Purpose of This Workshop

- DNA labs often encounter challenges when working with the many variable aspects of STR analysis and capillary electrophoresis separation/detection.
- This workshop will explore common challenges experienced by forensic laboratories and suggest solutions for fixing various problems.
- Participants are invited to suggest problems that they would like to have reviewed in advance of the workshop.
- Tried to have limited enrollment to encourage group discussion.

Presentation Plan

15 min
- Introductions
- Fundamentals of CE
  - sample prep, injection, separation, detection
  - setting instrument parameters and thresholds
    - applying validation data
    - mixture interpretation

30 min (John)
- Troubleshooting strategies and solutions
- Review of some specific problems
- Questions

BREAK

45 min (Bruce/John)

Our Backgrounds

John Butler
- NIST Fellow - National Institute of Standards and Technology
- PhD in Analytical Chemistry from University of Virginia (1995)
- Family: wife Terilynne and six children
- Hobbies: reading, writing, and making PowerPoint slides

Bruce McCord
- Professor of Analytical/Forensic Chemistry – Florida International University
- PhD in Analytical Chemistry from University of Wisconsin (1986)
- Family: wife Margie and three children
- Hobbies: dixieland jazz, windsurfing, sailing and editing John’s slides

Background of Participants…

Your name
Your organization
Instrumentation in use (e.g., ABI 310, 3100, 3130x)
What you hope to learn from this workshop

NIST and NIJ Disclaimer

Funding for John Butler: Interagency Agreement between the National Institute of Justice and NIST Office of Law Enforcement Standards

Funding for Bruce McCord: National Institute of Justice

Points of view are the presenters and do not necessarily represent the official position or policies of the US Department of Justice or the National Institute of Standards and Technology.

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

These workshop slides will be available at:
http://www.cstl.nist.gov/biotech/strbase/training.htm
Review of Fundamentals for STR Analysis and Capillary Electrophoresis

Steps in DNA Analysis

- Usually a 1-2 day process (minimum of ~5 hours)

1. Collection
2. Specimen Storage
3. Extraction
4. Quantitation
5. Multiplex PCR
6. STR Typing
7. Interpretation of Results
8. Database Storage & Searching
9. Calculation of Match Probability

Steps Involved

DNA Extraction

1. DNA profile
2. Population allele frequencies
3. Case report
4. Probability of a random match

DNA Separation and Sizing

- Technology
- Biology

Capillary Electrophoresis (CE)

- Fills with polymer solution
- Laser
- 5-20 kV
- DNA separation occurs in minutes
- Data Acquisition and Analysis

Typical Instruments Used for STR Typing

- GeneAmp 9700
- Thermal Cycler for PCR Amplification
- ABI 310
- ABI 3100
- Single capillary
- 16-capillary array

Review Article on STRs and CE


National Commission on the Future of DNA Evidence

- Report published in Nov 2000
- Asked to estimate where DNA testing would be in 2, 5, and 10 years

Conclusions

- STR typing is here to stay for a few years because of DNA databases that have grown to contain millions of profiles

http://www.ojp.usdoj.gov/nij/pubs-sum/183697.htm

- Report published in Nov 2000
- Asked to estimate where DNA testing would be in 2, 5, and 10 years

Conclusions

- STR typing is here to stay for a few years because of DNA databases that have grown to contain millions of profiles
Analytical Requirements for STR Typing


- Fluorescent dyes must be spectrally resolved in order to distinguish different dye labels on PCR products
- PCR products must be spatially resolved – desirable to have single base resolution out to >350 bp in order to distinguish variant alleles
- High run-to-run precision – an internal sizing standard is used to calibrate each run in order to compare data over time

Detection with Multiple Capillaries

(Irradiation for Capillary Arrays)

Capillary Array
Side irradiation (on-capillary)

Sheath flow detection
ABI 3100

Fixed laser, moving capillaries
MegaBACE

Process Involved in 310/3100 Analysis

- Separation
  - Capillary – 50um fused silica, 43 cm length (36 cm to detector)
  - POP-4 polymer – Polydimethyl acrylamide
  - Buffer - TAPS pH 8.0
  - Denaturants – urea, pyrolidinone
- Injection
  - electrokinetic injection process (formamide, water)
  - importance of sample stacking
- Detection
  - fluorescent dyes with excitation and emission traits
  - CCD with defined virtual filters produced by assigning certain pixels

Ohm’s Law

- \[ V = IR \] (where \( V \) is voltage, \( I \) is current, and \( R \) is resistance)
- Current, or the flow of ions, is what matters most in electrophoresis
- CE currents are much lower than gels because of a higher resistance in the narrow capillary
- CE can run a higher voltage because the capillary offers a higher surface area-to-volume ratio and can thus dissipate heat better from the ion flow (current)

Separation Issues

- Electrophoresis buffer –
  - Urea for denaturing and viscosity
  - Buffer for consistent pH
  - Pyrolidinone for denaturing DNA
  - EDTA for stability and chelating metals
- Polymer solution -- POP-4 (but others work also)
- Capillary wall coating -- dynamic coating with polymer
  - Wall charges are masked by methyl acrylamide
- Run temperature -- 60 °C helps reduce secondary structure on DNA and improves precision. (Temperature control affects DNA sizing)
What is in POP-4 and Genetic Analyzer Buffer?

POP-4 (4% poly-dimethylacrylamide, 8 M urea, 5% 2-pyrrolidinone)

Running buffer contains 100 mM TAPS and 1 mM EDTA (adjusted to pH 8.0 with NaOH)

TAPS = N-Tris-(hydroxymethyl)methyl-3-aminopropane-sulfonic acid

US Patent 5,552,028 covers POP-4 synthesis

See also Wenz et al. (1998) Genome Research 8: 69-80

Impact of Capillary Length and Polymer Concentration on DNA Sequencing Resolution

How to Improve Resolution?

