

# **POST GRADUATE (P.G.) CERTIFICATE COURSE IN BIOINFORMATICS**

(Syllabus for the session 2014-15 onwards)

**Duration:** 1 Year (Weekend) - 10 credits

**Intake:** 20 Seats

**Course fee:** Rs. 10,000/-

## **About Bioinformatics**

Bioinformatics at the interface between two of the most influential scientific fields i.e. an appreciation of computational and biological sciences, in particular, the terminology employed in both field and is essential for those working at such an interface. These courses aim to cover the commercial and academic perspectives on bioinformatics, the impact of bioinformatics on the methodologies used in biological sciences, complex and extensive public and private biological information sources to anticipate a deep change in methods for finding knowledge in the life sciences and the courses are designed to provide trained and skilled professionals that are required in this field and in the industry.

## **Career prospects**

The course gives an added value to the professionals of biological sciences research both academic level and at industry level. The jobs are available at scientist level in pharma industries, vaccine development, clinical research projects, academic research, etc.

## **Eligibility**

Any of the following (minimum 50% marks required in qualifying examination):

MBBS / B.D.S. / BAMS / B.H.M.S / B.Pharmacy / B.Tech-Biotech / BVSC / B.Sc. (Nursing) / M.Sc./ M.Pharm/ B.Sc (with minimum one year relevant industry/ academic research work experience).

Note: Any candidate pursuing post-graduate course in M.D. University, Rohtak, may also co-opt for this course along with their regular course.

## **Course Duration**

Post Graduate Certificate course in Bioinformatics will be covered within six months beginning with a foundation in bioinformatics and computational biology, followed by modules on structural bioinformatics and chemoinformatics. Medium of instruction will be English. The course is designed for postgraduates and those working in industry / academics and is scheduled on weekends only.

## **Teaching Methodology**

- Theory and practical lectures will be conducted over weekends at the venue by faculty approved by Centre for Bioinformatics, M.D. University, Rohtak.
- Projects: Each student will select the topic for his/her project within eight weeks of joining the course. Thus the projects can be taken by a group of students (not more than 5). Students will send their proposal to the Course Coordinator, who will help to nominate the guides. The project should involve about 10 days of research work.
- Note: This being a post graduate certificate course the students are expected to gain knowledge through interaction with the faculty, use of library /internet and self-study of course material.

## **Evaluation and Examination**

Students will be evaluated through periodic internal examination, evaluation of assignment and project work. At the end of the course students will appear for final examination conducted by Centre for Bioinformatics, M.D. University, Rohtak.

# **POST GRADUATE (P.G.) CERTIFICATE COURSE IN BIOINFORMATICS**

(Syllabus for the session 2014-15 onwards)

## **Program Specific Outcomes**

The students completing this Certificate course will be able to:

- PSO1** Serve the global scientific society with tools and techniques of Bioinformatics.
- PSO2** Work as skilled manpower in the field of Bioinformatics and IT.
- PSO3** Add the knowledge of Bioinformatics to their attained educational skills.
- PSO4** Become a part of Mission-skill India by disseminating the knowledge of Bioinformatics to next generation.
- PSO5** cover the commercial and academic perspectives on bioinformatics.

The students will be awarded as per absolute grading system, detailed below:

<b>Interval of Marks</b>	<b>Grade</b>	<b>Grade Points</b>
> or = 80 but <or =100	O (Outstanding)	10
> or = 70 but <80	A+(Excellent)	9
> or = 65 but <70	A (Very Good)	8
> or = 55 but <60	B+ (Good)	7
> or = 50 but <55	B (Above Average)	6
> or = 45 but <50	C (Average)	5
> or = 40 but <45	P (Pass)	4
Less than 40	F (Fail)	0
	Ab (Absent)	0

*Note: A student obtaining Grade F shall be considered failed and will be required to reappear in the examination.*

## **Course syllabus outline & Scheme of examination**

<b>S.No.</b>	<b>Paper code</b>	<b>Paper title</b>	<b>Internal evaluation</b>	<b>End term evaluation</b>	<b>Credits</b>	<b>Total marks</b>
1.	CBI 101	Basics of bioinformatics	10	40	2	50
2	CBI 102	Genomics & Proteomics	10	40	2	50
3	CBI 103	Structural bioinformatics	10	40	2	50
4	CBI 104	Biostatistics	10	40	2	50
5	CBI 105	Lab-course	-	50	-	50
6	CBI 106	Project work	-	50	2	50
<b>Total</b>			<b>40</b>	<b>260</b>	<b>10</b>	<b>300</b>

