NIST Validation Studies on the 3500 Genetic Analyzer

https://products.appliedbiosystems.com/ab/en/US/htdocs/productMgr/images/3500_family_big.jpg

Erica Butts

National Institute of Standards and Technology

Applied Biosystems HID University

Atlanta, GA

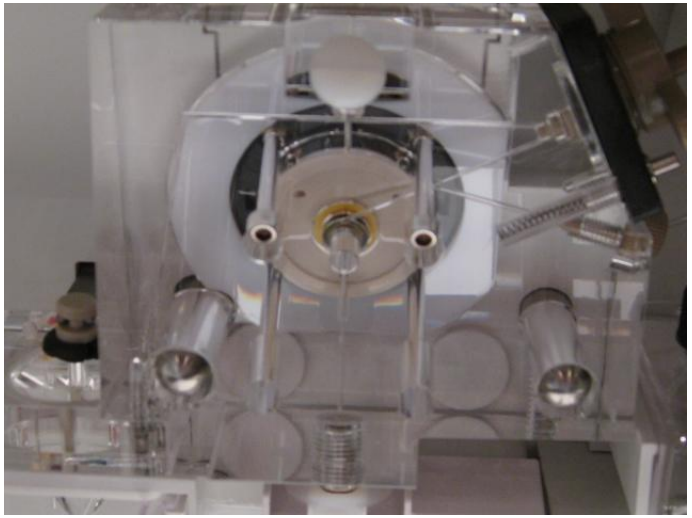
July 12, 2011

Outline

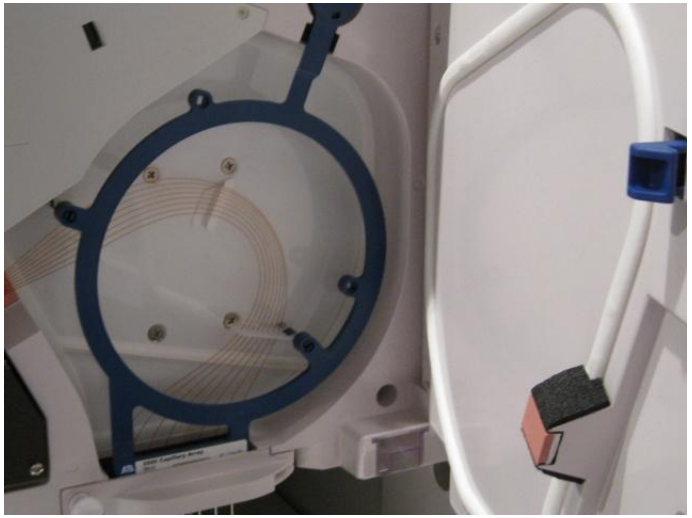
- Details of the new ABI 3500
- Validation design and results with Identifiler and Identifiler Plus
  - Precision
  - Injection parameters and reaction setup
  - Setting Analytical Thresholds
  - Setting Stochastic Thresholds
- Overview of signal normalization

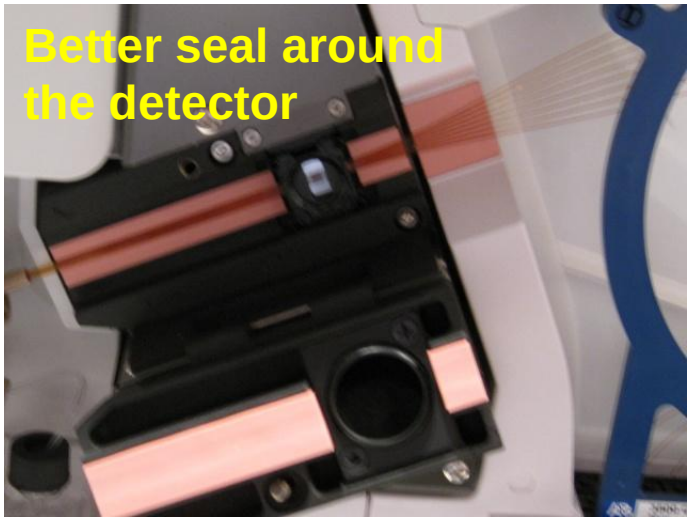
Details of the new ABI 3500

No lower pump block
(less polymer waste)



Improved sealing for better temperature control





Primary Differences Between 31xx and 3500

31xx Instruments

- Argon ion (Ar+) lasers with 488/514 nm wavelengths for fluorescence excitation
- 220V power requirement
- Optimal signal intensity 1500-3000 RFU
- Currently validated and operational in most forensic laboratories

3500 Instruments

- Single-line 505 nm, solid-state long-life laser
- Smaller footprint
- 110V power requirement
- Optimal signal intensity can approach 20,000-30,000 RFU
- Normalization of instrument-to-instrument signal variability
  - Ability to increase or decrease overall signal
- Requires the use of GeneMapper IDX v1.2

What is Validation?

Section 1.1 (SWGDAM Revised Validation Guidelines) Validation is the process by which the scientific community acquires the necessary information to:

(a) Assess the ability of a procedure to obtain reliable results.

(b) Determine the conditions under which such results can be obtained.

(c) Define the limitations of the procedure.

The validation process identifies aspects of a procedure that are critical and must be carefully controlled and monitored.

Reliability, Reproducibility, Robustness

Aspects of Validation

- Reliability
  - Size Standard Comparison
    - What is the difference between LIZ 500 and LIZ 600 v2.0?
  - Injection Parameters
    - What are the best injection parameters for typable data?
    - Adjustment alongside PCR reaction volume?
- Reproducibility
  - Precision
    - Is the precision comparable to the 3130*xl*?
  - Concordance
    - Are the correct allele calls being made?
- Robustness
  - Sensitivity
    - How do the thresholds affect the analysis of data?
  - Mixtures (current work in progress)
    - How often is the minor component identified?

Experimental Summary

Test	Types of Samples Used	Number Examined
Size Standard Comparison	16 Allelic Ladders per size standard (LIZ 500 vs. LIZ 600 v2.0)	32
Injection Parameters	3 samples heterozygous at all loci including Amelogenin 1 ng DNA input	3 samples/injection time Total 15 samples
Precision	Allelic Ladders	24
	3 samples heterozygous at all loci including Amelogenin	6
Sensitivity	Dilution series of 3 samples heterozygous at all loci including Amelogenin	4 Replicates of each dilution series Total 84 samples
Concordance	50 genomic DNA samples	60
	SRM 2391b: 10 genomic DNA samples	
Mixtures	Mixture dilution series of 2 samples heterozygous at all loci including Amelogenin	28
	Total Number of Samples	**249**

Identical experiments for Identifiler and Identifiler Plus

Reaction Setup

Identifiler

- PCR Volume: 12.5 µL
  - Primer Mix
  - Master Mix
  - Taq Gold Polymerase
- 28 cycles
- 1 ng DNA target input unless otherwise stated

Identifiler Plus

- PCR Volume: 12.5 µL
  - Primer Mix
  - Master Mix
  - No separate Taq/enzyme added
- 28 cycles
- 1 ng DNA target input unless otherwise stated

Identifiler Plus is optimized to overcome inhibition with an improved buffer mix
Cleaner baseline and improved heterozygote peak balance

Sample Preparation for Capillary Electrophoresis

- 17.4 µL HiDi + 0.6 µL LIZ 600 v2.0 per sample
- 2 µL Sample or Allelic Ladder
- Centrifuge for 1 min to mix
- Aliquot 10 µL into a separate plate
  - Centrifuge both plates 1 min
- Plates run on 3130*xl* and 3500 simultaneously

n=16 Allelic Ladders

Size Standard Comparison

	1	2
A	LIZ 500	LIZ 600 v2.0
B	LIZ 600 v2.0	LIZ 500
C	LIZ 500	LIZ 600 v2.0
D	LIZ 600 v2.0	LIZ 500
E	LIZ 500	LIZ 600 v2.0
F	LIZ 600 v2.0	LIZ 500
G	LIZ 500	LIZ 600 v2.0
H	LIZ 600 v2.0	LIZ 500

