


Potential Application of Forensic DNA Testing Methods to Cancer Diagnostics

Peter M. Vallone and Michael D. Coble
National Institute of Standards and Technology, Gaithersburg, MD

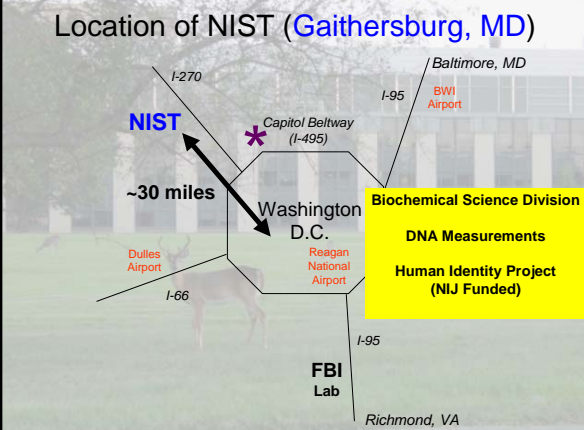
February 17, 2006
Fred Hutchinson Cancer Research Center
Seattle, WA

Outline

- STR loci and kits
- Real Time PCR for DNA quantitation
- SNP markers
- Case example
- Interlaboratory Studies
- NIST Standard Reference Materials (SRMs)
- miniSTRs
- Sample Quality Control

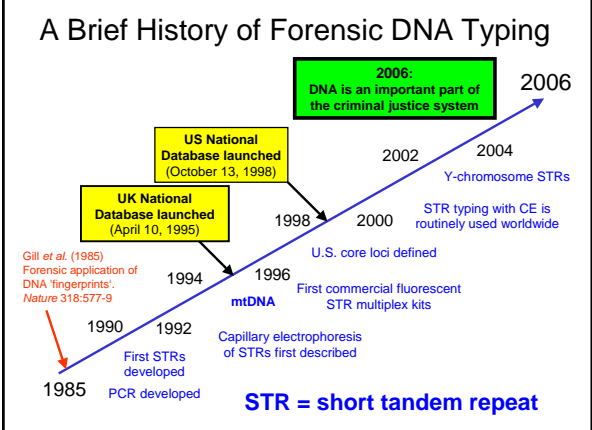


Location of NIST (Gaithersburg, MD)



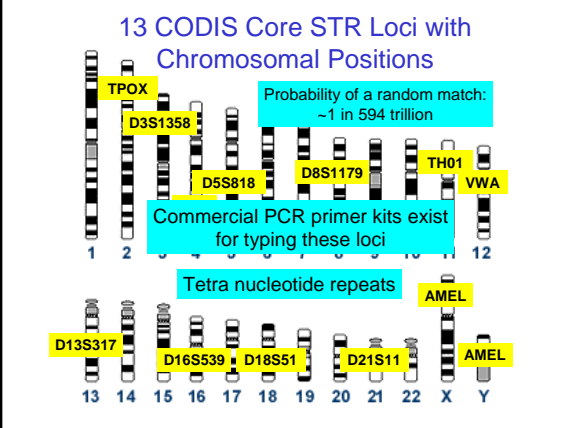
Biochemical Science Division
DNA Measurements
Human Identity Project (NIJ Funded)

A Brief History of Forensic DNA Typing



STR = short tandem repeat

13 CODIS Core STR Loci with Chromosomal Positions



Probability of a random match:
~1 in 594 trillion

Commercial PCR primer kits exist for typing these loci

Tetra nucleotide repeats

Forensic DNA Testing

- Probe subsets of genetic variation in order to differentiate between individuals
- DNA typing must be done efficiently and reproducibly (information must hold up in court)
- Typically, we are not looking at genes – little/no information about race, predisposal to disease, or phenotypical information (eye color, height, hair color) is obtained

Applications for Human Identity Testing

- Forensic cases - **matching suspect with evidence**
- Paternity testing - **identifying father**
- Historical investigations
- Missing persons investigations
- Mass disasters - **putting pieces back together**
- Military DNA "dog tag"
- Convicted felon DNA databases
- Genealogy (Y-STRs, Y-SNPs, mtDNA)

Common Challenges

- Sample extraction
- DNA Quantitation (0.5 - 2ng)
- PCR inhibition (humic acid, Ca⁺⁺)
- Degraded Samples (fragmented)
- Low copy number (stochastic effects)
- Contamination
- Mixtures
- **Other**

Steps in DNA Analysis

AT
TA

Collection
Extraction
Quantitation
Genotyping
Interpretation of Results
Database Storage & Searching

Collection & Storage (Blood Stain) DNA Extraction Slot Blot PicoGreen DNA Quantitation

Multiplex PCR Amplification

STR Typing

Male: 13,14-15,16-12,13-10,13-15,16

Interpretation of Results DNA Database

Identifiler STR Kit

Information is tied together with multiplex PCR and data analysis

Probability of a random match: 1 in $\sim 2.1 \times 10^{17}$

15 STR loci + Amelogenin

Run on ABI 3100

Companies Supply Allelic Ladders in STR Kits to Aid Interlaboratory Consistency

Profiler Plus kit allelic ladders (Applied Biosystems)

AMEL, D8S1179, D21S11, D18S51, D5S818, D13S317, D7S820

GS500 ROX internal standard

Same DNA sample run with Applied Biosystems STR Kits


Random Match Probability

Blue	1.0 x 10 ⁻³
Green I	7.8 x 10 ⁻⁴
Profiler™	9.0 x 10 ⁻¹¹
Profiler Plus™	2.4 x 10⁻¹¹
COfiler™	2.0 x 10 ⁻⁷
SGM Plus™	4.5 x 10 ⁻¹³

PCR Product Size (bp)

Outline

- STR loci and kits
- Real Time PCR for DNA quantitation
- SNP markers
- Case example
- Interlaboratory Studies
- NIST Standard Reference Materials (SRMs)
- miniSTRs
- Sample Quality Control



Steps in DNA Analysis

[Steps in DNA Analysis](#)

Collection

Extraction

Quantitation

Genotyping

Interpretation of Results

Database Storage & Searching

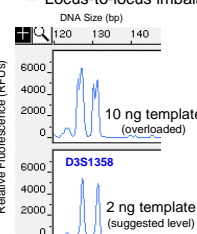
- It is important that the optimal amount of DNA is added to the PCR
- Typically 0.5 ng to 2 ng works well
- Too much or too little DNA will result in artifacts obscuring data interpretation
- DNA can be quantitated using UV spectroscopy, fluorescence (after the addition of an intercalation dye), hybridization with a probe, and **qPCR**

Impact of DNA Amount into PCR

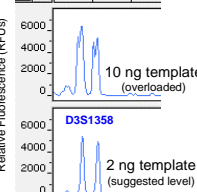
Reason that DNA Quantitation is Important Prior to Multiplex Amplification

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

- Too little DNA
 - Heterozygote peak imbalance
 - Allele drop-out
 - Locus-to-locus imbalance



10 ng template (overloaded)



2 ng template (suggested level)

Stochastic effect when amplifying low levels of DNA produces allele dropout.

