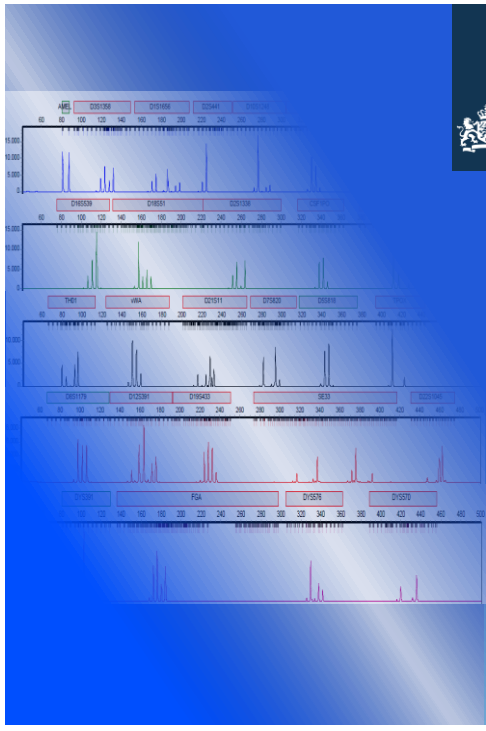


Netherlands Forensic Institute
Ministry of Justice and Security

PowerPlex Fusion 6C Profile analysis & interpretation

Corina Benschop

12/11/2020
Division Biological Traces

1

Netherlands Forensic Institute
Ministry of Justice and Security

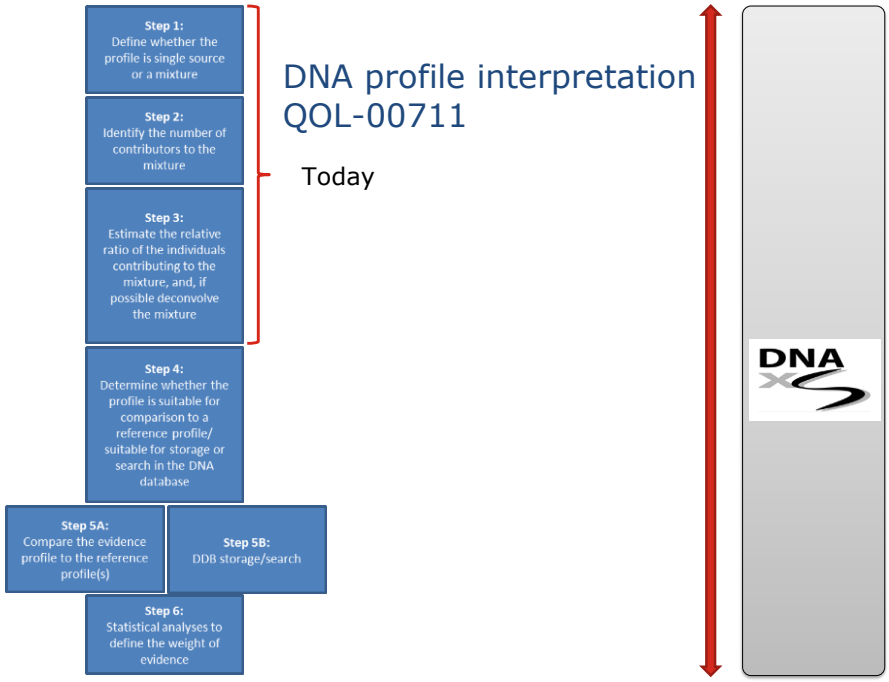
Interpretation of mixed DNA profiles

- Clayton *et al.* (1998) Forensic Science International 91: 55–70

Step 1	Identify the presence of a mixture
Step 2	Designate allele peaks
Step 3	Identify the number of potential contributors
Step 4	Estimate the relative ratio of the individuals contributing to the mixture
Step 5	Consider all possible genotype combinations
Step 6	Compare reference samples

2

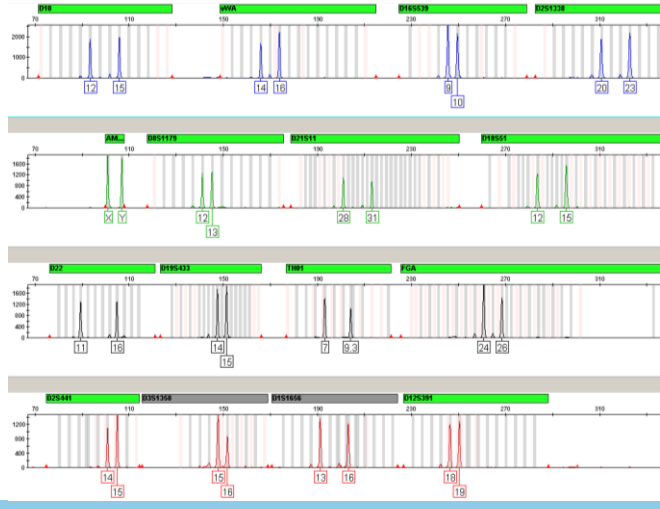
2



3

Netherlands Forensic Institute
Ministry of Justice and Security

DNA profile interpretation is easy!

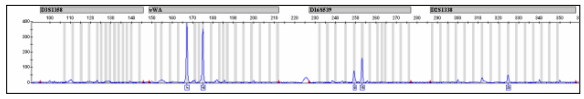


4

Netherlands Forensic Institute
Ministry of Justice and Security

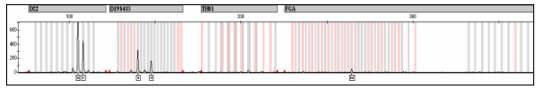
Unless…

- PCR inhibitors in DNA extract



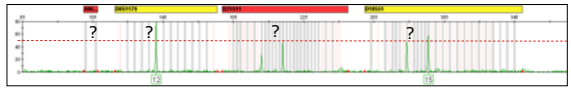
May reduce PCR amplification efficiency for some or all loci

- Degraded DNA (low quality)



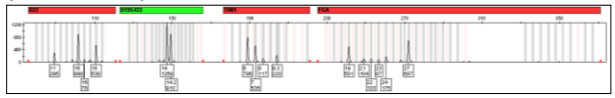
May make some allele targets unavailable (less intact target fragments)

- Low-template DNA (low quality)



DNA-profile suffers from stochastic variation

- (Unbalanced) DNA mixtures and/or relatives involved



Alleles may be masked in the electropherograms

5

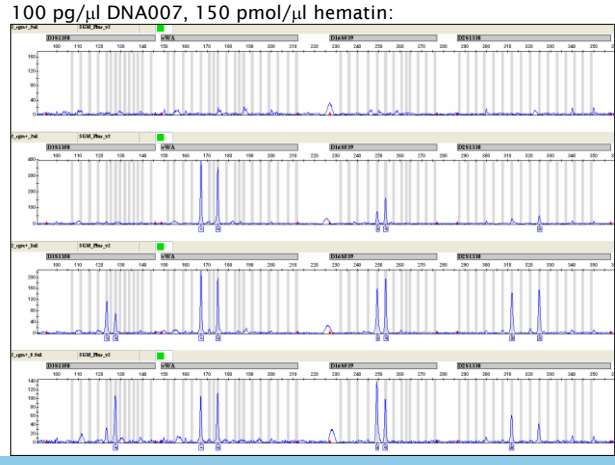
5

Netherlands Forensic Institute
Ministry of Justice and Security

PCR inhibitors

- PCR inhibitors interact with DNA or interfere with the DNA polymerase

100 pg/µl DNA007, 150 pmol/µl hematin:



