Analysis of an Internal Validation Dataset for the New Core STR Loci

Sarah Riman PhD, Erica L. Romsos MFS, Peter M. Vallone PhD

Applied Genetics Group

National Institute of Standards and Technology
Gaithersburg, Maryland

Disclaimer

Points of view in this presentation are mine and do not necessarily represent the official position of the National Institute of Standards and Technology or the U.S. Department of Commerce.

NIST Disclaimer Certain commercial products and instruments are identified in order to specify experimental procedures as completely as possible. In no case does such an identification imply a recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that any of these products are necessarily the best available for the purpose.

Joli Bregu,[1] B.S.; Danielle Conklin,[1] M.S.; Elisse Coronado,[1] M.S.; Margaret Terrill,[1] M.S.F.S; Robin W. Cotton,[1] Ph.D.; and Catherine M. Grgicak,[1] Ph.D.

Analytical Thresholds and Sensitivity: Establishing RFU Thresholds for Forensic DNA Analysis*,†

Forensic Science International: Genetics 3 (2009) 104–111

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

The *low-template-DNA* (stochastic) threshold—Its determination relative to risk analysis for national DNA databases

Peter Gill [a,b,*], Roberto Puch-Solis [c], James Curran [d]

Forensic Science International: Genetics 19 (2015) 107–122

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Research paper

Probabilistic characterisation of baseline noise in STR profiles

Ullrich J. Mönich [a,*], Ken Duffy [d], Muriel M
Catherine Grgicak [b]

Journal of the Royal Statistical Society
SERIES C Applied Statistics

Appl. Statist. (2010)
59, *Part* 5, *pp.* 855–874

Evaluating the weight of evidence by using quantitative short tandem repeat data in DNA mixtures

Torben Tvedebrink and Poul Svante Eriksen
Aalborg University, Denmark
and Helle Smidt Mogensen and Niels Morling

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]

❖ In no case does this presentation imply a recommendation or endorsement of any of the methods.

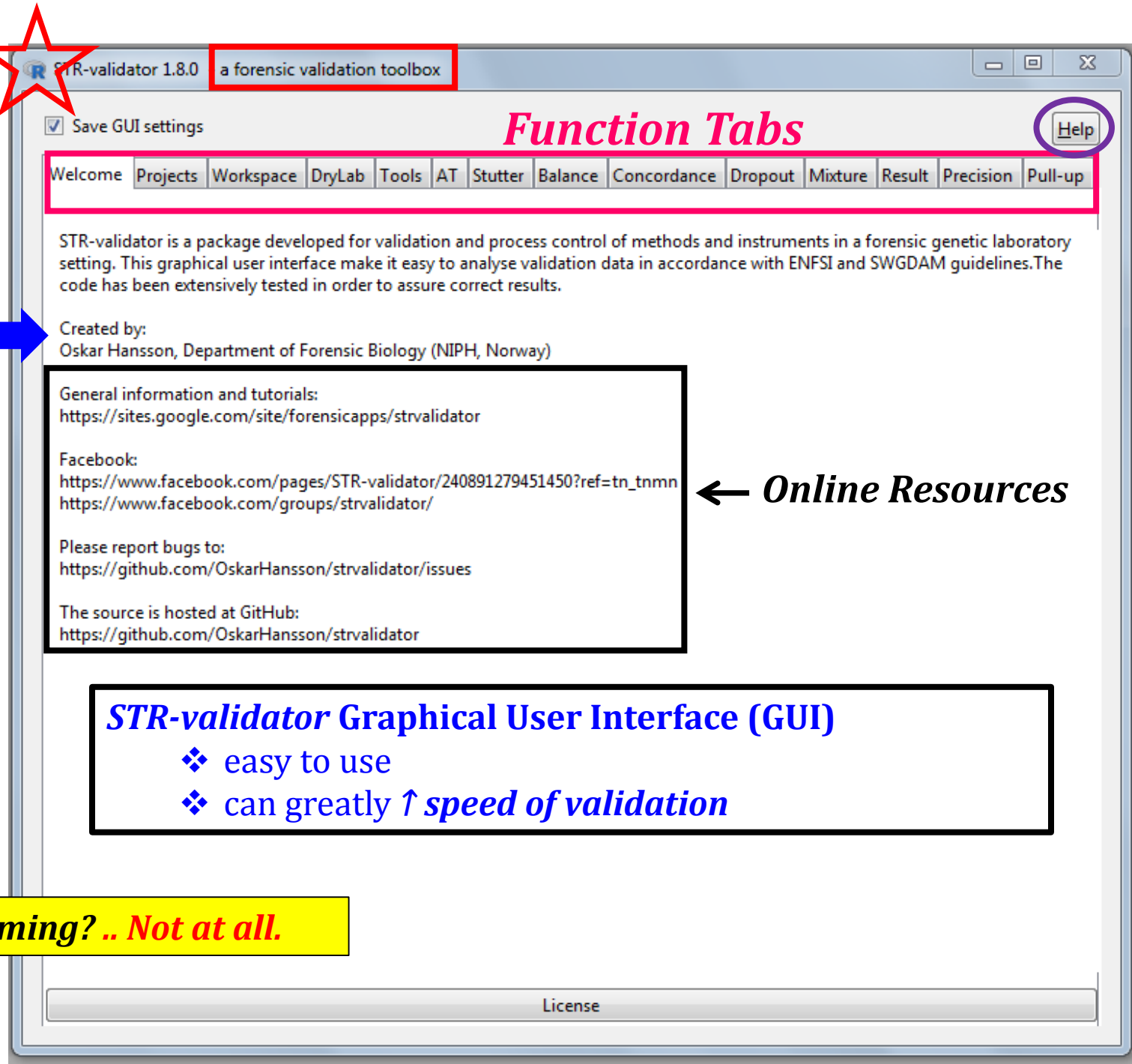
Presentation Outline

- Introduce the community to the availability of *STR-validator*

- Evaluate parameters in *STR-validator* such as:
  - Analytical thresholds
  - Stochastic thresholds
  - Stutter percentage
  - Peak height ratios
  - Base-pair sizing precision

- Report Analysis using
  - Plots
  - Histograms
  - Heatmaps
  - Boxplots
  - Tables

What is STR-Validator?



Creator of STR-Validator

← ***Online Resources***

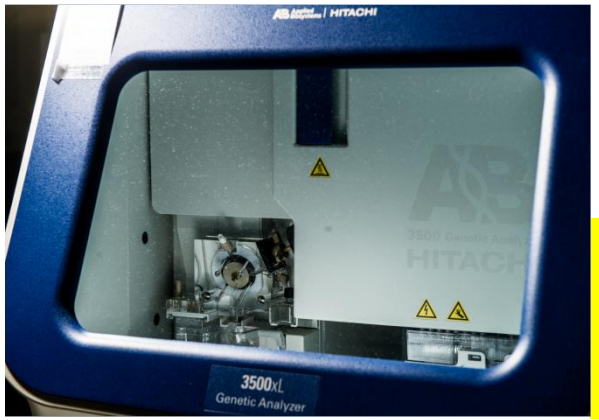
- A free and open source R-package
- Intended for:
  - Validating STR kits
  - Processing controls
  - Comparing methods and instrumentation

***STR-validator* Graphical User Interface (GUI)**

- easy to use
- can greatly ***↑ speed of validation***

- ***Should I be knowledgeable about programming? .. Not at all.***

3500xl Genetic Analyzer (24-Cappillary Array)

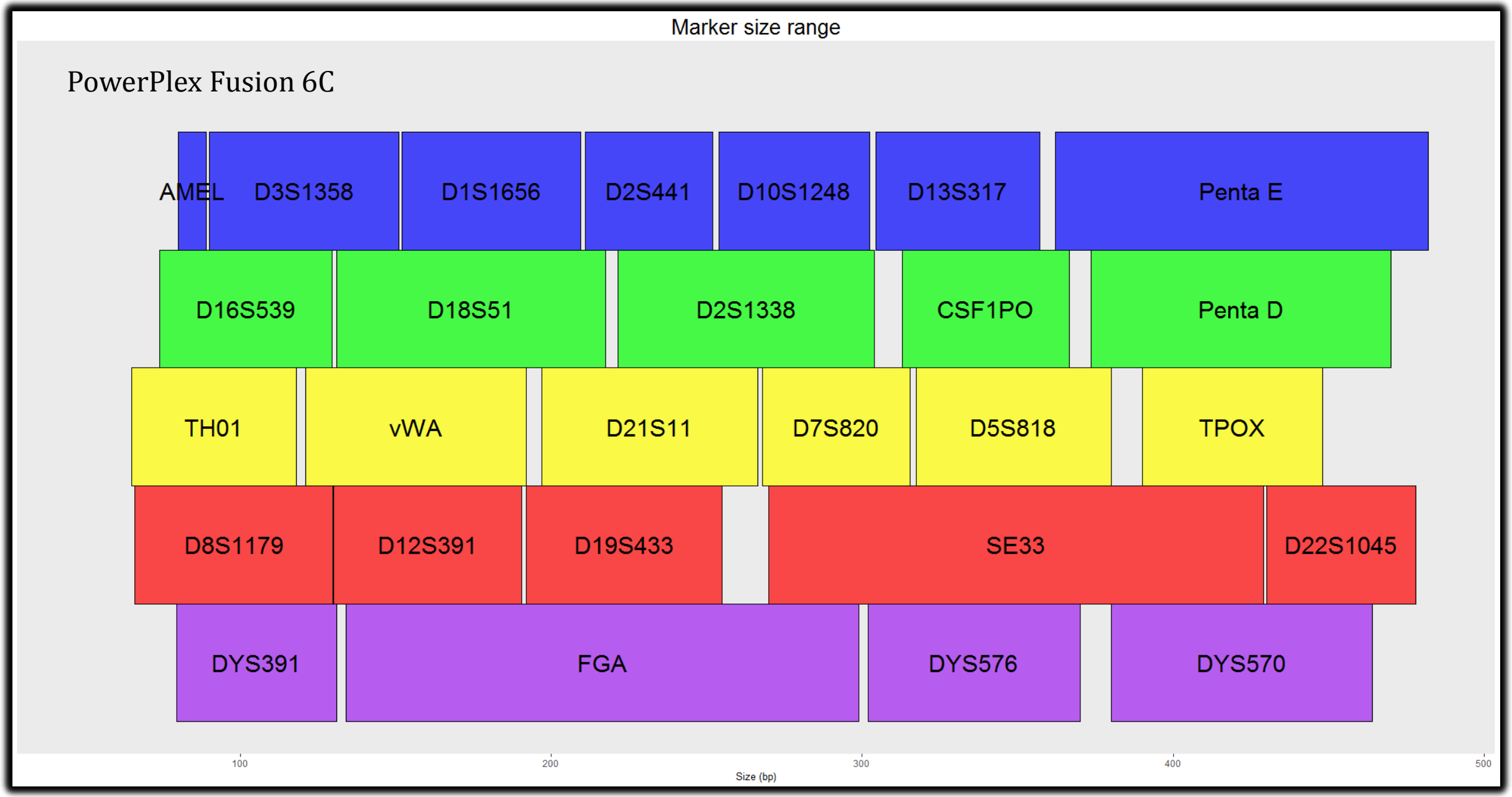


Conditions selected for internal validation testing:

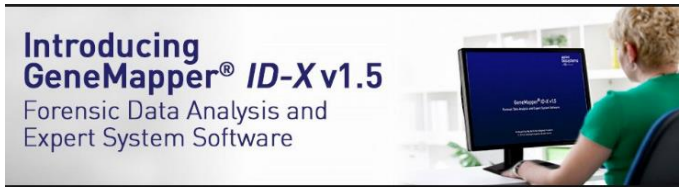
- **29 PCR cycles**
- **15 second injection at 1.2 kV**

Analysis of Internal Validation Study Experiments Using PowerPlex Fusion 6C and STR-Validator

Materials related to this presentation will be uploaded to *STRbase*
http://www.cstl.nist.gov/strbase/training.htm
When I return to NIST


Marker size range
PowerPlex Fusion 6C
AMEL
D3S1358
D1S1656
D2S441
D10S1248
D13S317
Penta E
D16S539
D18S51
D2S1338
CSF1PO
Penta D
TH01
vWA
D21S11
D7S820
D5S818
TPOX
D8S1179
D12S391
D19S433
SE33
D22S1045
DYS391
FGA
DYS576
DYS570
100
200
300
400
500
Size (bp)

How to Prepare the Data for Analysis?

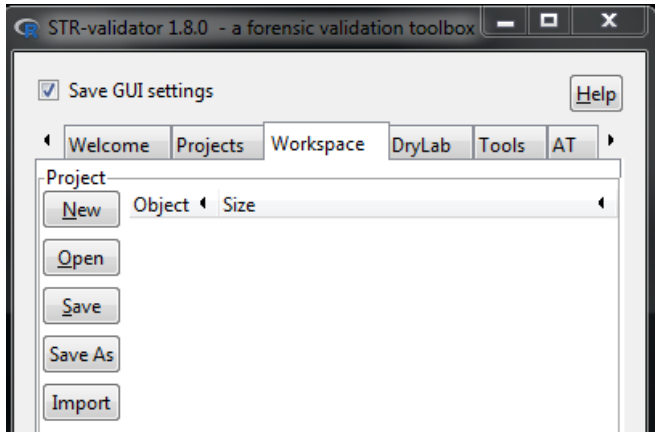


Export (.txt)

Sample Name	Marker	Dye	Allele 1	Allele 2	Allele 3	Allele 4	Allele 5	Allele 6	Allele 7	Allele 8	Allele 9	Allele 10	Allele 11	Allele 12	Allele 13	Allele 14	Allele 15	Allele
29_0.03ng	AMEL	B																
29_0.03ng	D3S1358	B	OL	OL	9	OL	OL	OL	17.2									
29_0.03ng	D1S1656	B	OL	OL	OL	OL	OL	OL	OL									
29_0.03ng	D2S441	B	OL	OL	11.3	OL	17											
29_0.03ng	D10S1248	B	OL	OL	16	OL	OL											
29_0.03ng	D13S317	B	OL	OL	8.1	OL	OL											
29_0.03ng	Penta E	B	OL	OL	OL	OL	OL	OL	OL	OL	OL	25	26					
29_0.03ng	D16S539	G	OL	OL	OL	11.3												
29_0.03ng	D18S51	G	OL	9.2	OL	15.2	OL	21.1	OL	OL	OL	OL						
29_0.03ng	D2S1338	G	OL	OL	23	OL	OL	OL										
29_0.03ng	CSF1PO	G																
29_0.03ng	Penta D	G	OL	7.4	OL	OL	17	OL	OL	OL								
29_0.03ng	TH01	Y	OL	7.3	8.3	OL	10	OL	OL	OL	OL							
29_0.03ng	vWA	Y	OL	17	OL	OL	OL											
29_0.03ng	D21S11	Y	OL	OL	30.3	OL	33.1	35										
29_0.03ng	D7S820	Y	6	8.1	9	10.1	12.1	13										
29_0.03ng	D5S818	Y	OL	OL	OL	OL	12	12.1	OL									

Semi-Wide type of table Format = *Unstacked Data*

Import (.txt)



Slim

Sample.Name	Marker	Dye	Allele	Size	Height
29_0.03ng	AMEL	B	NA	NA	NA
29_0.03ng	D3S1358	B	OL	92.06	8
29_0.03ng	D3S1358	B	OL	95.36	6
29_0.03ng	D3S1358	B	9	96.68	11
29_0.03ng	D3S1358	B	OL	101.95	9
29_0.03ng	D3S1358	B	OL	113.26	10
29_0.03ng	D3S1358	B	OL	130.44	6
29_0.03ng	D3S1358	B	17.2	133.66	4
29_0.03ng	D1S1656	B	OL	155.95	9
29_0.03ng	D1S1656	B	OL	158.42	8
29_0.03ng	D1S1656	B	OL	162.86	7
29_0.03ng	D1S1656	B	OL	164.83	8
29_0.03ng	D1S1656	B	OL	175.23	7
29_0.03ng	D1S1656	B	OL	179.7	7
29_0.03ng	D1S1656	B	OL	200.71	3
29_0.03ng	D2S441	B	OL	210.62	8
29_0.03ng	D2S441	B	OL	223.02	8
29_0.03ng	D2S441	B	11.3	227.1	8
29_0.03ng	D2S441	B	OL	240.29	8
29_0.03ng	D2S441	B	17	249.25	5
29_0.03ng	D10S1248	B	OL	257.96	6
29_0.03ng	D10S1248	B	OL	261.41	6
29_0.03ng	D10S1248	B	16	287.47	7
29_0.03ng	D10S1248	B	OL	290.03	7
29_0.03ng	D10S1248	B	OL	294.63	7
29_0.03ng	D13S317	B	OL	307.63	8
29_0.03ng	D13S317	B	OL	310.92	8

Semi-long narrow type of table Format = *Slim or Stacked data*

❖ *STR-validator* format

The Analytical Threshold

SWGDAM Autosomal STR Interpretation Guidelines

Analytical Threshold

- Peaks at and above this threshold can be reliably distinguished from background noise and are generally considered either artifacts or true alleles.

[1] SWGDAM Validation Guidelines for DNA Analysis Methods, (2016).

Experimental Design

- Three unique samples selected (29, 31, 32)
  - High levels of heterozygosity

- Run in triplicate
  - Three unique amplifications of the serial dilutions

- Dilution points
  - 2.0 ng, 1.0 ng, 0.5 ng, 0.25 ng, 0.125 ng, 0.0625 ng, 0.03 ng, 0.015 ng, and 0.008ng

	1	2	3	4	5	6	7	8	9
A	2.00	+	2.00	-	2.00	Ladder	2.00	Ladder	2.00
B	1.00	0.03	1.00	0.03	1.00	0.03	1.00	0.03	1.00
C	0.50	0.06	0.50	0.06	0.50	0.06	0.50	0.06	0.50
D	0.25	0.13	0.25	0.13	0.25	0.13	0.25	0.13	0.25
E	0.13	0.25	0.13	0.25	0.13	0.25	0.13	0.25	0.13
F	0.06	0.50	0.06	0.50	0.06	0.50	0.06	0.50	0.06
G	0.03	1.00	0.03	1.00	0.03	1.00	0.03	1.00	0.03
H	Ladder	2.00	-	2.00	+	2.00	-	2.00	+
	Sample 29			Sample 31			Sample 32		

	1	2	3	4	5	6	7	8	9
A	0.50	+	0.50	-	0.50	Ladder	0.50	Ladder	0.50
B	0.25	0.50	0.25	0.50	0.25	0.50	0.25	0.50	0.25
C	0.13	0.25	0.13	0.25	0.13	0.25	0.13	0.25	0.13
D	0.06	0.13	0.06	0.13	0.06	0.13	0.06	0.13	0.06
E	0.03	0.06	0.03	0.06	0.03	0.06	0.03	0.06	0.03
F	0.015	0.03	0.015	0.03	0.015	0.03	0.015	0.03	0.015
G	0.008	0.015	0.008	0.015	0.008	0.015	0.008	0.015	0.008
H	Ladder	0.008	-	0.008	+	0.008	-	0.008	+
	Sample 29			Sample 31			Sample 32		

- Analyzed in GeneMapper ID-X at 1 RFU in all dye channels

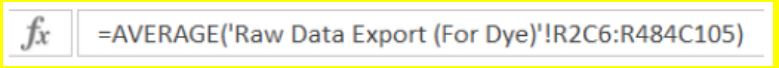
- Export the _SamplePlotSizingTable.txt from GeneMapper with at least the following information: "Dye/Sample Peak", "Sample.File.Name", "Marker", "Allele", "Height", and "Data.Point".

Analytical Threshold Most Commonly Determined by:

- Calculating baseline noise values from sensitivity dilution series
  - Threshold set at 1 RFU for all dye channels
  - Calls for all alleles and artifacts (stutter, n+4, pull-up, etc.) are removed

Sample Name	Amount	Marker	Dye	Height 1	Height 2	Height 3	Height 4	Height 5	Height 6	Height 7	Height 8	Height 9	Height 10	Height 11	Height 12	Height 13	Height 14	Height 15	Height 16	Height 17	Height 18	Height 19	Height 20	Height 21
A0.03	0.03	D3S1358	B	3	3	9	2	10	8	2	3	4												
A0.03	0.03	TH01	B	8	7	2	2	2	5	3	2													
A0.03	0.03	D21S11	B	3	2	7	8	6	2															
A0.03	0.03	D18S51	B	3	5	3	3	3	10	10	7	3	1	2										
A0.03	0.03	Penta E	B	9	3	5	3	3	2	3	3	2	2	2	1	2								
A0.03	0.03	D5S818	G	3	11	9	3	2	4															
A0.03	0.03	D13S317	G	12	7	2																		
A0.03	0.03	D7S820	G	3	3	16	4																	
A0.03	0.03	D16S539	G	8																				
A0.03	0.03	CSF1PO	G	6	5	2	4	2																
A0.03	0.03	Penta D	G	2	2	3	3	7	7	3	1	4	2											
A0.03	0.03	AMEL	Y																					
A0.03	0.03	vWA	Y	5	9	3	2	3																
A0.03	0.03	D8S1179	Y	3	3	2	21	2	3															
A0.03	0.03	TPOX	Y	1	3	4	3	15																
A0.03	0.03	FGA	Y	4	4	2	3	2	7	5	4	2	6	5	6	3	6	4	3	2	2	2	3	
A0.03	0.03	D3S1358	B	2	6	3																		
A0.03	0.03	TH01	B	3	2	3	2	1	2															
A0.03	0.03	D21S11	B	2	2	6	3	1	3															
A0.03	0.03	D18S51	B	2	2	2	2	1	2	2	4	3	3	1										
A0.03	0.03	Penta E	B	2	2	3	1	1	2															
A0.03	0.03	D5S818	G	2																				
A0.03	0.03	D13S317	G	2	2	2	4	3	4															
A0.03	0.03	D7S820	G	3	13	12	1																	
A0.03	0.03	D16S539	G	2	2	2																		
A0.03	0.03	CSF1PO	G	1	2	2	2																	
A0.03	0.03	Penta D	G	1	3	2	2	1	1	2	2													
A0.03	0.03	AMEL	Y	2																				
A0.03	0.03	vWA	Y	5	2	5	2	3	4	3	4	2												
A0.03	0.03	D8S1179	Y	6	4	3	9	2	3															
A0.03	0.03	TPOX	Y	3	2	3	5	12																
A0.03	0.03	FGA	Y	2	2	3	2	3	4	2	3	2	2	1	4	2	3							
A0.06	0.06	D3S1358	B	11	2	3	2																	
A0.06	0.06	TH01	B	3	2	3	2	2																
A0.06	0.06	D21S11	B	5	2	2	2																	
A0.06	0.06	D18S51	B	2	4	3	8	4	3	3	4	18	4	2										
A0.06	0.06	Penta E	B	1	3	3	3	2	4	2	2	1	2	2	2	2	3	3	1					

- Export 100 alleles per locus for all samples into an excel sheet
- Calculate the AVERAGE and STDEV of noise per dye channel

