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

Development of New miniSTR Loci for Improved Analysis of Degraded DNA Samples

Michael D. Coble,
Peter M. Vallone and John M. Butler
National Institute of Standards and Technology

NIST Human Identity Project Team


John Butler

Margaret Kline

Jan Redman

Pete Vallone

Dave Duewer

Amy Decker

Jill Appleby

Mike Coble

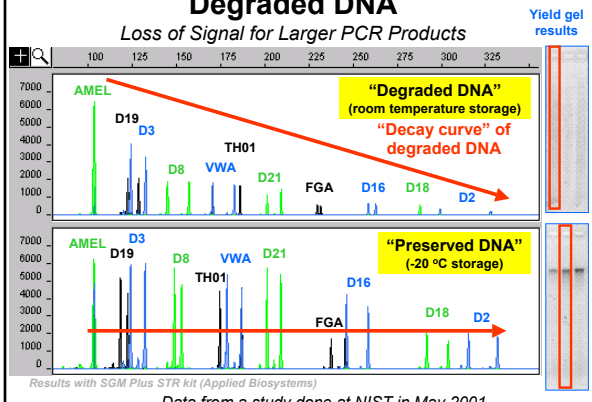
Former (Honorary) Project Team Members

Rich Schoske

Christian Ruitberg


Degraded DNA

Loss of Signal for Larger PCR Products



Data from a study done at NIST in May 2001

STR Size Reduction

Through Moving Primer Positions Closer to Repeat

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

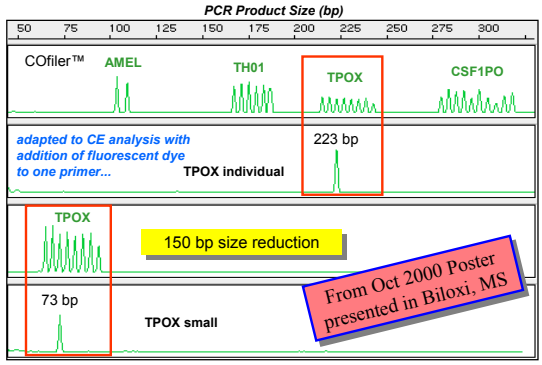
Advantages of Approach:
Size reduction enhances success rate with degraded DNA
Retains same marker information (database compatibility)
Uses highly polymorphic STR loci (high discriminatory power)

Selected References on STRs with Degraded DNA

- Whitaker, J.P., et al. (1995) Short tandem repeat typing of bodies from a mass disaster: high success rate and characteristic amplification patterns in highly degraded samples. *BioTechniques* 18: 670-677
- Clayton, T.M., et al. (1995) Further validation of a quadruplex STR DNA typing system: a collaborative effort to identify victims of a mass disaster. *Forensic Sci.Int.* 76: 17-25
- Yoshida, K., et al. (1997) Evaluation of new primers for CSF1PO. *Int.J.Legal Med.* 110: 36-38
- Schmerer, W.M., et al. (1999) Optimized DNA extraction to improve reproducibility of short tandem repeat genotyping with highly degraded DNA as target. *Electrophoresis* 20: 1712-1716
- Wiegand, P. and Kleiber, M. (2001) Less is more--length reduction of STR amplicons using redesigned primers. *Int.J.Legal Med.* 114: 285-287

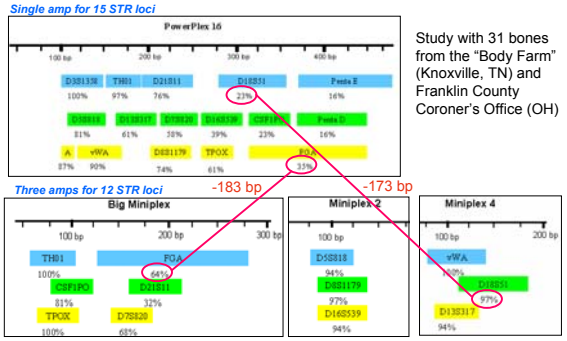
Smaller PCR product size improves success rates with degraded DNA

