National Institute of Standards and Technology
. . .working with industry to develop and apply technology, measurements and standards

Forensic DNA Projects at NIST

John M. Butler
EDNAP-ENFSI
Cyprus
April 15, 2004

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


Current Areas of NIST Research Effort

- Resources for "**Challenging Samples**" (degraded DNA or mixtures)
  - miniSTRs (reduced size amplicons) – J. Forensic Sci. 2003, 48, 1054-1064
  - Autosomal SNP assays and U.S. population information – 70 Orchid SNPs
  - mtDNA coding SNP multiplex assays with AFDIL mtGenome sequencing
- **Y-Chromosome** Information, Assays, and Standards
  - Y-STR 20plex, 11plex and U.S. population data including multi-copy DYS464
  - Y-SNP U.S. population data - J. Forensic Sci., in press (July 2004)
  - Human Y-Chromosome Standard Reference Material (SRM 2395)
- **DNA Quantitation**
  - Interlab study and standard reference material (SRM 2372)
  - Real-time PCR comparisons
- **Resource Information to Aid Forensic Laboratories**
  - STRBase (STR fact sheets and variant allele lists)
  - Forensic SNP site
- **Work with New Typing Technologies**
  - Microchip CE DNA separations

miniSTRs

(Reduced Size Amplicons)

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)

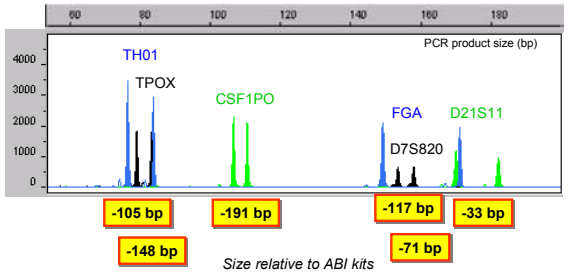
J. Forensic Sci. Sept 2003 issue

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



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	3	
	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

↓

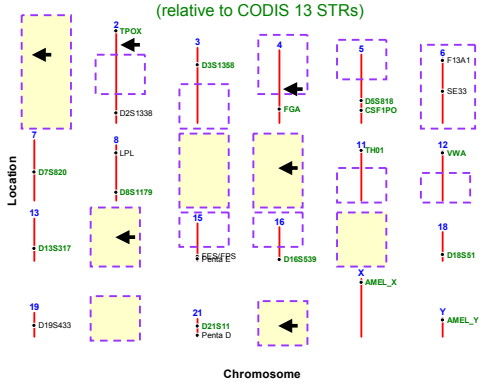
18 markers for initial testing

↓

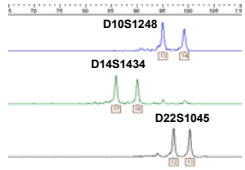
2 three loci miniplexes

Locations of Focus for New miniSTR Loci

(relative to CODIS 13 STRs)



