**NAME: OGUNBIYI SOFIYYAH MODUPE**

**DEPARTMENT: PHYSIOLOGY**

**COLLEGE: MEDICINE AND HEALTH SCIENCES**

**MATRIC NO: 18/MHS02/129**

**COURSE CODE: CSC 262**

**COURSE TITLE: BIOINFORMATICS**

**SOLUTIONS TO THE GIVEN ASSIGNMENT**

**1ai. Keyword/ Basic search:** It searches one field and useful to determine how big the collection of articles related to the topic is. It is also helpful when doing a known item search, for example; when the exact file is known.

 **a ii. Advance search:** It is a move refined search. Multiple search terms in multiple fields can be searched for simultaneously. It is used when the collection of interested article is large and needs to be narrowed.

**1b. BLAST:** Is an algorithm and program for comparing primary biological sequence information such as the amino acid sequence of protein or the nucleotides of the DNA/RNA sequences. It enables researchers to compare a subject protein or nucleotide sequence(called a query) with a database of sequences and identify library sequence that resemble the query sequence above a certain threshold.

**2a.** **PAM:**

* Are used to score alignments between closely related protein sequence
* Based on global alignments
* Alignment have high similarity than BLOSUM
* Mutation in global alignment are very significant
* Higher numbers in the PAM matrix naming denotes greater evolutionary distance

**BLOSUM:**

* Are used to score alignment between evolutionary divergent protein sequences
* Based on local alignment
* Alignment have low similarity than PAM
* Based on highly conservered stretches of alignments
* Higher numbers in the BLOSUM matrix naming denotes higher sequences similarly and smaller evolutionary distance.

 **2b. Heuristic Database Searching:** It refers to a search strategy that attempts to optimize a problem by iteratively improving the solution based on a given heuristic function or a cost measures. It does not always guarantee to find an optimal or the best solution, but may instead find a good or acceptable solution within a reasonable amount of time and memory space.

**3ai.Sequence Homology:** Is the biological homology between DNA, RNA or protein sequence defined terms of shared ancestry in the evolutionary history of life.

 **aii.Sequence Similarity:** Is a measure of an empirical relationship between sequences.

 **aiii.Sequence Identity:** Is the amount of characters which match exactly between two different sequences.

**3b. Methods of Alignment Algorithm**

* Dot matrix
* Dynamic Programming
* Word Alignment Algorithm

**3bii. Pairwise Sequence Alignment:** Is used to identify region of similarity that may indicate functional, structural and evolutionary relationship between two biological sequence( protein or nucleic acid)

**4a.** The difference between Local Alignment and Global Alignment is that; in local alignment, you may try to match your query with a substring(a portion of your subject) while in global alignment, you perform an end to end alignment with the subject.