2012 Mixture Interpretation Workshop:

Mixtures Using SOUND Statistics, Interpretation, & Conclusions



Profile 1 Validation & Research: Impacts on Interpretation of Low-Template Mixtures Catherine M. Grgicak

October 15, 2012 Nashville, TN





Steps in DNA Interpretation



Principles Behind Thresholds

Thresholds (example values)	Principles Behind (if properly set based on lab- & kit-specific empirical data)
Analytical Threshold (e.g. 50 RFU)	Below this value, observed peaks cannot be reliably distinguished from noise
Limit of Linearity (e.g. 5000 RFU)	Above this value, the CCD can become saturated and peaks may not accurately reflect relative signal quantities (e.g., flat-topped peaks) and lead to pull-up/bleed-through between dye color channels
Stochastic Threshold (e.g. 250 RFU)	Above this peak height value, it is reasonable to assume that allelic dropout of a sister allele of a heterozygote has not occurred at that locus; single alleles above this value is single-source samples are assumed homozygous
Stutter Threshold (e.g. 15%)	Below this value, a peak in the reverse (or forward) stutter position can be designated as a stutter artifact with single-source samples or some mixtures (often higher with lower DNA amounts)
Peak Height Ratio Threshold (e.g. 60%)	Above this value, two heterozygous alleles can be grouped as a possible genotype (often lower with lower DNA amounts)
Major/Minor Ratio (e.g. 4:1)	When the ratio of contributors is closer than this value in a two-person mixture, it becomes challenging and often impossible to correctly associate genotype combinations to either the major or minor contributor

Steps in DNA Interpretation



What analytical (RFU) threshold do you use?: Data from 107 responses

ISHI Mixture Workshop (Oct 2012)



Analytical Threshold



Method 1, 2, 3 and 4 - Negatives

-Negative sample run with an internal size standard (not shown) using manufacturer's recommended protocol

Negative = extraction or amplification negative



Method 1, 2 - Negatives

$$\begin{split} AT_{M1} &= \overline{Y}_{bl} + ks_{bl} \\ AT_{M1} &= 3.11 + (3*1.14) = 6.53 \\ AT_{M1} &= 3.11 + (3*1.14) = 6.53 \\ AT_{M1} &= 7 \end{split}$$

Method 3 and 4 - Negatives

$$AT_{M3} = 2(Y_{\max} - Y_{\min})$$

$$AT_{M3} = 2(9-0) = 18$$
$$AT_{M3} = 18$$

NOTE: Because we are NOT using raw data (but analyzed GeneMapper data), data below 0 RFU is not 'observed' and therefore, the number calculated is smaller than expected!!! HOWEVER, the calculated AT is still larger than either Method 1 or 2!

RFU Signal of Blank	No. of observ ations	Percent Rank	AT _{M4} at rank >=99%
1	206	3.87	
2	1481	31.73	
3	1884	67.16	
4	1161	89.00	
5	453	97.51	6
6	110	99.59	
7	18	99.92	
8	3	99.98	
9	1	100	

Methods 5 & 6 – Positives (Standard Curves)

- Regression of positive samples (i.e. single source samples)
- Amplified 0.0625-4ng dilution series, injected 5s using manufacturer's recommended protocol
- Plot of Input DNA (ng) versus average peak height (per color) with error bars
 - If a peak was homozygous, the RFU was divided by 2



- The points at 2 and 4 ng fall off the line (PCR efficiency approaching a plateau)!
- The error bars become larger with increased DNA input!

 A weighted linear regression is within the linear range (i.e. 0.0625 – 1 ng) was used.

- b (y-intercept) = -2.30
- S_y (standard error of regression) = 10.77

$$AT_{M4} = b + 3S_y$$

$$AT_{M4} = 31$$

- b (y-intercept) = -2.30
- S_y (standard error of regression) = 10.77
- t-stat (n-1=4) and alpha of 99% t=3.75

$$AT_{M5} = b + t_{n-1,\alpha}S_y$$

$$AT_{M5} = 39$$

Method	Origin	Analytical Threshold for green 5s injection example
1	Negatives	7
2	Negatives	4
3	Negatives	18
4	Negatives	6
5	DNA Series	31
6	DNA Series	39

Before you choose, consider the following slides...

