Advanced Topics in STR DNA Analysis

Validation Aspects to Consider in Bringing a New STR Kit "On-Line"

john.butler@nist.gov
NIST

AAFS 2006 Workshop #6
Seattle, WA
February 20, 2006

Dr. John M. Butler
Dr. Bruce R. McCord

mccordb@fiu.edu
FLORIDA INTERNATIONAL UNIVERSITY

Outline for This Section

Bruce
- Setting peak detection thresholds
- Measuring sensitivity, dynamic range, resolution, precision
- Development of data interpretation guidelines

John
- Validation definitions and requirements for documentation
- Determining the types of tests and numbers of samples to run
- Examples

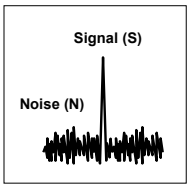
Setting thresholds for the ABI 310/3100

- Where do current ideas on instrument thresholds for the ABI 310/3100 come from?
- How do I set these values in my laboratory?
- Why might they vary from one instrument to the next?
- How do these thresholds affect data interpretation?

What is a true peak (allele)?

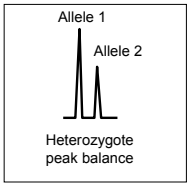
GeneScan function | *Genotyper function*

Peak detection threshold



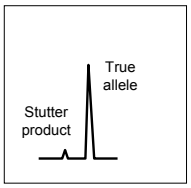
Signal > 3x sd of noise

Peak height ratio (PHR)



PHR consistent with single source Typically above 60%

Stutter percentage



Stutter location above 15%

Threshold Settings for the ABI 310/3100

Detection Limit: 3x the standard deviation of the noise.
Estimated using **2x peak to peak noise**. (approximately 35 - 50 RFUs)

Limit of Quantitation: 10x the standard deviation of the noise
Estimated using **7x peak to peak noise** (150-200 RFUs)
Below this point estimates of peak area or height are unreliable.

Dynamic Range: The range of sample quantities that can be analyzed from the lowest to the highest **(linear range is also important)**

Stochastic Threshold: Level of quantifiable DNA below which peaks can show severe imbalance (peak height ratios below 60%) Approximately 150 -200 RFUs. Enhanced stutter also occurs at these signal levels.

Will be covered more in the low copy number section of this workshop…

The Scientific Reasoning behind the Concept of an Analytical Threshold (limit of detection)

- This is fundamentally an issue of reliability
- **For a peak intensity three times the standard deviation of the noise there is a limited chance that such a signal is the result of a random fluctuation**
- This is because 99.7 percent of all noise signals fall below this value (from the definition of a Gaussian curve)
- **Below this point the very real possibility exists that what you think is a peak is simply a statistical fluctuation in the baseline noise**.

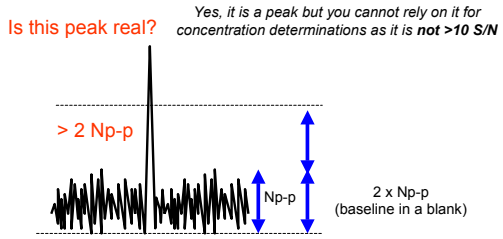
Sensitivity

- Limit of detection (**LOD**) – "the lowest content that can be measured with reasonable statistical certainty."
- Limit of quantitative measurement (**LOQ**) – "the lowest concentration of an analyte that can be determined with acceptable precision (repeatability) and accuracy under the stated conditions of the test."
- How low can you go?

EURACHEM Guide (1998) *The Fitness for Purpose of Analytical Methods: A Laboratory Guide to Method Validation and Related Topics*, p. 43; available at http://www.eurachem.ul.pt/guides/valid.pdf

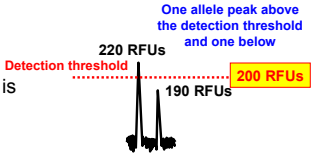
Limit of Detection (LOD)

- Typically 3 times the signal-to-noise (based on standard deviation of the noise) or 2x Np-p

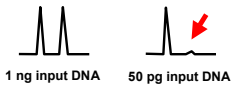


Types of Results at Low Signal Intensity
(Stochastic amplification potential)

Straddle Data
- Only one allele in a pair is above the laboratory stochastic threshold



Allelic Drop-out
- one or more sets of alleles do not amplify



TWGDAM validation of AmpFISTR Blue
Wallin *et al.* (1998) *J. Forensic Sci.* 43(4): 854-870

- Minimum cycle # (27-30 cycles examined)
- Amplification adjusted to 28 cycles so that quantities of DNA below 35pg gave very low peaks or no peaks (below the analytical threshold!)
- 35 pg is approx 5 cells
- (but is 35pg the analytical threshold?) Determining this value might be a useful goal of a validation study

TWGDAM validation of AmpFISTR Blue
Wallin *et al.* (1998) *J. Forensic Sci.* 43(4): 854-870

Determination of Minimum Sample

- Goal: avoid situations where peak imbalance results in only one detectable allele from a heterozygous pair.
- Perform serial dilution (1ng- 8pg) of 2 control samples which were heterozygous at all 3 loci
  - **Samples above 125pg had peak height RFUs above 150**
  - **Below 125pg peak heights were not significantly above background**
  - **At 31 pg peaks were very low or undetectable**
- **"Peaks below 150 RFU should be interpreted with caution"** **Why? Noise and stochastic fluctuation!**

Sensitivity of Detection
Moretti et al, JFS, 2001, 46(3), 661-676

- Different 310 instruments have different sensitivities; determination of stochastic threshold should be performed following in-house studies
  - Variations in quantitation systems
  - Variations in amplification systems
  - Variations in instrument sensitivity
- Peaks with heights below the threshold should be interpreted with caution
  - Caution should be used before modification of
    - Amplification cycles
    - Electrophoretic conditions

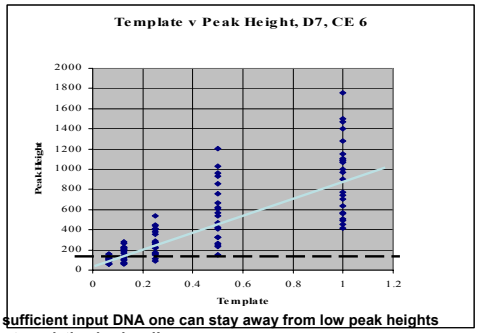
Sensitivity Study

(Debbie Hobson-FBI)

- 25 Individuals
  - 63 pg to 1 ng amplifications with Profiler Plus and Cofiler
  - amplicon run on five 310s
  - GeneScan Analysis threshold sufficient to capture all data
  - GenoTyper: category and peak height
- Import data into Excel
  - peak height ratios determined for heterozygous data at each locus

Sensitivity Study:
Profiler Plus

(Debbie Hobson-FBI)



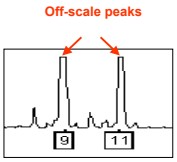
With sufficient input DNA one can stay away from low peak heights
But note variation in signal!

