CLASSROOM ACTIVITIES:

Quiz
Seminar presentations on selected topics
Completion of a MOOC course
Completion of a case study
Internal Test

LEARNING RESOURCES:

References:

1. Alpaydin E. (2020), “Introduction to machine learning”, MIT Press.
2. Heaton J. (2013), “Artificial intelligence for humans. Heaton Research Incorporated”,
3. Bostrom N. (2017), “Superintelligence”, Dunod.
4. Mueller J. P. & Massaron, L. (2016), “Machine learning for dummies”, John Wiley & Sons.
5. Kulkarni P. & Joshi P. (2015), “Artificial intelligence: building intelligent systems”, PHI Learning Pvt. Ltd.
6. Mitchell T. M. (2013), “Machine learning”, McGraw-Hill, International Student Edition.
7. Conway D. & White J. (2012), “Machine learning for hackers”, O'Reilly Media Inc.
8. Graham L. P., (2013). An introduction to Medicinal Chemistry. Oxford University Press, Inc
9. Graham L. P., (2015). BIOS instant notes, Medicinal chemistry. Garland Science.
10. Robert M. S., Janet Finer., Computational and structural approaches to drug discovery, RSC publishing, University of California, San Francisco, USA.
11. Charifson P. S. (1997). Practical application of computer-aided drug design. Marcel Dekker, Inc.
12. Liljefors T., Krogsgaard-Larsen P., & Madsen U. (Eds.). (2002). Textbook of drug design and discovery. CRC Press.
13. Propst C. L., & Perun T. (1989). Computer-aided drug design: methods and applications. Marcel Dekker, Inc.
14. Reddy M. R., & Erion M. D. (Eds.). (2001). Free energy calculations in rational drug design. Springer.
15. Silverman R. B., & Holladay M. W. (2014). The organic chemistry of drug design and drug action. Academic press.

On-line Courses/MOOCs:

1. Machine Learning & Deep Learning Tutorials (https://github.com/ujjwalkarn/Machine-Learning-Tutorials)
2. Drug Discovery: Coursera: https://www.coursera.org/learn/drug-discovery
3. Drug Development: Coursera: https://www.coursera.org/learn/drug-development

ASSESSMENT:

40% Continuous / Formative Assessment (see PG Regulations in Appendix).
60% End-semester/Summative Assessment: 3 hour written Exam.

Semester III	Course Code: BIN-CC-534	Credits: 3

BIOINFORMATICS LAB III

Programme Specific Outcome relevant to this course:

PSO 15: Basic *in- silico* laboratory skills relevant to CB & B

Course Outcome

On completion of the course, students should be able to:

CO1: Apply NGS techniques
CO2: Hands on experience in protein sequence databases
CO3: Hands on experience in protein structure databases
CO4: Hands on experience in small molecule databases and chemical format conversion
CO5: Understand drawing tools to draw the molecular structure
CO6: Hands on experience in protein structure prediction tools
CO7: Understand ADME prediction tools for molecular properties calculation
CO8: Use Python for developing ML models
CO9: Use Deep learning in Genomics and Biomedicine
CO10: Hands on experience in drug docking software
CO11: Apply R programming in solving typical use-cases in Bioinformatics

Semester III	Course Code: BIN-CC-534	Credits:3

BIOINFORMATICS LAB III
COURSE CONTENT

Module I: Next Generation Sequencing: Whole genome annotation of small genome: Pre-processing, Repeat masking, Gene prediction – Augustus, BUSCO Evaluation - Genes /Genome, Functional annotation – InterProScan, KEGG. RNA-seq data analysis- Tuxedo/New Tuxedo suite; Functional annotation and gene enrichment analysis of RNA-seq data-DAVID/PANTHER

Module II: Databases and tools: Uniprot, Sequence manipulation suite, Expasy, PDB, Protein sequence analysis: Perform pairwise sequence alignment of protein sequences using BLAST and EMBOSS. Protein structure prediction: Swiss modeling, ITASSER. Procheck: Ramachandran plot for structure stability.

Module III: Machine Learning Algorithms: Scikit learn libraries: sklearn.datasets, sklearn.preprocessing, sklearn.model_selection, sklearn. metrics, Implementation of MLPs, Implementation regression on fuel consumption dataset, Implementation of clustering methods and a recommender system, ML model for protein expression /Sequence classification.

Module IV: Deep Learning in Genomics and Medicine: Deep learning for Sequence analysis, protein structure prediction; Classify cytology images for diagnosis; Predict Protein-Protein Interaction in PPI networks

Module V: Docking software: Docking software: Autodock, Discovery studio: Protein small molecule docking, protein protein docking

Module VI: R Programming: RStudio, Objects, Data structures, control structures, functions, data handling, file handling, statistical operations, graphics generation, packages, bioconductor, regular expression.

BIN-CC-534 Bioinformatics Lab III: Module I

Course Outcomes relevant to this Module:

CO1: Apply NGS techniques

Module Outcome:

After Completion of this module, the student should be able to:
M01 Pre-processing of raw data – quality check and adaptor removal
M02 Masking of repeat region in the genome
M03 Predicting genes from whole genome assembly
M04 Evaluating the completeness of Gene prediction/Genome assembly
M05 Predicting the function and pathway mapping of Genes
MO6 Download and QC analysis of RNAseq data
MO7 Mapping, annotation and analysis of RNAseq data
MO8 Differential expression analysis and functional annotation of RNAseq data

The laboratory work will consist of 5-10 experiments.

1. Quality checking and adaptor removal of raw reads 2. Perform repeat masking using RepeatMasker webserver. 3. Identify genes from whole genome assembly using Augustus (ab initio method) 4. Evaluate gene prediction/ genome assembly using BUSCO 5. Use InterProscan and KEGG for predicting the function and pathway mapping of predicted genes. 6. Perform an RNAseq data analysis using Tuxedo/ New Tuxedo suite pipeline 7. Identify the differentially expressed genes from the RNA-seq data analysis performed 8. Use DAVID/ PANTHER for functional annotation and gene enrichment analysis

BIN-CC-534 Bioinformatics Lab III: Module II

Course Outcomes relevant to this Module:

CO2: Hands on experience in protein sequence databases
CO3: Hands on experience in protein structure databases
CO4: Hands on experience in small molecule databases and chemical format conversion
CO5: Understand drawing tools to draw the molecular structure
CO6: Hands on experience in protein structure prediction tools
CO7: Understand ADME prediction tools for molecular properties calculation

Module Outcome:

After Completion of this module, the student should be able to:
MO1 Retrieve amino acid sequence from UniProt and SwissProt
MO2 Calculate the molecular weight, percentage of amino acids composition, frequency of occurrence of amino acids with sequence manipulation suite
MO3 Retrieve the protein structure files from protein data bank
MO4 Analyse the structural hierarchy of protein using SCOP
MO5 Analyse protein structure with visualization tool : Swiss PDB viewer, PyMol to identify the interaction and various structural representation of protein
MO6 Analyse Protein sequence with ExPASy tools
MO7 Analyse the physicochemical features of amino acid using Protparam
MO8 Apply UCSF Chimera for protein analysis and visualization
MO9 Predict protein structure from sequence using I-TASSER
MO8 Retrieve small molecule in various chemical formats from PubChem, Drug bank, Chem spider
MO9 Sketch the chemical structure of small molecules using drawing tools Chemsketch, Chem draw
MO10 Predict and calculate the drug likeness of small molecules using Swiss ADME
MO11 Predict the active site of the protein using CASTp server
MO12 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**

The laboratory work will consist of 5-10 experiments.
1. Retrieve amino acid sequence from databases for sequence analysis: Uniprot 2. Analysis of amino acid sequence using various tools such as Sequence manipulation suite and Expasy 3. Retrieve protein structure from structure database: PDB 4. Calculate the physicochemical properties of protein using Protparam 5. Predict protein structure from sequence by Swiss modeling and I-TASSER. 6. Active site prediction of protein by CAST p server. 7. Retrieve small molecules from Chemical databases: Pubchem, Chem spider, Drug bank. 8. Use of Open Babel for Chemical format conversion 9. Sketch molecules using drawing tools: Chemsketch, Chemdraw. 10. Predict the molecular properties or drug likeness by Swiss ADME