SGM+ 28c:

5 µl input

2 µl input

1 µl input

0.5 µl input

6

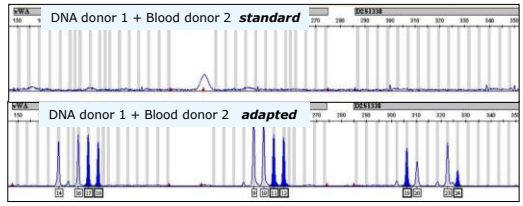
6

Netherlands Forensic Institute
Ministry of Justice and Security

Prevent PCR inhibition

- Avoid unnecessary collection of inhibitors
- Decrease the PCR input
- Purify the DNA extract
- Use a different DNA polymerase
- Use an STR kit with optimised buffer

e.g. NGM, PPF6C



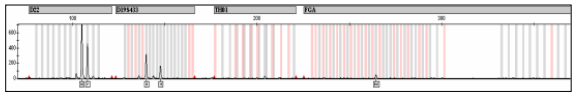
7

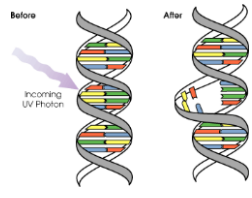
7

Netherlands Forensic Institute
Ministry of Justice and Security

DNA degradation

- The DNA may be highly fragmented or chemically modified
- Different DNA degradation mechanisms
- Due to *e.g.* long time exposure to UV-light, high temperatures





Possible solutions

- Multiple STR typing kits with different locus order
- mini-STRs, mtDNA, SNPs
- WGA prior to the PCR? Repair enzymes? Etc. …

9

9

Netherlands Forensic Institute
Ministry of Justice and Security

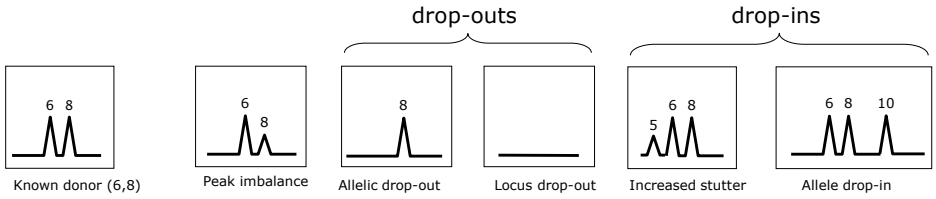
Low template (LT) DNA

- LT DNA/ trace DNA/ touch DNA:

`any sample which falls below recommended thresholds at any stage of the process and can not be defined by a precise picogram amount'

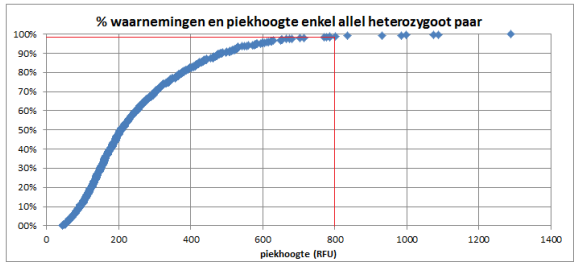
- DNA profiles with peaks below the stochastic threshold:

`threshold below which stochastic amplification artefacts can be expected'



Netherlands Forensic Institute
Ministry of Justice and Security

On the stochastic threshold



* Inception: VAL-00039 PowerPlex Fusion 6C sporen validatie

- Stochastic threshold of 800 rfus includes 98.9% of the single alleles

True genotype	Detected	Stochastic threshold risks
Homozygous {a,a}	1 allele {a}	Too high: allele is labeled {a,F} ⇒ risk of false inclusion
Heterozygous {a,b}	1 allele {a}	Too low: allele is labeled {a,a} ⇒ risk of false exclusion

Netherlands Forensic Institute
Ministry of Justice and Security

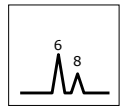
Heterozygote balance

Dependent on amount of DNA

High template

Balanced

Low-template



Peak imbalance can occur

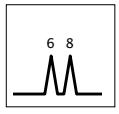
Pg input	gemiddelde piekbalans per input	stdev	# loci
63pg *(n=6)*	75%	15%	110
125pg *(n=6)*	79%	15%	120
250pg *(n=2)*	85%	12%	40
500pg *(n=2)*	89%	9%	40
1000pg *(n=2)*	91%	7%	40

* Inception: VAL-00039 PowerPlex Fusion 6C sporen validatie

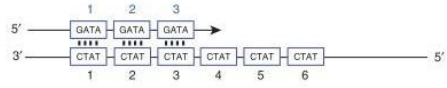
Netherlands Forensic Institute
Ministry of Justice and Security

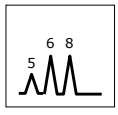
Stutter product formation

- Strand slippage during PCR



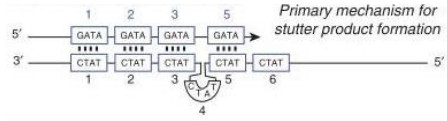
Normal replication

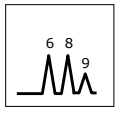




Increased -1 repeat unit (back stutter)

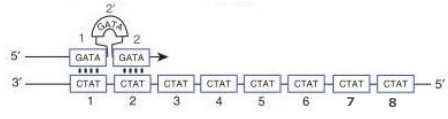
Deletion caused by forward slippage





Increased +1 repeat unit (forward stutter)

Insertion caused by backward slippage

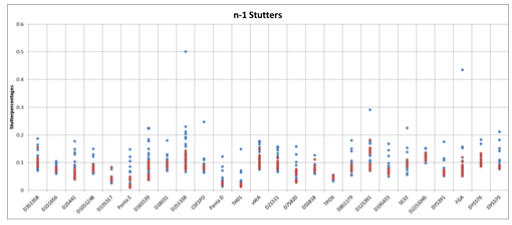


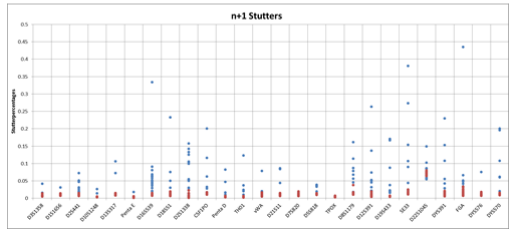
Netherlands Forensic Institute
Ministry of Justice and Security

