

The Central Dogma of Molecular Biology project comprises several classes designed to simulate and analyze various biological processes, including transcription, translation, and protein analysis. These classes interact to provide a comprehensive understanding of genetic information flow and protein synthesis. I am really interested in the field of bioinformatics, and I wanted to see if I can develop a modularized skeleton model of what undergoes every single organism on this planet in code. In the true biological version, there are enzymes and proteins that facilitate this process, but consider the computers processors and the execution of code as the enzymes. Below is a summary of what each part of the program does.

Transcription Class

The Transcription class serves as the entry point for converting DNA sequences into mRNA lists. It reads a DNA file, removes extraneous characters, and transcribes the DNA sequence into mRNA by replacing thymine (T) with uracil (U). Additionally, it provides methods to calculate the molecular weight of the sequence, determine GC content, and retrieve the sequence itself.

RNA Class

The RNA class extends the Transcription class by representing mRNA sequences as objects. It takes an mRNA list and converts it into a tuple of codons, facilitating subsequent translation processes. The class enables easy manipulation and analysis of mRNA sequences.

Protein Class

The Protein class represents the final step by translating mRNA sequences into protein chains and performing analyses on the resulting proteins. It translates codon dictionaries into peptide chains, calculates protein molecular weights, assesses absorbance at specific wavelengths, and determines the net negative charge of proteins. This class provides valuable insights into protein composition and characteristics about that protein.

The Central Dogma of Molecular Biology project utilizes the power of object-oriented programming to simulate and analyze a common genetic process. By implementing classes for transcription, RNA representation, and protein analysis, the project offers a beginner's toolkit for a broad view of characteristics about genes and the proteins that they encode for.