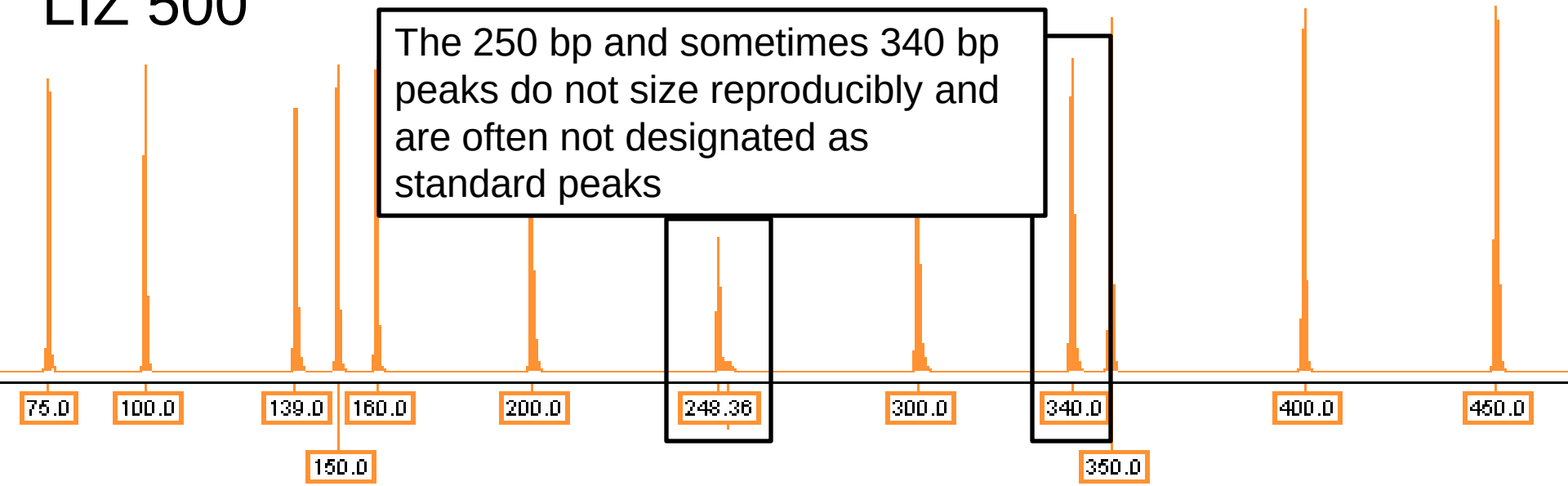
- Individual master mixes created for LIZ 500 and LIZ 600 v2.0 with Identifiler/Identifiler Plus allelic ladders
- Injected twice on 3130*xl*
  - Standard injection of 3 kV for 10 seconds
- Injected 3 times on 3500
  - Default Injection of 1.2 kV for 15 seconds

It is important to determine if one size standard can be used consistently on both the 3130xl and 3500 for proper comparison

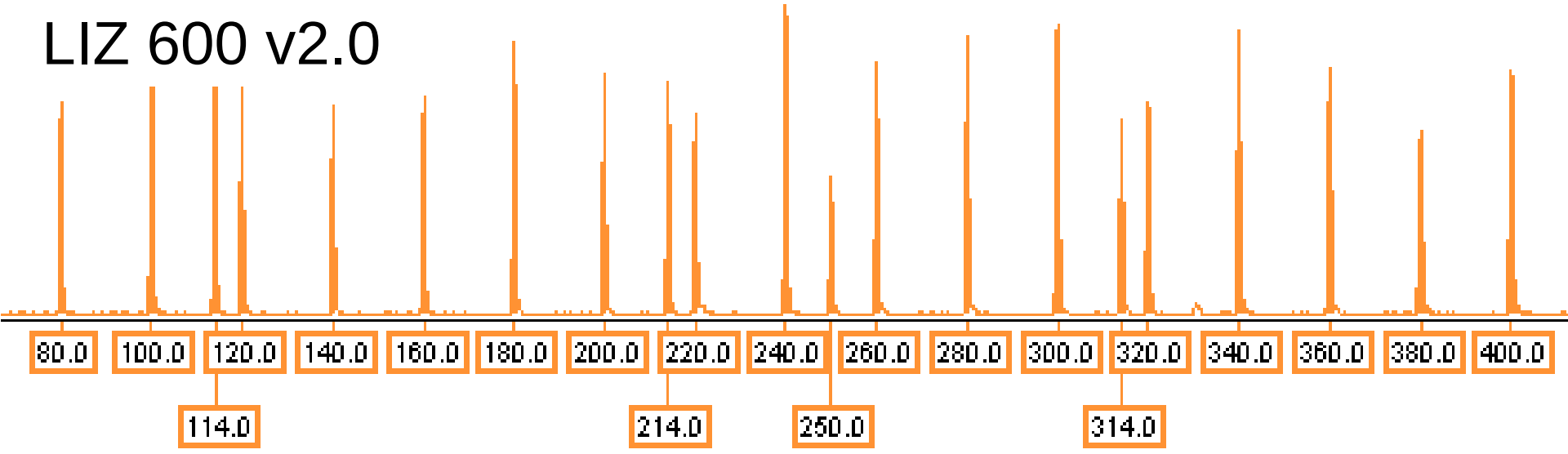
Size Standard Comparison

LIZ 500

The 250 bp and sometimes 340 bp peaks do not size reproducibly and are often not designated as standard peaks

