NIST Research Update

John M. Butler
Peter M. Vallone, Michael D. Coble, Amy E. Decker, Janette W. Redman, David L. Duewer, Margaret C. Kline

July 12, 2004
SWGDAM

NIST Human Identity Project Team

Funding:
Interagency Agreement between NIJ and NIST Office of Law Enforcement Standards

Presentation Outline

- Forensic DNA Typing, 2nd Edition
  - to be available in Jan 2005
- NIST Research Projects
  - Y-chromosome information, kits, and standards
  - New loci under development
  - miniSTRs
  - Autosomal SNPs
  - Performance with degraded DNA samples including hair shafts
  - DNA quantitation interlaboratory performance across 80 labs (NIST QS04)
  - STRBase updates and other tools to aid state and local labs

Forensic DNA Typing, 2nd Edition

Chapter 1: Overview & History of DNA Typing
Chapter 2: DNA Biology Review
Chapter 3: Sample Collection, Extraction, Quantitation
Chapter 4: PCR Amplification
Chapter 5: Common STRs and Commercial Kits
Chapter 6: Autosomal SNPs
Chapter 7: Forensic Issues
Chapter 8: Single Nucleotide Polymorphisms
Chapter 9: Y-Chromosome DNA Tests
Chapter 10: Mitochondrial DNA Analysis
Chapter 11: DNA Identification of Non-Human Forensics
Chapter 12: DNA Separation Methods
Chapter 13: DNA Detection Methods
Chapter 14: Instrumentation for STR Typing: ABI 310, ABI 3100, ABI 3130
Chapter 15: STR Genotyping Methods
Chapter 16: Additional STRs
Chapter 17: New Technologies, Automation, and Expert Systems
Chapter 18: CODIS and DNA Databanks
Chapter 19: Basic Genetic Principles and Statistics
Chapter 20: STR Database Analyst
Chapter 21: Proficiency Frequency Estimates
Chapter 22: DNA Analysis of Mixtures and Degraded DNA
Chapter 23: Human DNA Typing
Chapter 24: Mass Disaster DNA Victim Identification
Appendix A: International STR Allele Frequencies
Appendix B: U.S. Population Data STR Allele Frequencies
Appendix C: Requirements for DNA Analysis Equipment
Appendix D: Standards and Practices
Appendix E: DNA Recommendations on Statistics
Appendix F: Application of NRC II to STR Typing
Appendix G: Example DNA Cases

Academic Press plans to have available by January 2005

New Material:
10 additional chapters
Statistics (basics with examples)
Real-time PCR
Serology tests
Y-STRs and mtDNA
ABI 3100
Expert systems
Mass disasters including WTC
Example cases for training purposes
>500 new reference citations
50 new figures and 45 new tables
Manuscript is ~950 pages
Approximately double the size of the first edition

Current Areas of NIST Research Effort

- Y-Chromosome Information, Assays, and Standards
- Resources for “Challenging Samples” (SNPs and miniSTRs)
- DNA Quantitation (Interlab study, Real-time PCR comparisons)
- Tools to Aid State and Local Laboratories (e.g., STRBase)
- Aid to or Completion of Other NIJ Projects (e.g., LSBs)

Y-Chromosome
Information, Assays, and Standards

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
Selection of U.S. Core Loci:

DYS19, DYS385 a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS438, DYS439

(Minimal/standard haplotype)

DYS19, DYS389I/II, DYS390, DYS391, DYS392, DYS393

(White et al.)

DYS19, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS438, DYS439

(Ayok et al.)

DYS19, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS438, DYS439

(Sid et al.)

DYS19, DYS385 a/b, DYS390, DYS391, DYS392, DYS393, DYS438, DYS439

(Redd et al.)

DYS19, DYS385 a/b, DYS390, DYS391, DYS392, DYS393, DYS438, DYS439

(Bosch et al.)

DYS19, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS446, DYS447, DYS448, DYS449, DYS450, DYS452, DYS453, DYS454, DYS455, DYS456, DYS458

(Ayub et al.)

DYS19, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS438, DYS439

(DYS446-DYS645)

166 new Y STRs

(Manfred Kayser GDB entries)

A7.1 (DYS460)

A7.2 (DYS461)

A10

C4

H4

DYS385 a/b

DYS389I/II

DYS390

DYS391

DYS392

DYS393

DYS438

DYS439

DYS446-DYS645

Yfiler (Applied Biosystems)

Y-PLEX 5 (ReliaGene)

Y-PLEX 12 (ReliaGene)

PowerPlex Y (Promega)

YCAII a/b

YCAIII a/b

Y-PLEX 6 (ReliaGene)

Y-PLEX 12 (ReliaGene)

PowerPlex Y (Promega)

Yfiler (ABI)