1. Lower Field Strength
2. Increase Capillary Length
3. Increase Polymer Concentration
4. Increase Polymer Length

All of these come at the cost of longer separation run times

Electrokinetic Injection Process

Data collected at NIST by Tomohiro Takamaya (Japanese guest researcher, fall 2007)

http://www.cstl.nist.gov/biotech/strbase/training.htm
Sample Conductivity Impacts Amount Injected

\[ [\text{DNA}_{\text{inj}}] = \frac{E t (r^2)}{\lambda_{\text{sample}}} (\mu_{\text{ep}} + \mu_{\text{eof}}) [\text{DNA}_{\text{sample}}] (\lambda_{\text{buffer}}) \]

- \( [\text{DNA}_{\text{inj}}] \) is the amount of sample injected
- \( E \) is the electric field applied
- \( t \) is the injection time
- \( r \) is the radius of the capillary
- \( \mu_{\text{ep}} \) is the mobility of the sample molecules
- \( \mu_{\text{eof}} \) is the electroosmotic mobility

\( \lambda_{\text{sample}} \) is the buffer conductivity
\( \lambda_{\text{buffer}} \) is the sample conductivity

\( \text{DNA}_{\text{sample}} \) is the concentration of DNA in the sample
\( \text{DNA}_{\text{inj}} \) is the amount of DNA in the injection


Comments on Sample Preparation

- Use high quality formamide (<100 \( \mu \)S/cm)
- Denaturation with heating and snap cooling is not needed (although most labs still do it…)
- Post-PCR purification reduces salt levels and leads to more DNA injected onto the capillary

Steps Performed in Standard Module

- Capillary fill – polymer solution is forced into the capillary by applying a force to the syringe
- Pre-electrophoresis – the separation voltage is raised to 10,000 volts and run for 5 minutes
- Water wash of capillary – capillary is dipped several times in deionized water to remove buffer salts that would interfere with the injection process
- Sample injection – the autosampler moves to position A1 (or the next sample in the sample set) and is moved up onto the capillary to perform the injection; a voltage is applied to the sample and a few nanoliters of sample are pulled onto the end of the capillary; the default injection is 15 kV (kilovolts) for 5 seconds
- Water wash of capillary – capillary is dipped several times in waste water to remove any contaminating solution adhering to the outside of the capillary
- Water dip – capillary is dipped in clean water (position 2) several times
- Electrophoresis – autosampler moves to inlet buffer vial (position 1) and separation voltage is applied across the capillary; the injected DNA molecules begin separating through the POP-4 polymer solution
- Detection – data collection begins; raw data is collected with no spectral deconvolution of the different dye colors; the matrix is applied during Genescan analysis

Virtual Filters Used in ABI 310

Visible spectrum range seen in CCD camera

<table>
<thead>
<tr>
<th>Filter</th>
<th>Blue</th>
<th>Green</th>
<th>Yellow</th>
<th>Red</th>
<th>Orange</th>
<th>Used with These Kits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Filter A</td>
<td>FL</td>
<td>FAM</td>
<td>JOE</td>
<td>VIC</td>
<td>NED</td>
<td>PET</td>
</tr>
<tr>
<td>Filter C</td>
<td>FL</td>
<td>FL</td>
<td>JOE</td>
<td>VIC</td>
<td>NED</td>
<td>NED</td>
</tr>
<tr>
<td>Filter F</td>
<td>FL</td>
<td>FAM</td>
<td>JOE</td>
<td>VIC</td>
<td>NED</td>
<td>PET</td>
</tr>
<tr>
<td>Filter G5</td>
<td>FL</td>
<td>FAM</td>
<td>VIC</td>
<td>NED</td>
<td>PET</td>
<td>ROX</td>
</tr>
</tbody>
</table>

Commonly used fluorescent dyes

Arrows indicate the dye emission spectrum maximum

Fluorescent Emission Spectra for ABI Dyes

- 5-FAM
- JOE
- NED
- ROX
- NED is a brighter dye than TAMRA

Removal of Dye Artifacts Following PCR Amplification

<table>
<thead>
<tr>
<th>Filter Set</th>
<th>TH01</th>
<th>TPOX</th>
<th>CSF1PO</th>
<th>FGA</th>
<th>D7S820</th>
<th>D21S11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Filter A</td>
<td>FL</td>
<td>FL</td>
<td>JOE</td>
<td>VIC</td>
<td>NED</td>
<td>PET</td>
</tr>
<tr>
<td>Filter C</td>
<td>FL</td>
<td>FL</td>
<td>JOE</td>
<td>VIC</td>
<td>NED</td>
<td>NED</td>
</tr>
<tr>
<td>Filter F</td>
<td>FL</td>
<td>FAM</td>
<td>JOE</td>
<td>VIC</td>
<td>NED</td>
<td>PET</td>
</tr>
<tr>
<td>Filter G5</td>
<td>FL</td>
<td>FAM</td>
<td>VIC</td>
<td>NED</td>
<td>PET</td>
<td>ROX</td>
</tr>
</tbody>
</table>

Butler, J.M. (2001) Forensic DNA Typing, Figure 10.4, ©Academic Press

http://www.cstl.nist.gov/biotech/strbase/training.htm
Why Make a Matrix?

The matrix is the solution to a problem:

What's the contribution at any given wavelength (filter set) from each dye?