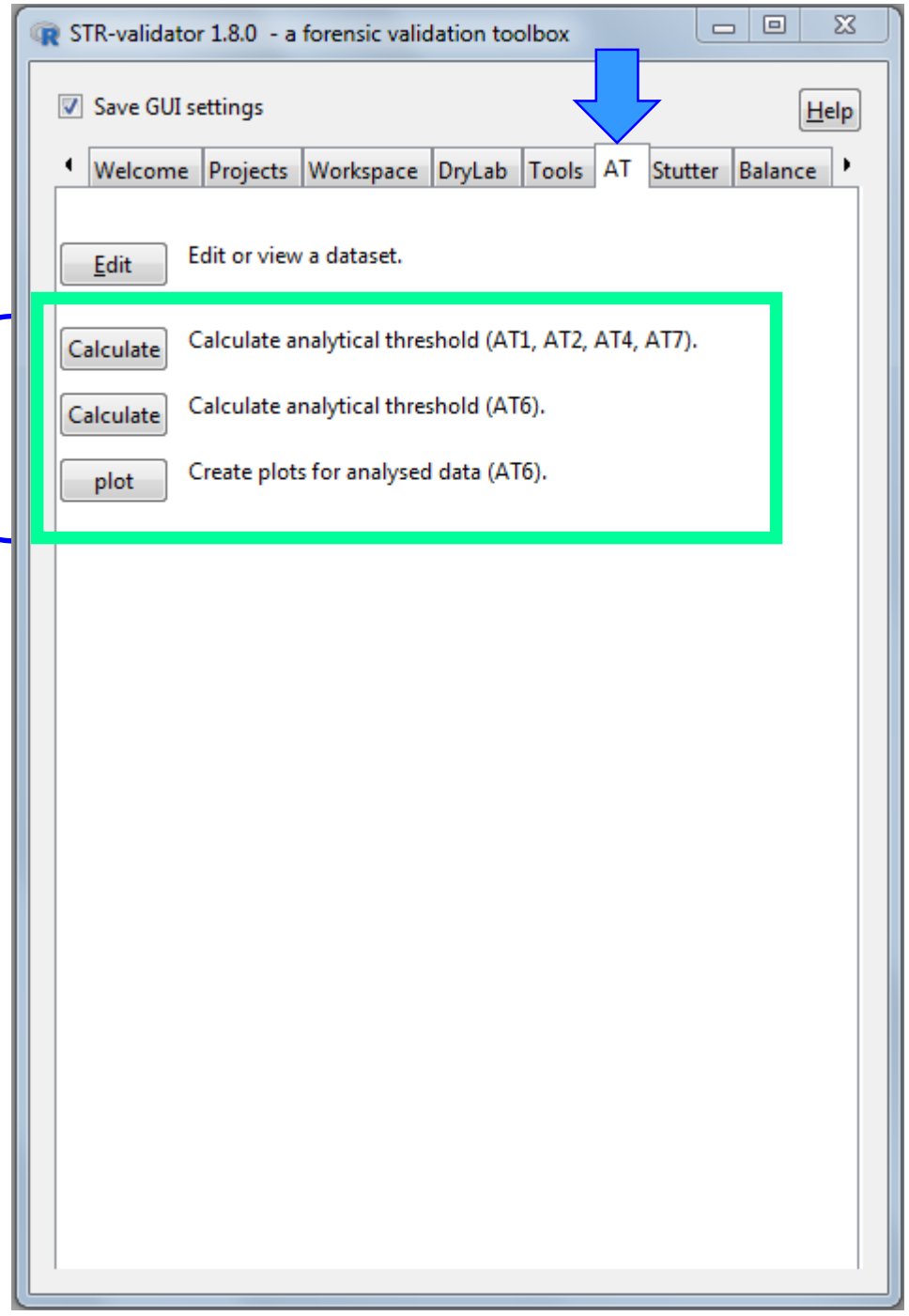


- AT = AVERAGE noise + 10 STDEV

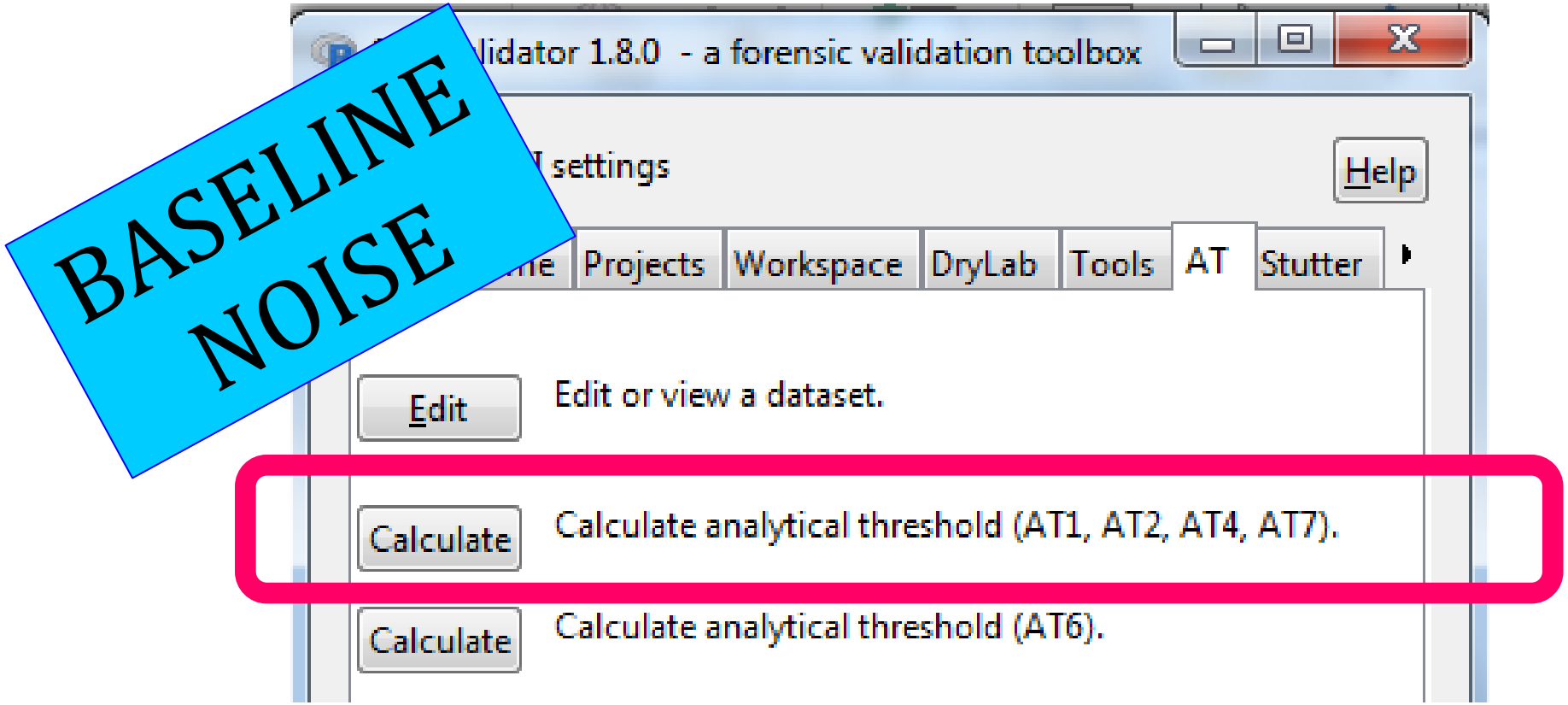
How does Analytical Threshold(s) get Calculated in STR-validator?

The Analytical Threshold in STR-validator

- Different methods for analytical threshold calculations
- Users can plot the analyzed data
- Methods 1, 2, 4, and 7 are calculated simultaneously (except for method 6)



The Analytical Threshold (Methods 1, 2, 4, 7)



JOURNAL OF FORENSIC SCIENCES

J Forensic Sci, January 2013, Vol. 58, No. 1
doi: 10.1111/1556-4029.12008
Available online at: onlinelibrary.wiley.com

TECHNICAL NOTE

CRIMINALISTICS

Joli Bregu,[1] B.S.; Danielle Conklin,[1] M.S.; Elisse Coronado,[1] M.S.; Margaret Terrill,[1] M.S.F.S.; Robin W. Cotton,[1] Ph.D.; and Catherine M. Grgicak,[1] Ph.D.

Analytical Thresholds and Sensitivity: Establishing RFU Thresholds for Forensic DNA Analysis*,†

ABSTRACT: Determining appropriate analytical thresholds (ATs) for forensic DNA analysis is critical to maximize allele detection. In this study, six methods to determine ATs for forensic DNA purposes were examined and compared. Four of the methods rely on analysis of the baseline noise of a number of negatives, while two utilize the relationship between relative fluorescence unit signal and DNA input in the polymerase chain reaction (PCR) derived from a dilution series ranging from 1 to 0.06 ng. Results showed that when a substantial mass of DNA (i.e., >1 ng) was amplified, the baseline noise increased, suggesting the application of an AT derived from negatives should only be applied to samples with low levels of DNA. Further, the number and intensity of these noise peaks increased with increasing injection times, indicating that to maximize the ability to detect alleles, ATs should be validated for each post-PCR procedure employed.

KEYWORDS: forensic science, minimum distinguishable signal, minimum discernible signal, forensic DNA analysis, analytical threshold, signal to noise

*AT1
*AT2
*AT4
*AT6

Forensic Science International: Genetics 19 (2015) 107–122

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Research paper

Probabilistic characterisation of baseline noise in STR profiles

Ullrich J. Mönich [a,*], Ken Duffy [d], Muriel Médard [a], Viveck Cadambe [c], Lauren E. Alfonse [b], Catherine Grgicak [b]

[a] Research Laboratory of Electronics, Massachusetts Institute of Technology, United States
[b] Biomedical Forensic Sciences, Boston University School of Medicine, United States
[c] Department of Electrical Engineering, Pennsylvania State University, United States
[d] Hamilton Institute, Maynooth University – National University of Ireland Maynooth, Ireland

ARTICLE INFO

Article history:
Received 27 February 2015
Received in revised form 22 May 2015
Accepted 2 July 2015
Available online 8 July 2015

Keywords:
Short tandem repeat
Noise
Peak height
Distribution
G-test
Stutter

ABSTRACT

There are three dominant contributing factors that distort short tandem repeat profile measurements, two of which, stutter and variations in the allelic peak heights, have been described extensively. Here we characterise the remaining component, baseline noise. A probabilistic characterisation of the non-allelic noise peaks is not only inherently useful for statistical inference but is also significant for establishing a detection threshold. We do this by analysing the data from 643 single person profiles for the Identifiler Plus kit and 303 for the PowerPlex 16 HS kit. This investigation reveals that although the dye colour is a significant factor, it is not sufficient to have a per-dye colour description of the noise. Furthermore, we show that at a per-locus basis, out of the Gaussian, log-normal, and gamma distribution classes, baseline noise is best described by log-normal distributions and provide a methodology for setting an analytical threshold based on that deduction. In the PowerPlex 16 HS kit, we observe evidence of significant stutter at two repeat units shorter than the allelic peak, which has implications for the definition of baseline noise and signal interpretation. In general, the DNA input mass has an influence on the noise distribution. Thus, it is advisable to study noise and, consequently, to infer quantities like the analytical threshold from data with a DNA input mass comparable to the DNA input mass of the samples to be analysed.

*AT7

AT_1

- *Numerical factor*
- *Chosen in accordance with the confidence level desired*

$$AT_{M1} = \overline{Y}_{bl} + ks_{bl}$$

The analytical threshold calculated using Method 1

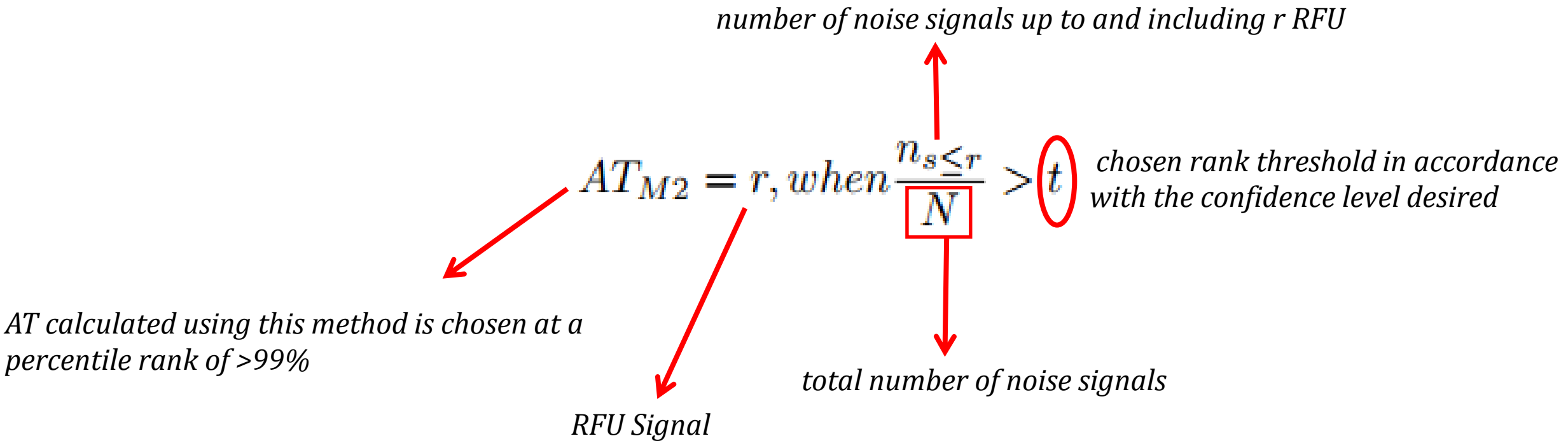
Average RFU signal

Standard deviation of the signal

REFERENCE

J. Bregu et al. Analytical Thresholds and Sensitivity: Establishing RFU Thresholds for Forensic DNA Analysis. JFS (2013) 1 pg 120-129.

AT_2



REFERENCE

J. Bregu et al. Analytical Thresholds and Sensitivity: Establishing RFU Thresholds for Forensic DNA Analysis. JFS (2013) 1 pg 120-129.

AT_4

$$AT_{M4} = \bar{Y}_{bl} + t_{\alpha,v}(1 + 1/n_{bl})^{1/2} s_{Y,bl}$$

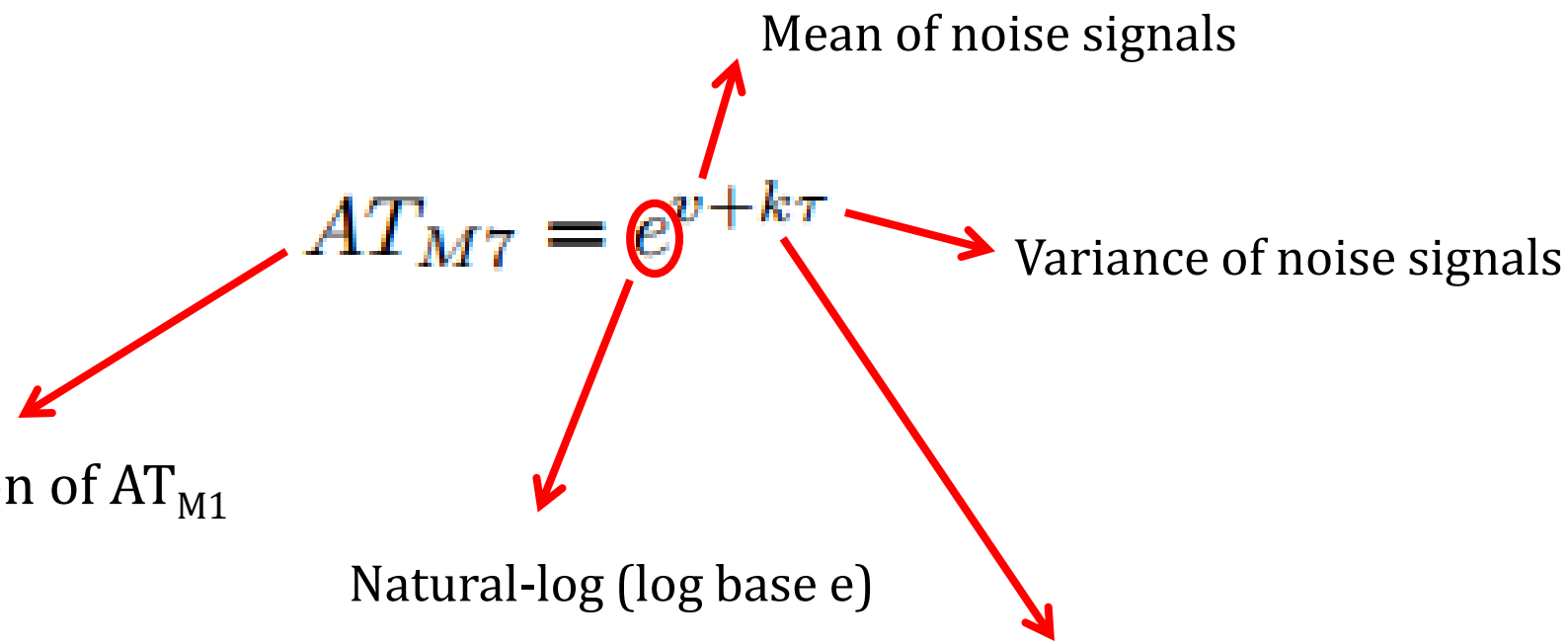
value obtained from the t-distribution for a given confidence interval "α"(one-sided)

number of samples

REFERENCE

J. Bregu et al. Analytical Thresholds and Sensitivity: Establishing RFU Thresholds for Forensic DNA Analysis. JFS (2013) 1 pg 120-129.

AT_7



REFERENCE

Ullrich J. Monich, Ken Duy, Muriel Medard, Viveck Cadambe, Lauren E. Alfonse, and Catherine Grgicak. Probabilistic characterisation of baseline noise in STR proles. Forensic Science International: Genetics

Mask Allele, Stutter, and Artifact Peaks

DATA MASKING

Options

- [x] Ignore case
- [ ] Add word boundaries
- [x] Mask high peaks — Mask all peaks above (RFU): 200
- [x] Mask sample alleles — Range (data points) around known alleles: 45
- [x] Mask sample alleles per dye channel
- [x] Mask ILS peaks — Range (data points) around known peak: 20

Confidence level 'k' (AT1, AT7): 3

Percentile rank threshold (AT2): 0.99

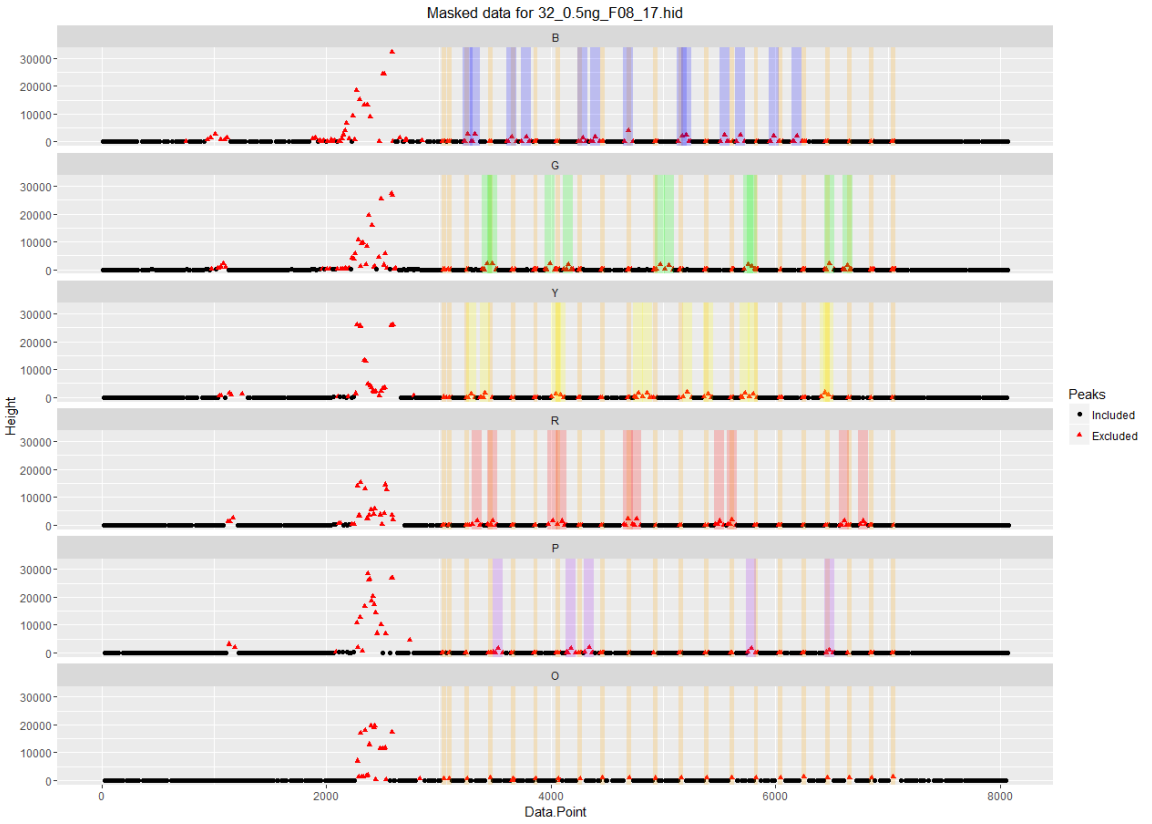
Upper confidence 'alpha' (AT4): 0.01

$$AT_{M1} = \bar{Y}_{bl} + ks_{bl}$$

$$AT_{M7} = e^{v+kr}$$

$$AT_{M2} = r, when \frac{n_{s \le r}}{N} > t$$

$$AT_{M4} = \bar{Y}_{bl} + t_{\alpha,v}(1 + 1/n_{bl})^{1/2} s_{Y,bl}$$



Analysis of AT1, AT2, AT4, and AT7 in *STR-validator*

DNA Dilution Series Data	AT1 (k=3)	AT1 (k=10)	AT2	AT4	AT7 (k=3)
Blue	62	185	108	67	53
Green	66	190	115	71	60
Yellow	51	154	93	55	34
Red	54	156	80	58	42
Purple	58	174	107	63	38

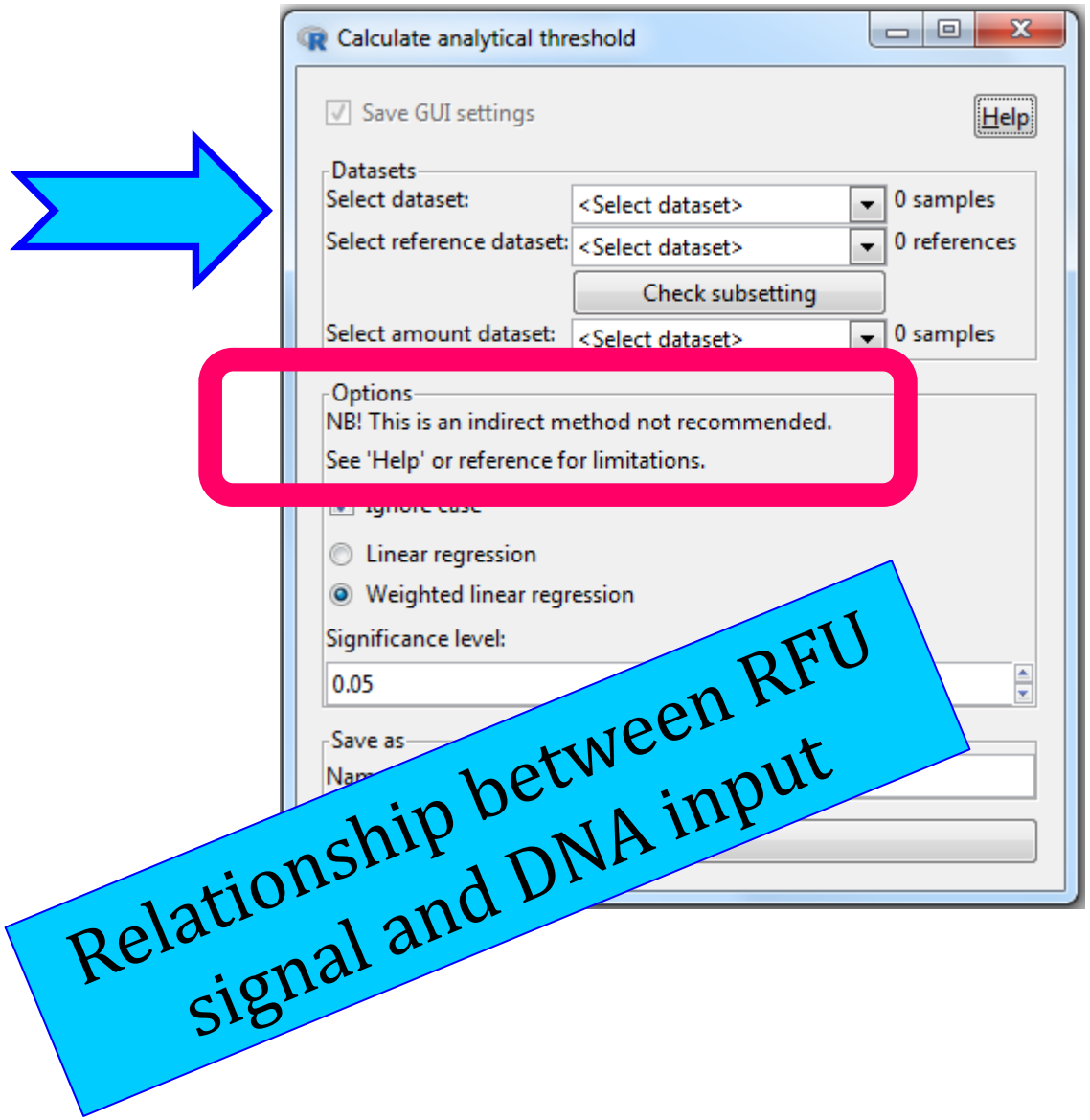
The Analytical Threshold

Method 6

1. Analyze samples in *GeneMapper at your AT*

2. Export -GenotypeTable.txt from GeneMapper with at least the following information: “Sample.Name”, “Marker”, “Allele” and “Height”.

