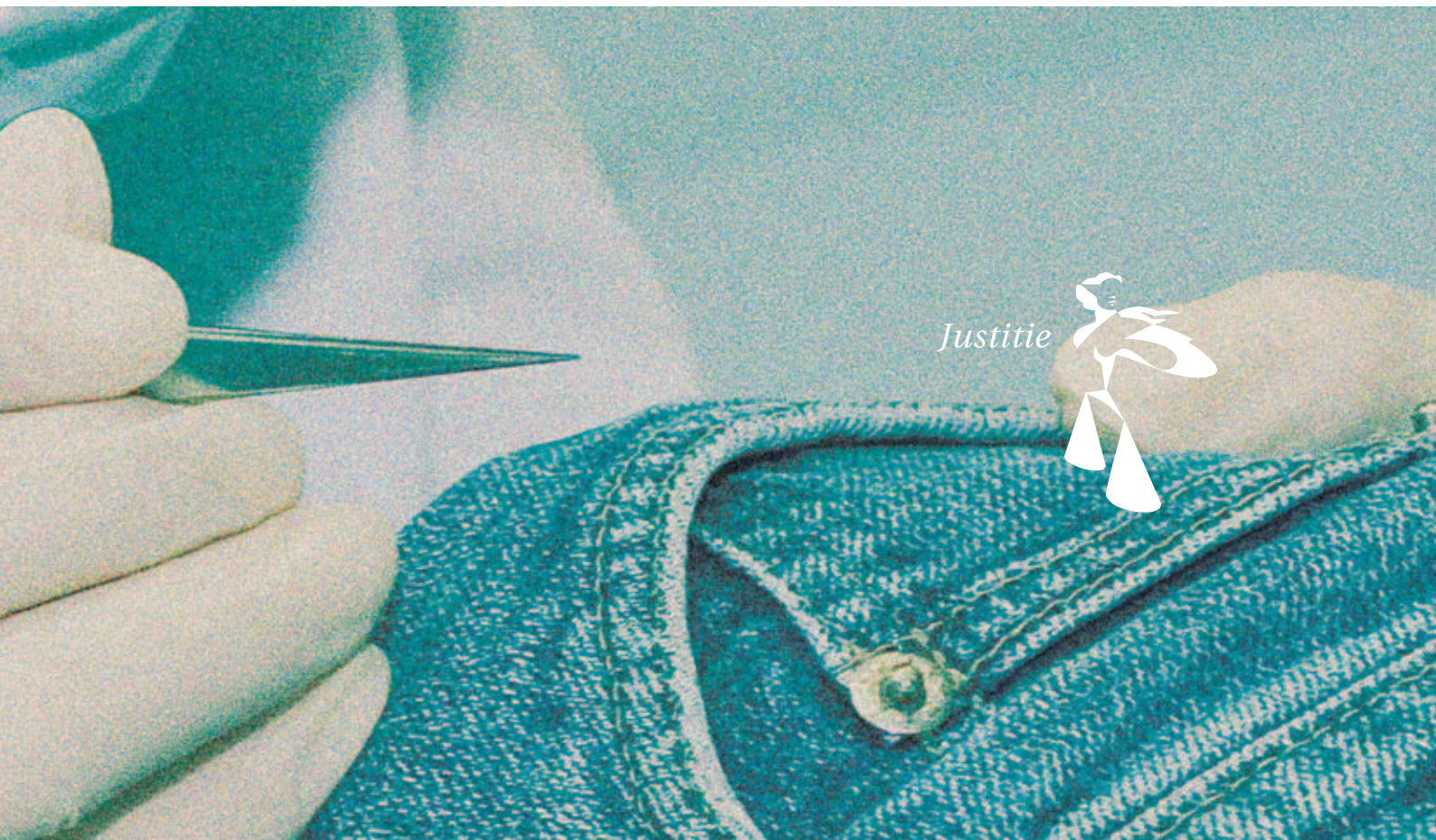


NEDERLANDS FORENSISCH INSTITUUT

De Essenties van forensisch DNA-onderzoek

6 Interpretatie van DNA-bewijs I
match en berekende frequentie



© 2007 Nederlands Forensisch Instituut

Alle rechten voorbehouden. Niets uit deze uitgave mag worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen, of enige andere manier, zonder voorafgaande schriftelijke toestemming van het Nederlands Forensisch Instituut.

Voor meer informatie

Nederlands Forensisch Instituut (NFI)

Bezoekadres

Laan van Ypenburg 6
2497 GB Den Haag
Telefoon (070) 888 66 66
Fax (070) 888 65 55

Frontdesk

Telefoon (070) 888 68 88

Postadres

Postbus 24044
2490 AA Den Haag

Emailadres

EssentiesDNA@nfi.minjus.nl

Inhoudsopgave

Drie mogelijkheden	4
Matchende DNA-profielen (kader)	5
Berekende frequentie	5
Kleiner dan één op één miljard	6
Statistische onderbouwing	7
Correctie voor subgroepen	7
Correctie voor grootte referentiedatabestand	8
Referentiedatabestand	8
Representatief	8
'Match' en 'berekende frequentie'	10
Conclusie bij gelijke volledige DNA-profielen	10
Geen absolute herkomstuitspraak	12
Wel uitsluiten	13
Vuistregel 1	13
Omgaan met de kans op fouten	14

Interpretatie van DNA-bewijs I

match en berekende frequentie

DNA-profielen kunnen een zeer sterke aanwijzing geven over de herkomst van een biologisch spoor. De zeldzaamheidswaarde van een DNA-profiel wordt uitgedrukt in de berekende frequentie van voorkomen van het DNA-profiel in de relevante populatie. Deze frequentie is voor een volledig DNA-profiel altijd kleiner dan één op één miljard. Als een volledig DNA-profiel van een verdachte of slachtoffer gelijk is aan dat van een spoor, is de kans dat deze match op toeval berust daardoor buitengewoon klein. Volledige DNA-profielen hebben daarom een zeer hoge identificerende waarde.

Drie mogelijkheden

Voor het beantwoorden van de vraag naar de herkomst van een biologisch spoor maakt het Nederlands Forensisch Instituut (NFI) gebruik van de DNA-technologie, die er op is gericht DNA-profielen te verkrijgen van celmateriaal. Het verkregen DNA-profiel van een biologisch spoor kan men vergelijken met DNA-profielen van andere biologische sporen of referentiemonsters van personen. Een voorwaarde voor vergelijkend DNA-onderzoek is dat het verkregen DNA-profiel hiervoor geschikt is. Dit hangt af van verschillende factoren.

In grote lijnen resulteert het DNA-onderzoek van een biologisch spoor in één van de volgende drie mogelijkheden:

- 1 Een DNA-profiel dat matcht met aan het DNA-profiel van een ander spoor of persoon in de zaak of in de Nederlandse DNA-databank voor strafzaken.
- 2 Een DNA-profiel dat niet matcht met een DNA-profiel van een ander spoor of persoon in de zaak of in de Nederlandse DNA-databank voor strafzaken.
- 3 Geen voor vergelijkend DNA-onderzoek geschikt resultaat.

