Low Copy Number (LCN) DNA Analysis

Becky Hill
National Institute of Standards and Technology

Advanced Forensic DNA Concepts Workshop

Mid-Atlantic Association of Forensic Sciences Meeting
Hunt Valley, MD
May 5, 2009

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

Introduction to Low Copy Number (LCN) DNA

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
- Other terms for LCN:
  - Low-level DNA
  - Trace DNA
  - Touch DNA

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

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

Low-Copy Number (LCN) Work

- **Early work on touched objects and single cells**:
  - van Oorschot, R. A. and Jones, M. K. (1997) DNA fingerprints from fingerprints. *Nature*. 387(6635): 767
  - Findlay, I., Taylor, A., Quirke, P., Frazier, R., and Urquhart, A. (1997) DNA fingerprinting from single cells. *Nature*. 389(6651): 555-556
- **Application to routine forensic casework was pioneered by the Forensic Science Service**:
  - Gill, P., Whitaker, J., Flaxman, C., Brown, N., and Buckleton, J. (2000) An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA. *Forensic Sci. Int.* 112(1): 17-40
  - Whitaker, J. P., Cotton, E. A., and Gill, P. (2001) A comparison of the characteristics of profiles produced with the AMPFISTR SGM Plus multiplex system for both standard and low copy number (LCN) STR DNA analysis. *Forensic Sci. Int.* 123(2-3): 215-223
  - Gill, P. (2001) Application of low copy number DNA profiling. *Croatian Medical Journal* 42(3): 229-32

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

1985-1995 **RFLP/VNTRs** **50 ng – 1000 ng**

↓

1991-present (kits since 1996) **PCR/STRs** **0.5 – 2 ng**

↓

1999-present **LCN/STRs** **<0.1 ng**

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

DNA Concentration in Samples: How Low Can You Go?

Where does low copy number start?

<100 pg template DNA

(Butler, 2001, Fregeau & Fourney 1993, Kimpton *et al* 1994)

Amount of DNA	~ # of cells
1 ng	152
0.5 ng	76
0.25 ng	38
0.125 ng	19
0.0625 ng	10

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

Robin Cotton, AAFS 2003 LCN Workshop
"Are we already doing low copy number (LCN) DNA analysis?"

Impact of DNA Amount into Multiplex PCR Reaction

We generally aim for 0.5-2 ng

DNA amount (log scale)

High levels of DNA create interpretation challenges (more artifacts to review)

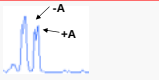
100 ng

10 ng

1 ng

0.1 ng

0.01 ng

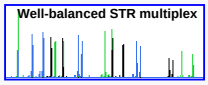


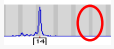
Too much DNA
- Off-scale peaks
- Split peaks (+/-A)
- Locus-to-locus imbalance

2.0 ng

STR Kits Work Best in This Range

0.5 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

Low copy number situations exist in many samples

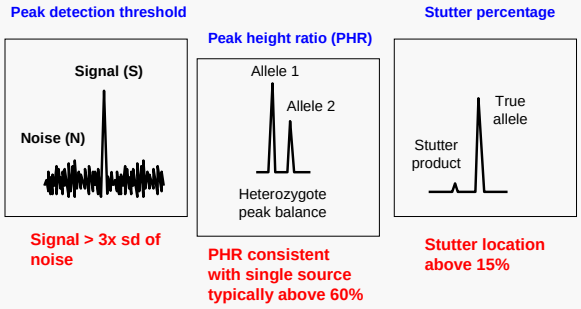
- In a 1:1 mixture, each DNA source is at LCN when the total amount of DNA in the amplification reaction is ~ 0.125 ng.
- In a 1:9 mixture, the minor component could be at LCN **even when the total amount of DNA in the amplification is 1 ng**.

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…

Robin Cotton, AAFS 2003 LCN Workshop
"Are we already doing low copy number (LCN) DNA analysis?"

Stochastic PCR Amplification

What is a true peak (allele)?



Signal > 3x sd of noise

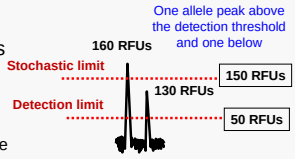
PHR consistent with single source typically above 60%

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**

Straddle data may be caused by degradation, inhibition and low copy issues.

