

Aplicarea algoritmilor Needleman-Wunsch și Smith-Waterman în alinierea secvențelor de aminoacizi

Procedura de calcul este la fel precum cea pe care am discutat-o și exersat-o la ulitmul curs/ LP față-în-față. Diferențele constau în utilizarea matricei PAM 250 (sau a altor matrici de substituție matrici similare) pentru calculul valorilor specifice pentru *match* și *mismatch*. În cazul secvențelor de aminoacizi (AA), pentru alinierea Smith-Waterman este permis *mismatch* la capete; în cazul secvențelor de nucleotide, nu sunt posibile *mismatch*-uri terminale.

1. Aliniere globală Needleman-Wunsch (NW)

Secvențe de AA:

S1 = YTIVW

S2 = YTTW

Sistem de scoruri: {PAM250; gap = -2}

		Y	T	I	V	W
	0	-2	-4	-6	-8	-10
Y	-2	10	8	6	4	2
T	-4	8	13	11	9	7
T	-6	6	11	13	11	9
W	-8	4	9	11	9	28

YTIVW

||*||

YTT_W

10 3 0 -2 17 Sc.Sim.A = +28

YTIVW

|||*|

YT_TW

10 3 -2 0 17 Sc.Sim.A = +28

Se observă că, pentru alinierea celor doua secvențe de aminoacizi cu sistemul de scoruri de mai sus, se obțin două aliniamente optime, cu același scor, adică **28**. Astfel de situații sunt rare în practică, dar nu imposibile, motiv pentru care sper să găsiți util acest exemplu.

2. Aliniere locală Smith-Waterman (SW)

Secvente de AA:

S1 = YTIVS

S2 = YTTS

Sistem de scoruri: {matricea PAM250; gap = -2}

		Y	T	I	V	S
	0	0	0	0	0	0
Y	0	+10	+8	+6	+4	+2
T	0	+8	+13	+11	+9	+7
T	0	+6	+11	+13	+11	+10
S	0	+4	+9	+11	+12	+13

YTIVS

||*|

YTT_S

10 3 0 -2 2 Sc.Sim.A = +13

YTIVS

||*|

YT_TS

10 3 -2 0 2 Sc.Sim.B = +13

YTI

||*

YTT

10 3 0 Sc.Sim.C = +13

YT

||

YT

10 3 Sc.Sim.D = +13

Se observă că, pentru alinierea celor două secvențe de aminoacizi cu sistemul de scoruri de mai sus, se obțin patru aliniamente optime cu același scor, adică **+13**.

În principiu, *software*-le care implementează algoritmi NW și SW raportează doar un singur aliniament optim, chiar dacă există mai multe astfel de aliniamente cu același scor. Un motiv în plus să deprindem modul în care “gândesc” acești algoritmi de aliniere.

Vă propun următoarele exerciții (pentru feed-back):

1. Aplicarea alinierii NW pentru secvențele de nucleotide:

S1: CGAATA

S2: CATAT

Sistemul de scoruri: {M = +2; mismatch = -1; gap = -3}

2. Aplicarea alinierii SW pentru secvențele de aminoacizi:

S1 = YTIVW

S2 = YTTW

Sistem de scoruri: {PAM250; gap = -5}

Reiterez recomandarea de la curs – după ce calculați aliniamentele optime cu ajutorul algoritmilor, vă rog să calculați scorul fiecărui aliniament și intuitiv, adică fără table NW și SW. Dacă lucrăm corect, cele două scoruri de similaritate/aliniere trebuie să fie identice.

Vă stau la dispoziție pentru orice întrebări/nelămuriri legate de utilizarea algoritmilor NW și SW. În cazul în care observați erori în exemplele prezentate, vă rog să mă atenționați.

MULT SUCCES!