Ad 1 Het verkregen DNA-profiel van het spoor matcht met het DNA-profiel van een ander spoor of persoon in de zaak of in de Nederlandse DNA-databank voor strafzaken. Matchen twee DNA-profielen dan vermeldt de gerechtelijke deskundige altijd hoe zeldzaam het DNA-profiel van het sporenmateriaal is.

Ad 2 Het verkregen DNA-profiel van het spoor matcht niet met een DNA-profiel van een ander spoor of persoon in de zaak of in de Nederlandse DNA-databank voor strafzaken. Dit betekent dat het spoor niet afkomstig is van de donor(en) van de andere sporen of van personen in de zaak of in de Nederlandse DNA-databank voor strafzaken. DNA-profielen van sporen die niet matchen met die van een persoon in de zaak neemt het NFI, conform de DNA-wet, op in de Nederlandse DNA-databank voor strafzaken.

Ad 3 Als er geen DNA-profiel is verkregen of het resultaat is ongeschikt voor vergelijkend DNA-onderzoek, dan is geen vergelijkend DNA-onderzoek mogelijk. Dit doet zich voor als er geen celmateriaal aanwezig is, of te weinig celmateriaal van voldoende kwaliteit. Ook als er wel voldoende DNA in het spoor aanwezig is, kan het voorkomen dat het resultaat van het DNA-onderzoek ongeschikt is voor vergelijkend DNA-onderzoek. Dit komt met enige regelmaat voor bij het onderzoek van mengsporen (zie Brontekst 7, *Interpretatie van DNA-bewijs II; onvolledige DNA-profielen en DNA-mengprofielen*).

Matchende DNA-profielen

Twee DNA-profielen zijn aan elkaar gelijk als voor elk onderzocht locus in het ene DNA-profiel de vastgestelde DNA-kenmerken gelijk zijn aan die van het desbetreffende locus in het andere DNA-profiel. Men spreekt dan van 'matchende' DNA-profielen.

Ook kunnen DNA-profielen met elkaar matchen zonder dat ze exact aan elkaar gelijk zijn. Dit betreft met name de volgende situaties (nader uiteengezet in Brontekst 7):

- als de DNA-kenmerken van een *onvolledig enkelvoudig DNA-profiel* (van een spoor) gelijk zijn aan de desbetreffende DNA-kenmerken van een *volledig enkelvoudig DNA-profiel* (van een referentiemonster van een persoon).
- als de DNA-kenmerken van een *enkelvoudig DNA-profiel* (van bijvoorbeeld een referentiemonster van een persoon) ook voorkomen in ('passen in') een *DNA-mengprofiel*.

De bewijswaarde van een match van twee DNA-profielen hangt af van hoe zeldzaam de waargenomen overeenkomst is. Bij twee exact gelijke enkelvoudige volledige DNA-profielen heeft deze match een zeer hoge bewijswaarde. Als het een match betreft met een onvolledig DNA-profiel of een DNA-mengprofiel is de bewijswaarde gerelateerd aan het aantal en de zeldzaamheid van de matchende DNA-kenmerken (zie Brontekst 7).

Berekende frequentie

Om de bewijswaarde van matchende DNA-profielen te bepalen is het van belang te weten hoe groot de kans is dat de DNA-profielen bij toeval matchen. Naarmate de kans hierop kleiner is, is het feit dat de DNA-profielen matchen opmerkelijker.

Hoe groot is nu de kans dat twee DNA-profielen, bijvoorbeeld dat van een spoor en dat van een verdachte, bij toeval gelijk zijn? Het antwoord op deze vraag wordt uitgedrukt in de 'berekende frequentie' van voorkomen van het DNA-profiel (van het sporenmateriaal) in de populatie. De frequentie van een DNA-profiel wordt berekend met speciaal daarvoor ontwikkelde, gevalideerde rekenprogramma's. De berekening is gebaseerd op populatiegenetische gegevens uit het referentiedatabestand van het NFI (zie pagina 8, 'Referentiedatabestand'). Met dit referentiedatabestand is voor elk DNA-kenmerk dat bij het DNA-onderzoek kan voorkomen berekend hoe vaak het in de populatie voorkomt. Dit is de verwachte frequentie van het DNA-kenmerk in de populatie. Het rekenprogramma berekent de frequentie van een DNA-profiel door de frequenties van de vastgestelde DNA-kenmerken met elkaar te vermenigvuldigen.

De berekening van de frequentie van het DNA-profiel omvat bovendien twee statistische correcties (zie 'Statistische onderbouwing' op pagina's 7 en 8). De eerste voor het feit dat binnen een populatie verschillende subgroepen voorkomen. De tweede voor het feit dat de gehanteerde frequenties van de afzonderlijke DNA-kenmerken zijn geschat op basis van een referentiedatabestand met een beperkt aantal personen.

De rekenmethode en de gehanteerde statistische correcties voor het berekenen van de frequentie van een DNA-profiel worden internationaal gebruikt.

De berekende frequentie is niet van toepassing op aan elkaar verwante personen. Voor bloedverwanten van de met het spoor matchende verdachte geldt dat de kans groter is dat hun DNA-profiel matcht met dat van het spoor, dan het DNA-profiel van een niet-verwante persoon.

Kleiner dan één op één miljard

Het NFI gebruikt voor het vaststellen van het DNA-profiel het zogenoemde SGM-Plus DNA-analysesysteem (SGM staat voor Second Generation Multiplex). Het SGM-Plus DNA-analysesysteem is een gevalideerd, internationaal gebruikt DNA-analysesysteem voor het vaststellen van DNA-profielen. Het is speciaal ontwikkeld voor forensisch DNA-onderzoek. Het SGM-Plus DNA-analysesysteem bepaalt van tien verschillende hypervariabele loci de DNA-kenmerken. Bovendien stelt het SGM-Plus DNA-analysesysteem het geslacht vast. Bij volledige DNA-profielen van tien loci vermenigvuldigt men de frequenties van tien paren DNA-kenmerken met elkaar (zie onderstaand 'Rekenvoorbeeld'). Dit resulteert in een frequentie die altijd kleiner is dan één op één miljard. Met andere woorden, de kans dat een willekeurig gekozen persoon dit DNA-profiel heeft is altijd kleiner dan één op één miljard. De combinatie van de DNA-kenmerken in elk volledig SGM-Plus DNA-profiel is dus in alle gevallen extreem zeldzaam.

