[Date]

**Shehab Mohamed**

**Task-2**

**DNA Sequence Analysis Assignment**

**Objective**

To identify all Open Reading Frames (ORFs) in a DNA sequence and filter them based on a minimum length.

**Implementation**

find\_orfs: Detects all ORFs in a sequence based on start and stop codons.

filter\_orfs: Filters ORFs shorter than a specified length.

**Example**

Input DNA: "ATGCGATAACTGAATGCGTTAATAG"

Output ORFs: [(0, 12), (13, 22)]

Filtered ORFs (length >= 6): [(0, 12)]

**Difference Between FASTA and FASTQ**

FASTA: Stores sequence data (DNA/protein) with headers.

FASTQ: Stores sequence data along with quality scores for sequencing.

**MutableSeq vs Seq**

MutableSeq allows in-place modifications of DNA sequences, unlike `Seq`, which is immutable.

**Example**

Input: "ATGCGA" -> Modify position 1 to "T" -> Output: "ATTGCA"

**Dot Plot:**

**Objective**

Visualize the similarity between two DNA sequences.

**Example**

Input: Sequence1 = "ATGCGA", Sequence2 = "ATGCTA"

Output: Dot plot visualization with matching characters as dots.

**Test Cases**

See the `test\_cases.py` file for automated testing scenarios.

**References:**

* Mount, D. W. (2004). *Bioinformatics: Sequence and Genome Analysis*. Cold Spring Harbor Laboratory Press.
* Cock, P. J. A., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., ... & de Hoon, M. J. (2009). *Biopython: Freely Available Python Tools for Computational Molecular Biology and Bioinformatics*. Bioinformatics, 25(11), 1422-1423.