
Gene finding X66236_en

Recall that in prokaryotic genomes, the sequence coding for a protein occurs as one contiguous open reading frame, and that an open reading frame begins with the start codon ATG (methionine) in most species and ends with a stop codon TAA, TAG, or TGA.

For example, the DNA sequence of Bacteriophage ϕ-X174, which was the first genome to be sequenced, has 117 open reading frames (11 of which are protein coding genes) within a circular single strand of 5,386 nucleotides.

Write code for the gene finding problem. The program must implement and use the GENE-FINDING function in the pseudocode discussed in class, which is iterative and is not allowed to perform input/output operations. Make one submission with Python code and another submission with C++ code.

Input

The input is a string s over the alphabet $\Sigma = \{A, C, G, T\}$.

Output

The output is a minimal substring of s (an open reading frame) from a start codon to a stop codon.

Sample input 1

```
GGTTTCTATGTGGCTAAATACGTTAACAAAAAGTCAGATATGGACCTTGCTGCTAAAGGTCTAGGAGCTAAAGAATGGAA
```

Sample output 1

```
ATGTGGCTAAATACGTTAACAAAAAGTCAGATATGGACCTTGCTGCTAAAGGTCTAG
```

Problem information

Author: Gabriel Valiente

Generation: 2026-01-25T17:35:23.538Z