Import, from one or several batches of sensitivity studies



Stochastic Threshold

SWGDAM Autosomal STR Interpretation Guidelines

Stochastic Threshold:

- Is the **RFU value above** which it is reasonable to assume that, at a given locus, **allelic dropout of a sister allele has not occurred**.
- Minimizes the chance of wrongly deciding a heterozygous locus as a homozygous one.

[1] SWGDAM Validation Guidelines for DNA Analysis Methods, (2016).

Experimental Design

- Three unique samples selected
  - High levels of heterozygosity

- Run in triplicate
  - Three unique amplifications of the serial dilutions

- Dilution points
  - 2.0 ng, 1.0 ng, 0.5 ng, 0.25 ng, 0.125 ng, 0.0625 ng, 0.03 ng, 0.015 ng, and 0.008ng

- Analyzed at your AT

- Export __GenotypeTable.txt from GeneMapper with at least the following information: "Sample.Name", "Marker", "Height", and "Allele".

	1	2	3	4	5	6	7	8	9
A	2.00	+	2.00	-	2.00	Ladder	2.00	Ladder	2.00
B	1.00	0.03	1.00	0.03	1.00	0.03	1.00	0.03	1.00
C	0.50	0.06	0.50	0.06	0.50	0.06	0.50	0.06	0.50
D	0.25	0.13	0.25	0.13	0.25	0.13	0.25	0.13	0.25
E	0.13	0.25	0.13	0.25	0.13	0.25	0.13	0.25	0.13
F	0.06	0.50	0.06	0.50	0.06	0.50	0.06	0.50	0.06
G	0.03	1.00	0.03	1.00	0.03	1.00	0.03	1.00	0.03
H	Ladder	2.00	-	2.00	+	2.00	-	2.00	+
	Sample 29			Sample 31			Sample 32		
Operator One									

	1	2	3	4	5	6	7	8	9
A	0.50	+	0.50	-	0.50	Ladder	0.50	Ladder	0.50
B	0.25	0.50	0.25	0.50	0.25	0.50	0.25	0.50	0.25
C	0.13	0.25	0.13	0.25	0.13	0.25	0.13	0.25	0.13
D	0.06	0.13	0.06	0.13	0.06	0.13	0.06	0.13	0.06
E	0.03	0.06	0.03	0.06	0.03	0.06	0.03	0.06	0.03
F	0.015	0.03	0.015	0.03	0.015	0.03	0.015	0.03	0.015
G	0.008	0.015	0.008	0.015	0.008	0.015	0.008	0.015	0.008
H	Ladder	0.008	-	0.008	+	0.008	-	0.008	+
	Sample 29			Sample 31			Sample 32		
Operator Two									

Stochastic Threshold Most Commonly Determined as:

The RFU value of the **highest** surviving false homozygous peak per dye channel

	Blue	Green	Yellow	Red	Purple
min	62	68	51	55	58
max	402	219	223	319	154
average	125	118	98	115	90
stdev	62	42	45	61	33

Sample.Name	Marker	Allele	Height	Dropout	Rfu	Dye
B_30pg_39	D3S1358	15	402	1	402	B
B_63pg_40	AMEL	X	297	1	297	B
B_63pg_33	D10S1248	16	265	1	265	B
B_0.03ng_1	Penta E	10	239	1	239	B
A_15pg_23	D3S1358	14	231	1	231	B

Sample.Name	Marker	Allele	Height	Dropout	Rfu	Dye
B_63pg_40	D16S539	10	219	1	219	G
A_63pg_32	D18S51	12	199	1	199	G
A_15pg_23	D16S539	11	195	1	195	G
A_0.03ng_3	D18S51	14	188	1	188	G
A_0.03ng_1	CSF1PO	12	183	1	183	G

Sample.Name	Marker	Allele	Height	Dropout	Rfu	Dye
B_63pg_33	vWA	14	223	1	223	Y
C_0.06ng_5	D7S820	8	219	1	219	Y
B_15pg_24	D21S11	30	211	1	211	Y
A_8pg_22	D5S818	10	206	1	206	Y
B_0.03ng_2	TPOX	6	200	1	200	Y

Sample.Name	Marker	Allele	Height	Dropout	Rfu	Dye
C_0.06ng_4	D8S1179	11	319	1	319	R
C_30pg_24	SE33	14	295	1	295	R
A_30pg_31	D12S391	21	286	1	286	R
A_30pg_38	SE33	28.2	258	1	258	R
C_63pg_39	D19S433	13	256	1	256	R

Sample.Name	Marker	Allele	Height	Dropout	Rfu	Dye
B_30pg_32	FGA	23	154	1	154	P
B_30pg_25	FGA	18	144	1	144	P
C_15pg_37	FGA	25	129	1	129	P
A_30pg_38	FGA	22	103	1	103	P
B_0.03ng_2	FGA	23	97	1	97	P

How does Stochastic Threshold(s) get Calculated in STR-validator?

Stochastic Threshold in STR-Validator

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]

Forensic Science International: Genetics 3 (2009) 104–111

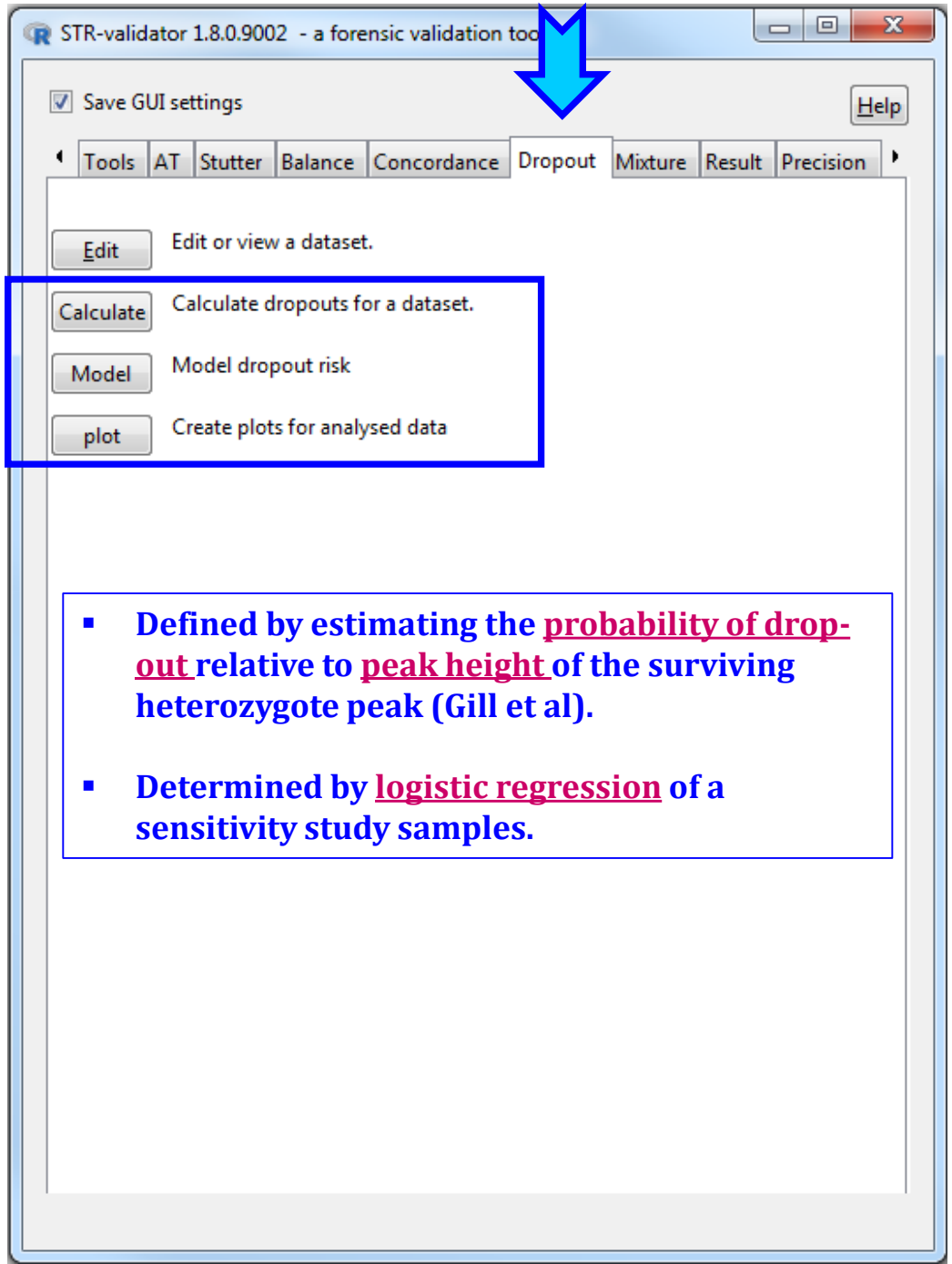
Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

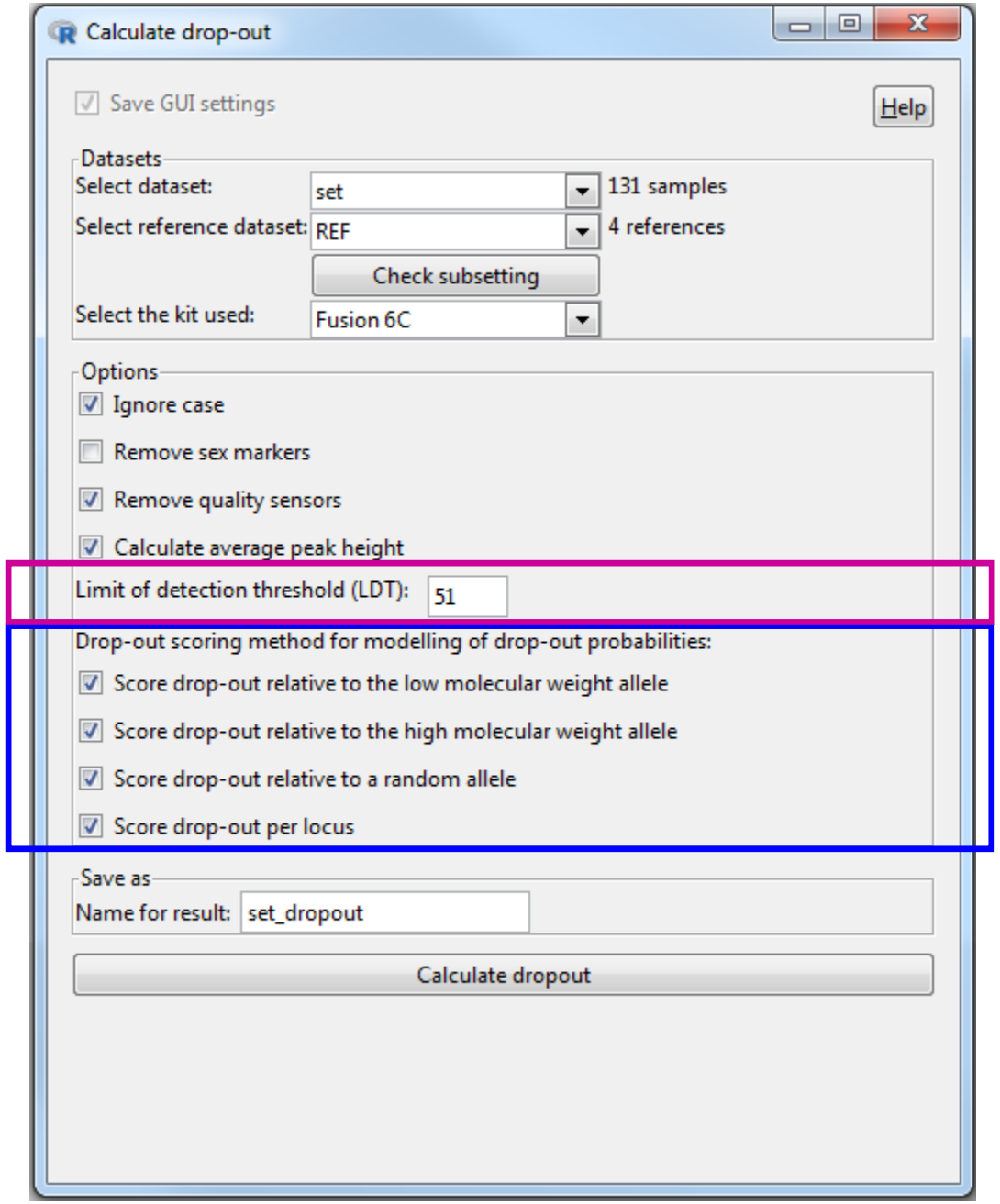
The *low-template-DNA* (stochastic) threshold—Its determination relative to risk analysis for national DNA databases

Peter Gill [a,b,*], Roberto Puch-Solis [c], James Curran [d]



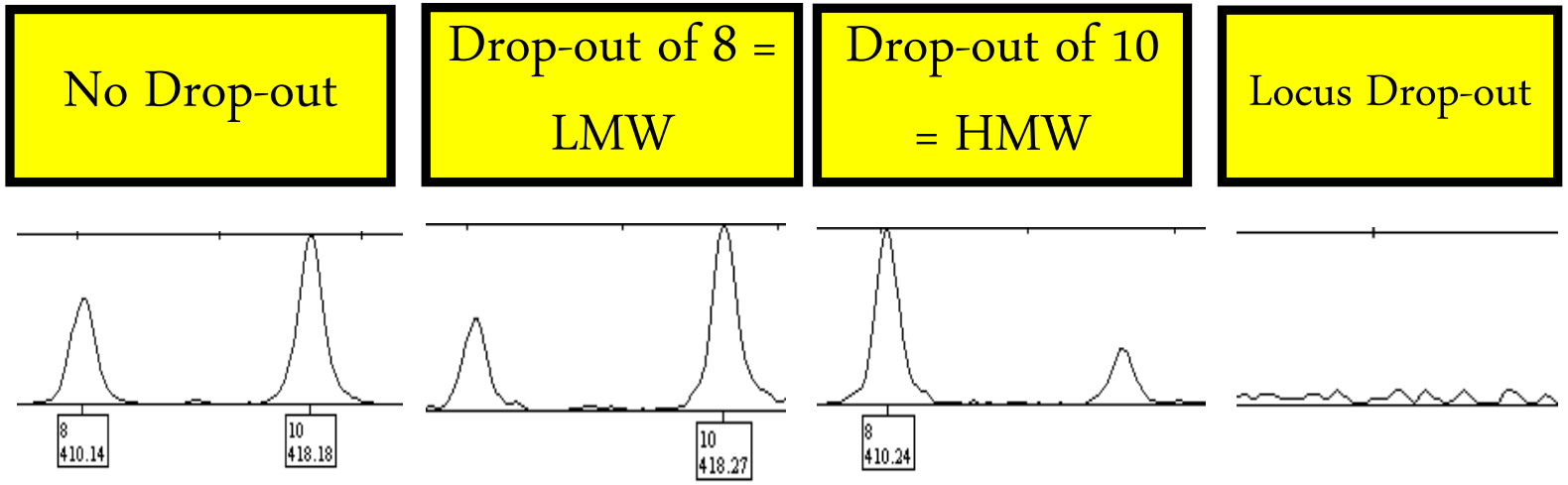
- **Defined by estimating the probability of drop-out relative to peak height of the surviving heterozygote peak (Gill et al).**
- **Determined by logistic regression of a sensitivity study samples.**

Four Methods to Score Drop-out Alleles



- Drop-Out= Allele with a peak height lower than the limit of detection threshold (LDT).
- LDT is not the AT. The lowest peak height in the dataset is automatically suggested in the 'Limit of Detection Threshold' field.

Four Methods to Score Drop-Out Alleles



Relative to Random Allele	***No-dropout = 0***	***No Dropout/Drop-out*** = 0/1	***No Dropout/Drop-out*** = 0/1	***N/A***
Relative to LMW	***No-dropout = 0***	***No Dropout = 0***	***Drop-out = 1***	***N/A***
Relative to HMW	***No-dropout = 0***	***Drop-out = 1***	***No Drop-out = 0***	***N/A***
Relative to Locus	***No-dropout = 0***	***Drop-out = 1***	***Drop-out = 1***	***Locus Drop-out*** = 2

Drop out Scoring Results

Scoring Methods discussed in the previous slide

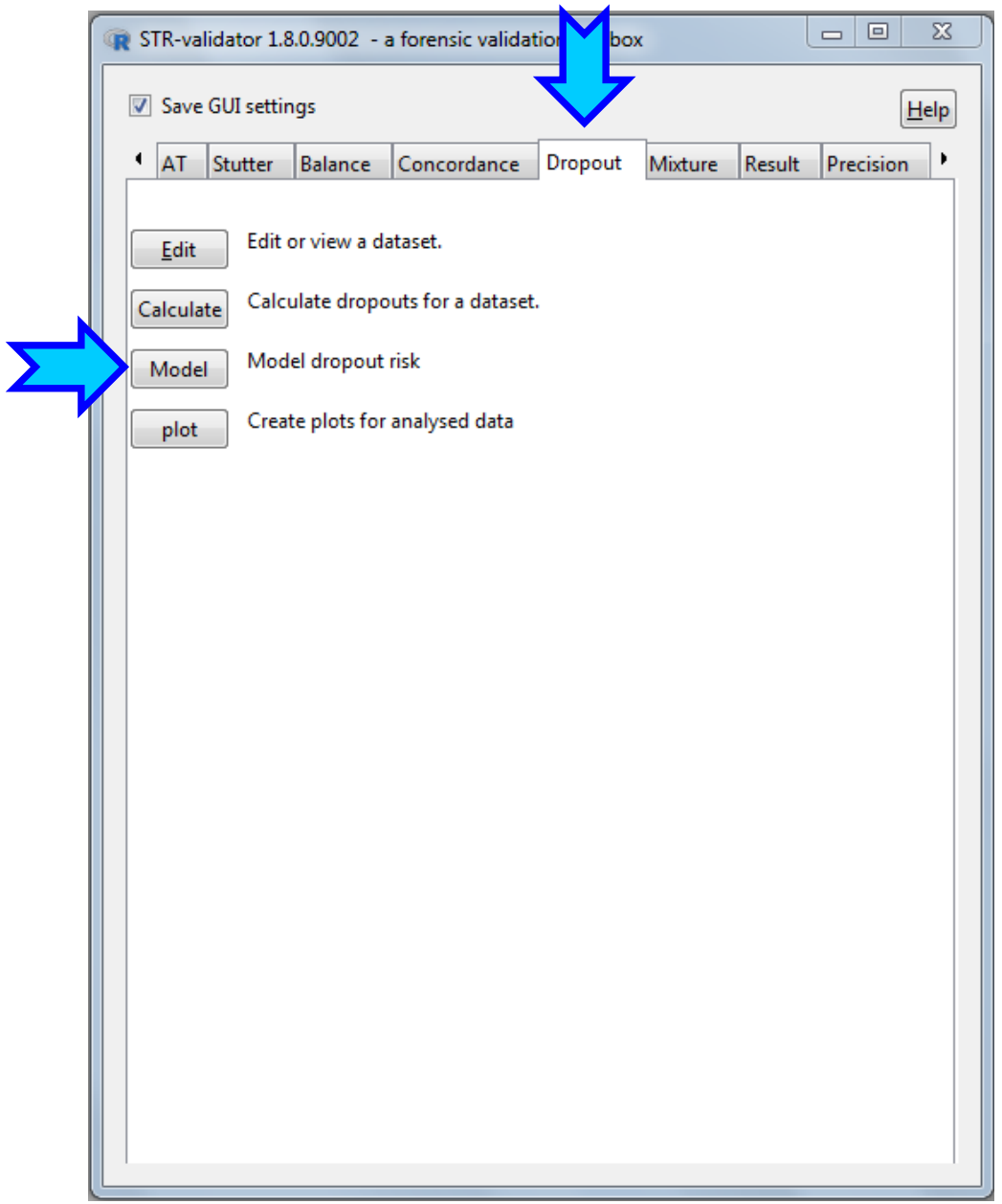
Sample.Name	Marker	Allele	Height	Dropout	Rfu	Heterozygous	MethodX	Method1	Method2	MethodL	MethodL.Ph	TPH	H	Peaks	Expected	Proportion
B_30pg_39	D3S1358	15	402	1	402	1	0	1	0	1	402	4662	116.55	37	50	0.74
C_0.06ng_4	D8S1179	11	319	1	319	1	1	1	0	1	319	12941	248.8654	48	50	0.96
B_63pg_40	AMEL	X	297	1	297	1	0	1	0	1	297	8824	173.0196	47	50	0.94
C_30pg_24	SE33	14	295	1	295	1	1	1	0	1	295	6720	146.087	42	50	0.84
A_30pg_31	D12S391	21	286	1	286	1	0	0	1	1	286	7016	152.5217	44	51	0.8627451
B_63pg_33	D10S1248	16	265	1	265	1	0	0	1	1	265	9742	187.3462	48	50	0.96
A_30pg_38	SE33	28.2	258	1	258	1	0	1	0	1	258	5648	125.5111	42	51	0.8235294
C_63pg_39	D19S433	13	256	1	256	1	0	1	0	1	256	10070	193.6538	48	50	0.96
B_0.03ng_1	Penta E	10	239	1	239	1	0	1	0	1	239	5058	126.45	38	50	0.76
A_15pg_23	D3S1358	14	231	1	231	1	1	1	0	1	231	3339	104.3438	30	51	0.5882353

TPH = Total Peak Height for each sample

H = sum(peakheights)/(n[het] + 2n[hom]

Profile Proportion = $\frac{\text{\# of Peaks}}{\text{\# of Expected Peaks}}$

Model Drop-out



Probability of Drop-out Modelled by Logistic Regression

Scored Drop-out Events
+
Peak Heights of the Surviving Allele

modelled →

by logistic regression
P(dropout|H) = B0 + B1*H

Stochastic Threshold

This data is used for logistic regression

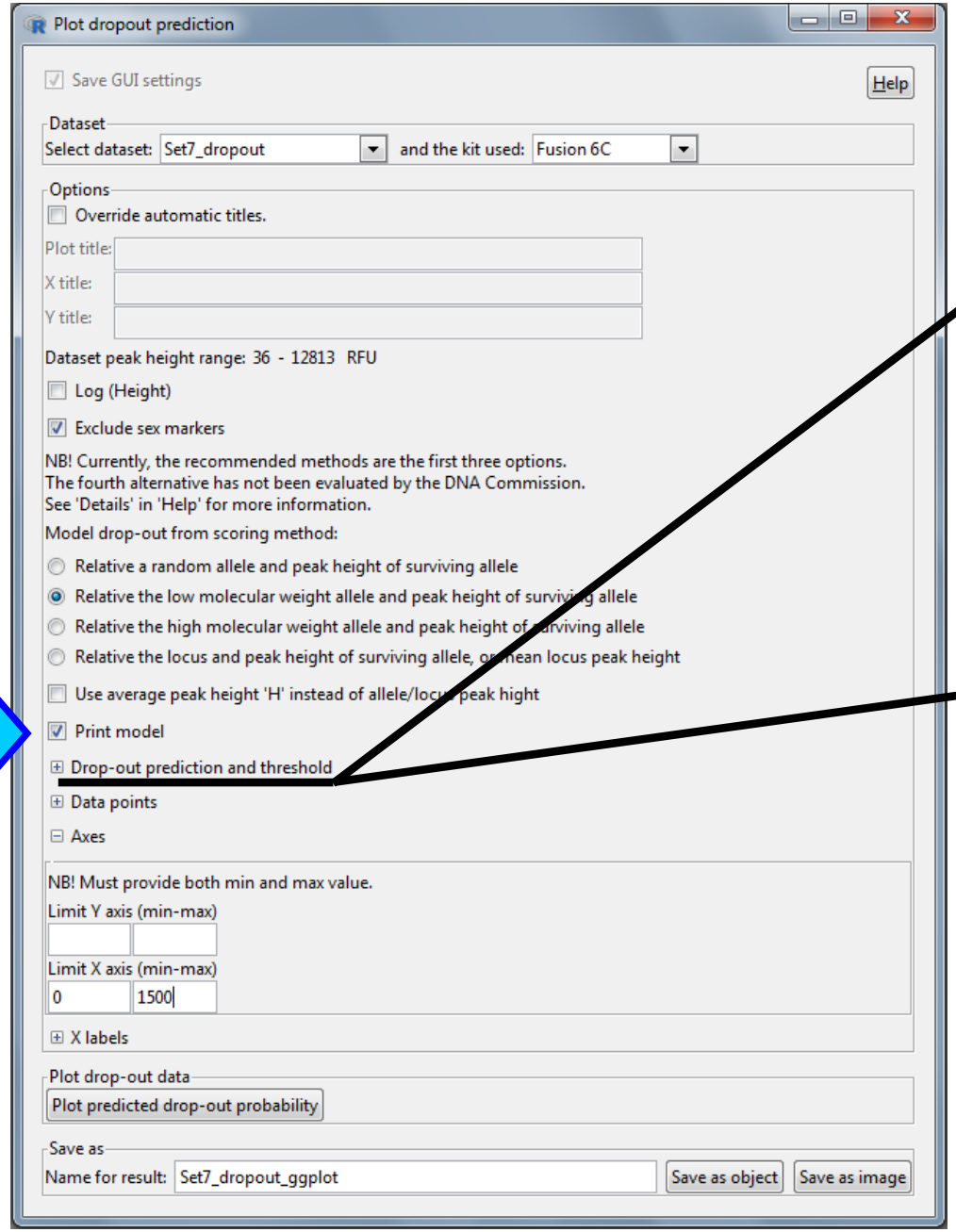
Logistic Model for the Estimation of Pr(D):