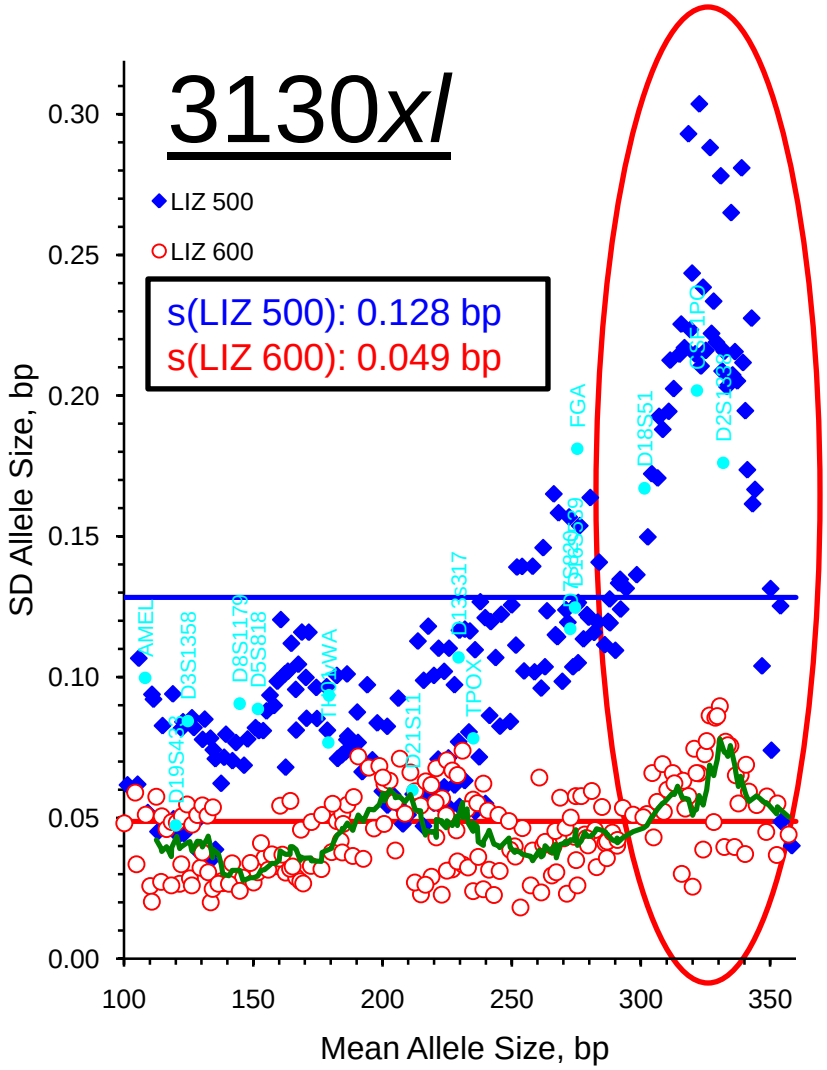


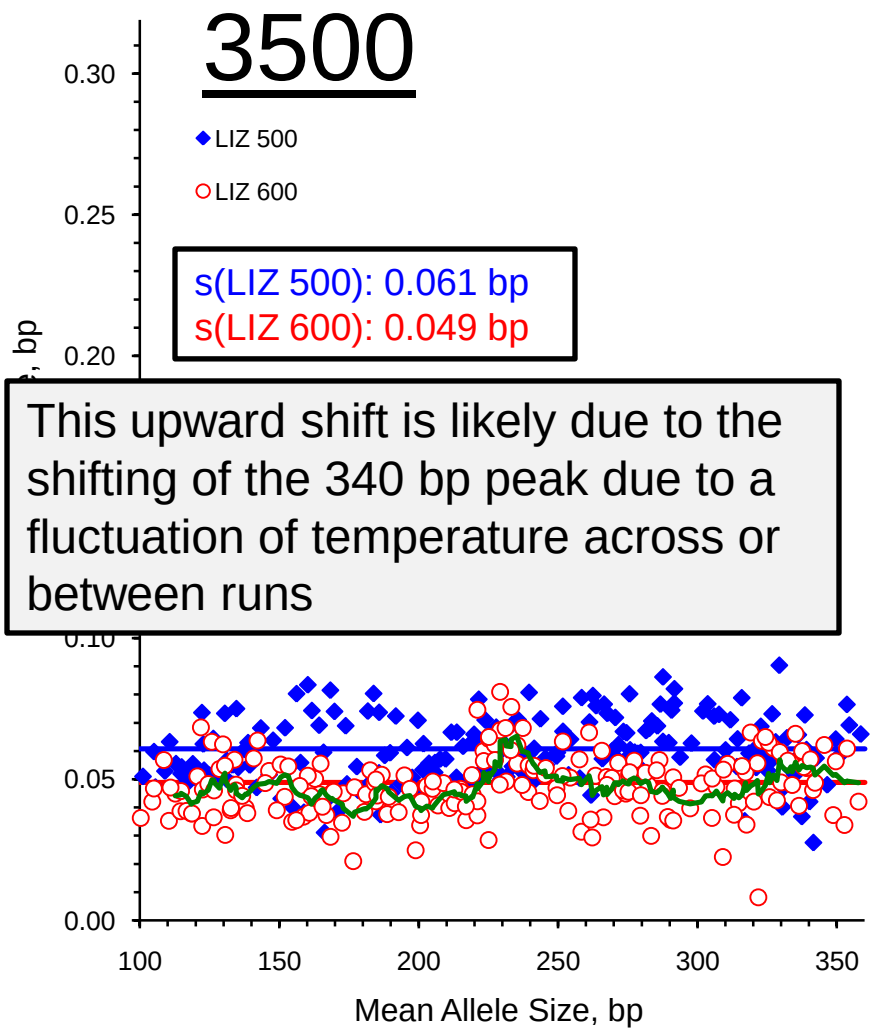
LIZ 600 v2.0



n=16 Allelic Ladders

Size Standard Comparison





LIZ 600 v2.0 generated the most linear results on both the 3130*xl* and 3500 and was used as the size standard on both instruments for remaining testing

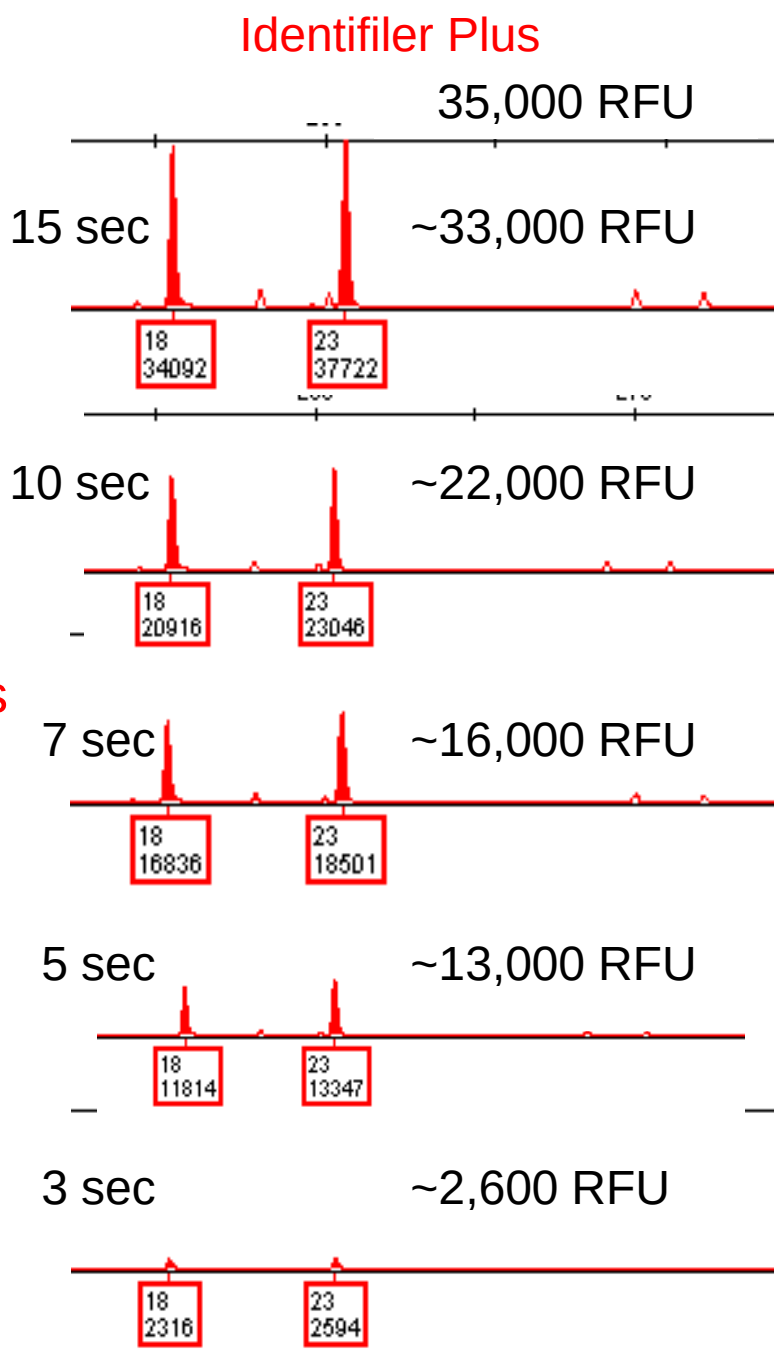
n=20: Identifiler n=15: Identifiler Plus

Injection Parameters

- Injection voltage/time:
  - 1.2 kV for 15 sec
  - 1.2 kV for 10 sec
  - 1.2 kV for 7 sec → Identifiler
  - 1.2 kV for 5 sec → Identifiler Plus
  - 1.2 kV for 3 sec

Standard injection parameters set based on samples with:

1. No pull-up present
2. No drop out present



n=24 Allelic Ladders, n=6 Samples

Sizing Precision

	Identifiler		Identifiler Plus	
	1	2	3	4
A	Identifiler	EB	Identifiler Plus	EB
B	Neg	Identifiler	Neg	Identifiler Plus
C	Identifiler	EB	Identifiler Plus	EB
D	Neg	Identifiler	Neg	Identifiler Plus
E	Identifiler	EB	Identifiler Plus	EB
F	Neg	Identifiler	Neg	Identifiler Plus
G	Identifiler	Sample	Identifiler Plus	Sample
H	Sample	Identifiler	Sample	Identifiler Plus