Offscale Data – Just as important as low signal intensity

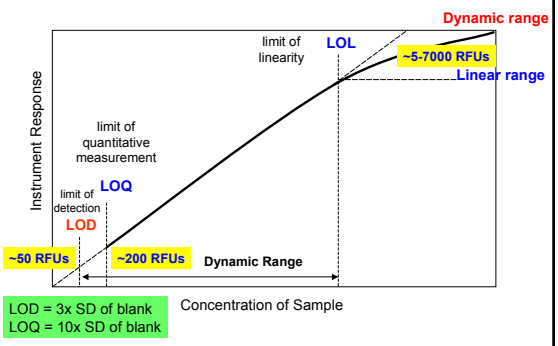
- Elevated baselines are seen with overloaded samples- Moretti et al, JFS 2001, 46(3)647-660
- Probably due to nonspecific amplification
- Stutter is artificially enhanced in such samples due to cutoff of peak top
- -A may also be apparent as a result of poor PCR conditions

Limit of Linearity (LOL)

- Point of saturation for an instrument detector so that higher amounts of analyte do not produce a linear response in signal (Linear range < Dynamic range)
- In ABI 310 or ABI 3100 detectors, the CCD camera saturates leading to flat-topped peaks.



Useful Range of an Analytical Method

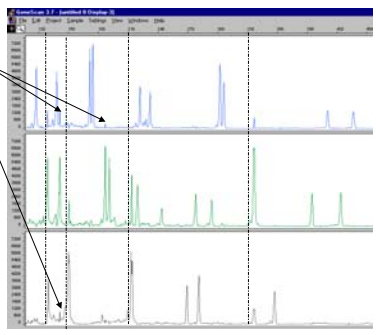


LOD = 3x SD of blank
LOQ = 10x SD of blank

Adapted from Figure 1-7 in Skoog, D.A., *et al.* (1998) *Principles of Instrumental Analysis* (5th Edition). Thomson Learning, Inc.

Matrix effects are caused by sampling outside of linear range. Overloaded samples stress the matrix calibration

The effect of pull-up on an overloaded sample



Setting Laboratory Thresholds

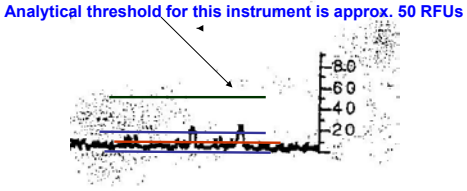
- Analytical Threshold – the minimum quantity that can be detected
- Dynamic Range – the range of sample quantities (highest and lowest) that can be detected
- Stochastic Threshold – the signal intensity at which a particular quantity of DNA can no longer reliably be detected
  - **Reliability can be defined by an increase in the standard deviation of peak height intensity or an increase in the standard deviation of signal intensity or both.**
  - **The stochastic threshold is greater than or equal to the analytical threshold**

How to Determine the Dynamic Range

1. Perform a series of amplifications of 5 different samples with 5.00, 2.00, 1.00, 0.50, 0.25, 0.13, 0.06, 0.03 ng DNA
2. Use your laboratory quantification system, your thermal cycler, and your 310.
3. Determine the average and standard deviation of each set of samples
4. Your dynamic range is the range of concentrations that are not overloaded. The linear range can be established by running concentration standards.

How to Set Thresholds

- First determine the analytical threshold for your particular laboratory using the signal intensity from one or several CE systems



How to determine the stochastic threshold

- Examine intensity and peak height ratio of 5 samples at three different low concentrations (e.g., 60, 75, and 125 pg)
- Observe variation in peak height ratio and peak intensity
- The stochastic threshold is the point at which this variation begins a rapid increase (change in slope of line relating std dev vs concentration)
- This can also be defined as the concentration at which a set percentage of peak height ratio values fall below 60%

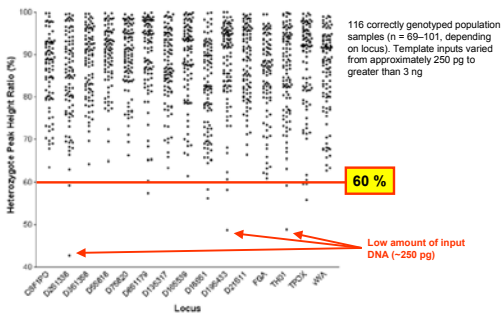
TWGDAM validation of AmpFlSTR BluePCR

Wallin et al.JFS, 1998 43(4) 854-870

- In approximately 80 heterozygous loci in population samples:
  - Average peak height ratio was 92% for each locus – D3, vWA, FGA
  - Standard deviation was 7%
- Thus 99.7% of all samples should show a peak height ratio (PHR) above 71%
- Those that have a PHR of <70% may result from mixtures, low [DNA], inhibition, degradation or poor primer binding

Heterozygote Peak Height Ratios

Identifiler STR Kit Developmental Validation



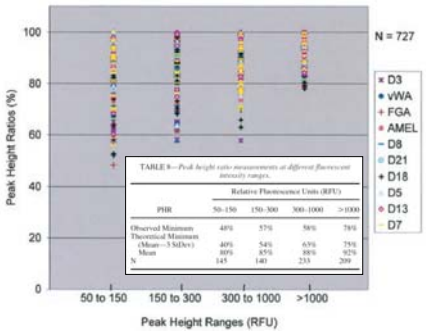
Collins PJ, Hennessy LK, Leibelt CS, Roby RK, Reeder DJ, Foxall PA. Developmental validation of a single-tube amplification of the 13 CODIS STR loci, D2S1338, D19S433, and amelogenin: the AmpFlSTR Identifiler PCR amplification kit. *J. Forensic Sci.* 2004; 49(6): 1265-1277.

Peak height ratios

Moretti et al., JFS 2001, 46(3) 647-660

- PP + Cofiler gave PHR >88% n= 230+ samples with a lower range PHR (-3sd) of 59%
- Suggest using 59% as a guide
- 2% of single source samples were below this value
- Many validation studies focus on 1ng input DNA. What happens with lower amounts?

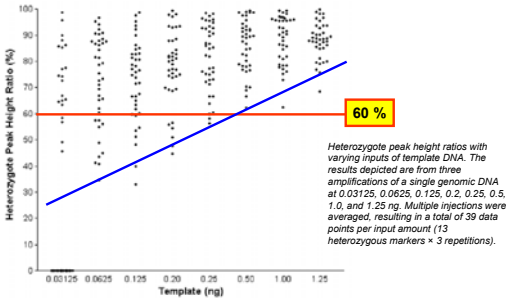
Heterozygote Peak Height Ratios



Holt CL, Buoncristiani M, Wallin JM, Nguyen T, Lazaruk KD, Walsh PS. TWGDAM validation of AmpFISTR PCR amplification kits for forensic DNA casework. *J Forensic Sci* 2002; 47(1): 66-96.

