

Teksten van boxen in hoofdstuk 10

Box 1. Zelf open reading frames zoeken

Een manier om potentiële genen in een genoom op te sporen is te zoeken naar *Open Reading Frames* (ORF). Dat zijn stukken DNA waarin een startcodon en een stopcodon zijn gelegen met daartussen voldoende nucleotiden voor het coderen voor een eiwit. Niet elk gevonden ORF hoeft ook daadwerkelijk een gen te zijn. Een aardige didactische tool om ORFs in DNA-sequenties te zoeken is de *geneboy*, een digitaal apparaat gebaseerd op het uiterlijk van een gameboy. Je kunt hier ook zelf sequenties invoeren en daar RNA-sequenties en complementaire sequenties van maken.

Bron: www.dnalc.org/resources/geneboy.html

Je kunt ook gewoon in Word met de functie 'zoeken en vervangen' naar open reading frames zoeken in een DNA-sequentie, door bij 'zoeken' de start- en stopcodons in te typen en bij 'vervangen' START of STOP. Hiervoor kunnen het beste genen worden gebruikt met weinig introns, zoals bacteriële genen.

Box 2. Ons genoom op internet

De meest actuele gegevens over het menselijk genoom zijn direct op een smartboard te tonen of door leerlingen op te zoeken bijvoorbeeld via de site

www.ensembl.org/Homo_sapiens/Location/Genome

Als je naar onderen scrollt is te lezen dat ons genoom bestaat uit ruim drie miljard basenparen en dat het aantal bekende eiwitcoderende genen bijna 21.000 bedraagt. Daarnaast blijken er meer dan 13.000 pseudogenen te vinden te zijn, en ruim 8000 RNA genen. Ook is te zien dat het aantal bekende transcripten op ruim 167.000 ligt, dus acht maal het aantal eiwitcoderende genen. De verklaring hiervan wordt gegeven in 10.2 en box 4.

