Guião Laboratórios de Bioinformática

Step 1 | General and Theoretical Research

The project began with detailed general and theoretical research on the chosen gene, TNF, which was essential for establishing a solid knowledge base about the mentioned gene. Platforms such as NCBI (National Center for Biotechnology Information) and PUBMED were used to access a wide range of information. The research contained data on the structure and function of the TNF gene, the different alleles associated with it, and the diseases related to its mutations.

Step 2 | Obtaining the 3D Image of the Protein

Regarding the Protein Data Bank (PDB) on the NCBI platform, it was possible to obtain the three-dimensional image of the protein associated with the TNF gene. As a result, this allowed us to visualize specific details, such as how different regions of the protein are organized in space.

Step 3 | Selection of Homologous Sequences

Through the NCBI platform, we obtained both the protein sequence of the TNF gene and selected 10 homologous protein sequences from various species.

The selection was crucial for comparing analyses (such as the phylogenetic tree, which will be discussed later). The sequences were obtained in FASTA format and grouped into a single text file, forming a set of 11 homologous sequences, which served as the basis for subsequent analyses.

Step 4 | Multiple Sequence Alignment (MSA)

The text file containing the 11 homologous sequences, was uploaded to the Clustal Omega platform, which processed the sequences and generated a Multiple Sequence Alignment (MSA). As a result, it was possible to identify conserved regions among the different species. Additionally, the platform produced an initial phylogenetic tree in its corresponding text format.

Step 5 | Phylogenetic Tree

The phylogenetic tree generated in Clustal Omega in text format was subsequently uploaded to the iTOL (Interactive Tree Of Life) platform. This tool allowed us to visualize the phylogenetic tree in different formats: rectangular, circular, and unrooted. In addition, we used iTOL's features to edit and highlight

the representations of the phylogenetic trees, enhancing important characteristics and improving the clarity of the data presentation.

Step 6 | Identification of Motifs

Using the Clustal Omega platform, we uploaded the file with the 11 homologous sequences and subsequently downloaded the alignment in FASTA format converted by Seqret, which was then uploaded to the WebLogo platform. Consequently, we obtained 2 motifs, one from positions 1-50 and another from 50-100.

Step 7 | Identification of Regulatory Elements

Trough the Genome Browser site, we searched for regulatory elements of the TNF gene, which was crucial for identifying DNA regions that control gene expression. The Genome Browser allowed us to visualize genomic data in a broader context, providing information on promoters, enhancers, and other regulatory elements that can influence the activity of the TNF gene.

Afonso Santos | *Carolina Jesus* | *Rafael Santos*

Link para o website:
https://sites.google.com/view/tnf-gene/tnf