

## Algoritmos Bioinformática /Bioinformática 2021/2022

### Exercícios Aula 3

For this exercise, there is no need to import other modules! Use only the code provided in classes.

1. Write the functions that are to be *all\_orfs()* and *all\_orfs\_ord()* in file *sequence\_functions.py* and test them properly by using the different test function in the main method.
2. Consider the DNA sequence *genomic\_dna.fa* provided in the exercise. Write a function, that may call other functions from *sequence\_functions.py*, in order to obtain the following information.
  - 2.1 Calculate the longest protein sequence that can be found in the DNA sequence.
  - 2.2 Indicate how many possible proteins are coded from the DNA sequence and the respective frame ((+1, +2, +3 or -1, -2, -3) from each these were coded. Write the proper code to implement this and add the result as a comment.
3. Write the function called *read\_fasta\_2dictionary*. The function should load the dictionary where the keys are the sequence identifiers and the values are the sequences. The output should be dictionary. Use the file PS000727.fasta to test. For the identifier use only the word after the ">" and before the first blank space. For instance, in the following header the identifier should be sp|P51173|APEA\_DICDI.  
  
>sp|P51173|APEA\_DICDI DNA-(apurinic or apyrimidinic site) lyase OS=Dictyostelium discoideum OX=44689 GN=apeA PE=2 SV=2
4. Complete the code for the class Phone Book in the file PhoneBook.py. The class should have as attributes of names and respective phone numbers. Implement the following methods:
  - Constructor (init)
  - Add phone entry
  - Search by name
  - Search by phone number
  - Print the entire book
  - Define a method to copy the phone book information.
  - Test the class with multiple instances and calling the different methods.

Complete the code in file PhoneBook.py and submit in Moodle.