

NIST Y Chromosome Standards and Multiplex Assays

John Butler
National Institute of Standards and Technology

International Y Chromosome User's Group
Porto, Portugal
November 7, 2002

Outline for Presentation

- New Y STR Markers Available
- NIST Y STR Assays
- Some Casework Results with NIST Multiplexes
- Y SNP Markers and Assays Under Evaluation
- NIST Standard Reference Material SRM 2395
- Nomenclature Issues for Y STR Alleles
- NIST Web Resources for Y Chromosome Information

- U.S. Efforts Towards Selection of Core Y STR and Y SNP Loci

Y STR Markers Available

History of Y STR Marker Discovery

1992 - **DYS19** (Roewer et al.) "Extended Haplotype"

1994 - YCAI a/b, YCAII a/b, YCAIII a/b, DXYS156 (Mathias et al.)

1996 - **DYS389II**, **DYS390**, **DYS391**, **DYS392**, **DYS393** (Roewer et al.)

1996 - DYF371, DYS425, DYS426 (Jobling et al.)

1997 - DYS288, DYS388 (Kayser et al.)

1998 - **DYS385 a/b** (Schneider et al.) "Minimal Haplotype"

1999 - A7.1 (DYS460), A7.2 (DYS461), A10, C4, H4 (White et al.)

2000 - DYS434, DYS435, DYS436, DYS437, DYS438, DYS439 (Ayub et al.)

2000 - G09411 (DYS462), G10123 (de Knijff unpublished)

2001 - DYS441, DYS442 (Iida et al.)

2002 - DYS443, DYS444, DYS445 (Iida et al.); DYS446, DYS447, DYS448, DYS449, DYS450, DYS452, DYS453, DYS454, DYS455, DYS456, DYS458, DYS459 a/b, DYS463, DYS464 a/b/c/d (Redd et al.)

2002 - DYS468-DYS596 ([129 new Y STRs](#); Manfred Kayser GDB entries)

New Y STR Markers (Redd et al.)

ARTICLE IN PRESS

To be published in *Forensic Sci. Int.* (Dec 2002)

Forensic Science International 149 (2002) 1–15
www.elsevier.com/locate/forensic

Forensic value of 14 novel STRs on the human Y chromosome

Alan J. Redd^{a,*}, Al B. Agelton^a, Veronica A. Kearney^a, Veronica A. Contreras^a, Tatiana Karafol^a, Hwayang Paik^a, Peter de Krig^b, John M. Butler^c, Michael F. Hammer^d

^aDivision of Biotechnology, University of Arizona, Tucson, P.O. Box 220, Tucson, AZ 85724, USA
^bResearch Laboratory for DNA Research, JRE, Department of Human and Clinical Genetics, London University Medical Centre, 61, Lincoln's Inn, London WC2A 3LJ, United Kingdom
^cGenome Research Centre, Texas Research Institute of Biotechnology and Biotechnology, P.O. Box 305, South Shore
^dNational Center of Human Genome Research, Biotechnology Division, Bethesda, MD 20895, USA

Received 11 February 2002; accepted 10 September 2002

eight tetranucleotide repeats (DYS449, DYS453, DYS454, DYS455, DYS456, DYS458, DYS459, and **DYS464**), five pentanucleotide repeats (DYS446, **DYS447**, DYS450, DYS452, and DYS463), and one hexanucleotide repeat (**DYS448**)

