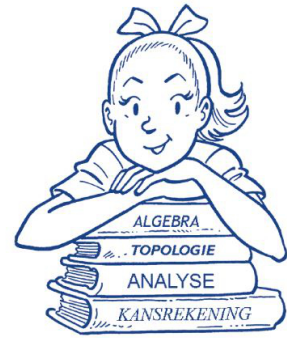


WISKUNNEND WISKE

DE DEFECTE DNA-KETENS

FINALE 2018 - RODE DRAAD B



Na de eerste bloedproef blijven er van de 3000 nog 20 sporters over die een positieve test aflegden. Een tweede bloedtest zou de lijst al korter kunnen maken, maar Wiske kwam met een ander idee. Op de spuit werden DNA-sporen gevonden. Deze kunnen vergeleken worden met het DNA dat uit de bloedstalen onttrokken kan worden. Door de eerste bloedproef zijn er helaas wel enkele fouten in de DNA-ketens geslopen. Gelukkig kunnen de fouten via indicatoren worden opgespoord.

Een korte opfrissing biologie:

1. DNA bestaat uit strengen opgebouwd met 4 nucleotiden, afgekort door de 4 letters A,C,G,T. Bijvoorbeeld ATACGT.
2. DNA heeft de vorm van een dubbele helix, dit betekent dat er voor elke streng een complementaire streng is. De nucleotiden zijn gepaard, dit betekent dat als op een bepaalde positie van een streng de letter A staat, dan staat in zijn complement de letter T op dezelfde positie. Analoog zijn C en G gepaard. De complementaire streng van het vorige voorbeeldje is dus TATGCA.



Een meetfout in het DNA wil zeggen dat een letter vervangen werd door een andere. De streng TTACGT kan zo een meting zijn met één fout van de correcte streng ATACGT.

Wiske tracht nu de dopingschuldige op te sporen op de volgende manier. Uit de bloedtesten worden voor elke atleet twee complementaire strengen gemeten, die elk **onafhankelijk** fouten kunnen bevatten. Het totaal aantal fouten in de twee strengen samen is gelukkig beperkt tot ten hoogste 3. De bedoeling is om de juiste strengen te reconstrueren, rekening houdende met de indicatoren, en te vergelijken met de DNA-streng gevonden op de spuit. Als de **gereconstrueerde streng of zijn complement** voorkomt als deelstreng, dan is die sporter een mogelijke verdachte. De streng ATACGT is bijvoorbeeld een deelstreng van GTC **ATACGT** AAA. De DNA-streng die Wiske aantrof op de spuit is de volgende:

ATGGCTAACTGTACGTTCTAG.

De gemeten DNA-strengen van de 20 sporters zijn weergegeven in de tweede kolom van onderstaande tabel. De laatste drie kolommen zijn indicatoren waarmee de meetfouten kunnen gedetecteerd en gecorrigeerd worden.

- De indicatoren #A en #T in kolom 3 en 4 kunnen de waarden 0 of 1 aannemen. De waarde van de indicator #A is 0 als A een even aantal keer voorkomt in de **bovenste** streng, en 1 anders. Analoog voor de indicator #T.
- De indicator #G+#C geeft de rest na deling door 3 van het aantal G en C samen in de **bovenste** streng.

Bijvoorbeeld, voor de streng ATACGT zijn de waarden voor de drie indicatoren respectievelijk 0, 0 en 2.

Opdracht (40 minuten)

Van welke sporters kan het DNA overeenkomen met de DNA-streng aangetroffen op de spuit? Geef een opsomming.

Noteer jullie antwoord ook op een apart blad en houd het bij tot de laatste opdracht van de dag. Je zal de resultaten nog nodig hebben.

Hint: Je zou 10 mogelijke verdachten moeten vinden.

ATGGCTAACTGTACGTTCTAG

Naam	DNA	#A (mod 2)	#T (mod 2)	#G+#C (mod 3)
Ine Tegraal	GCGAACTTT CGGTCGACA	0	0	2
Dries Hoek	CCACTATCG GTTGCTAGC	0	0	2
Piet Agoras	TTGCTATCG TTCGATAGC	1	0	1
Sien Metrie	GCCAAGATC ACGTCCTAG	1	0	1
Milo Garitme	GCTAACTAA CGATTGACA	0	1	1
Evi Erkant	CGATTTAGG GCTAAATCC	0	1	1
An Alyse	TGCTAAGGA AGGATTCCT	1	0	1
Raf Geleide	CCCATGAC GGGTAAGT	0	0	2
Cas Sociatief	TTGGCTAAC TACCTATTG	1	0	1
Debbie Nair	CCAAGTGA GATGGACAC	1	1	0

ATGGCTAACTGTACGTTCTAG

Naam	DNA	#A (mod 2)	#T (mod 2)	#G+#C (mod 3)
Pascal Driehoek	TTGGTTCTA AGCCAAGAA	0	0	0
Elfje Priem	CATGCAAGA GTAAATTTT	0	1	1
Frank Taal	GGCTCCTGA CCGAGGTCT	1	0	0
Ima Ginair	ACTCGGGAT TGAGGGGTA	0	0	2
Mo Dulo	AAGTGTACG TTCACATGC	0	0	0
Victor Ruimte	CTATACGTT CACATGTAA	1	0	1
Bas Is	AACCGATTG ATCCCTAAC	0	1	1
Lien Iaal	TGAAAGATC ACCTTATAG	1	0	1
Petra Pezium	ACCTTATAT TGAAAGATA	0	0	0
Emma Triks	CGCTACGTA GCGATCCAT	0	0	2