If there are 5 dyes

- Remember algebra from high school?
- To solve a problem with 4 unknowns, you need 4 equations

For Example

\[ I_{540} = b x_d + g y_d + y z_d + r w_d \]
\[ I_{560} = b x_g + g y_g + y z_g + r w_g \]
\[ I_{580} = b x_y + g y_y + y z_y + r w_y \]
\[ I_{610} = b x_r + g y_r + y z_r + y w_r \]

Where

\( b \) is the %blue labeled DNA
\( g \) is the %green labeled DNA, etc.
\( x, y, z, w \) are the numbers in the matrix (sensitivity to each color)

If you solve \( x, y, z, w \) for each dye individually
Then you can determine dye contribution for any mixture

5 x 5 matrix for 5-dye analysis on ABI 310

Deciphering Artifacts from the True Alleles

ABI Genetic Analyzer Usage at NIST

- ABI 310 x 2 (originally with Mac, then NT)
  - 1st was purchased in 1996
  - 2nd was purchased in June 2002
  - Not really used any more

- ABI 3100 (Data collection v1.0.1)
  - Purchased in June 2002
  - Original data collection software retained
  - Jan 2007 – upgraded to 3130xl with data collection v3.0

- ABI 3130xl upgrade (Data collection v3.0)
  - Purchased in April 2001 as ABI 3100
  - Upgraded to ABI 3130xl in September 2005
  - Located in a different room

http://www.cstl.nist.gov/biotech/strbase/training.htm
NIST Use of the ABI 3100/3130xl

- Data collection software:
  - version 1.0.1 (Apr 2001-Jan 2007)
  - version 3.0 (Jan 2007-present)

- POP-6 with 36 cm capillary array
  - STR kits and in-house assays for autosomal STRs, Y-STRs, and miniSTRs
  - SNaPshot assays: mtSNPs, Y-SNPs, and autosomal SNPs
  - DNA sequencing for mtDNA and STR repeat sequencing

We can routinely get more than 400 runs per capillary array by not changing the polymer between applications.

NIST ABI 3100 Analysis Using POP-6 Polymer

- High Resolution STR Typing
- SNaPshot SNP Typing
(Coding Region mtSNP 11plex minisequencing assay)

- mtDNA Sequencing (HV1)

Maintenance of ABI 310/3100/3130

- Syringe – leaks cause capillary to not fill properly
- Capillary storage & wash – it dries, it dies!
- Pump block – cleaning helps insure good fill
- Change the running buffer regularly

YOU MUST BE CLEAN AROUND A CE!

Protocols Used for STR Typing

- Most forensic DNA laboratories follow PCR amplification and CE instrument protocols provided by the manufacturer
- Comments
  - Lower volume reactions may work fine and reduce costs
  - No heat denaturation/snap cooling is required prior to loading samples into ABI 310 or ABI 3100
  - Capillaries do not have to be thrown away after 100 runs
  - POP-4 polymer lasts much longer than 5 days on an ABI 310
  - Validation does not have to be an overwhelming task

Questions?

- What are your biggest challenges with keeping your ABI 310/3100/3130 CE systems running?
- What kind of signal intensity variation are you seeing between your different instruments?
- Have you seen uneven injection across a sample plate?

Setting Instrument Parameters and Thresholds

Bruce & John

http://www.cstl.nist.gov/biotech/strbase/training.htm
Spinal Tap Video

- The problem of instrument sensitivity
- Exists everywhere and is fundamental to the concept of signal to noise

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?

Fundamental Ideas behind 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) Peaks below this level may be random noise.

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 and may not be reliable indicators of mixture ratios.

Stochastic Threshold: Level of quantifiable DNA below which peaks can show severe imbalance (peak height ratios below 60%). Approximately 150 - 200 RFUs. (always greater than the LOQ) Variance in peak height ratio is the sum of variance due to the stochastic amplification and instrumental noise.

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

- This is fundamentally an issue of reliability
  - For a peak intensity below the LOD there is a very real chance that such a signal is the result of a random fluctuation.
  - For peaks below the LOQ, the variation in peak height from one run to the next is excessive. These results should not be used in mixture calculations.

Abracadabra! It’s an allele
So how to set thresholds?

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

**Analytical threshold for this instrument is approx. 50 RFUs, LOQ is approximately 180 RFUs.**

Alternate way to define LOD

- Run a blank with peak detection threshold set at 1.
- Determine the standard deviation of the peak heights of all the noise peaks.
- LOD is average intensity + 3x SD of the average intensity. LOQ is the baseline +10x average intensity.
- This technique may produce lower estimate than the previous one.

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.
- In ABI 310 or ABI 3100 detectors, the camera saturates leading to flat-topped peaks.

Useful Range of an Analytical Method

- 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.
The effect of pull-up on an overloaded sample

Matrix Effects
overloaded samples stress the matrix calibration
Matrix problems also affect low level mixtures. Even with the best matrix you can expect 2-5% pull-up!

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

Observation: Peak height variation increases with concentration

Scientific Reasoning behind the Stochastic Threshold
• When stochastic fluctuation is present, interpreting data becomes problematic due to the potential for:
  – Allele dropout
  – Poorly defined mixture ratios
  – Low copy # DNA
• Bottom line: Input levels of DNA should be sufficiently high to avoid straddle data. Mixture interpretation must be done cautiously on low level data as peak intensities are highly variable.

Stochastic Statistical Sampling
Copies of allele 1
Copies of allele 2
True amount
What might be sampled by the PCR reaction…
OR
Resulting electropherogram
>20 copies per allele
6 copies per allele (LCN)

Types of Results at Low Signal Intensity
(Stochastic amplification potential)
Straddle Data
• Only one allele in a pair is above the laboratory stochastic threshold
At low levels of input DNA, the potential for straddle data is high.
The issue is best avoided by re-amplifying the sample at higher input DNA
Otherwise straddle data makes locus inconclusive
Straddle data may be caused by degradation, inhibition and low copy issues.

How low can you go?

http://www.cstl.nist.gov/biotech/strbase/training.htm
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 ratios fall below 60%

Heterozygote Peak Height Ratios

In spite of the improved sensitivity, peak balance is poor at low template concentration

Alternative Procedure (Mass State Police)

1. Since most estimates for LCN show up from 100-250 pg DNA, select a low level sample - say 150 pg as your stochastic limit.
2. Amplify 2 or more samples at a range of concentrations (1.0-0.005) ng multiple times and score the intensity
3. The stochastic limit is the intensity (RFUs) at which half the alleles have intensity above this value and half are below
4. In this way you define straddle data as at the point 50% of your alleles will be above this mark

http://www.cstl.nist.gov/biotech/strbase/training.htm
What else can go wrong?

