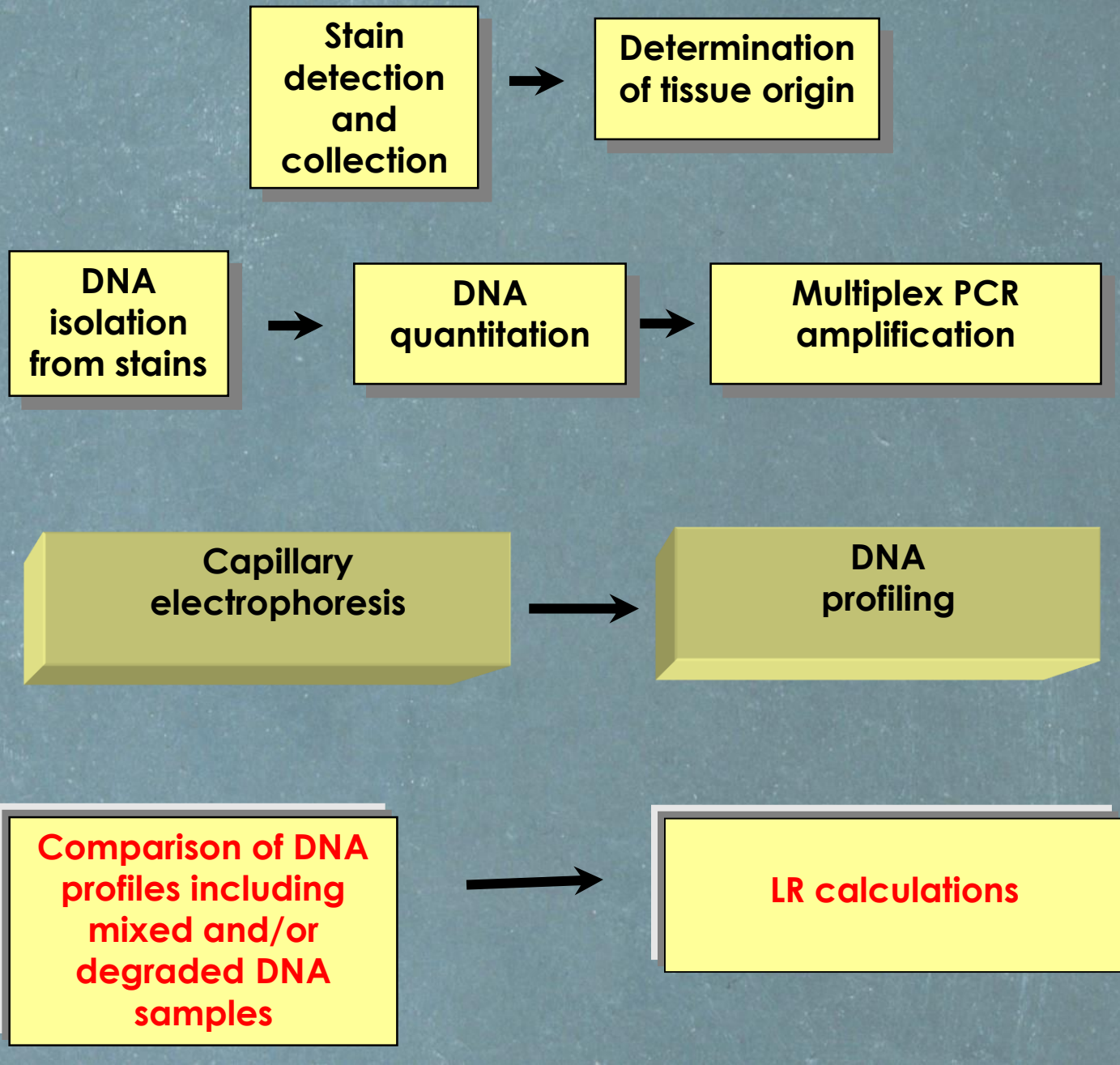
Forensic Genetics and Legal Medicine 2019-2020

29th April 2020

Advanced match interpretation

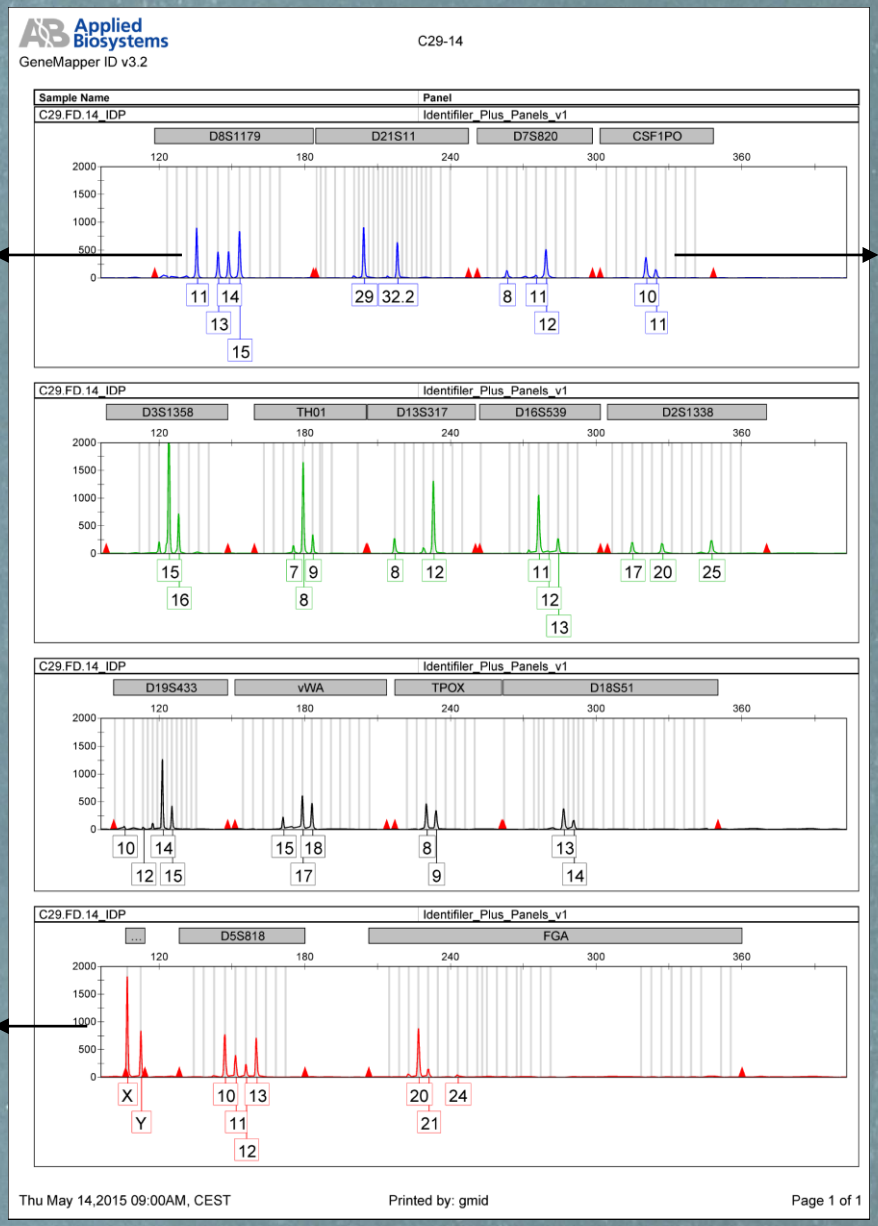

Stain detection and collection
Determination of tissue origin
DNA isolation from stains
DNA quantitation
Multiplex PCR amplification
Capillary electrophoresis
DNA profiling
Comparison of DNA profiles including mixed and/or degraded DNA samples
LR calculations

