JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -3 EXAMINATION- 2023

M.Tech-1 Semester (BT)

COURSE CODE (CREDITS): 13M11BT112 (3) MAX. MARKS: 35

COURSE NAME: Advanced Bioinformatics

COURSE INSTRUCTORS: Dr. Raj Kumar 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 five important resources and tools in NGS data analysis? [3]

Q2. Discuss why indexing is crucial in sequence alignment in next-generation sequencing (NGS) data analysis? [3]

Q3. Write the steps and commands to align the FASTQ files to the indexed database. [3]

Q4. Discuss the challenges in assembly of Repetitive DNA. [3]

Q5. Explain the term "mapping quality" in the context of sequence alignment. Why is it important in NGS data analysis? [3]

Q6. Write down the CIGAR string for the given examples: [5]

(a)	(b)	(c)	(d)	(e)
0123456789 AAGTC TAGAA (ref) GTCGATAG (query)	0123456789 AAGTC TAGAA (ref) GTCGATAG (query)	0123456789 AAGTCTAGAA (ref) GAATAG (query)	0123456789 AAGTCTAGAA (ref) GT TAG (query)	01234567890123456 CCCTACGTCCCAGTCAC (ref) TAC TCAC (query)

Q7. What information is typically included in the header section of a SAM file? [3]

Q8. Explain the following fields that provide information about the paired-end reads and their alignments: [3×1]

a) RNEXT
b) PNEXT
c) TLEN

Q9. Write down some of important applications of genetic variants identified in genomic data. [3]

Q10. The Variant Call Format (VCF) is a standard file format used in bioinformatics to store information about genetic variants identified in genomic data. Explain the mandatory columns of a VCF file? [3]

Q11. How are single nucleotide polymorphisms (SNPs) and insertions/deletions (indels) represented in VCF files? [3]