Heterozygote Peak Height Ratios

Identifiler STR Kit Developmental Validation



Heterozygote peak height ratios with varying inputs of template DNA. The results depicted are from three amplifications of a single genomic DNA at 0.03125, 0.0625, 0.125, 0.2, 0.25, 0.5, 1.0, and 1.25 ng. Multiple injections were averaged, resulting in a total of 39 data points per input amount (13 heterozygous markers × 3 repetitions).

Collins PJ, Hennessy LK, Leibelt CS, Roby RK, Reeder DJ, Foxall PA. Developmental validation of a single-tube amplification of the 13 CODIS STR loci, D2S1338, D19S433, and amelogenin: the AmpFISTR Identifiler PCR amplification kit. *J. Forensic Sci.* 2004; 49(6): 1265-1277.

Peak Height Ratio Guidelines

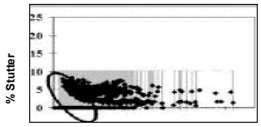
- One way to approach concentration dependence
- Profiler Plus
  - 200 to 300 RFU: 55 to 60%
  - 300 to 1000 RFU: 60 to 65%
  - above 1000 RFU: 65 to 70%
- Cofiler
  - 200 to 300 RFU: 60%
  - 300 to 1000 RFU: 60 to 65%
  - above 1000 RFU: 70% to 75%

TWGDAM validation of AmpFISTR Blue

Wallin et al.JFS, 1998 43(4) 854-870

- Stutter increases with allele size:
- Greatest amount was 8.6% w/ sd of 0.6, thus 8.6 + 3(0.6) = 10.4 or approx. 11% is the stutter threshold
- Stutter increases at low copy #

Leclair et. al (2004) Systematic analysis of stutter percentages and allele peak height and peak area ratios at heterozygous STR loci for forensic casework and database samples. J. Forensic Sci. 49(5): 968-980



Issues with Data below the Stochastic threshold and above the analytical threshold

- PCR artifacts and stutter become prevalent
- Low levels of bleed through are possible
- Instrument spikes are more numerous
- -A peaks may appear
- Dye blobs become more significant in overall e-gram
- Low level 2nd contributors may show peaks

Typically between 50 and 200 RFU – depends on validation studies

Fuzzy Logic in Data Interpretation

- The ABI 310 is a dynamic system
- Sensitivity varies with
  - Allele size
  - Injection solvent
  - Input DNA
  - Instrument factors
  - Presence of PCR inhibitors
  - Gel matrix

- Thus interpretation must be conservative and data from these studies yields guidelines, not rules. The results and their significance cannot be dissociated from the overall facts of the case.

So why examine low level data at all?

•Detection of straddle data in which one allele is above threshold and the other is below

•Detection of the presence of low level mixtures

•Clues to the presence of inhibited samples or poor injections

•Aids in determination if a suspect is excluded as a contributor

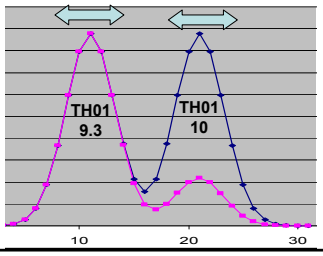
Other Analytical Factors to Consider

- Precision
- Resolution
- Sizing Algorithm

Precision (Resolution isn't enough!)

Microvariants can appear in the 4 base repeat motif present in these STRs

If 3 X Std. Dev. is greater than 0.5, then a certain number of 9.3 peaks will be labeled 10.



+/- 3σ = 99.7

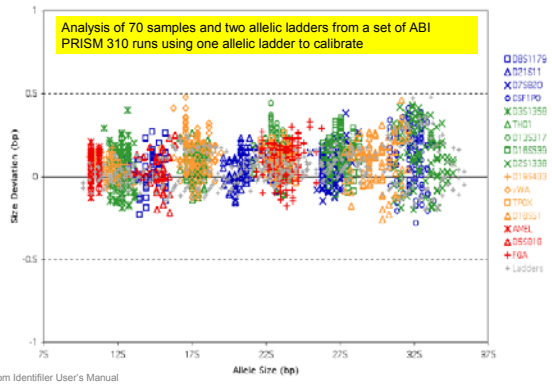
Precision

Current values in the literature range from 0.12 to 0.24 depending on the system and type of repeat. Most papers in the forensic literature report values under 0.15.

What affects precision?

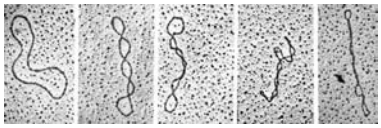
Obtaining good precision is not magic

Run-to-Run Precision on an ABI 310



Temperature Effects and DNA Secondary Structure

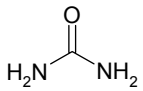
- Even under highly denaturing conditions DNA can self associate
- Differences in conformation can affect migration time
- Increase precision by limiting this effect?



http://biochem.stanford.edu/biochem201/Slides/DNA%20Topology/065%20Supercoiled%20DNA%20jpeg.JPG

How to avoid 2° Structure Effects

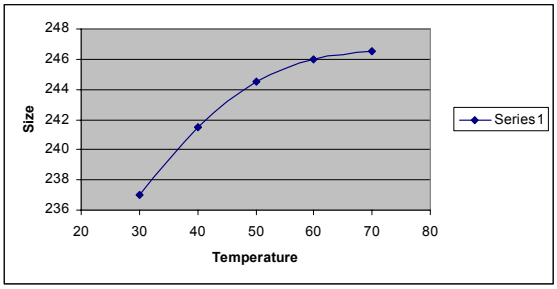
- **Elevate Temperature to 60°**
- **Add Strong Denaturants**
  - **7-8M Urea**
  - **Pyrrolidinone**
- **Examine response of 250, 340 peaks in ROX ladder**



urea

Rosenblum et al., Nucleic Acids Res., 1997

Change in size of GS 250 peak with Temperature (Tamra Std)



4% pDMA with 8M urea and 5% 2-pyrrolidinone
***Rosenblum et al., Nucleic Acids Res.(1997) 25,19, 2925**

Precision and Resolution

Elevated temperatures melt out DNA 2° structure, increasing the precision of the analysis. However, resolution is lost as a result of decreased viscosity.