Rekenvoorbeeld

De berekende frequentie van de DNA-kenmerkencombinatie per locus

Stel, een bepaald locus heeft de DNA-kenmerken 4 en 6. DNA-kenmerk 4 komt in de relevante populatie voor met een frequentie van 10% (0,1 ofwel 1 op 10). DNA-kenmerk 6 komt voor met een frequentie van 15% (0,15 ofwel 1 op 6,7). De berekening van de frequentie van voorkomen van deze combinatie van DNA-kenmerken van dat locus, is als volgt:

$$10\% (0,1) \times 15\% (0,15) = 1,5\% (0,015).$$

Omdat er voor deze DNA-kenmerkencombinatie twee mogelijkheden zijn: 4/6 en 6/4 (kenmerk 4 is van de vader geërfd en kenmerk 6 van de moeder, of andersom: kenmerk 4 is van de moeder geërfd en kenmerk 6 van de vader) is een vermenigvuldiging met een factor twee nodig.

De berekende frequentie van de DNA-kenmerkencombinatie van dat locus is dan $2 \times 1,5\% = 3\%$ (of: $2 \times 0,015 = 0,03$).

De berekende frequentie van het DNA-profiel

Het rekenprogramma berekent de frequentie van het DNA-profiel door de frequenties van de DNA-kenmerkencombinaties die van de afzonderlijke loci zijn bepaald met elkaar te vermenigvuldigen. Voor een volledig DNA-profiel, dat bestaat uit tien loci, zou dat er als volgt uit kunnen zien (de frequentie van de DNA-kenmerkencombinatie van het eerste locus (3% of 0,03) is hierboven berekend):

frequentie van de DNA-kenmerkencombinatie van locus 1 x

frequentie van de DNA-kenmerkencombinatie van locus 2 xx

frequentie van de DNA-kenmerkencombinatie van locus 10:

$$3\% \times 10\% \times 20\% \times 10\% \times 10\% \times 25\% \times 5\% \times 30\% \times 10\% \times 7\%, \text{ ofwel:}$$

$$0,03 \times 0,1 \times 0,2 \times 0,1 \times 0,1 \times 0,25 \times 0,05 \times 0,3 \times 0,1 \times 0,07 = 0,000000001575.$$

De frequentie van het DNA-profiel is 0,000000001575 (dus 0,0000001575%); dit is ongeveer één op 6,3 miljard (berekening: 1 gedeeld door 0,000000001575).

Bij het berekenen van de frequentie voert het rekenprogramma de voorgeschreven twee statistische correcties uit (zie 'Statistische onderbouwing' op pagina 7 en 8).

Op basis van al de mogelijke DNA-kenmerken die op de tien verschillende loci kunnen voorkomen, is het aantal theoretisch mogelijke combinaties vele malen groter dan de wereldbevolking. Hieruit volgt dat lang niet elk theoretisch mogelijk DNA-profiel van tien loci (en dus twintig DNA-kenmerken) in de wereldbevolking voorkomt.

Indien noodzakelijk, bijvoorbeeld bij het DNA-identificatieonderzoek van onbekende overleden personen, kunnen naast de tien hypervariabele gebieden (loci) die met het SGM-Plus DNA-analysesysteem worden bepaald ook de DNA-kenmerken van andere loci worden vastgesteld. Het NFI maakt hiervoor gebruik van verschillende gevalideerde DNA-analysesystemen.

Statistische onderbouwing

Bij de berekening van de frequentie van het DNA-profiel vermenigvuldigt men de frequenties van de DNA-kenmerken met elkaar (door toepassing van de zogenoemde 'productregel'). Om de frequenties van de DNA-kenmerken van de verschillende loci met elkaar te kunnen vermenigvuldigen gaat men uit van twee aannamen:

- 1 De DNA-kenmerken van *één locus* zijn onafhankelijk van elkaar (het zogenoemde 'Hardy-Weinberg evenwicht').
- 2 De DNA-kenmerken van de *verschillende loci* zijn onafhankelijk van elkaar (het zogenoemde 'linkage evenwicht').

Deze aannamen houden in dat de aanwezigheid van het ene DNA-kenmerk geen invloed heeft op de kans op de aanwezigheid van het andere DNA-kenmerk, zowel voor hetzelfde locus, als voor alle andere loci. Uit de populatiegenetica is bekend dat deze beide aannamen niet voor honderd procent opgaan. Voor het berekenen van de frequentie van het DNA-profiel (volgens de 'productregel') is daarom nog wel een kleine rekenkundige correctie nodig: de correctie voor subgroepen. Daarnaast is er nog een kleine rekenkundige correctie nodig omdat de gehanteerde frequenties waarmee de DNA-kenmerken in de populatie voorkomen, zijn vastgesteld op basis van een referentiedatabestand: de correctie voor de grootte van het referentiedatabestand. De twee rekenkundige correcties worden hieronder nader toegelicht.

Correctie voor subgroepen

In een populatie komen in de regel verschillende subgroepen voor. Een populatie is immers, uitzonderingen daargelaten, een dynamische gemeenschap die onderhevig is aan interne en externe invloeden. Immigratie en emigratie hebben invloed op de samenstelling van de populatie. Hierbij speelt de mate waarin verschillende subgroepen met elkaar integreren een belangrijke rol. Bovendien kunnen binnen een populatie in zekere mate traditioneel-geïsoleerde subgroepen voorkomen, zoals kan gelden voor eilandbewoners of bepaalde godsdienstige groeperingen. Zo is ook de blanke Nederlandse populatie opgebouwd uit verschillende subgroepen die meer of minder met elkaar integreren en van elkaar verschillen. De mate waarin de DNA-kenmerken in de verschillende subgroepen voorkomen, kan daardoor iets verschillen. Bij de berekening van de frequenties van DNA-profielen houdt men rekening met deze populatiesubstructuren en past men hiervoor een statistische correctie toe. De door het NFI gebruikte populatiegenetische correctie is de zogenoemde theta-correctie, ook wel Fst-correctie genoemd.

Correctie voor grootte referentiedatabestand

Een andere statistische correctie heeft betrekking op de frequenties van de DNA-kenmerken. Omdat de frequenties waarmee de afzonderlijke DNA-kenmerken voorkomen in de populatie zijn geschat op basis van een referentiedatabestand met een beperkt aantal personen (zie hieronder) vindt er hiervoor een correctie plaats (de zogenoemde 'size bias correctie').

Referentiedatabestand

De frequenties waarmee de verschillende DNA-kenmerken van een locus voorkomen in de populatie zijn door het NFI berekend met gegevens uit een referentiedatabestand met DNA-profielen van 231 blanke Nederlanders. Een veelgestelde vraag is of een referentiedatabestand van 'slechts' 231 personen voldoende basis biedt voor berekende frequenties (van volledige DNA-profielen) van kleiner dan één op één miljard. Rekenmodellen laten zien dat voor het schatten van de frequenties van de verschillende DNA-kenmerken die per locus kunnen voorkomen een referentiedatabestand van 231 personen groot genoeg is. Frequenties van DNA-kenmerken liggen over het algemeen tussen 0,2% en 30%. Deze orde van grootte kan men betrouwbaar schatten met referentiedatabestanden van minstens 200 personen. De schattingen zullen niet veel anders zijn wanneer het referentiedatabestand uit meer personen zou bestaan. Het rekenprogramma corrigeert voor deze eventueel aanwezige kleine afwijkingen.

