Capillary Electrophoresis in DNA Analysis

Real-time PCR and miniSTRs

NEAFS Workshop
Mystic, CT
September 29-30, 2004

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

Outline for Workshop

- Introductions
- STR Analysis
- Introduction to CE and ABI 310
- Data Interpretation
  - Additional Topics – Real-time PCR and miniSTRs
- Higher Throughput Approaches
- Troubleshooting the ABI 310 (Participant Roundtable)
- Additional Topics – Y-STRs, validation, accuracy
- Review and Test

Why is Accurate DNA Quantitation Important in Forensic DNA Testing?

- Limited amount of DNA available
  - Usually cannot perform multiple tests for quantity
  - Want to preserve DNA for STR testing
- Optimal signal from multiplex STR reactions is only in a tight concentration range (usually 0.5-2 ng)
  - Too much DNA leads to split peaks, off-scale peaks, and bleed through between dye colors
  - Too little DNA leads to loss of loci or alleles due to stochastic effects

Calculation of DNA Quantities in Genomic DNA

Table values for calculations:

- 1 bp = 618 g/mol A, 313 g/mol T, 304 g/mol G, 289 g/mol C
  - A-T base pairs = 617 g/mol

1 genome copy = ~3 x 10^9 bp = 23 chromosomes (one member of each pair)

1 mole = 6.02 x 10^23 molecules

Standard DNA typing protocols with PCR amplification of STR markers typically ask for 1 ng of DNA template. How many actual copies of each STR locus exist in 1 ng?

1 genome copy = (3 x 10^9 bp) x (618 g/mol/bp) = 1.85 x 10^12 g/mol
= (1.85 x 10^12 g/mol) x (1 mole/6.02 x 10^23 molecules)
= 3.08 x 10^-12 g = 3.08 picograms (pg)

Since a diploid human cell contains two copies of each chromosome, then
each diploid human cell contains ~6 pg genomic DNA

∴ 1 ng genomic DNA (1000 pg) = ~333 copies of each locus (2 per 167 diploid genomes)


Introduction

- What is rtPCR or qPCR?
- How does it work?
- How does it compare to traditional methods of Human DNA quantitation?
- What techniques are available?
- What systems are available?

History

- rtPCR is a very recently developed technique
  - Developed by Higuchi in 1993
  - Used a modified thermal cycler with a UV detector and a CCD camera
  - Ethidium bromide was used as intercalating reporter As [dsDNA] increased fluorescence increased

  Higuchi, R.; Fockler, C.; Dollinger, G.; Watson, R. “Kinetic PCR analysis: real-time monitoring of DNA amplification reactions” Biotechnology (N Y). 1993 Sep;11(9):1026-30

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
**Figure 3. Mechanism of PCR**

D. Voet, J. Voet, and C. Pratt, Fundamentals of Biochemistry, 1999

**PCR amplification**

- Theoretically the quantity of PCR template T doubles with each cycle.
- After 2 cycles the quantity of product is $2T$.
- After N cycles the quantity of product is $P = (2)^N T$.
  - Thus there is an exponential relationship between the original quantity of product and the amount of template.

**What is qPCR?**

- To use PCR as a Quantitative technique, the reaction must be clearly defined.
- In fact there are several stages to a PCR reaction:
  - Baseline stage
  - Exponential stage
  - Plateau stage

**PCR plateaus**

- PCR product can not double forever
  - Limited by
    - Amount of primer
    - Taq polymerase activity
    - Reannealing of product strands
- Reach plateau
  - No more increase in product
- End point detection
  - Run for fixed # cycles and then quantify on agarose gels
Problem #2: For endpoint detection, how many cycles should you do?
Different wells reach plateau at different cycle numbers. When you look, changes what you see.