Real Time qPCR

- Various methods/kits exist for qPCR
 - Currently evaluating
- SYBR green and TaqMan
- Single and multi copy loci
- All results rely on a Calibrant of a "known" concentration
- Duplex/Triplex for nuclear, mito, and Y
- We are in the process of developing a DNA quantitation reference material

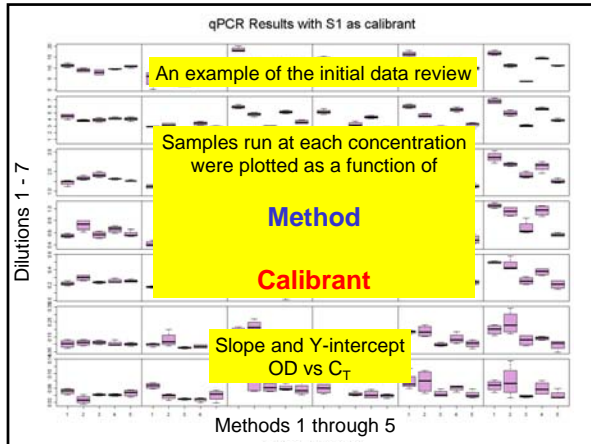
qPCR Methods Evaluated at NIST

- Quantifiler Human (TaqMan MGB)
- Quantifiler Y Male (TaqMan MGB)
- Alu (SYBR Green)
- CA DOJ nDNA (TaqMan BHQ)
- CFS HumTH01 (TaqMan MGB)

1. Quantifiler™ Human DNA Quantification Kit PN4343895
2. Quantifiler™ Y Human Male Quantification Kit PN4343906
3. Nicklas J, Buel E. J Forensic Sci 2003; 48:936-944.
4. Timken M, Swango K, Orrego C, Buoncrisiani M. J Forensic Sci, (Sept 2005)
5. Richard ML, Frappier RH, Newman JC. J Forensic Sci 2003;48:1041-1046.

qPCR Methods

Method	Amplicon (bp)	Target
Quantifiler Human	62	Human telomerase reverse transcriptase gene (hTERT), 5p15.33
Quantifiler Y Male	64	Sex determining region Y gene (SRY)
Alu	124	Alu , Ya5 Subfamily (multi copy)
CA DOJ	170-190	TH01, 11p15.5
CFS HumTH01	62	Flanking region of TH01, 11p15.5



Outline

- STR loci and kits
- Real Time PCR for DNA quantitation
- SNP markers
- Case example
- Interlaboratory Studies
- NIST Standard Reference Materials (SRMs)
- miniSTRs
- Sample Quality Control

SNP Typing Instrumentation

PCR & primer extension

Luminex Beads hybridization

Primer Extension

TaqMan

Allele-Specific Primer Extension

SNP Primer is extended by one base unit

“tail” used to vary electrophoretic mobility

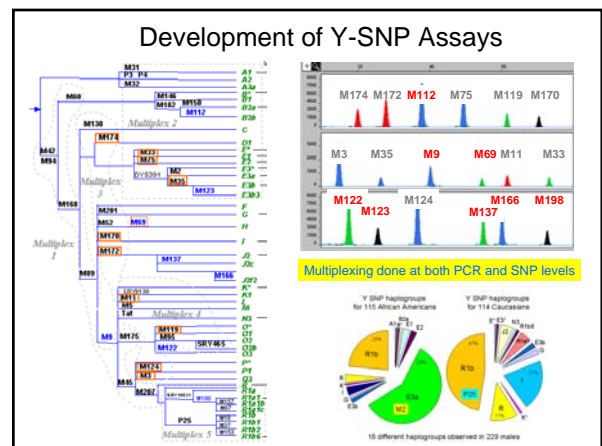
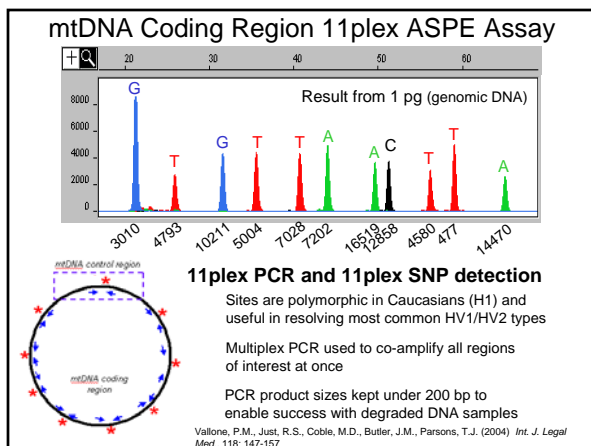
Oligonucleotide primer 18-28 bases

Fluorescently labeled ddNTPs + polymerase

PCR Amplified DNA Template

Multiplex PCR is used to amplify 1-2 ng of gDNA

Loci are typed with multiplex primer extension reactions




U.S. Population Dataset

260 Caucasians, 260 African Americans, 140 Hispanics, 3 Asians = **663 males**

DNA extracted from whole blood (anonymous; self-identified ethnicities) received from Interstate Blood Bank (Memphis, TN) and Millennium Biotech Inc. (Ft. Lauderdale, FL)

To date: (>100,000 allele calls)

- Identifiier (15 autosomal markers + Amelogenin) (10,608)
- Roche Linear Arrays (HV1/HV2 10 regions) (6,630)
- Y STRs 22 loci—27 amplicons (17,388)
- Y STRs 22 new loci (14,535)
- Yfiler kit 17 loci (11,237)
- Y SNPs 50 markers on sub-set of samples (11,498)
- Orchid 70 autosomal SNPs on sub-set (13,230)
- miniSTR testing-new loci and CODIS concordance (9,228)
- New miniSTR loci – for 11 loci, 7,293 genotypes
- mtDNA full control region sequences by AFDIL**




