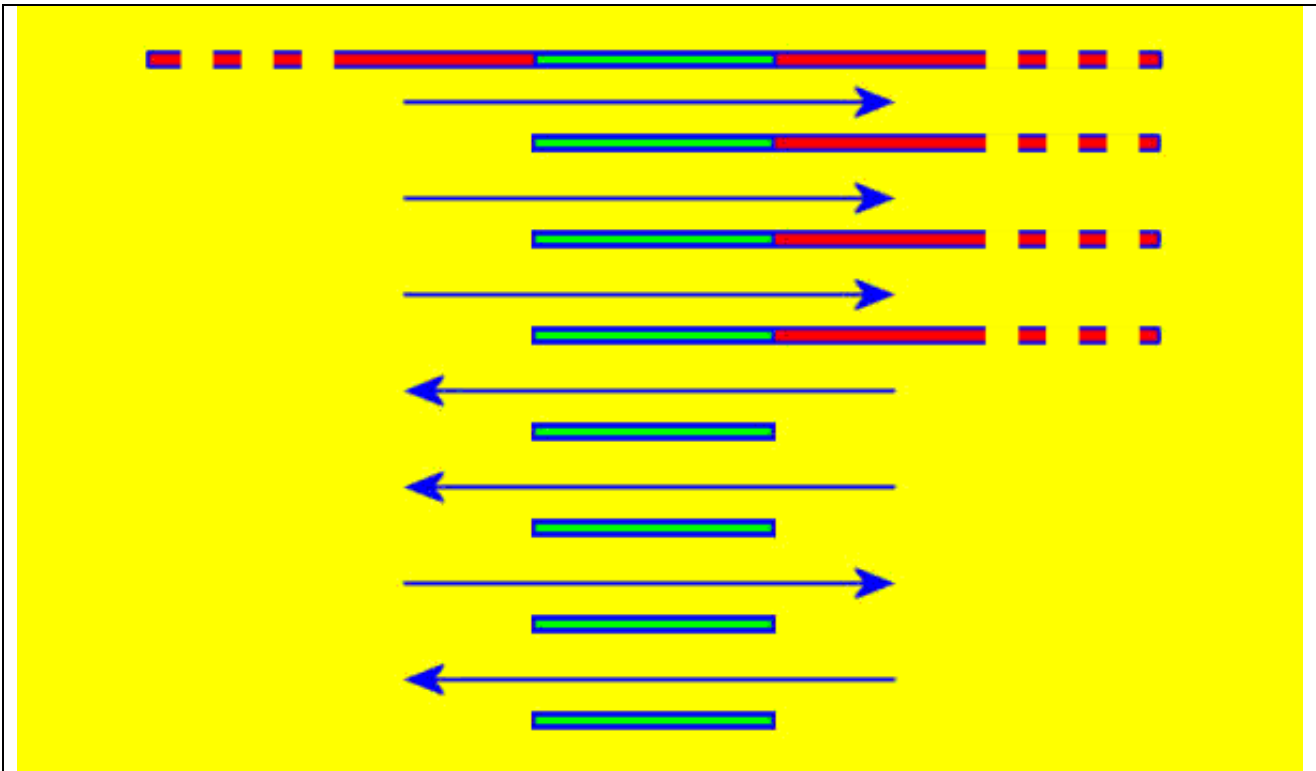


Om dat te verduidelijken moeten we eerst de PCR onder de loep nemen.

