Outline of Topics to Discuss

- Introduction to Low Copy Number (LCN) DNA: What is LCN DNA?
- DNA concentration of samples: How low can you go?
- The effects of stochastic PCR amplification
- Challenges and limitations with LCN DNA testing
- LCN data and Peak Height Ratios (PHR)
- Other methods for higher sensitivity and signal enhancements

Some Definitions of Low-Copy Number (LCN)

- Work with $<100$ pg genomic DNA (~15-17 diploid copies of nuclear DNA markers such as STRs)
- Below stochastic threshold level where PCR amplification is not as reliable (determined by each laboratory; typically 150-250 RFUs)
- Enhancing sensitivity of detection (34 cycles instead of 28 cycles)
- Too few copies of DNA template to ensure reliable PCR amplification

LCN is dependent on the amount of DNA present NOT the number of PCR cycles performed; LCN conditions may exist with 28 or 34 cycles

Low-Copy Number (LCN) Work

- Early work on touched objects and single cells:
- Application to routine forensic casework was pioneered by the Forensic Science Service:

LCN is not a "new" technique…

- 1996 – Taberlet et al. describe "reliable genotyping of samples with very low DNA quantities using PCR"
- 1997 - single cell STR analysis reported
- 1999 – Forensic Science Service begins LCN casework in UK (as an alternative to mtDNA)
- 2001 – Budowle and FBI co-authors urge caution with using LCN

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
Why attempt LCN? …

- Improved success rates with high sensitivity DNA testing vs. standard procedures
- Volume crime samples (burglary)
- Bone samples to provide improved matching statistics over mtDNA analysis

Amounts of DNA Required

<table>
<thead>
<tr>
<th>Period</th>
<th>Technique</th>
<th>Amount Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>1985-1995</td>
<td>RFLP/VNTRs</td>
<td>50 ng – 1000 ng</td>
</tr>
<tr>
<td>1991-present</td>
<td>PCR/STRs</td>
<td>0.5 – 2 ng</td>
</tr>
<tr>
<td>1999-present</td>
<td>LCN/STRs</td>
<td>&lt;0.1 ng</td>
</tr>
</tbody>
</table>

LCN extends the range of samples that may be attempted with DNA testing

Where does low copy number start?

<100 pg template DNA


<table>
<thead>
<tr>
<th>Amount of DNA</th>
<th>~ # of cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 ng</td>
<td>152</td>
</tr>
<tr>
<td>0.5 ng</td>
<td>76</td>
</tr>
<tr>
<td>0.25 ng</td>
<td>38</td>
</tr>
<tr>
<td>0.125 ng</td>
<td>19</td>
</tr>
<tr>
<td>0.0625 ng</td>
<td>10</td>
</tr>
</tbody>
</table>

Values for # of cells adjusted to reflect updated DNA quantitation numbers

Impact of DNA Amount into Multiplex PCR Reaction

We generally aim for 0.5-2 ng

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.

Two different amplifications would be useful with a 1:9 mixture situation:

- Normal level of total DNA (e.g., 1 ng) so that major component is on-scale.
- High level of total DNA (e.g., 5 ng) so that minor (e.g., ~500 pg) is out of LCN realm – yes, the major component will be off-scale...

DNA Concentration in Samples: How Low Can You Go?

~ # of cells

10
0.0625 ng
1 ng
5 pg template
2.0 ng
STR Kits Work Best in This Range
50 pg template
0.125 ng
0.01 ng
Too little DNA
Heterozygote peak imbalance
Allele drop-out
Locus-to-locus imbalance
Stochastic effects when amplifying low levels of DNA can produce allele dropout

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
Stochastic PCR Amplification

What is a true peak (allele)?

Peak detection threshold

Peak height ratio (PHR)

Stutter percentage

Noise (N)

Signal (S)

Signal > 3x sd of noise

PHR consistent with single source typically above 60%

Stutter location above 15%

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

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 Fluctuation Effects

• Unequal sampling of the two alleles present in a heterozygous individual can occur when low levels of input DNA are used (results in allele drop-out)

• PCR reactions with <100 pg (~17 diploid copies)

• Walsh et al. (1992) – propose avoiding stochastic effect by adjusting the number of PCR cycles in an assay so that the sensitivity limit is around 20 or more copies of target DNA (i.e., a full profile is obtained with ~125 pg)