Identifiler and Identifiler Plus allelic ladders in checkerboard pattern

Neg: PCR blank
- PCR primers + water

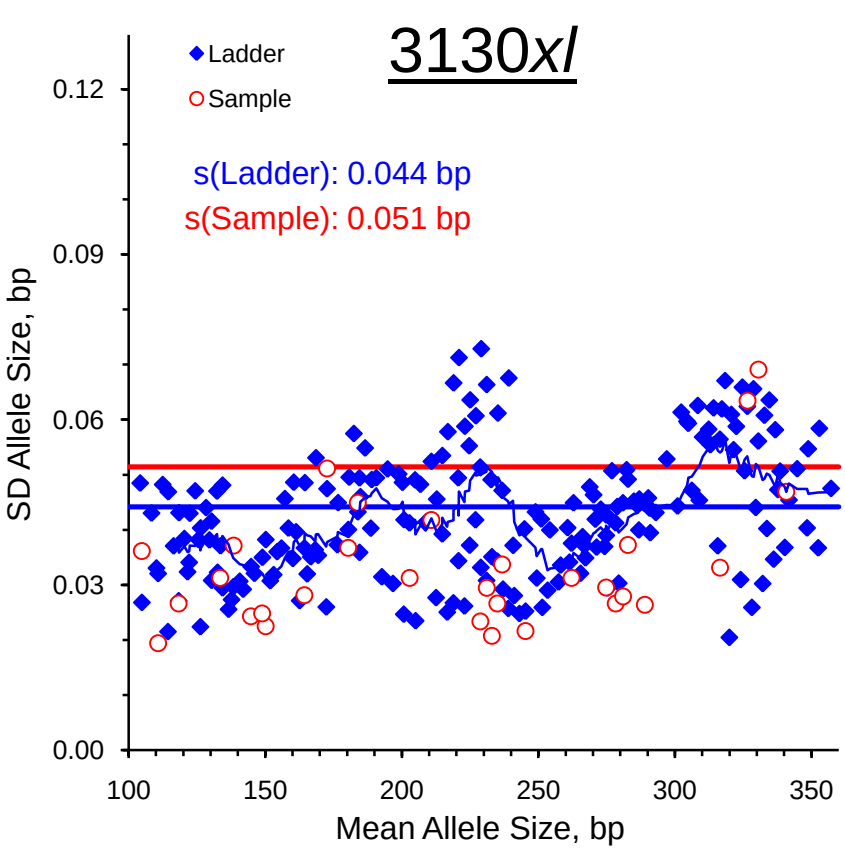
EB: Extraction blank
- PCR primers + extraction eluent

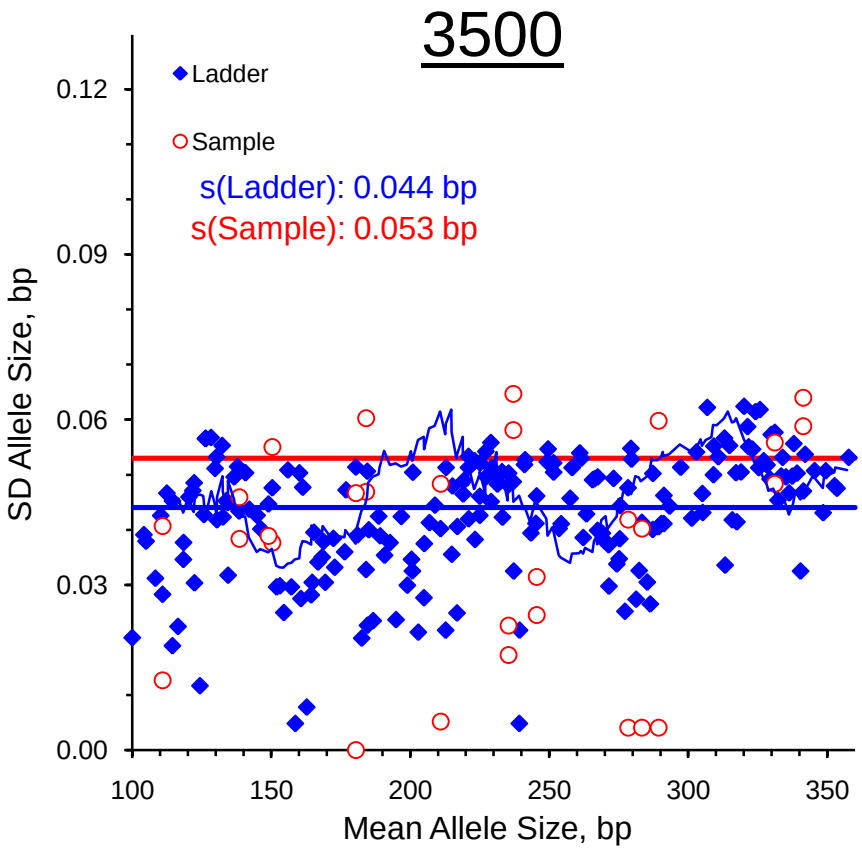
Sample: 1 ng heterozygous sample at 15 loci plus Amelogenin

Injected 3 times with the newly determined injection parameters

n=24: Allelic Ladders, n=6 Samples

Precision of Base Pair Sizing





No significant difference between 3130*xl* and 3500

No significant difference between Identifiler and Identifiler Plus

Setting Analytical Thresholds

- Analytical Threshold (AT)
  - Minimum threshold for data comparison and peak detection in the DNA typing process

Butler, J.M. (2009) *Fundamentals of Forensic DNA Typing*. Elsevier Academic Press: San Diego

- AT values can be calculated using negative controls
  - Analyze with threshold set at 1 RFU
  - Calculate average RFU noise per dye channel
- AT values can be calculated using a DNA dilution series
  - Analyze with threshold at 1 RFU
  - Remove calls for all alleles and artifacts (stutter, n+4, pull-up, etc)

Methods For Calculation

- Method 1 (LOD): Average RFU + (3 x Standard Deviation)
- Method 2: Average RFU + ($t_{1-\alpha,v}$ from student t-table x Standard Deviation)
- Method 3: 2 x ($Y_{max} - Y_{min}$)
- Method 4 (LOQ): Average RFU + (10 x Standard Deviation)

Calculations Using Negative Controls

Identifiler								
	Average RFU	**Stdev**	**Min RFU**	**Max RFU**	**Method 1**	**Method 2**	**Method 3**	**Method 4**
Blue	9	3.3	2	22	19	19	44	42
Green	13	3.6	5	27	24	23	54	49
Yellow	20	4.9	8	31	35	34	62	69
Red	27	7.1	10	50	49	48	100	99

Identifiler Plus								
	Average RFU	**Stdev**	**Min RFU**	**Max RFU**	**Method 1**	**Method 2**	**Method 3**	**Method 4**
Blue	9	3.1	3	20	18	18	40	39
Green	13	3.4	4	26	23	23	52	47
Yellow	20	5.1	7	37	36	35	74	72
Red	28	7.2	11	54	49	48	108	99

If calculating analytical threshold using negative controls:
Identifiler: 100 RFU

If calculating analytical threshold using negative controls:
Identifiler Plus: 100 RFU

Calculations Using DNA Dilution Series

Identifiler								
	Average RFU	**Stdev**	**Min RFU**	**Max RFU**	**Method 1**	**Method 2**	**Method 3**	**Method 4**
Blue	9	8.4	1	66	34	33	132	93
Green	13	11.5	3	84	48	47	168	128
Yellow	22	11.6	4	88	57	56	176	138
Red	28	8.8	10	80	54	53	160	116

Identifiler Plus								
	Average RFU	**Stdev**	**Min RFU**	**Max RFU**	**Method 1**	**Method 2**	**Method 3**	**Method 4**
Blue	10	4.6	3	68	23	23	136	55
Green	16	5.6	3	78	33	32	156	72
Yellow	24	7.9	7	63	48	47	126	103
Red	31	8.9	7	81	57	56	162	120

If calculating analytical threshold using a DNA dilution series
Identifiler: 140 RFU

If calculating analytical threshold using a DNA dilution series
Identifiler Plus: 120 RFU