Elevated stutter products

Red = high input
Blue = low input

Elevated stutters occur more often with LT than with HT DNA

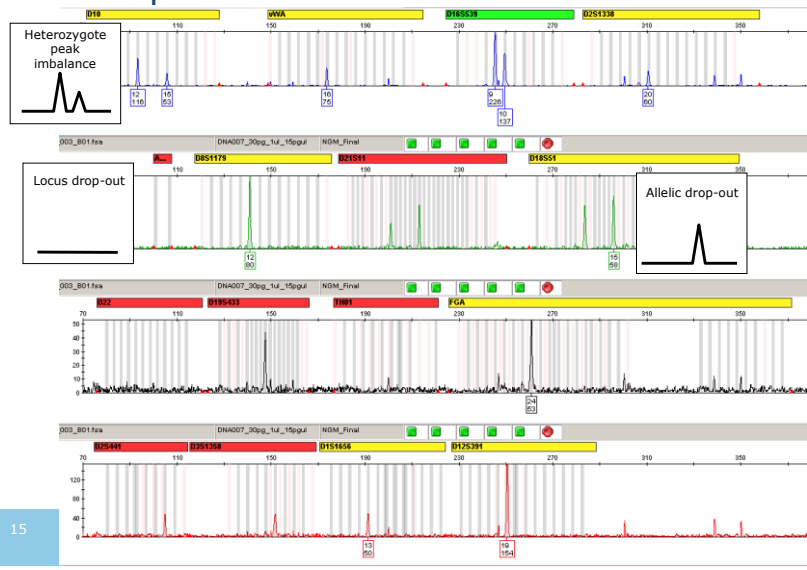




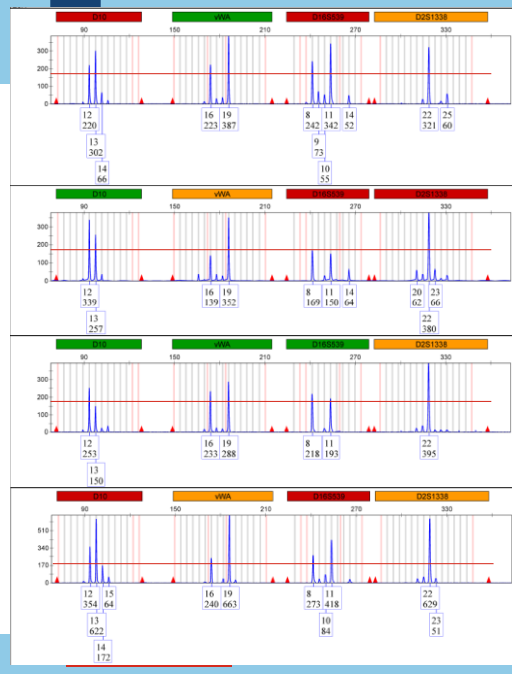
* Inception: VAL-00039 PowerPlex Fusion 6C sporen validatie

Netherlands Forensic Institute
Ministry of Justice and Security

LT DNA profile

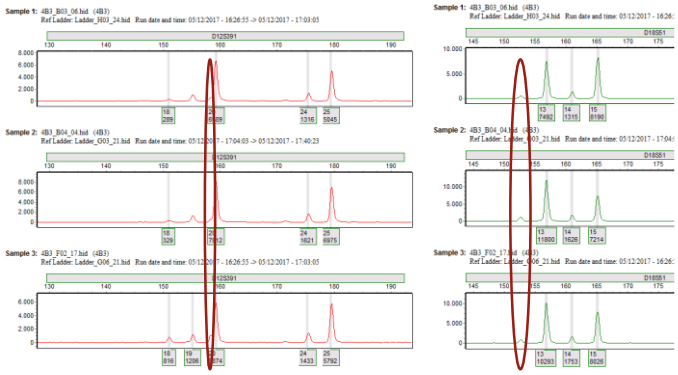


Replicate analysis



Netherlands Forensic Institute
Ministry of Justice and Security

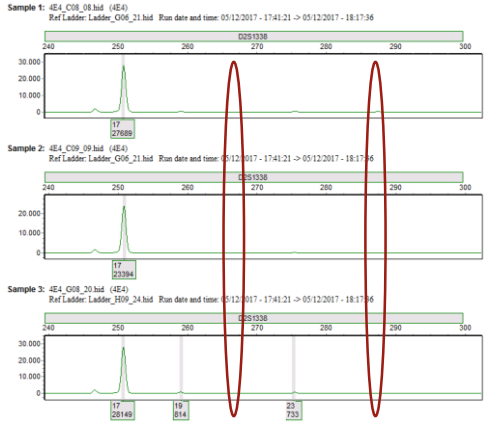
Drop-outs in replicates



Donor 1				20		25
Donor 2	18		19,3			
Donor 3		19			24	

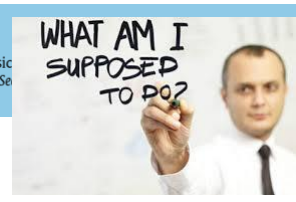
Donor 1		13		15
Donor 2	12	13		
Donor 3			14	

Netherlands Forensic Institute
Ministry of Justice and Security

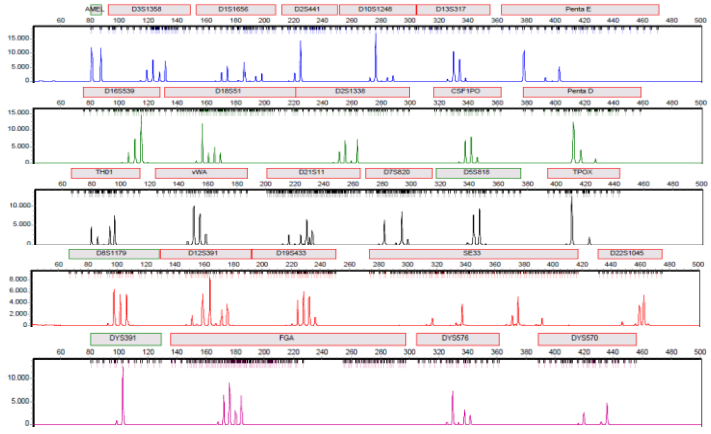


Donor 1	17				
Donor 2	17				26
Donor 3	17			23	
Donor 4		19	20		

Netherlands Forensic
Ministry of Justice and Se



Next: profile interpretation



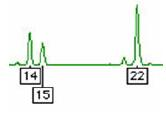
Identify the presence of a mixture

- **By extra bands**
- **By peak imbalance**

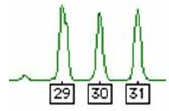
- Exceptions: Genetic aberrations
  - Trisomy - three peaks of equal size
  - Somatic mutations - three peaks of unequal size
    - Can be tissue specific
    - Hairs have a higher rate
    - Many somatics will be missed as stutters
    - Not inherited (unless gametic)
  - Unbalanced heterozygote - e.g. XXY

How often are genetic mutations observed?
Trisomy –
http://cstl.nist.gov/biotech/strbase/tri_tab.htm

Locus	Frequency
Amelogenin	0.1%
D21S11	0.0015%
D18S51	0.001%
D8S1179	0.004%
FGA	0.002%
vWA	0.0013%
TH0	0.00002%



Type 1: Sum of heights of two peaks is equal to third peak



Type 2: Balanced peak heights

Identify the number of potential contributors

➢ Maximum allele count (MAC) approach:

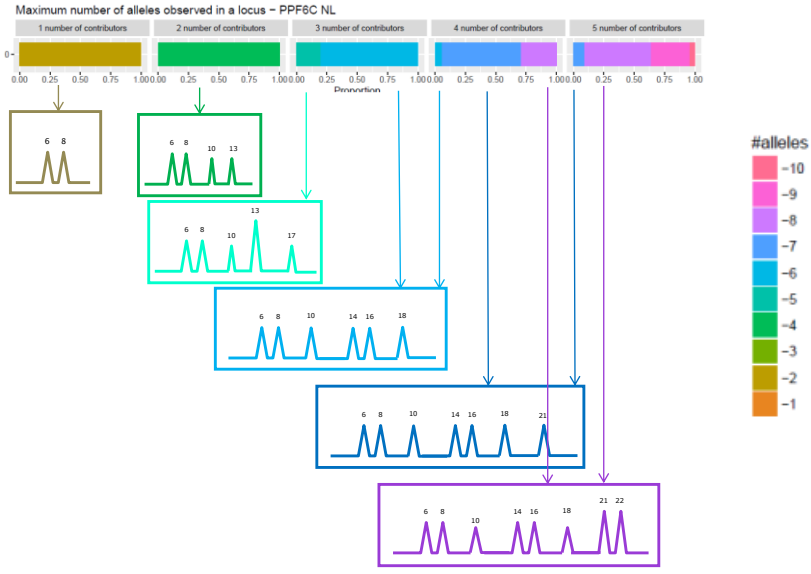
- At most 3 or 4 alleles/locus ➡ at least two donors
- At most 5 or 6 alleles/locus ➡ at least three donors
- >6 alleles/locus ➡ the exact number of donors cannot be reliably determined

➢ Mischaracterisation rates with MAC can be high, specifically with high order mixtures (≥3 contributors)

➢ The number of loci examined and their respective discriminating power, can reduce the mischaracterisation rates

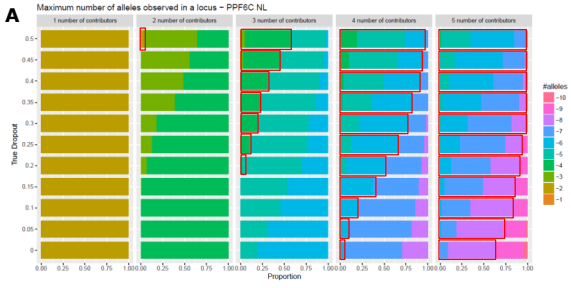
Netherlands Forensic Institute
Ministry of Justice and Security

Maximum allele count (MAC)

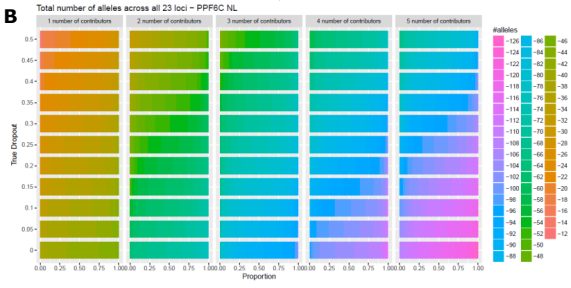


22

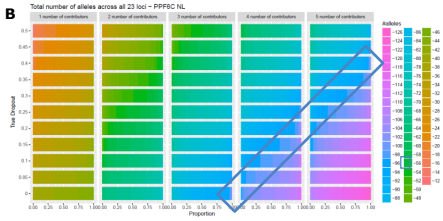
Maximum allele count (MAC)

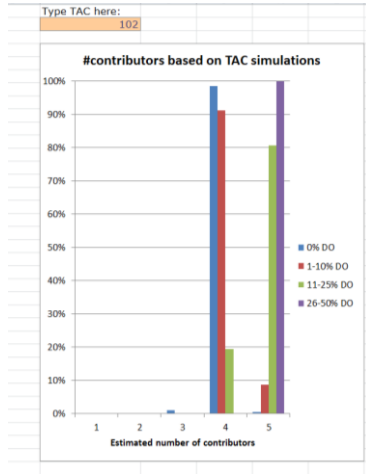


Total allele count (TAC)



23



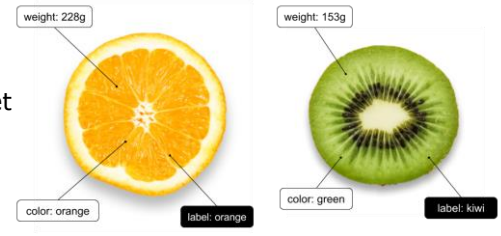


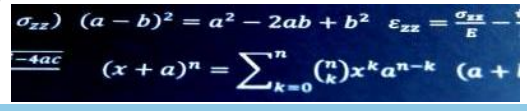
G:\BI\BI_DigiD

Excel file 'PPF6C_TAC_nC-tool_v1'

NOC-tool based on Machine Learning

- Features
  - Characteristics of the data
- Training, test and hold-out set
- Algorithms
- Model = algorithm + features
  - Performance





Netherlands Forensic Institute
Ministry of Justice and Security

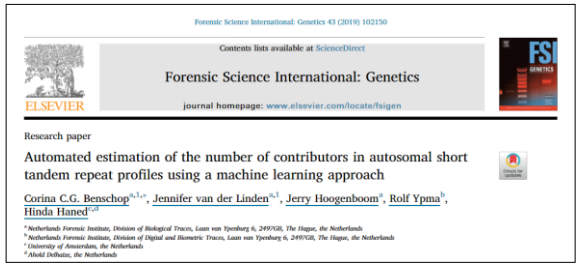
Machine learning approach

- Data collection
- Features calculation
- Feature selection
- Parameter optimization
- Model selection
- Model validation

Netherlands Forensic Institute
Ministry of Justice and Security

Two NOC models

- PPF6C model (RFC19)
- Generic Model (2.0-Generic_Model, RFC11)



Forensic Science International: Genetics 43 (2019) 102150

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsigen

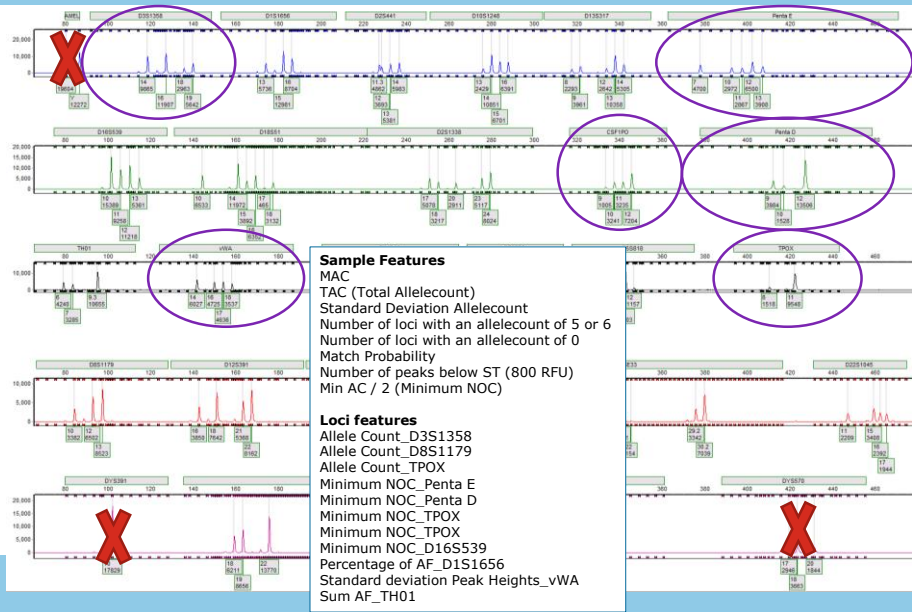
Research paper

Automated estimation of the number of contributors in autosomal short tandem repeat profiles using a machine learning approach