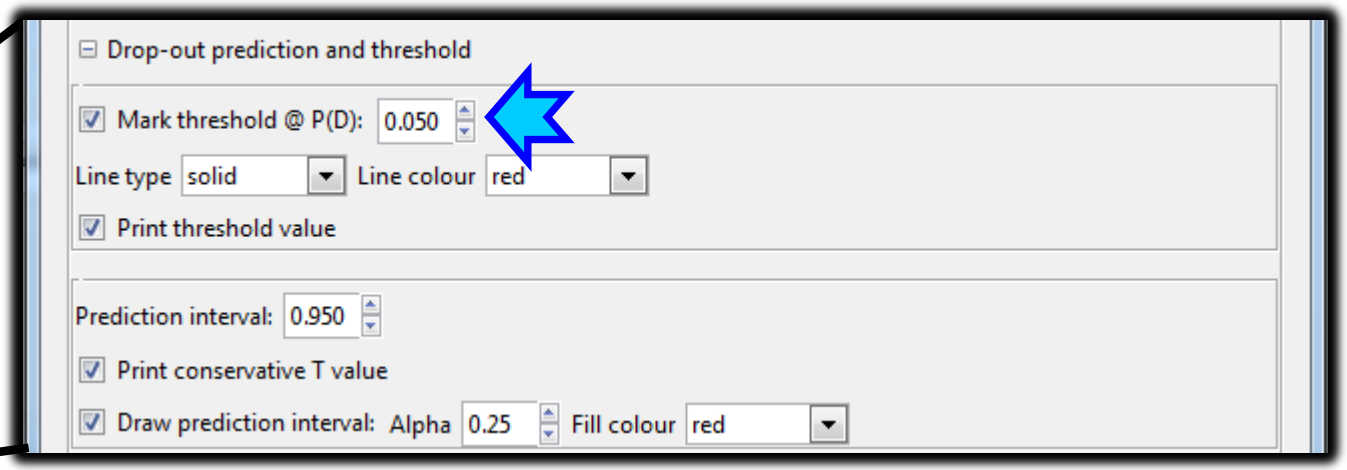
- P(dropout|H) = B0 + B1*H

- H = is the height of the partner allele in RFU
- P(dropout|H) = is the drop-out indicator of the allele of interest (either 0 or 1)
- B0 and B1 are estimated via logistic regression

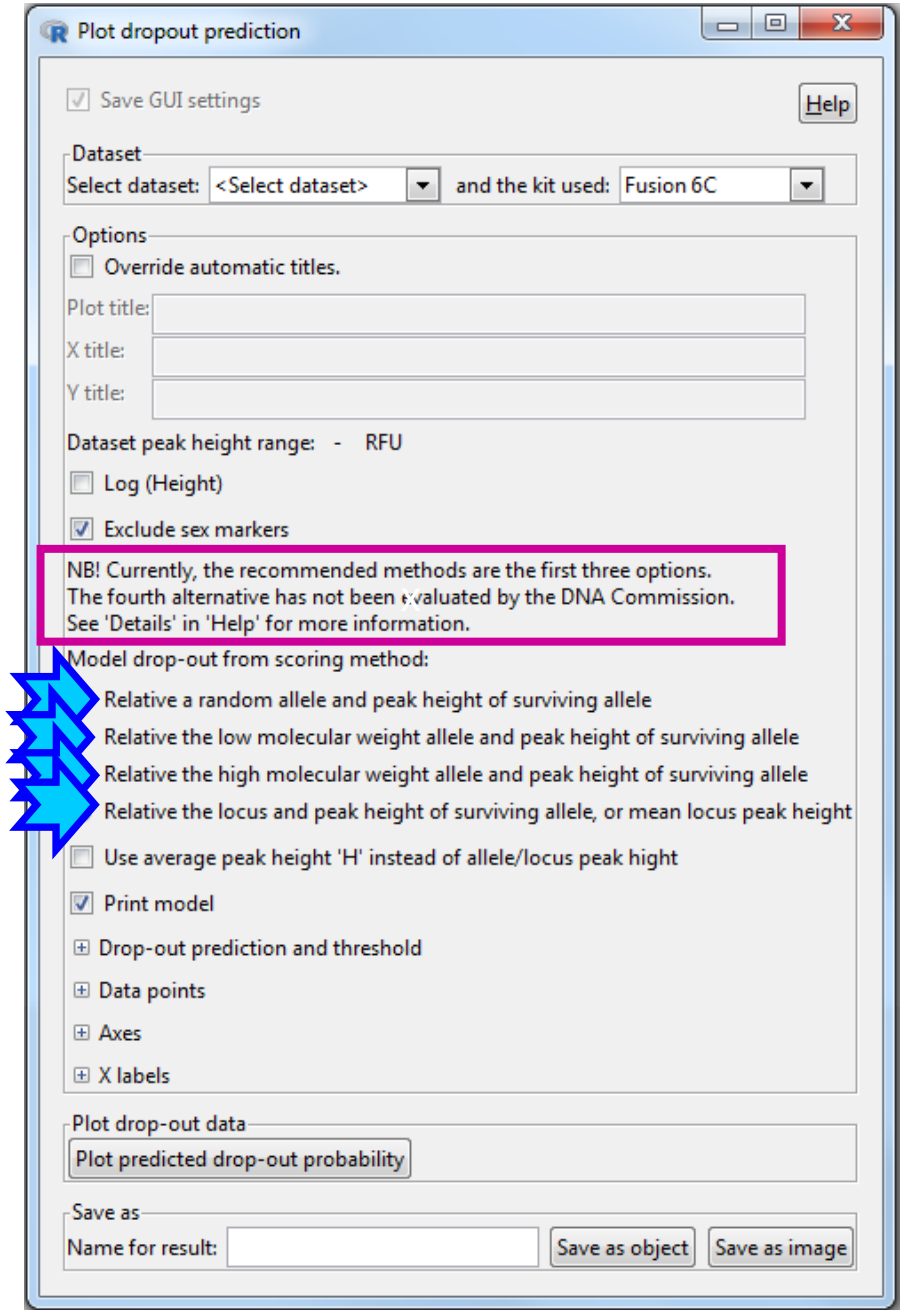
Sample.Name	Marker	Allele	Height	MethodX
B_30pg_39	D3S1358	15	402	0
C_0.06ng_4	D8S1179	11	319	1
B_63pg_40	AMEL	X	297	0
C_30pg_24	SE33	14	295	1
A_30pg_31	D12S391	21	286	0
B_63pg_33	D10S1248	16	265	0

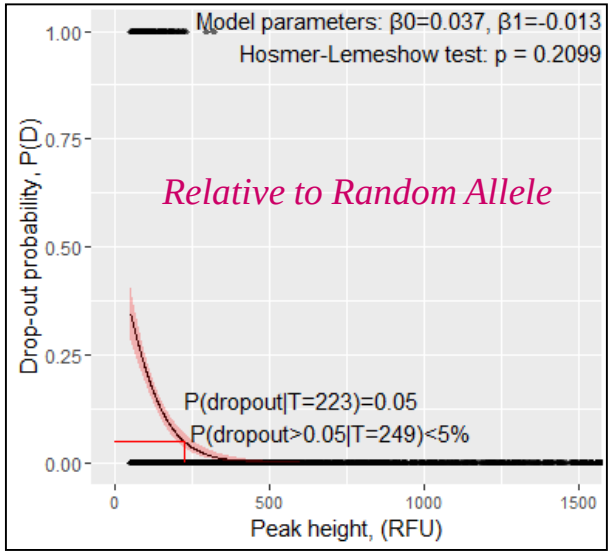
Estimated Thresholds for 5% Risk of Drop-out

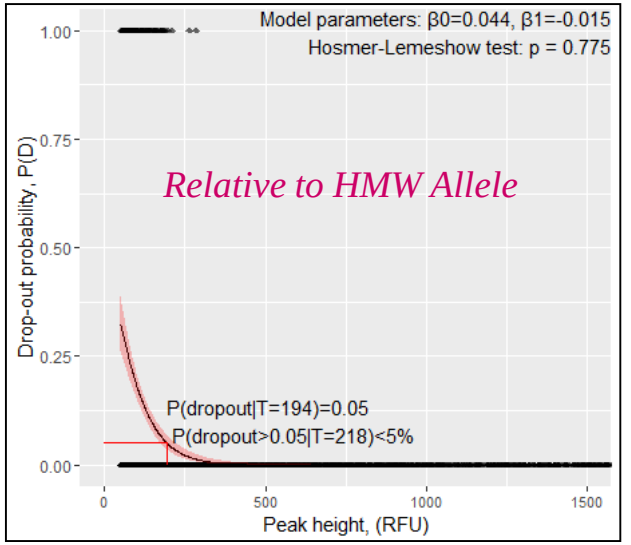


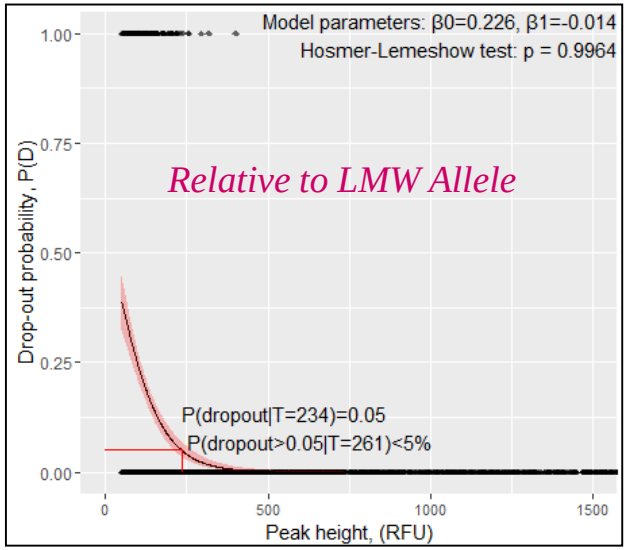


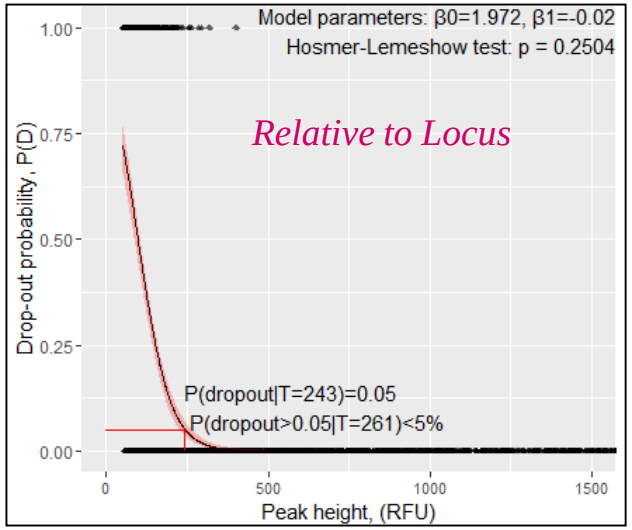
Estimated Thresholds for 5% Risk of Drop-out As A Function of Present-Allele Height



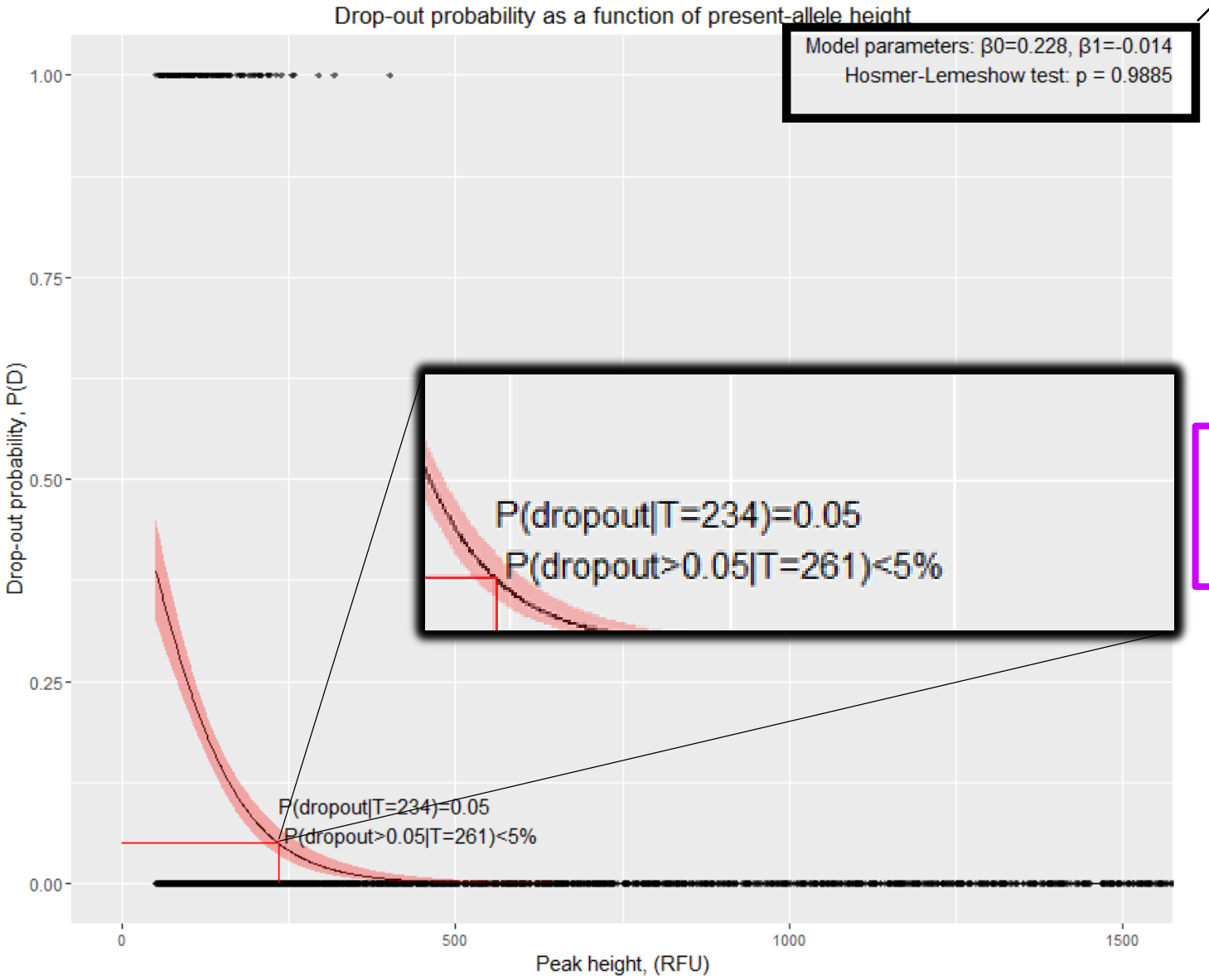








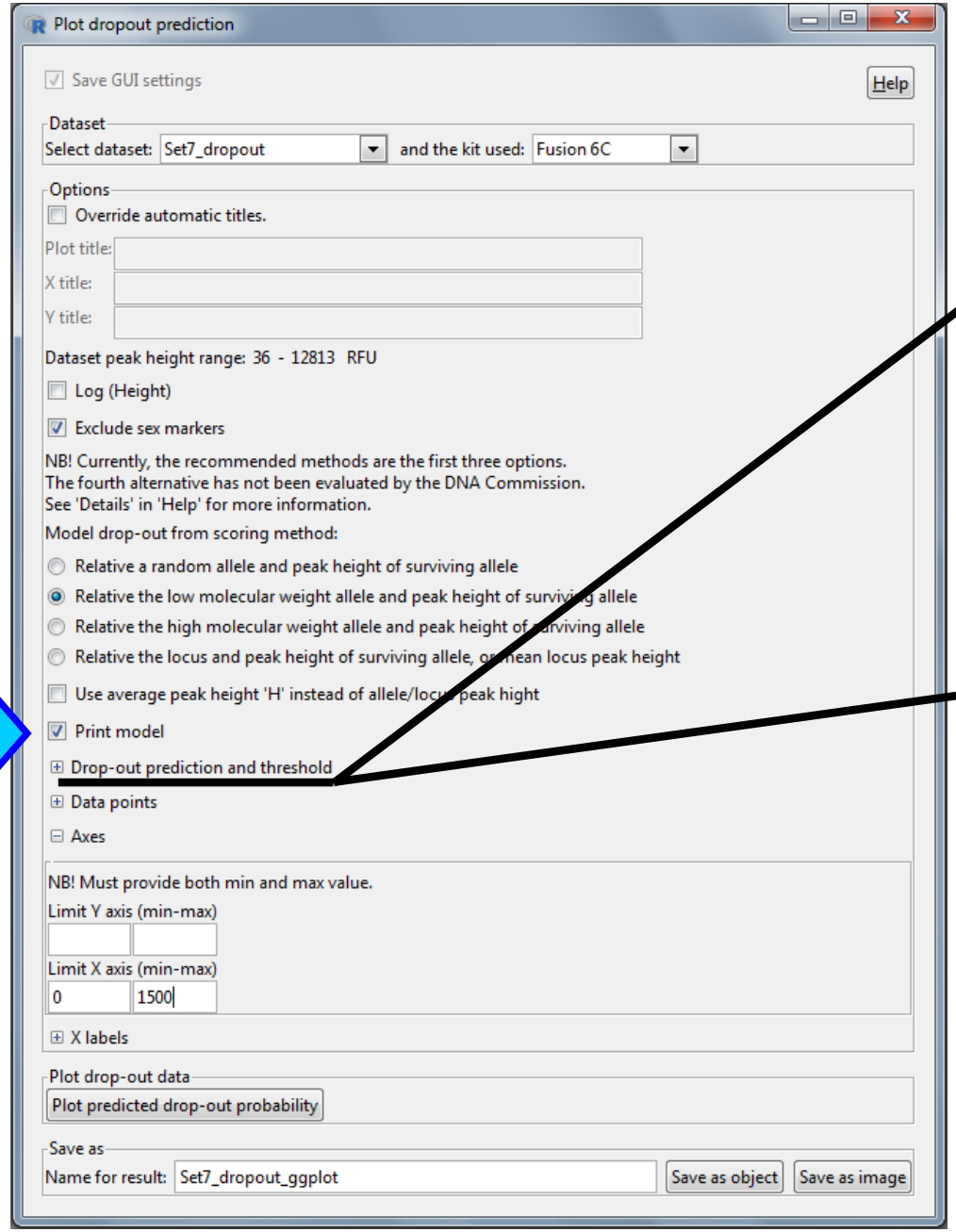
Plot showing logistic regression for the drop-out data

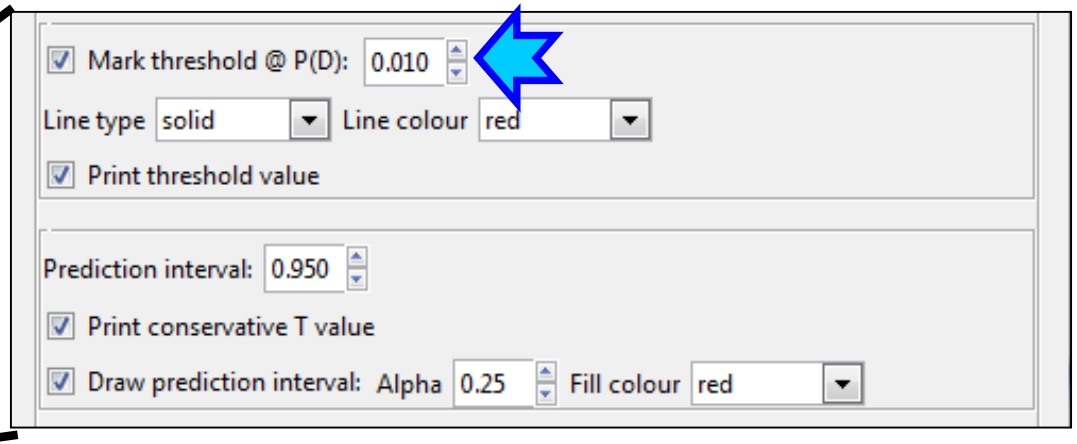


- Hosmer-Lemeshow evaluates the goodness of fit for the logistic regression.
- This indicates a good fit to observed data.

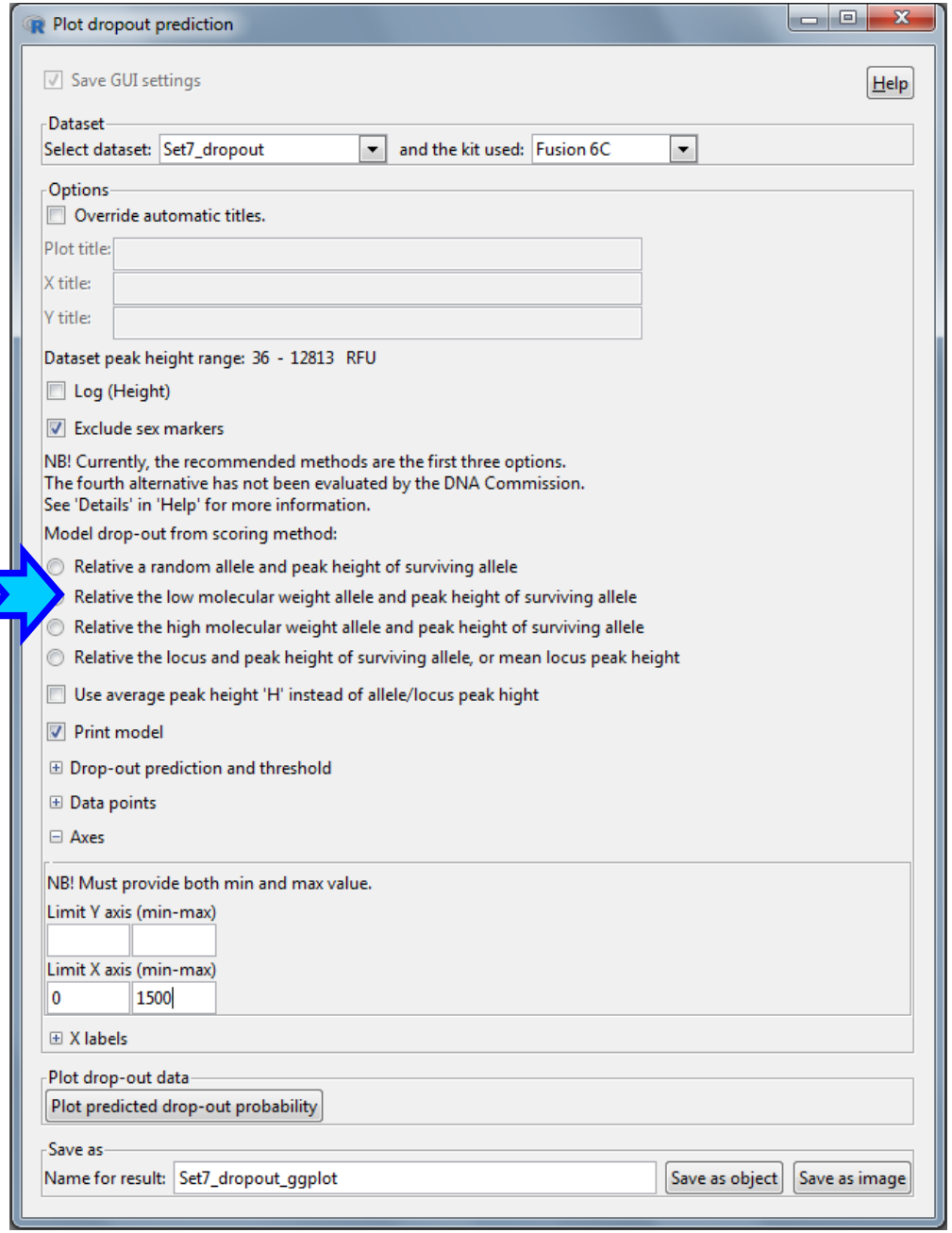
- Probability of drop-out is 5% at 234 RFU.
- A conservative threshold is 261 RFU with < 5% risk for getting drop-out probability > 5 %.

