

## Hoe leven werkt



Foto Rein Janssen

**'Big data en genomics', is dit jaar het thema van de publiekslezing De Anatomische Les, gehouden door de Amerikaanse geneticaexpert Eric Lander. We geven u vast een spoedcursus, aan de hand van vijf vragen die u niet durft te stellen omdat het vast domme vragen zijn (dat zijn het niet).**

MAARTEN KEULEMANS

### Wat is dat, 'genomics'?

Ze zeggen dat het de wereld gaat veranderen. Zoals de 20ste eeuw de eeuw was van de natuurkunde, van ongelimiteerde controle over elektronen en materialen, zo wordt dit de eeuw van de biologie, de eeuw waarin we levende weefsels zullen leren beheersen als nooit tevoren. Verwacht futuristische zaken als weer aangroeiende ledematen en in het lab gekweekte organen, verwacht bacteriën die ons niet ziek maken maar juist genezen en virussen die ons niet verzwakken maar ons dna herstellen. Allemaal dankzij de uitdijende oerknal van begrip rond datgene wat ooit nog gewoon 'genetica' of 'erfelijkheidsleer'

heette.

Maar eerst de basics.

Een menselijk lichaam is opgebouwd uit een slordige 37 biljoen cellen. Zoom in op die cellen en je zult binnenin de celkern vinden, met daarin een soort kluwen wol, het dna. Verdeeld over 46 pakketjes, die vooral tijdens de celdeling zichtbaar worden als een soort worstjes, de chromosomen.

Het dna - desoxyneucleïnezuur voor kenners - is het erfelijkheidsmolecuul. Hier ligt vast wat we zijn en nog kunnen worden; hier ligt de belofte van ons nageslacht, maar ook de erfenis van voorouders. Verpakt in de gedaante van een eindeloos lange sliert, met de iconische wenteltrapvorm.

Het dna-molecuul zelf bestaat uit twee lange, om elkaar kronkelende strengen, verbonden door 'traptreden'. Die bestaan steeds uit vier moleculen: adenine, guanine, cytosine en thymine, doorgaans afgekort met hun beginletters A, G, C en T. De treden zijn losjes met elkaar verbonden: de G klikt altijd vast aan de C, en de A altijd aan de T. Omdat de basen per trede verschillen, ontstaat er langs de wenteltrap een eindeloos lange letterreeks: A, C, T, T, G, G, T, C, A, A, A, A; ga zo maar door. Ziedaar de befaamde 'code van het leven' waarin cellen hun erfelijke informatie opslaan, opgeschreven in A's, T's, G's en C's.

Zo'n drie miljard tredes ofwel 'basenparen' telt de volledige dna- volgorde van de mens, en het is goed te beseffen hoe enorm dat eigenlijk is. Als je de genetische lettercode van een mens zou uitschrijven en afdrukken in Sir Edmund, zou je er 11 duizend nodig hebben, een stapel zo hoog als een tien verdiepingen hoog flatgebouw. Het genoom voorlezen zou 46 jaar kosten (zonder pauze

of nachtrust); de dna-code afdrukken in het lettertype waarin u dit leest, zou een sliert letters opleveren van hier tot in Kazachstan.

En dat dus 37 biljoen keer, in elke cel opnieuw. Zó ontzagwekkend is het 'genoom', het woord voor de totale genetische informatie van de mens. 'Genomics' is losjes gezegd: alles wat gaat over dat genoom. De genoomkunde.

### Waar zitten de genen eigenlijk?

In onze cellen zit uiteraard niet alleen dna; cellen zitten volgepropt met eiwitten, ingewikkelde moleculen die iets 'doen'. Zo'n 100 duizend soorten menselijke eiwitten zijn bekend. Die zwengelen chemische reacties aan, klitten samen tot bepaalde celonderdelen, plakken aaneen tot een celwand of tot vezels of moleculaire steunbalkjes, en vormen zodoende het hele moleculaire binnenwerk van een cel. Uiteindelijk bestaat ons lichaam volledig uit krioelende en aan elkaar plakkende eiwitten, aangelengd met water en vet en wat andere losse moleculen.