Corina C.G. Benschop[a,1,*], Jennifer van der Linden[a,1], Jerry Hoogenboom[a], Rolf Ypma[b], Hinda Haned[c,d]

[a] Netherlands Forensic Institute, Division of Biological Traces, Laan van Ypenburg 6, 2497GB, The Hague, the Netherlands
[b] Netherlands Forensic Institute, Division of Digital and Biometric Traces, Laan van Ypenburg 6, 2497GB, The Hague, the Netherlands
[c] University of Amsterdam, the Netherlands
[d] Ahold Delhaize, the Netherlands

Netherlands Forensic Institute
Ministry of Justice and Security



Sample Features
MAC
TAC (Total Allelecount)
Standard Deviation Allelecount
Number of loci with an allelecount of 5 or 6
Number of loci with an allelecount of 0
Match Probability
Number of peaks below ST (800 RFU)
Min AC / 2 (Minimum NOC)

Loci features
Allele Count_D3S1358
Allele Count_D8S1179
Allele Count_TPOX
Minimum NOC_Penta E
Minimum NOC_Penta D
Minimum NOC_TPOX
Minimum NOC_TPOX
Minimum NOC_D16S539
Percentage of AF_D1S1656
Standard deviation Peak Heights_vWA
Sum AF_TH01

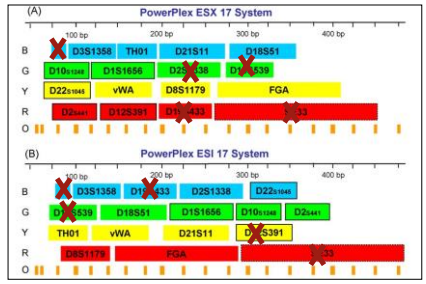
28

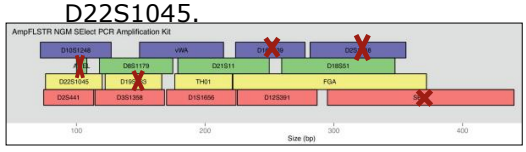
Netherlands Forensic Institute
Ministry of Justice and Security

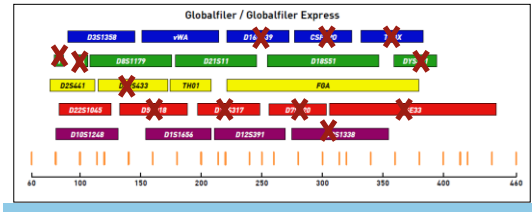
2.0 Generic Model (kit independent)

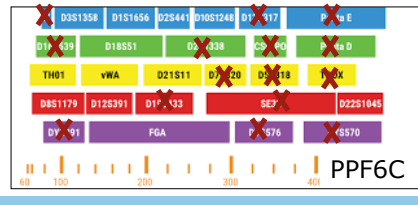
ESS core loci:

FGA, TH01, VWA, D1S1656, D2S441, D3S1358, D8S1179, D10S1248, D12S391, D18S51, D21S11, D22S1045.









29

29

Netherlands Forensic Institute
Ministry of Justice and Security

MAC vs PPF6C RFC19 model

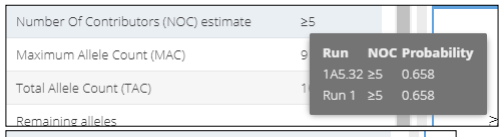
	MAC			PPF6C model		
True number of contributors (*n* per method)	**Under-estimated**	**Correctly estimated**	**Over-estimated**	**Under-estimated**	**Correctly estimated**	**Over-estimated**
2 (n=90)	0%	67%	33%	0%	100%	0%
3 (n=88)	0%	97%	3%	3%	97%	0%
4 (n =89)	20%	77%	3%	17%	83%	0%
5 (n =87)	63%	37%	0%	40%	60%	not applicable
Total (n =354)	21%	69%	10%	15%	85%	0%

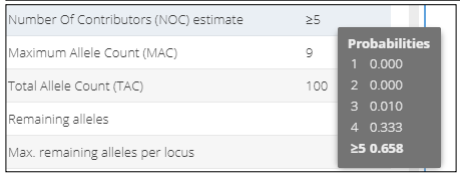
MAC vs Generic RFC11 Model

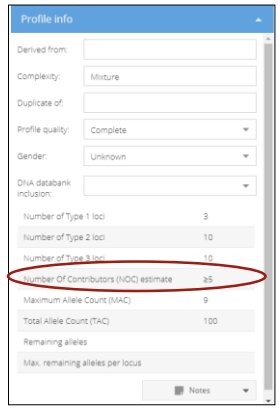
True NOC	**Percentage of correct predictions**											
	PPF6C (NL)			NGMSE (AU)			GlobalFiler (FR)			GlobalFiler (SL)		
	n	MAC	Generic Model	n	MAC	Generic Model	n	MAC	Generic Model	n	MAC	Generic Model
Total	48	64.6%	95.8%	42	61.9%	71.4%	42	71.4%	78.6%	47	51.1%	70.2%
% increase correct predictions MAC vs Model	31.2%			9.5%			7.2%			19.1%		

Netherlands Forensic Institute
Ministry of Justice and Security

Implementation DNAxs



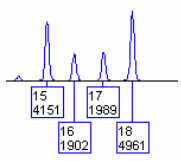




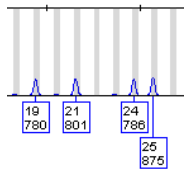
Netherlands Forensic Institute
Ministry of Justice and Security

Step 3:
Estimate the relative ratio of the individuals contributing to the mixture, and, if possible deconvolve the mixture

Estimate relative ratio of the contributors, if possible deconvolve



Indiv 1: 15/18
Indiv 2: 16/17



Indiv 1: 19/21
Indiv 2: 24/25
Or
Indiv 1: 19/24
Indiv 2: 21/25
Or
Indiv 1: 21/24
Indiv 2: 19/25
Or …

Netherlands Forensic Institute
Ministry of Justice and Security

Deconvolution

Options for deconvolution/inference of a major contributor's profile, e.g.:

- Manual

- GeneMapper ID-X
- MasterMix
- **LoCIM-tool**

- Probabilistic using e.g.:
  - **EuroForMix**
  - DNAmixtures
  - STRmix
  - TrueAllele


ELSEVIER

Forensic Science International
91 (1998) 41–53

Forensic Science International

Interpreting simple STR mixtures using allele peak areas

P. Gill[a,*], R. Sparkes[a], R. Pinchin[a], T. Clayton[b], J. Whitaker[b], J. Buckleton[c]

[a]Forensic Science Service, Priory House, Gooch Street North, Birmingham B5 6QQ, UK
[b]Forensic Science Service, Wetherby Laboratory, Sandbeck Way, Audby Lane, Wetherby, West Yorkshire LS22 4DN, UK
[c]ESR Forensic, Mt Albert Science Centre, Private Bag 92-021, Auckland, New Zealand