Estimated Thresholds for 1% Risk of Drop-out

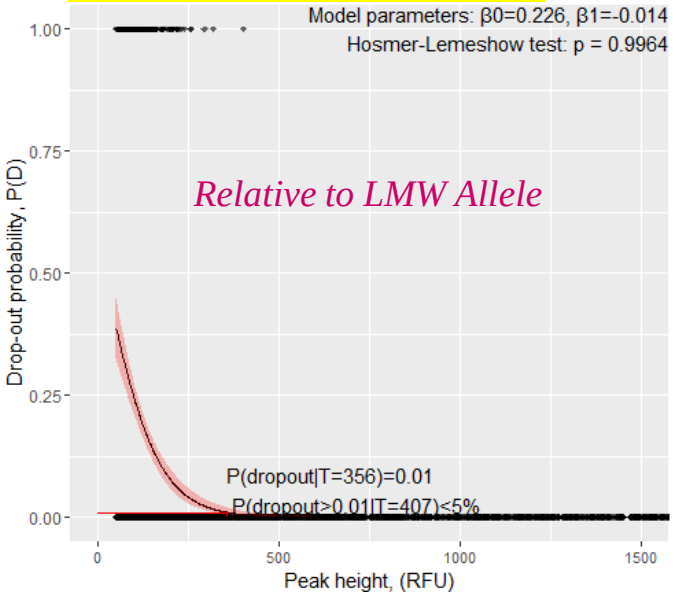




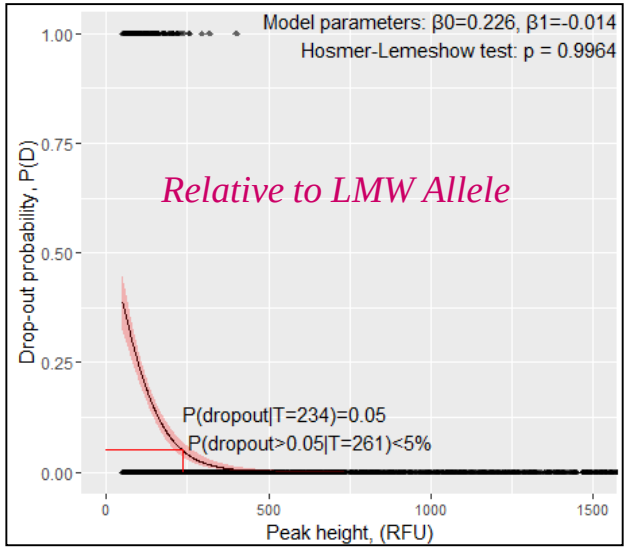
Estimated Thresholds for 1% Risk of Drop-out As A Function of Present-Allele Height



1 % Risk of Drop-Out



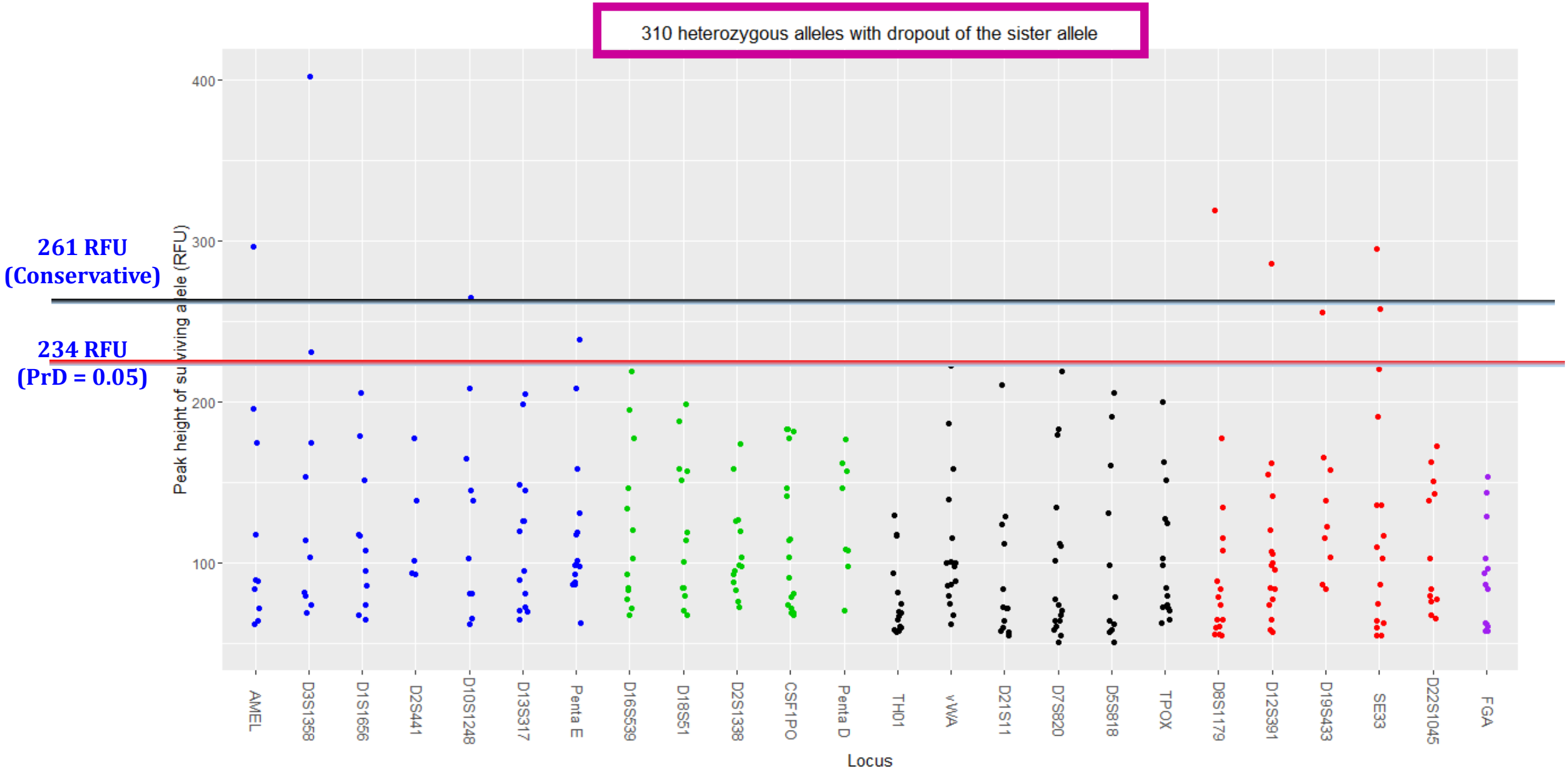
5 % Risk of Drop-Out



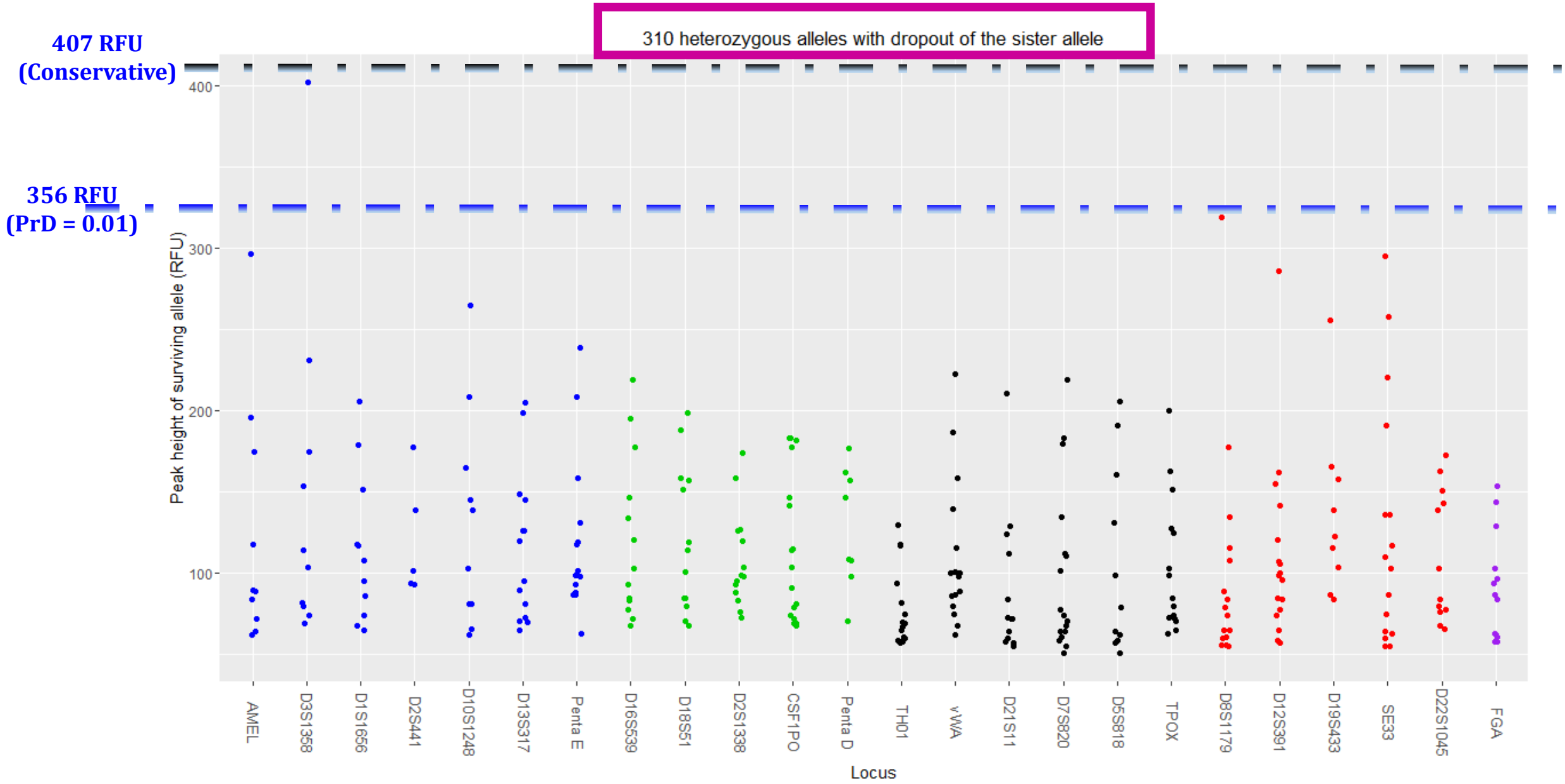
- Probability of drop-out is 1% at 356 RFU.
- A conservative threshold is 407 RFU with < 5% risk for getting drop-out probability > 1%.

How do these Thresholds Compare to the Observed Data?

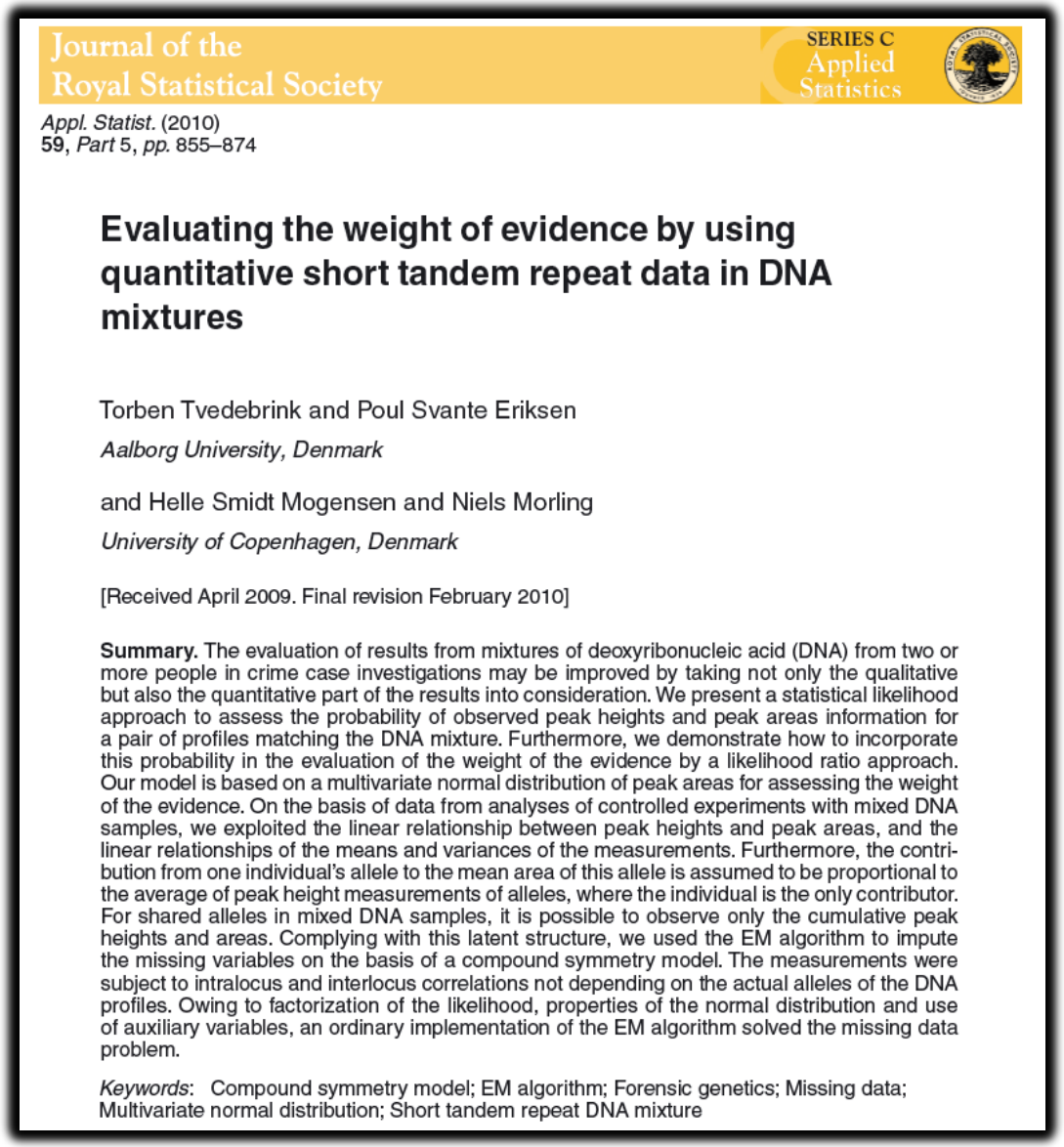
Drop out Events by Marker



Drop out Events by Marker



Stochastic Threshold in STR-Validator



Journal of the Royal Statistical Society
SERIES C Applied Statistics
Appl. Statist. (2010)
59, *Part* 5, *pp.* 855–874

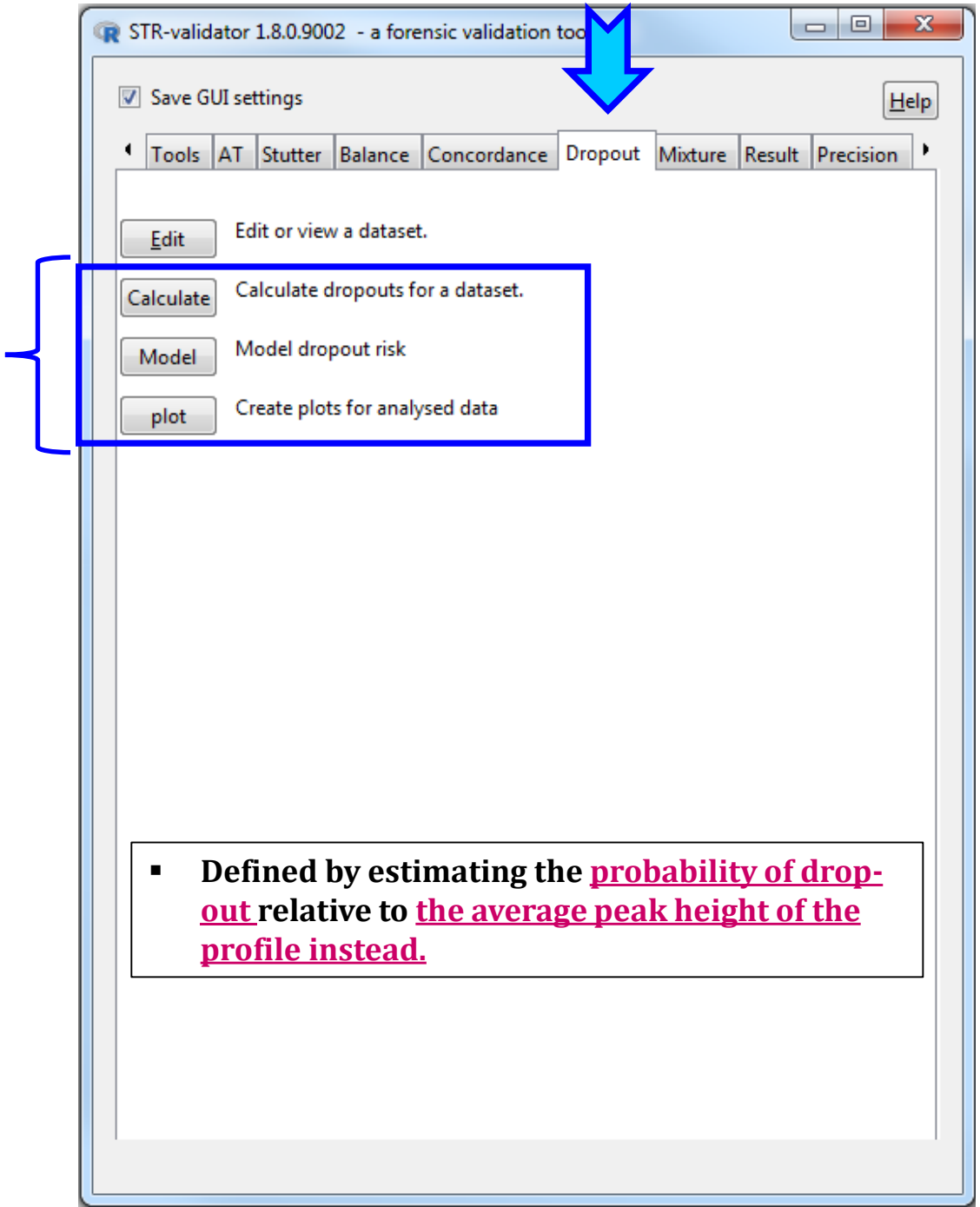
Evaluating the weight of evidence by using quantitative short tandem repeat data in DNA mixtures

Torben Tvedebrink and Poul Svante Eriksen
Aalborg University, Denmark
and Helle Smidt Mogensen and Niels Morling
University of Copenhagen, Denmark

[Received April 2009. Final revision February 2010]

Summary. The evaluation of results from mixtures of deoxyribonucleic acid (DNA) from two or more people in crime case investigations may be improved by taking not only the qualitative but also the quantitative part of the results into consideration. We present a statistical likelihood approach to assess the probability of observed peak heights and peak areas information for a pair of profiles matching the DNA mixture. Furthermore, we demonstrate how to incorporate this probability in the evaluation of the weight of the evidence by a likelihood ratio approach. Our model is based on a multivariate normal distribution of peak areas for assessing the weight of the evidence. On the basis of data from analyses of controlled experiments with mixed DNA samples, we exploited the linear relationship between peak heights and peak areas, and the linear relationships of the means and variances of the measurements. Furthermore, the contribution from one individual's allele to the mean area of this allele is assumed to be proportional to the average of peak height measurements of alleles, where the individual is the only contributor. For shared alleles in mixed DNA samples, it is possible to observe only the cumulative peak heights and areas. Complying with this latent structure, we used the EM algorithm to impute the missing variables on the basis of a compound symmetry model. The measurements were subject to intralocus and interlocus correlations not depending on the actual alleles of the DNA profiles. Owing to factorization of the likelihood, properties of the normal distribution and use of auxiliary variables, an ordinary implementation of the EM algorithm solved the missing data problem.

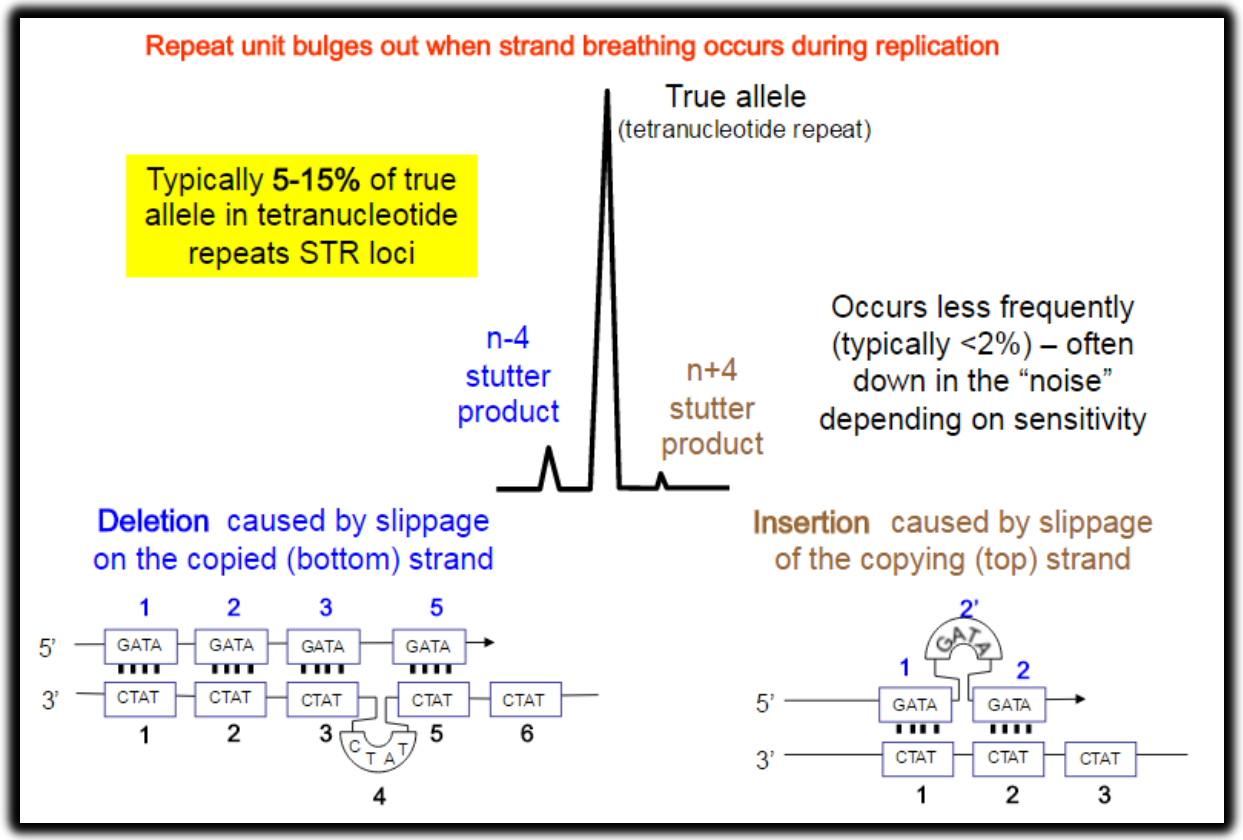
Keywords: Compound symmetry model; EM algorithm; Forensic genetics; Missing data; Multivariate normal distribution; Short tandem repeat DNA mixture



- **Defined by estimating the probability of drop-out relative to the average peak height of the profile instead.**

Stutter

- Is a well-characterized PCR artifact.
- Appears as a minor peak one or more repeat units upstream or downstream from a true allele.
- Results from strand slippage during the amplification process



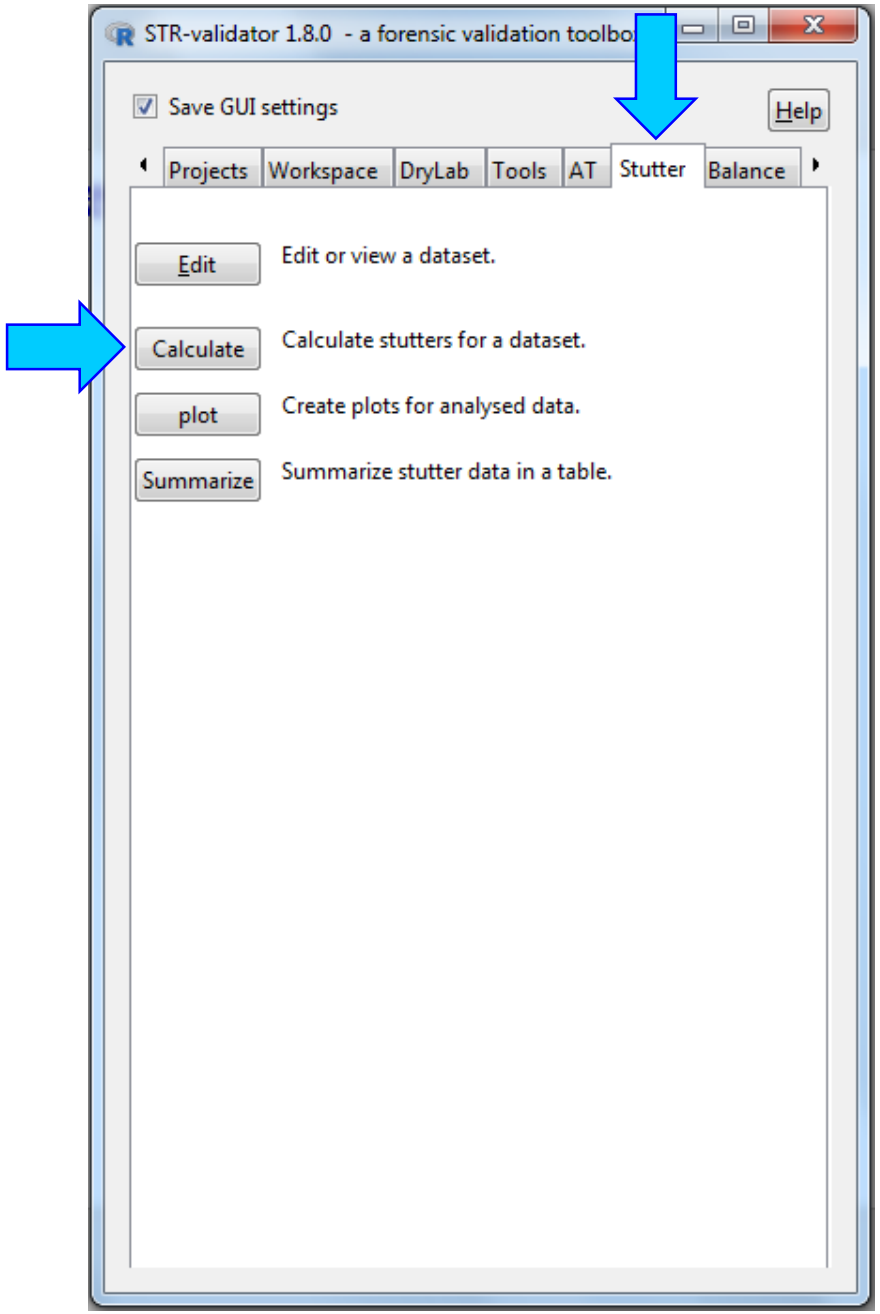
[1] SWGDAM Validation Guidelines for DNA Analysis Methods, (2016).