Stochastic Statistical Sampling


http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
Stochastic Effect

- Sometimes called “preferential amplification” – not really a correct term since either allele may be amplified if the other drops-out…not related to allele size
- Stutter product amounts may go up…
  - If in an early cycle of PCR, the stutter product is amplified more (due to sampling effect)
- Contaminating DNA can also be amplified giving rise to allele “drop-in” or a mixture

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

Challenges and limitations with LCN DNA testing

Challenges of LCN

- Increased chance for contamination (want a sterile lab environment to reduce staff contamination)
- Data interpretation is more complicated (due to stochastic variation during PCR amplification):
  - Allele drop-out
  - Allele drop-in
  - Increased stutter products
  - Heterozygote peak imbalance

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

Problems with Obtaining Correct Allele Calls at Low DNA Levels

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

Suggestions for Optimal Results with LCN

- At least two* PCR amplifications from the same DNA extract (if enough DNA is present to do more than 4-5 amplifications, then most likely a single aliquot would be run under standard STR typing conditions)
- 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
  *five is better; results are typically viewed as investigative

Typical LCN Procedure

Extract DNA from stain

Perform 3 Separate PCR Amplifications

Interpret Alleles Present

Develop a Consensus Profile (based on replicate consistent results)

New Interpretation Rules Required for LCN

An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA

Peter Gill*, Jonathan Wathkens*, Christine Flexman*, Nick Brown*, John Buckleton*

*Forensic Science Service, Percy House, South Street Yard, Birmingham B34 6BB, UK
FSS, Forensic Bug (2002), Auckland, New Zealand

Received 3 December 1999, received in revised form 12 February 2000; accepted 13 February 2000

Replicate LCN Test Results from FSS

<table>
<thead>
<tr>
<th>Allele</th>
<th>D1</th>
<th>D2</th>
<th>D3</th>
<th>D4</th>
<th>D5</th>
<th>D6</th>
<th>D7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>3.5</td>
<td>15.5</td>
<td>9.5</td>
<td>19.5</td>
<td>28.5</td>
<td>14.5</td>
<td>17.5</td>
</tr>
<tr>
<td>Sample</td>
<td>1</td>
<td>15.5</td>
<td>15.5</td>
<td>9.5</td>
<td>19.5</td>
<td>28.5</td>
<td>14.5</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>15.5</td>
<td>15.5</td>
<td>9.5</td>
<td>19.5</td>
<td>28.5</td>
<td>14.5</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>15.5</td>
<td>15.5</td>
<td>9.5</td>
<td>19.5</td>
<td>28.5</td>
<td>14.5</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>15.5</td>
<td>15.5</td>
<td>9.5</td>
<td>19.5</td>
<td>28.5</td>
<td>14.5</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>15.5</td>
<td>15.5</td>
<td>9.5</td>
<td>19.5</td>
<td>28.5</td>
<td>14.5</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>15.5</td>
<td>15.5</td>
<td>9.5</td>
<td>19.5</td>
<td>28.5</td>
<td>14.5</td>
</tr>
</tbody>
</table>

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")

Example LCN Data
**Experimental Design**

- 3 samples (Caucasian, African American, and Hispanic) that are heterozygous for all loci tested (2 peaks for each locus)
- DNA templates tested: 100 pg, 50 pg, and 10 pg
- Tested in triplicate
- Identifiler kit was used (1/2 reactions)
- Tested with 2 different cycles: 31 and 34

**Data Analysis**

- Determining consensus profile – 2 out of 3 times the allele is observed
- Concordance analysis with samples run with “normal” parameters (non-LCN conditions – 28 cycles) and higher concentrations (at 1 ng DNA)
- Summarizing incorrect allele calls, heterozygote peak imbalance, allele drop-out, locus drop-out, stutter percentages, and non-specific artifacts
Becky Hill – Low-Copy Number DNA Analysis

May 5, 2009

100 pg template DNA with 34 cycles of PCR

Identifier data

(100 pg template DNA with 34 cycles of PCR)

50 pg template DNA with 34 cycles of PCR

Identifier data

(50 pg template DNA with 34 cycles of PCR)

10 pg template DNA with 34 cycles of PCR

Identifier data

(10 pg template DNA with 34 cycles of PCR)

Peak Height Ratio Comparisons

Samples at <1 ng tested in triplicate (std dev shown)

Peak Height Ratio Measurements

Identifier STR Kit – only FGA shown

Signal aided with 31 PCR cycles

Peak Heights (RFUs)

