NIST Validation Studies on the 3500 Genetic Analyzer

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

Erica L.R. Butts

National Institute of Standards and Technology

Innsbruck, Austria

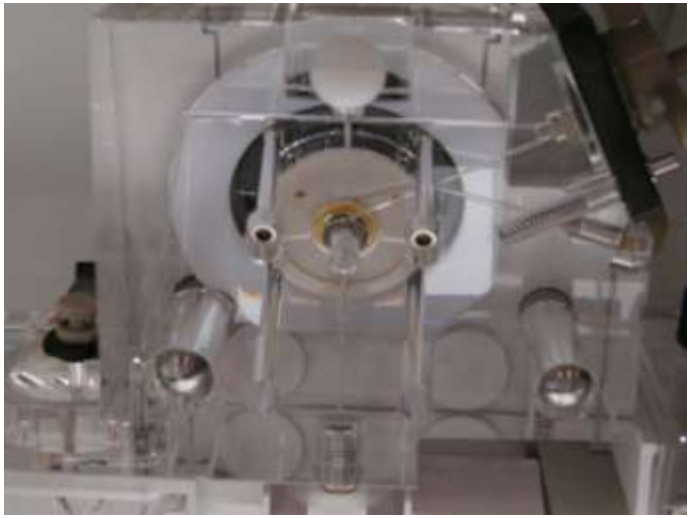
September 5th, 2011

Outline

- Details of the ABI 3500 Genetic Analyzer
- Validation design and results with Identifiler and Identifiler Plus
  - Injection parameters and reaction setup
  - Precision and size standard comparison
  - Concordance and mixture evaluation
- Methodology of setting analytical and stochastic thresholds
- Brief overview of signal normalization

Details of the ABI 3500

No lower pump block
(Fewer air bubbles)



Improved sealing for better temperature control

Primary Differences

	31xx Platforms	**3500 Platforms**
Laser	Argon ion (AR+) with 488/514 nm wavelength	Single-line 505 nm, solid-state, long-life laser
Power Requirement	220V	110V
File Generated	.fsa files	.hid files
Normalization	None	Instrument-to-instrument; only with AB kits
Optimal Signal Intensity	1500-3000 RFU	4x greater than 31xx platforms

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

SWGDAM: FBI Laboratory's Scientific Working Group on DNA Analysis Methods

Experimental Summary

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

Identical experiments for Identifiler and Identifiler Plus

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

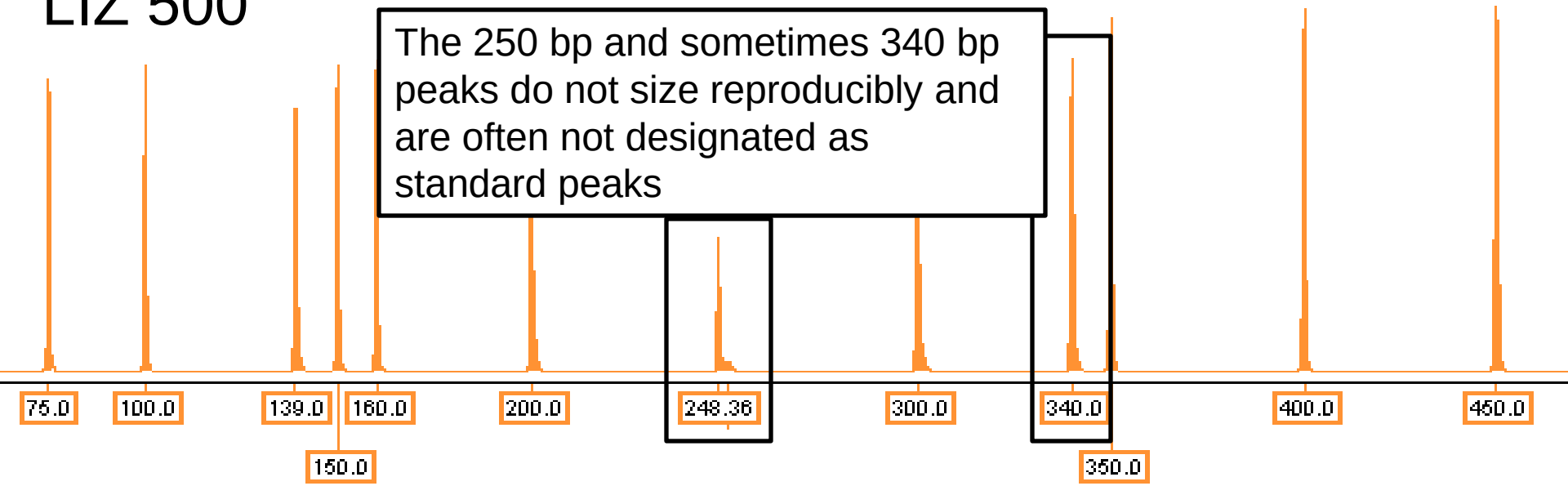
- Individual master mixes created for LIZ 500 and LIZ 600 v2.0 with Identifiler/Identifiler Plus allelic ladders
- Injected twice on 3130*xl*
  - Standard injection of 3 kV for 10 seconds
- Injected 3 times on 3500
  - Default Injection of 1.2 kV for 15 seconds

