

TOELICHTING M.B.T. MASSIVELY PARALLEL SEQUENCING

Algemene definities

PCR: Afkorting van Polymerase Chain Reaction, de chemische reactie waarmee specifieke delen van het DNA (in forensisch DNA-onderzoek meestal STRs) vele miljoenen malen worden gekopieerd. Deze kopieën worden PCR-producten genoemd.

Sequentie: Een DNA-profiel waarin de exacte volgorde van de bouwstenen (nucleotiden) van een stukje DNA is weergegeven. De nucleotiden worden afgekort tot **A** (Adenine), **T** (Thymine), **C** (Cytosine) en **G** (Guanine).

STR: Afkorting van Short Tandem Repeat; een gebied op het DNA waarin een bepaalde korte sequentie zich een aantal malen herhaalt (bv. **CCATCCATCCATCCAT**). Het aantal herhalingen (repeats) vertoont grote variatie tussen verschillende personen, resulterend in lengteverschillen welke allelen worden genoemd. Daarnaast kan de sequentie van de STR (meestal kleine) verschillen vertonen, bijvoorbeeld **CCATCCATCGATCCAT** in plaats van **CCATCCATCCATCCAT**. Ook rondom de STR kunnen dergelijke kleine verschillen optreden.

SNP: Afkorting van Single Nucleotide Polymorphism, een verandering van 1 nucleotide binnen een sequentie. Het hierboven gegeven voorbeeld van mogelijke variatie in de sequentie van een STR (een **G** in de derde repeat i.p.v. een **C**) betreft een SNP.

Verschillen tussen capillaire elektroforese (CE) en massively parallel sequencing (MPS) bij de bepaling van een (forensisch) DNA-profiel.

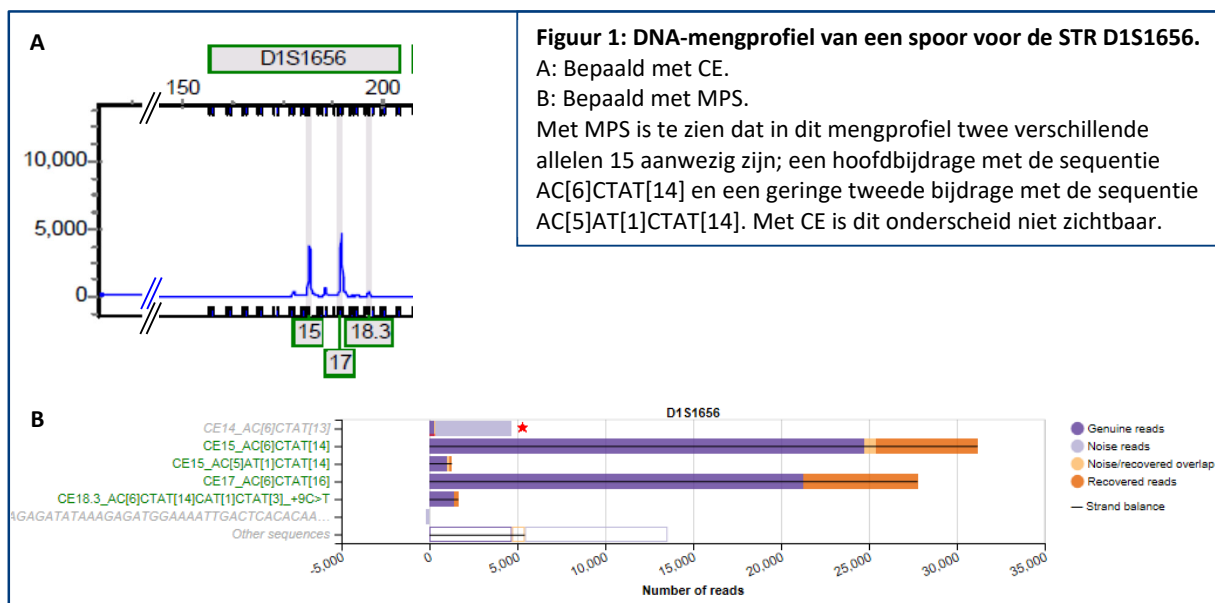
Groter onderscheidend vermogen (zie ook figuur 1).