Development of miniSTRs: Past Work



New primer sets are intended to aid with typing degraded DNA samples as well as future microchip CE and mass spectrometry applications…

Comparison of PCR Amplification Success Rates with Commercial Kit vs. miniSTR Assays



Study with 31 bones from the "Body Farm" (Knoxville, TN) and Franklin County Coroner's Office (OH)

Chung, et al., The application of miniplex primer sets in the DNA profiling of human skeletal remains, submitted

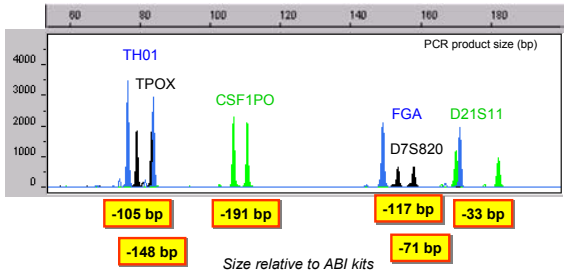
J. Forensic Sci. Sept 2003 issue

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	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), *in press.*

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), *in press*.

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

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

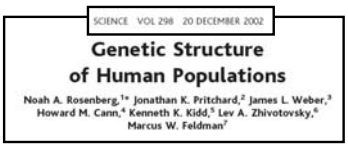
Candidate STR marker selection

http://research.marshfieldclinic.org/genetics/

http://www.decode.com/

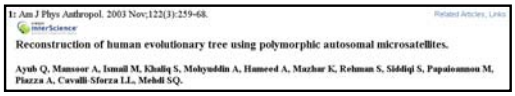
http://www.cidr.jhmi.edu/

Candidate STR marker selection


SCIENCE VOL 298 20 DECEMBER 2002

Genetic Structure of Human Populations

Noah A. Rosenberg,[1]* Jonathan K. Pritchard,[2] James L. Weber,[3] Howard M. Cann,[4] Kenneth K. Kidd,[5] Lev A. Zhivotovsky,[6] Marcus W. Feldman[7]


Am J Phys Anthropol. 2003 Nov;122(3):259-68.

Reconstruction of human evolutionary tree using polymorphic autosomal microsatellites.

Ayub Q, Mansoor A, Ismail M, Khaliq S, Mohyuddin A, Hameed A, Mazhar K, Rehman S, Siddiqi S, Papaioannou M, Piazza A, Cavalli-Sforza LL, Mehdi SQ.

Characterization of New miniSTR Loci

Rosenberg *et al.* 2002 – 1062 samples; 377 STRs; diverse populations

Locus name	Alternate name	Heterozygosity	Number of alleles	Chromosome
D6S1017	GCAT3H10	0.748	9	6
D6S1009	GATA32B03	0.748	13	6
D8S261	AFM123XG5	0.747	17	8
D18S1390	18QTEL11	0.747	15	18
D13S1807	GATA11C08	0.747	9	13
D12S1052	GATA28D02	0.747	8	12
D16S2616	ATA41E04	0.746	11	16
D2S1780	GATA72G11	0.746	13	2
NA-D1S-3	GATA133A08	0.745	12	1
NA-D8S-2	GAAT1A4	0.745	9	8
D4S1625	GATA107	0.745	10	4
D18S851	GATA6D09	0.745	12	18
D8S1136	GATA41A01	0.745	11	8
NA-D5S-1	ATA52D02	0.744	27	5
D3S3038	GATA73D01	0.743	14	3
D15S1515	GATA195H10	0.743	9	15
D16S2624	GATA81D12	0.742	8	16
D2S2972	GATA176C01	0.741	14	2
D13S895	GCAA22C01	0.740	11	13
D11S1998	GATA23E06	0.740	9	11

Focus on:

