National Institute of Standards and Technology

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

The Future of Forensic DNA Typing:
Y-STRs, SNPs, and Real-time PCR

John M. Butler

CODIS State Administrators
FBI Academy
May 25, 2004

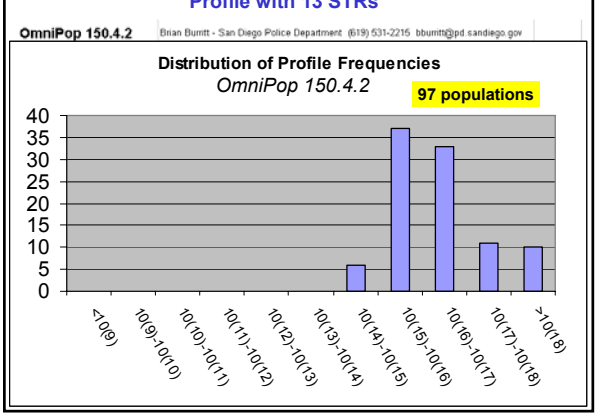
Presentation Outline

- Some general interest information
  - OmniPop 150.4.2 (166 Profiler Plus, 97 full CODIS loci pops)
  - *Forensic DNA Typing*, 2nd Edition (to be available in Jan 2005)
- Y-STRs
  - Advantages of Y-chromosome
  - Available Y-STR kits
  - Comparison to Y-SNPs
- Autosomal SNPs
  - Forensic SNP web page (http://www.cstl.nist.gov/biotech/strbase/SNP.htm)
  - Challenges ahead for SNPs
- miniSTRs
  - CODIS loci made small
  - New loci under development
  - Performance with degraded DNA samples
- DNA Quantitation
  - Real-time PCR methods
  - Performance across 80 laboratories (NIST interlab study QS04)

OmniPop 150.4.2

- Published allele frequencies
  - from 97 populations containing all 13 CODIS loci
  - From 166 populations with 9 loci (Profiler Plus)
- From 64 publications
- Available from Brian Buritt (San Diego Police Dept)
  - (619) 531-2215
  - bburritt@pd.sandiego.gov

Profile with 13 STRs



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 Biology of STRs
Chapter 7 Forensic Issues
Chapter 8 Single Nucleotide Polymorphisms
Chapter 9 Y-Chromosome DNA Tests
Chapter 10 Mitochondrial DNA
Chapter 11 Non-Human DNA
Chapter 12 DNA Separation Methods
Chapter 13 DNA Detection Methods
Chapter 14 Instrumentation for STR Typing: ABI 310, ABI 3100, FMBIO systems
Chapter 15 STR Genotyping Issues
Chapter 16 Lab Validation
Chapter 17 New Technologies, Automation, and Expert Systems
Chapter 18 CODIS and DNA Databases
Chapter 19 Basic Genetic Principles and Statistics
Chapter 20 Checking Population Data: Hardy-Weinberg, linkage equilibrium
Chapter 21 Profile Frequency Estimates: match probabilities, likelihood ratios
Chapter 22 Approaches to Degraded DNA and Mixture Interpretation
Chapter 23 Kinship and Paternity Testing
Chapter 24 Mass Disaster DNA Testing
Appendix I Reported STR Alleles
Appendix II U.S. Population Data-STR Allele Frequencies
Appendix III Suppliers of DNA Analysis Equipment
Appendix IV DAB QA Standards
Appendix V DAB Recommendations on Statistics
Appendix VI Application of NRC II to STR Typing
Appendix VII Example DNA Case from Start to Finish

National Institute of Justice
The Research, Development, and Evaluation Agency of the U.S. Department of Justice

Current Areas of NIST Research Effort

- Resources for "**Challenging Samples**" (degraded DNA or mixtures)
- **Y-Chromosome** Information, Assays, and Standards
- **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)

NIST Human Identity Project Team


John Butler


Margaret Kline


Jan Redman


Pete Vallone


Dave Duewer


Amy Decker


Jill Appleby


Mike Coble

Former (Honorary) Project Team Members

Rich Schoske

Christian Ruitberg

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?

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
  - *Forensic DNA Typing*, 2nd Edition (Dec 2004/Jan 2005)

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

Variant Alleles Cataloged in STRBase

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

Off-Ladder Alleles

224 total variants reported as of 04/21/04

- CSF1PO (10)
- D2S1338 (3)
- D3S1358 (16)
- D5S818 (5)
- D7S820 (20)
- D8S1179 (4)
- D13S317 (8)
- D16S539 (10)
- D18S51 (28)
- D19S433 (6)
- D21S11 (21)
- FESFPS (1)
- FGA (62)
- HUMTH01 (4)
- Penta D (8)
- Penta E (6)
- TPOX (7)
- VWA (5)

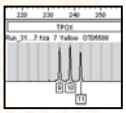
Currently **224** at 13/13 CODIS loci

Tri-Allelic Patterns

56 total patterns reported as of 04/09/04

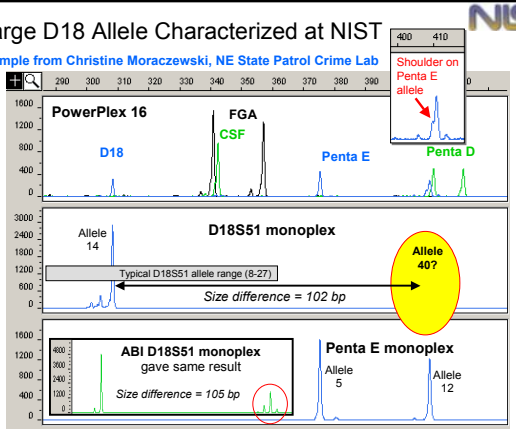
- CSF1PO (2)
- D3S1358 (4)
- D5S818 (1)
- D7S820 (3)
- D8S1179 (5)
- D13S317 (3)
- D16S539 (1)
- D18S51 (4)
- D21S11 (4)
- FGA (9)
- HUMTH01 (1)
- TPOX (12)
- VWA (7)