DNA MIXTURES

✓ Identify DNA mixtures

Multiple loci with >2 peaks (exclude PCR artifacts such as stutter, non-template addition, etc. and CE artifacts such as spikes, pull-ups)

Imbalanced signal for sex Chromosomes (exclude aneuploidies, e.g. 47,XXY)



Imbalanced 2-peak loci (in mixtures from relatives sharing alleles or very partial profiles no >2 peak loci may be found)

The effect of varying quantities of DNA from a genetically normal male and female on the peak area ratios

Mixture ratio		Dosage of alleles observed		Ratio of peak areas X:Y
Male (XY)	Female (XX)	X	Y	X:Y
10	1	12	10	1.2:1
5	1	7	5	1.4:1
4	1	6	4	1.5:1
3	1	5	3	1.6:1
2	1	4	2	2:1
1	1	3	1	3:1
1	2	5	1	5:1
1	3	7	1	7:1
1	4	9	1	9:1
1	5	11	1	11:1
1	10	21	1	21:1

Clayton et al., FSI 1998

Available online at www.sciencedirect.com
SCIENCE DIRECT
Forensic Science International 160 (2006) 90–101
Forensic Science International
www.elsevier.com/locate/forsciint

DNA commission of the International Society of Forensic Genetics: Recommendations on the interpretation of mixtures

P. Gill[a,*], C.H. Brenner[b], J.S. Buckleton[c], A. Carracedo[d], M. Krawczak[e], W.R. Mayr[f], N. Morling[g], M. Prinz[h], P.M. Schneider[i], B.S. Weir[j]

Probability of exclusion (PE) a.k.a. Random man not excluded (RMNE)

✓ Probability that a randomly chosen individual will carry at least one allele not present in the DNA mixture

pA1 = 0.1
pA2 = 0.1
pA3 = 0.1
q = 0.7

$PE = 2pq + q^2 = 1 - p^2 = 0.91$

$PE_n = 1 - [(1-PE_1) \times (1-PE_2) \ldots \times (1-PE_n)]$

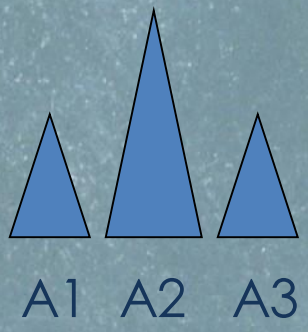
- it doesn't require an a priori definition of the number of contributors to the stain
- it doesn't require prior knowledge of the suspect's genotype
- it can't take account of stochastic events (drop out), so it's use is limited to unambiguous STR profiles

LR – <u>unrestricted</u> combinatorial approach

It can be expanded to include drop out /drop in events (semicontinuous models)

TO BE CONTINUED...

✓ Hp («prosecution hypothesis»): the suspect contributed to the stain
✓ Hd («defense hypothesis»): the suspect did not contribute to the stain



pA1 = 0.1
pA2 = 0.1
pA3 = 0.1

In general, the minimal number of contributors required to explain the mixed profile (i.e. max 4 alleles = 2 contributors, max 6 alleles = 3 contributors., etc.) is the one that maximizes the probability of the evidence under both Hp and Hd

- It needs an a priori definition of the number of contributors
- It needs the STR profile of the suspect

Let's assume 2 contributors with suspect's profile A2/A2

✓ Hp = 1 x 2pA1pA3
✓ Hd = 12pA1pA2pA3 (pA1 + pA2 + pA3)
✓ **LR** = 1 / 6pA2 (pA1 + pA2 + pA3) = **5,55**

A1/A2 + A1/A3	2pA1pA2 x 2pA1pA3
A1/A2 + A2/A3	2pA1pA2 x 2pA2pA3
A1/A2 + A3/A3	2pA1pA2 x $pA3^2$
A1/A3 + A1/A2	2pA1pA3 x 2pA1pA2
A1/A3 + A2/A3	2pA1pA3 x 2pA2pA3
A1/A3 + A2/A2	2pA1pA3 x $pA2^2$
A2/A3 + A1/A2	2pA2pA3 x 2pA1pA2
A2/A3 + A1/A3	2pA2pA3 x 2pA1pA3
A2/A3 + A1/A1	2pA2pA3 x $pA1^2$
A1/A1 + A2/A3	$pA1^2$ x 2pA2pA3
A2/A2 + A1/A3	$pA2^2$ x 2pA1pA3
A3/A3 + A1/A2	$pA3^2$ x 2pA1pA2

LR - restricted combinatorial approach

→ But It can be expanded to include drop out /drop in events (continuous models)

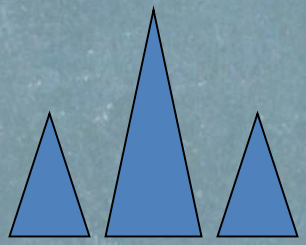
TO BE CONTINUED...

✓ It takes account of peak height
✓ It assumes that (within loci) peak heights obtained through PCR are proportional to the initial DNA inputs of the contributors

Let's assume again a 2 contributors mixture with suspect's profile A2/A2 at locus A.
Additional locus B provide further information regarding DNA ratios between the two contributors (in this example 1:1)

↓

- ✓ It can be done through approximation, or by more robust methods that calculate least squares residuals across loci
- ✓ It is hampered by expected allele imbalance in degraded DNA



Locus A

Locus B

✓ $Hp = 1 \times 2pA1pA3$

✓ $Hd = 2pA2^2 (4pA1pA3 + 1)$

✓ **LR** $= 1 / 6pA2^2 =$ **16,66**

~~A1/A2 + A1/A3~~	~~$2pA1pA2 \times 2pA1pA3$~~
A1/A2 + A2/A3	$2pA1pA2 \times 2pA2pA3$
~~A1/A2 + A3/A3~~	~~$2pA1pA2 \times pA3^2$~~
~~A1/A3 + A1/A2~~	~~$2pA1pA3 \times 2pA1pA2$~~
~~A1/A3 + A2/A3~~	~~$2pA1pA3 \times 2pA2pA3$~~
A1/A3 + A2/A2	$2pA1pA3 \times pA2^2$
A2/A3 + A1/A2	$2pA2pA3 \times 2pA1pA2$
~~A2/A3 + A1/A3~~	~~$2pA2pA3 \times 2pA1pA3$~~
~~A2/A3 + A1/A1~~	~~$2pA2pA3 \times pA1^2$~~
~~A1/A1 + A2/A3~~	~~$pA1^2 \times 2pA2pA3$~~
A2/A2 + A1/A3	$pA2^2 \times 2pA1pA3$
~~A3/A3 + A1/A2~~	~~$pA3^2 \times 2pA1pA2$~~