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

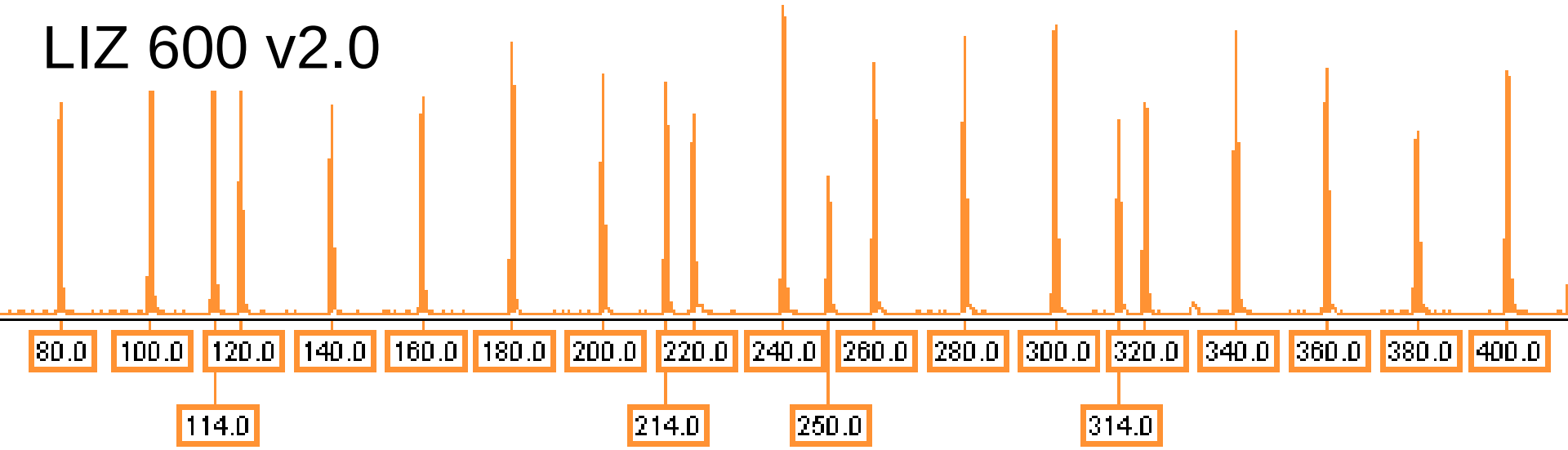
Size Standard Comparison

LIZ 500

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

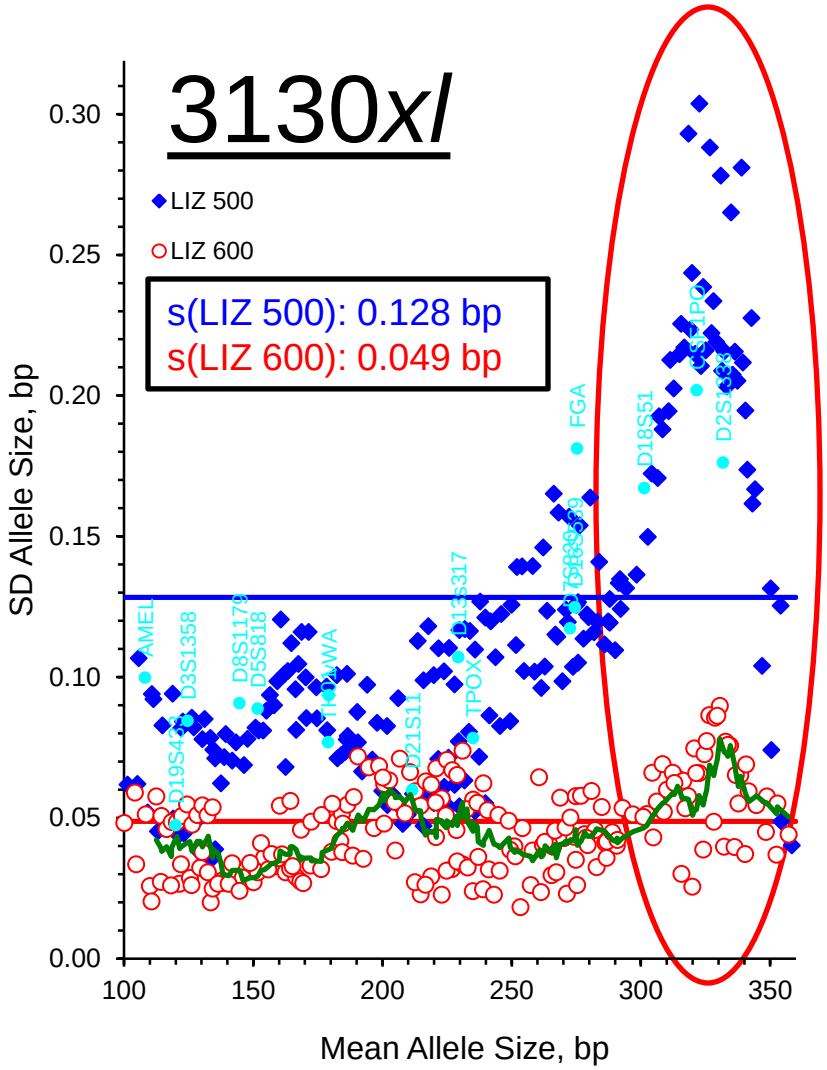


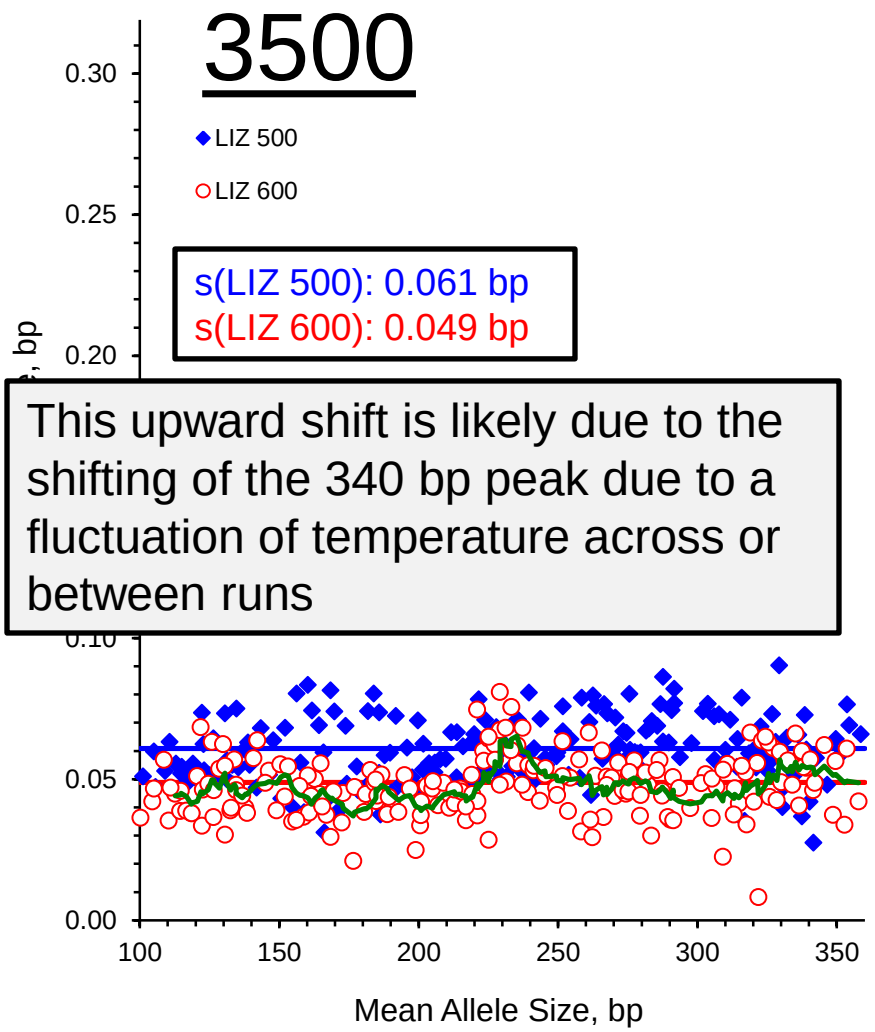
LIZ 600 v2.0



n=16 Allelic Ladders

Size Standard Comparison





This upward shift is likely due to the shifting of the 340 bp peak due to a fluctuation of temperature across or between runs