Scientific Reasoning behind the **Stochastic Threshold**

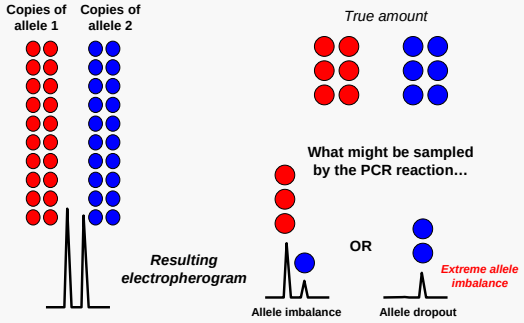
- **When stochastic fluctuation is present, interpreting data becomes problematic due to the potential for:**
  - **Allele dropout**
  - **Poorly defined mixture ratios**
  - **Low copy # DNA**
- Bottom line: Input levels of DNA should be sufficiently high to avoid straddle data. Mixture interpretation 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)

Walsh PS, Erlich HA, Higuchi R. Preferential PCR amplification of alleles: Mechanisms and solutions. *PCR Meth Appl* 1992; 1:241-250.

Stochastic Statistical Sampling



>20 copies per allele

6 copies per allele **(LCN)**

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

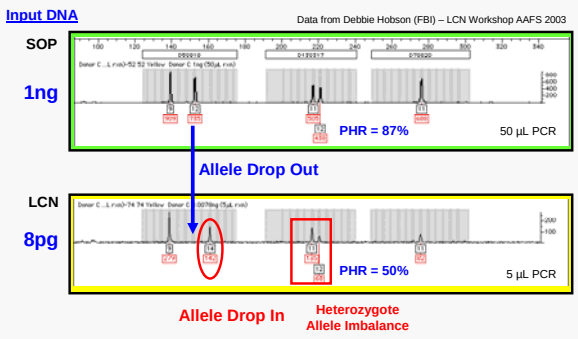
Challenges of LCN

Gill, P. (2001) *Croatian Med. J.* 42(3): 229-232

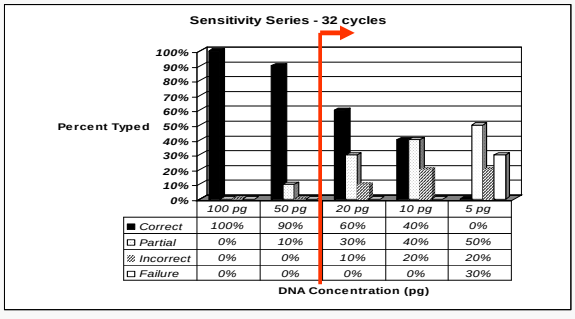
- Increased chance for contamination (want a sterile lab environment to reduce staff contamination)
- Data interpretation is more complicated (due to stochastic variation during PCR amplification):
  - Allele drop-out
  - Allele drop-in
  - Increased stutter products
  - Heterozygote peak imbalance

Comparison of STR Kit Amplification SOP with LCN

Using the Same DNA Donor



Problems with Obtaining Correct Allele Calls at Low DNA Levels



Coble, M.D. and Butler, J.M. (2005) *J. Forensic Sci.* 50: 43-53

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

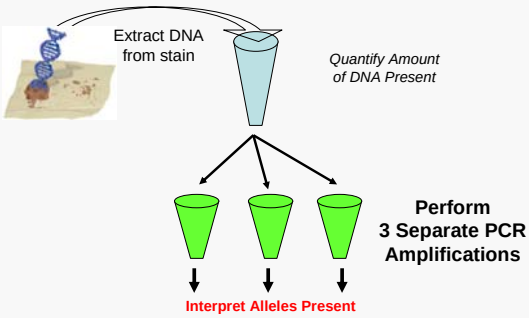
•Aids in determination if a suspect is excluded as a contributor

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



Develop a Consensus Profile
(based on replicate consistent results)

New Interpretation Rules Required for LCN

ELSEVIER

Forensic Science International
112 (2000) 17–40

Forensic Science International

www.elsevier.com/locate/forsciint

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

Peter Gill[a,*], Jonathan Whitaker[a], Christine Flaxman[a], Nick Brown[a], John Buckleton[b]

[a]Forensic Science Service, Priory House, Gooch Street North, Birmingham B56QQ, UK
[b]ESR, Private Bag 92021, Auckland, New Zealand

Received 9 December 1999; received in revised form 12 February 2000; accepted 13 February 2000

Replicate LCN Test Results from FSS

Gill, P. (2002) Role of short tandem repeat DNA in forensic casework in the UK--past, present, and future perspectives. *BioTechniques* 32(2): 366-385.