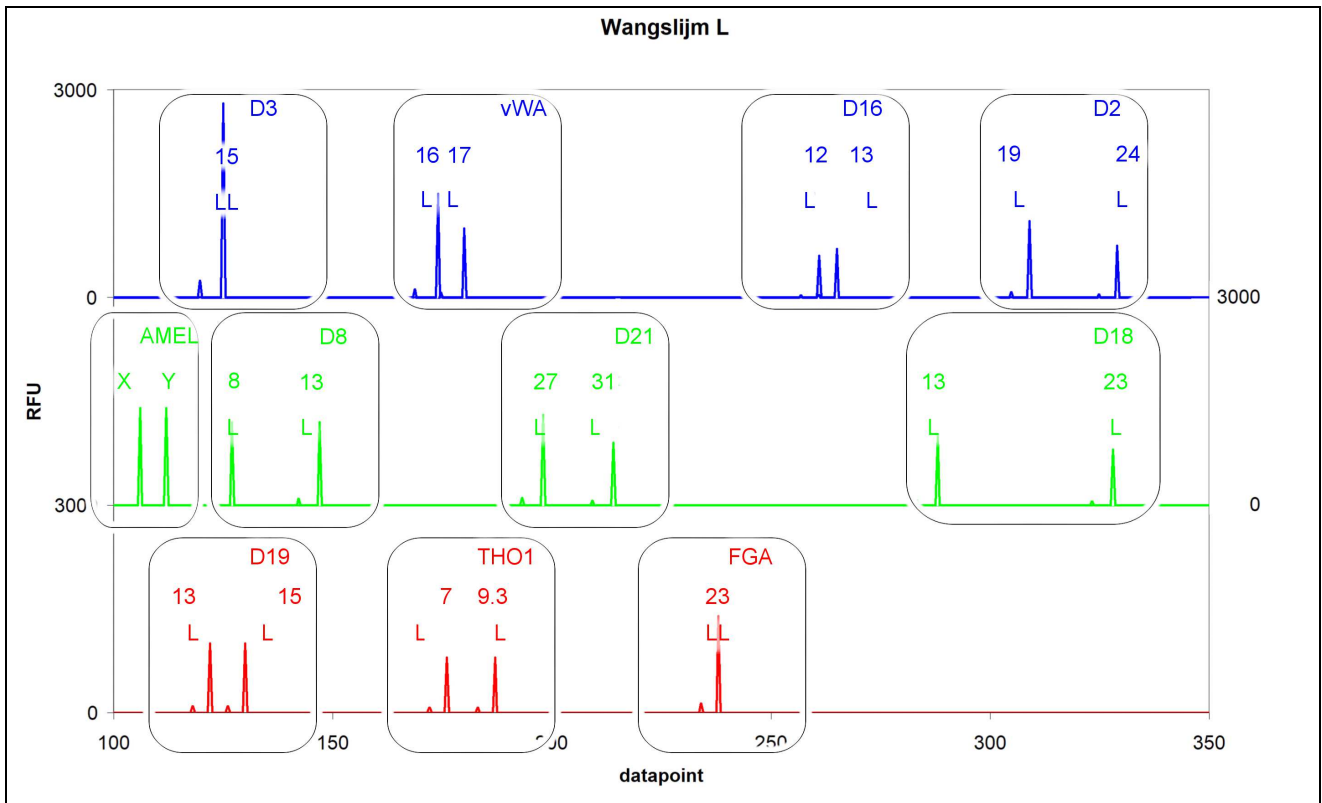


Het kopiëren van een DNA-fragment voor forensisch onderzoek (STR schematisch in groen, rest van het chromosoom in rood) verloopt omslachtig; eerst wordt een kopie gemaakt vanaf de scheiding tussen rood en groen vanaf één van de uiteinden van de STR, maar de andere helft van het chromosoom gaat dan gewoon mee in de kopie. Pas als in één van de vervolgstappen de kopieerbeweging van de tegenovergestelde kant komt, wordt de beoogde STR geïsoleerd. Hoeveel stappen dat vergt is dan onzeker. Als maar kleine aantallen van hetzelfde chromosoom aanwezig is, kan dat leiden tot onvoorspelbare uitkomsten wat betreft te verwachten kopiëren.

PCR voor forensische doeleinden bestaat uit het kopiëren van een klein stukje DNA-keten (STR = short tandem repeat, hier in het groen) dat verstopt zit in een chromosoom, een complete DNA-keten (hier rood - groen - rood). De verhoudingen zijn veel extremer dan dit diagram suggereert; zo'n STR is rond de 200 eenheden (DNA-monomeren) lang, de lengte een chromosoom ligt in de orde van grootte van 100 miljoen eenheden. Voor de forensische analyse wordt een stukje DNA (STR) geselecteerd dat twee eigenschappen toont:

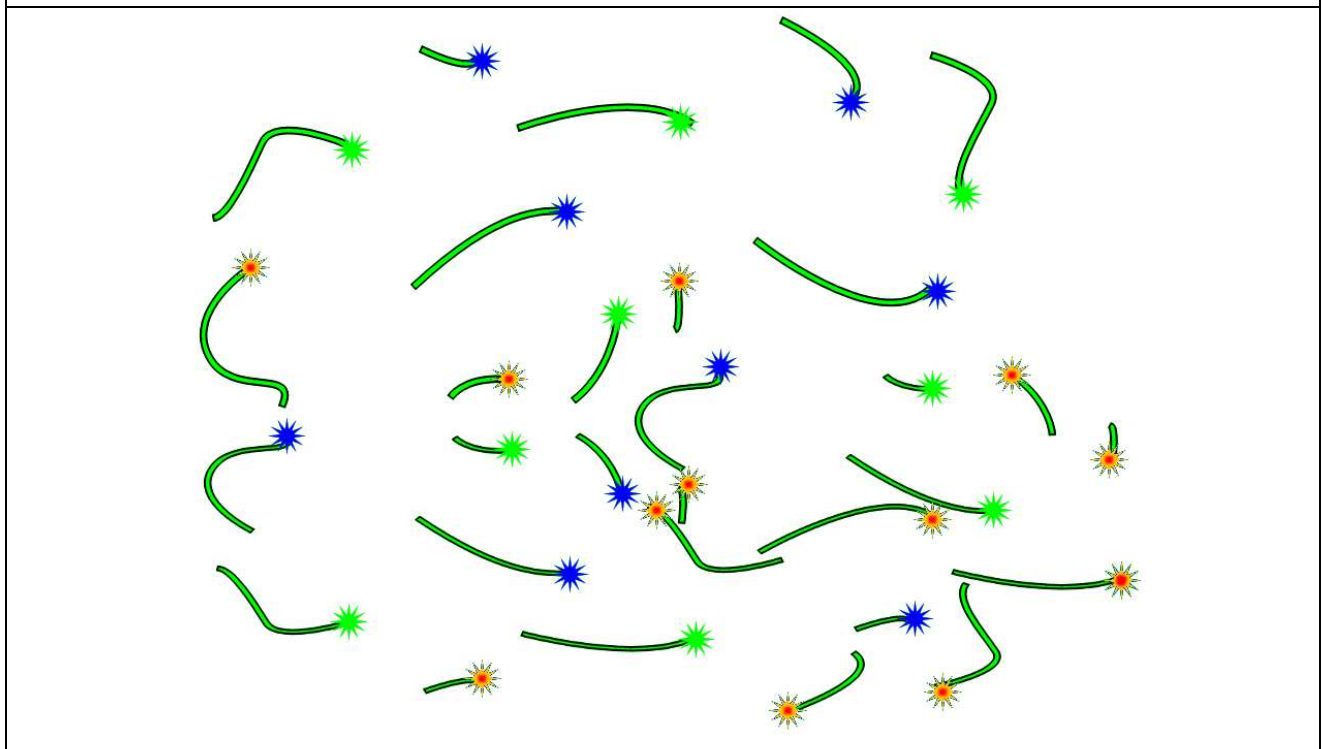
1. Het bestaat uit een vrij monotone herhaling van groepjes van doorgaans vier DNA-monomeren (de zgn. tandems, tandems kunnen hier dus ook uit vier eenheden bestaan).
2. Het aantal van die herhalingen is per individu juist zeer verschillend.