High Heterozygosity
Small # of Alleles
Tetranucleotide Repeats

Characterization of New miniSTR Loci

Rosenberg *et al.* 2002 – 1062 samples; 377 STRs; diverse populations

Locus name	Alternate name	Heterozygosity	Number of alleles	Chromosome
D6S1017	GCAT3H10	0.748	9	6
D6S1009	GATA32B03	0.748	13	6
D8S261	AFM123XG5	0.747	17	8
D18S1390	18QTEL11	0.747	15	18
D13S1807	GATA11C08	0.747	9	13
D12S1052	GATA28D02	0.747	8	12
D16S2616	ATA41E04	0.746	11	16
D2S1780	GATA72G11	0.746	13	2
NA-D1S-3	GATA133A08	0.745	12	1
NA-D8S-2	GAAT1A4	0.745	9	8
D4S1625	GATA107	0.745	10	4
D18S851	GATA6D09	0.745	12	18
D8S1136	GATA41A01	0.745	11	8
NA-D5S-1	ATA52D02	0.744	27	5
D3S3038	GATA73D01	0.743	14	3
D15S1515	GATA195H10	0.743	9	15
D16S2624	GATA81D12	0.742	8	16
D2S2972	GATA176C01	0.741	14	2
D13S895	GCAA22C01	0.740	11	13
D11S1998	GATA23E06	0.740	9	11

Focus on:

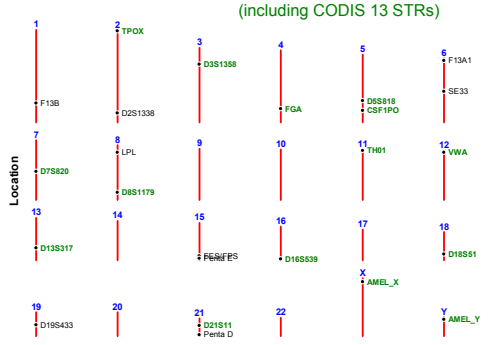
High Heterozygosity
Small # of Alleles
Tetranucleotide Repeats

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

Commercial STR Kit Loci Positions

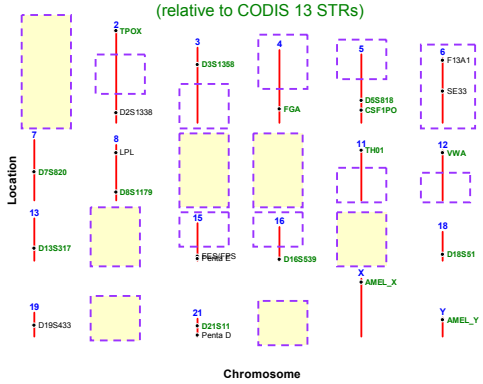
(including CODIS 13 STRs)



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

Locations of Focus for New miniSTR Loci

(relative to CODIS 13 STRs)



