

Thema 1, script 2

Verplichte onderdelen:

- Zoek informatie over C-C chemokine receptor type 5 en diens rol in HIV infecties. Download de referentie-sequentie behorende bij dit *eiwit* in FASTA-formaat.
- Schrijf een programma dat dit FASTA-bestand inleest en telt uit hoeveel aminozuren het eiwit bestaat. De naam van het te lezen bestand dient te worden gespecificeerd in een string-variabele. Zorg dat de juiste uitkomst wordt weergegeven.
- Laat het programma automatisch bepalen of een gegeven FASTA-bestand nucleotiden of aminozuren bevat, en meld dit aan de gebruiker. Test dat je programma juist werkt voor FASTA-bestanden met zowel DNA/mRNA- als eiwit-sequenties. (Hint: als de sequentie uit alleen A's, T's, C's en G's bestaat mag je ervan uitgaan dat het nucleotiden betreft.)
- Breid je programma uit zodat het het afzonderlijke aantal van elk van de voorkomende nucleotiden/aminozuren telt, en deze aantallen weergeeft op het scherm. (Hint: gebruik een dictionary.)
- Als het om nucleotide-sequenties gaat, geef dan eveneens het zogenaamde "G-C" percentage weer. (Hint: <https://en.wikipedia.org/wiki/GC-content>.)

Bonus-onderdelen:

- Zoek op de NCBI-website ook soortgelijke eiwit referentie-sequenties behorende bij andere aandoeningen (bv: taaislijmziekte, sikkelcelanemie) en test de werking van je programma hiermee.
- Pas je programma aan zodat de naam van het FASTA-bestand door de gebruiker kan worden ingevoerd. (Hint: gebruik de functie `input()`.)
- Pas je programma aan zodat het een staafdiagram tekent van de nucleotide/aminozuur-samenstelling (d.w.z. de percentages van de verschillende aminozuren). Gebruik het hash-teken om de staven te maken: honderd "#" op een rij betekent "100%" van de nucleotiden/aminozuren in het bestand is van dat type. Teken de staven op een regel van links naar rechts, en zet aan het begin van elke staaf de afkorting. Bv.
A ###
L ####
Q ##
T ####
R #
...
- Breid je programma uit zodat de gebruiker meerdere bestanden kan invoeren. Je programma dient van elk bestand afzonderlijk de inhoud te geven.

Lever je programma in via BlackBoard in de vorm van een python-bestand genaamd **script2.py**.

Let hierbij op dat je code:

- ✓ op tijd wordt ingeleverd;
- ✓ zonder foutmeldingen uitvoerbaar is;
- ✓ de (verplichte) onderdelen van de opdracht correct en volledig verricht;
- ✓ liefst zo efficiënt mogelijk geïmplementeerd is;
- ✓ en, een netjes leesbare programmeerstijl gebruikt.