In het referentiedatabestand van het Forensisch Laboratorium voor DNA Onderzoek (FLDO) van de Universiteit Leiden blijken de frequenties van DNA-kenmerken ruwweg gelijk aan die van het NFI. Dit referentiedatabestand kan daarom worden gezien als een vorm van validatie van de frequenties in het referentiedatabestand van het NFI.

Ook internationaal zijn de verschillen over het algemeen klein. Zo komt bijvoorbeeld DNA-kenmerk 17 op locus D3S1358 met een frequentie van 0,193 (19,3%) voor in het NFI-referentiedatabestand. Dit DNA-kenmerk heeft in het Engelse referentiedatabestand (dat een onderverdeling in subpopulaties kent) een frequentie van 0,195 (19,5%) voor blanke mensen, 0,232 (23,2%) voor negroïde mensen en 0,220 (22%) voor Aziatische mensen.

Representatief

Is het referentiedatabestand dat het NFI gebruikt ook representatief voor personen van niet-West-Europese afkomst, bijvoorbeeld een Turk of een Marokkaan? Uit vergelijkingen met andere referentiedatabestanden in binnen- en buitenland blijkt dat de geschatte frequenties van de verschillende DNA-kenmerken van de loci iets kunnen verschillen per referentiedatabestand (zie hierboven). Echter, de consequenties hiervan zijn verwaarloosbaar klein voor volledige DNA-profielen.

Dit komt doordat men bij het berekenen van de frequentie van het DNA-profiel van bijvoorbeeld tien loci (met elk twee -verschillende of gelijke- DNA-kenmerken) de frequenties van twintig DNA-kenmerken met elkaar vermenigvuldigt. Vermenigvuldiging van al deze frequenties resulteert in een frequentie van het volledige DNA-profiel die altijd kleiner is dan één op één miljard, ongeacht het gebruikte referentiedatabestand. Hierdoor hebben eventuele verschillen in de frequenties van DNA-kenmerken tussen verschillende populaties een verwaarloosbare invloed op de uiteindelijke uitkomst. Met andere woorden, de berekende frequentie van elk volledig DNA-profiel van tien loci is, ongeacht de populatie, altijd kleiner dan één op één miljard.

Alleen als het een onvolledig DNA-profiel betreft dat uit DNA-kenmerken van slechts enkele loci bestaat kan de berekende frequentie wezenlijk verschillen tussen verschillende populaties.

In die gevallen kan men besluiten gebruik te maken van een referentiedatabestand van de desbetreffende populatie. Zo heeft het NFI bijvoorbeeld ook een referentiedatabestand van de Antilliaanse populatie.

Dat volledige DNA-profielen extreem lage frequenties hebben, blijkt ook uit ervaringen in landen met zeer grote DNA-databanken met DNA-profielen van sporen en personen. Dat geldt bijvoorbeeld voor de Engelse DNA-databank met vier miljoen DNA-profielen. Hier is bij het vergelijken van DNA-profielen het slechts een enkele keer voorgekomen dat bij toeval de DNA-profielen van twee verschillende personen aan elkaar gelijk waren. Het betrof in deze gevallen in het verleden gegenereerde DNA-profielen, die nog bestonden uit DNA-kenmerken van slechts zes loci, of DNA-profielen die afkomstig waren van twee broers. In die gevallen waarbij gelijke DNA-profielen op basis van zes loci werden waargenomen gaf een later uitgevoerd DNA-onderzoek op tien loci wel een verschillend DNA-profiel.

Aandachtspunten berekende frequentie

De berekende frequentie is de maat voor de zeldzaamheid van een vastgesteld DNA-profiel (van een biologisch spoor) in de populatie. Het is feitelijk de kans dat een willekeurig gekozen persoon hetzelfde DNA-profiel heeft als dat van het spoor. Het NFI berekent deze frequentie van voorkomen met populatiegenetische gegevens uit een referentiedatabestand. De berekening van de frequentie omvat bovendien een statistische correctie voor subgroepen en voor de grootte van het referentiedatabestand.

Andere populaties

Volledige SGM-Plus DNA-profielen van tien loci hebben een berekende frequentie die altijd kleiner is dan één op één miljard. Ook wanneer men deze berekening zou baseren op populatiegenetische gegevens uit een referentiedatabestand van een willekeurig andere populatie (bijvoorbeeld Antilliaanse of Engelse), resulteert dit in een berekende frequentie die altijd kleiner is dan één op één miljard. Onvolledige DNA-profielen hebben een hogere berekende frequentie (en dus een kleinere bewijswaarde) dan volledige DNA-profielen. Als een onvolledig DNA-profiel een relatief hoge berekende frequentie heeft, dan kan die frequentie nog aanzienlijk verschillen als men de berekening baseert op een ander referentiedatabestand.

Bloedverwanten

De berekende frequentie is niet van toepassing op aan elkaar verwante personen. Voor bloedverwanten van een met het spoor matchende persoon geldt dat de kans groter is dat hun DNA-profiel matcht met dat van het spoor dan het DNA-profiel van een niet-verwante persoon. Zeker bij onvolledige DNA-profielen (met een hogere frequentie en een lagere bewijswaarde) moet men hiermee rekening houden.

Een in Nederland gevonden spoor

Zoals hierboven beargumenteerd, is het voor volledige DNA-profielen niet nodig frequenties te berekenen op basis van referentiedatabestanden van specifieke populaties. Belangrijk is te realiseren dat men moet uitgaan van het biologische spoor. Dit is een in Nederland gevonden spoor van een onbekende celdonor, mogelijk de dader van een misdrijf. Omdat dit spoor in Nederland is gevonden (en hoogstwaarschijnlijk daar is ontstaan) wordt hiervoor het referentiedatabestand van het NFI gebruikt als basis voor de berekening van de frequentie van het DNA-profiel. Alleen indien men vrij zeker weet (bijvoorbeeld op basis van tactische informatie) dat de celdonor van het spoor uit een bepaalde bevolking(sgroep) komt, kan men overwegen de statistische berekening te baseren op een referentiedatabestand van de desbetreffende populatie. Dit heeft overigens alleen zin indien van het spoor slechts een onvolledig DNA-profiel beschikbaar is. Bij volledige DNA-profielen zal de frequentie ook in deze populatie kleiner zijn dan één op één miljard.

'Match' en 'berekende frequentie'

Het NFI rapporteert de conclusie van een vergelijkend autosomaal DNA-onderzoek dat is gericht op de vraag of een biologisch spoor afkomstig is van een verdachte (of slachtoffer of andere betrokkene) in de vorm van een tweeledige uitspraak: *een uitspraak of de vergeleken DNA-profielen matchen en een uitspraak over de zeldzaamheid, de zeldzaamheidswaarde, van het vastgestelde DNA-profiel van het spoor, uitgedrukt in de berekende frequentie van voorkomen van het DNA-profiel in de populatie.*

Hoe lager de berekende frequentie, hoe sterker het DNA-bewijs. Bij een volledig SGM-Plus DNA-profiel, dat bestaat uit de DNA-kenmerken van tien loci, is de berekende frequentie altijd kleiner dan één op één miljard. Betreft het een DNA-profiel dat uit minder dan tien loci bestaat, een onvolledig DNA-profiel, dan is de berekende frequentie hoger en de bewijswaarde dus kleiner.