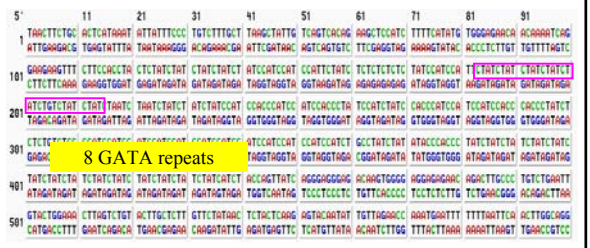
Characterization of New miniSTR Loci

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

PCR Primer Design

http://frodo.wi.mit.edu/cgi-bin/primer3/primer3.cgi/

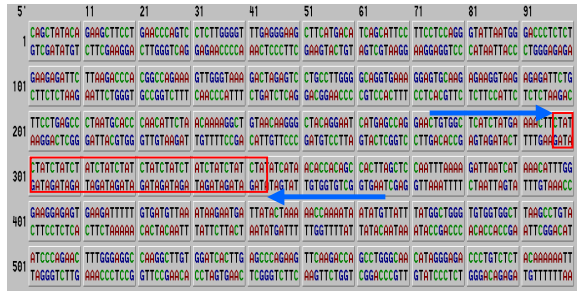
PCR Primer Design



PCR Primer Design

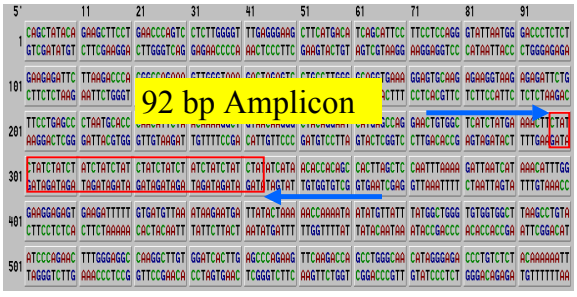
PCR Primer Design

D2S441



PCR Primer Design

D2S441



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

107 potential markers

↓

61 markers with “clean” flanking regions

↓

43 markers with amplicon size < 125bp

↓

18 markers for initial testing

↓

2 three loci miniplexes

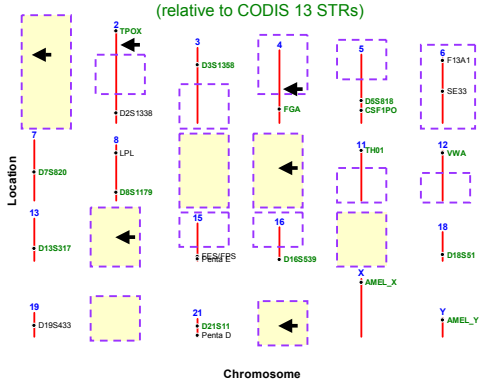
Initial Testing Results

Miniplex01- mD10S1248 - FAM
mD14S1434 - VIC
mD22S1045 - NED

Miniplex02- mD4S2364 - FAM
mD2S441 - VIC
mD1S1677 - NED

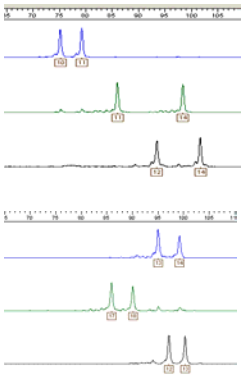
Locations of Focus for New miniSTR Loci

(relative to CODIS 13 STRs)



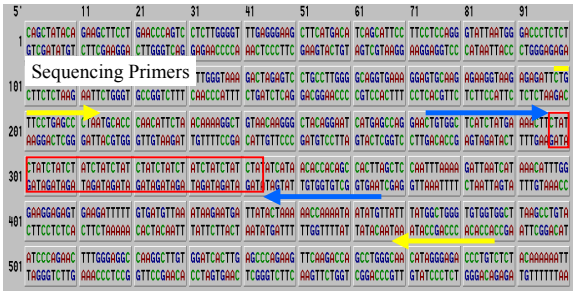
Characterization of New miniSTR Loci

Chr.	Marker Name	Het	Allele Size Range	(Motif)	Ref. Repeat	Primer distance from repeat	Amplicon Size	
10	D10S1248	0.77	20 bp	TETRA	13	1	102	
	GGAA23C05N			GGAA		0		
14	D14S1434	0.72	20 bp	TETRA	10	1	88	Miniplex01
	GATA168F06			GATA		0		
22	D22S1045	0.77	18 bp	TRI	13	3	105	
	ATA37D06			ATA		6		
1	D1S1677	0.74	35 bp	TETRA	15	0	103	
	GGAA22G10N			GGAA		0		
2	D2S441	0.74	18 bp	TETRA	12	0	92	Miniplex02
	GATA8F03			GATA		0		
4	D4S2364	0.64	18 bp	TETRA	7	2	78	
	GAAT1F09			GAAT		1		