Table 2. Results of Six Replicate PCR Tests of a Sample Under Low Copy Number Analysis Conditions Compared to the Control Sample

	Amelo	D19	D3	D8	THO	VWA	D21	FGA	D16	D18	D2
CONTROL	X X	14,14	18,18	15,15	7 9.3	19,19	28 32.2	20,23	9,12	12,16	17,23
Sample											
1	--	14 F'	--	15 F'	--	--	28 32.2	20 F'	--	16 F'	--
2	X F'	--	18 F'	15 F'	--	19 F'	--	--	12 F'	--	--
3	X F'	--	--	15 F'	--	--	--	--	--	--	17 F'
4	X F'	14 F'	18 F'	--	--	--	--	--	9 12	--	--
5	X F'	--	18 F'	--	--	18 F'	--	--	--	--	--
6	X F'	14 F'	--	--	--	19 F'	28 32.2	20 F'	--	12 F'	--
Consensus	X F'	14 F'	18 F'	15 F'	--	19 F'	28 32.2	20 F'	12 F'	--	--

The consensus result is reported, provided that an allele is observed at least twice. If only one allele is observed, then an F' designation is given to denote the possibility of allele drop-out.

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

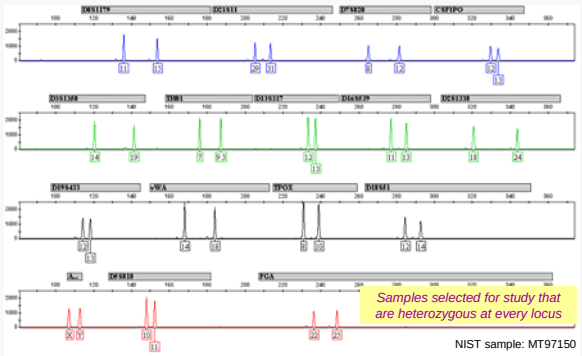
- 3 samples (Caucasian, African American, and Hispanic) that are heterozygous for all loci tested (2 peaks for each locus)
- DNA templates tested: 100 pg, 50 pg, and 10 pg
- Tested in triplicate
- Identifiler kit was used (1/2 reactions)
- Tested with 2 different cycles: 31 and 34

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

1 ng template DNA, 28 cycles

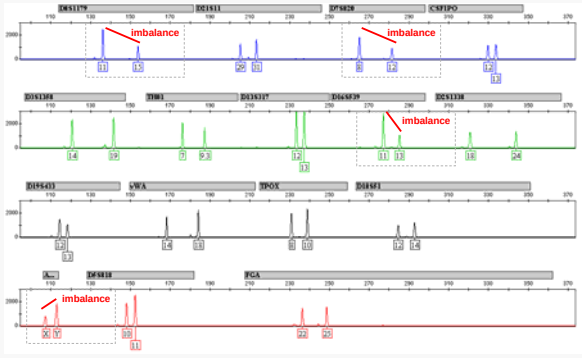
Identifiler data (full profile)



Peak Height Ratios (PHRs) all >0.80

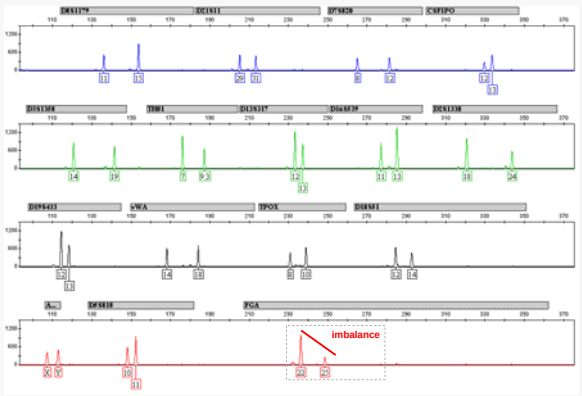
100 pg template DNA with 31 cycles of PCR

