# *Multiple Sequence Alignment*

*In this class, we will talk about the generalization of the sequence alignment problem to three or more sequences. We will show how this can be performed with an implementation of a progressive alignment strategy. Finally, an implementation of a simple version of MSA algorithm will be performed.*

**Learning Objectives**

- MSA is achieved by performing editing operations over the sequences to maximize the number of similar characters per column of the alignment.
- The complexity of MSA increases with relation to the number of sequences to align.
- Efficient and clever approaches are required to solve the MSA problem.

1. Review the slides “Multiple Sequence Alignment”.

Quiz

***Task 1 – Perform Multiple Sequence Alignment***

- *Go to Clustal Omega website https://www.ebi.ac.uk/jdispatcher/msa/clustalo.*
- *Upload the sequences in the provided file msa_test.fas*
- *Explore the alignments with colors and the trees.*
- *What is the meaning of the symbols : * and space in the bottom of the alignment?*
- *Of the four sequences, which one are more related and which one is the most different?*

***Task 1 – Complete the code***

1. *In the class MyAlign complete the function Consensus.*

2. *Test all the examples in each of the classes, except the MultipleAlignment class.*

***Task 2 – Complete class MultipleAlignment***

3. *Write a method ScoreColumn(self, charsCol) to add to the MultipleAlignment class that allows to calculate the score of each column in the alignment (charsCol is a list of characters from the column of the alignment that can be retrieved using the method column from the class MyAlign). The score is calculated using the Sum of Pairs (SP) approach, i.e. the score will be the sum of the scores of each pair of characters in the alignment. If two gaps are found in each pair then the score will be zero.*

4. *Using the existing methods, namely ScoreColumn, develop a new method called scoreSP(self, alignment) that returns the score of sum of pairs from a complete alignment.*