Issues for quantitation by non RT-PCR methods
- In spite of its use in mixture resolution, PCR is not technically a quantitative technique
- The time and rate at which plateau appears varies with temperature, tube position, inhibitors, matrix
- Once plateau appears, increase in product concentration is non linear
- Standards can be added but they must have the same primer binding sites and similar sequence to target

Solution
- Use data when still in exponential phase
  - PCR product proportional to initial template
- Need to look at PCR product each cycle
  - Use fluorescent detection, where fluorescence is proportional to PCR product
- Use real time PCR machine which records fluorescence for each well at each cycle

Quantitation using the PCR Reaction
- PCR proceeds exponentially doubling each cycle:
  \[ Y_n = Y_{n+1}(1+E_c) \]
  Where \( E_c \) is the efficiency (\( E_c = 1 \) for a perfect amplification) and \( Y_n \) is the yield of product for a particular cycle
- During the exponential stage of the reaction
  \( E_c \) is relatively constant and the reaction yield \( Y \) is a function of the quantity of input DNA, \( X \)
  \[ Y = X(1+E_c)^n \]

Effect of efficiency on [DNA]
- \( E_c \) is a function of:
  - Hybridization efficiency
  - Quantity of reactants/target DNA
  - Temperature

Real Time PCR
- Quantitation of DNA is a based on the number of cycles required to reach a threshold intensity, \( C_t \)
- The greater the amount of starting DNA, the sooner this threshold value is reached.
Quantitation using $C_t$

- The log of DNA template concentration vs $C_t$ is plotted using a series of stds yielding a calibration curve.
- The unknown is then run and the number of cycles required to reach threshold, $C_t$ is compared to the calibration curve.

Development of a standard curve

The output data is plotted on a log scale and the fractional # cycles required to reach $C_t$ is measured.

Standard curve

Plot the cycle # at threshold CT vs concentration

Concentration = $10^{(-0.297 \times CT + 4.528)}$

Detection Methods

- Fluorescent intercalating dye - SYBR Green
  - Fluorescence increases with concentration of dsDNA
- Taqman probes
  - Fluorescence increases as quenched probe is digested
- Molecular beacons
  - Fluorescence increases as quenched probe hybridizes to template

SYBR green product detection

- Easy
  - Fluorescence only with dsDNA
  - Use with existing PCR primers
- Generic
  - Detects all double stranded products, including primer dimers
  - However, can be very specific with proper primer design
- Singleplexed
  - Multiple probes cannot be used
Molecular beacons

- Consist of ssDNA with an internal complementary sequence that keeps reporter and quencher dyes close → No fluorescence
- Following denaturation, beacon anneals to template, separating both dyes and yielding fluorescence proportional to PCR product concentration

Molecular Beacons

- Improved specificity and multiplexing
  - Non-specific amplification will not produce a signal
  - Can multiplex several probes (quantify nuclear, Y, int std.)
- Can be tricky to design
  - Loop portion – binds to DNA template
  - Stem portion – must be complementary to other stem
  - Probe must denature from template below 72º so Taq polymerase does not chew it up during extension step

Above T_m loop structure reforms and probe leaves template

Tanneal< T_m < T_ext

Taqman

Probe also binds to PCR product during extension but is always quenched
- 5'–3' exonuclease activity of Taq polymerase digests probe and frees reporter dye from quencher
- Free dye accumulates with PCR product

Probes vs SYBR Green

- SYBR Green
  - Singleplex probes (Alu)
  - If no sample, amplification of contaminants occurs at high cycle #
  - If inhibition, no result or poor efficiency curve
- Probes (Taqman, Mol. beacons)
  - Multiplex targeted probes – Quant Y, nuclear DNA, int std
  - Inhibition and no sample can yield no result (if single locus probe)
  - to check for inhibition, an internal std. is used
- Choice: Simplicity (SYBR green) vs Multiplexing (probes)

Single vs Multilocus Targets

- SYBR Green – Multilocus Probe
  - Alu inserts occur at multiple locations throughout the genome - sensitive
  - If no sample, amplification of contaminants occurs at high cycle #
  - Syber green requires no special kit -inexpensive
- Probes (Taqman, Mol. beacons)
  - Single location in genome
  - an internal std. is used to check for amplification and correct for changes in efficiency
  - Lower sensitivity due to noise at low copy number
- Choice: Sensitivity (SYBR green) vs Internal Standard Precision (probes)

