

Comparação entre o algoritmo de Smith-Waterman e Needleman-Wunsch

O algoritmo Smith-Waterman foi desenvolvido por T.F. Smith e M.S Waterman é uma adaptação do algoritmo de Needleman-Wunsch utilizada para se obter o alinhamento local entre sequências, isto é, o alinhamento ótimo entre subsequências.

É considerada a técnica mais popular utilizada para alinhar sequências, pois permite buscar similaridades em partes da sequência - diferente do alinhamento global que busca sempre alinhar a sequência como um todo e possui complexidade $O(m*n)$. É um bem útil quando a similaridade entre sequências se estende por toda sua extensão. No caso dos exemplos de sequências disponibilizadas, todas as sequências são referentes a glicoproteína da proteína S (Spike) da família dos Coronavírus em diferentes espécies.

- Needleman-Wunsch

```
-----MLFVFIL-FLPSCL-GY--IG-DF--R--CIQT-VNYNGN--NASA-PSISTEA-VDVSK--GLGT
Y-YVL-D-RVYL-NA-TLLLTGY---YPV---D-GS-NYRNL-ALTGTN-TL-SLTWFKPP---FLSEF-N
DGI--FAK---VQ-----NLKT-N---TP-TGAT--SY-FPTIVIGSLF--GN-T-----SYT-VVLE
P--Y-NNIIM-A-SVCTYTIC-QL-P----YTPCKPNTNGNR-VIGF--WHTDVKPP-I--CLLKR--NFT
FNVNAP-WLY--F-HFYQ-QGG-TFY-AY-Y--A-DKP--S--A-T-TF-----LF-S-VY--I--GDILT
QYF-VLP-F--I---CT--P-T--A--G--STLLPLYWVTPL-L-K-RQ---YLFNFN-EKGV-IT-S-AV
DCA-SSYISEIK-CKT-QSL-LPSTGVYDLSGYT-VQP-V-GVVYRRVP-NLPDCKIE-E-WLTA-KSVPS
P-L-NWERR-TF-QNCNFNLSLLR--YVQAESLSCNNID-ASKVYGM--CFGs-VSVDK-FAI-PR-S-R
QIDLQIGNSG-FL-QTA-NYKIDTA-AT-SCQLYY---SLPKNN-VTINNYPSSWNRRYGFNDAGVFGKS
KHDVAYAQQCFIVRPSYCPCAQPDIVSACTSQT-KPMSAYCPTGTIHRCSLWNGPHLRSARVSGS-TYT-
C-EC-TCKPNP--F-DT--YDL-RCG-QIKT-IVNVGDHCEGLG-VL-EDKCGNSDPHKGC-S-CAHDSF-
IGWSHD-TCLVNDHSQIFANILLNG-INS GTTCST-D-LQL-P-NTEVAT--GV-CVRYDLYGITGQGVFK
E---VKADY-YNWQA-LLYDV-NGNLN-GFR--DLTT--N-KTYT-IRSC-YS--GRVSA-AY-HKEAPE
P--AL-LYR-NI-NCSYV-F-T-NN-ISRE-ENP-L-----NY--FDS-YLGCVVN-ADNRPD---EALPN
CD-LR-MG-AGL-CV-DYSK-SRRAR---RSV-ST-G-Y-RL-T-TFEPYMPML-VNDSVQSVG-GLYEMQ
IPTNFTIGHH--EEFIQ--IRAPKVTIDC-AA-FVC-G-DNAACRQQLV-EYGSFCDNV-NAILNE-VNNL
-L-DNMQL-QV-ASA-LM--QGVT-ISSRLPDGISGPIDDINFSPLL-G-CIGS-TCAEDGNGPS-AMRGR
SAIEDLLFDKVK-LSD-VGFVEAYNNCT---G-GQEV RDLLCVQ--SFNG-IKVLPPVLSE-S-QISGY-T
A--GA-TAAAMFPP-W-T--AA-AGVPFSL-NV-QYRINGLGVTMNVLSENQKMIASA-FNNALGAIQ-EG
F-DATNSAL-GKIQSVVNA-NAEALNNL-LNQLSNR-FGAISASLQ-EILT-RLDR-VEAK-AQIDRLING
RLTALNAYISKQ-L-SDSTLIK-FSA--AQAIKVNKCVKSQTT-R-INFCGNGNHILS-LVQNAP-YGLC
-FIHF-SYV-PTS-FKTANVSP--GLCI-SGD--RG-LAPKA-GYFVQD-NG-----EWKFTGSNYYY-PE
```

PITDKN-SV--VMISCA--VNY----TKA-PE--VFLNNSIP-NLPDFKEELDKWFKN-QTSIAPDL-SL-
DFEKLNVTFDLTYE-MNRIQDAIK-KLNESYINLKEVGTYEM-YV--KWPWYVWL-LI-GLAGVAVCVLL
FFIC-CCTGCGSC-CF-R-KCGS-CCDEYGGHQDSIVIYNISA-HED----

- **Smith-Waterman**

('M-VF-LFL-V-K-N-AL-G-G-DLLFDKV-A-Q-GF-N-A-A-AL-GAISAS-IDRLINGRLT-LNA-C
VK-Q-G-G-HI-S-V-NAP-GL-F-H-Y-WL-L-A-V-FICC-TGCG')
('M-FI-LPS-C-Y-S-AT-K-D-IDDINFS-R-N-LG-N-L-N-AL-SALGKI-QLSNRFGAIS-SLQ-L
NA-I-T-I-FS-A-A-EKV-EC-K-Q-N-SY-V-T-V-WPYV-WLLI')