JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST-3 EXAMINATIONS-2022

B.Tech-VIII Semester (BI)

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

COURSE NAME: NGS Data Analysis and Applications

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

***Note:** All questions are compulsory. Marks are indicated against each question in square brackets.*

Q1. What are the differences between first generation sequencing and next generation sequencing technologies? What were the drawbacks of first generation sequencing and advantages of next generation sequencing? [4 marks] (CO-1 & CO-2)

Q2. Explain the term "genome annotation". How to perform structural and functional annotation of genomes? Mention the software's used for structural and functional annotation of genomes. [3 marks] (CO-2 & CO-3)

Q3. Suppose, the sequencing of wheat under sodicity stress was performed and paired end have been received by RNA sequencing. What will be the further steps used for RNA-Seq data analysis. What are the pipelines used for *denovo* and reference based assembly. [5 marks] (CO-1 & CO-3)

Q4. Now-a-days next generation sequencing is widely used in multiple fields of biosciences. Describe in detail the fields in which next generation sequencing is used. [4 marks] (CO-4)

Q5. Exome comprises of approximately what % of the human genome? Complex diseases result from a combination of genetic and environmental factors, many of which are not understood. State the advantage associated with exome sequencing in the case of complex diseases? [3 marks] (CO-2)

Q6. Explain sequence formats, feature formats and alignment formats commonly used in different steps of next generation sequencing. [3 marks] (CO-2 & CO-3)

Q7. Describe the principle of Illumina sequencing technology? How to identify differentially expressed genes between a normal and a stressed condition? [4 marks] (CO-1 & CO-2)

Q8. What do understand by Phred score in a fastq file? What is the minimum requirement of Phred score to start further downstream analysis? For a sequence, if Phred score is 30, what will be the percentage inaccuracy? [3 marks] (CO-3)

Q9. What are the technologies of single cell gene expression analysis? Explain in detail. [3 marks] (CO-4)

Q10. Explain the following terms – [3 marks] (CO-4)

i. When we found a SNP in a genome, it might be of two types i.e., transition or transversion. Explain the terms transition and transversion with example.

ii. After annotation, the next step is to find the hub genes involved in a specific pathway. What is pathway analysis and how do we perform it?

iii. Metagenomics