Identifiler data (full profile)



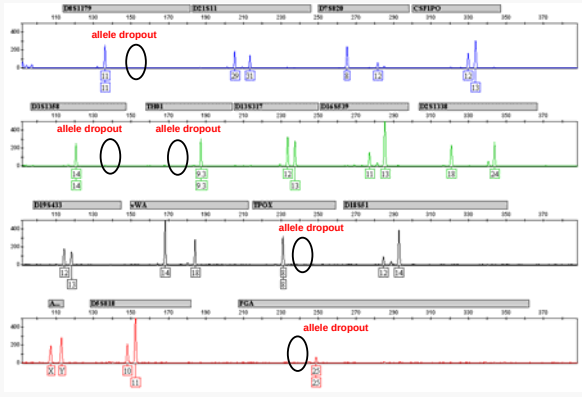
50 pg template DNA with 31 cycles of PCR

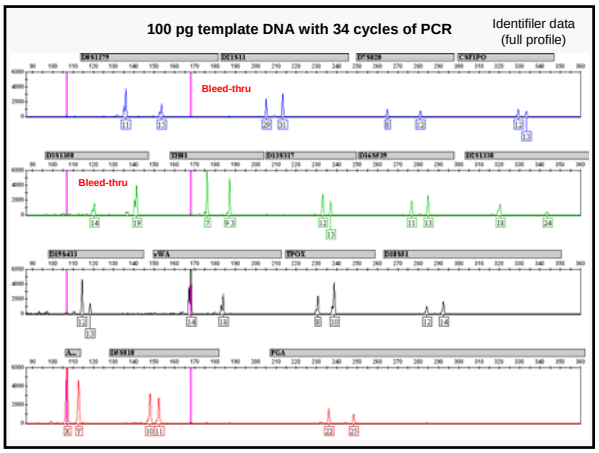
Identifiler data (full profile)

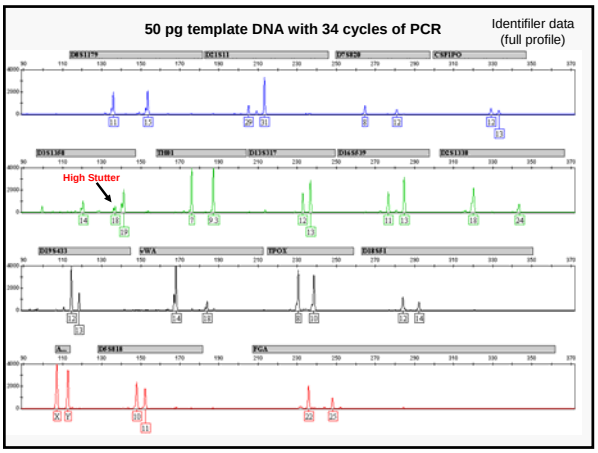


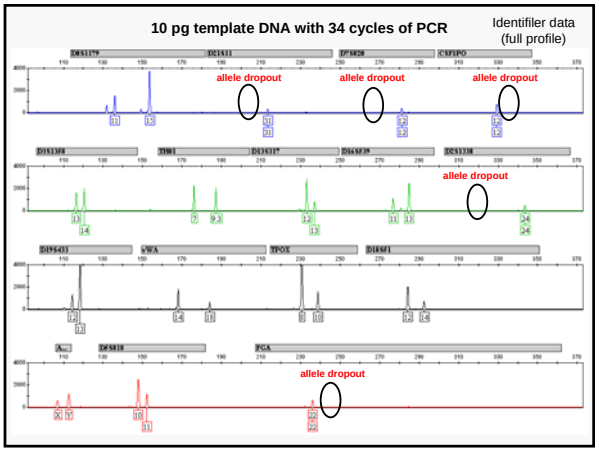
10 pg template DNA with 31 cycles of PCR

