STELLA MARIS COLLEGE (AUTONOMOUS) CHENNAI –600 086
(For candidates admitted from the academic year 2019 – 2020 & thereafter)

M. Sc. DEGREE EXAMINATION, APRIL 2024
BIOINFORMATICS
FOURTH SEMESTER

COURSE : CORE
PAPER : ADVANCES IN BIOINFORMATICS
SUBJECT CODE : 19BI/PC/AB44
TIME : 3 HOURS MAX. MARKS: 100

SECTION – A

ANSWER ALL QUESTIONS (20 x 1=20)

1. Which of these projects would be best suited for Next Generation Sequencing?
   a) To determine if a tumour sample contains a common missense mutation
   b) To find the transcriptome of a tumour sample
   c) To genotype ten genomic DNA samples for a known single nucleotide polymorphism
   d) All of the above.
2. Once the sequences are obtained from your Next Generation Sequencing experiment what is the first thing you should do?
   a) Perform a bioinformatics analysis of your data
   b) Check your data using a different method
   c) Publish your results
   d) Further investigate the sequences of interest.
3. Personalised medicine has the potential to yield plenty of health and economic benefits. Which of the following would not be a benefit of personalised medicine?
   a) Increased number of medical jobs b) Improved medical decision making
   c) Delivery of most effective therapies d) Optimise disease prevention strategies
4. The most commonly occurring variant in the human genome is
   a). tandem-repeat polymorphism. b) premature stop codon.
   c). nucleotide base insertion. d) single-nucleotide polymorphism
5. Which of the following statements are not true - Microarrays Are used for analysis of
   i. transcriptomes ii. Contain RNA sequences
   iii. Contain DNA sequences iv. Are smaller than DNA chips

   a) i and ii correct b) ii and iii correc c) i and iii correct d) ii and iv correct

6. In terms of understanding the pathways between genes and behaviour, it is fairly safe to say that
   a) we know more about the environment than the genes
   b) we know more about the genes than the environment
   c) the new field of molecular genetics is the best way to gain a full understanding of the gene/environment interactions in the pathways
   d) QTL analysis has actually hindered the understanding of these pathways by suggesting so many separate gene contributions to basic behaviour processes
7. When you scan the microarray in the scanner, the data show some black spots. What do these represent?
   a)The DNA that has been replicated in healthy cells.
   b) The mRNA that was washed away in the washing solution.
   c) The DNA that was not transcribed and expressed in healthy cells.
   d) The mRNA that was not bound by Oligo-d-tails in the beads.

8. The two most common processes that lead to production of multiple functional proteins from the same DNA sequence are:
    a) RNA editing and alternative splicing.
    b) Protein folding and posttranslational covalent modifications.
    c) Alternative splicing and posttranslational covalent modifications.
    d) Posttranslational covalent modification and transcriptional regulation.
9. In an analysis of eukaryotic gene, you identify several non-overlapping open reading frames, but they are not all in the same frame. Which explanation makes the most sense?
    a) By random chance, a second reading frame within the gene also has an open reading frame.
    b) This gene includes introns which are not multiples of three.
    c) This is a mutant allele that has had several small insertions.
    d) All of these.
10. In which situation would you expect the mutation frequency to increase over time?
    a) The deleterious effect of the mutation is balanced with the mutation rate.
    b) A lethal allele is created at a high mutation rate.
    c) A neutral allele is created at low mutation rate.
    d) All of the above.
11. Explain about the UCSC genome.
12. Explain the term "Gene regulatory dynamics".
13. Write importance of Crispr.
14. Write a short notes on FastQC.
15. Explain the applications of targeted mutagenesis.
16. What is regulatory sequence motifs?
17. Define NCBI refseq.
18 . Write a short notes on FASTQ file formats.
19. Explain the methods adopted for quantifying RNA.
20. How to create the expression table?

SECTION – B

ANSWER ANY FOUR QUESTIONS. EACH ANSWER SHOULD NOT EXCEED 500 WORDS. ALL QUESTIONS CARRY EQUAL MARKS. (4 x 10 = 40)

21. Explain the principles of Next generation sequencing and write its importance in medicine.
22. Describe in details about the alpha and beta diversity of metagenomic studies.
23. List out the Linux commands and its functions.
24. Discuss about the repair and data analysis of the edited genome.
25. Explain the epigenetic modifications.
26. Summarize the file formats to analyze NGS data.
27. Write a case study to perform RNA seq analysis.

SECTION – C

ANSWER ANY TWO QUESTIONS. EACH ANSWER SHOULD NOT EXCEED 1200 WORDS. ALL QUESTIONS CARRY EQUAL MARKS. (2 x 20 = 40)

28. Write an essay on analysis of metagenome data and logical steps for metagenome analysis.
29. Describe the RNA dynamics at the level of transcription and post-transcriptional processing.
30. Explain in details about the salient features of transcriptomics and write the logical steps for analysing RNA seq data.
31. Briefly explain Crispr-Cas9 mechanism and its applications.

****