<table>
<thead>
<tr>
<th>Allele PHR imbalance</th>
<th>Average PHR</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 pg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FGA-22</td>
<td>1662</td>
<td>1517</td>
</tr>
<tr>
<td>FGA-23</td>
<td>1915</td>
<td>864</td>
</tr>
<tr>
<td>FGA-25</td>
<td>1239</td>
<td>909</td>
</tr>
<tr>
<td>50 pg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FGA-22</td>
<td>900</td>
<td>260</td>
</tr>
<tr>
<td>FGA-23</td>
<td>1423</td>
<td>419</td>
</tr>
<tr>
<td>FGA-25</td>
<td>896</td>
<td>805</td>
</tr>
<tr>
<td>10 pg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FGA-22</td>
<td>54</td>
<td>107</td>
</tr>
<tr>
<td>FGA-23</td>
<td>130</td>
<td>219</td>
</tr>
</tbody>
</table>

All levels performed in triplicate...

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
Becky Hill – Low-Copy Number DNA Analysis

May 5, 2009

10:1 Female: Male

Input DNA

Identifier Results: NEST I1, I2, I3, I4 (varying input DNA)

Minor component amount

<table>
<thead>
<tr>
<th>Input DNA</th>
<th>Minor components drop out at low levels due to stochastic effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5 ng</td>
<td>150 pg</td>
</tr>
<tr>
<td>1.0 ng</td>
<td>100 pg</td>
</tr>
<tr>
<td>0.5 ng</td>
<td>50 pg</td>
</tr>
<tr>
<td>0.25 ng</td>
<td>25 pg</td>
</tr>
</tbody>
</table>

Data courtesy of Amy Christen, Marshall University NEST Project Team

Other methods for higher sensitivity and signal enhancements

Improving Sensitivity

- Improved recovery of biological material and DNA extraction
- Longer injection on CE
- Salt removal from CE sample – enhances electrokinetic injection
- Reduced volume PCR – concentrates amplicon
- Increase number of cycles in PCR and/or TaqGold concentration
- Use miniSTRs – shorter amplicons amplify better; MiniFiler
- Use mtDNA – higher copy number per cell

Higher Sensitivity with More Polymerase and Cycle Numbers

<table>
<thead>
<tr>
<th>Minor component amount</th>
<th>28 cycles – 1U Taq</th>
<th>32 cycles – 2U Taq</th>
</tr>
</thead>
<tbody>
<tr>
<td>200 pg</td>
<td>miniSTR assay for D10S1248</td>
<td>Allele dropped due to stochastic effects (poor statistical sampling of available chromosomes)</td>
</tr>
<tr>
<td>100 pg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50 pg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20 pg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 pg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 pg</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

From Coble and Butler (2009); J. Forensic Sci. 50: 43-53

Signal Enhancements

- Higher PCR cycles
- Lower PCR volume (problems with inhibitors)
- Brighter fluorescent dyes
- Longer CE injection
  - 10 s @ 3 kV = 30
  - 5 s @ 2 kV = 10
- Microcon cleanup to remove salts that interfere with electrokinetic injection

Modifications in DNA Analysis Process to Improve LCN Success Rates

- Collection – better swabs for DNA recovery
- DNA Extraction – into smaller volumes
- DNA Quantitation – qPCR helps with low DNA amounts
- PCR Amplification – increased number of cycles
- CE Detection – longer electrokinetic injection; more sensitive fluorescent dyes
- Interpretation – composite profile from replicate analyses with at least duplicate results for each reported locus
- Match – is it even relevant to the case?

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
Reduced Volume PCR

• Possibility of lower volume PCR to effectively concentrate the amount of DNA in contact with the PCR reagents

• Can samples be concentrated or can extraction volume be reduced?

miniSTRs and LCN

• miniSTR assays are typically more sensitive than conventional STR kits currently in use

• Labs will start “pushing the envelope” in order to try and get a result with more sensitive assays including future miniSTR assays and kits

• Labs may move into the LCN realm without realizing it or adopting the careful LCN interpretation rules such as replicate analyses with duplicate results prior to reporting alleles

LCN Summary

• LCN often defined as <100-200 pg input DNA

• Typically involves increasing the number of PCR cycles when performing multiplex PCR to amplify DNA with conventional STR kits (e.g., 34 cycles instead of 28 cycles)

• Enables lower amounts of DNA to be detected with STR markers but is prone to contamination

• Cautious data interpretation rules must be adopted as allele drop-out and drop-in may occur due to stochastic amplification effects

Thank you for your attention…

Our team publications and presentations are available at:
http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
http://www.cstl.nist.gov/biotech/strbase
Becky.hill@nist.gov
301-975-4275

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