New Y-STR paper

June 2004 issue of American Journal of Human Genetics

A Comprehensive Survey of Human Y Chromosomal Microsatellites

Manfred Kayser1,2, Kari Költö3, Axel Fréel1,4, Anita Hedman5, Andrew K. Loo1, Asha Mohd-Mokhlis4,5, S. Qadam Ali6, Zohar Rosen3, Mark Stenekes3, Mark A. Jablonski3, Arnt S. Sjønaes2, and Chris S. Skelton3

1Department of Forensic Genetics, 2Pathology Institute for Forensic Science, 3Laboratory of Genetics, 4Department of Genetics, University of Uppsala, Uppsala, Sweden; 5Department of Anthropology, University of Queensland, Brisbane, Australia; 6Department of Human Genetics, University of Manchester, Manchester, UK; and the Interdisciplinary Centre for Molecular Medicine, Uppsala, Sweden

• Searched for all regions with ≥ 8 consecutive repeats and 2,3,4,5, or 6 bp repeat units

• Discovered 139 new polymorphic Y-STR loci (166 male-specific)

• Only studied so far in 8 different samples

U.S. Population Data on 22 Y-STRs

Available online at http://www.biotechnology.nist.gov/strbase/NISTPub.htm

Y-Chromosome Standard

NIST SRM 2395

Human Y-Chromosome DNA Profiling Standard

- 5 male samples + 1 female sample (neg. control)

- 100 ng of each (50 µL at ~2 ng/µL)

- 22 Y STR markers sequenced

- 9 additional Y STR markers typed

- 42 Y SNPs typed with Marligen kit

Certified for all loci in commercial Y-STR kits:

Y-PLEX 6

Y-PLEX 5

Y-PLEX 12

PowerPlex Y

Yfiler

SWGDAM recommended loci

Yfiler - adds DYS635 (C4), now sequenced

帮助 meet DAB Standard 9.5 (and ISO 17025)...traceability to a national standard

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
Y-STRs in Casework

Utility of the Y-STR Typing Systems Y-PLEX™ 6 and Y-PLEX™ 5 in Forensic Casework and 11 Y-STR Haplotype Database for Three Major Population Groups in the United States*

<table>
<thead>
<tr>
<th>Case</th>
<th>Date</th>
<th>Population</th>
<th>Number</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Thoughts on Y-Chromosome Issues

- Core loci are selected, commercial kits are now available
- Y-STRs need to be put into greater use with forensic casework to demonstrate their value

Research Issues

- Nomenclature for Y-STR alleles in new loci
- Impact of additional loci to resolve most-common types
- Publicly available databases for additional loci
- Statistical issues with combining autosomal and Y-STR information

Resources for “Challenging Samples” (degraded DNA or mixtures)

- miniSTRs
  - CODIS loci (JFS 2003, 48, 1054-1064) – “BodePlexes”; WTC IDs; McCord collaboration
  - New loci (Coble, AAFS Feb 2004) – non-CODIS loci: unlinked; optimal for small amplicons and size ranges; <120 bp

- Autosomal SNPs
  - Validated Orchid 70 SNP markers (60-80 bp); population typing

- Mitochondrial DNA SNP Assays
  - Improve ease of use – Roche LINEAR ARRAY testing
  - Improve power of discrimination – AFDIL coding region SNPs

- Y-STRs
  - Improve evaluation of some extreme female-male mixtures?

Why go beyond CODIS loci

“STRs have proven to be highly successful [for mass disasters] in the past e.g. Waco disaster and various air disasters. However, even if the DNA is high quality there are occasions when there are insufficient family members available to achieve a high level of confidence with an association.”


Why go beyond CODIS loci

- Desirable to have markers unlinked from CODIS loci (different chromosomes) for some applications
- Small size ranges to aid amplification from degraded DNA samples

Why evaluate new markers?

• Highly Degraded samples (fragmented, questionable DNA quantity, inhibitors?)
• Telogenic/shed hairs (few copies)
• Low copy number cases (few copies)
• Siblings/Closely related individuals (paternity)

The primary characteristic of the assays for typing these new markers is their short PCR amplicon size (60 –150 base pairs)

STR Size Reduction
Through Moving Primer Positions Closer to Repeat

Focus on previously characterized STR markers with:
- High Heterozygosity
- Relatively small allele range
- "Clean" flanking regions for primer design adjacent to target repeat

Characterization of New miniSTR Loci

• Candidate STR marker selection
• Chromosomal locations and marker characteristics
• PCR primer design
• Initial testing results
• Population testing
• Allelic ladder construction
• Miniplex assay performance

Initial Testing Results

>900 potential markers
61 markers with “clean” flanking regions
43 markers with amplicon size < 125bp
18 markers for initial testing
2 three loci miniplexes