# **POST GRADUATE (P.G.) CERTIFICATE COURSE IN BIOINFORMATICS**

(Syllabus for the session 2014-15 onwards)

**Time: 3 Hours**

**Students completing this course will be able to:**

- CO1** Access the world of Bioinformatics, different types of biological data and databases.
- CO2** Access and explain the tools and techniques of analyzing DNA and protein sequences.
- CO3** Discuss local and global alignment (Smith Waterman algorithms, Needleman Wunsch algorithm).
- CO4** Utilize the different information retrieval system and sequence alignment tools (BLAST, FASTA, TCOFFEE).

**Note: Nine questions are to be set in all and the candidates are required to attempt five questions including compulsory question. Question no. I is compulsory covering the whole syllabus. Out of remaining eight questions, two questions are to be set from each unit. Candidate is required to attempt four questions, selecting one question from each unit.**

## **UNIT I**

### **Overview of Bioinformatics and Information Technology**

History, Scope and application, Internet and World Wide Web; Generation of computers; Concept of Networking; Introduction to Data Mining; Application of data mining in Bioinformatics.

## **UNIT II**

### **Bioinformatics Resources**

Biological databases, Basic classification – Sequence & Structure; Generalized & Specialized; Primary & Secondary, with example databases (Genbank, EMBL, DDBJ, INSDC, Swiss Prot, PIR, PDB, NDB, BLOCKS, Pfam, ProSITE, etc.); Literature databases.

## **UNIT III**

### **Bioinformatics techniques**

Sequence comparison and alignment; Local and Global Alignment – (Smith Waterman Algorithm; Needleman Wunsch Algorithm); Concept of Gap, Gap Penalty & Scoring Matrices (PAM, BLOSSUM); Dot Plot Analysis.

## **UNIT IV**

### **Bioinformatics tools**

Information retrieval system (Entrez, SRS); Sequence alignment tools (BLAST, FASTA, CLUSTAL-W/X, MUSCLE, TCOFFEE), Variants of BLAST (BLASTn, BLASTp, PSI-BLAST, PHI-BLAST, etc)

# **POST GRADUATE (P.G.) CERTIFICATE COURSE IN BIOINFORMATICS**

**(Syllabus for the session 2014-15 onwards)**

## **Suggested Readings:**

1. Essential Bioinformatics by Jin Xiong.
2. Bioinformatics : Sequence and Genome Analysis by David W. Mount.
3. Bioinformatics: Concepts, Skills and applications by R. C. Rastogi.
4. Introduction to Bioinformatics by Singhal and Singhal.
5. Bioinformatics: A Practical Guide to analysis of Genes and Genomes by Andreas D. Baxeavanis and b. F. Francis Ouellette.
6. Bioinformatics Basics by Hooman H. Rashidi, Lukas K Buehler.
7. Bioinformatics: A modern approach by Vittal R. Srinivas.
8. Bioinformatics: Database and Systems by Stanley Latovsky.

# **POST GRADUATE (P.G.) CERTIFICATE COURSE IN BIOINFORMATICS**

(Syllabus for the session 2014-15 onwards)

**CBI 102: Genomics and Proteomics MM: Th 80 + IA:20**

**Time: 3 Hours**

**Students completing this course will be able to:**

- CO1** Explain the concept of genomes, genomics and their networks (in context to HGP).
- CO2** Discuss about comparative genomics and metagenomics.
- CO3** Apply learnt proteomics technique for structure prediction, peptide fingerprinting and protein protein interactions.
- CO4** Define the concept of genomics and structural proteomics, peptide microarray- based technology and proteomics in drug discovery and toxicology.

**Note: Nine questions are to be set in all and the candidates are required to attempt five questions including compulsory question. Question no. I is compulsory covering the whole syllabus. Out of remaining eight questions, two questions are to be set from each unit. Candidate is required to attempt four questions, selecting one question from each unit.**

## **UNIT I**

### **Introduction to Genomics**

Introduction to the concept of genome, gene networks: basic concepts, Prediction of genes, promoters, splice sites, regulatory regions: basic principles, application of methods genome projects. Tools for genome analysis-RFLP, DNA fingerprinting, RAPD, PCR, Linkage and Pedigree analysis-physical and genetic mapping.