Courtesy Dr. John M. Butler

Experimental Procedure for Stutter Ratio

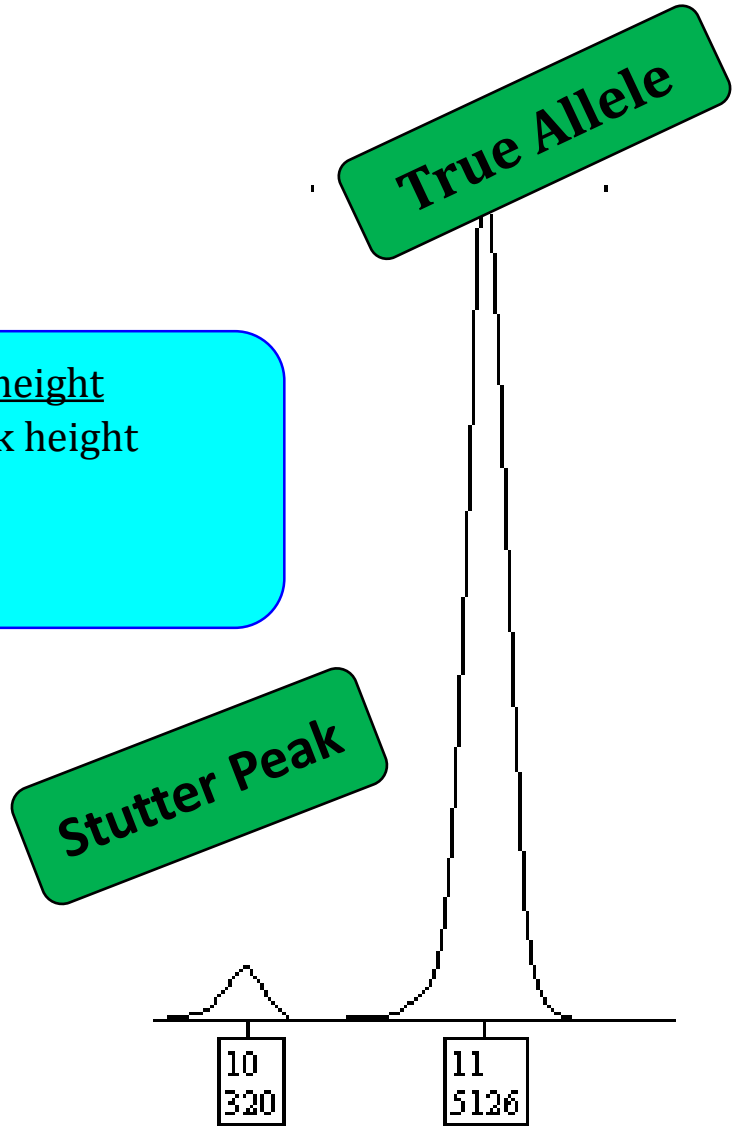
- 95 single source samples at 1.0 ng of DNA input included in stutter ratio calculation
- Analyzed at AT=1 in all dye channels with stutter filters turned off
- Export __GenotypeTable.txt from GeneMapper with at least the following information: “Sample.Name”, “Marker”, “Allele”, and “Height”.

How are Stutters Calculated in *STR-Validator*?