"Deconvoluting" a mixed stain

✓ "If a sample contains a predominance of one individual's DNA, that individual's DNA profile may be determined"

Scientific Working Group on DNA Analysis Methods (SWGDAM) Interpretation Guidelines

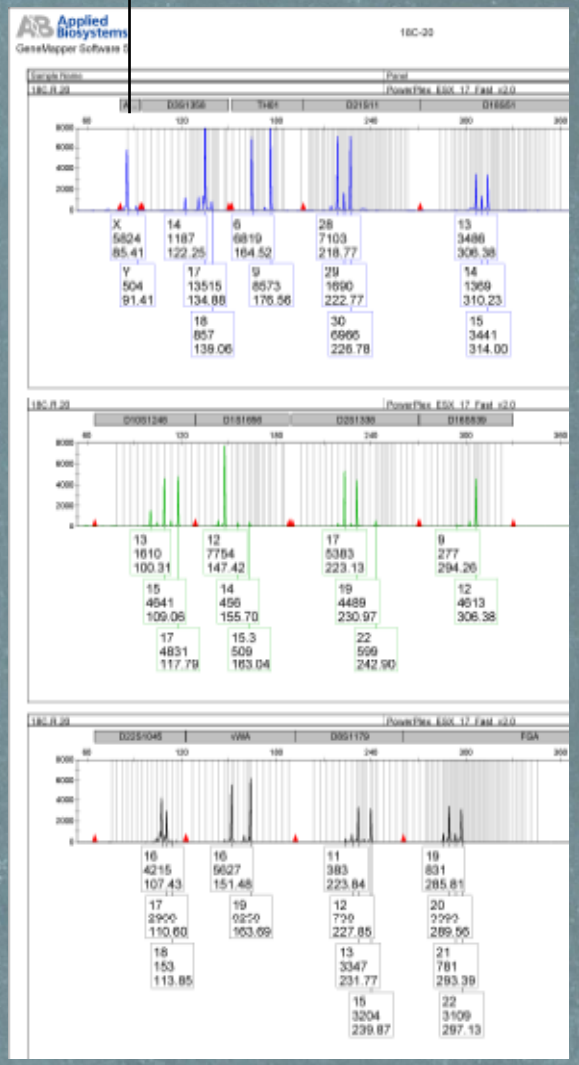
✓ "The conclusion of a major DNA profile from a single contributor in a mixed stain shall only be drawn if a peak height ratio of at least 4:1 (major vs minor component) is observed across all heterozygous DNA systems. In this case, the major DNA profile can be considered equivalent to that of a stain originating from a single person, and all calculations can be performed accordingly"

German Stain Commission

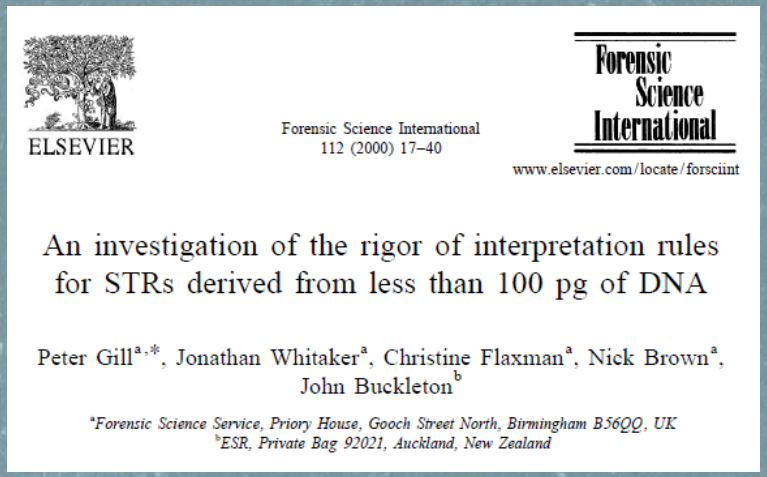
✓ For "second level" searches DNA profiles must be single source. A mixed DNA profile can be considered single source if a peak height ratio of at least 3:1 (major vs minor component) is observed across all heterozygous DNA systems

Italian DNA database DPR 7 april 2016 n 87

X:Y ratio ~11:1
female:male ratio ~5:1



2000 Low copy number (LCN) DNA – «biological model»



ELSEVIER

Forensic Science International
112 (2000) 17–40

Forensic Science International

www.elsevier.com/locate/forsciint

An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA

Peter Gill[a,*], Jonathan Whitaker[a], Christine Flaxman[a], Nick Brown[a], John Buckleton[b]

[a]Forensic Science Service, Priory House, Gooch Street North, Birmingham B56QQ, UK
[b]ESR, Private Bag 92021, Auckland, New Zealand

Following papers also suggested that also extension of electrokinetic injection time could be beneficial in LCN-DNA analysis by increasing peak height and thus reducing drop-out occurrence

✓ authors demonstrate that by increasing PCR cycle number (from standard ~28 up to 34) DNA profiles could be obtained even from PCR DNA input templates < 100 pg (LCN-DNA). Above 34 cycles no evident benefit could be seen due to loss of polymerase processivity

✓ LCN-DNA analysis is associated with specific interpretation challanges: risk of allele imbalance leading to possible drop out; risk of «contamination» (drop in); increase in stutter ratio

✓ Stochastic events (drop-out, drop-in) can be dealth with through replication of results

✓ Apparent mis-matches between crime-stain and a suspect DNA profile do not necessarily result in an exclusion

Sample	Amel	VWA	THO	D8	FG	D21	D18	D19	D3	D16	D2
R_1	XY	16,19	6,7	12,14	20,24	28,30	12,F	13,17	15,16	11,13	17,20
R_2	XY	16,19	6,F	12,14	20,24,25	28,30	12,F	13,17	15,16	11,13	17,20
Consensus	XY	16,19	6,F	12,14	20,24	28,30	12,F	13,17	15,16	11,13	17,20
Suspect	XY	16,19	6,7	12,14	20,24	28,30	12,12	13,17	15,16	11,13	17,20
Negative 1	X	14	–	14,15	–	–	–	15	15	–	–
Negative 2	X	14	–	–	–	–	–	14	16	5	–