Locations of Focus for New miniSTR Loci (relative to CODIS 13 STRs)

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

<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),_</td>
<td>9-18</td>
<td>81-117</td>
</tr>
<tr>
<td>D5S441</td>
<td>(CTCTA),_</td>
<td>9-17</td>
<td>78-110</td>
</tr>
<tr>
<td>D4S2364</td>
<td>(GAAT)(GGAT)(GAAT),_</td>
<td>8-12</td>
<td>67-83</td>
</tr>
<tr>
<td>D10S1248</td>
<td>(GGAA),_</td>
<td>10-20</td>
<td>83-123</td>
</tr>
<tr>
<td>D14S1434</td>
<td>(GATA)(GACA),_</td>
<td>13-20</td>
<td>70-98</td>
</tr>
<tr>
<td>D22S1045</td>
<td>(TAA),_</td>
<td>5-16</td>
<td>76-109</td>
</tr>
</tbody>
</table>

Coble and Butler, JFS, manuscript submitted

### Population Testing – Miniplexes vs. Identifiler

<table>
<thead>
<tr>
<th>Marker</th>
<th>Heterozygosity</th>
<th>N = 474 Individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>D2S1338</td>
<td>0.8784</td>
<td>164 African Americans</td>
</tr>
<tr>
<td>D18S51</td>
<td>0.8753</td>
<td>170 Caucasians</td>
</tr>
<tr>
<td>FGA</td>
<td>0.8710</td>
<td>140 Hispanics</td>
</tr>
<tr>
<td>D5S818</td>
<td>0.8393</td>
<td></td>
</tr>
<tr>
<td>D19S433</td>
<td>0.8245</td>
<td></td>
</tr>
<tr>
<td>D7S820</td>
<td>0.8170</td>
<td></td>
</tr>
<tr>
<td>D14S1434</td>
<td>0.7590</td>
<td></td>
</tr>
<tr>
<td>D22S1045</td>
<td>0.7548</td>
<td></td>
</tr>
</tbody>
</table>

### miniSTR Assay

#### Allelic Ladders

- **Miniplex01**
  - D10S1248
  - D14S1434
  - D22S1045

- **Miniplex02**
  - D4S2364
  - D5S441
  - D1S1677

**miniSTR Assay Sensitivity (D10S1248)**

- 200 pg: 28 cycles – 1U Taq
- 100 pg: 32 cycles – 2U Taq
- 50 pg: 28 cycles – 2U Taq
- 20 pg: 28 cycles – 2U Taq
- 10 pg: 28 cycles – 2U Taq
- 5 pg: 28 cycles – 2U Taq

### SNP Typing at NIST

- STRBase is the official ISFG/EDNAP/ENFSI repository of forensic SNP information

- We are cataloging SNP information with the goal to standardize assays and speed validation of markers

- We will continue to explore various SNP typing technologies to provide information to the forensic DNA typing community – primary focus on SNaPshot

- We are beginning to evaluate SNP performance directly against miniSTRs for analysis of degraded DNA - collaborative study planned with EDNAP

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
Forensic SNP Site now a part of STRBase

**SNP characteristics**

- 70 Loci – sites from Orchid – C/T bi-allelic
- Present on 20 of 22 autosomal chromosomes (all but 3, 16, and X, Y)
- Amplicon size range 59 - 108 bp (average 69 bp)
- Markers are typed by allele-specific primer extension assays (ABI SNaPshot)
- Level of multiplexing (6-12-plexes)
- Web page for SNP site info
  

**Allele-Specific Primer Extension**

SNP Primer is extended by one base unit

ABI PRISM® SNaPshot™ Multiplex System

Fluorescently labeled ddNTPs + polymerase

PCR Amplified DNA Template

6-plex PCR was used to amplify 2 ng of gDNA (0.5 U Taq Gold)

6-plex primer extension was used to type the loci

**SNP typing results for a single individual**

The assays still require signal balancing, but the genotypes were unambiguous for our databasing

- 12 assays/70SNPs

**SNP Assay Results**

- 70 were typed for 189 U.S. samples (self identified ethnicities)
- 74 Caucasians + 71 African Americans AA + 44 Hispanics

- Total of 13,230 possible genotypes
- One marker failed across all samples (13,041-98.6%)
- 42 Samples were re-injected to confirm ambiguous results (99.7 %) success rate on first pass
- Results described in manuscript (Forensic Sci. Int.)