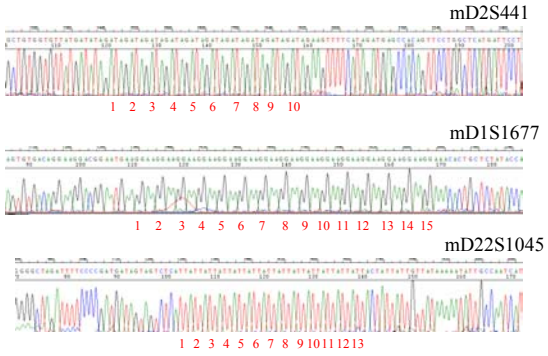


Characterization of New miniSTR Loci

D2S441



Characterization of New miniSTR Loci



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

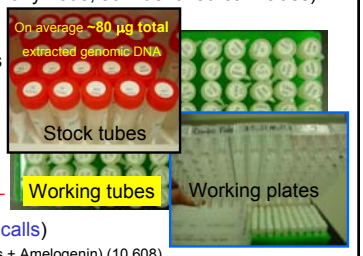
NIST U.S. Population Samples

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)

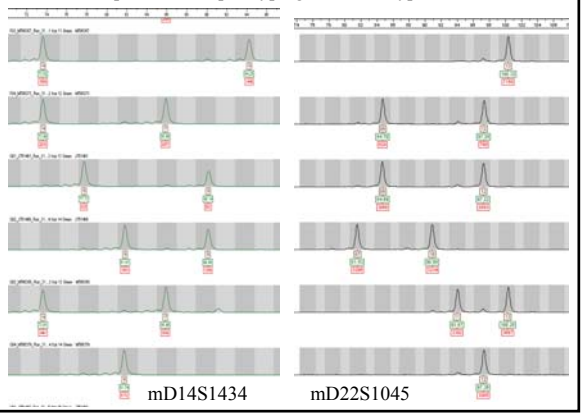
Working tubes/plates 1 ng/uL



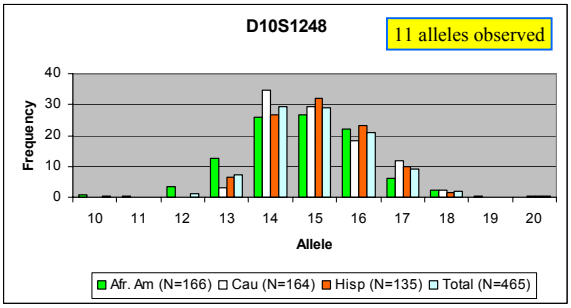
To date: (~50,000 allele calls)
Identifiler (15 autosomal markers + Amelogenin) (10,608)
Roche Linear Arrays (HV1/HV2 10 regions) (6,630)
Y STRs 22 loci—27 amplicons (17,388)
Y SNPs 50 markers on sub-set of samples (11,498)

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

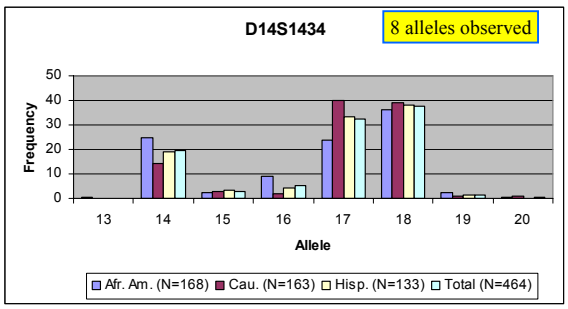
Population Sample Typing with Genotyper Macos