Received 13 May 1997; received in revised form 9 October 1997; accepted 27 October 1997


Forensic Science International: Genetics 11 (2014) 154–165

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

ELSEVIER

FSI GENETICS

LoCIM-tool: An expert's assistant for inferring the major contributor's alleles in mixed consensus DNA profiles

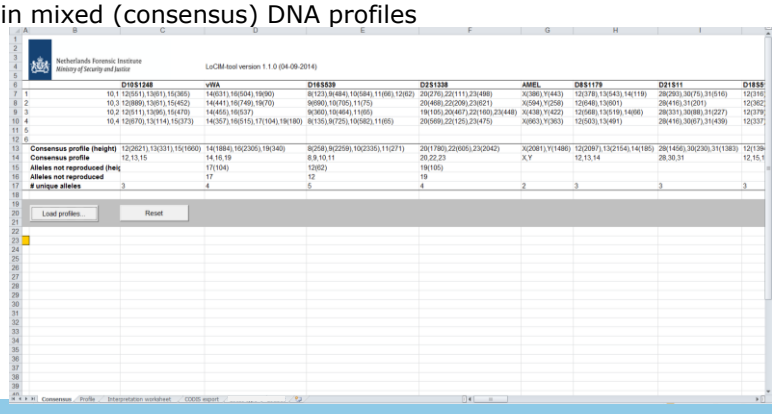
Corina C.G. Benschop*, Titia Sijen

Department of Human Biological Traces, Netherlands Forensic Institute, P.O. Box 24044, 2490 AA The Hague, The Netherlands

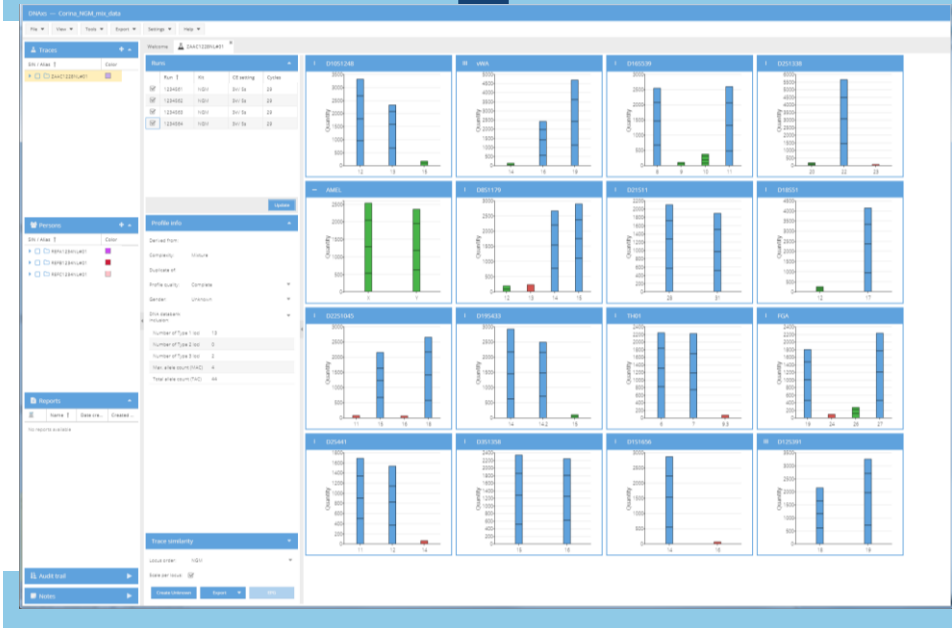
Netherlands Forensic Institute
Ministry of Justice and Security

LoCIM–tool

- **Lo**cus **C**lassification & **I**nference of the **M**ajor
- Automated tool to deduce the major component's alleles in mixed (consensus) DNA profiles



Netherlands Forensic Institute
Ministry of Justice and Security



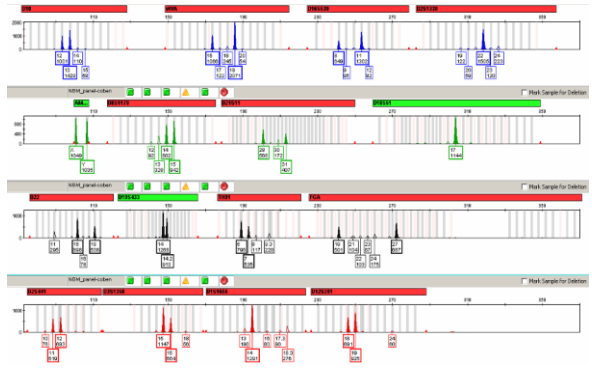
Locus Classification

- Locus classification into three types:

- Type 1
- Type 2
- Type 3

(-Locus drop-out)

Increasing complexity for inferring the major



Locus Classification

- Locus classification based on:
  - PHs
  - Stochastic threshold
  - Heterozygote balance (HB)
  - Major to minor(s) ratio

> Theoretical PHs are assigned to the alleles in the consensus using the sum of the PHs in the individual amplifications.

	Alleles detected at locus D16S539				
	8	9	10	11	12
PHs amplification 1	123	484	584	66	62
PHs amplification 2	-[1]	690	705	75	-
PHs amplification 3	-	360	464	65	-
PHs amplification 4	135	725	582	65	-
Alleles in *n/2* consensus	8	9	10	11	-
PHs assigned to alleles in the consensus (sum of PHs)	258	2259	2335	271	-

[1] The allele is not detected or below the detection threshold of 50 rfus.

Netherlands Forensic Institute
Ministry of Justice and Security

Locus Classification - criteria

PPF6C 29c 1.2kV24s	PH alleles major	HB alleles major	M^{hetero}:m ratio	M^{homo}:m ratio
Type 1	All ≥multiplied stochastic threshold	≥0.60	≥4:1	≥8:1
Type 2	-	≥0.60	≥2:1	≥4:1
Type 3	-	-	-	-

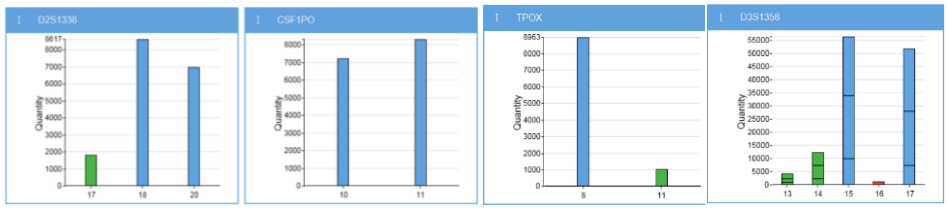
Increasing complexity for inferring the major

- Inferences include the allele with the highest summed PH (in the consensus) plus the alleles that have a PH within a specific percentage
  - Type 1 and type 2 loci: 50%
  - Type 3 loci: 33%