100mM TBE 2% HEC, DB-17 Capillary

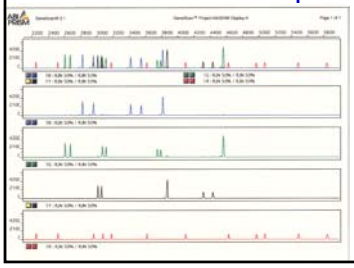
Temp.	Resolution (bp)	allele size	Std. Dev.
30	1.3	197.4	0.20
45	1.6	196.0	0.08
60	1.7 (n=7)	195.6 (n=7)	0.07 (n=200+)

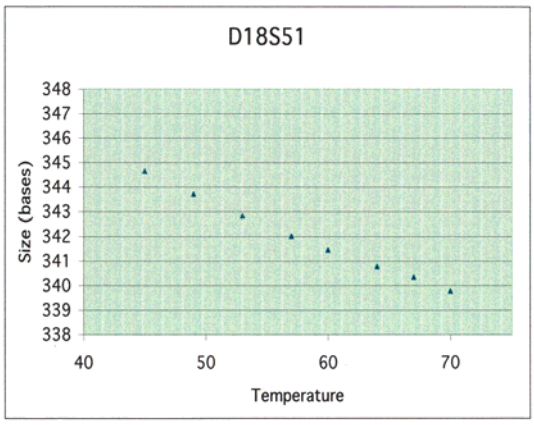
What is the effect of Temperature on Profiler+?

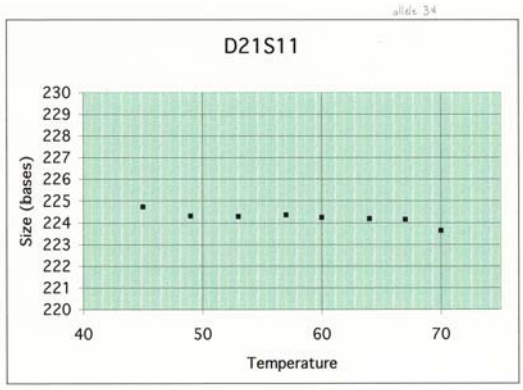
Could it affect precision, allele size?

Why do some band shifts occur at only one locus?

Examine various alleles at temperatures from 40-70 °C







Effect of Temperature on Allele Size

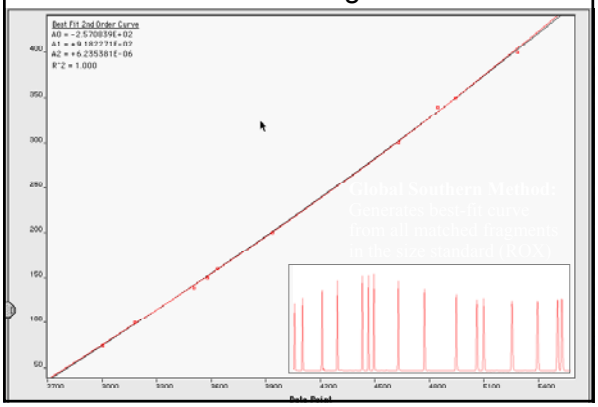
POP4, pH 8, 350V/cm, 45-70°C

STR	Allele	Size*	Slope+	Std. Dev
D3S1358	12	111.2	-0.10	0.01
vWA	21	194.9	-0.07	0.02
FGA	30	264.7	-0.14	0.02
Amel.	X	103.5	-0.13	0.01
D8S1179	19	170.4	-0.16	0.02
D21S11	36	232.4	-0.03	0.01
D18S51	26	341.9	-0.18	0.01
D5S818	7	131.2	-0.09	0.01
D13S317	8	205.0	-0.12	0.01
D7S820	15	292.8	-0.09	0.01

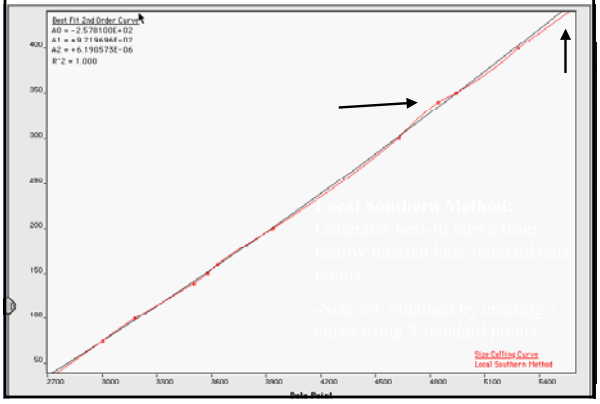
Effect of Operator Chosen Sizing Method

- **Global Southern Method:** Generates best-fit curve from all matched fragments in the size standard
- **Local Southern Method:** Generates best-fit curve from only nearby internal lane standard data points

Global Southern Sizing Method



Local Southern Sizing Method



Effect of Operator Chosen Sizing Method

Global Southern Sizing		
allele #		
CSF1PO	average slope	SD of ave.
7	-0.052	0.01
10	-0.050	0.01
12	-0.047	0.01
14	-0.048	0.01
VWA		
14	-0.085	0.002
15	-0.087	0.004
17	-0.089	0.006

Local Southern Sizing		
allele #		
CSF1PO	average slope	SD of ave.
7	-0.027	0.01
10	-0.135	0.02
12	-0.103	0.02
14	-0.156	0.01
VWA		
14	-0.060	0.009
15	-0.059	0.009
17	-0.097	0.008

- **Global Southern:**
  - **Similar slopes within a locus**
  - **Differential response in slopes between loci**
- **Local Southern:**
  - **Differential response between and within loci**
  - **Many slopes significantly larger (-0.156 vs. –0.104)**

Hartzell, Muncy, McCord, Forensic Science International, 2003, *133*, 228-234.,

Implications of Temperature Studies

Temperature affects precision through sample denaturation

New studies indicate there is a variable response to temperature especially between loci

The effect is far more pronounced in local southern

Temperature control is important because it affects both precision and resolution.

Band shifts are a natural consequence of differential response to temperature

How Does Your Laboratory Derive Its Interpretation Rules?

From your Validation Studies or Others?

- **Peak detection threshold** – set to 50 RFU or 150 RFU based on your lab data or what FBI or manufacturer has done? Do you use S/N >3 for determining if something is a true peak?
- **Peak height ratio threshold** – Set at 70% due to suggestion by manufacturer? Or 50-70% based on other data?
- **Stutter product threshold** – are Genotyper macros set to 15%, manufacturer values, or adjusted based on your validation? Does it matter? How do these values play into your mixture interpretation guidelines?
- **Sample Cleanup** - Post PCR concentration a sample may also remove salts artificially enhancing injection. Will this move results into stochastic range?

Example of an Interpretational Guideline

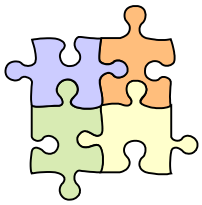
1or 3 ul product:	microcon enhancement:	Action:
50RFU < peak < 150RFU	peak > 150RFU	Report peak
peak <50RFU	peak > 150RFU	Report activity (A) only
peak <50RFU	50RFU < peak < 150RFU	Report activity (A) only
no peak detected	peak > 150RFU	Not reported
no peak detected	50RFU < peak < 150RFU	Not reported

What is **Validation** and Why Should It Be Done?