Yarrr, Take care mates!

- Most validation studies are performed on pristine samples derived from clean sources.
- DNA degradation will result in dropped alleles from larger sized amplicons.
- DNA inhibition will result in dropped alleles from any location and the effects are difficult to predict.
- Inhibition and degradation can produce stochastic effects – peak balance issues and allele dropout.

What can go wrong?

- DNA degradation
- DNA inhibition
- Inhibition and degradation

qPCR Humic Acid Inhibition

- Shift in take off cycle
- Change in melting curve
- No efficiency of amplification change
- Size effects on melt curve

Conclusion: Sequence specific Inhibitor

qPCR Calcium Inhibition

- No shift in take off cycle
- No change in melting curve
- Efficiency of amplification affected
- No difference for size or Tm

Conclusion: Taq Inhibitor

miniSTR Amplification w/ increasing Humic Acid

- D5 Most Affected

miniSTR Amplification w/ increasing Calcium

- D5 Least Affected
Fuzzy Logic in Data Interpretation

- The ABI 310/3100 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. These guidelines must be based on in-house validation. In addition the interpretation and its significance cannot be dissociated from the overall facts of the case.

Instrument factors

1. Because only signal is measured (RFUs) in forensic DNA analysis, many labs find that one instrument or another is more sensitive
2. There are also differences in sensitivity based on injection parameters, capillary illumination (single vs multiple) and laser intensity
3. Lastly the variation in qPCR sensitivity affects the output of any system
4. These differences should be corrected by proper setting of threshold parameters.

ABI 310 vs 3100

Sample
- 310 1.5 µL in 24 µL formamide
- 31xx 1 µL in 10 µL formamide

Injection
- 310 5s @15kV = 75 kVs
- 3130 (4 cap) 5s @ 3kV = 15 kVs
- 3100 (16 cap) 10s @ 3kV = 30 kVs

Irradiation
- 310 direct
- 3130 (4 cap) side
- 3100 (16 cap) both sides

**Bottom line:** you would expect to see
1. an approximate 3 fold difference in RFUs between a 310 and a 3130 (4 cap)
2. an approximate 2 fold difference between a 310 and a 3130xl (16 cap)

Additional Issues

- **Detection Threshold (ABI)**
  - 310 50 RFUs
  - 31xx 30 RFUs

- **Stochastic Threshold**
  - 310 150 RFUs
  - 31xx 90 RFU

- **Dynamic Range**
  - 310 4500 RFUs
  - 31xx 3500 RFUs

**Bottom line:** 310 will appear more sensitive with a wider dynamic range unless proper validations are performed.

Issues with Data

Below the Stochastic 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

Bottom Line

Validate each class of instrument and expect differences in sensitivity/signal to noise
Compensate for differences by choosing appropriate thresholds
Validate at 2 or more injection levels so that injection time can be increased- remembering that longer injections risk drifting into LCN regime

Calling thresholds involve sensitivity, dynamic range and the necessity to avoid low-copy number (LCN) data
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

Comparison of STR Kit Amplification SOP with LCN Using the Same DNA Donor


Stutter and Peak Height Ratios change at LCN

- Stutter increases with allele size:
- Stutter increases at low copy #

Low copy number situations exist in many samples

- In a 1:1 mixture, each DNA source is at LCN when the total amount of DNA in the amplification reaction is ~ 0.125 ng.
- In a 1:9 mixture, the minor component could be at LCN even when the total amount of DNA in the amplification is 1 ng.

Robin Cotton, AAFS 2003 LCN Workshop

“Are we already doing low copy number (LCN) DNA analysis?”

http://www.cstl.nist.gov/biotech/strbase/training.htm
Some interpretational guidelines with LCN

- At least two* PCR amplifications from the same DNA extract
  *five is better; results are investigative
- An allele cannot be scored (considered real) unless it is present at least twice in replicate samples
- Extremely sterile environment is required for PCR setup to avoid contamination from laboratory personnel or other sources

Meatloaf Principle

- I want you
  I need you
  But — there ain't no way I'm ever gonna love you
  Now don't be sad
  cause two out of three ain't bad
  — Meatloaf

- You see an allele twice in 3 runs
- What if the the 4th measurement shows no allele?
- Is seeing an allele 50% of the time a measure of reliability. Is 66% ok?

Typical LCN Procedure

- Extract DNA from stain
- Perform 3 Separate PCR Amplifications
- Quantify Amount of DNA Present
- Interpret Alleles Present
- Develop a Consensus Profile (based on replicate consistent results)

Catch 22

- Note the Catch 22. Are two amplifications of 50pg better than 1 of 100pg?
- Are 3 amplifications of 17pg better than one of 50?
- Data shows that the lower the amount of the DNA amplified the more likely allele dropout and false alleles occur
- This somewhat calls in to question the idea that a sample should be split and run multiple times

Problems with Obtaining Correct Allele Calls at Low DNA Levels

<table>
<thead>
<tr>
<th>Percent Typed</th>
<th>DNA Concentration (pg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td>100 pg</td>
</tr>
<tr>
<td>10%</td>
<td>50 pg</td>
</tr>
<tr>
<td>20%</td>
<td>20 pg</td>
</tr>
<tr>
<td>30%</td>
<td>10 pg</td>
</tr>
<tr>
<td>40%</td>
<td>5 pg</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Control</th>
<th>100%</th>
<th>90%</th>
<th>60%</th>
<th>40%</th>
<th>0%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partial</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Failure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sensitivity Series - 32 cycles


Replicate LCN Test Results from FSS


F’ used to designate that allele drop-out of a second allele cannot be discounted when only a single allele is observed (OCME uses “Z”)
Low Copy Number Limitations (cont):
From Bruce Budowle (2005) 1st International Human Identification E-Symposium

- Tissue source cannot be determined
- DNA may not be relevant – casual contact/transfer
- If victim and suspect have any common access...
- Old cases may not be viable – handling
- Not for post conviction analysis
- Rarely useful for database searching
- An intelligence tool


The Report

- No nuclear profile due to insufficient or excessively degraded DNA
- Suspect is excluded based on results for 2 of 17 Y STR markers.
- Huh?!
- My comments
  - 1. The result is clearly at low copy
  - 2. The pattern of alleles is not consistent with degradation as the cause of dropout.
  - 3. At low copy a scientist cannot express a strong opinion about where DNA arrived at the site where it was recovered. This DNA could just as easily come from thin air as it could come from the suspect.
- Bottom line: Why was this sample even run?