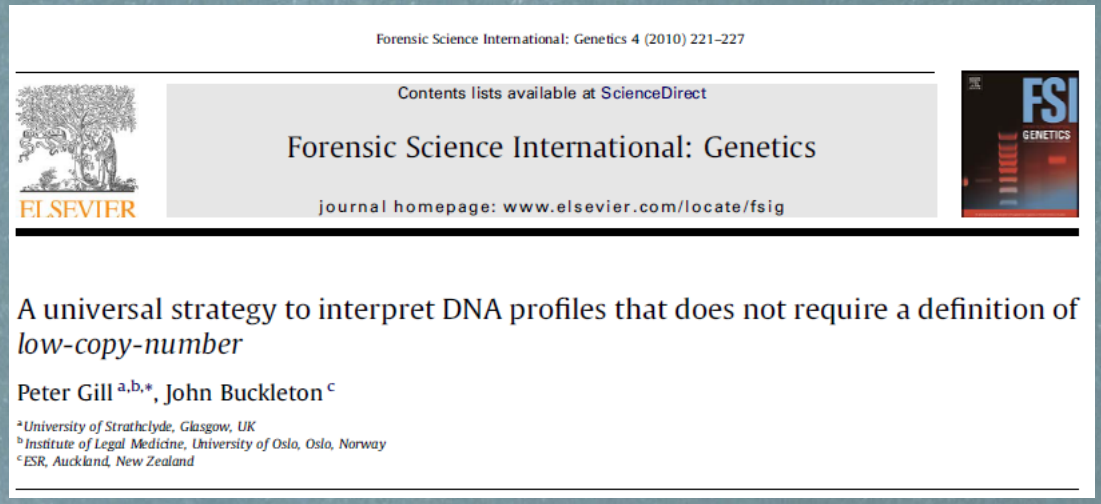
A consensus profile is built by reporting only alleles that are duplicated in PCR replicates

- in the provided example allele 25 at locus FG is not reported (drop-in)
- For loci in which a single allele is duplicated (TH0, D18) it cannot be excluded that any allele (F) have dropped out from the genotype
- For loci of such type (TH0, D8), whenever the suspect is heterozygous with an allele in common with the stain's duplicated allele, he should not be excluded as contributor and the LR calculated as 1 / 2p (where p is the population frequency of the allele observed in the stain)

✓ The biological model was developed in order to facilitate reporting of LCN-DNA profiles in the absence of software solutions which came later.

✓ Authors provided a statistical demonstration that validated the biological model, but that was cumbersome from practical casework in the absence of dedicated software.

2010 Low template (LT) DNA – towards a fully «statistical model»



Forensic Science International: Genetics 4 (2010) 221–227

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

A universal strategy to interpret DNA profiles that does not require a definition of *low-copy-number*

Peter Gill [a,b,*], John Buckleton [c]

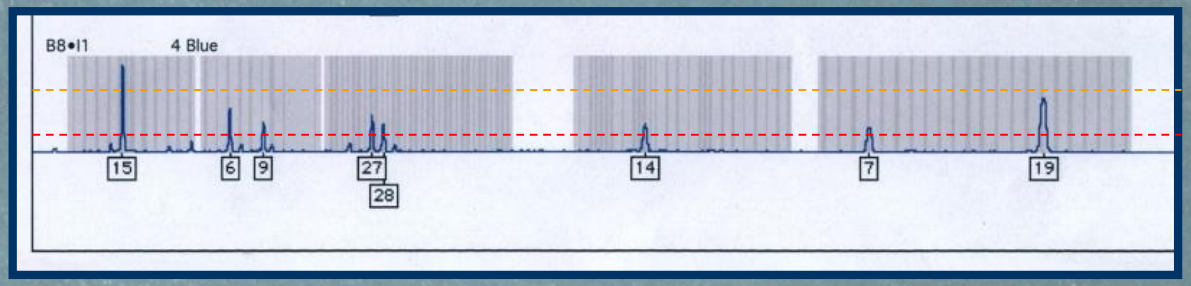
[a] University of Strathclyde, Glasgow, UK
[b] Institute of Legal Medicine, University of Oslo, Oslo, Norway
[c] ESR, Auckland, New Zealand

✓ the likely magnitude of stochastic events does not depend from the estimated quantity of DNA input in PCR, but from the actual peak height observed in epgs (in general, the smaller the peak height, the higher the risk that stochastic events occurred in the DNA profile)

Stochastic events (allele drop out, allele drop in)

When do I need to expect drop out?
When at an homozygous locus (single donor stains) or at any locus (in mixed stain) peaks below **stochastic threshold** are preserved

Can be empirically derived (e.g.average noise level + 3 SD)



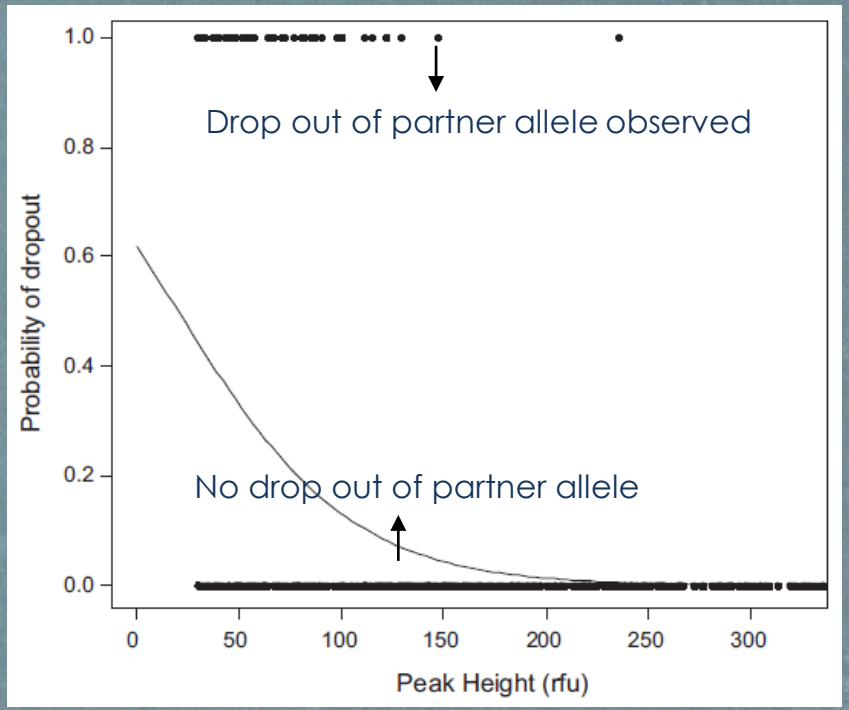
Analytical threshold (AT9 or limit of detection (LDT) signal below AT/LDT is background noise
Stochastic threshold (ST): highly unlikely that a locus showing a peak above ST will be affected by drop out

- ✓ Pr of drop in can be empirically calculated (e.g. counting the occurrende of peaks >LOD in PCR negative controls in a set of experiments); ~5% is a good operative estimate.
- ✓ Pr of drop out for any given peak height can be mathematically derived (e.g. by logistic regression through a set of experiments using series of samples of known genotypie and of varying quantity). See Gill et al. Forensic Sci Int Genet 2012 provided as supplementary material.