Effects of Inhibitors on Alu Assay

- Use Alu sequence, present at 1,000’s of copies/cell
  - Assay is sensitive to ambient human DNA in air and water
  - Normal Reagent blanks have a Ct at about 27-29 cycles
- If inhibitors are present – no amplification occurs or efficiency is altered
  - Thus low level ambient DNA serves as an internal control for inhibitors
- For non Alu based RtpCR, an internal standard is required to detect inhibition
Slot blot versus real time PCR

Slot Blot
- 2 days of rinses, incubations, pipettings, washes, exposures, and developments
- Semiquantitation by manual comparison or through scanner
- Quantity obtained may not reflect final result due to variations in PCR efficiency
- 2 hours setup and run time
- Automated quantitation
- Quantity obtained reflects amplifiable result

Big issue – sensitivity and dynamic range

Quantiblot-ECL 40 pg - 2.0 ng
ACES 2.0 40 pg - 4.0 ng
No longer available
(ACES tended to work better on degraded DNA)

Real Time PCR 1.0 pg - 16 ng

RTPCR has lower detection limit and larger dynamic range

Work in OhioU/FIU Laboratory
- Development of miniplex STRs for degraded DNA typical sizes 60-120 bp.
- Slot blot works poorly on these samples
- Current assay has 124bp Alu product – reduction to 84bp testing in progress

Comparison Studies Slot blot vs RTPCR

<table>
<thead>
<tr>
<th>Sample</th>
<th>Slot Blot</th>
<th>Real Time PCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blood on stick</td>
<td>0.32</td>
<td>0.50</td>
</tr>
<tr>
<td>Blood on metal</td>
<td>0.40</td>
<td>0.50</td>
</tr>
<tr>
<td>Blood on concrete</td>
<td>0.40</td>
<td>0.50</td>
</tr>
<tr>
<td>Blood on leaves</td>
<td>0.08</td>
<td>0.20</td>
</tr>
<tr>
<td>Blood on cardboard</td>
<td>0.27</td>
<td>0.24</td>
</tr>
<tr>
<td>Blood on cloth</td>
<td>0.54</td>
<td>0.68</td>
</tr>
<tr>
<td>Blood on denim</td>
<td>0.25</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Calibration studies in our lab with experimental primers

Validation work of Jan Nicklas and Eric Buel

RT-PCR Instruments Cited
- Corbett Research Rotorgene
  - Phenix Research, Hayward, CA
- ABI 7000 Sequence Detection System
- ABI 7700 (discontinued)
- ABI 7900HT Sequence Detection System
  - Applied Biosystems Foster City, CA
Real-Time PCR Efforts

- Eric Buel – Alu system (JFS 2003, 48(5):936-944)
- Centre for Forensic Sciences – nuclear; TH01 flanking region (JFS 2003, 48(5):1041-1046)
- John Hartmann – Alu system (SWGDAM Jan 2003)
- CA-DOJ – TH01 assay (NIST DNA Grants June 2003)
- SYBR Green assay – human-specific with right PCR
- Quantifiler kit (ABI) – separate nuclear and Y assays

NIST Lessons Learned from Real Time-PCR Assays

- Results are RELATIVE to standards used
- Single source and mixed source samples with same UV concentrations differ with RT-PCR assays
- Need to keep instrument clean to avoid background fluorescence problems
- Assay reagent costs:
  - Quantifiler: $2.46/sample (only permits 2 µL/sample)
  - SYBR Green: $0.80/sample (up to 10 µL/sample)
  - QuantBlot: $0.54/sample (5 µL/sample)