The Bottom Line:

1. Low signal levels are bad because:
   a. They may indicate low copy # DNA = inconsistent or misleading results
   b. They often coincide with peak imbalance
   c. PCR and instrumental artifacts appear at these levels
2. Relying on signal level to determine DNA quantity can be misleading
   a. There is wide variation in signal strength of amplified DNA
   b. Inhibitors and mixtures complicate interpretation
      1. peak imbalance can occur even in single source samples due to inhibition and degradation
      2. instruments can vary in sensitivity

Conclusions

- Be conservative in interpretation
  - Set thresholds based on signal to noise and stochastic amplification (2 thresholds). Base these numbers on controlled in-house experiments
  - Understand that different instruments may vary in sensitivity – set thresholds high enough to encompass this variation
  - Understand that even with such guidelines issues such as degradation and inhibition can skew results.
- Leave room for the facts of the sample in your interpretation

Threshold Illustration

Example values (empirically determined based on own internal validation)

150 RFUs
Peak real, can be used for CPE

97 RFUs
Peak should not be called

50 RFUs
Peak not considered reliable

107 RFUs
Peak real, but not used for CPE

Interpretation Threshold (Dropout/Stochastic/LOQ/Reporting)

Analytical Threshold (Reporting/Noise Limit-of-Detection)

Lab interpretation threshold = 125 RFU

Data from Brad Bannon (Duke lacrosse player defense attorney)
Troubleshooting: Strategies and Solutions

2. Sample Issues

- Formamide Conductivity
- Excessive salt in sample due to evaporation
- Metal ion contamination
- Sensitivity issues with Microcon cleanup (salt removal)
- Dye “blobs” – artifacts from primer synthesis

http://www.cstl.nist.gov/biotech/strbase/training.htm
Effect of bad Hi-Dye Formamide - Note broad peaks and extra bands

Problem disappears if same sample dissolved in pure water instead of formamide

What are these extra peaks doing in this ROX Sample? Could it also be formamide?

Here it gets worse!
How does it affect the sample?

In sample - shadow peaks appearing to left of allelic peaks

What does ABI Say?

Incomplete denaturation of standard due to excessive salt in sample or in formamide

ds DNA migrates faster and over time with this set of runs ds DNA replaced the ssDNA
Post PCR manipulation

- Reprocessing post PCR to concentrate samples can improve signal but be careful
  - PCR sample is concentrated but:
    - Spin filtration may result in removal of background salts,
    - This can greatly enhance sensitivity due to the stacking process
    - Best idea- remake sample up in buffer, not water to avoid reading stochastic effects.

3. External Factors

- Room temperature
  - Variations in room temperature can cause mobility shifts with band shifts and loss of calibration
  - Temperature is also important due to effects of high humidity on electrical conductance

- Cleanliness
  - Urea left in sample block can crystallize and catalyze further crystal formation causing spikes, clogs and other problems.
  - Best bet is to keep polymer in system and not remove or change block until polymer is used up.

Temperature effects

- Viscosity – mobility shift
  \[ \mu_{\text{ep}} = \frac{q}{6\pi\eta r} \]
- Diffusion – band broadening
- Conformation – DNA size based sieving
  \[ \mu_{\text{ep}} = \frac{qD}{6\eta} \]
- Current – Power
  \[ P = VI = I^2R \]
  - Increased current \(\Rightarrow\) internal temperature rise \(\Rightarrow\) diffusion \(\Rightarrow\) band broadening

Due to its structure and its non-calibration, the “250” peak can be used to indicate stability

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
"OL Alleles" - look at the 250 peak

"OL allele re-injected"

And the 250 peak...
True off-ladder alleles

Monitoring Room Temperature Over Time

Temperature Probes

Monitoring Instrument Room Temperature Fluctuations

Poor Temperature Control Causes DNA Sizing Imprecision

http://www.cstl.nist.gov/biotech/strbase/training.htm
Temperature plotting of Room A 230

1 hr 15 min cycle

Rapid changing is affecting data collection

Use of Second Allelic Ladder to Monitor Potential Match Criteria Problems

1st Injection (standard for typing)

15th Injection (treated as a sample)

These alleles have drifted outside of their genotyping bins due to temperature shifting over the course of the sample batch

What to do if calibration is lost?
The 310 only calibrates to the first run ladder this ladder sample may have been run at a different temperature!

- If protocol permits
  - Go to the next ladder
  - Rerun sample
  - Check current
  - Check allelic ladder

- Always check the ROX size standard
  - Look for extra bands
  - Check peak height
  - Check parameters and alignment

Cleanliness

- Urea sublimates and breaks down to ionic components - these find a path to ground
- Similarly wet buffer under a vial creates paths to ground
- Capillary windows must be clear or matrix effects will occur
- Laser will often assist in this process
- Vial caps will transfer low levels of DNA to capillary

Carbon Trails

High Humidity or wet buffer vials can create other paths to ground

Keep Your System Clean and Dry!

