**NAME: AUGOYE OMESIRI**

**DEPARTMENT: ANATOMY**

**MATRIC NO: 18/MHS01/094**

**COURSE: CSC262**

1a

* Sensitivity
* Specificity
* Speed

1. Sensitivity is the ability to find as many correct hits as possible. It is measured by the extent of inclusion correctly identified sequence of the same family. These correct hits are considered “true positives” in the database searching exercise.
2. Specificity refers to the ability to exclude incorrect hits. These incorrect hits are unidentified sequence mistakenly in database searching and are considered “false positives”.
3. Speed is the time it takes to get a result from a database search. Depending on the size of the database speed can be a primary concern. Ideally, one wants to have the greatest sensitivity, specificity and speed in a database.

1b

BLAST: basic local alignment search tool is a algorithm and program for comparing primary biological sequence information, such as the amino acid sequence of protein and/ or nucleotides of DNA and / or RNA sequences.

2a

|  |  |
| --- | --- |
| PAM | BLOSUM |
| They are used to score alignment between closely related protein sequences. | They are used to score alignment between evolutionary diverge protein sequences. |
| Based on global alignment mutations global alignment are very significant. | Based on local alignment, based on highly conserved stretches of alignment. |
| Examples are: PAM250 is used for more distant sequences than | BLOSUM 80 is more used for closely related sequences than |
| PAM 120. | BLOSUM 62 |

2b

Heuristic is a technique designed for solving a problem more quickly when classic method are too slow, or for finding an approximate solution when classic method fail to find any exact solution.

3a

* Sequence homology is the biological homology between DNA, RNA or protein sequences, defined in term of shared ancestry in the evolutionary history of life.
* Sequence similarity is a measure of an empirical relationship between sequences. A common objective of sequence similarity calculations is establishing the sequence homology.
* Sequence identity is a way to measure the similarity brtween two sequences.

3b

1. DOT matrix method
2. Dynamic programming method
3. Word method

ii. Pairwise sequence alignment is used to identify regions of similarity that may indicate functional, structural and evolutionary relationships between two biological sequences protein and nucleic acid.

4a

|  |  |
| --- | --- |
| GLOBAL ALINGMENT | LOCAL ALINGMENT |
| Contains all letters from both the query and target sequences. | Aligns a substring of the query sequence to a substring of target sequence |
| Suitable for aligning two closely relate sequence. | Suitable for aligning more divergent sequence or distantly related sequences. |
| A global technique is the NEEDLE MAN WUNSCH ALGORITHM. | A local technique SMITH WATERMAN ALGORITM. |
| EXAMPLES: EMBROSS NEEDLE  NEEDLE MAN WUNSCH | BLAST  EMBROSS WATER |

4b

|  |  |
| --- | --- |
| SIMILARITY | HOMOLOGY |
| It does not imply homology | It usually implies similarity |
| It refers to the likeness or % identity between two sequence | It refers to share ancestry |

|  |  |
| --- | --- |
| SIMILARITY | IDENTITY |
| It does not depends on their identity | It depends on their identity |
| It shows the extent to which residues is aligned | Gaps are not counted ad the measurement is relational to the shorter of the two sequences. |