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

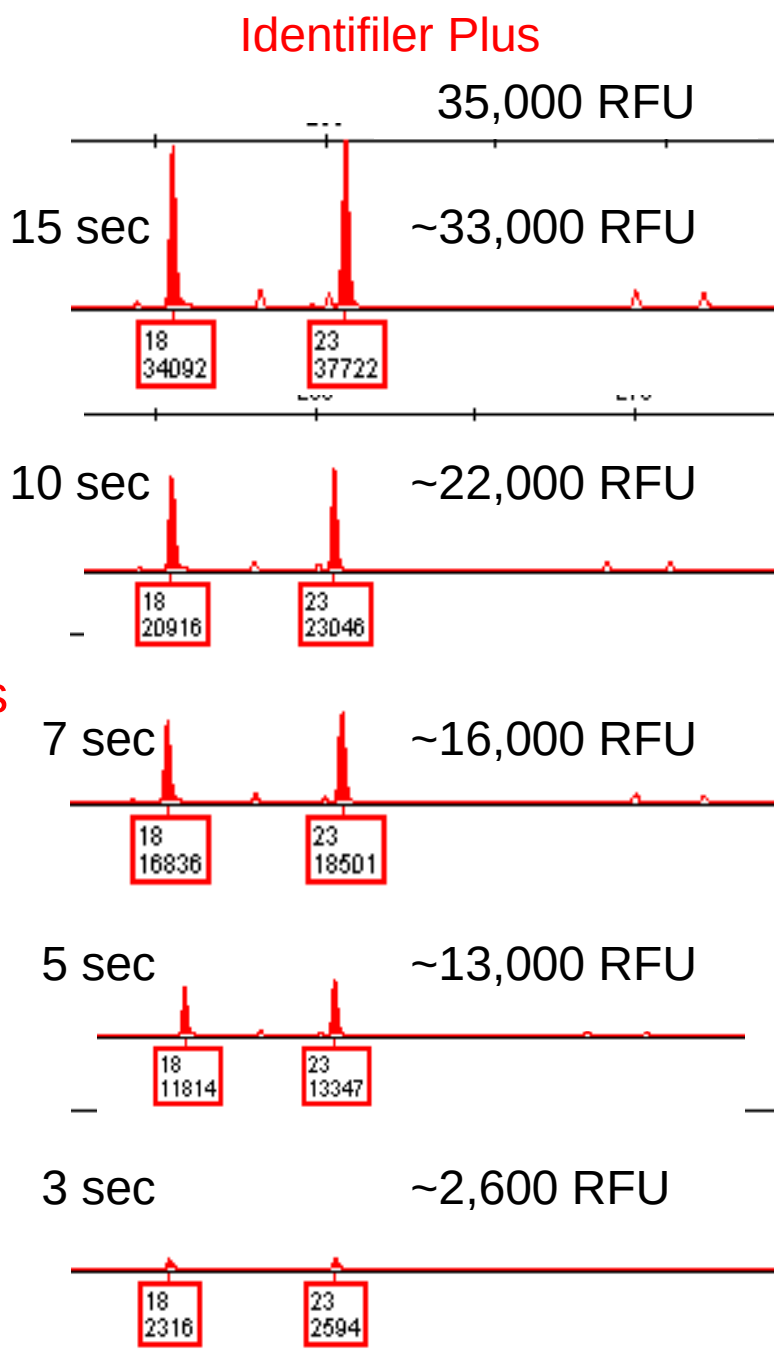
n=20: Identifiler n=15: Identifiler Plus

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

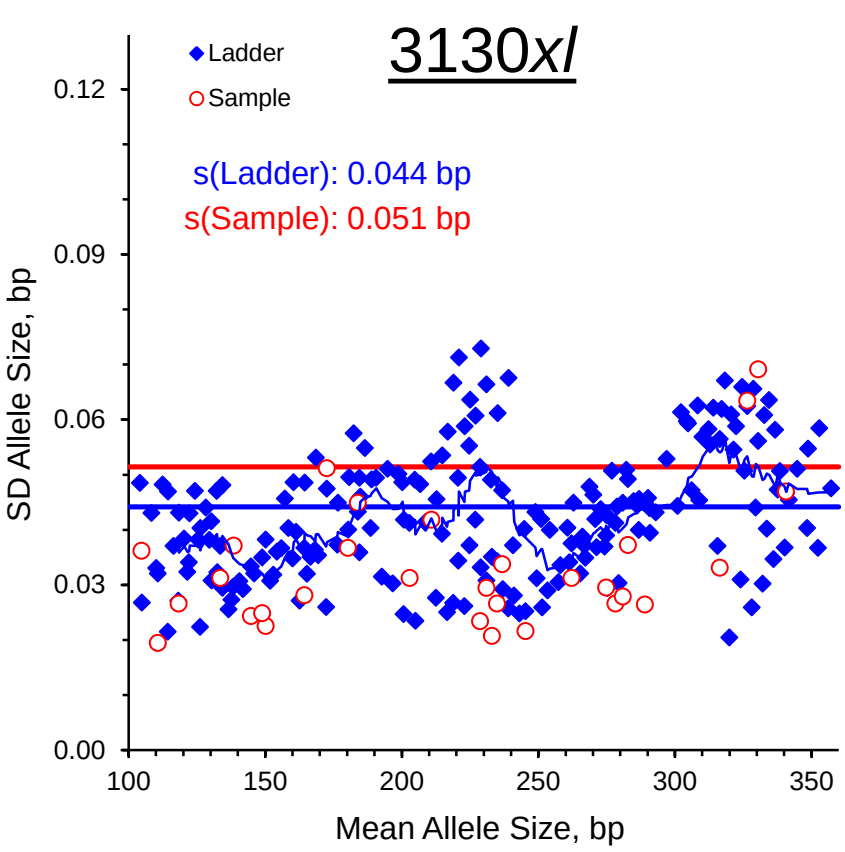
EB: Extraction blank
- PCR primers + extraction eluent