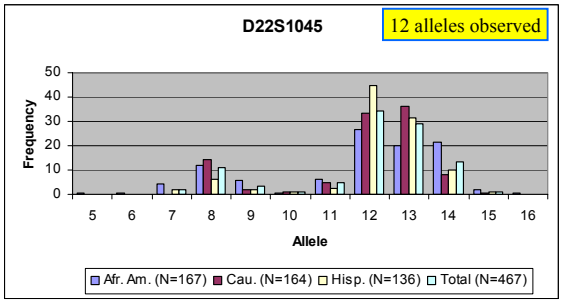
Population Data Analysis



Population Data Analysis



Population Data Analysis



Population Testing –Miniplex01 vs. Identifiler

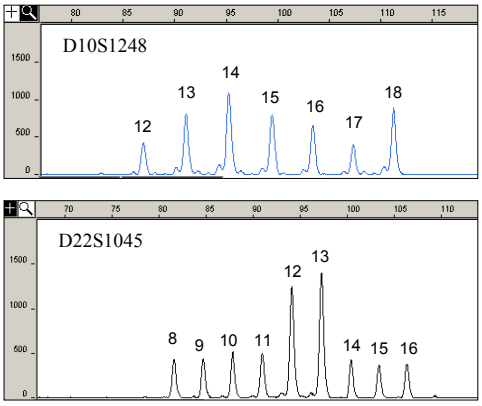
Loci	Heterozygosity
D18S51	0.914
FGA	0.886
D21S1338	0.871
D7S820	0.864
VWA	0.850
D2S1338	0.843
D13S317	0.843
D16S539	0.793
D8S1179	0.786
D19S433	0.764
THO1	0.764
D3S1358	0.757
CSF1PO	0.743
D10S1248	0.733
D5S818	0.729
D22S1045	0.721
TPOX	0.679
D14S1434	0.662

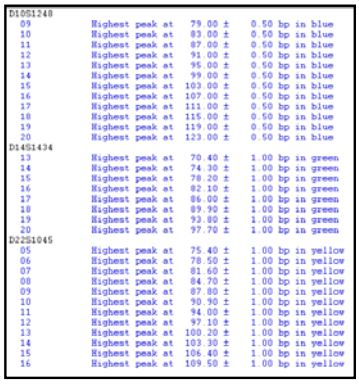
N = 140 Hispanics (Identifiler)

N = 135 Hispanics (Mini01)

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



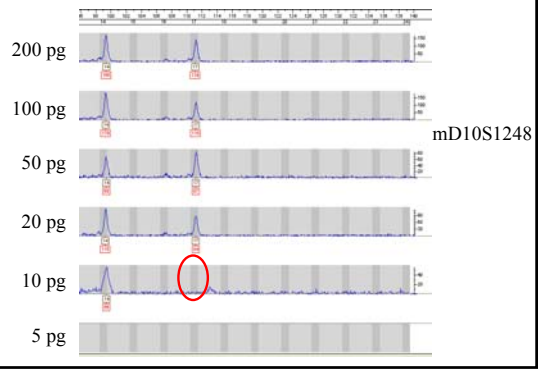


Macros for analysis have been developed

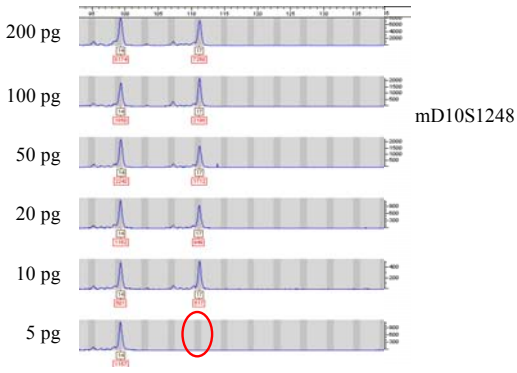
Characterization of New miniSTR Loci

- Candidate STR marker selection
- Chromosomal locations and marker characteristics
- PCR primer design
- Initial testing results
- Population testing
- Allelic ladder construction
- Miniplex assay performance
  - Sensitivity
  - Inheritance with family samples
  - Allele sizing precision
  - Locus stutter percentage characterization
  - Analysis on real-world samples