- Part of overall quality assurance program in a laboratory
- **We want the correct answer when collecting data…**
  - We want **analytical measurements made in one location to be consistent with those made elsewhere** (without this guarantee there is no way that a national DNA database can be successful).
- If we fail to get a result from a sample, we want to have confidence that the sample contains no DNA rather than there might have been something wrong with the detection method…

Elements for Guaranteeing Quality Results in Forensic DNA Testing

- Accepted Standards and Guidelines for Operation
- Laboratory Accreditation
- Proficiency Testing of Analysts
- Standard Operating Procedures
- **Validated Methods**
- Calibrated Instrumentation
- Documented Results
- Laboratory Audits
- **Trustworthy Individuals**



Assumptions When Performing Validation

- The equipment on which the work is being done is broadly suited to the application. It is clean, well-maintained and within calibration.
- The staff carrying out the validation are competent in the type of work involved.
- There are no unusual fluctuations in laboratory conditions and there is no work being carried out in the immediate vicinity that is likely to cause interferences.
- The samples being used in the validation study are known to be sufficiently stable.

Roper, P., *et al.* (2001) *Applications of Reference Materials in Analytical Chemistry.* Royal Society of Chemistry, Cambridge, UK, pp. 110-111.

How do you validate a method?

- Decide on analytical requirements
  - Sensitivity, resolution, precision, etc.
- **Plan a suite of experiments**
- **Carry out experiments**
- Use data to assess fitness for purpose
- Produce a statement of validation
  - Scope of the method

Roper, P., *et al.* (2001) *Applications of Reference Materials in Analytical Chemistry*. Royal Society of Chemistry, Cambridge, UK, pp. 108-109.

Tools of Method Validation

- Standard samples
  - positive controls
  - NIST SRMs
- Blanks
- Reference materials prepared in-house and spikes
- Existing samples
- Statistics
- **Common sense**

Roper, P., *et al.* (2001) *Applications of Reference Materials in Analytical Chemistry*. Royal Society of Chemistry, Cambridge, UK, p. 110.

Common Perceptions of Validation

The goal is not to experience every possible scenario during validation…

"You cannot mimic casework because every case is different."

Effort

Lots of experiments are required

Many labs are examining far too many samples in validation and thus delaying application of casework and contributing to backlogs…

Significant time is required to perform studies

Time

Validation Workshop (Aug 24-26, 2005 at NFSTC)

http://www.cstl.nist.gov/biotech/strbase/validation/validationworkshop.htm

COURSE CONTENTS

Day #1
- Validation Overview (John)
- Introduction to DAB Standards (Robyn & John)
- Developmental Validation (John)

Day #2
- Inconsistency in Validation between Labs (John)
- Internal Validation (Robyn)
- Method Modifications and Performance Checks (Robyn)

Day #3
- Practical Exercises (Robyn)

Was filmed and is being made into a training DVD as part of the President's DNA Initiative…

Number of Samples Needed

Relationship between a sample and a population of data

How do you relate these two values?

Data collected in your lab as part of validation studies → All potential data that will be collected in the future in your lab

Student's *t*-Test associates a sample to a population

"Sample" of Typical Data

"Population" of All Data Obtained

Student's *t*-Tests

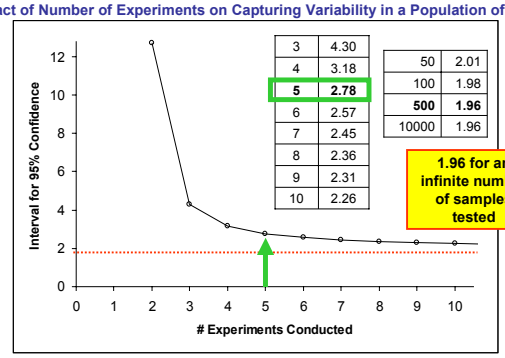
"Student" (real name: W. S. Gossett [1876-1937]) developed statistical methods to solve problems stemming from his employment in a brewery.

Student's *t*-test deals with the problems associated with inference based on "small" samples: the calculated mean (X_{avg}) and standard deviation (σ) may by chance deviate from the "real" mean and standard deviation (i.e., **what you'd measure if you had many more data items: a "large" sample**).

http://www.physics.csbsju.edu/stats/t-test.html

Student's t-Test Curve

Impact of Number of Experiments on Capturing Variability in a Population of Data



3	4.30
4	3.18
5	**2.78**
6	2.57
7	2.45
8	2.36
9	2.31
10	2.26

50	2.01
100	1.98
500	**1.96**
10000	1.96

1.96 for an infinite number of samples tested

Revised SWGDAM Validation Guidelines
(July 2004)

http://www.fbi.gov/hq/lab/fsc/backissu/july2004/standards/2004_03_standards02.htm

Forensic Science Communications July 2004 – Volume 6 – Number 3
Standards and Guidelines

Revised Validation Guidelines

Table of Contents
Back Issues
Search
Editors
About FS
Instructio
Authors

Scientific Working Group on DNA Analysis Methods (SWGDAM)

3. Internal Validation
…a total of **at least 50 samples**
(some studies may not be necessary…)

Program for DNA Analysis by the Technical Working Group on DNA Analysis Methods (*Crime Laboratory Digest* 1995;22(2):21-43) has been revised due to increased laboratory experience, the advent of new technologies, and the issuance of the Quality Assurance Standards for Forensic DNA Testing Laboratories by the Director of the FBI (*Forensic Science Communications* available: www.fbi.gov/hq/lab/fsc/backissu/july2000/codis2a.htm).

The document provides validation guidelines and definitions approved by SWGDAM July 10, 2003.

Design of Experiments Conducted for Validation Studies

- Before performing a set of experiments for validation, ask yourself:
  - **What is the purpose of the study?**
  - **Do we already know the answer?**
  - **Can we write down how we know the answer?**
- Think before you blindly perform a study which may have no relevance (e.g., extensive precision studies)
- **Too often we do not differentiate learning, validation, and training**

Points for Consideration

- Remove as many variables as possible in testing an aspect of a procedure
  - e.g., create bulk materials and then aliquot to multiple tubes rather than pipeting separate tubes individually during reproducibility studies
- Who can do (or should do) validation…
  - Outside contractor?
  - Summer intern
  - Trainee
  - Qualified DNA analyst

What are the goals of validation studies involving a new STR typing kit?

- Stutter product amounts
  Why?: aids in mixture interpretation guidelines (how often does your laboratory call peaks below 15% of an adjacent allele?)
- Precision studies
  Why?: aids in defining allele bin windows (in reality does anyone ever change the ±0.5 bp from the Genotyper macro?)
- Sensitivity studies
  Why?: aids in defining lower and upper limits
- Mixture studies
  Why?: aids in demonstrating the limits of detecting the minor component
- Concordance studies
  Why?: to confirm that new primer sets get the same results as original primer sets – potential of polymorphism causing allele dropout…
- Peak height ratio studies
  Why?: aids in mixture interpretation guidelines (how often does your laboratory call peaks below a 60% heterozygote peak height ratio?)