Een voorbeeld van een dergelijke STR-locatie is TH01. Deze varieert van  $n=3$  tot en met  $n=12$  tandems met als samenstelling  $[AATG]_n$ ; daarnaast worden op deze locatie ook nog een aantal minder uniforme varianten gevonden. Door het vaststellen van de lengte kan men het DNA van verschillende individuen van elkaar onderscheiden, mits men maar genoeg van deze STR-locaties gebruikt. Men gebruikt in verschillende systemen 7 tot 20 van deze locaties (loci). Omdat een celkern van ieder chromosoom twee exemplaren bezit (overerfd van de ouders), levert dit in principe 14 tot 40 verschillende STR's op per cel. Deze komen dus in paren, zie het referentieprofiel van Louwes' wangslim hieronder. Maar soms wordt er maar één gevonden; de lengtes van beide STR's op een bepaald chromosomenpaar kunnen gelijk zijn (homozygoot versus heterozygoot). Dat zien we per individu zo'n twee tot zes keer gebeuren. Dat levert overigens wel bruikbare informatie, we zien dan een dubbel zo sterke uitkomst van de meting van zo'n homozygote STR. In dit geval zien we dat op de loci D3-15 en FGA-23 optreden.



Boven: de drie - gekleurde - signalen (blauw, groen en geel) die aan het eind van iedere bepaling worden opgevangen, worden in de overeenkomstige kleuren geplot in een profiel. Om voor de hand liggende redenen wordt geel vervangen door rood. De - hier elf - verschillende locaties (of loci) in het profiel zijn benadrukt.

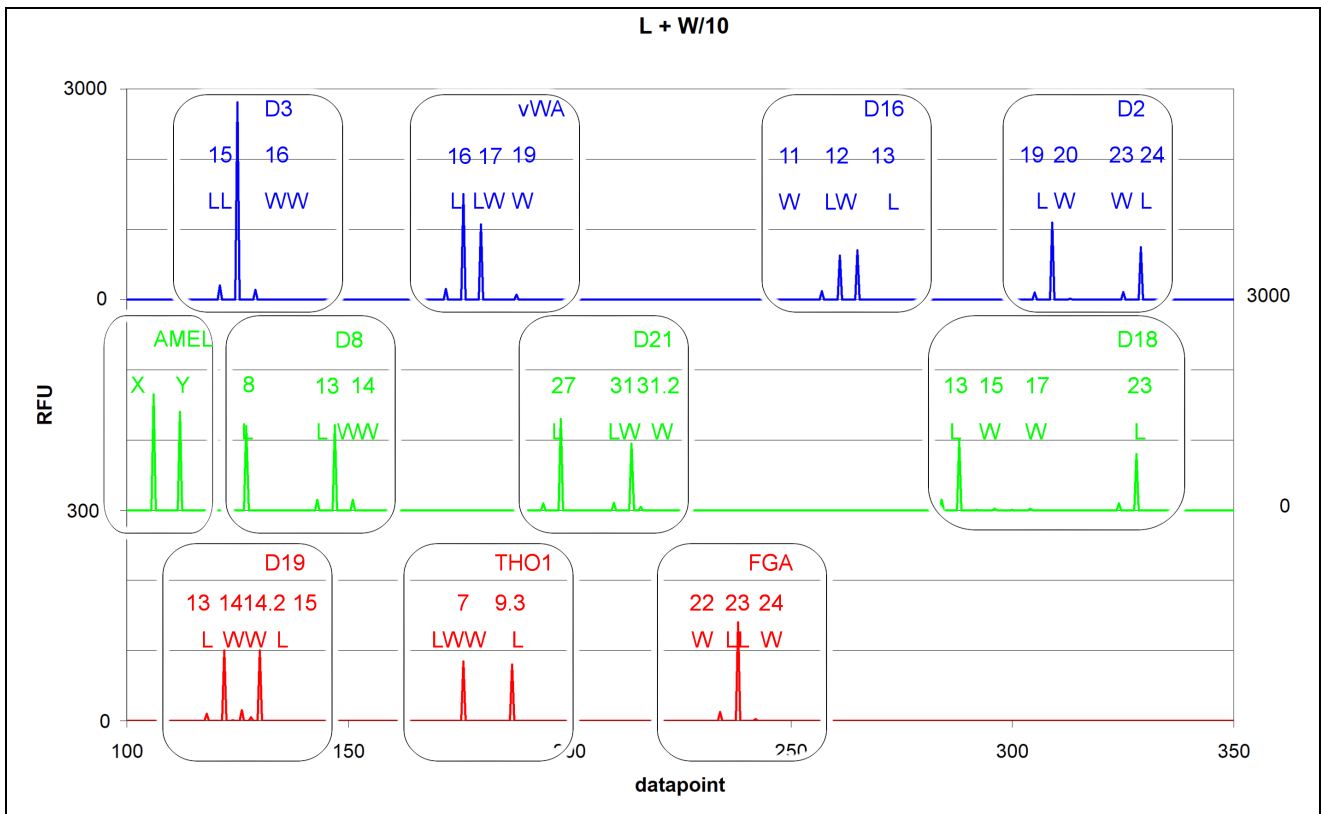
Onder: dit resultaat wordt verkregen met een slimme truc: tijdens het kopiëren van de STR's wordt een speciale groep aan de STR gekoppeld. Voor de locaties D3 t/m D2 geeft die blauw licht af, voor de locaties D8 t/m D18 en de geslachtskenmerken groen licht en de overige geel licht. Hierdoor is het mogelijk het profiel te verdelen over drie verschillende diagrammen, waardoor de analyse veel compacter kan worden uitgevoerd.