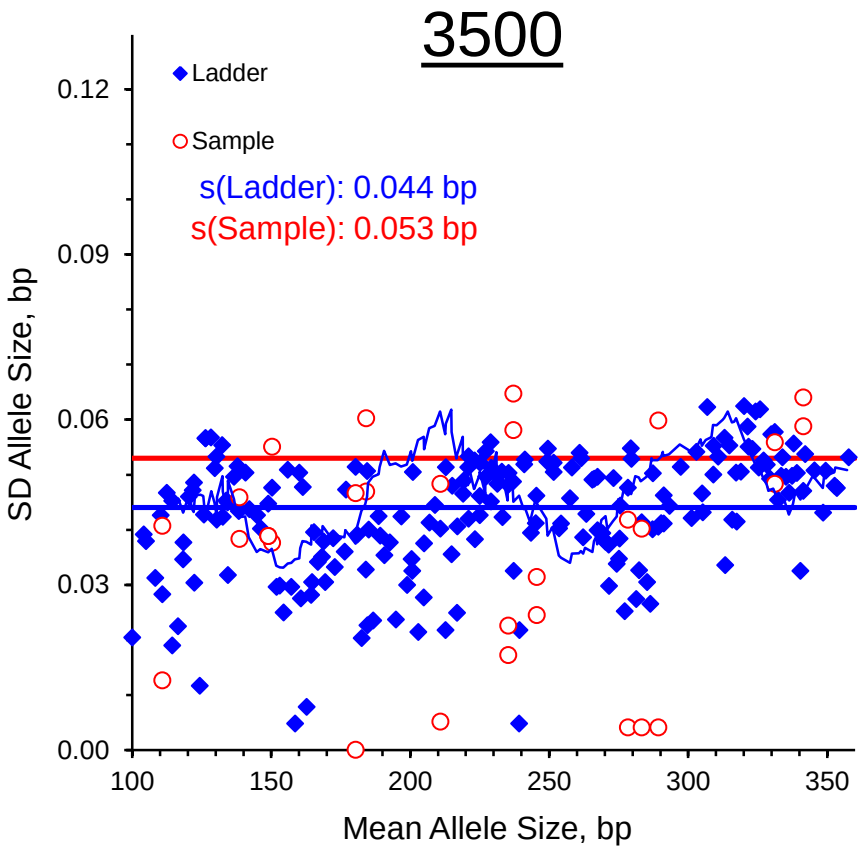
Sample: 1 ng heterozygous sample at 15 loci plus Amelogenin

Injected 3 times with the newly determined injection parameters

n=24: Allelic Ladders, n=6 Samples

Precision of Base Pair Sizing



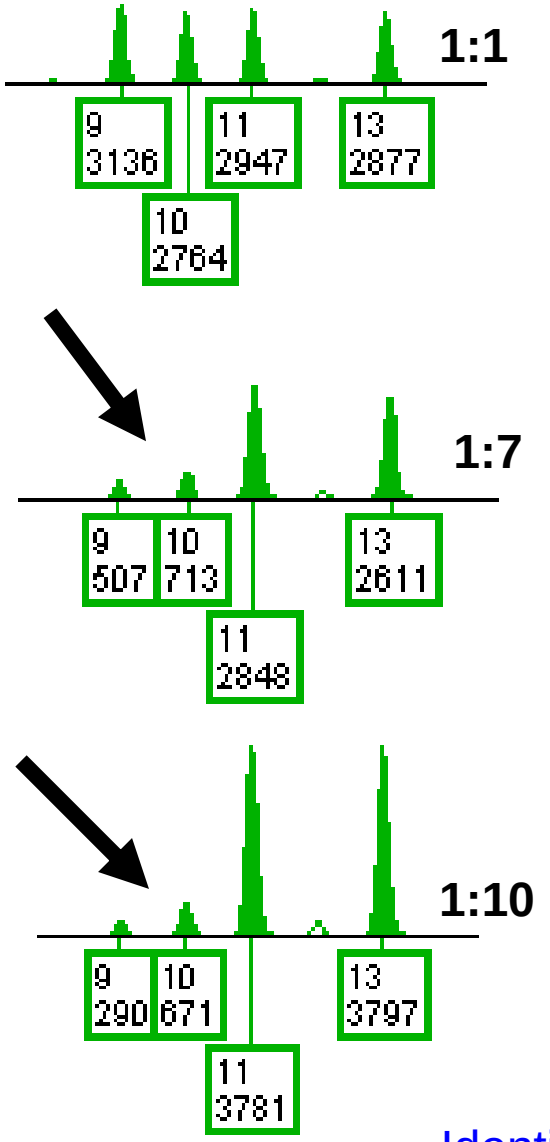


No significant difference between 3130*xl* and 3500

No significant difference between Identifiler and Identifiler Plus

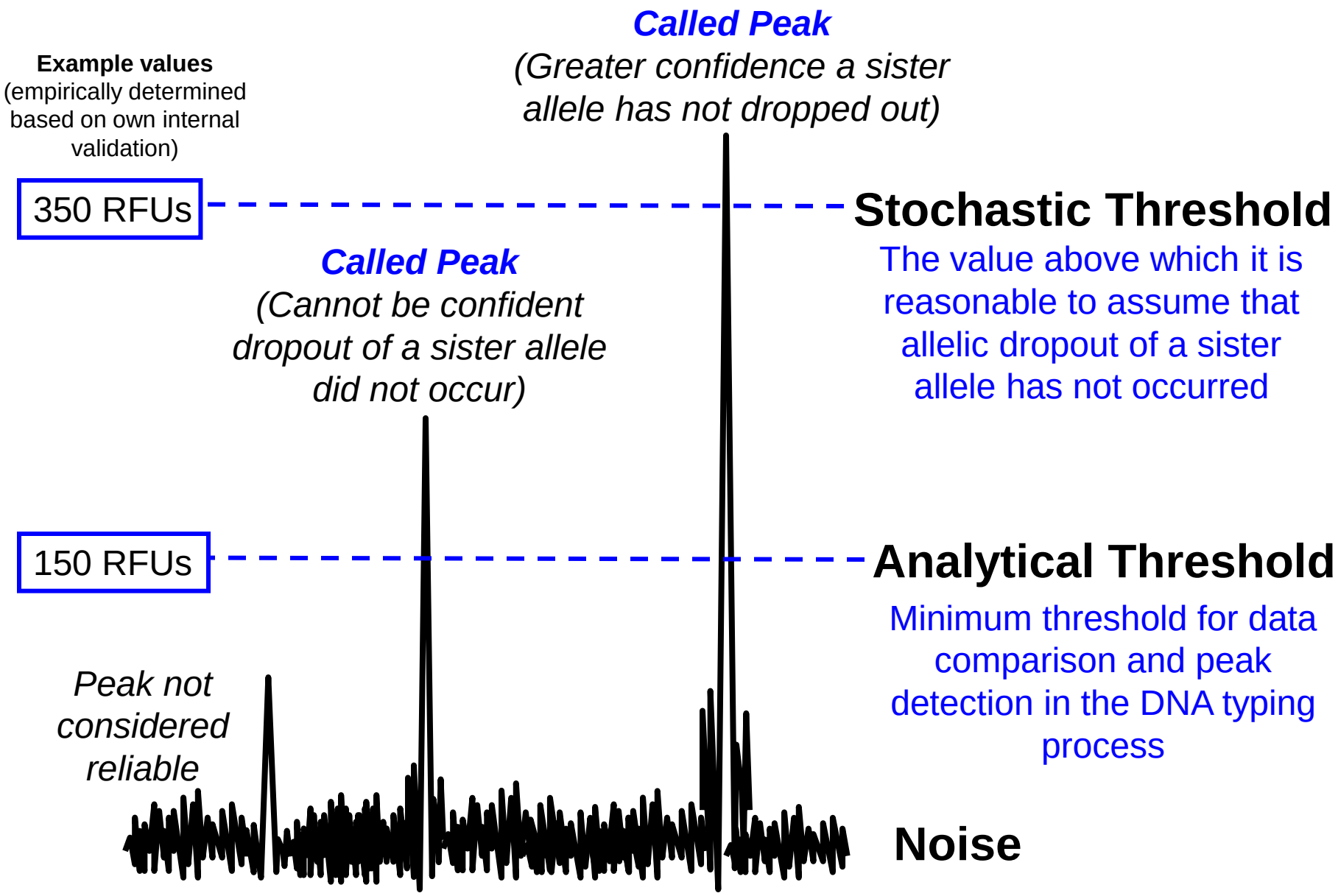
Concordance and Mixtures

- 60 samples concordant between 3130*xl* and 3500
  - Total of 1689 alleles examined
- Minor component identified correctly in a 1:10 mixture ratio