Maar die eiwitten moeten ergens vandaan komen, en daar duikt het genoom op. Op onze lettersliert van hier naar Kazachstan zitten ongeveer 23 duizend stukken 'tekst' die de instructies voor eiwitten bevatten. Gemiddeld bestaat zo'n instructie uit 30 duizend letters, ofwel zo'n zeven pagina's Sir Edmund vol A's, G's, T's en C's. Maar in de praktijk loopt de lengte van genen sterk uiteen, van 500 basenparen (een lettersliert van een halve meter, in het lettertype waarin u dit leest) tot 2,4 miljoen (3 kilometer lang).

Grof gezegd leest de cel in de genen af hoe je een eiwit moet maken door de wenteltrap open te scheuren, en het gen over te schrijven op een ander molecuul,

een kladblaadje genaamd rna. Dat brengt de boodschap over naar het ribosoom, een fabriekje dat de boodschap vertaalt en er eiwitten van maakt. Dat aflezen gaat met drietallen letters tegelijk, waarbij elk 'triplet' staat voor één aminozuur. ATG betekent 'methionine', GTA staat voor 'valine' enzovoorts - ons lijf werkt met 20 aminozuren. Het ribosoom klikt de aminozuren aan elkaar in de volgorde die de code voorschrijft. 'A-T-G-G-T-A-C-C-A': klik een bouwsteen methionine vast aan valine aan proline.

Dat klinkt nog redelijk overzichtelijk: genen die instructies bevatten voor eiwitten. Maar in de praktijk wordt het al snel zeer ingewikkeld.

Zo bestaat liefst 95 procent van ons dna helemaal niet uit genen, maar uit instructies voor allerlei schakelaars, springplanken en schuifjes die de cel gebruikt om de genen goed tot uitdrukking te brengen. We hebben immers zo'n 200 typen cellen en in verschillende organen en weefsels zijn verschillende eiwitten nodig. Welke eiwitten nodig zijn, verschilt ook nog eens per tijdstip.

Ook de genen zelf zitten vol trucs. Zo worden ze onderbroken door stukken andere informatie, de zogeheten 'intronen', alsof je in een zin HGRTFDGRDFEE hier en daar RGRFRGFHRHERRHT opeens HGRRFRRQWUYT stukken wartaal tegenkomt. Bovendien hebben de meeste genen meerdere 'startpunten' waar het aflezen kan beginnen: één gen kan zodoende verschillende eiwitten afleveren.

Zo is er nog veel meer. Genen die in een bepaald weefsel niet worden gebruikt, zijn 'afgeplakt' met eiwitten, als de bladzijden van een kookboek die aan elkaar plakken in hoofdstukken die je niet vaak leest. Ook bij de bouw van de eiwitten kan er van alles misgaan (of juist hersteld worden). En regelmatig verstoren genen elkaars signaal, of versterken ze die juist.

Het resultaat is een kakofonie van eiwitten, signalen, versterkers,

dimmers en plakkertjes. 'Het is een fabriek, het is constant in beweging en het reageert allemaal op de omgeving', zegt hoogleraar statistische genetica Daniëlle Posthuma van de VU Amsterdam. 'Echt een onvoorstelbaar proces eigenlijk. En zonder dat er iemand aan de knoppen zit, het gaat allemaal vanzelf.'

Het tomeloze tumult in de levende cel, waar alles tegelijk en dwars door elkaar heen gebeurt, staat zodoende in schril contrast met de steriele schematische plaatjes uit de biologieboeken. Met soms verrassende gevolgen, vertelt Posthuma. 'Vroeger dacht ik dat de relatie tussen wat er in het dna geschreven staat en het eiwit één op één was. Dit is de code, daar komt dit eiwit uit. Maar in werkelijkheid wordt de kans dat een gen ook echt wordt vertaald naar het betreffende eiwit eerder geschat op 1 op 3. Dan zie je opeens dat een patiënt een drager is van een defect gen, maar toch gewoon een goed eiwit maakt.'