BIN-CC-534-BIOINFORMATICS LAB-III: Module III

Course Outcomes relevant to this Module:

CO8: Use Python for developing ML models

Module Outcome:

After Completion of this module, the student should be able to:
MO1 Build ML models using Scikit-learn
MO2 Do Regression using Scikit-learn
MO3 Do Clustering using Scikit-learn
MO4 Develop Recommender Systems

Module Syllabus: The laboratory work will consist of 10-15 experiments 1. Use of Scikit-learn libraries to build machine learning models. 2. Use Scikit-learn pre-processing library to pre-process the given dataset 3. Implement Feed forward Network 4. Visualise activation functions 5. Computing gradient descent using backpropagation 6. Computing of cross entropy loss and mean square error 7. Fine tuning of Hyper parameters 8. Implement MLP using Scikit-learn 9. Implement Regression using Scikit-learn 10. Implement Clustering using Scikit-learn 11. Implement Recommender System 12. Implement ML model to predict the protein expression/sequence classification

BIN-CC-534-BIOINFORMATICS LAB-III: Module IV

Course Outcomes relevant to this Module:

CO9: Use Deep learning in Genomics and Biomedicine

Module Outcome:

After Completion of this module, the student should be able to:
MO1 Use deep learning for sequence analysis
MO2 Use deep learning for protein structure prediction
MO3 Use deep learning to predict target gene expression
MO4 Classify cytology images for medical diagnosis
MO5 Predict protein-protein interactions in PPI networks

The laboratory work will consist of 5-10 experiments 1. Use of CNN and RNN to predict the non-coding DNA sequence function. 2. Use of deep learning for structure prediction 3. Implement a Deep neural net to predict target gene expression 4. Perform diagnosis with ResNet on the X-ray images.4. Perform Diagnosis with microscopy images 5. Perform RNA-protein binding sites prediction with CNN 6. Perform graph embedding and predict protein protein interactions in PPI network.

BIN-CC-534 Bioinformatics Lab III: Module V

Course Outcomes relevant to this Module:

CO10: Hands on experience in drug docking software

Module Outcome:

After Completion of this module, the student should be able to:
MO1 Perform molecular docking using docking Software: Autodock and Discovery studio
MO2 Analyse and interpret the interaction of small molecules with its targets
MO3 Perform protein-protein docking studies and interpret the result

The laboratory work will consist of 5-10 experiments.

1. Retrieve the structure of target and ligand files in various databases 2. Prepare proteins and ligands for molecular docking 3. To perform molecular docking using various softwares: protein small molecule docking and protein protein docking 4. To calculate the binding energy of the desired pose

BIN-CC-534 Bioinformatics Lab III: Module VI

Course Outcomes relevant to this Module:

CO11: Apply R programming in solving typical use-cases in Bioinformatics

Module Outcome:

After Completion of this module, the student should be able to:
MO1 Write R programs involving the usage of data structures
MO2 Write R programs involving the usage of control structures
MO3 Use the common built in functions in R based on the requirement
MO4 Create user defined functions in R
MO5 Handle files using R scripts and read/write data in different format to/from the files
MO6 Perform basic statistical operations using the built-in functions in R
MO7 Use the basic high level and low-level graphic functions in R
MO8 Install and use R packages, particularly for the analysis of biological data
MO9 Apply regular expressions for the manipulation of biological sequence data

The laboratory work will consist of 5-10 experiments. 1. Use different data structures in R based on the requirement 2. Use different control structures provided by R. 3. Write an R script that uses a built-in function to display (i) the sequence of 10 integers from 1 to 1000 (ii) sequence of integers from 1 to 1000 with an interval 10. 4. Write an R script to prompt the user to enter an integer and then display its square using the print( ) function. 5. Write an R script to reverse a user given numerical vector and to add a new element to the end of the vector. 6. Define function in R to calculate the mean, median, range, minima and maxima, variance, standard deviation, correlation coefficient, covariance of a given set of data separately.7. Write an R program to create a vector of numbers from 1 to 10 and write it to a CSV file. Read the contents in the file and display it. 8. Write R scripts to generate scatterplot, boxplot, barplot, histogram and pie chart for the given set of data. 9. Write R scripts to plot time series data and categorical data. 10. Write an R program to input a given DNA sequence. Read the sequence into an R object and remove the GC content. Display the resultant string and display the substrings in which adenine is present. 11. Write an R script that reads the name of the user and then returns the number of alphabets in it as "Your name contains .... alphabets". 12. Write an R program to find the reverse of a user defined DNA sequence and then display its complement sequence. 13. Perform the analysis of given biological sequence using the given functions in seqinr and stringr package. 15. Apply Bioconductor packages for sequence alignment problems and phylogenetic tree construction.

Semester III	Course Code: BIN-DE-535	Credits: 2

PROGRAMMING IN R

Programme Specific Outcome relevant to this course:

PSO 7: Advanced skill in selected industry-required programming language

---

Course Outcome

On completion of the course, students should be able to:

CO1: Demonstrate understanding of basic features of R language
CO2: Do data and file management in R
CO3: Do basic graphics in R
CO4: Use selected packages in R
CO5: Apply R to selected Bioinformatics problems

Semester III	Course Code: BIN-DE-535	Credits: 2

Programming in R

COURSE CONTENT:

Module I: **Introduction**- Installing R; Features and significances of R; IDEs for R: Introducing Rstudio; **Programming with R:** Objects- Create/list/remove, Basic datatypes, Operators- arithmetic, relational, logical, assignment, Data Structures- vector, matrix, array, factor, list, data frames; Control Statements- if, for, repeat, while; Functions- Built-in functions and user defined functions.

Module II: Data and File Management: Loading data- load( ), data( ), save( ); Import data using Rstudio Interface; Built-in functions for reading data- scan( ), readLines( ), read.table( ), read.csv( ); Built-in functions for writing data- write( ), writeLines( ), write.table( ), write.csv( ); Getting and setting directory- getwd( ) and setwd( ); File handling in R- Opening a file, creating a file, editing a file, renaming a file, removing a file. **Statistics using R-** Basic statistical operations: mean, median, range, minima and maxima, variance, standard deviation, correlation coefficient, covariance.

Module III. Graphics in R- Basic graphics functions- high level functions (plot( ), barplot( ),boxplot( ), hist(), matplot( ), pie( ), ts.plot( )) and low-level functions (points( ), abline( ), segments( ), arrows( ), curve( ), rect( ), polygon( ), text( ), mtext( ), legend( ), axis( ), lines( ), title( )); Interacting with graphics: locator( ), identify() Saving graphical output;

Module IV: Packages in R: CRAN, Installing packages, loading packages, unloading packages, listing packages; seqinr package for analysing biological sequence-Reading sequence data, Writing sequence data, analysing sequence- base composition, GC Content, word count; Bioconductor - overview, features, installing Bioconductor Packages- Biostrings as an example; Regular expression for handling biological sequence data (using stringr package and its functions-pattern matching, substitution, replacement etc.)