Bij CE (de meest gebruikte techniek bij het genereren van een forensisch DNA-profiel) wordt alleen de lengte van de PCR-producten in kaart gebracht. Allelen met dezelfde lengte maar verschillende sequenties, kunnen met behulp van CE niet van elkaar worden onderscheiden. Bijvoorbeeld: de allelen **CCATCCATCGATCCAT** en **CCATCCATCCATCCAT** hebben beiden lengte 4 en zien er in een met CE bepaald DNA-profiel volledig identiek uit. De SNP die het verschil tussen deze allelen bepaalt, is met CE niet zichtbaar.

Bij MPS wordt naast de lengte ook de exacte sequentie van de PCR-producten zichtbaar gemaakt. Met behulp van MPS is het daarom ook mogelijk om verschillen tussen allelen van gelijke lengte zichtbaar te maken. De naamgeving van de allelen in een MPS-DNA-profiel verschilt om deze reden van de naamgeving van de allelen in een CE-DNA-profiel. Iedere allelnaam begint altijd met CEZZ, waarbij ZZ staat voor de lengte van het allel zoals deze in een CE-analyse waargenomen zou worden. De informatie daarna is een verkorte weergave van de exacte sequentie-variantie.

De MPS-allelnaam CE17_GGAA[11]GGCA[6]_-35C>A moet bijvoorbeeld als volgt worden gelezen:

- CE17 = In een CE-DNA-onderzoek wordt dit allel benoemd als allel 17.
- GGAA[11]GGCA[6] = De repeatstructuur van dit allel bestaat uit 11 maal GGAA, gevolgd door 6 maal GGCA.
- -35C>A = Op een afstand van 35 nucleotiden voorafgaand aan de STR is een A aanwezig in plaats van de C die daar, volgens de informatiebronnen m.b.t. deze STR, aanwezig zou moeten zijn.



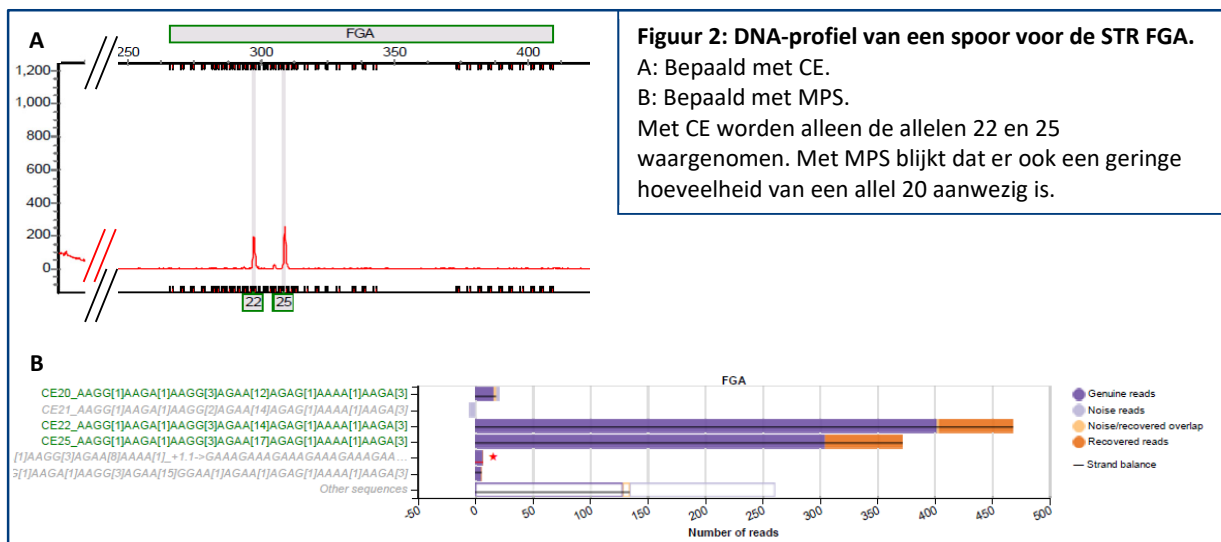
Betere inschatting van het minimale aantal donoren (zie ook figuur 2).