FBI DNA Quality Assurance Audit Developmental Validation Scorecard

Discussion

Developmental validation must precede the introduction of a novel methodology for forensic DNA analysis. A novel methodology may include an existing technology or testing procedure that has been developed for a specific technology (e.g., medical testing, genetic analysis) that is not currently applied to forensic DNA analysis. Citations in peer-reviewed scientific journals that provide the underlying scientific basis for a novel methodology should be available.

		Yes	No	N/A
8.1.2	Have novel forensic or database DNA methodologies used by the laboratory undergone developmental validation to ensure the accuracy, precision, and reproducibility of the procedure?			
8.1.2.1	Is there documentation and is it available that defines and characterizes each locus?			
8.1.2.2(FO)	Have species' specificity, sensitivity, stability, and mixture studies been conducted?			
8.1.2.3(FO)	Does the laboratory have access to a population database that is documented and available for use in population statistics?			
8.1.2.3.1(FO-a)	Where appropriate, has the database been tested for independence expectations?			
8.1.2.3.1(FO-b)	Does the database information include allele and frequency distributions for the locus or loci obtained from relevant populations?			

FBI DNA Quality Assurance Audit Document
Issue Date 07/04 (Rev. #6)
28

DNA Advisory Board Quality Assurance Standards

Section 2. Definitions

- (ff) Validation is a process by which a procedure is evaluated to determine its **efficacy and reliability** for forensic casework analysis *(DNA analysis)* and includes:

Manufacturer

  - (1) Developmental validation is the acquisition of test data and determination of conditions and limitations of a new or novel DNA methodology for use on forensic samples;
  - (2) Internal validation is an accumulation of test data within the laboratory to demonstrate that established methods and procedures perform as expected in the laboratory.

Forensic Lab

Validation Section of the DNA Advisory Board Standards

issued October 1, 1998 and April 1999; published in *Forensic Sci. Comm.* July 2000

STANDARD 8.1 The laboratory shall use validated methods and procedures for forensic casework analyses *(DNA analyses)*.

8.1.1 Developmental validation that is conducted shall be appropriately documented.

8.1.2 Novel forensic DNA methodologies shall undergo developmental validation to ensure the accuracy, precision and reproducibility of the procedure. The developmental validation shall include the following:

8.1.2.1 Documentation exists and is available which defines and characterizes the locus.

8.1.2.2 Species specificity, sensitivity, stability and mixture studies are conducted.

8.1.2.3 Population distribution data are documented and available.

8.1.2.3.1 The population distribution data would include the allele and genotype distributions for the locus or loci obtained from relevant populations. Where appropriate, databases should be tested for independence expectations.

8.1.3 Internal validation shall be performed and documented by the laboratory.

8.1.3.1 The procedure shall be tested using known and non-probative evidence samples *(known samples only)*. The laboratory shall monitor and document the reproducibility and precision of the procedure using human DNA control(s).

8.1.3.2 The laboratory shall establish and document match criteria based on empirical data.

8.1.3.3 Before the introduction of a procedure into forensic casework *(database sample analysis)*, the analyst or examination team shall successfully complete a qualifying test.

8.1.3.4 Material modifications made to analytical procedures shall be documented and subject to validation testing.

8.1.4 Where methods are not specified, the laboratory shall, wherever possible, select methods that have been published by reputable technical organizations or in relevant scientific texts or journals, or have been appropriately evaluated for a specific or unique application.

FORENSIC SCIENCE COMMUNICATIONS JULY 2000 VOLUME 2 NUMBER 3

DNA Advisory Board Standards

(***Forensic Sci. Comm.* July 2000**)

8.1.3 **Internal validation shall be performed and documented by the laboratory.**

8.1.3.1 The procedure shall be tested using known and non-probative evidence samples *(known samples only)*. The laboratory shall monitor and document the reproducibility and precision of the procedure using human DNA control(s).

8.1.3.2 The laboratory shall establish and document match criteria based on empirical data.

8.1.3.3 Before the introduction of a procedure into forensic casework *(database sample analysis)*, the analyst or examination team shall successfully complete a qualifying test.

8.1.3.4 Material modifications made to analytical procedures shall be documented and subject to validation testing.

Overview of Internal Validation Studies

3. Internal Validation: The internal validation process should include the studies detailed below encompassing **a total of at least 50 samples**. Some studies may not be necessary due to the method itself.

   3.1 Known and nonprobative evidence samples
   3.2 Reproducibility and precision
   3.3 Match criteria
   3.4 Sensitivity and stochastic studies
   3.5 Mixture studies
   3.6 Contamination
   3.7 Qualifying test

SWGDAM Revised Validation Guidelines
http://www.fbi.gov/hq/lab/fsc/backissu/july2004/standards/2004_03_standards02.htm

Suggestions for an Internal Validation of an STR Kit

- Standard samples (3.1) **Between 1 and ~20 samples**
  - Verify correct type with positive control or NIST SRM samples
  - Concordance study with 5-10 (non-probative casework) samples previously typed with other kit(s)
- Precision samples (3.2) **5-10 samples**
  - Run at least 5-10 samples (allelic ladder or positive control)
- Sensitivity samples (3.4) **14 samples**
  - Run at least 2 sets of samples covering the dynamic range
  - 5 ng down to 50 pg—e.g., 5, 2, 1, 0.5, 0.2, 0.1, 0.05 ng
- Mixture samples (3.5) **10 samples**
  - Run at least 2 sets of samples
  - Examine 5 different ratios—e.g., 10:1, 3:1, 1:1, 1:3, 1:10

>50 samples

Additional Suggestions for Meeting the SWGDAM Revised Validation Guidelines

- Match Criteria (3.3)
  - As part of running a batch of samples (e.g., 10 or 96), run one allelic ladder at the beginning and one at the end
  - If all alleles are typed correctly in the second allelic ladder, then the match criteria (i.e., precision window of +/-0.5 bp) has likely been met across the entire size range and duration of the run
- Contamination Check (3.6)
  - Run negative controls (samples containing water instead of DNA) with each batch of PCR products
- Qualifying Test (3.7)
  - Run proficiency test samples

Steps Surrounding "Validation" in a Forensic Lab

Effort to Bring a Procedure "On-Line"

This is what takes the time…

- **Installation** – purchase of equipment, ordering supplies, setting up in lab
- **Learning** – efforts made to understand technique and gain experience troubleshooting; can take place through direct experience in the lab or vicariously through the literature or hearing talks at meetings
- **Validation of Analytical Procedure** – tests conducted in one's lab to verify range of reliability and reproducibility for procedure
- **SOP Development** – creating interpretation guidelines based on lab experience
- **QC of Materials** – performance check of newly received reagents
- **Training** – passing information on to others in the lab
- **Qualifying Test** – demonstrating knowledge of procedure enabling start of casework
- **Proficiency Testing** – verifying that trained analysts are performing procedure properly over time