Currently **56** at 13/13 CODIS loci



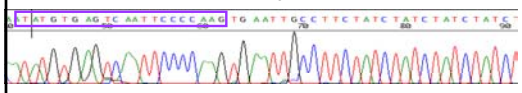
Large D18 Allele Characterized at NIST

Sample from Christine Moraczewski, NE State Patrol Crime Lab

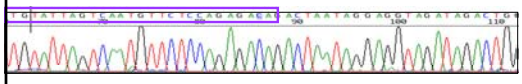


Sequence Analysis of STR Alleles

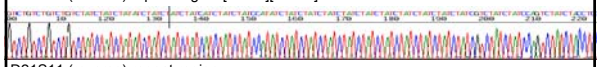
D21S11 PowerPlex 16 forward primer region [ATATGTGAGTCAATTCCCCAAG]



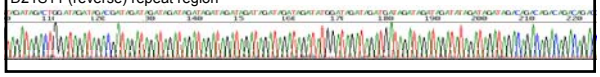
D21S11 PowerPlex 16 reverse primer region [TGTATTAGTCAATGTTCTCCAGAGAC]



D21S11 (forward) repeat region [TCTA][TCTG]



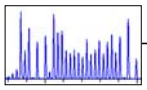
D21S11 (reverse) repeat region



NIST QA/QC Software

Tool being developed by Dave Duewer for STR Process Control

Will track allelic ladders and positive controls



NCBI Program

Peak Height, Area, Size

X | Y
103 | 436
104 | 569

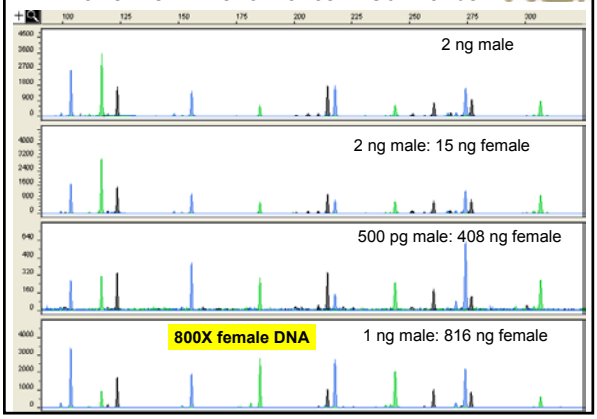
NIST ControlCharter → Date vs Sensitivity, Resolution, Precision

NIST ControlCharter → Date vs Signal/Noise

Y-Chromosome STRs

Male-specific amplification
(can handle extreme mixtures of male and female DNA)
Enables tracing paternal lineages
(permits extension of possible reference samples)

PowerPlex Y Performance in Our Hands



SWGDAM Sub-Committee on the Y Chromosome

- Formed in July 2002
- Members
  - Jack Ballantyne (UCF) – chair
  - Mecki Prinz (NYC) – co-chair
  - Bruce Budowle (FBI)
  - **John Butler (NIST)**
  - Ann Gross (MN)
  - John Hartmann (Orange Co.)
  - Laura Kienker (FBI Academy)
  - Carll Ladd (CT)
  - Demris Lee (AFDIL)
  - Phil Kinsey (OR)
  - Barb Koons (FBI Academy)
  - Tim Kupferschmid (ME)
  - Gary Sims (CA DOJ)
- **U.S. CORE Y-STR LOCI selected in January 2003**
- 60 sample set selected for screening markers and initial testing
- Testing of Y-PLEX 6 and Y-PLEX 5 kits in all labs
  - *All results completed agreed with NIST results sent to participating labs in Dec 2002*
- Jack Ballantyne's lab and John Butler's lab to examine additional Y STR and Y SNP markers in the same sample set

European and U.S. Core Y-STR Loci

Marker Name	Allele Range (repeat numbers)	Repeat Motif
DYS19	10-19	TAGA
DYS385 a/b	7-28	GAAA
DYS389 I **DYS389 II**	I: 9-17 II: 24-34	(TCTG) (TCTA) (TCTG) (TCTA)
DYS390	17-28	(TCTA) (TCTG)
DYS391	6-14	TCTA
DYS392	6-18	TAT
DYS393	8-17	AGAT
YCAII a/b	11-25	CA
DYS438	6-14	TTTTC
DYS439	8-15	AGAT

Minimal haplotype (Europe)

U.S. haplotype

Extended haplotype (Europe)

New Y-STR paper

June 2004 issue of American Journal of Human Genetics

Am. J. Hum. Genet. 74:1183–1197, 2004

A Comprehensive Survey of Human Y-Chromosomal Microsatellites

Manfred Kayser,[1,*] Ralf Kittler,[1,4] Axel Erler,[1,4] Minttu Hedman,[2] Andrew C. Lee,[3] Aisha Mohyuddin,[4,5] S. Qasim Mehdi,[5] Zoë Rosser,[3] Mark Stoneking,[1] Mark A. Jobling,[3] Antti Sajantila,[2] and Chris Tyler-Smith[4,6]