Stock tubes

↓

Genotypes with various human identity testing markers

Outline

- STR loci and kits
- Real Time PCR for DNA quantitation
- SNP markers
- Case example
- Interlaboratory Studies
- NIST Standard Reference Materials (SRMs)
- miniSTRs
- Sample Quality Control



Case Example: Typing SNPs for a Collaborator

Typing Autosomal SNPs to determine F_{ST} values

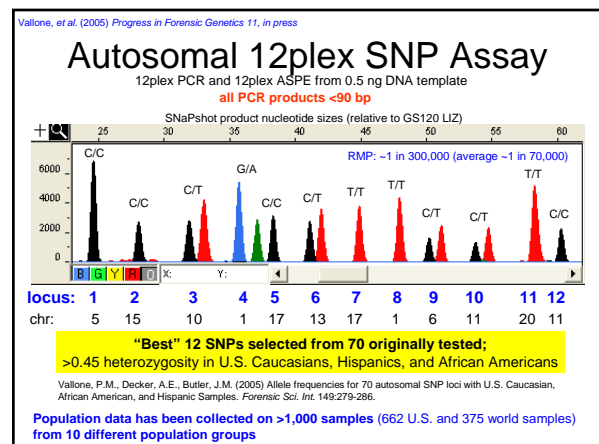
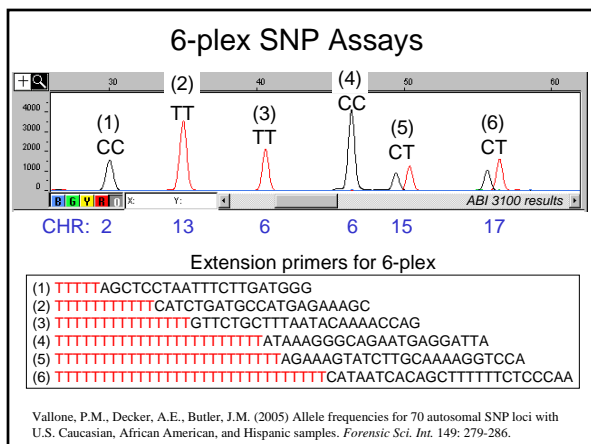
- 12-plex assay
- 12 multiplex panels (70 SNPs)

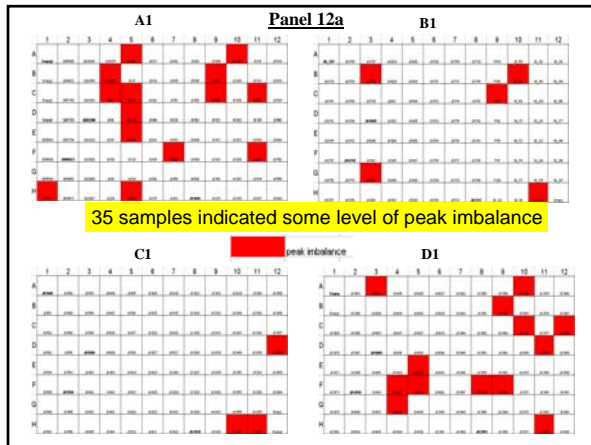
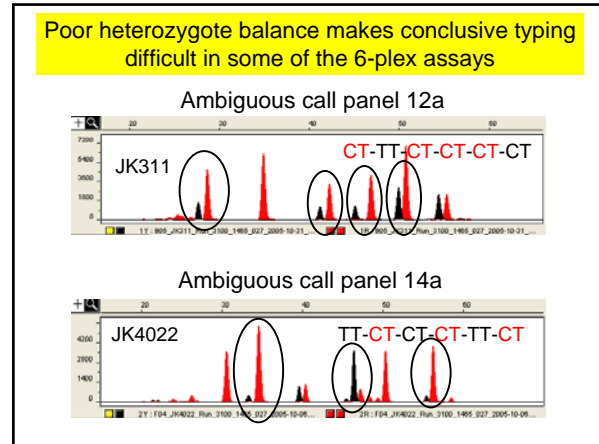
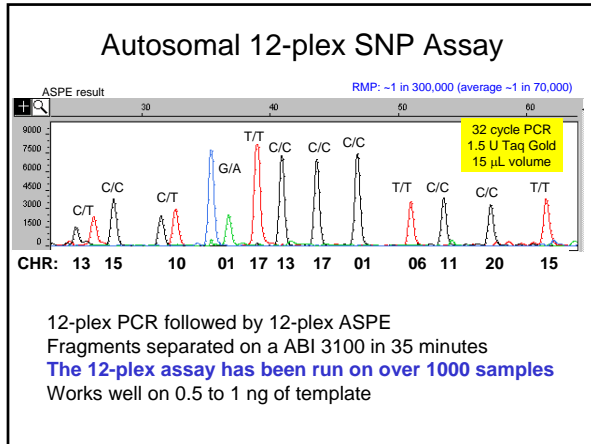
The samples were from different world population groups

- Biaka Pygmy
- Cambodian
- Chinese_Taiwan
- Hausa
- Ibo
- Maya
- Euro_American

Summary of Samples Sent and NIST Sample Handling

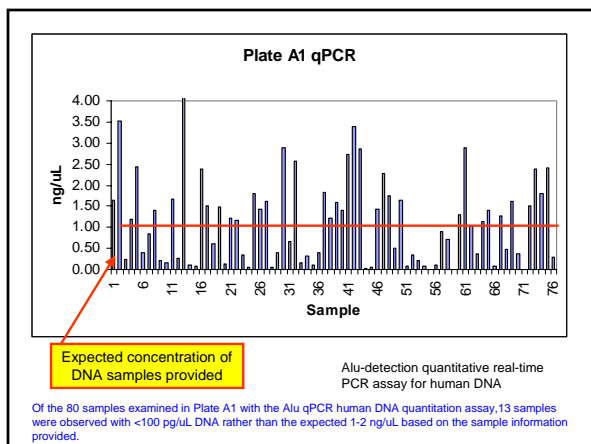
- Four 96-well sample plates containing 375 samples
- The samples were rehydrated in 100 μ L of nanopure water, centrifuged and allowed to equilibrate overnight
- Samples were handled with care to avoid contamination
- Based on information provided from collaborator, the nominal DNA template concentrations should be 1-2 ng/ μ L





Overview of Sample Quality and Quantity Issues

- Autosomal STR typing experiments exhibited mixtures that may indicate well to well contamination. Some of the STR profiles indicated a 1:1 mixture ratio.
- In addition, qPCR human DNA quantitation experiments conducted at NIST indicate that the actual concentrations varied from 100 pg to almost 10 ng.
- The low concentration samples are a concern because of the possibility of stochastic imbalance during PCR (allele drop-in or drop-out) and thus potential unreliable SNP results.