Module V: (Flexi module- Only for Internal Assessment. Lecturers may expand and/ or interpret the syllabus to update it or suit the particular cohort): Applications of R in Computational Biology and Bioinformatics: Case studies of bioinformatics and computational biology problems using the packages from Bioconductor: Sequence alignment using Needleman Wunch and Smith Waterman algorithms; Phylogenetic tree construction; Computational gene finding by identifying ORFs; Generating biological sequence using Hidden Markov Model

BIN-DE-535 - PROGRAMMING IN R: Module I

Course Outcomes relevant to this Module:

CO1: Demonstrate understanding of basic features of R language

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Independently install R in a computer system	Ap
MO2 Articulate the features and significances of R programming language	Un
MO3 List the names of popular IDEs for R	Re
MO4 Use RStudio interface for running R programs	Ap
MO5 Create/list/remove objects from R environment	Ap
MO6 Articulate and differentiate the basic data types in R	Un
MO7 Apply operators appropriately for operating the R objects	An
MO8 Articulate, create and use different data structures given a requirement	Ap
MO9 Explain how each control structure regulate the flow of R programs	Un
MO10 Use control structures based on requirement	Ap
MO11 Predict output given scripts with if/for/repeat/while statements	An
MO12 Choose and use built-in functions in R, given a requirement	Ev
MO13 Create/load/call user defined functions in R programs	Ap
MO14 Create R scripts to solve simple problems	Ap
MO15 Predict output given an R programs	An
MO16 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **Introduction**- Installing R; Features and significances of R; IDEs for R: Introducing Rstudio; **Programming with R:** Objects- Create/list/remove, Basic data types, Operators-arithmetic, relational, logical, assignment, Data Structures- vector, matrix, array, factor, list, data frames; Control Statements- if, for, repeat, while; Functions- Built-in functions and user defined functions.

BIN-DE-535 - PROGRAMMING IN R: Module II

Course Outcomes relevant to this Module:

CO2: Do data and file management in R

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Use load( ), data( ) and save( ) functions appropriately in R programs	Ap
MO2 Import dataset using Rstudio interface	Ap
MO3 Choose and use built-in function for reading data in R	Ev
MO4 Choose and use built-in functions for writing data in R	Ev
MO5 Get and set working directory using built-in functions in R	Ap
MO6 Open, create, edit, rename and remove files using built-in functions in R	Ap
MO7 Compute mean, median, range, minima, maxima, variance, standard deviation of a given dataset using built-in functions in R	Ap
MO8 Compute correlation coefficient and covariance of a given dataset using built-in functions in R	Ap
MO9 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: Data and File Management: Loading data- load( ), data( ), save( ); Import data using Rstudio Interface; Built-in functions for reading data- scan( ), readLines( ), read.table( ), read.csv( ); Built-in functions for writing data- write( ), writeLines( ), write.table( ), write.csv( ); Getting and setting directory- getwd( ) and setwd( ); File handling in R- Opening a file, creating a file, editing a file, renaming a file, removing a file. **Statistics using R-** Basic statistical operations: mean, median, range, minima and maxima, variance, standard deviation, correlation coefficient, covariance.

BIN-DE-535 - PROGRAMMING IN R: Module III

Course Outcomes relevant to this Module:

CO3: Do basic graphics in R

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Understand the graphics features in R	Un
MO2 Apply graphics packages for visualization purposes	Ap
MO3 Generate scatterplot, boxplot, barplot, histogram & pie chart using built-in functions in R	Ap
MO4 Plot time series data and categorical data using built-in functions in R	Ap
MO5 Plot columns of matrix using built-in functions in R	Ap
MO6 Use low-level functions in R to enhance visualization	Ap
MO7 Interact with graphics output from R using built-in functions in R	Ap
MO8 Save graphical outputs from R environment	Ap
MO9 Achieve other outcomes dynamically evolved by the teacher during curricular transaction	

Module Syllabus: **Graphics in R-** Basic graphics functions- high level functions (plot( ), barplot( ), boxplot( ), hist(), matplot( ), pie( ), ts.plot( )) and low-level functions (points( ), abline( ), segments( ), arrows( ), curve( ), rect( ), polygon( ), text( ), mtext( ), legend( ), axis( ), lines( ), title( )); Interacting with graphics: locator( ), identify() Saving graphical output

BIN-DE-535 - PROGRAMMING IN R: Module IV

Course Outcomes relevant to this Module:

CO4: Use selected packages in R

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Articulate the objective of CRAN	Un
MO2 Install new packages in R independently	Ap
MO3 List all installed packages in R environment	Re
MO4 Load/unload packages to/from the working environment	Ap
MO5 Apply seqinr package for various analysis of biological sequences	Ap
MO6 Articulate about Bioconductor package, its applications and features	Un
MO7 Install Bioconductor and its packages (eg. Biostrings)	Ap
MO8 Apply regular expressions in R for handling bio-sequence data	Ap
MO9 Use stringr package for handling biological sequences in R	Ap
MO10 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **Packages in R:** CRAN, Installing packages, loading packages, unloading packages, listing packages; seqinr package for analysing biological sequence-Reading sequence data, Writing sequence data, analysing sequence- base composition, GC Content, word count; Bioconductor - overview, features, installing Bioconductor Packages- Biostrings as an example; Regular expression for handling biological sequence data (using stringr package and its functions-pattern matching, substitution, replacement etc.)

BIN-DE-535 - PROGRAMMING IN R: Module V

Course Outcomes relevant to this Module:

CO5: Apply R to selected Bioinformatics problems

Module Outcome:

After Completion of this module, the student should be able to:	TL
MO1 Articulate the applications of R in Computational Biology and Bioinformatics	Un
MO2 Choose and apply Bioconductor packages for sequence alignment problem	Ev
MO3 Choose and apply Bioconductor packages for phylogenetic tree construction	Ev
MO4 Identify ORFs using R Scripts	Ap
MO5 Create R scripts for generating bio-sequences using HMMs	Ap
MO6 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **(Flexi module- Only for Internal Assessment. Lecturers may expand and/ or interpret the syllabus to update it or suit the particular cohort):** Applications of R in Computational Biology and Bioinformatics: Case studies of bioinformatics and computational biology problems using the packages from Bioconductor: Sequence alignment using Needleman Wunch and Smith Waterman algorithms; Phylogenetic tree construction; Computational gene finding by identifying ORFs; Generating biological sequence using Hidden Markov Model.

BIN-DE-535 - PROGRAMMING IN R
ACTIVITIES, LEARNING RESOURCES & ASSESSMENT

SUGGESTED CLASSROOM ACTIVITIES:

Program Portfolio: Students shall submit a portfolio with at least 20 R scripts executed in RStudio along with results.
Complete a MOOC on R Programming
Quiz and assignments

LEARNING RESOURCES:

1. Cotton R. (2015), "Learning R", O'Reilly.
2. Crawley M. J. (2013), "The R Book", Wiley.
3. Adler J. (2010), "R in a nutshell: A desktop quick reference", O'Reilly Media.
4. Gardener M. (2012), "Beginning R: The Statistical Programming Language", Wiley India Pvt. Ltd.
5. Lander J. P. (2015), "R for everyone- Advanced Analytics & Graphics", Pearson Education.
6. Gentleman R. (2008), "R programming for Bioinformatics", CRC Press.
7. Vries A., Meys, J. (2012), "R Programming for Dummies", Wiley India.
8. Mani S. R., Suresh K. T. V., Kasa M., Manvi S. K. S. (2017), "Programming with R", Cengage India.

On-line courses/MOOCs:

1. http://cran.r-project.org/doc/manuals/R-intro.pdf
2. www.cs.ukzn.ac.za/~murrellh/bio/docs/a-little-book-of-r-for-bioinformatics.pdf
3. https://www.datacamp.com/courses/free-introduction-to-r
4. https://www.coursera.org/learn/r-programming
5. http://manuals.bioinformatics.ucr.edu/home/R_BioCondManual
6. https://cran.r-project.org/doc/contrib/Seefeld_StatsRBio.pdf

ASSESSMENT:

40% Continuous / Formative Assessment (see PG Regulations in Appendix).
60% End-semester/Summative Assessment: 3 hour written Exam.

Semester III	**Course Code: BIN-DE-536**	**Credits: 2**

SEMINAR

Programme Specific Outcome relevant to this course:

PSO 12: Knowledge about scientific method and skill in research process
PSO22: Enhanced skills and attitudes for becoming a better learner, thinker, professional and a human being
PSO 17: Awareness of emerging trends and concepts in CB & B

Course Outcome

On completion of the course, students should be able to:

CO1: Experience in technical presentation & science communication

Semester III	Course Code: BIN-DE-536	Credits:2

BIN-DE-536 SEMINAR

Course Outcome (General)

CO1: Experience in technical presentation & science communication

Course Outcome (Specific)

1. Do a professional technical presentation of 10-15-minute duration
2. Develop a multimedia content (PPT/ Open office) to support presentation.
3. Write a 3-5-page seminar report which will confirm in style to a science magazine article, in own words
4. Demonstrate understanding of the topic presented.