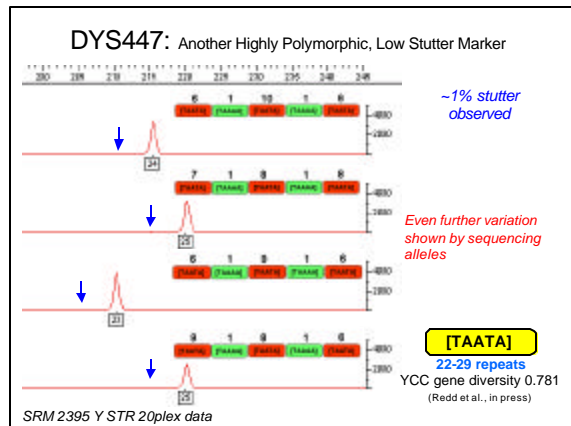
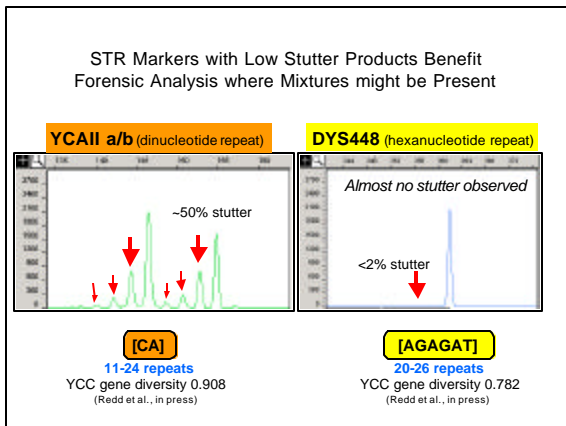
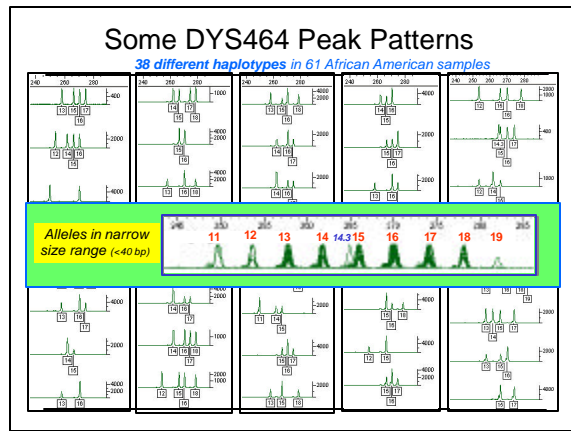
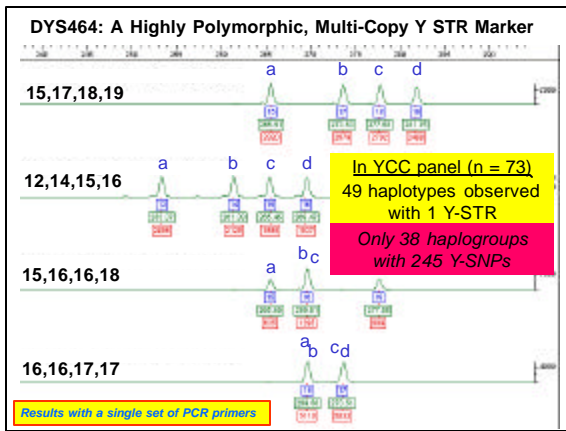
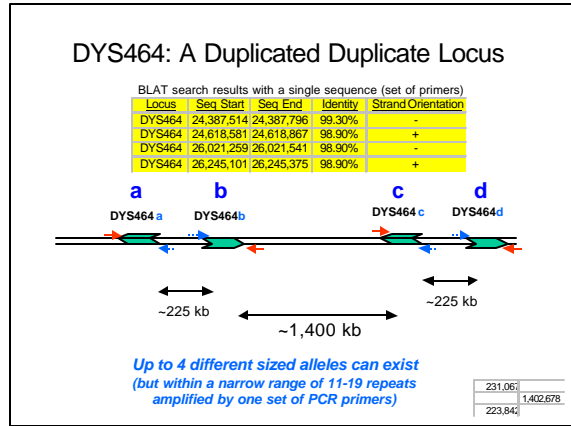
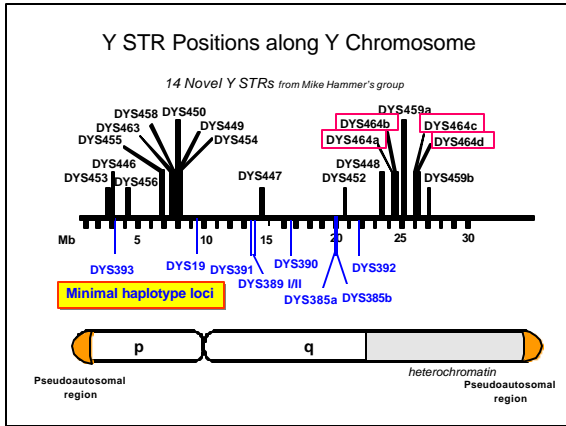
Properties of 14 Novel Y-STRs In YCC

STR	repeat type	repeat motif	avg. # of repeats	allele size range (bp)
DYS449	tetra-	(TTTC) _n GG(TTTC) _n	30	343–383
DYS458	tetra-	(GAAA) _n	17	111–139
DYS456	tetra-	(AGAT) _n	15	141–161
DYS464	tetra-	(CCTT) _n	14	250–282
DYS463	tetra-	(AAAT) _n	11	123–139
DYS454	tetra-	(AAAT) _n	11	196–204
DYS455	tetra-	(AAAT) _n	11	168–184
DYS459*	tetra-	(TAAA) _n	9	140–152
DYS452	penta-	(TATAC) ₂ (TGAC) ₂ (TATAC) _n ...	30	221–251
DYS447	penta-	(TAATA) ₂ (TAAA) ₂ (TAATA) _n ...	25	206–241
DYS463	penta-	(AAGG) ₂ (AAGGG) ₂ (AAGAA) ₂	20	224–269
DYS446	penta-	(TCTCT) _n	14	288–328
DYS450	penta-	(TTTAA) _n	9	356–371
DYS448	hexa-	(AGAGAT) _n GG(AGAGAT) _n	20	382–318

* multicopy Y-STR

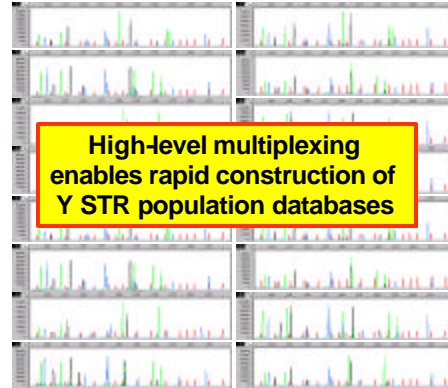
[Redd et al. *Forensic value of fourteen novel Y-STRs on the human Y chromosome. Forensic Science International (in press)*]

To be published in *Forensic Sci. Int.* (Dec 2002)

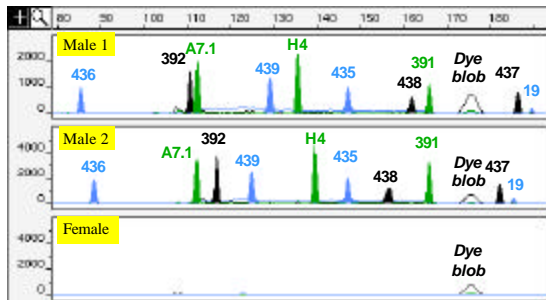


Y STR Assays and Commercial Kits

High-throughput Y STR Typing on the ABI 3100 (16-capillary array)