Identifiler

Different Threshold Overview



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

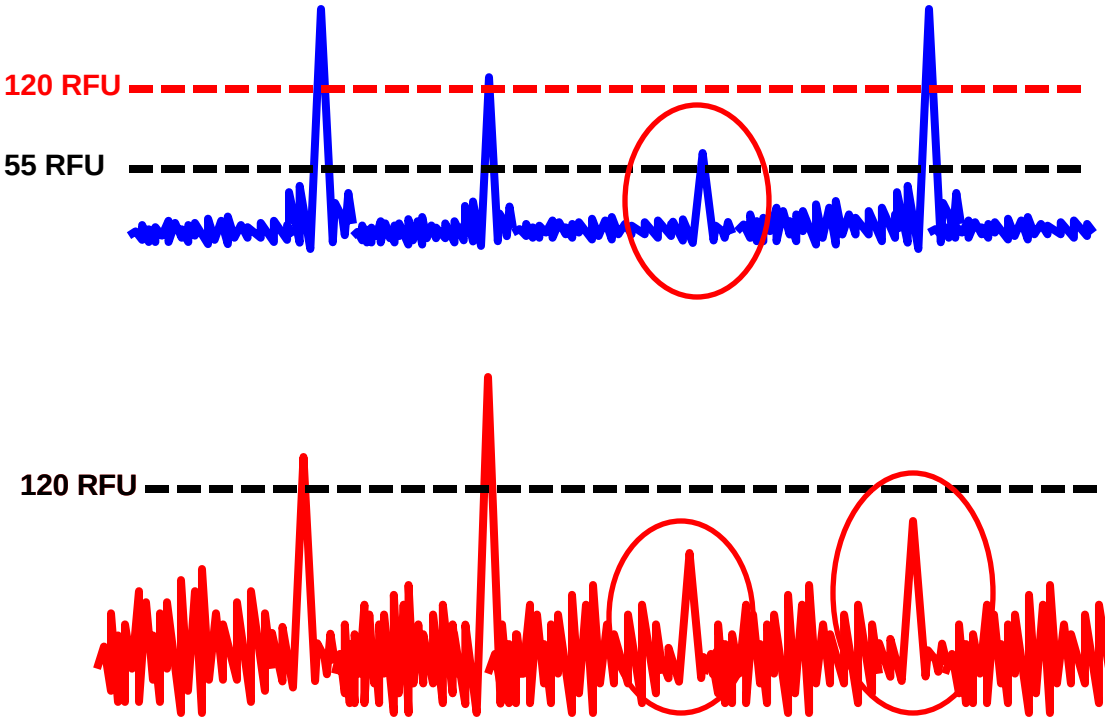
Analytical Threshold Methodology

- Baseline noise values calculated with data from the sensitivity study (DNA dilution series)
  - Threshold set at 1 RFU for all dye channels
  - Remove calls for all alleles and artifacts (stutter, n+4, pull-up, etc.)

- 4 methods for evaluation of analytical thresholds calculated

- Analytical Threshold: Average RFU + (10 x Standard Deviation)

Different Thresholds



Single thresholds for all dye channels assumes all dye channels have the **same** amount of noise

Can cause data to fall below the analytical threshold and not be called

Dye-specific thresholds take into consideration that all dye channels do not have the same level of noise

Can increase the amount of data that is callable

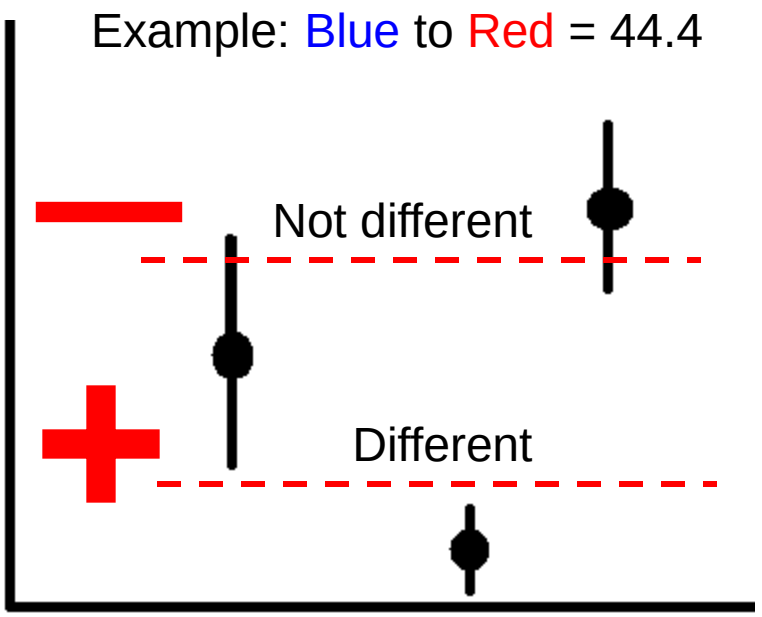
One Threshold vs. Dye Specific Thresholds

- Evaluation of 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



n=84 samples

Analytical Threshold Calculation

Identifiler					
Dye Channel	**Average RFU**	**Stdev**	**Min RFU**	**Max RFU**	**Calculated Noise (RFU)**
Blue	9	8.4	1	66	93
Green	13	11.5	3	84	128
Yellow	22	11.6	4	88	138
Red	28	8.8	10	80	116

Single Threshold: 140 RFU

Dye-Specific: Rounded to nearest 5 RFU

Identifiler Plus					
Dye Channel	**Average RFU**	**Stdev**	**Min RFU**	**Max RFU**	**Calculated Noise (RFU)**
Blue	10	4.6	3	68	55
Green	16	5.6	3	78	72
Yellow	24	7.9	7	63	103
Red	31	8.9	7	81	120