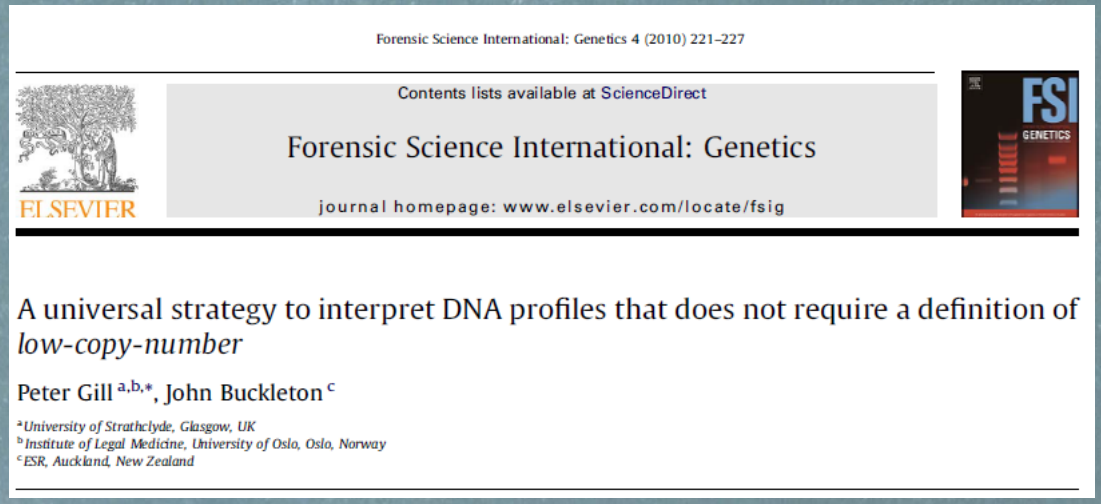
Sample no.	Allele designation	Allele peak height	Allele designation	Allele peak height	Drop-out state*
1	17	135	25	193	0
2	11	30	13	80	0
3	29	157	30	160	0
4	14	30	16	142	0
5	13	319	14	117	0
6	6	150	9.3	36	1
7	21	56	23	30	1

* Drop-out state = 0 means no drop-out of companion allele and drop-out state = 1 means drop-out is observed. All drop-out states are conditioned on alleles in the fourth/fifth columns.

- ✓ According to the present example, a ST of 200 rfu, above which the risk of drop out is negligible can be empirically set
- ✓ STR profiles with all peaks > 200 rfu can be analyzed disregarding drop out risk



2010 Low template (LT) DNA – towards a fully «statistical model»



Forensic Science International: Genetics 4 (2010) 221–227

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

A universal strategy to interpret DNA profiles that does not require a definition of *low-copy-number*

Peter Gill [a,b,*], John Buckleton [c]

[a] University of Strathclyde, Glasgow, UK
[b] Institute of Legal Medicine, University of Oslo, Oslo, Norway
[c] ESR, Auckland, New Zealand

✓ while the consensus approach tries to infer genotypes from replicates, the statistical model assesses the probability of DNA profiles under Hp and Hd from all possible genotypes, taking stochastic events (drop-in, drop-out) into account

✓ extension of the model to multiple replicates and mixed samples is straightforward

✓ dedicated software that overcomes the simplified consensus strategy is becoming available

2012

Forensic Science International: Genetics 6 (2012) 679–688

Contents lists available at SciVerse ScienceDirect

Forensic Science International: Genetics

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

DNA commission of the International Society of Forensic Genetics: Recommendations on the evaluation of STR typing results that may include drop-out and/or drop-in using probabilistic methods

P. Gill [a,b,*], L. Gusmão [c], H. Haned [d], W.R. Mayr [e], N. Morling [f], W. Parson [g], L. Prieto [h], M. Prinz [i], H. Schneider [j], P.M. Schneider [k], B.S. Weir [l]

[a] Norwegian Institute of Public Health, Oslo, Norway
[b] University of Oslo, Oslo, Norway
[c] IPATIMUP, Institute of Molecular Pathology and Immunology of the University of Porto, Portugal
[d] Netherlands Forensic Institute, Department of Human Biological Traces, The Hague, The Netherlands
[e] Department of Blood Group Serology and Transfusion Medicine, Medical University of Vienna, Austria
[f] Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark
[g] Institute of Legal Medicine, Innsbruck Medical University, Innsbruck, Austria
[h] Comisaría General de Policía Científica, University Institute of Research in Forensic Sciences (IUICP), Madrid, Spain
[i] Office of the Chief Medical Examiner, Department of Forensic Biology, New York, USA
[j] Hessisches Landeskriminalamt, Wiesbaden, Germany
[k] Institute of Legal Medicine, Faculty of Medicine, University of Cologne, Germany
[l] University of Washington, Department of Biostatistics, Seattle, USA

2016

Forensic Science International: Genetics 25 (2016) 191–197

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Research paper

DNA Commission of the International Society for Forensic Genetics: Recommendations on the validation of software programs performing biostatistical calculations for forensic genetics applications

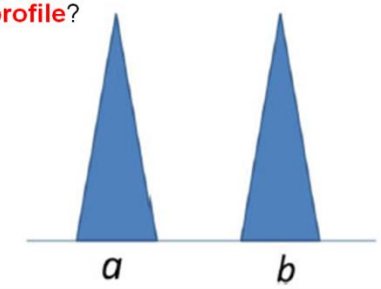
M.D. Coble[a,*], J. Buckleton[b,c], J.M. Butler[d], T. Egeland[e], R. Fimmers[f], P. Gill[g,h], L. Gusmão[i,j,k], B. Guttman[l], M. Krawczak[m], N. Morling[n], W. Parson[o,p], N. Pinto[j,k,q,r], P.M. Schneider[s], S.T. Sherry[t], S. Willuweit[u], M. Prinz[v]

[a] National Institute of Standards and Technology, Applied Genetics Group, Gaithersburg, MD, USA
[b] ESR, Private Bag 92021, Auckland 1142, New Zealand
[c] National Institute of Standards and Technology, Statistical Engineering Division (Guest Researcher), Gaithersburg, MD, USA
[d] National Institute of Standards and Technology, Special Programs Office, Gaithersburg, MD, USA
[e] Norwegian University of Life Sciences, Oslo, Norway
[f] Institute for Medical Statistics, Informatics, and Epidemiology, University Bonn, Germany
[g] Norwegian Institute of Public Health, Oslo, Norway
[h] University of Oslo, Oslo, Norway
[i] State University of Rio de Janeiro (UERJ), Rio de Janeiro, Brazil
[j] IPATIMUP, Institute of Molecular Pathology and Immunology of the University of Porto, Portugal
[k] Instituto de Investigação e Inovação em Saúde, University of Porto, Portugal
[l] National Institute of Standards and Technology, Software and Systems Division, Gaithersburg, MD, USA
[m] Institute of Medical Informatics and Statistics, Christian-Albrechts University of Kiel, Germany
[n] Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark
[o] Institute of Legal Medicine, Medical University of Innsbruck, Innsbruck, Austria
[p] Forensic Science Program, The Pennsylvania State University, PA, USA
[q] Institute for Research and Innovation in Health (I3S), University of Porto, Porto, Portugal
[r] Centre of Mathematics of the University of Porto, Porto, Portugal
[s] Institute of Legal Medicine, Faculty of Medicine, University of Cologne, Germany
[t] National Center for Biotechnology Information, National Library of Medicine, NIH, Bethesda, MD, USA
[u] Department of Forensic Genetics, Institute of Legal Medicine and Forensic Sciences, Charité—Universitätsmedizin, Berlin, Germany
[v] John Jay College of Criminal Justice, New York, USA

Trace:

✓ **Single profile**?

✓ Mixture?