Type II Error – False non-labeling of alleles (Drop-out)

Single source 0.125ng, 1ul 3130 prep volume



Type II Error – False non-labeling of alleles



-As AT increases, locus DO increases, while allele DO stabilizes after 50 RFU then starts to decrease after AT of ~150 RFU.

-Although a higher AT (i.e. >150 RFU) begins to decrease the number of loci where allele DO occurs (less stochastic variation),

-Locus DO increases, resulting in an overall increase in DO with AT for Low-template samples

Balancing Type I and Type II Errors – < 0.5ng



-AT's have a large effect on the ability to detect/label alleles.

- -Red = high level of allele drop-out, blue=low levels of allele drop-out.
- To take a 'conservative' approach and utilize high AT values leads to a substantial level of Type II errors for low-level samples (i.e. <1000RFU).

Baselines Positives *≠* **Baselines Negatives**



Type I Error – False Labeling of Noise Peaks

- This is not instrument baseline/noise
- Single source DNA data amplified from 0.0625 2 ng
 - Differentiated 'noise' from artifact
 - -A, pull-up, stutter (+ or -), spikes, dye artifacts
- Plotted RFU of the known/expected peak versus the highest 'noise' peak
- High noise with >0.5 ng of DNA, higher AT needed for higher-template samples



What analytical (RFU) threshold do you want to apply to the mixture?:

- One that is derived by analyzing baseline from negatives
- 2. One that is derived from analyzing standard curve
- 3. 50 RFU
- 4. 150 RFU because I want 14% to minimize stochastic effects

Data from 92 responses

ISHI Mixture Workshop (Oct 2012)



The AT is.....

#2.

- AT of 30 RFU (AT_{\rm M5} 95% confidence) based on samples that contained DNA.

Color	AT _{M5} 95% confidence	RFU Threshold Applied for ISHI workshop
Blue	19	
Green	24	
Yellow	16	<u>30 RFU</u>
Red	13	

 NB: 30 RFU for all colors was used for simplicity and ATs applied on a per color basis is recommended

Steps in DNA Interpretation



Stochastic Threshold - Method 1 (Max height)



Stochastic Threshold - Method 2 (Pr(D))



Gill et al. FSI Genetics, 2009, 3, 104-111.

This method minimizes the chance of wrongly deciding or concluding a heterozygous locus is homozygous

Stochastic Threshold - Method 3

Minimizing the error of wrongly deciding or concluding a heterozygous locus is homozygous

Minimizing the error of wrongly deciding or concluding a homozygous locus is heterozygous

At various STs, for all heterozygous loci, determine proportion of heterozygous loci falsely labeled as homozygous For all homozygous loci, determine the proportion of homozygous loci falsely considered possible heterozygotes. Plot the proportions against each other.



Method	ST	Description	Stochastic Threshold for ISHI workshop
1	160	Max peak height observed where sister allele is < AT	
2	150	Pr(D) < 0.01	<u>150 RFU</u>
3	146	Lowest overall error rate	

Steps in DNA Interpretation



Peak Height Ratio Thresholds

- Evaluate PHRs at various DNA template levels (e.g., dilution series of DNA).
- Different PHR expectations at different peak height ranges may be established.
- PHR requirements should be based on empirical data for interpretation of DNA typing results from single-source samples. Different PHR expectations can be applied to individual loci; alternatively, a single PHR expectation can be applied to multiple loci (e.g., 60%)."

Peak Height Ratios



Power Plex 16 data kindly provided by NIST,>8000 alleles

Peak Height Ratios



<0.5 ng	>0.5 ng
(500RFU)	(500RFU)
0.2	0.5

How would you determine peak height ratio information for casework use?