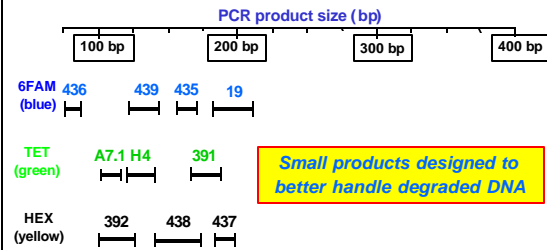


Original NIST Y STR 10plex



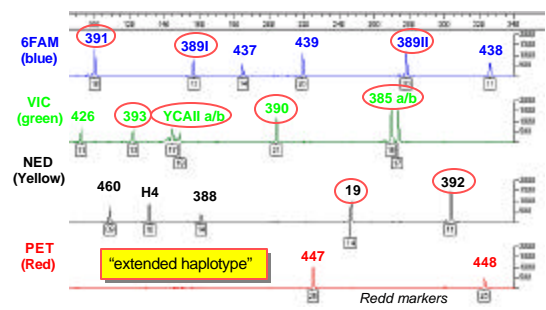
*Poster at Promega meeting in Biloxi, MS (Oct 2000)

Schematic of Loci in NIST Y STR 10plex



Requires filter C on ABI 310

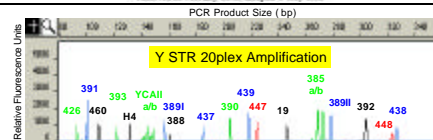
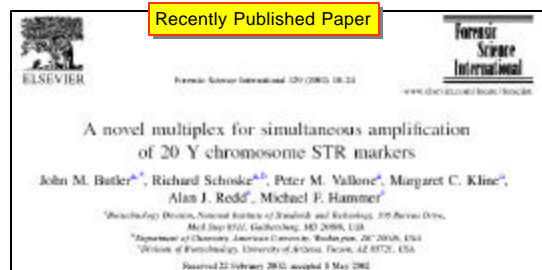
NIST Y STR 20plex Assay

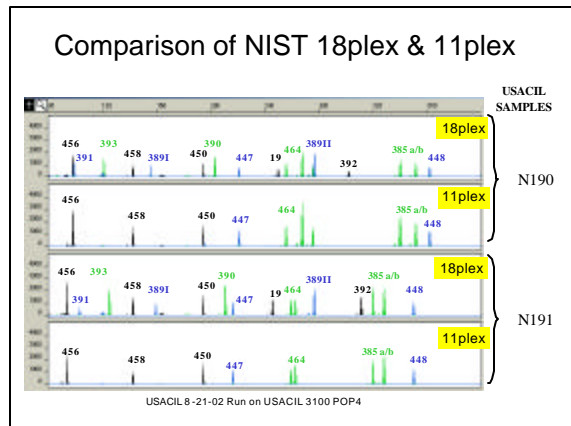
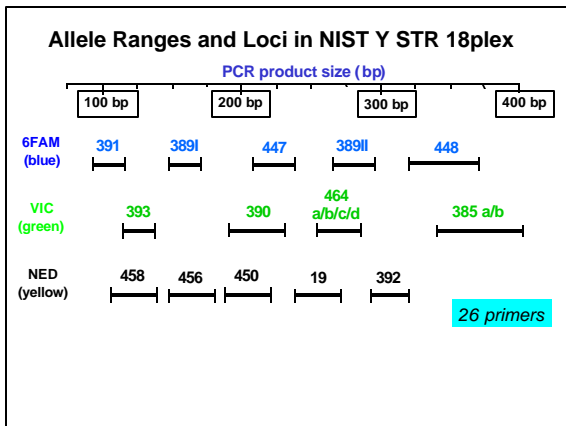
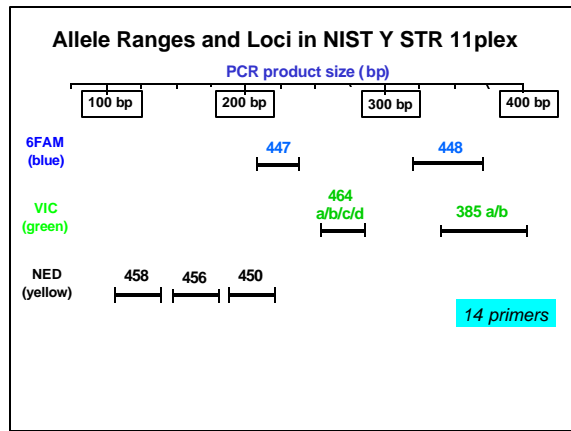
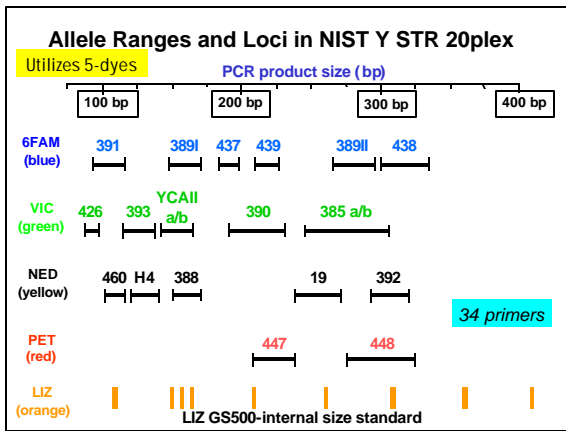
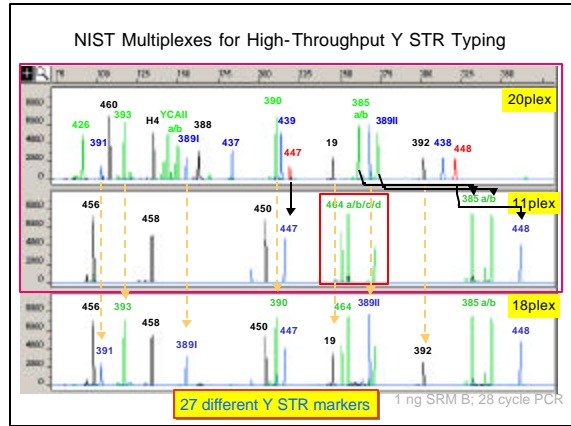
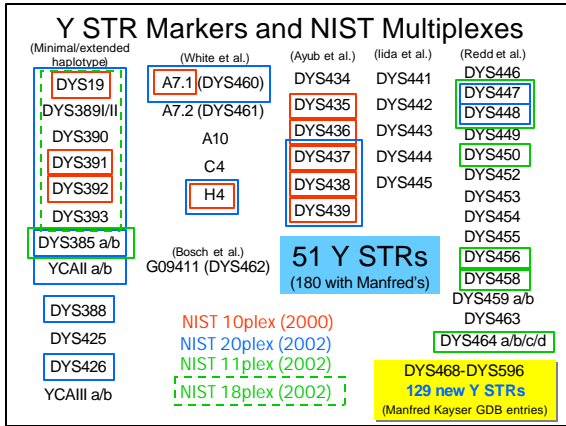