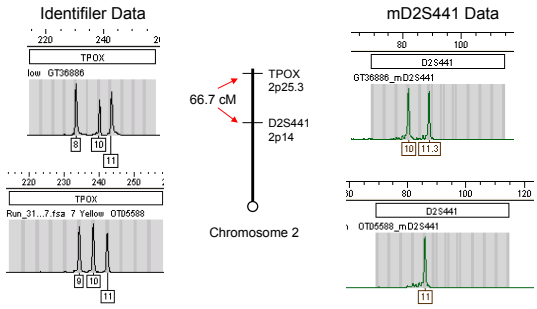
Sensitivity – 28 cycles



Sensitivity – 32 cycles – 2U Taq

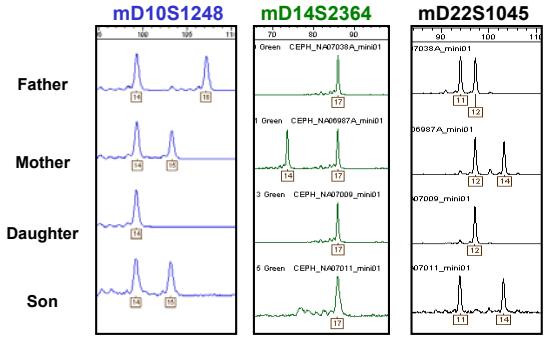


Concordance on Tri-Allelic Patterns for TPOX

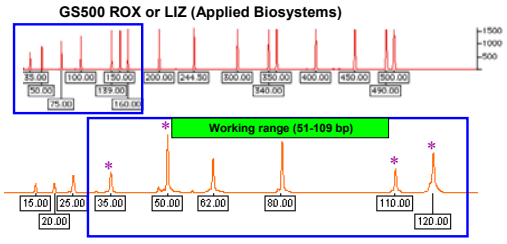


miniSTR primers for TPOX also gave tri-allelic patterns

CEPH Family miniSTR Results



Internal Sizing Standards



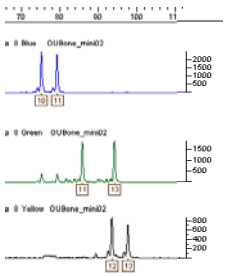
Local Southern sizing methods can be used with either GS500 LIZ or GS120 LIZ (normally for SNaPshot)

Preliminary Testing of Bone Extract

- DNA extract from a decomposing bone at the Body Farm supplied from Bruce McCord's lab

Mini02 Data

32 cycles; 2U Taq



"STR typing of human telogen hairs -- a new approach"

- Hellmann, *et al.* (2001) *Int. J. Legal Med.* 114(4-5): 269-273
- Primer pairs with annealing positions close to the repeat units of the STR loci FES/FPS, TPOX, and TH01 were used for amplification.
- **Shed telogen hairs could be typed!**

Future Plans

Testing and characterization of more markers.

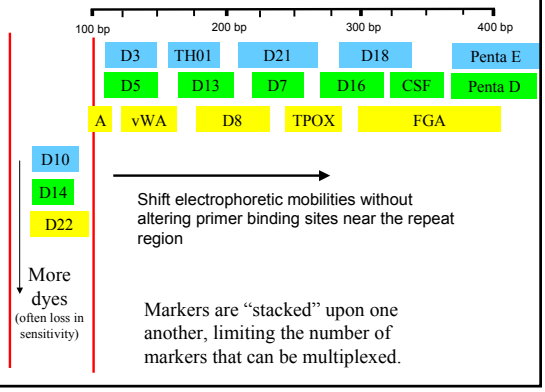
Population Databasing.

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

Expanding Multiplex-ability



Funding and Collaborations

We are funded by an Interagency Agreement between National Institute of Justice and NIST Office of Law Enforcement Standards

Our publications and presentations are made available at:
http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm

Human Identity Project Team
John Butler (Project Leader)
Margaret Kline
Jan Redman
Peter Vallone
David Duewer
Jill Appleby
Amy Decker
Mike Coble

Collaborators (also funded by NIJ):
Bruce McCord and students (Ohio U) for miniSTR work