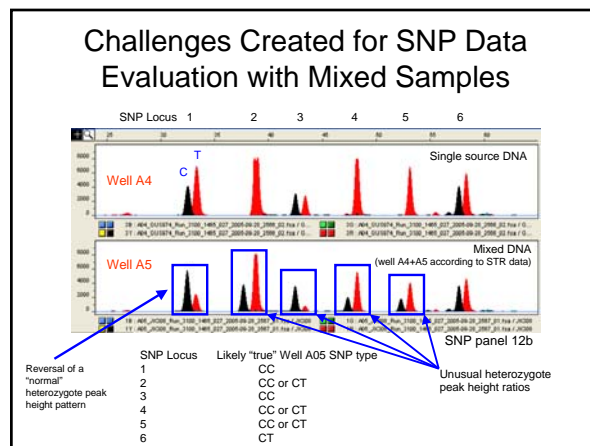
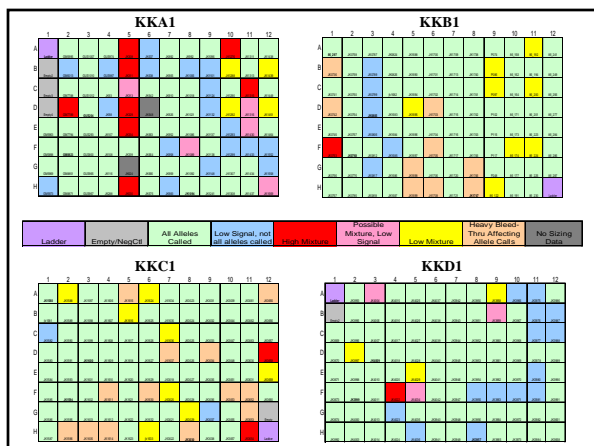
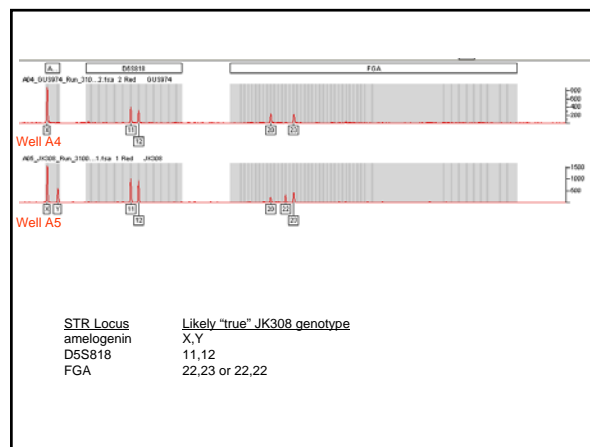
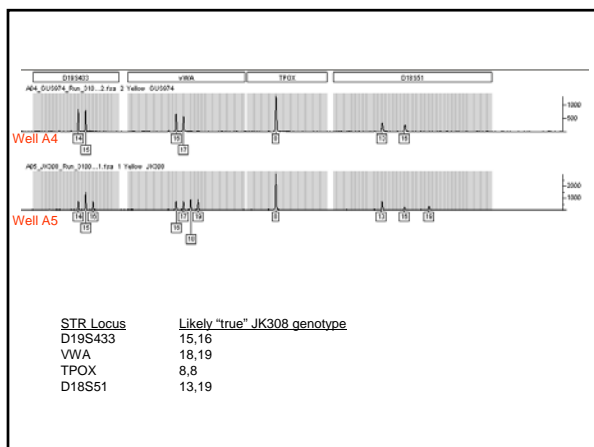
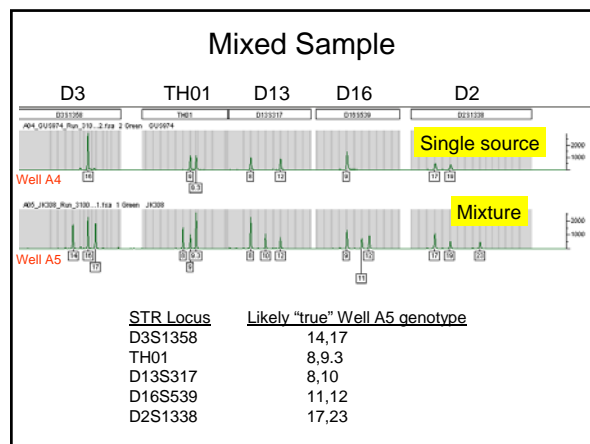
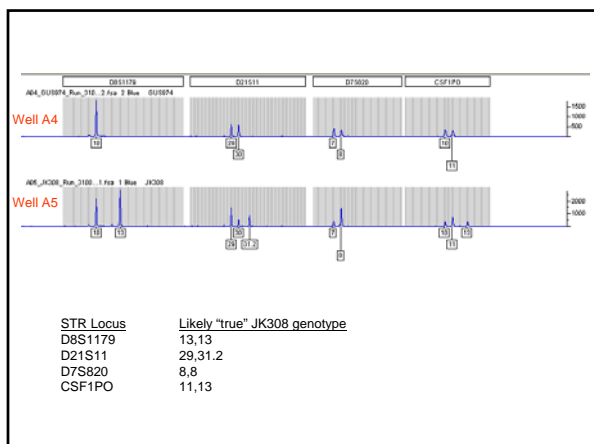


