

Standaard DNA sequentie moet identificatie vereenvoudigen

# Schimmels krijgen genetische barcode

Het Centraalbureau voor Schimmelcultures in Utrecht wil van alle schimmels en andere micro-organismen uit zijn collectie een genetische barcode vaststellen. Een database met al deze DNA sequenties moet de waarde van de collectie vergroten en de identificatie van onbekende schimmels vereenvoudigen.

Door Els van den Brink, gepubliceerd in het *Laboratorium Magazine* nr 8, november 2009.

Identificatie van micro-organismen zoals schimmels gebeurt traditioneel via een mix van morfologische en biochemische technieken. Experts van het Centraalbureau voor Schimmelcultures (CBS-KNAW) in Utrecht kunnen op basis van bepaalde kenmerken - zoals groeikarakteristieken en de manier van spoorvorming - vaststellen om welke soort het gaat. De laatste jaren wordt steeds duidelijker dat lang niet alle soorten op deze manier van elkaar zijn te onderscheiden. Een schimmel die morfologisch één soort lijkt te zijn, kan in werkelijkheid bestaan uit een complex van verschillende soorten. Soms wel meer dan tien. Moleculair biologische technieken op basis van DNA-sequencing geven veel meer duidelijkheid. Vandaar dat het CBS heeft besloten om hiervan een speerpunt te maken. Het CBS gaat van alle schimmels uit de collectie een genetische barcode vaststellen. Deze 'barcode of life' moet de waarde van de schimmelcollectie vergroten en de identificatie een stuk vereenvoudigen.

## Barcodeproject

Onderzoekers die op zoek zijn naar een bepaalde schimmel, of in hun eigen laboratorium geconfronteerd worden met een onbekende schimmel, kunnen daarvoor terecht bij het CBS. Het laboratorium heeft in de loop der jaren, sinds de oprichting in 1904, een grote collectie schimmels opgebouwd. Die bevat thans ruim 80.000 stammen. Daarnaast beheert het CBS ook de *Netherlands Culture Collection of Bacteria* (NCCB). De NCCB is ontstaan uit de fusie van kleinere collecties, en bevat in totaal zo'n 10.000 bacteriestammen, 450 plasmides en 600 bacteriofagen. Op zich maakt het CBS al langer gebruik van moleculair biologische technieken zoals DNA sequencing. "Het probleem is dat er nu nog geen goede referentiedatabase is, om sequence/data van een onbekende schimmel mee te vergelijken. Dat is wat we nu dus willen opzetten", vertelt Ursula Eberhardt, postdoctoraal onderzoeker en één van de kartrekkers van het barcode project binnen het CBS.

## Vereisten

Een genetische barcode is in feite niet meer dan de DNA/sequentie van een standaard stukje DNA uit het genoom van een schimmel of ander organisme, meestal de sequentie van één specifiek gen. Voor gebruik als barcode moet dit gen binnen één soort weinig variëren, terwijl er juist wel veel variatie moet voorkomen tussen de overeenkomstige gensequenties van verschillende soorten. Praktische vereisten zijn dat de het gen niet langer moet zijn dan de lengte die een standaard DNA sequencer in één run kan sequencen. Ook mag het niet te veel in lengte variëren door de aanwezigheid van introns. Een intron (een samentrekking van *intra-genic region*) is een stukje DNA dat zich bevindt in een gen maar niet dient om het eiwit te coderen. Voordat van alle schimmelsoorten een barcode kan worden vastgesteld, moeten onderzoekers dus eerst een gen uitkiezen dat geschikt is om als barcode te fungeren.

## Genselectie

Bij soortgelijke barcode-projecten voor dieren was eerder al het *Cytochrome C oxidase* (CO1)-gen geselecteerd als barcode. Dit is een gen dat voorkomt op het DNA in mitochondriën. Omdat een cel honderd tot duizend mitochondriën kan bevatten, zijn er relatief weinig cellen nodig voor een goede analyse. Voor schimmels bleek dit gen echter niet geschikt als barcode, omdat het verstoord wordt door meerdere introns. Eberhardt en haar collega's kozen daarom voor het *Internal Transcribed Spacer region* (ITS)-gen, een gen dat al langer wordt gebruikt voor identificatie van schimmels. Het ITS-gen is onderdeel van het genoom in de celkern. Het is een van de ribosomale genen, die ook in meerdere kopieën in de cel voorkomen, net als het mitochondrieel DNA. "Een belangrijk voordeel

van dit gen is dat het begin en eind bij de meeste schimmelsoorten ongeveer gelijk is, waardoor we voor de vermenigvuldiging van het DNA meestal dezelfde primers kunnen gebruiken”, vertelt Eberhardt. Het CBS heeft ervoor gekozen om behalve het ITS gen ook het naastgelegen ribosomale gen (LSU) te sequencen. Het LSU-gen is namelijk minder variabel, waardoor het makkelijker is om sequenties op een goede manier langs referentiedata te leggen en te vergelijken. Daarnaast biedt het LSU gen meer mogelijkheden om evolutionaire verwantschappen te bestuderen. Eberhardt verwacht echter dat dit steeds minder noodzakelijk wordt bij de verdere uitbreiding van de database.