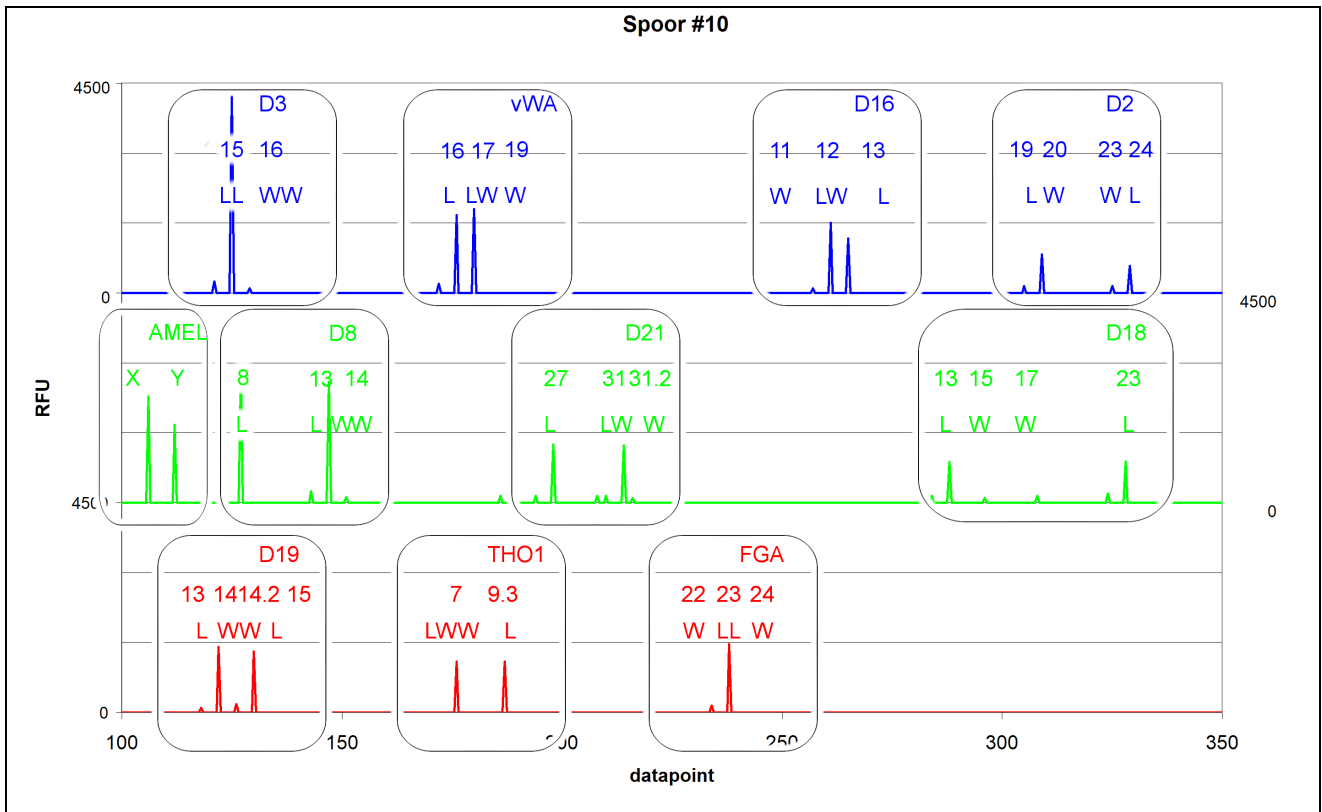
Als meerdere personen hun DNA op dezelfde plek hebben achtergelaten lopen de profielen toch door elkaar en vertonen ze meer pieken. Maar ook komt het voor dat pieken uit de oorspronkelijke profielen samensmelten.

Daar komt een lastige complicatie bij. Het hiervoor geproduceerde profiel toont namelijk veel meer pieken dan ik voorspelde. Hier zien we 20 pieken die vrij groot zijn met twee daarvan dubbel zo groot, want homozygoot. Dus eigenlijk 22 pieken, zoals de theorie hier voorschrijft (er worden twee geslachtskenmerken [X-Y] meegenomen in de analyse) voor het zogenaamde SGM-plus systeem dat het NFI in 2003 toepaste. Maar er zitten ook nog 16 snippertjes in, eigenlijk slippertjes. Bij het kopiëren ontstaan die slippertjes, in het jargon worden ze stutters genoemd. Deze stutters vinden we voortdurend aan de linker kant van iedere STR-piek.

Als nu een tweede persoon DNA bijdraagt en die persoon veel minder DNA heeft achtergelaten - in de orde van hoeveelheid als de genoemde stutters - wordt het heel ingewikkeld de DNA-bijdragen van die tweede persoon te onderscheiden van de pieken en stutters in het profiel van de eerste persoon. Het kan tegenwoordig wel met krachtige computeranalyses (zgn. continuous forensic DNA mixture interpretation), maar die worden pas een jaar of tien toegepast, dus ver nadat de analyses in de Deventer Moordzaak werden gemaakt.