General Steps for Internal Validation

- **Review literature and learn the technique**
- Obtain equipment/reagents, if necessary
- Determine necessary validation studies (there can be overlap and you only need to run a total of 50 samples)
- Collect/obtain samples, if necessary
- **Perform validation studies maintaining all documentation**
- Summarize the studies and submit for approval to Technical Leader
- Write-up the analytical procedure(s). Include quality assurance (controls, standards, critical reagents and equipment) and data interpretation, as applicable
- Determine required training and design training module(s)
- Design qualifying or competency test

From Robyn Ragsdale (FDLE), Validation Workshop (Aug 24-26, 2005 at NFSTC)
http://www.cstl.nist.gov/biotech/strbase/validation/validationworkshop.htm

Other DAB Standards to Consider:

9.1.1 The laboratory shall have an ***standard protocol*** for each analytical technique used.

9.1.2 The procedures shall include ***reagents, sample preparation, extraction, equipment and controls***, which are standard for DNA analysis and data interpretation.

9.2.3 The laboratory shall identify ***critical reagents*** (if any) and evaluate them prior to use in casework……

9.4 The laboratory shall monitor the analytical procedures using appropriate ***controls and standards***.

10.2 The laboratory shall identify ***critical equipment*** and shall have a documented program for calibration of instruments and equipment.

10.3 The laboratory shall have a ***documented program*** to ensure that instruments and equipment are properly maintained.

From Robyn Ragsdale (FDLE), Validation Workshop (Aug 24-26, 2005 at NFSTC)
http://www.cstl.nist.gov/biotech/strbase/validation/validationworkshop.htm

Example with Identifiler STR Kit

- **Your lab is currently running ProfilerPlus/COfiler and wants to switch to Identifiler**. What is needed for your internal validation?
- **What is different between Identifiler and ProfilerPlus/COfiler**?
  - **Two new STR loci: D19S433 and D2S1338**
  - **Different fluorescent dyes**
  - **Additional fluorescent dye (5-dye vs 4-dye)**
  - **Different dye on internal size standard**
  - **More loci being amplified in the multiplex**
  - **Mobility modifiers to move allele sizes**

Different
Loci (2 extra STRs)
Dyes
Mobility Modifiers
Software (5-dye)

- PCR primer sequences are the same so potential allele discordance due to primer binding site mutations should not be an issue
- **What has been reported in terms of developmental validation for Identifiler?**

ABI Kit Validation Papers

J. Forensic Sci. 2002; 47(1): 66-96

Cydne L. Holt,[1] Ph.D.; Martin Buoncristiani,[2] M.P.H.; Jeanette M. Wallin,[1] M.P.H.; Theresa Nguyen,[1] B.S.; Katherine D. Lazaruk,[1] Ph.D.; and P. S. Walsh,[1] M.P.H.

TWGDAM Validation of AmpFℓSTR™ PCR Amplification Kits for Forensic DNA Casework

J. Forensic Sci. 2004; 49(6): 1265-1277

Patrick J. Collins,[1] B.A.; Lori K. Hennessy,[1] Ph.D.; Craig S. Leibelt,[1] A.B.; Rhonda K. Roby,[1,†] M.P.H.; Dennis J. Reeder,[1] Ph.D.; and Paul A. Foxall,[1] Ph.D.

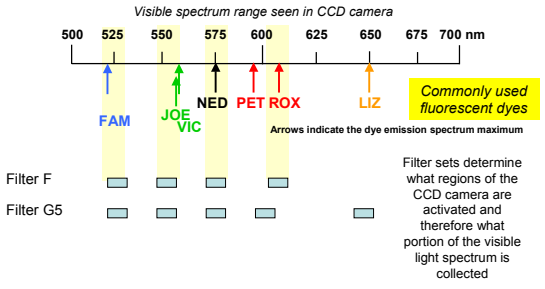
Developmental Validation of a Single-Tube Amplification of the 13 CODIS STR Loci, D2S1338, D19S433, and Amelogenin: The AmpFℓSTR® Identifiler® PCR Amplification Kit

AmpFℓSTR® Identifiler™
PCR Amplification Kit
User's Manual

Population Studies with D2S1338 and D19S433

- These STR loci are part of the widely used SGM Plus kit
- Included in profile frequency calculator using 24 European populations and 5,700 individuals: http://www.str-base.org/calc.php
- Budowle, B. (2001) Genotype profiles for five population groups at the short tandem repeat loci D2S1338 and D19S433. *Forensic Sci. Comm.* 3(3); available at http://www.fbi.gov/hq/lab/fsc/backissu/july2001/budowle1.htm
- Budowle, B., *et al.* (2001) Population data on the STR loci D2S1338 and D19S433. *Forensic Sci. Comm.* 3(3); available at http://www.fbi.gov/hq/lab/fsc/backissu/july2001/budowle2.htm
- Butler, J.M., *et al.* (2003) Allele frequencies for 15 autosomal STR loci on U.S. Caucasian, African American, and Hispanic populations. *J. Forensic Sci.* 48(4):908-911; genotypes available at http://www.cstl.nist.gov/biotech/strbase/NISTpop.htm

Different Fluorescent Dyes



	Blue	Green	Yellow	Red	Orange	Used with These Kits
Filter F	5FAM	JOE	NED	ROX		Profiler Plus
Filter G5	6FAM	VIC	NED	PET	LIZ	Identifiler

Mobility Shift with Non-Nucleotide Linker "Mobility Modifiers"

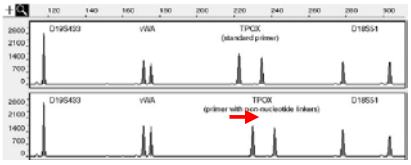


FIG. 1—*NED dye labeled loci from two amplifications of a single sample using TPOX primers both with and without non-nucleotide linkers. The X-axis indicates base pair size and the Y-axes RFU. The top panel depicts the amplification without non-nucleotide linkers. Sizes for the TPOX alleles for this panel were 222.93 and 234.81 bp. Sizes for the TPOX alleles in the amplification using the modified primer, depicted in the bottom panel, were 229.85 and 241.71 bp, indicating an average shift of 6.91 bp. Peaks heights, intralocus balance, and intracolor balance were similar in both amplifications.*

Collins PJ, Hennessy LK, Leibelt CS, Roby RK, Reeder DJ, Foxall PA. Developmental validation of a single-tube amplification of the 13 CODIS STR loci, D2S1338, D19S433, and amelogenin: the AmpFISTR Identifiler PCR amplification kit. *J. Forensic Sci.* 2004; 49(6): 1265-1277.