Syllabus

Seminar is an opportunity for self-study of a chosen microscopic topic under Computational Biology and Bioinformatics or allied areas and articulating the understanding through a presentation using multimedia support.

Guidelines

1. Each student shall be allotted a supervisor who shall be a full-time teacher and also a mentor, from among PDFs/ Research scholars
2. Student shall propose 3 topics for seminar and one of it shall be finalized in consultation with supervisor, with changes, if any. They may choose a seminal/classic research paper as their basis for presentation.
3. Student shall ensure that the whole seminar report is in own words and plagiarisms shall be checked and self-certified
4. Power point/ multimedia presentation shall not be made by pasting large pictures of text from the report, but shall be freshly made with telegraphic text only. Specific style guide issued by Dept shall be followed.
5. Student shall show evidence of critical self-reviews and peer-reviews of seminar report and presentations.
6. Student shall demonstrate knowledge of the topic of the seminar during presentation by responding to discussions and questions.

Semester III	Course Code: BIN-DE-537	Credits:2

BIN-DE-537 SOFT SKILLS

Programme Specific Outcome relevant to this course:

PSO22: Enhanced skills and attitudes for becoming a better learner, thinker, professional and a human being

Course Outcome

On completion of the course, students should be able to:

CO1: Awareness about various aspects of personality
CO2: Skills and motivation to develop personality.
CO3: Healthy attitude towards communicating in English
CO4: Better communication skills in English.
CO5: Better personal management skills
CO6: Better living skills
CO7: Better Creative and Critical Thinking skills

Course content and other modules same as that of BIN-DE-515

Semester IV

Course Title	Course Code
NGS Data Analytics III	**BIN-CC-541**
Project and Viva Voce	**BIN-CC-542**

Semester IV	Course Code: BIN-CC-541	Credits:3

NGS Data Analytics-III

Programme Specific Outcome relevant to this course:

PSO 8: Thorough knowledge of basic concepts underpinning Bioinformatics & NGS Data
PSO 14: Knowledge and skill in NGS Data Analytics

Course Outcome

On completion of the course, students should be able to:

CO1: Understanding structural and functional annotation of genes
CO2: Understand reference based differential gene expression analysis
CO3: Familiarize with the pipeline and terms used in differential gene expression
CO4: Explore the visual representation of DEGs and perform the biological interpretation of differential gene expression analysis
CO5: Understand the different mutations and human disorders
CO6: Understand different types of variants
CO7: Execute variant calling pipeline
CO8: Understand the relevance of ChIP-seq and when is it used
CO9: Understand the relevance of Hi-C technology and when is it used
CO10: Understand the general analysis steps in ChIP-seq and Hi-C.
CO11: Understand the need for metagenomics technology
CO12: Understand the general analysis steps in metagenomics
CO13: Understand the usage of Galaxy for NGS analysis
CO14: Be able to independently determine the analysis of a given dataset

Semester IV	Course Code: BIN-CC-541	Credits:3

NGS Data Analytics -III

COURSE CONTENT

Module I: Gene prediction & functional annotation: gene structure, exon, intron, ORF, CDS, UTR, alternative splicing, translating DNA sequence, codon usage, difference between prokaryote and eukaryote gene structure, Gene prediction methods – Ab initio, Homology & transcriptome/EST, Non coding RNA annotation – methods – types –tRNA, rRNA, miRNA, and snRNA. Functional annotation - Sequence Homology Based Annotation & Gene Ontology/ Pathway Mapping- Swiss-Prot,, KEGG ENZYME database.

Module II: Differential gene expression: Reference based- Tuxedo/ New Tuxedo suite, differential gene expression – control, treated, replicates & its importance, quantification - count table, normalization – FPKM, RPKM, TPM, upregulated genes, down regulated genes, FDR. Clustering algorithms – PCA, Hierarchical clustering, Visualization of DEGs (heatmap, volcano plot), Biological interpretation – Pathway and GO annotation system, Gene set enrichment analysis, DAVID, PANTHER.

Module III: Clinical genomics: Mutations and human disorders, Type of variants: single nucleotide polymorphism (SNPs), single nucleotide variations (SNVs), smaller insertions or deletions (indels), trinucleotide repeat disorders (TRDs), copy number variations (CNVs) and chromosomal aberrations, Human Gene *Mutation Database. OMIM, Whole genome and whole exome variant calling pipeline – SAMTools, GATK,* SIFT, PolyPhen. File formats – VCF, BED. Application of variant calling – Genetic screening and diagnosis, understanding the cause of rare diseases, cancer.

Module IV: Epigenomics: Introduction to ChIP-seq, Overview of ChIP-seq data analysis pipeline- Read mapping, Read count, Peak calling, Motif calling. Tools - MACS2, HOMER. Transcription factor binding site analysis. Histone binding analysis. Hi-C data analysis.

Module V: Metagenomics: Microbial Diversity and Species Richness, Pure culture and in consortium; Cultivable and Non-cultivable microbial analysis; Amplicon Sequencing, Whole genome Metagenomics, OTU - Binning Strategy, Metagenomics assembly methods, Sequence-based Metagenomics Analysis; Function based Metagenomics Analysis; Phylogenetic analysis and Comparative genomics. Tools – MGRAST, QIIME, metaSPAdes. Application and challenges of metagenomics.

Module VI: (Flexi module- Only for Internal Assessment. Lecturers may expand and/ or interpret the syllabus to update it or suit the particular cohort in any way): Cloud based NGS analysis – Galaxy. Case studies – any one analysis pipeline of choice.

BIN-CC-541- NGS Data Analytics III: Module I

Course Outcomes relevant to this Module:

CO1: Understanding structural and functional annotation of genes

Module Outcome relevant to this Module:

After Completion of this module, the student should be able to:	TL
MO1 Articulate on structural annotation of genes	Un
MO2 Differentiate between prokaryote and eukaryote gene structure	Ev
MO3 Articulate on different types of gene prediction methods	Un
MO4 Explain different types of Non coding RNA	Un
MO5 Articulate on non-coding RNA annotation methods	Un
MO6 Articulate and exemplify functional annotation	Un
MO7**Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **Gene prediction & functional annotation:** gene structure, exon, intron, ORF, CDS, UTR, alternative splicing, translating DNA sequence, codon usage, difference between prokaryote and eukaryote gene structure, Gene prediction methods – Ab initio, Homology & transcriptome/EST, Non coding RNA annotation – methods – types –tRNA, rRNA, miRNA, and snRNA. Functional annotation - Sequence Homology Based Annotation & Gene Ontology/ Pathway Mapping- Swiss-Prot,, KEGG ENZYME database.

BIN-CC-541- NGS Data Analytics III: Module II

Course Outcomes relevant to this Module:

CO2: Understand reference based differential gene expression analysis
CO3: Familiarize with the pipeline and terms used in differential gene expression
CO4: Explore the visual representation of DEGs and perform the biological interpretation of differential gene expression analysis

Module Outcome relevant to this Module:

After Completion of this module, the student should be able to:		TL
MO1	Articulate on differential gene expression	Un
MO2	Discuss the steps involved in a reference based differential gene expression analysis using Tuxedo/New Tuxedo suite pipeline	Un
MO3	Discuss the significance of replicates in the RNA sequencing experiment	Un
MO4	Explain the terms count table, normalization, upregulated & down regulated genes, FDR	Un
MO5	Illustrate the clustering algorithms- PCA and Hierarchical clustering in relation to RNAseq data analysis	Un
MO6	Explain the significance of visual representation of DEGs and infer the differentially expressed genes from the plots generated (volcano plot/heatmap)	Un
MO7	Discuss the importance of biological interpretation of the RNAseq data by GO annotation and pathway analysis	Un
MO8	Perform functional annotation and gene enrichment analysis using DAVID/ PANTHER	Ap
MO9	**Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **Differential gene expression:** Reference based- Tuxedo/ New Tuxedo suite, differential gene expression – control, treated, replicates & its importance, quantification - count table, normalization – FPKM, RPKM, TPM, upregulated genes, down regulated genes, FDR. Clustering algorithms – PCA, Hierarchical clustering, Visualization of DEGs (heatmap, volcano plot), Biological interpretation – Pathway and GO annotation system, Gene set enrichment analysis, DAVID, PANTHER.