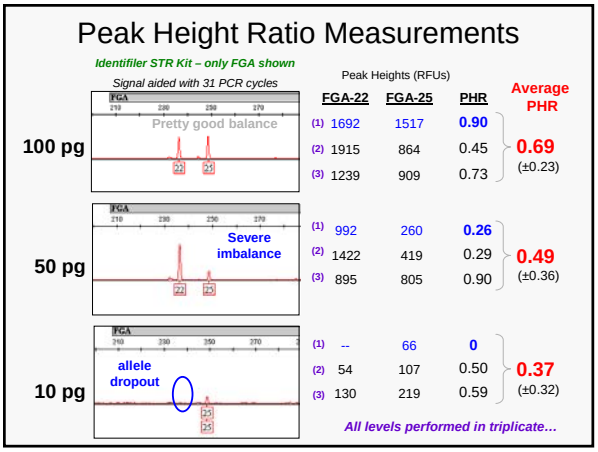
Identifiler data (full profile)

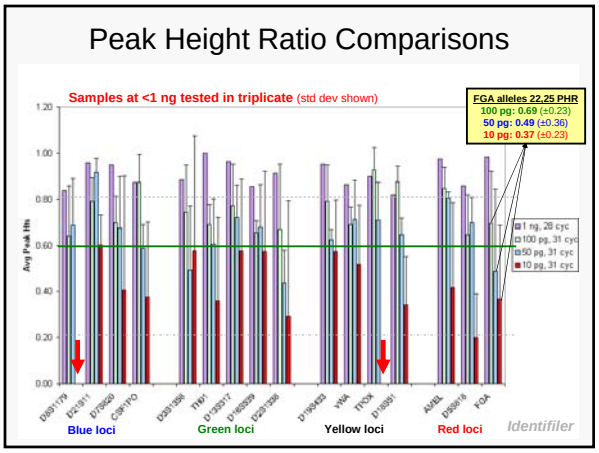


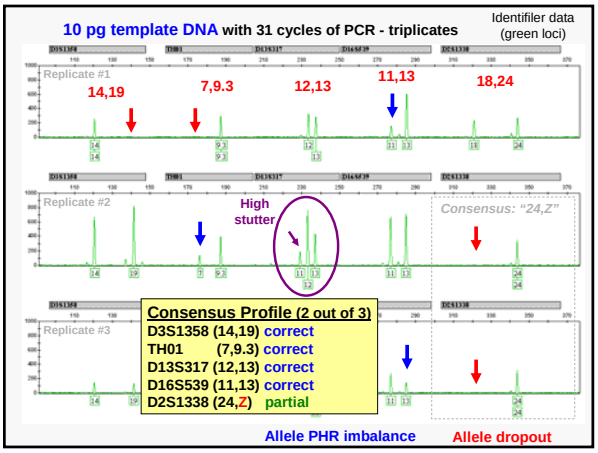

100 pg template DNA with 34 cycles of PCR
Identifiler data (full profile)
Bleed-thru
Bleed-thru

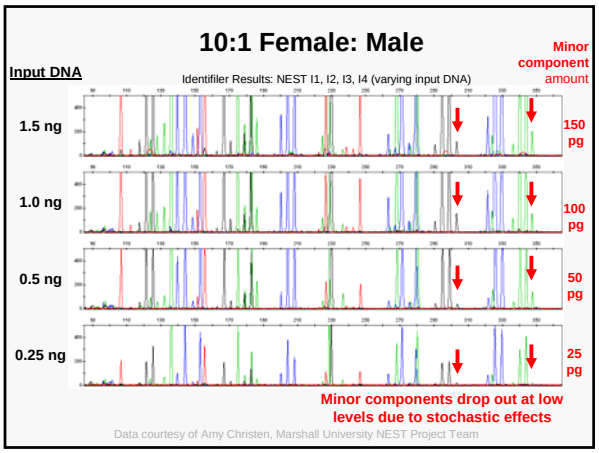

50 pg template DNA with 34 cycles of PCR
Identifiler data (full profile)
High Stutter


10 pg template DNA with 34 cycles of PCR
Identifiler data (full profile)
allele dropout
allele dropout
allele dropout
allele dropout
allele dropout


Peak Height Ratio Measurements
Identifiler STR Kit – only FGA shown
Signal aided with 31 PCR cycles
Peak Heights (RFUs)
FGA-22 FGA-25 PHR
Average PHR
100 pg
Pretty good balance
(1) 1692 1517 0.90
(2) 1915 864 0.45
(3) 1239 909 0.73
0.69 (±0.23)
50 pg
Severe imbalance
(1) 992 260 0.26
(2) 1422 419 0.29
(3) 895 805 0.90
0.49 (±0.36)
10 pg
allele dropout
(1) -- 66 0
(2) 54 107 0.50
(3) 130 219 0.59
0.37 (±0.32)
All levels performed in triplicate...