### **Wat zit er in je genen als het in de genen zit?**

In de jaren tachtig en negentig kon je de krant niet openslaan of er was nieuws: hét gen voor intelligentie ontdekt, hét gen voor astma gevonden. Het 'O GOD'-tijdperk, zeggen genetica achteraf grappend: 'One Gene, One Disease', elke ziekte zijn gen. Het waren schoten voor de boeg; de openingszetten op het ontluikende veld van de moderne genetica.

Zeker, omdat alles draait om eiwitten en de recepten voor die eiwitten in de genen zitten, kun je daar opzoeken wat ons kenmerkt. Dat geldt voor zaken als oogkleur, lengte of de vorm van onze oren, maar ook voor veel subtielere zaken als temperament, politieke voorkeur, religiositeit en zelfs het al dan niet hebben van een vaste relatie. Mensen verschillen onderling ongeveer één op de duizend dna-letters van elkaar: het verschil tussen blauwe of bruine ogen, of tussen wel of geen aanleg

voor migraine.

Nou ja, een beetje dan. De pest is namelijk dat genen en ziektes maar heel losjes op elkaar passen. Wat in de genen staat, bepaalt maar een beetje wat in de praktijk gebeurt.

Neem de aandoening 'depressie'. Er zijn talloze genen die de kans op depressiviteit verhogen, maar wie depressief is, heeft niet automatisch een van die genen. Evenmin is het zo dat mensen met een 'gen voor depressie' ook altijd depressief worden. Vergelijk het met een genetische aanleg voor zwakke botten: ook iemand zonder de aanleg kan zijn been breken, net zoals iemand mét de aanleg best een heel leven lang botbreukvrij kan blijven.

Vertaald naar de enorme aantallen van de genetica, levert dat een complexe statistische dans op met kansen en patronen en associaties. 'Bij zaken als schizofrenie, autisme en overgewicht zijn misschien wel duizenden genen betrokken', zegt Posthuma. 'Terwijl de meeste genen an sich een klein effect hebben en misschien maar een paar procent van het risico op ziekte verklaren.'

Soms blijken achter een bepaald ziektebeeld verschillende genetische oorzaken te zitten. Zoals kanker: steeds duidelijker is dat er talloze manieren zijn waarop cellen tot de ongeremde celdeling kunnen komen die we kanker noemen. Met als gevolg dat iedere patiënt misschien baat heeft bij een medicijn op maat, een 'gepersonaliseerde behandeling'.

### **Eric Lander is pionier in het ontrafelen van het menselijke genoom. Hoe doe je dat?**

Tussen 1990 en 2003 maakte Lander deel uit van het grote internationale onderzoeksconglomeraat HUGO, dat het menselijke genoom als eerste 'uitlas', met een combinatie van chemische en computertechnieken. Het was

wereldnieuws.

Tegenwoordig kan het in enkele dagen, geheel geautomatiseerd, en voor maar een paar duizend euro. Het genoom wordt daartoe in stukjes gehakt van een paar honderd basenparen - een regel of vijf à zes, uitgedrukt in de tekst die u nu leest. Die stukjes worden letter voor letter uitgelezen, met behulp van een machine die de slijtjes probeert te koppelen aan A's, T's, G's en C's die bij elke 'match' een lichtsignaalje afgeven. Vervolgens worden de losse stukjes code weer aan elkaar gepuzzeld. Een legpuzzel vergelijkbaar met het weer in elkaar zetten van 11 duizend Sir Edmunds die door de papierversnipperaar zijn gehaald.

Ook burgers kunnen hun dna, of stukken daarvan, laten uitlezen, bij een groeiende waaier van bedrijven en bedrijfjes die een veelheid van tests bieden. 'Je spuugt in een flesje en je stuurt het op', zegt Annemieke Aartsma-Rus, hoogleraar translationele genetica aan het LUMC. Maar, benadrukt ze, zonder professionele begeleiding heb je daaraan weinig. 'Het zal meer vragen oproepen dan beantwoorden. Stel, je ziet een variatie in een bepaald gen dat betrokken is bij kanker. Nou én? Alle kans dat het een variatie is die we niet kennen en dat we geen idee hebben wat het precies betekent. Maar zoiets kan iemand wel ongerust maken.'