Information from Quantifiler Kit Manual

<table>
<thead>
<tr>
<th>Sample</th>
<th>Sex</th>
<th>$A_{260}$ Result (µg/µL)</th>
<th>Obq Result (µg/µL)</th>
<th>% DIFF from $A_{260}$</th>
<th>% DIFF from Obq</th>
<th>Result (µg/µL)</th>
<th>% DIFF from $A_{260}$</th>
<th>% DIFF from Obq</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>M</td>
<td>17.5</td>
<td>29</td>
<td>6.99</td>
<td>61.7</td>
<td>66.6</td>
<td>10.13</td>
<td>41.9</td>
</tr>
<tr>
<td>2</td>
<td>M</td>
<td>14.9</td>
<td>44</td>
<td>14.3</td>
<td>7.1</td>
<td>28.5</td>
<td>16.78</td>
<td>9.0</td>
</tr>
<tr>
<td>3</td>
<td>M</td>
<td>33.9</td>
<td>39</td>
<td>15.45</td>
<td>11.4</td>
<td>40.4</td>
<td>14.30</td>
<td>2.9</td>
</tr>
<tr>
<td>4</td>
<td>M</td>
<td>11.4</td>
<td>19</td>
<td>6.69</td>
<td>61.7</td>
<td>66.6</td>
<td>10.13</td>
<td>41.9</td>
</tr>
<tr>
<td>5</td>
<td>M</td>
<td>16.3</td>
<td>20</td>
<td>13.9</td>
<td>7.1</td>
<td>28.5</td>
<td>16.78</td>
<td>9.0</td>
</tr>
<tr>
<td>6</td>
<td>M</td>
<td>13.9</td>
<td>19</td>
<td>15.45</td>
<td>11.4</td>
<td>40.4</td>
<td>14.30</td>
<td>2.9</td>
</tr>
<tr>
<td>7</td>
<td>M</td>
<td>11.8</td>
<td>17</td>
<td>6.69</td>
<td>61.7</td>
<td>66.6</td>
<td>10.13</td>
<td>41.9</td>
</tr>
</tbody>
</table>

The different methods produced similar quantification results.

Table 6-11 Average differences from $A_{260}$ and Quantifiler kit

- Average (%)

Table 6-12 Comparison with $A_{260}$ and Quantifiler Kit

Conclusions

- RTPCR is a homogeneous PCR based method for human specific quantification
  - Is easily automated, provides electronic storage of data
  - SYBR green or targeted probes can be used
- Results give quantity of amplifiable DNA
  - not necessarily overall quantity
  - Inhibition can be detected
  - Multiplexing can be used
- Big advantages are speed and dynamic range

Acknowledgements

- Jan Nicklas and Eric Buel - Vermont Crime Laboratory
- Jiri Drabek
- Denise Chung, Kerry Opel
- Nancy Tatarek
- John Butler, Yin Shen
- Major support provided by:
  - The National Institute of Justice
  - The OU Provost’s Undergraduate Research Fund
  - Ohio University Research Incentive Fund

References

- On-line
  - http://www.med.sc.edu:85/pcr/realtime-home.htm
  - http://info.realtimeprimers.org/
  - http://dna.chars.med.ucw.edu/realtime.htm
  - http://dorakmt.tripod.com/chemistry/realtime.htm

- In Print
  - Jordan, J. Real time detection of PCR products and microbiology, Trends in microbiology 2000, 12, pp. 61-66
miniSTRs

STR Size Reduction
Through Moving Primer Positions
Closer to the Repeat Region

miniSTR Work

- miniSTRs (a.k.a. BodePlexes) are being used successfully in WTC effort
- Collaboration between John Butler and Bruce McCord (NIJ-funded) to further develop reduced size STR amplicons
- Mike Coble (NRC postdoc) at NIST is developing new miniSTR loci that are unlinked to CODIS loci

Describes new primer sequences for all CODIS loci and initial assays developed


Large Multiplex Kits provide Efficient and Rapid Analysis of Convicted Offender Samples

But what about degraded DNA?

Such samples present a special challenge

The Miniplex Approach to Degraded DNA

1. Redesign primers to make each STR amplicon as short as possible.
2. Avoid overlap by having only 1 STR locus in each dye lane.
3. Provide an alternative to mtDNA for degraded DNA template.
4. Develop of specialized STR systems for degraded DNA.