4. Instrumental Factors

- Optical System
  - Sensitivity changes with age, capillary diameter, capillary cleanliness, instrument calibration

- Fluidic System
  - Effects of bubbles, dust, urea crystals, leaks in syringe and capillary ferrule

- Matrix Calculations
  - Changes in buffer, optics, sample dye can alter the software calibrations

- Capillary Problems
  - Chemisorbed materials on capillary surface can produce osmotic flow, DNA band broadening and inconsistent resolution (meltdowns)
Consider the Optical System


Issues with the Optical System

- Argon ion lasers outgas and eventually lose intensity; take note of laser current and monitor it over time
- Fluorescence expression:
  \[ I = I_0kbc\phi \]
  - changes in input intensity: \( I_0 \)
  - changes in capillary diameter: \( b \)
  - cleanliness of capillary, optics: \( k \)
  - All these things directly affect peak RFUs, however, baseline noise is more affected by detector.
- Thus by monitoring signal to noise, you can get a better picture of your optical system.

The Detection Window

Make sure that the capillary window is lined up (if it is not, then no peaks will be seen)

Window may need to be cleaned with ethanol or methanol

Buffer Issues

- The buffer and polymer affect the background fluorescence- affecting the matrix
- Urea crystals and dust may produce spikes
- High salt concentrations may produce reannealing of DNA
- High salt concentrations affect current
- Low polymer concentrations affect peak resolution

Remove all bubbles from the channels

Bubbles in the channels can prevent flow of ions and are usually exhibited by zero current when the voltage is applied
Separation problems, bubbles in capillary

These spikes resulted from buffer dilution with poor water. The problem disappeared when the HPLC grade water was purchased to dilute buffer and samples.

Beware of Urea Crystals

Urea crystals have formed due to a small leak where the capillary comes into the pump block.

Urea sublimates and can evaporate to appear elsewhere.

Use a small balloon to better grip the ferrule and keep it tight.

Pump block should be well cleaned to avoid problems with urea crystal formation.

Current Spikes

Generally appear in all lanes and are sharper than regular peaks.

These are a natural consequence of the application of high voltage in CE and may also be caused by particulates in buffer.

Storage when ABI 310 is not in use

- Keep inlet of capillary in water...if it dries out then urea crystals from the polymer will clog the opening.
- The waste vial (normally in position 3) can be moved into position.
- A special device can be purchased from Suppelco to rinse the capillary off-line.
- Store in distilled water.
- Note that the water in the open tube will evaporate over time...Also this will destroy the electrode if turned on without removing the tube.

Remember that the water in the open tube will evaporate over time...Also this will destroy the electrode if turned on without removing the tube.

Matrix Problems

- A poor matrix can lead to raised baseline and therefore calling of too many peaks.
- Larger sized alleles will not be identified as peaks because the GeneScan table for a particular dye color has filled up.
Butler/McCord – Promega 2008 Workshop  
Troubleshooting Common Laboratory Problems  
October 16, 2008

**Capillary Meltdowns**

(A) Good resolution  
(B) Poor resolution

**ABI Solution to Polymer Problem**

The transmission rate of our investigation suggests that most of the POP polymer may be contributing to some of the repeat problems and, as such, additional efforts have been directed at synthesizing a polymer that is not as sensitive to metal ions and that can be used over a longer period of time. We have also added to the polymer a solution of triethanolamine (TEA) to improve the resolution of the polymer. This helps to reduce the amount of contamination in subsequent analyses.

**Effect of contaminant in reference sample**

Contamination results in problems in subsequent analyses.  
Effect is transitory

**Metal ions in the Sample**

DNA clumps and injects poorly. Effect is pH and EDTA dependent

**Meltdowns can be the result of**

- Bad formamide
- Bubbles in the sample vial
- Water in the polymer buffer
- Syringe leak or bottom out
- Poisoned capillary
- Conductive polymer buffer due to urea degradation
- Crack/shift in capillary window
- Detergents and metal ions in sample
A permanent loss of resolution may mean:

- Adsorptive sites on a capillary
- Initiation of electroosmotic flow
- Conductivity changes in buffer/polymer
- Wrong buffer formulation
- Bad formamide or internal lane standard
- Contaminated syringe

5. Troubleshooting benchmarks

- Monitor run current
- Observe syringe position and movement during a batch
- Examine ILS (ROX) peak height with no sample
- Observe "250 bp" peak in GS500 size standard
- Monitor resolution of TH01 9.3/10 in allelic ladder and size standard peak shapes
- Keep an eye on the baseline signal/noise
- Measure formamide conductivity
- Reagent blank – are any dye blobs present?
- See if positive control DNA is producing typical peak heights (along with the correct genotype)

Measurement of Current

- \( V/I = R \) where \( R \) is a function of capillary diameter, [buffer], and buffer viscosity

- In a CE system the voltage is fixed, thus changes in resistance in the capillary will be reflected in the current observed

- Air bubbles, syringe leaks, alternate paths to ground, changes in temperature, changes in zeta potential, and contamination, will be reflected in the current observed

- A typical current for a CE system with POP4 buffer is 8-12 µA (microamps)

Syringe Travel

- The ABI 310 instrument also keeps track of the position of the syringe (in the log file)

- Depending on the resistance to flow, the syringe will travel different lengths

- Syringe leaks may be reflected in a longer distance traveled prior to each injection

- These leaks occur around the barrel of the syringe and at the connection to the capillary block
ROX Ladder QC procedures

- A recommended sequence for initial operation of the 310
  - Rox ladder – initial injection - throwaway
  - Rox ladder- QC to test peak intensity and look for problems in blank
  - Allelic ladder- to determine resolution and to provide standard
  - 10-15 samples
  - Allelic ladder
  - 10-15 samples
  - Allelic ladder

Measurement of Signal and Noise Ratio

- You can also use the ROX size standard to keep track of sensitivity
  - For a given set of runs determine the average peak height of the ROX standard
  - Monitoring this signal level will help determine if any major loss of sensitivity has occurred
  - You can also measure the P-P noise level in the same way and compare the two values.

Measuring Formamide Conductivity

The key is to measure the bottle when it comes in or buy the good stuff and immediately pipette it out into small tubes with or without ROX already added. Then freeze the tubes.

