JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -3 EXAMINATION- MAY-2023

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

COURSE NAME: NGS Data Analysis & 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 is differential gene expression? What are the thresholds that are used for identification of differentially expressed genes? [3 marks] (CO-3)

Q2. Describe the applications of Next-generation sequencing technology in the field of healthcare and agriculture. [4 marks] (CO-4)

Q3. 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)

Q4. 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 40, what will be the percentage inaccuracy? [3 marks] (CO-3)

Q5. Explain the following terms – [6 marks] (CO3 & CO-4)

i. Biological and Technical Replicates.
ii. Genome annotation
iii. Metagenomics

Q6. 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-1 & CO-3)

Q7. Suppose you received raw data from Illumina sequencer, how will you the quality of this data? What parameters must be considered for quality checking? [4 marks] (CO-2 & CO-3)

Q8. Define N50. What is *denovo* assembly and explain the steps of *denovo* assembly using flowchart [4 marks] (CO-1 & CO-3)

Q9. Name the sequencing techniques used for first, second and third generation sequencing and explain Illumina sequencing in detail. [5 marks] (CO-1 & CO-2)