Single Threshold: 120 RFU

Dye-Specific: Rounded to nearest 5 RFU

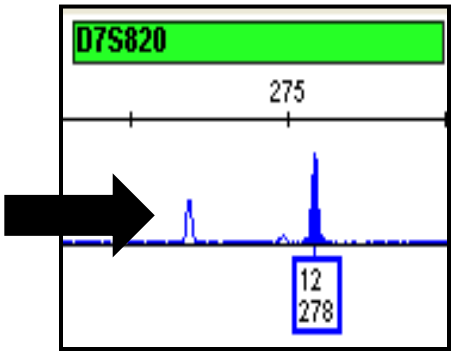
- Statistical difference was calculated between dye channels using a z-test
- Statistically each dye channel is <u>different</u> for both Identifiler and Identifiler Plus
  - Must be treated independently

n=560 alleles

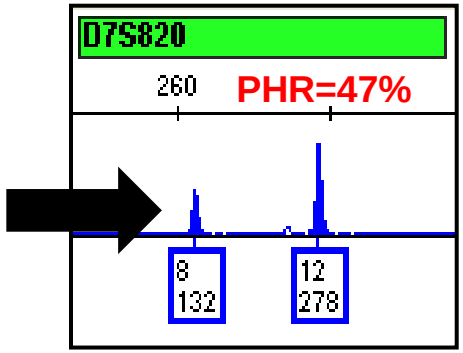
Threshold Comparison

Total of 560 alleles examined (50 pg, 30 pg, and 10 pg) where dropout was observed

Single Threshold (140 RFU)



Dye-Specific Thresholds



14.8% of the total possible allele calls were lost using a single threshold rather than using dye-specific thresholds with **Identifiler**

22.0% of the total possible allele calls were lost using a single threshold rather than using dye-specific thresholds with **Identifiler Plus**

Setting Stochastic Methodology

- Calculated with data from the sensitivity study (DNA dilution series) analyzed with dye specific analytical thresholds

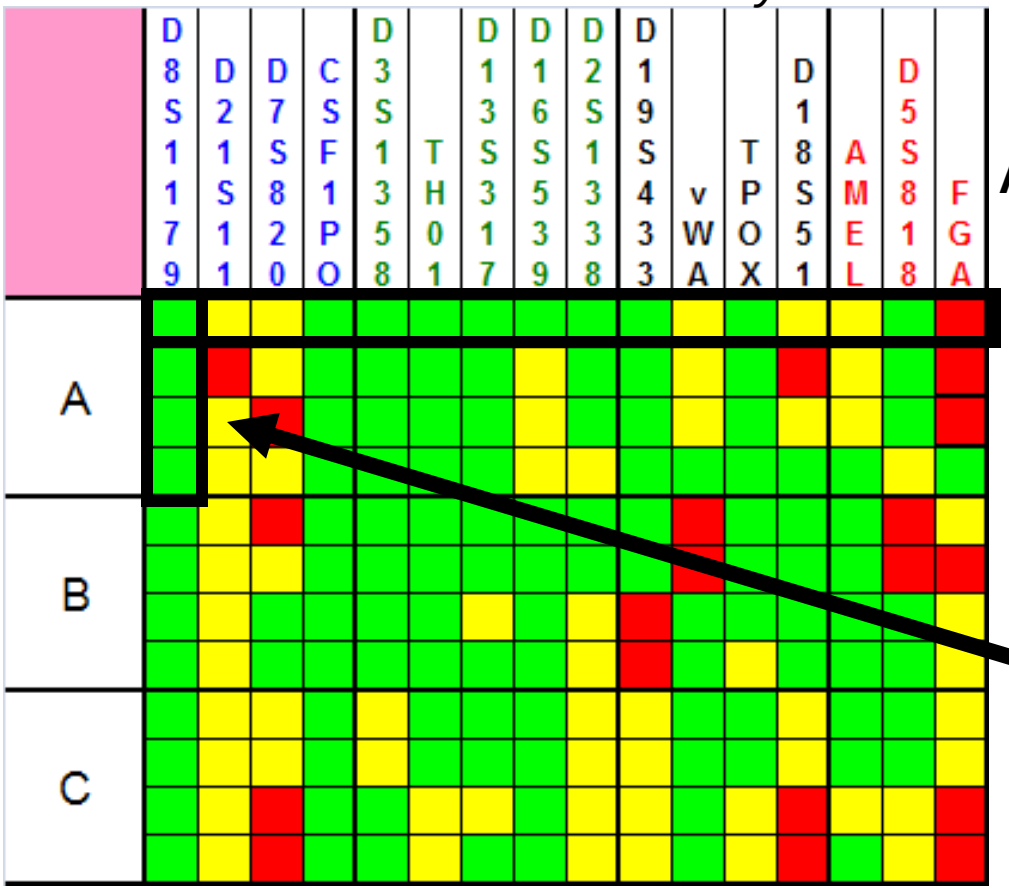
- Examination of sample amounts where dropout is observed (50 pg, 30 pg, 10 pg for Identifiler and Identifiler Plus)
  - Focus on sample amounts with dropout present to examine stochastic effects including severe imbalance of heterozygous alleles and allele dropout

- Stochastic Threshold: The RFU value of highest surviving false homozygous peak per dye channel

Heat Map Explanation

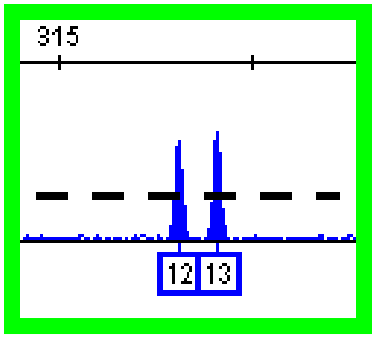
Green = full (correct) type
Yellow = allele dropout
Red = locus dropout