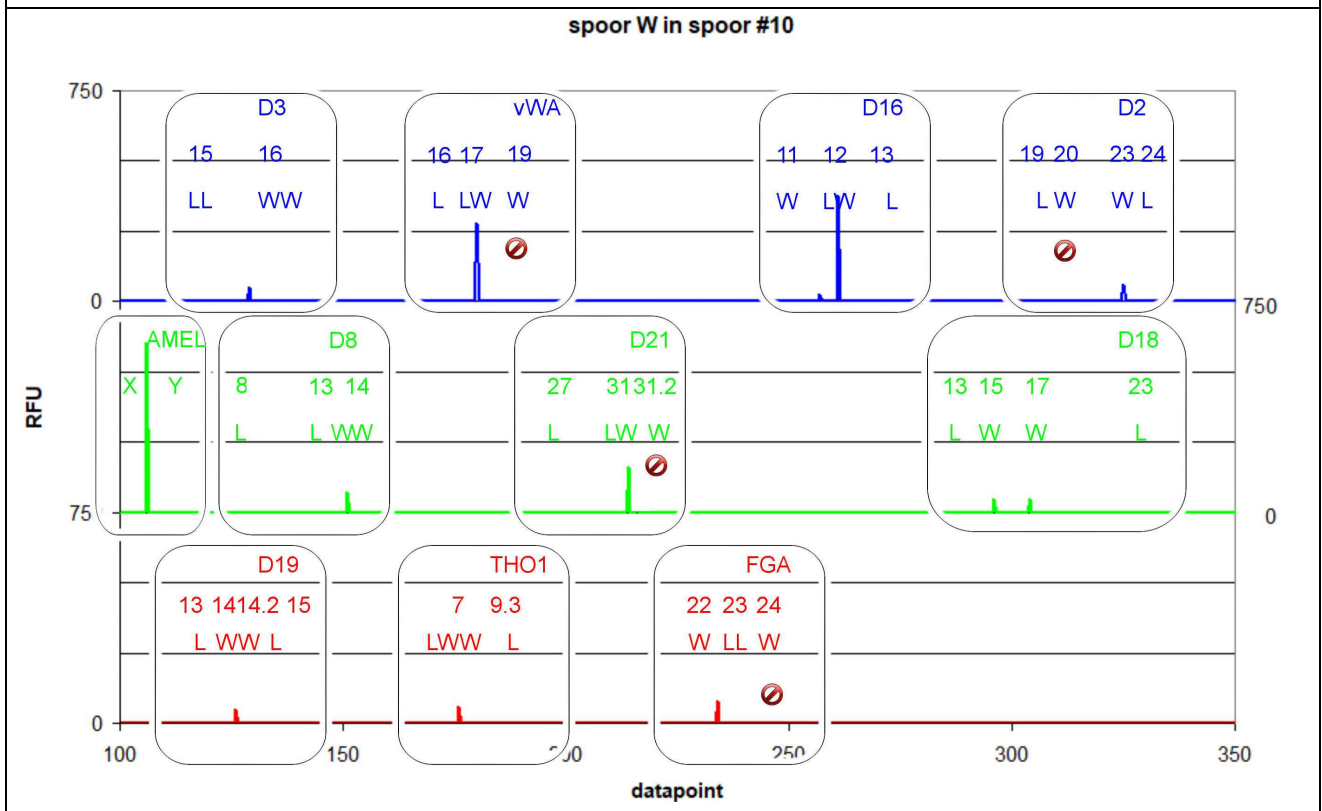


Hier zien we het resultaat van het samengaan van twee profielen op de wijze, zoals hierboven geschetst. Het resultaat hiervan is een profiel dat als twee druppels water lijkt op het resultaat dat onder vuur ligt, namelijk het profiel van spoor #10 dat door het NFI alleen aan Louwes werd toegedicht. Vrijwel alle kenmerken van W zijn op de een of andere manier - losstaand of samengesmolten - in het profiel van spoor #10 aanwijsbaar, waardoor de bewijswaarde van dit profiel ondanks de mening van NFI en OM nihil is. Omdat de sterkte van de referentieprofielen arbitrair is, behoeft de verticale as niet gelijk te zijn aan die van spoor #10.



Boven: het enige opvallende verschil tussen het voorgaande mengprofiel en het profiel van spoor #10 is de aanwezigheid van enige piekjes in de 'groene straat' die noch aan Louwes, noch aan Wittenberg kunnen worden toegeschreven en mogelijk op een profiel van de echte dader zouden kunnen wijzen.

Onder: Het is ook mogelijk het referentieprofiel van Louwes in mindering te brengen op het profiel in spoor #10. Normaal wordt dit tegenwoordig uitgevoerd via een computerprogramma. Het hieronder gepresenteerde profiel is op soortgelijke wijze handmatig verkregen en bevat het vrijwel volledige profiel van mevrouw Wittenberg (de ontbrekende pieken zijn gemarkeerd; het voorkomen van enige drop-outs is bij een dergelijk zwak profiel normaal.).