## **UNIT II**

### **Genome sequencing projects**

Microbes, plants and animals; Accessing and retrieving genome project information from web; Comparative genomics, Identification and classification using molecular markers-16S rRNA typing/sequencing, ESTs and SNPs. Human Genome Project. Large scale genome sequencing strategies, Genome assembly and annotation. Genome databases of Plants, animals and pathogens, Metagenomics: Concept and applications.

## **UNIT III**

### **Proteomics**

Concept of proteome, Protein analysis (includes measurement of concentration, amino-acid composition, N-terminal sequencing); 2-D electrophoresis of proteins; Microscale solution isoelectricfocusing; Peptide fingerprinting; LC/MS-MS for identification of proteins and modified proteins; MALDI-TOF; SAGE and Differential display proteomics, Protein-protein interactions, Yeast two hybrid system.

# **POST GRADUATE (P.G.) CERTIFICATE COURSE IN BIOINFORMATICS**

**(Syllabus for the session 2014-15 onwards)**

## **UNIT IV**

### **Functional genomics and proteomics**

Analysis of microarray data; Protein and peptide microarray-based technology; PCR-directed protein in situ arrays; Structural proteomics. Deriving function from sequence, Proteomics in drug discovery and toxicology.

#### **Suggested Readings:**

- 1 Bioinformatics: sequence and Genome Analysis by David Mount, second Edition. Cold Spring
- 2 Discovering Genomics, Proteomics and Bioinformatics, 2/E by A. Malcolm Campbell and Laurie J. Heyer, Publisher: Benjamin Cummings (2007)
- 3 Principles of Gene Manipulation and Genomics by S.B. Primrose and R.M. Twyman
- 4 Genomes by T.A. Brown
- 5 Proteomics: From Protein Sequence to Function by S. R. Pennington, Michael J. Dun

# **POST GRADUATE (P.G.) CERTIFICATE COURSE IN BIOINFORMATICS**

(Syllabus for the session 2014-15 onwards)

**CBI 103: Structural Bioinformatics**

**MM: Th 80 + IA:20**

**Time: 3 Hours**

Students completing this course will be able to:

- CO1 Account for the structure of proteins, DNA and RNA.
- CO2 Explain basic principles of experimental methods for the determination of the structure of macromolecules.
- CO3 Use computer programs to visualise three-dimensional structures and analyse the relationship between structure and function.
- CO4 Explore protein complex modelling approaches.

**Note: Nine questions are to be set in all and the candidates are required to attempt five questions including compulsory question. Question no. I is compulsory covering the whole syllabus. Out of remaining eight questions, two questions are to be set from each unit. Candidate is required to attempt four questions, selecting one question from each unit.**

## **UNIT I**

### **Macromolecular Structure**

Conformational Analysis of protein, Forces that determine protein structure, polypeptide chain geometries, Ramachandran Map, potential energy calculations, observed values for rotation angles, structure comparison and alignment. Conformational analysis of nucleic acids and carbohydrates, general characteristics of nucleic acid structure – geometries, glycosidic bond, rotational isomers and ribose puckering - forces stabilizing ordered forms – base pairing, base stacking.

## **UNIT II**

### **Computational structural biology**

Overview of molecular modelling - Introduction and challenges; Molecular modelling methods; Conformational searching, Potential energy maps, Ramachandran maps, *Ab-initio* methods, Semi-empirical methods; Empirical methods, Conformational analysis-Introduction and Methods: Molecular fitting, Energy Minimisation. Molecular dynamics simulations of Bio-macromolecules.

## **UNIT III**

### **Molecular visualization tools**

Visualization of tertiary structures, quaternary structures, architectures and topologies of proteins and DNA using molecular visualization softwares such as RasMol, Cn3D, SPDBV, Chime, Mol4D, etc.

## **UNIT IV**

### **Structure prediction tools and homology modelling**

Prediction of secondary structures of proteins using at least 5 different methods with analysis and interpretation of the results. Comparison of the performance of the different methods for various classes of proteins. Internet based modeling tools.

