Systems Biology and Network Analysis for Enabling Research in Personalized Genomics

Overview

The course introduces data integration and statistical methods used in contemporary Systems Biology, Bioinformatics and Systems Pharmacology research. It covers methods to process raw data from genome-wide mRNA expression studies (microarrays and RNA-seq) including data normalization, differential expression, clustering, enrichment analysis, network construction, PPI, and accessing, visualizing and predicting macromolecule structure data. The course contains practicals and tutorials for using tools and setting up pipelines, but it also covers the mathematics behind the methods applied within the tools. Few advanced topics related to data analysis will also be covered.

The course is mostly appropriate for beginning graduate students and advanced undergraduates majoring in fields such as biology, math, physics, chemistry, computer science, biomedical and electrical engineering. The course should be useful for researchers who encounter large datasets in their own research. The course presents various software tools that will be helpful for the participants for conducting their research work.

Dates	**December 18 - December 22, 2023** 14 hrs lectures and 10 hrs tutorials **Number of participants for the course will be limited to fifty.**
You Should Attend If…	▪ You are an experimentalist who wants to analyze your own data, a computational biologist or a computer science/IT engineer interested to learn advance algorithms to analyze biological data. ▪ You are a student (BTech/MTech/MSc/PhD) or faculty or a person from the industry interested in learning how to do research on "big-data" problems in personalized genomics.
Fees	The participation fees for taking the course is as follows: **Participants from abroad :** US $500 **Industry/ Research Organizations (Within India):** Rs 10,000 **Academic Institutions (Within India)** BTech/MTech/MSc Students: Rs. 1,500 PhD/Post-Doctoral Fellows: Rs. 3,000 Faculty/Researchers/Scientists/Staff from Academic Institutions: Rs. 4,000 The above fee include all instructional materials, computer use for tutorials and assignments, refreshments during breaks in the program and free internet facility. The participants are encouraged to bring their laptops. Accommodation may be provided on availability and on payment basis. At the completion of the course a certificate of participation will be issued. **Last date to apply:** December 10, 2023
Mode of Registration	All prospective participants need to do web registration for the course on GIAN (http://www.gian.iitkgp.ac.in/) portal. After the mandatory web registration, participants should share the registration details with the course coordinator by email (sdey@iitj.ac.in). The shortlisted participants will be informed by email to register for the course by making full payment of the course registration fee.

The Faculty

Professor Animesh Ray earned his PhD in microbial genetics from Monash University in Melbourne, Australia. His PhD research led to the identification of a gene for efficient plasmid maintenance in Escherichia coli and a method for generating a multi-copy infectious plasmid that is packageable inside a virus coat an early example of synthetic biology. He subsequently conducted research at the Institute of Molecular Biology, University of Oregon, and the Department of Biology, Massachusetts Institute of Technology; University of Rochester. In the late 1990s, Dr. Ray, along with a computer scientist colleague Dr. Mitsunori Ogihara, published a series of papers on experimental and theoretical investigations on designing massively parallel computing devices using solution phase DNA chemistry. He currently teaches courses on molecular systems biology that includes molecular mechanisms of human diseases and pharmacogenomics. He was KGI's faculty chair (2010-2016) and director of KGI's PhD programs (2006-2016). His current research work involve systems biology of Huntington's disease, chromosome instability, non-coding RNAs in cancers, and cancer drug resistance mechanisms.

Dr. Sucharita Dey received her PhD degree from University of Calcutta, Bose Institute Kolkata,. After obtaining her PhD degree in 2013, she worked as a post doctoral fellow at the National University of Singapore (June 2013 to Feb 2015). Later, she joined the Weizmann Institute of Science for her second postdoctoral research (Feb 2015 to March 2021). Currently, she is an assistant professor in the Department of Bioscience and Bioengineering, Indian Institute of Technology Jodhpur, India. Her research interests include computational biology and structural bioinformatics. She has published 20 papers so far in international journals and conferences, and 1 book chapter. Dr. Dey is a recipient of 2021 Ramalingaswami re-entry fellowship and a research grant from the Department of Biotechnology, Government of India.