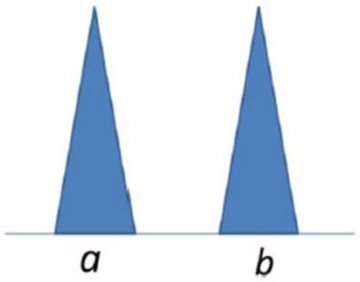
$P_a = 0{,}135$

$P_b = 0{,}200$

RMP = 2ab = 0.054

$$LR = \frac{Pr(E|Hp)}{Pr(E|Hd)} = 1 / 2ab = \mathbf{18.5}$$

Suspect

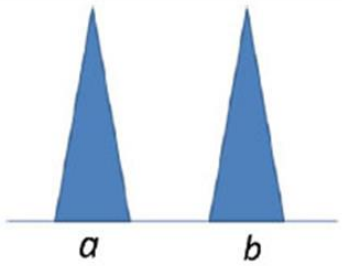


Statistical (probabilistic) model to approach LT-DNA

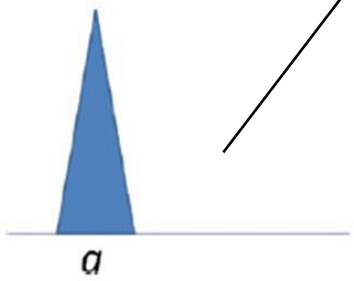
A simple solution is the "2p rule": admitting the possibility of drop out, the stain can carry any allele F (with frequency F=1)

$$2p_a p_F = 2p_a$$

$$LR = 1 / 2p_a$$



Reference profile (S)



Crime stain profile (E)

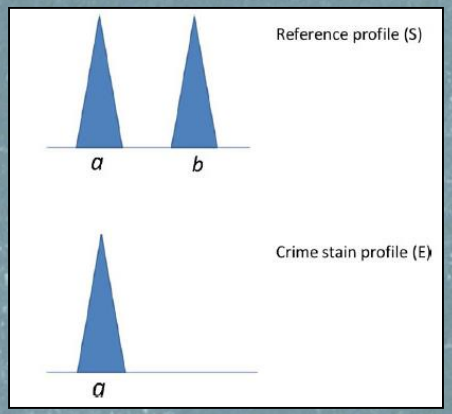
In a fully binary model, were an allele is rather present or not present (Pr drop out = 0), then RMP (and LR) = 0

In a fully probabilistic model the probability of the prosecutor's hypothesis is

$$0 < x < 1$$

$$LR = \frac{Pr(E|Hp)}{Pr(E|Hd)}$$

Semicontinuous model (peak height not incorporated in the model)



2p rule →

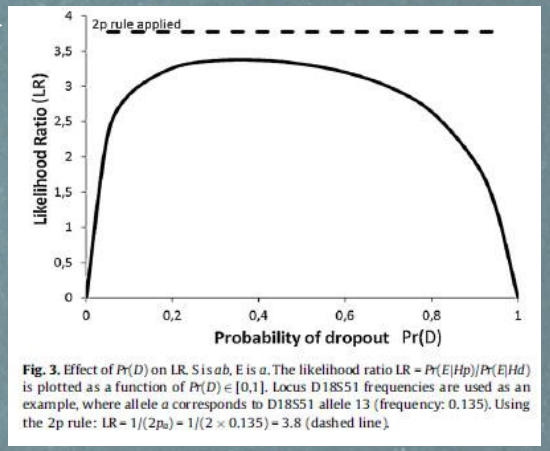


Fig. 3. Effect of $Pr(D)$ on LR. S is ab, E is a. The likelihood ratio $LR = Pr(E|Hp)/Pr(E|Hd)$ is plotted as a function of $Pr(D) \in [0,1]$. Locus D18S51 frequencies are used as an example, where allele a corresponds to D18S51 allele 13 (frequency: 0.135). Using the 2p rule: $LR = 1/(2p_a) = 1/(2 \times 0.135) = 3.8$ (dashed line).

Pr of drop out

Pr of no drop out

Pr of no drop in

$$LR = \frac{Pr(\bar{D})Pr(D)Pr(\bar{C})}{p_a^2 Pr(\bar{D}_2)Pr(\bar{C}) + 2p_a p_Q Pr(\bar{D})Pr(D)Pr(\bar{C}) + p_Q^2 Pr(D_2)Pr(\bar{C})p_a + 2p_Q p_{Q'} Pr(D)^2 Pr(C)p_a}$$

1) aa, no locus drop out*, no drop in

2) aQ, no drop out of heterozygous genotype, drop out of heterozygous genotype, no drop in

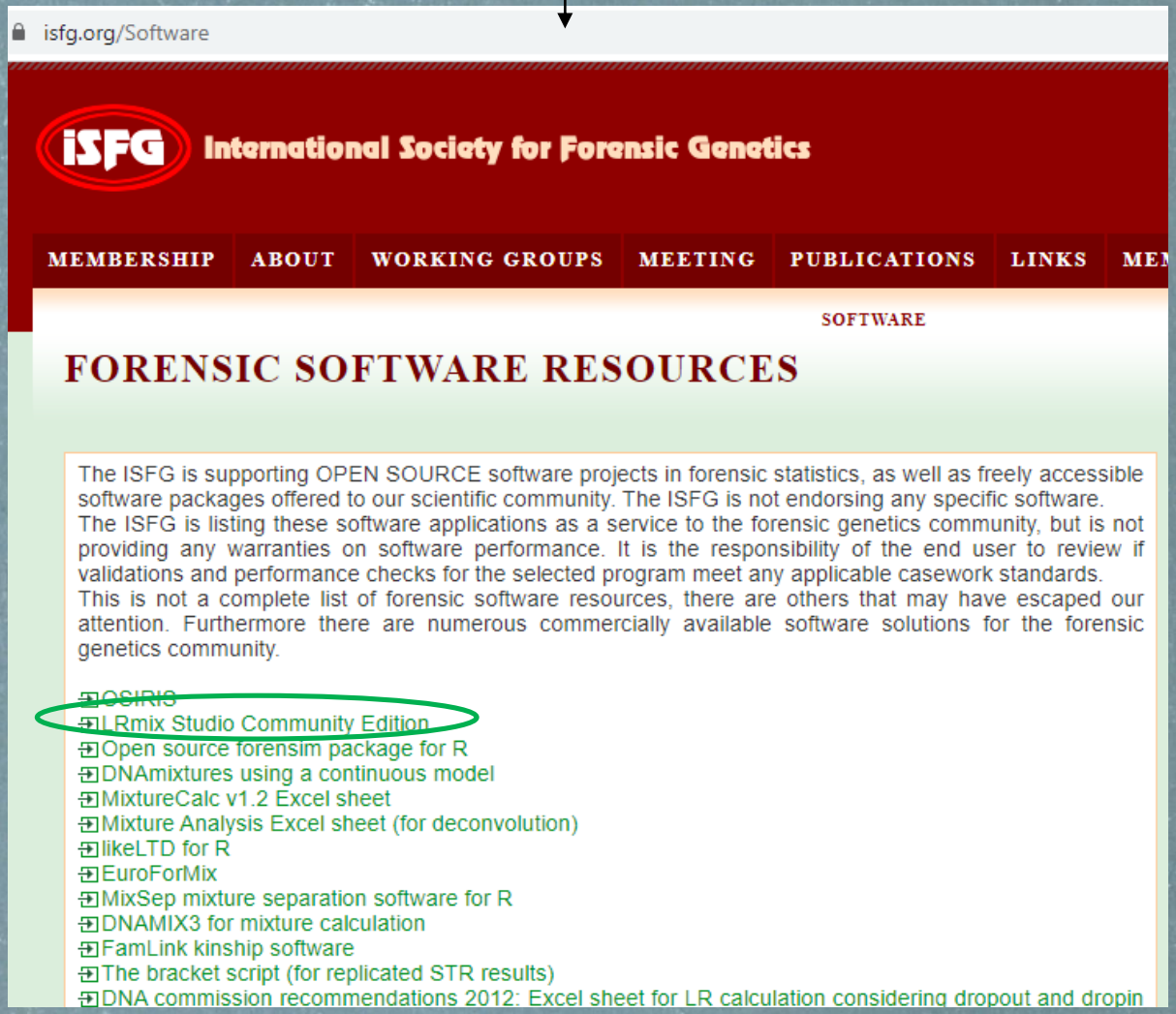
3) QQ' (with Q ≠ a and Q ≠ Q'), drop out of both alleles of heterozygous genotype, drop in of allele a