BIN-CC-541- NGS Data Analytics III: Module III

Course Outcomes relevant to this Module:

CO5: Understand the different mutations and human disorders
CO6: Understand different types of variants
CO7: Execute variant calling pipeline

Module Outcome relevant to this Module:

After Completion of this module, the student should be able to:		TL
MO1	Articulate the different types of mutations and human disorders	Un
MO2	Articulate the different types of variants like SNP, SNV, TRD, indels, CNV	Un
MO3	Use different variant databases like HGMD, OMIM	Ap
MO4	Execute Whole genome and whole exome variant calling pipeline	Ap
MO5	Use different variant calling tools like SAMTools, GATK	Ap
MO6	Articulate different variant file formats – VCF, BED	Un
MO7	Apply variant calling pipeline for genetic screening and diagnosis, finding rare genetic diseases, cancer etc	Ap
MO8 Achieve other outcomes dynamically evolved by the teacher during curricular transaction		

Module Syllabus: **Clinical genomics:** Mutations and human disorders, Type of variants: single nucleotide polymorphism (SNPs), single nucleotide variations (SNVs), smaller insertions or deletions (indels), trinucleotide repeat disorders (TRDs), copy number variations (CNVs) and chromosomal aberrations, Human Gene *Mutation Database. OMIM, Whole genome and whole exome variant calling pipeline – SAMTools, GATK,* SIFT, PolyPhen. File formats – VCF, BED. Application of variant calling – Genetic screening and diagnosis, understanding the cause of rare diseases, cancer.

BIN-CC-541- NGS Data Analytics III: Module IV

Course Outcomes relevant to this Module:

CO8: Understand the relevance of ChIP-seq and when is it used
CO9: Understand the relevance of Hi-C technology and when is it used
CO10: Understand the general analysis steps in ChIP-seq and Hi-C

Module Outcome relevant to this Module:

After Completion of this module, the student should be able to:		TL
MO1	Articulate on basic concepts and terminology of ChIP-seq technology	Un
MO2	Summarize the general steps involved in the ChIP-seq data analysis pipeline	Un
MO3	Differentiate between transcription factor binding and histone binding analysis from ChIP-seq data	Un
MO4	Summarise the general steps involved in Hi-C data analysis pipeline	Un
MO5	**Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **Epigenomics:** Introduction to ChIP-seq, Overview of ChIP-seq data analysis pipeline- Read mapping, Read count, Peak calling, Motif calling. Tools - MACS2, HOMER. Transcription factor binding site analysis. Histone binding analysis. Hi-C data analysis.

BIN-CC-541- NGS Data Analytics III: Module V

Course Outcomes relevant to this Module:

CO11: Understand the need for metagenomics technology
CO12: Understand the general analysis steps in metagenomics

Module Outcome relevant to this Module:

After Completion of this module, the student should be able to:		TL
MO1	Define metagenomes and articulate its basic concepts and terminologies	Un
MO2	Outline the steps involved in metagenomic data analysis pipeline	Un
MO3	Articulate microbial diversity and differentiate between cultivable and non-cultivable microbes	Un
MO4	Summarize the steps in whole genome metagenomics analysis pipeline	Un
MO5	Differentiate between sequence-based and function-based metagenomics analysis	Un
MO6	Articulate five different application of metagenomics	Un
MO7	**Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: **Metagenomics:** Microbial Diversity and Species Richness, Pure culture and in consortium; Cultivable and Non-cultivable microbial analysis; Amplicon Sequencing, Whole genome Metagenomics, OTU - Binning Strategy, Metagenomics assembly methods, Sequence-based Metagenomics Analysis; Function based Metagenomics Analysis; Phylogenetic analysis and Comparative genomics. Tools – MGRAST, QIIME, metaSPAdes. Application and challenges of metagenomics.

BIN-CC-541- NGS Data Analytics III: Module VI

Course Outcomes relevant to this Module:

CO13: Understand the usage of Galaxy for NGS analysis
CO14: Be able to independently determine the analysis of a given dataset

Module Outcome relevant to this Module:

After Completion of this module, the student should be able to:	TL
MO1 Outline the steps involved in running a NGS data analysis pipeline in Galaxy	Un
MO2 Elucidate the course of analysis for a given NGS dataset	Un
MO3 **Achieve other outcomes dynamically evolved by the teacher during curricular transaction**	

Module Syllabus: (Flexi module- Only for Internal Assessment. Lecturers may expand and/ or interpret the syllabus to update it or suit the particular cohort in any way): Cloud based NGS analysis – Galaxy. Case studies – any one analysis pipeline of choice.

BIN-CC-541- NGS Data Analytics III

ACTIVITIES, LEARNING RESOURCES & ASSESSMENT

SUGGESTED CLASSROOM ACTIVITIES:

Quiz
Seminar presentations on selected topics
Completion of a MOOC course

LEARNING RESOURCES:

References:

1. Brown S. M. (Ed.) (2015), "Next-generation DNA sequencing informatics", New York, NY: Cold Spring Harbor Laboratory Press.
2. Ye S. Q. (Ed.) (2016), "Big data analysis for bioinformatics and biomedical discoveries", CRC Press.
3. Wang X. (2016), "Next-generation sequencing data analysis", CRC Press.
4. Sablok G., Kumar S., Ueno S., Kuo J., & Varotto C. (Eds.) (2015), "Advances in the understanding of biological sciences using next generation sequencing (NGS) approaches", Springer International Publishing.
5. Mandoiu I., & Zelikovsky A. (2016), "Computational methods for next generation sequencing data analysis", John Wiley & Sons.
6. Korpelainen E., Tuimala J., Somervuo P., Huss M., & Wong G. (2014), "RNA-seq data analysis: a practical approach", CRC press.
7. Soh J., Gordon P. M., & Sensen C. W. (2012), "Genome annotation", CRC Press.

On-line Courses/MOOCs:

1. Introduction to Genomic Technologies, https://www.coursera.org/learn/introduction-genomics
2. Whole genome sequencing of bacterial genomes - tools and applications, https://www.coursera.org/learn/wgs-bacteria
3. Command Line Tools for Genomic Data Science, https://www.coursera.org/learn/genomic-tools
4. Big Data, Genes, and Medicine, https://www.coursera.org/learn/data-genes-medicine
5. Bioconductor for Genomic Data Science, https://www.coursera.org/learn/bioconductor

ASSESSMENT:

40% Continuous / Formative Assessment (see PG Regulations in Appendix).
60% End-semester/Summative Assessment: 3 hour written Exam.