Stutter Ratio = Stutter peak height / True allele peak height

= 320 / 5126

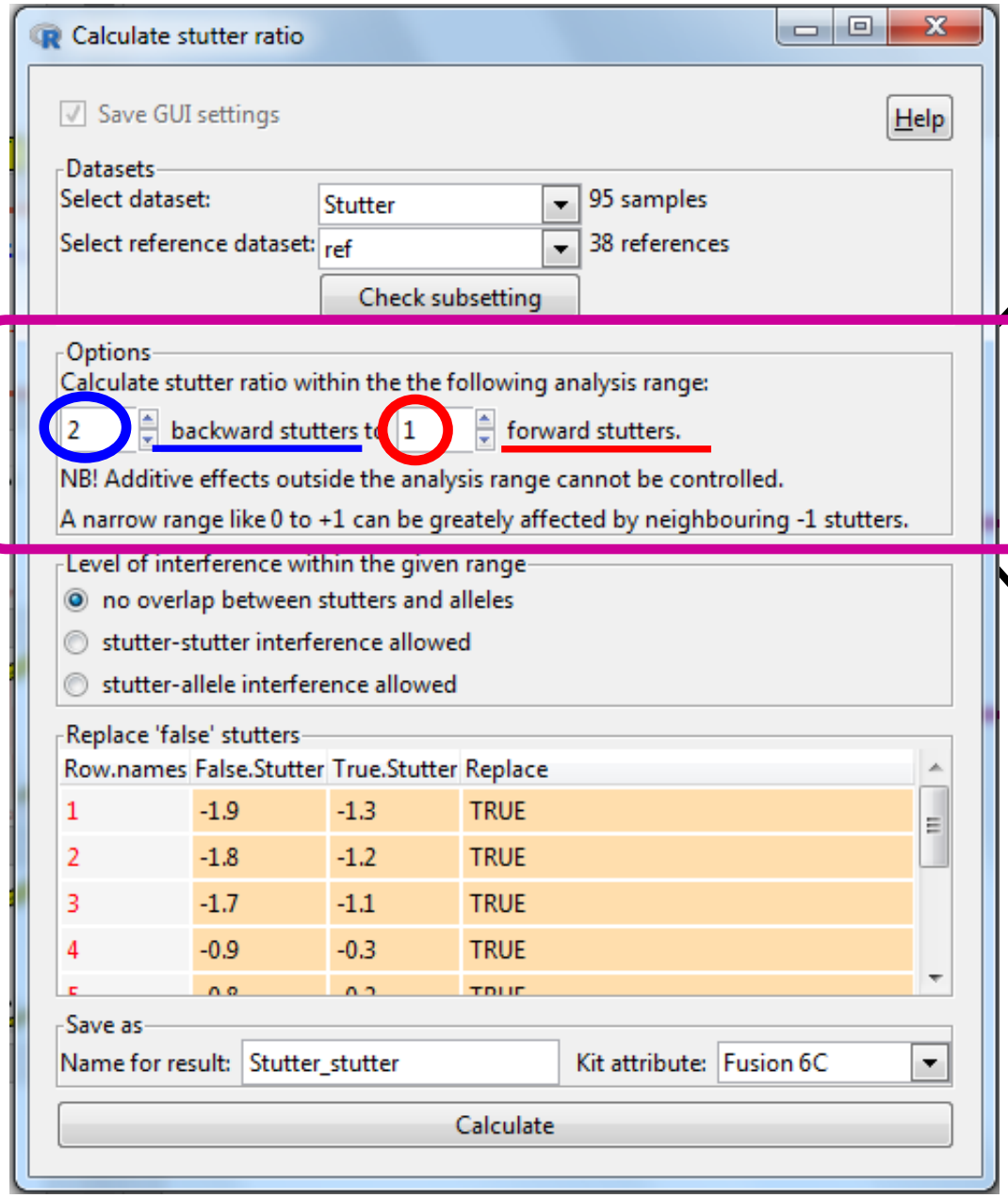


Stutter peak designation – True Allele designation

= 10 - 11

= -1 type of stutter

Analysis Range of Stutter Ratio



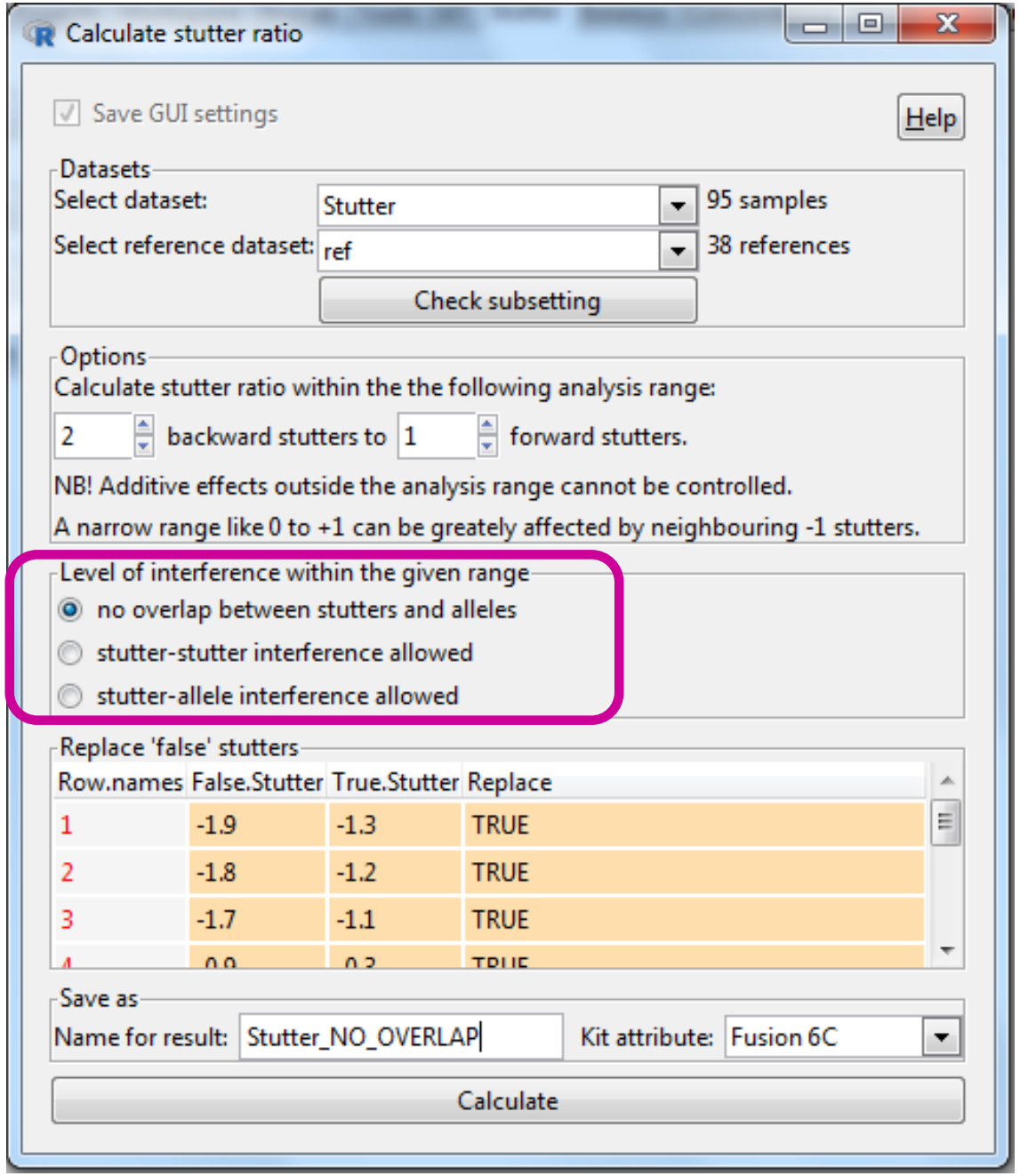
❖Number of backward stutters =2

an i.e. max repeat difference 2 = n-2 repeats

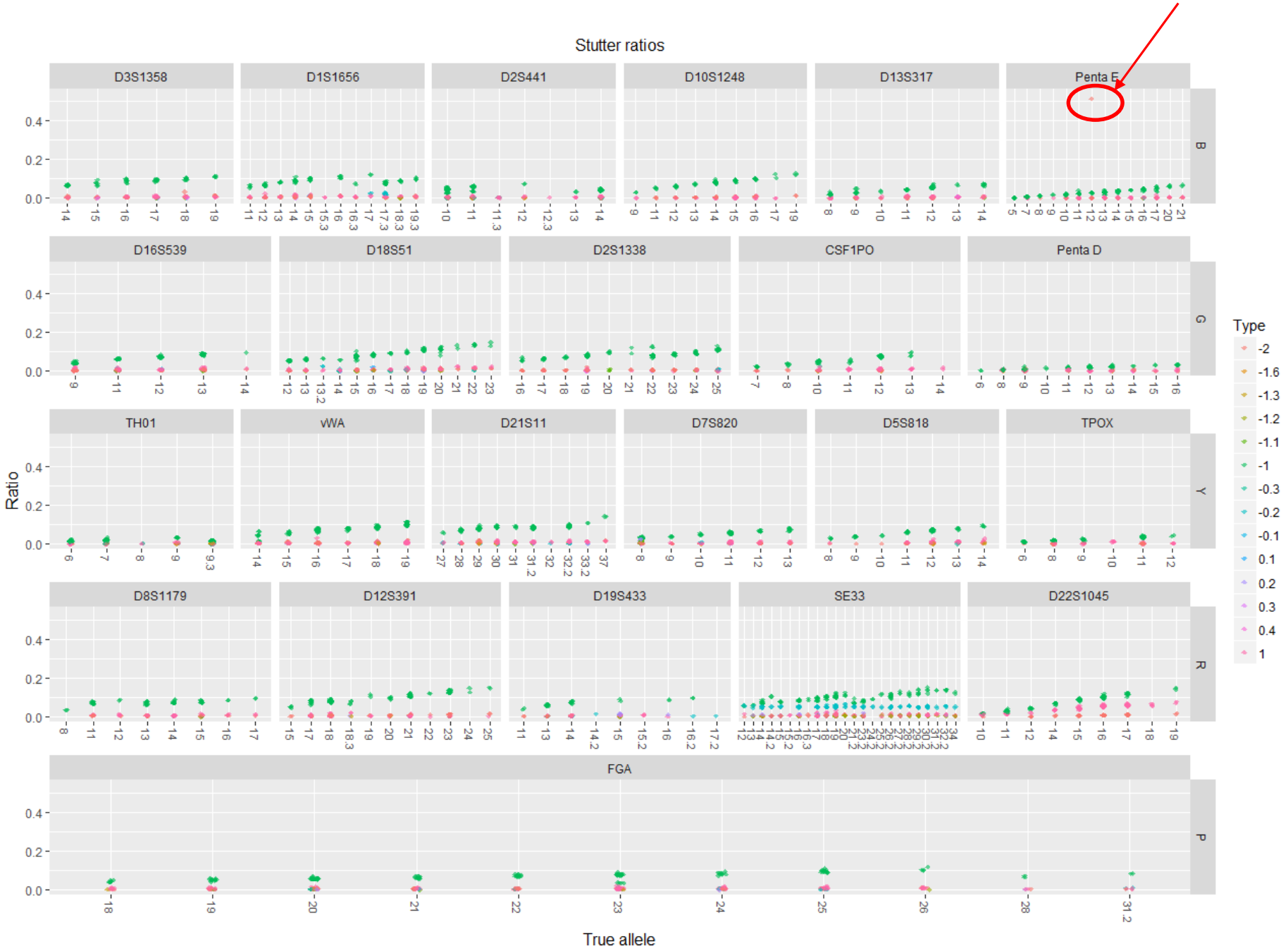
❖Number of forward stutters = 1

an i.e. max repeat difference 1 = n+1 repeats

Level of Interference

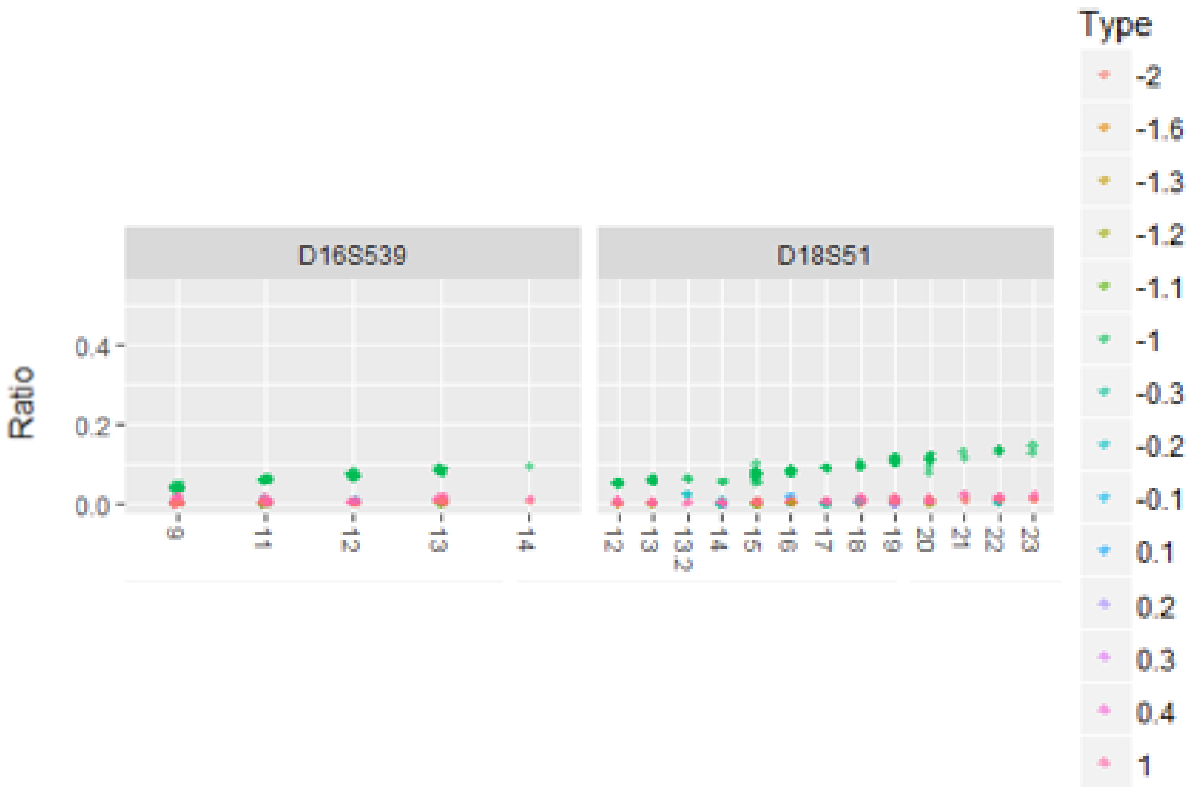


Stutter Ratio as a Function of Parent Allele

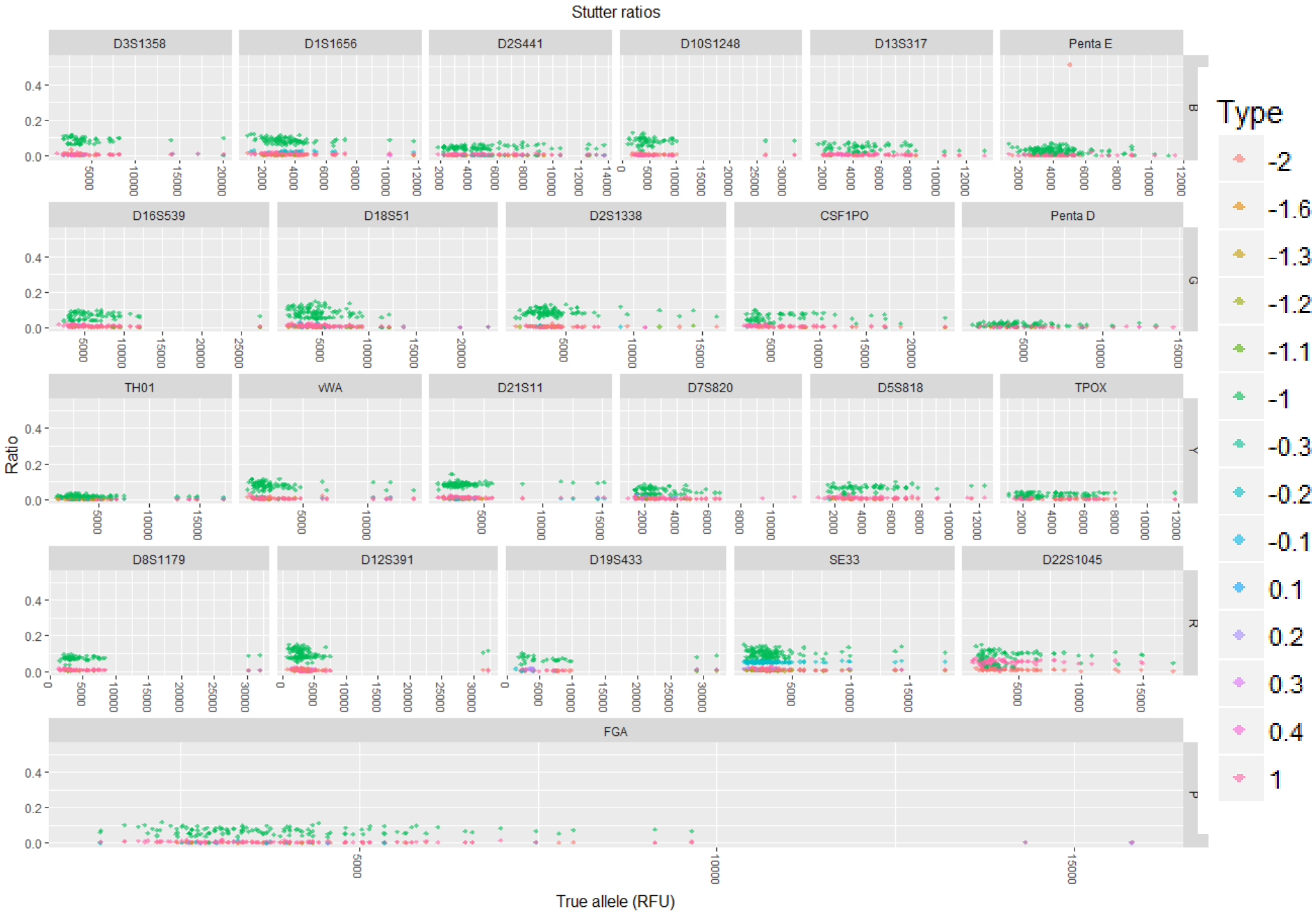


Stutter Ratio increases as the number of repeats increases

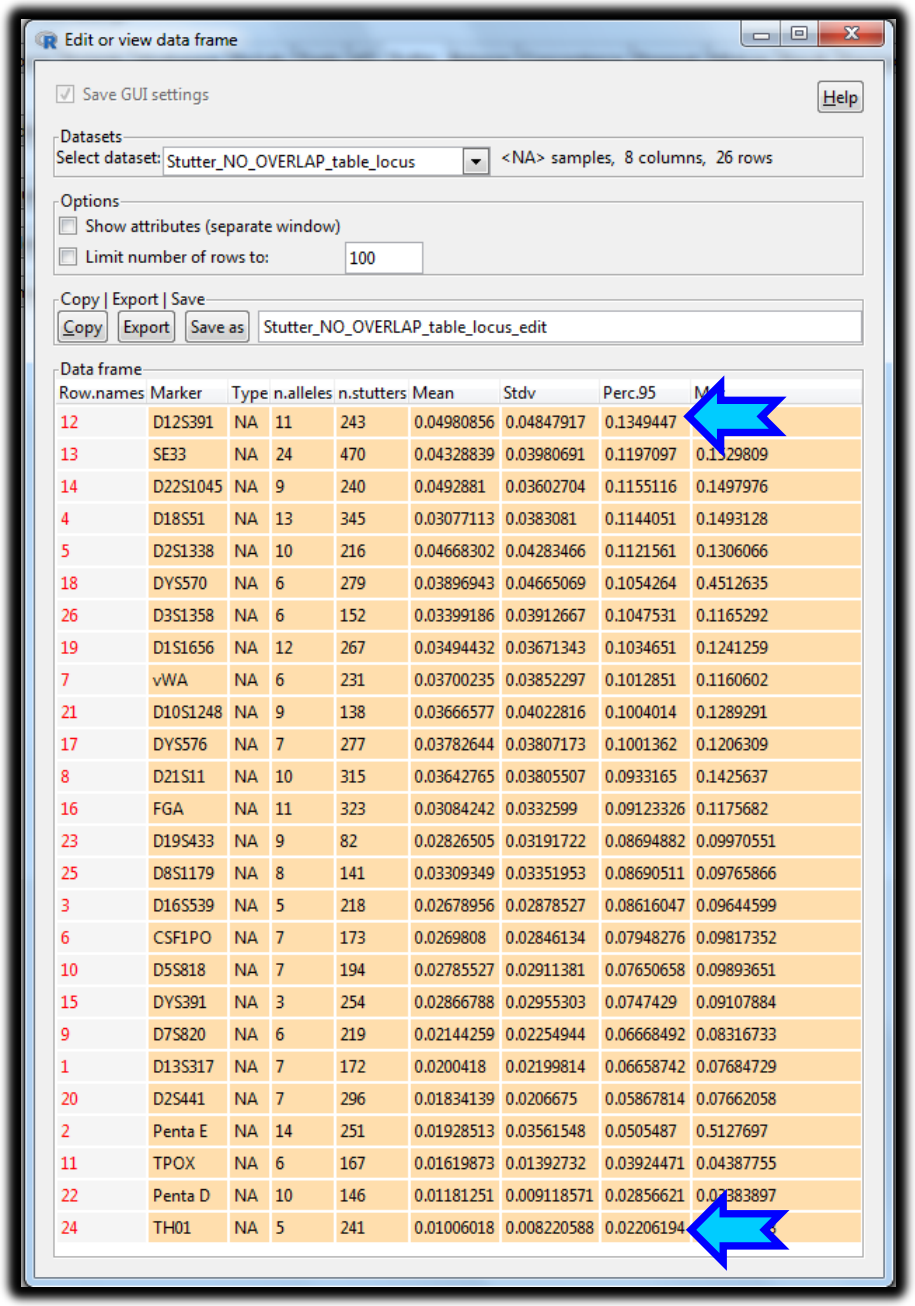
Stutter Ratio as a Function of Parent Allele



Stutter Ratio as a Function of Peak Height



Stutter Percentage at Each Locus



Row.names	Marker	Type	n.alleles	n.stutters	Mean	Stdv	Perc.95	M
12	D12S391	NA	11	243	0.04980856	0.04847917	0.1349447	
13	SE33	NA	24	470	0.04328839	0.03980691	0.1197097	0.1329809
14	D22S1045	NA	9	240	0.0492881	0.03602704	0.1155116	0.1497976
4	D18S51	NA	13	345	0.03077113	0.0383081	0.1144051	0.1493128
5	D2S1338	NA	10	216	0.04668302	0.04283466	0.1121561	0.1306066
18	DYS570	NA	6	279	0.03896943	0.04665069	0.1054264	0.4512635
26	D3S1358	NA	6	152	0.03399186	0.03912667	0.1047531	0.1165292
19	D1S1656	NA	12	267	0.03494432	0.03671343	0.1034651	0.1241259
7	vWA	NA	6	231	0.03700235	0.03852297	0.1012851	0.1160602
21	D10S1248	NA	9	138	0.03666577	0.04022816	0.1004014	0.1289291
17	DYS576	NA	7	277	0.03782644	0.03807173	0.1001362	0.1206309
8	D21S11	NA	10	315	0.03642765	0.03805507	0.0933165	0.1425637
16	FGA	NA	11	323	0.03084242	0.0332599	0.09123326	0.1175682
23	D19S433	NA	9	82	0.02826505	0.03191722	0.08694882	0.09970551
25	D8S1179	NA	8	141	0.03309349	0.03351953	0.08690511	0.09765866
3	D16S539	NA	5	218	0.02678956	0.02878527	0.08616047	0.09644599
6	CSF1PO	NA	7	173	0.0269808	0.02846134	0.07948276	0.09817352
10	D5S818	NA	7	194	0.02785527	0.02911381	0.07650658	0.09893651
15	DYS391	NA	3	254	0.02866788	0.02955303	0.0747429	0.09107884
9	D7S820	NA	6	219	0.02144259	0.02254944	0.06668492	0.08316733
1	D13S317	NA	7	172	0.0200418	0.02199814	0.06658742	0.07684729
20	D2S441	NA	7	296	0.01834139	0.0206675	0.05867814	0.07662058
2	Penta E	NA	14	251	0.01928513	0.03561548	0.0505487	0.5127697
11	TPOX	NA	6	167	0.01619873	0.01392732	0.03924471	0.04387755
22	Penta D	NA	10	146	0.01181251	0.009118571	0.02856621	0.03383897
24	TH01	NA	5	241	0.01006018	0.008220588	0.02206194	

Stutter percentages ranged from 2.2 % (TH01) to 13.4 % D12S391

Comparison of the n-1 Type Stutter with the Manufacturer Stutter Percentages

PPF6C_INTERNAL VALIDATION_STRVALIDATOR			
Marker	Mean	Stdv	Mean + 3*Stdv (%)
D22S1045	0.077	0.036	18.5
D12S391	0.099	0.024	17.2
SE33	0.101	0.021	16.5
D18S51	0.085	0.026	16.4
vWA	0.078	0.023	14.8
D10S1248	0.083	0.020	14.4
D2S1338	0.088	0.017	14.0
D3S1358	0.087	0.016	13.6
D16S539	0.064	0.019	12.1
CSF1PO	0.060	0.020	12.1
D21S11	0.086	0.012	12.0
D19S433	0.070	0.015	11.5
D5S818	0.065	0.016	11.2
D8S1179	0.074	0.011	10.8
D13S317	0.046	0.018	9.9
D7S820	0.048	0.017	9.8
DYS391	0.067	0.007	8.8
D2S441	0.047	0.010	7.8
Penta E	0.027	0.016	7.6
TPOX	0.027	0.010	5.7
Penta D	0.017	0.008	4.2
TH01	0.018	0.005	3.3

Stutter percentages ranged from 3.3 % (TH01) to 18.5 % D22S1045

PPF6C_DEVELOPMENTAL VALIDATION			
Marker	Mean	Stdv	Mean + 3*Stdv (%)
D12S391	0.091	2.8	17.5
D22S1045	0.072	3.2	16.8
SE33	0.1	2	16
D18S51	0.081	2.2	14.7
D1S1656	0.085	2	14.5
vWA	0.069	2.5	14.4
D2S1338	0.085	1.7	13.6
D3S1358	0.09	1.5	13.5
FGA	0.071	1.8	12.5
D19S433	0.07	1.7	12.1
D16S539	0.064	1.8	11.8
D8S1179	0.073	1.5	11.8
CSF1PO	0.055	1.9	11.2
D5S818	0.061	1.6	10.9
D13S317	0.049	1.8	10.3
D7S820	0.047	1.6	9.5
DYS391	0.067	0.9	9.4
D2S441	0.048	1.4	9
Penta E	0.027	1.5	7.2
TPOX	0.025	1	5.5
TH01	0.022	0.9	4.9
Penta D	0.019	0.9	4.6

Stutter percentages ranged from 4.6 % (PentaD) to 17.5 % D12S391

[1] M.G. Ensenberger, K.A. Lenz, L.K. Matthies, G.M. Hadinoto, J.E. Schienman, A.J. Przech, M.W. Morganti, D.T. Renstrom, V.M. Baker, K.M. Gawrys, M. Hoogendoorn, C.R. Steffen, P. Martin, A. Alonso, H.R. Olson, C.J. Sprecher, D.R. Storts, Developmental validation of the PowerPlex((R)) Fusion 6C System, Forensic science international. Genetics 21 (2016) 134-44.

Peak Balance

SWGDAM Guidelines

Peak Height Ratio (PHR)

- Establish potential expectations for allele pairing to define genotypes for mixed samples. It is an indication of which alleles may be heterozygous pairs.

- To express the PHR as a percentage: divide the peak height of an allele with a lower relative fluorescence unit (RFU) value by the peak height of an allele with a higher RFU value, and then multiplying this value by 100

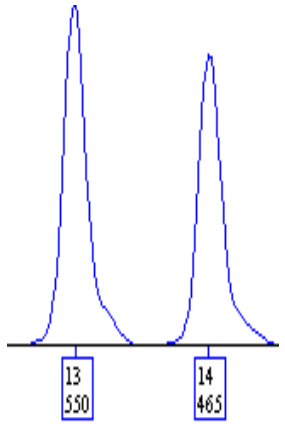
[1] SWGDAM Validation Guidelines for DNA Analysis Methods, (2016).

Experimental Design for Peak Height Ratio Analysis

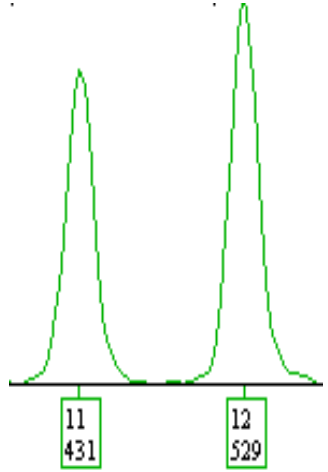
- 95 single source samples at 1.0 ng of DNA input included in PHR calculation
- Analyzed at your AT in all dye channels
- Export __GenotypeTable.txt from GeneMapper with at least the following information: “Sample.Name”, “Marker”, “Allele”, and “Height”.

Calculation of Intra-locus Peak Balance in STR-validator

D10S1248



CSF1PO



Hb = Peak Height HMW / Peak Height LMW = 465 / 550 = 0.85

Hb = Peak Height LMW / Peak Height HMW = 550 / 465 = 1.18

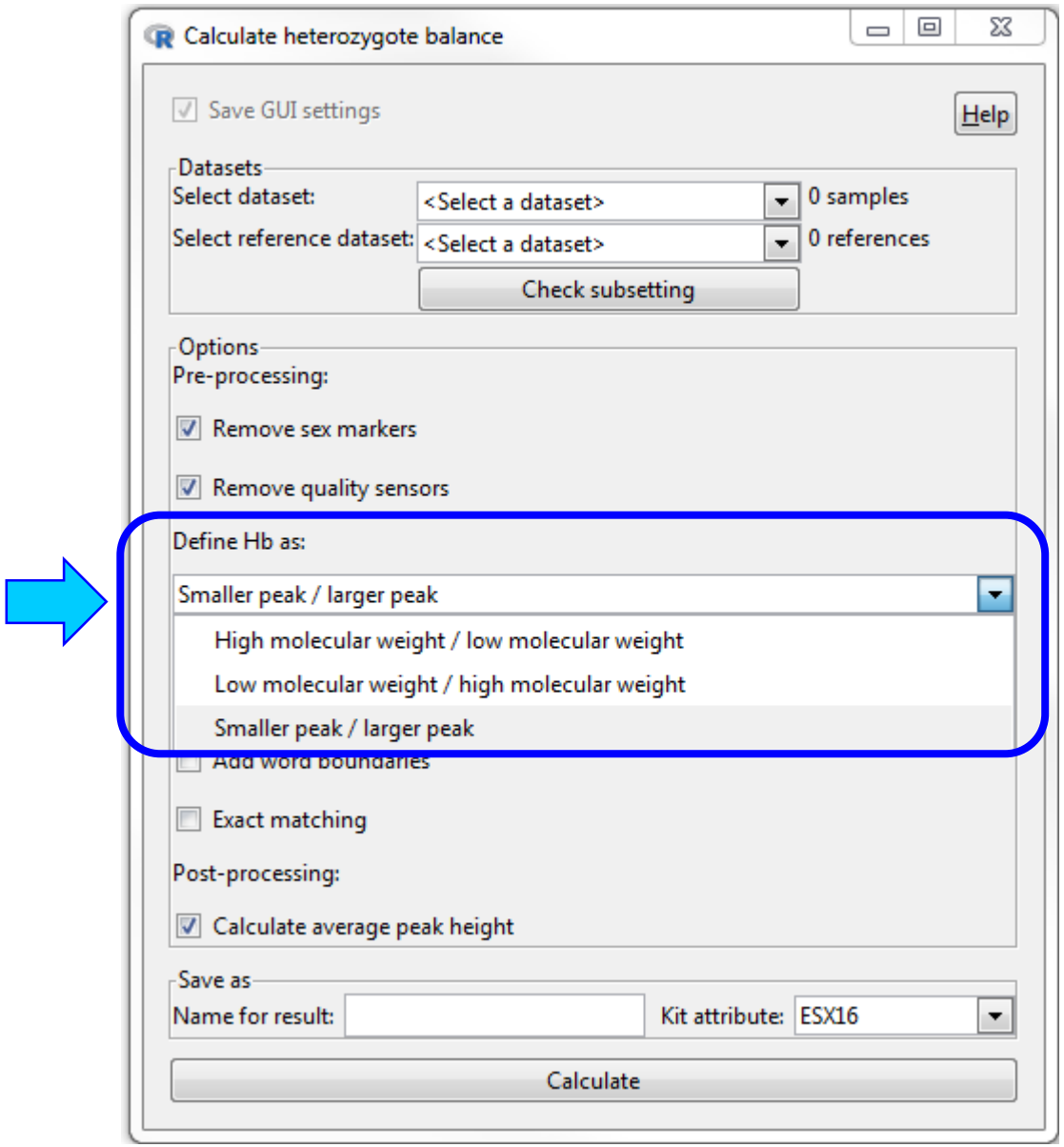
Hb = Peak Height smaller / Peak Height larger = 465 / 550 = 0.85

Hb = Peak Height HMW / Peak Height LMW = 529 / 431 = 1.2

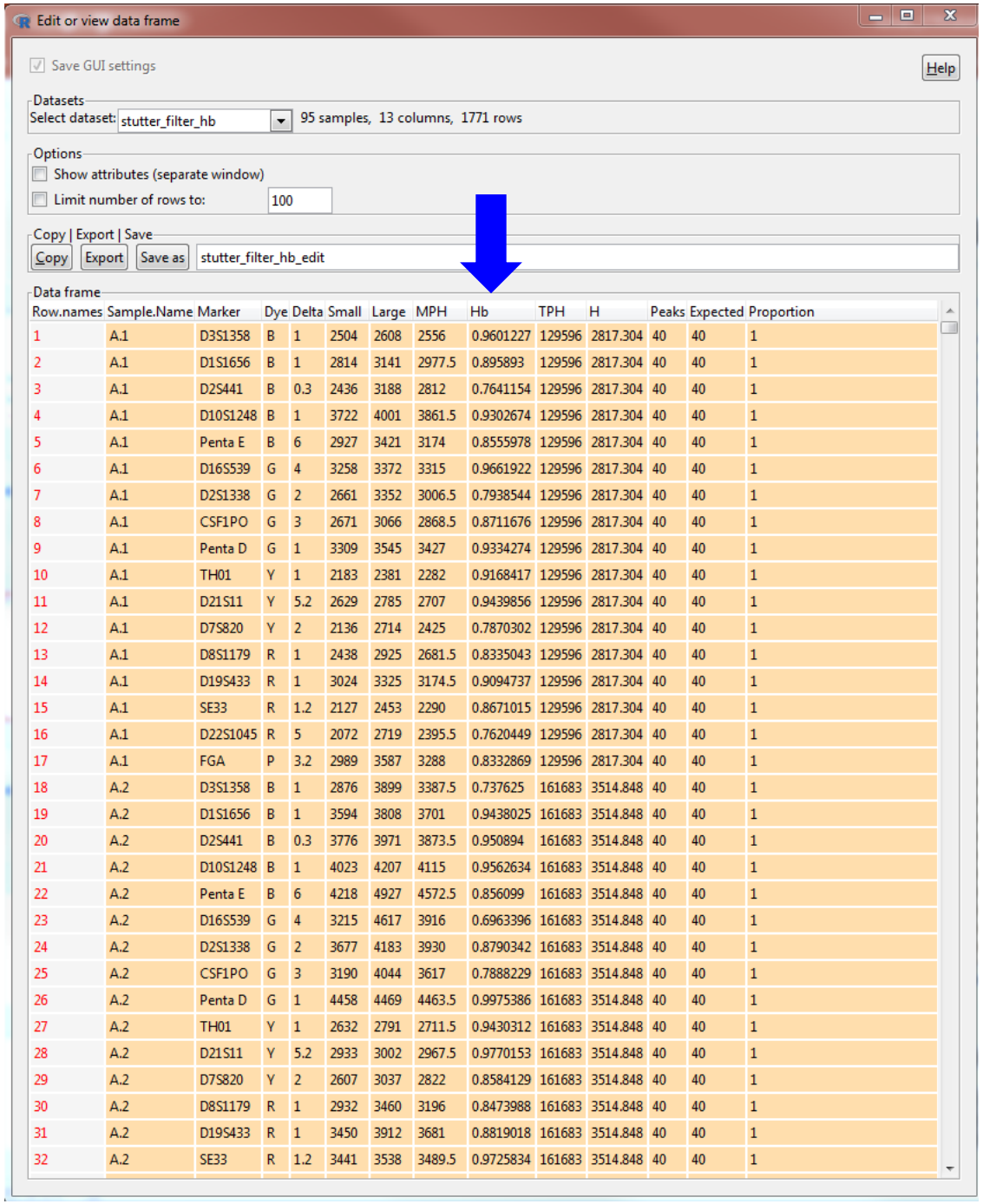
Hb = Peak Height LMW / Peak Height HMW = 431 / 529 = 0.81

Hb = Peak Height smaller / Peak Height larger = 431 / 529 = 0.81

Calculate Balance

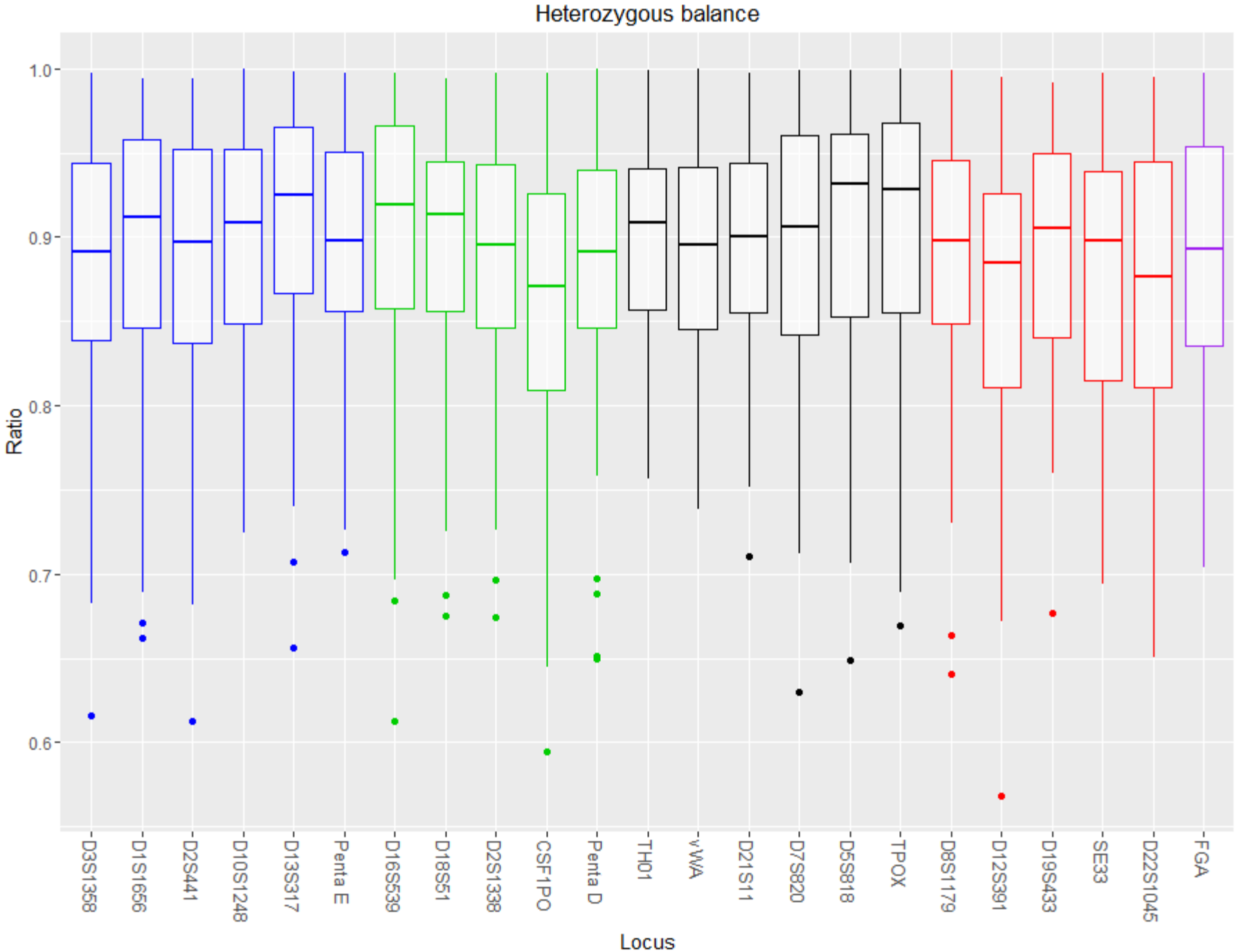


Results of Hb Analysis



Row.names	Sample.Name	Marker	Dye	Delta	Small	Large	MPH	Hb	TPH	H	Peaks	Expected	Proportion
1	A.1	D3S1358	B	1	2504	2608	2556	0.9601227	129596	2817.304	40	40	1
2	A.1	D1S1656	B	1	2814	3141	2977.5	0.895893	129596	2817.304	40	40	1
3	A.1	D2S441	B	0.3	2436	3188	2812	0.7641154	129596	2817.304	40	40	1
4	A.1	D10S1248	B	1	3722	4001	3861.5	0.9302674	129596	2817.304	40	40	1
5	A.1	Penta E	B	6	2927	3421	3174	0.8555978	129596	2817.304	40	40	1
6	A.1	D16S539	G	4	3258	3372	3315	0.9661922	129596	2817.304	40	40	1
7	A.1	D2S1338	G	2	2661	3352	3006.5	0.7938544	129596	2817.304	40	40	1
8	A.1	CSF1PO	G	3	2671	3066	2868.5	0.8711676	129596	2817.304	40	40	1
9	A.1	Penta D	G	1	3309	3545	3427	0.9334274	129596	2817.304	40	40	1
10	A.1	TH01	Y	1	2183	2381	2282	0.9168417	129596	2817.304	40	40	1
11	A.1	D21S11	Y	5.2	2629	2785	2707	0.9439856	129596	2817.304	40	40	1
12	A.1	D7S820	Y	2	2136	2714	2425	0.7870302	129596	2817.304	40	40	1
13	A.1	D8S1179	R	1	2438	2925	2681.5	0.8335043	129596	2817.304	40	40	1
14	A.1	D19S433	R	1	3024	3325	3174.5	0.9094737	129596	2817.304	40	40	1
15	A.1	SE33	R	1.2	2127	2453	2290	0.8671015	129596	2817.304	40	40	1
16	A.1	D22S1045	R	5	2072	2719	2395.5	0.7620449	129596	2817.304	40	40	1
17	A.1	FGA	P	3.2	2989	3587	3288	0.8332869	129596	2817.304	40	40	1
18	A.2	D3S1358	B	1	2876	3899	3387.5	0.737625	161683	3514.848	40	40	1
19	A.2	D1S1656	B	1	3594	3808	3701	0.9438025	161683	3514.848	40	40	1
20	A.2	D2S441	B	0.3	3776	3971	3873.5	0.950894	161683	3514.848	40	40	1
21	A.2	D10S1248	B	1	4023	4207	4115	0.9562634	161683	3514.848	40	40	1
22	A.2	Penta E	B	6	4218	4927	4572.5	0.856099	161683	3514.848	40	40	1
23	A.2	D16S539	G	4	3215	4617	3916	0.6963396	161683	3514.848	40	40	1
24	A.2	D2S1338	G	2	3677	4183	3930	0.8790342	161683	3514.848	40	40	1
25	A.2	CSF1PO	G	3	3190	4044	3617	0.7888229	161683	3514.848	40	40	1
26	A.2	Penta D	G	1	4458	4469	4463.5	0.9975386	161683	3514.848	40	40	1
27	A.2	TH01	Y	1	2632	2791	2711.5	0.9430312	161683	3514.848	40	40	1
28	A.2	D21S11	Y	5.2	2933	3002	2967.5	0.9770153	161683	3514.848	40	40	1
29	A.2	D7S820	Y	2	2607	3037	2822	0.8584129	161683	3514.848	40	40	1
30	A.2	D8S1179	R	1	2932	3460	3196	0.8473988	161683	3514.848	40	40	1
31	A.2	D19S433	R	1	3450	3912	3681	0.8819018	161683	3514.848	40	40	1
32	A.2	SE33	R	1.2	3441	3538	3489.5	0.9725834	161683	3514.848	40	40	1