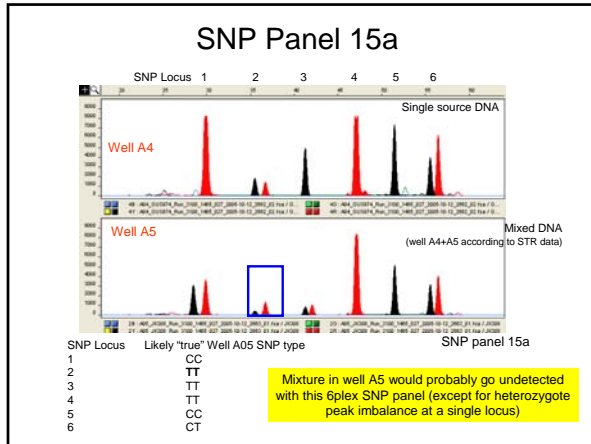
Autosomal STR Typing Data Confirm the Presence of Mixtures

- Obvious mixtures were identified
- Partial profiles indicate low template concentration
- Low mixtures and low signal/mixtures were also identified

Mixtures (major component) were observed

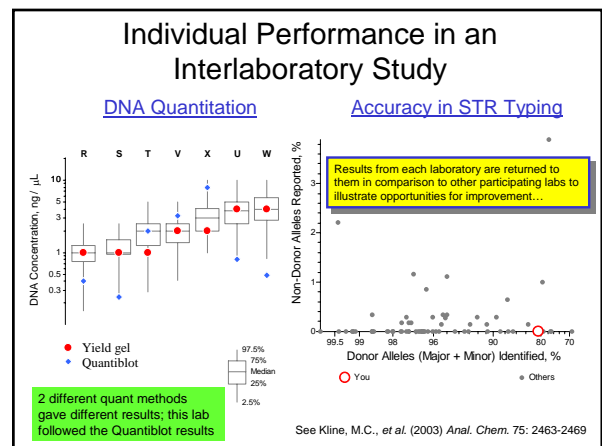
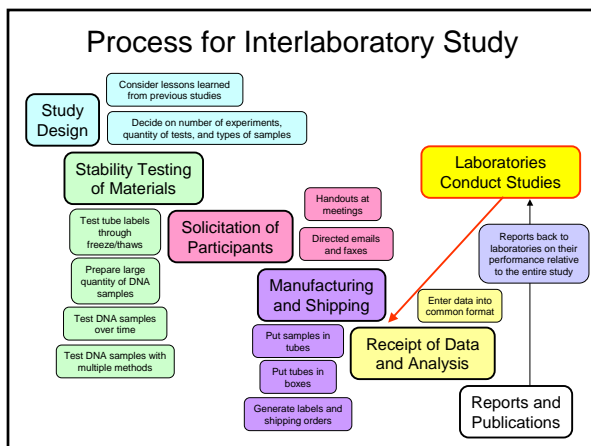
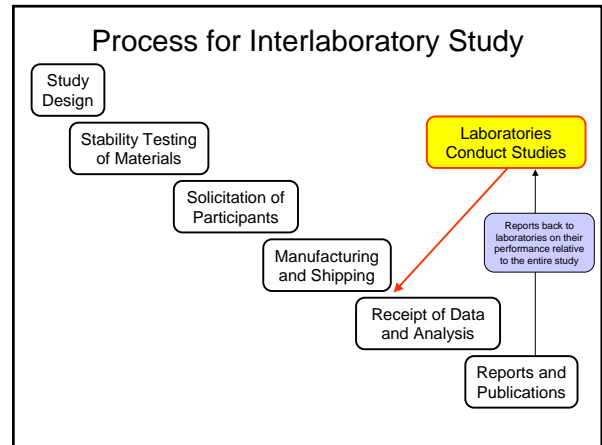
- 8 A1 (18 low, 5 possible mixture, 4 low/mixture)
- 1 B1 (6 low, 0 possible mixture, 8 low/mixture)
- 2 C1 (2 low, 0 possible mixture, 8 low/mixture)
- 1 D1 (14 low, 3 possible mixture, 3 low/mixture)






- ### Lessons Learned
- Quantitate all received samples prior to typing
 - Run autosomal STRs to confirm that samples are from a single source and are unique
 - Resolving mixed samples is time consuming and results are sometimes still ambiguous (SNPs more so than STRs)
 - Use caution when running a new assay on samples of unknown quality
 - By following this strategy we could have saved time and expense

- ### Outline
- STR loci and kits
 - Real Time PCR for DNA quantitation
 - SNP markers
 - Case example
 - Interlaboratory Studies
 - NIST Standard Reference Materials (SRMs)
 - miniSTRs
 - Sample Quality Control
-



Outline

- STR loci and kits
- Real Time PCR for DNA quantitation
- SNP markers
- Case example
- Interlaboratory Studies
- NIST Standard Reference Materials (SRMs)
- miniSTRs
- Sample Quality Control



NIST Standard Reference Materials (SRMs)

Standard Reference Materials Program

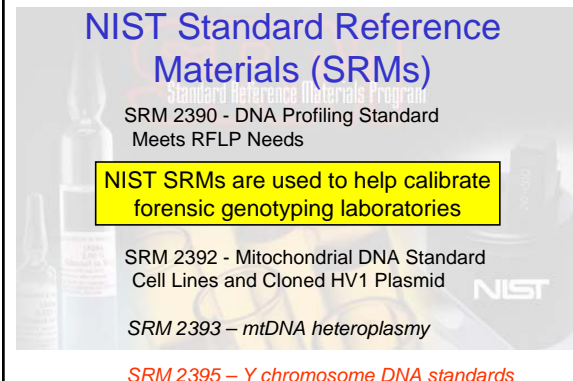
SRM 2390 - DNA Profiling Standard
Meets RFLP Needs

NIST SRMs are used to help calibrate forensic genotyping laboratories

SRM 2392 - Mitochondrial DNA Standard
Cell Lines and Cloned HV1 Plasmid

SRM 2393 – mtDNA heteroplasmy

SRM 2395 – Y chromosome DNA standards



STRBase

National Institute of Standards and Technology
Short Tandem Repeat DNA Internet Database

... working with industry to develop and apply technology, measurements and standards


Recent Additions

- Forensic SNP Information (will be official site for ISFG SNP information) .../SNP.htm
- NIST publications and presentations as pdf files .../NISTpub.htm

