

# Algorithms for Bioinformatics - lect 6-7

Francesco Penasa

March 16, 2020

2020 03 16

## 1 Previous lectures

1. Global sequence alignment
2. Local sequence alignment
3. substitution matrices (PAM and BLOSUM)

## 2 Issue about the complexity

**Needleman-Wunsch** time

$$O(mn)$$

space

$$O(n)$$

$m$  and  $n$  are the length of the two sequences

**Smith-Waterman** time

$$O(mn^2)$$

$n^2$  because gaps of variable length are considered during the computation of each element of the matrix.  
It is possible to reduce Local alignment in  $O(mn)$

Is the complexity low enough for practical uses? No. Use Heuristics methods, sacrifice exact solution in order to gain in terms of time complexity.

## 3 FASTA

FASTA is an heuristic method for global/local sequence alignment.

1. Look-up table: between the sequences
2. Best initial regions found: through the ktup matrix
3. Connect the best initial regions heuristically
4. Computing an exact alignment limiting the search to a promising area

<https://www.ebi.ac.uk/Tools/sss/fasta/>

## 4 BLAST

BELLO