Semester IV	Course Code: BIN-CC-542	Credits: 12

PROJECT AND VIVA-VOCE

Programme Specific Outcome relevant to this course:

PSO 12: Knowledge about scientific method and skill in research process
PSO 15: Basic *in- silico* laboratory skills relevant to CB & B
PSO 16: Skill in designing & implementing medium-scale projects in NGS Data Analytics
PSO 17: Awareness of emerging trends and concepts in CB & B

Course Outcome

On completion of the course, students should be able to:

CO1: Experience in solving a real-life problem using bioinformatics concepts, with attempted innovation

CO2: Experience in Professional Project Management and Scientific Reporting

CO3: Experience in Popular Science Communication in mother tongue

Semester IV	Course Code: BIN-CC-541	Credits: 12

PROJECT AND VIVA-VOCE

Course Outcome (General)

CO1: Experience in solving a real-life problem using bioinformatics concepts, with attempted innovation

CO2: Experience in Professional Project Management and Scientific Reporting

CO3: Experience in Popular Science Communication in mother tongue

Specific Outcome

MO1	A dissertation embodying a project work (30-50 pages) as per prescribed format
MO2	Experience of professional work environment in either R & D/ Industry (optional)
MO3	Evidence of compliance with project management system (lab notebooks, progress reports, plagiarism report, peer-review reports)
MO4	Evidence of attempt to publish/patent the solution developed in the project
MO5	(a) Critical analysis of project work and (b) thorough knowledge related to the area of the project (as articulated in a viva voce)
MO6	Oral presentation of the project work (using multi-media support) in a professional manner, (a) in 20 minutes and (b) in 3 minutes
MO7	Summary of the project work in research paper format (3-4 pages)
MO8	Summary of the project in popular science style, in mother tongue

Syllabus

The student shall do an individual project work spanning 90 working days (including reviews, evaluation etc.), preferably in an external R & D or industry organization, the result of which shall be embodied in a dissertation of (30-50 pages) in prescribed format. The project work shall relate to state-of-the-art Computational Biology or Bioinformatics or allied areas and shall involve an attempt to innovate. *In silico* work is intended, but students may add *in vitro*/*in vivo* components. Project management shall be done professionally as per prescribed guidelines issued by the Dept. and shall include project planning in the previous semester. The student shall demonstrate technical/scientific writing skills and critical mind in compiling the dissertation and articulating the same. The student shall demonstrate professional presentation skills in presenting the work in a viva-voce. They shall also demonstrate overall knowledge related to the area of the project during the viva-voce. They shall summarize the work in a research paper format and produce the same along with dissertation. A summary in popular science format in mother tongue shall also be submitted.

Guidelines

1. Projects should be selected at least 6 months prior to commencement. Planning Docket (Appendix A) should be used for this.

2. Students are expected to spend a minimum of clear 8 hours per day, ideally 10-12 hours on the project work. This is sure to reflect on the quality and quantity of work.
3. Students are encouraged to do their project in an external organization, to expose themselves to professional R&D work culture.
4. Students shall submit 3 Project Progress Reporting dockets (Appendix B) along with Work Reports (around 5 pages) as indicated in the weekly planner.
5. Students shall maintain Lab Notebooks, with brief notes for each day.
6. Thesis writing should be done in a distributed manner and not in haste after finishing work.
7. Collection of reference materials used should be produced during viva/discussion with supervisors, if required.
8. Follow Week based Planner (Appendix C)
9. For student Peer Review, Interim Review, Self-Evaluation Report, Final External Evaluation reports, use form in Appendix D, E, F, G (i) & G (ii) respectively.
10. Project Report shall conform to Green Charter of the University of Kerala.
11. Summary of the project, prepared in research paper format (3-4 pages) shall be submitted along with the project report.
12. A popular-science-style article based on the project, written in mother tongue shall be submitted along with the project report.

Appendix A: PROJECT PLANNING DOCKET
DEPARTMENT OF COMPUTATIONAL BIOLOGY & BIOINFORMATICS
UNIVERSITY OF KERALA
(To be filled in prior to project selection)

Name of the Student	
1. What broad area would you like to work? (put a tick mark)	1. Programming based 2. Including theory/ Mathematics 3. Bioinformatics including Wetlab 4.
2. List five areas of your Interest add (a mind map for each the topic as appendix)	1. 2. 3. 4. 5.
4. What are the general objectives of a Project work?	

5. List 5 skill sets you have in relation to your project work?

1.

2.

3.

4.

5.

6. List three external Institution and guides in consideration (If more than one person in an institution add additional sheets) *

* Those who are constrained to do the project in the department may submit a detailed justification for the same.

Full address	
PIN	
Email	
Phone	
Web	

Name of Guide	
His/Her current area of work	
Summary of their work that interests you(attach as appendix if lengthy)	
Skill sets required	
Constraints*	

Full address	
PIN	
Email	
Phone	
Web	

Name of Guide	
His/Her current area of work	
Summary of their work that interests you(attach as appendix if lengthy)	
Skill sets required	
Constraints*	

Full address	
PIN	
Email	
Phone	
Web	

Name of Guide	
His/Her current area of work	
Summary of their work that interests you(attach as appendix if lengthy)	
Skill sets required	
Constraints*	

Does the institution have date/ skill/ fee or other constraints? Briefly note

7. In case of institutions other than DCB, please mention logistics		
a. Arrangements for stay		
b. Living Expenditure and/or Fees (affordable?)		
c. Support from family		
d. Mention 3 Key advantages of doing project in the concerned institution and three challenges that you envisage		
8. Do you plan for a publication? Mention your dream in this regard(title of the paper, one target journal)		
9. How many hours of work /day are you prepared to put in during the project? How will you raise so many hours? What changes will you make in your current routine		
10. Have you read the guidelines for writing project proposals? What are the key observations that you made?		
11. Have you read any past M. Phil / MSc. thesis ? Give a brief summary of it and add 3 critical observations		
12. List key resource persons you would like to consult regarding your project		
13. Your choice of (i)Internal supervisor and (ii)Assistant supervisor (ii shall be Post doc/Project fellow/ Research Scholar)	Internal supervisor	
	Assistant supervisor	
FOR OFFICE USE: Remarks by Project advisory Committee		
External Guide / Supervisor		
Internal Guide / Supervisor		
Proposed Assistant Guide / Supervisor		
Dated signature of HOD		

Appendix B: PROJECT PROGRESS REPORTING DOCKET
DEPARTMENT OF COMPUTATIONAL BIOLOGY AND BIOINFORMATICS
UNIVERSITY OF KERALA

Regular documentation is required to ensure that meaningful monitoring and management of project work is carried out. This docket is to be submitted thrice during the project period as indicated in **Weekly Planner**. You are also encouraged to record key discussions and minute them as appendix.

Docket Submission Details

Name of the Student	
Serial No of Submission	
Date of Submission	
Dated Signature of Candidate	

Mark the Start Date & Estimated Project Submission Dates below (Put tick mark on column)

YEAR											

1	Name of Student:
2	Name(s) of Guide(s) :
3	No. of weeks elapsed after beginning of the project:
4	List 2 keywords to describe the project:
5	List a couple of papers published in journals which directly relate to the area of proposed work. 1. 2.
6	List 2 areas of basic knowledge that are directly related to the project work. 1. 2.
7	List 2 tools/ skills required to do the project 1. 2.

8 List 2 prominent researchers in the field who are currently active in the area.

1.

2.

9 List 2 journals in which papers of this area are appearing

1.

2.

List some books which deal with the basics of your project area.

1.

2.

10 List 2 upcoming conferences you would like to attend.

1.

2.

11 List 2 strengths you have to do this project.

1.

2.

12 List 2 areas of weakness in doing this project (*also say how you propose to overcome them*).

1.

2.

13 List a possible title of paper that you could publish on your project.

14 List 2 possible titles of your thesis.
1.

2.

15 List the possible titles of your thesis chapters.

1.	4.
2.	5.
3.	6.

16 List 2 deliverables of your project.
1.

	2.
17	Write in one sentence the up to date Synopsis of your project.
18	Write in one paragraph (4-5 sentences), the up to date Synopsis of your project.
19	Attach hand out / print out of the following PPTs (Appendix). Indicate version and date in the first page. 1. 1 slide PPT describing your project (This should capture the essence of the work in compact graphics – a sample is available on request) 2. 10 slide PPT describing your project.
20	What are your major achievements in the reporting period?
21	What are your immediate milestones and estimated dates to reach them?
22	One important journal paper read in the last month with a 2-3 sentence comment in your own words and also mentioning the open questions identified.
23	Lectures / Conferences / Training attended with 2-3 sentence summary/outcome.
24	Write an appreciation of your work during last month.

25 Write a criticism of your work during last month.

26 Describe your recent library usage. Mention the books you have read in the following areas

Subject Books	General Books	Online Resources

27 Is there any query you would like to make to your guide/ supervisor?

28 What are some of the issues (*intellectual as well as otherwise*) you face in project?
(*Also mention personal circumstances which prevent full output.*)

Rate your motivation level now: poor Average Good Very Good Excellent

STAR RATING OF PROGRESS (BEST IS 10 STARS * * * * * * * * * *)	
Your Own Rating of Progress since the last report	
Guide's Rating of Progress since the last report	
Your own Rating of **Total Progress**	
Guide's Rating of **Total Progress**	

Signature of Student Signature
of Guide

Official Data:

Item	**Data**
Days and dates of absence during reporting period	
Have you done bi weekly CD backup?	
Are you keeping a lab notebook?	

