

Needleman-Wunsch

- Sequence alignment algorithm
- “How different/similar are two DNA/RNA-sequences?”
- Gives optimal alignment, but there can be multiple ones
- Use of dynamic programming

- Input: Two sequences a and b
- $|a| = m$ ;  $|b| = n$
- Data structure: A matrix with  $m * n$
- Runtime and space complexity:  $O(m*n) \sim O(n^2)$

	$\epsilon$	$b_0$	$b_1$		$b_j$		$b_n$
$\epsilon$	0	1	2		$D_{0,j}$	...	$D_{0,n}$
$a_0$	1						
$a_1$	2						
$a_3$	3						
	...						
$a_i$	$D_{i,0}$						
	...						
$a_m$	$D_{m,0}$						$D_{m,n}$

$$D_{i,j} = \begin{cases} D_{i-1,j} + w(a_i, -) \\ D_{i-1,j-1} + w(a_i, b_j) \\ D_{i,j-1} + w(-, b_j) \end{cases}$$



$$D_{0,j} = D_{0,j-1} + w(-, b_i)$$



$$D_{i,0} = D_{i-1,0} + w(a_i, -)$$

# Cost-function

$$w(a, b) = \begin{cases} 0, & a = b \\ 1, & a \neq b \end{cases}$$

$$w(a, -) = 1$$

$$w(-, b) = 1$$

# Example

	<b>€</b>	<b>A</b>	<b>C</b>
<b>€</b>	0	1	2
<b>A</b>	1		
<b>G</b>	2		
<b>C</b>	3		

$$D_{i,j} = \begin{cases} D_{i-1,j} + w(a_i, -) \\ D_{i-1,j-1} + w(a_i, b_j) \\ D_{i,j-1} + w(-, b_j) \end{cases}$$

# Example

	<b>€</b>	<b>A</b>	<b>C</b>
<b>€</b>	0	1	2
<b>A</b>	1	0	
<b>G</b>	2		
<b>C</b>	3		

$$D_{i,j} = \begin{cases} D_{i-1,j} + w(a_i, -) \\ D_{i-1,j-1} + w(a_i, b_j) \\ D_{i,j-1} + w(-, b_j) \end{cases}$$

# Example

	<b>€</b>	<b>A</b>	<b>C</b>
<b>€</b>	0	1	2
<b>A</b>	1	0	1
<b>G</b>	2	1	1
<b>C</b>	3	2	1

$$D_{i,j} = \begin{cases} D_{i-1,j} + w(a_i, -) \\ D_{i-1,j-1} + w(a_i, b_j) \\ D_{i,j-1} + w(-, b_j) \end{cases}$$



# Example — Traceback

	€	A	C
€	0	1	2
A	1	0	1
G	2	1	1
C	3	2	1

$$D_{i,j} = \begin{cases} D_{i-1,j} + w(a_i, -) \\ D_{i-1,j-1} + w(a_i, b_j) \\ D_{i,j-1} + w(-, b_j) \end{cases}$$

$$D_{3,2} = 1 \neq D_{3-1,2} + 1$$



$$D_{3,2} = 1 \neq D_{3,2-1} + 1$$

$$D_{3,2} = 1 = D_{3-1,2-1} + 0$$

# Example — Traceback

	€	A	C
€	0	1	2
A	1	0	1
G	2	1	1
C	3	2	1

$$D_{i,j} = \begin{cases} D_{i-1,j} + w(a_i, -) \\ D_{i-1,j-1} + w(a_i, b_j) \\ D_{i,j-1} + w(-, b_j) \end{cases}$$

 Insert gap in b  
 Insert gap in a

Alignment:    A G C  
                   A - C

# List of references

- Lecture Bioinformatics I, Number 3, Dr. Dominic Rose. [http://www.bioinf.uni-freiburg.de//Lehre/Courses/2014\\_SS/V\\_Bioinformatik\\_1/lecture3-1.pdf](http://www.bioinf.uni-freiburg.de//Lehre/Courses/2014_SS/V_Bioinformatik_1/lecture3-1.pdf)  
Date: 07/11/2014