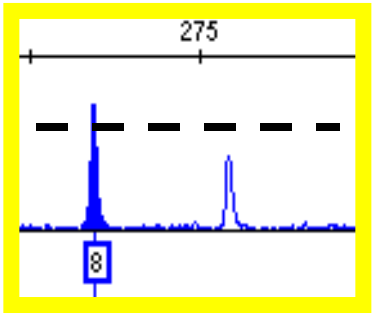
Results broken down by locus

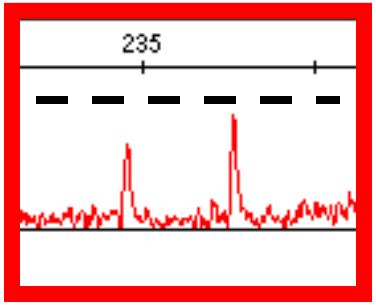


A single profile slice

A replicate slice







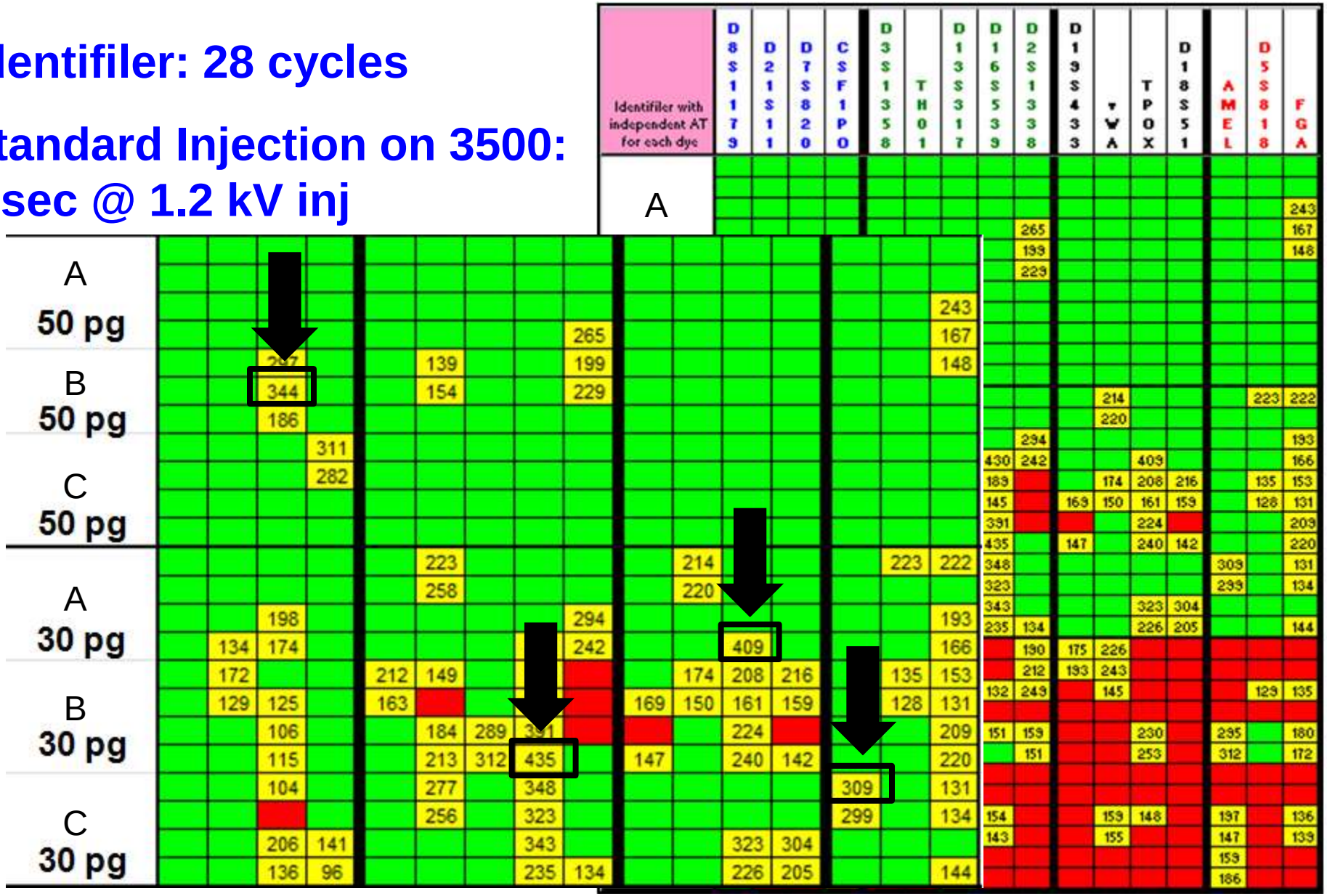
This is an easy way to look at a lot of data at once

n=84 Samples

Stochastic Threshold

Identifiler: 28 cycles

Standard Injection on 3500:
7 sec @ 1.2 kV inj

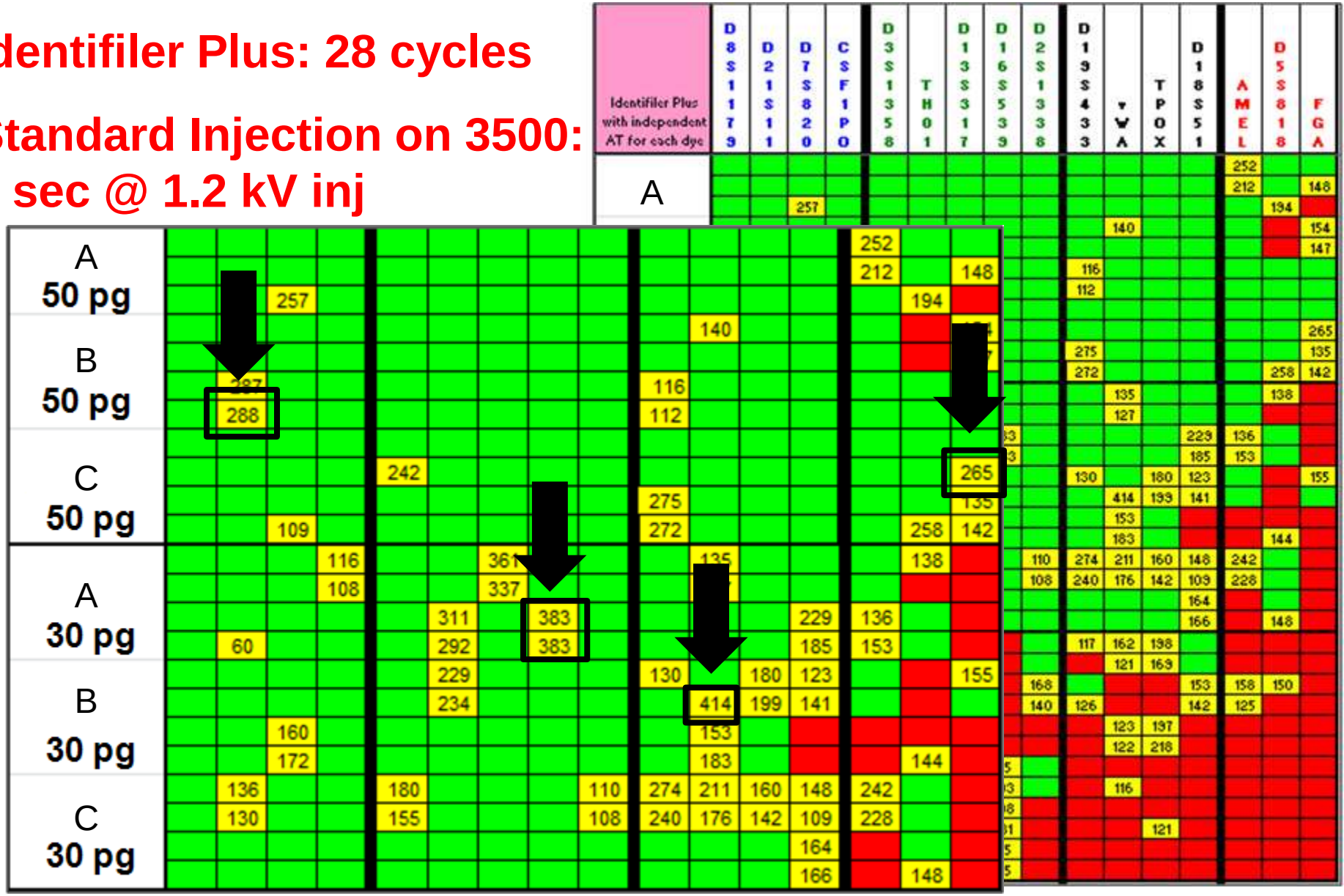


n=84 Samples

Stochastic Threshold

Identifiler Plus: 28 cycles

Standard Injection on 3500:
5 sec @ 1.2 kV inj



n=84 samples

Summary of Thresholds

Both AT and ST values rounded to the nearest 5 RFU value

Expected peak height ratio (PHR) is assuming the possibility of having one peak at the AT and one peak at the ST