Single Analytical Threshold Summary

	Negative Controls	Positive Controls
Identifier	100 RFU	140 RFU
Identifiler Plus	100 RFU	120 RFU

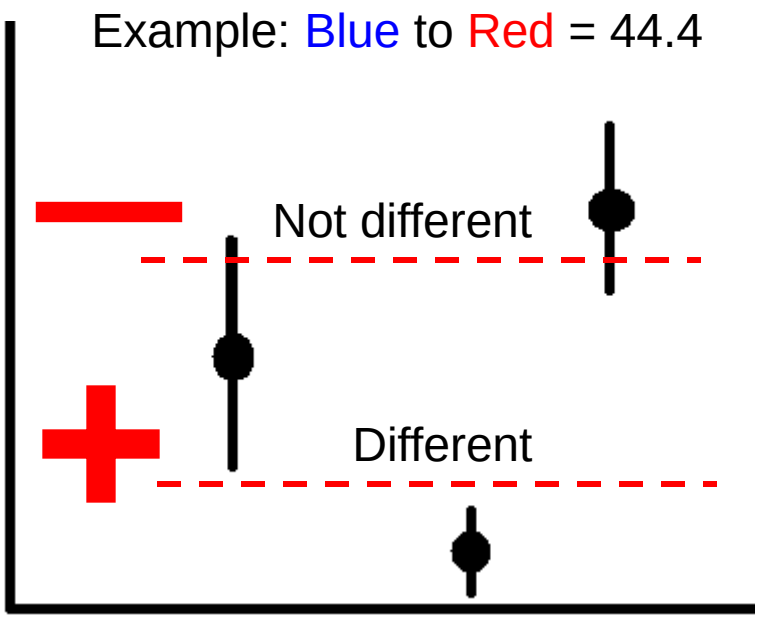
One Threshold vs. Dye Specific Thresholds

- Evaluation of Method 4 of the DNA dilution series data to determine the statistical difference between dye channel analytical thresholds

- Calculated statistical difference using a z-test

  - If negative: **Not** statistically different
    - Error bars overlap
    - One standard analytical threshold can be applied to all dyes

  - If positive: Statistically different
    - Error bars do not overlap
    - Dye specific analytical thresholds need to be applied



Dye Specific Thresholds

- Statistically each dye channel is different for both Identifiler and Identifiler Plus
  - Need to be treated independently

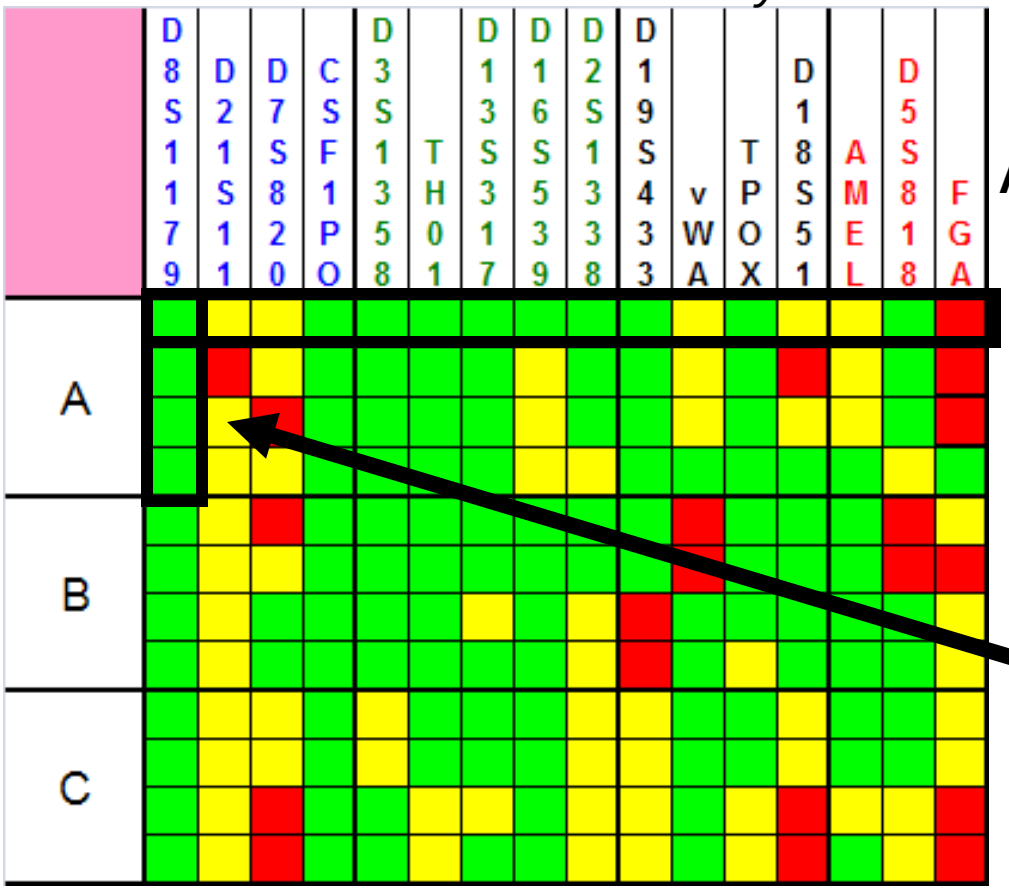
Identifiler		
	Method 4 (RFU)	AT (RFU)
Blue	93	95
Green	128	130
Yellow	138	140
Red	116	120

Identifiler Plus		
	Method 4 (RFU)	AT (RFU)
Blue	55	55
Green	72	75
Yellow	103	105
Red	120	120

Heat Map Explanation

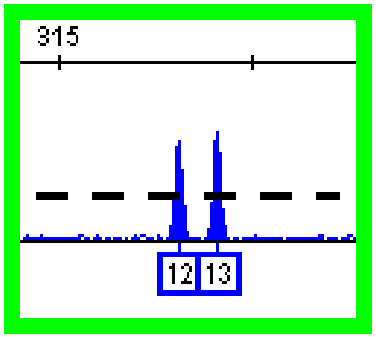
Green = full (correct) type
Yellow = allele dropout
Red = locus dropout