Course Co-ordinator

Dr. Sucharita Dey
Department of Bioscience and Bioengineering
Indian Institute of Technology Jodhpur
NH-62 Nagaur Road Karwar
Jodhpur 342037
Rajasthan
Phone: 0291-2801218
E-mail: sdey@iitj.ac.in

……

http://www.gian.iitkgp.ac.in/

NIHRD Scheme on GIobN Initiative on Academic Network (GIAN)

SYSTEMS BIOLOGY AND NETWORK ANALYSIS FOR ENABLING RESEARCH IN PERSONALIZED GENOMICS

1.0 Overview

The course introduces data integration and statistical methods used in contemporary Systems Biology, Bioinformatics and Systems Pharmacology research. It covers methods to process raw data from genome-wide mRNA expression studies (microarrays and RNA-seq) including data normalization, differential expression, clustering, enrichment analysis, network construction, and accessing, visualizing and predicting macromolecule structure data. The course contains practical tutorials for using tools and setting up pipelines, but it also covers the mathematics behind the methods applied within the tools. Few advanced topics related to data analysis will also be covered.

The course is mostly appropriate for beginning graduate students and advanced undergraduates majoring in fields such as biology, math, physics, chemistry, computer science, biomedical and electrical engineering. The course should be useful for researchers who encounter large datasets in their own research. The course presents various software tools that will be helpful for the participants for conducting their research work.

2.0 Objectives

The primary objectives of the course are as follows:

i) Illustrate “big-data” problems in personalized genomics by discussing recent publications,
ii) Provide training in high dimensional biological data analysis,
iii) Familiarize the participants with the use of statistical tools and principles for analyzing genomic/epigenomic data,
iv) Familiarize with macromolecule structure data
v) It will enable participants to utilize the methods presented in this course for analyzing their own research data for their own projects. For those participants that do not work in the field, the course introduces the current research challenges faced in the field of computational systems biology.

3.0 Course details Schedule

Day1

Lecture 1– 1 hr: AR 9 to 9.50 AM

Course Overview and Introductions: Design Principles of Complex Systems, Introduction to Genomics and Systems Biology

Lecture 2– 2 hrs : AR 10 to 12 PM (2 x 10 mins break in between)

Topological and Network Evolution Models: Small-World and Scale-Free Networks, Network Motifs, Large Size Motifs and Complex Models of Network Evolution, Network Properties of Biological Networks.

Student Project and Team assignment

LUNCH

Tutorial 1– 2 hrs: AR 3 to 5 PM

Problem solving session with examples:
R programming language
Working with R
Drawing simple networks
Generate networks from association data
Network from adjacency matrix

Formal Classes end at 5PM

Students retire & start working on team projects at night

Day 2

Lecture 3 : 1.5 hrs: AR 9 to 10.30 AM

Types of Biological Networks: Genetic, physical and metabolic Networks

Lecture 4: 1.5 hrs: AR 10.30 to 12 PM

Functional Enrichment Analysis (GO Enrichment)

Follow up on Student Projects and Team

LUNCH

Tutorial 2: 2 hrs: AR 3 to 5 PM

Problem solving session with examples:
Handling Biological networks
Network Visualization
Network topological properties
Random networks of distinct topologies

Real networks

Formal Classes end at 5PM
Students retire & work on team projects at night

Day 3

Lecture 5 : 1.5 hrs: SD 9 to 10.30 AM

Data Processing and Identifying Differentially Expressed Genes: Data Normalization, Univariate and multivariate methods for identifying differentially expressed gene sets - Part 1, Univariate methods (SAM, Welch's t-test, LIMMA, DEseq) - Part 2,

Lecture 6: 1.5 hrs: AR 10.30 to 12.00 PM

Multivariate method (Principal Component Analysis; Gene Set Enrichment Analysis; Supervised and unsupervised machine learning for data classification, including Clustering, — Part 3 (Characteristic Direction Method)

LUNCH

Tutorial 3: 2 hrs: SD and AR 3 to 5 PM
Problem solving session with examples
Discuss Team Project progress

Formal Classes end at 5PM
Students retire & work on team projects at night

Day 4

Lecture 7 : 1 hrs: SD 9 to 10 AM
Introduction to Biomacromolecules: Protein structure, function, databases

Lecture 8: 1.5 hrs: SD 10 to 11:30 PM
Protein Structure modelling: Homology modeling, ab-intio modeling (AlphaFOLD)

Tutorial 4: 2 hrs: SD 3 to 5 PM
Problem solving session with examples:
PDB
PYMOL
SWISS MODEL
ALPHAFOLD
Discuss Team Project progress

Formal Classes end at 5PM
Students retire & work on team projects at night

Day 5 9 to 12 AM
Lecture 9– AR 1 hr: Prediction of protein-protein interaction using machine learning

Lecture 10–SD 1hr lecture by external expert on MD simulation, exact title: to be decided

Teams work on final team project presentations

LUNCH

Evaluation: 3 to 6 PM
Presentations by students teams: 20mins talk + 5 min Q/A per team
Evaluated by IIT faculty invites

6:30-7:30 PM–Announcement of winning teams and celebration