

**JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT**

**TEST -2 EXAMINATION- 2024**

**B.Tech-VIII Semester (BT/BI)**

**COURSE CODE (CREDITS): 18B1WBI834**

**MAX. MARKS: 25**

**COURSE NAME: NGS Data Analysis & Applications**

**COURSE INSTRUCTORS: Dr. Shikha Mittal**

**MAX. TIME: 1 Hour 30 Minutes**

*Note: (a) All questions are compulsory.*

*(b) Marks are indicated against each question in square brackets.*

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

**Q1.** Denovo assembly is a method for constructing genomes from a large number of DNA fragments with no prior knowledge of order of fragments. Explain the steps involved in denovo assembly with the help of a flowchart. (3 marks) [CO-1 & CO-2]

**Q2.** How to check the quality of raw data generated from a sequencer. What are the parameters that can be checked for the quality of sequences? (4 marks) [CO-1]

**Q3.** Briefly explain the significance of 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-1]

**Q4.** Explain various types of formats commonly used in different steps of next generation sequencing. (3 marks) [CO-2 & CO-3]

**Q5.** 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]

**Q6.** Briefly explain - [8 marks] [CO-1, CO-2 & CO-3]

- (a) FASTQ format
- (b) Genome annotation
- (c) Biological and technical replicates
- (d) Synonymous and Non-synonymous SNPs