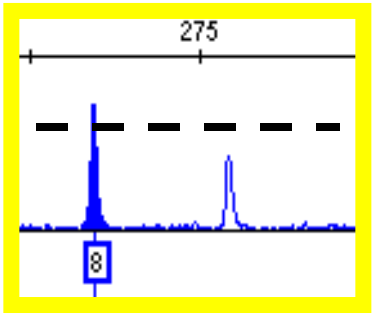
Results broken down by locus

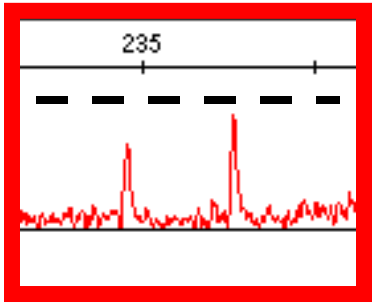


A single profile slice

A replicate slice





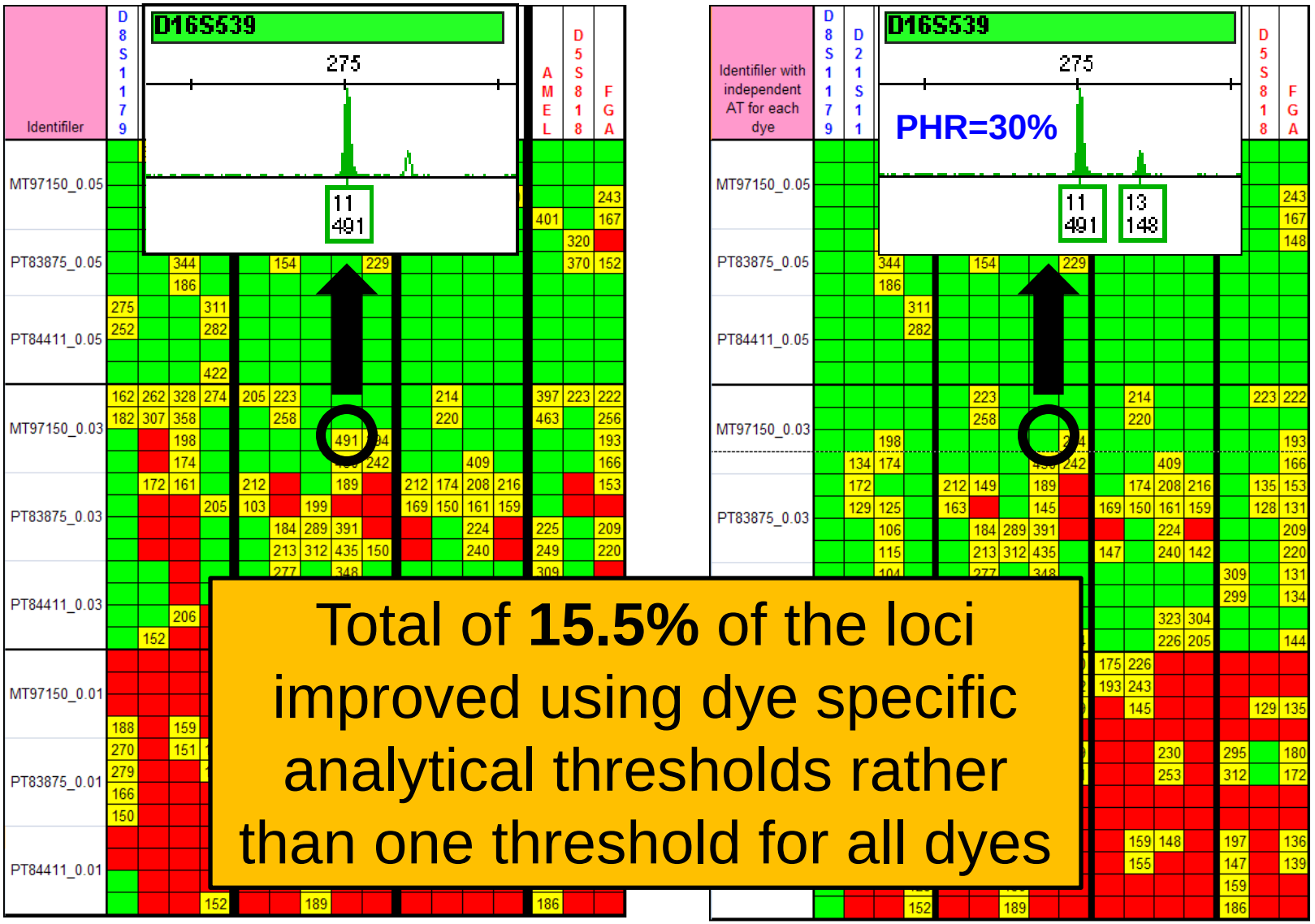


This is an easy way to look at a lot of data at once

Heat Map Comparison: Identifiler

One Threshold

Dye Specific Thresholds

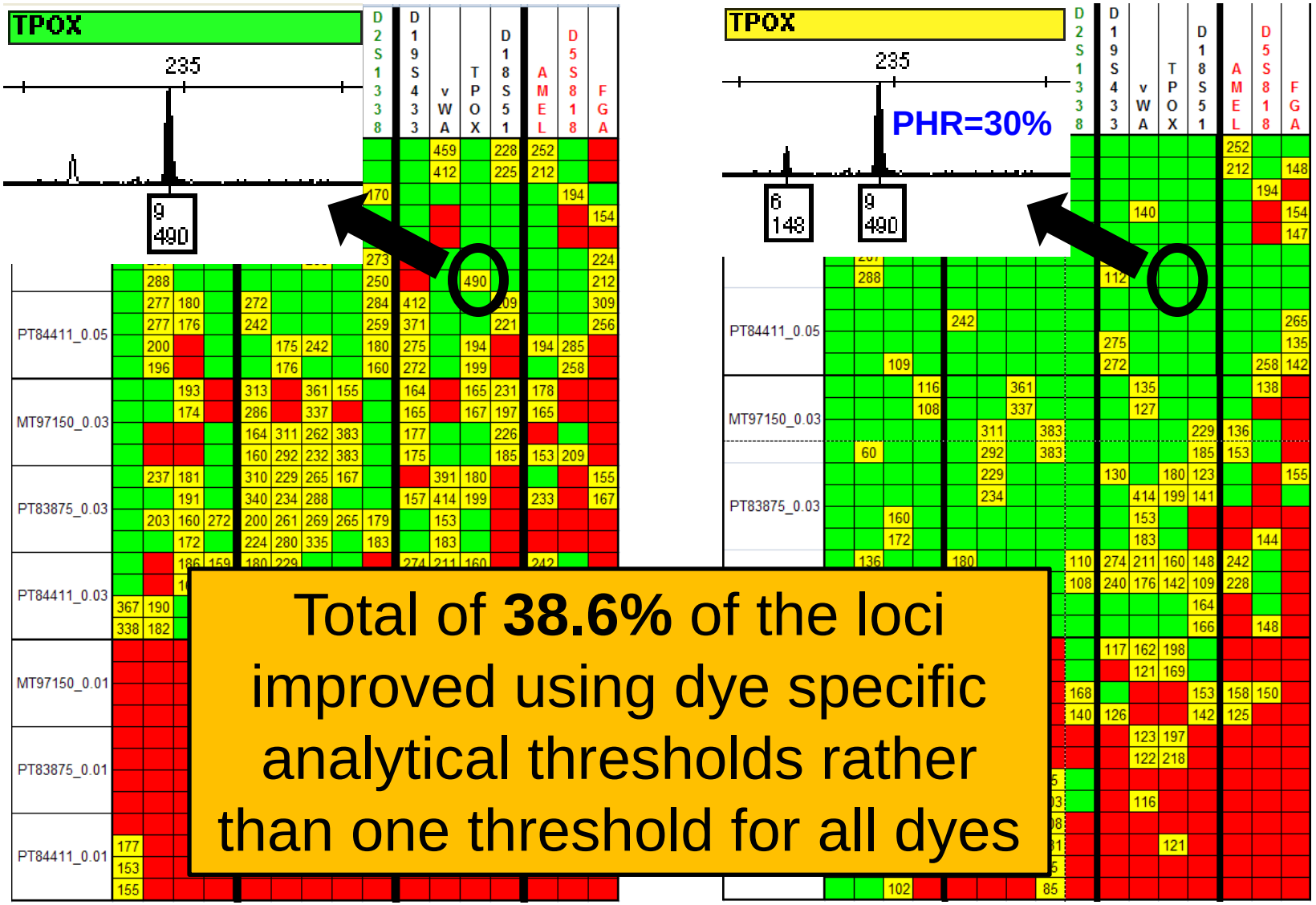


Total of **15.5%** of the loci improved using dye specific analytical thresholds rather than one threshold for all dyes

Heat Map Comparison: Identifiler Plus

One Threshold

Dye Specific Thresholds



Setting Stochastic Thresholds

Stochastic Threshold (ST)

- Detection level on an instrument (31xx or 3500) where a potential sister allele of detected peak may fall below the analytical threshold
- The value above which it is reasonable to assume that allelic dropout of a sister allele has not occurred

Butler, J.M. (2009) *Fundamentals of Forensic DNA Typing*. Elsevier Academic Press: San Diego

Setting Stochastic Thresholds

- Dilution series of three heterozygous samples at 15 loci plus Amelogenin to evaluate where drop out is first observed

- Total DNA input: 1.0 ng, 0.5 ng, 0.25 ng, 0.10 ng, 50 pg, 30 pg,10 pg in 4 replicates