## MPS betreedt het toneel

Intussen zijn dit soort STR-profielen geheel gebaseerd op de eigenschap van de STR's dat ze veel uiteenlopende lengtes vertonen. Maar hoe zit het met die eerste eigenschap van de STR's, de monotone herhaling van een uniforme samenstelling? Laten we eerst nog eens een voorbeeld van die samenstelling bekijken; hier volgt de STR met de afkorting FGA met een lengte van 15 repeats. Een enkele repeat bestaat steeds uit vier DNA-monomeren; hier hebben we er 15, hetgeen de gehele lengte dus feitelijk brengt op 60 eenheden, maar we gebruiken hier altijd de waarde 15:

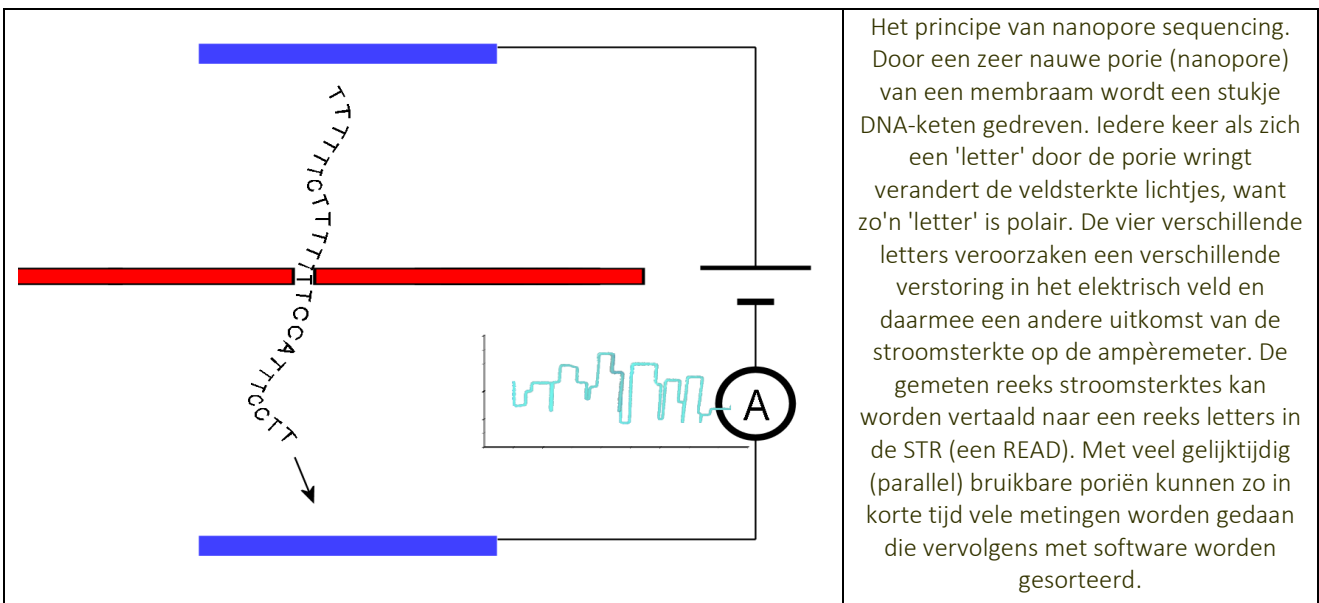
[TTTC]<sub>3</sub> TTTT TTCT [CTTT]<sub>7</sub> TTCC]<sub>2</sub>

De genoemde lengte van 60 wordt verder nooit genoemd, de werkelijke lengte van een STR is namelijk groter doordat tijdens het kopiëren nog meer DNA wordt toegevoegd, zij het steeds voor elke STR op zo'n locatie. Er bestaan maar vier soorten monomeren die elk gezegend zijn met een eigen letter: A, T, C, G. Veelal zijn er in dit soort repeats maar twee in gebruik, maar de eerder genoemde TH01 heeft er al drie.

Zoals hier te zien is zo'n repeat niet altijd zo monotoon als het voorbeeld TH01. Maar nu komt het: ook deze volgorde van letters is niet bij ieder individu hetzelfde, ook al is de STR even lang en op dezelfde locatie. Zelfs bij eenenige tweelingen zit er toch nog wel ergens een verschilletje - kwestie van een andere letter. Zo'n verschilletje heet een mutatie, meer specifiek een single-nucleotide polymorphism (SNP), dus vrijwel letterlijk: één verwisselde letter. Als we nu eens aanvullend van iedere gemeten STR de echte volgorde inclusief de opgetreden mutaties konden opmeten...