En denk aan de immense complexiteit van de genetica. De cel is een orkest dat een symfonie speelt. Het hoeft voor de symfonie niet direct iets uit te maken als een van de instrumenten - een van de genen - soms een valse noot aanslaat.

### En hoe zit het nou met die big data?

Inmiddels begint het dna-lezen een beetje uit de hand te lopen. Van zo'n 250 duizend mensen is het volledige genoom bekend, van nog eens honderden miljoenen meer zijn delen van de genetica bekend.

Tallose ziektes, eigenschappen en bevolkingsgroepen zijn genetisch 'in kaart' gebracht, aan de hand van honderdduizenden of miljoenen genprofielen per subgroep.

En dan is er nog het epigenoom (het overzicht van plakkerjes op de genen), het transcriptoom (het geheel van rna-boodschappenbriefjes), het proteoom (de eiwitten), het metabooloom (de routekaart van alle biochemische reacties in een cel) en verwante gegevensbanken zoals het microbioom (de genetica van de bacteriën die in onze darmen leven) en het interactoom (wat eiwitten met elkaar doen). Het is een succes en toch ook weer niet, zegt Aartsma-Rus: 'We kunnen nu veel meer. Maar een bijkomstigheid is wel dat we meer data genereren dan we aankunnen.'

Krankzinnige toestanden, levert dat soms op. Posthuma schetst het soort bestanden waarmee ze werkt: een spreadsheet met zo'n 10 miljoen kolommen (de dna-letters) en 100 duizend rijen (patiënten die hun dna hebben afgestaan). Om een indruk te geven: uitgeprint zou dat een vel papier opleveren waar de gemeente Eindhoven in zijn geheel onder past. 'Dat kun je niet meer openen in Wordpad', zegt Posthuma. 'Je moet ervoor inloggen bij een clustercomputer, waarna je met Unixcommando's met zo'n bestand kunt werken.'

En dan is de datatsunami nog nauwelijks op gang gekomen. Deze zomer bestudeerde een internationaal onderzoeksteam onder leiding van Gene Robinson van het Carl R. Woese-instituut voor genoombiologie in Illinois de data-aanwas voor de komende tien jaar. In 2025, schatten de onderzoekers in, zal de genomics 1.000.000.000.000.000.000 letters aan code per jaar in kaart brengen, een tekstsliert die je 10 duizend keer om het zonnestelsel kunt winden. De aanwas van genetische informatie gaat tegen die tijd tot 400 keer sneller dan de aanwas aan filmpjes op YouTube nu, becijfert Robinson.

Dat zal technische hobbels opleveren, maar ook onderzoeksproblemen, schetst Posthuma. 'De rol van de bio-informatici, maar ook die van de biologen wordt steeds belangrijker.' Die moeten immers bestuderen wat eventuele verdachte genetische patronen in de praktijk betekenen, doorgaans door de genetische code in kwestie in te bouwen in een bakje cellen of een proefdier en te bestuderen wat dan precies gebeurt.

De weg van de genomics loopt voorlopig bergopwaarts. 'Hoe meer data we hebben, hoe meer we kunnen zeggen,' zegt Aartsma-Rus. 'We kunnen niet anders dan vooruit. Maar het zal ontzettend veel analyse vergen om greep te krijgen op wat we hebben gevonden.'

### LEZERSAANBIEDING

Op donderdag 5 november houdt Eric Lander in het Concertgebouw in Amsterdam de 22ste Anatomische Les over de kansen op een nieuwe doorbraak in de genetica. Big data is daarbij het sleutelbegrip.

De Anatomische Les, aangeboden door het Academisch Medisch Centrum (AMC), het VU medisch centrum (VUmc) en de Volkskrant, is een jaarlijkse publiekslezing op het snijvlak van geneeskunde en maatschappij, door een internationaal toonaangevende spreker.

Zolang de voorraad strekt kunnen lezers van de Volkskrant toegangsbewijzen (à €5,-) online bestellen op [amc.nl/al-aanmeldingen](http://amc.nl/al-aanmeldingen).

Geen kaartverkoop aan de zaal, ook niet op 5 november. De lezing is in het Engels en begint om 13.45 uur.