

# Algoritmos para Bioinformática / Bioinformática

## 2021/2022

### *Aula 8 – Sequence Search - Blast*

#### **Task 1 – Complete the code**

1. Revise the code from the slides and complete the methods for the *MyBlast.py* class.
2. Describe the meaning for each of the parameters in the function *best\_alignment*?
3. Write a test function for the query sequences in *query1.fasta* and *query2.fasta* that finds the most similar sequence in *seqBlast.txt* (use as identifier the number of the sequence in the file). Print the respective score of the best alignment.
4. Develop a similar test function as in 3) with the *glyco\_sequences*, using the query and the *db* for the search of the most similar sequence. For the most similar sequence perform the global and local alignment and provide as output of the function.

#### **Task 2 – Code extension**

5. Consider the function *get\_hits* provided in the slides. Create a variant of this function that given an additional parameter mismatch *mismatch*, allows at most *mismatch* characters to be different between the query and the sequence words. Test that *mismatch* is smaller than the length of the word.
6. Adapt the appropriate functions to allow returning a ranking of the best alignments and not only the sequence that scores highest.