Peak Height Ratio Comparisons
Samples at <1 ng tested in triplicate (std dev shown)
FGA alleles 22,25 PHR
100 pg: 0.69 (±0.23)
50 pg: 0.49 (±0.36)
10 pg: 0.37 (±0.23)
Avg Peak Hts
1 ng, 28 cyc
100 pg, 31 cyc
50 pg, 31 cyc
10 pg, 31 cyc
Blue loci
Green loci
Yellow loci
Red loci
Identifiler


10 pg template DNA with 31 cycles of PCR - triplicates
Identifiler data (green loci)
Replicate #1
14,19
7,9.3
12,13
11,13
18,24
Replicate #2
High stutter
Consensus: "24,Z"
Replicate #3
Consensus Profile (2 out of 3)
D3S1358 (14,19) correct
TH01 (7,9.3) correct
D13S317 (12,13) correct
D16S539 (11,13) correct
D2S1338 (24,Z) partial
Allele PHR imbalance
Allele dropout

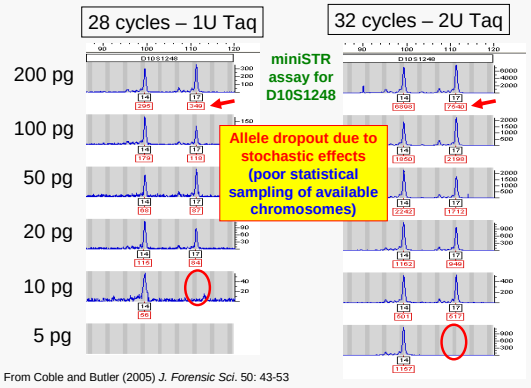


Other methods for higher sensitivity and signal enhancements

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



From Coble and Butler (2005) *J. Forensic Sci*. 50: 43-53

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?

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

Reduced Volume PCR

- Possibility of lower volume PCR to effectively concentrate the amount of DNA in contact with the PCR reagents
  - Gaines *et al*. (2002) *J. Forensic Sci*. 47(6):1224-1237
  - Leclair *et al*. (2003) *J. Forensic Sci*. 48: 1001-1013
- Can samples be concentrated or can extraction volume be reduced?

Leclair et al. (2003) JFS 48:1001-1013

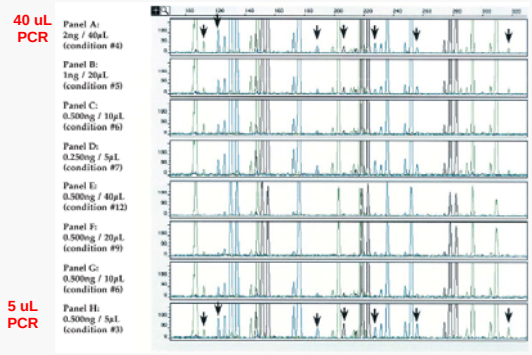


FIG. 5—*Effects of a reduction of PCR reaction volume and DNA template concentration on amplification of a casework sample with a minor profile representing 2% of the total mixture.*

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

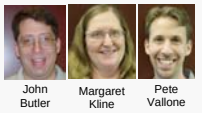
LCN Summary

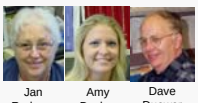
- LCN often defined as <100-200 pg input DNA
- Typically involves increasing the number of PCR cycles when performing multiplex PCR to amplify DNA with conventional STR kits (e.g., 34 cycles instead of 28 cycles)
- Enables lower amounts of DNA to be detected with STR markers but is prone to contamination
- 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


John Butler | Margaret Kline | Pete Vallone


Jan Redman | Amy Decker | Dave Duewer

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