[1]Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig; [2]Department of Forensic Medicine, University of Helsinki, Helsinki; [3]Department of Genetics, University of Leicester, Leicester, United Kingdom; [4]Department of Biochemistry, University of Oxford, Oxford; [5]Biomedical and Genetic Engineering Laboratories, Islamabad; and [6]The Wellcome Trust Sanger Institute, Hinxton, Cambridge, United Kingdom

- **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**

Commercial Y-STR Kits Available

- ReliaGene Technologies (New Orleans, LA)
  - **Y-PLEX™ 6**: DYS19, DYS389II, DYS390, DYS391, DYS393, DYS385 a/b
  - **Y-PLEX™ 5**: DYS389I/II, DYS392, DYS438, DYS439
  - **Y-PLEX™ 12**: DYS19, DYS385 a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS438, DYS439, amelogenin
- Promega Corporation (Madison, WI)
  - **PowerPlex® Y**: DYS19, DYS385 a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS438, DYS439, DYS437

  Applied Biosystems (Y-filer: 17plex with 5-dye chemistry)
- Serac (Germany)
  - **genRES® DYSplex-1**: DYS389I/II, DYS390, DYS391, DYS385 a/b, amelogenin
  - **genRES® DYSplex-2**: DYS19, DYS389I/II, DYS392, DYS393
- GKT Inc. (South Korea); *silver-stain kits*
  - **GeneKin® Y-STR Systems I**: DYS388, DYS19, DYS392
  - **GeneKin® Y-STR Systems II**: DYS393, DYS390, DYS391, DXY391X
  - **GeneKin® Y-STR Systems III**: DXYS156X, DXYS156Y, DYS389I/II
  - **GeneKin® Y-STR Systems IV**: DXYS156X, DXYS156Y, DYS385 a/b

Commercial Y-STR Kits

(Minimal/extended haplotype)
DYS19
DYS389I/II
DYS390
DYS391
DYS392
DYS393
DYS385 a/b
YCAII a/b
DYS388
DYS425
DYS426
YCAIII a/b

(White et al.)
A7.1 (DYS460)
A7.2 (DYS461)
A10
C4
H4

(Bosch et al.)
G09411 (DYS462)

(Ayub et al.)
DYS434
DYS435
DYS436
DYS437
DYS438
DYS439

(Iida et al.)
DYS441
DYS442
DYS443
DYS444
DYS445

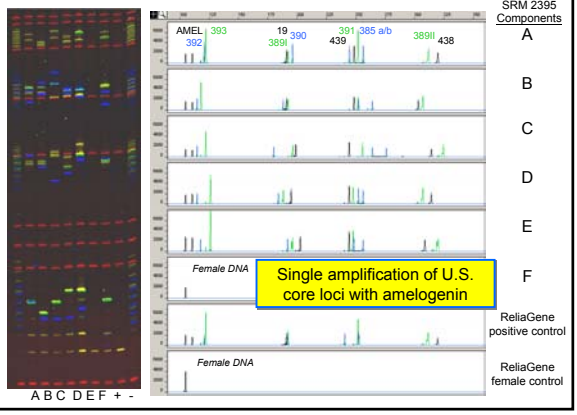
(Redd et al.)
DYS446
DYS447
DYS448
DYS449
DYS450
DYS452
DYS453
DYS454
DYS455
DYS456
DYS458
DYS459 a/b
DYS463
DYS464 a/b/c/d

43 (51) Y-STRs
(217 with Manfred's)

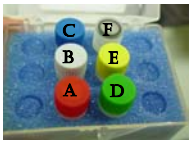
Y-PLEX 6 (ReliaGene)
Y-PLEX 5 (ReliaGene)
Y-PLEX 12 (ReliaGene)
PowerPlex Y (Promega)
Yfiler (Applied Biosystems)

DYS468-DYS645
166 new Y STRs
(Manfred Kayser GDB entries)

FMBIO III and ABI 3100 Results with Y-PLEX™ 12



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

$248

Certified for all loci in commercial Y-STR kits:
Y-PLEX 6
Y-PLEX 5
Y-PLEX 12
PowerPlex Y

SWGDAM recommended loci: DYS19, DYS385 a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS438, DYS439

Y-filer - adds DYS635 (C4)

Helps meet DAB Standard 9.5 (and ISO 17025)…traceability to a national standard

http://www.yhrd.org

As of 5/24/04: 22,872 haplotypes

Metapopulation	# Haplotypes	# distinct Haplotypes	# Populations	Avg. # haplotypes per population
Worldwide	22,873	11,573	200	114.365
Eurasian MP	18,885	8,020	160	118.03125
Eurasian MP / European MP	16,309	7,733	121	134.78513
Eurasian MP / Altaic MP	357	300	5	71.4
Eurasian MP / Caucasian MP	502	395	13	38.615385
Eurasian MP / Uralic MP	399	145	1	399.0
Eurasian MP / Arabian MP	100	91	1	100.0
Eurasian MP / Indo Iranian MP	1,000	564	17	58.82353
Eurasian MP / Indian MP	218	143	2	109.0
East Asian MP	2,105	1,781	16	131.5625
East Asian MP / Korean MP	316	284	1	316.0
East Asian MP / Japanese MP	394	316	3	131.33333
East Asian MP / Sino Tibetan MP	613	591	4	153.25
East Asian MP / Austroasiatic MP	209	190	1	209.0
East Asian MP / Thai MP	71	68	1	71.0
East Asian MP / Austronesian MP	473	377	5	94.6
East Asian MP / Indo Pacific MP	29	26	1	29.0
Australian Aboriginal MP	0	0	0	NaN
African MP	1,389	923	17	81.70588
African MP / Subsaharan MP	320	209	4	80.0
African MP / Afro-Asiatic MP	83	79	1	83.0
African MP / Afro-American MP	733	542	11	66.63636
African MP / Afro-Caribbean MP	253	229	1	253.0
Amerindian MP	163	112	5	32.6
Eskimo Aleut MP	69	45	1	69.0

