PAST SOLVED PAPERS OF BIF 401

Q.. Name three hydrophobic amino acids?

Ans: hydrophobic amino acids are those with side chain that do not like to reside in an aqueous medium. There are some hydrophobic amino acids,

Alanine …. Valine …. Leucine …. Isoleucine …. proline …. Methionine … tryptophan.

Q.. What are pseudoknots?

Ans: when in RNA, the unpaired bases of 3´ end fold up to itself within tertiary structure, it may result in formation of abnormal folding called **pseudoknots.**

Q.. what is systematic biology? Also mention its applications.

Ans: systematic biology define as the scientific study of kinds and diversity of living organisms , and relationship among them.

There are some applications of it,

It helps us, to produce

classification in taxonomy in

nomenclature making databse

Q.. what are the needs of bioinformatics?

Ans: The need for bioinformatics is on a rapid rise as biological data is rapidly increasing and becoming available online, free of any cost. The number of online tools for processing genomics and proteomics information are rapidly increasing. This is just a reflection of the need for bioinformatics in modern day biology.

Q.. What is the method of clustering?

Ans: . The method is generally attributed to Sokal and Michener.

In this method two sequences with the shortest evolutionary distance between them are considered and these sequences will be the last to diverge, and represented by the most recent internal node. **Q.. Write down the types of RNA.**

Ans: There are many types of RNA according to their funtions like:

- Messenger RNA (mRNA)
- Transfer RNA (tRNA)

- Ribosomal RNA (rRNA)
- Micro RNAs (miRNA)
- Small Interfering RNA (siRNA)

Q.. differentiate between local and global alighnment. Ans:

GLOBAL ALIGHNMENT	LOCAL ALIGHNMENT
• Maximizing sequence matches over the entire length of two sequences, by introducing gaps.	• Find regions between two sequences that have the strongest similarity, excluding less similar regions.
• Used to determine overall similarity, conservation	• Finding similar domains, motifs, detecting distant homology
• maximizes the number of matches between the query and source sequences along the entire length of both the sequences.	• Local alignment - gives the highest scoring local match between both query and sequences.

Q.. what is the role of 5´ cap and poly A tail?

Ans: the 5´ end of molecule caped with 7 methyl guanosine tri phosphate and 3´ end poly A tail cap, both caps play major role to translate mRNA to ribosome, also protect mRNA.

1. The 5' cap helps identify mRNAs at the Ribosomes
2. The 5' cap also acts as a shield against 5' exonuclease, thus leading to an increase in mRNA stability
3. The 3'-end of mRNA has a polyA tail (around 30-200 adenylate residues) which help shield against 3' exonucleases

Q.. Write down the objectives of comparing sequence.

ANS: There are millions sequences on GenBank and UniProt what will happen if we will compare them? By comparing sequences of DNA, RNA and Proteins we can get

- Similarity among sequences
- There might be some specific difference due to some disease or mutation

  There may be some evolutionary relationship.

As there nucleotides can be similar or differ from each other. While UniProt is used in case of amino acids sequence comparison.

By comparison of nucleotides and Amino acids of any DNA, RNA and protein sequence we can find many evolutionary facts and relations among species. **Q. write the benefits of Dot Plot.**

Dot plots provides us the Global similarity between the two sequences and helps us to visualize the alignments of sequences and sequence repeats appear as diagonal stacks in plot.

Q. Differentiate between pairwise sequence alignment and multiple sequence alignment?

Ans: Alignment between two sequences by sliding across each other is known as pair wise sequence alignment while multiple sequence alignment involves the aligmnet of more than 3 sequences or multiple sequences sliding across each other from top to bottom. **Q. define open reading frame?**

In molecular genetics, an **open reading frame** (**ORF**) is the part of a **reading frame** that has the potential to code for a protein or peptide. An **ORF** is a continuous stretch of codons that do not contain a stop codon (usually UAA, UAG or UGA). **Q.. What is scoring matrics?**

To build the Scoring Matrices we analyze the amino acids and nucleotides which are substituted in single gene and protein sequence.

Scoring Matrices have both values +ve and –ve. Positive value for matches and negative value for mismatches.

Q.. Write the progressive alignment for MSA.

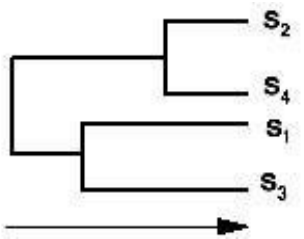
MSA involves progressive alignment of sequences. Doing so many progressive alignments can be slow.

STEPS:

- Step 1 : Pairwise Alignment of all sequences

Example: S_1, S_2, S_3, S_4, so that is 6 pairwise comparisons.

- Step 2: Construct a Guide Tree (Dendogram) using a *Distance Matrix.*



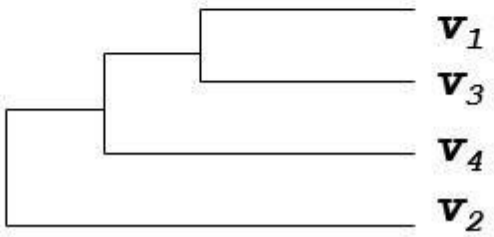
$$Distance = 1 - \frac{Matches}{Comparisons}$$

$$Similarity = \frac{Matches}{Comparisons}$$

- Step 3: Progressive alignment following branching order in tree.

$$Similarity = \frac{Matches}{Comparisons}$$

	v_1	v_2	v_3	v_4
v_1	–			
v_2	.17	–		
v_3	.87	.28	–	
v_4	.59	.33	.62	–



Q.. What is gibs free energy? What is the role of its in RNA formation?