Bij CE worden de PCR-producten voorzien van fluorescerende labels in verschillende kleuren (ieder PCR-product krijgt 1 label). Deze labels worden vervolgens met behulp van een laser en lichtsensor gedetecteerd en omgezet in een DNA-profiel: Iedere keer dat de lichtsensor fluorescentie van een bepaalde kleur heeft waargenomen, komt overeen met een piek in het CE-DNA-profiel.

De fluorescentie moet echter wel van een bepaalde intensiteit zijn voordat deze door de lichtsensor waargenomen kan worden. Niet ieder fluorescerend label heeft een gelijkwaardige intensiteit, waardoor het kan voorkomen dat de lichtsensor pas licht waarneemt wanneer een groot aantal labels van één kleur gelijktijdig fluoresceren. In de praktijk kan dit betekenen dat bijvoorbeeld 500 PCR-producten van allel N (en dus 500 labels die gelijktijdig fluoresceren) nog steeds onvoldoende licht produceren om door de sensor te worden waargenomen. Het uiteindelijke DNA-profiel bevat in dat geval geen piek voor allel N, terwijl dit allel wel na de PCR aanwezig was.

De bij MPS gebruikte detectiemethode heeft geen gevoeligheidsgrens. Daardoor worden alle allelen die na de PCR aanwezig zijn, ook daadwerkelijk in het DNA-profiel opgenomen.

Hierdoor kan het voorkomen dat een geringe bijdrage van DNA van een persoon met CE niet boven de detectiegrens uit komt, maar met MPS wel wordt waargenomen.



Betere inschatting van mengverhoudingen.

Mede doordat de lichtsensor bij CE een gevoeligheidsgrens heeft en niet alle labels een identieke lichtintensiteit hebben, geeft de piekhoogte bij CE (uitgedrukt in relatieve fluorescente eenheden, RFU) enkel een grove indicatie van het onderliggende aantal PCR-producten. In geval van een sterk fluorescerend label kan een piek van bijvoorbeeld 500 RFU daadwerkelijk gelijk staan aan ongeveer 500 PCR-producten, terwijl voor 500 RFU bij een zwak fluorescerend label mogelijk wel 1500 of meer PCR-producten nodig zijn.

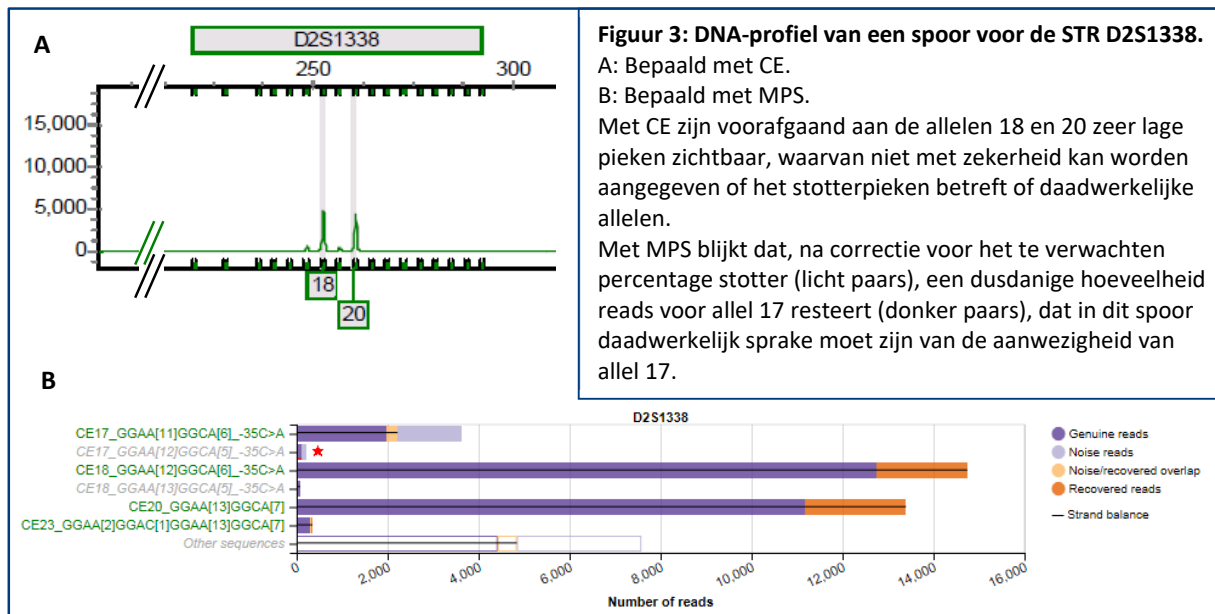
Bij MPS wordt ieder PCR-product apart geteld als 1 read. Indien in een MPS-DNA-profiel dus bijvoorbeeld staat dat voor een allel 20.000 reads zijn waargenomen, houdt dit in dat er daadwerkelijk 20.000 PCR-producten van dat allel zijn waargenomen.