Run only with minimal haplotype

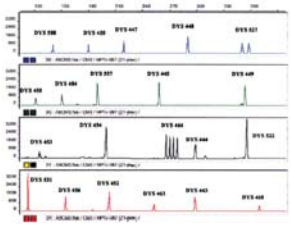
DYS19
DYS389I/II
DYS390
DYS391
DYS392
DYS393
DYS385 a/b

US haplotype requires 2 additional loci:
DYS438
DYS439

J Forensic Sci, Jan. 2004, Vol. 49, No. 1
Paper ID JFS2003209_491
Available online at: www.astm.org

Erin K. Hanson,[1] M.S. and Jack Ballantyne,[1,2,3] Ph.D.

A Highly Discriminating 21 Locus Y-STR "Megaplex" System Designed to Augment the Minimal Haplotype Loci for Forensic Casework*



Y-STRs in Casework

July 2004 *issue of Journal of Forensic Sciences*

J Forensic Sci, July 2004, Vol. 49, No. 4
Paper ID JFS2003246
Available online at: www.astm.org

Sudhir K. Sinha,[1] Ph.D.; Bruce Budowle,[2] Ph.D.; Ranajit Chakraborty,[3] Ph.D.; Ana Paunovic,[1] B.S.; Robin DeVille Guidry,[1] B.S.; Chris Larsen,[1] M.S.; Amrita Lal,[1] M.S.F.S.; Megan Shaffer,[1] Ph.D.; Gina Pineda,[1] M.S.; Siddhartha K. Sinha,[1] B.S.; Elaine Schneida,[1] B.S.; Huma Nasir,[1] B.S.; and Jaiprakash G. Shewale,[1] Ph.D.

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 1—*Y-STR cases using the Y-PLEX™ 6 and Y-PLEX™ 5 kits that have been accepted in U.S. courts.*

Case	Date	Jurisdiction	Docket No.	Notes
State of LA vs. Samuel Williams	10/23/01	Orleans Parish	416-355	Criminal paternity case
State of MS vs. Leon Felder	6/26/01	Pike County	00-537-KA	Sexual assault case-also had other STRs, Y-STR produced no result
State of GA vs. Ali R. Shabazz	7/31/02	Dekalb County	01-CR-4002	Sexual assault case
United States vs. Spc. Michael Kelly	10/16/02	Ft. Knox	...	Sexual assault case
State of OH vs. Chackie Unsworth	4/16/03	Lucas County	G-4801-CR-200301510	Daubert Hearing

U.S. Population Data on 22 Y-STRs

Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Forensic Science International
Forensic Science International 139 (2004) 107–121
www.elsevier.com/locate/forsciint

High-throughput Y-STR typing of U.S. populations with 27 regions of the Y chromosome using two multiplex PCR assays

Richard Schoske[a,b], Peter M. Vallone[a], Margaret C. Kline[a], Janette W. Redman[a], John M. Butler[a,*]
[a]*Biotechnology Division, National Institute of Technology, 100 Bureau Drive, Mail Stop 8311, Gaithersburg, MD 20899, USA*
[b]*Department of Chemistry, American University, Washington, DC 20016, USA*

Received 29 April 2003; received in revised form 25 September 2003; accepted 1 October 2003

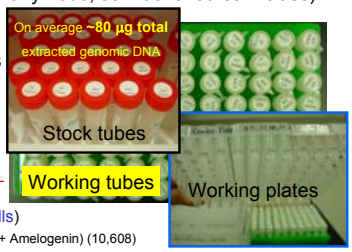
pdf file available at http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm

NIST U.S. Population Samples

As of 06/2003 666 males (anonymous; self-identified ethnicities)

286 Caucasians
252 African Americans
128 Hispanics

Whole blood received from Interstate Blood Bank (Memphis, TN)



Working tubes/plates 1 ng/uL

To date: (~70,000 allele calls)
Identifiler (15 autosomal markers + Amelogenin) (10,608)
Roche Linear Arrays (HV1/HV2 10 regions) (6,630)
Y STRs 22 loci—27 amplicons (17,388)
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)

Samples supplied to **OhioU** for miniSTR typing and **AFDIL** for whole mtGenome sequencing

Typing Beyond the U.S. Core Y-STR Loci

- **U.S. Core Y-STRs** (selected Jan 2003 by SWGDAM Y-chromosome subcommittee): DYS19, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS438, DYS439, DYS385 a/b
- First examination of 11 U.S. haplotype loci vs. new ones:
  647 U.S. males:
  - 260 African Americans (including 20 from Carll Ladd)
  - 244 Caucasians (including 20 from Carll Ladd)
  - 143 Hispanics (including 20 from John Hartmann)
- Types generated at NIST for 22 Y-STRs and 50 Y-SNPs
- Performance of U.S. haplotype vs. European "extended" haplotype (DYS438 and DYS439 vs. stutter-prone dinucleotide YCAII a/b)
- Resolution of most common types with additional markers

Schoske *et al.* (2004) High-throughput Y-STR typing of U.S. populations…, *Forensic Sci. Int., 139:107-121*

