Course Code	BIO541
Course Name	Data sciences in Genomics
Credits	4
Course Offered to	UG/PG
Course Description	The field of Genomics is expanding its horizon with help of high throughput technologies. Scientist are trying to answer fundamental questions related to health, society and human survival outside earth using genomics. With such increase of applications, it constantly needs computational experts for systematic analysis of data and acheiving meaning full insights. Infact the computational experts have now started taking lead in genomic projects. Hence this course is meant to guide students for data analysis approach and steps involved in computational genomics and make them familiar with latest development in genomics.

Pre-requisites

Pre-requisite (Mandatory)	Pre-requisite (Desirable)	Pre-requiste (Other)
None		None

Post Conditions

CO1	CO2	CO3	CO4
students can understand source and properties of different kinds of genomic data	students can use tools for regular processing different kinds of genomics data	Students can undertsand the statistical approach used by tools for processing genomics data	Students can perform few complex post processing of data to make meaningful clinical and biological inference

Weekly Lecture Plan

Week Number	Lecture Topic	Cos met	Assignments/Lab/Tutorials
Week 1	Introduction to molecular cell biology, central dogma, measurements using microarray and ngs technologies;	CO1	Assignment: (i) Paper reading on cellular functions and biological processes and writing in own words
Week 2-3	Introduction to Genetics/Genomics; Applications to clinical diagnostics;	CO1,CO2	
Week 4	NGS concepts and applications; DNA alignment algorithms and assessment; Data QC and Filtering	CO1, CO2	Assignment2 : R and Bioconductor; R packages for annotations, short reads handling, microarray analysis and visualizations;
Week 5	Applications of NGS in single cell study, bias, drop-out and noise in single cell data		
Week 6-7	Unsupervised (Clustering and Biclustering) and supervised (Classification) , statistical approach in Normalization and Differential gene expression (microarray, bulk RNA-seq and single cell RNA-seq)	C02, CO3, CO4	
Week 8	Network inference methods, WGCNA, visualisation	CO3, CO4	
Week 9-10	introduction to epigenome and analysis of epigenome ngs data, handling bias and noise, application of epigenome	C02, CO3	Assignment3: Processing ChIP-seq and RNA-seq datasets to make clinical inference
Week 11	Denovo assembly; Assembly assessment; Assembly refinement; Variant calling (SNPs, MNPs, small InDels)	C02, CO3	
Week 12-13	Variant annotations and interpretation, statistical method for effect prediction and QTL; Visualization of genomic data, population Genomics	CO3, CO4	Assignment 4: QTL analysis
	Project presentations		

Assessment Plan

Type of Evaluation	% Contribution in Grade
Mid-sem	20
End-sem	20
Assignments	20
Project	40

Resource Material

Type	Title
Textbook	B. Pierce, Genetics: A conceptual approach,
Textbook	https://en.wikibooks.org/wiki/Next_Generation_Sequencing_(NGS)
Textbook	Xinkun Wang, Next-Generation Sequencing Data Analysis,
Publication	Ryuichiro Nakato Katsuhiko Shirahige, Brief Bioinform (2017) 18 (2): 279-290 ,Recent advances in ChIP-seq analysis: from quality management to whole-genome annotation
Publication	Ryuichiro Nakato and Katsuhiko Shirahige, Recent advances in ChIP-seq analysis: from quality management to whole-genome annotation, Brief Bioinform. 2017 Mar; 18(2): 279–290.
Publication	Juliana Costa-Silva, Douglas Domingues, Fabricio Martins Lopes , RNA-Seq differential expression analysis: An extended review and a software tool PLOS one.
Publication	Robert Ekblom and Jochen B W Wolf, A field guide to whole-genome sequencing, assembly and annotation, Evol Appl. 2014 Nov; 7(9): 1026–1042.