PowerPlex 16 Result on Aged Blood Stain
(15 years at room temperature storage)

"Decay curve" of degraded DNA

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
DNA Degradation

1. Strand breakage
2. Pyrimidine dimers
3. Chemical oxidation and hydrolysis
4. Bacterial degradation and metal contamination

The Problem with Degraded DNA

Degraded DNA

1. Fragmentation due to the environment
2. The presence of PCR inhibitors

Result

1. Poor amplification efficiency
2. Peak imbalance and allele dropout

Current Miniplex Loci and reduction in size relative to ABI kits

<table>
<thead>
<tr>
<th>Dye Label</th>
<th>Blue</th>
<th>Green</th>
<th>Yellow</th>
</tr>
</thead>
<tbody>
<tr>
<td>Miniplex 1</td>
<td>TH01 -105</td>
<td>CSF1P0 -191</td>
<td>TPOX -148</td>
</tr>
<tr>
<td>Miniplex 2</td>
<td>D5S818 -53</td>
<td>D8S1179 -37</td>
<td>D16S539 -152</td>
</tr>
<tr>
<td>Miniplex 3</td>
<td>FGA -71</td>
<td>D21S11 -33</td>
<td>D7S820 -117</td>
</tr>
<tr>
<td>Miniplex 4</td>
<td>vWA -64</td>
<td>D18S51 -151</td>
<td>D13S317 -105</td>
</tr>
<tr>
<td>Miniplex 5</td>
<td>Penta D -282</td>
<td>Penta E -299</td>
<td>D2S1338 -198</td>
</tr>
</tbody>
</table>

Miniplex Primer Sets

<table>
<thead>
<tr>
<th>Miniplex</th>
<th>TH01</th>
<th>CSF1P0</th>
<th>TPOX</th>
<th>D5S818</th>
<th>D8S1179</th>
<th>D16S539</th>
<th>D13S317</th>
</tr>
</thead>
<tbody>
<tr>
<td>Big Miniplex</td>
<td>FAM</td>
<td>VIC</td>
<td>NED</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Miniplex 1</td>
<td>TH01</td>
<td>CSF1P0</td>
<td>TPOX</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Miniplex 2</td>
<td>D5S818</td>
<td>D8S1179</td>
<td>D16S539</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Miniplex 3</td>
<td>FGA</td>
<td>D21S11</td>
<td>D7S820</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Miniplex 4</td>
<td>vWA</td>
<td>D18S51</td>
<td>D13S317</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

MINI 2 PROFILE - BLOOD

Primer-Dimers

Dye Blobs

Poor Primer Binding

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
The Big Question:
How does it work with degraded DNA?

Approach: Examine the effect of template size on DNA amplification
1) Extract large quantities of DNA from liquid blood
2) Digest with DNaseI
3) Cut sections at different size ranges and amplify
4) Compare with a commercial multiplex kit

DNA Degraded With DNase I

Results

Results of Allele Concordance between Miniplex Primers and Commercial STR Kit Primers

- 532 U.S. population samples (Caucasians, African Americans, and Hispanics) examined with four miniplexes (including all CODIS loci except D8S1108)
- Results compared to identifier STR kit (6,384 genotypes compared)
- 15 discrepancies (0.23%) – most of which occur in 2 loci (D13S317 and VWA) and involve deletions between the two primer sites
- Representative samples from each discrepant locus have been sequenced to locate the polymorphic nucleotides causing the allele dropout