Genotyping performed by size rather than comparison to allelic ladders

Butler et al. (2002) Forensic Sci. Int., 129:10-24

Recently Published Paper





PCR Primer Quality Control

Dye labeled oligos



6FAM (yellow), VIC (orange), NED (red)

- UV Spec to determine concentration
- HPLC to evaluate purity
- TOF-MS to confirm correct sequence
- CE (ABI 310) to determine presence of residual dye molecules

Butler et al. (2001) *Forensic Sci. Int.* 119: 87-96

RELIAGENE TECHNOLOGIES, INC. 18801 HAVEN VALLEY RD. GARLAND, TX 75042

Y-PLEX™ 6

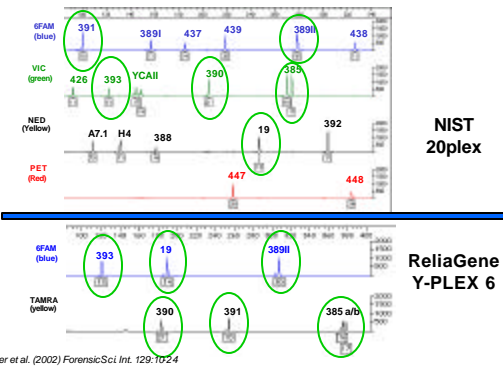
Minimal Haplotype + DYS438, DYS439

RELIAGENE TECHNOLOGIES, INC. "Instruction Manual v5"

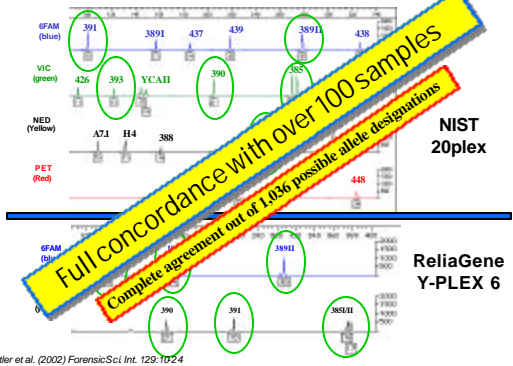
Y-PLEX™ 5

RELIAGENE TECHNOLOGIES, INC. Instruction Manual v1

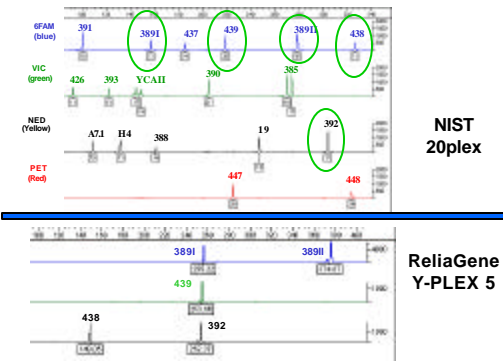
Comparison of Y STR 20plex to Commercial Y-Plex™ 6 Kit (ReliaGene)



Comparison of Y STR 20plex to Commercial Y-Plex™ 6 Kit (ReliaGene)

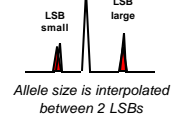
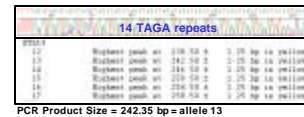
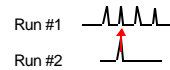


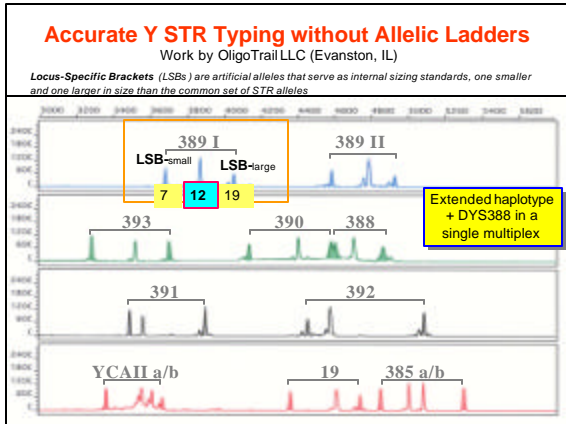
Comparison of Y STR 20plex to Y-Plex™ 5 Prototype Kit (ReliaGene)



Methods for Typing STR Samples

- Allelic Ladders
 - Available in commercial kits
- Precise Sizing (and a sequenced sample)
- Locus-specific brackets (LSBs)