US haplotype (Reliagene kits)
Yfiler (ABI)
PowerPlex Y (Promega)
+C4

Y-STR	Pooled Population STR diversity (N=647)	Rank	African American STR diversity (N=260)	Rank	Caucasian STR diversity (N=244)	Rank	Hispanic STR diversity (N=143)	Rank
DYS464 a/b/c/d	0.956	1	0.954	1	0.934	1	0.937	1
DYS385 a/b	0.912	2	0.942	2	0.838	2	0.901	2
YCAII a/b	0.790	3	0.797	3	0.701	5	0.772	4
DYS458	0.765	4	0.758	5	0.743	3	0.793	3
DYS390	0.764	5	0.664	10	0.701	5	0.665	13
DYS447	0.747	6	0.767	4	0.683	7	0.748	5
DYS389II	0.736	7	0.722	6	0.675	8	0.734	6
DYS448	0.721	8	0.722	6	0.595	11	0.704	8
DYS456	0.700	9	0.671	9	0.731	4	0.695	9
DYS438	0.691	10	0.569	15	0.594	12	0.690	10
DYS19	0.676	11	0.722	6	0.498	19	0.672	12
DYS439	0.656	12	0.636	11	0.639	9	0.717	7
DYS437	0.637	13	0.499	17	0.583	13	0.624	14
H4	0.611	14	0.612	12	0.562	14	0.609	15
DYS392	0.609	15	0.434	20	0.596	10	0.673	11
DYS460	0.570	16	0.568	14	0.555	15	0.556	18
DYS389I	0.549	17	0.531	16	0.538	17	0.596	16
DYS391	0.534	18	0.447	19	0.552	16	0.577	17
DYS426	0.519	19	0.375	21	0.482	20	0.522	19
DYS450	0.489	20	0.487	18	0.177	22	0.414	21
DYS393	0.485	21	0.586	13	0.363	21	0.448	20
DYS388	0.365	22	0.246	22	0.501	18	0.312	22

Schoske *et al.* (2004) High-throughput Y-STR typing of U.S. populations…, *Forensic Sci. Int., 139:107-121*

Single Nucleotide Polymorphisms (SNPs)

Autosomal SNPs
Y-SNPs
mtDNA SNPs (control region & coding region)

Forensic Utility of SNPs

The short PCR amplicons required for typing SNPs may result in success with degraded samples and possibly higher sensitivity – **but this has not been demonstrated yet in real-world samples…**

For serious forensic usage, parallel high-throughput methods and multiplex amplification will be required for typing low amounts of DNA

Forensic Utility of SNPs

Short tandem repeat (STR)
CTGATGCTA**(GATA)$_n$**GACTACTTA

n = 5 to 15 = 66 possible allelic combinations

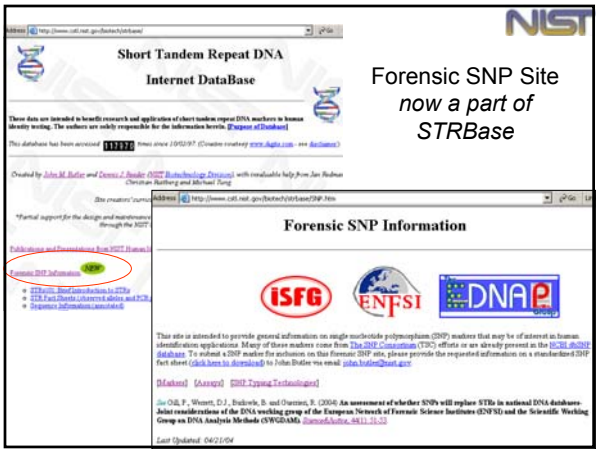
Single Nucleotide Polymorphism (SNP)
CTGATGCTA**(G/A)**GACTACTTA

3 possible allelic combinations

For human ID purposes more SNPs would be needed than STRs
Multiplexing is essential

SNP Typing at NIST

- STRBase is the official ISFG repository of forensic SNP information
  - Gill *et al. Science & Justice* 2004, 44, 51-53
  - **http://www.cstl.nist.gov/biotech/strbase/SNP.htm**
- 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



Forensic SNP Site
now a part of STRBase

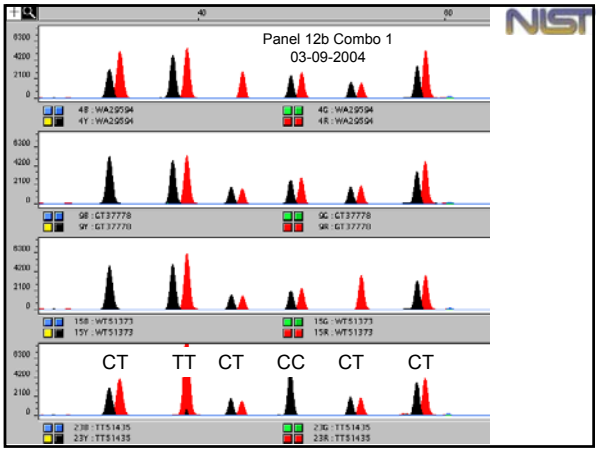
Autosomal SNPs

Orchid Cellmark provided their panel of 70 SNPs (C/T) located throughout the human genome

We validated these markers with SNaPshot assays for 8 CEPH samples in July 2001 for the WTC investigation as part of KADAP (Kinship and Data Analysis Panel)

We are evaluating these markers in U.S. populations (N=189 so far)

Marker info now on STRBase forensic SNP site:
http://www.cstl.nist.gov/biotech/strbase/SNP.htm