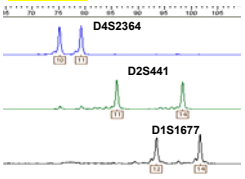
Miniplex 1



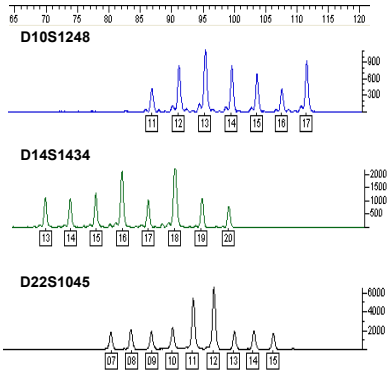
Some Marker Characteristics

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

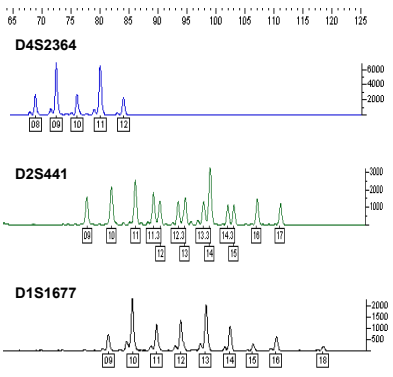
Miniplex 2



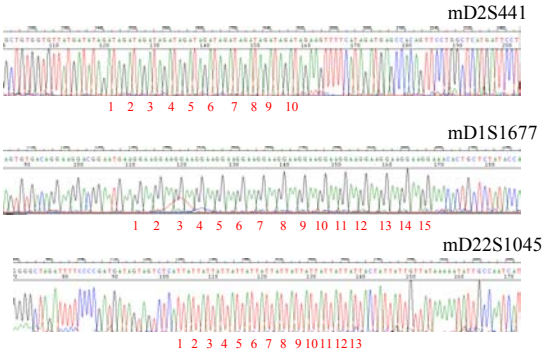
Allelic Ladders for New Miniplex01



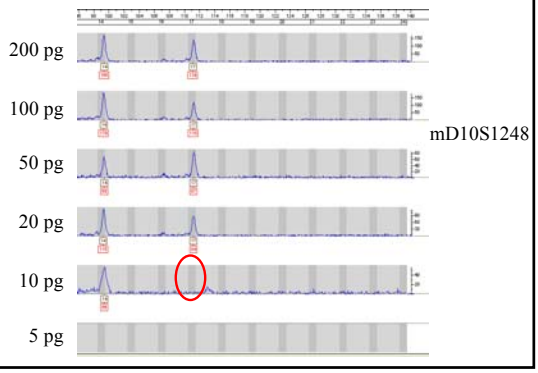
Allelic Ladders for New Miniplex02



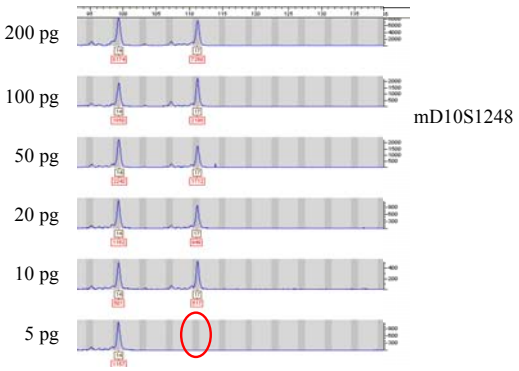
Characterization of New miniSTR Loci



Sensitivity – 28 cycles



Sensitivity – 32 cycles – 2U Taq

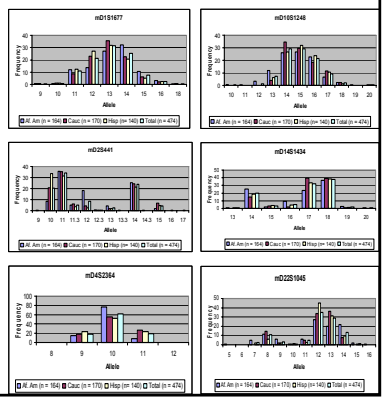


Heterozygosity Values

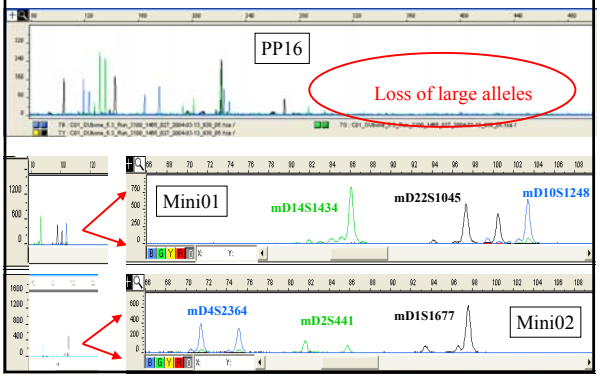
Heterozygosity	Marker
0.8784	D2S1338
0.8753	D18S51
0.8710	FGA
0.8393	D21S11
0.8245	vWA
0.8076	D7S820
0.7970	D19S433
0.7759	mD10S1248 - mini01
0.7759	D16S539
0.7674	mD22S1045 - mini01
0.7674	D8S1179
0.7590	mD2S441 - mini02
0.7548	D3S1358
0.7526	D13S317
0.7463	mD1S1677 - mini02
0.7378	CSF1PO
0.7378	THO1
0.7294	D5S818
0.7146	TPOX
0.6765	mD14S1434 - mini01
0.5307	mD4S2364 - mini02

Size range 79-95 bp

Allele Frequencies (n = 473)



Sensitivity - Degraded DNA from an OU Bone Sample



Future Plans

Testing and characterization of more markers.