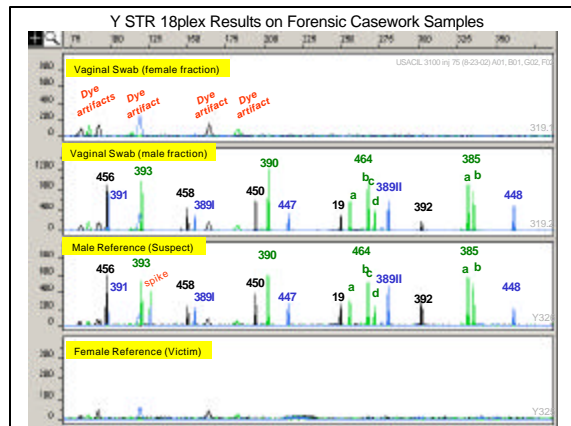
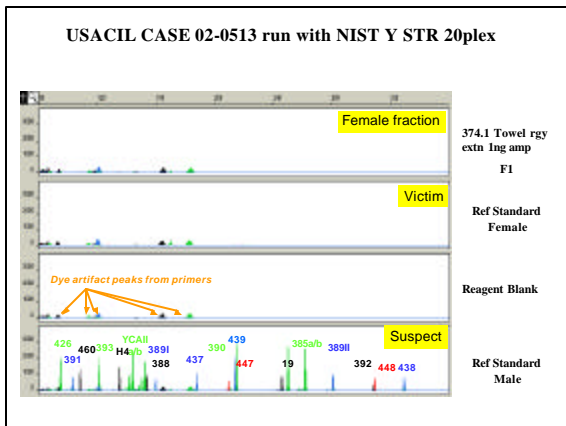
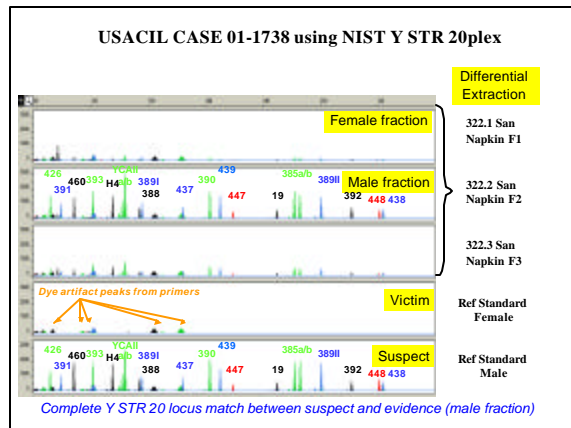
OligoTrail Y Multiplex and LSBs

Locus (Dye Color)	LSB _{small}	Common Alleles	LSB _{large}	Size Range
DYS389 I	7	9-17	19	145-193 bp
DYS389 II	24	25-34	36	265-313 bp
DYS393	6	9-17	19	97-149 bp
DYS390	15	17-28	30	202-262 bp
DYS388	8	10-18	20	270-306 bp
DYS391	5	6-14	16	106-150 bp
DYS392	4	6-17	19	243-279 bp
YCA II a/b	8	11-25	26	100-136 bp
DYS19	8	10-19	21	236-288 bp
DYS385 a/b	5	7-23	25	303-383 bp

Performance of NIST Multiplexes on Forensic Casework Samples

Work done with US Army Crime Lab (Atlanta)

Samples from 6 previously tested forensic cases were examined with NIST 20plex, 9plex (minimal haplotype multiplex), 11plex, 18plex and compared to Y-PLEX 6 and Y-PLEX 5 kits



Y SNP Markers and Assays Used at NIST

Instrumentation Required for SNP Assays

SNaPshot



Multi-Color Capillary Electrophoresis (ABI 310 or 3100)

Luminex Beads



Luminex 100 Flow Cytometer

Primer Extension



Time-of-Flight Mass Spectrometer

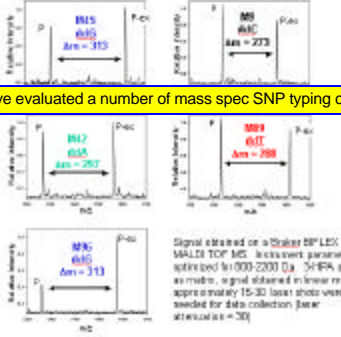
TaqMan



Roche LightCycler

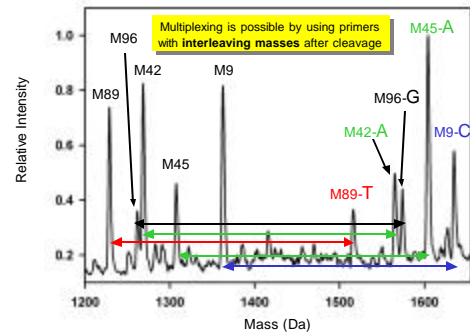
MALDI-TOF Mass Spectrometry

Strategies: spectra for each of the five Y SNP markers



We have evaluated a number of mass spec SNP typing chemistries

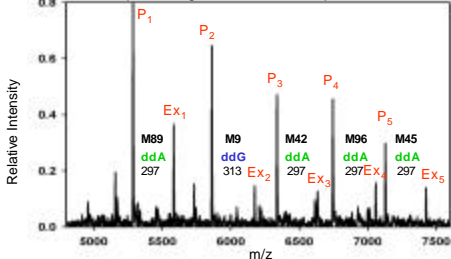
Y SNP 5plex using UV Photocleavable Extension Primers



Vallone et al. Poster presented at ASMS June 2002

Y SNP Multiplex Using MALDI-TOF Mass Spectrometry

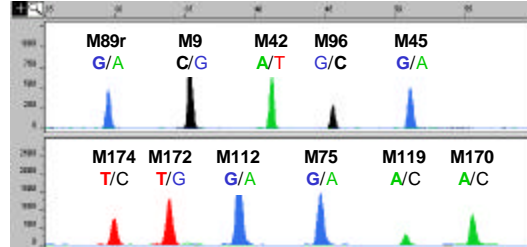
Pentaplex using natural extension primers



- Challenging to optimize multiplex detection (due to ionization differences)
- Requires expensive instrumentation that will probably not be widely adopted

