

Indraprastha Institute of Information Technology Delhi (IIITD)
Department of Computational Biotechnology

BIO213 – Introduction to Quantitative Biology

Quiz-1 (February 21, 2024)

Total time: 40 mins, Total marks: 30

1. In a comparative genomics study, you are interested in identifying regions of high conservation among distantly related species. Which alignment method would provide the most relevant information? **(1 mark)**
 - a. Global Alignment
 - b. **Local Alignment**

2. Which level of protein structure is determined by the sequence of amino acids? **(1 mark)**
 - a. **Primary**
 - b. Secondary
 - c. Tertiary
 - d. Quaternary

3. Which BLAST variant compares an amino acid sequence against a protein sequence database? **(2 mark)**
 - a. Blastn
 - b. **tblastn**
 - c. blastx
 - d. tblastx

4. What does PAM 250 stand for in the context of database searching methods? **(1 mark)**
 - a. Pairwise Alignment Matrix with 250 residues
 - b. Progressive Alignment Model with 250 iterations
 - c. Point Accepted Mutations with 250 sequences
 - d. **Point Accepted Mutations with 250 percent divergence**

5. Which of the following is more divergent? **(1 mark)**
 - a. PAM1
 - b. **BLOSUM45**

6. Which of the following is the best suited for finding orthologous proteins (proteins with conserved function) in different species? **(1 mark)**
 - a. **Global alignment**
 - b. Local alignment
 - c. End-free alignment

7. State one advantage and one limitation of progressive alignment approach over dynamic programming for multiple sequences. (2 marks)

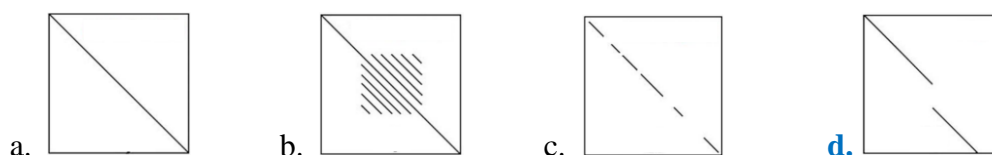
Advantage: Computationally less expensive in comparison to DP

Limitation: The order in which the sequences are compared together for the alignment dictates the final alignment in Progressive alignment.

8. Which of the following best describes local alignment? (1 mark)

- It aligns the entire sequence from start to end.
- It finds the best alignment of subsequences within the sequences.**
- It aligns sequences based only on identity.

9. Which of the following DOT plots indicates the presence of indel in the sequences? (1 mark)



10. Show all the possible sequence alignments corresponding to the paths shown with black arrows in the image below? (6 marks)

	-	A	G	A	A	C	A	A	G	G	C	G	T
-	0	-1	-2	-3	-4	-5	-6	-7	-8	-9	-10	-11	-12
A	-1	1	0	-1	-2	-3	-4	-5	-6	-7	-8	-9	-10
C	-2	0	0	-1	-2	-1	-2	-3	-4	-5	-6	-7	-8
A	-3	-1	-1	1	0	-1	0	-1	-2	-3	-4	-5	-6
A	-4	-2	-2	0	2	1	0	1	0	-1	-2	-3	-4
G	-5	-3	-1	-1	1	1	0	0	2	1	0	-1	-2
A	-6	-4	-2	0	0	0	2	1	1	1	0	-1	0
C	-7	-5	-3	-1	-1	1	1	1	0	0	2	1	0
A	-8	-6	-4	-2	0	0	2	2	1	0	1	1	2
G	-9	-7	-5	-3	-1	-1	1	1	3	2	1	2	1
C	-10	-8	-6	-4	-2	0	0	0	2	2	3	2	1
G	-11	-9	-7	-5	-3	-1	-1	-1	1	3	2	4	3
T	-12	-10	-8	-6	-4	-2	0	0	0	2	2	3	5

AGAACA-AGGCGT
| | | | |
ACAAGACA-GCGT

AGAACA-AGGCGT
| | | | |
ACAAGACAG-CGT

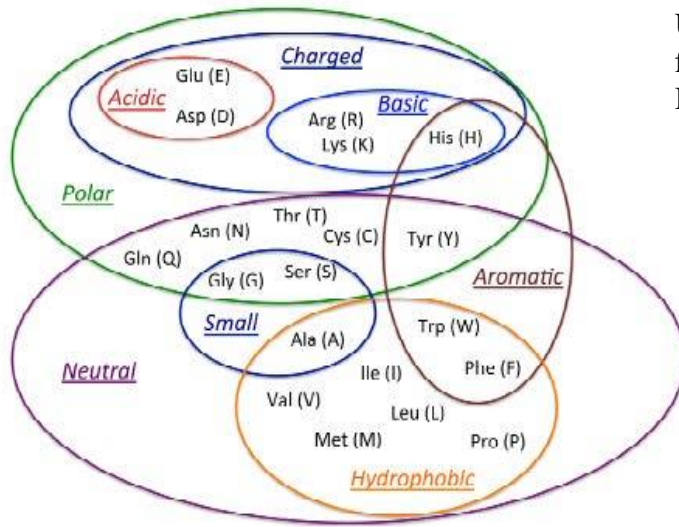
(3 marks each)

11. Find the score of the following alignment (Match = 4, Mismatch = -1, Gap opening penalty = -2, gap extension penalty = -1) (2 marks)

T A G - A - - C G A C - T
| | | | | | | | | | |
T A G C A T G C G - C A T

Score = 23

12. The provided image shows the categorization of different amino acids on the basis on their physicochemical properties. (3 marks)



Using the given information, state whether the following statements are CORRECT or INCORRECT.

- a. Substitution of Pro with Ala in a protein molecule will not make a significant structural difference as both are hydrophobic in nature.

Incorrect

- b. Log odd ratio of Glu→Asp is expected be higher in comparison to Glu→Arg.

Correct

- c. Arg, Lys and His are positively charged amino acids.

Correct

13. Get the log out ratio for P→Q. Fill in all the blank values in the section below that are required to estimate the final answer. (8 marks)

PQPP
QQQP
QQPP
QPPP
QQQP

	P	Q
P	13	14
Q		13

$$P(P) = \frac{10}{20} = 0.5, P(Q) = \frac{10}{20} = 0.5$$

$$e_{PQ} = 0.5 * 0.5 = 0.25$$

$$q_{PQ} = \frac{14}{40} = 0.35$$

$$\text{lod} = \log_2(q_{PQ}/e_{PQ}) = \frac{\log_2(0.35/0.25)}{1} = \log_2(1.4) = 0.485$$

Answer is log is also acceptable.

3 marks for the matrix, 1 mark for each value in the right panel.