4) QQ, locus drop out*, drop in of allele a

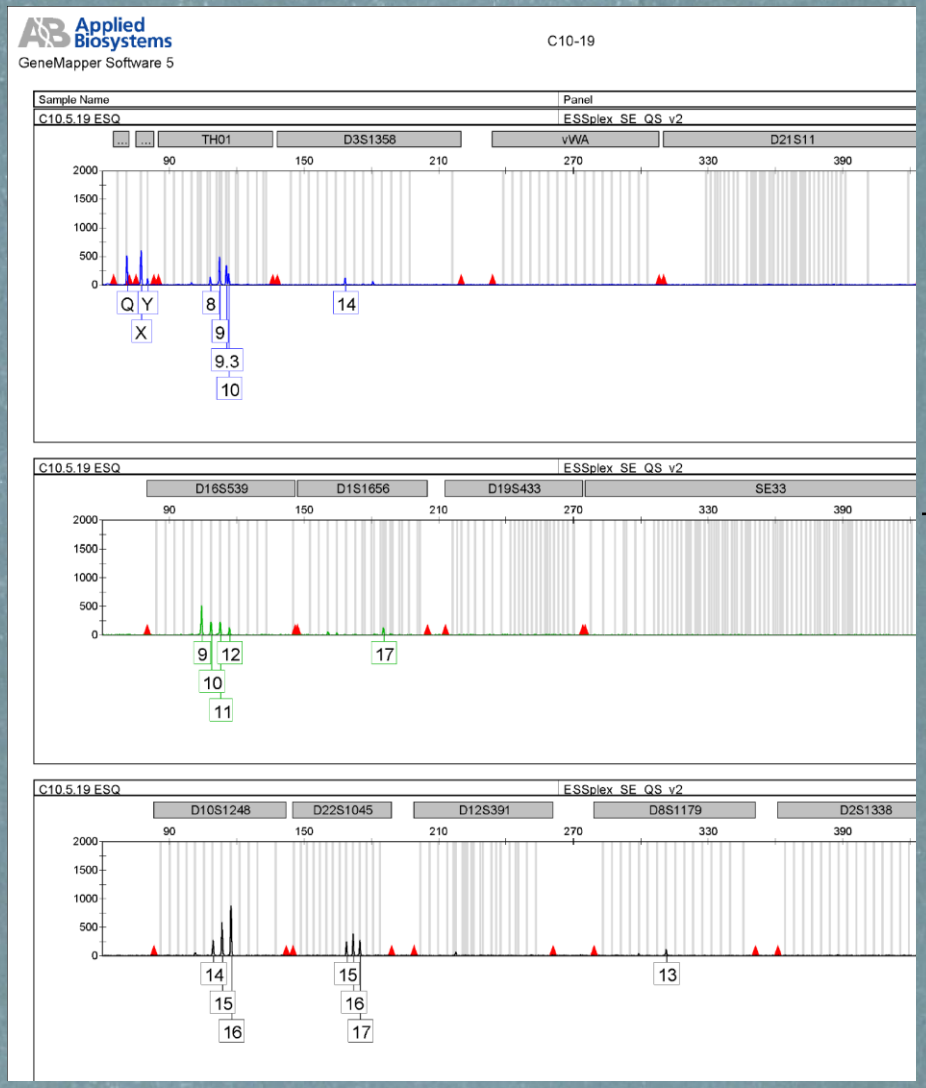
* ~ $Pr(D)^2$

Semicontinuous model (peak height not incorporated in the model)

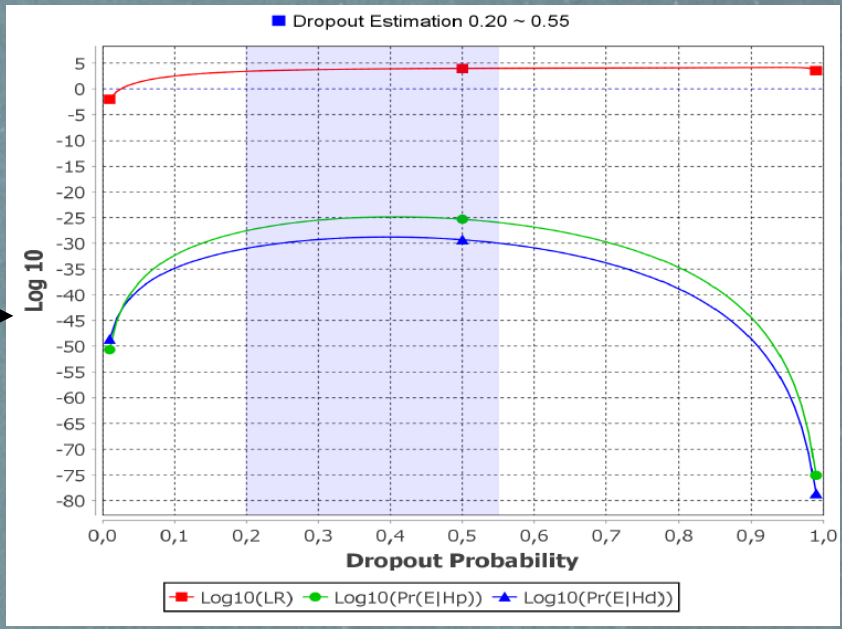
- ✓ the same model can be extended to mixed DNA profiles
- ✓ Commercial and open source dedicated software available



- ✓ LR values can be calculated for any Pr drop out, ranging from 0 to 100%
- ✓ A range of most likely Pr of drop out can be derived from the empirical distribution of the drop-out probability conditioned on the expected number of alleles observed relative to the genotype of the hypothesised contributors



Hp: suspect + 1 unknown
Hd: 2 unknowns



3 persons mixture (victim's DNA is assumed to be present in the stain) with 33 distinct alleles

Step 1: Simulate 1000 mixtures

Hp

Victim, Suspect } Fixed → 36 alleles

Unknown } By randomly sampling from the Dutch DNA frequencies → 17 alleles

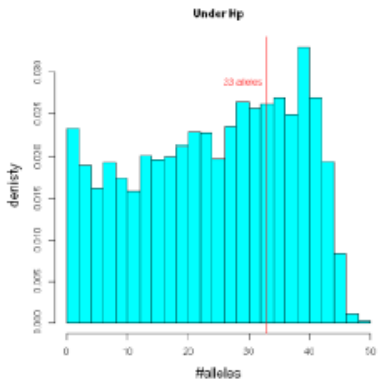
Mixture #1
51 distinct alleles

Mixture #2
50 distinct alleles

...