Do not ever open a cold bottle of formamide. Water will condense inside and aid in the formation of conductive formic acid.

Conclusion:

Troubleshooting is more than following the protocols

It means keeping watch on all aspects of the operation
1. Monitoring conductivity of sample and formamide
2. Keeping track of current and syringe position in log.
3. Watching the laser current
4. Watching and listening for voltage spikes
5. Monitoring room temperature and humidity
Review of Some Specific Problems

John & Bruce

Specific Problems

- Post-PCR oligo hybridization to eliminate vWA artifacts with PowerPlex 16 results
- Allele dropout with MiniFiler D16S539 (vs Identifier)
- Resolution loss due to 3130xl pump failure
- Signal loss (data gap) from 3100 laser shutter sticking
- Some variant and tri-alleles

Recent Promega Solution to Eliminating vWA Artifacts in PowerPlex 16 Results

Several laboratories have reported the occurrence of a split or n-1 peak at the vWA locus in PowerPlex 16... The root cause of this artifact is post-PCR reannealing of the unlabeled, unincorporated vWA primer to the 3’-end of the tetramethylrhodamine (TMR)-labeled strand of the vWA amplicon. This reannealing occurs in the capillary post-electrokinetic injection. The split peak is eliminated by incorporation into the loading cocktail of a sacrificial hybridization sequence (SHS) oligonucleotide that is complementary to the vWA primer. The SHS preferentially anneals to the primer instead of the TMR-labeled strand of the vWA amplicon...

Impact of Added Oligos to vWA Amplicon Peaks

Artifacts are eliminated by addition of complementary oligonucleotides after PCR

Different Genetic Tests Can Give Different Results Based on PCR Primer Positions

Mutations in the DNA Sequence (impact PCR primer annealing)

"Null" Allele from Allele Dropout
New Null Allele Section on STRBase

http://www.cstl.nist.gov/biotech/strbase/NullAlleles.htm


MiniFiler Discordance SRM 2391b Genomic 8 with D16S539

All allele calls with MiniFiler for CSF1PO, D7S820, D13S317, D18S51, D21S11, FGA, and D16S539 (with the exception noted below) match previously certified values.

PowerPlex 16

All allele calls with MiniFiler for CSF1PO, D7S820, D13S317, D18S51, D21S11, FGA, and D16S539 (with the exception noted below) match previously certified values.

D16S539 SRM 2391b Genomic 8

T→C mutation 34 bp downstream of the repeat

Position of the T→C probably affects the reverse primer of MiniFiler and is the 3rd base found the 5'end of the Reverse PP16 primer. This could explain the imbalance of the allele seen when using PP16.

What we call “melt downs”… probably due to an incompletely filled capillary

Does the capillary need to be replaced?

No! The next injection looks fine…

http://www.cstl.nist.gov/biotech/strbase/training.htm
Identifiler Allelic Ladder
March 14, 2007

Data from Becky Hill (NIST)

Examination of Resolution in TH01 Region

Data from Amy Decker (NIST)

Examine the Size Standard...

Processed Data (GS500 LIZ size standard)

Raw Data (Identifiler allelic ladder)

Data from Becky Hill (NIST)

When Pump Failed...

3130xl before pump failure
TH01 allelic ladder
3130xl after pump change

Data from Amy Decker (NIST)

http://www.cstl.nist.gov/biotech/strbase/training.htm
Poor Resolution and Peak Tailing

Data from Amy Decker (NIST)

Diluted Sample in Fresh Hi-Di Formamide

Data from Amy Decker (NIST)

Comparison of GS500 LIZ Size Standard

Sample diluted in Old Hi-Di Formamide

Data from Amy Decker (NIST)

Comparison of ABI 3100 Data Collection Versions

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

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

http://www.cstl.nist.gov/biotech/strbase/training.htm
3100 phenomena

what happened on one instrument during the last year...(Sept 2006-2007)

Data from Walter Parson’s Lab (Innsbruck, Austria)

3100 phenomena

• “Septa” – phenomenon
  – buffer and water reservoir septa
  – septa of the 96-well plates

• “Shift” – phenomenon
  – slow runs
  – fast runs

• “Shining through” – phenomenon
• “Carry-over” – phenomenon
• “Data gap” – phenomenon

Data from Walter Parson’s Lab (Innsbruck, Austria)

“Septa” – phenomenon

• with buffer and water reservoir septa
  – bad CE after about three injections

1. run
2. run
3. run

Data from Walter Parson’s Lab (Innsbruck, Austria)

“Septa” – phenomenon

• septum of the 96-well plate + buffer and water reservoir septa: bad matrix runs (spectral calibration)

with septa

without septa

Data from Walter Parson’s Lab (Innsbruck, Austria)

“Shift” – phenomenon

slow runs

• Shifted runs each 3. and/or 5. run (up to 400 data points measured at the 50bp-Std)
• Problem: “off ladder alleles”
• possible cause: bad POP6 Lot

regular run
slow run

Data from Walter Parson’s Lab (Innsbruck, Austria)

“Shift” – phenomenon

slow runs

• 1. run after starting the 3100 are shifted (up to 400 data points measured at the 50bp-Std)
• Problem: “off ladder alleles”
• cause: not explained yet – but no temperature problem

1. run (slow)
2. run (regular)

Data from Walter Parson’s Lab (Innsbruck, Austria)
"Shift" – phenomenon

- 1. run after starting the 3100 are shifted (up to 400 data points measured at the 50bp-Std)
- Problem: "off ladder alleles"
cause: Capillary fill not possible – sources of errors: syringes, check valve, capillary sleeve are leaking but no error message by the software.