Population Databasing.

Testing on degraded materials.

Information will be posted on STRBase website and published as these loci are characterized

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

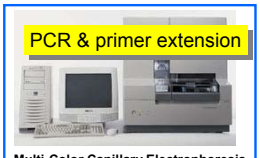
SNPs

Autosomal, Y, and mtDNA

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**

SNP Typing Instrumentation



Multi-Color Capillary Electrophoresis (ABI 310 or 3100)

Luminex 100 Flow Cytometer

Time-of-Flight Mass Spectrometer

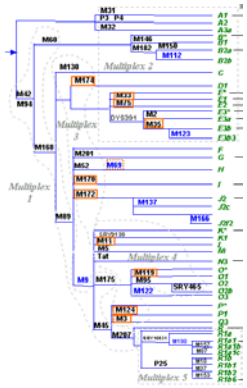
ABI 7000 SDS

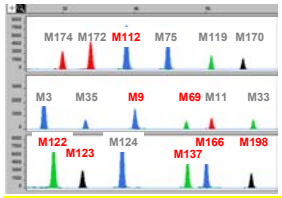
Forensic Utility of SNPs

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…**

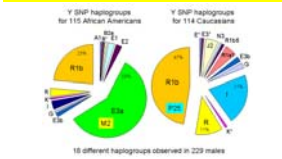
For serious forensic usage, parallel high-throughput methods and multiplex amplification will be required for typing low amounts of DNA

Y-SNP Assays and U.S. Populations





Multiplexing done at both PCR and SNP levels



Vallone & Butler, *J. Forensic Sci., in press (July 2004)*

Forensic Utility
51 Y-SNPs versus 1 Y-STR

For N = 211 male samples

	51Y-SNPs	Y-STR DYS464
Amount of sample consumed	10ng	as low as 50 pg
Number for types observed	**18**	**62**
Analysis	Multiple	1 reaction
Degraded samples	+	?

As a stand alone forensic assay
1 Y-STR is better than 51 Y-SNPs

Typing mtSNPs

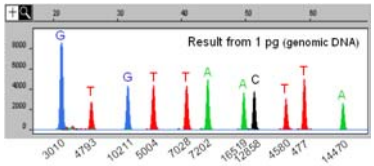
Coding Region SNPs

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

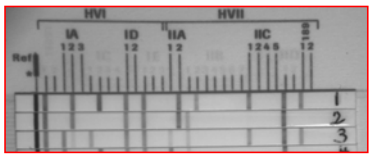
NIST mtDNA Work



Coding Region mtSNP 11plex (minisequencing assay)

Developed with AFDIL to resolve mtDNA most common types

Vallone et al., IJLM, in press



Roche Linear Arrays (probes for HVI/HVII)

Kline et al., submitted

Beta-test/Population Study

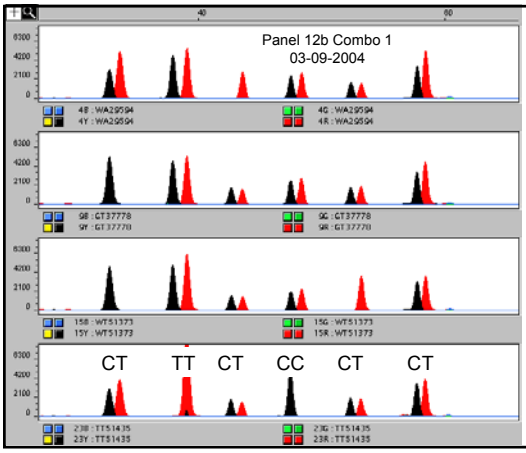
Autosomal SNPs

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

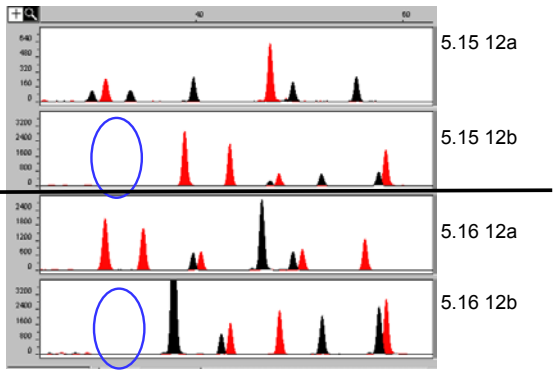
We validated these markers with SNaPshot assays for 8 CEPH samples in July 2001 for the WTC investigation as part of KADAP (Kinship and Data Analysis Panel)