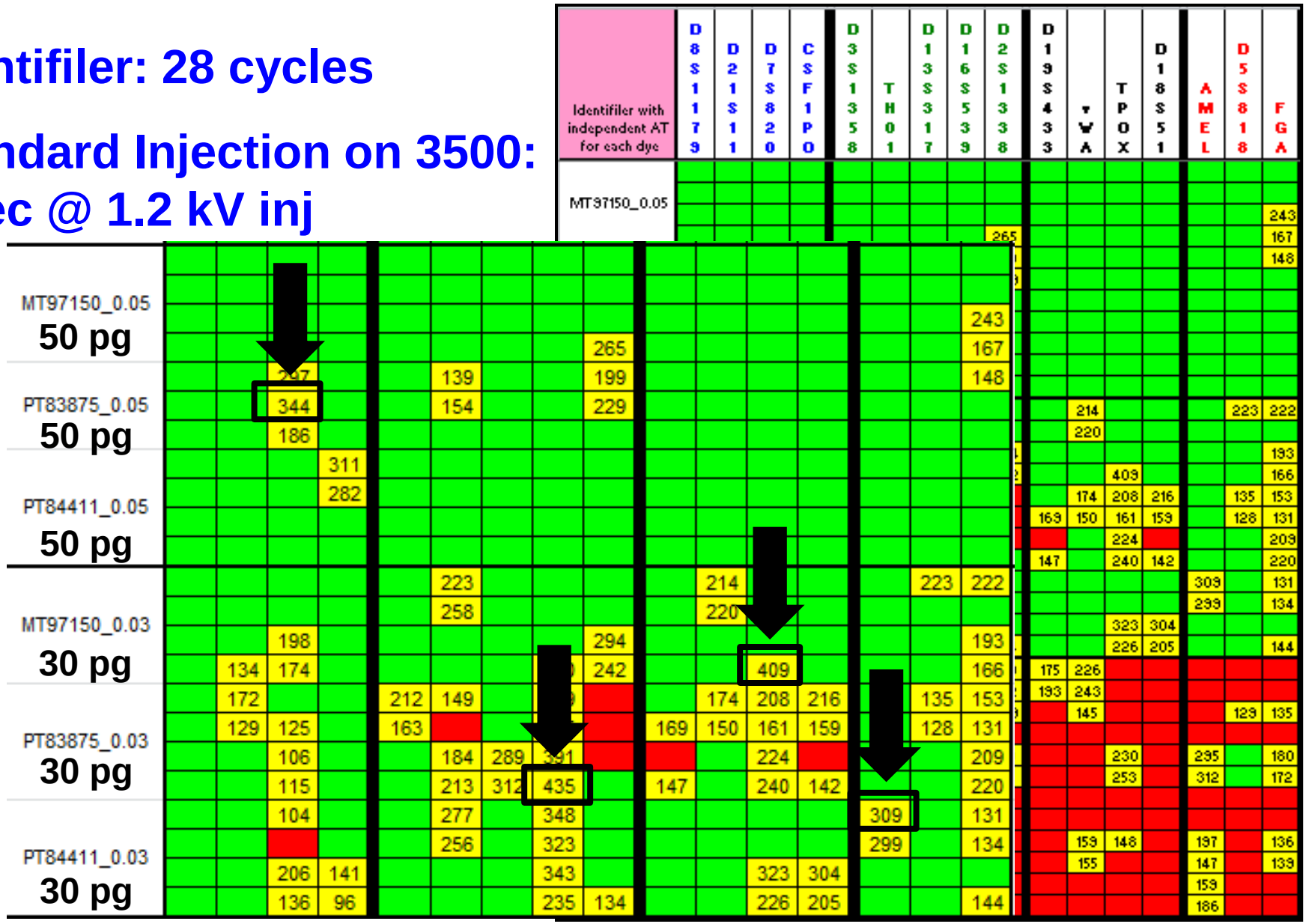
- Determine RFU value of highest surviving false homozygous peak per dye channel