Netherlands Forensic Institute
Ministry of Justice and Security

Typical examples

Type 1 loci

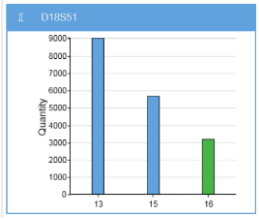


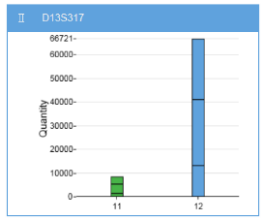
PPF6C 29c 1.2kV24s	PH alleles major	HB alleles major	M^{hetero}:m ratio	M^{homo}:m ratio
Type 1	All ≥multiplied stochastic threshold	≥0.60	≥4:1	≥8:1
Type 2	-	≥0.60	≥2:1	≥4:1
Type 3	-	-	-	-

Netherlands Forensic Institute
Ministry of Justice and Security

Typical examples

Type 2 loci





PPF6C 29c 1.2kV24s	PH alleles major	HB alleles major	M^{hetero}:m ratio	M^{homo}:m ratio
Type 1	All ≥multiplied stochastic threshold	≥0.60	≥4:1	≥8:1
Type 2	-	≥0.60	≥2:1	≥4:1
Type 3	-	-	-	-

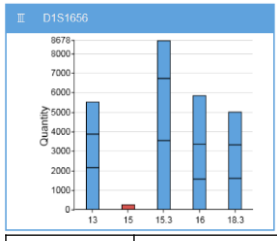
41

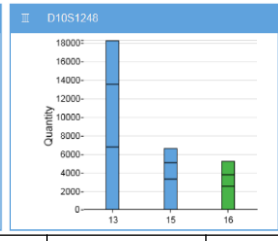
41

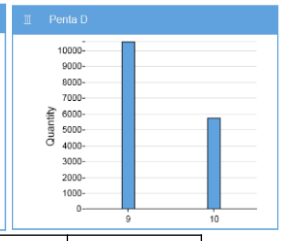
Netherlands Forensic Institute
Ministry of Justice and Security

Typical examples

Type 3 loci







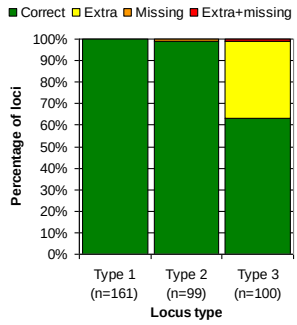
PPF6C 29c 1.2kV24s	PH alleles major	HB alleles major	M^{hetero}:m ratio	M^{homo}:m ratio
Type 1	All ≥multiplied stochastic threshold	≥0.60	≥4:1	≥8:1
Type 2	-	≥0.60	≥2:1	≥4:1
Type 3	-	-	-	-

42

42

Netherlands Forensic Institute
Ministry of Justice and Security

Inference of the Major

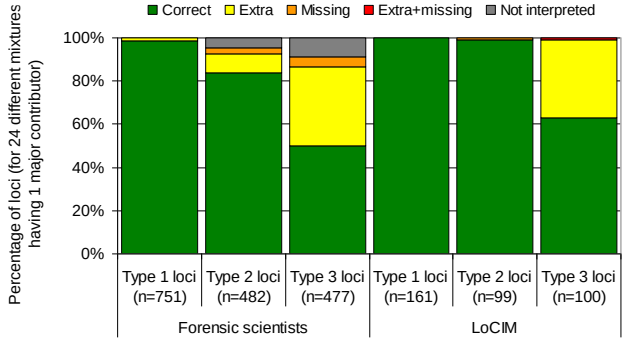


- Results for inference hold for consensus profiles generated from less amplifications and/or different CE setting, though with little shifts in the numbers for type 1, 2 and 3 loci
- LoCIM can be applied to single amplifications, though when an allele is not represented for the major the inference will be incorrect

Netherlands Forensic Institute
Ministry of Justice and Security

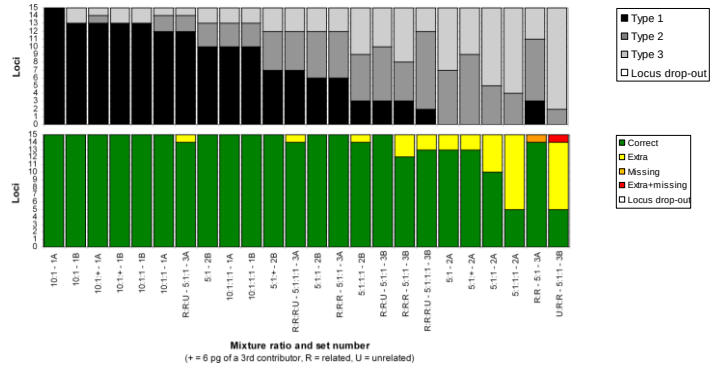
Manual vs automated inference

- Results of the survey among 19 forensic scientists



Netherlands Forensic Institute
Ministry of Justice and Security

Locus classification and inference results related



45

Netherlands Forensic Institute
Ministry of Justice and Security

Boundaries of the LoCIM-tool?

- Test set of 131 DNA mixtures:

A) 2p HT mixtures in ratios 1:2 – 1:6
B) 2p LT mixtures in ratios 1:2 – 1:6
C) 2p mixtures (1:5, 1:10, 1:15) with relatives or degraded DNA
D) 3p and 4p degraded DNA mixtures having one major that is 2, 5 or 10 times more abundant than the other contributors
E) 2p, 3p and 4p mixtures having no major contributor
F) 3p and 4p mixtures representing multiple HT contributors

- Examined performance of the LoCIM-tool
- Explored whether the tool can handle more types of mixtures
- Examined consequences when applying the criteria to mixtures having no major and only LT contributors

46

Netherlands Forensic Institute
Ministry of Justice and Security

LoCIM test set - Inference results

- Type 1 loci: 99.9% correct inference
  0.1% missing allele (1/1003)

- Type 2 & 3 loci: Inference results as expected

  LT and degraded DNA: When alleles are not detected (in the consensus) they cannot be inferred

Netherlands Forensic Institute
Ministry of Justice and Security

Reliability of inferences

- Type 1 loci inferences are reliable

- Type 2 loci inferences are reliable if one can assume that:
  1) the mixture includes one major contributor
  2) the major contributor has no drop-out in the consensus
  3) the DNA profile shows no (severe) degradation
  4) PHs and balance are not (severely) affected by allele sharing

- Type 3 loci inferences may include extra alleles

But also: Regard overall profile

Netherlands Forensic Institute
Ministry of Justice and Security

Reliability of inferences ➡ database actions?

