

DNA is opgebouwd uit fosforgroepen, suikergroepen en stikstofbase. Er zijn 5 soorten stikstofbase: Thymine, Adenine, Cytosine, Guanine en Uracil. Voordat er eiwitten gemaakt kunnen worden moet het DNA doormiddel van transcriptie omgezet worden naar RNA, dat vervolgens door translatie wordt omgezet in eiwitten. Het eiwit dat in elke week terug komt is CDK 1. De data die gebruikt is is afkomstig van de NCBI database, de PDB database en uniprot. Alle data is opgeslagen in FASTA bestanden, de resultaten zijn verwerkt in een Excel bestand. In week 2 is er een script geschreven dat het aantal nucleotiden kan tellen en het GC% van sequenties kan berekenen. In week 3 is er een script geschreven dat de lengte van een DNA sequentie en eiwit sequentie kan vergelijken en in een eiwit sequentie het aantal aminozuren D,E,R en K kan tellen. Week 4 ging het er om om met behulp van een script het molecuulgewicht van een eiwitsequentie uit te rekenen. Week 5 is er een script geschreven dat het identity-percentages van 2 eiwit sequenties kan berekenen. Een van de meest opvallende resultaten was dat de eiwitsequenties niet 3 x zo klein zijn als de DNA sequenties. Al zou je dit wel verwachten omdat 1 eiwit gevormd wordt door 3 stikstofbase. Maar uit de geschreven scripts kwam een ander resultaat: de DNA sequenties zijn groter dan 3 keer de lengte van een eiwit sequentie. Dit zou betekenen dat er stukjes DNA zijn die niet coderen voor eiwitten. Een ander opvallend resultaat kwam uit week 2, namelijk dat het GC% erg verschilt per sequentie. Dit komt doordat het GC% indicator is voor hoe stabiel een DNA sequentie is. Voor een onderzoeker kan dit erg nuttig zijn voor als hij een bepaald stabiel (of juist instabiel) stuk DNA nodig heeft want dan hoeft hij alleen maar naar het GC% te kijken. Een beperking van de scripts is dat er gebruik wordt gemaakt van input bestanden. In principe handig voor de gebruiker maar het nadeel is dat het script in dezelfde map moet staan als de input bestanden anders werkt het script niet. De vraagstelling voor een vervolg onderzoek zou kunnen zijn: kunnen we voorspellen hoe eiwitten met een bepaalde lading zich zullen gaan vouwen.

Aantal woorden: 371