n=84 Samples

Stochastic Threshold

Identifiler: 28 cycles

Standard Injection on 3500:
7 sec @ 1.2 kV inj

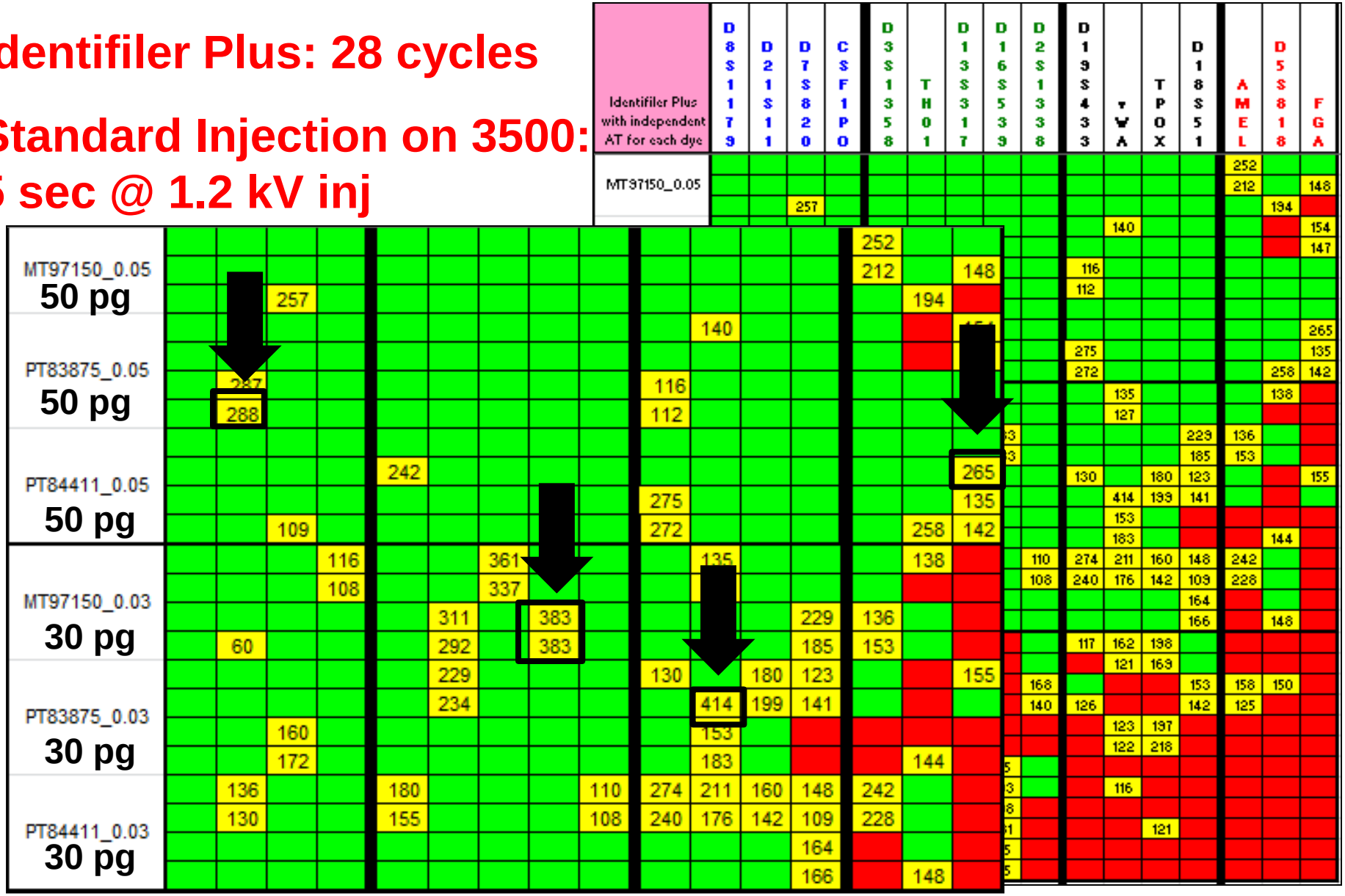


n=84 Samples

Stochastic Threshold

Identifiler Plus: 28 cycles

Standard Injection on 3500: 5 sec @ 1.2 kV inj



n=84 Samples

Stochastic Thresholds

Identifiler: 28 cycles

Standard Injection on 3500:
7 sec @ 1.2 kV inj

Identifiler	
	ST
Blue	345
Green	435
Yellow	410
Red	310

Identifiler Plus: 28 cycles

Standard Injection on 3500:
5 sec @ 1.2 kV inj

Identifiler Plus	
	ST
Blue	290
Green	385
Yellow	415
Red	265

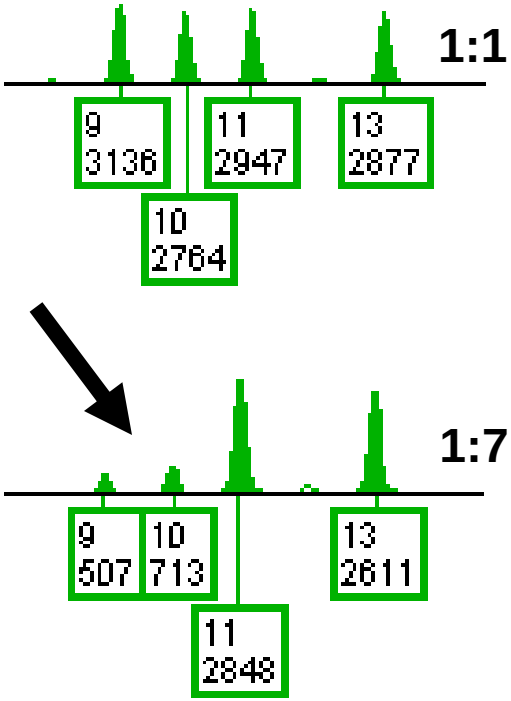
Concordance

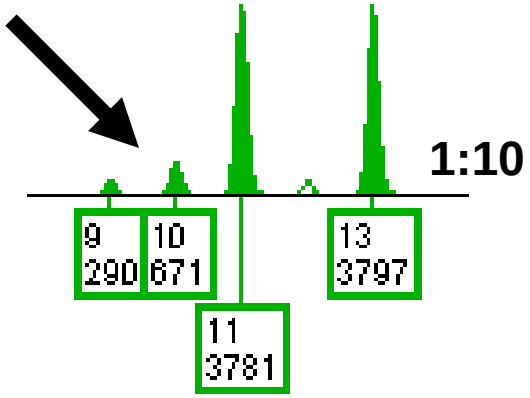
- 50 unique male samples
- SRM 2391b: 10 genomic DNA Samples
- All 60 samples concordant between 3130*xl* and 3500
- Total of 1689 alleles examined

n=28 Mixtures

Mixture Experimental Design

	1	2	3	4	5	6	7	8
A	Ladder	Ladder	Ladder	Ladder	Ladder	Ladder	Ladder	Ladder
B	1:1	1:1	1:1	1:1	1:1	1:1	1:1	1:1
C	2:1	2:1	1:2	1:2	2:1	2:1	1:2	1:2
D	3:1	3:1	1:3	1:3	3:1	3:1	1:3	1:3
E	5:1	5:1	1:5	1:5	5:1	5:1	1:5	1:5
F	7:1	7:1	1:7	1:7	7:1	7:1	1:7	1:7
G	9:1	9:1	1:9	1:9	9:1	9:1	1:9	1:9
H	10:1	10:1	1:10	1:10	10:1	10:1	1:10	1:10
	Identifiler				Identifiler Plus			





- 2 samples heterozygous at 15 loci plus Amelogenin
- Mixture ratios from 1:1 to 1:10 (and inverse)
- Samples were injected twice

Minor component identified correctly in a 1:10 mixture ratio

What is Normalization and how does it work?

Normalization of Data

- Recommended to compare signal between instruments
- Motivation mainly for large laboratories with many instruments
  - Correct for signal variation between instruments
- Can be used with a single instrument
  - Correct for signal variation between single and multiple injections

Normalization Definitions

- Normalization Target (NT)
  - Requires the use of LIZ 600 v2.0 size standard
  - Average peak heights of 11 peaks within LIZ 600 v2.0 selected for peak height consistency across lots
    - Applied within data collection software prior to running samples

- Normalization Factor (NF)
  - Adjustment needed for individual samples to reach the Normalization Target value
  - Full signal adjustment (baseline, peaks, artifacts, etc)
    - Either **increase** or **decrease** signal

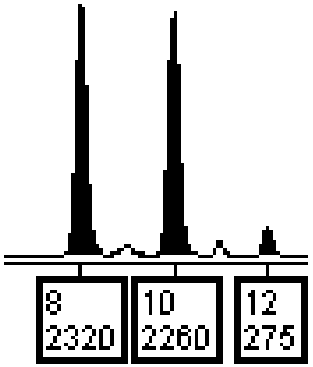
Normalization Example

Theoretical Normalization Target: 2000 RFU

Without Normalization

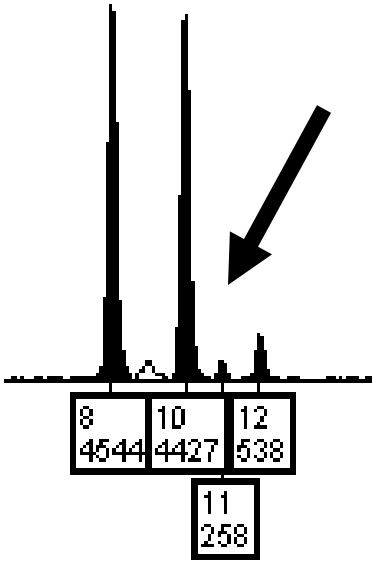
With Normalization

LIZ 600 v2.0 Peak Height Average: **1021 RFU**

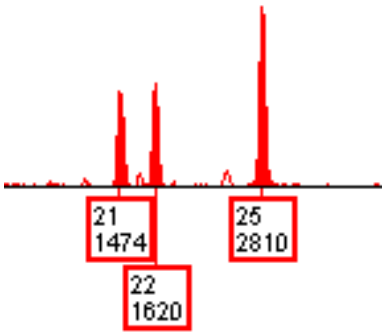


NF=1.959

Signal increases almost by a factor of 2

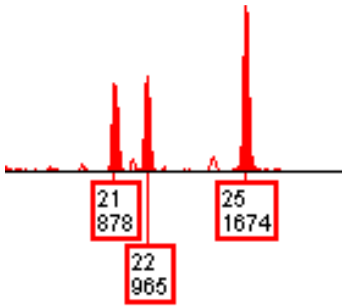


LIZ 600 v2.0 Peak Height Average: **3192 RFU**



NF= 0.596

Signal decreases by almost half



Conclusions

- The 3500 has proven to be reliable, reproducible and robust
  - Out of 498 samples between Identifiler and Identifiler Plus only 5 required reinjection
  - Precision within about 0.05 base pairs
- Dye specific analytical thresholds result in less allelic and full locus dropout than applying one analytical threshold to all dyes
- Stochastic thresholds are linked to analytical thresholds
  - If the analytical threshold is adjusted, the stochastic threshold should be reevaluated along with expected peak height ratios
- Minor contributor successfully identified in as low as a 1:10 mixture

Future Work

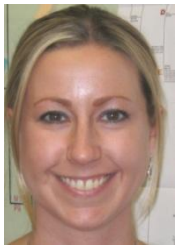
- Validation of additional kits
- More extensive review of the impact of thresholds on interpretation
  - Interaction between analytical and stochastic thresholds alongside peak height ratios
- More extensive review of normalization
  - Do thresholds change when employing normalization?

Acknowledgments

Forensic DNA Team


John Butler


Mike Coble

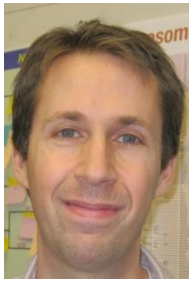

Becky Hill


Margaret Kline

Data Analysis Support


Dave Duewer

DNA Biometrics Team


Pete Vallone


Kristen Lewis O'Connor

Funding from the **National Institute of Justice (NIJ)** through NIST Office of Law Enforcement Standards

Funding from the **FBI S&T Branch** through NIST Information Access Division

Contact Info:

erica.butts@nist.gov

301-975-5107

NIJ National Institute of Justice

applied biosystems™

DEPARTMENT OF JUSTICE FEDERAL BUREAU OF INVESTIGATION