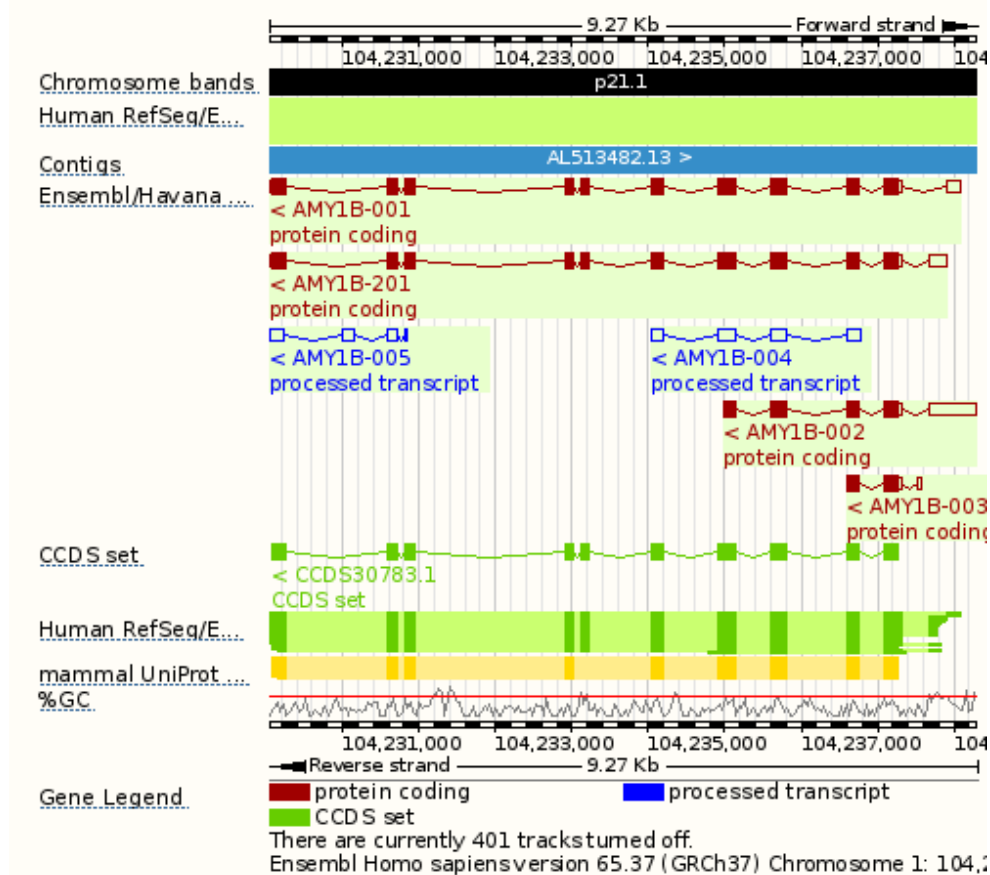
Box 3. Het genotype als fenotype

In sommige gevallen worden kenmerken van organismen bestudeerd die het DNA zelf betreffen. Dat is het geval bij de DNA-profielen zoals hierboven beschreven. Het aantal short tandem repeats op een bepaalde locus is het kenmerk dat wordt bepaald; het genotype is hier tevens het fenotype.

Een ander voorbeeld is de identificatie van parasitaire slakkensoorten die in koralen leven. Deze zijn uiterlijk identiek en alleen herkenbaar aan DNA-kenmerken, die wel precies correleren met het type koraal waar de slak op parasiteert (Bionieuws 2 april 2011).

Box 4. De amylase-genen in ons genoom

Ga naar de ensembl-databank: http://www.ensembl.org/Homo_sapiens/Info/Index?type='amylase' in het vakje, klik vervolgens op 'gene', dan op 'Homo sapiens'. Er blijken meerdere amylasegenen te vinden te zijn.



Amylase transcripten, exons en introns.

Het gen 'AMYP1' uit deze rij is een pseudogen, de genen voor speekselamylase worden aangeduid als AMY1, de genen voor pancreasamylase met AMY2. Aan de getallen is te zien dat deze genen in chromosoom 1 liggen, en wel vlakbij elkaar. Deze amylasegenen zijn ooit door verdubbeling uit een gezamenlijk 'vooroudergen' ontstaan.

Klik op AMY1B. Dit gen blijkt op zijn beurt zes transcripten te hebben. Het eerste transcript, AMY1B001 heeft 1787 basenparen, maar slechts 511 aminozuren. Er lijken wat basenparen 'teveel' te zijn, als je het aantal aminozuren x 3 berekent (1533). Dat komt doordat het RNA-polymerase altijd een paar nucleotiden voor het startcodon start en na het stopcodon eindigt. Die stukjes worden in de translatie niet omgezet in aminozuren en worden daarom *Untranslated Regions* (UTR) genoemd. Ook is op die webpagina te zien waar het gen ligt: in chromosoom 1, en wel tussen de basenparen 104.230.037 en 104.239.302. Hier ligt een heel groot verschil in basenparen, namelijk de 1787 die de lengte van het mRNA- transcript aangeeft en de afstand tussen begin en eind van de locatie, die meer dan 25000 basenparen bedraagt. Dit grote verschil in basenparen tussen de locatie en het mRNA ligt aan de introns, die wel meetellen in het bepalen van het begin en eindpunt in het chromosoom, maar niet in het mRNA.

Als je op de link klikt [Chromosome 1: 104,230,037-104,239,302](#) kun je onder elkaar drie afbeeldingen zien die de omgeving van het amylasegen aangeven. De bovenste afbeelding geeft weer waar het gen op het chromosoom te vinden is. De afbeelding daaronder laat zien dat het gen gelegen is in een gebied waar wel van alles in het DNA gelegen is, maar nauwelijks eiwitcoderende delen. De onderste afbeelding laat de exons (blokjes) en introns (lijnen) zien van elk transcript.

Box 5. TP53 als voorbeeld van een genennetwerk

TP53 is codeert voor het eiwit P53, dat een hoofdrol speelt in apoptose, 'gereguleerde celdood'. Mutaties in dit gen zijn belangrijke oorzaken van kanker, doordat apoptose een van de manieren is waarop het lichaam ontspoorde cellen kwijtraakt. TP53 staat daarom bekend als tumorsuppressorgen.

