

BTS 504 RESEARCH METHODOLOGY AND BIOINFORMATICS
(SOFT CORE COURSE)

Hours: 40

Course outcome

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

- CO 1. Acquire knowledge about basic concepts of research, scientific writing and paper publications
- CO 2. Use statistical measures such as dispersion, normal, binominal and poisson distribution, student's t-test, ANOVA, chi-square test etc.
- CO 3. Use databases, sequence alignment programs, BLAST and FASTA along with algorithms and applications.
- CO 4. Construct a phylogenetic tree and carry out protein structure analysis, protein prediction tools
- CO 5. Perform Computer Aided Drug Design (CADD) and apply it to design new drugs.

UNIT I (13 hrs)

Empirical science, scientific method, literature review, research gaps, questions, objectives, design, lab notebook. Search engines. Sampling - Experiments and controls. Data collection, quantitative and qualitative analysis. Determining outcomes and results. Ethics in research, scientific misconduct. Plagiarism. Technical writing skills, scientific papers. Referencing. Statistics – Definition, application of statistics in Biosciences, Classification and tabulation, Graphical representation of data, Histogram, frequency polygon, frequency curve. Measures of central tendency, Measures of dispersion. Normal distribution, Binomial, Poisson, Probability, non-parametric statistics, Correlation and regression; Sign test, Rank sum test, Rank correlation. Testing of hypothesis: Significance of t-test and ANOVA, Multiple range test, Chi-square test. Experimental designs. Diversity measures and evenness (e.g. Simpson and Shannon-Wiener). Statistical packages.

UNIT II (13 hrs)

Introduction to Bioinformatics. Basics of UNIX OS and PERL Programming. Biological databases: Nucleotide and protein sequence and structure (primary and secondary) databases, File formats, Molecular visualization softwares. Sequence analysis. Sequence Alignment: Gap penalties, scoring matrices, Alignment algorithms - Global and Local alignments, Dynamic programming and Heuristic methods (BLAST, FASTA). Multiple Sequence Alignment: Tree alignment, Star alignment, Progressive alignment methods and tools. Stand alone packages for sequence alignment: GCG Wisconsin and EMBOSS package.

UNIT III (14hrs)

Phylogenetics. Representation of phylogeny. Methods of phylogeny: Maximum Parsimony, Maximum Likelihood, Distance method, UPGMA. Softwares for phylogenetic analysis: PHYLIP, CLUSTAL, Tree viewing and editing softwares. Nucleotide sequence and structure prediction methods and tools: Promoter Scan, Gen Scan, CENSOR, Repeat Masker. Whole genome analysis. Genome sequencing strategies, Restriction mapping, Primer designing. Gene Expression analysis - microarray techniques. Protein sequence and structure prediction, Molecular modeling softwares and servers, Protein folding, Threading. Computer-aided Drug Designing: Molecular Docking. Distributed computing approach: Genome@home, Folding@home.

References

1. Research Methodology Methods and Techniques. Kothari, C.R., New Age Publishers, New Delhi 2004

2. Beginning Perl for Bioinformatics. Tisdall, J.D., San Val Pub., 2001
3. Bioinformatics: Sequence and Genome Analysis. Mount, D.W., CSHL Press, 2004
4. Bioinformatics: Methods and protocols. Misener, S., & Krawetz, S. A., Humana Press, 2000
5. Fundamental Concepts of Bioinformatics. Krane, D.E. & Raymer, M.L., Pearson Ed., 2002
6. Introduction to Protein Structure. Branden, C.-I. & Tooze, J., Garland Pub., 1999
7. Introduction to Bioinformatics. Attwood, T. & Parry-Smith, D., Prentice Hall Pub., 1999
8. Introductory Statistics for Biology. Parker, R.E., Hodder Arnold Pub., 1979
9. Statistics for Biological Sciences. Scheffler, W. C., Addison Wesley Pub., 1979
10. Biostatistical Analysis. Zar, J. H. Prentice Hall, 2010
11. Biostatistics. Lewis, A. E.. Prentice Hall, 2010

