JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -3 EXAMINATION- 2025

B.Tech-VIII Semester (BT/BI)

COURSE CODE (CREDITS): 18B1WBI834 (3) **MAX. MARKS: 35**

COURSE NAME: NGS Data Analysis & Applications

COURSE INSTRUCTORS: Dr. Shikha Mittal **MAX. TIME: 2 Hours**

Note: (a) All questions are compulsory.

(b) The candidate is allowed to make Suitable numeric assumptions wherever required for solving problems

Q.No	Question	CO	Marks
Q1	Compare reference-based genome assembly and de novo genome assembly.	[CO-1, CO-3]	(5)
Q2	You have obtained 10 million 150 bp paired-end reads. Calculate the total number of base pairs sequenced.	[CO-3]	(3)
Q3	In a differential gene expression analysis, 1200 genes were found to be significantly differentially expressed with FDR < 0.05. If 25,000 genes were analyzed, calculate the percentage of significant genes.	[CO-2, CO-4]	(3)
Q4	**a)** Discuss the role of variant detection in personalized medicine **b)** You detect 450 SNPs from a 1 Mb region of the genome. What is the SNP density (SNPs per kb)?	[CO-4]	(4)
Q5	**a)** Discuss the role of base quality scores (Phred scores) in read filtering and trimming **b)** You have an NGS dataset with a mean Phred quality score of 35. What is the probability of a base call being incorrect?	[CO-1, CO-3]	(4)
Q6	You are provided with a raw FASTQ file from an RNA-seq experiment. Describe the step-by-step pipeline (with specific software tools) to obtain differentially expressed genes, including data QC, mapping, quantification, and statistical testing.	[CO-1, CO-4]	(5)
Q7	Describe the principle and workflow of Sanger Sequencing. How does it differ from Next Generation Sequencing (NGS)?	[CO-1, CO-2 & CO-3]	(3)
Q8	**a)** Exome sequencing **b)** Metagenomics **c)** Genome annotation **d)** Variant detection	[CO-2, CO-4]	(8)