- 1. Use one value for all profiles
- 2. Use average-3SD
- 3. Use min. observed
- 4. Use 2 values: based on amount amplified
- 5. Use locus specific values



Peak Height Ratio Imbalance



When assuming that a mixture of DNA from only 2 contributors is present, the Peak Height Ratio may aid in the interpretation of the profile data when used to pair heterozygous alleles

The PHR Threshold is.....

Method	ST	Description	PHR Threshold for ISHI workshop
	0.2	Min peak height ratio observed at any target (ng)	
1	0.2 (<500RFU) and 0.5 (>500RFU)	Min peak height ratio – target dependent	<u>0.2</u> (<500RFU) AND
2	0.4	Average – 3SD	<u>0.5</u>
	0.3 (<500RFU) And 0.5 (>500RFU)	Average – 3SD -target dependent	<u>(>500RFU)</u>

Steps in DNA Interpretation



Stutter

3'

CTAT

1



Walsh, P.S., et al. (1996). Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus vWA. *Nucleic Acids Research, 24*, 2807-2812.

CTAT

6



CTAT

3

CTAT

5

CTAT

2

Insertion caused by slippage of the copying (top) strand





STR_StutterFreq!

Welcome to STR_StutterFreq!

Version <04-Jan-10>

STR_StutterFreq is a specialty analysis tool for characterizing stutter frequency... Development of **STR_StutterFreq** was funded in part by the National Institute of Justice.

 Program developed by Dave Duewer (NIST) to rapidly calculate stutter frequencies.

Stutter Filters



Fig. 4. c and d are unambiguous alleles, b is a minor allele in a stutter position and a is an unambiguous minor allele.

Gill *et al.* (2006) DNA Commission of the International Society of Forensic Genetics: Recommendations on the interpretation of mixtures. *Forensic Sci. Int.* 160: 90-101

Stutter Filters



Fig. 2. A two person mixture with major peaks C, D and minor peaks A. There is an additional peak present in a stutter position (B).



Fig. 3. A two person mixture with major peaks C, D and minor peaks A, B, where B is in a stutter position.

Possibly AB

(heterozygote)

Could also be AC, AD, AA, or A,? (dropout)

Stutter Filters Minor component is probative.....

	D8S1179	D21S11	D7S820	CSF1PO
Minor (w/ filters)	13,14 or 14,14 or 14,16	30,33.2	10,F	10,11
Minor (w/out filters)	13,14 or 14,14 or 14,15 or 14,16	30,32.2	10,F	10,11
Standard	14,15	30,33.2	10,10	10,11

Included/Excluded/Inconclusive

Analyze without stutter filters



Stutter Threshold for ISHI will be max observed from manufacturer's validation data

Locus	Stutter Threshold
CSF1PO	9.2%
D2S1338	11.1%
D3S1358	10.7%
D5S818	6.8%
D7S820	8.2%
D8S1179	8.2%
D13S317	8.0%
D16S539	10.4%
D18S51	17.0%
D19S433	13.3%
D21S11	9.4%
FGA	14.7%
TH01	5.1%
TPOX	4.8%
vWA	12.6%

Given our validation....Our interpretation scheme is.....

AT	30RFU	
ST	150RFU	
Stutter Filter	Off if examining minor contributors	Max observed when on
PHR	0.2 (<500RFU)	0.5 (>500RFU)
Major:Minor	4:1	

Profile (stutter filter off)



Profile 1. – Minor Genotype Possibilities

Description	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	TH01	D13S317	D16S539	D2S1338
Inferred									
Genotypes of									
Minor									
Suspect 1	13,14	28,29	10,11	12,14	14,15	9,9.3	12,12	8,11	16,18
Suspect 2	13,14	28,29	10,11	10,14	14,16	9,9.3	12,13	11,13	18,19

Description	vWA	TPOX	D18S51	AMEL	D5S818	FGA
Inferred						
Genotypes of						
Minor						
Suspect 1	17,17	8,8	15,15	X,Y	11,11	21,24
Suspect 2	14,15	8,11	15,20	X,Y	11,13	22,23

In your opinion should **Suspect 1** be considered a potential contributor to Profile 1?:

- Yes, suspect 1 should be included as a potential contributor and a match statistic should be determined
- 2. No, suspect 1 should be excluded
- 3. Inconclusive...I can't tell.



