JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -3 EXAMINATIONS- 2024

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)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. Discuss the major steps involved in the NGS workflow, including library preparation, sequencing, and data analysis. **[CO-2 & CO-3] (4 marks)**

Q2. Given the following information from an RNA-seq experiment: **[CO-3] (4 marks)**

Total number of fragments mapped to all genes: 1,000,000
Length of gene A: 2,000 base pairs
Number of fragments mapped to gene A: 10,000

Calculate the FPKM (Fragments Per Kilobase Million) value for gene A.

Q3. Explain the principles behind Illumina sequencing technology and describe how it differs from other NGS platforms. **[CO-1 & CO-2] (4 marks)**

Q4. Describe the process of quality control and preprocessing of NGS data, including steps for read trimming, adapter removal, and filtering for high-quality reads. **[CO-3] (5 marks)**

Q5. Discuss the statistical methods commonly used to assess differential gene expression, including their strengths and limitations. **[CO-4] (3 marks)**

Q6. In a sequencing experiment, a base call has a Phred score of 30. What is the probability that this base call is incorrect? Provide the formula used to calculate this probability and show the calculation. **[CO-2] (3 marks)**

Q7. Briefly explain the following – **[CO-1, CO2, CO-3 & CO-4] (2X6=12 marks)**

a. Exome sequencing
b. VCF file format
c. Difference between biological and technical replicates
d. Difference between Sanger sequencing and Maxam-gilbert sequencing
e. Salient features of HGP
f. GO annotation