We Regularly Update


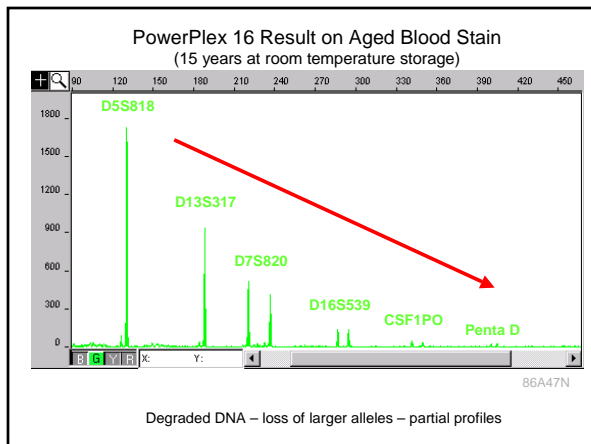
- Reference List
- Variant Alleles
- Addresses for Scientists
- Links to Other Web Sites
- Y-STR Information

We will continue to add downloadable PowerPoint files that can be used for training purposes




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

Outline

- STR loci and kits
- Real Time PCR for DNA quantitation
- SNP markers
- Case example
- Interlaboratory Studies
- NIST Standard Reference Materials (SRMs)
- miniSTRs
- Sample Quality Control

World Trade Center Towers (Sept 11, 2001)

Wreckage at Ground Zero

Highly degraded DNA was recovered

Identification of PCR Primers

```

No mapping library specified
Using 1-based sequence positions
CLUSTAL  I:  148  149  150  151  152  153  154  155  156  157  158  159  160  161  162  163  164  165  166  167  168  169  170  171  172  173  174  175  176  177  178  179  180  181  182  183  184  185  186  187  188  189  190  191  192  193  194  195  196  197  198  199  200
LEFT PRIMER  240  27  58.75  33.33  8.00  2.00  CATGAGTTTCAGAACTACTATTTCAC
RIGHT PRIMER  339  23  57.06  26.69  6.00  0.00  TTTTAAATTTTCAGAAATTTCA
SEQUENCE SIZE: 990
INCLUDED REGION SIZE: 990
SEQUENT SIZE: 990  PAIR ART CORPL: 7.00, PAIR 3' CORPL: 3.00
TANDETT  TPCNT, 1460  *  081,36
1  CCTACAGAAATCTAGAGTATCTCACTCGAAAATAAAAATATTTATTAAGGATAGAA
61  AATTCATCAAGGATCAAAATCTCTGTAATTAATTAATGACAGAAATAAAACAAATAC
121  AAAAACTACTTACTGTGATATAAACTATTCTTAATGAGAAATCAAAAGATGAAAC
181  AAACAAAAGTTTACTATTATGACTTTTTTGGTATLATTGGATATTTTGGATTGG
241  AAGTTA  CATGAGTTTCAGAACTACTATTTCAC  CCGAATATTATTAATTAATTA
XXXXXXXXXXXXXXXXXXXXXXXXXXXX
301  TTATTATTATTATTA  TGGAAATTTGAGAAATTAATA  TAATTTTGTCTGTGTGT
XXXXXXXXXXXXXXXXXXXXXXXXXXXX
    
```

PCR Primer Design

9 GATA repeats

TAGACAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATA

GAGAGATAGAGAGAGAGAGAGAGATGGGTTTTGGGGTTTTTTT

TGTTTGTGGTTTTTCAGACAGGATCTTAACGTGTAGTGGC

PCR Primer Design

9 GATA repeats

TAGACAGATAGATAGATAGATAGATAGATAGATAGATAGATA

GAGAGATAGAGAGAGAGAGAGAGATGGGTTTTGGGGTTTTTTT

TGTTTGTGGTTTTTCAGACAGGATCTTAACGTGTAGTGGC

REJECT!

PCR Primer Design

13 GGAA Repeats

AACCTGAGCATTAGCCCCAGGACCAATCTGGTCACAAACATA

TTAATGAAATTGACAAATGAGTGAAGTGGAAAGGAAAGGAA

GGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAATGAAG

ACAAACAAACAGAGTTGTTCCCTTAATAACAAAGCAAGGGA

AAAAGAGAACTCTCAGAAATAGTGTAAATATAATATCCAGG

PCR Primer Design

D10S1248

AACCTGAGCATTAGCCCCAGGACCAATCTGGTCACAAACATA

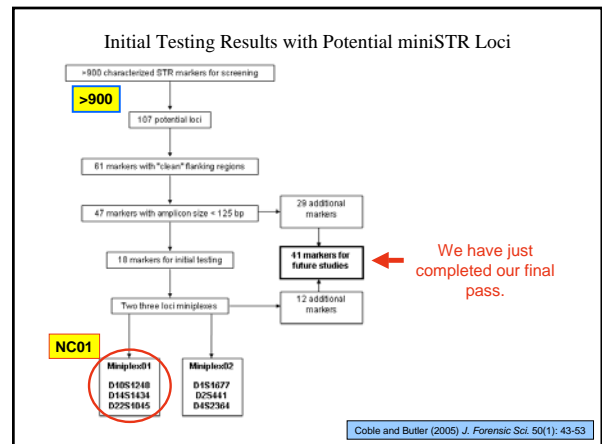
TTAATGAAATTGACAAATGAGTGAAGTGGAAAGGAAAGGAA