Appendix C:
General Guidelines

1. Students should maintain Lab Notebooks; with one-page brief report for each day. For example, if you spend time in Library, you should mention journals / books read. On web, note down URLs, while doing experiments note settings and intermediate results. Lab notebooks shall form a component for evaluation and shall be presented to the external examiner, if required.
2. Students are expected to spend a minimum of clean (that is, excluding phone calls, Facebook, tea/ lunch break etc) 8 hours per day, ideally 10-12 hours. This is sure to reflect on the quality and quantity of work.
3. Students should photocopy/ print important reference material and file them for constant reference and make notes on them as they read. They are expected to read key references many times over (10 or even 20 times is not unexpected). The collection of reference materials should be produced during viva / discussion with supervisors, if required.
4. Students shall submit 3 Project docket reports along with Work Reports (about 5 pages) as indicated in the weekly planner. Along with it, draft chapters as indicated should also be submitted.
5. Thesis writing is to be done in a distributed manner and not in haste after finishing work. To obtain quality feedback, submit draft chapters regularly as indicated in the week plan.
6. Copying text from net or from any other resources amounts to plagiarism. Understand the concept and write it in your own words, citing references.
7. 'Cambria'/ any similar font is recommended for typesetting. For text, font size should be 12 and for main headings font size can be 14 (bold) and for sub headings 12 (bold). Chapter headings can be in 18/20 points (bold).
8. Page margins recommended are: top: 2 cms, bottom: 2 cms, left: 3 cms, right: 2cms
9. Line spacing shall be 1.5 and paragraph spacing shall be 6 pt, before and after.
10. All figures/ pictures/ graphs shall be labeled in the following way. Label must start with 'Fig.' followed by 'Serial no.' For example, for the second figure in chapter 4 is labelled 'Fig. 4.2: Description of the figure' in point size 11, italics. Tables are also to be labeled likewise.
11. Page number should appear in the bottom centre position of every page, beginning with chapter 1 (previous title/content pages etc. may be numbered in roman numerals as i, ii, iii etc.
12. Project Report should conform to the Green Charter of the University of Kerala.

Appendix C:
PPT Guidelines

1. Open Viva-voice of dissertation requires a crisp and professional multi-media presentation.
2. Presentation is not a text to be read out. It is a prompt for the speaker to speak and a structure to help listeners to listen.
3. The PPT needs to be well planned to make the presentation in 15-20 minutes. Slides should be limited to 20-30. Students are expected to organize mock presentations and ensure length of presentation within limits.
4. The intention is not to convey ideas simply as in a lecture class, but to impress on the quantum and quality of work in a scholarly manner.
5. General introductory material should be limited to 10-15%.
6. Cite references in the slide where required and use graphics as much as you can.

Appendix C
WEEKLY PLANNER FOR PROJECT

Week	Date	Documentation Activity of Student	Activity of Supervisor/ Assistant supervisor	Remarks
Week 1				
Week 2				
Week 3				
Week 4		**Submit Project docket 1 + Draft of chapter 1 and 2+ work report 1**		
Week 5			**Issues a project review report**	
Week 6		**Revise chapter 1 & 2**		
Week 7				
Week 8		**Submit Project docket 2 + Partial Draft of chapter 3 and 4 + work report 2**		
Week 9			**Return drafts with detailed review**	
Week 10		**Revise chapter 3 & 4**		
Week 11				
Week 12		**Submit project docket 3+ dream Synopsis, partial result + work report 3**		
Week 13				
Week 14		**Submit 90 % complete draft of thesis* and Synopsis final* Pre submission docket +peer review**	**Send out final project checklist**	
Week 15		**Submission of Thesis**		
Week 16		**Preparation of OUP format paper & PPT & peer review**	**Review of paper & PPT**	
Report back to Department				

Appendix D

DEPARTMENT OF COMPUTATIONAL BIOLOGY AND BIOINFORMATICS
UNIVERSITY OF KERALA
PART-D: FORM FOR STUDENT PEER REVIEW OF PROJECT

1. Your Name:	
2. Title of project you are reviewing:	
3. Name of Project student:	
4. Give five positive aspects of the project. 1. 2. 3. 4. 5.	
5. Give 5 aspects that need improvement 1. 2. 3. 4. 5.	
6. Do you find the quantum of work comparab your own?	YES/NO
7. Do you find the quantum of work comparab your own?	YES/NO
8. Do you find the quantum of references compared to your own?	YES/NO
9. What are your suggestions for enhancing the quality of work?	
10. Are there books/web resources/journal papers/persons that you would like to suggest to the project student for drawing knowledge?	

Date:

Signature:

Appendix E

DEPARTMENT OF COMPUTATIONAL BIOLOGY AND BIOINFORMATICS
UNIVERSITY OF KERALA
PART-E: INTERIM REVIEW REPORT BY INTERNAL SUPERVISOR

Name: Date:
Title of Project:

Quick Remarks	
Satisfactory Progress?	Yes/No
Technical correctness of methodology?	Yes/No
Sufficient use of knowledge resources (References)?	Yes/No
Good library usage?	Yes/No
Good work habits?	Yes/No
Satisfactory Keeping of lab Notebook?	Yes/No
Student Peer Review and use?	Yes/No
Satisfactory Attendance?	Yes/No
Remarks on Technical Progress of Work:	
General Advices on bettering project work:	
Specific reply to questions, if any, raised by student in Progress Report	

Any '**To-Do**' before Next Review:

Counter signed by Internal Supervisor Signed by Asst.
Supervisor

Head of the Department

Appendix F: PROJECT SELF EVALUATION REPORT
DEPARTMENT OF COMPUTATIONAL BIOLOGY AND BIOINFORMATICS
UNIVERSITY OF KERALA
(DATA TO BE FILLED-IN BY STUDENT)

1. Name of student:	
2. Register Number:	
3. Title of the Thesis:	
4. Supervisors: Internal: External: Assistant:	
5.Total no. of pages in Dissertation	
6. No of pages of the final chapter (Conclusions and future work):	
7. Total no of references: 8. No. of journal cited:	
9. Have you taken peer reviews from other students? 10. Have they been used:	YES/NO YES/NO
11. Give three key achievements in your project work: 1. 2. 3.	
12. Give three weak aspects of project work and dissertation.	
13. If you were asked to say in one sentence, the abstract of your project, what would you say?	
14. What were the unanswered questions you identified at the end of the project?	
15. Have you attended any seminars/references? Give details.	
16. Have you ensured that every sentence in your dissertation is your own?	**YES/NO**
17. Have you cited any long paragraph as such with/without giving references?	**YES/NO**

18. Are all images, tables, etc. your own or sources cited clearly?	**YES/NO**
19. Have you spell checked the whole dissertation.	**YES/NO**
20. Have you verified the format of the dissertation based on instructions?	**YES/NO**
21. Have you separately checked the punctuation formatting, including the reference section?	**YES/NO**

Dated Signature of Student

Enclosures to be provided by students

1. *One-page synopsis including key references* ***2.*** *Lab Notebook.* ***3.*** *Three project progress reports and corresponding assistant supervisors review report* ***4.*** *Peer review reports by fellow students*

Appendix G

DEPARTMENT OF COMPUTATIONAL BIOLOGY AND BIOINFORMATICS
UNIVERSITY OF KERALA
PART-G: M.Sc. FINAL EVALUATION REPORT FOR BIN-CC-541
(2-6 TO BE FILLED IN BY BOTH EXAMINERS AND AVERAGE AWARDED)