Mobility Modifiers

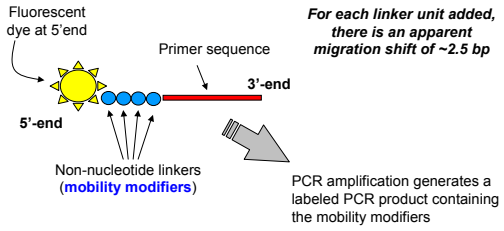


Figure 5.7, J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition © 2005 Elsevier Academic Press

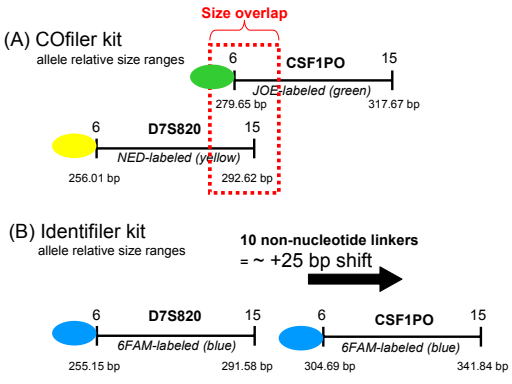


Figure 5.8, J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition © 2005 Elsevier Academic Press

Sizing Precision with Non-Nucleotide Linkers

TABLE 1—*Sizing shift and sizing precision data for loci incorporating non-nucleotide linkers on the ABI PRISM 310 Genetic Analyzer.*

Locus	Increase in Detected Size (bp)	Range of Standard Deviation of Alleles for Identifiler Kit (bp)	Range of Standard Deviation of Alleles Previous Kit (bp)
CSF1PO	26	0.08–0.13*	0.03–0.10[‡]
D2S1338	16	0.05–0.12*	0.02–0.15[‡]
D13S317	12	0.05–0.09*	0.02–0.09[§]
D16S539	23	0.06–0.09*	0.01–0.08[‡]
TPOX	7	0.03–0.08*	0.02–0.07[†]

No apparent significant decrease in precision with mobility modifiers…

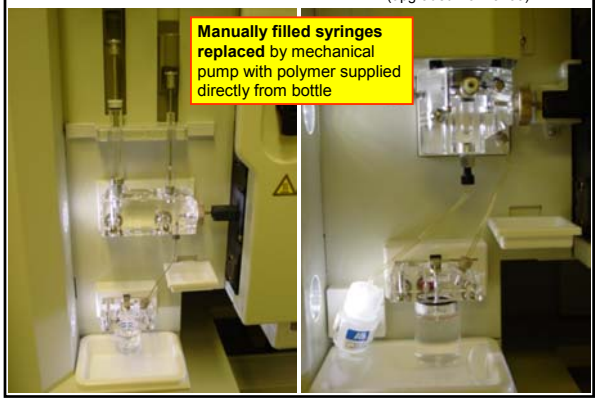
Collins PJ, Hennessy LK, Leibelt CS, Roby RK, Reeder DJ, Foxall PA. Developmental validation of a single-tube amplification of the 13 CODIS STR loci, D2S1338, D19S433, and amelogenin: the AmpFISTR Identifiler PCR amplification kit. *J. Forensic Sci.* 2004; 49(6): 1265-1277.

Instrument/Software Upgrades or Modifications

- What should be done to "validate" new upgrade?
  - ABI 7000 to ABI 7500
  - ABI 3100 to ABI 3130xl
  - GeneScan/Genotyper to GeneMapper*ID*
- Try to understand what is different with the new instrument or software program compared to the one you are currently using (e.g., ask other labs who may have made the switch)
- If possible, try to retain your current configuration for comparison purposes for the validation period

Run the same plate of samples on the original instrument/software and the new one

ABI 3100 | ABI 3130xl (upgraded from 3100)



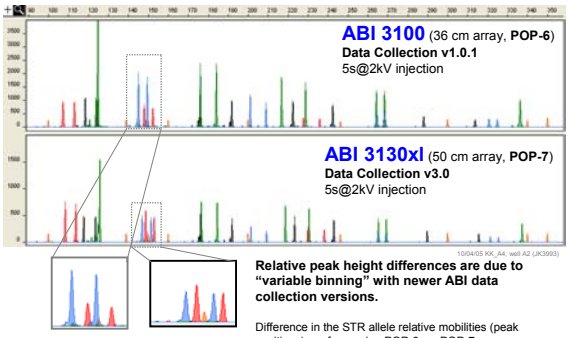
ABI 3130xl vs ABI 3100

What NIST did to "validate" a 3130xl upgrade

- Ran plates of samples on both instruments with same injection and separation parameters and compared results
  - Data Collection version 1.0.1 (3100) vs 3.0 (3130xl)
  - POP-6 (3100) vs POP-7 (3130xl)
  - 36 cm array (3100) vs 50 or 80 cm array (3130xl)
- Ran several plates of Identifiler samples and compared allele calls (noticed a sensitivity difference with equal injections and relative peak height differences between dye colors) – **all obtained allele calls were concordant**
- Ran a plate of Profiler Plus samples and compared sizing precision – **precision was not significantly different**
- Also examined SNaPshot products and mtDNA sequencing data

Comparison of ABI 3100 Data Collection Versions

Same DNA sample run with Identifiler STR kit (identical genotypes obtained) GeneScan display



Relative peak height differences are due to "variable binning" with newer ABI data collection versions.

Difference in the STR allele relative mobilities (peak positions) are from using POP-6 vs. POP-7.

Validation Section of the DNA Advisory Board Standards

issued July 1998 (and April 1999); published in *Forensic Sci. Comm.* July 2000

STANDARD 8.1 The laboratory shall use validated methods and procedures for forensic casework analyses *(DNA analyses)*.

8.1.1 Developmental validation that is conducted **shall be appropriately documented**.

8.1.3 Internal validation **shall be performed and documented by the laboratory**.

FORENSIC SCIENCE COMMUNICATIONS JULY 2000 VOLUME 2 NUMBER 3

Why is Documentation of Validation Important?

9. Documentation of Validated Methods

9.1 Once the validation process is complete it is important to document the procedures so that the method can be clearly and unambiguously implemented. There are a number of reasons for this. **The various assessments of the method made during the validation process assume that,in use, the method will be used in the same way each time**. If it is not, then the actual performance of the method will not correspond to the performance predicted by the validation data. Thus the **documentation must limit the scope for introducing accidental variation to the method**. In addition, proper documentation is necessary for auditing and evaluation purposes and may also be required for contractual or regulatory purposes.

9.2 Appropriate documentation of the method will help **to ensure that application of the method from one occasion to the next is consistent.**

EURACHEM Guide (1998) *The Fitness for Purpose of Analytical Methods: A Laboratory Guide to Method Validation and Related Topics*, p. 37; available at http://www.eurachem.ul.pt/guides/valid.pdf

Laboratory Internal Validation Summaries