Data from 78 responses

In your opinion should **Suspect 2** be considered a potential contributor to Profile 1?:

- Yes, suspect 2 should be included as a potential contributor and a match statistic should be determined
- 2. No, suspect 2 should be excluded
- 3. Inconclusive...I can't tell.



Data from 85 responses ISHI Mixture Workshop (Oct 2012)

	AT 30		Ū]		
D21511	ST 150		=U				
	Stutter Filter		6				
	PHR	0.2 (<500	RFU)	0.5 (>500RFU)			
215	Major:Minor 4						
(24-600) 29,120,13.7%	= 100(0.2) to	100	Pose	sible genoty	pe co	mbinatio	ns
(20-500) (438-1752)	Ster 20	0.2		(20,29,	,30,31)	
28 30 PH _{si}	$s_{ster28} = 20 \text{ to } 500$	RFU	P	erson 1		Person 2	
100 876 Since F (441-1766) then 28	PH ₂₉ =120 is btw 2 3,29 is possible co	0-500, ombinatior	n ^r	28,29		30,31	
31 883 Since PH ₃₀ Then 28,30	=876 is NOT btw) is NOT a possibl	l 20-500 le combina	ation	28,30		29,31	
 Is the 29 an allele (rem threshold is 9.4%)? Yes Possible non-observed 			28,31		29,30		
contributors?				29,30		28,31	
No			29,31		28,30		
List out all alleles: 28*, 29*, 30, 31 and consi	le	30.31			28,29		
pairs – assuming 2 contrib						<u>.</u>	







		AT		30RFU						
D10001		ST		150RFU						
		Stutte	er Filter	17%						
	L	Pl	HR	0.	2 (<500RFU)	0.5 (>500RFU)	Possible genetype			
A AA	L	Major	r:Minor		4:1			rossible genotype		
13,44,6.2% 19,72,10.8% 14 20		Poss geno	sible type		Pos geno	sible otype ations if	and 19 are stutter (14,15,20)			
713 668	COI	mbina 3 con	ations i tained	t	19 contained		Person 1	Person 2		
15 240	allele, 0 DO			allele, 0 DO		14,14	15,20			
2.10	(1	(13,14,15,20)			(14,15	,19,20)		14,20 or		
1. Are the 13 and 19	Per: 1	son I	Perso 2	n	Person 1	Person 2	14,15	15,20 or 20,20		
Maybe 2. Possible non-	13,	,14	15,20)	14,15	19,20	(14,20)	14,15 or 15,15 or		
observed alleles if 2	13,	,15	14,20)	14,19	15,20		15,20		
No.	13,	20	14,15	5	14,20	15,19	15,15	14,20		
List out all possible alleles: <u>13*</u> , 14, 15, <u>19*</u> , 20 and consider all possible	14,	,15	13,20)	15,19	14,20	15,20	14,14 or 14,15 or 14,20		
	14,	,20	13,15)	15,20	14,19		11,20		
pairs and scenarios– assuming 2 contributors	15,	,20	13,14	Ĺ	19,20	14,15	20,20	14,15		



Profile 1. – Minor Genotype Possibilities

Given our ISHI Thresholds and interpretation standard operating procedures, the genotypes of the minor are.....