Name	
Register Number	
EVALUATION COMPONENTS & **SUB COMPONENTS** (Weightage out of 100 in brackets)	**GRADE** A-Outstanding, B-Excellent, C-Good, D-Average, E-Below, F-Fail.
1. PROJECT MANAGEMENT (10)(To be given by internal Supervisor)	
Lab Notebook is regular & detailed (2.5)	**A/B/C/D/E/F**
Detailed & Precise Progress Reports (2.5)	**A/B/C/D/E/F**
Regular Peer Review/Supervisor review & Action (5)	**A/B/C/D/E/F**
2. PROJECT REPORTING (10)	
Scientific Reporting Standards, Formatting (5)	**A/B/C/D/E/F**
Citing Practice, Avoidance of Plagiarism (5)	**A/B/C/D/E/F**
3. TECHNICAL WORK (35)	
Quantum of work (10)	**A/B/C/D/E/F**
Meetings Objectives (10)	**A/B/C/D/E/F**
Demonstration of Results (15)	**A/B/C/D/E/F**
4. SCHOLARSHIP (15)	
Demonstration of Critical Analysis (5)	**A/B/C/D/E/F**
Contextualizing the work (5)	**A/B/C/D/E/F**
Knowledge as demonstrated in Viva (5)	**A/B/C/D/E/F**
5. COMMUNICATION SKILLS (10)	
Presentation Skills in Viva (5)	**A/B/C/D/E/F**
Use of language in Dissertation (5)	**A/B/C/D/E/F**
6. Publishing (15)	
Summary of Project in Research paper format (5)	**A/B/C/D/E/F**
Publication/Attempt for Publication in journals/conferences approved by the Department council (10)	**A/B/C/D/E/F**
7.Weightage for project done in external organization (5)	
Total Marks (out of 100*)	

Internal Examiner External Examiner

---

*Converting Grades to Marks

MAX	A	B	C	D	E	F
2.5	2.5	2	1.5	1	0.5	0
5	5	4	3	2	1	0
10	10	8	7	5	3	0
15	15	12	9	6	4	0

Appendix

Assessment and Evaluation

(Extracts from CSS Regulations)

6.1 Grading of Student Performance

Performance of students in individual courses shall be evaluated and assigned grades to indicate level of achievement of objectives. The grading scale shall be the same as the national pattern recommended by the UGC. Each grade shall be indicated by a letter as in the table below:

PG Programmes in University Depts. use 10-point scale recommended by UGC as follows

Grade	Grade Point
O (Outstanding)	10
A+ (Excellent)	9
A (Very good)	8
B+ (Good)	7
B (Above average)	6
C (Average)	5
D (Pass)	4
F (Fail)	0
Ab (Absent)	0

The 9 point grading scale for course, semester and programme grading

Each grade shall have a corresponding grade point which serves as a means of aggregating letter grades and are not marks or scores.

6.2 Types of Assessment

In the PG programmes, every taught course shall be assessed through: *(a) continuous (in-sem) assessment with 40% weightage (b) end-sem assessment with 60% weightage.*

6.3 Continuous (In-Sem) Assessment for Taught Courses

Continuous assessment is aimed at giving continuous feedback about teaching and learning to teachers and students respectively. Out of 40% weightage allotted for continuous assessment, the sub components shall normally be:

Component -1: Weightage 20%
Component -2: Weightage 10%
Component -3: Weightage 10%

20% shall be a written test. The other components designed by the course teacher are to be announced in advance as per provision 5.5. Component 2 may typically be an assignment involving a piece of scholarly critical work with critical reflection. Component 3 may typically be group discussion, fieldwork, presentations, case studies, etc.

6.3.1 Conduct of Assessment

Tests and other class assessments shall normally be conducted as per academic calendar and shall always be notified to students well in advance. Additional chance to take tests/examination or to submit assignments under extenuating circumstances or to take part in arts/sports events recognized by the University, shall be duly considered by the teacher.

6.3.2 Abnormal Test Results

Any test for which class average score is less than 40% shall not be reckoned for assessment and a re-test shall be administered for the whole class.

6.3.3 Articulating Grade Requirements & Late Submission Policies

For assignments/term papers/tasks, students shall be provided with clear indication of the effort required to achieve specific grades or score bands. Late submission policy for assignments shall be clearly stated. For each day of delay, not more than 10% of the maximum score may be deducted.

6.3.4 Receipts for Assignments

Students shall be issued with receipts for assignment if they so desire. Time stamped electronic communications shall also be considered as proof of submission.

6.3.5 Marking of Scripts

Teachers may offer written remarks on tests and assignments while returning marked scripts to students so as to enable them to understand the areas for improvement. This shall be mandatory for scores below 40%.

6.3.6 Plagiarism

Awareness shall be given by each course teacher to students about plagiarism and its unacceptability. Students may be encouraged to check plagiarism with software tool recommended by the UGC from time to time for assignments, term papers, and dissertations.

6.3.7 Communicating Results & Addressing Grievances

Each teacher shall formally communicate the score awarded for tests and assignments to every student. Student grievances, if any, shall be first addressed by the teacher. If not resolved, the same may be raised in the DC through the head of the department. Any appeal on the decision of the DC shall lie with student grievance subcommittee of CSSAC which shall give due hearing to the student(s) and teacher(s) and attempt to resolve the grievance, failing which the Syndicate of the University shall take a final decision in the matter.

6.4 End-Semester Assessment for Taught Courses

6.4.1 Nature of End-sem Exam

Every taught course shall be assessed through a written end-sem exam of maximum 3 hours duration (3 hours for 3 & 4 credits and 2 hours for 2 credits). The end-sem exams shall be of summative nature and aimed at testing the achievement of course objectives as stated in the syllabus. Every student shall be issued, in the first week of the semester, a model question paper for the end-sem examinations. The end-sem exam shall normally consist of at least 20% of questions which involve higher order critical thinking. End-sem examination shall not constrain itself to knowledge recall.

6.4.2 Conduct of Exams for Even Semesters

Question papers of end-sem examination shall be set by the teacher(s) teaching the course. The teacher shall set two sets of question papers and it shall be peer-reviewed by the DC. The Head of the Department shall select one of the QPs and arrange to conduct the exam with the same. Answer scripts shall be valued by the teacher who taught the course and one external examiner from a panel of examiners recommended by the DC and approved by the VC. If the grades awarded differ by more than one level, then a third examiner from the earlier panel shall be asked to re-mark and the average of all three scores shall be awarded to the candidate. Grievances if any shall be addressed as per provisions 6.3.7.

6.4.3 Conduct of Exams for Odd Semesters

Question papers of odd-sem examination shall be set by panel of examiners recommended by the DC and approved by the VC, based on the model question paper set by the teacher teaching the course. The Office of the CSSAC shall arrange to set the QP and conduct the exam. Answer scripts shall be marked by the teacher who taught the course and one external examiner from a panel of examiners recommended by the DC and approved by the VC. If the grades awarded differ by more than one level, then a third examiner from the earlier panel shall be asked to re-mark and the average of all three scores shall be awarded to the candidate. Grievances if any shall be addressed as per provisions of 6.3.7.

6.4.4 Grace Marks

Students who participate in any event listed in schedule III shall be eligible for grace marks of 15 marks per academic year and 30 marks for the whole programme. Grace mark shall be added to the total score achieved in any one or more courses of choice of the student, in the whole programme.

6.5 Assessment of Non-Taught Courses

6.5.1 Seminar/Field Work/Case Studies/Internship/Media Production etc

The assessment methodology and time lines shall be clearly defined by the teacher through the detailed course brochure issued to students. Evaluation rubric shall be proposed by the teacher and approved by the DC, and communicated clearly to the students.

6.5.2 Projects

Detailed guidelines on project objectives and detailed evaluation rubric for the same shall be approved by the DC and communicated effectively to the students. Project Management System shall be developed by each Department to enable effective monitoring and support, for projects with credits above 4. Evaluations rubric of projects with credit above 4 shall include innovation as one component.

6.5.3 Laboratory Courses

Detailed guidelines on conduct and assessment of laboratory courses and detailed evaluation rubric for the same shall be approved by the DC and communicated effectively to the students.