Allele Discordance Resulting from Mutations in Miniplex Primer Binding Site

<table>
<thead>
<tr>
<th>Locus</th>
<th>Origin</th>
<th>Miniplex</th>
<th>Identifier</th>
<th>PP16</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>D13S317</td>
<td>AA</td>
<td>11,13</td>
<td>10,13</td>
</tr>
<tr>
<td>2</td>
<td>D13S317</td>
<td>H</td>
<td>9,14</td>
<td>8,14</td>
</tr>
<tr>
<td>3</td>
<td>D13S317</td>
<td>AA</td>
<td>19,11</td>
<td>9,11</td>
</tr>
<tr>
<td>4</td>
<td>D13S317</td>
<td>H</td>
<td>10,11</td>
<td>9,11</td>
</tr>
<tr>
<td>5</td>
<td>D13S317</td>
<td>H</td>
<td>9,14</td>
<td>9,14</td>
</tr>
<tr>
<td>6</td>
<td>D5S818</td>
<td>AA</td>
<td>11,11</td>
<td>11,12</td>
</tr>
<tr>
<td>7</td>
<td>wVA</td>
<td>AA</td>
<td>16,16</td>
<td>12,16</td>
</tr>
<tr>
<td>8</td>
<td>wVA</td>
<td>AA</td>
<td>18,15</td>
<td>13,18</td>
</tr>
<tr>
<td>9</td>
<td>wVA</td>
<td>AA</td>
<td>15,15</td>
<td>14,15</td>
</tr>
<tr>
<td>10</td>
<td>wVA</td>
<td>AA</td>
<td>15,15</td>
<td>14,15</td>
</tr>
<tr>
<td>11</td>
<td>wVA</td>
<td>AA</td>
<td>17,17</td>
<td>14,17</td>
</tr>
<tr>
<td>12</td>
<td>wVA</td>
<td>AA</td>
<td>17,17</td>
<td>14,17</td>
</tr>
<tr>
<td>13</td>
<td>wVA</td>
<td>AA</td>
<td>19,19</td>
<td>14,19</td>
</tr>
<tr>
<td>14</td>
<td>wVA</td>
<td>AA</td>
<td>19,19</td>
<td>14,19</td>
</tr>
<tr>
<td>15</td>
<td>wVA</td>
<td>AA</td>
<td>19,19</td>
<td>14,19</td>
</tr>
</tbody>
</table>

Point Mutation?

- Original Primers
- Redesigned Primers
- No Mutation = Concordance
- Peak Imbalance
- Allele dropout

Primer Design - Concordance?

- Original Primers
- Redesigned Primers
- No Mutation = Concordance
- Peak Imbalance
- Allele dropout
Proposed mechanism

Deletion outside Miniplex primers-limited region
(samples 1-5)

Mutation/polymorphism in Miniplex primer binding region
(samples 6 - 15)

Examination of Concordance:
African American sample ZT79305

NIST Identifiler data

Ohio U miniSTR data

A deletion outside the miniSTR primers causes the kit produced allele to appear one repeat smaller.
### VWA Primers and Mutations

- **Allele drop out with Miniplex primer**
- **PPlex 16 primers still bind because there is only one base pair difference in the adjacent sequence**
- **ABI reported this as a C-T polymorphism (Lazaruk et al. 2001)**
- **There is actually 2 repeats (8 bp) more than the actual allele called in commercial kits**

### VWA Mutation Region

**CCATCAT CCATCAT**

- Allele drop out with Miniplex primer
- PPlex 16 primers still bind because there is only one base pair difference in the adjacent sequence
- ABI reported this as a C-T polymorphism (Lazaruk et al. 2001)
- There is actually 2 repeats (8 bp) more than the actual allele called in commercial kits

---

### New miniSTR (non-CODIS) Loci Under Investigation

<table>
<thead>
<tr>
<th>Locus</th>
<th>Sequence Motif</th>
<th>Allele Size Range (bp)</th>
<th>Observed Heterozygosity</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1S1677</td>
<td>(GGAA)$_n$</td>
<td>9-18</td>
<td>61-117</td>
</tr>
<tr>
<td>D2S441</td>
<td>(CTA)$_n$</td>
<td>9-17</td>
<td>78-110</td>
</tr>
<tr>
<td>D4S1364</td>
<td>(GAAT)(GGAT)(GAAT)$_n$</td>
<td>8-12</td>
<td>67-83</td>
</tr>
<tr>
<td>D10S1248</td>
<td>(GGAA)$_n$</td>
<td>10-20</td>
<td>83-123</td>
</tr>
<tr>
<td>D14S1434</td>
<td>(GATA)$_n$</td>
<td>13-20</td>
<td>70-98</td>
</tr>
<tr>
<td>D22S1045</td>
<td>(TAA)$_n$</td>
<td>5-16</td>
<td>76-109</td>
</tr>
</tbody>
</table>