Conclusie bij gelijke volledige DNA-profielen

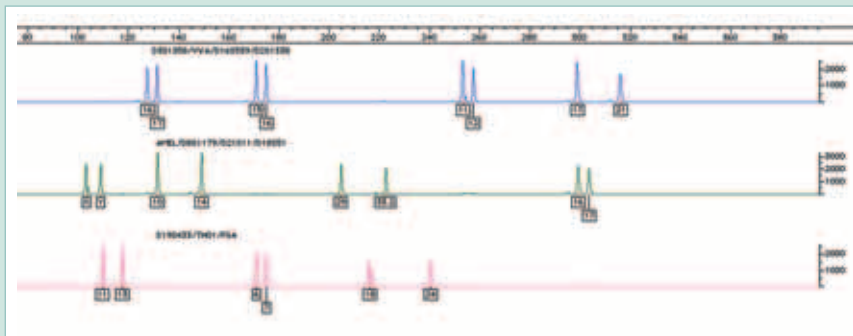
Als volledige DNA-profielen in een bepaald onderzoek aan elkaar gelijk zijn, geeft de deskundige dit resultaat samen met de berekende frequentie van het DNA-profiel weer in het deskundigenrapport. Bijvoorbeeld in het geval dat van het onderzochte bloedspoor op de broek van verdachte V een volledig DNA-profiel is verkregen en dit DNA-profiel is gelijk aan het DNA-profiel van het slachtoffer S (en niet gelijk aan dat van verdachte V) (zie illustratie 1): *'Van het DNA in de bemonstering [ABC001]#1 van het bloedspoor, aangetroffen op de broek van de verdachte V, is een DNA-profiel verkregen van een man. Dit DNA-profiel matcht met het DNA-profiel van het slachtoffer S [RRR001]. Dit betekent dat het bloed in de bemonstering [ABC001]#1 van het bloedspoor, aangetroffen op de broek van de verdachte V, afkomstig kan zijn van het slachtoffer S. De berekende frequentie van het DNA-profiel van het bloed in de bemonstering [ABC001]#1 van het bloedspoor is kleiner dan één op één miljard. Ofwel, de kans dat het DNA-profiel van een willekeurig gekozen man matcht met dit DNA-profiel is kleiner dan één op één miljard.'*

[ABC001]#1 en [RRR001] zijn de DNA-identiteitszegels die aan respectievelijk de desbetreffende bemonstering en het desbetreffende referentiemonster zijn toegekend.

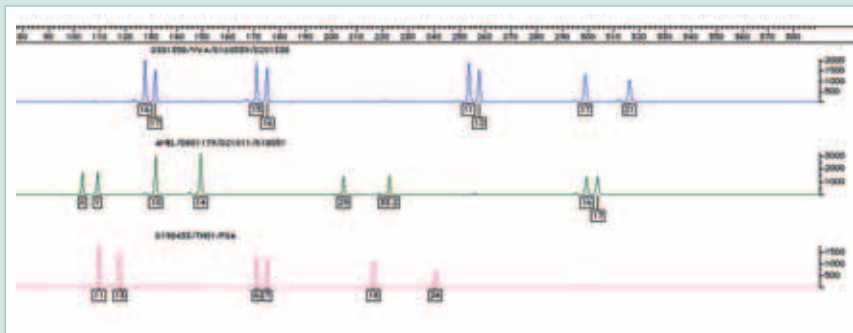
N B De berekende frequentie is niet van toepassing op aan elkaar verwante personen. Voor bloedverwanten van de met het spoor matchende persoon geldt dat de kans groter is dat hun DNA-profiel matcht met dat van het spoor dan het DNA-profiel van een niet-verwante persoon.

De frequenties van volledige DNA-profielen (van tien loci) zijn altijd kleiner dan één op één miljard. Het NFI rapporteert voor de frequentie van volledige DNA-profielen de bovengrens van één op één miljard. De reden hiervoor is dat het berekenen van frequenties van volledige DNA-profielen over het algemeen leidt tot extreem lage waarden, zoals één op 34.000 miljard. Dit is niet meer te bevatten. Bovendien zijn dergelijke extreem lage frequenties empirisch niet meer te valideren en daardoor feitelijk minder betrouwbaar dan de vermelding van de bovengrens van één op één miljard.

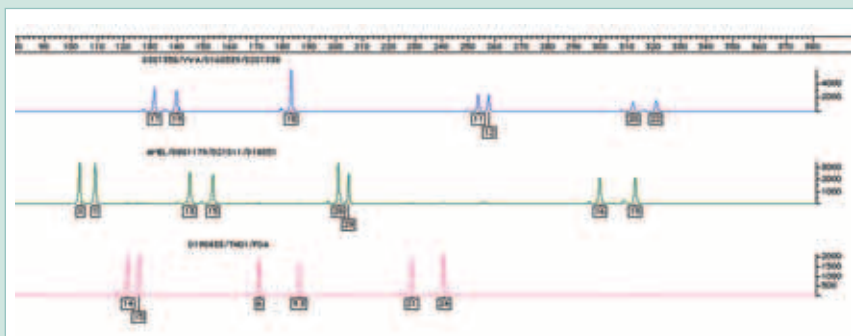
DNA-profiel van het bloedspoor [ABC001]#1



DNA-profiel van het slachtoffer S [RRR001]



DNA-profiel van de verdachte V [RVE666]



Illustratie 1: Van het DNA in de bemonstering [ABC001]#1 van het bloedspoor, aangetroffen op de broek van de verdachte V, is een volledig DNA-profiel verkregen, zie het bovenste DNA-profiel. Dit DNA-profiel matcht met het DNA-profiel van het slachtoffer S [RRR001], zie middelste DNA-profiel, en matcht niet met het DNA-profiel van de verdachte V [RVE666], zie onderste DNA-profiel.

Geen absolute herkomstuitspraak

Als de frequentie van een volledig DNA-profiel kleiner is dan één op één miljard, waarom kan men dan op basis hiervan geen absolute uitspraak doen over de herkomst van het biologische spoor? Per slot van rekening is dat de vraag die de rechter beantwoord wil zien. Hiervoor zijn drie belangrijke redenen:

1 'extreem zeldzaam' betekent niet 'uniek'

Ook al is een DNA-profiel nog zo extreem zeldzaam, het is nooit te achterhalen of dit DNA-profiel echt uniek is. Het is immers onmogelijk na te gaan of er in de gehele wereldbevolking daadwerkelijk niemand is die ditzelfde DNA-profiel heeft. Met andere woorden: ook al is het DNA-profiel statistisch gezien extreem zeldzaam, men kan nooit helemaal uitsluiten dat de DNA-profielen van twee niet verwante personen bij toeval gelijk zijn. Mensen winnen soms de hoofdprijs in de Lotto, ook al is die kans praktisch gesproken voor iedere persoon verwaarloosbaar klein. Logisch gezien kan ook een gerechtelijke deskundige dus niet de stap nemen naar de stelling dat het DNA-profiel uniek is en er maar één persoon is die dit specifieke DNA-profiel heeft.