Beter onderscheid tussen stotter-artefacten en daadwerkelijke allelen (zie ook figuur 3).

Stotter-artefacten zijn 'fouten' die optreden bij het uitvoeren van een PCR-reactie voor STRs: Bij het vermeerderen van de repeat ontstaan een klein aantal PCR-producten die één repeat korter (-1 stotter) of langer (+1 stotter) zijn dan het daadwerkelijk aanwezige allel. Over het algemeen komt -1 stotter vaker voor dan +1 stotter.

Doordat bij MPS het aantal reads overeenkomt met het daadwerkelijke aantal aanwezige PCR-producten is het mogelijk om voor iedere STR te bepalen in welke percentages stotter-allelen worden waargenomen ten opzichte van het daadwerkelijke allel. Om dit te bepalen zijn van een groot aantal referentiepersonen MPS-DNA-profielen gegenereerd, waarna van ieder allel in ieder MPS-DNA-profiel het percentage stotter is bepaald.

Deze waarden zijn opgenomen in de software die wordt gebruikt bij de analyse van MPS-DNA-profielen. Indien een allel wordt waargenomen wat op een -1 of +1 positie ligt ten opzichte van een ander allel voor dezelfde STR, wordt door de software een correctie toegepast waarbij het gemiddelde percentage stotter van het waargenomen aantal reads wordt afgetrokken.



Belangrijk

De Nederlandse DNA-databank voor strafzaken en andere DNA-databanken bevatten enkel CE-DNA-profielen. Dit komt doordat tot enkele jaren geleden enkel CE beschikbaar was voor forensisch DNA-onderzoek en doordat de software welke door de DNA-databanken wordt gebruikt nog geen MPS-DNA-profielen kan verwerken. Onder andere om deze reden wordt MPS enkel in uitzonderlijke gevallen door het FLDO gebruikt.

Een MPS-DNA-profiel bevat altijd, zoals ook aangegeven onder Groter onderscheidend vermogen, de lengte-informatie zoals deze met behulp van CE zou zijn verkregen.

Daardoor is het mogelijk om de lengte-informatie uit MPS-DNA-profielen aan te bieden voor opname of een eenmalige zoekactie in de van toepassing zijnde DNA-databank.

Daarnaast is het mogelijk om CE-DNA-profielen te vergelijken met de lengte-informatie uit MPS-DNA-profielen. Wanneer MPS-DNA-profielen van sporen beschikbaar zijn, maar van de referentiepersonen enkel CE-DNA-profielen beschikbaar zijn, kan op basis van een dergelijke vergelijking een eerste screening worden uitgevoerd m.b.t. mogelijk donorschap:

- Als het CE-DNA-profiel van een referentiepersoon niet overeenkomt met de MPS-DNA-profielen van de sporen, is het niet nodig om van deze persoon een MPS-DNA-profiel te bepalen, want ook op basis van MPS-DNA-profielen zal deze persoon dan geen overeenkomsten met het spoor laten zien.
- Als het CE-DNA-profiel van een referentiepersoon wel (een groot aantal) overeenkomsten vertoont met de MPS-DNA-profielen van de sporen, kan het zinvol zijn om een MPS-DNA-profiel van deze persoon te bepalen. Dit geldt zeker in geval van partiële DNA-profielen of DNA-mengprofielen. Tenslotte is het mogelijk dat een persoon weliswaar allelen van dezelfde lengte bezit als de allelen die in het spoor zijn waargenomen, maar dat de sequentie van deze allelen verschilt.