Y-SNP Typing Conclusions

- Different technologies yield the same Y-SNP type
  - Full concordance was observed between hybridization and primer extension technologies on 18 different Y-SNPs (>3,800 allele calls)
- Y-SNPs will have limited value for individualizing a sample
  - 18 different types observed in 229 individuals
- Current Y-SNPs appear to have limited value for ethnic differentiation in U.S. populations
  - One exception: M2 only in African Americans; not in Caucasians

Vallone, P.M. and Butler, J.M. (2004) *J. Forensic Sci., in press* (July issue)

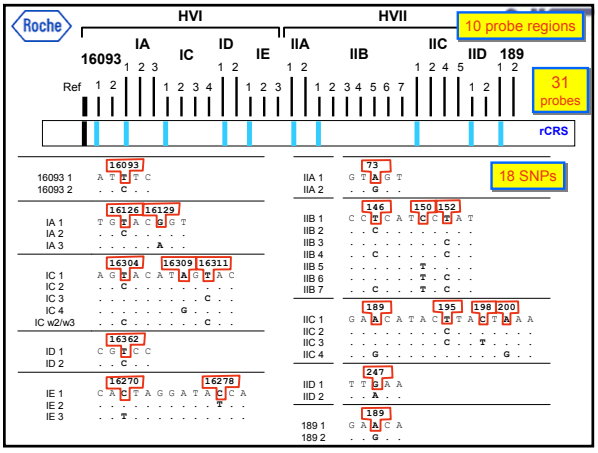
Typing mtSNPs

Coding Region SNPs

Collaboration with AFDIL (Tom Parsons and Mike Coble)
Develop an 11-plex assay for typing SNPs outside the control region
The 11 SNP sites are thought to help resolve Caucasians with the most common mitotype (~7%)

Control Region SNPs

Typing population samples with Roche linear arrays (Cassandra Calloway)
Probe 10 regions (18 SNPs) within HVI and HVII
Evaluate assay performance and ability to resolve U.S. population samples



miniSTRs

(Reduced Size Amplicons)

CODIS loci
New miniSTR loci

STR Size Reduction

Through Moving Primer Positions Closer to Repeat

Primer positions define PCR product size
Repeat information is independent of amplicon size

Advantages of Approach:
Size reduction enhances success rate with degraded DNA
Retains same marker information (database compatibility)
Uses highly polymorphic STR loci (high discriminatory power)

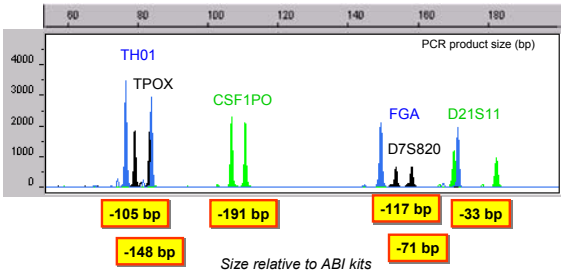
J. Forensic Sci. Sept 2003 issue

J Forensic Sci, September 2003, Vol. 48, No. 5
Paper ID JFS2003043_485
Available online at: www.astm.org

John M. Butler,[1] Ph.D.; Yin Shen,[2,3] Ph.D.; and Bruce R. McCord Ph.D.[2]

The Development of Reduced Size STR Amplicons as Tools for Analysis of Degraded DNA*

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



Size relative to ABI kits

Reduction in PCR Product Size

Locus	Size Difference (relative to ABI kits)
TH01	-105 bp
FGA	-71 bp
CSF1PO	-191 bp
D21S11	-33 bp
TPOX	-148 bp
D7S820	-117 bp

Not as much size reduction as other STR loci…

How Close Can a Stable Primer be Designed to the STR Repeat Region?

Locus		Distance 3'end from Repeat	Comment
CSF1PO	F	14	*partial repeat just 5' of repeat*
	R	6	
FGA	F	8	
	R	23	*partial repeat just 3' of repeat*
TH01	F	0	
	R	1	
TPOX	F	-4	
	R	5	
VWA	F	0	
	R	0	
D3S1358	F	-1	
	R	-1	
D5S818	F	4	
	R	-5	
D7S820	F	4	
	R	65	*polyA stretch just 3' of repeat*

Problems with Large Allele Spreads

STR Locus	GenBank Accession	GenBank Allele	Allele Range	Allele Spread
CSF1PO	X14720	12	6–16	40 bp
FGA	M64982	21	12.2–51.2	156 bp
TH01	D00269	9	3–14	44 bp
TPOX	M68651	11	5–14	36 bp
vWA	M25858	18	10–25	60 bp
D3S1358	NT_005997	18	8–20	48 bp
D5S818	AC008512	11	7–16	36 bp
D7S820	AC004848	13	5–15	40 bp
D8S1179	AF216671	13	7–19	48 bp
D13S317	AL353628	11	5–16	44 bp
D16S539	AC024591	11	5–15	40 bp
D18S51	AP001534	18	7–27	80 bp
D21S11	AP000433	29	24–38.2	58 bp
Penta D	AP001752	13	2.2–17	73 bp
Penta E	AC027004	5	5–24	95 bp
D2S1338	AC010136	20	15–28	52 bp

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."

Gill, P., Werrett, D.J., Budowle, B. and Guerrieri, R. (2004) **An assessment of whether SNPs will replace STRs in national DNA databases-Joint considerations of the DNA working group of the European Network of Forensic Science Institutes (ENFSI) and the Scientific Working Group on DNA Analysis Methods (SWGDAM).** *Science&Justice*, 44(1): 51-53.

Why go beyond CODIS loci

"To achieve this purpose, either new STRs could be developed, or alternatively, existing STRs could be supplemented with a SNP panel."