2 bloedverwanten; eeneiige twee- of meerlingen

Personen die aan elkaar zijn verwant, zoals familieleden, hebben onderling meer overeenkomsten in hun DNA dan niet-bloedverwante personen. Daarom moet men zich bij het interpreteren van het DNA-bewijs realiseren dat de kans groter is dat het DNA-profiel van een bloedverwant (van de met het spoor matchende persoon) matcht met het DNA-profiel van het spoor dan het DNA-profiel van een niet-verwante persoon. Bovendien is er een groep bloedverwanten met hetzelfde DNA: de eeneiige twee- of meerlingen. Het DNA van deze personen is, in alle chromosomen, identiek. Dit geldt dus ook voor hun DNA-profielen.

3 kans op fouten

De match kan berusten op een fout in het laboratorium of elders in de keten, van plaats delict tot deskundigenrapport. Geaccrediteerde forensische laboratoria, zoals in Nederland het NFI en het FLDO, werken volgens strikte protocollen en kwaliteitsnormen (conform de Europese norm NEN-EN ISO/IEC 17025). Ook elders in de keten zijn stringente maatregelen genomen om fouten te vermijden. Daardoor zal de kans op een fout in de onderzoeksketen klein zijn (zie ook 'Omgaan met de kans op fouten' op pagina 14-18).

Zoals hierboven blijkt, kan men in een specifieke strafzaak op basis van met elkaar matchende DNA-profielen geen absolute uitspraak doen over de herkomst van het celmateriaal (spoor). Wel geldt in het algemeen (onder voorbehoud van bovenstaande punten 2 en 3) dat in de praktijk bij gelijke volledige DNA-profielen van een spoor en een persoon het zeer waarschijnlijk is dat het celmateriaal van het spoor afkomstig is van deze persoon.

Wel uitsluiten

Hoe laag de berekende frequentie van een bepaald DNA-profiel ook is, een absolute uitspraak over de herkomst van het spoor, alleen gebaseerd op het DNA-profiel, is nooit mogelijk. Daar tegenover staat dat men wél een absolute uitspraak kan doen wanneer het DNA-profiel van de verdachte en dat van het spoor niet met elkaar matchen. Dan kan men de verdachte uitsluiten als donor van het celmateriaal van het biologisch spoor, zij het dat ook hier het voorbehoud geldt van een mogelijke fout in de onderzoeksketen.

Vuistregel 1

- **Een deskundige kan zelfs op grond van gelijke volledige DNA-profielen *geen absolute uitspraak doen over de herkomst van een biologisch spoor*; dus niet met absolute zekerheid concluderen dat de desbetreffende persoon de celdonor is van het spoor.**

Als de volledige DNA-profielen, bestaande uit tien (of meer) loci, van een spoor en een persoon aan elkaar gelijk zijn, dan is dit zeer sterk bewijs voor de stelling dat deze persoon de donor is van het celmateriaal van het spoor. Andere verklaringen (een ander persoon met dit DNA-profiel, bijvoorbeeld een eeneiige tweelingbroer of iemand die toevallig dit DNA-profiel heeft, of een fout in de onderzoeksketen) zijn echter nooit met zekerheid uit te sluiten. Want ook al is een DNA-profiel extreem zeldzaam, dan betekent dit nog niet dat het DNA-profiel ook daadwerkelijk uniek is.

- **Als het DNA-profiel van een persoon niet matcht met het DNA-profiel van het celmateriaal van het biologische spoor, dan kan men deze persoon uitsluiten als donor van het celmateriaal van het spoor** (onder voorbehoud van fouten in de onderzoeksketen).

Omgaan met de kans op fouten

Forensisch onderzoek in strafzaken is mensenwerk. Ondanks dat dit onderzoek met de grootst mogelijke zorgvuldigheid gebeurt, is niet uit te sluiten dat er in de onderzoeksketen (van het sporenonderzoek tot en met de uitspraak van de rechter) fouten worden gemaakt. Zo kunnen op de plaats delict sporen worden gemist of sporen worden veiliggesteld die niets met het delict te maken hebben. Ook kan tijdens het onderzoek op de plaats delict verwisseling of contaminatie van veiliggestelde sporen plaatsvinden of kan daarna, bij het onderzoek aan de sporen op het laboratorium, iets fout gaan. Tot slot kunnen er aan het eind van de keten misvattingen ontstaan over de interpretatie van de onderzoeksresultaten. Het omgaan met de kans op fouten is daarom voor de gehele keten van belang. Het NFI onderkende dat bij de introductie van het DNA-onderzoek in 1994 en heeft de wetgever geadviseerd dat er waarborgen moesten komen voor het geval er fouten worden gemaakt bij het DNA-onderzoek. De wetgever heeft daarop besloten dat een laboratorium dat in opdracht van justitie DNA-onderzoek doet geaccrediteerd moet zijn. Bovendien is geregeld dat een verdachte recht heeft op een contra-expertise door een tweede geaccrediteerd, onafhankelijk laboratorium.

Geaccrediteerd laboratorium

Bij de wettelijke regeling van het DNA-onderzoek in Nederland is de voorwaarde neergelegd dat het forensisch DNA-onderzoek dient te worden uitgevoerd in een geaccrediteerd onderzoekslaboratorium. In een geaccrediteerde omgeving wordt volgens strikte protocollen en kwaliteitsnormen gewerkt om de kans op fouten tijdens de analyse zo klein mogelijk te maken. De controlemechanismen in een geaccrediteerde omgeving zijn erop gericht onregelmatigheden in een zo vroeg mogelijk stadium op te kunnen sporen. Het NFI heeft zijn kwaliteitsmanagementsysteem ingericht conform de Europese norm NEN-EN ISO/IEC 17025. Deze internationale norm (standaard) beschrijft alle voorwaarden waaraan een test- en kalibratielaboratorium moet voldoen om aan te tonen dat het volgens dit kwaliteitssysteem werkt, competent is en technisch valide gegevens en resultaten levert. Naast de Nederlandse forensische laboratoria voldoen ook diverse buitenlandse forensische laboratoria aan deze standaard. De Raad voor Accreditatie is in Nederland de instantie die een 'accreditatie' (kwaliteitskeurmerk) in de publieke sector verleent. Het NFI is sinds 1994 geaccrediteerd. Deze accreditatie is een formele erkenning voor het kwaliteitssysteem van het NFI en wordt internationaal erkend. Elk jaar toetst de Raad voor Accreditatie het NFI. De toetsing gebeurt mede aan de hand van audits door onafhankelijke deskundigen. Tijdens deze audits bekijken deze deskundigen ook hoe het laboratorium heeft gepresteerd bij onderzoeken van 'blind cases' en in 'ringonderzoeken'. Bij een blind case wordt een door derden samengestelde casus aan het onderzoekslaboratorium aangeboden, zodanig dat het de laboratoriummedewerkers onbekend is dat het een niet werkelijk gebeurd misdrijf betreft. Pas nadat het onderzoek is afgerond en gerapporteerd wordt dit bekend gemaakt, tezamen met de waardering voor het uitgevoerde onderzoek. Andere belangrijke vakbekwaamheidstesten zijn de internationale ringonderzoeken. Hierbij ontvangen de deelnemende laboratoria allen, gelijktijdig hetzelfde uitgangsmateriaal van een door derden opgezette casus. De onderzoeksresultaten die de verschillende laboratoria hebben verkregen worden naderhand met elkaar vergeleken en in speciale bijeenkomsten besproken. In 2006 heeft het NFI nog een maatregel genomen om de kwaliteit van het onderzoek en de rapportage verder te waarborgen. Dit betreft de zogenoemde twijfel of tegenspraakprocedure en is een regeling voor situaties waarin verschillende deskundigen het met elkaar oneens zijn over de conclusie van een onderzoek.

