Name of the Programme: M. Sc (Botany) Course Code: BOT-605 Title of the Course: Lab in Genome Informatics Number of Credits: 1 (30 hours) Effective from AY: 2022-23

Prerequisites	Basic knowledge of molecular biology and computers.	
Objective(s):	To provide practical experience in using standard computational	
Objective(s):	to provide practical experience in using standard computational tools and databases. To facilitate the investigation of molecular biology and evolution-related concepts. To train in modern methods of biological analysis.	
<u>Content:</u>	 Exploring National Centre for Biotechnology Information (NCBI) database, PUBMED and GenBank databases, (NCBI), European Bioinformatics Institute (E.B.I.) server, and searching the European Molecular Biology Laboratory (EMBL) Nucleotide database, Entrez (Global Query Cross-Database Search System), SWISSPROT & UniProtKB. 	2 hours
	2. Sequence retrieval of D.N.A. and Protein from different databases.	2 hours
	3. Homology searches using different versions of the Basic Local Alignment Search Tool (BLAST) and interpretation of the results to derive the biologically significant relationships of the query sequences (proteins/D.N.A.) with the database sequences.	2 hours
	 Use of scoring matrices, Pair-wise local alignments of protein, and D.N.A. sequences using the Smith-Waterman algorithm and interpretation of results. 	2 hours
	 Multiple sequence alignments of sets of sequences using the web-based and stand-alone version of CLUSTAL. Interpretation of results to identify conserved and variable regions and correlate them with physico-chemical and structural properties. 	2 hours
	6. Search and retrieval: genomic data at NCBI server, Interpreting D.N.A. and Protein microarray data.	2 hours
	 Use of gene prediction methods (Genscan,/Glimmer), various primer designing, and restriction site prediction tools. 	2 hours
	8. Promoter analysis of different genes and T.F. binding sites.	2 hours
	9. Use of different protein structure prediction databases Protein data bank (PDB), SCOP Structural Classification of Protein Database (SCOP), Protein Structure Classification Database (CATH).	2 hours
	 Exploring and using the derived databases: PROSITE, PRINTS, BLOCKS, Pfam, and Prodom for pattern searching, domain searches, etc.). 	2 hours
	11. Protein-protein interaction study tools.	2 hours

	12. Construction and study of protein structures using RASMOL/Deepview/PyMol. Homology modelling of proteins. Use of tools for mutation and analysis of protein structures.	2 hours
	 Phylogenetic analysis of protein and nucleotide sequences, tree building, databases for barcoding. 	2 hours
	14. Use of galaxy tool for D.N.A. sequence analysis and NGS data.	2 hours
	15. Use of R language in data analysis.	2 hours
Pedagogy:	Internet-based tools, hands-on and group exercises, videos, moodle guided exercises, and expert lectures.	
References/ Readings:	 moodle guided exercises, and expert lectures. Andrew Leach. (2001). Molecular Modelling: Principles and Applications, Prentice Hall. Antao, T. (2018). Bioinformatics with Python Cookbook: Learn how to use modern Python bioinformatics libraries and applications to do cutting-edge research in computational biology. Packt Publishing Ltd. Attwood, T. K., Parry-Smith, D. J. and Phukan S. (2022). Introduction to Bioinformatics Delhi. Pearson Education (Singapore) Pvt. Ltd. Bajorath J. (2004) Chemoinformatics: Concepts, Methods, and Tools for Drug Discovery (Methods in Molecular Biology), Humana Press. Baxevanis, A.D., Davison, D.B., Page, R.D.M., and Petsko, G.A. (2004) Current Protocols in Bioinformatics by, New York, John Wiley & Sons Inc. Bourne Philip E. and Weissig Helge (2003). Structural Bioinformatics - Methods of biochemical Analysis V. 44. New Jersey. Wiley-Liss. Bujnicki, J. M., Droogmans, L., Grosjean, H., Purushothaman, S. K., and Lapeyre, B. (2008). Practical Bioinformatics. Springer. Davies T.M. (2016). The Book of R: A First Course in Programming and Statistics. No Starch Press. New Zealand. 	
	 Dov Stekel, (2003) Microarray Bioinformatics; Cambridge University Press. Fasman, G.D. (1989). Prediction of protein structure and the principles of protein conformation. New York. Plenum Press. Friesner, R.A. Ed., Prigogine, L. Ed. and Rice, S.A. (2002). Computational methods for protein folding: Advances in chemical physics vol. 120. New York. John Wiley and Sons, Inc. Publication. Gimona, G. Cesareni and Yaffe, M. Sudol (Eds.). (2004). Modular protein domains, USA, Wiley-vch Verlag gmbh and co. 3-527-30813-X. Gundertofte, K. and Jorgensen, F.S. (2000). Molecular modelling and prediction of bioactivity, New York. Kluwer Academic Publishers. 	

	Mount, David. (2004). Bioinformatics: Sequence and Genome	
	Analysis. New York, Cold Spring Harbor Laboratory Press.	
	Rastogi, S.C., Medirattia, N. and Rastogi. P. 4 ed (2013).	
	Bioinformatics, methods and applications, genomics,	
	proteomics, and drug discovery, Prentice Hall of India, Pvt.	
	Ltd., New Delhi.	
	Solomon K. A. (2011). Molecular Modelling and Drug Design. MJ publisher.	
	 Stephen Misener and Stephen Krawetz. (2004). Bioinformatics, methods and protocols, methods in molecular biology, Volume 132, Humana Press, New Jersey, Third Indian reprint Webster, D. M. Ed. (2000). Protein structure prediction: methods and protocols. Totowa Humana Press 	
	Public domain database/tools/resources	
	 DBGET-http://www.genome.jp/dbget/ 	
	 LinkDB-http://www.genome.jp/dbget/linkdb.html 	
	Fgeneshttp://www.softberry.com/berry.phtml?topic=products	
	GeneBuilder-http://www.itb.cnr.it/sun/webgene/	
	 GeneSCAN-http://genes.mit.edu/GENSCAN.html 	
	• GRAIL-http://compbio.ornl.gov/Grail-1.3/	
	• CLC Free Workbench http://www.clcbio.com/index.php?id=28	
	BioEditor-http://bioeditor.sdsc.edu/	
	• CN3D 4.1 -	
	http://www.ncbi.nlm.nih.gov/Structure/CN3D/cn3d.shtml	
	Protein	
	Explorerhttp://www.umass.edu/microbio/chime/pe_beta/pe/pr	
	otexpl/f rntdoor.htm	
	 Chimera-http://www.cgl.ucsf.edu/chimera/ 	
	• Yasara-http://www.yasara.comhttp://www.yasara.com)	
	 Ribosome builder-http://rbuilder.sourceforge.net/ 	
	 ArrayExpress-www.ebi.ac.uk/arrayexpress/ 	
	• EPICLUST-http://ep.ebi.ac.uk/EP/	
Learning	1. Will enable to work with computational tools and to gain	
Outcomes:	practical knowledge.	
	2. Will enable to analyze Next Generation Sequencing (NGS) raw	
	data.	
	3. Will help to be better equipped to investigate specific	
	contemporary biological questions.	