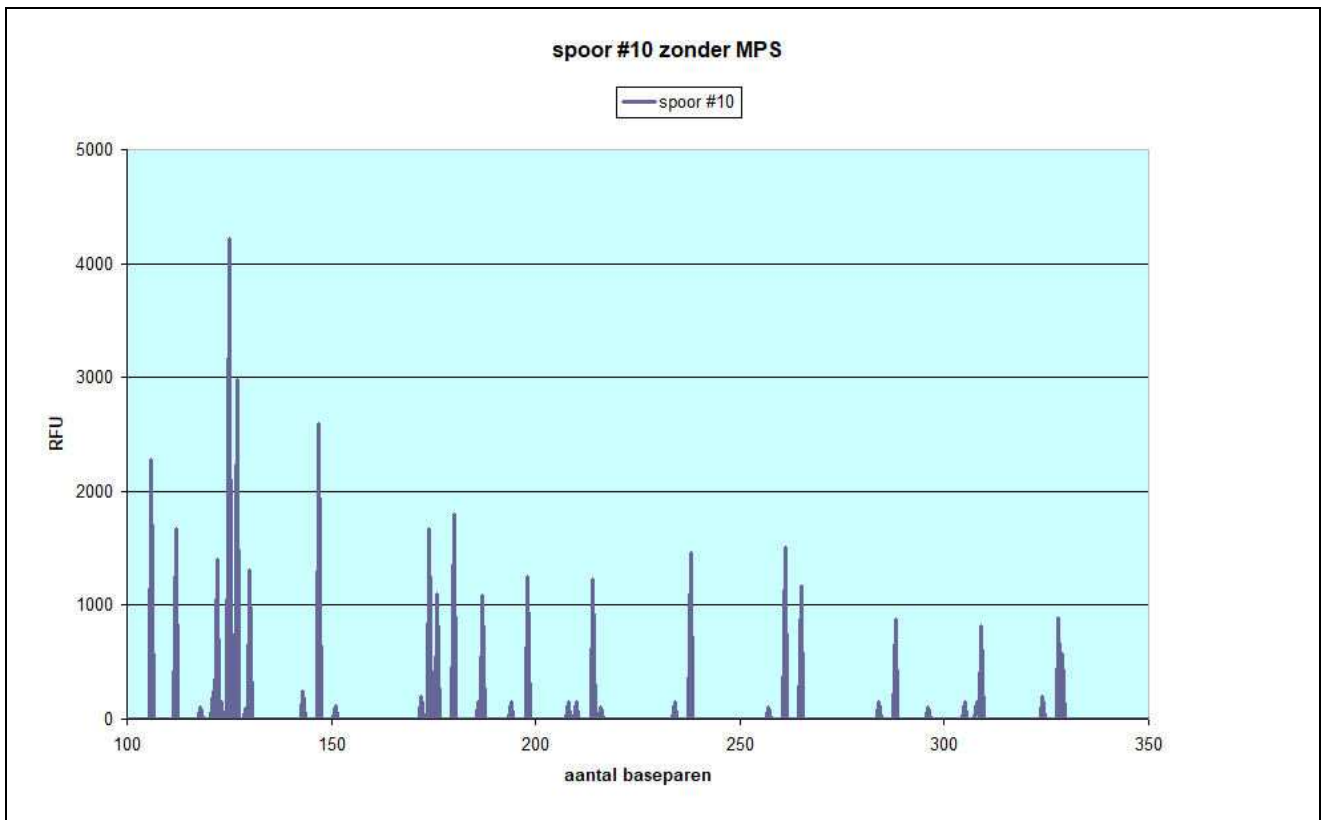
Nu dat kan gewoon, met Massive(ly) Parallel Sequencing. Het vaststellen van een bepaalde lettervolgorde = sequencing, is allang een standaardtechniek in de medische wereld met vele varianten en een enorme ontwikkeling in de tijd. Het wordt steeds routinematiger en ... goedkoper.

Het bijzondere is dat voor gebruik in medische toepassingen de chromosomen eerst in stukjes worden gehakt, vervolgens de brokstukjes worden uitgelezen en het resultaat (de letters) in een computer gebracht. Deze moet dan de resultaten (sliertjes letters) weer op een logische wijze aan elkaar breien: shotgun sequencing. Dit hoeft in de forensische setting helemaal niet, want de stukjes hebben we al dat zijn de STR's, verkregen tijdens de PCR!



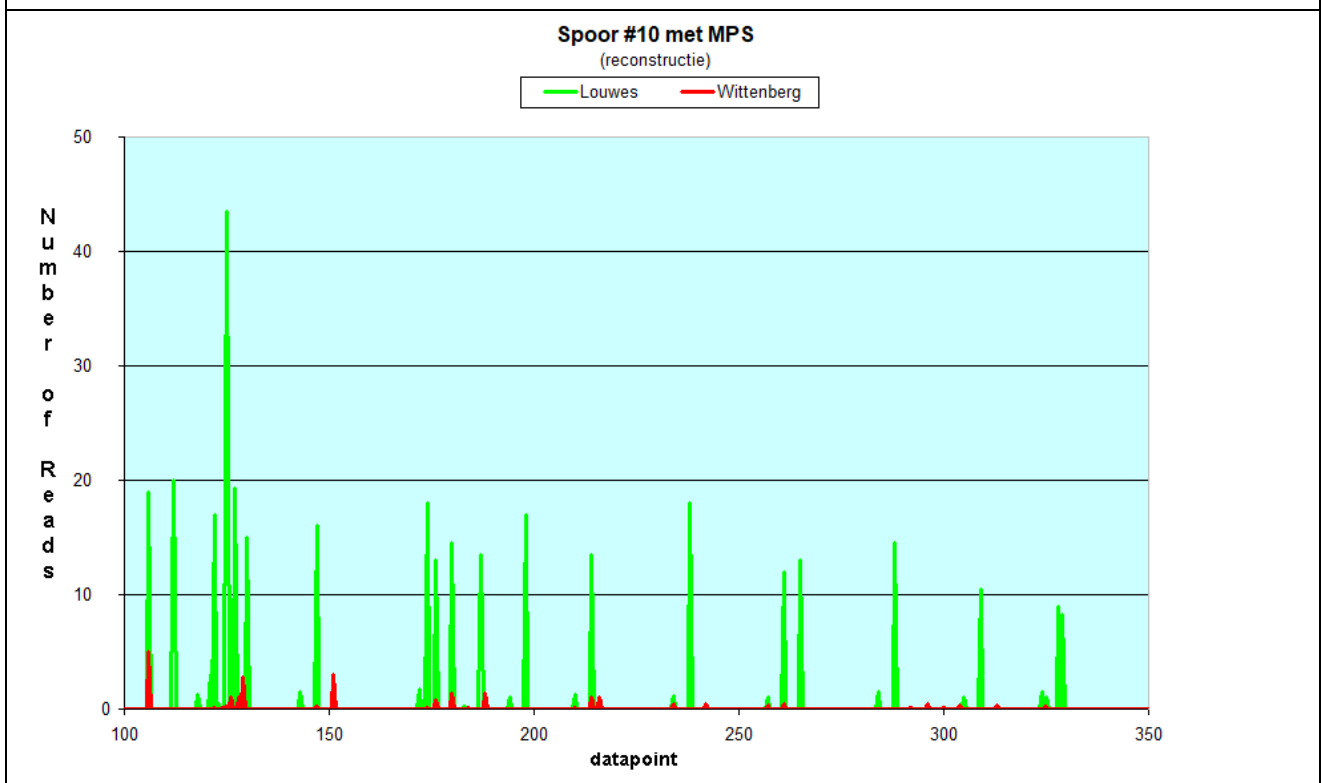
En omdat dit wordt uitgeoefend op overzichtelijk kleine fragmentjes DNA, is het relatief een peulenschil. Zodoende kunnen per onderscheidenlijke STR (zoals bijvoorbeeld FGA-15, als deze erbij zit) nu twee zaken in één keer bepaald worden na de nodige foutcorrecties (onvolledige READS en eenmalige afwijkingen bijvoorbeeld):