"Gibbs Free Energy" is the free energy of an RNA molecule available for reaction. The smaller, the better!. RNA structure formation lowers the free energy **Energy of RNA structures**
RNA structures have the lowest (or close) quantity of free energy. In cases where RNA can take two structural forms, one can select the one with lower energy state.

Q.. What is the role of gibs energy in formation of secondary structure of RNA?

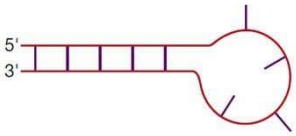
- 5 nucleotides formed H-Bonds
- This bond formation released energy (-12.0 kcal/mol)
- RNA molecule took up a 2' structure
- Hence became more stable

Q.. describe hair pin structure of RNA.

RNA 2' Structures

- The second 2' structure
- is the **hairpin loop**

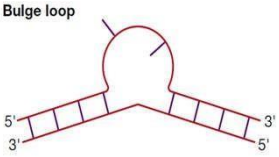
C. Stem and loop or hairpin loop.



RNA 2' Structures

- The loop of the hairpin must at least four bases long to avoid steric hindrance with base-pairing in the stem part of the structure.
- Note that hairpins reverse the chemical direction of the RNA molecule.

D. Bulge loop



Q.. what are bulges?

- The third type of 2' structure is the bulge loop

Q.. How are Bulges Formed?

- Bulges, are formed when a double-stranded region cannot form base pairs perfectly.

- Bulges can be asymmetric with varying number of base pairs on one side of the loop.

Q.. What is the use of Blastx, tBlastx and tBlastn?

- **Blastx:**

Compares a nucleotide query sequence against a protein sequence database.

Helps find potential translation products of unknown nucleotide sequences

- **tblastn:**

Compares a protein query sequence against a nucleotide sequence database

Nucleotide sequence dynamically translated into all reading frames

- **tblastx:**

Compares the six-frame translated proteins of a nucleotide query sequence against the sixframe translated proteins of a nucleotide sequence database. **Q.. What are the types of MS base proteomics?**

Q.. Differentiate between top down and bottom up proteomics.

- Bottom up proteomics measures the mass of peptides
- Peptides result after enzymatic digestion of precursor proteins
- Peptides are searched in protein databases

While,

- TDP measures the mass of intact proteins
- Mass of post translational modifications can be easily measured.

Q.. Briefly describe Expacy and its fields.

- Developed by Swiss Bioinformatics Institute (SIB)
- Website provides access to databases and tools
- Proteomics, Genomics, Phylogeny, Systems biology, Population genetics, transcriptomics etc are the basic fields of Expacy.

- Expasy provides access to a variety of online databases and tools.

- Depending upon your requirement, you find sequence information from Expasy.

Q.. Write down the method of contructing phylogenetic trees *Phylogenetics:*

"Phylogenetics is the study of evolutionary relationship and refers to creation of relationship trees between various species of bacteria, archaea and eukaryote".

Rooted and unrooted trees can be used to show phylogenetic relationship between sequence.

Phylogenetic analysis algorithms:

There are many types of algorithms which are basically classified into two classes,

1. Clustering approaches 2. Objective base method

Clustering ap proach	***Objective base method***
UPGMA	Least Square Distance
WPGMA	Maximum Likelihood
Neighbor joining	Maximum Parsimony
Single linkage	
Complete li nkage	

UPGMA:

UPGMA is a simple agglomerative method. Which is the abbreviation of `Unweighted pair group Method with Arithmetic means`.

In this method two sequences with the shortest evolutionary distance are considered and last to diverge, and represented by the most internal node.

Least Square Distance Method:

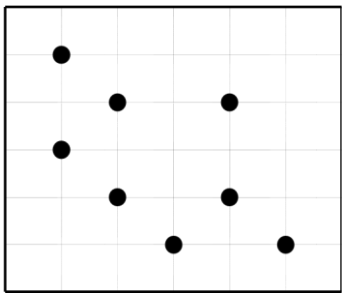
It is often used to represent the "Observed" distance between sequence. Applicable branch length.

Q.. How can we modify the Dot Plot to consider Gap and Mismatch?

- Dot plots cannot capture insertions, deletions and gap indications. How can we deal with them? We Modify the dot plot by this method.

Matches are labelled by +1 instead of a dot
Indels (gaps) and substitutions (mutations) can be included in the dot plot as -1.

	A	C	G	C	G
A	1	-1	-1	-1	-1
C	-1	1	-1	1	-1
A	1	-1	-1	-1	-1
C	-1	1	-1	1	-1
G	-1	-1	1	-1	1



Q.. what are matrics? Also mention its types

1. .Ans: Alignment matrices scoring is very useful method to score the sequences alignment for match and mismatch. There are two types of scoring matrices. PAM and BLOSUM .

Q.. write down the steps of PAM matrices.

2. Align the protein sequence which are 1-PAM Unit diverge.
3. Let Ai,i be the number of times Ai is substituted by Ai.
4. Compute the frequency fi of amino acid Ai.

Then, PAM1=pii= $\frac{A_{ij}}{\sum_k A_{ik}}$

PAM 'n'= $(PAM1)^n$

Q.. what is BLOSSOM matrice? Also mention its steps.

BLOSUM matrices is also called the Block substitution matrix without any gap although it has mismatches in sequences.

There are three steps to compute the BLOUSM Matrices.

Step 1: Eliminate sequences that are identical in x% positions

Step 2: Compute observed frequency $f_{i,j}$ of aligned pair A_i to A_j. Hence, $f_{i,j}$ becomes the probability of aligning A_i and A_j in the selected blocks.

Step 3: Compute f_i which is the frequency of observing A_i in the entire block.

Q… Describe the 3D-1D Bowie algorithm.

Q… Differentiate between fingerprinting and short gun.

Q… What are substitution and indel?

Q… Why RNA has short life span?