

ITBC Bi5 Proteomics & Programmeren in Java

Weektaak 1

OWE5 Proteomics

Taak 1 – Sequentieniveau

De aminozuursequentie bepaalt de driedimensionale structuur van een proteïne, de cellulaire locatie, functie en interactie met andere proteïnen (Anfinsen, 1973). De vraag is hoe de biologische functie en structuur van een proteïne door de aminozuursequentie gecodeerd wordt.

Er zijn een aantal methoden om eigenschappen (features) van proteïnen op basis van de aminozuursequentie te voorspellen zoals het molecuulgewicht, iso-electrisch punt, secundaire structuur, transmembraan segmenten, en de solvent accessibility aan de hand van de hydrofobiciteit. Er zijn ook methoden om de functie van proteïnen op basis van sequentie-informatie te voorspellen. Hierbij horen methoden om profielen en patronen in proteïnen te herkennen als ook de voorspelling van cellulaire lokalisatie.

Opdrachten

Om je voor te bereiden op het project dat in week 3 begint deze inleidende opdrachten. Maak de opdrachten individueel en vat dan samen met de projectgroep. Doen is hier heel belangrijk om later je weg goed te kunnen vinden tussen alle eiwitsites. En natuurlijk studeer je op deze manier al voor de thematoets.

Opdracht 1

Beantwoord de volgende vragen / werk de opdrachten uit. Bespreek deze tijdens het tutor uur.

1. Geef de definities van de volgende begrippen:
 1. proteïne domein
 2. proteïne familie
 3. consensus sequence
 4. pattern, signature, fingerprint, motif.
 5. position specific scoring matrix (PSSM)
 6. Hidden Markov Model (HMM)
 7. profile

2. Wat is Gene Ontology (GO)? Hoe is GO opgebouwd? Noem 2 voorbeelden waarvoor je Gene Ontology kan gebruiken?
3. Wat zijn kenmerken van transmembraan segmenten? Noem 3 verschillende manieren/algorithmes om een transmembraan segment te voorspellen

Opdracht 2

Onderzoek met je projectgroep één van de sequenties (zie bestanden). Verwijs naar de gebruikte **bronnen** en vermeld de **accessiecodes**:

1. Functie (moleculaire functie, cellulaire locatie en biologisch proces)
2. Transmembraan segmenten / topologie (= vouwing in het membraan, waar zit de N- en waar de C-terminus). Gebruik 3 tools.
3. Matches met proteïnefamilies
4. Geconserveerde domeinen
5. Profiles.

Vat je resultaten in een presentatie samen en bediscussieer of de resultaten betrouwbaar zijn. Wat is nu de functie van dit eiwit?

Maak voor deze opdracht gebruik van:

- Voor onderzoek naar structuur en functie van het eiwit: BLAST (hypothese mogelijke functie), NCBI/Uniprot (sequenties met annotatie), PubMed (literatuur via NCBI), CDD (conserved domains database via NCBI), GO (gene ontology), Interpro (protein families), Prosite (protein domains), Pfam (protein families), MOTIF search (vinden van patterns en profiles)
- ExPASy tools voor **voorspellen** van transmembraan domeinen, topologie en cellulaire locatie. Let goed op het onderschrift van de figuren. Een figuur moet zelfstandig 'leesbaar' zijn. Hoe is het figuur tot stand gekomen en wat laat het zien. Geef in de discussie een beschrijving van de mogelijke functie van deze sequentie. Gebruik daarvoor de informatie die je hebt gevonden.

Let op. De eerste Blast hit is niet identiek aan de sequentie, hij lijkt er alleen veel op. Voor verder onderzoek kun je beginnen met wat er bekend is over de sequentie gevonden met Blast. Ter verificatie kun je eigenschappen voorspellen. Komen de voorspellingen overeen met de sequentie gevonden met Blast? Verwijs in je presentatie naar de bijbehorende bronnen en vermeld de relevante accessiecodes. Bij tools en databases verwijzen naar het artikel behorend bij de tool of database. Accessiecodes alleen vermelden in de tekst (PDB: 1ABC).

Literatuur

- Definities: Pevsner, Bioinformatics 3rd edition, glossary + blz 553-578.
- Proteins and transmembraan prediction: Lehninger, Biochemistry, 5th edition: hfst 2.2-3, 3, 4 en transmembraan predictie blz 378-379 + ~1000.
- Eiwit translocatie: Lehninger, Biochemistry, 5th edition: hfst 27.3 blz 1100 – 1106.

- Artikel over GO: Plessis, Skunca and Dessimoz, The what, where, how and why of gene ontology - a primer for bioinformaticians. Briefings in bioinformatics. vol 12(6), p 723-735 2011. Via [HAN Quest](#) of via [deze link](#)
- Topology prediction tools: <http://expasy.org/resources/search/querytext:topology%20prediction>
- Gene Ontology: <http://www.geneontology.org/>
- Conserved domain database: <http://www.ncbi.nlm.nih.gov/cdd>
- Interpro: <http://www.ebi.ac.uk/interpro/>
- Pfam: <http://pfam.sanger.ac.uk/>
- Prosite: <http://prosite.expasy.org/>
- Motif search: <http://www.genome.jp/tools/motif/>
- Motif scan: <http://hits.isb-sib.ch/cgi-bin/PFSCAN>
- Expasy: <http://www.expasy.ch/proteomics>

Resultaten

Een presentatie van de uitwerking van de opdrachten (deel 1 & 2).

Copyright 2018 @ I.Paffen