Description	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	TH01	D13S317	D16S539	D2S1338
Inferred Genotypes of Minor	13,14	28,29	9,11 or 10,11 or 11,11 or 11,0	9,14 or 10,14 or 11,14 or 12,14 or 14,14 or 14,0	14,14 or 14,15 or 14,16	9,9 or 9,9.3 or 9.3,9.3 or 9,0 or 9.3,0 or 0,0	9,12 or 10,12 or 12,12 or 12,13	8,11 or 9,11 or 11,11 or 11,12 or 11,13	15,18 or 16,18 or 18,18 or 18,19 or 18,0
Suspect 1	13,14	28,29	10,11	12,14	14,15	9,9.3	12,12	8,11	16,18
Suspect 2	13,14	28,29	10,11	10,14	14,16	9,9.3	12,13	11,13	18,19

Description	vWA	ΤΡΟΧ	D18S51	AMEL	D5S818	FGA
Inferred Genotypes of Minor	14,17 or 15,17 or 17,17	8,8 or 8,11 or 11,11 or 8,0 or 11,0 0,0	14,15 or 15,15 or 15,19 or 15,20	X,Y	11,11 or 11,12 or 11,13	20,22 or 21,22 or 22,22 or 22,23 or 22,0
Suspect 1	17,17	8,8	15,15	X,Y	11,11	21,24
Suspect 2	14,15	8,11	15,20	X,Y	11,13	22,23

O= any other allele (not observed)

What about now.....In your opinion should Suspect 1 be considered a potential contributor to Profile 1?: Data from 91 responses ISHI Mixture Workshop (Oct 2012)

- Yes, suspect 1 should be included as a potential contributor and a match statistic should be determined
- 2. No, suspect 1 should be excluded
- 3. Inconclusive...I can't tell.

41% 36% 23%

2

3

1

What about now....In your opinion should Suspect 2 be considered a potential contributor to Profile 1?: Data from 70 responses

- Yes, suspect 2 should be included as a potential contributor and a match statistic should be determined
- 2. No, suspect 2 should be excluded
- 3. Inconclusive...I can't tell.



Profile 1. – Minor Genotype Possibilities

Given our ISHI Thresholds and interpretation standard operating procedures, the genotypes of the minor are.....

Description	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	TH01	D13S317	D16S539	D2S1338
Inferred Genotypes of Minor	13,14	28,29	9,11 or 10,11 or 11,11 or 11,0	9,14 or 10,14 or 11,14 or 12,14 or 14,14 or 14,0	14,14 or 14,15 or 14,16	9,9 or 9,9.3 or 9.3,9.3 or 9,0 or 9.3,0 or 0,0	9,12 or 10,12 or 12,12 or 12,13	8,11 or 9,11 or 11,11 or 11,12 or 11,13	15,18 or 16,18 or 18,18 or 18,19 or 18,0
Suspect 1	13,14	28,29	10,11	12,14	14,15	9,9.3	12,12	8,11	16,18
Suspect 2	13,14	28,29	10,11	10,14	14,16	9,9.3	12,13	11,13	18,19

Description	vWA	ΤΡΟΧ	TPOX D18S51		D)5S818	FGA
Inferred Genotypes of Minor	14,17 or 15,17 or 17,17	8,8 or 8,11 or 11,11 or 8,0 or 11,0 0,0	14,15 or 15,15 or 15,19 or 15,20	X,Y	1 [.] 1 [.]	1,11 or 1,12 or 11,13	20,22 or 21,22 or 22,22 or 22,23 or 22,0
Suspect 1	17,17	8,8	15,15 X,Y			11,11	21,24
Suspect 2	14,15	8,11	15,20	X,Y		11,13	22,23
O= any other allele (not observed)		BU Adv. DN suspec	IA Class, t 1	/9 novice analysts		Truth	
		Includ	bed	5			
		Exclu	ded	1		Included	
		Inconcl	usive	3			

Conclusions

- Given our ISHI Thresholds and interpretation standard operating procedures, the genotypes of the minor can be deduced
- As seen by the large number of genotype possibilities, there is a level of uncertainty associated with this deduction
- All scenarios need to be considered when determining genotype possibilities
- Peak height ratio thresholds can be used to examine possible genotype combinations
- Validation data must be used to establish these thresholds
- Thresholds obtained for a given method must be applied only to evidence obtained using the same method (i.e. kit, injection time, etc).