Aan dit gen en eiwit is een hele site gewijd: <http://p53.bii.a-star.edu.sg/index.php>. Bij het openen van deze site kun je onder andere kiezen voor *transcriptional targets*. Dat zijn de ruim 60 genen waarop het eiwit P53 invloed heeft. Door op deze genen te klikken is te zien welke functie die genen weer hebben. Je kunt op de site ook kiezen voor *transcriptional factors*. Dat zijn de ruim 20 eiwitten die de transcriptie van TP53 kunnen beïnvloeden, waarbij ook aangegeven is of ze deze activeren of remmen. Er is dus sprake van een netwerk van relaties; het gen TP53 wordt aan- of uitgezet door veel stoffen in de cel, en het genproduct (het eiwit P53) kan weer veel andere genen aan- of uitzetten.

Dit voorbeeld laat ook zien dat een mutatie die een van de *transcriptional factors* uitschakelt, niet automatisch een grote invloed op apoptose hoeft te hebben, omdat er andere factoren zijn die een soortgelijk effect op TP53 hebben. Anderzijds is duidelijk dat een mutatie die P53 uitschakelt, een hele lawine van effecten teweeg zal brengen.

Box 6. Regulerende elementen en kanker

Kanker wordt niet alleen veroorzaakt door gemuteerde genen die voor verkeerde eiwitten coderen, maar ook vaak door genen die voor het juiste eiwit coderen maar te veel of te weinig actief zijn, of actief op het verkeerde moment. Dat betekent dat de meeste systemen die betrokken zijn bij de regulatie van genen ook terug te vinden zijn als factoren die betrokken zijn bij kanker. Zo blijkt het patroon van micro-RNA's (een van de non-coding RNA vormen) in kankercellen af te wijken van het patroon in gezonde cellen en bruikbaar te zijn in de diagnose van het type kanker (Meltzer, 2005). Overmatige methylering kan tot kanker leiden door het afremmen van tumoronderdrukkende genen. Ook afwijkende methylering / acetylering van histonen kan ertoe leiden dat genen op een ongewenst moment actief worden. Het koppelen en losmaken van acetylgroepen aan histonen gebeurt door de enzymen *histon acetyltransferase* (HAT) en *histon deacetylase* (HDAC), die op hun beurt weer geremd kunnen worden door enzymremmende stoffen. Sommige van deze enzymremmers, die dus indirect invloed hebben op de activiteit van genen, worden nu ontwikkeld als medicijn tegen bepaalde vormen van kanker (Marks e.a. 2000).

Box 7. Er bestaan geen kankergenen

Mediakoppen als 'nieuwe kankergenen ontdekt' versterken het beeld dat er sprake zou zijn van 'slechte genen' die in ons lichaam de rol hebben van ziekteverwekker. Slordig taalgebruik houdt dat in stand, maar ook in wetenschappelijke communicatie worden gemakshalve termen als oncogen gebruikt. Je zou dit foutieve beeld kunnen vergelijken door in een ontwerp voor een auto het gaspedaal een 'ongelukspedaal' te noemen omdat vastzittende gaspedalen vaak tot ongelukken leiden. In werkelijkheid coderen 'oncogenen' in normale cellen voor functies zoals het opvangen van groeisignalen. Vaak wordt een gen of de functie ervan pas zichtbaar bij verstoring, en wordt daardoor dit gen vernoemd naar het afwijkende verschijnsel of de ziekte.

Box 8. Kapotte genen versus verstoorde balans

Sikkelcelanemie is een sprekend voorbeeld hoe een verandering van één nucleotide invloed kan hebben op de eiwitstructuur en daardoor op fysiologische eigenschappen. Maar in veel gevallen ontstaan problemen niet zozeer doordat een gen voor een afwijkend product codeert, maar doordat dit product te veel, te weinig, of op een verkeerd moment ontstaat. Het meeste bekende voorbeeld is het Downsyndroom, veroorzaakt door trisomie 21. Hier is geen sprake van kapotte genen, maar van een verstoorde balans door de aanwezigheid van een extra kopie van de genen in chromosoom 21. Ook bij kanker zijn vaak chromosoommutaties betrokken zoals duplicaties, deleties en translocaties. Daardoor wordt de balans tussen remmende en stimulerende invloeden op de celdeling verstoord. Een normale embryonale ontwikkeling hangt ook af nauwkeurig op tijd en hoeveelheid afgestemde productie.