1. Hoeveel verschillende STR's (= READS) van dezelfde lengte en zelfde locatie naast elkaar in het profiel zitten, ofwel hoeveel individuen op die plek aan het profiel hebben bijgedragen
2. met welke frequentie die STR's (READS) daar voorkomen.



Boven: zo ziet een profiel eruit, als men alleen de aantallen STR-lengtes benut als informatiebron. Het is een verzameling van tientallen anonieme pieken.

Onder een reconstructie van datzelfde profiel, waarin het aantal READS dat ieder individu oplevert in een bepaald monster in een andere kleur is geplot. Deze reconstructie is gebaseerd op de analyse van spoor #10 gebaseerd op een handmatige toepassing van continuous interpretation, rekening houdend met stutterwaarden en piekbalansen. Er worden nu automatisch meerdere profielen zichtbaar als meerdere personen een bijdrage hebben geleverd.



Omdat dit op alle locaties wordt uitgevoerd kan hiermee met een grote mate van zekerheid een conclusie worden getrokken omtrent het aantal personen dat aan het hele profiel heeft bijgedragen en in welke mate dat heeft plaatsgevonden. Mengsels tot een 99-1 verhouding bleek correcte resultaten op te leveren (Gaag, van der et al. 2016)

Intussen is al nagegaan in hoeverre de resultaten die langs deze weg worden verkregen overeenstemmen met de resultaten die met de eerder genoemde meer digitale methoden - de continuous forensic DNA mixture interpretation - worden verkregen (NFI, FLDO en UvA; Benschop et al. 2021). Het blijkt dat de beide methoden sterk overeenkomstige uitkomsten bieden, zeker als men dit vergelijkt met de resultaten van de ouderwetse zgn. binaire aanpak die voorheen werd toegepast.

Indien we zouden willen, zouden we een computer gewoon kunnen programmeren om in plaats van één profiel van het spoor, meerdere profielen te plotten, voor iedere persoon die bijdraagt één. Inmiddels zijn de eerste resultaten van een dergelijke aanpak al te signaleren, cf. Hoogenboom et al. 2017, Bennett et al. 2018.

## Conclusies

Op basis van 1. kan nu haarzuiver worden bepaald, hoeveel personen aan een profiel bijdragen en op basis van 2. kan ook het volume van hun bijdragen worden vastgesteld.

## Naschrift

Om het fenomeen MPS uiteen te zetten moet ik steeds balanceren op het scherp van de snede tussen begrijpelijkheid en volledigheid. Daarom zijn er aan de volledigheid flinke concessies gedaan. Maar wellicht ook aan de begrijpelijkheid.

## Literatuur

Bennet et al. 2019. Mixture deconvolution by massively parallel sequencing of microhaplotypes. <https://doi.org/10.1007/s00414-019-02010-7>

Benschop et al. 2021. Application of a probabilistic genotyping software to MPS mixture STR data is supported by similar trends in LR compared with CE data. doi:10.1016/j.fsigen.2021.102489

Bruijns et al. 2018. Massively parallel sequencing techniques for forensics: A review. <https://doi.org/10.1002/elps.201800082>

Gaag, van der et al. 2016. Massively parallel sequencing of short tandem repeats—Population data and mixture analysis results for the PowerSeq™ system. [https://www.fldo.nl/artikelen/van%20der%20Gaag%20et%20al,%20Forensic%20Sci%20Int%20Genet%2024%20\(2016\)%20p.86-96.pdf](https://www.fldo.nl/artikelen/van%20der%20Gaag%20et%20al,%20Forensic%20Sci%20Int%20Genet%2024%20(2016)%20p.86-96.pdf)

Hoogenboom et al. 2017. FDSTools: A software package for analysis of massively parallel sequencing data with the ability to recognise and correct STR stutter and other PCR or sequencing noise. [https://www.fldo.nl/artikelen/Hoogenboom%20et%20al,%20Forensic%20Sci%20Int%20Genet%2027%20\(2017\)%20p.27-40.pdf](https://www.fldo.nl/artikelen/Hoogenboom%20et%20al,%20Forensic%20Sci%20Int%20Genet%2027%20(2017)%20p.27-40.pdf)

Knijff, Peter de. 2019. From next generation sequencing to now generation sequencing in forensics. [https://www.fldo.nl/artikelen/de%20Knijff,%20Forensic%20Sci%20Int%20Genet%2038%20\(2019\)%20p.175-180.pdf](https://www.fldo.nl/artikelen/de%20Knijff,%20Forensic%20Sci%20Int%20Genet%2038%20(2019)%20p.175-180.pdf)