"Shining through" phenomenon

the samples shine through into other capillaries systematically:
- blue box: sample (can 8000 rfu)
- blue letters: numbers of capillaries that show the profile "shining through"
cause: the laser shutter opens again wrongly for a short time after closing

"Carry-over" – phenomenon

cause: repeated injections using the same 96-well plate septum;
other causes possibly involved (e.g. no array filling because of defective syringes)
“Data gap” - phenomenon

STRs

Y-filer sample with “invisible” gap

Y-filer sample with “detectable” gaps (zoomed in)

Data from Walter Parson’s Lab (Innsbruck, Austria)

“Data gap” - phenomenon

sequences

“invisible” gap

Data from Walter Parson’s Lab (Innsbruck, Austria)

There are gaps in the data records

cause: the laser shutter is closed too long

these gaps can be detectable in the profile or are concealed by the software (“invisible” gaps)

Data from Walter Parson’s Lab (Innsbruck, Austria)

Signal Loss Due to Laser Power Supply Shutdown


SGM Plus Kit Results

PowerPlex 16

vWA 16,16

TH01 8,8,9.3,9.3

vWA 16,17

TH01 8,8,9.3

A double discordance discovered?

The laser power turned off for ~1 sec at the precise time that the vWA allele 17 and TH01 allele 8 migrated past the detector. The data collection software spliced together the .fsa file without any warning messages to the user.

ABI 3100 Avant with GeneMapperID software

Variant Alleles

• “Off-ladder” Alleles within a locus

• “Off-ladder” Alleles between loci

• Tri-Allelic Patterns

http://www.cstl.nist.gov/biotech/strbase/training.htm
An Example of an “Off-Ladder” Microvariant at the Yfiler Locus DYS635

http://www.cstl.nist.gov/biotech/strbase

Lab Resources and Tools
- Addresses for scientists working with STRs
- Training Materials
- STR Allele Sequencing

STRbase has a summary of alleles that have been submitted and sequenced, if the submitting agency agrees to share the information.

We require a minimum of 10 ng for the sequencing.

We request copies of the electropherograms demonstrating the variant allele.

The more information we have up front the better.

Please have patience we will get to your samples!

Sample Submissions
- For those that desire more assurances of confidentiality we can have MOUs signed.
- We generally re-type the samples at NIST prior to starting sequencing.
- We may run a monoplex assay (single locus).
- We return results as PowerPoint slides.
- We thank all of those agencies that have used this free service (thanks to NIJ!)

Contact Margaret Kline: margaret.kline@nist.gov

Characterizing a Variant Allele That Occurs Between Two Loci
- Use a different multiplex STR kit with different locus combinations
- Test singleplex for each putative locus
- Example: Identifier D16S539 and D2S1338

Penta D 10, Variant Allele 19

All sequenced bases align before and after the repeat region. The 19 allele has been previously reported in STRBase. The Penta D ladder has Alleles 2.2, 3.2, 5, 7 – 17 represented.
Steps to Detection of Which Locus an Out-of-Range Allele Belongs With...

- Consider locus heterozygosities – heterozygote is likely from locus with higher heterozygosity (e.g., D16 = 0.766 while D2 = 0.882)

- Remember that tri-allelic patterns and homozygotes are less common than heterozygotes – thus two heterozygotes are more likely than a homozygote next to a tri-allelic pattern

- Check STRBase for variant alleles reported previously by other labs (e.g., D16 has no >16 alleles while D2 has several <15 alleles)

- Consider genotype frequencies observed for the various possible combinations (e.g., D16 11,11 = 10.7% while D2 20,20 = 0.92%)

---

Three-Peak Patterns


D18S51

\[ \text{Type 1} \]

Sum of heights of two of the peaks is equal to the third

Most common in D18S51 and TPOX

TPOX

\[ \text{Type 2} \]

Balanced peak heights

Most common in TPOX and D21S11

D21S11

---

TPOX Tri-Allelic Patterns

FSI Genetics 2008; 2(2): 134-137

The nature of tri-allelic TPOX genotypes in African populations

A.B. Luke


Approximately 2.4% of indigenous South Africans have three rather than two TPOX alleles. Data collected during routine paternity testing revealed that the extra allele is almost always allele 10 and that it segregates independently of those at the main TPOX locus. Approximately twice as many females as males have tri-allelic genotypes which suggested that the extra allele is on an X chromosome.

---

TPOX Tri-Allelic Patterns Reported on STRBase

http://www.cstl.nist.gov/biotech/strbase/var_TPOX.htm#Tri

- 6,8,10 (4x)
- 6,9,10 (5x)
- 6,10,11 (4x)
- 7,8,10 (2x)
- 7,9,10 (1x)
- 7,10,11 (2x)
- 8,9,10 (14x)
- 8,10,11 (19x)
- 8,10,12 (4x)
- 8,11,12 (3x)
- 9,10,11 (11x)
- 9,10,12 (2x)
- 10,10,11 (1x)
- 10,11,12 (4x)
- 10,10,12 (1x)
- 10,11,12 (4x)
- 10,11,12 (4x)

In 78 observations of 16 different TPOX tri-allelic patterns, only 4 times (5%) is allele “10” not present

---

http://www.cstl.nist.gov/biotech/strbase/training.htm
Variant Alleles Cataloged in STRBase
http://www.cstl.nist.gov/biotech/strbase/var_tab.htm

Off-Ladder Alleles

Tri-Allelic Patterns

Currently 457 at 13/13 CODIS loci
- F13A01, FES/PPS, Penta D, Penta E, D2S1338, D19S433

Currently 176 at 13/13 CODIS loci
- FES/PPS, Penta D, Penta E, D2S1338, D19S433

Thank you for your attention!

Contact Information

John Butler
- NIST (Gaithersburg, MD)
- Phone: 301-975-4049
- Fax: 301-975-8505
- Email: john.butler@nist.gov
- Website: http://www.cstl.nist.gov/biotech/strbase/

Bruce McCord
- FIU (Miami, FL)
- Phone: 305-348-7543
- Fax: 305-348-3772
- Email: mccordb@fiu.edu
- Website: http://www.fiu.edu/~mccordb/

http://www.cstl.nist.gov/biotech/strbase/training.htm