## **POST GRADUATE (P.G.) CERTIFICATE COURSE IN BIOINFORMATICS**

**(Syllabus for the session 2014-15 onwards)**

Prediction of tertiary structures of proteins using Homology Modeling approach: SWISSMODEL, SWISS-PDB Viewer Prediction of tertiary structures of proteins using at least 3 methods for fold recognition along with analysis and interpretation of results. Structure Databases – PDB, NDB, CCD - Structural Classification – SCOP, CATH, FSSP.

### **Suggested Readings:**

1. Protein structure prediction: methods and protocols by Webster, D.M., Ed. Totowa Humana Press.
2. Molecular protein domains by Gimona, G. Cesareni, & Yaffe, M. Sudol(EDS.), USA., Wiley-vech verlag gmbh & co., 3-527-30813-X.
3. Molecular modeling: basic principles and applications by Holtje, H.D. & Folkers, G., Weinheim, VCH.
4. Molecular Modeling: Basic Principles and application by Hans Dieter & Didier Rognan, Wiley VeH Gmbh and Co.KGA.
5. Principles of Protein X ray Crystallography by Jan Drenth (2nd Edition)
6. Bioinformatics for comparative proteomics, C.H. Wu, C.Chen, Springer, ISBN 978-1-60761-976-5
- 7 Proteome Bioinformatics, S. J. Hubbard, A. R. Jones, Springer, ISBN 978-1-60761-443-2
- 8 Structural Bioinformatics edited by Jenny Gu, Philip E. Bourne
- 9 Structural bioinformatics: an algorithmic approach By F. J. Burkowski



# **POST GRADUATE (P.G.) CERTIFICATE COURSE IN BIOINFORMATICS**

(Syllabus for the session 2014-15 onwards)

**CBI 104: Biostatistics**

**MM: Th 80 + IA:20**

**Time: 3 Hours**

**Students completing this course will be able to:**

- CO1** Explain & differentiate probability concept and distribution for analyzing large biological data.
- CO2** Describe Boolean network modeling , bayesian belief network and modelling of metabolic network.
- CO3** Exploration of Low Dimensional projections, One Dimensional histogram ordering and Two Dimensional Scatterplot ordering.
- CO4** Analyze in detail with reference to genetics following techniques & solve questions based on them:  $\chi^2$  -test, t and f tests confidence intervals.

**Note: Nine questions are to be set in all and the candidates are required to attempt five questions including compulsory question. Question no. I is compulsory covering the whole syllabus. Out of remaining eight questions, two questions are to be set from each unit. Candidate is required to attempt four questions, selecting one question from each unit.**

## **UNIT I**

Probability Concepts and Distributions for Analyzing Large Biological Data, Basic Concepts, Conditional Probability and Independence, Random Variables, Expected Value and Variance, Distributions of Random Variables, Joint and Marginal Distribution, Multivariate Distribution, Sampling Distribution.

## **UNIT II**

Clustering: Unsupervised Learning In Large Biological Data, Measures of Similarity, Clustering, Assessment of Cluster Quality, Statistical Network Analysis For Biological Systems And Pathways, Boolean Network Modeling, Bayesian Belief Network, Modeling of Metabolic Networks.

## **UNIT III**

Multidimensional Analysis and Visualization on Large Biomedical Data, Classical Multidimensional Visualization Techniques, Two-Dimensional Projections, Issues and Challenges, Systematic Exploration of Low-Dimensional Projections, One-Dimensional Histogram Ordering, Two-Dimensional Scatterplot Ordering.

## **UNIT IV**

Samples and Sampling Distribution, Standard Error, significance level, Degrees of freedom, Tests of significance, tests for proportion, t and F tests Confidence Intervals, Contingency tables of  $\chi^2$  (Chi square) tests of goodness of fit and homogeneity, Correlation: Simple, Partial and Multiple Correlation, Methods of averages and least squares, polynomial fitting, Regression Analysis. Analysis of variance for one and two way classification.

# **POST GRADUATE (P.G.) CERTIFICATE COURSE IN BIOINFORMATICS**

**(Syllabus for the session 2014-15 onwards)**

## **Suggested Readings:**

1. Basic Biostatistics By B. Burt Gerstma
2. Biostatistics: The Bare Essentials By Geoffrey R. Norman, David L. Streiner
3. Fundamentals of Biostatistics By Bernard Rosner
4. Briley N.J.T : Statistical Methods in Biology
5. Sokal R.R., Rohlf E.J : Introduction to Biostatistics