

# DNA in de bieb

Amerikaanse onderzoekers zijn erin geslaagd het record voor **dataopslag in DNA** te breken. Ze codeerden er een boek mee.

SHURAILA STOPPEL

Onlangs codeerden Amerikaanse onderzoekers een boek, met elf illustraties, in DNA. De genetische code herbergt al miljoenen jaren biologische informatie. Synthetisch biologen kunnen sinds enkele jaren ook stukjes niet-biologische informatie in DNA opslaan. Tot nu toe lukte dat nooit eerder in grote hoeveelheden. Marcel Reinders, hoogleraar bioinformatica van de TU Delft is onder de indruk van het werk: "Dit is een echte mijlpaal. De schaal waarop het is gedaan is indrukwekkend."

Dataopslag in DNA is nu nog een probleem vanwege hoge kosten, maar de prijzen voor DNA-synthese en sequensen verminderen exponentieel. Volgens de onderzoekers dalen de kosten zelfs sneller dan die voor elektronische opslag-media. Vooral door de voortschrijdende ontwikkeling van *next-generation sequencing* lukt het om grotere hoeveelheden DNA steeds sneller en goedkoper te lezen. "Kleine simpele DNA-sequencers, nu in ontwikkeling, maken het straks makkelijker om DNA-gecodeerde informatie te lezen", voorspelt Reinders. "Daardoor kan ik me voorstellen dat we deze techniek over 10 jaar kunnen gebruiken."

## HOUDBAAR

De onderzoekers, Church, Kosuri en Gao van het onderzoeksinstituut Wyss in Boston en de Johns Hopkins Universiteit in Baltimore, begonnen hun uitdaging met een html-bewerking van het boek *Regenesis*. Daarvan maakten ze een bitversie. In die bitversie codeerden DNA-basen A en C voor 0; T en G codeerden voor 1. Zo kun je voor elke 0 of 1 een van de twee basen willekeurig kiezen. Daarmee zorgden ze ervoor dat er niet te veel moeilijke stukken zouden voorkomen, zoals extreem veel GC-gebieden of secundaire structuren. Met een computerprogramma maakten ze de benodigde DNA-segmenten. Voor het boek waren 54.898 datablokken nodig



**'Voor langetermijn-dataopslag is het zeker een alternatief'**

van 96 bits, ofwel 96 nucleotiden.

De informatiedichtheid, stabiliteit en efficiëntie in energieverbruik zijn, volgens de onderzoekers, de grote voordelen van dataopslag in DNA. De informatiedichtheid van de bits is 5,5 petabits per mm<sup>3</sup>. Ter vergelijking: op de huidige harddiskdrives is de informatiedichtheid in de orde van gigabits per mm<sup>3</sup>, een miljoen keer minder. Bovendien is DNA stabiel en in theorie eeuwenlang houdbaar.

Toch kunnen we nu niet direct onze harde schijf omwisselen voor DNA.

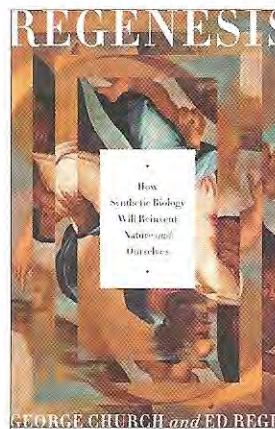
"Daarvoor is deze manier van opslag niet geschikt. Er komen te veel handelingen bij kijken. Maar voor langetermijnopslag van data is het zeker een alternatief", meent Reinders.

## KORTE STRENGEN

Pogingen in het verleden om data in DNA te coderen, stuitten op problemen met schrijven en teruglezen van lange stukken DNA. Daarom gebruikten de onderzoekers korte strengen DNA, van 159 nucleotiden. Aan 96 nucleotidendata-informatie, kwam een 19-nucleotiden-adres om de re-assemblering in goede banen te leiden. Verder kwam, aan beide kanten van de DNA-fragmenten, een 22-nucleotidesequentie voor amplificatie en next-generation sequencing.

Opslag van DNA kan op diverse manieren, ingevroren, op een chip of in een organisme. In deze studie kozen de onderzoekers ervoor het gesynthetiseerde DNA op het oppervlak van een microchip aan te brengen. Opslag in organismen als bacteriën had niet hun voorkeur. Organismen bewaren vergt onderhoud en ze kunnen nog doodgaan ook.

De wetenschappers maakten, op de microchip, zeer veel kopieën van de fragmenten. Dus zelfs als er tijdens de synthese van de DNA-fragmenten of tijdens de sequencing foutjes in de data sluipen, gebeurt dat willekeurig en zal elke kopie voor de andere corrigeren. Zo gaat de originele boodschap niet verloren.



Het 'vertaalde' boek.