"There also efforts for modifying existing STR panels by decreasing the size amplicons by designing new primers."

Gill, P., Werrett, D.J., Budowle, B. and Guerrieri, R. (2004) **An assessment of whether SNPs will replace STRs in national DNA databases-Joint considerations of the DNA working group of the European Network of Forensic Science Institutes (ENFSI) and the Scientific Working Group on DNA Analysis Methods (SWGDAM).** *Science&Justice*, 44(1): 51-53.

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

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

↓

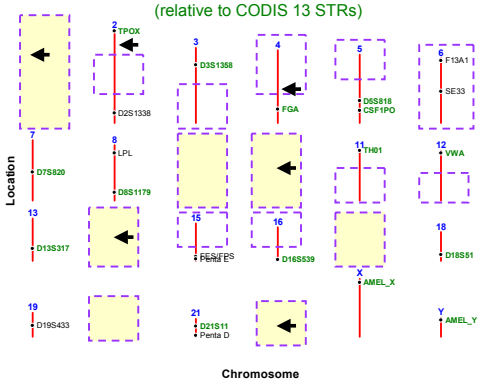
18 markers for initial testing

↓

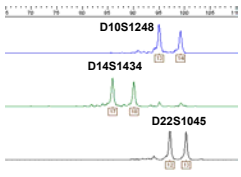
2 three loci miniplexes

Locations of Focus for New miniSTR Loci

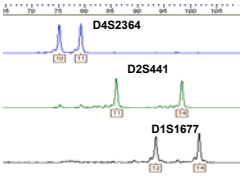
(relative to CODIS 13 STRs)



Miniplex 1



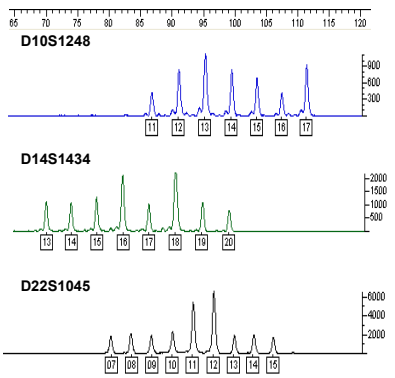
Miniplex 2



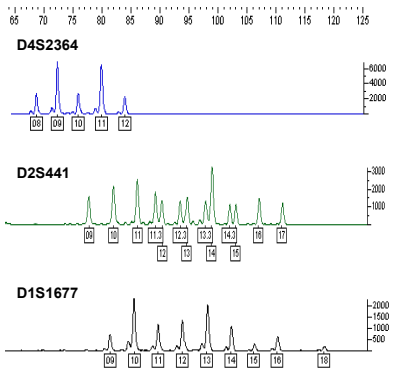
Some Marker Characteristics

Chr.	Marker Name	(Motif)	Ref. Repeat	Amplicon Size	Primer distance from repeat
10	D10S1248	TETRA	13	102	1
	GGAA23C05N	GGAA			0
14	D14S1434	TETRA	10	88	1
	GATA168F06	GATA			0
22	D22S1045	TRI	13	105	3
	ATA37D06	ATA			6
1	D1S1677	TETRA	15	103	0
	GGAA22G10N	GGAA			0
2	D2S441	TETRA	12	92	0
	GATA8F03	GATA			0
4	D4S2364	TETRA	7	78	2
	GAAT1F09	GAAT			1

