

A Protein-Wave Alignment Tool for Frequency Related Homologies Identification in Polypeptide Sequences

Authors : Victor Prevost, Solene Landerneau, Michel Duhamel, Joel Sternheimer, Olivier Gallet, Pedro Ferrandiz, Marwa Mokni

Abstract : The search for homologous proteins is one of the ongoing challenges in biology and bioinformatics. Traditionally, a pair of proteins is thought to be homologous when they originate from the same ancestral protein. In such a case, their sequences share similarities, and advanced scientific research effort is spent to investigate this question. On this basis, we propose the Protein-Wave Alignment Tool ("P-WAT") developed within the framework of the France Relance 2030 plan. Our work takes into consideration the mass-related wave aspect of protein biosynthesis, by associating specific frequencies to each amino acid according to its mass. Amino acids are then regrouped within their mass category. This way, our algorithm produces specific alignments in addition to those obtained with a common amino acid coding system. For this purpose, we develop the "P-WAT" original algorithm, able to address large protein databases, with different attributes such as species, protein names, etc. that allow us to align user's requests with a set of specific protein sequences. The primary intent of this algorithm is to achieve efficient alignments, in this specific conceptual frame, by minimizing execution costs and information loss. Our algorithm identifies sequence similarities by searching for matches of sub-sequences of different sizes, referred to as primers. Our algorithm relies on Boolean operations upon a dot plot matrix to identify primer amino acids common to both proteins which are likely to be part of a significant alignment of peptides. From those primers, dynamic programming-like traceback operations generate alignments and alignment scores based on an adjusted PAM250 matrix.

Keywords : protein, alignment, homologous, Genodic

Conference Title : ICBBE 2022 : International Conference on Bioinformatics and Biological Engineering

Conference Location : Paris, France

Conference Dates : October 27-28, 2022