Expected PHR = AT/ST

Identifiler: 7 sec @ 1.2 kV (28 cycles)				
	AT (RFU)	**Highest Surviving Peak (RFU)**	**ST (RFU)**	**Expected PHR**
Blue	95	344	345	28%
Green	130	435	435	30%
Yellow	140	409	410	34%
Red	120	309	310	39%

Identifiler Plus: 7 sec @ 1.2 kV (28 cycles)				
	AT (RFU)	**Highest Surviving Peak (RFU)**	**ST (RFU)**	**Expected PHR**
Blue	55	288	290	19%
Green	75	383	385	19%
Yellow	105	414	415	25%
Red	120	265	265	45%

Consumable RFID Tracking Limits

	RFID Hard Stops	Usage Comments From a Research Laboratory Standpoint
Array	None	1. Very easy to change between HID and sequencing 2. Array from validation was stored at least twice and reinstalled on 3500 during validation
Buffer	Expiration Date 7 Days on Instrument # Injections	1. Can no longer use in-house buffer 2. Very easy to change on the instrument (snap-and-go)
Polymer	Expiration Date # Samples # Injections	1. Hard stop with the expiration date has caused us to discard unused polymer we would have otherwise kept on the instrument 2. ~50% of total polymer remains in the pouch after "consumption" 3. Expiration dates have changed purchasing strategy (smaller batches, based on ongoing project needs)

Validation Conclusions

- The 3500 has proven to be reliable, reproducible and robust
  - Out of 498 samples between Identifiler and Identifiler Plus only 5 required reinjection
- Dye specific analytical thresholds resulted 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
    - Requires consideration for overall interpretation workflow which we are still evaluating
- RFID tracking decreases flexibility in our research experience

What is Normalization and how does it work?

Normalization of Data

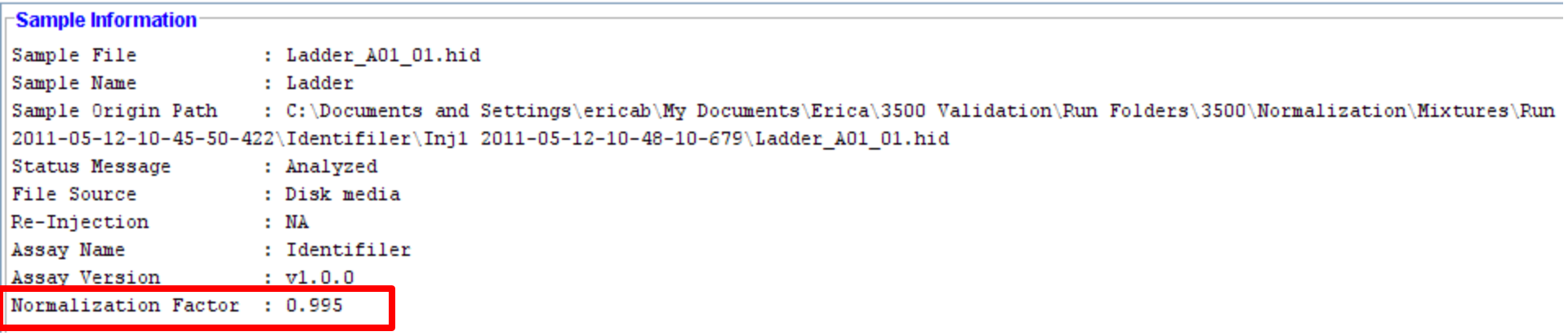
- Recommended to compare signal between instruments
- Motivation mainly for large laboratories with many instruments
  - Correct for signal variation between instruments
- Can be used with a single instrument
  - Correct for signal variation between single and multiple injections

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 Definitions

- 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



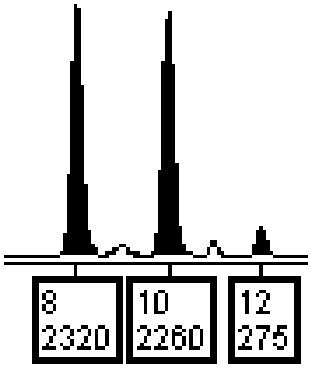
Normalization Example

Theoretical Normalization Target: 2000 RFU

Without Normalization

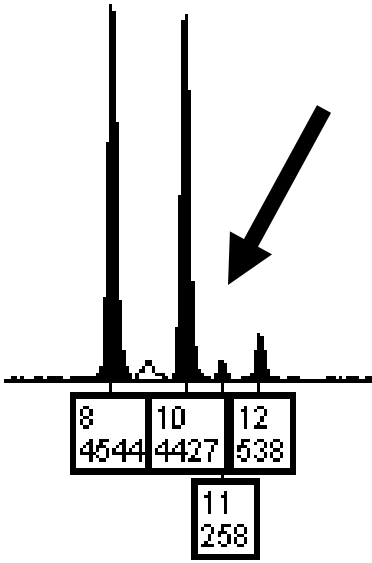
With Normalization

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

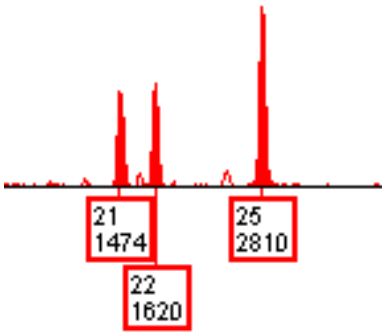


NF=1.959

Signal increases almost by a factor of 2

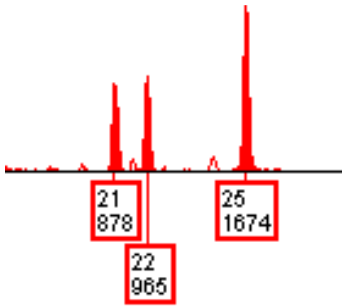


LIZ 600 v2.0 Peak Height Average: **3192 RFU**



NF= 0.596

Signal decreases by almost half



Future Work

- Validation of additional kits (Promega)
- 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 Biometrics Center of Excellence** 'Forensic DNA Typing as a Biometric Tool'

Jeff Sailus-Applied Biosystems

Contact Info:
erica.butts@nist.gov
+1-301-975-5107