Contra-expertise

In de Nederlandse DNA-wet is in 1994 een belangrijke maatregel opgenomen om fouten op te sporen die een verdachte ten onrechte kunnen belasten: het recht van de verdachte om een contra-expertise te laten uitvoeren door een tweede onafhankelijk laboratorium. Artikel 7 lid 1 van het DNA-besluit noemt hiervoor het Forensisch Laboratorium voor DNA-Onderzoek (FLDO) van de Universiteit Leiden, maar biedt in principe ook de mogelijkheid om dit elders te laten doen, bijvoorbeeld in het buitenland. Wel dient het laboratorium dat de contra-expertise uitvoert geaccrediteerd te zijn en deskundig te zijn op het gebied van forensisch DNA-onderzoek. Voor een laboratorium dat gevestigd is in het buitenland geldt dat dit door een met de Raad voor Accreditatie vergelijkbare instantie geaccrediteerd moet zijn. De resultaten van de contra-expertises door het FLDO geven inzicht in de meest ernstige fout die in de keten kan optreden: het DNA-spoor wordt ten onrechte met de verdachte geassocieerd, zonder dat de fout door het NFI wordt ontdekt. Uiteraard werkt dit vangnet alleen wanneer verdachten die door een DNA-onderzoek op het NFI ten onrechte als donor van sporen materiaal zijn aangemerkt daadwerkelijk een contra-expertise aanvragen.

Kans op een fout in het onderzoekslaboratorium

Hoewel de handelingen tijdens het forensisch DNA-onderzoek steeds verder worden geautomatiseerd en gerobotiseerd, blijven (manuele) interventies door de onderzoeksmedewerkers noodzakelijk. Het NFI voert alle handelingen bij het DNA-onderzoek met de grootst mogelijke zorgvuldigheid uit. Dat neemt niet weg dat, ondanks het kwaliteitssysteem, fouten in het forensisch DNA-onderzoek nooit volledig zijn uit te sluiten. Een relevante vraag is dan hoe vaak er fouten worden gemaakt. Een geaccrediteerd onderzoekslaboratorium zoals het NFI registreert en analyseert alle waargenomen fouten en alle andere onregelmatigheden die tot een fout zouden kunnen leiden: zowel de fouten die in de loop van het proces zelf zijn waargenomen (en meestal kunnen worden hersteld), als fouten die door derden zijn geconstateerd. Fouten met betrekking tot het forensisch DNA-onderzoek publiceert het NFI in het jaarverslag. Bij elke fout wordt bekeken wat de consequenties zijn en of er structurele maatregelen voor verbetering genomen dienen te worden. Maar hoe goed het kwaliteitssysteem ook is, er zullen altijd fouten zijn die niet worden ontdekt. Dit aantal is niet zonder meer te relateren aan het aantal fouten dat wel wordt ontdekt. Dat maakt het vaststellen van de kans op een fout erg moeilijk. Daarbij komt dat men leert van gemaakte fouten. Maatregelen moeten er immers voor zorgen dat geconstateerde fouten in de toekomst niet meer voorkomen, waardoor de kans op die fouten kleiner wordt. Daarnaast kunnen nieuw geïntroduceerde onderzoekstechnieken of processen tot minder of mogelijk tot nieuwe fouten leiden. Het kan daarom een vertekend beeld geven om op basis van waargenomen fouten een getalsmatige uitspraak te doen over de kans dat een fout optreedt. De kwaliteit van de resultaten van het laboratoriumonderzoek van stukken van overtuiging en sporen is mede afhankelijk van hoe deze zijn gevonden, veiliggesteld, opgeslagen en vervoerd. Daarom is minstens zo belangrijk te weten hoe groot de kans is op een fout in de rest van de keten. Onderzoek hiernaar is heel complex, maar nadere informatie hierover zal voor alle betrokkenen bijdragen aan een beter inzicht. Met deze kennis kunnen potentiële foutenbronnen in de keten worden opgespoord en kunnen maatregelen worden genomen.

Vakbekwaamheidstesten

Vakbekwaamheidstesten zoals ringonderzoeken, waarbij een groot aantal laboratoria aan hetzelfde uitgangsmateriaal van een vooraf opgezette casus onderzoek verrichten, zijn een goede en objectieve maatstaf voor de kwaliteit en performance van een laboratorium. Om op basis van deze momentopnamen betrouwbaar vast te kunnen stellen wat de kans op een fout onderzoeksresultaat voor een bepaald laboratorium is, zal dat laboratorium een zeer groot aantal ringonderzoeken moeten uitvoeren. Pas op basis van vele duizenden ringonderzoeken zou een verantwoorde statistische uitspraak kunnen worden gedaan over het aantal fouten dat een laboratorium per duizend onderzoeken rapporteert. Het spreekt voor zich dat dit in de praktijk geen reële optie is.

Om dit probleem enigszins te omzeilen door de resultaten van de vakbekwaamheidstesten van verschillende laboratoria bij elkaar te voegen ('industry-wide error rate') heeft een ander nadeel. Goed presterende laboratoria zijn dan de dupe van hun slecht presterende collega's en de slechte resultaten van de onvoldoende presterende laboratoria worden verbloemd.

Kans op een fout in een specifieke zaak

Waar het in de praktijk om gaat is of in een specifieke zaak de kans bestaat dat er in *die* zaak een fout is opgetreden, en wat het gevolg daarvan kan zijn. Een indicatie van de orde grootte van de kans op een fout bij het laboratoriumonderzoek kan verkregen worden uit de statistieken van het kwaliteitssysteem van het onderzoekslaboratorium. Daarnaast geldt uiteraard dat ook fouten elders in de onderzoeksketen in beschouwing moeten worden genomen. Het blijkt dat de kans op een fout sterk verschilt tussen verschillende typen zaken. Zo zal de kans op een match die het gevolg is van een fout groter zijn in een situatie waarbij slechts één minuscuul biologisch spoor beschikbaar is dat een onvolledig DNA-profiel oplevert, dan in de situatie waarin er een groot aantal sporen beschikbaar zijn waarvan de DNA-profielen allemaal aan elkaar gelijk zijn.

