Metódy v bioinformatike CB #3 Zarovnávanie sekvencií

Jana Černíková

FMFI UK

9/10/2025

Globálne zarovnanie

Uvažujme skórovanie zhoda +3, nezhoda -1, medzera -2 Reťazce TAACGG a CACACT

$$A[i,j] = \max \left\{ egin{array}{l} A[i-1,j-1] + s(x_i,y_j), \ A[i-1,j] - 2, \ A[i,j-1] - 2 \end{array}
ight.$$

$$s(x_i, y_j) = 3$$
 ak $x_i = y_j$,
 $s(x_i, y_j) = -1$ ak $x_i \neq y_j$

$$A[i,0] = -2i$$
$$A[0,j] = -2j$$

Globálne zarovnanie

		С	Α	С	Α	C	Т
	0	-2	-4	-6	-8	-10	-12
Т	-2						
Α	-4						
Α	-6						
С	-8						
G	-10						
G	-12						

Globálne zrovnanie

C A C A C 0 0 -2 -4 -6 -8 -10 1 T -2 -1 -3 -5 -7 -9 2 A -4 -3 2 0 -2 -4	6	5	4	3	2	1	0		
0 0 -2 -4 -6 -8 -10 1 T -2 -1 -3 -5 -7 -9 2 A -4 -3 2 0 -2 -4	Т	С	Α	С	Α	С			
1 T -2 -1 -3 -5 -7 -9 2 A -4 -3 2 0 -2 -4	-12	-10	-8	-6	-4	-2	0		0
2 A -4 -3 2 0 -2 -4	-7	-9	-7	-5	-3	-1	-2	Т	1
	-6	-4	-2	0	2	-3	-4	Α	2
3 A -6 -5 0 1 3 1	-1	1	3	1	0	-5	-6	Α	3
4 C -8 -3 -2 3 1 6	4	6	1	3	-2	-3	-8	С	4
5 G -10 -5 -4 1 2 4	5	4	2	1	-4	-5	-10	G	5
6 G -12 -7 -6 -1 0 2	3	2	0	-1	-6	-7	-12	G	6

CACACT-

TA-ACGG

alebo

CACAC-T

TA-ACGG

Lokálne zarovnanie

Uvažujme skórovanie zhoda +3, nezhoda -1, medzera -2 Reťazce TAACGG a CACACT

$$A[i,j] = \max \left\{ egin{array}{l} 0, \ A[i-1,j-1] + s(x_i,y_j), \ A[i-1,j] - 2, \ A[i,j-1] - 2 \end{array}
ight.$$

$$s(x_i, y_j) = 3$$
 ak $x_i = y_j$,
 $s(x_i, y_j) = -1$ ak $x_i \neq y_j$

$$A[i,0] = 0,$$

$$A[0,j] = 0$$

Lokálne zarovnanie

		С	Α	С	Α	С	Т
	0	0	0	0	0	0	0
Т	0						
Α	0						
Α	0						
С	0						
G	0						
G	0						

Lokálne zarovnanie

		0	1	2	3	4	5	6
			С	Α	С	Α	С	Т
0		0	0	0	0	0	0	0
1	T	0	0	0	0	0	0	3
2	Α	0	0	3	1	3	1	1
3	Α	0	0	3	2	4	2	0
4	С	0	3	1	6	4	7	5
5	G	0	1	2	4	5	5	6
6	G	0	0	0	2	3	4	4

 ACAC

A-AC