Hb vs Marker


Heterozygous balance
Ratio
1.0
0.9
0.8
0.7
0.6
Locus
D3S1358
D1S1656
D2S441
D10S1248
D13S317
Penta E
D16S539
D18S51
D2S1338
CSF1PO
Penta D
TH01
vWA
D21S11
D7S820
D5S818
TPOX
D8S1179
D12S391
D19S433
SE33
D22S1045
FGA

Hb Summary Statistics by Locus

Edit or view data frame

Save GUI settings | Help

Datasets
Select dataset: stutter_filter_hb_table_locus <NA> samples, 7 columns, 23 rows

Options
Show attributes (separate window)
Limit number of rows to: 100

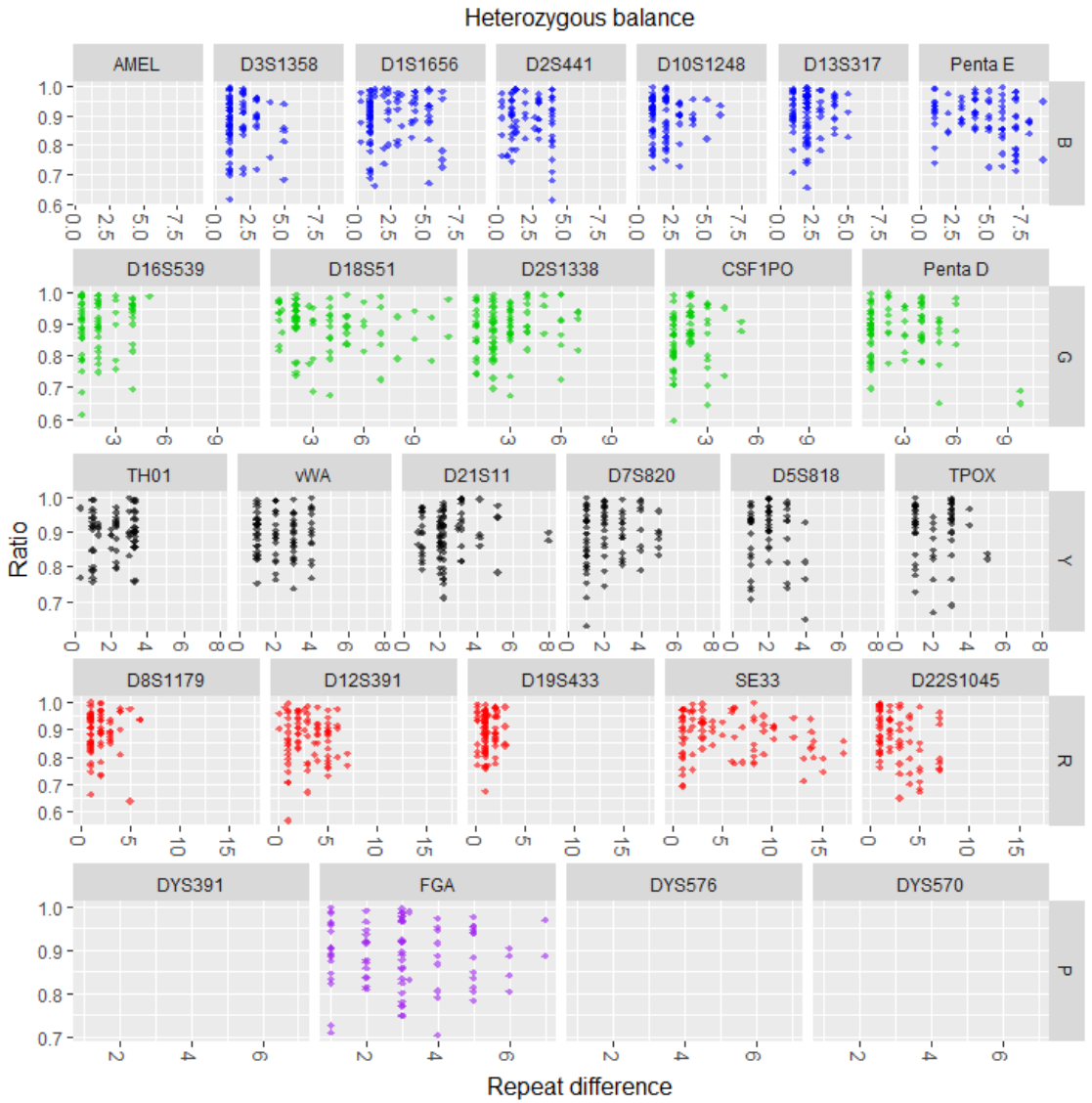
Copy | Export | Save
Copy | Export | Save as | stutter_filter_hb_table_locus_edit

Data frame

Row.names	Marker	Hb.n	Hb.Min	Hb.Mean	Hb.Sd	Hb.Max	Hb.Perc.5
16	D22S1045	71	0.6503783	0.8717967	0.09215057	0.9953015	0.7053714
1	D3S1358	79	0.6157654	0.8760157	0.08802503	0.9971048	0.7061884
8	CSF1PO	67	0.594947	0.8630976	0.0871485	0.9975135	0.7082371
2	D1S1656	89	0.6619263	0.8918997	0.08517202	0.9944644	0.7159104
21	D12S391	81	0.5685678	0.8689462	0.08372068	0.9947577	0.7301872
7	D2S1338	89	0.6741703	0.8877156	0.07447339	0.9974315	0.7384348
23	D5S818	59	0.6486384	0.8998864	0.08525826	0.9989265	0.7390588
13	D8S1179	70	0.6403179	0.8893265	0.07678957	0.9993565	0.7402201
5	Penta E	85	0.7128883	0.8901279	0.07284199	0.9975642	0.7457725
18	D18S51	80	0.6751799	0.8934292	0.07425412	0.9942959	0.7465895
3	D2S441	70	0.6128105	0.8854903	0.08221411	0.9938511	0.7476048
15	SE33	89	0.6940328	0.8813438	0.07648513	0.997509	0.7497926
6	D16S539	70	0.612762	0.8980979	0.08689724	0.9971216	0.7501955
17	FGA	87	0.703799	0.8887843	0.07407832	0.9977156	0.7559048
20	TPOX	58	0.6695573	0.9055004	0.08049648	1	0.7581212
9	Penta D	82	0.649378	0.8813575	0.07957125	0.9995117	0.7584861
4	D10S1248	77	0.7240583	0.8944208	0.0703462	0.999473	0.7668952
10	TH01	67	0.7567347	0.8966153	0.06506874	0.9986888	0.768737
14	D19S433	79	0.6768735	0.8928187	0.06995027	0.9918119	0.7692973
19	vWA	77	0.7381534	0.8940821	0.06415032	0.999591	0.7762276
12	D7S820	79	0.6301075	0.894619	0.07576705	0.998855	0.7785241
22	D13S317	79	0.6561222	0.9073231	0.07401154	0.9981813	0.780273
11	D21S11	87	0.7107579	0.8937022	0.06685845	0.9977768	0.7813653

Hb ≥ 0.7

Hb and Distance Between the Two Alleles



Precision Analysis

SWGDAM Guidelines

Precision

- Characterizes the degree of mutual agreement among a series of individual measurements/values and results.
- Depends only on the distribution of random errors and does not relate to the true value or specified value.
- Is usually expressed in terms of imprecision and computed as a standard deviation of the test results.

How to measure the precision of your instrument?

All measured alleles should fall within a ± 0.5 bp window around the measured size for the corresponding allele in the allelic ladder.

[1] SWGDAM Validation Guidelines for DNA Analysis Methods, (2016).

Experimental Procedure for Precision Analysis

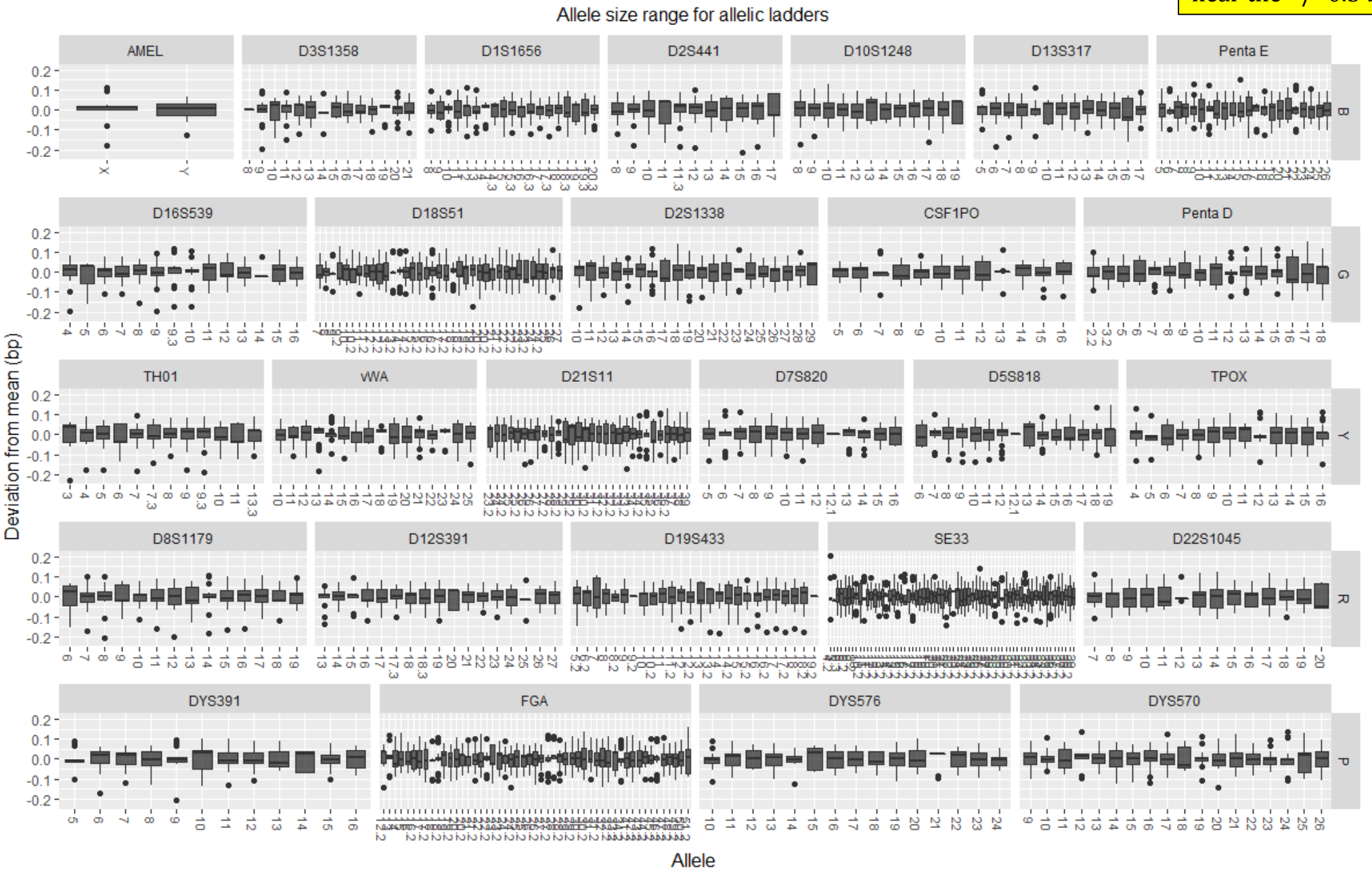
One injection of 24 ladders performed

- 1 ladder assigned as the “ladder”
- 22 ladders assigned as samples (A-V)
- Analyzed at your Analytical Threshold (AT)
- Export _GenotypeTable.txt from GeneMapper with at least the following information: “Sample.Name”, “Marker”, “Allele” and “Size”.

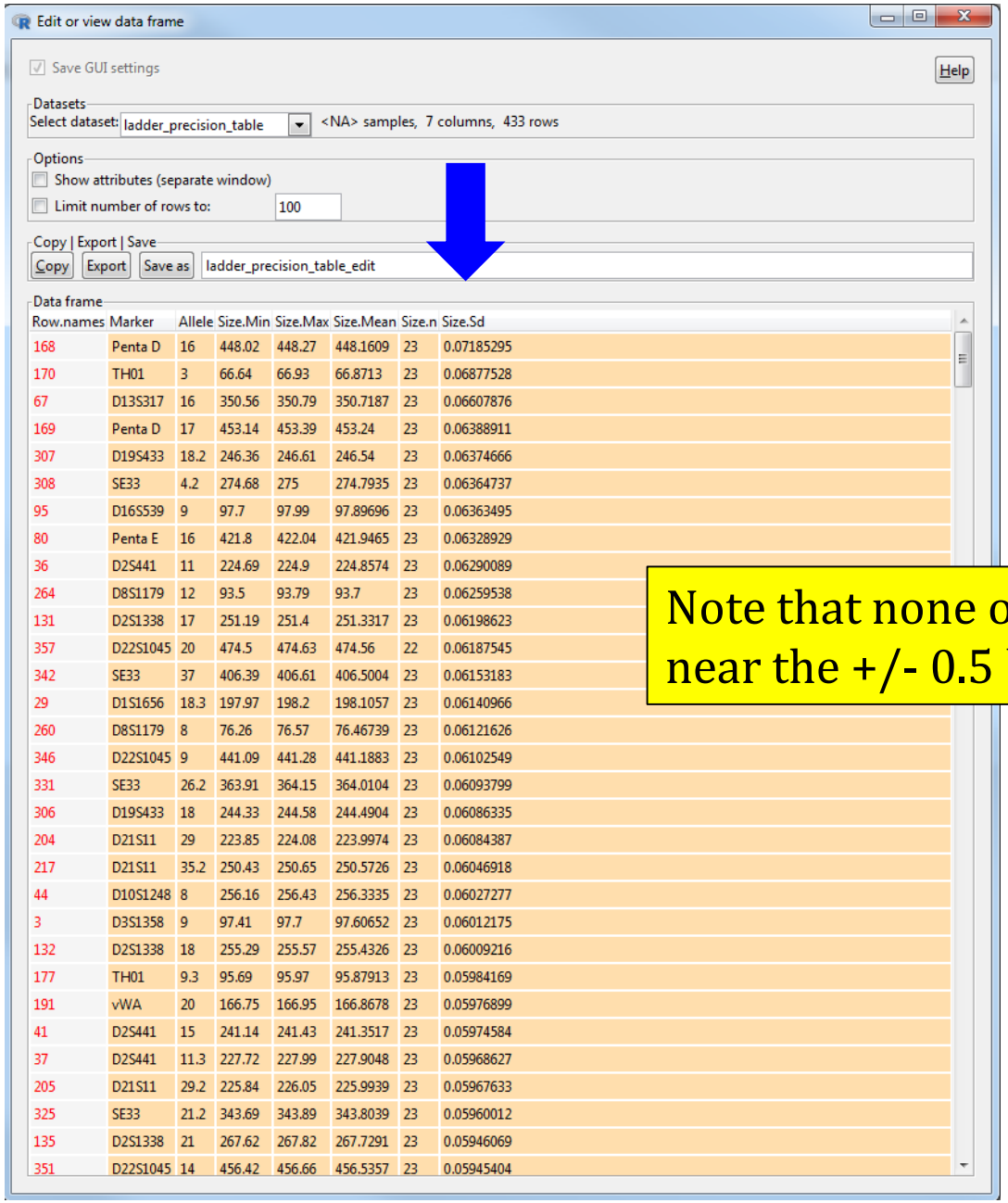
	10	11	12
A	Ladder	Ladder	Ladder
B	Ladder	Ladder	Ladder
C	Ladder	Ladder	Ladder
D	Ladder	Ladder	Ladder
E	Ladder	Ladder	Ladder
F	Ladder	Ladder	Ladder
G	Ladder	Ladder	Ladder
H	Ladder	Ladder	Ladder

Size Precision Boxplot for the Allelic Ladders by Allele

Note that none of the intervals extend near the +/- 0.5 bp range



Summary Statistics for Precision Data, "Size.Sd" has been Sorted by Descending Order



Row.names	Marker	Allele	Size.Min	Size.Max	Size.Mean	Size.n	Size.Sd
168	Penta D	16	448.02	448.27	448.1609	23	0.07185295
170	TH01	3	66.64	66.93	66.8713	23	0.06877528
67	D13S317	16	350.56	350.79	350.7187	23	0.06607876
169	Penta D	17	453.14	453.39	453.24	23	0.06388911
307	D19S433	18.2	246.36	246.61	246.54	23	0.06374666
308	SE33	4.2	274.68	275	274.7935	23	0.06364737
95	D16S539	9	97.7	97.99	97.89696	23	0.06363495
80	Penta E	16	421.8	422.04	421.9465	23	0.06328929
36	D2S441	11	224.69	224.9	224.8574	23	0.06290089
264	D8S1179	12	93.5	93.79	93.7	23	0.06259538
131	D2S1338	17	251.19	251.4	251.3317	23	0.06198623
357	D22S1045	20	474.5	474.63	474.56	22	0.06187545
342	SE33	37	406.39	406.61	406.5004	23	0.06153183
29	D1S1656	18.3	197.97	198.2	198.1057	23	0.06140966
260	D8S1179	8	76.26	76.57	76.46739	23	0.06121626
346	D22S1045	9	441.09	441.28	441.1883	23	0.06102549
331	SE33	26.2	363.91	364.15	364.0104	23	0.06093799
306	D19S433	18	244.33	244.58	244.4904	23	0.06086335
204	D21S11	29	223.85	224.08	223.9974	23	0.06084387
217	D21S11	35.2	250.43	250.65	250.5726	23	0.06046918
44	D10S1248	8	256.16	256.43	256.3335	23	0.06027277
3	D3S1358	9	97.41	97.7	97.60652	23	0.06012175
132	D2S1338	18	255.29	255.57	255.4326	23	0.06009216
177	TH01	9.3	95.69	95.97	95.87913	23	0.05984169
191	vWA	20	166.75	166.95	166.8678	23	0.05976899
41	D2S441	15	241.14	241.43	241.3517	23	0.05974584
37	D2S441	11.3	227.72	227.99	227.9048	23	0.05968627
205	D21S11	29.2	225.84	226.05	225.9939	23	0.05967633
325	SE33	21.2	343.69	343.89	343.8039	23	0.05960012
135	D2S1338	21	267.62	267.82	267.7291	23	0.05946069
351	D22S1045	14	456.42	456.66	456.5357	23	0.05945404

Note that none of the intervals extend near the +/- 0.5 bp range

Interest in further training on how to use the software

ISHI INTERNATIONAL SYMPOSIUM ON HUMAN IDENTIFICATION
SEATTLE, WA • OCTOBER 2–5, 2017

Exhibitor Information | Attendee Regstration

Search

HOME | NEWS | PROGRAM | ATTENDEES | EXHIBIT | VENUE | PRESS

Efficient Validation Using STR-Validator

Chair: Oskar Hansson, Section for Forensic Biology, Oslo University Hospital, Department for Forensic Medical Services, Oslo University Hospital

Cost: $50 Fees include course materials and break

Description: Validation of new STR typing kits are often time consuming which can delay implementation of new technology. STR-validator is a free open-source software tool designed to help forensic genetic laboratories evaluate the data. Lectures and exercises will provide the necessary skills for experimental design and efficient evaluation. Validation according to ENFSI recommendations/SWGDAM guidelines is exemplified using real data. Practical examples will highlight quality control of the data and result. The goal is to provide hands-on experience with the software.

Acknowledgments:

Applied Genetics Group-Forensic Genetics

- Peter Vallone
- Michael Coble
- Erica Romsos

Norwegian Institute of Public Health

- Oskar Hansson

Contact Info:

sarah.riman@nist.gov

(301) 975-4162

References

1. https://sites.google.com/site/forensicapps/strvalidator
2. https://github.com/OskarHansson/strvalidator
3. https://cran.r-project.org/web/packages/strvalidator/index.html
4. O. Hansson, P. Gill, T. Egeland, STR-validator: An open source platform for validation and process control, Forensic Science International: Genetics 13 (2014) 154–166.
5. P. Gill, L. Gusmao, H. Haned, W. Mayr, N. Morling, W. Parson, L. Prieto, M. Prinz, H. Schneider, P. Schneider, B. Weir, 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, Forensic Science International: Genetics 6 (2012) 679–688.
6. Peter Gill, Roberto Puch-Solis, James Curran, The low-template-DNA (stochastic) threshold-Its determination relative to risk analysis for national DNA databases, Forensic Science International: Genetics, Volume 3, Issue 2, March 2009, Pages 104-111
7. Torben Tvedebrink, Poul Svante Eriksen, Helle Smidt Mogensen, Niels Morling, Evaluating the weight of evidence by using quantitative short tandem repeat data in DNA mixtures Journal of the Royal Statistical Society: Series C (Applied Statistics), Volume 59, Issue 5, 2010, Pages 855-874,
8. J. Bregu et al. Analytical Thresholds and Sensitivity: Establishing RFU Thresholds for Forensic DNA Analysis. JFS (2013) 1 pg 120-129.
9. Ullrich J. Monich, Ken Duy, Muriel Medard, Viveck Cadambe, Lauren E. Alfonse, and Catherine Grgicak. Probabilistic characterisation of baseline noise in STR proles. Forensic Science International: Genetics.
10. J.-A. Bright, J. Turkington, J. Buckleton. Examination of the variability in mixed DNA profile parameters for the Identifiler™ multiplex. Forensic Sci. Int. Genet., 4 (2) (2010), pp. 111–114.

FORENSIC SCIENCE ERROR MANAGEMENT

INTERNATIONAL FORENSICS SYMPOSIUM

July 24-28, 2017 @NIST, Gaithersburg, MD

July 24-28 @NIST, Gaithersburg, MD

Technical Tracks

- Crime Scene
- Death Investigation
- Human Factors
- Legal Factors
- Quality Assurance
- Laboratory Management
- Criminalistics
- Digital Evidence

go.usa.gov/x9yEK

Or search for “NIST 2017 forensic error management”