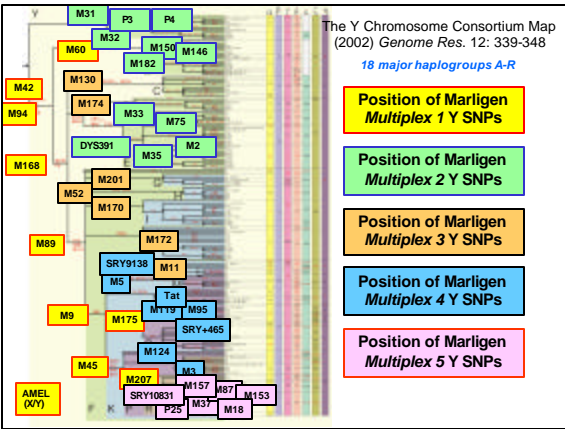
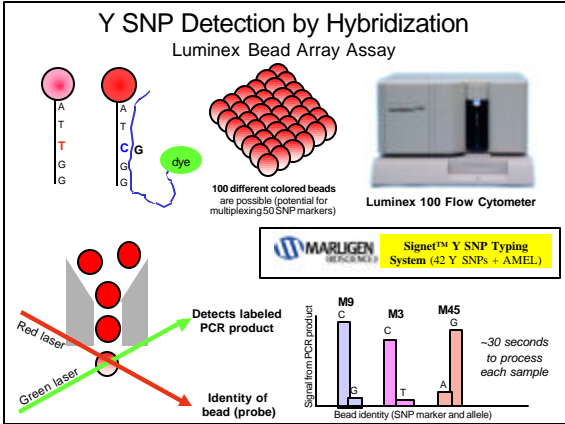
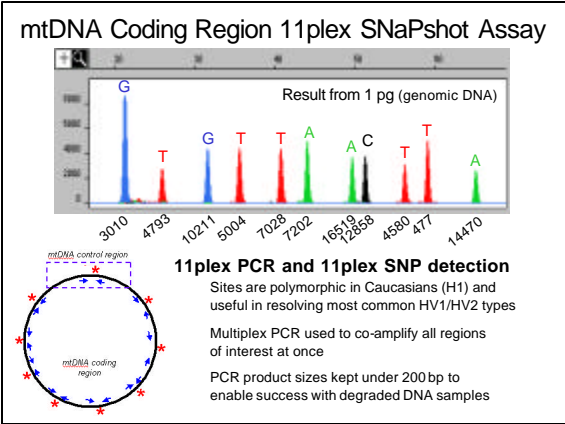
Examples of NIST Y SNaPshot Assays

Poly(T) tails used to space Y SNP alleles



Currently 29 different Y SNPs are in 5 different multiplexes

Equal multiplexing done at both PCR and SNP levels



42 Y SNPs Typed with Luminex Assay

Multiplex 1									
AMEL	M168	M175	M207	M42	M45	M60	M89	M94	
XX or XY	(C/T)	(+/-)	(A/G)	(A/T)	(A/G)	(+/-)	(C/T)	(A/C)	

Multiplex 2											
DYS391	M146	M150	M182	M2	M31	M32	M33	M35	M75	P3	P4
(C/G)	(A/C)	(C/T)	(C/T)	(A/G)	(C/G)	(C/T)	(A/C)	(C/G)	(A/G)	(C/T)	(A/G)

Multiplex 3						
M11	M130	M170	M172	M174	M201	M52
(A/G)	(C/T)	(A/C)	(G/T)	(C/T)	(G/T)	(A/C)

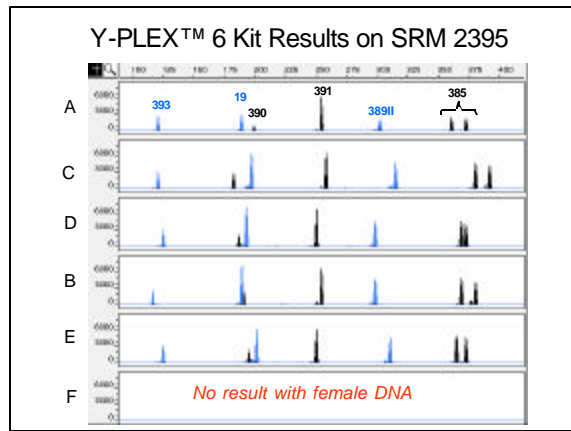
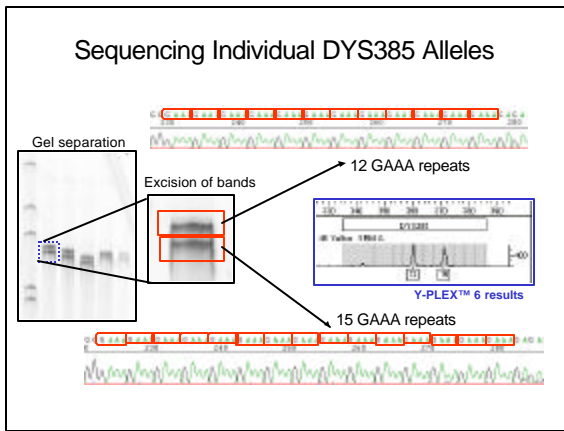
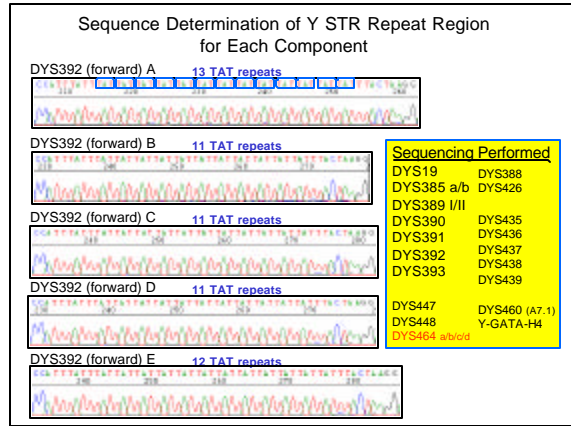
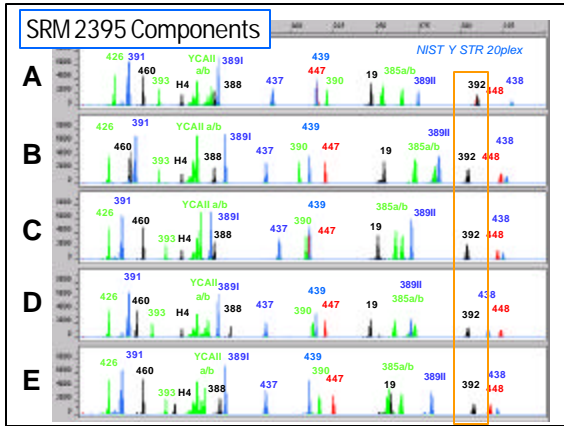
Multiplex 4							
M119	M124	M3	M5	M95	SRY465	SRY9138	Tat
(A/C)	(C/T)	(C/T)	(C/T)	(C/T)	(C/T)	(C/T)	(C/T)