### **Winst**

Hoewel de keuze voor het ITS als primair barcode-gen al is gemaakt, bekijkt het CBS wel of er nog andere, nog geschiktere, genen zijn. Er valt zeker nog winst te behalen. Eberhardt verwacht op basis van het ITS-gen zo'n zeventig procent van alle schimmelsoorten te kunnen identificeren. Het zou natuurlijk mooi zijn als het in de toekomst lukt om op basis van andere genen nog hoger uit te komen. Maar daar gaan Eberhardt en haar collega's niet op wachten.

### **Logistieke uitdaging**

Het barcode-project binnen het CBS is pas net van start gegaan. De aftrap was in juni 2009. Op dit moment bestaat het projectteam alleen nog uit Eberhardt en - sinds kort - twee analisten. Het is de bedoeling dat het team tot vijf mensen wordt uitgebreid om het project binnen drie jaar te kunnen afronden. De grootste uitdaging is de praktische uitvoering van het geheel, met name logistiek. Eberhardt vertelt: “Het moleculair biologische labwerk is vrij standaard, en in principe ook makkelijk te automatiseren, al hebben we daar op dit moment nog geen geld voor. We proberen wel steeds meer te werken in 96-well platen in plaats van in losse buisjes, om het proces wat meer te stroomlijnen. De fase die daaraan vooraf gaat, het voorbereiden van de schimmelstammen, zal echter altijd handwerk blijven. We zullen moeten proberen om alles zo te organiseren, dat we desondanks toch steeds voldoende materiaal aanleveren aan het moleculair biologische lab om door te kunnen gaan met de DNA-sequencing.”

### **Kwaliteitscontrole**

Na afronding van het barcodeproject wordt de identificatie van schimmels een stuk eenvoudiger. Onderzoekers hoeven dan alleen nog maar DNA te isoleren, een PCR reactie uit te voeren en het ITS-gen te sequencen. Een vergelijking met de database zal dan duidelijk maken om welke soort het gaat. Eberhardt verwacht dat het project ook veel nieuwe data zal opleveren. “Een soort die in het verleden alleen morfologisch is gekarakteriseerd, kan uiteindelijk een mix van verschillende soorten blijken te zijn. Sommige schimmelsoorten uit de collectie zijn alleen nog gekarakteriseerd in een tijd dat microscopen een stuk beperkter waren en er nog geen biochemische analysemethoden beschikbaar waren. Al met al zal de waarde van onze collectie sterk toenemen”, zegt Eberhardt. Daarnaast biedt de technologie ook mogelijkheden voor verbeterde kwaliteitscontrole. Het gaat dan vooral om de schimmels die niet als drooggevroren poeder of in de vriezer bij -180°C bewaard kunnen worden. Deze schimmels worden gehouden als actief groeiende cultuur en moeten periodiek worden doorgeënt. Daarbij kan de kwaliteit van de schimmels wel achteruitgaan. “We kunnen straks de barcode gebruiken als interne kwaliteitscontrole voor deze en andere stammen”, legt Eberhardt uit.

### **Ouderwets**

Hoewel in het project grote woorden worden gebruikt, kan Eberhardt het wel relativeren. “De relevantie van de barcode is maar tijdelijk. Over een jaar of tien is het waarschijnlijk alweer ouderwets. Maar tot die tijd is het zeker waardevol en zinvol om te doen. Bovendien, stel dat je zou overstappen op ‘next generation sequencing’, of de opvolger daarvan - al heb ik geen idee wat dat wordt - dan heb je voor de data-analyse nog altijd een referentiegenoom nodig. Dan is het wel handig als je markergenen al hebt gesequenced, zoals deze barcode.”

### **Barcodes in de rimboe**

Wereldwijd zijn tal van onderzoeksgroepen bezig met onderzoek naar de 'barcode of life'. Ze hebben zich verenigd in diverse internationale organisaties met klinkende namen, zoals de *European Consortium for the barcode of life* (ECBoL) en het *Barcoding of Life Initiative* (BOLI). Terwijl het CBS zich vooral richt op schimmels, houden andere groepen zich bezig met de barcodes van planten of dieren. Uiteindelijk is het idee om van alle bekende organismen een barcode vast te stellen en op te slaan in een centrale database. Het liefst zouden ze daarbij dan ook nog een draagbaar apparaatje ontwikkelen, een *Life Barcoder*, waarmee een bioloog in de rimboe direct de barcode van een nieuw ontdekt organisme kan bepalen. Het is de bedoeling dat er in Nederland twee speciale laboratoria komen voor barcode onderzoek, een voor barcodes van macro-organismen en een voor micro-organismen. Het eerste lab komt waarschijnlijk in Leiden, als onderdeel van nieuwe Nederlands Centrum voor Biodiversiteit, en het tweede lab in Utrecht, als onderdeel van het CBS.

### **Van soort tot stam**

Bij veel organismen is het niet alleen van belang om welke soort het gaat, maar ook om welke stam (*strain*). De onderlinge verschillen tussen deze stammen kunnen grote consequenties hebben. De ene stam kan bijvoorbeeld infectieuzer zijn dan de andere. Ook voor laboratoriumonderzoek is het belangrijk om te weten om welke stam het precies gaat. Een genetische barcode zal daarover echter geen uitsluitsel geven. In principe is identificatie op stamniveau mogelijk op basis van de DNA sequentie van meerdere genen, oftewel DNA fingerprinting. In de praktijk wordt ook veel gebruik gemaakt van andere technieken, variërend van klassieke biochemische analyses tot moderne *matrix-assisted laser desorption ionisation time of flight* (MALDI-TOF) massaspectrometrie of Fourier-transformatie infraroodspectroscopie (FT-IR).