But also: Regard overall profile

- Type 1 loci inferences are reliable
  - *Can be stored*

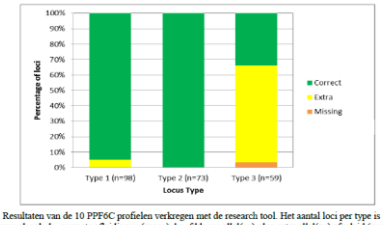
- Type 2 loci inferences are reliable if one can assume that:
  1) the mixture includes one major contributor
  2) the major contributor has no drop-out in the consensus
  3) the DNA profile shows no (severe) degradation
  4) PHs and balance are not (severely) affected by allele sharing
  - *Can be stored if the above assumptions hold*

- Type 3 loci inferences may include extra alleles
  - *Advise against storage in the DDB, if based on LoCIM solely*

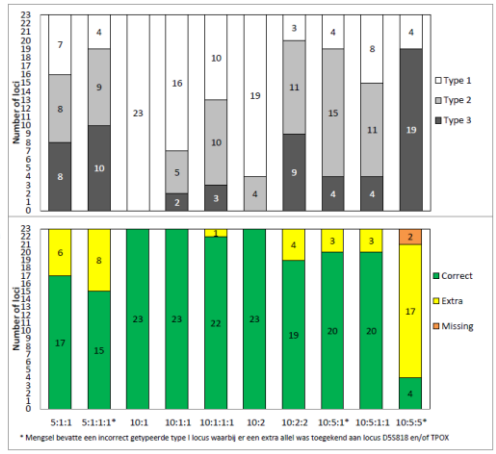
49

Netherlands Forensic Institute
Ministry of Justice and Security

PPF6C LoCIM performance

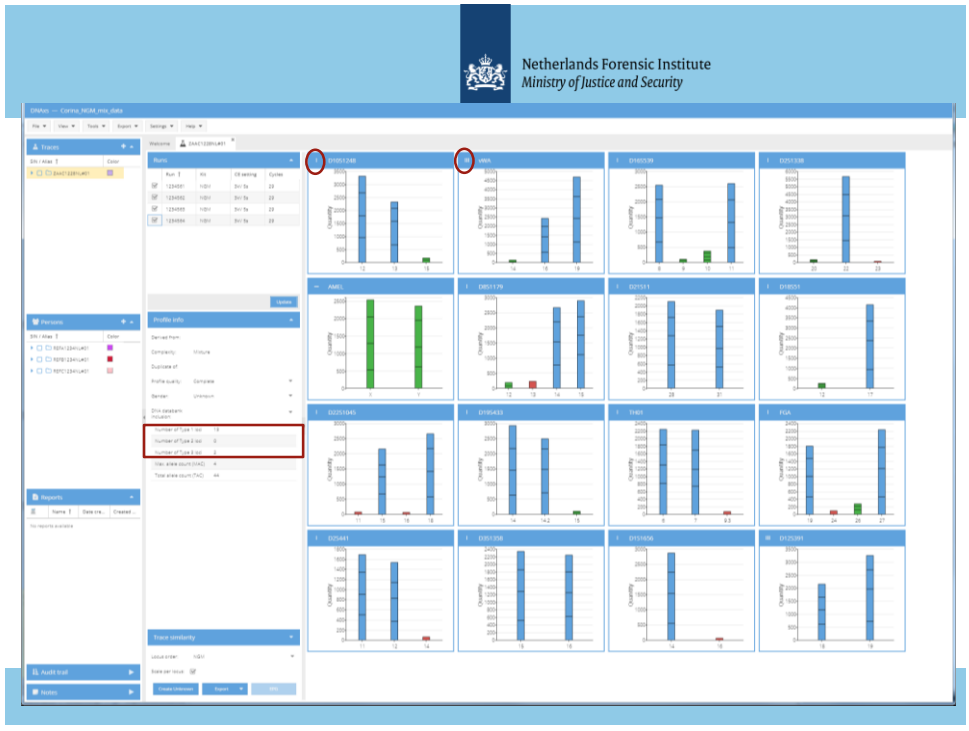


Figuur 8. Resultaten van de 10 PPF6C profielen verkregen met de research tool. Het aantal loci per type is weergegeven alsook de correcte afleidingen (groen), hoofddonor allel(en) plus extra allel(en) afgeleid (geel), en incomplete afleidingen (oranje).

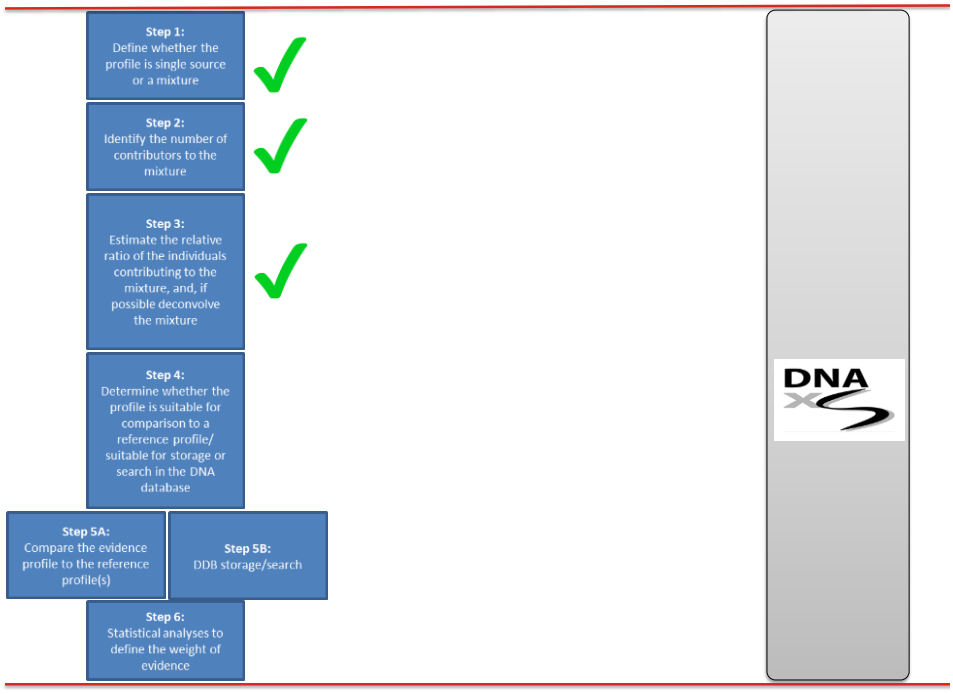


Figuur 9. Overzicht van de resultaten van de 10 PPF6C mengsels verkregen met de research tool. 1= 20 pg. De bovenste grafiek geeft het soort locus typeringen per mengsel aan (type 1, type 2 en type 3) en de onderste grafiek het aantal correcte afleidingen (groen), hoofddonor allel(en) plus extra allel(en) afgeleid (geel), en incomplete afleidingen (oranje) per mengsel.

50


Netherlands Forensic Institute
Ministry of Justice and Security

52


Step 1:
Define whether the profile is single source or a mixture
Step 2:
Identify the number of contributors to the mixture
Step 3:
Estimate the relative ratio of the individuals contributing to the mixture, and, if possible deconvolve the mixture
Step 4:
Determine whether the profile is suitable for comparison to a reference profile/ suitable for storage or search in the DNA database
Step 5A:
Compare the evidence profile to the reference profile(s)
Step 5B:
DDB storage/search
Step 6:
Statistical analyses to define the weight of evidence
DNA

53

Netherlands Forensic Institute
Ministry of Justice and Security

Next: Read and practice

To read in advance:

- QOL-00434 Validatierapport PPF6C sporen
- QOL-00711 Interpretatie van autosomale STR DNA-profielen
  (t/m 7. Deconvolutie functionaliteit in DNAxs)

To read after this training:

- Paper PPF6C LT
- Paper LoCIM-tool
- Paper en validatierapport NOC-tool

And have fun!