Multiplex 5						
M153	M157	M18	M37	M87	P25	SRY10831
(A/T)	(A/C)	(+/-)	(C/T)	(C/T)	(A/C)	(A/G)

17 Y SNPs overlap with current SNaPshot assays

NIST Y Chromosome Standard Reference Material (SRM 2395)

- ### Status of SRM 2395
- #### NIST Human Y Chromosome Standard
- 5 male samples + 1 female sample (neg. control)
 - Plan to supply 100 ng of each component
 - 22 Y STR markers sequenced to provide certified values (number of repeats)
 - 42 Y SNPs typed with Marligen kit
 - Interlaboratory testing conducted with ReliaGene Y-PLEX 6 and Y-PLEX 5 kits
 - Plan to have available from NIST Standard Reference Material office (www.nist.gov/srm) by Jan 2003
 - Can be used to verify results with any primer sets
 - Will help U.S. labs meet DAB/FBI Standards



Interlaboratory Confirmation of SRM 2395 Y STR Allele Calls

- ReliaGene Technologies Inc.
 - Y-PLEX 6: DYS19, DYS385 a/b, DYS389 II, DYS390, DYS391, DYS393
 - Y-PLEX 5: DYS389 I/II, DYS392, DYS438, DYS439
- OligoTrail LLC
 - Locus-specific brackets: DYS19, DYS385 a/b, DYS389 I/II, DYS390, DYS391, DYS392, DYS393, DYS388, YCAII a/b
- Peter de Knijff

Y SNP Results on SRM 2395

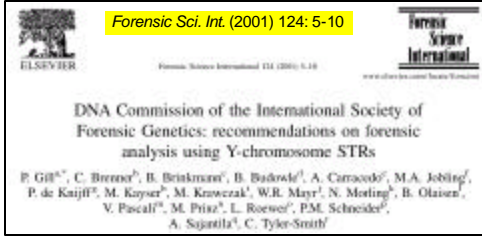
from Marligen Signet™ Multiplexes (Luminex bead assay)

SRM 2395	AMEL	M207	M45	M89	DYS391	M2	M170	M172	M201
	(A/G)	(A/G)	(C/T)	(C/G)	(A/G)	(A/C)	(G/T)	(G/T)	(G/T)
Component A	XY	G	A	T	C	A	A	T	G
Component B	XY	A	G	T	C	A	A	G	G
Component C	XY	A	G	C	G	G	A	T	G
Component D	XY	A	G	T	C	A	A	T	G
Component E	XY	A	G	T	C	A	C	T	G
Component F	XX								

SRM components are all distinguishable from one another with these Y SNPs

42 Y SNPs measured across all samples

Nomenclature Issues for Y STR Alleles



ISFG Guidelines for Y STRs

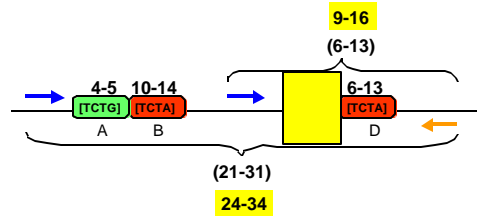
- Locus nomenclature should be DYS number if possible
- Allelic ladders should be used
- Allele nomenclature discussed...

ISFG Guidelines for Y STR Allele Nomenclature

Gill et al. (2001) *Forensic Sci. Int* 124: 5-10

- Number of complete repeats
- A partial repeat (variant allele) is designated by number of complete repeats separated by a dot followed by the number of bases in the incomplete repeat (e.g., 17.3)
- Some locus nomenclatures take into account the **total number** of repetitive units (non-variant plus variant) while others have taken into account **only the variable repetitive stretches**
 - "If a nomenclature is already in use, it is recommended that it should be continued. However, to encourage consistency for newly reported STRs, it is recommended that alleles should be named according to the **total number of repeat units** of the DNA that comprises **both variant and non-variant repeats**."
- Duplicated systems such as DYS385 have to be treated as genotypes and alleles should be separated by a hyphen (e.g., "11-14")

Nomenclature Issues with DYS389 I/II

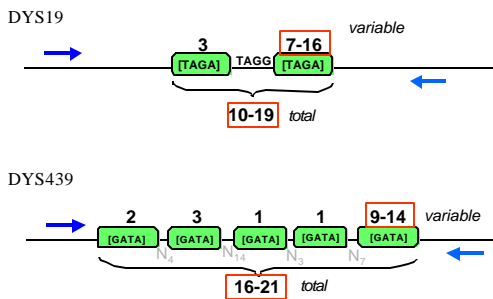


$$\text{DYS389I} = 3+11 = 14; [\text{TCTG}]_3[\text{TCTA}]_{11}$$

$$\text{DYS389II} = 5+12+3+11 = 31; [\text{TCTG}]_5[\text{TCTA}]_{12}[\text{TCTG}]_3[\text{TCTA}]_{11}$$