Mixture #1000
53 distinct alleles

Distribution of the numbers of alleles among 1000 mixtures

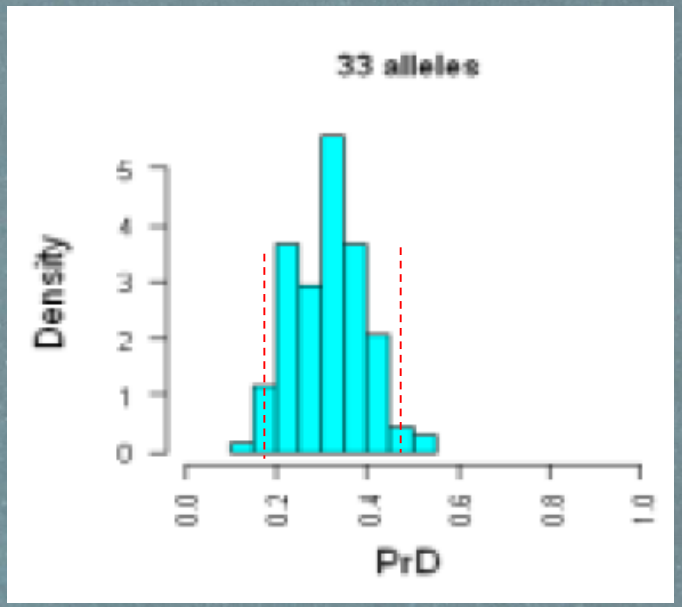


Mixture #1
51 alleles

Victim
Suspect
Unknown

drop out →

Pr(D)	# surviving alleles
0.01	50
0.02	49
...	...
0.50	25
...	
0.99	1

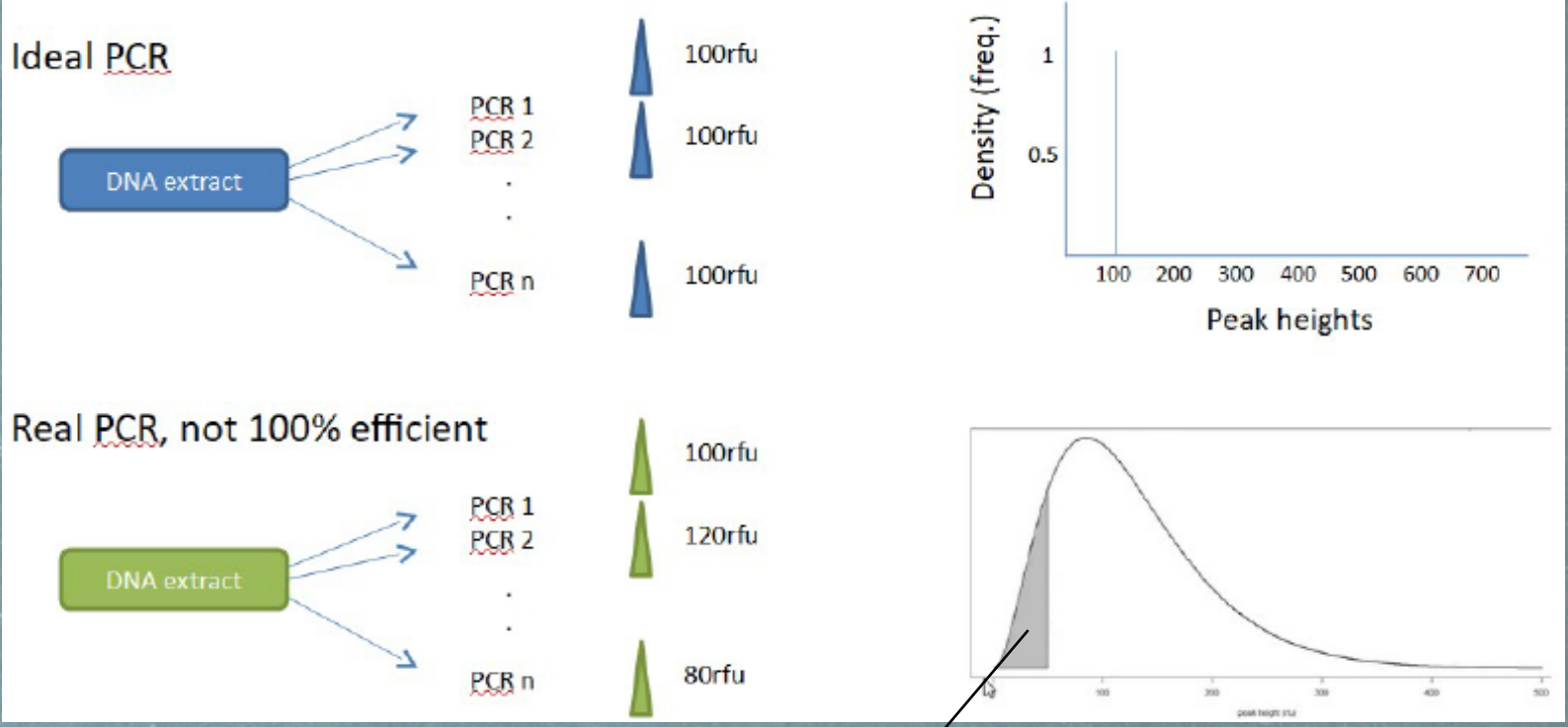


5%-95% percentile

Continuous model (peak height incorporated in the model)

- ✓ the same model can be extended to mixed DNA profiles
- ✓ Commercial and open source dedicated software available

- Assume that the distribution of peak areas arising in the PCR amplification of STRs follows a Gamma distribution



The shaded area corresponds to the drop out probability, that will change according to genotype and model parameters shaping the gamma distribution

Parameter	
K (number of contributors)	Set by the analyst after epg inspection
μ (expected peak height)	for a single heterozygote allele (without degradation)
σ (cv of peak heights)	coefficent of variation for a single heterozygote allele (without degradation)
m (mixture proportion)	
ξ (stutter proportion)	
β (degradation slope)	
C (drop-in probability)	

Software finds the combination of model parameters that maximizes the likelihood of observing the actual peak heght in the stain sample under each (Hp and Hd) hypotheses and thus provides LR of Hp/Hd

Analyst can choose to includine/exclude/combine stutter, degradation and drop-in in the model