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

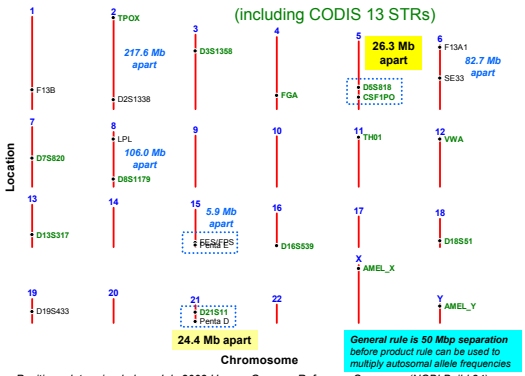
Marker info now on STRBase forensic SNP site:
http://www.cstl.nist.gov/biotech/strbase/SNP.htm



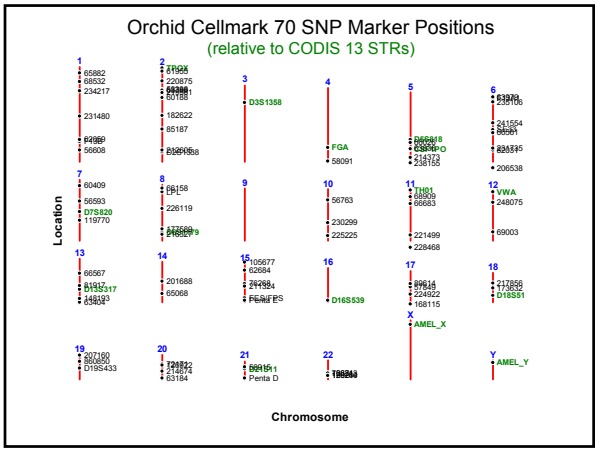
Results from Two Degraded Bone Samples

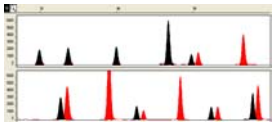


Commercial STR Kit Loci Positions (including CODIS 13 STRs)



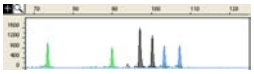
Positions determined along July 2003 Human Genome Reference Sequence (NCBI Build 34)



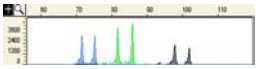


Typing panel 12a (6 loci)
we observed
143 unique genotypes
110 observed once
23 observed 2 times
7 observed 3 times
3 observed 4 times

Typing Panels 12a/b (12 loci)
we observed
189 unique genotypes



Miniplex01 (3 loci)
we observed
176 unique genotypes
164 observed once
11 observed 2 times
1 observed 3 times



Miniplex02 (3 loci)
we observed
163 unique genotypes
142 observed once
18 observed 2 times
2 observed 3 times
1 observed 5 times

DNA Quantitation

Intra and Inter Laboratory Studies

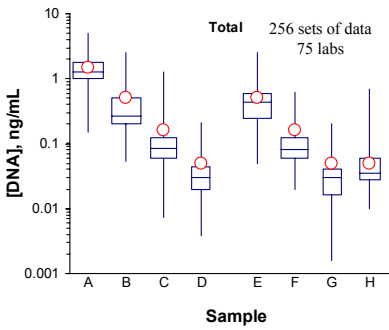
Preliminary Results NIST Quantitation Study 04

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 79 Labs (94%)
Total of 284 sets of data
Participants used 27 different quantification methods
(primarily variations on Quantiblot and Real-time PCR)

Preliminary NIST QS 04 Results



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

Length of the whiskers is the range of values submitted.

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

Circle is the target [DNA].

NOT FINAL DATA SET (UPDATE)