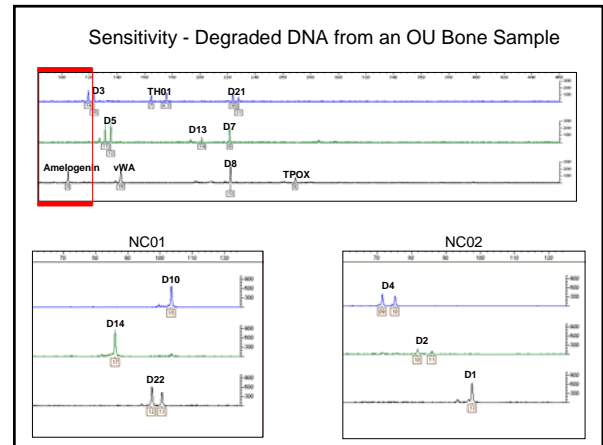
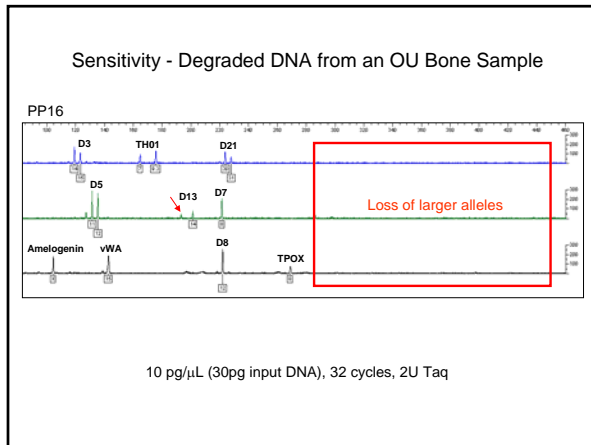
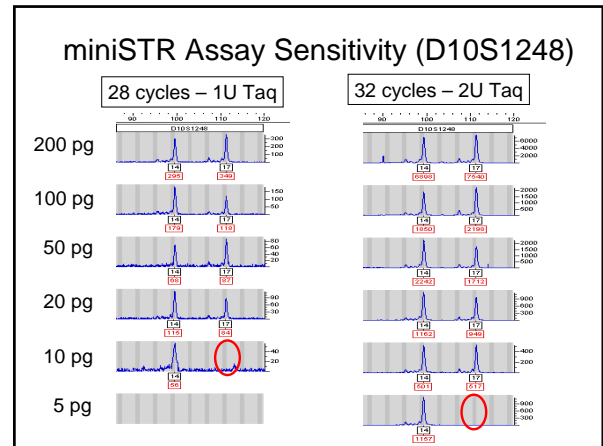
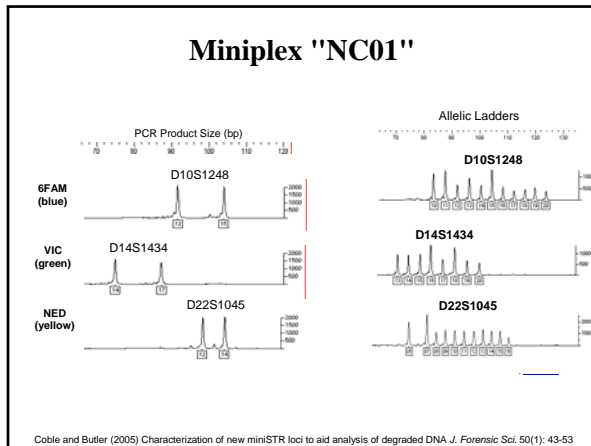
GGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAATGAAG

ACAAACAAACAGAGTTGTTCCCTTAATAACAAAGCAAGGGA

AAAAGAGAACTCTCAGAAATAGTGTAAATATAATATCCAGG

102 bp Amplicon





Global Impact of NC miniSTRs

FSI (2006) 156(2): 242-244

Short communication

The evolution of DNA databases—Recommendations for new European STR loci

Peter Gill^{a,b}, Lyn Fereday^b, Niels Morling^c, Peter M. Schneider^d

^a Forensic Science Service, Birmingham, UK
^b Forensic Science Service, London, UK
^c Department of Forensic Genetics, Institute of Forensic Medicine, University of Copenhagen, Denmark
^d Institute of Legal Medicine, University of Cologne, Germany

Received 25 May 2005; accepted 26 May 2005

...recommended that existing multiplexes are re-engineered to enable small amplicon detection, and that three new mini-STR loci with alleles <130 bp (D10S1248, D14S1434 and D22S1045) are adopted as universal. This will increase the number of European standard Interpol loci from 7 to 10.

(D14 has been replaced with D2S441 from NC02)

Characterization of New miniSTRs

Autosomal miniSTR Loci (Gene Symbol to be STR, Fast Short)

NC01: [D10S1248][D14S1434][D22S1045]

NC02: [D4S1345][D15S1169][D18S1477]

NC03: [D18S1477][D2S441][D20S482]

NC04:

NC05:

Link to Protocols

PCR Conditions

Preparation of Master Mix:

Preparation of individual PCR Reactions:

Thermal Cycling


Population Data D10S1248

Source	Number of Samples	Population
Coble and Butler (2005)	164	African Americans
Coble and Butler (2005)	170	U.S. Caucasians
Coble and Butler (2005)	140	U.S. Hispanics
Anzures et al. IJLIM in press	142	Japanese
Simon Len, in preparation	226	Chinese
Simon Len, in preparation	238	Malaysian
Simon Len, in preparation	238	East Indian
Prof. Bassani D'Amico, in preparation	100	Italian
total = 1418		

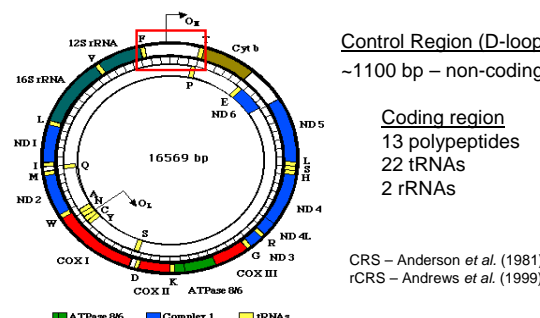
<http://www.cstl.nist.gov/biotech/strbase/newSTRs.htm>

Outline

- STR loci and kits
- Real Time PCR for DNA quantitation
- SNP markers
- Case example
- Interlaboratory Studies
- NIST Standard Reference Materials (SRMs)
- miniSTRs
- Sample Quality Control



The Mitochondrial Genome



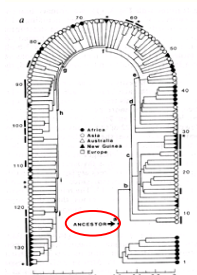
Control Region (D-loop)
~1100 bp – non-coding

Coding region
13 polypeptides
22 tRNAs
2 rRNAs

CRS – Anderson *et al.* (1981)
rCRS – Andrews *et al.* (1999)

<http://www.mitomap.org/>

mtDNA as a Genetic Marker

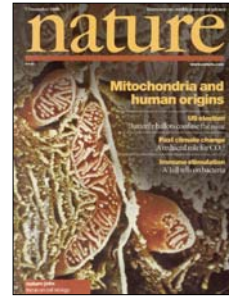


Cann *et al.* 1987

RFLP analysis of 134 mtDNA types from 148 individuals ~ 370 different restriction sites per individual.