Het wegvallen van een remming kan ook tot evolutieprocessen leiden. Veel genetische verschillen tussen mens en chimpansee blijken deleties te zijn, onder andere van een gebied dat betrokken is bij de regulatie van de groei van hersendelen (McLean e.a., 2011).

Box 8. Voelt u zich genomen?

Als je de advertenties van aanbieders van *personal genome tests* leest is ons genoom ons ware zelf. Dat blijkt al uit de sitenamen; 'Knome', 'Mycellf', 'DeCODEme'. Het genoom is het nieuwe beeld van de ziel; datgene dat we diep van binnen zijn, wat ons kenmerkt en wat ons overgeleverd is. Is dat beeld te rijmen met het gegeven dat we allerlei genomen in ons lichaam hebben? Is ons lichaam het exclusieve voertuig van onszelf, of een bus vol passagiers die elk kunnen claimen dat het hun bus is, en wij alleen de chauffeur? Als we voor 99,5% overeenstemmen met de persoon naast ons en voor 95% met de chimpansee, is het dan die 5% die ons mens maakt en die halve procent die ons onszelf maakt? Of is juist het wezenlijke van de mens dat de hele evolutionaire geschiedenis in de genen is terug te zien, inclusief fossiele genen en virusrestanten?

Box 9. Metagenomics

Metagenomics is een onderzoeksveld waarin zonder voorafgaande determinatie DNA wordt verzameld en geanalyseerd van alle organismen die aanwezig zijn in een monster van water, bodem of darm. Doordat verschillende soorten herkenbaar zijn aan specifieke DNA-kenmerken (bij bacteriën vooral ribosomaal RNA) kunnen soorten worden onderscheiden zonder isolatie vooraf. Deze techniek is vooral in het nieuws gekomen door onderzoek van Craig Venter, die al beroemd was geworden door de competitie rond het sequensen van het humane genoom. Hij verzamelde watermonsters uit de door hemzelf bevaren Sargassozee. Door het gehele metagenoom, dus de optelsom van de aanwezige genomen tegelijk te analyseren, heeft hij 1800 verschillende soorten op basis van hun genoom kunnen identificeren en bovendien 148 nieuwe soorten ontdekt (Venter e.a., 2004). In Wageningen wordt deze techniek onder andere toegepast in onderzoek naar bodems die ziekteverwekkende eigenschappen hebben.

Box 10. Een insulinstamboom maken

Insuline heeft een basale functie in dierlijke organismen en het is aannemelijk dat een gen voor insuline al in vroege voorouders aanwezig was. Via bioinformatica-programma's zoals BLAST en ClustalW kunnen niet alleen de peptidestructuren van insuline vergeleken worden, maar kan ook een evolutionaire stamboom worden gemaakt op basis van deze volgorden. Ga naar de website <http://mrs.cmbi.ru.nl> en zoek in de databank Swiss-Prot naar 'Insulin os:human'.

Selecteer de hit 'ins_human' onderaan de lijst.

1. Scroll helemaal naar beneden voor de aminozuurvolgorde. Kopieer deze in een Word-bestand.
2. Scroll weer naar boven en klik op Blast. En vervolgens op Run Blast.
3. Als het programma klaar is met rekenen, kun je op het resultaat klikken.
4. Je ziet nu een lijst van organismen met insuline. Verzamel de aminozuurvolgordes van verschillende organismen en kopieer die naar het Word-bestand. Begin iedere volgorde met >naam van het organisme.
5. Kopieer de inhoud van het gemaakte Word-bestand naar het tekstveld op de website www.ebi.ac.uk/Tools/msa/clustalw2.
6. Klik op submit. Als dit ClustalW2 programma klaar is met rekenen kun je klikken op 'guide tree'. Onderaan de pagina staat een stamboom op basis van insuline.

Zie voor dit soort activiteiten ook www.bioinformaticaindeklas.nl.

Box 11. De Gamma voor beta's

Er zijn internationale wedstrijden voor studenten om nieuwe varianten van bouwstenen van levende cellen te ontwikkelen (iGEM). Hiervoor is zelfs een soort Bouwmarkt-catalogus beschikbaar, de 'Registry of Standard Biological Parts' http://partsregistry.org/Main_Page. Deze catalogus geeft een aardig beeld van wat onder bouwstenen van de cel wordt verstaan en wat daarin te koop is.