Original paper (Kayser et al. (1997) *Int. J. Legal Med.* 110:141-149) defines allele nomenclature without repeat segment "C"; it has now been added in more recent nomenclatures thus making alleles +3 repeats larger

Y STR Allele Nomenclatures




Issues with DYS439 Nomenclature



Original description of DYS439 (only variable repeat used)



Alleles 9-14



STRBase

Short Tandem Repeat DNA Internet Database

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

Y-Chromosome STR Information Available

- ❖ Over 200 publications on Y STRs & SNPs cataloged
- ❖ Allele information on over 20 Y STR loci
- ❖ Downloadable PowerPoint on Y STRs and Y SNPs
- ❖ Links to other Y-chromosome sites
- ❖ Information on new Y STR multiplexes developed at NIST (published 20plex primers)
- ❖ Y STR mapped positions along chromosome

Example Y STR Fact Sheet from STRBase

www.cstl.nist.gov/biotech/strbase

We would like to collect variant alleles for Y STRs as they are discovered...

Allele	Set 1	Set 2	Set 3	Repeat Structure	Ref.
7	217	95	295	(TAT) ₇	959
8	248	98	306	(TAT) ₈	959
9	246	104	302	(TAT) ₉	959
10	248	107	305	(TAT) ₁₀	959
11	252	116	308	(TAT) ₁₁	959
12	255	112	311	(TAT) ₁₂	959
13	258	116	314	(TAT) ₁₃	959
14	261	119	317	(TAT) ₁₄	959
15	264	122	320	(TAT) ₁₅	2594
16	267	128	326	(TAT) ₁₆	959

Allele Ladder: Alleles 14-15 generated by SILEX 5 by Bioscience Technology

U.S. Efforts Towards Selection of a Core Set of Y STR and Y SNP Markers

- ### SWGDAM Y Group
- Formed in July 2002
 - Members
 - Jack Ballantyne
 - Bruce Budowle
 - John Butler
 - Ann Gross, *etc.*
 - 40 sample set selected for screening markers and initial testing
 - 20 African Americans, 20 Caucasians
 - Supplied by Carll Ladd (CT State Police Lab)
 - Testing of Y-PLEX 6 and Y-PLEX 5 kits in all labs
 - Jack Ballantyne's lab and John Butler's lab to examine additional Y STR and Y SNP markers in the same sample set

- ### Towards Optimal Y STR and Y SNP Marker Sets for U.S. Populations
- NIST results on 27 Y STRs and 42 Y SNPs
 - 20 African Americans and 20 Caucasians
 - **Minimal haplotype distinguishes all 40 samples**
 - Unique haplotypes can be produced with only 4 markers: DYS464, DYS385, DYS448, DYS390
 - All 42 Y SNPs only produced 8 different haplogroups
 - 3 Y SNPs most useful were M207, M89, M45

Unique Haplotypes Produced from 40 SWGDAM Samples

	464	+385	+448	+390
20 AA	16	20	20	20
20 C	15	18	20	20
40 total	29	37	39	40

↑ Multi-copy locus ↑ Hexanucleotide repeat (low-stutter)

All 40 samples can be fully resolved with only 4 Y STR primer pairs (2 new loci not in minimal haplotype)

Y STR ALLELE FREQUENCY ANALYSIS

- Allele frequency tables for each Y STR and respective ethnicity were generated
- The 10 most polymorphic loci for each ethnicity is given below (out of 27 Y STRs tested)

African American		Caucasian	
STR	Diversity Value	STR	Diversity Value
385 a/b	0.920	464 a/b/c/d	0.900
464 a/b/c/d	0.910	447	0.740

7 of top 10 polymorphic loci in both Ethnic Groups

385 a/b, 464 a/b/c/d, 447, 458, 456, 448, 389II

Only two of these 385 a/b, 389II are from minimal haplotype

Summary of NIST Y Chromosome Work

- Development of new Y STR multiplex assays (**Y STR 20plex, etc.**)
- Evaluation of SNP typing methodologies and development of **Y SNP assays** involving primer extension and the SNaPshot kit
- Creation of a Y Chromosome Standard Reference Material (**SRM 2395**)
- Standardization of information on Y chromosome markers with internet accessibility (**STRBase**)

Rich Schoske

PhD student from American University
Funded by Air Force



Developing Y STR
Megaplexes

Acknowledgments



Funding:

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

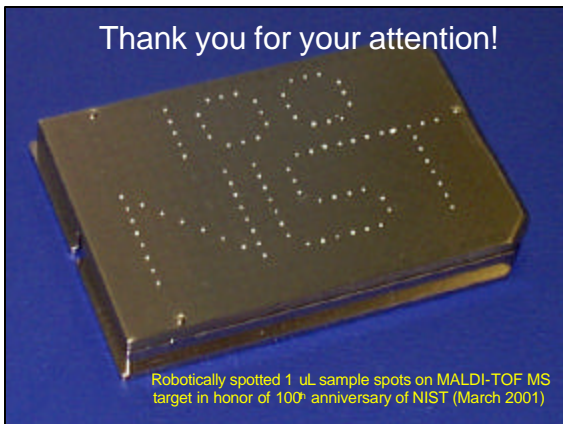
NIST Personnel:

John Butler (Project Leader)
Pete Vallone
Margaret Kline
Jan Redman
Rich Schoske (AU)
Gordon Spangler (AU)
Dave Duewer (Anal. Chem.)

Collaborators:

Mike Hammer and **Alan Redd** (U. Arizona) for Y STR sequences and samples
Mecki Prinz (NYC OCME)
Dave Carlson (Marligen) on Y SNP work with Luminex beads
Del Price and Clem Smetana (USACIL) for casework samples

Thank you for your attention!



Robotically spotted 1 uL sample spots on MALDI-TOF MS target in honor of 100th anniversary of NIST (March 2001)