- We are in the process of optimizing a 12-plex panel of SNPs
A subset of the 70 SNP markers

- Observed heterozygosity of >0.45 in each of the 3 populations
- All 189 samples were individualized using these 12 markers

**Can we get nuclear DNA from hair shafts?**

Yes…

But depends on the extraction method and assay used

---

**SNP Assay Sensitivity**

- 1.0 ng
- 0.5 ng
- 125 pg
- 63 pg
- 31 pg

**TNCa Buffer**

- Tris
- NaCl
- CaCl₂
- 2% SDS
- ProK
- DTT

Complete digestion of hair in about 1 hour based on method by


---

**SNP and STR Assays on Shed Hairs**

- 6 µL template
- 10 µL rxn
- 5 hairs (2.5 cm ea); 2 U Taq; 32 cycles

---

**PowerPlex 16 Genotype**

- 5 Hair Shafts (2.5 cm each, no roots)
- Loss of larger alleles

(5 markers typed – note: no Amelogenin )
SNP and STR Assays on Shed Hairs

1 hair (20 cm); 2 U Taq; 32 cycles

2 uL temp
10 uL rxn

Future directions with SNPs and miniSTRs

- Optimize 12-plex for SNPs
- Determine sensitivity of assays
- Examine data interpretation issues for LCN assays (e.g., allele drop out, RFU thresholds)
- Type on a “standard” degraded sample (compare to commercial kits)
- Mobility modifiers with miniSTRs (potential for greater multiplexing)

DNA Quantitation

Interlaboratory Study Results
SRM 2372 : Human DNA Quantitation Standard

NIST Quantitation Study 2004 (QS04)

Consisted of:
- 8 DNA extracts labeled A – H
- Shipped Dec 2003 – Jan 2004 to 84 laboratories for quantification.
- Labs were requested to use multiple methods / multiple analysts
- Last day for submission extended from 15 March to 5 April 2004

We received data from 80 Labs (95%)
Total of 287 sets of data
Participants used 19 different quantification methods (primarily variations on Quantiblot and Real-time PCR)

Summary of the Blot based Methods


http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
Summary of the RT-PCR Methods

![Graph showing summary of RT-PCR methods](Kline, et al., J. Forensic Sci., in preparation)

Among Participate Results

![Graph showing among participate results](Kline, et al., J. Forensic Sci., in preparation)

Table 2: The percent recovery rate for a single target.

<table>
<thead>
<tr>
<th>Target DNA</th>
<th>1.5</th>
<th>0.5</th>
<th>0.1</th>
<th>0.05</th>
<th>0.01</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
</tr>
<tr>
<td>Quantifier</td>
<td>77</td>
<td>88</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Other RT-PCR</td>
<td>23</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>“Alu RT-PCR”</td>
<td>54</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>AEC</td>
<td>13</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>PicoGreen</td>
<td>12</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>ECL</td>
<td>79</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>TMB</td>
<td>95</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Voxel grid</td>
<td>54</td>
<td>57</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Apparent Precision

- Concordance – is a direct multi-material analogue of bias
- Apparent precision – is analogous to precision but also incorporates sample-specific measurement differences or “matrix effects”.

![Graph showing apparent precision](Kline, et al., J. Forensic Sci., in preparation)

- A = AluQuant
- P = PicoGreen

Apparent Precision

- Concordance

![Graph showing apparent precision and concordance](Kline, et al., J. Forensic Sci., in preparation)

- A = ACES
- T = TMB
- E = ECL

Apparent Precision

- Concordance

![Graph showing apparent precision and concordance](Kline, et al., J. Forensic Sci., in preparation)

- 0 = Quantifier
- 1 = Alu RT-PCR
- 5 = BRCA1
- 6 = CSF-HUMRT

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
Publication of NIST QS04 Results

- Paper describing results is complete
- Being sent out for review by all 80 contributing laboratories along with certificates of participation
- Concurrently going through NIST internal review
- Following these reviews, the manuscript will be submitted to *J. Forensic Sci.*

*Results being used for SRM 2372 development*

Tools to Aid State and Local Laboratories

- **STRBase** – standard information source
- **Variant Alleles** – cataloging variants and tri-allelic patterns
- **NIST U.S. Population Samples and Database**
- **Quality Assurance Tool** – resolution monitor to track analytical performance over time
- **Validation Standardization Information**

**Training Materials**
- Downloadable PowerPoint files from STRBase
- Current Protocols in Human Genetics, Electrophoresis review article on STR analysis with ABI 310 and ABI 3100

Acknowledgements

**Funding:** Interagency Agreement between National Institute of Justice and NIST Office of Law Enforcement Standards

**NIST Project Team:**
- John Butler  Pete Vallone
- Margaret Kline  Jan Redman
- Jill Appleby  Amy Decker
- Mike Coble  Dave Duewer

AutoDimer – primer screening software is now freely available

http://www.cstl.nist.gov/biotech/strbase/AutoDimerHomepage/AutoDimerProgramHomepage.htm

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