Ook de gevolgen kunnen verschillen. Een fout kan in de ene zaak grote gevolgen hebben, terwijl dezelfde fout in een andere zaak geen grote consequenties heeft. Stel, in een zaak zijn de DNA-profielen van twee verdachten verwisseld. Dit zal dan geen gevolgen hebben als het DNA-profiel van het spoor niet met dat van één van de verdachten matcht. Echter, als het DNA-profiel van het spoor wel met het DNA-profiel van één van de verdachten matcht dan zal door de verwisseling het DNA-bewijs de verkeerde verdachte in verband brengen met het misdrijf (en de andere verdachte ten onrechte vrijpleiten).

De kans dat de match tussen twee DNA-profielen in een specifieke zaak op een fout berust hangt dus af van verschillende factoren, zoals de complexiteit van de zaak, de eigenschappen en kwaliteit van het te onderzoeken materiaal, het aantal bemonsteringen, eventueel herhaalde analyses en de kwaliteit van de betrokken onderzoekers. Het is daarom niet verstandig om op basis van de algemene statistieken de kans op een fout in een specifieke zaak in te schatten. Deze kans zal op zaakniveau moeten worden beoordeeld, in de context van al de beschikbare gegevens en resultaten van die zaak.

Indien noodzakelijk kunnen deskundigen worden verzocht een persoonlijke inschatting te geven van dit soort kansen. Deze inschatting zal altijd een subjectief oordeel van de deskundige zijn en geen goed onderbouwde getalsmatige kans.

Bewijswaarde van een match

Voor het antwoord op de ultieme vraag of in het geval van een match van het DNA-profiel van een verdachte met dat van een spoor, deze verdachte daadwerkelijk betrokken is bij het delict zijn de volgende kansen van belang:

- 1 De kans dat het spoor van een persoon is die daadwerkelijk bij het delict is betrokken (dader). Met andere woorden, wat is de delictgerelateerdheid van het spoor? Celmateriaal, en dus DNA kan om onschuldige redenen aanwezig zijn, of zelfs met kwade opzet moedwillig door iemand zijn geplaatst;
- 2 De kans dat het DNA-profiel van een bloedverwant van de met het spoor matchende verdachte matcht met DNA-profiel van het spoor;
- 3 De kans dat het DNA-profiel van een willekeurig gekozen, niet-verwante persoon matcht met het DNA-profiel van het spoor. Dit is een uitspraak over de zeldzaamheid van het vastgestelde DNA-profiel, uitgedrukt in de berekende frequentie van voorkomen van het DNA-profiel in een populatie;
- 4 De kans dat de match het gevolg is van een fout.

De deskundige rapporteert alleen de onder punt 3 genoemde kans, die voor volledige DNA-profielen altijd kleiner is dan één op één miljard, maar voor onvolledige DNA-profielen veel groter kan zijn. De zeldzaamheid van het DNA-profiel is voor de rechter van cruciaal belang om te bepalen in hoeverre hij rekening moet houden met de mogelijkheid dat de waargenomen match een 'toevallige' match met een onschuldige persoon is. Daarom moet de gerapporteerde berekende frequentie van het matchende DNA-profiel dus altijd worden meegewogen, ook al zijn in een specifieke zaak de kansen op de andere mogelijkheden die tot een foute conclusie kunnen leiden veel groter (of kleiner: hoe de verhouding ligt kan sterk per zaak verschillen).

Als de kans op een fout relatief groot is, dan is de bewijswaarde van de match veel minder dan de berekende frequentie van het DNA-profiel wellicht suggereert. Door sommigen wordt daarom voorgesteld om in de rapportage één gecombineerde kans te geven waarin zowel de berekende frequentie van het DNA-profiel als de kans op een fout is verwerkt (Thompson et al. 2003). Het NFI is echter van mening dat de rechter het beste is gediend met een apart antwoord op elke vraag. De zienswijze van het NFI wordt ook internationaal gedragen (National Research Council 1996 en Buckleton et al. 2005). Zie voor een toelichting hierover het onderstaande kader.

Zoals in Vuistregel 2 (Brontekst 8) gesteld, geldt dat DNA-bewijs altijd moet worden beschouwd in de context van de zaak. Zo kunnen ook gemaakte fouten worden herkend, bijvoorbeeld als het resultaat van het DNA-onderzoek niet aansluit bij het overige bewijs in de zaak. Bij gerede twijfel over de juistheid van een DNA-resultaat is het daarom van groot belang dat dit aan de deskundige wordt gemeld, zodat er aanvullend onderzoek of een contra-expertise kan plaatsvinden.

De kans op een fout verdisconteren in de kans op een match?

Thompson et al. (2003) laten in hun artikel 'How the probability of a false positive affects the value of DNA-evidence' zien dat de bewijswaarde van een match tussen twee DNA-profielen aanzienlijk vermindert als men hierin een relatief grote kans op een fout verdisconteert. Zij stellen dat voor het rapporteren van het DNA-bewijs de berekende frequentie van het DNA-profiel van het matchende spoor niet volstaat en dient te worden gecorrigeerd met de kans op een fout in de onderzoeksketen. Door een ingeschatte kans op een fout in de onderzoeksketen (een 'fout-positief' resultaat) van bijvoorbeeld één op 10.000 te combineren met de berekende frequentie van het matchende DNA-profiel van het spoor van bijvoorbeeld één op één miljard zal het DNA-bewijs veel minder sterk lijken dan de berekende frequentie alleen suggereert*.

Juist bij zeldzame DNA-profielen zal deze methode een grote invloed hebben op de bewijswaarde. Het probleem bij het verdisconteren van de kans op een fout in de berekende frequentie van het verkregen DNA-profiel is echter dat, zoals uit bovenstaande tekst blijkt, de grootte van deze kans niet exact is weer te geven. Het gaat er immers om wat de kans is dat in de desbetreffende zaak fouten met een bepaald gevolg zijn gemaakt en niet wat in algemene zin de kans op een fout is (Balding 2005). Voor een volledig beeld van de bewijswaarde zouden bovendien ook de kans dat het spoor van de dader is en de kans dat het spoor van een verwante is verdisconteerd moeten worden. Verdiscontering van dergelijke zeer subjectief gekozen kansen tot één alles overkoepelend getal zou echter een exactheid suggereren die geen onderbouwing heeft. Het NFI volstaat daarom met het rapporteren van de berekende frequentie en het attenderen op het belang van de overige kansen die een rol spelen.

* N B Overigens lost het verdisconteren van de kans op een fout het probleem niet op dat iemand door een fout ten onrechte kan worden veroordeeld. Ook bij verdiscontering van de kans op een fout zal de bewijswaarde nog relatief hoog zijn.