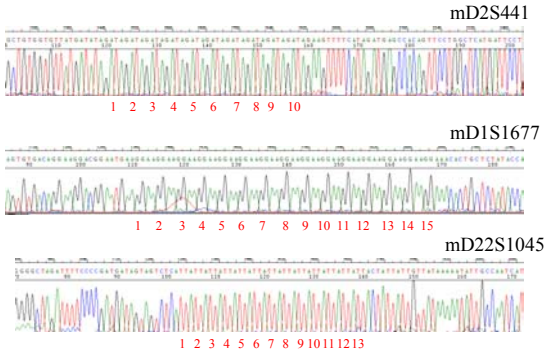
Allelic Ladders for New Miniplex01



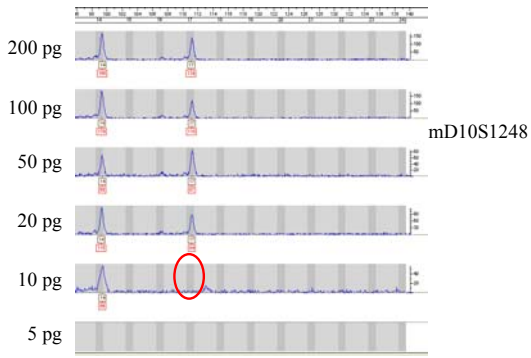
Allelic Ladders for New Miniplex02



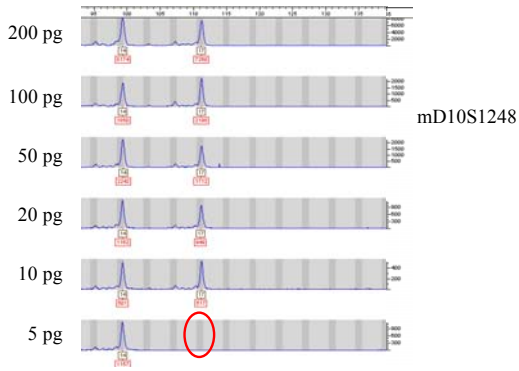
Characterization of New miniSTR Loci



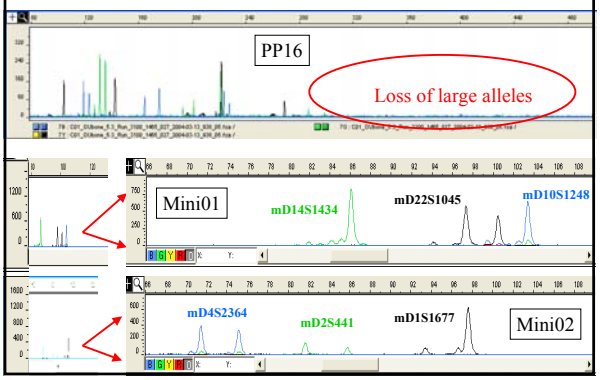
Sensitivity – 28 cycles



Sensitivity – 32 cycles – 2U Taq



Sensitivity - Degraded DNA from an OU Bone Sample



Future Plans with miniSTRs

Testing and characterization of more markers

Further population databasing

Testing on degraded materials (bones and hair)

Information will be posted on STRBase website and published as these loci are characterized (*article submitted to JFS for first set*)

We would welcome collaborations with those wishing to test some of these new miniSTR systems

DNA Quantitation

Real-time PCR methods
Interlaboratory Study Results

$ Cost per sample (20 µL – 25 µL)

Assay	$ PCR Master Mix	$ Primers	$ TaqMan probe	Total
Alu	0.80*	0.0025	NA	$0.8025
TH01	0.80*	0.0025	NA	$0.8025
CFS-HUMRT	0.73#	0.0025	0.17	$0.9025
RB1	0.73#	0.0025	0.17	$0.9025
mtDNA	0.73#	0.0025	0.17	$0.9025
Qfiler Human	NA	NA	NA	$2.50
Qfiler Y Male	NA	NA	NA	$2.50

* Platinum® SYBR® Green qPCR SuperMix UDG (Invitrogen, Carlsbad, CA)
# Platinum® Quantitative PCR SuperMix – UDG (Invitrogen, Carlsbad, CA)

Assay specifications tried at NIST TaqMan Probes

Assay	amplicon	GeneTarget	probe	#Cycles
CFS-HUMRT 11p15.5	62 bp	Human tyrosine hydroxylase gene	15 bp VIC	40
RB1 13	79 bp	Human retinoblastoma susceptibility gene	26 bp FAM	50
mtDNA	143 bp	tRNA lysine & ATP synthase 8, Coding Region	29 bp VIC	50
Qfiler Human 5p15.33	62 bp	Human telomase reverse transcriptase (hTERT)	? FAM	40
Qfiler Y Male Yp11.3	64 or 61 bp	Sex-determining region Y gene (SRY)	? FAM	40

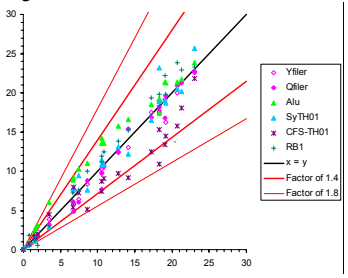
Comparison of RT-PCR assays

Following Published Protocols

Series of NIST population samples with a range of [DNA] from 40 pg to 23 ng

The same "Standard" was used for all methods (8 dilutions).

Time for the assay:
Alu-RT-PCR ~ 1.25 h
(fewer cycles required)
The rest ~ 1.75 h



X = median value for all methods
Y = measured value for the method

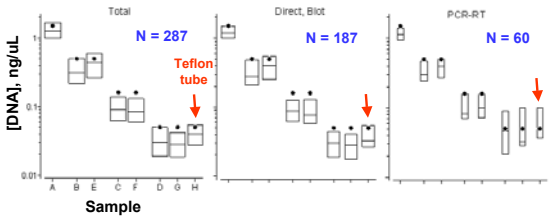
Results of NIST Quantitation Study 04

Consisted of:

8 DNA extracts labeled A – H Shipped Dec 2003 –Jan 2004
Shipped to 84 laboratories for quantification.
Labs asked 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)

NIST QS 04 Results-Box plots

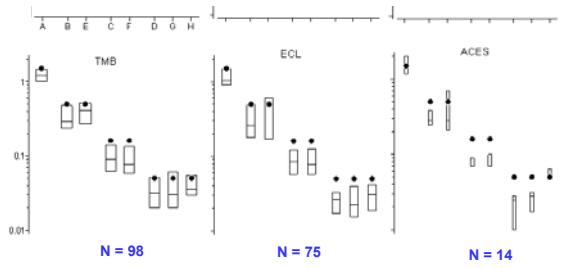


Width of the box is proportional to the number data points.

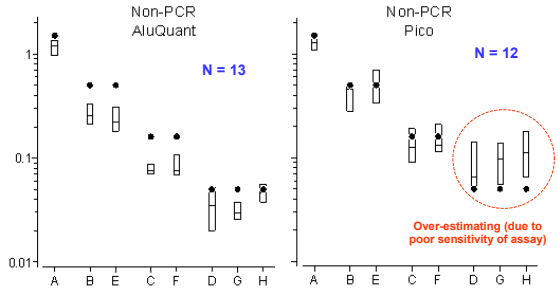
Line in the box is the median value.
The box represents 50% of the data submitted.

Dot is the target [DNA].

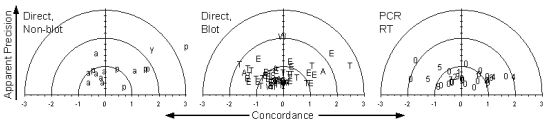
Comparison of the Blot Methods



AluQuant and PicoGreen



NIST QS 04 Results-Target Plots



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