Indraprastha Institute of Information Technology Delhi (IIITD)

Department of Computational Biotechnology

BIO213 – Introduction to Quantitative Biology

Quiz-3 (April 26, 2023)

Question 1. Which of the following is NOT true for GOR method of secondary structure prediction? **(1 mark)**

1. It was developed by Garnier, Osguthorpe and Robson.
2. It is an information theory-based method.
3. It considers a window size of 15- one central amino acid and 7 neighboring amino acids on each flanking side.
4. It predicts four states of secondary structure- helix, strand, turn and coil.

a) All of these
b) None of these
c) 3
d) 2 and 3
e) 1 and 4

Question 2. Which of following is NOT a measure to reduce computational complexity of *Ab initio* modelling? **(2 marks)**

a) Torsion angles are restricted to a finite set of values.
b) **Protein sequence is searched against fold database to look for regions of structural similarity**.
c) Only the polar hydrogens are given importance.
d) Bulky side chains are replaced by single pseudo-atoms.

Question 3. Which of the following is NOT true about template selection step in homology modeling of proteins? **(2 marks)**

a) The first step in protein structural modeling is to select appropriate structural templates.
b) This forms the foundation for rest of the modeling process.
c) Heuristic alignment search programs are used.
d) **The template selection involves searching the non-redundant protein database for homologous proteins with determined structures.**

Question 4. A major design consideration in an RNA-seq experiment is the incorporation of replicates. Select which one/s of the following statement/s are consequences of not including replicates in your experiment: **(2 marks)**

a) The number of false-positive errors in detecting differentially expressed genes will increase.
b) The number of false-negative errors in detecting differentially expressed genes will increase.
c) No statistical analysis of the significance of the observed changes will be possible.
d) **All of the above.**

Question 5. State whether the following statements are correct or incorrect? In case of incorrect statement, justify your answer. **(4 marks)**

a) Conformational search algorithm in *ab initio* protein structure modeling explores the potential energy surface and locate the local minimum. **INCORRECT - Conformational search algorithm locates the global minimum. The native structure of the protein is believed to have the least potential energy, therefore a conformation representing the global minimum of the potential energy landscape.**

b) Logs put positive and negative value of fold changes on a symmetric scale. **CORRECT**

Question 6. Which of the following is not a conformational search method? **(2 marks)**

a) Monte Carlo
b) Genetic Algorithm
c) Simulated annealing
d) **Maximum parsimony**

Question 7. Which of the following is true with respect to RNA-seq differential gene expression analysis? **(2 marks)**

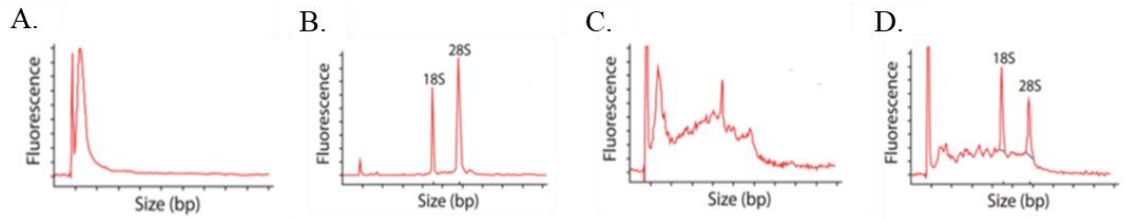
1. The chances of finding false positives increases with increasing number of genes.
2. For 40000 genes and p-value of 0.05, one would expect 2000 genes to have p-value<0.05 by chance.
3. p-value is adjusted using methods like false discovery rate.

a) 1 and 2
b) 1 and 3
c) 2 and 3
d) **All of these**
e) None of these

Question 8. Which of the following statistical tests are used if the data does not follow the normal distribution? **(2 marks)**

a) T-test and ANOVA
b) Wilcoxon test and ANOVA
c) ANOVA and Kruskal–Wallis test
d) **Kruskal–Wallis test and Wilcoxon test**

Question 9. Which of the following indicates the best and the worst quality RNA? **(2 marks)**



a) Best- A, Worst- D
b) **Best- B, Worst- A**
c) Best- B, Worst- C
d) Best- A, Worst- D

Question 10. Which of the following is the correct sequence of steps involved in RNA-seq analysis? **(2 marks)**

a) Library preparation, sequencing, quantification, read mapping, differential expression analysis
b) Quantification, library preparation, sequencing, read mapping, differential expression analysis
c) **Library preparation, sequencing, read mapping, quantification, differential expression analysis**
d) Quality check, library preparation, sequencing, read mapping, differential expression analysis

Question 11. Secondary structure is defined by _____. **(1 mark)**

a) **Hydrogen bonding**
b) Vander Waals forces
c) Covalent bonding
d) Ionic bonding

Question 12. Match the following: **(2 marks)**

1. Modeller	A. Structure validation
2. I-TASSER	B. Homology modelling
3. ERRAT	C. Secondary structure prediction
4. BHAGEERATH	D. Threading
5. Zpret	E. *Ab Initio* protein modelling

a) 1-B, 2-E, 3-A, 4-D, 5-C
b) **1-B, 2-D, 3-A, 4-E, 5-C**
c) 1-B, 2-D, 3-A, 4-C, 5-E
d) 1-E, 2-E, 3-C, 4-D, 5-A

Question 13. Find the incorrect match. **(1 marks)**

a) FASTQC- checks the quality of raw reads.
b) STAR- maps reads to the reference genome.
c) **CASAVA- quantifies the mapped reads.**
d) DEseq2- normalization of read counts and differential gene expression analysis.