“Out of Africa”

mtDNA as a Genetic Marker

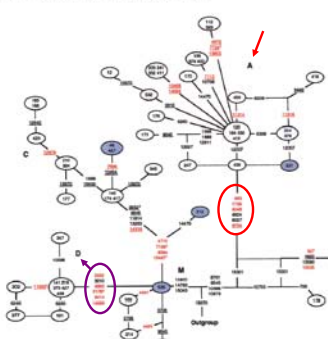


Ingman *et al.* (2000)

53 entire genome sequences from diverse global populations.

Confirmation for OOA.

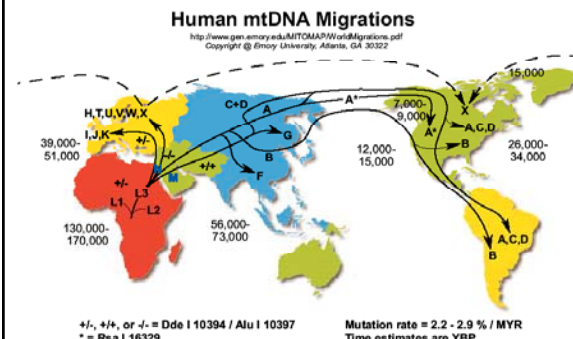
Mitochondrial Haplogroups



Haplogroup - A group of related haplotypes.

Each haplogroup cluster is defined by a set of specific, shared polymorphisms.

mtDNA Haplogroups (RFLP)



Human mtDNA Migrations

Mutation rate = 2.2 - 2.9 % / MYR
Time estimates are YBP

mtDNA as a Forensic Tool

Cases that have utilized mtDNA testing

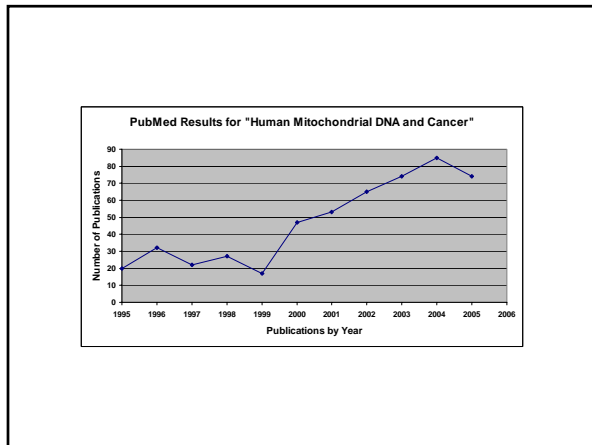
HUMAN GENOMICS, VOL. 2, NO. 4, 232-237, JANUARY 2006

Mitochondrial DNA as a potential tool for early cancer detection

Ryan L. Parr,^{1*} Gabriel D. Dakubo,¹ Robert E. Thayer,¹ Keith McKenney² and Mark A. Birch-Machin³

¹Genesis Genomics Inc, 1294 Balmoral Street, Thunder Bay, Ontario, P7B 5Z5, Canada
²School of Computational Sciences, George Mason University, GMU-PW, MS-4E3 10900 University Blvd., Manassas, VA 20110, USA
³Dermatological Sciences, University of Newcastle, Newcastle upon Tyne, NE2 4BFL, UK
 *Correspondence to: Tel: +1 807 346 8106; Fax: +1 807 346 8106; E-mail: Ryan.Parr@genomics.com

Date received: 14th September 2005

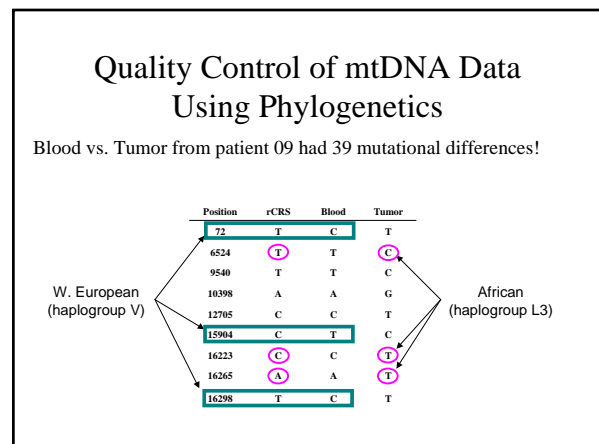


- ### Quality Control of mtDNA Data Using Phylogenetics
- NIST – Early Detection Research Network (EDRN) test site for cancer biomarker evaluation.
 - mtDNA as a cancer biomarker.
 - Samples from JHU (Patient lung tumor tissue and blood).
 - WERB – internal review of manuscripts

Quality Control of mtDNA Data Using Phylogenetics

Blood vs. Tumor from patient 09 had 39 mutational differences!

Position	rCRS	Blood	Tumor
72	T	C	T
6524	T	T	C
9540	T	T	C
10398	A	A	G
12705	C	C	T
15904	C	T	C
16223	C	C	T
16265	A	A	T
16298	T	C	T



FBI Forensic mtDNA database - USA.AFR.000942

HV1	HV2
16126-16187-16189-16223-16264	73-249d-263-290d-291d
16270-16278-16293-16311-16519	309.1C-315.1C-489

FBI Forensic mtDNA database - USA.AFR.000942

HV1	HV2
16126-16187-16189-16223-16264	73-249d-263-290d-291d
16270-16278-16293-16311-16519	309.1C-315.1C-489

African haplogroup L1b Asian haplogroup C1

“Artificial Recombination”

Lessons Learned

- Bandelt et al. (2004), Yao et al. (2004), Salas et al. (2005) – practical suggestions for improved quality control and methods to identify errors in data using the well-established mtDNA phylogeny.
- Increased interactions with experts in the mtDNA world (aDNA, population genetics, etc...)
- Sample quality – LCN, degradation can lead to erroneous results (e.g. artifacts, contamination) – need for improved protocols (controls) and established methods for analyses.

NIST Human Identity Project Team

John Butler (Project Leader) Margaret Kline Pete Vallone Mike Coble

Dave Duewer Jan Redman Amy Decker Becky Hill Chris DeAngelis
Anal. Chem. Division computer programmer

Funding: Interagency Agreement 2003-IJ-R-029 between National Institute of Justice (NIJ) and NIST Office of Law Enforcement Standards (OLES)